

A New Multi Kernel Hybrid Support Vector Machine Data Classification Algorithm

R.Anusha *, Dr.(Mrs)N.Rama **

*Research Scholar, University of Madras Chennai, Tamilnadu, India

** Assoc.Professor, Presidency College Chennai, Tamilnadu, India

Abstract- The objective of this paper is to propose a new multi kernel hybrid support vector machine data classification algorithm to classify data. This paper also focuses on comparing various data classification algorithms such as Support vector machine algorithm with linear kernel, Radial basis Kernel, Neural network feedforward back propagation algorithm with the proposed algorithm to measure the accuracy in classification of data collected for evaluating the website quality and usability of various online shopping websites such as Flipkart, Amazon, Jabong and Snapdeal by taking into consideration various quality parameters such as functionality, reliability, security, navigation, information quality, presentation, visual interface etc.

Index Terms- Data classification, Support vector machine, Website quality

I. INTRODUCTION

Businesses and researchers have been generating huge amount of data which has led to the need for fast, accurate and robust algorithms for analyzing the data for meaningful interpretations and inferences. Classification is one of the crucial tasks for various applications like text categorization, tone recognition, image classification, micro-array gene expression, proteins structure predictions, data Classification etc. The existing supervised classification methods can provide the best results only for finite samples. However the samples generated by real time applications such as credit card transactions etc are huge and unbalanced and this leads to non linear relationship between predictive variables and such samples are easily classified by support vector machine data classification algorithms. Support Vector Machine (SVM) is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data.

II. DATA CLASSIFICATION USING SVM

SVM is a useful technique for data classification. Even though it's considered that Neural Networks are easier to use than this, however, sometimes unsatisfactory results are obtained. A classification task usually involves with training and testing data which consist of some data instances¹. Each instance in the training set contains one target values and several attributes. The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the

attributes. Classification in SVM is an example of Supervised Learning. Known labels help indicate whether the system is performing in a right way or not². This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly³. Support Vector Machine algorithm supports various kernels like linear kernel, Radial basis kernel, Gaussian kernel etc to classify data.

III. RESEARCH FINDINGS

A new multi kernel hybrid data classification algorithm by combining linear and radial basis kernel was proposed and was used to classify the data collected for evaluating the website quality of various online shopping sites such as Flipkart, Amazon, Jabong and Snapdeal. The proposed algorithm was compared with support vector machine with linear kernel algorithm, support vector machine with radial basis kernel algorithm and feedforward back propagation neural network algorithm to test the accuracy of the data classification.

Pseudo Code for the proposed Multi kernel hybrid data classification Algorithm

Input: Training set $(v_1, I_1), \dots, (v_N, I_N)$

Output: Multiclass SVM and RBF Classifier

Training: Binary SVMs and graded relevance scores

For $j = 1$ to $(k - 1)$ do

STEP1: For all samples from C_1 to C_j classes, set labels to $(+1)$ and all samples from C_{j+1} to C_k , set labels to (-1)

STEP2: Train j th binary SVM

STEP 3: Classify the training samples

STEP 4: if $(j > 1)$, compute scores for all training samples classified as $(+1)$ and define $(j - 1)$ thresholds by splitting the curve of sorted relevance scores into equally spaced intervals.

STEP5: if $(j < k)$, compute scores for all training samples classified as (-1) and define $(k - j - 1)$ thresholds by splitting the curve of sorted relevance scores into equally spaced intervals.

End for

The results are tabulated below:

Comparative analysis of Performance of proposed multi kernel hybrid data classification algorithm, SVM linear Classification algorithm, RBFSVM algorithm and feed forward back propagation neural network algorithm for Functionality factor

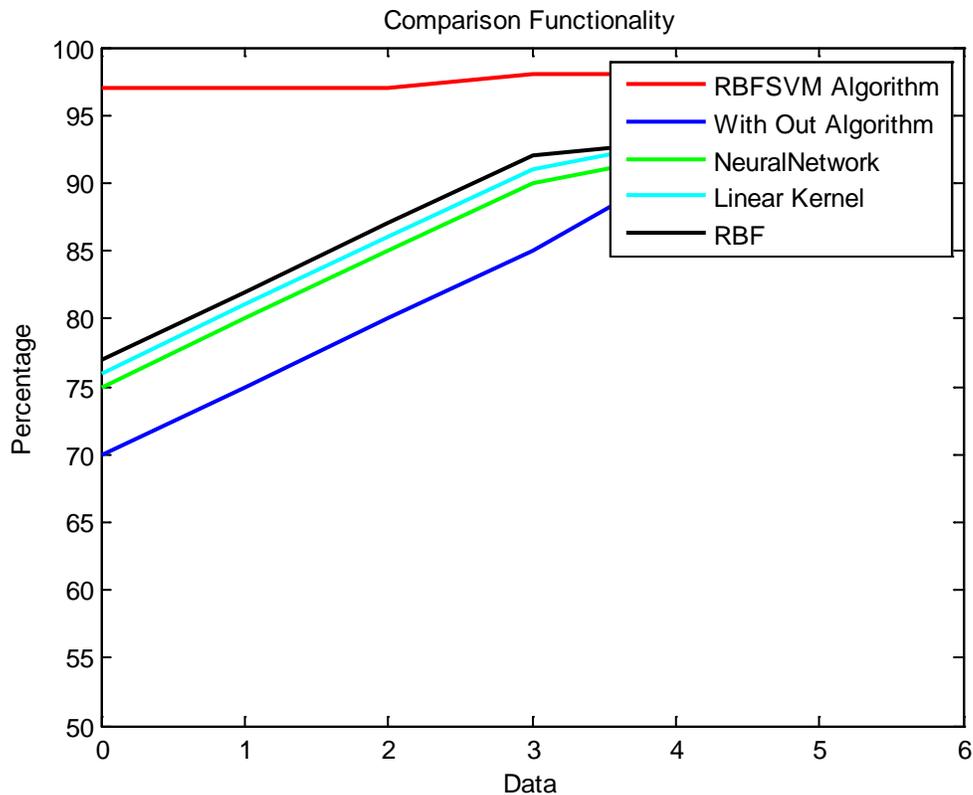


Figure 1: Graph –Comparative Performance Analysis of algorithms for Functionality Factor

TABLE 1: Comparison of accuracy of various data classification algorithms

Parameters	Accuracy measure of classification Algorithms (in Percentage)			
	Linear kernel	RBF Kernel	Neural Network	Proposed Multi Kernel Hybrid
Functionality	70	78	75	83
Reliability	77	78	75	95
Efficiency	77	78	75	95
Security	76	78	75	95
Navigation	76	78	75	94
Information Quality	77	78	75	95
Visual Interface	76	77	75	92
Interactivity	76	77	75	93
Presentation	76	78	75	92
Satisfaction	76	77	75	92
Trust	76	77	75	90
Usability	76	77	75	90

IV. CONCLUSIONS

The research findings indicate that the performance of the proposed multi kernel hybrid data classification algorithm in terms of accuracy was improved and better when compared to the other existing data classification algorithms such as linear kernel, radial basis kernel and feed forward back propagation neural network data classification algorithms.

REFERENCES

- [1] Lijuan Liu¹, Bo Shen^{1,2}, Xing Wang, "Research on Kernel Function of Support Vector Machine", Journal of Computers Vol. 25, No. 1, April 2014
- [2] Z. Yang, "Kernel-based support vector machines," Computer Engineer and Applications, Vol. 44, No. 33, pp.1-6, 2008.
- [3] Himani Bhavsar, Mahesh H. Panchal, "A Review on Support Vector Machine for Data Classification", International Journal of Advanced

Research in Computer Engineering & Technology (IJARCET) Volume 1,
Issue 10, December 2012

Second Author: Dr(Mrs)N.Rama,Assoc.Prof,Presidency
College,Chennai.Email:nramabalu@gmail.com

Correpondence author: R.Anusha,
email:rajamani_anusha@hotmail.com,ph:9500084482

AUTHORS

First Author: R.Anusha,Research Scholar,University of
Madras,email:rajamani_anusha@hotmail.com