Comparative Study of Support Vector Machine to a 3 layer Neural Network Using Titanic Datasets

Pratiksha Yadav
Machine Learning Engineer Intern
pratikshanehayadav@gmail.com

AI TECHNOLOGY& SYSTEMS
www.ai-techsystems.com

Abstract*: This Support Vector Machines (SVM) as well as 3-Layer Neural Networks (NN) is very powerful tools for classification, which can be used for pattern recognition and regression analysis. They are being used in a large array of different areas such as medicine, education, agriculture, climate etc. Numerical databases are widely used in almost every field of research. In this thesis, comparative study of Support Vector Machine with Neural Network using numerical datasets has been done. Basic information about NN, including description of pattern recognition networks using feed-forward Learning Algorithm, has also been introduced as well as information about various kernels of SVM. Finally, the classification comparison of numerical datasets using accuracy has been done. Classification has been done in two versions: using NN and SVM. Both versions have been trained and their results of comparison are compared and presented here.*

Keywords: support vector machine, neural network.

I. INTRODUCTION

The aim of this paper is to compare support vector machine and 3-layer neural network, investigating whether neural networks are competitive with support vector machine according to accuracy. Classification is one of the oldest and the most important method of data mining. There are two types of classification, one is supervised and other one is unsupervised. Supervised classification has many techniques. Support vector machine and 3-layer neural network are most popular techniques of classification. So, problem is to find out better technique between SVM and 3 layer NN.

Now a day, there are vast amount of data being stored in databases across the globe. Data mining offers promising ways to uncover hidden patterns from such amount of data. These hidden patterns can probably be used to predict future behaviour. Classification in the data mining is one of the tasks to uncover these hidden patterns. The input for the classification is the training datasets, whose class labels are already known. Many approaches have been introduced to solve the classification problem. SVM is considered as one of the most robust and efficient methods among all well-known algorithms for classification. It is the powerful algorithm used for supervised learning, and is widely used in classification problems. On the other hand, 3-layer neural network contains more than one layer of artificial neurons or nodes. They differ widely in design. It is important to note that single layer neural networks were useful early in the evolution of AI, the vast majority of networks used today have a multi layered model. Multi-layer can be setup in numerous ways. Typically, they have at least one input layer, one hidden layer, one output layer at the end which is also called as 3- layer neural network.

II. BACKGROUND

In data mining, classification is one of the most popular tasks with wide variety of applications. Many algorithms have been presented to produce an accurate and efficient classifier. All these algorithms are worked on single table as an input but in real world applications, data is stored on multiple tables. There have been many techniques for classification as Neural Networks and Support Vector Machines. However, they can only be applied to data in single table. The conversion of the multiple tables to single table is very difficult and expensive. So, Multi-relational classification which uses weighted voting technique can be applied to combine classifiers to get class label based on the contribution of tables.

SVM classifier is a powerful classifier for the classification task of data mining. SVM is based on statistical learning theory and used to find the optimal separating hyperplane between two classes. Optimal hyperplane is the one giving maximum margin between training examples of different classes.

SVM converts the original data point dimensional space and the data point is viewed as a dimensional vector. The main aim of the SVM is to predict which

class a new data point will be in. There are many hyperplanes that may classify the data. The best hyperplane is the one which has largest separation margin between two classes. We choose this hyperplane because the distance from the nearest point on the each side is maximized. This hyperplane is known as maximum-margin hyperplane or optimal hyperplane. By this hyperplane, the linear classifier is defined known as maximum margin classifier. Fig 1 is showing the concept of SVM.

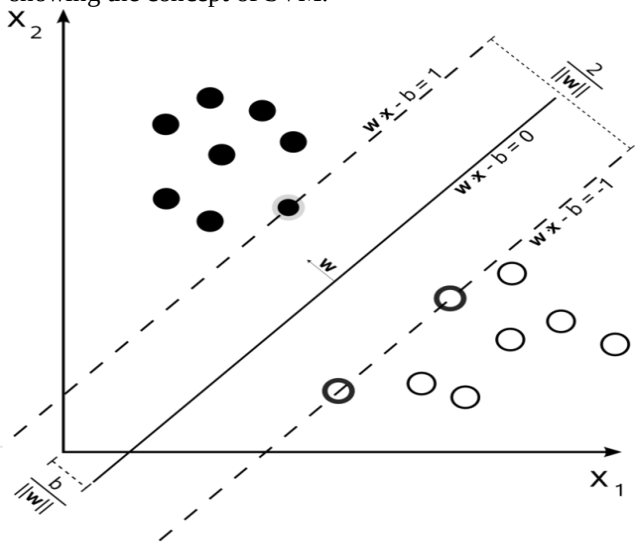

Fig.1-Support Vector Machine

3-layer neural network research is motivated by two desires: to obtain better understanding of the human brain, and to develop computers that can deal with the abstract and poorly defined problems. For example- conventional computers have trouble understanding speech and recognizing people's faces. In comparison, human do extremely well all these tasks. Many different NN structures have been tried, some based on initiating what a biologist sees under the microscope, some based on a more mathematical analysis of the problem. The most common structure is shown below in Fig 2. This NN is formed in three layer called input layer, hidden layer and output layer. Each layer consists of one or more nodes represented in this diagram by the small circles. The line between the nodes indicate the flow of information from one node to another.

The nodes of input layer are passive, meaning they do not modify the data. They receive a single value on their input and duplicate the value to their multiple outputs. In comparison, the nodes of the hidden and output layer are active. This means they modify the data. Each value from the input layer is duplicated and sent to all of the hidden nodes. This is called the fully interconnected structure.

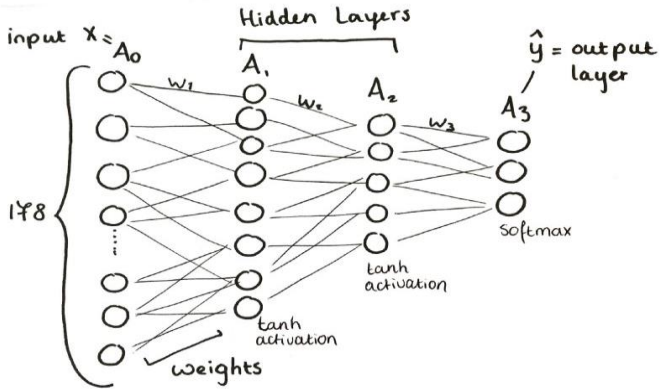

Fig. 2. 3-layer Neural Network

This the flow diagram of active node used in the hidden and output layers of NN. Each input is multiplied by weight and then summed up. This produce a single value that is passed through a 'S' shaped non-linear function called the sigmoid.

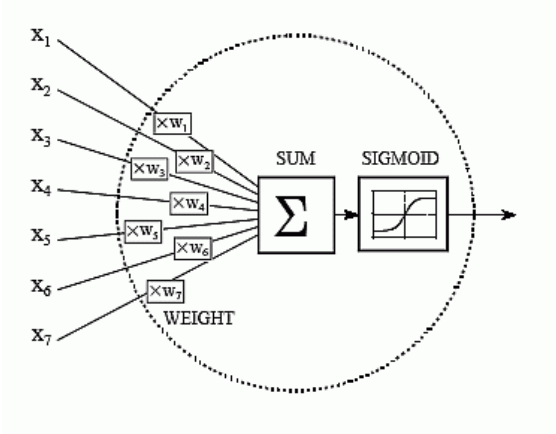

Fig.3 Flow diagram of neural network

III. COMPARING SVM AND 3 LAYER NEURAL NETWORK

A. Compared method

We focused on classification tasks for SVM and 3 layer NN.

B. Datasets

The titanic dataset is used which provide information on the fate of passengers on the Titanic.

C. Methodology

Step 1: I have taken the titanic datasets to compare the performances of SVM and 3-Layer Neural Network.
Step 2: For SVM, first train the dataset and then the extraction of features is done.
Step 3: Build the SVM Model.
Step 4: Results of SVM have been observed.
Step 5: For NN, initialization of data has been done.
Step 6: Then we propagate the data into forward and backward.
Step 7: After that train of data has been done.
Step 8: Results of NN have been observed.

Step 9: Comparison of both the results have been done by using the charts.

.

D. Experimental results

The Figures presents a comparison of the best results achieved by each method. It appears that much better results in classification were obtained using SVM than NN. It also seems that SVMs are more resistant to insufficient data amount, because even for small set of integer datasets results were satisfactory. That cannot be said about NN, which gives less accuracy as compared to SVM in integer datasets.

- **Accuracy**

Accuracy is the proximity of measurement results to the true value; precision, the repeatability, or reproducibility of the measurement. In the fields of science, engineering and statistics, the accuracy of a measurement system is the degree of closeness of measurements of a quantity to that quantity's true value. Accuracy of Support vector machine is sometimes better than neural network in classification of integer datasets because it depends of dataset. Comparison of Accuracy using SVM and NN is presented in the figure.

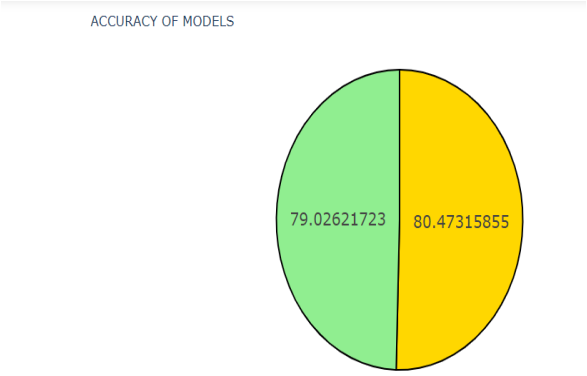

Fig. Comparison of Accuracy.

- **CPU Time**

CPU time (or process time) is the amount of time for which a central processing unit (CPU) was used for processing instructions of a computer program or operating system. The amount of time the CPU is actually executing instructions. During the execution of most programs, the CPU sits idle much of the time while the computer fetches data from the keyboard or disk, or sends data to an output device. This CPU time is the time used by the matlab to fetch data from datasets, process and gives results of classification.

- **Training Time**

Training time is the time in which the machine trains the network or prepares the model to perform

classification task. In figure 5, comparison of Support vector machine and artificial neural network is presented.

- **Testing Time**

Testing time is defined as the time taken by machine to predict the class label of the datasets. In this dissertation, we have used two class labels using SVM and NN.

SVM performed significantly better than Neural Network in the accuracy. Support Vector Machine classifies the data with the help of support vectors and Neural Network classifies the data with the help of hidden nodes. Number of support vector and Number of hidden nodes are equal.

IV. CONCLUSIONS

Experimental results indicate that support vector machine performed significantly better than Neural Network when talk about accuracy. It is observed that Support Vector Machine performed significantly better than Neural Network in the CPU time also. If the number of features or attributes is less, then Neural Network takes less time than SVM but if Number of features is large, then SVM takes less time than Neural Network. In the datasets above 5000 tuples, the CPU time taken by Support Vector Machine is always less than time taken by Neural Network.

In the Neural Network, number of hidden nodes is already decided but support vector machine generates number of support vectors by the model produced by training process itself. For the comparison of Neural Network and Support Vector Machine we used equal number of hidden nodes in Neural Network and number of Support Vectors in Support Vector Machine.

So we can conclude that SVM perform better then Neural Network in this titanic dataset.

REFERENCES

1. Titanic dataset from https://www.kaggle.com/c/titanic
2. Using Towards data science.comhttps://towardsdatascience.com/predicting-the-survival-of-titanic-passengers-30870ccc7e8
3. https://www.freecodecamp.org/news/building-a-3-layer-neural-network-from-scratch-99239c4af5d3/
4. https://www.researchgate.net/publication/285164623_An_Introduction_to_Convolutional_Neural_Networks
5. https://plot.ly/python/bar-charts/