A Pearson Correlation Based Approach for Assisting Recruitment Process

AmeghR

Department of CSE
Reva University, Bangalore, India
amegh.r@gmail.com

Abhik Shil

Department of CSE
Reva University, Bangalore,
Indiaabhikshil@gmail.com

Abhishek Kumar

Department of CSE
Reva University, Bangalore,
Indiaabhikr.9399@gmail.com

Ankur Ashuthosh Jha

Department of CSE
Reva University, Bangalore,
Indiaankurjha72@gmail.com

Prof. Supreeth S

Department of CSE
Reva University, Bangalore,
Indiasupreeth.s@reva.edu.in

Abstract:Currently, placement-related updates are being provided mainly through WhatsApp and Email, which causes ambiguity and miscommunication. This paper studies the Indian placement scenario of today and highlights the necessary changes to be brought about to ensure a smooth concise placement process. It looks at how potential recruiters interact with students and universities to hire desired candidates for various roles. It is based on a placement application built by the team for campus recruitment. A machine learning approach is used which predicts the required study material for the students with the help of Pearson Correlation method which considers the test score of a student as a deciding factor and compares it with the score assigned to the placement material for suggesting the resources in order to improve the knowledge related to a topic. This approach helps a

student in understanding their weakness and suggest the required solution for overcoming those by making use of the machine learning prediction model

Keywords: Placement Assistance, Pearson Correlation Approach, Placement Study material, Coding platform

I. INTRODUCTION

` Campus Placement is an absolutely complicated process. Right from the start of inviting companies for placements to making the students placement ready and getting them placed in a reputed organization. The entire process requires dealing with a lot of data and hence the systematic storage and availability of data are quite essential.This paper is aimed at developing an application for the Placement Department of a College. The system can be accessed throughout the organization with authentication. Once any company approaches for placements, the admin will upload details about the company which includes the eligibility criteria, package, number of openings, contact details, and other suitable details required to be shared with the students. The admin can filter students based on the eligibility criteria and share the info as well.The students have to option to evaluate themselves by taking up aptitude test and based on the score they receive the machine learning model implemented will predict the study materials required for the student using Pearson Correlation that will calculate the topic wise score and suggest resources related to that particular topic. Taking inputs [13], the personalized recommender system was created for making the required placement material prediction.The companies then can utilize the testing and coding platform to conduct their rounds using Cloud Computing [14]. Students will be updated regularly on the placement process and selection. Also features like filtering eligible students, chat-bots, and other statistical information will be available for making the user experience and decision-making better. The main motivation behind choosing this project idea came from the current scenario of undergoing campus placements, certain voids can be filled with our application.Workflow Diagram of Placement Assistance Application is shown in the Fig. 1.

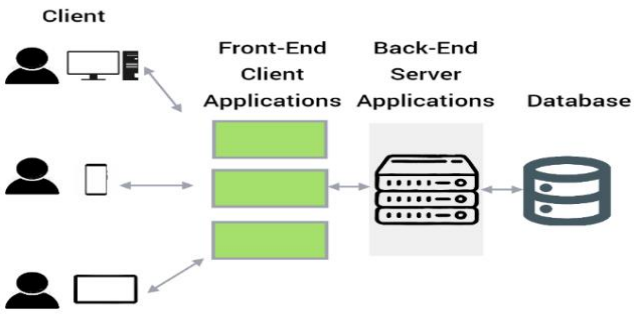

Fig 1. Workflow Diagram of Placement Assistance Application

Following are the objectives which are been addressed in this work.

1. Create a placement assistance application that helps in overcoming the problems faced in the traditional manual practice of conducting offline recruitment drives.
2. Implement a machine learning prediction model using Pearson Correlation method
3. Include features like resume builder, coding platform, chat-bot, study materials for student support
4. Collect data from the company to provide a job description for students as well as for the placement materials and to train chat-bot
5. Ensure the application is user friendly and is easily accessible for the students and the placement department team of a particular institution

II. Literature Survey

In [1] mainly address about the various advantages and benefits of online job portals. The mainidea behindof creating such a platform is to support thestudentsin finding the profession theywant. The existing web applications and not very efficient. To overcome the difficulties and make a platform which is efficient and provide accurate results, huge data was gathered, and a new website was created. This website will be useful for a student on that campus who is preparing to get placed in a reputed company as it gives job notifications based on the ability and skill of a candidateand the employers can filter the students based on the requirements of their company. So, this can help a student to exploremanyjobs and the recruiters can find suitable employees easily

In [2]addresses about the various issues that are faced by every employer because of the huge number of job applications it receives. The proposed model aids the HR department in collecting the required information about a candidate and shortlisting them as well as rearranging them based on the skills which are mentioned by the applicant in their CV or resume. The model also conducts an interview with the help of a chatbot in order to find out fake CVs and also to understand the personal skills and behaviour of the candidate by capturing facial expressions through the camera. The candidates that are interested to work in a company can check the job requirements and then apply for that post. Therefore, the virtual process will reduce the hectic workload of the recruitment team and helps them in saving a lot of time.

The literature review which is described in [3] is mainly intended for students whoare in the second last year perusing various courses to get placed in a company they dreamt of. It discusses regarding various predictions that will be helpful for students to realize the existing knowledge they have and the steps they need to take which can improve their skills. A device called placement suggester will examine the chances of a student getting selected for a company and many programs created which will be useful for an academic institution for providing training to improve the skills and knowledge of a candidate. By making use of various algorithms and by exploring data, a variety of predictive models can

be created which will help a candidate to travel on the right path in getting a satisfactory job.

In [4] demonstrates how a job prediction model can be created for students and also how data mining will be useful in structuring the educational system to get the best results for placements. The model will collect information from the students and based on that the students can be classified according to the type of companies in which the student is likely to get selected in like companies that recruit heavily, which pay more than average,and which pay high salary which is dream job. The described model will also suggest the various new skills and technologies in which a student should gain knowledge so that their chances of getting selected will be high. The various results, performance measures, as well as the real-time findings of the model,are mentioned in the paper which plays a major role in improving the strategies that an educational institution must follow to train a student to become a good candidate for the placement activities and also in real life.

In [5] describes the new strategy which is being followed for conducting recruitment of students from an institution. The various problems and disadvantages of the traditional method of conducting the drives where the recruiters visit the campus for shortlisting the students to find the best candidate who is suitable for the job are mentioned. This paper comes us with an ideal solution by making the placement drives online where the students can attend the various rounds of selection through their laptop or desktop from anywhere and get selected for the job. Recruiting a candidate by making use of the internet is given the term green placement. This makes it easy for the students to participate in the recruitment process and also for the selectors and placement department of a particular institution for conducting the shortlisting process easily and smoothly. This method helps in saving many resources like time, space for conducting drives water, electricity, cost, etc.

In [6], discusses about creating a job proficiency profile of a candidate which can be helpful in understanding various details about the candidate to help them in improving their knowledge and skillset. The eligibility of a candidate for applying to a specific job can be found with the help of a competence profile so that it becomes easier for the recruiters as they will be receiving applications only from the eligible students. The feedback received after taking various tests and skill up-gradationprograms will help the candidate understand where they stand and the necessary steps that need to be taken to improve themselves. In this way, a candidate can prepare with a proper strategy for building their skills and will be useful in getting placed in a reputed company.

In [7] described two mathematical models which recommend a suitable candidate for hire. It becomes a gruellingtask when the number of candidates who applies for a job is huge. Doing the work manually can consume a lot of time and effort and the chances of creating an error are also high. For understanding a candidate better, comparing a profile method can be used and shortlist them for further rounds and Naïve Bayes can be used to predict if a student has the chance of

getting placed or not. Bivariate correlation is used as an assessment factor for Profile Matching. The accuracy and also correlation data are provided in the paper which depicts the placement of a candidate with 100% efficiency by making use of Naïve Bayes method.

In [8] addressed the onlinehiring process canbe mainly based on the information that are provided on theresume. For finishing the process of selecting of a candidate for a particular job, most of the tasks should be automated. One of the main automation tasks in the selection process is to compare the different wordsthat are given in the candidate's profile with the ones that are mentioned in the job description.This method sometimes gives low exactness, and many suggestions are not dependable. Based on the meaningslatest approaches have been created but it has limitations because of the partial resources and lack of proper domain coverage.The system was deployed victoriously, and attestation is provided in the paper which shows that it becomes very much useful in the hiring process.

III. Methodology

The Placement Assistance Application is helping in overcoming such difficulties and making the process of placements much smoother which also helps in saving many resources like cost, time, paper, electricity, etc.

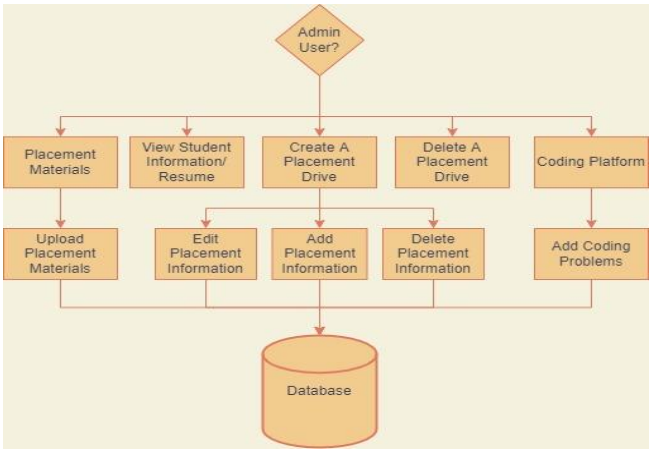

Fig 2. Block Diagram of Admin Functionalities

Fig 2. Depicted the block diagram of the Placement Assistance Application. The Placement Assistance Application consists of one Admin Module and a Student/User Module.

3.1 Admin Functionalities

Taking reference from Fig 2 we can see all the mapped activities of the admin. The admin can register themselves on the website by providing simple details which will include their name, email address, university, and password that are supposed to check

the admin box if registering as an admin. On successful submission of the details. The university authorities will verify the identity of the admin. On successful verification of admin, the details of admin are stored in the database. Here the password stored will be hashed using the sha256 hashing technique. Fig 3 shows the Block Diagram of the System

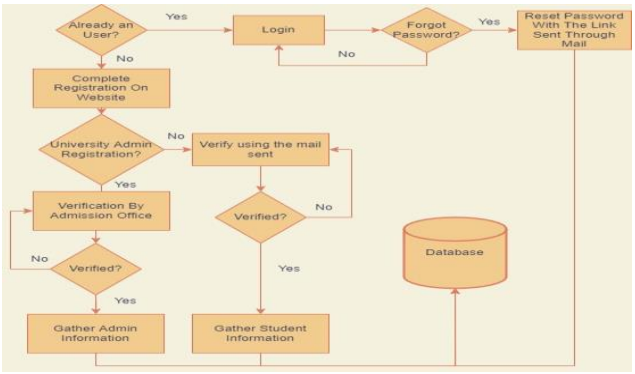

Fig 3. Block Diagram of Registration/Login System

The admin then can have the functionalities to add a placement. Here he/she is supposed to provide details which will include the company name. Once the company tab is created the admin can go into that tab and make further updates. There the admin can add/edit an update from the company or placement department by clicking add notes.The admin will also have a list of candidates with their resume details which the admin can view and provoke necessary actions and share them with the companies recruiting.

3.2 Student/User Functionalities

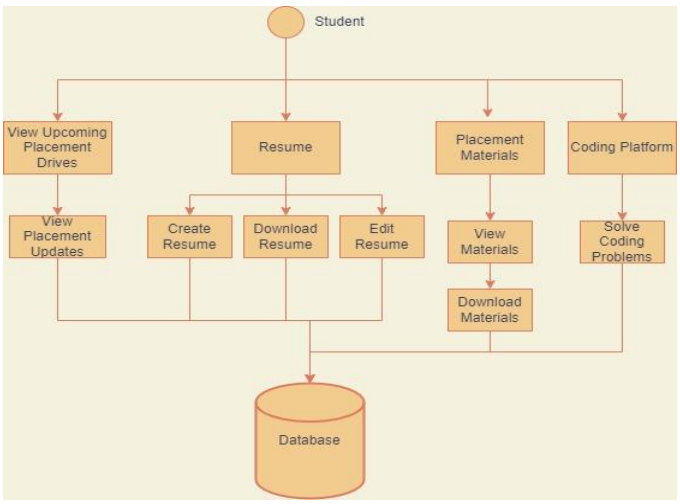

Fig 5. Block Diagram of Student Functionalities

Fig 5. Here depict the functionalities that a user can perform onthis website. The student can register themselves in the similar way it is mentioned in case of admin. The user will

receive a mail to verify themselves once done, they can revert to the application and login themselves.

The users here will have the functionalities of viewing all the upcoming placement drives. Once the user clicks on any recruiting company tab. He/she will be able to see all the updates regarding that hiring process and further updates from the company.Also, the users will be able to create their own resume by going to the resume section and clicking on create your resume. Here the user will be required to provide certain details like their 10^{th} class percentage, 12^{th} class percentage, skills, hobbies, etc. Once the user submits these details, they will be able to view their resume. By clicking on the download resume button, the user can download their resume in pdf format. And they can also edit their resume if required by the edit resume button.

3.3 Placement Materials

The users can head over to the placement materials tab, they can access all the placement materials required. They can filter them using the search tab.And they can head over to practice questions to test their skills. The admin can upload and update the necessary materials when required for the students to access and download from the application.

One of the differentiating factors in our application is the placement material recommendations. Users will be first prompted to give a test, based on the score they receive in this particular test, the users get recommended with study materials that they might want to refer.

For performing this particular task, we have used a machine learning algorithm using Pearson Correlation technique.Pearson Correlation Coefficient algorithm is widely used algorithm for various similarity measures as mentioned in [10]. Hence a suitable similarity measure key is needed to be chosen for the system to make proper recommendations.

Considering two users U and P respectively the Pearson Correlation Coefficients can be given as mentioned in [11]

$$w_{a,u}=\frac{\sum_{i=1}^{n}(r_{a,i}-\overrightarrow{r_a}).(r_{u,i}-\overrightarrow{r_u})}{\sqrt{\sum_{i=1}^{n}(r_{a,i}-\overrightarrow{r_a})}.\sqrt{\sum_{i=1}^{n}(r_{u,i}-\overrightarrow{r_u})^2}} \quad (1)$$

$$p_{a,i}=\frac{\sum_{u=1}^{m}(r_{u,i}-\overrightarrow{r_u}).w_{a,u}}{\sqrt{\sum_{i=1}^{n}(r_{a,i}-\overrightarrow{r_a})}.\sqrt{\sum_{i=1}^{n}(r_{u,i}-\overrightarrow{r_u})^2}} \quad (2)$$

Where**'a'** represents active user, **'u'** represents another user in the system, **'n'** represents the number of items both users recommended, '**ra**' represents the rating of active user and '**ru**' represents the rating of another user.

The correlation of two independent variables as given in [12] can be derived if we have two zero-meaned random variables. Let us consider x and y as two zero-meaned variables then the Pearson Correlation Coefficient can be given as [12]

$$\rho(x,y)=\frac{E(x,y)}{\sigma_x.\sigma_y} \quad (3)$$

Where **E(x,y)** represents the cross-correlation between x and y and $\sigma_x$and σ_y represents the variances of x and y respectively.

Algorithm 1: Placement Material Recommendation Using Pearson Correlation in Python

Input: User Test Score csv and Placement Materials csv

Result: Recommended placement materials for users to refer.

1. START
2. Fetch the user test score and placement materials csv files from the database.
3. import pandas and NumPy as pd and np respectively.
4. Set user_test_score using pd. read_csv(user_test_data.csv)
5. Set user_test_score_columns = ['user_id','aptitude_score','programming_score']
6. Set user_aptitude_score to user_test_score['aptitude_score'] and user_programming_score to user_test_score['programming_score']
7. Set placement_material_data using using pd. read_csv(placement_material.csv)
8. Set placement_material_data.columns = ['material_id','material_title','material_type','material_author','score']
9. Set aptitude_materials to all placement_materials which have material_type=="aptitude"
10. Set programming_materials to all placement_materials which have material_type=="programming"
11. Form a DataFrameaptitude_material_score_count using pd.DataFrame(aptitude_materials.groupby('material_id')['score']. count()) and sort it using sort_values().
12. Also Form a DataFrameprogramming_material_score_count using pd.DataFrame(programming_materials.groupby('material_id')['score'].count()) and sort it using sort_values().
13. Now using correlation between user_aptitude_score and aptitude_material_score_count using aptitude_material_score.corrwith(user_aptitude_score) and store it in corr_aptitude_material.
14. Also using correlation between user_programming_score and programming_material_score_count using programming_material_score.corrwith(user_programming_score) and store it in corr_programming_material.
15. Form DataFramesrecommended_aptitude_material and recommended_programming_material using pd.DataFrame(corr_aptitude_materials,columns=['pearsonRating']) and pd.DataFrame(corr_programming_materials,columns=['pearsonRating']).
16. Finally merge the recommended_aptitude_material with placement_material_data on column 'material_id'.
17. Also merge the recommended_programming_material with placement_material_data on column 'material_id'.
18. Now the both the datasets are ready to be recommended to the users.
19. STOP

3.4 The Coding Platform

The users can practice their coding skills in various programming languages.

Propercompany-specific coding problems with test cases will be provided to the users to practice their coding skills and get placement ready.

Various compiler for compiling languages like C, C++, python and Java need to be incorporated properly in order forimplementing the test cases accurately as mentioned in [14].

3.5 Remember Me feature (Cookie creation)

Algorithm 2 shows the cookie creation for implementing the remember me feature, so that the user will still be remembered even after the browser is closed.

Algorithm 2: Remember Me Cookie Creation

Input: Two Randomly Generated Strings Result: Cookie for remembering the user when they visit next time.
1. START 2.Check if user has checked the "Remember Me" box. 3.If yes, then create two variables authentificator1 and authentificator2 4. Assign authentificator1 to a randomly generated string of bytes using opensl_random_pseudo_bytes(10) function. 5.Convert authentificator1to hexadecimal using bin2hex(authentificator1) function. 6. Assign authentificator2 to a randomly generated string of bytes using openssl_random_pseudo_bytes(20) function. 7. Concatenate authentificator1 and authentificator2 and store it in a variable cookieValue. 8. Now set the cookie by passing the cookieValue variable in setCokkie(cookie_name,cookieValue,expirytime) function. 9. Now hash the authentificator2 variable using hash('sha256', authentificator2) function and store it in f2authentificator2 variable. 10. Finally run the SQL query to insert authentificator1, f2authentificator2, user_id and expiry_time to rememberme table. 11. STOP

3.8Database Connection and Population

The database is connected by providing certain details which include the database name, database password, database username. Once the connection is established, various MySQL operations are performed on need. Various select operations to fetch existing user and placements data. Also, various insert and update commands while managing

placement and user details.The Fig 6 below gives the complete flowchart on how our website works.

IV. Results and Discussion

The application is able to overcome few of the existing problems. Previously most of the colleges stores the student's details in excel sheets, for a student to prepare for placement activities they have to go to different websites for getting information and every time they need the internet for accessing it. The placement coordinators provide the details about the placement drives through WhatsApp or email. The students will be having so many queries in their mind for which they have to individually call the placement coordinators.

The proposed application is able to solve these difficulties. All the student's biodata will be stored in a database and can be viewed in the website in a user-friendly manner. The students can access the placement materials in one place and can download those and study them offline. The mentioned algorithms are able to predict the study material for the students and filter them based on the test scores which they can take for the web application. The chatbot is able solve all the doubts that the student is having and the students can practice coding with the help of the coding platform

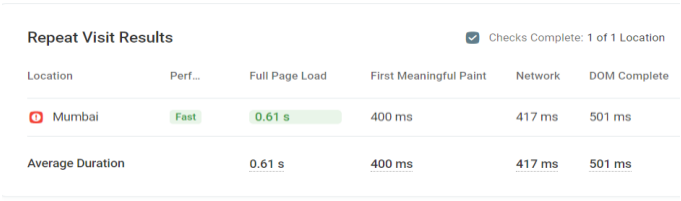

Fig 7. Throughput Analysis

Fig 6. depicts the website has a throughput of 0.81 seconds on first time load, and 0.61 seconds or repeat visit.

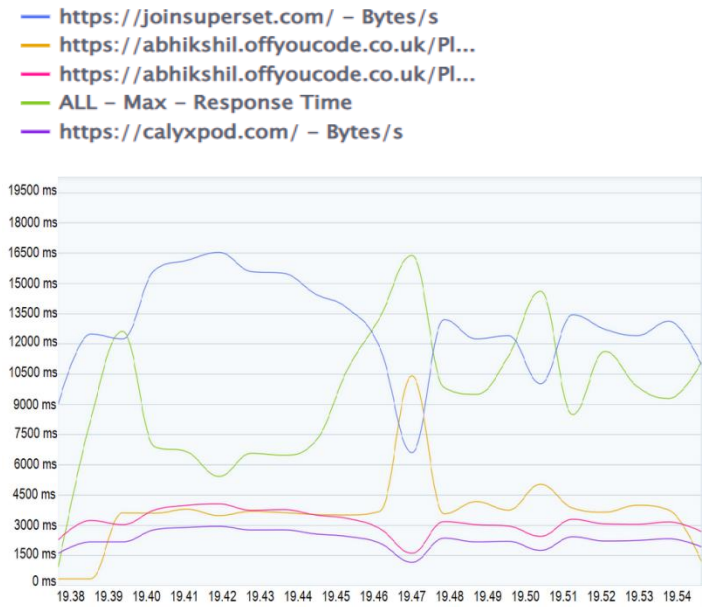

Fig 7. Throughput And Byte Analysis of Our application with clayxpod and superset

Comparison Table: -

Features	Placement Assistance Application	Superset	Calyxpod
Placement Support	Yes	Yes	Yes
University/College Management	Yes	Yes	Yes
Student Management	Yes	Yes	Yes
Resume Builder	Yes	Yes	Yes
Customized Placement Materials	Yes	No	No
Practice Coding Platform	Yes	No	No

Table 1: Feature Comparison Analyses

Fig 8. Placement Updates

V. Conclusion

The work was focused towards creating a web application that will be beneficiary to students, recruiters, and the coordinators of the placement department of a college in conducting the placement activity smoothly by organizing the student data and the details related to a job provided by the company in a well-structured manner. The application stores the details regarding login/signup, student's biodata, and job description in the database. The study materials support a student for preparation, the chatbot for user's help, and a coding platform for practicing programming questions. The algorithms used in the application will make predictions for the students regarding the study materials they need to refer based on their test scores. The recommendation system using Pearson correlation approach is able to provide the list of recommended placement materials in an efficient interval of time. The entire application is hence able to solve all the placement needs as it is a one stop solution and is providing a throughput of 0.61 seconds which is 32% faster as compared to the existing solutions for placement assistance.

VI. Acknowledgement

We would like to thank Reva University and our guide for providing us with the opportunity to research on the current placement scenario in India and build a dedicated assistance application to solve basic issues at campus level.

VII. References

[1] Aman Shakya, Shree Ram Rauniyar, Suman Khatri, Ujjwal Gewali, Vijay Yadav, "Smart Job Recruitment Automation: Bridging Industry and University", 2019 Artificial Intelligence for Transforming Business and Society DOI: 10.1109/AITB48515.2019.8947445

[2] Diaa Salama Abd Elminaam, Noha ElMasry, Youssef Talaat, Mohamed Adel, Ahmed Hisham, Kareem Atef, Abdelrahman Mohamed, Mohamed "HR-Chatbot: Designing and Building Effective Interview Chat-bots for Fake CV Detection", 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) DOI:10.1109/MIUCC52538.2021.9447638

[3] Ajay Shanker Singh, Thirunavukkarasu K, E. Rajesh, "Campus Placement Predictive Analysis using Machine Learning", 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) DOI:10.1109/ICACCCN51052.2020.9362836

[4] Abhishek S. Rao, Aruna Kumar S V, Pranav Jogi, Chinthan Bhat K, Kuladeep Kumar B, Prashanth Gouda, "Student Placement Prediction Model: A Data Mining Perspective for Outcome-Based Education System" 2019 International Journal of Recent Technology and Engineering DOI:10.35940/ijrte.C4710.098319

[5] Varun Shenoy, P. S. Aithal, "Green Placement – An Innovative Concept & Strategy in Campus Placement Model" IRA-International Journal of Technology & Engineering DOI: http://dx.doi.org/10.21013/jte.v4.n3.p3 ISSN 2455-4480; Vol.04, Issue 03 (2016)

[6] Soumya M.D, Thushara Sugathan, Kamal Bijlan, "Improve Student Placement using Job Competency Modeling and Personalized Feedback" 2017 ICACCI DOI: 10.1109/ICACCI.2017.8126097

[7] Fitra A. Bachtiar, Fajar Pradana, Rizkia Desi Yudiari, "Employee Recruitment Recommendation Using Profile Matching and Naive Bayes" 2019 International Conference on (SIET) DOI:10.1109/SIET48054.2019.8985988

[8] Aseel B. Kmail, Mohammed Maree, Mohammed Belkhatir, "MatchingSem: Online Recruitment System based on Multiple Semantic Resources" 2015(FSKD)DOI: 10.1109/FSKD.2015.7382376.

[9] Peng Yi, Cheng Yang,Chen Li, Yingya Zhang, "A Job Recommendation Method Optimized by Position Descriptions and Resume Information" DOI: 10.1109/IMCEC.2016.7867312

[10] L. Sheugh and S. H. Alizadeh, "A note on pearson correlation coefficient as a metric of similarity in recommender system," 2015 AI & Robotics (IRANOPEN), 2015, pp. 1-6, DOI: 10.1109/RIOS.2015.7270736.

[11] Dharaneeshwaran, S. Nithya, A. Srinivasan and M. Senthilkumar, "Calculating the user-item similarity using Pearson's and cosine correlation," 2017 International Conference on Trends in Electronics and Informatics (ICEI), 2017, pp. 1000-1004, DOI: 10.1109/ICOEI.2017.8300858.

[12] J. Benesty et al., Noise Reduction in Speech Processing, Springer Topics in

Signal Processing 2, DOI 10.1007/978-3-642-00296-0_5, © Springer-Verlag Berlin Heidelberg 2009

[13] Gowda N, Rekha K, "Implementation of cognitive approaches in question-answering system", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 5(10):2548-51, 2016.

[14] M. Zambrano et al., "Active Learning of Programming as a Complex Technology Applying Problem Solving, Programming Case Study and OnlineGDB Compiler," 2021(ICEIT), 2021, pp. 120-129, doi: 10.1109/ICEIT51700.2021.9375611.

Supreeth, S., & Patil, K. K. (2019). 'Virtual machine scheduling strategies in cloud computing—A review. *International Journal on Emerging Technologies*, *10*(03), 181-188.