# Chapter 01

**Introduction**

# 1.1 Introduction

Campus placement of a student plays a very important role in a college. Campus placement is a process where companies visit colleges and identify students who are talented and qualified, before they complete their graduation. Therefore taking a wise career decision regarding the placement after completing a particular course is crucial in a student’s life. An educational institution contains a large number of student records. Therefore, finding patterns and characteristics in this large pool of data, will help find parameters that are the most important for this placement procedure.

The prediction of engineering students, about where they can be placed, from the second year and onwards, will help to improve efforts of student for proper progress. It will help teachers to take proper attention towards the progress of the student during the course of time. It will help to build reputation of the institute for having such a sophisticated system in place which helps the students to train and practice for campus placements. The present study concentrates on helping the students, bridging the gap between the industry and the curriculum, and showing them the path to a better future. We apply data mining and machine learning techniques using Artificial Neural Networks, in order to interpret the potential of the student.

**1.2 Aim and Objectives**

Our project aims to create placement guidance system which will use the concept of Artificial Intelligence. We intend to combine both qualitative and quantitative parameters for the decision making process. To do so we consider the academic history of the student as well as their skill set like, programming skills, communication skills, analytical skills and teamwork, which are tested by the hiring companies during the recruitment process. Though many research has been done previously on placement prediction using different methods, none of them gave consideration to qualitative parameters to a large extent, which plays a vital role in placement of any student. Thus, by taking this into account our aim is to achieve a system with greater than 85% of accuracy. Predicting the placement of a student gives an idea to the placement office as well as the student on where they stand. Not all companies look for similar talents. If the strengths and weaknesses of the students are identified it would benefit the student in getting placed. The placement Office can work on identifying the weaknesses of the students and take measures of improvement so that the students can overcome the weakness and perform to the best of their abilities. Thus the key lies in assessing the capabilities of the student in the right areas and subjecting them to the right training which is essentially our objective behind creating such system.

**1.3 Scope**

The work currently deals with predicting the results based on last 5 year’s training and placement data, examination department’s data, and the extra curricular data of the students. The present work will be using a regression model which is training itself, so the accuracy of the system will increase over time, making the system more reliable over time.Even if the parameters change or the system will adapt to it over time. In future, this system will be extended on the web end to give better outputs and competitive insights and also providing more statistics to the corporate end.This system can be used for many years, as the system is adapting with the data, and the accuracy is increasing over time.

**1.3.1. Resume analysis**

The only input to the system is the resume, which is provided by the user. Further the resume is analysed for user’s skills, education qualifications, achievements, certifications and extra-curricular activities; using Natural Language Processing and Text Mining.

**1.3.2. Crawl Resume for different questions**

The questions and answers, pertaining to different technical skills, are asked to the users which are crawled from popular websites like [www.indiabix.com,](http://www.indiabix.com/)[www.geeksforgeeks.org](http://www.geeksforgeeks.org/) and [www.tutorialspoint.com.](http://www.tutorialspoint.com/) The crawling is done with the help of PHP-Spider. PHP spider is a crawler which supports two traversal algorithms: breadth-first and depth-first, supports crawl depth limiting, queue size limiting and max downloads limiting, supports adding custom URI discovery logic, based on XPath, CSS selectors, or plain old PHP.

**1.3.3. Difficulty Level identification of the questions**

For identifying the difficulty level of questions, we allocate a score to each question. This score is calculated based on manual scoring system. This is the base level of each question.

**1.3.4. Skill based questions**

Questions asked to the users are based on both, the skills retrieved during the resume analysis and the user’s score.

**1.3.5. Generated scores based on the answers to the questions**

As the user answers the question of the test, the difficulty level of both the questions as well as the level of user changes. If the user’s answer is correct, the level of user increases and correspondingly