A METHOD FOR PRODUCT DEFECTIVENESS PREDICTION WITH PROCESS ENACTMENT DATA IN A SMALL SOFTWARE ORGANIZATION

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF INFORMATICS INSTITUTE OF MIDDLE EAST TECHNICAL UNIVERSITY

BY

DAMLA SİVRİOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN THE DEPARTMENT OF INFORMATION SYSTEMS

JUNE 2012

A METHOD FOR PRODUCT DEFECTIVENESS PREDICTION WITH PROCESS ENACTMENT DATA IN A SMALL SOFTWARE ORGANIZATION

Submitted by **DAMLA SİVRİOĞLU** in partial fulfillment of the requirements for the degree of **Master of Science in Information Systems**, **Middle East Technical University** by,

Prof. Dr. Nazife Baykal	
Director, Informatics Institute	
Prof. Dr. Yasemin Yardımcı Çetin	
Head of Department, Information Systems	
Prof. Dr. Onur DEMİRÖRS	
Supervisor, Information Systems, METU	
Dr. Ayça TARHAN	
Co-Supervisor, Computer Engineering, Hacettepe Univer	sity
Examining Committee Members:	
Prof. Dr. Semih BİLGEN	
Electrical and Electronics Engineering, METU	
Prof. Dr. Onur DEMİRÖRS	
Information Systems, METU	
Dr. Ayça TARHAN	
Computer Engineering, Hacettepe University	
Assist. Prof. Dr. Aysu BETİN CAN	
Information Systems, METU	
Assoc. Prof. Dr. Altan KOÇYİĞİT	
Information Systems, METU	

Date: 28.06.2012

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Damla Sivrioğlu

Signature :

ABSTRACT

A METHOD FOR PRODUCT DEFECTIVENESS PREDICTION BY USING PROCESS ENACTMENTT DATA IN A SMALL SOFTWARE ORGANIZATION

Sivrioğlu, Damla M.Sc., Department of Information Systems Supervisor: Prof. Dr. Onur Demirörs Co-Supervisor: Dr. Ayça Tarhan

June 2012, 170 pages

As a part of the quality management, product defectiveness prediction is vital for small software organizations as for instutional ones. Although for defect prediction there have been conducted a lot of studies, process enactment data cannot be used because of the difficulty of collection. Additionally, there is no proposed approach known in general for the analysis of process enactment data in software engineering.

In this study, we developed a method to show the applicability of process enactment data for defect prediction and answered "Is process enactment data beneficial for defect prediction?", "How can we use process enactment data?" and "Which approaches and analysis methods can our method support?" questions. We used multiple case study design and conducted case studies including with and without process enactment data in a small software development company. We preferred machine learning approaches rather than statistical ones, in order to cluster the data which includes process enactment informationsince we believed that they are convenient with the pattern oriented nature of the data.

By the case studies performed, we obtained promising results. We evaluated performance values

of prediction models to demonstrate the advantage of using process enactment data for the prediction of defect open duration value. When we have enough data points to apply machine learning methods and the data can be clusteredhomogeneously, we observed approximately 3% (ranging from -10% to %17) more accurate results from analyses including with process enactment data than the without ones.

Keywords: software defect prediction, machine learning, software measurement, defectiveness, software process enactment.

ÖZ

KÜÇÜK BİR KURUMDA ÜRÜN HATALILIK TAHMİNİ İÇİN SÜREÇ İŞLETME VERİSİNİN KULLANILDIĞI BİR METOT

Sivrioğlu, Damla Yüksek Lisans, Bilişim Sistemleri Bölümü Tez Yöneticisi: Prof. Dr. Onur Demirörs Ortak Tez Yöneticisi: Dr. Ayça Tarhan

Haziran 2012, 170 sayfa

Kalite yönetiminin bir parçası olarak ürün hatalılığı tahmini kurumsal şirketlerde olduğu kadar küçük yazılım kurumları için de hayati önem taşır. Hata tahmini ile ilgili pek çok çalışma yürütülmüş olmasına rağmen süreç işletme verisi, toplama zorluğu nedeniyle kullanılamamaktadır. Buna ek olarak süreç işletme verisinin yazılım mühendisliğinde analizi için önerilen ve genel olarak bilinen herhangi bir yaklaşım yoktur.

Biz bu çalışmada, süreç işletme verisinin hata tahmini için uygulanabilirliğini gösteren bir metot geliştirdik ve "Süreç işletme verisinin kullanımı hata tahmini için yararlı mıdır?", "Süreç işletme verisini nasıl kullanabiliriz?" ve "Bizim geliştirdiğimiz metot hangi analiz metotlarını destekleyebilir?" sorularını cevapladık. Çoklu durum çalışması tasarımını kullandık ve küçük bir yazılım şirketinde süreç işletme verisinin kullanıldığı ve kullanılmadığı durumlar dahil olmak üzere dört durum çalışması için analizler gerçekleştirdik. Süreç işletme bilgisini içeren verinin gruplaması için istatistiksel yaklaşımlar yerine makine öğrenmesi yaklaşımlarını tercih ettik. Çünkü örüntü tanıma amaçlı olan makine öğrenmesi yöntemlerinin, örüntüye yönelik doğası gereği süreç işletme verisi için elverişli olduğunu değerlendirdik.

Yaptığımız durum çalışmaları ile ümit verici sonuçlar elde ettik. Hata açık kalma süresi değerinin tahmini için süreç işletme verisinin kullanımının avantajını göstermek için tahmin modellerinin performanslarını değerlendirdik. Makine öğrenmesi metotlarını uygulamak için yeterli veri noktamız olduğunda ve veri homojen olarak gruplanabildiğinde, süreç işletme verisinin dahil edildiği analiz sonuçlarının, dahil edilmemiş olanlara göre yaklaşık 3% (-10% ile 17% aralığında) daha doğru olduğunu gözlemledik.

Anahtar Kelimeler: yazılım hata tahmini, makine öğrenmesi, yazılım ölçümü, hatalılık, yazılım süreç işletme.

To my family, my friends and my life coach...

ACKNOWLEDGEMENTS

Special thanks should be given to Dr. Ayça Tarhan, my co-supervisor for her professional guidance, dedication and valuable support and to Prof. Onur Demirörs for his useful and constructive recommendations on this study.

I would like to express my very great appreciation to Assoc. Prof. Veysi İşler for his support in providing environment for scientific analyses. I wish to thank various people for their contribution to this study; Şafak Burak Çevikbaş, Dr. Aydın Okutanoğlu and Ahmet Birdal for their valuable technical support on this study; Gökhan Yılmaz, Emine Akman, Caner Kara, Behlül Kuran, Nurbanu Kından and Ayşegül Ünal, staff of Simsoft company, for their help in collecting the raw data. Their all assistance encouraged me to complete this study.

I send my deep thanks to my family and Muhammet Eren Aslan for their trust, patience and support.

TABLE OF CONTENTS

ABSTRACTiv
ÖZvi
ACKNOWLEDGEMENTSix
TABLE OF CONTENTSx
LIST OF TABLES xiii
LIST OF FIGURESxiv
LIST OF ABBREVIATIONS xviii
CHAPTER1
1. INTRODUCTION
1.1 Importance of Defect Data and Process Enactment Information Analysis2
1.2 Difficulty of Collecting Defect Data With Process Enactment
1.3 Aims of This Study
1.4 Approaches Used in This Study5
1.5 Roadmap
2. BACKGROUND
2.1 Defect Prediction Basics
2.2 Quantitative Analysis Methods Utilized for Defect Prediction
2.2.1 Statistical Methods
2.2.1.1 Reliability Models11
2.2.1.2 Hypothesis tests
2.2.1.3 Univariate analysis
2.2.1.4 Bivariate Analysis14
2.2.1.5 Multivariate Analysis: Regression Models, PCA, DA, CA15
Linear Regression Analysis15

Principal Component Analysis (PCA)15
Discriminant Analysis (DA)15
Correlation Analysis (CA)16
2.2.2 Machine Learning Methods
2.2.2.1 K Nearest Neighbor (kNN)16
2.2.2.2 C4.5 Decision Tree
2.2.2.3 Multilayer Perceptron (MLP)
2.2.2.4 Bayesian Belief Networks
2.2.2.5 Apriori
2.3 Defect Prediction Studies
2.3.1 Prediction Models without Process Data by Statistical Analysis Methods 20
2.3.2 Prediction Models without Process Data by Machine Learning Methods21
2.3.3 Prediction Models with Process Data by Statistical Analysis Methods23
2.3.4 Prediction Models with Process Data by Machine Learning Methods24
2.4 Methods to Collect Process Enactment Data
3. DESIGN OF CASE STUDIES
3.1 Goal-Question-Metric (GQM) Tree Approach
3.2 Metric Usability Questionnaire (MUQ)
3.3 Data Collection
3.3.1 Defect and Product Size Data Collection
3.3.2 Process Enactment Data Collection
3.3.2.1 Process Execution Record (PER) Asset
3.3.2.2 Process Similarity Matrix (PSM) Asset
3.4 Data Cleaning and Preprocessing
3.5 Clustering According to Process Enactment Data Approach
3.6 Analysis

4. CASE S	TUDIES	
4.1 Case S	tudy 1 (Project-1 Data)	
4.1.1 Case	Study 1A (Project-1)	48
4.1.2 Case	Study 1B (Project-1)	51
4.1.3 Resu	Its Comparison for Case Study 1 (Project-1)	
4.2 Case S	tudy 2 (Project-2 Data)	62
4.2.1 Case	Study 2A (Project-2)	62
4.2.2 Case	Study 2B (Project-2)	65
4.2.3 Resu	Its Comparison for Case Study 2 (Project-2)	71
5. CONCL	USION AND FUTURE WORK	75
REFEREN	ICES	
APPENDI	CES	
А.	CASE STUDY PLAN	
B.	DETAILS OF CASE STUDY 1A	
C.	DETAILS OF CASE STUDY 1B	
D.	DETAILS OF CASE STUDY 2A	
E.	DETAILS OF CASE STUDY 2B	

LIST OF TABLES

Table 3.1 GQM for This Study	33
Table 3.2 Defect and Product Related Metric Descriptions	41
Table 4.1 GQM for Case Study 1A	48
Table 4.2 Defect and Product Related Metric Descriptions for Case Study 1A	49
Table 4.3 GQM for Case Study 1B	52
Table 4.4 Process Enactment Metric Descriptions for Case Study 1B	55
Table 4.5 Process Attributes Patterns for Case Study 1B Clusters	56
Table 4.6 Results Comparison for Case Study 1	60
Table 4.7 GQM for Case Study 2A	62
Table 4.8 Defect and Product Related Metric Descriptions for Case Study 2A	63
Table 4.9 GQM for Case Study 2B	65
Table 4.10 Process Enactment Metric Descriptions for Case Study 2B	68
Table 4.11 Process Attributes Patterns for Case Study 2B Clusters	69
Table 4.12 Results Comparison for Case Study 2	72
Table A.1 Tasks of Study	88
Table A.2 Study Calendar	88

LIST OF FIGURES

Figure 3.1 Proposed Method
Figure 3.2 The Goal-Question-Metric Hierarchy [14]
Figure 3.3 Metric Usability Questionnaire and Rating for Basic Metrics
Figure 3.4 Metric Usability Questionnaire and Rating for Derived Metrics
Figure 3.5 Process Execution Record (PER)
Figure 3.6 Process Similarity Matrix (PSM) In Literature
Figure 3.7 Process Similarity Matrix (PSM) In This Study
Figure 4.1 PER for Case Study 1B53
Figure 4.2 PSM for Case Study 1B
Figure 4.3 Clustered Metric Files for Case Study 1B
Figure 4.4 PER for Case Study 2B
Figure 4.5 PSM for Case Study 2B67
Figure B.1 MUQ for "Source component" Basic Metric of Project-1
Figure B.2 MUQ for "Created Date" Basic Metric of Project-190
Figure B.3 MUQ for "Closed Date" Basic Metric of Project-191
Figure B.4 MUQ for "Test Type" Basic Metric of Project-192
Figure B.5 MUQ for "Product Version" Basic Metric of Project-193
Figure B.6 MUQ for "Product SLOC" Basic Metric of Project-194
Figure B.7 MUQ for "Product Complexity" Basic Metric of Project-195
Figure B.8 MUQ for "Reproducibility" Basic Metric of Project-196
Figure B.9 MUQ for "Project Phase" Basic Metric of Project-197
Figure B.10 MUQ for "Remaining Open Duration" Derived Metric of Project-1 98
Figure B.11 Weka View of Case Study 1A99
Figure B.12 Multilayer Perceptron Results of Case Study 1A100
Figure B.13 BayesNet Results of Case Study 1A101
Figure B.14 Logistic Results of Case Study 1A102

Figure B.15 J48 Results of Case Study 1A	103
Figure C.1 SimpleKMeans Clustering of Case Study 1B	104
Figure C.2 Weka View of Case Study 1B Cluster 0	105
Figure C.3 Multilayer Perceptron Results of Case Study 1B Cluster 0	106
Figure C.4 BayesNet Results of Case Study 1B Cluster 0	107
Figure C.5 Logistic Results of Case Study 1B Cluster 0	108
Figure C.6 J48 Results of Case Study 1B Cluster 0	109
Figure C.7 Weka View Results of Case Study 1B Cluster 1	110
Figure C.8 Multilayer Perceptron Results of Case Study 1B Cluster 1	111
Figure C.9 BayesNet Results of Case Study 1B Cluster 1	112
Figure C.10 Logistic Results of Case Study 1B Cluster 1	113
Figure C.11 J48 Results of Case Study 1B Cluster 1	114
Figure C.12 Weka View of Case Study 1B Cluster 2	115
Figure C.13 Multilayer Perceptron Results of Case Study 1B Cluster 2	116
Figure C.14 BayesNet Results of Case Study 1B Cluster 2	117
Figure C.15 Logistic Results of Case Study 1B Cluster 2	118
Figure C.16 J48 Results of Case Study 1B Cluster 2	119
Figure C.17 Weka View of Case Study 1B Cluster 3	120
Figure C.18 Weka View of Case Study 1B Cluster 4	121
Figure C.19 Weka View of Case Study 1B Cluster 5	122
Figure C.20 Weka View of Case Study 1B Cluster 6	123
Figure C.21 Multilayer Perceptron Results of Case Study 1B Cluster 6	124
Figure C.22 BayesNet Results of Case Study 1B Cluster 6	125
Figure C.23 Logistic Results of Case Study 1B Cluster 6	126
Figure C.24 J48 Results of Case Study 1B Cluster 6	127
Figure D.1 MUQ for "Detected SCU Name" Basic Metric of Project-2	128
Figure D.2 MUQ for "Source Component" Basic Metric of Project-2	129
Figure D.3 MUQ for "Created Date" Basic Metric of Project-2	130
Figure D.4 MUQ for "Closed Date" Basic Metric of Project-2	131
Figure D.5 MUQ for "Test Type" Basic Metric of Project-2	132
Figure D.6 MUQ for "Product Version" Basic Metric of Project-2	133

Figure D.7 MUQ for "Product SLOC" Basic Metric of Project-2	134
Figure D.8 MUQ for "Product Complexity" Basic Metric of Project-2	135
Figure D.9 MUQ for "Reproducibility" Basic Metric of Project-2	136
Figure D.10 MUQ for "Project Phase" Basic Metric of Project-2	137
Figure D.11 MUQ for "Remaining Open Duration" Derived Metric of Pr	oject-2
	138
Figure D.12 Weka View of Case Study 2A	139
Figure D.13 Decision Table Results of Case Study 2A	140
Figure D.14 BayesNet Table Results of Case Study 2A	141
Figure D.15 SimpleLogistic Table Results of Case Study 2A	142
Figure D.16 J48 Results of Case Study 2A	143
Figure E.1 SimpleKMeans Clustering of Case Study 2B	144
Figure E.2 Weka View of Case Study 2B Cluster 0	145
Figure E.3 DecisionTable Results of Case Study 2B Cluster 0	146
Figure E.4 BayesNet Results of Case Study 2B Cluster 0	147
Figure E.5 SimpleLogistic Results of Case Study 2B Cluster 0	148
Figure E.6 J48 Results of Case Study 2B Cluster 0	149
Figure E.7 Weka View of Case Study 2B Cluster 1	150
Figure E.8 DecisionTable Results of Case Study 2B Cluster 1	151
Figure E.9 BayesNet Results of Case Study 2B Cluster 1	152
Figure E.10 Simple Logistic Results of Case Study 2B Cluster 1	153
Figure E.11 J48 Results of Case Study 2B Cluster 1	154
Figure E.12 Weka View of Case Study 2B Cluster 2	155
Figure E.13 DecisionTable Results of Case Study 2B Cluster 2	156
Figure E.14 BayesNet Results of Case Study 2B Cluster 2	157
Figure E.15 SimpleLogistic Results of Case Study 2B Cluster 2	158
Figure E.16 J48 Results of Case Study 2B Cluster 2	159
Figure E.17 Weka View of Case Study 2B Cluster 3	160
Figure E.18 DecisionTable Results of Case Study 2B Cluster 3	161
Figure E.19 BayesNet Results of Case Study 2B Cluster 3	162
Figure E.20 SimpleLogistic Results of Case Study 2B Cluster 3	163

Figure E.21 J48 Results of Case Study 2B Cluster 3	164
Figure E.22 Weka View of Case Study 2B Cluster 4	165
Figure E.23 DecisionTable Results of Case Study 2B Cluster 4	166
Figure E.24 BayesNet Results of Case Study 2B Cluster 4	167
Figure E.25 SimpleLogistic Results of Case Study 2B Cluster 4	168
Figure E.26 J48 Results of Case Study 2B Cluster 4	169

LIST OF ABBREVIATIONS

ACC: Accuracy

AUC: Area Under the ROC Curve

BBN: Bayesian Belief Networks

CA: Correlation Model

CMMI: Capability Maturity Model Integrated

DA: Discriminant Analysis

FPR: False Positive Rate

GM: Gokhale and Mullen

GQM: Goal-Question-Metric

ISO/IEC:International Organisation for Standardization and International Electrotechnical Commission

kNN: k Nearest Neighbor

KLOC: Kilo Lines of Code

MAE: Mean Absolute Error

MLP: Multilayer Perceptron

MMRE: Mean Magnitude of Relative Error

MUA: Metric Usability Attribute

MUF: Metric Usability Factor

MUQ: Metric Usability Questionnaire

ODC: Orthogonal Defect Classification

PAP: Process Attribute Pattern

PCA: Principal Component Analysis

PER: Process Execution Record

PREC: Precision

PSM: Process Similarity Matrix

RMSE: Root Mean Square Error

ROC: Receiver Operating Characteristic

RRSE: Root Relative Square Error

S: Schneidewind

SCU: Software Configuration Unit

SLOC: Source Lines of Code

SVM: Support Vector Machines

TPR: True Positive Rate

CHAPTER 1

INTRODUCTION

As stated in Weinberg's definition "Software quality is conformance to customer requirements.". Neverthless, software quality is a very crucial feature of a product to gain acceptance from the customer. In this viewpoint, software quality needs continuous monitoring and controlling through the software project. The defectiveness of software is an important quality measure to interpret the status of the product quality. Therefore, software defectiveness should be focus point of researches and quality models. For example, process reference models such as CMMI [1] proposes defect metrics for measurement and analysis activities to achieve multiple process areas.

In this context, we first performed a case study for searching for analysis techniques to understand product defectiveness and affecting factors in a small organization [2]. We applied various statistical and machine learning analysis methods to our product data. By doing this, we collected defect related and product related metrics in different data sets. At the end, we presented our inferences in three categories based on their confidence. According to our evaluation findings, the statistical analysis used for product data results could be considered as confident if supported by new studies. In addition, Apriori machine learning analysis used for defect data results could also be considered as confident, since we observed 90% "correctly classified instances" value in Weka tool. In contrary to this, C4.5 decision tree and logistic regression machine learning analyses used for defect data results had approximately 50% "incorrectly classified instances" value.

We have argued as one of the reasons of this low accurateness rate, process enactment information had not been used for analysis. The CMMI mentioned above suggest after second maturity level the mapping between the product and process data and also suggest to take into account this mapping for process improvement.

Since obtaining process traces and combining them with defect data are not easy, the analyses together with defect and process enactment data are not applied. Machine learning techniques are commonly used for prediction purposes, whereas process enactment data is slightly used. Our proposal is that machine learning approach can interpret more accurate performance results when the process enactment data is used together with product data.

To validate this proposal, we used a method for defect prediction by using machine learning classification [3]. The method clusters the data by using defect data with the context of defect management process before building the prediction model. The data of a small software company, Simsoft, was used for validation. This thesis explains the method in detail and provides its results from four case studies in two different projects.

1.1 Importance of Defect Data and Process Enactment Information Analysis

In all software projects correcting of detected software errors in an attentive and timely manner is vital. If defect correction cannot be completed on time and as it should be, it causes some risks such as giving poor quality products to the customer and / or exceeding the project budget due to error correction labor costs called as rework effort in literature. To minimize these risks, analysis of defect data is required. Besides defect data investigation provides quality improvement and prevents injection of new defects by application of preventive actions to the quality [4]. CMMI's Causal Analysis and Resolution support process area at maturity level 5 suggests selecting defect data for cause analysis [1, 5]. Percentage of defects removed, defect escape rates and number, and density of defects are

suggested to be used as process-performance attributes in CMMI's Organizational Process Performance process area at maturity level 4. Historical defect data is suggested to be used for estimation of project planning parameters in CMMI's Integrated Project Management process area at maturity level 3. And finally, defect density derived measure is suggested to be used to address quality measurement objectives in CMMI's Measurement and Analysis process area at maturity level 2. On the other hand, percentage of defects is suggested to express process performance objectives in ISO/IEC 15504's performance management attribute [6]. Using defect density is suggested as process measurement attribute in ISO/IEC 15504's process measurement attribute.

Since software is different from other engineering disciplines, the information about executed software process during development constitutes importance for the quality and defectiveness of output product. What is the difference from other disciplines? Software production processes are not in a regular and static format as in a fabric production. For software development there are many ways for the production of process artifacts. And the results of applied processes show differences in different environment circumstances. Because of these reasons, evaluation of process knowledge with defect data might be so beneficial. In other words, without knowledge about the processes executed during developing the product, analyzing only defect data may not be sufficient to make decision and take preventive action. Process reference models like CMMI and assessment models like ISO 15504 address this issue over the concept of organizational maturity and process capability, and recommend applying prediction models at higher maturity/capability levels. But, we believe this should not be the only way to use such models.

1.2 Difficulty of Collecting Defect Data With Process Enactment

In recent years software defect data analysis has been a common research area [7, 8, 9]. But analysis and interpretation of software development process data are hard since software engineering is an area which is affected from multiple factors. For example, in some prediction studies [7, 10], authors suffer from the difficulty

of collecting process-related data and taking into account all relevant evidences to generate a prediction model.

In order to understand the context of the product development traces, the traces throughout process practices must be recorded and the analysis of these tracks is required. However, since the nature of software process is abstract and dynamic, and there are too many variables which affect software process directly or indirectly, the measurement of software process is not easy especially in emergent contexts. This difficulty has supported the assessment insight by measuring the performance of software process using the characteristics of the developed product [11].

Since the process related data (e.g. the activities performed, the roles taken, the experience of the process performers) is not stored in the same tools with defect data, the accessibility to the product data and the mapping of process enactment are difficult. The collection of data from a tool's database is categorized as a third degree data collection technique since collection by extracting data from database is independent of real development time [12]. Since this situation causes some issues in mapping product data into process data which will be analyzed to understand the software development process, the most of the organizations can not use these data for prediction models. The organizations which use models or not, need guiding and methods about defectiveness evaluation and prediction.

1.3 Aims of This Study

The data of some process factors such as test type and project phase are stored in the defect tracking tool databases and analyzed by companies [2]. But the data of process enactment can not be provided in most of the cases. We aim to analyze product data with process enactment and show the benefits, if any, of this way in our study.

To do that we investigated the difference in machine learning prediction results with process enactment data and without enacment data. We chose machine learning analysis because of its pattern oriented nature. We believe that the patterns between software processes executed during development and related defect data from a product can be recognized with machine learning techniques.

1.4 Approaches Used in This Study

In this study, we intended to answer the questions; "Is process enactment data beneficial for defect prediction?", "How can we use process enactment data?" and "Which approaches and analysis methods can our method support?".

We used defect open duration metric as dependent variable in our analyses since defect open duration metric could easily be calculated according to the created date and closed date information of the defect obtained from the issue tracking tool. That is to say, we set open duration attribute as class attribute in Weka Tool [13] during machine learning classification operation.

To answer the questions, we first decided which indicators and metrics would be useful for this study. Therefore, the Goal-Question-Metric (GQM) [14] method was applied. The GQM goal was set as follows: to understand the effect of process enactment on software product defectiveness.

We used data of two completed projects in an emergent organization. We grouped defect data used in three categories.

1. Defect data detected during test activities: This data set was obtained from issue tracking tool database.

2. Product version and product size data: This data was obtained from configuration management tool and combined with the defect data. After combination, we had one data set that shows which defect is detected in which product version and how much size the product version has.

3. Process enactment data of defect management process: This data shows the features of each execution of the defect management process. In other words, inputs, outputs, performed process steps (activities) from the start to the end of the process, personnel roles which work for the process, and tools

andtechniquesinformation for each detected defect in software during tests and created in issue trackin tool is process enactment data of defect management process. This data set was manually obtained by using Process Execution Record (PER) and Process Similarity Matrix (PSM) assets.

PER (Process Execution Record) forms [16] was filled by interviewing with process experts. PSM (Process Similarity Matrix) was filled by manuallyreviewing issue tracking tool.

WEKA tool's [13] clustering facility (on cluster tab) was used to cluster the whole dataset obtained by combination of two categorized datasets, and classification facility (on classify tab) was used to conduct machine learning prediction.

We evaluated and compared the accuracy of the analysis results from the data sets with process enactment and without process enactment in the case study A and case study B separately.

1.5 Roadmap

The remainder of this thesis is organized as follows. Section two provides an overview of studies about the techniques used for software defect analysis and prediction, and explains the most known analysis methods. Section three gives the organization of the case studies and their results. Section four discusses the effect of process enactment in defect management process by comparing the performance results of the case studies with process enactment data and without process enactment data. Section five provides overall conclusions and future work.

CHAPTER 2

BACKGROUND

Defect prediction models do not only predict how many latent defects the software contains, but also in which parts of the software they are. In addition to that, they give clues on how to improve the quality of software development processes such as design and implementation. In other words, they aim to show project attributes that are related to better quality or reliability.

According to process reference models defect prediction can be used as an indicator of cause prevention. Therefore, the detection of cause and its place are visualized for process stakeholders.

2.1 Defect Prediction Basics

A "mistake" or "fault" can be committed to the software at any stage during development [15]. When it cannot be detected, it causes unintended work of the software product.

Defect is a stage of the "mistake" cycle. In most cases defects cause fault and failures but this is not a must.

Defects are crucial for the quality of the product since it shows the nonconformance to the customer requirements [17, 18]. Less defective software is more reliable and reliability is an attribute of quality.

Defect detection, correction and verification have cost in the project, because some effort is spent to find, resolve and verify detected defects. These activities are required for quality management. The cost of defect correction and re-testing has positive relation with the latency of the detection [19]. In other words, how much late the defect is detected, that much more defect correction and re-testing cost is. Therefore, defect prevention and the analysis of remaining defects are two important terms for software quality management.

For the defects, open duration metric is important because it gives information about the cost of the defect and makes us understand the trend in process with respect to time. Defect prevention is important to take actions before a flaw does not occur. That not only decreases rework effort, but also establishes an improved quality management system.

One of the defect prevention methods is defect prediction [5]. Defect prediction provides estimating number, type of the defects and their place in the software. In software development projects, planning of quality assurance and test activities, personnel allocating and training, process improvement can be done according to defect prediction results.

In this study we chose to answer our questions in a way that we try to predict defect open duration by using it as a class attribute in machine learning classification techniques for defect prevention.

The meanings of the terms mentioned in this study are below;

Case study: A research strategy, an empirical investigation technique that investigates a phenomenon within its real-life context [20]. This research technique is commonly used in software related studies.

Class attribute:Dependent variable in statistics that is used for classification, you have to select one of your attributes manually before executing classification

analysis. Your data is classified according to your dependent variable and the tool gives you a model to be used for the prediction purpose with its performance evaluation values. Class attribute is called as classifier in some studies (i.e. [21]).

Defect:Software bug that causes an incorrect or unexpected result, or causes product to behave in unintended ways.

Defect open duration: The period that elapses from the detection and recording of the defect to the closure of it. It is in number of day unit.

Defect prediction: The analysis to forecast the behaviour of the defects in software product in future by various quantitative methods.

Defect prevention:The approach that avoids the defects from injection to the software. Defect prediction is only one of the activities that provide defect prevention [5].

Failure:The inability of software that does not perform its required functions within specified performance limits [18, 22].

Fault:An incorrect step, process, or data definition in a computer program which causes the program to perform in an unintended or unanticipated manner [23].

Machine learning: A scientific data mining discipline that concerns with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases [24]. Machine learning aims to recognize patterns and learn. Then, make intelligent decisions based on data after learning. For this purpose some part of whole data is separated as training data and remaining data is kept for test.

Machine learning classification: The techniques that are called as supervised. A classifier is identified for classification.

Machine learning clustering: The techniques that are called as unsupervised. A classifier is not identified for clustering.

McCabe cyclomatic complexity:A software product complexity size measure computed with the number of decision nodes in software product.

Metric: The quantitative indicator of the measurement. In software engineering we can categorize the metrics in three classes: product metrics, process metrics and project metrics [17]. Product metrics are directly measured from software such as size, complexity and defect density. Process metrics are measure of performance of software processes such as testing time, and reviewing time. Project metrics give information about project characteristics such as earned value, and number of skilled project personnel.

Nonconformance: Lack of meeting specified requirements.

Performance evaluation values: The values evaluated in order to determine the accuracy and reliability of a technique.

Process: The series of activities to transform inputs to outputs. In software engineering, processes constitute software development life-cycle.

Process Enactment Data:The workflow of activities that are performed during process execution. The elements of the workflow are inputs, outputs, activities, roles, and tools and techniques.

Source lines of code (SLOC):A measure that shows the lenght of the software product which is computed by counting of the code lines.

Software configuration unit (SCU): The each part of the software product identified in order to to provide management easiness.

Software reliability: The probability of successful operation of a computer program for a specified time in a specified environment.

Quality assurance: Systematic activities that are performed to determine whether product meets customer requirements.

Test:The software quality assurance activity that evaluates by running the code whether product meets customer requirements. This activity provides dynamic verification and validation of the software product.

2.2 Quantitative Analysis Methods Utilized for Defect Prediction

Both statistical and machine learning methods are used for the purpose of defect analysis and prediction. In addition to these studies there are reviews that assess the features and the technical characteristics of defect related measurement studies in literature. Before giving information about these previously performed studies, the analysis method commonly used in these researches are given in this section.

2.2.1 Statistical Methods

Before the discovery of data mining techniques, statistical methods are commonly used in software measurement and analysis like every other science. However, it is thought that statistical methods are insufficient to resolve complex patterns in high number of datasets. Common statistical methods used for defect analysis and prediction are given in the subsections below.

2.2.1.1 Reliability Models

Software reliability is a commonly used attribute of software quality for defect prediction. Software reliability models are based on defect data and the time between defect detected and resolved. They might be categorized in two types. One is called Rayleigh model which depicts the software development process beginning from project initiation to the end of maintenance phase. Second is called software reliability growth models and given with Jelinski-Moranda, Littlewood, Goel-Okumoto, Musa-Okumoto and S models in literature [17]. These second type models are based on exponential distribution approach.

Reliability models deal with several assumptions given below;

1. There are N unknown software faults at the start of testing.

2. Failures occur randomly (times between failures are independent).

3. All faults contribute equally to cause a failure.

4. Fix time is negligible.

5. Fix is perfect for each failure; there are no new faults introduced during collection.

6. Testing intervals are independent of each other.

7. Testing during intervals is reasonably homogeneous.

8. Numbers of defects detected during nonoverlapping intervals are independent of each other.

9. Test process is effective.

The accuracy of method is assessed according to the good-of-fit test results[25].

After data collection, below steps are performed.

Step 1: A model is selected.

Step 2: The parameters of the model are estimated.

Step 3: Fitted model is obtained by substituting the estimates of the parameters into the chosen model.

Step 4: A goodness-of-fit test is performed.

2.2.1.2 Hypothesis tests

The statistical method compares distribution characteristics such as mean and variance of two samples. Besides, whether there is the impact of an attribute on another attribute are searched with this analysis. According to the characteristics

of our data set, t-test, Z-test, Chi-square, ANOVA tests are some of the applied statistical techniques [26].

During analysis, below steps are performed [26];

Step 1: Null hypothesis and alternative hypotheses are stated.

Step 2: Significance level is set.

Step 3: The probability value are obtained by using a statistical package program. Step 4: The probability value is compared with significance level. If probability value is higher that significance level, null hypothesis is accepted.

2.2.1.3 Univariate analysis

With this analysis technique, defect classification and defect count understanding is easy. By analyzing representations, defect progress in future can be predicted, decision making are performed, and defect prevention is achieved [27].

Univariate analysis is carried out with the description of a single variable and its attributes of the applicable unit of analysis. If the variable defect data was the subject of the analysis, the researcher would look at how many subjects fall into a given defect data attribute categories. This analysis provides understanding with examined attribute of an object. Therefore, it is used for descriptive purposes. Variables could be either categorical or numerical.

A basic way of presenting univariate data is to create a frequency table which involves presenting the number of attributes of the variable studied for each case observed in the sample. Furthermore, graphical representation can be used to visualize data. Some of the mostly used graph types for defect data are Pareto Diagram, Histogram, Scatter Diagram and Control Chart.

Moreover, some quantitative measures called central tendency (mean, mode, median and dispersion) range, variance, max, min, quartiles, and standard deviation give information about the distribution of the attribute.

2.2.1.4 Bivariate Analysis

Bivariate analysis involves the analysis of two variables in order to determine the empirical relationship between them [27].

Bivariate analysis can be helpful in testing simple hypotheses of association and causality (checking to what extent it becomes easier to know and predict a value for the dependent variable if we know a case's value on the independent variable).Whereas the purpose of univariate analysis is describing, the purpose of bivariate analysis is explaining. It looks for the correlations, comparisons, relationships and causes between two variables.

During bivariate analysis, the steps given below are applied [28];

Step 1: The nature of the relationship whether the values of the independent variables relate to the values of the dependent variable or not is defined.

Step 2: The type and direction, if applicable, of the relationship are identified.

Step 3: It is determined if the relationship is statistically significant and generalizable to the population.

Step 4: The strength of the relationship is identified, i.e. the degree to which the values of the independent variable explain the variation in the dependent variable.

According to the measurement scales of our variables, statistical techniques that should be used are given below to understand the relationships between pairs of variables in a data set. When we called two variables as X and Y;

•If measurement scales of X and Y are interval and interval, and they are independently distributed, Pearson's correlation is used.

•If measurement scales of X and Y are ordinal and ordinal, and they are independently distributed, Kendall's Tau Spearman's Rho Wilcoxon Signed Test or Mann-Whittney Test are performed. •If measurement scales of X and Y are nominal and nominal, and they are independently distributed, Chi- square Lambda Test is performed.

•If measurement scales of X and Y are interval and interval, and one of them is dependent, simple linear regression is used.

•If measurement scales of X and Y are nominal and interval, and Y is independent, regression with dummy variables and one way analysis of variance are used.

2.2.1.5 Multivariate Analysis: Regression Models, PCA, DA, CA

Multivariate analysis involves observation and analysis of more than two statistical variables at a time.

Several mostly used multivariate analysis approaches are given below.

Linear Regression Analysis

In multivariate linear regression, several independent variables are used to predict one dependent variable. The relationship between dependent variable and independent variables are investigated [29].

Principal Component Analysis (PCA)

PCA decomposes a data table with correlated measurements into a new set of uncorrelated variables [30]. The importance of each component is expressed by the variance (i.e., eigenvalue) of its projections or by the proportion of the variance explained.

Discriminant Analysis (DA)

DA is used to predicting a nominal variable. The prediction of dependent variable is performed by looking for the relationships with the independent variables [29].

Correlation Analysis (CA)

Correlation analysis combines dependent variables to find pairs of new variables which have the highest correlation. However, new variables, even when highly correlated, do not necessarily explain a large portion of the variance of the original tables. This makes the interpretation of the new variable sometimes difficult [29].

2.2.2 Machine Learning Methods

Commonly used machine learning methods for defect prediction are given below.

2.2.2.1 K Nearest Neighbor (kNN)

There is no explicit training phase. K nearest neighbor algorithm searches for minimum distance from the query instance to the training samples to determine the K-nearest neighbors [31].

There is no assumption with data distribution [32]. kNN assumes that the data is in a feature space and the data points are in a metric space. The data can be scalars or possibly even multidimensional vectors. Since the points are in feature space, they have a notion of distance. This need not necessarily be Euclidean distance although it is the one commonly used.

During analysis, the steps given below are applied [33];

Step 1: Euclidean or Mahalanobis distance from target plot to those that were sampled is computed.

Step 2: Samples taking for account calculated distances are ordered.

Step 3: Optimal k-nearest neighbor according to performance value done by cross validation technique is heuristically chosen.
Step 4: An inverse distance weighted average with the *k*-nearest multivariate neighbors is calculated.

Its advantages are robustness to noisy training data and effectiveness if the training data is large.

Its disadvantages areneed to determine value of parameter k (number of nearest neighbors), distance based learning is not clear which type of distance to use and which attribute to use to produce the best results, computation cost is quite high because of the need to compute distance of each query instance to all training samples.

2.2.2.2 C4.5 Decision Tree

Given a set S of cases, C4.5 first grows an initial tree using the divide-andconquer algorithm as follows [34]:

• If all the cases in S belong to the same class or S is small, the tree is a leaf labeled with the most frequent class in S.

• Otherwise, choose a test based on a single attribute with two or more outcomes. Makethis test the root of the tree with one branch for each outcome of the test, partition S intocorresponding subsets S1, S2, . . . according to the outcome for each case, and apply thesame procedure recursively to each subset.

1. Check for base cases for each attribute a,

2. Find the normalized information gain (difference in entropy) from splitting on a,

3. Let a_best be the attribute with the highest normalized information gain,

- 4. Create a decision node that splits on a_best,
- 5. Recurse on the sublists obtained by splitting on a_best, and add those nodes as children of node.

Its advantages are creating decision trees need no tuning parameters [35], no assumptions about distribution of attribute values or independence of attributes,

no need for transformation of variables (any monotonic transformation of the variable will result in the same trees), the method automatically finds a subset of the features that are relevant to the classification, decision trees are robust to outliers as the choices of a split depends on the ordering of feature values and not on the absolute magnitudes of these values, and it can easily be extended to handle samples with missing values.

Its disadvantages are the need to construct a good classifier is proportional to the number of regions, complex view, and not a solution for all problems.

2.2.2.3 Multilayer Perceptron (MLP)

A learning rule is applied in order to improve the value of the MLP weights over a training set T according to a given criterion function [36].

This network has an input layer(on the left) with three neurons, one hidden layer(in the middle) with three neurons and an output layer(on the right) with three neurons. There is one neuron in the input layer for each predictor variable. In the case of categorical variables,N-1 neurons are used to represent the N categories of the variable.

Step 1: The number of hidden layers to use in the network is selected.Step 2: The number of neurons to use in each hidden layer is decided.Step 3: A globally optimal solution that avoids local minima is found.Step 4: It is converged to an optimal solution in a reasonable period of time.Step 5: The neural network is validated to test for overfitting.

Its advantages are generalization and fault tolerance.

Its disadvantages are being computationally expensive learning process, giving no guaranteed solution, not scaling up well from small research systems to larger real systems.

2.2.2.4 Bayesian Belief Networks

A Bayesian belief network is a model that represents the possible states of a given domain. A Bayesian belief network also contains probabilistic relationships among some of the states of the domain [37].

Its steps are;

1. Gather information regarding the way in which the topic under discussion is influenced by conducting interviews

2. Identify the factors (i.e. nodes) that influence the topic, by analyzing and coding the interviews

3. Define the variables by identifying the different possible states (state-space) of the variables through coding and direct conversation with experts

4. Characterize the relationships between the different nodes using the idioms through analysis and coding of the interviews

5. Control the number of conditional probabilities that has to be elicited using the definitional/synthesis idiom [38]

6. Evaluate the Bayesian belief network, possibly leading to a repetition of (a number of) the first 5 steps

7. Identify and define the conditional probability tables that define the relationships in the Bayesian belief network

8. Fill in the conditional probability tables, in order to define the relationships in the Bayesian belief network

9. Evaluate the Bayesian belief network, possibly leading to a repetition of (a number of) earlier steps

Its advantages are providing knowledge in the form of causal structures [39], understandable and extensible network, used easily with missing data.

Its disadvantages are fixed sized hypothesis space [40], underfit or overfit of the data that may not contain any good classifiers if prior knowledge is wrong.

2.2.2.5 Apriori

Apriori mines for associations among items in a large database [41].

Its steps are;

Step 1: It mines a set of execution traces where each has a support value greater than the minimum support threshold [42].

Step 2: It extracts the traces which are a superset of all generator traces.

Step 3: It filters the non-generator traces away, leaving behind a set of generator traces.

Its advantages are usage of large itemset property, easily parallelization, easiness of implementation.

Its disadvantages are assuming transaction database is memory resident, requiring many database scans.

2.3 Defect Prediction Studies

We categorized studies in five categories as using process enactment data or not, using statistical methods or machine learning ones, using assets to collect process enactment data.

2.3.1 Prediction Models without Process Data by Statistical Analysis Methods

Koru and Tian [43] have validated the relationship between complexity and defect count metrics by using statistical hypothesis tests. They have investigated in their study how high complexity affects defect count.

Salman [44] has presented a measurement framework for component oriented software systems as his PhD thesis. He has generated statistical regression models to predict size and effort metrics. The independent variables of his models are component oriented metrics such as number of components, number of connectors, and number of interfaces.

Sivrioğlu and Tarhan [2] have prepared a case study by analyzing same dataset with both statistical and machine learning techniques but dataset has not included process enactment data. The dataset is the data of a completed software project. At the end of the study they have suggested to use contextual data for more accurate results.

Manzoor [45] has tried code metric to estimate defect fix time. But the estimation results have not been found promising. Manzoor has explained the reasons of this inaccurate estimation. He has given 14 factors which affect badly parametric estimation methods performed by using size metrics such as SLOC and FP (function points). His factors are pointed out to the dependence of analysis results to development environment and applied processes.

Ohlsson et al [46] have built prediction models by using Principal Component Analysis (PCA) and Discriminant Analysis (DA) methods. They have used product design metrics for prediction. And they have divided software modules into two categories called as fault-prone and not-fault-prone.

This type studies ignore process related data while analyzing software defect and product data, and their generated models have no process knowledge scraps. Because development environment has high impact on these models, they are specific to the examined project.

2.3.2 Prediction Models without Process Data by Machine Learning Methods

Boetticher [47] has suggested nearest neighbor machine learning method to group data. He has used product related metric data to predict the class in terms of its defectiveness status in the software.

Sivrioğlu and Tarhan [2] have analyzed defect data with both statistical and machine learning methods. They have mentioned that the results of machine learning techniques are more accurate than the ones of statistical techniques, because machine learning gives better results when number of data is high than statistical hypothesis tests when sufficient data is supplied.

Sandhu et al [48] have recommended genetic algorithm technique to predict fault proneness of software modules. He has used requirements and code metrics called as product related metrics for his research.

Çatal and Diri [49] have reviewed software defect prediction studies in a systematical way. They have separated the studies to categories before review. The review states that the studies with using class-level, process-level and component-level measures are not sufficient. Besides, machine learning methods are suggested because they give better results than statistical analysis and expert view methods.

Ahsan et al [50] have conducted a study to estimate bug fix effort. R (Pearson correlation coefficient), MAE (Mean Absolute Error), RMSE (Root Mean Square Error), MMRE (Mean Magnitude of Relative Error) and RRSE (Root Relative Square Error) performance values of five machine learning methods are compared at the end of the study. Because the defect fix effort data are not available, defect fix days metric is used as independent variable for prediction. Product metrics such as number of functions, number of changed operators, SLOC and complexity are included in analysis as input attributes.

When number of data is high, machine learning techniques can give promising results for prediction. But without process enactment obtained, models can not be used for other projects or other development teams of same project.

2.3.3 Prediction Models with Process Data by Statistical Analysis Methods

It is slightly possible to find studies by using process data in literature. Jalote et al [51] have explained a defect prediction approach by performing quantitative quality management and statistical process charts.

Wahyudin et al [9] have presented a defect prediction model by using statistical hypothesis with a combination of product and process measures.

Dhiauddin [8] has generated a prediction model for testing phase in his master thesis. With this model he discovers the strong factors that contribute to the number of testing defects by using statistical methods such as regression analysis.

Gokhale and Mullen [52] have hypothesized a Laplace Transform of the Lognormal distribution model with defect repair times data in day unit. At the same time, they give several factors which are considered affecting defect repair time and causing a lognormal distribution in repair rates because of the factors' multiplicativeness.

Schneidewind [53] has explained the delay between fault detection and fault correction times with exponential distribution. To obtain this statistical empirical result, MSE (Mean Square Error) values of three operational increments have observed in a project. Failure rate, test time parameters are used as input attributes in model.

As mentioned in introduction section, process measures can not be used in most cases because of the collection difficulty. However, the studies which includes process related metrics and analyzed product metrics together with process metrics gave more reliable results for software projects.

If we use process enactment by taking a step forward of process related metrics, the models are going to give more reliable results and predictions can be used for similar projects or development teams with similar environment.

2.3.4 Prediction Models with Process Data by Machine Learning Methods

Fenton and Neil [54] have evaluated defect oriented software metrics and statistical models. They have specified that reliability can not be computed by using defect density because the defects which cause not working of software (its fault) can not be parsed and user oriented defects cannot be chosen. They have stated some inconsistent results that while there is positive correlation between number of defects and other metrics such as software size, in some studies there are negative regression. Regression models provide information only about the past and it does not indicate a prediction model for new data. To analyze average values in data does not explain raw data; therefore it does not give realistic results. The relationship between size and defect is so complex that simple models are insufficient to present these complicated relations. They suggest probabilistic methods such as Bayesian Belief Network (BBN) to present complicated relations between defect and the factors which affect it.

Leey et al [7] have developed a prediction model with micro interaction metrics which are supposed as process-related metrics. In this study, they have made comparisons between the accuracy results of the model of code metrics, the model of history metrics, and the combination of them. They use machine learning classification and regression techniques.

Fenton et al [55] have suggested Bayesian Belief Networks machine learning technique as prediction model. Process data is given for this model, again.

He et al [56] have generated models with J48 (C4.5), Naïve Bayes and SVM (Support Vector Machines) by using same metrics with previously mentioned two studies. The performance of the models has been evaluated by MAE, MMRE and comparison between minimum MAE and median values of data groups.

Song et al [57] have suggested association rule mining for defect correction effort prediction. Apriori accuracy values such as mean, median and standard deviation have compared with the ones of PART, C4.5 and Naïve Bayes approaches. Defect type metric has been used as input data. Also, false negative rate, false positive rate performance values have been reviewed for evaluation.

Zeng and Rine [58] have estimated defect fix effort by using dissimilarity matrix and Self Organizing Maps (Kohonen Networks) which is a type of Neural Networks method. With this data mining technique the data have been clustered for prediction. Model performance has been evaluated by magnitude of relative error (MRE) values of 6 grouped data sets. The input attributes of the model are defect fix time in hour unit, defect severity, the activity during which the defect is detected, system mode, defect category and SLOC (source lines of code) changed. Defect severity, detection activity, system mode and defect category attributes can be considered as contextual metrics.

Thaw et al [59] have performed a similar study with Zeng and Rine. They have concluded their study that prediction model gives accurate results for the projects which have same software development processes like product line projects.

Menzies et al [60] have presented a case study that compares defect analysis results between machine learning and manual analysis used human expertise. ODC (Orthogonal Defect Classification) technique has been used. They have found that manual domain expertise gives more accurate results than treatment learning. But manual analysis is insufficient when we have a complex and large dataset. They have specified that the application of both manual and machine learning analysis gives the most accurate results.

Weiss et al [61] have used the defects life-time phases gone through issue tracking tool as the attributes for defect fix effort prediction. They compared two types of Nearest Neighbor approaches called as with (α -kNN) and without thresholds (kNN). They used text mining for grouping the data before kNN analysis.

Hassouna and Tahvildari [62] have improved Weiss' study by adding 1. data enrichment to infuse additional issue information into the similarity-scoring procedure, 2. majority voting to exploit many of the similar historical issues repeating effort values, 3. adaptive threshold to automatically adjust the similarity threshold to ensure that they obtain only the most similar matches and 4. binary clustering to form clusters when the similarity scores are very low phases.

Hewett and Kijsanayothin [63] have penned down a comprehensive study regarding defect repair time prediction. Firstly, they have applied five different empirical machine learning approaches to two individual data sets with and without attribute selection. AUC (Area Under the ROC Curve), TPR (True Positive Rate, Recall, Sensitivity, Hit Rate), PREC (Precision), FPR (False Positive Rate, False Alarm Rate), ACC (Accuracy) and RMSE (Root Mean Square Error) values have been evaluated for performance. Secondly, they have applied three analytical models: S (Schneidewind) model [53], GM (Gokhale and Mullen) [52] model, their own proposed model and compared the results. Defect detected testing phase, defect severity, defect state and defect state update dates have been used as input attributes for prediction models.

Menzies et al [64] have pointed the importance of the models of similar regions than global ones in empirical studies. Two tools called WHERE to cluster algorithm that divides the data and WHICH learner to find treatments in clusters used to compare the treatments learned from global or local contexts.

It is seen that researchers' insight has been changing as clustering data before modeling. Therefore, we can obtain more local (specialized) results and accurate models for prediction. We will provide this clustering by using process enactment data in our study before applying machine learning techniques. The performance results of clustered dataset and not clustered will be compared.

2.4 Methods to Collect Process Enactment Data

Tarhan and Demirörs [65, 66] have emphasized the importance of process differences in software projects. They have defined and applied some assets such

as Metric Usability Questionaire (MUQ), Process Execution Record (PER), and Process Similarity Matrix (PSM) for data collection. They used MUQ for the decision of usable metrics, PER and PSM for collection and verification of process enactment data.

It is seen that researchers claim the benefits of process measures, machine learning methods, some data collection and grouping methods for defect prediction models one by one. However, none of them use several of these methods together for empirical studies. Combining defect data with process enactment and generating a model from combined data by using above quantitative measurement techniques, we believe, is a promising research topic.

2.5 Validation Methods in Machine Learning and Weka Tool

Machine learning validation methods provide assessing the performance of the models by estimating their accuracies. In other words, it can be evaluated how well the mining models perform against real data.

The descriptions of commonly used validation methods are given below.

Training and Testing Data Sets

In this method, the data set are separated into two sets for training and test. Mostly, training data set is bigger than the portion of the test set. After a model has been processed by using the training set, the model is tested by making predictions against the test set. Since, the data in the testing set already contains known values for the attribute that you want to predict, it is easy to determine whether the model's guesses are correct [72]. The splitting 66% of the data set for training set and remaining for test is a commonly used technique.

Cross Validation

The original data set is randomly partitioned into k sets. Of the k sets, a single set is retained as the validation data for testing the model, and the remaining k - 1 sets are used as training data. The cross-validation process is then repeated k times

(the folds), with each of the k sets used exactly once as the validation data. The k results from the folds then can be averaged (or otherwise combined) to produce a single estimation. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used for validation exactly once [72]. 10-fold cross-validation is commonly used type of cross validation.

In Weka tool, the models are validated by selecting one of the three options given below;

- Using all data set file classified as training set,
- Using another supplied data set file from classified data set as test data,
- k-fold cross validation,
- 66% percentage split.

At the end of classification and clustering executions some performance values are given as output in Weka. These performance values are "correctly classified instances", "incorrectly classified instances", kappa statistic, mean absolute error, root mean squared error, relative absolute error, root relative squared error, TPR, FPR, recall, precision, F-measure and ROC area.

CHAPTER3

DESIGN OF CASE STUDIES

As seen from literature search, previous studies generally do not include process metrics. Even though number of them is low, process metrics which measure test performance, defect resolution timeliness and reliability are analyzed in several studies. But, in this study we do not focus on process metrics directly. Instead of this, we assume that process enactment data give detailed information about process tracks. Therefore, we can investigate the advantages of process related data usage for analysis and prediction. Our motivations to choose process enactment to understand and predict defect data are detailed below;

1. Since the nature of the metric is subsequent, process metrics can be collected only after application of the process. They are performance values. In other words you cannot collect test effectiveness metric, before running any test. This situation causes late feedback in most cases. It means that we are late to prevention; we can only apply corrective action items. However, enactment data can be collected before process execution according to our planning, by taking into consideration previous similar project process applications or company process assets.

2. Process performance metric results are specific to product and project, because they are affected from many factors. These factors can be skills of the project staff, customer experience in domain area, programming language, number of personnel, suitable tool usage etc. On the other hand, process enactment data is more usable to generalize the analysis results. We can use the analysis results of one project for the prediction of other projects that apply same process attribute

patterns.

3. Process metrics cannot be collected and recorded automatically by tools. We need manual calculations after process implementations even though we gather data from databases. But enactment data that is used by this study had been recorded in real time while process was being implemented.

Our base questions waiting to be answered in our study are "Is process enactment data beneficial for defect prediction?", "How can we use process enactment data?" and "Which approaches and analysis methods can our method support?".

We applied case study method from empirical investigation techniques. There are four types of case studies according to objective aspect [12]: exploratory, descriptive, explanatory, or improving. Other categorization related with case study attributes are: 1) Single-case vs. multiple-case, and 2) Holistic vs. embedded.

In these case studies we have four cases and we do not have multiple units within a case since we can say that our case study design is compatible with multiplecase and holistic one. The purpose is descriptive in Case Study 1A and Case Study 2A since we give machine learning analysis results with the only defect data metrics' analysis results as is. On the other hand, the purposes of Case Study 1B and Case Study 2B are "exploratory" since we investigate what will happen when we use process enactmentmetrics together with defect metric. All four case studies are performed for an improving purpose. We intend to improve machine learning defect prediction aspect.

According to data collection aspect there are three categories of methods [12]: Direct (e.g. interviews), indirect (e.g. tool instrumentation) and independent (e.g. documentation analysis).

We used all of the three data collection approaches. Fully structured interviews were performed with process experts by filling Process Executions Records (PER). Issue tracking tool and configuration management tool were used as third degree archival data. The data had already stored in tools while the process was being executed. The quality of the data has improved by the support from expert opinions.

We analyzed data quantitatively with machine learning classification techniques. We interpreted results on comparative basis. We compared the validity results of the project data with process enactmentwith the one without process enactment. Also, the performance values which show classification model prediction accuracy in Weka output were evaluated for validation.

The variations between four case studies are listed below;

Case Study 1A: Project-1 data was collected based on defined metrics. We ignored process enactmentdata in this case study concept.

Case Study 1B: Project-1 data wascollected based on defined metrics. We took into account process enactmentdata in this case study concept and we included it in the analysis.

Case Study 2A: Project-2 data wascollected based on defined metrics. We ignored process enactmentdata in this case study.

Case Study 2B: Project-2 data wascollected based on defined metrics. We took into account process enactmentdata in this case study and we included it in the analysis.

Our proposed method consists of the sequential steps below(Figure 3.1);



Figure 3.1 Proposed Method

3.1 Goal-Question-Metric (GQM) Tree Approach

GQM [14] approach proposes a top-down measurement definition. The approach states that a goal-based measurement way provides opportunity to the organizations for specifying themselves and their project's goals, tracing the goals to the questions that ask what they should wonder for that goal and finally specifying the interpretation of metrics collected for those questions.



Figure 3.2 The Goal-Question-Metric Hierarchy [14]

In this study before analysis phase, to make analyses in terms of our goals, Goal-Question-Metric (GQM) method was applied. Firstly, our aims wereset; secondly the questionsweredefined for each goal; thirdly to answer the question, related metrics and analysis methods werespecified (Table 3.1).

GOAL	QUESTION NO	QUESTION	ANALYSIS METHOD	DERIVED METRIC	BASIC METRIC NO	BASIC METRIC	CASE STUDY NO
To understand if		How much	Bayesnet, Logistic, C4.5 Tree, Multilayer Perceptron Machine Learning Techniques	Defect Data: open duration (closed date-created date)	3.1.1	Defect and Product Data: detected module name, closed date, created date, detected test type, product version, product SLOC, product SLOC, product complexity, reproducibility, detected project phase	Case Study 1A (Project-1), Case Study 2A (Project-2)
there is effect of process enactment on software product defectiveness.	3.1	impact has process enactment on defect open duration prediction?		Defect Data: open duration (closed date-created date)	3.1.2	Defectand Product Data: detected module name, closed date, created date, detected test type, product version, product SLOC, product SLOC, product complexity, reproducibility, detected project phase Process Enactment Data: defect management process attributes	Case Study 1B (Project-1), Case Study 2B (Project-2)

Table 3.1 GQM for This Study

3.2 Metric Usability Questionnaire (MUQ)

MUQ is a form filled according to metric usability attributes [16]. Each form is filled for one metric. The questions and ratings are different for basic metrics (Figure 3.3) and derived metrics (Figure 3.4). Rating is quantitatively calculated according to metric usability factors (MUF) by dividing "Yes" answers to the all

number of questions. Obtained percentage value is qualitatively categorized according to the rules below.

• If the percentage value of factor is between %86-100, MUF is qualitatively categorized as fully statisfied (F).

• If the percentage value of factor is between %51-85, MUF is qualitatively categorized as largely statisfied (L).

• If the percentage value of factor is between %16-50, MUF is qualitatively categorized as partially statisfied (P).

• If the percentage value of factor is between %16-50, MUF is qualitatively categorized as not statisfied (N).

In rating phase, metric usability attributes (MUA) are ordered sequential to their criticality: 1) data metric identity, 2) data existence, 3) data verifiability, and 4) data dependability. If the regarding values of MUA-1 and MUA-2 are F and F; and MUA-3 and MUA-4 are F or L, the basic metric is "usable".

	Please rate each attribute in four scales, based on answers to questions as indicators:
Metric Name:	F: Indicators of the attribute are fully satisfied (%86-100)
Conceptual Definition:	L: Indicators of the attribute are largely satisfied (%51-85)
Assessed On:	P: Indicators of the attribute are partially satisfied (%16-50)
Assessed By:	N: Indicators of the attribute are not satisfied (%0-15)

Attributes		Answers	Rating	Expected Answers	
	Indic	ators		I	
Measure Identity			MUF-1	F	
	Q1	Which entity does the measure measure?			
	Q2	Which attribute of the entity does the measure measure?			
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)			Ratio, Absolute
	Q4	What is the unit of the measurement data?			
	Q5	What is the type of the measurement data? (integer, real, etc.)			
	Q6	What is the range of the measurement data?			
Data Existence			MUF-2	F	
	Q7	Is measurement data existent?			Available > 20
	Q8	What is the amount of overall observations?			
	Q9	What is the amount of missing data points?			
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)			
Data Verifiability			MUF-3	F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)			
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)			Yes
	Q14	Who is responsible for recording measurement data?			
	Q15	Is all measurement data recorded by the responsible body?			Yes
	Q16	How is measurement data recorded? (on a form, report, tool. etc.)			
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)			Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)			
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)			Yes
Data Dependability			MUF-4	F	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)			
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)			
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)			
	Q23	Are the frequencies for data generation, recording, and storing different?			No
	Q24	Is measurement data recorded precisely?			Yes
	Q25	Is measurement data collected for a specific purpose?			Yes
	Q26	Is the purpose of measurement data collection known by process performers?			Yes
	Q27	Is measurement data analyzed and reported?			Yes
	Q28	Is measurement data analysis results communicated to process performers?			Yes
	Q29	Is measurement data analysis results communicated to management?			Yes
	Q30	Is measurement data analysis results used as a basis for decision making?			Yes
Data Normalizability					
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)			
Data Integrability	1				
	Q32	Is measurement data integrable at project level?			
	Q33	Is measurement data integrable at organization level?			

(a) Metric Usability Questionnaire

Metric Name:		
Conceptual Definition:		
Assessed On:		
Assessed By:		
Metric Usability Attributes	Rating	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	F	L or F
Metric Usability Result	F	L or F (Usable) Not Usable otherwise

(b) Metric Usability Rating

Figure 3.3 Metric Usability Questionnaire and Rating for Basic Metrics

The difference of the derived metric rating from basic metric one is that MUF 3&4 values of the basic metrics should be F or L for a derived metric to be an "usable" derived metric.

	Please rate each attribute in four
	scales, based on answers to
	questions as indicators:
Metric	F: Indicators of the attribute are fully
Name:	satisfied (%86-100)
Conceptual Definition:	L: Indicators of the attribute are largely satisfied (%51-85)
Assessed On:	P: Indicators of the attribute are partially satisfied (%16-50)
Assessed By:	N: Indicators of the attribute are not satisfied (%0-15)

Attributes		Answers	Rating	Expected Answers	
	Indica	tors			
Measure Identity			MUF-1	F	
	Q1	Which entity does the measure measure?			
	Q2	Which attribute of the entity does the measure			
	Q3	What is the scale of the measurement data? (nominal, ordinal interval ratio absolute)			Ratio, Absolute
	Q4	What is the unit of the measurement data?			
	Q5	What is the type of the measurement data? (integer,			
	06	real, etc.)			
Data Eviatorea	Q6	What is the range of the measurement data?	MUEA	-	
Data Existence	07		WUF-2	F	A 11 LL 00
	Q7	Is measurement data existent?			Available > 20
	Q8	What is the amount of overall observations?			
	Q9	What is the amount of missing data points?			
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)			
Data Verifiability			MUF-3	F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)			
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)			Yes
	Q14	Who is responsible for recording measurement data?			
	Q15	Is all measurement data recorded by the responsible body?			Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)			
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)			Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)			
	Q19	Is all measurement data stored in the same place? (in a file database, etc.)			Yes
Data Dependability			MUF-4	F	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)			
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)			
	Q22	What is the frequency of storing measurement data?			
	Q23	Are the frequencies for data generation, recording, and storing different?			No
	Q24	Is measurement data recorded precisely?			Yes
	Q25	Is measurement data collected for a specific purpose?			Yes
	Q26	Is the purpose of measurement data collection known			Yes
	Q27	Is measurement data analyzed and reported?			Yes
	Q28	Is measurement data analysis results communicated			Yes
	Q29	Is measurement data analysis results communicated			Yes
	Q30	Is measurement data analysis results used as a basis for docision making?			Yes
Data					
Normalizability	Q31	Can measurement data be normalized by parameters			
Data Integrability		or measures? (If yes, please specify them)			
	Q32	Is measurement data integrable at project level?			
	Q33	Is measurement data integrable at organization level?	1		

(a) Metric Usability Questionnaire

Metric Usability Attributes	Rating	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	F	L or F
MUF-3&4 for basic metric-1	F	L or F
MUF-3&4 for basic metric-2	F	L or F
MUF-3&4 for basic metric-n	F	L or F
Metric Usability Result	F	L or F (Usable) Not Usable otherwise

(b) Metric Usability Rating

Figure 3.4 Metric Usability Questionnaire and Rating for Derived Metrics

In this study after defining the metrics, metric usability analysis for each basic metric has been performed to determine if the metric is applicable and available for our study. MUQ form was filled for each basic metric and the derived metric "defect open duration". During the examination of filled MUQ forms it was determined not to use number of requirements based on product version metric. Because "number of requirements" metric was collected on monthly basis instead of product version basis, this period was not applicable for our analysis goal.

3.3 Data Collection

In this study the two projects' data of Simsoft company is used. Simsoft Computer Technologies Co., Ltd. is a software development company established in 2006. It is especially experienced in simulation systems. Simsoft is conducting business as a university - industry Cooperation Company in Technology Development Center at Middle East Technical University Technopolis in METU Campus. It has 30 personnel, including Software Engineers, Modeling and Graphics Designers, and Quality Assurance Supporters. The company has developed software projects for a large number of institutes especially for defense industry by now. The organization has already ISO 9001 [67] certificate and executes documented process assets in compatible with CMMI Level 3. The company has a specific measurement process, in this concept obeying policies for analyzing the monthly data and reporting the results to high level management.

The projects whose defect data is used are listed below;

Project-1: The software product developed in the project has 2 Software Configuration Units (SCU) with 4 module types, and 6 personnel worked for 7 months project duration. At the end of the development, C++ source lines of code are 23 KLOC, number of requirements is 955, and the number of defects detected during tests is 296. This project's development phase was completed in January 2012.

Project-2: The product has 14 Software Configuration Units (SCU), and 15 personnel worked for 8 months project duration. At the end of the development C# source lines of code is 188 KLOC, number of requirements is 1492, and the number of defects detected during tests is 425. This project's development phase was completed in June 2011.

3.3.1 Defect and Product Size Data Collection

Since software testing is a must and a part of development, resolution of detected defects is a necessity. With this aspect for the tracking of defects in software, a tracking tool is used by lots of institutions contemporarily. With these tools a detected defect during any quality activity can be recorded and assigned to related personnel for resolution. After assignment; monitoring, verifying and closing activities are tracked over these tools. In addition to the tracking of defect status, the detailed information regarding the defect such as software module, product version where the defect is detected, test type and source project phase during which the defect is detected can be accessed at any time since defect information is stored with its history in the database. These tools store descriptions of the defects detected on software, detection dates and resolution status of defects.

While using issue tracking tool for the monitoring of the status of the defects detected in software product, in order to perform the updates on product in a controlled manner, organizations need configuration management tool. Configuration management tools provide a common environment to the developers to track the modifications in product. These tools do not allow multiple personnel to modify the product at the same time. The personnel can access whole update information beginning from first creation of the product in the tool. With the aid of configuration management tool, the important information about software product can be obtained historically since it stores all product versions in a historical manner and anyone can access versioned product at any time.

The defect related basic metrics' data; detected software configuration unit (SCU) name, created date, closed date, test type, product version and reproducibility were extracted from issue tracking tool database. Besides, the defect related derived metric called as "defect open duration" was manually calculated as the difference between the closed and the created dates.

The project phase process metric data is manually collected by filling "Project Phase" column in Excel sheet while directly interviewing with the process expert.

The product size basic metrics' data; product version size (logical source lines of code) and complexity (McCabe cyclomatic complexity), however, are obtained indirectly from the tool. We say "indirectly" because these metrics are calculated with LOCMetrics tool [71] by using the product version where the defect is detected from the information recorded in the tool. In other words, to collect SLOC and McCabe cyclomaticcomplexity, configuration management tool was used together with the product version information in issue tracking tool, and the total SLOC was counted by LOCMetrics and recorded manually. Metric descriptions are given in Table 3.2.

		Measurement
Metrics	Metric Description	Scale
	The time starting with the creation of the	
Remaining	defect and finishing with the closure of the	
Open Duration	defect. Calculated by the difference of detect	
0pm 2	closed date and defect created date. Unit is	
	number of days.	Absolute
Detected SCU	The name of the software configuration unit	
Name	(SCU) where the defect is detected. Entered by	
	developer to the issue tracking tool.	Nominal
	The date when the defect is detected. Filled by	
Created Date	the issue tracking tool automatically when the	
	tester record the defect.	Interval
	The date when the defect is closed. Filled by	
Closed Data	the issue tracking tool automatically when the	
Closed Date	project manager change the status of the defect	
	as "Closed".	Interval
	The name of test type during which the defect	
Test Type	is detected. Entered by tester to the issue	
	tracking tool.	Nominal
Droduct	The version of the software product which the	
Version	defect is detected. Entered by tester to the	
Version	issue tracking tool.	Ordinal
SLOC	The size of the product version where the	
(Source Lines	defect is detected. Collected from	
of Code)	configuration tool by using Locmetrics tool.	Absolute
	The McCabe complexity of the product	
	version where the defcet is detected. Collected	
Complexity	from configuration tool by using Locmetrics	
	tool.	Absolute
D 1 1 1 1 1	The repetability of the defect detected. Entered	
Reproducibility	by tester to the issue tracking tool.	Nominal
	The project phase where the defect detected	
Project Phase	Collected manually by domain expert.	Nominal

Table 3.2 Defect and Product Related Metric Descriptions

The raw data of regarding metrics are gathered in an Excel sheet.

3.3.2 Process EnactmentData Collection

Process Execution Record (PER) and Process Similarity Matrix (PSM) assets are utilized to gather process traces [16]. First, you decide on which process data is needed for your analysis. Then, PER is filled out for your regarding process and process attributes such as inputs, outputs, activities and tools. This knowledge is then entered to PSM Excel sheet for each process executions. For example, process execution might be each product version release for a configuration management process.

3.3.2.1 Process Execution Record (PER) Asset

PER is a form in Word file format (Figure 3.5) used to define all actual process values in process attributes basis. Inputs, outputs, roles, tools and techniques all are process attributes and with the help of PER form, all alternative values of them for process executions are recorded. Prepared list in PER are used to fill PSM.

N.	Name	Description		
rools	and Techniques: Please	e list the tools and te	echniques that are used to support process execu	tior
2				
1				
No	Name		Description	
Roles	: Please list the roles that	were allocated res	ponsibilities in process execution.	
4				
3				
2				
1				
No	Name	Description		
Activi	ties: Please list in sequer	nce the activities that	t were performed while executing the process.	
2		1		
1				
NO	Name	Description		
Jutpu	Its: Please list the outputs	s from the process of	execution.	
		<i>c</i>		
· ·				
1	Name	Description		
No	Name	Description	1011.	
nnut	Diagon list the inputs to	the process execut	ion	
Proc	ess Execution No:		Recorded By:	
1100	Coo Ivanic.		Recolueu Oli.	

Process Execution Record (Internal Attributes)

Figure 3.5 Process Execution Record (PER)

In this study the collection of defect management process enactment data was aimed in order to capture the traces of defect management process and combine it with defect related process data and product data for prediction analysis. PER forms were collected with expert opinions by interviewing.

3.3.2.2 Process Similarity Matrix (PSM) Asset

PSM is a spreadsheet in Excel file format (Figure 3.6) used to gather process attribute values for all process executions. Horizontally there are process attributes specified in PER before, vertically there are numbered process executions. The cells in matrix is filled by entering a circle sign if the process attribute is applicable for regarding process execution. After PSM is completed, the differences in columns are examined and clustering is manually performed.

		Pre	oces	ss E	xec	utio	ns																	
	Process Attributes	PE 1	PE 2	PE 3	PE 4	PE 5	PE 6	PE 7	PE 8	PE 9	PE 10	РЕ 11	PE 12	PE 13	PE 14	PE 15	PE 16	PE 17	PE 18	PE 19	PE 20	PE 21	PE 22	
	1.1 <input 1=""/>	0	o																					
1	1.2 <input 2=""/>	0	o	···· ···																				
_	2.1 <output 1=""></output>	0	ο	····																				
2	2.2 <output 2=""></output>		0																					
	3.1 <activity 1=""></activity>	0	о	····																				
3	3.2 <activity 2=""></activity>	0	o	····																				
ľ	3.3 <activity 3=""></activity>	0	0	 																				
	3.4 <activity 4=""></activity>		0	 																				
4	4.1 <role 1=""></role>	0	0	 																				
-	4.2 <role 2=""></role>	о	о	····																				
	5.1 < Tools and	0	0																					
5	5.2 <tools and<="" th=""><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></tools>																							

Figure 3.6 Process Similarity Matrix (PSM) In Literature

In this study we used PSM a little bit different from the utilization in literature (Figure 3.7). We transposed the matrix vectors. The process executions were horizontally collected since this structure was more convenient to combine with collected defect and product data. In other words, this way provided

straightforwardness since also in the spreadsheet that consisted of the defect and product data, the metric attributes were in column vector against which regarding process executions exist. Besides, we entered "1" or "0" instead of "o" or "". Thus, "1" and "0" scaling could be identified in numeric measurement scale by Weka tool. Process enactment data identified as numeric could be clustered by machine learning clustering technique. PSM sheets were collected from issue tracking tool by extracting historical defect management process data such as defect status updates and the roles of the personnel who had updated the defect status.

					F	Process A	ttributes					
	1 In	puts	2 Ou	tputs		3 Activ	/ities		4 R	oles	5 Too Techr	ls and liques
Proce ss Execu tions	1.1 <input 1></input 	1.2 <input 2></input 	2.1 <output 1></output 	2.2 <output 2></output 	3.1 <activity 1></activity 	3.2 <activity 2></activity 	3.3 <activity 3></activity 	3.4 <activi ty 4></activi 	4.1 <role 1></role 	4.2 <role 2></role 	5.1 <tools and Techniqu es 1></tools 	5.2 <tools and Techniqu es 2></tools
PE1	1	1	1	0	1	1	1	0	1	1	1	0
PE2	1	1	1	1	1	1	1	1	1	1	1	1
PE3												
PE4												
PE5												
PE6												
PE7												
PE8												
PE9												
PE10												
PE11												
PE12												
PE13												
PE14												
PE15												
PE16												
PE17												
PE18												
PE19												
PE20												
PE21												
PE22												

Figure 3.7 Process Similarity Matrix (PSM) In This Study

3.4 Data Cleaning and Preprocessing

In this phase the redundant data and attributes whether there are in data set are removed from data set to avoid from overfitting and multicollinearity during machine learning analysis techniques. The redundant data might be the rows that have missing values or attribute columns that give same information. Removing of redundant attributes is called as data reduction. Some approaches such as Principal Component Analysis (PCA) can be used for data reduction too [68]. By using PCA, redundant attributes are composed and attribute number decreases by providing new attributes, andat the end more meaningful and explanatory attributes can be obtained. Otherwise overfitting [69] problem is common in machine learning techniques.

In numeric scale, attribute data should be discretized before analysis to obtain more meaningful analysis results. Some machine learning classification approaches such as C4.5 decision tree does not accept a numeric scaled attribute as class attribute for classification analysis. There are several techniques used for discretization such as equal-width or equal-frequency [68] in Weka. Applying clustering before discretization is another way to determine discretized bin number. Because of these reasons data cleaning and preprocessing phase is important for machine learning techniques.

Since in this study we needed only data in "Defect" category for prediction model, the issues recorded as "Change" were removed from data set.

All defects detected during test activities are recorded to issue tracking tool although all defects detected during review activities are not stored in tool. Therefore, the detected defects except during test activities were removed from data set. Only defect data detected during tests was taken into account after data cleaning.

We had to discretize defect open duration attribute to set as class attribute in machine learning classification. We discretized this attribute by using equal-width

method before (i.e. 0-5, 5-10, 10-15). Before discretization operation, we clustered open duration data by using K-Means with Euclidean distance technique to display how many clusters would be better to contain. Screenshots of analysis views are provided in Appendix-C and Appendix-E.

3.5 Clustering According to Process EnactmentData Approach

In machine learning if the user has no idea about data set, s/he should use unsupervised methods for grouping of data. Since s/he does not know which attribute can be considered as independent variable to set as class attribute. One example of unsupervised methods is clustering. In clustering method, the user do not have to set an attribute as class attribute.

In this study the process enactment data was examined in Weka and by clustering, similar process attribute columns were removed. With Weka tool the row data regarding process executions that had same process attributes was separated in different clusters.SimpleKMeans approach was used and the difference between process executions was obtained. According to cluster number automatically given by Weka, the separate Excel sheets were manually prepared for each cluster. Clustering according to process enactment approach was applied only in case studies 1B and 2B since they were the only case studies that contain process enactment data for analysis. Screenshots of analysis views are provided in Appendix-C and Appendix-E.

3.6 Analysis

When evaluated with the presence of high volume data stored in software engineering tools, it has been observed that data mining applications over the software data are being increased especially in recent years [12].

Machine learning classification approaches are utilized for the purposes of generating prediction models. Mostly used techniques are Bayesian Belief Networks (BBN), Multilayer Perceptron, Logistics Regression and Decision Trees. Despite the fact that there are a lot number of divergent studies related with

using machine learning techniques for building prediction models, there is not any model technique defined as the best prediction approach or any way to apply in sequential manner described as the best method. Therefore, the studies in literature can be successful only by comparing their selected techniques among themselves and assuming the one that has the most accurate results as the best model.

Weka gives performance evaluation values for model validation. In addition, there are other validation methods such as using cross-validation or separating the data into training and test data sets [68].

In this study we chose defect open duration metric as dependent variable for classification analysis since this metric was directly related with defect management process and product quality status.

Bayesnet, Multilayer Perceptron, Logistic and C4.5 Tree machine learning analysis approaches [68] were performed by keeping defect open duration metric as class attribute (dependent variable). By selecting these approaches for analysis, we paid attention to apply machine learning techniques from different categories.

CHAPTER4

CASE STUDIES

4.1 Case Study 1 (Project-1 Data)

Case Study 1A was conducted with the data of Project-1 (for the characteristics of Project-1 please refer to Section 3.3). In this case study firstly, only defect and product data were used for analysis. After case study 1A had been completed, we performed case study 1B with applying same analysis approaches but this time we used both defect and product data, and process enactmentdata of Project-1.

4.1.1 Case Study 1A (Project-1)

GQM Tree was prepared as shown in Table 4.1 after the data fields which the basic metrics were tracing to our goal in issue tracking tool database had been examined. The metric descriptions are provided in Table 4.2.

GOAL	QUESTION NO	QUESTION	ANALYSIS METHOD	DERIVED METRIC	BASIC METRIC NO	BASIC METRIC
To understand if there is effect of enactment context on software product defectiveness.	4.1	What is software product defectiveness prediction accuracy without using process enactment data?	Bayesnet, Logistic, C4.5 Tree, Multilayer Perceptron Machine Learning Techniques	Defect Data: open duration (closed date- created date)	4.1.1	Defect and Product Data: source component, closed date, created date, detected test type, product version, product SLOC, product complexity, reproducibility, project phase

Table 4.1 GQM for Case Study 1A

Metrics	rics Metric Description			
	The time starting with the creation of the			
Remaining Open	defect and finishing with the closure of the			
	defect. Calculated by the difference of			
Duration	defect closed date and defect created date.			
	Unit is number of days.	Absolute		
Source Component	The component name of the defect detected.			
	Component name can be component-A,			
	component-B, component-C, component-D			
	or component-E. Filled by the issue			
	tracking tool automatically when the tester			
	record the defect.	Nominal		
	The date when the defect is detected. Filled			
Created Date	by the issue tracking tool automatically			
	when the tester record the defect.	Interval		
Closed Date	The date when the defect is closed. Filled			
	by the issue tracking tool automatically			
	when the project manager change the status			
	of the defect as "Closed".	Interval		
	The name of test type during which the			
Test Type	defect is detected. Entered by tester to the			
	issue tracking tool.	Nominal		
	The version of the software product which			
Product Version	the defect is detected. Entered by tester to			
	the issue tracking tool.	Ordinal		
SLOC	C The size of the product version where the			
(Source Lines of	defect is detected. Collected from			
Code)	configuration tool by using Locmetrics tool.	Absolute		
	The McCabe complexity of the product			
Complexity	version where the defect is detected.			
Complexity	Collected from configuration tool by using			
	Locmetrics tool.	Absolute		
Reproducibility	The repeatability of the defect detected.			
	Entered by tester to the issue tracking tool.	Nominal		
	The project phase where the defect			
Project Phase	detected. Collected manually by domain			
~	expert.	Nominal		

Table 4.2 Defectand Product Related Metric Descriptions for Case Study 1A

We filled MUQ shown in Figure 3.3 and 3.4 for basic and derived metrics (filled questionnaires are provided in Appendix-B). Afterrating results, we had idea about the usability of the metric. According to MUQ results, all basic metrics and

derived metrics of Project-1 were classified as "partially usable". Since MUA-1 is N, MUA-2 and MUA-3 are F, and MUA-4 is P.

Detected project phase data manually collected by using project's archival data such as project meeting minutes, and expert opinions. Source component, closed date, created date, test type, product version and reproducibility metrics' data had already been stored in issue tracking tool. These data directly extracted from tool database. Source lines of code (SLOC) and complexity metrics' data are calculated by LocMetrics and manually entered into spreadsheet that includes defect data. Open duration metric data was calculated in the one column of the spreadsheet. All defect and product data were recorded in an Excel file (Appendix-B).

Data Excel file converted to .csv file format to be analyzed in Weka.

Open duration attribute had to be discretized, in other words the continuous scale of this attribute had to be transformed to discrete scale to identify as class attribute (classifier). Before discretization operation, we clustered open duration data with K-Means technique to display how many clusters it contains. After trials with 3, 4, 5 and 6 number of clusters, we observed that the 5-clustered data set denotes the most frequency equivalent within clusters than others. Therefore, we discretized open duration data to five equal-width clusters as "0-27", "27-54", "54-81", "81-108", and "108-135" days. Screen views of the operation implemented in Weka are provided in Appendix-B.

After we transformed class attribute to nominal scale by discretization, we used Weka classification techniques by choosing defect open duration attribute as dependent attribute (class attribute). We applied Multilayer Perceptron, Bayesian Belief Networks, Logistic Regression and C4.5 Decision Tree (J48) machine learning techniques. We used 10-folds validation technique because of its high accurateness rate. Screen views of the operation implemented in Weka are provided in Appendix-B.

Findings from the study:

We observed that 296 data points are sufficient to obtain confident prediction results. Since Project-1 is newly completed and all personnel who had developed the project software still exist in company, expert opinions increased the reliability of the data and results.

Correctly classification performance values of the generated models are given below. The other performance values of the models are provided in Appendix-B. Multilayer perceptron gave the best performance values compared with other machine learning approaches.

- Multilayer perceptron machine learning technique validated with 10-folds gives 95% correctly classified instances value.
- Bayesian networks machine learning technique validated with 10-folds gives 85% correctly classified instances value.
- Logistic machine learning technique validated with 10-folds gives 82% correctly classified instances value.
- J48 decision tree machine learning technique validated with 10-folds gives 92% correctly classified instances value.

To complete this case study, we spent 5 person-days. The effort includes applying the approach, performing the analyses, and interpreting the results. If the product size and complexity metrics had previously been collected in the same Excel sheet with defect data and project phase metric had been recorded in real time during creating defect in issue tracking tool, spent effort for this case study could have been lower than now. The complete set of Weka outputs are provided in Appendix-B.

4.1.2 Case Study 1B (Project-1)

GQM Tree was prepared shown in Table 4.3.

Table 4.3 GQM for Case Study 1B

GOAL	QUESTION NO	QUESTION	ANALYSIS METHOD	DERIVED METRIC	BASIC METRIC NO	BASIC METRIC
To understand if there is effect of process enactment on software product defectiveness.	4.1	What is software product defectiveness prediction accuracy with using process enactment data?	Bayesnet, Logistic, C4.5 Tree, Multilayer Perceptron Machine Learning Techniques	Defect Data: open duration (closed date- created date)	4.1.2	Defect and Product Data: source component, closed date, created date, detected test type, product version, product SLOC, product complexity, reproducibility, project phase Process Enactment Data: defect management process attributes

We filled out PER to identify all alternative process attributes of the process executions (shown in Figure 3.5). PER form was filled by interviewing with Configuration Management Responsible personnel of the company. Defect management process is performed by issue tracking tool and with the monitoring and control of Configuration Management Responsible personnel in company.
Process Execution Record (Internal Attributes)

	Proce	ess Name:	Issue Management		Recorded On:	26.03.2012						
	Proce	ess Execution No:	N/A		Recorded By:	Damla Sivrioğlu						
1.	Inputs	Please list the inputs	to the process execution.									
	No	Name		Des	cription							
	1	Defects										
	2	Change requests										
2.	Outpu	ts: Please list the outp	uts from the process exec	ution.								
	No	Name		Des	Description							
	1	Updated product vers	ion									
3 🕮	Activit	ies. Please list in seru	oncothe activities that we	ro norfr	med while execut	ting the process						
J. <u>+</u>	No	Name	ence are acavated that we	Des	cription	ang are process.						
	1	Assign defect										
	2	Defect resolution (Fix	issues)									
	3	Not verified for secon	dtime									
	4	Defect verification										
	5	Close defect										
4	Roles	Please list the roles th	at were allocated respons	ibilities	in process executi	on						
	No	Name		Des	cription							
	1	ProjectManager		Tra	ckissues, Fix issue	s						
	2	Configuration Manag	er	Tra	ckissues							
	3	Developer		Fixi	ssues							
	4	Modelling and Graph	ics Designer	Fixi	Fixissues							
	5	Tester		Оре	Open issues							
5.	Tools	and Techniques: Plea	se list the tools and techn	iques th	at are used to sup	port process execution.						
	No	Name		Des	cription							

No	Name	Description
1	Redmine	Issue tracking tool
2	Excel	Version Description List is in Excel format.
3	SVN	Configuration management tool
	I	I

Figure 4.1 PER for Case Study 1B

After completing PER form, same process attributes were entered into PSM columns and process execution values were filled in PSM shown in Figure 3.7 for each defect. Process attributes were given with abbreviations starting with "dm" phrase, which means "defect management", in PSM in order to ease reading of data file when opened in Weka. Because of place constraint, only 21 of the 296 data points could be shown in Figure 4.2.

								Proc	cess At	tribu	tes						
		1 Inj	puts	2 Outp uts		3 /	Activiti	es			4	Role	s		5 Te	Tools ar echnique	nd es
		1.1 <in put 1></in 	1.2 <in put 2></in 	2.1 <outp ut 1></outp 	3.1 <acti vity 1></acti 	3.2 <acti vity 2></acti 	3.3 <acti vity 3></acti 	3.4 <acti vity 4></acti 	3.5 <acti vity 5></acti 	4.1 <r ole 1></r 	4.2 <r ole 2></r 	4.3 <r ole 3></r 	4.4 <r ole 4></r 	4.5 <r ole 5></r 	5.1 <tools and Techni ques 1></tools 	5.2 <tools and Techni ques 2></tools 	5.3 <tools and Techni ques 3></tools
Proce ss Execu tions	Def ect No	dmi 1	dml 2	dm01	dmA 1	dmA 2	dmA 3	dmA 4	dmA 5	dm R1	dm R2	dm R3	dm R4	dm R5	dmT1	dmT2	dmT3
PE1	1	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE2	2	1	0	0	1	1	0	0	1	1	0	1	0	1	1	1	1
PE3	3	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE4	4	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE5	5	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
PE6	6	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE7	7	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE8	8	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE9	9	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1
PE10	10	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE11	11	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE12	12	1	0	0	1	0	0	0	1	1	0	1	0	1	1	1	1
PE13	13	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE14	14	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE15	15	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE16	16	1	0	0	1	0	0	0	1	1	0	1	0	1	1	1	1
PE17	1/	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE18	18	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE19	20	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE20	20	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
PE22	22	1	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1
	23	1	0	0	1	0	0	0	1	1	0	1	0	1	1	1	1

Figure 4.2 PSM for Case Study 1B

To prevent multicollinearity during the analysis in Weka, we should remove redundant process attributes, if exists, from spreadsheet. When we examined PSM, we observed that dmA1, dmA5, dmR1, dmR3 and dmR5 had displayed same behaviors. In other words, assigning personnel and closing defect activities had been implemented for all 296 process executions, and project manager, developer and tester personnel had performed their roles in all 296 process executions. Since dmR3 had alone fulfilled the characteristics (differences among executions) of these process attributes, we kept only dmR3 from these sixattributes for the analysis. Additionally, dmI1 and dmI2 do not give any information for analysis. Since, they do not differ in values among executions. In other words, since process input called "defect" was only input category that we had taken into account for our study, we had ignored the process input execution data categorized as "change" request. Therefore, we do not include dmI2 for our analyses. After data cleaning, we had an Excel file that consists of dmO1, dmA2, dmA3, dmA4, dmR2, dmR3 and dmR4 process attributes described in Figure 4.1.

		Measurement
Metrics	Metric Description	Scale
dmO1	New software version is the output of defect	
unioi	management process.	Nominal
	Developer response is one of the activities of defect	
dmA2	management process. It means that developer has	
	resolved the defect.	Nominal
	Not verified is one of the activities of defect	
dmA3	management process. It means that tester has tested	
	resolved defect but can not verified for second time.	Nominal
	Defect verification is one of the activities of defect	
dmA4	management process. It means that tester has tested	
	resolved defect and verified.	Nominal
	Configuration manager personnel is one of the roles	
dmP2	of defect management process. This personnel is	
unnx2	responsible of configuration control of software	
	product versions.	Nominal
	Developer personnel is one of the roles of defect	
dmR3	management process. This personnel is responsible	
	of develop software product and fix the defects.	Nominal
	Graphic designer is one of the roles of defect	
dmR/	management process. This personnel is responsible	
ullix+	of developing graphics of software product and fix	
	the defects.	Nominal

Table 4.4 Process Enactment Metric Descriptions for Case Study 1B

We combined collected defect, product and process enactment data in an Excel file spreadsheet.

We used K-Means and Euclidean Distance clustering technique and clustered the data. We obtained seven clusters which were called as c0, c1, c2, c3, c4, c5 and c6

in the rest of the case study. The differences of clusters are provided in Table 4.5. Implemented clustering steps are provided in Appendix-C.

	-0											
Cluster Name	CU											
_	2 Outputs	2.2.4.1	3 Activ		4.2 (Dala	4 R	Dies					
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	12 dm01	2> dm/2	<u>حد</u> dm/3	42 dm//	dmR2	dmR3	42 dmR/					
	1	1	0	1	1	1	0					
	1	1	1	1	1	1	0					
FAF2	1	I	I	I								
Cluster Name	c1											
	2 Outputs		3 Activities									
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1> dm01	2> dm/\2	3> dm/3	4> dm \/	Z> dmP2	3> dmP2	4> dmP4					
	unioi	uniaz	uiiAS	unia	unitz	units	unin ,					
PAP1	1	1	0	1	1	0	1					
PAP2	1	1	1	1	1	0	1					
PAP3	1	1	1	1	0	0	1					
PAP4	0	0	1	1	1	0	1					
Cluster Name	c2											
	2 Outputs		3 Activities			4 Roles						
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1>	2>	3>	4>	2>	3>	4>					
Pattern (PAP)	dm01	dmA2	dmA3	dmA4	dmR2	dmR3	dmR4					
PAP1	1 1		0	1	0	1	0					
Cluster Name	с3											
	2 Outputs 3 Activities 4 Roles											
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1>	2>	3>	4>	2>	3>	4>					
Pattern (PAP)	dm01	dmA2	dmA3	dmA4	dmR2	dmR3	dmR4					
PAP1	1	1	0	1	1	0	0					
Cluster Name	c4											
	2 Outputs		3 Activities			4 Roles						
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1>	2> dm 0.2	3>	4>	2> dm02	3>	4>					
	dilioi	umaz	ulliAS	ulliA4	ullikz	uiiks	ullik4					
PAP1	1	1	1	1	1	0	0					
Cluster Name	c5											
	2 Outputs		3 Activities			4 Roles						
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1>	2>	3>	4>	2>	3>	4>					
Pattern (PAP)	dm01	dmA2	dmA3	dmA4	dmR2	dmR3	dmR4					
PAP1	0	1	0	0	0	1	0					
Cluster Name	Сб											
	2 Outputs		3 Activities			4 Roles						
Process	2.1 <output< th=""><th>3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<></th></output<>	3.2 <activity< th=""><th>3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<></th></activity<>	3.3 <activity< th=""><th>3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<></th></activity<>	3.4 <activity< th=""><th>4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<></th></activity<>	4.2 <role< th=""><th>4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<></th></role<>	4.3 <role< th=""><th>4.4 <role< th=""></role<></th></role<>	4.4 <role< th=""></role<>					
Attributes	1>	2>	3>	4>	2>	3> 4>						
Pattern (PAP)	dm01	dmA2	dmA3	dmA4	dmR2	dmR3	dmR4					
PAP1	0	0	0	0	0	1	0					
PAP2	0	0	1	0	0	1	0					

Table 4.5 Process Attributes Patterns for Case Study 1B Clusters

We separated data Excel sheet to clusters and prepared a separate .csv file for each cluster. At the end we obtained the files shown in Figure 4.3. Each of these files included defect, product and process enactment metrics of the related defects.

الله data_with_process_context_c0 الله data_with_process_context_c1 الله data_with_process_context_c2 الله data_with_process_context_c3 الله data_with_process_context_c4 الله data_with_process_context_c5 الله data_with_process_context_c6

Figure 4.3 Clustered Metric Files for Case Study 1B

After clustering, we applied Multilayer Perceptron, Bayesian Belief Networks, Logistic Regression and C4.5 Decision Tree (J48) machine learning techniques for each clusterseparately. During these analyses, we identified open duration as class attribute. Screen views of the operation implemented in Weka are provided in Appendix-C.

Findings from the study:

We observed that the history data stored by issue tracking tool is beneficial to collect process enactment data. We collected process enactment data by firstly filling PER to identify process attributes. These process attributes can be identified easier by reviewing history data in tool database since all process activity alternatives are stored with their dates and the personnel who perform the activity. For example, when any personnel updates the defect status as "verified", the tool constitutes a record that "Defect status was updated by <personnel name> on <date>." in database. This process history data is used to fill out PSM for each defect record, in other words for each process execution.

Correctly classification performance values of the generated models for cluster-0 are given below. The other performance values of the models and the clusters are

provided in Appendix-C. Bayesian networks gave the best performance values compared with other machine learning approaches.

- Multilayer perceptron machine learning technique validated with 10-folds gives 96% correctly classified instances value for cluster 0.
- Bayesian networks machine learning technique validated with 10-folds gives 97% correctly classified instances value for cluster 0.
- Logistic machine learning technique validated with 10-folds gives 95% correctly classified instances value for cluster 0.
- J48 decision tree machine learning technique validated with 10-folds gives 96% correctly classified instances value for cluster 0.

Since clusters 3, 4 and 5 include low number of data, we could not apply machine learning techniques to them. If the cluster number is decreased or we have more data points, this issue can be solved.

To complete this case study, we spent 10 person-days. The effort includes applying the approach, performing the analyses, and interpreting the results. If the process enactmentdata had previously been collected or the process history data could automatically be extracted by a query from issue tracking tool, spent effort for this case study could have been lower than now. In other words, the most important reason of high spent effort is that we have collected process enactment data by entering each of 296 defects in tool and recording the history data to Excel sheet. The complete set of Weka outputs are provided in Appendix-C.

4.1.3 Results Comparison for Case Study 1 (Project-1)

According to Table 4.5, the characteristics of clusters can be described as follows in terms of process attribute patterns;

• Cluster 0 includes the metrics of process executions through which an updated product version is obtained as output, defect resolution and defect verification activities are implemented, and configuration manager and developer perform their roles. But, modeling and graphics designer does not perform his role.

- Cluster 1 includes the metrics of process executions through which defect verification activity is implemented, and modeling and graphics designer perform his role. But, developer does not perform his role.
- Cluster 2 includes the metrics of process executions through which an updated product version is obtained as output, defect resolution and defect verification activities are implemented, and developer performs hisrole. But, configuration manager and modeling and graphics designer do not perform their roles.
- Cluster 3 includes the metrics of process executions through which an updated product version is obtained as output, defect resolution and defect verification activities are implemented, and configuration manager performs his role. But, developer and modeling and graphics designer do not perform their roles.
- Cluster 4 includes the metrics of process executions through which an updated product version is obtained as output, defect resolution, not verified for second time and defect verification activities are implemented, and configuration manager performs his role. But, developer and modeling and graphics designer do not perform their roles.
- Cluster 5 includes the metrics of process executions through which defect resolution activity is implemented, and developer performs his role. But, configuration manager and modeling and graphics designer do not perform their roles.
- Cluster 6 includes the metrics of process executions through which no activities documented in PER are implemented, and only developer performs his role. In only one of the 425 executions not verified for second time activity is implemented. It means that in one defect management process execution the defect in resolved status could not be verified during second test repetition by test specialist.

We observed that generally the analysis results of clustered data sets with process enactmentare more accurate than data set without process enactmentas shown in Table 4.6.

Number of												
instances			Correctly	Incorrectly	Kanna	Mean	Root mean	Relative				
(data points)	Data set	Method	Instances	Instances	statistic	error	error	error				
·	Cluster 0	Multilayer										
	Data	Perceptron	96,43%	3,57%	94,86%	1,70%	10,37%	6,06%				
112	Process	Bayesnet	97,32%	2,68%	96,16%	1,40%	10,45%	4,98%				
	Enactme	Logistic	94,64%	5,36%	92,28%	2,14%	14,64%	7,63%				
	nt)	J48	95,54%	Incorrectly fied ces Incorrectly Instances Kappa statistic Mean absolute error Root mean squared error 6 3,57% 94,86% 1,70% 10,37% 6 2,68% 96,16% 1,40% 10,45% 6 2,68% 92,28% 2,14% 14,64% 6 4,46% 93,55% 2,15% 11,47% 6 15,49% 79,06% 7,35% 24,39% 6 19,72% 73,61% 8,31% 27,57% 6 18,31% 75,58% 7,19% 26,46% 6 14,08% 80,95% 7,41% 21,94% 6 4,29% 92,13% 3,61% 14,75% 6 8,57% 83,48% 5,53% 21,94% 6 10,00% 81,04% 6,54% 25,37% 6 17,14% 64,87% 17,21% 31,64% 126 are between 81-108) Interverter Interverter Interverter Interverter 126 are between 81-108) Interver	7,64%							
	Cluster 1	Multilayer	Q4 E10/	15 400/	70.06%	7 250/	24 200/	61 410/				
of instances (data points) 112 71 70 26 5 5 1 1 11 11 296	(With	Perceptron	04,01%	15,49%	79,00%	7,35%	24,39%	01,41%				
	Process	Dayeshet	80,28%	19,72%	73,01%	8,31%	27,57%	09,87%				
	Enactme	Logistic	81,69%	18,31%	75,58%	7,19%	26,46%	23,84%				
-	nt) Cluster 2	J48 Multilavor	85,92%	14,08%	80,95%	7,41%	21,94%	Relative absolute error 6,06% 4,98% 7,63% 7,64% 61,41% 69,87% 23,84% 24,57% 9,76% 14,96% 17,70% 46,55% 9,16% 0,07% 6,58% 0,00% 7,80% 18,36% 22,78%				
70	Data	Perceptron	95.71%	4.29%	92.13%	3.61%	14.75%	9.76%				
	(With	Bavesnet	91.43%	8.57%	83.48%	5.53%	21.94%	14.96%				
10	Process	Logistic	90.00%	10.00%	81.04%	6.54%	25.37%	17,70%				
	Enactme	.148	82.86%	17 14%	64 87%	17 21%	31 64%	46 55%				
	Cluster 3	Multilayer	02,0070	17,1470	04,0170	17,2170	01,0470	40,0070				
	Data	Perceptron	N/A (all 26	are between 8	1-108)							
26	(With	Bayesnet	N/A (all 26	I/A (all 26 are between 81-108)								
	Fnactme	Logistic	N/A (all 26	are between 8	1-108)							
	nt)	J48	N/A (all 26									
5	Cluster 4	Multilayer										
	Data	Perceptron	N/A (only 5									
	Process	Bayesnet	N/A (only 5	data points)								
	Enactme	Logistic	N/A (only 5	data points)			Root mean squared error Relative absolute error 10,37% 6,06% 10,45% 4,98% 14,64% 7,63% 11,47% 7,64% 24,39% 61,41% 27,57% 69,87% 26,46% 23,84% 21,94% 24,57% 14,75% 9,76% 21,94% 14,96% 25,37% 17,70% 31,64% 46,55% 7,04% 9,16% 0,06% 0,07% 13,14% 7,80% 20,81% 18,86% 26,16% 22,78%					
	nt)	J48	N/A (only 5	data points)								
	Cluster 5	Multilayer	N/Δ (only 1	data noint 81	.108)		Root mean squared error Relative absolute error 10,37% 6,06% 10,45% 4,98% 14,64% 7,63% 11,47% 7,64% 24,39% 61,41% 27,57% 69,87% 26,46% 23,84% 21,94% 24,57% 14,75% 9,76% 21,94% 14,96% 25,37% 17,70% 31,64% 46,55% 1 1 1 1 1 1 1 1 1 1 25,37% 17,70% 31,64% 46,55% 1 1 1 1 1 1 1 1 1 1 1 1 20,00% 0,00% 0,10% 6,58% 0,00% 0,00% 13,14% 7,80% 20,81% 18,35%					
70 26 5 1 11	(With	Bayesnet	N/A (only 1	data point 81	108)							
1	s Method Correctly Classified Instances Incorrectly Classified Instances Kappa statistic Mean absolute error Cluster 0 Data (With Process Enactme Multilayer Perceptron 96,43% 3,57% 94,86% 1,70% With Process Enactme Logistic 94,64% 5,36% 92,28% 2,14% J48 95,54% 4,46% 93,55% 2,15% Cluster 1 Data (With Process Enactme Multilayer Perceptron 84,51% 15,49% 79,06% 7,35% Bayesnet 80,28% 19,72% 73,61% 8,31% 7,9% Cluster 1 Data (With Process Enactme Perceptron 95,71% 4,29% 92,13% 3,61% Proceptron 95,71% 4,29% 92,13% 3,61% Proceptron 90,00% 10,00% 81,04% 6,54% Logistic 90,00% 10,00% 81,04% 6,54% Data (With Process Enactme Logistic N/A (all 26 are between 81-108) Perceptron J48 N/A (all 26 are between 81-108) Perceptron N/A (all 26 are between											
	Enactme		N/A (only 1	data point 01	100)							
	nt) Cluster 6	J48 Multilaver	N/A (only 1	data point 81-	108)							
	Data	Perceptron	100,00%	0,00%	100,00%	4,02%	7,04%	9,16%				
112 71 70 26 5 1 1 11 296	(With	Bayesnet	100,00%	0,00%	100,00%	0.03%	0,06%	0.07%				
	Process	Logistic	100,00%	0,00%	100,00%	0,03%	0,10%	6,58%				
	nt)	J48	100.00%	0.00%	100.00%	0.00%	0.00%	0.00%				
	Data	Multilayer		-,,-	,	3,0070	-,,	5,0070				
	Without	Perceptron	94,93%	5,07%	93,38%	2,40%	13,14%	7,80%				
296	Process	Bayesnet	85,14%	14,86%	80,54%	5,79%	20,81%	18,86%				
	Enactme	Logistic	82,43%	17,57%	76,90%	7,00%	26,16%	22,78%				
112 71 70 26 5 1 1 11 296	nt	J48	91,55%	8,45%	88,87%	5,63%	17,03%	18,35%				

Table 4.6 Results Comparison for Case Study 1

The average of correctly classified intances values of the methods applied to cluster 0 data is 95,98%. On the other hand the average of correctly classified intances values of the methods applied to data without process enactment is 88,51%. The correctly classified rate is 7,47% higher in cluster 0 than the result of

the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 0 data is 11,73%. On the other hand the average of root mean squared error values of the methods applied to data without process enactment is 19,29%. The root mean squared error is 7,55% lower in cluster 0 than the result of the data set that do not include process enactment.

The average of correctly classified intances values of the methods applied to cluster 1 data is 83,10%. The correctly classified rate is 5,41% lower in cluster 1 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 1 data is 25,09%. The root mean squared error is 5,81% higher in cluster 1 than the result of the data set that do not include process enactment. We could not obtain promising results from this cluster, the reason of this is the noise in cluster patterns that is seen in Table 4.5. To avoid this noise and achieve more accurate prediciton for cluster 1, one more clustering operation can be performed within cluster 1 data.

The average of correctly classified intances values of the methods applied to cluster 2 data is 90,00%. The correctly classified rate is 1,49% higher in cluster 2 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 2 data is 23,43%. The root mean squared error is 4,14% higher in cluster 2 than the result of the data set that do not include process enactment. Although, average correctly classified instances is high, we obtained a high average error value. The reason of this is the low error rate in J48 (C4.5) decision tree method, since this machine learning method needs more data point for a more accurate prediction than the other machine learning methods.

The average of correctly classified intances values of the methods applied to cluster 6 data is 100,00%. The correctly classified rate is 11,49% higher in cluster 6 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 6 data is

1,80%. The root mean squared error is 17,49% lower in cluster 6 than the result of the data set that do not include process enactment.

4.2 Case Study 2 (Project-2 Data)

Case Study 2A conducted with the data of Project-2 (for the characteristics of Project-2 please refer to Section 3.3). In this case study firstly only defect and product data used for analysis. After case study 2A had been completed, we performed case study 2B with applying same analysis approaches but this time we used both defect and product data, and process enactmentdata of Project-2.

4.2.1 Case Study 2A (Project-2)

GQM Tree shown in Table 4.7 was prepared after the data fields that were basic metrics tracing to our goal in issue tracking tool database had been examined. The metric descriptions are provided in Table 4.8. As different from Case Study 1, we identified detected software configuration unit (SCU) metric for Case Study-2. Since Project-2 includes several SCUs in its developed software product, this data might give important information for the patterns in data set. Second difference from Case Study 1 is that we selected Decision Table technique rather than Multilayer Perceptron. And we included Simple Logistic Regression rather that Logistic Regression, since we again wanted to validate our proposed method for various machine learning techniques.

GOAL	QUESTION NO	QUESTION	ANALYSIS METHOD	DERIVED METRIC	BASIC METRIC NO	BASIC METRIC
To understand if there is effect of process enactment on software product defectiveness.	4.2	What is software product defectiveness prediction accuracy without using process enactment data?	Bayesnet, SimpleLogistic, C4.5 Tree, Decision Table, Multilayer Perceptron Machine Learning Techniques	Defect Data: open duration (closed date-created date)	4.2.1	Defect and Product Data: detected SCU name, detected module name, closed date, created date, detected test type, product version, product SLOC, product complexity, reproducibility, project phase, source component

Matrian	Matria Description	Measurement
Metrics	Metric Description	Scale
D 0	The time starting with the creation of the defect and finishing	
Remaining Open	with the closure of the defect. Calculated by the difference of	
Duration	defect closed date and defect created date. Unit is number of	A I I I.
	days.	Absolute
Detected SCU	The name of the software configuration unit (SCU) where the	
Name	defect is detected. Entered by developer to the issue tracking	
	tool.	Nominal
Created Date	The date when the defect is detected. Filled by the issue	
	tracking tool automatically when the tester record the defect.	Interval
	The date when the defect is closed. Filled by the issue	
Closed Date	tracking tool automatically when the project manager change	
	the status of the defect as "Closed".	Interval
Test Tune	The name of test type during which the defect is detected.	
rest rype	Entered by tester to the issue tracking tool.	Nominal
Product Version	The version of the software product which the defect is	
rioduct version	detected. Entered by tester to the issue tracking tool.	Ordinal
SLOC		
(Source Lines of	The size of the product version where the defect is detected.	
Code)	Collected from configuration tool by using Locmetrics tool.	Absolute
	The McCabe complexity of the product version where the	
Complexity	defect is detected. Collected from configuration tool by using	
	Locmetrics tool.	Absolute
Danroduaibility	The repetability of the defect detected. Entered by tester to	
Reproducionity	the issue tracking tool.	Nominal
Drainat Dhaga	The project phase where the defect detected. Collected	
Project Phase	manually by domain expert.	Nominal
	The component name of the defect detected. Component	
Source Commonant	name can be BusinessManager, Form, GMManager, Report,	
Source Component	DBManager, Table and Menu-Template. Manually collected	
	by domain expert.	Nominal

Table 4.8 Defect and Product Related Metric Descriptions for Case Study 2A

We filled MUQ shown in Figure 3.3 and Figure 3.4 for basic and derived metrics (filled questionnaires are provided in Appendix-D). After obtained rating results, we had idea about the usability of the metric. According to MUQ results, all basic metrics and derived metric of Project-2 were classified as "partially usable". Since MUA-1 is N, MUA-2 and MUA-3 are F, and MUA-4 is P.

Detected project phase and source component data manually collected by using project's archival data such as project meeting minutes, and expert opinions. Detected module name, closed date, created date, test type, product version and reproducibility metrics' data had already been stored in issue tracking tool. These data directly extracted from tool database. Source lines of code (SLOC) and complexity metrics' data are calculated by LocMetrics and manually entered into spreadsheet that includes defect data. Open duration metric data was calculated in the one column of the spreadsheet. All defect and product data were recorded in an Excel file (Appendix-D).

Data Excel file was converted to .csv file format to be analyzed in Weka.

We discretized open duration data to seven equal-width clusters as "0-20", "20-40", "40-60", "60-80", "80-100", "100-120" and "120-140" days. Since, there were not any open duration value in "80-100" range, this cluster had no data.

We applied Decision Table, Bayesian Belief Networks, Simple Logistic Regression and C4.5 Decision Tree (J48) machine learning techniques by selecting open duration as class attribute. Screen views of the operation implemented in Weka are provided in Appendix-D.

Findings from the study:

We observed that 425 data points are sufficient to obtain confident prediction results. Since Project-2 had been completed a long time ago and several personnel who had developed the project software do not work for the company anymore, we believe that the reliability of the data collected by interviews might be lower than Case Study 1.

Correctly classification performance values of the generated models are given below. The other performance values of the models are provided in Appendix-D. Decision Table, Bayesian Networks, Simple Logistic and J48 Decision Tree gave the best performance values compared with other machine learning approaches.

- Decision Table machine learning technique validated with 10-folds gives 88% correctly classified instances value.
- Bayesian networks machine learning technique validated with 10-folds gives 89% correctly classified instances value.
- Simple Logistic machine learning technique validated with 10-folds gives 86% correctly classified instances value.

• J48 decision tree machine learning technique validated with 10-folds gives 89% correctly classified instances value.

To complete this case study, we spent 10 person-days. The effort includes applying the approach, performing the analyses, and interpreting the results. If the source component metric had previously been collected in the same Excel sheet with defect data and project phase metric had been recorded in real time during creating defect in issue tracking tool, spent effort for this case study could have been lower than now. The complete set of Weka outputs are provided in Appendix-D.

4.2.2 Case Study 2B (Project-2)

GQM Tree was prepared shown in Table 4.9.

GOAL	QUESTION NO	QUESTION	ANALYSIS METHOD	DERIVED METRIC	BASIC METRIC NO	BASIC METRIC
To understand if there is effect of process enactment on software product defectiveness.	4.2	What is software product defectiveness prediction accuracy with using process enactment data?	Bayesnet, SimpleLogistic, C4.5 Tree, Decision Table, Multilayer Perceptron Machine Learning Techniques	Defect Data: open duration (closed date-created date)	4.2.2	Defect and Product Data: detected SCU name, closed date, created date, detected test type, product version, product SLOC, product complexity, reproducibility, project phase, source component Process Enactment Data: defect management process attributes

Table 4.9 GQM for Case Study 2B

We filled out PER to identify all alternative process attributes of the process executions (shown in Figure 3.5). PER form was filled by interviewing with Project Manager personnel of the project (Figure 4.4).

Process Execution Record (Internal Attributes)

	Proc	ess Name:	Issue Management	Recorded On:	27.03.2012						
	Proc	ess Execution No:	N/A	Recorded By:	Damla Sivrioğlu						
1	Innuts	Please list the inputs	to the process execution								
	No	Name	to the process exceduor.	Description							
	1	Defects									
	2	Changerequests		Improvements							
2	Outpu	its: Please list the out	outs from the process execution	1.							
	No	Name		Description							
	1	Modified software		Targetversion							
3	Activit	ties: Please list in seq	uence the activities that were p	erformed while executi	na the process.						
	No	Name		Description							
	1	Adding explanation t	o defect								
	2	Requesting more fee	dback	From test specia	list by developer						
	3	Defect resolution (Is	sue implementing)								
	4	Defect rejection		Defect is rejecte	d and is not resolved.						
	5	Defect have not been	n tried again								
	6	Status changed as "	Resolved" by test specialist								
	7	"Subject" field of def	ect is changed								
	8	Adding additional pic	ture for explanation								
4.	Roles:	Please list the roles t	nat were all ocated responsibilit	ies in process executio	n.						
I	No	Name		Description							
	1	Project Manager		Tracks issues							
	2	Configuration Respo	nsible	Tracks issues							
	3	Developer		Implements issu	Jes						
	4	Test specialist		Opens issues							
5.	Tools	and Techniques: Ple	ase list the tools and technique	s that are used to supp	ort process execution.						
	No	Name		Description	Description						
	1	Redmine		Issuetrackingto	Issuetrackingtool						
	2	SVN		Configuration m	anagementtool						
	3	Visual Studio		Developmenter	Development environment						

Figure 4.4 PER for Case Study 2B

After completing PER form, same process attributes were entered into PSM columns and process execution values were filled in PSM shown in Figure 3.7 for each defect. Process attributes were given with abbreviations starting with "dm" (defect management) phrase in PSM in order to ease reading of data file when opened in Weka. Because of place constraint, only 21 of the 425 data points could be shown in Figure 4.5.

		Process Attributes																	
		1	1	2 Outp										4			5]	Γools a	nd
		Inp	uts	uts				3 Acti	vities				Ro	les			Те	chniqu	es
													4.	4.	4.	4.	5.1	5.2	5.3
		1.1	1.2										1	2	3	4	<tool< th=""><th><tool< th=""><th><tool< th=""></tool<></th></tool<></th></tool<>	<tool< th=""><th><tool< th=""></tool<></th></tool<>	<tool< th=""></tool<>
		<ln< th=""><th><in< th=""><th>2.1</th><th>3.1</th><th>3.2</th><th>3.3</th><th>3.4</th><th>3.5</th><th>3.6</th><th>3.7</th><th>3.8</th><th><r< th=""><th><r< th=""><th><r< th=""><th><r< th=""><th>s and</th><th>s and</th><th>s and</th></r<></th></r<></th></r<></th></r<></th></in<></th></ln<>	<in< th=""><th>2.1</th><th>3.1</th><th>3.2</th><th>3.3</th><th>3.4</th><th>3.5</th><th>3.6</th><th>3.7</th><th>3.8</th><th><r< th=""><th><r< th=""><th><r< th=""><th><r< th=""><th>s and</th><th>s and</th><th>s and</th></r<></th></r<></th></r<></th></r<></th></in<>	2.1	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	<r< th=""><th><r< th=""><th><r< th=""><th><r< th=""><th>s and</th><th>s and</th><th>s and</th></r<></th></r<></th></r<></th></r<>	<r< th=""><th><r< th=""><th><r< th=""><th>s and</th><th>s and</th><th>s and</th></r<></th></r<></th></r<>	<r< th=""><th><r< th=""><th>s and</th><th>s and</th><th>s and</th></r<></th></r<>	<r< th=""><th>s and</th><th>s and</th><th>s and</th></r<>	s and	s and	s and
		t pu	pu t	out	ivity	ivity	ivity	ivity	ivity	ivity	ivity	ivity	e	e	e	e	nique	nique	nique
		1>	2>	1>	1>	2>	3>	4>	5>	6>	7>	8>	1>	2>	3>	4>	s 1>	s 2>	s 3>
Proc																			
ess Exec	De												Ь	Ь	Ь	Ь			
ution	t	dm	dm	dmO	dm	dm	dm	dm	dm	dm	dm	dm	m	m	m	m			
S	No	11	12	1	A1	A2	A3	A4	A5	A6	A7	A8	R1	R2	R3	R4	dmT1	dmT2	dmT3
PE1	1	1	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	1
PE2	2	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE3	3	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE4	4	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE5	5	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE6	6	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE7	7	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE8	8	1	0	1	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1
PE9	9	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE10	10	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
PE11	11	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
PE12	12	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE13	13	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE14	14	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE15	15	1	0	1	0	0	0	0	0	1	1	0	1	1	0	1	1	1	1
PE16	16	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE17	17	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE18	18	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE19	19	1	0	1	0	0	0	0	0	1	0	0	1	1	0	1	1	1	1
PE20	20	1	0	1	0	0	1	0	0	0	0	0	1	1	0	1	1	1	1
PE21	21	1	0	1	0	0	0	0	0	0	1	0	1	1	0	1	1	1	1
PE22	22	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1
	23	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	1	1	1

Figure 4.5 PSM for Case Study 2B

To prevent multicollinearity during the analysis in Weka, we should remove redundant process attributes, if exists, from spreadsheet. When we examined PSM, we observed that I1 and I2 are same in all 425 executions; we have removed them from analysis.Since O1, R2 and R4 showed same behavior for each defect and there was little difference when reviewed all rows, we removed them from analysis. Since A7 and A8 showed same behavior for each defect and these activities had low impact on defect management process, we removed them from analysis. We specify these activities as having low effect on independent variable, open duration metric, because changing "subject" field (a field to fill in issue tracking tool) in defect record or adding additional picture to defect record have no technical context on the quality of final product and they are executions that are rarely seen during defect management process of whole project. Since all row data were same in R1, in other words project manager has role in for all 296 process executions, it is redundant to include it in analysis. Therefore, we removed R1 process attribute from analysis. Since T1, T2, T3 were used in all defect management process executions, they did not give additional information about process change through defect management. Therefore, T1, T2 and T3 were removed from analysis because of being redundant. After data cleaning, we had an Excel file that consisted of dmA1, dmA2, dmA3, dmA4, dmA5, dmA6 and dmR3 process attributes described in Figure 4.10.

Metrics	Metric Description	Measurement Scale
	Adding explanation to defect is one of the activities of defect	
	management process. It means that developer and/or tester	
dmA1	fills "additional explanation" field to give more detailed	
	information about the defect.	Nominal
	Requesting more feedback is one of the activities of defect	
dmA2	management process. It means that developer needs more	
	information about the defect before resolving it.	Nominal
1	Defect resolution is one of the activities of defect management	
dinA5	process. It means that developer has resolved the defect.	Nominal
	Defect rejection is one of the activities of defect management	
dmA4	process. It means that the defect record is examined and	
	decided that it is not a defect actually.	Nominal
	Defect have not been tried again is one of the activities of	
dmA5	defect management process. It means that the defect	
	recorded can not be repeated.	Nominal
	Status changed as "Resolved" by test specialist is one of the	
dmA6	activities of defect management process. It means that test	
	specialist has verified the resolution of the defect.	Nominal
	Developer personnel is one of the roles of defect management	
dmR3	process. This personnel is responsible of develop software	
	product and fix the defects.	Nominal

Table 4.10 Process Enactment Metric Descriptions for Case Study 2B

We combined collected defect, product and process enactmentdata in an Excel file spreadsheet (Appendix-E).

We used K-Means and Euclidean Distance clustering technique and separated the data into five clusters which were called as c0, c1, c2, c3 and c4 in the rest of the case study. The differences of clusters are provided in Table 4.11. Implemented clustering steps are provided in Appendix-E.

Cluster	c0							Cluster	c2						
Name								Name	~~						
		3 Activities 4 Roles					3 Activities					4 Roles			
Process Attributes Pattern (PAP)	3.1 <activity 1=""> dmA1</activity>	3.2 <activity 2=""> dmA2</activity>	3.3 <activity 3=""> dmA3</activity>	3.4 <activity 4=""> dmA4</activity>	3.5 <activity 5=""> dmA5</activity>	3.6 <activity 6=""> dmA6</activity>	4.3 <role 3=""> dmR3</role>	Process Attributes Pattern (PAP)	3.1 <activity 1=""> dmA1</activity>	3.2 <activity 2=""> dmA2</activity>	3.3 <activity 3=""> dmA3</activity>	3.4 <activity 4=""> dmA4</activity>	3.5 <activity 5=""> dmA5</activity>	3.6 <activity 6=""> dmA6</activity>	4.3 <role 3=""> dmR3</role>
PAP1	0	1	0	0	0	1	1	PAP1	0	1	0	0	0	0	1
PAP2	0	1	0	0	0	0	1	PAP2	0	1	0	0	0	1	1
PAP3	1	1	0	0	0	1	1	PAP3	0	1	0	0	1	1	1
PAP4	1	1	0	0	0	0	1	PAP4	0	1	1	0	0	1	1
PAP5	1	1	0	1	0	1	1	PAP5	0	0	0	0	0	0	0
PAP6	1	1	1	0	0	0	1	PAP6	0	1	0	0	0	1	0
PAP7	1	1	0	0	1	1	1	PAP7	1	1	0	0	0	0	1
PAP8	1	1	0	0	1	0	1	PAP8	1	1	0	0	0	1	1
PAP9	1	0	1	0	0	0	1	PAP9	1	0	0	0	0	0	1
PAP10	0	0	0	0	0	1	1	PAP10	1	1	1	0	0	1	1
PAP11	0	1	1	0	0	0	1	PAP11	1	1	0	1	0	1	1
PAP12	0	1	0	0	1	1	1	PAP12	1	0	0	0	0	0	0
PAP13	0	0	1	0	0	0	0								
PAP14	0	0	0	0	0	1	0								
PAP15	0	0	0	0	0	0	1								
Cluster Name	c1							Cluster Name	c3						
Desses			3 Activ	vities			4 Roles	Deserves			3 Act	ivities			4 Roles
Attributor				24		36	4.2	Attributor	24			34	35	26	4.2
Pattors	5.1 «Artivitu 4»	3.2 (Artivity 3-	2,5 <artivity 25<="" td=""><td>2,4 cArtivity As</td><td>2,3 cArtivity Es</td><td>D.D (Activity Ex</td><td>4.5 <dala 25<="" td=""><th>Pattern</th><td>0.1 Activity (S</td><td>3.2 (Artivity 3:</td><td>D.D «Artivitu 2»</td><td>2,4 «Artivity As</td><td>0.0 «Artivity Es</td><td>D.D «Artivity 6»</td><td>4.5 (Dolo 7:</td></dala></td></artivity>	2,4 cArtivity As	2,3 cArtivity Es	D.D (Activity Ex	4.5 <dala 25<="" td=""><th>Pattern</th><td>0.1 Activity (S</td><td>3.2 (Artivity 3:</td><td>D.D «Artivitu 2»</td><td>2,4 «Artivity As</td><td>0.0 «Artivity Es</td><td>D.D «Artivity 6»</td><td>4.5 (Dolo 7:</td></dala>	Pattern	0.1 Activity (S	3.2 (Artivity 3:	D.D «Artivitu 2»	2,4 «Artivity As	0.0 «Artivity Es	D.D «Artivity 6»	4.5 (Dolo 7:
(PAP)	dmA1	dmA2	dmA3	dmA4	dmA5	dmA6	dmD2		dmA1	dmA2	dmA3	dmA4	dmA5	dmA6	dmD2
PAP1	0	0	0	0	0	0	0	PAP1	0	0	1	0	0	0	1
PAP2	Ő	ŏ	Ő	ő	ő	ő	1	PAP2	0	ő	0	1	0	0	1
PAP3	0	0	0	1	0	0	1	PAP3	1	0	0	1	0	0	1
PAP4	0	0	0	0	0	1	0	PAP4	0	1	1	0	0	0	1
PAP5	0	0	0	1	0	1	1	PAP5	0	1	0	0	1	1	1
PAP6	1	0	0	1	0	0	1	PAP6	1	0	1	0	0	0	1
PAP7	1	1	0	1	0	1	1	PAP7	0	0	1	0	0	0	0
PAP8	1	0	0	0	0	0	1	PAP8	0	0	0	1	0	1	1
								PAP9	1	0	1	1	0	0	1

Table 4.11 Process Attributes Patterns for Case Study 2B Clusters

Cluster	c4										
Name											
	3 Activities										
Process	3.1	3.2	3.3	3.4	3.5	3.6	4.3				
Attributes	<activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><role< th=""></role<></th></activity<></th></activity<></th></activity<></th></activity<></th></activity<></th></activity<>	<activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><role< th=""></role<></th></activity<></th></activity<></th></activity<></th></activity<></th></activity<>	<activity< th=""><th><activity< th=""><th><activity< th=""><th><activity< th=""><th><role< th=""></role<></th></activity<></th></activity<></th></activity<></th></activity<>	<activity< th=""><th><activity< th=""><th><activity< th=""><th><role< th=""></role<></th></activity<></th></activity<></th></activity<>	<activity< th=""><th><activity< th=""><th><role< th=""></role<></th></activity<></th></activity<>	<activity< th=""><th><role< th=""></role<></th></activity<>	<role< th=""></role<>				
Pattern	1>	2>	3>	4>	5>	6>	3>				
(PAP)	dmA1	dmA2	dmA3	dmA4	dmA5	dmA6	dmR3				
PAP1	1	1	0	0	0	1	1				
PAP2	0	1	0	0	0	1	1				
PAP3	0	1	0	0	0	1	0				
PAP4	1	1	0	0	0	0	1				
PAP5	0	0	0	0	1	1	1				
PAP6	0	1	1	0	0	1	1				
PAP7	0	1	0	0	1	1	1				
PAP8	0	0	0	0	0	0	0				
PAP9	0	0	0	0	0	1	0				
PAP10	0	1	0	0	0	0	1				
PAP11	0	1	0	1	0	0	1				
PAP12	1	1	0	1	0	1	1				
PAP13	0	1	1	0	0	0	1				

After preparing five Excel files for clustered data sets, we applied Decision Table, Bayesian Belief Networks, Simple Logistic Regressionand C4.5 Decision Tree (J48) machine learning techniques. Screen views of the operation implemented in Weka are provided in Appendix-E.

Findings from the study:

We mentioned that Project-2 had been completed more previously than Project-1, and therefore collecting process enactment data was harder than the first project. Additionally, we could interview with lower number of personnel who developed project's software product. Besides, this project is an old project, and executed processes are so changeable. This is observed with the variety of the process attribute patterns provided in Table 4.11. It is seen that the clusters are more noisy than the ones of project one.

Correctly classification performance values of the generated models for cluster-0 are given below. The other performance values of the models and the clusters are provided in Appendix-E. Decision Table, Bayesian Networks, Simple Logistic and J48 Decision Tree were applied and Bayesian Networks gave the best performance values compared with other machine learning approaches. 10-folds technique was used for validation.

- Decision Table machine learning technique validated with 10-folds gives 95% correctly classifies instances value for cluster 0.
- Bayesian Networks machine learning technique validated with 10-folds gives 97% correctly classifies instances value for cluster 0.

- Simple Logistic machine learning technique validated with 72% of all data points allocated for training data set gives 96% correctly classifies instances value for cluster 0.
- J48 Decision Tree machine learning technique validated with 10-folds gives 94% correctly classifies instances value for cluster 0.

To complete this case study, we spent 10 person-days. The effort includes applying the approach, performing the analyses, and interpreting the results. If the process enactmentdata had previously been collected or the process history data could automatically be extracted by a query from issue tracking tool, spent effort for this case study could have been lower than now. In other words, the most important reason of high spent effort is that we have collected process enactmentdata by entering each of 425 defects in tool and recording the history data to Excel sheet. The complete set of Weka outputs are provided in Appendix-E.

4.2.3 Results Comparison for Case Study 2 (Project-2)

According to Table 4.11, the characteristics of clusters can be described as follow in terms of process attribute patterns;

- Cluster 0 predominantly includes the metrics of process executions through which status changed as "Resolved" by test specialist activity is implemented, and developer performs hisrole. But, adding explanation to defect, defect rejection and not tried again activities are not implemented.Requesting more feedback activity is seen in the 13% some of the executions.
- Cluster 1 predominantly includes the metrics of process executions through which defect rejection is implemented, and developer performs his role. But, adding explanation to defect, requesting more feedback, defect resolution, not tried again activities, status changed as "Resolved" by test specialist are predominantly not implemented.
- Cluster 2 predominantly includes the metrics of process executions through which requesting more feedback and status changed as "Resolved" by test specialist activities are implemented, and developer performs his role. But,

adding explanation to defect, defect rejection and not tried again activities are not implemented.

- Cluster 3 predominantly includes the metrics of process executions through which defect resolution activity are implemented, and developer performs his role except 25% executions.
- Cluster 4 predominantly includes the metrics of process executions through which requesting more feedback and status changed as "Resolved" by test specialist activities are implemented, and developer performs his role. But, adding explanation to defect activity is implemented in 18% of executions.

We observed that generally the analysis results of clustered data sets with process enactmentare more accurate than data set without process enactmentshown in Table 4.12. However, we can not say the same thing for cluster 4. Although it has the highest number of data points, its performance values are lower than the analysis without process enactment data.

Number of instances (data points)	Data set	Method	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Relative absolute error
· /	Cluster 0	Decision	05.040/	4.000/	05 000/	0.000/	45 740/	00.00%
	Data	Table	95,04%	4,96%	85,03%	8,92%	15,71%	69,02%
101	(With	Bayesnet	96,69%	3,31%	90,35%	1,18%	10,37%	9,10%
121	Process	Logistic	95,87%	4,13%	87,71%	5,90%	12,46%	45,62%
	nt)	J48	94,21%	5,79%	82,01%	2,96%	13,04%	22,88%
	Cluster 1 Data (With Process Enactme nt)	Decision Table	88,46%	11,54%	0,00%	11,35%	20,26%	99,55%
		Bayesnet	96,15%	3,85%	78,33%	1,93%	11,53%	16,95%
26		Simple Logistic	96,15%	3,85%	78,33%	5,67%	13,07%	49,77%
		J48	88,46%	11,54%	0,00%	5,20%	18,98%	45,58%
109	Cluster 2	Decision Table	92,66%	7,34%	80,78%	11,99%	19,54%	91,13%
	Data	Bayesnet	91,74%	8,26%	79,28%	2,98%	15,78%	22,61%
	(With Process Enactmo	Simple Logistic	90,83%	9,17%	74,16%	3,59%	14,54%	27,27%
	Enactme nt)	J48	93,58%	6,42%	82,82%	3,29%	13,42%	25,01%

Table 4.12 Results Comparison for Case Study 2

32	Cluster 3	Decision Table	87,50%	12,50%	66,84%	13,07%	21,61%	81,67%
	Data	Bayesnet	93,75%	6,25%	83,42%	3,24%	15,70%	20,27%
	(With Process Enactme nt)	SimpleLo gistic	81,25%	18,75%	53,62%	6,62%	23,53%	41,39%
		.148	78 13%	21.88%	47 66%	8 30%	24 48%	51 86%
137	Cluster 4 Data (With Process Enactme nt)	Decision Table	75,18%	24,82%	52,97%	14,22%	24,42%	77,89%
		Bayesnet	72,26%	27,74%	48,45%	10,13%	27,61%	55,47%
		SimpleLo gistic	70,80%	29,20%	43,91%	12,42%	27,70%	68,03%
		J48	70,80%	29,20%	41,36%	14,13%	26,98%	77,38%
425	Data	Decision Table	88,47%	11,53%	83,26%	10,52%	18,92%	46,34%
	Without	Bayesnet	88,94%	11,06%	84,00%	4,02%	18,36%	17,71%
	Process Enactme	SimpleLo gistic	88,24%	11,76%	82,70%	4,64%	17,95%	20,42%
	nt	J48	88,94%	11,06%	83,03%	4,80%	16,66%	21,12%

The average of correctly classified intances values of the methods applied to cluster 0 data is 95,45%. On the other hand the average of correctly classified intances values of the methods applied to data without process enactment is 88,65%. The correctly classified rate is 6,08% higher in cluster 0 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 0 data is 12,90%. On the other hand the average of root mean squared error values of the methods applied to cluster 0 data is 12,90%. On the other hand the average of root mean squared error values of the methods applied to cluster 0 data is 12,90%. On the other hand the average of root mean squared error values of the methods applied to data without process enactment is 17,97%. The correctly classified rate is 5,08% lower in cluster 0 than the result of the data set that do not include process enactment.

The average of correctly classified intances values of the methods applied to cluster 1 data is 92,31%. The correctly classified rate is 3,66% higher in cluster 1 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 1 data is 15,96%. The correctly classified rate is 2,01% lower in cluster 1 than the result of the data set that do not include process enactment.

The average of correctly classified intances values of the methods applied to cluster 2 data is 92,20%. The correctly classified rate is 3,56% higher in cluster

2than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 2 data is 15,82%. The correctly classified rate is 2,15% lower in cluster 2 than the result of the data set that do not include process enactment.

The average of correctly classified intances values of the methods applied to cluster 3 data is 85,16%. The correctly classified rate is 3,49% lower in cluster 3 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 3 data is 21,33%. The correctly classified rate is 3,36% higher in cluster 3 than the result of the data set that do not include process enactment. We could not obtain promising results from this cluster. To investigate the reason of this we reviewed data and observed that the cluster noise based on between process attibute patterns and defect open duration metric are high in cluster 3, although the number of data points is low.

The average of correctly classified intances values of the methods applied to cluster 4 data is 72,26%. The correctly classified rate is 16,39% lower in cluster 4 than the result of the data set that do not include process enactment. The average of root mean squared error values of the methods applied to cluster 4 data is 26,68%. The correctly classified rate is 8,71% higher in cluster 4 than the result of the data set that do not include process enactment. We could not obtain promising results from this cluster. To investigate the reason of this we reviewed data and observed that the cluster noise based on between process attibute patterns and defect open duration metric are high in cluster 4.

CHAPTER 5

CONCLUSION AND FUTURE WORK

Defect data gives information related to the software quality. The accessibility to defect data is easy in most cases, since a detailed view of detected defect is recorded to issue tracking tools and all thedata is stored from the initiation of the project to the end of maintenance phase. When the defect data is analyzed by researchers, the understanding of the product environment and process execution is provided.

Quality models such as CMMI enforces in Level 5 that defect prevention is vital for mature process and product. When the cost effectiveness is considered, achieving defect prevention for emergent enterprises is as beneficial as for the institutional ones. One of the activities used for defect prevention is defect data analysis or defect prediction. In order to point out the usable techniques for the understanding of product defectiveness and the factors that have impact on it, we applied various statistical and machine learning analysis methods to our data in our first study. By doing this, we collected defect related and product related metrics in different data sets. At the end, we presented our inferences in three categories based on their confidence [2].

We aimed to understand the effect of process enactment on product defectiveness prediction. After literature search, we decided to use machine learning algorithms for prediction, since these algorithms are suitable for recognizing the patterns in process enactment data. In this context, we performed case studies by using two different software projects. Before conducting case studies, we needed a method in order to systematically plan and analyze case studies. Therefore, we developed a method shown in Figure 3.1. By this method, we achieved the collection of process enactment data, data preprocessing and machine learning analysis.

The method applied in this thesisuses GQM, MUQ, PER, PSM, clustering and classification approaches. Goal-Question-Metric was used to determine the metrics that should be collected. Metric Usability Questionnaires were used to determine usable metrics data. Process Execution Record and Process Similarity Matrix were used to capture process traces and collect process enactment data. Attribute discretization and data reduction were performed in data cleaning and preprocessing phase of the case studies.

To validate the method, we performed four case studies which are conducted on the data of completed two software projects in a small company. In the first case study (case study-1A), product size metrics and defect related metrics data of Project-1 was classified with machine learning approaches. In second case study (case study-1B), same metrics in Project-1 were combined with defect management process enactment attributes and machine learning approaches were repeated. After case-study-1, we observed that the performance values of prediction models with process enactment data are better than the ones without process enactment data. The implementation of case study-1 was repeated with the data of Project-2 in case study-2. We observed similar results in case study 2 with case study 1 except a roughness. The roughness is that two of the clustered data sets with process enactment data gave lower performance values than the analysis results of the data set without process enactment data.

Defect open duration metric was the classifier for all case studies. In other words, it was identified as dependent variable for prediction models. Clustering was applied only in case study 1B and 2B. Clustered defect data was split to separate data sets.

In case study 1 (when compared case studies 1A and 1B) we observed that the

data clustered according to process enactmentpatterns gives approximately3% more accurate results when the cluster has a low number of noisy process patterns (low number of pattern difference) and has sufficient data points to apply machine learning methods. The correctly classified instance values that are the performance evaluation value in machine learning approachesare ranging from - 10 to 17%.

In case study 2 (when compared case studies 2A and 2B) we observed that the data clustered according to process enactment patterns gives 3% more accurateresults in terms of defect open duration metric (ranging from -7% to 8%) when the cluster has a low number of noisy process patterns. The cause of this high noise is implementation of different activities during process execution. Since the project-2 data is so old that the development processes applied might have not been stable organization in these days. To decrease the noise several more clustering operations can be performed.

Another reason of the inconsistent result in case study-2 is that the data used for case study-2 might be retrospective, although the project in case study 1 is a newly completed one. This circumstance causes to collect unreliable data especially for process enactment in case study 2. The MUQs were filled via interviews with current data providers. However, since the providers of data have changed for Project-2 (most of the staffproviding data for the project does not work in the company anymore), theanswers to the questions in MUQ might not have reflected the realsituation for already stored data. Therefore a new part questioning the characteristics of the providers of data under evaluation might be good toadd to the MUQ.

While conducting case studies, we paid attention to take help from process experts by interviewing. But since several personnel of Project-2 were not working for the company anymore, we had to fill PER with the experts who knew only the second half of the development phase.On the other hand, for Project-1 we could easily collect data by using suggested assets. We can say that GQM provides a systematic way to determine the data that will be collected and the analysis methods. MUQ provides to obtain more accurate results by using more accurate data. PER and PSM provides to collect process enactment attributes. Especially the newly proposed usage ways for these assets provides more practical solutions to collect process enactment data. Aside from interviewing approach to fill PER, the usage of the historical process data in issue tracking tool was advantageous to fill PSM during or after process executions. Multilayer Perceptron and Bayesian Networks methods gave more accurate results than the other applied machine learning techniques in this study.

In conclusion, multiple case study implementations showed us that our method can be used if we access reliable PER data in emergent organizations. Our first question was "Is process enactmentdata beneficial for defect prediction?". To answer this, we assessed case study 1 and case study 2 results and the answer is yes. Thesecond question was "How can we use process enactmentdata?". For this question we applied several assets called PER, PSM [16] and clustering in Weka. The third question was "Which approaches or analysis methods can our method support?", and we explained the approaches applied in Section 3 in detail.

When we think of cost of quality [70], performing defect prediction approach costs 10 person-days for a project that shows similar features with the project of the case study 1B that has 296 defects detected. After applying the generated prediction models in new projects, we can calculate the decrease in defect management costs. Therefore, our proposed methodcan support and might be beneficial for the quality system of the organization.

We suggest using process enactmentpatterns for defect prediction operation and also we recommend methods to extract process enactmentdata. In other words, regardless of the analysis method applied, defect and product data must be tracked and assessed with its context to understand the product quality and process performance in turn.Since machine learning is a pattern oriented domain area, process enactmentdata is very convenient for pattern recognition.However, more studies should be performed for more evidence as a future work. Besides, we suggest coding a script to automatically extract historical process data from issue tracking tool, since manual collection of process enactment data for each defect management execution is costly.

The prediction model of defect open duration generated with the proposed way provides a basis for the estimation of the open period of a defect that has been detected in software. If the distribution of the defects is displayed, the trend of open duration for detected defects can be estimated within a project. However, this assumption is not verified in this thesis and might be subject to future work.

The factors that have impact on software product defectiveness can be considered in two categories:Environmental factors and internal process execution. The process enactmentdata which we have gathered for this study contains only inner processes. However, there are some outer factors, such as environmental impacts like personnel skills that affect the results. These outer process factors can be investigated and different collection methods might be discovered for the data. Additionally one more idea for future work is using of classification results to improve processes in organization. By observing the process patterns which give lower open duration values in PSM, organizational processes can be updated according to realized process attributes patterns that show better performance.

REFERENCES

[1] CMMI Product Team, "CMMI for Development, Version 1.3, Technical Report, SEI, 2010.

[2] Sivrioğlu, D. and Tarhan A., "Yazılım Modül Özelliklerine Göre Hatalılık Analizi: Bir Durum Çalışması", UYMS Bildiri Kitabı, February 2012.

[3] Sivrioğlu, D., Tarhan A. and Demirörs O., "An Analysis of Product Defectiveness and Affecting Factors In A Small Software Organization", under review by Euromicro Conference on Software Engineering and Advanced Applications (SEAA) Committee, April 2012.

[4] Jalote, P. and Agrawal, N., "Using Defect Analysis Feedback For Improving Quality and Productivity in Iterative Software Development", Information and Communications Technology, Enabling Technologies for the New Knowledge Society: ITI 3rd International Conference, p. 703 – 713, 2005.

[5] Chrissis, M.B., Konrad, M. and Shrum, S., "CMMI Second Edition Guidelines for Process Integration and Product Improvement", CMMI for Development, Version 1.2, SEI Series in Software Engineering, 2006.

[6] International Organisation for Standardization and International Electrotechnical Commission, "ISO/IEC 15504 Software Process Improvement and Capability Determination Model (SPICE)", 2004.

[7] Leey, T., Namx, J., Hanx D., Kimx S. and Iny H. P., "Micro Interaction Metrics for Defect Prediction", ESEC/FSE '11 Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering, 2011.

[8] Dhiauddin, M., "Defect Prediction Model For Testing Phase", Master Thesis, Universiti Teknologi Malaysia, Faculty of Computer Science and Information System, May 2009.

[9] Wahyudin, D., Schatten, A., Winkler, D., Tjoa, A. M. And Biffl S., "Defect Prediction using Combined Product and Project Metrics A Case Study from the Open Source "Apache" MyFaces Project Family", Software Engineering and Advanced Applications, 2008. SEAA '08. 34th Euromicro Conference, p.207-215, 3-5 Sept. 2008.

[10] Koru, A. G. and Liu, H., "Building Effective Defect-Prediction Models in Practice", IEEE Software, Vol. 22, No. 6, November/December 2005.

[11] Florac, A.W., Park, R.E., and Carleton A.D., Practical Software Measurement: Measuring for Process Management and Improvement. Guidebook: CMU/SEI-97-HB-003, 1997.

[12] Runeson, P. and Höst, M., "Guidelines for conducting and reporting case study research in software Engineering", Empir Software Eng (2009) 14:131–164.

[13] http://www.cs.waikato.ac.nz/~ml/weka/, last access date: 11th April 2012.

[14] Basili, V. R., Caldiera, G. ve Rombach, H. D., "Goal Question Metric Paradigm", Encyclopedia of Software Engineering – 2 Volume Set, ISBN#1-54004-8, 1994.

[15] http://en.wikipedia.org/wiki/Software_bug, last access date: 16th May 2012.

[16] Tarhan, A., "An Assessment Model For The Applicability of Statistical Process Control For Software Processes", PhD Thesis, Informatics Institute, Middle East Technical University, Turkey, 2006.

[17] Kan, S. H., "Metrics and Models in Software Quality Engineering", 2nd Edition, Pearson Education Inc, 2003.

[18] Department of Defense, "Military Standard Quality Assurance Terms and Definitions", MIL-STD-109C, 1994.

[19] Boehm, B., "Software Engineering Economics", Prentice Hall, 1981.

[20] Schell, C., "The Value of the Case Study as a Research Strategy", Manchester Business School, 1992.

[21] Challagulla, V. U. B., Bastani, F. B., Yen, I. and Paul, R. A., "Empirical Assessment of Machine Learning based Software Defect Prediction Techniques", WORDS '05 Proceedings of the 10th IEEE International Workshop on Object-Oriented Real-Time Dependable Systems, USA, 2005.

[22] International Organisation for Standardization and International Electrotechnical Commission, "ISO/IEC JTC 1/SC 7 - Software and Systems Engineering", 2009.

[23] International Organisation for Standardization and International Electrotechnical Commission, "ISO/IEC 24765 - Systems and software engineering - Vocabulary", 2010.

[24] http://en.wikipedia.org/wiki/Machine_learning#cite_ref-0, last access date: 17th May 2012.

[25] http://en.wikipedia.org/wiki/Goodness_of_fit, last access date: 17th May 2012.

[26] Freund, J., E., "Modern Elementary Statistics", 11th Edition, Pearson, 2004.

[27] Babbie, E., R., The Practice of Social Research", 12th edition, Wadsworth Publishing, 2009.

[28] http://en.wikipedia.org/wiki/Bivariate_analysis, last access date: 5th June 2012.

[29] Abdi, H., "Multivariate Analysis", The University of Texas at Dallas, 2003.

[30] http://support.sas.com/publishing/pubcat/chaps/56903.pdf, last access date: 17th May 2012.

[31] http://people.revoledu.com/kardi/tutorial/KNN/HowTo_KNN.html, last access date: 17th May 2012.

[32] http://saravananthirumuruganathan.wordpress.com/2010/05/17/adetailed-introduction-to-k-nearest-neighbor-knn-algorithm/, last access date: 17th May 2012.

[33] http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm, last access date: 5th June 2012.

[34] Malik, J., S., Goyal, P., Sharma, A., K., "Data Mining Problem Solving Algorithms & Their Comparative Study", Pioneer Journal, 12th National Conference,2011.

[35] http://cis.poly.edu/~mleung/FRE7851/f07/decisionTrees.pdf, last access date: 17th May 2012.

[36] http://www.uow.edu.au/~markus/teaching/CSCI323/Lecture_MLP.pdf, last access date: 17th May 2012.

[37] https://www.cra.com/pdf/bnetbuilderbackground.pdf, last access date: 17th May 2012.

[38] Neil, M., Fenton, N., Nielson, L, "Building large-scale Bayesian Networks", The Knowledge Engineering Review, Vol. 15, No. 3, pp. 257-284, 2000.

[39] http://www.aiai.ed.ac.uk/links/bn.html, last access date: 17th May 2012.

[40] http://www-scf.usc.edu/~csci567/11-12-Bayesian-Learning.pdf], last access date: 17th May 2012.

[41] http://css.engineering.uiowa.edu/~comp/Public/Apriori.pdf, last access date: 17th May 2012.

[42] http://ijrte.academypublisher.com/vol02/no03/ijrte02037377.pdf, last access date: 17th May 2012.

[43] Koru, A. G. and Tian, J., "An Empirical Comparison and Characterization of High Defect and High Complexity Modules, The Journal of Systems and Software, 67 (2003) 153–163, 2003.

[44] Salman, N., "A Measurement Framework for Component Oriented Software Systems", PhD Thesis, Computer Engineering, Middle East Technical University, Turkey, 2006.

[45] Manzoor, K., "A Practical Approach to Estimating Defect-Fix Time", http://homepages.com.pk/kashman/defectsEstimation.htm, last access date: 19th May 2012.

[46] Ohlsson, N., Zhao, M. and Helander, M., "Application of multivariate analysis for software fault prediction, Software Quality Journal 7,(1998) 51–66, 1996.

[47] Boetticher, G., D., "Nearest Neighbor Sampling for Better Defect Prediction", PROMISE '05 Proceedings of the 2005 workshop on Predictor models in software engineering, p. 1-6, USA, 2005.

[48] Sandhu, P. S., Khullar, S., Singh, S., Bains S. K., Kaur, M. and Singh, G., "A Study on Early Prediction of Fault Proneness in Software Modules using Genetic Algorithm", World Academy of Science, Engineering and Technology 72, 2010.

[49] Çatal, Ç. ve Diri, B., "A Systematic Review of Software Fault Prediction Studies", Expert Systems with Applications, 36 (2009) 7346-7354, 2009.

[50] Ahsan, S. N., Ferzund, J. and Wotawa, F., "Program File Bug Fix Effort Estimation UsingMachine Learning Methods for Open SourceSoftware Projects", IST-TR-2009-04-24, IST Technical Report, Institut für Softwaretechnologie, Australia, 2009.

[51] Jalote, P., Dinesh, K., Raghavan, S., Bhashyam R. and Ramakrishnan M., "Quantitative Quality Management through Defect Prediction and Statistical Process Control", http://www.iiitd.edu.in/~jalote/papers/2WCSQPaper.pdf, last access date: 19th May 2012.

[52] Gokhale, S. S., Mullen, R. E., "Software Defect Repair Times: A Multiplicative Model", PROMISE '08 Proceedings of the 4th international workshop on Predictor models in software engineering, p. 93-100, USA, 2008.

[53] Schneidewind, N. F., "Modelling the Fault Correction Process", ISSRE '01 Proceedings of the 12th International Symposium on Software Reliability Engineering, p. 185, USA, 2001.

[54] Fenton, N. and Neil, M., "A Critique of Software Defect Prediction Models, IEEE Transactions on Software Engineering, Vol. 25, No 3, p. 675-689, 1999.

[55] Fenton, N., Krause, M., Neil P., "A Probabilistic Model for Software Defect Prediction", For submission to IEEE Transactions in Software Engineering.

[56] He, Z., Shu F., Yang Y., Zhang, W. and Wang, Q., "Data Unpredictability in Software Defect-Fixing Effort Prediction", 2010 10th International Conference on Quality Software, 2010.

[57] Song, Q., Shepperd, M., Cartwright, M. and Mair, C., "Software Defect Association Mining and Defect Correction Effort Prediction", IEEE Transactions on Software Engineering, Vol. 32, NO. 2, 2006.

[58] Zeng, H. and Rine D., "Estimation of Software Defects Fix Effort Using Neural Networks", COMPSAC '04 Proceedings of the 28th Annual International Computer Software and Applications Conference - Workshops and Fast Abstracts – Vol. 02, p. 20-21, USA, 2004.

[59] Thaw, T., Aung, M. P., Wah, N. L., Nyein, S. S., Phyo, Z. L., K. Z., and Htun, K. Z., "Comparison for the Accuracy of Defect Fix Effort Estimation" http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05485813, last access date: 20th May 2012.

[60] Menzies, T., Lutz, R., Mikulski, C., "Better Analysis of Defect Data at NASA", Proceedings of the 15th International Conference on Software Engineering and Knowledge Engineering (SEKE'03), 2003.

[61] Weiss, C., Premraj, R., Zimmermann, T. and Zeller, A.," How Long will it Take to Fix This Bug?", MSR '07 Proceedings of the Fourth International Workshop on Mining Software Repositories, p. 1, USA, 2007. [62] Hassouna, A. and Tahvildari, T., "An Effort Prediction Framework for Software Defect Correction ", Information and Software Technology 52 (2010) 197–209, 2010.

[63] Hewett, R. and Kijsanayothin, P., "On Modeling Software Defect Repair Time", Empir Software Eng (2009) 14:165–186, 2008.

[64] Menzies, T., Butcher, A., Marcus, A., Zimmermann, T. and Cok, C., "Local vs. Global Models for Effort Estimation and Defect Prediction", ASE '11 Proceedings of the 2011 26th IEEE/ACM International Conference on Automated Software Engineering, p. 343-351, USA, 2011.

[65] Tarhan, A. and Demirörs, O., "Investigating the Effect of Variations in Test Development Process: A Case from a Safety-Critical System", Software Quality Journal, DOI: 10.1007/s11219-011-9129-8.

[66] Tarhan, A. and Demirörs, O., "Assessment of Software Process and Metrics to Support Quantitative Understanding: Experience from an Undefined Task Management Process", "Apply Quantitative Management Now", IEEE Software, DOI: 10.1109/MS.2011.91, 2011.

[67] ISO, ISO 9001: 2000: Quality Management Systems – Requirements.

[68] Witten, I. H. and Frank E., "Data Mining Practical Machine Learning Tools and Techniques", 2nd Edition, Elsevier, 2005.

[69] Langford, J., "Clever Methods of Overfitting", Machine Learning (Theory), http://hunch.net/?p=22, last access date: 20th May 2012, 2005.

[70] Demirörs, O., Yıldız, Ö. and Güceğlioğlu, A. S., "Using Cost of Software Quality for a Process Improvement Initiative", METU, 26th EUROMICRO Conference (EUROMICRO'00)-Volume 2, 2007.

[71] http://www.locmetrics.com/, last access date: 11th April 2012.

[72] http://msdn.microsoft.com/en-us/library/ms174493, last access date: 15th July2012.

APPENDICES

A. CASE STUDY PLAN

Design of Study	Purpose
Preparation of GQM tree	Define goals, metrics and statistical analysis
	methods.
Data collection from tools	Defect density and other factors data will be
	available to start analysis.
Data verification	Verify data before using for analysis and to
	decide on applicability for our analysis,
	Metric Usability Questionnaire Forms are
	filled for each basic and derived metrics. (A
	Sample Form given in Attachment-1)
Process data collection	Obtain Process Data. PER (Process Execution
	Report), PSM (Process Similarity Matrix) will
	be used.
Conduct interviews with	Obtain Process Data. PER and PSM will be
domain experts	used.
Data Analysis	Statistical and Machine Learning data
	analysis methods will be applied to data after
	data cleaning.
Presentation Preparation,	Analysis results will be documented.
Reporting	Observed interesting patterns will be shared,
	suggestions will be discussed.

Table A.1 Tasks of Study

Tasks	Start Date	Finish Date	Duration
Preparation of GQM tree	14.03.2012	31.03.2012	17 days
Data collection from tools	14.03.2012	29.03.2012	17 days
Data verification	30.03.2012	31.03.2012	2 days
Process data collection	01.04.2012	15.04.2012	14 days
Conduct interviews with			
domain experts	01.04.2012	15.04.2012	14 days
Data Analysis	16.04.2012	30.04.2012	14 days
Presentation Preparation, Make			
Corrections according to			
Review Items, Reporting	01.05.2012	28.05.2012	27 days

Table A.2 Study Calendar
B. DETAILS OF CASE STUDY 1A

Metric	Name	: Source component			
Attribu	ıtes		Answers	Rating	Expected Answers
	Indica	tors			
Measu	re Ide	ntity		N	
	Q1	Which entity does the measure measure?	Product		
	Q2	Which attribute of the entity does the measure measure?	Defective components of product		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Not applicable		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
	Q6	What is the range of the measurement data?	Five component types		
Data E	xisten	ice		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	296	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data V	/erifial	bility		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data D)epeno	fability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data N	orma	lizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data I	ntegra	bility			
	Q32	Is measurement data integrable at project level?	No		
	Q33	Is measurement data integrable at organization level?	No		

Figure B.1 MUQ for "Source component" Basic Metric of Project-1

Met	ric Nan	ne: Created date			
Attr	ibutes		Answers	Rating	Expected Answers
	Indicate	YS			
Mea	isure lo	lentity		N	
	Q1	Which entity does the measure measure?	Process		
	Q2	Which attribute of the entity does the measure measure?	The date of the defect record		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Time (dd.mm.yy hh:mm)		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Date		
	Q6	What is the range of the measurement data?	00.00.0000 00:00		
Dat	a Exist	ince		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	296	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Dat	a Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	N	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Dat	a Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	N	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	N	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Dat	a Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Dat	a Integ	rability			
	Q32	Is measurement data integrable at project level?	No		
	Q33	Is measurement data integrable at organization level?	No		

Figure B.2 MUQ for "Created Date" Basic Metric of Project-1

Metri	ic Nan	ne: Closed date			
Attrit	butes		Answers	Rating	Expected Answers
l	ndicato	5/5			
Meas	ure lo	lentity		N	
Q	1	Which entity does the measure measure?	Process		
Q)2	Which attribute of the entity does the measure measure?	Closed date of the defect record		
Q	23	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
Q	14	What is the unit of the measurement data?	Time (dd.mm.yy hh:mm)		
C	25	What is the type of the measurement data? (integer, real, etc.)	Date		
Q	26	What is the range of the measurement data?	00.00.0000 00:00		
Data	Exist	ence		F	
Q	17	Is measurement data existent?	Yes		
C	28	What is the amount of overall observations?	296	1	Available > 20
Q	29	What is the amount of missing data points?	0		
Q	210	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
Q	211	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data	Verifi	ability		F	
C	12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Later		
Q	213	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
Q	214	Who is responsible for recording measurement data?	Project Manager		
0	215	Is all measurement data recorded by the responsible body?	Yes	1	Yes
Q	216	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
Q	17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
Q	218	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
Q	219	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data	Depe	ndability		P	
Q	220	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
Q	221	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
Q)22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q	23	Are the frequencies for data generation, recording, and storing different?	No	1	No
Q	24	Is measurement data recorded precisely?	Yes	1	Yes
Q	25	Is measurement data collected for a specific purpose?	No	1	Yes
Q	28	Is the purpose of measurement data collection known by process performers?	No	1	Yes
Q	27	Is measurement data analyzed and reported?	No		Yes
Q	28	Is measurement data analysis results communicated to process performers?	No		Yes
Q	29	Is measurement data analysis results communicated to management?	No		Yes
C	230	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data	Norm	alizability			
Ç	31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data	Integ	rability			
C	32	Is measurement data integrable at project level?	No		
Q	33	Is measurement data integrable at organization level?	No		

Figure B.3 MUQ for "Closed Date" Basic Metric of Project-1

Metri	c Nan	ne: Test type			
Attrib	utes		Answers	Rating	Expected Answers
	Indic	ators			
Meas	ure lo	lentity		N	
	Q1	Which entity does the measure measure?	Process		
	Q2	Which attribute of the entity does the measure measure?	Defect detection rate of different test activities		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Not applicable		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
	Q6	What is the range of the measurement data?	Two test activity types		
Data	Exist	ence		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	296	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data	Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes		Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data	Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly,	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data	Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data	Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure B.4 MUQ for "Test Type" Basic Metric of Project-1

Metric Nam	e: Product version			
Attributes		Answers	Rating	Expected Answers
Indi	ators			
Measure Id	entity		N	
Q1	What is the measure formula? (please refer to related basic metrics)	Process		
Q2	Which attribute of the entity does the measure measure?	Frequency of product development updates		
Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
Q4	What is the unit of the measurement data?	Not applicable		
Q5	What is the type of the measurement data? (integer, real, etc.)	Text (x.y.z)		
Q6	What is the range of the measurement data?	0.0.1-2.0.8		
Data Existe	nce		F	
Q7	Is measurement data existent?	Yes		
Q8	What is the amount of overall observations?	298	1	Available > 20
Q9	What is the amount of missing data points?	0		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data Verifi	sbility		F	
Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
Q14	Who is responsible for recording measurement data?	Test Specialist		
Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data Deper	ndability		P	
Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
Q24	Is measurement data recorded precisely?	Yes	1	Yes
Q25	Is measurement data collected for a specific purpose?	Yes	1	Yes
Q26	Is the purpose of measurement data collection known by process performers?	Yes	1	Yes
Q27	Is measurement data analyzed and reported?	No		Yes
Q28	Is measurement data analysis results communicated to process performers?	No		Yes
Q29	Is measurement data analysis results communicated to management?	No		Yes
Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data Norm	alizability			
Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data Integ	ability			
Q32	Is measurement data integrable at project level?	No		
Q33	Is measurement data integrable at organization level?	No		

Figure B.5 MUQ for "Product Version" Basic Metric of Project-1

Me	etric Nan	ne: Product SLOC			
Att	ributes		Answers	Rating	Expected Answers
	Indicate	YS			
Me	asure lo	lentity		N	
	Q1	Which entity does the measure measure?	Product		
	Q2	Which attribute of the entity does the measure measure?	Size of the product version		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	Q4	What is the unit of the measurement data?	LOC, KLOC		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Integer		
	Qß	What is the range of the measurement data?	[8489,23425]		
Da	ita Exist	ence		F	
	Q7	Is measurement data existent?	No (collected manually)		
	Q8	What is the amount of overall observations?	11	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Da	ıta Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Monthly		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Project Manager		
	Q15	is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Report		
	Q17	is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	tool		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Da	ita Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Monthly		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	Yes	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	Yes	1	Yes
	Q27	Is measurement data analyzed and reported?	Yes		Yes
	Q28	Is measurement data analysis results communicated to process performers?	Yes		Yes
	Q29	Is measurement data analysis results communicated to management?	Yes		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	Yes		Yes
Da	ita Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	Yes (KLOC)		
Da	ta Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure B.6 MUQ for "Product SLOC" Basic Metric of Project-1

Me	tric Nar	ne: Product complexity			
Att	ributes		Answers	Rating	Expected Answers
	Indicate	VS			
Me	asure lo	lentity		N	
	Q1	Which entity does the measure measure?	Product		
	Q2	Which attribute of the entity does the measure measure?	Complexity of the product version		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Number of decision nodes in software		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Integer		
	Q6	What is the range of the measurement data?	(0]		
Da	ta Exist	ence		F	
	Q7	Is measurement data existent?	No (collected manually)		
	Q8	What is the amount of overall observations?	11	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Da	ta Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Never		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	No	1	Yes
	Q14	Who is responsible for recording measurement data?	No one		
	Q15	is all measurement data recorded by the responsible body?	No	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	No		
	Q17	is all measurement data recorded the same way? (on a form, report, tool, etc.)	No	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	No		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	No	1	Yes
Da	ta Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	Yes	1	No
	Q24	Is measurement data recorded precisely?	No	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Da	ta Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Da	ta Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure B.7 MUQ for "Product Complexity" Basic Metric of Project-1

Metric N	ame: Reproducibility	1		
Attribute	6	Inswers	Rating	Expected Answers
In	dicators	0120512	rating	Experies misters
Measure	Identity	·	N	<u> </u>
0	1 Which entity does the measure measure?	Propess		1
0	2 Which attribute of the entity does the measure measure?	Repeatability status of detected defects		
0	3 What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
0	4 What is the unit of the measurement data?	Not applicable	0	
0	5 What is the type of the measurement data? (integer real, etc.)	Text	-	
0	8 What is the rance of the measurement data?	Three status types		
Data Exi	stence		F	
0	7 Is measurement data existent?	Yes		
Q	8 What is the amount of overall observations?	296	1	Available > 20
0	9 What is the amount of missing data points?	0		
0	10 Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
0	11 Is measurement data time sequenced? (If no. please state how measurement data is sequenced)	Yes		
Data Ver	ifiability		F	1
0	2 When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
0	13 Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	V	Yes
0	14 Who is responsible for recording measurement data?	Test Specialist		
0	15 Is all measurement data recorded by the responsible body?	Yes	V	Yes
Q	18 How is measurement data recorded? (on a form, report, tool, etc.)	Tool		11050
0	17 Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
0	18 Where is measurement data stored? (in a file, database, etc.)	The tool's database	S	
0	19 Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	Ň	Yes
Data Dep	pendability		P	
0	20 What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q	21 What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly,	Synchronously		
0	22 What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously	1	
Q	23 Are the frequencies for data generation, recording, and storing different?	No	V	No
0	24 Is measurement data recorded precisely?	Yes	Ń	Yes
Q	25 Is measurement data collected for a specific purpose?	No	Ń	Yes
Q	28 Is the purpose of measurement data collection known by process performers?	No	V	Yes
Q	27 Is measurement data analyzed and reported?	No		Yes
Q	28 Is measurement data analysis results communicated to process performers?	No		Yes
Q	29 Is measurement data analysis results communicated to management?	No		Yes
0	30 Is measurement data analysis results used as a basis for decision making?	No		Yes
Data Nor	malizability			
0	31 Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data Inte	grability			
0	32 Is measurement data integrable at project level?	Yes		
Q	33 Is measurement data integrable at organization level?	Yes		

Figure B.8 MUQ for "Reproducibility" Basic Metric of Project-1

Metric Nar	ne: Project phase			
Attributes		Answers	Rating	Expected Answers
Indi	pators			
Measure In	lentity		N	
Q1	Which entity does the measure measure?	Process		
Q2	Which attribute of the entity does the measure measure?	Project phase of detected defects		
Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
Q4	What is the unit of the measurement data?	Not applicable		
Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
Q6	What is the range of the measurement data?	Two phase types		
Data Exist	ence		F	
Q7	Is measurement data existent?	No		
QS	What is the amount of overall observations?	Not applicable	N	Available > 20
Q9	What is the amount of missing data points?	296		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	296		
Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Not applicable		-
Data Verif	ability (After manual collection)		F	
Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Later		1
Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Later	V	Yes
Q14	Who is responsible for recording measurement data?	Project Manager		
Q15	Is all measurement data recorded by the responsible body?	Yes	V	Yes
Q16	How is measurement data recorded? (on a form, report, tool, etc.)	On a form		
Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
Q18	Where is measurement data stored? (in a file, database, etc.)	in a file	(1
Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	V	Yes
Data Depe	ndability		P	
Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	Yes	V	No
Q24	Is measurement data recorded precisely?	No	N	Yes
Q25	Is measurement data collected for a specific purpose?	No	N	Yes
Q26	Is the purpose of measurement data collection known by process performers?	No	N	Yes
Q27	Is measurement data analyzed and reported?	No		Yes
Q28	Is measurement data analysis results communicated to process performers?	No		Yes
Q25	Is measurement data analysis results communicated to management?	No		Yes
Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data Norn	alizability			
Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data Integ	rability	dala -		
Q32	Is measurement data integrable at project level?	Yes		
Q33	Is measurement data integrable at organization level?	Yes		

Figure B.9 MUQ for "Project Phase" Basic Metric of Project-1

etric Nai	me. Remaining open duration			
della de a		4	D-ft	5
tributes		Answers	Rating	Expected Answ
Inc	dicators			
easure k	identity		N	
Q1	1 What is the measure formula? (please refer to related basic metrics)	Created date, closed date		
02	2 What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
Q3	3 What is the unit of the measurement data?	day		
Q4	4 What is the type of the measurement data? (Integer, real, etc.)	Integer		
QS	5 What is the range of the measurement data?	[0-108]		
ita Exist	tence		F	
QE	6 Is measurement data existent?	Yes		
Q7	7 What is the amount of overall observations?	296	4	Avallable > 20
Q	8 What is the amount of missing data points?	0		
QS	9 Are data points missing in periods? (if yes, please state observation numbers for missing periods)	0		
Q1	10 Is measurement data time sequenced? (if no, please state how measurement data is sequenced)	Yes		
ita Verif	Nability		F	
Q1	11 How is the measure data calculated? (by a tool, manually, etc.)	By a tool		
Q1	12 Is all measurement data calculated with the same way? (by a tool, manually, etc.)	Yes	4	Yes
Q1	13 Is all measurement data calculated according to measure formula?	Yes		
Q1	14 Where is measurement data stored? (in a file, database, etc.)	in a file	4	Yes
Q1	15 Is all measurement data stored in the same place? (in a file, database, etc.)	Yes		
ita Depe	endability		P	
Q1	16 Is measurement data stored precisely?	Yes	4	Yes
Q1	17 Is measurement data stored for a specific purpose?	Yes	4	Yes
Q1	18 Is the purpose of measurement data collection known by process performers?	Yes	4	Yes
Q1	19 Is measurement data analyzed and reported?	No		Yes
02	20 Is measurement data analysis results communicated to process performers?	NO		Yes
02	21 Is measurement data analysis results communicated to management?	No		Yes
02	22 Is measurement data analysis results used as a basis for decision making?	No		Yes
ita Norn	malizability			
02	23 Can measurement data be normalized by parameters or measures? (if yes, please specify them)	NO		
ita integ	grability			
02	24 Is measurement data integrable at project level?	Yes		
Q2	25 Is measurement data Integrable at organization level?	Yes		

Figure B.10 MUQ for "Defect Open Duration" Derived Metric of Project-1

C 1								
Open file	Open URL	Open DB.	Gene	rate	Undo	Edit		Save
None								
relation				Selected attribute				
tion: 120409_raw_de	fect_data_table_v5_discretized-v	weka.filters.unsupervised.attribute	R Attributes: 10 Sum of weighter, 206	Name: Remain	OpenDuration	intinct: E	Type: Nominal	
DC5, 270			Juli of Weights, 250	Missilig, 0 (076)	U	isuici, J	unique, u (u /o)	
				No. Labe		Count	Weight	
All	None	Invert	Pattern	2 108-1	25	/8	/8.0	
				3 54-81		77	77.0	
Name				4 27-54		60	60.0	
1 Category				5 0-27		68	68.0	
2 UndatedDate								
4 TestType								
5 ProductVersion	n							
6 SLOC								
7 Complexity								
a line i data								
8 Reproducibility	/							
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration							
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)			▼ Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)			▼ Uis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77		▼ Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77		▼ Vis
8 Reproducbilty 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77		▼ Vis
8 Reproducbilty 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	60	• Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	60	▼ Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uraton			Class: RemainOpen	Duration (Nom)	77	60	• Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uraton			Class: RemainOpen	Duration (Nom)	77	80	• Vic
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	80	▼ Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	60	• Ve
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	60	▼ Vis
8 Reproducibility 9 ProjectPhase 10 RemainOpenD	uration			Class: RemainOpen	Duration (Nom)	77	60	• Vis
8 Reproducibility 9 ProjectPhase 10 RemainOperD	/ uration	900/4		Class: RemainOpen	Duration (Nom)	77	60	▼ Ve

Figure B.11 Weka View of Case Study 1A

🗿 Weka Explorer			-		· mark	100		
Preprocess Classify Cluster Associate Select attributes V	isualize							
Classifier								
Choose MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V	0 -5 0 -E 20 -H a							
Test options Classifier output								
Ise trainin set	.eu cross-variuaci							
Summary	===							
Supplied test set Set	laggified Instance	- 281		94 9324				
Cross-validation Folds 10 Incorrectly	Classified Instar	nces 15		5.0676	•			
Percentage split % 66 Kappa statis	utic	0.93	38					
More options Mean absolut	e error:	0.02	4					
Root mean so	uared error	0.13	14					
(Nom) RemainOpenDuration	olute error	7.80	146 8 17 8					
Coverage of	cases (0.95 leve)	L) 97.29	73 %					
Start Stop Mean rel. re	gion size (0.95]	level) 21.95	95 %					
Result list (right-dick for options) Total Number	of Instances	296						
00:09:27 - functions.MultilayerPerceptron								
=== Detailed	Accuracy by Clas	3 ===						
	TP Rate FP P	Rate Precision	Recall	F-Measure	ROC Area	Class		
	0.949 0.	018 0.949	0.949	0.949	0.982	81-108		
	0.692 0.	004 0.9	0.692	0.783	0.996	108-135		
	0.961 0.	.023 0.937	0.961	0.949	0.994	54-81		
	0.96/ 0.	1	0.907	0.945	1	27-54		
Weighted Avg	ı. 0.949 0.	.015 0.95	0.949	0.949	0.993	527		
=== Confusio	n Matrix ===							
	i o d classif	ind an						
74 1 2 1	i 0 a = 81-108) }						
4 9 0 0) 0 b = 108-13	35						
0 0 74 3	0 c = 54-81							
0 0 2 56	0 d = 27-54							
0011	.66 e = 0-27							
								-
			m					
Status								
OK								Log x 0

Figure B.12 Multilayer Perceptron Results of Case Study 1A

🥥 Weka Explorer	Conditioned by State	
Preprocess Classify Cluster Associate 5	lect attributes Visualize	
Classifier		
Choose BavesNet -D -O weka.classi	ers, baves.net.search.local.K2 P1 - S BAYES -E weka.classifiers.baves.net.estimate.SimpleEstimator	A0.5
Test options	Classifier output	
🔘 Use training set	=== Summarv ===	A
Supplied test set Set	,	
© Crass unklation - Falls - 10	Correctly Classified Instances 252 85.1351 %	
	Incorrectly Classified Instances 44 14.8649 %	
Percentage split % 66	Kappa statistic 0.8054	
More options	Mean absolute error 0.0579	
	Root mean squared error 0.2001 Delative absolute error 18.8618.8	
(Nom) RemainOpenDuration	Root relative squared error 53.1213 %	
	Coverage of cases (0.95 level) 96.2838 %	
Start Stop	Mean rel. region size (0.95 level) 24.1892 %	
Result list (right-click for options)	Total Number of Instances 296	
00:09:27 - functions.MultilayerPerceptron		
07:37:33 - bayes.BayesNet	=== Detailed Accuracy By Class ===	
	TD Date FD Date Dracision Decall F-Measure DOC &	Yaa (1200
	0.949 0.032 0.914 0.949 0.931 0.9	81 81-108
	0.462 0.004 0.857 0.462 0.6 0.9	92 108-135
	0.805 0.018 0.939 0.805 0.867 0.9	87 54-81
	0.733 0.064 0.746 0.733 0.739 0.9	71 27-54
	0.971 0.075 0.795 0.971 0.874 0.9	96 0-27
	Weighted Avg. 0.851 0.043 0.857 0.851 0.848 0.9	35
	=== Conflicton Matrix ===	
	a b c d e < classified as	=
	74 1 2 1 0 a = 81-108	
	7 6 0 0 0 b = 108-135	
	0 0 62 13 2 c = 54-81	
	$0 \ 0 \ 1 \ 44 \ 15 \ \ 0 \ = \ 27 \ -54$	
	0 0 1 1 00 E = 0-27	
	<pre>(</pre>	,
[]		
Status		
UN		Lug

Figure B.13 BayesNet Results of Case Study 1A

🗿 Weka Explorer		-	and the state	in the second	a the s				- 0 X
Preprocess Classify Cluster Associate Sel	ect attributes Visualize								
Classifier									
Choose Logistic -R 1.0E-8 -M -1									
	.								
Test options C	Jassifier output	alluation	-						
O Use training set	=== Summary ===								*
O Supplied test set Set									
Oroco.uplidation Eoldo 10	Correctly Classified I	nstances	244		82.4324	ł			
	Incorrectly Classified	Instances	52		17.5676	8			
Percentage split % 66	Kappa statistic		0.76	9					
More options	Mean absolute error		0.07						
	Root mean squared erro	r	0.26	16					
(Nom) RemainOpenDuration	Relative absolute erro	r 	22.78	48 8					
(rein) reinen open odraden	Coverage of cases (0 9	Error 5 level)	00./0	/05 295					
Start Stop	Mean rel region size	(0 95 level)	20 47	२ ६					
Pecult list (right-click for antions)	Total Number of Instan	ces	296						
00:09:27 - functions MultilaverDercentron									
07:37:33 - baves.BavesNet	=== Detailed Accuracy	By Class ===							
07:38:02 - functions.Logistic									
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
	0.923	0.041	0.889	0.923	0.906	0.956	81-108		
	0.462	0.011	0.667	0.462	0.545	0.781	108-135		
	0.935	0.087	0.791	0.935	0.857	0.974	54-81		
	0.5	0.017	0.882	0.5	0.638	0.957	27-54		
	0.941	0.075	0.79	0.941	0.859	0.98	0-27		
	Weighted Avg. 0.824	0.054	0.83	0.824	0.812	0.959			
	=== Confusion Matrix =	==							
	abcde <	classified as	3						
	72 3 2 1 0 a =	81-108							
	7 6 0 0 0 b =	108-135							
	1 0 72 3 1 c =	54-81							
	0 0 14 30 16 d =	27-54							
	103064 e =	0-27							
									Ξ
									Ψ.
	(+
Status									
ОК									Log 💉 X O

Figure B.14 Logistic Results of Case Study 1A

🔕 Weka Explorer	1000			-	-	100				- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visualize									
Classifier										
Choose 348 -C 0.25 -M 2										
Techen	Charles a back									
lest options	Classifier output	arruation	-							
O Use training set	=== Summary ===									*
Supplied test set Set										
Cross-validation Folds 10	Correctly Classified In	nstances	271		91.5541	8				
	Incorrectly Classified	Instances	25	_	8.4459	8				
Percentage split % 66	Kappa statistic		0.888	37						
More options	Mean absolute error		0.050	53 19						
	Root mean squared erro.		18 34	15 16 8						
(Nom) RemainOpenDuration 🔹	Root relative squared	error	43.482	29 8						
	Coverage of cases (0.9	5 level)	98.648	36 %						
Start Stop	Mean rel. region size	(0.95 level)	29.93	24 %						
Result list (right-dick for options)	Total Number of Instan	ces	296							
00:09:27 - functions.MultilayerPerceptron										
07:37:33 - bayes.BayesNet	=== Detailed Accuracy 1	By Class ===								
07:38:02 - functions.Logistic			_	.						
08:23:56 - trees.J48	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class			
	0.962	0.06	0.052	0.962	0.904	0.939	109_135			
	0.909	0.018	0.946	0.909	0.927	0.978	54-81			
	1	0.034	0.882	1	0.938	0.988	27-54			
	0.971	0	1	0.971	0.985	0.992	0-27			
	Weighted Avg. 0.916	0.027	0.879	0.916	0.896	0.967				
	=== Confusion Matrix ==	-								
	abcde <	classified as	3							
	12 0 0 0 0 0 b =	100 125								
	0 0 70 7 0 1 c =	54-81								
	0 0 0 60 0 1 d =	27-54								=
	0 0 2 0 66 e =	0-27								
										*
	•									Þ
Statue										
OK									L	Log 💉 x 0

Figure B.15 J48 Results of Case Study 1A

C. DETAILS OF CASE STUDY 1B

🖉 Weka Explorer	100						i fai			- 0 X
Preprocess Classify Cluster Associate Select attributes	Visualize									
Clusterer										
Choose SimpleKMeans -N 7 -A "weka.core.Euclidean	nDistance -R first-las	t" -1 500 -5 10								
Cluster mode	Clusterer output									
Ose training set	Number of 1 Within alua	terations: 2	manad annon		5051/795					*
Supplied test set Set	Missing val	ues globally	replaced wi	th mean/mode	30314720					
Percentage split % 66			•							
Charges to dustors availables	Cluster cen	troids:								
			Cluster#							
(Num) dmR4 v	Attribute	Full Data	(112)	1 (71)	2	3	4	5	6	
V Store clusters for visualization		(296)	(112)	(/1)	(70)	(20)	(5)	(1)	(11)	
	dm01	0.9561	1	0.9859	1	1	1	0	0	
Ignore attributes	dmA2	0.9595	1	0.9859	1	1	1	1	0	
	dmA3	0.098	0.0625	0.2254	0	0	1	0	0.0909	
Start Stop	dmA4	0.9595	1	1	1	1	1	0	0	
Result list (right-dick for options)	dmR2	0.7196	1	0.9859	0	1	1	0	0	
20:25:20 - SimpleKMeans	dmR3	0.6554	1	0	1	0	0	1	1	
20:25:37 - SimpleKMeans	dmR4	0.2399	0	1	0	0	0	0	0	
20:25:52 - SimpleKMeans										
20:47:46 - SimpleKMeans										
20:50:59 - SimpleKMeans 20:57:43 - SimpleKMeans										
20:58:01 - SimpleKMeans	Time taken	to build mode	l (full tra	ining data)	: 0.03 seco	nds				
20:58:20 - SimpleKMeans										
20:58:38 - SimpleKMeans	=== Model a	nd evaluation	on trainin	g set ===						Ξ
20:59:10 - SimpleKMeans										
20:59:49 - SimpleKMeans 21:02:16 - SimpleKMeans	Clustered I	nstances								
2102110 Simplefections	0 112	1 2081								
	1 71	(24%)								
	2 70	(24%)								
	3 26	(9%)								
	4 5	(28)								
	5 1	(0%)								
	6 11	(4%)								*
	•					III				 •
Chables										
OK										Log 💉 x O

Figure C.1SimpleKMeans Clustering of Case Study 1B

🕽 Weka Exp	lorer	<u> </u>			1.1		Contraction of the local division of the loc	and states		- 0 X
Preprocess	Classify Cluster Ass	ociate Select attributes	Visualize							
(Open file	Open URL		Open DB	Gene	rate	Undo	Edit		Save
Filter										
Choose	None									Apply
Current relat Relation: Instances:	ion 120410_dwpc_c0-weka 112	a.filters.unsupervised.attri	oute.Remove-R2-weka.1	filters.unsupe Sum (Attributes: 17 of weights: 112	Selected attribute Name: RemainO Missing: 0 (0%)	penDuration D	istinct: 5	Type: Nominal Unique: 1 (1%)	
Attributes						No. Label		Count	Weight	
	Al	None	Invert	Pa	itern	1 81-108	5	36 1	36.0 1.0	
No.	Name					3 54-81		31	31.0	
1	SourceComponent					4 27-54		39	39.0	
2	CreatedDate					002/			0010	
3	ClosedDate									
4	TestType ProductVorgion									
6	SLOC									
7	Complexity									
8	Reproducibility									
9	ProjectPhase									
10	dm01 dm02					Class: RemainOpenD	luration (Nom)			▼ Visualize Al
12	dmA2									
13	dmA4									
14	dmR2					26				39
15	dmR3					30				
16	dmR4							31		
1/	RemanOpenDuration	1								
		Re	move				1		5	
itatus										100
м										- vy

Figure C.2 Weka View of Case Study 1BCluster 0

🗿 Weka Explorer		_ 0 _ X
Preprocess Classify Cluster Associate S	Select attributes Visualize	
Classifier		
Choose MultilaverPercentrop J 0	0.3.M0.2.ME00.V.0.50.£20.H a	
Choose indicate refeeper on 2 o	0.0.1 0.0 10.0 10.0 10.0 10.0 10.0 10.0	
Test options	Classifier output	
💮 Use training set	=== Stratified cross-validation ===	*
Supplied test set Set	=== Summary ===	
Cross-validation Folds 10	Correctly Classified Instances 108 96.4286 %	
Percentage split % 66	Incorrectly Classified Instances 4 3.5714 %	
	Kappa statistic 0.9489	
More options	Mean absolute error 0.017	
	Root mean squared error 0.1037	
(Nom) RemainOpenDuration 🔹	Relative absolute error 6.0552 \$	
	KOUL FEIGLIVE Squared EFFOF 27.7515 \$	
Start Stop	Mean rel region gize (0.95 level) 21 25 %	
Result list (right-click for options)	Total Number of Instances 112	
21:48:38 - bayes.BayesNet		
21:49:53 - bayes.NaiveBayes	=== Detailed Accuracy By Class ===	
21:52:24 - functions.Logistic		
21:53:18 - functions.Logistic	IP Rate FP Rate Precision Recall F-Measure ROC Area Class	
21:54:46 - functions.MultilayerPerceptron	1 0.013 0.973 1 0.986 0.993 81-108	
	0 0 0 0 0.189 108-135	
	0.968 0.012 0.968 0.968 0.997 54-81	
	1 0.019 0.714 1 0.833 1 27-54	
	0.949 0 1 0.949 0.974 0.998 0-27	
	Weighted Avg. 0.964 0.008 0.961 0.964 0.961 0.989	
	=== Confusion Matrix ===	
	a b c d e < classified as	
	$36 \ 0 \ 0 \ 0 \ 1 \ a = 81-108$	
	$1 \ 0 \ 0 \ 0 \ b = 108-135$	
	0 0 30 1 0 c = 54-81	
	0 0 0 5 0 d = 27-54	
	0 0 1 1 37 e = 0-27	
		Ŧ
Chabus		
OK		.og 💉 x 0

Figure C.3 Multilayer Perceptron Results of Case Study 1BCluster 0

🥥 Weka Explorer				
Preprocess Classify Cluster Associate 5	Gelect attributes Visualize			
Classifier				
Choose BayesNet -D -Q weka.classi	iers.bayes.net.search.local.K2P 1 -S BAYES -E	weka.classifiers.bayes.net.esi	imate.SimpleEstimatorA 0.5	
Test options	Classifier output			
🔘 Use training set	=== Summary ===			•
Supplied test set Set	Correctly Classified Instances	109	97.3214 %	
Cross-validation Folds 10	Incorrectly Classified Instances	3	2.6786 %	
Percentane solit % 66	Kappa statistic	0.9616		
	Mean absolute error	0.014		
More options	Root mean squared error	0.1045 / 082/ S		
	Root relative squared error	27.9532 %		
(Nom) RemainOpenDuration	Coverage of cases (0.95 level)	97.3214 %		
Start Stop	Mean rel. region size (0.95 level) 21.0714 %		
Danula bat (rinka alia) far antiana)	Total Number of Instances	112		
Result list (right-click for options)	Datailed Meanman By Class	_		
21:46:36 - Dayes.BayesNet	Decarted Accuracy by Crass	-		
	TP Rate FP Rate	Precision Recall	F-Measure ROC Area Class	
	1 0.013	0.973 1	0.986 0.988 81-108	
	0 0	0 0	0 0.459 108-135	
	0.968 0.012	0.968 0.968	0.968 0.996 54-81	
	0.009	U.833 I 1 0.974	0.909 0.996 27-54	
	Weighted Avg. 0.973 0.008	0.966 0.973	0.969 0.989	-
	=== Confusion Matrix ===			
	$a \ b \ c \ d \ e \ < \ classified$	33		
	1 0 0 0 0 b = 108-135			E
	0 0 30 1 0 c = 54-81			
	0 0 0 5 0 d = 27-54			
	0 0 1 0 38 e = 0-27			
				_
			m	• • • • • • • •
Status				
OK				Log 📣 XU

Figure C.4 BayesNet Results of Case Study 1BCluster 0

🗿 Weka Explorer		-	-	-		-	Sumatrue .			0 X
Preprocess Classify Cluster Associate 5	Select attributes Visualize									
Classifier										
Choose Logistic -R 1.0E-8 -M -1										
Test options	Classifier output									
Use training set	Startified ences a	1/4/								*
Supplied test set Set	=== Summary ===	11080100								
Cross-validation Folds 10	Correctly Classified In		106		04 6420					
Percentage split % 66	Incorrectly Classified In	Instances	6		5.3571	•				
More ontions	Kappa statistic		0.92	28						
Plote options	Mean absolute error		0.02	14						
	Root mean squared error		0.14	64						
(Nom) KemainOpenDuration	Relative absolute error		7.63	11 %						
Start Stop	Root relative squared e	rror	39.16	93 %						
	Coverage of cases (0.93	(level)	94.04. 20	29 8						
Result list (right-click for options)	Total Number of Instand	0.53 IEVEL) 98	112							
19:11:06 - functions.MultilayerPerceptron										
1572 hzz - functionis/cogsuc	=== Detailed Accuracy H	y Class ===								
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class			
	1	0.013	0.973	1	0.986	0.992	81-108			
	0	0	0	0	0	0.306	108-135			
	0.935	0.025	0.935	0.935	0.935	0.953	27-54			
	0.949	0.003	0.949	0.949	0.949	0.996	0-27			
	Weighted Avg. 0.946	0.021	0.938	0.946	0.942	0.977	0 27			
	=== Confusion Matrix ==	=								
	abcde < o	lassified as								
	36 0 0 0 0 a =	81-108								
	1 0 0 0 0 b =	108-135								
	0 0 29 1 1 c =	54-81								
	0 0 0 4 1 d =	27-54								-
	0 0 2 0 37 e =	0-27								
										•
Status										
OK									Log	🐠 x0

Figure C.5 Logistic Results of Case Study 1BCluster 0

🥥 Weka Explorer									_ 0 _ X
Preprocess Classify Cluster Associate 5	Select attributes Visualize								
Classifier									
Churry 140 C 0 05 140									
Unoose J48 -C 0.25 -I11 2									
Test options	Classifier output								
O Use training set	=== Stratified cr	oss-validation ==							*
Supplied test set Set	=== Summary ===								
Cross-validation Folds 10	Correctly Classif:	ied Instances	107		95.5357	ł			
Percentage split % 66	Incorrectly Class	ified Instances	5		4.4643	1			
	Kappa statistic		0.93	55					
More options	Mean absolute erro	or	0.02	15					
	Root mean squared	error	0.11	17 7 e					
(Nom) RemainOpenDuration 🔹	Root relative gou	ared error	30.70	/ 76 7.0.5					
	Coverage of cases	(0.95 level)	97.32	14 %					
Start Stop	Mean rel. region	size (0.95 level)	21.60	71 %					
Result list (right-dick for options)	Total Number of In	nstances	112						
21:48:38 - bayes.BayesNet									
21:49:53 - bayes.NaiveBayes	=== Detailed Accu:	racy By Class ===	-						
21:52:24 - functions.Logistic									
21:53:18 - functions.Logistic	TP	Rate FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
21:54:46 - functions.MultilayerPerceptron		1 0.013	0.973	1	0.986	0.988	81-108		
22:00:37 - 0.955,340		0 0	0	0	0	0.356	108-135		
		0.968 0.037	0.909	0.968	0.937	0.976	54-81		
		0.6 0.009	0.75	0.074	0.007	0.993	27-54		
	Weighted Avg	0.9/4 0	1 946	0.9/4	0.96/	0.956	U-27		
	Weighten Avg.	0.555 0.015	0.540	0.555	0.55	0.570			
	=== Confusion Mat:	rix ===							
	abcde	< classified a	15						E
	36 0 0 0 0 1	a = 81-108							
	100001	b = 108-135							
		C = 54-81							
	0 0 2 3 0 1	a = 0_27							
	0 0 1 0 30 1	C = 0-27							
									•
Status									
ОК									Log 📣 x 0

Figure C.6 J48 Results of Case Study 1BCluster 0

Open file ar Choose None rent relation Relation: 120410_dwpc_c1.weka.filters.u stances: 71 ributes All Name 1 SourceComponent 2 CreatedDate 3 ClosedDate 4 TestType 5 ProductVersion 6 SLOC 7 Complexity 8 Reproducibility 9 ProjectPhase 10 dm01 11 dmA3	Open URL	Copen DE	3 Gr nsupe Attributes: 11 Sum of weights: 7?	Renerate Selected attribute Name: RemainOpe Missing: 0 (0%) No. Label 1 108-135 2 81-03 3 54-81 4 27-54 5 0-27	Undo Duration Distinct: 5 Count 2 12 21 14 22	Edt Type Unique 2 1 2 1 2 2 1 2 2 1 2 2	e: Nominal e: 0 (0%) Weight 2.0 21.0 21.0 21.0 22.0	·
er Choose None rent relation Relation: 120410_dwpc_c1-weka.fiters.u stances: 71 ributes All SourceComponent 2 CreatedDate 3 ClosedDate 4 TestType 5 ProductVersion 6 SLOC 7 Complexity 8 Reproducibility 9 ProjectPhase 10 dmO1 11 dmA2	nsupervised.attribute.F	Remove-R 19-weka. filters. u Invert	nsupe Attributes: 11 Sum of weights: 7: Pattern	Selected attribute Name: RemainOpe Missing: 0 (0%) No. Label 1 108-135 2 81-108 3 54-81 4 27-54 5 0-27	nDuration Distinct: 5 Count 2 12 21 14 22	Type Unique 2 1 2 1 2 2 2 2 2 2 2 2 2	e: Nominal e: 0 (0%) Weight 2.0 21.0 21.0 21.0 22.0	App
Choose None rent relation Relation: 120410_dwpc_c1-weka.fiters.u stances: 71 All All SourceComponent CoreateDate CoreateDate CoreateDate CoreateDate CoreateDate SourceComponent CoreateDateDate CoreateDateDate CoreateDateDate CoreateDateDate CoreateDateDate CoreateDateDateDate CoreateDateDateDate CoreateDateDateDateDateDateDateDate CoreateDateDateDateDateDateDateDateDateDate	Insupervised.attribute.P	Remove-R19-weka.filters.ur Invert	nsupe Attributes: 1 Sum of weights: 7: Pattern	Selected attribute Name: RemainOpe Missing: 0 (0%) No. Label 1 108-135 2 81-108 3 54-81 4 27-54 5 0-27	nDuration Distinct: 5 Count 2 12 21 14 22	Type Unique 2 1 2 1 2 2 2 2 2	e: Nominal e: 0 (0%) Weight 2.0 21.0 21.0 22.0	Αρ
rent relation Relation: 120410_dwpc_c1-weka.filters.u stances: 71 ributes All . Name 1 SourceComponent 2 CreatedDate 3 ClosedDate 4 TestType 5 ProductVersion 6 SLOC 7 Complexity 8 Reproducibility 9 ProjectPhase 10 dmO1 11 dmA2 12 dmA3	None	Remove-R19-weka.filters.ur Invert	nsupe Attributes: 1 Sum of weights: 7: Pattern	Selected attribute Name: RemainOpe Missing: 0 (0%) No. Label 1 108-135 2 81-108 3 54-81 4 27-54 5 0-27	nDuration Distinct: 5 Count 2 12 21 14 22 21 14 22	Type Unique 2 1 2 1 2 2 2 2 2	e: Nominal e: 0 (0%) Weight 2.0 21.0 21.0 21.0 22.0	
All All All SourceComponent CreatedDate Cr	None	Invert	Pattern	No. Label 1 108-135 2 81-108 3 54-81 4 27-54 5 0-27	Count 2 12 21 14 22	2 1 2 1 2 2 2	Weight 2.0 21.0 21.0 21.0 22.0 22.0	
Al Name Name 1 SourceComponent 2 CreatedDate 3 ClosedDate 4 TestType 5 ProductVersion 6 SLOC 7 Complexity 8 Reproducibility 9 ProjectPhase 10 dm01 11 dmA3	None	Invert	Pattern	1 108-135 2 81-108 3 54-81 4 27-54 5 0-27	2 12 21 14 22	2 1 2 1 2 2	2.0 12.0 21.0 14.0 22.0	
Name SourceComponent CreatedDate CreatedDate CreatedDate CosedDate Consection Succ Complexity Reproducbility ProjectPhase 10 dmO1 dmA3				3 54-81 4 27-54 5 0-27	21 14 22	2	21.0 14.0 22.0	
Iname Iname SourceComponent SourceComponent CreatedDate ClosedDate TestType ProductVersion SLOC Complexity Reproducbility ProjectPhase In dmA1 dmA2				4 27-54 5 0-27	14 22	1 2	14.0	
Source.omponent Source.omponent CreatedDate CreatedDate ClosedDate TestType ProductVersion SLOC Complexity Reproducibility ProjectPhase 10 dmO1 dmA2 dmA3				5 0-27	22	2	22.0	
2 Createduate 3 ClosedDate 4 TestType 5 ProductVersion 6 SLOC 7 Complexity 8 Reproducibility 9 ProjectPhase 10 dmO1 11 dmA2 12 dmA3								
11 dmA2								
12 IdmA3				Class: RemainOpenDura	ation (Nom)		•	Visualiz
13 dmA4 14 dmR2 15 dmR3 16 dmR4				-	21		2	
17 RemanOperDuration	Remove	e		2	12	14		

Figure C.7 Weka View Results of Case Study 1BCluster 1

🥥 Weka Explorer		- 10		-	A	-	Sugar Sector			- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visualize									
Classifier										
Choose MultilayerPerceptron -L 0	.3 -M 0.2 -N 500 -V 0 -S 0 -E 2)-Ha								
Test options	Classifier output									
🔘 Use training set	=== Stratified cross	s-validation ==	=							*
O Supplied test set Set	=== Summary ===									
Cross-validation Folds 10										
Percentage split % 66	Correctly Classifie	i instances	6U 11		54.50/ 15.402	5 6				
	Kappa statistic	teu instances	0.79	16	10.490	5				
More options	Mean absolute error		0.07	35						
	Root mean squared e:	ror	0.23	32						
(Nom) RemainOpenDuration 🔹	Relative absolute es	rror	24.392	26 %						
	Root relative square	ed error	61.40	5 8						
Start Stop	Coverage of cases ().95 level)	88.73	24 %						
Result list (right-click for options)	Mean rel. region si	ze (0.95 level)	25.91	55 %						
19:11:06 - functions.MultilayerPerceptron	Total Number of Ins	lances	71							
19:24:22 - functions.Logistic	Datailed Accura	w Ru Class								
19:25:34 - Dayes.BayesNet	Decarred Accura	.y by Ciass								
19:30:34 - functions.MultilaverPerceptron	TP R	ate FP Rate	Precision	Recall	F-Measure	ROC Area	Class			
	0	0	0	0	0	0.812	108-135			
	0.1	75 0.034	0.818	0.75	0.783	0.84	81-108			
	0.1	357 0.06	0.857	0.857	0.857	0.93	54-81			
	0.1	786 0.07	0.733	0.786	0.759	0.965	27-54			
	1	0.041	0.917	1	0.957	0.997	0-27			
	Weighted Avg. 0.	345 0.05	0.82	0.845	0.832	0.939				
	=== Confusion Matrix	(===								
	abcde <	classified a	3							
	020001	a = 108-135								
	09210]	o = 81-108								
	0 0 18 3 0 0	c = 54-81								
		1 = 27-54								
	0 0 0 0 22 1 1	= 0-27								
										*
Status										0
UK									LO	

Figure C.8 Multilayer Perceptron Results of Case Study 1BCluster 1

🗿 Weka Explorer	and the second second	Intel Property lighted and it	Acres on the second	Allowed as a second sec	
Preprocess Classify Cluster Associate S	ect attributes Visualize				
Classifier					
Chases BayesNet -D -O webs classi	are haven pet cearch local V2D.1 .S. RAVES .	F waka classifiers haves not es	timata SimpleEctimator		
Cinage payeance or 6 monorcrash	si shaqosin dasadirin nodan ke	E word, classifier stabayes in locies	emaceromplecound of the era		
Test options	Classifier output				
💮 Use training set	Stratified Cross-varidation				*
Supplied test set Cat	Summary				
O Suppled test set	Correctly Classified Instances	57	80.2817 %		
Cross-validation Folds 10	Incorrectly Classified Instances	14	19.7183 %		
Percentage split % 66	Kappa statistic	0.7361			
More ontions	Mean absolute error	0.0831			
	Root mean squared error	0.271			
(Nom) DomainOpenDuration -	Relative absolute error	27.5665 %			
	Root relative squared error	69.8671 %			
Start Stop	Mean rel region size (0.95 level)	07.3239 %			
Decult liet (right-click for options)	Total Number of Instances	71			
19:11:06 - functions.MultilaverPercentron					
19:24:22 - functions.Logistic	=== Detailed Accuracy By Class =	==			
19:25:34 - bayes.BayesNet					
19:26:19 - functions.Logistic	TP Rate FP Rate	Precision Recall	F-Measure ROC Area	Class	
19:30:34 - functions.MultilayerPerceptron	0 0	0 0	0 0.957	108-135	
19:37:37 - functions.Logistic	0.75 0.034	0.818 0.75	0.783 0.891	81-108	
19.00.10 boyca.boycanet	0.714 0.04	0.602 0.714	0.769 0.929	27-54	
	0.955 0.041	0.913 0.955	0.933 0.994	0-27	
	Weighted Avg. 0.803 0.058	0.8 0.803	0.794 0.947		
	=== Confusion Matrix ===				
	a b c d e < classified	83			=
	0 2 0 0 0 a = 106-135 0 9 2 1 0 b = 81-108				
	0 0 15 6 0 c = 54-81				
	0 0 0 12 2 d = 27-54				
	0 0 0 1 21 e = 0-27				
					T
	•		III		Þ
Status					
OK					Log x0

Figure C.9 BayesNet Results of Case Study 1BCluster 1

🗿 Weka Explorer				84	-	anal A				- 0 - X
Preprocess Classify Cluster Associate 5	Select attributes Visua	alize								
Classifier										
Choose Logistic -R 1.0E-8 -M -1										
Test options	Classifier output									
O Use training set	=== Stratified	i cross-val	idation ==	-						
O Supplied test set Set	=== Summary ==	-								
Cross-validation Folds 10	Correctly Clas	sified Ins	tances	58		81.6901	1			
Percentage split % 66	Incorrectly Cl	assified]	instances	13		18.3099	1			
More ontions	Kappa statisti	c		0.75	58					
	Mean absolute	error		0.07	19					
	Root mean squa	red error		0.26	16					
(Nom) RemainOpenDuration	Relative absol	ute error		23.84	13 %					
Chart Chan	Root relative	squared en	ror	68.23	58					
Start Stop	Coverage of ca	ises (0.95	level)	83.09	36 %					
Result list (right-click for options)	Mean rel. regi	.on size ((.95 level)) 20.56	34 ≹					
19:11:06 - functions.MultilayerPerceptron	Total Number of Instances			71						
19:24:22 - functions.Logistic	Detailed 3		. (1	_						
19:25:34 - Dayes.BayesNet	=== Decailed A	ecuracy by	C1833 ===	-						
19:20:19 - functions.Logistic		TD Date	VD Date	Presiden	Decell	E-Manaura	DOC Area	Class		
19:37:37 - functions Logistic		IF NAUC	0 014	0 CIECIPION	ACCALL	1-Measure	0 116	109-135		
1969-107 Torregonaleogade		0.75	0.014	0 75	0.75	0.75	0.833	81-108		
		0.714	0.02	0.938	0.714	0.811	0.877	54-81		
		0.857	0.105	0.667	0.857	0.75	0.882	27-54		
		1	0.041	0.917	1	0.957	0.985	0-27		
	Weighted Avg.	0.817	0.048	0.82	0.817	0.811	0.883			
	=== Confusion	Matrix ===								
	abcd	e < ci	assified a	15						
	0200	0 a = 1	.08-135							
	1911	0 b = 8	1-108							
	0 1 15 5	0 c = 5	i4-81							
	0 0 0 12	2 d = 2	7-54							
	0 0 0 0 2	22 e = ()-27							=
										-
Status										
OK										Log 🐠 X

Figure C.10 Logistic Results of Case Study 1BCluster 1

🥥 Weka Explorer		-	-		-	A	-	Sugar and			- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visu	alize									
Classifier											
Chases 149 C 0.25 M 2											
Unoose J46 -C 0.25 -M 2											
Test options	Classifier output										
🔘 Use training set	Stratifio	i aroaa wal	idation	_							*
© Cumplied test est	Summary	1 CIUSS-Val	11040100 ==	-							
Supplied test set	Julillary	-									
Oross-validation Folds 10	Correctly Clas	ssified Ins	stances	61		85.9155	ł				
Percentage split % 66	Incorrectly Cl	lassified]	Instances	10		14.0845	ł				
More ontions	Kappa statisti	ic		0.80	95						
- Hore options	Mean absolute	error		0.07	41						
	Root mean squa	ared error		0.21	94						
(Nom) RemainOpenDuration 🔹	Relative absol	lute error		24.56	78 %						
Ctart Stop	Root relative	squared en	ror	56.56	8 8						
Start Stop	Coverage of ca	ases (0.95	level)	90.14	8 8						
Result list (right-click for options)	Mean rel. regi	ion size (().95 level)	29.85	92 %						
19:11:06 - functions.MultilayerPerceptron	lotal Number o	or instance	3	71							
19:24:22 - functions.Logistic	Detailed /	Accuracy B	, Class								
19:25:34 - Dayes, DayesNet	becaried P	scouracy by	01035								
19:30:34 - functions.MultilaverPerceptron		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class			
19:37:37 - functions.Logistic		0	0	0	0	0	0.652	108-135			
19:38:15 - bayes.BayesNet		0.75	0.017	0.9	0.75	0.818	0.824	81-108			
19:38:54 - trees. J48		0.857	0.04	0.9	0.857	0.878	0.913	54-81			
		0.857	0.07	0.75	0.857	0.8	0.982	27-54			
		1	0.061	0.88	1	0.936	0.998	0-27			
	Weighted Avg.	0.859	0.048	0.839	0.859	0.846	0.931				
	=== Confusion	Matrix ===									
	abcd	e < cl	lassified a	15							
		1 a = 1	1 109								
	0 0 18 3	0 0 0 = 0	31-100 34-81								Ε
	0 0 0 12	2 d = 2	27-54								
	0 0 0 0 2	22 e = ()-27								
	L										
Status										ſ	100
UK										l	LUG VU

Figure C.11 J48 Results of Case Study 1BCluster 1

🧿 Weka Explorer	100 BOOK	1000	100.00	-	- North	The second second second second second second second second second second second second second second second s			5 0 X
Preprocess Classify Clust	er Associate Select attributes Visu	Jalize							
Open file	Open URL	Open DB	Gei	nerate		Undo	Edit	Save	
Filter									
Choose None									Apply
Current relation Relation: 120410_dwpc_ Instances: 70	_c2-weka.filters.unsupervised.attribute	e.Remove-R2-weka.filters.uns	uper Attributes: 16 Sum of weights: 70	Selecter Nam Missin	l attribute e: RemainOpenl g: 0 (0%)	Duration Distinct: 3	u u	Type: Nominal nique: 0 (0%)	
Attributes				No.	Label	Cour	ıt	Weight	
All	None	Invert	Pattern]	1 54-81	22		22.0 41.0	
No Name					3 0-27	7		7.0	
3 TestType 4 ProductVers 5 SLOC 6 Complexity 7 Reproducbil 8 ProjectPhase 9 dmA1 10 dmA2 11 dmA3 12 dmA4 13 dmR2 14 dmR3 15 dmR4	ion ity e Ouration Remov	YE		22	mainOpenDurat	tion (Nom) 41		•	Visualize Al
Status OK								Log] 🔊 x

Figure C.12 Weka View of Case Study 1BCluster 2

🗿 Weka Explorer	A feat from part of the contrast of Mindows	- 0 X
Prenrocess Classify Cluster Associate	Select attributes Visualize	
Classifier		
Chases MultilauerDercentres 0		
Choose Providayer refeeper on -c. o		
Test options	Classifier output	
Use training set	moue z	*
Supplied test set Set	Time silve as build model, 40.70 seconds	
Cross-validation Folds 10	Time caken to build model: 12.75 Seconds	
Percentage split % 66	=== Stratified cross-validation ===	
	=== Summary ===	
more opuons		
	Correctly Classified Instances 67 95.7143 %	
(Nom) RemainOpenDuration	Incorrectly Classified Instances 3 4.2857 %	
Charle Circo	Kappa statistic 0.9213	
Start Stop	Mean absolute error 0.0361	
Result list (right-click for options)	Root mean squared error 0.1475	
19:48:39 - functions.MultilayerPerceptron	Kelative absolute error 9.15/2 %	
	Root Francise Squared Filor S4.4332 %	
	Mean rel. region size (0.95 level) 37.1429 %	
	Total Number of Instances 70	
	=== Detailed Accuracy By Class ===	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	0.909 0.021 0.952 0.909 0.93 0.984 54-81	
	0.976 0.069 0.952 0.976 0.964 0.978 27-54	
	1 U 1 1 1 U-2/	
	Weighted Avg. 0.35/ 0.04/ 0.35/ 0.35/ 0.35/ 0.362	
	=== Confusion Matrix ===	
	a b c < classified as	
	20 2 0 a = 54-81	
	1 40 0 b = 27-54	
	0 0 7 c = 0-27	
		_
		Ψ.
Status		
ОК		Log 🛷 x O

Figure C.13 Multilayer Perceptron Results of Case Study 1BCluster 2

🗿 Weka Explorer	when have performed and the second se	
Preprocess Classify Cluster Associate S	elect attributes Visualize	
Classifier		
Choose BavesNet -D -O weka.classif	ers, baves, net. search, local, K2 P15 BAYES -E weka, classifiers, baves, net, estimate. SimpleEstimator A 0.5	
Test options	Classifier output	
🔘 Use training set		*
O Supplied test set Set	Time taken to build model: 0.02 seconds	
Cross-validation Folds 10	=== Stratified cross-validation ===	
Percentage split % 66	=== Summary ===	
More options	Correctly Classified Instances 64 91.4286 %	
	Incorrectly Classified Instances 6 8.5714 %	
(Nom) RemainOpenDuration	Kappa statistic 0.8348	
	Mean absolute error 0.0553	
Start	Root mean squared error 0.2194	
Result list (right-click for options)	Relative absolute error 14.9629 %	
19:48:39 - functions.MultilayerPerceptron	KOOT FELATIVE SQUARED EFFOR 51.2249 %	
19:54:31 - bayes.BayesNet	Coverage of Cases (0.95 level) 94.2657 %	
	Total Number of Instances 70	
	=== Detailed Accuracy By Class ===	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	0.818 0 1 0.818 0.9 0.973 54-81	
	1 0.207 0.872 1 0.932 0.96 27-54	
	0.714 0 1 0.714 0.833 1 0-27	
	Weighted Avg. 0.914 0.121 0.925 0.914 0.912 0.968	
	=== Confusion Matrix ===	
	a h c < classified as	F
	18 4 0 a = 54-81	
	0 41 0 b = 27-54	
	0 2 5 c = 0-27	
		*
	۲ (III	Þ
Status		
OK		Log 🐠 x 0

Figure C.14 BayesNet Results of Case Study 1BCluster 2

🗿 Weka Explorer	Alter Australia and a state of the second seco	_ 0 X
Preprocess Classify Cluster Associate S	Select attributes Visualize	
Classifier		
Choose Logistic -R 1.0E-8 -M -1		
Test options	Classifier output	
🔘 Use training set	riojettilase 0 7.4423043533000030210	*
Sunnlied test set		
	Time taken to build model: 0.02 seconds	
Cross-validation Folds		
Percentage split % 66	Stratified cross-validation	
More options	=== Summary ===	
	Correctly Classified Instances 63 90 %	
(Nonly Kendinopenburguon +	Incorrectly Classified Instances / 10 %	
Start Stop	Mean absolute error 0.0654	
Result list (right-click for options)	Root mean sourced error 0.2537	
19:48:39 - functions MultilaverPercentron	Relative absolute error 17.6976 %	
19:54:31 - bayes.BayesNet	Root relative squared error 59.2178 %	
19:55:00 - functions.Logistic	Coverage of cases (0.95 level) 91.4286 %	
	Mean rel. region size (0.95 level) 33.8095 %	
	Total Number of Instances 70	
	Deput fact Democra Bu 61-ce	
	=== Detailed Recuracy by Class ===	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	0.909 0.042 0.909 0.909 0.905 54-81	
	0.951 0.172 0.886 0.951 0.918 0.966 27-54	
	0.571 0 1 0.571 0.727 0.984 0-27	
	Weighted Avg. 0.9 0.114 0.905 0.9 0.896 0.965	
	Confusion Matrix	
	a b c ≺ classified as	
	$20 \ 2 \ 0 \ \ a = 54-81$	
	2 39 0 b = 27-54	
	0 3 4 c = 0-27	E
		Ŧ
Status		
OK		Log 💉 x O

Figure C.15 Logistic Results of Case Study 1BCluster 2

🖸 Weka Explorer	A fee has presented by the standing black the	- 🗊 🗙
Preprocess Classify Cluster Associate	Select attributes Visualize	
Classifier		
Choose 148 -C 0.25 -M 2		
Choose 340 *C 0.20*11 2		
Test options	Classifier output	
O Use training set	Jize ut the time time t	*
© Curried test out		
Suppled test set	Time taken to huild model: 0 seconds	
Cross-validation Folds 10		
Percentage split % 66	=== Stratified cross-validation ===	
More options	=== Summary ===	
	Correctly Classified Instances 58 82.8571 %	
(nom) kemanopenburauon •	Incorrectly Classified Instances 12 17.1429 %	
Start Stop	Kappa Statistic U.046/	
Degult list (right slick for options)	Root mean squared error 0.3164	
19:49:39 - functions MultilaverPercentron	Relative absolute error 46.5475 %	
19:54:31 - bayes.BayesNet	Root relative squared error 73.8687 %	
19:55:00 - functions.Logistic	Coverage of cases (0.95 level) 98.5714 %	
19:55:30 - trees.348	Mean rel. region size (0.95 level) 74.2857 %	
	Total Number of Instances 70	_
	Detriled January Dr. Class	
	Detailed Robitady by class	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	0.818 0.021 0.947 0.818 0.878 0.871 54-81	
	0.976 0.379 0.784 0.976 0.87 0.785 27-54	
	0 0 0 0 0.59 0-27	
	Weighted Avg. 0.829 0.229 0.757 0.829 0.785 0.792	=
	On-Environ Materia	-
	Confusion Matrix	
	a b c < classified as	
	18 4 0 a = 54-81	
	1 40 0 b = 27-54	
	0 7 0 c = 0-27	
		τ.
Status		
ОК		Log 💉 X O

Figure C.16 J48 Results of Case Study 1BCluster 2

🗿 Weka Explorer	A Real Property lies	alless vi			1.00		al design.	
Preprocess Classify Cluster	Associate Select attributes Visual	lize						
Open file	Open URL	Open DB	Gene	erate	U	ndo	Edit	Save
Filter								
Choose None								Appl
Current relation Relation: 120410_dwpc_c3-w Instances: 26	veka.filters.unsupervised.attribute.l	Remove-R2, 19	Attributes: 17 Sum of weights: 26	Selected a Name: Missing:	ttribute RemainOpenDuration 0 (0%)	Distinct:	Tyj 1 Uniq	pe: Nominal ue: 0 (0%)
Attributes				No.	Label	Co	unt	Weight
Al	None	Invert	Pattern		1 81-108	26		26.0
SourceComponen SourceComponen CreatedDate GatedDate GatedDate GatedDate GatedDate GatedDate GatedDate GatedDate GatedDate GatedDateDate GatedDateDate GatedDateDateDate GatedDateDateDateDateDateDateDateDateDateDate	t	2		Class: Rema	ainOpenDuration (Nom)			✓ Visualize /
Status OK								Log 🐠

Figure C.17 Weka View of Case Study 1BCluster 3

0	Inen file	Open LIPI	Open [R Ce	norato		740	Edit	Sava
	pen ne	open okc	opent	0	neidte			Luitin	adve
ose	None								
t relat	ion				Selected	l attribute			
ation:	120410_dwpc_c4-wek	a.filters.unsupervised.attri	ibute.Remove-R2,19	Attributes: 17	Name	e: RemainOpenDuration		Type:	Nominal
nces:	5			Sum of weights: 5	Missing	g: 0 (0%)	Distinct: 2	Unique:	0 (0%)
tes					No.	Label	Count	We	eight
_	A	None	Invert	Pattern		1 81-108	3	3.0	
						2 108-135	2	2.0	
	Name								
1	SourceComponent								
2	CreatedDate								
3	ClosedDate								
4	TestType								
5	ProductVersion								
6	SLOC Camplavity								
/	Reproducibility								
9	ProjectPhase								
10	dm01								
11	dmA2				Class: Re	mainOpenDuration (Nom)			▼ Visua
12	dmA3								
13	dmA4				3				
14	dmR2				-				
15	dmR4				-				
17	RemainOnenDuratio	1							
								2	
								ĺ	
		R	emove						

Figure C.18 Weka View of Case Study 1BCluster 4

🔕 Weka Explorer			Second Party		1000	San Income	all-bendler.		X
Preprocess Classify Cluster Ass	ociate Select attributes Visu	ualize							
Open file	Open URL	Open DB	Gene	erate		Undo	Edit	Save	
Filter									
Choose None									Apply
Current relation Relation: 120410_dwpc_c5-wek Instances: 1	a.filters.unsupervised.attribut	e.Remove-R2,19	Attributes: 17 Sum of weights: 1	Selected Name: Missing:	attribute RemainOpenDurat 0 (0%)	tion Distinct	T) : 1 Unit	pe: Nominal ue: 1 (100%)	
Attributes				No.	Label	C	Count	Weight	
Al	None	Invert	Pattern		1 81-108	1		1.0	
No. Name 1 SourceComponent 2 CreatedDate 3 ClosedDate 4 FestType 5 Productiversion 6 SLOC 7 Complexity 8 Reproductiversion 10 dm01 11 dm43 12 dm83 16 dmR4 17 RemainOper/Duration	n Remo	N6		Class: Ren	nainOpenDuration ()	Nom)		✓ Usua	alize All
Status OK								Log	у х О

Figure C.19 Weka View of Case Study 1BCluster 5

Onen file	Open LIDI	0 00		vata	Linda	E.44	Cau-
Open file	Upen UKL	Upen UB	Gene	sate	Undo	Ealt	Save
							,
se None							
relation				Selected attribute			
tion: 120410_dwpc_c6-we	a.filters.unsupervised.attrib	oute.Remove-R2,19	Attributes: 17	Name: RemainOpenE	Duration	T	ype: Nominal
ces: 11			Sum of weights: 11	Missing: 0 (0%)	Distinct: 2	Uni	que: 0 (0%)
es				No. Label	Count		Weight
Al	None	Invert	Pattern	1 108-135	8		8.0
				2 54-81	3		3.0
Name							
1 SourceComponent							
2 CreatedDate							
3 ClosedDate							
4 TestType							
5 ProductVersion							
6 SLOC							
7 Complexity							
8 Reproducibility							
9 ProjectPriase				Į			
11 dmA2				Class: RemainOpenDurat	ion (Nom)		▼ Vis
12 dmA3							
13 dmA4							
14 dmR2				8			
15 dmR3							
16 dmR4							
17 RemainOpenDurati	'n						
						3	
	Ua	move					

Figure C.20 Weka View of Case Study 1BCluster 6



Figure C.21 Multilayer Perceptron Results of Case Study 1BCluster 6
🕢 Weka Explorer	Alter Summer at the second States	- 0 X
Preprocess Classify Cluster Associate 5	Jelect attributes Visualize	
Classifier		
Choose BayesNet -D -Q weka.classi	iers, bayes, net, search, local, K2 P1 5 BAYES -E, weka, classifiers, bayes, net, estimate, SimpleEstimator A 0, 5	
Test options	Classifier output	
🔘 Use training set	LogScore AIC: -179.400339300687	*
O Supplied test set Set		
Cross-validation Folds 10		
Dercentage gelt	Time taken to build model: 0.02 seconds	
O Percentage spin 78 00	Stwatified opnes_uslidation	
More options	=== Summary ===	
(Nom) RemainOpenDuration	Correctly Classified Instances 11 100 %	
Chut Chu	Incorrectly Classified Instances 0 0 %	
Start Stop	Kappa statistic 1	
Result list (right-click for options)	Mean absolute error U.UUU3	
19:48:39 - functions.MultilayerPerceptron	Relative absolute error 0.0748 %	
19:55:00 - functions.Logistic	Root relative squared error 0.1262 %	
19:55:30 - trees. J48	Coverage of cases (0.95 level) 100 %	
19:58:50 - functions.MultilayerPerceptron	Mean rel. region size (0.95 level) 50 %	
19:59:42 - bayes.BayesNet	Total Number of Instances 11	
	Detailed Jeanwary Dr. Class	
	peratter working by class	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	1 0 1 1 1 108-135	
	1 0 1 1 1 54-81	
	Weighted Avg. 1 0 1 1 1 1	
	=== Confligion Matrix ===	Ξ
	a b < classified as	
	8 0 a = 108-135	
	0 3 b = 54-81	
		T
		•
Status		
OK		Log x0

Figure C.22 BayesNet Results of Case Study 1BCluster 6

🗿 Weka Explorer	And International Contraction of Con	_ 🗇 🗙
Preprocess Classify Cluster Associate	Select attributes Visualize	
Classifier		
Chases Legistic D 1 05 0 M 1		
Choose Logistic -R 1.02-0 -M-1		
Test options	Classifier output	
∩ Use training set	Treproducinities 0	
	ProjectPhase 0.0001	
O Supplied test set Set	dmA3 20.0855	
Cross-validation Folds 10		
Percentage split % 66	Time taken to build model. O sepanda	
	Time taken to build model: 0 seconds	
More options	Stratified cross-validation	
(Nom) RemainOpenDuration	· · · · · · · · · · · · · · · · · · ·	
	Correctly Classified Instances 11 100 %	
Start Stop	Incorrectly Classified Instances 0 0 %	
Result list (right-click for options)	Kappa statistic 1	
19:48:39 - functions.MultilaverPerceptron	Mean absolute error 0.0003	
19:54:31 - bayes.BayesNet	Root mean squared error 0.001	
19:55:00 - functions.Logistic	Relative absolute error 0.0658 %	
19:55:30 - trees. J48	Root relative squared error 0.199 %	
19:58:50 - functions.MultilayerPerceptron	Coverage of cases (0.95 level) 100 %	
19:59:42 - bayes.BayesNet	Mean rel. region size (0.95 level) 50 %	
20:00:15 - functions.Logistic	Total Number of Instances 11	
	=== Detailed Accuracy By Class ===	
	TO Date DD Date Descision Date 11 D Manuary DOG have Olars	
	IF RATE FF RATE FFECISION RECAIL F-MEASURE KUC AFEA CLASS	
	Neighted Burg 1 0 1 1 1 1 1	
	=== Confusion Matrix ===	
		=
	a b < classified as	
	8 0 a = 108-135	
	0 3 b = 54-81	
		-
Status		
UK		

Figure C.23 Logistic Results of Case Study 1BCluster 6

🗿 Weka Explorer	A fee has period at the second of the local second of the	- 0 X
Preprocess Classify Cluster Associate 5	elect attributes Visualize	
Classifier		
Choose 148 -C 0 25 -M 2		
Test options	Classifier output	
🔘 Use training set	TRUMPET OF DEGVES . 2	*
Supplied test set Set	Size of the tree : 3	
(i) Cross-validation Folds 10		
Percentage split % 66	Time taken to build model: 0 seconds	
More options	Stratifiad Arnag_uzlidation	
	=== Sutatile Gross valuation ===	
(Nom) RemainOpenDuration 🔹	e domine x]	
	Correctly Classified Instances 11 100 %	
Start Stop	Incorrectly Classified Instances 0 0 %	
Result list (right-click for options)	Kappa statistic 1	
19:48:39 - functions.MultilayerPerceptron	Mean absolute error 0	
19:54:31 - bayes.BayesNet	Root mean squared error 0	
19:55:00 - functions.Logistic	Relative absolute error 0 %	
19:55:30 - trees. J48	Root relative squared error 0 %	
19:58:50 - functions.MultilayerPerceptron	Coverage of cases (0.95 level) 100 %	
19:59:42 - bayes.BayesNet	Mean rel. region size (0.95 level) 50 %	
20:00:15 - functions.Logistic	Total Number of Instances 11	
	=== Detailed Accuracy By Class ===	
	TP Rate FP Rate Precision Recall F-Measure ROC Area Class	
	1 0 1 1 1 108-135	
	1 0 1 1 1 54-81	=
	Weighted Avg. 1 0 1 1 1 1	
	=== Confusion Matrix ===	
	a D < classified as	
	U 3 D = 54-81	
		•
Status		
ОК		2g 💉 🎻 x 0

Figure C.24 J48 Results of Case Study 1BCluster 6

D. DETAILS OF CASE STUDY 2A

Metric Na	me: Detected SCU name		_	
Attributes		Answers	Rating	Expected Answers
Indi	cators	201200-01		1
Measure I	dentity	0	N	
Q1	Which entity does the measure measure?	Product		
Q2	Which attribute of the entity does the measure measure?	Defective unis of product		1
Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
Q4	What is the unit of the measurement data?	Not applicable		
Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
QB	What is the range of the measurement data?	Fourteen configuration units of		
Data Exis	tence		F	
Q7	Is measurement data existent?	Yes		
QS	What is the amount of overall observations?	425	1	Available > 20
Q9	What is the amount of missing data points?	0		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		1
Data Veril	iability		F	
Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
Q14	Who is responsible for recording measurement data?	Test Specialist		100
Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		10. TT.
Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		1
Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	Ň	Yes
Data Dep	endability		P	
Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		1
Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No	N	No
Q24	Is measurement data recorded precisely?	Yes	N	Yes
Q25	Is measurement data collected for a specific purpose?	No	V	Yes
Q28	Is the purpose of measurement data collection known by process performers?	No	N	Yes
Q27	Is measurement data analyzed and reported?	No		Yes
Q28	Is measurement data analysis results communicated to process performers?	No		Yes
Q29	Is measurement data analysis results communicated to management?	No		Yes
Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data Norr	nalizability			a la construcción de la construc
Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		1
Data Inter	arability			1
Q32	Is measurement data integrable at project level?	No		1
Q33	Is measurement data integrable at organization level?	No		1

Figure D.1 MUQ for "Detected SCU Name" Basic Metric of Project-2

Metri	c Nan	ie: Source component			
Attri	utes		Answers	Rating	Expected Answers
	Indic	ators		risting	Expedice misters
Meas	ure lo	lentity		N	
incu.	01	Which entity does the measure measure?	Product	-	
	02	Which attribute of the entity does the measure measure?	Defective components of product	-	
-	03	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal	-	Ratio Absolute
-	Q4	What is the unit of the measurement data?	Not applicable	-	
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
	QB	What is the range of the measurement data?	Seven component types		
Data	Exist	ence		F	
	07	Is measurement data existent?	Yes		
	QS	What is the amount of overall observations?	425	1	Available > 20
	09	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0	-	
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		1
Data	Verifi	ability		F	1
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	V	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q18	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	Ň	Yes
Data	Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	022	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	N.	No
	Q24	Is measurement data recorded precisely?	Yes	N	Yes
	Q25	Is measurement data collected for a specific purpose?	No	N	Yes
	Q28	Is the purpose of measurement data collection known by process performers?	No	N	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data	Norm	alizability			
8	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data	Integ	rability			
1	Q32	Is measurement data integrable at project level?	No		
	Q33	Is measurement data integrable at organization level?	No		

Figure D.2 MUQ for "Source Component" Basic Metric of Project-2

Met	ric Na	me: Created date			
844	lleuter		An even of the second s	Dation	Evented Income
MU	Indicat	205	Answers	Rating	Expected Answers
Max	nura l	dantity		N	
mea	01	Which antity drac the measure measure?	Process		
	02	Which attribute of the entity does the measure measure?	The date of the defect record		
	03	What is the scale of the measurement data? (nominal ordinal interval ratio absolute)	Nominal		Ratio Absolute
	04	What is the unit of the measurement data?	Time (dd mm vy hh:mm)		none, mesenere
	05	What is the type of the measurement data? (integer real, etc.)	Date		
	08	What is the range of the measurement data?	00.00.0000 00:00		
Dat	a Exist	ence		F	
Π	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	425	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Dat	a Verif	iability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Dat	a Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Dat	a Norn	nalizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Dat	a Integ	rability			
	Q32	Is measurement data integrable at project level?	No		
	Q33	Is measurement data integrable at organization level?	No		

Figure D.3 MUQ for "Created Date" Basic Metric of Project-2

Me	tric Nan	e: Closed date			
Att	ributes		Answers	Rating	Expected Answers
	Indicate	75			
Me	asure lo	lentity		N	
	Q1	Which entity does the measure measure?	Process		
	Q2	Which attribute of the entity does the measure measure?	Closed date of the defect record		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Time (dd.mm.yy hh:mm)		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Date		
	Q6	What is the range of the measurement data?	00.00.0000 00:00		
Da	ta Exist	ince		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	298	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Da	ta Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Later		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Project Manager		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Da	ta Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Da	ta Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Da	ta Integ	rability			
	Q32	Is measurement data integrable at project level?	No		
	Q33	Is measurement data integrable at organization level?	No		

Figure D.4 MUQ for "Closed Date" Basic Metric of Project-2

Metri	c Nan	ne: Test type			
Attrib	utes		Answers	Rating	Expected Answers
	Indic	ators			
Meas	ure lo	lentity		N	
	Q1	Which entity does the measure measure?	Process		
	Q2	Which attribute of the entity does the measure measure?	Defect detection rate of different test activities		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Not applicable		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
	Q6	What is the range of the measurement data?	Two test activity types		
Data	Exist	ence		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	296	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data	Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	N	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	N	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data	Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly,	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	N	Yes
	Q25	Is measurement data collected for a specific purpose?	No	N	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	N	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data	Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data	Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure D.5 MUQ for "Test Type" Basic Metric of Project-2

Metric I	Metric Name: Product version						
Attribut	es		Answers	Rating	Expected Answers		
	Indica	tors					
Measur	e Ide	ntity		N			
(Q1	What is the measure formula? (please refer to related basic metrics)	Process				
(02	Which attribute of the entity does the measure measure?	Frequency of product development updates				
(Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute		
(Q4	What is the unit of the measurement data?	Not applicable				
(Q5	What is the type of the measurement data? (integer, real, etc.)	Text (x.y.z)				
(Q6	What is the range of the measurement data?	0.0.1-2.0.8				
Data Ex	isten	ce		F			
(Q7	Is measurement data existent?	Yes				
(80	What is the amount of overall observations?	298	1	Available > 20		
(Q9	What is the amount of missing data points?	0				
(Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0				
(Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes				
Data Ve	rifial	pility		F			
(Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start				
(Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes		
(Q14	Who is responsible for recording measurement data?	Test Specialist				
(Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes		
(Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool				
(Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes		
(Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database				
(Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes		
Data De	epend	fability		P			
(Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously				
(021	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously				
(022	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously				
(Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No		
(Q24	Is measurement data recorded precisely?	Yes	1	Yes		
(Q25	Is measurement data collected for a specific purpose?	Yes	1	Yes		
(Q26	Is the purpose of measurement data collection known by process performers?	Yes	1	Yes		
(Q27	Is measurement data analyzed and reported?	No		Yes		
(Q28	Is measurement data analysis results communicated to process performers?	No		Yes		
(Q29	Is measurement data analysis results communicated to management?	No		Yes		
(Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes		
Data No	ormal	izability					
(Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No				
Data In	tegra	bility					
(032	Is measurement data integrable at project level?	No				
(033	Is measurement data integrable at organization level?	No				

Figure D.6 MUQ for "Product Version" Basic Metric of Project-2

Met	tric Nar	ne: Product SLOC			
0.44	ibuter		Incurre	Dation	Expected Incurrent
MU	Indicat	ne .	Answers	Raung	Expected Answers
Ma	nura l	dantitu		N	
mea	01	Which entity does the measure measure?	Product		
	02	Which attribute of the entity does the measure measure?	Size of the product version		
	03	What is the scale of the measurement data? (nominal ordinal interval ratio shecilute)	Absolute		Ratio Absolute
	04	What is the unit of the measurement data?	LOC KLOC		Nally, nesvine
	05	What is the tune of the measurement data? (intener real etc.)	Integer		
	08	What is the range of the measurement data?	[283 51533]		
Dat	a Evist	ance	[600,01000]	F	
	07	le mascurament data avietant?	No (collected manually)		
	08	What is the amount of overall observations?	19	4	Available > 20
	09	What is the amount of missing data points?	0	,	
	010	Are data points mission in periods? (If yes, place state observation numbers for mission periods)	0		
	011	Is massurament data time seminaned? (If no, nlasse state doservation numbers for missing periods)	Vac		
Dat	a Verif	ability	100	F	
	012	When is measurement data recorded in the process? (at start, middle, and, later, etc.)	Monthly	- r	
	013	Is all measurement data recorded in the process: (at start, mode, etc., ater, etc.)	Vec	1	Vac
	014	Who is responsible for reporting massurament data?	Project Manager	,	165
	015	Is all massurement data reported by the responsible body?	Voe		Vec
	018	How is measurement data recorded by the responsible body :	Peport	,	165
-	017	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Vas	1	Vas
-	018	Where is measurement data recorded the same way r (on a rorm, report, oor, etc.)	teal	,	1 43
	010	Is all measurement data stored : (in a rile, database, etc.)	Var	-	Ver
Dat	a Dana	adability	105	P	165
Ua	020	Maddiniy What is the freewenny of generating management data? (seventhenewich, daily, weakly, monthly, ato)	Acynahranauch /	- r	
	6220	what is the nequency of generating measurement data: (asynchronodsiy, daily, weekly, monthily, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Monthly		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No	1	No
	Q24	Is measurement data recorded precisely?	Yes	1	Yes
	Q25	Is measurement data collected for a specific purpose?	Yes	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	Yes	1	Yes
	Q27	Is measurement data analyzed and reported?	Yes		Yes
	Q28	Is measurement data analysis results communicated to process performers?	Yes		Yes
	Q29	Is measurement data analysis results communicated to management?	Yes		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	Yes		Yes
Dat	ta Norn	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	Yes (KLOC)		
Dat	ta Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure D.7 MUQ for "Product SLOC" Basic Metric of Project-2

Met	tric Nar	ne: Product complexity			
			-		-
Attr	ibutes		Answers	Rating	Expected Answers
	Indicate	ors			
Mea	asure lo	lentity		N	
	Q1	Which entity does the measure measure?	Product		
	Q2	Which attribute of the entity does the measure measure?	Complexity of the product version		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Number of decision nodes in software		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Integer		
	Q6	What is the range of the measurement data?	[33,9180]		
Dat	ta Exist	ence		F	
	Q7	Is measurement data existent?	No (collected manually)		
	Q8	What is the amount of overall observations?	18	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Dat	ta Verifi	ability		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Never		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	No	1	Yes
	Q14	Who is responsible for recording measurement data?	No one		
	Q15	Is all measurement data recorded by the responsible body?	No	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	No		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	No	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	No		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	No	1	Yes
Dat	ta Depe	ndability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	Yes	1	No
	Q24	Is measurement data recorded precisely?	No	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Dat	ta Norm	alizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Dat	ta Integ	rability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure D.8 MUQ for "Product Complexity" Basic Metric of Project-2

Metrie	Nam	e: Reproducibility			
Attrib	utes		Answers	Rating	Expected Answers
	Indic	ators		, and a second	Laperica monero
Measu	ure Ide	ntity		N	
	Q1	Which entity does the measure measure?	Process		
	02	Which attribute of the entity does the measure measure?	Repeatability status of detected defects		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	04	What is the unit of the measurement data?	Not applicable		1
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
à 3	Q8	What is the range of the measurement data?	Four status types		
Data	Exister	nce		F	
	Q7	Is measurement data existent?	Yes		
	Q8	What is the amount of overall observations?	425	1	Available > 20
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	0		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Yes		
Data	Verifia	bility		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	At start		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1	Yes
	Q14	Who is responsible for recording measurement data?	Test Specialist		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	Tool		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	N	Yes
1	Q18	Where is measurement data stored? (in a file, database, etc.)	The tool's database	()	
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	N	Yes
Data	Depen	dability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly,	Synchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		-
	Q23	Are the frequencies for data generation, recording, and storing different?	No	N	No
	Q24	Is measurement data recorded precisely?	Yes	N	Yes
1	Q25	Is measurement data collected for a specific purpose?	No	N.	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	Ň	Yes
	Q27	Is measurement data analyzed and reported?	No	_	Yes
	Q28	Is measurement data analysis results communicated to process performers?	No	_	Yes
	Q29	Is measurement data analysis results communicated to management?	No	_	Yes
-	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data	Norma	lizability		-	-
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data	Integra	ability			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes	1	

Figure D.9 MUQ for "Reproducibility" Basic Metric of Project-2

Metric	Name	:: Project phase			
Attribu	utes		Answers	Rating	Expected Answers
	Indica	ators			
Measu	ire Ide	ntity		N	
	Q1	Which entity does the measure measure?	Process		
	Q2	Which attribute of the entity does the measure measure?	Project phase of detected defects		
	Q3	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Nominal		Ratio, Absolute
	Q4	What is the unit of the measurement data?	Not applicable		
	Q5	What is the type of the measurement data? (integer, real, etc.)	Text		
	Q6	What is the range of the measurement data?	Four phase types		
Data E	xister	ice		F	
	Q7	Is measurement data existent?	No		
	Q8	What is the amount of overall observations?	Not applicable	1	Available > 20
	Q9	What is the amount of missing data points?	425		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	425		
	Q11	Is measurement data time sequenced? (If no, please state how measurement data is sequenced)	Not applicable		
Data V	/erifial	bility (After manual collection)		F	
	Q12	When is measurement data recorded in the process? (at start, middle, end, later, etc.)	Later		
	Q13	Is all measurement data recorded at the same place in the process? (at start, middle, end, later, etc.)	Later	1	Yes
	Q14	Who is responsible for recording measurement data?	Project Manager		
	Q15	Is all measurement data recorded by the responsible body?	Yes	1	Yes
	Q16	How is measurement data recorded? (on a form, report, tool, etc.)	On a form		
	Q17	Is all measurement data recorded the same way? (on a form, report, tool, etc.)	Yes	1	Yes
	Q18	Where is measurement data stored? (in a file, database, etc.)	In a file		
	Q19	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes	1	Yes
Data D	Depen	dability		P	
	Q20	What is the frequency of generating measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q21	What is the frequency of recording measurement data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously		
	Q22	What is the frequency of storing measurement data? (asynchronously, daily, weekly, monthly, etc.)	Synchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	Yes	1	No
	Q24	Is measurement data recorded precisely?	No	1	Yes
	Q25	Is measurement data collected for a specific purpose?	No	1	Yes
	Q26	Is the purpose of measurement data collection known by process performers?	No	1	Yes
	Q27	Is measurement data analyzed and reported?	No		Yes
	Q28	Is measurement data analysis results communicated to process performers?	No		Yes
	Q29	Is measurement data analysis results communicated to management?	No		Yes
	Q30	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data N	lorma	lizability			
	Q31	Can measurement data be normalized by parameters or measures? (If yes, please specify them)	No		
Data I	ntegra	ibility			
	Q32	Is measurement data integrable at project level?	Yes		
	Q33	Is measurement data integrable at organization level?	Yes		

Figure D.10 MUQ for "Project Phase" Basic Metric of Project-2

Metric Name: Remain	ng open duration				
Attributor			Anewore	Pating	Expected Answers
Auributes	Indicators		MISHEIS	Kaung	Expected Allswers
Measure Identity				N	
	01	What is the measure formula? (please refer to related basic metrics)	Created date, closed date		
	02	What is the scale of the measurement data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	03	What is the unit of the measurement data?	day		
	Q4	What is the type of the measurement data? (integer, real, etc.)	Integer		
	Q5	What is the range of the measurement data?	[0-140]		
Data Existence				F	
	Q6	Is measurement data existent?	Yes		
	Q7	What is the amount of overall observations?	425	1	Available > 20
	Q8	What is the amount of missing data points?	0		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for	0		
	Q10	Is measurement data time sequenced? (If no, please state how measurement data is	Yes		
Data Verifiability				F	
	Q11	How is the measure data calculated? (by a tool, manually, etc.)	By a tool		
	Q12	Is all measurement data calculated with the same way? (by a tool, manually, etc.)	Yes	1	Yes
	Q13	Is all measurement data calculated according to measure formula?	Yes		
	Q14	Where is measurement data stored? (in a file, database, etc.)	In a file	V	Yes
	Q15	Is all measurement data stored in the same place? (in a file, database, etc.)	Yes		
Data Dependability				P	
8	Q16	Is measurement data stored precisely?	Yes	V	Yes
	Q17	Is measurement data stored for a specific purpose?	Yes	V	Yes
	Q18	Is the purpose of measurement data collection known by process performers?	Yes	V	Yes
	Q19	Is measurement data analyzed and reported?	No		Yes
	Q20	Is measurement data analysis results communicated to process performers?	No		Yes
	Q21	Is measurement data analysis results communicated to management?	No		Yes
0	Q22	Is measurement data analysis results used as a basis for decision making?	No		Yes
Data Normalizability					
	Q23	Can measurement data be normalized by parameters or measures? (If yes, please speci	No		
Data Integrability					
	Q24	Is measurement data integrable at project level?	Yes		
	Q25	Is measurement data integrable at organization level?	Yes		9

Figure D.11 MUQ for "Defect Open Duration" Derived Metric of Project-2

🗿 Weka Explorer				the last	and so the	-	-	Sec. 1				- I X
Preprocess Classify	Cluster Asso	ciate Select attributes Visual	ize									
Open file	e	Open URL	0	pen DB	Gene	erate		Undo		Edit		Save
Filter												
Choose None												Apply
Current relation Relation: hata ta Instances: 425	blosu_ham veri_c	discretized-weka. filters. unsupe	rvised.attribute.Rem	ove-R1-we Sum	Attributes: 11 of weights: 425	Selected Name: Missing:	attribute Remaining-Op 0 (0%)	en-Duration	Distinct: 6		Type: Nominal Unique: 1 (0%)	
Attributes						No.	Label		Count		Weight	
All		None	Invert	Pa	attern		1 120-140 2 100-120		1 95		1.0 95.0	
No. Nam	ne						3 40-60		42		42.0	
1 SCU-	Name						4 60-80 5 20-40		85		2.0	
2 Crea	ted-Date						6 0-20		200		200.0	
3 Close	ed-Date											
5 Produ	uct-Version											
6 Proje	ect-Phase											
7 Repr	oducibility											
9 MVG	JL											
10 Source	ce-Component											
11 🔲 Rema	aining-Open-Dura	tion				Class: Ren	naining-Open-D	Duration (Nom)				 Visualize All
												200
							9:	5			05	
											80	
									42	I.		
		Remove	2									
						1				2		
Status OK											L	og 📈 x O

Figure D.12 Weka View of Case Study 2A

🕢 Weka Explorer					-		-	-		- 0 X
Preprocess Classify Cluster Associate S	Gelect attributes Visu	alize								
Classifier										
Choose 348 -C 0.25 -M 2										
Test ontions	Classifier output									
Cat up to its									 	*
	Correctly Clas	ssified In:	stances	376		88.4706 %				
Supplied test set Set	Incorrectly CI	lassified : ic	Instances	49	26	11.5294 %				
Cross-validation Folds 10	Mean absolute	error		0.03	20 52					
O Percentage split % 66	Root mean squa	ared error		0.189	92					
More options.	Relative absol	lute error		46.33	76 %					
	Root relative	squared e	rror	56.243	32 %					
(Nom) Pamaining Open Duration	Coverage of ca	ases (0.95	level)	99.529	94 %					
(Non) Kendining-Open-Duration	Mean rei, regi Total Number o	ion size (of Instance	0.95 Ievel)	/6.31	37 8					
Start Stop	IUUII Maabei e	JI INSCANO		125						
Result list (right-click for options)	=== Detailed A	Accuracy By	y Class ===							
20:24:30 - rules.DecisionTable										
20:25:11 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20:25:44 - functions.SimpleLogistic		0	0.009	0 00	1	0 005	0.563	120-140		
20:35:50 - 0.663:340		0.833	0.003	0.972	0.833	0.897	0.965	40-60		
		0	0	0	0	0	0.576	60-80		
		0.871	0.094	0.698	0.871	0.775	0.925	20-40		
		0.86	0.049	0.94	0.86	0.898	0.962	0-20		
	Weighted Avg.	0.885	0.043	0.899	0.885	0.889	0.96			
	=== Confusion	Matrix ===	-							
	a b c	de:	t < Cla 1 a = 1	assified as						E
	0 95 0	0 0) a = 1) b = 1	120-140						
	0 0 35	0 7	0 c = 4	10-60						
	0 0 1	0 0	1 d = 6	50-80						
	1 0 0	0 74 1	0 e = 2	20-40						
	3 0 0	0 25 17:	2 f = (0-20						
	1									
Status OK										Log x0

Figure D.13 Decision Table Results of Case Study 2A

🗿 Weka Explorer	100 B	- 14	-			-		Sec.		- 0 X
Preprocess Classify Cluster Associate 5	Select attributes Visualiz	ze								
Classifier										
Choose 348 -C 0.25 -M 2										
Test options	Classifier output									
O Use training set									 	*
Sunnlied test set	Incorrectly Classi	ified inst ssified In	stances	378		11.0588 \$				
Construction Folds 10	Kappa statistic			0.84						
Cross-validation Folds 10	Mean absolute er	rror		0.040)2					
Percentage split % 66	Root mean square	ed error		0.183	36					
More options	Relative absolut	te error mared err	or	17.708	13 8 01 8					
	Coverage of case	es (0.95 1	Level)	90.823	35 %					
(Nom) Remaining-Open-Duration	Mean rel. region	n size (O.	.95 level)	18.745	51 %					
Ctart Stan	' Total Number of	Instances	3	425						
	Detailed law		Class							
Result list (right-click for options)	Decarred Acc	curacy by	C1033							
20:25:11 - bayes.BayesNet	J	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20:25:44 - functions.SimpleLogistic		0	0	0	0	0	0.705	120-140		
20:35:58 - trees. J48		1	0.006	0.979	1	0.99	0.999	100-120		
		0.833	0.005	0.946	0.833	0.886	0.988	40-60 60-80		
		0.906	0.1	0.694	0.906	0.786	0.944	20-40		
		0.855	0.031	0.961	0.855	0.905	0.972	0-20		
	Weighted Avg.	0.889	0.037	0.903	0.889	0.892	0.971			
	=== Confusion Ma	atrix ===								
	abco	def	≺ cla	ssified as						
	0 1 0 (0 0 0	a = 1	20-140						E
	0 95 0 0	0 0 0	b = 1	00-120						
	0 0 35 0	070	c = 4	0-60						
		0 1 0	d = 6	0-80						
		2 26 171	1 f = 0	-20						
										*
	•			III						Þ
Status										
ОК										Log 💉 X O

Figure D.14 BayesNet Table Results of Case Study 2A

🗿 Weka Explorer			2.0	the second	-	-	-	Sugar de la constante de la co		_ 0 <u>X</u>
Preprocess Classify Cluster Associate 9	Select attributes Visualiz	e								
Classifier										
Choose 348 -C 0.25 -M 2										
Test ontions	Classifier output									
Cuse training set										*
Cartalistat Cat	Correctly Classi	ified Inst	ances	375		88.2353 %				
Suppled test set	Kappa statistic	Dollied II	ISCAILCES	0.827	,	11./04/ 8				
Cross-validation Folds 10	Mean absolute en	rror		0.046	54					
Percentage split % 66	Root mean square	ed error		0.179	95					
More options	Relative absolut	te error		20.421	.98					
	Coverage of case	juared eri	or evel)	53.3/5 95.058	01 8 NA 8					
(Nom) Remaining-Open-Duration	Mean rel. region	n size (0.	.95 level)	21.647	18					
	Total Number of	Instances	3	425						
Start Stop		_								
Result list (right-click for options)	=== Detailed Acc	curacy By	Class ===							
20:24:30 - rules.DecisionTable	1	IP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20:25:44 - functions.SimpleLogistic		0	0	0	0	0	0.149	120-140		
20:35:58 - trees. J48		0.989	0.003	0.989	0.989	0.989	0.991	100-120		
		0.833	0.005	0.946	0.833	0.886	0.977	40-60		
		0 924	0 095	0 707	0 924	0 761	0.111	60-80 20-40		
		0.88	0.08	0.907	0.88	0.893	0.956	0-20		
	Weighted Avg.	0.882	0.056	0.883	0.882	0.881	0.956			
	=== Confusion Ma	atrix ===								
	abco	i e f	≺ cla	ssified as						
	0 1 0 0	0 0 0	a = 1	20-140						
	0 94 0 0	0 0 1	b = 1	00-120						
	0 0 35 0	0 5 2	c = 4	0-60						
	0 0 1 0	0 1 0	d = 6	0-80						=
) 70 15 1 23 176	e = 2	-40 -20						
	•			III						Þ
Status										
ОК										Log x0

Figure D.15 SimpleLogistic Table Results of Case Study 2A

🗿 Weka Explorer			1.0		-		100	and so its		
Preprocess Classify Cluster Associate 5	Select attributes Visu	alize								
Classifier										
Choose 348 -C 0.25 -M 2										
Test setion	Charaiffere austaunt									
est options	Classiner output									
() Use training set	Correctly Clas	sified Ins	stances	378		88.9412 🛚				<u>^</u>
Supplied test set Set	Incorrectly Cl	assified 1	Instances	47		11.0588 %				
() Cross-validation Folds 10	Kappa statisti Mean absolute	.C		0.83	13					
Percentage split % 66	Root mean squa	red error		0.16	56					
Mara antiana	Relative absol	ute error		21.124	12 %					
More options	Root relative	squared en	ror	49.53	L2 %					
Alter) Beneficies On a Develi	Coverage of ca	ises (0.95	level)	96.94	12 %					
(Nom) Remaining-Open-Duration 🔹	Mean rel. regi	on size (().95 level)	25.13	73 %					
Start Stop	Total Number o	if instance	18	425						
Decult liet (right-click for options)	=== Detailed A	ccuracy By	/ Class ===							
20:24:30 - rules.DecisionTable										
20:25:11 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20:25:44 - functions.SimpleLogistic		0	0	0	0	0	0.429	120-140		
20:35:58 - trees.J48		0.989	0.003	0.989	0.989	0.989	0.997	100-120		
		0.833	0	1	0.833	0.909	0.991	40-60		
		0.624	0.012	0.93	0.624	0.746	0.932	20-40		
		0.98	0.187	0.824	0.98	0.895	0.965	0-20		
	Weighted Avg.	0.889	0.091	0.894	0.889	0.881	0.965			
	=== Confusion	Matrix ===								
	a h c	det	F K cla	esified as						
	0 1 0	0 0 0) a = 1	120-140						
	0 94 0	0 0 1	l b=1	100-120						
	0 0 35	0 0 1	1 c = 4	10-60						
	0 0 0	0 0 2	2 d = 6	50-80						
	0 0 0	0 53 32	2 e = 2	20-40						E
	000	0 4 190) I = U	J-20						
										_
	(
Chakar										
OK										Log 🛷 x O

Figure D.16 J48Results of Case Study 2A

E. DETAILS OF CASE STUDY 2B

🔇 Weka Explorer	a s a special					- 0 X
Preprocess Classify Cluster Associate Select attributes	Visualize					
Clusterer						
Choose Simple/Means JUS 10 "webs core Manhatta	oDictance "D first-last" "I 500 "S 10					
Choise Simplex reals 14.3 A Weba.cole.mainatta	Indiscance PCTII Schast P1 300 P3 10					
Cluster mode	Clusterer output					
Output Use training set	GDuncellendigi-tarih	5/31/201110:28:00-AM !	5/31/201110:28:00-AM	5/31/201110:29:00-AM	7/4/20114:28:00-PM	5/31/201110:: 🔺
Supplied test set	Test-Tipi	Yaz?l?mKalifikasyon	Yaz?l?mKalifikasyon	Yaz?l?mKalifikasyon	Yaz?l?mKalifikasyon	Yaz?l?mM?h(
	UrDun-Versiyonu	0.14	0.14	0.02	0.02	
Percentage split % 66	Kaynakland⊔igi-asama Demmeduaibilitu	Kodlama	Kodlama	Kodlama	Kodiama	
Classes to clusters evaluation	L-SLOC	AIWAY5 18533	A1W4Y5 20710	326V 326V	0033 HIMGÅR	nave m
(Nom) Acik-kalma-suresi	MVG	2470	20710	470	957	
Store dustars for visualization	Kaynaklandigi-bilesen	Form	Form	Form	Business	
	dmA1	0	0	0	0	
Terror ethtle ter	dmA2	1	1	0	1	
Ignore attributes	dmA3	0	0	0	0	
Start Stop	dmA4	0	0	1	0	
June 1	dmA5	0	0	0	0	
Result list (right-dick for options)	dmA6	1	1	0	1	
19:02:36 - SimpleKMeans	dmR3	1	1	1	1	
19:12:21 - SimpleKMeans	Acik-kalma-suresi	0-20	100-120	0-20	0-20	
19:12:46 - SimpleKMeans						
19:13:12 - SimpleKMeans						
19:13:28 - SimpleKMeans						_
19:15:32 - SimpleKMeans	Tine solve as build and	1 (5.1)	10			
19:23:38 - SimpleKMeans	lime taken to build mode	i (full training data) : (J.13 Seconds			
19:27:26 - SimpleKMeans	Model and evaluation	on training set				
	Hoder and evaluation	on craining sec				
	Clustered Instances					
						=
	0 121 (28%)					-
	1 26 (68)					
	2 109 (26%)					
	3 32 (8%)					
	4 137 (32%)					
	,					
	4					
						r
Status					ſ	
ОК						Log 🛷 x O

Figure E.1 SimpleKMeans Clustering of Case Study 2B

er Choose None Tent relation Relation: 120430	ute taining-Open-Duration PK) Detinct: 3 abel Count 20-140 0 00-120 1 0-60 0 0-40 0 1-40 2 23	Type: Noninal Unique: 1 (4%) : Weight 0.0 1.0 0.0 0.0 0.0 2.0 2.0 2.0 2.0	
All None Selected attribute. Remove Attributes: 18 Name: Restances: 25 All None Insert Pattern 11 All None Insert Pattern 21 Name 1 Science 20 Science 20 Science 20 Name 1 None Insert Pattern 21 Name 1 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20 Science 20	tute taining-Open-Duration (%) Distinct: 3 abel Count 20-140 0 0-120 1 1 0-60 0 0-40 2 2-30 23	Type: Nominal Urique: 1 (4%) t. Uleight 0.0 1.0 0.0 0.0 2.0 2.0 2.0 2.0	
ret relation Selected atth Relation: 120430	tute naining-Open-Duration (%) Distinct: 3 abel Count 20-140 0 00-120 1 0-60 0 0-40 0 0-40 2 23	Type: Noninal Unique: 1 (4%) : Weight 0.0 1.0 0.0 0.0 0.0 2.0 2.0 2.0	
All None Name	ranng-Open-Duration 29(4) District: 3 abel Count 20-140 0 00-120 1 0-60 0 0-40 0 2-40 2 30 23	Type: Normal Unique: 1 (4%) : 0.0 1.0 0.0 0.0 2.0 2.0 2.0 2.0	
Al None Insert Pattern 11 Al None Insert Pattern 21 Name 1 21 21 21 Name 4 6 52 22 Created-Date 6 6 6 52 22 Created-Date 6 7 8 4 6 <td< td=""><td>abel Count 20-140 0 00-120 1 0-60 0 0-40 0 0-40 2 -20 23</td><td>t Weight 0.0 1.0 0.0 0.0 2.0 2.0 2.0</td><td></td></td<>	abel Count 20-140 0 00-120 1 0-60 0 0-40 0 0-40 2 -20 23	t Weight 0.0 1.0 0.0 0.0 2.0 2.0 2.0	
All None Inset Pattern 11 1 Name 31	20-140 0 100-120 1 0-60 0 0-40 2 2-30 23	0.0 1.0 0.0 2.0 2.0 2.0	
Name Ote Stret Patient 21 Name 3 3 4	00-120 1 0-60 0 0-80 0 0-40 2 -20 23	10 0.0 2.0 23.0	
Name 3 /4 1 SCUHame 5 /2 2 Created-Date 6 /0 3 Created-Date 6 /0 4 Test-Type 5 5 Product-Version 6 6 Project-Phase 7 7 Reproducibity 8 8 -SLOC 9 9 M/G 1	0.60 0 0.40 0 0.40 2 20 23	0.0 0.0 2.0 23.0	
Name 4 (4) 1 SCI-Hane S1 2 Created-Date 6 (0) 3 Cosed-Date 6 (0) 4 Test-Type 5 5 Product-Version 6 6 Project-Phase 7 7 Reproductivery 8 9 MVG 9	0-40 0 0-40 2 -20 23	0.0 2.0 23.0	
1 SCI-Hane Si 2 Created-Date 6 (3 Cosed-Date 6 (4 Test-Type 5 5 Product-Hersion 6 6 Project-Phase 7 7 Reproductifity 8 8 I-SLOC 9 9 MVG 5	0-40 2 -20 23	2.0 23.0	
2 Created-Date 6(3 Closed-Date 4 Text-Type 5 Product-Version 6 Project-Phase 7 Reproducbity 8 U-9.0C 9 MVG	-20 23	23.0	
3 Cosed-Date 4 Text-Type 5 Product-Version 6 Project-Phase 7 Reproduct/Nersion 8 L-9.0C 9 MVG			
5 Product-Version 6 Project-Phase 7 Reproducbility 8 L-9LOC 9 W/G			
6 Project-Phase 7 Reproducbity 8 L-9.0C 9 MVG			
7_H2PTGLCBH7 8_H2LOC 9_M/G			
9 M/G			
7 110			
10 Course / amount			
11 Test 11 Cass Remain	ro-Coen-Duration (Nom)		* Vis.
17 dm42			
13 dm43			
3 014			23
15 dnA5			
35 dnk6			
17 dmR3			
18 Remaining-Open-Duration			

Figure E.2 Weka View of Case Study 2B Cluster 0

🗿 Weka Explorer					-		-	and and		- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visua	alize								
Classifier										
Choose DecisionTable -X 1 -S "weka	a.attributeSelection.Be:	stFirst -D 1 -N	15"							
Test ontions	Classifier output									
Col Upa training out	Classifici ou put									
	Correctly Clas	sified In:	stances	115		95.0413 %				
Supplied test set Set	Incorrectly Cl	assified :	Instances	6		4.9587 %				
Cross-validation Folds 10	Kappa statisti	C		0.850	13					
Percentage split % 66	Root mean squa	red error		0.00	92 71					
	Relative absol	ute error		69.01	55 %					
More options	Root relative	squared e	rror	63.70	38 %					
	Coverage of ca	ses (0.95	level)	99.17	36 %					
(Nom) Remaining-Open-Duration	Mean rel. regi	on size (0.95 level)	85.67	19 %					
Start Stop	Total Number o	f Instance	es	121						
Result list (right-click for options)	=== Detailed A	ccuracy B	y Class ===							
19:07:57 - trees. J48										
19:12:34 - functions.SimpleLogistic		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
19:27:04 - bayes.BayesNet		0	0	0	0	0	?	120-140		
19:27:35 - rules.DecisionTable		1	0.185	0.949	1	0.974	0.907	100-120		
		0.955	0.01	0.955	0.955	0.955	0.99	40-60		
		0	0	0	0	0	0.429	20-40		
		0	0	0	0	0	0.86	0-20		
	Weighted Avg.	0.95	0.146	0.911	0.95	0.93	0.916			
	=== Confusion	Matrix ==:	=							
	abcd	ef≺-∙	- classifie	ed as						
	0 0 0 0	0 0 a	= 120-140							E
	09400	0 0 b	= 100-120							
	0 1 21 0	0 0 c	= 40-60							
	0 0 1 0	0 0 d	= 60-80							
	0 1 0 0	00 e	= 20-40							
	0300	UUII	= 0-20							
	4									· · ·
	×									P
Status OK										Log 💉 x O

Figure E.3 DecisionTable Results of Case Study 2B Cluster 0

🥥 Weka Explorer		1.0		-	Ar 10	-	States -		_ 0 <u>_ X</u>
Preprocess Classify Cluster Associate S	Select attributes Visualize								
Classifier									
Choose BayesNet -D -Q weka.classif	fiers.bayes.net.search.local.K2	P 1 -S BAYES -E	weka.classifiers.b	ayes.net.es	timate.SimpleEst	matorA 0.5			
Test options	Classifier output								
) Use training set	Commenting Classifier	Trataraa	117		06 6042				*
Sunnlied test set	Incorrectly Classified	ed Instances	4		3,3058	5 1			
Oppied totoet Optimized	Kappa statistic		0.90	35					
Cross-validation Folds 10	Mean absolute error		0.01	18					
Percentage split % 66	Root mean squared en	ror	0.10	37					
More options	Relative absolute en	ror d error	9.10	38 8 57 8					
	Coverage of cases ((.95 level)	96.69	42 %					
(Nom) Remaining-Open-Duration	Mean rel. region siz	e (0.95 level) 16.80	44 %					
Chart Chan	Total Number of Inst	ances	121						
	Detailed leaves	. Pr. Class	_						
Result list (right-click for options)	Decalled Accurat	y by class ==	-						
19:07:37 - dees.346 19:12:34 - functions.SimpleLogistic	TP Ra	te FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
19:27:04 - bayes.BayesNet	0	0	0	0	0	?	120-140		
	1	0.111	0.969	1	0.984	0.999	100-120		
	1	0.01	0.957	1	0.978	0.999	40-60 60-80		
	0	0	0	0	0	0.933	20-40		
	0.3	33 0	1	0.333	0.5	1	0-20		
	Weighted Avg. 0.9	67 0.088	0.952	0.967	0.955	0.996			
	Confusion Natria								
	=== Confusion Matrix								
	abcdef	< classifi	ed as						
	0 0 0 0 0 0	a = 120-140							
	0940000	b = 100-120							E
	0 0 22 0 0 0	C = 40-60							
	0 1 0 0 0 0	e = 20-40							
	020001	f = 0-20							
									τ.
	(III						+
Status									
OK								Lo	9 🛷 ×0

Figure E.4 BayesNet Results of Case Study 2B Cluster 0

🗿 Weka Explorer				Case, or	-	-	The second	-		- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visua	alize								
Classifier										
Choose SimpleLogistic -I 0 -M 500 -	-H 50 -W 0.0									
Test options	Classifier output									
O Use training set	C			44.6		05.0670.0				*
Sunnlied test set	Incorrectly Clas	sified in:	Stances Instances	116		95.8678 %				
	Kappa statisti	C		0.877	1					
Cross-validation Folds 10	Mean absolute	error		0.059)					
Percentage split % 66	Root mean squa	red error		0.124	16					
More options	Relative absol	ute error		45.620)5 %					
	Coverage of ca	squared en ses (0.95	level)	100	15 S					
(Nom) Remaining-Open-Duration	Mean rel. regi	on size (().95 level)	83.333	3 8					
	Total Number o	f Instance	25	121						
Start Stop	Detailed 1		. (1							
Result list (right-click for options)	=== Decarred A	ccuracy by	/ 01833 ===							
19:07:37 - trees.346		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
		0	0	0	0	0	?	120-140		
		1	0.148	0.959	1	0.979	0.88	100-120		
		1	0.01	0.957	1	0.978	0.99	40-60		
		0	0	0	0	0	0.000	20-40		
		õ	0	0	0	0	0.114	0-20		
	Weighted Avg.	0.959	0.117	0.919	0.959	0.938	0.867			
	=== Confusion	Matrix ===								
	abcd	ef <	- classifie - 120-140	d as						
	0 94 0 0	0 0 1 a 0 0 1 b	= 120-140							E
	0 0 22 0	0 0 c	= 40-60							
	0 0 1 0	0 0 d	= 60-80							
	0 1 0 0	0 0 e	= 20-40							
	0 3 0 0	0 0 f	= 0-20							
	(m						*
Chalue										
OK										Log 💉 x O

Figure E.5 SimpleLogistic Results of Case Study 2B Cluster 0

🔇 Weka Explorer			1.14	a base o	-	Art 20. 114		and and		_ 0 X
Preprocess Classify Cluster Associate Se	elect attributes Visua	alize								
Classifier										
Choose 348 -C 0.25 -M 2										
Test estions	Classifier output									
lice training cet										4
	Correctly Clas	sified In	stances	114		94.2149 %				
Supplied test set Set	Vanna statisti	assified .	Instances	7	11	5.7851 %				
Cross-validation Folds 10	Mean absolute	error		0.02	96					
Percentage split % 66	Root mean squa	red error		0.13	04					
More options	Relative absol	ute error		22.87	56 %					
	Root relative	squared e:	rror	52.85	94 %					
(Nom) Remaining-Open-Duration	Mean rel. regi	ses (0.95 on size (level) 1.95 level)	95.86 19.60	18 7 8					
	Total Number o	f Instance	28	121	/ 3					
Start Stop										
Result list (right-click for options)	=== Detailed A	ccuracy B	y Class ===							
19:07:57 - trees.J48		TD Date		Burnstation	B11		D00 3	61		
		IP Kate	n Rete	Precision	Recall 0	r-Measure	RUC Area	120-140		
		1	0.259	0.931	1	0.964	0.932	100-120		
		0.909	0	1	0.909	0.952	0.955	40-60		
		0	0	0	0	0	0.375	60-80		
		0	0	0	0	0	0.2	20-40		
	Weighted Aug	0 042	0	0	0 042	0 0.022	0.727	0-20		
	weighted Avg.	0.942	0.201	0.905	0.942	0.922	0.92			
	=== Confusion	Matrix ===	-							
	abcd	ef≺-	- classifie	ed as						
	0 0 0 0 0	u uja n nib	= 120-140							-
	0 2 20 0	0 0 1 c	= 40-60							=
	0 1 0 0	0 0 d	= 60-80							
	0100	00 e	= 20-40							
	0300	0 0 f	= 0-20							
	4									
										,
Status										
UN										 _ _ _

Figure E.6 J48 Results of Case Study 2B Cluster 0

nose None ent relation elation: 120430_	E o Terrenaturala 6					
hoose None entirelation elation: 120430_ terree: 26	i u Trimmaturda 6					
ent relation Telation: 120430	la contrata de la contrat					
kelation: 120430_	denematurala f			Selected attribute		
Contraction and	j (joberer toer	fiters.unsupervised.attribute.Ren	ove Attributes: 18 Sun of weights: 26	Name: Remaining-Open-Duration Missing: 0 (0%)	Distinct: 3	Type: Noninal Unique: 1 (4%)
iutes				No. Label	Count	Unit
				1 120-140	0	0.0
Al	None	inet	Pattern	2 100-120	1	1.0
				3 40-60	0	0.0
. Nane				4 60-80	0	0.0
1 SOLHar	ne			5 20-40	2	2.0
2 Created	Đate			6 0-20	23	23.0
3 Closed-C	late					
4 Test Typ	2ê					
5 Product-	Version					
6 Project 4	hase					
7 Reprodu	cbity					
3 1.400						
9 MVa	·					
11	Lonponent			Class: Remaining-Open-Duration (Nor	1	• Ve
12 4412				commenter of the second second second second second second second second second second second second second se	7	
12 4=42						
14 614						22
15 dmA5						
16 dm46						
17 dnR3						
18 Renain	no-Open-Duration					

Figure E.7 Weka View of Case Study 2B Cluster 1

🗿 Weka Explorer				diam'r	-	AL 84.00	10	and so its		- 0 X
Preprocess Classify Cluster Associate S	Select attributes Visu	alize								
Classifier										
Choose DecisionTable -X 1 -5 "weka	a.attributeSelection.Be	estFirst -D 1 -N	15"							
Test options	Classifier output									
🔘 Use training set	Correctly Clas	sified In:	stances	23		88.4615 %				*
Supplied test set Set	Incorrectly Cl	lassified (Instances	3		11.5385 🖁				
Cross-validation Folds 10	Kappa statisti	lc		0						
Percentage split % 66	Root mean sour	error ared error		0.11	26					
Nore options	Relative absol	lute error		99.55	31 %					
	Root relative	squared e	rror	99.85	51 %					
(Nom) Remaining-Open-Duration	Coverage of ca	ises (0.95	level)	100	8					
	Total Number (on size (of Instance	0.95 IEVEI) PS	26	55 8					
Start Stop	iour nuber (
Result list (right-click for options)	=== Detailed A	Accuracy B	y Class ===							
19:36:16 - rules.DecisionTable		70 D	TD Date		D11	E Maran	D00 1	(1)		
		1P Kate 0	IF Kate	Precision 0	0 Kecall	r-measure 0	xUC Area	120-140		
		0	0	0	0	0	0.1	100-120		
		0	0	0	0	0	?	40-60		
		0	0	0	0	0	?	60-80		
		0	0	0	0	0 020	0.229	20-40		
	Weighted Avg.	0.885	0.885	0.783	0.885	0.83	0.232	0-20		
	=== Confusion	Matrix ==	=							
	abcd	ef≺-	- classifie	ed as						
	0 0 0 0	0 0 a	= 120-140							
	0 0 0 0	01 b	= 100-120							
		0 0 0 0 0 d	= 40-60 - 60-80							
	0 0 0 0	021e	= 20-40							
	0 0 0 0	023 f	= 0-20							
										τ.
	(•
Status									ſ	100
OK										Log V

Figure E.8 DecisionTable Results of Case Study 2B Cluster 1

🥥 Weka Explorer			in the second	-	-	-	and the second	
Preprocess Classify Cluster Associate S	Select attributes Visualize							
Classifier								
Choose BayesNet -D -Q weka.classif	fiers.bayes.net.search.local.K2	-P 1 -S BAYES -E	weka.classifiers.b	ayes.net.es	timate.SimpleEst	matorA 0.5		
Test options	Classifier output							
O Use training set	Correctly Classified	Instances	25		06 1539	5		*
Supplied test set Set	Incorrectly Classifie	i Instances	1		3.8462	•		
Ornes-validation Endds 10	Kappa statistic		0.78	33				
Dercentane split % 66	Mean absolute error		0.019	93				
Percentage spit 70 00	Root mean squared err Relative absolute err	or	16.94	55 36 %				
More options	Root relative squared	error	56.818	36 %				
Alan) Demaining Ocean Demainer	Coverage of cases (0.	95 level)	96.15	38 %				
(vom) kemaining-Open-Duration •	Mean rel. region size	(0.95 level)	20.512	28 %				
Start Stop	TOCAL NUMBER OF THESE	1000	20					
Result list (right-click for options)	=== Detailed Accuracy	By Class ===						
19:36:16 - rules.DecisionTable		DD Date	Description	P	E Marana	D00 1	61	
19:36:35 - bayes.BayesNet	IF Rat	e FF Kate	Precision 0	0 Kecall	r-Measure	XUC Area	120-140	
	0	0	0	0	0	0.32	100-120	
	0	0	0	0	0	?	40-60	
	0	0	0	0	0	?	60-80 20-40	
	1	0.333	0.958	1	0.979	0.899	20-40	
	Weighted Avg. 0.96	0.295	0.925	0.962	0.943	0.884		
	=== Confusion Matrix							
	abcdef	(classifie	ed as					
	0 0 0 0 0 0 0	a = 120-140						
	0 0 0 0 0 1	b = 100-120						E
		c = 40-60 d = 60-80						
	0 0 0 0 2 0 1	e = 20-40						
	0 0 0 0 0 23	f = 0-20						
	4							
Chable								
OK								Log 💉 x O

Figure E.9 BayesNet Results of Case Study 2B Cluster 1

🗿 Weka Explorer			1.1	i han i	-	-	-			- 0 X
Preprocess Classify Cluster Associate 5	Select attributes Visu	alize								
Classifier										
Choose SimpleLogistic -I 0 -M 500	-H 50 -W 0.0								 	
Test options	Classifier output									
🔘 Use training set	Correctly Clas	sified In	stances	25		96.1538 8				*
Supplied test set Set	Incorrectly Cl	assified	Instances	1		3.8462				
Cross-validation Folds 10	Kappa statisti	.C		0.783	33					
© Percentane solit % 66	Mean absolute	error		0.050	67					
Fercentage spin 78 00	Relative absol	ute error		49.76	J/ 62 %					
More options	Root relative	squared e	rror	64.41	54 %					
	Coverage of ca	ses (0.95	level)	100	8					
(Nom) Remaining-Open-Duration	Mean rel. regi	on size (0.95 level)	74.359	98					
Start Stop	Total Number o	f Instanc	85	26						
	Detailed A	aguragu B	u Class							
Result list (right-click for options)	=== Decarred A	ccuracy b	y C1855 ===							
19:36:35 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
19:37:03 - functions.SimpleLogistic		0	0	0	0	0	?	120-140		
		0	0	0	0	0	0.2	100-120		
		0	0	0	0	0	?	40-60		
		0	0	0	1	1	?	60-80 20-40		
		1	0.333	1 0.958	1	0.979	0.71	0-20		
	Weighted Avg.	- 0.962	0.295	0.925	0.962	0.943	0.713			_
	=== Confusion	Matrix ==	-							
			-1							
		e I <	- CIASSIIIE - 120-140	a as						
	0 0 0 0	0 1 b	= 100-120							E
	0 0 0 0	0 0 c	= 40-60							
	0 0 0 0	0 0 d	= 60-80							
	0 0 0 0	20 e	= 20-40							
	0 0 0 0	023 f	= 0-20							
	(*
Status										DO ANAL XO
e										

Figure E.10 Simple Logistic Results of Case Study 2B Cluster 1

🔕 Weka Explorer			1.1	i han i		-	-	and the second		- 0 X
Preprocess Classify Cluster Associate 5	Select attributes Visu	ualize								
Classifier										
Choose J48 -C 0.25 -M 2										
Test onlines	Classifier output									
Clear training ant										
	Correctly Cla	ssified In:	stances	23		88.4615 4				
Supplied test set Set	Incorrectly C	lassified :	Instances	3		11.5385 §				
Cross-validation Folds 10	Kappa Statist	10		0 05'	,					
Percentage split % 66	Root mean sou	ared error		0.189	<u>.</u> 98					
Muuutuu	Relative abso	lute error		45.58	14 %					
More opuons	Root relative	squared en	ror	93.53	93 %					
	Coverage of c	ases (0.95	level)	88.463	15 %					
(Nom) Remaining-Open-Duration 🔹	Mean rel. reg	ion size ().95 level)	20.512	28 %					
Start Stop	Total Number	of instance	23	26						
Deput liet (right dick for options)	=== Detailed	Accuracy B	/ Class ===	1						
19:36:16 - rules DerisionTable										
19:36:35 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
19:37:03 - functions.SimpleLogistic		0	0	0	0	0	?	120-140		
19:43:15 - trees.348		0	0	0	0	0	0.04	100-120		
		0	0	0	0	0	?	40-60		
		0	0	0	0	0	2 0 917	00-00 20-40		
		1	1	0.885	1	0.939	0.623	0-20		
	Weighted Avg.	0.885	0.885	0.783	0.885	0.83	0.623			
	=== Confusion	Matrix ===	•							
			-1							
		0.01.a	- Classifie = 120-140	0 83						Ξ
	0 0 0 0	0 1 b	= 100-120							
	0 0 0 0	0 0 c	= 40-60							
	0 0 0 0	0 0 d	= 60-80							
	0 0 0 0	02 e	= 20-40							
	0 0 0 0	023 f	= 0-20							
	1									
										r
Status										
VI									Lug	

Figure E.11 J48 Results of Case Study 2B Cluster 1

	Open fie	Open URL	Open I	8 Gene	rateU	ndo Ed	t Save
g							
Choose	None						
rrent rela	stan				Selected attribute		
Relation:	120430_1	tj _deneme1-veka.	fiters unsupervised attribute R	enov Attributes: 18	Nane: Renaining-Open-Durato	n	Type: Noninal
nstances	: 109			Sum of weights: 109	Mesing: 0 (0%)	District: 3	Unque: 0 (0%)
titutes					No. Label	Count	Weight
_	A	None	Inet	Pattern	1 120-140	0	0.0
					2 100-120	0	0.0
a.	Name				3 40-60	13	13.0
	ESO Mare				4 60-80	0	0.0
	Created-Date				5 20-40	11	11.0
3	Closed-Date				0 0/20	63	02.0
4	Test-Type						
5	Product-Version						
6	Project-Phase						
1	Reproducibility						
8	1-900						
9	MIG						
	A REAL REAL REAL PROPERTY.				Class: Remaining-Open-Duration (No	m	• Ve
10							
10	dnA1					-1	-
10 11 12 13	dmA1 dmA2 dmA3					-1	
10 11 12 13 14	dnA1 dnA2 dnA3 dnA4					-1	
10 11 12 13 14 15	dnA1 dnA2 dnA3 dnA4 dnA5					7	6
10 11 12 13 14 15 15	dnA1 dnA2 dnA3 dnA4 dnA5 dnA5						
10 11 12 13 14 15 15 15 17	500 (E-04)(048)(dhA1 dhA2 dhA3 dhA4 dhA5 dhA5 dhA6 dhR3					-	

Figure E.12 Weka View of Case Study 2B Cluster 2

🗿 Weka Explorer			1.1	a bana	that said to	Act and only	-	and and		_ 0 _ X
Preprocess Classify Cluster Associate 5	Select attributes Visua	alize								
Classifier										
Choose DecisionTable -X 1 -S "web	a.attributeSelection.Be	estFirst -D 1 -N	5"							
Tect ontions	Classifier output									
Cite training set	Left-clic	k to edit pro	perties for this	object, right-cl	ick/Alt+Shi	ft+left-click for	menu			
	Correctly Clas	ssified In:	stances	101		92.6606	1			
Supplied test set Set	Vanna statisti	lassified . ic	Instances	0.80	79	7.3394	5			
Cross-validation Folds 10	Mean absolute	error		0.11	99					
Percentage split % 66	Root mean squa	ared error		0.19	54					
More options	Relative absol	lute error		91.12	34 %					
	Root relative	squared en	rror	78.67	58					
(Nom) Remaining-Open-Duration	Coverage of ca	ises (0.95	level)	100	8					
	Total Number o	of Instance	1.95 IEVEL) 28	109	11 8					
Start Stop										
Result list (right-click for options)	=== Detailed A	Accuracy By	y Class ===							
19:48:40 - rules.DecisionTable										
		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class 120-140		
		0	0.010	0	0	0	?	100-120		
		1	0	1	1	1	1	40-60		
		0	0	0	0	0	?	60-80		
		0.636	0.041	0.636	0.636	0.636	0.909	20-40		_
	Ned about 1 hours	0.953	0.083	0.976	0.953	0.964	0.964	0-20		
	weighted Avg.	0.927	0.069	0.945	0.927	0.935	0.963			
	=== Confusion	Matrix ===								
	a b c d	ef≺-	- classifie	d as						
	0 0 0 0	00 a	= 120-140							E
	0 0 0 0	a 100	= 100-120							
	0 0 0 0	0 0 d	= 60-80							
	2000	72 e	= 20-40							
	0 0 0 0	481 f	= 0-20							
										·
	•									•
Status OK										Log 🛷 x O

Figure E.13 DecisionTable Results of Case Study 2B Cluster 2

🔇 Weka Explorer				i lana	-	-	-	and the second s	
Preprocess Classify Cluster Associate S	Select attributes Visua	alize							
Classifier									
Choose BayesNet -D -Q weka.classif	fiers.bayes.net.search	.local.K2P	1 -S BAYES -E (weka.classifiers.b	ayes.net.es	timate.SimpleEst	imatorA 0.5		
Test ontions	Classifier output								
Clice training set									
	Correctly Clas	sified In	stances	100		91.7431	1		
Supplied test set Set	Incorrectly Cl	assified :	Instances	9 0.70	20	8.2569	1		
Cross-validation Folds 10	Mean absolute	error		0.02	29 98				
Percentage split % 66	Root mean squa	red error		0.15	78				
More options	Relative absol	ute error		22.60	8 8				
	Root relative	squared e	rror	63.55	32 %				
(Nom) Remaining-Open-Duration	Coverage of ca	ises (0.95	level)	93.57	38 148				
(rem) remaining open bereation	Total Number o	of Instance	0.90 IEVEL) PR	10.00	14 8				
Start Stop									
Result list (right-click for options)	=== Detailed A	Accuracy B	y Class ===						
19:48:40 - rules.DecisionTable									
19:49:02 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class	
		0	0	0	0	0	?	120-140	
		1	0	1	1	1	1	40-60	
		0	0	0	0	0	?	60-80	
		0.818	0.071	0.563	0.818	0.667	0.884	20-40	
		0.918	0.083	0.975	0.918	0.945	0.954	0-20	
	Weighted Avg.	0.917	0.072	0.936	0.917	0.924	0.953		
	=== Confusion	Matrix ===	=						
	a b c d	ef≺-	- classifie	d as					
	0 0 0 0	0 0 a	= 120-140						
	0 0 0 0	001b	= 100-120						E
	0 0 0 0	0 0 1 0	= 40-60						
	0 0 0 0	92 e	= 20-40						
	0 0 0 0	778 f	= 0-20						
									τ.
	•								•
Status									
ОК									Log 🛷 x 0

Figure E.14 BayesNet Results of Case Study 2B Cluster 2

🔇 Weka Explorer		100	the second second	No. of Concession, Name	-		
Preprocess Classify Cluster Associate S	Select attributes Visualize						
Classifier							
Choose SimpleLogistic -I 0 -M 500 -	-H 50 -W 0.0						
Test ontions	Classifier output						
Cities training set							A
	Correctly Classified In	nstances	99	90.8257 %			
Supplied test set Set	Incorrectly Classified	Instances	10	9.1743 8			
Cross-validation Folds 10	Mean absolute error		0.0359				
Percentage split % 66	Root mean squared erro:	c .	0.1454				
More options	Relative absolute error	c	27.2732 %				
	Root relative squared	error	58.5413 %				
(Nom) Remaining-Open-Duration	Coverage of cases (0.9	() A5 level)	98.1651 % 21.8654 %				
(non) remaining open boroton	Total Number of Instan	(U.95 IEVEI) 288	21.0034 % 109				
Start Stop							
Result list (right-click for options)	=== Detailed Accuracy 1	By Class ===					
19:48:40 - rules.DecisionTable							
19:49:02 - bayes.BayesNet	TP Rate	FP Rate Pr	ecision Recall	F-Measure	ROC Area	Class	
19:49:30 - functions.SimpleLogistic	0	0	0 0	0	?	120-140	
	1	0	1 1	1	1	40-60	
	0	0	0 0	0	?	60-80	
	0.455	0.041	0.556 0.455	0.5	0.792	20-40	
	0.953	0.25	0.931 0.953	0.942	0.904	0-20	
	Weighted Avg. 0.908	0.199	0.901 0.908	0.904	0.904		
	=== Confusion Matrix ==	=					
	abcdef <	classified a	3				
		a = 120 - 140 a = 100 - 120					
	0 0 13 0 0 0	c = 40-60					Ξ
	0 0 0 0 0 0 0 0	i = 60-80					
	0 0 0 0 5 6 6	e = 20-40					
	0 0 0 0 4 81 :	E = 0-20					
	1						• • •
							,
Status OK							Log x0

Figure E.15 SimpleLogistic Results of Case Study 2B Cluster 2

🗿 Weka Explorer			100	diam'r	-	AL INCOME	-	and the second		_ 0 _ X
Preprocess Classify Cluster Associate 5	Select attributes Visua	alize								
Classifier										
Choose J48 -C 0.25 -M 2										
Test selfine	Charal Constants									
lest options	Classifier output									
O Use training set	Correctly Clas	sified In:	stances	102		93.578	ł			^
Supplied test set Set	Incorrectly Cl	assified 3	Instances	7		6.422	8			
Cross-validation Folds 10	Kappa statisti	.C		0.828	32					
Percentage split % 66	Root mean squa	red error		0.032	12					
	Relative absol	ute error		25.006	56 %					
More options	Root relative	squared en	ror	54.025	57 %					
	Coverage of ca	ses (0.95	level)	97.247	17 %					
(Nom) Remaining-Open-Duration	Mean rel. regi	on size ().95 level)	18.807	3 8					
Start Stop	Total Number o	f Instance	5	109						
Result list (right-click for options)	=== Detailed A	ocuracy B	/ Class ===							
19:48:40 - rules.DecisionTable										
19:49:02 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
19:49:30 - functions.SimpleLogistic		0	0	0	0	0	2	120-140		
		1	0 0	1	1	1	1	40-60		
		0	0	0	0	0	?	60-80		
		0.727	0.041	0.667	0.727	0.696	0.844	20-40		
		0.953	0.125	0.964	0.953	0.959	0.917	0-20		
	Weighted Avg.	0.936	0.102	0.939	0.936	0.937	0.92			_
	=== Confusion	Matrix ===								
	a b c d	ef≺	- classifie	d as						
	0 0 0 0	00 a	= 120-140							
	0 0 0 0	001b	= 100-120							E
	0 0 13 0	0 0 1 0	= 40-60							
	0 0 0 0	831e	= 20-40							
	0 0 0 0	481 f	= 0-20							
										Ŧ
	•			II						Þ
Status										
ОК										Log 📣 x0

Figure E.16 J48 Results of Case Study 2B Cluster 2

(Open file	Open URL	Open D8	Gene	sate	Undo	Edt Save
2	None						
relat	ion				Selected attribute		
ton: CBS:	120430] Jr.J 32	deneme1-veka.fite	rs unsupervised attribute. Ren	rove Attributes: 18 Sum of weights: 32	Name: Remaining-Ope Mosing: 0 (0%)	en-Duration Distinct: 2	Type: Nominal Unique: 0 (0%)
8					No. Label	Count	Weight
-	4	None	Inet	Pattern	1 120-140	0	0.0
					2 100-120	0	0.0
	Name				3 40-60	0	0.0
-	Tene .				4 60-80	0	0.0
-	ocumane Crusted Date				5 20-40	23	23.0
1	Creat/Jate				6 0-20	9	9.0
1	Test-Tipe						
ŝ	Product-Version						
6	Project-Phase						
7	Reproducibility						
8	1-9.00						
9	M/G						
10	Source-Component					union Atom)	
11	dnA1				uass: kenaning-upen-uu	urason (viom)	• 13
-	0TA2						
12							20
12	414						
12 13 14 14	dnA4 dnA5						
12 13 14 15 16	án44 án45 án46						
12 13 14 15 16 17	ánA4 ánA5 ánA6 ánR3						
12 13 14 15 15 15 17	dmA4 dmA5 dmA6 dmR3 Remaining-Open-Duration						
12 13 14 15 15 17 13	dmA4 dmA5 dmA6 dmR3 Remaining-Open-Duration						
12 13 14 15 15 17 18	dmA4 dmA5 dmA6 dmR3 temaning-Open-Duraton						
12 13 14 15 15 17 18	dmA4 dmA5 dmA6 dmR3 Remaining-Open-Duration						
12 13 14 15 15 17 13	dn44 dn45 dn46 dn83 Besaring-Open-Duraton						
12 13 14 15 16 17 13	dn44 dn45 dn46 dn83 desaintrg-Open-Duraton						
12 13 14 15 15 17 13	dn44 dn45 dn46 dn83 Resaing-Open-Duraton						
12 13 14 15 15 17	dn44 dn45 dn46 dn83 Resaing-Open-Duraton						
12 13 14 15 16 17 18	dn44 dn45 dn46 dn83 Remaing-Open-Duraton	Reno	re				

Figure E.17 Weka View of Case Study 2B Cluster 3
🔇 Weka Explorer			100	diam'r	-	-	-			
Preprocess Classify Cluster Associate S	Select attributes Visua	lize								
Classifier										
Choose DecisionTable -X 1 -5 "weka	a.attributeSelection.Be:	stFirst -D 1 -N	5"							
Test options	Classifier output									
🔘 Use training set	Correctly Clas	sified In:	stances	28		87.5	\$			*
Supplied test set Set	Incorrectly Cl	assified :	Instances	4		12.5	8			
Cross-validation Folds 10	Kappa statisti	с		0.668	34					
	Mean absolute error			0.130	17					
Percentage split % bb	Root mean squa	red error		0.21	51					
More options	Relative absol	ute error		01.07. 91.62(14 8 10 8					
	Coverage of ca	ses (0.95	level)	100	8					
(Nom) Remaining-Open-Duration 🔹	Mean rel. regi	on size ().95 level)	86.979	92 8					
Start Stop	Total Number o	f Instance	25	32						
Result list (right-click for options)	=== Detailed A	ccuracy By	/ Class ===							
20:07:47 - rules.DecisionTable										
		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
		0	0	0	0	0	?	120-140		
		0	0	0	0	0	?	100-120		
		0	0	0	0	0	: 2	40-00 60-80		
		0.957	0.333	0.88	0.957	0.917	0.713	20-40		
		0.667	0.043	0.857	0.667	0.75	0.727	0-20		
	Weighted Avg.	0.875	0.252	0.874	0.875	0.87	0.717			
	=== Confusion 1	Matrix ===								
	abcd	ef <	- classifie	d as						
	0 0 0 0	00 a	= 120-140							=
	0 0 0 0	0 0 b	= 100-120							
	0 0 0 0	0 0 c	= 40-60							
	0 0 0 0	0 0 d	= 60-80							
	0 0 0 0 2	21 e	= 20-40							
		3011	= 0-20							
	(III						•
Chhu										
OK										Log x0

Figure E.18 DecisionTable Results of Case Study 2B Cluster 3

🥥 Weka Explorer				i land			-	and the second		
Preprocess Classify Cluster Associate S	Select attributes Visu	ualize								
Classifier										
Choose BayesNet -D -Q weka.dassif	fiers.bayes.net.searc	h.local.K2P	1 -S BAYES -E (weka.classifiers.b	ayes.net.es	timate.SimpleEs	timatorA 0.5			
Test options	Classifier output									
🔘 Use training set	Correctly Cla	ssified In:	stances	30		93.75	ł			*
Supplied test set Set	Incorrectly C	lassified :	Instances	2		6.25	ł			
Cross-validation Folds 10	Kappa statist	ic		0.83	12					
	Mean absolute error			0.03	24					
Percentage split % 00	Root mean squared error Relative absolute error			0.15	1					
More options	Relative abso	Relative absolute error Root relative squared error			5 57 8					
	Coverage of c	ases (0.95	level)	93.75	92 3 8					
(Nom) Remaining-Open-Duration 🔹	Mean rel. reg	ion size ().95 level)	17.70	3 8					
	Total Number	of Instance	28	32						
Start Stop										
Result list (right-click for options)	=== Detailed	Accuracy By	/ Class ===							
20:07:47 - rules.DecisionTable		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20108:15 - Dayes.Bayesvet		0	0	0	0	0	?	120-140		
		0	0	0	0	0	?	100-120		
		0	0	0	0	0	?	40-60		
		0	0	0	0	0	?	60-80		
		1	0.222	0.92	1	0.958	0.874	20-40		
		0.778	0	1	0.778	0.875	0.874	0-20		
	Weighted Avg.	0.938	0.16	0.943	0.938	0.935	0.874			
	=== Confusion	Matrix ===								
	abcd	e f <	- classifie	d as						
	0 0 0 0	0 0 a	= 120-140							
	0 0 0 0	0 0 b	= 100-120							E
	0 0 0 0	0 0 c	= 40-60							
	0 0 0 0	0 0 d	= 60-80							
	0 0 0 0	23 0 e	= 20-40							
	0000	2 / 1	= 0-20							
	1									•
Status OK										Log 🛷 X O

Figure E.19 BayesNet Results of Case Study 2B Cluster 3

🖸 Weka Explorer	d'rigites/	1								
Preprocess Classify Cluster Associate	Select attributes Visua	lize								
Classifier										
Choose SimpleLogistic -I 0 -M 500	-H 50 -W 0.0									
Test ontions	Classifier output									
Clise training set										A
	Correctly Class	sified In:	stances	26		81.25	8			
Suppled test set	Kappa statisti	assiiied . c	Instances	0.536	52	10./5	\$			
Cross-validation Folds 10	Mean absolute	error		0.066	52					
Percentage split % 66	Root mean squa	red error		0.235	53					
More options	Relative absolu	ute error		41.390	17 %					
	Root relative :	squared e	rror	88.880)4 %					
(Nom) Remaining-Open-Duration	Coverage of ca	ses (0.95	level)	87.5	8					
(winy remaining open bardaon +	Total Number of	on size (f Instanci	0.95 IEVEL)	20.312	25 8					
Start Stop	iooui numbei o.	I INStanto		52						
Result list (right-click for options)	=== Detailed A	ccuracy B	y Class ===							
20:07:47 - rules.DecisionTable										
20:08:15 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
20:09:05 - functions.SimpleLogistic		0	0	0	0	0	?	120-140		
		0	0	0	0	0	? 2	40-60		
		0 0	0	0	0	0 0	?	60-80		
		0.87	0.333	0.87	0.87	0.87	0.775	20-40		
		0.667	0.13	0.667	0.667	0.667	0.775	0-20		
	Weighted Avg.	0.813	0.276	0.813	0.813	0.813	0.775			
	=== Confusion 1	Matrix ==	-							
	abcd	ef≺-	- classifie	d as						
	0 0 0 0	0 0 a	= 120-140							
	0 0 0 0	0 0 b	= 100-120							
	0 0 0 0	0 0 c	= 40-60							
		0 0 d	= 60-80							=
		03 e 36 f	= 20-40							
			0 20							
	(III						•
Status										
OK										Log 💉 x0

Figure E.20 SimpleLogistic Results of Case Study 2B Cluster 3

Internet Security Guero Market Security Classified Instances 25 78.125 % Conservation <td< th=""><th>🥥 Weka Explorer</th><th>A</th><th></th><th>1.14</th><th>i lanci</th><th>-</th><th>-</th><th>-</th><th>and so is</th><th></th><th></th></td<>	🥥 Weka Explorer	A		1.14	i lanci	-	-	-	and so is		
Clearly Clearly Doors D4 < 0.57 M2 Tetrytors Occentify Classified Instances 75 © Ubertrangset Sec. © Occendation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Conservation Folds © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil % 650 © Precentage poil <	Preprocess Classify Cluster Associate S	Select attributes Visua	lize								
Choose 348 - 6.25 AH 2 Categoins Contently Classified Instance 25 78.125 % Supple faster Contently Classified Instance 7 21.075 % Supple faster Contently Classified Instance 7 21.075 % Maps statistic 0.4766 Maps statistic 0.4766 More option More option Contact y Classified Instances 7 21.075 % More option More option Statistic (globadic for option) Contact y Classified Instances 32 Statistic (globadic for option) Encot relative space for case (0.55 level) 7.042 Å Total Winter of Instances 32 Statistic (globadic for option) TP fance FP fance Precision Recult F-Measure ENC Areas Class Contact y Classified Instances 32 Statistic Statistic (globadic for option) Classified Instances 32 Statistic (globadic for option) Statistic (globadic for option) Descriptic Precision Recult F-Measure Statistic (globadic for option) Statistic (globadic for option) Statistic (globadic for option) Descriptic Precision Recult F-Measure Statististic (globadic for option) Statisti	Classifier										
Text uptions Classified instances 25 78.125 8 Suppled lenset Set Rappe statistic 0.4766 Mana aboutive error 0.083 Workstopen Public More contra Root statistic 0.4766 More contra Statistic 0.4766 More contra Root statistic 0.4766 More contra Statistic space for core 0.483 More contra Statistic space for core 0.4364 Non statistic space for core 0.4364 More contra Statistic space for core 0.4637 Non No No	Choose 348 -C 0.25 -M 2										
Tet of or Outs taining set Outs taining set Correctly Classified Instances 25 78.125 % Supplet last set. Decorrectly Classified Instances 7 21.075 % @ Oosr-widdom Fok [3] Decorrectly Classified Instances 7 21.075 % @ Perenting set % Decorrectly Classified Instances 7 21.075 % @ Oosr-widdom Fok [3] Decorrectly Classified Instances 0.4766 Mem options											
Oute training eff Correctly Classified Instances 25 78.125 % @ Consulation Fide B Description (1) Description (1) Description (1) @ Consulation Fide B Description (1) Description (1) Description (1) @ Consulation Fide B Description (1) Set B Description (1) @ More outrom. Set B Description (1) Description (1) @ More outrom. Set B Description (1) Description (1) @ Domain description (1) Set B Description (1) Description (1) @ Domain description (1) Set B Description (1) Description (1) @ Domain description (1) Set B Description (1) Description (1) @ Domain description (1) Description (1) Description (1) Description (1) @ Domain description (1) Description (1) Description (1) Description (1) Description (1) @ Domain description (1) Description (1) Description (1) Description (1) Description (1) @ Domain description (1) Description (1) Description (1) Description (1) Description (1) @ Domain description (1) Descriptio	Test options	Classifier output									
Suppled test set Set @ Cross-validator Fide @ Cross-validatorecro.X# Fide	Use training set	Correctly Class	sified In:	stances	25		78.125	8			A
0 Cosswaldstor vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt 5 6 vice lags solt Site vice lags solt Site vice lags solt Site	O Supplied test set Set	Incorrectly Cla	assified (Instances	7		21.875	ł			
Image: State More plans. Hean absolute error 0.2443 More plans Root nean squared error 92.4653 % More plans Coverage of cases (0.55 level) 90.625 % Non Remaing Oper Duration When plans is (0.55 level) 90.625 % Non Remaing Oper Duration When nell.region size (0.55 level) 90.625 % Non Average of cases (0.55 level) 90.625 % Hean rel. region size (0.55 level) 2007.97. rules Descontable 2007.97. rules Descontable 20 2009.05. fondrom Supplicit 0 0 0 2009.05. fondrom Supplicit 0 0 0 2 2009.05. fondrom Supplicit 0 0 0 2 2009.05. fondrom Supplicit 0 0 0 2 2009.05. fondrom Supplicit 0 0 0 0 2 2009.05. fondrom Supplicit 0 0 0 2 2 2009.05. fondrom Supplicit 0 0 0 2 2 200.05. fondrom Supplicit 0 0 0 </th <th>Cross-validation Folds 10</th> <th>Kappa statisti</th> <th>с</th> <th></th> <th>0.470</th> <th>56</th> <th></th> <th></th> <th></th> <th></th> <th></th>	Cross-validation Folds 10	Kappa statisti	с		0.470	56					
0 Field More quiters 0 0 0 deel adjuated error 0.2448 Realistive absolute error 51.807 % Realistive absolute error 20.4658 % Coverage of cases (0.35 level) 90.625 % Man Remaing Querdurator V Stat Stop Stat Stop Resultst(ghrddx for option) 20.0015-bayes Saymetet 20:0035-fordrom Saymeted 0 0 0 0 0 0 0 0 0 2 0 0 0 0 2 0 0 0 0 2 0 0 0 2 20-140 0 0 0 0 2 20-140 0 0 0 0 2 20-140 0 0 0 0 2 20-140 0 0 0 0 2 20-10 0.826 0.333 0.864 0.822 0.703 0-20 Heighted Avg. 0.781 0.226 0.703 0-20	Dercentres ofit	Mean absolute (error		0.08	3					
More options Packadure adjusted error 92.4658 % Non Remaining-OpenDuration Non relative adjusted error 92.4658 % Start Stop Start Stop Start Stop 200747-rules.DesignTable 20.078 % 200747-rules.DesignTable 0 200747-rules.DesignTable 0 200747-rules.DesignTable 0 200805-functions.produces 0 0 0	Percentage spin % 00	Root mean squa:	red error		0.244	10					
Nom/Remaining-Open-Duration Coverage of cases (0.45 level) 90.625 % Start Stop Start Stop Start Stop 2007.47-0462.065007.000 22.007.47 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 2007.47-0462.065007.000 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 <th>More options</th> <th>Root relative</th> <th>squared e</th> <th>mor</th> <th>92.46</th> <th>77 % 58 %</th> <th></th> <th></th> <th></th> <th></th> <th></th>	More options	Root relative	squared e	mor	92.46	77 % 58 %					
Num Remains OpenDuration Hean rel. region size (0.35 level) 27.6042 % Start Stop		Coverage of ca	ses (0.95	level)	90.62	58					
Stat Stop Realtist(myhtdx for options) == Detailed Accuracy By Class === 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2007:47 - rules.DecisionTable 0 0 0 0 2 100-120 0 0 0 0 0 2 40-60 0 0 0 0 2 40-60 0.826 0.333 0.864 0.826 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 == Confusion Matrix === a b c d d d d 0 0 0 0	(Nom) Remaining-Open-Duration	Mean rel. regio	on size ().95 level)	27.60	12 8					
Stat Stop Result is (ight-ddx for optons) === Detailed &couracy By Class === 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2008:15 - bayes.BayesNet 0 0 0 0 2 120-140 2009:05 - functions.Smplet.ogist 0 0 0 0 2 100-120 0 0 0 0 0 2 100-120 0 0 0 0 0 2 60-80 0.826 0.333 0.864 0.826 0.703 20-40 0.667 0.174 0.6 0.662 0.703 0-20 Weighted Avg. 0.781 0.288 0.709 0.701 0.705 0.703 === Confluxion Matrix == - a b c d e f - classified as 0 0 0 0 0 0 0 0 0 0 0 0 0		Total Number o	f Instance	3	32						
Result ist (ight-did for options) === Detailed Accuracy By Class ==== 2007:47 - rules.DecisionTable 0 0 0 0 2 120-140 2009:05: 5-bayes.BayesNet 0 0 0 0 0 2 120-140 2009:05: 5-bayes.BayesNet 0 0 0 0 0 2 100-120 0 0 0 0 0 0 2 100-120 0 0 0 0 0 2 40-60 0 0 0 0 0 2 40-60 0.826 0.333 0.864 0.826 0.703 20-40 0.667 0.174 0.6 0.667 0.703 20-40 0.677 0.174 0.6 0.667 0.703 20-40 0 0 0 0 0 10.785 0.703	Start Stop										
2007/47 - rdes DecisionTable TP Rate FP Rate Precision Recall F-Measure ROC Area Class 2008:15 - bayes.BayesNet 0 0 0 0 2 120-140 2009:05 - functions.SimpleLogistx 0 0 0 0 2 100-120 0 0 0 0 0 2 100-120 0 0 0 0 2 40-60 0 0 0 0 2 60-80 0.626 0.333 0.864 0.826 0.844 0.703 20-40 0.667 0.174 0.6 0.667 0.703 0.703 -20 Weighted Rvg. 0.781 0.785 0.703 = Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0.785 0.703 a b c d e f < classified as 0 0 0 0 0 0 0 0 0 0 0 0 0	Result list (right-dick for options)	=== Detailed A	ccuracy By	/ Class ===							
2008:15-bayes.8ayeshet 0 0 0 0 0 2009:05-functors.Simplet.ogstr. 2009:05-functors.Simplet.ogstr. 0 0 0 0 0 2 100-120 0 0 0 0 0 0 2 40-60 0 0 0 0 0 2 40-60 0 0 0 0 2 40-60 0 0 0 0 2 40-60 0.826 0.333 0.864 0.826 0.824 0.703 20-40 0.667 0.174 0.6 0.667 0.632 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 == Confusion Matrix === a b c d f < classified as 0 0 0 0 b 100-120 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 <t< th=""><th>20:07:47 - rules.DecisionTable</th><th></th><th></th><th></th><th></th><th>.</th><th></th><th></th><th></th><th></th><th></th></t<>	20:07:47 - rules.DecisionTable					.					
ZAVERUS - functions simple logsic 0 0 0 0 0 0 2 100-120 0 0 0 0 0 0 2 40-60 0 0 0 0 0 2 60-80 0.826 0.333 0.864 0.826 0.844 0.703 20-40 0.667 0.174 0.6 0.667 0.632 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 === Confusion Matrix === a b c d e f < classified as 0	20:08:15 - bayes.BayesNet		TP Rate	FF Rate	Precision	Recall	r-Measure	ROC Area	Class		
a b c	20:09:05 - functions.SimpleLogistic		0	0	0	0	0	2	120-140		
0 0 0 0 0 2 60-80 0.826 0.333 0.864 0.826 0.844 0.703 20-40 0.667 0.174 0.6 0.667 0.632 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 === Confusion Matrix === a b c d e f <	20.10.00 * 0 665.340		0	0	0	0	0	2	40-60		
0.826 0.333 0.864 0.826 0.844 0.703 20-40 0.667 0.174 0.6 0.667 0.632 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 === Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 d = 60-80 0 0 0 0 0 19 4 e = 20-40 0 0 0 0 0 3 6 f = 0-20			0	0	0	0	0	?	60-80		
0.667 0.174 0.6 0.667 0.632 0.703 0-20 Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 === Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 0 3 6 f = 0-20			0.826	0.333	0.864	0.826	0.844	0.703	20-40		
Weighted Avg. 0.781 0.288 0.789 0.781 0.785 0.703 === Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 0 3 6 f = 0-20			0.667	0.174	0.6	0.667	0.632	0.703	0-20		
<pre>=== Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20</pre>		Weighted Avg.	0.781	0.288	0.789	0.781	0.785	0.703			
<pre>=== Confusion Matrix === a b c d e f < classified as 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20</pre>											
a b c d e f < classified as 0 0 0 0 0 0 0 a = 120-140 0 0 0 0 0 0 0 b = 100-120 0 0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20		=== Confusion I	Matrix ===								
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		abcda	a f /	. classifie	d ae						
$\begin{array}{cccccccccccccccccccccccccccccccccccc$		0000	0 0 I a	= 120-140	4 45						=
0 0 0 0 0 0 c = 40-60 0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20		0 0 0 0	0 0 b	= 100-120							-
0 0 0 0 0 0 d = 60-80 0 0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20		0 0 0 0	0 0 c	= 40-60							
0 0 0 19 4 e = 20-40 0 0 0 0 3 6 f = 0-20		0 0 0 0	0 0 d	= 60-80							
0 0 0 0 3 6 f = 0-20		0 0 0 0 1	94 e	= 20-40							
		0 0 0 0 3	36 f	= 0-20							
											v
											ł
Status	Status										
OK Log 🛷	OK										Log 🛷 XO

Figure E.21 J48 Results of Case Study 2B Cluster 3

	en fie	Open URL	Open Di	B Gene	rate	Undo	Edit Save
	tona.						
25	NIC						
trelato	n				Selected attribute		
nces: 1	20430 37	_deneme1-weia.v	nters-unsupervised attribute.Re	nov Attributes: 18 Sun of weights: 137	Name: Remaining-Open-D Missing: 0 (0%)	Uration Distinct: 5	Type: Nomnal Unique: 2 (1%)
в					No. Label	Count	Weight
-	A	line	Inert	Dattern	1 120-140	1	10
	~	INC	2001	revitin	2 100-120	0	0.0
	lines				3 40-60	7	7.0
	nene Innun				4 60-80	1	10
-	putilitiane Countred Data				5 20-40	-8	48.0
-	Closed Date				6 0-20	80	80.0
-	Tect-Tipe						
5	Product-Version						
6	Project-Phase						
	-						
7	Reproducibility						
7	Reproducibility 1-51.0C						
7 8 9	Reproducibility L-SLOC M/G						
7 8 9 10	Reproducibility 1-9.00 M/G Source-Component				Place Davide Ana Deck	or Alam)	- 14-
7 8 9 10 11	Reproducibility L-SLOC MVG Source-Component dmA1				Class: Remaining-Open-Durað	on (Nem)	• [Visu
7 8 9 10 11 12	Reproducbilty L-SLOC MVG Source-Component dmA1 dmA2 dmA2				Cass: Renaining-Open-Duraő	on (Nan)	• Visu
7 8 9 10 11 12 13	Reproducbilty L-5LOC MVG Source-Component dmA1 dmA2 dmA3 dmA4				Class: Remaining-Open-Duratio	an (lian)	• Visu
7 8 9 10 11 12 13 14 15	Reproducibility L-5LOC MVG Source-Component dmA1 dmA2 dmA3 dmA4 dmA5				Class: Remaining-Open-Durab	an (lian)	• Visu
7 8 9 10 11 12 13 14 15 16	Reprodubility L-5LOC MIG Source-Component drik1 drik2 drik3 drik4 drik5 drik6				Class: Remaining Open-Durab	an (lian)	• Visu
7 8 9 10 11 12 13 14 15 15 17	Reprodubility L-6.00 MrG Source-Component drik1 drik2 drik3 drik4 drik5 drik6 drik3				Class: Remaining Open-Durab	an (lian)	• Visu
7 8 9 10 11 12 13 14 15 15 15 17 18	Reprodubility L-5.00 MiG Sourz-Component dri A1 dri A2 dri A3 dri A4 dri A5 dri A6 dri R3 Remoning Com Duratio	Ω			Class: Remaining-Open-Durab	on (Van)	- Veu
7 8 9 10 11 12 13 14 15 15 17 28	Reprodubility 1-5.00 MG Sourz-Component dinA1 dinA2 dinA3 dinA4 dinA5 dinA5 dinA6 dinA5 dinA6 dinA3 Source Quantum Duration	n			Class: Remaining-Open-Durab	on (Nan)	• Visu a
7 8 9 10 11 12 13 14 15 15 15 15 15 15 15	Reprodubility 1-5.00 MiG Sourze-Component dinA1 dinA2 dinA3 dinA4 dinA5 dinA5 dinA6 dinR3 Straning-CompUnition	n			Class: Remaining-Open-Durab	on (Nam)	• Vsu q
7 8 9 10 11 12 13 14 15 15 17 13	Reprodubility L-5.00 MiG Sourze-Component dinA1 dinA2 dinA3 dinA4 dinA5 dinA6 dinR3 Semaining-OpenDucato	0			Class: Remaining-Open-Durab	an (lion)	• Visu a
7 8 9 10 11 12 13 14 15 15 15 17 23	Reprodubility L-5LOC MIG Source-Component dnA1 dnA2 dnA3 dnA4 dnA5 dnA6 dnA5 dnA6 dnA6	n			Class: Remaining-Open-Durab	an (lian)	+ Visu a
7 8 9 10 11 12 13 14 15 15 15 15	Reprodubility L-5LOC MIG Source-Component dnA1 dnA2 dnA3 dnA4 dnA5 dnA6 dnA5 dnA6 dnA3 Semanng-CompUrate	n			Class: Remaining-Open-Durab	an (lian)	+ Visu S
7 8 9 10 11 12 13 14 15 15 15 17 12	Reprodubility L-5LOC MIG Source-Component driA1 driA2 driA3 driA4 driA5 driA5 driA6 driA3 Remaining-Com-Durato	n			Class: Remaining Open-Durab	an (lian)	e Visu
7 9 10 11 12 13 14 15 15 15 17 18	Reprodubility L-5LOC MIG Source-Component driA1 driA2 driA3 driA4 driA5 driA6 driA3 Remaining Open Duratio	n			Class: Remaining Open-Ourab	an (lian)	e Ver

Figure E.22 Weka View of Case Study 2B Cluster 4

🥥 Weka Explorer	1.0		1.14	diam'r	ad and the	Ad an orth	24	and the second		- 0 X
Preprocess Classify Cluster Associate 9	Select attributes Visua	lize								
Classifier										
Choose DecisionTable -X 1 -S "wek	a.attributeSelection.Be	stFirst -D 1 -N	5"							
Test options	Classifier output									
🔘 Use training set	Correctly Clas	sified In:	stances	103		75.1825 %				*
O Supplied test set Set	Incorrectly Cl	assified :	Instances	34		24.8175 %				
Cross-validation Folds 10	Kappa statisti	c		0.52	97					
Dercentane split % 66	Mean absolute	error		0.14	22					
Percentage spin 70 00	Relative absolu	rea error ute error		77.89	12) 4 %					
More options	Root relative	squared en	ror	81.80	21 %					
	Coverage of ca	ses (0.95	level)	99.27)1 %					
(Nom) Remaining-Open-Duration 🔹	Mean rel. regi	on size (().95 level)	74.57	2 8					
Start Stop	Total Number o	f Instance	25	137						
Deput to (State State Score)	Detailed A	coursey B	, (1200							
Result list (right-clock for options)	bebailed A	couracy by	01000							
20,10,07 Hules, Decision able		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
		0	0.007	0	0	0	0.893	120-140		
		0	0	0	0	0	?	100-120		
		0	0	0	0	0	0.726	40-60		
		0 833	0 258	0 635	0 833	0 721	0.423	60-80 20-40		
		0.788	0.175	0.863	0.788	0.824	0.861	0-20		
	Weighted Avg.	0.752	0.193	0.726	0.752	0.733	0.843			
	=== Confusion !	Matrix ===								
	abcd	ef<-∙	- classifie	d as						
	0 0 0 0	01 a	= 120-140							Ξ
	0 0 0 0	0 0 b	= 100-120							
	0 0 0 0	61 c	= 40-60							
		01 d	= 60-80							
	0 0 0 0 1	763 f	= 0-20							
										-
	•									Þ
Status										
OK									l	.og 📈 x0
									_	

Figure E.23 DecisionTable Results of Case Study 2B Cluster 4

🗿 Weka Explorer	1.1.1		1.1	a family	-	-	24	and and		_ 0 X
Preprocess Classify Cluster Associate S	Select attributes Visua	alize								
Classifier										
Choose BayesNet -D -Q weka.classif	fiers.bayes.net.search	local.K2P	1 -S BAYES -E	weka.classifiers.b	ayes.net.es	timate.SimpleEsti	matorA 0.5			
Test options	Classifier output									
Cities training set										*
	Correctly Clas	sified Ins	stances	99		72.2628	1			
Suppled test set	Kanna statisti	.assiiied . .c	Instances	30 0.48	15	21.1312	5			
Cross-validation Folds 10	Mean absolute	error		0.10	13					
Percentage split % 66	Root mean squa	ared error		0.27	51					
More options	Relative absol	ute error		55.46	59 %					
	Root relative	squared en	rror	92.48)1 %					
(Nom) Remaining-Open-Duration	Coverage of ca	ises (0.95	level)	86.86	138					
	Mean rei, regi Total Number o	on size () of Instance	0.95 IEVEL) 20	24.09	59 8					
Start Stop	IUCAI NUMBEI C	I INSCANCE		137						
Result list (right-click for options)	=== Detailed A	Accuracy By	y Class ===							
20:13:59 - rules.DecisionTable										
20:14:46 - bayes.BayesNet		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		
		0	0	0	0	0	0.912	120-140		
		0.143	0.008	0.5	0.143	0.222	2 0.943	40-60		
		0	0	0	0	0	0.919	60-80		
		0.813	0.315	0.582	0.813	0.678	0.813	20-40		
		0.738	0.158	0.868	0.738	0.797	0.837	0-20		
	Weighted Avg.	0.723	0.203	0.736	0.723	0.715	0.835			_
	=== Confusion	Matrix ===	-							
	contabion	HUUIIA	-							
	a b c d	ef≺-	- classifie	ed as						
	0 0 0 0	01 a	= 120-140							
	0 0 0 0	0 0 b	= 100-120							Ξ
	0 0 1 0	60 c	= 40-60							
	0 0 0 0	10 d	= 60-80 = 20-40							
		2159 f	= 20-40							
										*
	•									Þ
Status										
OK									(Log 🛷 x O

Figure E.24 BayesNet Results of Case Study 2B Cluster 4

🥥 Weka Explorer			1.1	diam'r	-	-	24	and the second		
Preprocess Classify Cluster Associate S	Select attributes Visu	alize								
Classifier										
Choose SimpleLogistic -10 -M 500	-H 50 -W 0.0									
Test selferer	Classifier subsub									
lest options	Classiner output									
O use training set	Correctly Clas	sified In:	stances	97		70.8029 🛚				^
Supplied test set Set	Incorrectly Cl	assified :	Instances	40		29.1971 %				
Cross-validation Folds 10	Kappa statisti	.C		0.439	91					
Percentane split % 66	Mean absolute	error		0.124	12					
Tercentage spik 70 00	Relative absol	ute error		68.033	, , ,					
More options	Root relative	squared en	rror	92.76	17 8					
	Coverage of ca	ises (0.95	level)	91.970	8 8					
(Nom) Remaining-Open-Duration 🔹	Mean rel. regi	on size ().95 level)	32,603	34 %					
	Total Number o	f Instance	28	137						
Start Stop										
Result list (right-click for options)	=== Detailed A	Iccuracy B	/ Class ===							
20:13:59 - rules.DecisionTable		TD Data	FD Data	Drasisian	Decell	F. Managuna	DOC 1	C1		
20:14:46 - bayes.BayesNet		IF Kate	nr Kate	Precision	N N N	r-measure	NUC Area	120-140		
20:15:56 - functions, simpleLogistic		0	0	0	0	0	2	100-120		
		0	0.008	0	õ	0 0	0.851	40-60		
		0	0	0	0	0	0.647	60-80		
		0.708	0.281	0.576	0.708	0.636	0.729	20-40		
		0.788	0.246	0.818	0.788	0.803	0.776	0-20		
	Weighted Avg.	0.708	0.242	0.68	0.708	0.691	0.757			
	Confineiro	Manada								
	=== Confusion	Matrix ==:	•							
	abcd	ef <	- classifie	d as						
	0 0 0 0	1 0 a	= 120-140							
	0 0 0 0	0 0 b	= 100-120							
	0 0 0 0	7 0 c	= 40-60							_
	0 0 0 0	1 0 d	= 60-80							=
	0 0 0 0 3	414 e	= 20-40							
	0 0 1 0 1	.663 f	= 0-20							
										•
	•									•
Status									ſ	
OK									l	

Figure E.25 SimpleLogistic Results of Case Study 2B Cluster 4

🥥 Weka Explorer			1.1	i han i	-	AL INCOME	24	and so its		
Preprocess Classify Cluster Associate 5	Select attributes Visua	alize								
Classifier										
Choose J48 -C 0.25 -M 2										
Test options	Classifier output									
🔘 Use training set	Correctly Clas	sified In	stances	97		70.8029				*
Supplied test set Set	Incorrectly Cl	lassified	Instances	40		29.1971				
Cross-validation Folds 10	Kappa statisti	ic		0.41	36					
	Mean absolute error			0.14	13					
0 Percentage spin % 00	Root mean squa	ared error		0.26	98 97 s					
More options	Root relative	squared e	rror	90.36	24 8 81 8					
	Coverage of ca	ises (0.95	level)	95.62)4 %					
(Nom) Remaining-Open-Duration	Mean rel. regi	ion size (0.95 level)	40.38	93 %					
	Total Number o	of Instanc	es	137						
Start Stop										
Result list (right-click for options)	=== Detailed A	Accuracy B	y Class ===							
20:13:59 - rules.DecisionTable		TD D -+-	ED Data		P11	P. M	D00 1	61		
20:14:46 - bayes.BayesNet		IF Kate	IF Kate	Precision	Recall	r-Measure	RUC Area	120-140		
20:15:58 - functions.simpleLogistic		0	0	0	0	0	2	100-120		
		0	0	0	0	0	0.583	40-60		
		0	0	0	0	0	0.408	60-80		
		0.646	0.191	0.646	0.646	0.646	0.744	20-40		_
		0.825	0.404	0.742	0.825	0.781	0.746	0-20		
	Weighted Avg.	0.708	0.303	0.659	0.708	0.682	0.732			
	Confination	Manufa								
	=== CONTRATON	Matrix ==	-							
	abcd	ef<-	- classifie	ed as						
	0 0 0 0	10 a	= 120-140							
	0 0 0 0	0 0 b	= 100-120							
	0 0 0 0	2 5 c	= 40-60							
	0 0 0 0	0 1 d	= 60-80							
	0 0 0 0 3	81 17 e	= 20-40							
	00001	14 00 I	= 0-20							
	1									· · ·
										,
Status OK									(Log 🛷 x O

Figure E.26 J48 Results of Case Study 2B Cluster 4



TEZ FOTOKOPİ İZİN FORMU

_

		••
FNS	ГІТ	U
LING		<u> </u>

Fen Bilimleri Enstitüsü	
Sosyal Bilimler Enstitüsü	
Uygulamalı Matematik Enstitüsü	
Enformatik Enstitüsü	
Deniz Bilimleri Enstitüsü	

<u>YAZARIN</u>

	Soyadı : Adı : Bölümü :
	TEZİN ADI (İngilizce) :
	TEZIN TÜRÜ : Yüksek Lisans Doktora
1.	Tezimin tamamı dünya çapında erişime açılsın ve kaynak gösterilmek şartıyla tezimin bir kısmı veya tamamının fotokopisi alınsın.
2.	Tezimin tamamı yalnızca Orta Doğu Teknik Üniversitesi kullancılarının erişimine açılsın. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)
3.	Tezim bir (1) yıl süreyle erişime kapalı olsun. (Bu seçenekle tezinizin fotokopisi ya da elektronik kopyası Kütüphane aracılığı ile ODTÜ dışına dağıtılmayacaktır.)
Yaz	zarın imzası