Decision Support System in Healthcare for Predicting Blood Pressure Disorders

Marwa Mustafa Owess
Department of Natural, Engineering, and Technology Sciences
Arab American University
Ramallah P600, Palestine
m.owess@student.aaup.edu

Amani Yousef Owda*
Department of Natural, Engineering, and Technology Sciences
Arab American University
Ramallah P600, Palestine
amani.owda@aaup.edu

Majdi Owda
Faculty of Data Science
Arab American University
UNESCO Chair for Data Science
Ramallah P600, Palestine
majdi.owda@aaup.edu

2023 International Conference on Information Technology (ICIT) | 979-8-3503-2006-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICIT58056.2023.10226098

***Abstract*— Blood pressure problems including hypertension and hypotension are considered common among the elderly population, especially hypertension. Recently it started being common among younger adults due to many factors related to unhealthy lifestyles, stress, or genetic factors. Besides the fact of considering hypertension as a chronic disease, it is also considered a primary or contributing cause of complicated risky health issues such as strokes, heart diseases, and chronic kidney failure. In many cases, hypertensive patients may not be aware of their problem because they don't experience any symptoms or warning signs. For this reason, it is essential to build a decision support system to identify individuals at high risk of blood pressure problems including raised blood pressure and low blood pressure. This paper proposes a decision support model for predicting blood pressure disorders by using input variables such as sex, age, body mass index (BMI), cholesterol level, heart rate, and glucose level. The decision model helps in early warning of the potential risk of hypertension or hypotension. As a result, people under potential risk are advised to measure their blood pressure regularly and take the needed precautions or medications to avoid or control this health issue. The proposed decision support model is based on using supervised machine learning classification algorithms, mainly Random Forest, Decision Tree, and XGBoost. The experimental results show that the model achieved the best performance when implemented using a Random Forest classifier with a 10-fold cross-validation method, with an accuracy of 85.81%.**

Keywords—Decision support system, blood pressure, classification, machine learning

I. Introduction

The advancement of technology has opened up new horizons for innovation on a large scale and in various fields, even in the decision-making process, especially during this era, with the huge amount of data produced from various information systems. In response to the need to make decisions in some complex situations, which might be based on data of large volume, that needs much analysis and effort to derive conclusions, scientists and researchers have managed to build computerized decision support systems.

The decision support system (DSS) is a knowledge-based computer system designed to facilitate the decision-making processes with a quick response to the changing input variables facing decision-makers [1].

Many businesses in different sectors have implemented computerized decision support systems in their organizations to improve their performance levels and increase their profitability [2].

This paper presents a novel decision-support system in healthcare using three machine learning algorithms to detect blood pressure (BP) disorders. The paper is organized as follows: Section II introduces a literature review of the decision support systems and discusses the related studies in this field. Section III presents the proposed models. Section IV presents the outcomes and results of the proposed models. Section V provides the conclusion and motivations for future directions.

II. Literature Review

Decision Support Systems (DSS) have gained the researcher's interest since the early 1960s when researchers started their studies to identify the key terms that a system needs to have, in order to support the decision-making process [3]. With the eager efforts of researchers and information experts, the decision-support systems started to be implemented and used in multiple areas such as marketing, education, health care, manufacturing, agriculture, and finance [4].

One of the fields in which decision support systems have been widely used in the healthcare and medical field, with the aim of helping healthcare professionals and practitioners to make decisions in complex situations, in light of this, some researchers have identified clinical DSS as computer software designed to assist the health professional in clinical decision making [5], [6]. Several studies and projects demonstrating the successful use of computerized decision support systems in the clinical field have been published. The authors of [7] proposed a clinical decision support model for identifying hospitalized patients who are at high risk of deep vein thrombosis without prophylactic treatment, by which the responsible doctor receives an alert of deep-vein thrombosis risk, based on the patient data in the hospital database. The model has proven its success in preventing deep-vein thrombosis. The study in [8] discussed the benefit of combining a decision support system with computerized physician order entry, for preventing the adverse effects of medications for patients with kidney impairment, by collecting retrospective cohort data for kidney patients from six hospitals in Boston and combining the collected data with pharmacy and lab data and renal dose checking. Similarly, in a cross-sectional, randomized study conducted in a Belgian university hospital intensive care unit, It was shown that patients with renal failure in a clinical decision support system with computerized physician order entry settings were threefold less likely to commit medication prescription errors than those in a paper-based setting in the hospital [9]. A computerized decision support system can be used as a medical intervention and recommendation tool to improve healthcare, as concluded by the study in [10], which was conducted on hospitalized patients with renal insufficiency, to examine the effectiveness of combining a computerized order entry system with a system application that adjusts the drug doses and frequency for those patients, in improving drug prescriptions and patient outcomes. As a result of using the computerized application intervention, clinically meaningful increases were observed in the proportion of prescriptions considered appropriate for renal insufficiency patients. The study in [11] discussed the effect of using a clinical decision support system as an alert system to improve compliance with the clinical guidelines for measuring plasma concentration during the therapy for a hospital in patients who were treated with gentamicin or vancomycin antibiotics. Improved compliance with guidelines and an optimized treatment with gentamicin and vancomycin for inpatients were achieved when applying the proposed clinical decision support system along with the daily reviewed patient list.

The functionalities provided by clinical decision support systems are not limited to patient safety, clinical management, or diagnostics support, it includes a wide scope of functions such as alarm systems, disease management, prescription, drug control, Cost containment, Patient decision support, and much more [12]. Using clinical decision support systems in their various functions can contribute to the healthcare sector in many ways, first, improving the quality of healthcare and patients' safety by reducing prescription errors and drug side-effects, as well as the direct following of evidence-based clinical instructions, increasing effectiveness by the faster process of orders, reducing medical test repetition and drug side-effects, adjusting drug consumption pattern, and drugs types to reduce the healthcare costs, finally improvement of healthcare professionals medical knowledge by the ease of access to scientific resources, presentation of reminders, and providing useful and critical information to desirable decision-making with minimum errors [13]. In contrast, decision support systems may have drawbacks and might be harmful if it is not implemented in a proper manner, not well integrated into existing workflows, or if users are not well-trained on using the system [14]. However, adopting a computerized clinical decision support system in healthcare centers might have challenges from several sides including financial challenges due to the expensive setup in terms of capital and human resources as well as the long-term cost-effectiveness that cannot be guaranteed. From the system maintenance side, maintaining the CDSS can be challenging as practice changes and knowledge rules change. Additionally, the challenges of CDSS transportability and interoperability related to integration issues with other hospital systems make it difficult to be disseminated and scaled with otherwise high-quality systems [15], [16].

In this study, a simple clinical decision support system that aims to predict the risk of blood pressure diseases was implemented using three different supervised machine learning algorithms.

Previous studies proposed several models to predict or estimate blood pressure, the proposed models were based on different types of variables and different approaches. The study in [17] proposed a model for predicting systolic and diastolic blood pressure based on regular biological measurements collected by a CM400 device and using Classification and Regression Trees machine learning algorithms with utilizing the cross validation method for better performance compared to other similar models that were based on using linear regression, ridge regression, the support vector machine and neural network , another study proposed a model for predicting increased blood pressure using classification tree machine learning technique based on weight, height, waist and hip circumference physical measurements collected from two groups of both sexes, the proposed model has good performance results compared to other existing models implemented using traditional logistic regression algorithm, due to the power of the classification tree technique in finding the cut-off values of the independent variables that maximize the prediction power of the target variable, however the authors of the proposed model acknowledged that the model results cannot be generalized as it has some limitations that need to be handled related to the sampling methodology in the data collection phase, but the proposed model can be adopted for development [18]. The authors of [19] proposed a machine learning model that is based on using artificial neural networks and variables related to (BMI, age, exercise, alcohol, and smoking) to predict systolic blood pressure only, which may contribute to detect hypertension and cardiovascular disease risks at an early stage and to support the healthcare worker in obtaining more accurate blood pressure measurements, the model performance was not outstanding but within the acceptable rate. The researchers in [20] presented a study for predicting hypertension using a large clinical dataset with more than 200,000 records containing a group of hypertension patients and another group of patients who are not having hypertension or other chronic diseases, the used dataset includes a set of significant risk factors such as gender, BMI, blood pressure measurements, LDL, HDL, triglyceride, and cholesterol blood tests results, the model was trained using ANN algorithm, the model was successful predicting hypertension at the accuracy of 82%. The author of [21] discussed the use of machine learning techniques for predicting hypertension and estimating blood pressure, the types of different datasets used in previous studies as ECG and PPG signals, demographic and clinical datasets, or using datasets combining both clinical and physiological signals data. The study in [22] proposed a model for early prediction of diabetes and hypertension using different machine learning techniques including DBSCAN clustering algorithms for detecting and handling outliers, SMOTE algorithm for handling class imbalance issues, and Random Forest for classifying detected cases, the model was applied on three different datasets collecting different risk factors that have been used by some previous studies, the highest accuracy was obtained by applying the model was 92.56% for one of the tested datasets. The authors of [23] presented a machine learning model for predicting hypertension and cardiovascular disease using a dataset collected by an ongoing study that includes personal interviews, physical examinations, and lab test results. The proposed model in this study have implemented by testing different machine learning classification algorithms such as logistic regression, SVM, random forest, and gradient boosting trees, then the tested algorithms were combined to implement a weighted ensemble model which contribute significantly to improving the accuracy results to be about

84%. The authors of [24] proposed an image-based deep learning approach for predicting hypertension, the increase in blood glucose levels, and the unbalanced levels of lipids in blood by using a cross-sectional dataset from retinal fundus images, the experimental results in this study concluded that using this approach was useful and effective in predicting diabetes, hypertension, and dyslipidemia diseases. The authors in [25] presented a linear regression model for blood pressure prediction using the measures of human pulse transit time (PTT), the proposed model yielded high accuracy and low error rates in estimating systolic and diastolic blood pressure with better results in predicting the diastolic blood pressure which its estimation was based on the systolic component. The study in [26] proposed a model for estimating blood pressure using data collected from fingerprint oximeter devices and using neural networks, although the experimental results were satisfactory, the proposed model has a limitation regarding real-time processing due to the slow learning pace of the underlying neural network used in the model. The authors of [27] proposed a machine learning model for blood pressure prediction, the proposed model was based on using recurrent neural networks with contextual layers and applied to data consisting of two parts, historical measurements, and contextual data. The historical measurements include systolic and diastolic blood pressure, heart rate, time of the measurement, and taking drugs, which was collected using a wireless blood pressure monitor that is synchronized with a mobile device and sent to a server in the cloud, the contextual data includes the BMI, gender, and age. While the model is reasonably accurate in predicting blood pressure, it requires historical data to be accumulated on the cloud server, which may burden the model's security.

The proposed model in this paper focused on predicting blood pressure diseases based on variables related to sex, age, BMI, cholesterol level, heart rate, and glucose level. It also investigated the correlations between these variables and their effect on blood pressure. Identifying potential blood pressure patients will contribute to the early detection and prevention of possible future cases and thus provide the needed intervention and medication, besides timely management of the related risk factors. The main contribution of the proposed model is the prediction and early detection of hypertension or hypotension patients, by building a machine-learning model that uses the main causative risk factors for these diseases as inputs for model learning. This will improve the health of individuals and society, reduce the premature death rate, and the treatment costs related to these diseases and their complications, assist healthcare providers to detect potential hypertension and diabetes patients automatically, and provide them with the attention and interventions they need to monitor their health. Several previous works proposed prediction models for detecting or estimating hypertension using machine learning and data mining techniques. Most of these models were proposed using data collected from live monitors connected to patients or historical data of multiple readings of blood pressure for a patient, other studies used data from ongoing cohort studies for building the model, also in previous studies the focus was on hypertension disease and the datasets in most cases were collected from hypertension patients. In this work, the proposed model aimed to predict disorders in blood pressure generally, not only hypertension in particular, using a simple dataset collecting the main risk factors contributing to this health issue, related to a person's demographics, physical measurements, and medical history.

III. Methodology

This section introduces a description of the used dataset and the detailed methodology used for building the blood pressure decision support model using supervised machine learning classification techniques. The working methodology consists of two important parts, first the data cleaning for handling any possible issues in the data, to produce well-prepared data for the next step of building the blood pressure (BP) disorders prediction model. The data preparation tasks and building the blood pressure model are done with Python, Python is one of the topmost common powerful, open-source programming languages widely used in data science applications due to its flexibility and its extensive collection of libraries for creating and managing data structures.

A. Dataset Description

The proposed models were demonstrated to a public dataset obtained from Kaggle, the "Framingham Heart Study" dataset, which is an ongoing cohort study, conducted in Framingham, Massachusetts, in cooperation between Boston University and National Heart, Lung, and Blood Institute in the USA, to identify characteristics that contribute to cardiovascular disease among adults [28]. The obtained dataset is structured data, involving over 4000 observations and 16 features, the features in the dataset are of different types, including categorical features and continuous numeric/quantitative features. The data collected includes the participants' age, gender, smoking status, and clinical data related to the participant's medical history, whether they take blood pressure medications or had previously high blood pressure, diabetes, or stroke, the heart rate, glucose, systolic and diastolic blood pressure measurements.

B. Data Preparation

Data preparation is an important step for modeling, which is needed to handle any issue in the data before applying the machine learning models, to obtain the best results by using clean and consistent data. As data quality is an important factor for producing useful and high-performance models, the first step used in implementing the proposed model is pre-processing the used Framingham dataset. The data preparation and pre-processing were applied to the Framingham dataset using the following steps.

- Handling missing values: by exploring the Framingham dataset it is found that it contains a number of missing values, the missing values related to clinical test results were handled by dropping their related observation from the dataset, and for other status columns such as the smoking status the missing values were handled by filling with zero.
- Checking outliers: the Framingham dataset contains medical data including features for lab test results, some of these features contain abnormal values. The visual approach was used for detecting the outlier values using the boxplots, these outlier values might be real values, and indicate possible health issues, so these values were not dropped from the dataset.
- The blood pressure in the used dataset consists of two features, systolic blood pressure, and diastolic blood pressure, which are used usually to measure blood pressure. Systolic blood pressure represents the pressure in the arteries when the heart beats or in other words when pushing the blood out from the heart.

While diastolic blood pressure is the pressure when the heart rests between beats. In this work, for modeling the blood pressure disorders prediction DSS, these two features were used to derive a new feature for the disorder of blood pressure, which is also the target feature in the proposed model, to indicate whether the blood pressure is within the normal range or not. The WHO hypertension guidelines are used to identify whether the blood pressure is within the normal range or not, which identifies hypertension when the systolic blood pressure is equal to or greater than 140 (mmHg) and the diastolic blood pressure is equal to or greater than 90 (mmHg) [29]. While for hypotension (low blood pressure) the systolic blood pressure is less than 90 (mmHg) and the diastolic blood pressure is less than 60 (mmHg) [30].

- Features selection and elimination: multiple methods were used for obtaining and selecting the most important features for modeling the BP disorders DSS. This step is highly important and necessary, as insignificant features can decrease the predictive power and learning rate of the machine learning algorithms [31], [32]. The first used approach is calculating the correlation matrix for all the features in the dataset and eliminating correlated features from the independent features list with correlation coefficients greater than 0.7, as highly correlated variables can affect the machine learning models stability and results. In this stage, the smoking status and number of cigarettes per day features are highly correlated, and the smoking status feature was eliminated. Also, the systolic and diastolic blood pressures were eliminated from the dataset as they are already used for deriving the blood pressure disorders feature. The second method used in feature elimination is checking the variables importance using Random Forest classifier, to identify the most significant independent features that contribute more to the predictive power of the model in classifying the target-dependent feature related to blood pressure disorder.
- Features scaling using normalization: the numerical features in the Framingham dataset are from different numerical scales, which may increase the computation difficulty during the predictive modeling process. Thus, the min-max normalization technique was applied to the numerical features in the dataset to reduce the complexity of the model computation process, by normalizing the features to be in the range between 0 and 1, according to the following normalization equation [33]:

$$Xnorm = \frac{\mathrm{X} - \mathrm{Xmin}}{\mathrm{Xmax} - \mathrm{Xmin}} \quad (1)$$

- Handling class imbalance: By exploratory data analysis for the processed dataset, it is found that the target feature in this work related to blood pressure disorder, has a class imbalance issue, having only 34.7% of the observations in the dataset holding value of 1 indicating blood pressure disorder and 65.3% of the observation with no disorder. This may affect the learning of the machine learning algorithms, as it is known that class imbalance is considered a major issue in machine learning classification algorithms [34]. To handle this issue, the overall sampling technique was used, by replicating the dataset observation from the minority class (BP disorder) to the same size as the majority class (no BP disorder) [35].

The final dataset dimensions after applying the cleaning steps and preparing the data are 11 features and 4896 records, the processed dataset is ready for use in the next stage of modeling the blood pressure DSS.

C. Modelling Blood Pressure Disorders Using DSS

This section illustrates the used machine learning algorithms for implementing the blood pressure disorder prediction model. The blood pressure prediction model is a binary classification model used to identify whether a patient may suffer from a disorder in blood pressure or not, by processing and analyzing the related risk factors that are represented as features in the used dataset. The main machine learning algorithms that were used for implementing the proposed blood pressure DSS is Random Forest, Decision Tree, and XGBoost algorithm, in addition to using cross-validation technique. Fig. 1 depicts the workflow of the proposed BP disorder prediction model.

Random Forest is one of the most powerful supervised machine-learning algorithms used for classification problems, that is based on generating a forest of multiple decision trees for a subset of the features, where the features in each tree are selected randomly, then the prediction results of all the decision trees are combined to get the final refined output. By using this ensemble approach, the Random Forest makes an advantage over the Decision Tree algorithm which has a known issue of high variance that is minimized using Random Forest [36].

Decision Tree is a supervised machine learning algorithm that can be used for building classification and regression models. Decision Tree is considered one of the simplest machine learning algorithms also it is widely used for building prediction models, its working methodology is based on a tree structure in that the features are organized as a tree, with splitting the features recursively according to selected impurity criteria that might be the entropy measure, Gini index, or information gain value [37].

XGBOOST algorithm which stands for Extreme Gradient Boosting, is a supervised-learning algorithm that is an improved version of the Gradient Boost algorithm, used commonly for regression and classification on large datasets due to its high efficiency in terms of getting accurate results and fast execution time. XGBOOST is a tree-based algorithm that works in a sequential manner, which means the outcome of a processed tree is passed into and used by the next tree [38].

K-Fold Cross Validation is a machine learning technique that works by random shuffling of the dataset and then dividing the dataset into a number of groups (K parameter), each fold is split into training and test data, and the model is applied to each fold separately, the results from all folds are combined to get the final result of the model. Using this approach can improve the performance of the model and reduce the likelihood of an overfitting problem, by training the model on unseen data, which as a result takes more time in training the model and need high computation power [39].

D. Experimental Parameters Settings

The hyperparameter configurations of each machine learning algorithm used in this work including the Decision Tree, Random Forest, and XGBoost are illustrated in Table I. The GridSearch approach was used to find the parameter settings that achieve the best performance for each algorithm.

TABLE I. BP DISORDER MODELS PARAMETERS CONFIGURATION.

Algorithm	Parameters Settings
Random Forest	n_estimators=3000, criterion="gini", max_depth=300, min_samples_split=2, min_samples_leaf=1
Decision Tree	criterion="gini", max_depth=22, min_samples_split=2, min_samples_leaf=1, random_state=40, class_weight= {0: 0.3, 1: 0.7}
XGBoost	max_depth=3, n_estimators=5000, learning_rate=0.1

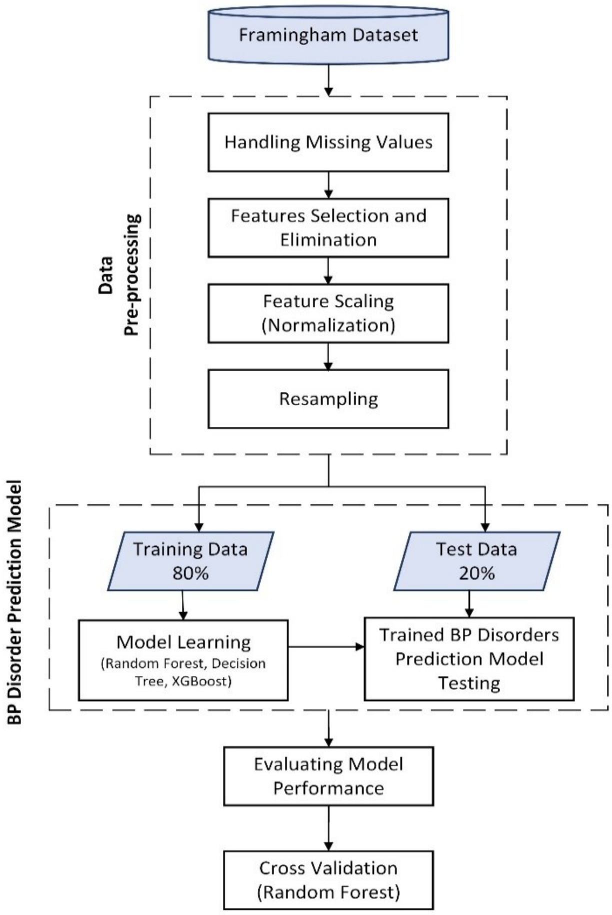

Fig. 1. BP Disorder Prediction Model Workflow.

IV. RESULTS AND DISCUSSION

A. Experimental Results

The performance of the used classification algorithms for building the BP disorder model is evaluated using various metrics. These include accuracy, precision, recall, F1-score, and ROC-AUC. This study utilized three robust supervised machine learning classifiers: Random Forest, Decision Tree, and XGBoost. Fig. 2 compares the ROC curves of the Random Forest, Decision Tree, and XGBoost classifiers. The AUC for Random Forest is 0.83, 0.80 for XGBoost, and 0.79 for Decision Tree. The BP disorder model using Random Forest classifier shows a higher AUC than the other used classifiers, which indicates that it is more suitable for the BP disorder prediction model than Decision Tree, and XGBoost algorithms.

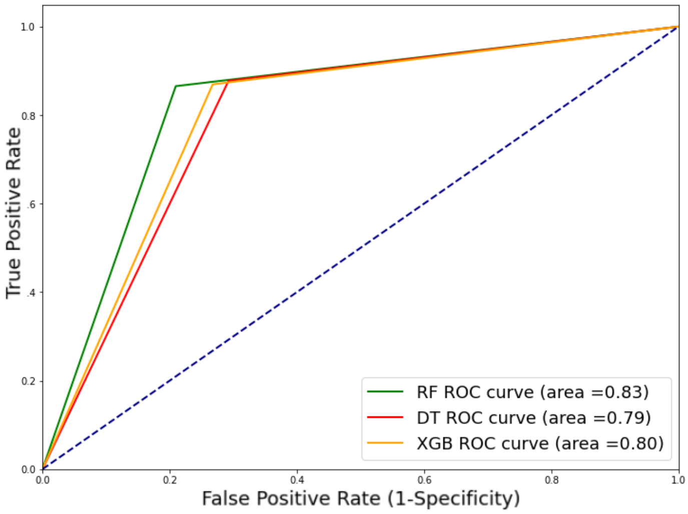

Fig. 2. BP Disorder Prediction Models ROC Curves Comparison

Table II presents the results of the performance evaluation using various measures including accuracy, recall, precision, and F1 Score. The BP disorder model using Random Forest algorithm has an accuracy score of 82.86%, which is higher than that of the XGBoost, and the Decision Tree. As a result, the Random Forest algorithm has better performance in predicting the disorders of blood pressure among individuals compared to the other algorithms.

TABLE II. PERFORMANCE EVALUATION.

Algorithm	Accuracy	Recall	Precision	F1-Score
Random Forest	82.86%	86.55%	81.02%	83.69%
XGBoost	80.20%	86.95%	77.05%	81.70%
Decision Tree	79.39%	87.75%	75.61%	81.23%

From the experimental results, it is obvious that the Random Forest model has achieved the best performance results in terms of accuracy, precision, and F1-score metrics. Therefore, the Random Forest algorithm was selected to be used for training the model another time in combination with the cross-validation technique. The evaluation results of applying the K-fold cross-validation technique on the Random Forest model give a slight improvement in the model performance when using K=5 with a mean accuracy of 83.31%, and higher performance when using 10 folds cross-validation that resulted in a mean accuracy of 85.81%.

V. CONCLUSION AND FUTURE WORK

In this research, a DSS was proposed for predicting blood pressure disorders using a selected set of powerful supervised machine learning classification algorithms, based on some related variables represented by risk factors in the used dataset. The proposed models are implemented using three main classification algorithms, Random Forest, Decision Tree, and XGBoost. The implemented model using Random Forest classifier yielded the best accuracy result of 82.86%, the accuracy of the Random Forest model improved when it was implemented in combination with 10-fold cross-validation to be 85.81%. As a plan for future work, it is recommended to use a bigger dataset that includes more related risk factors and features and to use more advanced machine learning algorithms such as deep learning to enhance the results. The importance of having such a decision model with high accuracy is supporting the healthcare providers and reducing the workload to improve patient health.

REFERENCES

[1] A. Yanie, A. Hasibuan, I. Ishak et al., "Web based application for Decision Support System with electre method," Journal of Physics: Conference Series, vol. 1028, no. 1, 2018.

[2] M. D. Goslar, G. I. Green, and T. H. Hughes, "Applications and implementation decision support systems: An empirical assessment for decision making," Decision Sciences, vol. 17, no. 1, pp. 79–91, 1986.

[3] S. Eom and E. Kim, "A survey of decision support system applications (1995–2001)," Journal of the Operational Research Society, vol. 57, no. 11, pp. 1264–1278, 2006.

[4] H. B. Eom and S. M. Lee, "Decision support systems applications research: a bibliography (1971–1988)," European Journal of Operational Research, vol. 46, no. 3, pp. 333–342, 1990.

[5] E. H. Shortliffe, "Computer programs to support clinical decision making," JAMA: The Journal of the American Medical Association, vol. 258, no. 1, pp. 61–66, 1987.

[6] I. Sim, P. Gorman, R. A. Greenes et al., "Clinical decision support systems for the practice of evidence-based medicine," Journal of the American Medical Informatics Association, vol. 8, no. 6, pp. 527–534, 2001.

[7] N. Kucher, S. Koo, R. Quiroz et al., "Electronic alerts to prevent venous thromboembolism among hospitalized patients," New England Journal of Medicine, vol. 352, no. 10, pp. 969–977, 2005.

[8] B. L. Hug, D. J. Witkowski, C. M. Sox et al., "Occurrence of adverse, often preventable, events in community hospitals involving nephrotoxic drugs or those excreted by the kidney," Kidney International, vol. 76, no. 11, pp. 1192–1198, 2009.

[9] K. Colpaert, B. Claus, A. Somers et al., "Impact of computerized physician order entry on medication prescription errors in the intensive care unit: a controlled cross-sectional trial," *Critical Care*, vol. 10, no. 1, pp.1–9, 2006.

[10] G. M. Chertow, J. Lee, G. J. Kuperman et al., "Guided medication dosing for inpatients with renal insufficiency," JAMA, vol. 286, no. 22, pp. 2839–2844, 2001.

[11] M. Pereboom, I. J. Mulder, S. L. Verweij, R. T. M. van der Hoeven, and M. L. Becker, "A clinical decision support system to improve adequate dosing of gentamicin and vancomycin," International Journal of Medical Informatics, vol. 124, pp. 1–5, 2019.

[12] R. T. Sutton, D. Pincock, D. C. Baumgart et al., "An overview of clinical decision support systems: Benefits, risks, and strategies for Success," npj Digital Medicine, vol. 3, no. 1, 2020.

[13] A. M. Shahsavarani, E. Azad Marz Abadi, M. Hakimi Kalkhoran, S. Jafari, and S. Qaranli, "Clinical decision support systems (CDSSs): state of the art review of literature," International Journal of Medical Reviews, vol. 2, no. 4, pp. 299–308, 2015.

[14] R. S. Castillo and A. Kelemen, "Considerations for a successful clinical decision support system," CIN: Computers, Informatics, Nursing, vol. 31, no. 7, pp. 319–326, 2013.

[15] J. Eichner and M. Das, "Challenges and barriers to clinical decision support (CDS) design and implementation experienced in the Agency for Healthcare Research and Quality CDS demonstrations," Rockville: Agency for Healthcare Research and Quality, 2010.

[16] W. Bonney, "Impacts and risks of adopting clinical decision support systems," in Efficient Decision Support Systems: Practice and Challenges in Biomedical Related Domain, Croatia: INTECH Open Access Publisher, 2011, pp. 21–30.

[17] B. Zhang, Z. Wei, J. Ren, Y. Cheng, and Z. Zheng, "An empirical study on predicting blood pressure using classification and regression trees," IEEE access, vol. 6, pp. 21758–21768, 2018.

[18] H. F. Golino, L. S. D. B. Amaral, S. F. P. Duarte et al., "Predicting increased blood pressure using machine learning," Journal of obesity, vol. 2014, 2014.

[19] T. H. Wu, G. K. H. Pang, and E. W. Y. Kwong, "Predicting systolic blood pressure using machine learning," 7th International Conference on Information and Automation for Sustainability, 2014.

[20] D. LaFreniere, F. Zulkernine, D. Barber, and K. Martin, "Using machine learning to predict hypertension from a clinical dataset," in 2016 IEEE symposium series on computational intelligence (SSCI), pp. 1–7, 2016.

[21] E. Martinez-Ríos, L. Montesinos, M. Alfaro-Ponce, and L. Pecchia, "A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data," Biomedical Signal Processing and Control, vol. 68, 2021.

[22] M. F. Ijaz, G. Alfian, M. Syafrudin, and J. Rhee, "Hybrid prediction model for type 2 diabetes and hypertension using DBSCAN-based outlier detection, synthetic minority over sampling technique (SMOTE), and random forest," Applied Sciences, vol. 8, no. 8, 2018.

[23] A. Dinh, S. Miertschin, A. Young, and S.D. Mohanty, "A data-driven approach to predicting diabetes and cardiovascular disease with machine learning," BMC medical informatics and decision making, vol. 19, no. 1, pp. 1–15, 2019.

[24] L. Zhang, M. Yuan, Z. An et al., "Prediction of hypertension, hyperglycemia and dyslipidemia from retinal fundus photographs via deep learning: A cross-sectional study of chronic diseases in central China," PloS one, vol. 15, no. 5, 2020.

[25] Y. Y. Hsieh, C. D. Wu, S. S. Lu, and Y. Tsao, "A linear regression model with dynamic pulse transit time features for noninvasive blood pressure prediction," in 2016 IEEE biomedical circuits and systems conference (BioCAS), pp. 604–607, 2016.

[26] C. Sideris, H. Kalantarian, E. Nemati, and M. Sarrafzadeh, "Building continuous arterial blood pressure prediction models using recurrent networks," In 2016 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 1–5, 2016.

[27] X. Li, S. Wu, and L. Wang, "Blood pressure prediction via recurrent models with contextual layer,". In Proceedings of the 26th International Conference on World Wide Web, pp. 685–693, 2017.

[28] National Heart Lung and Blood Institute, "BIOLINCC: Framingham heart study-cohort (FHS-cohort)," 2021, https://biolincc.nhlbi.nih.gov/studies/framcohort/.

[29] World Health Organization, "Hypertension," 2023, https://www.who.int/news-room/fact-sheets/detail/hypertension.

[30] L. Tritapepe, C. Nencini, G. Frasacco, and D. Tallarico, "Hypotension," in Textbook of Echocardiography for Intensivists and Emergency Physicians, United States: Springer, 2019, pp. 275–282.

[31] E. M. Karabulut, S. A. Özel, and T. Ibrikci, "A comparative study on the effect of feature selection on classification accuracy," Procedia Technology, vol. 1, pp. 323–327, 2012.

[32] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," in Data classification: Algorithms and applications, United States: CRC Press, 37, 2014, pp. 37–64.

[33] V. G. Raju, K. P. Lakshmi, V. M. Jain, A. Kalidindi, and V. Padma, "Study the influence of normalization/transformation process on the accuracy of supervised classification," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 729–735), 2020.

[34] P. Yildirim, "Chronic kidney disease prediction on imbalanced data by multilayer perceptron: Chronic kidney disease prediction," 2017 IEEE 41st annual computer software and applications conference (COMPSAC), vol. 2, pp. 193–198, 2017.

[35] Z. Zheng, Y. Cai, and Y. Li, "Oversampling method for imbalanced classification," Computing and Informatics, vol. 34, no. 5, pp. 1017–1037, 2015.

[36] A. Parmar, R. Katariya, and V. Patel, "A review on random forest: An ensemble classifier," in International Conference on Intelligent Data Communication Technologies and Internet of Things, pp. 758–763, 2018.

[37] Y. Y. Song and L. U. Ying, "Decision tree methods: applications for classification and prediction," Shanghai archives of psychiatry, vol. 27, no. 2, pp. 130–135, 2015.

[38] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785–794, 2016.

[39] T. T. Wong and P. Y. Yeh, "Reliable accuracy estimates from k-fold cross-validation," IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 8, pp. 1586-1594, 2019.