

DESKTOP BASED RECOMMENDATION SYSTEM FOR CAMPUS RECRUITMENT USING MAHOUT

¹Ronak V Patil, ²Sneha R Gadekar, ³Prashant P Chavan, ⁴Vikas G Aher

Department of Computer Engineering, University of Pune, Imperial College of
Engineering and Research Pune, India

¹ronak_patil@yahoo.co.in, ²sneha.gadekar5393@gmail.com,

³prashantc724@gmail.com, ⁴ahervikas780@gmail.com

Abstract: Data mining is the process of posing queries and extracting patterns, often previously unknown from large quantities of data using pattern matching or other reasoning techniques. This project presents a Desktop Based Recommendation System for Campus Recruitment Using Mahout, which can help college placement office to match the company's criteria with students' eligibility. This will reduce the need of human intervention. This system automates the activities of placement cells in colleges. Our main focus is on profile matching. The eligible candidates among all will be matched with company's criteria. For doing this we will use Mahout's Naïve Bayes Classification algorithm Based on profile similarity degree, the preference lists of companies and students are calculated, which serves as the input of two-sided matching. The new system embedding SMS-based interaction can raise the matching degree, shorten recruiting period and reduce cost. Furthermore, this recommendation service not only is applicable in the field of campus recruitment, but also can provide a framework for the field of mobile business with the extension to other domains such as hospital-intern and college-student matching and recommendation.

Keywords: Mahout, Pattern Matching

1. INTRODUCTION

Internet has been the most useful technology of the modern times which help us not only in our daily lives, but also our personal and professional lives developments. The internet helps us achieve this in several different ways. For the students and educational purposes the internet is widely used to gather information so as to do the research or add to the knowledge of any sort of subject they have. This led to develop a web based placement system. So that anyone with a proper logging can access it from anywhere. This system will help us to do all the placement tasks automatically. This system is developing to

improve the existing system. Drawbacks of existing system can be removed by this system. This system has improved functionality. It has the functionality of maintaining important data required for placement and students information very efficiently. With this system there are less chances of data loss. The time required for maintaining all the data manually will be reduced with this system. All the data will be maintained automatically with the help of this system. We can save a lot of time. Students can get information regarding the arriving companies with this. Students can get all the data at one central place. They can access to this information from anywhere with a proper logging. TPO can get the list of eligible students among all automatically with this. TPO can access this system with a proper logging. She/he can access all students' resumes. He can go through them all. As soon as he uploads the information regarding the upcoming placement drives he'll get a list of eligible students allowed to sit for the drive. This will be all automatically done. This system will reduce a lot of workload of TPO.

2. LITERATURE SURVEY

Now a day's every college's placement cell suffer from the problem of maintaining data required for placement. The data is so large in amount and it becomes difficult for maintaining and using it. There is a need to maintain this data wisely. Today the tasks done for placement recruitment in college campus are all manually. And this task is very tedious task. Lots of efforts need to be taken for carrying out this process. And there is a problem in maintaining the data as well. So we are proposing a system named "Desktop Based Recommendation System For Campus Recruitment Using Mahout" in which we are going to do all this task automatically. We are implementing Mahout's Naïve Based Classification algorithm for filtering the data. And we are also using MySQL database for storing the data.

2.1 E-recruitment [1]

In this paper author proposed that E-recruitment is a real revolution spreading over the world of job hunting & hiring. The term online recruitment, e-recruitment, cyber cruting, or internet recruiting, imply the formal sourcing of job information online. The first references to e-recruitment Appear in articles of the mid-1980s E-recruitment can be divided into two types of uses: corporate web site for recruitment and commercial jobs boards (such as monster.com) for posting job advertisements .Corporate websites are a company's own website with a link for job posting/career options where candidates can log into for current openings. If the company advertises its vacant positions on other website that specialize in recruitment such as - naukri.com, timesjob.com, monster.com, etc., the companies would be adopting commercial job boards for recruitment.

2.2 An Interactive Online Training and Placement System [2]

In this paper author proposed that OTaP system is an online application that can be accessed throughout the organization and outside as well with proper login provided. This system can be used as an application for the Training and Placement Officer (TPO) of the college to manage the student information with regards to placement. Students logging should be able to upload their information in the form of a CV. Visitors/Company representatives logging in may also access/search any information put up by Students. Computer based information system are designed to improve existing system. Whatever the information, TPO has to pass to the student and he or she can inform online. Improve accuracy in result. It has user-friendly interface having quick authenticated access to documents. It provides the facility of maintaining the details of the students. It will reduce the

paper work and utilize the maximum capabilities of the setup and organization as well as it will save time and money, which are spending in making reports and collecting data. It can be access throughout the organization and outside as well with proper login provided. This system can be use as an application for college to manage the student information concerning placement. Also helps company coming for campus recruitment to see student details. Before coming for campus, company can get information about eligible students along with interested students

3. CURRENT SYSTEM

Now a day's every college's placement cell suffer from the problem of maintaining data required for placement. The data is so large in amount and it becomes difficult for maintaining and using it. There is a need to maintain this data wisely. Today the tasks done for placement recruitment in college campus are all manually. And this task is very tedious. Lots of efforts need to be taken for carrying out this process. And there is a problem in maintaining the data as well. So we are proposing a system named "Desktop Based Recommendation System for Campus Recruitment Using Mahout" in which we are going to do all this task automatically. We are implementing Mahout's Naïve Based Classification algorithm for filtering the data.

4. PROPOSED SYSTEM

The proposed Placement system helps in making the task of Training and Placement officer easy. It eliminates all the human intervention. Human efforts are reduced. To access this system all you need to have is an authorized login id and internet. You can access this system anywhere from your campus or outside campus. All you need to have is your internet connection on. There are two main authorized users of this system. One is student login and other is Administrator login. Students can access the system by authorized student login and TPO and TnP staff can access the system by Admin login. Students can upload their resumes, can check upcoming placement drives and can also get study material from here. The other access to the system is for Administrator. TPO and TnP staffs are two main users of this system. They can access and make updates to the data in the system. TPO can view the system. He can view all students resume. He can upload information regarding upcoming drives. He can also upload information regarding the companies which will be coming. This information can be helpful for the students. TPO gets the highest priority among both the users. He can access as well as update the data. TnP staff can recheck all the information filled by the students. They approve the correctness of the data. Our key element is Mahout's Naïve Bayes Classification algorithm.

4.1 Introduction to Apache Mahout

The success of companies and individuals in the information age depends on how quickly and efficiently they turn vast amount of data into actionable information. Whether it's for processing hundreds or thousands of personal e-mail messages a day or divining user intent from petabytes of weblogs, the need for tool that can organize and enhance data has never been greater. Therein lies the need of Apache Mahout.

Apache Mahout is a new open source project by the Apache Software Foundation (ASF) with the primary goal of creating scalable machine learning algorithms that are free to use under the Apache license. Mahout contains implementations for clustering, categorization, CF and evolutionary programming.

In our proposed system we are using Mahout's Naïve Bayes algorithm for classifying the data. Classification algorithms can automatically classify documents, images, implement spam filters and in many other domains. Naïve Bayes algorithm works by using a training set which is set of documents already associated to a category. Using this set, the classifier determines for each word, the probability that it makes a document belong to each of the considered categories. The category with the highest probability is the one the document is most likely to belong to.

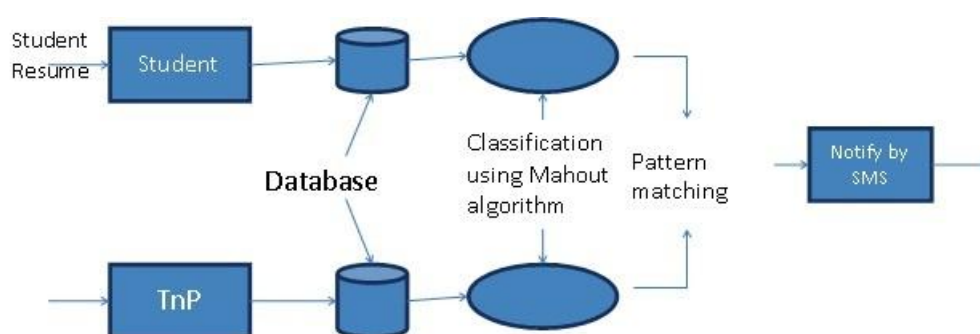


Fig.1: Architectural Model

5. FLOW DIAGRAM OF PROJECT

The flow diagram basically represents outline of any work. In our project following represents the flow diagram. We will go through it. Database is necessary for storing any kind of huge data. So, we have created database, where it would be easy to store and retrieve data. We are supposed to create two modules:

a) Student Module: This will be one of the main modules of our system. This will be a GUI on which student will upload their resumes. The resumes can be in any format. The student can also get information regarding the upcoming drives. They can also get aptitude questions and companies' information. Student can interact with the system using this module.

b) TPO Module: This is another important module of our system. TPO can interact with the system using this module. TPO can upload the information regarding the upcoming drives there. He has all the access to change or modify the data. He can access students' resumes too.

c) Admin Module: Admin has the central control of the system. He controls TPO. He can add as many TPOs as he wants. He can manage as well as make changes to the database.

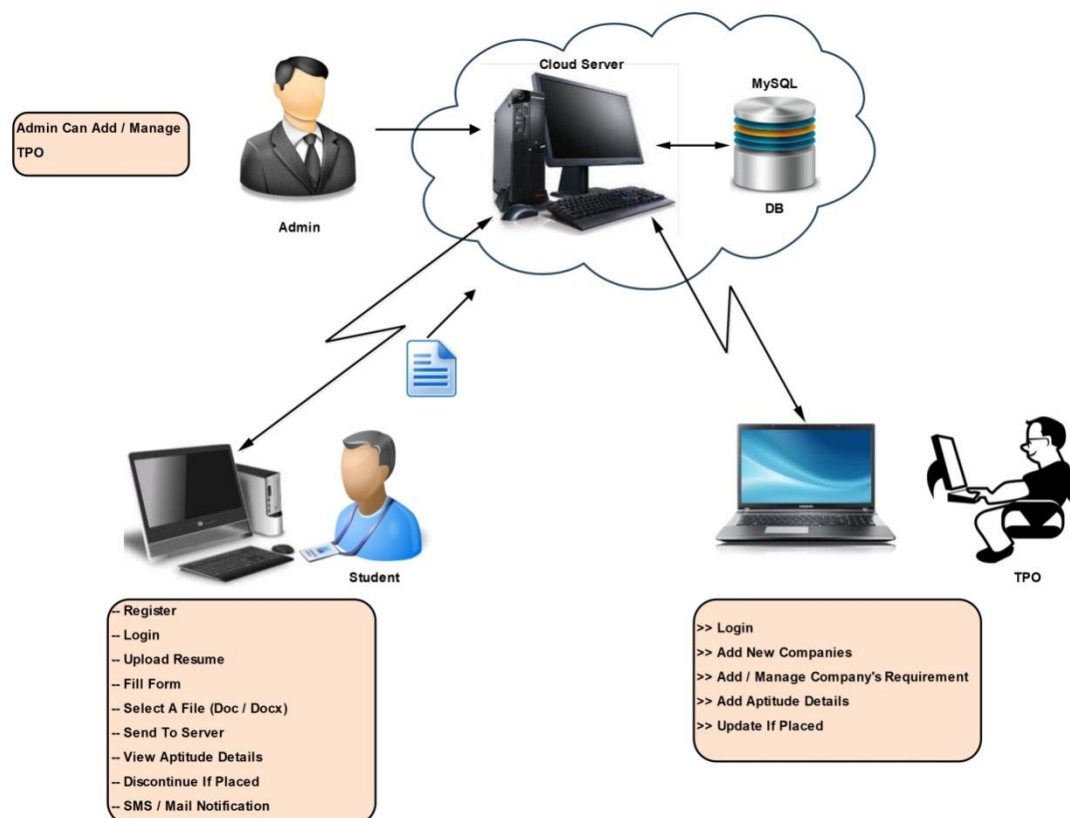


Fig. 2: Flow Diagram of Project

6. ADVANTAGES

1. Since the system uses data mining it depends on already developed code.
2. System is easy to handle.
3. Gives output in very less time so scalable and useful for large applications
4. Requires less human interference

7. DISADVANTAGES

- Need to update the data on regular basis.

8. CONCLUSION

Existing system is error prone and slow as it includes human intervention. There are more chances of errors and incorrect results. The proposed system will help college placement office to match the companies and students with higher precision and lower cost. We are mainly focusing on profile matching. For profile matching we are using Mahout's Naïve bayes Classification algorithm.

9. FUTURE SCOPE

The system has been designed at the maximum possible excellence. Still we accept drawbacks, as it is a human effort. The system cannot provide the sms integration. Hence, it can be modified. New functionality can be added without affecting the modules.

ACKNOWLEDGEMENT

It is with great pleasure that we acknowledge the enormous assistance excellent co-operation extended to by the following respected personalities. Firstly we would like to express our sincere thanks to our respected Principal Dr. Sachin Admane sir for providing us the necessary infrastructure, laboratories and requirement. We are greatly indebted to, Head of Computer Department, Prof. S.R.Todmal for providing valuable guidance. We would like to express our sincere thanks to our respected Guide Prof. Nilesh Sable for his valuable guidance.

REFERENCES

- [1]. E- Recruitment (International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-4, April 2012). Avinash S. Kapse, Vishal S. Patil, Nikhil V. Patil
- [2]. Dr Seema Shah, Vidyalaya Institute of Technology Wadala. Proceeding on An Interactive Online Training and placement system, available at www.ijarcsse.com.
- [3]. Talaba, D., Moja, a, Zirra, E., Guidelines towards a European standard for quality assurance of student placement, available in login space on www.q-planet.org.
- [4]. Zirra E., March F., Building University - Enterprise Cooperation for the Benefit of Students, Enterprises and Companies. EUI-Net workshop, Athens 28 September (2006). http://www.eui-net.org/Project_documents/.
- [5]. Tynjälä, P., Perspective into learning at the workplace, Educational Research Review, 3, 2008, pp. 130-154.
- [6]. Training and Placement Department of Rajiv Gandhi Institute of Technology.
- [7]. Talaba D., University-Industry cooperation in the Knowledge based society. Proceedings of the 2nd International EUI-Net Conference on: Teaching and Research Synergy, 4-6 May , Tallinn, pp.7-13. (2006).