

REPUBLIC OF TURKEY

AKDENİZ UNIVERSITY



**A FLEXIBLE COMPUTER BASED ASSESSMENT SYSTEM
IMPLEMENTATION USING JAVA WEB TECHNOLOGIES**

Alperen AKSOY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

JUNE 2019

ANTALYA

REPUBLIC OF TURKEY

AKDENİZ UNIVERSITY



**A FLEXIBLE COMPUTER BASED ASSESSMENT SYSTEM
IMPLEMENTATION USING JAVA WEB TECHNOLOGIES**

Alperen AKSOY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

JUNE 2019

ANTALYA

REPUBLIC OF TURKEY
AKDENİZ UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

**A FLEXIBLE COMPUTER BASED ASSESSMENT SYSTEM
IMPLEMENTATION USING JAVA WEB TECHNOLOGIES**

Alperen AKSOY

DEPARTMENT OF COMPUTER ENGINEERING

MASTER THESIS

This thesis unanimously accepted by the jury on 25/06/2019.

Prof. Dr. Melih GÜNEY (Supervisor)

Asst. Prof. Dr. Taner DANIŞMAN

Asst. Prof. Dr. Yılmaz Kemal YÜCE



ÖZET

JAVA WEB TEKNOLOJİLERİ KULLANILARAK ESNEK BİLGİSAYAR TABANLI DEĞERLENDİRME SİSTEMİ İMPLEMENTASYONU

Alperen AKSOY

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Melih GÜNAY

Haziran 2019; 61 sayfa

Bu tez kapsamında, Java Web teknolojileri kullanılarak web tabanlı esnek bilgisayar tabanlı değerlendirme sistemi geliştirilmiştir. Sistem tasarlanırken esnek olması ve mevcuttaki sistemlerin problemlerini giderecek şekilde olması sağlanmıştır. Geliştirilen sistem bir web sunucusuna konuşlandırıldıktan sonra otuz adet elektronik sınav Akdeniz Üniversitesi Mühendislik Fakültesi öğrencilerine uygulanmıştır.

Sisteme eğitmenler tarafından eklenen sorular ve oluşturulan testler, madde cevap kuramı ilkelerine göre analiz edilmektedir. Bu analizlerin sonuçları eğitime yansıtılarak, sorulmuş sorular hakkında geri dönüş sağlanmaktadır. Bu tez kapsamında referans seçilen programlamaya giriş dersinin soru bankası irdelenmiştir. Yapılan analizler sonucunda sınavlarda sorulan bazı soruların uygun olmadığı, bazılarının ise revize edilmesi gerektiği saptanmıştır.

Araştırmacılar bir çok çalışmada bilgisayar tabanlı değerlendirme sistemlerinin kullanımının öğrencilerin sınav başarılarına dair olumsuz etkisinin olup olmadığını araştırmışlardır. Çoktan seçmeli soruların sorulduğu sınavların elektronik ortamda yapılmasının öğrencilerin sınav başarısına bir etkisinin olmadığı ortaya konya da açık uçlu sorular için yapılmış bir çalışma bulunmamaktadır. Bu tez kapsamında Doğa Bilimleri (Natural Science) dersi bu karşılaştırmayı yapmak için referans ders olarak seçilmiştir. 2017-2018 güz döneminde dersi sınavları kağıt ortamında açık uçlu sorular üzerinden sınavlar yapılmıştır. 2018-2019 gün döneminde ise yine açık uçlu sorular ile elektronik ortamda sınavlar yapıldı. Ortalama sınav puanları üzerinden yapılan bağımsız student t-test analizine göre iki grup arasında anlamlı bir fark bulunmadığı anlaşılmıştır ($p= 0,441$). Elde edilen bulgular ile elektronik sınav sistemi kullanımının açık uçlu sorulardan oluşan sınavlarda da öğrencilerin başarısına dair olumsuz bir etkisinin olmadığı kanıtlanmıştır.

Elektronik sınavların yapıldığı ortamlar genellikle okulların bilgisayar laboratuvarlarıdır. Bu ortamlar ders verilmek için tasarlandığından dolayı, sınav esnasında öğrencilerin birbirlerinin ekranlarına bakmaları çok zor olmamaktadır. Bilgisayarların yerleşim planında değişiklik yapmak her zaman mümkün olmasa da sınavlardan sonra öğrenci cevaplarının birbirine olan benzerlikleri araştırılarak kopya çeken öğrencileri tespit etmek mümkündür. Geliştirilen sistem kullanılarak yapılan sınavlarda, Jaccard ve Kosinüs benzerlik analizi yöntemleri kopya çeken öğrencilerin tespitinde kullanılmıştır. Bu iki yöntem elde ettiği sonuçlar bakımından kıyaslanmıştır. Kosinüs benzerlik analizi yönteminin kullanımının, kopya tespiti konusunda Jaccard benzerlik yöntemine göre daha uygun olduğu yapılan çalışmalar sonucunda görülmüştür.

Tüm bu yapılan çalışmaların sonunda, açık kaynak kodlu teknolojiler kullanılarak eğitim kurumlarının kullanımına uygun olarak tasarlanan ve tüm ölçme ve değerlendirme süreçlerini kapsayan, web ortamında çalışan bilgisayar tabanlı değerlendirme sistemi geliştirilmiştir.

ANAHTAR KELİMELER: Bilgisayar Tabanlı Değerlendirme, E-sınav, Madde Cevap Kuramı, Java Programlama Dili, Spring Çatısı, Web Teknolojileri

JÜRİ: Prof. Dr. Melih GÜNAY

Dr. Öğr. Üyesi Taner DANIŞMAN

Dr. Öğr. Üyesi Yılmaz Kemal YÜCE

ABSTRACT

A FLEXIBLE COMPUTER BASED ASSESSMENT SYSTEM IMPLEMENTATION USING JAVA WEB TECHNOLOGIES

Alperen AKSOY

MSc Thesis in Computer Engineering

Supervisor: Prof. Dr. Melih GÜNAY

June 2019; 61 pages

In the scope of this thesis, a web based flexible computer based assessment system was developed using Java web technologies. The system is designed to be flexible in order to eliminate the problems of existing systems. After the deployment on a web server of the developed system, thirty electronic exams were applied to the Akdeniz University Engineering Faculty students.

The questions and tests added by the instructors to the system are analyzed according to the principles of item-response theory. The results of these analyses are presented the instructor and feedback are given about the questions. Within the scope of this thesis, introduction to programming course was selected as a reference course and its question bank was examined. According to the analysis results, it was determined that some questions were inappropriate for asking and some questions need to be revised.

Researchers have investigated in many studies, whether the use of computer based assessment system has a negative impact on the student's exam success. Although it is proved that the exams in which multiple-choice questions are asked in the electronic environment have no effect on the student's exam success, there are no studies for open-ended questions. In the scope of this thesis, Natural Science course was selected as a reference course. In 2017-2018 fall semester, exams of this course were held in the paper environment with the open-ended questions. In 2018-2019 fall semester, exams were held in the electronic environment with the open-ended questions. According to the independent student t-test analysis performed on the average exam scores, there was no significant difference between the two groups ($p=0,441$). The findings of the study revealed that the use of the electronic exam system did not have a negative effect on student's success in the exams which consisted of open-ended questions.

Electronic exams are usually held in computer laboratories of schools. These laboratories are designed to teach courses. Because of this, it is not difficult for students to look at each other's screens. Although it is not always possible to make changes in the settlement of the computers, it is possible to detect the students who are copying by examining the similarities of the student's answers after the exams. In the exams conducted using the developed system, the methods of Jaccard and Cosine similarity analysis were used to detect the students who were copying. These two methods were compared in terms of the results. It has been seen that the use of Cosine similarity analysis method is more suitable than the Jaccard similarity analysis method for copy detection.

As a result of all these studies, by using open source technologies, a computer based assessment system designed in accordance with the use of educational institutions and covering all measurement and evaluation processes is developed.

KEYWORDS: Computer Based Assessment, E-exam, Item Response Theory, Java Programming Language, Spring Framework, Web Technologies

COMMITTEE: Prof. Dr. Melih GÜNAY

Asst. Prof. Dr. Taner DANIŞMAN

Asst. Prof. Dr. Yılmaz Kemal YÜCE

ACKNOWLEDGEMENTS

Firstly, I would like to thank Prof. Dr. Melih GÜNAY my thesis supervisor. He always helps me any problem with my thesis and other academic issues during my master education. Secondly, I would like to thank the staff of Akdeniz University Computer Science Research and Application Center and Akdeniz University IT department for their help during the development process of the developed system.

I would also like to thank Akdeniz University Department of Computer Engineering's staff and students for using the electronic examination system which has been developed in the scope of this thesis.

Finally, I would like to thank my family since they have supported me all of my life.



LIST OF CONTENTS

ÖZET	i
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
TEXT OF OATH	viii
LIST OF ABBREVIATIONS	ix
LIST OF FIGURES	x
LIST OF TABLES	xi
1. INTRODUCTION	1
2. LITERATURE REVIEW	4
2.1. Computer Assisted Assessment	4
2.2. Computer Based Assessment	4
2.3. Evaluation of Quality of Questions	7
2.3.1. Classical test theory	7
2.3.2. Item response theory	8
2.3.3. Application of item response theory	10
2.4. Similarity Analysis Techniques	12
2.4.1. Jaccard similarity analysis technique	13
2.4.2. Cosine similarity analysis technique	13
2.5. Adaptive Testing	14
3. MATERIAL AND METHOD	15
3.1. Design Considerations	15
3.2. Implementation Phase of CBA	16
3.2.1. Java programming language	16
3.2.2. Spring framework	17
3.2.3. Model view controller pattern	17
3.2.4. Mysql database	18
3.2.5. Centos operating system	19
3.2.6. MyBatis object relational mapping framework	19
3.2.7. Bootstrap framework	20
3.2.8. Apache maven	21
3.2.9. Git version management system	21
3.3. CBAEntities Project	22
3.4. CBAInstructor Project	23
3.5. CBASStudent Project	26

3.6. Deployment Environment	30
4. RESULTS AND DISCUSSION	32
4.1. Survey	35
4.2. Question Quality	36
4.3. Resource Consumption	36
4.4. Comparision with the PPA	37
4.5. Cheat Detection	38
5. CONCLUSION.	42
6. REFERENCES	43
CURRICULUM VITAE	

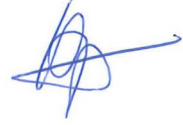


TEXT OF OATH

I declare that this study "A FLEXIBLE COMPUTER BASED ASSESSMENT SYSTEM IMPLEMENTATION USING JAVA WEB TECHNOLOGIES", which I present as master thesis, is in accordance with the academic rules and ethical conduct. I also declare that I cited and referenced all material and results that are not original to this work.

25/06/2019

Alperen AKSOY



LIST OF ABBREVIATIONS

Abbreviations:

ANOVA	: Analysis of Variance
CAA	: Computer Assisted Assessment
CBA	: Computer Based Assessment
CENTOS	: Community Enterprise Operating System
CTT	: Classical Test Theory
GRE	: Graduate Record Examination
GMAT	: Graduate Management Admission Test
IRT	: Item Response Theory
NASA-TLX	: National Aeronautics and Space Administration Task Load Index
ORM	: Object Relational Mapping
OSYM	: Assesment Selection and Placement Center
PPA	: Paper and Pencil Based Assessment
TOEFL	: Test of English as a Foreign Language

LIST OF FIGURES

Figure 2.1.	Coverage of terms for computer usage in assessment	5
Figure 2.2.	Lower and upper scoring group	11
Figure 2.3.	Sample item characteristic curve	12
Figure 2.4.	An example for transforming string to vector	13
Figure 3.1.	The structure of MVC pattern	18
Figure 3.2.	Screenshot from CWP user interface	19
Figure 3.3.	MyBatis mapper sample usage with XML	20
Figure 3.4.	MyBatis mapper usage via annotations	20
Figure 3.5.	Bootstrap grid system	20
Figure 3.6.	Example usage of Bootstrap toolkit	21
Figure 3.7.	Structure of CBAEntities project	22
Figure 3.8.	Structure of CBAInstructor project	23
Figure 3.9.	Screenshot of CBAInstructor project main screen	24
Figure 3.10.	Data transfer in controller to view part	24
Figure 3.11.	Automatic test question creation process	25
Figure 3.12.	One by one question selection process	26
Figure 3.13.	View of test format	27
Figure 3.14.	Exam creation page	27
Figure 3.15.	Assess free answer question page	28
Figure 3.16.	Exam result page	28
Figure 3.17.	Question analysis page	29
Figure 3.18.	Multiple-choice question screen in exam	30
Figure 3.19.	Open-ended question screen in exam	31
Figure 3.20.	Question with image in exam screen	31
Figure 4.1.	Schematic representation of the question bank	32
Figure 4.2.	An example of a shuffling operation by directly used on questions	33
Figure 4.3.	An example of shuffling operation used on question groups	34

LIST OF TABLES

Table 2.1.	Meanings of discrimination parameter	11
Table 4.1.	Questionnaire results from examinees	35
Table 4.2.	Distribution of question discrimination parameter for 117 questions	36
Table 4.3.	Resource consumption of developed CBA system	36
Table 4.4.	Information about the student groups	37
Table 4.5.	T-Test results for total final points	38
Table 4.6.	Similarity rates for a few copying examples detected by the system	39
Table 4.7.	Comparision of features of TCExam and developed CBA system	40



1. INTRODUCTION

Due to the rapid development of computer technology, the use of computers is rapidly increasing in all areas, even the area that previously made little use of computers. In education especially, this use is very beneficial. The use of computers in education is generally focused on distance education systems but also computerized testing systems are taking attention. Currently, TOEFL, GMAT and GRE exams are the best examples of computerized testing systems abroad, while in Turkey, OSYM created electronic examination center in Ankara, Istanbul and Izmir cities. OSYM has been started electronic exams in 2013 with YDS exam.

Assessment is an important concept for the learning process. Reasons for assessing students include helping students with their learning process and monitoring the effectiveness of current teaching and learning strategies (Chalmers and McCausland 2002).

Currently, six types of assessment are known from the literature. (i) Diagnostic: Assess student's performance before the instruction process. This type of process ensures necessary information about students' ability level. (ii) Formative: Assess student performance during instruction process. Generally taken by students' via quizzes, midterm exams etc. (iii) Summative: Assess student performance after the instruction process. Students are generally taken by via final exams etc. (iv) Norm-referenced: Comparison of performance of students with other students. It doesn't matter in this type of assessment have taken by students before or after the instruction process. Generally implemented via national or internal tests. In Turkey, the biggest example of this type of assessment is YGS exam taken from OSYM. (v) Criterion – Referenced: Measure students skills against a specific objective. For example, foreign language proficiency exams etc. Generally, students take these exams before the instruction process for deciding they should take the course or not. (vi) Benchmark: Measure students' performance at interval periods. For example, weekly quizzes etc. By giving these exams, instructors can estimate students' ability level of the final exam (Anonymous 1 2018).

Advantages of using computerized tests are (I) saving paper and printing costs: In addition to saving money by eliminating the need to print, it also prevents tree cutting. (II) Saving time: After the creation of question bank, instructors need only select exam questions by clicking. Completed exam assessment requires only a few seconds for each

exam, and strong question and test analyses are done automatically by the system.

The main purpose of this thesis is to measure and evaluate students via computers. For faster and more effective measurement and evaluation with less cost in a short time, we implemented the Computer Based Assessment system with the following objectives, using JAVA web technologies:

- More than one lecturers are able to work on the same question bank.
- Questions are kept safely in electronic environment and flexibly categorized for re-use with multi-level hierarchy.
- Necessary student and staff information are easily integrable with the other software systems (student information system, staff information system). Hence instructors can focus exclusively on the instruction thereby reducing workload and avoiding information inaccuracies.
- Multiple-choice questions and open-ended questions can be in the same exam booklet together.
- Copy detection mechanism between examinee is established.
- Security issues are handled.
- The system is optimized for multi-user usage and ensure that fast response and fluently usage.
- The system is designed as an outcome focused model to ensure student and teacher performance improvements. With this model, students and instructors can see their topic based success.
- The system is designed to prevent data loss in some undesired situations (power cut etc.).
- The system is provided feedback to parties to improve achievement for individualized education via reported topic based success rates for both students and instructors.
- Question and test analyzes are done with the fundamentals of IRT. With these analyzes, instructors can get feedback about the asked questions and quality of the question is determined by the system automatically.

In the scope of this thesis, computer based assessment system was developed using JAVA web technologies. Thirty electronic exams were held in Akdeniz University Engineering Faculty by using developed CBA system. Questions and tests were analyzed by using item response theory. To detect copy between examinees, two similarity analysis techniques were implemented and compared. To show there is no significant difference between paper based and computer based exams on student's success, independent student t-test was applied.



2. LITERATURE REVIEW

2.1. Computer Assisted Assessment

Computer Assisted Assessment is a term that refers to the use of computers to assess student's performance during the learning process. Bull and McKenna (2004) defined CAA as the use of computers for assessing student learning (Conole and Warburton 2005). CAA is the most comprehensive term for the usage of computers in the assessment. For example, after the multiple choice exams, optical mark recognition devices can be used for assessment. This is a typical example for CAA (Chalmers and McCausland 2002).

Preprinted paper tests are taken by the student using pen or pencil and fed into OMR at rates of up to 2200 forms per hour. The most common test format for CAA multiple choice questions or true, false question whose correct answer is predetermined. Bull and McKenna ordered reasons for using CAA are shown below (Bull and McKenna 2004).

- To increase the frequency of assessment.
- To broaden the range of knowledge assessed.
- To increase feedback to students and lecturers.
- To extend the range of assessment methods.
- To increase objectivity and consistency.
- To decrease marking loads.
- To aid administrative efficiency

2.2. Computer Based Assessment

CBA is one of the common term for assessment but quite confused term with CAA. CBA is practice of giving quizzes and test on computer environment, instead of using pencil and paper. In this concept, all assessment processes are carried out by computers. So CAA covers CBA. Coverage of terms in assessment about computers is shown in Figure 2.1 .

CBA is cover stand-alone assessment and network based assessment. In stand-alone assessment, during the examination, computer doesn't need to connect with others. For

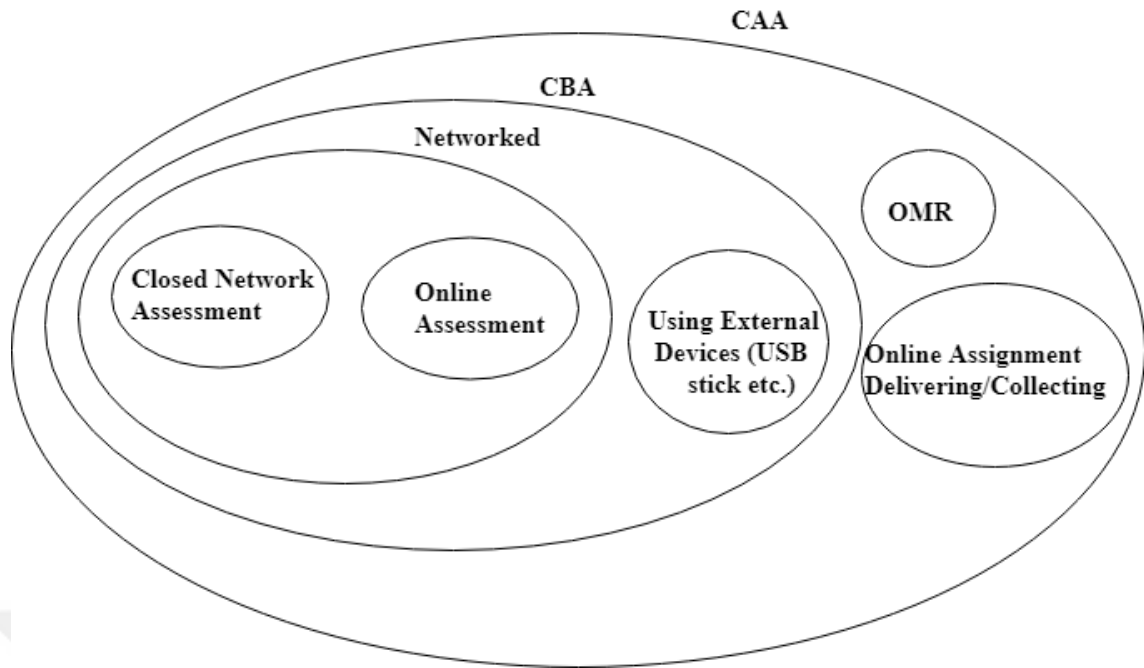


Figure 2.1. Coverage of terms for computer usage in assessment

example, students take the test in their home by using their own computer to test yourself before summative exams or students take the test in school by using the computer without network. After the exam instructor collects the answers from computers' local storage.

Nowadays, many exams have been applied in electronic environment. Common examples are TOEFL exam, GRE exam and GMAT exam, which have been applied in electronic environment for more than 10 years.

One of the greatest concerns about the usage of CBA is that proficiency with computers may affect the exam results. Clariana and Wallace (2002) handle this topic, paper-based versus computer-based assessment: key factors associated with the test mode effect. This investigation seeks to confirm several key factors that influence performance, such as content familiarity, computer familiarity, competitiveness and gender. This study used a post-test design with one factor, which was the test mode (computer-based and paper-based). Dependent variables included student's scores on 100-item multiple choice test and student's self-report on a distance learning survey. This study divides the students into two parts for testing, one of them for CBA and the others for PPA. For showing that the two groups' ability levels don't have any significant difference, they use ANOVA analysis. For results, student achievement comparison used ANOVA statistic method. Ac-

According to their results, content attainment is the most important factor when content attainment rate is high students can take more than content attainment is less students but it is possible that student who content familiarity is less than other student and select computer based assessment method he or she can take better score. According to results, computer familiarity is an insignificant factor for computer based assessment.

Noyes et al. (2004) based their study on the study by Clariana R. and Wallace P. in 2002. Clariana and Wallace sought to confirm the key factors which influence paper-based and CBA on the academic performance of university undergraduates. Noyes et al studied about the workload is an another test mode effect? They used NASA-TLX method for calculate the workload. They selected 30 undergraduate students, half of which were men with the others being women. They assigned students randomly to either paper based or CBA groups. The comprehension task comprised reading an article and answering 10 multiple-choice question about the article. The results showed that there were no significant difference in the comprehension or the overall workload scores for two groups, but significant difference in the perceived effort sub-scale across the two conditions. According to this finding, computer-based test required more effort than the paper-based comprehension test. They suggested workload should be considered when administering computer based assessment.

The use of CBA is advantageous in many ways. One of them is that students access their exam results immediately. Daniels and Gierls (2017) handle effects of students accessing their score immediately by using CBA system. According to the results, immediate test score reporting had no effect on student's anger. Furthermore, score reporting immediately reduces feelings of anxiety and shame. According to the survey results of this study, 98% of students prefer to learn their scores immediately.

Bloom et al. (2018) handle their institution's experience of using CBA system. For comparing the effects of student's performance on paper based versus CBA, they took 2 years exam results using PPA system and took 2 years exam results by using CBA system. According to results, no consistent performance difference between the paper and computer-based formats. They also collect the students feedback after electronic exams and categorized these feedbacks. 46% of these feedbacks related to the improving CBA system. 38% of these feedbacks are related to the problems during the exam. The most common problem is the students' computer during the exam. It shows that examination

environment is one of the important key factor of success of CBA system.

Akdemir and Oguz's (2008) study compared paper based exam and CBA. In this study, 24 undergraduate students were divided into two groups. One of them took Turkish Language and Literature course's exam with CBA system. Other group took Pencil and Paper based version of the same exam. The exam had only thirty multiple-choice questions. After the exam, results were analyzed using statistical methods. According to the results, student's scores were not statistically different between paper based and computer based groups. In this study, all of the students already took basic computer course, so all of the students have at least basic computer usage skills. Therefore computer usage familiarity factor wasn't considered as a factor in this study.

Pawauskas et al's (2014) study is handle their adopting CBA process and also they compare cost of CBA and PPA. Their one year PPA cost was approximately \$23.500. Their course has approximately 130 students. The university uses Sakai software for its learning system. This allows faculty to post their course material, administer unsecured quizzes, tests, or homework. Previously, the primary method of administering examinations, quizzes and homework assignments was predominantly paper based. This method requires either manual grading or utilization of an optical mark recognition scanner. Exam module of Sakai software doesn't have any intelligent shuffling mechanism and require internet connection for examination. Because of these disadvantages, they declared that they need new intelligent CBA system.

2.3. Evaluation of Quality of Questions

2.3.1. Classical test theory

When the measurement history is investigated, two basic theories are found. First of them is classical test theory. Parameters of this theory are, item difficulty level, item discrimination parameter, reliability index. Difficulty parameter is the number of candidates who answered the question correctly is the ratio of the total number of candidates. This ratio is getting value between 0 and 1. Item discrimination parameter is looked question what level distinguish for candidates. From the information given, the parameter of CTT are depends on the group which taking test. The measurement tools developed by the CTT are generally aimed at individuals with a medium ability level. In this theory, valid and

safe test for a group may not be valid and safe for another group(Lord 1980, Embretson 1999).

2.3.2. Item response theory

Item response theory was introduced in the middle of the 20th century in order to eliminate the deficiencies of CTT. According to this theory, there is a relationship between the direct observations of individuals in a given field and their responses to the questioning field and this relationship can be expressed mathematically. IRT is based on following hypothesis. Unidimensionality, local independence, invariance and item characteristic curve (Crocker and Algina 1986, Wainer and Thiisen 2001, Celen 2008).

Item response theory models are divided into two basic models. One-dimensional models are used for measure single implicit property, multidimensional models are used for measure more than one implicit property. In these models, one-dimensional logistic models which are the most widely used for adaptive testing applications. Logistic models are divided into three. These are, one parameter, two parameter and three parameter logistic models.

In one parameter logistic model, probability of given correct response of candidate is shown in Equation 2.1. In this equation θ denotes ability level and b_i parameter denotes difficulty parameter of i th question. This model also called Rasch model in literature (Reid et al. 2007).

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1 + e^{\theta-b_i}} \quad i = 1, 2, 3, 4, \dots, n \quad (2.1)$$

In this model, gradients of item characteristic curves are equal because of the only difficulty levels are different from each other.

In two parameter logistic model, item characteristic curve has two parameter. Difference between one parameter logistic model and this model while calculating the probability of given correct response, question discrimination parameter is also considered in two parameter logistic model. Chance factor also taken 0 in this model. According to this model, the probability of a person at θ level to answer the question i correctly given in Equation 2.2. In this equation a_i parameter denotes the question discrimination (Birnbaum 1968).

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad i = 1, 2, 3, 4, \dots, n \quad (2.2)$$

In addition to one parameter model, this model, which also uses the discrimination parameter of the question, is more suitable if the questions don't indicated equally implicit properties (Embretson and Reise 2000).

Three parameter logistic model obtained by adding chance factor to two parameter logistic model. Chance factor is probability of correct response of low ability level person by without any information about question. According to this model, the probability of a person at level to answer the question i correctly given in Equation 2.3. In this equation, c_i denotes the chance factor (Crocker and Algina 1986, Embretson and Reise 2000).

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad i = 1, 2, 3, 4, \dots, n \quad (2.3)$$

Before asking question in exam, response can be estimated by using models shown above. If exams implemented compatible with IRT, by using 50% - 90% less question can estimate individual's ability level more precisely (Weiss 1985).

Another important concept in IRT is item information function is given in Equation 2.4.

$$I_i(\theta) = \frac{P'_i(\theta)^2}{P'_i(\theta)(1 - P_i(\theta))} \quad (2.4)$$

In this function, I_i denotes distribution of i^{th} question for person at θ ability level. The graph of this function is generally bell-shaped. $P'(\theta)$ function is first derivative function of $P(\theta)$. If this formula is interpreted, while discrimination parameter increase also value of the I function is increased. The value of I function also increases as the difficulty parameter (b) moves closer to the θ parameter. While the chance factor parameter is decreasing, the value of the I function is getting higher. The chance factor parameter (c) is inversely proportional to the value of the I function.

The sum of the item information function in test for all items gives the test information function. The formula of the test information function given in Equation 2.5 (Crocker and Algina 1986).

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \quad i = 1, 2, 3, \dots, n \quad (2.5)$$

The standard error for ability estimation calculates as formula given in Equation 2.6.

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (2.6)$$

According to Equation 2.6, while increasing value of test information function, standard error is decreasing. Standard error parameter is generally used in adaptive test concept for termination criteria.

The formula used for ability estimation is given in Equation 2.7. In this formula, θ_n denotes current ability after the answered n questions. θ_{n+1} denotes new estimated ability level.

$$\theta_{n+1} = \theta_n + \frac{\sum_{i=1}^n S_i(\theta_n)}{\sum_{i=1}^n I_i(\theta_n)} \quad (2.7)$$

By using formula given in Equation 2.7, new estimation parameter can be calculated. In this equation calculation of S_i parameter is given in Equation 2.8.

$$S_i(\theta) = (\mu_i - P_i(\theta)) \frac{P'_i(\theta)}{P'_i(\theta)(1 - P_i(\theta))} \quad (2.8)$$

2.3.3. Application of item response theory

By using fundamental principles of IRT, question and test analysis are done automatically. First step of these analysis, calculation of the item parameters such as question difficulty (b), question discrimination (a) and chance factor (c) parameters.

Question difficulty parameter is calculated as the ratio of those who answered the question correctly to all respondents. This parameter is getting value between 0 and 1. Calculation formula is given in Equation 2.9.

$$b = \frac{\text{correctly answered respondents count}}{\text{all respondents count}} \quad (2.9)$$

According to Equation 2.9, as the question becomes easier, question difficulty parameter approaches 1. Vice versa, the question difficulty parameter approaches 0.

The calculation of question discrimination parameter formula is given in Equation 2.10.

$$a = p_{upper} - p_{lower} \quad (2.10)$$

p_{upper} denotes proportion of examinees responded correctly to the question in upper scoring group and p_{lower} denotes proportion of examinees responded correctly to the same question in lower scoring group. To determine lower and upper scoring group, examinee list is ordered descending by their scores. First 27% examinees are grouped as an upper scoring group, last 27% examinees are grouped as a lower scoring group. Other examinees are ignored (Kelley 1939).

Exam Score List Ordered By Score Descendent



Figure 2.2. Lower and upper scoring group

To interpret discrimination factor, a guideline was developed by Ebel. According to this guideline meanings of a values are shown in Table 2.1 (Ebel 1965).

Table 2.1. Meanings of discrimination parameter

Discrimination Parameter	Meaning
>0.40	Item is performing satisfactorily, no revision required
>0.30 and <0.39	Little or no revision required
>0.20 and <0.29	Need revision
<0.19	This question shouldn't be used.

Chance factor (c) parameter denotes, probability of given correct response by examinee without any knowledge about the question. This parameter was taken 0.3 for calculation.

In education, examinees ability level named as θ . Theoretically, this parameter getting value between $-\infty$ and $+\infty$. In this work, range of this value set -3 and 3 . This value is directly proportional to the examinee's ability level. The regression of item point on the θ parameter named "Item Characteristic Function". The curve of this function is named "Item Characteristic Curve" (Embretson and Reise 2000). Calculation of item characteristic function is given in Equation 2.3. Sample item characteristic curve is shown in Figure 2.3.

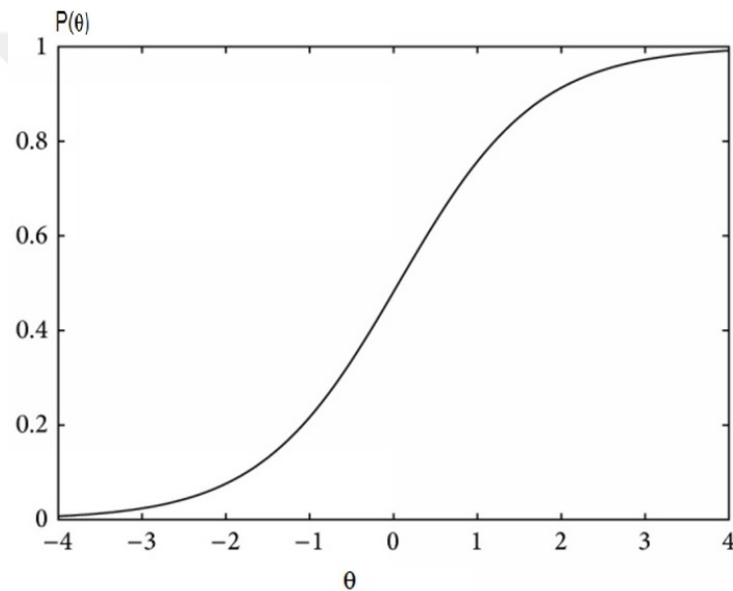


Figure 2.3. Sample item characteristic curve

2.4. Similarity Analysis Techniques

Similarity analysis techniques are used to find similarity rate between two documents. In the scope of this thesis, two similarity analysis techniques was investigated. These are Jaccard similarity analysis technique and Cosine similarity analysis technique.

2.4.1. Jaccard similarity analysis technique

Jaccard similarity analysis is technique for calculate similarity rate for two document based on common words. Calculation of two document's similarity rate by using Jaccard Similarity algorithm is given in Equation 2.11.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2.11)$$

By using formula given in Equation 2.11, similarity rate can be getting value between 0 and 1.

2.4.2. Cosine similarity analysis technique

Another similarity analysis technique which considered the scope of this work is Cosine similarity analysis technique. In this technique documents are represented as a vector. Cosine of angle of between two vectors is denote similarity rate. First step of this algorithm is transform answer string to the vector. In this transformation, strings are transformed to integer arrays. The size of each array is equal to total distinct word count of all documents. Each index is matched with one word. The content of array is including count for matched word. One example is shown in Figure 2.4.

s1 = ali can play football
s2 = ahmet can play basketball

String	ali	can	play	football	basketball	ahmet
s1	1	1	1	1	0	0
s2	0	1	1	0	1	1

Vector of s1 = [1,1,1,1,0,0]
Vector of s2 = [0,1,1,0,1,1]

Figure 2.4. An example for transforming string to vector

Second step of cosine similarity algorithm is calculating cosine value between vectors. The formula of the calculation is shown in Equation 2.12.

$$\cos(s1, s2) = \frac{\vec{s1} \cdot \vec{s2}}{\|\vec{s1}\| \|\vec{s2}\|} = \frac{\sum_1^n s1_i s2_i}{\sqrt{\sum_1^n s1_i^2} \sqrt{\sum_1^n s2_i^2}} \quad (2.12)$$

The cosine value of two vector is cosine similarity metric. This value can get value between -1 and 1. If this value approaches 1, similarity increases but if it approaches -1, similarity is decreases.

2.5. Adaptive Testing

Adaptive tests, an application of item response theory is form of computer based tests that adapts to examinees' ability level. In adaptive tests, questions are determined by examinees' past answers. Until examinees' ability level recognized, CBA system ask question automatically. If examinee answered previous question correctly, next question is selected harder by the CBA system. The vice versa, next question is selected easier. (Anonymous 2 2019)

First application of adaptive tests was carried out by Alfred Binet and Theodore Simon in 1916 named as a Stanford-Binet Intelligence test. This test is begun with estimation of starting ability level before the adaptive concept. Questions in this test are divided into difficulty levels as age groups. If examinee answered correctly all questions in beginner level, all questions in below levels are assumed answered correctly. Tests continue until examinee answered all questions incorrectly in a certain ability level. At this level determined as a ceiling level of examinee and test is terminated (Bain and Allin 2005).

3. MATERIAL AND METHOD

In the scope of this thesis, a web based flexible computer based assessment system was developed by using open source technologies. A web platform was selected as the application platform. Because, only a web browser is needed in order to run web applications. The first step of this work is a determine the design considerations.

3.1. Design Considerations

Identified design considerations of the CBA system is listed below.

- More than one lecturers should able to work on the same question bank.
- Questions should be kept safely in electronic environment and flexibly categorized for re-use with multi-level hierarchy.
- Able to handle shuffling for consecutive group questions (paragraph question).
- Support for simultaneously testing with time synchronization.
- System should provide feedback to parties to improve achievement for individualized education via reported topic based success rates for both students and instructors.
- During the exam, network and/or electricity may be cut off. The system should prevent data loss.
- System should support both open-ended and multiple-choice questions in the same booklet.
- Questions may be edited and asked at different times. Therefore, all versions of the questions should be kept in the database.
- Questions may include multi-media content image, audio, and video.
- During testing, it is desired that the test taker see the progress of the exam, remaining time, percent answered, bookmarked questions and more.
- System should be support importing questions/test takers at various formats from external sources. Further, system should also allow exporting of answers and other statistics for further analysis.

- Online and off-line copy detection mechanism should be included in the system through anomaly detection data mining techniques.
- System should adaptively decide on questions based on test taker performance during the examination.
- System should allow flexible test booklet generation from the question bank constraint by topic/module and type. These booklets may be configured to be per user or per exam. In case custom question booklet is generated per user then the system should ensure fairness of the booklet.

3.2. Implementation Phase of CBA

Implementation of this computer based assessment system, web platform was selected. In the web platform, Model, View and Controller (MVC) pattern was selected. In this pattern, controller receives all requests and get information from model and send to this information to the view part. For the software development process, incremental software development model was selected. In incremental development model, the whole requirement is divided into small parts. For each part, the software life cycle is used. Each cycle includes the following processes: (i) requirement analysis, (ii) design, (iii) implementation, (iv) test, (v) deployment, and (vi) maintenance. (Larman and Basili 2003)

The system is divided into the following three sub projects:

- CBAEntities
- CBAInstructor
- CBAStudent

Each project was implemented using the JAVA programming language. Spring MVC framework was selected for the Web front-end GUI technology.

3.2.1. Java programming language

Java is general-purpose, open source, object oriented programming language. James Gosling, Mike Sheridan and Patrick Naughton started Java Language Project in June 1991. Java 1.0 released by Sun Microsystems in 1996. Java is cross-platform language. It means same source code can work more than one platform (Windows OS, Linux OS,

Android OS, MacOSx). It came up with “Write Once, Run Anywhere” slogan (Anonymous 3 2019).

Java was originally designed for interactive television, but it was too advanced for digital cable television industry at the time. This language initially named Oak. Later the project the name was changed as a Green and finally named JAVA.

In most programming languages, compiler generates code can execute specific target machine. But java compiler doesn't create an object file, instead of object file it creates a bytecode file which can run on a JVM (Java Virtual Machine). This bytecode can be used any platform which is java installed. Consequently, owing to this feature java is called cross-platform language (Carroll 2019).

3.2.2. Spring framework

Spring framework is an open source java framework developed based on principles of dependency injection and inversion of control released in 2003 under the Apache license. The advantages of spring framework is listed below (Chand 2019).

- **Lightweight:** It doesn't consume a lot of resource
- **Aspect Oriented Programming:** By separating application business logic from system services, Spring Framework supports Aspect Oriented Programming.
- **Container:** Life cycle of objects managed by Spring Framework.
- **MVC Framework:** Spring supports MVC pattern and isolate business logic and interface.
- **Transaction Management:** Spring Framework provides a generic abstraction layer.
- **JDBC Exception Handling:** Spring has JDBC abstraction layer and simplifies the error handling strategy.

3.2.3. Model view controller pattern

MVC is a popular software pattern which separate data, business logic and interface. The data layer contains raw data named Model. In controller layer, data can be read from model layer and handled. In presentation layer named view, processed data can presented to user via interfaces.

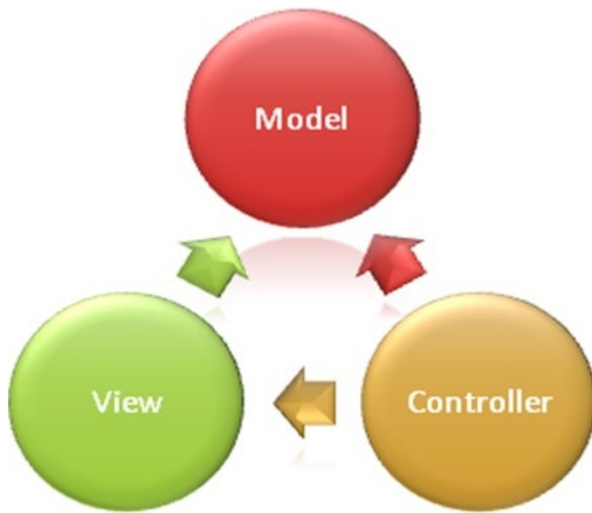


Figure 3.1. The structure of MVC pattern

The structure of MVC pattern is given in Figure 3.1. Advantages of MVC pattern are listed below (Anonymous 4 2019).

- **Faster development:** It supports parallel development. It means while one programmer work on the view part, other can work on the controller part.
- **Modifications doesn't affect the entire model:** MVC pattern isolate layers from each other. Therefore any changes in the model part doesn't take the controller and view part.
- **Asynchronous Technique Support:** MVC support asynchronous techniques. It helps develop an application that loads very fast.

3.2.4. Mysql database

MySQL is an Oracle-backed open source relational database management system based on structured query language. MySQL can work on all platforms such as Windows, Linux, Unix etc (Rouse 2019).

It is based on client-server model. The core of MySQL is named MySQL Server. It handles all of the database instructions. For security, it uses an access privilege and encrypted password system enables host-based verification. It also supports lots of client and utility programs such as MySQL Workbench, Phpmyadmin etc.

3.2.5. Centos operating system

CentOS is a well-known Linux distribution from Red Hat Enterprise Linux. CentOS stands for Community Enterprise Operating System. It uses the Red Hat source code but unlike Red Hat, CentOS doesn't require a paid subscription (Anonymous 5 2019).

CentOS has own free web hosting control panel designed for quick and easy server management via web browser. It is called CWP (Centos Web Panel.) CWP supports some open source firewalls, MySQL Database, Apache Web Server, Nginx Reverse Proxy, File browser, some third party applications etc. After installation, server management process is done via only web browser (Anonymous 6 2019).

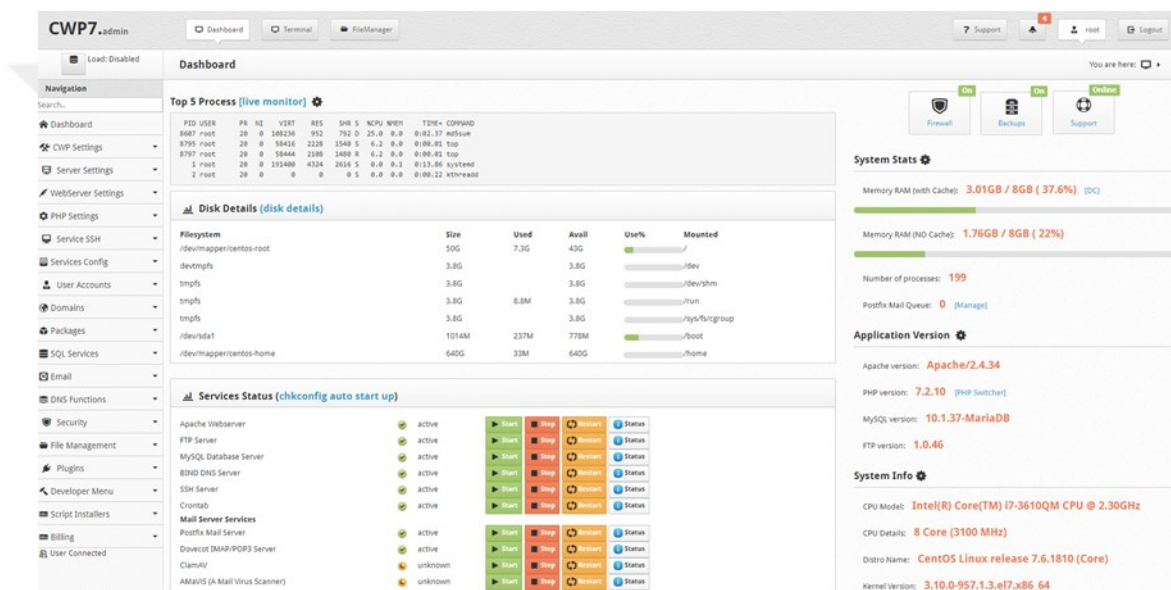


Figure 3.2. Screenshot from CWP user interface

3.2.6. MyBatis object relational mapping framework

ORM is technique for converting data between incompatible type systems using object-oriented programming languages. Generally done by frameworks such as MyBatis, Hibernate, Entity Framework etc. (Anonymous 7 2019).

By using ORM, database tables and object oriented classes mapped automatically. The development process is accelerating thanks to this. MyBatis is ORM for JAVA programming languages. It can be fetched data from database via mapper files. The contents of mapper files can be XML codes or annotations. XML usage is shown in Figure 3.3.

```
<select id="selectPerson" parameterType="int" resultType="hashmap">
  SELECT * FROM PERSON WHERE ID = #{id}
</select>
```

Figure 3.3. MyBatis mapper sample usage with XML

MyBatis usage with annotations is shown in Figure 3.4.

```
@Select("SELECT * FROM Kisi WHERE KisiID = #{kisiId}")
public Kisi getKisiById(@Param("kisiId") Long kisiId);
```

Figure 3.4. MyBatis mapper usage via annotations

3.2.7. Bootstrap framework

Bootstrap is the most popular open source HTML, CSS and javascript development toolkit. It released from Twitter in 2011. Bootstrap has own grid system. This grid system divides pages into 12 columns. By using this grid system, objects in web pages can be located easily. In Figure 3.5, the page divided 3 part by using bootstrap grid system, this grid system supports also responsive development (Anonymous 8 2019).

.col-4	.col-4	.col-4
.col-sm-4	.col-sm-4	.col-sm-4
.col-md-4	.col-md-4	.col-md-4

Figure 3.5. Bootstrap grid system

Bootstrap framework has lots of useful components such as colorful buttons, alerts, dropdowns, navbars, etc. The usage of these components are very easy by using class names. The example usage is shown in Figure 3.6.

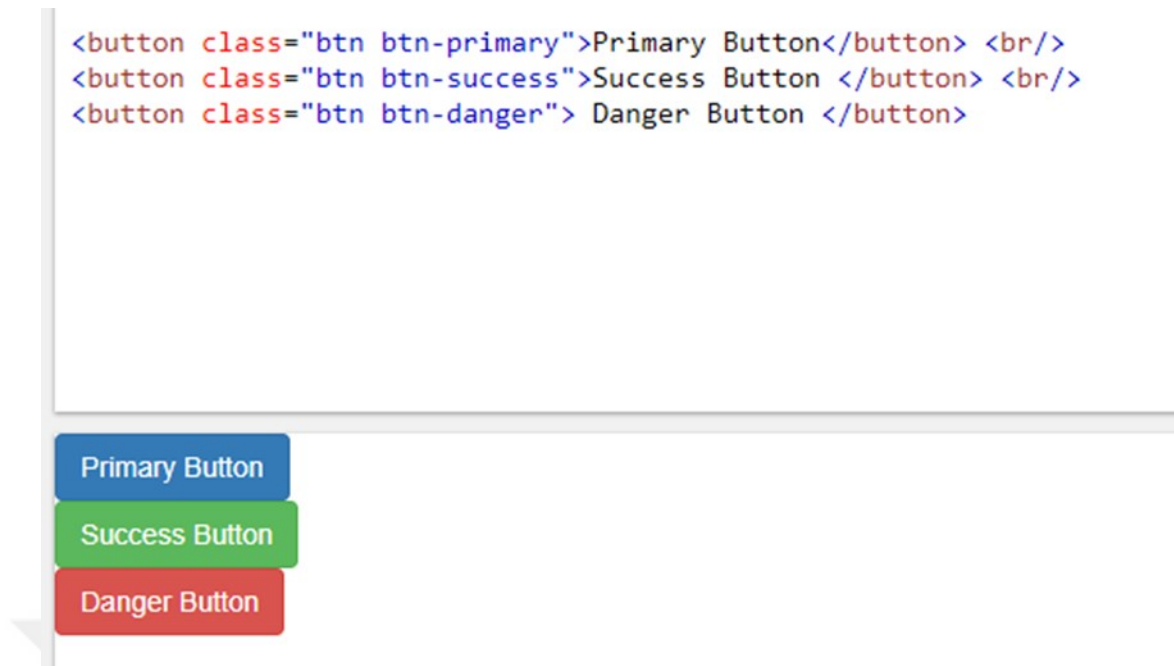


Figure 3.6. Example usage of Bootstrap toolkit

3.2.8. Apache maven

Apache Maven is a software project management and comprehension tool. Dependents sources and add-ons can be managed by POM files easily. Maven, stored library files own repository servers. When library files added project, Maven firstly search this files in local repository. If library files can't find second step of maven is search library files in Maven's own repository servers. If library files couldn't find second step, last step for Maven is search library files in defined server address from programmer. If libraries depends on other libraries, maven can find and load required libraries automatically (Anonymous 9 2019).

3.2.9. Git version management system

Git is free and open source version control system designed to handle everything from small to very large projects with speed and efficiency. It provides more than one programmer can work on the same project simultaneously (Anonymous 10 2019).

Git has branch system. When repositories initialized from git, master branch automatically created. Basic commands of git are clone, add, push, and pull. This commands

used via command line interpreter (Anonymous 10 2019).

Some well-known repositories are very popular based on Git. For example, GitHub is the most famous git repository. If project codes shared publicly, GitHub is free but for private repositories it requires cost. Generally open source projects managed in GitHub.

Another well-known repository based on Git is named GitLab. GitLab is free open source project version management system. GitLab can be installed any machine. Generally companies uses GitLab for manage their projects.

3.3. CBAEntities Project

CBAEntities project contains entities and mappers which is common for other projects. This project is a layer between the database and other projects. Contents of this projects are java class files and java interface files. To work with Mybatis ORM, mappers are saved as an interface and entities are saved as a class file. The structure of this project is shown in Figure 3.7.

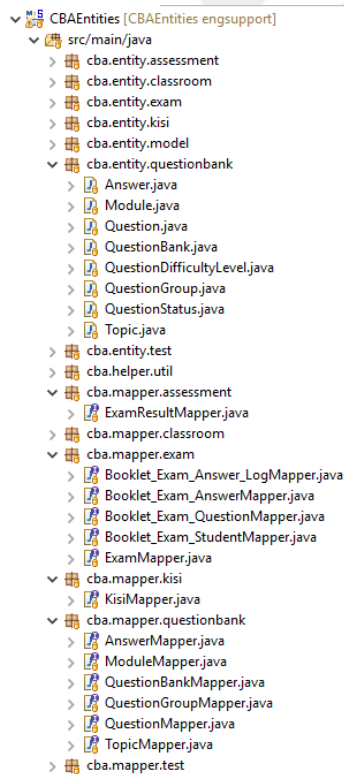


Figure 3.7. Structure of CBAEntities project

3.4. CBAInstructor Project

CBAInstructor project is a Spring MVC project. This interface project was developed for instructors. The output of CBAEntities project (CBAEntities.jar) is included in this project as a model layer of MVC pattern. The structure of this project is shown in Figure 3.8.

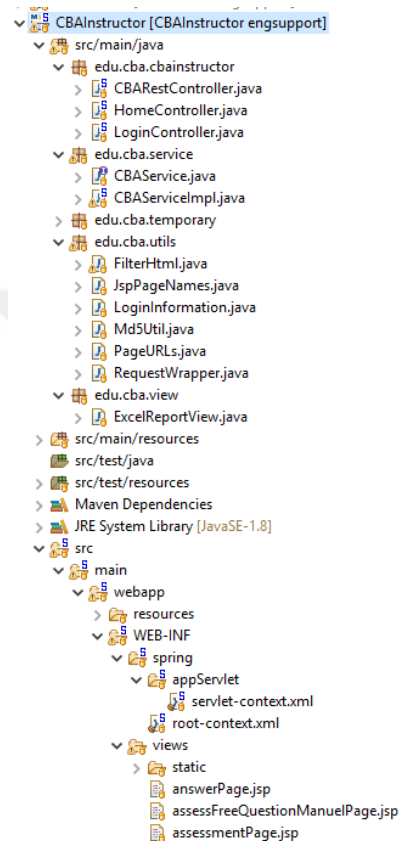


Figure 3.8. Structure of CBAInstructor project

The content of this project is divided into four part. The first part is named my question bank, the second part is named my tests, the third part is named my classrooms and last part is named as my exams. The screenshot of the main screen is shown in Figure 3.9.

To design of user interface, Bootstrap CSS and JavaScript framework is used. In order to transfer data to the view part, services are called in the controller part. Called service makes a connection to the database and fetch data from the database. After the fetching process, fetched data are returned to the controller. In controller, fetched data is bound via model.addAttribute method. In view part, the bound data can be accessed through JSTL

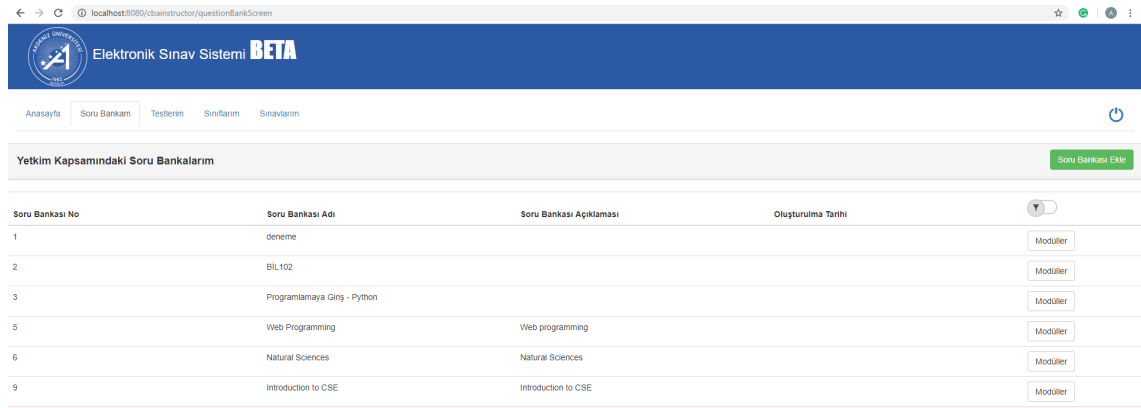


Figure 3.9. Screenshot of CBAInstructor project main screen

methods. This usage is shown in Figure 3.10.



Figure 3.10. Data transfer in controller to view part

In question bank part, instructors can add own question banks, modules, topics, questions, and answers. Question statistical analyzes are displayed in this part.

In the test part, instructors can create tests with selected questions in question banks. This selection is done in two way. In the first way, only topics, question difficulty level, and question count are selected by the instructor. The system selects questions from selected topics automatically. This process is shown in 3.11.

In this type of selection, automatically created questions can be changed manually. In the second way, test questions are selected one by one from the instructor. This selection procedure is shown in Figure 3.12.

After the selection procedure, the format of the test is shown in Figure 3.13. This classification greatly simplifies topic-based achievement analysis. Instructors can select questions and answers will be shuffled or not for the test booklet creation.

Figure 3.11. Automatic test question creation process

In classroom part, instructors can create classrooms and add students to these classrooms. Classroom entity has language property. Developed CBA system has multi language support. If classroom language is selected Turkish, students can display questions in Turkish language. If classroom language is selected English, students can display same questions in English language. While questions are inserting, instructors have an option that insert question text in second language.

In exam part, instructors can create exams, assess exams, display exam results and assess open-ended questions. In the exam creation process, test name, classroom name, exam date, examination time, exam creation type, multiple choice question count, total multiple choice question point and exam entrance code are entered by instructor. As an exam creation type, two type exam creation are implemented. In first type, all of the students take the test with same questions. If shuffling selected, all of the examinees take same questions and answers but different order. In second type, all of the examinees take test with same format but their booklets may contain different questions. Exam creation page is shown in Figure 3.14.

After the exam, assessment procedures are performed exam part of the system. If all

The screenshot shows a web interface for selecting questions. At the top, there are four dropdown menus for filtering: 'Soru Bankası' (Programlamaya Giriş - Python), 'Modül' (Programlamaya Giriş-Python), 'Konu' (Değişkenler), and 'Soru Zorluk Derecesi' (KOLAY). Below these, there are three sections, each starting with 'Aşağıdaki soruyu cevaplayınız.' (Answer the following question). The first section contains a question: 'Python yorumlayıcısına 10**2/(9-2) yazıldığında elde edilecek değer hangisidir?' (Which value will be obtained when 10**2/(9-2) is written in the Python interpreter?). The options are 2, 2.8, 14, 14.2, and 0.2. The second section contains a question: 'Python yorumlayıcısına str(5 * 2) + "10" yazıldığında elde edilecek değer hangisidir?' (Which value will be obtained when str(5 * 2) + '10' is written in the Python interpreter?). The options are '1010', 100, 20, '100', and 10. Each question has an 'Ekle' (Add) button to its right.

Figure 3.12. One by one question selection process

of the questions are multiple choice, directly click of the assess button is enough to reach exam results. If exam contains open-ended questions, first step of assessment procedure is assessing open-ended questions. This process is handled in assess open-ended question section. In this section, instructors give a point for questions to students one by one. This screen is shown in Figure 3.15.

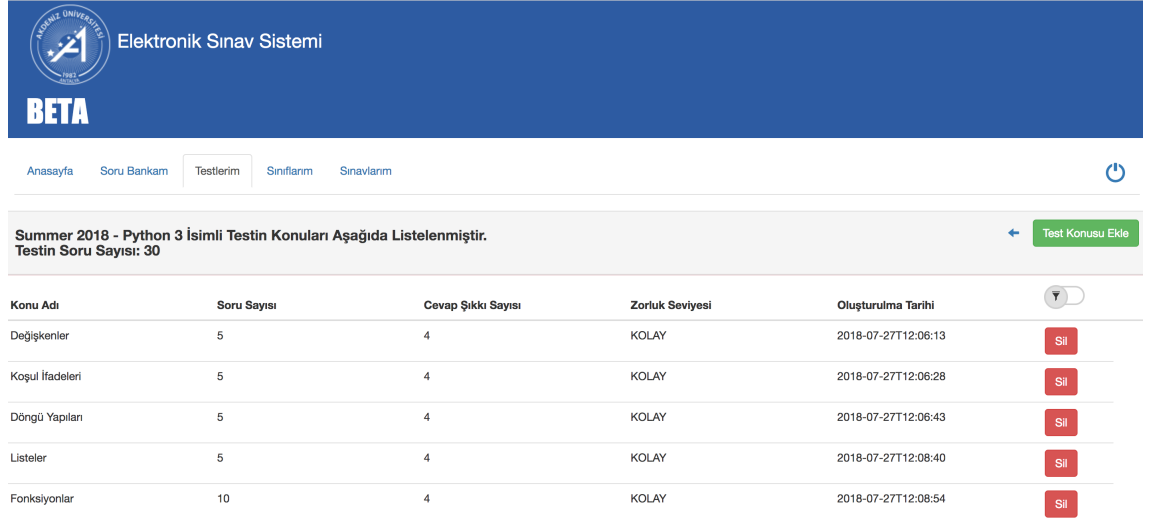
After the assessing open-ended question process, in the exam screen after clicking assess button, assessment of multiple choice questions is done by the system and results are displayed on the screen. This screen is shown in Figure 3.16.

Analysis of questions are done in the analysis page as shown in Figure 3.17. Through the analysis of questions using Item Response Theory, questions may be identified for withdrawal for future use or enhanced by the instructor.

3.5. CBASStudent Project

CBASStudent project is a web based Spring MVC project. This project was developed for examinees. The output of CBAEntities project (CBAEntities.jar) is included in this project as a model layer of MVC pattern.

Examinees can access this project via web browser. They can login with their student



Elektronik Sınav Sistemi
BETA

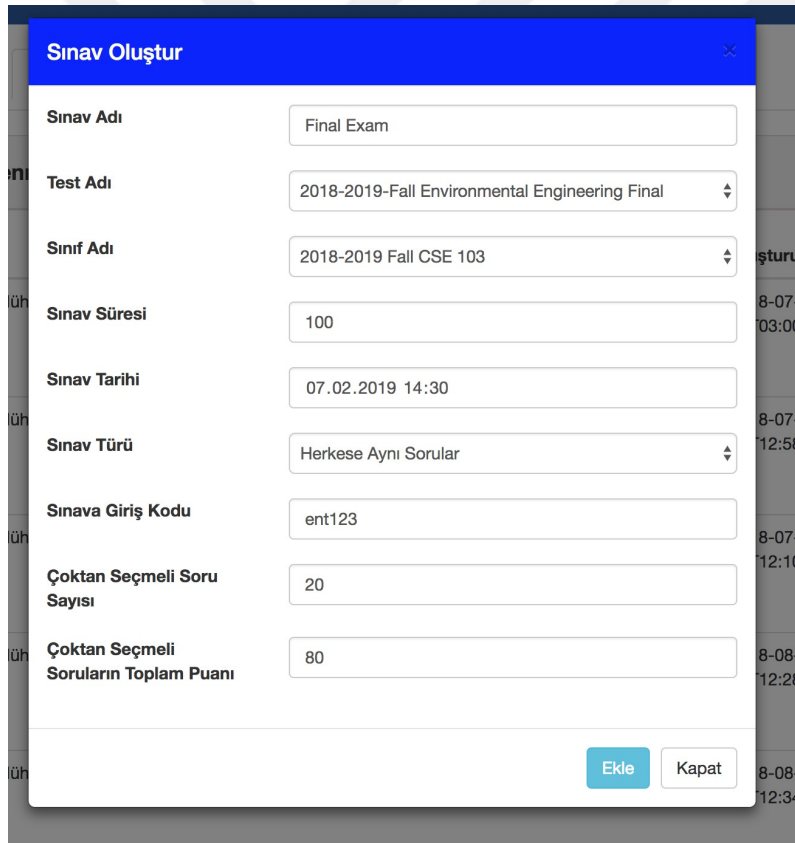
Anasayfa Soru Bankam Testlerim Sınıflarım Sınavlarım

Summer 2018 - Python 3 İsimli Testin Konuları Aşağıda Listelenmiştir.
Testin Soru Sayısı: 30

Test Konusu Ekle

Konu Adı	Soru Sayısı	Cevap Şıkkı Sayısı	Zorluk Seviyesi	Oluşturulma Tarihi	
Değişkenler	5	4	KOLAY	2018-07-27T12:06:13	Sil
Koşul İfadeleri	5	4	KOLAY	2018-07-27T12:06:28	Sil
Döngü Yapıları	5	4	KOLAY	2018-07-27T12:06:43	Sil
Listeler	5	4	KOLAY	2018-07-27T12:08:40	Sil
Fonksiyonlar	10	4	KOLAY	2018-07-27T12:08:54	Sil

Figure 3.13. View of test format



Sınav Oluştur

Sınav Adı: Final Exam

Test Adı: 2018-2019-Fall Environmental Engineering Final

Sınıf Adı: 2018-2019 Fall CSE 103

Sınav Süresi: 100

Sınav Tarihi: 07.02.2019 14:30

Sınav Türü: Herkese Aynı Sorular

Sınava Giriş Kodu: ent123

Çoktan Seçmeli Soru Sayısı: 20

Çoktan Seçmeli Soruların Toplam Puanı: 80

Ekle Kapat

Figure 3.14. Exam creation page

number and password. In the main screen, assigned exams are listed. Students can enter exam or display exam results in this screen.

DETA

Geri Dön Adı Soyadı: [REDACTED] Öğrenci No: [REDACTED] Sınav Adı : 2018-2019 Fall CSE 103 Final

Answer the following question.

What is SQL and where do you use it ?

Öğrenci Cevabı

if there are multiply users and if we use excell or metin belgesi,its take big space and computer get hard, we use sqj in database, and sqj make easy to save data for instance bank information or students grade

Verilen Puan: 7,0 Kaydet

<< Önceki Öğrenci Üniversitesi. Her hakkı saklıdır. Sonraki Öğrenci >>

Figure 3.15. Assess free answer question page

2018-2019 Fall CSE 103 Final Adlı Sınavın Sonuçları Aşağıdadır. Değerlendirmeyi İptal Et

Ortalama Puan = 44.608695652173914

Sınava Giren Kişi Sayısı = 69

#	Öğrenci No	Öğrenci Adı Soyadı	Doğru Sayısı	Yanlış Sayısı	Çoktan Seçmeli Puan	Açık Uçlu Puan	Toplam Puan
1	[REDACTED]	[REDACTED]	3	2	18.0	12.0	30.0
2	[REDACTED]	[REDACTED]	3	2	18.0	22.0	40.0
3	[REDACTED]	[REDACTED]	4	1	24.0	38.0	62.0
4	[REDACTED]	[REDACTED]	2	3	12.0	22.0	34.0
5	[REDACTED]	[REDACTED]	3	2	18.0	34.0	52.0
6	[REDACTED]	[REDACTED]	4	1	24.0	26.0	50.0
7	[REDACTED]	[REDACTED]	3	2	18.0	35.0	53.0
8	[REDACTED]	[REDACTED]	3	2	18.0	52.0	70.0
9	[REDACTED]	[REDACTED]	0	5	0.0	21.0	21.0
10	[REDACTED]	[REDACTED]	4	1	24.0	26.0	50.0
11	[REDACTED]	[REDACTED]	1	4	6.0	39.0	45.0
12	[REDACTED]	[REDACTED]	3	2	18.0	34.0	52.0
13	[REDACTED]	[REDACTED]	3	2	18.0	12.0	30.0
14	[REDACTED]	[REDACTED]	1	4	6.0	17.0	23.0
15	[REDACTED]	[REDACTED]	0	0	0.0	0.0	0.0

Figure 3.16. Exam result page

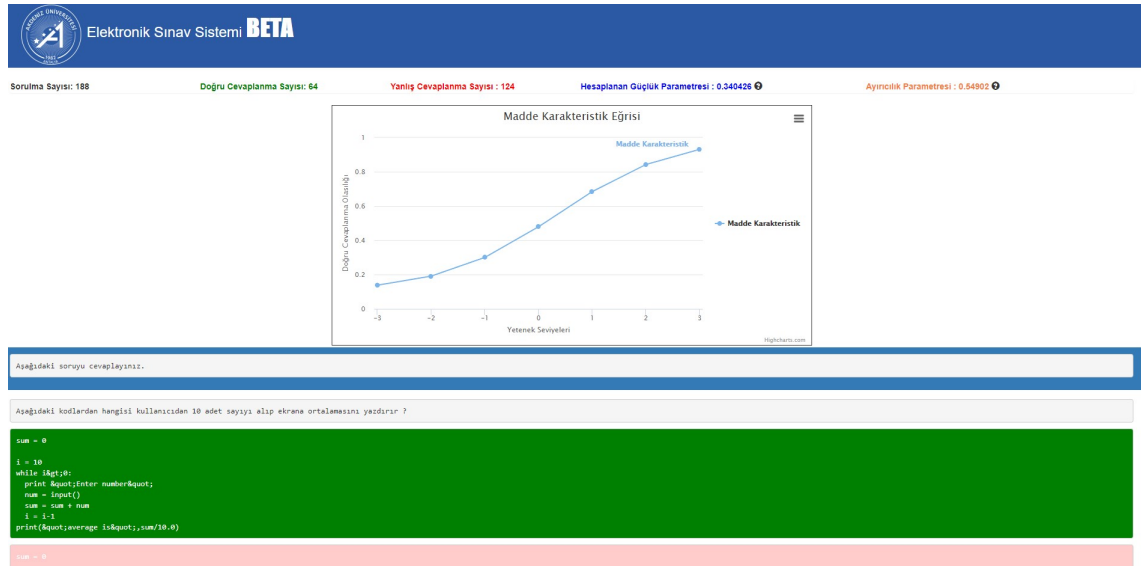


Figure 3.17. Question analysis page

In order to start exam, "go to exam" button is clicked. After clicking, system asks for the exam entrance code. If exam entrance code is entered correctly and current time is on the exam date time, exam screen is appears. The exam screen is divided into two parts, according to the type of a question. The first screen is designed for multiple-choice questions. This screen was designed using the bootstrap CSS and JavaScript framework. In this screen, top section displays student information such as student number, name and surname, department and exam name. The progress bar, which is located below the information section indicates what percent of exam questions have been answered. Under the progress bar, "previous question", "clear my answer", "check again" and "next question" buttons section are located. This section is also located below the answers section. On the left-center of the screen, the question information section is located. For multiple-choice questions, answers are listed below the question. In the right section of the screen, remaining exam time is located. The remaining time is taken from the server. Below the remaining time, a question list menu is located. In this menu, questions of the exam are listed and selected question is stripped. Right of the question, the status of each question is written. These statuses are, didn't displayed, answered, displayed and check again. Examinees can change question through clicking the question in the list. This screen is shown in Figure 3.18.

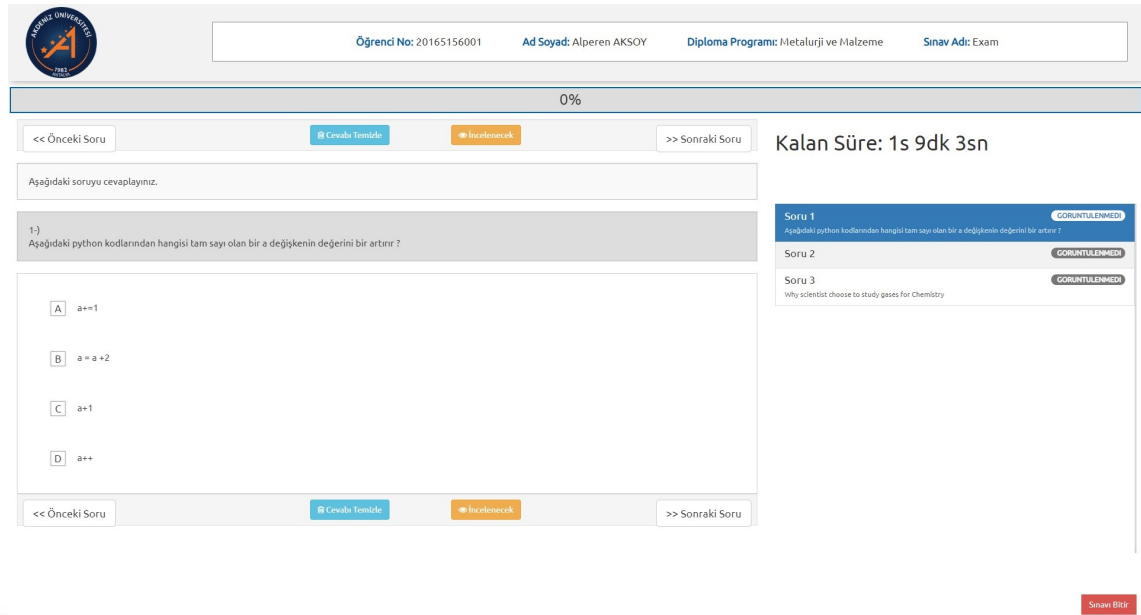


Figure 3.18. Multiple-choice question screen in exam

In this screen, answers are selected via clicking by examinee. By clicking, JavaScript Ajax function is triggered and this function send selected answer to the server.

Open-ended question screen is shown in Figure 3.19. In this screen, answers are typed in a text box. In order to send an answer to the server, "Save Answer" button is clicked. By clicking, a JavaScript Ajax function is triggered and this function sends answer to the server. Other parts of the screen are same as the multiple-choice question screen.

In proposed CBA system, images may be attached to the questions. In exam screen, these images are located on the question text section. This screen is shown in Figure 3.20.

3.6. Deployment Environment

The developed projects were deployed on a computer with the following specifications:

- Intel i7-3610QM 2.30 Ghz CPU
- 8 GB of RAM
- 750 GB 7200 RPM disk
- CentOS 7 operating system
- Java 1.8 on the server

Figure 3.19. Open-ended question screen in exam

Figure 3.20. Question with image in exam screen

- Apache Tomcat 8 application deployed
- MySQL for database

4. RESULTS AND DISCUSSION

The structure of our design is divided into 6 components; question bank, test, exam, classroom, student and instructor.

The question bank part is comprised of 6 entities:

- QuestionBank
- Module
- Topic
- QuestionGroup
- Question
- Answer

The question bank can include one or more modules, with each module being divided into topics. The question group is deployed as a layer between topics and questions. Each question must have a question group. Question groups can include more than one question. For example, paragraph questions can be grouped by using this structure. A schematic representation of this structure is shown in Figure 4.1.

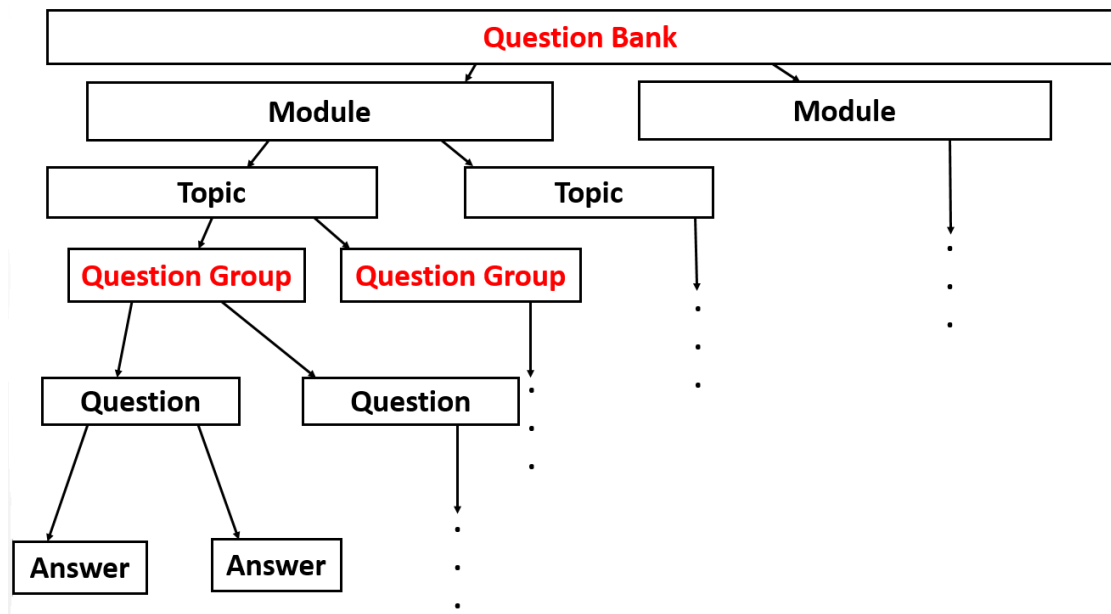


Figure 4.1. Schematic representation of the question bank

CBA systems generally shuffle questions for creating more than one exam booklet by using same questions. This shuffling may lead to some problems if the shuffling is done directly through the questions. For example, if question bank include paragraph questions or cloze test questions, after the shuffling process, these questions may not be located sequentially. An example of this case is shown in Figure 4.2.

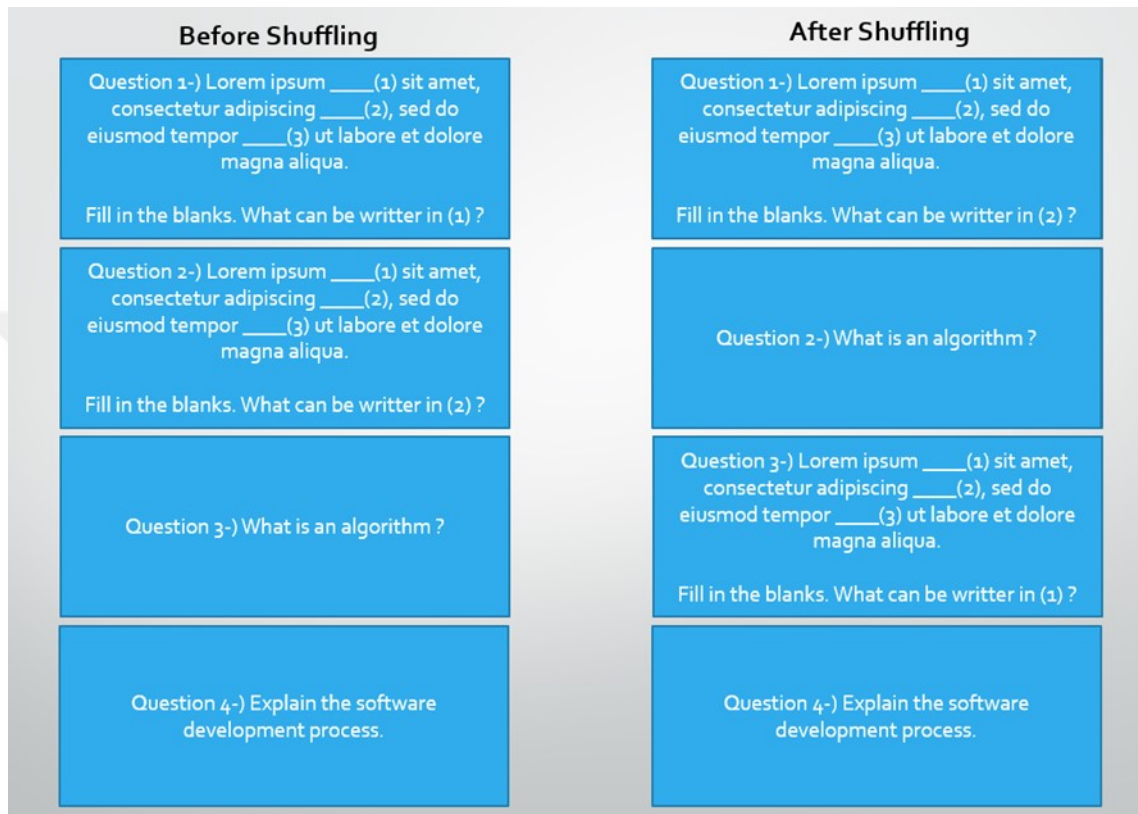


Figure 4.2. An example of a shuffling operation by directly used on questions

In Figure 4.2, shuffling operation was directly applied to the questions. In this question set, we have a close test questions. Before the shuffling, question 1 and question 2 was located sequentially. There is a bond between first and second questions, they have the same paragraph text and examinee can't respond to the second question before the first question. After shuffling, the second question was located the first question of a new booklet. This shuffling operation gives adversely affects for examinee exam process.

To Prevent the scenario given in Figure 4.2, questions can be grouped by using question group entity. For the solution of this problem, each question must belong to the one question group. Question group can have more than one question. Once these require-

ments met, the shuffling operation can be applied to the question groups. This operation ensures questions in the question group are located sequentially after the shuffling process in a newly created booklet.

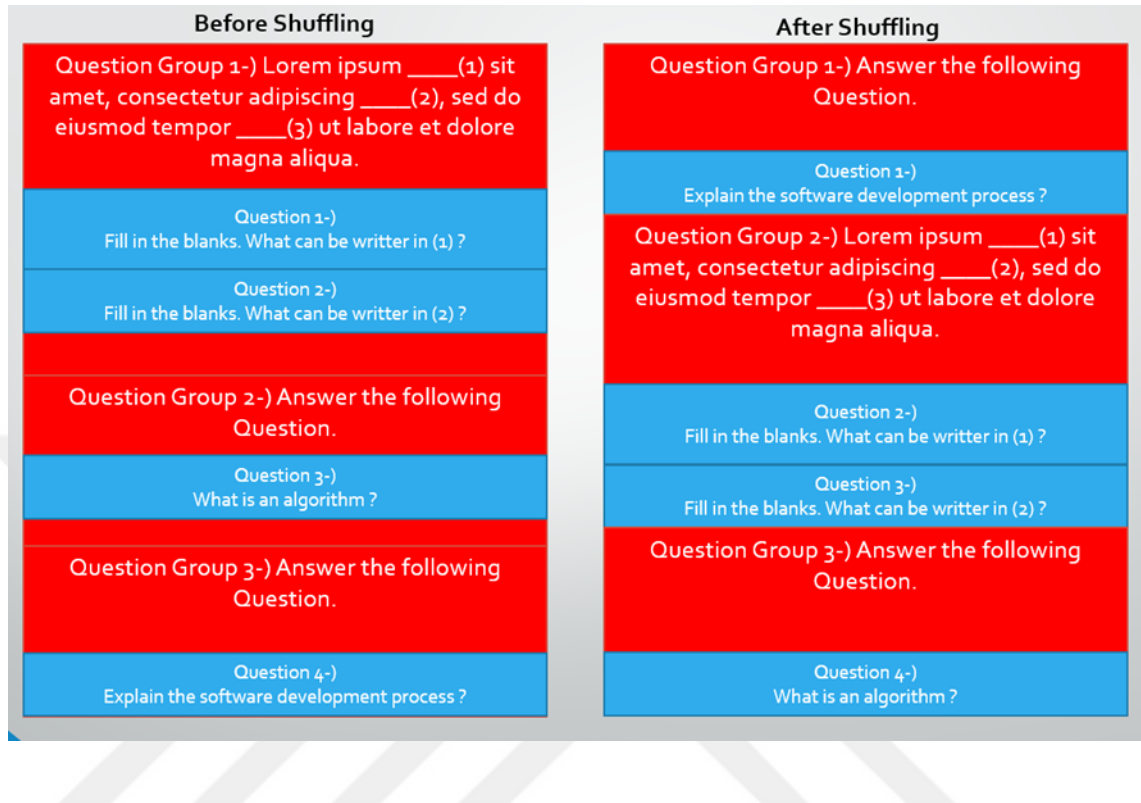


Figure 4.3. An example of shuffling operation used on question groups

In Figure 4.3, shuffling operation was applied to the question groups. After shuffling, dependent questions were located sequentially. Question group has property named isShuffle. If this property is selected true, questions in the question group can be shuffled between each other.

The process of selection of questions for creating a test booklet is called the test creation process. While creating a test in CBA system, users have two options. In the first option, users can select questions and answers in the test booklet manually. In this option, users can filter by question bank, module, topic and question difficulty level when selecting questions. In other option, only question counts, topics and difficulty level are selected. Questions are selected automatically by the system according to the given information. Automatically selected questions can be changed manually via web based user interface.

Student entity denotes to the examinees. This entity keeps student number is also used

as a login username for the student. Other properties of this entity are name, surname, and department. The instructor can register students one by one or collectively. Collectively insertion is done via Excel files.

Classroom entity is related to the student entity. Each student must belong to at least one classroom in the CBA system. The task of the this entity being is to group student.

Exam entity is used to assign a test to students via classroom entity. Each exam must be assigned to the classroom. Each student in the selected classroom takes a test in selected test format. If a user selects different questions mode, the system assigns a test to students different questions but the same format. In other mode, students take the same questions and answers but they are shuffled. For each student, novel booklet is created by the system. This entity has exam entrance code property which is required for students to join the exam.

4.1. Survey

After the development stage, 30 electronic exam was handled by using developed CBA system at Akdeniz University Engineering Faculty. Exams was conducted computer laboratories of Engineering Faculty. Introduction to Programming, Web Programming, Natural Sciences and Introduction to Computer Engineering courses' exams was conducted in electronic environment.

After the exams, a questionnaire was applied to 261 bachelor students of Engineering Faculty. 47.5% of these students from the mechanical engineering department, 36% are from the computer engineering department and 16.5% was environmental engineering department. 77% of the participants are male, 23% of the participants are female. The age range of participants is 18-27. Results are given in Table 4.1.

Table 4.1. Questionnaire results from examinees

Question	Strongly Disagree	Disagree	Partially Agree	Agree	Strongly Agree
Electronic exams reduce my performance	29.5%	40.2%	17.6%	7.3%	5.4%
It is easier to cheating on electronic exams	18%	36%	25%	13%	8%
Electronic exam system shouldn't use in Universities	44.4%	36%	12.6%	5%	1.9%
Electronic exam system should use in multiple choice question tests	5.7%	5.7%	14.6%	28.4%	45.6%
Electronic exam system should use in open-ended question tests	28.4%	18.8%	24.1%	18%	10.7%
Exam interface was sufficient	3.8%	6.1%	20.3%	39.1%	30.7%
I think there are vulnerabilities in the electronic exam system	18%	36%	24.5%	13.5%	8%

4.2. Question Quality

Question bank for Introduction to Python course has 117 questions. These questions were created by the committee from the Computer Engineering department of Akdeniz University. Question discrimination parameters, question difficulty parameters, and other item parameters were calculated by the developed system. Distribution of question discrimination parameter for 117 questions is given in Table 4.2.

Table 4.2. Distribution of question discrimination parameter for 117 questions

Question Count	Question Discrimination Parameter
54 (42.2%)	> 0.4
25 (21.4%)	Between 0.3 - 0.4
22 (18.8%)	Between 0.2 - 0.3
16 (13.6%)	< 0.19

According to the question analyzes, 16 questions were removed which are question discrimination parameter is lower than 0.20. Because these questions can't enough to distinguish students with their ability level. Total of 22 questions was revised to increase the question discrimination parameter.

4.3. Resource Consumption

In the deployment environment, resource consumption of the developed CBA system is given in Table 4.3. The measurements were recorded in an exam of 117 participants.

Table 4.3. Resource consumption of developed CBA system

Metric	Result (Mean)
CPU Usage	3.4%
Hits Per Minute	288 Hits
Used Physical Memory	390 Megabyte
Response Time	66 Millisecond
Used Java Memory	195 Megabyte

According to Table 4.3, developed CBA system doesn't consume much resource.

Mean response time is 66 millisecond which is reasonable. Examinees don't experience any loss of time changing questions. Only 3.4% mean usage for CPU and 390 Megabyte physical memory consumed. More basic systems can be enough to run developed CBA system. Any complaint not recorded about the speed of the developed system.

4.4. Comparison with the PPA

In this work, Natural Science course was selected for analysis. This course is given at the Akdeniz University, Department of Computer Engineering. First-year students take this course. To assess a student's academic success, two midterm exams and one final exam were applied. In these exams, all of the questions are open-ended question type. It means that students write their own words to answer these questions. In 2017, these exams were held in paper and pencil based environment. In 2018, students took their exams in the developed web based computer environment. The ability level of students was almost equal. Because they were selected in an exam by given from student selection and placement center in Turkey and their scores almost the same.

The information about these two group's total student count, average point of natural science course, pass rate, standard deviation, standard error and variance values are given in Table 4.4.

Table 4.4. Information about the student groups

Metric	2017 (Paper Based)	2018 (Computer Based)
Total Student	101	118
Average Point (Out of 100)	51,86	49,35
Pass Rate	53.5%	50.0%
Standard Deviation	24,28	23,61
Standard Error	2,41	2,17
Variance	589,61	557,80

Independent student t-test was applied for these two group's final total points. This test is generally applied to determine whether the two independent samples are different from each other in terms of a given variable. To apply this test, IBM SPSS Statistic 23 package software was used.

The results of independent student t-test is given in Table 4.5.

Table 4.5. T-Test results for total final points

F Value Signature	T Value	P Value
0,825	0,772	0,441

According to the independent student t-test, there is no significant difference about examinees' total final points between computer based exam group and paper based exam group (p value > 0.05).

4.5. Cheat Detection

In developed CBA system, in order to prevent cheating between examinees questions and answers are shuffled. For multiple-choice questions, this shuffling operation is sufficient but if question count of exam is less and questions are open-ended type, shuffling isn't sufficient. The settlement order of computer labs aren't designed for exams. They generally designed for courses. To catch copying answers between examinees, similarity analysis techniques can be used after the exam. In the scope of this thesis, two similarity analysis technique are handled. Jaccard and Cosine similarity analysis techniques.

By using these techniques, several copies of cases detected. Few examples are shown below. Student answers were given pure form.

- **Question: What is SQL and where do you use it?**
- **Student Answer 1:** SQL is database. we can find space any websites or programmes so they can be done non error.if we found any space we can enter program or website.
- **Student Answer 2:** Sql is database. We can find space any websites or programmes so they can be done non error. If we found any space, we can enter this program or websites.
- **Question: What does the software development lifecycle? Explain steps of the development lifecycle.**
- **Student Answer 3:** It use to programming. thanks to it,programs is better than

before. it is whole systematic rules. understand the problem writing a program implementing testing. algorithms with flowcharts pseudocode

- **Student Answer 4:** It use programming thanks to lifecycle.it is whole very systemic rules. it programs is better than before.
- **Question: Technology sometimes develops ahead of our scientific understanding of the technological processes. What does this mean. You may give an example.**
- **Student Answer 5:** They didnt have enough theroic knowledge, they invented something but they thoght there should be something for this inventions to work and then they find some theoric knowledge. For example they didn't know anything about thermodinamics but they invented steam trains. Another example, Galileo invented his own telescope with 30X zoom but he didn't know anything about optics
- **Student Answer 6:** for examplen try didnt know thermodinamick butb they invented steam trains.they didnt have enough theoric knowledge they invented sometething but they thought there should be sometimes for this inventions to work. Another examplr Galileo invnted his own telescope but he didnt know anythin about optics. That is why he bought mathematic, trigonometri books and he read a lot of scientist examining.

Table 4.6. Similarity rates for a few copying examples detected by the system

Answer Set	Jaccard Similarity Rate	Cosine Similarity Rate
Answer 1-2	0.67	0.94
Answer 3-4	0.28	0.76
Answer 5-6	0.35	0.70

Some copying examples and their similarity rates were given in Table 4.6. According to the this table, Cosine similarity technique gives better results than the Jaccard similarity technique. If the threshold value of Jaccard similarity analysis is decreased, it gives many answer set as a copy which isn't copied from other examinees. If examinee steal only subset of other examinee's answer, Jaccard similarity rate is decreased and system isn't

detect this answer as a copy. However, in this case, the cosine similarity metric isn't decreasing too much and developed system detect these answer set as a copy.

Nowadays, there is not many open source CBA systems. The most widely adapted open source CBA system is web based TCEXAM. It was developed by using PHP programming language and MySQL database. The comparison of features of TCEXAM software and developed CBA system is given in Table 4.7.

Table 4.7. Comparison of features of TCEXAM and developed CBA system

Feature	TCEXAM	Developed CBA System
Developed by using open-source technologies	✓	✓
Support of open-ended and multiple-choice questions	✓	✓
Attachment of multimedia content to the questions	✓	✓
Question analysis with IRT	X	✓
Intelligent shuffling mechanism for group questions	X	✓
Import/export students and questions from excel file	✓	✓
IP-based security restrictions	✓	✓
Support of cheating analysis	X	✓
Question bank sharing mechanism between instructors	X	✓

During the exam, there is no complaint recorded about the speed of the developed system. Computer startup times were quite long for some computers on the computer labs of Akdeniz University Engineering Faculty. To prevent this obstacle, 15 minutes before the exam all computers were set for the exam.

According to the questionnaire results, 69.7% of respondents think that the usage of CBA doesn't give badly affect their performance. 80.4% of examinees think that developed CBA system should use for exams in universities. While 88.6% of respondents think that the developed CBA system should be used in multiple-choice question tests, while 52.8% of examinees think it should be used in open-ended question tests. But 40% of the participants didn't see the open-ended question test environment. All of the questions about their exams were multiple-choice questions. The reason for this low rate may be considered to be prejudiced or generally, students don't want to reply open-ended questions. 90.1% of respondents think that the exam interface was sufficient.

For comparison CBA and PPA on student's academic achievement, Natural Science

course was selected. In 2017 exams of this course was handled in paper based environment. In 2018 these exams were handled in computer based environment. According to the statistical analysis of final grades, there was no significant difference between two years of grades. These findings are shown that the previous studies were correct. Computer usage familiarity couldn't be compared, because of all of the examinees was already taken computer usage courses.



5. CONCLUSION

Within the scope of this thesis, the electronic examination system, which is a data collection tool that can be used in academic studies, has been developed by using open source web technologies. The developed system still in use at Akdeniz University Engineering Faculty. Questions and tests are analyzed item-response theory principles. Adaptive tests which are an important application of item-response theory are compatible with the developed system. The system can be used as a test environment for adaptive test algorithms.

By using the question analysis module, exam questions were analyzed and revised. By means of this, the revised questions were used next exams and better measurement and assessment environment was provided.

Web-based CBA system was developed with open source Java web technologies. After the development phase, thirty exams were applied in electronic environment. Approximately 15.000 paper and many inks saved thanks to the developed CBA system.

Electronic exams were held in computer laboratories of Akdeniz University Engineering Faculty. These laboratories are designed for courses. Because of this design, the settlement of computers was inappropriate for exams. Although exams are conducted with supervisors, it isn't sufficient to prevent copying. To detect copy cases, two similarity analysis technique was implemented after the exams and these methods were compared. According to the findings, the Cosine similarity analysis technique yielded better results than the Jaccard similarity analysis technique. For this problem, it is more convenient to use the Cosine similarity analysis technique.

As future studies, the electronic exam system can be further developed to make real adaptive exams available through adaptive algorithms tested. There can also be studies to increase the question quality of the exams by sharing the questions with their analysis results between the instructors who give the same course.

6. REFERENCES

- Akdemir O. and Oguz A. 2008. Computer-based testing: An alternative for the assessment of Turkish undergraduate students. *Computers & Education*, 51 (3): 1198-1204.
- Anonymous 1: 6 Types of Assessment of Learning. https://www.teachthought.com/pedagogy/6_types_assessment_learning [Last access date: 01.12.2018].
- Anonymous 2: Computerized Adaptive Testing Implementations. https://en.wikipedia.org/wiki/Computerized-adaptive_testing [Last access date: 10.05.2019].
- Anonymous 3: Java Language (programming language). [https://en.wikipedia.org/wiki/Java_\(programming-language\)](https://en.wikipedia.org/wiki/Java_(programming-language)) [Last access date: 3.4.2019].
- Anonymous 4: Model View Controller Design Pattern. <https://www.geeksforgeeks.org/mvc-design-pattern> [Last access date: 6.5.2019].
- Anonymous 5: What is CentOS ?. <https://www.computerhope.com/jargon/c/centos.htm> [Last access date: 7.5.2019].
- Anonymous 6: CentOS Web Panel. <http://centos-webpanel.com> [Last access date: 6.5.2019].
- Anonymous 7: Object-relational mapping technologies. https://en.wikipedia.org/wiki/Object_relational_mapping [Last access date: 7.5.2019].
- Anonymous 8: Bootstrap. <https://getbootstrap.com> [Last access date: 4.4.2019].
- Anonymous 9: Apache Maven. <https://maven.apache.org> [Last access date: 5.5.2019].
- Anonymous 10: About-Git. <https://git-scm.com/about> [Last access date: 6.4.2019].
- Bain, S.K. and Allin, J.D. 2005. Book review: Stanford-Binet intelligence scales, fifth edition. *Journal of Psychoeducational Assessment*, 23 (1): 87-95.
- Birnbaum, A. 1968. Some latent trait models and their use in inferring examinee's ability. In: Lord, F.M. and Novick, M.R. (Ed.), *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, 397-479.

- Blom J. et al. 2018. Perceptions and performance using computer-based testing: One institution's experience. *Currents in Pharmacy Teaching and Learning*, 10 (2): 235-242.
- Bull, J. and McKenna, J. 2004. *Blueprint for Computer-Assisted Assessment*. Routledge, London and New York, 208 pp.
- Carroll, J. How Java Works. <http://www.cs.cmu.edu/jcarroll/15-100-s05/supps/basics/history.html> [Last access date: 5.5.2019].
- Celen, U. 2008. Comparison of validity and reliability of two tests developed by classical test theory and item response theory. *Elementary Education Online*, 7 (3):758-768.
- Chalmers, D. and McCausland, D. 2002. Computer-Assisted Assessment. In: *The Handbook for Economics Lecturers*. Glasgow Caledonian University. Bristol , 21 pp.
- Chand, S. What is the benefit of using spring framework. <https://www.quora.com/What-is-the-benefit-of-using-spring-framework> [Last access date: 6.5.2019].
- Clariana R. and Wallace P. 2002. Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, 33 (5): 593-602.
- Conole, G. and Warburton, B. 2005. A review of computer-assisted assessment. *Research in Learning Technology*, 13 (1): 17-31.
- Crocker, L. and Algina, J. 1986. *Introduction Classical and Modern Test Theory*. Wadsworth Publication, USA.
- Daniels L.M. and Gierl M.J. 2017. The impact of immediate test score reporting on university students' achievement emotions in the context of computer-based multiple-choice exams. *Learning and Instruction*, 52 (1):27-35.
- Ebel, R.L. 1965. *Measuring educational achievement*. Prentice-Hall. New Jersey. 481 pp.

- Embretson, S.E. 1999. Issues in the measurement of cognitive abilities. In S.E. Embretson S.E. and Hershberger L. (Eds.), *The new rules of measurement: What every psychologist and educator should know*. Psychology Press, New Jersey, pp. 1-15.
- Emretson, S.E. and Reise, S.P. 2000. *Item response theory for psychologists*. Psychology Press, New Jersey, 384 pp.
- Kelley, T.L. 1939. The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, 30 (1):17-24.
- Larman, C. and Basili, V. 2003. Iterative and Incremental Development: A Brief History. *IEEE Computer*, 36 (6): 47-56.
- Lord, F. 1980. *Applications of item response theory to practical test problems*. Lawrence Erlbaum Associates. New Jersey. 274 pp.
- Noyes, J. et al. 2004. Paper-based versus computer-based assessment: is workload another test mode effect ?. *British Journal of Educational Technology*, 35 (1):111-113.
- Pawasauskas J. et al. 2014. Transitioning to computer-based testing. *Currents in Pharmacy Teaching and Learning*, 6 (2): 289-297.
- Reid, C.A. et al. 2007. Modern psychometric methodology: Applications of item response theory. *Rehabilitation Counselling Bulletin*, 50 (3):177-188.
- Rouse, M. What is MySQL ?. <https://searchoracle.techtarget.com/definition/MySQL> [Last access date: 6.5.2019].
- Wainer, H. and Thissen, D. 2001. True score theory: The traditional method. In: *Test Scoring*. Lawrence Erlbaum Associates, New Jersey, pp. 23-72.
- Weis, D.J. 1985. Adaptive testing by computer. *Journal of Consulting and Clinical Psychology*, 53 (6): 774-789.

CURRICULUM VITAE

ALPEREN AKSOY

alperenaksoy41@hotmail.com



EDUCATION

Master of Science 2016-2019	Akdeniz University Institute of Natural Sciences, Department of Computer Engineering, Antalya
Bachelor of Science 2012-2016	Erciyes University Engineering Faculty, Department of Computer Engineering, Kayseri

WORK EXPERIENCE

Research Assistant 2018 -	Akdeniz University Institute of Natural Sciences, Department of Computer Engineering, Antalya
Research Assistant 2017-2018	Alanya Alaaddin Keykubat University Engineering Faculty, Department of Computer Engineering, Antalya

PUBLICATIONS

Articles published in other peer reviewed international journals

- 1- Aksoy A., Aslan S., Karaboga D. (2018). Order Based Emigrant Creation Strategy for Parallel Artificial Bee Colony Algorithm. *Journal of Selcuk Technic*, Special Issue 2018, 58-75.
- 2- Aksoy A., Ozturk C. (2018). Mobile Based Traffic Offences Report System. *Journal of Intelligent Systems with Applications*, 1:175-179.

- 3- Aslan S., Karaboga D., Aksoy A. (2018). A New Emigrant Creation Strategy with Randomized Sources of Parallel Artificial Bee Colony Algorithm. *Journal of Intelligent Systems with Applications*, 1:81-86.

Papers delivered in international conferences and printed as a proceedings

- 1- Aslan S., Aksoy A., Gunay M., Performance of Parallel Artificial Bee Colony Algorithm on Solving Probabilistic Sensor Deployment Problem. International Conference on Artificial Intelligence and Data Processing (IDAP), MALATYA, TURKEY, 28-30 September 2018, pp. 1-5.

- 2- Aslan S., Aksoy A., Solving Wireless Sensor Deployment Problem with Parallel Artificial Bee Colony Algorithm. 26th Signal Processing and Communications Applications Conference, İZMİR, TÜRKİYE, 2-5 May 2018, pp.1-4.

- 3 – Aslan S., Aksoy A., Gunay M., Performance Analysis of Artificial Bee Colony Algorithm for High Dimensional Problems on ARM Based Mobile Platform. 1st International Symposium on Multidisciplinary Studies and Innovative Technologies. ANTALYA, TURKEY, 11-13 April 2018, pp. 93-93.

- 4 – Karaboga D., Aslan S., Aksoy A., Finding DNA Motifs with Collective Parallel Artificial Bee Colony Algorithm. International Conference on Artificial Intelligence and Data Processing (IDAP). MALATYA, TURKEY, 28-30 September. pp.1-7.

- 5 – Aslan S., Aksoy A., Gunay M., Performance Analysis of Graphical Processors with Calculation of Sum of Pairs Score. 26th IEEE Signal Processing and Communications Applications Conference. İZMİR, TURKEY, 2-5 May 2018. pp.1-4.

- 6 – Karakoc M. Aksoy A., Sen E., Gunay M., Outcome Focused Computer Based Assessment System using Intelligent Shuffling Algorithm. 26th International Congress of Educational Sciences (ICES-UEBK 2017), ANTALYA , TURKEY, 21-23 April 2017. pp. 1283-1284.

- 7- Aslan S., Karaboga D., Aksoy A., Performance Analysis of Artificial Bee Colony Algorithm on ARM Based Mobile Platform. International Conference on Computer Science and Engineering (UBMK), ANTALYA, TURKEY, 5-8 October 2017, pp. 155-159.

- 8 – Aksoy A., Gunay M., Karakoc M., A Flexible Question Storage Hierarchy for Computer Based Assessment Systems. International Conference on Engineering Technologies (ICENTE 17). KONYA, TURKEY, 7-9 December 2017, pp.17-17.