Design of Multi-Layer Perceptron for the Diagnosis of Diabetes Mellitus Using Keras in Deep Learning

Muni Kumar N[1] , Manjula R[2],

[1] Research Scholar, [2] Associate Professor,
SCOPE, VIT University, Vellore, Tamil Nadu,
{muni.kumarn2014, rmanjula}@vit.ac.in

Abstract. Diabetes Mellitus is one of the deadly diseases growing at a rapid rate in the developing countries. Diabetes has affected over 246 million people worldwide. According to the World Health Organization (WHO) report, by 2025 this number is expected to rise to over 380 million. If the diabetes mellitus left untreated, it will result in blindness, amputation, kidney failures, heart attack and stroke. So, the timely diagnosis of diabetes is a great boon to the people around the world. This study has taken the diabetes data set and applied the Deep Learning method for the diagnosis of diabetes using Keras library. It is observed that, with the increase in the number of layers and nodes per layer in the MLP would result in greater diagnosing accuracy.

Keywords: Diabetes Mellitus, Artificial Neural Networks, Deep Neural Networks, Multi Layer Perceptron, Keras, Machine Learning, Deep Learning.

1 Introduction

Diabetes Mellitus [1] is one of the deadly diseases affecting over 400 million people worldwide and expected to rise to 640 million by 2040. It is a chronic disease that normally occurs when the pancreas is no longer able to produce insulin and the other scenario where the body cannot utilize the produced insulin. The first scenario in which pancreas does not produce insulin is called Type-I diabetes and the second scenario is in which the pancreas doesn't produce enough insulin is called Type-II diabetes. According to The National Diabetes Statistics Report 2017, Estimates of Diabetes and its burden in the United States, it is estimated that 30.3 million people of all ages of the U.S population had diabetes in 2015, of which 7.2 million people are not aware that they had diabetes. The percentage of adults with diabetes is increasing with age. Diabetes was the seventh leading cause of death in the United States in 2015. The total direct and indirect estimated cost of diagnosed diabetes in the United States in 2012 was $245 billion and average medical expenditure for people with diagnosed diabetes were about $13,700 per year. The average medical expenditures among people with diagnosed diabetes were about 2.3 times higher than expenditures for people without diabetes.

An estimated 33.9% of U.S adults aged 18 years or older had pre-diabetes in 2015, nearly half of adults aged 65 years or older had pre-diabetes. Also, the survey presented that, in 2011-2014, comparatively more men (36.6%) than women (29.3%)

had pre-diabetes. The number of people with diabetes has nearly quadrupled since 1980. Prevalence is increasing worldwide, particularly in low-and middle- income countries. Diabetes of all types can lead to complications in many parts of the body and increase the risk of premature deaths.

2 Related Works

Emrane et.al [2] developed an expert clinical decision support system to predict diabetes mellitus. In this system, they used Decision trees and K-Nearest Neighbour (KNN) algorithms and experimental results proved C4.5 classification model provides better accuracy for diabetes disease prediction. The experiment was conducted on the PIMA Indians Diabetes dataset. Anjali and Varun [3] attempted to develop a diabetes prediction method using Support Vector Machine (SVM) on different datasets with more than 1 lakh instances. The experiment was carried on WEKA environment and achieved 72% accuracy for diabetes prediction. Aparimita et.al [4] proposed the prediction and classification of diabetes mellitus using Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS).

Veena and Anjali [5] carried out a machine learning approach for the prediction and diagnosis of diabetes mellitus. The results proved that the Ada Boost algorithm with decision stump as base classifier performs better compared to that of Support Vector Machine, Naïve Bayes and Decision Trees. Ayush and Divya [6] designed a diabetes prediction system based on the personal daily life style activities. This paper used classification and regression trees (CART) algrorithm for the prediction of diabetes and obtained 75% accuracy in prediction. Durairaj & Kalaiselvi [7] proposed a back propagation algorithm for the prediction of diabetes on the PIMA Indian dataset. The classification accuracy of back propagation network was proved to be better than other methods.

Sapna and Pravin Kumar [8] proposed the diagnosis of diabetes disease from clinical big data using Neural network. They compared Quasi-Newton back propagation and Resilient back propagation methods and proved Resilient back propagation performed better. Zhilbert et.al [9] developed an Intelligent system for diabetes prediction. They focused on the joint implementation of the Support Vector Machine (SVM) and Naïve-Bayes statistical modeling and proved the joint implementation performs better compared to individual implementations.

Olaniyi and Adnan [10] proposed a back propagation network to predict the diabetes mellitus given the various input attributes to the network. Sigmoid transfer function was used in the hidden layers and the output layers. They proved Artificial Neural network trained with back propagation gave 82% accuracy. Linli [11] proposed a weight-adjusted voting approach for the diagnosis of diabetes. In this approach, the author combined three classifiers, support vector machines, Artificial Neural Network and Naïve Bayes to diagnose diabetes. Here, the PIMA Indian diabetes dataset was used and achieved an accuracy of 80.0%.

Jetri and Suharjito [12] implemented a method of neural network called Extreme Learning Machine (ELM) and proved the accuracy of ELM is better than the accuracy of Back Propagation in the diagnosis of diabetes mellitus.

Deepthi Jain & Divakar Singh [13] presented a feed forward neural network approach for the estimation of diabetes risk factors based on the body characteristics. They used neural network tool box of MATLAB on the dataset with 50,788 records provided by the SAS Data mining Shootout competition.

Muhammad et.al [14] presented a study on the prediction of diabetes using different supervised learning algorithms on the diabetes patients between 25 to 78 years old and their prediction accuracies were evaluated.

3 METHODOLOGY

3.1 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) [15] is one of the soft computing techniques that simulates the behavior of human brain and has become popular in its applications in the field of data mining, fault detection, image processing, pattern recognition, weather forecasting, job scheduling, medical diagnosis, etc. These networks are capable of learning through examples (train data), remember their past experiences and perform the parallel processing. The learning activities in these types of networks are possible because of the neurons that could receive and process the information as similar to that of human brain.

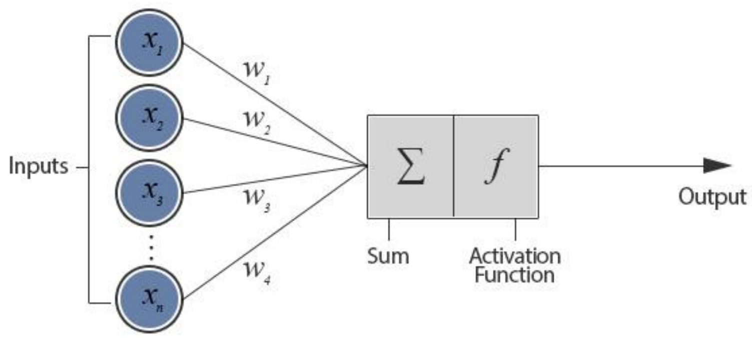

Fig. 1. A Simple Neural Network with 2-Layers. This shows a figure consisting of simple neural network. The input layer with n-neurons, activation functions and an output layer.

A very basic neural network with 2 layers ie., input layer and processing or output layer is depicted in Fig.1. The input layer receives input from the sensors and passes the input to the next layer for processing. The processing layer consists of summation and activation functions where the activation function is checked with the threshold value and generates the output signal of the Neural Network.

3.2 Multi Layer Perceptron (MLP)

A Multi-Layer Perceptron is a feed forward network with atleast three layers: input layer, hidden layer and output layer as depicted in Fig.2. In MLP, except the input layer, other layers use non-linear activation functions. MLP uses a supervised learning called Back Propagation for training. The MLP consists of minimum of 3 layers with non-linear activation functions, which makes this a Deep Neural Network (DNN). Every node in the MLP contains certain weights w_{ij}, which are connected to the nodes in the following layers. MLP learns in its learning phase with the change of weights after the processing of every node. Based on the output of the unit compared to the expected result, the error is calculated and then the weights gets updated which is carried out through back propagation method.

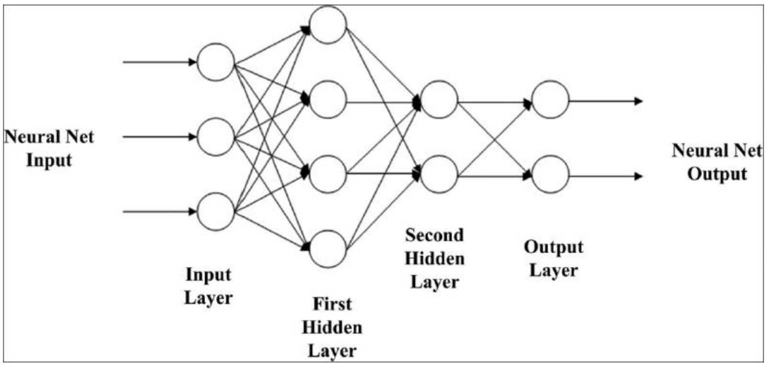

Fig. 2. A Multi-Layer Perceptron. This shows a figure with one input layer, two hidden layers and an output layer.

3.3 Proposed Model

In our proposed model, we installed the Theano and Tensorflow libraries as pre-requisite for using the Keras library. Keras allows for easier development of neural networks that can run both on CPU's and GPU's for faster computations and parallel processing seamlessly. The major task in the design of deep neural networks is organizing the layers. Using Keras it is easier to build any complex neural network with minimal number of steps.

Keras allows for easier development of neural networks that can run both on CPU's and GPU's for faster computations and parallel processing seamlessly. Using Keras it is easier to build any complex neural network with minimal number of steps as model is the core data structure in Keras which is used for organizing layers.

The proposed model shown in Fig.3. lists all the steps for the design and construction of a Multi-Layer Perceptron (MLP) network. Initial steps to begin the model is the installation of required packages and carrying out the data preprocessing steps like encoding the categorical data, independent variables and splitting the data into training set and test set. After the completion of pre-processing, the ANN must be initialized and the required number of layers should be added to build the model.

While building layers, suitable activation function should be supplied along with number of nodes in each layer. Among the available activation functions, ReLu activation function is preferred for the hidden layers and Sigmoid for the output layer. Our proposed model also uses Relu and Sigmoid activation functions.

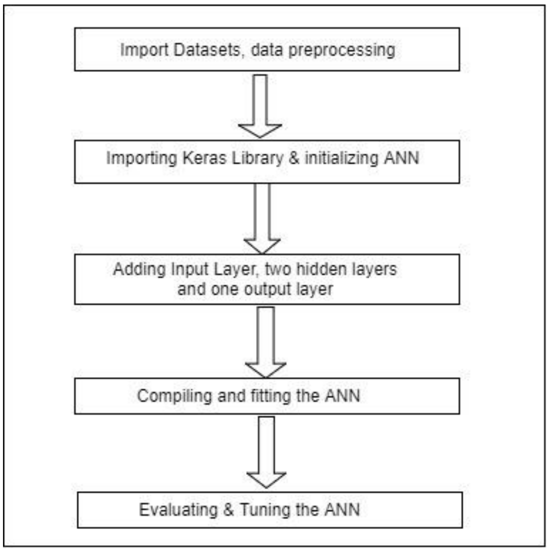

Fig. 3. Proposed Model for the design of MLP. This shows a figure with the proposed model starting with the dataset imports, import libraries, model, compilation and evaluation of ANN.

3.4 Pseudocode

The pseudocode for the proposed design of Multi-Layer Perceptron (MLP) network for the diagnosis of diabetes using PIMA Indian diabetes dataset is as follows.

1. Import the libraries for data pre-processing and the PIMA Indian diabetes dataset downloaded from the UCI respository.
2. Preprocess the input data like encoding categorical data, independent variables, splitting the dataset into training set & test set and perform feature scaling if required.
3. Import the Theano, Tensorflow and Keras library and required python packages for setting up the python environment.
4. Initialize the ANN
5. Add the input layer and the first hidden layer
6. Add the extra hidden layers, apply the activation functions
7. Add the output layer along with activation function
8. Compile the ANN
9. Fit the ANN to the training set
10. Predict the test results
11. Evaluate the ANN
12. Improve and Tune the ANN

4 EXPERIMENT

The proposed work is implemented using the Python environment on Windows 10 operating system with intel corei5 processor and 8 GB of RAM. The experiment requires some initial pre-requisites for setting up the environment which will be discussed as environmental setup in the following section. For the experiment, we have choosen PIMA Indian diabetes dataset from UCI data repository. The experiment was conducted with the python source code for the above given pseudocode.

4.1 Environmental Setup

Our proposed model requires the any operating system with python 2.7 installed. In python environment, the libraries can be installed with the following commands.

```
$ pip theano install          # installation of theano
$ pip install tensorflow      # installation of tensorflow
$ pip install keras           # installation of keras
```

Before using the above commands, the pip must be installed on python system. In our experimental setup, we have installed Anaconda environment for the execution of python programs in spider editor. Anaconda is an open source distribution of Python language for large scale data processing, predictive analytics and scientific computing which simplifies package management and deployment.

4.2 Data Set

The dataset used for this experiment was downloaded from the University of California, Irvine (UCI) Repository of Machine Learning databases. The name of the dataset is PIMA Indians dataset, which consists of 768 instances of 9 attributes belonging to the Indian people living in Arizona, USA.

5 Results and Discussion

The classification accuracies obtained by this research work on the PIMA Indian diabetes dataset for different MLP (Multi-Layer Perceptrons) were presented in the following table. The experimental results are tabulated for 3 Perceptrons as follows:

The notation 8-12-8-1 is used for the Multi Layer Perceptron indicating 8 nodes in the input layer, 12 and 8 nodes in the hidden layers and one node in the output layer. Similarly three more MLP networks are designed with the increase in nodes in the hidden layers as 8-32-32-1, 8-64-64-1 and 8-128-128-1. We have used the ReLU activation function in the input layer and the hidden layers and Sigmoid activation function in the output layer. The following table lists the results obtained by executing each network for 10 runs with 150 epochs in the each model.

Table 1. Comparison of the Performances of 4 different MLP Networks

	Performance of 4 different Multi Layer Perceptrons			
S.No.	8-12-8-1	8-32-32-1	8-64-64-1	8-128-128-1
1	78.91%	82.68%	83.33%	86.46%
2	79.56%	84.11%	82.81%	86.33%
3	77.99%	82.16%	83.72%	87.11%
4	78.78%	82.81%	84.77%	87.89%
5	79.43%	83.20%	84.51%	86.72%
6	79.69%	82.68%	85.29%	84.77%
7	78.26%	83.98%	84.64%	85.29%
8	79.17%	83.72%	85.68%	88.15%
9	78.26%	82.16%	84.38%	87.24%
10	80.47%	83.20%	83.46%	86.42%
Average	79.05%	83.02%	84.26%	86.67%

The results are collected in 10 runs for 4 different MLP networks for 150 epochs each network and compared in Table.1. A graph was drawn with the results obtained in Table.1 as shown in Fig.4. From the above results, it is evident that with the increase in number of nodes in the hidden layers.

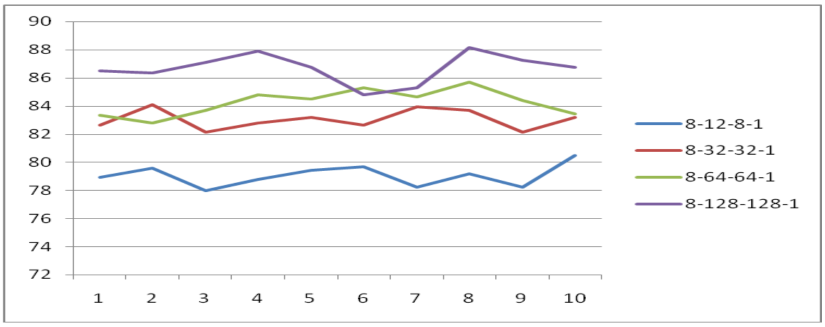

Fig. 4. Graph with classification accuracies of 4 different MLP Networks. This figure shows the comparisons of the MLP Networks drawn in separate color for each MLP.

6 Conclusion

This paper presented a new methodology for the design and development of a Multi-Layer Perceptron network using Keras. We have used PIMA Indian diabetes dataset for the diagnosis of diabetes mellitus. It was observed that, the designed model achieved greater accuracy by designing deep neural networks with the increase in the number of hidden layers and nodes per layer. This work can be further extended to obtain an accuracy of more than 90% with deep neural networks so that, the diabetes can be accurately diagnosed during the early stages.

Acknowledgments

I would like to take this opportunity to express my profound gratitude and deep regard to my guide, Prof. R. Manjula, SCOPE, VIT University, for her exemplary guidance, valuable feedback and constant encouragement in completing this paper. Her valuable suggestions were of immense help in getting this work done. Also, I would like to extend my sincere gratitude to my parents, my wife and my lovable daughter for their constant support and encouragement in completing this paper.

References

1. International Diabetes Federation. [Online]: 2017. Available from: https://www.idf.org/about-diabetes/what-is-diabetes (2017)
2. Emrana Kabir Hashi., Md. Shahid Uz Zaman., Md. Rokibul Hasan.: An Expert Clinical Decision Support System to Predict Disease Using Classification Techniques: Intl. Conf. on Electrical, Computer and Comm. Engineering (ECCE), pp. 16-18, Bangladesh (2017).
3. Anjli Negi., Varun Jaiswal.: A First Attempt to Develop a Diabetes Prediction Method Based on Different Global Datasets: Fourth IEEE Intl. Conf. on Parallel, Distributed and Grid Computing (PDGC), pp. 237-240, (2016).
4. Aparimita Swain., Sachi Nandan Mohanty., Ananta Chandra Das.: Comparative Risk Analysis on Prediction of Diabetes Mellitus Using Machine Learning Approach.: IEEE Intl. Conf. on Electrical, Elec. and Optimization Techniques (ICEEOT).,pp. 3312-3317., (2016)
5. Veena Vijayan V., Anjali C.: Prediction and Diagnosis of Diabetes Mellitus – A Machine Learning Appraoch.: IEEE Recent Advances in Intelligent Computational Systems (RAICS)., Trivandrum., (2015).
6. Ayush Anand., Divya Shakti.: Prediction of Diabetes Based on Personal Lifestyle Indicators.: IEEE 1st Intl. Conf. on Next Generation Computing Technologies (NGCT), Dehradun, pp. 673-676. (2015).
7. Durairaj., Kalaiselvi.: Prediction of Diabetes Using Back Propagation Algorithm.: Intl. Journal of Emerging Technology and Innovative Eng., Vol.1, Issue 8, pp.21-25, (2015)
8. Sapna S., Pravin Kumar M.: Diagnosis of Disease from Clinical Big Data Using Neural Network.: Indian Journal of Science and Technology, Vol 8(24)., pp. 1-7., (2015).
9. Zhilbert Tafa., Nerxhivane Pervetical., Bertran Karahoda.: An Intelligent System for Diabetes Prediction: 4th Mediterranean Conf. on Embedded Computing.pp. 378-382. (2015).
10. Ebenezer Obaloluwa Olaniyi., Khashman Adnam.: Onset Diabetes Diagnosis Using Artificial Neural Network.: Intl. Journal of Scientific & Eng. Research, pp. 754-759. (2014).
11. Lin Li.: Diagnosis of Diabetes using a Weight-Adjusted Voting Approach.: IEEE 14th Intl. Conf. on Bioinformatics and Bioengineering, pp. 320-324. (2014)
12. Jefri Junifer Pangaribuan., Suharjito.: Diagnosis of Diabetes Mellitus Using Extreme Learning Machine.: IEEE Intl. conf. on Information Technology Systems and Innovation (ICITSI), pp. 33-38. (2014).
13. Deepthi Jain., Divakar Singh.: A Neural Network based Approach for the Diabetes Risk Estimation.: International Journal of Computer Applications, pp. 1-4. (2013).
14. Mohammad Akmal Sapon., Khadijah Ismail., Suehazlyn Zainudin.: Prediction of Diabetes by using Artificial Neural Network.: Intl. Conf. on Circuits, Systems and Simulations (IPCSIT), Singapore, pp. 299-303. (2011).
15. B.K. Tripathy., J. Anuradha.: Soft Computing – Advances and Applications.: Cengage Learning (2015).