

Hepatosellulyar Karsinoma üzrə Milli Bazanın MICE Metodu Əsasında Təkmilləşdirilməsi ilə Proqnozlaşdırma Məsələsinin Həlli

Məsumə Məmmədova¹, Zərifə Cəbraylova², Lalə Qarayeva³, Mehriban Hüseynova⁴

^{1,2,3}ETM İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan

⁴Azərbaycan Tibb Universiteti, Bakı, Azərbaycan

¹mmg51@mail.ru, ²djabrailova_z@mail.ru, ³karayevalala.01@gmail.com, ⁴mehribanhuseynova9@gmail.com

Xülasə— Məqalədə hepatosellulyar karsinoma (HSK) üzrə milli verilənlər bazasındakı boşluqların doldurulması və maşın təlimi alqoritmləri əsasında proqnozlaşdırma məsələsinin həlli təklif edilmişdir. Proqnozlaşdırma məsələlərinin maşın təlimi alqoritmləri ilə həlli bazanın təkmilləşdirilməsini və mövcud boşluqların doldurulmasını tələb edir. Tədqiqatda MICE metodu əsasında HSK üzrə milli bazadakı boşluqların doldurulması mərhələləri təqdim edilmişdir. Milli bazadakı məlumatlar əsasında kliniki xəstələrin vəziyyəti ilə bağlı proqnoz nəticələrin alınması üçün onun hissə-hissə Kaggle platformasından götürülmüş, proqnoz nəticələri (“ölüm-0” və ya “sağqalma-1”) məlum olan *HCC datasetinə* birləşdirilməsi və genişləndirilmiş bazada *Random Forest* və *XGBClassification* maşın təlimi alqoritmlərinin tətbiqi verilmişdir.

Açar sözlər— *hepatosellulyar karsinomanın proqnozlaşdırılması; maşın təlimi alqoritmləri; verilənlər bazasındakı boşluqların doldurulması; MICE metodu.*

I. GİRİŞ

Təcrübəli həkim-ekspertlərin biliklərinin toplanması, saxlanması, manipulyasiyası, eləcə də, hər bir konkret verilənlər toplusu üzrə xəstəliyin müəyyən edilməsi və adekvat qərarların qəbul olunması üçün daha səmərəli vasitə biliklərə əsaslanan intellektual sistemlərdir. Bu sistemlərin əsasını tibbin konkret predmet sahəsində olan xəstəliklər, onların mümkün səbəbləri, inkişaf müddəti, kliniki təzahürləri, müşahidə olunan əlamətləri, simptomları və s. təşkil edir. Bu sistemlər diaqnoz qoyulması, daha effektiv müalicə üsulunun seçilməsi, proqnozlaşdırma, uyğun vəziyyətlərin (presedentlərin) axtarışı, terapiyaya nəzarət, təsvirlərin tanınması və interpretasiyası, dərman vasitələrinin kliniki-farmakoloji xüsusiyyətlərinin (toksikliyinin) monitorinqi və s. məsələlərin həllində tətbiq olunur.

Xərçəng növlərindən biri olan HSK çoxlu sayda kliniki göstəricilər, əlamətlərlə təyin olunan kritiki vəziyyətlərlə özünü biruzə verir. Onun diaqnozu və proqnozlaşdırılması üçün dəqiq, birmənalı meyarlar mövcud deyil [1, 2]. HSK-nın kritiki vəziyyətlərinin müxtəlif tipli, strukturlaşdırılmamış, çoxsaylı göstəricilərlə xarakterizə olunması HSK-nın diaqnozu, proqnozlaşdırılması ilə bağlı qərarların qəbulunda həkim səhvlərinə səbəb olur [3, 4]. Nəticədə göstəriciləri xarakterizə edən kliniki əlamətlərin müəyyən kombinasiyaları ilə təyin

edilən həkim qərarlarının qəbulunda xətalər qaçılmaz olur. Bu HSK-nın diaqnostikası və proqnozlaşdırılması üçün süni intellekt metodlarının tətbiqini, kliniki qərarların qəbuluna dəstək sistemlərinin yaranmasını şərtləndirir. Hazırda elmi ədəbiyyatda qaraciyər xəstəliklərinin aşkarlanması, diaqnostikası və müalicəsi üçün intellektual sistemlərin işlənilməsi istiqamətində tədqiqatlara rast gəlinir [5-8].

Hazırkı məqalədə maşın təlimi alqoritmlərinə istinad etməklə milli baza əsasında HSK-nın ilkin diaqnozu və proqnozlaşdırılması məsələsinin həlli imkanları araşdırılmışdır. Qeyd edək ki, maşın təlimi alqoritmləri əsasında verilənlərin analitik təhlili tam verilənlər toplusunun olmasını tələb edir, əsaslandığımız HSK üzrə milli bazada bir sıra boşluqların olması bu baxımdan problemlər yaradır. Adətən ənənəvi proqnozlaşdırmada bazadakı boşluqların doldurulması üçün uyğun sahə üzrə mütəxəssislərin rəylərinə istinad edilir və onlar bazadakı verilənlərin necə toplandığını yaxşı bildikləri üçün müvafiq düzəlişləri edə bilirlər. Lakin son dövrlərdə bazadakı boşluqların doldurulması üçün ekspert-mütəxəssislərin istinad etdiyi müxtəlif mürəkkəb prinsiplər əsasında yaradılmış alqoritmlərdən istifadə olunur [5, 9, 10]. Beləliklə, hazırkı məqalədə HSK üzrə milli bazadakı boşluqların doldurulması ilə bazada olan kliniki xəstələrin vəziyyətinə uyğun proqnoz nəticələrin alınması alqoritm təklif edilmişdir.

II. PROBLEMİN QOYULUŞU

HSK üzrə milli baza əsasında maşın təlimi alqoritmlərinin tətbiqi ilə proqnozlaşdırma məsələsinin həlli bu bazadakı çoxsaylı boşluqların olması, proqnoz nəticələrin göstərilməməsi kimi problemlərin aradan qaldırılmasını şərtləndirir. Beləliklə, HSK üzrə milli verilənlər bazasındakı kliniki xəstələrin vəziyyəti ilə bağlı proqnoz nəticələrin alınması üçün aşağıdakı məsələlərin həlli nəzərdə tutulur:

- HSK üzrə milli verilənlər bazasında boşluqların doldurulması;

- HSK üzrə milli verilənlər bazasına daxil olan kliniki xəstələrin vəziyyətinə uyğun proqnoz nəticələrin alınması üçün daha mükəmməl bazanın seçilməsi və verilənlərin ilkin emalı;

- seçilmiş HSK üzrə mükəmməl bazada proqnozlaşdırma məsələsinin maşın təlimi alqoritmlərinin tətbiqi ilə həlli və

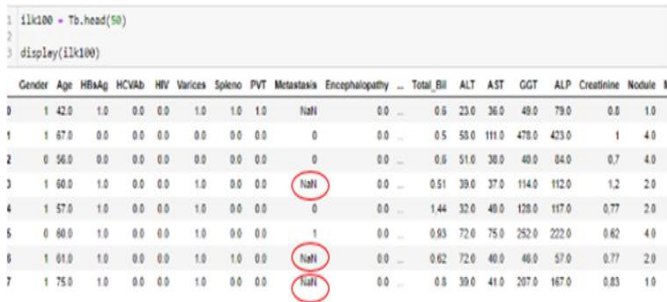
dəqiqlir meyarlarına görə daha yaxşı nəticə göstərən metodun seçilməsi;

- milli bazanın hissə-hissə daha mükəmməl bazaya köçürülməsi və alınan genişləndirilmiş bazada seçilmiş maşın təlimi algoritminin tətbiqi ilə kliniki xəstələrin vəziyyəti ilə bağlı proqnoz nəticələrin alınması.

III. PROBLEMİN HƏLLİ

a. HSK üzrə milli verilənlər bazasında boşluqların doldurulması

Azərbaycan Tibb Universitetinin I cərrahi xəstəliklər kafedrasının Türkiyənin Malatya İnönü Universitetinin Qaraciyər Nəqli İnstitutu ilə birgə fəaliyyəti əsasında yaradılmış HSK üzrə kliniki xəstələr haqqında verilənlər bazası (milli HCC Dataset) tədqiqat mənbəyi seçilmişdir. Verilənlər bazası 27 göstərici üzrə 556 kliniki xəstənin məlumatları əsasında formalaşdırılmışdır (şəkil 1) (bazadakı boşluqlar şəkil 1-də qırmızı halqalarla göstərilmişdir).



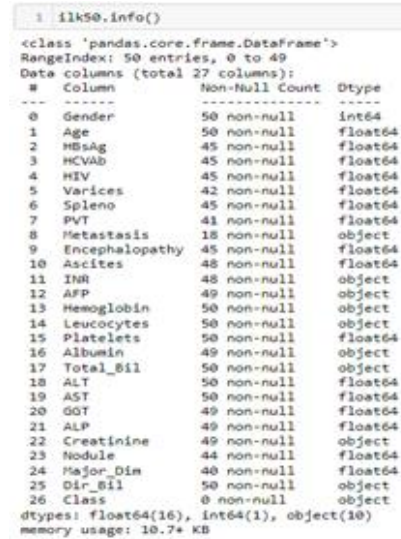
	Gender	Age	HBsAg	HCVAb	HIV	Varices	Spleno	PVT	Metastasis	Encephalopathy	Total_Bil	ALT	AST	GGT	ALP	Creatinine	Nodule	Class
0	1	42.0	1.0	0.0	0.0	1.0	1.0	1.0	NaN	0.0	0.0	23.0	36.0	48.0	79.0	0.8	1.0	0
1	1	67.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0.0	0.0	5.0	58.0	111.0	478.0	423.0	1	4.0
2	0	56.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0.0	0.0	0.0	51.0	30.0	48.0	84.0	0.7	4.0
3	1	60.0	1.0	0.0	0.0	1.0	0.0	0.0	NaN	0.0	0.0	0.51	39.0	37.0	114.0	112.0	1.2	2.0
4	1	57.0	1.0	0.0	0.0	1.0	0.0	0.0	0	0.0	0.0	1.44	32.0	40.0	128.0	117.0	0.77	2.0
5	0	60.0	1.0	0.0	0.0	1.0	0.0	0.0	1	0.0	0.0	0.93	72.0	75.0	252.0	222.0	0.82	4.0
6	1	61.0	1.0	0.0	0.0	1.0	1.0	0.0	NaN	0.0	0.0	0.62	72.0	40.0	46.0	57.0	0.77	2.0
7	1	75.0	1.0	0.0	0.0	1.0	0.0	0.0	NaN	0.0	0.0	0.0	39.0	41.0	207.0	167.0	0.83	1.0

Şəkil 1. Milli HCC Dataset-inin fragmenti.

Qeyd edək ki, bazadakı 556 nümunədən (row) yalnız 50 nümunədə proqnoz nəticələr (0 (ölüm) və 1 (yaşayan)) göstərilmişdir. Bu baza əsasında kliniki xəstələrin vəziyyətinin proqnozlaşdırılması (yəni, hər bir nümunə üzrə proqnoz nəticənin alınması) üçün ilk növbədə istifadəçinin müdaxiləsi ilə bazadakı müxtəlif tipli verilənlər eyni tipə (float) gətirilməli, bazadakı boşluqlar doldurulmalıdır.

Boşluqların doldurulması üçün istifadə edilən metodların təhlili MICE (ing. *Multivariate Imputation by Chained Equations*) algoritminin seçilməsi ilə tədqiqatın aparılmasına imkan yaratmışdır [11, 12]. Bu algoritim ilə boşluqlarda olacaq dəyişənlər üçün posterior ehtimal paylamalarını əldə etmək məqsədilə Markov sxemi üzrə Monte-Karlo metodu istifadə edilir. Bu metod verilənlər bazasında çoxlu sayda dəyişənlər üzrə boşluqlar olduğu halda istifadə üçün yararlıdır. İlk növbədə buradakı boşluqların doldurulmasını tələb edir və boşluqların doldurulması mərhələləri aşağıdakı addımlarla yerinə yetirilir.

I mərhələ. Baxılan verilənlər bazasında bir-biri ilə əlaqəli olmayan müxtəlif tipli (object, int, float) verilənlər müəyyənləşdirilir (şəkil 2).



#	Column	Non-Null Count	Dtype
0	Gender	50 non-null	int64
1	Age	50 non-null	float64
2	HBsAg	45 non-null	float64
3	HCVAb	45 non-null	float64
4	HIV	45 non-null	float64
5	Varices	42 non-null	float64
6	Spleno	45 non-null	float64
7	PVT	41 non-null	float64
8	Metastasis	18 non-null	object
9	Encephalopathy	45 non-null	float64
10	Ascites	48 non-null	float64
11	INR	48 non-null	object
12	AFP	49 non-null	object
13	Hemoglobin	50 non-null	object
14	Leucocytes	50 non-null	object
15	Platelets	50 non-null	float64
16	Albumin	49 non-null	object
17	Total_Bil	50 non-null	object
18	ALT	50 non-null	float64
19	AST	50 non-null	float64
20	GGT	49 non-null	float64
21	ALP	49 non-null	float64
22	Creatinine	49 non-null	object
23	Nodule	44 non-null	float64
24	Major_Dim	40 non-null	float64
25	Dir_Bil	50 non-null	object
26	Class	0 non-null	object

Şəkil 2. Milli HCC Datasetində olan verilənlər və onların tipləri.

Verilənlərin vahid formaya gətirilməsi üçün onların ilkin emalı həyata keçirilir, istifadəçinin müdaxiləsi ilə bazadakı müxtəlif tipli verilənlər eyni tipə (float) gətirilir.

II mərhələ. MICE algoritmi ilə boşluqların doldurulması üçün boşluqların olduğu sütunların adları kodlaşdırılaraq proqrama daxil edilmir.

Kodlaşdırma əsasında MICE metodunun həll algoritminə uyğun olaraq boşluqların olduğu sütunlar, məsələn “Varices” sütununun aşağıdakı modeli qurulur:

$$Varices_g \approx f(Age_g, Gender_g, HBsAg_g, HCVAb_g, HIV_g, \dots). (1)$$

Növbəti mərhələdə bu sütundakı boşluqlara doldurulacaq qiymətlərin təxminlərinə görə sütunun aşağıdakı yeni modeli qurulur:

$$\overline{Varices}_g = \hat{f}(Age_g, Gender_g, HBsAg_g, HCVAb_g, HIV_g, \dots). (2)$$

“Varices” sütununda təxmin edilən qiymətlər milli bazadakı müvafiq çatışmayan qiymətlərlə əvəz olunur və əməliyyat növbəti çatışmayan veriləni olan sütun üçün təkrarlanır. Şəkil 3-dən göründüyü kimi, bu milli bazadakı “PVT” sütunudur və onun modeli aşağıdakı kimidir:

Bu sütundakı boşluqlara doldurulacaq qiymətlərin təxminlərinə görə sütunun aşağıdakı modeli qurulur:

$$PVT_g \approx f(Age_g, Gender_g, HBsAg_g, HCVAb_g, HIV_g, \dots). (3)$$

Beləliklə, hər bir çatışmayan veriləni olan sütun üçün eyni prinsip tətbiq olunur və hər iterasiyada, əvvəlki iterasiyadan əldə olunan nəticələr nəzərə alınır.

$$\overline{PVT}_g = \hat{f}(Age_g, Gender_g, HBsAg_g, HCVAb_g, HIV_g, \dots). (4)$$

III mərhələ. MICE algoritminin tətbiqi ilə milli HCC Datasetdəki boşluqların doldurulmasından sonrakı təsviri şəkil 3-də verilmişdir.

	Gender	Age	HBsAg	HCVAb	HIV	Varices	Spleno	PVT	Metastasis	Encephalopathy	...	Total_Bil	ALT	AST	GGT	ALP	Creatinine	Nodu
0	1	42.0	1.0	0.0	0.0	1.000000	1.0	1.000000	0.271354	0.000000	...	0.60	23.0	36.0	49.0	79.0	0.800000	1.000000
1	1	67.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	...	0.50	58.0	111.0	478.0	423.0	1.000000	4.000000
2	0	56.0	0.0	0.0	0.0	0.000000	0.0	0.000000	0.000000	0.000000	...	0.60	51.0	36.0	40.0	84.0	0.700000	4.000000
3	1	60.0	1.0	0.0	0.0	1.000000	0.0	0.000000	0.283040	0.000000	...	0.51	39.0	37.0	114.0	112.0	1.200000	2.000000
4	1	57.0	1.0	0.0	0.0	1.000000	0.0	0.000000	0.000000	0.000000	...	1.44	32.0	49.0	128.0	117.0	0.770000	2.000000
5	0	60.0	1.0	0.0	0.0	1.000000	0.0	0.000000	1.000000	0.000000	...	0.93	72.0	75.0	252.0	222.0	0.620000	4.000000
6	1	61.0	1.0	0.0	0.0	1.000000	1.0	0.000000	0.267000	0.000000	...	0.62	72.0	40.0	46.0	57.0	0.770000	2.000000
7	1	75.0	1.0	0.0	0.0	1.000000	0.0	0.000000	0.438200	0.000000	...	0.80	39.0	41.0	207.0	167.0	0.830000	1.000000

Şəkil 3. MICE algoritminin tətbiqi ilə boşluqların doldurulmasından sonra milli HCC Dataset-indən bir fraqmentin təsviri.

MICE metodunun tətbiqi ilə HCC Datasetdəki boşluqların doldurulması bir daha bu metodun paraktiki məsələlərin həllində reallaşdırılması imkanını göstərir. Baxılan halda alınan nəticənin effektivliyi HSK üzrə milli bazadakı boşluqların aradan qaldırılmasına nail olunmasıdır ki, bu da maşın təlimi alqoritmləri əsasında proqnozlaşdırma məsələsinin həllinə imkan yaradır.

B. HSK üzrə milli verilənlər bazasına daxil olan kliniki xəstələrin vəziyyətinə uyğun pronoz nəticələrin alınması üçün daha mükəmməl bazanın seçilməsi və verilənlərin ilkin emalı

HSK-nın proqnozlaşdırılmasında maşın təlimi alqoritmlərinin tətbiqi üçün Kaggle şirkətinin HCC Dataset [13] adlı açıq verilənlər bazasından istifadə olunmuşdur. Verilənlər bazası Portuqaliya Universiteti Xəstəxanasında HCC-dən əziyyət çəkən 165 klinik xəstənin məlumatları əsasında formalaşdırılmışdır. Verilənlər bazasında Qaraciyərin Tədqiqi üzrə Avropa Assosiasiyası – Xərçəngin Tədqiqi və Müalicəsi üzrə Avropa Təşkilatı (eng. European Association for the Study of the Liver - European Organisation for Research and Treatment of Cancer) tərəfindən tövsiyə edilən 49 xüsusiyyət (kliniki əlamət) yer almışdır.

Bu baza əsasında maşın təlimi alqoritmlərinin tətbiqi ilə HSK-nın proqnozlaşdırılması üçün verilənlərin ilkin emalı yerinə yetirilir. Bu mərhələdə bazada bir-biri ilə əlaqəli olmayan verilənlərin olub-olmaması araşdırılır, səpələnmiş verilənlər müəyyən edilir, onların vahid formaya gətirilməsi həyata keçirilir. Bu mərhələ verilənlər bazasının təmizlənməsi prosesi də adlandırılır və bilavasitə istifadəçinin müdaxiləsi ilə yerinə yetirilir. Verilənlər bazasını faydalı etmək, verilənin işlənməsini sadələşdirmək üçün Pandas [14] və NumPy [15] kitabxanalarından istifadə edilir. Kodların tətbiqi ilə müxtəlif tip verilənlər eyni tip verilənlərə çevrilir və bir-biri ilə əlaqəli xüsusiyyətlər təyin edilir. Verilənlər arasında həm xətti, həm də qeyri-xətti əlaqələri tapmaq üçün Correlation heatmaps istifadə edilir. Məlumatları normallaşdırmaq üçün minimum-maksimum miqyaslama bir və ya bir neçə xüsusiyyət sütununa tətbiq olunur. Minimum-maksimum miqyaslama əsasında verilənlərin normallaşdırılmaq üçün şəkil 4-dəki kodlaşdırmadan istifadə olunur. Daha sonra, HCC Dataset-də olan verilənlərin bütün xüsusiyyətləri təhlil olunmuş, buradakı 49 kliniki əlamət/atributdan 23-ünün kəmiyyət, 26-sının keyfiyyət xarakterli olduğu müəyyən edilmişdir. Hədəf sinfi (class) – birillik hesabata görə, 0 (ölən) və 1 (yaşayan) kimi ikili dəyişəndən ibarət olan proqnoz nəticədir.

C. HSK-nın proqnozlaşdırılması üçün maşın təlimi alqoritmlərinin tətbiqi və dəqiqlik meyarlarına görə daha yaxşı nəticə göstərən metodun müəyyənəşdirilməsi

HCC Dataset-də təsnifləndirməni aparmaq üçün Logistic Regression (LR), Support Vector Machine (SVM) və Random Forest (RF) maşın təlimi alqoritmləri istifadə olunmuşdur. Maşın təlimində klassifikatorların aşkarlama performansının qiymətləndirilməsi üçün həssaslıq, tamlıq, yanlış pozitiv hallar, doğru pozitiv hallar, f-ölçü, dəqiqlik meyarlarından istifadə olunmuşdur [16, 17].

```
1 std = MinMaxScaler()
2 X = pd.DataFrame(std.fit_transform(X) , columns=X.columns)
3 X
```

	Gender	Symptoms	Alcohol	HBsAg	HBsAb	HCVAb	Cirrhosis	Endemic	Smoking	...	
0	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	...	
1	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	1.0	...	
2	1.0	0.0	1.0	1.0	0.0	1.0	0.0	1.0	0.0	1.0	...
3	1.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	...	
4	1.0	1.0	1.0	1.0	0.0	1.0	0.0	1.0	0.0	1.0	...
...	...	...	...	...	...	...	...	...	...	...	
199	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	...	
200	0.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	...	
201	1.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	...	
202	1.0	1.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	...	
203	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	

204 rows x 49 columns

Şəkil 4. Qaraciyər xərçəngi verilənlər bazasının MinMaxScaler əsası miqyaslanması.

Həssaslıq (P) həqiqi müsbətlərin sayı kimi müəyyən edilir və aşağıdakı kimi təyin edilir:

$$P = \frac{Tp}{(Tp + Fp)} \quad (5)$$

Burada: Tp – doğru təsnif edilmiş proqnozlaşdırma ilə əlaqəli verilənlərin sayı, Fp – səhv təsnif edilmiş proqnozlaşdırma ilə əlaqəli olmayan verilənlərin sayıdır.

Tamlıq (R) həqiqi müsbətlərin sayı kimi müəyyən edilir və aşağıdakı düsturla hesablanır:

$$R = \frac{Tp}{(Tp + Fn)} \quad (6)$$

Burada: F_n – səhv kimi təsnif edilmiş proqnozlaşdırma ilə əlaqəli olmayan verilənlərin sayıdır.

F1-ölçü (F1-Score) və geri çağırmanın harmonik ortası kimi müəyyən edilir və aşağıdakı düsturla hesablanır:

$$F1 = \frac{2 * (P * R)}{(P + R)} \quad (7)$$

IV. NƏTİCƏ

HSK üzrə milli bazadakı boşluqların doldurulması üçün mövcud yanaşmalar, alqoritmlər şərh edilmiş, kliniki bazalarda boşluqların doldurulması üçün daha məqsədəuyğun hesab edilən MICE alqoritmının riyazi əsasları yerinə yetirilən əməliyyatlar ardıcılığına uyğun şəkildə təqdim edilmişdir.

- Milli HCC Dataseti əsasında proqnozlaşdırma məsələsinin maşın təlimi metodlarının tətbiqi ilə həlli üçün HSK üzrə daha mükəmməl baza olaraq Kaggle saytından götürülmüş HCC Dataseti seçilmişdir;

- HCC Datasetin-dəki yazıların 90% təlim və 10%-i test yazısı kimi seçilmiş, LR, SVM və RF maşın təlimi alqoritmləri tətbiq edilmiş, RF alqoritmının ən yüksək dəqiqliyə malik nəticə göstərdiyi təyin olunmuşdur;

- HSK üzrə milli bazadakı 50 nümunə (proqnoz nəticələri məlum olan) Kaggle saytından götürülmüş HCC Datasetinə əlavə edilmiş və baza genişləndirilmişdir. Genişləndirilmiş HCC Datasetində proqnoz nəticələrin alınması üçün RF alqoritmı və XGB klassifikatoru tətbiq edilmiş, milli HCC Datasetindəki hər bir nümunənin “0 (ölüm)” və “1 (sağqalma)” sinifləri üzrə paylanması göstərilmişdir;

- RF alqoritmı və XGB klassifikatoru üzrə alınan nəticələrin bir-birilə və retrospektivdə müqayisəsi verilmiş, əsaslandırılmışdır ki, milli HCC Dataset-indəki digər məlumatların da hissə-hissə (genişləndirilmiş HCC Dataset-indəki nümunələr hər dəfə 80% qəbul edilməsi, milli bazadan 20% nümunənin seçilməsi) genişləndirilmiş HCC Datasetinə əlavə edilməsi və XGB klassifikatorunun tətbiqi əsasında kliniki xəstələrin vəziyyəti ilə bağlı proqnoz nəticələrin alınması mümkündür.

ƏDƏBİYYAT

- [1] N. Y. Bayramov, *Surgical Diseases of the Liver*. Baku: Qismet, 2012.
- [2] D. Q. Huang, A. Tran, E. X. Tan, S. N. Nerurkar, R. Teh, M. L. P. Teng, E. J. Yeo, B. Zou, C. Wong, C. O. Esquivel, C. A. Bonham, and M. H. Nguyen, “Characteristics and outcomes of hepatocellular carcinoma patients with macrovascular invasion following surgical resection: a meta-analysis of 40 studies and 8,218 patients,” *Hepatobiliary Surg. Nutr.*, vol. 11, no. 6, pp. 848–860, 2022.
- [3] J. R. Jennifer and L. Brit, “Suffering in silence: Medical error and its impact on health care providers,” *J. Emerg. Med.*, vol. 54, no. 4, pp. 402–409, 2017.
- [4] B. Attia, A. K. Rehan, and W. R. Ahsan, “Medical errors: causes, consequences, emotional response and resulting behavioral change,” *Pak. J. Med. Sci.*, vol. 32, no. 3, pp. 523–528, 2016.

- [5] S. Aman and P. Babita, “An efficient diagnosis system for detection of liver disease using a novel integrated method based on principal component analysis and K-nearest neighbor (PCA-KNN),” *Int. J. Healthc. Inf. Syst. Inform.*, vol. 11, no. 4, pp. 56–61, 2016.
- [6] J. S. Sartakhti, M. H. Zangoeei, and K. Mozafari, “Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA),” *Comput. Methods Programs Biomed.*, vol. 108, no. 2, pp. 570–579, 2015.
- [7] M. G. Mammadova, N. Y. Bayramov, Z. G. Jabrayilova, M. I. Manafli, and M. R. Huseynova, “Knowledge transformation in the intelligent system for hepatocellular carcinoma staging,” in *Proc. 8th Int. Conf. on Control and Optimization with Industrial Applications (COIA)*, Baku, Azerbaijan, Aug. 24–26, 2022, vol. 1, pp. 318–321.
- [8] M. H. Məmmədova, N. Y. Bayramov, Z. Q. Cəbrayilova, M. İ. Manafli, M. R. Hüseynova, “Hepatosellulyar karsinomanın diaqnostikası üçün həkim qərarlarının dəstəklənməsi sistemi,” *Azerbaijan Medical Journal*, 2024, №4, pp.163-169.
- [9] M. H. Mammadova, Z. G. Jabrayilova, and L. Garayeva, “Selecting a machine learning algorithm for the prediction of hepatocellular carcinoma with the determination of key indicators,” *5th International Conference on Problems of Cybernetics and Informatics*, Baku, 2023, pp. 1-4.
- [10] M. H. Mammadova, Z. G. Jabrayilova, and L. Garayeva, “Prognosing Hepatocellular Carcinoma Based on a National Database Using Machine Learning Algorithms,” *9th International Conference on Control and Optimization With Industrial Applications*, Istanbul, Turkey, 2024, pp. 108-113.
- [11] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, John Wiley & Sons, 408 p., 2014.
- [12] Z. Zhang, “Multiple Imputation with Multivariate Imputation by Chained Equation (MICE) Package,” *Annals of Translational Medicine*, vol. 4, no. 2, pp. 30, 2016.
- [13] Santos, P. A. Henriques, P. Garcia-Laencina, A. Simao, and A. Carvalho, “A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients,” *J. Biomed. Inform.*, vol. 58, pp. 49–59, 2015.
- [14] S. Yadav and S. Shukla, “Analysis of k-fold cross-validation over holdout validation on colossal datasets for quality classification,” in *Proc. 2016 IEEE 6th Int. Conf. Adv. Comput. (IACC)*, 2016, pp. 78–83.
- [15] C. R. Harris, K. J. Millman, et al., “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. [Online].M.
- [16] M. H. Mammadova, Z. G. Jabrayilova, A. A. Ahmadova, and L. Garayeva, “Prediction of hepatocellular carcinoma using a machine learning algorithm,” *The 16th IEEE International Conference on Application of Information and Communication Technologies (AICT’2022)*, Washington, DC, 2022, pp. 1–4.
- [17] M. H. Məmmədova, Z. Q. Cəbrayilova, and L. A. Qarayeva, “Hepatosellulyar karsinoma üzrə milli bazada maşın təlimi alqoritmlərinin tətbiqi ilə proqnozlaşdırma məsələsinin həlli,” *The International Conference on Artificial Intelligence and Digital Development: Current Realities and Future Perspectives*, Baku, 2024, pp. 13–19.