

# AZƏRBAYCAN RESPUBLİKASI

*Əlyazması hüququnda*

## **STATİSTİK VERİLƏNLƏRDƏN ƏLDƏ EDİLƏN BİLİKLƏRƏ ƏSASLANAN SİSTEMLƏRİN YARADILMASI ÜÇÜN METODİKANIN İŞLƏNMƏSİ**

İxtisas: 3338.01 – “Sistemli analiz, idarəetmə və  
informasiyanın işlənməsi”

Elm sahəsi: Texnika elmləri

İddiaçı: **Vüqar Elçin oğlu Mirzəxanov**

Fəlsəfə doktoru elmi dərəcəsi almaq üçün  
təqdim edilmiş dissertasiyanın

### **AVTOREFERATI**

**Bakı – 2022**

Dissertasiya işi Azərbaycan Dövlət Neft və Sənaye Universitetinin  
“Kompüter mühəndisliyi” kafedrasında yerinə yetirilmişdir.

Elmi rəhbər: texnika elmləri doktoru, professor  
**Lətafət Abbas qızı Qardaşova**

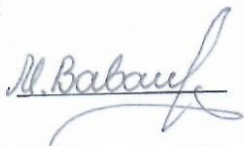
Rəsmi opponetlər: texnika elmləri doktoru, professor  
**Məhəmməd Nurməhəmməd oğlu Nuriyev**

texnika elmləri doktoru, professor  
**Ramin Rza oğlu Rzayev**

texnika elmləri doktoru, professor  
**Cavanşir Firudin oğlu Məmmədov**

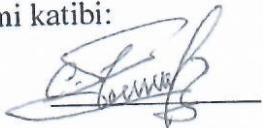
Azərbaycan Respublikasının Prezidenti yanında Ali Attestasiya  
Komissiyasının Azərbaycan Dövlət Neft və Sənaye Universiteti  
nəzdində fəaliyyət göstərən ED 2.02 Dissertasiya şurası

Dissertasiya şurasının  
sədri:



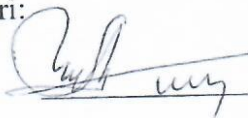
texnika elmləri doktoru, professor  
**Mustafa Baba oğlu Babanlı**

Dissertasiya şurasının  
elmi katibi:



texnika elmləri namizədi, dosent  
**Tahir Qaffar oğlu Cabbarov**

Elmi seminarın  
sədri:



texnika elmləri doktoru, professor  
**Tərlan Səməd oğlu Abdullayev**



## **İŞİN ÜMUMİ XARAKTERİSTİKASI**

**Mövzunun aktuallığı.** İndiki vaxtda insan fəaliyyətinin bir çox sahələrində Data Mining çoxşaxəli elmi qolu ilə əhatə olunan verilənlərin intellektual təhlili metodlarının tətbiqi və istismarına maraq artmaqdadır. Data Mining-in əsas məqsədi yığcam və faydalı bilik elementlərini müəyyən əhəmiyyətli miqdarda ilkin verilənlərdən çıxarmaq üçün metod və vasitələrin dizaynı və tədqiqidir.

Data Mining sahəsi ilə yanaşı, populyarlıq qazanan başqa bir elmi sahə də var – əsas məqsədi biliklərə əsaslanan sistemlərin dizaynı və tədqiqi olan Knowledge-Based Systems. Biliklərə əsaslan sistem (BƏS) tapşırıqları həll etmək üçün müvafiq tətbiq sahəsinə aid açıq-aşkar bilik elementlərindən istifadə edən aparat və/və ya program kompleksidir. Elmi sahənin başlanğıc inkişaf mərhələlərində biliklərə əsaslanan sistemlərin əsas nümayəndələri ekspert və ya müvafiq tətbiq sahəsindəki mütəxəssis qrupunu bilik elementlərinin mənbəyi hesab edən ekspert sistemləri idi. Hal-hazırda, bəzi müəlliflər hələ də biliklərə əsaslanan sistemlərin və ekspert sistemlərinin eyni olduğunu düşünsələr də, cari tendensiya müvafiq verilənlər toplusunun avtomatlaşdırılmış işlənməsindən əldə olunan bilik elementlərini tətbiq etməklə BƏS-i dizayn etməkdir. Beləliklə, müasir biliklərə əsaslanan sistemlər Data Mining alətlərinin tətbiqinə əhəmiyyətli dərəcədə yönəldilmişdir.

Knowledge-Based Systems-də Data Mining metod və vasitələrinin tətbiqi aşağıdakı problemin yaranmasına səbəb olur. Data Mining prosedurunun əsas nəticəsi ilkin verilənləri kompakt və əsaslı şəkildə təsvir edən bir bilik elementidir, halbuki BƏS-in dizayn prosedurunun əsas nəticəsi müvafiq tətbiq sahəsində bəzi müəyyən tipli tapşırıqları avtomatlaşdırılmış şəkildə həll etməyə qadir olan təsviri və ya proqnozlaşdırma sistemidir. Beləliklə, Data Mining alətlərinin tətbiqi biliklərə əsaslanan sistemlərin dizaynı üçün kifayət etmir, bu da biliklərə əsaslanan sistemlərin dizaynında müvafiq Data Mining metod və vasitələrinin tətbiqi məsələlərinə dair tədqiqatların aparılmasında obyektiv tələbat yaradır.

**Tədqiqat obyektı və mövzusu.** Tədqiqat obyektı 2011-ci ildə təqdim olunan Wu–Mendel yanaşmasıdır. Wu–Mendel yanaşması qeyri-səlis Data Mining metodlarından biridir. Wu–Mendel yanaşmasının bir neçə spesifik xüsusiyyəti var: bunların arasında həm təsviri, həm də proqnozlaşdırma məsələlərinin həllində tətbiq olunma xüsusiyyəti əsaslı şəkildə fərqlənir. Tədqiqat mövzusu biliklərə əsaslanan sistemlərin dizaynında Wu–Mendel yanaşmasının tətbiqi məsələləridir.

**Tədqiqatın məqsəd və vəzifələri.** Dissertasiyanın məqsədi Wu–Mendel yanaşmasının BƏS-in dizaynında praktik tətbiqini təmin etmək üçün kompleks və hərtərəfli təhlil, modifikasiya və tamamlanmasıdır. Göstərilən məqsədə çatmaq üçün aşağıdakı əsas məsələləri həll etmək lazımdır:

- Wu–Mendel yanaşmasının təhlili: aşkar olunan problem və məsələlərin həlləri yanaşmanın BƏS-in dizaynında tətbiq olunmasını əsaslandırılmalı və ya gücləndirməlidir.
- İlk verilənlər toplusunun atributlarına qarşı tətbiq olunan avtomatlaşdırılmış fazifikasiya texnikasının sintezi.
- Wu–Mendel yanaşmasının istifadəçi tərəfindən təyin olunan parametrlərdən əhəmiyyətli asılılığının azaldılması.
- Wu–Mendel yanaşmasının özünəməxsus əhəmiyyətli hesablama dəyərinin azaldılması.
- Proqnozlaşdırma məsələlərində yanaşmanın tətbiqini yaxşılaşdırmaq üçün Wu–Mendel yanaşmasındakı keyfiyyət ölçmə düsturlarının modifikasiyası.
- Wu–Mendel yanaşmasının dinamik ilkin verilənlər toplusuna uyğunlaşdırılmasını təmin edən uyğunlaşma alqoritminin dizaynı.
- Wu–Mendel yanaşmasını proqnozlaşdırma məsələlərində tətbiq etməklə əldə olunan BƏS-in fəaliyyətini yaxşılaşdırmaq üçün optimallaşdırma prosedurunun dizaynı.

**Tədqiqat metodları.** Qoyulmuş məsələlər qeyri-səlis çoxluq nəzəriyyə, linqvistik xülasələşdirmə, klasterizasiya, stoxastika elementləri, genetik alqoritm, assosiativ qaydaların çıxarılmasından istifadə etməklə həll edilir.

**Müdafiyyə çıxarılan əsas müddəalar.** Dissertasiyanın aşağıdakı əsas məqamları və nəticələri müdafiə üçün irəli sürülür:

- BƏS-in dizaynında Wu–Mendel yanaşmasının tətbiqi problemlərini aşkar etmək üçün yanaşmanın kompleks təhlili aparılıb.
- Xülasələşdirilən verilənlər toplusunun atributlarına qarşı tətbiq olunan avtomatlaşdırılmış fazifikasiya proseduru təklif olunub.
- Müvafiq avtomatlaşdırılmış proseduru sintez etməklə Wu–Mendel yanaşmasının istifadəçi tərəfindən təyin olunan parametrlərdən asılılığı azaldılıb.
- Bir neçə müvafiq həlli sintez etməklə Wu–Mendel yanaşmasının hesablama dəyəri xeyli aşağı salınıb.
- Wu–Mendel yanaşmasında ƏGƏR–ONDA tipli qaydaların yeni keyfiyyət ölçmə düsturları təklif olunub.
- Wu–Mendel yanaşmasında keyfiyyət ölçmə düsturlarına tətbiq olunan tənzimləmə proseduru təklif olunub ki, bu da yanaşmanın dinamik xülasələşdirilən verilənlər toplusuna uyğunlaşdırılmasını təmin edir.
- Wu–Mendel yanaşmasından istifadə edərək dizayn olunan BƏS-ə tətbiq olunan genetik optimallaşdırma proseduru təklif olunub.

**Elmi yenilik.** Dissertasiyanın elmi yeniliyi aşağıdakı kimi iddia olunur:

1. İlk dəfə Wu–Mendel yanaşmasının kompleks və hərtərəfli təhlili aparılıb.
2. İlk dəfə Wu–Mendel yanaşmasının hərtərəfli modifikasiyası aparılıb ki, bu da hesablama dəyərinin azalması, istifadəçi tərəfindən müəyyən etdiyi parametrlərdən asılılığın azalması, yanaşmanın effektivliyinin əhəmiyyətli dərəcədə artmasına və bir neçə başqa nəticələrə gətirir.
3. İlk dəfə həm proqnozlaşdırıcı, həm də təsviri BƏS-lərin dizaynında Wu–Mendel yanaşmasının tətbiqetmə məsələləri araşdırılıb və həll olunub.

**Nəzəri və praktiki əhəmiyyət.** Dissertasiyada təsviri/proqnozlaşdırma məsələləri həll edən biliklərə əsaslanan

sistemlərin dizaynına imkan verən kompleks metodika təklif olunur. Bu metodikanın üstünlükləri onun avtomatlaşdırılmış istifadəsi; əhəmiyyətli miqdarda verilənləri emal etmək qabiliyyəti; dizayn olunan BƏS-in dinamik ilkin verilənlər toplusuna uyğunlaşdırılma qabiliyyəti; Wu–Mendel yanaşmasının xülasələşdirilən verilənlərin linqvistik təsvirinə yönəldilməsi səbəbi ilə metodikanın tətbiqi prosesinin və nəticələrinin aydın formada təsvir olunmasıdır. Təklif olunan metodika müxtəlif tətbiq sahələrini təsvir edən bir neçə verilənlər toplusundan istifadə edərək təsdiqlənir; metodikanın əksər komponentləri üçün proqram vasitələri təqdim edilmiş, qısa istifadəçi təlimatları ilə təmin edilmiş və ictimaiyyətə təqdim edilmişdir (onlayn şəkildə).

**İşin aprobasiyası.** Dissertasiyanın əsas məqamları aşağıdakı konfranslarda təqdim olunub:

- 13<sup>th</sup> International Conference on Theory and Application of Fuzzy Systems and Soft Computing, (ICAFS-2018), Warsaw, 2018.
- 2019 IEEE International Conference on Fuzzy Systems, (FUZZ-IEEE 2019), New Orleans, 2019.

**Çap olunmuş elmi əsərlər.** Dissertasiya mövzusu üzrə 8 elmi məqalə dərc edilmişdir: 6 məqalə resenziya olunan elmi jurnallarda, 2 məqalə isə konfrans məqalələrinin toplularında dərc edilmişdir. Göstərilən əsərlərin əsas indeksləşdirilməsi aşağıdakı kimi bildirilir:

- 3 məqalə Scopus və Web of Science-də indekslənmiş jurnallar [Science Citation Index Expanded] tərəfindən nəşr edilmişdir.
- 2 məqalə Scopus və Web of Science-də indeksləşdirilmiş konfrans məqalələrinin toplularında dərc edilmişdir.
- 1 məqalə AR AAK tərəfindən tövsiyə olunan bir jurnalda dərc edilmişdir.
- 2 məqalə “РИИЦ” tərəfindən indekslənmiş və RF AAK tərəfindən tövsiyə olunan jurnalda dərc edilmişdir.

**Dissertasiya işinin yerinə yetirildiyi təşkilatın adı.** Dissertasiya işi Azərbaycan Dövlət Neft və Sənaye Universitetinin “Kompüter mühəndisliyi” kafedrasında yerinə yetirilmişdir.

**İşin quruluşu və həcmi.** Dissertasiya işi 160 səhifədə təqdim olunub; giriş, 3 bölmə, nəticə, 314 adda istinad siyahısından ibarətdir; 42 şəkil və 15 cədvəldən istifadə edir. Səkillər, cədvəllər və istinad siyahısı xaricində dissertasiyanın ümumi həcmi 174556 simvoldur.

## İŞİN ƏSAS MƏZMUNU

Mövzunun aktuallığı əsaslandırılmış və dissertasiyanın qısa təsviri **girişdə** verilmişdir.

**Birinci bölmə** əsasən Data Mining elmi qolun tədqiq və təhlilinə həsr olunub və 3 alt bölmədən ibarətdir.

*Birinci alt bölmə*, ilk növbədə, Data Mining-in əsas anlayışlarının, qurulma şərtlərinin və tarixinin araşdırılmasına həsr olunmuşdur:

- “İnformasiya cəmiyyəti” termininin tarixi, tərfi və təhlili təqdim olunub.
- “Verilənlər”, “informasiya” və “bilik” anlayışlarının qısa təhlili, birinci terminin qalan iki termindən kəmiyyət üstünlüyünün əsaslandırılması ilə birlikdə, aparılıb.
- Data Mining-in qurulması və inkişafının praktik məqsəduyğunluğu nümayiş etdirilib.
- “Data Mining” termininin tarixi, tərfi və təhlili təqdim olunub.
- Data Mining-də tətbiq olunan əsas anlayışların tərfi verilib; bunların içərisində ən vacibləri Data Mining metodu, vasitəsi və modelidir.
- Data Mining və Knowledge-Based Systems arasında əlaqə nəzərdən keçirilib.

*İkinci alt bölmə* Data Mining metodlarının tədqiq və təhlilinə həsr edilmişdir; və, əsas Data Mining metodlarından əlavə, bölmədə bəzi digər məsələlərə və *Knowledge Discovery in Databases* prosesinin müvafiq metodlarına da diqqət yetirilir: bunun səbəbi Data Mining-in *Knowledge Discovery in Databases* prosesinin bir addımı kimi təsvir olunma imkanındır. Əsas Data Mining metodları aşağıdakı 5 kateqoriyaya bölünərək bölmədə təsnif edilir:

- Təsnifat və reqressiya (TvR) metodları.

- Xarakterizə və ayrışeçkilik (XvA) metodları.
- Klasterizasiya (K) metodları.
- Xarici analiz (XA) metodları.
- Tez-tez təkrarlanan nümunələrin, assosiasiyaların və əlaqələrin mədən olması (TTNAvƏMO).

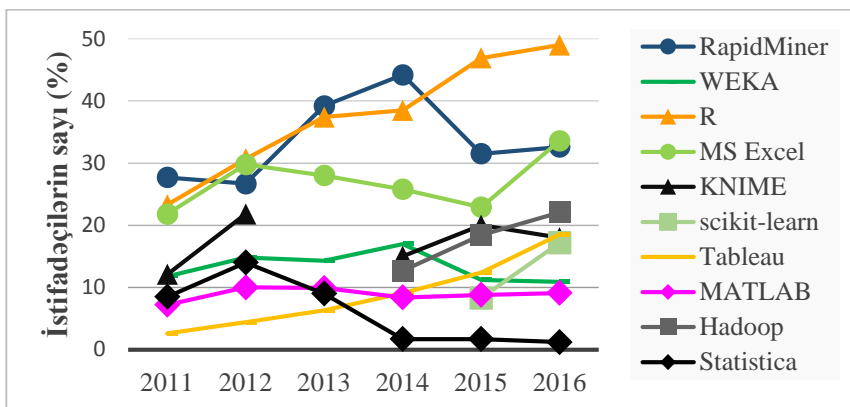
Alt bölmədə əsas diqqət cədvəl 1-də göstərilən 10 əsas Data Mining metodunun tədqiq və təhlilinə yetirilmişdir. Hər baxılan Data Mining metodunun tədqiqi və təhlili aşağıdakı əsas məzmunlu öz alt hissəsini təşkil edir: metodun tərifi, qısa tarixi (metodun qurulması və inkişafı), əsas xüsusiyyətləri, metodun riyazi təsviri və dəyişmələrinin təsnifatı [mümkündürsə], digər Data Mining metodları və Soft Computing komponentləri ilə birlikdə istifadə olunma qabiliyyəti, metodun insan fəaliyyətinin müxtəlif sahələrində praktik tətbiqi.

**Cədvəl 1**  
**Əsas Data Mining metodlarının təsnifatı**

Metodlar	TvR	XvA	K	XA	TTNAvƏMO
Neyron şəbəkələr	+				
Linqvistik xülasələşdirmə				+	+
Assosiativ qaydaların çıxarılması					+
Regressiya analizi	+	+			
Verilənlər kubu		+			
Naive Bayes təsnifatı	+				
Klasterizasiya			+	+	
Genetik alqoritm	+				
Qərar ağacları	+				
Diskriminant analizi	+				

*Üçüncü alt bölmə* müasir Data Mining proqram vasitələrinin tədqiq və müqayisəli təhlilinə həsr edilmişdir. Alt bölmədə əsas diqqət son illər ərzində Data Mining icmasında olduqca populyar olan 10 proqram vasitəsinə yetirilmişdir (şəkil 1).





**Şəkil 1 Data Mining vasitələrinin Data Mining icmasında populyarlığa görə müqayisəsi (illik KDnuggets Software Sorğu Nəticələrinə əsasən)**

Alt bölmədə nəzərdən keçirilən bütün Data Mining vasitələri üç kateqoriyaya bölünür:

- Əsas Data Mining vasitələri kimi istifadə olunan və Data Mining icmasında populyar olan əhəmiyyətli Data Mining vasitələri (cədvəl 2).
- Data Mining-in bəzi addımları zamanı ikincil Data Mining vasitələri kimi istifadə olunan əlavə Data Mining vasitələri (cədvəl 3).
- Əsas Data Mining vasitələri kimi istifadə olunan, lakin Data Mining icmasında populyar olmayan əhəmiyyətsiz Data Mining vasitələri (cədvəl 4).

**İkinci bölmə** biliklərə əsaslanan sistemlərin dizaynı üzrə metodikanın işlənmə məsələsinin tərtibinə, təhlilinə və həllinə həsr edilmişdir. Metodikanın əsas hissəsi 2011-ci ildə təqdim edilmiş və aşağıda qısa təsvir olunan qeyri-səlis Data Mining metodu – Wu–Mendel yanaşmasıdır<sup>1</sup>.

<sup>1</sup> Wu, D., Mendel, J. Linguistic Summarization Using IF–THEN Rules and Interval Type-2 Fuzzy Sets // IEEE Transactions on Fuzzy Systems, – 2011. № 1, – p. 136–151.

**Cədvəl 2****Əhəmiyyətli Data Mining vasitələri**

<b>Vasitənin adı</b>	<b>Xidmət dövrü</b>	<b>Müəllif və ya sahibkar</b>
RapidMiner	2001 – ****	RapidMiner, ABŞ və Almaniya
WEKA	1994 – ****	University of Waikato, Yeni Zelandiya
IBM SPSS Modeler	1994 – ****	IBM Corp., ABŞ
KNIME	2006 – ****	KNIME.com AG, İsveçrə və Almaniya
Orange	1997 – ****	University of Ljubljana, Sloveniya
R	1991 – ****	Ross Ihaka və Robert Gentleman
scikit-learn (Python)	2007 – ****	David Cournapeau və s.
SAS Enterprise Miner	1999 – ****	SAS, ABŞ
Oracle Data Mining	2002 – ****	Oracle Corp., ABŞ
SPM software suite	2010 – ****	Salford Systems, ABŞ
Apache Hadoop	2011 – ****	Apache Software Foundation, ABŞ
H2O	2011 – ****	H2O.ai, ABŞ
MATLAB	1984 – ****	MathWorks, ABŞ
Azure Machine Learning Studio	2014 – ****	Microsoft Corp., ABŞ

**Cədvəl 3****Əlavə Data Mining vasitələri**

<b>Vasitənin adı</b>	<b>Xidmət dövrü</b>	<b>Müəllif və ya sahibkar</b>
Microsoft Excel	1985 – ****	Microsoft Corp., ABŞ
Google Sheets	2006 – ****	Google Inc., ABŞ
EDM Workbench	2011 – ****	Ateneo Laboratory for the LS, Filippin
Jupyter Notebook (Python)	2014 – ****	NumFOCUS Foundation, ABŞ
Tableau	2003 – ****	Tableau Software, ABŞ
D3.js	2011 – ****	Mike Bostock və s.
MS SQL Server	1989 – ****	Microsoft Corp., ABŞ

## Əhəmiyyətsiz Data Mining vasitələri

Vasitənin adı	Xidmət dövrü	Müəllif və ya sahibkar
DataDetective	1991 – ****	Sentient Information Systems, Niderland
GhostMiner	2002 – ****	FQS Poland/Fujitsu Ltd., Yaponiya
KEEL	2004 – ****	Spanish National Projects, İspaniya
Angoss Knowledge STUDIO	1997 – ****	Angoss Software Corp., Kanada
ADAPA	2007 – ****	Zementis, ABŞ
Viscovery suite	2002 – ****	Viscovery Co., Avstriya
Coheris SPAD	1996 – ****	Coheris SA, Fransa
TANAGRA	2004 – 2013	Lumière University Lyon 2, Fransa
GGobi	2001 – ****	Deborah Swayne və s.
Statistica	1984 – ****	Statsoft, ABŞ Dell Software, ABŞ Francisco Partners, ABŞ
ELKI	2008 – ****	Erich Schubert, Arthur Zimek və s.

Wu–Mendel yanaşmasında tətbiq olunan ilkin verilənlər toplusu (İVT) aşağıdakı kimi müəyyən edilir:

$$D = \{d_j^i | i = 1, \dots, I; j = 1, \dots, J\}, \quad (1)$$

burada  $d_j^i$  –  $i$ -ci qeyddəki  $j$ -ci atributun dəyəri;  $I$  və  $J$ , müvafiq olaraq qeydlərin və atributların sayıdır.

Wu–Mendel yanaşmasının tətbiqində məqsəd  $D$ -i təsvir edən bütün mümkün linqvistik qaydaların siyahısını çıxarmaqdır:

$$\begin{aligned} \text{ƏGƏR } A_1 = S_1 \text{ və } \dots \text{ və } A_m = S_m, \text{ ONDA} \\ A_{m+1} = S_{m+1} \text{ və } \dots \text{ və } A_{m+n} = S_{m+n} [Q], \end{aligned} \quad (2)$$

burada  $A_j$  ( $j = 1, \dots, m + n$ ) ilkin verilənlər toplusunun atributudur;

$S_j$  ( $j = 1, \dots, m + n$ ) –  $A_j$  atributunun tip-1 (T1) və ya interval tip-2 (İT2) qeyri-səlis çoxluğu (QŞÇ) ilə modelləşdirilmiş bir termdir;  $Q \in [0, 1]$  – qaydanın nə qədər keyfiyyətli olduğunu göstərən bir ölçüdür.

Wu–Mendel yanaşması aşağıdakı 5 keyfiyyət ölçülərini təyin edir:

- Doğruluq dərəcəsi ( $T$ ) – çıxarılan qaydanın dəqiqliyini təyin edən keyfiyyət ölçüsüdür.
- Əhatə dərəcəsi ( $C$ ) – çıxarılan qaydanın ümumiliyini təyin edən keyfiyyət ölçüsüdür.
- Etibarlılıq dərəcəsi ( $R$ ) – çıxarılan qaydanın etibarlılığını təyin edən keyfiyyət ölçüsüdür.
- Yenilik dərəcəsi ( $U$ ) – çıxarılan qaydanın yeniliyini təyin edən keyfiyyət ölçüsüdür.
- Sadəlik dərəcəsi ( $S$ ) – çıxarılan qaydanın mürəkkəbliyini təyin edən keyfiyyət ölçüsüdür.

T1 QŞÇ-lar tətbiq edilərsə *doğruluq dərəcəsi* aşağıdakı kimi hesablanır:

$$T = \frac{\sum_{i=1}^l \min(\mu_{S_1}(d_1^i), \dots, \mu_{S_{m+n}}(d_{m+n}^i))}{\sum_{i=1}^l \min(\mu_{S_1}(d_1^i), \dots, \mu_{S_m}(d_m^i))}. \quad (3)$$

T1 QŞÇ-lar tətbiq edilərsə *əhatə dərəcəsi* aşağıdakı kimi hesablanır:

$$C = \begin{cases} 0, & r_c \leq r_c^{\min} \\ 2 \left( \frac{r_c - r_c^{\min}}{r_c^{\max} - r_c^{\min}} \right)^2, & r_c^{\min} < r_c < \frac{r_c^{\min} + r_c^{\max}}{2} \\ 1 - 2 \left( \frac{r_c^{\max} - r_c}{r_c^{\max} - r_c^{\min}} \right)^2, & \frac{r_c^{\min} + r_c^{\max}}{2} \leq r_c < r_c^{\max} \\ 1, & r_c \geq r_c^{\max} \end{cases}, \quad (4)$$

burada  $r_c^{\min}$  və  $r_c^{\max}$  istifadəçi tərəfindən təyin olunan parametrlərdir;  $r_c$  aşağıdakı kimi hesablanır:

$$r_c = \frac{\sum_{i=1}^l t_i}{l}, \quad (5)$$

burada  $t_i$  aşağıdakı kimi hesablanır:

$$t_i = \begin{cases} 1, & \mu_{S_1}(d_1^i) > 0 \text{ və ... və } \mu_{S_{m+n}}(d_{m+n}^i) > 0 \\ 0, & \text{əks halda} \end{cases}. \quad (6)$$

*Etibarlılıq dərəcəsi* iki əsas keyfiyyət ölçülərindən biridir və aşağıdakı kimi hesablanır:

$$R = \min(T, C). \quad (7)$$

*Yenilik dərəcəsi* iki əsas keyfiyyət ölçülərindən biridir və aşağıdakı kimi hesablanır:

$$U = \begin{cases} \min(\max(T, 1 - T), 1 - C), & T > 0 \\ 0, & T = 0 \end{cases}. \quad (8)$$

*Sadəlik dərəcəsi* aşağıdakı kimi hesablanır:

$$S = 2^{2-l}, \quad (9)$$

burada  $l = m + n$  – qaydada istifadə olunan atributların ümumi sayıdır.

Linqvistik xülasələşdirmə (LX) metodu kimi təqdim olunmasına baxmayaraq, Wu–Mendel yanaşması mövcud olan linqvistik xülasələşdirmə metodlarının arasında fərqlənməyə imkan verən bir sıra spesifik xüsusiyyətlərə malikdir.

Birincisi, Wu–Mendel yanaşması ƏGƏR–ONDA tipli qaydanın beş fərqli xassəsini qiymətləndirmək üçün beş keyfiyyət ölçüsü tətbiq edir, halbuki başqa mövcud olan LX metodlarının əksər hissəsi xülasənin yalnız bir xassəsini qiymətləndirir. Wu–Mendel yanaşmasındakı keyfiyyət ölçüləri çıxarılan linqvistik qaydalarının dəqiqliyi, ümumiliyi, etibarlılığı, yeniliyi və sadəliyini ölçmək üçün

istifadə olunur. Bu cür müxtəliflik qaydaları daha əhatəli şəkildə qiymətləndirməyə imkan verir: məsələn, eyni dərəcədə dəqiq olan iki qaydanı daha asan başa düşülən bilik elementini tapmaq üçün sadəlik ölçüsü üzrə müqayisə etmək olar.

İkincisi, Wu–Mendel yanaşması qeyri-səlis məntiqin müxtəlif növlərini dəstəkləyir: məqalə<sup>1</sup>-də T1 və İT2 qeyri-səlis məntiq tiplərinin tətbiqi nümayiş etdirilir.

Üçüncüsü və ən maraqlısı, Wu–Mendel yanaşmasının tətbiq olunması nəticəsində əldə edilən linqvistik xülasə ƏGƏR–ONDA tipli qaydaların toplusudur, halbuki LX-də aparılan tədqiqatların əksər hissəsi ƏGƏR–ONDA tipli qaydaların əvəzinə kəmiyyət cümlələri tətbiq edir. ƏGƏR–ONDA tipli qaydaların istifadə edilməsinin əsas üstünlüyü, Wu–Mendel yanaşmasının BƏS-lərin ən geniş yayılmış növləri arasında olan *qaydalara əsaslanan sistemlərin* dizaynında bilavasitə istifadə olunma qabiliyyətidir.

Yuxarıda göstərilənə baxmayaraq qeyd etmək lazımdır ki, əvvəlcədən LX metodu kimi təqdim edildiyi üçün orijinal Wu–Mendel yanaşması proqnozlaşdırma məsələlərinin həllinə yox, təsviri məsələlərin həllinə yönəldilmişdir; halbuki biliklərə əsaslanan sistemlər əsasən proqnozlaşdırma məsələlərinin həllinə yönəldilib. Yuxarıda göstərilənlərə görə, statistik verilənlərdən əldə edilən biliklərə əsaslanan sistemlərin dizaynı üçün metodikanın işlənməsi iki əsas tədqiqat istiqamətinə malikdir:

- Wu–Mendel yanaşmasının təsviri məsələlərdə tətbiq olunmasının təkmilləşdirilməsi.
- Wu–Mendel yanaşmasının proqnozlaşdırma məsələlərində tətbiq olunması imkanının yaradılması.

Biliklərə əsaslanan sistemlərin yaradılması üçün təklif olunan metodika aşağıdakı 4 addımın ardıcılığı şəklində təqdim olunur (şəkil 2): verilənlərin emalına ilkin hazırlama, linqvistik xülasələşdirmə, BƏS-in sintezi, BƏS-in optimallaşdırılması.

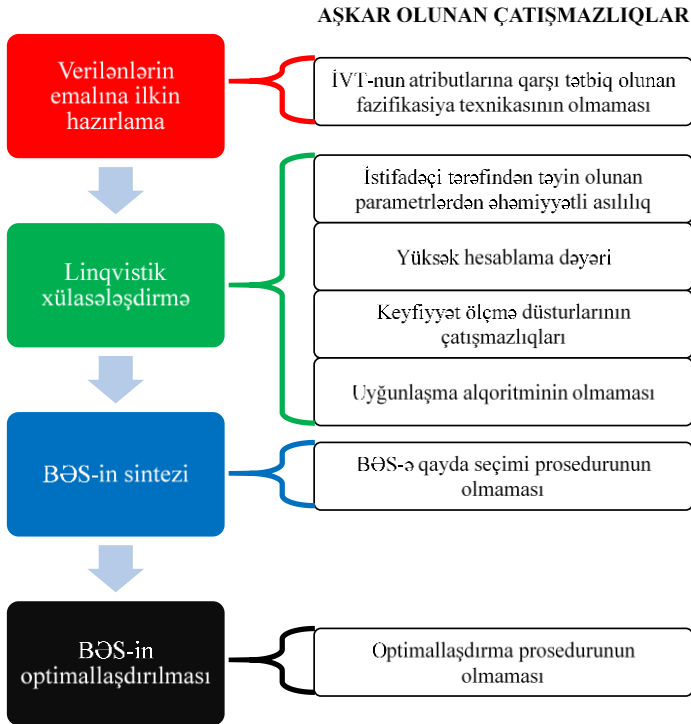
*Verilənlərin emalına ilkin hazırlama.* Bu addım zamanı sonrakı linqvistik xülasələşdirmə prosesinə hazır olmaq üçün verilənlər toplusu ilkin emaldan keçirilir.

*Linqvistik xülasələşdirmə.* Bu addım zamanı İVT-nu təsvir edən ƏGƏR–ONDA tipli qaydaların siyahısı hazırlanır.

*BƏS-in sintezi.* Bu addım zamanı Mamdani tipli qeyri-səlis çıxarış sisteminin (QSÇS) qaydalar bazasına ƏGƏR–ONDA tipli qaydalar seçilir.

*BƏS-in optimallaşdırılması.* Bu addım zamanı sintez edilmiş Mamdani tipli QSÇS işini yaxşılaşdırmaq üçün optimallaşdırılır.

BƏS-in dizaynının bütün yuxarıda göstərilən addımları halında çatışmazlıqlar aşkar edilib, və bu çatışmazlıqların aradan qaldırılması metodikanın adekvat tətbiq olunmasını təmin edir (şəkil 2). Aşkar edilmiş hər bir çatışmazlıq biliklərə əsaslanan sistemlərin yaradılma metodikasının işlənməsi üçün müvafiq məsələni doğurur; və aşkar edilmiş hər bir məsələnin tərfi, təhlili və həlli, aşağıda müxtəsərcə təsvir edilmiş məzmunu ilə, dissertasiyada öz alt bölməsini təşkil edir.



**Şəkil 2 BƏS-in yaradılması üçün təklif olunan metodikanın əsas addımları və çatışmazlıqları**

2.3-cü alt bölmədə, LX metodlarının əksəriyyətində İVT-nun atributlarının fazifikasiya məsələsinə kifayət qədər diqqət yetirilmədiyi bildirilir: bu prosedur ya tədqiqat sahəsindən kənarında olduğuna görə araşdırılmır, ya da istifadəçi/mütəxəssis tərəfindən müəyyən edilmiş prosedur kimi göstərilir. Wu–Mendel yanaşmasında bu çatışmazlığı aradan qaldırmaq üçün alt bölmədə İVT-nun atributlarının istifadəçi tərəfindən müəyyənləşdirilmiş fazifikasiyasını tənzimləmək üçün avtomatlaşdırılmış fazifikasiya tənzimləməsi üsulu təklif olunur: İVT-nun atributlarını təsvir edən bitişik (qonşu) linqvistik termlərin hər cütünün bir-birinə oxşarlığı təhlil olunur və yüksək oxşarlığa malik olduğu təqdirdə həmin cütlər vahid ortaqlıq linqvistik termlərlə əvəz olunurlar. Təklif olunan prosedur keyfiyyət baxımından əhəmiyyətli dərəcədə azaltma doğurmadan linqvistik termlərin və ƏGƏR–ONDA tipli qaydaların ümumi sayını azaltmaqla çıxarılan linqvistik xülasəni sadələşdirir.

Baxılan alt bölmədə oxşarlıq xassəsi  $O \in [0, 1]$  üst-üstə düşmə dərəcəsi ilə təmsil olunur. Üst-üstə düşmə dərəcəsi  $T$ -dən əldə edilən bir keyfiyyət ölçüsüdür, və bu tərifə aşağıdakı kimi başa düşmək olar: üst-üstə düşmə dərəcəsi yalnız Wu–Mendel yanaşması ilə linqvistik xülasələşdirmə aparıldıqdan sonra hesablanır və təbiiq edilə bilər. Üst-üstə düşmə dərəcəsi aşağıdakı kimi hesablanır:

$$O(S_j^{z_j}, S_j^{z_j+1}) = \frac{\sum_{w=1}^W (1 - |T(S_j^{z_j}, P_w) - T(S_j^{z_j+1}, P_w)|)}{W}, \quad (10)$$

burada  $S_j^{z_j}, S_j^{z_j+1}$  ( $j = 1, \dots, m + n$ ) –  $A_j$  atributunun bitişik (qonşu) linqvistik termləridir;  $(S_j^{z_j}, P_w) - S_j^{z_j}$  termini istifadə edilən  $w$ -ci qaydadır;  $P_w = P_w(S_j^{z_j}, S_j^{z_j+1}) - S_j^{z_j}/S_j^{z_j+1}$  termlərinin istifadəsində fərqlənən iki qaydanın qalıq hissəsidir;  $W - S_j^{z_j}/S_j^{z_j+1}$  termlərinin istifadəsində fərqlənən qayda cütlərinin sayıdır.

Təqdim olunan oxşarlıq analizinə əsaslanan fazifikasiya tənzimləməsi texnikası optimallaşdırma üsuludur, və bu tərifə aşağıdakı kimi başa düşmək olar: texnika artıq yerinə yetirilmiş



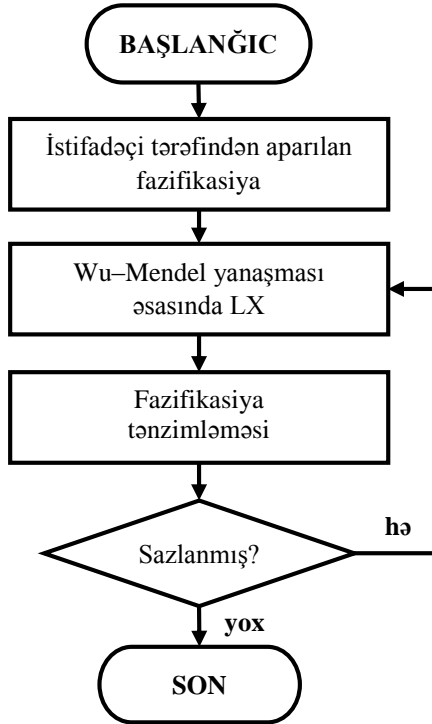
fazifikasiya nəticələrini tənzimləyir. Texnika İVT-nun bütün atributlarının qonşu termlərinin hər cütü üçün üst-üstə düşmə dərəcəsini hesablayır, və  $O \geq O^*$  xassəsinə malik olan cütlər vahid ümumi linqvistik termlərlə əvəz olunmalıdır, burada  $O^*$  – istifadəçi tərəfindən müəyyən edilmiş bir dəyərdir. Ümumiyyətlə,  $O^*$  dəyəri nə qədər az olarsa, əvəz olunan term cütlərinin sayı o qədər də çox olar. Ancaq, qeyd olunmalıdır ki, linqvistik xülasədə tətbiq olunan linqvistik termlərin çox az olması onun keyfiyyətinə mənfi təsir göstərir. Beləliklə,  $O^*$  dəyərini bəzi ağlabatan hədudlarda saxlamaq vacibdir. Dissertasiyada  $O^* \in [0.9, 0.95]$  şərtindən istifadə edilir. Lakin, qeyd olunmalıdır ki, bəzi konkret hallarda, haçan xülasənin sadəliyi onun keyfiyyətindən daha çox vacibdir,  $O^*$  dəyərini aşağı salınması mümkündür. Bundan əlavə,  $O \geq O^*$  xassəsinə malik olan term sırasında 1-dən çox qonşu term cütlüyü varsa, yalnız ən böyük  $O$  dəyərində malik olan cütlük birləşdirilir: bu eyni linqvistik termlərə və yüksək  $O$  dəyərində malik olan cütlüklər olduqda münaqişələrin idarə olunmasına kömək edir.

Linqvistik xülasələşdirmədə təklif olunan fazifikasiya tənzimləməsi üsulunun mümkün tətbiqi şəkil 3-də göstərilib. Şəkil 3-dən görüldüyü kimi, təklif olunan tətbiq prosesi aşağıdakı addımlardan ibarətdir:

- *İstifadəçi tərəfindən aparılan fazifikasiya.* İlk verilənlər toplusunun hər atributu üçün qeyri-səlis linqvistik dəyişənin müəyyənləşdirilməsi istifadəçi tərəfindən həyata keçirilir.
- *Wu–Mendel yanaşması əsasında LX.* İlk verilənlərin linqvistik xülasələşdirməsi Wu–Mendel yanaşmasından istifadə etməklə aparılır.
- *Fazifikasiya tənzimləməsi.* Təyin olunan linqvistik dəyişənləri tənzimləmək üçün təklif olunan fazifikasiya tənzimləməsi üsulu tətbiq olunur.
- *“Saxlanmış?” qərar addımı.* Baxılan iterasiyada heç bir tənzimləmə aparılmırsa, bu  $O \geq O^*$  xassəsinə malik olan term cütlərinin olmaması deməkdir, onda bütün iterativ tətbiq prosesi başa çatır. Əks təqdirdə, baxılan iterasiyada tənzimləmə yerinə yetirilsə, onda LX ilkin verilənlər

toplusunun atributlarının t nziml nmif qeyri-s lis d yiif nl rind n istifad  etm kl  yened n h yata ke irilir.

2.4-c  alt b lmədə, orijinal Wu–Mendel yanaifmasının istifad ci/ekspert t r find n veril n q rarlardan  h miyy tli d rəcədə asılılıęa sahib olduęu ifadə edilir: x susil , (4)-d   $r_c^{min}$  v   $r_c^{max}$  parametrl ri istifad ci t r find n m  yyn edilir, halbuki beif keyfiyy t ol l rind n  c  ( $C, R, U$ ) onlardan asılıdır. Bu  atıifmazlıęı aradan qaldırmaq  c n alt b lmədə  $r_c^{min}$  v   $r_c^{max}$  parametrl rin daha optimal d y rl rini hesablamaq  c n b zi d sturlar t klif olunur.



**Őakil 3 T klif olunan fazifikasiya t nziml m si  sulunun linqvistik x lasel fdirmədə t tbii**

Baxılan alt b lmədə,  $r_c^{max}$  parametrl nin istifad ci t r find n m  yynl fdirilm si  v zin , aŐaęıdakı m lahizel r  uyęun olaraq hesablanmas  t klif olunur. (2)-d  t svir olunan qaydanın  $r_c^{max}$

parametri müvafiq linqvistik termlərdən asılıdır, buna görə  $r_c^{max}$  parametri  $S_1, S_2, \dots, S_{m+n}$  termlərinin funksiyası kimi təyin edilə bilər. Alt bölmədə,  $r_c^{max}$  parametrinin aşağıdakı kimi hesablanması təklif olunur:

$$\begin{aligned} & r_c^{max}(S_1, S_2, \dots, S_{m+n}) = \\ & = r_c^{max}(S_1) \times r_c^{max}(S_2) \times \dots \times r_c^{max}(S_{m+n}), \end{aligned} \quad (11)$$

burada  $r_c^{max}(S_j)$  ( $j = 1, \dots, m + n$ ) parametrinin hesablanması  $A_j$  atributunun tipindən asılıdır və aşağıdakı kimi müəyyənləşdirilir.

$A_j$  kəsilməz ədədi atribut olduğu halda,  $r_c^{max}(S_j)$  ( $j = 1, \dots, m + n$ ) aşağıdakı kimi hesablanır:

$$r_c^{max}(S_j) = \frac{r_{S_j} - l_{S_j}}{r_{A_j} - l_{A_j}} \times \frac{\int_{l_{S_j}}^{r_{S_j}} \mu_{S_j}(A_j) dA_j}{(r_{S_j} - l_{S_j}) \times 1} = \frac{\int_{l_{S_j}}^{r_{S_j}} \mu_{S_j}(A_j) dA_j}{r_{A_j} - l_{A_j}}, \quad (12)$$

burada  $[l_{S_j}, r_{S_j}]$  –  $S_j$  terminin daşıyıcısıdır (hesablamalarda tətbiq olunmur və burada yalnız düsturun alınma prosesinin aydınlaşdırılması üçün verilir);  $[l_{A_j}, r_{A_j}]$  –  $A_j$  atributunun qiymət intervalıdır.

$A_j$  diskret ədədi və ya kateqoriyalı atribut olduğu halda,  $r_c^{max}(S_j)$  ( $j = 1, \dots, m + n$ ) aşağıdakı kimi hesablanır:

$$r_c^{max}(S_j) = \frac{1}{\text{card}(A_j)}, \quad (13)$$

burada  $\text{card}(A_j)$  –  $A_j$  atributunu təsvir edən linqvistik termlərin ümumi sayıdır.

(11) və (12)-dən görüldüyü kimi,  $r_c^{max}$  parametri fərqli uzunluğa və fərqli tipli istifadə olunan linqvistik termlərə sahib olan qaydalar üçün eyni deyil. Beləliklə, LX-də  $r_c^{max}$  parametrinin vahid ümumi dəyərindən istifadə etmək üçün aşağıdakı məhdudiyətlər qəbul edilməlidir:

- LX-də nəzərdən keçirilən qaydalar eyni uzunluğa və eyni istifadə olunan atributlara malik olmalıdır.
- Bir atributu təsvir edən linqvistik termlərin toplusu eyni tipli mənsubiyyət funksiyaları ilə təmsil olunmalıdır.

$r_c^{min}$  baxılan alt bölmədə aşağıdakı kimi hesablanır:

$$r_c^{min}(S_1, S_2, \dots, S_{m+n}) = d_r \times r_c^{max}(S_1, S_2, \dots, S_{m+n}), \quad (14)$$

burada  $d_r$  – istifadəçi tərəfindən təyin edilmiş parametrdir.

Baxılan alt bölmədə  $d_r$  parametrinin 0.1-ə bərabər olması təklif olunur, və bu aşağıdakıları göstərir:  $C > 0$  xassəsinə malik olan qayda  $r_c = r_c^{max}$  və  $C = 1$  xassələrinə malik olan qayda ilə əhatə olunan qeydlərin 10% -dan çoxunu əhatə etməlidir.

2.5-ci alt bölmədə, (6) və (7) düsturlarının tətbiqi, bəzi hallarda, linqvistik xülasələşdirmənin qeyri-kafi nəticələrinə səbəb ola biləcəyi ifadə edilir. Wu–Mendel yanaşmasındakı bu çatışmazlığı aradan qaldırmaq üçün alt bölmədə  $C$  və  $R$  ölçüləri üçün yeni düsturlar təklif olunur.

(6)-ni əvəz etmək üçün aşağıdakı düstur təklif olunur:

$$t_i = \mu_{S_1}(d_1^i) \times \dots \times \mu_{S_{m+n}}(d_{m+n}^i). \quad (15)$$

(7)-ni əvəz etmək üçün aşağıdakı düstur təklif olunur:

$$R = T \times C. \quad (16)$$

2.6-cı alt bölmədə, Wu–Mendel yanaşmasının hesablama dəyərinin kompleks təhlili həm nəzəri, həm də praktik baxımdan aparılır. Hesablama dəyərinin ölçüsü kimi linqvistik xülasələşdirmənin vaxt üzrə uzunluğu olan işləmə müddəti (İM) tətbiq olunur. Wu–Mendel yanaşmasının İM-nin azaldılması üçün üç nəzəri və iki praktik (proqram) həll təklif olunur.

Wu–Mendel yanaşmasının hesablama dəyərini azaltmaq üçün nəzəri həllər aşağıdakılardır:

- *DRC (Data Reduction by Clustering)*: K-means<sup>2</sup> klasterləşdirmə yolu ilə ilkin verilənlərin kompressiyasını həyata keçirir və İM-ni azaldır.
- *RRA (Rule Reduction by Apriori)*: Apriori-nin<sup>3</sup> sadələşdirilmiş versiyası əsasında işlənmiş qaydaların sayını azaldaraq İM-ni azaldır.
- *DRRS (Data and Rule Reduction by Stochastics)*: ilkin mərhələdə "mövcud olmayan" qaydaların stoxastik əsaslı kəşfi və xülasədən ləğv olunması yolu ilə İM-ni azaldır; burada "mövcud olmayan" xassəsi  $T$  və  $C$  dəyərlərinin son dərəcə aşağı olması deməkdir.

Wu–Mendel yanaşmasının hesablama dəyərini azaltmaq üçün praktik həllər aşağıdakılardır:

- *Wu–Mendel yanaşmasının paralelləşməsi*: LX-nin kompüter klasterinin bir neçə qurğusunda eyni vaxtda yerinə yetirilməsinə imkan verir.
- *Mənsubiyyət matris sütunlarının birdəfəlik hesablanması*: xülasədəki hər qayda üçün  $Q$ -nin hesablanması zamanı ilkin verilənlərin lazımsız siklik işlənməsinin zəruriliyini aradan qaldırır (həll tətbiq olunarsa, İVT LX-nin əvvəlində yalnız bir dəfə işlənir).

Wu–Mendel yanaşmasının hesablama dəyərini azaltmaq üçün sadalanan həllər olduqca mürəkkəb təsvirə malik olduğundan, onların analitik təqdimatı və təhlili burada buraxılmışdır.

2.7-ci alt bölmədə, orijinal Wu–Mendel yanaşmasının çatışmazlıqlarından birinin dinamik İVT-na uyğunlaşma olmaması olduğu bildirilir: xülasələşdirilən ilkin verilənlər toplusu vaxtaşırı yeni qeydlərlə tamamlanırsa, xülasəni yeniləmək üçün LX bütün verilənlər toplusu tətbiq olunaraq təkrar-təkrar yerinə yetirilməlidir.

---

<sup>2</sup> Lloyd, S. Least squares quantization in PCM // IEEE Transactions on Information Theory, – 1982. № 2, – p. 129-137.

<sup>3</sup> Agrawal, R., Srikant, R. Fast algorithms for mining association rules // Proceedings of the 20th VLDB Conference, – San Francisco: Morgan Kaufmann Publishers Inc., – 1994, – p. 487-499.

Baxılan alt bölmədə, Wu–Mendel yanaşmasında mövcud olan bu çatışmazlığı aradan qaldırmaq üçün Wu–Mendel yanaşmasını inkrementasiya xüsusiyyəti ilə tamamlayan və, nəticədə, bir linqvistik xülasənin dinamik ilkin verilənlər toplusuna uyğunlaşmasını təmin edən alqoritm təklif olunur: inkrementasiya xüsusiyyəti, bütün dinamik ilkin verilənlər toplusunun əvəzinə, yalnız əlavə edilmiş qeydləri işləyərək xülasə qaydalarının keyfiyyət ölçülərini ( $T, C, R, U, S$ ) tənzimləmək qabiliyyətini ifadə edir.

$S$  ölçüsünün inkremental şəkildə dəyişdirilməsinə ehtiyac yoxdur, çünki yalnız qayda uzunluğundan asılıdır, bu da xülasədəki müəyyən qayda üçün eynidir və yeni verilənlər daxil olduqda dəyişmir.  $R$  və  $U$  keyfiyyət ölçüləri  $T$  və  $C$ -nin funksiyalarıdır, buna görə doğruluq və əhatə dərəcələrinin inkremental şəkildə dəyişdirilmiş dəyərlərindən istifadə etməklə birbaşa hesablanı bilər. Beləliklə, Wu–Mendel yanaşmasının inkremental versiyası yalnız iki keyfiyyət ölçüsünün inkremental şəkildə yenilənməsini tələb edir:  $T$  və  $C$ .

Qəbul edək ki, ilkin verilənlər toplusu  $D$  iki ardıcıl verilənlər toplularının (qeydlərinin sayı  $I_1$  olan  $D_1$  verilənlər toplusunun və qeydlərinin sayı  $I_2$  olan  $D_2$  verilənlər toplusunun) birləşməsidir. LX yanaşması,  $D_1$  toplusunun LX-dən sonra  $D_1$ -i yenidən işlətmədən ( $D_1 + D_2$ ) toplusunun LX-ni yerinə yetirə bilərsə, inkrementasiya xüsusiyyətinə malik hesab olunur.

$T$  keyfiyyət ölçüsünün inkrementasiya xüsusiyyətinə malik olması düstur (3)-ün yenidən nəzərdən keçirilməsini nəzərdə tutur. (3)-ə görə,  $D_1$  toplusunun LX zamanı və ( $D_1 + D_2$ ) toplusunun LX zamanı  $T$  keyfiyyət ölçüsü aşağıdakı kimi hesablanır:

$$T(D_1) = \frac{\sum_{i=1}^{I_1} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_{m+n}}(a_{m+n}^i))}{\sum_{i=1}^{I_1} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_m}(a_m^i))} = \frac{c_{D_1}(S_1, \dots, S_{m+n})}{c_{D_1}(S_1, \dots, S_m)}. \quad (17)$$

$$\begin{aligned} T(D_1 + D_2) &= \\ &= \frac{\sum_{i=1}^{I_1} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_{m+n}}(a_{m+n}^i)) + \sum_{i=I_1+1}^{I_1+I_2} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_{m+n}}(a_{m+n}^i))}{\sum_{i=1}^{I_1} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_m}(a_m^i)) + \sum_{i=I_1+1}^{I_1+I_2} \min(\mu_{S_1}(a_1^i), \dots, \mu_{S_m}(a_m^i))} = \\ &= \frac{c_{D_1}(S_1, \dots, S_{m+n}) + c_{D_2}(S_1, \dots, S_{m+n})}{c_{D_1}(S_1, \dots, S_m) + c_{D_2}(S_1, \dots, S_m)}. \end{aligned} \quad (18)$$

(17) və (18)-dən göründüyü kimi,  $T$  keyfiyyət ölçüsü inkremental şəkildə aşağıdakı kimi hesablanıla bilər:

1.  $D_1$  toplusunun LX zamanı, hesablanmış  $T$  linqvistik xülasəyə  $c_{D_1}(S_1, \dots, S_{m+n})$  və  $c_{D_1}(S_1, \dots, S_m)$  dəyərləri ilə birlikdə əlavə olunur.
2.  $(D_1 + D_2)$  toplusunun ardıcıl LX zamanı,  $c_{D_2}(S_1, \dots, S_{m+n})$  və  $c_{D_2}(S_1, \dots, S_m)$  dəyərləri hesablanır, və  $T$  keyfiyyət ölçüsü hesablanmış  $c_{D_2}(S_1, \dots, S_{m+n})$  &  $c_{D_2}(S_1, \dots, S_m)$  və yadda saxlanmış  $c_{D_1}(S_1, \dots, S_{m+n})$  &  $c_{D_1}(S_1, \dots, S_m)$  dəyərlərinin əsasında hesablanır.

$C$  keyfiyyət ölçüsünün inkrementasiya xüsusiyyətinə malik olması düstur (5)-in yenidən nəzərdən keçirilməsini nəzərdə tutur. (5)-ə görə,  $D_1$  toplusunun LX zamanı və  $(D_1 + D_2)$  toplusunun LX zamanı  $r_c$  dəyəri aşağıdakı kimi hesablanır:

$$r_c(D_1) = \frac{\sum_{i=1}^{I_1} t_i}{I_1}. \quad (19)$$

$$r_c(D_1 + D_2) = \frac{\sum_{i=1}^{I_1} t_i + \sum_{i=I_1+1}^{I_1+I_2} t_i}{I_1 + I_2}. \quad (20)$$

(19) və (20)-dən göründüyü kimi,  $C$  keyfiyyət ölçüsü inkremental şəkildə aşağıdakı kimi hesablanıla bilər:

1.  $D_1$  toplusunun LX zamanı,  $r_c$  əsasında hesablanmış  $C$  linqvistik xülasəyə  $\sum_{i=1}^{I_1} t_i$  və  $I_1$  dəyərləri ilə birlikdə əlavə olunur.
2.  $(D_1 + D_2)$  toplusunun ardıcıl LX zamanı,  $\sum_{i=I_1+1}^{I_1+I_2} t_i$  dəyəri hesablanır, və  $C$  keyfiyyət ölçüsü hesablanmış  $\sum_{i=I_1+1}^{I_1+I_2} t_i$  &  $I_2$  və yadda saxlanmış  $\sum_{i=1}^{I_1} t_i$  &  $I_1$  dəyərlərinin əsasında hesablanır.

2.8-ci alt bölmədə, BƏS-in dizaynına LX nəticələrinin tətbiqi prosesi nəzərdən keçirilir və Mamdani tipli QSÇS-nin qaydalar bazasına ƏGƏR–ONDA tipli qaydaların seçilmə proseduru təklif olunur.

Baxılan alt bölmədə, BƏS-in qaydalar bazasına etibarlılığı bir kritik dəyəərə nisbətən daha çox və ya bərabər olan qaydaların seçilməsi təklif olunur:  $R \geq R_{cv}$ , burada  $R_{cv}$  etibarlılıq dərəcəsinin kritik dəyəridir.

$R$ -in kritik dəyəərini hesablayan bir prosedurun işlənmə məsələsi dissertasiyada uğurla həll olunmamışdır: [2]-də təklif olunan yeganə həll qeyri-kafi olduğundan bəri göstərilməyib. Beləliklə, dissertasiyada  $R$ -in kritik dəyəəri metodikanın istifadəçisi tərəfindən müəyyən edilir.

Qeyd etmək lazımdır ki,  $R \geq R_{cv}$  ilə bəzi ziddiyyətli qaydalar varsa, qaydalar bazasına yalnız ən yüksək  $R$  dəyəərinə malik qayda seçilir. Üstəlik, qaydalar bazasına seçilmiş qaydalar  $R$  dəyəərlərinə bərabər olaraq təyin olunmuş çəkilərə malikdirlər ki, bu da az etibarlı qaydaların QSÇS-nin fəaliyyətinə çox etibarlı qaydalardan daha az təsir göstərməsinə səbəb olur.

*2.9-cu alt bölmədə*, LX nəticələrinin tətbiqi ilə tərtib edilmiş BƏS-ə tətbiq olunan bir optimallaşdırma prosesi nəzərdən keçirilir: baxılan optimallaşdırma prosesi proqnozlaşdırma məsələlərdə dizayn edilmiş BƏS-in performansını (dəqiqliyini) yaxşılaşdırmalıdır.

Baxılan alt bölmədə, BƏS-in girişlərini təsvir edən linqvistik termlər optimallaşdırma obyektini kimi təklif olunur. Linqvistik termlərin parametrləri optimallaşdırma zamanı tənzimlənən parametrlərdir (məsələn, üçbucaqlı mənsubiyyət funksiyalarından istifadə olunursa, hər linqvistik term üç tənzimlənən parametr ilə təsvir olunur). Optimallaşdırma genetik alqoritm vasitəsi ilə həyata keçirilir və optimallaşdırmada tətbiq olunan fitness funksiyası kimi ilkin verilənlər toplusunun bir alt hissəsində BƏS-in nəticələrinin dəqiqliyini təyin edən OKX (orta kvadratik xəta) seçilib.

**Üçüncü bölmə** biliklərə əsaslanan sistemlərin dizaynı üçün təklif olunan metodikanın eksperimental verifikasiyasını təsvir edir. Eksperimental verifikasiya MATLAB (R2016a və R2017b) proqram təminatı vasitəsi ilə həyata keçirilir:

- Bölmədə təsvir olunan hər bir verifikasiya proseduru üçün müvafiq proqram skriptləri hazırlanıb.



- Bölmədə təsvir olunan verifikasiya prosedurlarının əksəriyyəti üçün müvafiq proqram tulboksları hazırlanıb: bu tulbokslar müvafiq istifadəçi təlimatları ilə tamamlanıb və MATLAB icmasının rəsmi veb saytında təqdim olunub<sup>4</sup>.

Təklif olunan metodikanın eksperimental verifikasiyasını həyata keçirmək üçün aşağıdakı avadanlıqdan istifadə olunur: MS Windows 8.1 əməliyyat sistemi əsasında işləyən və eyni konfigurasiyaya malik iki Dell Latitude E6420 noutbuku. Qeyd etmək lazımdır ki, “*Wu–Mendel yanaşmasının paralelləşməsi*” hesablama dəyərini azaltmaq üçün praktik həlldən başqa, bölmənin qalan bütün alt bölmələrində noutbuklardan yalnız biri istifadə olunur.

Aparılan eksperimental verifikasiya BƏS-in dizaynında aşağıdakı ilkin verilənlər toplularını tətbiq edir:

- “*Combined Cycle Power Plant*” – buxar-qaz turbinli stansiyanın statistik parametrlərini təsvir edir (cədvəl 5).
- “*Banknote Authentication*” – saxta və orijinal banknotların statistik parametrlərini təsvir edir (cədvəl 6).
- “*Skin Segmentation*” – insan dərisinin rəngini təsvir edən/etməyən RGB tipli piksellərdən ibarətdir (cədvəl 7).
- “*HTRU 2*” – pulsarların (döyünən radio ulduzların) statistik parametrlərini təsvir edir (cədvəl 8).

**Cədvəl 5**

**“Combined Cycle Power Plant” toplusunun atributları**

<b>№</b>	<b>Ad</b>	<b>Tip</b>	<b>Dəyişmə diapazonu</b>
1	AT	ant.	[1.81, 37.11]
2	V	ant.	[25.36, 81.56]
3	AP	ant.	[992.89, 1033.3]
4	RH	ant.	[25.56, 100.16]
5	EP	kon.	[420.26, 495.76]

<sup>4</sup> Mirzaxanov, V. MATLAB Central [Elektron resurs] / 2019. URL: <https://www.mathworks.com/matlabcentral/fileexchange/?term=authorid%3A283483>.

**Cədvəl 6****“Banknote Authentication” toplusunun atributları**

<b>№</b>	<b>Ad</b>	<b>Tip</b>	<b>Dəyişmə diapazonu</b>
1	Variance	ant.	[-7.0421, 6.8248]
2	Skewness	ant.	[-13.7731, 12.9516]
3	Kurtosis	ant.	[-5.2861, 17.9274]
4	Entropy	ant.	[-8.5482, 2.4495]
5	Class	kon.	{0, 1}

**Cədvəl 7****“Skin Segmentation” toplusunun atributları**

<b>№</b>	<b>Ad</b>	<b>Tip</b>	<b>Dəyişmə diapazonu</b>
1	Red	ant.	[0, 255]
2	Green	ant.	[0, 255]
3	Blue	ant.	[0, 255]
4	Class	kon.	{1, 2}

**Cədvəl 8****“HTRU 2” toplusunun atributları**

<b>№</b>	<b>Ad</b>	<b>Tip</b>	<b>Dəyişmə diapazonu</b>
1	MP	ant.	[5.8125, 192.6172]
2	DP	ant.	[24.7720, 98.7789]
3	KP	ant.	[-1.8760, 8.0695]
4	SP	ant.	[-1.7919, 68.1016]
5	MC	ant.	[0.2132, 223.3921]
6	DC	ant.	[7.3704, 110.6422]
7	KC	ant.	[-3.1393, 34.5398]
8	SC	ant.	≈ [-1.9770, 1191]
9	Class	kon.	{0, 1}

Hər bir eksperimental verifikasiya proseduru dissertasiyada öz alt bölməsini təşkil edir və aşağıda qısaca təsvir olunub:

- 3.2-ci alt bölmədə, təklif olunan avtomatlaşdırılmış fazifikasiya tənzimləməsi üsulunun eksperimental verifikasiyası aparılır ki, bu da üsulun keyfiyyət baxımından əhəmiyyətli bir itki olmadan linqvistik xülasənin əhəmiyyətli dərəcədə sadələşdirməsi qabiliyyətini göstərir.

- 3.3-cü alt bölmədə, Wu–Mendel yanaşmasının istifadəçi tərəfindən müəyyən etdiyi parametrlərdən asılılığını azaltmaq üçün təklif olunan prosedurun eksperimental verifikasiyası aparılır ki, bu da  $r_c^{min}$  və  $r_c^{max}$  dəyərlərinin hesablanması üçün təklif olunan düsturların adekvatlığını sübut edir.
- 3.4-cü alt bölmədə,  $C$  və  $R$  keyfiyyət ölçülərinin hesablanması üçün təklif olunan düsturların eksperimental verifikasiyası aparılır ki, bu da onların orijinal Wu–Mendel yanaşmasında istifadə olunan düsturlara nisbətən daha yüksək keyfiyyət təmin etdiyini göstərir.
- 3.5-ci alt bölmədə, Wu–Mendel yanaşmasının hesablama dəyərini azaltmaq üçün təklif olunan həllərin eksperimental verifikasiyası aparılır ki, bu da təklif olunan həllərdən hər hansı birini tətbiq edərkən İM-nin əhəmiyyətli dərəcədə azaldığını göstərir. Əlavə, alt bölmədə bəzi həllərin birlikdə tətbiq olunma qabiliyyəti göstərilir.
- 3.6-cı alt bölmədə, LX nəticələrinin dinamik İVT-na uyğunlaşmasını təmin edən təklif olunan alqoritmin eksperimental verifikasiyası həyata keçirilmişdir ki, bu da modifikasiya olunmuş Wu–Mendel yanaşmasının dinamik verilənlər toplusunun linqvistik xülasəsini inkremental şəkildə yeniləmə qabiliyyətini nümayiş etdirir.
- 3.7-ci alt bölmədə, BƏS-ə tətbiq olunan optimallaşdırma prosedurun eksperimental verifikasiyası aparılır ki, bu da ilkin verilənlərin antedent tipli atributlarını təsvir edən linqvistik termlərin genetik tənzimlənməsi nəticəsində BƏS-in proqnozlaşdırma dəqiqliyinin artdığını göstərir.

## ƏSAS NƏTİCƏLƏR

Dissertasiyada aşağıdakı tədqiqatlar aparılıb və nəticələr əldə edilib:

1. Yanaşmanın BƏS-in dizaynında bir neçə tətbiqetmə məsələsini və problemlərini ortaya qoyan Wu–Mendel yanaşmasının kompleks təhlili aparılıb.

2. İlk verilənlər toplusunun avtomatlaşdırılmış fazifikasiya tənzimləməsi üsulu sintez olunub ki, bu da alınan linqvistik xülasənin keyfiyyətində mütənasib əhəmiyyətli itkisiz ilkin verilənlər atributlarının linqvistik təsvirini sadələşdirməyə imkan verir.
3. Wu–Mendel yanaşmasında ƏGƏR–ONDA tipli qaydaların keyfiyyət ölçülərinin hesablanmasına yenidən baxılaraq, yanaşmanın istifadəçi tərəfindən təyin olunan parametrlərdən asılılığı aşağı salınıb.
4. Wu–Mendel yanaşmasının hesablama dəyərinin kompleks analizi aparılıb ki, bu da Wu–Mendel yanaşmasını əhəmiyyətli dərəcədə sürətləndirən bir neçə həllin təklif edilməsi ilə nəticələnib.
5. Wu–Mendel yanaşmasında ƏGƏR–ONDA tipli qaydaların bəzi keyfiyyət ölçülərini hesablamaq üçün yeni düsturlar təklif olunub ki, bu da nəticədə alınan linqvistik xülasənin keyfiyyətini əhəmiyyətli dərəcədə artırır.
6. Wu–Mendel yanaşmasını inkrementasiya xüsusiyyəti ilə tamamlayan bir linqvistik xülasənin dinamik ilkin verilənlər toplusuna uyğunlaşmasını təmin edən alqoritm təklif olunub: alqoritm Wu–Mendel yanaşmasında qaydaların keyfiyyət ölçülərinə qarşı tətbiq olunan inkremental şəkildə yenilənmə prosedurundan ibarətdir.
7. Wu–Mendel yanaşmasının tətbiqi ilə dizayn olunmuş BƏS-in dəqiqliyini artırmaq üçün genetik optimallaşdırma proseduru təklif olunub.

Qeyd etmək lazımdır ki, yuxarıda göstərilən məqamlardan əlavə, dissertasiyada aparılan araşdırmalar hələlik həll olunmamış qalan bəzi məsələləri də üzə çıxartdı:

- Biliklərə əsaslanan sistemlərin yaradılması üçün işlənən metodikanın vacib addımlarından biri dizayn olunan BƏS-in qaydalar bazasına ƏGƏR–ONDA tipli qaydaların seçilməsidir. Dissertasiyada qayda seçiminin etibarlılıq dərəcəsi ( $R$ ) əsasında həyata keçirildiyi bildirilir; lakin  $R$ -in kritik (sərhəd) dəyərini hesablamaq üçün müvafiq effektiv prosedur sintez olunmayıb.

- Dissertasiyada bir çox prosedurlar və həllər təklif olunub ki, əksəriyyətinin bir-birinə uyğun olmasına baxmayaraq, bəzilərinin birləşməsi əlavə tədqiqatlar tələb edir.
- Biliklərə əsaslanan sistemlərin yaradılması üçün işlənilən metodika Data Mining və BƏS sahələrində dəyərini sübut edə biləcək kompleks praktik yoxlamaları tələb edir.
  - Biliklərə əsaslanan sistemlərin yaradılması üçün işlənilən metodika digər müasir yanaşma və metodikalar ilə müqayisə olunmalıdır.
  - Təklif olunan metodikanın insan fəaliyyətinin müxtəlif sahələrində tətbiq olunmasına dair tədqiqatlar aparılmalıdır ki, tədqiqatlar sadəcə metodikanın praktik tətbiqetmə imkanını (dissertasiyada olduğu kimi) yox, həm də tətbiq olunan sahələrdə metodikanın üstünlüyə malik olub/olmamasını müəyyənləşdirməlidir. Bu cür araşdırmalar təklif olunan metodikanın bəzi tanınan yanaşma və metodikalarla yanaşı tətbiq edilməsini də motivasiya edəcəkdir.

**Dissertasiyanın əsas nəticələri aşağıdakı elmi əsərlərdə əks olunmuşdur:**

1. Мирзаханов, В. Краткий обзор современных средств Data Mining // Автоматизация. Современные технологии, – 2018. № 7, – с. 330-336.
2. Mirzakhanov, V., Gardashova, L. Application of linguistic summarization in banknote authentication // Herald of the Azerbaijan Engineering Academy, – 2018. № 3, – p. 118-124.
3. Mirzakhanov, V., Gardashova, L. The Incrementality Issue in the Wu-Mendel Approach for Linguistic Summarization Using IF-THEN Rules // 13th International Conference on Theory and Application of Fuzzy Systems and Soft Computing — ICAFS 2018, – Cham: Springer, – 2019, – p. 293-300.
4. Мирзаханов, В. Синтез графической системы распознавания участков человеческой кожи посредством

лингвистического реферирования // Автоматизация. Современные технологии, – 2019. № 1, – с. 37-41.

5. Mirzakhonov, V., Gardashova, L. Modification of the Wu-Mendel approach for linguistic summarization // Journal of Experimental & Theoretical Artificial Intelligence (JETAI), – 2019. № 1, – p. 77-97.
6. Mirzakhonov, V. The fuzzification issue in the Wu–Mendel approach for linguistic summarisation using IF-THEN rules // JETAI, – 2019. № 1, – p. 117-136.
7. Mirzakhonov, V., Gardashova, L. Wu–Mendel Approach for Linguistic Summarization: Practical Considerations and Solutions // Proceedings of the 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), – New Orleans: IEEE, – 2019, – p. 1-8.
8. Mirzakhonov, V. Value of fuzzy logic for data mining and machine learning: A case study // Expert Systems with Applications, – 2020. – p. 1-11.

Həmmüəlliflərlə nəşr olunmuş əsərlərdə iddiaçının şəxsi rolu: [2], [3], [5], [7] – nəzəri tədqiqatlar, program təminatının hazırlanması/tətbiqi və eksperimental verifikasiya, tədqiqatların elmi məqalələr şəklində təqdimatı.

Dissertasiyanın müdafiəsi 29 04 2022 il tarixində saat 13:00 Azərbaycan Dövlət Neft və Sənaye Universitetinin nəzdində fəaliyyət göstərən ED 2.02 Dissertasiya şurasının iclasında keçiriləcək.

Ünvan: AZ1010, Bakı, Azadlıq prospekti, 20. Azərbaycan Dövlət Neft və Sənaye Universiteti, əsas bina, otaq 250.

Dissertasiya ilə Azərbaycan Dövlət Neft və Sənaye Universitetinin kitabxanasında tanış olmaq mümkündür.

Dissertasiya və avtoreferatın elektron versiyaları Azərbaycan Dövlət Neft və Sənaye Universitetinin rəsmi internet saytında yerləşdirilmişdir.

Avtoreferat 15 03 2022 il tarixində zəruri ünvanlara göndərilmişdir.

Çapa imzalanıb: 10.03.2022

Kağızın formatı: A5

Həcm: 44465

Tiraj: 70