Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Optimizing diabetes prediction with MLP neural networks and feature selection algorithms

Majd Mohammad A. Al-Hawamdeh[a*]

[a]*School of Computer Sciences and Information Technology, Department of Computer Science, Jerash University, Jordan*

C H R O N I C L E

Article history:
Received: July 18, 2024
Received in revised format: August 10, 2024
Accepted: August 22, 2024
Available online: August 22, 2024

Keywords:
Diabetes forecast
Multilayer Perceptron Neural Network (MLPNN)
Memetic algorithm (MA)
Arithmetic Optimization Algorithm (AOA)

A B S T R A C T

In this research, the goal was to improve diabetes prediction by combining Multilayer Perceptron Neural Network (MLPNN) with Memetic Algorithm (MA) and Arithmetic Optimization Algorithm (AOA). The method suggested used a preprocessing step to choose a representative subset of attributes from the initial set. Next, the method suggested utilized a combination of the MA and AOA algorithms to optimize feature selection, resulting in a refined dataset that served as input for the Neural Network. Ultimately, the suggested approach utilized the multilayer perceptron neural network (MLPNN) to train the network with hidden layer neurons. The experimental findings indicated a 95% high accuracy rate was achieved. Machine learning classifiers achieved better accuracy compared to classifiers in previous studies, with Decision Tree and Logistic Regression classifiers each reaching 93.57% and 93.33% accuracy, respectively.

1. Introduction

Due to the huge growth of current technology, a large amount of data is generated, especially in the medical field (Turnbull et al., 2004). This data can include the patients' information, labs, imaging studies, and diseases. That is why these data need to be collected, sorted, classified, and analyzed. The process of classification is considered as one of the most important parts, which is difficult when the data being processed has many dimensions; this is where the role of machine learning comes in by mining the data. Approximately (422) million individuals globally are diagnosed with diabetes, with the majority residing in low-income nations, and diabetes directly leads to (1.6) million deaths annually. The number of individuals with diabetes grew from (108) million in (1980) to (422) million in (2014). From 2000 to 2016, there was a 5% rise in premature death due to diabetes. The occurrence of diabetes and the number of cases have both been on the rise in recent decades. In Jordan, diabetes affects (12.9%) of males and (13.5%) of females, and is responsible for (7%) of all fatalities. Data extraction should be done through a specific method while taking into consideration the precise features needed from the dataset. The benefit of data selection is to improve prediction and classification while at the same reducing the volume of the given data. Many risk factors contribute to the development of Diabetes Mellitus (DM) including obesity, family history, race or ethnicity, or age. The mined medical data will be used by healthcare professionals to reach diagnostic decisions which will eventually contribute to lowering the overall mortality rate. The mortality rate caused by DM can be reduced by using preventative countermeasures including lifestyle modifications, dietary changes, and exercise for instance (Al-Hawamdeh, & Alshaer, 2022). There are a lot of studies related to the development of the Machine Learning (ML) methods which focus on the prediction and early diagnosis of DM. However, many improvements need to be done to increase the prediction accuracy most of which previously focused only on feature preprocessing. In this study, a benchmark dataset from the UCI repository will be used. To realize the highest accuracy, the dataset will undergo preprocessing which will determine the features needed to achieve the most accurate prediction. This can be done by developing a hybrid algorithm consisting of a multi-layer perception neural network with backpropagation (MLP) neural network (NN) with Memetic

* Corresponding author.
E-mail address Majd.alhawamdeh@jpu.edu.jo (M. M. A. Al-Hawamdeh)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2024.8.021

algorithm (MA) (Jaradat et al., 2018) and Arithmetic Optimization Algorithm (AOA) (Abualigah et al., 2021) can be adapted for features selection problem within the prediction model. Such adaptation results in MLPNN-MA and MLPNN-AOA to filter irrelevant features, discover the most accurate features, via data ranking and correlation, and thus enhance the prediction model's accuracy.

With the main objective of the suggested research in mind, the following interconnected research inquiries will shape the study's structure:

1. How do the proposed model (MA) and (AOA) help to select the most pertinent features?
2. How does the proposed model help get rid of the problems of overfitting and underfitting in the prediction model for DM?
3. How do the MA and AOA affect the accuracy of MLPNN for the DM prediction model?

Drawing from the research inquiries, the aims of this investigation are outlined as follows:

- To develop algorithms that combine a multi-layer perceptron neural network and backpropagation (MLP) neural network (NN) with two optimization algorithms, MA and AOA, to preprocess the dataset and determine which features have the greatest impact on DM classification between these algorithms, and use it to predict the presence or absence of diabetes.
- To utilize the MLPNN-MA and MLPNN-AOA algorithms to discard the problem of overfitting and underfitting the DM prediction model.
- To evaluate the MLPNN-MA and MLPNN-AOA algorithms by many measurement tools to choose which one is more accurate than the other in the classification task.

Making a diagnosis, predicting a prognosis, applying a treatment plan, and preventing disease are the cornerstones that constitute what the medical field is all about. Therefore, compelling results can be produced using ML techniques in the medical field which will significantly help reduce the cost of diagnostic tests (Alzubi et al., 2016). Develop algorithms consisting of a multi-layer perception neural network and backpropagation (MLP) neural network (NN) with two optimization algorithms named MA and AOA to preprocess the dataset and determine which features have the biggest effect between these algorithms by which one has high accuracy, and use it on the prediction of presence or absence of diabetes. In the proposed model, an MLP neural network (NN) with MA and AOA will process the given data and select the basic attributes that can affect the accuracy of the DM prediction model.

2. Literature review

Diabetes mellitus is a highly risky illness that has no cure. If the person is impacted by this illness, it will be a lifelong condition. Additionally, an excess of glucose in the bloodstream can lead to health issues. The widely recognized forms of diabetes mellitus include: a) type 1 diabetes mellitus; b) type 2 diabetes mellitus, and c) gestational diabetes (Rajeswari & Prabhu, 2019). The sheer amount of research done in this field presents two main obstacles for researchers and developers aiming to construct models for forecasting type 2 diabetes. Initially, there was notable diversity in the machine learning (ML) methodologies employed in earlier research, making it challenging to identify the most effective one. Secondly, the lack of clarity on the processes utilized to train models hinders their interpretability, a crucial aspect for medical professionals. (Fregoso-Aparicio et al., 2021). The emergence of artificial intelligence and related technologies has led to the application of computational methods in real-time detection models across various fields. The complexity of learning new methods that can enhance existing methods has been greatly reduced with the utilization of data mining, deep learning, machine learning, and computer vision technologies (Sharma & Shah, 2021). Machine Learning is a division of AI where the machine tries to forecast a result using certain data and past outcomes. There exist two categories of ML. Supervised learning involves data serving as a teacher to form the model around the dataset. Unsupervised learning is the second type, where data is self-trained to identify and categorize patterns within the dataset. (Saxena et al., 2022). Classification is typically necessary for organizing the large amount of business and health data sets. Classification is a type of data mining that organizes items in a group into specific categories. Achieving the anticipated level of accuracy involves categorizing the data sets of individuals with diabetes (Abu-Alaish et al., 2021). For instance, SVM, J48, Naive Bayes, Logistic Regression (LR), Decision Tree (DT), artificial neural network (ANN), and so on are more effective in diagnosing different illnesses (Rajeswari & Prabhu, 2019).

The primary goal is to categorize the data as either diabetic or non-diabetic and enhance classification accuracy. The focus of machine learning in diabetes diagnosis is primarily on analyzing patterns within the diabetes dataset that will be supplied. Machine learning has continuously advanced as a trusted and supportive technology in healthcare recently (Saxena et al., 2022). Choosing the right features and classifier is the key challenge in the ML approach (Khanam & Foo, 2021).

2.1 Diabetes Prediction using Classical Machine Learning

Multiple studies and research have focused on improving diabetes prediction through traditional machine learning techniques. This segment introduces the latest research focusing on diabetes prediction using classical ML methods like KNN, RF, and SVM. In the same setting, in the study by Jaggi et al. (2021), they introduced a model designed to recommend an expert system capable of accurately predicting if a patient does/does not have diabetes. The suggested approach involved the application of an artificial neural network (ANN) with six dense layers. The findings indicated that the model had an accuracy rate of (77 %) in its predictions. The researchers determined that the utilized model is very effective and dependable. In the research by Darabi and Tarokh (2018), a model was suggested to assess the likelihood of developing diabetes mellitus using data from lab tests, lifestyle, and family background, utilizing machine learning algorithms. The study tested eight different ML algorithms: LR, Nearest Neighbor, Decision Tree (DT), RF, SVM, Naive Bayesian, KNN, and Gradient Boosting. The findings indicated that the model utilizing the gradient boosting algorithm exhibited the highest level of performance, achieving a prediction accuracy of 95.50%. It was determined that this model is suitable for diagnosing diabetes. Table 1 displays a comparison of prior research including the methods, dataset, and accuracy achieved in each study.

Table 1
Research summary of Diabetes prediction using ML techniques

Authors	Methods	Dataset	Accuracy
Mujumdar & Vaidehi (2019)	SVM, RF, DT, Extra Tree Classifier, Ada Boost, Perceptron, Linear Discriminant Analysis algorithm, LR, KNN, GNB, Bagging, Gradient Boost	Collected dataset	AdaBoost = 98.8 % Gradient Boost = 98.1 %
Jian et al. (2021)	LR, SVM, decision tree, RF, AdaBoost, and XGBoost.	Rashid Center for Diabetes and Research	ACU= 90.7 %
Alanazi & Mezher (2020)	SVM RF	Security Force Primary Health Care	SVM = 97 % RF = 98 %
Jaggi et al. (2021) Khaleel & Al-Bakry (2023) Gupta & Goel (2023) Gowthami et al. (2024) Wee et al. (2024)	ANN A variety of ML and deep learning techniques	Collected dataset Totally different datasets other than UCI's collected from various medical sources	ACU= 77% Highly accurate with brief description and experimental settings

2.2 Diabetes Prediction using MLP Neural Network

Several studies aimed to enhance the prediction of the diabetes disease using MLPNN. Table (2.2) shows a comparison between previous studies that contains each of the methods used, dataset, and the accuracy reached for each study. In this section, we tried to show the most common methods used for diabetes disease prediction, and how they have used the different classical machine learning methods and the Multilayer Perceptron Neural Network. Table 1 reports a comparative study of the different classical machine learning methods and techniques that are used in general for diabetes disease prediction.

Table 2
Research summary of Diabetes prediction using MLP neural networks

Authors	Methods	Dataset	Accuracy
Bani-Salameh et al. (2021)	MLP	Collected dataset	77.6%
Mohapatra et al. (2019)	MLP	Pima Indian Diabetes (PID)	77.5%
Verma et al., (2020)	MLP	PID	82%
Karthiga et al. (2020)	LR, DT, RF, KNN and ANN – MLP	PID	MLP = 86% LR = 78 % DT = 78 %
Theerthagiri (2021)	KNN, DT, Naive Bayes, Extra Trees, Radial Basis Function, MLP	PID	KNN = 71.7% DT = 66.8% Naive Bayes = 77.2% Extra Trees = 72.4% Radial Basis Function = 68.2% MLP = 80.6%
Bukhari et al. (2021)	ABP-SCGNN	PID	ABP-SCGNN (Training)= 94.3% ABP-SCGNN (Validation)= 92.7%

Table 2 reports the different methods and techniques used in the Multilayer Perceptron Neural Network. A different method for diabetes disease prediction has been used with the different datasets, and we can observe the differences in results between them.

In this research, the new method attempted to create a fresh approach to improve the prediction of diabetes disease. This method aims to improve the efficiency and accuracy of predicting diabetes by utilizing the Memetic algorithm (MA) and the Arithmetic Optimization Algorithm (AOA) for feature selection.

3. Research methodology (MLP-AMA)

This part introduces the method suggested for predicting diabetes with the Multilayer Perceptron Neural Network (MLPNN) optimized by combining the Memetic algorithm (MA) and Arithmetic Optimization Algorithm (AOA) for feature selection. The initial step in the suggested approach involves gathering important features for diabetes prediction to compile the raw dataset. In the following stage, the suggested approach applied preprocessing on the raw data to choose a relevant subset of features from the initial set of attributes. During this stage, the method being proposed made use of the Correlation-based Feature Selection method, Principal Component Analysis, Information Gain Ratio based feature selection, and the Minimum Redundancy Maximum Relevance. The goal of this phase is to create a set that accurately represents the data by selecting only the significant and crucial attributes from the original dataset. In the following stage, the suggested technique employed a combination method to enhance feature selection by alternating between the MA and AOA algorithms. This stage produces an improved dataset that will be used as the input for the Neural Network in the upcoming step. In the end, the method employed MLP neural networks for training by utilizing hidden layer neurons.

3.1 Overall Structure

Fig. 1 displays the structure and overall framework of the suggested approach:

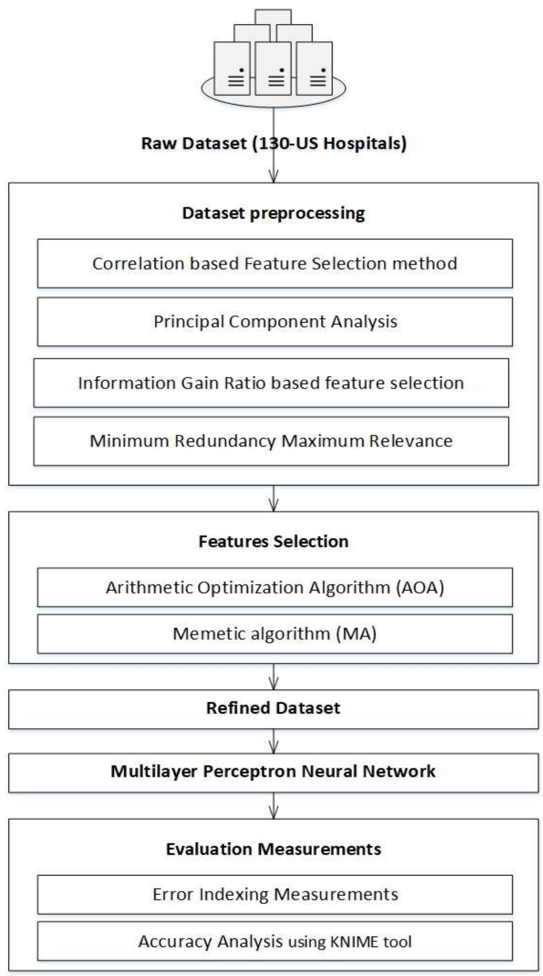

Fig. 1. General framework

Generally, the proposed method contains several phases including raw dataset, dataset preprocessing, features selection, refined dataset, MLPNN, and evaluation measurements.

3.1.1 Dataset

In this research, the method suggested utilized a dataset from 130 US hospitals for the years 1999-2008 focusing on diabetes. The dataset includes a span of (10) years from 1999 to 2008, capturing clinical care at (130) US hospitals and integrated delivery networks. The dataset contains more than (50) characteristics that reflect patient and hospital results. The data includes patient ID, ethnicity, sex, age, admission method, length of hospital stay, admitting physician's specialty, number of lab tests, HbA1c result, diagnosis, number of medications, diabetic meds, previous year's outpatient, inpatient, and emergency visits, etc. Table 3 provides an overview of the dataset used.

Table 3
Dataset Description

Data Set Characteristics	Multivariate
Number of Instances	100000
Area	Life
Attribute Characteristics	Integer
Number of Attributes	55
Date Donated	2014-05-03
Associated Tasks	Classification, Clustering
Missing Values?	Yes
Number of Web Hits	395505

In the proposed method, we modified the values for the dataset before we used it. The modification process aims to convert the string values to an integer to be more compatible with MATLAB. It is used for extracting features from the original dataset, then it is used for building the classifier model for testing the results.

3.1.2 Dataset Preprocessing

Dataset preprocessing involves implementing methods to simplify the dataset by removing irrelevant values and unnecessary attributes, ultimately reducing its complexity (Hamid et al. 2016). This study focuses on the dataset preprocessing stage, which involves choosing all relevant and suitable attribute sets from the raw attributes or raw dataset (specifically the US hospitals dataset). The dataset's representative set works by retaining important attributes while removing any irrelevant ones. In this research, the prediction model utilized the Information Gain Ratio (IGR) feature selection method to choose relevant data processing features from the original dataset, making data visualization and understanding easier. IGR is utilized to divide the distribution of attribute patterns into categories, with the attribute's gain ratio decreasing as the split information value rises.

3.1.3 Features Selection

In this phase, the proposed method used each of the MA and AOA algorithms as a features selection method. This phase aims to build a refined dataset that contains the most effective attributes from the raw dataset. For implementing the hybrid approach between the AOA and MA algorithms, the proposed method used the MA algorithm as a fitness function for the AOA algorithm. The MA algorithm is an evolutionary algorithm that employs local search instead of global search algorithms. Metaheuristic algorithms (MAs) use local search methods to improve individuals in an evolutionary way. By merging global and local searches, we achieve a global optimization procedure. Therefore, an effective algorithm is necessary for optimal routing (Ramadan et al., 2018). The refined dataset resulting from this stage will serve as the input data for the MLPNN phase.

3.1.4 Multilayer Perceptron Neural Network

A Multi-layer Perceptron (MLP) is the term used for a fully connected neural network according to Janke et al. in 2019. The MLP algorithm is an addition to and type of feed-forward neural network (Gumbarević et al., 2020). It consists of three types of layers: input layer, output layer, and hidden layer. The input layer is busy receiving the input signal for processing. The output layer carries out tasks like prediction and classification as stated by Al-Saif et al. (2021). The real computational power of the MLP lies in having a random number of hidden layers situated between the input and output layers. Just like in a feed-forward network within a multi-layer perceptron, the information moves forward from the input to the output layer. The back propagation learning algorithm is used to train the neurons in the MLP. MLPs are created to estimate any continuous function and have the ability to address issues that are not linearly separable (Saha et al., 2021). MLP is primarily used for pattern classification, recognition, prediction, and approximation.

3.1.5 Evaluation Measurements

For data analysis and interpretation, and the well-known error indexing measurements are used for evaluation for the proposed prediction model.

3.1.5.1 Error Indexing Measurements

Calculation of error percentage among various methods is the main focus of error indexing measurements, utilizing Mean absolute error (MAE), Mean relative error (MRE), Mean square error (MSE), and Mean square percentage error (MSPE) as outlined by Xiong et al. (2019). Eq. (1) computes the Mean absolute error (MAE), which primarily shows the average absolute difference between the actual value (classification in the dataset) and the forecasted value (results from system methods).

$$MAE = \frac{1}{N} \sum |Yreal - Y\text{scenario}| \tag{1}$$

Eq. (2) is used to determine the Mean relative error (MRE). MRE is employed to evaluate how much the predicted value differs from the actual value. The smaller the value is, the less the difference between the predicted value (system methods results) and the actual value (dataset classification), resulting in a better classification effect.

$$MRE = \frac{1}{N} \{ \sum \frac{|Yreal - Ypredict|}{Yreal} \} \times 100\% \tag{2}$$

Eq. (3) is applied for determining the Mean Square Error (MSE), which represents the distribution of errors. The error distribution will be more concentrated and the classification effect will be better when the value of the error is smaller.

$$MSE = \frac{1}{N} \sum (Yreal - Ypredict)^2 \tag{3}$$

In conclusion, Eq. (4) is employed to compute the Mean Square Percentage Error (MSPE), which indicates the error distribution and the difference between the predicted and actual values to some degree.

$$MSPE = \frac{1}{N} \{ \sum (\frac{Yreal - Ypredict}{Yreal})^2 \} \times 100\% \tag{4}$$

3.1.5.2 Accuracy Analysis

The purpose of the accuracy analysis phase is to utilize ML methods to determine the accuracy and error rates of the chosen features in the improved dataset. This is to assess the reliability of the prediction model by analyzing the outcomes from the confusion matrix. Essentially, the goal of the prediction model is to create a confusion matrix through the testing of results with machine learning methods. Each of these methods will produce their own confusion matrix showing correct classifications, incorrect classifications, accuracy rates, and error rates. Prior to beginning, it is necessary to change the format of the data results file from Excel to CSV in order to work with the prediction model developed on the MATLAB platform. The confusion matrix is a common tool for evaluating a classifier's performance on test data with known true values. The confusion matrix is easy to comprehend, but the associated vocabulary may be perplexing. The accuracy measures are taken from (Markoulidakis et al., 2021), and can be described in the following way. The ensuing details the top measurement commonly used for creating the confusion matrix:

- The precision value (Pr) or known as the Positive Predictive value, is the ratio of correctly classified wrong flows (TP), in front of all the classified flows (TP+CF).

$$\text{Pr} = \frac{TP}{TP + CF} \tag{5}$$

- Recall (Rc), is the ratio of correctly classified wrong flows (TP), in front of all generated flows for all experiments (TP+FN).

$$\text{Rc} = \frac{TP}{TP + Fn} \tag{6}$$

- F-Measure (F1), is a hybrid combination of the Pr and Rc into one measure.

$$\mathrm{F1} = \frac{2}{\frac{1}{Pr} + \frac{1}{Rc}} \tag{7}$$

- The accuracy or percentage of correct classification (PCC), can be calculated using the formula below:

$$\mathrm{PCC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

- Sensitivity is the (Number of true positive assessments)/ (Number of all positive assessments) (Zhu et al., 2010):

$$\mathrm{Sensitivity} = \frac{TP}{TP + Fn} \tag{9}$$

- The Specificity is the (Number of true negative assessments)/ (Number of all negative assessments) (Zhu et al., 2010):

$$\mathrm{Specificity} = \frac{TN}{TN + FP} \tag{10}$$

TP represents samples correctly classified as true, TN represents samples correctly classified as false, FP represents falsely classified samples as false, and FN represents falsely classified samples as true. Through experimental analysis, the prediction model under consideration utilizes a validation approach involving both training and testing sets. The training set comprises 70% of the data, while the testing set consists of 30% of the total data.

4. Results and analysis

This part showcases the outcomes of the suggested approach, with the outcomes categorized into three sections. In the initial part, the outcomes of the experiment with different MLP, AOA, and MA algorithms were analyzed based on the MATLAB results. The outcomes of the suggested prediction model are included in the second section. Ultimately, the concluding part presents the findings of the proposed method compared to another research.

4.1 Testing and trials

The experiment outcomes for the suggested approach are presented here, employing the Multi-Layer Perceptron (MLP) algorithm with multiple neurons and epochs at each layer based on the mean square error. In order to conduct the tests, we utilized five experiments (Exp_1, Exp_2, Exp_3, Exp_4, and Exp_5). 400 instances were trained in each experiment. Table 4 displays various experiment characteristics that impact the creation of the refined dataset and training of the network.

Table 4
Experiments Attributes

	Exp_1	**Exp_2**	**Exp_3**	**Exp_4**	**Exp_5**
Search Number	20	40	40	40	100
Iteration	5	10	40	40	100
Max Iteration	20	40	40	40	100
Epochs	10	20	40	80	80
Hidden Neurons	10	20	40	80	80

The performance of the AOA and MA algorithms is directly influenced by the search numbers, iterations, and maximum iterations. The goal is to determine the best values for selecting top features from the initial dataset in order to construct the enhanced dataset. Alternatively, both the epochs and the number of hidden layers have an impact on the performance of the MLP algorithm. The purpose is to discover the best values for predicting diabetes disease.

4.1 Findings from the Experiment

In this part, we will outline the outcomes for each individual experiment, including the training error percentage, execution time, and accuracy rates.

4.1.1 Trial (1)

The initial characteristics were followed in the first trial as demonstrated in Table (4.1). The findings of the initial experiment are displayed in Table 5.

Table 5
Experiment_1 Initial Results

MSE	0.2476 %
MSPE	0.0012 %
Execution Time for Data Processing	5.0573
Execution Time for Hybrid algorithms (AOA, MA)	46.4817
Execution Time for MLP algorithm	9.6788

According to Table 5, we can notice the different results for the first experiment, where the MSE reached (0.2476 %), while the MSPE reached (0.0012 %). For the executions time, the data processing needs the lowest time reached (5.0573) millisecond, with (9.6788) millisecond for the MLP algorithm, but the hybrid between each of the AOA and MA needs the most of the time that reached (46.4817) millisecond. For the prediction results, Table 6 shows the true and false predictions for the (400) instances based on the first experiment:

Table 6
Experiment_1 Prediction Results

TRUE Prediction	364
FALSE Prediction	36
Accuracy	**91 %**

According to Table 6, the true prediction for (400) instances in the first experiment reached (364) instances, with (36) instances as a false prediction. The overall accuracy rate reached (91 %) for the first experiment.

4.1.1 Experiment (2)

The second experiment worked according to the initial attributes as shown in Table 5. Table 7 shows the results of the second experiment:

Table 7
Experiment_2 Initial Results

MSE	0.1726
MSPE	0.001
Execution Time for Data Processing	5.7879
Execution Time for Hybrid algorithms (AOA, MA)	92.2675
Execution Time for MLP algorithm	9.8868

According to Table 7, we can notice the different results for the second experiment, where the MSE reached (0.1726 %), while the MSPE reached (0.001 %). For the executions time, the data processing needs the lowest time reached (5.7879) millisecond, with (9.8868) millisecond for the MLP algorithm, but the hybrid between each of the AOA and MA needs most of the time that reached (92.2675) millisecond. For the prediction results, Table 8 shows the true and false predictions for the (400) instances based on the second experiment:

Table 8
Experiment_2 Prediction Results

TRUE Prediction	371
FALSE Prediction	28
Accuracy	**93 %**

According to Table 8, the true prediction for (400) instances in the second experiment reached (371) instances, with (28) instances as a false prediction. The overall accuracy rate reached (93 %) for the second experiment.

4.1.2 Experiment (3)

The third experiment worked according to the initial attributes as shown in Table 5. Table 9 shows the results of the third experiment:

Table 9
Experiment_3 Initial Results

MSE	0.1265
MSPE	8.9136
Execution Time for Data Processing	5.4483
Execution Time for Hybrid algorithms (AOA, MA)	104.9378
Execution Time for MLP algorithm	11.5949

According to Table 9, we can notice the different results for the third experiment, where the MSE reached (0.1265 %), while the MSPE reached (8.9136 %). For the executions time, the data processing needs the lowest time reached (5.4483) millisecond, with (11.5949) millisecond for the MLP algorithm, but the hybrid between each of the AOA and MA needs the most of the time that reached (104.9378) millisecond. For the prediction results, Table 10 shows the true and false predictions for the (400) instances based on the third experiment:

Table 10
Experiment_3 Prediction Results

TRUE Prediction	364
FALSE Prediction	36
Accuracy	**91 %**

According to Table 9, the true prediction for (400) instances in the third experiment reached (364) instances, with (36) instances as a false prediction. The overall accuracy rate reached (91 %) for the third experiment.

4.1.3 Experiment (4)

The fourth experiment worked according to the initial attributes as shown in Table 5. Table 11 shows the results of the fourth experiment:

Table 11
Experiment_4 Initial Results

MSE	0.0994
MSPE	7.9015
Execution Time for Data Processing	6.2579
Execution Time for Hybrid algorithms (AOA, MA)	105.1819
Execution Time for MLP algorithm	11.8501

According to Table 11, we can notice the different results for the fourth experiment, where the MSE reached (0.0994 %), while the MSPE reached (7.9015 %). For the executions time, the data processing needs the lowest time reached (6.2579) millisecond, with (11.8501) millisecond for the MLP algorithm, but the hybrid between each of the AOA and MA needs the most of the time that reached (105.1819) millisecond. For the prediction results, Table 12 shows the true and false predictions for the (400) instances based on the fourth experiment:

Table 12
Experiment_4 Prediction Results

TRUE Prediction	366
FALSE Prediction	34
Accuracy	**92 %**

According to Table 12, the true prediction for (400) instances in the fourth experiment reached (366) instances, with (34) instances as a false prediction. The overall accuracy rate reached (92 %) for the fourth experiment.

4.1.4 Experiment (5)

The fifth experiment worked according to the initial attributes as shown in Table 5. Table 13 shows the results of the fifth experiment:

Table 13
Experiment_5 Initial Results

MSE	0.1113
MSPE	8.3609
Execution Time for Data Processing	5.738
Execution Time for Hybrid algorithms (AOA, MA)	245.2663
Execution Time for MLP algorithm	10.7953

According to Table 13, we can notice the different results for the fifth experiment, where the MSE reached (0.1113 %), while the MSPE reached (8.3609 %). For the executions time, the data processing needs the lowest time reached (5.738) millisecond, with (10.7953) millisecond for the MLP algorithm, but the hybrid between each of the AOA and MA needs the most of the time that reached (245.2663) millisecond. For the prediction results, Table 14 shows the true and false predictions for the (400) instances based on the fifth experiment:

Table 14
Experiment_5 Prediction Results

TRUE Prediction	378
FALSE Prediction	22
Accuracy	**95 %**

According to Table 14, the true prediction for (400) instances in the fifth experiment reached (378) instances, with (22) instances as a false prediction. The overall accuracy rate reached (95 %) for the fifth experiment.

4.2 Comparing Experiment Results

In this section, we will present the results for all experiments and compare them, which contain the error percentage for training, the run time executions, and the accuracy rates. Fig. 2 shows the comparing results for all experiments according to the MSE results:

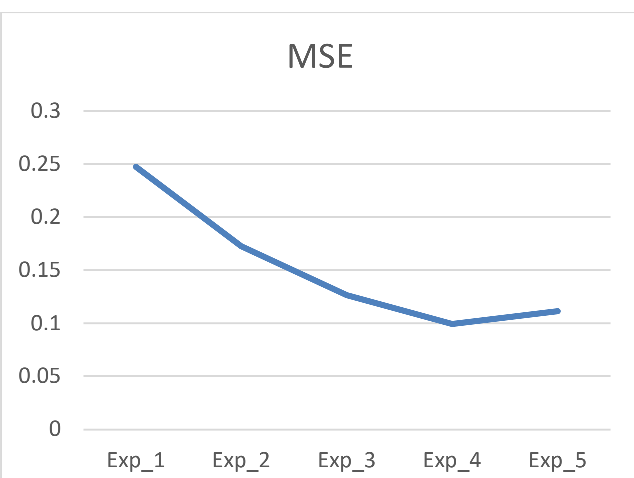

Fig. 2. Comparing MSE results

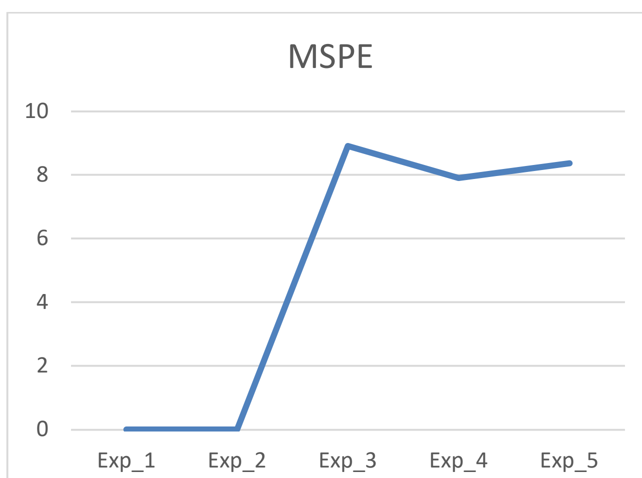

Fig. 3. Comparing MSPE results

According to Fig. 2, we can notice that the first experiment has a high MSE compared to other experiments, where it reached (0.2476 %). While the fourth experiment has a low MSE compared to other experiments, where it reached (0.0994 %). Fig. 3 shows the comparing results for all experiments according to the MSPE results:

We can notice that the first and second experiments have a high MSPE compared to other experiments because of the initial attributes (the number of iterations), where it reached (0.0012 %) for the first experiment, with (0.001 %) for the second experiment. Fig. 4 shows the comparing results for the run time for each of the data processing, hybrid algorithms execution (AOA, MA), and the MLP algorithm execution:

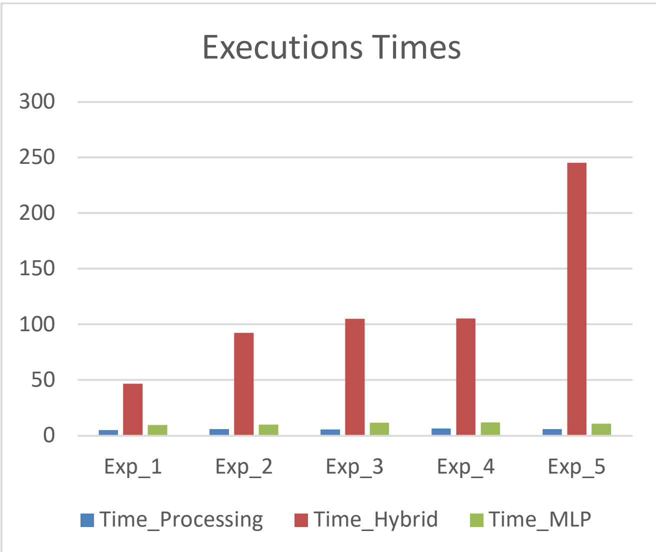

Fig. 4. Comparing Executions Times

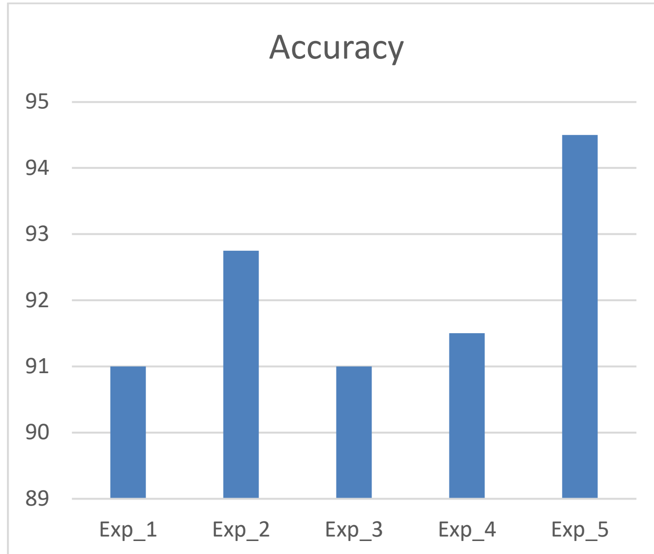

Fig. 5. Comparing Final Accuracy

According to Fig. 4, we can notice that the fifth experiment has a high run time for the execution of the hybrid algorithm (AOA, MA), because of each of the search numbers, iteration numbers, and the max iteration numbers. Finally, Fig. 5 shows the comparing results for the final accuracy rates for all experiments. According to Fig. 5, we can notice that the fifth experiment has a high accuracy compared to other experiments, where it reached (95 %).

4.3 Prediction Model's Results

The suggested forecasting model produces a confusion matrix detailing correct classifications, misclassifications, accuracy, and error. The suggested model includes four classifiers and metrics that aim to evaluate the effectiveness and precision of the proposed technique. Each classifier produces a confusion matrix file using the scorer node, with each matrix based on various measures outlined in Table 15.

Table 15
Confusion Matrix of four classifiers

Classifier	Correct Classified	Wrong Classified	Accuracy	Error
SVM	541	59	90.167 %	9.833 %
Decision Tree	588	42	93.571 %	6.429 %
Logistic Regression	560	40	93.333 %	6.667 %
Naive Bayes	552	48	92.857 %	7.143 %

Based on Table 15, the outcomes for the four classifiers closely resemble the ultimate results for the suggested approach as depicted in Fig. 5, with SVM achieving an accuracy of 90.167% and an error rate of 9.833%. The accuracy of the Naive Bayes classifier was 92.857% with an error rate of 7.143%, while the Logistic Regression classifiers achieved an accuracy of 93.333% with an error rate of 6.667%. Eventually, the Decision Tree classifier achieved an accuracy rate of 93.571% with a minimum error rate of 6.429%.

4.4 Comparing Results

In this section, we compared the final results of the proposed approach with other studies that used the MLP algorithm for diabetes prediction. Table 16 shows the collecting models and classifiers for all previous studies with the proposed method results for our study:

Table 16
Comparing Results

Study	Measurements	Accuracy
(Mohapatra et al., 2019)	MLP	77.50 %
(Verma et al., 2020)	MLP	82.00 %
(Karthiga et al., 2020)	MLP	86.00 %
(Bani-Salameh et al., 2021)	KNN	68.40 %
(Bani-Salameh et al., 2021)	SVM	66.60 %
(Bani-Salameh et al., 2021)	MLP	71.90 %
(Theerthagiri, 2021)	MLP	81.00 %
(Bukhari et al., 2021)	ABP-SCGNN	93.00 %
MLP-AMA	**SVM**	**90.16 %**
MLP-AMA	**Decision Tree**	**93.57 %**
MLP-AMA	**Logistic Regression**	**93.33 %**
MLP-AMA	**Naive Bayes**	**92.85 %**

According to the Table 16, we can notice the preference for the proposed method accuracy compared to other studies, each of the Decision Tree, and Logistic Regression classifiers reached a higher accuracy compared to all classifiers in other studies, where the accuracy reached for both of classifiers respectively (93.57 %), and (93.33 %).

5. Conclusions and Future Works

Diabetes mellitus is an extremely serious illness that is incurable. If this illness impacts the person, it will be lifelong. Researchers and developers face two main challenges in building models to predict type 2 diabetes due to the high number of studies conducted in this field. Initially, there was a notable diversity in the machine learning (ML) techniques utilized in earlier research, posing challenges in identifying the best approach. Secondly, there is a lack of transparency in the algorithms used to train models, diminishing their interpretability. This study aimed to improve diabetes prediction by using an MLPNN with optimization algorithms for feature selection through a hybrid approach combining MA and AOA. The initial stage of the suggested approach is to gather significant features for predicting diabetes in order to gather the raw dataset. In the following stage, the method suggested utilized preprocessing on the raw dataset to choose a distinct collection of attributes from the original set of attributes. During this stage, the suggested technique applied IGR to select features, which is crucial for generating a set that accurately represents the data by including only significant attributes. In the following stage, the method proposed employed a hybrid approach to optimize feature selection by alternating between the MA and AOA algorithms. This stage produces an improved dataset that will be used as the input for the Neural Network in the following step. In the final stage, the suggested approach employed MLPNN to train the network with hidden layer neurons. The proposed method's experiments involved employing a group of multiple perceptrons/neurons and epochs at each layer within the MLP algorithm based on the mean square error. For the experiments, we utilized five sets, with each set containing (400) instances for training. The AOA and MA algorithm's performance is directly impacted by the search numbers, iterations, and max iterations. The goal is to determine the best values to extract top features from the original dataset in order to create the improved dataset. However, both the epochs and the number of hidden layers have an impact on the MLP algorithm. The objective is to determine the best values for predicting diabetes. The results of the experiment indicated that the initial trial had a significantly higher MSE of 0.2476% in comparison to the other experiments. Although the fourth experiment achieved an MSE of 0.0994%, it is lower than the MSE of other experiments. Additionally, the first and second experiments demonstrated a high MSPE compared to other experiments due to their initial attributes (number of iterations), with the first experiment reaching 0.0012% and the second experiment reaching 0.001%. The fifth experiment shows a long run time for executing the hybrid algorithm (AOA, MA) due to the search numbers, iteration numbers, and max iteration numbers. In terms of accuracy rate results, the fifth experiment achieved a high accuracy of 95% compared to the other experiments.

We utilized four classifiers (SVM, Decision Tree, Logistic Regression, and Naive Bayes) from the suggested prediction model for the evaluation metrics. These classifiers are evaluating the effectiveness and precision of the suggested approach. The final results for the proposed method in Figure (4.4) show that the accuracy of SVM reached 90.167%, with an error rate of 9.833%, which is similar to the results of the four classifiers. Next, the Naive Bayes classifier achieved an accuracy rate of 92.857 % and an error rate of 7.143 %, whereas the Logistic Regression classifiers had an accuracy rate of 93.333 % and an error rate of 6.667 %. In the end, the Decision Tree classifier achieved a high accuracy of 93.571% with a minimal error rate of 6.429%.

At last, we evaluated the outcomes of our proposed method against prior research employing the MLP algorithm for predicting diabetes. Based on the comparison results, it is evident that the proposed method is favored for its accuracy over other studies. Both Decision Tree and Logistic Regression classifiers achieved higher accuracy rates (93.57% and 93.33% respectively) when compared to all other classifiers in previous studies. In terms of future research, it is advised to explore alternative methods for predicting diabetes, such as utilizing decision support systems with added dataset attributes. Furthermore, there is an expectation for enhanced system scalability through the latest advancements in the field, allowing for the exploration of new scenarios using various algorithms commonly used in the broader domain of artificial intelligence to assess alternative selection options.

References

Abu-Alaish, A., Jaradat, G., Al-Shugran, M., & Alodat, I. (2021). Automating SWOT Analysis Using Machine Learning Methods. *International Journal of Advances in Soft Computing & Its Applications, 13*(2).

Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H. (2021). The Arithmetic Optimization Algorithm. *Computer Methods in Applied Mechanics and Engineering, 376*, 113609. https://doi.org/10.1016/j.cma.2020.113609.

Alanazi, A. S., & Mezher, M. A. (2020). Using Machine Learning Algorithms for Prediction of Diabetes Mellitus. *2020 International Conference on Computing and Information Technology, ICCIT 2020, 02*, 55–57. https://doi.org/10.1109/ICCIT-144147971.2020.9213708.

Al-Hawamdeh, M. M., & Alshaer, S. A. (2022). Artificial intelligence applications as a modern trend to achieve organizational innovation in Jordanian commercial banks. *The Journal of Asian Finance, Economics and Business, 9*(3), 257-263.

Al-Saif, A. M., Abdel-Sattar, M., Aboukarima, A. M., & Eshra, D. H. (2021). Application of a multilayer perceptron artificial neural network for identification of peach cultivars based on physical characteristics. *PeerJ, 9,* 1–22. https://doi.org/10.7717/peerj.11529.

Alzubi, J. A., Almomani, O., Alzubi, O. A., & Al-shugran, M. (2016). Intelligent and dynamic neighbourhood entry lifetime for position-based routing protocol using fuzzy logic controller. *International Journal of Computer Science and Information Security, 14*(1), 118.

Bani-Salameh, H., Alkhatib, S. M., Abdalla, M., Al-Hami, M., Banat, R., Zyod, H., & Alkhatib, A. J. (2021). Prediction of diabetes and hypertension using multi-layer perceptron neural networks. *International Journal of Modeling, Simulation, and Scientific Computing, 12*(2). https://doi.org/10.1142/S1793962321500124.

Bukhari, M. M., Alkhamees, B. F., Hussain, S., Gumaei, A., Assiri, A., & Ullah, S. S. (2021). An Improved Artificial Neural Network Model for Effective Diabetes Prediction. *Complexity, 2021*. https://doi.org/10.1155/2021/5525271.

Darabi, P. K., & Tarokh, M. J. (2018). Type 2 Diabetes Prediction Using Machine Learning Algorithms. *Jorjani Biomedicine Journal*, *8*(3), 4-18. https://doi.org/10.29252/jorjanibiomedj.8.3.4.

Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. *Diabetology and Metabolic Syndrome, 13*(1). https://doi.org/10.1186/s13098-021-00767-9.

Gowthami, S., Reddy, R. V. S., & Ahmed, M. R. (2024). Exploring the effectiveness of machine learning algorithms for early detection of Type-2 Diabetes Mellitus. *Measurement: Sensors, 31,* 100983.

Gumbarević, S., Milovanović, B., Gaši, M., & Bagarić, M. (2020). Application of Multilayer Perceptron Method on Heat Flow Meter Results for Reducing the Measurement Time. 29. https://doi.org/10.3390/ecsa-7-08272.

Gupta, S. C., & Goel, N. (2023). Predictive modeling and analytics for diabetes using hyperparameter tuned machine learning techniques. *Procedia Computer Science, 218*, 1257-1269.

Hamid, H., Sugumaran, M., and Journaux, L. (2016). Machine Learning Techniques for Intrusion Detection: A Comparative Analysis*. Conference: The International Conference*. DOI: 10.1145/2980258.2980378.

Janke, J., Castelli, M., & Popovič, A. (2019). Analysis of the proficiency of fully connected neural networks in the process of classifying digital images: Benchmark of different classification algorithms on high-level image features from convolutional layers. *Expert Systems with Applications, 135*, 12–38. https://doi.org/10.1016/j.eswa.2019.05.058.

Jaradat, G., Abu Al Aish, A., Ayob, M., Alsmadi, M., & Almarashdeh, I. (2018). An Elite Pool-Based Big Bang-Big Crunch Metaheuristic for Data Clustering. *Journal of Computer Science, 14*. https://doi.org/10.3844/jcssp.2018.1611.1626.

Jian, Y., Pasquier, M., Sagahyroon, A., & Aloul, F. (2021). A machine learning approach to predicting diabetes complications. *Healthcare (Switzerland), 9*(12). https://doi.org/10.3390/healthcare9121712.

Karthiga, S., Abirami, A. M., Indira, K., & Angeline, C. N. (2020). Efficient machine learning technique to predict diabetics using multilayer perceptron. *PalArch's Journal of Archaeology of Egypt/Egyptology*, *17*(7), 4816-4824.

Khaleel, F. A., & Al-Bakry, A. M. (2023). Diagnosis of diabetes using machine learning algorithms. Materials Today: Proceedings, 80, 3200-3203.

Khanam, J. J., & Foo, S. Y. (2021). A comparison of machine learning algorithms for diabetes prediction. *ICT Express, 7*(4), 432–439. https://doi.org/10.1016/j.icte.2021.02.004.

Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *Technologies, 9*(4), 81. https://doi.org/10.3390/technologies9040081.

Mohapatra, S. K., Swain, J. K., & Mohanty, M. N. (2019). Detection of Diabetes Using Multilayer Perceptron. *In Advances in Intelligent Systems and Computing (Vol. 846, Issue September). Springer Singapore.* https://doi.org/10.1007/978-981-13-2182-5_11.

Mujumdar, A., & Vaidehi, V. (2019). Diabetes Prediction using Machine Learning Algorithms. *Procedia Computer Science, 165*, 292–299. https://doi.org/10.1016/j.procs.2020.01.047.

Rajeswari, M., & Prabhu, P. (2019). A Review of Diabetic Prediction Using Machine Learning Techniques. *International Journal of Engineering and Techniques, 5*(July 2019).

Ramadan, R. M., Gasser, S. M., El-Mahallawy, M. S., Hammad, K., & El Bakly, A. M. (2018). A memetic optimization algorithm for multi-constrained multicast routing in ad hoc networks. *PLoS ONE, 13*(3), 1–17. https://doi.org/10.1371/journal.pone.0193142.

Saha, S., Paul, G. C., Pradhan, B., Abdul Maulud, K. N., & Alamri, A. M. (2021). Integrating multilayer perceptron neural nets with hybrid ensemble classifiers for deforestation probability assessment in Eastern India. *Geomatics, Natural Hazards and Risk, 12*(1), 29–62. https://doi.org/10.1080/19475705.2020.1860139.

Saxena, R., Sharma, S. K., Gupta, M., & Sampada, G. C. (2022). [Retracted] A Comprehensive Review of Various Diabetic Prediction Models: A Literature Survey. *Journal of Healthcare Engineering*, *2022*(1), 8100697.

Sharma, T., & Shah, M. (2021). A comprehensive review of machine learning techniques on diabetes detection. *Visual Computing for Industry, Biomedicine, and Art, 4*(1). https://doi.org/10.1186/s42492-021-00097-7.

Theerthagiri, P. (2021). Diagnosis and Classi cation of the Diabetes Using Machine Learning Algorithms.

Turnbull, C. H. S. and D., Dolnicar, S., Chapple, A., Adorno, T., Horkheimer, M., Negus, K., Pickering, M., Leckenby, J., & Li, H. (2004). Scholar (11). In Why We Need the Journal of Interactive Advertising (Vol. 10, Issue 10, pp. 349–383).

Verma, G., & Verma, H. (2020). A multilayer perceptron neural network model for predicting diabetes. *International Journal of Grid and Distributed Computing*, *13*(1), 1018-1025. https://doi.org/10.13140/RG.2.2.23203.89126.

Wee, B. F., Sivakumar, S., Lim, K. H., Wong, W. K., & Juwono, F. H. (2024). Diabetes detection based on machine learning and deep learning approaches. *Multimedia Tools and Applications, 83*(8), 24153-24185.

Xiong, Z., Xue, Y., Gao, D., and Fang, C. (2019). Research on short-term traffic flow prediction method based on real-time traffic status, *IOP Conf. Series: Materials Science and Engineering,* doi:10.1088/1757-899X/569/5/052062.

Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations. Northeast SAS Users Group 2010: Health Care and Life Sciences, 1–9.