**IOSR Journal of Engineering (IOSRJEN)** **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 19-23

# An ANN Based EDM Approach to Explore the Reasons for Scarcity of Expected Skilled Graduates in Employability

Dr. M. Subha
*Assistant Professor, Department of Computer Science (PG), K.S.Rangasamy College of Arts and Science (Autonomous), Tiruchengode – 637 215, Namakkal District, Tamil Nadu, India.*
*Mobile No.: 9043450255, E-Mail ID.: Subhasangeeth2015@gmail.com*

***Abstract:*** *Every year vacancies are tremendously increasing in both government and private sector. The number of fresh graduates is also increasing. Un-employment is also increasing every year. Latest research in India, percentage of un-employability increased. But there is scarcity in the recruitment of qualified and skilled graduates. Why un-employability after the immediate graduation? The main objective of this paper is to give the concepts about the reasons of unemployability by using soft computing techniques. Also construct the student model to make them qualitative for recruitment. It is essential to analyze the reason of increasing percentage of un-employability. The academic achievement of higher education is a deciding factor in the life of any student. Higher learning specializations that are offered by universities and higher technological educational institutes are best. This approach is useful for the educators to evaluate and improve the structure of their course context to meet the demand. Educational data mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational context. Based on the past 5 years graduates dataset and the un-employability dataset from few private and public sector, the reasons, expectations and demands of employability can be predicted by using the ANN based EDM. Hence, the best training algorithm for constructing an accurate prediction student model will constructed, that will predict student characteristics or performances in educational institutions.*

***Keywords:*** *ANN, EDM, Student Model.*

## I. INTRODUCTION

### 1.1 Unemployability

New Delhi, August 30, 2018: AI and machines have made their way into almost every area of recruitment evaluation today. Himanshu Aggarwal elaborates this in his recent article featured in the Silicon India. To overcome this scenario analyse the data and improve the student efficiency to face the interview in the higher education context.

### 1.2 Educational Data Mining

Now a days, data in educational institutes' is growing rapidly and in voluminous amount. To process large volume of data available in databases of educational institutes', provided with data mining to figure out the relation between various attributes and to predict the result. Data mining when applied on educational data is called education data mining [1]. Education data mining uses techniques which are analytical tools for extracting and processing data available for data analysis. Education data mining is a powerful technique to make the data available in the data repositories of institutes' useful. By using previous data for future prediction, data mining can help a lot in raising the institutes' value. By using data mining focused on to form a system which can assist placement cell of the institute to know the student performance and corresponding to that, particular inputs can be given to the students which can enhance placements. Education data mining is used is to find out various placements related issues which are as follow:

To find out the correlation between various attributes available in mined dataset.
To find out how particular student is likely to pass a certain level of the interview process.
To find out what input or subject a student needs more attention in to get placed.

EDM application areas are:Analysis and visualization of data, Providing feedback for supporting instructors, Recommendations for students, Predicting student performance, Student modeling Detecting undesirable student behaviors, Grouping students, Social network analysis, Developing concept maps, Constructing courseware.

### 1.3 Artificial Neural Network (ANN)

Neural networks are parallel computing devices, which are basically an attempt to make a computer model of the brain. The main objective is to develop a system to perform various computational tasks faster than the traditional systems.

Artificial Neural Network (ANN) is an efficient computing system whose central theme is borrowed from the analogy of biological neural networks. ANNs are also named as "artificial neural systems," or "parallel distributed processing systems," or "connectionist systems." ANN acquires a large collection of units that are interconnected in some pattern to allow communication between the units. These units, also referred to as nodes or neurons, are simple processors which operate in parallel.

Every neuron is connected with other neuron through a connection link. Each connection link is associated with a weight that has information about the input signal. This is the most useful information for neurons to solve a particular problem because the weight usually excites or inhibits the signal that is being communicated. Each neuron has an internal state, which is called an activation signal. Output signals, which are produced after combining the input signals and activation rule, may be sent to other units.

#### 1.3.1 Model of Artificial Neural Network

The following diagram represents the general model of ANN followed by its processing.

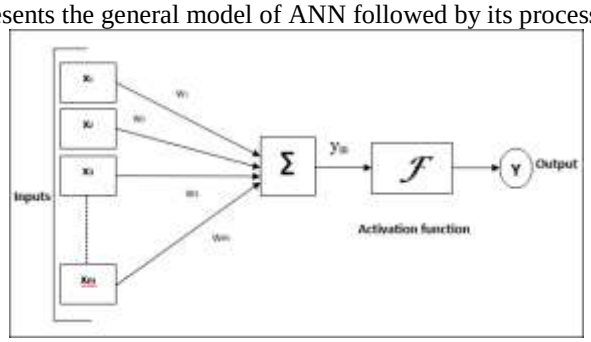



***Figure 1.*** *General Model of ANN*

For the above general model of artificial neural network, the net input can be calculated as follows −

Yin = x1.w1+x2.w2+x3.w3…xm.wm

i.e., Net input yin = ∑mixi.wi

The output can be calculated by applying the activation function over the net input. Y = F(yin)

Output = function (net input calculated)

#### 1.3.2 Machine Learning in ANNs

ANNs are capable of learning and they need to be trained. There are several learning strategies −

- **Supervised Learning** − It involves a teacher that is scholar than the ANN itself. For example, the teacher feeds some example data about which the teacher already knows the answers.
  For example, pattern recognizing. The ANN comes up with guesses while recognizing. Then the teacher provides the ANN with the answers. The network then compares it guesses with the teacher's "correct" answers and makes adjustments according to errors.
- **Unsupervised Learning** − It is required when there is no example data set with known answers. For example, searching for a hidden pattern. In this case, clustering i.e. dividing a set of elements into groups according to some unknown pattern is carried out based on the existing data sets present.
- **Reinforcement Learning** − Strategy built on observation. The ANN makes a decision by observing its environment. If the observation is negative, the network adjusts its weights to be able to make a different required decision the next time.

## II. RELATED WORK

The adoption of learning management systems in education has been increasing in the last few years. Various data mining techniques like prediction, clustering and relationship mining can be applied on educational data to study the behavior and performance of the students.

## III. EDM PROCESS AND METHODOLOGY OF STUDENT MODELING

Educational Data mining is a technique of extraction hidden predictive information from large databases; it is a powerful new technology with great potential to help higher learning Universities or institutions to focus on the most important information in their data warehouses. Data mining tools predict

future trends and behavior patterns, allowing institution to make proactive, knowledge-driven and appropriate decisions. Few EDM tools and algorithms: Machine Learning, Artificial Intelligence, Emulating human intelligence, Neural Networks for prediction, Biological models and psychological models, SLIQ (Supervised Learning in Quest).

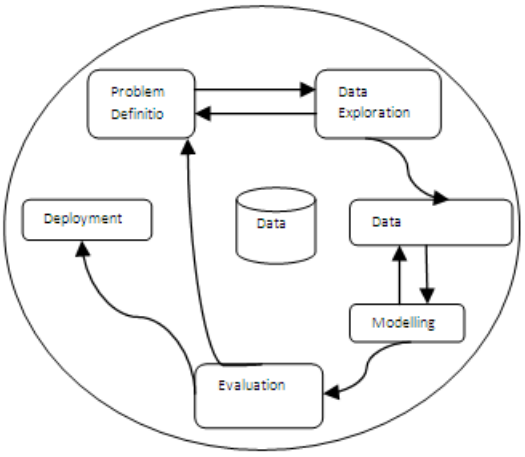



***Figure 2.*** *EDM Process*

### 3.1 Phases of EDM

EDM is an iterative process that typically involves the following phases: Problem definition, Data exploration, Data preparation, Modeling, Evaluation. Few tools of data collection & analysis are needed for project are for analyzing data, designing, implementation and some developing software tool such as: MYSQL Datbase , Excel, Weka Data Mining Tool, etc.

## IV. METHODOLOGY OF STUDENT MODELING

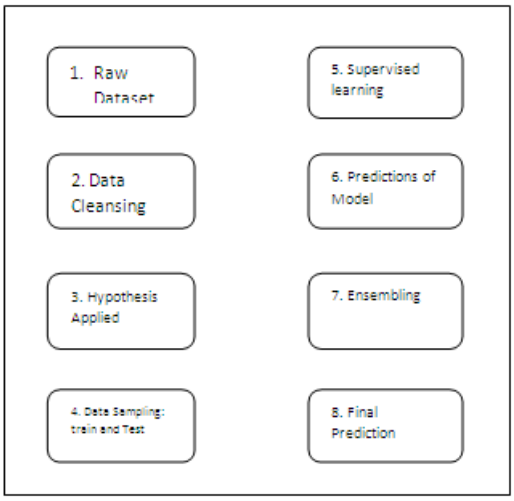



***Figure 3.*** *Methodology of Student Model*

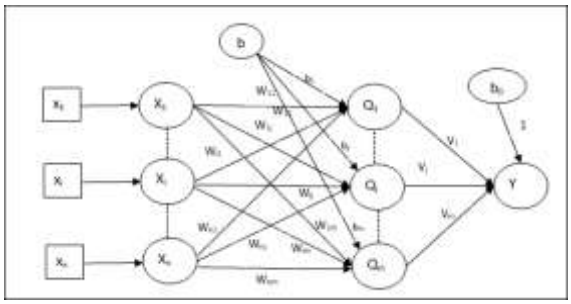
## V. TRAINING ALGORITHM FOR STUDENT MODEL

### 5.1 Multiple Adaptive Linear Neuron (Madaline)

Madaline which stands for Multiple Adaptive Linear Neuron, is a network which consists of many Adalines in parallel. It will have a single output unit. Some important points about Madaline are as follows −

- It is just like a multilayer perceptron, where Adaline will act as a hidden unit between the input and the Madaline layer.
- The weights and the bias between the input and Adaline layers, as in we see in the Adaline architecture, are adjustable.
- The Adaline and Madaline layers have fixed weights and bias of 1.
- Training can be done with the help of Delta rule.

**5.1.2 Architecture**



The architecture of Madaline consists of **"n"** neurons of the input layer, **"m"** neurons of the Adaline layer, and 1 neuron of the Madaline layer. The Adaline layer can be considered as the hidden layer as it is between the input layer and the output layer, i.e. the Madaline layer.

**5.2 Training Algorithm**

Only the weights and bias between the input and the Adaline layer are to be adjusted, and the weights and bias between the Adaline and the Madaline layer are fixed.

**Step 1** − Initialize the following to start the training −

- Weights
- Bias
- Learning rate αα

For easy calculation and simplicity, weights and bias must be set equal to 0 and the learning rate must be set equal to 1.

**Step 2** − Continue step 3-8 when the stopping condition is not true.
**Step 3** − Continue step 4-6 for every bipolar training pair **s:t**.
**Step 4** − Activate each input unit as follows − xi=si(i=1ton)
**Step 5** − Obtain the net input at each hidden layer, i.e. the Adaline layer with the following relation − Qinj=bj+∑inxiwijj=1tom
Here **'b'** is bias and **'n'** is the total number of input neurons.
**Step 6** − Apply the following activation function to obtain the final output at the Adaline and the Madaline layer −

f(x)={1−1ifx⩾0ifx<0

Output at the hidden (Adaline) unit
Qj=f(Qinj)
Final output of the network y=f(yin)
**i.e.** yinj=b0+∑mj=1Qjvj
**Step 7** − Calculate the error and adjust the weights as follows −
**Case 1** − if **y** □ **t**and **t = 1** then,

wij(new)=wij(old)+α(1−Qinj)xi
bj(new)=bj(old)+α(1−Qinj)

In this case, the weights would be updated on $\mathbf{Q_j}$ where the net input is close to 0 because **t = 1**.
**Case 2** − if **y** □ **t**and **t = -1** then,

wik(new)=wik(old)+α(−1−Qink)xi
bk(new)=bk(old)+α(−1−Qink)

In this case, the weights would be updated on $\mathbf{Q_k}$ where the net input is positive because **t = -1**.

Here **'y'** is the actual output and **'t'** is the desired/target output.

**Case 3** − if **y = t** then
There would be no change in weights.
**Step 8** − Test for the stopping condition, which will happen when there is no change in weight or the highest weight change occurred during training is smaller than the specified tolerance.

## VI. CONCLUSION

Proposed paper is a content paper for reducing the unemployability percentage and increasing the students' placement during their higher education. Using ANN based EDM is an idea to analyse the percentage of the unemployability in the current scenario. Using unsupervised algorithm in ANN a student model can be

created and applied in the education Institutions to improve the skills of student. Instead of quantitative the higher education can be improved qualitatively.

## REFERENCES


[1]. *"EducationalDataMining.org", 2013. Retrieved 2013-07-15.*
[2]. R. Baker (2010) Data Mining for Education. In McGaw, B., Peterson, P., Baker, E. (Eds.) International Encyclopedia of Education (3rd edition), vol. 7, pp. 112-118. Oxford, UK: Elsevier.
[3]. G. Siemens, R.S.j.d. Baker (2012). "Learning analytics and educational data mining: towards communication and collaboration". Proceedings of the 2nd International Conference on Learning Analytics and Knowledge: 252–254.
[4]. C. Romero, S. Ventura. Educational Data Mining: A Review of the State-of-the-Art. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 40(**6**), 601-618, 2010.
[5]. "http://educationaldatamining.org/EDM2008/" Retrieved 2013-09-04
[6]. Baker, Ryan. "Data Mining for Education"(PDF). oxford, UK: Elsevier. Retrieved 9 February 2014.
[7]. Baker, R.S.; Yacef, K (2009). "The state of educational data mining in 2009: A review and future visions". JEDM-Journal of Educational Data Mining. **1** (1): 2017.
[8]. Romero, Cristobal; Ventura, Sebastian (JAN-FEB 2013). Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. **3** (1): 12–27.
[9]. Romero, Cristobal; Ventura, Sebastian (2007). "Educational data mining: A survey from 1995 to 2005". Expert Systems with Applications. **33** (1): 135–146.
[10]. "Assessing the Economic Impact of Copyright Reform in the Area of Technology-Enhanced Learning".
[11]. Azarnoush, Bahareh, et al. "Toward a Framework for Learner Segmentation." JEDM-Journal of Educational Data Mining 5.2 (2013): 102-126.
[12]. U.S. Department of Education, Office of Educational Technology. "Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief".
[13]. Handbook of educational data mining. CRC Press. 2010.
[14]. "Big Data in Education". Retrieved 30 March 2014, 13 October 2015, 17 July 2018.
[15]. "Learning Analytics | Teachers College Columbia University". www.tc.columbia.edu.
[16]. "How Can Educational Data Mining and Learning Analytics Improve and Personalize Education?".
[17]. Educational Data Mining: A Literature Review. Available from: https://www.researchgate.net/publication/313066371_Educational_Data_Mining_A_Literature_Review [accessed Sep 28 2018].