**BABES-BOLYAI UNIVERSITY CLUJ-NAPOCA**

**FACULTY OF MATHEMATICS AND COMPUTER SCIENCE**

**SPECIALIZATION COMPUTER SCIENCE ENGLISH**

**DIPLOMA THESIS**

**Job Recruitment with Intelligent Recommender Engine**

**Supervisor**

Lecturer Dr. Lupsa Radu

*Author*

*Ciont Cristian-Lucian*

# **Abstract**

Contents

[**Abstract** 1](#_Toc71492659)

[**Introduction** 3](#_Toc71492660)

[**Chapter II** 5](#_Toc71492661)

[**The Domain of E-Recruitment** 5](#_Toc71492662)

[**2.1 Backgrounds** 5](#_Toc71492663)

[**2.2 Definition** 6](#_Toc71492664)

[**2.3 Advantages** 8](#_Toc71492665)

[**2.4 Disadvantages/Challenges** 9](#_Toc71492666)

[**2.5 Players in E-Recruitment** 10](#_Toc71492667)

[**2.6 The Hiring Process** 12](#_Toc71492668)

[**2.7 Traditional recruitment vs E-Recruitment** 13](#_Toc71492669)

[**2.8 E-Recruitment in the context of the application** 14](#_Toc71492670)

[**Chapter III** 15](#_Toc71492671)

[**Artificial Intelligence in E-Recruitment** 15](#_Toc71492672)

[**3.1 Artificial Intelligence and Machine Learning** 15](#_Toc71492673)

[**3.2 Uses of Artificial Intelligence in E-Recruitment** 18](#_Toc71492674)

[**3.3 Artificial Intelligence in the context of the application** 22](#_Toc71492675)

[**Chapter IV** 23](#_Toc71492676)

[**Application Description** 23](#_Toc71492677)

[**4.1 Requirements** 23](#_Toc71492678)

[**4.2 Design & Specification** 25](#_Toc71492679)

[**4.2.1 Application Flow Diagram** 25](#_Toc71492680)

[**4.2.2 Use Case Diagram** 25](#_Toc71492681)

[**4.2.3 Use case specification** 27](#_Toc71492682)

[**4.2.4 Database Diagram** 28](#_Toc71492683)

[**4.2.5 Class Diagram** 29](#_Toc71492684)

[**4.3 Technologies & Implementation** 30](#_Toc71492685)

[**4.4 Testing & Validation** 30](#_Toc71492686)

[**4.5 Execution** 30](#_Toc71492687)

[**Chapter V** 31](#_Toc71492688)

[**Conclusions** 31](#_Toc71492689)

[**Bibliography:** 32](#_Toc71492690)

**Chapter I**

# **Introduction**

E-recruitment, also known as online recruitment, represents the modern equivalent of finding various job candidates using web-based technology, in order to enhance its efficiency compared to the classic method of face-to-face/in-person encounters.

Nowadays, due to the exceptional and spontaneous growth of technology in the last decade, almost every company out there broke in the area of online recruitment. This doesn’t mean that the new approach needs to replace the old, but to work alongside each other to extend the growth and standards of a company.

Some of the most important reasons that propelled investors and entrepreneurs in taking another step towards the future of their businesses are the minimized hiring costs (less HR-human resources required during the process), broader scope of audience (enables targeting far wider, locally or abroad without paying extra taxes) and the accessibility (job ads and campaigns can be published through various social media sites, 3rd party websites, in order to reach the target group). It is also good practice to take into consideration that disadvantages also exist, such as a much larger number of unsuitable candidates applying to the offer versus the standard method, the lack of knowledge about the personality traits of the applicants (as these online elements tend to focus more on the technical side) and complications regarding the senior positions that are not very accustomed with technology.

The paper is centered around the concept of finding a job, both from the perspective of people on the lookout for new opportunities and from the perspective of a company wanting to recruit/expand. It is structured in 5 chapters that will be shortly described below.

The first chapter summarizes the main theme of the thesis and how it is structured.

The second chapter dwells in the domain of E-Recruitment, as the title suggests. It starts off with an introduction, followed by a couple of definitions. As recruitment and technology always evolve and change, a perfect definition that covers the entire concept doesn’t exist. Then, some advantages and challenges that commonly appear upon interacting with online environments are stated. In continuation, the players that are interacting with such system are presented (jobseeker, company, recruitment market and administrator) followed by how the hiring process takes place from beginning to end. The chapter ends with a brief comparison between e-recruitment and the traditional counterpart.

The third chapter enters the world of A.I. (Artificial Intelligence) and describes how intelligent agents can improve and shape the interaction with online recruitment in order to achieve more realistic results related to each individual’s personal preferences.

The fourth chapter breaks down the technologies used in the application.

The last chapter contains conclusions after completing the thesis.

# **Chapter II**

# **The Domain of E-Recruitment**

# **2.1 Backgrounds**

Recruitment is mainly composed of practices and activities carried out by the firms with the primary purposes of identifying/attracting potential employees, sorting through the applicants, contacting the potential fits and closing the deal with the best candidates. (Breaugh & Starke, 2000).

E-recruitment has grown exponentially in the last decade, as the Internet became the most popular tool in hiring and is now widely used by both recruiters and job seekers across the world (Cober & Brown, 2006). In order to draw human capital, the process is split in mainly two categories: corporate websites and commercial job boards. Corporate websites are private web pages that belong to companies where users can browse for current openings, while job boards are third party websites that contain offers from multiple recruiters that paid a certain fee for publicity. (Kapse & Patil, 2012). Lately, social networks are increasingly being used because of their efficiency as a middle agent between the other two. (involving a free third-party website where each company can create a personal page with all the information needing to attract human resources).

Technology can ease the selection of employees over the traditional method, and so every website evolved into a complex interactive engine that tries to automate every aspect of cyber recruitment, making it accessible to a wider audience over longer distance. ( Galanaki, 2005; Joe Dysart, 2006). The world wide web became a powerful tool for recruiting in the mid-1990s and was advertised by the media as a handler behind a “recruiting revolution” caused by all the potential improvements that it could bring to the table over the more traditional method. Easier access to information and reduced costs in human resources resulted in more sophisticated and interactive methods of finding the right person for a specific opening. (Boydell,2002; Freeman & Autor, 2002).

When we are talking about what are the highest chances of finding the most suitable candidates for a job opening, the flexibilities and capacities of e-recruitment are standing out. It can offer a global geographical spread with unlimited access at any time. People from all over the world can find attractive job offers, greatly increasing the reputation and brand of company. Also, it can be a time saver, as more and more automated tools and engines are cutting out the work for employers. Nevertheless, everything has some flaws. In this particular case, the lack of internet in some undeveloped countries, the low technical knowledge of senior people that prevent a best experience of an online environment and the insufficiency of accessible website to allow people with disabilities equal chances.

Given all that, as technology always evolves and changes, the cracks that come with e-recruitment have a full potential to be fixed and all the domain to be improved.

# **2.2 Definition**

Recruitment is a process that implies looking for potential candidates to fill available vacancies in organizations. With the help of effective human capital work and following a linear process composed of multiple steps, the services of the most suitable and capable personnel are being secured. The rapid growth of technology suddenly changed the way companies are conducting business and lured almost all of them to more non-traditional forms of enrolling practices, precisely “e-recruitment”.(Erica, 2007).

Vidot (2000) suggested that “e-recruitment is the use of internet to attract high quality candidates, screening of suitable profiles, streamlining the application and selection process.”

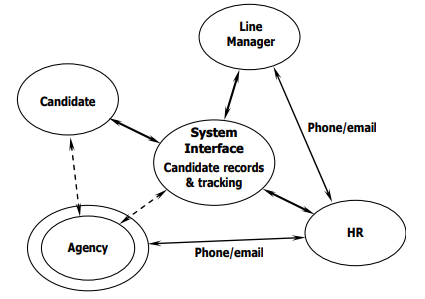


Figure 1: Landscape of e-recruitment (Source: IES, 2003).

Galanaki (2002) said that “The words e-recruitment, online recruitment, cyber-recruiting, or internet recruiting are synonymous. They imply formal sourcing of jobs online.”

Edwin B. Flippo stated that “Recruitment is the process of searching the candidates for employment and stimulating them to apply for jobs in the organization”, meaning that recruitment stands as a link between employers and job seekers, as **Figure 1** shows. The candidate interacts with the System Interface, a website in our case, and the application information is processed by the HR (Human Resources) department, either directly or through phone/email, then finally reaching the organization.

E-recruitment can also be defined as “a way of implementing [recruitment] strategies, policies, and practices in organizations through a conscious and directed support of and/or with the full use of web-based channels” (Girard & Fallery, 2010, p. 1). This statement affirms that in order for a recruitment process to be categorized as e-recruitment, it must have a “significant portion” taking place online, as rejecting candidates through emails and chatting applications or inviting them to take a quiz/survey by providing a link doesn’t mean you are not using a variant of traditional methods.

To put it shortly, web-based recruitment means the use of technology in supplementing and improving the classic method, without trying to eliminate it. To show how dominant this trend became, Taleo conducted a research in 2003, finding out that over 93% of the Global (that operate and distribute in many countries around the world) 500 companies were already using online recruitment. Another investigation, also by Taleo, showed that a large number of firms are only allowing the job seekers to apply in an online environment (from 27% in 2000 to 77% in 2005).

# **2.3 Advantages**

More and more companies started to invest in their own websites or using 3rd party services because of the various benefits the online environment could bring to their businesses. The most important feature would be the availability, as the Internet provides unlimited and continuous access to the possible future employees at any time and almost anywhere and by doing so it broadens the timetables that would be provided by the HR (Human Resources) workers.

Another important facility are the reduced costs that come hand in hand with going online. First of all, posting a new job vacancy on your own website is free, relatively cheap on an external service compared to a recruitment agency or getting a headline in a newspaper. Second of all, as processing applications online means that an automated algorithm/system is doing all the work for you, this implies that the administration burden of the HR department can be greatly reduced as well. (Kerrin M 2005).

Wider access results in a higher chance of finding the most suitable candidates in a fair and efficient way. Online advertising provides a larger pool of local, national, and international offers. Having people from all around the world encourages diversity, different opinions, and various knowledge on a common subject. (Kerrin M 2005). By using an online recruiting tool, information about all the applications is stored in a database and can be used later for future openings and thus giving a second chance to the people that did not manage to close the deal in the first place.

An additional very important feature is time saving. Completing online quizzes, surveys, forms and attaching your CVs (Curriculum vitae) can now be processed by engines in a matter of minutes, saving valuable time. This is particularly useful in larger companies where some vacancies can attract a large volume of job seekers. All the steps from preparing the job description and specification to the pre-selection of candidates to the interviews and the decision making can be tailored to fit the best interests of both the user and the organization. Receiving quick responses to questions and information on certain subjects are easily achieved, as a lot of companies often have forums where anyone can post their concerns and get an answer from a qualified worker. (Elkington T 2005; IDS Study Plus 2000).

The reputation and brand can also be a deciding factor when done properly. Lots of websites have sections where they describe the organizations accomplishments, the most impressive and large projects, the most skilled and renown employees that walked through their door, the culture and maybe the future goals. Lately, it is common for “virtual tours”, “employee of the month” and “a day in the life of” videos to be presented in order to immerse in the feeling of how it would be like to work there and how you would be treated by others. Having more features and putting work to describe your firm can greatly increase how an outsider is viewing and formulating his first opinion about you, resulting in a better image. (Kerrin M 2005).

Some other smaller advantages of e-recruitment are the attraction of passive jobseekers, as having a very attractive job post can stimulate even someone who is not necessarily on the hunt to take a peek, the reduction of unqualified candidates, as the offers can be very precisely tailored to very specific needs, more opportunities for smaller firms, as any application can become popular overnight due to recommendation systems and it can often be less intrusive than the traditional method, as some people are more comfortable in front of a screen rather than in front of a committee.

# **2.4 Disadvantages/Challenges**

As any other domain, E-recruitment has its own share of challenges and shortcomings, but they are not written in stone and can always take a turn for the better.

The main problem is related to the number of applications and their authenticity that need to be processed. Even though an online platform generates a high amount of suitable candidates and also reduces the number of unsuitable candidates, their number can still be quite large, especially for a popular job post. However, automation can only take you so far and afterwards it turns into a time costly operation of sorting out unwanted offers. Another difficulty is keeping a clean record of both online and offline offers, which can lead to a chaotic internal organization and can affect the state of the company. In some cases, overloading of the website’s capacity due to high traffic can create some technical, IT-related problems, which again need time and resources to solve. (IRS 2005).

A really important issue is the lack of Internet in a lot of underdeveloped countries around the world, which affects the chances of their inhabitants to get an equal chance of obtaining a job as someone in a more thriving place. Also, companies need to take into consideration the population that is not accustomed to technology and usually encounter technical difficulties when interacting with an online environment, thus firms shouldn’t solely depend on online recruitment methods. Various disabilities can often limit or even negate the use of a website and a lot of employers are unaware of this implication and are not providing an accessible environment or an alternative medium depending on the situation. (IRS 2005).

Smaller complications also exist, like the preference of individuals for face-to-face traditional recruitment in the case of a company where only online applications are allowed and the absence of “personal touch” of a website. To elaborate, the online platforms provide a more technical resume of a candidate, often neglecting their personality, temper and emotions thus obtaining a robotic environment that cannot learn how a possible future employee will behave.

# **2.5 Players in E-Recruitment**

* Job Seeker: An individual that is looking for new opportunities. On a website, they have permissions to create an account, participate in various quizzes and tests organized by companies, browse the available vacancies in the job posts and of course to submit their CV/resume to any organization they see fit. They can be split in two categories:

a) Active job seekers: Candidates that are searching regularly for jobs that can offer them a change of scenery in another domain, higher salaries, better opportunities for growth and many other personal or professional reasons.

b) Passive job seekers: Candidates that are not deliberately in need of a change, but they are browsing the internet in their spare time. A well-made and attractive job post can definitely catch their eye, so a new opportunity, even if not intentional, can always appear.

Diagram

Description automatically generated

Figure 2: The components of e-recruitment (Source: [4])

* *Administrator* (Optional): The administrator is an optional candidate here, as he only appears in the case of 3rd party software like commercial job boards. He owns a website that holds multiple offers from multiple companies, which need to pay a fee in order to advertise their vacancies. The administrator has the ability to remove both users and companies from his platform and to view several reports containing information from clients.
* Employers (Companies/Firms): The final target of the job seeker where they can close the deal based on the job requirements and their experience and skill. On a website, companies have the ability to post new job vacancies, modify the details and to process using their own policies, all the data received from applicants.
* Recruitment Market: The recruitment market is the mediator between all the components that is responsible for keeping everything together. It is the main piece where e-recruitment revolves around. During time the online environment took different shapes like corporate websites, commercial job boards/portals, social-media sites, games, blogs and lately a lot of them have been upgraded to sustain a recommender engine that allows for more precise and efficient results fit for every user’s personal preferences.

As **Figure 2** shows, all the information the recruitment market collects from jobseekers, companies, and administrators, depending on the situation, is stored in a database so that it is safe, validated, current, organized and available at any time.

# **2.6 The Hiring Process**

The process of hiring is composed of multiple consecutive tasks. It usually begins with identifying applicants in a specific domain, making an interesting eye-catching job announcement and continued by assessing the incoming applications to communicating with the most suitable candidates in an interview.

* Identification: Identification requires the least amount of Internet usage, as seen in **Table 1**, because it mainly consists of creating the wanted candidate profiles with the appropriate qualifications. A suitable job description for the chosen segment of the working market should be created and then distributed is some type of online environment. (Newell 2009)
* Attraction: Attraction is the most important and Internet dependent step towards the expansion and growth of a company. The brand and reputation of a firm tends to be a very relevant factor for driving jobseekers to their doors. Corporate website, social media sites and job boards are the most popular and efficient software tools to reach the target audience. The principal ways of submitting your information are through CV’s resumes and application forms. (IRS 2005). Lately, a lot of website are using verification engines to assure the authenticity of the incoming applications.

|  |  |
| --- | --- |
| Recruitment phase | E-recruitment techniques, tools and sources |
| Identification of candidates | None |
| Attracting applicants | E-recruitment system, Internet job boards and career portals, corporate career website, social networking sites, various (third-party) websites, industry-specific web resources, electronic mail and electronic mailing lists |
| Processing and pre-screening applications | E-recruitment system |
| Communicating with candidates | E-recruitment system, corporate career website, electronic mail and electronic mailing lists |

Table 1: The use of e-recruitment sources, techniques, and tools in the case companies at different recruitment phases (Source: [2])

* Selection: Selection implies reducing the large amount of received applications by eliminating the unsuitable candidates for each vacancy. This can be achieved in the form of technical tests, evaluations, quizzes or simply by removing candidates whose qualifications or studies/degrees are not matching the given requirements. Because of issues of security, confidentiality and not knowing if the right individual is taking the assessment, a lot of companies are hesitant to use online testing tools and prefer the more physical/in-person way. (IRS 2005). Having an automated engine to perform the processing of candidates is essential in ensuring an organized and valid system for storing data for further refinement.
* Communication: Communication is considered a critical part of the hiring process because it determines whether you get dropped off or proceed further towards an interview. There is risk of receiving negative feedback from candidates whose resumes were rejected too early and it can appear as if the process was rushed without enough consideration and this can eventually lead to damaging the firm’s reputation. However, it can be easily avoided by giving quick and on point feedback to the applicants, explaining what they got wrong and how can it be improved for future job interactions. Communication can be achieved by email, phone calls or even face-to-face, where the distance is not a problem.

# **2.7 Traditional recruitment vs E-Recruitment**

Traditional or paper-based recruitment revolves around the exact same concept as e-recruitment: finding the right person to fill an available vacancy. The main difference between the two comes at how the hiring process takes place. As **Figure 3** shows, the steps of both concepts are almost the same. However, online recruitment comes with an advantage at selection and communication with the applicants, as it allows these processes to happen concurrently (at the same time). Technology made it possible, by always allowing access to candidates’ information, for HR department and the employees to keep the users in touch with their progress during the whole recruiting process. Running an online environment comes with the responsibility of upkeeping and constantly improving it, so it can provide a pleasant and different experience to the user compared to the classical method.

Diagram

Description automatically generated

Figure 3: The design and sequence of tasks in traditional paper-based recruitment process vs. the (new) recruitment process using e-recruitment. (Source: [2])

The traditional method involved publicity by newspapers, placement agencies, job fairs, employee referrals, educational institutes, so its geographical spread was greatly reduced compared to the online method, implying a smaller number of candidates and a higher cost for the administration and processing required by the human resources departments.

As a conclusion, paper-based recruitment should never be replaced, it should be supplemented. The flaws of online recruitment can be patched by traditional methods and the volume, speed will be greater, costs lower, due to modern technologies.

# **2.8 E-Recruitment in the context of the application**

The application can be categorized as a commercial job board, providing an online environment consisting of multiple job offers across several companies. The program itself is acting as a recruitment market and allows interaction from both jobseekers and organizations.

In the perspective of the hiring process, it revolves around the attraction area. Candidates can browse and apply to a wide variety of job vacancies advertised by different firms. A small portion of the selection process is also present in the application, as the companies can visualize the applicants for each individual offer, but it does not provide any automation tool to allow easy elimination of the unsuitable candidates.

# **Chapter III**

# **Artificial Intelligence in E-Recruitment**

# **3.1 Artificial Intelligence and Machine Learning**

Artificial Intelligence or A.I. is a domain of computer science that aims at providing machines the ability to mimic human behaviour and perform complex tasks like learning, decision-making, problem-solving and planning. Even though A.I. is centered around simulating human intelligence, it is also used widely in areas that involve more computation than people can handle.[8]

John McCarthy stated that “It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”

There are countless real-word applications nowadays that are benefiting from A.I., the most common being:

* Computer Vision: Is a technology that enable machines to obtain and process information from digital photos, videos and various visual data. Applications can vary from image recognition to radiology imaging in healthcare to self-driving cars that are aware of the surrounding environment.[14]
* Speech Recognition: Also referred as speech-to-text, is a technology that uses natural language processing (NLP) to transform human speech into a written format understood by a computer. It is particularly used in systems revolving voice search, as Siri, Alexa, Google, and to provide more accessibility for texting.[14]
* Customer Service: Customer services provide 24/7 assistance by replacing human resources with virtual assistants, also known as chatbots. They are specifically trained to answer the FAQ (Frequently Asked Questions) and to provide suggestions and feedback for users around a specific media platform. [14]

As the ultimate goal of A.I. is to reach and possibly surpass the human intelligence in machines, this can only be achieved through learning algorithms that know how the human brain behaves. Thus, a very popular and widely used branch of A.I. emerged, called Machine Learning (ML). Machine learning relies on working with datasets by examining and comparing the data to find common patterns and to improve over time, without being explicitly programmed to do so.[9]

Tom M. Mitchell said that “Machine learning is the study of computer algorithms that allow computer programs to automatically improve through experience.”. A more formal definition, also by him, affirmed that “a computer program is said to learn from experience (E) with respect to some task (T) and some performance measure (P), if its performance on T, as measured by P, improves with experience E then the program is called a machine learning program.”.[9]

Machine learning requires the human intervention for providing starting data and gives a response about what is most likely to happen when some new data is introduced. This process can be done in several ways, as shown in **Figure 4**:

A picture containing text, businesscard, screenshot

Description automatically generated

Figure 4: The three ways a machine can learn (Source: [12])

* Supervised learning: Supervised learning is the most commonly used technique of learning as it requires labeled data, meaning that each input data has an answer or result already given, usually by humans. This can help the machine to reduce the number of expected output results and thus to improve the overall correctness of the program. In order to achieve the best prediction, this technique compares the computed result with the available labels and learns by adjusting an error to get as close to the expected output as possible. Depending of what you want to predict, supervised learning can be used to solve regression or classification problems. Classification problems have a limited number of label values, usually only two, meaning that it responds to a random question like “Is this an animal” with a binary response (for example 1 for “yes” and 0 for “no”). Regression problems are used with continuous label values that have no specific value limits and they can be used for example to predict the height of a man or the price of a house.[9][12]
* Unsupervised learning: Unsupervised learning uses unlabeled data and relies on the machines learning on their own to detect patterns and group them. Depending on what you are trying to group, there are mainly to ways to achieve this: clustering or association. Clustering is looking for similarities in the data and grouping the together accordingly. Association tries to find the connections between the formed groups and to understand the rules and meaning behind why they were grouped in that particular way.[9][12]
* Reinforcement learning: Reinforcement learning is usually a long-term process that continuously learns based on a reward/penalty system, meaning that the machine is rewarded when it gives a correct output and is penalized when it gives an erroneous one.[9][12]

# **3.2 Uses of Artificial Intelligence in E-Recruitment**

After an in-depth comparison presented in chapter two between the traditional way of recruiting and E-Recruitment, it was clear that the last vastly expanded the scope of finding new talent by operating in a more widely geographical space and thus bringing a much larger volume of candidates. Of course, having a larger pool of applicants also meant needing more human resources and a lot of time required for processing each individual offer. Thus, the need of artificial intelligence and machine learning to speed up the entire process and enhance the user experience has increased or even became mandatory for an efficient management.

Automation has been included in various areas of the E-Recruitment domain. The first such area is job advertisement. Especially in the case of big companies, this process can become very time consuming and tedious to manually create offers for all the available openings, many times not even guarantying optimal results from applicants. As such, machine learning algorithms have been developed that are able to use pattern recognition on data that the hiring person has provided to find out why some job adverts are popular, and others are not. After gathering enough information, they can give you ways of remodeling your job offers so that it can attract a broader group of candidates. [15]

Another popular area of automation is related to candidate sourcing and selection. If a company does well in promoting and advertising the vacant openings, all the incoming applications will need to be processed and analyzed. Due to the fact that manual CV screening is no longer an option for the human resources departments in case of large amount of candidates, machine learning can be used to significantly reduce time and resources, without exterior intervention. Such programs can use NLP (Natural language processing) in order to identify the keywords in the resumes that correlate with the necessary skills and experience required for the job, providing you with the top candidates in a matter of seconds.[15]

Talent engagement and relationship management is extremely important in keeping clients interested in your company. As a result of websites being available internationally, candidates can browse at any given time, with high chances of being outside of the timetable of the employees. To avoid such situations, machine learning can be used to create voice assistants or chatbots that are available all the time. They can be trained to answer the frequently asked questions and to keep to users informed during the recruitment process, becoming more efficient as they receive more data about everyone’s personal preferences. Voice assistants are particularly useful in case of persons with disabilities that are unable to type, by providing the same services using speech-to-text technologies.[15]

Lately the most popular machine learning automation, particularly where huge amounts of data circles every day, are the recommendation engines. They are systems that suggest products or information to users based on factors such as the characteristics of an item, the interests of the user or the behaviour of similar users. From the perspective of a company, a well-defined engine can significantly boost their revenues by providing a higher customer satisfaction. Instead of browsing through hundreds of job offers, a recommender system can narrow down the selection to a smaller amount of applications that the user is most likely to enjoy.

Diagram

Description automatically generated

Figure 5: The recommendation process (Source: [10])

These systems need to learn so that they can provide valid responses. Therefore, they take information from your past browsing activities, such as the ratings you gave a job opening, whether you liked it or not, what kind of reviews does it have, personal information about you (age, gender, your experience in the industry, what are you interested in, etc.). As **Figure 5** shows, the engine takes data from the specific items you are looking into, in our case the features of a job opening and combines them with the feedback taken from other users in order to create a prediction for each job that is measured by the usefulness of a certain job to a certain user. As a result, the system usually gives a ranking with the highest estimated rating entities that the current user has not discovered yet.[10][11]

As any other machine learning algorithms, the more information you provide the recommendation engines, the more effective they become in giving you accurate suggestions. Therefore, data collection is the most critical part of this process. In the domain of E-Recruitment the available information needs to be gained in a bidirectional way, considering the preferences not only of the recruiter but also of the jobseeker. The obtained insight can be:

* Personal information from your profile, such as age, gender, location.
* Information about your educational background: the school/universities you attended, bachelor or master’s degrees obtained, the fields of education.
* Insight about your past and current professional background: companies you worked for, positions you occupied, your experience, what technologies you know best, if you like to work full-time or part-time, if you prefer to work from home or from the company headquarters, etc.
* The features of the job offers: minimum experience required, job type, developer type, necessary levels of foreign languages, salary, etc.
* The actions you and other users performed on the website: the ratings given to jobs, the act performed on a job offer (applying, liking, disliking, sending an email with your CV), the number of revisits on each job opening, the read time on job descriptions, available reviews of the job post and so on.[7][13]

Based on what information the recommender system receives and processes, there can be distinguished three different techniques for giving predictions: content-based filtering, collaborative filtering, and hybrid systems.

Content-based filtering is based on the metadata collected from a single user’s interactions. The given suggestions will depend solely on their choices, preferences, and background. This approach will become more efficient and accurate as the user provides more inputs and makes actions. Combining the received feedback with the features of the items (jobs in our case), the system builds a user profile and according to it recommends the most suitable job applications. Since the returning information is only based on one individual, content-based filtering is considered to be a safe technique that can’t exceed the boundaries of the user’s comfort zone.[5][10]

Collaborative filtering is a technique that expands its scope to the other users’ interactions and history. It exploits similarities between your actions and theirs to derive the most accurate suggestions for you. Because it collects taste information from many users to find what you will most likely enjoy, it is considered more diverse than content-based filtering. This brings high chances of receiving suggestions that are different from your ordinary searches. New opportunities can always arise but so can recommendations that you are not inclined towards, thus making it a more aggressive approach. [10][11]

Hybrid systems, as their name suggest, are trying to combine features with good synergy from the above two techniques to obtain a more enhanced system with better performance and less disadvantages.[10]

Since it can be very tedious to work with textual data, all the information that is passed to the recommender system needs to be vectorized so that it can be processed in an efficient way, as shown in the formula below. For instance, if you have the features “Is this job full-time?”, “Is java required?” and “Rating”, a possible such vector would look like [1, 0, 4.5], meaning that the job is full-time, java is not a technology that is required and it has a user rating of 4.5 points. The resulting vectors can be easily used with similarity algorithms (Euclidean distance, cosine similarity, Jaccard measure) to assess the best jobs tailored to the users’ interactions and preferences.[13]

Feature

Equation 1: Vectorization of a feature (Source: [13])

It is also worth mentioning that recommendation systems come with challenges and flaws, like any other algorithms. The cold start problem is a very common issue when it comes to new users or items, as they miss interaction and feedback. This affect mostly the collaborative filtering systems because they are dependent on both sides and cannot make useful suggestions in case of new items or users, meanwhile content-based systems can manage only with the characteristics of an item as they are not solely dependent on previous interactions. Another common issue is scalability which means accommodating the system with the increasing amount of data in a graceful and subtle manner. The constant increasing in popularity of a website can provide an overwhelming amount of information that could lead to slower systems or inaccurate results. High volume of data can also lead to the sparsity problem, which affects mainly collaborative filtering systems, since they are hinging on the feedback of other users. In a large pool of information that is continuously growing, the available ratings become scarcer and scarcer as users tend to not rate most of the accessed items. For content-based filtering recommendation engines, the over specialization problem can occur, as these systems tend to give similar results to what the users already know and defined in their profiles. This could prevent them from discovering new items and discourages diversity.[10]

# **3.3 Artificial Intelligence in the context of the application**

The application uses a dataset consisting of various job openings posted in 2019 on an Indian website called Naukri. It features jobs from multiple domains and industries, which have been filtered to the ones related to IT, as the program is intended only for that domain. The dataset contains information such as the title of the job, the minimum experience required, the technologies that you are expected to know, what kind of position is available, etc.

The application features two recommendation engines, both situated in the category of content-based filtering.

The first one uses purely an item-based approach, meaning is takes information only about the features of the job offers. It provides the user with similar suggestions to the job he is currently browsing, and it could be called a “More like this” approach.

The second one takes full advantage of the content-based systems technique, as it considers the feedback provided by the current user and his/her personal background. When a new account is made, the user is asked various information about his preferences (favorite technologies, his experience, job type) and after the account is registered, they can like and dislike any job opening. This metadata is processed to form the user’s profile which is used in contingency with the information collected from the available jobs to give relevant suggestions tailored to personal interests. The more feedback the user provides, the more accurate the results will be.

## 

# **Chapter IV**

# **Application Description**

# **4.1 Requirements**

The application represents an online environment that allows jobseekers to easily browse through various offers to find a suitable job and companies to recruit in an organized and efficient manner. It also incorporates an intelligent system to further enhance the user experience by providing similar job opportunities to his preferences. Thus, there are mainly two types of users, clients/jobseekers, and organizations/companies. In addition, there is a third type of user, the administrator who has full privileges over the entire software.

As a client you can benefit from the following functionalities:

* Visualize/browse available job offers.
* Search/Filter jobs on various criteria.
* Like/Dislike an offer.
* Apply to jobs.
* Visualize/browse similar jobs to a specific one.
* Login / Change profile settings.
* Get recommendations based on your personal background and the ratings you give.
* Visualize the distance in kilometers between you and the headquarters of the organization the current job opening belongs to (if both you and the company have marked their locations on the map).

As a company you are able to:

* Login / Change profile settings.
* Enlist your organization / Modify organization data.
* Post job openings / Manage the details of your own offers / Remove applications.
* Visualize candidates and their information that applied to your openings.

As an administrator you have the power to:

* Remove any applications / organizations / accounts.

As an unregistered user (no account) you can do the following:

* Visualize/browse available job offers.
* Visualize/browse similar jobs to a specific one.
* Register account.
* Search/Filter jobs on various criteria.
* Visualize various charts/graphs with statistics.

# **4.2 Design & Specification**

## **4.2.1 Application Flow Diagram**

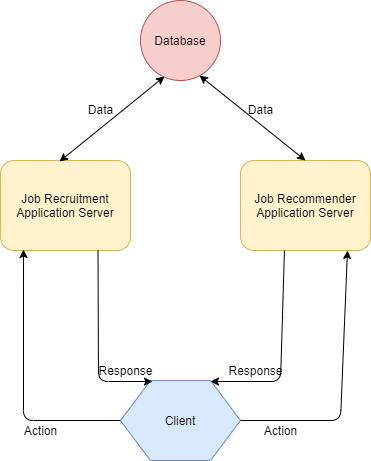


Figure 6: Application flow schematic

As Figure 6 shows, the application is structured in four main components. All the actions that are performed on the website are sent to one of the two servers, depending on the context and all the received data is persisted in a shared database amongst the two.

## **4.2.2 Use Case Diagram**

Diagram, schematic

Description automatically generated

Figure 7: Use case diagram

## **4.2.3 Use case specification**

From a more technical perspective, the software requirements of the application can be translated to specific use cases for a better understating of the flow of events. Below are detailed the most important functionalities of the program:

* **Use case: View similar jobs.**

Description: The user can browse the top offers that are the most related to the one he is currently visualizing.

Input: The characteristics of a job.

Pre-conditions: The user does not necessarily need to be logged in for this functionality. It will use a content-based filtering approach that will process only information related to the features of the available jobs.

Post-conditions: The top ranking of the jobs that are similar to the one received as input will be presented to the user.

Data handling: All the data required will be retrieved from a database.

* **Use case: View recommended jobs.**

Description: The user will receive appropriate job suggestions to their personal preferences provided in the profile background and to the past interactions with the offers.

Input: The identifier of a user.

Pre-conditions: It is mandatory for the user to have an account and to be logged in. When a new client account is created, they will be prompted to introduce information about their educational background (what programming languages they prefer, what position they like to occupy, their coding experience, etc.). The clients will also be able to like or dislike a job offer. The content-based filtering engine will combine the information in the user profile background to the features of the jobs to return suggestions appropriate to the users’ interests. Also, the feedback received from interacting with the offers (likes or dislikes) will also be taken into consideration to increase the accuracy of the recommendations. This functionality will be disabled in case of company or administrator accounts.

Post-conditions: The top ranking of the unseen job offers that are best matched to the users’ actions and preferences will be shown.

Data handling: All the data required will be retrieved from a database.

* **Use case: View offers.**

Description: The companies can visualize a list with their available offers and details about the candidates that applied to them.

Input: The identifier of a company.

Pre-conditions: The user needs to be logged in with an account with organization permissions and to have the company already registered. All the job openings posted by the firm will appear with the possibility of showing personal information about the candidates that applied to any individual offer.

Post-condition: A list with the jobs and the corresponding user information.

Data handling: All the data required will be retrieved from a database.

* **Use case: Register company.**

Description: Users with company credentials can register information about their company so they can start posting job offers.

Input: The characteristics of a company.

Pre-conditions: The user needs to be logged in with an account with organization permissions. The features of the company will be validated prior to sending the request. Any firm is obligated to submit their logo (in .png format) and the approximate number of employees. Optionally, they can add a marker on the map with the location of the headquarters.

Post-conditions: Not required.

Data handling: The resulting company data will be persisted to a database.

* **Use case: Browse job offers.**

Description: The user can navigate through all the available job offers with the possibility of filtering and sorting the data by various criteria.

Input: Not required.

Pre-conditions: It is not necessary to be logged in. The received data will be displayed using pagination, allowing the user to go back and forth between pages and to select how many entities can one page hold. Furthermore, the jobs can be filtered by any available feature and they can be sorted alphabetically, by salary or by the most recent ones. In case of the filtering feature, the use of pagination is no longer mandatory due to small amount of data.

Post-conditions: All the data that matched the given criteria.

Data handling: All the data required will be retrieved from a database.

## **4.2.4 Database Diagram**

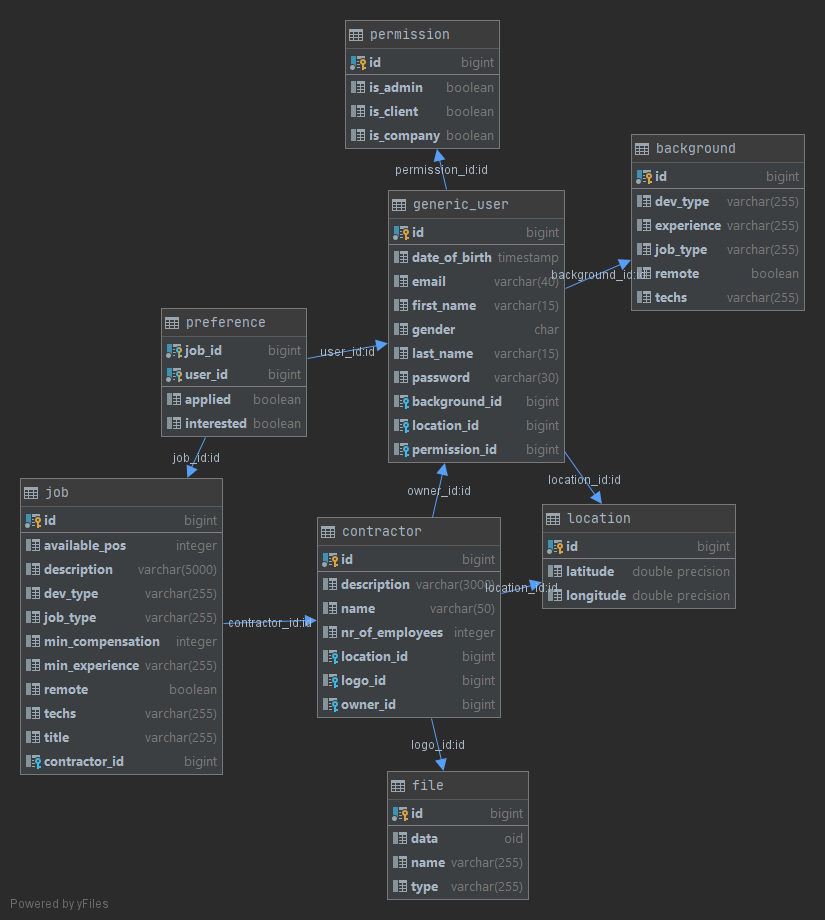


Figure 8: Database diagram

## **4.2.5 Class Diagram**

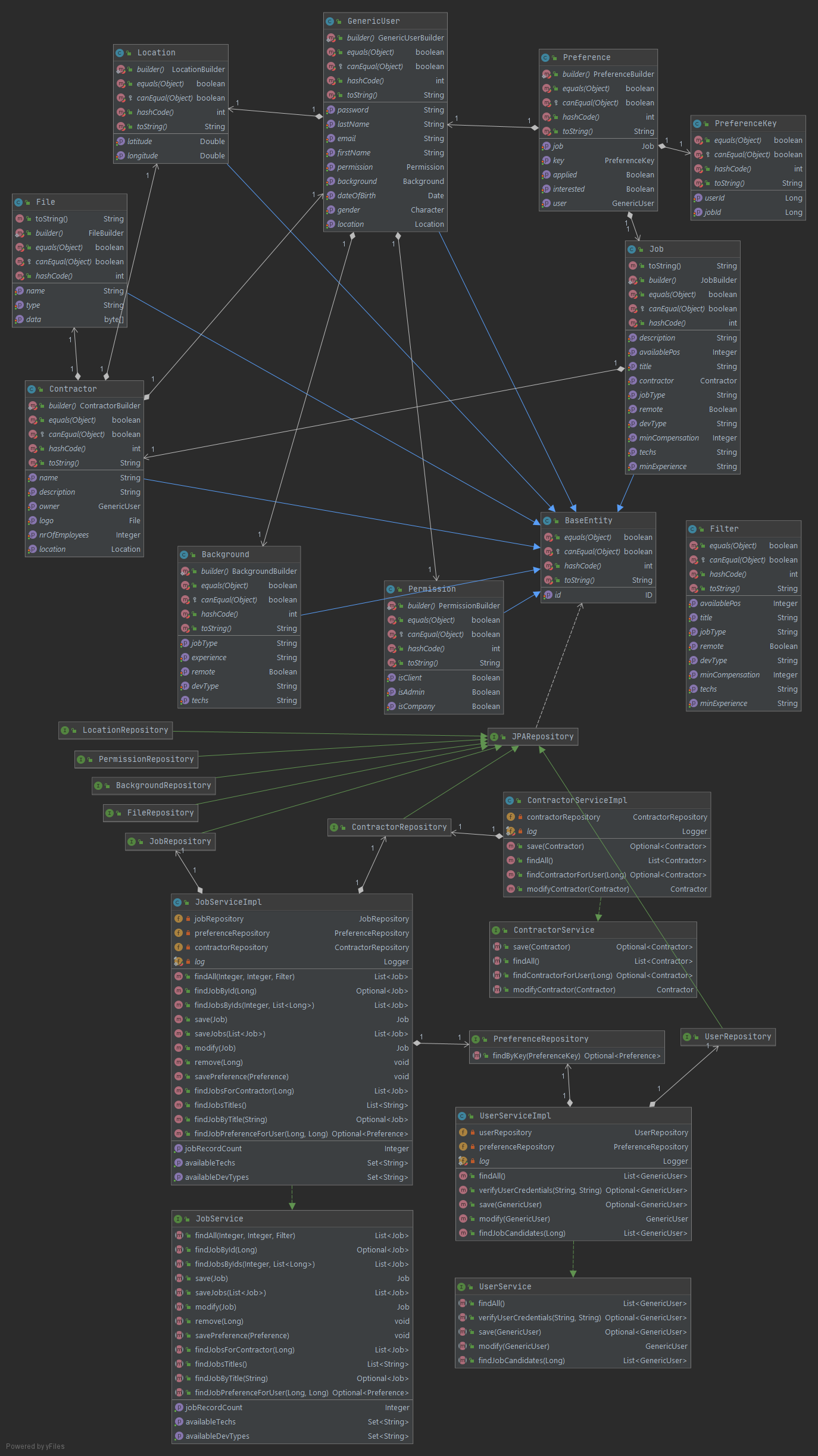


Figure 9: Spring server class diagram

# **4.3 Technologies & Implementation**

# **4.4 Testing & Validation**

# **4.5 Execution**

# **Chapter V**

# **Conclusions**

# **Bibliography:**

[1] Dhamija, Pavitra. "E-recruitment: a roadmap towards e-human resource management." *Researchers World* 3.3 (2012): 33.

[2] Holm, Anna B. "E-recruitment: towards an ubiquitous recruitment process and candidate relationship management." *German Journal of Human Resource Management* 26.3 (2012): 241-259.

[3] Barber, Linda. "E-recruitment Developments.", Technical Report (2006), https://www.employment-studies.co.uk/system/files/resources/files/mp63.pdf.

[4] Kapse, Avinash S., Vishal S. Patil, and Nickil V. Patil. "E-recruitment." *International Journal of Engineering and Advanced Technology* 1.4 (2012): 82-86.

[5] SHUVAYAN, DAS. “Beginners Guide to learn about Content Based Recommender Engines”. <https://www.analyticsvidhya.com/blog/2015/08/beginners-guide-learn-content-based-recommender-systems/>. August 11, 2015.

[6] Woźniak, Jacek. "The use of gamification at different levels of e-recruitment." *Management Dynamics in the Knowledge Economy* 3.2 (2015): 257-278.

[7] Rafter, Rachael, Keith Bradley, and Barry Smyth. "Automated collaborative filtering applications for online recruitment services." *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer, Berlin, Heidelberg, 2000.

[8] McCarthy, John. "What is artificial intelligence." (2007): 2020.

[9] Das, Sumit, et al. "Applications of artificial intelligence in machine learning: review and prospect." *International Journal of Computer Applications* 115.9 (2015).

[10] Sharma, Lalita, and Anju Gera. "A survey of recommendation system: Research challenges." *International Journal of Engineering Trends and Technology (IJETT)* 4.5 (2013): 1989-1992.

[11] Roger Chua, “A simple way to explain the Recommendation Engine in AI”. <https://medium.com/voice-tech-podcast/a-simple-way-to-explain-the-recommendation-engine-in-ai-d1a609f59d97>, Jun 26, 2019.

[12] Randy Lao, “A Beginner’s Guide to Machine Learning”, <https://medium.com/@randylaosat/a-beginners-guide-to-machine-learning-dfadc19f6caf>, Jan 22, 2018.

[13] Al-Otaibi, Shaha T., and Mourad Ykhlef. "Job recommendation systems for enhancing e-recruitment process." *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2012.

[14] IBM Cloud Education, “Artificial Intelligence (AI)”, <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence#toc-types-of-a-q56lfpGa>, 3 June 2020.

[15] Neelie Verlinden, “Machine Learning In Recruitment & How To Do It Right”, <https://harver.com/blog/machine-learning-in-recruitment/>, October 23, 2018.