



**T.C.
İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**



DOKTORA TEZİ

**MAKİNE ÖĞRENMESİ İLE UZAKTAN EĞİTİM
ÖĞRENCİLERİNİN PERFORMANSLARININ
DEĞERLENDİRİLMESİ**

Osman YILDIZ

Enformatik Anabilim Dalı

Enformatik

Danışman

Prof. Dr. Sevinç GÜLSEÇEN

II. Danışman

Doç. Dr. Abdullah BAL

Şubat,2014

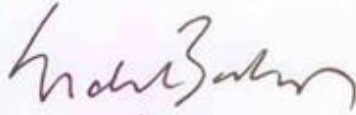
İSTANBUL

Bu çalışma 25/02/2014 tarihinde aşağıdaki jüri tarafından Enformatik Anabilim Dalı Enformatik programında Doktora Tezi olarak kabul edilmiştir.

Tez Jürisi:



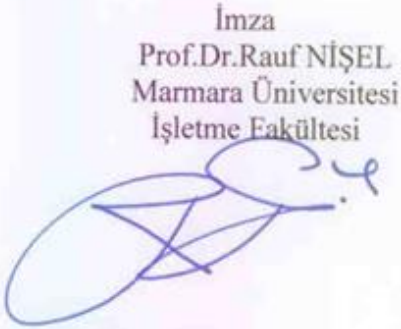
İmza
Prof.Dr.Sevinç GÜLSEÇEN (Danışman)
İstanbul Üniversitesi
Enformatik Bölümü



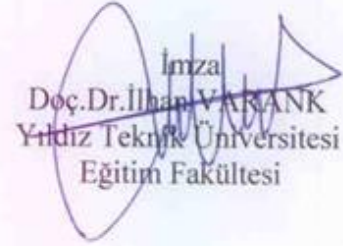
İmza
Prof.Dr.Erdal BALABAN
İstanbul Üniversitesi
İşletme Fakültesi



İmza
Prof.Dr.Hülya ÇALIŞKAN
İstanbul Üniversitesi
Eğitim Fakültesi



İmza
Prof.Dr.Rauf NİŞEL
Marmara Üniversitesi
İşletme Fakültesi



İmza
Doç.Dr.İlhan VARANK
Yıldız Teknik Üniversitesi
Eğitim Fakültesi

ÖNSÖZ

Doktora eğitimimin başlangıcından tez çalışmasının sonuçlandırılmasına kadar olan her aşamada desteğini ve yardımlarını benden esirgemeyen, tavsiyeleri ve yorumları ile daima yanımda olan, bilgisini ve zamanını benimle paylaşan danışmanım Sayın Prof. Dr. Sevinç GÜLSEÇEN'e, yine doktora tez çalışmalarım süresince 2.danışmanım olarak bana hep örnek olan, ışık tutan araştırma konusunun seçiminde, çalışmalarında, değerlendirilmesinde ve yazılmasında yardımını daima gördüğüm değerli hocam Doç.Dr.Abdullah BAL'a çok teşekkür ederim.

Tez çalışmasının gelişmesindeki değerli katkıları ile çalışmaya yön veren tez izleme komitesi üyeleri Prof.Dr. Erdal BALABAN'a ve Doç.Dr. İlhan VARANK'a çok teşekkür ederim.

Üniversite yaşamımda tüm zorluklara katlanan, her koşulda sınırsız destek sağlayan ve her zaman yardımlarını sunan vefakâr eşim Nilgün YILDIZ'a ve yoğun çalışmalarım sırasında sabırlı ve anlayışlı davranan sevgili oğlum Ahmet Eren YILDIZ'a, Tüm yorgunluğumu unutturan ve tekrar çalışabilmem için gerekli enerjiyi veren kızım Elif Berra'ya çok teşekkür ederim.

Şubat, 2014

Osman YILDIZ

İÇİNDEKİLER

Sayfa No

ÖNSÖZ.....	i
İÇİNDEKİLER	ii
ŞEKİL LİSTESİ.....	v
TABLO LİSTESİ	viii
SİMGE VE KISALTMA LİSTESİ	xi
ÖZET.....	xii
SUMMARY	xiii
1. GİRİŞ.....	1
1.1 ARAŞTIRMANIN AMACI VE ÖNEMİ	4
1.2 TEZİN BÖLÜMLERİ	6
2. GENEL KISIMLAR	7
2.1 LİTERATÜR ARAŞTIRMASI.....	7
2.2 LİTERATÜR DEĞERLENDİRMESİ	19
2.2.1 Kullanılan Veri Tipi	19
2.2.2 Verilerin Zaman Aralığı.....	20
2.2.3 Kullanılan Metotlar	20
2.2.4 Kategorik ya da Sayısal Veri Tahmini.....	21
2.3 ÖNERİLEN MODEL.....	21
3. MALZEME VE YÖNTEM	24
3.1 BULANIK MANTIK.....	25
3.1.1 Bulanık Modelleme Tipleri	28
3.1.1.1 Mamdani Tipi Bulanık Çıkarım Sistemi.....	29
3.1.1.2 Takagi -Sugeno Tipi Bulanık Çıkarım Sistemi.....	30
3.1.1.5 Mamdani Bulanık Mantık Çıkarım Sistemi ile Takagi-Sugeno Bulanık Mantık Sisteminin Karşılaştırılması.....	35
3.2 GENETİK ALGORİTMA.....	36
3.2.1 Genetik Algoritma Örneği	40
3.3 GENE-BULANIK MANTIK MELEZ YÖNTEMİ	44
3.4 KÜMELEME YÖNTEMLERİ	44

3.4.1	Çok Değişkenli Aykırı Veri Belirleme Yöntemleri	46
3.4.1.1	<i>Mahalanobis Uzaklığı</i>	46
3.4.1.2	<i>Cook Uzaklığı</i>	47
3.4.1.3	<i>Kaldıraç Nokta (Leverage Değeri)</i>	47
3.4.2	K-Ortalamalar Kümeleme Yöntemi.....	48
3.4.3	Bulanık C-Ortalama Yöntemi	48
3.4.4	Çıkarımlı Kümeleme Yöntemi.....	49
4.	BULGULAR	51
4.1	VERİ KÜMESİ	51
4.1.1	Zaman, Frekans, Süre (ZFS) Veri Kümesi.....	51
4.1.2	ZFS, Küçük Sınav ve Ara Sınav Veri Kümesi.....	51
4.2	UYGULAMA 1:ZFS VERİLERİ KULLANILARAK OLUŞTURULAN BULANIK MANTIK YÖNTEMİ İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME... 52	52
4.3	UYGULAMA 2: ZFS, QUIZ VE ARASINAV VERİLERİ KULLANILARAK OLUŞTURULAN GENE-BULANIK MANTIK HİBRİT YÖNTEMİ İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME	55
4.3.1	Klasik Bulanık Mantık Modeli	55
4.3.2	Uzman Görüşüne Dayalı Bulanık Mantık Modeli	59
4.3.3	Gene-Fuzzy Modeli.....	62
4.4	UYGULAMA 3: TAKAGI-SUGENO TİPİ BULANIK MANTIK İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME	68
4.4.1	Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin Kantil Yöntemine Göre Tahmin Edilmesi.....	68
4.4.1.1	<i>Aykırı Veri Analizi</i>	68
4.4.1.2	<i>Verilerin Normalize Edilmesi</i>	72
4.4.1.3	<i>K-Ortalamalar Kümeleme Yöntemi</i>	73
4.4.1.4	<i>Bulanık C-Ortalamalar Kümeleme Yöntemi</i>	84
4.4.1.5	<i>Çıkarımlı (Subtractive) Kümeleme Yöntemi</i>	92
4.4.1.6	<i>Modelin Geçerliliği</i>	100
4.4.1.7	<i>Sonuçların Karşılaştırılması</i>	105
4.4.2	Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin Kantil Yöntemine Göre Tahmin Edilmesi.....	110
4.4.2.1	<i>K-Ortalamalar Kümeleme Yöntemi</i>	110
4.4.2.2	<i>Bulanık C-Ortalamalar Kümeleme Yöntemi</i>	113

4.4.2.3	<i>Çıkarımlı Kümeleme Yöntemi</i>	116
4.4.2.4	<i>Sonuçların Karşılaştırılması</i>	120
5.	TARTIŞMA VE SONUÇ	121
	KAYNAKLAR	124
	EKLER	130
	ÖZGEÇMİŞ	150

ŞEKİL LİSTESİ

Sayfa No

Şekil 2.1: Öğretim Yönetim Sistemi tarafından tutulan öğrencilere ait kayıt örneği.	20
Şekil 3.1: Bulanık mantık sisteminin çalışma yapısı [56].	28
Şekil 3.2: Bulanık VE ve VEYA İşlemleri için Mamdani Tipi Bulanık Çıkarım Sisteminde Minimizasyon ve Maksimizasyon İşlemleri [57].	30
Şekil 3.3: Sugeno Bulanık Çıkarım Sistemi.	33
Şekil 3.4: Genetik algoritma ile biyolojik evrim süreci arasındaki analogi [18].	37
Şekil 3.5 Genetik algoritma adımları.	37
Şekil 3.6: 1-1 Bireylerin çaprazlama gösterimi.	42
Şekil 3.7: 2-4 Bireylerin çaprazlama gösterimi.	43
Şekil 3.8: Bulanık mantık üyelik fonksiyonları.	44
Şekil 4.1: ZFS Bulanık Mantık Modeli.	52
Şekil 4.2: Zaman üyelik fonksiyonu.	53
Şekil 4.3: Frekans üyelik fonksiyonu.	53
Şekil 4.4: Süre üyelik fonksiyonu.	54
Şekil 4.5: Klasik bulanık mantık üyelik fonksiyonu.	56
Şekil 4.6 Normal dağılım z değerleri.	56
Şekil 4.7: Süre üyelik fonksiyonu aralık değerleri.	57
Şekil 4.8: Girdi değişkenlerine ait üyelik fonksiyonları.	58
Şekil 4.9: Çıkış Değişkenine ait üyelik fonksiyonu.	58
Şekil 4.10: Akademik performans ve tahmin değerleri karşılaştırma grafiği.	59
Şekil 4.11: Girdi değişkenlerine ait üyelik fonksiyonları.	61
Şekil 4.12: Akademik performans ve tahmin değerleri karşılaştırma grafiği.	62
Şekil 4.13: Üyelik fonksiyonu aralık değerleri.	63

Şekil 4.14 Genetik algoritma kromozom yapısı.	63
Şekil 4.15: Gene-Bulanık mantık yöntemi ile elde edilen üyelik fonksiyonları.....	65
Şekil 4.16: Akademik performans üyelik fonksiyonu.	66
Şekil 4.17: Eğitim verilerine ait akademik performans ile tahmin değerleri karşılaştırma grafiği.....	66
Şekil 4.18: Test verilerine ait akademik performans ile tahmin değerleri karşılaştırma grafiği..	67
Şekil 4.19: K-Ortalamlar kümeleme yöntemi üyelik fonksiyonları grafikleri.....	76
Şekil 4.20: K-ortalamlar kümeleme yöntemi ile kurulan bulanık model e ait veri tipi bazında akademik performans ve tahmin karşılaştırma grafikleri.	78
Şekil 4.21: K-ortalamlar kümeleme yöntemi Matlab ekran görüntüsü.....	80
Şekil 4.22: K-ortalamlar kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	81
Şekil 4.23: K-ortalamlar kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	81
Şekil 4.24: K-ortalamlar kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	82
Şekil 4.25: K-ortalamlar kümeleme yöntemi küçük sınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	82
Şekil 4.26: K-Ortalamlar kümeleme yöntemi arasınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	83
Şekil 4.27: K-Ortalamlar kümeleme yöntemi akademik performans çıktısı kurallara ait Matlab ekran görüntüsü.	83
Şekil 4.28: Bulanık C-Ortalamlar kümeleme yöntemi üyelik fonksiyonları grafikleri.	85
Şekil 4.29: Bulanık C-Ortalamlar kümeleme yöntemi ile kurulan bulanık model sonuçları.....	87
Şekil 4.30: Bulanık C-ortalamlar kümeleme yöntemi Matlab ekran görüntüsü.	88
Şekil 4.31: Bulanık C-ortalamlar kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	89
Şekil 4.32: Bulanık C-ortalamlar kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	89
Şekil 4.33: Bulanık C-ortalamlar kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	90
Şekil 4.34: Bulanık C-ortalamlar kümeleme yöntemi ödev girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	90

Şekil 4.35: Bulanık C-ortalamlar kümeleme yöntemi arasınavgirdisi üyelik fonksiyonları Matlab ekran görüntüsü.	91
Şekil 4.36: Bulanık C-ortalamlar kümeleme yöntemi akademik performans çıktısı,kurallara ait Matlab ekran görüntüsü.	91
Şekil 4.37: Çıkarımlı kümeleme yöntemi üyelik fonksiyonları grafikleri.	93
Şekil 4.38: Çıkarımlı kümeleme yöntemi ile kurulan bulanık model sonuçları.	95
Şekil 4.39: Çıkarımlı kümeleme yöntemi Matlab ekran görüntüsü.	96
Şekil 4.40: Çıkarımlı kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	97
Şekil 4.41: Çıkarımlı kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	97
Şekil 4.42: Çıkarımlı kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	98
Şekil 4.43: Çıkarımlı kümeleme yöntemi küçük sınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	98
Şekil 4.44: Çıkarımlı kümeleme yöntemi arasınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.	99
Şekil 4.45: Çıkarımlı kümeleme yöntemi akademik performans çıktısı,kurallara ait Matlab ekran görüntüsü.	99
Şekil 4.46: Bulanık C-Ortalamlar kümeleme yöntemi üyelik fonksiyonları grafikleri.	111
Şekil 4.47: K-Ortalamlar kümeleme yöntemi ile kurulan bulanık model sonuçları.	112
Şekil 4.48: Bulanık C-Ortalamlar kümeleme yöntemi üyelik fonksiyonları grafikleri.	114
Şekil 4.49: Bulanık C-Ortalamlar kümeleme yöntemi ile kurulan bulanık model sonuçları...115	115
Şekil 4.50: Çıkarımlı kümeleme yöntemi üyelik fonksiyonları grafikleri.	117
Şekil 4.51: Bulanık C-Ortalamlar kümeleme yöntemi ile kurulan bulanık model sonuçları...119	119

TABLO LİSTESİ

Sayfa No

Tablo 2.1: Literatürde ve çalışmada kullanılan metotların karşılaştırılması.	21
Tablo 3.1: Bireylerin uygunluk değerleri.....	41
Tablo 3.2: Seçim sonrası eşleşecek bireyler.	41
Tablo 3.3: Eşleşecek bireylerin ikilik düzende gösterimi.	42
Tablo 4.1: Klasik bulanık mantık modeline bağlı doğruluk oranları.	54
Tablo 4.2: Uzman görüşlü bulanık mantık modeline bağlı doğruluk oranları.	54
Tablo 4.3: Süre girdisi için ortalama ve standart sapma.	56
Tablo 4.4: Eşleştirme hata oranları.	64
Tablo 4.5: Metotlara ait doğruluk (accuracy) oranları.	67
Tablo 4.6: Gene-Bulanık test verileri doğruluk değeri.	67
Tablo 4.7: Mahalanobis uzaklığı metoduna göre aykırı verinin değerleri.	69
Tablo 4.8: Cook Uzaklığı Metodu ile bulunan eğitim ve test verileri içerisinde ki aykırı verilere ait değerler.	70
Tablo 4.9: Cook uzaklığı metodu ile bulunan doğrulama verileri içerisinde ki aykırı verilere ait değerler.	70
Tablo 4.10: Kaldıraç Nokta Metodu ile bulunan eğitim ve test verileri içerisinde ki aykırı verilere ait değerler.	71
Tablo 4.11: Kaldıraç Nokta Metodu ile bulunan doğrulama verileri içerisinde ki aykırı verilere ait değerler.	71
Tablo 4.12: Normallik testi.	72
Tablo 4.13: Eğitim verisi küme sayılarına göre doğruluk değerleri.....	73
Tablo 4.14: K-ortalamalar kümeleme yöntemi ile bulunan küme merkezleri.	74
Tablo 4.15: K-ortalamalar metodu ile bulunan değişkenlere ait gauss eğri denklemleri.	75
Tablo 4.16: Kurallara ait parametreler.	77
Tablo 4.17: Verilere ait adet bilgisi.....	79

Tablo 4.18: Verilere ait adet doğruluk oranları.....	79
Tablo 4.19: Eğitim verisi küme sayılarına göre doğruluk değerleri.....	84
Tablo 4.20: Bulanık C-Ortalamalar kümeleme yöntemi ile bulunan küme merkezleri.	84
Tablo 4.21: Kurallara ait parametreler.	86
Tablo 4.22 Verilere ait adet bilgisi.....	87
Tablo 4.23 Verilere ait doğruluk oranları.....	87
Tablo 4.24: Çıkarımlı kümeleme yöntemi ile bulunan küme merkezleri.....	92
Tablo 4.25: Kurallara ait parametreler.	94
Tablo 4.26: Verilere ait adet bilgisi.....	95
Tablo 4.27: Verilere ait doğruluk oranları.	95
Tablo 4.28: Normallik Test Sonuçları.....	100
Tablo 4.29: Sıralama değer sonuçları.....	101
Tablo 4.30: KOKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.	101
Tablo 4.31: Sıralama değer sonuçları.....	102
Tablo 4.32: BCOKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.	102
Tablo 4.33: Sıralama değer sonuçları.....	103
Tablo 4.34: ÇKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.	103
Tablo 4.35: KOKY ve BCOKY modellerinin varyans eşitlik test sonuçları.	104
Tablo 4.36: Gerçek Değerlerle KOKY için Spearman Korelasyon Katsayısı.	104
Tablo 4.37: Gerçek Değerlerle BOCKY için Spearman Korelasyon Katsayısı.....	105
Tablo 4.38: Akademik performansın sayısal olarak tahminine dayalı doğruluk oranları.	105
Tablo 4.39: Akademik performansın “geçti”-“kaldı” tahminine dayalı doğruluk oranları.....	106
Tablo 4.40: Kümeleme metotlarına göre sayısal tahmin sonuçları.....	106
Tablo 4.41: Kümeleme Metotlarına Göre Kategorik Tahmin Sonuçları.....	108
Tablo 4.42: Verilere ait adet bilgisi.....	110
Tablo 4.43: Bulanık C-Ortalamalar kümeleme yöntemi ile bulunan küme merkezleri.	110

Tablo 4.44 Kurallara ait parametreler.	111
Tablo 4.45: Verilere ait doğruluk oranları.	112
Tablo 4.46: Bulanık C-Ortalamalar kümeleme yöntemi ile bulunan küme merkezleri.	113
Tablo 4.47: Kurallara ait parametreler.	114
Tablo 4.48: Verilere ait doğruluk oranları.	116
Tablo 4.49: Çıkarımlı kümeleme yöntemi ile bulunan küme merkezleri.....	116
Tablo 4.50: Kurallara ait parametreler.	118
Tablo 4.51: Verilere ait doğruluk oranları.	119
Tablo 4.52: Akademik performansın sayısal olarak tahminine dayalı doğruluk oranları.	120
Tablo 4.53: Akademik performansın “geçti”-“kaldı” tahminine dayalı doğruluk oranları.....	120

SİMGE VE KISALTMA LİSTESİ

Simgeler	Açıklama
$\mu_A(x)$: Bulanık Mantık Üyelik Fonksiyonu
\hat{y}	: Bulanık Model Çıktısı
c_i	: Küme Merkezi
D_i	: Yoğunluk Değeri

Kısaltmalar	Açıklama
ÖYS	: Öğretim Yönetim Sistemi
KOKY	: K-Ortalamalar Kümeleme Yöntemi
BCOKY	: Bulanık C-Ortalamalar Kümeleme Yöntemi
ÇKY	: Çıkarımlı Kümeleme Yöntemi
BM	: Bulanık Mantık
RFM	: Recency, Frequency, Monetary
ZFS	: Zaman, Frekans, Süre
T-S	: Takagi-Sugeno
OMYH	: Ortalama Mutlak Yüzde Hata
AP	: Akademik Performans
KCD	: Kritik Cook Değeri
KD	: Kaldıraç Nokta Değeri

ÖZET

DOKTORA TEZİ

MAKİNE ÖĞRENMESİ İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN PERFORMANSLARININ DEĞERLENDİRİLMESİ

Osman YILDIZ

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Enformatik

Danışman: Prof. Dr. Sevinç GÜLSEÇEN

II.Danışman: Doç. Dr. Abdullah BAL

Uzaktan eğitimdeki öğrencilerin yıl sonu akademik performanslarının önceden tahmin edilmesi ve bu tahmin sonucunda elde edilecek bilginin kullanılmasıyla gerekli önlemlerin alınması çok önemlidir. Özellikle öğrencilerin akademik performanslarının iyileştirilmesine ve böylece eğitim kalitesinin artırılmasına olanak sağlanabilir. Bu çalışmada uzaktan eğitimdeki öğrencilerin ilk 6 haftalık öğretim yönetim sistemi verileri kullanılarak, yıl sonu akademik performanslarını tahmin etmeye yönelik yeni bir matematiksel model geliştirilmiştir. Klasik bulanık ve uzman görüşüne dayanılarak oluşturulan bulanık modeller kurulmuş, sonrasında genetik algoritma kullanılarak bulanık modele ait üyelik fonksiyon aralıkları optimize edilmiştir. Daha sonrasında kümeleme yöntemleri kullanılarak bulanık mantık ile birlikte melez bir model oluşturulmuştur. Bu çalışmada kullanılan veriler veriler açık kaynak kodlu öğretim yönetim sistemi olan Moodle üzerinden elde edilmiştir. 2011-2012 öğretim yılı Temel Bilgisayar Bilimleri dersine kayıtlı 218 öğrencinin verileri kullanılarak model oluşturulmuştur. 2012-2013 öğretim yılı Temel Bilgisayar Bilimleri dersine kayıtlı 95 öğrenciye ait veriler de doğrulama verisi olarak kullanılmıştır. Yapılan çalışmada bir öğrencinin sisteme en son ne zaman girdiği, ne sıklıkla girdiği, en son oturumda ne kadar süre kaldığı gibi 6 haftalık verilere ilave olarak, 4.haftada yapılan küçük sınav notu ve 8.haftada yapılan ara sınav notu bilgileri giriş verisi olarak kullanılmıştır. Bu veriler ışığında oluşturulan modellerle öğrencilerin akademik performans tahminleri yapılmış ve elde edilen sonuçlar karşılaştırılmıştır.

Şubat 2014,152 sayfa

Anahtar kelimeler:Bulanık, Çıkarımlı, Genetik, FCM, Kümeleme

SUMMARY

Ph.D. THESIS

EVALUATING DISTANCE LEARNING STUDENTS' PERFORMANCE BY MACHINE LEARNING

Osman YILDIZ

Istanbul University

Graduate School of Science and Engineering

Department of Informatics

Supervisor: Prof. Dr. Sevinç GÜLSEÇEN

Co-Supervisor: Assoc. Prof. Dr. Abdullah BAL

It is essential to predict distance education students' year-end academic performance early during the semester and to take relevant measures on the basis of such prediction. In this way, it could be possible to enhance students' academic performance and thus the quality of education they are provided with. The present study is forced on designing a new mathematical model to predict distance education students' year-end academic performance in reference to the 6-week data kept in the learning management system. Particular fuzzy models were formed on the basis of classical fuzzy and learned opinion, and the membership function ranges for the fuzzy model were optimized through a genetic algorithm. Next, a hybrid model was formed with fuzzy logic using clustering methods. The data were collected on Moodle, an open-source learning management system. The model was based on the data on a total of 218 students registering for the course Basic Computer Sciences during the 2011-2012 Academic Year. In addition, one more dataset, which was concerning another 95 students registered for the course during the 2012-2013 Academic Year, was used as verification data. The input data were comprised of five components, namely the last time when students logged on to the system, the frequency at which they logged on to the system, the amount of time spent online in the last session, one quiz administered in Week 4, and one midterm exam administered in Week 8. The models formed on the basis of these data were used to predict students' academic performance and to make particular comparisons.

February 2014, 152 pages

Anahtar kelimeler: Fuzzy, Subtractive, Genetic, FCM, Clustering

1.GİRİŞ

Şahin [1] tarafından belirtildiği gibi, yirmi birinci yüzyılda, bilgi patlaması ile birlikte bütün toplumların eğitime olan gereksiniminin arttığı görülmektedir. Geleneksel eğitim kuruluşları artan eğitim talebini karşılayamamaktadırlar. Ortaya çıkan eğitim açığı her geçen gün giderek artmaktadır. Bu gereksinim, toplumları geleneksel eğitime alternatif arama çabasına itmiş, “uzaktan eğitim” kavramı ortaya çıkmıştır [1].

Uzaktan eğitim, ilgileri, yetenekleri, yaşları, öğrenim düzeyleri, çalışma koşulları çok çeşitli olan ve coğrafi koşulları nedeni ile eğitim engeli bulunan kitlelerin eğitim ihtiyaçlarını karşılamak, bireysel çalışmaya, kendi kendine öğrenmeye dayandığı için, öğrenme sorumluluğunu büyük ölçüde bireye bırakan, insanlara değişik eğitim seçeneği sağlamak gibi belirgin imkanlar sağlayan, öğrencilerin kendi kendine karar verme ve girişimcilik yeteneklerinin gelişmesine yardımcı olmayı amaçlar [2].

Bilgisayar destekli eğitim ortamlarında öğrenci davranışları, üzerinde durulması gereken bir konudur. Geleneksel eğitim ortamı olan sınıflarda öğretmenler ders süresi içerisinde öğrenci davranışlarını gözlemleyebilmektedirler. Bu gözlemler neticesinde öğrencinin gereksinimleri belirlenebilmektedir. Eğitim ortamının düzenlenmesinde, ders içeriğinin ve öğretim yönteminin belirlenmesinde bu gözlemlerden yararlanılmaktadır. Ölçme ve değerlendirme aşamasında da, öğrenci davranışlarının gözlemlenmesi kullanılan bir yöntemdir. Geleneksel eğitim ortamlarında incelenen öğrenci davranışları, uzaktan eğitimde takip edilmek istenildiğinde değişik yöntemlere ihtiyaç duyulmaktadır. Teknolojiye dayalı eğitim ortamı öğretmen için, öğrenci davranışlarını inceleme açısından zorluklar içermektedir. Öğrenci ders çalıştığı zamanlarda öğretmeninden uzakta olmaktadır. Uzaktan eğitimde öğrenci davranışları bilgisayar kayıtları olarak veri dosyalarında saklanmaktadır. Bu verilerden öğrenci davranışları ile ilgili yorum yapılmak istenildiğinde analizlere gereksinim duyulmaktadır [3].

Uzaktan Eğitimde geleneksel eğitim ortamı olmadığı için öğrencilerin başarılarının da takip edilmesi zor olmaktadır. Uzaktan eğitimin avantajlarının yanında, dezavantajlarından biri olarak öğrencinin genel olarak bir işte çalışıyor olması ve

kendine ya da ailesine ayırması gereken zamanı ders çalışarak geçirmesi gerçeği gösterilebilir. Derse gerekli zaman ayrılmadığında telafisi zor sonuçlara sebep olabilmektedir. Bu sonuçlardan bazıları öğrencinin derste başarısız olması, yeniden dersi almak durumunda kalması ya da daha kötüsü öğrencinin eğitimini tamamen bırakmak zorunda kalmasıdır. Uzaktan eğitim sisteminin ortaya çıkmasındaki en önemli etkenlerden biri yer ve zamandan bağımsız olarak, eğitimin herkes tarafından alınabilmesidir. Geleneksel eğitimde olduğu gibi, uzaktan eğitimde de ulaşılması amaçlanan hedef öğrencinin başarılı bir şekilde bulunduğu programdan mezun olmasıdır.

Uzaktan eğitimde yapılan tüm yenilikler, güncelleştirmeler, teknolojinin kullanımı, öğrenim kuramlarının ders içeriklerinde kullanılması tek bir amaç için gerçekleşmektedir; bu da öğrencinin hedeflenen eğitim ve öğretim amaçlarına etkin bir şekilde ulaşmasıdır. Bu amacın yerine getirilmesinde öğrencinin akademik performansının dönem başında değerlendirilmesi son derece büyük katkılar sağlayacaktır. Uzaktan eğitimde ki bir öğrencinin nihai akademik performansı dönem boyunca alacağı notların birleşiminden oluşmaktadır. Dönem sonu itibari ile öğrencinin başarılı ya da başarısız olduğu değerlendirilmesi yapılabilmektedir. Bu aşamadan sonra yapılan tüm değerlendirmeler, geçmiş eğitim dönemini kapsamaktadır.

Geleneksel eğitimde eğitmen, öğrenci davranışlarını gözlemleyerek ölçme değerlendirmeye katkı sağlayabilmektedir. Uzaktan eğitimde ise gözlemeleme metodu mümkün olmamaktadır. Bu çalışma, bu probleme çözüm getirmeyi varsaymaktadır. Uzaktan eğitimde öğrencinin performansının takibi öğretim yönetim sistemi içerisindeki loglarla tutulmaktadır. Log kelimesi İngilizce bir kelimedir. Türk Dil Kurumu'nun Bilişim Terimleri Sözlüğü'ne göre anlamı *"Bir bilgisayar dizgesinin kullanıldığı tüm iş geçişlerine ve görevlere ilişkin her türlü bilgiyi, her birimin kullanım süreleri, kullanıcıların kimlikleri vb. tüm ayrıntılarıyla içeren tutanaklar. Bu tutanakların birikmesiyle oluşan kütük; günlük kütüğü"* şeklindedir. Bu loglar da, öğrencinin ne kadar süre öğretim materyallerini çalıştığı, sistemde ne kadar süre aktif olduğu, yapılan küçük sınavlara (quiz) ait başarı ve başarısızlıkları, konu ile ilgili açılan forumlarda ne kadar aktif olduğu, ne kadar mesaj atıp, ne kadar mesaj okuduğu bilgileri tutulabilmektedir. Bu loglar analiz edilerek bir öğrencinin performansı final sınavından önce tahmin edilebilir.

Öğrencinin performansının tahmin edilmesi, eğitimi veren kuruluşa ve eğitmene birçok fayda sağlar. Akademik yılın ortalarında öğrencilerin performanslarının tahmini, risk grubundaki öğrenciler için, ileride karşılaşılabilecekleri kötü sonuçların olmaması adına erken önlemler alınabilmesini sağlar.

Bu çalışma ile hedeflenen, öğrencinin akademik performansını dönem bitiminden haftalar önce tahmin edilebilmek ve öğretim görevlilerine, öğretim kurumlarına, içerik sağlayıcılarına iyileştirmeler ve güncellemeler ile ilgili katkılar sağlayabilmektir.

Son yıllarda eğitim ve öğretim süreçlerinde geleneksel eğitimden, çevrimiçi eğitime doğru anlamlı derecede bir dönüşüm yaşanmaktadır. Bu dönüşümün en büyük nedenleri olarak, uzaktan eğitimin erişim kolaylığı, esneklik, kişisel öğrenme ve güçlü geri besleme gibi özelliklere sahip olması gösterilmektedir [4]. Uzaktan eğitim sistemi, çevrimiçi eğitim içeriklerini görsel ve işitsel olarak sunma imkanı verir. Sürekli geliştirilmeye açıktır. Her yerden ulaşılma imkanı vardır. Forum ve sohbet gibi özellikleri ile öğrenciyi merkeze taşır. Farklı eğitim arayüzü ve modüllerle kişiye özel eğitim avantajı sağlar.

Education Dynamic [5], çevrimiçi eğitim ile ilgili yapmış olduğu bir anket çalışmasında uzaktan eğitim öğrencilerinin eğitimlerini neden yarıda bıraktıklarına cevaplar aranmıştır. Sonuçlar 5 neden üzerinde toplanmıştır. Ekonomik problemler (%41), hayat şartları (%31), sağlık problemleri (%23), kişisel motivasyon eksikliği (%21) ve kurumla etkileşim eksikliği (% 21) [5]. Farklı bir açıdan bakıldığında öğrencilerin kişisel motivasyon eksikliği (%21) ve kurumla etkileşim eksikliği (%21) uzaktan eğitimi veren kurumu ilgilendirir. Kurum bu sebepleri ne kadar erken teşhis edebilirse ona göre erken önlemler alabilir ve böylece hem öğrencinin eğitimini yarıda bırakmasını engelleyebilir, hem de başarının artmasını sağlayabilir.

Education Dynamics [5] tarafından yapılmış olan anket çalışmasında ortaya çıkan nedenlerin dışında uzaktan eğitimin dezavantajları arasında gösterilen bir neden daha vardır. Bu da geleneksel eğitimde var olan gözlemleme becerisinin eksikliğidir. Geleneksel eğitimde öğretmen öğrenci davranışlarını gözlemleyebilir ve gözlemleme sonucuna göre iyileştirici önlemler alabilir. Ne var ki bu uzaktan eğitimde neredeyse imkansızdır.

Uzaktan eğitim genel olarak bir öğretim yönetim sistemi (ÖYS) ortamında verilir. Bu ortamlar da çok büyük miktarlarda veriler bulunmaktadır. Öğrencilerin ortam üzerindeki her hareketi izlenir ve veri tabanlarında log olarak kayıtları tutulur. Bu verilerin analizi ile çok değerli bilgilere ulaşılabilir [6].

1.1 ARAŞTIRMANIN AMACI VE ÖNEMİ

Son yıllarda sanal öğrenme ortamları veya uzaktan eğitim platformları katlanarak büyümektedir. Bu sistemler güçlükleri ortadan kaldırma potansiyeline sahip olmakla birlikte esneklik, sürekli güncellenen materyaller, kişiselleştirilebilen öğrenme ortamları, sürekli güncellenen materyaller ve hızlı geri bildirim gibi özellikler sunmaktadır [6]. Katılımcı sayısındaki artış ile birlikte internet altyapısının da gelişmesiyle uzaktan eğitim, geleneksel eğitime alternatif olarak hızla gelişmektedir.

Uzaktan eğitimde eğitmenin özel bir yeri vardır. Genel itibari ile görevi öğretmek, değerlendirmek ve sürekli olarak öğrencilerini desteklemektir. Öğrenci ile aralarındaki iletişim, postayla, telefonla, mail ile ya da danışmanlık görüşmeleri ile gerçekleşir. Bu iletişim eğitmenlere sahip oldukları karışık rolleri için yardımcı olur. Her durumda eğitmen, öğrencinin eğitimsel problemini çözmeli, arkadaşça yaklaşmalı, eğitimlerini devam ettirebilmeleri için cesaretlendirmeli, zorluklarını anlayabilmeli ve verimli şekilde öğrencilerini desteklemelidir. Ayrıca eğitmenler, öğrencilerin çalışmalarına not verir, yorum yapar ve tavsiyede bulunurlar [7].

Geleneksel eğitim ortamı olan sınıflarda eğitmenler ders süresi içerisinde öğrenci davranışlarını gözlemleyebilmektedirler. Bu gözlemler neticesinde öğrencinin gereksinimleri belirlenebilmektedir. Eğitim ortamının düzenlenmesinde, ders içeriğinin ve öğretim yönteminin belirlenmesinde bu gözlemlerden yararlanılmaktadır. Ölçme ve değerlendirme aşamasında, öğrenci davranışlarının gözlemlenmesi kullanılan bir yöntemdir.

Bilgisayar destekli eğitim ortamlarında öğrenci davranışları, üzerinde durulması gereken bir konudur. Geleneksel eğitim ortamlarında incelenen öğrenci davranışları, uzaktan eğitimde takip edilmek istenildiğinde değişik yöntemlere ihtiyaç duyulmaktadır. Teknolojiye dayalı eğitim ortamı eğitmen için, öğrenci davranışlarını

inceleme açısından zorluklar içermektedir. Öğrenci ders çalıştığı zamanlarda eğitmeninden uzakta olmaktadır. Uzaktan eğitimde öğrenci davranışları bilgisayar kayıtları olarak veri dosyalarında saklanmaktadır. Bu verilerden öğrenci davranışları ile ilgili yorum yapılmak istenildiğinde analizlere gereksinim duyulmaktadır.

Geleneksel eğitimde eğitmen, öğrenci davranışlarını gözlemleyerek ölçme değerlendirmeye katkı sağlayabilmektedir. Uzaktan eğitimde ise gözlemeleme metodu mümkün olmamaktadır. Bu çalışma, bu probleme çözüm getirmeyi varsaymaktadır.

Günümüzde uzaktan eğitim veren kurumlarda genel olarak değerlendirme deneme yanılma yöntemine dayanmaktadır. Önce öğrenci belirlenen içerikle eğitimini almakta, dönem sonu itibari ile yapılan değerlendirmeler sonucunda gerekli görüldüğü takdirde değişiklikler yapılmaktadır. Bu değerlendirmenin yapılabilmesi için başlangıçta öngörülen eğitim süresinin bitmesi gerekmektedir. Çünkü öğrencinin eğitim süresi sonunda almış olduğu başarı notu ile yapılan değerlendirmeler neticesinde bir sonraki dönemde yapılacak eğitim ile ilgili kararlar alınabilmektedir. Bu şekilde yapılan bir değerlendirme sürecinin bazı eksiklikleri vardır. Yeni uygulanacak yöntemler ya da kararlar bu sürede değerlendirilmeye tutulan öğrenciler için hiçbir önem taşımamaktadır; ancak yeni gelecek öğrencileri ilgilendirmektedir, ya da, başarısız olup aynı dersi tekrar almak zorunda kalan öğrencileri ilgilendirebilir.

Bu çalışma ile uzaktan eğitimdeki bir öğrencinin 6 haftalık öğrenci yönetim sisteminde kaydedilen logların analizi ile yıl sonundaki başarı notu tahmin edilmeye çalışılmaktadır. Böylece 6 hafta sonunda elde edilen enformasyon, kurumlar için önemlidir. Bu enformasyonun kullanılmasıyla eğitimi veren kurum, kritik düzeyde olan öğrenciler için farklı pedagojik metotlar uygulayabilir ya da içerikle alakalı genel değişikliğe gidebilir. Bu çalışma dersi veren eğitmen ve kurum için bir çeşit acil uyarı sistemi görevini yürütecek enformasyon çıktılarını sağlayacaktır.

Öğrencinin performansının tahmin edilmesi, eğitimi veren kuruluşa ve eğitmene birçok fayda sağlar. Akademik yılın başlarında öğrencilerin performanslarının tahmini, risk grubundaki öğrenciler için, ileride karşılaşılabilecekleri kötü sonuçlar olmaması adına erken önlemler alınabilmesini sağlar.

1.2 TEZİN BÖLÜMLERİ

Sunulan tez çalışması ile ilgili olarak ikinci bölümde araştırma konusu olan uzaktan eğitim öğrencilerinin akademik performansının ölçülmesi, kurulan modellerin hangi makine öğrenme yöntemleri ile kurulduğu, hangi veri tipleri ile çalışmaların yapıldığına dayalı literatür çalışması yer almaktadır. Üçüncü bölümde model kurmada kullanılan yöntemler olarak bulanık mantık, genetik algoritma ve kümeleme yöntemleri anlatılmaktadır. Bölüm dört, uygulama modelini ve başarımlarını anlatmaktadır. Araştırmanın genel sonuçları ve çıkarımları beşinci bölümde sunulmuştur.

2.GENEL KISIMLAR

2.1 LİTERATÜR ARAŞTIRMASI

İlgili literatüre bakıldığında konu ile ilgili benzer çalışmalar olduğu görülmüştür. Uluslararası yayınlarda öğrenci performans tahminleri makine öğrenmesi teknikleri kullanılarak tahmin edilmeye çalışılmıştır. Uluslararası yayınlarda ortak olan nokta, performans kriteri olarak “geçti” ve “kaldı” şeklinde kategorik bazda tahminlerin yapılmasıdır [6, 7]. Bu çalışmalar sadece öğrencinin geçip geçmediğini tahmin eden çalışmalardır. Çalışmaların sonuç bölümünde, ileride yapılabilecek çalışmalar arasında öğrencilerin notlarının tahmin edilebileceği ayrıca final sınavından önce asgari ne kadar zamandaki verilerin tahmin için yeterli olacağı ile ilgili çalışmalar yapılabileceği belirtilmiştir. Ayrıca yapılan çalışmalarda ortak olan diğer bir nokta da; yapılan çalışmaların genellikle tek bir metot ile yapılmış olmasıdır. Birden fazla metodun birleştirilmesi ile oluşan bir algoritmanın kullanılmadığı görülmüştür.

Pena-Reyes ve Sipper [8] tarafından, 2009 yılında yapmış oldukları, “A Fuzzy –Genetic Approach to Breas Cancer Diagnosis” adlı çalışması göğüs kanseri teşhisini konu almaktadır. Öncelikle hastalık parametrelerine bağlı olarak bulanık mantık kuralları oluşturulmuş ve bu kurallar baz alınarak bulanık mantık sistemi modeli kurulmuştur. Sonrasında bulanık mantık parametreleri genetik algoritma ile bulunabilecek şekilde algoritma geliştirilmiştir. 120 iterasyon sonucunda %97 oranında başarı elde edilmiştir [8].

Aydın ve Biroğul [9], 2008 yılında Bilişim Teknolojileri Dergisi’nde yayınlanan, “E-öğrenmede Açık Kaynak Kodlu Öğretim Yönetim Sistemleri” adlı çalışmalarında açık kaynak kodlu öğretim sistemleri hakkında bilgi vermiş, mevcut yazılımlar ile ilgili karşılaştırmalar yapılmıştır. Karşılaştırma sonucunda Moodle, öğretim yönetim sistemleri arasında barındırdığı özelliklerden dolayı ön plana çıktığı görülmüştür [9].

“A New Method for Constructing Membership Functions and Fuzzy Rules from Training Examples” isimli 1999 yılında yayınlanan çalışmada, Wu ve Chen [10], bir

çok çalışmada bulanık mantıkta sayısal veriden kural oluşturulabileceğinden bahsetmişler, kendi çalışmalarının alfa-cut ile kuralların oluşturulduğunu ve bu yönüyle çalışmanın diğerlerinden farklı olduğu belirtilmiştir. Ayrıca yapılan uygulama bilgisayarda Matlab ortamında geliştirici bir araç haline getirilmiş ve iris sınıflandırma probleminde uygulanmıştır. Sonuçlar daha önce yapılan Hong ve Lee'nin çalışması ile karşılaştırılmıştır. Wu ve Chen bu çalışmada, doğruluk oranını yaklaşık %1'lik artırarak %96,21 e çıkardığı görülmektedir [10].

Taylan ve Karagözoğlu [11] tarafından yapılan çalışma, "An Adaptive Neuro-Fuzzy Model For Prediction Of Students' Academic Performance" ismi ile 2009 yılında yayınlanmıştır. Adı geçen çalışma öğrencilerin akademik performanslarının tahmini ile ilgili bir model sunmaktadır. Bu model, yapay sinir ağlarına dayanan, bulanık mantık sisteminin dizaynına, sistematik bir yaklaşım sergilemektedir. Bu yöntemin bulanık mantık sistemlerinin eksik kalan taraflarından olarak gösterilen öğrenebilme ve adaptif yeteneklerine katkı sağlayacağı belirtilmiştir. Yapılan çalışmada öğrencilerin akademik performansı ağırlıklı toplam değerlendirme aracı ile hesaplanmıştır. Sonrasında bulanık mantık ve yapay sinir ağları hibrid edilerek model kurulmuş ve sonuçlar değerlendirilmiştir [11].

"Real-Valued Genetic Algorithm for Fuzzy grey Prediction System" isimli, 1997'de yayınlanan çalışmada, Huang ve Huang [12], genetik algorithmda bir kodlama çeşidi olan gri kodlamaya göre gerçek değerleri kullanarak bir model geliştirmişlerdir. Genetik algorithmda yaygın kullanım olan ikilik düzene kodlama yapıp, sonrasında seçme, kopyalama ve mutasyon işlemleri yapılmaktadır. Kurulan modelde ikilik düzene çevrilmeden gerçek sayılarla, genetik algoritmanın adımları işlemde geçirilmiştir. Sonrasında model Missouri bölgesine ait aylık sıcaklık değerlerini tahmin etmek üzere denenmiş ve sonuçlar değerlendirilmiştir [12].

Yusof ve diğ. [13], 2009 yılında yapmış oldukları "Evaluation of Students' Performance and Learning Efficiency Based of ANFIS" adlı çalışmada, Programming Technique kursu öğrencilerinin performanslarının tahmine odaklanmıştır. Öğrencilerin ortalama skorları, ne kadar zaman harcadıkları, yardıma ihtiyaç duyup duymadıkları gibi parametreler, giriş parametreleri kabul edilmiştir. Öncesinde sadece bulanık mantık kuralları uygulanarak yapılan çalışma, yapay sinir ağları da dahil edilerek ANFIS

modeli kurulmuştur. Uzman görüşüne dayanılarak oluşturulan bulanık mantığın eksiklikleri yapay sinir ağları ile giderilmeye çalışılmıştır [13].

Cho ve diğ. [14] tarafından 1997 yılında yayınlanan, “Fuzzy-PID Hybrid Control Automatic Rule Generation Using Genetic Algorithm” isimli makalede, genetik algoritma kullanılarak bulanık mantık kurallarının otomatik olarak oluşturulabileceği gösterilmiştir. PID kontrolleri için oluşturulan kural tablosu simülasyon sonuçları göstermiştir ki, genetik algoritma en uyguna yakın kuralların çıkarılması ya da aranması ile ilgili etkili bir metottur [14].

“A new Method to Construct Membership Functions and Generate Fuzzy Rules from Training Instances” isimli, Chen ve Tsai [15] tarafından yapılan makale 2005 yılında yayınlanmıştır. Adı geçen makalede son yıllarda iris sınıflandırma ile ilgili bir çok çalışma yapıldığı belirtilip, ilgili çalışmada iris sınıflandırma problemi üyelik fonksiyonlarının oluşturulması ve kuralların genetik algoritma ile çıkarılmasına dayanan yeni bir model yapıldığı belirtilmiştir. Bulanık mantık sınıflandırmada en önemli görevin üyelik fonksiyonlarının oluşturulması ve uygun bir kural sınıfının bulunması olduğu ifade edilen makalenin sonucunda %97,33 doğruluk payı bulunmuştur. Bu sonuç iris sınıflandırma örneğini daha önce farklı yöntemlerle yapan çalışmalara göre daha yüksektir [15].

Kaptanoğlu ve Özok [16], “Akademik Performans Değerlendirmesi için Bir Bulanık Model” isimli çalışmalarında, akademik yükseltme ve atanma ölçütleri adı altında sıralanan ölçütler ve bunların puanlarından oluşan sistemlerdeki eksikliklerden yola çıkmışlardır. Bu kapsamda bir çok ölçütü bulanık karar verme modeli ile modellemişler ve bulanık analitik hiyerarşi prosesi esaslı bir çalışma yapmışlardır. Chang’ın bulanık analitik hiyerarşi prosesi modelinin temel alındığı çalışmada üç ayrı bulanık sıralama yöntemi kullanılmış ve sonuçlar tartışılmıştır [16].

“Determination of Fuzzy Logic Membership Functions Using Genetic Algorithm” adlı çalışma Arslan ve Kaya [17] tarafından hazırlanmış olup 2001 yılında yayınlanmıştır. İlgili yayında bulanık mantık için dezavantajlardan biri olarak kabul edilen üyelik fonksiyonları, genetik algoritma kullanılarak oluşturulmuştur. Çalışmada bulanık mantık üyelik fonksiyonu ikili üçgen şeklinde ve üçlü üçgen şeklinde kullanılmış ve sonuçlar değerlendirilmiştir [17].

Herrera [18], 2008 yılında yayınlanan “Genetic Fuzzy Systems: Taxonomy, Current Research Trends Prospects” isimli çalışmasında bulanık mantığın geniş bir alanda farklı çalışma dallarında kullanıldığından bahsetmiş, genetik algoritmanın bulanık mantığa adapte edilmesi ile birlikte, öğrenme ve adaptif yönlerinin de eklenebileceği belirtilmiştir. Çalışmada kısaca klasik bulanık mantık anlatılmış ve genetik bulanık mantık sistemleri hakkında bilgi verilmiştir [18].

Chiu [19], “Fuzzy Model Identification Based on Cluster Estimation” adlı makalesinde çıkarımlı kümeleme analizinin algoritmalarını belirtmiş, bu algoritmalar sonucu küme merkezlerinin nasıl bulunacağını göstermiştir. Bulunan küme merkezleri ile Takagi-Sugeno tipi bulanık modellemede kuralların oluşturulması ve bu oluşturulan kurallara ait parametrelerin bulunması üzerinde durulmuştur [19].

Bulanık kümelerde dezavantaj olarak gösterilen kuralların uzman kişilerce belirlenmesi konusu araştırmacılar tarafından geliştirilmesi gereken önemli bir konu olarak görülmüştür. Priyono ve diğ. [20], 2005 yılında Jurnal Teknologi isimli dergide yayınladıkları “Generation of Fuzzy Rules With Subtractive Clustering” adlı makalede Chiu’nun geliştirdiği çıkarımlı kümeleme metodu ile model kurulmuş ve uygunluk limit değeri genetic algoritma kullanılarak hesaplanmıştır ve sonuçlar yorumlanmıştır.

Küçük mobile robotların davranışları için geliştirilen bulanık kuralların sunulduğu, Kim ve Kong [21] tarafından 1997 de sunulan bildiriye, yine Chiu tarafından geliştirilen yöntemle bulanık kurallar oluşturulup, mobil robotlarda kullanılması anlatılmaktadır.

“Introduction to Five Data Clustering Algorithms” adlı yayında Moertini [22], K-Means Kümeleme, Fuzzy C-Means Kümeleme, Mountain Kümeleme, Subtractive Kümeleme ve Partition Simplification Fuzzy C-Means Kümeleme ile ilgili giriş seviyesinde bilgiler vermektedir. Sonuç kısmında bu algoritmalar özetlenmektedir. Bulanık Kümeleme algoritmalarında sıkça kullanılan Bulanık C-Ortalamalar kümeleme yöntemi ile Çıkarımlı Kümeleme yöntemi melez edilmesini konu alan “A modified Hybrid Fuzzy Clustering Algorithm for Data Partitions” adlı makalede deneysel sonuçlara yer verilmiştir. Melez model kullanılmadan yapılan kümeleme ile melez model kullanılarak kümeleme arasındaki farklar gösterilmiştir.

“Benzetilmiş Tavlama ve Çıkarımlı Kümeleme Tabanlı Yeni Bir Bulanık Sınıflandırıcı” adlı makalede Torun ve Tohumoğlu [23], sınıflandırıcı amaçlı bir bulanık sistemin üyelik fonksiyonlarının oluşturulmasında ve kural tabanı kurulmasında çıkarımlı kümeleme yöntemini kullanmıştır. Bulanık sınıflandırıcının küme kural sayısının ve mimarisine doğrudan etkisi olan çıkarımlı kümeleme yönteminin küme komşuluk yarıçapı, yeni bir benzetilmiş tavlama eniyileme yöntemiyle elde edilmiştir. Bu çalışmada ortaya konulan melez sınıflandırıcının diğer sınıflandırıcılara göre daha iyi bir performansa sahip olduğu gözlenmiştir.

Lykourantzou ve diğ. [24] 2009 yılında yaptığı çalışmada farklı 3 metottan gelen sonuçların değerlendirilmesi ile bir öğrencinin dersi ya da okulu bırakması “multiple genetic algorithm” yöntemi ile tahmin ediliyor. Çalışmada test sonuçları, proje değerlendirmeleri ve demografik veriler kullanılmıştır.

Vandamme ve diğ. [25] 2007 yılında yaptıkları çalışmada öğrencilerin demografik verileri ile sosyo-ekonomik geçmişi, akademik geçmişi gibi verilerle öğrencileri low, medium, high gibi risk gruplarına sınıflandırarak, kimin sınıfta kalacağı ya da okulu bırakacağı yapay sinir ağları metodu ile tahmin ediliyor.

Kaya [26], (2003) yılında “Credit Risk Modeling Using Machine Learning Techniques” adlı yüksek lisans tezinde yapay sinir ağları, destek vektör makineleri ve k-en yakın komşu sınıflandırma algoritması gibi çeşitli makine öğrenmesi teknikleri kullanarak kredi skorlaması sonuçları üzerinde durmuştur. Lojistik regresyonu da karşılaştırma amacıyla kullanmıştır. Sonuçlara göre DVM kredi skorlaması için iyi bir alternatif ve DVM-Reject bu çalışmadaki en başarılı model olmuştur.

Tuğ [27], “Genetik Algoritmalar İle Tıbbi Veri Madenciliği” adlı tezinde öncelikle kısaca Genetik Algoritma yöntemi açıklanmıştır. Daha sonra veri madenciliğinden bahsedilmiştir. Çalışmada Doğal Bağışıklık Sisteminden esinlenerek oluşturulmuş bir metod olan Yapay Bağışıklık Sistemi önerilmiştir. Burada SAND olarak bilinen bir yaklaşım sunulmuştur. Bu yaklaşım hatayı minimize eden bir arama ile veri kümesinde en iyi başlangıç popülasyonunu üretmeyi hedeflemektedir. Ayrıca daha hassas sonuçlar üretmek için bir yöntem olarak Bulanık Mantıktan yararlanılmıştır. Bu çalışmada genetik algoritma kullanılarak bir veri madenciliği uygulaması geliştirilmiştir. Bu

uygulamada tıbbi veriler kullanılmıştır. Kanser hastalarına ait veriler içinde hastalığa sebep olan etken grupları araştırılmaktadır. Amaç sık karşılaşılan hastalık sebeplerini bulmaktır.

“Tıbbi Veriler Üzerinde Makine Öğrenme Algoritmaları ve Bulanık Mantık ile Kurallar Öğrenme” isimli yüksek lisans tezinde Uzun [28]; ameliyat sonrası hastalardan alınan tıbbi veriler üzerinde sınıflandırıcı makine öğrenme algoritmaları (OneR, Naive Bayes, Jrip, Ridor, SMO, J48, LMT, Conjunctive Rule, Decision Tables, NNge, Kstar, IBk, PART) ve bulanık mantık ile sınıflandırma analizleri yapmıştır. Tıbbi veri kümesi 9 özellik ve 90 veriden oluşmaktadır. Bulanık Mantık sınıflandırmayı kıyaslamak için OneR algoritması kullanılmış. Bulanık Mantık ile yapılan sınıflandırma değerinin OneR algoritması ile elde edilen sınıflandırma değerinden daha iyi olduğu tespit edilmiştir.

Deniz [29], 2006 yılında “Bulanık Mantık Tabanlı Tahmin Modeli ve Uygulaması” adlı yüksek lisans tezinde bulanık mantığın ileriye yönelik tahmin (öngörü) konusundaki belirsizlikler için önerdiği yaklaşımları analiz etmiştir. Tez tahmin kavramına genel bir bakışla başlamakta olup tahmin kavramı hakkında önce genel bir bilgi verilmektedir. Ardından bulanıklık kavramının genel bir çerçevesi sunulmuş. Daha sonra bulanık mantığı temel alan tahmin metotları açıklanmıştır. Bu tezde zaman serilerinde ortaya çıkabilecek mevsimselliğin tahmini için bir bulanık tahmin metodu geliştirilmiş ve bu metodun uygulanması ele alınmıştır. Ortaya konulan metot Türkiye’ye gelen turist sayısının öngörüsüne uygulanmıştır.

Çiftçi [3], 2006 yılında “Uzaktan Eğitimde Öğrencilerin Ders Çalışma Etkinliklerinin Log Verileri Analiz Edilerek İncelenmesi” isimli tezinin temel amacı uzaktan eğitimde öğrencilerin bilgisayar destekli öğretim ile ders çalışma alışkanlıklarını değerlendirmektir. Araştırma, Uzaktan Sürekli Eğitim (e-USE) projesi log verileri dikkate alınarak geliştirilen anket formuna dayalı olarak gerçekleştirilmiştir. Araştırmanın temel bulgularının elde edilmesinde anket formu verileri ile log verilerinin karşılaştırmalı analizleri esas alınmıştır. Araştırmada bulguların yorumlanmasında anket ile log verilerinin analizleri dikkate alınmıştır

Aydın [30], “Veri Madenciliği Ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama” adlı 2007 yılında yapılan tezinde, Anadolu Üniversitesi Uzaktan Eğitim Sisteminde eğitim gören öğrencilere ilişkin farklı kaynaklardaki veriler bir araya

getirilerek veri madenciliği uygulaması gerçekleştirmiştir. Uzaktan Eğitim Sisteminin planlama faaliyetlerine katkı sağlayabilecek öğrenci performansını tahmin etmeye yönelik model geliştirilmiş ve mezun olan öğrencilerin profillerini belirlemeye yönelik kümeleme çalışması yapılmıştır. Öğrenci başarısını tahmin etmeye yönelik çalışmada C5.0 karar ağacı algoritmasının kullanıldığı bir tahmin modeli önerilmiştir. Önerilen modelin karar kuralları sisteme entegre edilerek öğrenci başarı tahmini amacıyla kullanılabilmesi öngörülmüştür. Mezun olan öğrencilere yönelik çalışmada “K-means” algoritması kullanılarak beş küme elde edilmiştir. Kümeleme analizi ile elde edilen bilgilerin bilgisayar kullanımı ve öğrenci başarısı arasındaki ilişkiyi doğrular nitelikte olduğu görülmüştür

2008 yılında yapılan “Bulanık Mantıklı Sürücü Modeli İle Hız Tahmini Ve En Kısa Yol Belirleme” adlı tezinde Ekmekçi [31], sürücü davranışlarına etki eden; eğitim, ortalama viraj yarıçapı, kaplama cinsi gibi yolun fiziksel ve geometrik özellikleri ile yoldaki trafik sıkışıklığı ve kamyon, otomobil, otobüs gibi araç tipleri ele alınmış ve Mamdani metotları kullanılarak sürücü hızını tahmin eden bulanık model oluşturulmuştur. Elde edilen sonuçlar, Karayolları Genel Müdürlüğü’nün hassas ölçümlerle elde ettiği araç hızları verileriyle kıyaslanmış, kamyon için % 2.22, otobüs için % 2.22 ve otomobil için % 2.21’lik ortalama hata oranları tespit edilmiştir.

“Hibrit Bulanık Sinir Ağını Kullanarak Bir Sınıflandırma Ve Kural Çıkarma Sisteminin Geliştirilmesi” adlı doktora tezinde Kahramanlı [32], 2008 yılında, tıp alanında sınıflandırma yapılabilmesi için melez bulanık bir sinir ağı önermiş ve bir algoritma geliştirmiştir. Çalışmada bu güne kadar yapılandırılmış farklı olarak, ayrık nitelikler ikili olarak kodlanırken, gerçel nitelikler bulanıklaştırılmıştır. YSA ve BSA’nın birleşiminden oluşan bir melez ağ oluşturularak kodlanmış veriler melez ağa sunulmuştur. Eğitim algoritması olarak, Geri Yayılım kullanılmıştır. Bu ağın iki farklı veri tabanında uygulaması gerçekleştirilerek sistemin performansı 10 kez çapraz doğrulama ve “eğitim”+”test” yöntemleri kullanılarak test edilmiştir.

Bu çalışmada ikinci olarak, melez sistemden kural çıkarmak için bir yöntem sunulmuştur. Bunun için sinir ağı eğitimi bitirdikten sonra oluşan fonksiyon çıkarılmış ve optimize edilmiştir. Optimizasyon için bir yapay bağımsızlık algoritması olan Opt-

aiNET kullanılmıştır. Yöntem iki veri tabanında test edilmiş ve başarılı olduğu gözlenmiştir.

“Yapay Sinir Ağları Ve Bulanık Mantık Yöntemleri Kullanılarak Tava Buharlaştırma Tahmini Yapılması” adlı tezinde Afşar [33], 2008 yılında, Kayseri, Kırşehir, Nevşehir ve Yozgat illerindeki dört istasyondan alınan günlük ortalama sıcaklık, minimum sıcaklık, maksimum sıcaklık, ortalama nem ve buharlaştırma verileri kullanmıştır. MATLAB programında hazırlanan kodlar yardımıyla Çok Katmanlı Yapay Sinir Ağları (ÇKYS), Radyal Tabanlı Yapay Sinir Ağları (RTYS), Genelleştirilmiş Regresyon Yapay Sinir Ağları (GRYS), Adaptif Ağ Yapısına Dayalı Bulanık Çıkarım Kuralı (ANFIS) ve Çoklu Doğrusal Regresyona (ÇDR) değişik giriş kombinasyonları uygulayarak buharlaştırma tahminleri yapmış ve her bir değişkenin buharlaşmaya olan etkisini incelemiştir.

“Yeni makine Öğrenmesi Metotları ve İlaç Tasarımına Uygulamaları” adlı doktora tezinde Amasyalı [34] 2008 yılında yapılan tezinde; tezin uygulama alanı olarak seçilen ilaç tasarımı için çeşitli makine öğrenmesi algoritmaları geliştirmiştir. İlaçların insan sağlığına olan olumlu etkilerinden bahsederek, gelecek 20 yılda TÜBİTAK raporuna dayandırılarak ilaç tasarımının teknolojik faaliyet konuları içerisinde yer alacağı vurgusunda bulunulmuştur. İlaç tasarım sürecinde sınıflandırma, kümeleme, özellik seçimi/çıkarımı, regresyon (eğri uydurma) problemleri yer almaktadır. Bu tezde amaç bu tarz problemlere çözüm üretmeyi amaçlayan makine öğrenmesi metotları yardımıyla ilaç tasarımının süresi ve maliyeti azaltılabilmektedir. Bu nedenle de tezde makine öğrenmesinin birçok alanını kapsayacak bir çalışma gerçekleştirilmiştir. Sınıflandırma problemleri için Cline adı altında bir algoritma ailesi tasarlanmıştır. Geliştirilen algoritmalar temelde karar ağacı oluşturma algoritmalarıdır. Karar ağaçları, yüksek performansları ve ürettikleri kuralların verinin yapısına ait çıkarımlar yapmayı kolaylaştırması sebebiyle oldukça popüler olmuş makine öğrenmesi algoritmalarındandır. Yapılan denemelerde geliştirilen algoritmaların basitliklerine rağmen UCI ve ilaç veri kümelerinde mevcut algoritmalarla yarışabilecek performansta algoritmalar oldukları görülmüştür. Sınıflandırıcı komiteleri literatürdeki birçok çalışmada tekil sınıflandırıcılardan daha başarılı sonuçlar üretmiştir. Bu çalışmada da buna paralel sonuçlar alınmış ve Cline algoritma ailesine Cline karar ormanları eklenmiştir. UCI ve ilaç veri kümeleri üzerinde Cline karar ormanları mevcut

algoritmalarından çok daha iyi sonuçlar sergilemiştir. Cline karar ağacı ve karar ormanları algoritmaları ClineToolbox adlı bir yazılımla kullanıcıların hizmetine sunulmuştur.

Özellik seçimi problemleri için karar ağaçları ve karar ormanlarından yararlanan bir yaklaşım geliştirilmiş ancak tatmin edici sonuçlar elde edilememiştir. Kümeleme problemleri için Clusline adı altında bir algoritma ailesi geliştirilmiş ve mevcut algoritmalarla çeşitli kümeleme performans kriterlerine göre yarışan sonuçlar elde edilmiştir.

Coşkun [35], 2010 yılında “Veri Madenciliği Algoritmaları Karşılaştırılması” adlı yüksek lisans tezinde veri madenciliğinin genel bir tanımı ile başlayarak; veri madenciliği metotları ve algoritmaları hakkında bilgi vermiş, model oluşturma basamakları ve oluşturulan modellerin karşılaştırılması için kullanılan metriklerden bahsedilmiş ve bu teorik bilgiler ışığında bir veri kaynağı üzerinde veri madenciliğinde yaygın olarak kullanılan birkaç algoritmanın karşılaştırmalı değerlendirmesi yapılmıştır. Çalışmanın teorik kısmında veri madenciliği uygulamasının bir veri kaynağı üzerinde hangi aşamalardan geçtiği üzerinde durulmuş, kullanılmakta olan metotlar ve bu metotlara ait algoritmalar anlatılmış ve oluşturulan modeli değerlendirme kriterleri hakkında bilgi verilmiştir.

Çalışmanın uygulama kısmında ise, teorik kısımda anlatılmış olan bilgiler ışığında; J48, NaiveBayes, Lojistik Regresyon ve KStar algoritmalarının karşılaştırması yapılmıştır. Karşılaştırma yapılırken veri ön işlemeden başlamak üzere, hangi bilgilerin kullanıldığı, nasıl bir metot izlendiği, algoritmaların oluşturduğu modellerin istatistiksel sonuçları ve bu sonuçların nasıl değerlendirildiği detaylı bir şekilde anlatılmıştır.

Yapılan karşılaştırma sonucuna göre eldeki veri üzerinde çalıştırılan dört farklı algoritmadan J48 algoritmasının ürettiği modelin en iyi karşılaştırma ölçütlerine sahip olduğu sonucuna ulaşılmış, ancak modeller arasında belirgin bir farklılık oluşmadığı sonucu da vurgulanmıştır.

Akbilgiç [36], 2011 yılında “Hibrit Radyal Tabanlı Fonksiyon Ağları İle Değişken Seçimi Ve Tahminleme: Menkul Kıymet Yatırım Kararlarına İlişkin Bir Uygulama” adlı doktora tezinde yapay sinir ağlarının özel bir biçimi olarak değerlendirilen Radyal Tabanlı Fonksiyon Ağlarının eksiklikleri olan gizli katman nöron sayısının deneme

yanılma yolu ile belirlenmesi, merkez ve yayılım parametrelerinin uzun zaman alabilen iteratif yöntemlerle belirlenmesi ve tasarım matrisinin tekilliği gibi ve bağımsız değişkenlerinin hangilerinin bağımlı değişken ile ilişkili olduğunun belirlenememesi gibi sorunlara çözüm bulma amaçlanmıştır. Çalışmada radyal tabanlı fonksiyon ağları sözü edilen sorun ve eksikliklere çözüm getirecek şekilde uygun istatistik yöntemlerle entegre edilerek, Hibrit Radyal Tabanlı Fonksiyon Ağları modeli oluşturulmuştur. Oluşturulan Hibrit Radyal Tabanlı Fonksiyon Ağının performansı ve geçerliliği, İMKB Ulusal 100 endeksinin yönünün belirlenmesi üzerine bir çalışma ile test edilmiştir.

2011 yılında yayınlanan “Genetik Algoritma Tabanlı Optimal Adaptif Fuzzy PID Kontrolcü Tasarımı” isimli tezinde Çelebi [37], birbirinden farklı metodlar geliştirilmesine rağmen doğrusal olmayan kontrol sistemlerinde kullanılan klasik PID kontrolör performanslarının, doğrusal kontroldeki başarıları ile örtüşmediğini gözlemlemiştir. Doğrusal olmayan sistemler için doğrusal olmayan çıkışlar üretebilen ve uygulaması çok basit olan bulanık mantık kontrolörler (BMK) klasik PID performansına göre daha başarılı sonuç vermekte olduğunu ifade etmiştir. Bulanık kontrolörün avantajlarına rağmen; kontrolör tasarım parametrelerinin (giriş/çıkış fonksiyonları, fonksiyon sayısı, kural tabloları) sadece uzman bilgisi tarafından ya da deneme yanılma yöntemi ile belirlenme gibi bir dezavantajı olduğu belirtilmiştir. Bu sebepten dolayı; bulanık kontrol parametrelerinin bulunmasında genetik algoritma kullanıldığı belirtilmiştir.

Işık [38], 2006 yılında basılan “Genetik Algoritma Tabanlı Bulanık Kontrolün Uçuş Kontrol Sistem Tasarımına Uygulanması” isimli doktora tezinde problem olarak, uçağın uçuş zarfı içinde istenen bir uçuş durumunda uzun süre tutulması üzerine odaklanmıştır. İç ve dış ortam uyarılarından kaynaklanan bozucu etkilerden dolayı, uçağın dinamik durumu kısa bir süre içinde değişebilmektedir. Bundan dolayı herhangi bir uçuş durumu için tasarlanmış bir kontrol sistemi (klasik kontrol sistemleri), bu denge durumundan sapma halinde istenen kararlılık ve performans özelliklerini vermeyebilmektedir. Işık literatür incelemesinde uçuş kontrol sistem tasarımında farklı kontrol uygulamalarının kullanıldığını görmüştür. Bu kontrol uygulamalarından olan bulanık kontrol, sistem transfer fonksiyonunun tanımlanmadığı fakat uygulanacak denetimin sözel olarak ifade edildiği sistemlerde kullanılmaktadır. Ayrıca, transfer fonksiyonu bilinen sistemlerde klasik denetleyicilerin yerine

sistem performansını iyileştirme amacıyla da uygulanabilmektedir. Bulanık sistemlerin performanslarının artırılması için genetik algoritmaların kullanılması da oldukça yaygın bir yöntemdir. Bulanık uçuş kontrol sistem tasarımı konusunda yapılan çalışmaların büyük çoğunluğu ise benzetim seviyesindedir. Bu çalışmada uçuş kontrol sisteminde önemli bir yer tutan irtifa kontrolünün genetik bulanık PD denetleyici kullanılarak tasarımı incelenmiş ve geniş gövdeli, dört motorlu jet yolcu uçağına ait elde edilen uygulama sonuçları, klasik PD denetleyici kullanılarak elde edilen tepkilerle birlikte değerlendirilmiştir.

Altay [39], “Genetik Algoritma ve Bir Uygulama ” isimli, 2007 yılında basılan tezinde İstanbul Teknik Üniversitesi son sınıf öğrencilerinin teorik derslerde aldıkları bilgilerin uygulama alanında tecrübeye dönüştürebilmeleri için uygun şirketlerle eşleştirilmeleri üzerine çalışmıştır. Sanayideki şirketlerin her biri kendi çalışacakları konuları belirlemiş, buna göre öğrencilerin konuları ile eşleştirme sağlanması bu çalışmadaki problem cümlesi olmuştur. Bu eşleştirme için Altay genetik algoritma kullanmıştır. Genetik algoritma ile çeşitli amaç fonksiyonlarının karşılaştırılması yapılacak ve hangi amaç fonksiyonu daha verimli çalıştığı belirlenmesi amaçlanmıştır. Bu amaç doğrultusunda öncelikle hazırlanan adımlar işlenerek, Microsoft Excel Visual Basic Editor’de yazılan bir makro ile iki amaç fonksiyonu kullanılarak GA ile amaç fonksiyonlarının etkinlikleri belirlenmiş ve sonuçlar değerlendirilmiştir.

2007 yılında Gündoğdu [40], tarafından yüksek lisans tezi olarak sunulan, “Veri Madenciliğinde Genetik Algoritmalar” isimli çalışmada 2003-2004 girişli öğrencilerin verileri ile genetik algoritma kullanılarak ortaya çıkan kurallar ve parametreler aracılığıyla yapılan bazı değişikliklerin sonuçlara olan etkisi incelenmiştir. Veri madenciliğinde sınıflandırma algoritması olarak kullanılan genetik algoritmanın performansı ve hızı gözlemlenmeye çalışılmış ayrıca algoritmanın veri madenciliğinde kullanım şekli iyileştirilmeye çalışılmıştır.

2006 yılında yayınlanan “Bulanık Kümeleme Algoritmalarının Karşılaştırılmalı Analizi ve Bilgisayar Uygulamaları” isimli tezinde Avcı [41], temel kümeleme yöntemlerini ele alarak küme sayısı bilinmediğinde parametre ayarlaması yapılmaksızın optimum kümeleme sonucunun nasıl elde edileceğini incelemiştir. Literatürde bulunan kümeleme metotlarından DBSCAN yöntemi detayları ile analiz edilmiş ve yeni

önerilen bir yöntem olan ve bulanık kümeleme yöntemleri ile yoğunluk tabanlı kümeleme düşüncesinin melez edildiği FJP yöntemi de incelenmiştir. DBSCAN ve FJP yöntemlerinin karşılaştırılabilirlikleri için bilgisayar programları oluşturulmuş ve sonuçlar yorumlanmıştır.

Erilli [42], 2009 yılında basılan “Kümeleme Analizinde Bulanık Yaklaşım Algoritmaları ve Uygulamaları” isimli yüksek lisans tezinde kümeleme ve bulanık kümeleme yöntemleri tanıtılmış, aralarındaki farklar belirtilmiştir. Uygulama kısmında Devlet Planlama Teşkilatının, illerin gelişmişlik göstergelerinde kullanılan sosyo-ekonomik veriler yardımıyla, Türkiye’deki 81 il kümelerine ayrılmıştır. Elde edilen sonuçların, gerçek ekonomik yapılarla tutarlı olduğu görülmüştür.

Güler [43], “Bulanık Kümeleme Analizi ve Bulanık Modellemeye Uygulamaları” isimli, 2006 yılında basılan tezinde, kümeleme analizi ile ilgili teorik bilgiler verildikten sonra hiyerarşik olmayan kümeleme yaklaşımına dayanan ve veri setindeki veri grupları arasında kesin ayrımın söz konusu olmadığı durumlarda başarıyla uygulanan bulanık kümeleme algoritmalarından Bulanık c-ortalamlar ve bulanık c-regresyon kümeleme algoritmalarının bulanık modellemeye uygulamalarını ele almıştır. Bu bağlamda, Mamdani, Takagi-Sugeno ve Sugeno-Yasukawa gibi çok iyi tanınan bulanık modellerin yapılması ve kurulması için gerekli aşamalar anlatılmıştır. Sonrasında Takagi-Sugeno modelinin yeni bir versiyonunun üzerinde durulmuş ve modelin tanımlanması için algoritma adımları verilmiştir. Ayrıca tezde bulanık çoklu regresyon modelinin iyileştirilmesi konusu ele alınmış ve bu amaçla yeni modelleme algoritması sunulmuştur.

2011 yılında Gün [44] tarafından yüksek lisans tezi olarak sunulan, “Bulanık Kümeleme Analiziyle Görüntüdeki Yüzün Tanınması” isimli çalışmada yüz tanıma ve tanımlamanın kullanım alanlarına ve önemine vurgu yapılarak, görüntülerden yüz resimlerinin elde edilebilmesi için bazı ön işlemlerin sonrasında öznitelik vektörleri Ayrık Dalgacık Dönüşümü metodu ile edilmiş ve Yapay Sinir Ağları, K-Ortalamlar ve Bulanık C-Ortalamlar algoritmalarından faydalanılarak sınıflandırılmıştır. Kişilerin tanınması işlemi sınıflandırma yapılarak gerçekleştirilmiştir. Tasarlanmış bu sistem sayesinde kimlik tespiti, kişinin tanınması gibi işlemler kolaylıkla gerçekleştirilebileceği savunulmuştur.

Ulutağay [45], 2004 yılında tamamlanan “Bulanık c-Ortalamlar Kümeleme Analizi ve Uygulamaları” isimli tezinde bulanık kümeleme ile ilgili deneyimler sağlayacak visual-interaktif işlem sistemi oluşturmuştur. Bazı sık kullanılan bulanık kümeleme, küme sayısı belirleme ve başlangıç kümelerinin seçilebilmesi, sonuçların görsel olarak ekrana verilebilmesi, veri setinin görsel olarak oluşturulabilmesi, optimal küme sayısının otomatik olarak ayarlanabilmesi gibi işlemlerin yapılabilme imkanı sağlanmıştır.

2.2 LİTERATÜR DEĞERLENDİRMESİ

Yapılan çalışmaların değerlendirilmesi bir kaç noktadan yapılabilir. Uzaktan eğitimdeki öğrencilerin akademik performanslarının değerlendirilmesi üzerine yapılan çalışmalarda kullanılan verilerin tipi, verilerin geçerli olduğu zaman aralığı, kullanılan metotlar ve tahmin sonuçlarının kategorik ya da sayısal olma durumuna göre 4 kategoride değerlendirme yapmak mümkündür.

2.2.1 Kullanılan Veri Tipi

Yapılan çalışmalarda kullanılan veri tipleri incelendiğinde genel olarak demografik veri kullanımı yaygındır. Öğrencilerin cinsiyet, nerede yaşadığı, iş tecrübesi, eğitim seviyesi gibi veriler kullanılmıştır. Bunun yanında bir çok çalışmada çoktan seçmeli sorulardan oluşan testler, projeler, ödevler gibi veriler de kullanılmıştır. Çalışmamızda daha önce akademik performans tahmininde hiç kullanılmayan 3 farklı veri tipi kullanılmıştır. Bu verileri şu şekilde açıklamak mümkündür;

- Zaman (Recency); Öğretim elemanının dersi öğretim yönetim sistemine yükledikten sonra öğrencinin kaç gün sonra sistem içerisinde o derse girdiğini gösterir veri tipidir.
- Frekans (Frequency); Öğrencinin bir sonraki ders yükleninceye kadarki sisteme giriş çıkışlarının sayıldığı veri tipidir.
- Süre (Monetary); Öğrencinin o derste toplam kaç dakika süre harcadığını gösterir veri tipidir.

Şekil 2.1’de uzaktan eğitimin verildiği platform olan öğretim yönetim sistemi tarafından tutulan kayıtlara ait görüntü bulunmaktadır.

Cum 14 Haziran 2013, 13:20	78.188.214.82	Yücel Kader	course view	Temel Bilgisayar Bilimleri
Cum 14 Haziran 2013, 07:12	78.188.214.82	Yücel Kader	page view	Matlab Grafikler
Cum 14 Haziran 2013, 07:12	78.188.214.82	Yücel Kader	quiz view	Quiz 6
Cum 14 Haziran 2013, 07:12	78.188.214.82	Yücel Kader	course view	Temel Bilgisayar Bilimleri
Prş 13 Haziran 2013, 23:54	88.240.35.217	sinan kara	page view	Matlab Döngüler
Prş 13 Haziran 2013, 23:53	88.240.35.217	sinan kara	course view	Temel Bilgisayar Bilimleri
Prş 13 Haziran 2013, 23:18	88.241.195.122	Güner Yiğit Küner	page view	Uygulama Tekrar
Prş 13 Haziran 2013, 22:10	88.241.195.122	Güner Yiğit Küner	page view	Matlab Grafikler
Prş 13 Haziran 2013, 21:49	88.241.195.122	Güner Yiğit Küner	page view	Uygulama Tekrar
Prş 13 Haziran 2013, 21:34	88.240.35.217	sinan kara	page view	Matlab Grafikler
Prş 13 Haziran 2013, 21:33	88.240.35.217	sinan kara	course view	Temel Bilgisayar Bilimleri
Prş 13 Haziran 2013, 20:45	88.241.195.122	Güner Yiğit Küner	page view	Matlab Döngüler
Prş 13 Haziran 2013, 20:07	88.241.195.122	Güner Yiğit Küner	page view	Matlab Diziler Vektörler, Polinomlar, Matrisler

Şekil 2.1: Öğretim Yönetim Sistemi tarafından tutulan öğrencilere ait kayıt örneği.

2.2.2 Verilerin Zaman Aralığı

Yapılan çalışmalarda kullanılan veriler dönem içerisine yayılan aktivitelerden elde edilmiştir. Bu aktiviteler sınav olabileceği gibi, herhangi bir ödevin değerlendirilmesi sonucu verilen not olabilmektedir. Genel olarak incelendiğinde dönem içerisindeki aktivitelerden elde edilen verilerle birlikte yıl sonuna doğru yapılan final sınavı da tahminlerde kullanılmıştır. Sadece bir çalışmada 10.hafta sonunda elde edilen verilerle tahmin yapılmıştır. Bu çalışmayı diğerlerinden ayıran en önemli özelliklerden bir tanesi zaman aralığının olabildiğince kısa tutulmaya çalışılmasıdır. Bu çalışmada öğretim yönetim sistemi içerisinde tutulan logların 6 haftalık kısmı kullanılmıştır. Bu zaman dilimi içerisinde yapılan küçük sınav sonucu ve 8.haftada yapılan ara sınav sonucu ile birlikte 5 farklı değişken kullanılmıştır.

2.2.3 Kullanılan Metotlar

Yapılan çalışmalarda makine öğrenmesi tekniklerinden bir çoğunun kullanıldığı görülmüştür. Yapılan çalışmaların ortak özelliklerinden biri de genel olarak tek bir metot ile modelin kurulması ve sonuçların diğer metotlar ile karşılaştırılması şeklinde olmuştur. Makine öğrenmesi metotlarında bir metodun diğerine göre üstün ya da zayıf yanları olabilmektedir. Bu tür durumlarda zayıflıklar farklı metotlarla birlikte kullanılarak avantaja dönüştürülebilir. Literatürde yer alan çalışmalardan farklı olarak

bu çalışmada bulanık mantık kullanılmış ve bulanık mantık yönteminin eksiklikleri, melez yöntemler kullanılarak giderilmeye ve daha yüksek doğruluk oranları elde edilmeye çalışılmıştır.

Tablo 2.1: Literatürde ve çalışmada kullanılan metotların karşılaştırılması.

Kullanılan Metotlar	
Literatürde	Çalışmada
<ul style="list-style-type: none"> • OneR, Naive Bayes, Jrip, Ridor, SMO, J48, LMT, Conjunctive Rule, NaiveBayes, Lojistik Regresyon ve KStar • Birçok makine öğrenme yöntemi • Genel olarak tek metod 	<ul style="list-style-type: none"> • Bulanık Mantık • Genetik Algoritma • Gene-Bulanık Mantık • Bulanık Mantık-Kümeleme Melez Yöntemi

2.2.4 Kategorik ya da Sayısal Veri Tahmini

Yapılan çalışmalarda uzaktan eğitimdeki öğrencilerin akademik performansları, genel olarak “GEÇTİ”,”KALDI” ya da risk seviyesine göre kategorik olarak tahmin edilmiştir. Öğrencilerin yıl sonu akademik performansları sayısal olarak tahmin edilmemiştir. Bu çalışmada öğrencilerin akademik performansları hem sayısal olarak hem de kategorik olarak tahmin edilmiştir.

2.3 ÖNERİLEN MODEL

RFM (Recency, Frequency, Monetary) analizi, pazarlamada müşteri davranışlarını ölçmede, en çok bilinen ve kullanılan bir bölütleme yöntemlerinden biridir. Recency, Frequency ve Monetary kelimelerinin baş harflerinden oluşur. Recency, müşterinin en son alışveriş güncelliğine, frequency işlem sıklığına, monetary, toplamda müşteri tarafından harcanan miktarı belirtir. RFM analizi ideal müşterinin; en kısa zamanda, çok sık ve çok para harcayan müşteriler olduğu varsayımına dayanır [46]. RFM analizi bir müşterinin davranışını belirleyebilmek için en iyi bilinen metottur [47]. Bu çalışmada uzaktan eğitimin verildiği, öğretim yönetim sistemi bir pazarlama alanı ve öğrenciler de bu alan içerisindeki müşteriler gibi düşünülmüştür. Bu şekli ile bir uzaktan eğitim öğrencisi için şu sorular soruldu;

- Sisteme en son ne zaman girdi?
- Ne sıklıkta girdi?
- Sistemde ne kadar zaman kaldı?

Bu sorularla nasıl ki pazarlamada bir müşterinin karakteristik özellikleri belirlenebiliyorsa, aynı şekilde bir öğrencinin de davranışları sayısal bir kimliğe dönüştürülmeye çalışılmıştır. Bu değerler biraz daha detaylandırıldığında;

- Zaman (Recency) (Z) Değeri: Konunun öğretim elemanı tarafından sisteme yüklenmesinden itibaren, öğrencinin kaç gün sonra o konuya bakmak için girdiğini gün olarak gösteren değerdir.
- Frekans (Frequency) (F) Değeri: Bir sonraki konu sisteme yükleninceye kadar geçen süre içerisinde konuya bakma sıklığını gösteren değerdir.
- Süre (Monetary) (S) Değeri: Her konu için toplamda sistemde kalma süresini dakika olarak gösteren değerdir.

Zaman, Frekans, Süre (ZFS) değerleri 6 haftalık veriler ile sınırlıdır. Model için bu girdi değerlerine ilaveten 2 farklı girdi değişkeni de sisteme eklenmiştir. Bunlar; 4.haftada öğretim yönetim sistemi üzerinden çevrimiçi olarak yapılan küçük sınav sonucu ve 8.haftada sınıf ortamında yapılan 1.arasınav sonucudur.

Bu araştırma Yıldız Teknik Üniversitesinde öğrenim gören Temel Bilgisayar Bilimleri dersini alan öğrencilerin oluşturduğu 312 kişilik bir çalışma grubu ve 16 haftalık bir uygulama ile sınırlıdır. Araştırmanın uygulama süreci Temel Bilgisayar Bilimleri dersinin tüm konularını kapsamaktadır. 312 kişilik grubun 218'i, 2011-2012 öğretim yılına aittir. 95 kişilik kalan kısım ise 2012-2013 öğretim yılına ait verilerdir.

Bu ders için kullanılan öğretim yönetim sistemi açık kaynak kodlu Moodle olmuştur. Moodle internet tabanlı ders sitesi oluşturmak için kullanılan bir yazılım paketidir. Moodle günümüzde bir çok ülkede kullanılmakta olup, seksenin üzerinde ayrı dil desteği vermekte ve yüz binin üzerinde kayıtlı kullanıcısı bulunmaktadır. Moodle php tabanlı açık kaynak kodlu çevrimiçi eğitim sistemidir. Dersler modüller halinde kurulmaktadır. Linux, Unix, Windows ve Mac OS X işletim sistemlerine destek vermektedir [48]. Moodle ile öğrencilerin içeriklere en son ne zaman baktıkları , ne

sıklıkla girdikleri ve içerikte ne kadar zaman harcadıklarına dair veriler loglar da tutulabilmektedir.

Yapılan çalışmalar genel olarak özetlendiğinde; kullanılan yöntemler çeşitli makine öğrenme yöntemleridir. Genellikle farklı yöntemlerle modeller kurulmuş ve sonuçlar karşılaştırılmıştır. Bu çalışmada ana yöntem olarak bulanık mantık seçilmiştir. Bulanık mantık seçilme sebeplerinin başında kesin ve belirgin bilgilerin olmamasıdır. Özellikle kullanılan 3 değişken (Zaman, Frekans, Süre) bilgisayar başındaki öğrencilerin sistemde tutulan logları kullanılarak yapılmıştır. Öğrencinin sisteme giriş yaptıktan sonraki davranışları belirsizlikler içermektedir. Akademik performansın ölçülmesi durumu belirsizlikler taşır. Bulanık mantık bu belirsizlikler altında etkili ve verimli bir metottur [49]. Fakat her yöntemin kendine göre eksiklikleri olabileceği gibi, bulanık mantığın da kendine göre eksiklikleri vardır. Bu eksiklikler genetik algoritma ve kümeleme metotları kullanılarak giderilmeye çalışılmıştır. Kurulan melez model ile eğitim öğretim döneminin yarısında uzaktan eğitimdeki bir öğrencinin yıl sonu akademik performansı yüksek doğruluk oranları ile tahmin edilmiştir. Bu çalışmanın eğitim bilimine katkısının önemli olduğu düşünülmektedir.

3. MALZEME VE YÖNTEM

Makine öğrenmesi metotları geçmişteki verileri kullanarak veriye en uygun modeli bulmaya çalışırlar. Yeni gelen verileri de bu modele göre analiz ederler. Büyük miktarda verinin incelenip onun içinden işe yarayan bilginin (modelin) elde edilmesi işlemine veri madenciliği (data mining) de denilmektedir. Farklı uygulamaların analizlerden farklı beklentileri olmaktadır. Makine öğrenmesi metotlarını bu beklentilere göre sınıflandırmak mümkündür [50].

1. Sınıflandırma: Geçmiş bilgilerin hangi sınıflara ait olduğu verildiğinde yeni gelen verinin hangi sınıfa dahil olduğunun bulunması işlemidir. Örnek olarak iki tahlil sonucuna göre bir kişinin hasta olup olmadığı belirlenmeye çalışılırsa önceki hasta ve sağlam kişilerin tahlil sonuçları kullanılır.
2. Kümeleme: Geçmiş bilgilerin sınıflarının / etiketlerinin verilmediği / bilinmediği durumlarda verilerden birbirine benzerlerin yer aldığı kümelerin bulunması işlemidir.
3. Eğri Uydurma (Regresyon): Geçmiş verilerin sınıflarının sürekli sayılar olduğu durumlarda kullanılır. Örneğin bir hisse senedinin değeri sürekli bir sayıdır ve bu senede ait model bu değeri tahmin etmeye yönelik bir eğri uydurma işlemi olacaktır.
4. Özellik Seçimi / Çıkarımı: Veriye ait birçok özellikten, verinin kümesini / sınıfını / değerini belirleyen özelliklerinin hangileri olduğu bilinmeyebilir. Bu durumda tüm özellik kümesinin bir alt kümesi seçilir (özellik seçimi) ya da bu özelliklerin birleşimlerinden yeni özellikler elde edilir (özellik çıkarımı).
5. İlişki Belirleme: Bir süpermarkette X ürününü alan müşterilerden %80'i Y ürününü de alıyorsa, X ürününü alıp Y ürününü almayan müşteriler, Y ürününün potansiyel müşterileridir. Müşterilerin sepet bilgilerinin (bir alışverişte alınan ürün bilgileri) bulunduğu bir veritabanında potansiyel Y müşterilerini bulma işlemi türündeki problemler ilişki belirleme metotlarıyla çözülmektedir.

Sonuç olarak makineler insanlığın işgücüne sağladıkları katkıyı, makine öğrenmesi metotları sayesinde insanlığın beyin gücüne de sağlamaya başlamışlardır. Her tür uygulama için çok miktarda verinin analiz edilerek gelecekle ilgili varsayımlar geliştirmemize, kararlar vermemize yardımcı olan makine öğrenmesi metotları önemlerini ve katkılarını her geçen gün arttırmaktadır [34]. Bu çalışmada makine öğrenme metotlarından bulanık mantık, genetik algoritma, kümeleme metotları kullanılarak modeller oluşturulmuştur.

3.1 BULANIK MANTIK

İnsanoğlu günlük hayatını sürdürürken pek çok sorun ile karşılaşmakta ve geçmişte edindiği bilgi ve deneyimlerden yararlanarak çözümler üretmektedir. Bu sorunların bir kısmı tamamen belirli olmakta ve kolayca tanımlanabildiğinden dolayı çözüm getirilmesi de kolay olmaktadır. Bunun yanı sıra, belirsizlikler içeren veya tam olarak tanımlanamayan sorunların çözümü ise nispeten zor olmakta ve görecelilik arz etmektedir.

Belirsizliğin bu önemli rolünün anlaşılması, geleneksel anlayıştan belirsizliği temel alan modern anlayışa geçişi sağlayan dönüşümü başlatmıştır. Bu anlayış literatüre ilk girdiği anda deterministik bir anlayıştaki klasik görüşe ters düşmesinden dolayı çok sıcak karşılanmamıştır. Bu alandaki gelişme sürecinde, ihtimal teorisinden farklı olarak belirsizlikle ilgili birçok teorisinin ortaya çıktığı görülmüştür [51].

İlk defa Azeri matematikçi Lotfi A. Zadeh (1965) tarafından ortaya atılan Bulanık Mantık (BM), daha sonraları hızlı bir gelişme göstermiştir. Tam ve kesin bilginin bulunmadığı, belirsizliğin hakim olduğu durumlarda karar verme esasına dayanan BM, karar verme gibi çok geniş bir alanda kullanıldığı için bilim ve teknolojinin hemen her alanında uygulanmaktadır. BM’de amaç, belirsizlik ifade eden, tanımlanması güç veya anlamsal olarak zor kavramlara üyelik derecesi belirleyerek değerleri daha belirgin hale getirmektir [52].

Bulanık Mantığın tanımı iki anlamda yapılabilir; Dar kapsamda düşünüldüğünde BM, 1865 yılında Cantor (Alman matematikçi) tarafından ortaya koyulan klasik ikili mantık sistemine dayanan kümeler teorisinin genelleştirilmiştir. Daha geniş kapsamda

bakıldığında ise BM kesin mantığı da içine alan bulanık kümeler teorisi ile eş anlamdadır [36]. Günümüzde bulanık mantık denildiğinde, geniş anlamda kullanılan tanıma ait bulanık mantık anlaşılmaktadır.

Bulanık mantığın en geçerli olduğu iki durumdan biri, incelenen olayın çok karmaşık olması ve bu olayla ilgili yeterli bilginin bulunmaması durumunda kişilerin görüş ve değer yargılarına yer verilmesi, ikincisi ise insan kavrayış ve yargısına gerek duyulan hallerdir. İnsan düşüncesinde sayısal olmasa bile belirsizlik, yararlı bir bilgi kaynağıdır. İşte bu tür bilgi kaynaklarının, olayların incelenmesinde kullanılmasında bulanık mantık ilkelerinden faydalanılır [53].

Klasik mantıkta bir önerme doğru ise 1, yanlış ise 0 değeri olmak üzere, yalnızca iki değer alabilir. BM’da ise bir önerme, doğruluk derecesine göre kapalı aralığında sonsuz farklı değer alabilmektedir. Bu da gerçek hayata uygulandığında çok keskin kararlar yerine daha esnek kararlar verilmesini sağlamaktadır. Örnek olarak “boyu 1.85 m üzerinde olanlar uzundur” dediğimizde, klasik mantığa göre boyu 1.85 m olan uzundur ama 1.84 olan uzun değildir. Bulanık mantık yaklaşımında ise boyu 1.84 olan bir kişi 0.9 üyelik derecesi ile uzun 0.1 üyelik derecesi ile orta boylu olabilmektedir. Hayatımızda her şey klasik küme de olduğu gibi 1 ve 0’lardan oluşmamakta aksine bir çok olay belirsiz özelliklere sahip olabilmektedir. Günümüzde belirsizliğin hakim olduğu bir problemin modellenerek çözülmesinde BM oldukça sık kullanılmaktadır.

Bulanık kümeler, kullandığımız sözel ifadeleri bilgisayara aktarabilmek için oluşturduğumuz matematiksel modellerdir. Örneğin; uzun boylu insanlar kümesi, pahalı otomobiller kümesi, yakın sürüş mesafesi, güvenilir kar araçları, birden çok büyük sayıların oluşturduğu küme gibi kavramlar klasik küme teorisinin öngördüğü şekilde incelenemezler. Bir bulanık küme, çalışma yapılan alana ait her bir elemana matematiksel olarak kümedeki üyelik derecesini temsil eden bir değer atayarak tanımlanır. Bu değer, elemanın bulanık küme tarafından ifade edilen kavrama üyelik derecesini ifade eder. Bundan dolayı bireylerin kümeye ait olması farklılaşır. Üyelik dereceleri 0 ile 1 arasındaki gerçel sayılarla temsil edilirler. Tam üye olma ve üye olmama durumu, bulanık kümede sırasıyla 1 ve 0 değerleriyle karşılanır. Bundan dolayı da, klasik küme kavramı bulanık küme kavramının bu iki değere kısıtlanmış özel bir

şekli olarak görülebilir [51].

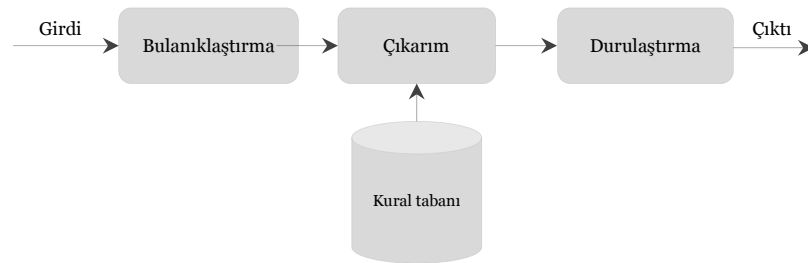
Bulanık Mantığın sağladığı avantajlar:

- İnsan düşünce sistemine ve tarzına yakındır.
- Uygulanmasında mutlaka matematiksel bir modele gereksinim duymaz.
- Yazılımın basit olması nedeniyle, sistem daha ekonomik olarak kurulabilir.
- Bulanık Mantık kavramını anlamak kolaydır.
- Üyelik değerlerinin kullanımı sayesinde, diğer kontrol tekniklerine göre daha esneklerdir.
- Kesinlik arz etmeyen bilgiler kullanılabilir.
- Doğrusal olmayan fonksiyonların modellenmesine izin verir.
- Sadece uzman kişilerin tecrübeleri ile kolaylıkla bulanık mantığa dayalı bir model ya da sistem tasarlanabilir.
- Geleneksel kontrol teknikleriyle uyum halindedir [54]

Bulanık mantıkla ilgili yöntem ve tekniklerin sıklıkla kullanıldığı alanlar arasında görüntü işleme, sinyal işleme, denetleyici sistemler, uzman sistemler, veri tabanları ve veri madenciliği sıralanabilir [53].

Bulanık mantık 3 adımdan oluşur. Fuzzyfication, bulanıklaştırıcı adımdır. Bu adımda girdi olarak sisteme giren gerçek değerler bulanıklaştırılır. Bir elemanın herhangi bir bulanık kümeyle olan üyeliğini belirlemek için üyelik fonksiyonları kullanılır. Bulanık kümeler $\mu_A(x)$ üyelik fonksiyonu ile temsil edilir. $\mu_A(x)$, üyelik fonksiyonundaki bir x noktasının A bulanık kümesindeki üyelik derecesidir. $\mu_A(x)=1$ olduğunda x elemanı A kümesinin kesin bir elemanı demektir. $\mu_A(x)=0$ olduğunda x elemanı kesinlikle A bulanık kümesinin bir elemanı değildir demektir. $0 < \mu_A(x) < 1$ arasındaki her değer, x 'in A bulanık kümesinin bu aralıkta bir üyelik derecesi ile bir elemanı olduğu anlamı taşımaktadır [55]. Burada ki her bir girdi değerine bir üyelik değeri atanıp, dilsel yapıya dönüştürülür. Bir büyüklüğün niceliğini ya da niteliğini anlatmak için günlük dilde farklı kavramlar kullanılır. Bu kavramlar birden fazla özelliği aynı anda içerebileceği gibi açık ve net tek bir anlamı da yoktur. Örneğin “iyi” sıfatı bir arabayı nitelerken bir çok anlamı içinde barındırabilir. Aynı sıfat bir kişi için kullanıldığında kişiye ait ahlaksal özellikler ile beraber sosyal yönünü de yansıtabilir. Aynı terim bir yiyecek için

kullanıldığında ise lezzetine vurgu yapıldığı anlaşılabilir. Günlük hayatta bir çok ifadeyi keskin sınırlarla ayırmak mümkün değildir. Sıcak, ılık, soğuk gibi kavramlar ya da uzun kısa gibi kavramlar net olmayan ifadelerdir ve bu ifadeleri açıklayabilmek için bulanık kümelerden faydalanılır. Bulanık küme kuramının mühendislik veya diğer alanlarda uygulanmasındaki amaç, kesin olmayan bilgiler ışığında tutarlı sonuçlar çıkarabilmektir. İnsan belleğinde yaşam boyunca tecrübe edilen enformasyonların, yorumlanma, anlamlandırılma ve değerlendirilmeleri sonucu oluşan bir sözel kurallar dizisi vardır. Karar vermede çoğu zaman bu kurallar dizisi çalışır ve karar verilir. Makinelerinde bir olay karşısında karar verebilmeleri için o durumla ilgili bilgi ve tecrübelerden oluşan bir kural tabanını sisteme girmek gerekmektedir. İşte bu kısım, bulanık mantığın 2.kismini yani kuralların işlendiği “Bulanık karar verme” adımını oluşturur. Bu kısımda kurallar “eğer öyleyse” şeklinde bulunur. Kurallar dilsel terimlerden oluşur. Örneğin bir kural “Eğer bulut yoksa ve sıcaklık orta ise pikniğe gidilebilir” şeklindedir. Girdi olarak sisteme giren giriş verileri üyelik derecelerine dönüştürülür ve kural tabanındaki önermelerle karşılaştırılır ve yine dilsel terimlere dönüştürülür. Kural tabanı, karar verme işleminde kullanılan bir çok paralel kurallardan ve sistem değişkenlerinden oluşmaktadır. Bu kurallar sistemin girişi ile çıkışı arasındaki ilişkiyi oluşturmaktadır. Makine çıktı olarak başka bir sisteme bilgi yolluyorsa, bulanık çıkışlar sayısal değerlere dönüştürülmesi gerekir. Bu kısma durulaştırma ya da berraklaştırma denir. bulunur. Durulaştırmada bulanık halde bulunan değerler gerçel sayılara dönüştürülür. Şekil 3.1’de genel bir bulanık mantık sistemine ait yapı görülmektedir.



Şekil 3.1: Bulanık mantık sisteminin çalışma yapısı [56].

3.1.1 Bulanık Modelleme Tipleri

Bulanık Mantık sonuçlarının elde edilmesinde yaygın olarak 2 tip çıkarım sistemi kullanılmaktadır. Bunlar Mamdani ve Takagi-Sugeno tipi bulanık çıkarım sistemleridir.

3.1.1.1 Mamdani Tipi Bulanık Çıkarım Sistemi

Mamdani tipi bulanık çıkarım sistemi, yaygın olarak kullanım alanı olan, uzman bilgisi gerektiren ve her türlü problemin çözümüne uygulanabilen bir bulanık mantık yöntemidir. Mamdani sisteminde hem girdi değişkenleri hem de çıktı değişkeni kapalı formdaki üyelik fonksiyonları ile ifade edilir.

Mamdani tipi bulanık çıkarım sistemi çok kolay oluşturulur, insan davranışlarına çok uygundur. Bu nedenle yaygın bir kullanıma sahiptir ve diğer bulanık mantık modellerin temelini oluşturur. İlk defa bir buhar motorunun insan tecrübelerinden elde edilen sözel kontrol kuralları yardımıyla kontrolü amacıyla kullanılmıştır.

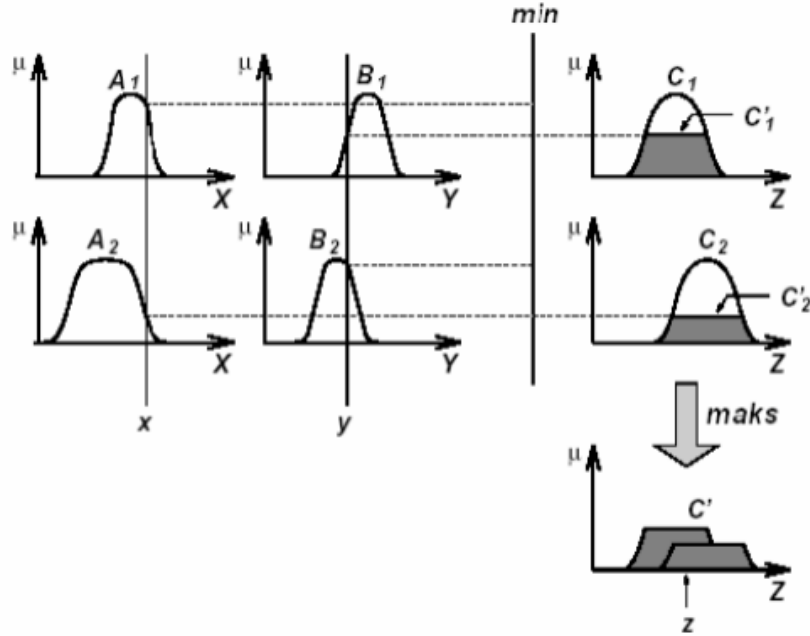
Bu modelde hem girdi değişkenleri hem de çıktı değişkeni kapalı formdaki üyelik fonksiyonları ile ifade edilir. Mamdani tipi bir bulanık model aşağıdaki 5 adımda oluşturulur:

- a) Girdilerin bulanıklaştırılması: Öncül kısımdaki bütün bulanık ifadeleri kullanarak girdi değişkenlerine ait 0 ile 1 arasında değişen üyelik derecelerinin belirlenmesi.
- b) Bulanık mantık işlemlerini kullanarak kural ağırlıklarının belirlenmesi
- c) Bulanık küme mantıksal işlemcilerin (“ve”, “veya”) uygulanması
- d) Sonuçların toplanması: Her bir kuralın çıktısını temsil eden bulanık kümelerin birleştirilmesi
- e) Durulaştırma: Tek bir sayıya dönüştürülmüş toplam bulanık küme sonuçlarının durulaştırılması [31].

Mamdani bulanık modeller girdi olarak klasik(kesin) değerler kullanırlar ve bulanık kümeyi klasik değere dönüştürmek için durulaştırma işlemini kullanırlar.

Mamdani tipi bulanık modelin avantajlarını özetlemek gerekirse;

- Modelin oluşturulması basittir.
- Diğer bulanık mantık modellemenin temelini oluşturur.
- İnsan davranış ve duyularına uygundur [43].



Şekil 3.2: Bulanık VE ve VEYA İşlemleri için Mamdani Tipi Bulanık Çıkarım Sisteminde Minimazasyon ve Maksimazasyon İşlemleri [57].

3.1.1.2 Takagi -Sugeno Tipi Bulanık Çıkarım Sistemi

Takagi-Sugeno tipi bulanık mantık çıkarım sistemi ya da Sugeno bulanık mantık ilk kez 1985 yılında kullanılmaya başlanmıştır. Mamdani bulanık çıkarım sisteminin bir uyarlamasıdır. Mamdani yöntemi genelde uzman sistemlerin geliştirilmesi için kullanılırken, veriden hareketli yaklaşımda Takagi-Sugeno modeli daha çok uygulanmaktadır [53]. Girdi değişkenlerinin bulanıklaştırılması ve bulanık mantık işlemleri Mamdani bulanık modelleme ile tamamen aynıdır. İki yöntem arasındaki fark çıktı üyelik fonksiyonlarındadır. Takagi-Sugeno tipi bulanık modellemede çıktı üyelik fonksiyonları lineerdir.

Bir başka deyişle bu modelleme yaklaşımında, model bulanık ifadelerden ve kurallar tarafından oluşturulur. Kuralın “Eğer” kısmı bulanık değişkenleri içeren bulanık ifadeleri, her bir bulanık ifadenin “öyleyse” kısmı ise modelin girdi değişkenlerinin parametrelerden oluşan eşitliğini gösterir. i . bulanık ifade için R^i

$$R^i = \text{Eğer } x_1 \rightarrow A_1 \text{ ve } x_2 \rightarrow A_2, \dots \text{ ve } x_k \rightarrow A_k \text{ ise } y^i = a_0^i + a_1^i x_1 + \dots + a_k^i x_k$$

Burada R^i ($i=1,2,\dots,c$) bulanık kuralı, x_i ($i=1,2,\dots,k$) girdi değişkenlerini, y^i bulanık kuralın çıktısını, a_j^i 'ler ise modelleme sürecinden elde edilen model parametrelerini göstermektedir. Modelin yapısı girdi değişkenlerin sayısı, bulanık değişkenlerin sayısı

veya herbir girdi değişkenin üyelik değerleri ile belirlenir. Modelin yapısı, modellenecek sistem hakkındaki ön bilgiye dayanarak önceden belirlenebilir.

Basit olarak bu sistem çok-girdili-tek-çıkıtlı sistem olarak varsayılır. Çok çıktı olması durumunda, her kural için $y_1^i, y_2^i, y_3^i, \dots, y_n^i$ şeklinde çok sayıda çıktı değişkeni olur. Ayrıca kuraldaki $A_1^i, A_2^i, \dots, A_k^i$, değişkenleri, bulanık alt uzayı temsil eden üçgen, yamuk, çan eğrisi veya diğer üyelik fonksiyonları ile birlikte kullanılan bulanık değişkenlerdir.

Bulanık modelin çıktısı \hat{y} , aşağıdaki gibi gösterilir.

$$\hat{y} = \frac{\sum_{i=1}^m w^i y^i}{\sum_{i=1}^m w^i} \quad (3.1)$$

Burada m kural sayısını, k ise girdi değişkenlerinin boyutunu göstermektedir. T-S dinamik ve statik lineer olmayan sistemleri modellemek için kullanılır ve girdi uzayının bulanık bölünmesine dayanır. T-S bulanık model lineer olmayan sistemleri birkaç farklı lineer sistemin kombinasyonu olarak varsayar. Bulanık modelleme yapılırken ilk önce kuralın sol tarafındaki öncül parametreler (premise parameter) elde edilir daha sonra, verilen öncül parametrelere göre kuralın sağ tarafındaki soncul parametreler (consequent parameter) ayarlanır. Bu işlem kompleks algoritma tarafından iteratif olarak devam eder [43].

T-S tipi bulanık modelin avantajları aşağıda sıralanmıştır;

- Lineer olmayan sistemlerin kontrol edilmesi için lineer teknikler kullanılabilir.
- Optimizasyon ve uyarlanabilir (adaptive) tekniklerle birlikte iyi çalışır ve çıktı parametrelerini optimize ederek sonuçları iyileştirir.
- Çıktı uzayında sürekliliği garantiler.
- Matematiksel analiz için uygundur.
- Hesaplama için çok uygundur.

T-S tipi bulanık modelin dezavantajları ise ;

- Yüksek derecedeki Sugeno bulanık modelleme kullanıldığında oldukça kompleks

bir yapıya sahip olur.

- Girdi ve alt küme sayılarının artması verilerin eğitilmesini zorlaştırır, sonuçların elde edilmesi için belirlenmesi gereken soncul parametrelerin sayısı artar.
- İnsan sezgilerine çok uygun değildir [43].

3.1.1.3 Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin En Küçük Kareler Yöntemine Göre Tahmin Edilmesi

Takagi-Sugeno tipi bulanık mantık sistemine ait kurallar eşitlik 3.2 de gösterilmektedir.

$$y^i = a_0^i + a_1^i x_1 + \dots + a_k^i x_k \quad (3.2) \quad (3.2)$$

Kurallarda yer alan katsayılar en küçük kareler yöntemi kullanılarak bulunur. Aşağıda bu parametrelerin bulanık mantık sisteminde bulunma adımları gösterilmiştir.

Her bir x_i girdisi için

$$\begin{aligned} f(x_1) = & \beta_1(a_0^1 + a_1^1 x_{11} + \dots + a_k^1 x_{k1}) + \\ & \beta_2(a_0^2 + a_1^2 x_{12} + \dots + a_k^2 x_{k2}) + \dots + \\ & \beta_n(a_0^n + a_1^n x_{1n} + \dots + a_k^n x_{kn}) \end{aligned}$$

Yukarıdaki denklemde m.kural için

$$\beta_m = \frac{\mu_m(x_0)}{\sum Mx_m} \quad (3.3) \quad (3.3)$$

$$\sum Mx_m = \sum_{j=1}^n \mu_{jm}(x_0) \quad (3.4) \quad (3.4)$$

Tüm kurallar için aynı işlemler yapıldığında

$$\begin{aligned} f_1(x_1) = & \beta_{11}(a_0^1 + a_1^1 x_{11} + \dots + a_k^1 x_{k1}) + \\ & \beta_{21}(a_0^2 + a_1^2 x_{12} + \dots + a_k^2 x_{k2}) + \dots + \\ & \beta_{n1}(a_0^n + a_1^n x_{1n} + \dots + a_k^n x_{kn}) \\ f_2(x_1) = & \beta_{12}(a_0^1 + a_1^1 x_{11} + \dots + a_k^1 x_{k1}) + \\ & \beta_{22}(a_0^2 + a_1^2 x_{12} + \dots + a_k^2 x_{k2}) + \dots + \\ & \beta_{n2}(a_0^n + a_1^n x_{1n} + \dots + a_k^n x_{kn}) \end{aligned}$$

$$f_w(x_1) = \beta_{1w}(a_0^1 + a_1^1 x_{11} + \dots + a_k^1 x_{k1}) + \beta_{2w}(a_0^2 + a_1^2 x_{12} + \dots + a_k^2 x_{k2}) + \dots + \beta_{nw}(a_0^n + a_1^n x_{1n} + \dots + a_k^n x_{kn})$$

$$\begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} \beta_{11}\mu_{11} & \beta_{11}\mu_{12} & \dots & \beta_{11}\mu_{n1} & \beta_{11} & \dots & \beta_{m1}\mu_{11} & \beta_{m1}\mu_{12} & \dots & \beta_{m1}\mu_{n1} & \beta_{m1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ \beta_{1w}\mu_{11} & \beta_{1w}\mu_{12} & \dots & \beta_{1w}\mu_{n1} & \beta_{1w} & \dots & \beta_{mw}\mu_{1m} & \beta_{mw}\mu_{2m} & \dots & \beta_{mw}\mu_{nm} & \beta_{mw} \end{bmatrix} \begin{bmatrix} a_0^1 \\ a_1^1 \\ \vdots \\ a_n^1 \\ a_0^m \\ a_1^m \\ \vdots \\ a_n^m \end{bmatrix} \quad (3.5)$$

Burada β_{ik} parametresi i. eleman ve k. küme merkezi x_k için

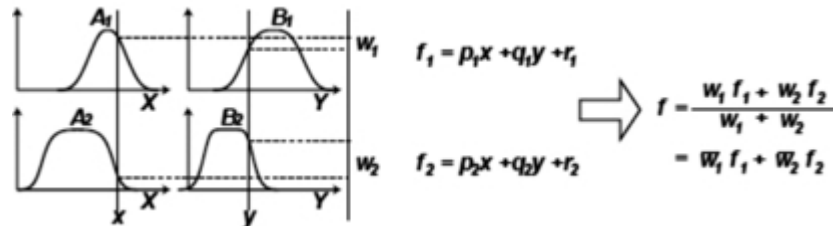
$$\beta_{ik} = \frac{\mu_{ik}(x_0)}{\sum M x_{ik}} \quad (3.6)$$

$$\sum M x_{ik} = \sum_{i=1}^w \mu_i(x_0) \quad (3.7)$$

Yukarıdaki matris çarpımını $B=AX$ şeklinde ifade edersek, A matrisi sabit sayıları, B matrisi çıktıları ve X matrisi de tahmin edilecek parametre değerlerini göstermektedir. En küçük kareler yöntemi kullanılarak problem çözülmektedir [58].

$$X = (A^T A)^{-1} A^T B \quad (3.8)$$

En küçük kareler yöntemi ile parametreler bulunduktan sonra kurallara ait denklemler oluşturulur.



Şekil 3.3: Sugeno Bulanık Çıkarım Sistemi.

3.1.1.4 Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin Kantil Yöntemine Göre Tahmin Edilmesi

Alma ve Vupa [59], yaptıkları çalışmada regresyon denkleminin parametre tahmininde en küçük kareler yöntemi ile farklı yöntemleri karşılaştırmışlardır. Veri setinde birden fazla sapan değer bulunduğunda, bu değerlerin birbirini maskeleyebildiği ve güvenilir verilerin sapan değer olarak görülmesine sebep olabildiği tespit edilmiştir. Bu nedenle hata terimlerinin normal dağılmadığı veya bağımlı değişkenin sapan değer içermesi durumlarında, küçük örneklem için regresyon modelinde, en küçük kareler yönteminin diğer yöntemlere göre daha fazla etkilendiğini belirtmişlerdir. Parametrelerin tahmin edilmesinde, aykırı veri seti etkilenmeyen Kantil yöntemi, alternatif bir yöntem olarak çalışmada kullanılacaktır.

Mostseller ve Tukey ; EKK regresyon eğrisinin dağılım kümeleri hakkında eksik bilgi verdiğini, Regresyon eğrisinin x 'lerin kümesine karşılık gelen dağılımların ortalamasının özet bilgisini vermek olduğunu, EKK yönteminden farklı olarak daha fazla bilgi için dağılımların çeşitli yüzde puanlarına denk gelen farklı regresyon eğrilerinin hesaplanabileceğini, bu sayede de kümenin daha detaylı bilgisi elde edilebileceğini belirtmişlerdir. [60].

Bir seriye ait gözlem değerleri küçükten büyüğe doğru sıralandığında, serinin toplam frekansını iki, dört, on veya yüz eşit kısma ayıran değerler, genelde kantil olarak tanımlanmaktadır [61]. Klasik doğrusal regresyonda, regresyon eğrisi gözlem noktalarının ortalarından veya ağırlık merkezlerinden geçer. Kantil Regresyon (KR) eğrileri ise kantillerinden geçmektedir. KR'de eğrinin izlediği yol asimetric olarak belirlenebilir. Örneğin, kantil değeri 0.10 olarak belirlendiğinde gözlemlerin %10'u regresyon doğrusunun altında bulunur. Bu sayede verilerinin birikimli dağılımının daha ayrıntılı bir görüntüsü elde edilebilir. Yani, 0.10 lik kantil düzeyiyle ilgileniliyorsa noktaların büyük çoğunluğu KR eğrisinin üzerinden geçecektir. Kantil regresyonda ortalama yerine kantillerin kullanılmasının önemli sebepleri vardır. Bu sebepler aşağıdaki gibi sıralanabilir.

- Dağılımın herhangi bir noktasının analiz edilmesini sağlar.
- Sağlam tahminler verir.
- Çarpık verilerin analizinde daha kullanışlıdır.

- Temsili bir değerle ilgilenildiğinde doğru bir yaklaşımdır.
- Dağılımın kuyruğuyla ilgilenildiğinde kullanılmalıdır.
- Tahmin edici, değişkenlerin etkilerinin hem konum hem de ölçek
- parametreleri yönünden incelenmesini sağlar.
- Hata terimlerinin parametrik dağılımı hakkındaki varsayımlardan kaçınan yarı parametrik bir yaklaşıma sahiptir. Bu durum, değişen varyanslı örneklemelerde KR'nin kullanımını sağlar.

Kantil regresyon eşitlik 3.9 şeklinde ifade edilir.

$$Y_i = X_i' \beta(\tau) + u_i(\tau) \quad (3.9)$$

$$Q_\tau(Y_i|X_i) = X_i' \beta(\tau) \quad (3.10)$$

Burada;

Y_i : Bağımlı değişkeni;

X_i : Bağımsız değişkenleri;

$u_i(\tau)$: hata terimlerini;

$\beta(\tau)$: dilim ile değişen parametre vektörünü;

$\tau \in (0,1)$: analiz edilen dilimi;

$Q_\tau(\cdot)$: Y_i 'nin birikimli koşullu dağılım fonksiyonu $F(\cdot)$ 'nin tersi olarak tanımlanan dilim fonksiyonudur.

Parametrelerin tahmini aşağıdaki minimizasyon probleminin çözümünü gerektirmektedir.

$$\text{Min}_{\beta \in \mathbb{R}^k} [\sum_{i \in \{i: Y_i < X_i \beta\}} \tau |Y_i - X_i \beta| + \sum_{i \in \{i: Y_i \geq X_i \beta\}} (1 - \tau) |Y_i - X_i \beta|] \quad (3.11)$$

Y_i 'nin τ . koşullu dilimi $Q_\tau(Y_i|X_i) = X_i' \beta(\tau)$ olduğundan, onun tahmini $\widehat{Q}_\tau(Y_i|X_i) = X_i' \beta(\tau)$ ile elde edilir. 3.11 numaralı denklemde $\tau = \frac{1}{2}$ hata terimlerinin eşit olarak ağırlıklandırıldığı medyan regresyonu verir. J.bağımsız değişken için dilim regresyon parametresi 3.12 eşitliği ile hesaplanır [62]:

$$\frac{\partial Q_\tau(Y_i|X_i)}{\partial X_{ij}} = \beta_j \quad (3.12)$$

3.1.1.5 Mamdani Bulanık Mantık Çıkarım Sistemi ile Takagi-Sugeno Bulanık Mantık Sisteminin Karşılaştırılması

Mamdani bulanık mantık çıkarım sisteminde;

- Modelin oluşturulması basittir.
- Diğer bulanık mantık modellemenin temelini oluşturur.
- İnsan davranış ve duyularına uygundur.

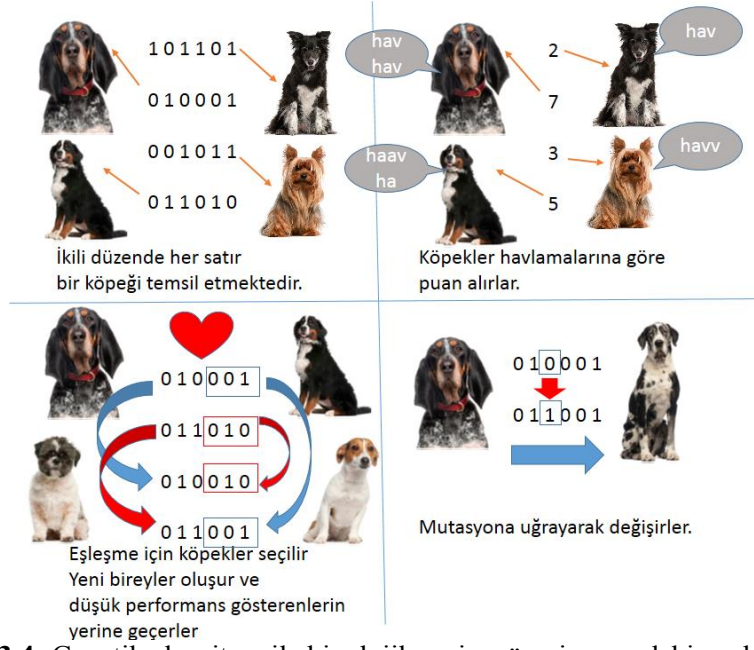
Takagi-Sugeno bulanık mantık çıkarım sistemine bakıldığında;

- Mamdani yöntemi genelde uzman sistemlerin geliştirilmesi için kullanılırken, veriden hareketli yaklaşımda Takagi-Sugeno modeli daha çok uygulanmaktadır .
- İki yöntem arasındaki fark çıktı üyelik fonksiyonlarındadır. Takagi-Sugeno tipi bulanık modellemede çıktı üyelik fonksiyonları lineerdir.

3.2 GENETİK ALGORİTMA

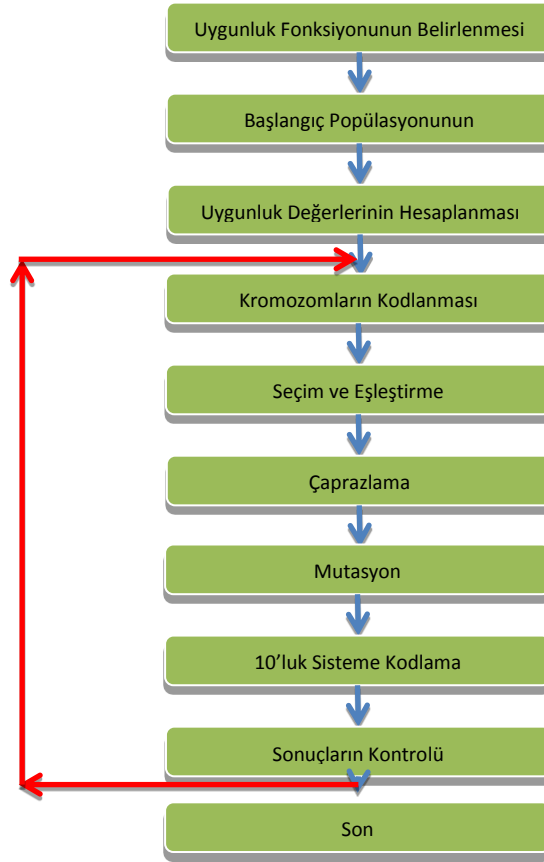
Genetik algoritma ilkeleri genetiğe ve doğal seleksiyona dayanan bir optimizasyon ve arama tekniğidir. Genetik algoritma ilk olarak 1975 yılında John Holland tarafından geliştirilmiştir. Daha sonra öğrencisi olan David Goldberg tarafından bir çok probleme uygulanmıştır.

Şekil 3.4'de biyolojik evrim ile ikilik sayı sistemindeki genetik algoritma arasındaki benzeşim gösterilmektedir. Başlangıçta rasgele bireylerden oluşan bir popülasyon bulunmaktadır. Her satırdaki sayılar okla gösterilen köpeğin yüksek sesle havlaması ile ilgili özelliklerini temsil etmektedir. Yüksek sesle havlayan köpekler aldıkları puanlara göre damızlık olarak tutulur. Damızlık köpeklerin yüksek havlama özellikleri bulunmaktadır. Dolayısıyla bu iki köpekten meydana gelecek yavru köpeklerinde yüksek havlama özelliğine sahip olması oranı oldukça yüksektir. Bu iki damızlık köpekten 2 adet yavru meydana gelir. İkili düzende, yavru köpekler, ebeveynlerinin kodlarının bazı bölümlerine sahiptirler. Bu iki yavru köpek başlangıç popülasyonunda seçilemeyen, yüksek havlama gücüne sahip olamayan köpeklerin yerine geçer. Böylece ilk popülasyonun değerine ulaşılmış olunur. Son aşamada oluşan bireylerde, ebeveynlerinin genlerinin dışında özellikler barındıran değişiklikler olabilmektedir [18].



Şekil 3.4: Genetik algoritma ile biyolojik evrim süreci arasındaki analogi [18].

Şekil 3.5'te genetik algoritmanın adımları görülmektedir.



Şekil 3.5 Genetik algoritma adımları.

1.Adım:Uygunluk Fonksiyonunun Belirlenmesi

Uygunluk fonksiyonu girdi değerlerini kullanarak çıktı değerleri oluşturur. Uygunluk fonksiyonu matematiksel bir fonksiyon olabileceği gibi, bir deney ya da oyun da olabilir. Amaç giriş değişkenlerinde bazı değişiklikler yaparak optimal sonuçlar elde edebilmektir.

2.Adım:Başlangıç Popülasyonunun Belirlenmesi:

Genetik algoritma bir grup kromozomla işlemlere başlar. Kromozomlar popülasyonu oluşturan bireylerdir. Kromozomun yapısı aşağıdaki gibidir.

$$c=c_1 c_2 c_3... c_n$$

burada $c_i, i=1,2,...,n$ genleri göstermektedir. Başlangıç popülasyonu rastgele kromozomlardan oluşabileceği gibi, hazır bir popülasyon ile de başlayabilir.

3.Adım:Uygunluk Değerlerinin Hesaplanması

Başlangıç popülasyonunu oluşturan kromozomlar uygunluk fonksiyonunda yerine konularak uygunluk değerleri hesaplanır. Bireylerin uygunluk değerleri bu bireylere genetik işlemler uygulanıp uygulanmayacağı konusunda yardımcı olur [40].

4.Adım:Kromozomların Kodlanması

Genetik algoritmada yaygın olarak ikilik sayı düzeninde kodlama kullanılmaktadır. Bu aşamada başlangıç popülasyonundaki bireyler ikilik sayı düzenine çevrilir. Çaprazlama ve mutasyon gibi genetik işlemlerde oldukça kullanışlıdır [39].

5.Adım:Seçim ve Eşleştirme

Seçim Yöntemi genetik algoritmanın en temel işlemlerindendir. İyi bireylerin bir sonraki nesle aktarılması ve genetik operatörler için seçilecek olan kromozomların uygunluk değerlerinin yüksek olması ile çözüme ulaşmak kolaylaşır. Seçim yöntemleri probleme göre seçilebilmektedir. Genel olarak rulet tekerleği, elitizm, turnuva seçimi gibi yöntemler bulunmaktadır.

Rulet tekerleği seçiminde, kromozomlar uygunluk değerlerinin yüzdelik oranlarında göre bir rulet tekerleğine yerleştirilirler. Uygunluk değeri oranı yüksek olan birey, rulet tekerleği üzerinde daha fazla bir alana sahiptir. Sonrasında rulet tekerleği üzerinde rastgele seçim yapılır.

Elitizim 1975’de Kenneth De Jong tarafından geliştirilen bir yöntemdir. Bu yöntemle en iyi sonucu veren birey, seçim, çaprazlama ve mutasyon işlemlerinin etkilerinden korunmuş olarak bir sonraki nesle aktarılır. Ne zamanki yeni bir birey uygunluk değeri noktasında daha iyi bir değer alır, o zaman bir önceki elitizim korumasındaki kromozom normal bir birey kabul edilir ve elitizim koruması yeni bireye geçer.

Turnuva seçiminde popülasyondan rastgele 2 birey seçilir. $[0,1]$ aralığında rastgele bir r sayısı seçilir. Eğer r sayısı k ($k=0.75$ olsun) sayısından küçükse bu iki birey ebeveyn olarak kabul edilir, değilse popülasyon havuzuna tekrar gönderilir ve bir sonraki aşamada tekrar seçilebilme hakları olur [40].

6.Adım:Çaprazlama

Seçim ve eşleştirme adımından sonra , seçilen ebeveynler ile yeni bireylerin oluşturulma sürecidir. [63]. Genetik algoritma çaprazlama işlemini uygunluk ölçütlerine göre seçilmiş iki ebeveyn, iyi özellikte yeni bireyler elde etmek için kullanılır. Çaprazlama ile birlikte popülasyon en iyiye doğru yavaş yavaş ilerlemektedir [37].

Çaprazlama operatörü iki ebeveyn kromozomun belli genlerinin birbirleri ile değişimi sonucu yeni bireylerin oluşması işlemidir. Hangi genden itibaren, genlerin kesilip değiş tokuş yapılacağı rastgele olarak belirlenir. Çaprazlama işlemi iki şekilde yapılabilmektedir; tek noktalı çaprazlama ve iki noktalı çaprazlama.

Tek noktalı çaprazlamada rastgele olarak maksimum ebeveynlerin gen sayısınınca bir sayı üretilir. Sayı üretildikten sonra 1.ebeveynin bu sayıya kadar olan genleri ile 2.ebeveynin bu sayıdan sonraki genleri birleştirilerek birinci yeni birey oluşturulur. Benzer şekilde 2.ebeveynin bu sayıya kadar olan genleri ile 1.ebeveynin bu sayıdan sonraki genleri birleştirilerek 2.birey oluşturulur.

Çift noktalı çaprazlamada 2 adet rastgele sayı üretilir. Bu sayılar en fazla ebeveynlerin gen sayısı kadar olabilir. Bu iki sayının arasındaki genlerin yer değiştirilmesi ile yeni bireyler oluşturulur.

7.Adım:Mutasyon

Çaprazlamadan sonra oluşan yeni popülasyon içerisindeki genlerden küçük bir kısmının değiştirilmesi işlemidir [38]. Mutasyon en iyilemeye doğru giden algorithmada çözümün belli bir bireyde kilitlenmesini engeller. Mutasyon oranının büyük olması iyileştirmeyi bozucu yönde etkiler. Bu yüzden mutasyon oranının küçük olması tercih edilir.

3.2.1 Genetik Algoritma Örneği

Bu örnekte $[0,15]$ aralığında $f(x)=x^2+1$ fonksiyonunu minimum yapan değerleri genetik algoritma kullanarak gösterilecektir. Aşağıda Genetik Algoritmanın adımları bu örnek için uygulanmıştır.

1.Adım: Uygunluk Fonksiyonunun Tanımlanması

Bu örnekte uygunluk fonksiyonu $f(x)=x^2+1$ 'dir. Amaç bu fonksiyonu minimum yapan değeri bulmaktır.

2.Adım:Başlangıç Popülasyonunun Belirlenmesi

Başlangıç popülasyonu hazır bir küme olarak verilebileceği gibi, program tarafından istenen sayıda rastgele oluşturulabilir. Bu örnekte $[0,15]$ aralığında rastgele 4 adet sayı belirlenmiştir. Aşağıda başlangıç popülasyonuna (P) ait elemanlar görülmektedir.

$P=[1,4,7,9]$;

3.Adım: Uygunluk Değerlerinin Hesaplanması

Uygunluk değerleri popülasyondaki bireylerin uygunluk fonksiyonundaki değerleridir. Tablo 3.1'de bireylerin uygunluk değerlerine göre aldıkları yüzde oranları görülmektedir.

Tablo 3.1: Bireylerin uygunluk deęerleri.

Birey No	Birey	Uygunluk Deęeri	Yüzde Oranı (%)
1	3	10	56
2	4	17	33
3	9	82	7
4	12	145	4
Toplam			100

4.Adım: Seçim İşlemi

Seçim işlemi için farklı yöntemler bulunmaktadır. Bu örnekte 100 elemanlı bir vektör oluşturulmuştur. Her birey sahip olduğu yüzde oranı kadar bu vektör içerisinde temsil edilir. En iyi sonucu veren birey bir sonraki nesle aynen aktarılır. Buna elitizm denir. Kalan bireyler 1-100 aralığında rastgele üretilen sayılarla belirlenir. Üretilen sayılar vektörün eleman sıra sayısıdır. Bu örnek için seçilen elemanlar aşağıda seçim (S) vektöründe gösterilmiştir.

$$S=[3,3,4,12];$$

5.Adım: Çaprazlama

Çaprazlamaya başlamadan önce bireylerin eşleştirilmesi gerekmektedir. Eşleştirme, çaprazlanacak adayların belirlenme işlemidir. En iyi sonucu veren bireyi bir sonraki nesle aktarmak için kendisi ile eşleştirilir. Bu örnek için kalan diğer eşleştirme rastgele üretilen sayılar sonucu bulunur. Toplamda 2 eşleştirme yapılacak ve bunun sonucunda toplam 4 birey oluşacaktır. Bu örnekte 1.-1. ve 2.-4. Eşlenerek çaprazlama yapılacaktır. Tablo 3.2’de seçim sonrası eşleşen bireyler görülmektedir.

Tablo 3.2: Seçim sonrası eşleşecek bireyler.

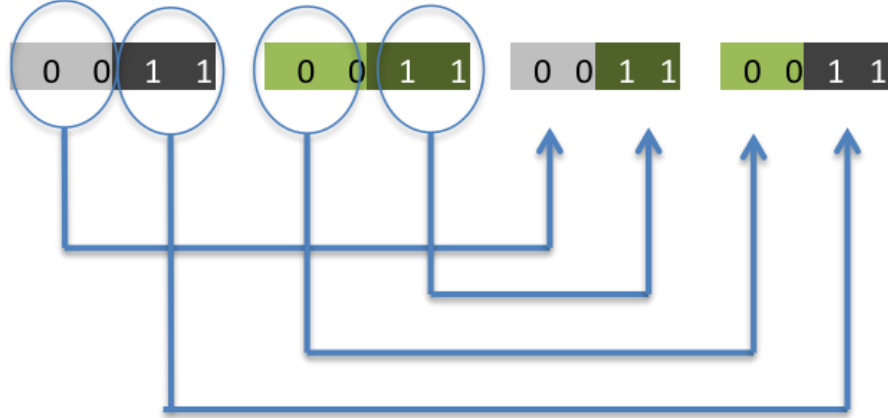
Birey No	Bireyler
1-1	3-3
2-4	3-12

Bu aşamadan sonra çaprazlama adımına geçilir. Çaprazlama yapılabilmesi için öncelikle tüm bireylerin ikilik sayı sistemine çevrilmesi gerekmektedir. Tablo 3.3’de eşleşme yapacak bireylerin ikilik sistemdeki karşılıkları görülmektedir.

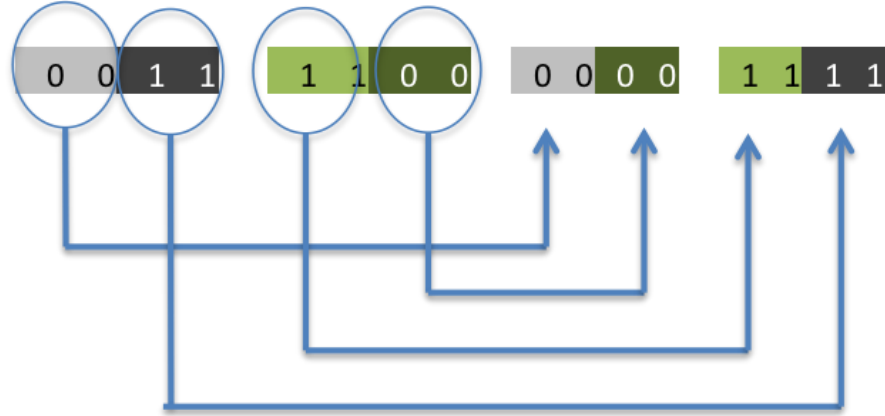
Tablo 3.3: Eşleşecek bireylerin ikilik düzende gösterimi.

Birey No	Bireyler
1-1	0011 - 0011
2-4	0011 - 1100

Çaprazlama tek noktalı ya da çift noktalı olarak yapılabilmektedir. Bu örnekte tek noktalı çaprazlama yapılacaktır. Çaprazlamanın yerinin belirlenebilmesi için öncelikle ikilik sistemdeki sayıların boyutuna bakılır. Burada 4 bitten oluşan sayılar bulunmaktadır. Bu durumda 1-4 aralığında rastgele bir sayı üretilir ve bu sayının olduğu kısımdan itibaren sayı bloğu kesilir. Bloğun solunda kalan kısma diğer sayının sağında kalan kısım eklenir. Aynı şekilde bloğun sağında kalan kısım diğer sayının solunda kalan kısma eklenerek eşleştirme yapılmış olunur. Bu örnek için üretilen rastgele sayımız 2’dir. Şekil 3.6 ve şekil 3.7’de eşleştirme sonrası yapılan çaprazlama işlemleri gösterilmiştir.



Şekil 3.6: 1-1 Bireylerin çaprazlama gösterimi.



Şekil 3.7: 2-4 Bireylerin çaprazlama gösterimi.

Çaprazlama sonrası oluşan yeni bireyler matris formunda aşağıda verilmiştir.

$$CS = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

6.Adım: Mutasyon

Mutasyon adımı genler içerisinde rastgele olarak bazı sayılar değişmektedir. Örneğimize baktığımızda CS matrisinin boyutları 4x4'dür. Matriste toplam 16 adet 0 ve 1'lerden oluşan elemanlar vardır. Her genetik algoritmada başlangıçta mutasyon değeri belirtilir. Mutasyon değeri, mutasyona girecek elemanların sayısını verecek olan orandır. Bu örnekte mutasyon oranı 0.15 olarak alınmıştır. Bunun anlamı toplam 16 sayının %15'i kadarının mutasyona uğrayacağıdır. Bu örnekte yaklaşık olarak 2 adet sayı mutasyona uğrayacaktır. Mutasyona uğrayacak sayıların belirlenmesi de yine rastgele sayı üretilerek yapılmaktadır. Satır ve sütun sayılarına göre rastgele 2 adet ikili üretilir. Bu sayıların birleşimi ile matrisde mutasyona uğrayacak sayının yeri belirlenmiş olur. Mutasyonun anlamı ilgili konumdaki sayı 1 ise 0, 0 ise 1 ile değiştirilmesidir. Bu örnekte rastgele olarak belirlenen sayı çiftleri (2,2) ve (4,3) dür. Mutasyon sonucu oluşan yeni nesil (YN) aşağıdaki matriste verilmiştir.

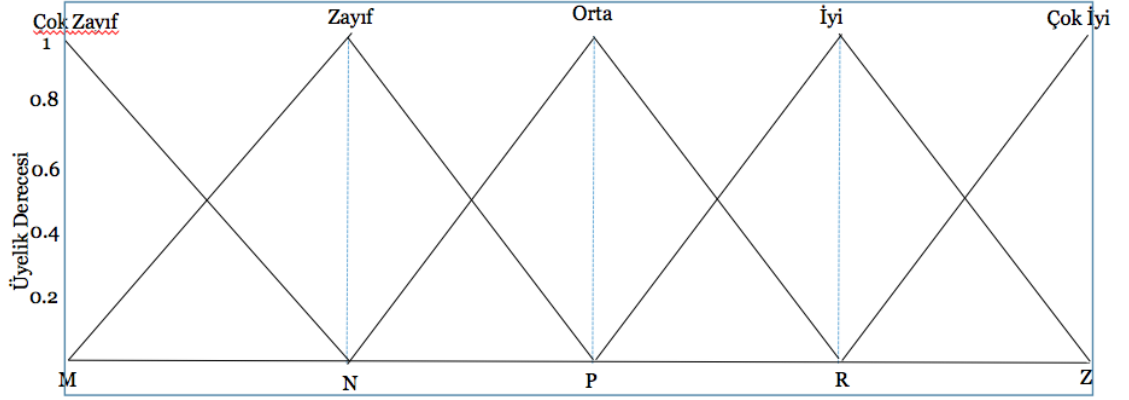
$$YN = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

Yeni nesile ait elemanları 10'luk sayı sistemine çevirdiğimizde, yeni nesli oluşturan elemanlar 3,7,0,13 olmaktadır. Bu sayılar arasında ki 0 değeri minimum değeri vereceği için algoritma 0 sayısından daha düşük sonuç veren bir sayı bulamadıkça, sayıyı bir sonraki nesle aktaracaktır. Algoritma bu sayılarla adımları tekrar eder. Ne zaman ki

istenen sonuca ya da belirlenen bir bitirme iterasyonuna ulaşırsa program çalışmasına son verir.

3.3 GENE-BULANIK MANTIK MELEZ YÖNTEMİ

Bulanık mantığın en önemli bileşenlerinden biri üyelik fonksiyonlarıdır. Şekil 3.8’de üçgen şeklinde üyelik fonksiyonları görülmektedir.



Şekil 3.8: Bulanık mantık üyelik fonksiyonları.

Üyelik fonksiyonu kendi içinde dilsel terimlerin oluşturduğu şekillerden oluşur. Bu şekillerden en çok kullanılanları üçgen, yamuk, parabol şekilleridir. Bu çalışmada üçgen şekli kullanılmıştır. Her dilsel terimin belli bir aralık değeri vardır. Klasik bulanık mantık da bu aralıklar matematiksel bir kaç işlem sonucunda bulunur. Uzman görüşüne dayanan sistemde ise deneyimler ön plana çıkmaktadır. Gene-bulanık metodu ise bu aralıkların başlangıç ve bitiş değerlerini her bir üyelik fonksiyonu için, genetik algoritmayı oluşturan adımlar işlenerek bulunur. Optimum sonuç bulunana kadar algoritma döngüye sokulur. Bulunan aralıklar bulanık mantık sistemine aktarılır. Bu yöntemin basit matematiksel hesaba dayanan klasik bulanık mantıktan veya deneyimlerin ön planda olduğu uzman görüşüne dayalı bulanık mantıktan den daha iyi sonuç vereceği varsayılmaktadır.

3.4 KÜMELEME YÖNTEMLERİ

Kümeleme analizi nesnelere birbirlerine olan benzerliklerine göre gruplara ayırır. Elde bulunan veriler incelenerek birbirine benzeyenler bir kümeye, benzemeyenler ise bir başka kümeye toplanırlar [64]. Kümeleme analizi genellikle sınıflandırma ile karıştırılabilmektedir. Sınıflandırma analizinde önceden belirlenmiş sınıflar bulunmaktadır. Elde bulunan veri bu sınıflara göre ayrılmaktadır. Kümeleme analizinde

ise daha öncesinde tanımlanmış herhangi bir sınıf olmaksızın benzer özellik gösteren verileri aynı kümede toplayarak veriyi anlamlı ve faydalı gruplara ayırır [65].

Kümeleme analizinin asıl amacı, gruplanmamış verileri benzerliklerine göre gruplandırmak ve araştırmacıya özetleyici bilgiler sunmaktır. Kümeleme analizi bundan başka gerçek tiplerin belirlenmesi, gruplar için ön tahmin, hipotez testi, veriler yerine kümelerin değerlendirilmesi ve aykırı değerlerin bulunması gibi amaçlarla da kullanılmaktadır [66].

Kümeleme analizinde, istatistiki yöntemlerde önem taşıyan normallik, doğrusallık ve eş varyansa sahip olma gibi varsayımlar fazla dikkate alınmamaktadır. Ancak kümeleme analizinde örneklemin anakütleyi temsil gücü ve çoklu doğrusal bağlantı dikkate alınması gereken iki önemli konu olmaktadır. Araştırmacı öncelikle elindeki örneklemin anakütleyi yeterli düzeyde temsil ettiği konusunda emin olmalıdır. Çünkü kümeleme analizinde kümeler anakütlenin yapısını göstermek için türetilir. Araştırmacılar kümeleme analizinin, örneklemin temsil yeteneğinin ne ölçüde iyi olduğuna bağlı olarak başarı sağlayan bir yöntem olduğunu bilmelidirler [67]. Diğer çok değişkenli istatistiksel yöntemlerin bazılarında çoklu bağıntı probleminin var olması durumunda, aralarında çoklu doğrusal bağlantı bulunan değişkenlerin gerçek etkilerini ayırt etmek genellikle önemli bir sorun olarak ortaya çıkmaktadır. Ancak kümeleme analizinde bu sorun önem taşımamaktadır. Çünkü kümeleme analizinde, çoklu doğrusal bağlantı gösteren değişkenler üstü kapalı olarak daha fazla ağırlıklandırılarak bu sorun ortadan kaldırılmaktadır [68].

Küme sayısının belirlenmesi konusunda son yıllarda yoğun çalışmalar vardır. Ancak halen çok da güvenilir olmayan ve 1970'lerde bulunmuş olan teknikler kullanılmaktadır. En pratik yöntem olarak küme sayısı (k);

$$k \cong \sqrt{(n / 2)} \quad (3.13)$$

biçiminde belirlenmektedir. Fakat örneklem sayısı büyüdüğünde bu yöntemle iyi sonuçlar alınamamaktadır. K-ortalamlar kümeleme metodunda, birimlerin incelenmesi ile k=2'den başlayarak ve küme sayılarını her defasında birer arttırarak, deneysel bir biçimde en uygun kümelemeyi bulmayı amaçlayarak da uygulanabilmektedir [66].

Kümeleme yöntemleri genel olarak iki ana gruba ayrılmaktadır. Bunlar hiyerarşik kümeleme yöntemleri ve hiyerarşik olmayan kümeleme yöntemleridir.

Hiyerarşik kümeleme yönteminde başlangıçta her birey bir küme olarak görülmekte, daha sonra en yakın iki küme birleştirilerek, küme sayısı indirgenmekte ve bu şekilde devam edilerek en uygun kümeleme işlemi yapılmaktadır.

Hiyerarşik olmayan kümeleme yöntemleri ise küme sayısı hakkında bir ön bilgi ya da araştırmacının küme sayısını belirttiği durumlarda kullanılır [42].

Öte yandan kümeleme analizinde küme sayısı önceden bilinmemektedir. Sadece verilerin mevcut durumlarından yola çıkılarak kümeler elde edilmektedir. Bu nedenle, elde edilen sonuçlar gelecek tahmininde kullanılamaz. Ayrıca kümeleme analizinde verilere ilişkin olarak normallik varsayımı olmasına rağmen sadece uzaklık değerlerinin normalliği yeterli görülmektedir. Kovaryans matrisine ilişkin herhangi bir varsayım da bulunmamaktadır [69].

Kümeleme analizi metotlarından önce kümeleme analizi öncesinde yapılması gereken analizlerden çok değişkenli aykırı veri belirleme yöntemleri ve verinin normallik testi gibi yöntemler anlatılmıştır. Daha sonrasında çalışmada kullanılan hiyerarşik olmayan kümeleme yöntemlerinden K-Ortalamlar Yöntemi, Bulanık C-Ortalamlar Yöntemi ve Çıkarımlı Kümeleme Yöntemi incelenmiştir.

3.4.1 Çok Değişkenli Aykırı Veri Belirleme Yöntemleri

Çok değişkenli aykırı veri belirlemede 3 yöntemden faydalanılacaktır. Bu yöntemler Mahalanobis Uzaklığı , Kaldıraç Noktası ve Cook Uzaklığıdır.

3.4.1.1 Mahalanobis Uzaklığı

Tek değişkenli uzaklık tekniğiyle yakın ilişkili olmasının yanı sıra, çoklu normal ve geniş örneklerde kullanılan Mahalanobis uzaklığı, değişken sayısına bağlı olarak elde edilen serbestlik dereceli ki-kare (χ^2) dağılımına uyar ve örneklem sayısı arttıkça da bu kestirim daha iyi hale gelir. Mahalanobis uzaklığı değerinin, bağımsız değişken sayısını serbestlik derecesi alan χ^2 tablo değeriyle karşılaştırılması gerekir. Hesaplanan Mahalanobis uzaklığı değeri, bulunan tablo değerinden küçük ise, çok değişkenli

normallik varsayımının karşılandığı söylenebilir. χ^2 tablo değerinin üzerinde bir Mahalanobis uzaklığı değerine sahip olan denekler, aykırı veri olarak belirlenmektedir ve verilerden çıkartılabilmektedir. Bu işlemler için 0.01 ya da daha tutucu olmak isteniyorsa 0.001 manidarlık düzeyinin dikkate alınması önerilmektedir [70].

3.4.1.2 Cook Uzaklığı

Cook (1977) tarafından sunulan Cook uzaklığı genel olarak, belirli bir modele göre bir gözlemin etkisini göstermek üzere geliştirilmiş bir yöntemdir. Cook uzaklığı her bir birimin tüm tahmin değeri (n sayıdaki) üzerindeki etkileri ölçmektedir. Cook uzaklığı değeri 1'den büyük olan birimlerin etkili birimler olduğuna karar verilir fakat bazı istatistikçiler $4/(n-2)$ kritik değerini önermektedir [71].

Cook Uzaklığı'nın 1'den büyük (Cook Uzaklığı >1) değer aldığı gözlemler, uç değer içeren gözlem olarak tanımlanmaktadır. Bu durumda bu gözlem veri kümesinden ya atılır, ya model yeniden ele alınır ya da modele yeni değişkenler eklenir [72]. Fakat çalışmalarda genellikle $4/n$ (n gözlem sayısı) değeri ile karşılaştırılarak karar verilir. $4/n$ değerinden büyük değerlerin uç değer olduğu ifade edilir [73].

3.4.1.3 Kaldıraç Nokta (Leverage Değeri)

Veri noktasının merkezinden oldukça uzakta olan nokta, regresyon doğrusunu büyük bir güçle çektiğinden dolayı bu nokta Kaldıraç Noktası (leverage point) olarak adlandırılırlar [73].

Mahalanobis Uzaklığı ve Cook'un Uzaklığı gibi popüler indeksler, regresyon çizgisinin değerini tahmin ederek Kaldıraç Noktasını hesaplayan ve sıklıkla kullanılan yöntemlerdir. Kaldıraç Noktası, çok değişkenli veri setlerinde uç değerleri belirlemek için kullanılan istatistiksel bir ölçüdür. Kaldıraç Noktası, Mahalanobis Uzaklığı ile ilişkilidir; ancak farklı bir ölçekte ölçüldüğü için ki kare dağılımına dayalı manidarlık testleri uygulanamaz. Kaldıraç Noktası, 0-1 aralığında değişen bir değer almaktadır. Kaldıraç Noktası=0, gözlemler üzerinde hiçbir etkisi yok, Kaldıraç Noktası=1 gözleme tahmin edilenden çok daha fazla etki ettiği anlamına gelir. Bir başka deyişle, Kaldıraç Noktası 1'e yaklaştıkça, gözlemin uç değer içermeye olasılığı artar (Field, 2009). Kaldıraç gücü $2p/n$ 'den (p : değişken sayısı, n gözlem sayısı) büyük olduğunda, Belsley'e (1980) göre yüksek kaldıraç (uç deger)'dir. Küçük örneklem için Vellman ve Welsch (1981) $3p/n$ 'den (p : değişken sayısı, n : kişi sayısı) ölçütünü önerir [72].

Mahalanobis Uzaklığı, Cook Uzaklığı ve Kaldıraç Noktası regresyon analizi çatısında geliştirilmişlerdir ve birçok istatistiksel paket program tarafından kolaylıkla uygulanabilmektedir.

3.4.2 K-Ortalamalar Kümeleme Yöntemi

İlk olarak 1965’de Forgy ve 1967’de MacQueen tarafından ortaya atılan hata kareler azaltma ilkesinden sonra, 1979’da Hartigan ve Wong tarafından tanımlanmıştır. K-Ortalamalar algoritması, geniş çaplı kullanılan bir kümeleme algoritmasıdır. Adını, küme merkezi olarak tanımlanan k tane kümenin her birini temsil eden noktaların ortalamalarından alır [42].

K-ortalamar yöntemi, ilk önce n adet nesneden rasgele k adet nesne seçer ve bu nesnelerin her biri, bir kümenin merkezini veya orta noktasını temsil eder. Geriye kalan nesnelere her biri kendisine en yakın olan küme merkezine göre kümelere dağılırlar. Yani bir nesne hangi kümenin merkezine daha yakın ise o kümeye yerleşir. Ardından her küme için ortalama hesaplanır ve hesaplanan bu değer o kümenin yeni merkezi olur. Bu işlem tüm nesnelere kümelere yerleşinceye kadar devam eder [41].

3.4.3 Bulanık C-Ortalama Yöntemi

Kümeleme aslında, veri kümesinin uygun benzerlik ölçülerine göre homojen sınıflara ayrılması ile ilgilidir. Herhangi bir kümeye ait olan birimler benzer, farklı kümelerdeki birimler ise mümkün olduğunca farklıdır. Klasik kümeleme analizinde, farklı kümelerin sınırları keskindir ve bir birim yalnızca bir kümeye ait olabilir. Pratikte, veriler genelde iyi bir dağılıma sahip olmazlar, yani sınırlar kesin olarak tanımlanamaz. O zaman, bir birim farklı üyelik dereceleri ile birden fazla kümeye ait olabilir. Bu durum klasik kümeleme analizine benzer olarak bulanık kümelemede de farklı yöntemler ile ifade edilebilir [74].

Bulanık c-means algoritması da amaç fonksiyonu temelli bir metottur. Bulanık c-means metodu, nesnelerin iki veya daha fazla kümeye ait olabilmesine izin verir. Bulanık mantık prensibi gereği her veri, kümelerin her birine $[0,1]$ arasında değişen birer üyelik değeri ile aittir. Bir verinin tüm sınıflara olan üyelik değerleri toplamı “1” olmalıdır. Nesne hangi küme merkezine yakın ise o kümeye ait olma üyeliği diğer kümelere ait

olma üyeliğinden daha büyük olacaktır. Amaç fonksiyonun belirlenen minimum ilerleme değerine yakınsaklaşmasıyla kümeleme işlemi tamamlanır.

Algoritma, en küçük kareler yönteminin genellemesi olan aşağıdaki amaç fonksiyonunu öteleyerek minimize etmek için çalışır.

$$Jm = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad 1 \leq m \leq \infty \quad (3.14)$$

U üyelik matrisi rasgele atanarak algoritma başlatılır. İkinci adımda ise merkez vektörleri hesaplanır. Merkezler 3.15 nolu eşitlik ile hesaplanır .

$$c_i = \frac{\sum_{i=1}^N u_{ij}^m x_i}{\sum_{i=1}^N u_{ij}^m} \quad (3.15) \quad (3.15)$$

Hesaplanan küme merkezlerine göre, U matrisi 3.16 nolu eşitlik kullanılarak yeniden hesaplanır. Eski U matrisi ile yeni U matrisi karşılaştırılır ve fark ε 'dan küçük olana kadar işlemler devam eder .

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_i\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad (3.16) \quad (3.16)$$

Kümeleme işlemi sonucunda bulanık değerler içeren U üyelik matrisi kümelemenin sonucunu yansıtır. İstenirse, berraklaştırma yapılarak bu değerler yuvarlanıp 0 ve 1'lere dönüştürülebilir [75].

3.4.4 Çıkarımlı Kümeleme Yöntemi

Çıkarımlı Kümeleme Yöntemi (ÇKY), veri noktalarının yoğunluk ölçütüne dayanır. Bu yöntem, yüksek yoğunluklu veri noktalarının özellik uzayındaki bölgelerinin saptanmasına dayanır. En yüksek sayıda komşusu olan nokta, bir küme için merkez olarak seçilir. ÇKY'de, her bir veri noktası x_i , küme merkezi için aday olduğundan dolayı birinci küme merkezi için yoğunluk ölçütü eşitlik 3.17'de tanımlanmıştır:

$$D_i = \sum_{j=1}^n e^{-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}} \quad (3.17) \quad (3.17)$$

Burada, r_a komşuluk yarıçapını gösteren pozitif bir sabittir. Buna göre, bir veri noktası çok sayıda komşu veri noktalarına sahip ise, bu veri noktası yüksek bir yoğunluğa sahip

olacaktır. En büyük yoğunluk değeri D_{c1} 'e sahip nokta, birinci küme merkezi x_{c1} olarak seçilir. Her bir veri noktasının yoğunluğu eşitlik 3.18 ile elde edilir:

$$D_i = D_i - D_{ci} e^{-\frac{\|x_i - x_{ci}\|^2}{(r_b/2)^2}} \quad (3.18) \quad (3.18)$$

Buradaki r_b , yoğunluk değerindeki yeterli azaltmaya sahip komşuluğu tanımlayan pozitif bir sabittir. Bu yüzden, birinci küme merkezi x_{c1} 'e yakın veri noktaları önemli ölçüde azaltılmış yoğunluk değerine sahip olacaktır [76]. Bir sonraki küme merkezinin seçilebilmesi için aşağıdaki kriterlerin doğrulanması gerekmektedir.

Eğer $D_{ci} > \bar{\epsilon} D_1$ ise

x_{ci} yeni küme merkezi olarak kabul edilir ve süreç devam eder.

Eğer $D_{ci} < \underline{\epsilon} D_1$ ise

x_{ci} yeni küme merkezi olarak kabul edilmez ve süreç devam eder.

Eğer ikiside değilse

$d_{\min=x_{ci}}$ ile önceki bulunan küme merkezleri arasında en kısa mesafe alınır

Eğer $\frac{d_{\min}}{r_a} + \frac{D_{ci}}{D_1} \geq 1$ ise x_{ci} yeni küme merkezi olarak kabul edilir.

Değilse x_{ci} küme merkezi olarak kabul edilmez ve bir sonraki en büyük yoğunluğa sahip küme merkezi adayı seçilir ve tekrar yukarıdaki testlerden geçirilir. Burada $\bar{\epsilon}$ bir eşik değeridir ve yoğunlukla beraber bu değerden fazla ile veri noktası küme merkezi olarak seçilir. $\underline{\epsilon}$ değeri de küme merkezi olmama durumu için kullanılan eşik değeridir. Bu değerler genellikle $\bar{\epsilon}=0,5$ ve $\underline{\epsilon} = 0,15$ dir [77].

4. BULGULAR

4.1 VERİ KÜMESİ

4.1.1 Zaman, Frekans, Süre (ZFS) Veri Kümesi

ZFS veri kümesi, 2010-2011 öğretim yılı, Yıldız Teknik Üniversitesi'nde web tabanlı olarak verilen Matlab dersine aittir. Toplam 62 kayıtlı öğrenci bulunmaktadır. 16 öğrencinin herhangi bir etkinliği olmadığından dolayı 46 öğrenci ile ilgili 6 haftalık ZFS verilerinden oluşmaktadır. Bu veriler dersin verilmiş olduğu uzaktan eğitim platformu olan Moodle üzerinden alınmıştır.

4.1.2 ZFS, Küçük Sınav ve Ara Sınav Veri Kümesi

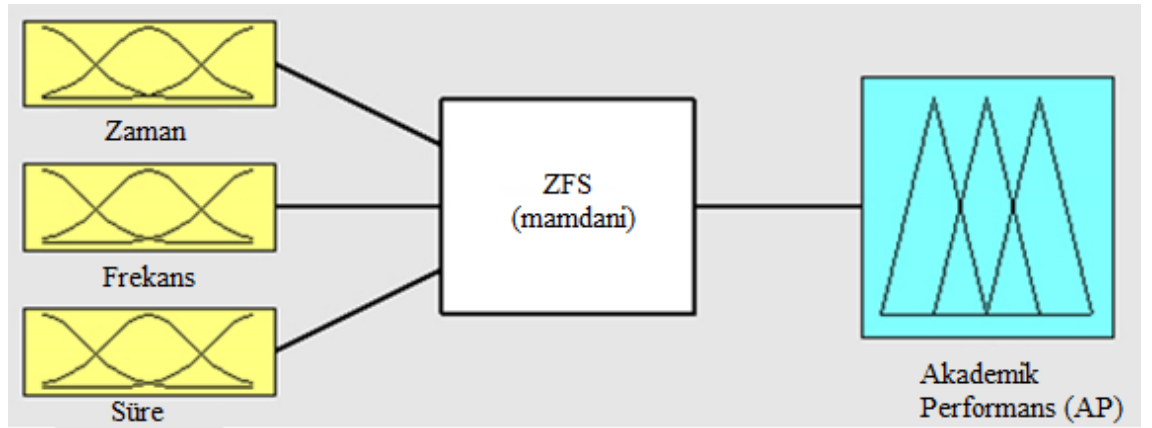
ZFS veri kümesi kullanılarak yapılan model sonuçlarında görülmüştür ki daha yüksek doğruluk oranları elde etmek için ZFS verileri tek başına yeterli olmamaktadır. Bu sebeple sisteme ilave olarak iki girdi değişkeni daha eklenmiştir. Bunlar küçük sınav ve ara sınav sonuçlarıdır. 46 kişilik ZFS verisinde küçük sınav ve ara sınav sonuçları bulunmadığından yeni bir veri kümesi çalışması yapılmıştır.

ZFS, küçük sınav ve ara sınav veri Kümesi, 2011-2012 öğretim yılı ve 2012-2013 öğretim yılı, Yıldız Teknik Üniversitesi'nde verilen Temel Bilgisayar Bilimleri dersine kayıtlı 337 öğrenciye aittir. 337 öğrenciden 24'ü bir dönem boyunca uzaktan eğitimdeki hiçbir etkinliğe katılmadığından dolayı kapsam dışı bırakılmış ve çalışma 313 öğrenci ile yapılmıştır. Bu çalışmada demografik özellikler analize katılmamıştır. Çalışma kapsamında 5 adet giriş verisi ile birlikte 1 adet çıkış verisi kullanılmıştır. Giriş verilerinden ZFS verileri, dersin verilmiş olduğu uzaktan eğitim platformu olan Moodle üzerinden alınmıştır. Dönem başından itibaren 6 haftalık veriler önce log dosyası olarak Moodle üzerinden alınmıştır. Log dosyası yaklaşık olarak 55bin satırdan oluşmaktadır. Bu veri içerisinde Zaman, Frekans ve Süre değerlerinin her öğrenci için ayrı ayrı hesaplanması, Matlab ortamında yazılan yazılımla sağlanmıştır. Dördüncü veri olarak 4.haftada moodle üzerinden online olarak yapılan küçük sınav sonucu alınmıştır. Sekizinci haftada örgün olarak uygulanan arasınnav sonucu beşinci girdi olarak alınmıştır. Dersin kapsamında bir dönem boyunca öğrenciler üç adet online küçük

sınav, iki adet arasınava ve bir adet final sınavı olmaktadır. Bir öğrencinin yıl sonu akademik başarısına etki eden bu sınavların ağırlık dereceleri; üç adet online küçük sınav %20, iki adet arasınava %40 ve final sınavı %40 değerinde olmaktadır.

4.2 UYGULAMA 1: ZFS VERİLERİ KULLANILARAK OLUŞTURULAN BULANIK MANTIK YÖNTEMİ İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME

Bu uygulamada uzaktan eğitim öğrencilerinin akademik performansları ZFS veri kümesi ile bulanık mantık yöntemi ile tahmin edilmeye çalışılmıştır. Model kurulurken dilsel terimler kullanılmıştır. ZFS verileri için “zayıf”, “orta”, “iyi” dilsel terimleri belirlenmiştir. Şekil 4.1, kurulan modelin yapısını göstermektedir.



Şekil 4.1: ZFS Bulanık Mantık Modeli.

Bulanık modelde giriş değişkenleri olarak Zaman, Frekans ve Süre olmak üzere üç giriş değişkeni kullanılmıştır. Merkezde yer alan kısımda kurallar dizisi bulunmaktadır. Toplamda 27 kural belirlenmiştir. Bu kurallar

“Eğer Zaman zayıfsa ve Frekans zayıfsa ve Süre zayıfsa, öyleyse akademik performans zayıftır” şeklindedir. Bu kural biraz daha uzun şekilde tarif edilmek istendiğinde; uzaktan eğitimdeki bir öğrenci ders sisteme yüklendikten çok sonraları derse bakmaya başlamışsa, sisteme girme çıkma sıklığı çok azsa ve girdiğinde de sistemde çok az kalmışsa, bu öğrencinin tahmini akademik performansı zayıf olacaktır.

Model üç girdi değişkenine karşılık sadece tek çıktı değişkeni sağlamaktadır ki bu da akademik performansa denk gelmektedir.

Bulanık modelde kurallardan sonraki en önemli kısım üyelik fonksiyonlarının aralıklarının belirlenmesidir. Üyelik fonksiyonlarına ait aralıklar iki şekilde

belirlenmiştir. Birinci olarak aralıklar 3 eşit parçaya bölünmüştür. Örneğin Zaman bir öğrencinin, konu sisteme girildikten kaç gün sonra sisteme girdiğini gösteren değişkendir ve bu değişken 1 ile 36 arasında değişmektedir. Bu aralık üç eşit parçaya bölüldüğünde;

- 1-12 arası: İyi
- 12-24 arası: Orta
- 24-36 arası: Zayıf

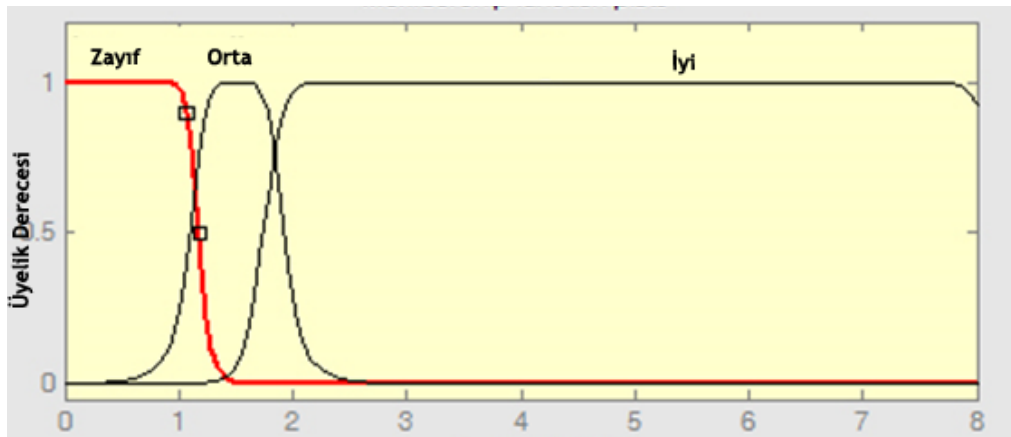
şeklinde olmaktadır. Aynı aralık uzman görüşüne göre bölüldüğünde;

- 1-7 arası: İyi
- 7-14 arası: Orta
- 14-36 arası: Zayıf

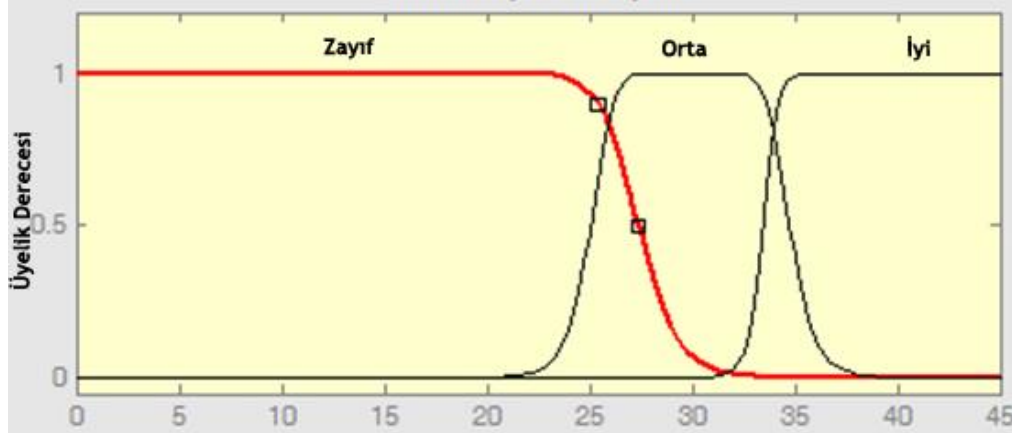
Şeklinde olmaktadır. Şekil 4.2,4.3,4.4'de ZFS üyelik fonksiyonlarının grafikleri görülmektedir.



Şekil 4.2: Zaman üyelik fonksiyonu.



Şekil 4.3: Frekans üyelik fonksiyonu.



Şekil 4.4: Süre üyelik fonksiyonu.

Klasik ve uzman görüşüne dayanılarak kurulan modelde ayrıca üyelik fonksiyon tipi değiştirilerek de model çalıştırılmış ve anlamlı farklılık olup olmadığına bakılmıştır. Klasik şekilde kurulan modele ait sonuçlar tablo 4.1'de gösterilmiştir.

Tablo 4.1: Klasik bulanık mantık modeline bağlı doğruluk oranları.

Üyelik Fonksiyonu	Doğruluk Oranı (%)
Trimf	65,25
Gauss2mf	64,81
Pimf	64,62
Gbellmf	64,28

Klasik bulanık mantık modeline bağlı üyelik fonksiyonları ile yapılan çalışmalarda en iyi sonucu üçgen üyelik fonksiyonu sağlamıştır. Tablo 4.2'de uzman görüşüne dayalı kurulan bulanık mantık modeline bağlı sonuçlar bulunmaktadır.

Tablo 4.2: Uzman görüşlü bulanık mantık modeline bağlı doğruluk oranları.

Üyelik Fonksiyonu	Doğruluk Oranı (%)
Gbellmf	74,40
Gauss2mf	73,40
Pimf	73,29
Trimf	72,90

Tablo 4.2'ye göre en yüksek doğruluk oranı geliştirilmiş çan eğrisi üyelik fonksiyonu (Gbellmf) ile alınmıştır. 6 haftalık ZFS verileri ile yapılan çalışmada üyelik fonksiyon aralıklarının değiştirilmesi, sonucu anlamlı derecede değiştirdiği görülmektedir.

4.3 UYGULAMA 2: ZFS, QUIZ VE ARASINAV VERİLERİ KULLANILARAK OLUŞTURULAN GENE-BULANIK MANTIK HİBRİT YÖNTEMİ İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME

Bu uygulamada uzaktan eğitim öğrencilerinin akademik performansları üç metot ile tahmin edilmeye çalışılmıştır. Bunlar Klasik Bulanık Mantık, Uzman Görüşüne Dayalı Bulanık Mantık ve Genetik Algoritma ile birlikte bulanık mantık yönteminin melez olarak kullanıldığı Gene-Bulanık Mantık Modelidir.

4.3.1 Klasik Bulanık Mantık Modeli

Bu çalışmada oluşturulan bulanık Mantık Sistemine ait beş girdi ve bir çıktı vardır. Zaman, Frekans, Süre, Küçük sınav ve Arasnav notu girdi değerleridir. Öğrencinin yıl sonu akademik performans notu çıktı değeridir. Her girdi ve çıktı değeri için bir adet üçgen üyelik fonksiyonu oluşturulmuştur. Tüm veriler normal dağılım kullanılarak maksimum ve minimum noktalar tespit edilmiştir. Bu noktalar kullanılarak (M,N,P,R,Z) üyelik fonksiyonu aralıkları tespit edilmiştir.

X =Normal Dağılımda %95 değeri

Y =Normal Dağılımda %5 değeri

Z =Her bir veri içinde en yüksek değer

T =Her bir veri içinde en düşük değer

Max =En büyük(X,Z)

Min =En Küçük(Y,T)

a =($Max-Min$)/4

M = Min

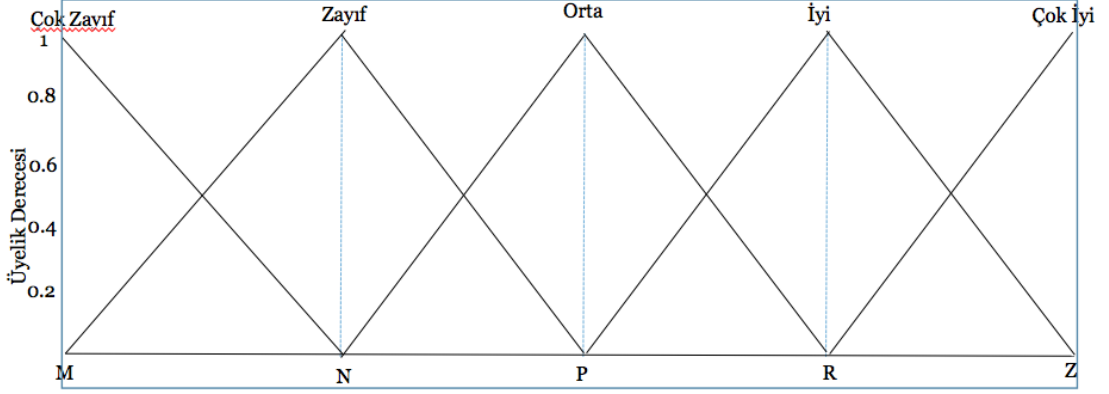
N = $Min+a$ veya $Max-3a$

P = $Min+2x$ veya $Max-2x$

R = $Min+3x$ veya $Max-x$

Z=Max

Tüm bu hesaplamalar sonucunda şekil 4.5'te görülen üyelik fonksiyonu elde edilmiştir.

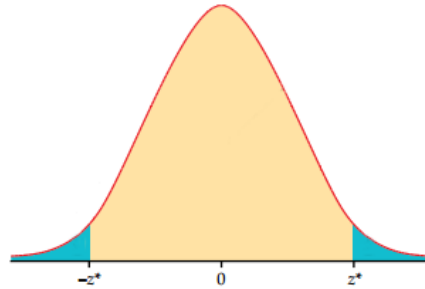


Şekil 4.5: Klasik bulanık mantık üyelik fonksiyonu.

Bu çalışma 218 adet öğrenci verisi ile yapılmıştır. Bu verilerin %70'i rastgele seçilerek eğitim verisi olarak, kalan %30'luk veri de test verisi olarak kullanılmıştır. Her bir girdi ve çıktı için üyelik fonksiyonu aralıkları yukarıda adımları verilen işlemler sonucunda bulunmuştur. Tablo 4.3'te örnek olarak Süre girdi değerlerine ait ortalama ve standart sapma değerleri görülmektedir.

Tablo 4.3: Süre girdisi için ortalama ve standart sapma.

Ortalama	22.16
Standart Sapma	15.79



Şekil 4.6 Normal dağılım z değerleri.

Süre değişkenine ait aralıkların bulunması şu şekilde olmaktadır; Öncelikle normal dağılıma göre ilk ve son %5'lik dilime karşılık gelen değerler $Z = \frac{x-\mu}{\sigma}$ formülü kullanılarak bulunur. İlk %5'lik dilime gelen z^* değeri -1.65, son %5'lik dilime karşılık gelen z^* değeri +1.65'dir [78]. Formül uygulandığında $x=-3.90$ ve $x=48.23$ değerleri

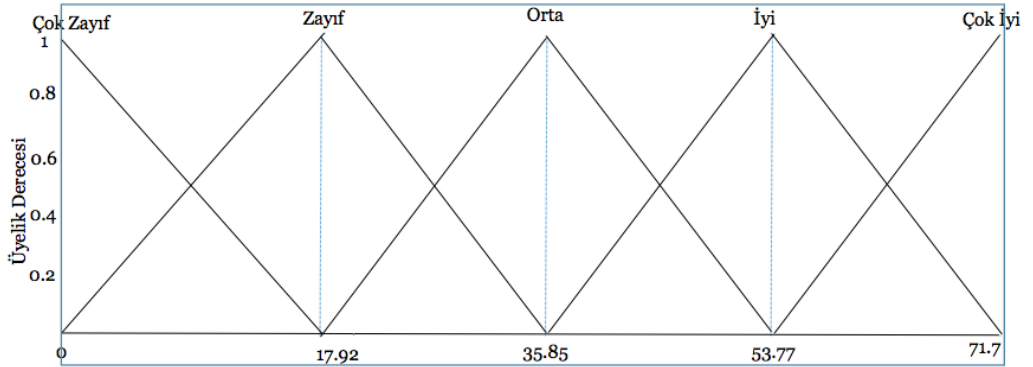
bulunur. Değerlere bakıldığında minimum değer negatif olduğundan, sınır 0 kabul edilir. Üst değere bakıldığında Süre verileri arasında maksimum değer 48.23 den büyük ve 71.7'dir. Bu yüzden maksimum değer 71.7 kabul edilir. Bulanık Mantık Üyelik fonksiyonunun oluşturulabilmesi için zayıf, orta, iyi sınırlarının belirlenmesi gereklidir. Bu nedenle üst sınırdan alt sınır çıkarılır ve 4'e bölünerek aralık değerleri bulunur.
 $x=(71.7-0)/4=17.92$

$$\text{Zayıf aralığı (N)}=0+17.92=17.92$$

$$\text{Orta aralığı (P)}=17.92+17.92=35.85$$

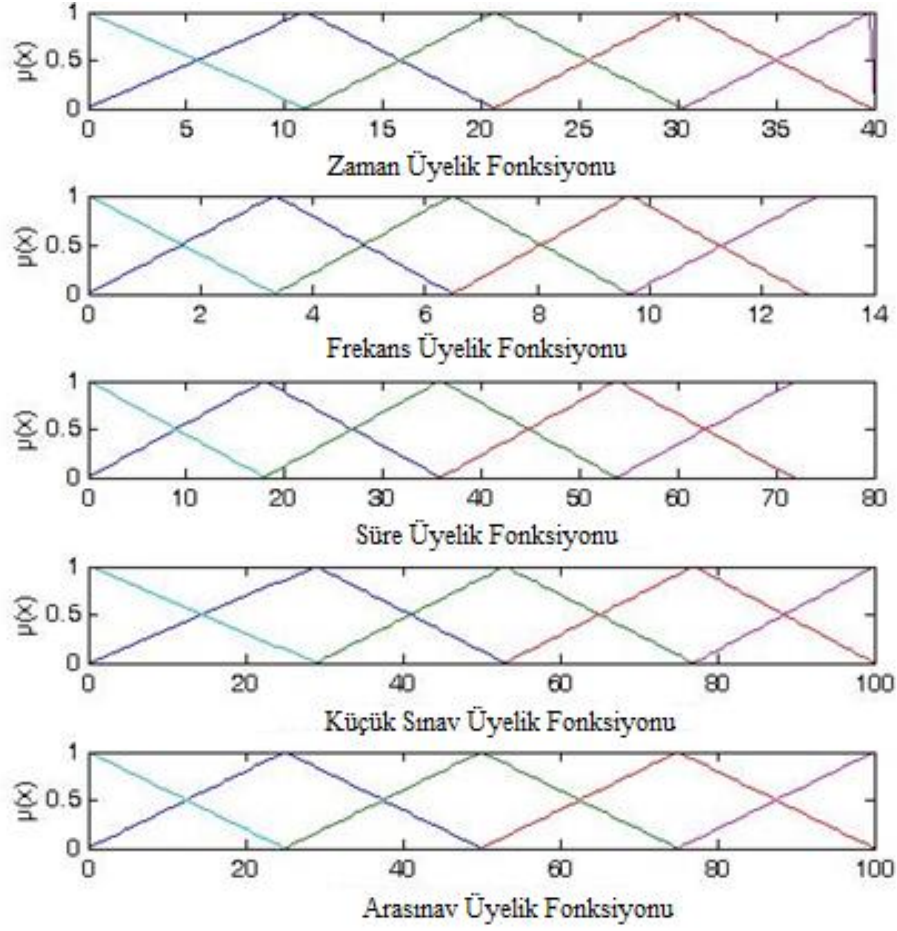
$$\text{İyi aralığı (R)}=35.85+17.92=53.77$$

olarak hesaplanır. Hesaplanan verilerle süre'ye ait üyelik fonksiyonları şekil 4.7'de oluşturulmuştur.



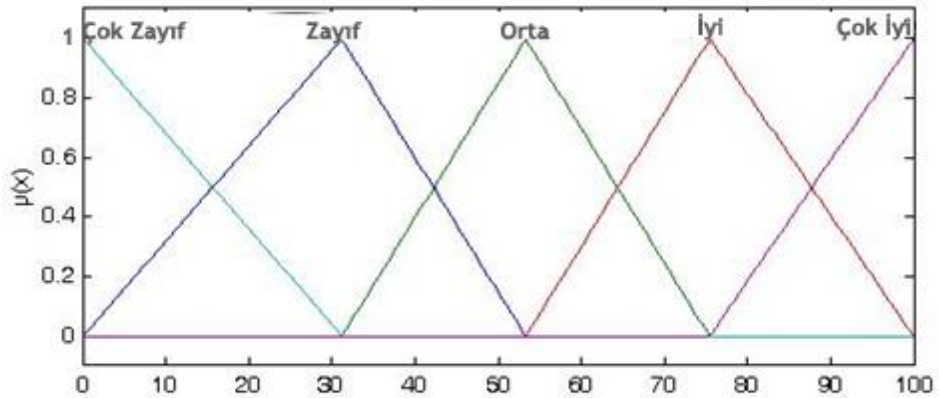
Şekil 4.7: Süre üyelik fonksiyonu aralık değerleri.

Bu işlem diğer veriler için de uygulandığında üyelik fonksiyonlarının grafikleri şekil 4.8'de görüldüğü gibi elde edilmiştir.



Şekil 4.8: Girdi değişkenlerine ait üyelik fonksiyonları.

Çıkış değişkenine ait üyelik fonksiyonu şekil 4.9'da verilmiştir.



Şekil 4.9: Çıkış Değişkenine ait üyelik fonksiyonu.

Girdiler sisteme giriş olarak verilir ve tahminler yapıldığında doğruluk oranı ortalama mutlak yüzde hata (mean absolute percentage error-MAPE) metodu ile bulunmuştur.

OMYH'ya ait eşitlik 4.1'de bulunmaktadır. Hesaplama sonucunda doğruluk oranı yaklaşık %72 bulunmuştur.

$$OMYH = \frac{\sum_{i=1}^n \left| \frac{y_i - f_i}{y_i} \right|}{n} \quad (4.1)$$

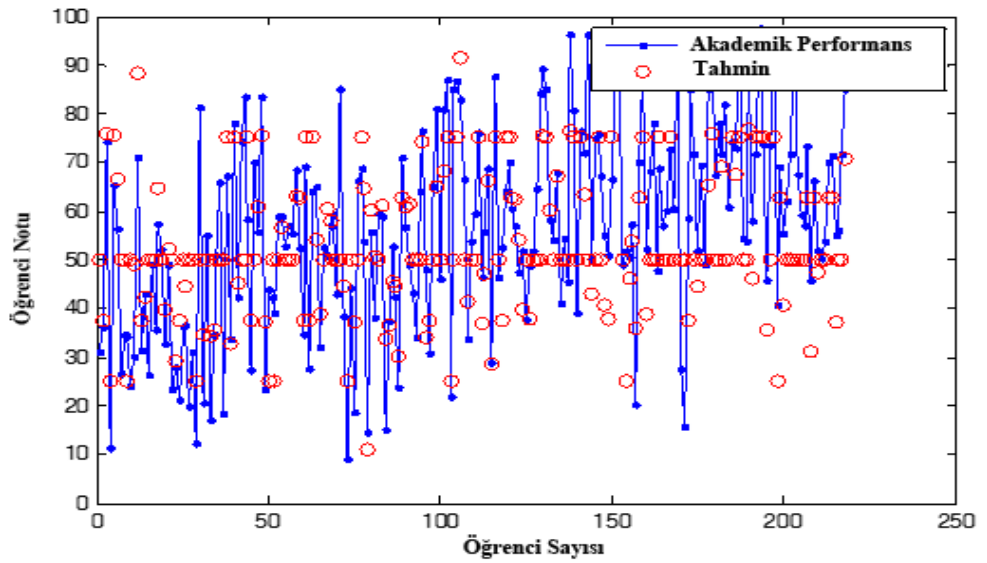
burada;

y_i : gerçek değer

f_i : tahmin edilen değer

n : eleman sayısı

Şekil 4.10'da bulunan sonuçlara ait karşılaştırma grafiği bulunmaktadır.



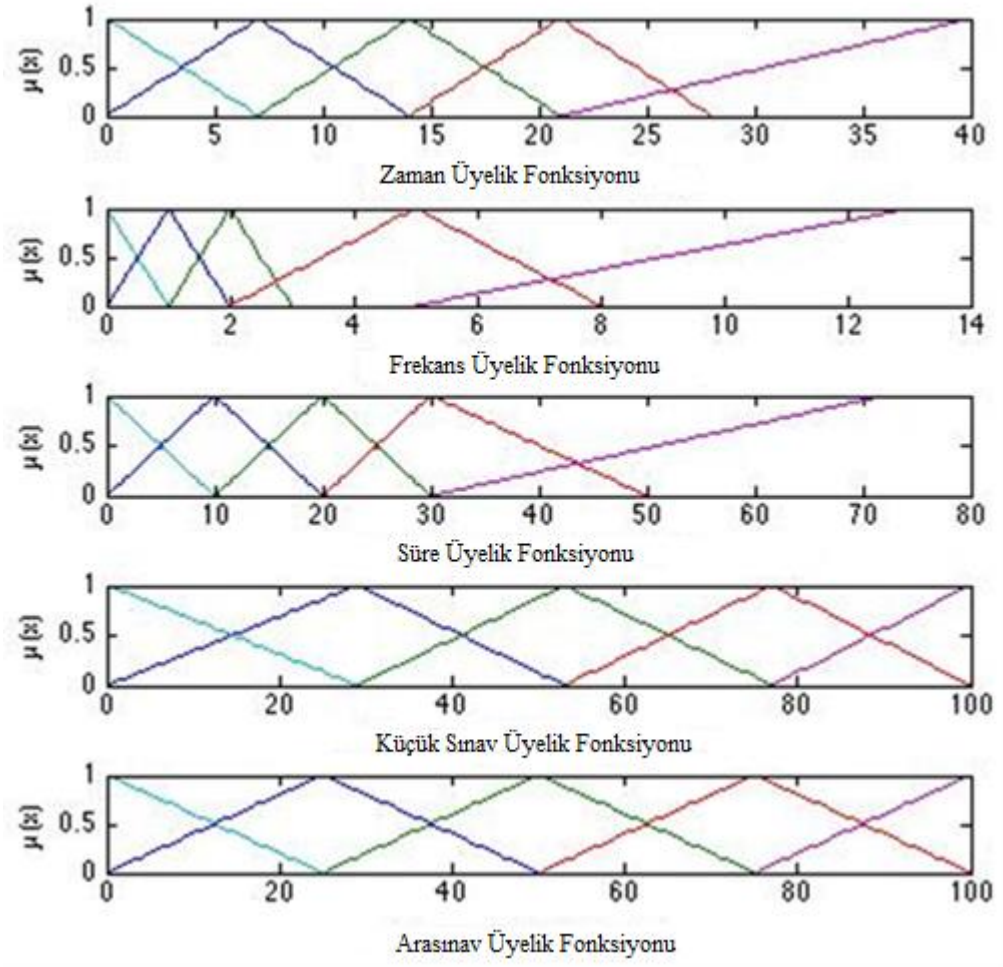
Şekil 4.10: Akademik performans ve tahmin değerleri karşılaştırma grafiği.

4.3.2 Uzman Görüşüne Dayalı Bulanık Mantık Modeli

Bulanık Mantık modeli kurulurken en önemli kriterler üyelik fonksiyonları ve kuralların oluşturulma sürecidir. Üyelik fonksiyonlarının seçimi deneme yanılma yoluyla yapılır ki bu uzun bir zaman alabilir. Ayrıca üyelik fonksiyonunun seçiminden sonra girdilere ait aralıkların belirlenmesi en önemli adımlardan biridir. İşte bu aşamada deneyimler oldukça önemlidir. Uzman görüşüne dayanan bulanık mantık modellemesinde sonuçlar

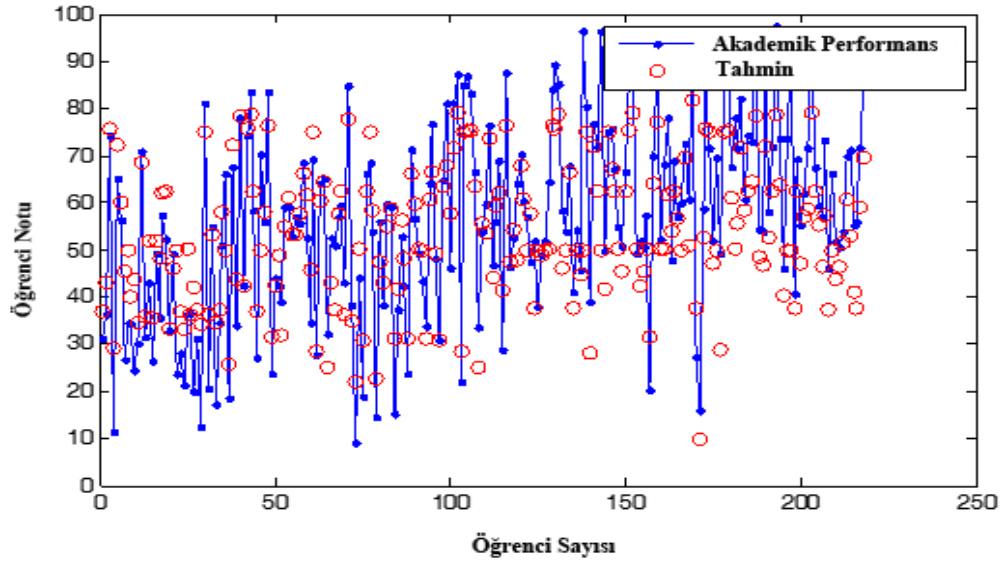
daha doğru bir şekilde tahmin edilebilir. Bu çalışmada klasik yolla kurulan bulanık mantık modelinden sonra deneyimlere dayanan bir model kurulmuştur. Örneğin, öğrencinin dersin sisteme yüklenmesinden kaç gün sonra sisteme girdiği kaydı tutan zaman değerinin sınırları [0,42] dir. Bunun anlamı, öğrenciler ders sisteme yüklendikten itibaren 0 ile 42 gün arasında sisteme girmiş ve konuya bakmışlardır. Klasik bulanık mantık da “iyi”,”orta”,”zayıf” aralıkları belirlenirken, maximum ile minimum arasındaki fark 4’e bölünmek sureti ile aralıklar belirlenmiştir. Bu durumda “iyi” aralığının sınırı [0 11 20] olarak belirlenmiştir. Bu aralıklar okunduğunda bir öğrenci ders sisteme yüklendikten 0-20 gün arasında sisteme giriş yaparsa, bu öğrencinin “iyi” sınıfındaki üyelik derecesi 0-1 arasında değişecektir. Bir öğrenci en geç 7 gün içinde dersi tekrar ederse dersi daha iyi anladığı kabul edilerek, deneyime dayanan modelde “iyi” sınıfının aralığı [0 7 14] olarak değiştirilmiştir. Bu yöntemle zaman üyelik fonksiyonu ile birlikte, frekans ve süre üyelik fonksiyonlarında da değişiklikler yapılarak model kurulmuş ve sonuçlar alınmıştır. Bulunan sonuçlara bakıldığında çalıştırılan model sonucu alınan doğruluk oranlarının ortalaması %78.62 bulunmuştur.

Şekil 4.11’de uzman görüşüne dayanan modele ait üyelik fonksiyonları bulunmaktadır.



Şekil 4.11: Girdi değişkenlerine ait üyelik fonksiyonları.

Şekilde uzman görüşüne dayanarak yapılan verilerle bulunan sonuçlara ait karşılaştırma şekil 4.12'de gösterilmiştir.

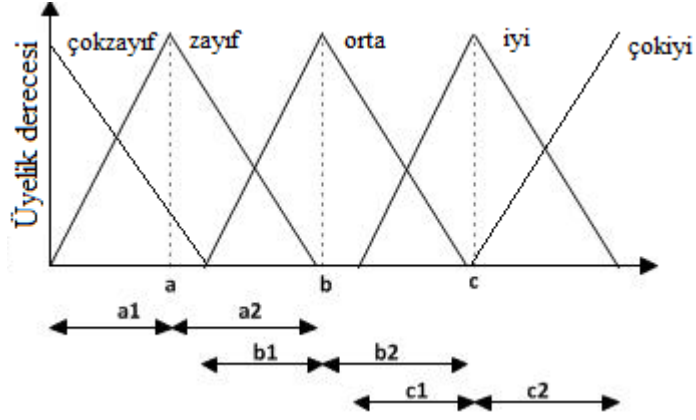


Şekil 4.12: Akademik performans ve tahmin değerleri karşılaştırma grafiği.

4.3.3 Gene-Fuzzy Modeli

Üyelik fonksiyonlarının aralıklarının belirlenmesinde deneme yanılma yöntemi ile uzman görüşüne dayanan yöntemde sonuçların doğruluk yüzdesini artırmak oldukça zordur. Bu nedenle aralıkların optimizasyonu, temeli doğal seleksiyon ve evrime dayanan en iyileme metotlarından biri olan genetik algoritma kullanılarak yapılmıştır. Burada ki amaç rastgele üretilen sayılarla oluşan popülasyonda en iyi sonucu veren aralıkların korunarak daha iyi aralıklar üretmek ve en iyi sonucu veren aralığa ulaşmaktır. Kısaca genetik algoritma adımları aşağıda verilmiştir.

1. Rastgele toplam 8 adet, Zaman, Frekans ve Süre girdilerinden her birine karşılık 3 adet 0-128 aralığında sayı üret. Bu sayılar şekil 4.13’de görüldüğü üzere her bir girdi değeri için üretilecektir. Örneğin “zayıf” üçgenine ait a_1 , a_2 ve a_3 değerleri üretilecektir. Bu kısımda küçük sınav ve arasınay üyelik fonksiyonları optimizasyon işlemine sabit aralıklarla dahil edilmiştir.



Şekil 4.13: Üyelik fonksiyonu aralık değerleri.

2. Daha önceki klasik bulanık mantık ve uzman görüşüne dayalı bulanık mantık modellerine ait girdi aralık değerlerini oluşturulan popülasyonun içine ekle. Toplam 10x45 lik matrisden oluşan başlangıç popülasyon matrisi her satırı 5 adet girdinin üyelik fonksiyon aralıklarını temsil etmektedir. Şekil 4.14'de bir kromozoma ait genler görülmektedir.

$$\text{kromozom} = [1110011010001 \dots 1001011]$$

gen1 gen2 gen45

Şekil 4.14 Genetik algoritma kromozom yapısı.

3. Başlangıç popülasyon matrisindeki değerler bulanık modele uygulanarak çalıştırılır ve uygunluk fonksiyonuna göre en az hata değerini veren değerler sıralanır. Bu çalışmada uygunluk fonksiyonu hata oranı fonksiyonudur.

4. %50 seçim oranı kullanılarak en iyi sonucu veren 5 popülasyon üyesi hayatta kalır [63].

5. Ağırlıklı rulet seçimi yöntemi, bireylerin performanslarının ağırlıklı ortalamalarına göre seçilmelerini esas alır. 5 popülasyon üyesinin ortalama mutlak hata oranları bulunur.

$$OMYH = \frac{\sum_{i=1}^n \frac{|y_i - f_i|}{y_i}}{n} \quad (4.2)$$

5 adet aralığa ait hata oranları 100 üzerinden oranlanmış ve çıkan sonuçlar aşağıdaki tablo 4.4'te verilmiştir.

Tablo 4.4: Eşleştirme hata oranları.

No	Hata Oranı	Yaklaşık Popülasyon Sayısı
1	0,177622	19,85411
2	0,177116	19,91076
3	0,17663	19,96559
4	0,175561	20,0871
5	0,174732	20,18244
Toplam	28,35658	100

Bu sonuçlara göre %17.76 hata oranına sahip 1. popülasyon 100 popülasyondan oluşan bir kümede 20 adet yere sahip olacak, aynı şekilde %17.47 hata oranına sahip en iyi hata sonucunu veren popülasyon üyesi de 20 adet yere sahip olacaktır. Bu şekilde yeni popülasyon havuzu oluşturulur.

6. Popülasyon havuzundan toplam 9 adet yeni üye seçilir. En iyi değeri veren popülasyon üyesi aynı şekilde kopyalanarak en iyi aralık korunur.

7. İkilik Sayı düzeninde kodlama: 10'luk düzendeki sayılar 2lik düzene çevrilerek kromozomlar oluşturulur.

8. Eşleştirme Aşaması: 1 ile 5 arasındaki sayılar ile bir dizi oluşturulur. 6-10 arasındaki sayıların rastgele sıralanması sonucu oluşan yeni dizi ile eşleştirilir.

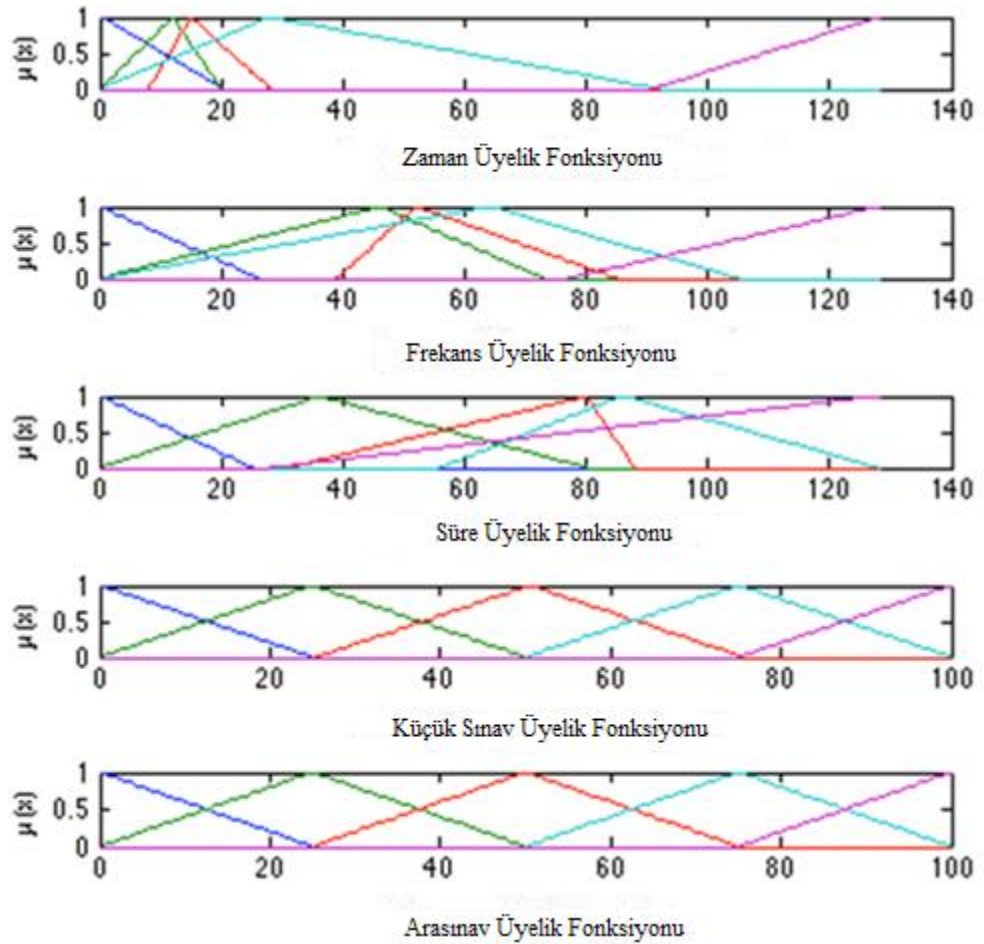
9. Çaprazlama Aşaması: Eşleştirilen iki birey örneğin 0110010 ve 0111001 olsun. Bu noktada 1-7 arasında rastgele bir sayı üretilir. Bu çaprazlamanın ikilik katarın hangi basamağında olacağını belirtir. Bu sayının 5 olduğunu varsayalım. Birinci bireyin ilk beş basamağı ile ikinci bireyin son 2 basamağı, ikinci bireyin ilk beş basamağı ile

birinci bireyin son 2 basamağı birleştirilerek yeni bireyler oluşturulur. Bu oluşumun sonunda birinci birey 0110001, ikinci birey 0111010 olur.

10. Mutasyon Aşaması: Mutasyonun amacı, popülasyondaki tüm bireylerin aynı sonucu verecek hale gelmesinin önüne geçmektir. Mutasyon işlemi çaprazlama sonucu oluşan yeni bireye ait genlerden rastgele seçilerek değiştirilme işlemidir.

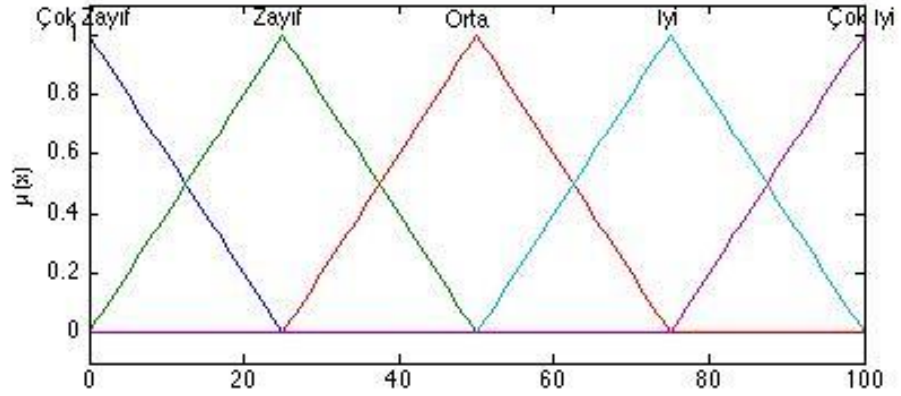
11. Tekrarlama Aşaması: 4. Adımdan itibaren en iyi sonuç elde edilinceye kadar iterasyon yapılır.

Şekil 4.15'te iterasyon sonucu aralıkların aldığı son duruma ait grafikler görülmektedir.



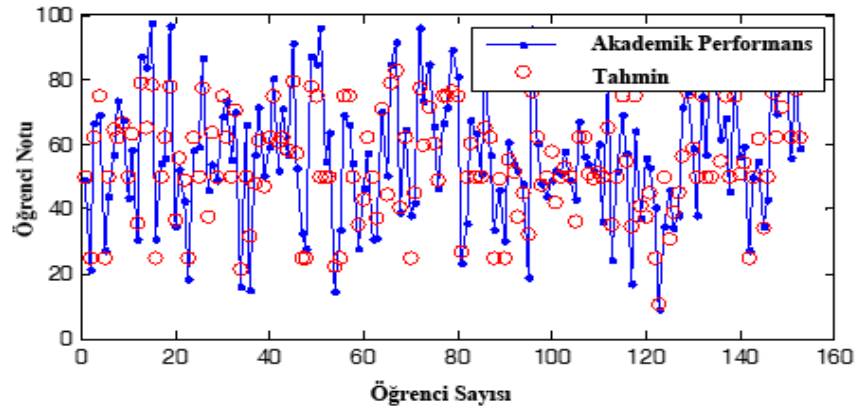
Şekil 4.15: Gene-Bulanık mantık yöntemi ile elde edilen üyelik fonksiyonları.

Şekil 4.16'da çıkış (akademik performans) değerlerine ait üyelik fonksiyonu grafiği bulunmaktadır.



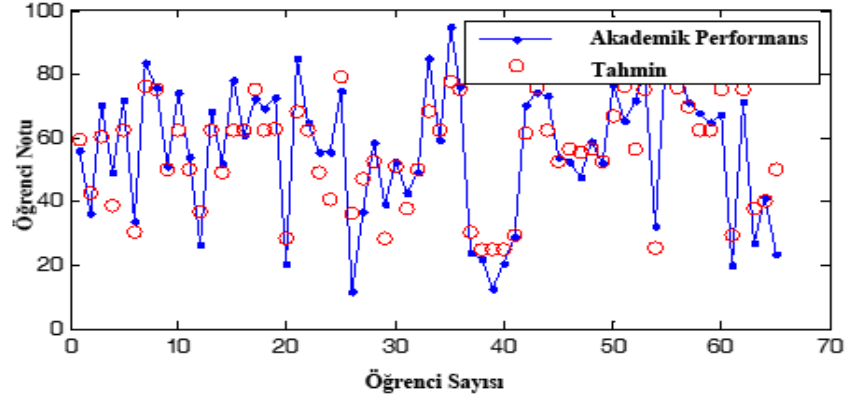
Şekil 4.16: Akademik performans üyelik fonksiyonu.

Şekil 4.17’de %70’lik rastgele seçilen optimizasyon verileriyle bulunan sonuçlara ait karşılaştırma grafiği bulunmaktadır. Yapılan çalışma sonucunda 1000 kez iterasyon yapılmış ve iterasyon sonucunda %70 rastgele seçilen verilerle elde edilen doğruluk oranı %82.5 bulunmuştur.



Şekil 4.17: Eğitim verilerine ait akademik performans ile tahmin değerleri karşılaştırma grafiği.

Şekil 4.18’de kalan %30’luk verilerle bulunan sonuçlara ait karşılaştırma grafiği bulunmaktadır. %30’luk test verileri ile elde edilen doğruluk oranı %81.11 olarak bulunmuştur.



Şekil 4.18: Test verilerine ait akademik performans ile tahmin değerleri karşılaştırma grafiği. Bu çalışmada uygulanan 3 metota ait doğruluk değerleri tablo da verilmiştir.

Rastgele seçilen %70 veri seti ile yapılan tahminlere ait doğruluk oranları tablo 4.5'te verilmiştir.

Tablo 4.5: Metotlara ait doğruluk (accuracy) oranları.

Model	Doğruluk Oranı (%)
Klasik Bulanık Mantık	72
Uzman Bulanık Mantık	78
Gene-Bulanık Mantık	82,50

Geriye kalan %30'luk veri seti ile aynı model üzerinde yapılan doğruluk değerleri tablo 4.6'da verilmiştir.

Tablo 4.6: Gene-Bulanık test verileri doğruluk değeri.

Model	Doğruluk Oranı (%)
Gene-Bulanık Mantık	81

4.4 UYGULAMA 3: TAKAGI-SUGENO TİPİ BULANIK MANTIK İLE UZAKTAN EĞİTİM ÖĞRENCİLERİNİN AKADEMİK PERFORMANSLARINI TAHMİN ETME

Bu tez çalışmasında Takagi-Sugeno tipi bulanık mantık metodunun kullanılmasının en önemli sebebi, bulanık mantık modelinde aralıkların ayarlanması, kuralların oluşturulmasında yoruma dayalı olmaktan kurtarıp tamamen modelin kendisinin bu bileşenleri ayarlamasıdır. Kümeleme analizi ile veri seti benzer özellik taşıyan elemanlara sahip kümelere ayrılır. Üyelik fonksiyonu ve kuralların model tarafından oluşturulmasıyla sisteme girdiler yüklenerek çıktılar alınır.

Bu uygulama da parametre tahmininde iki yöntem kullanılmıştır; birincisi en küçük kareler yöntemi ve ikincisi kantil regresyon yöntemidir. En küçük kareler yöntemi aykırı veri setinden etkilendiği için kümeleme analizi öncesinde kullanılacak olan verilere önce aykırı veri analizi yapılmalı, sonra da normal dağılıma sahip olup olmadığının testi yapılmalıdır. Çünkü bir çok kümeleme analizinde veriler ham haliyle direk kullanılmazlar. Verilerin bazı dönüştürücü araçları kullanılarak normalize ya da standardize edilmeleri gerekmektedir [65]. Kantil yönteminde ise aykırı veriden etkilenme oranı çok daha az olduğu için aykırı veri analizi yapılmadan tüm veri ile sonuçlar bulunmuştur.

4.4.1 Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin Kantil Yöntemine Göre Tahmin Edilmesi

4.4.1.1 Aykırı Veri Analizi

Aykırı veri analizi bölüm 3'de ayrıntıları verilen 3 metot ile yapılmıştır. Bunlar Mahalanobis Uzaklığı, Cook Uzaklığı ve Kaldıraç (Leverage) Noktası metotlarıdır. Bu metotların uygulanmasında SPSS 17 paket programı kullanılmıştır. Aykırı veri analizi öncelikle 2011-2012 öğretim yılı kapsamında 218 öğrencinin eğitim ve test verisi kullanılarak yapılmıştır. Daha sonra da, 2012-2013 yılında aynı dersi alan 95 öğrenci verisi için yapılmıştır.

4.4.1.1.1 Mahalanobis Uzaklığı

Mahalanobis Uzaklıkları, bir noktanın verinin merkezinden (aritmetik ortalama) ne kadar uzak olduğunu skaler ölçüsünü verir. Mahalanobis uzaklığını kullanarak, verinin merkezinden uzakta olan noktaları bulabilir ve bu noktaları aykırı gözlem olarak nitelendirebiliriz. [79]. İlk aşamada eğitim ve test verisi olan 218 öğrenciye ait Mahalanobis uzaklığı SPSS paket programı yardımı ile bulunmuştur. Mahalanobis uzaklık değerine sahip olan veri tablo 4.7’de gösterilmiştir.

Tablo 4.7: Mahalanobis uzaklığı metoduna göre aykırı verinin değerleri.

Sıra No	Zaman	Frekans	Süre	Küçük sınav	Arasnav 1	AP	Mahalanobis Uzaklığı	P
90	13,50	4,70	21,40	55,00	35,00	46,40	0,41042	0

Mahalanobis uzaklık metodu doğrulama verisi içinde uygulanmıştır. Fakat bu uygulamanın sonucunda bu kritere uyan herhangi bir veri bulunamamıştır.

4.4.1.1.2 Cook Uzaklığı

Bir verinin aykırı veri olup olmadığına genellikle $4/n$ (n : gözlem sayısı) değeri ile hesaplanarak karar verilir. $4/n$ değerinden büyük değerlerin aykırı değer olduğu ifade edilir [73]. Kritik Cook değeri (KCD) $4/n$ ile hesaplanır.

$$KCD=4/218=0.0183$$

Eğitim ve Test verileri için Cook Uzaklığı hesaplandıktan sonra KCD değeri ile karşılaştırma yapılır. KCD değerinden büyük olan değerler aykırı veri olarak nitelendirilir. Tablo 4.8’de KCD değerine göre Eğitim ve Test verileri içindeki aykırı veriler bulunmaktadır.

Tablo 4.8: Cook Uzaklığı Metodu ile bulunan eğitim ve test verileri içerisinde ki aykırı verilere ait değerler.

Sıra No	Zaman	Frekans	Süre	Küçük sınav	Arasınav 1	AP	Cook Uzaklığı
194	11,20	5,80	34,60	100,00	95,00	74,40	0,02022
118	18,30	2,30	6,50	98,00	79,00	65,20	0,03048
105	15,30	9,30	40,60	44,00	38,00	56,00	0,02
162	7,00	5,80	33,30	33,00	50,00	23,60	0,03033
204	9,00	9,00	33,60	81,00	10,00	67,60	0,05508
86	12,50	0,80	1,70	15,00	45,00	52,00	0,09486
119	16,20	12,80	50,30	100,00	100,00	81,20	0,03233
207	2,00	3,00	50,00	75,00	95,00	85,00	0,02941
106	9,80	8,50	62,60	83,00	23,00	64,00	0,02433

Cook uzaklığı ile doğrulama verisi üzerinde de analizler yapılmıştır. Toplamda 95 veri olduğundan KCD değeri 0,042 olmuştur. Bu değerin üzerinde olan veriler tablo 4.9'da gösterilmiştir.

Tablo 4.9: Cook uzaklığı metodu ile bulunan doğrulama verileri içerisinde ki aykırı verilere ait değerler.

Sıra No	Zaman	Frekans	Süre	Küçük sınav	Arasınav 1	AP	Cook Uzaklığı
21	3,20	3,67	68,02	57,00	65,00	49,40	0,33736
88	10,20	5,00	55,00	81,00	28,00	69,80	0,08183
23	12,00	14,67	13,00	87,00	69,00	57,20	0,0699
26	30,00	1,00	7,63	35,00	5,00	8,00	0,06239
8	9,40	17,00	15,37	37,00	85,00	41,40	0,05393

4.4.1.1.3 Kaldıraç Nokta (Leverage Değeri)

Kaldıraç nokta değeri (KD) $2p/n$ 'den (p: değişken sayısı, n: gözlem sayısı) büyük olduğunda, Belsley'e (1980) göre yüksek kaldıraç (uç değer)'dir. Küçük örneklem için Vellman ve Welsch (1981) $3p/n$ (p: değişken sayısı, n: kişi sayısı) ölçütünü önerir [72].

$$KD=2p/n=2*5/218=0,045$$

Eğitim ve Test verileri için Leverage değeri hesaplandıktan sonra KD değeri ile karşılaştırma yapılır. KD değerinden büyük olan değerler aykırı veri olarak

nitelendirilir. Tablo 4.10'da KD değerine göre Eğitim ve Test verileri içindeki aykırı veriler bulunmaktadır.

Kaldıraç değeri ile doğrulama verisi üzerinde de analizler yapılmıştır. Toplamda 95 veri olduğundan KD değeri 0,1052 olmuştur. Bu değer üzerinde olan veriler tablo 4.11'da gösterilmiştir.

Tablo 4.10: Kaldıraç Nokta Metodu ile bulunan eğitim ve test verileri içerisinde ki aykırı verilere ait değerler.

Sıra No	Zaman	Frekans	Süre	Küçük sınav	Arasınav 1	AP	Kadıraç Değeri
119	16,20	12,80	50,30	100,00	100,00	81,20	0,06045
125	30,00	5,80	43,60	47,00	53,00	43,40	0,06245
210	14,30	8,30	66,70	85,00	62,00	64,90	0,07712
185	18,80	6,20	22,80	31,00	90,00	55,50	0,06629
207	2,00	3,00	50,00	75,00	95,00	85,00	0,14088
106	9,80	8,50	62,60	83,00	23,00	64,00	0,07775
102	6,30	9,00	71,70	30,00	28,00	27,50	0,11158
83	39,80	1,80	9,70	5,00	18,00	15,80	0,10585
55	30,50	12,80	66,70	100,00	100,00	96,40	0,08459

Tablo 4.11: Kaldıraç Nokta Metodu ile bulunan doğrulama verileri içerisinde ki aykırı verilere ait değerler.

Sıra No	Zaman	Frekans	Süre	Küçük sınav	Arasınav 1	AP	Kadıraç Değeri
42	38,00	1,00	2,92	5,00	20,00	4,00	0,18989
62	13,60	18,17	40,15	34,00	43,66	35,33	0,18066
8	9,40	17,00	15,37	37,00	85,00	41,40	0,15837
20	4,00	1,00	0,86	10,00	10,00	2,00	0,14738
67	34,80	1,50	9,24	18,00	0,00	6,60	0,14616
21	3,20	3,67	68,02	57,00	65,00	49,40	0,14362

Aykırı veri olarak Eğitim ve Test verilerine bakıldığında 106,119 ve 207 sıra numaralı veri, hem Cook uzaklığı yöntemi ile hem de Kaldıraç değeri yöntemi ortak olarak tespit edilmiştir. Bu çalışmada farklı 3 metot ile tespit edilen toplam 15 veri, veri kümesinden çıkarılmıştır. Aykırı veriler çıkarıldıktan sonra toplamda 218 olan Eğitim ve Test verisi 203 kalmıştır. Aynı şekilde doğrulama verileri de aykırı veriler çıkarıldıktan sonra 85 adet kalmıştır.

4.4.1.2 Verilerin Normalize Edilmesi

Kümeleme analizinde verilerin normal dağılıma sahip olması varsayılır. Bundan dolayı kümeleme analizine girecek olan veri seti için Shapiro Wilk testi yapılmıştır. Tablo 4.12’de Shapiro-Wilk testi sonuçları görülmektedir.

Tablo 4.12: Normallik testi.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Zaman	,059	143	,200 [*]	,989	143	,309
Frekans	,080	143	,026	,977	143	,015
Süre	,076	143	,043	,957	143	,000
Küçük sınav	,081	143	,022	,974	143	,009
Arasınava	,083	143	,017	,957	143	,000
sqrtSüre	,080	143	,025	,978	143	,020
SqrtArasınava1	,054	143	,200 [*]	,975	143	,009
AP	,062	143	,200 [*]	,982	143	,055

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Beş farklı bağımsız değişkenin normalliğini test etmek için öncelikle hipotezler kurulur.

Zaman için hipotezler;

H₀:%95 güvenle veriler normal dağılımlıdır.

H₁:%95 güvenle veriler normal dağılımlı değildir.

Frekans için hipotezler;

H₀:%95 güvenle veriler normal dağılımlıdır.

H₁:%95 güvenle veriler normal dağılımlı değildir.

Süre için hipotezler;

H₀:%95 güvenle veriler normal dağılımlıdır.

H₁:%95 güvenle veriler normal dağılımlı değildir.

Küçük sınav için hipotezler;

H_0 :%95 güvenle veriler normal dağılımlıdır.

H_1 :%95 güvenle veriler normal dağılımlı değildir.

Arasınav için hipotezler;

H_0 :%95 güvenle veriler normal dağılımlıdır.

H_1 :%95 güvenle veriler normal dağılımlı değildir.

Burada "Shapiro-Wilk" testine göre p değerinin 0.05' den büyük olduğundan Zaman, Frekans ve Küçük sınav için H_0 hipotezleri kabul edilir. Yani bu bağımsız değişkenler için "%95 güvenle veriler normal dağılımlıdır." denilebilir. Fakat Süre ve Arasınav için H_0 reddedilir ve "%95 güvenle veriler normal dağılımlıdır." denilemez. Normal dağılıma sahip olmayan veriler için bazı dönüşüm teknikleri vardır. Bu dönüşüm tekniklerinden karekök alma işlemi Süre ve Arasınav e uygulanıp aynı test yapıldığında p değerlerinin 0.05 den büyük olduğu görülmektedir ve bu değerlerle normal dağılıma sahip bir veri seti elde edilmiş olur.

Bu çalışmada 3 farklı kümeleme yöntemi ile uzaktan eğitim öğrencilerinin akademik performansı tahmin edilmeye çalışılmıştır. Bu yöntemler K-Ortalamalar, Bulanık C-Ortalamalar ve Çıkarımlı (Subtractive) kümeleme yöntemleridir

4.4.1.3 K-Ortalamalar Kümeleme Yöntemi

K-Ortalamalar kümeleme yönteminde küme sayısı kullanıcı tarafından girilmektedir. Bu nedenle modelde döngü yapılmış ve en uygun değer bulunduğu döngü durdurulmuştur. K-ortalamar kümeleme yöntemi ile eğitim verileri üzerinde yapılan analizler sonucunda elde edilen doğruluk oranları tablo 4.13'te verilmiştir.

Tablo 4.13: Eğitim verisi küme sayılarına göre doğruluk değerleri.

Eğitim Verisi Küme Sayılarına Göre Doğruluk Değerleri								
Küme Sayısı	2	3	4	5	6	7	8	9
Doğruluk Oranı	87,44	87,91	88,83	88,87	89,63	89,39	90,93	90,11

Tablo 4.13'te ki eğitim verisine göre en iyi doğruluk oranı küme sayısı 8 olan modele göre bulunmuştur. Küme sayısı ve doğruluk oranı tablo 4.13'de değerler gri dolgu ile

doldurulup belirgin hale getirilmiştir. Bu aşamadan sonra yapılacak olan işlemlerde küme sayısı 8'e göre yapılacaktır.

K-ortalamlar kümeleme yöntemi ile bulunan küme merkezleri tablo 4.14'de verilmiştir.

Tablo 4.14: K-ortalamlar kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük sınav	Arasnavı
1	20,51	3,82	3,76	51,64	5,91
2	14,92	5,16	4,40	80,31	6,98
3	14,41	4,22	3,98	42,94	4,25
4	10,67	5,31	5,01	67,65	6,91
5	21,15	4,48	4,15	65,25	5,67
6	13,00	2,13	2,61	27,89	4,27
7	8,97	3,77	3,59	51,91	5,64
8	12,80	5,34	4,93	95,71	8,87

Sugeno tipi bulanık mantık uygulamasında gauss eğrisi kullanılır. Gauss eğrisini oluşturan 2 parametre vardır [80]. Burada

c :eğrinin merkezi

σ :standart sapma

$$gauss(x; c; \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \quad (4.3)$$

Tablo 15'de yer alan küme merkezlerine göre gauss eğrileri oluşturulmuştur. Frekans girdisine ait 1 numaralı küme merkezinin gauss eğrisi aşağıda verilmiştir.

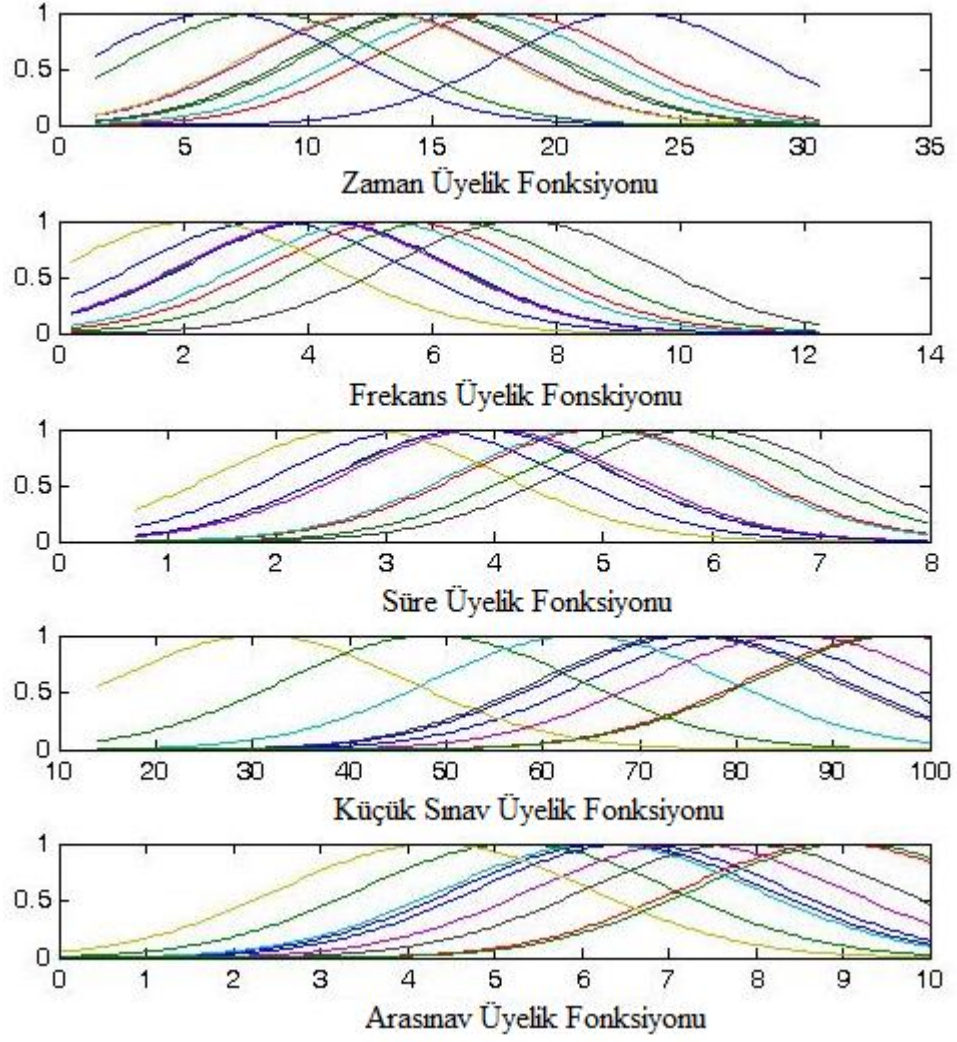
$$gauss(x; 20.51; 5.13) = e^{-\frac{1}{2}\left(\frac{x-20.51}{5.13}\right)^2}$$

K-Ortalamlar kümeleme metodu ile bulunan küme merkezleri kullanılarak, her bir değişkene ait gauss eğri denklemi bulunur. Tablo 4.15'te girdi değişkenlerine ait gauss eğri denklemleri bulunmaktadır.

Tablo 4.15: K-ortalamlar metodu ile bulunan deęişkenlere ait gauss eğri denklemleri.

No	Zaman	Frekans	Süre	Küçük sınav	Arasınav
μ_1	$e^{-\frac{1}{2}\left(\frac{x-20.51}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-3.82}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-3.76}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-51.64}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.91}{1.77}\right)^2}$
μ_2	$e^{-\frac{1}{2}\left(\frac{x-14.92}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.16}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.40}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-80.31}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.25}{1.77}\right)^2}$
μ_3	$e^{-\frac{1}{2}\left(\frac{x-14.41}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.22}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-3.98}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-42.94}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.25}{1.77}\right)^2}$
μ_4	$e^{-\frac{1}{2}\left(\frac{x-10.67}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.31}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.01}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-67.65}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-6.91}{1.77}\right)^2}$
μ_5	$e^{-\frac{1}{2}\left(\frac{x-21.15}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.48}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.15}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-65.25}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.67}{1.77}\right)^2}$
μ_6	$e^{-\frac{1}{2}\left(\frac{x-13.00}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-2.13}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-2.61}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-27.89}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.27}{1.77}\right)^2}$
μ_7	$e^{-\frac{1}{2}\left(\frac{x-8.97}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-3.77}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-3.59}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-51.91}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.64}{1.77}\right)^2}$
μ_8	$e^{-\frac{1}{2}\left(\frac{x-12.80}{5.13}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-5.34}{2.12}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-4.93}{1.29}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-95.71}{15.20}\right)^2}$	$e^{-\frac{1}{2}\left(\frac{x-8.87}{1.77}\right)^2}$

Üyelik fonksiyonlarının grafikleri şekil 4.19’da verilmiştir.



Şekil 4.19: K-Ortalamalar kümeleme yöntemi üyelik fonksiyonları grafikleri.

Üyelik fonksiyonları bulunduktan sonra sugeno tipi bulanık mantık da kurallara ait fonksiyonların bulunması gerekmektedir. Sugeno tipi bulanık mantık da kurallar “eğer-öyleyse” şeklindedir. Eğer kısmında girdi değişkenleri sisteme giriş yaparken öyleyse kısmında kuralların olduğu bir sistem bulunmaktadır. Her kurala ait sonuçlar daha sonra formül 4.3 ile sonuçlandırılır.

Eğer $x_1 \in \text{Zaman}$ ve $x_2 \in \text{Frekans}$ ve $x_3 \in \text{Süre}$ ve $x_4 \in \text{Küçük Sınav}$ ve $x_5 \in \text{Arasnav}$

$$\text{öyleyse } Z \rightarrow w^k = \beta_0^k + \beta_1^k x_1 + \beta_2^k x_2 + \beta_3^k x_3 + \beta_4^k x_4 + \beta_5^k x_5 \quad (4.4)$$

$$Z = \frac{\sum_{k=1}^5 \mu_k W_k}{\sum_{k=1}^5 \mu_k} \quad (4.5)$$

Kurallara ait parametreler en küçük kareler yöntemi kullanılarak bulunur. Tablo 4.16'da kurallara ait parametreler görülmektedir.

Tablo 4.16: Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	145,78	-0,70	-18,89	7,80	-3,10	6,17
2	23,29	-1,47	-1,02	-3,02	0,26	3,02
3	-52,41	3,87	4,51	0,82	-0,33	1,45
4	-1,26	0,50	-1,10	-3,66	0,98	4,27
5	292,38	-2,17	9,26	-31,74	-1,09	11,43
6	-5,03	-0,86	2,51	-0,54	0,44	4,98
7	-13,01	4,60	-1,47	-7,82	0,54	4,16
8	58,99	1,83	-0,42	-0,61	-0,36	5,62

Parametreler bulunduktan sonra kurallara ait denklemler;

$$w^1 = 145,78 - 0,70x_1 - 18,89x_2 + 7,80x_3 - 3,10x_4 + 6,17x_5$$

$$w^2 = 23,29 - 1,47x_1 - 1,02x_2 - 3,02x_3 + 0,26x_4 + 3,02x_5$$

$$w^3 = -52,41 + 3,87x_1 + 4,51x_2 + 0,82x_3 - 0,33x_4 + 1,45x_5$$

$$w^4 = -1,26 + 0,50x_1 - 1,10x_2 - 3,66x_3 + 0,98x_4 + 4,27x_5$$

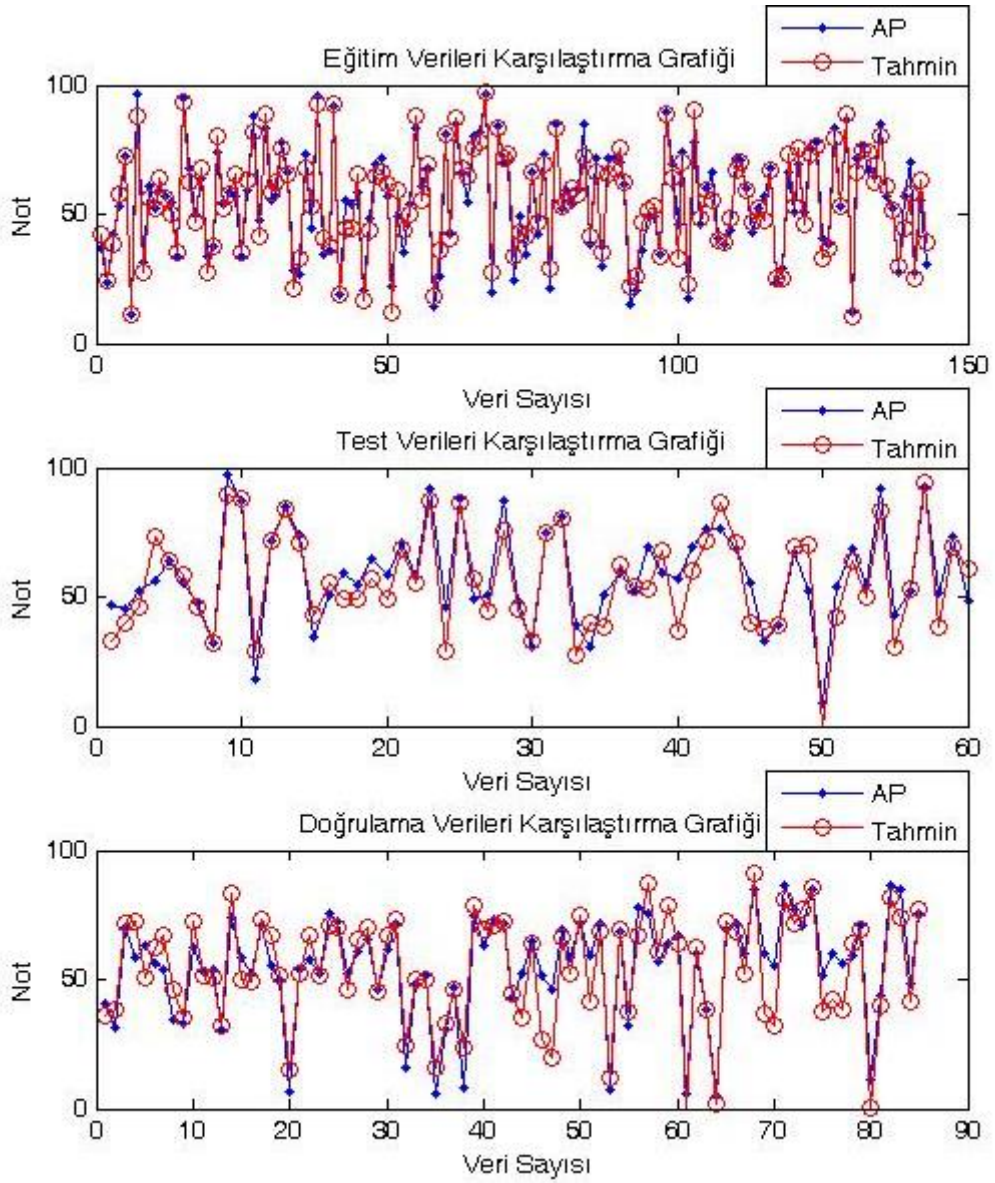
$$w^5 = 292,38 - 2,17x_1 + 9,26x_2 - 31,74x_3 - 0,55x_4 + 8,02x_5$$

$$w^6 = -5,03 - 0,86x_1 + 2,51x_2 - 0,54x_3 + 0,44x_4 + 4,98x_5$$

$$w^7 = -13,01 + 4,60x_1 - 1,47x_2 - 7,82x_3 + 0,54x_4 + 4,16x_5$$

$$w^8 = 58,99 + 1,83x_1 - 0,42x_2 - 0,61x_3 - 0,36x_4 + 5,62x_5$$

Şekil 4.20'de eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.



Şekil 4.20: K-ortalamalar kümeleme yöntemi ile kurulan bulanık model e ait veri tipi bazında akademik performans ve tahmin karşılaştırma grafikleri.

K-ortalamalar kümeleme metoduna göre bulunan, küme merkezleri baz alınarak oluşturulan modele ait doğruluk değeri %90.93 olarak bulunmuştur.

Eğitim verisini test etmek için kullanılan test verisi girdi olarak modele uygulandığında %84.00 doğruluk oranı elde edilmiştir. Doğrulama verisi modele girdi olarak uygulandığında sonuç %80.21 bulunmuştur. Tablo 4.17’de 3 veri tipine ait veri sayısı görülmektedir. Tablo 4.18’de ise 3 veri tipine göre doğruluk oranları verilmiştir.

Tablo 4.17: Verilere ait adet bilgisi.

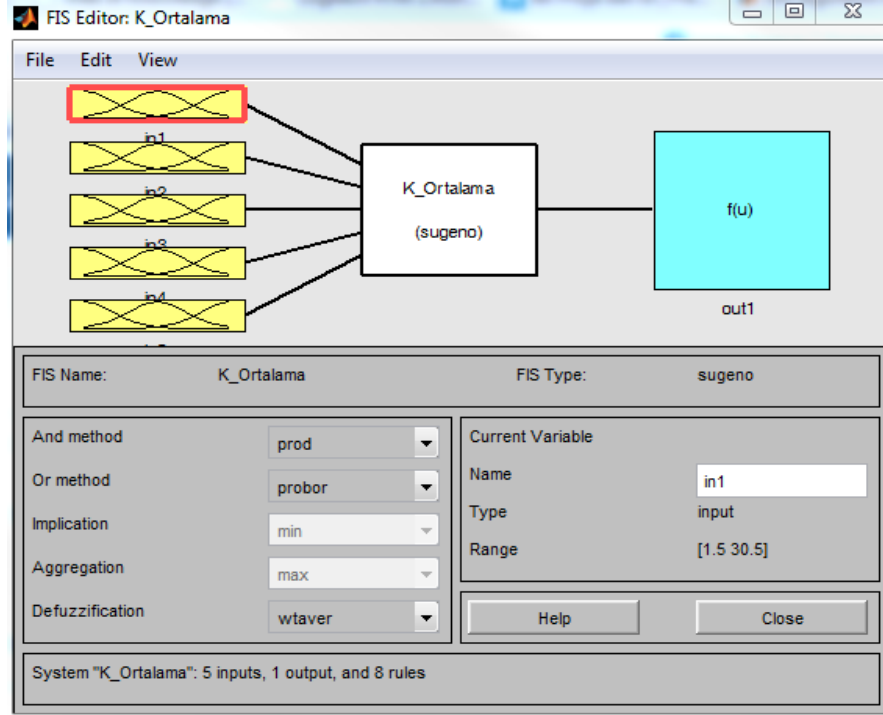
	Eđitim	Test	Dođrulama	Toplam
Veri Sayısı	143	60	85	288

Tablo 4.18: Verilere ait adet dođruluk oranları.

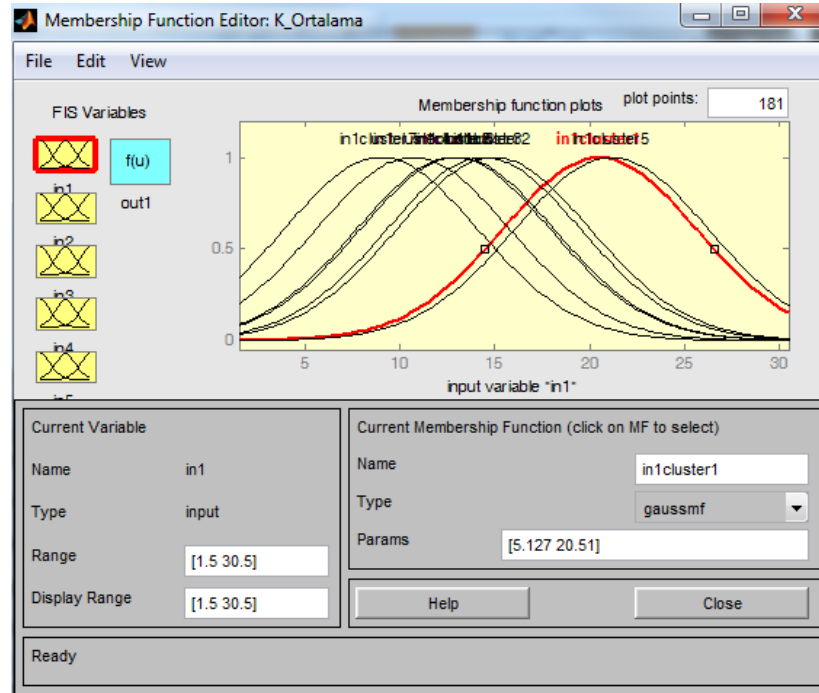
Küme Sayısı=8	Eđitim Verisi	Test Verisi	Dođrulama Verisi	Ortalama
Dođruluk Oranı	90,93	84,00	80,21	85,05

Aykırı veri analizi sonrasında, aykırı veriler sistemden çıkarılmadığında yapılan analizlerde eğitim verilerinde başarı oranı %90,03 çıkarken dođrulama verilerinde bu oran %70,91'e düşmüştür. Bu sonuçlarda aykırı verilerin çıkarılmasıyla modelde anlamlı derecede farklılık meydana geldiđini göstermektedir.

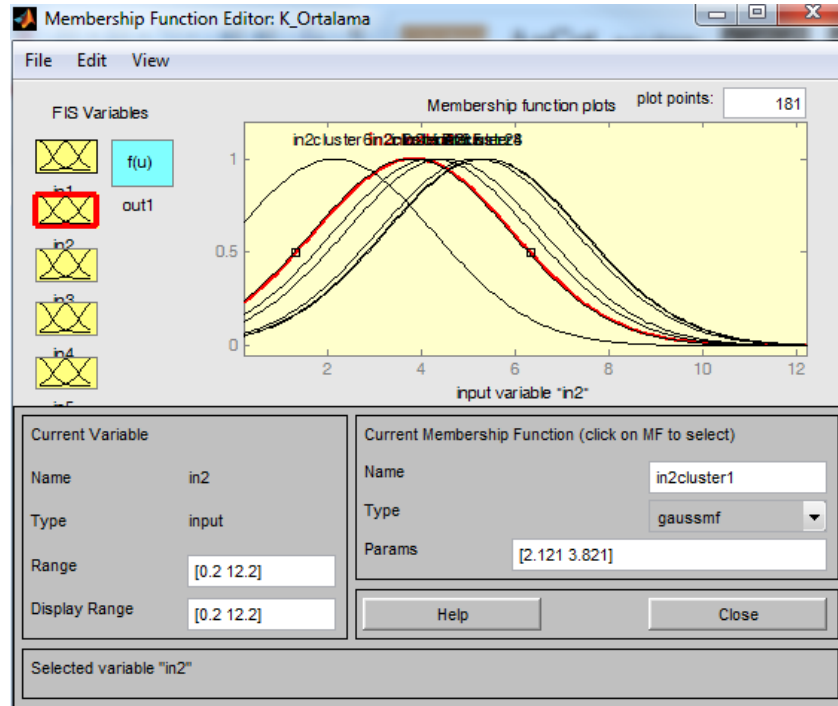
Elde edilen tüm sonuçlar için kurulan model Matlab programı 2011 versiyonu ile elde edilmiştir. Kodlama yapılarak kurulan bulanık mantık modeli aynı zamanda Matlab'de hazır bulanık mantık model kurma aracı olan fis editör dosya formatı şeklinde de kaydedilmiştir. Kayıt işleminden sonra Matlab ortamında açılan dosyanın ekran görüntüleri aşağıda verilmiştir.



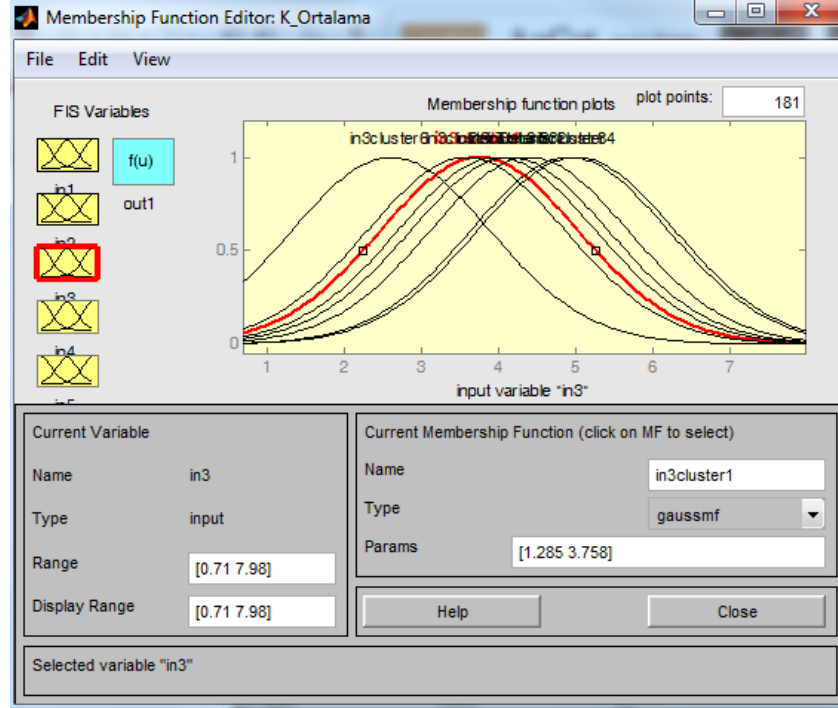
Şekil 4.21: K-ortalamlar kümeleme yöntemi Matlab ekran görüntüsü.



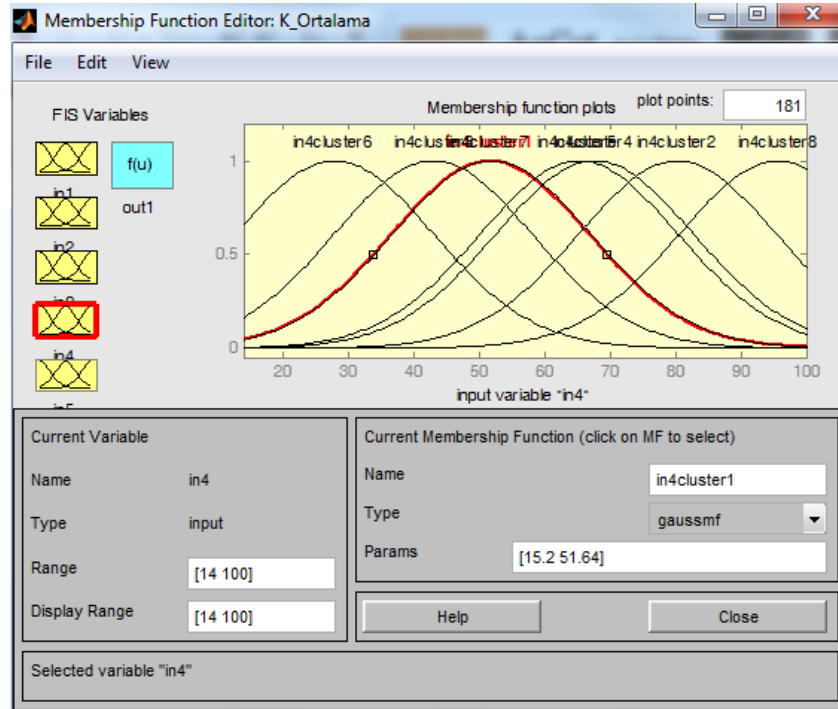
Şekil 4.22: K-ortalamlar kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



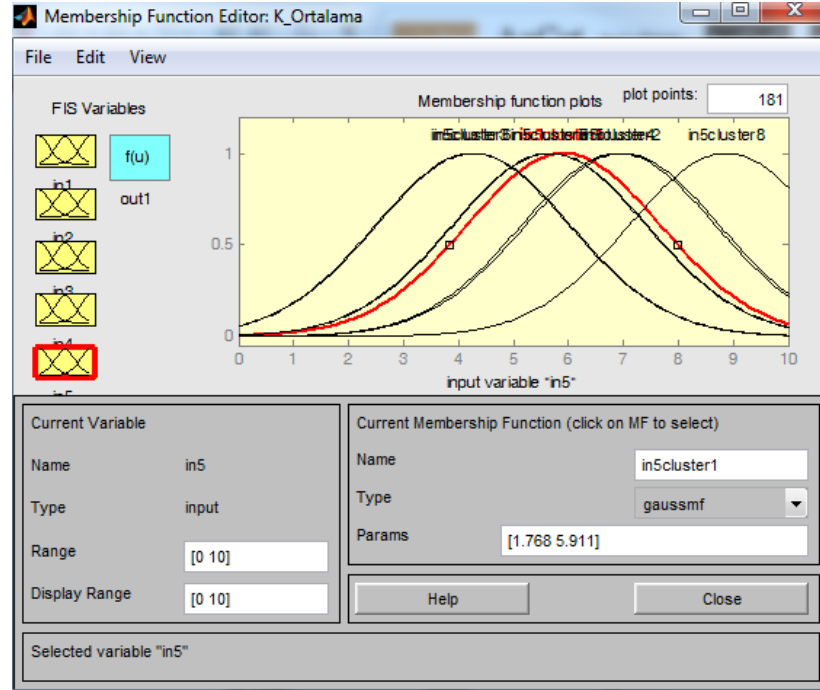
Şekil 4.23: K-ortalamlar kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



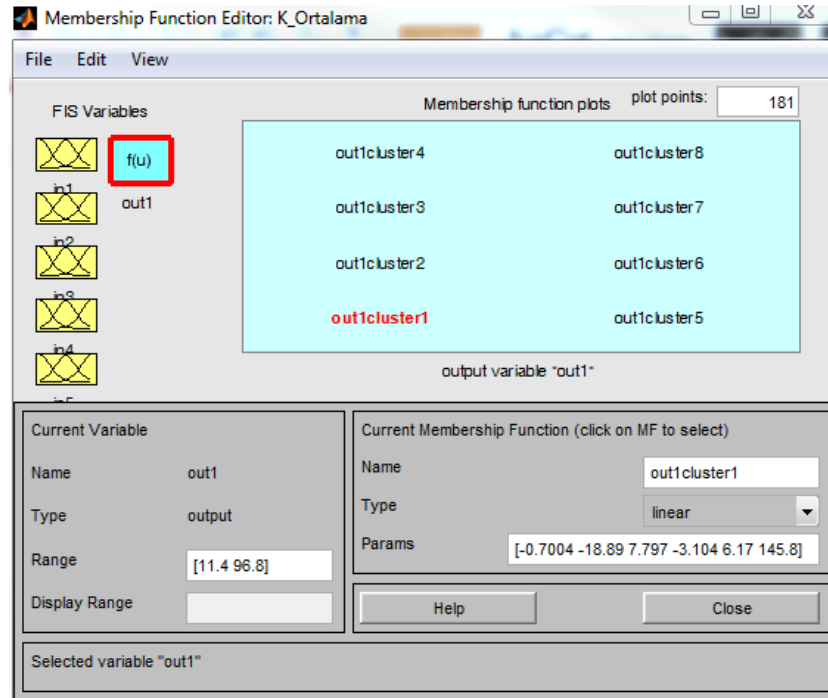
Şekil 4.24: K-ortalamalar kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.25: K-ortalamalar kümeleme yöntemi küçük sınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.26: K-Ortalamlar kümeleme yöntemi arasınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.27: K-Ortalamlar kümeleme yöntemi akademik performans çıktısı kurallara ait Matlab ekran görüntüsü.

4.4.1.4 Bulanık C-Ortalamalar Kümeleme Yöntemi

Bulanık C-Ortalamalar Kümeleme metoduna göre bulunan küme merkezleri baz alınarak oluşturulan modele ait doğruluk değerleri Tablo 4.19’da verilmiştir.

Bulanık C-Ortalamalar kümeleme yönteminde küme sayısı, K-ortalamlar kümeleme yönteminde olduğu gibi kullanıcı tarafından girilmektedir. Bu nedenle modelde döngü yapılmış ve en uygun değer bulunduğu anda döngü durdurulmuştur. Bulanık C-ortalamlar kümeleme yöntemi ile eğitim verileri üzerinde yapılan analizler sonucunda elde edilen doğruluk oranları tablo 4.19’da verilmiştir.

Tablo 4.19: Eğitim verisi küme sayılarına göre doğruluk değerleri.

Eğitim Verisi Küme Sayılarına Göre Doğruluk Değerleri									
Küme Sayısı	2	3	4	5	6	7	8	9	10
Doğruluk	87,44	87,97	88,72	88,67	88,76	89,52	89,90	91,09	90,76
Oranı (%)									

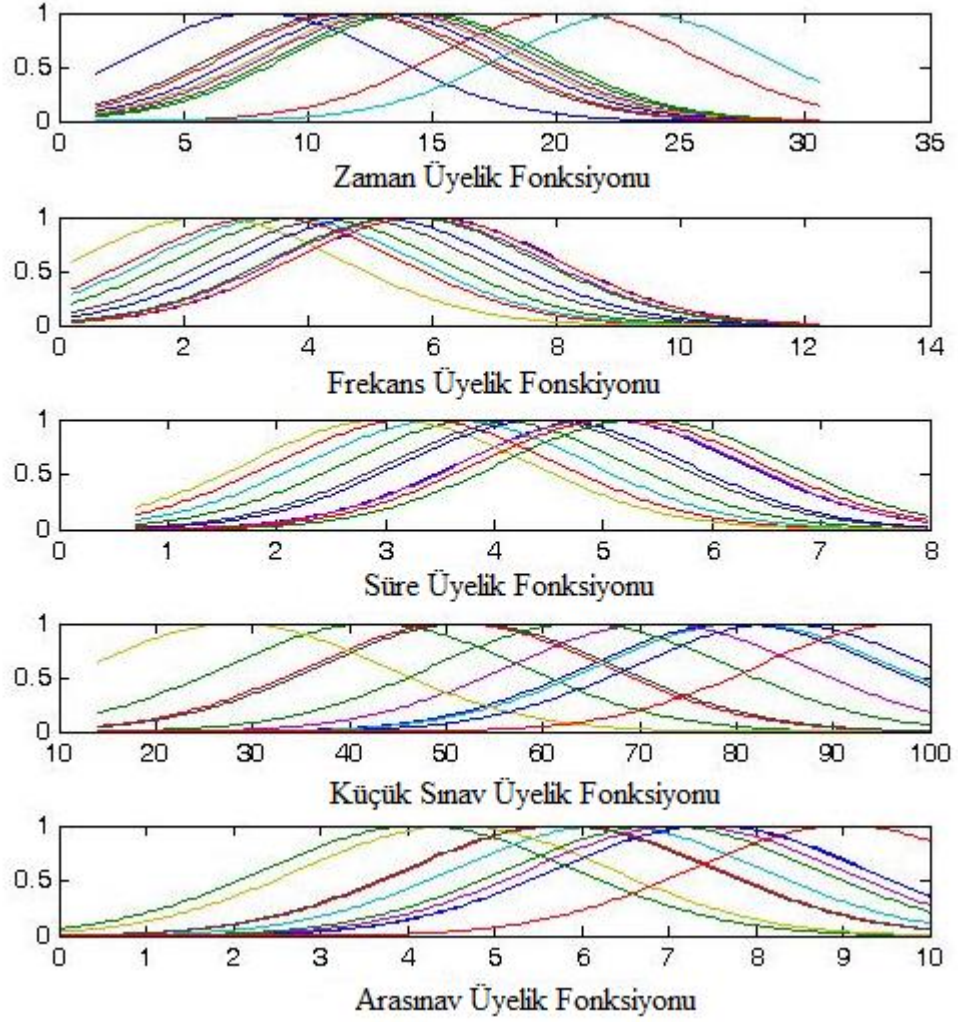
Tablo 4.20: Bulanık C-Ortalamlar kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük sınav	Arasınav
1	9,75	5,63	4,82	80,78	7,35
2	11,57	5,70	5,19	67,26	7,16
3	16,21	4,64	4,45	92,70	8,65
4	9,65	6,12	5,38	97,68	9,03
5	15,04	3,95	3,81	43,29	4,23
6	20,21	3,79	3,57	77,35	6,53
7	13,08	2,38	2,98	28,57	4,55
8	19,87	4,28	4,13	60,63	6,06
9	13,55	4,41	4,24	51,84	5,80

Tablo 4.19’da ki eğitim verisine göre en iyi doğruluk oranı küme sayısı 9 olan modele göre bulunmuştur. Bu aşamadan sonra yapılacak olan işlemlerde küme sayısı 9’a göre yapılacaktır.

Bulanık C-ortalamalar kümeleme yöntemi ile bulunan küme merkezleri tablo 4.20’de verilmiştir.

Üyelik fonksiyonlarının grafikleri şekil 4.28’de verilmiştir.



Şekil 4.28: Bulanık C-Ortalamalar kümeleme yöntemi üyelik fonksiyonları grafikleri.

Bulanık C-Ortalamalar Kümeleme yönteminden elde edilen üyelik fonksiyonları sonucunda en küçük kareler yöntemi ile hesaplanan kurallara ait parametreler tablo 4.21’de verilmiştir.

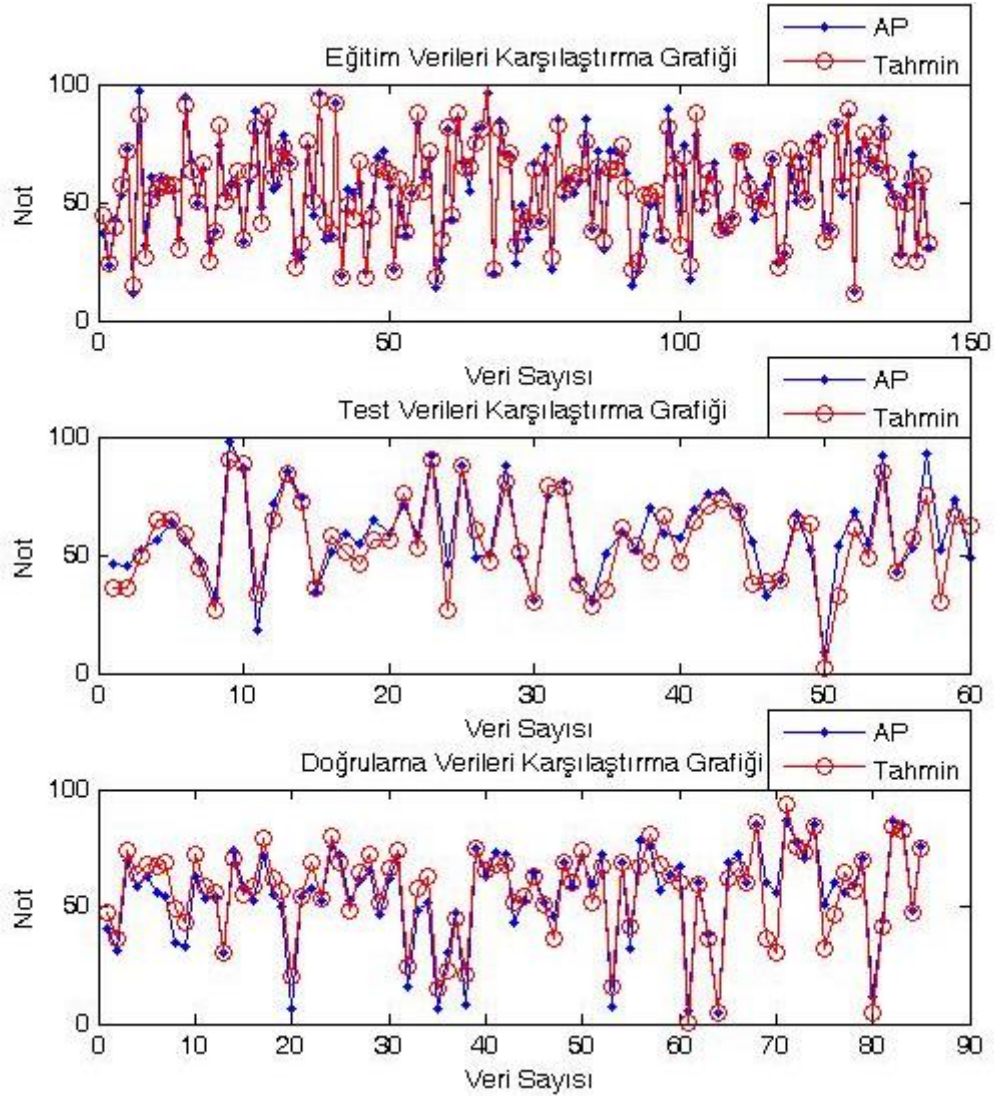
Tablo 4.21: Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	-156,10	-0,05	9,94	-4,71	2,62	1,68
2	215,18	1,27	-7,50	-0,54	-1,70	0,21
3	-244,46	-0,79	1,92	1,78	0,64	27,48
4	16,91	2,00	-3,77	9,94	-0,89	9,56
5	-41,12	0,73	1,65	9,30	0,05	1,60
6	-22,20	0,11	-2,92	8,54	0,40	8,14
7	2,43	-0,01	-0,69	-0,08	0,43	2,28
8	3,15	0,09	8,19	-5,65	0,33	-1,55
9	91,53	4,74	-8,62	-25,80	0,22	4,76

Şekil 4.29’da eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.

Bulanık C-ortalamalar kümeleme metoduna göre bulunan, küme merkezleri baz alınarak oluşturulan modele ait doğruluk değeri %91.09 olarak bulunmuştur.

Eğitim verisini test etmek için kullanılan test verisi girdi olarak modele uygulandığında %88.07 doğruluk oranı elde edilmiştir. Doğrulama verisi modele girdi olarak uygulandığında sonuç %83.94 bulunmuştur. Tablo 4.22’de 3 veri tipine ait veri sayısı görülmektedir. Tablo 4.23’te ise 3 veri tipine göre doğruluk oranları verilmiştir.



Şekil 4.29: Bulanık C-Ortalamalar kümeleme yöntemi ile kurulan bulanık model sonuçları.

Tablo 4.22 Verilere ait adet bilgisi.

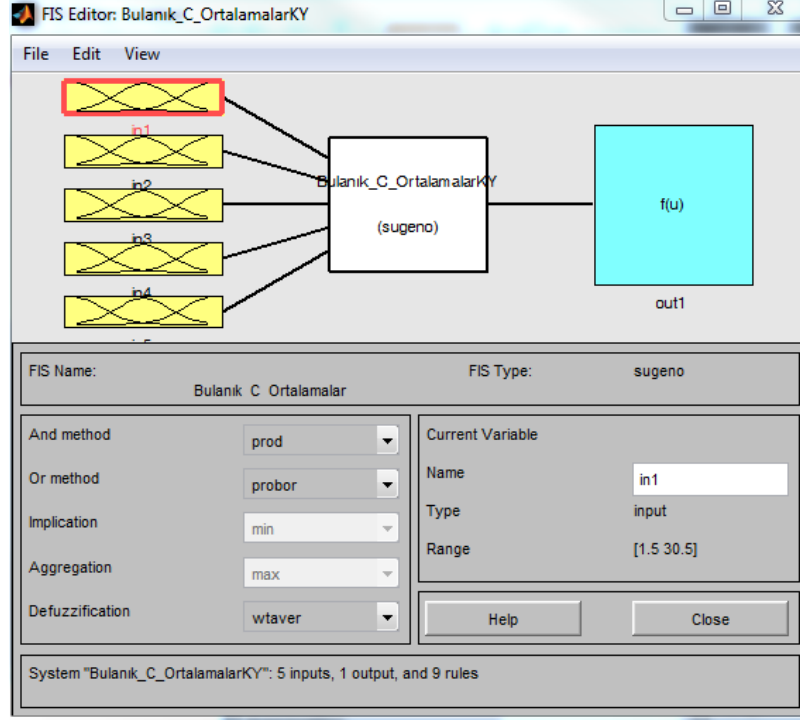
	Eğitim	Test	Doğrulama	Toplam
Veri Sayısı	143	60	85	288

Tablo 4.23 Verilere ait doğruluk oranları.

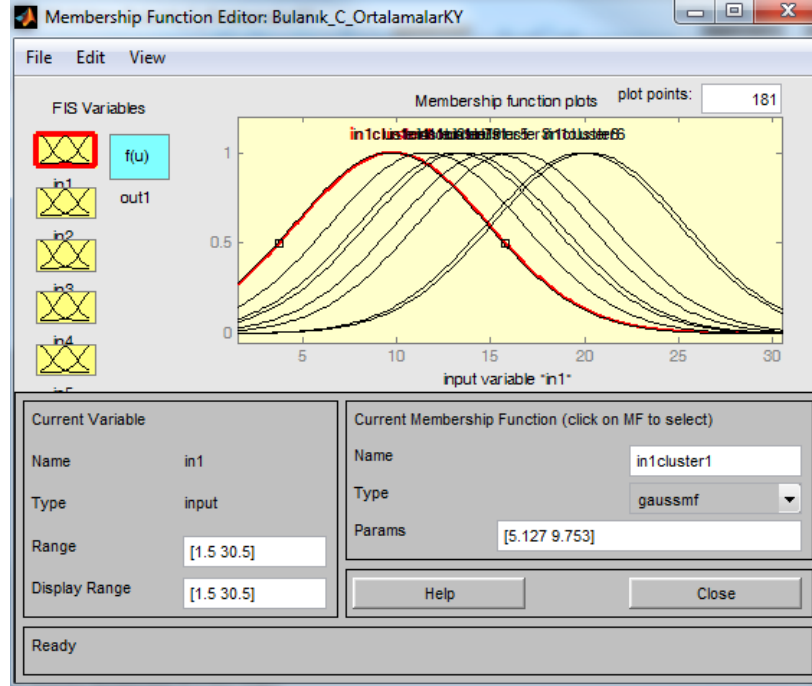
Küme Sayısı=9	Eğitim	Test	Doğrulama	Ortalama
Doğruluk Oranı (%)	91,09	88,07	83,94	87,70

Aykırı veri analizi sonrasında, aykırı veriler sistemden çıkarılmadığında yapılan analizlerde eğitim verilerinde başarı oranı %91,15 çıkarken doğrulama verilerinde bu oran %65,14'e düşmüştür. Bu sonuçlarda aykırı verilerin çıkarılmasıyla modelde anlamlı derecede farklılık meydana geldiği görülmektedir.

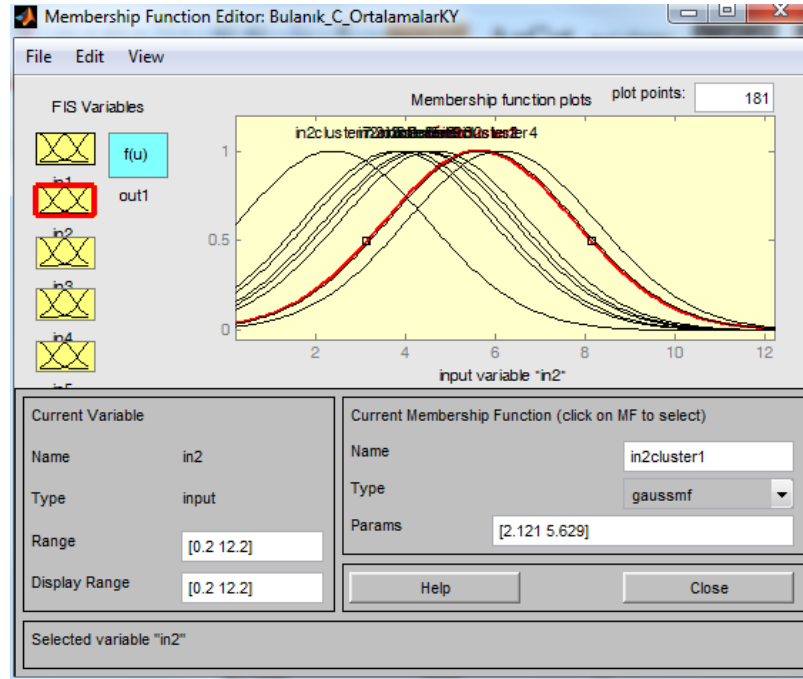
Matlab ortamında hazırlanan algoritma sonucunda Fis Editör’de açılabilir şekilde kaydedilen dosyaya ait ekran görüntüleri aşağıda verilmiştir.



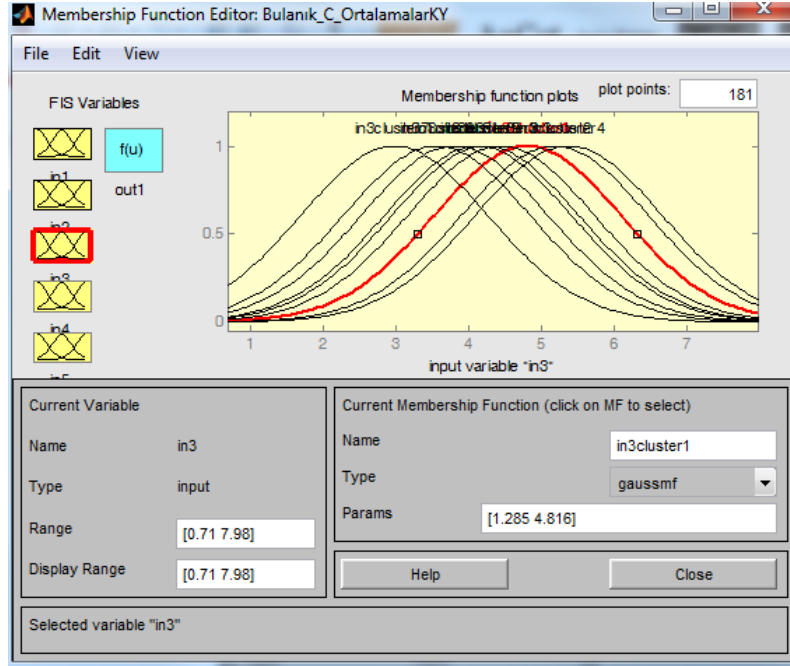
Şekil 4.30: Bulanik C-ortalamar kümeleme yöntemi Matlab ekran görüntüsü.



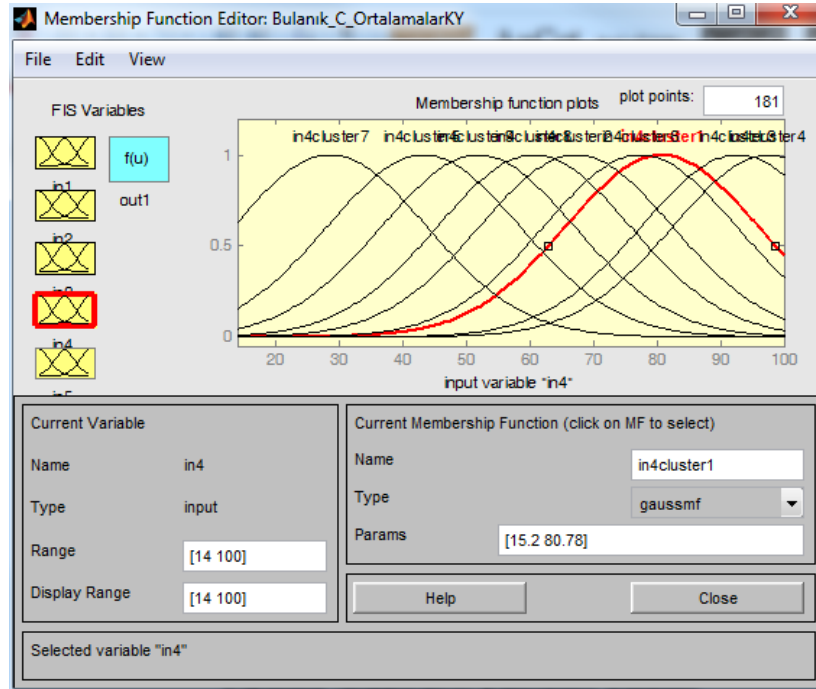
Şekil 4.31: Bulanık C-ortalamalar kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



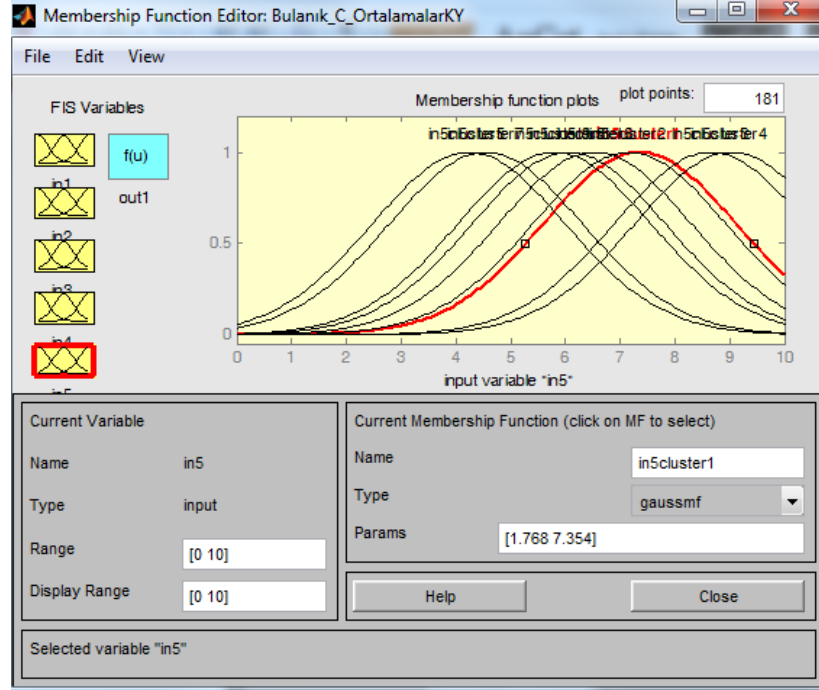
Şekil 4.32: Bulanık C-ortalamalar kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



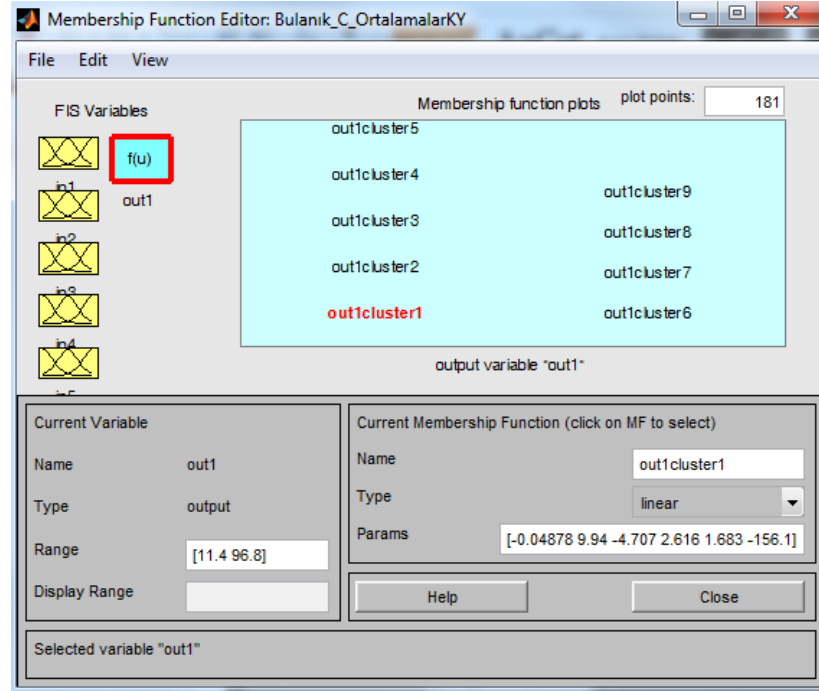
Şekil 4.33: Bulanık C-ortalamlar kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.34: Bulanık C-ortalamlar kümeleme yöntemi ödev girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.35: Bulanık C-ortalamlar kümeleme yöntemi arasınavgirdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.36: Bulanık C-ortalamlar kümeleme yöntemi akademik performans çıktısı, kurallara ait Matlab ekran görüntüsü.

4.4.1.5 Çıkarımlı (Subtractive) Kümeleme Yöntemi

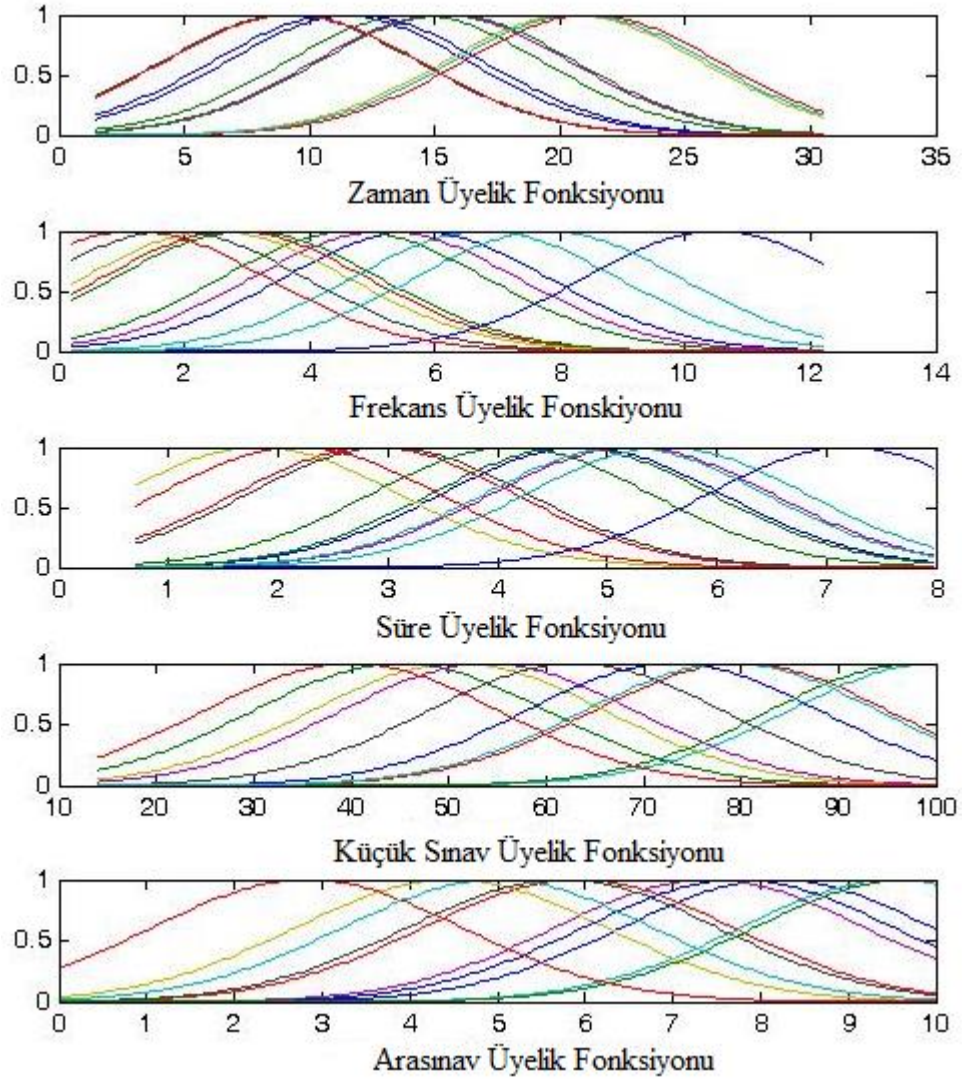
Çıkarımlı Kümeleme Yönteminde , K-Ortalamlar ve Bulanık C-Ortalamlar kümeleme yönteminden farklı olarak, küme sayısının girilmesine gerek yoktur. Bu yöntem sahip olduğu algoritma ile küme sayısını kendi belirlemektedir.

Çıkarımlı kümeleme metoduna göre bulunan küme merkezleri Tablo 4.24’de verilmiştir.

Tablo 4.24: Çıkarımlı kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük sınav	Arasınav
1	11,20	5,70	4,77	73,00	8,19
2	14,20	4,70	4,67	45,00	5,00
3	21,20	2,80	2,88	80,00	5,92
4	9,20	6,80	5,54	100,00	9,54
5	15,30	5,20	5,23	54,00	7,42
6	20,50	2,50	1,82	51,00	4,47
7	9,20	1,80	3,00	63,00	5,74
8	11,80	10,50	7,20	73,00	7,75
9	15,50	3,00	4,12	98,00	9,70
10	9,30	1,20	2,19	40,00	2,83
11	20,80	7,80	5,16	79,00	5,00

Çıkarımlı kümeleme yönteminde, algoritma 11 adet küme merkezi bulmuştur. Küme merkezlerine göre oluşan üyelik fonksiyonlarının grafikleri şekil 4.37’de verilmiştir.



Şekil 4.37: Çıkarımlı kümeleme yöntemi üyelik fonksiyonları grafikleri.

Çıkarımlı kümeleme yönteminden elde edilen üyelik fonksiyonları sonucunda hesaplanan kurallara ait parametreler tablo 4.25’de verilmiştir.

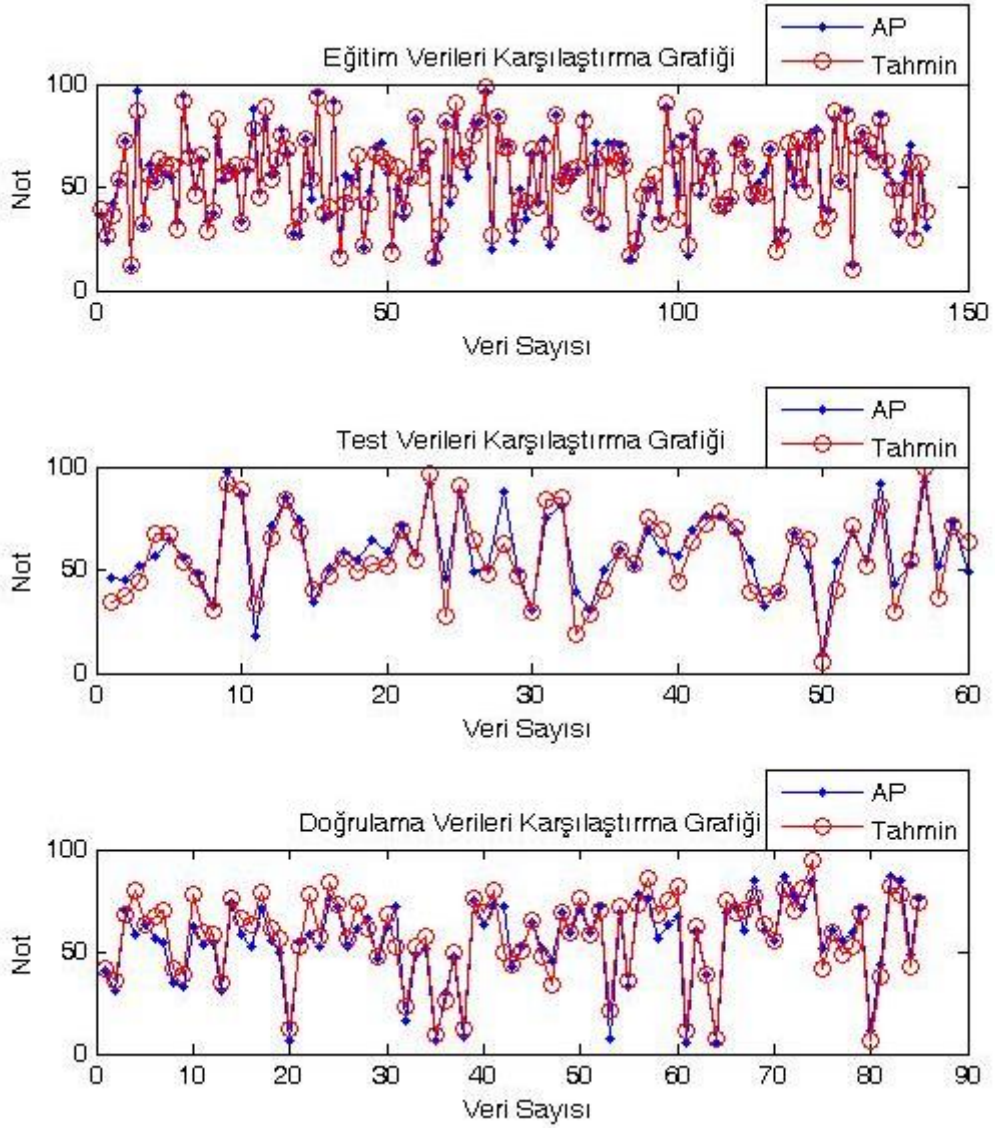
Tablo 4.25: Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	198,42	-1,50	1,16	-9,80	-0,89	-1,81
2	1,45	0,85	1,69	-2,12	0,13	3,18
3	-9,98	1,39	5,90	-5,53	0,28	4,01
4	362,30	2,98	-11,14	-4,14	-1,54	-5,66
5	-86,32	-0,24	-0,40	7,84	0,49	10,75
6	50,80	-1,96	-1,81	7,56	-0,11	3,72
7	-17,35	3,12	-3,60	-4,03	0,65	4,09
8	-95,29	-0,56	1,27	2,00	1,01	8,34
9	-93,02	1,45	0,69	-1,02	-0,55	21,75
10	28,40	-1,49	-2,12	0,00	0,12	-0,79
11	19,71	0,13	5,15	-1,07	-0,18	2,65

Şekil 4.38’de eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.

Bulanık C-ortalamlar kümeleme metoduna göre bulunan, küme merkezleri baz alınarak oluşturulan modele ait doğruluk değeri %91.33 olarak bulunmuştur.

Eğitim verisini test etmek için kullanılan test verisi girdi olarak modele uygulandığında %86.28 doğruluk oranı elde edilmiştir. Doğrulama verisi modele girdi olarak uygulandığında sonuç %83.24 bulunmuştur. Tablo 4.26’da 3 veri tipine ait veri sayısı görülmektedir. Tablo 4.27’de ise 3 veri tipine göre doğruluk oranları verilmiştir.



Şekil 4.38: Çıkarımlı kümeleme yöntemi ile kurulan bulanık model sonuçları.

Tablo 4.26: Verilere ait adet bilgisi.

	Eğitim	Test	Doğrulama	Toplam
Veri Sayısı	143	60	85	288

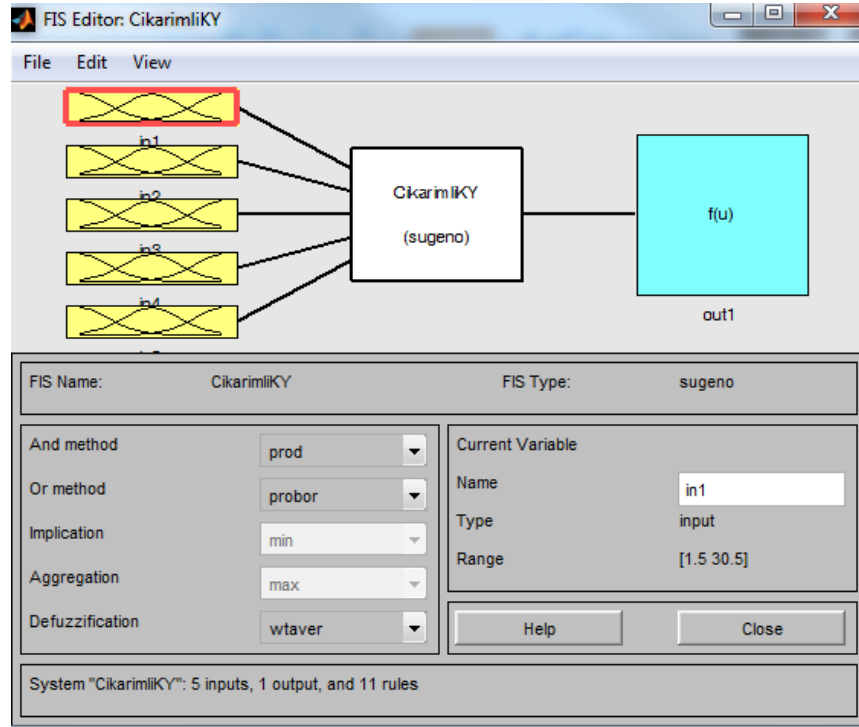
Tablo 4.27: Verilere ait doğruluk oranları.

Küme Sayısı=11	Eğitim	Test	Doğrulama	Ortalama
Doğruluk Oranı (%)	91,33	86,28	83,24	86,95

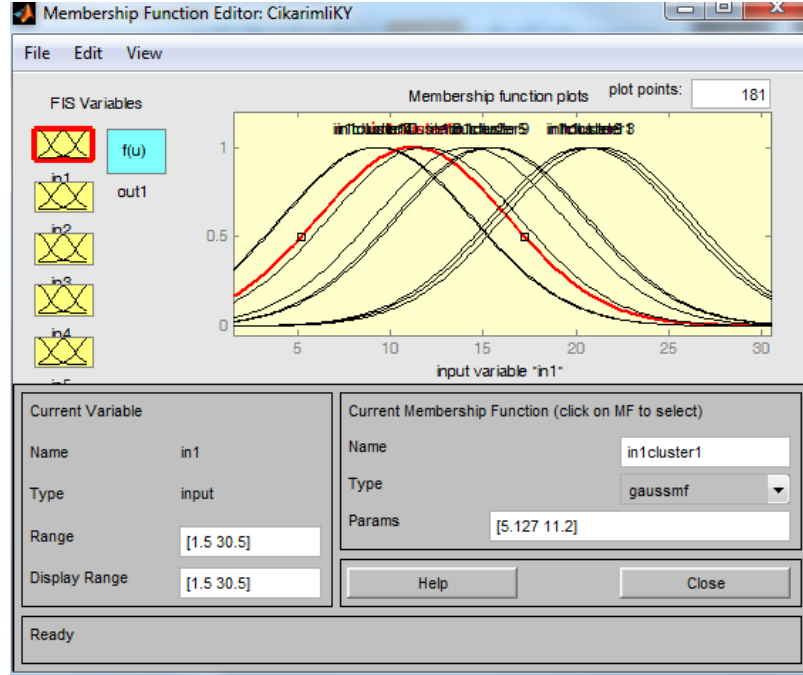
Aykırı veri analizi sonrasında, aykırı veriler sistemden çıkarılmadığında yapılan analizlerde eğitim verilerinde başarı oranı %89,75 olarak hesaplanırken, doğrulama

verilerinde bu oran %76,40'a düşmüştür. Bu sonuçlarda aykırı verilerin çıkarılmasıyla modelde anlamlı derecede farklılık meydana geldiğini göstermektedir.

Matlab ortamında hazırlanan algoritma sonucunda fis editörde açılabilir şekilde kaydedilen dosyaya ait ekran görüntüleri aşağıda verilmiştir.

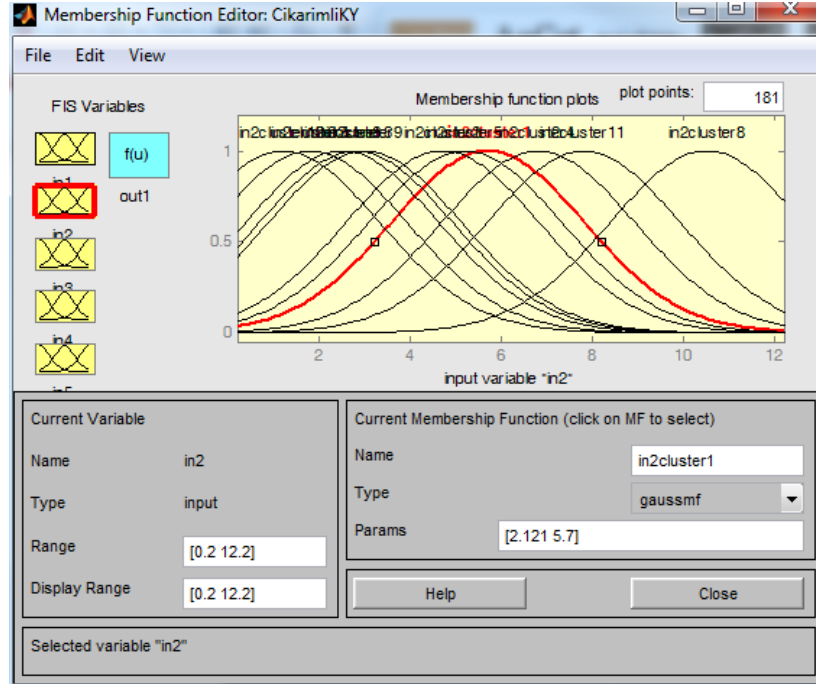


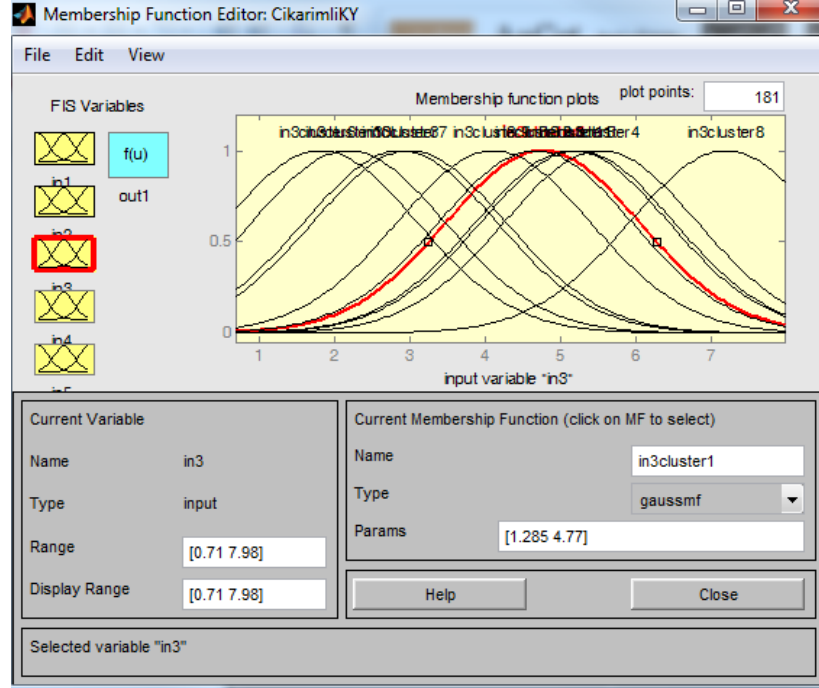
Şekil 4.39: Çıkarımlı kümeleme yöntemi Matlab ekran görüntüsü.



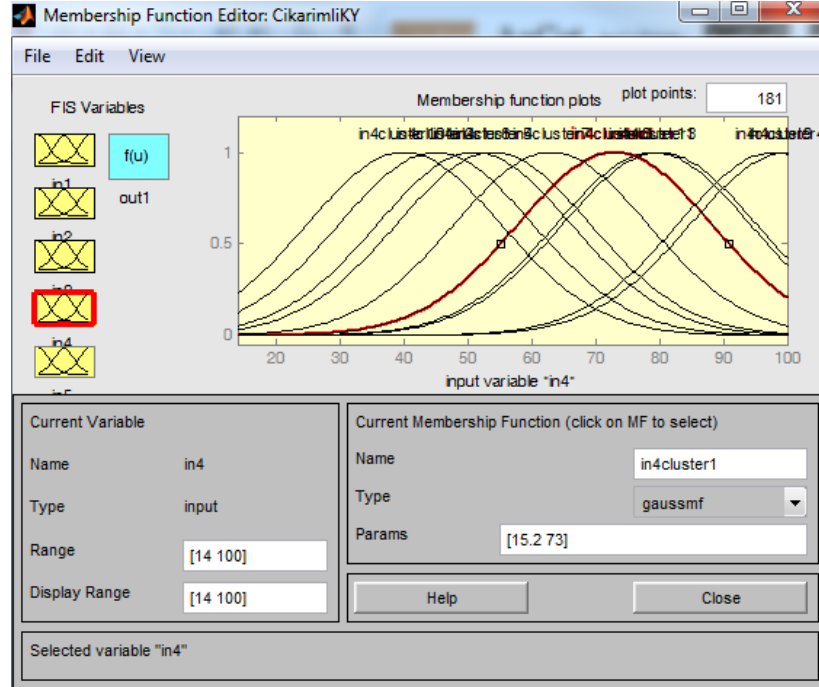
Şekil 4.40: Çıkarımlı kümeleme yöntemi zaman girdisi üyelik fonksiyonları Matlab ekran görüntüsü.

Şekil 4.41: Çıkarımlı kümeleme yöntemi frekans girdisi üyelik fonksiyonları Matlab ekran görüntüsü.

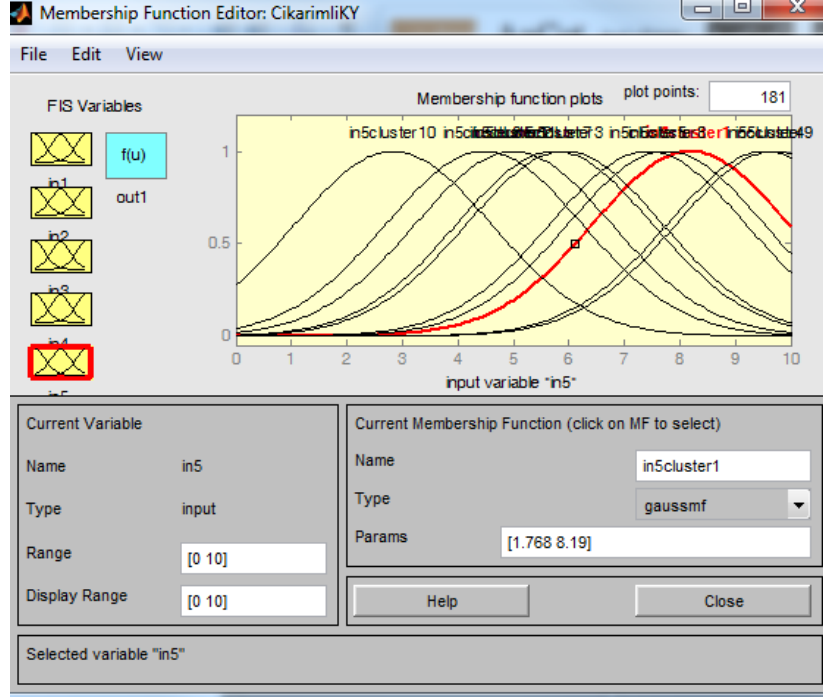




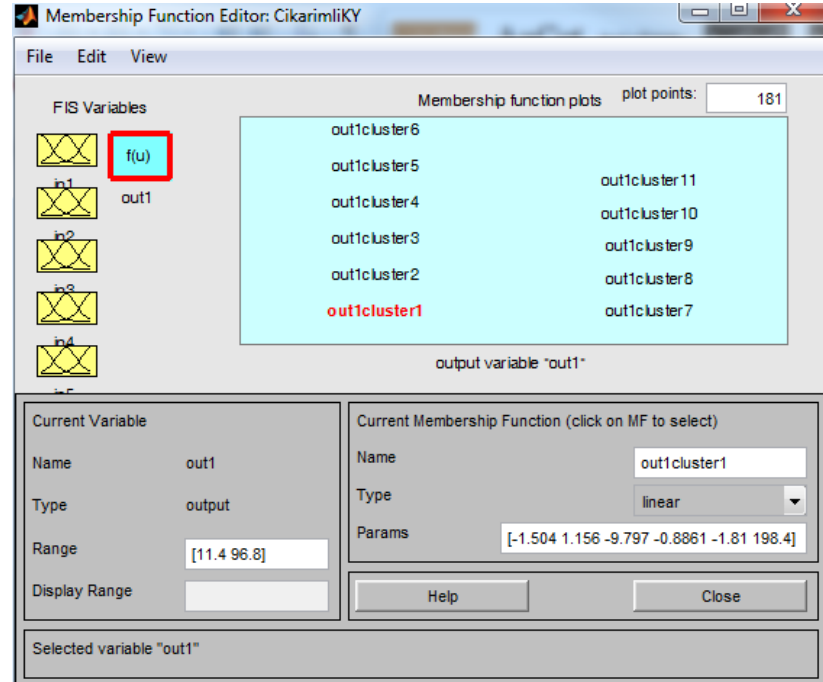
Şekil 4.42: Çıkarımlı kümeleme yöntemi süre girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.43: Çıkarımlı kümeleme yöntemi küçük sınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.44: Çıkarımlı kümeleme yöntemi arasınav girdisi üyelik fonksiyonları Matlab ekran görüntüsü.



Şekil 4.45: Çıkarımlı kümeleme yöntemi akademik performans çıktısı, kurallara ait Matlab ekran görüntüsü.

4.4.1.6 Modelin Geçerliliği

Geçerlilik, bir testin ölçmek istediği özelliği ölçmesi ile ilgili bir kavramdır. Bu bağlamda, bir test ölçmek istediği özelliği doğru ve diğer özelliklerle karıştırmadan ölçüyor ise bu testin geçerli olduğu söylenir [81]. Bu çalışmada oluşturulan modeller ile öğrencilerin yıl sonu akademik performans notu tahmin edildi. Gerçek değerler ile tahmin edilen değerler arasında anlamlı fark olup olmadığına dair iki eş arasındaki farkın önemlilik testi (paired t-test) uygulanmaktadır. Bu test bağımlı iki gruba ilişkin ortalamaların karşılaştırılmasında kullanılan bir testtir. Ancak farkların normal dağılım göstermesi gerekir. Eğer farklar normal dağılım göstermiyorsa veriler parametrik olmayan Wilcoxon T testi ile test edilir [82].

Test yöntemlerinden hangisinin kullanılacağını bilebilmek için öncelikle normallik testi olan Shapiro-Wilk testi yapılmıştır. Bu bölümdeki tüm testler SPSS 17 paket programı ile elde edilmiştir. Normalliği test etmek için kurulan hipotezler;

H0:%95 güvenle veriler normal dağılımlıdır.

H1:%95 güvenle veriler normal dağılımlı değildir.

Tablo 4.28'de Normallik Test sonuçları görülmektedir.

Tablo 4.28: Normallik Test Sonuçları.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Gerçek	,139	85	,000	,912	85	,000
Kortalamalar	,138	85	,000	,950	85	,003
BulanıkKortalamalar	,111	85	,011	,936	85	,000
Çıkarımlı	,110	85	,012	,931	85	,000

a. Lilliefors Significance Correction

Burada "Shapiro-Wilk" testine göre p değeri 0.05' den küçük olduğundan H0 hipotezi reddedilir ve veriler normal dağılıma sahip değildir sonucu çıkarılmıştır.

Veriler normal dağılmadığı için iki bağımlı grup arasında anlamlı fark olup olmadığına dair parametrik olmayan veriler için kullanılan wilcoxon testi uygulanmıştır. Test öncesinde kurulan hipotez testleri şu şekildedir;

H0:%95 güvenle gerçek değerler ile tahmin değerleri arasında anlamlı fark yoktur.

H1:%95 güvenle gerçek değerler ile tahmin değerleri arasında anlamlı fark vardır.

Tablo 4.29'da sıralama değer sonuçları ve Tablo 4.30'da K-Ortalamalar Kümeleme Yöntemi (KOKY) ile elde edilen tahmin değerleri ile gerçek değerlere ait Wilcoxon T-Testi sonuçları verilmiştir.

Tablo 4.29: Sıralama değer sonuçları.

		Ranks		
		N	Mean Rank	Sum of Ranks
Kortalamalar - Gerçek	Negative Ranks	43 ^a	44,74	1924,00
	Positive Ranks	42 ^b	41,21	1731,00
	Ties	0 ^c		
	Total	85		

a. Kortalamalar < Gerçek

b. Kortalamalar > Gerçek

c. Kortalamalar = Gerçek

Tablo 4.30: KOKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.

Test Statistics ^a	
	Kortalamalar - Gerçek
Z	-,423 ^b
Asymp. Sig. (2-tailed)	,672

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

p değeri $0,672 > 0,05$ olduğundan H0 hipotezi kabul edilir yani K-Ortalamalar Kümeleme Metodu ile elde edilen tahmin değerleri ile gerçek değerler arasında anlamlı olarak herhangi bir fark yoktur.

Tablo 4.31'de sıralama değer sonuçları ve Tablo 4.32'de Bulanık C-Ortalamalar Kümeleme Yöntemi (BCOKY) ile elde edilen tahmin değerleri ile gerçek değerlere ait Wilcoxon T-Testi sonuçları verilmiştir.

Tablo 4.31: Sıralama değer sonuçları.

		Ranks		
		N	Mean Rank	Sum of Ranks
BulanıkCortalamalar - Gerçek	Negative Ranks	40 ^a	35,90	1436,00
	Positive Ranks	45 ^b	49,31	2219,00
	Ties	0 ^c		
	Total	85		

a. BulanıkCortalamalar < Gerçek

b. BulanıkCortalamalar > Gerçek

c. BulanıkCortalamalar = Gerçek

Tablo 4.32: BCOKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.

Test Statistics ^a	
	BulanıkCortalamalar - Gerçek
Z	-1,715 ^b
Asymp. Sig. (2-tailed)	,086

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

p değeri $0,086 > 0,05$ olduğundan H_0 hipotezi kabul edilir yani Bulanık C-Ortalamlar Kümeleme Metodu ile elde edilen tahmin değerleri ile gerçek değerler arasında anlamlı olarak herhangi bir fark yoktur.

Tablo 4.33'de sıralama değer sonuçları ve Tablo 4.34'de Çıkarımlı Kümeleme Yöntemi (ÇKY) ile elde edilen tahmin değerleri ile gerçek değerlere ait Wilcoxon T-Testi sonuçları verilmiştir.

Tablo 4.33: Sıralama değer sonuçları.

		Ranks		
		N	Mean Rank	Sum of Ranks
Çıkarımlı - Gerçek	Negative Ranks	28 ^a	40,71	1140,00
	Positive Ranks	57 ^b	44,12	2515,00
	Ties	0 ^c		
	Total	85		

a. Çıkarımlı < Gerçek

b. Çıkarımlı > Gerçek

c. Çıkarımlı = Gerçek

Tablo 4.34: ÇKY ile elde edilen tahmin sonuçları ile gerçek değerlere ait Wilcoxon T-Test Sonuçları.

Test Statistics ^a	
	Çıkarımlı - Gerçek
Z	-3,012 ^b
Asymp. Sig. (2-tailed)	,003

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

p değeri $0,003 < 0,05$ olduğundan H_0 hipotezi reddedilir, yani Çıkarımlı Kümeleme Metodu ile elde edilen tahmin değerleri ile gerçek değerler arasında anlamlı fark vardır.

Çalışmada kullanılan modellerden KOKY ile BCOKY ile elde edilen tahmin değerleri ile gerçek değerler arasında anlamlı bir fark bulunamamıştır. Bu iki model için aynı zamanda varyansların eşitliği testi yapılarak, varyans ortalamaları arasında anlamlı fark olup olmadığı test edilmiştir. Verilerimiz parametrik olmadığı için Mann-Whitney U testi kullanılmıştır. Mann-Whitney U testinde H_0 hipotezi; iki dağılım arasında fark yoktur. H_1 hipotezi ise iki dağılım arasında fark vardır ya da bir grubun değerleri diğerinden farklıdır [81]. Tablo 4.35’de gerçek değerler ile KOKY ve BCOKY modellerine ait varyans eşitlik test sonuçları görülmektedir.

Tablo 4.35: KOKY ve BCOKY modellerinin varyans eşitlik test sonuçları.

Test Statistics ^a		
	KOKY	BCOKY
Mann-Whitney U	3475,000	3444,000
Wilcoxon W	7045,000	7099,000
Z	-,299	-,396
Asymp. Sig. (2-tailed)	,765	,692

a. Grouping Variable: VAR00008

Tablo 4.35'deki p değerleri $0,765 > 0,05$ ve $0,692 > 0,05$ olduğundan H_0 hipotezi kabul edilir, yani KOKY ve BCOKY modelleri ile gerçek değerlerin dağılımları arasında fark yoktur, sonucu kabul edilir.

Korelasyon katsayıları değişkenler arasındaki ilişkinin kuvveti ve yönü hakkında bilgi veren ölçülerdir. Çalışmada KOKY ve BCOKY modelleri anlamlılık açısından geçerli modellerdir. Bu modellerle elde edilen tahmin değerlerinin gerçek sonuçlarla olan ilişkisinin ölçmek için Spearman korelasyon katsayısı incelenmiştir [81].

Tablo 4.36: Gerçek Değerlerle KOKY için Spearman Korelasyon Katsayısı.

		Gerçek	Kortalamalar
Spearman's rho	Correlation Coefficient	1,000	,897**
	Gerçek		
	Sig. (2-tailed)	.	,000
	N	85	85
	Kortalamalar		
	Sig. (2-tailed)	,000	.
	N	85	85

** . Correlation is significant at the 0.01 level (2-tailed).

Tablo 4.37: Gerçek Değerlerle BOCKY için Spearman Korelasyon Katsayısı.

		Gerçek	BulanıkCortalamar
Gerçek	Correlation Coefficient	1,000	,902**
	Sig. (2-tailed)	.	,000
	N	85	85
Spearman's rho	Correlation Coefficient	,902**	1,000
	Sig. (2-tailed)	,000	.
	N	85	85

** . Correlation is significant at the 0.01 level (2-tailed).

Tablo 4.36 ve Tablo 4.37 sonuçlarına bakıldığında Bulanık Ortalamalar Kümeleme Yöntemi (BOCKY) ile elde edilen tahmin değerlerinin korelasyon katsayısının, K-Ortalamlar Kümeleme Yöntemi ile elde edilen tahmin değerlerinin korelasyon katsayısından büyük olduğu görülmüştür. Modellerin tahmin değerleri ile gerçek değerler arasında pozitif yönlü güçlü bir korelasyon ($r=0,897$ ve $r=0,902$) vardır.

4.4.1.7 Sonuçların Karşılaştırılması

Tüm metotlara ait eğitim verileri ve test verileri kullanılarak bulunan doğruluk oranları tablo 4.38’de verilmiştir.

Tablo 4.38: Akademik performansın sayısal olarak tahminine dayalı doğruluk oranları.

Method	Doğruluk Oranı (%)				
	Küme Sayısı	Eğitim Verisi	Test Verisi	Doğrulama Verisi	Ortalama
K-Ortalamlar Kümeleme Yöntemi	8	90,93	84,00	80,21	85,05
Bulanık C-Ortalamlar Kümeleme Yöntemi	7	91,09	88,07	83,94	87,70

Çıkarımlı Kümeleme Yöntemi	9	91,33	86,28	83,24	86,95
-----------------------------------	---	-------	-------	-------	-------

Kategorik olarak öğrenciler sadece geçti kaldı şeklinde sınıflandırıldığında doğruluk oranları Tablo 4.39’da verilmiştir.

Tablo 4.39: Akademik performansın “geçti”-“kaldı” tahminine dayalı doğruluk oranları.

Kategorik Sınıflandırma Doğruluk Oranı (Geçti,Kaldı) (%)					
Methot	Küme Sayısı	Eğitim Verisi	Test Verisi	Doğrulama Verisi	Ortalama
K-Ortalamlar Kümeleme Yöntemi	8	92,31	83,33	90,59	88,74
Bulanık C-Ortalamlar Kümeleme Yöntemi	7	93,01	88,33	95,29	92,21
Çıkarımlı Kümeleme Yöntemi	9	93,71	86,67	96,47	92,28

Tablo 4.40’da kullanılan kümeleme yöntemleri ile kurulan modelde ki doğrulama verileri kullanılarak tahmin edilen uzaktan eğitim öğrencilerinin akademik performans sonuçları, 3 farklı metot sonucuna göre görülmektedir. Sonuçlar sayısal olarak bulunmuştur.

Tablo 4.40: Kümeleme metotlarına göre sayısal tahmin sonuçları.

No	Zaman	Frekans	Süre	Küçük sınav	Arasınav	AP	Tahmin		
							K-Ortalamlar Kümeleme Yöntemi	Bulanık C-Ortalamlar Kümeleme Yöntemi	Çıkarımlı Kümeleme Yöntemi
1	10,00	4,00	7,42	59,00	5,21	41,02	36,43	47,57	41,27
2	6,60	4,33	3,65	44,00	6,42	31,03	38,14	36,34	35,78
3	7,40	2,00	7,53	79,00	7,81	70,00	71,74	74,08	68,19
4	6,80	6,67	4,30	68,00	8,37	58,60	72,75	64,01	79,46
5	6,20	2,83	6,30	77,00	7,28	63,00	50,56	68,18	63,24
6	20,80	2,33	2,10	87,00	5,79	56,11	62,60	67,27	67,02
7	15,60	5,67	3,08	66,00	8,32	54,03	67,42	68,43	70,24
8	6,80	7,67	2,35	48,00	5,57	34,80	46,33	49,10	41,98
9	12,40	4,00	2,85	37,00	6,52	33,10	34,95	42,98	38,97
10	7,80	7,67	4,41	91,00	7,38	62,70	72,39	72,24	78,08
11	19,60	3,50	5,55	53,00	7,94	53,20	51,24	58,15	59,57

12	14,60	2,50	5,64	59,00	7,53	54,13	50,71	56,17	58,35
13	11,20	5,83	4,62	33,00	6,31	30,76	31,97	29,97	35,14
14	11,60	2,50	4,11	70,00	9,42	73,75	83,45	70,53	75,82
15	12,20	13,33	4,79	81,00	6,93	58,80	50,37	55,41	66,03
16	14,60	1,00	0,75	48,00	8,06	52,60	49,51	58,77	63,22
17	11,20	5,50	4,45	98,00	8,28	71,31	73,55	79,44	78,62
18	13,00	3,33	3,80	75,00	6,92	55,38	67,01	61,75	61,36
19	19,40	4,83	6,11	49,00	7,42	49,80	51,59	57,12	55,95
20	10,00	6,00	2,45	33,00	2,24	6,60	15,22	19,69	12,27
21	12,00	7,83	5,79	70,00	5,66	54,00	52,70	53,90	52,55
22	15,80	6,00	2,09	68,00	8,46	57,92	67,26	68,39	77,99
23	18,40	3,83	1,31	59,00	7,36	52,44	51,63	53,35	57,11
24	15,20	5,00	2,07	86,00	9,21	75,55	70,96	79,75	83,61
25	7,40	3,17	4,17	71,00	9,00	72,40	70,18	69,20	73,19
26	4,00	11,50	5,06	56,00	8,12	52,40	46,35	48,09	56,83
27	7,20	10,67	4,53	71,00	8,66	61,20	65,31	64,77	73,66
28	12,20	3,83	3,19	88,00	7,60	65,77	70,24	71,83	61,34
29	15,20	4,00	5,26	58,00	6,40	46,19	45,47	51,39	47,20
30	6,20	9,67	5,48	76,00	7,49	62,01	66,69	66,02	68,41
31	4,80	10,33	3,68	98,00	8,28	71,71	73,06	73,96	52,90
32	13,60	3,67	4,18	32,00	3,54	15,90	24,54	24,66	22,81
33	17,40	2,67	4,40	58,00	7,07	48,60	49,70	57,53	52,58
34	14,00	4,17	1,06	51,00	7,62	51,80	50,14	62,93	57,46
35	11,80	2,00	2,09	20,00	3,21	6,07	16,24	14,71	9,60
36	5,40	2,67	4,28	38,00	6,19	30,27	32,65	22,93	26,22
37	14,60	6,50	5,20	39,00	7,55	47,20	46,22	45,04	49,62
38	8,00	4,00	1,96	31,00	3,34	8,43	23,70	21,00	12,00
39	8,60	6,33	5,24	75,00	8,94	75,00	78,79	74,43	76,29
40	14,00	5,50	3,40	75,00	7,87	62,97	69,11	65,62	69,00
41	5,60	5,00	4,08	72,00	8,83	73,00	70,70	67,83	79,57
42	2,60	13,33	3,31	91,00	8,60	72,00	72,35	67,85	49,36
43	19,60	3,50	3,76	43,00	6,78	42,80	44,85	51,36	44,04
44	12,80	12,17	5,24	55,00	8,25	52,60	35,16	54,08	50,41
45	11,40	6,33	5,96	67,00	7,82	64,64	64,20	62,67	65,10
46	9,20	14,00	4,59	51,00	8,25	51,80	26,81	50,82	48,04
47	12,20	12,50	5,80	43,00	7,55	46,00	20,11	36,38	33,69
48	12,80	6,00	4,86	88,00	7,56	69,22	66,56	69,14	69,45
49	10,60	10,00	5,89	62,00	8,28	58,91	52,22	60,21	59,11
50	9,20	8,00	6,03	87,00	7,79	71,32	74,75	73,91	76,39
51	7,40	3,00	5,28	62,00	7,21	59,60	41,62	51,70	57,89
52	19,00	5,17	5,00	64,00	9,22	71,80	68,41	67,50	70,24
53	23,40	1,00	2,24	31,00	2,24	7,20	11,82	15,88	20,95
54	12,20	5,50	2,24	83,00	7,16	68,84	69,63	68,90	72,45
55	20,20	2,50	4,44	27,00	5,92	32,40	37,65	41,60	35,96
56	19,80	4,17	5,92	70,00	9,17	78,20	67,25	67,51	72,80
57	6,40	7,00	2,19	89,00	8,66	75,80	87,55	80,36	85,36
58	17,60	6,50	2,43	59,00	8,37	56,80	60,80	68,26	68,18
59	9,00	5,83	3,15	71,00	7,55	63,60	78,81	63,92	75,15
60	2,60	17,50	4,87	72,00	8,83	67,00	64,01	60,67	81,73
61	21,40	0,50	0,41	5,00	3,03	5,83	-1,25	0,25	11,02
62	11,60	6,83	4,28	68,00	6,58	60,25	62,26	59,21	62,67
63	12,40	1,17	3,44	47,00	5,00	38,40	38,10	36,04	38,85
64	12,00	1,33	1,64	5,00	2,24	5,00	1,62	4,64	7,26
65	4,20	11,00	4,29	74,00	8,49	69,20	72,44	61,60	74,69
66	20,60	3,17	3,49	86,00	8,00	72,00	68,30	66,21	69,05
67	9,40	10,33	4,28	65,00	7,87	60,40	52,08	60,00	69,80
68	16,33	12,50	3,31	87,00	9,90	85,00	91,28	85,50	76,55
69	5,00	5,17	6,56	47,00	7,62	60,00	37,00	35,94	62,92
70	4,00	6,17	6,32	49,00	6,68	55,72	32,21	30,20	55,15
71	13,00	7,17	2,47	100,00	9,17	86,80	81,13	93,49	81,20
72	5,40	9,83	6,16	94,00	6,94	77,22	72,05	75,39	69,72
73	12,40	11,67	6,24	88,00	8,25	70,62	77,06	73,36	81,03
74	7,00	4,50	7,60	92,00	9,11	85,00	85,64	84,02	94,17
75	9,00	6,67	6,20	44,00	6,08	51,20	37,79	31,76	41,95
76	6,40	10,33	5,38	47,00	8,94	60,40	42,05	46,75	58,90
77	14,80	3,33	7,55	54,00	5,93	55,83	38,14	64,16	48,19
78	9,00	4,67	3,01	64,00	5,93	59,04	64,01	56,64	51,99
79	11,20	10,00	6,71	79,00	7,75	70,80	69,08	69,77	69,45
80	19,80	2,17	3,08	23,00	0,00	11,40	0,05	4,82	5,91
81	19,80	6,50	4,76	45,00	5,20	44,00	39,77	41,38	37,89
82	7,00	5,50	3,13	92,00	9,00	86,60	81,80	84,11	82,10
83	4,80	3,67	4,06	93,00	8,71	84,98	73,77	82,08	78,41

84	17,20	2,00	4,65	55,00	5,66	48,40	41,85	47,16	42,91
85	5,60	9,67	6,51	69,00	9,22	75,80	77,20	74,69	73,71

Tablo 4.40: (devam): Kümeleme Metotlarına Göre Kategorik Tahmin Sonuçları.

Tablo 4.41’de kullanılan kümeleme yöntemleri ile kurulan modelde ki doğrulama verileri kullanılarak tahmin edilen uzaktan eğitim öğrencilerinin akademik performans sonuçları geçti kaldı olarak bulunmaktadır.

Tablo 4.41: Kümeleme Metotlarına Göre Kategorik Tahmin Sonuçları.

No	Zaman	Frekans	Süre	Küçük sınav	Arasınav	AP	Tahmin		
							K- Ortalamalar Kümeleme Yöntemi	Bulanık C- Ortalamalar Kümeleme Yöntemi	Çıkarımlı Kümeleme Yöntemi
1	10	4	7,42	59	5,21	KALDI	KALDI	GEÇTİ	KALDI
2	6,6	4,33	3,65	44	6,42	KALDI	KALDI	KALDI	KALDI
3	7,4	2	7,53	79	7,81	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
4	6,8	6,67	4,3	68	8,37	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
5	6,2	2,83	6,3	77	7,28	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
6	20,8	2,33	2,1	87	5,79	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
7	15,6	5,67	3,08	66	8,32	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
8	6,8	7,67	2,35	48	5,57	KALDI	GEÇTİ	GEÇTİ	KALDI
9	12,4	4	2,85	37	6,52	KALDI	KALDI	KALDI	KALDI
10	7,8	7,67	4,41	91	7,38	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
11	19,6	3,5	5,55	53	7,94	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
12	14,6	2,5	5,64	59	7,53	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
13	11,2	5,83	4,62	33	6,31	KALDI	KALDI	KALDI	KALDI
14	11,6	2,5	4,11	70	9,42	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
15	12,2	13,33	4,79	81	6,93	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
16	14,6	1	0,75	48	8,06	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
17	11,2	5,5	4,45	98	8,28	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
18	13	3,33	3,8	75	6,92	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
19	19,4	4,83	6,11	49	7,42	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
20	10	6	2,45	33	2,24	KALDI	KALDI	KALDI	KALDI
21	12	7,83	5,79	70	5,66	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
22	15,8	6	2,09	68	8,46	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
23	18,4	3,83	1,31	59	7,36	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
24	15,2	5	2,07	86	9,21	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
25	7,4	3,17	4,17	71	9	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
26	4	11,5	5,06	56	8,12	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
27	7,2	10,67	4,53	71	8,66	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
28	12,2	3,83	3,19	88	7,6	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
29	15,2	4	5,26	58	6,4	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
30	6,2	9,67	5,48	76	7,49	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
31	4,8	10,33	3,68	98	8,28	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
32	13,6	3,67	4,18	32	3,54	KALDI	KALDI	KALDI	KALDI
33	17,4	2,67	4,4	58	7,07	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
34	14	4,17	1,06	51	7,62	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
35	11,8	2	2,09	20	3,21	KALDI	KALDI	KALDI	KALDI

36	5,4	2,67	4,28	38	6,19	KALDI	KALDI	KALDI	KALDI
37	14,6	6,5	5,2	39	7,55	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
38	8	4	1,96	31	3,34	KALDI	KALDI	KALDI	KALDI
39	8,6	6,33	5,24	75	8,94	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
40	14	5,5	3,4	75	7,87	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
41	5,6	5	4,08	72	8,83	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
42	2,6	13,33	3,31	91	8,6	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
43	19,6	3,5	3,76	43	6,78	KALDI	KALDI	GEÇTİ	KALDI
44	12,8	12,17	5,24	55	8,25	GEÇTİ	KALDI	GEÇTİ	GEÇTİ
45	11,4	6,33	5,96	67	7,82	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
46	9,2	14	4,59	51	8,25	GEÇTİ	KALDI	GEÇTİ	GEÇTİ
47	12,2	12,5	5,8	43	7,55	GEÇTİ	KALDI	KALDI	KALDI
48	12,8	6	4,86	88	7,56	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
49	10,6	10	5,89	62	8,28	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
50	9,2	8	6,03	87	7,79	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
51	7,4	3	5,28	62	7,21	GEÇTİ	KALDI	GEÇTİ	GEÇTİ
52	19	5,17	5	64	9,22	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
53	23,4	1	2,24	31	2,24	KALDI	KALDI	KALDI	KALDI
54	12,2	5,5	2,24	83	7,16	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
55	20,2	2,5	4,44	27	5,92	KALDI	KALDI	KALDI	KALDI
56	19,8	4,17	5,92	70	9,17	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
57	6,4	7	2,19	89	8,66	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
58	17,6	6,5	2,43	59	8,37	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
59	9	5,83	3,15	71	7,55	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
60	2,6	17,5	4,87	72	8,83	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
61	21,4	0,5	0,41	5	3,03	KALDI	KALDI	KALDI	KALDI
62	11,6	6,83	4,28	68	6,58	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
63	12,4	1,17	3,44	47	5	KALDI	KALDI	KALDI	KALDI
64	12	1,33	1,64	5	2,24	KALDI	KALDI	KALDI	KALDI
65	4,2	11	4,29	74	8,49	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
66	20,6	3,17	3,49	86	8	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
67	9,4	10,33	4,28	65	7,87	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
68	16,33	12,5	3,31	87	9,9	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
69	5	5,17	6,56	47	7,62	GEÇTİ	KALDI	KALDI	GEÇTİ
70	4	6,17	6,32	49	6,68	GEÇTİ	KALDI	KALDI	GEÇTİ
71	13	7,17	2,47	100	9,17	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
72	5,4	9,83	6,16	94	6,94	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
73	12,4	11,67	6,24	88	8,25	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
74	7	4,5	7,6	92	9,11	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
75	9	6,67	6,2	44	6,08	GEÇTİ	KALDI	KALDI	KALDI
76	6,4	10,33	5,38	47	8,94	GEÇTİ	KALDI	GEÇTİ	GEÇTİ
77	14,8	3,33	7,55	54	5,93	GEÇTİ	KALDI	GEÇTİ	GEÇTİ
78	9	4,67	3,01	64	5,93	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
79	11,2	10	6,71	79	7,75	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
80	19,8	2,17	3,08	23	0	KALDI	KALDI	KALDI	KALDI
81	19,8	6,5	4,76	45	5,2	KALDI	KALDI	KALDI	KALDI
82	7	5,5	3,13	92	9	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
83	4,8	3,67	4,06	93	8,71	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ
84	17,2	2	4,65	55	5,66	GEÇTİ	KALDI	GEÇTİ	KALDI
85	5,6	9,67	6,51	69	9,22	GEÇTİ	GEÇTİ	GEÇTİ	GEÇTİ

Tablo 4.41: (devam): Kümeleme Metotlarına Göre Kategorik Tahmin Sonuçları.

4.4.2 Takagi-Sugeno Tipi Bulanık Mantık Çıkarım Sisteminde Kurallara Ait Parametrelerin Kantil Yöntemine Göre Tahmin Edilmesi

Kantil yönteminde aykırı veri analizi yapılmadan model kurulmuştur. Bu modelde kullanılan verilerin sayısı tablo 4.42’de verilmiştir.

Tablo 4.42: Verilere ait adet bilgisi.

	Eğitim	Test	Doğrulama	Toplam
Veri Sayısı	153	65	95	313

En küçük kareler yöntemi kullanılarak yapılan uygulamada toplam veri sayısı 288’dir. Aradaki fark olan 15 veri aykırı veri olarak model dışında bırakılmıştır. Kantil yöntemi kullanılarak yapılan uygulamada bu 15 veri model içinde tutulmuştur. Bu yöntem ile önceki yöntemde olduğu gibi aynı kümeleme yöntemleri ile modeller kurulmuştur. Bu kümeleme yöntemleri K-Ortalamlar, Bulanık C-Ortalamlar ve Çıkarımlı kümeleme yöntemidir.

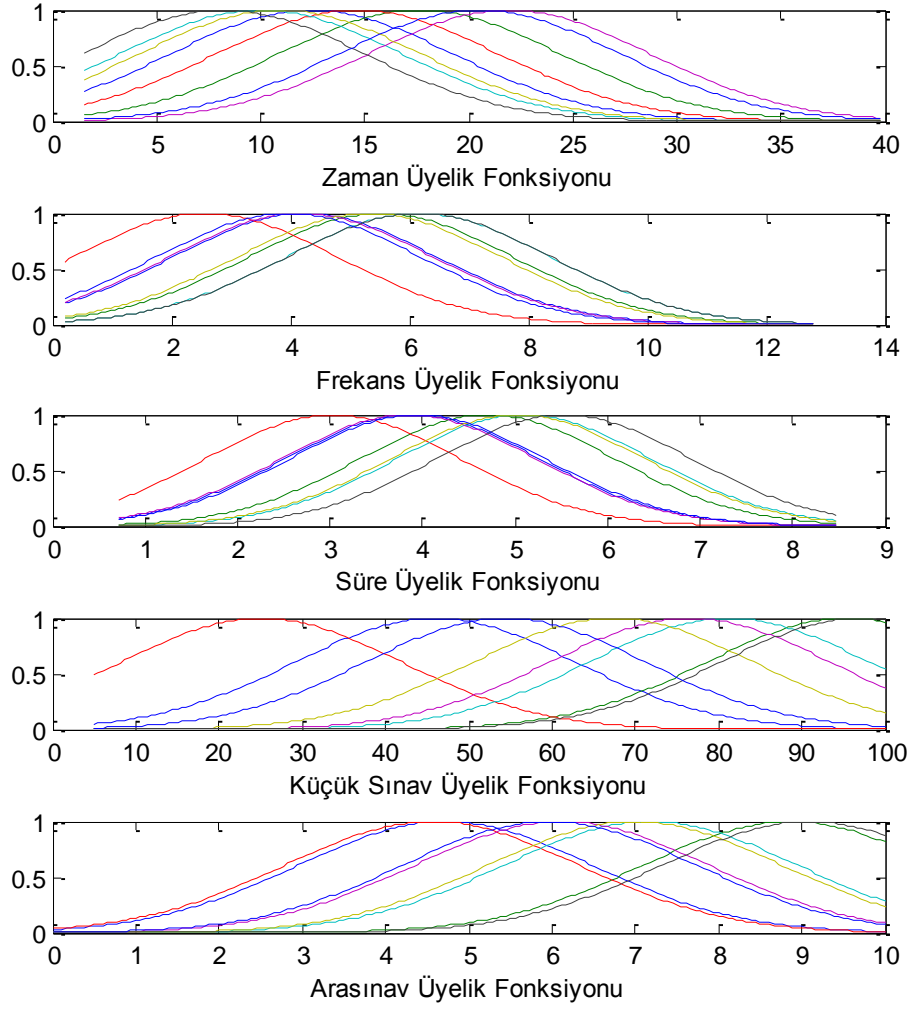
4.4.2.1 K-Ortalamlar Kümeleme Yöntemi

K-Ortalamlar Kümeleme metoduna göre bulunan küme merkezleri Tablo 4.43’te verilmiştir.

Tablo 4.43: Bulanık C-Ortalamlar kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük Sınav	Arasınav
1	12,45	4,24	3,88	45,97	4,72
2	17,63	5,52	4,70	95,44	8,86
3	14,63	2,57	3,03	25,08	4,56
4	10,01	6,15	5,10	81,25	7,19
5	21,93	4,18	3,84	76,56	6,12
6	10,94	5,33	5,03	67,42	6,99
7	8,21	6,16	5,53	97,09	9,09
8	20,52	3,96	3,95	54,39	5,97

Eğitim verisine göre en iyi doğruluk oranı küme sayısı 8 olan modele göre bulunmuştur. Üyelik fonksiyonlarının grafikleri şekil 4.46’da verilmiştir.



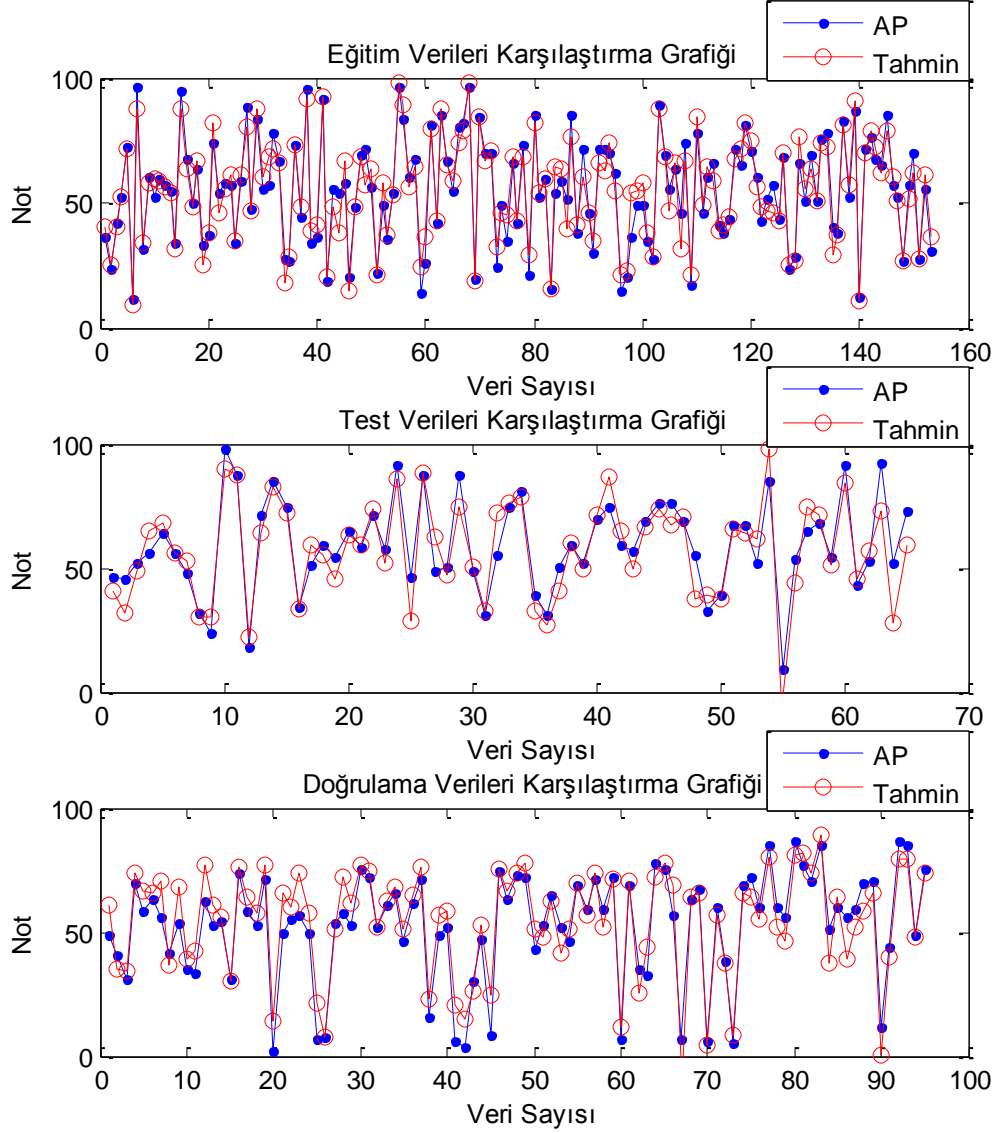
Şekil 4.46: Bulanık C-Ortalamalar kümeleme yöntemi üyelik fonksiyonları grafikleri.

K-Ortalamalar Kümeleme yönteminden elde edilen üyelik fonksiyonları sonucunda kantil yöntemi ile hesaplanan kurallara ait parametreler tablo 4.44'te verilmiştir.

Tablo 4.44 Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	145,78	-0,70	-18,89	7,80	-3,10	6,17
2	23,29	-1,47	-1,02	-3,02	0,26	3,02
3	-52,41	3,87	4,51	0,82	-0,33	1,45
4	-1,26	0,50	-1,10	-3,66	0,98	4,27
5	292,38	-2,17	9,26	-31,74	-1,09	11,43
6	-5,03	-0,86	2,51	-0,54	0,44	4,98
7	-13,01	4,60	-1,47	-7,82	0,54	4,16
8	58,99	1,83	-0,42	-0,61	-0,36	5,62

Şekil 4.47’de eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.



Şekil 4.47: K-Ortalamalar kümeleme yöntemi ile kurulan bulanık model sonuçları.

Tablo 4.45’te veri tiplerin göre doğruluk oranları görülmektedir. Öğrencilerin akademik performansları sayısal olarak ve “geçti”, “kaldı” şeklinde kategorik olarak tahmin edilmiştir.

Tablo 4.45: Verilere ait doğruluk oranları.

Küme Sayısı=8	Eğitim	Test	Doğrulama	Ortalama
Sayısal Tahmin Doğruluk Oranı (%)	0,90	0,87	0,69	0,82
Kategorik Tahmin Doğruluk Oranı	0,92	0,83	0,96	0,90

 (%)

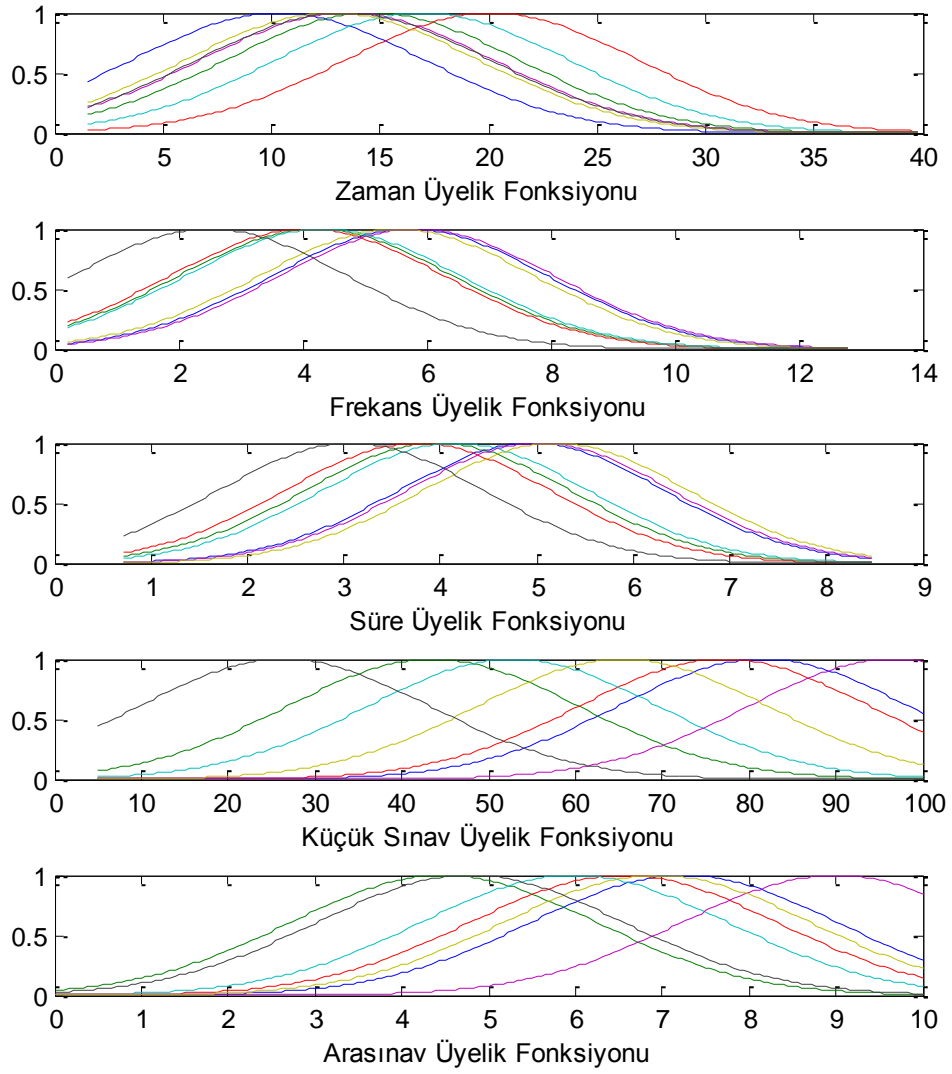
4.4.2.2 Bulanık C-Ortalamalar Kümeleme Yöntemi

Bulanık C-ortalamalar kümeleme metoduna göre bulunan küme merkezleri Tablo 4.46'da verilmiştir.

Tablo 4.46: Bulanık C-Ortalamalar kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük Sınav	Arasınav
1	10,29	5,73	4,94	81,62	7,25
2	14,65	4,21	3,97	43,60	4,48
3	20,16	4,07	3,76	77,17	6,55
4	16,95	4,33	4,16	52,81	5,99
5	13,46	5,84	5,03	96,68	8,99
6	12,77	5,52	5,20	65,44	6,94
7	13,31	2,50	3,07	26,49	4,79

Eğitim verisine göre en iyi doğruluk oranı küme sayısı 7 olan modele göre bulunmuştur. Üyelik fonksiyonlarının grafikleri şekil 4.48'de verilmiştir.



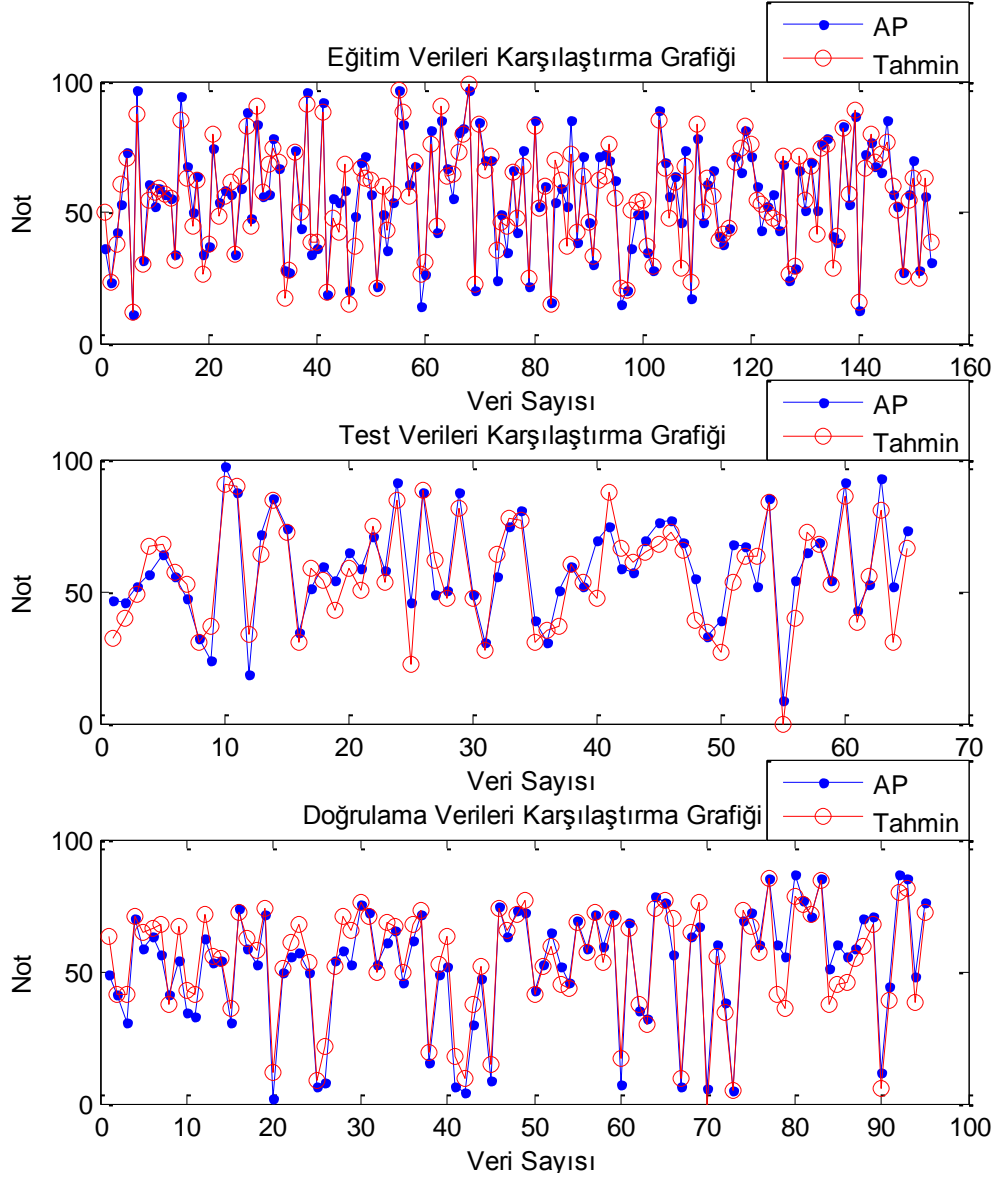
Şekil 4.48: Bulanık C-Ortalamalar kümeleme yöntemi üyelik fonksiyonları grafikleri.

Bulanık C-ortalamalar kümeleme yönteminden elde edilen üyelik fonksiyonları sonucunda kantil yöntemi ile hesaplanan kurallara ait parametreler tablo 4.47'de verilmiştir.

Tablo 4.47: Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	131,97	-1,92	-0,17	3,60	-0,31	-0,20
2	-87,93	3,06	3,89	1,24	0,91	0,89
3	58,05	0,72	-0,73	-3,56	0,08	-0,28
4	-38,44	-3,84	5,86	-4,07	1,12	13,89
5	27,04	1,11	-1,83	-0,73	-0,56	12,30
6	-51,46	3,27	-1,10	6,76	-0,42	3,33
7	-17,62	-0,48	-1,82	8,97	0,49	6,72

Şekil 4.49’da eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.



Şekil 4.49: Bulanık C-Ortalamalar kümeleme yöntemi ile kurulan bulanık model sonuçları.

Tablo 4.48’de veri tiplerin göre doğruluk oranları görülmektedir. Öğrencilerin akademik performansları sayısal olarak ve “geçti”, “kaldı” şeklinde kategorik olarak tahmin edilmiştir.

Tablo 4.48: Verilere ait doğruluk oranları.

Küme Sayısı=7	Eğitim	Test	Doğrulama	Ortalama
Sayısal Tahmin Doğruluk Oranı (%)	0,90	0,85	0,76	0,83
Kategorik Tahmin Doğruluk Oranı (%)	0,90	0,86	0,95	0,90

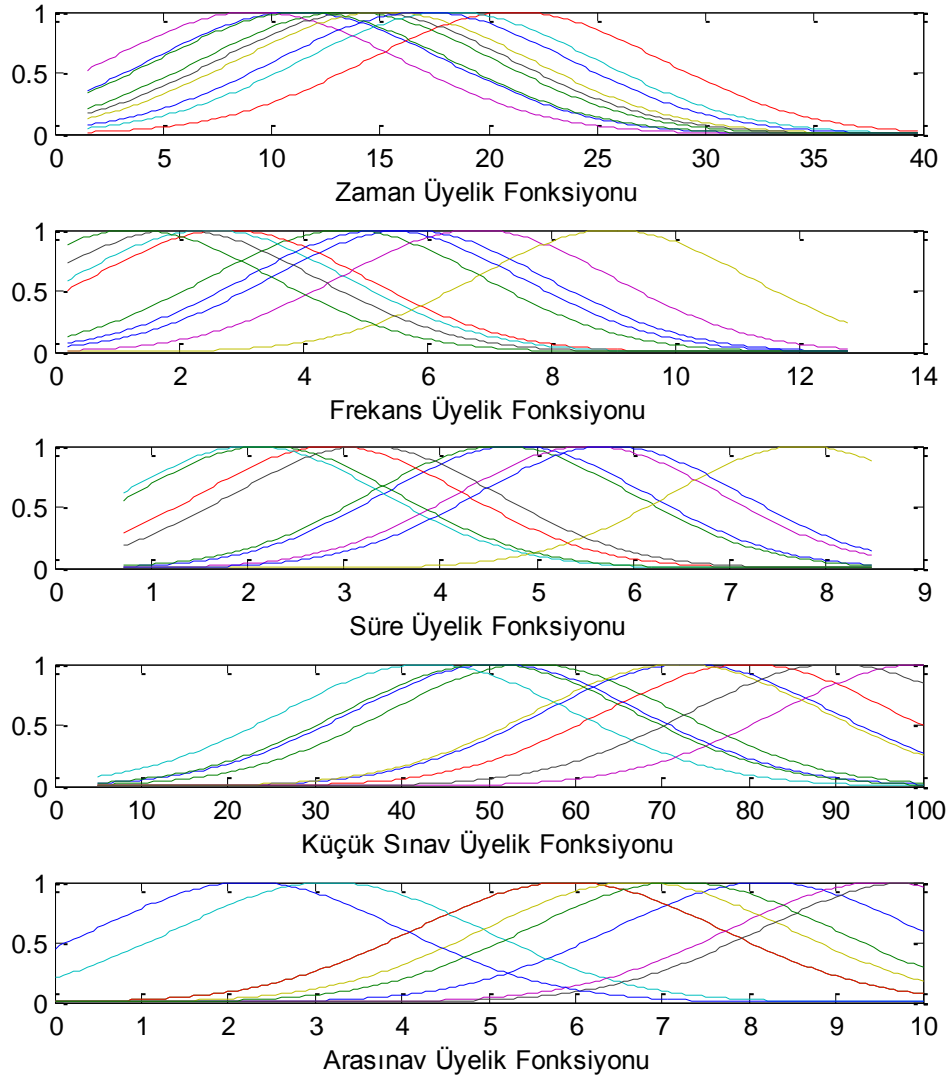
4.4.2.3 Çıkarımlı Kümeleme Yöntemi

Çıkarımlı kümeleme metoduna göre bulunan küme merkezleri Tablo 4.49'da verilmiştir.

Tablo 4.49: Çıkarımlı kümeleme yöntemi ile bulunan küme merkezleri.

Küme Merkezi	Zaman	Frekans	Süre	Küçük Sınav	Arasınav
1	11,20	5,70	4,77	73,00	8,19
2	13,50	4,70	4,63	55,00	5,92
3	21,20	2,80	2,88	80,00	5,92
4	18,20	2,50	2,05	43,00	3,16
5	9,20	6,80	5,54	100,00	9,54
6	15,20	9,00	7,76	72,00	6,71
7	14,30	2,00	3,22	90,00	9,90
8	17,00	5,30	5,75	51,00	2,24
9	11,50	1,30	2,19	50,00	7,21

Çıkarımlı kümeleme yöntemi en uygun küme sayısını 9 olarak bulmuştur. Bu modele göre üyelik fonksiyonlarının grafikleri şekil 4.50'de verilmiştir.



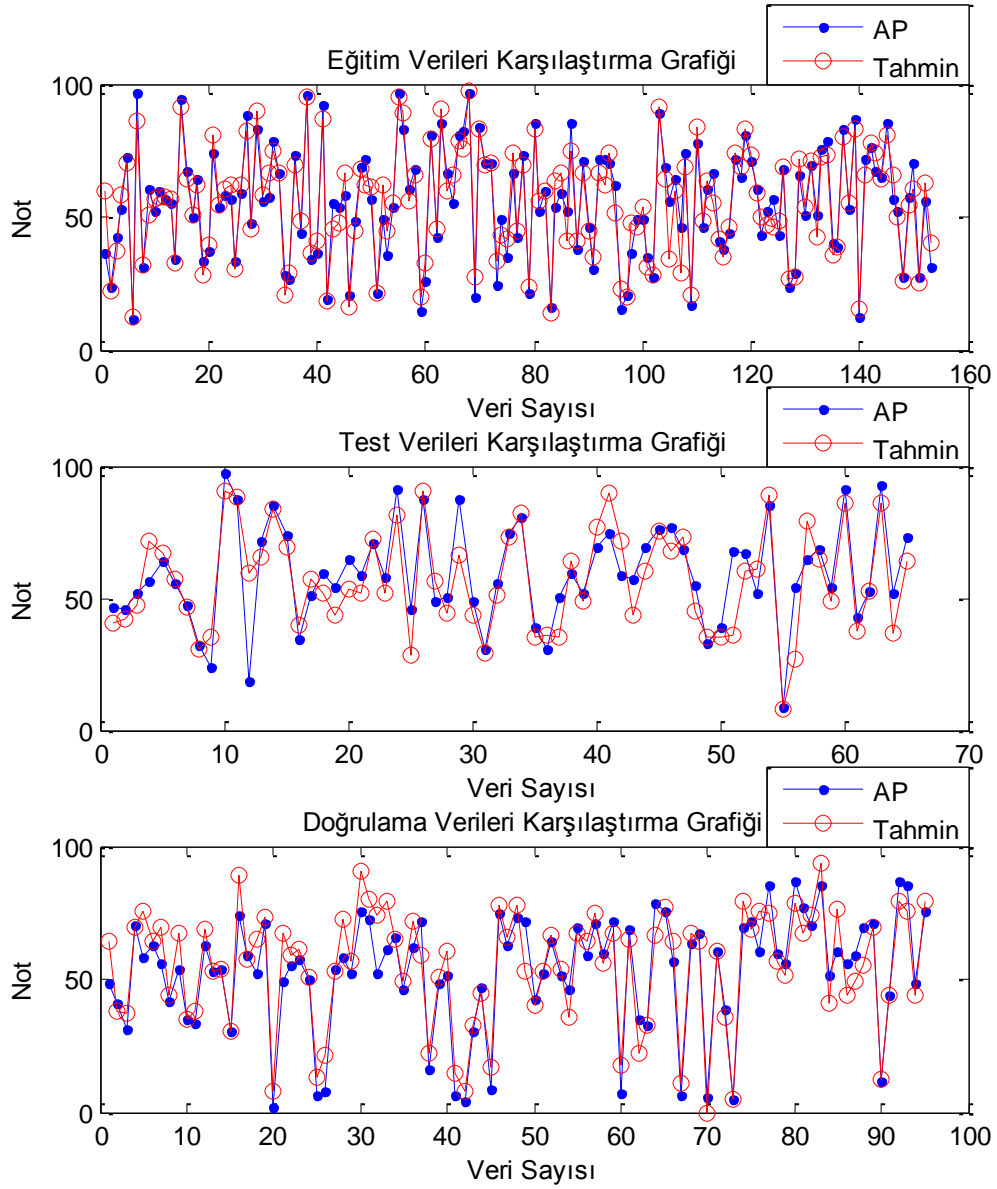
Şekil 4.50: Çıkarımlı kümeleme yöntemi üyelik fonksiyonları grafikleri.

Çıkarımlı kümeleme yönteminden elde edilen üyelik fonksiyonları sonucunda kantil yöntemi ile hesaplanan kurallara ait parametreler tablo 4.50’de verilmiştir.

Tablo 4.50: Kurallara ait parametreler.

Kural No	β_0	β_1	β_2	β_3	β_4	β_5
1	187,10	-1,04	1,33	-2,93	-1,10	-2,68
2	3,49	1,23	-0,63	3,98	0,50	-4,18
3	16,18	-0,01	-1,75	-1,25	0,65	0,99
4	-8,58	-0,10	-1,36	5,71	0,41	2,17
5	203,03	1,57	-5,79	0,22	-1,21	3,22
6	-81,43	-0,62	0,49	5,68	0,85	5,85
7	40,88	0,03	6,29	-0,27	-1,46	17,32
8	-86,74	2,82	10,37	8,43	-0,77	-0,18
9	-1,90	-0,67	0,77	-2,31	0,27	8,13

Şekil 4.51’de eğitim, test ve doğrulama verilerine ait uzaktan eğitim öğrencilerinin dönem sonunda almış oldukları Akademik Performans notu ile bulanık model ile bulunan sonuçların yer aldığı grafikler görülmektedir.



Şekil 4.51: Bulanık C-Ortalamalar kümeleme yöntemi ile kurulan bulanık model sonuçları.

Tablo 4.51’de veri tiplerin göre doğruluk oranları görülmektedir. Öğrencilerin akademik performansları sayısal olarak ve “geçti”, “kaldı” şeklinde kategorik olarak tahmin edilmiştir.

Tablo 4.51: Verilere ait doğruluk oranları.

Küme Sayısı=9	Eğitim	Test	Doğrulama	Ortalama
Sayısal Tahmin Doğruluk Oranı (%)	0,90	0,84	0,77	0,83
Kategorik Tahmin Doğruluk Oranı (%)	0,93	0,78	0,95	0,89

4.4.2.4 Sonuçların Karşılaştırılması

Tüm metotlara ait eğitim verileri ve test verileri kullanılarak bulunan doğruluk oranları tablo 4.52’de verilmiştir. Karşılaştırmalar en küçük kareler (EKK) ve kantil yöntemlerine göre aykırı veri analizi yapılmadan alınan sonuçlara göre yapılmıştır. Akademik performansın tahmininde Takagi-Sugeno Tipi Bulanık Mantık Modelini kurmak için kümeleme yöntemleri olarak K-Ortalamalar Kümeleme Yöntemi (KOKY), Bulanık C-Ortalamalar Kümeleme Yöntemi (BCOKY) ve Çıkarımlı Kümeleme Yöntemi (ÇKY) kullanılmıştır.

Tablo 4.52: Akademik performansın sayısal olarak tahminine dayalı doğruluk oranları.

Veri Tipi	EKK			Kantil		
	KOKY	BCOKY	ÇKY	KOKY	BCOKY	ÇKY
Eğitim	0,90	0,91	0,90	0,90	0,90	0,90
Test	0,87	0,87	0,85	0,87	0,85	0,84
Doğrulama	0,71	0,62	0,75	0,69	0,76	0,77
Ortalama	0,83	0,80	0,83	0,82	0,84	0,84

Sonuçlara bakıldığında sayısal olarak akademik performansın tahmininde kantil yöntemi ile doğruluk oranların daha iyi olduğu görülmüştür. Kategorik olarak öğrenciler sadece geçti kaldı şeklinde sınıflandırıldığında doğruluk oranları Tablo 4.53’te verilmiştir.

Tablo 4.53: Akademik performansın “geçti”-“kaldı” tahminine dayalı doğruluk oranları.

Veri Tipi	EKK			Kantil		
	KOKY	BCOKY	ÇKY	KOKY	BCOKY	ÇKY
Eğitim	0,90	0,91	0,91	0,92	0,90	0,93
Test	0,84	0,84	0,79	0,83	0,86	0,78
Doğrulama	0,94	0,93	0,95	0,96	0,95	0,95
Ortalama	0,89	0,89	0,88	0,90	0,90	0,89

Kategorik olarak, öğrencilerin akademik performansları EKK ve Kantil yöntemlerine göre tahmin edildiğinde kantil yöntemi ile bulunan sonuçların doğruluk oranlarının daha iyi olduğu görülmüştür.

5.TARTIŞMA VE SONUÇ

Teknolojinin her geçen gün baş döndürücü bir hızla gelişmesi, hayatımızda daha önce var olmayan, yeni alanlar oluşturduğu gibi, varolan alanlar üzerinde de olumlu katkıları göz ardı edilemez. Bu alanlardan biri hiç şüphesiz eğitimidir. Teknoloji ile birlikte geleneksel eğitimden uzaktan eğitime doğru bir yöneliş olmuştur ve bu yönelişe ait ivme, her geçen gün artmaktadır. Uzaktan eğitimin avantajları ile birlikte dezavantajları da bulunmaktadır. Bu dezavantajlardan biri de öğrencilerin performanslarının geleneksel eğitimde olduğu gibi gözlemlenememesidir. Bu çalışma ile, uzaktan eğitimde geleneksel eğitimdeki gibi gözlemlenememesi için, onun yerine ÖYS içerisinde tutulan loglar kullanılarak matematikse gözlem modelleri oluşturulmuştur. Bu çalışmanın eğitim bilimine, özellikle uzaktan eğitimde eğitmen ve yöneticilere büyük katkı sağlayacağı düşünülmektedir.

Pazarlamada müşteri profili analizinde RFM (Recency, Frequency, Monetary) metodu sıklıkla kullanılmaktadır. Kısaca RFM metodu ile bir müşteri en son ne zaman işlem yaptı, ne sıklıkla işlem yaptı ve ne kadar harcadı sorularına puanlar verilerek müşteri profili çıkarılır.

Bu çalışmada da, bir öğrenci ÖYS'ne “en son ne zaman girdi?”, “ne sıklıkla girdi?” ve “sistemde ne kadar süre kaldı?” sorularının cevapları veritabanından alınarak bir veriseti haline getirilmiştir. Yine bu çalışmamızda, öğrencilerin ÖYS içerisindeki davranışları belirsizlikler içerdiğinden, belirsizliklere karşı en iyi sonucu veren modellerden bulanık mantık metodu ile matematiksel model kurulmuştur. Klasik bulanık mantık ve uzman görüşüne dayanan bulanık modeller ile sonuçlar bulundu. Klasik bulanık mantık metodu ile bulunan doğruluk oranı yaklaşık %65 iken, uzman görüşüne dayalı bulanık mantık ile bu doğruluk oranı yaklaşık %74 bulunmuştur. Bulanık mantık modellemede en önemli iki bileşen üyelik fonksiyonlarının aralıklarının tespiti ve kuralların oluşturulmasıdır. Bu nedenle bulanık mantığın üyelik fonksiyonlarının aralıklarının optimizasyonu için genetik algoritma kullanılarak hibrit bir model oluşturulmuştur. Ayrıca bu model için daha fazla sayıda örnek sisteme katıldı ve modelin girdilerine, tahmin oranlarının artması için iki girdi daha ilave edildirmiştir. Bu girdiler 4. haftada

çevrimiçi, çoktan seçmeli olarak yapılan kısa sınav ve 8.haftada sınıf ortamında yapılan arasınnav sonuçlarıdır. Genetik algoritma ile melez olarak kurulan bulanık model (Gene-bulanık mantık) ile bulunan sonuçlarda doğruluk oranı yaklaşık %81 olarak bulunmuştur.

Bulanık Mantık'da çıkarım sistemi genellikle Mamdani ve Takagi-Sugeno tipi olarak iki şekilde yapılır. Bu çalışmada hem Mamdani ile hem de Takagi-Sugeno tipi bulanık mantık çıkarım modeli ile sonuçlar bulunmuştur. Mamdani çıkarım sistemi ile kurulan bulanık mantık bazlı modellere bakıldığında; Klasik Bulanık Mantık, Uzman Görüşüne Dayalı Bulanık Mantık ve üyelik fonksiyon aralıklarının optimizasyonuna dayanan Gene-Bulanık Mantık Modelleridir. Bu modellerle uzaktan eğitimdeki bir öğrencinin akademik performansı %65-%81 aralığında tahmin edilmiştir. En iyi sonucu gene-bulanık mantık melez modeli sağlamıştır.

Gene-bulanık modele alternatif olarak kümeleme metotları kullanılarak melez bulanık modeller oluşturulmuştur. Bu modeller Takagi-Sugeno tipi bulanık çıkarım sistemleri üzerinde kurulmuştur. Böylece kümeleme yöntemleri ile küme merkezleri elde edilmiş, elde edilen küme merkezleri ile üyelik fonksiyonları oluşturulmuş, oluşturulan üyelik fonksiyonları ile kurallar kümesi meydana gelmiştir. Kurallara ait parametreler sıklıkla kullanılan metotlarda en küçük kareler yöntemi ile bulunmaktadır. En küçük kareler yönteminin dezavantajı olarak aykırı veri setinden etkilendiği görülmüştür. Bundan dolayı modellere ait kural parametrelerinin bulunmasında hem en küçük kareler yöntemi kullanılmış ve alternatif olarak aykırı verilerden çok etkilenmeyen parametre tahmin yöntemlerinden kantil yöntemi kullanılmıştır.

En küçük kareler yöntemi ile parametrelerin bulunacağı uygulamada öncelikle verilere önce aykırı veri analizi yapılmış ardından verilerin normal dağılımın sağlamak için bazı dönüştürücü araçlar kullanılmıştır. Kümeleme yöntemleri olarak Çıkarımlı Kümeleme, K-Ortalamlar, ve Bulanık C-Ortalamlar kullanılmıştır. Modelde eğitim verisi, test verisi ve doğrulama verisi olmak üzere 3 farklı veri tipi kullanılmıştır. Ortalamada en yüksek doğruluk oranı Bulanık c-ortalamlar kümeleme yöntemi ile bulunmuştur. Bulunan doğruluk oranı %87.70 dir. Bu oranın anlamı, uzaktan eğitimdeki öğrencilerin yıl sonu akademik performansları sayısal olarak tahmin edildiğinde yaklaşık %12 hata payı vardır. Ayrıca yapılan bu tahmin 8 haftalık veri ile yapılmıştır. Sayısal tahminle

birlikte, kategorik olarak “geçti”, “kaldı” şeklinde de tahmin yapılmıştır. En iyi sonuç %92,28’dir ve çıkarımlı kümeleme yöntemi ile bulunmuştur.

Kantil metodu ile aykırı veri analizi yapılmadan sonuçlar bulunmuş ve aykırı veri analizi yapılmadan en küçük kareler yöntemi ile bulunan sonuçlar karşılaştırılmıştır. Uzaktan eğitim öğrencilerinin akademik performanslarının sayısal olarak tahmininde kantil yöntemi ile bulunan sonuçlar ortalama bazında %1-%4 arasında daha yüksek bulunmuştur. Kategorik bazda yapılan tahminlerde bu oran yaklaşık %1 oranında daha yüksek çıkmıştır. Genel olarak değerlendirildiğinde aykırı veri analizi ile yapılan tahminlerdeki hata oranları, aykırı veri analizi yapılmadan yapılan tahminlere göre daha düşüktür.

Bu çalışma Temel Bilgisayar Bilimleri dersine online kayıt yaptıran öğrencilerin verileri ile yapılmıştır. Gelecekte farklı derslerle alakalı veriler toplanarak sonuçlar bulunup karşılaştırmalar yapılabilir. Bu çalışmada öğrencilerin demografik özellikleri modele katılmamıştır. Gelecekteki çalışmalara öğrencilerin demografik özellikleri de katılarak model kurulabilir. Bulunan sonuçlar şimdiki sonuçlarla karşılaştırılabilir. Ayrıca Gelecek çalışmada bulanık modele ait kuralların parametrelerinin tahmininde kullanılan en küçük kareler yöntemi yerine alternatif olarak sadece kantil metodu ile parametre tahmini yapılmıştır, gelecekte başka yöntemler kullanılarak sonuçlarda iyileştirilme yapılabilir.

KAYNAKLAR

- [1] Şahin, M. C., 2005, *İnternet Tabanlı Uzaktan Eğitimin Etkililiği: Bir Meta Analiz Çalışması*, Yüksek Lisans Tezi, Sosyal Bilimler Enstitüsü, Çukurova Üniversitesi
- [2] G. Ekici, «Uzaktan eğitim ortamlarının seçiminde öğrencilerin öğrenme stillerinin önemi,» *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, cilt 24, pp. 48-55, 2003.
- [3] Çiftçi, S., 2006, *Uzaktan Eğitimde Öğrencilerin Ders Çalışma Etkinliklerinin Log Verileri Analiz Edilerek İncelenmesi*, Yüksek Lisans Tezi, Eğitim Bilimleri Enstitüsü, Gazi Üniversitesi.
- [4] Chou S.-W. , Liu C.-H., 2005, Learning Effectiveness İn A Web-Based Virtual Learning Environment: A Learner Control Perspective, *Journal of Computer Assisted Learning*, pp. 65-76.
- [5] EducationDynamics, «EducationDynamics Survey: Many Online Learners Never Seek Help Before Dropping Out,» 01 06 2013. [Çevrimiçi]. Available: <http://www.educationdynamics.com/News/Press-Releases/2009-Press-Releases/EducationDynamics-Survey--Many-Online-Learners-Nev.aspx>.
- [6] Zafra, A., Ventura, S., 2009, Predicting Student Grades in Learning Management Systems with Multiple Instance Genetic Programming,» *Educational Data Mining* , pp. 307-314.
- [7] Kotsiantis, S., Pierrakeas, C., Pintelas,P., 2004, Predicting Students' Performance in Distance Learning Using Machine Learning Techniques, *Applied Artificial Intelligence*, Cilt 18, No. 5, pp. 411-426.
- [8] Rena-Reyes,C. A., Sipper, M.,1999, A Fuzzy- Genetic Approach to Breast Cancer Diagnosis, *Artificial Intelligence in Medicine*, pp. 131-155.
- [9] Aydın, C. Ç., Biroğul, S., 2008, E-öğrenmede Açık Kaynak Kodlu Öğretim Yönetim Sistemleri, *Bilişim Teknolojileri Dergisi*, Cilt 1, No. 2, pp. 31-36.
- [10] Wu, T.-P., Chen, S.-M.,1999, A New Method for Constructing Membership Functions and Fuzzy Rules from Training Examples, *IEEE Transactions on Systems*, Cilt 29, No. 1, pp. 25-40.
- [11] Taylan, O., Karagozoglu, B.,2009, An Adaptive Neuro-Fuzzy Model for Prediction of Student's Academic Performance, *Computers & Industrial Engineering*, pp. 732-741.

- [12] Huang, Y.-P., Huang, C.-H., 1997, Real-Valued Genetic Algorithm for Fuzzy grey Prediction System, *Elsevier Fuzzy Sets and Systems*, pp. 265-276.
- [13] Yusof, N., Ariffin, N., Noraniah, M. Z., Yassin, M., Samsuri, P., 2009, Evaluation of Students' Performance and Learning Efficiency Based of ANFIS, *International Conference of Soft Computing and Pattern Recognition*, 4-7 Dec, Malacca, 460-465.
- [14] Cho, H.-J., Cho, K.-B. ve Wang, B. H., 1997, Fuzzy-PID Hybrid Control Automatic Rule Generation Using Genetic Algorithm, *Information and Management Sciences*, pp. 305-316.
- [15] Chen, S.-M., Tsai, F.-M., 2005, A New Method to Construct Membership Functions and Generate Fuzzy Rules from Training Instances, *Information and Management Sciences*, Cilt 16, No. 2, pp. 47-72.
- [16] Kaptanoğlu, D., Özok, A. F., 2008, Akademik Performans Değerlendirmesi için Bir Bulanık Model, *İtü Dergisi*, pp. 10-14.
- [17] Arslan, A., Kaya, M., 2001, Determination of Fuzzy Logic Membership Functions Using Genetic Algorithm, *Elsevier Fuzzy Sets and Systems*, Cilt 118, No. 2, pp. 297-306.
- [18] Herrera, F., 2008, Genetic Fuzzy Systems: Taxonomy, Current Research Trends Prospects, *Evolutionary Intelligence*, Cilt 1, pp. 27-46.
- [19] Chiu S., 1994, Fuzzy Model Identification Based on Cluster Estimation, *Journal of Intelligent and Fuzzy Systems*, Cilt 2, pp. 209-219.
- [20] Priyono A., Muhammad Ridwan A. J. A. ve Atiq R., 2005, Generation of Fuzzy Rules with Subtractive Clustering, *Jurnal Teknologi*, Cilt 43, pp. 143-153, 2005.
- [21] Kim M.-S. ve Kong S.-G., 1997, Fuzzy Rule Generation using the clustering Algorithm for Cooperative Behavior of Small Robots, *Proceedings of the 2nd Asian Control Conference*, July 22-25, Seoul, pp. 623-626.
- [22] Moertini, V. S., 2002, Introduction to Five Data Clustering Algorithms, *Journal of Integral*, Cilt 7, No. 2 pp. 87-96.
- [23] Torun, Y., Tohumoğlu, G., 2009, A New Fuzzy Classifier Based On Simulated Annealing And Subtractive Clustering, *Signal Processing and Communications Applications Conference*, 9-11 April, Antalya, 460-463.
- [24] Lykourantzou, I., Giannoukos, I., Nikolopoulos, V., Mpardis, G., Loumos, V., 2009, Dropout Prediction in E-Learning Courses Through the Combination of Machine Learning Techniques, *Computers & Education*, Cilt 53, No.3, pp. 950-965, 2009.
- [25] Vandamme, J. -P., Meskens, N., Superby, J. -F., 2007, Predicting Academic Performance by Data Mining Methods, *Education Economics*, Cilt 15, No. 4, pp.

405-419.

- [26] Kaya, M. E., 2006, *Credit Risk Modeling Using Machine Learning Techniques*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Bogazici Üniversitesi.
- [27] Tuğ, E., 2005, *Genetik Algoritmalar İle Tıbbi Veri Madenciliği*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Selçuk Üniversitesi.
- [28] Uzun, Y., 2005, *Tıbbi Veriler Üzerinde Makine Öğrenme Algoritmaları ve Bulanık Mantık ile Kurallar Öğrenme*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Selçuk Üniversitesi.
- [29] Deniz, E., 2006, *Bulanık Mantık Tabanlı Tahmin Modeli ve Uygulaması*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Muğla Üniversitesi.
- [30] Aydın, S., 2007, *Veri Madenciliği Ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama*, Doktora Tezi, Fen Bilimleri Enstitüsü, Anadolu Üniversitesi.
- [31] Ekmekçi, D., 2008, *Bulanık Mantıklı Sürücü Modeli İle Hız Tahmini Ve En Kısa Yol Belirleme*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Sakarya Üniversitesi.
- [32] Kahramanlı, H., 2008, *Hibrit Bulanık Sinir Ağını Kullanarak Bir Sınıflandırma Ve Kural Çıkarma Sisteminin Geliştirilmesi*, Doktora Tezi, Fen Bilimleri Enstitüsü, Selçuk Üniversitesi.
- [33] Afşar, S., 2008, *Yapay Sinir Ağları Ve Bulanık Mantık Yöntemleri Kullanılarak Tava Buharlaştırma Tahmini Yapılması*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Erciyes Üniversitesi.
- [34] Amasyalı, F., 2008, *Yeni Makine Öğrenmesi Metotları ve İlaç Tasarımına Uygulamaları*, Doktora Tezi, Fen Bilimleri Enstitüsü, Yıldız Teknik Üniversitesi.
- [35] Coşkun, C., 2010, *Veri Madenciliği Algoritmaları Karşılaştırılması*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Diyarbakır: Dicle Üniversitesi, 2010.
- [36] Akbilgiç, O., 2011, *Hibrit Radyal Tabanlı Fonksiyon Ağları İle Değişken Seçimi Ve Modelleme: Menkul Kıymet Yatırım Kararlarına İlişkin Bir Uygulama*, Doktora Tezi, İstanbul Üniversitesi.
- [37] Çelebi, M., 2011, *Genetik Algoritma Tabanlı Optimal Adaptif Fuzzy PID Kontrolcü Tasarımı*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Yıldız Teknik Üniversitesi.
- [38] Işık, Y., 2006, *Genetik Algoritma Tabanlı Bulanık Kontrolün Uçuş Kontrol Sistem Tasarımına Uygulanması*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Anadolu Üniversitesi.
- [39] Altay, A., 2007, *Genetik Algoritma Ve Bir Uygulama*, Yüksek Lisans Tezi, Fen

Bilimleri Enstitüsü, İstanbul Teknik Üniversitesi.

- [40] Gündoğdu, E., 2007, *Veri Madenciliğinde Genetik Algoritmalar*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Kocaeli Üniversitesi.
- [41] Avcı, U., 2006, *Bulanık Kümeleme Algoritmalarının Karşılaştırılmalı Analizi ve Bilgisayar Uygulamaları*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Ege Üniversitesi.
- [42] Erilli, N. A., 2009, *Kümeleme Analizinde Bulanık Yaklaşım Algoritmaları ve Bir Uygulama*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Ondokuz Mayıs Üniversitesi.
- [43] Güler, N., 2006, *Bulanık Kümeleme Analizi ve Bulanık Modelleme Uygulamaları*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Muğla Üniversitesi.
- [44] Gün, A. M., 2011, *Bulanık Kümeleme Analiziyle Görüntüdeki Yüzün Tanınması*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Haliç Üniversitesi.
- [45] Ulutagay, G., 2004, *Bulanık c-Ortalamlar Kümeleme Analizi ve Uygulamaları*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Ege Üniversitesi.
- [46] Gül, G., 2013, *RFM Analizi*, <http://www.kurumsalzeka.com/?p=22>, [Ziyaret Tarihi: 01 01 2013].
- [47] Henry, C., Chai, C., 2005, Online Auction Customer Segmentation Using a Neural Network, *International Journal of Application Science*, Cilt 3, No.2, pp. 101–9.
- [48] Elmas, Ç., Doğan, N., Biroğul, S., Koç, M. S., 2008, Moodle Eğitim Yönetim Sistemi İle Örnek Bir Dersin Uzaktan Eğitim Uygulaması, *Bilişim Teknolojileri Dergisi*, Cilt 1, No. 2, pp. 53-62.
- [49] Ogor, E. N., 2007, Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques, *Fourth Congress of Electronics, Robotics and Automotive Mechanics*, 25-28 Sept., Morelos, pp. 354-359.
- [50] Alpaydın, E., 2004, *Introduction to Machine Learning*, MIT Press, London.
- [51] M, Sarı., Murat, Y.Ş., Kırbalı, M., 2005, *Fuzzy Modelling Approach and Application*, <http://www3.dpu.edu.tr/akademik/fbe/wp-content/uploads/2005/12/09.pdf>. [Ziyaret tarihi: 05 Mayıs 2013].
- [52] Türkşen, İ. B., 1985, Bulanık Kümeler Kuramı ve Uygulamaları, *Yöneylem Araştırma Dergisi*, 4(1), pp. 1-15, 1985.
- [53] Baykal, N., Beyan, T., 2004, *Bulanık Mantık Uzman Sistemler ve Denetleyiciler*, Bıçaklar Kitapevi, Ankara.
- [54] Çiftçi, H., 2002, *Fuzzy Logic Function Approximation for Some Mathematical*

Functions, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Osman Gazi Üniversitesi.

- [55] Klir, G. J., Yuan, B., 2004, *Fuzzy Sets and Fuzzy Logic-Theory and Applications*, Prentice Hall, New Jersey
- [56] Osofisan, P. B., 2013, *Fuzzy Logic Control of the Syrup Mixing Process in Beverage Production*, http://ljs.academicdirect.org/A11/093_108.htm, [Ziyaret tarihi:15 Nisan 2013.
- [57] Yalçın, Ş., 2007, *Bulanık Mantık Metodolojisi ile Taşınmaz Değerlenme Modelinin Geliştirilmesi ve Uygulaması:Konya Örneği*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Selçuk Üniversitesi.
- [58] Ren, Q., 2006, *Type-2 Takagi-Sugeno-Kang Fuzzy Logic Modeling Using Subtractive Clustering*, Thesis (Ph.d), University of Montreal.
- [59] ALMA, Ö., VUPA, Ö., 2008, Regresyon Analizinde Kullanılan En Küçük Kareler ve En Küçük Medyan Kareler Yöntemlerinin Karşılaştırılması, *Sdü Fen Edebiyat Fakültesi Fen Dergisi (E-Dergi)*, 3(2), 219-229.
- [60] Altındağ, İ., 2010, *Quantile Regresyon ve Bir Uygulama*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Selçuk Üniversitesi.
- [61] Keskin, B., 2012, *Sağlam Bir Çıkarsama Yöntemi:Kantil Regresyon*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Muğla Sıtkı Koçman Üniversitesi.
- [62] Çolak, Ö. F., Öztürkler, H., Tokathoğlu, İ., 2008, Türkiye'de Tüketim Fonksiyonunun Dilim Regresyon Yöntemi İle Tahmini, *İktisat İşletme Finans*, 23(268), 63-93.
- [63] Haupt, R. L., Haupt, S. E., 2004, *Practical Genetic Algorithms*, John Wiley & Sons New Jersey.
- [64] Silahtaroglu, G., 2008, *Kavram ve Algoritmalarıyla Veri Madenciliği*, Papatya, İstanbul.
- [65] Gan, G., Ma,C., Wu, J.,2007, *Data Clustering*, American Statical Association and The Society for Industrial and Applied Mathematics, Alexandria.
- [66] Berberoğlu, B., 2011, 2008 Global Krizinin Türkiye Ve Avrupa Birliği'ndeki Etkilerinin Kümeleme Analizi İle İncelenmesi, *Anadolu üniversitesi Sosyal Bilimler Dergisi*, 11(1),105-130.
- [67] Grimm, L. G., Yarnold P., 1994, *Reading and Understanding More Multivariate Statistics*, American Psychological Association, Washington.
- [68] Hair, J. F., Anderson, R. E., Tatham, R. L., Black, W. C., 1995, *Multivariate Data Analysis*, Prentice- Hall Inc, Upple Saddle Niver,NJ.

- [69] Tatlıdil, H., 2002, *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Akademi Matbaası, Ankara.
- [70] Büyüköztürk, Ş., 2009, *Veri Analizi El Kitabı*, Pegem Yayıncılık, Ankara.
- [71] Albayrak, A. S., 2006, *Uygulamalı Çok Değişkenli İstatistik Teknikleri*, Asil yayın Dağıtım, Ankara.
- [72] Koğar, H., 2010, *Farklı Örneklem Büyüklüklerinde Uç Değerlerle Baş Etme Yöntemlerinin Puanların Geçerlilik ve Güvenirlik Kanıtlarına Etkisi*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Ankara Üniversitesi.
- [73] Vural, A., 2007, *Aykırı Değerlerin Regresyon Modellerine Etkileri ve Sağlam Kestiriciler*, Yüksek Lisans Tezi, Sosyal Bilimler Enstitüsü, Marmara Üniversitesi.
- [74] Ulutağay, G., 2004, *Bulanık C-Ortalamalar Küme Analizi ve Bir Uygulama*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Ege Üniversitesi.
- [75] Işık, M., Çamurcu A., 2007, K-Means, K-Medoids ve Bulanık C-Means Algoritmalarının Uygulamalı Olarak Performanslarının Tesbiti, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(11), 31-45.
- [76] Hekim, M., Orhan, U., Durusoy, G., 2006, Veri Kümeleme Üzerine Hibrit Yaklaşım, *ELECO 2006 Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu*, Bursa, 41-45.
- [77] Chiu S. L., 1997, *Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification*, Fuzzy Information Engineering: A Guided Tour of Applications, John Wiley and Sons.
- [78] Moore, D. S., McCabe, G. P., Craig, B. A., 2009, *Introduction to the Practice of Statistics*, W. H. Freeman and Company, New York.
- [79] Gümüş, E., 2013, *Çok Değişkenli Veride Aykırı Gözlemlerin Tespiti için En Küçük Kovaryans Determinantına Dayalı Test İstatistiğinin 1. Tip Hata Bakımındaki Sağlamlığının İrdelenmesi*, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Gazi Üniversitesi.
- [80] Jang, R., Sun, T., Mizutani., 1997, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall.
- [81] ALPAR, R., 2012, *Uygulamalı İstatistik ve Geçerlilik Güvenirlik*, Detay Yayıncılık Ankara.
- [82] Özdamar, K., 2004, *Paket Programlar ile İstatistiksel Veri Analizi*, Kaan Kitapevi, Eskişehir

EKLER

EK 1. Genetik Algoritma İle Üyelik Fonksiyon Aralıklarının Optimizasyonu Matlab

Kodu

```

clear all
psize=10
last=10;
psize=10;
mo=0.15;
[RFMnum,RFMtxt,RFMraw]=xlsread('fuzzyrfm2013.xlsx');
SRFM11=zeros(10,46);
[v bb]=size(RFMnum);
rs=ceil(v*0.70);
test1=v-rs;
raa=randperm(v);
for i=1:rs
    train1(i,:)=RFMnum(raa(i),:);
end
for i=1:test1
    test11(i,:)=RFMnum(raa(i+rs),:);
end

[Rulenum,ruletxt,ruleraw]=xlsread('fuzzyrule3.xlsx');
RFM=train1(:,1:3);

fis11=readfis('rfm1285liyenii2');
fis11=addrule(fis11,Rulenum);
fisexpert=readfis('rfmfuzzy5liexpert');
fisklasik=readfis('rfmfuzzy5li');
girisexpert=getfis(fisexpert,'inmfparams');
girisklasik=getfis(fisklasik,'inmfparams');

r1=sort((round(10+(110-10).*(rand(psize-2,3))))),2); %satırları soldan
sağa sıralanmış 3 adet a ile b arasında rastgele sayı üretir
f1=sort((round(10+(110-10).*(rand(psize-2,3))))),2); %satırları soldan
sağa sıralanmış 3 adet a ile b arasında rastgele sayı üretir
m1=sort((round(10+(110-10).*(rand(psize-2,3))))),2); %satırları soldan
sağa sıralanmış 3 adet a ile b arasında rastgele sayı üretir
zerol=zeros(size(r1));
ones1=128*ones(size(r1));
rr=[zerol(:,1) zerol(:,1) r1(:,1) zerol(:,1) r1(:,1) r1(:,2) r1(:,1)
r1(:,2) r1(:,3) r1(:,2) r1(:,3) ones1(:,1) r1(:,3) ones1(:,1)
ones1(:,1)];
ff=[zerol(:,1) zerol(:,1) f1(:,1) zerol(:,1) f1(:,1) f1(:,2) f1(:,1)
f1(:,2) f1(:,3) f1(:,2) f1(:,3) ones1(:,1) f1(:,3) ones1(:,1)
ones1(:,1)];
mm=[zerol(:,1) zerol(:,1) m1(:,1) zerol(:,1) m1(:,1) m1(:,2) m1(:,1)
m1(:,2) m1(:,3) m1(:,2) m1(:,3) ones1(:,1) m1(:,3) ones1(:,1)
ones1(:,1)];

for it=1:100

```

```

    if it==1
    for i=1:psize
        if i==9
            giris1=girisexpert;
            fis11.input(1).mf(1).params=[giris1(1,1) giris1(1,2) giris1(1,3)];
            fis11.input(1).mf(1).params=[giris1(2,1) giris1(2,2) giris1(2,3)];
            fis11.input(1).mf(1).params=[giris1(3,1) giris1(3,2) giris1(3,3)];
            fis11.input(1).mf(1).params=[giris1(4,1) giris1(4,2) giris1(4,3)];
            fis11.input(1).mf(1).params=[giris1(5,1) giris1(5,2) giris1(5,3)];

            fis11.input(2).mf(1).params=[giris1(6,1) giris1(6,2) giris1(6,3)];
            fis11.input(2).mf(1).params=[giris1(7,1) giris1(7,2) giris1(7,3)];
            fis11.input(2).mf(1).params=[giris1(8,1) giris1(8,2) giris1(8,3)];
            fis11.input(2).mf(1).params=[giris1(9,1) giris1(9,2) giris1(9,3)];
            fis11.input(2).mf(1).params=[giris1(10,1) giris1(10,2) giris1(10,3)];

            fis11.input(3).mf(1).params=[giris1(11,1) giris1(11,2) giris1(11,3)];
            fis11.input(3).mf(1).params=[giris1(12,1) giris1(12,2) giris1(12,3)];
            fis11.input(3).mf(1).params=[giris1(13,1) giris1(13,2) giris1(13,3)];
            fis11.input(3).mf(1).params=[giris1(14,1) giris1(14,2) giris1(14,3)];
            fis11.input(3).mf(1).params=[giris1(15,1) giris1(15,2) giris1(15,3)];
            rfmrandz=giris1(:)';
            rfmrand(9,:)=rfmrandz(1:45);
            elseif i==10
                giris2=girisklasik;
                fis11.input(1).mf(1).params=[giris2(1,1) giris2(1,2) giris2(1,3)];
                fis11.input(1).mf(1).params=[giris2(2,1) giris2(2,2) giris2(2,3)];
                fis11.input(1).mf(1).params=[giris2(3,1) giris2(3,2) giris2(3,3)];
                fis11.input(1).mf(1).params=[giris2(4,1) giris2(4,2) giris2(4,3)];
                fis11.input(1).mf(1).params=[giris2(5,1) giris2(5,2) giris2(5,3)];

                fis11.input(2).mf(1).params=[giris2(6,1) giris2(6,2) giris2(6,3)];
                fis11.input(2).mf(1).params=[giris2(7,1) giris2(7,2) giris2(7,3)];
                fis11.input(2).mf(1).params=[giris2(8,1) giris2(8,2) giris2(8,3)];
                fis11.input(2).mf(1).params=[giris2(9,1) giris2(9,2) giris2(9,3)];
                fis11.input(2).mf(1).params=[giris2(10,1) giris2(10,2) giris2(10,3)];

                fis11.input(3).mf(1).params=[giris2(11,1) giris2(11,2) giris2(11,3)];
                fis11.input(3).mf(1).params=[giris2(12,1) giris2(12,2) giris2(12,3)];
                fis11.input(3).mf(1).params=[giris2(13,1) giris2(13,2) giris2(13,3)];
                fis11.input(3).mf(1).params=[giris2(14,1) giris2(14,2) giris2(14,3)];
                fis11.input(3).mf(1).params=[giris2(15,1) giris2(15,2) giris2(15,3)];
                rfmrand1(1,:)=giris2(:)';
                rfmrand(10,:)=rfmrand1(1:45);
            else

                fis11.input(1).mf(1).params=[rr(i,1) rr(i,2) rr(i,3)];
                fis11.input(1).mf(2).params=[rr(i,4) rr(i,5) rr(i,6)];
                fis11.input(1).mf(3).params=[rr(i,7) rr(i,8) rr(i,9)];
                fis11.input(1).mf(4).params=[rr(i,10) rr(i,11) rr(i,12)];
                fis11.input(1).mf(5).params=[rr(i,13) rr(i,14) rr(i,15)];

                fis11.input(2).mf(1).params=[ff(i,1) ff(i,2) ff(i,3)];
                fis11.input(2).mf(2).params=[ff(i,4) ff(i,5) ff(i,6)];
                fis11.input(2).mf(3).params=[ff(i,7) ff(i,8) ff(i,9)];
                fis11.input(2).mf(4).params=[ff(i,10) ff(i,11) ff(i,12)];
                fis11.input(2).mf(5).params=[ff(i,13) ff(i,14) ff(i,15)];

                fis11.input(3).mf(1).params=[mm(i,1) mm(i,2) mm(i,3)];

```

```

fis11.input(3).mf(2).params=[mm(i,4) mm(i,5) mm(i,6)];
fis11.input(3).mf(3).params=[mm(i,7) mm(i,8) mm(i,9)];
fis11.input(3).mf(4).params=[mm(i,10) mm(i,11) mm(i,12)];
fis11.input(3).mf(5).params=[mm(i,13) mm(i,14) mm(i,15)];
    end

```

```

rfmrand(1:psize-2,:)= [rr ff mm];
RFM(:,(4:5))=train1(:,(4:5));
tahmin=evalfis(RFM,fis11);
gsonuc=train1(:,6);
giris=getfis(fis11,'inmfparams');%Giriş değerlerini verir
cikis=getfis(fis11,'outmfparams');
girisaralik=fis11.input(1).mf(2).params;
d=abs(tahmin-gsonuc);
c=length(d);
accuracy(i)=sum(d./gsonuc)/c

```

```

    end

```

```

rfmrand(1:psize-2,:)= [rr ff mm];

```

```

RFM(:,(4:5))=train1(:,(4:5));
tahmin=evalfis(RFM,fis11);
gsonuc=train1(:,6);
giris=getfis(fis11,'inmfparams');
cikis=getfis(fis11,'outmfparams');
girisaralik=fis11.input(1).mf(2).params;
d=abs(tahmin-gsonuc);
c=length(d);
accuracy(i)=sum(d./gsonuc)/c

```

```

    else

```

```

        for i=1:psize

```

```

fis11.input(1).mf(1).params=[rr(i,1) rr(i,2) rr(i,3)];
fis11.input(1).mf(2).params=[rr(i,4) rr(i,5) rr(i,6)];
fis11.input(1).mf(3).params=[rr(i,7) rr(i,8) rr(i,9)];
fis11.input(1).mf(4).params=[rr(i,10) rr(i,11) rr(i,12)];
fis11.input(1).mf(5).params=[rr(i,13) rr(i,14) rr(i,15)];

```

```

fis11.input(2).mf(1).params=[ff(i,1) ff(i,2) ff(i,3)];
fis11.input(2).mf(2).params=[ff(i,4) ff(i,5) ff(i,6)];
fis11.input(2).mf(3).params=[ff(i,7) ff(i,8) ff(i,9)];
fis11.input(2).mf(4).params=[ff(i,10) ff(i,11) ff(i,12)];
fis11.input(2).mf(5).params=[ff(i,13) ff(i,14) ff(i,15)];

```

```

fis11.input(3).mf(1).params=[mm(i,1) mm(i,2) mm(i,3)];
fis11.input(3).mf(2).params=[mm(i,4) mm(i,5) mm(i,6)];
fis11.input(3).mf(3).params=[mm(i,7) mm(i,8) mm(i,9)];
fis11.input(3).mf(4).params=[mm(i,10) mm(i,11) mm(i,12)];
fis11.input(3).mf(5).params=[mm(i,13) mm(i,14) mm(i,15)];
rfmrand=[rr ff mm];

```

```

RFM(:,(4:5))=train1(:,(4:5));
tahmin=evalfis(RFM,fis11);

```

```

gsonuc=train1(:,6);
cikis=getfis(fis11,'outmfparams');
girisaralik=fis11.input(1).mf(2).params;
d=abs(tahmin-gsonuc);
c=length(d);
accuracy(i)=sum(d./gsonuc)/c ;

    end
end

sort(accuracy)
[err,ind]=sort(accuracy)

for i=1:ceil(psize/2)
    popsec(i,:)=rfmrand(ind(i),:);
end

toplam=sum(1./err(1:ceil(psize/2)));

oran=round(100*(1./err(1:ceil(psize/2))./toplam));

n1=1;
m1=oran(1);

[ro co]=size(popsec);
popsecoran=zeros(sum(oran),co);
    for i=1:ceil(psize/2)
        for j=1:co
            popsecoran(n1:m1,j)=popsec(i,j);
        end
        n1=m1+1;
        if i<round(psize/2)
            m1=m1+oran(i+1);
        end
    end
    secim2(1,:)=popsec(1,:);

for i=2:ceil(psize)
    secim2(i,:)=popsecoran(ceil(1+98.*(rand(1,1))),:);

end
popsecoran1=secim2;
R1=popsecoran1(1:10,1:15);
F1=popsecoran1(1:10,16:30);
M1=popsecoran1(1:10,31:45);
for i=1:psize

fis11.input(1).mf(1).params=[R1(i,1) R1(i,2) R1(i,3)];
fis11.input(1).mf(2).params=[R1(i,4) R1(i,5) R1(i,6)];
fis11.input(1).mf(3).params=[R1(i,7) R1(i,8) R1(i,9)];
fis11.input(1).mf(4).params=[R1(i,10) R1(i,11) R1(i,12)];
fis11.input(1).mf(5).params=[R1(i,13) R1(i,14) R1(i,15)];

fis11.input(2).mf(1).params=[F1(i,1) F1(i,2) F1(i,3)];
fis11.input(2).mf(2).params=[F1(i,4) F1(i,5) F1(i,6)];
fis11.input(2).mf(3).params=[F1(i,7) F1(i,8) F1(i,9)];
fis11.input(2).mf(4).params=[F1(i,10) F1(i,11) F1(i,12)];
fis11.input(2).mf(5).params=[F1(i,13) F1(i,14) F1(i,15)];

```

```

fis11.input(3).mf(1).params=[M1(i,1) M1(i,2) M1(i,3)];
fis11.input(3).mf(2).params=[M1(i,4) M1(i,5) M1(i,6)];
fis11.input(3).mf(3).params=[M1(i,7) M1(i,8) M1(i,9)];
fis11.input(3).mf(4).params=[M1(i,10) M1(i,11) M1(i,12)];
fis11.input(3).mf(5).params=[M1(i,13) M1(i,14) M1(i,15)];

```

```

tahmin=evalfis(RFM,fis11);
gsonuc=train1(:,6);
giris=getfis(fis11,'inmfparams');
cikis=getfis(fis11,'outmfparams');
girisaralik=fis11.input(1).mf(2).params;
d=abs(tahmin-gsonuc);
c=length(d);
accuracy1(i)=sum(d./gsonuc)/c;

end
sort(accuracy1)
[err1,ind1]=sort(accuracy1);
SR1=R1;
SF1=F1;
SM1=M1;
for i=1:psize
    SR1(i,:)=R1(ind1(i),:);
    SF1(i,:)=F1(ind1(i),:);
    SM1(i,:)=M1(ind1(i),:);
end
SRFM=[SR1 SF1 SM1];
SRFM11=[SRFM11;([accuracy' SR1 SF1 SM1])];
AX=SRFM11;
AX1=sort(AX,'descend');
for i=1:ceil(psize/2)
    popsec(i,:)=rfmrand(ind(i),:);
end

toplam=sum(1./err(1:ceil(psize/2)));

oran=round(100*(1./err(1:ceil(psize/2))./toplam));

n1=1;
m1=oran(1);

[ro co]=size(popsec);
popsecoran=zeros(sum(oran),co);
for i=1:ceil(psize/2)
    for j=1:co
        popsecoran(n1:m1,j)=popsec(i,j);
    end
    n1=m1+1;
    if i<round(psize/2)
        m1=m1+oran(i+1);
    end
end
end

```

```

        secim2(1,:)=popsec(1,:);
for i=2:ceil(psize)
    secim2(i,:)=popsecoran(ceil(1+98.*(rand(1,1))),:);

end
    popsecoran1=secim2;
    R1=popsecoran1(1:10,1:15);
    F1=popsecoran1(1:10,16:30);
    M1=popsecoran1(1:10,31:45);

for i=1:psize

    fis11.input(1).mf(1).params=[R1(i,1) R1(i,2) R1(i,3)];
    fis11.input(1).mf(2).params=[R1(i,4) R1(i,5) R1(i,6)];
    fis11.input(1).mf(3).params=[R1(i,7) R1(i,8) R1(i,9)];
    fis11.input(1).mf(4).params=[R1(i,10) R1(i,11) R1(i,12)];
    fis11.input(1).mf(5).params=[R1(i,13) R1(i,14) R1(i,15)];

    fis11.input(2).mf(1).params=[F1(i,1) F1(i,2) F1(i,3)];
    fis11.input(2).mf(2).params=[F1(i,4) F1(i,5) F1(i,6)];
    fis11.input(2).mf(3).params=[F1(i,7) F1(i,8) F1(i,9)];
    fis11.input(2).mf(4).params=[F1(i,10) F1(i,11) F1(i,12)];
    fis11.input(2).mf(5).params=[F1(i,13) F1(i,14) F1(i,15)];

    fis11.input(3).mf(1).params=[M1(i,1) M1(i,2) M1(i,3)];
    fis11.input(3).mf(2).params=[M1(i,4) M1(i,5) M1(i,6)];
    fis11.input(3).mf(3).params=[M1(i,7) M1(i,8) M1(i,9)];
    fis11.input(3).mf(4).params=[M1(i,10) M1(i,11) M1(i,12)];
    fis11.input(3).mf(5).params=[M1(i,13) M1(i,14) M1(i,15)];

    tahmin=evalfis(RFM,fis11);
    gsonuc=train1(:,6);
    cikis=getfis(fis11,'outmfparams');
    girisarik=fis11.input(1).mf(2).params;
    d=abs(tahmin-gsonuc);
    c=length(d);
    accuracy1(i)=sum(d./gsonuc)/c;

end
    sort(accuracy1)
    [err1,ind1]=sort(accuracy1);
    SR1=R1;
    SF1=F1;
    SM1=M1;
for i=1:psize
    SR1(i,:)=R1(ind1(i),:);
    SF1(i,:)=F1(ind1(i),:);
    SM1(i,:)=M1(ind1(i),:);

end
    SRFM=[SR1 SF1 SM1];
    pair1=(2:ceil(psize/2));
    pair1=[1 pair1];
    pair2=randperm(psize-1-ceil(psize/2)-1+1)+ceil(psize/2)+1-1;
    pair2=[1 pair2];
    R1bin=dec2bin(SR1');

```

```

F1bin=dec2bin(SF1');
M1bin=dec2bin(SM1');
[Rro Rc]=size(R1bin);
[Fro Fc]=size(F1bin);
[Mro Mc]=size(M1bin);
Rr=ceil(1+(Rc-1).*(rand(1,1)));
Fr=ceil(1+(Fc-1).*(rand(1,1)));
Mr=ceil(1+(Mc-1).*(rand(1,1)));
for i=1:length(pair1)
    CRbin(15*i-14:15*i,1:Rc)=[R1bin(pair1(i)*15-14:pair1(i)*15,1:Rr)
R1bin(pair2(i)*15-14:pair2(i)*15,Rr+1:Rc)];
    CRbin2(15*i-14:15*i,1:Rc)=[R1bin(pair2(i)*15-14:pair2(i)*15,1:Rr)
R1bin(pair1(i)*15-14:pair1(i)*15,Rr+1:Rc)];
    CFbin(15*i-14:15*i,1:Fc)=[F1bin(pair1(i)*15-14:pair1(i)*15,1:Fr)
F1bin(pair2(i)*15-14:pair2(i)*15,Fr+1:Fc)];
    CFbin2(15*i-14:15*i,1:Fc)=[F1bin(pair2(i)*15-14:pair2(i)*15,1:Fr)
F1bin(pair1(i)*15-14:pair1(i)*15,Fr+1:Fc)];

    CMbin(15*i-14:15*i,1:Mc)=[M1bin(pair1(i)*15-14:pair1(i)*15,1:Mr)
M1bin(pair2(i)*15-14:pair2(i)*15,Mr+1:Mc)];
    CMbin2(15*i-14:15*i,1:Mc)=[M1bin(pair2(i)*15-14:pair2(i)*15,1:Mr)
M1bin(pair1(i)*15-14:pair1(i)*15,Mr+1:Mc)];

end
s=1;
for i=1:length(pair1);
    CRb(s*15-14:s*15,:)=CRbin(i*15-14:i*15,:);
    CRb((s+1)*15-14:(s+1)*15,:)=CRbin2(i*15-14:i*15,:);

    CFb(s*15-14:s*15,:)=CFbin(i*15-14:i*15,:);
    CFb((s+1)*15-14:(s+1)*15,:)=CFbin2(i*15-14:i*15,:);

    CMb(s*15-14:s*15,:)=CMbin(i*15-14:i*15,:);
    CMb((s+1)*15-14:(s+1)*15,:)=CMbin2(i*15-14:i*15,:);

    s=s+2;

end
Rmut=ceil((Rro-1)*Rc*mo);
Rmutr=ceil(18+(Rro-18)*rand(1,Rmut));
Rmutc=ceil(Rc*rand(1,Rmut));

Fmut=ceil((Fro-1)*Fc*mo);
Fmutr=ceil(18+(Fro-18)*rand(1,Fmut));
Fmutc=ceil(Fc*rand(1,Fmut));

Mmut=ceil((Mro-1)*Mc*mo);
Mmutr=ceil(18+(Mro-18)*rand(1,Mmut));
Mmutc=ceil(Mc*rand(1,Mmut));
for i=1:Rmut
    if CRb(Rmutr(i),Rmutc(i))==0
        CRb(Rmutr(i),Rmutc(i))='1';
    else
        CRb(Rmutr(i),Rmutc(i))='0';
    end
end
for i=1:Fmut
    if CFb(Fmutr(i),Fmutc(i))==0

```

```

        CFb(Fmutr(i),Fmutc(i))='1';
    else
        CFb(Fmutr(i),Fmutc(i))='0';
    end
end
for i=1:Mmut
    if CMb(Mmutr(i),Mmutc(i))==0
        CMb(Mmutr(i),Mmutc(i))='1';
    else
        CMb(Mmutr(i),Mmutc(i))='0';
    end
end

MR1=bin2dec(CRb);
MF1=bin2dec(CFb);
MM1=bin2dec(CMb);
[MuR1,padded] = vec2mat(sort(vec2mat(MR1,3),2),15);

[MuF1,padded] = vec2mat(sort(vec2mat(MF1,3),2),15);

[MuM1,padded] = vec2mat(sort(vec2mat(MM1,3),2),15);
[Rm,Cm]=size(MuR1);

MuR2=MuR1;
MuF2=MuF1;
MuM2=MuM1;
[Mr,Mc]=size(MuR2);

for kk=3:Mr
    for cc=1:3:Mc
    switch cc
        case 1
            if MuR2(kk,1)==0 & MuR2(kk,1)==0 & MuR2(kk,1)==0
                MuR2(kk,3)= sort((round(10+(64-10).*(rand(1,1))))),2)
            end
            if MuF2(kk,1)==0 & MuF2(kk,1)==0 & MuF2(kk,1)==0
                MuF2(kk,3)= sort((round(10+(64-10).*(rand(1,1))))),2)
            end
            if MuM2(kk,1)==0 & MuM2(kk,1)==0 & MuM2(kk,1)==0
                MuM2(kk,3)= sort((round(10+(64-10).*(rand(1,1))))),2)
            end
        case 4
            if MuR2(kk,cc)==0 & MuR2(kk,cc+1)==0 & MuR2(kk,cc+2)==0
                ras=sort((round(10+(64-10).*(rand(1,3))))),2)
                MuR2(kk,cc)=ras(1);
                MuR2(kk,cc+1)=ras(2);
                MuR2(kk,cc+2)=ras(3);
            elseif MuR2(kk,cc)==0 & MuR2(kk,cc+1)==0
                ras=sort((round(1+(MuR2(kk,cc+2)-1).*(rand(1,2))))),2)
                MuR2(kk,cc)=ras(1);
                MuR2(kk,cc+1)=ras(2);
            end
            if MuF2(kk,cc)==0 & MuF2(kk,cc+1)==0 & MuF2(kk,cc+2)==0
                ras=sort((round(10+(64-10).*(rand(1,3))))),2);
                MuF2(kk,cc)=ras(1);
                MuF2(kk,cc+1)=ras(2);
                MuF2(kk,cc+2)=ras(3);
            end
        end
    end
end

```



```

elseif MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0
    ras=sort((round(1+(MuF2(kk, cc+2)-1).*(rand(1,2))))),2);
    MuF2(kk, cc)=ras(1);
    MuF2(kk, cc+1)=ras(2);
end
if MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0 & MuM2(kk, cc+2)==0
    ras=sort((round(10+(64-10).*(rand(1,3))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
    MuM2(kk, cc+2)=ras(3);
elseif MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0
    ras=sort((round(1+(MuM2(kk, cc+2)-1).*(rand(1,2))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
end
case 7
if MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0 & MuR2(kk, cc+2)==0
    ras=sort((round(30+(80-30).*(rand(1,3))))),2);
    MuR2(kk, cc)=ras(1);
    MuR2(kk, cc+1)=ras(2);
    MuR2(kk, cc+2)=ras(3);
elseif MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0
    ras=sort((round(10+(MuR2(kk, cc+2)-10).*(rand(1,2))))),2);
    MuR2(kk, cc)=ras(1);
    MuR2(kk, cc+1)=ras(2);
end
if MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0 & MuF2(kk, cc+2)==0
    ras=sort((round(30+(80-30).*(rand(1,3))))),2);
    MuF2(kk, cc)=ras(1);
    MuF2(kk, cc+1)=ras(2);
    MuF2(kk, cc+2)=ras(3);
elseif MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0
    ras=sort((round(10+(MuR2(kk, cc+2)-10).*(rand(1,2))))),2);
    MuF2(kk, cc)=ras(1);
    MuF2(kk, cc+1)=ras(2);
end
if MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0 & MuM2(kk, cc+2)==0
    ras=sort((round(30+(80-30).*(rand(1,3))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
    MuM2(kk, cc+2)=ras(3);
elseif MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0
    ras=sort((round(10+(MuM2(kk, cc+2)-10).*(rand(1,2))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
end
case 10
if MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0 & MuR2(kk, cc+2)==0
    ras=sort((round(50+(100-50).*(rand(1,3))))),2);
    MuR2(kk, cc)=ras(1);
    MuR2(kk, cc+1)=ras(2);
    MuR2(kk, cc+2)=ras(3);
elseif MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0
    ras=sort((round(10+(MuR2(kk, cc+2)-10).*(rand(1,2))))),2);
    MuR2(kk, cc)=ras(1);
    MuR2(kk, cc+1)=ras(2);
end
if MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0 & MuF2(kk, cc+2)==0
    ras=sort((round(50+(100-50).*(rand(1,3))))),2);
    MuF2(kk, cc)=ras(1);

```

```

        MuF2(kk, cc+1)=ras(2);
        MuF2(kk, cc+2)=ras(3);
    elseif MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0
        ras=sort((round(10+(MuF2(kk, cc+2)-10).* (rand(1,2))))),2);
        MuF2(kk, cc)=ras(1);
        MuF2(kk, cc+1)=ras(2);
    end
if MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0 & MuM2(kk, cc+2)==0
    ras=sort((round(50+(100-50).* (rand(1,3))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
    MuM2(kk, cc+2)=ras(3);
elseif MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0
    ras=sort((round(10+(MuM2(kk, cc+2)-10).* (rand(1,2))))),2);
    MuM2(kk, cc)=ras(1);
    MuM2(kk, cc+1)=ras(2);
end
case 13
    if MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0 & MuR2(kk, cc+2)==0
        ras=sort((round(64+(100-64).* (rand(1,3))))),2);
        MuR2(kk, cc)=ras(1);
        MuR2(kk, cc+1)=128;
        MuR2(kk, cc+2)=128;
    elseif MuR2(kk, cc)==0 & MuR2(kk, cc+1)==0
        ras=sort((round(64+(128-64).* (rand(1,2))))),2);
        MuR2(kk, cc)=ras(1);
        MuR2(kk, cc+1)=ras(2);
        MuR2(kk, cc+2)=128;
    end
    if MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0 & MuF2(kk, cc+2)==0
        ras=sort((round(64+(100-64).* (rand(1,3))))),2);
        MuF2(kk, cc)=ras(1);
        MuF2(kk, cc+1)=128;
        MuF2(kk, cc+2)=128;
    elseif MuF2(kk, cc)==0 & MuF2(kk, cc+1)==0
        ras=sort((round(64+(128-64).* (rand(1,2))))),2);
        MuF2(kk, cc)=ras(1);
        MuF2(kk, cc+1)=ras(2);
        MuF2(kk, cc+2)=128;
    end
    if MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0 & MuM2(kk, cc+2)==0
        ras=sort((round(64+(100-64).* (rand(1,3))))),2);
        MuM2(kk, cc)=ras(1);
        MuM2(kk, cc+1)=128;
        MuM2(kk, cc+2)=128;
    elseif MuM2(kk, cc)==0 & MuM2(kk, cc+1)==0
        ras=sort((round(64+(128-64).* (rand(1,2))))),2);
        MuM2(kk, cc)=ras(1);
        MuM2(kk, cc+1)=ras(2);
        MuM2(kk, cc+2)=128;
    end
end

end
end
end
MuR1=MuR2;
MuF1=MuF2;
MuM1=MuM2;

rr=MuR1;

```

```

ff=MuF1;
mm=MuM1;
hold on
figure (1)
plotmf(fis11, 'input', 1)
hold on
figure (2)
plotmf(fis11, 'input', 2)
hold on
figure (3)
plotmf(fis11, 'input', 3)

end
figure (4)
plotmf(fis11, 'input', 1)

figure (5)
plotmf(fis11, 'input', 2)

figure (6)
plotmf(fis11, 'input', 3)
figure (7)
plotmf(fis11, 'input', 4)
figure (8)
plotmf(fis11, 'input', 5)
karsilastir=[tahmin gsonuc];

writefis(fis11, 'rfm1285lioptyeni');
fis2=readfis('rfm1285lioptyeni');
tahmin1=evalfis(test11(:, 1:5), fis2);
gsonuc1=test11(:, 6);
d1=abs(tahmin1-gsonuc1);
c1=length(d1);
accuracy1=sum(d1./gsonuc1)/c1
figure
plot(gsonuc, 'b.-')
hold on
plot(tahmin, 'ro')
xlabel('Öğrenci No')
ylabel('Notlar')
legend('Akademik Performans, 'Tahmin')
hold off
figure
plot(gsonuc1, 'b.-')
hold on
plot(tahmin1, 'r.')
xlabel('Öğrenci No')
ylabel('Notlar')
legend('Akademik Performans, 'Tahmin')
hold off

```

EK 2. Kümeleme Yöntemleri İle Akademik Performans Tahmini Matlab Kodu

```

clc
clear all

load('son_notout_KM_2_Tnotout_1321_2711.mat')

[numData,numInp] = size(Xin);
[numData2,numOutp] = size(Xout);
radii=0.5;
options=[];
xBounds=[];
if numData ~= numData2

    if numData == numOutp

        Xout = Xout';
        numOutp = numData2;
    else
        error('Mismatched input and output data matrices');
    end
end

m=menu('Kümeleme Analizi','Kmeans','FCM','Subtractive');
switch m
    case 1
        ks=input('küme sayısını giriniz');
        for kk=2:ks
            [IDX,Kc,sumd,Dist] = kmeans(Xin,k);
            centers=Kc;
            Mt='KM';
        end
    case 2
        ks=input('küme sayısını giriniz');
        for kk=2:ks

            [fcmc, fcmU, fcmobj_fcn] = fcm(Xin,k);
            centers=fcmc;
            Mt='FCM';
        end

    case 3
        [centers,sigmas] = subclust([Xin Xout],radii);
        Mt='SC';
end

X = [Xin Xout];
if isempty(xBounds)
    minX = min(X);
    maxX = max(X);

    index = find(maxX == minX);
    minX(index) = minX(index) - 0.0001*(1 + abs(minX(index)));
    maxX(index) = maxX(index) + 0.0001*(1 + abs(maxX(index)));
else

```

```

        minX = xBounds(1,:);
        maxX = xBounds(2,:);
        if length(minX) ~= size(X,2)
            error('xBounds contains the wrong dimensions for input
data X');
        elseif any(maxX == minX)
            error('xBounds has a data range of zero');
        end
    end
    sigmas = (radii .* (maxX - minX)) / sqrt(8.0);

    centers = centers(:,1:numInp);
    sigmas = sigmas(:,1:numInp);

    distMultp = (1.0 / sqrt(2.0)) ./ sigmas;

    [numRule,foo] = size(centers);
    sumMu = zeros(numData,1);
    muVals = zeros(numData,1);
    dxMatrix = zeros(numData,numInp);
    muMatrix = zeros(numData,numRule * (numInp + 1));

    for i=1:numRule

        for j=1:numInp
            dxMatrix(:,j) = (Xin(:,j) - centers(i,j)) * distMultp(j);
        end

        dxMatrix = dxMatrix .* dxMatrix;
        if numInp > 1
            muVals(:) = exp(-1 * sum(dxMatrix'));
        else
            muVals(:) = exp(-1 * dxMatrix');
        end

        sumMu = sumMu + muVals;

        colNum = (i - 1)*(numInp + 1);
        for j=1:numInp
            muMatrix(:,colNum + j) = Xin(:,j) .* muVals
        end

        muMatrix(:,colNum + numInp + 1) = muVals;

    end

    sumMuInv = 1.0 ./ sumMu;
    for j=1:(numRule * (numInp + 1))
        muMatrix(:,j) = muMatrix(:,j) .* sumMuInv;
    end

    outEqns = muMatrix \ Xout;

    numInDigit = floor(log10(numInp)) + 1;
    numOutDigit = floor(log10(numOutp)) + 1;
    numRuleDigit = floor(log10(numRule)) + 1;

```

```

numRow = 11 + (2 * (numInp + numOutp)) + (3 * (numInp + numOutp) *
numRule);
numCol = numInp + numOutp + 2;
strSize = 3 + numInDigit + numOutDigit;
numCol = max(numCol, strSize);
strSize = 4 + numInDigit + numRuleDigit;
numCol = max(numCol, strSize);
strSize = 5 + numOutDigit + numRuleDigit;
numCol = max(numCol, strSize);
numCol = max(numCol, 7);

theStr = sprintf('sug%g%g', numInp, numOutp);
fismat.name=theStr;
% FIS type
fismat.type = 'sugeno';
fismat.andMethod = 'prod';
fismat.orMethod = 'probor';
fismat.impMethod = 'prod';
fismat.aggMethod = 'max';
fismat.defuzzMethod = 'wtaver';

rowIndex = 11;
for i=1:numInp
    theStr = sprintf('in%g', i);
    strSize = length(theStr);
    fismat.input(i).name = theStr;
end

for i=1:numOutp
    theStr = sprintf('out%g', i);
    strSize = length(theStr);
    fismat.output(i).name = theStr;
end

if length(xBounds) == 0
    minX = min(Xin);
    maxX = max(Xin);
else
    minX = xBounds(1, 1:numInp);
    maxX = xBounds(2, 1:numInp);
end
ranges = [minX ; maxX]';
for i=1:numInp
    fismat.input(i).range = ranges(i, :);
end

if length(xBounds) == 0
    minX = min(Xout);
    maxX = max(Xout);
else
    minX = xBounds(1, numInp+1:numInp+numOutp);
    maxX = xBounds(2, numInp+1:numInp+numOutp);
end
ranges = [minX ; maxX]';
for i=1:numOutp
    fismat.output(i).range = ranges(i, :);
end

```

```

for i=1:numInp
    for j=1:numRule
        theStr = sprintf('in%gcluster%g',i,j);
        fismat.input(i).mf(j).name = theStr;
    end
end

for i=1:numOutp
    for j=1:numRule
        theStr = sprintf('out%gcluster%g',i,j);
        fismat.output(i).mf(j).name = theStr;
    end
end

for i=1:numInp
    for j=1:numRule
        fismat.input(i).mf(j).type = 'gaussmf';
    end
end

for i=1:numOutp
    for j=1:numRule
        fismat.output(i).mf(j).type = 'linear';
    end
end

colOfOnes = ones(numRule
for i=1:numInp
    for j=1:numRule
        fismat.input(i).mf(j).params = [sigmas(i) centers(j, i)];
    end
end

for i=1:numOutp
    for j=1:numRule
        outParams = reshape(outEqns(:,i),numInp + 1,numRule);
        fismat.output(i).mf(j).params = outParams(:,j)';
    end
end

colOfEnum = [1:numRule]';

for j=1:numRule
    for i=1:numInp
        fismat.rule(j).antecedent(i)=colOfEnum(j);
    end
    for i=1:numOutp
        fismat.rule(j).consequent(i) = colOfEnum(j);
    end

    fismat.rule(j).weight=1;
    fismat.rule(j).connection=1;
end
tahmin=evalfis(Xin,fismat)
d=abs(tahmin-Xout);
c=length(d);

```

```

accuracy(k)=sum(d./Xout)/c;

writefis(fismat, 'sugenodeneme')
end
figure1 = figure('PaperType','a4letter',...
    'PaperSize',[20.98404194812 29.67743169791],...
    'Color',[1 1 1]);
[x,mf] = plotmf(fismat, 'input',1);
subplot(5,1,1), plot(x,mf);
xlabel('Zaman Üyelik Fonksiyonu');
[x,mf] = plotmf(fismat, 'input',2);
subplot(5,1,2), plot(x,mf);
xlabel('Frekans Üyelik Fonksiyonu');
[x,mf] = plotmf(fismat, 'input',3);
subplot(5,1,3), plot(x,mf);
xlabel('Süre Üyelik Fonksiyonu');
[x,mf] = plotmf(fismat, 'input',4);
subplot(5,1,4), plot(x,mf);
xlabel('Küçük Sınav Üyelik Fonksiyonu');
[x,mf] = plotmf(fismat, 'input',5);
subplot(5,1,5), plot(x,mf);
xlabel('Arasınav Üyelik Fonksiyonu');
figure2=figure('PaperSize',[20.98404194812 29.67743169791],...
    'Color',[1 1 1]);
subplot(3,1,1)
plot(Xout, 'b.-')
hold on
plot(tahmin, 'ro-')
xlabel('Veri Sayısı')
ylabel('Not')
legend('AP', 'Tahmin');
title('Eğitim Verileri Karşılaştırma Grafiği')
axis([0 160 0 100])
hold off
subplot(3,1,2)
plot(Xoutt, 'b.-')
hold on
plot(tahmintest, 'ro-')
xlabel('Veri Sayısı')
ylabel('Not')
legend('AP', 'Tahmin')
title('Test Verileri Karşılaştırma Grafiği')
axis([0 70 0 100])
subplot(3,1,3)
plot(Xouttv, 'b.-')
hold on
plot(tahminv, 'ro-')
axis([0 100 0 100])
title('Doğrulama Verileri Karşılaştırma Grafiği')
xlabel('Veri Sayısı')
ylabel('Not')
legend('AP', 'Tahmin')

```

EK 3. Kantil Regresyon Parametre Tahmini Matlab Kodu

```

clear all
%load son_notout_KM_8_Tnotout_0997_2909.mat

```



```

%load son_notout_KM_8_xtv.mat
load son_notout_FCM_7_Tnotout_1069_2480.mat
%load son_notout_SC_1_Tnotout_1025_2360.mat
[bbb,h,exitflag,hhat] = quantile1(0.5,Xout,muMatrix);
outEqns=bbb;
numInDigit = floor(log10(numInp)) + 1;
numOutDigit = floor(log10(numOutp)) + 1;
numRuleDigit = floor(log10(numRule)) + 1;

numRow = 11 + (2 * (numInp + numOutp)) + (3 * (numInp + numOutp) *
numRule);
numCol = numInp + numOutp + 2;          % number of columns required for
the rule list
strSize = 3 + numInDigit + numOutDigit; % size of name
'sug[numInp][numOutp]'
numCol = max(numCol,strSize);
strSize = 4 + numInDigit + numRuleDigit; % size of
'in[numInp]mf[numRule]'
numCol = max(numCol,strSize);
strSize = 5 + numOutDigit + numRuleDigit; % size of
'out[numOutp]fm[numRule]'
numCol = max(numCol,strSize);
numCol = max(numCol,7); % size of 'gaussmf' is 7

theStr = sprintf('sug%g%g',numInp,numOutp);
fismat.name=theStr;
fismat.type = 'sugeno';
order, klu)
fismat.andMethod = 'prod';
fismat.orMethod = 'probor';
fismat.impMethod = 'prod';
fismat.aggMethod = 'max';
fismat.defuzzMethod = 'wtaver';

rowIndex = 11;

for i=1:numInp
    theStr = sprintf('in%g',i);
    strSize = length(theStr);
    fismat.input(i).name = theStr;
end

for i=1:numOutp
    theStr = sprintf('out%g',i);
    strSize = length(theStr);
    fismat.output(i).name = theStr;
end

if length(xBounds) == 0
    minX = min(Xin);
    maxX = max(Xin);
else
    minX = xBounds(1,1:numInp);
    maxX = xBounds(2,1:numInp);
end
ranges = [minX ; maxX]';
for i=1:numInp
    fismat.input(i).range = ranges(i,:);
end

```

```

if length(xBounds) == 0
    minX = min(Xout);
    maxX = max(Xout);
else
    minX = xBounds(1,numInp+1:numInp+numOutp);
    maxX = xBounds(2,numInp+1:numInp+numOutp);
end
ranges = [minX ; maxX]';
for i=1:numOutp
    fismat.output(i).range = ranges(i,:);
end

% Set the input membership function labels
for i=1:numInp
    for j=1:numRule
        theStr = sprintf('in%gcluster%g',i,j);
        fismat.input(i).mf(j).name = theStr;
    end
end

for i=1:numOutp
    for j=1:numRule
        theStr = sprintf('out%gcluster%g',i,j);
        fismat.output(i).mf(j).name = theStr;
    end
end

for i=1:numInp
    for j=1:numRule
        fismat.input(i).mf(j).type = 'gaussmf';
    end
end

for i=1:numOutp
    for j=1:numRule
        fismat.output(i).mf(j).type = 'linear';
    end
end

colOfOnes = ones(numRule,1); % a column of ones
for i=1:numInp
    for j=1:numRule
        fismat.input(i).mf(j).params = [sigmas(i) centers(j, i)];
    end
end

for i=1:numOutp
    for j=1:numRule
        outParams = reshape(outEqns(:,i),numInp + 1,numRule);
        fismat.output(i).mf(j).params = outParams(:,j)';
    end
end

colOfEnum = [1:numRule]';

for j=1:numRule
    for i=1:numInp
        fismat.rule(j).antecedent(i)=colOfEnum(j);
    end
    for i=1:numOutp
        fismat.rule(j).consequent(i) = colOfEnum(j);
    end
end

```

```

end

    fismat.rule(j).weight=1;
    fismat.rule(j).connection=1;
end
tahmin=evalfis(Xin,fismat)
d=abs(tahmin-Xout);
c=length(d);
accuracy(k)=sum(d./Xout)/c;

modelsonuc=[Xout tahmin Xout-tahmin];
[err,ind]=sort(accuracy);

tahmintest=evalfis(Xint,fismat);
dt=abs(tahmintest-Xoutt);
ct=length(dt);
accuracytest(k)=sum(dt./Xoutt)/ct;

testsonuc=[Xoutt tahmintest Xoutt-tahmintest];
[errt,indt]=sort(accuracytest)
aaa=[Xin Xout tahmin];aat=[Xint Xoutt tahmintest];
TT='Tnormal';
tahminv=evalfis(Xintv,fismat);
dtv=abs(tahminv-Xouttv);
ctv=length(dtv);
accuracyv(k)=sum(dtv./Xouttv)/ctv
%sort(accuracyv(k));

    a1=num2str(accuracy(k));
    a2=num2str(accuracyv(k));

```

EK 4. Süre Girdi Değişkeninin Veri Setinden Okunması Matlab Kodu

```

[num,txt,raw]=xlsread('rfmosm.xlsx');%rfmosm.xlsx okuma
[uniqueC,~,idx] = unique(raw(:,4));%her öğrencinin kaç satırda yer
aldığını bulur
counts = accumarray(idx(:,1),1,[],@sum);
[adsoyadn,adsoyadr,adsoyadt]=xlsread('adsoyad.xlsx');
adsoyadn(:,2)=counts;
[tarih]=xlsread('tariharalik.xlsx');%tarih araliklari okundu
ekle=0;
Tliste=adsoyadr;
Tfrekans=raw;
Tfrndata=num;
[Slr,Slc]=size(Tliste);
[Sfr,Sfc]=size(Tfrekans);

recency=[];
for i=1:Slr

s1=1;s2=1;s3=1;s4=1;s5=1;s6=1;s7=1;s8=1;s9=1;s10=1;s11=1;s12=1;s13=1;s
14=1;s15=1;
b1=[];b2=[];b3=[];b4=[];b5=[];b6=[];b7=[];b8=[];b9=[];b10=[];b11=[];b1
2=[];b13=[];b14=[];
    for j=1:Sfr
        if strcmp(Tliste(i,2),Tfrekans(j,4))==1
            if Tfrndata(j,5)==1

```


ÖZGEÇMİŞ



Kişisel Bilgiler

Adı Soyadı	Osman YILDIZ
Uyruğu	T.C
Doğum tarihi, Yeri	19.06.1975 / Wertahl
Telefon	05327715507
E-mail	osmyildiz@gmail.com
Web adres	

Eğitim

Derece	Kurum/Anabilim Dalı/Programı	Yılı
Doktora	/ /	
Yüksek Lisans	Marmara Ü. Sosyal Bilimler Enstitüsü/ Ekonometri	2005
Lisans	Marmara Üniversitesi, Matematik Eğitimi	1998
Lise	Kartal Meslek Lisesi	1992

Makaleler / Bildiriler

<p>Bildiri ve Makaleler</p> <p>Yıldız, O., Bal, A., & Gülseçen, S., 2013, Clustering based fuzzy model for academic performance estimation in distance education. <i>Journal of Pensee</i>, 75(12).</p> <p>Yıldız, O., Bal, A., & Gülseçen, S., 2013, Improved Fuzzy Modelling to Predict the Academic Performance of the Distance Education Students'. <i>The International Review of Research in Open and Distance Learning</i>, 14(5), 144-</p>

165.

Yıldız, O., Bal, A., Gülseçen, S., & Kentli, F. D., 2012, A Genetic-Fuzzy Based Mathematical Model to Evaluate The Distance Education Students' Academic Performance. *Procedia - Social and Behavioral Sciences*, Vol:55,p:409-418.

Yildiz, O, Bal, A., Gülseçen, S., & Kentli, F. D., 2014, Rules Optimization Based Fuzzy Model for Predicting Distance Education Students' Grades, *International Journal of Information and Education Technology*, 4(1), 59-62.

