iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** VII **Month of publication:** July 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Comparison Of Various Kernels Of Support Vector Machine

Supriya Pahwa[1], Deepak Sinwar[2]
[1]M.Tech Scholar (CSE), [2]Assistant Professor (CSE)
BRCM CET, Bahal, Bhiwani, Haryana, India

Abstract— As we know, classification plays an important role in every field. Support vector machine is the popular algorithm for classification and prediction. For classification and prediction by support vector machine, LIBSVM is being used as a tool. Support vector machine classifies the data points using straight line. Some datasets are impossible to separate by straight line. To cope with this problem kernel function is used. The central idea of kernel function is to project points up in a higher dimensional space hoping that separability of data would improve. There are various kernels in the LIBSVM package. In this paper, Support Vector Machine (SVM) is evaluated as classifier with four different kernels namely linear kernel, polynomial kernel, radial basis function kernel and sigmoid kernel. Several datasets are being experimented to find out the performance of various kernels of support vector machines. Based on the best performance result, linear kernel is capable of classifying datasets accurately with the average accuracy 88.20 % of correct classification and faster with 4.078 sec of prediction time. Radial basis function Kernel is capable of taking less training time compared to other kernels that is 4.92675 sec.
Keywords— SVM, Libsvm, Kernels, classification

I. INTRODUCTION

As we know, Data mining is the process of analysing large amounts of data to extract the patterns and the useful information, and classification is one of the techniques of data mining. In the area of classification, Support vector machine (SVM) plays an important role as the classifier. Basically SVM works on the binary classification. SVM is used to classify the class label by separating the data point with straight line. But in some datasets, it is not possible to separate the data point by one straight line. To cope up with this problem some kernel functions are introduced [5, 9]. Kernel functions project the data points up in a higher dimensional space so that the data points would easily separate by straight line.

Motivation: Consider a bank scenario, having two types of customers. One customer is trustworthy so that loan could be given to him and other is fraud, to whom bank doesn't give loan. Manager of bank does not have time to manually check the details and put the customer in one of the two categories. The manager may adopt support vector machine to automatically predict the class of a new customer by the old databases of the customers. As customers increase, it takes long time to predict with better accuracy. So, we can use different kernels and parameters to reduce the time and improve the accuracy. So, classification will be accurate and fast. In this paper, various basic kernel functions such as linear kernel, radial basis function kernel, polynomial kernel and sigmoid kernel has been compared by the classification accuracy, training time and prediction time. The remainder of this paper is organized as follows. Section II will review SVM and various kernels. Section III, will describe comparison of different kernel; and Section 4 reports conclusion.

II. BACKGROUND

In the last few years, data mining has been widely used as a powerful data-analysis tool in a various fields: not only in the computer science, but also in medicine, health, sociology, physics etc [9]. Data mining is also used to predict labor market needs. For this three techniques can be implemented which are Decision Trees, Naïve Bayes Classifiers and Decision Rules techniques. The tables are created to perform data mining task which are known as training tables. The sets of these tables were generated by using various factors such as instances, class label, features etc. For prediction the training tables are used to predict the classification of other instances those are unclassified, and tabulate the results of the unknown instance for classification [6]. Data Stream Mining is also one of the areas gaining lot of practical significance and is progressing in various fields with new methodologies. And find in various applications related to computer science, medicine, bioinformatics and stock market prediction, weather forecast, audio and video processing, text etc [8]. Classification is one of the oldest, popular and the most important technique of data mining.

©IJRASET 2015: All Rights are Reserved

Classification is of two types: supervise and unsupervised. Supervise classification means learning from data that is already classified correctly, and using the pre-built model to find the classification accuracy for the new data or the testing set. In the unsupervised classification, the pre-build model is used to predict the new data with no class labels. There are four criteria which can be used to compare between the classifiers: the accuracy or the percent of class labels that are classified correctly, the speed or the computational cost of both training model and testing process, the robustness or the ability to cope with missing or noisy data and the scalability or the ability to handle very large amounts of data [3]. Support vector machine (SVM) is the popular and most important technique of classification and was developed by Vladimir Vapnik. It is based on statistical learning theory. In the classification of small datasets, SVM has yielded excellent performance that is hardly provided by any other method and able to solve practical problems such as high dimension, over learning and local minima. The standard Support vector machine algorithm is leads to a quadratic optimization problem with bound constraints and one linear equality constraint. However, when the datasets are large with large number of data points, the quadratic programming solvers become very difficult, because their time requirements and memory are highly dependent on the size of the training datasets [7]. This is the only limitation of support vector machine. Support vector machine works on the kernel function which is used to project the data points to higher dimensions for better accuracy of classification. SVM which is kernel based algorithms have achieved considerable success in various problems in the classification where all the training data is available in advance. Support vector machine combines the kernel trick with the large margin idea. Various kernels are used to classify the data by support vector machine such as linear kernel, sigmoid kernel, polynomial kernel, radial basis function kernel etc. Support vector machines (SVMs) and kernel methods (KMs) have become very popular as techniques for learning. New kernel expert system is also introduced by R.Zhang, W.Wang for better classification performance [4]. These kernels basically depend on the number of support vectors. There are some kernels present in the literature those are independent of number of support vectors namely: intersection kernel, chi-squared kernel, additive kernel [5, 1].

SVMs (Support Vector Machines) are the efficient technique for data classification and prediction. It works on the principle of supervised learning. As we discussed that kernel function plays an important role in the classification by support vector machine. Kernels are used to project the data points in the higher dimensions for better classification of the datasets as shown in fig.1. Some kernel functions are present in support vector machine algorithm are based on neural networks. Support vector machine is considered easier to use than neural networks but time taken by support vector machine is more compared to neural network [2]. The radial basis kernel, polynomial kernel and sigmoid kernel of support vector machine is used for non-linear separation and works on the principle of neural networks.

A. Kernels

As we know, kernel function is used to project the data points to higher dimensional space for better classification. There are various kernels which are used to improve the performance of classification by support vector machine namely: linear kernel, Radial basis function kernel, polynomial kernel and sigmoid kernel. Kernel is used to project the data point to higher dimensional space to improve its ability to find best hyperplane to separate the data points of different classes. The kernel function used are described below-

Linear Kernel- Linear kernel separates the data points linearly by using straight line as shown in fig. 2. Linear kernel is good at classifying two classes at a time. Mapping of data points to higher dimension is not required.

Polynomial Kernel- The polynomial kernel is a kernel function commonly used with support vector machines (SVMs) and other kernelized models, that is similar to vectors (training samples) in a feature space over polynomials of the original variables. It generally works with non-linearly separable data.

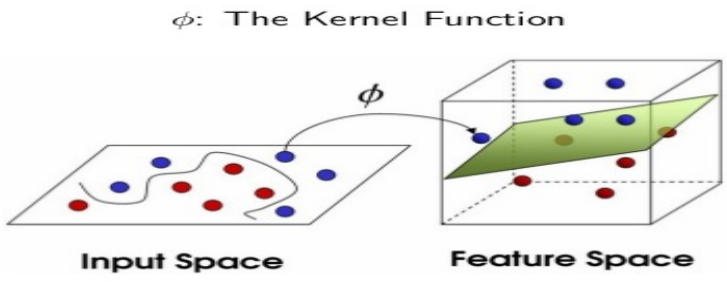


Fig 1. Function of kernels in support vector machines [10].

©IJRASET 2015: All Rights are Reserved

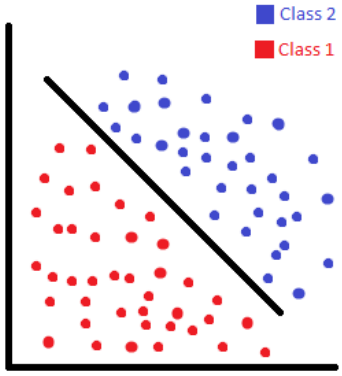


Fig 2. Separation of data points by straight line.

Radial basis function kernel- The radial basis function network is an artificial neural network that uses radial basis functions as activation functions. It mainly used to classify non-linear data. Radial basis function networks have many uses, including time series prediction, function approximation, classification and system control.

Sigmoid Kernel- The sigmoid kernel is very popular for support vector machines due to its origin from neural networks. As we know neural network is another approach for the classification with less time consumption.

III. COMPARISON OF DIFFERENT KERNELS

A. Compared Method

We focused on classification tasks. For SVMs, we tested the various datasets by using different kernel with using similar parameter. Four different kernels are compared on the basis of accuracy, training time and prediction time.The result has shown in table I, table II and table III.

Linear kernel gives the best result for the accuracy than other kernels. As we know, linear kernel separates the data linearly and separates the data by straight line. Result is shown in table I.

Radial basis kernel takes less training time than other kernels. It is capable of separating the data non-linearly. Results showing training time taken are shown in table II.

Linear kernel takes less prediction time compared to other kernel. Results are shown in table III.

Datasets	Prediction Time by Linear Kernel	Prediction Time by Polynomial Kernel	Predic -tion Time by Radial basis function Kernel	Predic -tion Time by Sigmoid Kernel
a1a	2.11	2.998	3.042	4.024
a6a	11.964	16.906	14.338	19.296
w7a	2.23	3.412	5.295	5.197
Australia	0.008	0.018	0.08	0.066

TABLE I: ACCURACY BY USING VARIOUS KERNEL

©IJRASET 2015: All Rights are Reserved

Datasets	Accuracy by Linear Kernel	Accuracy by Polynomial Kernel	Accuracy by Radial basis function Kernel	Accuracy by Sigmoid Kernel
a1a	83.9085	76.0267	83.6686	82.157
a6a	84.7242	75.8727	84.1713	84.0307
w7a	98.683	97.0507	97.3461	97.1585
Australia	85.5072	55.5072	55.5072	74.7826

TABLE III: TRAINING TIME BY USING VARIOUS KERNEL

Datasets	Training Time by Linear Kernel	Training Time by Polynomial Kernel	Training Time by Radial basis function Kernel	Training Time by Sigmoid Kernel
a1a	0.298	0.241	0.252	0.315
a6a	12.589	12.417	11.745	13.727
w7a	7.46	4.012	7.606	8.554
Australia	21.066	85.892	0.104	0.059

TABLE IIIII PREDICTION TIME BY USING VARIOUS KERNEL

B. Software

We used Intel-Core i3-370M Processor 2.40 GHz with window 7 professional (32-bit), 4GB RAM and 500 GB Hard Disk. We used LIBSVM 3.20 and MATLAB r2010a for Classification by support vector machine.

C. Methodology

1) *Preprocessing:* Convert numerical dataset into sparse format and if the data sets given is not in the sparse format then it is first converted to the libsvm format because libsvm package could not work on the format other than sparse.
2) *Kernel Function:* There are various kernel in the libsvm package which can be chosen by changing the value of '–t' parameter which is denoted as the type of kernel used. SVM models were obtained using the linear kernel function, polynomial kernel function, radial basis function kernel function or sigmoid kernel function.
3) *Parameters And Model Selection*: c-svm (support vector machine with cost function) is the model used and different kernels are selected to find the best accuracy, training time and prediction time for the datasets used.
4) *Model Training And Testing*: The models were trained and tested over four training and testing data set.

D. Results

Result is shown in Fig 3 and Fig 4. There are a graph to describe the average result of comparison of accuracy and CPU Time of four different kernels. These graphs show the performance of four kernel functions namely: linear kernel, radial basis function kernel, polynomial kernel and sigmoid kernel. Fig 3 shows that linear kernel with different datasets gives high accuracy. Fig 4 shows that Radial basis function kernel gives best performance for training time in classification by SVM and Linear kernel gives best performance for prediction time.

©IJRASET 2015: All Rights are Reserved

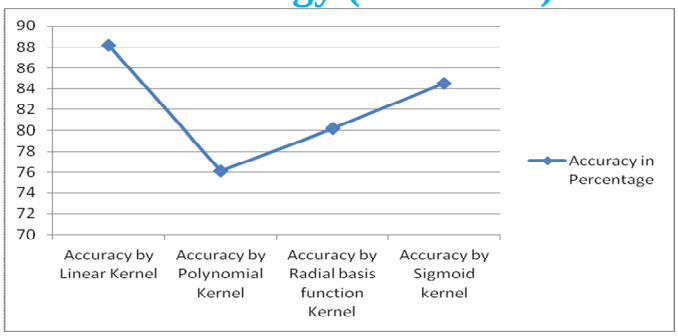


Fig. 3 Result of average accuracy by different Kernels.

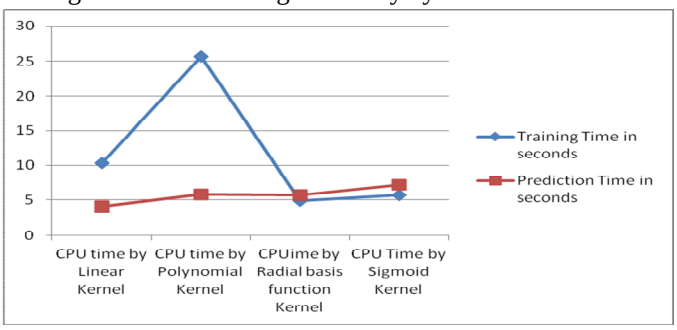


Fig. 4 Result of average prediction time and training time by different Kernels

E. Data Sets

Several benchmark datasets from the LIBSVM website were used for the comparison, namely a1a, a6a and w7a. These datasets were already in sparse format. One dataset were used from UCI repository. This dataset was numerical and we converted his dataset to sparse format or libsvm format. All datasets have two class labels or binary classes.

IV. CONCLUSIONS

SVM classification performed its task successfully. The Support vector machine classifiers have been tested on several binary datasets. Different types of kernel function namely: linear, radial basis function kernel, polynomial Kernel and sigmoid kernel are used to perform classification task and all four give different results. Linear kernel gives the best performance in average of 88.20% correct classification as compared to the other types of Kernel function and prediction time with average of 4.078 sec . The experiment shows that linear kernel is good at prediction time and classification accuracy but in the case of training time, Radial basis kernel gives best results with the average of 4.92675 sec, which is less time taken by other kernels..

REFERENCES

[1] S Maji, A Berg & J Malik (2008). Classification using intersection kernel support vector machines is efficient. In proceeding of IEEE conference on computer vision and pattern recognisation, (pp. 1-8).

[2] C.W Hsu, C.C Chang, and C.J Lin(2010). A Practical Guide to Support Vector Classification. National Taiwan University, Taipei 106, 1-16.

[3] E.Al-Daoud (2010). Integration of Support Vector Machine and Bayesian Neural Network for Data Mining and Classification. World Academy of Science, Engineering and Technology Vol: 4 , 145-150.

[4] R Zhang, W Wang (2011). Facilitating the applications of support vector machine by using a new kernel. Expert Systems with Applications 38, 14225–14230.

[5] S Maji, A.C Berg & J Malik (2013). Efficient classification for additive kernel SVMs. IEEE transactions on pattern analysis and machine intelligence. Vol. 35, NO. 1, 66-77

[6] Y.A Alsultanny(2013). Labor Market Forecasting by Using Data Mining. Procedia Computer Science 18, 1700 – 1709.

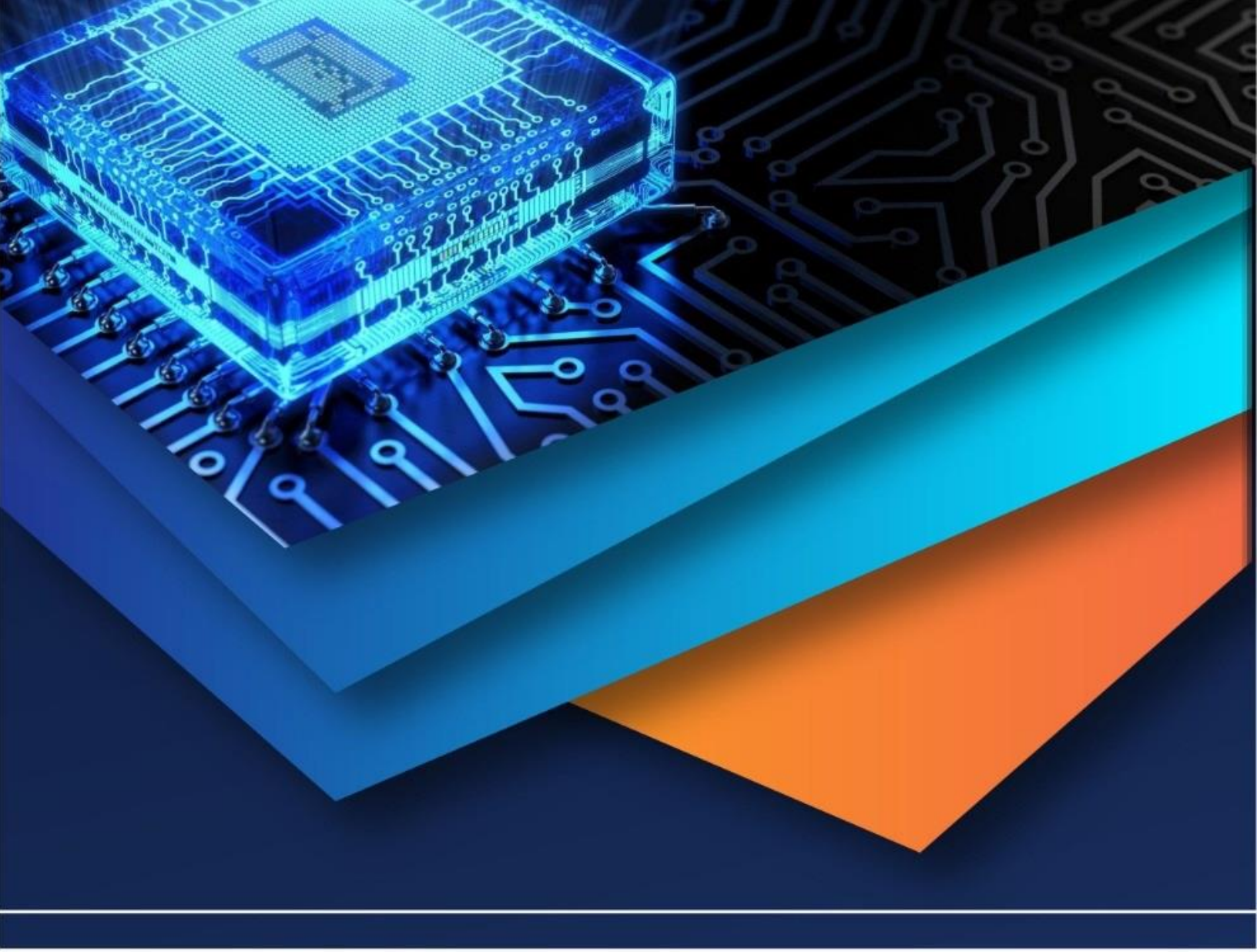
[7] G Lixin, X Weixin, P Jihong (2014). Design of SVM based on radial basis function Neural Networks pre-partition.ICSP2014 Proceedings. 1480-1483, IEEE.

[8] M.S.B PhridviRaja, C.V GuruRaob (2014). Data mining – past, present and future – a typical survey on data streams. Procedia Technology 12, 255 – 263.

[9] F.Gullo (2015). From Patterns in Data to Knowledge Discovery: What Data Mining Can Do. Physics Procedia 62 ,18 – 22.

[10] Identification-of-relevant-sections-in-web-pages-using-a-machine-learning-approach."http://image.slidesharecdn.com/presentation-121109113051 phpapp02/95/identification-of-relevant-sections-in-web-pages-using-a-machine-learning-approach-11638.jpg?cb=1352461178".visited on June,22,2015

©IJRASET 2015: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)