

## A Study on Employability of Engineering Graduates Using Data Mining Techniques

Selvi.R<sup>1</sup>, Anitha.P<sup>2</sup>, Padmini.P<sup>3</sup>

Under Graduate<sup>1</sup>, Under Graduate<sup>2</sup>, Assistant Professor<sup>3</sup>

Department of Computer Science / Arunesa Arts and Science College for Women, Tiruvannamalai-606 611

**Abstract:** The constant flow of new graduates each year makes the Indian job market highly competitive. Enabling these new job seekers to acquire a decent employment is a challenge to stakeholders in education. Career counselling at early age in school is essential students are made aware of the possibilities and realities of each stream. Unemployment has effected engineering education in such as way that millions of engineering graduates are struggling to get placed in the related jobs. Choosing the right career path is becoming more and more important for young graduates today. Irrespective of the branch of study, maximum engineers opt for jobs that don't rely on the specific subjects they have studied. The present study aimed to focus the employability of engineering graduates and to determine the possibilities of enhancing employability skills deals with the adequate teaching methodologies.

**Keywords:** Data mining, Classification techniques, Engineering graduates employability, J48.

### I. Introduction

"Over the years, a large number of programmes and schemes have been formulated and implemented for improving the quality and standards of technical education in India. They include programmes related to faculty development, curriculum development, apprenticeship training, community polytechnics, development of rural technology, modernization and removal of obsolescence, institutional networking, technical man power information system, advanced technician courses, continuing education, research and development, industry-institution interaction, and so on"[Tripti Mishra et al., "Student's Employability Prediction Model through Data Mining", International Journal of Applied Engineering Research, Volume 11, Number 4 (2016).]. Further several innovative programmes have been introduced as part of the implementation of National Policy on Education-1986. But because of inadequate resources the scope and dimensions of these programmes have been reduced considerably. The emphasis on higher education in India can be understood by the number of universities currently present in India and the quality of education they provide. As of 2014, there are 677 universities, 37,204 colleges and 11443 stand-alone institutions in India, as per the latest statistics from the website of India's HRD ministry. India has achieved the production capacity of 1.5 million engineers every year with mammoth 4000 institutes..!!!. Over the last ten years, There is a almost 200 % increment in the intake and pass out of (so called) engineers. Right now, India is producing engineers more than us and China producing engineers together.

Growth of INTAKE of Technical Institutions in the Country (UNDER GRADUATE)

Year	Engg	Phar	Arch	HMCT	Total	Added in Year
2006-07	659717	76030	5085	5840	746672	30432
2007-08	701214	77582	5189	5959	789944	43272
2008-09	753910	78763	5268	6050	843991	54047
2009-10	1093380	80370	5375	6174	1185299	341308
2010-11	1219347	81594	5457	6268	1312666	127367
2011-12	1386083	83041	6894	6295	1482313	169647
2012-13	1565722	85461	8874	6355	1666412	184099
2013-14	1634596	86444	8614	6520	1736174	69762

**Table 1.** Growth of Intake of Technical Institutions in the Country

State/Union Territory	Number of Engineering Institutes
Total	4000
Andhra Pradesh	960
Tamil Nadu	934
Maharashtra	739
Uttar Pradesh	700
Haryana	342
Rajasthan	338
Madhya Pradesh	285-310
Punjab	221
Karnataka	210
Kerala	198
West Bengal	155
Orissa	125
Gujarat	120
Chhattisgarh	75
Himachal Pradesh	66
Delhi	37
Jharkhand	33
Bihar	30
Jammu & Kashmir	28
Assam	23
Puducherry	21
Goa	10
Chandigarh	9
Meghalaya	4
Uttarakhand	4
Arunachal Pradesh	3
Sikkim	3
Manipur	2
Tripura	2

**Table 2.** List of Number of Engineering Institutes

The constant flow of new graduates each year makes the Indian job market highly competitive. Enabling these new job seekers to acquire a decent employment is a challenge to stakeholders in education

## II. Related Works

Padmini(2012) report that, Education and training create assets in the form of knowledge and skills which increases everyone productive capacity of manpower and this is referred to a human capital.

MasuraRahmat et al., (2012) says that study aims to measure the employability of the FIST (Faculty of Information Science and Technology) graduates, Measurement carried out is based on the skills possessed by graduates during their studies at the faculty, skills are assessed based on their basic knowledge of programming system development, soft skills, and entrepreneur skills.

Tripti Mishra et al., suggested that early prediction of employability of Master of Computer Applications (MCA) students and Predicting student employability can help identify the students who are at risk of unemployment and management can intervene timely and take essential steps to train the students to improve their performance.

BangsukJantawan and Cheng-Fa Tsai (2014) found that comparing the accuracy of classification model under Bayesian methods for identifying the factors underlying graduate employability. This information may provide in-depth data valuable for the ministry of education to monitor and improve various aspects of an administrative system in the institutions of higher education.

MeshramS.andDeveshDubey (2015) report that the skill sets required by young graduates to enter the IT industry for their sustainability and to assess how there can be a value creation for better chances of survival in the tough global job market. The study concludes education focuses on the fundamentals, concepts in different subjects which brings out with excellent basics and strong foundation. But we should provide skills both technical and soft skills to students and facilitate their employability and play a role in empowering employable India.

## III. Methods

Methodology is a process, which reveals all those methods and techniques used by the researchers during the course of studying their research problem. The role of methodology is to carry out the research work in a scientific and valid manner. Adaptation of a suitable methodology can raise the efficiency and dignity of the research work. The success of any research mainly depends on the tools and techniques and the proper methods adopted in the research process. To raise the quality of the research work, a suitable methodology should be adopted. The data collected from different sources require proper method of extracting knowledge from large warehouses for better decision making. Data mining aims at the discovery of useful information from large collections data.

#### IV. Hypothesis Of The Study

Hypothesis testing: A procedure, based on sample evidence and probability theory, use to determine whether the hypothesis is reasonable statement and should not be rejected, or is unreasonable and should be rejected.

A research hypothesis is a predictive statement capable of being tested by scientific methods that relates an independent variable to some independent variable. The hypothesis is capable of being objectively verified and tested. Thus we may conclude that a hypothesis states what we are looking for and it is a proposition which can be put to test to determine its validity. The following hypotheses were formulated for the present study.

**Hypothesis – I:** There is a significant Association between Engineering Branch and the working specialization.

**Hypothesis – II:** There is a significant Association between Engineering Branch and Students placed in campus interview.

**Hypothesis – III:** There is a significant Association between Engineering Branch and job satisfaction of an employee.

**Hypothesis – IV:** There is a significant Association between the graduate working on the Specialization and job satisfaction.

**Hypothesis – V:** There is a significant Association between placement opportunities in campus interview and CGPA of a student.

**Hypothesis – VI:** There is a significant association between Salary received per annum and satisfaction with job profile

The data collected from different sources require proper method of extracting knowledge from large warehouses for better decision making. Data mining aims at the discovery of useful information from large collections data.

#### 4.1 Phases Of The Present Investigation

Methodology is a process, which reveals all those methods and techniques used by the researchers during the course of studying their research problem. The role of methodology is to carry out the research work in a scientific and valid manner. Adaptation of a suitable methodology can raise the efficiency and dignity of the research work. The success of any research mainly depends on the tools and techniques and the proper methods adopted in the research process. To raise the quality of the research work, a suitable methodology should be adopted. The data collected from different sources require proper method of extracting knowledge from large warehouses for better decision making. Data mining aims at the discovery of useful information from large collections data.

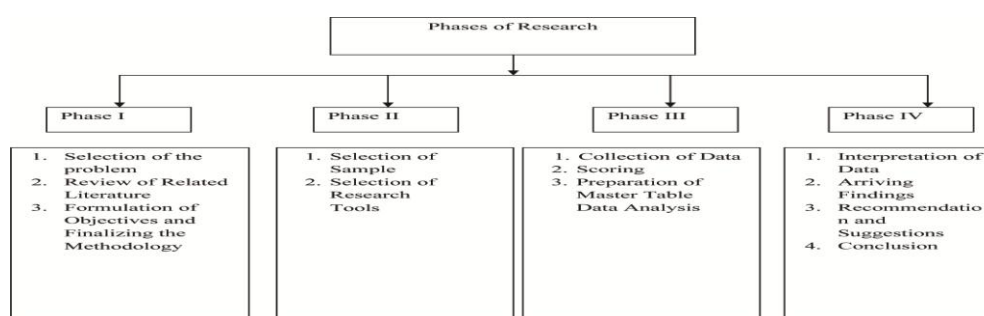


Figure 1: Phases Of Present Investigation

#### 4.2 Data Mining Techniques

##### Classification Algorithm

Now that you have loaded a dataset, it's time to choose a machine learning algorithm to model the problem

##### Naïve Bayes

This method is based on probabilistic knowledge. This method goes by the name Naïve Bayes, because it's based on Bayes's rule and "naively" assumes independence- it is only valid to multiply probabilities when the events are independent. The goal is to predict the class of the test cases with class information that is provided in the training data. Two types of Naïve Bayes algorithms are mentioned as Naïve Bayes (NB) and Simple Naïve Bayes (SNB)

## Decision Tree

The decision tree used in WEKA is termed as J48 which is a modification of the C4.5 algorithm. Classification of data and the confusion matrix will be displayed in the classifier output screen.

## J48

J48 is an extension of ID3. The additional features of J48 are accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc.

## FT

Functional Trees (FT) are classification trees that could have logistic regression functions at the inner nodes and/or leaves. The algorithm that uses FT can deal with binary and multi- class target variables, numeric and nominal attributes and missing values.

## 4.3 Data Mining Techniques

### Cluster Analysis

#### Simple k-means

The k-means algorithm is an algorithm to cluster  $n$  objects based on attributes into  $k$  partitions, where  $k < n$ . It is similar to the expectation-maximization algorithm for mixtures of Gaussians in that they both attempt to find the centers of natural clusters in the data. It assumes that the object attributes form a vector space. An algorithm for partitioning (or clustering)  $N$  data points into  $K$  disjoint subsets  $S_j$  containing data points so as to minimize the sum-of-squares criterion. Where,  $x_n$  is a vector representing the  $n$ th data point and  $u_j$  is the geometric centroid of the data points in  $S_j$ . Simply speaking k-means clustering is an algorithm to classify or to group the objects based on attributes/features into  $K$  number of group.  $K$  is positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid.

### Association Rule

In data mining, association rule learning is a popular and well researched method for discovering interesting relations between variables in large databases.

### Apriori

Association rule generation is usually split up into two separate steps:

1. First, minimum support is applied to find all frequent itemsets in a database.
2. Second, these frequent itemsets and the minimum confidence constraint are used to form rules.

While the second step is straight forward, the first step needs more attention. Finding all frequent itemsets in a database is difficult since it involves searching all possible itemsets (item combinations). The set of possible itemsets is the power set over  $I$  and has size  $2^n - 1$  (excluding the empty set which is not a valid itemset).

## V. Result And Discussion

After framing the research questions and research objectives various kinds of related articles, journals have to be go through. Graduates studied from various engineering colleges were selected for the research study and sent them a questionnaire. Then, the formulated set of hypotheses was tested with the collected data by SPSS tool using Chi square Test. Attribute selection, Classification, Association and Cluster data mining functionalities are done by using WEKA tool. A total of 300 records were taken for the analysis. The following figure shows that test dataset.

Nominal class attribute	Weka Classifiers		
	FT	J48	Naïve Bayes
Engineering Branch	100	98.67	99.67

Table 3: Weka Classifiers

This chapter summarizes the findings of the evaluation, including summarizing the test results for each of the evaluation hypotheses. Each hypothesis is identified as either “accepted,” or “rejected.” A total of 6 evaluation hypotheses were identified for testing and 5 hypotheses were accepted and 1 hypothesis is rejected. The following are the attributes and the corresponding hypothesis to verify the relationship between the attributes.

#### Testing Hypothesis

Attributes	Hypothesis
Engineering Branch and the working specialization.	Ho accepted
Engineering Branch And Students placed in campus interview.	Ho accepted
Engineering Branch And job satisfaction of an employee.	Ho rejected
Graduates Working on the Specialization and job satisfaction.	Ho accepted
Placement opportunities in campus interview and CGPA of a student .	Ho accepted
Salary received per annum and satisfaction with job profile.	Ho accepted

**Table 4:** Attributes/Hypothesis

## VI. Conclusion And Future Work

The current situation of the engineering educational system and the quality of all its individual components it can be noticed that the awareness of the importance of education as a foundation for the growth and development of the country is not strong enough. Engineering graduates passing out from educational institutions have to fulfil modern and high standard requirements that are needed by industry. Need to focus on developing employability skills in our engineering graduates. Future light has been thrown for enlarging the experiments to more colleges, in order to enrich the student databases. For future work the experiment can be extended with different tools for more algorithms and questionnaire be elaborated with increased attributes.

## References

- [1]. Padmini, “Education Vs Employability- the Need to Bridge the Skills Gap among the Engineering and Management Graduates in Andhra Pradesh”, International Journal of Management & Business studies, Vol. 2, Issue 3, July - Sept 2012.
- [2]. Tripti Mishra et al., “ Student’s Employability Prediction Model through Data Mining”, International Journal of Applied Engineering Research, Volume 11, Number 4 (2016).
- [3]. Meshram S. A and Devesh Dubey, Employability of engineering graduates in IT sector – (A case study of College of Engineering, Pune COEP), Asian Journal of Management Research, Volume 5 Issue 4, 2015.
- [4]. Bangsuk Jantawan<sup>1</sup> and Cheng-Fa Tsai<sup>2</sup>, A Classification Model on Graduate Employability Using Bayesian Approaches: A Comparison, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 6, June 2014.
- [5]. Azami zaharim et al., “Engineering Employability Skills Required by Employers in Asia”, Proceedings of the 6th WSEAS International Conference on Engineering Education.
- [6]. G.Swamy, “The need and importance of employability skills at ug level: a study of private degree colleges of telangana state”, International Journal of English Language, Literature and Translation Studies, Vol.2. Issue 3, 2015 ( July-Sept ).
- [7]. A. Ravichandran et al., “A Study on Job Satisfaction of Employees of Manufacturing Industry in Puducherry, India”, International Journal of Innovative Research and Development, Vol 4, Issue 2, February, 2015.
- [8]. Chandra Sekhar Patro<sup>1</sup> and B.Lohit<sup>2</sup>, “Impact of Unemployment on Engineering Graduates in Recession”, IOSR Journal of Business and Management, Volume 16, Issue 2. Ver. III (Feb. 2014).
- [9]. N. B. Awandekar et al., “Employability skills assessment amongst pharmacy post graduates in nagpur region”, International Journal of Pharmaceutical Sciences and Business Management, Vol.4, Issue. 10, October – 2016.
- [10]. Priyadrsan Dumbre and Vinod Shrirang Ingawale, “Student Profiling: Importance in Enhancing StudentEmployability”, International Journal of Multifaceted and Multilingual Studies, Volume-iii, issue-IX, 1st September 2016 .