IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (4) Part (2): 252-262

UKJAES

University of Kirkuk Journal For Administrative and Economic Science

ISSN:2222-2995 E-ISSN:3079-3521
University of Kirkuk Journal For Administrative and Economic Science

Kirkuk University - College of Administration & Economics - جامعة كركوك - كلية الإدارة والاقتصاد 1428 2007

Khalil Badiaa Rahman, Ahmed Farhad Ali & Salih Samira Muhamad. Classifying and predicting diabetic patients using Multinomial Logistic Regression and Multilayer perceptron model. *University of Kirkuk Journal For Administrative and Economic Science* **(2025) 15 (4) Part (2):252-262.**

Classifying and predicting diabetic patients using Multinomial Logistic Regression and Multilayer perceptron model

Badiaa Rahman Khalil [1], Farhad Ali Ahmed [2], Samira Muhamad Salih [3]

[1,2,3] *Sulamani University: College Administration & Economics, Statistics & Informatics Department, Sulaimani, Iraq*

badiaa.khalil@univsul.edu.iq [1]
farhad.ahmed@univsul.edu.iq [2]
samira.salih@univsul.edu.iq [3]

Abstract: Diabetes is a chronic medical condition that affects millions of people worldwide and poses a serious health risk if not diagnosed early and treated effectively. These datasets consist of ten clinical variables and patient characteristics associated with the diagnosis of diabetes. The data used in this study were collected from patients of Sulaimani Diabetes Hospital for the years 2022 to 2023 and 2024, from which we were able to record information on 192 patients. Each row represents one patient record. The register consists of 10 attributes, of which one is a predictive Y-valued attribute (diabetic, prediabetic, and nondiabetic). Other features are used for the prediction part of the algorithm. Feature selection techniques with two classification algorithms are implemented using Weka programming. The aim of this study is to predict and classify patients with diabetes using two machine learning algorithms: the Multinational Logistic Regression and the Multilayer Perceptron (MLP) model. K-fold cross-validation is used when K = 5, 7, 9, and 10. From the results of this study, we find that the MLR classifier at k=9 achieves the highest accuracy (90.1042%), as in MLP, which has an accuracy of 88.0208% when k=9.

Keywords: Diabetes Prediction, Polynomial Logistic Regression (MLR), Multilayer Perceptron (MLP), Machine Learning Classification.

تصنيف مرضى السكري والتنبؤ بهم باستخدام الانحدار اللوجستي متعدد الحدود ونموذج الإدراك الحسي متعدد الطبقات

م.م. بديعه رحمان خليل [١]، د.فرهاد على احمد [٢]، م.د. سميره محمد صالح [٣]

[١،٢،٣] *جامعة السليمانية: كلية الإدارة والاقتصاد، قسم الإحصاء والمعلوماتية، السليمانية، العراق*

المستخلص: داء السكري هو حالة طبية مزمنة تؤثر على ملايين الأشخاص في جميع أنحاء العالم وتشكل خطرًا صحيًا خطيرًا إذا لم يتم تشخيصها مبكرًا وعلاجها بشكل فعال. تتكون مجموعات البيانات هذه من عشرة متغيرات سريرية وخصائص المريض المرتبطة بتشخيص داء السكري. تم جمع البيانات المستخدمة في هذه الدراسة من

مرضى مستشفى السليمانية للسكري للأعوام من ٢٠٢٢ إلى ٢٠٢٣ و٢٠٢٤، والتي تمكنا من خلالها من تسجيل معلومات عن ١٩٢ مريضًا. يمثل كل صف سجلًا واحدًا للمريض. يتكون السجل من ١٠ سمات، إحداها سمة تنبؤية بقيمة Y (مصاب بالسكري، ومقدمات السكري، وغير مصاب بالسكري). يتم استخدام ميزات أخرى لجزء التنبؤ من الخوارزمية. يتم تنفيذ تقنيات اختيار الميزات مع خوارزميتين للتصنيف باستخدام برمجة Weka. الهدف من هذه الدراسة هو التنبؤ بمرضى السكري وتصنيفهم باستخدام خوارزميتين للتعلم الآلي: الانحدار اللوجستي متعدد الجنسيات ونموذج متعدد الطبقات Perceptron (MLP). يُستخدم التحقق المتبادل متعدد الأضعاف (K-fold) عند K = 5، ٧، ٩، و١٠. بناءً على نتائج هذه الدراسة، نجد أن مُصنِّف MLR عند k = 9 يُحقق أعلى دقة (٩٠,١٠٤٢٪)، كما هو الحال في MLP، الذي تبلغ دقته ٨٨,٠٢٠٨٪ عند k = 9.

الكلمات المفتاحية: التنبؤ بمرض السكري، الانحدار اللوجستي متعدد الحدود (**MLR**)، المُدرك متعدد الطبقات (**MLP**)، تصنيف التعلم الآلي.

Corresponding Author: E-mail: badiaa.khalil@univsul.edu.iq

Introduction

Diabetes is a chronic, progressive condition that affects the body and how it uses blood sugar. When blood sugar levels are too high, it can cause serious problems in the heart, kidneys, eyes and nerves [1]. With the incidence of diabetes increasing worldwide, timely diagnosis and accurate classification of patients are important. These steps will help us get treatment, improving patient health outcomes, and improving the health care system[4]. While traditional methods of diagnosing diabetes are useful, they often use rules that allow patients to miss complex patterns in the data. New developmental researchers in machine learning are offering better ways to aid in the use of accurate predictions [6]. In this study, we use two machine learning methods that classify many polynomial logistic regression (MLR) and multilayer perceptron (MLP) regions and predict whether the subject is diagnosed or prediabetic or not. MLR is a type of logistic domain that works well for classifying objects into multiple groups and shows a clear relationship between causes and outcomes. MLP, on the other hand, is a type of artificial neural network that can detect complex and inconsistent patterns in data, making it a powerful tool for medical classification. To test these two models, we used a data set from Sulaimani Diabetes Hospital that included 192 patient records. We used K-Fold Cross-Validation with values (K=5, 7, 9, and 10) to ensure that the results are reliable. We measured how well the model performed with accuracy, precision, recall and F1 score. We want to find out which of these two models best classifies and predicts people with diabetes. The results of this study can help doctors to diagnose the disease early and provide appropriate treatment. Machine learning models can also analyse medical data accurately and effectively, improving the quality of healthcare systems.

1st: Related Works

This section describes several studies that used machine learning algorithms to predict and classify people with diabetes using medical data.

Zoe et al. (2018) used a medical dataset with 14 clinical features. was obtained from hospital physical examinations in Luzhou, China, to predict diabetes using decision trees, random forests (RF), and neural networks. The study used five-fold cross-validation and corrected cluster imbalances by taking five random samples from the data (68,994 records) and averaging the results across all experiments. They used PCA and mRMR to reduce the number of dimensions, and the results showed that RF had the best prediction accuracy [10].

Ahuja et al. (2019). Ahuja and colleagues (2019) in this study used 768 records from the Pima Indian diabetes dataset. After correcting missing values using the median, linear discriminant analysis is performed. Five classification methods are used in conjunction with feature selection strategies using the Python programming language: SVM, MLP, RF, DT, and logistic regression. The aim of this study is to compare machine learning algorithms in order to accurately predict diabetes in patients. and use K-fold cross validation. When k = 2, 4, 5, and 10. In their analysis, MLP produced the best accuracy of 78.7%, recall of 61.26%, precision of 72.45%, and F1 score of 65.97% when k = 4 [2].

Verma G. and Verma H. (2020). In these papers they propose a machine learning model based on multilayer perceptual neural network (MLP), which can detect diabetic patients. In addition to MLP, 5-fold cross-validation is also used to obtain better results on test data after training. Experimental results of the model show 82% accuracy in prediction, which is quite better. The results of the proposed model are also compared with some existing state of the art [11].

Ahmed et al. (2021). Using data collected from 553 patients at the Federal Medical Centre in Yola, this study compared the classification techniques of logistic regression and multilayer perceptron for diabetes prediction. Insulin, body weight, glucose, diastolic blood pressure, and patient age were the primary factors taken into account. The MLP model was trained on 70% of the data, tested on 20%, and validated on the remaining 10%. Both models were optimized to perform at their peak using SPSS. With an accuracy of 90.6%, the LR model identified diabetic cases 91.0% of the time. Despite the fact that both models performed satisfactorily, MLP outperformed LR in terms of overall classification rate [3].

2nd: Classification Algorithms

1- SoftMax Function

The SoftMax feature is a mathematical operation used to convert raw prediction points (logits) to multimodal logistic regression and opportunities for classifying multiple classes. This ensures that all possibilities make yoga for 1. Each option is between 0 and 1[12].

2- Multinomial Logistic Regression

Multinomial Logistic Regression is an extension of binary logistic regression used for multi-class classification (when the target variable has 3 or more unordered categories) [5].
Suppose the response variable Y has J unordered categories (classes), labelled as:
$Y \in \{1,2,\ldots\ldots,J\}$
$Let\ X = (x_1, x_2, \ldots, x_3)$. Be the vector pf predictor variables.
We choose one category (usually the last, J) as the reference class.
Then, for each of the other classes $j = 1,2,\ldots,J-1$,the model estimates the log odds as:

$$log\left(\frac{P(Y=j/X)}{P(Y=J/X)}\right) = \beta_{j0} + \beta_{j1}X_1 + \cdots + \beta_{jp}X_p = X^T\beta_j \quad (1)$$

This can also be written more compactly as:

$$log\left(\frac{P(Y=j/X)}{P(Y=J/X)}\right) = X^T\beta_j \quad for\ \ j = 1,2,\ldots,J-1 \quad (2)$$

To compute the probabilities for each class:

$$P(Y = \mathrm{j}/\mathrm{X}) = \frac{e^{X^T}\beta_j}{1+\sum_{k=1}^{J-1} e^{X^T}\beta_j} \quad (3)$$

3- A multi-layer perceptron (MLP)

An artificial neural network with multiple layers of neurons is called a multilayer perceptron (MLP) [9]. Complex patterns in network data are possible because MLP neurons often do not employ linear activation properties. Because MLP machines are capable of learning non-relationships in data, they are valuable learning tools for tasks like pattern recognition, regression, and classification. The human nervous system served as the model for the multilayer Perceptron [9]. MLP has the following advantages: (i) it is tolerant of excessive flaws, meaning that even if neurons and their relationships fail, they still function; and (ii) it is not linear by nature, making it appropriate for a wide range of real-world issues [2].

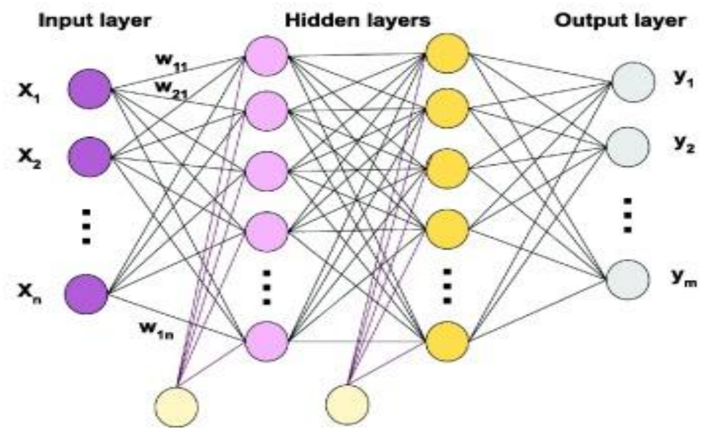

Figure (1): Structure of a multilayer perceptron

Input feature vector $\boldsymbol{x} = [\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots\ldots \boldsymbol{x_p}]^{\boldsymbol{T}}$ (4)

Suppose we have one hidden layer with h neurons

$\boldsymbol{Z} = \boldsymbol{f}(\boldsymbol{w}^{(1)}\boldsymbol{x} + \boldsymbol{b}^{(1)})$ (5)

$w^{(1)} = weight\ matrix\ for\ hidden\ layer\ (size\ h \times p)$

$b^{(1)} = bias\ vector\ (size\ h \times p)$

$f = activation\ function$

Compute raw output (logits) for each class

$\boldsymbol{O} = \boldsymbol{w}^{(2)}\boldsymbol{Z} + \boldsymbol{b}^{(2)}$ (6)

Were

$\boldsymbol{w}^{(2)} = weight\ matrix\ for\ output\ layer\ (size\ 3 \times h)$

$\boldsymbol{b}^{(2)} = bias\ vector\ (size\ 3 \times 1)$

Compute probability for each class

$P_j = \frac{e^{o^j}}{\sum_{K=1}^{3} e^{o^K}}$ (7)

$for\ j = 1(A), 2(B), 3(C)$ choose class with highest probability P_j

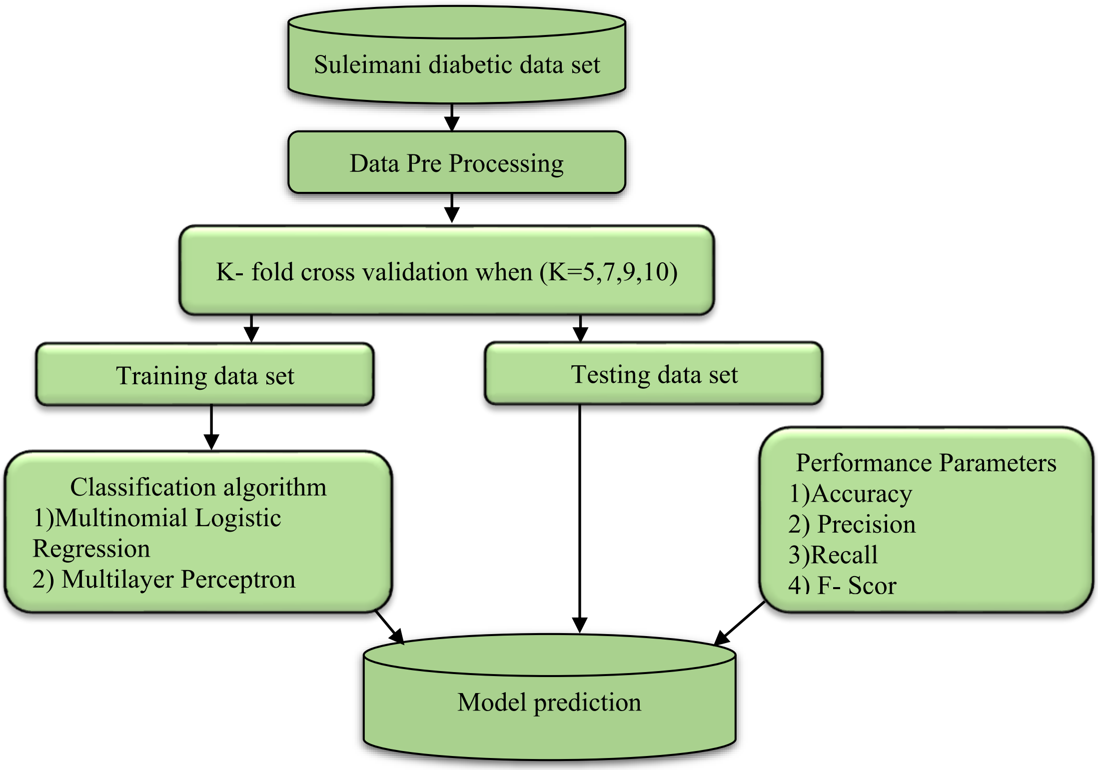

Figure (2): Explains in full the classification and prediction procedure.

3rd: Experimental Configuration

1- Parameters of Performance

The following four evaluation parameters are taken into account:

A. Accuracy

This serves as the foundation for evaluating any future model's quality. The accuracy was determined by calculating the proportion of accurate predictions for all data points. Following the practical selection and application of the folding technique, this paper presents the highest accuracy attained by two machine learning models. The accuracy equation is given by equation (8) [2].

$$\text{Accuracy} = \frac{(TP+TN)}{TP+TN+FP+FN)} * 100\% \quad \ldots \quad (8)$$

Were,
TP = True Positive , FP = False Positive
TN = True Negative , FN = False Negative

B. Precision

The exactness the percentage of pertinent events among the retrieved events is a model's precision. An alternative name for it is a positive predictive value. It is determined by dividing the number of true positives in a model by the total number of positives. To put it simply, a high precision algorithm produces more relevant results than irrelevant ones. The accuracy equation is given by equation (10) [2].

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

C. Recall

Another name for recall is the model's sensitivity. It is a subset of the quantity of pertinent events that have been extracted. With a high recall, the majority of events were recalled. For the total of real positivity and false negative in (10) [2], it is quantified as a relationship between actual positivity.

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

D. F-Score

By calculating its harmonic mean, the F-Score is a metric that combines recall and precision. If they are near, it is roughly the mean of the two; if not, it is the harmonic mean. The ratio of the square of the geometric mean to the arithmetic mean is known as the harmonic mean. Precision and recall are equally weighted in the F1 metric, as stated in (11)[2].

$$F - Score = \frac{Precion*Recall}{Precion+Recall} \quad (11)$$

2- K-Fold Cross Validation

The 192-case original data set is divided into equalized sub-segments for the K-fold cross-validation procedure [10]. The number of segments is determined by the value of k, which in our case is 5, 7, 9, or 10. Out of all the sub-segments, nine were utilized for training data, while one was used for "testing data." Every sub-partition is used as the testing data at least once during the [k] repetitions of this cross-validation approach. These outcomes from the aforementioned iterations are averaged or combined in some other way to generate a single estimate. The benefit of employing this validation technique is that each dataset entry is used at least once to validate the outcome, and every single data point is used for both training and testing the model [7]. As a result, the model's accuracy is improved [2].

4th: dataset

The dataset used in this thesis was collected from patients of Sulaimani Diabetes Hospital between the years (2022-2023-2024). This dataset consists of a total of 192 patient records. Each row represents one patient record. The register contains 10 attributes, one of which is a predictive attribute labelled Y whose value indicates the type (no diabetic, pre_ diabetic and diabetic). Other features are used in the prediction part of the algorithm. All 10 attributes are categorical attributes. The table below shows the datasets used in this thesis.

Table: (5): The attribute descriptions given below

No	Attribute Name	Attribute description	Value
1	Age	Age of the person	No particular range
2	Gender	Gender of the person	Female=0 Male=1
3	Weight	Binary value	Number
4	High	High of the person	Number
5	Smoker	Nominal value	No=0 Ex=1 Yes=2
6	Family history of diabetes	Nominal value	Negative=0 Positive=1
7	Alcohol	Nominal value	No=0 Ex=1 Yes=2
8	Hypertension	Binary value	No=0 Yes=1
9	Treatment	Nominal Value	Insulin=1 Tablet=2 Tablet and, insulin=3 No medication=4
10	Hbc1A	Y= Result	A- Diabetes (at risk) B- Pre diabetes(higher) C-No diabetic

1- Result

In this study, two learning algorithms—multilayer perceptron (MLP) and multinomial logistic regression (LR) were used to identify and forecast diabetes patients based on 192 data sets of Sulaimani patients. K-fold cross-validation was used when k = 5, 7, 9, and 10. F1 score, accuracy, recall, and accuracy were performance metrics. Table 1 summarizes the findings from the two accuracy-based classification tests.

Table (6): Comparing classifiers with varying k-fold values based on accuracy.

K-Fold Cross Validation	Multinomial Logistic Regression	Multi -Layer perceptron
K=5	89.0625 %	88.0208 %
K=7	89.0625 %	86.9792 %
K=9	90.1042 %	88.0208 %
K=10	90.1042 %	86.4583 %

As shown in Table 6, the result suggests that multinational logistics area improved multilayer perceptron (MLP) in terms of accuracy at all K-fold values. The multinational logistics area achieved the highest accuracy (K = 90.10%) and the lowest accuracy (89.06% at K = 5 and 7), with an area accuracy of (89.58%) and (1.04%) accuracy, indicating a stable performance. In contrast, MLP achieves the highest accuracy (88.02%) at K = 5 and 9 and the lowest accuracy (86.46% at K = 10), with an average accuracy of (87.37%) and a wide range of (1.56%), more variability. The multinational logistics area provided better and more consistent classification performance than all K-folded cross-folded validation settings.

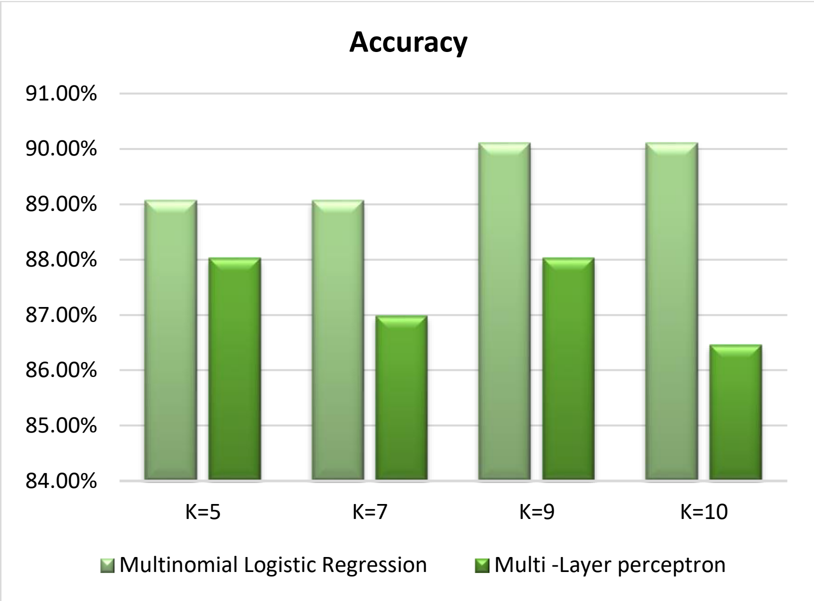

Figure (5): Accuracy comparison of classifiers with varying k-fold values

Table (7): Comparing classifiers with varying k-fold values using precision, recall, and F-score

Multinomial Logistic Regression			
K-Fold Cross Validation	**Precision**	**Recall**	**F-score**
K=5,7	0.910	0.977	0.942
K=9,10	0.911	0.989	0.945
Multilayer Perceptron			
K-Fold Cross Validation	**Precision**	**Recall**	**F-score**
K=5	1.056	1.73	1.036
K=7	0.908	0.959	0.930
K=9	0.908	0.945	0.930
K=10	0.912	0.949	0.930

Multinomial Logistic regression has slightly better F-score and recall than MLP (all but K=5). Both models are performing wonderfully, but logistic regression is slightly more stable at high folds (K=9, 10). Multinomial logistic regression is superior to the multilayer perceptron (MLP) in all the metrics (precision, recall, and F-score). Its highest F-score (0.945) at K=9,10 shows excellent performance. But MLP has severe anomalies at K=5 (metrics > 1.0) suggesting data or model errors. For K≥7, MLP stabilizes but performs poorer than logistic regression (F-score: 0.930 vs. 0.945).

Figure (6): Comparing classifiers with varying k-fold values using precision, recall, and F-score

Table (8): Multinomial Logistic Regression Confusion Matrix when k-fold = 9

Multinomial Logistic regression	**Predicted Class**			
Actual Class	Diabetes	Pre diabetes	No diabetic	Sum
Diabetes	173	13	4	190
Pre diabetes	1	0	0	1
No diabetic	1	0	0	1
Sum	175	13	4	192

The confusion matrix below provides several different details about the results obtained from the evaluation of the Multinomial logistic regression. As mentioned earlier, we have used cross-validation when k=9, which gives us the best results for polynomial logistic regression. There are 192 test cases, or data from 192 patients were used to test the evaluated method. In fact, 175 patients had diabetes mellitus, the remaining 13 patients had prediabetes, and the remaining 4 patients had no diabetes mellitus. Diabetes category of 175 patients, 173 were correctly classified and 2 were misclassified. The prediabetic category correctly classified all 13 cases, and the nondiabetic category correctly classified all 4 cases.

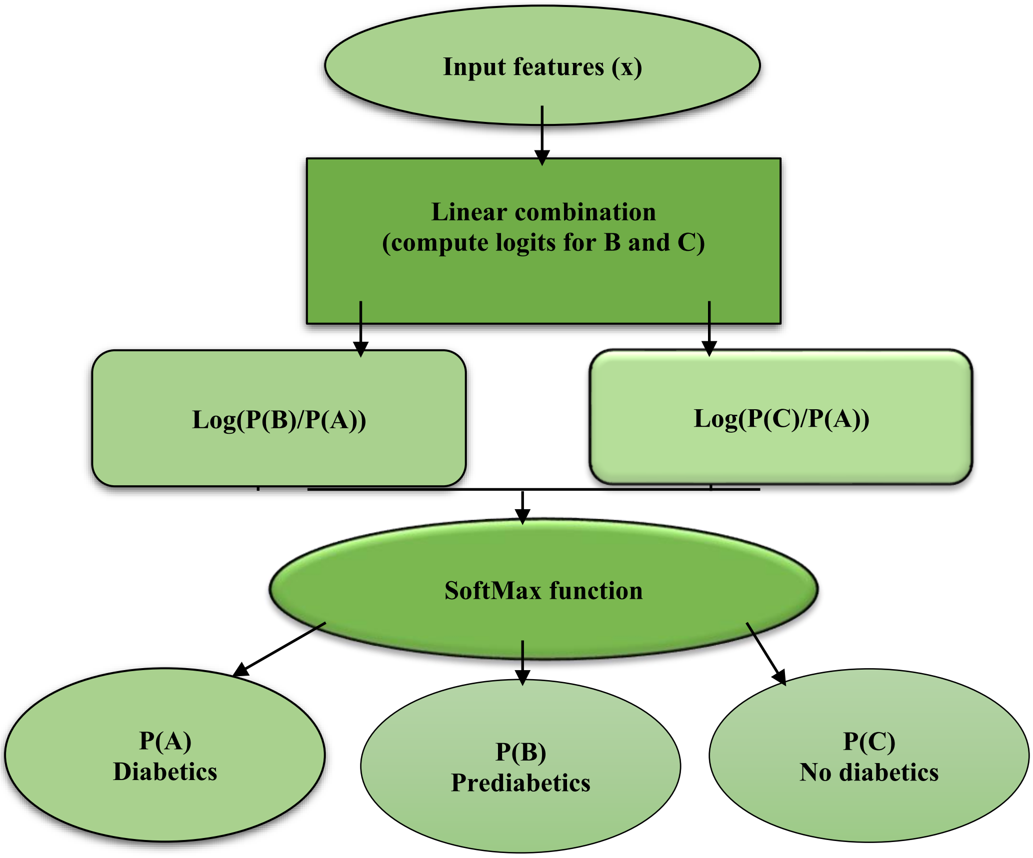

Figure (7): Multinomial logistic regression for three classes is shown in this diagram.

Table (8): Multilayer Perceptron Confusion Matrix when k-fold = 9

Multilayer Perceptron	**Predicted Class**			
Actual Class	Diabetes	Pre diabetes	No diabetic	Sum
Diabetes	168	12	4	184
Pre diabetes	6	1	0	7
No diabetic	1	0	0	1
Sum	175	13	4	192

A variety of information regarding the outcomes of the multilayer perceptron evaluation can be found in the confusion matrix below. We have employed cross-validation for k=9, as previously stated, and this yields the best outcomes for the multilayer perceptron. The evaluated method is tested on 192 test cases, which are data from 192 patients. Actually, 175 of the patients had diabetes mellitus, 13 of them had prediabetes, and 4 of the patients had neither condition. Seven patients in the 175-patient diabetic category were misclassified, whereas 168 patients were correctly classified. Twelve of the thirteen cases in the pre-diabetes category were correctly categorized, while one instance was mistakenly classified. All four cases in the non-diabetic category were accurately classified.

IRAQI Academic Scientific Journals

University of Kirkuk Journal For Administrative and Economic Science (2025) 15 (4) Part (2): 252-262

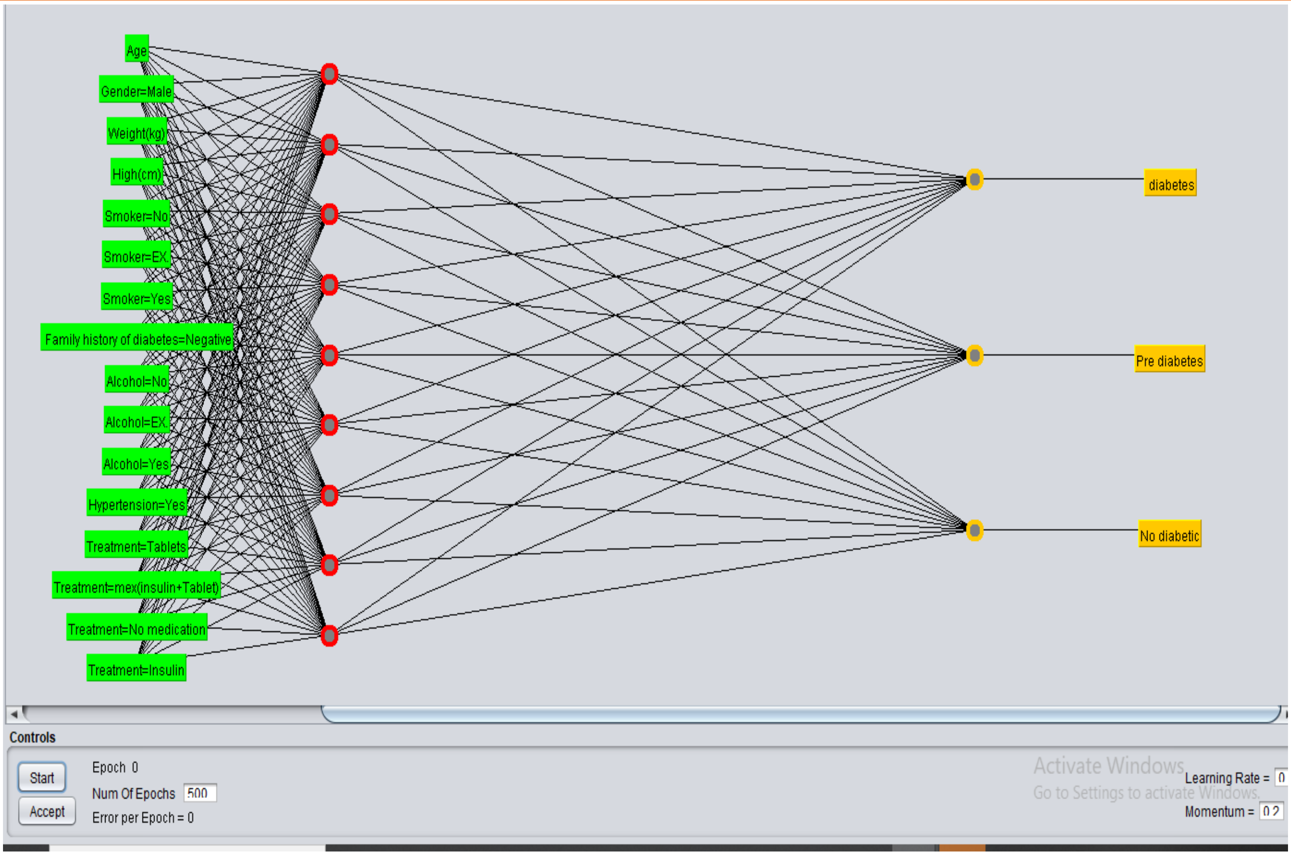

Figure (8): This diagram describes Multilayer Perceptron for 3 classes

5th: Conclusions

This study considered a data set of diabetes mellitus collected from 192 registers in Sulaimani Diabetes Hospital. Two machine learning classifiers, polynomial logistic regression (MLR) and (MLP), were applied using (k-fold) cross-validation with k = 5, 7, 9, and 10 to determine their performance. The results showed that MLR produced the highest prediction performance with the highest accuracy (90.10% at k = 9). The MLP classifier has lower accuracy (88.0208%). From the results, it is also observed that the polynomial logistic regression (MLR) compared to the multilayer perceptron classification algorithm was superior in terms of all other performance parameters such as accuracy, recall and F1 score.

6th: Recommendation

Despite achieving high accuracy, Multinomial Logistic Regression (MLR) was unable to distinguish between pre-diabetic and non-diabetic cases because of class imbalance. Future research ought to use methods like class weighting or SMOTE, or gather balanced data. Reliability will increase with sample size, particularly for minority classes. It is advised to adjust hyperparameters and test additional models (such as Random Forest and SVM). Evaluation and interpretation can be improved by using metrics such as the F1 score and explainable AI tools (e.g., SHAP, LIME).

Reference

1- A.3 Multinomial logistic regression | Notes for Predictive Modelling
2- Ahmed, H., M. B. Mohammed, and I. A. Baba. "On Comparing Multi-Layer Perceptron and Logistic Regression for Classification of Diabetic Patients in Federal Medical Center Yola, Adamawa State." *International Journal of Engineering Technologies and Management research* 8, no. 6 (2021): 24-40.
3- Ahuja, Ravinder, Subhash C. Sharma, and Maaruf Ali. "A diabetic disease prediction model based on classification algorithms." *Annals of emerging technologies in computing (AETiC)* 3, no. 3 (2019): 44-52.
4- Ahuja, Ravinder, Vishal Vivek, Manika Chandna, Shivani Virmani, and Alisha Banga. "Comparative study of various machine learning algorithms for prediction of Insomnia." In *Research anthology on machine learning techniques, methods, and applications*, pp. 776-799. IGI Global Scientific Publishing, 2022.
5- American Diabetes Association. "Diagnosis and classification of diabetes mellitus." *Diabetes care* 33, no. Supplement_1 (2010): S62-S69.
6- Care, Diabetes. "Standards of care in diabetes—2023." *Diabetes care* 46 (2023): S1-S267.
7- García Portugués, Eduardo. "Notes for predictive modeling." (2023).
8- https://en.wikipedia.org/wiki/Softmax_function
9- Kasula, Balaram Yadav. "Machine learning applications in diabetic healthcare: A comprehensive analysis and predictive modeling." International Numeric Journal of Machine Learning and Robots 7, no. 7 (2023): 1-9.
10- Reinhardt, Astrid, and Tim Hubbard. "Using neural networks for prediction of the subcellular location of proteins." Nucleic acids research 26, no. 9 (1998): 2230-2236.
11- Rodriguez, German. "Generalized linear models." Lecture notes [acessado em 1 Mai 2010]. Disponível em: http://data. princeton. edu/wws509/notes (2007).
12- Verma, Garima, and Hemraj Verma. "A multilayer perceptron neural network model for predicting diabetes." International Journal of Grid and Distributed Computing 13, no. 1 (2020): 1018-1025.
13- Zou, Quan, Kaiyang Qu, Yamei Luo, Dehui Yin, Ying Ju, and Hua Tang. "Predicting diabetes mellitus with machine learning techniques." Frontiers in genetics 9 (2018): 515.