Job Portal Analysis and Salary Prediction System

Disha Achrekar[1], Mayuri Pawade[2], Griffin Mathias[3], Bhanu Tekwani[4]

[1,2,3]Student, Information Technology, Vidyalankar Institute of Technology, Mumbai, India
[4]Assistant Professor, Information Technology, Vidyalankar Institute of Technology, Mumbai, India

---***---

Abstract - *The amount of information about educational and job-related sources spread over the internet is huge and diverse like Tech-blogs and Job portals like Naukri, Monster, Times Job etc. are gaining momentum. With the usage of all these various sites and resources present on the internet finding or choosing the right platform has not only become difficult but also time consuming. This paper reviews an effective and efficient job portal analysis and salary prediction system with which any student (fresher) or professional could choose the right platform to build a career in by knowing exactly which skills are trending in the IT Market.*

Key Words: Job Analysis, Web Scraping, Prediction, Visualization, Machine Learning, Regression Analysis.

1. INTRODUCTION

There are several job-portal sites that proposes jobs for individuals by the recruiter or crawled from other job portals or websites. These jobs posted include job title, location, experience and various skills needed but only few disclose the salary. Also, individual profiles of employees are rarely modeled for salary. The main objective of these sites is to manage the details of Employer, Employee and ensure to make a wide range of job available to the user. However, it fails to analyze the skills in demand with emerging technologies in the industry. Determining which skill is trending and contains a greater number of jobs and having the highest salary turns out to be a challenging task. Such a complex problem with a certain amount of uncertainty is always a challenging proposition. In this paper we concentrate on individual interested skill, company or location and a salary model with predictions. It is efficient and ethically according to the latest market policy and trends, we need to take the reference of large amounts of data from various job sites taking all the features is huge task and is not possible for a human being which is why we have tried to make it easy by web scraping and taking certain specific attributes and introducing this system. Such a prediction engine [3] takes certain key features as input and gives the most favorable outcome by analyzing all the factors based on the previous data.

2. LITERATURE SURVEY

The work of such projection has been studied by Sananda Dutta, Airiddha Halder and Kousik Dasgupta [1] also has emphasized on the problem of predicting salary for job advertisements in which salary are not mentioned and also tried to help fresher to predict possible salary for different companies in different locations. Himanshi and Komal Kumar Bhatia [3], they compared the profile of current students with the former graduated student presenting a salary prediction system using data mining techniques. Shaun Jackman and Graham Reid [4] have worked Predicting Job Salaries from Text Descriptions, they have tested a variety of regression models including maximum-likelihood regression, lasso regression, artificial neural networks and random forests. Singh. R [5] has done a detailed study to find the salary determinants of fresh undergraduate engineers in Indian job markets.

3. PROPOSED SYSTEM

There are two modules in this system: Data visualization has been implemented to display the trending skill in a company or location in the form of a graph and a prediction model to predict an average salary.

A. Data Collection

The system needs data from various job portals to perform analysis and prediction. Two main datasets needed for the system have been collected and stored separately.

1. To identify the trends in skills: With the help of online web scraper tool data has been scraped into .csv file (comma-separated-value) from the I.T section of the Times Job portal.

2. For Salary Prediction: Data was collected analyzing various valid sources such as Naukri, Monster etc. and these salaries were used to train a salary-activity model to predict a salary.
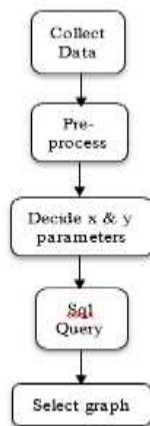
B. Data Visualization

Fig -3.1: Steps for Visualization

The dataset consists of Location, Company name, and Skills.73251 tuples and 3 attributes have been visualized selecting x parameters as skills and y parameter as location or company, which displays a graph after execution of an sql query emphasizing which skills have the greatest number of jobs based on the parameters to improve the experience of people searching for jobs, and help employee and job seekers to figure out the latest market worth.
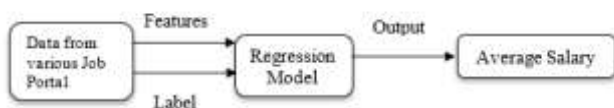
C. Salary Prediction

Fig -3.2 Prediction Model

Further, to determine the salary of the employee based on their interested Domain, Job title, skills, and experience. Hopefully, knowing their predicted salary will urge them to pay more attention to studying with a clearer future direction and outcome. Machine Learning technique was a core component of the prediction system. The module is implemented using Regression algorithm, which is a statistical approach to find the relationship between variables. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. It performs the task to predict a dependent variable(label) value (y) based on a given independent variable(features) (x). Considering the dataset:

1. Independent variable (x) - Domain, Job title, Experience, Skill.

2. Dependent variable (y) - Salary.
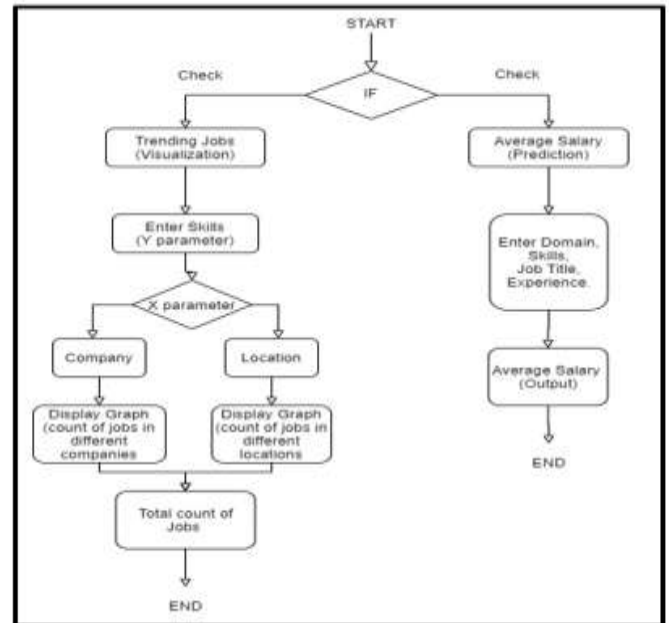
4. METHODOLOGY

Fig -4.1: Flowchart

- The System consist of two Modules.
- If the user is a fresher and wants to check the trending technology/ skill would have to select x and y parameters which would further display two graphs.
- Y parameter: Skill
- X parameter has two option, i.e. if the user wants to check the number of jobs available in various company should select X parameter as company, similarly if the user wants to check the number of jobs available in various location should select X parameter.
- 2nd Graph would display total number of jobs available in the selected skill.
- And, if the user is experienced and wants to check the salary in a particular domain should enter the domain name, Job title i.e. the designation the user wants to work in, Skills related to the job designation the user knows and the experience the user has in that domain.
- The input goes to the Prediction Model which generates an annual salary.

5. CONCLUSION

In this paper we have examined the data collected from Times-Job, Naukri, Monster, etc. This system will help people to analyze the current job trend and will aid freshers w.r.t

the skills, location and company in the market. We have also identified the features required to predict salary in an efficient manner. The system will use regression algorithm to analyze and predict average salary efficiently.

6. REFERENCES

[1] "Design of a novel Prediction Engine for predicting suitable salary for a job",Sananda Dutta ,Airiddha Halder and Kousik Dasgupta, Department of Computer Science and Engineering Kalyani Government Engineering College Kalyani, India.

[2] Bauman, "What is a prediction engine?" MindMeister, 20-Nov-2013. [Online]. Available: https://www.mindmeister.com/349147807/whatis-a-prediction-engine. [Accessed: 14-Oct-2018].

[3] "Prediction Model for Under-Graduating Student's Salary Using Data Mining Techniques",Himanshi1, Komal Kumar Bhatia2,Department of Computer Engineering, YMCA University, Faridabad, Haryana, India.

[4] Jackman, S., & Reid, G. (2013), "Predicting Job Salaries from Text Descriptions", Doctoral dissertation, University of British Columbia.

[5] Singh, R. (2016), "A Regression Study of Salary Determinants in Indian Job Markets for Entry Level Engineering Graduates", Masters Dissertation. Dublin Institute of Technology.