International Journal of Engineering & Technology, 7 (2.4) (2018) 178-181

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper



Trend analysis of university placement by using machine learning algorithms

Chandrasekhar Kumbhar [1] *, Dr. S. S. Sridhar [1]

[1] Department of Computer Science Engineering, Kattankulathur Campus, SRM Institute of Science and Technology
**Corresponding author E-mail: 1chandrashekhar_b@srmuniv.edu.in*

Abstract

Machine learning is a method of data analysis that automates analytical model building. These models help you to make a trend analysis of university placements data, to predict a placement rate for the students of an upcoming year which will help the university to analyze the performance during placements. Many students look at universities as a means of investment which can help them make a great future by getting placed in good companies and which will relieve their stress and unease from their lives before graduating from the university. The trend will also help in giving the companies reasons as to why they should visit university again and again. Some attributes play the very important role while analyzing the student for e.g. Student's name, Department, Company, Location and Annual package. So, classification can help you to classify those data and clustering helps to make the clusters department wise. In this paper we have used neural networks to predict the upcoming student placement and got 77% of accuracy while testing were iteration are 1000. Through extensive trend analysis of varies complex data collected from different sources, we can demonstrate that our analysis can provide a good pragmatic solution for future placement of students.

Keywords: *Support Vector Machine; Decision Tree and Artificial Neural Network.*

1. Introduction

This paper focuses on the trend analysis of university placement. A trend analysis means an aspect of technical analysis that tries to predict future movement on past data. Students are the main stakeholders of universities and their performance plays a significant role in a country's social and economic growth by producing creative graduates, innovators and entrepreneurs. a millions of university and colleges provides the technical education nowadays. But few of them land in good jobs as compared to the total. A university placement plays a very important role to get more admission and maintain the reputation in the world. Authors consider some of the important attributes of students like company, package, location etc. a machine learning plays the most important role to find out the performance, define the trend or predict the placement for the upcoming year.
The dataset considered for this work is the CSE department in which we have initially considered four attribute of each student i.e. 10th marks, 12th / Diploma marks, B.Tech and M.Tech marks. With the help of deep neural network classifier and neural networks have been known to successfully predict the future of student, out of 50 records 37 are used for training and 13 used for testing purpose and we got 77% accuracy while testing.

2. Literature survey

1) The author presents the development of placement predictor system (PPS) using logistic regression model. Based on the student scores in matriculation, senior secondary, subjects in various semesters of technical education and demographics, PPS predicts the placement of a student in upcoming recruitment session. The author used logistic regression and Gradient Descent Algorithm to implement placement predictor system.
2) The student placement is identified by using Naïve Bayesian Classification, Multilayer Perception, and C4.5 Tree. The best algorithm based on the placement data is Naïve Bayes Classification with an accuracy of 86.15%. This paper focuses on the classification of the student for the placement process in an engineering college. So author suggests a system which makes the work of prediction of student's placement easy. For this, they suggest to use ID3 algorithm- data mining classification algorithm. They have compared CHAID, C4.5 and ID3.
3) Using the student datasets and attributes, predictions are made using the Data Mining Algorithm "Fuzzy logic" and "K nearest neighbor (KNN)". The results obtained from each approach are then compared with respect to their "performance" and "accuracy".
4) This paper focuses on Predicting the performance of the M.sc student. author collect 5 Different department data and apply Naïve Bayes and 1-NN algorithms on It. They have got accuracy 85.71%. Parameter they have considered as sex, age group, marital status, bachelor, another master. They have used weka tool to find out the accuracy.
5) Writer proposed the concept of the teaching of machine learning techniques using software Weka as a possible solution to integrate this issue into the study plan based on the example of problem-based learning. Techniques used linear regression Logistic regression Factor analysis Cluster analysis (hierarchical and non-hierarchical) Naive Bayes algorithm k-nearest neighbors.
6) Predicting the success or failure of a student in a course or program like how many students got cleared course or pro-

Copyright © 2018 Chandrasekhar Kumbhar, Dr. S. S. Sridhar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

gram in given time and the first attempt is a problem that has recently been addressed using data mining techniques k-Nearest Neighbors, Random Forest, Classification and Regression Trees, Support Vector Machines, Naive Bayes and for regression they used Ordinary Least Squares.

7) Author implemented a system which was used to analyze the performance of students using k-means clustering algorithm. They cluster the student marks according to their percentage. i.e. 75% Dist and above, 60% first class. The parameter considered are 10th 12th B.Tech, project, internship and skill set. The algorithm used is K-means and Rapidminer studio is used.

3. System architecture

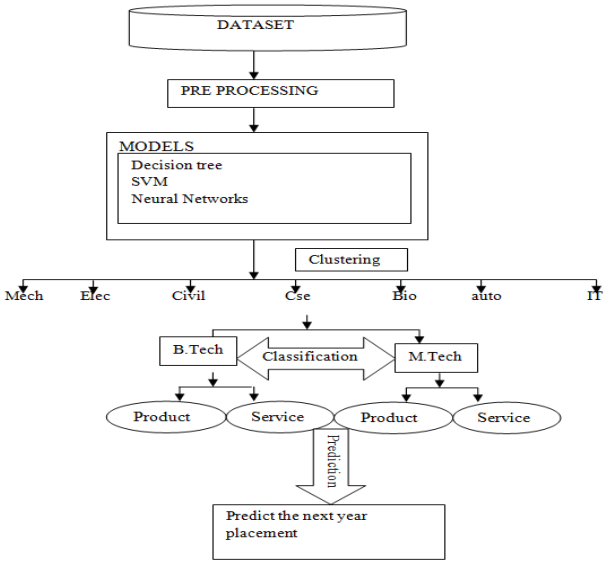


Fig. 1: System Architecture.

The architectural design consists of the various algorithms/ Methods in machine learning. In the first phase student's data will be collected from the university. In the second phase, data will be cleaned manually by removing duplicate and unnecessary columns. In the third phase the varies techniques are applied to discover the knowledge. To predict the student's placement for upcoming year, the following algorithms are used to build the prediction model.

- Decision tree
- Support vector machine
- Neural networks

4. Dataset construction

The dataset was collected from Department of CSE of an engineering college in Pune. A dataset has 50 entries in which we have considered 37 entries for training purpose and 13 for testing. The attribute we have considered i.e. 10th marks, 12th/Diploma marks, UG and PG marks. A programming language python and open source software library Tensor flow are playing important roll to predict the student whether he/she will be placed or not in upcoming year placement.

- Learning Parameters

Table 1: Parameter Used for Model

No of layers 3	
No of hidden layers	No of Nodes
1	10
2	20
3	10

Table 2: Classification Accuracy

Classifier used DNN	
No of titration	Accuracy
1000	71
2000	77
5000	93

5. Methodology

Decision tree
A decision tree in which each branch node represents a choice between a number of alternatives and each leaf node represents a decision. It is type of supervised learning algorithm. That is mostly used in classification problems and works for both categorical and continuous input and output variables. ID3, C4.5, CART are the

5.1. Decision tree algorithms

5.2. Support vector machine

A SVM is a discriminative classifier formally defined by a separating hyperactive plane. The goal is to design a hyperactive plane that classifies all training vectors in two classes. The best choice will be the hyperactive plane that leaves the maximum margin from both classes. If data is linearly separable then we use SVM and if data is non-linear, we use to apply kernel methods to classify the data.

5.3. Neural networks

An Artificial neural network (ANN) is an information-processing paradigm that is inspired by way biological nervous system. Its designed as an interconnected system of processing element each with a limited numbers of input and output rather than being programmed this system learns to recognize pattern. Neural networks are constructed for a specific application, such as patter recognition through a learning process. In biological system, learning involves adjustment to the synaptic connections between neurons.

6. Result and screenshots training

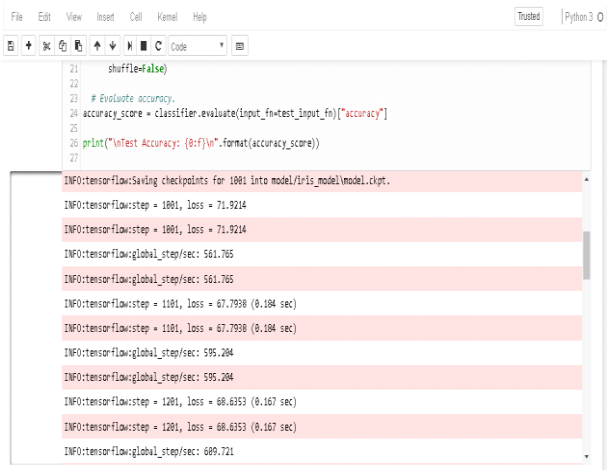


Fig. 2: Training and Data Iteration.

The above figure 2 shows the training of data where number of iteration are 1000. We got 71% accuracy.

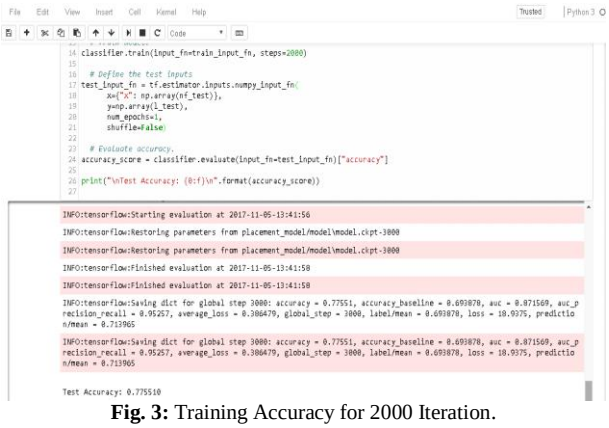


Fig. 3: Training Accuracy for 2000 Iteration.

The above fig 3 shows the 77% accuracy where number of iterations is 2000.

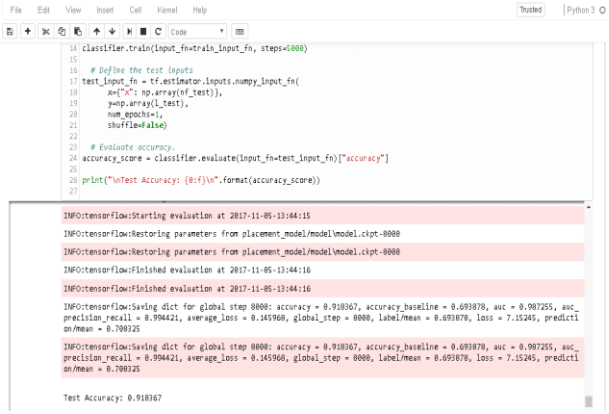


Fig. 4: Training Accuracy for 5000 Iteration.

The above figure 4 shows the accuracy of 93% and number of iterations are 5000.

6.1. Accuracy

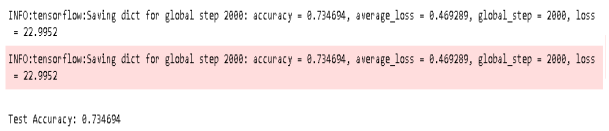


Fig. 5: Accuracy.

6.2. Testing

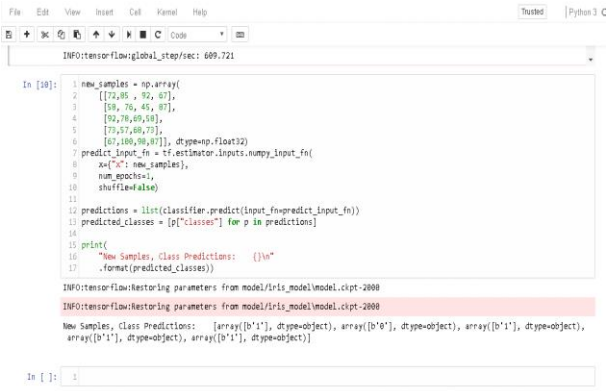


Fig. 6: Testing of Result.

The above fig shows the testing prediction of trend analysis of university placement where we got the output by using DNN classifier that a new student will get placed or not. here 0 indicates he/she will not get placed and 1 indicates get placed.

The iteration sometimes called as steps in Deep neural network classifier from artificial neural networks decides the how steps we have to perform on data. Minimum steps will give you less accuracy and highest steps high accuracy but sometimes high number of steps with high accuracy leads to wrong output. That's why here we used minimum 1000 and max 5000 steps. The amount of steps depends upon the size of your dataset.

7. Future work

A machine learning algorithms are play a very important role while predicting something, in future student performance will tell the institute and student weather he/she will clear the subject or not.

8. Conclusion

The algorithms of machine learning we have discussed are can used to find the trend of placement, which will be helpful for university to get more admission in future. We compared the algorithm and find out the accuracy by considering some of attributes of students. Here we used deep neural network classifier with the 1000, 2000, 5000 iteration with 71%, 77%, and 91% of accuracy.

References

[1] Mangasuli Sheetal B1, Prof. Savita Bakare2 "Prediction of Campus Placement Using Data Mining Algorithm-Fuzzy logic and K nearest neighbour" International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 6, June 2016

[2] Ajay Kumar Pal Research Scholar, Sai Nath University, Ranchi, Jharkhand ,Saurabh Pal Head, Department of MCA, VBS Purvanchal University, Jaunpur, India"classification model of prediction for placement of students" I.J.Modern Education and Computer Science, 2013, 11, 49-56 Published Online in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijmecs.2013.11.07

[3] Namita Puri , Deepali Khot , Pratiksha Shinde , Kishori Bhoite , Prof. Deepali Maste,Department of Computer Engineering, Atharva College of Engineering, Malad,Maharashtra,India "Student Placement Prediction Using ID3 Algorithm" International Journal for Research in Applied Science & Engineering Technology Volume 3 Issue III, March 2015

[4] Amanpreet Singh, Narina Thakur, Aakanksha Sharma "A Review of Supervised Machine Learning Algorithms" IEEE 978-9-3805-4421-2/16/$31.00c 2016

[5] ELSEVIER "A comparison of machine learning techniques for customer churn prediction" T. Vafeiadis a, K.I. Diamantaras b, G. Sarigiannidis a, K.Ch. Chatzisavvas International Conference on Communication Technology and System Design 2011

[6] H.B. Kazemian , S. Ahmed ELSEVIER "Comparisons of machine learning techniques for detecting malicious a Webpages"

[7] Elaf Abu Amrieh Thair Hamtini " Mining Educational Data to Predict Student's academic Performance using Ensemble Methods" International Journal of Database Theory and Application Vol.9, No.8 (2016), pp.119-136 ,2016

[8] Seema Sharma1, Jitendra Agrawal2, Shikha Agarwal3, Sanjeev Sharma4 "Machine Learning Techniques for Data Mining: A Survey" IEEE International Conference on Computational Intelligence and Computing Research, 978-1-4799-1597-2/13/$31.00 ©2013

[9] Ajay Shiv Sharma1, Swaraj Prince2, Shubham Kapoor3 Keshav Kumar4 "PPS - Placement Prediction System using Logistic Regression" IEEE 978-1-4799-6876-3/14/$31.00c 2014

[10] Madan Somvanshi, Shital Tambade, Pranjali Chavan, S.V. Shinde "A Review of Machine Learning Techniques using Decision Tree and Support Vector Machine" IEEE International Conference on Computational Intelligence and Computing Research 2016

[11] Rakesh Kumar Arora Dr. Dharmendra Badal " Placement Prediction through Data Mining" International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 7, July 2014

[12] Pedro Strecht Luís Cruz Carlos Soares João Mendes-Moreira Rui Abreu "A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance" " IEEE 978-1-4799-6876-3/14/$31.00c 2015

[13] J. Bucko*, L. Kakalejčik "Machine Learning Techniques in the Education Process of Students of Economics" MIPRO 2017, May 22-26, 2017, Opatija, Croatia
[14] M. Mayilvaganan D. Kalpanadevi "Comparison of Classification Techniques for predicting the performance of Students Academic Environment" International Conference on Communication and Network Technologies (ICCNT) 2014