

Deep Learning Architecture based on Behavioral and Academic Perspectives for Classifying Students in a University

Srivats S. Ramanujam, J. Naren, Nivedha Jayaseelan,
Vijayalakshmi and Dr.G. Vithya

Abstract--- *Research on the academic records of students has become increasingly popular. Recent research results advocates' prospects that a student's lifestyle and personality can have a major impact on his/her ethos. It has become increasingly common for students studying in universities to not fully realize the potential and prospects in a professional sphere. Technology can aid students in realizing their caliber and prosper in the professional domain. In the proposed work, a system has been suggested which examines the feasibilities of the impact that personality features can have on a student's career in addition to academic performance.*

Keywords--- *Deep Learning, Students Academic Performance.*

I. INTRODUCTION

Education is a basic right that every citizen of a country deserves. The guidance that students get to make right choices is vital for shaping careers. Many University students often fail to realize their caliber and are thus confused about their employability. The above factor has a negative impact on student's career decisions which makes every individual end up in jobs that are unfit, and regret for the decisions taken later in life. Many companies in the industry judge students solely based on the academic performance and fail to take the capabilities of individuals into consideration. Technologies like Big Data and Deep Learning has made a huge contribution in solving such problems. Applications of Machine Learning, Deep learning in particular can aid students by giving valid career plans. Unlike Machine Learning, Deep Learning has several layers of algorithms, each specifying a distinct analysis before coming to a conclusion. Various career choices can be generated by feeding student data into Deep Learning classifier that might aid students in building a better future for them.

II. RELATED WORK

Academic Performance Prediction

MuslihahWooket.al [1] used personality traits in addition to academic data into consideration to predict academic conduct and used Artificial Neural Network techniques for doing the same. Academic performance prediction using a Multivariate Regression Prediction system, incorporating characteristics like adaptability, time

*Srivats S. Ramanujam, B. Tech Computer Science and Engineering, School of Computing, SASTRA Deemed University, Thanjavur, India.
E-mail: 120003312@sastra.ac.in*

*J. Naren, Asst Professor, School of Computing, SASTRA Deemed University, Tirumalaisamudram, Thanjavur, Tamil Nadu, India.
E-mail: naren.jeeva3@gmail.com*

*Nivedha Jayaseelan, B. Tech Computer Science and Engineering, School of Computing, SASTRA Deemed University, Thanjavur, India.
E-mail: nivijay98@gmail.com*

*Vijayalakshmi, B. Tech Computer Science and Engineering, School of Computing, SASTRA Deemed University, Thanjavur, India.
E-mail: 120003363@sastra.ac.in*

Dr.G. Vithya, Professor, School of Computing, KL University, Vijayawada. E-mail: vithyamtech@gmail.com

management, sports participation, and analytical thinking were taken into account by S. Chaitanya Kumar [2] et.al for determining academic efficiency.

Good fit or Bad fit for classification

Muhammad Fahim Uddin et.al obtained personality data from social networking sites such as Facebook, LinkedIn, used Stochastic Probability Distribution and Bayesian Networks and found whether a student has chosen a career appropriate to his personality or not. [10]

The relationship between academic prowess and student psychology

Radhika R Halde [3] et.al studied the relationship between academic prowess and student psychology and achieved the same using Decision Trees and Neural Networks.

Determination of social patterns in Students using Data mining techniques

Data Mining was used to predict behavioral patterns in students and newer methods to do the same using Machine Learning and Scientific Computing were suggested. [5].

Correlation between personality and academic performance

Academic data in combination with the behavior of students was used to determine academic prowess and was achieved using Fuzzy clustering and Multivariate regression [9]

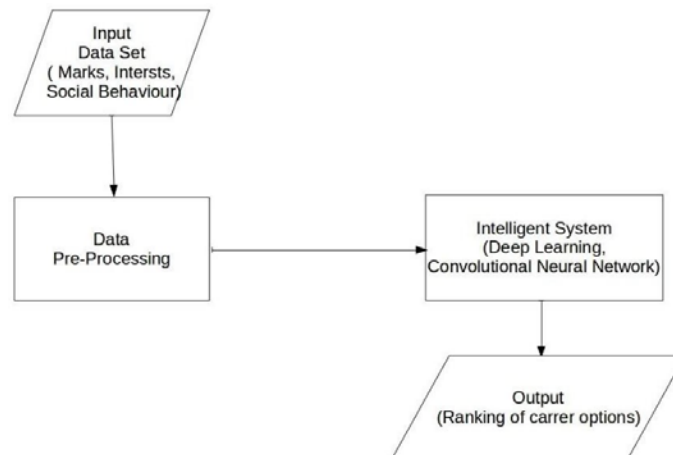


Fig. 1.1: Architecture for student classification based on behavioral and academic perspectives

The input data set which is given as to the system is pre-processed using sufficient algorithms. The system takes the processed data and classifies it using deep learning algorithms in order to generate the model. Similarly, the test data is processed and classify it. The system provides the appropriate output

III. DESCRIPTION

The inputs to the system are training examples and data which are to be classified. The data has various attributes that can be used to classify a person/student based on behavioural perspectives, academic performance such as CGPA, reaction of a student in various situations etc.

The training data has attributes that determine the behavioural and academic perspectives of a person. The Marks, interests, hobbies and other social behaviour are the features which are on a scale of 1 to 10 and can be gathered using Google Forms. The data can be collected from people who are working as researchers, corporate employee, corporate manager, start-up employee, start-up manager and teacher. The training examples contain the expected output, since the algorithm used here is supervised machine learning.

```

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances      31          100 %
Incorrectly Classified Instances    0            0 %
Kappa statistic                     1
Mean absolute error                 0.0137
Root mean squared error             0.0214
Relative absolute error             4.1505 %
Root relative squared error         5.3098 %
Total Number of Instances          31

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
          1.000   0.000   1.000     1.000   1.000     1.000  1.000    1.000    Working in a Corporate Company, Technical
          1.000   0.000   1.000     1.000   1.000     1.000  1.000    1.000    Working in a Startup, Technical Job
          1.000   0.000   1.000     1.000   1.000     1.000  1.000    1.000    Working in a Non Technical Job
          1.000   0.000   1.000     1.000   1.000     1.000  1.000    1.000    Doing Research
Weighted Avg.  1.000   0.000   1.000     1.000   1.000     1.000  1.000    1.000

=== Confusion Matrix ===

 a  b  c  d  <-- classified as
16  0  0  0 | a = Working in a Corporate Company, Technical Job
 0  7  0  0 | b = Working in a Startup, Technical Job
 0  0  5  0 | c = Working in a Non Technical Job
 0  0  0  3 | d = Doing Research
    
```

Fig 1.2: A classification algorithm used for implementation

Pre-processing of data

The original data can be pre-processed by removing the noisy data. It is an important step in data mining. In data pre-processing, the data can be cleaned and isolated. Proper pre-processing techniques must be used for efficient use for producing a better output.

IV. IMPLEMENTATION

Feature Selection

Identification of related features from the dataset is made by removing irrelevant features which doesn't contribute much on the outcome is termed as feature selection. Many feature selection techniques are available. Some of the techniques are: Filter Methods and Wrapper Methods. Filter method is used to estimate the features, cost efficient but sometimes inefficient to find the right set of features. Wrapper method is very expensive but provides more accurate features. Support Vector machine is a Wrapper method.

Tan Metal [4] made several enhancements to the Support Vector Machine method including the idea of a Feature Generating Machine to deal with ultra-high dimensional feature selection techniques which includes complex group structures and optimization schemes to reduce the bias in feature selection. Deep Learning is basically a subset of Machine Learning that constitutes many layers, and each of the layers provides a disparate analysis of the data that is fed.

A characteristic extractor is built manually to get required information from the input in a traditional model for pattern detection. It does so by discarding the extraneous features. Features extracted are later separated into disparate categories in order to form multiple layers of fully connected networks. Convolution Neural Network is preferred as it reduces the need for explicitly performing feature reduction. The input encounters many layers of feature extractors at the beginning and back propagation algorithm is used. First few prime features are extracted by neurons in the network and are subsequently connected with higher layers. Testing data can be used after the training data is passed into the Convolutional Neural Network. Different areas of work in the field are ranked by linking training data with the CNN model. The percentage of matching data is subsequently found.

Output

Disparate functioning areas in the field are listed in rank order that is best suited for the given input is given as the output. Corporate Employee, Start-up Manager, Corporate Manager, Start-up Employee, Researcher, Teacher, is the different fields of work taken into consideration. Data, which is obtained from Google Forms, is passed through a Deep Learning Model and is subsequently classified.

V. CONCLUSION AND FUTURE WORK

The given system can be extremely effective in guiding students to make great career decisions. [5].The proposed system can be effectively used by companies as a standard during the hiring process and would also pave the way for companies in hiring the most employable students from the entire lot. The proposed system could be further extended by helping students in finding the company that is best suited for them.

REFERENCES

- [1] Wook, Muslihah, Yuhanim Hani Yahaya, Norshahriah Wahab, Mohd Rizal Mohd Isa, Nor Fatimah Awang, and Hoo Yann Seong "Predicting NDUM Student's Academic Performance Using Data Mining Techniques." Second International Conference on Computer and Electrical Engineering, 2009. ICCEE'09, vol. 2, pp. 357-361. IEEE, 2009.
- [2] Kumar, S. Chaitanya, E. Deepak Chowdary, S. Venkatrama phani kumar, and KV Krishna Kishore. "M5P model tree in predicting student performance: A case study." IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 1103-1107. IEEE, 2016.
- [3] Halde, Radhika R., Arti Deshpande, and Anjali Mahajan. "Psychology assisted prediction of academic performance using machine learning." IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT), pp. 431-435. IEEE, 2016.
- [4] Tan M, Tsang IW, Wang L. "Towards ultrahigh dimensional feature selection for big data", J Mach Learn Res. 2014; 15:1371-1429.
- [5] Elakia, Gayathri, and Naren J, Aarthi. "Application of data mining in educational database for predicting behavioural patterns of the students." (IJCSIT) International Journal of Computer Science and Information Technologies 5, no. 3 (2014): 4649-4652.
- [6] Dutt, Ashish, Maizatul Akmar Ismail, and Tutut Herawan."A Systematic Review on Educational Data Mining."

- IEEE Access (2017) vol.5, pp: 15991 – 16005.
- [7] Hidayah, Indriana, Adhistya Erna Permanasari, and Ning Ratwastuti. "Student classification for academic performance prediction using Neuro-fuzzy in a conventional classroom", 2013 International Conference on Information Technology and Electrical Engineering (ICITEE) (2013) pp.1-5.
 - [8] Shanmugarajeshwari,V., and R. Lawrance."Analysis of students' performance evaluation using classification techniques." International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), pp. 1-7. IEEE, 2016.
 - [9] ZhenpengLi, Changjing Shang and Qiang Shen, (2016),"Fuzzy- clustering embedded regression for predicting student academic performance", IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp.1-8.
 - [10] Muhammad Fahim Uddin, Jeongkyu Lee, (2016),"Utilizing Relevant Academic and Personality Features from Big Unstructured Data to Identify Good and Bad Fit Students", Procedia Computer Science, Volume 95, 2016, Pages 383-391.
 - [11] Schmidhuber, Jorgen."Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.