

**İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ**

**BULANIK MANTIĞIN VERİ MADENCİLİĞİNE  
UYGULANMASI**

**YÜKSEK LİSANS TEZİ  
Mat. Müh. Selahattin BOSTANCI**

**Anabilim Dalı : MÜHENDİSLİK BİLİMLERİ**

**Program : SİSTEM ANALİZİ**

**MAYIS 2003**

**BULANIK MANTIK İN VERİ MADENCİLİĞİNE  
UYGULANMASI**

**YÜKSEK LİSANS TEZİ**

**Mat. Müh. Selahattin BOSTANCI**

**509011204**

**Tezin Enstitüye Verildiği Tarih: 5 Mayıs 2003**

**Tezin Savunulduğu Tarih: 28 Mayıs 2003**

**Tez Danışmanı : Yrd. Doç. Dr. Ali ERCENGİZ**

**Diğer Jüri Üyeleri Doç. Dr. Şakir KOCABAŞ**

**Doç. Dr. Metin Orhan KAYA**

**MAYIS 2003**

## ÖNS ÖZ

Tezi mi hazırlarken bana her konuda yardımcı olan ve manevi desteğini hiçbir zaman esirgemeyen saygın hocam Yrd. Doç. Dr. Ali ERCENGİ Z'e;

Beni veri madenciliği konusuna yönlendiren, bilgisini ve desteğini benimle her zaman paylaşmış olan saygın hocam Prof. Dr. Gazanfer ÜNAL'a;

Beni hiçbir zaman yalnız bırakmayan ve benim bu noktaya gelmemde en büyük paya sahip olan anne me, baba ma ve abla ma;

Bu tezi hazırlanırken yardımlarını benden esirgemeyen canım sevgilim E. Aygül KAYNAK'a çok teşekkür ederim

Mayıs 2003

Selahattin BOSTANCI

<b>İÇİNDEKİLER</b>	
<b>KISALTMALAR</b>	<b>v</b>
<b>TABLO LİSTESİ</b>	<b>vi</b>
<b>ŞEKİL LİSTESİ</b>	<b>vii</b>
<b>SEMBOL LİSTESİ</b>	<b>viii</b>
<b>ÖZET</b>	<b>ix</b>
<b>SUMMARY</b>	<b>x</b>
<b>1. VERİ MADENCİLİĞİ</b>	<b>1</b>
1.1 Giriş	1
1.2 Veri Madenciliği Modelleri	2
1.2.1 Sınıflama ve Regresyon Modelleri	3
1.2.2 Kümeleme Modelleri	4
1.2.3 Birliklik Kuralları ve Ardışık Zamanlı Öüntüler	4
1.3 Örnek Uygulamalar	4
1.4 Metodoloji	6
<b>2. VERİ MADENCİLİĞİNDE BULANIK MANTIK YAKLAŞIMI</b>	<b>8</b>
2.1 Giriş	8
2.2 Bulanık Kümeler	10
2.2.1 Üyelik Fonksiyonları	11
2.2.2 Bulanık Değişkenler	16
2.2.3 Temel Bulanık Küme İşlemleri	17
2.2.4 Bulanık Kümenin Özellikleri	19
2.3 Bulanık Mantık	19
2.3.1 Bulanık Mantığın Tarihi	20
2.3.2 Bulanık Mantığın Özellikleri	20
2.3.3 Bulanık Mantık ile Klasik Mantığın Farklılıkları	21
2.3.4 Dilsel Değişkenler	21
<b>3. BULANIK SORGULAMA İLE BİLGİ KEŞFİ</b>	<b>23</b>
3.1 Giriş	23
3.2 Bulanık Sorgulama	24
3.3 Dilsel Özetler	26

3.4 Bulamık Fonksiyonel ve Dereceli Fonksiyonel Bađlılıklar	29
3.4.1 Bulamık Fonksiyonel Bađlılık	30
3.4.2 Dereceli Fonksiyonel Bađlılık	32
3.4.3 Bulamıklık Operatörü	38
<b>4. GELİŐTİRİLEN UYGULAMANIN GENEL YAPISI</b>	<b>39</b>
4.1 Giriő	39
4.2 Veritabanları	39
4.2.1 Őirket Veritabanı	39
4.2.2 Borsa veritabanı	41
4.3 Geliőtirilen Araç	42
4.3.1 Bulamık Sorgulama	43
4.3.2 Dösel Özet	44
4.3.3 Bulamık Fonksiyonel Bađlılık ve Dereceli Fonksiyonel Bađlılık	44
4.3.4 Yeni Bulamık Alanlar Tanımlamak	45
4.4 Borsa Veritabanından Üretilen Kurallar	47
4.4.1 BFB Kullanılarak Üretilen Kurallar	47
4.4.2 DFB Kullanılarak Üretilen Kurallar	49
4.4.3 Borsa Veritabanının Deđerlendirilmesi	56
<b>SONUÇLAR VE ÖNERİLER</b>	<b>58</b>
<b>KAYNAKLAR</b>	<b>59</b>
<b>EKLER</b>	<b>60</b>
EK A: GELİŐTİRİLEN UYGULAMA	60
<b>ÖZGEÇMİŐ</b>	<b>61</b>

## **KISALTMALAR**

<b>İMKB</b>	: İstanbul Menkul Kıymetler Borsası
<b>VT</b>	: Veri Tabanı
<b>FB</b>	: Fonksiyonel Bağlılık
<b>BFB</b>	: Bulanık Fonksiyonel Bağlılık
<b>DFB</b>	: Dereceli Fonksiyonel Bağlılık

## TABLO LİSTESİ

<b>Tablo 2.1</b>	Veri Tabanı Teknolojisinin Evrimi . . . . .	8
<b>Tablo 3.1</b>	Örnek veritabanı . . . . .	25
<b>Tablo 3.2</b>	Uzun boylu insanlar sorgusu . . . . .	25
<b>Tablo 3.3</b>	Uzun ve şişman insanlar sorgusu . . . . .	26
<b>Tablo 3.4</b>	Uzun boylu insanların üyelik dereceleri . . . . .	27
<b>Tablo 3.5</b>	Uzun boylu ve şişman insanların üyelik dereceleri . . . . .	28
<b>Tablo 3.6</b>	Örnek veri tabanı . . . . .	31
<b>Tablo 3.7</b>	Bilimsel Fonksiyonel Bağlılık Değerleri . . . . .	32
<b>Tablo 3.8</b>	Örnek veritabanı . . . . .	35
<b>Tablo 3.9</b>	Dereceli Fonksiyonel Bağlılık Değerleri . . . . .	37

## ŞEKİL LİSTESİ

Şekil 1.1	: Veri madenciliği çalışmasında kullanılan metodoloji. . . . .	7
Şekil 2.1	: 3'e yakın sayıların bulanlık kümelere farklı gösterimi . . . . .	12
Şekil 2.2	: (a) Yürek fonksiyonu (b) Üçgen fonksiyonu . . . . .	12
Şekil 2.3	: Üçgen fonksiyonda (a) Artan yapı (b) Azalan yapı . . . . .	13
Şekil 2.4	: Sigmoid fonksiyonda (a) S yapı (b) Z yapı . . . . .	14
Şekil 2.5	: $\pi$ yapı . . . . .	15
Şekil 2.6	: A ve B kümelerinin üyelik fonksiyonlarında a) Birleşimi b) Kesişimi . . . . .	18
Şekil 2.7	: A ve A' üyelerinin üyelik fonksiyonları ile gösterimi . . . . .	18
Şekil 3.1	: Bulanık kavram olarak (a) uzun boylu insanlar (b) şişman insanlar . . . . .	24
Şekil 3.2	: Çoğu bulanık fonksiyonu. . . . .	27
Şekil 4.1	: Şirket veritabanının içerdiği tablolar . . . . .	39
Şekil 4.2	: Yürek Fonksiyonu . . . . .	40
Şekil 4.3	: Bulanık_T Tablosunun içeriği . . . . .	40
Şekil 4.4	: Nicelci Tablosunun içeriği . . . . .	41
Şekil 4.5	: Kayıtlar Tablosunun içeriği . . . . .	41
Şekil 4.6	: VTabanı_D Tablosunun içeriği . . . . .	41
Şekil 4.7	: Hisse Senetleri Tablosunun içeriği . . . . .	42
Şekil 4.8	: Açılış Ekranı . . . . .	42
Şekil 4.9	: Bulanık sorgulama ekranı . . . . .	43
Şekil 4.10	: Dışel Özet ekranı . . . . .	44
Şekil 4.11	: BFB ve DFB ekranı . . . . .	45
Şekil 4.12	: Yeni Bulanık Alan Ekleme . . . . .	46
Şekil 4.13	: Bulanık Alan Değerlerini Değiştirme . . . . .	46



## SEMBOL LİSTESİ

- $\mu_A$  : A bulan k kümesi ni n üyeli k derecesi  
 $Q$  : Veritabanı ndaki herhangi bir nesne  
 $A_j$  : Veritabanı ndaki alan lar  
 $Q.A_j$  :  $Q$  nesnesi ni n  $A_j$  alanı ndaki değeri  
 $S$  : Özetleyici  
 $Q$  : Ncele yici  
 $\tau$  : Bir kural ın doğruluk değeri  
 $\approx$  : Bul an k benzerlik operat ör ü  
 $G$  : Dereceli terim

## ÖZET

### BULANIK MANTIĞIN VERİ MADENCİLİĞİNE UYGULANMASI

Son yıllarda, veri tabanı sistemlerinin kullanım gittikçe artmaktadır. Bu artış beraberinde potansiyel değerli veriler içerebilecek veri yığınlarını yaratmıştır. Geleneksel sorgu veya raporlama araçlarının veri yığınları karşısında yetersiz kalması, veritabanlarında bilgi keşfi yada veri madenciliği adı altında, sürekli ve yeni arayışlara neden olmaktadır.

Örüntü tanıma ve sınıflama problemleri üzerinde yoğunlaşan yapay zeka ve istatistik disiplinlerindeki gelişmeler veri madenciliğinin temellerini oluşturmaktadır. Ayrıca veri madenciliği, uzman sistemler, veri tabanları, makine öğrenimi, optimizasyon, görselleştirme, yüksek performanslı paralel işlemler gibi çeşitli disiplin ve teknolojilerdeki gelişmelerden de etkilenmektedir.

Klasik kümeleme üye olma ve üye olmama ilişkisi çerçevesinde geliştirilmiştir. Bu tür kümelelerdeki elemanların üyeliklerini ifade edebilmek için 0 veya 1 değerlerinden biri kullanılırken, bulanık küme teorisi bu üyelik değerleri 0 ile 1 arasında değişmektedir. Bulanık kümeleme bu özelliği sayesinde günlük hayatta kullanılan kümeleme matematiksel olarak ifade edebiliriz. Bunun bir sonucu olarak da veritabanlarından daha esnek sorgulamalar yapabiliriz. Ayrıca yine bulanık kümeleme kullanarak veri tabanlarının alanları arasındaki ilişkileri inceleyip kuralları çıkarabiliriz.

Bu tezde, bulanık mantık ve bulanık kümeleme teorisinden yararlanılarak bir uygulamaya geliştirilmiştir. Bu uygulama Windows platformunda, Microsoft Visual Basic 6.0 ve Microsoft Access kullanılarak yazılmış olup, kesin olmayan, belirsiz dilsel sorgu cümleleriyle, sıradan veritabanı sistemleri üzerinde bulanık sorgulamayı yapabilmektedir. Ayrıca, bu yeni sorgulama diliyle benzer örüntüler aranabilir, bulanık kuralları çıkarılabilir, bulanık fonksiyonel bağımlılık ve dereceli fonksiyonel bağımlılıklar hesaplanabilir.

## **SUMMARY**

### **APPLICATION OF FUZZY LOGIC ON DATA MINING**

In recent years, the use of database systems has been increasing rapidly. This rise has created mountains of data that contain potentially valuable knowledge. Because of the insufficiency of conventional query systems and reporting tools against these data stacks, a new field has grown to satisfy this need which called data mining or knowledge discovery in databases.

The development in artificial intelligence and statistics disciplines which become dense in pattern recognition and classification problems, form the fundamentals of data mining. Besides these disciplines, data mining has also been affected by the development in expert systems, databases, machine learning, optimization, visualization and massively parallel processing systems.

Classical sets have been developed with the relation of being a member of the set or not being a member of the set. In these kind of sets a value of 1 or 0 is being used to express the membership, while in fuzzy set theory this membership value is between 1 and 0. This property of the fuzzy sets, makes easier to express daily used sentences as mathematical expressions. As a result of this property, we can make more flexible queries in databases. Also, with the help of fuzzy set theory, we can easily find the relation between the fields of databases and turn them into rules.

In this thesis, a fuzzy tool has been developed by using fuzzy logic and fuzzy sets. Microsoft Visual Basic 6.0 and Microsoft Access is used in Windows platform for developing this application. With this tool, we can define indefinite, uncertain linguistic query sentences as queries in conventional databases. Furthermore this new query language makes it possible to search for patterns like typical values, fuzzy rules, fuzzy functional dependencies and gradual functional dependencies.

# 1. VERİ MADENCİLİĞİ

## 1.1 Giriş

Günümüzde, teknolojinin çok hızlı bir şekilde gelişmesi ve firmalar arasındaki rekabetin sadece kaliteli ürün üretmekle sınırlı kalmasıyla bir sonucu olarak, firmalar piyasada daha iyi bir yer edinebilmek için müşteri memnuniyetini en üst düzeye çıkarmayı hedeflemektedirler. Bunu yapabilmek için de, bilgisayar sistemlerinin geçmişe oranla çok daha ucuz ve güçlü olmasıyla da etkisiyle bilgisayar sistemlerinden en etkin biçimde faydalanmaktadır.

Örneğin eski den süper marketteki kasa basit bir toplama makinesinden ibaretken ve sadece müşterinin o anda satın aldığı malların toplamını hesaplamak için kullanılırken, günümüzde kasa yerine kullanılan satış noktası terminaleri sayesinde bu hareketlerin bütün detayları saklanabilmektedir. Saklanan bu binlerce malın ve binlerce müşterinin hareket bilgileri sayesinde her malın zaman içindeki hareketleri ve eğer müşteriler bir müşteri numarası ile kodlanmışsa bir müşterinin zaman içindeki verilerine ulaşmak ve analiz etmek de kolay hale gelmiştir.

Bu analizlerin sonucu olarak her mal için bir sonraki ayın satış tahminleri çıkarılabilir; müşteriler satın aldıkları mallara bağlı olarak gruplanabilir; yeni bir ürün için potansiyel müşteriler belirlenebilir; müşterilerin zaman içindeki hareketleri incelenerek onların davranışları ile ilgili tahminler yapılabilir. Binlerce malın ve müşterinin olabileceği düşünülürse bu analizin elle ve gözle yapılamayacağı, otomatik olarak yapılmasının gerektiği ortaya çıkmaktadır. Veri madenciliği burada devreye girer. Veri madenciliği, büyük miktarda veri içinden gelecekle ilgili tahmin yapabilmeyi sağlayacak bağlantı ve kuralların bilgisayar programları kullanarak aranmasıdır.

Veri madenciliği, astronomi, biyoloji, finans, pazarlama, sigorta, tıp ve birçok başka dalda uygulanmaktadır. Son 20 yıldır Amerika Birleşik Devletleri'nde çeşitli veri

madenciliđ algoritmalarının gizli dinlemeden, vergi kaçakçılıklarının ortaya çıkartılmasına kadar çeşitli uygulamalarda kullanıldığı bilinmektedir.

Bununla birlikte günümüzde veri madenciliđ teknikleri özellikle işletmelerde çeşitli alanlarda başarı ile kullanılmaktadır. Bu uygulamaların başlıcaları ilgili alanlara göre özetlenmiştir.

#### *Pazarlama*

- Müşterilerin satın alma örüntülerini belirlemesi,
- Müşterilerin demografik özellikleri arasındaki bağlantıların bulunması,
- Posta kampanyalarında cevap verme oranının artırılması,
- Mevcut müşterilerin elde tutulması, yeni müşterilerin kazanılması,
- Pazar sepeti analizi (*Market Basket Analysis*)
- Müşteri ilişkileri yönetimi (*Customer Relationship Management*)
- Müşteri değerlendirme (*Customer Value Analysis*)
- Satış Tahmini (*Sales Forecasting*).

#### *Bankacılık*

- Farklı finansal göstergeler arasında gizli korelasyonların bulunması,
- Kredi kartı dolandırıcılıklarının tespiti,
- Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi,
- Kredi taleplerinin değerlendirilmesi.

#### *Sigortacılık*

- Yeni poliçe talep edecek müşterilerin tahmin edilmesi,
- Sigorta dolandırıcılıklarının tespiti,
- Riskli müşteri örüntülerini belirlemesi.

## **1.2 Veri Madenciliđ Modelleri**

Veri madenciliđinde kullanılan modeller, tahmin edici (*Predictive*) ve tanımlayıcı (*Descriptive*) olmak üzere iki ana başlık altında incelenmektedir.

Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak sonuçları bilinmeyen veri kümeleri için sonuç değerleri tahmin edilmesi amaçlanmaktadır. Örneğin bir banka önceki dönemlerde vermiş olduğu kredilere ilişkin gerekli tüm verilere sahip olabilir. Bu verilerde bağımsız değişkenler kredi alan müşterinin özellikleri, bağımlı değişken değeri ise kredinin geri ödenip ödenmediğidir. Bu verilere uygun olarak kurulan model, daha sonraki kredi taleplerinde müşteri özelliklerine göre verilecek olan kredinin geri ödenip ödenmeyeceğinin tahmininde kullanılmaktadır.

Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılabilecek mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır. X Y aralığında geliri ve iki veya daha fazla arabası olan çocuklu aileler ile, çocuğu olmayan ve geliri X Y aralığından düşük olan ailelerin satın alma örüntülerinin birbirlerine benzerlik gösterdiğinin belirlenmesi tanımlayıcı modellere bir örnektir.

Veri madenciliği modellerini gördükleri işlemlere göre,

- Sınıflama (*Classification*) ve Regresyon (*Regression*),
- Kümeleme (*Clustering*),
- Birliklik Kuralları (*Association Rules*) ve Ardışık Zamanlı Örüntüler (*Sequential Patterns*),

olarak üzere üç ana başlık altında incelemek mümkündür. Sınıflama ve regresyon modelleri tahmin edici, kümeleme, birliklik kuralları ve ardışık zamanlı örüntü modelleri tanımlayıcı modellerdir.

### 1.2.1 Sınıflama ve Regresyon Modelleri

Mevcut verilerden hareket ederek geleceğin tahmin edilmesi için faydalanan ve veri madenciliği teknikleri içerisinde en yaygın kullanıma sahip olan sınıflama ve regresyon modelleri arasındaki temel fark, tahmin edilen bağımlı değişkenin kategorik veya süreklilik gösteren bir değere sahip olmasıdır. Ancak çok terimli lojistik regresyon (*multinomial logistic regression*) gibi kategorik değerlerin de tahmin edilmesine olanak sağlayan tekniklerle, her iki model giderek birbirine yaklaşmakta ve bunun bir sonucu olarak aynı tekniklerden yararlanılması mümkün olmaktadır. Sınıflama ve regresyon modellerinde kullanılan başlıca teknikler,

- Karar Ağaçları (*Decision Trees*),

- Yapay Sınır Ağları (*Artificial Neural Networks*),
- Genetik Algoritmalar (*Genetic Algorithms*),

K-En Yakın Komşu (*K-Nearest Neighbor*),

- Nai ve- Bayes.

### 1.2.2 Kümeleme Modelleri

Kümeleme modellerinde amaç, üyelerinin birbirlerine çok benzediği, ancak özelliklerinin birbirinden çok farklı olduğu kümelere bulunması ve veri tabanındaki kayıtların bu farklı kümelere bölünmesidir. Başlangıç aşamasında veri tabanındaki kayıtların hangi kümelere ayrılacağı veya kümelere hangi değişken özelliklerine göre yapılacağı bilinmemekte, konunun uzmanı olan bir kişi tarafından neler olacağı tahmin edilmektedir.

### 1.2.3 Birliklik Kuralları ve Ardışık Zamanlı Örneklemler

Bir alışveriş sırasında veya birbirini izleyen alışverişlerde müşterinin hangi mal veya hizmetleri satın almaya eğilimi olduğunun belirlenmesi, müşteriye daha fazla ürünün satılmasını sağlama yollarından biridir. Satın alma eğiliminin tanımlanmasını sağlayan birliklik kuralları ve ardışık zamanlı örneklemler, pazarlama amaçlı olarak pazar sepeti analizi adı altında veri madenciliğinde yaygın olarak kullanılmaktadır. Bununla birlikte buteknikler tıp finans ve farklı olayların birbirleri ile ilişkili olduğunun belirlenmesi sonucunda değerli bilgi kazanımının söz konusu olduğu ortamlarda da önem taşımaktadır.

### 1.3 Örnek Uygulamalar

Yukarıda anlatılan modellerde kullanılan tekniklerin bir kısmı ile ilgili örnekler aşağıda verilmiştir.

- BAĞINTI: “Çocuk bezi alan müşterilerin %30’u bira da satın alır.”

Sepet analiziinde müşterilerin beraber satın aldığı malların analizi yapılır. Buradaki amaç mallar arasındaki pozitif veya negatif korelasyonları bulmaktır. Çocuk bezi alan müşterilerin mama da satın alacağını veya bira satın alacakları varsayarak tahmin edebiliriz. Ama sadece otomatik bir analiz bütün olasılıkları gözönüne alır ve kolay düşünülmeyecek, örneğin çocuk bezi ve bira arasındaki bağıntıları da bulur.

- SİNFLANDIRMA: “Genç kadınlar küçük araba satın alır, yaşlı, zengin erkekler büyük, lüks araba satın alır.”

Amaç bir malın özellikleri ile müşteri özelliklerini eşleştirmektir. Böylece bir müşteri için ideal ürün veya bir ürün için ideal müşteri profili çıkarılabilir. Örneğin bir otomobil şirketi geçmiş müşteri hareketlerinin analizi ile yukarıdaki gibi iki kural bulursa; genç kadınların okuduğu bir dergiye reklam verirken, küçük modelli arabalarını bu dergiye koymayı tercih edecektir.

- REGRESYON: “Ev sahibi olan, evli, aynı iş yerinde beş yıldan fazladır çalışan, geçmiş kredilerinde geç ödemesi bir ayı geçmemiş bir erkeğin kredi skoru 825’dir.”

Başvuru skorlarında (*application scoring*) bir finans kurumuna kredi için başvuran kişi ile ilgili finansal güvenilirliğini notlayan örneğin 0 ile 1000 arasında bir skor hesaplanır. Bu skor kişinin özelliklerine ve geçmiş kredi hareketlerine dayalı olarak hesaplanır.

- ZAMAN İÇİNDE SİRALI ÖRÜNTÜLER: “İlk üç taksitinden iki veya daha fazlasını geç ödemiş olan müşteriler %60 olasılıkla kanuni taksit beğeniyor.”

Davranış skoru (*behavioral score*), başvuru skorundan farklı olarak kredi almış ve taksitleri ödeyen bir kişinin sonraki taksitlerini ödeme/geçiktirme davranışını notlamayı amaçlar.

- BENZER ZAMAN SİRALARI: “X şirketinin hisse fiyatları ile Y şirketinin hisse fiyatları benzer hareket ediyor.”

Amaç zaman içindeki iki hareket serisi arasında bağıntı kurmaktır. Bunlar iki malın zaman içindeki satış miktarları olabilir. Örneğin dondurma satışları ile kola satışları arasında pozitif, dondurma satışları ile sahlep satışları arasında negatif bir bağıntı beklenebilir.

- İSTİSNALAR (FARK SAPTANMASI): “Normalden farklı davranış gösteren müşterilerim var mı?”

Amaç önceki uygulamaların aksine kural bulmak değil, kurala uymayan istisnai hareketleri bulmaktır. Bu da olası sahtekarlıkların saptanmasını (*fraud detection*) sağlayabilir. Örneğin Visa kredi kartı için yapılan CRI S sisteminde, bir yapay sinir ağı kredi kartı hareketlerini takip eder ve müşterinin normal davranışına uymayan



hareketler gözleđinde müşteri nın bankası ile teması geçerek müşteri onayı isteyebilir.

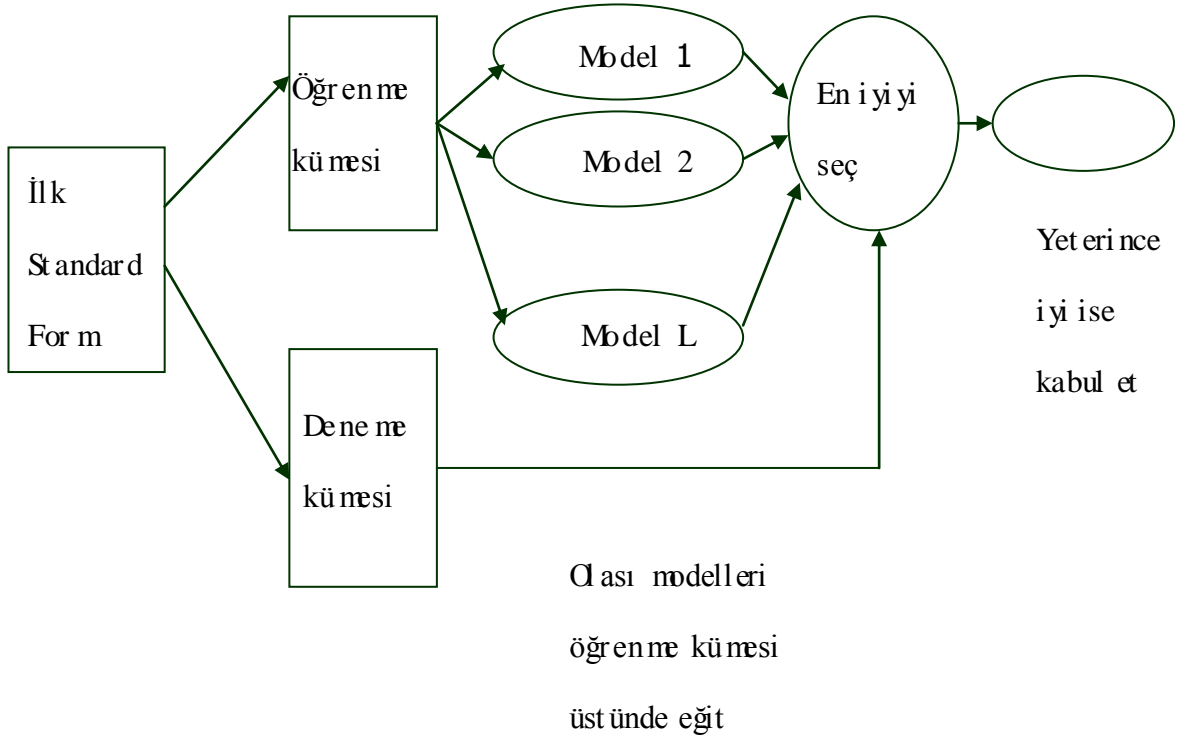
- DÖKÜMAN MADENCİLİ Ğ: “Arşivi nde (veya internet üzerinde) bu dökümana benzer hangi dökümanlar var?”

Amaç dökümanlar arasında ayrıca elle bir tasnif gerekmeden benzerlik hesaplayabilmektir (*text mining*). Bu genelde otomatik olarak çıkarılan anahtar sözcüklerin tekrar sayısı sayesinde yapılır.

#### **1.4 Metodoloji**

Bir veri madenciliđ çalışmasında kullanılan metodoloji Şekil 1.1’de verilmiştir. Standart formiçinde verilen veri, öğrenme ve deneme olmak üzere ikiye ayrılır. Her uygulamada kullanılacak birden çok teknik vardır ve önceden hangisinin en başarılı olacağını kestirmek olası değildir. Bu yüzden öğrenme kümesi üzerinde L değişik teknik kullanılarak L tane model oluşturulur. Sonra bu L model deneme kümesi üzerinde deneyerek en başarılı olanı, yani deneme kümesi üzerindeki tahmin başarısı en yüksek olanı seçilir.

Eğer bu en iyi model yeterince başarılıysa kullanılır, aksi takdirde başa dönülerek çalışma tekrarlanır. Tekrar sırasında başarısız olan örnekler incelenerek bunlar üzerindeki başarının nasıl arttırılabileceđi araştırılır. Örneğin standart forma yeni alanlar ekleyerek programa verilen bilgi arttırılabilir; veya olan bilgi değişik bir şekilde kodlanabilir; veya amaç daha değişik bir şekilde tanımlanabilir.



Şekil 1.1 : Veri madenciliği çalışmasında kullanılan metodoloji.

## 2. VERİ MADENCİLİĞİNDE BULANIK MANTIK YAKLAŞIMI

### 2.1 Giriş

Veri tabanı sistemleri, oluşumlarından beri pek çok konuya başlık olmuş ve pek çok teknolojinin gelişmesine sebep olmuştur. 1970'li yıllarda veritabanları sadece kayıtlarımızı tutmaya ve onlardan basit raporlar alabilmemize yararırken, günümüzde teknolojinin de gelişmesiyle çok daha detaylı sorgulara cevaplar alabileceğimiz, gelecek hakkında bize fikirler verebilecek sistemler haline dönüşmüşlerdir. Tablo 2.1 de bu teknolojik evrim daha detaylı bir şekilde gösterilmektedir.

**Tablo 2.1** Veri Tabanı Teknolojisinin Evrimi

Evri msel Adım	Örnek İşletme Sorusu	Çanak Sağlayan Teknoloji	Ürün Sunan Firmalar	Özellikler
Veri Toplama (1970)	Son 5 yıl daki geliri mne kadar?	Veri Tabanı	IBM	Statik yapı, geriye bakmayı sağlıyor.
Veri Erişimi (1980)	Geçen yıl mart ayında satışları mne oldu?	İlişkisel Veri Tabanları, SQL	Oracle, Microsoft, Informix	Dinamik yapı, kayıt seviyesinde veriye erişim
Veri Anbarı (1995)	Son üç ayda Ankara'daki X mağazada Y ürününden ne kadar satmış m?	OLAP	IBM Cognos	Detaylı sorgulama, geçmişe ve şimdiki zaman dönmek
Veri Madenciliği	Gelecekte Ankara'daki satışları mne kadar olacak?	İleri algoritmalar, istatistik, yapay zeka	IBM SPSS	Veriler arasındaki gizli benzerliklerin keşfi.

Burada di kkat edil mesi gereken şey, ister veritabanından çok basit bir sorgu yaparken, ister de bir OLAP ( *Online Analytical Process*) aracı ile çok karmaşık bir sorgu la ma yaparken, yapılan bütün sorgu la ma ların kesin ve katı kurallar içer mesi dir.

Fakat günlük hayatı mızda kur duğu mız cümlelere di kkat edilirse, böyle kesin ve katı kurallar yada kararlar yoktur. Bunlar yerine daha çok belirsizlik içeren cümleler vardır. M esela, bir basketbol koçu takı mna oyuncu seçerken uzun boylu insanları seçmek isteyecektir. Fakat oyuncuların kayıtları bilgisayarda tutuluyorsa ve veritabanından oyuncular seçilecekse yapılacak SQL sorgusu katı kurallar içerecektir. Yapılabilecek sorgu boyu 1.9m den uzun oyuncuları görüntüle gi bi bir sorgu olacaktır. Bu durumda da 1.89m yada 1.88m boyunda olanlar sorgu la ma sonucunda görüntülenmeyecektir. Böylece belki de çok iyi oyuncular basketbol koçunun gözünden kaçabilecektir. Bunun yerine bulanık mantık ve bulanık kü neleri kullanarak veritabanından uzun boylu insanları göster di ye bir sorgu yapılırsa bütün herkes üyelik dereceleri sırasıyla koçun karşısına çıkacaktır ve böylece de koç takı mna oyuncu seçerken daha az hata yapmış olacaktır.

Gör üldüğü gi bi insan dili belirsizlikler içerektedir. Belirsizliğe, istenmeyen bir durum anlayışıyla yaklaşmak yerine, büyük bir fayda alanı açan ve üzerinde çalışıl ması gereken önemli bir durum anlayışıyla yaklaşmak daha doğrudur.

Yukarıdaki basketbol örneğinde de olduğu gi bi sorgu la ma lara belirsizliği katarak veritabanlarından bulanık sorgu la ma lar yapmak daha iyi sonuçlar verir. Bununla birlikte, sadece bulanık sorgu la ma yaparken değil, veritabanlarından bilgi keşfi yaparken yada alanlar arasındaki ilişkiler incelenirken de bulanık mantık etkili bir biçimde kullanılabilir.

Bu çalış mada, bulanık mantık ve bulanık kü ne teorisi nden yararlanılarak, Windows platformunda, Microsoft Visual Basic 6.0 ve Microsoft Access kullanılarak geliştirilmiş, kesin olmayan, belirsiz dilsel sorgu cümleleriyle, sıradan veritabanı sistemleri üzerinde bulanık sorgu la ma yapabilen, veritabanlarından kurallar çıkarabilen ve veritabanlarının alanları arasındaki ilişkileri inceleyen bir araç oluşturul muştur.

Programın son kısmında bahsedilen ve alanlar arasındaki ilişkileri bulanık mantık ve bulanık kü ne teorisi ne göre inceleyen bölümde örnek veritabanı olarak İMKB' deki (İstanbul Menkul Kıymetler Borsa) hisse senetlerinin günlük kapanış değerleri

kullanılmıştır. Bu uygulamanın sonucu olarak “X hisse senedi nin fiyatı artarken, Y hisse senedi nin de fiyatı artar” yada “X hisse senedi ile Y hisse senedi nin hareketleri arasında %80’lik bir benzerlik vardır” gibi kurallar üretilebilecektir.

Programın yapısı, işleyişi ve örnek uygulamalar Bölüm 4’te ayrıntıları ile açıklanacaktır.

## 2.2 Bulanık Kümeler

Bulanık küme kavramını anlayabilmek için klasik anlamdaki küme kavramının, elinizdeki bilgilerin değerlendirilmesinde yetersiz kaldığı yerler üzerinde biraz düşünmemiz yeterlidir. Bir örnek vermek gerekirse,  $X \subset \mathcal{R}$  olsun ve  $X$  sifıra yakın reel sayıların kümesini göstereyin. Klasik küme tanımından yola çıktığımızda herhangi bir atama değerini, kümeye ait olma ya da olmama özelliği olarak tespit edip, bu aralıktaki bütün sayıları  $X$ ’in elemanı, dışındakileri de  $X$ ’in elemanı değil şeklinde nitelendirme durumundayız. Sorun bu değer nasıl belirleneceğinden ziyade belirlendikten sonra, belirlenen değer çok yakındaki sayılara haksızlık yapıldığı konusundaki itirazlar olacaktır. Yani  $\pm 2$  atama değeri olarak seçilirse  $\pm 2.001$  niye kümenin elemanı olmasın. Bu aynı zamanda kümenin yeterince iyi tanımlanması anlamına da gelmektedir. İşte bu tip sorunları ortadan kaldirabilmek için bulanık küme teorisi kullanılabilir.

Klasik kümeler üye olma ve üye olmama ilişkisi çerçevesinde geliştirilmiştir. Bu tür kümeleri ifade etmek için özel bir fonksiyon tanımlanabilir ve bu fonksiyona karakteristik fonksiyon denilir. Karakteristik fonksiyon her bir elemana 1 ve 0 değerlerinden birini üyelik durumuna göre atayarak evrensel küme üzerinde tanımlanan ve bizim ilgilendiğimiz özelliğe sahip olan elemanların oluşturduğu kümeyi belirler.

Örneğin,  $X$  evrensel kümesi üzerinde belirli bir özelliği taşıyan elemanları ayırarak oluşturduğumuz  $A$  kümesini karakteristik fonksiyon yardımıyla,  $\forall x \in X$  için

$$X_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} \quad (2.1)$$

verilir. Fonksiyon  $A$  kümesine ait elemanlara 1 değerini, ait olmayan elemanlara ise 0 değerini atamaktadır.

Bu fonksiyon evrensel kümenin elemanlarına belirli bir aralıkta olmak üzere ve gözlem altındaki kümenin elemanlarının üyelik derecelerini ifade edecek biçimde genelleştirilebilir. Daha yüksek değerler üyelik derecesinin yüksekliğini gösterir. Bu fonksiyona üyelik fonksiyonu ve bu fonksiyonun oluşturduğu kümeye bulanık küme denir.

$X$  boş olmayan bir küme olsun.  $X$  deki bir Bulanık  $A$  kümesi  $\forall x \in X$  için;  $A: X \rightarrow [0, 1]$  fonksiyonu ile verilir.  $A$  ya bulanık kümeye karşılık gelen üyelik fonksiyonu adı verilir. Bulanık  $A$  kümesi ise  $X$  deki her elemanın üyelik derecesiyle birlikte oluşturduğu kümedir.  $x$ 'in  $A$  ya ait olma veya üyelik derecesi  $A(x)$  olarak okunur. ( $\mu_A$  olarak da gösterilebilir)

Üyelik fonksiyonunun değer aralığı için yaygın olarak  $[0, 1]$  aralığı kullanılır. Bu durumda, her üyelik fonksiyonu bir klasik evrensel kümenin elemanlarını bu aralıktaki bir sayıya karşılık getiren bir fonksiyondur.

Bulanık küme kavramının temsili için genel olarak;

$$\mu_A: X \rightarrow [0, 1] \quad (2.2)$$

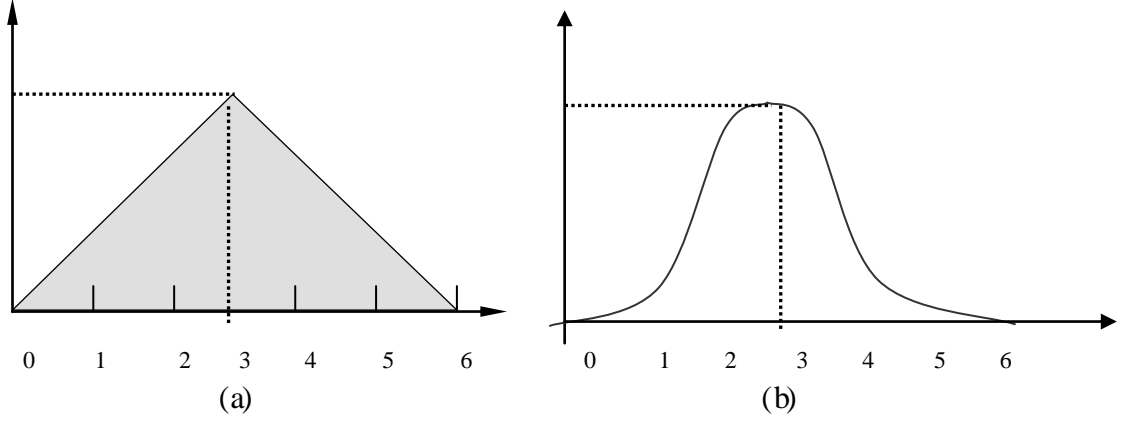
gösterimi kullanılır.

Bulanık kümeler daha önce de ifade edildiği üzere doğal dildeki belirsiz ve bulanık kavramları temsil etmeye ve onları matematiksel olarak ifade etmeyi mümkün kılarlar. Bu temsil işlemi sadece kavramın kendisine değil kavramın kullanıldığı alana da bağlıdır. Örneğin, “yüksek ısı” kavramı hava olayları için kullanıldığı anda oluşturulacak bulanık kümeyle, bir nükleer reaktör içindeki ıssıyı ifade etmede kullanıldığı anda oluşturulacak bulanık kümenin birbirinden tamamen farklı olması gerektiği açıktır. Hatta aynı alan içerisinde ki aynı kavramın farklı bulanık kümelerle ifade edilebileceği durumlarının olması inkan dahilindedir. Fakat, böylesi durumlarda bu farklı kümelerin bazı temel nitelikler bakımından birbirine benzerliği gerekir.

## 2.2.1 Üyelik Fonksiyonları

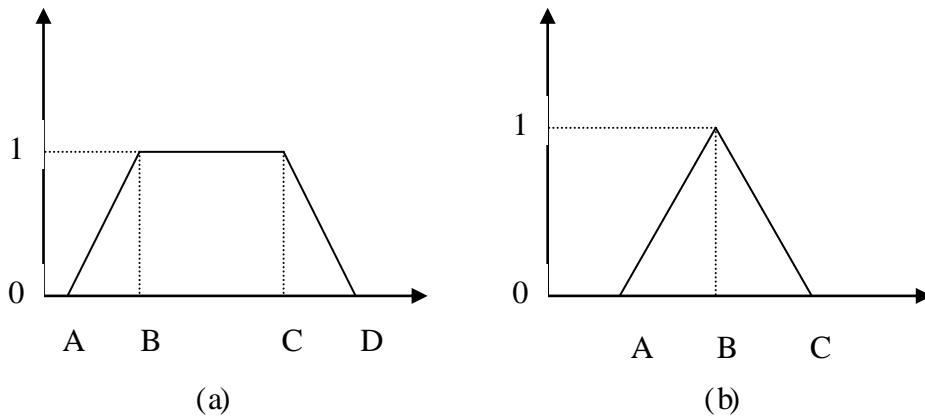
Üyelik fonksiyonları bir çok farklı şekillerde olabilir. Özel bir şeklin uygun olup olmayacağını tespit etmek çalışılan uygulamaya alan tarafından elde edilen verilerle belirlenir. Fakat, birçok uygulamaya bu tür şekil değişikliklerine karşı çok fazla duyarlılık göstermezler. Hesaplama açısından getirdiği kolaylıklar göz önüne

alı narak istenilen şekilde üyelik fonksiyonunun seçilmesi, bulanık küme teorisinin esnekliğini yansıttığına örneğin bir durumdur. Çoğu durumda, üçgen ve yamuk üyelik fonksiyonları işimizi görecektir niteliklere sahiptir.



**Şekil 2.1** : 3'e yakın sayıların bulanık kümelerinde farklı gösterimleri

Şekil 2.1'de, "3'e yakın reel sayılar kümesi" kavramını temsil eden üyelik fonksiyonuna ait iki örnek vardır. Açıkça görülmektedir ki bulanık kümelerin kullanılabilirliği büyük oranda bizim farklı kavramlara uygun üyelik derecesi fonksiyonlarını oluşturabilme becerimize dayanmaktadır. Aşağıda bazı çok kullanılan fonksiyon tipleri ve bunların analitik olarak nasıl hesaplanabileceği gösterilmiştir.

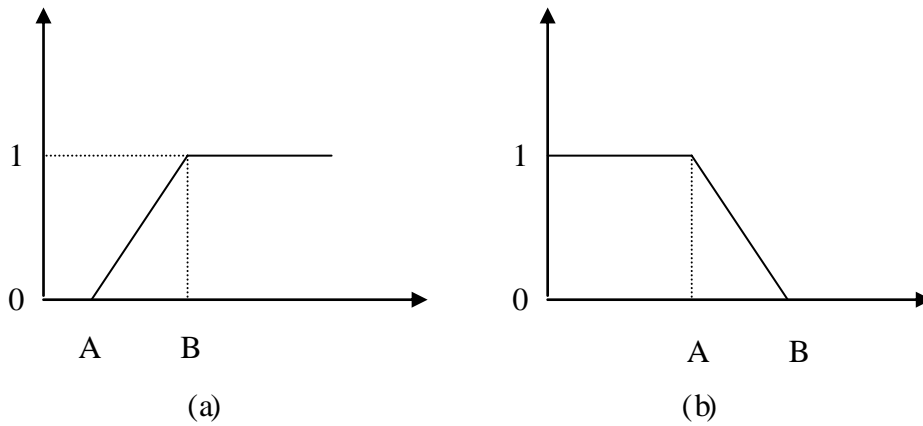


**Şekil 2.2** : (a) Yamuk fonksiyonu (b) Üçgen fonksiyonu

$$\text{Ya mık}(x; A, B, C) : \begin{cases} 0 & ; x < A \\ \frac{x-A}{B-A} & ; A \leq x < B \\ 1 & ; B \leq x \leq C \\ \frac{D-x}{D-C} & ; C \leq x < D \\ 0 & ; x \geq D \end{cases} \quad (2.3)$$

$$\text{Üçgen}(x; A, B, C) : \begin{cases} 0 & ; x < A \\ \frac{x-A}{B-A} & ; A \leq x < B \\ 1 & ; x = B \\ \frac{D-x}{D-C} & ; C \leq x < D \\ 0 & ; x \geq D \end{cases} \quad (2.4)$$

Ya mık ve üçgen yapıdaki üyelik fonksiyonlarının for mülleleri sırasıyla (2.3) ve (2.4), şekilleri ise Şekil 2.2'de gösterilmiştir. Ya mık yada üçgen yapıdan sadece artan yada sadece azalan yapılarda geliştirilebilir. Bu yapılar Şekil 2.3'de gösterilmiştir.



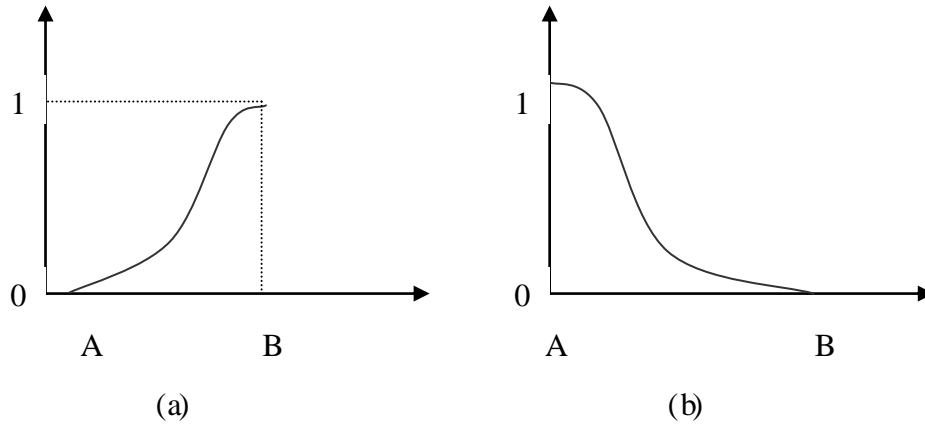
**Şekil 2.3** : Üçgen fonksiyonda (a) Artan yapı (b) Azalan yapı



$$LS(x; A, B) : \begin{cases} 0 & ; x < A \\ \frac{x-A}{B-A} & ; A \leq x < B \\ 1 & ; x \geq B \end{cases} \quad (2.5)$$

$$LZ(x; A, B) : \begin{cases} 1 & ; x < A \\ \frac{B-x}{B-A} & ; A \leq x < B \\ 0 & ; x \geq B \end{cases} \quad (2.6)$$

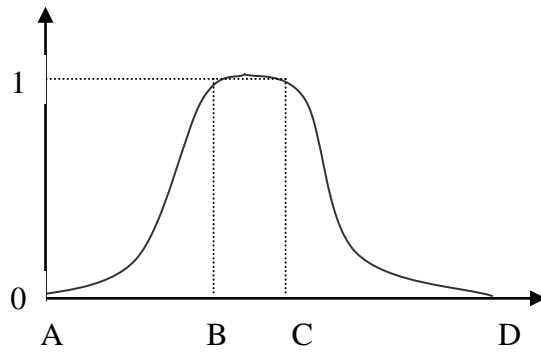
Bir başka çok kullanılan fonksiyon tipi de sigmoid fonksiyon tipidir. Sigmoid fonksiyonda S yapısı ve Z yapısı Şekil 2.4'de, hesaplamaları ise sırasıyla (2.7) ve (2.8) de gösterilmiştir. Bir başka sigmoid fonksiyon tipi olan  $\pi$  yapısı Şekil 2.5'de, hesaplanması ise (2.9) da gösterilmiştir.



Şekil 2.4 : Sigmoid fonksiyonda (a) S yapısı (b) Z yapısı

$$S(x; A, B) : \begin{cases} 0 & ; x \leq A \\ 2 * \left( \frac{x-A}{B-A} \right)^2 & ; A < x \leq (A+B)/2 \\ 1 - 2 * \left( \frac{x-B}{B-A} \right)^2 & ; (A+B)/2 < x \leq B \\ 1 & ; x > B \end{cases} \quad (2.7)$$

$$Z(x; A, B) : \begin{cases} 1 & ; x \leq A \\ 1 - 2^* \left( \frac{x-A}{B-A} \right)^2 & ; A < x \leq (A+B)/2 \\ 2^* \left( \frac{x-B}{B-A} \right)^2 & ; (A+B)/2 < x \leq B \\ 0 & ; x > B \end{cases} \quad (2.8)$$



Şekil 2.5 :  $\pi$  yapı

$$\pi(x; A, B, C, D) : \begin{cases} S(x; A, B) & ; x < B \\ 1 & ; B < x \leq C \\ Z(x; C, D) & ; x > C \end{cases} \quad (2.9)$$

Yukarıda gösterilen bütün fonksiyon tipleri ve bunlarla oluşturulabilecek bulanık kümeler sonsuz elemana sahiptirler. Fakat bazı durumlarda bulanık kümeler sonlu elemana da sahip olabilirler. Bu durumda bu kümeler şu şekilde ifade edilebilirler:

$$A = \sum_{x \in X} \mu_A(x)/x;$$

Örnek olarak uzun boylu insanları bir sonlu bulanlık kümesi olarak şu şekilde ifade edebiliriz

$$\text{Uzun boylu insanlar} = \{ 0 / 160 + 0.2 / 165 + 0.5 / 175 + 0.8 / 185 + 1 / 190 \}$$

Elemanlarla üyelik dereceleri arasındaki bölümlenme işareti ve toplama işareti gerçek kullanımdaki anlamında değildir. Sadece elemanlar arası ilişkileri ve elemanlarla üyelik derecesi arasındaki ilişkiyi ortaya koymak için kullanılırlar. Yukarıdaki örnek için 165cm boyundaki insanlar kümesine 0.2 üyelik derecesiyle dahil edilirken, 185cm boyundaki insanlar kümesine 0.8 üyelik derecesiyle dahil edilirler. Fakat küme sonlu sayıda elemana sahip olduğu için kümeye ait olmayan boy uzunluklarının üyelik dereceleri 0 olur. Yani 188cm yada 202cm boyundaki insanlar kümesine ait olmadıkları için üyelik dereceleri 0'dır.

## 2.2.2 Bulanık Değişkenler

Günlük kullanımda ait olan düşük, orta seviye, yüksek ve bunun gibi kavramları temsil eden çeşitli bulanık kümeler bir değişkenin durumlarını tanımlamak amacıyla kullanılırlar. Bu değişkenlere bulanık değişkenler ve onun alt durumlarına da bulanık terimler denilir. Örneğin “ısı” kavramı kendi içinde çok düşük, düşük, orta seviye, yüksek ve çok yüksek gibi durumlarla nitelenebilen bulanık bir değişken olarak alınabilir.

Bulanık değişkenlerin öneki kavrama ait bu durumlar arasındaki geçişi yansıtabilmesiindeki kolaylıkta yatar. Bu durumda da belirsizlik altında yapılan yöntem ve ölçümlerle ilgilenme ve onları ifade etme diğer yöntemlerden daha başarılı sonuçlar verir. Geleneksel klasik değişkenler ise bu kapasiteden yoksundurlar. Bir durumun klasik değişkenler yardımıyla tanımlanması matematiğin olarak doğru olduğu halde, kaçınılmaz ölçüm hataları içerdiğinden gerçeğe uygunluk gösteremez. Bu şekilde tanımlanmış durumların arasında ki sınırların civarındaki değerler için yapılacak bir ölçüm sadece durumların birini desteklediği bir gözlem olarak alınır. Böylesi bir karar kaçınılmaz olarak belirsizlik içerse dahi matematiğin tanımlar ve hesaplamalar buna göre yapılır. Herhangi bir ölçümün sınırın her iki tarafındaki durumlar için de destekleyici bir gözlem olarak kabul edilmesi gerektiği yer olan bu değerler de belirsizlik en yüksek düzeye ulaşır. Fakat, klasik değişkenler yoluyla işlemler yapılırken bu durumlar dahi gözmezlikten gelir ve ihmal edilir. Yapılan şey, durumlar arası sınır değerini keyfi matematiğin

tanımlanmalarla sadece tek bir duruma aitmiş gibi göstermek ve deneysel ölçümlerde bu değerin tek bir duruma ait destekleyici bir değer olarak kabul edilmesidir.

Bulanık değişkenler, belirsizlikleri deneysel verilerin bir parçası olarak ele aldıklarından dolayı, gerçeğe daha uygundur ve bize olgular hakkında klasik değişkenlere dayanan bilgilerden daha doğru bilgiler verirler. Ünlü fizikçi Einstein bu durumu şu şekilde ifade etmiştir: “Matematiğin kavramları kesin oldukları sürece gerçeği yansıtmazlar, gerçeği yansıttıkları sürece de kesin değillerdir.”

Bulanık kümeler üzerine kurulan matematiksel yapı, klasik matematiğe en fazla açıklayıcı bir güce sahip olmasına rağmen kullanılabilirliği uygulama alanlarında karşımıza çıkan kavramlar için uygun üyelik fonksiyonlarının inşa edilmesine bağlıdır. Fakat, bu işlemi tatmin edici düzeyde gerçekleştirmesi bir çok ek araştırma gerektirir.

### 2.2.3 Temel Bulanık Küme İşlemleri

Bulanık kümelerde, sıradan kümelereki gibi temelde üç işleme sahiptir. Bu işlemler; birleşme, kesişme ve tüme nedir. Bulanık kümeler üyelik fonksiyonları ile tanımlandıklarından bu kümeler üzerindeki işlemler de üyelik fonksiyonları yardımıyla yapılır.

**Birleşme:** A ve B bulanık kümelerin birleşimi  $A \cup B$  kümesi aşağıdaki üyelik fonksiyonu ile tanımlanır. Birleşme işlemi semantik olarak VEYA'ya karşılık gelir. Bu işlem grafiksel olarak Şekil 2.6 (a)'da gösterilmiştir.

$$\mu(A \cup B(x)) = \mu(A(x)) \vee \mu(B(x)) = \text{Max}(\mu(A(x)), \mu(B(x))) \quad (2.10)$$

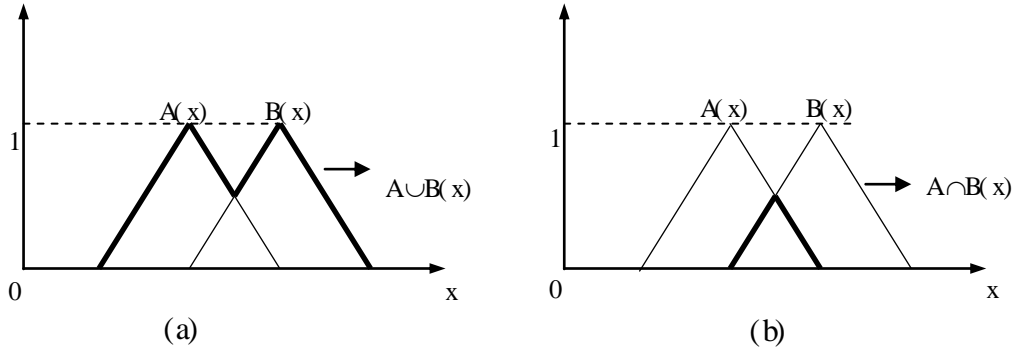
**Kesişme:** A ve B bulanık kümelerin kesişimi  $A \cap B$  kümesi aşağıdaki üyelik fonksiyonu ile tanımlanır. Kesişme işlemi semantik olarak VE'ye karşılık gelir. Bu işlem grafiksel olarak Şekil 2.6 (b)'da gösterilmiştir.

$$\mu(A \cap B(x)) = \mu(A(x)) \wedge \mu(B(x)) = \text{Min}(\mu(A(x)), \mu(B(x))) \quad (2.11)$$

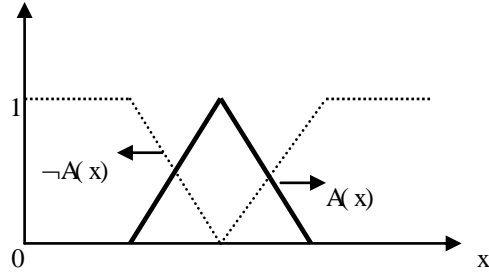
**Tüme ne:** A kümesinin tüme neyi  $A'$  aşağıdaki üyelik fonksiyonu ile tanımlanır. Tüme ne işlemi semantik olarak DEĞİL'e karşılık gelir. Bu işlem grafiksel olarak Şekil 2.7'de gösterilmiştir.

$$\mu(A'(x)) = 1 - \mu(A(x)) \quad (2.12)$$

M n ve max operatörlerini birleşme özelliği dolayısıyla butunlar herhangi sayıda bulanık küme üzerine de genişletilebilirler.



Şekil 2.6 : A ve B kümelerinin üyelik fonksiyonlarında a) Birleşimi b) Kesişimi



Şekil 2.7 : A ve A tünleyen kümelerinin üyelik fonksiyonları ile gösterimi

İki sonlu bulanık küme arasındaki birleşmeyi, kesişmeyi ve bir kümenin tünleyeni ni bir örnekle açıklayalım

$$A = \{0.6/-2 + 0.3/-1 + 0.6/0 + 1.0/1 + 0.6/2 + 0.3/3 + 0.0/4\}$$

$$B = \{0.1/-2 + 0.3/-1 + 0.9/0 + 1.0/1 + 0.9/2 + 0.3/3 + 0.2/4\}$$

ise,

$$A \cup B = \{0.6/-2 + 0.3/-1 + 0.9/0 + 1.0/1 + 0.9/2 + 0.3/3 + 0.2/4\}$$

$$A \cap B = \{0.1/-2 + 0.3/-1 + 0.6/0 + 1.0/1 + 0.6/2 + 0.3/3 + 0.0/4\}$$

$$\neg A = \{0.4/-2 + 0.7/-1 + 0.4/0 + 0.0/1 + 0.4/2 + 0.7/3 + 1.0/4\}$$

$$\neg B = \{0.9/-2 + 0.7/-1 + 0.1/0 + 0.0/1 + 0.1/2 + 0.7/3 + 0.8/4\}$$

olur.

## 2.2.4 Bulanık Kümenin Özellikleri

Klasik kümeler için bilinen tüm işlemler aşağıdaki iki durum hariçinde bulanık kümeler için de geçerlidir.

$$1. A \cup \neg A = X \quad \text{ve} \quad 2. A \cap \neg A = \emptyset$$

Bu iki durum bulanık kümeler için geçerli değildir. Bulanık kümelerde geçerli olan diğer durumlar ise aşağıda sunulmuştur;

A, B ve C Xevrensel kümesi üzerinde tanımlı bulanık kümeler olmak üzere

$$1- A \cup B = B \cup A, \quad A \cap B = B \cap A$$

$$2- (A \cup B) \cup C = A \cup (B \cup C), \quad (A \cap B) \cap C = A \cap (B \cap C)$$

$$3- A \cap A = A, \quad A \cup A = A$$

$$4- \neg \neg A = A$$

$$5- A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$6- \neg(\neg A) = A$$

$$7- \left. \begin{array}{l} \neg(A \cap B) = \neg A \cup \neg B \\ \neg(A \cup B) = \neg A \cap \neg B \end{array} \right\} \text{De' Morgan Kuralları}$$

## 2.3 Bulanık Mantık

Bulanık mantık, adından da anlaşılacağı gibi kesin çıkarımlardan çok yaklaşık çıkarımlar elde edilen mantıktır. Önermeler doğal dille ifade edildiğinde, çoğunlukla doğal dilin yapısı gereği belirsizlikler taşırlar. Dolayısıyla bu çeşit önermelerden kesin çıkarım yapmak olanaklı olmaz. İnsanlar çoğu zaman kesin çıkarım yapmak yerine yaklaşık çıkarım yaparlar. Örneğin insanların boyunun kaç cm ve kaç mm olduğunu bilmeden ve ilgilenmeden, insanların uzun boylu, orta boylu, kısa boylu vs. gibi kümelere ayırırlar. Yani insan düşüncüsü ve çıkarım bulanıktır demek hemen hemen doğrudur.

### 2.3.1 Bulanık Mantığın Tarihi

Matematiğin doğruluğundaki ve bütünlüğündeki başarısında Aristoteles'in ve onun izinden giden düşünürlerin büyük katkısı olmuştur. Onların mantık teorisini oluşturma çabaları ile matematik gelişmiş ve "Düşüncenin Yasaları" oluşturulmuştur. Bu yasalardan biri her önermenin "Doğru" yada "Yanlış" olması gerektiğini öngörmüştür. Bu kavram Perminedes ilk ortaya attığı zaman bile (yaklaşık MÖ 400) karşı görüşlerin oluşması uzun sürmemiştir. Heraclitus bazı şeylerin aynı anda hem doğru olmasının hem de doğru olmasının mümkün olabileceğini savunmuştur.

Bulanık Mantık oluşturacak temel düşünceyi Plato, "Doğru" ve "Yanlış'ın" iç içe girdiği üçüncü bir durumu belirterek oluşturdu. Hegel ve Marx gibi modern düşünürler bu düşünceyi destekledi ancak ilk kez Lukasiewicz Aristoteles'in iki-değerli mantığına sistematik bir alternatif getirdi.

Lukasiewicz 1900'lerin başında 3. bir değer ortaya attı: "olası". Lukasiewicz daha sonra 4., 5., 6. vs. gibi değerleri de oluşturdu ve "Doğru" ile yanlış arasında sonsuz farklı değerler atanabileceğini gösterdi. Lukasiewicz ve onu izleyen diğer matematikçiler bu değerleri nümerik olarak ifade etmiş olsalarda, 1965 yılında Lotfi A. Zadeh, bu değerleri  $[0, 0, 1, 0]$  aralığındaki sayılarla ifade ettiği teorisini "Bulanık Mantık" adlı çalışmasında tanımlayana dek, sonsuz-değerli mantık uygulamada başarılı olamamıştır.

### 2.3.2 Bulanık Mantığın Özellikleri

Bulanık mantığın bazı karakteristik özelliklerini şöyle sıralayabiliriz

- Bulanık mantıkta, kesin çıkarım yaklaşık çıkarımın özel bir hali olarak görülür.
- Herhangi bir mantıksal sistem bulanıklaştırılabilir.
- Bulanık mantıkta bilgi, bir değişkenler kümesi üzerinde tanımlanan bulanık kısıtlar olarak yorumlanır.
- Çıkarım mekaniizması, bulanık kısıtların sağlanması işlemi olarak görülür.

### 2.3.3 Bulanık Mantık ile Klasik Mantığın Farklılıkları

Bulanık mantık, geleneksel mantık sistemlerinden he mana düşünüş şekli olarak he m de detayda farklılıklar gösterir. Bu farklılıkların birkaçını şöyle sıralayabiliriz:

Doğruluk (Truth): İki değerli sistemlerde doğruluk sadece iki değer alabilir. Çok değerli sistemlerde ise doğruluk genellikle sonlu bir kümenin elemanıdır. Bulanık mantıkta doğruluk, verilen kümenin  $[0, 1]$  aralığındaki bir bulanık alt kümesidir.

Yüklemeler (Predicates): İki değerli sistemlerde yüklemeler kesindir. Mesela “Ankara Türkiye’nin başkentidir” klasik mantıkta basit bir önermedir. Fakat günlük hayatta kullandığımız çoğu cümle ni yüklem *aslında* bulanıktır. (büyük, küçük..)

Yüklem Değiştiriciler (Predicate modifiers): Geleneksel sistemlerde kullanılan tek modifikatör *değil* değiştiricisi dir. Bulanık mantıkta buna ek olarak pek çok değiştirici kullanılabilir. Örneğin *çok*, *aşağı yukarı*, *aşırı* vs.

Niceleyiciler (Quantifiers): Geleneksel sistemlerde sadece iki niteleyen kullanılır, *her* ve *bazı*. Bulanık mantıkta ise pek çok niceleyici kullanılabilir. Örneğin *birkaç*, *çoğu*, *sıksık*, *pekçok* vs.

### 2.3.4 Dilsel Değişkenler

Bulanık mantık ve yaklaşık çıkarıma ihtiyaç duyulmasının nedeni daha önce bahsettiğimiz gibi kesin ve belirli bir anlam olmayan veya birden fazla anlama gelebilen dilsel değişkenlerle karşılaşmamızdır. Dilsel değişkenler belirli bir kümenin bulanık alt kümeleri olarak ele alınabilir. Değiştiriciler içi n sıkça kullanılan ve kabul görmüş bazı matematiksel modellerden konsantrasyon, yoğunlaştırma ve genişleme modelleri sırasıyla aşağıdaki gibidir.

- $\mu_{kon}(A)(x) = \mu_A(x)^2$
- $\mu_{yog}(A)(x) = 2 * (\mu_A(x))^2$
- $\mu_{gen}(A)(x) = (\mu_A(x))^{0.5}$

Bu modelleri kullanarak şu değiştiricileri oluşturabiliriz:

- ÇOK (A) = kon(A)
- BİR AZ (A) = gen(A)



- HAFİFÇE A=(BİRAZ A ve DEĞİL ÇOK)

Bu dilsel deęişkenlere örnek olarak, bir kiři 0.6 üyelik derecesiyle çalışanlar kümesine üye ise, “çok çalışkan” ifadesi o kiři için 0.64 lük bir doğruluk payına sahiptir.

### 3. BULANIK SORGULAMA İLE BİLGİ KEŞFİ

#### 3.1 Giriş

Günümüz modern insanının her alışverişinde, her bankacılık işleminde, her telefon görüşmesinde kaydedilen, uzaktan algılayıcılardan, uydulardan toplanan, devlet ve işletme yönetiminde yapılan işlemler sonucunda saklanan veriler her an inanılmaz boyutlarda artmaktadır.

1995 yılında birincisi düzenlenen *Knowledge Discovery In Databases* konferansı bildiri kitabı sunuşunda, enformasyon teknolojilerinin oluşturduğu veri dağları aşağıdaki cümleler ile vurgulanmıştır.

*“Dünyadaki enformasyon miktarının her 20 ayda bir ikiye katlandığı tahmin edilmektedir. Bu ham veri seli ile ne yapmamız gerekmektedir? İnsan gözleri bunun ancak çok küçük bir kısmını görebilecektir. Bilgisayarlar bilgelik pınarı olmayı vaat etmekte, ancak veri selleri ne neden olmaktadır.”*

Veri tabanı sistemlerinin artan kullanımı ve hacimindeki bu olağanüstü artış, organizasyonları elde toplanan bu verilerden nasıl faydalanabileceği problemi ile karşı karşıya getirmiştir. Bunun sonucu olarak veri madenciliği adı altında birçok yeni yöntem ortaya çıkmıştır. Bu çalışmada, bulanık mantık ve bulanık küme kullanılarak, Dan Rasmussen ve R.R. Yager'in 1997 yılında yazdığı "*Fuzzy Query Language For Hypothesis Evaluation*" adlı makale temel alınarak bir uygulama geliştirilmiştir.

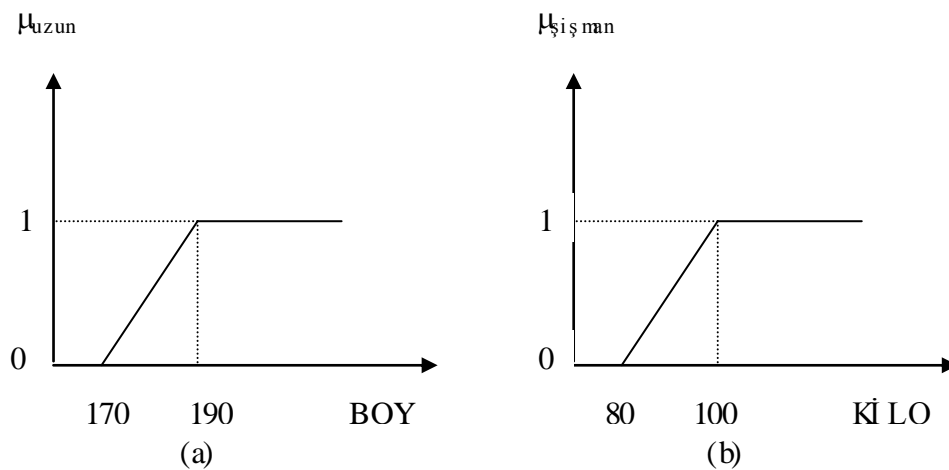
Öncelikle veritabanından bulanık sorgulamaların nasıl yapılacağı ve veritabanlarından kuralların nasıl çıkarılabileceği anlatılacak, daha sonra da, bulanık fonksiyonel ve dereceli bağılılık adı altında, veritabanlarındaki alanlar arasındaki ilişkilerin nasıl hesaplandığı anlatılacaktır.

### 3.2 Bulanık Sorgulama

Veritabanlarından bilgi çekerken kullandığımız sorgu cümleleri bazı katı kuralları içerir. Mesela “Maaşı 1 Milyardan fazla olanları görüntüle” sorgusu standart bir SQL sorgusudur. Bulanık sorgulamalarda ise kısıt kısmı farklıdır. Bir bulanık sorgulamada “uzun, şişman, yaklaşıklık” gibi dilsel ifadeler kullanabiliriz. Böylece oluşturulan sorgunun sonuçları daha işe yarar hale gelir.

Örneğin bir sınıftaki öğrencilerin başarı durumlarını sorgularken “Ortalaması  $\geq 70$  ve devam  $< 10$ ” sorgusu yerine “ortalaması iyi ve devam normal” şeklinde bir sorgu daha kullanışlıdır. Böylece sorgulamanın sonuçlarında 69 ortalamalı ve 9 devamlı bir öğrenci gözümüzden kaçmamış olur.

Bulanık sorgulamanın veri tabanlarında nasıl yapıldığını anlamak için bir örnek vermek gerekirse; VT veritabanı olsun, Q veritabanındaki bir nesne ve  $A = \{a_1, a_2, \dots, a_n\}$  veritabanındaki alanların özellikleri olsun. Bu durumda  $Q.a_j$  bize Q nesnesinin  $a_j$  özelliğinin değerini verecektir. Bunun dışında ayrıca bulanık kısıtları tanımlamamız gerekir.  $\mu_{uzun}(Boy)$  bize Boy alanındaki uzun kavramını tanımlar.  $\mu_{uzun}(1.8)=0.5$  ise 1.8m boyun uzun bulanık kümesine 0.5 üyelik derecesiyle ait olduğunu gösterir. Şekil 3.1(a) uzun boylu insanlar, Şekil 3.1(b) şişman insanlar bulanık kavramlarını tanımlar.



Şekil 3.1 : Bulanık kavram olarak (a) uzun boylu insanlar (b) şişman insanlar

Veritabanındaki her bir bulanık sorgu, veritabanındaki her nesnenin bir üyelik derecesi alması ile sonuçlanır. Burada nesnelere üyelik derecelerine göre sıralamak sonuçları daha iyi ve etkin görmek açısından faydalıdır. Veritabanımız Tablo 3.1'deki gibi üç elemandan oluşsun.

**Tablo 3.1** Örnek veritabanı

İs m	Boy	Kı lo
Can	180	100
Ah met	170	90
Meh met	190	85

“Uzun boyl u insanlar ı göster” sor gusu řu řekilde gösterilir. Üyelik dereceleri Tablo 3.2’de gösteril miřtir.

$$Q_{\text{boy=uzun}}(VT) = \{ Q \in VT \mid \mu_{\text{uzun}}(Q, \text{Boy}) \} \quad (3.1)$$

$$= \{ 1/( Mehmet, 190, 85), 0.5/( Can, 180, 100), 0/( Ah met, 170, 90) \}$$

**Tablo 3.2** Uzun boyl u insanlar sor gusu

İs m	Boy	Kı lo	Üyelik Der.
Meh met	190	85	1
Can	180	100	0.5
Ah met	170	90	0

Sor gu cümleleri bir den fazla cümleden de oluşabilir. Bu durumda cümleler VE yada VEYA bağlaçları ile birleştirilir. Cümleler temel bulanık küme işlemleri di kkat e alınarak birleştirilir. Örnek olarak “uzun ve řiř man insanlar ı göster” sor gusu řu řekil dedir. Sor gunun sonucunda oluş an üyelik dereceleri Tablo 3.3’de gösteril miřtir.

$$Q_{\text{boy=uzun VE kilo=yüksek}}(VT) = \{ Q \in VT \mid \mu_{\text{uzun}}(Q, \text{Boy}) \wedge \mu_{\text{yüksek}}(Q, \text{Kilo}) \}$$

$$Q_{\text{boy=uzun VE kilo=yüksek}}(VT) = \{ Q \in VT \mid \min(\mu_{\text{uzun}}(Q, \text{Boy}), \mu_{\text{yüksek}}(Q, \text{Kilo})) \}$$

$$= \{ \min(1/(\text{Mehmet}), 0.5/(\text{Can}), 0/(\text{Ahmet})), \min(1/(\text{Can}), 0.5/(\text{Ahmet}), 0.25/(\text{Mehmet})) \}$$

$$= \{ 0.5/(\text{Can}), 0.25/(\text{Mehmet}), 0/(\text{Ahmet}) \}$$

**Tablo 3.3** Uzun ve şişman insanlar sorgusu

İs m	Boy	Kilo	Üyelik Der.
Can	180	100	0.5
Mehmet	190	85	0.25
Ahmet	170	90	0

### 3.3 DİSEL ÖZETLER

Dİsel özetler, bir veritabanındaki bilgilerden hareket ederek o veritabanına ait genel çıkarımlar yapmak için kullanılır. Dİsel özetlere örnek olarak “*Bu veritabanındaki çoğu insan uzundur*” veya “*Bu veritabanındaki yüksek maaşlı insanlar şişmandır*” verilebilir.

Bİçimsel olarak tanımladığımızda dİsel özet, format olarak sorgu cümlesi içinde, “*VT’deki S nesne Q (τ doğruluk değeri ile)*” şeklinde gösterilebilir. Burada S özetleyici, Qi se sorgunun uyumuşmasındaki ölçüt olarak adlandırılabilir. Bulunan dİsel özet τ ile gösterilen ve [0, 1] aralığında olan bir doğruluk değeri ne sahiptir. τ ‘nun alacağı değer bize o kuralın veritabanında ne kadar geçerli olduğunu gösterir.

BİR dİsel özetin τ’nu hesaplayabilmek için  $VT = \{O_1, O_2, \dots, O_n\}$  veritabanını oluşturan kayıtlar kümesi olsun. S özetleyici olarak kullanılan bulanık bir ifade, Q bulanık nicelleyici, A’da VT veritabanındaki bir alan olarak tanımlansın. VT veritabanındaki kayıtlar için yapılacak dİsel özetleme işlemi aşağıdaki gibi tanımlanabilir;

1. Veritabanındaki her kayıt için S özetleyici değeri hesaplanır.

$$2. r = \frac{1}{n} \sum_{i=1}^n \mu_S(Q), S'yi \text{ sađlayan } VT \text{ deki kayıtların oranları bulunur.} \quad (3, 2)$$

3.  $\tau = \mu_Q(r)$ , Bulunan  $r$  değeri  $Q$  niceliyi cisinde yerine konarak  $\tau$  değeri hesaplanır.

Dİsel özete örnek olarak “X veritabanındaki çođu insan uzundur” sorgusunun  $\tau$  doğruluk değeri ni hesaplayalım

X veritabanı ve her kayıt için  $\mu_{uzun}$  değeri Tablo 3.4’ de gösterilmiştir.

**Tablo 3.4** Uzun boylu insanların üyeli k dereceleri

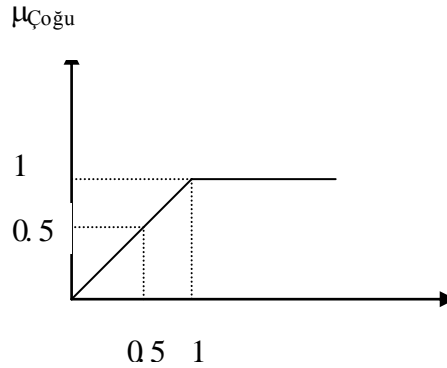
İs i m	Boy	Kilo	$\mu_{uzun}$
Can	180	100	0.5
Ah met	170	90	0
Meh met	190	85	1

İkinci adı molarak bütün  $\mu_{uzun}$  değeri ni toplar ve toplam kayıt sayısına böl ünür.

$$r = (0.5 + 1 + 0) / 3$$

$$r = 0.5$$

Daha sonra bulunan bu  $r$  değeri  $\mu_{çođu}$  bulanık fonksiyonunda yerine konur ve  $\tau$  doğruluk değeri bulunur.  $\mu_{çođu}$  bulanık fonksiyonu Şekil 3.2 de gösterilmiştir.



**Şekil 3.2** : Çođu bulanık fonksiyonu

$$r = 0.5 \text{ ise}$$

$$\tau = \mu_{\text{çoğ u}}(0.5)$$

$\tau = 0.5$  olarak bulunur.

Bu sorgulama tipine ek olarak daha spesifik bir sorgulama yapmak için dilsel özetlemeyi şu şekilde de kullanabiliriz

“X veritabanındaki **çoğ u şiş man** insan **uzundur.**”

Bu sorgulamada, ilk sorgulamaya ek olarak şişman insanlar kümesi *bir alt küme* olarak eklenmiştir. Bu sorgulama tipinin genel hali ise şu şekilde gösterilir:

**VT** ‘deki **Q Ri** insan **S** ‘dir ( $\tau$  doğruluk değeri ile)

R : alt küme

Bu şekildeki bir dilsel özetlemenin doğruluk değeri nin hesabı ise şu şekilde yapılır:

1. Veritabanındaki her kayıt için S özetleyici değeri hesaplanır.

$$2. r = \frac{\sum_{i=1}^n \tau(\mu_S(Q), \mu_R(Q))}{\sum_{i=1}^n \mu_R(Q)} \quad (3.3)$$

3.  $\tau = \mu_Q(r)$ , Bulunan r değeri Q nicel eyici sinde yeri ne konarak  $\tau$  değeri hesaplanır.

2. basamakta ortaya çıkan T, bir t-norm işlemidir. T-norm minimum ve çarpım altında tanımlı ve VE operatörünü sağlayan bir grup ikili operatördür.

Az önce verdiğimiz örnek için hesaplanmış  $\mu_{\text{uzun}}$  ve  $\mu_{\text{şiş man}}$  değerleri Tablo 3.5 de gösterilmiştir.

**Tablo 3.5** Uzun boylu ve şişman insanların üyelik dereceleri

İs m	Boy	Kilo	$\mu_{\text{şiş man}}$	$\mu_{\text{uzun}}$
Can	180	100	1	0.5
Ah met	170	90	0.5	0
Meh met	190	85	0.25	1

$\tau$  değerini hesap için Tablo 3.5'deki  $\mu_{uzun}$  ve  $\mu_{şişman}$  değerlerini her kayıt için  $\mu$  ni mu nları alır ve alt küme olarak tanımladığımız  $\mu_{şişman}$  değerlerini toplamına bölünür.

$$r = [\min(0.5, 1) + \min(0, 0.5) + \min(1, 0.25)] / (1 + 0.5 + 0.25)$$

$$r = 0.428$$

$$\tau = \mu_{çoğ u(r)} = 0.428$$

Sonuç olarak “**X** veritabanındaki **çoğ u şişman** insan **uzundur.**” kuralı veritabanımızda %42.8 lik bir doğruluk değeri ile hesaplanmış olur.

### 3.4 Bulanık Fonksiyonel ve Dereceli Fonksiyonel Bağlılıklar

1980’li yıllarda ortaya çıkan ilişkisel veritabanı sistemleriyle birlikte, veritabanındaki gereksiz verileri yok etmek için fonksiyonel bağlılıklar (FB) kullanılmaya başlanmıştır. Yine fonksiyonel bağlılıklar veritabanı yönetim sistemlerinde bütünlük kısıtları olarak da önemli rol oynarlar. “*Aynı zaman içinde bir arabanın tek bir sahibi vardır*” cümlesi fonksiyonel bağlılık için örnek olarak verilebilir. Bir fonksiyonel bağlılık matematiksel olarak şöyle ifade edebiliriz  $X \rightarrow Y$ , ( $X$  in  $Y$  ye fonksiyonel bağlı olması için)  $Q \in VT$  için,  $\forall Q, Q|Q.X=Q.Y=Q$ . Yani  $X$  in her değeri için  $Y$  de aynı değerin karşılık gelmesi gereklidir.

Bulanık veritabanı kavramının oluşmasıyla birlikte, bulanık fonksiyonel bağlılık (BFB) kavramı da ortaya çıkmıştır. Fakat FB’lerin aksine, BFB’ler veri bütünlüğü saklamaktan çok veritabanlarından bilgi keşfi için kullanılmaktadırlar. BFB’ye örnek olarak “*Aynı tip arabaya sahip olanların gelir düzeyleri de birbirine benzer*” verilebilir. BFB’de amaç, iki yada daha fazla alan arasındaki benzerliğin bulunmasıdır.

Bir başka fonksiyonel bağlılık tipi de dereceli fonksiyonel bağlılıktır (DFB). Dereceli fonksiyonel bağlılık veritabanlarındaki nesnelere hareketleri arasındaki monotonluğunu inceler. DFB’ye örnek olarak “*Bir kişinin yaşı arttıkça, maaşı da artar*” verilebilir.



Sonuç olarak BFB ve DFB'leri kullanarak veritabanlarından kurallar üretebiliriz. Ürettiğimiz bu kurallar daha öncede bahsettiğimiz bütün bulanık küme ve bulanık mantık kriterlerine uyarlar ve daha önce de olduğu gibi ürettiğimiz her kuralın  $[0, 1]$  aralığında bir  $\tau$  doğruluk değeri vardır.

### 3.4.1 Bulanık Fonksiyonel Bağlılık

FB'de yukarıda da anlatıldığı gibi bir alan başka bir alana ya fonksiyonel bağlıdır yada değildir, yani alanın üyelik derecesi 0 yada 1'dir. BFB'de ise bu üyelik derecesi 0 ile 1 arasındadır ki bu rakam 1 ise alan için fonksiyonel bağlıdır da diyebiliriz.  $X$  ve  $Y$  arasındaki BFB'yi göstermek için FB'deki  $Q \cdot X = Q \cdot Y$  bağıntısının yerini  $Q \cdot X \approx Q \cdot Y$  bağıntısı alır. Burada  $\approx$  bulanık benzerlik operatörüdür ve  $\text{Agg}(X_1 \approx X_1, \dots, X_n \approx X_n)$  fonksiyonuyla tanımlanır.

BFB'de,  $X$  ve  $Y$  arasındaki benzerliği bulabilmek için her nesne için birer bulanık kural üretilmelidir. Bir bulanık kuralın genel formu şu şekildedir.

$$R_k : \text{HER } Q \in VT, \text{ EĞER } Q \cdot X \approx Q \cdot Y \text{ İSE } Q \cdot Y \approx Q \cdot Y \quad (3.4)$$

Oluşturulan her kural için bir de doğruluk değeri hesaplanmalıdır. Bu doğruluk değerinin hesabı ise şöyledir.

$$\tau_k = \sum_d (Q \in VT | \sum_d (Q \in \{ Q \in VT | Q \cdot X \approx Q \cdot Y | Q \cdot Y \approx Q \cdot Y \})) \quad (3.5)$$

$\{ Q \in VT | Q \cdot X \approx Q \cdot Y \}$ ;  $X$  baz alınarak oluşturulmuş bulanık bir altküme, id bulanık nicelendirici  $d(\tau) = \tau$  ve  $\approx$ , bulanık benzerlik operatörü  $\text{Max}((2 - |x - y|) / 2, 0)$ 'dir. Bulanık benzerlik operatörü  $\text{Agg}(X_1 \approx X_1, \dots, X_n \approx X_n)$  fonksiyonunda " $\approx$ " yerini " $=$ " alarak ve toplama fonksiyonu " $\text{Agg}$ ", t-norm operatörüne eşit olduğu kabul edilerek oluşturulur.

Her bir kural için  $\tau_k$  değerleri hesaplandığında, BFB  $X \rightarrow_F Y$  ni  $\tau$  değeri bulunmuş olur. Oluşturulan kural, "Eğer veritabanındaki herhangi iki nesne  $X$  için benzer değerlere sahipse,  $Y$  içinde benzer değerlere sahiptir" şeklindedir.  $X$  in  $Y$  ye BFB 'sinin doğruluk değerini bulmak için bütün kuralların doğruluk değerleri toplanır ve toplam kural sayısına bölünür.

$$\tau = \frac{\sum_{i=1}^n \tau_k}{n} \quad (3,6)$$

Bir örnek verilmek gerekirse; veritabanımız Tablo 3.6'daki gibi olsun ve A alanının B'ye ne kadar BFB olduğunu bulalım

**Tablo 3.6** Örnek veritabanı

Nesne	A	B
O <sub>1</sub>	1	1
O <sub>2</sub>	2	4
O <sub>3</sub>	4	4

R<sub>1</sub> : EĞER 1 ≈ Q . A İSE 1 ≈ Q . B dir.

R<sub>2</sub> : EĞER 2 ≈ Q . A İSE 4 ≈ Q . B dir.

R<sub>3</sub> : EĞER 4 ≈ Q . A İSE 4 ≈ Q . B dir.

Her kural için doğruluk değerlerini hesaplayalım

$$1 \approx Q . A \text{ için } 1 \approx 1 = \max((2 - |1 - 1|) / 2, 0) = 1$$

$$1 \approx 2 = \max((2 - |1 - 2|) / 2, 0) = 0.5$$

$$1 \approx 4 = \max((2 - |1 - 4|) / 2, 0) = 0$$

$$1 \approx Q . B \text{ için } 1 \approx 1 = \max((2 - |1 - 1|) / 2, 0) = 1$$

$$1 \approx 4 = \max((2 - |1 - 4|) / 2, 0) = 0$$

$$1 \approx 4 = \max((2 - |1 - 4|) / 2, 0) = 0$$

$$\tau_1 = \frac{\min(1, 1) + \min(0.5, 0) + \min(0, 0)}{1 + 0.5 + 0} = \frac{2}{3}$$

R<sub>1</sub> için doğruluk değeri 0.67 olarak bulunur. Hesaplanan bütün ilişkiler Tablo 3.7'de gösterilmiştir.

**Tablo 3.7** Bulanık Fonksiyonel Bağlılık Değerleri

Nesne	A	B	$1 \approx Q$	$2 \approx Q$	$4 \approx Q$	$1 \approx Q$	$4 \approx Q$	$4 \approx Q$
			.A	.A	.A	.B	.B	.B
$Q_1$	1	1	1	0.5	0	1	0	0
$Q_2$	2	4	0.5	1	0	0	1	1
$Q_3$	4	4	0	0	1	0	1	1

$$\tau_2 = \frac{\min(0,5,0) + \min(1,1) + \min(0,1)}{0,5 + 1 + 0} = \frac{2}{3}$$

$R_2$  için doğruluk değeri  $0,67$  olarak bulunur.

$$\tau_3 = \frac{\min(0,0) + \min(0,1) + \min(1,1)}{0 + 0 + 1} = 1$$

$R_3$  için doğruluk değeri  $1$  olarak bulunur.

$$\{A\} \rightarrow_F \{B\} = \frac{2/3 + 2/3 + 1}{3} = \frac{7}{9} \text{ bulunmuş olur.}$$

### 3.4.2 Dereceli Fonksiyonel Bağlılık

DFB, veritabanındaki nesnelere arasındaki ilişkiyi ve monotonluğunu yansıtır. Aynı BFB'de olduğu gibi DFB'de de bulunan kuralın bir doğruluk derecesi vardır. DFB'ye örnek olarak “*Bir kişinin yaşı arttıkça, maaşı da artar*” yada “*X hisse senedi nin fiyatı azaldıkça, Y hisse senedi nin fiyatı artar*” verilebilir.

DFB'deki monotonluk dereceli terimlerle belirlenir. Artan dereceli terimler için “*arttıkça, daha yüksek, daha büyük*” gibi terimler kullanılırken, azalan dereceli terimler için “*azaldıkça, daha alçak, daha az*” örnekleri verilebilir. DFB'nin genel yapısı; “Veritabanındaki her nesne  $G_1$  ise,  $G_2$  dir” şeklindedir.  $G_1$  ve  $G_2$  dereceli terimleri ifade eder. Üstteki örnek için  $G_1$  yaşın artması,  $G_2$  maaşın artması şeklindedir.

DFB’da kuralın doğruluk değeri hesaplanırken, istenen özellikler dışında, o özelliklerin zıtlarının da hesaplanması gereklidir. Yani “*Bir kişinin yaşı arttıkça, maaşı da artar*” kuralının doğruluk değeri hesaplanırken, yaş artarken maaşın artması hesaplanacağı gibi yaş azalırken maaşın azalması da hesaplanması gerekir. Bir veri tabanındaki bütün nesnelere için oluşturulan kuralların doğruluk değerleri hesaplanır. Bu kuralların doğruluk değerlerinin toplamı toplam kural sayısına bölüldüğünde aranan DFB bulunmuş olur. Her nesne için oluşturulan kurallar şu formdadır:

$$\tau_k = \sum_d (Q \in VT | (G_1(Q_k, Q) \wedge G_2(Q_k, Q)) \vee (G_1(Q_k, Q) \wedge G_2(Q_k, Q))) \quad (3.7)$$

$G(Q_k, Q)$  aranan dereceli terimler olurken,  $G_i(Q_k, Q)$  aranan dereceli terimlerin zıt karşılıklarıdır. Örneğin olarak “*Bir kişinin yaşı arttıkça, maaşı da artar*” kuralında  $G_1 =$  “yaşın artması”,  $G_2 =$  “maaşın artması”,  $G_1 =$  “yaşın azalması”,  $G_2 =$  “maaşın azalması”,dır. Bu kuralı tekrar ifade edecek olursak

$$\tau_k = \sum_d (Q \in VT | (Q_k.yaş < Q.yaş \wedge Q_k.maaş < Q.maaş) \vee (Q_k.yaş > Q.yaş \wedge Q_k.maaş > Q.maaş))$$

şeklini alır. Eğer  $G_1$  dereceli terimi “yaşın artması” yerine “yaşın azalması” olsaydı, oluşacak kural şu şekilde olacaktır.

$$\tau_k = \sum_d (Q \in VT | (Q_k.yaş < Q.yaş \wedge Q_k.maaş > Q.maaş) \vee (Q_k.yaş > Q.yaş \wedge Q_k.maaş < Q.maaş))$$

$G_i$  nin içinde birden fazla dereceli terimde olabilir. Mesela  $G_1 =$  “yaşın artması ve boyun artması” olsaydı

$$G_1(Q_k, Q) = (Q_k.yaş < Q.yaş \wedge Q_k.boy < Q.boy),$$

$$G_1(Q_k, Q) = (Q_k.yaş > Q.yaş \wedge Q_k.boy > Q.boy)$$

şeklinde gösterilirdi.  $G$ deki bütün bağıntılar görüldüğü gibi ikili sistemdedir ve 1 yada 0 değerini alırlar. Bu bağıntıları bulmaklaştırmak için  $\approx$  ve  $\approx$  operatörleri kullanılır.  $x \approx y$  ilişkisi şu şekilde tanımlanır:

$$\text{M}ax ( \text{M}ax ( ( 2 - | x - y | ) / 2, 0 ), x < y ) \quad (3.8)$$

$x \approx y$  ilişkisi ise

$$\text{M}ax ( \text{M}ax ( ( 2 - | x - y | ) / 2, 0 ), x > y ) \quad (3.9)$$

şeklinde dir. Bu bağ lam da “*Bir ki ş i ni n ya ş ı art t ık ç a, maa ş ı da art ar*” kural ı n ın son hali a ş a ğ ı daki ğ i bi d ur:

$$\tau_k = \sum_d ( Q_k \in VT | ( Q_k.ya ş < Q.ya ş \wedge Q_k.maa ş \approx Q.maa ş ) \vee ( Q_k.ya ş > Q.ya ş \wedge Q_k.maa ş \approx Q.maa ş )$$

Yukarı daki özet yapı bir bul an ık kural olarak da gösterilebilir:

$$\text{EĞ ER} ( Q_k.ya ş < Q.ya ş \vee Q_k.maa ş > Q.maa ş ) \text{İ SE} ( Q_k.ya ş < Q.ya ş \wedge Q_k.maa ş \approx Q.maa ş ) \vee ( Q_k.ya ş > Q.ya ş \wedge Q_k.maa ş \approx Q.maa ş )$$

Bazı örnek DFB yapı lar ı n ı ş u ş e kil de gösterebiliriz:

- **EĞ ER** *A art ar* **İ SE** *B art ar.*

$$\text{EĞ ER} ( Q_k.A < Q.A \vee Q_k.A > Q.A ) \text{İ SE} ( Q_k.A < Q.A \wedge Q_k.B \approx Q.B ) \vee ( Q_k.A > Q.A \wedge Q_k.B \approx Q.B )$$

“*X hisse senedi ni n fiyat ı art t ık ç a Y hisse senedi ni n fiyat ı da art ar.*”

- **EĞ ER** *A az d ır* **İ SE** *B art ar.*

$$\text{EĞ ER} ( Q_k.A < Q.A \vee Q_k.A > Q.A ) \text{İ SE} ( Q_k.A < Q.A \wedge Q_k.B \approx Q.B ) \vee ( Q_k.A > Q.A \wedge Q_k.B \approx Q.B )$$

“*İ nsan lar ı n boyu art t ık ç a, kilo su da art ar.*”

- **EĞ ER** *A art ar* **İ SE** *B az d ır.*

$$\text{EĞ ER} ( Q_k.A > Q.A \vee Q_k.A < Q.A ) \text{İ SE} ( Q_k.A > Q.A \wedge Q_k.B \approx Q.B ) \vee ( Q_k.A < Q.A \wedge Q_k.B \approx Q.B )$$

“*İ nsan lar ı n kilo su art t ık ç a, ç evi k li k leri az d ır.*”

- **EĞ ER** *A az d ır* **İ SE** *B az d ır.*

**EĞER** ( $Q_k. A > Q. A \vee Q_k. A < Q. A$ ) **İSE** ( $Q_k. A > Q. A \wedge Q_k. B \gg Q. B$ )  $\vee$   
 ( $Q_k. A < Q. A \wedge Q_k. B \approx Q. B$ )

“*X hisse senedi ni n fiyatı azal dıkça Y hisse senedi ni n fiyatı da azal ır.*”

- **EĞER** A art ar VE B art ar **İSE** C art ar.

**EĞER** ( $Q_k. A > Q. A \wedge Q_k. B > Q. B$ )  $\vee$  ( $Q_k. A < Q. A \wedge Q_k. B < Q. B$ )

**İSE** ( $Q_k. A > Q. A \wedge Q_k. B > Q. B \wedge Q_k. C \gg Q. C$ )  $\vee$

( $Q_k. A < Q. A \wedge Q_k. B < Q. B \wedge Q_k. C \approx Q. C$ )

“*Arabal arı n beygır gücü ve mot or gücü art tıkça, hız lar ı da art ar.*”

- **EĞER** A azal ır VE B art ar **İSE** C azal ır.

**EĞER** ( $Q_k. A < Q. A \wedge Q_k. B > Q. B$ )  $\vee$  ( $Q_k. A > Q. A \wedge Q_k. B < Q. B$ )

**İSE** ( $Q_k. A < Q. A \wedge Q_k. B > Q. B \wedge Q_k. C \approx Q. C$ )  $\vee$

( $Q_k. A > Q. A \wedge Q_k. B < Q. B \wedge Q_k. C \gg Q. C$ )

“*X hisse senedi ni n fiyatı düş erken borsa art arsa Y hisse senedi ni n fiyat ı düş er.*”

Tablo 3.8 deki verilere dayanarak “A özelli ğ artarken, B özelli ğ azal ır” kural ı n n doğr ul uk değ eri ni hesap layal ım

**Tablo 3.8** Ö nek veritaban ı

Nesne	A	B
$Q_1$	1	1
$Q_2$	2	4
$Q_3$	4	4

Her nesne iç i n hesap layaca ğ ımız kural ı n genel formu ař a ğ ı daki ğ i bi dır.

$$R_k = \text{EĞER} (Q_k. A < Q. A \vee Q_k. A > Q. A) \text{ İSE} (Q_k. A < Q. A \wedge Q_k. B \approx Q. B) \vee \\ (Q_k. A > Q. A \wedge Q_k. B \approx Q. B)$$

$$R_1 = \text{EĞER} (1 < Q. A \vee 1 > Q. A) \text{ İSE} (1 < Q. A \wedge 1 \approx Q. B) \vee \\ (1 > Q. A \wedge 1 \approx Q. B)$$

$$R_2 = \text{EĞER} (2 < Q. A \vee 2 > Q. A) \text{ İSE} (2 < Q. A \wedge 4 \approx Q. B) \vee \\ (2 > Q. A \wedge 4 \approx Q. B)$$

$$R_3 = \text{EĞER} (4 < Q. A \vee 4 > Q. A) \text{ İSE} (4 < Q. A \wedge 4 \approx Q. B) \vee \\ (4 > Q. A \wedge 4 \approx Q. B)$$

Birinci kuralın hesabı aşağıdaki gibidir.

$(1 < Q. A \wedge 1 \approx Q. B)$  için

$$(1 < 1 \wedge 1 \approx 1) = 0 \wedge \text{Max} ( \text{Max} ( (2 - |1 - 1|) / 2, 0), 1 > 1) \\ = 0 \wedge \text{Max} ( 1, 0) \\ = 0$$

$$(1 < 2 \wedge 1 \approx 4) = 1 \wedge \text{Max} ( \text{Max} ( (2 - |1 - 4|) / 2, 0), 1 > 4) \\ = 1 \wedge \text{Max} ( 0, 0) \\ = 0$$

$$(1 < 4 \wedge 1 \approx 4) = 1 \wedge \text{Max} ( \text{Max} ( (2 - |1 - 4|) / 2, 0), 1 > 4) \\ = 1 \wedge \text{Max} ( 0, 0) \\ = 0$$

$(1 > Q. A \wedge 1 \approx Q. B)$  için

$$(1 > 1 \wedge 1 \approx 1) = 0 \wedge \text{Max} ( \text{Max} ( (2 - |1 - 1|) / 2, 0), 1 < 1) \\ = 0 \wedge \text{Max} ( 1, 0) \\ = 0$$

$$(1 > 2 \wedge 1 \approx 4) = 0 \wedge \text{Max}(\text{Max}((2 - |1 - 4|)/2, 0), 1 < 4)$$

$$= 0 \wedge \text{Max}(0, 1)$$

$$= 0$$

$$(1 > 4 \wedge 1 \approx 4) = 0 \wedge \text{Max}(\text{Max}((2 - |1 - 4|)/2, 0), 1 < 4)$$

$$= 0 \wedge \text{Max}(0, 1)$$

$$= 0$$

Sonuç olarak bulunan değerleri bir bulan kural halinde aşağıdaki gibi gösterebiliriz

EĞER 0 İSE 0  $\vee$  0

EĞER 1 İSE 0  $\vee$  0

EĞER 1 İSE 0  $\vee$  0

Bütün kuralların sonuçları Tablo 3.9 da gösterilmiştir.

**Tablo 3.9** Dereceli Fonksiyonel Bağlılık Değerleri

Nesne	A	B	$1 < Q.A$	$2 < Q.A$	$4 < Q.A$	$(1 < Q.A$	$(2 < Q.A$	$(4 < Q.A$
			$\vee$ $1 > Q.A$	$\vee$ $2 > Q.A$	$\vee$ $4 > Q.A$	$\wedge$ $1 \approx Q.B$	$\wedge$ $4 \approx Q.B$	$\wedge$ $4 \approx Q.B$
						$\vee$ $(1 > Q.A$	$\vee$ $(2 > Q.A$	$\vee$ $(4 > Q.A$
						$\wedge$ $1 \approx Q.B$	$\wedge$ $4 \approx Q.B$	$\wedge$ $4 \approx Q.B$
$Q_1$	1	1	0	1	1	0	0	0
$Q_2$	2	4	1	0	1	0	0	1
$Q_3$	4	4	1	1	0	0	1	0



$$\tau_1 = \frac{\min(0,0) + \min(1,0) + \min(1,0)}{0+1+1} = 0$$

$$\tau_2 = \frac{\min(1,0) + \min(0,0) + \min(1,1)}{1+0+1} = 0.5$$

$$\tau_3 = \frac{\min(1,0) + \min(1,1) + \min(0,1)}{1+1+0} = 0.5$$

$$S_{GFD} = \frac{\tau_1 + \tau_2 + \tau_3}{\text{Nesne Sayısı}}$$

$$S_{GFD} = \frac{0 + 0.5 + 0.5}{3} = 0.33$$

“A özelliđi artarken, B özelliđi azalır” kuralı %3 doğrulukla geçerlidir.

### 3.4.3 Bulanıklık Operatörü

İster BFB olsun, ister de DFB olsun bulanıklık operatörlerini kullanırken dikkat etmeniz gereken bir nokta vardır. BFB’da ki operatör olan  $\text{Max}((2 - |x - y|) / 2, 0)$  ve DFB’de ki operatörler  $\text{Max}(\text{Max}((2 - |x - y|) / 2, 0), x < y)$  ve  $\text{Max}(\text{Max}((2 - |x - y|) / 2, 0), x > y)$  olan bulanıklık operatörlerinde yer alan 2 sayısı, kayıtlar arasındaki benzerliđi bulurken çok büyük öneme sahiptir. 2 sayısı iki sayının benzerliđine bakarken ki aralıđı belirler. Yani 2 sayısı operatördeyken 1 ile 2 sayılarının benzerliđi 0.5 olurken, operatörün değeri 20 olursa 1 ile 2 sayılarının benzerliđi 0.95 olur. Bu yüzden aralık değeri nin uygulama dan uygulama ya değışmesi kaçınılmazdır. Geliştirilen borsa örneğinde aralık değeri kullanıcı tarafından değıştirilebilir. Birçok deneme ve uzman fikirlerinden sonra en azından İMKB deki hisse senetleri için aralık değeri nin 500 olarak kabul edilmesi uygun görülmüştür.

## 4. GELİŞTİRİLEN UYGULAMANIN GENEL YAPISI

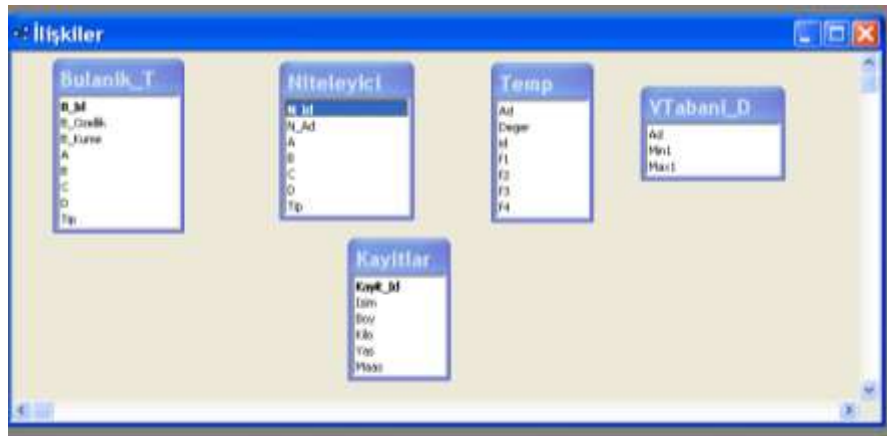
### 4.1 Giriş

Uygulama, Windows platformunda Visual Basic 6.0 ve MS Access kullanılarak geliştirilmiştir. Uygulamanın daha iyi temsil edilebilmesi için iki farklı veritabanı kullanılmıştır. Birinci veritabanında bir şirkette çalışan kişilerin boyları, kıdaları, yaşları ve maaşları tutulmuş ve bu veritabanı üzerinde daha önce anlatılan “bulanık sorgulama” ve “dilsel özet” kısımları uygulanmıştır. İkinci veritabanında ise İMKB’deki 5 hisse senedi nin Eylül 2002 ile Nisan 2003 arasındaki günlük kapanış değerleri tutulmuş ve bu veritabanı üzerinde “bulanık fonksiyonel bağıklık” ve “dilsel fonksiyonel bağıklık” kısımları uygulanmıştır.

### 4.2 Veritabanları

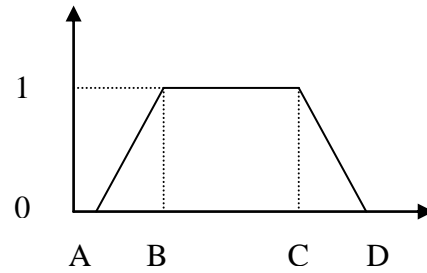
#### 4.2.1 Şirket Veritabanı

Bu veritabanı toplam 5 tablodan oluşmaktadır. Tablolar ve içerdikleri bilgiler Şekil 4.1’de gösterilmiştir.



Şekil 4.1 : Şirket veritabanının içerdiği tablolar

Bulanık\_T tablosu, veritabanında o ana kadar tanımlanmış bütün bulanık kısıtları içermektedir. B\_Ozellik, tanımlanan bulanık kısıtın adıdır, B\_Kueme, o bulanık kısıtın ait olduğu bulanık alan, Tip, bulanık kısıt için kullanılan fonksiyonun tipi ve A, B, C, D alanları o fonksiyonun hesabı için gerekli olan değerleri belirtmektedir. Bulanık\_T tablosu ve içeriği Şekil 4.3’de gösterilmiştir. Programın içinde yeni bulanık kısıtlar eklenebilir, silinebilir yada fonksiyon tipleri ve değerleri değiştirilebilir. Kullanılabilecek toplam 7 çeşit fonksiyon tipi vardır. Bunlar S, P, Z, lineer artan(LS), lineer azalan(LZ), üçgen ve yamuk( TRAP)’tur. Bu fonksiyonların tipleri bulanık kümeler bölümünde anlatılmıştır. Dikkat edilirse S, Z, LS ve LZ fonksiyonlarının C ve D değerleri 0’dır. Bunun nedeni, o fonksiyonların hesabı için iki değer (başlangıç ve bitiş) yeterli olmasıdır. Yamuk ve P fonksiyonu içinse başlangıç ve bitiş değerlerinin yanı sıra fonksiyonun 1 olduğu durumdaki aralığın belirlenebilmesi için iki değere daha ihtiyaç vardır. Örnek olarak normal kilolu insanlar için tanımlanan yamuk fonksiyonu Şekil 4.2’de gösterilmiştir. Fonksiyon B ve C değerleri olan 67 ve 77 kilolarındaki insanlara 1 üyelik derecesini vermekte 60 kilodan az ve 85 kilodan fazla insanlara ise 0 üyelik derecesini vermektedir.



Şekil 4.2 : Yamuk Fonksiyonu

Yamuk fonksiyonun özel bir durumu olarak B ve C alanlarına aynı değerler girilerek fonksiyon üçgen fonksiyon tipi alınır. Buna örnek olarak Orta Boylu insanlar bulanık fonksiyonu verilebilir.

B_id	B_Ozellik	B_Kueme	A	B	C	D	Tip
1	Uzun	Boy	170	190	0	0	0.1S
2	Orta	Boy	150	170	170	190	TRAP
3	Kısa	Boy	150	170	0	0	0.1Z
4	İğirman	Kilo	60	100	0	0	0.1S
5	Sarımsık	Kilo	60	67	77	85	TRAP
6	Zayıf	Kilo	50	70	0	0	0.1Z
7	Yaşlı	Yaş	50	60	0	0	0.1
8	Orta yaşlı	Yaş	25	30	40	50	P
9	Genc	Yaş	25	35	0	0	0.2
10	Yüksek	Maas	000	1400	0	0	0.1
11	Düşük	Maas	250	600	0	0	0.2
12	Ucuca	Boy	125	150	0	0	0.1Z
13	İnci	Kilo	40	60	0	0	0.1Z
14	(Zaman/Sayı)		0	0	0	0	

Şekil 4.3 : Bulanık\_T Tablosunun içeriği

Nceleyici tablosu, dilsel özetleme yapılırken kullanılan niceleyicileri içerir. N\_Ad, niceleyicinin adını, Tip, niceleyici için belirlenen fonksiyonun tipini ve A, B, C, D alanları da o fonksiyonun hesabı için gerekli olan değerleri belirtir. Nceleyici tablosunun içeriği Şekil 4.4’de gösterilmiştir.

N_Ad	N_Ad	A	B	C	D	Tip
Coğu		0	1	0	0	0,05
Hansen Hansen I		0,75	0,95	0	0	0,15

Şekil 4.4 : Nceleyici Tablosunun içeriği

Kayıtlar tablosu, bulanık sorgulamanın ve dilsel özetin yapılacağı bütün kayıtları içerir. Bu tabloda kişilerin özellikleri tutulmaktadır. Bu kayıtlar Şekil 4.5’de gösterilmiştir.

Kayıt_Id	İsim	Boy	Kilo	Yas	Maas
1	Hans	180	100	22	495
2	Peter	170	90	25	875
3	Ben	100	85	26	1270
4	John	195	96	24	800
5	Jake	191	97	27	1270
6	Harry	190	99	22	900
7	Jerry	187	96	32	1500
8	Sam	185	94	30	1200
9	Burt	185	96	23	650
10	Dustin	177	87	37	1500
11	Dan	175	93	40	1750
12	Adam	173	80	39	1500
13	Michael	170	78	37	1200
14	Eric	168	75	34	1200
15	Drew	164	70	30	900

Şekil 4.5 : Kayıtlar Tablosunun içeriği

VTabam\_D tablosu, kullanıcı fonksiyon tipini belirlerken gerekli olan A, B, C, D değerlerini bilmediği yada kullanmak istemediği zamanlarda otomatik olarak veritabanından uygun değerleri üretmek için kullanılır. İçeriği Şekil 4.6’dadır.

Ad	Min1	Max1
Boy	170	190
Kilo	85	100
Maas	405	1270
Yas	22	40

Şekil 4.6 : VTabam\_D Tablosunun içeriği

Temp tablosu, program masasında geçici değerlerin tutulması için kullanılır.

#### 4.2.2 Borsa veritabanı

Daha önce de anlatıldığı üzere, BFB ve DFB’ni daha iyi anlaşılabilirliği için ve çok daha iyi sonuçlar alınabildiği için BFB ve DFB uygularken İ MKB’deki hisse

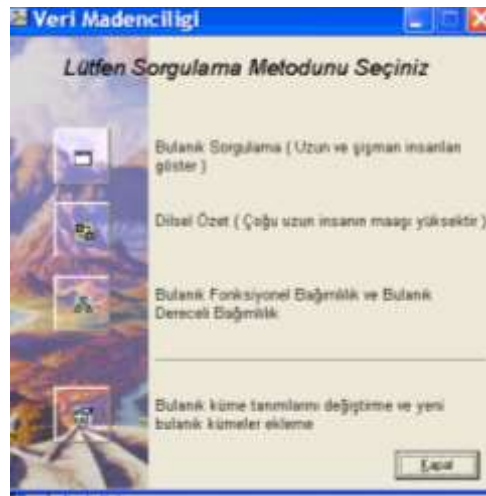
senetlerini günlük kapanış değerlerini kullanılması daha uygun görünmüştür. BFB ve DFB hesaplamalarında herhangi bir bulanık kriter olmadığı ve sadece alanlar arasındaki ilişkileri incelendiği için borsa veritabanı tek bir tablodan oluşmaktadır. Bu tablo Şekil 4.7'de gösterilmiştir. Bu tabloda tarih birincil anahtar olarak kullanılmış ve Akbank, Garanti Bankası, İş Bankası (C), Koç Holding ve Sabancı Holding hisse senetlerini günlük değerleri tutulmuştur.

Hisse_Senetleri : Tablo					
Tarih	Akbank	Garanti Bankası	Is Bankasi C	Koc Holding	Sabancı Holding
01.09.2002	4700	1575	3050	15750	3050
01.09.2002	4800	1575	3000	15000	3050
04.09.2002	4800	1625	3700	15750	3050
05.09.2002	4800	1550	3050	15750	3050
06.09.2002	5150	1600	3000	16250	4000
06.09.2002	5100	1500	3750	15750	3050
10.09.2002	4950	1450	3750	15000	3700
11.09.2002	5000	1500	3750	16000	3750
12.09.2002	5000	1500	3750	15750	3700
13.09.2002	4800	1500	3750	15750	3000
16.09.2002	4850	1475	3700	16000	3700
17.09.2002	4900	1500	3750	16000	3800
18.09.2002	4900	1500	3600	16250	3800
19.09.2002	4900	1500	3750	15900	3850
20.09.2002	4950	1475	3750	16250	3900
23.09.2002	5000	1500	3750	16250	3900
24.09.2002	4900	1450	3350	15750	3800
25.09.2002	4800	1425	3400	15750	3700
26.09.2002	4850	1400	3350	15900	3750
27.09.2002	4750	1400	3400	15750	3800
30.09.2002	4750	1450	3350	15900	3750
01.10.2002	4750	1525	3400	15750	3800
02.10.2002	4800	1475	3400	16000	3800
03.10.2002	4700	1500	3300	15750	3800
04.10.2002	4650	1525	3250	15750	3800
07.10.2002	4800	1550	3300	15750	3800
08.10.2002	4800	1550	3250	15900	3800
09.10.2002	4800	1675	3350	15750	3900
10.10.2002	4800	1700	3400	16250	4000

Şekil 4.7 : Hsse Senetleri Tablosunun içeriği

### 4.3 Çeşitirilen Araç

Programın açılış ekranı Şekil 4.8'de gösterilmiştir. Kullanıcı bulanık sorgulama, dilsel özet, BFB, DFB ve yeni bulanık kısıt ekleme, değiştirme yapabilmektedir.



Şekil 4.8 : Açılış Ekranı

### 4.3.1 Bulanık Sorgulama

Bulanık sorgulama ekranı Şekil 4.9'da gösterilmiştir. Sol taraftaki karma kutularından öncelikle hangi küme üzerinde sorgulama yapılacağı seçilir. Buradaki kümeler veritabanındaki Bulanık\_T tablosunun B\_Küme değerlerinden otomatik olarak çekilir. Bu değerler şu an için boy, kilo, maaş ve yaştır. Bir değer seçildikten sonra özellik karma kutusu aktif hale gelir ve seçilen değerle ilgili bütün özellikleri gösterir. Ayrıca yüklem değiştirici karma kutusu ile çok, biraz gibi yüklem önüne yüklem değiştiriciler getirilebilir. Bu seçimler yapıldıktan sonra "Ekle" düğmesine basılarak istenen değerler Sorgu kutusuna yazılır. Eğer sorgu yeterliyse "Hesapla" düğmesine basılarak sorgu hesaplanır. Birden fazla sorgu aynı anda yapılması isteniyorsa "Ekle" düğmesine basıldıktan sonra sol üstte bulunan "Ve", "Veya" bağlaçları kullanılarak aynı işlemler tekrarlanabilir.

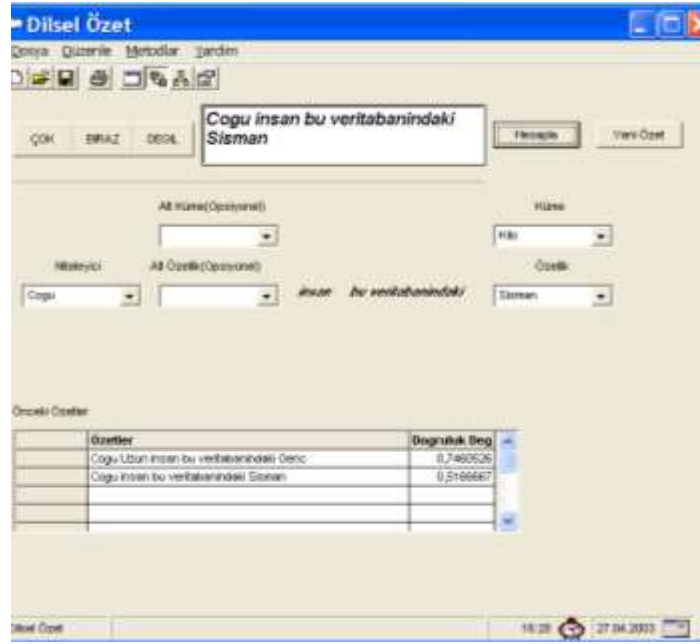


Şekil 4.9 : Bulanık sorgulama ekranı

Kayıtlar üyelik derecesinin büyükten küçüğe sıralanmasıyla gösterilmektedir. Böylece sorguya uyan kurallar daha önce gözükür. Ayrıca kullanıcı bir eşik değeri belirleyerek eşik değerinin üzerindeki kayıtların farklı renkte görünmesini sağlayabilir. Programda standart olarak gelen eşik değeri 0.25'dir. Fakat bu değer programda kullanıcı tarafından değiştirilebilir.

### 4.3.2 Dİsel Özet

Dİsel özet ekranı Şekil 4.10'da gösterilmiştir. Öncelikle nicelleyici seçilir. Daha sonra isteğe bağlı olarak bir alt küme seçilebilir ve son olarak asıl küme ni z ve kısıtlı m z seçilir. Sol üstte “çok”, “değil” ve “biraz” yüklem değıştircileri kısıtlar seçilmeden önce kullanılabilir. “Hesapla” düğmesine basılarak istenen özeti n veri tabanı ndaki doğruluk değeri bulunur. Alt tarafta bulunan ızgara (grid) sayesinde daha önceki özetler ve doğruluk değeri de izlenebilir.



Şekil 4.10 : Dİsel Özet ekranı

Örnek olarak iki tane dİsel özet verilmiştir. Bunlardan ilki bir alt küme de içer mektedir ve veritabanında %74.6'lık doğruluk değeriyle geçerli dir. İkinci özet ise alt küme içer me mektedir ve veritabanında %51.55'lik doğruluk değeriyle geçerli dir.

### 4.3.3 Bulanık Fonksiyonel Bağlılık ve Dereceli Fonksiyonel Bağlılık

BFB ve DFB ekranı Şekil 4.11'de gösterilmiştir. Daha önce de bahsedildiği üzere, BFB ve DFB için İ MKB'deki hisse senetlerinin bir kısmının günlük kapanış değerlerinin tutulduğu bir veritabanı kullanılmıştır. Aralık değeri, uzmanlara da danışılarak 500 olarak kabul edilmiştir. Fakat “Değıştir” düğmesi ile aralık değeri değıştirilebilir. BFB için EĞER Xİ SE Yşeklinde bir yapı kullanılmıştır. Bu yapı ile X hisse senedi ile Y hisse senedinin hareketleri arasındaki benzerlik oranı Bölüm 3.4.1'deki tanımlara göre hesaplanmaktadır. DFB için X hisse senedi

*art ar(azdır)ken* Y hisse senedi *art ar(azdır)*, şeklinde bir yapı kullanılmıştır. Bu yapının hesabı için Bölüm 3.4.2 deki tanımın kullanılmaktadır. Yine bu hesaplamalar yapılırken aralık değeri 500 olarak kabul edilmiştir. En alttaki “Test” bölümüyle iki sayı arasındaki benzerlik aralık değerine göre hesaplanabilmektedir. Böylece aralık değerinin değişiminin ne gibi farklılıklar yaratacağı kolayca gözlemlenebilir.

Kuralın Adı	Derecelik Değeri
Albank ile Sabancı Holding benzer hareket eder	0.7251191757964196
Albank Artakim Sabancı Holding Artar	0.9501516329656110

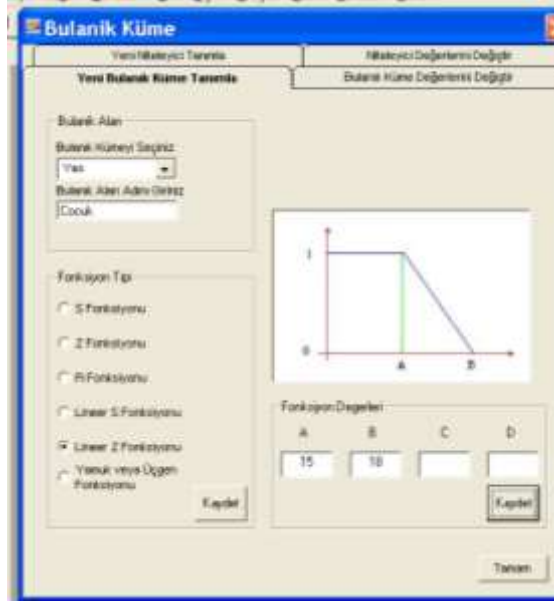
Şekil 4.11 : BFB ve DFB ekranı

Daha önce hesaplanan kurallar bir ızgara (grid) yardımıyla görülebilmektedir.

#### 4.3.4 Yeni Bulanık Alanlar Tanımlamak

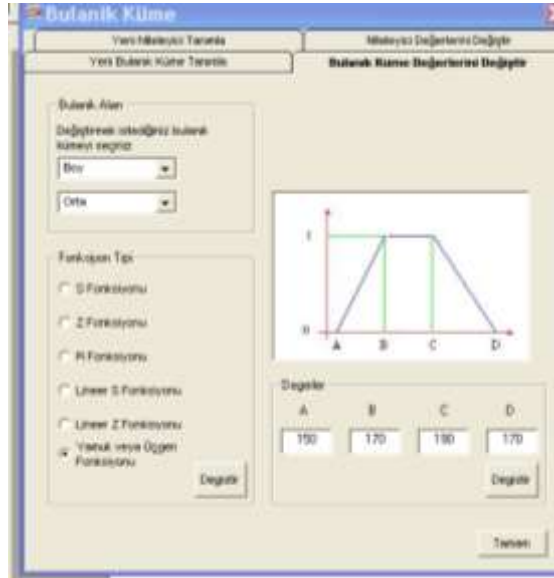
Giriş sayfasının en altında bulunan “Bulanık küme tanımlarını değiştirme ve yeni bulanık kümeler eklenme” düğmesiyle programa yeni bulanık kümeler eklenebilir yada varolan bulanık kümeler değiştirilebilir. Şekil 4.12 yeni bulanık alanlarının nasıl eklendiğini göstermektedir.





**Şekil 4 12 : Yeni Bulanık Alan Ekleme**

Şekil 4. 12’deki örnekte yaş bulanık kümesi ne “Çocuk” isimli yeni bir bulanık alanın eklenmesi gösterilmiştir. Küme seçilip alan ismi girildikten sonra yeni alan için uygun olan fonksiyon tipi seçilir. Bu seçim esnasında kullanıcıya kolaylık olması açısından fonksiyon tipleri şekillere de ifade edilmiştir. Uygun fonksiyon tipi seçildikten sonra bulanık alanın hesaplanması için gerekli olan değerler girilir ve böylece yeni bulanık alan kaydedilmiş olur. Şekil 4. 13’te daha önceden tanımlanmış olan bulanık alanların nasıl değiştirilebileceği gösterilmiştir.



**Şekil 4 13 : Bulanık Alan Değerlerini Değiştirme**

Değiştirilmek istenen bulanık alanın ait olduğu bulanık küme seçildikten sonra otomatik olarak o kümeye ait bütün alanlar liste halinde kullanıcıya gösterilir. Kullanıcı istediği alanı seçtikten sonra, o alan için daha önceden tanımlanmış fonksiyon tipi ve fonksiyonun aldığı değerler ekrana gelir. Kullanıcı buradan istediği şekilde fonksiyon tipini yada değerlerini değiştirebilir.

#### **4.4 Borsa Veritabanından Üretilen Kurallar**

Üretilen kurallar, İ MKB'deki Akbank, Garanti Bankası, İş Bankası Ç Koç Holding ve Sabancı Holding hisse senetlerinin 01 Eylül 2002 ile 30 Nisan 2003 tarihleri arasındaki günlük kapanış değerlerinden hesaplanmıştır. Üretilen bu kurallar Bölüm 4.4.1 ve Bölüm 4.4.2'de gösterilmiştir. Kurallar doğruluk değeri en yüksekte en düşüğe göre sıralanmıştır.

##### **4.4.1 BFB Kullanılarak Üretilen Kurallar**

- Koç Holding hisse senedinin hareketleriyle Garanti Bankası hisse senedinin hareketleri birbirine benzer.

Doğruluk değeri: 0.794

- İş Bankası C hisse senedinin hareketleriyle Garanti Bankası hisse senedinin hareketleri birbirine benzer.

Doğruluk değeri: 0.784

- Akbank hisse senedinin hareketleriyle Garanti Bankası hisse senedinin hareketleri birbirine benzer.

Doğruluk değeri: 0.718

- Koç Holding hisse senedinin hareketleriyle Sabancı Holding hisse senedinin hareketleri birbirine benzer.

Doğruluk değeri: 0.715

- Sabancı Holding hisse senedinin hareketleriyle Garanti Bankası hisse senedinin hareketleri birbirine benzer.

Doğruluk değeri: 0.665

- Koç Holding hisse senedi nin hareketleriyle İş Bankası C hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.595

- İş Bankası C hisse senedi nin hareketleriyle Sabancı Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.592

- Koç Holding hisse senedi nin hareketleriyle Akbank hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.575

- İş Bankası C hisse senedi nin hareketleriyle Akbank hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.551

- Akbank hisse senedi nin hareketleriyle Sabancı Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.538

- Garanti Bankası hisse senedi nin hareketleriyle Akbank hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.529

- Garanti Bankası hisse senedi nin hareketleriyle Sabancı Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.513

- Sabancı Holding hisse senedi nin hareketleriyle Akbank hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.481

- Garanti Bankası hisse senedi nin hareketleriyle İş Bankası C hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.428

- Sabancı Holding hisse senedi nin hareketleriyle İş Bankası C hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.411

- Akbank hisse senedi nin hareketleriyle İş Bankası C hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.409

- İş Bankası C hisse senedi nin hareketleriyle Koç Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.288

- Sabancı Holding hisse senedi nin hareketleriyle Koç Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.265

- Akbank hisse senedi nin hareketleriyle Koç Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.222

- Garanti Bankası hisse senedi nin hareketleriyle Koç Holding hisse senedi nin hareketleri birbirine benzer.

Doğruluk değeri: 0.217

#### **4.4.2 DFB Kullanılarak Üretilen Kurallar**

- “X artarken, Y artar”; “X azalırken, Y azalır” yapısındaki kurallar

- Akbank hisse senedi nin değeri artarken, Garanti Bankası hisse senedi nin değeri de artar.

Akbank hisse senedi nin değeri azalırken, Garanti Bankası hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.973

- Koç Holding hisse senedi nin değeri artarken, Garanti Bankası hisse senedi nin değeri de artar.

Koç Holding hisse senedi nin değeri azalırken, Garanti Bankası hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.967

- İş Bankası C hisse senedi nin değeri artarken, Garanti Bankası hisse senedi nin değeri de artar.

İş Bankası C hisse senedi nin değeri azalırken, Garanti Bankası hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.958

- Garanti Bankası hisse senedi nin değeri artarken, Akbank hisse senedi nin değeri de artar.

Garanti Bankası hisse senedi nin değeri azalırken, Akbank hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.956

- Koç Holding hisse senedi nin değeri artarken, Sabancı Holding hisse senedi nin değeri de artar.

Koç Holding hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.929

- Garanti Bankası hisse senedi nin değeri artarken, İş Bankası C hisse senedi nin değeri de artar.

Garanti Bankası hisse senedi nin değeri azalırken, İş Bankası C hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.907

- Koç Holding hisse senedi nin değeri artarken, Akbank hisse senedi nin değeri de artar.

Koç Holding hisse senedi nin değeri azalırken, Akbank hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.907

- Garanti Bankası hisse senedi nin değeri artarken, Koç Holding hisse senedi nin değeri de artar.

Garanti Bankası hisse senedi nin değeri azalırken, Koç Holding hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.905

- Akbank hisse senedi nin değeri artarken, Sabancı Holding hisse senedi nin değeri de artar.

Akbank hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.902

- Sabancı Holding hisse senedi nin değeri artarken, Garanti Bankası hisse senedi nin değeri de artar.

Sabancı Holding hisse senedi nin değeri azalırken, Garanti Bankası hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.900

- Garanti Bankası hisse senedi nin değeri artarken, Sabancı Holding hisse senedi nin değeri de artar.

Garanti Bankası hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.897

- İş Bankası C hisse senedi nin değeri artarken, Akbank hisse senedi nin değeri de artar.

İş Bankası C hisse senedi nin değeri azalırken, Akbank hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.884

- Akbank hisse senedi nin değeri artarken, İş Bankası C hisse senedi nin değeri de artar.

Akbank hisse senedi nin değeri azalırken, İş Bankası C hisse senedi nin değeri de azalır.

Dođrul uk deđeri: 0 875

- Sabancı Hol dı ng hisse senedi ni n deđeri artarken, Akbank hisse senedi ni n deđeri de artar.

Sabancı Hol dı ng hisse senedi ni n deđeri azalırken, Akbank hisse senedi ni n deđeri de azalır.

Dođrul uk deđeri: 0 874

- Sabancı Hol dı ng hisse senedi ni n deđeri artarken, Koç Hol dı ng hisse senedi ni n deđeri de artar.

Sabancı Hol dı ng hisse senedi ni n deđeri azalırken, Koç Hol dı ng hisse senedi ni n deđeri de azalır.

Dođrul uk deđeri: 0 869

- Akbank hisse senedi ni n deđeri artarken, Koç Hol dı ng hisse senedi ni n deđeri de artar.

Akbank hisse senedi ni n deđeri azalırken, Koç Hol dı ng hisse senedi ni n deđeri de azalır.

Dođrul uk deđeri: 0 860

- Koç Hol dı ng hisse senedi ni n deđeri artarken, İş Bankası C hisse senedi ni n deđeri de artar.

Koç Hol dı ng hisse senedi ni n deđeri azalırken, İş Bankası C hisse senedi ni n deđeri de azalır.

Dođrul uk deđeri: 0 840

- İş Bankası C hisse senedi ni n deđeri artarken, Koç Hol dı ng hisse senedi ni n deđeri de artar.

İş Bankası C hisse senedi ni n deđeri azalırken, Koç Hol dı ng hisse senedi ni n deđeri de azalır.

Dođrul uk deđeri: 0 809

- İş Bankası C hisse senedi ni n deđeri artarken, Sabancı Hol dı ng hisse senedi ni n deđeri de artar.

İş Bankası C hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.772

- Sabancı Holding hisse senedi nin değeri artarken, İş Bankası C hisse senedi nin değeri de artar.

Sabancı Holding hisse senedi nin değeri azalırken, İş Bankası C hisse senedi nin değeri de azalır.

Doğruluk değeri: 0.720

- “X artarken Y azalır”, “X azalırken Y artar” yapısındaki kurallar

- İş Bankası C hisse senedi nin değeri artarken, Sabancı Holding hisse senedi nin değeri azalır.

İş Bankası C hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri artar.

Doğruluk değeri: 0.531

- Sabancı Holding hisse senedi nin değeri artarken, İş Bankası C hisse senedi nin değeri azalır.

Sabancı Holding hisse senedi nin değeri azalırken, İş Bankası C hisse senedi nin değeri artar.

Doğruluk değeri: 0.465

- Sabancı Holding hisse senedi nin değeri artarken, Garanti Bankası hisse senedi nin değeri azalır.

Sabancı Holding hisse senedi nin değeri azalırken, Garanti Bankası hisse senedi nin değeri artar.

Doğruluk değeri: 0.443

- Garanti Bankası hisse senedi nin değeri artarken, Sabancı Holding hisse senedi nin değeri azalır.

Garanti Bankası hisse senedi nin değeri azalırken, Sabancı Holding hisse senedi nin değeri artar.

Doğruluk değeri: 0.400



- Akbank hisse senedi nin deęeri artarken, Sabancı Hol d i ng hisse senedi nin deęeri azalır.

Akbank hisse senedi nin deęeri azalırken, Sabancı Hol d i ng hisse senedi nin deęeri artar.

Doęrul uk deęeri: 0 396

- İş Bankası C hisse senedi nin deęeri artarken, Caranti Bankası hisse senedi nin deęeri azalır.

İş Bankası C hisse senedi nin deęeri azalırken, Caranti Bankası hisse senedi nin deęeri artar.

Doęrul uk deęeri: 0 389

- Koç Hol d i ng hisse senedi nin deęeri artarken, Caranti Bankası hisse senedi nin deęeri azalır.

Koç Hol d i ng hisse senedi nin deęeri azalırken, Caranti Bankası hisse senedi nin deęeri artar.

Doęrul uk deęeri: 0 369

- Sabancı Hol d i ng hisse senedi nin deęeri artarken, Akbank hisse senedi nin deęeri azalır.

Sabancı Hol d i ng hisse senedi nin deęeri azalırken, Akbank hisse senedi nin deęeri artar.

Doęrul uk deęeri: 0 365

- Akbank hisse senedi nin deęeri artarken, Caranti Bankası hisse senedi nin deęeri azalır.

Akbank hisse senedi nin deęeri azalırken, Caranti Bankası hisse senedi nin deęeri artar.

Doęrul uk deęeri: 0 359

- İş Bankası C hisse senedi nin deęeri artarken, Akbank hisse senedi nin deęeri azalır.

İş Bankası C hisse senedi nin deęeri azalırken, Akbank hisse senedi nin deęeri artar.

Dođrul uk deđeri: 0 356

- Koç Hol di ng hisse senedi ni n deđeri artarken, Sabancı Hol di ng hisse senedi ni n deđeri azalır.

Koç Hol di ng hisse senedi ni n deđeri azalırken, Sabancı Hol di ng hisse senedi ni n deđeri artar.

Dođrul uk deđeri: 0 355

- Koç Hol di ng hisse senedi ni n deđeri artarken, İş Bankası C hisse senedi ni n deđeri azalır.

Koç Hol di ng hisse senedi ni n deđeri azalırken, İş Bankası C hisse senedi ni n deđeri artar.

Dođrul uk deđeri: 0 340

- Koç Hol di ng hisse senedi ni n deđeri artarken, Akbank hisse senedi ni n deđeri azalır.

Koç Hol di ng hisse senedi ni n deđeri azalırken, Akbank hisse senedi ni n deđeri artar.

Dođrul uk deđeri: 0 327

- Akbank hisse senedi ni n deđeri artarken, İş Bankası C hisse senedi ni n deđeri azalır.

Akbank hisse senedi ni n deđeri azalırken, İş Bankası C hisse senedi ni n deđeri artar.

Dođrul uk deđeri: 0 312

- İş Bankası C hisse senedi ni n deđeri artarken, Koç Hol di ng hisse senedi ni n deđeri azalır.

İş Bankası C hisse senedi ni n deđeri azalırken, Koç Hol di ng hisse senedi ni n deđeri artar.

Dođrul uk deđeri: 0 291

- Caranti Bankası hisse senedi ni n deđeri artarken, Akbank hisse senedi ni n deđeri azalır.

Garanti Bankası hisse senedi nin değeri azalırken, Akbank hisse senedi nin değeri artar.

Doğruluk değeri: 0.279

- Garanti Bankası hisse senedi nin değeri artarken, İş Bankası C hisse senedi nin değeri azalır.

Garanti Bankası hisse senedi nin değeri azalırken, İş Bankası C hisse senedi nin değeri artar.

Doğruluk değeri: 0.278

- Akbank hisse senedi nin değeri artarken, Koç Holding hisse senedi nin değeri azalır.

Akbank hisse senedi nin değeri azalırken, Koç Holding hisse senedi nin değeri artar.

Doğruluk değeri: 0.240

- Sabancı Holding hisse senedi nin değeri artarken, Koç Holding hisse senedi nin değeri azalır.

Sabancı Holding hisse senedi nin değeri azalırken, Koç Holding hisse senedi nin değeri artar.

Doğruluk değeri: 0.230

- Garanti Bankası hisse senedi nin değeri artarken, Koç Holding hisse senedi nin değeri azalır.

Garanti Bankası hisse senedi nin değeri azalırken, Koç Holding hisse senedi nin değeri artar.

Doğruluk değeri: 0.196

#### **4.4.3 Borsa Veritabanının Değerlendirilmesi**

Bölüm 4.4.1 ve 4.4.2’de veritabanından oluşturulabilecek bütün sonuçlar gösterilmiştir. Fakat bulunan her sonucu bir kural olarak değerlendirerek yanlış olur. Bir sonucun kural olabilmesi için doğruluk değerinin belli bir eşik değerinin üzerinde olması gerekir. Bu eşik değeri, her yapı için farklı olacaktır. Bu eşik değerinin neden

farklı olması gerektiği DFB'deki iki yapının sonuçlarının incelenmesiyle kolayca anlaşılabilir.

DFB için “X artarken, Y artar” yapısı incelendiğinde bulunan bütün sonuçların % 70'in üzerinde doğruluk değerine sahip olduğu görülmektedir. Aynı şekilde ikinci yapı olan “X artarken, Y azalır” yapısında da bulunan doğruluk değerleri % 50'nin üzerine pek çıkmamaktadır.

Bunun nedeni, borsanın genel hareketlerinin biraz incelenmesiyle anlaşılabilir. Borsanın yükselmesi yada düşmesi için çoğu hisse senedini aynı gün içinde aynı yönde hareket etmesi gereklidir. Bu yüzden aynı yönlü hareketler için doğruluk değerleri oldukça yüksek, zıt yönlü hareketler içinse doğruluk değerleri düşük çıkmaktadır.

Bunları göz önüne alarak, ilk yapı için doğruluk değerini %90'ın üzerinde, ikinci yapı içinse doğruluk değerini %40'ın üzerinde seçerek daha sağlıklı kurallar elde edebiliriz.

## SONUÇLAR VE ÖNERİLER

Veri madenciliği, içinde bulunduğumuz yüzyılın en değerli teknolojisi olan bilgi teknolojilerinin ulaştığı son noktadır ve gün geçtikçe bu konu hakkında yeni fikirler üretilmektedir.

Bir veri madenciliği çalışması için öncelikle çok miktarda kaliteli veri gereklidir. Amaç bu veri içinde saklı, gelecekle ilgili tahmin yapmakta kullanılacak kural ve bağlantıların çıkarılmasıdır. Böyle bir çalışmanın başarılı olabilmesi için uygulamadaki uzmanların veri tabanları ve veri madenciliği konusundaki uzmanlarla beraber çalışması gerekir.

Bu tez çalışması ile, klasik veri madenciliği algoritmalarına alternatif bir algoritma üretilmiştir. Bu algoritma sayesinde bulanık küme teorisi kullanılarak veri tabanı sistemlerinden bilgi keşfi yapılabilmektedir.

## KAYNAKLAR

- [1] **Akpınar, H** 2000. Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği. İ. Ü İşletme Fakültesi Dergisi, Sayı 1, s 1-22
- [2] **Aksoy, Y,** 1995. Modern Mantık ( Sembolik Mantık), Y.T.Ü Yayın No. 300, İstanbul
- [3] **Alpaydın, E,** 2000. Bilişim2000 Veri Madenciliği Eğitimi Semineri
- [4] **Chen, S M,**1988. A New Approach to Handling Fuzzy Decision Making Problems, IEEE Transactions on Systems, Man, and Cybernetics- Vol. 18 No. 6
- [5] **Chen, S M,** 1996. A Fuzzy Reasoning Approach for Rule- Based Systems Based on Fuzzy Logic, IEEE Transactions on Systems, Man, And Cybernetics- Part B Vol. 26 No. 5
- [6] **Kir, J G,** 1988. Fuzzy Sets, Uncertainty and Information. Prentice- Hall
- [7] **Raschia, G And Mouaddib, N** 2001. Saint Etienne A Fuzzy Set- Based Approach to Database Summarization. Elsevier Science B V.– Fuzzy Sets and Systems 129 (2002) 137-162
- [8] **Rasmussen, D And Yager, R R,** 1996. Fuzzy Query Language for Hypothesis Evaluation,
- [9] **Rasmussen, D And Yager, R R,** 1997. Finding Fuzzy and Gradual Dependencies with SummarySQL. Elsevier Science B V. Fuzzy Sets and Systems 106 (1999) 131-142
- [10] **Zadeh, Lutfi A,** 1973. Outline of a New Approach to the Analysis of Complex Systems and Decision Process, University of California, Berkeley

## **EKLER**

### **EK A: GELİŞTİRİLEN UYGULAMA**

Geliştirilen uygulama ve program kodu tezin arka iç kapağındaki cd'de bulunmaktadır.

## ÖZGEÇMİŞ

Bu tezi hazırlayan Selahattin BOSTANCI, 1979 yılında Karabük'te doğmuştur. Lise eğitimi İstek Özel Belde Lisesi'nde tamamlamıştır. 1997 Yıldız Teknik Üniversitesi Matematik Mühendisliğinde lisans eğitime başlamış ve 2001 yılında bu bölümden mezun olmuştur. Aynı yıl içinde İstanbul Teknik Üniversitesi Mühendislik Bilimleri Anabilim Dalı Sistem Analizi Bölümü'nde yüksek lisans eğitime başlamıştır.