Undergraduate engineering students employment prediction using hybrid approach in machine learning

Vinutha Krishnaiah, Yogisha Hullukere Kadegowda
[1]Department of Information Science and Engineering, BMS Institute of Technology and Management, Bangalore, India
[2]Department of Information Science and Engineering, Ramaiah Institute of Technology (Affiliated to Visvesvaraya Technological University), Bangalore, India

ABSTRACT

The knowledge discovery from student's data can be very useful in predicting the employment under different categories. The machine learning is helping in this regard up to the great extent. In this paper, a hybrid model of machine learning has proposed to predict the jobs categories, students may get in their campus placement. The considered groups of students are from undergraduate courses from engineering stream having the semester's scheme in their academic. The mapping of jobs has predicted based on their previous seven semesters marks as well as their personality index. The proposed hybrid model consists of three different model based on multilayer feed forward architecture, radial basis function neural network and K-means based clustering method. The proposed model provided the relative chances of available each job category with high accuracy and consistency.

Corresponding Author:

Vinutha Krishnaiah
Department of Information Science and Engineering, BMS Institute of Technology and Management
Doddaballapur Main Road, Avalahalli, Yelahanka, Bengaluru, Karnataka 560064, India
Email: vinuthak_ise2014@bmsit.in

1. INTRODUCTION

Machine learning based system of predictive analytics, can play a very important role in assisting student employability outcomes. Powered by policymakers, policy promoters, paternities, society, and currently sights higher education as the most important means to an entry-level job [1]. As an outcome, education Institutions are progressively more held responsible for employability product-making it an importance for career centers to find actual and accessible methods for student employability governance. The capability to expect which students will and will not be hired and what category of job will assign to individual can help a lot to recruiter as well as to students itself. Employability predictions can advise career advisors on accurately how much internship, non-collegiate activities, other types of activities, and most likely subjects which lead to a job offer and how all these aspects work together. Prediction of employability of a student's is happening today. However, we have just begun to use to the new common of data analytics, the game is changing. One quick evolution is machine learning: a method of artificial intelligence that takes the intuitions from data through pattern recognition to predict upcoming outcomes. Machine learning can study huge, difficult datasets while providing correct results, faster than human. The initial step in machine learning is determining the information that you what you have to learn from your data. Later, collect and clean the data so that it can be analyzed. After this the analyzed data are loaded to a machine learning model. As soon as the machine learning model identifies patterns, fresh data can be given to it to forecast an outcome that is presently unknown. The machine learning capability of prediction has now become part of our day to day lives. The technology is used as online advertisements recommenders, controls driverless cars,

detects banking scam, and predicts medical diagnoses. It is likely that machine learning will become typical in higher education, but the question is: "Can we accurately predict which students will be employed at graduation?" Resume audit study simply has a low capability to predict employability of student before graduation; mainly, students having higher academic performance and good skills are reflected as more employable. Machine learning provides an alternate and more powerful method of analysis when compared to static resume audits. Predicting results with career center data is still in its early stages. The research is composite and chaotic because of huge quantity of data which requires analysis and the newness of applying higher education data to the models. However, there are huge opportunities to predict results in employment and outside, including the factors like time to graduation, student satisfaction, retention and more. The potentials are boundless and holding machine learning within institution career centers is a normal starting point.

In this paper students doing the undergraduate course in the different branch of engineering college have considered for their employability in the three different categories of jobs. The past seven semesters percentage marks and the parameter capturing the personality point have considered as the inputs for the predictive models. To make the prediction more robust, rather than having a single predictive model in the solution domain, two different model based on neural network multilayer perception feed forward architecture and radial basis function neural network have considered. K-means based clustering approach has also included making the solution free from mapping approximation.

In section 2, related work in the area of use of machine learning towards employment has discussed. Section 3 contains the detail of proposed approach while section 4 carries the experimental results and analysis. Conclusion has presented at the end.

Academic inspiration provided to students plays very important role in their academic performance. For any of the educational institutes, it is very important to identify the students with less academic performance at the earliest and motivate them to improve their performances. Development of classification model for predicting the student academic motivation built on their performance in learning management system (LMS) courses, aims to establish a link between predicted student academic motivation and their performance in the LMS course [1]. Cai [2] has suggested the global higher education boards on how to motivate and produce graduates who meets the requirements and expectations of employers.

Schumacher et al. [3] had investigated a study which is designed to forecast the students who successfully completed the enrolled actuarial program using logistic regression to find the probability of dropouts or successfully completing actuarial program. The relation of social hierarchy to individual's career choice self-efficacy is discovered in [4] using an ability of social hierarchy. Determining the elements of job switch and guessing the job switching time of an individual are the most inevitable methods for learning the careers of any individuals. But still it is very difficult to predict the job switching time due to advancement of workforce division and universalization which made career more dynamic.

In [5], the association between day to day and professional life based on the experience of workers and their check in records is presented. Effort has given to expose to what level the job switch can be predicted using day to day activities of workers and their job agility. Paparrizos and Cambazoglu [6] presented query recommender system using supervised machine learning techniques which recommends new suitable job to the job seekers. Wang et al. [7] also suggested an enquiry recommender model which assists the users to find the data they require. With the vast use of information technology in ambient assisted living system, the data management model used the 3 V features of big data like volume, velocity and variety to find a comprehensive data generalization and to construct more universal systems.

Mao et al. [8] it has proposed the method of big data abstraction, with metric space as a universal abstraction for ambient assisted living data types. Cognitive modelling can determine the hidden features of examinees for forecasting their scores on each problem. Cognitive modelling plays a very significant role in various applications, for example customized remedy recommendation; some of them are also designed for providing solutions for literature. Still extraction data from both specific and objective problems to obtain accurate and understandable cognitive analysis is underexplored. Wu et al. [9] has proposed a fuzzy cognitive diagnosis framework (fuzzy CDF) for victim cognitive modelling using both specific and objective problems.

Thakar and Mehta [10] presented an experimental study that relates various classification algorithms on two datasets of masters in computer application learners composed from different affiliated colleges of state universities in India. Shahiria et al. [11], [12] has provided a summary on the data mining techniques which are used for prediction of student performance. It also provided information regarding how the prediction algorithms can be used to find the most significant features in student data. Nie et al. [13] proposed a data-madriven structure for predicting student job option upon completion of graduation based on their cognitive skills, soft skills that they exhibit in campus and thereby playing an vital role in assisting career guidance and counselling. Al-Sudani and Palaniappan [14] has used a mixture of institutional,

emotional, educational, demographic and commercial features in predicting the performance of students using multi layered neural network which classified students 'degree into either basic or good degree. In [15] external features that affects the students' performance is studied and classification model is developed using XGBoost classifier. The conservative numerical evaluations are limited in giving good predictions of educational quality of Institutions. Lau *et al.* [16] has presented a method with both conservative numerical analysis and neural network modelling for prediction of students performance. It [17], [18] tried to forecast the placements and results of programming competitions through machine learning without complicated evaluation. Descriptive variables used for machine learning are classified into different categories.

2. PROPOSED WORK

Neural network is a supervised learning method which is often used because of the curve fitting method. The network is provided with training pairs which consists of a vector from subordinate input house in addition with the preferred network reaction. From a defined learning formula, the network accompanies the variations of its weights so error between the actual and anticipated reaction is condensed compared to some optimization conditions. Once trained, the network performs the interruption within the output vector house. A nonlinear mapping between the input and also the output vector areas are often achieved with multi-layer perceptron neural network (MLPNN) and radial basis function neural network [19], [20]. This both network have been considered as universal approximation.

2.1. Multi-layer perceptron neural network (MLPNN)

For the Back Propagation the weight update equation is the amount of the updates which is proportional to the derivative of the nonlinear activation functions. An activation function for neurons is of sigmoid type with bell-shaped derivatives in MLPNN. Explicitly, the weights must be selected so as to lessen the performance criterion

$$E_{q} = \frac{1}{2}\left(d_q - x_{out}^{(s)}\right)^T\left(d_q - x_{out}^s\right) = \frac{1}{2}\sum_{h=1}^{n}\left(d_{qh} - x_{out,h}^s\right)^2 \tag{1}$$

where s means the number of layers present in the network and $d_q \in \Re^{n \times 1}$ and x_{out}^s are the derived and actual output, respectively of the network due to q^{th} training pattern. Taken an activation function of the sigmoid $type(g)$ the learning algorithm for MLP neural network having the three layers can be presented as followed.
a. Adjust the weights in the network as per the typical initialization process.
b. From set, the set of training data, develop the network response.
c. Match the chosen network output with the fixed output of the network and the local error is calculated according to

$$\text{For output layer: } \delta_i^s = \left(d_q - x_{out,i}^s\right)g(u_i^s) \tag{2}$$

$$\text{For hidden layer: } \delta_i^s = \sum_{h=1}^{n2} \delta_h^{s+1} w_{hi}^{s+1} \, g(u_i^s) \tag{3}$$

d. The weights of the network can be updated as

$$w_{ij}^s(t+1) = w_{ij}^s(t) + \mu\delta_i^s x_{out,j}^s + \alpha\left[w_{ij}^s(t) - w_{ij}^s(t-1)\right] \tag{4}$$

e. If network converged then Stop the iteration, else go back to step 2.

2.2. Radial basis function neural network (RBFNN)

The radial basis function neural network (RBFNN) design consists of 3 layers: an input layer, one layer of hidden layer which is nonlinear process neurons and also the output layer. The RBFNN output is calculated consistent with (5):

$$y_i = f_i(x) \sum_{k=1}^N W_{ik} \, \phi_k(x, c_k) = \sum_{k=1}^N W_{ik} \, \phi_k(\|x - c_k\|_2), \forall \, i = 1, 2, \dots m \tag{5}$$

where $x \epsilon \Re^{nx1}$ is an input vector, $\phi_k(.)$ is a function from $\Re^+ \, to \, \Re, \|.\|_2$ denotes the Euclidean norm, W_{ik} are the weights in the output layer. N is the number of neurons in the hidden layer, and $c_k \epsilon \Re^{nx1}$ is the RBF centers in the output space. The Euclidean distance for each neuron in the hidden layer is calculated between its related centers and the input to the network. Nonlinear function of the distance is the output of hidden

layer. At last, the weighted sum of the hidden layer outputs is calculated as the output of the network [21]. The functional form of $\phi_k(.)$ is assumed to be given and is mostly Gaussian function as given by (6),

$$\phi(x) = exp(-x^2/\sigma^2)$$

(6)

where σ parameter provides the "width" of radial basis function and is normally stated as spread parameter. The centers are the identified points which are expected to accomplish a suitable test group of the input vector space. They are typically selected as a subset of the input data [22], [23]. In Gaussian radial basis function, the spread parameter σ is commonly set according to the below heuristic relationship.

$$\sigma = \frac{d_{max}}{\sqrt{k}}$$

(7)

where d_{max} is the maximum Euclidean distance between the identified centers and K is the number of centers. Using (8) gives RBF of a neuron in the hidden layer of the network

$$\phi(x, c_k) = exp\left(-\frac{k}{d_{max}^2}\|x - c_k\|^2\right)$$

(8)

error value operate as;

$$j(n) = \frac{1}{2}|e(n)|^2 = \frac{1}{2}[y_d(n) - \sum_{k-1}^N w_k(n)\phi\{x(n), c_k(n)\}]$$

(9)

when the chosen RBF is Gaussian, (9) becomes;

$$J(n) = \frac{1}{2}\left[y_d(n) - \sum_{k-1}^N w_k(n)exp\left(-\frac{\|x(n) - c_k(n)\|_2^2}{\sigma_k^2(n)}\right)\right]^2$$

The equations for updating the network parameters are given by (10);

$$w(n + 1) = \left[w(n) - \mu_w \frac{\partial}{\partial w}J(n) >_{w=w(n)}\right] = [w(n) + \mu_w e(n)\psi(n)]$$

(10)

where $e(n) = \widehat{y(n)} - y_d(n), y_d(n)$, $y_d(n)$ is the preferred network output, and μ_w is appropriate learning parameters.

For both, MLPNN and RBFNN, initially, with the predefined parameters values, through the modeling equations quite lot sets of dissimilar kinds of patterns have created. Pre-processing has done with respect to bring down the rate of patterns around the range of [0 1]. The variety of pattern set designed, where every set brings the different number of patterns. An appropriate architecture is used to study the knowledge of pattern classification by assuming the objective as the maximum value equal to one for conforming pattern class and minimum value equal to zero for pattern classes [24], [25]. The mean square error is evaluated for the error function to select the quality of learning. To change the values of weights Gradient method is used.

2.3. K-means

K means algorithm is one of the unsupervised machine learning techniques and a popular clustering algorithm which is used to cluster data points. It selects the random points which are the center of the clusters denoted as K called centroids, where value of K is predetermined. Later compute the distance of each data points with the identified centroids are and cluster them accordingly. Repeat the process of recalculating the centroids and find distance between the data points and updated centroids and readjust the clusters until none of the data points change the clusters. The algorithm:

1. Initialize Cluster Centroids $\mu_1, \mu_2, \mu_3, \ldots \ldots \mu_k \in \mathbb{R}^n$ Arbitrarily
2. Reiterate the process up to convergence: {
   For each i, fix
   $$c^{(i)} = \arg\min_j\|x^{(i)} - \mu_j\|^2$$
   For every j, fix
   $$\mu_j = \frac{\sum_{i=1}^m\{c^{(i)}=j\}x^{(i)}}{\sum_{i=1}^m\{c^{(i)}=j\}}$$

3.   RESEARCH METHOD

The function block diagram of the proposed work which consists of three different algorithms has shown in Figure 1 First from the complete data set, three different data set has formed and a normalized version (TDS1 and TDS2) has applied to MLPNN and RBFNN architecture correspondingly while raw data has applied for K-means algorithm. The hidden layer and output layer weights (wh and wo) are being updated through the gradient decent algorithm by minimizing the mean square error value. The output layer weight (wo) of RBFNN have been upgraded from a separated learning process which is also based on gradient decent method. Once terminating criteria has been satisfied, the weight values for both algorithms have been stored. The K-means algorithm has applied to obtain the three different centroids corresponding to each category of job. Once the learning process are over for all the three algorithms, in the test phase, the stored values of weights and centroids values have been utilized to predict the outcomes from the individual process. The outcome of each process has combined to form the final prediction for individual test cases as shown in Figure 2.
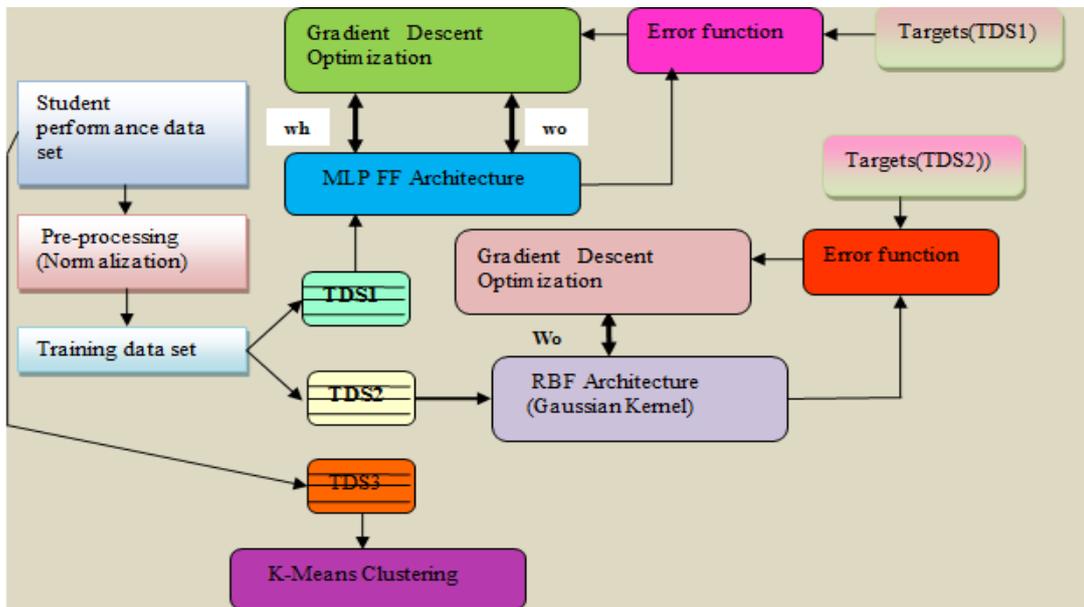

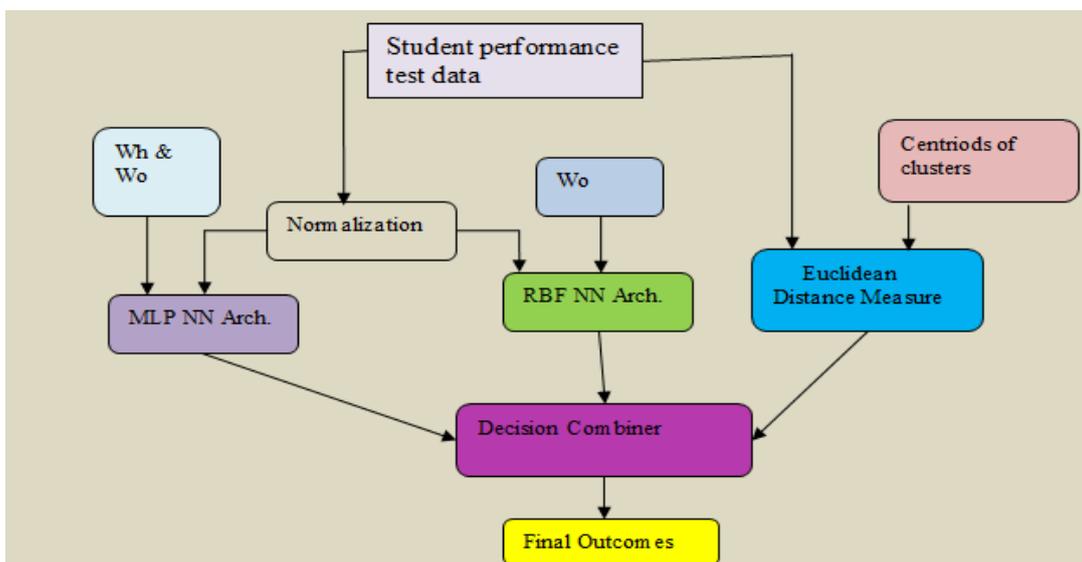Figure 1. Function block diagram of hybrid model in learning phase

Figure 2. Function block diagram of hybrid model in test phase

Undergraduate engineering students employment prediction using hybrid approach ... (Vinutha Krishnaiah)

4. RESULTS AND DISCUSSION

There was total 300 graduate student record of an engineering college who has finished their 1st to 7th semester courses and their marks have been considered to predict the possibilities of their employment in different categories of jobs. There is one more parameter has included here and called as personality point (PP) which encompass several qualities of an individual personality like presentation skill, reaction skill, team work skill and other extra-curricular activities. The score of PP has considered in the range of 1 to 10. The features of PP are not very key parameters to make someone very suitable for the core jobs of their engineering discipline but surely a high value can help to be a better employee so chance to get a better job increased. Here three different categories of jobs have considered like Type I (very good category job: >8L/annum), Type II (good category job: 5 to 8 L/annum) and Type III (average category: <5L/annum). The marks were taken from student's record while PP value has been assigned by the recruiters and the employment towards jobs has been taken from the history of recruitment over 2 year of time span by different companies visited over the campus. Sample of data set has shown in Table 1. Through uniform random process, 150 student's data have considered as training purposes while left remaining 150 students data have considered for the test cases. With this process, three different data set has created for training and test purpose and utilized by the three algorithms as shown in Figure 3 to predict the possibilities of category of employment.

Table1. Sample data from data set

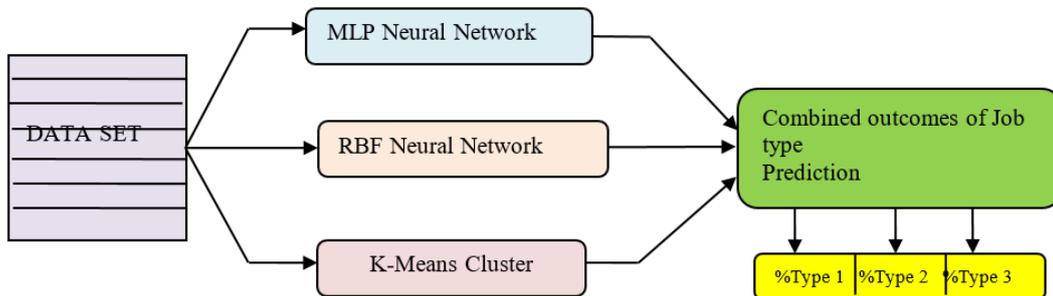| Stu	Sm1	Sm2	Sm3	Sm4	Sm5	Sm6	Sm7	PP	Obtained job
1	76.1	74.3	65.2	69.1	80.2	80.1	78.4	1	Type II
2	80.3	76.2	78.1	79.4	82.2	84.1	80.3	6	Type I
3	68.3	79.8	76.1	72.6	73.2	70.0	73.0	2.0	Type II
4	73.1	76.7	77.5	78.6	79.1	75.7	75.6	6.0	Type II
5	67.8	64.7	69.5	65.4	64.5	67.8	61.3	6.0	Type III
6	63.7	61.7	68.8	69.2	67.3	62.9	61.5	4.0	Type III
7	69.9	69.9	69.9	68.6	69.8	66.4	69.6	9.0	Type II

Figure 3. Abstract form of proposed solution

From the data set of 300 students record, through random sampling process three training data set of same size (150) has been formed. The test data set has formed by including all the data left behind after defining the training data set. For the MLPNN and RBFNN, the normalized data have been applied to prevent the operation in saturation region. For K-means clustering the normal data (without normalization) has been applied to increase the inter cluster distance. There are very high possibilities that all the three data set will have the similarity as well as diversity in their contents because of random sampling process. For MLPNN and RBFNN, the optimal numbers of hidden layer nodes have decided through the estimation of performances over different possibilities of hidden nodes like 2, 5, 8 and 10. In both case the size of input layer contains 8 nodes while output layer has 3 nodes. The output node representing the outcomes of different job classes. The learning rate and momentum constant in gradient decent process for MLPNN have considered as 0.9 and 0.2 while for RBF the learning parameter has considered as 0.1. The obtained performances have shown in Figure 4. The whole process has been implemented in the MATLAB environment.

It is clear from Figures 4(a) to (d) that MLPNN has shown better performance when the numbers of nodes were lesser than 10. In fact, there was high level of consistency can also observed in the MLPNN performance in compare to RBFNN. When the nodes number was 10 in the hidden layer performance of both

neural networks were same. Hence for MLPNN the number of hidden nodes has considered as 5 while for RBFNN it is 10. It is also interesting to note that there were low mean square errors (MSE) in the case of MLPNN in compare to RBFNN. The convergence characteristics for all the case of hidden nodes have shown in Figure 4. It can observe that there was a faster convergence for RBFNN in compare to MLPNN but quality of convergence was better for MLPNN.

       For the three different training data set, the performance of MLPNN, RBFNN and K-means algorithms have obtained for 7 independent trials to understand the consistency in delivered outcomes. The performance over training and test data along with overall performances over the whole data set has estimated. To obtain the benefit of proposed hybrid model, the outcomes over all data set have estimated. The obtained results have shown in Table 2. The statistical benefit of each method has estimated through the mean and standard deviation. It can observe that the hybrid model has delivered the high percentage of right result along with minimal variation in the performances over different trials.
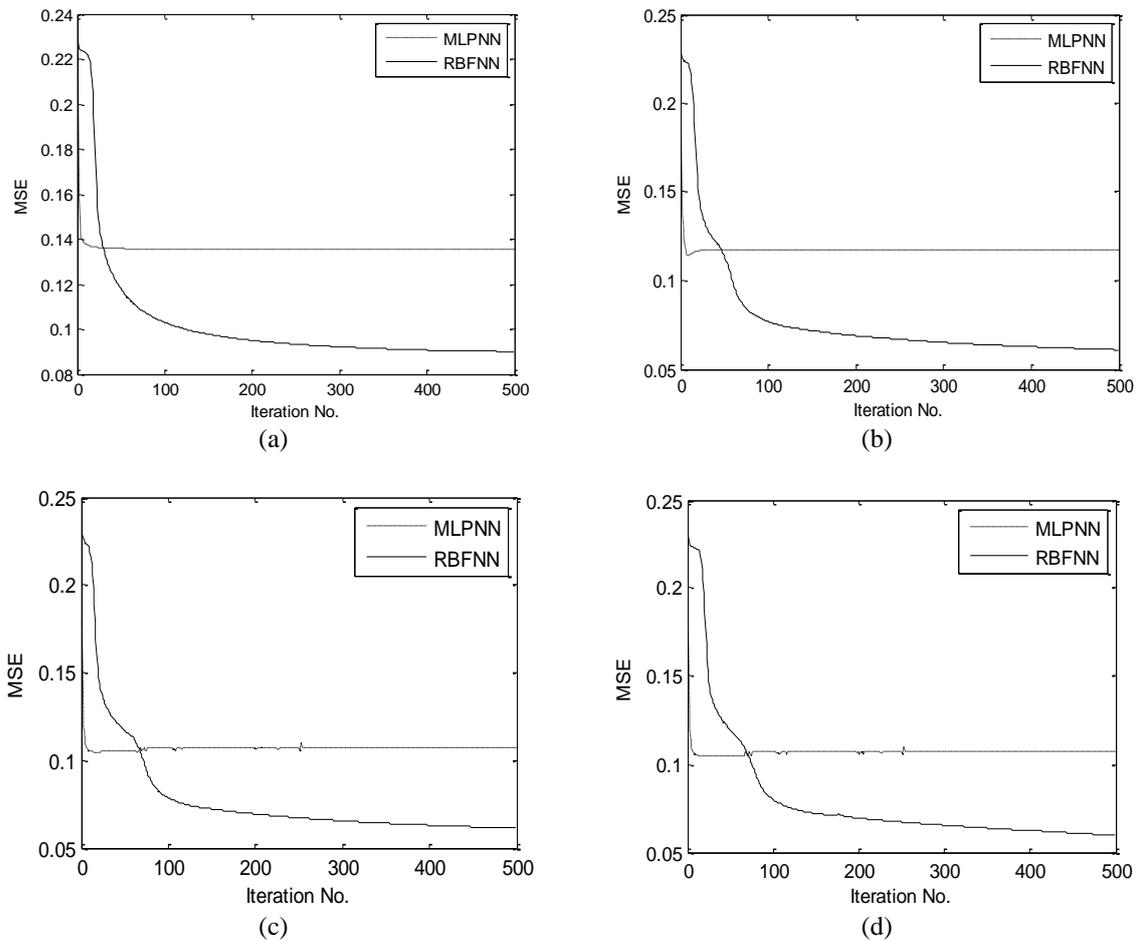

Figure 4. Convergence characteristics of MLPNN and RBFNN with (a) 2 hidden nodes, (b) 5 hidden nodes, (c) 8 hidden nodes, and (d) 10 hidden nodes

Table 2. Performance of different algorithms under 7 independent trials

Trials	MLPNN			RBFNN			K-MEANS CLUSTER			Hybrid model overall
	Tr.	Test	Overall	Tr.	Test	Overall	Tr.	Test	Overall	
1	86.67	90.67	88.67	86.67	90.67	88.67	86.67	90.67	88.67	88.67
2	86.67	90.67	88.67	77.33	76.00	76.67	86.67	90.67	88.67	88.67
3	88.67	89.33	89.00	84.00	81.33	82.67	91.33	86.67	89.00	89.00
4	88.00	88.67	88.33	85.33	88.67	87.00	88.67	88.00	88.33	88.33
5	88.67	80.00	84.33	86.67	74.67	80.67	87.33	88.67	88.00	86.67
6	88.00	84.00	86.67	85.33	92.00	88.67	47.33	50.00	48.67	88.33
7	80.67	76.00	78.33	88.00	88.00	88.00	91.33	87.33	89.33	89.00
Mean	86.76	85.62	86.28	84.76	84.47	84.62	82.76	83.14	82.95	88.38
(Std.Dev)	2.60	5.35	3.58	3.51	7.10	4.69	15.75	14.69	15.12	(0.8027)

Performance for various individual student test cases have shown in Table 3. The beauty of proposed method is it just not predicting the possibility of single job category but also declare the chance of other category jobs if possible. For example, for the student1 the predicted job category was Type II by all the three algorithms hence the chance of getting the Type II job is nearly 100% while there is also choice job available of Type III job. For student3, there is outcome of 33.33% chance of job having Type I while there is 66.33% chance of Type II job appeared because of 1 algorithm has given the favor of Type I while other 2 algorithms have given favor of Type II. Such kind of prediction makes the recruitment process optimal for the recruiters.

Table 3. Prediction for individuals carrying the details about possibilities of employment

Stu	Student performance								Chance of job type		
	Sm1	Sm2	Sm3	Sm4	Sm5	Sm6	Sm7	PP	Type I	Type II	Type III
1	76	74	65	69	80	80	78	1	×	√(100%)	√
2	80	76	78	79	82	84	80	6	√(100%)	√	√
3	68.3	79.8	76.1	72.6	73.2	70.0	73.0	2.0	√(33.33%)	√(66.33%)	√
4	73.1	76.7	77.5	78.6	79.1	75.7	75.6	6.0	√(33.33%)	√(66.33%)	√
5	67.8	64.7	69.5	65.4	64.5	67.8	61.3	6.0	×	√(33.33%)	√(66.33%)
6	63.7	61.7	68.8	69.2	67.3	62.9	61.5	4.0	×	√(33.33%)	√(66.33%)
7	69.9	69.9	69.9	68.6	69.8	66.4	69.6	9.0	×	√(66.33%)	√(33.33%)

5.   CONCLUSION

Predicting the employment of graduate students in engineering college is a very challenging task specially when there is limited number of parameters included. The challenges increases further level when there is need of high consistency and efficiency. The proposed solutions of job prediction have used the three different models in combine form to deliver the final prediction. The input-output mapping relation has been archived with MLPNN and RBFNN. It was observed that MLPNN needed lesser number of hidden nodes in compared to RBFNN for the same performance. Convergence characteristic was also better in MLPNN. The hybrid model has delivered the predicting accuracy of 88.38% while the variations in performance were limited to 0.8 % which is very appealing. The proposed model is having the simplicity in design and can be extended to use for other types of streams of academic.

REFERENCES

[1]   I. Đ. Babić, "Machine learning methods in predicting the student academic motivation," *Croatian Operational Research Review*, vol. 8, no. 2, pp. 443–461, Dec. 2017, doi: 10.17535/crorr.2017.0028.
[2]   Y. Cai, "Graduate employability: a conceptual framework for understanding employers' perceptions," *Higher Education*, vol. 65, no. 4, pp. 457–469, Apr. 2013, doi: 10.1007/s10734-012-9556-x.
[3]   P. Schumacher, A. Olinsky, J. Quinn, and R. Smith, "A comparison of logistic regression, neural networks, and classification trees predicting success of actuarial students," *Journal of Education for Business*, vol. 85, no. 5, pp. 258–263, May 2010, doi: 10.1080/08832320903449477.
[4]   M. N. Thompson and L. M. Subich, "The relation of social status to the career decision-making process," *Journal of Vocational Behavior*, vol. 69, no. 2, pp. 289–301, Oct. 2006, doi: 10.1016/j.jvb.2006.04.008.
[5]   H. Xu, Z. Yu, H. Xiong, B. Guo, and H. Zhu, "Learning career mobility and human activity patterns for job change analysis," in *2015 IEEE International Conference on Data Mining*, Nov. 2015, pp. 1057–1062, doi: 10.1109/ICDM.2015.122.
[6]   I. Paparrizos, B. B. Cambazoglu, and A. Gionis, "Machine learned job recommendation," in *Proceedings of the fifth ACM conference on Recommender systems - RecSys '11*, 2011, p. 325, doi: 10.1145/2043932.2043994.
[7]   J. Wang, J. Zhexue Huang, J. Guo, and Y. Lan, "Recommending high-utility search engine queries via a query-recommending model," *Neurocomputing*, vol. 167, pp. 195–208, Nov. 2015, doi: 10.1016/j.neucom.2015.04.076.
[8]   R. Mao, H. Xu, W. Wu, J. Li, Y. Li, and M. Lu, "Overcoming the challenge of variety: big data abstraction, the next evolution of data management for AAL communication systems," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 42–47, Jan. 2015, doi: 10.1109/MCOM.2015.7010514.
[9]   R. Wu *et al.*, "Cognitive modelling for predicting examinee performance," in *IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence*, 2015, pp. 1017–1024.
[10] P. Thakar, P. Dr., and D. Manisha, "Role of secondary attributes to boost the prediction accuracy of students' employability via data mining," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 11, 2015, doi: 10.14569/IJACSA.2015.061112.
[11] A. M. Shahiri, W. Husain, and N. A. Rashid, "A review on predicting student's performance using data mining techniques," *Procedia Computer Science*, vol. 72, pp. 414–422, 2015, doi: 10.1016/j.procs.2015.12.157.
[12] R. Ade and P. R. Deshmukh, "Efficient knowledge transformation system using pair of classifiers for prediction of students career choice," *Procedia Computer Science*, vol. 46, pp. 176–183, 2015, doi: 10.1016/j.procs.2015.02.009.
[13] M. Nie *et al.*, "Advanced forecasting of career choices for college students based on campus big data," *Frontiers of Computer Science*, vol. 12, no. 3, pp. 494–503, Jun. 2018, doi: 10.1007/s11704-017-6498-6.
[14] S. Al-Sudani and R. Palaniappan, "Predicting students' final degree classification using an extended profile," *Education and Information Technologies*, vol. 24, no. 4, pp. 2357–2369, Jul. 2019, doi: 10.1007/s10639-019-09873-8.
[15] B. M. Rao and B. V. R. Murthy, "Prediction of student's educational performance using machine learning techniques," in *Data*

Engineering and Communication Technology. Advances in Intelligent Systems and Computing, 2020, pp. 429–440.

[16] E. T. Lau, L. Sun, and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," *SN Applied Sciences*, vol. 1, no. 9, p. 982, Sep. 2019, doi: 10.1007/s42452-019-0884-7.

[17] R. Ishizue, K. Sakamoto, H. Washizaki, and Y. Fukazawa, "Student placement and skill ranking predictors for programming classes using class attitude, psychological scales, and code metrics," *Research and Practice in Technology Enhanced Learning*, vol. 13, no. 1, p. 7, Dec. 2018, doi: 10.1186/s41039-018-0075-y.

[18] V. K and Yogisha H K, "Employability prediction of engineering graduates using machine learning algorithms," *International Journal of Recent Technology and Engineering*, vol. 8, no. 5, pp. 4521–4524, Jan. 2020, doi: 10.35940/ijrte.E6823.018520.

[19] B. Hu, "Teaching quality evaluation research based on neural network for University physical education," in *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, May 2017, pp. 290–293, doi: 10.1109/ICSGEA.2017.155.

[20] U. Kose and A. Arslan, "Optimization of self-learning in computer engineering courses: an intelligent software system supported by artificial neural network and vortex optimization algorithm," *Computer Applications in Engineering Education*, vol. 25, no. 1, pp. 142–156, Jan. 2017, doi: 10.1002/cae.21787.

[21] E. A. Villaseñor, R. Arencibia-Jorge, and H. Carrillo-Calvet, "Multiparametric characterization of scientometric performance profiles assisted by neural networks: a study of Mexican higher education institutions," *Scientometrics*, vol. 110, no. 1, pp. 77–104, Jan. 2017, doi: 10.1007/s11192-016-2166-0.

[22] H. M. H. Mustafa, A. Al-Hamadi, S. A. Al-Ghamdi, M. M. Hassan, and A. A. Khedr, "On assessment of students' academic achievement considering categorized individual differences at engineering education (Neural Networks Approach)," in *2013 12th International Conference on Information Technology Based Higher Education and Training (ITHET)*, Oct. 2013, pp. 1–10, doi: 10.1109/ITHET.2013.6671003.

[23] Qi-Jun Zhang, K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for rf and microwave design-from theory to practice," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, no. 4, pp. 1339–1350, Apr. 2003, doi: 10.1109/TMTT.2003.809179.

[24] R. Bajaj and V. Sharma, "Smart Education with artificial intelligence based determination of learning styles," *Procedia Computer Science*, vol. 132, pp. 834–842, 2018, doi: 10.1016/j.procs.2018.05.095.

[25] A. A. Kardan, H. Sadeghi, S. S. Ghidary, and M. R. F. Sani, "Prediction of student course selection in online higher education institutes using neural network," *Computers {\&} Education*, vol. 65, pp. 1–11, Jul. 2013, doi: 10.1016/j.compedu.2013.01.015.

BIOGRAPHIES OF AUTHORS

Vinutha Krishnaiah ⓘ 🅶 sc Ⓟ is working as Assistant professor in Dept of ISE, BMSIT&M. She has completed her B.E in Channabasaveshwara Institute of Technology Tumkur in 2011 and M.Tech from AMC Engineering College 2013 in CSE. She is having 7 years of teaching experience and pursuing research in the field of prediction models using Machine learning. Her area of interest is Machine learning and Data mining. She is a life member of "Indian Science Congress Association". She is a Research Scholar at VTU Belgaum. She can be contacted at email: Vinutha.k90@gmail.com.

Yogisha Hullukere Kadegowda ⓘ 🅶 sc Ⓟ is working as Professor and Head of the Department of Information Science and Engineering. He completed his B.E in Computer Science and Engineering from PESCE, Mandya, Mysore University, and M.Tech in Computer Engineering from SJCE Mysore, VTU Belagavi and Ph.D. in Computer Science at Bharathiar University, Coimbatore. He has published and presented 57 papers at National/International Conferences/Journals among ELEVEN papers were indexed by SCOPUS, EIGHT IEEE, TWO papers in DBLP. He has authored three text-books He is a life time member of various Professional bodies like FIE, LMISTE, CSI and IAENG. He has also worked as Corporate Trainer for TWO years, trained almost 800 Tech Mahindra employees at CMC Ltd and at Uttara Info Solutions. He can be contacted at email: yogishhk@gmail.com.