

TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DİYABET AKILLI DESTEK SİSTEMİ: Bilgi Erişim Teknikleri ile
Kan Şekeri Seviyesinin Tahmini

YÜKSEK LİSANS TEZİ

Adil Kondilođlu

Elektrik ve Bilgisayar Anabilim Dalı

Elektrik ve Bilgisayar Mühendisliđi Programı

MART 2015

TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DİYABET AKILLI DESTEK SİSTEMİ: Bilgi Erişim Teknikleri ile
Kan Şekeri Seviyesinin Tahmini

YÜKSEK LİSANS TEZİ

Adil Kondilođlu
1203617004

Elektrik ve Bilgisayar Anabilim Dalı

Elektrik ve Bilgisayar Mühendisliđi Programı

Tez Danışmanı : Yrd. Doç. Dr. Tayfun Küçükylmaz

MART 2015

Türk Hava Kurumu Üniversitesi Fen Bilimleri Enstitüsü'nün 1203617004 numaralı Yüksek Lisans öğrencisi, Adil KONDİLOĞLU ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı DİYABET AKILLI DESTEK SİSTEMİ: Bilgi Erişim Teknikleri ile Kan Şekeri Seviyesinin Tahmini başlıklı tezini, aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : Yrd. Doç. Dr. Tayfun KÜÇÜKYILMAZ
Türk Hava Kurumu Üniversitesi

Jüri Üyeleri : Yrd. Doç. Dr. Engin DEMİR
Türk Hava Kurumu Üniversitesi

Yrd. Doç. Dr. Selen PEHLİVAN
TED Üniversitesi

Tez Savunma Tarihi : 18.03.2015

TÜRK HAVA KURUMU ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

Yüksek Lisans Tezi olarak sunduğum, DİYABET AKILLI DESTEK SİSTEMİ: Bilgi Erişim Teknikleri ile Kan Şekeri Seviyesinin Tahmini adlı çalışmamın, tarafımdan akademik etik ve kurallara aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım kaynakların kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

18.03.2015

Adil KONĐİLOĐLU

TEŐEKKÖRLER

Yüksek Lisans tez çalışma sürecinde beni yönlendiren, karşılaştığım zorlukları bilgi ve tecrübesi ile aşmamda yardımcı olan desteğini ve yardımını hiçbir zaman esirgemeyen tez danışmanım değerli Yrd. Doç. Dr. Tayfun KÜÇÜKYILMAZ' a teşekkürlerimi sunarım.

Hayatım boyunca her zaman yanımda olan, maddi ve manevi desteklerini hiçbir zaman esirgemeyen aileme teşekkürlerimi, sevgi ve saygılarımı sunarım.

Mart, 2015

Adil KONDİLOĞLU

İÇİNDEKİLER

TEŞEKKÜRLER	ii
İÇİNDEKİLER	iii
TABLolar	v
ŞEKİLLER	vii
KISALTMALAR	ix
ÖZET	x
ABSTRACT	xii
BİRİNCİ BÖLÜM	1
1. GİRİŞ	1
İKİNCİ BÖLÜM	3
2. LİTERATÜR ÖZETİ	3
2.1 Tıbbi Çalışmalar	3
2.2 Raporlama ve İzleme Çalışmaları	4
2.3 Kan Şekeri Tahmin Çalışmaları	4
ÜÇÜNCÜ BÖLÜM	7
3. MATERYAL ve YÖNTEM	7
3.1 Veri Seti Yapısı ve Özellikleri	8
3.2 Veri Dosyaları Yapısı	8
3.2.1 İnsülin Verisi Özellikleri	9
3.2.2 Kan Şekeri Verisi Özellikleri	10
3.2.3 Yemek Verisi Özellikleri	12
3.2.4 Aktivite Verisi Özellikleri	12
3.3 Veri Dosyaları İşlemleri ve Veritabanı Yapısı	14
3.3.1 Verileri Temizlenme ve İndirgeme	14
3.3.2 Ana Veri (Corpus)	15
3.4 Öznitelik Oluşturma	16
3.4.1 Öznitelik ve Öznitelik Dosyaları	16
3.4.2 Zaman Dilimi Özellikleri	16
3.4.3 Öznitelikler ve Özellikleri	17
3.4.3.1 Genel öznitelikler	20
3.4.3.2 Kan şekeri öznitelikler	20
3.4.3.3 Zaman Dilimi öznitelikleri-1	21
3.4.3.4 Yemek öznitelikleri	22
3.4.3.5 İnsülin öznitelikleri	22
3.4.3.6 Aktivite öznitelikleri	23
3.4.3.7 Tanımlanamayan kan şekeri öznitelikleri	23
3.4.3.8 Zaman Dilimi öznitelikleri-2	24
3.4.4 Öznitelikler Dosya Yapısı	26

3.5 Veri Madenciliği Yöntemleri.....	26
3.5.1 Sınıflama ve Regresyon	27
3.5.1.1 Destek vektör makineleri (SVM).....	28
3.5.1.2 Karar ağaçları öğrenmesi (Decision Tree Learning)	29
3.5.1.3 K-en yakın komşu (KNN)	29
3.5.1.4 Yapay sinir ağları (YSA)	30
3.5.2 Kümeleme	31
3.5.3 Birliktelik Kuralları	32
3.6 Veri Madenciliği Programları.....	32
3.7 WEKA Programı ve Eğitim-Test Dosyaları	32
3.8 Eğitim ve Test Dosyalarının Oluşturulması	33
DÖRDÜNCÜ BÖLÜM	37
4. DENEYSEL ÇALIŞMALAR.....	37
4.1 Kullanılan Algoritmalar	37
BEŞİNCİ BÖLÜM	41
5. DENEY SONUÇLARI.....	41
5.1 Test-1 Deneyleri ve Sonuçları	41
5.2 Test-2 Deneyleri ve Sonuçları	51
5.3 Test-3 Deneyleri ve Sonuçları	59
ALTINCI BÖLÜM	67
6. SONUÇLAR ve ÖNERİLER	67
KAYNAKLAR	68
EKLER	72
ÖZGEÇMİŞ	87

TABLolar

Tablo 3.1	Hasta Dosyasındaki Deęerlerin Adları ve Kod Numaraları	9
Tablo 3.2	İnsülin Çeşitleri ve Etki Süreleri	10
Tablo 3.3	Değişkenler Arasındaki İlişki Çeşitleri	27
Tablo 4.1	Korelasyon Katsayısı Aralığı ve İlişki Gücü	39
Tablo 4.2	Algoritmaların Adları ve Kısaltmaları	40
Tablo 5.1	Data-1 ile Eğitilen Tüm Algoritmaların Maksimum Korelasyon Deęerleri.....	41
Tablo 5.2	Hasta' 5 in Data-1 Verisi ile Eğitilen DS Algoritmasının Kan Şekeri Ölçüm ve Tahmin Deęerleri	42
Tablo 5.3	Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritmasının Ölçüm ve Tahmin Deęerleri	43
Tablo 5.4	Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritmasının Ölçüm ve Tahmin Deęerleri	44
Tablo 5.5	Hasta5' in 02.09.1990 Tarihli Verileri	49
Tablo 5.6	Hasta5' in 06.09.1990 Tarihli Verileri	50
Tablo 5.7	Farklı Boyutlarda Veriler ile Eğitilen M5P Algoritmasının En Yüksek Korelasyon Deęerleri	51
Tablo 5.8	Farklı Boyutlarda Veriler ile Eğitilen DS Algoritmasının En Yüksek Korelasyon Deęerleri	52
Tablo 5.9	Farklı Boyutlarda Veriler ile Eğitilen RT Algoritmasının En Yüksek Korelasyon Deęerleri	52
Tablo 5.10	Hasta5' in Eğitim Verisi Parçalarının Veri Miktarı ve M5P Algoritması Korelasyon Deęerleri	53

Tablo 5.11	Hasta5' in 10.09.1990 Tarihli Verileri.....	54
Tablo 5.12	Hasta5' in Data-1 ve Data-2 Verileri ile Eğitilen M5P Algoritmasının Kan Şekeri Tahmin Değerleri.	56
Tablo 5.13	Algoritmaların Data-1 ile Eğitimleri Sonucundaki Korelasyon Değerleri.....	57
Tablo 5.14	Algoritmaların Data-5 ile Eğitimleri Sonucundaki Korelasyon Değerleri.....	58
Tablo 5.15	Test-3 Deneylerindeki Algoritmaların En Yüksek Korelasyon Değerleri.....	59

ŞEKİLLER

Şekil 3.1	Veri Dosyalarının Yapısı	8
Şekil 3.2	İnsülin Verisi Çeşitleri.....	9
Şekil 3.3	Kan Şekeri Ölçüm Verisi Çeşitleri	11
Şekil 3.4	Yemek Verisi Çeşitleri	12
Şekil 3.5	Aktivite Verisi Çeşitleri.....	12
Şekil 3.6	Tüm Hastaların Veri Sayıları ve Gözlem Süreleri.....	13
Şekil 3.7	Veri Madenciliği Süreci Akış Diyagramı.....	14
Şekil 3.8	Zaman Damgası Hesaplama Formülü	15
Şekil 3.9	Veritabanı Yapısı	15
Şekil 3.10	Bir Günün Zaman Dilimleri.....	16
Şekil 3.11	Öznelik Grupları ve Özellikleri	18
Şekil 3.12	Öznelik Dosya Yapısı.....	26
Şekil 3.13	SVM Örnek Gösterim.....	28
Şekil 3.14	Karar Ağacı Örnek Gösterim.....	29
Şekil 3.15	KNN Örnek Gösterim.....	29
Şekil 3.16	Yapay Sinir Ağı Hücresini Oluşturan Elemanlar	30
Şekil 3.17	Yapay Sinir Ağında Kullanılan Aktivasyon Fonksiyonları.....	31
Şekil 3.18	Arff Dosya Yapısı.....	33
Şekil 3.19	Eğitim Verilerinin Parçalara Ayrılması	34

Şekil 3.20	Test Dosyaları ve Farklı Boyutlardaki Train Dosyaları Oluşumu Akış Diyagramı	35
Şekil 5.1	Hasta5' in Data-1 Verisi ile Eğitilen DS Algoritmasının Kan Şekeri Ölçüm ve Tahmin Grafiği.....	43
Şekil 5.2	Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritmasının Kan Şekeri Ölçüm ve Tahmin Grafiği.....	44
Şekil 5.3	Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritmasının Kan Şekeri Ölçüm ve Tahmin Grafiği.....	45
Şekil 5.4	Hasta5' in Data-1 Verisi ile Eğitilen DS Algoritması Tahminlerinin Karar Ağacı Şekli	46
Şekil 5.5	Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritması Tahminlerinin Karar Ağacı Şekli	47
Şekil 5.6	Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritması Tahminlerinin Lineer Modelleri	47
Şekil 5.7	Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritması Tahminlerinin Karar Ağacı Şekli	48
Şekil 5.8	Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritması Tahminlerinin Karar Ağacı.....	48
Şekil 5.9	Test3 Deneylerindeki Hasta40' ın RT Algoritması Karar Ağacı	61
Şekil 5.10	Test-3 Deneylerindeki Hasta5' in LR Modeli	63
Şekil 5.11	Test-3 Deneylerindeki Hasta5' in M5P Algoritması Modeli	64

KISALTMALAR

AR	: AdditiveRegression [1]
DS	: DecisionStump [2]
GP	: GaussianProcesses [3]
GPL	: General Public License (Genel Kamu Lisansı)
KNN	: K-Near Neighbours (K-En Yakın Komşu)
LR	: LinearRegression [4]
MP	: MultilayerPerceptron [5]
M5P	: M5P [6]
NPH	: Neutral Protamine Hagedorn [7]
RT	: RepTree [8]
SMOR	: SMOReg [9]
SVM	: Support Vektor Machine (Destek Vektör Makineleri)
UCI	: Univercity of California, Irvine
WEKA	: Waikato Environment for Knowledge Analysis
YSA	: Yapay Sinir Ağları

ÖZET

DİYABET AKILLI DESTEK SİSTEMİ: Bilgi Erişim Teknikleri ile Kan Şekeri Seviyesinin Tahmini

KONDILOĞLU, Adil

Yüksek Lisans, Elektrik ve Bilgisayar Anabilim Dalı

Tez Danışmanı: Yrd. Doç. Dr. Tayfun KÜÇÜKYILMAZ

Mart-2015, 87 sayfa

Diyabet hastalığı tedavisi olmayan ve hayat boyu kontrol altında tutulması gereken bir hastalıktır. Bu tez çalışmasında diyabet hastalarının kan şekeri değerlerinin tahmin edilmesi ve hasta sağlığı üzerinde olumsuz bir durum oluşmaması için hastaların önceden bilinçlendirilmesine katkıda bulunmak amaçlanmıştır.

Eğitim ve tahmin işlemleri diyabet hastalarının kullandıkları insülin dozu değerleri, yemek porsiyon değerleri, aktivite değerleri ve kan şekeri ölçüm değerleri kullanılarak yapılmıştır. Bu amaçla her hastanın kan şekeri değerinin doğru tahminini ve tam olarak ifade edilmesini sağlayacak öznelik (feature) veri setleri oluşturulmuştur. Oluşturulan veri setleri kullanılarak bir veri madenciliği modeli oluşturulmuştur. Kan şekeri değeri tahmini için uygun olan makine öğrenme metotları seçilmiş ve yapılan üç test ile metotlar denenmiştir. Tüm hastalar için test verileri ve farklı boyutlarda eğitim verileri oluşturulmuştur. İlk testte, tüm metotlar her hastanın en büyük boyutlu eğitim verisi ile eğitilerek kan şekeri tahminleri yapılmıştır. Böylelikle tahminin kişiselliği ve metotlar arasındaki farklar incelenmiştir. İkinci testte, tüm hastaların farklı boyutlardaki eğitim verileri ile tüm metotlar eğitilerek kan şekeri tahmini yapılmıştır. Böylelikle tahminin kişiselliği, eğitim verisi boyutunun tahmin üzerindeki etkisi ve metot sonuçları incelenmiştir.

Üçüncü testte, her hastanın verisi test verisi olarak kullanılmış ve diğer hastaların verileri birleştirilerek metotlar eğitilmiştir. Böylelikle kan şekeri tahmininin kişisellikten çıkarılarak tüm hastalar için genelleme yapıp yapılamayacağı incelenmiştir.

Bu çalışmalar sonucunda diyabet hastalarının kan şekeri tahminlerinin genellenebilir özeliğinin olduğu fakat başarı oranının yeterince yüksek olmadığı, hastaya özel tahminlerde ise başarı oranının daha yüksek olduğu görülmüştür. Ayrıca metot sonuçları karşılaştırılmış, hastaların kan şekeri değerlerini öngörebilmek için hangi özniteliklerin etkili olduğu saptanmış, dolayısı ile kan şekeri tahmininde insülin dozu, yemek değeri, aktivite değeri ve kan şekeri ölçüm değerlerinin etkinliği belirlenmiştir.

Anahtar Kelimeler : Diyabet, kan şekeri, kan şekeri değeri, kan şekeri değeri tahmini, veri madenciliği, makine öğrenmesi.

ABSTRACT

DIABETES INTELLIGENT SUPPORT SYSTEM: Prediction of Blood Sugar Levels Using Information Retrieval Techniques

KONDILOĞLU, Adil

Master, Department of Elektrical and Computer

Thesis Supervisor: Yrd. Doç. Dr. Tayfun, KÜÇÜKYILMAZ

March–2015, 87 page

Diabetes is an incurable disease and should be kept under control through a patient's life. The aim of this work is to raise patient's awareness by estimating the blood sugar levels in order to avoid miscalculations/misinterpretations of blood sugar measurements due to human error.

In order to make a prediction on the patient's blood sugar levels, the medication levels, food consumption and post blood sugar measurements of several patients are used as the dataset. An extensive set of features is extracted from this dataset in order to create a blood sugar level corpus, which is used to train a machine learning model for predicting the future blood sugar measurements. These models are then tested on the same patient set, for evaluating the performance of the proposed technique.

In order to evaluate the effectiveness of the proposed technique, three test settings are constructed. In the first test, the patient's timeline is divided into two: past and future. The past data for each patient is used to train a separate machine learning model for each patient; and then the established model is evaluated using the future blood sugar measurements of the same patient. The aim of the first set is threefold: To establish whether there are individual factors that dictate a person's blood sugar levels or not, to establish what those factors are if they exist, and to create a baseline for the upcoming tests. In the second test, we limit the size of the training set by leaving out the oldest measurements in the corpus. The aim of this test is to determine the shortest training time where our technique can be useful.

In the third test, the training sets for the whole patient database is combined in order to create a single large training data for training a single machine learning model which will be used to predict each patient's blood sugar levels. The aim of this last test is to determine whether the factors that define the blood sugar levels are in fact personal or not.

The results of our studies show that it is possible to find and use a generalized model for predicting blood sugar levels of the diabetes patients; though the prediction rates are not high enough for practical use. However, our tests prove that the blood sugar level prediction rates with personalized models are rather high. In this work, a wide range of new contributions is also provided such as comparison of different machine learning algorithms with respect to their prediction quality and analysis on a large range of attributes for predicting blood sugar levels such as insulin dose, food consumption and activity rates.

Keywords: diabetes, blood sugar, blood sugar level, blood sugar level estimation, data mining, machine learning.

BİRİNCİ BÖLÜM

GİRİŞ

Diyabet, pankreas adlı salgı bezinin yeterli miktarda insülin hormonu üretmemesi ya da ürettiği insülin hormonunun etkili bir şekilde kullanılamaması nedeni ile oluşan ve ömür boyu süren bir hastalıktır. Sonuç olarak kişi, yediği besinlerden kana geçen şekeri yani glikozu verimli olarak kullanamaz ve kan şekeri yükselir [10]. Kan şekerinin düzenli olarak regüle edilememesi birçok soruna neden olabilir [11] [12]. Ayrıca kan şekerinin düşüklüğü de vücut fonksiyonlarının sağlıklı yerine getirilememesine sebep olmaktadır. Yani diyabet tedavisinde amaç; kandaki şekerin istenilen düzeyde kalmasını sağlamaktır. Bilindiği gibi diyabet tedavisi bulunan bir hastalık değildir [13]. Ortaya çıkması durumunda ömür boyu takip edilmesi ve kontrol altında tutulması zorunludur.

Kan şekeri oluşumunda yenilen yemeklerin glikoz değerleri direkt olarak etkili olan etkidir. Yemeklerdeki glikozun kana geçmesi ve kandan hücrelere geçmesine kadar olan zaman içindeki sindirim, dolaşım, insülin miktarı ve vücut hücreleri de kan şekerini etkileyen biyolojik etkenlerdir. Bu biyolojik etkenler kişiden kişiye farklılık göstermekle birlikte kişinin farklı durumlarında da farklılık göstermektedir. Mesela kişi heyecanlandığında, korktuğunda ya da dolaşım hızını değiştiren bir aktivite yaptığında fizyolojik etkenler de değişmekte ve kan şekerini etkilemektedir. Bu değişim, dolaşım ve sindirim hızını etkilemekte, kandaki insülin miktarı, insülin oluşumu ve insülinin etkisi de değişmektedir. Biyolojik etkenlerin kişiden kişiye farklılık göstermesinin yanında kişide meydana gelen bu tür değişiklikler de kan şekeri değerinin tahmin edilmesini çok zor bir problem haline getirmektedir.

Bu çalışmanın özünde diyabet hastalarının kan şekeri değerlerinin tahmini için en uygun veri madenciliği algoritmaları, algoritmaların tahminlerinde eğitim verisi boyutunun etkisi ve tahminin kişiselliği ile genellenebilirliği araştırılmaktadır.

Çalışmamız genel olarak beş kısımdan oluşmaktadır. Birinci kısımda diyabet hastaları ve kan şekeri tahmini ile ilgili literatür taraması yapılmış, bu çalışmaların genel olarak amacı açıklanmıştır. İkinci kısımda çalışmamızda kullandığımız materyaller ve kullanılan yöntemler anlatılmakta, deneysel çalışmalarda verilerimizin kullanımı ile ilgili bilgiler verilmektedir. Üçüncü kısımda çalışmamızda yaptığımız deneysel çalışmalar ve sonuçları anlatılmakta, deneysel çalışmalardan çıkarımlar yapılmaktadır. Dördüncü kısımda çalışmamız ile ulaşılan sonuçlar, karşılaşılan zorluklar ve öneriler anlatılmaktadır.

İKİNCİ BÖLÜM

LİTERATÜR ÖZETİ

Diyabet hastalığı tedavisi bilinmediğinden ve dünya üzerinde bu hastalık ile yaşamaya çalışan milyonlarca insanın bulunması [14] [15] dolayısı ile üzerinde birçok araştırma yapılan hastalıktır. Bu araştırmalar incelendiğinde genel olarak üç başlık altında toplamak mümkündür. Bunlar; sağlık alanında yapılan araştırmalar, kan şekeri ölçüm-raporlama araştırmaları ve kan şekeri tahmin araştırmaları olarak adlandırılabilir. Diyabet üzerinde yapılan çalışmaların son derece fazla olmasından dolayı bu araştırmada sadece yaptığımız çalışma ile direkt olarak ilgili olan ve literatürde belli başlı önem arz eden çalışmaları sunacağız.

2.1.Tıbbi Çalışmalar

Bu çalışmalar genel olarak diyabet hastası kişilerin tedavi süreçleri, diyabet hastalığının diğer hastalıklarla ilişkisi, diyabet hastası kişilerin kullandıkları ilaçlar ve dozları gibi hastalığın tıbbi boyutu ile ilgili yapılan çalışmalardır. Bu alanda yapılan çalışmalardan bir tanesinde [16] kan şekeri kontrol altında olan ve olmayan hastaların psikiyatrik belirtilerinin kan şekeri kontrolü ile ilişkisi araştırılmıştır. 42 hasta üzerinde yapılan araştırmada kan şekeri kontrolü olan hastalardan 13' ünde ve kan şekeri kontrol altında olmayan hastalardan da 5' inde psikososyal stres tespit edilmiştir. Yapılan bir diğer çalışmada [17] besinlerin kan şekeri üzerindeki etkileri araştırılmıştır. Besinlerin tüketildikten sonraki kan şekerini artırıcı etkileri göz önüne alınarak pirinç yerine bulgurun tercih edilmesi gerektiği, beyaz ekmek yerine tam buğday ekmeğinin tercih edilmesi gerektiği, meyve suları yerine meyvelerin kendilerinin tüketilmesinin gerektiği gibi bilgiler verilmektedir.

2.2. Raporlama ve İzleme Çalışmaları

Bu çalışmalar ile diyabet hastalarının kan şekeri ölçüm sonuçlarının kaydedilmesi ve kan şekeri değişiminin izlenmesi amacı ile yapılan raporlama çalışmalarıdır. Bu çalışmalar kan şekeri ölçüm cihazları ve bu cihazlar ile birlikte kullanılan kan şekeri takip programlarıdır. Bu programlar bazı kan şekeri ölçüm cihazı üreticileri tarafından cihaz ile birlikte kullanılması için verilmektedir. Apple cihazlar ile kullanılan kablosuz kan şekeri ölçüm ürünü [18] kan şekeri ölçümü yapmakta ve ölçüm sonuçlarını raporlamaktadır. Kan şekeri ölçüm cihazları olan MyGlucoHealt [19] ve One Touch Verio [20] kan şekeri ölçümü yapmakta ve ölçüm sonuçlarını raporlamaktadır.

2.3. Kan Şekeri Tahmin Çalışmaları

Yapılan literatür araştırmaları sonucunda kan şekeri tahmin etmek için yapılan birçok çalışma bulunduğu tespit edilmiştir. Bu çalışmalar kan şekeri tahmin edilebilirliğini farklı yöntemler ile analiz etmektedir. Diyabet izleme, kayıt ve yorumlama sistemi olarak yapılan bir çalışmada [21] hastaların yaşam tarzları ile ilgili veriler toplanmaktadır. Mobil cihazlardaki programlar, internet üzerinden kullanılan programlar ve kan şekeri ölçüm cihazları ile elde edilen hasta verilerine hastanın hastane bilgileri de eklenerek bir veritabanı oluşturulmaktadır. Bu veritabanı üzerinde analiz ve veri madenciliği yöntemleri kullanılarak hastanın metabolizma modeli oluşturulmakta ve bu model ışığında gerekli durumlarda hastaya geri bildirim yapılmaktadır. Bu çalışmada analiz işlemleri belirli yaş aralığında kadın ve erkek hasta verileri kullanılarak yapılmakta, sonuçlar hastaya özgü olarak oluşturulmaktadır.

Bir diğer çalışmada [22] plazma glikoz seviyesi, plazma insülin konsantrasyonu, vücut ağırlığı başına plazma kan şekeri dağılım hacmi gibi bilgiler kullanılarak oluşturulan matematiksel denklemler ile yemek sonrasında oluşacak kan şekerinin tepe değerinin tahmin çalışması yapılmaktadır.

Yakın zamanda yapılmış olan başka bir çalışmada [23] diyabet hastalarının kan şekeri ölçüm verileri kullanılarak hastalarda hipoglisemi (kan şekeri yüksekliği) tahmini çalışması yapılmıştır. Bu çalışmada diyabet hastalarının kan şekeri değerlerini tahmin etmek için J4.8, RepTree karar ağacı algoritmaları ve torbalama (bagging) yöntemi kullanılmıştır. Her hasta için kullanılan yöntem ve algoritmaların duyarlılık,

özgüllük, Youden indeks ve doğruluk değerleri incelenmiştir. Son iki kan şekeri ölçümü ve bu ölçümler arasındaki fark kullanılarak tahmin için model oluşturulmuştur.

Başka bir çalışmada [24] yazarlar diyabet hastalarının kan şekeri kontrolüne yardımcı olmak ve gerektiğinde doktorlar ile diyabet hastalarının bilgilerinin paylaşımını sağlayacak bir sistem tasarımı yapmıştır. Bu sistem ile diyabet hastalarının; yedikleri yemek görüntüleri, aldıkları insülin miktarları, enerji tüketimi ve kan şekeri ölçüm sonuçları toplanarak hasta ile ilgili bir veri bankası oluşturulmaktadır. Bu veri bankası kullanılarak hastaların yemek görüntülerinden yemeğin toplam kalorisinin ve besin değerlerinin tespiti yapılmakta, kan şekeri ölçüm sonuçları ve diğer veriler de incelenerek kısa bir sürede geri dönüş yapılmaktadır.

Kan şekeri tahmini üzerine yapılan bir diğer çalışmada [25] kan şekeri tahmini için yemek ve insülin değerleri kullanılarak kan şekeri değeri tahmini yapılabilecek matematiksel model oluşturulmuştur. Bu model: yenilen karbonhidrat kütleleri, gıda glisemik indeksi, vücut kitle indeksi, yiyeceğin tüketim zamanı, insülin tüketim zamanı, kullanılan insülin dozu gibi değerler kullanılarak kan şekeri tahmininin yapıldığı matematiksel fonksiyondur.

Tip iki diyabet hastaları üzerinde yapılan bir çalışmada [26], cep telefonu üzerinde çalışan bir yazılım ile kan şekeri ölçüm sonuçları ve egzersiz bilgileri kullanan bir algoritma ile kan şekeri sonuçlarını tahmin edilmeye çalışılmıştır.

Bizim çalışmamızın literatürdeki diğer çalışmalardan ayıran özellikleri bulunmaktadır. Bu özelliklerden ilki; çalışmamızda kan şekeri tahmini yalnızca kişisel olarak yapılmamakta ayrıca kan şekeri tahmini genelleme çalışması yapılmaktadır. Her hasta için hasta verilerinden bağımsız genel bir eğitim veri seti kullanılarak algoritmalar eğitilmekte ve hastaların kan şekeri değerleri tahmin edilmektedir. Böylelikle kan şekeri tahmini için genel bir model olup olamayacağı araştırılmaktadır. İkinci farklı özellik olarak kan şekeri tahmini için kullandığımız veri setimiz birçok özellik içermektedir. Veri setimiz ile yaptığımız tahmin çalışmaları neticesinde bu özelliklerin kan şekeri üzerindeki etkisi araştırılmaktadır. Böylelikle kan şekeri üzerinde etki yapan öncelikli özelliklerin tespiti yapılmaktadır. Üçüncü olarak, çalışmamızda her hasta verisi belirlenen 9/10 en yüksek oranında parçaya ayrılarak en büyük boyuttaki eğitim verisi oluşturulmakta, algoritmalar eğitilmekte ve test işlemleri yapılmaktadır. Böylelikle en büyük boyutta eğitim verisi ile eğitilen algoritmaların tahminleri karşılaştırılmaktadır. Dördüncü olarak, her hastanın eğitim verisi farklı boyutlarda parçalara ayrılarak algoritmalar eğitilmekte ve test işlemleri yapılmaktadır.

Böylelikle eğitim verisi boyutunun algoritma tahminleri üzerindeki etkisi incelenmekte ve farklı algoritmaların aynı eğitim verisi sonucundaki tahminleri karşılaştırılmaktadır. Ayrıca birçok farklı özneliğin birleştirilmesi sonucunda oluşturulacak bir modeli kullanmayı planlayan araştırmamız, kullanılan tekniğin literatürde bir ilk olması nedeni ile de önem teşkil etmektedir. Literatürde kan şekerinin birçok farklı faktöre bağlı olduğu bilinmektedir [27] [17]. Çalışmamızın başarılı olması, kan şekeri tahmini üzerinde büyük katkılara yol açacaktır.

ÜÇÜNCÜ BÖLÜM

MATERYAL VE YÖNTEM

Bu çalışmada amaç; diyabet hastalarının insülin, yemek, aktivite ve kan şekeri değerleri kullanılarak hastaların oluşacak kan şekeri değerlerini tahmin edilmesine çalışılmaktadır. Diyabet hastalarının kullandıkları; insülin dozu ve özelliği, yemek miktarı, aktivite miktarı ve kan şekeri ölçüm sonuçları kullanılarak her hasta için öznitelikler oluşturulmaktadır. Kan şekeri tahmin çalışmaları üç test aşamasında yapılmaktadır.

Test-1 çalışmalarında tüm hastaların en büyük boyuttaki eğitim verileri ve test verileri oluşturulmaktadır. Her hasta için bireysel eğitim ve test işlemleri yapılarak tahmin çalışmaları yapılmaktadır. Böylelikle kan şekeri tahminin bireysel olarak tahmin edilebilirliği araştırılmakta ve kişisel tahmin için belirleyici nitelikler tespit edilmektedir.

Test-2 çalışmalarında tüm hastaların verilerinden test verisi ve farklı boyutlarda eğitim verileri oluşturulmaktadır. Oluşturulan farklı boyutlardaki eğitim verileri ile ayrı ayrı eğitilen algoritmaların yaptıkları tahminler algoritmalar ve veri boyutları olarak karşılaştırılmaktadır. Böylelikle hangi algoritmanın en yüksek tahmin değerini hangi veri boyutu ile elde ettiği ve tahmin için eğitim verisinin boyutunun etkisi araştırılmaktadır.

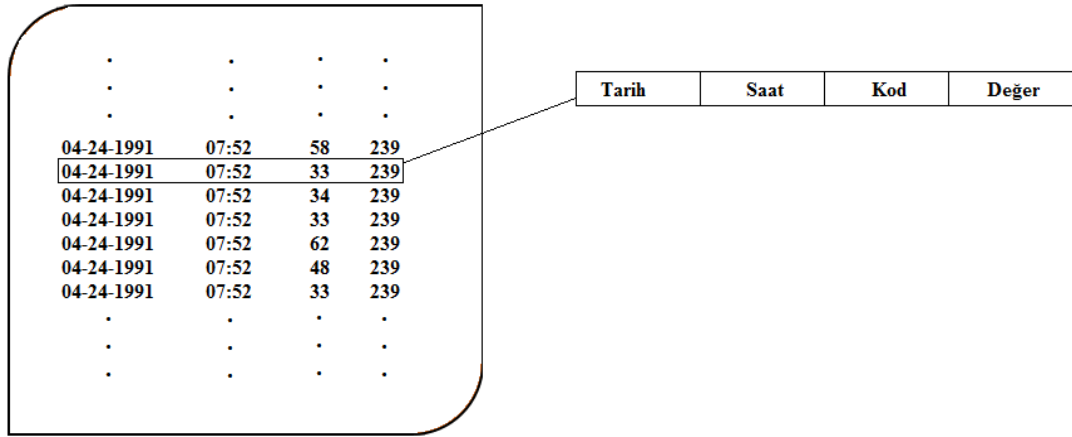
Test-3 çalışmalarında her hastanın verileri test verisi olarak kullanılmakta, diğer hastaların verileri birleştirilerek algoritmalar eğitilmektedir. Her hasta için yapılan bu işlemler sonucunda kan şekeri tahmini için hastadan bağımsız eğitim verisinin etkisi incelenmekte ve algoritmalar karşılaştırılmaktadır.

3.1. Veri Seti Yapısı ve Özellikleri

Çalışmada kullanılan veri seti [28] UCI veri deposundan [7] alınmıştır. Bu veri setinde 70 diyabet hastasının verileri bulunmaktadır. Bu veriler üzerinde öncelikli işlemler (preprocessing) yapılarak eğitim ve test işlemleri için uygun hale getirildikten sonra çalışmalarımızda kullanılmaktadır.

3.2. Veri Dosyaları Yapısı

Veri seti dosyaları hastaların; kan şekeri ölçümü, kullanılan insülin dozu ve türü, yemek değeri ve aktivite bilgilerinden oluşmaktadır. Bu bilgiler için belirli bir kod değerleri kullanılarak veri dosyası oluşturulmuştur.



Tarih	Saat	Kod	Değer
.	.	.	.
.	.	.	.
.	.	.	.
04-24-1991	07:52	58	239
04-24-1991	07:52	33	239
04-24-1991	07:52	34	239
04-24-1991	07:52	33	239
04-24-1991	07:52	62	239
04-24-1991	07:52	48	239
04-24-1991	07:52	33	239
.	.	.	.
.	.	.	.
.	.	.	.

Şekil 3.1: Veri Dosyalarının Yapısı.

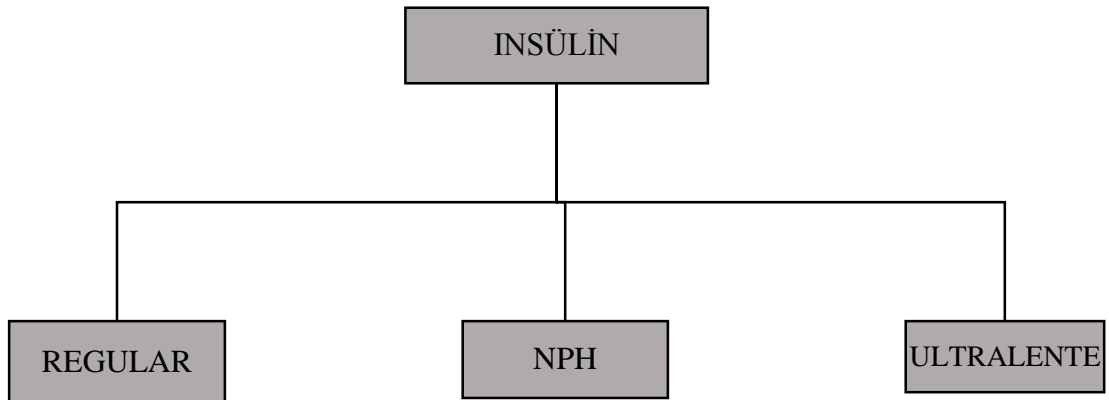
Şekil 3.1’ de veri dosyalarının yapısı gösterilmektedir. Her bir veri Tarih, Saat ve Kod olarak üç değer ile birlikte verilmektedir. Birinci değer verinin tarihidir. Tarih formatı olarak MM-DD-YYYY kullanılmıştır. İkinci değer verinin saatidir. Saat bilgisi HH:MM olarak 24 saat formatında verilmektedir. Üçüncü değer olarak da verinin hangi bilgi olduğunu belirten kod değeri olarak verilmektedir. Tablo 3.1’ de kod değerleri ve bu değerlerin belirttiği özellikler verilmektedir.

Tablo 3.1: Hasta Dosyasındaki Değerlerin Adları ve Kod Numaraları [7].

KOD	DEĞER ADI
33	Regular insulin dose (Normal insülin dozu)
34	NPH insulin dose (NPH insülin dozu)
35	UltraLente insulin dose (Ultralente insülin dozu)
48	Unspecified blood glucose measurement (Belirtilmemiş kan şekeri ölçümü)
57	Unspecified blood glucose measurement (Belirtilmemiş kan şekeri ölçümü)
58	Pre-breakfast blood glucose measurement (Kahvaltı öncesi kan şekeri ölçümü)
59	Post-breakfast blood glucose measurement (Kahvaltı sonrası kan şekeri ölçümü)
60	Pre-lunch blood glucose measurement (Öğle yemeği öncesi kan şekeri ölçümü)
61	Post-lunch blood glucose measurement (Öğle yemeği sonrası kan şekeri ölçümü)
62	Pre-supper blood glucose measurement (Akşam yemeği öncesi kan şekeri ölçümü)
63	Post-supper blood glucose measurement (Akşam yemeği sonrası kan şekeri ölçümü)
64	Pre-snack blood glucose measurement (Aperatif öncesikan şekeri ölçümü)
65	Hypoglycemic symptoms (Hipoglisemik semptomlar)
66	Typical meal ingestion (Normal yemek yeme)
67	More-than-usual meal ingestion (Normalden daha fazla yemek yeme)
68	Less-than-usual meal ingestion (Normalden daha az yemek yeme)
69	Typical exercise activity (Normal egzersiz etkinliği)
70	More-than-usual exercise activity (Normalden daha fazla egzersiz etkinliği)
71	Less-than-usual exercise activity (Normalden daha az egzersiz etkinliği)
72	Unspecified special event (Belirtilmemiş özel etkinlik)

Veri dosyası; insülin verileri, glikoz verileri, hipoglisemi verileri, yemek verileri, aktivite verileri ve belirtilmemiş veriler olarak gruplanmaktadır.

3.2.1. İnsülin Verisi Özellikleri



Şekil 3.2: İnsülin Verisi Çeşitleri [7].

İnsülin verinin çeşitleri Şekil 3.2’ de gösterilmektedir. İnsülin verileri etkisi bakımından üçe ayrılmaktadır. Bunlar Regular, NPH ve Ultralente insülin verileridir [7]. Regular insülin kısa etki düzeyli insülin, NPH ortalama etki düzeyli insülin ve Ultralente insülin de uzun etki düzeyli insülinidir. İnsülinlerin etki süreleri ile ilgili bilgi veri setinde [7] belirtilmiştir. Tablo 3.2’ de de görüleceği üzere her bir insülin çeşidinin etki başlangıç, tepe değer ve etki süresi olmak üzere üç farklı değeri vardır.

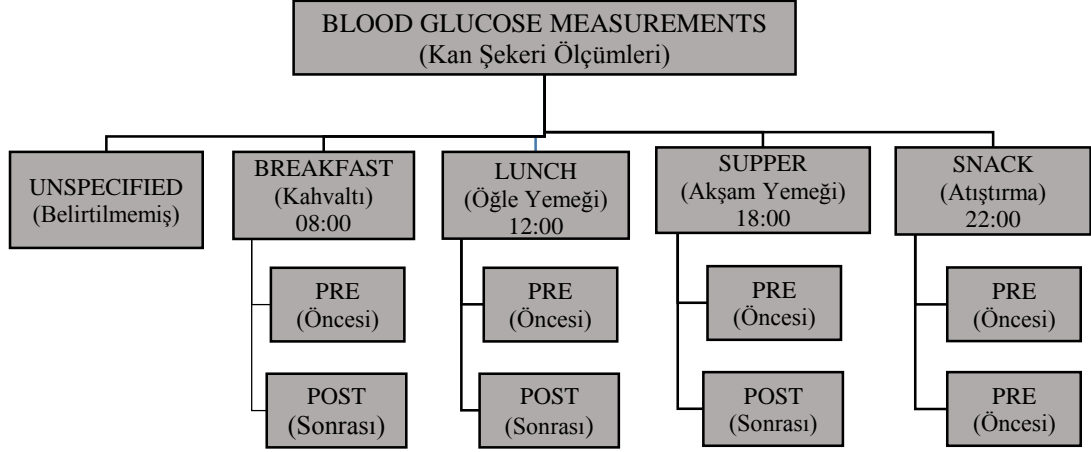
Tablo 3.2: İnsülin Çeşitleri ve Etki Süreleri [7].

	Etki Başlangıç Süresi	Tepe Değeri Süresi	Etki Süresi
Regular İnsülin	15-45 dakika	1-3 saat	4-6 saat
NPH İnsülin	1-3 saat	4-6 saat	10-14 saat
Ultralente İnsülin	2-5 saat	Tepe süresi yok	24-30 saat

Etki başlangıç değeri alınan insülinin kan plazmasına geçtikten sonra etkisinin başlangıcı için geçen zamanı göstermektedir. Tepe değeri ise alınan insülinin kan plazmasına geçtikten sonra maksimum etkiyi göstereceği tepe değere ulaşması için geçen süreyi göstermektedir. Etki süresi ise alınan insülinin kan plazmasına geçtikten sonra etkisini göstereceği süreyi göstermektedir. Öznitelikler oluşturulurken insülin çeşitlerinin etki süreleri dikkate alınarak oluşturulmaktadır. Kullanılan veri setinde insülin değerleri hastaya uygulanan doz olarak verilmekte ve çalışmamızda bu şekilde kullanılmaktadır.

3.2.2. Kan Şekeri Verisi Özellikleri

UCI veri setinde [7] kahvaltı saati 08:00, öğle yemeği saati 12:00, akşam yemeği saati 18:00 ve yatma zamanı da saat 22:00 olarak verilmektedir. Bu verileri altı başlık altında toplamak mümkündür. Kan şekeri ölçüm verisinin çeşitleri Şekil 3.3’ de gösterilmektedir.



Şekil 3.3: Kan Şekeri Ölçüm Verisi Çeşitleri [7].

Birinci grup; Unspecified blood glucose measurement (Belirtilmemiş kan şekeri ölçümleri) dir. Bu ölçüm sonuçları çeşitli nedenlerden dolayı belirli bir gruba dahil olamayan ölçümlerdir. Örneğin breakfast, lunch, supper ve snack ölçüm saatlerinde olmayan, genel olarak saat 23:00’ dan sonra ya da 00:00’ dan sonra yapılan ölçümler bu gruba dahil edilmektedir.

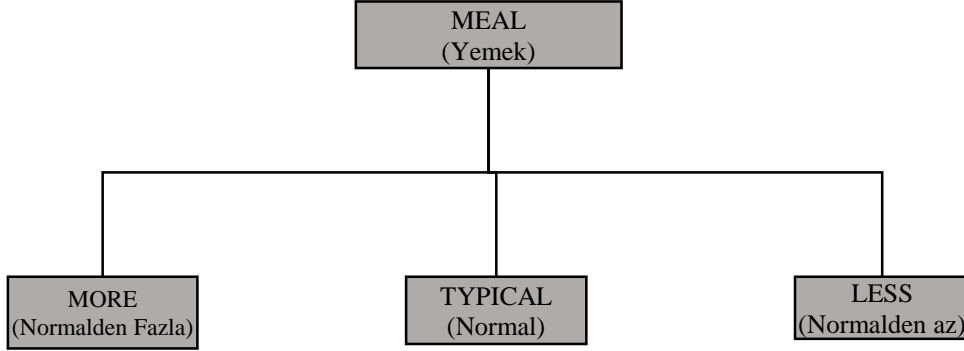
İkinci grup; Breakfast (Kahvaltı) ölçümleridir. Bu ölçümler pre-breakfast blood glucose measurement (kahvaltı öncesi kan şekeri ölçümü) ve post-breakfast blood glucose measurement (kahvaltı sonrası kan şekeri ölçümü) olarak ikiye ayrılmaktadır. Öncesinde ölçümler yemek saatleri öncesinde yapılan ölçümler olduğundan dolayı bir önceki yemeğin kan şekeri üzerindeki etkisinin düşmeye başladığı zamana denk gelmektedir. Sonrasında ölçümler ise yemekten sonra olduğundan dolayı kan şekerinin yükselme eğiliminde olduğu zamana denk gelmektedir.

Üçüncü grup; Lunch (öğle yemeği) ölçümleridir. Bu ölçümler pre-lunch blood glucose measurement (öğle yemeği öncesi kan şekeri ölçümü) ve post-lunch blood glucose measurement (öğle yemeği sonrası kan şekeri ölçümü) olarak ikiye ayrılmaktadır.

Dördüncü grup; Supper (akşam yemeği) ölçümleridir. Bu ölçümler pre-supper blood glucose measurement (akşam yemeği öncesi kan şekeri ölçümü) ve post-supper blood glucose measurement (akşam yemeği sonrası kan şekeri ölçümü) olarak ikiye ayrılmaktadır.

Beşinci grup pre-snack blood glucose measurement (atıştırma öncesi kan şekeri ölçümü) dür. Hasta verileri arasında bu grubun post ölçümleri bulunmamaktadır.

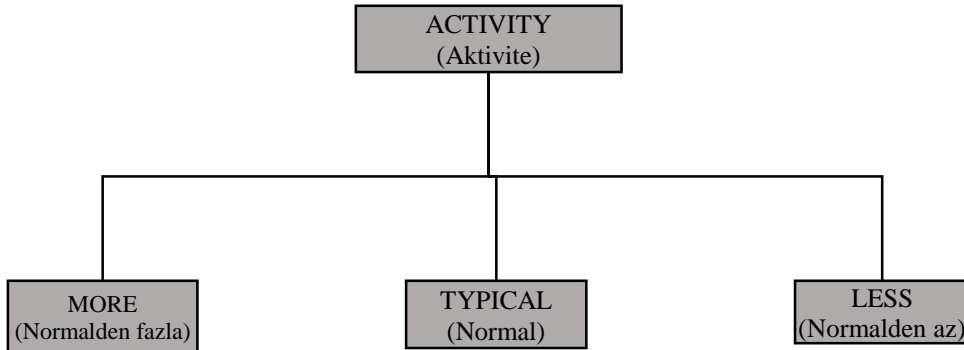
3.2.3. Yemek Verisi Özellikleri



Şekil 3.4: Yemek Verisi Çeşitleri [7].

Yemek verisi Tablo 3.4' de görüldüğü üzere typical meal ingestion (normal yemek yenmesi), more-than-usual meal ingestion (normalden daha fazla yemek yenmesi), less-than-usual meal ingestion (normalden daha az yemek yenmesi) şeklinde üç grupta belirtilmektedir. Kullanılan veri setinde yemek değerleri ile ilgili bir bilgi verilmemektedir. Yemek bilgileri kod değerleri ile birbirinden ayrılmaktadır. Çalışmalarımızda *Less* yemek değeri 1, *Typical* yemek değeri 2 ve *More* yemek değeri 3 olarak kullanılarak işlemler yapılmaktadır.

3.2.4. Aktivite Verisi Özellikleri



Şekil 3.5: Aktivite Verisi Çeşitleri [7].

Aktivite verisi Şekil 3.5’ de görüldüğü üzere typical exercise activity(normal egzersiz etkinliği), more-than-usual exercise activity (normalden daha fazla egzersiz etkinliği), less-than-usual exercise activity (normalden daha az egzersiz etkinliği) şeklinde üç grupta belirtilmektedir. Bu grupların herhangi biri altında belirtilemeyecek olan aktiviteler de Unspecified special event(Belirtilmemiş özel olay) kodu ile verilmektedir. Kullanılan veri setinde egzersiz değerleri ile ilgili bir bilgi verilmemektedir. Egzersiz bilgileri kod değerleri ile birbirinden ayrılmaktadır. Çalışmalarımızda *Less* egzersiz değeri 1, *Typical* egzersiz değeri 2 ve *More* egzersiz değeri 3 olarak kullanılarak işlemler yapılmaktadır.

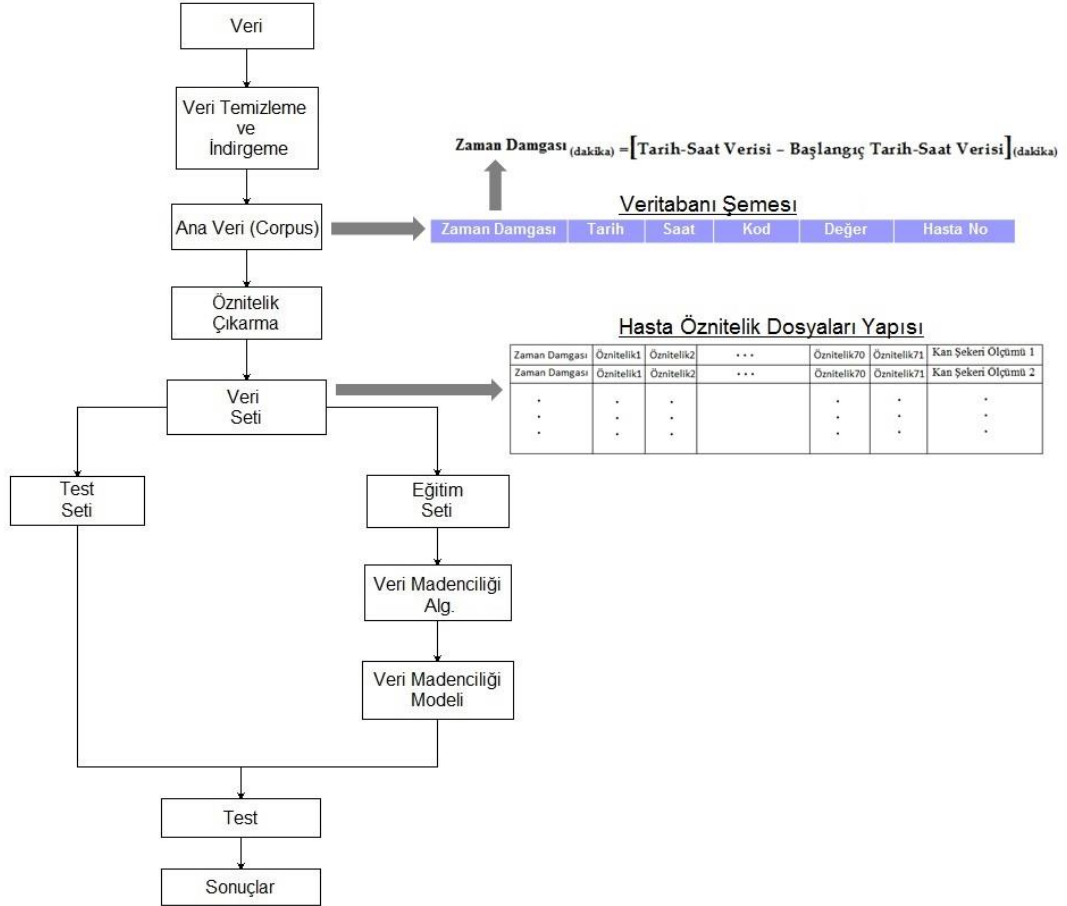
Şekil 3.6’ da veri setinde bulunan tüm hastaların insülin, kan şekeri, yemek, aktivite verileri ve bu veriler oluşturulurken yapılan gözlem süreleri verilmektedir.

Hasta No	Gözlem Süresi (Gün)	İnsülin Verisi Sayısı	Glikoz Verisi Sayısı	Yemek Verisi Sayısı	Aktivite Verisi Sayısı	Hasta No	Gözlem Süresi (Gün)	İnsülin Verisi Sayısı	Glikoz Verisi Sayısı	Yemek Verisi Sayısı	Aktivite Verisi Sayısı
1	136	523	369	0	0	36	31	103	121	15	5
2	96	381	373	0	0	37	35	151	131	6	2
3	39	134	139	10	0	38	30	171	115	3	3
4	39	145	141	6	0	39	31	181	119	0	0
5	41	148	144	2	0	40	8	32	31	33	9
6	25	84	65	0	0	41	63	251	243	0	0
7	35	129	113	0	0	42	66	287	254	0	0
8	37	11	139	8	3	43	31	170	112	1	1
9	31	100	99	3	3	44	31	188	103	0	0
10	34	140	107	0	0	45	29	194	101	0	0
11	28	109	67	25	8	46	35	202	95	0	0
12	42	145	89	23	8	47	33	189	101	0	0
13	34	138	97	21	20	48	35	189	104	0	0
14	38	99	58	38	18	49	46	0	197	6	16
15	54	165	93	19	4	50	61	0	266	1	21
16	57	176	93	20	3	51	40	116	154	10	15
17	27	102	100	0	18	52	48	103	174	11	10
18	33	142	132	0	17	53	44	174	123	0	0
19	32	146	125	1	20	54	103	409	362	8	0
20	135	548	454	0	0	55	149	714	606	3	0
21	64	310	207	0	0	56	125	510	499	9	0
22	45	195	103	0	0	57	28	19	100	5	3
23	50	188	100	5	2	58	30	156	116	14	6
24	47	183	117	0	0	59	31	157	120	15	2
25	31	16	91	1	0	60	31	164	112	15	8
26	79	178	305	0	0	61	33	169	120	6	5
27	119	465	458	0	0	62	34	154	132	10	3
28	122	476	475	0	0	63	35	155	138	3	1
29	159	633	626	0	0	64	24	0	82	0	4
30	148	588	591	0	0	65	137	545	546	25	0
31	85	335	335	0	0	66	20	78	87	70	0
32	14	58	55	15	18	67	120	473	467	9	0
33	30	119	119	17	28	68	166	166	527	0	0
34	30	140	120	9	16	69	14	0	49	0	0
35	30	139	119	13	5	70	61	232	109	0	0

Şekil 3.6: Tüm Hastaların Veri Sayıları ve Gözlem Süreleri.

3.3. Veri Dosyaları İşlemleri ve Veritabanı Yapısı

Çalışmamızda kullandığımız veriler birinci elden elde edilen ham verilerdir. Bu ham veriler bazı işlemlere tabi tutularak işlem yapılacak veriler haline getirilmektedirler.



Şekil 3.7: Veri Madenciliği Süreci Akış Diyagramı.

Şekil 3.7' de çalışmamızın veri madenciliği süreci görülmektedir. Hasta verileri şekilde gösterilen işlem süreçlerinden geçirilerek en son basamakta gösterilen sonuçlara ulaşılmaktadır.

3.3.1. Verileri Temizleme ve İndirgeme

Veri setimiz incelendiğinde bazı hastaların bazı verilerinin istenmeyen şekilde olduğu tespit edilmiştir. Örnek vermek gerekirse; hastaların kan şekeri ölçüm sonuçlarının sayısal değer olması gerekmektedir. Veri seti incelendiğinde bazı hasta

verilerinde kan şekeri ölçüm sonuçlarının ölçülememesi ya da veri dosyası oluşturulurken kaynaklanan hatalar sonucunda bu verilerin sayısal değerler olmadığı tespit edilmiştir. Bu veriler üzerinde yapılacak matematiksel işlemlerin düzgün yapılabilmesi için istenmeyen özellikteki bu veriler veri kümemizden temizlenmiştir.

Verilerimiz üzerinde daha hızlı ve istediğimiz formata uygun işlemler yapabilmemiz için verilerimizin tarih ve saat değerleri üzerinde bir indirgeme işlemi yapılmaktadır. Şekil 3.8’ de Zaman Damgası (Time Stamp) hesaplamak için kullanılan formül gösterilmektedir.

$$\text{Zaman Damgası}_{(\text{dakika})} = \left[\text{Tarih-Saat Verisi} - \text{Başlangıç Tarih-Saat Verisi} \right]_{(\text{dakika})}$$

Şekil 3.8: Zaman Damgası Hesaplama Formülü.

Zaman Damgası değeri oluşturma işleminden önce veri dosyalarının tarih formatı DD.MM.YYYY şekline çevrilmiştir. Bu tarih bilgileri saat bilgileri ile birleştirilerek formüldeki Tarih-Saat verisi oluşturulmuştur. Birleştirilmiş verinin formatı DD.MM.YYYY hh:mm şeklindedir. Bu formatta DD gün, MM ay, YYYY yıl, hh saat ve mm dakikayı belirtmektedir. Bu veri ile Başlangıç Tarih-Saat Verisi yani 00.00.0000 00:00 formatlı veri arasındaki fark dakika olarak alındığında oluşacak sonuç Zaman Damgası değeri olmaktadır. Zaman Damgası değeri her bir veri için ayrı ayrı oluşturulmaktadır.

3.3.2. Ana Veri (Corpus)

Veriler, temizleme ve indirgeme işlemlerinden sonra öznitelik çıkarma işlemlerinin daha rahat yapılabilmesi için belirli bir formatta MySql veri tabanı üzerinde birleştirilerek hasta veritabanı oluşturulmaktadır. Oluşturulan veri tabanı şeması Şekil 3.9’ de gösterilmektedir.

Zaman Damgası	Tarih	Saat	Kod	Değer	Hasta No
---------------	-------	------	-----	-------	----------

Şekil 3.9 : Veritabanı Yapısı.

Her bir hasta dosyamız ve bu dosyaların her bir verisi için oluşturulan Zaman Damgası değerleri ile birlikte Şekil 3.9’ de gösterilen veritabanı yapısı oluşturularak veriler birleştirilmektedir. Veritabanı kullanılarak her hasta için öznitelikler oluşturularak ve öznitelik dosyaları oluşturulmaktadır.

3.4. Öznitelik Oluşturma

Verilerimiz üzerinde veri madenciliği yöntemlerini kullanabilmemiz için verilerin kullanacağımız veri madenciliği algoritmalarına uygun hale getirilmektedir. Bunun için öncelikle her hastanın verisi için öznitelik değerleri hesaplanmakta ve her hasta için ayrı ayrı öznitelik dosyaları oluşturulmaktadır. Bu öznitelik dosyaları kullanılarak her hasta için test ve eğitim dosyaları oluşturulmaktadır. Eğitim dosyaları veri madenciliği algoritmaları kullanılarak eğitilmektedir. Eğitilen algoritmalar test dosyaları ile test edilmekte ve test sonuçları yorumlanmaktadır. Eğitim, test dosyalarının oluşturulması ve algoritmalar ile yapılan testler deneysel çalışmalar bölümünde daha detaylı olarak anlatılmaktadır.

3.4.1. Öznitelikler ve Öznitelik Dosyaları

Kan şekeri değerlerinin tahmin edilebilmesi için her bir kan şekeri ölçüm değeri için bu değeri ifade etmeye yarayan öznitelik değerleri kullanılmaktadır. Bu öznitelik değerlerinin çokluğu ve doğruluğu ile kan şekeri ölçüm değeri daha iyi ifade edilmektedir. Ayrıca bu değerlerin kan şekeri değerleri üzerindeki belirleyici etkileri de incelenerek kan şekeri oluşumundaki belirleyici özellikler ortaya çıkarılmaktadır.

3.4.2. Zaman Dilimi Özellikleri

Diyabet hastalarının kan şekeri değişiminin daha anlaşılır incelenebilmesi için hastanın bir günü eş zaman dilimine ayrılmıştır. Bu zaman dilimleri Şekil 3.10’ de gösterilmektedir.

00:00 – 24:00			
00:00 – 06:00	06:00 – 12:00	12:00 – 18:00	18:00 – 24:00
06:00 – 18:00			

Şekil 3.10: Bir Günün Zaman Dilimleri.

Öznitelik oluřturma iřlemlerinde bu zaman dilimleri de kullanılmaktadır. Bylelikle kan řekeri deęiřimi daha iyi analiz edilmekte ve kan řekeri tahminin yksek doęrulukta yapılmasına alıřılmaktadır.

3.4.3. znitelikler ve zellikleri

Veri dosyaları genel olarak; inslin deęerleri, kan řekeri lmleri, yemek deęerleri ve aktivite deęerlerinden oluřmaktadır. Bu deęerler kendi ierisinde; inslin deęerleri , kan řekeri lm sonuları on, yemek deęerleri  ve aktivite deęerleri  farklı deęere ayrılmaktadır. Veri setinde toplam olarak on dokuz farklı deęer kullanılmaktadır. Tm farklı deęerler farklı zellikler oluřturularak znitelik deęerleri oluřturulmaktadır. znitelik deęerlerini yedi bařlık altında toplamak mmkndr. Bunlar; Genel znitelikler, Kan řekeri znitelikleri, Zaman Dilimi znitelikleri, Yemek znitelikleri, İnslin znitelikleri, Aktivite znitelikleri ve Tanımlanamayan kan řekeri znitelikler olarak gruplanmaktadır. řekil 3.11' da bu znitelik grupları ve zellikleri grlmektedir.

GENEL ÖZİNTELİKLER									
Özınitelik Adı	Özınitelik 1	Özınitelik 2	Özınitelik 3	Özınitelik 4	Özınitelik 5	Özınitelik 6	Özınitelik 7	Özınitelik 8	Özınitelik 9
Özınitelik Açıklaması	Tüm Eski Kan Şekeri Değerleri Ortalaması	Eski Unspecified Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile Aynı Gruptaki(pre-post) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile Aynı Gruptaki(break-fast-lunch-supper-snack-unspecified) Kan Şekeri	Eski Ölçüm ile Tümüyile Aynı Gruptaki Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile Aynı Zaman Dilimi Ortalaması	Eski Ölçüm ile Aynı Zaman Dilimi ve pre-post Ortalaması	Eski Ölçüm ile Aynı Zaman Dilimi ve Grup(breakfast-lunch-supper-snack-unspecified)	Eski Ölçüm ile Herşeyi ile Aynı Gruptaki Kan Şekeri Değerleri Ortalaması
KAN ŞEKERİ ÖZİNTELİKLERİ									
Özınitelik Adı	Özınitelik 10	Özınitelik 11	Özınitelik 12	Özınitelik 13	Özınitelik 14	Özınitelik 15	Özınitelik 16	Özınitelik 17	Özınitelik 18
Özınitelik Açıklaması	Eski pre-breakfast Kan Şekeri Değerleri Ortalaması	Eski pre-lunch Kan Şekeri Değerleri Ortalaması	Eski pre-supper Kan Şekeri Değerleri Ortalaması	Eski pre-snack Kan Şekeri Değerleri Ortalaması	Eski pre-tüm Kan Şekeri Değerleri Ortalaması	Eski post-breakfast Kan Şekeri Değerleri Ortalaması	Eski post-lunch Kan Şekeri Değerleri Ortalaması	Eski post-supper Kan Şekeri Değerleri Ortalaması	Eski post-tüm Kan Şekeri Değerleri Ortalaması
ZAMAN DİLİMİ ÖZİNTELİKLERİ - 1									
Özınitelik Adı	Özınitelik 19	Özınitelik 20	Özınitelik 21	Özınitelik 22	Özınitelik 23				
Özınitelik Açıklaması	Tüm Saat 00:00-06:00 arasındaki KanŞekeri Değerleri Ortalaması	Tüm Saat 06:00-12:00 arasındaki KanŞekeri Değerleri Ortalaması	Tüm Saat 12:00-18:00 arasındaki KanŞekeri Değerleri Ortalaması	Tüm Saat 18:00-24:00 arasındaki KanŞekeri Değerleri Ortalaması	Tüm Saat 06:00-18:00 arasındaki KanŞekeri Değerleri Ortalaması				
YEMEK ÖZİNTELİKLERİ									
Özınitelik Adı	Özınitelik 24	Özınitelik 25	Özınitelik 26	Özınitelik 27	Özınitelik 28	Özınitelik 29	Özınitelik 30	Özınitelik 31	Özınitelik 32
Özınitelik Açıklaması	Son 4 saat içerisindeki Yemek Değerleri Ortalaması	Son 6 saat içerisindeki Yemek Değerleri Ortalaması	Son 1,5 saat içerisindeki Yemek Değerleri Ortalaması	Son 12 saat içerisindeki Yemek Değerleri Ortalaması	Son 24 saat içerisindeki Yemek Değerleri Ortalaması	Son 2 saat içerisindeki Yemek Değerleri Ortalaması	Son 1 saat içerisindeki Yemek Değerleri Ortalaması	Son Yemekten Sonra Geçme Süre(dakika)	Son Yemekten Sonra Geçme Süre(saat)

Şekil 3.11: Özınitelik Grupları ve Özellikleri.

İNSÜLİN ÖZNETELİKLERİ										
Öznitelik Adı	Öznitelik 33	Öznitelik 34	Öznitelik 35	Öznitelik 36	Öznitelik 37	Öznitelik 38				
Öznitelik Açıklaması	Son 6 saat içindeki Regular insülin Değerleri Ortalaması	Son 14 saat içindeki NPH insülin Değerleri Ortalaması	Son 24 saat içindeki Ultralente insülin Değerleri Ortalaması	Son 6 saat içindeki Regular insülin Max Değeri	Son 14 saat içindeki NPH insülin Max Değerleri	Son 24 saat içindeki Ultralente insülin Max Değerleri				
AKTİVİTE ÖZNETELİKLERİ										
Öznitelik Adı	Öznitelik 39	Öznitelik 40	Öznitelik 41	Öznitelik 42	Öznitelik 43	Öznitelik 44	Öznitelik 45	Öznitelik 46	Öznitelik 47	Öznitelik 48
Öznitelik Açıklaması	Son 2 saat içindeki Aktivite Toplamı	Son 4 saat içindeki Aktivite Toplamı	Son 6 saat içindeki Aktivite Toplamı	Son 12 saat içindeki Aktivite Toplamı	Son 24 saat içindeki Aktivite Toplamı	Son 2 saat içindeki Aktivite Ortalaması	Son 4 saat içindeki Aktivite Ortalaması	Son 6 saat içindeki Aktivite Ortalaması	Son 12 saat içindeki Aktivite Ortalaması	Son 24 saat içindeki Aktivite Ortalaması
TANIMLANAMAYAN KAN ŞEKERİ ÖZNETELİKLER										
Öznitelik Adı	Öznitelik 49	Öznitelik 50	Öznitelik 51	Öznitelik 52	Öznitelik 53	Öznitelik 54	Öznitelik 55	Öznitelik 56		
Öznitelik Açıklaması	Eski Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı gruptaki(pre-post) Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı gruptaki(breakfast-lunch-supper-snack) Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı gruptaki Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı zaman dilimindeki Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı zaman dilimindeki ve gruptaki (pre-post) Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı zaman dilimindeki ve gruptaki (breakfast-lunch-supper-snack) Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması	Eski Ölçüm ile aynı zaman dilimindeki ve herşeyi ile aynı gruptaki Unspecified(Yemekten sonra 2 saat geçmiş ise) Kan Şekeri Değerleri Ortalaması		
ZAMAN DİLİMİ ÖZNETELİKLERİ - 2										
Öznitelik Adı	Öznitelik 57	Öznitelik 58	Öznitelik 59	Öznitelik 60	Öznitelik 61	Öznitelik 62	Öznitelik 63	Öznitelik 64	Öznitelik 65	Öznitelik 66
Öznitelik Açıklaması	Eski Saat 00:00-06:00 arasındaki KanŞekeri Değerleri Ortalaması	Eski Saat 06:00-12:00 arasındaki KanŞekeri Değerleri Ortalaması	Eski Saat 12:00-18:00 arasındaki KanŞekeri Değerleri Ortalaması	Eski Saat 18:00-24:00 arasındaki KanŞekeri Değerleri Ortalaması	Eski Saat 06:00-18:00 arasındaki KanŞekeri Değerleri Ortalaması	Eski Saat 00:00-06:00 arasındaki Yemek Değerleri Ortalaması	Eski Saat 06:00-12:00 arasındaki Yemek Değerleri Ortalaması	Eski Saat 12:00-18:00 arasındaki Yemek Değerleri Ortalaması	Eski Saat 18:00-24:00 arasındaki Yemek Değerleri Ortalaması	Eski Saat 06:00-18:00 arasındaki Yemek Değerleri Ortalaması
Öznitelik Adı	Öznitelik 67	Öznitelik 68	Öznitelik 69	Öznitelik 70	Öznitelik 71					
Öznitelik Açıklaması	Eski Saat 00:00-06:00 arasındaki Aktivite Değerleri Ortalaması	Eski Saat 06:00-12:00 arasındaki Aktivite Değerleri Ortalaması	Eski Saat 12:00-18:00 arasındaki Aktivite Değerleri Ortalaması	Eski Saat 18:00-24:00 arasındaki Aktivite Değerleri Ortalaması	Eski Saat 06:00-18:00 arasındaki Aktivite Değerleri Ortalaması					

Şekil 3.11 (Devam): Öznitelik Grupları ve Özellikleri.

3.4.3.1. Genel öznitelikler

Öznitelik 1 :Kan şekeri ölçüm tarihinden önceki tüm kan şekeri ölçüm değerleri ortalaması

Öznitelik 2 :Kan şekeri ölçüm tarihinden önceki tüm unspecified olarak tanımlanmış kan şekeri ölçüm değerleri ortalaması

Öznitelik 3 :Kan şekeri ölçümünün pre veya post olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 4 :Kan şekeri ölçümünün breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 5 :Kan şekeri ölçümünün pre, post, breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 6 :Kan şekeri ölçümünün zaman dilimi ile aynı olan ölçüm tarihinden önceki kan şekeri ölçümleri ortalaması

Öznitelik 7 :Kan şekeri ölçümünün zaman dilimi, pre veya post olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 8 :Kan şekeri ölçümünün zaman dilimi, breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 9 :Kan şekeri ölçümünün zaman dilimi, pre, post, breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

3.4.3.2. Kan şekeri öznitelikleri

Öznitelik 10 :Kan şekeri ölçümünün pre ve breakfast olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 11 :Kan şekeri ölçümünün pre ve lunch olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 12 :Kan şekeri ölçümünün pre ve supper olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 13 : Kan şekeri ölçümünün pre ve snack olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 14 : Kan şekeri ölçümünün pre olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 15 : Kan şekeri ölçümünün post ve breakfast olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 16 : Kan şekeri ölçümünün post ve lunch olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 17 : Kan şekeri ölçümünün post ve supper olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

Öznitelik 18 : Kan şekeri ölçümünün post olarak tanımlanmasına bakılarak ölçüm tarihinden önceki aynı ölçüm değerlerinin ortalaması

3.4.3.3. Zaman dilimi öznitelikleri - 1

Öznitelik 19 : Kan şekeri ölçümünün 00:00-06:00 saatleri arasında ölçülmüş olmasına bakılarak aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

Öznitelik 20 : Kan şekeri ölçümünün 06:00-12:00 saatleri arasında ölçülmüş olmasına bakılarak aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

Öznitelik 21 : Kan şekeri ölçümünün 12:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

Öznitelik 22 : Kan şekeri ölçümünün 18:00-24:00 saatleri arasında ölçülmüş olmasına bakılarak aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

Öznitelik 23 : Kan şekeri ölçümünün 06:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

3.4.3.4. Yemek öznitelikleri

Öznitelik 24 : Kan şekeri ölçüm zamanından önceki 4 saat içerisindeki yemek değerleri ortalaması

Öznitelik 25 : Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki yemek değerleri ortalaması

Öznitelik 26 : Kan şekeri ölçüm zamanından önceki 5 saat içerisindeki yemek değerlerinin 1-5 saat içerisindeki yemek değerleri ortalaması

Öznitelik 27 : Kan şekeri ölçüm zamanından önceki 12 saat içerisindeki yemek değerleri ortalaması

Öznitelik 28 : Kan şekeri ölçüm zamanından önceki 24 saat içerisindeki yemek değerleri ortalaması

Öznitelik 29 : Kan şekeri ölçüm zamanından önceki 2 saat içerisindeki yemek değerleri ortalaması

Öznitelik 30 : Kan şekeri ölçüm zamanından önceki 1 saat içerisindeki yemek değerleri ortalaması

Öznitelik 31 : Kan şekeri ölçüm zamanından önceki son yemek zamanına kadar olan sürenin dakika değeri

Öznitelik 32 : Kan şekeri ölçüm zamanından önceki son yemek zamanına kadar olan sürenin saat değeri

3.4.3.5. İnsülin öznitelikleri

Öznitelik 33 : Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki Regular insülin dozlarının ortalaması

Öznitelik 34 : Kan şekeri ölçüm zamanından önceki 14 saat içerisindeki NPH insülin dozlarının ortalaması

Öznitelik 35 : Kan şekeri ölçüm zamanından önceki 24 saat içerisindeki Ultralente insülin dozlarının ortalaması

Öznitelik 36 : Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki maksimum Regular insülin dozu

Öznitelik 37 : Kan şekeri ölçüm zamanından önceki 12 saat içerisindeki maksimum NPH insülin dozu

Öznitelik 38 :Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki maksimum Ultralente insülin dozu

3.4.3.6. Aktivite öznitelikleri

Öznitelik 39 :Kan şekeri ölçüm zamanından önceki 2 saat içerisindeki aktivite değerleri toplamı

Öznitelik 40 :Kan şekeri ölçüm zamanından önceki 4 saat içerisindeki aktivite değerleri toplamı

Öznitelik 41 :Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki aktivite değerleri toplamı

Öznitelik 42 :Kan şekeri ölçüm zamanından önceki 12 saat içerisindeki aktivite değerleri toplamı

Öznitelik 43 :Kan şekeri ölçüm zamanından önceki 24 saat içerisindeki aktivite değerleri toplamı

Öznitelik 44 :Kan şekeri ölçüm zamanından önceki 2 saat içerisindeki aktivite değerleri ortalaması

Öznitelik 45 :Kan şekeri ölçüm zamanından önceki 4 saat içerisindeki aktivite değerleri ortalaması

Öznitelik 46 :Kan şekeri ölçüm zamanından önceki 6 saat içerisindeki aktivite değerleri ortalaması

Öznitelik 47 :Kan şekeri ölçüm zamanından önceki 12 saat içerisindeki aktivite değerleri ortalaması

Öznitelik 48 :Kan şekeri ölçüm zamanından önceki 24 saat içerisindeki aktivite değerleri ortalaması

3.4.3.7. Tanımlanamayan kan şekeri öznitelikleri

Öznitelik 49 :Kan şekeri ölçüm zamanından önceki 2 saat içerisindeki kan şekeri ölçümleri ortalaması

Öznitelik 50 :Kan şekeri ölçümünün pre veya post olarak tanımlanmasına bakılarak ölçüm tarihinden önceki 2 saat içerisindeki aynı ölçüm değerlerinin ortalaması

Öznitelik 51 : Kan şekeri ölçümünün breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki 2 saat içerisindeki aynı ölçüm değerlerinin ortalaması

Öznitelik 52 : Kan şekeri ölçümünün pre, post, breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki 2 saat içerisindeki aynı ölçüm değerlerinin ortalaması

Öznitelik 53 : Kan şekeri ölçümünün zaman dilimi ile aynı olan ölçüm tarihinden önceki tüm yemek değerlerinden sonraki 2 saat içerisindeki kan şekeri ölçümleri ortalaması

Öznitelik 54 : Kan şekeri ölçümünün zaman dilimi ile aynı olan ve kan şekeri ölçümün pre veya post olarak tanımlanmasına bakılarak ölçüm tarihinden önceki tüm yemek değerlerinden sonraki 2 saat içerisindeki kan şekeri ölçümleri ortalaması

Öznitelik 55 : Kan şekeri ölçümünün zaman dilimi ile aynı olan ve kan şekeri ölçümünün breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki tüm yemek değerlerinden sonraki 2 saat içerisindeki kan şekeri ölçümleri ortalaması

Öznitelik 56 : Kan şekeri ölçümünün zaman dilimi ile aynı olan ve kan şekeri ölçümünün pre, post, breakfast, lunch, supper, snack ve unspecified olarak tanımlanmasına bakılarak ölçüm tarihinden önceki tüm yemek değerlerinden sonraki 2 saat içerisindeki kan şekeri ölçümleri ortalaması

3.4.3.8. Zaman dilimi öznitelikleri - 2

Öznitelik 57 : Kan şekeri ölçümünün 00:00-06:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

Öznitelik 58 : Kan şekeri ölçümünün 06:00-12:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması

- Öznitelik 59** : Kan şekeri ölçümünün 12:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması
- Öznitelik 60** : Kan şekeri ölçümünün 18:00-24:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması
- Öznitelik 61** : Kan şekeri ölçümünün 06:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm kan şekeri ölçümlerinin ortalaması
- Öznitelik 62** : Kan şekeri ölçümünün 00:00-06:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm yemek değerlerinin ortalaması
- Öznitelik 63** : Kan şekeri ölçümünün 06:00-12:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm yemek değerlerinin ortalaması
- Öznitelik 64** : Kan şekeri ölçümünün 12:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm yemek değerlerinin ortalaması
- Öznitelik 65** : Kan şekeri ölçümünün 18:00-24:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm yemek değerlerinin ortalaması
- Öznitelik 66** : Kan şekeri ölçümünün 06:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm yemek değerlerinin ortalaması
- Öznitelik 67** : Kan şekeri ölçümünün 00:00-06:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm aktivite değerlerinin ortalaması
- Öznitelik 68** : Kan şekeri ölçümünün 06:00-12:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm aktivite değerlerinin ortalaması

Öznitelik 69 : Kan şekeri ölçümünün 12:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm aktivite değerlerinin ortalaması

Öznitelik 70 : Kan şekeri ölçümünün 18:00-24:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm aktivite değerlerinin ortalaması

Öznitelik 71 : Kan şekeri ölçümünün 06:00-18:00 saatleri arasında ölçülmüş olmasına bakılarak ölçüm zamanından önceki aynı zaman dilimindeki tüm aktivite değerlerinin ortalaması

3.4.4. Öznitelik Dosya Yapısı

Öznitelik dosyaları her hasta için öznitelik özelliklerine uygun olarak oluşturulmuştur. C# programlama dilinde yazılan program ile oluşturduğumuz hasta veritabanı kullanılarak öznitelik dosyaları oluşturulmaktadır. Öznitelik dosya yapısı Şekil 3.12' de gösterilmektedir.

Zaman Damgası	Öznitelik1	Öznitelik2	...	Öznitelik70	Öznitelik71	Kan Şekeri Ölçümü 1
Zaman Damgası	Öznitelik1	Öznitelik2	...	Öznitelik70	Öznitelik71	Kan Şekeri Ölçümü 2
.
.
.

Şekil 3.12 : Öznitelik Dosya Yapısı.

Her kan şekeri ölçümü için 71 öznitelik değeri hesaplanmakta ve Şekil 3.12' de gösterildiği gibi öznitelik dosyaları her hasta için ayrı ayrı oluşturulmaktadır.

3.5. Veri Madenciliği Yöntemleri

Veri madenciliği konusunda çok sayıda yöntem ve algoritma geliştirilmiştir. Bu yöntemler Sınıflama ve Regresyon, Kümeleme ve Birliktelik Kuralları olarak üç grupta toplanabilmektedir. Sınıflama ve regresyon modelleri tahmin edici, kümeleme ve birliktelik kuralları modelleri tanımlayıcı modellerdir [29].

3.5.1. Sınıflama ve Regresyon

Sınıflama ve regresyon, önemli veri sınıflarını ortaya koyan veya gelecek veri eğilimlerini tahmin eden modelleri kurabilen iki veri analiz yöntemidir [29]. Sınıflandırma yönteminde verilerin belirli bir bölümü eğitim amacı ile kullanılarak sınıflandırma kurallarının oluşturulması sağlanır, sonra bu kurallar yardımıyla yeni bir durum ortaya çıktığında nasıl karar verileceği belirlenir [30]. Üzerinde durulan değişkenlerden birinin bağımlı(y) diğerinin(x) bağımsız olması durumunda y' nin x' in bir fonksiyonu olarak ifade edilen ilişkiye de bağlanım denir. Bağımlı değişken ile bir ya da daha çok bağımsız değişken arasındaki ilişkiyi incelemek amacıyla kullanılan bir analiz yöntemine de regresyon analizi denir. Regresyon analizi neden-sonuç ilişkisini bulmamıza imkan veren bir analiz yöntemidir. Regresyon başlıca 2'ye ayrılır. Tek Değişkenli Regresyon Analizi; Tek değişkenli regresyon analizi bir bağımlı değişken ve bir bağımsız değişken arasındaki ilişkiyi inceler. Tek değişkenli regresyon analizi ile bağımlı ve bağımsız değişkenler arasındaki doğrusal ilişkiyi temsil eden bir doğrunun denklemi formüle edilir. Çok Değişkenli Regresyon Analizi; İçinde bir adet bağımlı değişken ve birden fazla bağımsız değişkenin bulunduğu regresyon modelleri çok değişkenli regresyon analizi olarak bilinir [31]. Değişkenler arasındaki ilişki çeşitleri Tablo 3.3' de gösterilmektedir.

Tablo 3.3: Değişkenler Arasındaki İlişki Çeşitleri [31].

1	$y = a + bx$	Doğru Denklemi
2	$y = a + bx + cx^2$	Parabolik İlişki
3	$y = ab^x$ veya $\log y = \log a + x \log b$	Üssel Eğri
4	$y = ax^b$ veya $\log y = \log a + b \log x$	Geometrik Eğri
5	$y = \frac{1}{a+bx}$ veya $\frac{1}{y} = a + bx$	Hiperbolik İlişki

Değişkenler arasındaki ilişki Tablo 3.3' de gösterilen ilk denklem yani doğru denklemi olarak ifade edilebiliyorsa bu regresyona doğrusal regresyon denir. Diğer ilişki çeşitleri olarak ifade edilebiliyorsa doğrusal olmayan regresyon denir.

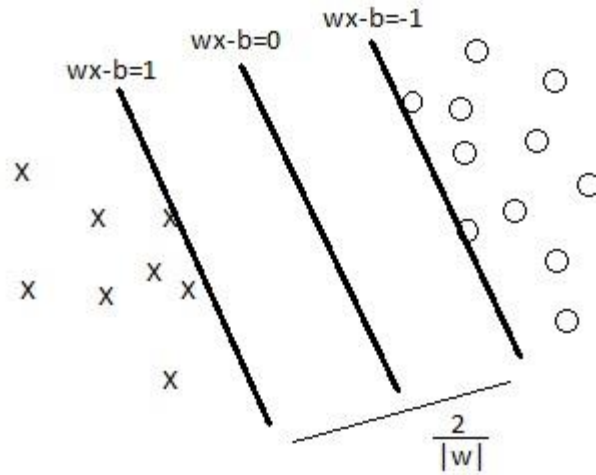
Sınıflama kategorik değerleri tahmin ederken, regresyon süreklilik gösteren değerlerin tahmin edilmesinde kullanılır [29].

Sınıflama ve regresyon modellerinde kullanılan başlıca teknikler şunlardır [32] [29].

- 1 - Destek Vektör Makineleri (SVM)
- 2 - Karar Ağaçları Öğrenmesi (Decision Tree Learning)
- 3 - K-En Yakın Komsu (KNN)
- 4 - Yapay Sinir Ağları (YSA)

3.5.1.1. Destek vektör makineleri (SVM)

Sınıflandırma için düzlem üzerindeki iki grubu birbirinden ayırmak için kullanılmaktadır. İki grup arasındaki en uzak noktaya çizilen sınır ile grupları ayırmak mümkün olmaktadır. Şekil 3.13' de iki boyutlu düzlem üzerinde gösterilen iki grubun SVM ile ayrılması gösterilmektedir.

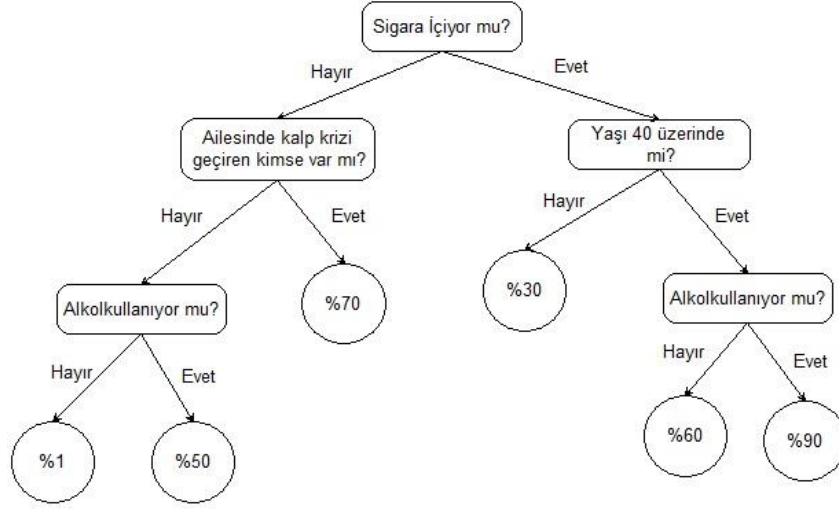


Şekil 3.13 : SVM Örnek Gösterimi [32].

Bu yöntemde iki gruba da yakın ve birbirlerine paralel sınır çizgileri çizilir. Bu sınır çizgileri birbirlerine yaklaştırılarak ortak sınır çizgisi bulunur. Bu bulunan sınır çizgisi iki grubu birbirinden ayıran en uzak sınır olmaktadır.

3.5.1.2. Karar ağaçları öğrenmesi (Decision Tree Learning)

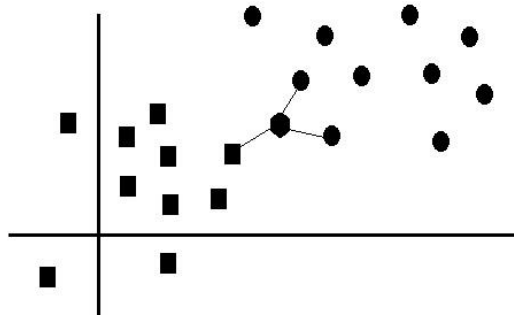
Karar Ağacı Öğrenmesi makine öğrenmesi yöntemlerinden bir tanesidir. Karar ağaçları yöntemlerinde bir ağaç yapısı oluşturularak ağacı yaprakları seviyesinde sınıf etiketleri ve bu yapraklara giden ve başlangıçtan çıkan kollar ile de özellikler üzerindeki işlemler ifade edilmektedir [32]. Şekil 3.14’ de karar ağacı örneği görülmektedir.



Şekil 3.14 : Karar Ağacı Örnek Gösterimi [32].

3.5.1.3. K-En yakın komşu (KNN)

Sınıflandırmada kullanılan bu algoritmaya göre sınıflandırma sırasında çıkarılan özelliklerden sınıflandırılmak istenilen yeni elemanın daha önceki elemanlardan k tanesine yakınlığına bakılmaktadır [32]. Şekil 3.15’ de KNN algoritması iki boyutlu düzlemde gösterilmektedir.

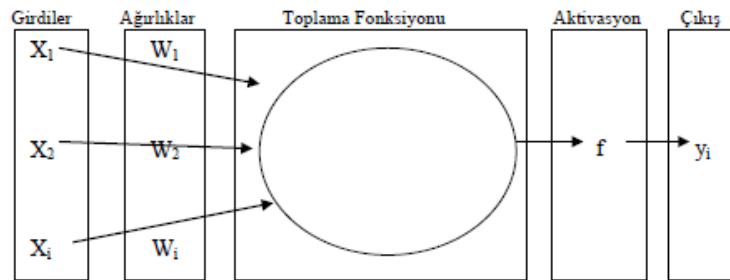


Şekil 3.15 : KNN Algoritması Örnek Gösterimi [32].

KNN algoritmasında $k=3$ için yeni bir elemanın sınıflandırılması işleminde eski sınıflandırılmış elemanlardan en yakın 3 elemana bakılarak yeni elemanın sınıflandırılması yapılmaktadır [32].

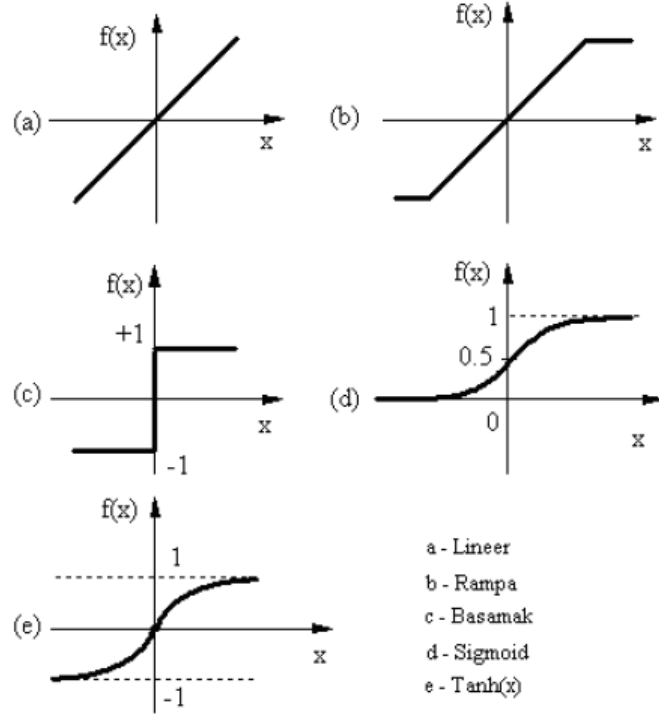
3.5.1.4. Yapay sinir ağları (YSA)

Yapay Sinir Ağları veya kısaca YSA insan beyninden esinlenerek geliştirilmiş, ağırlıklı bağlantılar aracılığı ile birbirine bağlanan işlem elemanlarından oluşan paralel ve dağıtılmış bilgi işleme yapılarıdır. En önemli özelliği, deneyimlerden (tecrübe) yararlanarak öğrenebilmesidir. Yapay sinir ağları, insan beyninin özelliklerinden olan öğrenme yolu ile yeni bilgiler türetebilme, yeni bilgiler oluşturabilme ve keşfedebilme gibi yetenekleri herhangi bir yardım almadan otomatik olarak gerçekleştirmek amacı ile geliştirilmişlerdir [33]. Yapay Sinir Ağlarını oluşturan sinir hücrelerinde girdiler, ağırlıklar, toplam fonksiyonu, aktivasyon fonksiyonu ve çıktı olmak üzere 5 ana kısım vardır [34]. Şekil 3.16' da yapay sinir ağı hücresinin yapısı gösterilmektedir.



Şekil 3.16 : Yapay Sinir Ağı Hücresini Oluşturan Elemanlar [34].

Aktivasyon fonksiyonu, etkinlik işlevi olarak da adlandırılır. Bir aktivasyon fonksiyonunun kullanım amacı hücreye gelen net girdiyi işleyerek hücrenin bu girdiye karşılık üreteceği çıktıyı belirler. Toplama fonksiyonunda olduğu gibi aktivasyon fonksiyonu olarak da çıktıyı hesaplamak için de değişik formüller kullanılır. Aktivasyon fonksiyonu olarak kullanılan diğer fonksiyonlarda Şekil 3.17' de gösterilmektedir [34].



Şekil 3.17 : Yapay Sinir Ağında Kullanılan Aktivasyon Fonksiyonları [34].

3.5.2. Kümeleme

Kümeleme, verilerin kendi aralarındaki benzerlikleri göz önüne alınarak gruplandırılması işlemidir. Bu özelliği nedeni ile pek çok alanda kullanılabilir [30].

Genel olarak başlıca kümeleme yöntemleri şu şekilde sınıflandırılabilir [29]:

- 1 - Bölme yöntemleri (Partitioning methods)
- 2 - Hiyerarşik yöntemler (Hierarchical methods)
- 3 - Yoğunluk tabanlı yöntemler (Density-based methods)
- 4 - Izgara tabanlı yöntemler (Grid-based methods)
- 5 - Model tabanlı yöntemler (Model-based methods)

3.5.3. Birliktelik Kuralları

Birliktelik kuralları birlikte gerçekleşen olayları inceleyerek olaylar arasındaki ilişkileri bulmak için kullanılmaktadırlar. Birliktelik kurallarının kullanıldığı en tipik örnek market sepeti uygulamasıdır. Bu işlem, müşterilerin yaptıkları alışverişlerdeki ürünler arasındaki birliktelikleri bularak müşterilerin satın alma alışkanlıklarını analiz eder [29].

3.6. Veri Madenciliği Programları

Veri madenciliği algoritmalarını kendi yazacağımız program ile uygulayabileceğimiz gibi veri madenciliği algoritmalarını kullanabileceğimiz ücretsiz programlar da bulunmaktadır. Orange Canvas [35], Knime Analytics Platform [36] ve WEKA [37] programları bu programlara örnek olarak gösterilebilir. Projemizde kullanılmak üzere WEKA programı seçilmiştir. WEKA programının açık kaynak kodlu Java dili ile yazılması ve GPL lisansı ile dağıtılması bu programı seçme sebeplerimizdendir. Ayrıca WEKA programı ile yapılmış birçok çalışmanın bulunması, programın sürekli geliştirme aşamasında olması, program ve algoritmaları ile ilgili bilgilerin çok fazla ve açık olması da çalışmamızda WEKA programını tercih etme sebeplerimizdendir.

3.7. WEKA Programı ve Eğitim-Test Dosyaları

WEKA programı makine öğrenmesi amacı ile Waikato Üniversitesinde geliştirilmiştir. Java dili ile geliştirilmesi ve algoritmaların jar dosyası şeklinde olması sebebi ile Java programlama dili ile yazılan projelere kolaylıkla entegre edilebilmektedir. Program üzerinde temel olarak Sınıflandırma (Classification), Kümeleme (Clustering) ve İlişkilendirme (Association) veri madenciliği işlemleri yapılabilmektedir. Ayrıca Veri Ön işleme (Data Pre-Processing) ve Görselleme (Visualization) işlemleri de yapılabilmekte ve birçok hazır algoritma bulunmaktadır. WEKA programı kendisi için özel olarak geliştirilen Attribute Relationship File Format kelimesinin baş harflerinden oluşan .arff dosya tipini kullanmaktadır. Şekil 3.18' de arff dosya yapısı gösterilmektedir.


```
@relation Hasta1  
@attribute Feature1 numeric  
@attribute Feature2 numeric  
@attribute Feature3 numeric  
@Data  
100,200,150  
122.11,144.44,145  
147,123,156
```

Şekil 3.18: Arff Dosya Yapısı.

Bu yapıya göre; arff dosyaları @relation ifadesinden sonra yazılan isim ile başlar. Burası dosyanın isim alanını oluşturduğundan algoritmalarda işleme tabi değildir.

@attribute ifadesi dosyamızda kullanılacak değişkeni tanımlamak için kullanılır. Bu ifadeden sora bir boşluk bırakılarak değişken adı, bir boşluk bırakılarak da değişkenin tipi yazılır. Attribute ifadeleri dosyamızda kullanılacak değişkenler göre artıp azalacaktır. Değişkenlerimiz sayısal değer olduğundan numeric ifadesi kullanılmıştır.

@Data ifadesi değişken değerlerinin yazıldığı kısımdır. Belirtilen değişkenlerin değerleri aralarında virgül olacak şekilde yazılır. Eğer numeric değişkenlerde ondalıklı kısım yazılacak ise ondalık ayracı olarak nokta kullanılır.

3.8. Eğitim ve Test Dosyalarının Oluşturulması

Kan şekeri tahmini için üç test yapılmaktadır. Test-1 ve Test-2 deneyleri kan şekeri tahminin kişisel veriler üzerinden yapılabilirliğini incelemek için yapılmaktadır. Test-3 ise kan şekerinin kişisel veriler kullanılmadan tahminini incelemek için yapılmaktadır.

Test-1 de her hasta için oluşturulacak en yüksek miktardaki eğitim verisi ile testler yapılacaktır. Test-2 de ise eğitim verileri belirli kurallara bağlı kalınarak parçalanarak her hasta için ayrı büyüklüklerde dört tane daha eğitim dosyası oluşturularak her hasta için ayrı ayrı kendi verileri ile testler yapılacaktır. Böylece farklı boyuttaki eğitim verilerinin tahmin üzerindeki etkileri incelenecektir.

Ayrıca test sonuçları her algoritma için birleştirilerek her algoritmanın farklı veri boyutunda performansı karşılaştırılacaktır. Test-3 de ise kan şekeri tahmininin kişisel verilerden bağımsız olarak yapılabilirliği incelenmektedir.

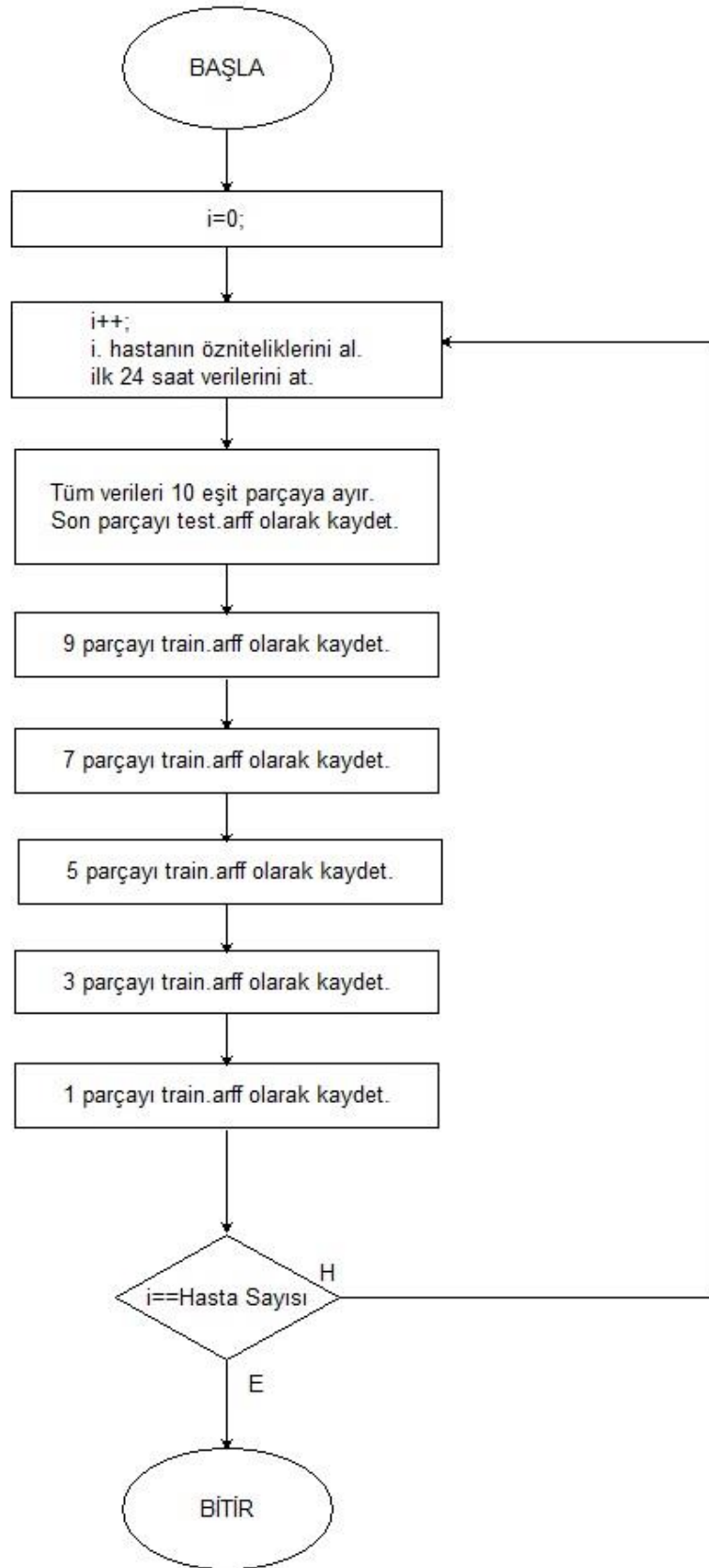
Test-1 ve Test-2 de kullanılacak eğitim ve test dosyalarının oluşturulması için hasta verilerinin nasıl bölüneceği Şekil 3.19' da gösterilmektedir.

ilk 24 saat verisi	9. Parça 1/10	8. Parça 1/10	7. Parça 1/10	6. Parça 1/10	5. Parça 1/10	4. Parça 1/10	3. Parça 1/10	2. Parça 1/10	1. Parça 1/10	1/10' luk test verisi
--------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	-----------------------

Şekil 3.19: Eğitim Verilerinin Parçalara Ayrılması.

Öncelikli olarak Şekil 3.19' da görüldüğü gibi her hasta verisinden ilk 24 saatlik veri kısmı atılacaktır. Çünkü öznitelik değeri oluşturulurken öznitelik değeri hangi kan şekeri ölçümü için oluşturuluyorsa o ölçümden önceki ölçüm değerleri kullanılmaktadır. Hasta verilerinin ilk 24 saat ölçümlerinden önce hiçbir veri olmadığından dolayı ilk 24 saat ölçümleri için oluşturulacak öznitelik değerleri sıfır olacaktır. Fakat ilk 24 saatten sonraki kan şekeri ölçümleri için oluşturulacak öznitelik değerleri sıfır olmayacaktır. Bundan dolayı ilk 24 saatlik kısım atılarak eğitim dosyası oluşturulmaktadır.

Test-1 ve Test-2 için verinin en son tarihli 1/10' luk kısmı hasta için test verisi olarak kullanılmaktadır. Diğer 9 parça beş farklı şekilde birleştirilerek eğitim verileri oluşturulmaktadır. Geri kalan verinin 9 parçasının tamamı alınarak ilk eğitim verisi, 7 parçası alınarak ikinci eğitim verisi, 5 parçası alınarak üçüncü eğitim verisi, 3 parçası alınarak dördüncü eğitim verisi ve 1 parçası alınarak beşinci eğitim verisi oluşturulmaktadır. Hasta verilerinin; 9 parçası kullanılarak oluşturulan eğitim verisi Data-1, 7 parçası alınarak oluşturulan eğitim verisi Data-2, 5 parçası alınarak oluşturulan eğitim verisi Data-3, 3 parçası alınarak oluşturulan eğitim verisi Data-4, 1 parçası alınarak oluşturulan eğitim verisi da Data-5 olarak adlandırılmaktadır. Tüm eğitim verilerinin ve test verilerinin oluşturulmasını gösteren akış diyagramı Şekil 3.20' de gösterilmektedir.



Şekil 3.20: Test Dosyaları ve Farklı Boyutlardaki Eğitim Dosyaları Oluşumu Akış Diyagramı.

Test-1 tahmin çalışmalarında en büyük oranda veriye sahip olan Data-1 eğitim verileri ile her hasta eğitilerek yapılmaktadır. Genel olarak kan şekeri tahminin yapılabilirliği incelenmektedir.

Test-2 tahmin çalışmalarında farklı büyüklükte oluşturulan her eğitim verisi ile tahmin çalışmaları yapılmaktadır. Böylelikle eğitim verilerinin farklı büyüklüklere bölünmesinin tahminler üzerindeki etkisi incelenmektedir. Ayrıca algoritmalar da karşılaştırılarak farklı büyüklüklerde veri ile her algoritmanın tahminleri karşılaştırılmaktadır.

Test-3 tahmin çalışmalarında her hasta öznelik verilerinin ilk 24 saati atılarak test dosyası oluşturulmaktadır. Diğer hasta öznelik dosyalarının da ilk 24 saatlik kısmı atılarak birleştirilmekte ve eğitim dosyası olarak kullanılmaktadır. Tahmin sonuçları incelenerek kan şekeri tahminin genel bir kurala bağlanabilirliği araştırılmakta, kişisel verilerden bağımsız olarak kan şekeri tahminin yapılabilirliği araştırılmaktadır.

DÖRDÜNCÜ BÖLÜM

DENEYSEL ÇALIŞMALAR

Kan şekeri tahmini için oluşturulan eğitim ve test dosyalarımız hastaların insülin kullanımı, yemek, aktivite ve kan şekeri ölçüm bilgilerinden oluşmaktadır. Kan şekeri tahminini hastanın geçmiş kan şekeri ölçümlerini hasta verileri ile ilişkilendirerek yapmaktadır. Geçmiş hasta kan şekeri değerleri kullanılarak kan şekeri tahmininin yapılması, değer tahmini içerdiğinden dolayı regresyon analizi çalışmamızda kullanacağımız yöntem olacaktır. Çalışmamızda regresyon analizinin yanında Karar Ağaç Algoritmaları da kullanılarak hastaların kan şekeri tahmini üzerinde etkili olan özniteliklerin belirlenmesi amaçlanmaktadır. Bu şekilde etkili olan özniteliklerin incelenmesi ve kan şekeri tahmini üzerinde etkili olan insülin bilgisi, yemek bilgisi, aktivite bilgisi ve kan şekeri ölçüm bilgisi gibi hasta verilerinin belirlenmesi de çalışmamızın diğer amaçlarındanıdır.

4.1. Kullanılan Algoritmalar

Çalışmamızda veri madenciliği aracı olarak WEKA kullanılmıştır. WEKA içerisinde kullanılan algoritmalar :

- 1 - AdditiveRegression
- 2 - DecisionStump
- 3 - GaussianProcesses
- 4 - LinearRegression
- 5 - M5P
- 6 - MultilayerPerceptron
- 7 - RepTree
- 8 - SMOReg

Çalışmamızda kullanılan algoritmalar regresyon analizini destekleyen sınıflama algoritmaları arasından seçilmiştir. DecisionStump, M5P ve RepTree karar ağaç yöntemlerini kullanan sınıflama algoritmalarıdır. AdditiveRegression, GaussianProcesses, LinearRegression ve SMOReg regresyon analizini kullanan sınıflama algoritmalarıdır. MultilayerPerceptron Yapay Sinir Ağı sınıflama algoritmasıdır.

Decision Stump algoritması, tek seviyeli bir karar ağacı oluşturmaktadır. Bu yöntem ile oluşturulan ağaçta kök düğüm, yaprak düğümlere doğrudan bağlıdır ve karar işlemini doğrudan tek bir öznitelik değerine dayalı olarak gerçekleştirmektedir. Decision Stump algoritması genellikle boosting yöntemleri ile birlikte kullanılmaktadır [2] [38]. Çalışmamızda boosting yöntemi kullanılmamıştır.

M5P algoritması, M5 algoritmasının regresyon modelleri ile ağaç yapısına uyarlanmasıdır. M5P algoritması M5 algoritmasını temel alarak genel karar ağacı yapısı ile Linear Regression fonksiyonlarını birleştirmektedir [6].

REPTree algoritması, hızlı karar ağacı sınıflandırma algoritmalarından biridir. Algoritma, karar ya da regresyon ağacının oluşturulmasında bilgi kazancı ölçütünü kullanmakta ve oluşan ağacı, azaltılmış hata budaması yöntemine dayalı olarak budama işlemine tabi tutmaktadır. REPTree algoritmasında, yalnızca nümerik özniteliklerin sıralanması söz konusudur. Eksik değerler için ise C4.5 algoritmasının örneklere karşılık gelen parçalara ayırma yaklaşımı uygulanmaktadır [8] [38].

AdditiveRegression algoritması, doğrusal olmayan zaman serilerinin tahmininde kullanılan yöntemlerdir [39]. Standart regresyon modelinden daha esnektir. Tahminler sınıflayıcı tahminlerine ekleyerek gerçekleştirilir. Bu sebeple hatalar birbirinden bağımsızdır. Bu hata bağımsızlığı bir düzeltme etkisi yaratmaktadır fakat tahmin süresini uzatmaktadır [1].

GaussianProcesses algoritması, sınıflama yaparken gauss süreçlerini kullanır. Öğrenilen modele göre gauss uzaklığı formülünü kullanarak sınıflama yapar [3].

LinearRegression algoritması, değişkenler arasındaki ilişkiyi fonksiyonel olarak açıklamaktadır. Akaike bilgi kriterini kullanmaktadır [4].

SMOReg algoritması, SVM üzerinde regresyon uygulamaktadır [32]. Çeşitli parametreler kullanarak öğrenme işlemini gerçekleştirmektedir. En popüler olanı RegSMOImproved algoritmasıdır [9].

MultilayerPerceptron algoritması, yapay sinir ağı algoritmasıdır ve insan beyninin sinir hücresinden oluşmuş yapısını bilgisayar ortamına aktarmaktadır [5]. Algoritma özneliklerin yarısı kadar gizli katman oluşturacak şekilde çalıştırılmaktadır.

Algoritmalar ile yapılan tahmin işlemi sonucunda WEKA programı algoritmaların türlerine göre matematiksel model ya da karar ağacı yapısı oluşturmaktadır. Bu karar ağacı yapısına bakılarak kan şekeri tahmininde önemi olan öznelikler belirlenmektedir. Ayrıca WEKA programı her algoritma için Korelasyon Katsayısı (Correlation coefficient), Ortalama Mutlak Hata (Mean absolute error), Ortalama Karesel Hatanın KareKökü (Root mean squared error), Göreli Mutlak Hata (Relative absolute error) ve Göreli Karesel Hatanın KareKökü (Root relative squared error) değerlerini üretmektedir.

Korelasyon; iki değişken arasındaki ilişkiyi veya bir değişkenin daha fazla değişken ile ilişkisini tespit etmek, varsa yönünü ve şiddetini bulmak için kullanılan yöntemdir. Korelasyon değeri -1 ile +1 arasında değişmektedir. Değerin -1' e yaklaşması ilişkinin ters yönde, +1' e yaklaşması ise ilişkinin aynı yönde olduğunu göstermektedir. Korelasyon katsayısı ile katsayıya karşılık gelen ilişkinin gücü Tablo 4.1' de gösterilmektedir.

Tablo 4.1: Korelasyon Katsayısı Aralığı ve İlişki Gücü [40].

Korelasyon Katsayısı	İlişki Gücü
0,00 – 0,25	Çok zayıf ilişki
0,26 – 0,49	Zayıf ilişki
0,50 – 0,69	Orta ilişki
0,70 – 0,89	Yüksek ilişki
0,90 – 1,00	Çok yüksek ilişki

Algoritma tahminleri sonucunda üretilen korelasyon değerleri kan şekeri ölçüm değerleri ile kan şekeri tahmin değerleri arasındaki korelasyonu göstermektedir. Kan şekeri tahmininin doğru yapıldığının belirlenmesi için pozitif korelasyon değerleri dikkate alınmaktadır. Çünkü korelasyon değerlerinin +1' e yakınlığı kan şekeri tahminlerinin kan şekeri ölçümlerine yakınlığını göstermektedir.

Korelasyon deęerinin negatif olması durumunda ise kan řekeri tahminleri ile kan řekeri ölçümleri arasında zıt bir ilişkinin olduęu söylenebilir. Çalışmamız kan řekerinin doęru tahmini ve kan řekeri oluşumunda etkili olan özellikler üzerine olduğundan algoritmaların tahminleri sonucunda oluşan pozitif korelasyon deęerleri üzerinden deęerlendirmeleri yapılmaktadır. Algoritmaların parametreleri deęiştirilmemiş, WEKA tarafından varsayılan parametreler ile işlemler yapılmıştır.

Çalışmamızda kullandığımız algoritmalar kısaltmalar kullanılarak belirtilmektedir. Algoritmaların kısaltmaları Tablo 4.2' de gösterilmektedir.

Tablo 4.2: Algoritmaların Adları ve Kısaltmaları.

Algoritma Adı	Algoritma Kısaltması
AdditiveRegression	AR
DecisionStump	DS
GaussianProcesses	GP
LinearRegression	LR
M5P	M5P
MultilayerPerceptron	MP
RepTree	RT
SMOReg	SMOR

DeneySEL çalışmalarımızda yapılan testlerin sonucunda algoritmaların korelasyon deęerlerinin akışa ve sayfa düzenine uygun olarak gösterilmesi için algoritma adları yerine Tablo 4.2' de gösterilen algoritma kısaltmaları kullanılmaktadır.

BEŞİNCİ BÖLÜM

DENEYLER VE SONUÇLAR

Bu bölümde Test-1, Test-2 ve Test-3 başlıkları altında yapılan farklı tahmin çalışmaları incelenmektedir. Bu tahminlerin sonuçları incelenmekte ve tahmin çalışması ile ilgili değerlendirmeler yapılmaktadır. Değerlendirmeler algoritmaların korelasyon katsayıları incelenerek yapılmaktadır.

5.1. Test-1 Deneyleri ve Sonuçları

Test-1, algoritmaların dördüncü bölümde anlatılmış olan Data-1 eğitim dosyaları ile eğitilmesi sonucundaki testlerdir. Tüm hastaların verileri Şekil 3.18’ de gösterildiği gibi parçalara ayrılarak Data-1 eğitim verileri oluşturulmaktadır. Data-1 verileri kişisel eğitim verileri için oluşturulan en büyük boyuttaki verilerdir. Tüm hastaların Data-1 verileri ile eğitilen algoritmalar ile kan şekeri tahmini yapılmaktadır. Tüm hastalar için yapılan testlerde her algoritma için elde edilen en yüksek sonuçlar Tablo 5.1’ de verilmektedir. Tüm algoritmaların Data-1 ile eğitilmesi sonucundaki tahmin değerleri Ek A Tablo A-1’ de verilmiştir.

Tablo 5.1: Data-1 ile Eğitilen Tüm Algoritmaların Maksimum Korelasyon Değerleri.

AR	DS	GP	LR	MSP	MP	RT	SMOR
0,6976	0,8354	0,6356	0,6922	0,8537	0,7833	0,8354	0,7037

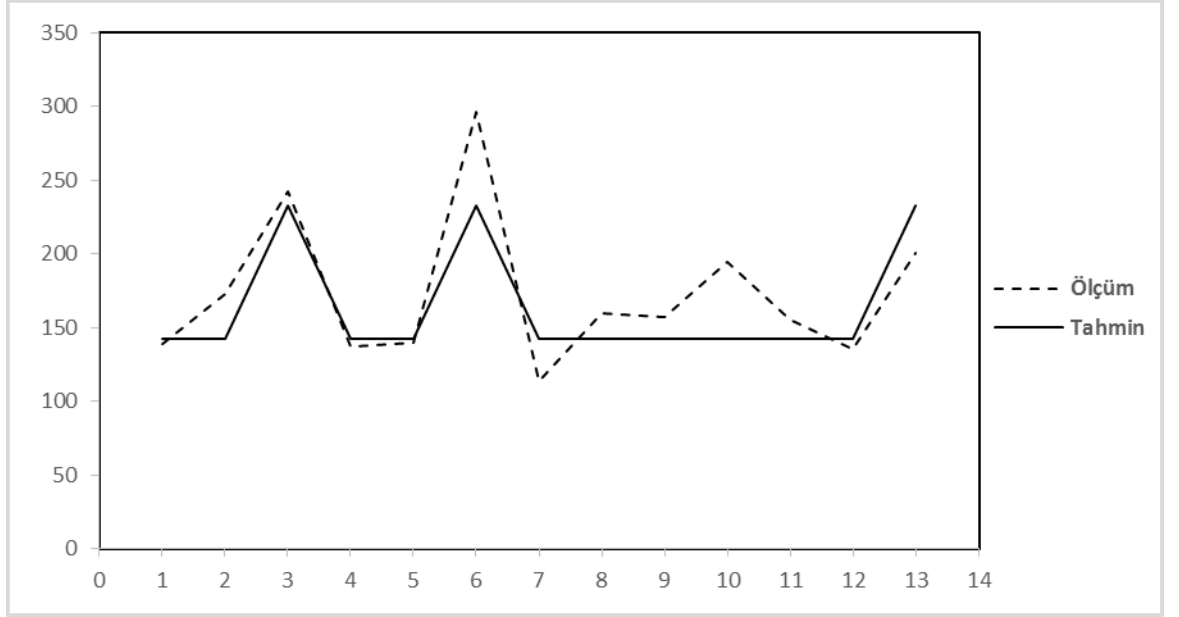
Tablo 5.1’ deki korelasyon değerlerine bakıldığında yüksek korelasyon değerine sahip algoritmaların olduğu görülmektedir. Bu da kurulan sistemin tamamen hatalı olmadığını göstermektedir. Ancak bazı algoritmaların korelasyon değerleri düşük çıkmıştır. Üçüncü bölüm Şekil 3.6’ dan da görüleceği üzere hastaların insülin, kan şekeri ölçüm, yemek, aktivite verileri ve gözlem süreleri farklıdır. Verilerdeki bu

eksikliklerden dolayı algoritmaların korelasyon değerlerinin düşük olması muhtemeldir. Ayrıca hasta verilerine karışmış olan hatalardan dolayı da hasta algoritmaların korelasyon değerleri düşük çıkmaktadır. Hasta verilerindeki hatalar Test-2 deneylerinde daha detaylı incelenmektedir.

Test-1 deneylerindeki algoritmaların sonuçlarına bakacak olursak en yüksek korelasyon değeri Hasta5' in DS, M5P ve RT algoritmasına aittir. En yüksek korelasyon değerine sahip üç algoritmalarından DS' nin kan şekeri ölçüm ve tahminleri Tablo 5.2' de, M5P' nin kan şekeri ölçüm ve tahminleri Tablo 5.3' de ve RT' nin kan şekeri ölçüm ve tahminleri Tablo 5.4' de verilmektedir. Test-1 deneyleri ölçüm ve tahminlerinden DS algoritmasının ölçüm-tahmin grafiği Şekil 5.1' de, M5P algoritmasının ölçüm-tahmin grafiği Şekil 5.2' de ve RT algoritmasının ölçüm-tahmin grafiği Şekil 5.3' de verilmektedir. Tahmin tabloları ve grafikler incelendiğinde algoritmaların korelasyon değerlerinin yanında kan şekerinin değişim anlarını nasıl yorumladığı anlaşılmaktadır. Kan şekerinin ölçümünün azaldığı veya arttığı durumlarda DS, M5P ve RT algoritmaları kan şekeri ölçümünün bu değişimini doğru tahmin etmektedirler. Kan şekeri ölçümünün artması durumunda algoritmaların tahminleri artmakta, kan şekeri ölçümünün azalması durumunda da algoritmaların tahminleri azalmaktadır. Algoritmaların kan şekeri ölçümünün değişimini doğru tahmin etmesi algoritmaların grafiklerinden daha net anlaşılmaktadır.

Tablo 5.2: Hasta5' in Data-1 Verisi ile Eğitilen DS Algoritmasının Kan Şekeri Ölçüm ve Tahmin Değerleri.

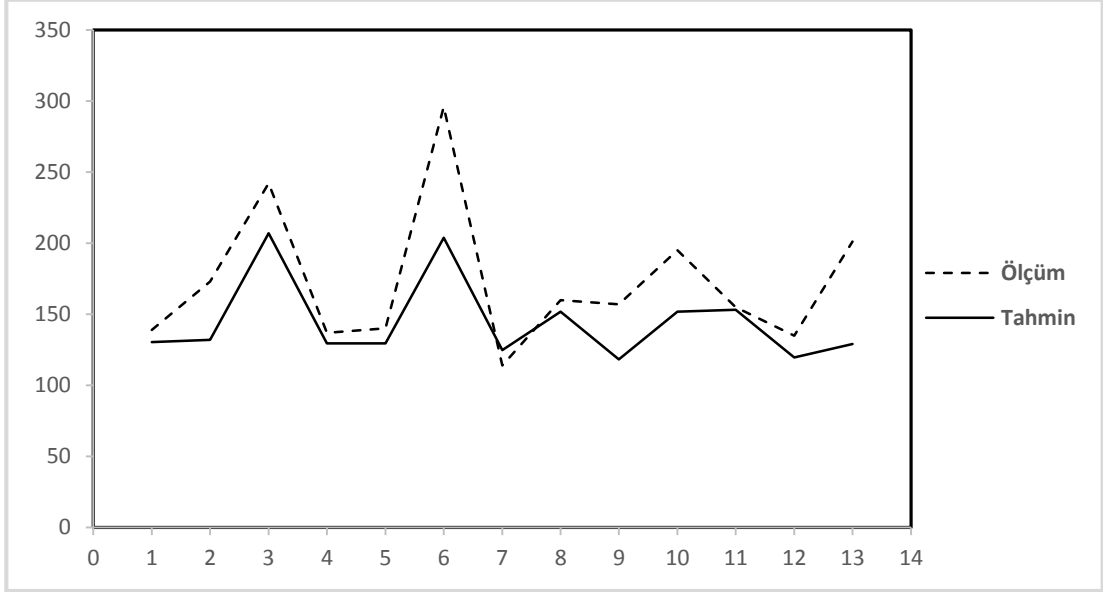
Ölçüm	Tahmin
139	142,670213
173	142,670213
242	233,181818
137	142,670213
140	142,670213
296	233,181818
114	142,670213
160	142,670213
157	142,670213
195	142,670213
155	142,670213
135	142,670213
201	233,181818



Şekil 5.1: Hasta5' in Data-1 Verisi ile Eğitilen DS Algoritmasının Kan Şekeri Ölçüm ve Tahmin Grafiği.

Tablo 5.3: Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritmasının Kan Şekeri Ölçüm ve Tahmin Değerleri.

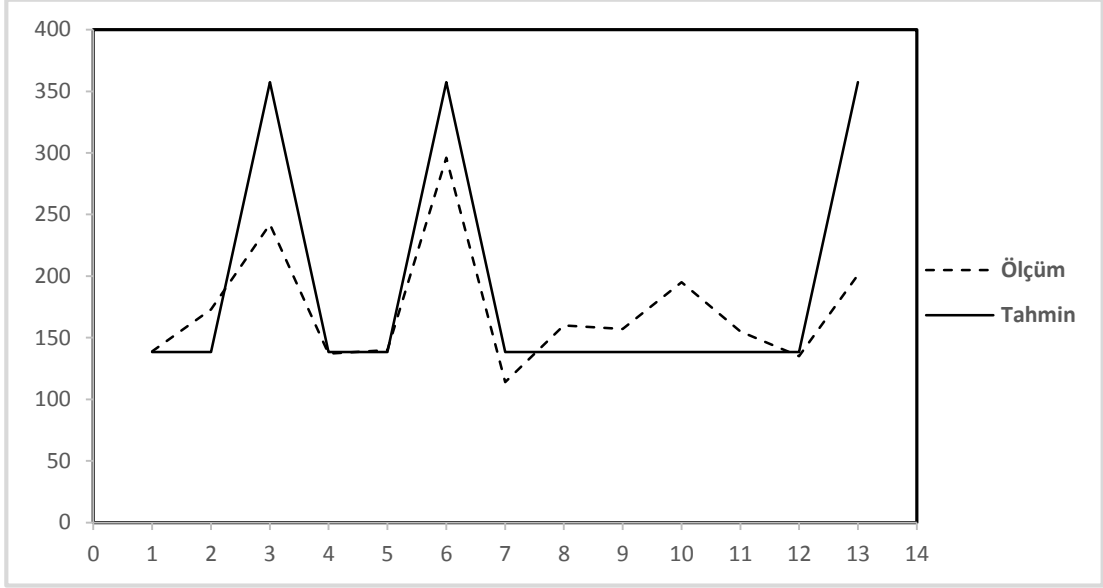
Ölçüm	Tahmin
139	130,392866
173	132,034137
242	207,066219
137	129,412955
140	129,414268
296	203,720029
114	124,773284
160	151,838947
157	118,182059
195	151,891598
155	153,071126
135	119,560513
201	129,142125



Şekil 5.2: Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritmasının Kan Şekeri Ölçüm ve Tahmin Grafiği.

Tablo 5.4: Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritması Ölçüm ve Tahmin Değerleri.

Ölçüm	Tahmin
139	138,367816
173	138,367816
242	357,333333
137	138,367816
140	138,367816
296	357,333333
114	138,367816
160	138,367816
157	138,367816
195	138,367816
155	138,367816
135	138,367816
201	357,333333



Şekil 5.3: Hasta5' in Data-1 Verisi ile Eğitilen RT Algoritması Ölçüm-Tahmin Grafiği.

Tahminlerin yürütüldüğü veri hastaların tüm verilerinin zamana bağlı bölünmesi ile elde edilmektedir. Tahmin değerleri Hasta5' e ait olduğundan dolayı Tablo 5.2, Tablo 5.3 ve Tablo 5.4' deki ölçüm verileri aynıdır. Tablo 5.2, Tablo 5.3 ve Tablo 5.4 incelendiğinde kan şekeri ölçüm değerlerinde yükselme ve alçalmanın yaşandığı noktalarda tahmin değerlerinin de ölçüm değerleri ile paralel olarak hareket ettiği görülmektedir. Değerlerdeki bu paralellik algoritmaların kan şekeri değişimini doğru öngördüğünü göstermektedir. Kan şekeri ölçüm değerleri ve algoritma tahminlerindeki bu paralellik Şekil 5.1, Şekil 5.2 ve Şekil 5.3' de daha net görülmektedir. Tahmin değerlerinin analizi yapıldığında ölçüm ve tahmin arasında hata payının olduğu görülmektedir. Bu hata payı sadece modelden dolayı değil verilerde bulunan hatalardan da olabilir. Çünkü kan şekeri ölçüm verilerine katılması olası birçok hata ihtimali bulunmaktadır. Bu hata ihtimallerini hastadan ve ölçücüden kaynaklı hatalar olarak iki gruba ayırmak mümkündür. Hastadan kaynaklı olan hatalarda, verilebilecek örneklerden biri hastanın biyolojik ve psikolojik durumudur. Kan şekeri değeri üzerinde etkili olan birçok biyolojik etmen vardır [41]. Hatta, biyolojik her etkinin kan şekeri üzerinde etkili olabileceği ihtimalinin olduğunu söylemek doğrudur. Örnek vermek gerekirse; hastanın ölçüm yapılmadan önce heyecanlanması ya da nefes alış hızının artması dahi kan şekeri ölçüm değerlerini değiştirecektir. Bu durumda ölçülen kan şekeri değeri olası durumun dışında oluşan

bir kan şekeri ölçümü olacaktır. Hastanın kan şekeri ölçümü sırasında kan şekeri değerinin yanlış ölçülmesi ya da hastanın kan şekeri değerini değiştirecek ve kayıt altına alınmamış olası bir aktivitenin bulunması da kan şekeri ölçümünün veri setinde tam olarak ifade edilememesine sebebiyet verecektir. Bu hataların hasta verilerine katılması algoritmamızın kan şekeri tahmininde hata payı olarak karşımıza çıkabilir. Araştırmamızda yaptığımız tahmin çalışmaları Data-1 ile yani en yüksek orandaki eğitim verileri ile yapılmaktadır. Dolayısı ile hasta verilerine hata karışma ihtimali var ise boyutu en büyük olan veri grubu bu grup olduğundan dolayı en fazla hatanın karışması muhtemel veriler de bunlar olacaktır. Dikkat çekmek istediğimiz bir nokta, kan şekeri tahmininde hata payı olmasına rağmen kan şekeri değişim durumlarını yüksek doğruluk oranı ile yapılmıştır. Bu sonuca bakılarak kullanılan algoritmaların kan şekeri değişimini doğru analiz ettiği söylenebilmektedir. Kan şekeri değişiminin doğru analizinden dolayı oluşturulan öznelik değerlerinin de kan şekeri tahmini için uygun olduğu söylenebilir. En yüksek üç tahmin değerini veren algoritma da Karar Ağacı algoritmalarıdır. Bu algoritmaların tahminlerinin ağaç yapıları incelendiğinde bu tahminin hangi özellikler üzerinde odaklanarak gerçekleştirildiği görülebilir. Yani kan şekeri tahmin edebilmekte en etkili olan özellikler ve bu özelliklerin sıralamaları ağaç yapılarından görülebilmektedir. Şekil 5.4’ de DS, Şekil 5.5’ de M5P, Şekil 5.7’ de RT algoritmasının karar ağacı şekli görülmektedir. Ayrıca Şekil 5.6’ da M5P algoritmasının oluşturduğu Lineer Modeller ve Şekil 5.7’ de de RT algoritmasının karar ağacı görülmektedir.

```
Decision Stump
```

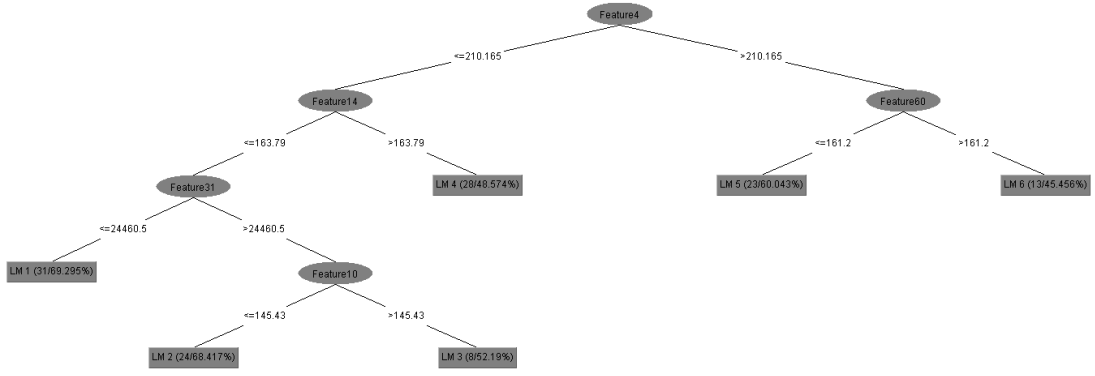
```
Classifications
```

```
Feature8 <= 219.75 : 142.67021276595744
```

```
Feature8 > 219.75 : 233.1818181818182
```

```
Feature8 is missing : 166.18897637795277
```

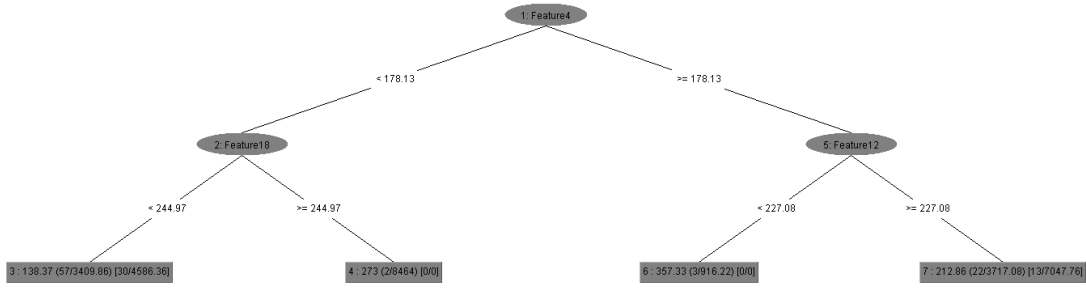
Şekil 5.4: Hasta5’ in Data-1 Verisi ile Eğitilen DS Algoritması Tahminlerinin Karar Ağacı.



Şekil 5.5: Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritması Tahminlerinin Karar Ağacı Şekli.

<p>LM num: 1 Glikoz =</p> <ul style="list-style-type: none"> -1.7687 * Feature1 - 0.0251 * Feature2 - 0.0452 * Feature3 + 0.0972 * Feature4 + 0.2671 * Feature10 - 0.2699 * Feature11 - 0.278 * Feature12 + 0.6272 * Feature13 + 0.4404 * Feature14 - 0.0006 * Feature31 - 1.8823 * Feature34 + 0.5487 * Feature58 - 0.1724 * Feature59 - 0.4452 * Feature60 + 392.9969 	<p>LM num: 2 Glikoz =</p> <ul style="list-style-type: none"> -3.3185 * Feature1 - 0.0251 * Feature2 - 0.454 * Feature3 + 0.0972 * Feature4 + 0.2509 * Feature10 - 0.2699 * Feature11 - 0.278 * Feature12 + 1.5103 * Feature13 + 0.4404 * Feature14 - 0.0006 * Feature31 - 1.862 * Feature34 + 0.5487 * Feature58 - 0.1724 * Feature59 - 0.4565 * Feature60 + 607.8863 	<p>LM num: 3 Glikoz =</p> <ul style="list-style-type: none"> -4.42 * Feature1 - 0.0251 * Feature2 - 0.3042 * Feature3 + 0.0972 * Feature4 + 0.2509 * Feature10 - 0.2699 * Feature11 - 0.278 * Feature12 + 2.1247 * Feature13 + 0.4404 * Feature14 - 0.0006 * Feature31 - 3.4277 * Feature34 + 0.5487 * Feature58 - 0.1724 * Feature59 - 0.4565 * Feature60 + 662.2865
<p>LM num: 4 Glikoz =</p> <ul style="list-style-type: none"> -0.7636 * Feature1 - 0.0251 * Feature2 - 0.082 * Feature3 + 0.0972 * Feature4 - 0.4968 * Feature10 + 1.009 * Feature11 - 0.278 * Feature12 + 0.6272 * Feature13 + 1.2417 * Feature14 - 0.0003 * Feature31 - 1.3277 * Feature34 + 0.5487 * Feature58 - 0.1724 * Feature59 - 1.2133 * Feature60 + 155.3258 	<p>LM num: 5 Glikoz =</p> <ul style="list-style-type: none"> 1.1478 * Feature1 - 0.2819 * Feature2 - 1.162 * Feature4 - 1.1336 * Feature5 - 0.4562 * Feature7 - 1.9134 * Feature10 - 0.5609 * Feature11 - 0.5778 * Feature12 - 0.305 * Feature13 - 1.1307 * Feature14 - 0.0005 * Feature31 - 0.9103 * Feature34 + 1.1405 * Feature58 - 2.0528 * Feature59 - 0.0006 * Feature60 + 1684.11 	<p>LM num: 6 Glikoz =</p> <ul style="list-style-type: none"> 2.4905 * Feature1 - 0.3639 * Feature2 - 1.6491 * Feature4 - 2.228 * Feature10 - 0.5609 * Feature11 - 0.5778 * Feature12 - 0.8795 * Feature13 - 1.1307 * Feature14 - 0.0005 * Feature31 - 0.9103 * Feature34 + 3.2895 * Feature58 - 2.6579 * Feature59 + 0.5066 * Feature60 + 1213.0314

Şekil 5.6: Hasta5' in Data-1 Verisi ile Eğitilen M5P Algoritması Tahminlerinin Linear Modelleri.



Şekil 5.7: Hastas5' in Data-1 Verisi ile Eğitilen RT Algoritması Tahminlerinin Karar Ağacı Şekli.

```

REPTree
=====

Feature4 < 178.13
|   Feature18 < 244.97 : 138.37
|   Feature18 >= 244.97 : 273
Feature4 >= 178.13
|   Feature12 < 227.08 : 357.33
|   Feature12 >= 227.08 : 212.86

Size of the tree : 7

```

Şekil 5.8: Hastas5' in Data-1 Verisi ile Eğitilen RT Algoritması Tahminlerinin Karar Ağacı.

En yüksek korelasyon sonuçlarını veren algoritmalarından M5P, DS ve RT algoritmalarının karar ağaç modelleri incelendiğinde genel olarak kan şekeri öznitelikleri üzerinden karar verildiği görülmektedir. Yani algoritmaların karar ağaçlarını oluşturan özniteliklerin özellikleri kan şekeri ölçümü ile ilgilidir. M5P algoritmasının karar ağacına bakıldığında Öznitelik4, Öznitelik14, Öznitelik10, Öznitelik31 ve Öznitelik60 değerlerinin karar için öncelikli değerler olduğu görülmektedir.

Karar ağacının oluşturduğu altı tane Lineer Model incelendiğinde de her modelde Öznitelik1, Öznitelik2, Öznitelik3, Öznitelik4, Öznitelik10, Öznitelik11, Öznitelik12, Öznitelik13, Öznitelik14, Öznitelik31, Öznitelik34, Öznitelik58, Öznitelik59 ve Öznitelik60 değerlerinin olduğu görülmektedir. Yukarıda belirtilen özniteliklerin açıklamaları üçüncü bölümde anlatılmıştır. Algoritmasının seçtiği öznitelikler genel olarak kan şekeri ölçümü ile ilgili öznitelikler olduğu görülmektedir.

Buna ek olarak yemek miktarı (Öznitelik31) ve insülin dozu (Öznitelik34) ile ilgili öznitelikler de belirleyici etki göstermiştir. Özniteliklerin özellikleri göz önüne alınarak kan şekeri tahmininde en yüksek önemi önceki kan şekeri ölçüm bilgilerinin gösterdiği görülmektedir. Karar ağaçlarının oluşumunda kan şekeri ölçümü ile ilgili özniteliklerin daha fazla ve yukarılardaki düğümlerde olmalarından dolayı algoritmaların kan şekeri tahmininde kan şekeri ölçüm bilgilerine yüksek önem verdiği söylenebilir.

DS algoritmasının karar ağacına bakıldığında Öznitelik8 değerinin karar için referans alındığı görülmektedir. Bu öznitelik önceki ölçümlerden aynı zaman dilimi ve aynı grupta olan kan şekeri ölçüm değerleri ile ilgilidir. Seçilen bu özniteliğe bakılarak bu algoritmanın kan şekeri tahmininde önceki ölçümleri referans aldığını göstermektedir.

RT algoritmasının karar ağacına bakıldığında ise karar ağacındaki tüm öznitelik değerlerinin kan şekeri ölçümü ile ilgili olduğu görülmektedir. Bu algoritma da tahmin için önceki kan şekeri ölçümlerini kullanmaktadır.

Üç algoritmanın sonuçları incelendiğinde kan şekeri tahminlerinin önceki kan şekeri ölçümlerine bakılarak yapıldığı görülmektedir. İnsülin ve yemek değerlerinin de tahmin aşamasında etkisi olmuştur. Bu üç algoritma tahmininin incelenmesi Hasta5 üzerinde gerçekleşmiştir. Bu tahminler sonucundaki hata payının hasta verilerinden kaynaklandığının daha net gösterilmesi için Hasta5' in verilerinin belirli parçaları alınmıştır.

Tablo 5.5: Hasta5' in 02.09.1990 Tarihli Verileri.

Tarih	Saat	Kod	Data
02.09.1990	08:01	58	66
02.09.1990	08:03	33	2
02.09.1990	08:03	34	10
02.09.1990	11:40	60	136
02.09.1990	11:42	67	3
02.09.1990	17:15	56	330
02.09.1990	17:17	62	319
02.09.1990	17:18	67	3
02.09.1990	17:18	33	2
02.09.1990	17:18	34	8
02.09.1990	21:57	64	227

Tablo 5.6: Hasta5' in 06.09.1990 Tarihli Verileri.

Tarih	Saat	Kod	Data
06.09.1990	06:12	58	60
06.09.1990	06:14	33	3
06.09.1990	06:14	34	9
06.09.1990	11:47	60	85
06.09.1990	17:27	62	341
06.09.1990	17:29	33	1
06.09.1990	17:29	34	8
06.09.1990	22:06	64	137

Tablo 5.5' de Hasta5 ile ilgili 02.09.1990 tarihli saat 11:42' deki yiyecek tüketimi ve ayrıntıları ile aynı gün saat 17:17' deki kan şekeri ölçümü koyu yazı stili ile gösterilmektedir. Bu verilere bakılarak kan şekeri ölçümünde kan şekeri değerin 319 olarak kaydedilmesinin beş saat önceki aşırı yemek ile alakası olduğu düşünülmektedir.

Hasta ile ilgili Tablo 5.6 incelendiğinde koyu yazı stili ile gösterilen kan şekeri ölçüm verisinde ise 06.09.1990 tarihinde saat 17:27' de Tablo 5.6' nın Data kolonunda kan şekeri değerinin 341 olarak kaydedildiği görülmektedir. Fakat Tablo 5.6' da tüm gün verisi içerisinde hiçbir yemek verisi kaydı bulunmamaktadır. Kan şekerinin değişiminin yemek ile ilişkisi bilindiğinden [27] [17] bu verideki kan şekeri yüksekliğinin nedeni olarak yemek verisinin kaydedilmemesi ya da ölçüm sonuçlarına katılmış olabilecek hatalar olduğu düşünülmektedir. Bu şekilde hasta verilerine karışması muhtemel birçok hatanın olması neticesinde algoritmalarımız ile yapılan tahminler sonucunda hata payları oluşabileceği gösterilmek ve ispatlanmak istenmiştir. Yukarıdaki tablo verilerinin dikkatli analizi kullanılan veri setinde hatalar olma olasılığının güçlü olduğunu tasdik eder niteliktedir.

5.2. Test-2 Deneyleri ve Sonuçları

Test-2, her hasta için oluşturulan farklı boyutlardaki eğitim verileri ile algoritmalar ayrı ayrı eğitilerek çalıştırılmıştır. Test-1 çalışmasında en yüksek korelasyon değerini veren M5P, DS ve RT algoritmalarının farklı veri boyutları ile yapılan testler sonucunda farklı veri boyutları için ulaşılmış oldukları en yüksek korelasyon değerleri Tablo 5.7, Tablo 5.8 ve Tablo 5.9’ da gösterilmektedir. Sunumu ve akışı bozmamak için tablolarda sadece farklı boyuttaki her veri için en yüksek tahmin değerine sahip hastaların değerleri gösterilmektedir. Ek B Tablo B-1’ de tüm hastaların M5P algoritması ile tahmin sonuçları, Ek B Tablo B-2’ de DS algoritması ile tahmin sonuçları ve Ek B Tablo B-3’ de tüm hastaların RT algoritması ile tahmin sonuçları bulunmaktadır.

Tablo 5.7: Farklı Boyutlarda Veriler ile Eğitilen M5P Algoritmasının En Yüksek Korelasyon Değerleri.

M5P					
Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
.
.
.
5	0,8537	0,6691	0,7991	0,5192	0,4237
.
.
.
11	0,5986	0,7156	0,5634	0,623	0,6679
.
.
.
31	0,2799	0,6128	0,6064	0,6049	0,6969
.
.
.
46	0,2139	0,0325	0,0682	0,6745	0
.
.
.

Tablo 5.8: Farklı Boyutlarda Veriler ile Eğitilen DS Algoritmasının En Yüksek Korelasyon Değerleri.

DecisionStump					
Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
.
.
.
3	0,5274	0,5274	0,5274	0,6484	0,0659
.
.
.
5	0,8354	0,8354	0,8354	0,1695	0,3142
.
.
.
46	0,282	0,282	0,5893	0,373	0,753
.
.
.

Tablo 5.9: Farklı Boyutlarda Veriler ile Eğitilen RT Algoritmasının En Yüksek Korelasyon Değerleri.

RepTree					
Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
.
.
.
3	0,6222	0,2584	0	0,6484	0
.
.
.
5	0,8354	0,8903	0	0	0
.
.
.
7	0	0	0	0	0,5093
.
.
.
35	0,3666	0	0,594	0	0
.
.
.

Büyük boyutlu eğitim verileri ile eğitilen algoritmaların düşük boyutlu eğitim verileri ile eğitilen algoritmalarından daha düşük korelasyon değeri üretmesi beklenmektedir. Fakat algoritma sonuçlarına bakıldığında bazı durumlarda beklenti gerçekleşmemekte, düşük boyutlu veri ile eğitilen algoritmaların yüksek korelasyon değeri ürettiği görülmektedir. Yüksek korelasyon değeri üreten algoritmalarımız olduğundan dolayı hataların sistemimizden kaynaklı olmadığı, hasta verisindeki hatalar yada veri eksikliğinden dolayı düşük veri ile eğitilen algoritmaların yüksek veri ile eğitilen algoritmalarından daha yüksek korelasyon değeri üretmektedir. Bu durum Hasta5' in M5P algoritması eğitim verileri üçüncü bölüm Şekil 3.18' de gösterildiği gibi parçalara ayrılarak her parçanın insülin verisi, kan şekeri ölçüm verisi, yemek verisi, aktivite verisi ve her parça ile eğitilen M5P algoritmasının test verisi ile yapılan test sonucundaki korelasyon değerleri Tablo 5.10' da verilmektedir.

Tablo 5.10: Hasta5' in Eğitim Verisi Parçalarının Veri Miktarı ve M5P Algoritması Korelasyon Değerleri.

Veri Parçası	İnsülin Verisi Sayısı	Kan Şekeri Ölçümü Sayısı	Yemek Verisi Sayısı	Aktivite Verisi Sayısı	Korelasyon Değeri
1. Parça	14	15	0	0	0,4237
2. Parça	19	19	0	0	0,792
3. Parça	8	9	0	0	-0,0733
4. Parça	14	14	0	0	0,4721
5. Parça	14	14	0	0	0,6437
6. Parça	16	14	0	0	0,4976
7. Parça	15	14	0	0	0,8032
8. Parça	16	14	0	0	0,7858
9. Parça	14	14	1	0	0,8152

Tablo 5.10 incelendiğinde en yüksek korelasyon değerine sahip olan parça 9. Parçadır. Bu parçada 14 insülin, 14 kan şekeri ölçüm ve 1 yemek verisi bulunmaktadır. İkinci en yüksek korelasyona sahip parça 7. Parçadır. Bu parçada 15 insülin ve 14 kan şekeri ölçüm verisi bulunmaktadır. Üçüncü en yüksek korelasyon değerine sahip veri parçası 2. Parçadır. Bu parçada 19 insülin ve 19 kan şekeri ölçüm verisi bulunmaktadır. En düşük korelasyon değerine sahip parça ise 3. Parçadır. Bu parçada da 8 insülin ve 9 kan şekeri ölçüm verisi bulunmaktadır. Bu sonuçlar incelendiğinde korelasyon

değeri en yüksek parçanın veri çeşitliliğinin diğer parçalardan daha fazla olduğu görülmektedir. Veri çeşitliliğinin fazla olması kan şekeri değişiminin daha iyi örneklenmesine sebep olmakta, algoritmanın oluşturduğu model de daha doğru tahmin üretebilecek bir model olmakta ve dolayısı ile de daha doğru tahminler üretilmektedir. Veri çeşitliliğinin ve veri sayısının düşük olmasında ise kan şekeri değişimi iyi örneklenememekte, algoritmanın oluşturduğu model de sağlıklı tahmin üretebilecek bir model olmamaktadır. Veri çeşitliliği aynı olan fakat veri sayıları farklı olan 2. ve 7. Parçaya baktığımızda ise en yüksek veriye sahip parçanın 2. Parça olduğu görülmektedir. Bu parçanın korelasyon değeri ondan daha az veriye sahip olan 7. Parçadan daha düşük korelasyon değerine sahip olmasının sebebi ise veri içerisindeki hatalardır. Tablo 5.7’ de Hasta5’ in Data-1 korelasyon değerlerinin Data-2 korelasyon değerlerinden düşük çıktığı görülmektedir. Data-1 oluşturulurken 9 parçanın tamamı birleştirilmiş, Data-2 oluşturulurken de 7 parçanın tamamı birleştirilmiştir. Data-2’ de olmayan 9. Parça veri çeşitliliği en fazla, en fazla korelasyona sahip olan parçadır. Bu parçanın eksikliği nedeni ile Data-2’ nin korelasyon değeri Data-1’ den düşük çıkmaktadır. Data-3’ ün korelasyon değeri ise Data-2’ nin korelasyon değerinden daha yüksek çıkmıştır. Data-3’ ün korelasyon değeri ise Data-2’ nin korelasyon değerinden yüksek çıkmıştır. Data-2 içerisindeki düşük korelasyona sahip 6. Parçanın Data-2’ de olamamasından dolayı korelasyon değerinde böyle bir artış oluşmuştur. Parçaların korelasyon değerlerinin düşüklüğünün daha iyi anlaşılması için Hasta5’ in eğitim verisinin kesiti incelenmektedir. Hastanın 10.09.1990 tarihli tüm verileri Tablo 5.11’ de gösterilmektedir. Tablodaki veri kod açıklamaları üçüncü bölüm Tablo 3.1’ de bulunmaktadır.

Tablo 5.11: Hasta5’ in 10.09.1990 Tarihli Verileri.

Tarih	Saat	Kod	Data
10.09.1990	05:58	58	142
10.09.1990	06:00	33	3
10.09.1990	06:00	34	9
10.09.1990	17:34	62	265
10.09.1990	17:36	33	1
10.09.1990	17:36	34	8
10.09.1990	21:41	64	195

Tablo 5.11’ de kan şekeri ölçüm verileri Kod değeri 58, 62 ve 64 olan veriler, insülin verileri Kod değeri 33 ve 34 olan verilerdir. Bu hasta için yemek ve aktivite verisi olmadığı için bu değerler Tablo 5.10’ da görülmemektedir. Hastanın 10.09.1990 tarihli saat 05:58’ deki Data kolonundaki kan şekeri ölçümüne bakıldığında ölçümün 142 olduğu görülmektedir. Saat 6:00’ da alınan insülin dozundan sonra saat 17:34’ te kan şekeri ölçümü yapılmış ve bu saate yapılan ölçüm sonucu 265 olarak görülmüştür. Bu bölümden yapılabilecek çıkarım insülin dozunun kan şekeri üzerinde artırıcı etkisinin olduğudur. Gün içerisindeki saat 17:36’ da alınan insülin dozundan sonra saat 21:41 deki kan şekeri ölçümünde ölçüm daha düşük kaydedilmiştir. Bu bölümden yapılabilecek çıkarım ise alınan insülinin kan şekeri değeri üzerinde azaltıcı bir etkisinin olduğudur. Hastanın verilerine bakılarak insülin dozu için yapılan bu farklı iki yorumdan bir tanesinin hatalı olacağı açıktır. Bilindiği üzere insülin kan şekeri üzerinde düşürücü etkiye sahiptir [17] [42]. İlk incelemedeki kan şekeri yükselişinin sebebinin insülin kaynaklanmayacağı aşikardır. Fakat hasta verilerinde kan şekeri yükseltici etkiye sahip olacak yemek verisi de olmadığından bu bölüm hasta verisi üzerindeki hata/eksik olduğu söylenebilir. Kan şekerini tanımlamak için oluşturduğumuz öznitelikler kan şekeri ölçümü öncesindeki bu değerlerin kullanılarak oluşturulmasından dolayı bu öznitelikler ile eğitilen algoritmamızda bu tip veri eksikliklerinden dolayı yapılacak tahminlerde hatalar oluşması muhtemeldir. Örneğin, renk tanımak için eğitilmiş bir algoritmamız olduğunu düşünelim. Bu algoritmayı dokuz tanesi kırmızı renkte, bir tanesi beyaz renkte olan eğitim verisi ile eğittiğimizi düşünürsek. Eğitilmiş bu algoritmaya turuncu renkte bir test örneğinin renginin tahmini istendiğinde algoritma bu veriyi %90 kırmızı veya %10 beyaz olarak sınıflandırması muhtemeldir. Bu eğitim verilerinden beş tanesinin turuncu renkte olmasına rağmen yanlışlıkla kırmızı olarak kayıt altına aldığını düşündüğümüzde ise bu hata payının %50 olarak doğru sınıflandırılacak olan sonucu etkilediği düşünülebilir. Tablo 5.11’ de Hasta5’ in Data-1 ve Data-2 verileri ile M5P algoritmasının yaptığı tahmin sonuçları gösterilmektedir.

Tablo 5.12: Hasta5' in Data-1 ve Data-2 Verileri ile Eğitilen M5P Algoritmasının Kan Şekeri Tahmin Değerleri.

Kan Şekeri Ölçümleri	Data-1 Tahminleri	Data-2 Tahminleri
139	170,29607	135,006938
173	170,356389	136,908274
242	213,531216	193,535573
137	129,007154	135,415911
140	153,517611	135,958609
296	201,406912	188,002844
114	127,939844	132,538875
160	129,540861	157,90235
157	78,138692	125,539963
195	128,951236	162,373602
155	130,228364	165,487783
135	80,16854	133,575692
201	164,913053	136,657489

Tablo 5.12' deki tahmin değerlerine bakıldığında Data-1 ile yapılan tahminlerin Data-2 ile yapılan tahminlerden çoğunlukla daha yüksek olduğu görülmektedir. Tablo 5.10' daki eğitim verisi parçalarının korelasyon değerlerindeki düşüklük ve bu düşüklüğün sebebi olarak hasta verisindeki hatalar Hasta5' in M5P algoritmasının Data-1 üzerindeki test sonuçlarının Data-2 üzerindeki test sonuçlarından düşük çıkması olarak karşımıza çıkmaktadır.

Tablo 5.13' de tüm algoritmaların Data-1 üzerindeki testlerin korelasyon değerleri gösterilmektedir. Tablo 5.14' de de tüm algoritmaların Data-5 üzerindeki testlerin korelasyon değerleri gösterilmektedir. Sunumu ve akışı bozmamak için Tablo 5.13 ve Tablo 5.14' de ilk ve son on hastanın korelasyon değerleri verilmiştir. Ek C Tablo C-1' de tüm hastaların Data-1 testlerinin tüm algoritma sonuçları ve Ek C Tablo C-2' de tüm hastaların Data-5' testlerinin tüm algoritma sonuçları bulunmaktadır.

Tablo 5.13: Algoritmaların Data-1 ile Eğitimleri Sonucundaki Korelasyon Değerleri.

Hasta No	AR Data-1	DS Data-1	GP Data-1	LR Data-1	M5P Data-1	MP Data-1	RT Data-1	SMOR Data-1
1	-0,054	-0,01	-0,117	0,163	0,0301	-0,179	-0,01	-0,046
2	0,2866	0,2405	0,2719	0,0713	0,0704	0,3056	0,3939	0,001
3	0,6439	0,5274	0,3718	0,6329	0,6716	-0,086	0,6222	0,6512
4	-0,016	0,2237	0,1603	0,3218	0,2399	0,0113	0,2603	0,335
5	0,6686	0,8354	0,6354	0,5148	0,8537	0,3025	0,8354	0,4582
6	0,2457	-0,299	-0,181	-0,004	-0,28	-0,253	-0,087	0,0533
7	0,2434	-0,116	0,2437	0,3902	0,4692	-0,103	0	0,5336
8	0,033	0	0,2476	-0,205	-0,207	0,3463	0	-0,252
9	0,1636	0	-0,179	0,5728	-0,166	-0,187	0	0,1902
10	-0,625	-0,625	-0,62	-0,659	-0,636	-0,265	0	-0,697
.
:	:	:	:	:	:	:	:	:
.
60	0,4432	0,5637	0,372	0,2959	0,5532	0,6131	0	0,6757
61	-0,209	-0,193	-0,299	-0,07	0,0764	-0,057	0	-0,003
62	0,4067	0,3101	0,2944	0,4282	0,501	0,04	0,1553	0,324
63	0,4738	0,4772	0,3837	0,4828	0,3803	0,4114	0,5697	0,4592
64	0,6976	0,6948	0,1802	0,2604	-0,357	0,1623	0	0,3022
65	0,3092	0,2586	0,2835	0,1692	0,3459	0,1243	0,184	0,1918
66	0,5706	0	0,6177	0,5061	0	0,7833	0,2675	-0,332
67	0,37	0,4874	0,4899	0,354	0,4312	0,2575	0,5022	0,4021
68	0,2507	0,2455	0,2514	0,2581	0,2544	0,2466	0,2455	0,2485
69	0,0376	0	0,4435	0,5836	0,4307	0,1159	0	0,1543
70	0,5904	0,6779	0,3381	0,312	0,2821	0,4172	0	0,2997

Tablo 5.14: Algoritmaların Data-5 ile Eğitimleri Sonucundaki Korelasyon Değerleri.

Hasta No	AR Data-5	DS Data-5	GP Data-5	LR Data-5	M5P Data-5	MP Data-5	RT Data-5	SMOR Data-5
1	-0,143	0	-0,041	-0,08	-0,214	-0,013	0	-0,089
2	0,2393	0	0,0008	0,0441	-0,053	-0,031	0	0,0849
3	0,0158	0,0659	0,4103	0	0,6506	-0,024	0	-0,162
4	0,3006	0,2237	-0,135	0,5117	0,4823	0,5033	0	0,4346
5	0,5697	0,3142	-0,125	0	0,4237	-0,168	0	-0,077
6	0,0171	0,3269	-0,243	-0,03	0,3137	0,1374	0	-0,173
7	0,1517	0,1861	0,4223	0	0,4035	-0,033	0,5093	-0,029
8	0,0553	-0,021	-0,369	0,4316	-0,004	0,8398	0	0,5149
9	-0,453	-0,324	0,6936	0,3483	-0,324	0,5061	0	-0,112
10	0,4568	0,5515	-0,081	0	0,1186	0,5512	0	0,2725
.
.
.
60	0,2765	0,2552	0,3256	0,6439	0,1913	0,7028	0	0,6728
61	0,3292	-0,058	-0,053	0,614	0,3173	0,5982	-0,249	0,6906
62	0,2542	0,2941	0	0,7006	0,4754	0,688	0	0,6588
63	0,2168	0	-0,392	0,3576	0,4256	-0,035	0	0,5112
64	0,1638	0,1052	0,0302	0	0,1951	-0,335	0	0,2376
65	0,1174	-0,066	-0,062	0,2151	0,1057	0,0663	0	0,3626
66	0,0567	0	0,7127	-0,554	0	-0,232	0	-0,223
67	0,0665	0,289	0,1238	0,1531	0,2899	0,21	0	0,2982
68	0,1466	0	0,2898	0,2935	0,3191	0,0911	0	0,3224
69	0,4056	0,4958	0,1848	0	0	0,3821	0	0,363
70	0,1287	0,1399	0,2104	0	0,5181	-0,331	0,1399	-0,465

Hastaların Tablo 5.13’ deki algoritmanın korelasyon değeri ile Tablo 5.14’ deki aynı algoritmanın korelasyon değerleri karşılaştırıldığında Data-1 ile yapılan test sonuçları genel olarak Data-5 ile yapılan testlerden daha yüksek olduğu görülmektedir. Bu teoriden dolayı, eğitim verilerindeki hataların arındırılması sonucunda doğru tahmin oranının yükseleceği de söylenebilir. Tahmin modellerinin hatalardan arındırılması ile doğruluk oranı daha da yükseleceğinden dolayı 0,8 korelasyon değeri üzerindeki veriler yani Data-1 ve Data-2 boyutlarında yapılan tahmin işlemleri en yüksek başarıyı sağlayacaktır. Bu tahmin aralığı genişletilmek istenildiğinde Data-3 ile yapılacak tahminler de tahmin çalışmalarına dahil edilebilir. Fakat diğer veriler ile yapılan eğitim çalışmalarının her zaman düşük çıktığı görülmektedir. Bu verilerden hataların arındırılmış olması bile bu veri boyutlarının yüksek başarılı modeller eğitilmesi için yeterli olmayacaktır.

Ayrıca üçüncü bölüm Şekil 3.6' da tüm hastaların veri sayılarından görüldüğü üzere genel olarak hastaların yemek ve aktivite verileri kan şekeri ve insülin verilerinden daha azdır. Hastaların bu verilerinin az olması algoritmaların tahmin oranlarını etkilemektedir. Bazı besinlerin kan şekerini artırdığı yapılan çalışmalardan [27] [17] bilinmektedir. Ayrıca yapılan çalışmalar [43] [44] doğru egzersizin de kan şekeri üzerinde düşürücü etkisinin olduğunu göstermektedir. Bundan dolayı hasta verilerinde bu verilerin de olması durumunda algoritma tahminlerimizin daha iyi olacağı düşünülmektedir. Bu değerlerin eksikliği hasta verilerimizdeki hata olarak nitelendirilebilir.

5.3. Test-3 Deneyleri ve Sonuçları

Test-3, kan şekeri değerinin kişisel veriler kullanılmadan tahmin edilebilirliğinin araştırılması çalışmalarıdır. Bu çalışmalarda her hasta için oluşturulacak test verileri için hastaların kendi öznelik dosyaları kullanılmaktadır. Deneylerin eğitim verileri için ise hasta dışında kalan tüm hastaların öznelik verileri birleştirilerek kullanılmaktadır. Tablo 5.15' de sunumu ve akışı bozmamak için sadece her algoritma için en yüksek tahmin değerine sahip hastaların değerleri gösterilmektedir. Ek D Tablo D-1' de tüm hastaların tüm algoritmalar ile tahmin sonuçları bulunmaktadır.

Tablo 5.15: Test-3 Deneylerinde Algoritmaların En Yüksek Korelasyon Değerleri.

Hasta No	AR	DS	LR	M5P	MP	RT	SMOR
.
.
.
4	0,4052	0,4289	0,418	0,4162	0,0863	0,4036	0,422
5	0,4804	0,4088	0,5349	0,5304	0,3853	0,3244	0,5279
.
.
.
40	0,0755	0	-0,0708	0,1446	0,0472	0,6121	0,1989
.
.
.

Tablo 5.15’ de yapılmış olan tahminlerden de görüldüğü gibi algoritmaların tahmin korelasyon değerleri düşük çıkmaktadır. Hasta verilerinin birleştirilmesi ile eğitilmiş makine modellerinin en yüksek tahmin korelasyon değeri 0,6121 ile RT yani RepTree algoritması ile Hasta40’ a ait olduğu görülmektedir. Şekil 5.9’ da Hasta40’ ın RT algoritmasının karar ağacı, Şekil 5.10’ de Hasta5’ in LR algoritmasının oluşturduğu model ve Şekil 5.11’ de M5P algoritmasının oluşturduğu model gösterilmektedir.

Linear Regression Model

Glikoz =

```
-0.0458 * Feature1 +
 0.0786 * Feature3 +
 0.0592 * Feature4 +
 0.1469 * Feature5 +
 0.2188 * Feature6 +
-0.0979 * Feature7 +
 0.1903 * Feature8 +
-0.0333 * Feature9 +
-0.0454 * Feature11 +
-0.1007 * Feature12 +
 0.072 * Feature14 +
-0.0285 * Feature15 +
-0.0341 * Feature17 +
 0.0277 * Feature19 +
 0.1513 * Feature20 +
 0.1569 * Feature21 +
 0.3847 * Feature22 +
 0.2879 * Feature23 +
-5.9878 * Feature24 +
 8.2688 * Feature26 +
-2.7964 * Feature28 +
-0.1664 * Feature34 +
-0.7307 * Feature35 +
-0.5481 * Feature36 +
-0.1858 * Feature37 +
 0.3487 * Feature38 +
-10.8629 * Feature39 +
 10.7721 * Feature40 +
 7.5864 * Feature41 +
 2.6069 * Feature43 +
-16.1758 * Feature45 +
-5.3618 * Feature46 +
 3.7737 * Feature47 +
-5.2824 * Feature48 +
 0.0318 * Feature49 +
-0.0262 * Feature50 +
-0.0488 * Feature51 +
 0.0808 * Feature53 +
-0.0949 * Feature54 +
 0.0344 * Feature55 +
-0.0219 * Feature56 +
-0.0263 * Feature57 +
-0.0614 * Feature58 +
-0.1704 * Feature60 +
-0.1616 * Feature61 +
 1.8835 * Feature63 +
 1.2618 * Feature64 +
 1.6601 * Feature65 +
-0.3265
```

Şekil 5.10: Test-3 Deneylerindeki Hasta5' in LR Modeli.


```

LM num: 1
Glikoz =
-0.0549 * Feature1
+ 0.0758 * Feature3
+ 0.0607 * Feature4
+ 0.1472 * Feature5
+ 0.2139 * Feature6
- 0.092 * Feature7
+ 0.1945 * Feature8
- 0.039 * Feature9
- 0.05 * Feature11
- 0.1062 * Feature12
- 0.0161 * Feature13
+ 0.087 * Feature14
- 0.0277 * Feature17
+ 0.0277 * Feature19
+ 0.1499 * Feature20
+ 0.1554 * Feature21
+ 0.3812 * Feature22
+ 0.2934 * Feature23
+ 3.2404 * Feature27
- 3.7132 * Feature28
- 0.1686 * Feature34
- 0.7506 * Feature35
- 0.5118 * Feature36
- 0.1967 * Feature37
+ 0.3585 * Feature38
- 8.2257 * Feature40
+ 2.6788 * Feature42
+ 0.0349 * Feature49
- 0.025 * Feature50
- 0.0528 * Feature51
+ 0.0804 * Feature53
- 0.0986 * Feature54
+ 0.0179 * Feature55
- 0.0254 * Feature57
- 0.068 * Feature58
- 0.1604 * Feature60
- 0.1539 * Feature61
+ 1.6421 * Feature63
+ 1.2893 * Feature64
+ 1.6715 * Feature65
- 0.9748

```

Şekil 5.11: Test-3 Deneylerindeki Hasta5' in M5P Algoritması Modeli.

Bu çalışmada kullanılan eğitim dosyaları 13000 den fazla örneğe (instance) sahiptir. En fazla örneğe sahip eğitim verileri bu testte kullanılmıştır. Tahmin sonuçlarının fazla yüksek çıkmaması, eğitim verisi uzunluğunun artırılarak tahmin kalitesinin artırılamayacağını göstermektedir. Ayrıca eğitim verileri oluşturulurken hasta dışındaki hastaların verileri kullanıldığından dolayı kan şekeri tahmini için hastadan bağımsız verilerin kullanılması ile yüksek tahmin oranının yakalanamayacağı görülmektedir.

En yüksek tahmin oranını veren RT algoritmasının oluşturduğu karar ağacı incelendiğinde karar ağacını oluştururken Öznitelik2-12, Öznitelik14, Öznitelik16, Öznitelik18, Öznitelik20, Öznitelik22, Öznitelik23, Öznitelik28, Öznitelik31, Öznitelik33-36, Öznitelik53, Öznitelik60, Öznitelik61, Öznitelik67 ve Öznitelik70 özniteliklerini kullandığı görülmektedir. Özniteliklerin özellikleri incelendiğinde kan şekeri ölçümü, insülin değeri ve yemek değeri ile ilgili oldukları görülmektedir. RT algoritmasının tahmin işleminde bu özellikleri dikkate aldığı söylenebilir.

LR algoritmasının oluşturduğu model incelendiğinde Öznitelik1, Öznitelik3-9, Öznitelik11, Öznitelik12, Öznitelik14, Öznitelik15, Öznitelik17, Öznitelik19, Öznitelik20-24, Öznitelik26, Öznitelik28, Öznitelik34-41, Öznitelik43, Öznitelik45-51, Öznitelik53-58, Öznitelik60, Öznitelik61, Öznitelik63-65 özniteliklerini kullanarak model oluşturduğu görülmektedir. Bu öznitelikler incelendiğinde özniteliklerin kan şekeri ölçüm sonuçları, insülin değerleri, yemek değerleri, aktivite değerleri ve tanımlanamayan kan şekeri değerleri ile ilgili olduğu görülmektedir. Dolayısıyla algoritmanın tahmininde bu özelliklerin etkili olduğu söylenebilir.

M5P algoritması incelendiğinde oluşturduğu karar ağacında Öznitelik1, Öznitelik3-9, Öznitelik11-14, Öznitelik15, Öznitelik17, Öznitelik19-23, Öznitelik27, Öznitelik28, Öznitelik34-38, Öznitelik40, Öznitelik42, Öznitelik49-51, Öznitelik53-55, Öznitelik57, Öznitelik58, Öznitelik60, Öznitelik61, Öznitelik63-65 özniteliklerinin olduğu görülmektedir. Bu öznitelikler incelendiğinde özniteliklerin kan şekeri ölçüm sonuçları, insülin değerleri, yemek değerleri, aktivite değerleri ve tanımlanamayan kan şekeri değerleri ile ilgili olduğu görülmektedir. Dolayısıyla algoritmanın tahmininde bu özelliklerin etkili olduğu söylenebilir. Bu algoritma Lineer Model ürettiğinden dolayı algoritma sonuçları LR ile benzerlik göstermektedir.

Bu test ile yapılan tahmin sonuçları kişisel veriler ile yapılan test sonuçlarından daha düşük çıktığından dolayı kan şekeri tahmininde kişisel veriler kullanılarak daha iyi sonuçlar elde edildiğinden dolayı kan şekeri tahmininde kişisel verilerin kullanılması ile daha iyi sonuçlar elde edileceği söylenebilir.

Ayrıca hasaların kişisel veri dosyalarında aktivite ve yemek verilerinin kan şekeri ölçümü ve insülin verileri kadar fazla olmamasından dolayı kişisel tahminlerde algoritmaların karar ağaçlarında bu öznitelikler bulunmamaktadır. Bu çalışmada hasta verileri birleştirilerek algoritmalar eğitildiğinden dolayı algoritmaların karar ağaçlarında yemek ve aktivite öznitelikleri de bulunmaktadır.

ALTINCI BÖLÜM

SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında UCI veri deposundan alınan 70 diyabet hastasının insülin, yemek, aktivite ve kan şekeri ölçüm verileri kullanılarak kan şekeri tahmini yapılmıştır. Kan şekeri tahmini için kan şekeri ölçümlerini ifade eden öznitelikler oluşturulmuştur. Her kan şekeri ölçümü için 71 öznitelik değeri oluşturularak kan şekeri oluşumundaki özelliklerin ifade edilmesi sağlanmıştır.

Tahmin çalışmaları kişisel veriler kullanılarak ve kişisel veriler kullanılmadan eğitilen algoritmalar ile yapılmıştır. Bu testler sonucunda kişisel veriler kullanılarak eğitilen algoritmalar ile daha yüksek sonuçlar elde edilmiştir. Ayrıca bu çalışmalarda eğitim verileri belirli oranlarda bölünerek eğitim gerçekleştirilmiştir. Tahmin sonuçlarına göre 9/10, 7/10 ve 5/10 oranında veriler ile eğitilen algoritmalar daha yüksek sonuç vermektedir. Kişisel veriler ile yapılan çalışmalarda bu oranlar dikkate alınarak yapılmalıdır. Ayrıca düşük korelasyon sonuçlarının nedenleri araştırılmış ve eğitim verilerinin parçaları incelenmiştir. Eğitim verilerine karışan hataların olduğu tespit edilmiştir. Eğitim verisinin hatalardan arındırılması ile kan şekeri tahmininden yüksek başarı elde edilebilir.

Kişisel veriler kullanılmadan yapılan çalışma sonuçlarında algoritmaların tahmin oranları düşüktür. Bundan dolayı tahmin işlemlerinin kişisel veriler ile yapılması daha iyi sonuçlar alınması için önemlidir. Ayrıca bu tahmin sonuçlarında algoritmaların oluşturdukları karar ağaçlarında kan şekeri, insülin, yemek ve aktivite özellikleri olan öznitelikler bulunmaktadır. Kişisel tahmin çalışmalarında kişisel verilerde aktivite ve yemek verilerinin yeteri kadar olmamasından dolayı bu değerler kullanılamamıştır. Hasta verileri kayıt altına alınırken bu değerlerin tam ve hatadan arınık olarak alınması durumunda kan şekeri tahmini daha iyi yapılabilecektir.

KAYNAKLAR

- [1] <http://weka.sourceforge.net/doc.dev/weka/classifiers/meta/AdditiveRegression.html> (01.01.2015).
- [2] <http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/DecisionStump.html> (01.01.2015).
- [3] <http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/GaussianProcesses.html> (01.01.2015).
- [4] <http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/LinearRegression.html> (01.01.2015).
- [5] <http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/MultilayerPerceptron.html> (01.01.2015).
- [6] <http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/M5P.html> (01.01.2015).
- [7] <https://archive.ics.uci.edu/ml/datasets/Diabetes> (01.07.2014).
- [8] <http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/REPTree.html> (01.01.2015).
- [9] <http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMOreg.html> (01.01.2015).
- [10] <http://www.turkdiab.org/page.aspx?u=1&s=12> (01.01.2015).
- [11] <http://www.turkdiab.org/page.aspx?u=1&s=26> (01.01.2015).
- [12] Turhan, H., "Tip 2 Diabetes Mellitus' lu Hastalarda Tedavi Şekline ve Hastalık Süresine Göre Depresyon ve Anksiyete", *Yayımlanmış uzmanlık tezi*, Taksim Eğitim ve Araştırma Hastanesi, İstanbul, 2007.
- [13] <http://www.turkdiab.org/page.aspx?u=1&s=17> (01.01.2015).
- [14] <http://ailehekimligi.gov.tr/kronik-hastalklar/514-diyabet.html> (01.01.2015).

- [15] Kara, K. Çınar, S., “Diyabet Bakım Profili ile Metabolik Kontrol Değişkenleri Arasındaki İlişki”, *Kafkas Tıp Bilimleri Dergisi*, 1 (2), 2011. 57-63.
- [16] Kuloğlu, M., Karaoğlu, A., Atmaca, M., Özkan, Y., Tezcan, A. E., “Tip II Diyabetik Hastalarda Psikiyatrik Belirtiler ve Kan Şekerini Kontrolü”, *Düşünen Adam*, 13 (1), 19-23, 2000.
- [17] Çiftçi, H., Akbulut, G., Yıldız, E., Mercanlıgil, S. M., “Kan Şekerini Etkileyen Besinler”, 2008.
<http://sbu.saglik.gov.tr/Ekutuphane/kitaplar/t68.pdf>
- [18] <http://www.ihealthlabs.com/glucometer/wireless-smart-gluco-monitoring-system/> (01.01.2015).
- [19] <http://www.myglucohealth.net/> (01.01.2015).
- [20] <http://www.onetouch.com/> (01.01.2015).
- [21] Georga, E., Protopappas, V., Guillen, A., Fico, G., Ardigo, D., Arredondo, M. T., Exarchos, T. P., Polyzos, D., Fotiadis, D. I.: “Data Mining for Blood Glucose Prediction and Knowledge Discovery in Diabetic Patients: The METABO Diabetes Modeling and Management System”, *Proceedings of IEEE, Annual International Conference on Engineering in Medicine and Biology Society, EMBC*, 5633 – 5636, 2009.
- [22] Islam, M. S., Leech, J., Lin, C. C. Y., Chrostowski, R.: “Peak Blood Glucose Prediction Algorithm Following in a Meal Intake”, *Proceedings of IEEE, Canadian Conference on Electrical and Computer Engineering, CCECE*, 579 – 582, 2007.
- [23] Eljil, K. S., Quadah, G., Pasquer, M.: “Predicting Hypoglycemia in Diabetic Patients Using Data Mining Techniques”, *Proceedings of IEEE, 9th International Conference on Innovations in Information Technology, IIT*, 130 – 135, 2013.
- [24] Duke, D. L., Thorpe, C., Mahmoud, M., Zirie, M.: “Intelligent Diabetes Assistant: Using Machine Learning to Help Manage Diabetes”, *Proceedings of IEEE/ACS, International Conference on Computer Systems and Applications, AICCSA*, 913 – 914, 2008.
- [25] Bazaev, N., Pletenev, A., Pozhar, K.: “Classification of Factors Affecting Blood Glucose Concentration Dynamics”, *Biomedical Engineering*, Vol. 47, No. 2, 100 – 103, 2014.
- [26] Chemlal, S., Colberg, S., Satin-Smith, M., Gyuricsko, E., Hubbard, T., Scerbo, M. W., McKenzie, F. D.: “Blood Glucose Individualized Prediction for Type 2 Diabetes using iPhone Application”, *Proceedings of IEEE, 37th Annual Northeast Bioengineering Conference, NEBEC*, 1 – 2, 2011.
- [27] <http://www.turkdiab.org/haber2.aspx?h=23> (01.01.2015).

- [28] Kahn, M.: “The Diabetes Dataset”, *AI in Medicine Symposium AIM-94*, 1994, <https://archive.ics.uci.edu/ml/datasets/Diabetes> (01.07.2014).
- [29] Özekes, S.: “Veri Madenciliği Modelleri ve Uygulama Alanları”, *İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi*, 2 (3), 65-82, 2003.
- [30] Özkan, Y.: *Veri Madenciliği Yöntemleri*, 2.bs., Papatya Yayıncılık Eğitim, İstanbul, Türkiye, 2013.
- [31] Ada, N., Regresyon Analizi, 2012, <http://ormanweb.sdu.edu.tr/dersler/scarus/regresyon.pdf> (01.01.2015).
- [32] Şeker, Ş. E.: *İş Zekası ve Veri Madenciliği*, Cinius Yayınları, İstanbul, Türkiye, 2013.
- [33] Uğur, A. ve Kınacı, A. C.: “Yapay Zeka Teknikleri ve Yapay Sinir Ağları Kullanılarak Web Sayfalarının Sınıflandırılması”, TOBB Ekonomi ve Teknoloji Üniversitesi, Ankara, 2006.
- [34] Can, M.: “Yapay Sinir Ağları İle Akım Tahmini: Mahmudiye Göleti Örneği”, *Yüksek Lisans Tezi*, Fen Bilimleri Enstitüsü, İstanbul Kültür Üniversitesi, İstanbul, 2012.
- [35] <http://orange.biolab.si/> (01.01.2015).
- [36] <https://www.knime.org/> (01.01.2015).
- [37] <http://www.cs.waikato.ac.nz/ml/weka/> (01.01.2015).
- [38] Onan, A.: “Şirket İflaslarının Tahmin Edilmesinde Karar Ağacı Algoritmalarının Karşılaştırmalı Başarım Analizi”. *Bilişim Teknolojileri Dergisi*, 8 (1), 2015.
- [39] Aydın, D., İşçi, Ö.: “Doğrusal Olmayan Otoregresif Zaman Serileri Modellerinin Kestirimi”, *Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 28, 205-218, 2012.
- [40] Subaşı, H.: ”Korelasyon Hakkında Bilgi Sunumu”, 2014, <http://www.slideshare.net/HasanSuba/korelasyon> (01.01.2015).
- [41] Bozkaya, G.: “Kan Şekerini Değiştiren Etkenler Nelerdir”, 2014, <http://www.hayatisaglik.com/saglik-bilgiler/kan-sekerini-degistiren-etkenler-nelerdir.html> (01.01.2015).
- [42] <http://www.turkdiab.org/page.aspx?u=1&s=19> (01.01.2015).
- [43] Kızıltan, G., Arpacı, S.: “Fiziksel Aktivitenin Diyabetik Hastaların Antropometrik Ölçümleri ve Beslenme Durumları Üzerine Etkisi”, *Fiziksel Aktivite Beslenme ve Sağlık Kongresi Bildiri kitabı*, Başkent Üniversitesi, Ankara, 17, 2009.

- [44] Yetkin, M. K.: "Diyabette Doğru Egzersiz İlaçtır",
<http://www.diabetcemiyeti.org/c/diyabette-dogru-egzersiz-ilactir>
(01.01.2015).

EKLER

Ek A Test-1 Deneylerindeki Algoritmalarının Korelasyon Deęerleri

Tablo A-1: Tüm Hastaların Test-1 Deneylerindeki AR, DS, GP, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Deęerleri

Ek B Test-2 Deneylerindeki M5P, DS ve RT Algoritmalarının Korelasyon Deęerleri

Tablo B-1: Tüm Hastaların Farklı Boyutlardaki Eęitim Verileri ile Eęitilen M5P Algoritmasının Korelasyon Deęerleri

Tablo B-2: Tüm Hastaların Farklı Boyutlardaki Eęitim Verileri ile Eęitilen DS Algoritmasının Korelasyon Deęerleri

Tablo B-3: Tüm Hastaların Farklı Boyutlardaki Eęitim Verileri ile Eęitilen RT Algoritmasının Korelasyon Deęerleri

Ek C Tüm Algoritmaların Data-1 ve Data-5 ile Eęitimleri Sonucundaki Korelasyon Deęerleri

Tablo C-1: Tüm Hastaların Data-1 Verileri ile Eęitilen Algoritmaların Korelasyon Deęerleri

Tablo C-2: Tüm Hastaların Data-5 verileri ile Eęitilen Algoritmaların Korelasyon Deęerleri

Ek D Test-3 Deneylerindeki Algoritmaların Korelasyon Deęerleri

Tablo D-1: Tüm Hastaların Test-3 Deneylerindeki AR, DS, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Deęerleri

Ek A Test-1 Deneylerindeki Algoritmalarının Korelasyon Değerleri

Tüm hastaların Data-1 eğitim verileri ile AR, DS, GP, LR, M5P, MP, RT ve SMOR algoritmalarının eğitilmesi sonucundaki algoritmanın korelasyon değerleri Tablo A-1’ de gösterilmektedir.

Tablo A-1: Tüm Hastaların Test-1 Deneylerindeki AR, DS, GP, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Değerleri.

Hasta No	AR	DS	GP	LR	M5P	MP	RT	SMOR
1	-0,054	-0,0097	-0,1173	0,163	0,0301	-0,1789	-0,0097	-0,046
2	0,2866	0,2405	0,2719	0,0713	0,0704	0,3056	0,3939	0,001
3	0,6439	0,5274	0,3718	0,6329	0,6716	-0,0855	0,6222	0,6512
4	-0,0159	0,2237	0,1603	0,3218	0,2399	0,0113	0,2603	0,335
5	0,6686	0,8354	0,6354	0,5148	0,8537	0,3025	0,8354	0,4582
6	0,2457	-0,2989	-0,1807	-0,0037	-0,2797	-0,2534	-0,0868	0,0533
7	0,2434	-0,1157	0,2437	0,3902	0,4692	-0,1025	0	0,5336
8	0,033	0	0,2476	-0,2054	-0,2072	0,3463	0	-0,2523
9	0,1636	0	-0,179	0,5728	-0,1659	-0,1871	0	0,1902
10	-0,6254	-0,6254	-0,6199	-0,6594	-0,6357	-0,2652	0	-0,6972
11	0,3188	0,6331	-0,3625	0,6394	0,5986	0,2398	0	0,3646
12	0,6614	0,7907	-0,091	0,6241	0,8029	-0,221	-0,527	0,4669
13	0,5962	0,3707	0,4282	0,664	0,3164	0,1984	0,3707	0,4782
14	-0,0565	0,2777	0,3165	0	0,2202	0,0655	0,2777	0,1531
15	0,2954	0,2954	0,4619	-0,416	0,322	0,4264	0	-0,2497
16	0,1772	0	-0,0543	0,0748	0,3698	-0,1528	-0,2758	-0,242
17	-0,2166	0	0,6356	0,1809	0,5778	0,2292	0	0,4085
18	0,3293	0,2591	0,069	0,4468	0,0928	0,1328	-0,2467	0,45
19	0,3265	-0,0572	0,2063	0,326	0,0826	0,0273	0,2477	0,1162
20	0,018	0	-0,0846	-0,1611	-0,1172	0	0	-0,0689
21	0,2072	0	-0,0253	-0,0378	0,5498	0,0338	0	-0,0382
22	0,5585	0	0,5523	0,2265	0,2568	0,2406	0	0,4195
23	-0,016	0	-0,0475	0,5967	0,3226	0,0716	0	0,6526
24	-0,5213	0	0,2923	0,2191	-0,6178	0,2074	-0,6585	0,1749
25	0,3508	0	0,3407	0,4141	0,5111	0,2084	0	0,4946
26	-0,1567	0	0,1605	-0,1336	0,4221	-0,0823	0,3361	0,0503
27	0,2289	0,0773	0,2357	0,1547	-0,1035	0,3602	0,2763	0,2074
28	0,1897	0,2473	0,1572	0,3026	0,251	0,2911	0,1076	0,2745
29	0,0567	0,368	0,3676	0,352	0,352	0,3733	0,368	0,3971
30	0,2701	0	0,3625	0,3376	0,34	0,3617	0,1097	0,3218
31	-0,6196	0	0,2377	0,339	0,2799	0,7549	0	0,2884
32	0,3639	0,5342	0,2894	0	0,0204	0,4489	0,3603	0,3368
33	0,0804	0	-0,2161	0,055	-0,013	-0,2929	-0,2582	0,2106
34	0,537	0,5116	0,5647	0,6922	0,5583	0,4928	-0,182	0,4396
35	0,6099	0,2917	0,3511	0,3839	0,3937	0,349	0,3666	0,4694

Tablo A-1(Devam): Tüm Hastaların Test-1 Deneylerindeki AR, DS, GP, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Değerleri.

36	-0,348	0	0,0408	-0,1855	-0,0776	-0,1203	-0,3928	-0,03
37	0,0023	0	0,4707	0,0017	-0,0192	0,2096	0	-0,0197
38	0,2039	0	0,1464	0,1676	0,0251	0,3719	0	0,2931
39	0,4999	0,4304	0,4354	0,2596	0,3608	0,1585	0	0,2736
40	0,3445	0	0,3911	0	0,2387	0,3377	0	0,2866
41	0,2271	0,2379	0,1073	0,1462	0,0883	-0,4255	0,387	-0,022
42	0,4227	0,3229	0,1971	0,1847	0,2451	0,2597	0	0,2476
43	0,4573	0	0,136	0,0424	0,4088	0,1005	0	0,0605
44	0,139	0	0,4173	0,4861	0,1328	0,1851	0	0,287
45	-0,3067	0	-0,0525	-0,1458	0,5724	-0,0683	0	-0,2134
46	0,2108	0,282	0,3815	0,1981	0,2139	-0,0247	0,3479	-0,0633
47	-0,2225	0	-0,0456	-0,0887	-0,3017	-0,1252	0	0,3516
48	-0,1233	0	-0,6032	0,3402	0,2585	-0,1099	0	-0,2587
49	0,4159	0	0,0488	0,3322	0,1184	-0,2389	0	0,4958
50	0,2869	0	0,0613	0,3235	0,2173	0,0174	0,0203	0,2761
51	0,1121	0,006	0,1819	0,3201	0,3715	0,1693	0	0,0745
52	0,679	0	0,1145	0,4801	0,6907	0,2292	0	0,7037
53	0,4756	0,4756	0,2898	0,4958	0,4902	-0,0344	0,4756	0,3723
54	0,2942	0,5193	0,1434	0,4759	0,4304	0,3165	0,5193	0,4229
55	-0,1658	0	-0,0587	0,1735	0,167	0,232	0,1818	0,2712
56	0,127	0	0,3914	0,0893	0,1255	0,3653	0,0188	0,126
57	0,3049	0	-0,0464	-0,2463	-0,2502	-0,0585	0	-0,2845
58	0,6397	0,5752	0,6023	0,5374	0,5272	0,4533	0	0,5505
59	0,5898	0,4745	0,3737	0,617	0,3822	-0,0389	0	0,5123
60	0,4432	0,5637	0,372	0,2959	0,5532	0,6131	0	0,6757
61	-0,209	-0,1929	-0,2987	-0,0696	0,0764	-0,0572	0	-0,0033
62	0,4067	0,3101	0,2944	0,4282	0,501	0,04	0,1553	0,324
63	0,4738	0,4772	0,3837	0,4828	0,3803	0,4114	0,5697	0,4592
64	0,6976	0,6948	0,1802	0,2604	-0,357	0,1623	0	0,3022
65	0,3092	0,2586	0,2835	0,1692	0,3459	0,1243	0,184	0,1918
66	0,5706	0	0,6177	0,5061	0	0,7833	0,2675	-0,3319
67	0,37	0,4874	0,4899	0,354	0,4312	0,2575	0,5022	0,4021
68	0,2507	0,2455	0,2514	0,2581	0,2544	0,2466	0,2455	0,2485
69	0,0376	0	0,4435	0,5836	0,4307	0,1159	0	0,1543
70	0,5904	0,6779	0,3381	0,312	0,2821	0,4172	0	0,2997

Ek B – Test-2 Deneylerindeki M5P, DS ve RT Algoritmalarının Korelasyon Değerleri

Tüm hastaların farklı boyutlardaki eğitim verileri ile M5P algoritmasının eğitilmesi sonucundaki algoritmanın korelasyon değerleri Tablo B-1’ de gösterilmektedir.

Tablo B-1: Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen M5P Algoritmasının Korelasyon Değerleri.

Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
1	0,0301	0,136	-0,103	-0,1883	-0,2141
2	0,0704	0,0574	0,1056	0,0771	-0,0529
3	0,6716	0,4293	0,6127	-0,0374	0,6506
4	0,2399	0,2969	0,3419	0,5522	0,4823
5	0,8537	0,6691	0,7991	0,5192	0,4237
6	-0,2797	-0,1751	-0,2275	0,0964	0,3137
7	0,4692	-0,0394	0,4129	0,4508	0,4035
8	-0,2072	-0,242	0,165	0,4925	-0,0043
9	-0,1659	0,2564	0,0701	-0,2125	-0,3236
10	-0,6357	-0,6907	-0,6838	-0,7676	0,1186
11	0,5986	0,7156	0,5634	0,623	0,6679
12	0,8029	0,0852	0,3031	0,5031	0
13	0,3164	0,3834	0,2003	0,2263	0,5318
14	0,2202	0,2131	0	0	0
15	0,322	0	0,3667	0,156	-0,5602
16	0,3698	-0,1065	0	0,443	0,4021
17	0,5778	0,5256	0,505	-0,0099	0,482
18	0,0928	0,1712	-0,1368	0,1857	0,0696
19	0,0826	0,0517	-0,0695	0	-0,2312
20	-0,1172	-0,097	-0,2077	0,0794	-0,2033
21	0,5498	0,1565	-0,2131	0,0464	0,2129
22	0,2568	0,1029	0,0368	0,212	0,0368
23	0,3226	0,5415	0,2006	0,4554	0,0943
24	-0,6178	-0,5392	0,0489	-0,1374	-0,1558
25	0,5111	0,4765	0,4056	0,6154	0
26	0,4221	0,5203	0,4847	0,4522	0
27	-0,1035	-0,0587	0,2407	0,1802	0,3145
28	0,251	0,2788	0,2929	0,1396	0,3152
29	0,352	0,2535	0,2944	0,0845	0,3403
30	0,34	0,321	0,0698	0,1384	0,006
31	0,2799	0,6128	0,6064	0,6049	0,6969
32	0,0204	-0,6188	0,3938	0,608	0
33	-0,013	0,4165	0,4303	0,3607	0,1814
34	0,5583	0,5413	0,5532	0,5916	0,6809
35	0,3937	0,6014	0,348	0,2308	0,2852

Tablo B-1(Devam): Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen M5P Algoritmasının Korelasyon Değerleri.

36	-0,0776	0,2592	0,2258	-0,1164	0
37	-0,0192	-0,0642	-0,205	0,0268	0,19
38	0,0251	0,0874	0,3205	0,0301	0,3803
39	0,3608	0,4284	0,3938	0,4579	0,0786
40	0,2387	0,1999	0,1192	-0,0434	0
41	0,0883	-0,2975	-0,0982	0,2617	0,2218
42	0,2451	0,1675	0,339	0,2247	0,1074
43	0,4088	0,2055	0,177	-0,1772	0,1514
44	0,1328	0,5027	0,5297	0,5412	0,5212
45	0,5724	0,3094	0,7515	-0,1302	0
46	0,2139	0,0325	0,0682	0,6745	0
47	-0,3017	-0,3319	-0,141	0,1439	0,1359
48	0,2585	-0,4255	0,4899	0,2084	0,3349
49	0,1184	0,2645	0,1135	0,156	0
50	0,2173	0,1412	0,038	-0,0248	-0,0671
50	0,2173	0,1412	0,038	-0,0248	-0,0671
51	0,3715	0,3513	0,4476	-0,1258	0,1853
52	0,6907	0,6735	0,5959	0,3434	0,0336
53	0,4902	0,4527	-0,0484	0,0049	0
54	0,4304	0,4892	0,4392	0,5341	0,4787
55	0,167	-0,0703	0,1237	-0,1088	0,234
56	0,1255	0,1477	0,0936	0,1893	0,1675
57	-0,2502	0,0579	-0,2447	0,0278	-0,1857
58	0,5272	0,5402	0,616	0,5599	0,5117
59	0,3822	0,3135	0,6698	0,3705	0,2516
60	0,5532	0,2477	0,5577	0,4506	0,1913
61	0,0764	0,4705	0,0344	0,0603	0,3173
62	0,501	0,4367	0,4472	0,4512	0,4754
63	0,3803	0,4852	0,473	0,4633	0,4256
64	-0,357	0,2028	-0,4134	0,066	0,1951
65	0,3459	0,3488	0,0932	0,113	0,1057
66	0	0,3682	0,7387	0,6658	0
67	0,4312	0,4537	0,5417	0,2135	0,2899
68	0,2544	0,2457	0,2752	0,2066	0,3191
69	0,4307	0	0	-0,3096	0
70	0,2821	0,5273	0,3349	0,532	0,5181

Tüm hastaların farklı boyutlardaki eğitim verileri ile DS algoritmasının eğitilmesi sonucunda algoritmanın korelasyon değerleri Tablo B-2’ de gösterilmektedir.

Tablo B-2: Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen DS Algoritmasının Korelasyon Değerleri.

Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
1	-0,0097	-0,0097	0	0	0
2	0,2405	0	0,028	0	0
3	0,5274	0,5274	0,5274	0,6484	0,0659
4	0,2237	0,2237	0,2237	0,2237	0,2237
5	0,8354	0,8354	0,8354	0,1695	0,3142
6	-0,2989	-0,2989	-0,2989	0	0,3269
7	-0,1157	-0,1157	-0,05	-0,05	0,1861
8	0	-0,423	-0,0087	-0,1951	-0,0206
9	0	-0,0928	-0,0928	-0,3236	-0,3236
10	-0,6254	-0,6254	-0,6254	-0,6254	0,5515
11	0,6331	0,3743	0	0,3743	0,3278
12	0,7907	0,7907	0,7907	0,3402	0,281
13	0,3707	0,3707	0,1088	0,1088	0,5329
14	0,2777	0,2777	0,2777	0,2777	0,1361
15	0,2954	0,2954	0,0582	0,0582	-0,6242
16	0	-0,2758	0,2758	0	0,3864
17	0	-0,0479	0	0,2805	0,3988
18	0,2591	0	0	0	0,3177
19	-0,0572	0,2477	0	0,2477	-0,2503
20	0	0	0	0,0077	-0,0432
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0,4377	-0,1248
24	0	0	0	-0,2695	0
25	0	0	0,2674	0	0
26	0	0	0	0	0,4458
27	0,0773	0,0773	0,0958	0,2198	0
28	0,2473	0,2473	0,2473	0,2473	0,2473
29	0,368	0,368	0,368	0,368	0,368
30	0	0	0	0,1554	0,1554
31	0	0	0	0	0
32	0,5342	0	0	0,5342	0,2983
33	0	0	0	0	-0,1137
34	0,5116	0,5116	0,5116	0,5116	0,1592
35	0,2917	0,2917	0	0,3865	0,3498

Tablo B-2(Devam): Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen DS Algoritmasının Korelasyon Değerleri.

36	0	0	0	0	0
37	0	0,0464	0,0464	0,2494	0,2494
38	0	0	0	0	0
39	0,4304	0,4304	0,4304	0,0947	0,3834
40	0	0	0	0	0
41	0,2379	0,2379	0,3283	0,098	-0,1124
42	0,3229	0,0766	0,0554	0,0114	0
43	0	0	0	0,0801	0,1783
44	0	0	0	0	0
45	0	0	-0,0957	-0,0957	0
46	0,282	0,282	0,5893	0,373	0,753
47	0	-0,2098	-0,2098	-0,1494	-0,2098
48	0	0	0	0,018	0,3302
49	0	0	0	0	-0,1556
50	0	0,0203	0,2433	0,0203	0,1386
51	0,006	0,006	0	0,006	0
52	0	0,2896	0,2896	0,2703	0
53	0,4756	0	0,1668	0,0848	0
54	0,5193	0,5193	0,5193	0,414	-0,2751
55	0	0	0	0,1462	-0,0132
56	0	0	0,0188	0,2513	-0,0245
57	0	-0,1985	0	0,2801	0,1516
58	0,5752	0,5752	0,5752	0,5752	0,4779
59	0,4745	0,4745	0,4745	0,2529	-0,3702
60	0,5637	0,4633	0,5843	0	0,2552
61	-0,1929	0,2372	-0,1929	0,2372	-0,0578
62	0,3101	0,3101	0,1553	0,1553	0,2941
63	0,4772	0,5305	0,4772	0,5305	0
64	0,6948	0,6948	0	0,1052	0,1052
65	0,2586	0,1417	0,1417	0,0743	-0,066
66	0	0	0	0,6007	0
67	0,4874	0,4874	0,4874	0,4874	0,289
68	0,2455	0,2455	0,2455	0,1784	0
69	0	0	0	0	0,4958
70	0,6779	0,6779	0,1399	-0,3154	0,1399

Tüm hastaların farklı boyutlardaki eğitim verileri ile RT algoritmasının eğitilmesi sonucunda algoritmanın korelasyon değerleri Tablo B-3' de gösterilmektedir.

Tablo B-3: Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen RT Algoritmasının Korelasyon Değerleri.

Hasta No	Data-1	Data-2	Data-3	Data-4	Data-5
1	-0,0097	-0,0097	0	0	0
2	0,3939	0,0305	0,1951	0	0
3	0,6222	0,2584	0	0,6484	0
4	0,2603	0,0949	0,2237	0	0
5	0,8354	0,8903	0	0	0
6	-0,0868	0,3766	0	0	0
7	0	0	0	0	0,5093
8	0	0	0	0	0
9	0	0	0	0,0053	0
10	0	0	-0,5223	0	0
11	0	0	0	0	0
12	-0,527	0	0	0	0
13	0,3707	0	0	0,3937	0
14	0,2777	0	0	0	0
15	0	0	0	0,0582	0
16	-0,2758	0	0	0	0
17	0	0,1466	0	0	0,3988
18	-0,2467	0	0	0,2554	0
19	0,2477	0	0	0	0
20	0	-0,2926	0	-0,0667	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	-0,3089	-0,2696	0
24	-0,6585	0	0	0	0
25	0	0	0	0	0
26	0,3361	0	0,0835	0	0
27	0,2763	0,2929	0,0503	0,2198	0,2089
28	0,1076	0,2473	0,2473	0,1445	0,2473
29	0,368	0,2949	0,3659	0,368	0
30	0,1097	0	0	-0,3989	0,1287
31	0	0	0	0	0
32	0,3603	0	0	0	0
33	-0,2582	0,4959	0,0469	0	0
34	-0,182	0,5116	0	0	0,4771
35	0,3666	0	0,594	0	0

Tablo B-3(Devam): Tüm Hastaların Farklı Boyutlardaki Eğitim Verileri ile Eğitilen RT Algoritmasının Korelasyon Değerleri.

36	-0,3928	0	0	0	0
37	0	0	0	-0,1546	0
38	0	0	0,2192	0	0
39	0	0	0,3907	0	0,2872
40	0	0	0	0	0
41	0,387	0,2567	0,3283	0,0503	0
42	0	0,1461	0,0114	0,199	0,285
43	0	0	0,4388	0	0
44	0	-0,0507	0,2073	0,1695	0
45	0	0	0	0	0
46	0,3479	-0,0448	0,5893	0	0
47	0	-0,2098	-0,2098	0	0
48	0	0	0	0,018	0
49	0	0	0	0	0
50	0,0203	0	0,1531	0	0,1386
51	0	0	0	0	0
52	0	0,6759	0,4227	0,2703	0
53	0,4756	0	0	0	-0,2313
54	0,5193	-0,0281	0,5193	0,3845	0,1254
55	0,1818	-0,3802	0	-0,0132	0
56	0,0188	0,2768	-0,0245	0	0
57	0	0	0	0	0
58	0	0	0	0,5752	0
59	0	0	0	0	-0,3133
60	0	0	0	0,5248	0
61	0	0,0065	-0,0578	0	-0,2489
62	0,1553	0,3101	0,1553	0	0
63	0,5697	0,323	0,5101	0,5305	0
64	0	0	0	0	0
65	0,184	0	0,1417	0,1213	0
66	0,2675	0	0	0,6007	0
67	0,5022	0,5885	0,4676	0,5411	0
68	0,2455	0,3063	0,0755	0	0
69	0	0	0	0	0
70	0	0,7141	0,3972	0	0,1399

Ek C – Tüm Algoritmaların Data-1 ve Data-5 ile Eğitimleri Sonucundaki Korelasyon Değerleri

Tüm hastaların Data-1 eğitim verileri ile tüm algoritmaların eğitilmesi sonucunda algoritmaların korelasyon değerleri Tablo C-1’ de gösterilmektedir

Tablo C-1: Tüm Hastaların Data-1 Verileri ile Eğitilen Algoritmaların Korelasyon Değerleri.

Hasta No	AR Data-1	DS Data-1	GP Data-1	LR Data-1	M5P Data-1	MP Data-1	RT Data-1	SMOR Data-1
1	-0,054	-0,01	-0,117	0,163	0,0301	-0,179	-0,01	-0,046
2	0,2866	0,2405	0,2719	0,0713	0,0704	0,3056	0,3939	0,001
3	0,6439	0,5274	0,3718	0,6329	0,6716	-0,086	0,6222	0,6512
4	-0,016	0,2237	0,1603	0,3218	0,2399	0,0113	0,2603	0,335
5	0,6686	0,8354	0,6354	0,5148	0,8537	0,3025	0,8354	0,4582
6	0,2457	-0,299	-0,181	-0,004	-0,28	-0,253	-0,087	0,0533
7	0,2434	-0,116	0,2437	0,3902	0,4692	-0,103	0	0,5336
8	0,033	0	0,2476	-0,205	-0,207	0,3463	0	-0,252
9	0,1636	0	-0,179	0,5728	-0,166	-0,187	0	0,1902
10	-0,625	-0,625	-0,62	-0,659	-0,636	-0,265	0	-0,697
11	0,3188	0,6331	-0,363	0,6394	0,5986	0,2398	0	0,3646
12	0,6614	0,7907	-0,091	0,6241	0,8029	-0,221	-0,527	0,4669
13	0,5962	0,3707	0,4282	0,664	0,3164	0,1984	0,3707	0,4782
14	-0,057	0,2777	0,3165	0	0,2202	0,0655	0,2777	0,1531
15	0,2954	0,2954	0,4619	-0,416	0,322	0,4264	0	-0,25
16	0,1772	0	-0,054	0,0748	0,3698	-0,153	-0,276	-0,242
17	-0,217	0	0,6356	0,1809	0,5778	0,2292	0	0,4085
18	0,3293	0,2591	0,069	0,4468	0,0928	0,1328	-0,247	0,45
19	0,3265	-0,057	0,2063	0,326	0,0826	0,0273	0,2477	0,1162
20	0,018	0	-0,085	-0,161	-0,117	0	0	-0,069
21	0,2072	0	-0,025	-0,038	0,5498	0,0338	0	-0,038
22	0,5585	0	0,5523	0,2265	0,2568	0,2406	0	0,4195
23	-0,016	0	-0,048	0,5967	0,3226	0,0716	0	0,6526
24	-0,521	0	0,2923	0,2191	-0,618	0,2074	-0,659	0,1749
25	0,3508	0	0,3407	0,4141	0,5111	0,2084	0	0,4946
26	-0,157	0	0,1605	-0,134	0,4221	-0,082	0,3361	0,0503
27	0,2289	0,0773	0,2357	0,1547	-0,104	0,3602	0,2763	0,2074
28	0,1897	0,2473	0,1572	0,3026	0,251	0,2911	0,1076	0,2745
29	0,0567	0,368	0,3676	0,352	0,352	0,3733	0,368	0,3971
30	0,2701	0	0,3625	0,3376	0,34	0,3617	0,1097	0,3218
31	-0,62	0	0,2377	0,339	0,2799	0,7549	0	0,2884
32	0,3639	0,5342	0,2894	0	0,0204	0,4489	0,3603	0,3368
33	0,0804	0	-0,216	0,055	-0,013	-0,293	-0,258	0,2106
34	0,537	0,5116	0,5647	0,6922	0,5583	0,4928	-0,182	0,4396
35	0,6099	0,2917	0,3511	0,3839	0,3937	0,349	0,3666	0,4694

Tablo C-1(Devam): Tüm Hastaların Data-1 Verileri ile Eğitilen Algoritmaların Korelasyon Değerleri.

36	-0,348	0	0,0408	-0,186	-0,078	-0,12	-0,393	-0,03
37	0,0023	0	0,4707	0,0017	-0,019	0,2096	0	-0,02
38	0,2039	0	0,1464	0,1676	0,0251	0,3719	0	0,2931
39	0,4999	0,4304	0,4354	0,2596	0,3608	0,1585	0	0,2736
40	0,3445	0	0,3911	0	0,2387	0,3377	0	0,2866
41	0,2271	0,2379	0,1073	0,1462	0,0883	-0,426	0,387	-0,022
42	0,4227	0,3229	0,1971	0,1847	0,2451	0,2597	0	0,2476
43	0,4573	0	0,136	0,0424	0,4088	0,1005	0	0,0605
44	0,139	0	0,4173	0,4861	0,1328	0,1851	0	0,287
45	-0,307	0	-0,053	-0,146	0,5724	-0,068	0	-0,213
46	0,2108	0,282	0,3815	0,1981	0,2139	-0,025	0,3479	-0,063
47	-0,223	0	-0,046	-0,089	-0,302	-0,125	0	0,3516
48	-0,123	0	-0,603	0,3402	0,2585	-0,11	0	-0,259
49	0,4159	0	0,0488	0,3322	0,1184	-0,239	0	0,4958
50	0,2869	0	0,0613	0,3235	0,2173	0,0174	0,0203	0,2761
51	0,1121	0,006	0,1819	0,3201	0,3715	0,1693	0	0,0745
52	0,679	0	0,1145	0,4801	0,6907	0,2292	0	0,7037
53	0,4756	0,4756	0,2898	0,4958	0,4902	-0,034	0,4756	0,3723
54	0,2942	0,5193	0,1434	0,4759	0,4304	0,3165	0,5193	0,4229
55	-0,166	0	-0,059	0,1735	0,167	0,232	0,1818	0,2712
56	0,127	0	0,3914	0,0893	0,1255	0,3653	0,0188	0,126
57	0,3049	0	-0,046	-0,246	-0,25	-0,059	0	-0,285
58	0,6397	0,5752	0,6023	0,5374	0,5272	0,4533	0	0,5505
59	0,5898	0,4745	0,3737	0,617	0,3822	-0,039	0	0,5123
60	0,4432	0,5637	0,372	0,2959	0,5532	0,6131	0	0,6757
61	-0,209	-0,193	-0,299	-0,07	0,0764	-0,057	0	-0,003
62	0,4067	0,3101	0,2944	0,4282	0,501	0,04	0,1553	0,324
63	0,4738	0,4772	0,3837	0,4828	0,3803	0,4114	0,5697	0,4592
64	0,6976	0,6948	0,1802	0,2604	-0,357	0,1623	0	0,3022
65	0,3092	0,2586	0,2835	0,1692	0,3459	0,1243	0,184	0,1918
66	0,5706	0	0,6177	0,5061	0	0,7833	0,2675	-0,332
67	0,37	0,4874	0,4899	0,354	0,4312	0,2575	0,5022	0,4021
68	0,2507	0,2455	0,2514	0,2581	0,2544	0,2466	0,2455	0,2485
69	0,0376	0	0,4435	0,5836	0,4307	0,1159	0	0,1543
70	0,5904	0,6779	0,3381	0,312	0,2821	0,4172	0	0,2997

Tüm hastaların Data-5 eğitim verileri ile tüm algoritmaların eğitilmesi sonucunda algoritmaların korelasyon değerleri Tablo C-2’ de gösterilmektedir.

Tablo C-2: Tüm Hastaların Data-5 Verileri ile Eğitilen Algoritmaların Korelasyon Değerleri.

Hasta No	AR Data-5	DS Data-5	GP Data-5	LR Data-5	M5P Data-5	MP Data-5	RT Data-5	SMOR Data-5
1	-0,143	0	-0,041	-0,08	-0,214	-0,013	0	-0,089
2	0,2393	0	0,0008	0,0441	-0,053	-0,031	0	0,0849
3	0,0158	0,0659	0,4103	0	0,6506	-0,024	0	-0,162
4	0,3006	0,2237	-0,135	0,5117	0,4823	0,5033	0	0,4346
5	0,5697	0,3142	-0,125	0	0,4237	-0,168	0	-0,077
6	0,0171	0,3269	-0,243	-0,03	0,3137	0,1374	0	-0,173
7	0,1517	0,1861	0,4223	0	0,4035	-0,033	0,5093	-0,029
8	0,0553	-0,021	-0,369	0,4316	-0,004	0,8398	0	0,5149
9	-0,453	-0,324	0,6936	0,3483	-0,324	0,5061	0	-0,112
10	0,4568	0,5515	-0,081	0	0,1186	0,5512	0	0,2725
11	0,4825	0,3278	-0,209	0,432	0,6679	0,6044	0	0,628
12	0,095	0,281	-0,389	0	0	0,0369	0	0,2371
13	0,4172	0,5329	0,5904	0,6856	0,5318	0,5177	0	0,6497
14	0,0639	0,1361	0,4814	-0,289	0	0,201	0	-0,045
15	-0,46	-0,624	-0,733	0	-0,56	-0,089	0	-0,139
16	0,3259	0,3864	0,0525	0	0,4021	0,2106	0	0,2774
17	0,3586	0,3988	0,1326	0,4983	0,482	0,4381	0,3988	0,5572
18	0,4536	0,3177	0,0315	0,2224	0,0696	0,1989	0	0,1939
19	0,0629	-0,25	0	0,1652	-0,231	0,3021	0	0,1658
20	-0,041	-0,043	-0,096	0,0853	-0,203	0,0535	0	-0,147
21	0,0228	0	-0,044	0,2101	0,2129	0,0287	0	0,3108
22	0,2702	0	-0,236	0	0,0368	-0,008	0	0,2324
23	0,105	-0,125	-0,043	0	0,0943	-0,008	0	0,2542
24	-0,415	0	-0,072	-0,446	-0,156	-0,407	0	-0,639
25	-0,408	0	0,2056	0	0	0,3209	0	0,4534
26	0,2769	0,4458	0,4397	0,4059	0	0,0488	0	0,4534
27	-0,029	0	-0,249	0,2535	0,3145	0,0758	0,2089	0,2166
28	-0,03	0,2473	-0,143	0,2216	0,3152	-0,307	0,2473	0,16
29	0,0771	0,368	0,2705	0,2554	0,3403	-0,103	0	0,2697
30	0,1082	0,1554	0,332	0,006	0,006	0,0814	0,1287	0,0314
31	-0,576	0	-0,545	0,6732	0,6969	-0,58	0	0,7362
32	0,1264	0,2983	0,0334	0,0133	0	-0,081	0	0,4095
33	-0,044	-0,114	-0,048	0,2523	0,1814	0,0901	0	0,2643
34	0,5867	0,1592	-0,232	0,3962	0,6809	0,3466	0,4771	0,4426
35	-0,036	0,3498	0,2716	0,3489	0,2852	0,2654	0	0,3755

Tablo C-2(Devam): Tüm Hastaların Data-5 Verileri ile Eğitilen Algoritmaların Korelasyon Değerleri.

36	0,1148	0	0,234	0	0	0,2301	0	0,0198
37	0,4005	0,2494	-0,009	0	0,19	0,2133	0	0,214
38	0,2092	0	-0,13	0,497	0,3803	0,4599	0	0,3351
39	0,4591	0,3834	0,0417	0	0,0786	-0,001	0,2872	-0,101
40	0	0	-0,265	0	0	0,1744	0	0,1731
41	0,0539	-0,112	-0,329	0,1348	0,2218	0,2826	0	0,2816
42	0,3943	0	0,3358	0,1074	0,1074	0,2124	0,285	0,3041
43	0,0561	0,1783	-0,197	0,3757	0,1514	0,2021	0	0,3572
44	0,0256	0	-0,385	0,2196	0,5212	0,2553	0	0,2686
45	0,4001	0	0	-0,134	0	-0,247	0	-0,177
46	0,6909	0,753	0,6989	0	0	0,435	0	0,7052
47	-0,399	-0,21	-0,13	0,1985	0,1359	0,2254	0	0,178
48	0,2611	0,3302	-0,172	0	0,3349	-0,152	0	0,0074
49	-0,177	-0,156	0,3281	0	0	0,0183	0	0,0548
50	0,2348	0,1386	-0,226	0	-0,067	0,1541	0,1386	-0,097
51	0,2146	0	0,1268	0	0,1853	-0,103	0	-0,029
52	-0,242	0	0,2841	0	0,0336	-0,26	0	0,4477
53	-0,312	0	0,2207	0	0	-0,209	-0,231	-0,43
54	-0,252	-0,275	-0,004	0,2773	0,4787	0,1806	0,1254	0,2223
55	-0,081	-0,013	-0,056	-0,025	0,234	-0,027	0	0,0937
56	-0,117	-0,025	0,308	0,269	0,1675	0,2266	0	0,2591
57	0,705	0,1516	-0,122	0	-0,186	0,5811	0	0,4987
58	0,6072	0,4779	0,5074	0,4383	0,5117	0,4556	0	0,3999
59	-0,018	-0,37	0	-0,079	0,2516	0,1686	-0,313	-0,178
60	0,2765	0,2552	0,3256	0,6439	0,1913	0,7028	0	0,6728
61	0,3292	-0,058	-0,053	0,614	0,3173	0,5982	-0,249	0,6906
62	0,2542	0,2941	0	0,7006	0,4754	0,688	0	0,6588
63	0,2168	0	-0,392	0,3576	0,4256	-0,035	0	0,5112
64	0,1638	0,1052	0,0302	0	0,1951	-0,335	0	0,2376
65	0,1174	-0,066	-0,062	0,2151	0,1057	0,0663	0	0,3626
66	0,0567	0	0,7127	-0,554	0	-0,232	0	-0,223
67	0,0665	0,289	0,1238	0,1531	0,2899	0,21	0	0,2982
68	0,1466	0	0,2898	0,2935	0,3191	0,0911	0	0,3224
69	0,4056	0,4958	0,1848	0	0	0,3821	0	0,363
70	0,1287	0,1399	0,2104	0	0,5181	-0,331	0,1399	-0,465

Ek D – Test-3 Deneylerindeki Algoritmaların Korelasyon Değerleri

Her hasta için diğer hastaların verileri kullanılarak eğitim verileri oluşturulmuştur. Tüm hastaların eğitim verileri ile eğitilen algoritmaların korelasyon değerleri Tablo D-1’ de gösterilmektedir.

Tablo D-1: Tüm Hastaların Test-3 Deneylerindeki AR, DS, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Değerleri.

Hasta No	AR	DS	LR	M5P	MP	RT	SMOR
1	0,0345	0,0263	0,085	0,0853	0,0965	0,012	0,0964
2	-0,0442	0,007	0,099	0,0989	-0,0922	0,1187	0,0892
3	0,4472	0,3515	0,4018	0,4113	0,233	0,3978	0,4069
4	0,4052	0,4289	0,418	0,4162	0,0863	0,4036	0,422
5	0,4804	0,4088	0,5349	0,5304	0,3853	0,3244	0,5279
6	0,1342	0,0639	0,2003	0,2007	0,3074	0,2208	0,2184
7	0,1632	0,1022	-0,0115	-0,0111	-0,1311	0,0384	-0,0227
8	-0,0271	-0,009	-0,0606	-0,0485	-0,0957	-0,0932	-0,1293
9	-0,1633	-0,1364	-0,012	-0,0067	0,0283	0,0015	-0,0341
10	0,0454	0,0748	0,113	0,1211	0,0317	-0,0055	0,1134
11	0,1493	-0,0507	0,2206	0,2156	0,2257	0,057	0,1683
12	0,1381	0,1113	0,3364	0,3376	-0,211	0,0621	0,3178
13	0,217	0,2468	0,3936	0,3987	0,067	0,192	0,4061
14	0,0758	0,1222	0,0912	0,0997	0,0052	-0,0236	0,0433
15	-0,0183	-0,0493	0,0121	0,0234	-0,0549	-0,2202	0,031
16	-0,1555	-0,0487	-0,1431	-0,128	-0,0571	-0,0298	-0,1321
17	0,133	0,0849	0,1142	0,105	0,0744	0,0945	0,1287
18	0,218	0,2323	0,1876	0,1868	0,0074	0,1927	0,2033
19	0,0672	0,0281	0,1059	0,0624	0,0411	0,0987	0,1527
20	0,1038	0,0652	0,1219	0,1216	0,0851	0,0971	0,1269
21	0,0173	-0,0513	-0,0084	-0,0083	0,0818	-0,0549	-0,0173
22	-0,1022	-0,0973	0,0363	0,0356	-0,017	-0,0924	-0,0133
23	-0,0381	-0,0037	0,0304	0,0428	0,1132	0,0716	0,104
24	-0,0475	-0,0756	0,0112	0,0093	-0,1268	0,0066	0,0024
25	0,0088	0	0,0658	0,1048	-0,1464	0,0327	0,0085
26	0,1575	0,166	0,0985	0,0952	0,1124	0,0359	0,1128
27	0,2635	0	0,2403	0,2402	0,0971	0,2602	0,2539
28	0,2996	0,3276	0,3468	0,3467	0,2651	0,3258	0,3437
29	0,3145	0,293	0,1702	0,1741	0,1431	0,1087	0,1678
30	0,252	0,2318	0,2055	0,207	0,1002	0,2153	0,1913
31	0,1283	0	0,127	0,1282	0,0253	0	0,1349
32	-0,1149	-0,0315	-0,167	-0,1823	-0,0684	0,0134	-0,1871
33	0,0783	0,0996	0,1878	0,1753	0,1361	0,0745	0,1831
34	0,3017	0,3196	0,2704	0,2495	0,1472	0,1818	0,264
35	0,2735	0,2086	0,2497	0,2781	0,0813	0,3246	0,2345

Tablo C-1(Devam): Tüm Hastaların Test-3 Deneylerindeki AR, M5P, DS, LR, M5P, MP, RT ve SMOR Algoritmalarının Korelasyon Değerleri.

36	0,1605	0,1525	0,1543	0,1541	-0,1057	0,0645	0,1497
37	0,0954	0,1485	0,1524	0,1605	0,0828	0,0219	0,1956
38	0,0401	0,0632	0,1152	0,0849	-0,1904	-0,0742	0,0999
39	0,1651	0,1759	0,2364	0,2355	0,0316	0,275	0,2667
40	0,0755	0	-0,0708	0,1446	0,0472	0,6121	0,1989
41	0,3341	0,3589	0,3025	0,2993	0,2913	0,0561	0,312
42	0,2223	0,1955	0,2591	0,2644	-0,0043	0,2379	0,2571
43	0,0477	0,1053	0,0101	-0,0243	0,099	0,0455	-0,0537
44	-0,0839	-0,0246	0,1427	0,1482	0,1268	0,073	0,1557
45	0,0019	0,1406	0,1605	0,1561	0,093	0,1583	0,1701
46	0,321	0,2959	0,3755	0,3786	-0,0969	0,2968	0,3624
47	0,191	0,1301	0,0497	0,0491	-0,0745	0,0334	0,0467
48	0,1246	0,0835	0,1652	0,1659	0,1361	0,1143	0,1563
49	-0,0007	-0,0519	0,0612	0,0296	0,1034	-0,0195	0,0542
50	0,0788	0,0314	0,0875	0,0933	-0,1904	0,0052	0,1832
51	0,2275	0,1211	0,2377	0,2438	0,1135	0,2392	0,1817
52	0,2972	0,3074	0,2943	0,3042	0,0094	0,4215	0,3502
53	0,2244	0,1959	0,0377	0,034	-0,0189	0,1296	0,063
54	0,2886	0,2344	0,3027	0,3126	0,1949	0,2551	0,2875
55	0,2531	0,0664	0,1783	0,1798	0,161	0,2077	0,2024
56	0,2504	0,2122	0,2642	0,2642	0,1923	0,2237	0,2336
57	0,0961	0,1084	-0,0215	-0,0206	-0,1312	0,1151	0,0094
58	0,2299	0,1218	0,3071	0,2996	0,0671	0,2767	0,266
59	0,1972	0,1818	0,2877	0,2778	0,1586	0,2343	0,246
60	0,3365	0,2251	0,2037	0,2342	0,0316	0,2832	0,1495
61	0,0302	-0,0606	0,0249	0,0212	0,1589	0,0045	0,0291
62	0,4439	0,1719	0,4379	0,4532	0,09	0,2366	0,43
63	0,1373	0,1747	0,2432	0,2451	0,1529	0,1414	0,2793
64	0,0737	-0,0505	0,1283	0,1371	0,0627	0,137	0,0939
65	0,0107	-0,0183	0,1596	0,16	0,0477	0,0576	0,1516
66	-0,0224	0,0579	0,0065	-0,1036	-0,1056	-0,0406	-0,0457
67	0,3912	0,1958	0,4778	0,4793	0,2621	-0,3206	0,4583
68	0,2519	0,0399	0,2731	0,2726	0,0492	0,2513	0,2456
69	0,2288	-0,1607	0,1559	0,1558	-0,0938	0,2021	0,0832
70	0,1246	-0,0886	0,0968	0,119	0,1308	0,1546	0,1258

ÖZGEÇMİŞ



KİŞİSEL BİLGİLER

Adı Soyadı: Adil KONDILOĞLU

Uyruğu: T.C.

Doğum Yeri ve Tarihi: Giresun - 27.09.1983

Medeni Hali: Bekar

Adres: Çıtlakkale Mh. 26 Evler 124 Nolu Sk. No:1 Daire:7 Giresun

E-Posta Adresi: adil.kondiloglu@gmail.com

İletişim (Telefon) : 0 544 870 75 70

EĞİTİM

Lise : Anadolu Teknik Lisesi/Bilgisayar Yazılımı (Giresun) - 2001

Lisans : Marmara Üniversitesi/Bilgisayar ve Kontrol Teknolojisi
Eğitimi (İstanbul) - 2009

Yüksek Lisans : Türk Hava Kurumu Üniversitesi/Elektrik ve Bilgisayar
Mühendisliği (Ankara) – Devam Ediyor

TEZ ve PROJELER

Labirent Robotu, Lisans Bitirme Tezi, 2009

C# ile görüntü işleme, Atmel89C52 programlama, RF haberleşme

PWM ile Motor Hız Kontrolü, Sayısal Kontrol Projesi, 2009

Atmel89C52 programlama, LCD ve tuş takımı kontrolü

İŞ DENEYİMİ

Çoruh Edaş, Giresun

Yaz Stajı 02.07.2007 - 03.08.2007

V.B.6.0 ile Yazılım Geliştirme, Teknik Destek

Ülker, İstanbul

Yaz Stajı 07.07.2008 - 08.08.2008

V.B.6.0 ile Yazılım Geliştirme, Teknik Destek

Ada Elektrik ve Otomasyon Sistemleri, Giresun

Yetkili 21.05.2012 - 01.10.2012

Yazılım geliştirme, web sitesi tasarımı, teknik destek