

**VERİ MADENCİLİĞİ ALANINDA  
BİR BULANIK MANTIK UYGULAMASI**

**YÜKSEK LİSANS TEZİ**

**Mat. Müh. Sezin Selen İNCELER**

**Anabilim Dalı: Mühendislik Bilimleri**

**Program: Sistem Analizi**

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

**MAYIS 2003**

**VERİ MADENCİLİĞİ ALANINDA  
BİR BULANIK MANTIK UYGULAMASI**

**YÜKSEK LİSANS TEZİ**  
**Mt. Mih. Sezin Selen İNCELER**  
**(509011207)**

**Tezin Enstitüye Verildiği Tarih: 05 Mayıs 2003**  
**Tezin Savunulduğu Tarih: 28 Mayıs 2003**

**Tez Danışmanı : Yr. Doç. Dr. Ali ERCENGİZ**  
**Diğer Jüri Üyeleri Doç. Dr. Şakir KOCABAŞ**  
**Doç. Dr. Metin Orhan KAYA**

**MAYIS 2003**

## ÖNSÖZ

Bu çalışmayı hazırlarken yaptığım değerli katkılardan ve manevi desteğinden ötürü Sayın Hocam Y. Doç. Dr. Ali Ercengiz'e ve daha sonra ise bu çalışmanın oluşum sürecinde büyük emeği bulunan Sayın Hocam Prof. Dr. Gazanfer Ünal'a teşekkür ve şükranlarımı sunar, çalışmada kaynak olarak kullanılan verileri temin etmemdeki yardımlarından dolayı Vatan Hastanesi Grup Laboratuvarları Koordinatörü Dr. Yavuz Zöngür'e, bana olan güven ve sevgilerini her zaman yanımda hissettiğim sevgili aileme çok teşekkür ederim

Mayıs 2003

Mt. Mih. Sezin Selen İNCELER

## İÇİNDEKİLER

<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. GİRİŞ</b>	<b>1</b>
1.1. Belirsizlik	1
1.2. Çerçelme ve Kurallar	2
1.3. Belirgin Bağlılık	3
1.3.1. Belirgin Bağlılık İşlemleri	4
1.3.2. Belirgin Bağlılık Kompozisyonu	5
1.4. Bulanık Bağlılık	5
1.4.1. Bulanık Bağlılık İşlemleri	6
1.4.2. Bulanık Bağlılık Kompozisyonu	7
1.5. Bulanık Mantık	7
1.5.1. Bulanık Mantık Yaklaşımıyla Çıkartım	7
1.5.2. Bulanık Kümelere ve Üyelik Fonksiyonları	9
1.5.3. Dilsel Değişkenler, Dilsel Pekiştiriciler ve Değil Değiştiricisi	12
1.5.4. Bulanık Önerme ve Dilsel Niteleyicilerin Bulanık Önermedeki Yeri	14
1.6. Basit Bulanık Önerme ve Bileşik Bulanık Önerme	16
<b>2. HİPOTEZ ÇIKARIMI İÇİN ESNEK BİR BULANIK SORGULAMA DİLİ</b>	<b>18</b>
2.1. Giriş	18
2.2. Bulanıklaştırma Modülü	19
2.3. Bulanık Soru Cümlesi Analizi	20
2.3.1. Basit ve Bileşik Bulanık Soru Cümlecikleri	20
2.3.2. Dilsel Niteleyiciler İle Bulanık Sorgulama	23
2.4. Bulanık Kural Çıkartım Mekaniizması	24
2.5. Bulanık Fonksiyonel Bağlılık Kural Çıkartım Mekaniizması	26
2.6. Dereceli Bulanık Fonksiyonel Bağlılık Kural Çıkartım Mekaniizması	30

<b>3. BULANIK SORGU ARACI UYGULAMASI</b>	<b>35</b>
3.1. Genel Yapı	35
3.2. Veritabanı Yapısı	36
3.3. Bulanıklaştırma Modülü	39
3.4. Soru Modülleri	40
3.4.1. Bulanık At Küme İle Bleşik Sorulama	40
3.4.2. Fonksiyonel Soru	41
3.4.3. Dereceli Soru	43
3.5. Hasta Bilgi Veritabanı Üzerinde Sorulama	44
3.5.1. Hasta Bilgi VT Üzerinde Bulanık Sorulama Sonuçları	45
3.5.2. Hasta Bilgi VT Üzerinde Basit ve Bleşik BFB Sorulama Sonuçları	47
3.5.3. Hasta Bilgi VT Üzerinde DFB Sorulama Sonuçları	52
<b>4. SONUÇLAR VE ÖNERİLER</b>	<b>55</b>
<b>KAYNAKLAR</b>	<b>56</b>
<b>ÖZGEÇMİŞ</b>	<b>57</b>

## **KISALTMALAR**

<b>VT</b>	: Veritabanı
<b>BFB</b>	: Bulanık Fonksiyonel Bağlılık
<b>DFB</b>	: Dereceli Bulanık Fonksiyonel Bağlılık

## TABLO LİSTESİ

<b>Tablo 2.1</b>	Kaynak VT Tablosu.....	27
<b>Tablo 2.2</b>	Bulank $V_{f\ G1}$ .....	27
<b>Tablo 2.3</b>	Bulank $V_{f\ G2}$ .....	27
<b>Tablo 2.4</b>	BFB Önek Tablosu.....	29
<b>Tablo 2.5</b>	GFB Önek Tablosu.....	33

## ŞEKİL LİSTESİ

Şekil 1.1	: Kural Yapısı ve Gerçeklerin kullanılış şekli.....	2
Şekil 1.2	: R Bağlılığı temsil matrisi.....	4
Şekil 1.3	: $A \times B$ Bağlılığı temsil matrisi.....	6
Şekil 1.4	: S_Fonksiyon.....	10
Şekil 1.5	: Z_Fonksiyon.....	11
Şekil 1.6	: Trapezoidal Fonksiyon.....	11
Şekil 1.7	: Uzun bulanlık alt kümesini üyeli fonksiyonu.....	13
Şekil 1.8	: Uzun ve Çok Uzun bulanlık alt kümelerini üyeli fonksiyonları	14
Şekil 2.1 a	: Uzun Dilsel Değişkeni üyeli fonksiyonu.....	21
Şekil 2.1 b	: Ağır Dilsel Değişkeni üyeli fonksiyonu.....	21
Şekil 2.2	: Çoğu Dilsel Niteleyicisi üyeli fonksiyonu.....	24
Şekil 3.1	: Sorgu Aracı Çalışma Prensipleri.....	36
Şekil 3.2	: Hızla Bilgi Tablosu.....	36
Şekil 3.3	: TableParameters Tablosu.....	37
Şekil 3.4	: Cholesterol Değeri için tanımlı üyeli fonksiyonları.....	38
Şekil 3.5	: Powers Tablosu.....	38
Şekil 3.6	: Quantifiers Tablosu.....	38
Şekil 3.7	: Bulanlık Küme Tanımlama Menüü.....	39
Şekil 3.8	: Bulanlık Sorgulama Menüü.....	40
Şekil 3.9	: Bir Kısmi Dilsel Niteleyicisi için Bulanlık Sorgulama Menüü.....	41
Şekil 3.10	: Fonksiyonel Sorgulama Menüü.....	41
Şekil 3.11	: Bileşik Fonksiyonel Sorgulama Menüü.....	43
Şekil 3.12	: Dereceli Sorgulama Menüü.....	44



## SEMBOL LİSTESİ

- $\mu_A(x)$  :  $x$  elemanın  $A$  bulanık kümesine olan üyelik derecesi  
 $q$  : Veritabanı elemanları  
 $S$  : Özetleyici  
 $Q$  : Dilsel nitelleyici  
 $r$  : Bulanık kural doğruluk derecesi  
 $\tau$  : Dilsel nitelleyiciye olan üyelik derecesi  
 $\approx$  : Benzerlik ifadesi

## VERİ MADENCİLİĞİ ALANINDA BİR BULANIK MANTIK UYGULAMASI

### ÖZET

Sıradan veritabanı sistemlerinde bilgiler çoğaldıkça ve bilgi kaynakları arttıkça, toplanan verilerin işlenmesi de bir o kadar zorlaşır. Verilerin işlenmesindeki en önemli sorunlardan biri de eldeki veriden azami fayda sağlamaadaki zorluktur. Bilgi toplama, bilgi den sonuç üretme ve üretilen sonuçtan faydalanma süreçlerinde kesin ve belirgin bilgileri analiz edebilme kadar, bulanık kavramı çeren bilgileri de analiz edebilme önemlidir. Bulanıklık kesin olmaktan kaynaklanan belirsizlik durumudur ve doğal dilde kullanılan pek çok ifade de karşımıza çıkmaktadır. Klasik mantık ve bulanık küme yaklaşımıyla sıradan veritabanlarında sözsel sorgu yapabilen esnek bulanık sorgulama sistemleri, doğal dildeki ne olabilirdiğince yakın sorgu cümleleri kullanır. Bu sorgu cümleleri ile verileri işleyerek, yapılarındaki gizli gerçekleri açığa çıkartmak bir veri madenciliği tekniği olarak kabul edilmiştir.

1965 yılında L. A. Zadeh tarafından tanımlanan bulanık kümeler teorisi, kesin olmayan belirsiz bilgiyi işleme özelliğine sahip olup, uygulamada pek çok avantajları da beraberinde getirmiştir. Klasik mantıkta bir önerme ya doğru ya yanlıştır ve belirsizliğe yer verilmemiştir. Oysa Zadeh'in yaklaşımı ile  $\{0, 1\}$  değerlerini alabilen iki değerli mantık,  $[0, 1]$  aralığında değer alabilen sürekli değerli mantığa dönüşürülmüştür. Bu teori de bulanık önermeler, bulanık küme yaklaşımıyla oluşturulur ve verilerin bu küme üyesi olma durumunu ifade eden üyelik derecelerinin  $[0, 1]$  aralığında hesaplanması ile önermenin doğruluğu hesaplanır.

Bir bulanık sorgulama aracı tekniği olan SummarySQL, doğal dildeki ifadelere çok yakın ifadeler ile sorgulama yapma imkanı sağladığı için bu tez çalışmasında baz alınmış ve bir veri madenciliği tekniği olarak kullanılmıştır. SummarySQL tekniği ile oluşturulan bulanık sorgulama aracı sayesinde, veriler bulanık kümeler yardımıyla bulanıklaştırma sürecine tabi tutulur ve kural tabanlı sorgulamalar yapmaya olanak sağlar. Dilsel değişkenler sınıfı olarak da adlandırılan (uzun, yüksek, kısa) bulanık kümeler ile oluşturulan sorgular; bulanık sorgulama, fonksiyonel bulanık sorgulama ve dereceli bulanık sorgulama yapısında oluşturulmuş ve her bir sorgu bir kural olarak kabul edilmiştir. Kural çıkarım mekanizması sonucu hesaplanan kuralların doğruluk dereceleri, verilerin yorumunun yapılması aşamasında kullanıcıya yol gösterecek ve sorgulanan veritabanı yapısını göz önüne serecektir.

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

### **SUMMARY**

As data accumulates and number of data source increases, processing of data gets more complicated in usual database systems. One of the most important problem of today's data technology is the difficulty of getting maximum benefit from existing data. In the process of collecting knowledge, getting result and using benefits, fuzzy concept of knowledge should also be analyzed as well as clear and certain concept of the process. Fuzziness is a kind of uncertainty which hasn't got certain structure and as people use daily language, fuzziness comes out in many expressions. Fuzzy query systems use these fuzzy expressions by getting the approach of traditional logic and fuzzy sets theory. In that case fuzzy query language gets the form of structure as the language people use everyday. Processing data and observing hidden structure of data is a kind of data mining technique that can also be accepted as the technique of fuzzy queries.

Fuzzy sets theory which was defined by L. A. Zadeh in 1965, has the capability of processing fuzzy sets that are not crisp. A traditional logic, one proposition is either true or false, in addition it hasn't got the structure that handles uncertainty. But at Zadeh's approach, it uses continuous valued logic which has the values in the range 0 to 1 instead of the two valued logic which has the values either 1 or 0. Also fuzzy propositions are considered as fuzzy sets and every member of these sets gets membership degree in the case of partial or full membership happens. At conclusion, all these processes will give the truth value of the fuzzy proposition.

Since SummarySQL which is one of the fuzzy query tool, uses daily life expressions in its structure, it was preferred to be the base of this study and named as data mining technique. Fuzzy query tool which uses the base facts of SummarySQL, has the process of getting fuzzy expressions from data with the support of fuzzy sets and helps to introduce rule based queries. Fuzzy sets can be also used to construct linguistic term sets or vocabularies like tall, high...etc. By using these terms, we can form fuzzy queries, fuzzy functional queries and gradual functional queries that have indeed rule structure. The truth value of every used rule'll give the validity of queries and the clue about the database which conjecture these rules. The validity of such a summary can be seen as the confirmation of the rule based query.

# 1. GİRİŞ

## 1.1 Belirsizlik

Verinin saklanması için kullanılan veri yapıları, bilginin temsili sürecinde yararlanılan en önemli kaynaklardan biridir. Verinin işlenip, çeşitli çıkarım mekanizmaları ile açığa çıkartılması için kullanılan çok çeşitli yöntemler vardır. Bu aşamada verinin içerdiği belirsizlik yada netliği, seçtiğimiz yönteme bağlı olarak nasıl kullanacağımıza karar vermeliyiz. Çünkü bilginin temsil edilmesi ve çıkarım sürecinde kullanılabilirliği için biçimlendirilmesi, şekillendirilmesi gerekmektedir. Bilgiye belirsizlik ise, bilginin her zaman kesin ve net ifade edilemediği durumlarda ortaya çıkar. Örneğin “Çok Uzun”, “Her men Her men Ağır” gibi ifadeler belirsizlik içeren önermelerdir.

Belirsizlik pek çok nedenden kaynaklanabilir. Karmaşıklık, Çok Yönlülük, Rastgele Seçim Bilgiye Yetersizlik ve Doğal Dildeki Bulanıklık bu nedenlerden birkaçıdır. Öyleyse belirsizlik altında karar vermeye gereken durumlarda ne yapmalıyız? Sistemin davranışı hakkında genel ve detaylı bilgiler bize yol gösterebilir.

Sayısal verilerin çok olduğu kimi durumlarda ise, verileri bulanık ifadeler yardımıyla biçimlendirerek, o veri topluluğu hakkında daha genel bilgilere ulaşmamız mümkündür. SummarySQL [6] de veriyi bu şekilde biçimlendirmek için geliştirilmiş bir Bulanık Sorgulama Aracı Tekniğidir. Bulanık ifadelerle dönüştürülmüş veriden kurallar üretmek suretiyle bir takım sonuçlar çıkarır.

Günlük hayatta karşılaştığımız pek çok ifade aslında içinde belirsizlik barındıran önermelerdir. Bulanık ifade olarak da adlandırabileceğimiz bu önermeler belirsiz bilgi içeren kalıplardır. İnsanlar kendi aralarında kullandıkları bulanık ifadeleri anlamakta pek zorluk çekmezlerken, bilgisayar sistemleri için bu durum geçerli değildir. Bulanık Sorgulama Araçları bu noktada devreye girerek, net ve belirgin tanımlı verileri bulanıklaştırma sürecine tabi tutarak, bilgisayar sistemlerinin sorgu işlemlerinin bu veriyi işleme sürecine olanak sağlar.

Veri Madenciliği ile SummarySQL'i buluşturan en önemli nokta, Veri Madenciliğinin temelinde yatan kural çıkarım mekanizmalarından biri olarak SummarySQL'in kullanılabilir olmasıdır.

## 1.2 Gerçekler ve Kurallar

Bilginin temsil edilmesinde kullanılan gerçekler ve kurallar, bilgi yapıları olarak adlandırılırlar. Gerçekler bilginin temel parçacıkları olarak nitelendirilebilirler. Örneğin, Veritabanından elde ettiğimiz “X cisminin hızı 40 km'dir” ifadesi bir gerçektir. Gerçekler çıkarım mekanizmasında tek başlarına kullanılmazlar. Gerçeklerden çıkarım mekanizması sonucu elde ettiğimiz yeni gerçekler için bir kural yapısı tanımlanmalıdır. Genel olarak kural yapısını tanımlayacak olursak;

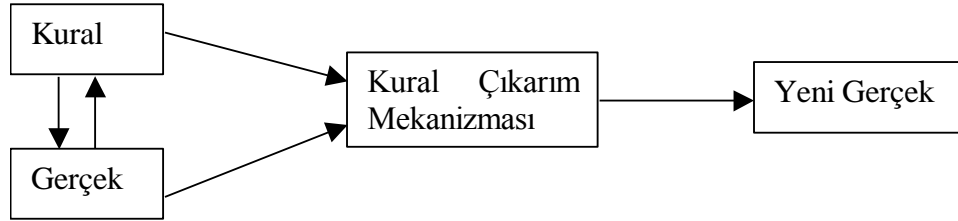
IF KOŞUL THEN SONUÇ

KOŞUL kısmı VE, VEYA operatörleri ile, SONUÇ kısmı da yine VE, VEYA operatörleri ile bileşik biçimde ifade edilebilir.

KOŞUL: a (VE/ VEYA) b

SONUÇ: c (VE/ VEYA) d

Kural yapısı ve Gerçeklerin birlikte kullanılış şekli Şekil (1.1)'de aşağıda gösterilmiştir.



Şekil(1.1) Kural yapısı ve Gerçeklerin kullanılış şekli

### 1.3 Belirgin Bağntı

$A_1, A_2, \dots, A_n$  Belirgin Kü meler ini di kkat e alalım  $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$  ol mak üzere  $(a_1, a_2, \dots, a_n)$  sıralı d iziliş ler i ni n alt kü mesi ne  $A_1, A_2, \dots, A_n$  kü meler i ni n kartezyen ç ar pı m denir.

$A_1 \times A_2 \times \dots \times A_n$  ile gösterilir.

$n=2$  ol mak üzere  $A_1 \times A_2$  kartezyen ç ar pı mı nın her alt kü mesi i kili bağntı (bi nary relation) olarak tanı mlamır.  $X$  ve  $Y$  i ki evrensel kü me ol mak üzere, bu uzay lar ın kartezyen ç ar pı m;

$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$  şekli nde gösterilir.

Ve  $X$  evrensel kü mesi ni n her elemanı  $Y$  evrensel kü mesi ni n her elemanı ile tam olarak ilişkilidir.

Daha belirgin  $R \subseteq X \times Y$  bağntısı ise; 0 veya 1 deęerlerini alabilen  $\chi_R$  karakteristik fonksiyonu ile ifade edilir. 1 deęeri  $x$  ve  $y$  eleman lar ı nın bağntı üzerinden tam olarak ilişkilili, 0 deęeri  $x$  ve  $y$  eleman lar ı nın bağntı üzeri nden ilişkilili ol ma dı kl ar ı nı gösterir.

$\chi_R$  karakteristik fonksiyonu (1.1) for mül ünde gösteril miştir.

$$\chi_R(x, y) = \begin{cases} 1, & (x, y) \in R \\ 0, & (x, y) \notin R \end{cases} \quad (1.1)$$

Örneęin;  $X = \{2, 3\}$

$Y = \{4, 9\}$  dsun.

Ve  $R$  bağntısı da “Karekökü Ol ma” ilişkisi olarak tanı mlansın Evrensel kü me eleman lar ı nın sonlu ol ma sı durumunda  $R$  bağntısı Şekil(1.2)’deki matris şekli nde temsil edilir.

R	2	3
4	1	0
9	0	1

Şekil(1.2) R Bağntısı temsil matrisi

### 1.3.1 Belirgin Bağntı İşlemleri

R ve S,  $X \times Y$  kartezyen çarpım üzerinde tanımlı iki ayrı bağntı olsun. Evrensel küme elemanlarının sonlu olduğu durumu göz önüne alırsak Belirgin Bağntı işlemleri şu şekilde tanımlanır.

1) Boş Bağntı

$$0 = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}$$

2) Tam Bağntı

$$E = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$$

3) Birleşim

$$\chi_{R \cup S}(x, y) = \text{Max}[\chi_R(x, y), \chi_S(x, y)]$$

4) Kesişim

$$\chi_{R \cap S}(x, y) = \text{Min}[\chi_R(x, y), \chi_S(x, y)]$$

5) Tümleneyen

$$\chi_{\neg R}(x, y) = 1 - \chi_R(x, y)$$

6) İçerme

$$R \subseteq S = \chi_R(x, y) \leq \chi_S(x, y)$$

7) Özdeşlik

Q E Tam Bağlı ya da Boş Bağlı olma durumudur.

### 1.3.2 Belirgin Bağlı Kompozisyonu

R; X ve Y

S; Y ve Z evrensel kümeleri üzerinde tanımlı bağlantılar olsun

X ve Z evrensel kümelerinde tanımlı T bağlantısı şu şekilde ifade edilir.

$T = R \circ S$

$$\chi_T(x, z) = \bigvee_{y \in Y} (\chi_R(x, y) \wedge \chi_S(y, z)) \quad (1.2)$$

### 1.4 Bulanık Bağlı

Bulanık Bağlılar da Belirgin Bağlılar gibi, iki evrensel kümenin elemanlarını birbiriyle ilişkilendirir. Ve bağlantı  $[0, 1]$  aralığında değer alabilen üyelik fonksiyonu ile ifade edilir. (1.3) formülünde gösterilmiştir.

A, X evrensel kümesinde,

B, Y evrensel kümesinde tanımlı bulanık bağlantılar olsun

R bulanık bağlantısı  $X \times Y$  Kartezyen çarpımından  $[0, 1]$  aralığında değer alabilen bir bağlantıdır.

$\underline{R} = \underline{A} \times \underline{B} \subseteq X \times Y$

$$\mu_{\underline{R}}(x, y) = \mu_{\underline{A} \times \underline{B}}(x, y) = \min[\mu_{\underline{A}}(x), \mu_{\underline{B}}(y)] \quad (1.3)$$

Özetleyecek olursak;

X evrensel kümesinde tanımlı A bulanık bağlantısı üzerinde, X evrensel kümesinin bir elemanı olan x elemanının alacağı değer ve Y evrensel kümesinde tanımlı B bulanık bağlantısı üzerinde, Y evrensel kümesinin bir elemanı olan y elemanının alacağı değer arasındaki  $\min$  değer bize R Bulanık Bağlısının değerini verecektir.



$\mu_A(x)$ : X evrensel kümesini n elemanlarının n A Bulanık Bağntısı üzeri nde alabileceği deęerler dan üyeli k derecesi deęerleri dir.

$\mu_B(y)$ : Y evrensel kümesini n elemanlarının n B Bulanık Bağntısı üzeri nde alabileceği deęerler dan üyeli k derecesi deęerleri dir.

Örneęi ni ;

$$X = \{65, 80, 100\}$$

$$Y = \{1, 5, 12\}$$

$$\underline{A} = \{ \text{Yüksek Not Ortalama sı} \} = \{0/65, 1/80, 1/100\}$$

$$\underline{B} = \{ \text{Fazla Çalışma Saati} \} = \{0/1, 0.5/5, 1/12\}$$

şekli nde tanı mlarırsa;

$\underline{R} = \underline{A} \times \underline{B}$  matrisi Şekil(1.3)'de ki gibi ifade edilir.

R	0	1	1
0	0	0	0
0.5	0	0.5	0.5
1	0	1	1

Şekil(1.3)  $A \times B$  Bağntısı tensil matrisi

#### 1.4.1 Bulanık Bağntı İşlemleri

$\underline{R}$  ve  $\underline{S}$  X ve Y evrensel kümeleri nde tanı mlı iki bulanık bağntı olsun. Bu dur umda Bulanık Bağntı işlemleri şu şekil de tanı mlanır;

1) Birleşim

$$\mu_{\underline{R} \cup \underline{S}}(x, y) = \text{Max}[\mu_{\underline{R}}(x, y), \mu_{\underline{S}}(x, y)]$$

2) Kesişim

$$\mu_{\underline{R} \cap \underline{S}}(x, y) = \text{Min}[\mu_{\underline{R}}(x, y), \mu_{\underline{S}}(x, y)]$$

3) Tümleneyen

$$\mu_{\neg \underline{R}}(x, y) = 1 - \mu_{\underline{R}}(x, y)$$

4) İçerme

$$\underline{R} \subseteq \underline{S} = \mu_{\underline{R}}(x, y) \leq \mu_{\underline{S}}(x, y)$$

## 1.4.2 Bulanık Bağntı Kompozisyonu

$\underline{R}$ : X ve Y

$\underline{S}$ : Y ve Z evrensel kümeleri üzerinde tanımlı iki bulanık bağntı olsun

X ve Z evrensel kümelerinde tanımlı T bağntısı şu şekilde ifade edilir.

$\underline{T} = \underline{R} \circ \underline{S}$

$$\mu_{\underline{T}}(x, z) = \bigvee_{y \in Y} (\mu_{\underline{R}}(x, y) \wedge \mu_{\underline{S}}(y, z)) \quad (1.4)$$

$\mu_{\underline{T}}(x, z)$ ; X ve Z evrensel kümelerinin her bir elemanının T bulanık bağntısı üzerinde alacağı üyelik derecesi dir.

## 1.5 Bulanık Mantık

Klasik bir deęer adıyla iki deęerli mantığın temel prensibi “her önermenin doğru veya yanlış olmasıdır”. Önermelerin alacağı deęerler 0 yada 1 deęerleriyle kısıtlandırılmıştır. Daha sonra ortaya atılan çok deęerli mantık da ise, önermelerin deęerlerinin mutlak doğrulukta 1, mutlak yanlışlıkta 0 ve belirsiz durumlarda 0.5 deęerlerini alabileceęi gösterilmiştir. Bulanık Mantık ise 60’lı yılların ortalarında Azeri Profesör Lütü Zadeh[1] tarafından üç deęerli mantığın bir uzantısı olarak ortaya atılmıştır.

Bulanık Mantığın temel amacı, bulanık küme teorisini bir araç olarak kullanarak yaklaşım çıkarım yapma sürecine olanak sağlamaktır. Yaklaşım çıkarım yapma sürecinde, belirsiz önermeler ile çıkarım yapmak söz konusu olduğundan, bulanık mantığın odak noktası doğal dildir.

Bulanık Mantık belirsizlik altında çıkarım yapabilmek için

“Uzun”, “Kısa”, “Güçlü”, “Güçsüz” gibi Bulanık Yükl emlerden,

“Genellikle”, “Çoęu”, “Nadiren”, “Bırkısı m”, gibi Bulanık Nceleycilerden ve

Hesaplamalar sonucu bulunan Bulanık Doğruluk Deęerlerinden faydalanır.

### 1.5.1 Bulanık Mantık Yaklaşım Yık arım

Konuşma diliyle ifade edilen pekçok söylem birer bulanık önerme olarak kabul edilir.

X evrensel kümesini tanımlayalım

$X = \{ \text{km cinsinden mesafeler} \} = \{ 1 \text{ km}, 15 \text{ km}, 100 \text{ km} \}$

A Bulanık Bağıntısını tanımlayalım

$A = \{ \text{Uzun Mesafe} \} = \{ 0/1, 0.3/15, 0.7/100 \}$

15 km uzunluğundaki bir mesafe için P önermesini

P = "15 km Uzun Mesafedir" olarak tanımlarsak;

$$\mu_A(x) = 0.3$$

X evrensel kümesini bir elemanı olan 15 km mesafesinin A Bulanık Bağıntısına olan üyelik derecesi 0.3'dür. 0.3 değeri aynı zamanda P önermesini doğruluk değeri dir.

$$\mathbb{T}(P) = 0.3$$

$$\mathbb{T}(P) = \mu_A(x), \quad 0 \leq \mu_A(x) \leq 1 \quad (1.5)$$

Bulanık Mantık (1.5) formülünden de anlaşılacağı gibi önermelerin doğruluk değerlerini aynı zamanda da Bulanık Bağıntılara olan üyelik derecesi değerlerini  $[0, 1]$  aralığında kısıtlamıştır.

P ve Q Bulanık Önermeler olmak üzere aşağıdaki ifadelerin doğruluk değerleri şu şekilde hesaplanır;

1)  $P \vee Q$  (P veya Q)

$$\mathbb{T}(P \vee Q) = \text{Max}(\mathbb{T}(P), \mathbb{T}(Q))$$

2)  $P \wedge Q$  (P ve Q)

$$\mathbb{T}(P \wedge Q) = \text{Min}(\mathbb{T}(P), \mathbb{T}(Q))$$

3)  $\neg P$  (değil P)

$$\mathbb{T}(\neg P) = 1 - \mathbb{T}(P)$$

4)  $P \rightarrow Q$  (P gerektirir Q)

$$\mathbb{T}(P \rightarrow Q) = \mathbb{T}(\neg P \vee Q) = \text{Max}(\mathbb{T}(\neg P), \mathbb{T}(Q))$$

P gerektirir Q ifadesi aynı zamanda bir kural ortaya koymaktadır.

X ve Y evrensel kümelerinde  $A \subseteq X$  ve  $B \subseteq Y$  olmak üzere;

P:  $x \in A$  ve Q:  $y \in B$  şeklinde tanımlanmış bulanık önermeler olsun

Bu durumda  $P \rightarrow Q$  ifadesi (P gerektirir Q, IF  $x \in A$  THEN  $y \in B$  kuralına karşılık gelir.

### 1.5.2 Bulanık Kümeler ve Üyelik Fonksiyonları

Bulanık Küme [2] teorisi, sınırları keskin tanımlarla yapılanayan sınıflamaları, semantik yollar kullanarak ifade etmeye çalışır. “Genç Kadın”, “Küçük Araba”, “Dar Sokak” gibi bulanık ifadeler bu sınıflamalara örnek gösterilebilir.

Bulanık Kümeler, Bulanık olmayan evrensel kümelerden türetilir. Evrensel Küme U, bir odadaki objeler kümesi, telefon listesindeki isimler kümesi, 1 ile 100 aralığındaki tamsayılar kümesi olarak tanımlanabilir. U nun Bulanık alt kümesi; A, B, C, D gibi büyük harf sembolleri ile ifade edilirse Bulanık Alt Küme A bir üyelik fonksiyonu ile ifade edilir.

$$\mu_A: U \rightarrow [0, 1] \quad (1.6)$$

A'nın üyelik fonksiyonu değerini 0.5 yapan U evrensel kümesinin elemanı “Crossover Point” olarak adlandırılır.

U evrensel kümesini [0,100] aralığındaki Yaşlar için oluşturursak, U kümesinin her bir elemanının “Yaşlı” Bulanık Alt kümesine olan üyelik derecesi, aşağıdaki (1.7) formülünde gösterilen “Yaşlı” Bulanık alt kümesini üyelik fonksiyonu ile bulunur.

$$\mu_A = \begin{cases} 0 & 0 \leq u \leq 50 \\ \frac{1}{2} \left( 1 + \left( \frac{u-50}{2} \right)^{-2} \right)^{-1} & 50 \leq u \leq 100 \end{cases} \quad (1.7)$$

Bu durumda “Yaşlı” Bulanık Alt kümesini Cross Over Point değeri 55 olur. 60 yaş değeri nin “Yaşlı” bulanık alt kümesine olan üyelik derecesi, yani “Yaşlı” bulanık alt kümesini üzerinde aldığı değer;

$$\mu_{YAŞLI}(60) = 0.8 \text{ dir.}$$

BULANIK KÜME ifadeleri aşağıdaki formlerde tanımlanmıştır.

İlk önce bir sonlu Belirgin Küme ifadesini ele alalım

$$U = \{u_1, u_2, \dots, u_n\} \quad (1.8)$$

$$U = u_1 + u_2 + \dots + u_n \quad (1.9)$$

$$U = \sum_{i=1}^n u_i \quad (1.10)$$

(1.9) 'daki + ifadesi aritmetik toplam operatöründen çok birleşimi ifade eder ve (1.10) 'daki ifade ile aynı anlamdadır.

Bir Sonlu Bulanık Küme ifadesi ise (1.11) ve (1.12) formlerinde gösterilmiştir.

$$A = \mu_1 u_1 + \dots + \mu_n u_n \quad (1.11)$$

$$A = \sum_{i=1}^n \mu_i u_i \quad (1.12)$$

$\mu_i$ ,  $u_i$  elemanının A'daki üyelik derecesini gösterir. Bir ayrıca seperatör kullanarak çok anlamlı ifadeyi engellemek mümkündür. (1.13) formülünde gösterilmiştir.

$$A = \mu_1 / u_1 + \dots + \mu_n / u_n$$

$$A = \sum \mu_i / u_i \quad (1.13)$$

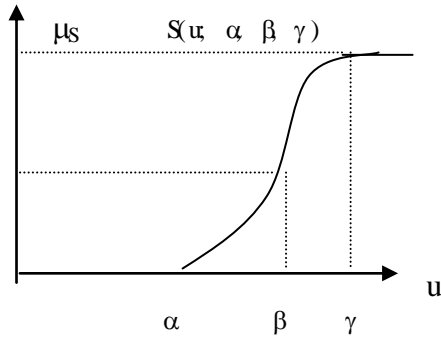
Bir Sonsuz Bulanık Kümeyi ise integral formda aşağıdaki gibi ifade edebiliriz;

$$A \triangleq \int_U \mu_A(u) / u \quad (1.14)$$

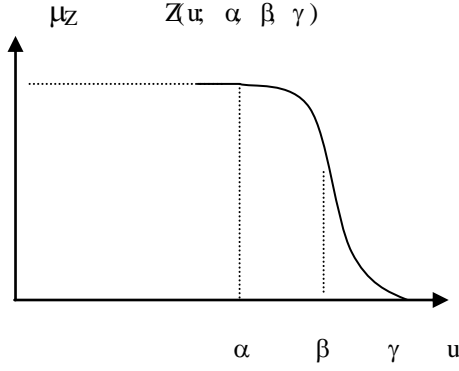
$\mu_A(u)$ ; u elemanının A bulanık kümesine olan üyelik derecesi ve

$\mu_A(u) / u$ ;  $u \in U$  herbir bulanık elemanın birleşimini, integral ifade ile tanımlayabiliriz

En yakın çözümleri sunan üyelik fonksiyonları S\_Fonksiyon,  $\pi$ \_Fonksiyon Z\_Fonksiyon ve Trapezoidal Fonksiyonlardır. S\_Fonksiyon Şekil(1.4)'de, Z\_Fonksiyon Şekil(1.5)'de, Trapezoidal Fonksiyon Şekil(1.6)'da açıklanmaya çalışılmıştır.

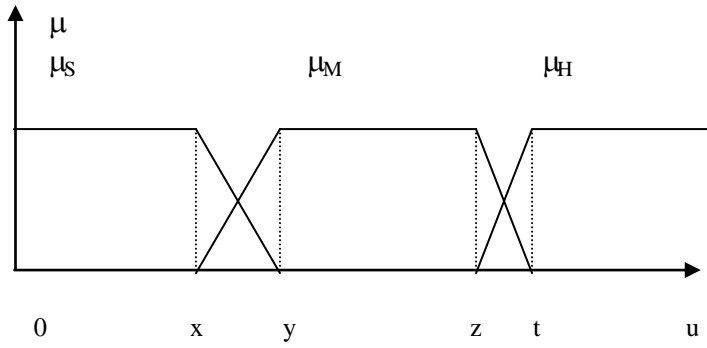


Şekil(1.4) S\_Fonksiyon



Şekil

(1.5) Z\_Fonksiyon



Şekil(1.6) Trapezoidal Fonksiyon

S\_Fonksiyonunu formüle edecek olursak;

$$\begin{aligned}
 S(u; \alpha, \beta, \gamma) &= 0 & u \leq \alpha \\
 &= 2 \left( \frac{u - \alpha}{\gamma - \alpha} \right)^2 & \alpha \leq u \leq \beta \\
 &= 1 - 2 \left( \frac{u - \alpha}{\gamma - \alpha} \right)^2 & \beta \leq u \leq \gamma \\
 &= 1 & u \geq \gamma
 \end{aligned} \tag{1.15}$$

$\pi$ \_Fonksiyonunu formüle edecek olursak;

$$\begin{aligned}
 \pi(u; \beta, \gamma) &= S(u; \gamma - \beta, \gamma - (\beta/2), \gamma) & u \leq \gamma \\
 &= 1 - S(u; \gamma, \gamma + (\beta/2), \gamma + \beta) & u \geq \gamma
 \end{aligned} \tag{1.16}$$

Z\_Fonksiyonunu formüle edecek olursak;

$$\begin{aligned}
Z(u; \alpha \beta \gamma) &= 1 & u \leq \alpha \\
&= 1 - 2((u - \alpha) / (\gamma - \alpha))^2 & \alpha \leq u \leq \beta \\
&= 2((u - \alpha) / (\gamma - \alpha))^2 & \beta \leq u \leq \gamma \\
&= 0 & u \geq \gamma
\end{aligned} \tag{1.17}$$

Trapezoidal Fonksiyonu formüle edecek olursak;

$$\begin{aligned}
\mu_S(u) &= 1 & u \leq x \\
&= (y - u) / (y - x) & x \leq u \leq y \\
\mu_M(u) &= (x - u) / (x - y) & x \leq u \leq y \\
&= 1 & y \leq u \leq z \\
&= (t - u) / (t - z) & z \leq u \leq t \\
\mu_H(u) &= (z - u) / (z - t) & z \leq u \leq t \\
&= 1 & t \leq u
\end{aligned} \tag{1.18}$$

### 1.5.3 Dilsel Değişkenler, Dilsel Pekiştiriciler ve Değil Değiştiricisi

Dilsel Değişkenler[3], bulanık mantıkta karşımıza çıkan bulanık alt küme, bulanık bağlantı kavramlarını, günlük hayatta konuşma dilinde çok sıkça kullandığımız sıfatları kullanarak ifade eder. "Ayşe uzun boyludur" önermesini dikkate alalım Bu önerme bulanık bilgi içermektedir, bunun sebebi "uzun" bulanık alt kümesini kişi den kişiye farklı anlam taşımasıdır. Buradaki "uzun" bulanık alt kümesini artık bir dilsel değişken olarak adlandırabiliriz. Bulanık kümeler olarak da tanımladığımız dilsel değişkenler yardımıyla matematiksel olarak temsil edilen bilgi den başka bilgi parçacıkları elde edilebilir. Bu şekilde elde etme işlemi, dilsel pekiştiriciler ve dilsel değişkenlerin Değil değiştiricisi kullanılarak yapılabilir.

Dilsel değişkenlerin, dilsel pekiştiriciler ve Değil değiştiricisi ile birlikte kullanılabilir olmaları bize bulanık mantıkta bulanık alt küme değerlerini etkileyen pek çok etken olabileceğini gösterir. Halbuki geleneksel sistemlerde kullanılan tek modifikatör Değil değiştiricisi dir.

“Ayşe uzun boyludur” önermesinde A bulanık alt kümesi “Uzun Boy”dur. “Uzun Boy”u artık bir dilsel değişken olarak kabul edebileceğimizden bahsetmiştik. Dilsel pekiştiriciler ise, dilsel değişkenlerin değerlerini pekiştirici özellikler taşıyan terimlerdir. Ve dilsel değişkenler oluşturulurken önemli rol oynarlar. Örneğin; “Çok”, “Çok Çok”, “Yaklaşık” gibi terimler dilsel pekiştiricilere örnek olarak verilebilirler.

“Ayşe uzun boyludur” önermesinde Ayşe’nin sırasıyla “Uzun”, “Çok Uzun”, “Çok Çok Uzun”, “Yaklaşık Uzun”, “Biraz Uzun” olarak tanımlanan, dilsel pekiştiriciler kullanılarak tanımlanmış bulanık alt kümelerine olan üyelik derecesi sırasıyla (1.19)’da gösterilmiştir.

$$\mu_{UZUN\ BOY}(Ayşe) = 0.95$$

$$\mu_{ÇOK\ UZUN\ BOY}(Ayşe) = \mu_{UZUN\ BOY}^2(Ayşe) = (0.95)^2$$

$$\mu_{ÇOK\ ÇOK\ UZUN\ BOY}(Ayşe) = \mu_{UZUN\ BOY}^4(Ayşe) = (0.95)^4$$

$$\mu_{YAKLAŞIK\ UZUN\ BOY}(Ayşe) = \mu_{UZUN\ BOY}^{0.5}(Ayşe) = (0.95)^{0.5}$$

$$\mu_{BİR AZ\ UZUN\ BOY}(Ayşe) = \mu_{UZUN\ BOY}^{1.25}(Ayşe) = (0.95)^{1.25} \quad (1.19)$$

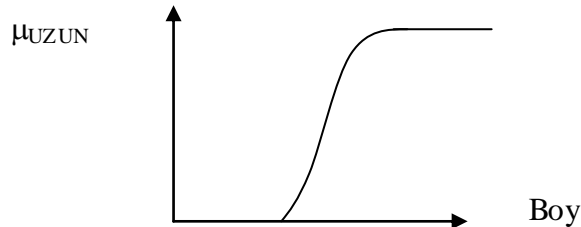
Görüldüğü gibi dilsel pekiştiriciler için kabul görmüş bazı matematiksel modeller olan yoğunlaştırma, biraz yoğunlaştırma ve yayma modellerinin kullanıldığı açıktır.

$$\mu_{yog}(A)(x) = (\mu_A(x))^2$$

$$\mu_{bir}(A)(x) = (\mu_A(x))^{1.25}$$

$$\mu_{yay}(A)(x) = (\mu_A(x))^{0.5}$$

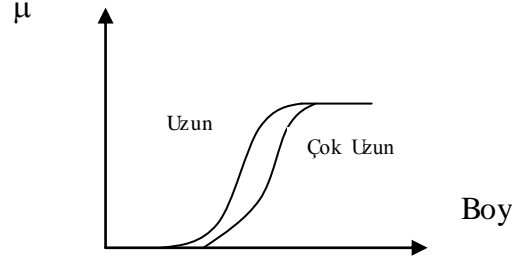
“Uzun” bulanık alt kümesinin üyelik fonksiyonu Şekil(1.7)’deki gibi tanımlanır;



Şekil(1.7) Uzun bulanık alt kümesinin üyelik fonksiyonu



Dilsel pekiştirici ile tamamlanmış “Çok Uzun” bulanık alt kümesini üyelik fonksiyonu Şekil(1.8)’deki gibidir.



Şekil(1.8) Uzun ve Çok Uzun bulanık alt kümelerinin üyelik fonksiyonları

Bulanık alt küme yada pekiştirilmiş bulanık alt küme kavramlarıyla elde ettiğimiz üyelik derecesi değerlerinin  $[0, 1]$  aralığında değerler alabileceğinden bahsetmiştik. Bu değerleri 1 taşımasından çıkarttığımız zaman elde edeceğimiz sonuç, bize Değil değıştiricisini nasıl davrandığı hakkında bilgi verecektir. ”Ayşe çok uzun boylu değildir” önermesini dikkate alalım Ayşe’nin “Çok Uzun Boy” bulanık alt kümesine olan üyelik derecesi şu şekilde gösterilmiş iken,

$$\mu_{\text{ÇOK UZUN BOY}}(\text{Ayşe}) = \mu^2_{\text{UZUN BOY}}(\text{Ayşe}) = (0.95)^2$$

Ayşe’nin ”Değil Çok Uzun Boy” bulanık alt kümesine olan üyelik derecesi aşağıdaki gibidir.

$$\mu_{\text{DEĞİL ÇOK UZUN BOY}}(\text{Ayşe}) = (1 - \mu^2_{\text{UZUN BOY}}(\text{Ayşe})) = 1 - (0.95)^2$$

#### 1.5.4 Bulanık Önerme Ve Dilsel Niteleyicilerin Bulanık Önermedeki Yeri

Geleneksel sistemlerde sadece iki nitelleyen “her” ve “bazı” kullanılırken, bulanık mantıkta pek çok nitelleyen kullanılabilir. Örneğin, “Bırkaç”, “Bırkısı mı”, ”Çoğu”, “Azbırkısı mı”, ”Pek az” gibi terimler dilsel niteleyiciler olarak adlandırılabilirler. Bulanık mantıkta dilsel niteleyiciler birer bulanık alt küme olarak ele alınırlar. Dilsel niteleyicilerin nasıl kullanıldıklarını tam olarak anlayabilmek için, bulanık önerme tanımına tekrar göz atalım

Bulanık yüklem içeren önermelere Bulanık Önerme[4] denir. Günlük hayatta karşımıza çıkan pek çok ifadenin aslında birer bulanık önerme olduğundan bahsetmiştik. Örneğin, “Elif kısa saçlıdır” veya “x büyük bir sayıdır” ifadeleri birer bulanık önermedir. Ve genel olarak şu şekilde ifade edilirler.

$X \text{ F' dir}$

Buradaki  $X$  bir nesnelere kümesi,  $F$  ise bulanık bir alt küme yada dilsel değişken olarak adlandırılabilir. Dilsel değişkenlerin bulanık alt küme olarak da adlandırılmasının sebebi, bu değişkenlerin bulanık alt kümeler yardımıyla ifade edilebiliyor olmasıdır.

Her  $x \in X$  için  $x \text{ F' dir}$  önermelerinin doğruluğu saptanabilirse, genel olarak  $X \text{ F' dir}$  ifadesinin doğruluğunu saptamış oluruz.

$X = \{x_1, x_2, \dots, x_n\}$  ise,  $x \text{ F' dir}$  önermelerinin doğruluk dereceleri (1.20) formülünde gösterilmiştir.

$$\text{Dogruluk } (x \text{ F}) = \mu_F(x) \quad i=1..n \quad (1.20)$$

Genel olarak  $X \text{ F' dir}$  önermesini doğruluk derecesi ise ancak tüm  $x$  değerleri için  $x \text{ F' dir}$  önermelerinin doğruluk derecesi bulunduğundan sonra hesaplanabilir.

$$r = (1/n) \sum_{i=1}^n \mu_F(x_i) \quad (1.21)$$

Dilsel pekiştiriciler ile tamamlanmış bulanık alt kümeler ile oluşturulan bulanık önerme niz “Eğer çok kısa saçlıdır” şeklinde ifade edilirse, bulanık önermenin genel yapısı şu şekilde olur.

$X \text{ mF}$

Dilsel nicelleyicilerin bulanık önermelerde nasıl kullanıldığını inceleyecek olursak; ilk önce yapısında dilsel nicelleyici barındıran bir önerme tanımlamalıyız. “Çoğu insanlar çok kısa saçlıdır” bulanık önermesini dikkate alalım. Dilsel nicelleyici ile oluşturduğumuz önermenin genel yapısı şu şekilde olacaktır.

$QX \text{ mF}$

Böyle bir önermenin doğruluğunu saptamak için önce  $X \text{ mF' dir}$  önermesini doğruluk derecesi olan  $r$  bulunur.

$$\text{Dogruluk } (x \text{ mF}) = \mu_{mF}(x) \quad i=1..n$$

$$r = (1/n) \sum_{i=1}^n \mu_{mF}(x_i) \quad (1.22)$$

Dilsel nicelleyicilerin daha önce bulanık alt kümeler olarak ele alındığından bahsetmiştik. Bulduğumuz  $r$  doğruluk derecesinin, dilsel nicelleyiciyi tanımlayan üyelik fonksiyonunda aldığı değer ise bize,  $QX = \mu_F^r$  dir önermesini doğruluk derecesini verecektir. Önermenin doğruluk derecesi  $r^*$  ile ifade edilecek olursa (1.23) formülünü elde ederiz.

$$r^* = \mu_Q(r) \quad (1.23)$$

## 1.6 Basit Bulanık Önerme Ve Bileşik Bulanık Önerme

Basit bulanık önermeler için “Ayşe uzun boyludur”, “Ayşe çok ağır değildir” bulanık önermelerini örnek verebiliriz. Önermelerin genel ifade edilmiş tarzları sırasıyla gösterilmiştir.

$X = F$  dir

$X = G$  dir

Basit bulanık önermeleri, “ve” yada “veya” mantıksal operatörleriyle birleştirerek oluşturulan yeni bulanık önermeler ise bir bileşik bulanık önerme olacaktır. Ve aşağıda belirtilen ifadeler elde edilecektir.

$X = F$  dir ve  $X = G$  dir

$X = F$  dir veya  $X = G$  dir

“Ayşe uzun boyludur ve Ayşe çok ağır değildir” bulanık önermesini dikkate alalım. Bileşik bulanık önerme farklı dilsel değişkenlere ait bulanık kümeler içeriyorsa, bileşik önermenin doğruluk derecesi daha önce Bulanık Bağlantı İşlemleri bölümünde anlatıldığı gibi şu şekilde hesaplanır.

Ve mantıksal operatörlerle oluşturulan bileşik bulanık önerme şu şekilde olsun.

$X = D$  dir ve  $Y = E$  dir

Her  $x \in X$  ve  $y \in Y$  için önermenin doğruluk derecesi (1.24) formülündeki gibi olacaktır.

$$\text{Doğruluk}(x \in D \text{ ve } y \in E) = \min[\mu_D(x), \mu_E(y)] \quad i=1 \dots n$$

$$r = (1/n) \sum_{i=1}^n \min[\mu_D(x_i), \mu_E(y_i)] \quad (1.24)$$

Veya mantıksal operatörlerle oluşturulan bileşik bulanık önerme ise aşağıdaki gibi olsun.

$X$  D'dir veya  $Y$  E'dir

Bu durumda her  $x \in X$  ve  $y \in Y$  için önermenin doğruluk derecesi (1.25) formülündeki gibi olacaktır.

$$\text{Doğruluk}(x \text{ D veya } y \text{ E}) = \max[\mu_D(x_i), \mu_E(y_i)] \quad i=1, \dots, n$$

$$r = (1/n) \sum_{i=1}^n \max[\mu_D(x_i), \mu_E(y_i)] \quad (1.25)$$

## 2. HİPOTEZ ÇIKARIMI İÇİN ESNEK BİR BULANIK SORGULAMA DİLİ

### 2.1 Giriş

Günümüzde Veritabanı konusu, her dakika gelişmekte olduğundan ve bilgilerin saklanması da kullanılan en etkin yöntemi oluşturduğundan dolayı incelenmeye değer bilimsel konuların başında gelmektedir. Veritabanları konusundaki araştırmalar, Bilişim yönetimi sistemlerinde çokça kullanılmakta ve karar verme aşamasında, veritabanının barındırdığı verileri analiz edebilmek için veritabanı yönetimi sistemlerinin desteğine ihtiyaç duyulmaktadır. Veritabanı yönetimi sistemlerine karar verebilmek yeteneğinin kazandırılması, veritabanı üzerindeki sorgulamanın verimliliğini artırıcı bir etken olmuştur. Kullanıcı tarafından kullanıcı arayüzleriyle oluşturulacak sorgulama, veritabanındaki verilerin nasıl saklandığı konusunda oldukça bağlantılıdır. Bu bağlantıyı kullanıcının oluşturabilmesi için, veritabanının kullandığı veri modeli, veritabanındaki alanlar ve bunların yapısı, alanlar üzerindeki indeksler ve kısıtlar hakkında yeterli bilgiye kullanıcı sahip olmalıdır. Fakat kullanıcıların her zaman için bu bilgilere yeterli derecede sahip olmadıkları düşünülürse, burada devreye Esnek Bulanık Sorgulama Dilleri girmektedir. Kullanıcının net ve belirgin olarak ifade edemediği alanlar için bir bulanıklaştırma süreci uygulayan ve bu süreç sonunda kullanıcının günlük dildeki ifadeleri kullanarak sorgulama yapmasına olanak sağlayan esnek bulanık sorgulama dillerinden biri olan SummarySQL bu tezde incelenmeye çalışılmıştır. SummarySQL dili ile sorgu cümlecikleri; Koşul ve Sonuç bölümlerinde bulanık önermeler içeren kurallar şeklinde oluşturulmuş ve bu kuralların veritabanında sağlanma dereceleri olan doğruluk dereceleri ile verilerin kural içinde geçen bulanık önermeleri sağlama dereceleri hesaplanmaya çalışılmıştır.

Veritabanında göz önüne çıkmayan gerçekleri, verinin yapısında gizli kalmış özellikleri açığa çıkarmak için bir takım kurallar yardımıyla, kural çıkarım mekanizmaları oluşturan bir veritabanı sorgulama aracı olan Veri madenciliği tekniği, bize bulanık sorgulama araçlarının bu tekniğe ne kadar yakın olduğunu göstermektedir.

SummarySQL ile Veri madenciliğini buluşturan en önemli nokta kural çıkarım mekanizmaları sayesinde verinin içinde ilgi çekici ve beklenmeyen, tahmin edilemeyen özellikleri göz önüne seriyor olmasıdır. Bu bakımdan SummarySQL ile oluşturduğumuz bir bulanık sorgulama dili, bir Veri madenciliği tekniği olarak sayılabilir.

1991 yılında Yager[5] tarafından ortaya atılan dilsel değişkenler sınıfının bu dil ile yorumlanması, tündengelim mantığına dayandırılmaktadır. SummarySQL dili bulanık önermeleri yorumlayarak, kesin ifadeler ile oluşturulan kuralların daha esnek bir biçimde değerlendirilmesini sağlar.

## 2.2 Bulanıklaştırma Süreci

Dilsel değişkenlerin kullanılması suretiyle oluşturduğumuz bulanık önermeler, esnek bulanık sorgulama dili olan SummarySQL dilinde kurallar yardımıyla ifade edilir. Oluşturulan kurallar ve bulanık fonksiyonel bağımlılık özellikleri kullanılarak ortaya çıkan çıkarım mekanizması sürecinde, kullanılabilir ve bizi sonuca götürebilecek veriler elde edilebilir. Günümüzde artık her alanda kullanım ihtiyacı hissedilen veritabanı yönetim sistemleri sayesinde, saklanan veriye özgü nitelikler işlenerek açığa çıkartılabilir ve veriyi kullanan kullanıcıya sunulmaktadır. Soru cümleciklerinin yapısındaki kesin ve belirgin ifadeler çoğu zaman net olarak ifade edilebilirler. Bu soru cümleciklerinin yapısındaki kesin sınırları daha esnek bir hale getirip kullanıcıya daha yakın arayüzler oluşturabilmemiz için bir bulanıklaştırma süreci gerekmektedir.

Kesin ifadelerle oluşturulmuş bir kural olan “EĞER kişinin boyu>2.0 m İSE kişinin mesleği Basketbolcu” kuralı katı sınırlar içermektedir. Bu sınırları daha esnek bir hale getirmemizi sağlayan dilsel değişkenler olan “Küçük”, ”Orta”, “Büyük”, “Hızlı”, “Yavaş”, “Uzun”, “Kısa” vs ve dilsel nicelleyiciler olan “Çoğu”,

“Az bir kısım”, “Pekçoğu”, “He men he men hepsi” vs gibi ifadeler sayesinde bulanık önermelerle ifade edebildiğimiz şu ifadeyi oluşturabiliriz

“Bir kısım uzun boylu insanların mesleği basketbolcudur”

Bulanık önermeyi bulanık kural haline getirirsek ise şu ifadeyi elde ederiz

EĞER Bir kısım {insanların boyu=uzun} İSE {insanların mesleği basketbolcu}

Bulanık kuralda karşımıza çıkan dilsel nicelleyici “Bir kısım”, bize kuralın doğruluk derecesi hakkında bilgi verir. Bulanık kuralın doğruluk derecesini, kesin ifadelerle oluşturulan kuralı veritabanında gerçekleyen verilerin yüzdesi ile kıyaslayabiliriz

## 2.3 Bulanık Sorgu Cümlesi Analizi

### 2.3.1 Basit ve Bileşik Bulanık Sorgu Cümleleri

Veritabanında verileri yorumlarken kullandığımız bir veritabanı sorgulama dili olan SQL ‘de sorgu cümlelerini yapıları şu şekildedir.

```
SELECT Bütün insanlar
```

```
WHERE boy >= 1.8m
```

Oysa bulanık önermeler ile oluşturduğumuz bir SummarySQL sorgu cümlesinin yapısı şu şekilde olacaktır.

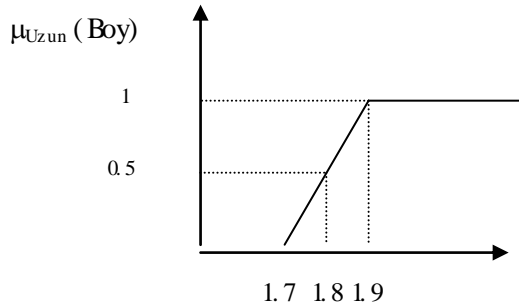
```
SELECT Bütün insanlar
```

```
WHERE Boyu uzun
```

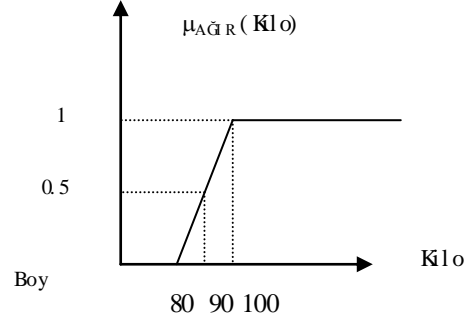
Sorgu cümleciği yapısındaki “uzun” kullanımını tanımladığı dilsel değişkenler kümesini bir elemanıdır. Ve bu kümenin her bir elemanı için tanımlanan üyelik derecesi fonksiyonları sayesinde bulanık kuralın doğruluk derecesine ulaşmak mümkündür.  $V_T$  tanımladığımız veya oluşturduğumuz bir veritabanı,  $q_i$  ise bu veritabanına ait bir eleman olsun. Her  $q_i \in V_T$  veritabanındaki bir kaydı oluşturacaktır. A ise veritabanında tanımladığımız alanların özellikler kümesi olsun.  $A = \{a_1, \dots, a_n\}$  şeklinde ifade edebileceğimiz bu özellikler kümesi, tanımladığımız dilsel değişkenler domain’i içinde geçerli olacaktır.  $q_i.a_j$  ifadesi ise  $q_i$  elemanının  $a_j$  özelliği olarak tanımlanabilir.

Bulanık sorgu analizinde dilsel değişkenlerin üyelik fonksiyonları önemli rol oynamaktadır. Uzun dilsel değişkenini, Boy domain’i üzerinde tanımlayan bir üyelik

fonksiyonu  $\mu_{\text{Uzun}}(\text{Boy})$  şeklinde ifade edilirken,  $\mu_{\text{Uzun}}(1.8)$  ifadesi 1.8 m'nin “Uzun” dilsel değişkeni için tanımlanan üyelik fonksiyonunda aldığı değeri verecektir. Şekil(2 1a)  $\mu_{\text{Uzun}}(\text{Boy})$ , Şekil(2 1 b)  $\mu_{\text{AĞIR}}(\text{Kilo})$  üyelik fonksiyonunu ifade eder.



Şekil (2 1a)



Şekil(2 1b)

Bul anık sorgula nın sonucunda bir bul anık kü me oluşur ve bu kü menin her bir ele manının üyelik derecesi aslında o ele manın bul anık sorguyu sağ la ma derecesi dir. 3 kayıttan oluş an bir VT oluştur alı m

$$VT = \{ (Ayşe, 1.8m, 100kg), (Ahmet, 1.7m, 90kg), (Hasan, 1.9m, 85kg) \}$$

VT' deki alan özelliklerinin oluşturduğu kü me ise;

$$A = \{ Ad, Boy, Kilo \}$$

VT' deki uzun boylu insanlar için oluşturacağımız basit bir bul anık sorgu cümlesi artık (2 1) for mül ündeki gibi ifade edilebilir.

$$Q_{BOY=UZUN}(VT) = ? \quad (2.1)$$

(2 1)'deki bul anık sorgunun genel yapısını daha açık ifade edecek olursak (2.2)'deki gibi bir ifade oluşturabiliriz.

$$Q_{BOY=UZUN}(VT) = \{ q \in VT \mid \mu_{\text{Uzun}}(q.Boy) \} \quad (2.2)$$

Bul anık sorgu ifadesindeki  $\mu_{\text{Uzun}}(q.Boy)$ , VT' deki her bir kayıtlı n Boy domain'inde tanı mlı olan verisinin, bir dilsel değişken olan “Uzun” üyelik fonksiyonu üzerindeki üyelik derecesi dir.

Bul anık sorgu sonucunda oluş an bul anık kü me ise aşağıdaki gibi olacaktır.

$$\{ 1/(Hasan, 1.9m, 85kg), 0.5/(Ayşe, 1.8m, 100kg), 0/(Ahmet, 1.7m, 90kg) \}$$



Basit sorgu cümlesinin bulanık önerme şeklindeki ifadesi “VT’deki uzun boylu insanlar” olarak tanımlanabilir.

“Ve”, “Veya” mantıksal operatörleriyle oluşturulan bulanık sorgu cümleleri ise bileşik bulanık sorgu cümleleri olarak adlandırılırlar. VT’deki uzun boylu ve ağır kilolu insanlar için oluşturacağımız Bileşik bulanık sorgu cümlesi (2.3)’deki gibi ifade edilebilir.

$$Q_{BOY=UZUN \wedge KILLO=AĞIR} (VT) =? \quad (2.3)$$

(2.3)’deki bulanık sorgunun genel yapısını daha açık ifade edecek olursak (2.4)’deki gibi bir ifade oluşturabiliriz

$$Q_{BOY=UZUN \wedge KILLO=AĞIR} (VT) \\ = \{q \in VT \mid \mu_{Uzun}(q \cdot Boy) \wedge \mu_{Ağır}(q \cdot Kilo)\} \quad (2.4)$$

Ve mantıksal operatörünün Mn operatörü olarak kullandığını bu bölümün başında belirtmiştik (2.5)’de Mn operatörü kullanılarak oluşturulan bulanık sorgu, (2.4)’deki ifade ile aynı anlama gelmektedir.

$$Q_{BOY=UZUN \wedge KILLO=AĞIR} (VT) \\ = \{q \in VT \mid \text{Min}(\mu_{Uzun}(q \cdot Boy), \mu_{Ağır}(q \cdot Kilo))\} \quad (2.5)$$

Bileşik sorgu sonucunda oluşan bulanık küme ise aşağıdaki gibi olacaktır.

$$\{0.5/(Ayşe, 1.8 \text{ m} 100\text{kg}), 0.25/(Hasan, 1.9 \text{ m} 85\text{kg}), 0/(Ahmet, 1.7 \text{ m} 90\text{kg})\}$$

Bileşik sorgu cümlesinin bulanık önerme şeklindeki ifadesi “VT’deki uzun boylu ve ağır kilolu insanlar” olarak tanımlanabilir.

Aynı şekilde “Veya” mantıksal operatörü Max operatörü olarak kullanılır ve “VT’deki uzun boylu veya ağır kilolu insanlar” önermesi için oluşturulan bileşik bulanık sorgu cümlesi (2.6)’daki gibi ifade edilir.

$$Q_{BOY=UZUN \vee KILLO=AĞIR} (VT) =? \quad (2.6)$$

(2.6)’daki bulanık sorgunun genel yapısını daha açık ifade edecek olursak (2.7)’deki gibi bir ifade oluşturabiliriz

$$Q_{BOY=UZUN \vee KILLO=AĞIR} (VT) \\ = \{q \in VT \mid \mu_{Uzun}(q \cdot Boy) \vee \mu_{Ağır}(q \cdot Kilo)\} \quad (2.7)$$

(2 8)'de  $\text{Max}$  operatörü kullanılarak oluşturulan bulanık sorgu (2 7)'deki ifade ile aynı anlamı gelmektedir.

$$Q_{\text{BOY=UZUN} \wedge \text{KILLO=AĞIR}}(\text{VT}) \\ =\{q \in \text{VT} \mid \text{Max}(\mu_{\text{Uzun}}(q \cdot \text{Boy}), \mu_{\text{Ağır}}(q \cdot \text{Kilo}))\} \quad (2.8)$$

Basit ve Bileşik bulanık sorgu cümlelerinin doğruluk değerlerini bulunması için aşağıdaki adımları izlenmelidir.

$$\text{VT}=\{q_1, \dots, q_n\}$$

S=dilsel değişken

A=VT' deki alan özellikleri

r=Bulanık sorgu cümlesinin doğruluk derecesi olmak üzere,

Önce her bir  $q \in \text{VT}$  için  $\mu_S(q \cdot A)$  hesaplanır.  $\mu_S(q \cdot A)$ , VT' deki her bir  $q$  kaydının dilsel değişkenlerin tanımlı olduğu üyelik fonksiyonuna olan üyelik derecesi dir.  $r$  doğruluk derecesinin bulunması için ise, dilsel değişkenlerin üyelik fonksiyonuna bir şekilde üye olan kayıtların, tüm VT' deki kayıt sayısına olan oranının hesaplanması gerekir.

### 2.3.2 Dilsel Nceleyiciler İle Bulanık Sorgulama

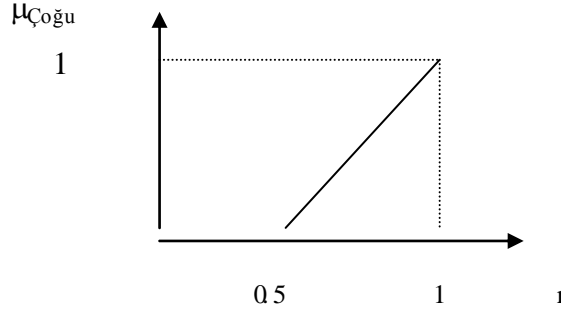
“VT' deki çoğu insan uzun boyludur” bulanık önermesi “Çoğu” gibi bir dilsel nceleyici içermektedir. Bu bulanık önerme için oluşturacağımız sorgu cümlesinin doğruluk derecesini hesaplayabilmek için, önce (2 9)'daki sorgunun doğruluk derecesi hesaplanmalıdır.

$$Q_{\text{BOY=UZUN}}(\text{VT}) \\ =\{q \in \text{VT} \mid \mu_{\text{Uzun}}(q \cdot \text{Boy})\} \quad (2.9)$$

$r$ , (2 9)'daki sorgu cümlesinin doğruluk derecesi olarak tanımlanacak olursa şu şekilde hesaplanır.  $n$ , VT' deki sorgulama yapmak istediğimiz kayıtların sayısıdır.

$$r=(1/n) \sum_{i=1}^n \mu_{\text{Uzun}}(q_i \cdot \text{Boy}) \quad (2.10)$$

“Çoğu” dilsel nicelleyicisinin üyelik fonksiyonu Şekil(2.2)’deki gibi tanımlansın.



Şekil (2.2) Çoğu Dilsel Nicelleyicisi Üyelik Fonksiyonu

r doğruluk derecesinin “Çoğu” dilsel nicelleyicisinin üyelik fonksiyonu üzerinde alacağı değer “VT’deki Çoğu insan uzun boyludur” bulanık önermesini doğruluk derecesinin verecektir. (2.11)’de gösterilmiştir.

$$\tau = \mu_{\text{Çoğu}}(r) \quad (2.11)$$

#### 2.4 Bulanık Kural Çıkarma Mekanizması

Bulanık kuralların genel yapısı “EĞER X İSE Y’dir” şeklinde ifade edilebilir. Bir bulanık kuralı aşağıdaki şekilde tanımlanabilir.

“EĞER Koşul İSE Sonuç”

Bulanık kuralların Koşul ve Sonuç kısımlarını bulanık önermeler ile oluşturmak mümkündür. Basit bulanık önermeleri ve, veya mantıksal operatörleri ile birleştirerek bileşik koşul ve bileşik sonuçlu kurallar elde edilebilir. “VT’deki uzun boylu insanlar ağırlıdır” bulanık önermesinden bir kural türetilecek olursa şu ifade oluşur.

“EĞER Uzun Boylu İSE Ağırlıdır”

Oluşturulan kuralın geçerliliğinin ispatına çalışalım. Bunun için kuralın başına bir dilsel nicelleyici de getirilebilir. Böylece “VT’deki Çoğu uzun boylu insanlar ağırlıdır” bulanık önermesinden türeteceğimiz kuralı ifade edecek olursak;

“EĞER Çoğu Uzun Boylu İSE Ağırlıdır” ifadesini elde ederiz

Bu kural ile uzun boylu insanlar kümesindeki çoğu insanın ağırlı olduğunu iddia ediyoruz. Uzun boylu insanlar kümesi artık bulanık bir veritabanı oluşturur. Ve

bu bulanık veritabanındaki her bir eleman “Uzun” dilsel deęişkeninin üyelik fonksiyonuna geçerli bir üyelik derecesi ile üyedir. Yani  $\mu_{Uzun}(q.Boy)$ ,  $[0, 1]$  aralığındaki deęerler için geçerlidir.  $Q_{BOY=UZUN}(VI) = ?$  Sorusu sonucu artık bulanık bir veritabanı olacaktır. Ve kuralın genel olarak ifade edilmiş biçimi (2.12)’deki gibi olacaktır.

$$\begin{aligned} & \sum_{\text{Çoęu}} (q \in \text{Uzun Boylu insanlar} \mid q.Kilo=Aęır) (\text{Uzun İnsanlar}) \\ & = \sum_{\text{Çoęu}} (q \in \{ q \in VT \mid \mu_{Uzun}(q.Boy) \} \mid \mu_{Aęır}(q.Kilo) ) \end{aligned} \quad (2.12)$$

“EĞER Çoęu Uzun Boylu İSE Aęır Kıldur” bulanık kuralının doğruluk derecesini bulanık önermeyi oluşturarak hesaplamaya çalışalım

“VT’deki QR insan S’dir” bulanık önermesi için r doğruluk derecesi (2.13)’deki gibi olacaktır.

$$r = \frac{\sum_{i=1}^n T(\mu_S(q), \mu_R(q))}{\sum_{i=1}^n \mu_R(q)} \quad (2.13)$$

$T_{\text{Norm}}$  operatörünü  $M_{\text{in}}$  operatörü olarak kabul edersek, her bir kaydın S ve R dilsel deęişkenlerinin tamamı olduğu üyelik fonksiyonlarına olan üyelik derecelerinin  $M_{\text{in}}$ ’larının toplamının, her bir kaydın R dilsel deęişkeninin tamamı olduğu üyelik fonksiyonuna olan üyelik derecelerinin toplamına olan oran r doğruluk derecesini verecektir.

R dilsel deęişkeninin oluşturduğu bulanık kümeyi bulanık veritabanı olarak kabul edersek, r doğruluk derecesi (2.14)’deki gibi ifade edilir.

$$r = \frac{\sum_{i=1}^n T(\mu_S(q), \mu_{VT1}(q))}{\sum_{i=1}^n \mu_{VT1}(q)} \quad (2.14)$$

Son olarak r doğruluk derecesinin “Çoęu” dilsel nicelleyicisinin üyelik fonksiyonuna olan üyelik derecesi bulunarak kuralın doğruluk derecesine ulaşılmış olur.

$$\tau = \mu_{\text{Çoęu}}(r) \quad (2.15)$$

## 2.5 Bulanık Fonksiyonel Bağlılık Kural Çıkartım Mekanizması

Birbirine benzer verilerin, bulanık alt kümeler yardımıyla ifade edilmesi bulanık fonksiyonel bağılılık olarak adlandırılır. "VT'deki Çoğu insanın geliri birbirine benzerdir" bulanık önermesinin doğruluk derecesi, birbirine benzer geliri olan kayıtların VT'deki ölçümünü verecektir. Sa, VT'deki Gelir alan özelliği olarak tanımlanırsa, bulanık fonksiyonel bağılılıkların genel olarak yapısı (2.16)'daki gibi ifade edilir.

$$\sum (q \in VT | q.Sa \approx Sa) \quad (2.16)$$

İfadedeki  $\approx$  sembolü benzerlik yakınlık fonksiyonu olarak tanımlanır ve iki nesne arasındaki benzerlik ölçüsünün değerini verir. BFB'de VT'deki her bir kayıt potansiyel bir karşılaştırma değeridir. Eğer  $q_i$  kaydı için benzerlik ölçüsü oluşturmaya çalışırsak;

1- Benzerlik ölçüsünü hangi domain üzerinde seçeceğimize karar vermeliyiz

$q.Sa$

2- VT'deki her bir kaydın seçtiğimiz özelliğinin, VT'deki diğer tüm kayıtlarla olan benzerlik ölçüsünü hesaplamalıyız.

$$\sum (q \in VT | q.Sa \approx q.Sa)$$

Sonuç olarak bulanık fonksiyonel bağılılık için tanımladığımız "VT'deki insanların geliri birbirine benzerdir" bulanık önermesinin, bulanık sorgu cümlesi (2.17)'deki gibidir.

$$Q_{\text{Gelir}} = Sa \wedge \sum (q_j \in VT | q_j.\text{Gelir} \approx SA) \quad (VI)$$

$$= \{q \in VT | \sum (q_j \in VT | \mu_k (q.\text{Gelir}, q_j.\text{Gelir}))\} \quad (2.17)$$

(2.17)'deki  $\mu_k (q.\text{Gelir}, q_j.\text{Gelir})$  ifadesi  $q_j$  elemanının gelirinin ve  $q$  elemanının gelirinin benzerlik fonksiyonunda yerine konulması ile bulunan benzerlik ölçüsünün değeridir. En yüksek benzerlik ölçüsüne sahip kayıtlar, birbirine en yakın gelire sahip kayıtlar olacaktır.

"VT'deki insanların yaşı birbirine benzerdir" bulanık önermesinin bulanık fonksiyonel bağılılıkta ne anlama geldiğini anlamak için Yaş ve Gelir alan özelliklerinden oluşan bir VT'yi Tablo(2.1)'deki gibi tanımlayalım

Tablo(2.1) Kaynak VT tablosu

Id	Yaş	Gelir
Q <sub>1</sub>	55	500,000,000 TL
Q <sub>2</sub>	50	450,000,000 TL
Q <sub>3</sub>	52	750,000,000 TL
Q <sub>4</sub>	53	550,000,000 TL

Ve bu VT tablosundaki her bir elemanı için bulunan birer VT oluşturalım Tablo(2.2) Q<sub>1</sub> elemanı için, Tablo(2.3) Q<sub>2</sub> elemanı için oluşturulan bulunan VT'lerdir.

Tablo(2.2) Bulanık VT<sub>f<sub>α1</sub></sub>

Id	Yaş	Gelir
Q <sub>1</sub>	55	500,000,000 TL
Q <sub>3</sub>	52	750,000,000 TL

Tablo(2.3) Bulanık VT<sub>f<sub>α2</sub></sub>

Id	Yaş	Gelir
Q <sub>2</sub>	50	450,000,000 TL
Q <sub>3</sub>	52	750,000,000 TL
Q <sub>4</sub>	53	750,000,000 TL

VT<sub>f<sub>oi1</sub></sub> bulunan veritabanındaki;

-q<sub>1</sub> nesnesinin bu VT'ye olan üyelik derecesi, o<sub>i1</sub>. Yaş'ın q<sub>1</sub>. Yaş'a olan benzerlik ölçüsü

-q<sub>3</sub> nesnesinin bu VT'ye olan üyelik derecesi, o<sub>i3</sub>. Yaş'ın q<sub>1</sub>. Yaş'a olan benzerlik ölçüsüdür.

VT<sub>f<sub>oi2</sub></sub> bulunan veritabanındaki;

-q<sub>2</sub> nesnesinin bu VT'ye olan üyelik derecesi, o<sub>i2</sub>. Yaş'ın q<sub>2</sub>. Yaş'a olan benzerlik ölçüsü

-q<sub>3</sub> nesnesinin bu VT'ye olan üyelik derecesi, o<sub>i3</sub>. Yaş'ın q<sub>2</sub>. Yaş'a olan benzerlik ölçüsü

-q<sub>4</sub> nesnesinin bu VT'ye olan üyelik derecesi, o<sub>i4</sub>. Yaş'ın q<sub>2</sub>. Yaş'a olan benzerlik ölçüsüdür.

Ve her bir bulanık  $VT_f$  için şu özetlemeyi yapmak mümkündür. "  $VT_f$  veritabanındaki insanların yaşı  $q_i$ . Yas'a benzerdir". Buradanda asıl VT için oluşturulan "VT'deki insanların yaşı birbirine benzerdir" bulanık önermesini nasıl yorumlandığı açıklanmış oldu.

BFB' de bulanık önermeler yardımıyla hazırlanan kuralların ifade ediliş biçimleri (2.18)'de açıklanmaya çalışılmıştır.

"EĞER Yaş Birbirine Benzer İSE GELİR de Birbirine Benzerdir"

"EĞER Yaş Benzer  $q_i$ . Yas İSE GELİR Benzer  $q_i$ . Gelir"

"EĞER  $q_k$ .  $X \approx q_i$ . X İSE  $q_k$ .  $Y \approx q_i$ . Y"

" $q \in VT$  EĞER  $\forall i, j \mid q_i$ .  $X \approx q_i$ . X İSE  $q_i$ .  $Y \approx q_j$ . Y" (2.18)

(2.18)'deki kuralların sorgu cümlesi (2.19)'da belirtilmiştir.

$Q_{Yas} = A \wedge Gelir = SA \wedge \Sigma Coğu (q_j \in Q_{yas} \approx A (VT) \mid q_j$ . Gelir  $\approx SA$ ) (VT)

$= \{q \in VT \mid \Sigma Coğu (\{q_j \in VT \mid \mu_k(q_i$ . Yas,  $q_i$ . Yas)  $\} \mid \mu_k(q_i$ . Gelir,  $q_i$ . Gelir))

(2.19)

BFB' de benzerlik fonksiyonu  $\mu_k(x, y) = \max(0, (t - |x - y|) / t)$  olarak tanımlanmıştır. Fonksiyondaki t değeri bulanık sorgulamaya yapan kullanıcının kayıtlar arasında benzerlik ölçüsünü hesaplamak istediğı her bir alan özelliğı için ayrı ayrı değerlendirilmeli ve bu alan için benzerlik değerinin max hangi aralıkta geçerli olabileceğı hesaplandıktan sonra t değeri atanmalıdır. Örneğın; VT'deki Gelir alan özelliğinin birbirine benzerliğı ölçülebilir düşüncesi ile t değeri ne 200,000,000 ml TL olan kayıtların benzerliğı ölçülebilir düşüncesi ile t değeri ne 200,000,000 ml TL atanabilir.

Tablo(2.1)'deki VT için "EĞER Yaşı Birbirine Benzer İSE Geliri de Birbirine Benzerdir" kuralının doğruluk derecesini hesaplayacak olursak, VT'deki her bir kayıt için bir kural oluşturulur. Ve bu kuralların doğruluk dereceleri hesaplanır.

$\tau_1 = EĞER 55 \approx q_i$ . Yas İSE 500.000.000  $\approx q_i$ . Gelir

$\tau_2 = EĞER 50 \approx q_i$ . Yas İSE 450.000.000  $\approx q_i$ . Gelir

$\tau_3 = EĞER 52 \approx q_i$ . Yas İSE 750.000.000  $\approx q_i$ . Gelir

$\tau_4 = EĞER 53 \approx q_i$ . Yas İSE 550.000.000  $\approx q_i$ . Gelir

Yaş için

$$\mu_{\approx}(q, q) = \text{Max}(|5 - q \cdot \text{Yas} - q \cdot \text{Yas}| / 5, 0) \quad (2.20)$$

Gelir için

$$\mu_{\approx}(q, q) = \text{Max}(|100000000 - q \cdot \text{Gelir} - q \cdot \text{Gelir}| / 100000000, 0) \quad (2.21)$$

Yaş ve Gelir için (2.20) ve (2.21)'deki benzerlik fonksiyonları tam anlamıyla kayıtların benzerlik ölçüleri Tablo(2.4)'deki gibi olurken, her bir kuralın doğruluk derecesi de (2.22)'de gösterilmiştir.

Tablo(2.4) BFB Örnek Tablosu

Yaş	Gelir	55≈ Q. Yas	50≈ Q. Yas	52≈ Q. Yas	53≈ Q. Yas	500 ml TL≈ Q. Gelir	450 ml TL≈ Q. Gelir	750 ml TL≈ Q. Gelir	550 ml TL≈ Q. Gelir
55	500 ml TL	1	0	0.4	0.6	1	0.5	0	0.5
50	450 ml TL	0	1	0.6	0.4	0.5	1	0	0
52	750 ml TL	0.4	0.6	1	0.8	0	0	1	0
53	550 ml TL	0.6	0.4	0.8	1	0.5	0	0	1

$$\tau_1 = \text{Min}(1, 1) + \text{Min}(0, 0.5) + \text{Min}(0.4, 0) + \text{Min}(0.6, 0.5) / (1+0+0.4+0.6) = 0.75$$

$$\tau_2 = \text{Min}(0, 0.5) + \text{Min}(1, 1) + \text{Min}(0.6, 0) + \text{Min}(0.4, 0) / (0+1+0.6+0.4) = 0.50$$

$$\tau_3 = \text{Min}(0.4, 0) + \text{Min}(0.6, 0) + \text{Min}(1, 1) + \text{Min}(0.8, 0) / (0.4+0.6+1+0.8) = 0.35$$

$$\tau_4 = \text{Min}(0.6, 0.5) + \text{Min}(0.4, 0) + \text{Min}(0.8, 0) + \text{Min}(1, 1) / (0.6+0.4+0.8+1) = 0.53$$

(2.22)

Son olarak “EĞER Yaşı Birbirine Benzer İSE Geliri de Birbirine Benzerdir” kuralının doğruluk derecesi,  $\tau = (0.75+0.50+0.35+0.53) / 4 = 0.53$  olarak bulunur. Ve (2.23)'deki gibi ifade edilir.

$$\{\text{Yaş}\} \rightarrow \{\text{Gelir}\} = 0.53 \quad (2.23)$$



## 2.6 Dereceli Bulanık Fonksiyonel Bağlılık Kural Çıkarım Mekanizması

DFB' de birbirine benzer verilerin sorgulanması için bir kural çıkarım mekanizması oluşturulurken DFB' de kademeli olarak artan yada azalan verilerin benzerliğinin sorgulandığı bir kural çıkarım mekanizması mevcuttur. DFB' de kullanılan bir bulanık önerme şu şekilde ifade edilebilir.

“VT de daha fazla Yaşa sahip insanlar, daha fazla Gelire sahiptirler ”

DFB' de “Daha fazla”, ”Daha az” gibi ifadeler yardımıyla oluşturulan bulanık önerme, VT deki verilerin monoton artma yada azalma derecelerine göre sorgulanmasına imkan sağlamaktadır. Aşağıda DFB için geçerli kural yapıları gösterilmiştir.

“EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha fazladır ”

“EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha azdır ”

“EĞER insanların Yaşları Daha az İSE Gelirleri Daha fazladır ”

“EĞER insanların Yaşları Daha az İSE Gelirleri Daha azdır ”

“EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha fazladır ” kuralı için DFB' de, VT deki her bir kaydı n Gelir alan özelliğinden Gelir alan özellikleri büyük olan kayıtların  $x \lesssim y$  ile ifade edilen benzerlik fonksiyonuna olan üyelik dereceleri hesaplanır.

$$D_1(\alpha_k, q) = \alpha_k \cdot \text{Yas} < q \cdot \text{Yas}$$

$$\dot{D}_1(\alpha_k, q) = \alpha_k \cdot \text{Yas} > q \cdot \text{Yas}$$

$$D_2(\alpha_k, q) = \alpha_k \cdot \text{Gelir} < q \cdot \text{Gelir}$$

$$\dot{D}_2(\alpha_k, q) = \alpha_k \cdot \text{Gelir} > q \cdot \text{Gelir} \text{ olarak tanımlanırsa,}$$

Her bir kayıt için DFB' de oluşturulacak kural;

$$\text{EĞER } D_1(\alpha_k, q) \vee \dot{D}_1(\alpha_k, q)$$

$$\text{İSE } (D_1(\alpha_k, q) \wedge D_2(\alpha_k, q)) \vee (\dot{D}_1(\alpha_k, q) \wedge \dot{D}_2(\alpha_k, q)) \quad (2.24)$$

(2.24)'deki kuralı daha ayrıntılı şekilde ifade edecek olursak

$$\tau_k = \sum_d (q \in \{ q \in VT \mid D_1(\alpha_k, q) \vee D_2(\alpha_k, q) \} \mid D_1(\alpha_k, q) \wedge D_2(\alpha_k, q)) \\ \vee (D_1(\alpha_k, q) \wedge D_2(\alpha_k, q))$$

$$S_{DFB} = \sum_d (\alpha_k \in VT \mid \tau_k)$$

$$= \sum_d (\alpha_k \in VT \mid \sum_d (q \in \{ q \in VT \mid D_1(\alpha_k, q) \vee D_2(\alpha_k, q) \} \\ (D_1(\alpha_k, q) \wedge D_2(\alpha_k, q)) \vee (D_1(\alpha_k, q) \wedge D_2(\alpha_k, q)))$$

$$S_{DFB} = \sum_d (\alpha_k \in VT \mid \sum_d (q \in \{ q \in VT \mid \alpha_k. \text{Yaş} < q. \text{Yaş} \vee \alpha_k. \text{Yaş} > q. \text{Yaş} \} \\ (\alpha_k. \text{Yaş} < q. \text{Yaş} \wedge \alpha_k. \text{Gelir} < \approx q. \text{Gelir}) \\ \vee (\alpha_k. \text{Yaş} > q. \text{Yaş} \wedge \alpha_k. \text{Gelir} > \approx q. \text{Gelir}))$$

DFB' de "Daha fazla" ifadesi için  $x < \approx y$  ilişkisini oluşturan benzerlik fonksiyonu (2.25)'de tanımlanmıştır.

$$\text{Max}(\text{Max}((t - |x - y|)/t, 0), x < y) \quad (2.25)$$

"Daha az" ifadesi için  $x > \approx y$  ilişkisini oluşturan benzerlik fonksiyonu (2.26)'da tanımlanmıştır.

$$\text{Max}(\text{Max}((t - |x - y|)/t, 0), x > y) \quad (2.26)$$

olarak tanımlanmıştır. Fonksiyonlarda geçen t değeri BFB' de olduğu gibi yine kullanıcının arasında dereceli artan yada azalan monotonluk özelliğini ölçmek istediği alanlar için ayrı ayrı tanımlanmalıdır. Tanımlanan fonksiyonlardan da anlaşılacağı gibi DFB kural çıkarım mekanizması boolean 1 (doğru) yada 0 (yanlış) değerlerini döndürür.

"EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha fazladır" bulan kuralı için Tablo (2.5)'deki değerleri ele alalım

“Daha fazla Gelir” ifadesi için  $x \approx y$  ilişkisini oluşturan benzerlik fonksiyonu

$$\text{Max}(\text{Min}(((100000000-|x-y|)/100000000), 0), x < y) \quad (2.28)$$

(2.27) ve (2.28)'deki gibi olsun

Şu halde VT'deki her bir kayıt için oluşacak kurallar (2.29)'daki gibi olacaktır.

$$\tau_1 = \text{EĞER } 55 < q. \text{ Yaş Veya } 55 > q. \text{ Yaş İSE}$$

$$(55 < q. \text{ Yaş Ve } 500m \text{ TL} < \approx q. \text{ Gelir})$$

Veya

$$(55 > q. \text{ Yaş Ve } 500m \text{ TL} > \approx q. \text{ Gelir})$$

$$\tau_2 = \text{EĞER } 50 < q. \text{ Yaş Veya } 50 > q. \text{ Yaş İSE}$$

$$(50 < q. \text{ Yaş Ve } 450m \text{ TL} < \approx q. \text{ Gelir})$$

Veya

$$(50 > q. \text{ Yaş Ve } 450m \text{ TL} > \approx q. \text{ Gelir})$$

$$\tau_3 = \text{EĞER } 52 < q. \text{ Yaş Veya } 52 > q. \text{ Yaş İSE}$$

$$(52 < q. \text{ Yaş Ve } 750m \text{ TL} < \approx q. \text{ Gelir})$$

Veya

$$(52 > q. \text{ Yaş Ve } 750m \text{ TL} > \approx q. \text{ Gelir})$$

(2.29)

VT'deki her bir kayıt için oluşan kuralların doğruluk dereceleri

$$\tau_1 = \text{Min}(0, 0) + \text{Min}(1, 1) + \text{Min}(1, 0) / (0+1+1) = 0.5$$

$$\tau_2 = \text{Min}(1, 1) + \text{Min}(0, 0) + \text{Min}(1, 0) / (1+0+1) = 0.5$$

$$\tau_3 = \text{Min}(1, 0) + \text{Min}(1, 1) + \text{Min}(0, 0) / (1+1+0) = 0.5$$

olacaktır.

Tablo(2.5) GFB Örnek Tablosu

Yaş	Gelir	55< Q. Yas ∨ 55> Q. Yas	50< Q. Yas ∨ 50> Q. Yas	52< Q. Yas ∨ 52> Q. Yas	((55<Q. Yas) ∧ (500mTL<≈ Q. Gelir)) ∨ ((55>Q. Yas) ∧ (500mTL>≈ Q. Gelir))	((50<Q. Yas) ∧ (450mTL<≈ Q. Gelir)) ∨ ((50>Q. Yas) ∧ (450mTL>≈ Q. Gelir))	((52<Q. Yas) ∧ (750mTL<≈ Q. Gelir)) ∨ ((52>Q. Yas) ∧ (750mTL>≈ Q. Gelir))
55	500 m TL	0	1	1	0	1	0
50	450 m TL	1	0	1	1	0	1
52	750 m TL	1	1	0	0	0	0

Ve son olarak “EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha fazladır ” bulan kuralının VT’deki doğruluk derecesi;

$$\tau = (0.5+0.5+0.5) / 3 = 0.5 \text{ olarak bulunur.}$$

DFB bulan sorgu yapısında kuralları daha açık ifade edecek olursak;

-“EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha fazladır ” kuralı için;

$$S_{DFB} = \sum_d (o_k \in VT | \sum_d (q \in \{ q \in VT | o_k. Yas < q. Yas \vee o_k. Yas > q. Yas \})$$

$$(o_k. Yas < q. Yas \wedge o_k. Gelir < \approx q. Gelir)$$

$$\vee (o_k. Yas > q. Yas \wedge o_k. Gelir > \approx q. Gelir)$$

-“EĞER insanların Yaşları Daha fazla İSE Gelirleri Daha azdır ” kuralı için;

$$S_{DFB} = \sum_d (o_k \in VT | \sum_d (q \in \{ q \in VT | o_k. Yas < q. Yas \vee o_k. Yas > q. Yas \})$$

$$(o_k. Yas < q. Yas \wedge o_k. Gelir > \approx q. Gelir)$$

$$\vee (o_k. Yas > q. Yas \wedge o_k. Gelir < \approx q. Gelir)$$

-“EĞER insanların Yaşları Daha az İSE Gelirleri Daha fazladır ” kuralı için;

$$S_{DFB} = \sum_d (\alpha_k \in VT | \sum_d (q \in \{ q \in VT | \alpha_k. \text{Yas} > q. \text{Yas} \vee \alpha_k. \text{Yas} < q. \text{Yas} \})$$

$$(\alpha_k. \text{Yas} > q. \text{Yas} \wedge \alpha_k. \text{Gelir} < \approx q. \text{Gelir} )$$

$$\vee (\alpha_k. \text{Yas} < q. \text{Yas} \wedge \alpha_k. \text{Gelir} > \approx q. \text{Gelir} )$$

-“EĞER insanların Yaşları Daha az İSE Gelirleri Daha azdır ” kuralı için;

$$S_{DFB} = \sum_d (\alpha_k \in VT | \sum_d (q \in \{ q \in VT | \alpha_k. \text{Yas} > q. \text{Yas} \vee \alpha_k. \text{Yas} < q. \text{Yas} \})$$

$$(\alpha_k. \text{Yas} > q. \text{Yas} \wedge \alpha_k. \text{Gelir} > \approx q. \text{Gelir} )$$

$$\vee (\alpha_k. \text{Yas} < q. \text{Yas} \wedge \alpha_k. \text{Gelir} < \approx q. \text{Gelir} )$$

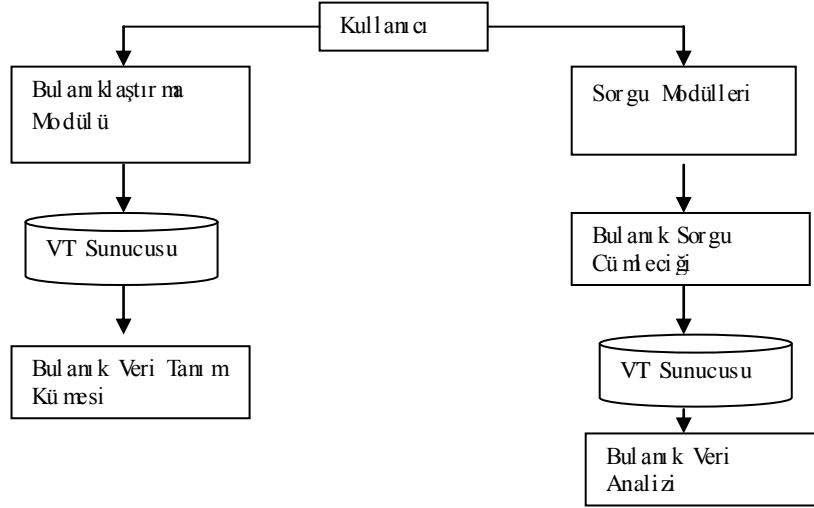
### 3. BULANIK SORGU ARACI UYGULAMASI

#### 3.1 Genel Yapı

Bulanık mantık ve bulanık küme teorilerinin kullanımdaki en önemli avantajlarından biri doğal dildeki ifadeleri kullanarak sorgulama yapma mza olanak sağlama arıdır. Yapay Zeka gibi bilgi gösteri yöntemlerinin belirsizlik, bulanıklık içeren bilgiyi temsil etmede yetersiz kaldığı noktalarda bulanık mantık ve bulanık küme teorisi önemli rol oynamaktadır. Bu tez çalışmasının temelindeki en önemli gaye, sıradan veritabanı sistemlerinde bir standart olan SQL sorgulama diline bir alternatif olarak bulanık bir sorgulama aracı geliştirebilmek ve verideki bilinmeyen gerçekleri bu yöntemle ortaya çıkartabilmektir. Bu amaçla sıradan veritabanı sistemlerinde belirsiz dilsel sorgu cümleleriyle bulanık sorgulama yapabilen bir araç geliştirilmiştir.

Sorgu aracının geliştirilmesinde kullanıcı arayüzü olarak Visual Basic 6.0, veritabanı sunucusu olarak ise SQL Server 7.0 kullanılmıştır. Sorgu aracı veritabanı sunucusuna ADO Connection kullanarak bağlanmakta ve SQL Server 7.0'ın hız ve performans açısından programa kazandırdığı avantajlar kullanılmaktadır. Sorgu aracının geliştirilmesinde Visual Basic 6.0'ın nesne yönelimli programlama yapısı, sınıf yapıları ve ActiveX gibi özellikleri kullanılmıştır.

Sorgu aracı temelde iki ana bölümü kullanır. Birinci bölüm Bulanıklaştırma Modülü kullanımı, ikinci bölüm Sorgu Modülleri kullanımıdır. Sorgu aracı çalışma prensibi Şekil(3.1)'de gösterilmiştir.



Şekil(3.1) Sorgu Aracı Çalışma Prensi bi

### 3.2 Veritabanı Yapısı

Bu tez çalışmasında Kadıköy Vatan Hastanesi Laboratuvar Bölümünde kayıtlı olan hastaların kan testi laboratuvar sonuçları kullanılmış ve hastaların kanlarındaki ki myasal yağ bileşim oranları ve yaşları arasındaki ilişkiyi sorgulamak için bir VT yaratılmıştır. VT'deki "Hasta Bilgi" tablosu, hastaların Kolesterol, Trigliserid, HDLC2 değerlerini ve yaşlarını tutan temel tablolardan biridir. Kalp ve damar hastalıkları risk faktörünü oluşturan kandaki ki myasal yağ bileşim oranları genel de belli bir yaştan sonra ortaya çıktığı için hasta popülasyonu olarak 45 yaş üstü hastalar seçilmiştir. Şekil(3.2) "Hasta Bilgi" tablosunu göstermektedir.

HastaId	Yas	Cholesterol	Trigliserid	HDLC2
1	45	226	175	48
2	49	333	143	57
3	52	242	186	50
4	64	182	122	57
5	82	212	236	52
6	49	167	211	34
7	95	214	209	46
8	62	192	50	71
9	96	246	415	32
10	87	221	161	60
11	75	310	175	51
12	74	119	142	41
13	92	165	122	42
14	63	310	290	50
15	55	265	247	40
16	88	157	136	42
17	66	239	238	31

Şekil(3.2) Hasta Bilgi Tablosu

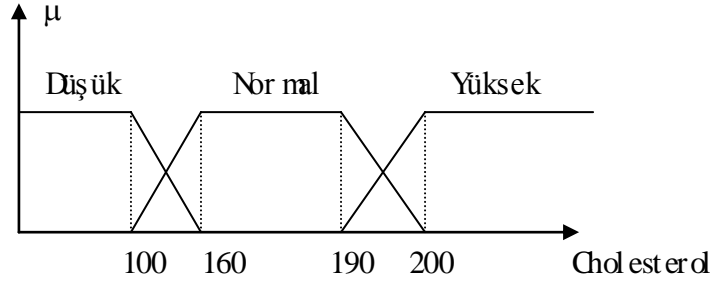
VT’deki “TableParameters” tablosu ana tablo olan “Hasta Bilgi” tablosundaki bulan klastırma süreci ne tabii tutacağı mız alanlarının dilsel deęişkenleri ve bu dilsel deęişkenlerin üyelik fonksiyonları deęerlerini tutar. Şekil(3.3)’de “TableParameters” tablosu gösterilmiştir. “TableParameters” tablosundaki TableName ve UniqueColumn alanları ilgili oldukları dilsel deęişkenin VT’deki hangi tablo ile ilişki de olduklarını belirten önemli alanlardan ikisidir. Ana tablo ile dilsel deęişkenleri ve üyelik derecelerini ilişkilendiren bu bağlantı sayesinde VT’de sorgulamayı yapmak isteyeceğimiz birden fazla ana tablo yaratabilir ve birbirlerinden bağımsız olarak herbiri için ayrı ayrı dilsel deęişken ve bu dilsel deęişkenlerin üyelik fonksiyonlarını tanımlayabiliriz “Hasta Bilgi” ve “TableParameters” tablolarından da görüldüğü gibi “Hasta Bilgi” tablosunun bir alanı olan Cholesterol için, “TableParameters” tablosunda Düşük, Normal, Yüksek dilsel deęişkenleri tanımlanmıştır. Bu dilsel deęişkenlerin üyelik fonksiyonları FunctionName alanı altında belirtilmiş ve Trapezoidal fonksiyon olarak tanımlanmıştır.

AttributeName	FuzzyLabel	X1Min	X1Max	X2Max	X2Min	FunctionName	TableName	UniqueColumn
Cholesterol	Dusuk	0	0	100	160	Azalan	HastaBilgi	HastaId
Cholesterol	Normal	100	160	190	200	Yamuk	HastaBilgi	HastaId
Cholesterol	Yukse	190	200	0	0	Artan	HastaBilgi	HastaId
Triglisericid	Dusuk	0	0	140	150	Azalan	HastaBilgi	HastaId
Triglisericid	Normal	140	150	190	200	Yamuk	HastaBilgi	HastaId
Triglisericid	Yukse	190	200	0	0	Artan	HastaBilgi	HastaId
HDLC2	Dusuk	0	0	30	45	Azalan	HastaBilgi	HastaId
HDLC2	Normal	30	45	50	55	Yamuk	HastaBilgi	HastaId
HDLC2	Yukse	50	55	0	0	Artan	HastaBilgi	HastaId
Yas	Orta	45	50	55	60	Yamuk	HastaBilgi	HastaId
Yas	Yasli	55	60	0	0	Artan	HastaBilgi	HastaId
*								

Şekil(3.3) TableParameters Tablosu

Şekil(3.4) ise “TableParameters” tablosunda tanımlı Cholesterol deęeri için geçerli üyelik fonksiyonlarını göstermektedir.





Şekil(3.4) Cholesterol Değeri için tam üyelik fonksiyonları

VT'deki "Powers" tablosu Çok, Yaklaşık gibi dilsel pekiştiricilerin tam üyeli olduğu, "Quantifiers" tablosu ise dilsel niceleyicilerin ve üyelik fonksiyonlarının tam üyeli olduğu tablolarıdır. Şekil(3.5) "Power" tablosunu, Şekil(3.6) "Quantifiers" tablosunu göstermektedir.

PowerName	PowerValue
Cok	2
Yaklasik	0.5

Şekil(3.5) Powers Tablosu

QuantifierName	X1Min	X1Max	X2Max	X2Min	FunctionName
Cogu	0.6	0.75	0	0	Artan
BirKisim	0.4	0.5	0.6	0.75	Yamuk
Azinlik	0	0	0.4	0.5	Azalan

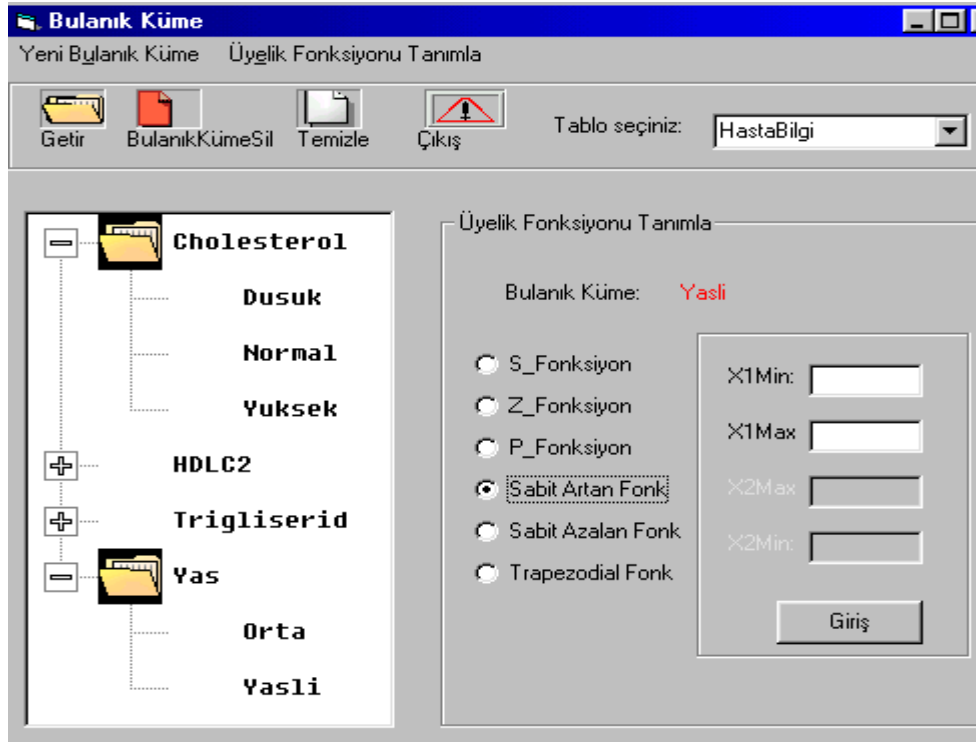
Şekil(3.6) Quantifiers Tablosu

VT'de tutulan geçici fakat her farklı veritabanı uygulaması için oluşturulması gereken "IfThenTable", "ListTable", "TempTable" tabloları ise bulanık sorgulama sürecinde oluşan ara sonuçları tutmak için kullanılan geçici tablolarıdır. Bu tez çalışmasında geliştirilen bulanık sorgulama aracı ana tablodan bağımsız olarak kullanıcının "TableParameters", "Powers", "Quantifiers" tablolarında tanımladığı değerler sayesinde ana tablo ile ilişki kurmakta ve bu sayede çeşitli veritabanları ile uyumlu olarak çalışabilmektedir.

### 3.3 Bulanıklaştırma Modülü

Sorgulama aracında kullanıcıya sunulmak üzere dört ana menü geliştirilmiştir. Bunlardan ilki olan Değişken Tanımla menüsü Bulanıklaştırma Modülünü içeren diğer Bulanık Sorgu, Fonksiyonel Sorgu, Dereceli Sorgu menüleri ise Sorgu Modüllerini içeren menülerdir.

Bulanık sorgulama aracının temel modüllerinden biri olan Bulanıklaştırma Modülü dilsel değişkenlerin ve dilsel değişkenlerin üyelik fonksiyonlarının kullanıcı tarafından tanımlanmasına olanak sağlayan bir modüldür. Şekil(3.7)'de kullanıcı arayüzü gösterilmiştir. Hasta Bilgi tablosunun Kolesterol, Trigliserid, HDLC2 ve Yaş alanları için tanımlanan dilsel değişkenleri Tree View şeklinde gösterilmiş ve kullanıcıya tanımlı üyelik derecelerini değiştirebilme yetkisi verilmiştir. Arayüzdeki Yeni Bulanık Küme menüsü ile ise Kolesterol, Trigliserid, HDLC2 ve Yaş alanları için yeni dilsel değişken tanımlamaya olanak sağlanmıştır. Üyelik fonksiyonu tanımlanırken seçilen fonksiyon çeşidine göre girilmesi gereken alanlar X1Min, X1Max, X2Max ve X2Min alanları kendiliğinden aktif hale gelerek kullanıcıya yol göstermektedir.



Şekil(3.7) Bulanık Küme Tanımlama Menüsü

### 3.4 Sorgu Mdlleri

#### 3.4.1 Bulanık Alt Kme İle Bileşik Sorgulama

Bulanık Sorgu mns ile ekrana gelen kullanıcı arayz “IF ( Koşul ) Then ( Şart)” kural yapısına gre bir bulanık kural oluřturmakta ve kuralın dođruluk derecesini hesaplamaktadır. Kural yapısının koşul kısmı bileşik bulanık önermeler ierebildiđi gibi şart kısmında aynı şekilde bileşik bulanık önermeler ierebilmektedir. Bileşik yada basit önermelerle oluřturulan kuralın Koşul kısmı bir bulanık VT oluřturur. Sorgu iřlemesi artık bu bulanık VT stnden kuralın şart kısmını arařtıracaktır. Sorgu iřlemesi bulanık alt kme ile bileşik sorgulamaya yaparken, kuralın bileşik koşul ve şart kısımlarını basit alt cmlelere blnerek analiz edilir. Herbir kayıt iin oluřturulan her bir basit alt cmle VT’de geici tablolarda korunarak, hem kuralın koşul hem de şart kısımlarını iin Ve, Veya operatrleri ile iřleme sokulur ve bileşik cmle yapısının eřleşme derecesi bulunmuř olur.

Şekil(3.8) ‘deki Bulanık Sorgulama Mns ile “EĐER Kolesterol Deđil Yaklařık Dřk Ve Trigliserid Normal İSE Yař Yaklařık Yařlıdır” bulanık kuralı iin sorgulanmuř ve kuralın dođruluk derecesi 0.66 olarak bulunmuřtur. Gridde her bir kayıt iin kuralın koşul ve şart kısmını sađlama dereceleri ve kuralın o kayıt iin geerli dođruluk derecesi gsterilmiřtir.

Yas	Cholesterol	Trigliserid	HDLc2	IfPart	ThenPart	ConclusionPart	ID
84	205	74	89	0	1	0	55
85	284	192	34	0.8	1	0.8	56
95	166	48	64	0	1	0	57
58	233	170	48	1	0.77	0.77	58
68	229	164	51	1	1	1	59
61	150	186	39	0.59	1	0.59	60
54	237	90	103	0	0	0	61

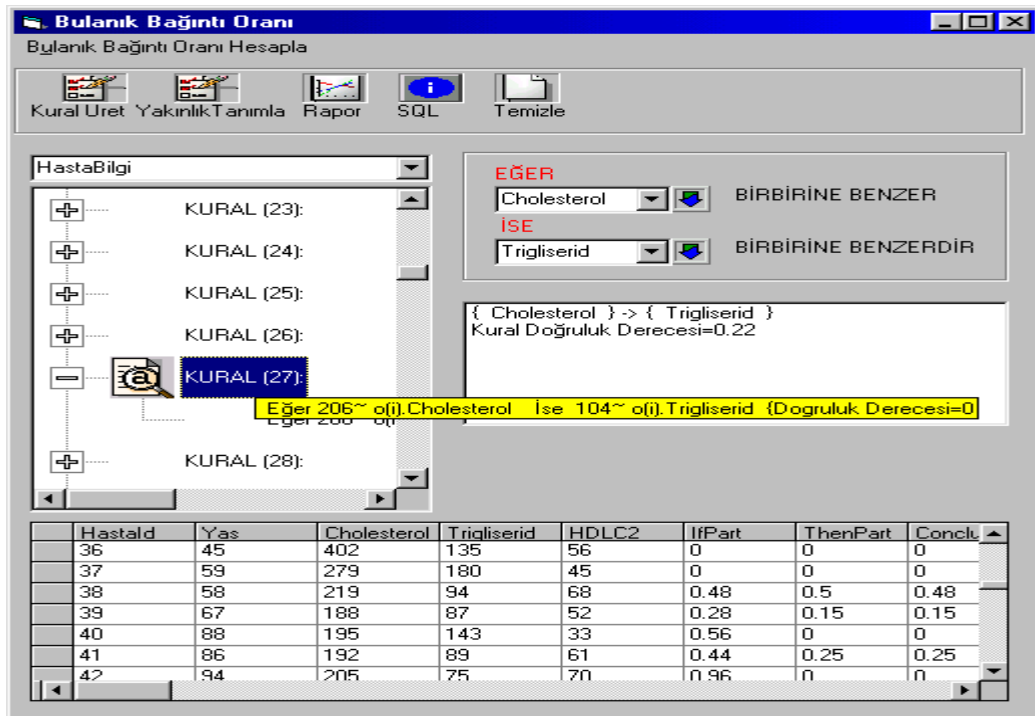
Şekil(3.8) Bulanık Sorgulama Mns

Menüde herhangi bir dilsel nicelleyici seçilmediği için kuralın bir dilsel nicelleyiciye olan doğruluk derecesi hesaplanmamıştır. Aynı kural için Bir Kısım dilsel nicelleyicisini seçecek olsaydı “EĞER VT’deki Bir Kısım insanların Cholesterol’ü Değil Yaklaşık Düşük Ve Trigiserid’i Normal İSE Yaş’ı Yaklaşık Yaşlıdır” bulan kuralını oluşturmuş olacaktık ve kuralın doğruluk derecesini 0.63 olarak bulacaktık Şekil (3.9) Bir Kısım dilsel nicelleyicisini seçilmiş şeklini göstermektedir.



Şekil(3.9) Bir Kısım Dİsel Ncelleyicisi için Bulanık Sor gul a ma Menüü

### 3.4.2 Fonksiyonel Sorgu

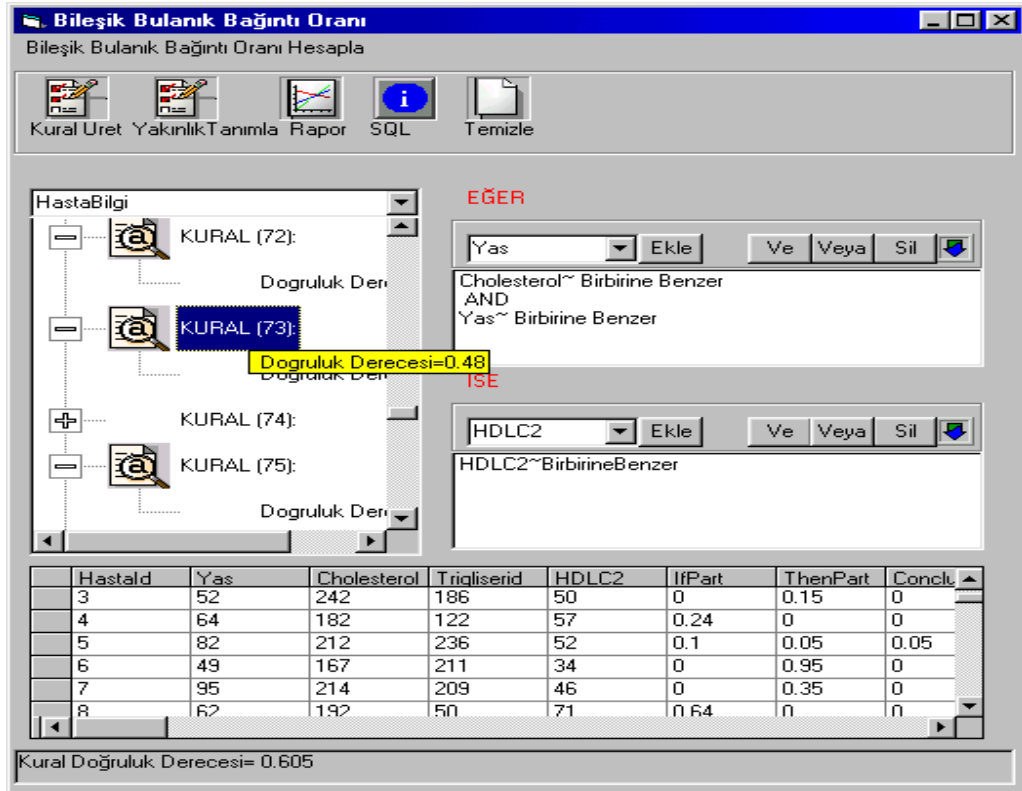


Şekil(3.10) Fonksiyonel Sor gul a ma Menüü

Şekil(3.10), Fonksiyonel Sorgulama menüsünden seçeceğimiz basit bulanık fonksiyonel bağıllık oranı sorgulama ekranını göstermektedir. “EĞER VT’de Kolesterol Birbirine Benzer İSE Tri gliserid de Birbirine Benzerdir” bulanık kuralı için yapılan sorgu analizi sonunda bulanık kuralın doğruluk derecesi 0.22 olarak bulunmuştur.

Ve Tree View’da VT’deki Hasta Bilgi tablosundaki herbir kayıt için oluşturulmuş kurallar ve doğruluk dereceleri gösterilmiştir. Gridde ise, Tree View’dan seçilecek olan kural için VT’deki tüm kayıtların alacağı eşleşme değerleri gösterilmektedir. Şekil(3.10)’da Hasta Bilgi tablosundaki 27. kayıt için “EĞER VT’de Kolesterol Değerleri 206 ‘ya Benzer İSE Tri gliserid değerleri de 104’e Benzerdir” bulanık kuralı oluşturulmuş ve doğruluk derecesi 0.16 olarak hesaplanmıştır. Sorgu işlemini çalıştırmadan önce Yakınlık Tanımla butonu ile Bulanık Fonksiyonel Bağıllık’da kullanılan yakınlık fonksiyonu için Kolesterol ve Tri gliserid değerlerinin benzerlik aralığını tanımlamamız gerekmektedir. Sorguda Kolesterol benzerlik aralığı 25 mg/dl, Tri gliserid benzerlik aralığı 20 mg/dl olarak tanımlanmıştır. Herbir kayıt için tek tek ilgili kuralları bulunarak hesaplanan doğruluk derecelerinin toplam kayıt sayısına oranı ise bize “EĞER VT’de Kolesterol Birbirine Benzer İSE Tri gliserid de Birbirine Benzerdir” bulanık kuralının tüm VT’de geçerli doğruluk derecesini verecektir.

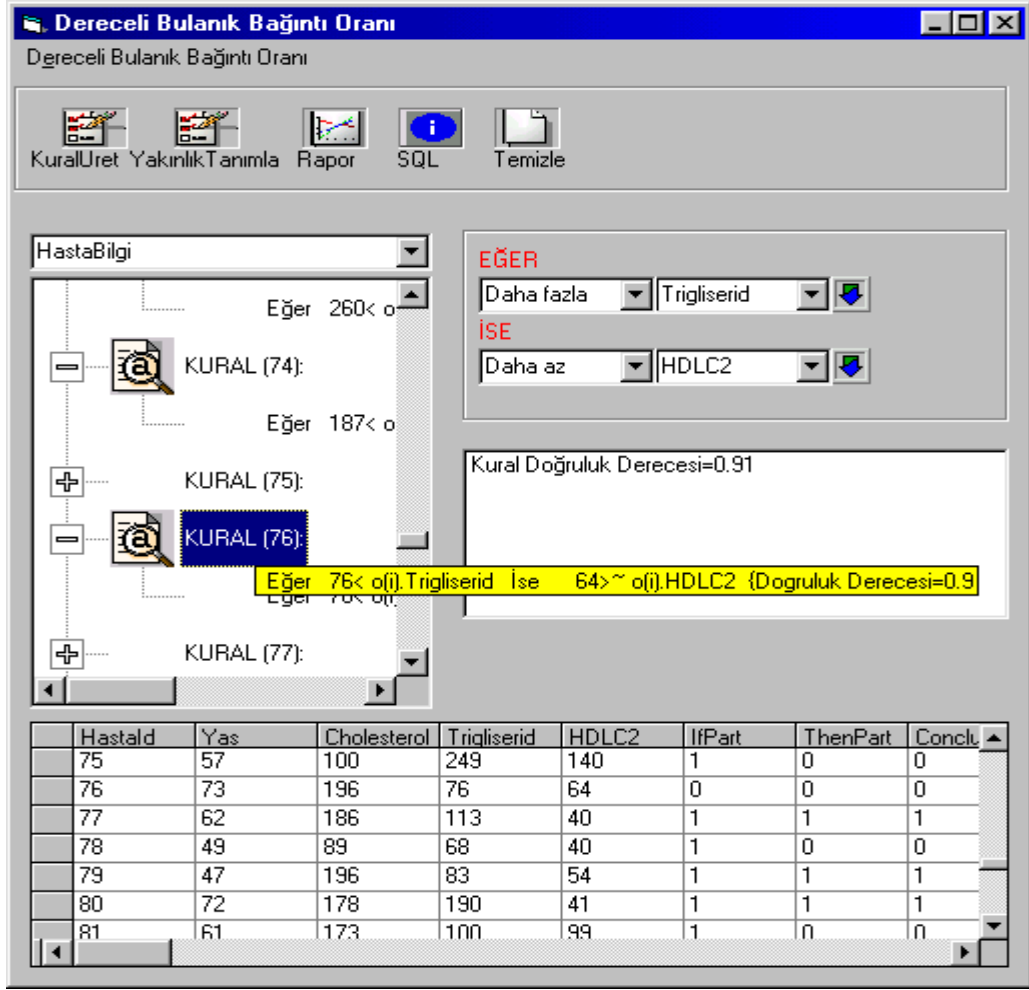
Fonksiyonel Sorgulama menüsünden seçeceğimiz Bileşik Bulanık Fonksiyonel Bağıllık oranı sorgulama ekranı ise Şekil(3.11) ‘de gösterilmiştir. “EĞER VT’de Kolesterol Birbirine Benzer VE Yaş Birbirine Benzer İSE HDLC2 de Birbirine Benzerdir” bulanık kuralı için oluşturulmuş sorgu sonucu kuralın tüm VT’de geçerli doğruluk derecesi 0.60 olarak bulunmuştur. Sorgu işlemini çalıştırmadan önce Yakınlık Tanımla butonu ile Bulanık Fonksiyonel Bağıllık’da kullanılan yakınlık fonksiyonu için Kolesterol, Yaş ve HDLC2 değerlerinin benzerlik aralığını tanımlamamız gerekmektedir. Şekil(3.11) ‘deki sorgu için Kolesterol benzerlik aralığı 25 mg/dl, Yaş benzerlik aralığı 20, HDLC2 benzerlik aralığı 20 mg/dl olarak tanımlanmıştır.



Şekil(3.11) Bileşik Fonksiyonel Sorgulama Menüü

### 3.4.3 Dereceli Sorgu

Şekil(3.12), Dereceli Sorgu menüsünden seçeceğimiz Dereceli bulanık fonksiyonel bağılılık oranı sorgulama ekranını göstermektedir. “EĞER VT’de hastalar daha fazla Trigliserid değerlerine sahip İSE daha az HDLC2 değerlerine sahiptir” bulanık kuralı sorgulanmış ve kuralın doğruluk derecesi olarak 0.91 bulunmuştur. YakınlıkTanımla butonu ile HDLC2 değerlerinin benzerlik aralığı 20 mg/dl olarak tanımlanarak sorgu işlemcisinin benzerlik fonksiyonunda bu değeri dikkate alması sağlanmıştır. Hasta Bilgi tablosundaki her bir kayıt için Trigliserid ve HDLC2 değerleri dikkate alınarak bulanık kural oluşturulmuş ve bu kuralların doğruluk dereceleri hesaplanarak Tree View’da gösterilmiştir. Şekil(3.12) ‘de Hasta Bilgi tablosundaki 76. kayıt için Trigliserid ve HDLC2 değerleri dikkate alınarak “EĞER VT’deki hastaların Trigliserid değerleri 76 ‘dan daha fazla İSE HDLC2 değerleri de 64 ‘den daha azdır ” bulanık kuralı üretilmiş ve kuralın doğruluk derecesi de 0.95 olarak bulunmuştur. Sorgu işlemcisi her bir kayıt için bu tür bulanık kuralları üreterek tüm VT’de geçerli olacak “EĞER VT’de hastalar daha fazla Trigliserid değerlerine sahip İSE daha az HDLC2 değerlerine sahiptir” bulanık kuralının doğruluk derecesini hesaplayacaktır.



Şekil(3.12) Dereceli Sorgulama Menüü

### 3.5 Hasta Bilgi Veritabanı Üzerinde Sorgulama

Kadıköy Vatan Hastahanesi Laboratuvarından elde edilen hastaların kan değerleri ne ilişkin test sonuçlarının tutulduğu verilerle oluşturulan veritabanında, hastaların Kolesterol, Triglisericid ve HDLC2 değerleri mg/dl birimi bazında verilmiştir. Kolesterol, Triglisericid ve HDLC2 değerleri, kandaki ki myasal yağ bileşimleri olup, sağlıklı insanlarda Kolesterol ve HDLC2 değerlerini ters orantılı bir artış ilişkisi içinde olması beklenmektedir. Kolesterol ve Triglisericid değerlerini yüksek olması, HDLC2 değerini ise düşük olması insan sağlığı açısından kalp ve damar hastalıkları risk faktörünü artırırken, Kolesterol ve Triglisericid değerlerinin düşük olması, HDLC2 değerini ise yüksek olması bu riski azaltmaktadır. Bu bilgiler ışığında, Hasta Bilgi VT de tutulan veriler üzerinde, hastaların yaşı, Kolesterol, Triglisericid ve HDLC2 değerleri arasındaki ilişkiyi inceleyecek olan bulanık sorgulama yapılmış

ve bu deęerler arasında nasıl bir ilişki olduęu kurallar oluşturulmak suretiyle incelenmeye çalışılmıştır. Hasta Bilgi VT'sinden türetilmiş olan kurallar ve bu kuralların doğruluk dereceleri bulanık sorgulama, bulanık fonksiyonel baęlılık dereceli bulanık fonksiyonel baęlılık kural çıkarım mekanizmaları ile oluşturulmuştur.

### **3.5.1 Hasta Bilgi VT Üzerinde Bulanık Sorgulama Sonuçları**

-“EĞER Kolesterol Düşük Veya HDLC2 Normal İSE Yaş Orta'dır”

Kural Doğruluk Derecesi:0.27

-“ EĞER Kolesterol Normal Ve Trigliserid Deęil Yaklaşık Normal İSE HDLC2 Çok Yüksek'tir”

Kural Doğruluk Derecesi:0.5

-“ EĞER Trigliserid Deęil Normal İSE Yaş Çok Orta'dır”

Kural Doğruluk Derecesi:0.18

-“ EĞER Kolesterol Çok Yüksek Veya Trigliserid Düşük İSE HDLC2 Yüksek'tir”

Kural Doğruluk Derecesi:0.48

-“EĞER Yaş Yaşlı Veya HDLC2 Normal İSE Kolesterol Deęil Yüksek Ve Trigliserid Düşük'tür”

Kural Doğruluk Derecesi:0.24

-“EĞER Kolesterol Çok Normal Ve Trigliserid Düşük İSE HDLC2 Yaklaşık Yüksek'tir”

Kural Doğruluk Derecesi:0.6

-“ EĞER Yaş Orta Veya HDLC2 Düşük İSE Kolesterol Yüksek'tir”

Kural Doğruluk Derecesi:0.63



-“ EĞER Yaş Değil Yaşlı İSE Cholesterol Yüksek'tir”

Kural Doğruluk Derecesi: 0.72

-“ EĞER Çoğu hastalar için Yaş Değil Yaşlı İSE Cholesterol Yüksek'tir”

Kural Doğruluk Derecesi: 0.77

-“ EĞER Cholesterol Değil Çok Normal Veya Trigliserid Yaklaşık Düşük İSE HDLC2 Çok Yüksek'tir”

Kural Doğruluk Derecesi: 0.46

-“ EĞER Azınlıktaki hastalar için Cholesterol Değil Çok Normal Veya Trigliserid Yaklaşık Düşük İSE HDLC2 Çok Yüksek'tir”

Kural Doğruluk Derecesi: 0.41

-“ EĞER Bir kısmı hastalar için HDLC2 Düşük Veya Yaş Orta İSE Trigliserid Normal Veya Cholesterol Değil Yüksek'tir”

Kural Doğruluk Derecesi: 1

-“ EĞER HDLC2 Düşük İSE Cholesterol Yüksek'tir”

Kural Doğruluk Derecesi: 0.55

-“ EĞER Cholesterol Değil Düşük Ve Trigliserid Yaklaşık Düşük İSE HDLC2 Değil Çok Normal Veya Yaş Çok Orta'dır”

Kural Doğruluk Derecesi: 0.87

-“ EĞER Trigliserid Değil Düşük Veya HDLC2 Normal İSE Yaş Değil Orta Ve Cholesterol Düşük'tür”

Kural Doğruluk Derecesi: 0.04

-“ EĞER Azınlıktaki hastalar için Trigliserid Değil Düşük Veya HDLC2 Normal İSE Yaş Değil Orta Ve Cholesterol Düşük'tür”

Kural Doğruluk Derecesi: 1

### 3.5.2 Hast a Bl gi VT Üzeri nde Basit ve Bleşik BFB Sorgulama Sonuçları

BFB' de kural doğruluk derecesi hesaplanırken tam nıanması gereken benzerlik aralığı her bir kural için ayrı ayrı tam nıanmış ve kural çıkarım mekanizması nda bu değerler dikkate alınmıştır.

-“EĞER Kolesterol değerleri Birbirine Benzer İSE HDLC2 Değerleri de Birbirine Benzer’ dir”

Kolesterol Değeri için Benzerlik Aralığı =5 mg/dl

HDLC2 Değeri için Benzerlik Aralığı =5 mg/dl

Kural Doğruluk Derecesi: 0.40

-“EĞER Kolesterol değerleri Birbirine Benzer İSE Tri gliserid Değerleri de Birbirine Benzer’ dir”

Kolesterol Değeri için Benzerlik Aralığı =5 mg/dl

Tri gliserid Değeri için Benzerlik Aralığı =5 mg/dl

Kural Doğruluk Derecesi: 0.35

-“EĞER Kolesterol değerleri Birbirine Benzer İSE Yaş Değerleri de Birbirine Benzer’ dir”

Kolesterol Değeri için Benzerlik Aralığı =5 mg/dl

Yaş Değeri için Benzerlik Aralığı =10

Kural Doğruluk Derecesi: 0.43

-“EĞER Yaş değerleri Birbirine Benzer İSE Kolesterol Değerleri de Birbirine Benzer’ dir”

Yaş Değeri için Benzerlik Aralığı =10

Kolesterol Değeri için Benzerlik Aralığı =10 mg/dl

Kural Doğruluk Derecesi: 0.12

-“EĞER Yaş değerleri Birbirine Benzer İSE Tri gliserid Değerleri de Birbirine Benzer’ dir”

Yaş Değeri için Benzerlik Aralığı =5

Tri gliserid Değeri için Benzerlik Aralığı =20 mg/dl

Kural Doğruluk Derecesi: 0.21

-“EĞER Yaş değerleri Birbirine Benzer İSE HDLC2 Değerleri de Birbirine Benzer’dir”

Yaş Değeri için Benzerlik Aralığı =7

HDLC2 Değeri için Benzerlik Aralığı =15 mg/dl

Kural Doğruluk Derecesi: 0.41

-“EĞER Tri gliserid değerleri Birbirine Benzer İSE Yaş Değerleri de Birbirine Benzer’dir”

Tri gliserid Değeri için Benzerlik Aralığı =15 mg/dl

Yaş Değeri için Benzerlik Aralığı =5

Kural Doğruluk Derecesi: 0.30

-“EĞER Tri gliserid değerleri Birbirine Benzer İSE Kolesterol Değerleri de Birbirine Benzer’dir”

Tri gliserid Değeri için Benzerlik Aralığı =20 mg/dl

Kolesterol Değeri için Benzerlik Aralığı =20 mg/dl

Kural Doğruluk Derecesi: 0.32

-“EĞER Tri gliserid değerleri Birbirine Benzer İSE HDLC2 Değerleri de Birbirine Benzer’dir”

Tri gliserid Değeri için Benzerlik Aralığı =25 mg/dl

HDLC2 Değeri için Benzerlik Aralığı =25 mg/dl

Kural Doğruluk Derecesi: 0.66

-“EĞER HDLC2 değerleri Birbirine Benzer İSE Kolesterol Değerleri de Birbirine Benzer’dir”

HDLC2 Değeri için Benzerlik Aralığı =5 mg/dl

Kolesterol Değeri için Benzerlik Aralığı =10 mg/dl

Kural Doğruluk Derecesi: 0.21

-“EĞER HDLC2 değerleri Birbirine Benzer İSE Tri gliserid Değerleri de Birbirine Benzer’ dir”

HDLC2 Değeri için Benzerlik Aralığı =10 mg/ dl

Tri gliserid Değeri için Benzerlik Aralığı =15 mg/ dl

Kural Doğruluk Derecesi: 0.18

-“EĞER HDLC2 değerleri Birbirine Benzer İSE Yaş Değerleri de Birbirine Benzer’ dir”

HDLC2 Değeri için Benzerlik Aralığı =5 mg/ dl

Yaş Değeri için Benzerlik Aralığı =3 mg/ dl

Kural Doğruluk Derecesi: 0.22

-“EĞER Cholesterol değerleri Birbirine Benzer Ve HDLC2 değerleri de Birbirine Benzer İSE Tri gliserid Değerleri Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =25 mg/ dl

HDLC2 Değeri için Benzerlik Aralığı =20 mg/ dl

Tri gliserid Değeri için Benzerlik Aralığı =25 mg/ dl

Kural Doğruluk Derecesi: 0.39

-“EĞER Cholesterol değerleri Birbirine Benzer Ve Yaş Değerleri Birbirine Benzer İSE Tri gliserid Değerleri de Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =25 mg/ dl

Yaş Değeri için Benzerlik Aralığı =5

Tri gliserid Değeri için Benzerlik Aralığı =25 mg/ dl

Kural Doğruluk Derecesi: 0.56

-“EĞER HDLC2 değerleri Birbirine Benzer Ve Tri gliserid değerleri Birbirine Benzer İSE Cholesterol Değerleri Birbirine Benzer Ve Yaş Değerleri de Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =20 mg/ dl

HDLC2 Değeri için Benzerlik Aralığı =20 mg/ dl

Yaş Değeri için Benzerlik Aralığı =7

Tri gliserid Değeri için Benzerlik Aralığı =10 mg/ dl

Kural Doğruluk Derecesi: 0.45

-“EĞER Yaş değerleri Birbirine Benzer Ve HDLC2 değerleri Birbirine Benzer İSE Cholesterol Değerleri Birbirine Benzer Veya Tri gliserid Değerleri Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =10 mg/ dl

HDLC2 Değeri için Benzerlik Aralığı =15 mg/ dl

Yaş Değeri için Benzerlik Aralığı =5

Tri gliserid Değeri için Benzerlik Aralığı =20 mg/ dl

Kural Doğruluk Derecesi: 0.46

-“EĞER Yaş değerleri Birbirine Benzer Ve Cholesterol değerleri Birbirine Benzer İSE HDLC2 Değerleri Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =15 mg/ dl

HDLC2 Değeri için Benzerlik Aralığı =15 mg/ dl

Yaş Değeri için Benzerlik Aralığı =10

Kural Doğruluk Derecesi: 0.65

-“EĞER Cholesterol değerleri Birbirine Benzer Ve HDLC2 değerleri Birbirine Benzer Ve Tri gliserid Değerleri Birbirine Benzer İSE Yaş Değerleri Birbirine Benzer’ dir”

Cholesterol Değeri için Benzerlik Aralığı =20 mg/ dl

HDLC2 Değeri için Benzerlik Aralığı =20 mg/ dl

Tri gliserid Değeri için Benzerlik Aralığı =20 mg/ dl

Yaş Değeri için Benzerlik Aralığı =5

Kural Doğruluk Derecesi: 0.65

-“EĞER Cholesterol değerleri Birbirine Benzer Veya HDLC2 değerleri Birbirine Benzer Ve Tri gliserid Değerleri Birbirine Benzer İSE Yaş Değerleri Birbirine Benzer’ dir”

Cholesterol Deęeri iin Benzerlik Aralıęı =5 mg/ dl

HDLC2 Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Tri gliserid Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Yaş Deęeri iin Benzerlik Aralıęı =5

Kural Doğruluk Derecesi: 0.84

-“EĞER Cholesterol deęerleri Birbirine Benzer İSE HDLC2 deęerleri Birbirine Benzer Ve Tri gliserid Deęerleri Birbirine Benzer Ve Yaş Deęerleri Birbirine Benzer’ dir”

Cholesterol Deęeri iin Benzerlik Aralıęı =5 mg/ dl

HDLC2 Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Tri gliserid Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Yaş Deęeri iin Benzerlik Aralıęı =5

Kural Doğruluk Derecesi: 0.33

-“EĞER HDLC2 deęerleri Birbirine Benzer İSE Cholesterol deęerleri Birbirine Benzer Veya Tri gliserid Deęerleri Birbirine Benzer’ dir”

Cholesterol Deęeri iin Benzerlik Aralıęı =10 mg/ dl

HDLC2 Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Tri gliserid Deęeri iin Benzerlik Aralıęı =7 mg/ dl

Kural Doğruluk Derecesi: 0.23

-“EĞER Yaş deęerleri Birbirine Benzer Veya Cholesterol deęerleri Birbirine Benzer İSE Tri gliserid Deęerleri Birbirine Benzer Veya HDLC2 Deęerleri Birbirine Benzer’ dir”

Cholesterol Deęeri iin Benzerlik Aralıęı =10 mg/ dl

HDLC2 Deęeri iin Benzerlik Aralıęı =5 mg/ dl

Tri gliserid Deęeri iin Benzerlik Aralıęı =8 mg/ dl

Yaş Deęeri iin Benzerlik Aralıęı =3

Kural Doğruluk Derecesi: 0.24

-“EĞER Kolesterol deęerleri Birbirine Benzer İSE HDLC2 deęerleri Birbirine Benzer Veya Tri gliserid Deęerleri Birbirine Benzer Veya Yaş Deęerleri Birbirine Benzer’ dir”

Kolesterol Deęeri için Benzerlik Aralıęı =5 mg/ dl

HDLC2 Deęeri için Benzerlik Aralıęı =10 mg/ dl

Tri gliserid Deęeri için Benzerlik Aralıęı =5 mg/ dl

Yaş Deęeri için Benzerlik Aralıęı =5

Kural Doğruluk Derecesi: 0.54

### **3.5.3 Hast a Bilgi VT Üzeri nde DFB Sorgul a ma Sonuęları**

DFB’ de kural doğruluk derecesi hesaplanırken tanımlanması gereken benzerlik aralıęı her bir kural için ayrı ayrı tanımlanmış ve kural çıkarım mekanizmasında bu deęerler dikkate alınmıştır.

-“EĞER hastalar Daha Fazla Kolesterol Deęerine sahip İSE Daha Az HDLC2 Deęerine Sahiptir”

HDLC2 Deęeri için Benzerlik Aralıęı =10 mg/ dl

Kural Doğruluk Derecesi: 0.61

-“EĞER hastalar Daha Fazla Kolesterol Deęerine sahip İSE Daha Az Tri gliserid Deęerine Sahiptir”

Tri gliserid Deęeri için Benzerlik Aralıęı =5 mg/ dl

Kural Doğruluk Derecesi: 0.42

-“EĞER hastalar Daha Fazla Kolesterol Deęerine sahip İSE Daha Az Yaş Deęerine Sahiptir”

Yaş Deęeri için Benzerlik Aralıęı =5

Kural Doğruluk Derecesi: 0.60

-“EĞER hastalar Daha Az Kolesterol Deęerine sahip İSE Daha Az Tri gliserid Deęerine Sahiptir”

Tri gliserid Deęeri için Benzerlik Aralıęı =5 mg/ dl

Kural Doğruluk Derecesi: 0.61

-“EĞER hastalar Daha Az Kolesterol Değeri ne sahip İSE Daha Az Yaş Değeri ne Sahiptir”

Yaş Değeri için Benzerlik Aralığı =10

Kural Doğruluk Derecesi: 0.62

-“EĞER hastalar Daha Fazla HDLC2 Değeri ne sahip İSE Daha Fazla Kolesterol Değeri ne Sahiptir”

Kolesterol Değeri için Benzerlik Aralığı =10 mg/dl

Kural Doğruluk Derecesi: 0.61

-“EĞER hastalar Daha Fazla HDLC2 Değeri ne sahip İSE Daha Fazla Trigliserid Değeri ne Sahiptir”

Trigliserid Değeri için Benzerlik Aralığı =15 mg/dl

Kural Doğruluk Derecesi: 0.4

-“EĞER hastalar Daha Fazla HDLC2 Değeri ne sahip İSE Daha Fazla Yaş Değeri ne Sahiptir”

Yaş Değeri için Benzerlik Aralığı =3 mg/dl

Kural Doğruluk Derecesi: 0.59

-“EĞER hastalar Daha Fazla Trigliserid Değerine sahip İSE Daha Az Kolesterol Değeri ne Sahiptir”

Kolesterol Değeri için Benzerlik Aralığı =20 mg/dl

Kural Doğruluk Derecesi: 0.52

-“EĞER hastalar Daha Fazla Trigliserid Değerine sahip İSE Daha Az HDLC2 Değeri ne Sahiptir”

HDLC2 Değeri için Benzerlik Aralığı =5 mg/dl

Kural Doğruluk Derecesi: 0.74

-“EĞER hastalar Daha Fazla Trigliserid Değeri ne sahip İSE Daha Az Yaş Değeri ne Sahiptir”

Yaş Değeri için Benzerlik Aralığı =5

Kural Doğruluk Derecesi: 0.61



-“EĐER hastalar Daha Az Yaş Deęeri ne sahip İSE Daha Az Cholesterol Deęeri ne Sahiptir”

Cholesterol Deęeri için Benzerlik Aralıęı =15 mg/dl

Kural Doğruluk Derecesi: 0.56

-“EĐER hastalar Daha Az Yaş Deęeri ne sahip İSE Daha Az HDLC2 Deęeri ne Sahiptir”

HDLC2 Deęeri için Benzerlik Aralıęı =10 mg/dl

Kural Doğruluk Derecesi: 0.72

-“EĐER hastalar Daha Az Yaş Deęeri ne sahip İSE Daha Az Tri gliserid Deęeri ne Sahiptir”

Tri gliserid Deęeri için Benzerlik Aralıęı =5 mg/dl

Kural Doğruluk Derecesi: 0.48

#### 4. SONUÇLAR VE ÖNERİLER

Bu yüksek lisans tez çalışmasının konusu olan, bir bulanık sorgulama aracı geliştirme sürecinde, SummarySQL sorgulama tekniği baz olarak alınmış ve kullanıcıların doğal dildeki ne çok yakın ifadelerle veritabanları üstünde sorgulama yapmalarına olanak sağlanmıştır. Sorgulama aracı kullanıcının ekran girdisini bir kural olarak alıp, bu kuralın işlenen veritabanı üzerindeki doğruluk derecesini hesaplayarak kuralın geçerliliği hakkında bilgi vermektedir. Uygulama sürecinde doğal dildeki ifadeleri bulanık kümeler olarak ele alıp, bilgisayarda temsil ve işlenmesi ile ilgili bir yaklaşım geliştirilmiş ve bu çalışmayı temel kabul ederek daha ileri seviyeye götürecek olan çalışmalara zemin hazırlanmaya çalışılmıştır.

Günümüzde her kullanıcının bir uzman bilgisine sahip olmadığı düşünülürse, bu çalışma kullanıcılara esnek sorgulama yapma imkanı verirken fakat dilsel ifadelerin yorumlanacağı üyelik fonksiyonları için bir uzman görüşü ve bilgisi beklenmektedir.

Çalışmanın temel sorgularından olan bulanık fonksiyonel bağılıklık oranı ve dereceli fonksiyonel bağılıklık oranlarının hesaplanması sırasında, veritabanı büyümesini performans etkileyecek bir durum oluşturduğu gözlemlenmiş ve veritabanı bağlantılarının mümkün olduğunca azaltılması sağlanarak, matris yapılarının kullanılması suretiyle performans iyileştirmesine gidilmiştir.

Bir verimlilik sorgulama tekniği olarak düşünülmesi gereken bulanık sorgulama araçları, verideki gizli gerçekleri kurallar türetmek yardımıyla açığa çıkartabilen ve bu aşamada bilgi teknolojilerinin faydalanabileceği yeni bir kapı oluşturmaktadır.

## **KAYNAKLAR**

- [1] **Yen, J, Langari, R, Zadeh,** 1994. Industrial Applications of Fuzzy Logic and Intelligent Systems, Texas A&M University, New York
- [2] **Zimmerman, HJ,** 1991. Fuzzy Set Theory and Its Applications, Kluwer Academic Publishers, Boston
- [3] **Kir, GJ, Yuan, B,** 1996. Fuzzy Sets, Fuzzy Logic and Fuzzy Systems, Binghampton University, USA
- [4] **Chiang, W, Lee, J,** 1995. Fuzzy Logic For The Applications to Complex Systems, National Central University, Taiwan
- [5] **Dubois, D, Prade, H, Yager, RR,** 1997. Fuzzy Information Engineering Wiley Computer Publishing, New York
- [6] **Rasmussen, D, Yager, RR,** 1997. Finding Fuzzy and Gradual Functional Dependencies with SummarySQL, Elsevier Science B V, 131-142

## ÖZGEÇMİŞ

Bu çalışmayı hazırlayan Sezin Selen İNCELER, 1978 yılında İstanbul'da doğmuştur. Lise öğrenimini Kadıköy Anadolu Lisesi'nde tamamladıktan sonra, 1997 yılında Yıldız Teknik Üniversitesi Matematik Mühendisliği Bölümü'nde lisans eğitime başlamıştır. 2001 yılında lisans eğitimi tamamlayıp aynı yıl içinde İstanbul Teknik Üniversitesi Mühendislik Bilimleri Anabilim Dalı Sistem Analizi Bölümünde yüksek lisans eğitime başlamıştır. 2002 yılında Banksoft Bankacılık Yazılım Sistemleri İt d.Şti.'nde Kredi Kartları üzerine uygulamalar geliştirmeğe üzere göreve başlayıp halen bu görevini sürdürmektedir.

