Support Vector Machine Modeling In Education Datamining – A Survey

N. Ayyanathan

Director(MCA), KLN College of Information Technology, Madurai

A. Bavani Rekha

*Guest Lecturer, Department of English Language Studies,
Madurai Kamaraj University, Madurai.*

Abstract

Learning outcome of students sample size 'n' is to be assessed based on several parameters. Traditional statistical measures are outperformed by neural network classifiers and support vector machine models. This paper analyses the specific application of support vector machine modeling in education datamining which helps as a decision support system in Education Management.

Keywords: Support vector machine, Education datamining, Neural network classifiers, Predictive datamining.

Introduction

Academic performance was confined to only scoring marks in the final exams and was considered as the only preferred parameter till recently.

The key essentials of campus placement performance of an institution changed this totally and the performance measurement is now a days based on other soft skill parameters, along with major metric called 'MARKS'.

Placement performance is just a major milestone in a student's career, whereas learning out come of the same student is very unpredictable and nonlinear in nature.

The development of computational model with support vector machine is hence justified in education datamining.

The authors preferred to do a detailed review of literature on support vector machine in education data management in this survey paper.

Literature Review

Datamining is on the interface of Computer Science and statistics, utilizing advances in both disciplines to make progress in extracting information from large databases. (Clark Glymour, et al., 1997).

Learning in neural networks are of two kinds namely supervised and unsupervised. The result of supervised learning is a combined estimate. The result of unsupervised learning is quantification of data clusters and classification (Satishkumar, P.157, 2009).

The simplest perception learning rule that adjusts weights of neurons to respond correctly to a two class classification problem, updates the weight vector by respectively adding or subtracting a mis-classified input vector scaled by learning rate. This amounts to a movement perpendicular to the pattern hyperplane in weight space (Satish Kumar, P.156, 2009).

The universal approximation theory proves that there exists a feed forward neural network architecture that can approximate any continuos function to an arbitrary degree of accuracy provided enough hidden neurons are incorporated into the architecture. (Satish Kumar, P.212, 2009).

For high degree of accuracy, multilayer architecture must be employed. Neurons must arrive from linear to sigmoidal regime. The degree of accuracy depends on number of hidden neuron. (Satish Kumar, P.213, 2009).

The design of SVM (Support Vector Machine) is based on quadratic optimization of a Lagrangian in the dual space. SVM can be used for both classification and regression. (Satish Kumar, P.340, 2009).

Data of Methodology

The periodic assessment of students' performance in the internal tests, semester examination, co-curricular entities, aptitude tests are the predictors of the data set , the target vector being the placement performance in the campus selection by the corporate companies.

Neural network models based on the above predictors with which many authors applied support vector machine modeling out performed traditional methods.

However employability still metrics need to be further classified and assessed under the major sub division of soft skills viz., communicative skills, which include verbal and non-verbal communication, interpersonal speaking, hierarchal communication, verbal etiquette and public speaking.

The next important skill set is classified under interpersonal skills, stress management, time management, confidence level, leadership, team work,

attitude, motivation, critical thinking, need to be further evaluated for any given sample for the research study.

Along with above skill set variables, problem solving skill, decision making capacity, honesty, integrity, adaptability, analytical thinking need to be assessed and proper weightage must be given for the over all evaluation.

Support Vector Machines

A set of features that describes one case (i.e a row of predictor values) is called a vector. Predictor variable is called an attribute. A transformed attribute that is used to define the hyperplane is called a feature. The task of choosing the most suitable representation is known as feature selection. (www.dtreg.com).

The goal of SVM modeling is to find the optimal hyperplane that separates clusters of vectors in such a way that cases with one category of the target variable are one side of the plane and cases in the other category are on the other side of the plane. The vectors near the hyperplane are support vectors. (www.dtreg.com)

Conclusion

The research design for the classification of employability skill metrices for Indian engineering students is proposed by the authors based on the degree of reachability and degree of validity in the development of neural classification models.