

ANN AND SVM IS USED FOR PC DATABASE



#1 Harshali S. Lidhure.

¹harshali_3691@rediffmail.com

#1 Department of Electronics and Communication.

S.S.G.B.C.O.E.T, BHUSAWAL.
Dist-Jalgoan, Maharashtra, India.

ABSTRACT

Once a network has been planned for a particular application, that network is ready to be trained. Then, the training, or learning, begins. In this paper we offer methods that select a separation of data for SVM training. Several methods have been proposed, one of which is to decrease the training set size called training data selection. Recently, there has been great study on data selection for SVM training.

However, the training of neural networks can take many hours on a PC. In this paper we suggest a technique for training difficult neural networks for a general purpose computer. Our system allows the training to be done with minimal CPU utilization time. This allows the user to carry out other tasks while the training is in growth. We balance several avenues of neural network training on a general purpose computer. In this paper we propose methods that select a separation of data for SVM training. To speed up training thereby reduction the time for model selection, several methods have been proposed, one of which is to reduce the training set size called training data selection. Recently, there has been great research on data selection for SVM training.

Neural networks can be used for prediction with various levels of success. The advantage of them includes automatic learning of dependencies only from measured data without any need to add further information (such as type of dependency like with the regression). The neural network is qualified from the past data with the hope that it will determine unseen dependencies and that it will be able to use them for predicting into future

Keywords:-Neural network, artificial neural network, support vector machine.

ARTICLE INFO

Article History

Received: 5th November 2016

Received in revised form :

5th November 2016

Accepted: 7th November 2016

Published online :

8th November 2016

I. INTRODUCTION

This paper predicting the missing data, which was collected from traffic of computer using Artificial neural network (ANN) and Support vector machine (SVM). It starts with the detail report of a neat study of hourly traffic data of one year and its comparable numerical model by using two techniques. The Neural Network and SVM are skilled with the irregular data of 13 weeks over four years. The missing data is interpreted with quite reasonable accuracy with both the methods. The accuracy of the methods is compared and it is found that SVM method is much more accurate than the ANN technique. Both the ANN and SVM models could capture the variety of trends of the traffic data reasonably accurately as has been depicted in the paper.

Artificial neural networks (ANNs) have been used for a selection of applications together with pattern matching, data

mining and data prediction. Training requirements big training sets and a lot processing time. An ANN is made from a gathering of artificial neurons where each neuron consists of an opening function defined according to the application. There are various types of ANNs in the literature and can be classified into networks based on supervised and unsupervised learning. In the supervised method, the neural network has the input data as well as the target output whereas in the case of unsupervised learning, only the input data is provided.

To speed up training decrease the time for model selection, several methods have been proposed, one of which is to reduce the training set size called data selection. Recently, there has been large research on data selection for SVM training. This study seeks to study the effect of

applying a system based on a support vector machine (SVM) to predict the results of data in general purpose computer.

Data mining is the term used to explain the process of extracting value from a database. Data mining involves the use of complicated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. Accordingly, data mining consists of more than collecting and managing data, it also includes analysis and prediction. Some of these changes include the growth of computer networks, which can be used to connect databases.

II. RELATED WORK

Neural network method is used for classification, clustering, feature mining, prediction and pattern recognition.

Data mining process can be composed by three main phases:

- A. data preparation,
- B. data mining,
- C. expression and interpretation of the results,

Data mining process is the reiteration of the three phases. In the neural network, there are two methods to training data-supervised and unsupervised. Supervised training involves a system of given that the network with the providing the desired outputs with the inputs. Unsupervised training is where the network has to make logic of the inputs without outside help.

2.1 Supervised Training.

In supervised training, both the inputs and the outputs are needed. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the control the network. The set of data which enables the training is called the "training set." In the training of a network the same set of data is processed many times as the connection are ever refined.

2.2 Unsupervised, or Adaptive Training.

The other type of training is called unsupervised training. In unsupervised training, the network is provided with inputs but not with desired outputs.

2.3 SVM Prediction System Design

Time-series data: The time-series data create a sequence of data that is together at regular intervals over a period of time.

Data pre-processing: Two data pre-processing features are introduced to the system; namely, "replace missing value" and "normalization". "Replace missing value" is used to replace missing values since SVM does not support missing value.

III. METHODOLOGY

3.1 Neural Network

Generally neural networks have layers of interconnected nodes where each node producing a non-linear function of its input and input to a node may come from other nodes or directly from the input data. Also, some nodes are recognized with the output of the network.

3.2 Artificial Neural Network

ANN has been widely known as a methodology for classification of complex datasets. It has the ability to efficiently describe the non-linear connection of datasets. ANN structure is shown in fig 1 with an input layer, hidden layer and output layer each consisting of nodes. There might be one or more hidden layers, depending upon the problem.

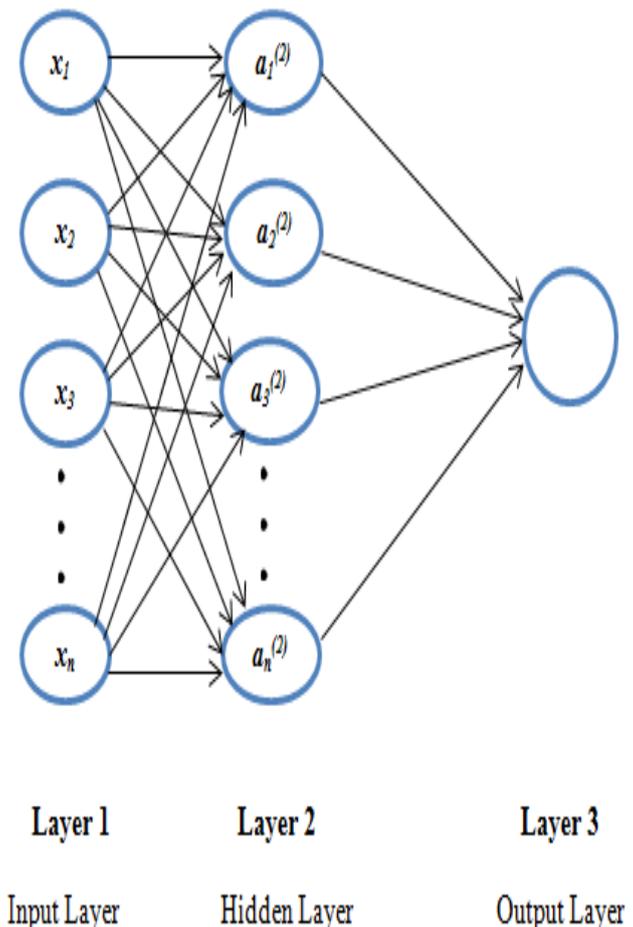


Fig 1 Structure of Artificial Neural Network

3.3 SVM Prediction System Design and Implementation

The design of the proposed system and the step-by-step implementation of the five modules are shown in Fig. 2.

Time-series data: The time-series data compose a sequence of data that is collected at regular intervals over a period of time

Data pre-processing: Two data pre-processing features are introduced to the system; namely, “replace missing value” and “normalization”. “Replace missing value” is used to replace missing values since SVM does not support missing value.

Predictive model: This model, as shown in Figure 2, describes an SVM predictive model.

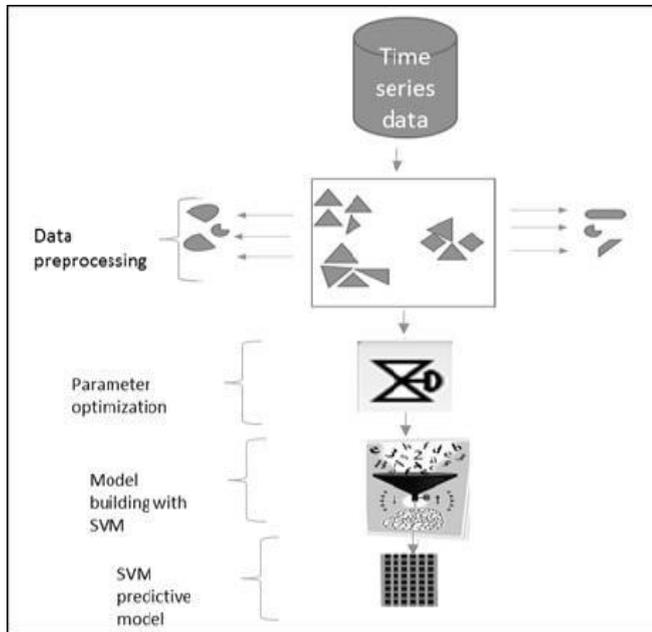


Fig 2: implementation for the SVM prediction system

IV.CONCLUSION

The selection optimum values of the parameters for SVM are an important step in SVM design. When applied to a large data set, however, it requires a long time for training so the model selection task and its performance can be degraded a long time. To reduce the time for model selection, in this page we propose a training data selection method then apply the model selection on reduced training set.

At present, neural network is very suitable for solving the problems of data mining because its characteristics of good robustness, self-organizing adaptive, parallel processing, distributed storage and high degree of fault tolerance. Hence, the use of neural networks in data mining is a promising field of research especially given the ready availability of large mass of data sets and the reported ability of neural networks to detect and assimilate relationships between a large numbers of variables

Neural networks can be used for prediction with various levels of success. The advantage of then includes automatic learning of dependencies only from measured data without any need to add further information (such as type of dependency like with the regression). The neural network is qualified from the past data with the hope that it will determine unseen dependencies and that it will be able to use them for predicting into future.

The main purpose of this study was to investigate the applicability and capability of the SVM methods for modeling of time-series forecasting. To verify the application of this approach, the benchmarked data are used in this study.

The well known data are five well-known data sets that always handled in real life time series application. There are the chemical process concentration, the IBM common stock closing prices, the chemical process temperature, the Wolf's sunspot data and the international airline passengers. The prediction result by SVM method is compared with those by ANN.

ACKNOWLEDGEMENT

I wish to express my profound thanks to all who helped us directly or indirectly in making this paper.

Finally I wish to thank to all our friends and well-wishers who supported us in completing this paper successfully

I am especially grateful to our guide **Prof. G. A. Kulkarni** Sir for him time to time, very much needed, valuable guidance. Without the full support and cheerful encouragement of my guide, the paper would not have been completed on time.

REFERENCES

- [1] Agrawal, R., Imielinski, T., Swami, A., “Database Mining: A Performance Perspective”, IEEE Transactions on Knowledge and Data Engineering, pp. 914- 925, December 1993.
- [2] Berry, J. A., Lindoff, G., Data Mining Techniques, Wiley Computer Publishing, 1997 (ISBN 0-471-17980-9).
- [3] Berson, “Data Warehousing, Data-Mining & OLAP”, TMH.
- [4] Haykin, S., Neural Networks, Prentice Hall International Inc., 1999.
- [5] Khajanchi, Amit, Artificial Neural Networks: The next intelligence.
- [6] Zurada J.M., “An introduction to artificial neural networks systems”, St. Paul: West Publishing (1992).
- [7] Y. Bengio, J. M. Buhmann, M. Embrechts, and J.M. Zurada. Introduction to the special issue on neural networks for data mining and knowledge discovery. IEEE Trans. Neural Networks.
- [8] M. W. Craven and J. W. Shavlik. Using neural networks for data mining. Future Generation Computer Systems, 13:211–229, 1997.
- [9] Abe, S., Inoue, T. Fast training of support vector machines by extracting boundary data. In: Proceedings of the International Conference on Artificial Neural Networks (ICANN) (2001) 308–313.
- [10] Almeida, M. B., Braga, A. P., Braga, J. P. SVM-KM: speeding SVMs learning with a priori cluster selection and kmeans. In Proceedings of the 6th Brazilian Symposium on Neural Networks (2000) 162–167.
- [11] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, A practical guide to Support Vector Classification, Libsvm: a library for support vector machines, 2005.

[12] Raul Rojas. Neural Networks - A Systematic Introduction, Springer.1996.

[13] Owens J D et al. A survey of general purpose computation on graphics hardware.Computer Graphics Forum. Vol 26, 2007.