Artificial Intelligence Research and Development
T. Alsinet et al. (Eds.)

doi:10.3233/FAIA240414

Fuzzy Logic-Based Variable Encoding for Improved Diabetic Retinopathy Prediction

Jordi PASCUAL-FONTANILLES [a,1] Aida VALLS [a,c] Antonio MORENO [a,c]
Pedro ROMERO-AROCA [b,c]

[a] *ITAKA, Dept. Enginyeria Informàtica i Matemàtiques Universitat Rovira i Virgili, Tarragona, Catalonia, Spain*
[b] *Servei d'Oftalmologia, Hospital Universitari Sant Joan de Reus, Catalonia, Spain*
[c] *Institut d'Investigació Sanitària Pere Virgili, Tarragona, Catalonia, Spain*

ORCiD ID: Jordi Pascual-Fontanilles https://orcid.org/0000-0002-7528-5819, Aida Valls https://orcid.org/0000-0003-3616-7809, Antonio Moreno https://orcid.org/0000-0003-3945-2314, Pedro Romero-Aroca https://orcid.org/0000-0002-7061-8987

Abstract. Electronic Health Records (EHRs) contain valuable historical information for building clinical decision support systems. In this study, we focus on exploring novel techniques for improving the prediction of the severity degree of Diabetic Retinopathy (DR) in Diabetes Mellitus patients. In a previous paper, we evaluated the behaviour of different classifiers using the patients' retrospective EHR data to assess their current level of DR, achieving good results. Continuing that work, we now focus on studying different methods for encoding numerical variables, in order to improve the accuracy of these predictions. We propose three normalization methods based on fuzzy sets for encoding numerical data. Because of the inherent uncertainty of medical data, using fuzzy logic to represent the numerical variables can enhance the accuracy of a classifier. The results of the experimental tests, conducted on a dataset of 2108 patients, show that for low-complexity classifiers (such as KNN or CNN) a classical fuzzification technique works the best, while for more complex architectures (like TapNet or ResNet) a fuzzy two-hot encoding gives the best performance. The final aim of the research is to build a clinical decision support system that can make an accurate and personalised prediction of DR evolution.

Keywords. Time series classification, Fuzzy logic, Variable encoding, Diabetic retinopathy

1. Introduction

Diabetic retinopathy (DR) is a severe ocular complication resulting from uncontrolled sugar levels in blood that suffer some diabetes mellitus patients. This dis-

[1] Corresponding Author. Email: jordi.pascual@urv.cat.

ease can damage the retina of the eye. If left untreated at early stages, DR can cause vision loss and even blindness [1]. According to the medical ETDRS standard classification [2], DR can be categorized into four stages: healthy ($DR = 0$), mild ($DR = 1$), moderate ($DR = 2$), and severe ($DR = 3$).

The number of diabetic patients worldwide was estimated at 537 million people in 2021. In Spain alone, it is expected that by 2030, approximately 11.1% of the population will be diagnosed with diabetes, reaching 3.8 million inhabitants [3]. Moreover, a recent study found that the DR prevalence of type-2 diabetic patients in Spain was 15.28%, including 1.92% at the severe level ($DR = 3$) [1].

For this reason, computer-based methods to assist in DR diagnosis are being developed. Many of them perform the risk assessment through the analysis of eye fundus images. However, not all computer-based analysis of eye fundus images are solely focused on DR classification. Lesion detection (such as microaneurysms, exudates and hemorrhages) and segmentation (including blood vessels and the optical disc) are also studied. Dubey et al. conducted an analysis of image-based DR techniques from the state-of-the-art [4]. For binary classification, the accuracies ranged from 94% to 97%, and for multiclass classification from 94% to 98%. Despite the promising results, given the cost associated with obtaining eye fundus images, other clinical decision support systems (CDSS) based on clinical and analytical data available in the Electronic Health Records (EHR) are on the focus of research.

Several studies in the literature have analysed and compared various classifiers on DR EHR datasets, including models such as random forests, XGBoost, logistic regression, support vector machines, and k-nearest neighbors [5,6,7,8,9,10]. However, the results of these studies vary, highlighting a lack of consensus on the best classifier for predicting DR. Retiprogram was presented as a CDSS for DR prediction that considers the patient's current conditions and analytical data from the last blood analysis [11,12]. It is based on a Fuzzy Random Forest classifier, which achieved an accuracy of 81%, sensitivity of 80%, and specificity of over 84%. It was initially developed as a binary CDSS [12]. Later, it was extended to handle ordinal multi-class classification, enabling the detection of different levels of DR severity [13] with an accuracy of 73%.

When a patient is diagnosed with DR, he/she usually starts some treatments in order to improve some clinical factors; therefore, for the patients under treatment it is more challenging to distinguish their DR grade only by observing the values of a unique visit. Our hypothesis was that a retrospective analysis could be more adequate to have an overall view of the patient's conditions evolution and could improve the grading of DR, especially for those long-term patients for which classical machine learning models have more difficulties. For this work, we have a real dataset about Type-2 diabetic people, with data between 2010 and 2021. The dataset contains clinical and analytical variables of different types (numerical and categorical), which were extracted from their EHRs. The purpose of the work was to study the complexities of the available temporal data. In [14] we designed a CDSS that includes a pre-processing stage to build homogeneous data series and the use of a multi-variate series classifier to take advantage of the historical information available. In this paper, we want to study the effects of using approximate reasoning techniques in the input data codification step at the pre-processing stage. Instead of providing precise numerical values, the models will take advantage of the fuzzy linguistic model that represents the domain knowledge of the medical specialists to encode numerical variables.

The rest of this paper is structured as follows. In Section 2 we summarise the DR classifier that incorporates temporal risk factors for DR grading. Next, Section 3 introduces the fuzzy logic based techniques we propose to encode numerical variables. Section 4 presents the experimental results of our proposed encoding methods. Finally, Section 5 summarises the conclusions and future work.

2. DR classifier based on temporal risk factors

In this section, we describe the DR classification model using multivariate temporal data. At each visit to the ophthalmologist, some clinical and analytical data about the diabetic patient are collected. The data are stored in his/her EHR. Nine relevant risk factors for DR were selected by the experts. They consist of six numerical and three categorical variables [11]. Those variables and the number of labels defined by the ophthalmologists for each of them are shown in Table 1.

Table 1. Diabetic Retinopathy risk factors

Variable	Description	Type	Range	# Labels
Age	Current age in years	Numerical	0 - 100	7
EVOL	Duration in years of Type-2 diabetes	Numerical	0 - 30	5
HbA1c	Concentration of glycated hemoglobin present in the bloodstream	Numerical	4 - 12	5
CKDEPI	Measure for the glomerular filtration rate of the kidney	Numerical	10 - 100	5
MA	Level of albumin in the urine	Numerical	29 - 32	2
BMI	Body Mass Index	Numerical	18 - 40	7
Gender	The sex of the patient	Categorical	-	2
TTM	Prescribed treatment for DM	Categorical	-	3
HTAR	Control of arterial hypertension	Categorical	-	2
DR	Degree of Diabetic Retinopathy	Categorical	-	4

After some years, each patient has a sequence of values for those variables, which are stored in his/her EHR, together with the DR diagnosis value assigned by the ophthalmologist in each visit ($DR = \{0, 1, 2, 3\}$). As it was studied in [15], this kind of data presents several challenges that must be solved by applying several pre-processing techniques on the raw EHR data to obtain an appropriate time series. These pre-processing steps are depicted in Figure 1. In this paper, we have focused on the encoding step. The DR diagnosis is also included as a categorical variable in the training dataset, except for the last entry of the sequence, since this is the value the system must predict. In this way, we can use the previous DR evolution to train the time series classifier. Moreover, we can use the last DR value as ground truth to validate the output value.

The periodicity of control visits of diabetic people ranges from 6 to 24 months. In each visit, a new value is obtained for each of the presented variables (Ta-

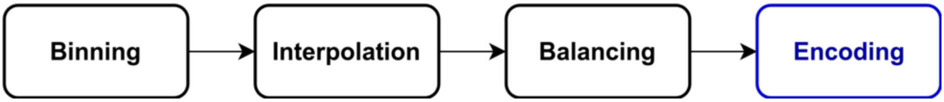

Figure 1. Pre-processing steps for temporal EHR series

ble 1). In this framework, several challenges have been identified: the short data sequences, as EHR sequences are much shorter than usual time series; irregular visits frequency, as the data points are not spaced at regular time intervals; different data alignment, because the visit frequency of each patient is highly variable; missing data on the EHR; labelling mistakes due to human error and data imbalance, as 82.1% of the data belong to the negative ($DR = 0$) class, 9.9% to mild ($DR = 1$), 6.3% to moderate ($DR = 2$), and just 1.7% to severe ($DR = 3$).

To overcome the existing challenges, in [14] we proposed tailored pre-processing solutions that take into account the medical knowledge related to the DR disease. After observing the collected data and considering the medical expertise, the desired length of the sequences was fixed to 10 years. In order to obtain a sequence with data of 10 consecutive years for each patient, for each of the variables the following 4 pre-processing steps were defined:

1. **Binning**: if the patient has more than one value for a given variable in the same year, the most severe outcome is selected.
2. **Interpolation**: if the patient has no visit on a certain year, the value is obtained by a double interpolation procedure taking into account the values of the rest of the years of the sequence.
3. **Balancing**: a new fuzzy-based sample generation method was also proposed to generate new fictive but feasible data values for the minority classes (i.e., the positive ones). Those new examples were used to balance the training dataset.
4. **Encoding**: to have the same scale in all variables, one-hot encoding is used for categorical values, and standard normalization for numerical ones.

This pre-processing method is able to transform the EHR data into fixed frequency and equal length series. The resulting dataset can then be used to train any time series classifier with multivariate data capabilities. In the next section, we will explain some alternative techniques for encoding the data using domain knowledge about the semantics of the medical numerical variables.

3. Fuzzy logic-based variable encoding

Numerical variables in medical datasets can have large differences between their scales, which can affect the performance of AI models. To address this issue, numerical variable normalization is used to transform all numerical variables into a common scale. In this way, each feature contributes equally to the modeling process.

One common way of normalizing numerical variables is to standardize them by removing the mean (i.e., the mean value will be 0) and scaling to unit variance, Eq. 1 (SE).

$$z = \frac{x-\mu}{\sigma},\ \mu = \frac{1}{N}\sum_{i=1}^{N}(x_i),\ \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}, \tag{1}$$

where μ is the mean and σ is the standard deviation of a given variable.

Medical data has an inherent uncertainty that can be attributed to incomplete or missing values, and to imprecise measurements or diagnoses. Moreover, in the diagnosis process, medical doctors reason qualitatively on attribute values when assessing patients' conditions. For instance, instead of considering the precise numerical age or body mass index values, they consider more general levels such as child/young/old or underweight/normal/overweight. A difference of one year in age or of one kilogram makes no difference on the health diagnosis.

Fuzzy logic is a well-known paradigm to reason qualitatively [16]. Therefore, in this work, we take advantage of the fuzzy linguistic model that represents the domain knowledge of the medical specialists to encode numerical variables. Instead of providing the precise numerical values, the models will use a set of labels provided by the doctors.

In the context of fuzzy logic, a fuzzy partition divides a universe of discourse into overlapping regions or subsets based on their membership values. In contrast to binary or crisp partitions, fuzzy partitions allow for gradual transitions between partitions. A standardized fuzzy partition $A^i = \{A^i_1, \ldots, A^i_n\}$ on a numerical continuous variable X_i, is defined by a group of fuzzy sets A^i_j with a membership degree $0 \leq \mu_{A^i_j}(x) \leq 1$, and satisfies the following property: $\forall x \in X_i, \sum_j \mu_{A^i_j}(x) = 1$. Then, a linguistic variable is defined by a fuzzy partition A^i and a set of p linguistic labels $L^i = \{L_1, \ldots, L_p\}$.

Figure 2 depicts the fuzzy sets that were defined by experts for the numerical variable CKDEPI. Five sets were defined, with the linguistic variables Normal, LowNormal, ModeratelyLow, VeryLow and ExtremelyLow.

Taking advantage of fuzzy linguistic variables, we will study three methods to encode the numerical data using the fuzzy sets provided by the medical specialists. Each numerical variable will be transformed into p dummy variables, one for each value in L^i.

In all the methods, the first step is obtaining the fuzzy sets where the numerical value has any membership. Their corresponding labels will determine the values that will be set in each dummy variable. The proposed methods are the following ones, which for an input x will determine the value $v(L^i_j)$.

- **Fuzzy One-Hot Encoding (FOH)**:

$$v(L^i_j) = \begin{cases} 1 & \text{if } \mu_{A^i_j}(x) \geq \mu_{A^i_k}(x) \forall k \neq j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

- **Fuzzy Two-Hot Encoding (FTH)**:

$$v(L^i_j) = \begin{cases} 1 & \text{if } \mu_{A^i_j}(x) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

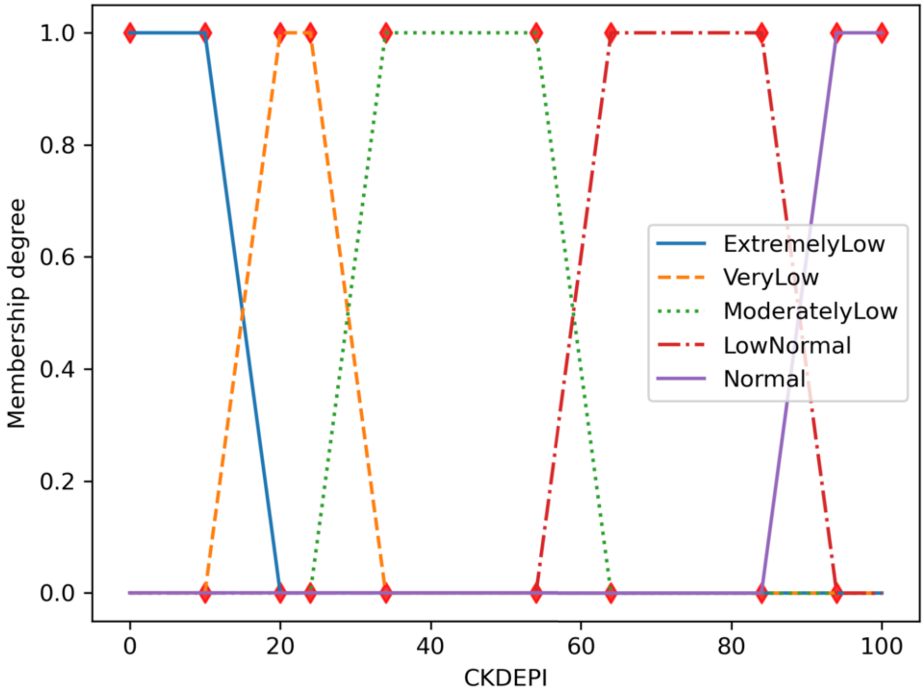

Figure 2. Definition of linguistic labels for CKDEPI

- **Fuzzy Membership (FM)**:

$$v(L_j^i) = \mu_{A_j^i}(x) \tag{4}$$

FOH and FTH generate binary dummy variables, while FM generates a variable with Real values in the range $[0, 1]$. In the next section, these three methods will be tested and compared to the standard scaling.

4. Experimental results

This section presents the experimental results obtained using the data presented in Section 2. The DR dataset has 2108 sequences of 10 entries, after conducting the 2 first stages of the pre-process (binning and interpolation) [14]. The dataset was divided into two: training and testing, with 70% and 30% of the data, respectively (Table 2). It can be clearly seen that there is a high imbalance towards the negative class ($DR = 0$), making it more challenging to predict the classes with higher DR risk. To obtain classes of equal size, we apply a balancing method based on generating new realistic fictitious examples, obtained from short sequences of real data [14]. This third pre-processing step introduces additional examples to the three minority classes in the training dataset, until there are a total of 1212 examples for each class.

Table 2. Distribution of the Diabetic Retinopathy time series data in training/testing

Class/Dataset	**Training (70%)**	**Testing (30%)**	**Total**
DR = 0	1212 (82.2%)	518 (81.8%)	1730 (82.1%)
DR = 1	148 (10%)	61 (9.64%)	209 (9.9%)
DR = 2	92 (6.2%)	41 (6.5%)	133 (6.3%)
DR = 3	23 (1.6%)	13 (2.1%)	36 (1.7%)
Total	1475	633	2108

In the fourth step, we applied data standardization or encoding, before training the classifiers. Categorical variables have been encoded using one-hot encoding (OHE), while numerical ones were either standardized (SE) or encoded with one of the three fuzzy-based methods proposed in Section 3.

Next, we introduce the five multivariate multiclass time series classifiers we have tested. They are the best performing classifiers according to the review of Pasos et al. [17] with the addition of ResNet. They are representatives of different kinds of approaches to classification. The hyperparameter selection for the classifiers was tuned in [14] and validated with a 10-fold cross validation. The classifiers are the following:

- **KNN**: a well-known, simple, yet good performing, distance-based classifier. In time series, it is usually paired with the Dynamic Time Warping distance.
- **ROCKET**: combines convolutional kernel transforms with a linear classifier. The randomly chosen kernels are used to create feature maps, which are used to train the classifier [18].
- **TapNet**: an architecture that combines three components: dimension permutation, embedding learning, and attentional prototype learning to generate class prototypes [19].
- **CNN**: an adaptation of convolutional neural networks for time series [20].
- **ResNet**: a type of CNN that uses residual connections to ease the training process by allowing gradients to flow more easily through the network [21].

We repeated the experiments five times for each classifier and encoding method. The obtained results for each test have been aggregated using the mean. Table 3a displays the overall performance metrics of accuracy and quadratic weighted kappa. Quadratic weighted kappa is particularly relevant for this ordinal multiclass problem, because it penalizes mistakes according to the distance between the ground truth and the predicted class. In medical decision support, a short difference between the correct class and the predicted one is crucial in order to not affect the health of the patient. Hence, we aim to minimise it as much as possible. In contrast, Table 3b presents precision and recall metrics. Both of them focus on the positive examples, which are the most relevant to be detected on medical diagnosis problems. Precision measures how often examples are correctly identified as belonging to the positive classes. A high precision indicates a low rate of false positives, which is essential to minimise unnecessary treatments. On the other hand, recall measures the proportion of actual positive examples being correctly identified. A high recall means a low number of false negatives, reducing the likelihood of missed diagnoses.

Table 3. DR temporal classification results

(a) Accuracy and Kappa

Classifier	**Metric**	**SE**	**FOH**	**FTH**	**FM**
KNN	*Acc.*	0.866	0.874	0.821	**0.896**
	Kappa	0.654	0.678	0.413	**0.737**
CNN	*Acc.*	0.929	0.937	0.934	**0.942**
	Kappa	0.868	0.871	0.867	**0.888**
TapNet	*Acc.*	0.924	0.924	**0.942**	0.902
	Kappa	0.858	0.859	**0.877**	0.845
ROCKET	*Acc.*	0.909	**0.911**	**0.911**	**0.911**
	Kappa	0.736	**0.741**	0.737	0.740
ResNet	*Acc.*	0.929	0.898	**0.934**	0.891
	Kappa	0.860	0.745	**0.863**	0.788

(b) Precision and Recall

Classifier	**Metric**	**SE**	**FOH**	**FTH**	**FM**
KNN	*Prec.*	0.654	0.687	0.510	**0.764**
	Recall	0.703	0.753	0.473	**0.761**
CNN	*Prec.*	0.873	0.885	0.861	**0.905**
	Recall	0.809	**0.839**	0.796	0.819
TapNet	*Prec.*	0.866	0.867	**0.886**	0.853
	Recall	**0.813**	0.810	0.812	0.811
ROCKET	*Prec.*	**0.916**	0.914	**0.916**	0.913
	Recall	0.667	**0.673**	0.665	0.671
ResNet	*Prec.*	0.847	0.759	**0.867**	0.795
	Recall	0.802	**0.813**	0.804	0.804

From the obtained results, several conclusions can be drawn by comparing the standard encoding (SE, Eq. 1) with the three fuzzy-based methods proposed in this paper. Firstly, we observe that the choice of the encoding approach has a limited impact on the performance of the ROCKET classifier. Although all proposed methods outperform SE in accuracy and kappa, their improvements are marginally better in this classifier. The Kappa and accuracy of the fuzzy encoding outperform the standard normalization in the rest of the classifiers. The improvement of Fuzzy One-Hot Encoding is generally small in comparison with SE. FTH improves the performance on the classifiers using complex architectures (TapNet and ResNet), but it is surpassed by FM (Fuzzy Membership) for simpler classifiers (CNN and KNN). All those observations are supported by the precision measure, as well. In the case of recall, there is more variation in the obtained results. The recall is higher with FOH for CNN, ROCKET and ResNet.

In general, the best classification performance is obtained by CNN+FM with accuracy of 0.942 and kappa of 0.888. It also has a precision of 0.905 and a recall of 0.819. ROCKET is slightly better in precision, but it is much worse in recall (below 0.7). The second-best classifier is TapNet+FTH, with similar accuracy and kappa, but with lower precision and recall.

The highest improvements in the use of fuzzy encoding are observed in KNN for all metrics, which is an expected result given its lower baseline performance. The other classifiers had much better performance from the beginning, resulting in smaller improvements.

In summary, no single encoding method consistently outperforms the others across all classifiers. However, as a general guideline, Fuzzy Membership seems to work better for architecturally simple classifiers, and Fuzzy Two-Hot for architecturally complex ones.

5. Conclusions and future work

In this paper, we investigated the impact of different encoding methods on the performance of various multiclass time series classifiers in assessing diabetic retinopathy risk from EHR data. Our results demonstrate that while standard encoding can be effective for some classifiers, the fuzzy-based encoding approaches we propose can lead to improved performance.

The findings suggest that Fuzzy Membership is a good choice for simple architectures like CNN and KNN, whereas Fuzzy Two-Hot Encoding may be more effective for complex models such as TapNet and ResNet. However, no single encoding method consistently outperforms others across all classifiers, emphasizing the importance of testing different encoding approaches.

By selecting the most optimal encoding method for each specific classifier, we can improve the overall performance of DR prediction models and ultimately contribute to better healthcare decisions. Moreover, these fuzzy-based encoding techniques are not limited to DR risk assessment, as they could also be applied to other domains where historical information is available, or even on non-temporal datasets.

Future work could involve exploring other fuzzy-based encoding methods or combining them with traditional encoding approaches to further enhance model performance. We should also test the proposed methods with other EHR datasets and other time series classifiers to confirm the validity of these results in different situations. The possibility of using or adapting the proposed method for non-temporal datasets could also be studied.

Acknowledgements

This work was supported by Instituto de Salud Carlos III (ISCIII) [project PI21/00064] and co-funded by the European Union; Universitat Rovira i Virgili (URV) [projects number 2023PFR-URV-114, 2022PFR-URV-41]; ITAKA funding from AGAUR [2021-SGR-00114].

References

[1] Romero-Aroca P, López-Galvez M, Martinez-Brocca MA, Pareja-Ríos A, Artola S, Franch-Nadal J, et al. Changes in the Epidemiology of Diabetic Retinopathy in Spain: A Systematic Review and Meta-Analysis. Healthcare (Switzerland). 2022 7;10:1318.

[2] Wilkinson CP, Ferris FL, Klein RE, Lee PP, Agardh CD, Davis M, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology. 2003;110:1677-82.

[3] Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes research and clinical practice. 2010;87(1):4-14.

[4] Dubey S, Dixit M, dixitmits MD. Recent developments on computer aided systems for diagnosis of diabetic retinopathy: a review. Multimedia Tools and Applications. 2023;82:14471-525.

[5] Zhao Y, Li X, Li S, Dong M, Yu H, Zhang M, et al. Using Machine Learning Techniques to Develop Risk Prediction Models for the Risk of Incident Diabetic Retinopathy Among Patients With Type 2 Diabetes Mellitus: A Cohort Study. Frontiers in Endocrinology. 2022 5;13:885.

[6] Tsao HY, Chan PY, Su ECY. Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms. BMC Bioinformatics. 2018 8;19:111-21.

[7] Ogunyemi OI, Gandhi M, Tayek C. Predictive Models for Diabetic Retinopathy from Non-Image Teleretinal Screening Data. AMIA Summits on Translational Science Proceedings. 2019;2019:472.

[8] Sun Y, Zhang D. Diagnosis and Analysis of Diabetic Retinopathy Based on Electronic Health Records. IEEE Access. 2019;7:86115-20.

[9] Ogunyemi O, Kermah D. Machine Learning Approaches for Detecting Diabetic Retinopathy from Clinical and Public Health Records. AMIA Annual Symposium Proceedings. 2015;2015:983.

[10] Reddy GT, Bhattacharya S, Ramakrishnan SS, Chowdhary CL, Hakak S, Kaluri R, et al. An Ensemble based Machine Learning model for Diabetic Retinopathy Classification. International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020. 2020 2.

[11] Romero-Aroca P, Valls A, Moreno A, Sagarra-Alamo R, Basora-Gallisa J, Saleh E, et al. A Clinical Decision Support System for Diabetic Retinopathy Screening: Creating a Clinical Support Application. Telemedicine and e-Health. 2019;25:31-40.

[12] Saleh E, Błaszczyński J, Moreno A, Valls A, Romero-Aroca P, de la Riva-Fernández S, et al. Learning ensemble classifiers for diabetic retinopathy assessment. AI in Medicine. 2018;85:50-63.

[13] Pascual-Fontanilles J, Lhotska L, Moreno A, Valls A. Adapting a Fuzzy Random Forest for Ordinal Multi-Class Classification. Frontiers in Artificial Intelligence and Applications. 2022 10;356:181-90.

[14] Pascual-Fontanilles J, Valls A, Romero-Aroca P. Multivariate data binning and examples generation to build a Diabetic Retinopathy classifier based on temporal clinical and analytical risk factors. Knowledge-Based Systems. 2024;300:112154.

[15] Pascual-Fontanilles J, Valls A, Moreno A, Romero-Aroca P. Challenges in the Exploitation of Historical Clinical Data for the Classification of Diabetic Retinopathy Patients. Frontiers in Artificial Intelligence and Applications. 2023 10:204-7.

[16] Zadeh LA. Knowledge representation in fuzzy logic. In: An introduction to fuzzy logic applications in intelligent systems. Springer; 1992. p. 1-25.

[17] Ruiz AP, Flynn M, Large J, Middlehurst M, Bagnall A. The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Mining and Knowledge Discovery. 2021;35:401-49.

[18] Dempster A, Petitjean F, Webb GI. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. Data Mining and Knowledge Discovery. 2020;34:1454-95.

[19] Zhang X, Gao Y, Lin J, Lu CT. TapNet: Multivariate Time Series Classification with Attentional Prototypical Network. Proceedings of the AAAI Conference on Artificial Intelligence. 2020 4;34:6845-52.

[20] Zhao B, Lu H, Chen S, Liu J, Wu D. Convolutional neural networks for time series classification. Journal of Systems Engineering and Electronics. 2017 2;28:162-9.

[21] Wang Z, Yan W, Oates T. Time series classification from scratch with deep neural networks: A strong baseline. In: 2017 International Joint Conference on Neural Networks (IJCNN); 2017. p. 1578-85.