



CurEval - Curriculum Evaluation

LILIANA CRISTINA DE LEMOS NOVAIS

julho de 2023

CurEval - Curriculum Evaluation

Liliana Novais

**A dissertation submitted in partial fulfillment of
the requirements for the degree of Master of Science,
Specialisation Area of Information and Knowledge Systems**

Supervisor: Dr. Constantino Martins
Co-Supervisor: Vladimiro Ferreira

Evaluation Committee:

President:

Dr. João Paulo Pereira, Professor, DEI/ISEP

Members:

Dr. Ana Almeida, Professor, DEI/ISEP

Porto, July 26, 2023

Statement of Integrity

I hereby declare having conducted this academic work with integrity.

I have not plagiarised or applied any form of undue use of information or falsification of results along the process leading to its elaboration.

Therefore the work presented in this document is original and authored by me, having not previously been used for any other end.

I further declare that I have fully acknowledged the Code of Ethical Conduct of P.PORTO.

ISEP, Porto, July 26, 2023

Dedictory

This thesis work is dedicated to my daughter, Isabel Novais, who has been a constant source of support and encouragement during the challenges that arose from pursuing this Master's course and the adaptations during the Covid-19 pandemic when we were at home trying hard to follow the classes and deal with the *new normal*. Isabel, remember that the future is just around the corner and no-one knows what wonders the future holds.

I am truly thankful for having you in my life. This work is also dedicated to my parents, Rosa Novais and Avelino Novais, that believed in my capabilities and helped during the most complicated stages.

Abstract

Efficiently screening and evaluating curricula in **recruitment processes** is a critical task that often requires substantial time and effort from Human Resources professionals. This work presents CurEval, an algorithm developed to automate the **evaluation and screening of curricula** based on vacancy requirements. The algorithm utilizes a predefined set of **keywords** and a CSV file format for input, facilitating easy data structuring and processing. To validate the algorithm's performance and address privacy concerns, synthetic curricula were generated using templates with slight variations in personal data. The algorithm's results were compared with evaluations made by a Human Resources collaborator and external paid recruitment platforms.

The study's findings indicate that CurEval effectively filters out irrelevant curricula, reducing the screening workload for HR professionals. The algorithm aligns with human evaluations, ensuring accurate classification of curricula according to **vacancy requirements**. Additionally, bias analysis revealed no evidence of discriminatory **bias** in the algorithm or human evaluations in the sample data.

Further improvements for CurEval include expanding the list of keywords, incorporating **natural language processing** techniques, and integrating **machine learning** to enhance accuracy and adaptability. Real-time data integration, feedback loops with HR professionals, and integration with Applicant Tracking Systems are suggested to streamline the recruitment process. Multi-lingual support, performance metrics, and ongoing ethical considerations are also essential for refining and maintaining the algorithm's effectiveness and fairness.

CurEval offers promising potential to revolutionize the curricula evaluation process, enabling faster and more efficient screening while ensuring fairness and equal opportunity. Future work should focus on enhancing the algorithm's capabilities, addressing biases, and continuously validating and improving its performance through collaboration and feedback from HR professionals.

Keywords: Curricula Evaluation, Artificial Intelligence, Classification, Evaluation, Standards

Resumo

A automação da **análise e classificação de currículos** tem sido alvo de estudo e destaque nas últimas décadas, guiado pela evolução e aperfeiçoamento dos algoritmos de **Inteligência Artificial** e da **Machine Learning**. Nesta dissertação vai ser abordado o processo de **análise e classificação** destes assim como as **questões éticas** e **bias** associados ao processo que advém da natureza humana e das vivências individuais do recrutador. De forma a se evitar que estes ocorram durante o processo de recrutamento foi desenvolvido um algoritmo de análise e classificação dos currículos de acordo com a vaga em questão. Para além deste serão criados **standards** para a classificação e análise dos currículos, independentemente da sua origem e dos formatos.

O algoritmo utiliza um conjunto pré-definido de palavras-chave e um formato de arquivo CSV para entrada, facilitando a estruturação e processamento dos dados. Para validar o desempenho do algoritmo e abordar preocupações de privacidade, currículos sintéticos foram gerados usando modelos com pequenas variações nos dados pessoais. Os resultados do algoritmo foram comparados com avaliações feitas por um colaborador de Recursos Humanos e plataformas externas de recrutamento pagas.

Os resultados do estudo indicam que o CurEval filtra efetivamente currículos irrelevantes, reduzindo a carga de trabalho de triagem para os profissionais de RH. Este está alinhado com as avaliações humanas, garantindo a classificação precisa dos currículos de acordo com os requisitos das vagas. Além disso, a análise de viés discriminatórios revelou que não há evidências da existência dos mesmos no algoritmo ou nas avaliações humanas para a amostragem.

Melhorias futuras para o CurEval incluem a expansão da lista de palavras-chave, a incorporação de técnicas de processamento de linguagem natural e a integração de Machine Learning para aprimorar a precisão e adaptabilidade. Integração de dados em tempo real, ciclos de feedback com profissionais de RH e integração com Sistemas de Acompanhamento de Candidatos são sugeridos para otimizar o processo de recrutamento. Suporte a múltiplos idiomas, métricas de desempenho e considerações éticas contínuas são essenciais para refinar e manter a eficácia e equidade do algoritmo.

Acknowledgement

Words cannot express my gratitude to Dr. Constantino Martins, Vladimiro Ferreira for their invaluable patience and feedback.

I am also grateful to my colleges, especially my office mates and the Human Resources Collaborators, Ana Lúcia and Rita Cruz, for their editing help, feedback sessions, and moral support.

Lastly, I would be remiss in not mentioning my family, especially my parents, sister and children. Their belief in me has kept my spirits and motivation high during this process. I would also like to thank my cats for all the entertainment and emotional support.

Contents

List of Figures	xv
List of Tables	xvii
Listings	xvii
List of Acronyms	xix
1 Introduction	1
1.1 Context	1
1.2 Problem	1
1.3 Objectives & Research Question	2
1.4 Contributions	3
1.5 Research Methodology	3
1.6 Document Structure	4
2 Recruitment & Selection	5
2.1 Types of recruitment	6
2.2 The Recruitment Process	6
2.2.1 Motivation for recruitment	8
2.2.2 The Selection Process	9
Steps on the Selection process	9
Issues in the Selection Process	10
2.3 Ethics in the selection process	11
2.3.1 Conscious and Unconscious Bias	11
2.3.2 Referral Pressure	13
2.3.3 The Curriculum Classification Problem	13
2.4 Manual Curriculum Analysis	14
2.5 Automated Solution for Curriculum Analysis	15
2.5.1 Commercial Solutions	15
2.5.2 Non-commercial solutions	16
2.5.3 Manual versus Computational Process	16
2.5.4 Ethical Questions in using AI Methods	17
2.5.5 General Data Protection Regulation	17
3 State of the art	19
3.1 Paper selection Process	19
3.2 Papers Review	20
3.3 Conclusion	23
4 Analysis and Design	25

4.1	Robotic Process Automation	26
4.2	Requirements	26
4.2.1	Functional Requirements	26
4.2.2	Non Functional Requirements	26
	Adaptability	26
	Maintenance	27
	Parameterization	27
	Technologies	27
	Security	27
4.3	Process To-Be	27
4.3.1	Proof of concept	30
4.4	Summary	31
5	Value analysis	33
5.1	The concept of value	33
5.1.1	Value For Human Resources Department	34
5.1.2	Value for the Company	34
5.1.3	Value for the Candidates	34
5.2	Opportunity Identification and Analysis	35
5.3	Comparison between automated solutions	35
5.4	Quality Function Deployment	36
6	Experiments, Evaluation and Discussion	39
6.1	Research Hypothesis	39
6.2	Indicators and Information Sources	39
6.3	Methodology	40
6.4	Sample Data	41
6.5	Applicant data - Synthetic data	42
6.6	Results Comparison	44
7	Conclusions	49
	Bibliography	53
A	Appendix A	57
B	Appendix B	61

List of Figures

1.1	Design Science Research Methodology	4
2.1	Steps of the Recruitment Process.	7
2.2	Steps of the Selection Process.	10
4.1	Curricula Evaluation Process As-Is	25
4.2	Curricula Evaluation Process To-Be - general overview	28
4.3	Sub-process Email Reception	28
4.4	Sub-process Curriculum Evaluation	29
4.5	Algorithm CurEval	29
4.6	Sub-process Intermediary Emailing	30
5.1	QFD for CurEval	38
6.1	Screenshot from one of the synthetic data analysed in Recruitee	45
6.2	Screenshot from one of the synthetic data analysed in Workable	45

List of Tables

5.1	Comparison between Comercial software and CurEval	36
6.1	Keywords used for the open vacancy to be used as an example	42
6.2	Synthetic Data Generated for the curricula	43
6.3	Comparison between CurEval and HR Collaborator's results	46

List of Acronyms

AI	Artificial Intelligence.
HR	Human Resources.
ML	Machine Learning.
NLP	Natural Language Processing.
POC	Proof of Concept.

Chapter 1

Introduction

In this Chapter we are going to give contextualize the problem that we are trying to solve, as well as the objectives of this dissertation. The research question and the methodology are presented.

1.1 Context

The Recruitment Process is one of the most important processes in a Company, since its main objective is to aggregate talent, as stated by Chiavenato 2010, with the objective of improving its results.

In recent decades there were major improvements and developments in the fields of Natural Language Processing (NLP) as refereed by Harsha et al. 2022 and, also, in the subject of Machine Learning (ML) mentioned in the paper by Sudha et al. 2021. Taking advantage of those advancements new possibilities arise in several different fields that had very complicated questions to be solved, being one of such fields the curriculum vitae classification.

The main goal of this dissertation is to develop a Proof of Concept (POC) that will perform the evaluation of curricula that are received by a company. This project will be called CurEval. This algorithm will perform automatically the classification of them using several criteria such as skills, education, work experience and achievements by searching keywords from the words from the vacancy.

The project for this dissertation was developed in SEG Automotive Portugal, Unipessoal Lda, for the Human Resources (HR) Department. The purpose of the project is to develop a working POC to explore subjects as NPL and ML. It is also a main concern the possible existence of Bias and Peer Pressure during the recruitment process, and the CurEval project is expected to remove these factors from the recruitment process.

1.2 Problem

According to Chiavenato 2010 and Cardoso 2005 one of the keys for a Company's success is to attract and retain talent. Select the correct person to a given position is the focus of the recruitment process. But, sometimes, as mentioned by Chiavenato 2010 and Cardoso 2005, the choice made is not the best match and the company's results suffer.

The recruitment process is very long and it comprises several stages that start even before the vacancy is published, as can be seen in Cardoso 2005. After publishing a given vacancy, it is expected the reception of applications, so, the first step after this publication is the evaluation of the curricula received, also the reception of these can also result from a spontaneous application and not to an application to a given vacancy as mentioned by

Chiavenato 2010.

The emergence of the Covid-19 pandemic forced, for around two years, changes in our daily routines and also on the way we work and interact with our co-workers. So, behaviours had to be altered and adapted to a more digital and remote work environment, Ozimek 2020. For this reason, the recruitment process also became more digital, the vacancies are published on websites like, for instance, Linked-In, the applications are received via e-mail, also the evaluation process is performed digitally. Hence, the volume of curricula to be analyzed is larger than before, this is due to the fact that the possibility of remote work is now valued and this allowed that a collaborator can be working from anywhere in the world.

This increased the time spent analyzing all applications for a given vacancy, making the entire recruitment process longer and more complex, increasing significantly the costs for the company, since the collaborator responsible for that particular recruitment is going to be performing the screening of curricula for a larger amount of time becoming unavailable to perform other tasks during that same period.

It is also important to take into account that when the recruitment process is performed by an Human Resources Collaborator, he/her may be driven by its own bias about the candidate as refereed by Roxburgh and Hansen 2015, resulting in several types of discrimination regarding one or more aspects such as nationality, sex, age and so on. Also a main issue is the possibility of peer pressure for the selection of a given candidate due to referral programs that can influence the selection process.

The reduction of costs related to recruitment and the increase of efficiency of this process are essential and urgent for any company, so the digitization and automation of the curriculum classification process makes sense for large companies with a considerable amount of open positions.

The importance of a good selection process is high due to the fact that when the selected person is not the best match for the vacancy and the company, resulting that the expected goals may not be achieved or even the selected candidate may have higher expectations about the company, causing him/her to leave it, forcing a new recruitment sooner than anticipated. Preventing this situation is very important such as suggested by Chiavenato 2010 and Cardoso 2005, but this endeavour is not always easy to achieve. It is also very important to understand that is very rare to have only one recruitment process occurring in a given time frame.

In Seg Automotive Portugal, Unipessoal Lda, the recruitment process is all manual, so the implementation of an automate process has to be made from scratch.

1.3 Objectives & Research Question

In this dissertation it was intended to develop an working POC, named CurEval, that will evaluate applications to a given open vacancy by classifying each curricula received, according to a set of keywords previously defined for that vacancy that described the preferred requirements that a candidate should present in order to be a perfect fit for the vacancy. this will help to choose wich of the applicants is the best candidate.

The solution should implement the following tasks:

- Analysis of the manual process.
- Drawing of the automated solution.
- Development of the solution.
- Robotic Process Automation (RPA) solution for saving curriculum's received via email.

- RPA solution to send rejection, acceptance and interview appointment emails, finally it would create a folder in a share and archive all curriculum's and close the recruitment process.
- Curriculum classification based on NPL.
- Creation of Standards in evaluation of curriculum's so that all evaluations follow the same guidelines and ensuring homogeneous treatment of all applications.

This dissertation intends to respond a question that arose from a conversation with the IT and RH departments:

The application of NLP techniques can help to decrease the time spent in the screening of curricula and reduce ethical questions, such as bias that may arise during its performance?

This question highlights the ethical concerns surrounding the use of automated systems in the recruitment process, particularly in the context of potential biases and discrimination against certain groups of candidates. It also acknowledges the need for strategies to address these concerns, such as the development of fair and transparent algorithms, the use of diverse and inclusive data sets, and the involvement of human oversight and feedback in the evaluation process.

With this project it is intended, firstly, to decrease the workload on the HR collaborator that is responsible by the recruitment process. Secondly we intend to create standard and reduce possible bias as well as possible peer pressure in the recruitment process.

1.4 Contributions

This dissertation provides a brief study of the recruitment process, with a special focus on the selection sub-process on the curriculum selection phase.

In the state of the art the recruitment process is explored and explained with all its characteristics and peculiarities. Diving into more detail in the selection process, the main issues and ethical questions that can occur in both the manual and the automated process. Possible non- and commercial solutions are presented and a review of recent papers on the subject are presented.

After this the a value analysis of Automating the process in terms of the benefits that can bring, in a general perspective, to the company, employees and applicants, and opportunity identification and analysis are specified for a better understanding of the process.

1.5 Research Methodology

In Figure 1.1 is the visual representation of the process described by Peffers et al. 2007, it is iterative flexible, and composed of six activities.

1. **Problem identification and motivation:** Define the problem and the importance and value of the solution.
2. **Define the objectives of a solution:** Infer the objectives of the solution from the problem definition and knowledge of what is possible and feasible.
3. **Design & Development:** Create the artifact.
4. **Demonstration:** Demonstrate the use of the artifact to solve one or more instances of the problem.

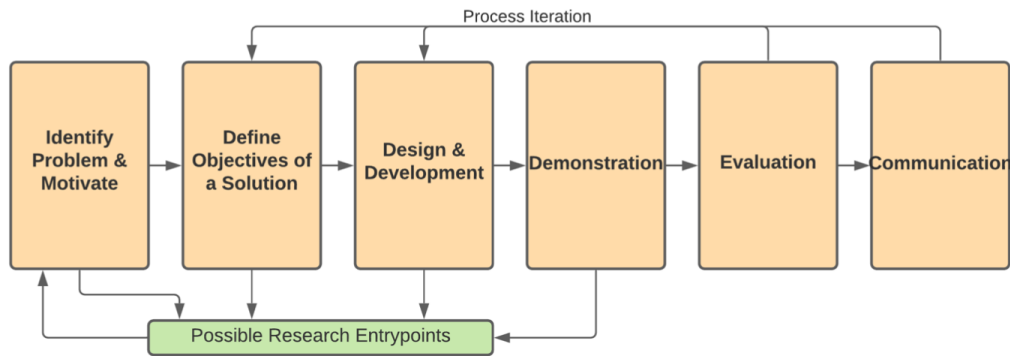


Figure 1.1: Design Science Research Methodology

5. **Evaluation:** Observe and measure how well the artifact supports the solution to the problem. This activity involves comparing the objectives of the solution to actual observed results from the use of the artifact in the demonstration.
6. **Communication:** Communicate the problem and its importance, the artifact, its utility and novelty, the rigor of its design, and its effectiveness.

The research methodology used in this dissertation is the Design Science Research Methodology because it is used to solve identified organisational problems by designing and evaluation IT artefacts.

The first activity, problem identification and motivation, consists of the analysis of the internal process of analysis and classification of curriculum's. In the second activity, define the objectives of a solution, the objective is to evaluate and select what needs to be improved and the objectives with the improvement of the manual process and subsequent automation. The third activity is Design and Development. The design phase is where an architecture solution is developed, and the project plan should be complete. In development is where the developer builds the automated algorithm according to the requirements, functional documentation, and the project plan defined in the design phase. In the four activity, Demonstration, there is only the part of dealing only with process exceptions. The fifth activity, Evaluation, is to evaluate the results comparing the process without automation and the automated one. The sixth and last activity is Communication, and it will be through this dissertation.

1.6 Document Structure

This document comprises the current context for the curriculum analysis process as well as the solution that will be developed. In Chapter 2 a small theoretical introduction on the recruitment process is presented.

In Chapter 3, we present the current state of the art for the dissertation subject.

In chapter 4 the current process was analysed and the requirements form the algorithm were compiled. After the context is given, in Chapter 5, the Value Analysis is made and presented in order to reach a solution for the problem presented. Following, in chapter 6 the experimental and methodology are presented and the results are discussed. And, finally, in chapter 7 the preliminary conclusions are presented.

Chapter 2

Recruitment & Selection

The Recruitment Process is an important process in a company as mentioned in Chapter 1. Several authors have addressed it for decades and published books regarding several parts of the process as well as about several recruitment techniques available.

The two first questions that comes to mind when talking about this process is: What is the Recruitment Process? How is it defined?

Searching in dictionaries one can also find the definition of recruitment, visiting the Oxford English Dictionary website *Oxford dictionary* n.d. the definition of recruitment found is "the action or process of finding and attracting suitable people to fill a vacancy or job."

In Cambridge English Dictionary *Cambridge English Dictionary* n.d., the definition found for recruitment is "the process of finding people and employing them, especially for a job in a company or organization."

Searching the Human Resources literature, one can find other definitions that are less general than in dictionaries.

In his book, Chiavenato 2010 presents several accepted definitions for recruitment:

"**Recruitment** is a set of activities drawn to attract qualified candidates to an organization.

Recruitment is a set of techniques and procedures that aims to attract potentially qualified candidates and capable of occupying positions inside the organization. Is basically an information system, through which the organization divulges and offers the Human Resources Market the job opportunities that it intends to fulfill.

Recruitment is the process to attract a set of candidates to a particular job position. it should announce the availability of the position to the market and attract qualified candidates to dispute it. The market in which the organization tries to find candidates can be internal or external or a mix of both. In other words, the organization must find candidates within it self, outside or in both contexts.

Recruitment is a set of policies and actions meant to attract and aggregate talent to the organization to gift then with the competences required to their success."

Cardoso 2005 defends that the "Recruitment and selection is a chained process of procedures that culminate in the final decision to choose the candidate who will perform the available function, and must be preceded by actions aimed at identifying and characterizing

the job to be filled, performance to the fundamental and which conditions all the following steps."

All the definitions presented highlight the main aspects of recruitment, which is the process of identifying job vacancies, attracting and sourcing suitable candidates, evaluating their qualifications and suitability, and ultimately selecting the best fit for the role.

According to Martins and Cruz 2019, potential candidates try to identify if a given company is the correct for themselves. So it can also be viewed as a two ways evaluation, since the company is searching for the best applicant and also the applicant is searching for the best place to work.

2.1 Types of recruitment

When engaging on a recruitment process there are several concepts to keep in mind and decisions to be made. For instance. The type of recruitment to be used is very important. This decision will influence the outcome of a recruitment process.

Society for Human Resource Management n.d. (SHRM), Cardoso 2005, Camara, Guerra, and Rodrigues 1997 and Chiavenato 2010, all agree that there are two major types of recruitment, the internal and the external.

The **Internal Recruitment** occurs when an employee that already works in the company is selected to occupy a new position, this is also known as *internal promotion*. On the other hand, when the recruitment is based on searching for the best candidate for that vacancy outside of the company, it is called **External Recruitment**.

There are several sub-types of external recruitment, such as:

- **Direct Recruitment** - consists in announcing employment vacancies directly to the public by using job boards, company websites or social media.
- **Indirect Recruitment** - is the recruitment of candidates through employees recommendations, job fairs, campus recruitment, or recruitment agencies.
- **Passive Recruitment** - reaching candidates that are not actively looking for a new job, but may be opened to a new challenge.
- **Campus Recruitment** - consists on recruiting students or new graduates directly from faculties, universities or polytechnic institutes.
- **Diversity Recruitment** - this type of recruitment occurs when all endeavours are gathered into hiring a diversified group of candidates, including under-represented groups, such as, women, coloured people, people with disabilities.

The recruitment types described above can be used individually or combined, depending on the organization's needs and the specific requirements for the vacancy.

2.2 The Recruitment Process

According to *Society for Human Resource Management* n.d. and in the book Several 2009, the recruitment process has several steps that have to be performed to select the perfect candidate for the position available. The process is more of a cycle since it is always ongoing depending on the available vacancies and the new ones that can be opened when the need

arises.



Figure 2.1: Steps of the Recruitment Process.

In Figure 2.1, the Recruitment Process is summarized. As we can observe the process consists of the following steps:

1. **Need for Recruitment** - The first step in the recruitment process is to analyze the demand for the recruitment according to the opening of a position and the company's needs.
2. **Job Analysis** - In this step it is required to understand the job requirements, responsibilities, and specifications.
3. **Create the Job description** - create all the requirements and specifications for the job.
4. **Select the Selection Process** - Decide the methods to perform the selection of candidates, e.g. interview, aptitude tests, and so on.
5. **Decide how to attract candidates** - In this step it is required to choose how the vacancy will be publicized, and what the company offers to the possible candidates.
6. **Divulge the vacancy** - Publish the vacancy in the selected channels.
7. **Sourcing Candidates** - After a previously defined time interval, all the applications are analysed and sorted in order to verify and select a short list the better candidates.
8. **Interviewing** - In this step the face-to-face interviews are conducted, as well as behavioral interviews, technical interviews, or other assessments to evaluate the candidate's suitability for the job.
9. **Selection** - At this step the most suitable candidate is chosen based in their skills, experience and qualifications.
10. **Offer and Acceptance** - here the job proposal is offered to the selected candidate, which accepts (or rejects) it.

11. **Onboarding** - Integration of the new hire into the organization, introducing them to the company's culture and work environment.

The recruitment process aims to find the best match between the job requirements and the candidate's qualifications, skills, and experience.

2.2.1 Motivation for recruitment

There are several motives to start a recruitment process. Consulting *Society for Human Resource Management* n.d. and Martins and Cruz 2019, the main motives to initiate a recruitment process:

1. **Growth and expansion** - The recruitment of new employees supports the growth and expansion of their business.
2. **Replacement of departing employees** - New employees can be recruited to replace employees that have resigned, retired, or been terminated.
3. **New job creation** - Newly created positions can be filled by new employees, due to increase business demand or the launch of new products or services.
4. **Improved skills and capabilities** - Organizations can recruit new employees in order to bring in new skills and capabilities that are not currently present within itself.
5. **Diversification** - Organizations may recruit new employees to promote diversity and inclusiveness within the workplace.
6. **Cost savings** - Organizations may recruit new employees to reduce labor costs, or to reduce labor costs, or even to reduce the workload of existing employees.
7. **Legal compliance** - Organizations may recruit new employees in order to comply with equal opportunity and anti-discrimination laws, or to fill affirmative action goals.

Organizations recruit employees for a variety of reasons, and it is important for them to have a well-defined and structured recruitment process in order to attract and retain top talent and make informed hiring decisions.

The recruitment process can be triggered by the various reasons mentioned above, but it is convenient to dwell a little on the motivations for an employee to leave the company. According to Chiavenato 2010, there are two different kinds for termination: dismissal by the employee, or dismissal by the company.

The second case can occur for several reasons, for instance, inadaptability of the person to the position.

In the case that the termination occur by initiative of the employee is important to analyse the motivations in detail, to prevent a similar situation for occurring. Consulting the page *Society for Human Resource Management* n.d. and Martins and Cruz 2019 we can find some of those reasons:

1. **Lack of growth opportunities** - Employees may feel that they have reached a dead end in their current job and are looking for new challenges and opportunities for growth.
2. **Poor management or workplace culture** - When there are negative relationships with managers or co-workers, or even a toxic workplace culture, can be major drivers for employees to seek a new employment.

3. **Better compensation or benefits** - Another reason for an employee to leave a company is the search for better pay or benefits at another organization.
4. **Lack of work-life balance** - An excessive workload or long hours can lead employees to seek employment that better accommodates their personal and family responsibilities.
5. **Career change** - Some employees may leave their current job to pursue a different career path or switch to a new industry.

It is important for organizations to understand why employees leave, as high turnover can impact morale, productivity, and bottom-line results. Addressing the reasons that employees leave can help organizations to retain their top talent and improve overall employee satisfaction.

2.2.2 The Selection Process

As mentioned previously, the selection process can be seen as a sub-process inside the bigger Recruitment process.

In Chiavenato 2010, the author presents several definitions for the Selection Process:

"Selection is the process of choosing the best candidate for the job.

Selection is the process through which a company chooses from a list of candidates the person that better fulfills the selection criteria for the available position, considering the current market conditions.

Selection consists on obtaining and using information about the candidates recruited externally to choose which of them should receive the job offer.

Selection is the decision-making process based on reliable data to add talents and skills capable of contributing to the long-term success of the organization."

Accessing the *Cambridge English Dictionary* n.d. website, we can find the definition for selection as "the process of choosing someone or something from a group of people or things after careful thought."

While in the *Oxford dictionary* n.d. we can find the definition to be: "the process of choosing somebody/something from a group of people or things, usually according to a system".

From all the definitions above, we can conclude that the selection process final objective is to find the right person for the right place.

Steps on the Selection process

After understanding what is the selection process, it is important to dwell in the process itself and see in detail which steps are taken in it.

According to Several 2009, *Society for Human Resource Management* n.d. and Chiavenato 2010 an overview of this process can be drawn, has showed in figure 2.2.



Figure 2.2: Steps of the Selection Process.

From the analysis of the references, we concluded that the steps of the selection process can be described as:

- **Reviewing curricula and applications** - Initial screening of candidate's curricula and applications to determine if they meet the minimum qualifications for the job.
- **Interviewing** - Candidates are interviewed to assess their skills, qualifications, and fit for the job. Interviews can take different forms, including face-to-face interviews, telephone interviews, or video interviews.
- **Testing** - Candidates may be required to take aptitude tests, skills assessments, or personality tests to determine their strengths and weaknesses.
- **Checking references** - Organizations check candidate's references to verify information provided in their application and during the interview process.
- **Final selection and offer** - Based on the information gathered during the selection process, the organization makes a final decision on the best-suited candidate and extends a job offer.

The selection process is a critical part of the hiring process, since it helps organizations to identify the most qualified and suitable candidate for the job. The process should be well-designed, objective, and transparent to ensure that the best candidate is selected for the job.

Issues in the Selection Process

The Selection process depends on humans and human decision's, so it can be challenging for organizations and several difficulties may , which make the process complex and with different nuances, and sometimes the most qualified person is not selected due to several factors external to the process itself. These difficulties, according to *Harvard Business Review* n.d., *Society for Human Resource Management* n.d. and Several 2009, include the following:

1. **Bias** - Conscious and unconscious bias can impact the selection process resulting in the possible selection of less qualified candidates.

2. **Lack of diversity** - The selection process can also perpetuate a lack of diversity if organizations are not intentional about seeking out and considering diverse candidates.
3. **Inaccurate assessments** - Tests and evaluations used in the selection process may not accurately reflect candidates' skills and abilities, leading to the selection of the wrong person for the job.
4. **Time-consuming** - The selection process can be time-consuming and resource-intensive, especially if it involves multiple interviews, tests, and background checks.
5. **Legal compliance** - Organizations must ensure that their selection process complies with all applicable laws and regulations, such as anti-discrimination laws and privacy laws.

These are some of the difficulties that organizations may face during the selection process. To mitigate these difficulties, organizations should implement best practices, such as regular training to address biases, intentional efforts to seek out and consider diverse candidates, and using validated and reliable assessment tools. For the purpose of this dissertation, it is important to address the Bias subject and also peer pressure due to referral of candidates these two subjects can, in some cases, apply pressure in the Human Resources collaborator responsible for a given recruitment process.

2.3 Ethics in the selection process

In this subsection we are going to explain, how ethics in the selection process ensures fair and just evaluation and decision-making procedures for employment and other opportunities. It involves equal opportunity, transparency, merit-based selection, privacy, non-discrimination, consistency, feedback, and continuous improvement. Upholding these principles promotes fairness, diversity, and inclusion, leading to better outcomes for individuals and organizations.

2.3.1 Conscious and Unconscious Bias

Both in *Harvard Business Review* n.d. and Consul et al. 2021, there can be conscious and unconscious Bias present in the decisions made during the selection process, this addresses the influence that personal beliefs and attitudes can have on decision making. Both bias can be present during a selection process and it is important to understand them in order to improve the process and try to eliminate them.

A Conscious Bias occurs when a decision-maker is aware of his or her biases and consciously makes decisions based on those same preconceptions. For instance, when a decision-maker has a preference for recruiting male candidates and consciously prioritize those candidates over other.

Unconscious bias, on the other hand, refers to biases that a person may hold without being aware of them. These biases can be deeply rooted in a person's subconscious and can affect their decision-making capabilities, without they being aware of it, even in situations where they are trying to make objective decisions, as mentioned by Roxburgh and Hansen 2015. For instance, a decision-maker can subconsciously associate certain physical characteristics to the ability to perform a given job and make decisions based on them.

Both conscious and unconscious biases can negatively impact the selection process by leading to discrimination and the selection of less qualified candidates.

The study of both types of bias is very important to the evolution of the selection process and to improve the companies performances, as mentioned in Roxburgh and Hansen 2015.

Starting with the article published by Spears and Schmader 2021, and the books Burdick 2021 and Fuller, Murphy, and Chow 2020, we can see that there are several types of unconscious bias that one individual can present. The following types can be the most common:

- **Confirmation Bias** - This bias refers to a tendency to favor information that supports the existing beliefs or values, while disregarding information that contradicts them.
- **Implicit Bias** - In Banaji 2013 the authors describe this bias as a systematic error that is driven by automatic attitudes, beliefs, and stereotypes that affect behavior and decision-making. As for instance, in the example of the book, the person was not recommended for a cholesterol test due to the idea that the doctor had that women had less probability of high cholesterol.
- **Cognitive Bias** - Assuming the definition present in book Kahneman 2011, this bias consists in systematic errors that occur during the decision-making and thinking processes that are often driven by unconscious factors (e.g., prior experiences, personal belief systems).
- **Selection Bias** - In the book Cooper, Hedges, and Valentine 2009, the authors refer that this bias refers to a systematic error in the selection of participants or data when conducting an experiment or a study, which can affect the validity and generality of the results.
- **Stereotyping** - Assuming the book Aronson, Wilson, and Akert 2012, the authors indicate that this bias refers to the tendency to make judgements about individuals based on their belonging into a given group, without considering their personal characteristics.
- **Halo effect** - Assuming the definition given in Myers 2009, this bias is the phenomenon where a favorable impression about a person in a given domain affects the evaluation of their abilities and qualities in other domains.
- **Overconfidence bias** - In the book Sharot 2011 it is described as the tendency to overestimate a person's own abilities or its accuracy of its beliefs and predictions.

All the different types bias above described can affect several aspects of one's life, and are particularly true when the recruitment process is made. So, in *Harvard Business Review* n.d. and in *Society for Human Resource Management* n.d. it is important to study bias in the selection process for several reasons, such as:

- **Improving fairness** - The study of bias can help organizations to identify and eliminate sources of bias in the selection process, leading to a more fair and equitable hiring decisions.
- **Enhancing diversity** - Trying to understand and mitigate the sources for bias in the selection process can lead to a more diverse and better representation of different backgrounds and perspectives.
- **Better hiring decisions** - When reducing the impact of bias, organizations can make better hiring decision and select the most qualified candidates for the job.
- **Increased productivity** - When a company hires the right person for the job can lead to increased productivity and better employee engagement.

- **Legal compliance** - Addressing bias in the selection process is also important for legal compliance, as discrimination in hiring is illegal under many laws and regulations.
- **Improved organizational reputation** - A workplace that is free from bias can enhance the reputation of an organization, attract better candidates, and retain existing employees.

Studying bias in the selection process is an ongoing process and requires ongoing efforts to identify and eliminate sources of bias. Organizations should implement best practices, such as regular training to address biases, blind resume reviews, and using validated and reliable assessment tools.

2.3.2 Referral Pressure

It is a common practice in organizations to ask for referrals to current collaborators in order to speed up the recruitment process, but this action can also have a negative impact on the process itself. Referral pressure in the selection process can have significant impacts on the fairness and objectivity of this process. According to *Harvard Business Review* n.d. the negative impacts that may occur are:

- **Sources of pressure** - Referral pressure can come from current employees, friends, family members, or even external recruiters. In many cases, the referring person may have a personal or professional relationship with the candidate and want to see them succeed
- **Lack of transparency** - Referral pressure can often occur behind the scenes and may not be transparent to the organization. This can make it difficult to identify and address the issue.
- **Impacts on merit-based selection** - Referral pressure can undermine the merit-based selection process and lead to the selection of less qualified candidates over more suitable ones.
- **Reduced diversity** - Referral pressure can limit the diversity of the workplace, as referrals are often made among people with similar backgrounds and perspectives.
- **Decreased employee engagement** - Hiring unqualified candidates through referral pressure can negatively impact the morale of existing employees and lead to decreased engagement.

Organizations can minimize the impact of referral pressure by implementing a structured and objective selection process, setting clear criteria for selection, and encouraging employees to refer a diverse pool of candidates.

2.3.3 The Curriculum Classification Problem

After analysing the topics mentioned above, it is possible to conclude that the curriculum classification can lead to several problems:

- **Unconscious bias** - The use of degrees, majors, and institutions as criteria for selection can introduce unconscious bias into the process and exclude qualified candidates who have non-traditional educational backgrounds.

- **Stereotyping** - The use of educational criteria can lead to stereotyping and assumptions about a candidate's abilities and potential, regardless of their actual qualifications or experience.
- **Missed talent** - By relying solely on traditional educational criteria, organizations may miss out on talented candidates who have acquired skills and experience through other means, such as self-study, work experience, or online courses.
- **Reduced diversity** - The use of educational criteria can limit the diversity of the workplace and exclude candidates from diverse backgrounds and perspectives.

To avoid these problems, organizations should adopt a more holistic and diverse approach to the selection process, evaluating candidates based on a wide range of skills, experience, and qualifications, and not just their educational background. This can help to reduce bias and increase the diversity of the workplace.

2.4 Manual Curriculum Analysis

The manual process of screening Curriculum's and evaluating them consists of the following steps as mentioned in Cardoso 2005, Chiavenato 2010, Camara, Guerra, and Rodrigues 1997, Martins and Cruz 2019:

- **Establish Screening Criteria** - before reviewing curriculum's, it is essential to establish clear screening criteria based on the job requirements. This includes qualifications, experience, skills, certifications, and any other specific criteria outlined in the job description.
- **Initial Scan** - Recruiters or hiring managers begin by quickly scanning each CV to get an overview of the candidate's background, qualifications, and experience. This initial scan helps to identify the most relevant information and narrow down the candidate pool.
- **Detailed Analysis** - After the initial scan, the curriculum's that pass the initial screening are subjected to a more detailed analysis. Recruiters carefully review each section of the curricula, including the candidate's personal details, education, work experience, skills, achievements, certifications, and any additional information provided.
- **Assessment of Qualifications** - Recruiters evaluate whether the candidate possesses the necessary qualifications and credentials for the role. This includes assessing educational degrees, certifications, and any specific training or courses relevant to the job requirements.
- **Work Experience Evaluation** - Recruiters closely analyze the candidate's work experience, looking for relevance, progression, and achievements. They assess the candidate's past roles, responsibilities, projects, and accomplishments to determine the level of expertise and suitability for the position.
- **Skill Assessment** - Recruiters assess the candidate's skills mentioned in the curricula, such as technical skills, language proficiency, software knowledge, and other relevant competencies required for the role. They look for evidence or examples that demonstrate the candidate's proficiency in these areas.
- **Achievement Recognition** - Recruiters pay attention to notable achievements mentioned in the curricula, such as awards, recognition, or significant contributions in

previous roles. These achievements provide insight into the candidate's capabilities and potential impact within the organization.

- **Formatting and Presentation** - Recruiters also evaluate the overall formatting and presentation of the CV. They consider factors such as clarity, structure, grammar, spelling, and overall professionalism. A well-organized and error-free CV indicates attention to detail and professionalism.

2.5 Automated Solution for Curriculum Analysis

In the previous section, the manual process for Curriculum Analysis was described, but in recent years the automation for the curriculum analysis gain strength with the development of more robust AI algorithms.

In the following section we will address the automated solution available and compare them with the manual process, to find the commercial solutions the HR collaborator gave me some names and a search on Google The keyword "Recruitment Software" and for the free ones the keyword was "algorithm curriculum vitae evaluation" and the same keywords on B-On website.

2.5.1 Commercial Solutions

With the increasing interest and with the improvement of Artificial Intelligence (AI) algorithms several commercial solutions become available in the market that evaluated curriculum's. Some popular solutions are:

- **Lever** *Lever* n.d. - It is a talent acquisition platform that provides a complete recruitment process, including curriculum evaluation. It offers a flexible and customizable evaluation system, allowing organizations to assess candidates based on their specific requirements.
- **Workable** *Workable* n.d. - It is a recruitment platform that provides tools for automating the recruitment process, including curriculum evaluation. It offers a customizable evaluation system, allowing organizations to assess candidates based on their specific requirements.
- **JazzHR** *JazzHR* n.d. - It is a cloud-based recruitment platform that provides a suite of tools for automating the recruitment process, including curriculum evaluation. It offers a flexible and customizable evaluation system, allowing organizations to assess candidates based on their specific requirements.
- **Jobvite** *Jobvite* n.d. - It is a cloud-based recruitment platform that provides tools for automating the recruitment process, including curriculum evaluation. It offers a customizable evaluation system, allowing organizations to assess candidates based on their specific requirements.
- **SmartRecruiters** *SmartRecruiters* n.d. - It is a cloud-based recruitment platform that provides a variety of tools for sourcing, evaluating, and hiring candidates, including tools for evaluating candidate curriculum's.
- **iCIMS** *iCIMS* n.d. - It is a cloud-based recruitment platform that provides tools for managing the entire recruitment process, including evaluating candidate curriculum's.

These are just a few examples of commercial solutions available for curriculum evaluation. Organizations should research and compare these and other solutions to determine which one is best suited to their specific requirements and goals.

2.5.2 Non-commercial solutions

Besides the commercial solutions there are several types of algorithms that can perform the Curriculum Evaluation according to the requirements for a given vacancy. The next algorithms can perform the evaluation, and are being used in recent years:

- **Natural Language Processing (NLP)** - This algorithm was applied by Harsha et al. 2022, in a Python application to perform the evaluation of curriculum's. This method processes the curriculum's as a human, resulting in an learning process.
- **Machine Learning** - The previous case is an example of ML, but there is more in this subject, for instance, the authors Sudha et al. 2021 used ML algorithms to perform personality prediction.
- **Decision Trees** - In the article El-Ariss, Daher, and Elhadj 2021, the authors used decision trees to perform the evaluation of curriculum's based on their features and attributes.
- **Clustering** - The case of clustering was used in the article by Bondielli and Marcelloni 2021, these algorithms can be used to group the candidate's curriculum's based on similarities and differences, making it easier to compare and evaluate.
- **Artificial Neural Networks** - Mridha et al. 2021 used ANN to classify candidate curriculum's based on their skills, qualifications, and experience, making it easier to determine their suitability for a particular vacancy.

These algorithms can provide organizations with valuable insights and information about candidate curriculum's, helping them make more informed decisions about which candidates to pursue for job opportunities.

2.5.3 Manual versus Computational Process

From the previous sections there are several methods to address the curriculum selection process. The automated process is faster and efficient processing the curriculum's than the manual process, since it processes a large amount of data in a fraction of time than the manual process. When the process is based in algorithms, the classification of process becomes more objective, eliminating personal biases, providing a more objective approach to the curriculum selection, on the other hand the manual process allows a better interaction between the recruiter and candidate allowing a better understanding of the candidate's personal background, experience and qualifications.

The usage of of automatic processes allows that the classification remains consistent since it applies the same algorithms and data analysis methods to all curricula, while in the case of the manual process, the decisions may vary based on personal decisions. In the case of manual selections the recommendations don't have the same level of customization as it is limited by the selector's personal knowledge and experience, while the automated process allows a high level customization.

The automatic curriculum selection is inflexible in the sense that is based on a set of rule

and algorithms, and for this reason does not bring a personal perspective and experience to the process, that could make it more flexible.

2.5.4 Ethical Questions in using AI Methods

Has in the case of the manual process the use of AI methods raise several ethical questions that need to be addressed. In the article Soleimani et al. 2021, the authors address the possible perpetuation of biases that can lead to the the exclusion of certain groups based on their race, gender, or other personal characteristics due to bias that exist previously on the dataset and that can be unavoidably coded into the algorithms.

Since there is a large amount of personal information in the curriculum's analysis process, the usage of that information in AI raises concerns and the potential for misuse of this information. Also the AI algorithms may be difficult to understand and interpret, leading to a lack of transparency in the recruitment process and making it difficult for candidates to challenge the results.

The usage of AI algorithms in curriculum analysis can have unintended consequences, such as the exclusion of highly qualified candidates or the selection of less qualified candidates based on factors that are not related to their ability to perform the job.

Also when using AI algorithms that raises a question regarding the accountability and responsibility for the results of the recruitment process and the decision reached: Is it the HR Department, is it the person who codifies AI, or the one who passes information from the recruitment process? These ethical questions highlight the importance of responsible and ethical use of AI in the recruitment process, and the need for organisations to implement safeguards and best practises to mitigate these risks.

2.5.5 General Data Protection Regulation

The General Data Protection Regulation (RGPD) is a comprehensive data protection law implemented by the European Union (EU). It was adopted on April 14, 2016, and its enforcement began on May 25, 2018. The regulation is designed to safeguard the privacy and personal data of individuals within the EU and provides them with greater control over their information. It can be found in the EU website *RGPD* n.d. and in *RGPD in all EU languages* n.d.

The RGPD applies to all organizations that process the personal data of EU residents, regardless of their location. It also encompasses companies outside the EU that offer goods or services to individuals within the EU. The regulation sets out several fundamental principles and requirements that organizations must adhere to when handling personal data.

According to RGPD, personal data should be processed in a lawful, fair and transparent way. Organizations are required to clearly communicate the purpose and methods of data processing to individuals. The regulation emphasizes purpose limitation, meaning that personal data should only be collected for specific and legitimate purposes and not used in a manner incompatible with those purposes.

Another key principle is data minimization, which calls for organizations to collect and retain only the necessary personal data for their intended purposes. Accuracy is also highlighted, with organizations expected to maintain accurate and up-to-date information, rectifying any inaccuracies promptly.

The RGPD emphasises the importance of storing personal data for no longer than necessary and implementing measures to ensure the integrity and confidentiality of the data. It also grants individuals various rights, such as the right to access their data, the right to rectify

incorrect information, the right to be forgotten, the right to data portability, and the right to object to data processing.

Failure to comply with the RGPD can lead to severe penalties. Non-compliant organizations may face fines of up to 4% of their global annual revenue or €20 million, whichever is higher, depending on the nature and severity of the violation.

Chapter 3

State of the art

In today's competitive job market, the manual screening of Curriculum Vitae (CV) poses a time-consuming and biased challenge for recruiters. However, the integration of Natural Language Processing (NLP) and Machine Learning (ML) has revolutionized CV screening. By leveraging NLP techniques, valuable information can be extracted from CVs, while ML models can analyze and compare candidates' qualifications, skills, and experience. This automation not only reduces the effort and time involved but also ensures fair evaluation and enables recruiters to focus on strategic tasks. As research progresses, we can expect further advancements in CV screening systems, transforming the way employers identify and evaluate potential talent.

In this Chapter we are going to address some current research on this subject.

3.1 Paper selection Process

In order to find relevant work regarding the recruitment process, a search in several sources was made. As a starting point the B-On search feature available on the website was selected since it aggregates a large repository of papers and using the EduVPN one can access easily the different journal websites that can be read freely in The institute.

The first search that was made to obtain information about the recruitment process was made in the tab "Pesquisar Periódicos/ E-books" under "Conteúdos" with the following query:

```
Title contains 'Recruitment'
```

With this query all the e-books are selected that have the the term "Recruitment" in the Title. Using this query, 109 publications addressing it where found. The publications span from several decades so, additional keywords had to be added since the amount of publications has too large to read in its entirety. The new query used in order to decrease the number of publications was:

```
Title contains 'Recruitment and selection'
```

This new keywords selection resulted in only 12 e-books and publications in the subject, making it easier to verify if the results are suited for the project under development.

After having the required background, the advance search of b-on website was used to obtain the publication on the subject. The first query used was:

```
Recuitment TX Todo o Texto
```


And Selection TX Todo o Texto

Using this query, 1171788 results were obtained. This number is impossible to read and to analyse so, further refinement was required. First, all papers that didn't have the full text available were discarded. Followed by the language selection, only documents in English should be taken into account. And finally the date, only papers that were published in the last 2 years should be pertinent due to the fast growth in this area. This reduced the volume to 175835, which is still too high. So, refinement of the query should be added, instead of searching the all document, the key words must be in the abstract and the word Curriculum or Resume, must be added. This resulted in a new query:

Selection TI Título

And Resume AB Resumo

to 87. After this the abstracts were read in order to understand if the papers are relevant for the project at hand.

After this filter was applied an alert was created in order to receive information on new published papers that fall into the search parameters. After searching in the B-on website also the Google Scholar search was used to identify some articles in the references of the selected previously for analysis. The questions that were intended to find answers were:

- What are the benefits and challenges of implementing an automated recruitment system?
- What are the most effective techniques and algorithms for extracting and analyzing information from resumes and job applications?
- How can machine learning algorithms be used to improve the recruitment process?
- What are the ethical and legal considerations when using automation in recruitment processes?
- What are some examples of successful implementations of automated recruitment processes in different industries?
- What are the potential future developments and advancements in automated recruitment processes, and how can organizations prepare for these changes?

3.2 Papers Review

As mentioned in Chapter 1, there have been several developments in the AI area that occurred recently can help to approach a solution to the automatic curriculum analysis as mentioned by Harsha et al. 2022 and also in the paper by Sudha et al. 2021.

The usage of NLP and ML is being applied to the curriculum's analysis problem.

Harsha et al. 2022 present a solution for automated resume screening using NLP and ML algorithms. The authors address the research question regarding the benefits and challenges of implementing an automated recruitment system. Their proposed solution, the "Resumé Screening" web application, follows the Model-View-Controller (MVC) architecture. It efficiently analyzes a large number of resumes by cleaning the documents and assigning scores to relevant skills. The application accepts Word or PDF format files as input, automatically screens the content, and provides output in various formats such as PDF, PNG, or CSV. This approach aims to improve the efficiency and effectiveness of resume screening, enabling

organizations to process a substantial volume of resumes in a shorter time frame.

Mridha et al. 2021 contribute to the research question regarding the most effective techniques and algorithms for analyzing resumes and job applications. Their study focuses on implementing Convolutional Neural Networks (CNNs), a deep learning technique, for screening individual job profiles. The authors demonstrate the CNN-based model's effectiveness in extracting common fields from job descriptions and resumes, overcoming the challenge of heterogeneous terminologies. The proposed model can accurately rank resumes based on relevant skills and experience. The authors suggest that advancements in technology can further enhance the system's performance by addressing the need for annotating a significant number of resumes to train the algorithm.

Kinge et al. 2022 provide insights into the research question regarding the application of machine learning algorithms to improve the recruitment process. Their resume screening system leverages ML and NLP techniques to minimize the cost of hiring new candidates and address key challenges in the hiring process. The authors use NLP and text classification methods, such as n-grams, along with ML algorithms like Random Forest, Multinomial Naive Bayes, Logistic Regression, and Linear Support Vector Machine Classifier. By employing these techniques, the system can effectively rank job applications, identify suitable candidates, and offer suggestions for improving resumes. This implementation showcases the potential of ML algorithms to streamline the recruitment process and assist companies in finding the most suitable talent.

Raghavan et al. 2020 contribute to the research question concerning the ethical and legal considerations of using automation in recruitment processes. Their study focuses on algorithmic pre-employment assessments and the associated biases. The authors highlight the heterogeneity in vendor practices and emphasize the importance of transparency and monitoring metrics beyond disparate impact to mitigate bias. They recommend considering differential validity and advocate for transparency and fairness in algorithmic systems. The findings of this study provide valuable insights into the ethical and legal dimensions of using automation in recruitment and emphasize the need for responsible implementation and ongoing monitoring.

Köchling and Wehner 2020 address the research question regarding the ethical and legal considerations of using automation in recruitment and HR development. Their study emphasizes the potential biases and discrimination introduced by algorithmic decision-making. The authors stress the importance of auditing algorithms to ensure fairness and advocate for considering both objective fairness and subjective fairness perceptions of applicants and employees. They highlight the persistent issue of racial inequality in recruitment and employment and recommend techniques such as diversity training and the availability of positive role models to address implicit biases. The study underscores the need for a balanced approach that integrates human judgment and ethical principles in algorithmic decision-making.

A very recent paper written by Pimpalkar et al. 2023, the authors conclude that natural language processing and machine learning techniques demonstrate their effectiveness in analyzing curricula and extracting relevant information. Curricula parsers employing semantic search and predefined criteria can transform unstructured data into structured data, streamlining the analysis process. However, challenges such as recognizing different writing styles, understanding curricula wording and grammar, and contextual search remain. Machine learning algorithms, including Naïve Bayes, Support Vector Machine, Random Forest, Decision Tree, and KNN, can improve the resume screening and selection processes, automating the task and reducing errors. Future developments should focus on enhancing accuracy by considering writing styles and grammar, exploring advanced NLP techniques, and integrating

other technologies like facial recognition and voice analysis. Overall, the research emphasizes the potential of these technologies while highlighting the need to address challenges and explore further research directions in resume analysis.

The article by Sanyal et al. 2017, presents a model that aims to parse information from unstructured resumes using natural language processing (NLP) techniques. The model extracts essential information such as name, contact details, work experience, education, skills, and more from resumes and transforms it into a structured JSON format. The parsed information is then stored in a database for further use. The model utilizes NLP concepts such as lexical analysis, syntactic analysis, and semantic analysis to effectively process resumes. The main objective is to provide an automated system that can analyze and rank resumes based on job descriptions, helping employers find the most relevant candidates. The article highlights the evolution of hiring processes and the need for intelligent algorithms to handle the increasing volume of resumes. Overall, the research focuses on developing a resume parsing system that can handle diverse resume formats and extract valuable information for efficient recruitment processes.

In the research by Alsharif et al. 2023, the authors study explored the efficiency of text-similarity measures in automated resume screening for recruitment. The Cosine, Sqrt-Cosine, and Improved Sqrt-Cosine (ISC) similarity measures were evaluated in comparison to the decisions made by an expert hiring manager. The findings demonstrated that ISC and Sqrt-Cosine measures were closer to the expert-human decision, outperforming the Cosine similarity measure. These text-similarity algorithms proved effective in shortlisting candidates for various positions, including high-level roles, and exhibited rapid execution time on standard CPU processors. The study suggests that these algorithms can efficiently facilitate decision-making in recruitment and serve as viable alternatives to expert hiring managers.

The bi-directional recommendation system described by Alsaif et al. 2022 utilizes various natural language processing (NLP) techniques to analyze job descriptions and candidate resumes, aiming to improve the matching process between job seekers and job opportunities. The system begins by employing web scraping techniques to gather a large amount of job data from various online sources. This data is then subjected to pre-processing steps to clean and categorize the textual information, ensuring that the subsequent analysis is performed on high-quality and well-organized data.

To extract meaningful information from the job descriptions and resumes, a named entity recognition (NER) model is trained using the spaCy library. The NER model is trained on annotated data to identify relevant entities such as names, locations, skills, and educational qualifications. This step is crucial in understanding the context and requirements of both job listings and candidate profiles.

Once the NER model is trained, it is utilized to match job listings with suitable resumes. The matching process involves calculating similarity measures between the job descriptions and candidate resumes. To achieve this, the Word2Vec model is employed to convert text into distributed word embeddings, representing the semantic meaning of words and phrases. By comparing the embeddings of the job descriptions and resumes, the system can measure the similarity between the two and identify potential matches.

Cosine similarity is commonly used to determine the similarity between vectors, and it is applied in this system to compute the similarity scores between the job descriptions and resumes. A higher cosine similarity score indicates a stronger match between the requirements of a job and the qualifications of a candidate.

By combining the NER model and similarity measures, the bi-directional recommendation system provides a comprehensive approach to recommend relevant job opportunities to candidates and identify suitable candidates for job vacancies. This system not only considers

the skills and qualifications mentioned in the job descriptions but also takes into account contextual information such as location and educational background. The aim is to enhance the efficiency and accuracy of the job matching process, benefiting both job seekers and employers.

By examining these articles in detail, we gain a comprehensive understanding of the benefits and challenges of implementing automated recruitment systems, effective techniques and algorithms for resume analysis, the ethical and legal considerations associated with automation in recruitment processes, successful implementations across different industries, and potential future developments in the field. These insights are crucial for organizations to navigate the complexities of automated recruitment and ensure fair and effective practices.

3.3 Conclusion

After reading the articles above described, one can draw several conclusions.

Automated recruitment systems offer the potential for efficient screening of a large number of resumes, saving time and resources for organizations. They can assist in identifying qualified candidates and ranking applications based on specific criteria.

Implementing automated recruitment systems requires careful consideration of potential biases and discrimination. Ethical and legal considerations must be taken into account to ensure fairness and transparency in the hiring process.

NLP and ML techniques have shown promise in extracting and analyzing information from resumes and job applications. Techniques such as Convolutional Neural Networks (CNNs) and prediction classifiers like Logistic Regression and Random Forest Classification have been employed to evaluate qualifications and predict personality traits.

Machine learning algorithms can enhance the recruitment process by automating resume screening, ranking applications, and identifying suitable candidates for specific job roles. These algorithms enable efficient analysis of large volumes of data, reducing manual effort and improving accuracy.

The use of automation in recruitment processes raises ethical and legal considerations. It is crucial to address potential biases, ensure fairness, and minimize discrimination. Transparency in algorithmic decision-making, monitoring of metrics, and auditing of algorithms are important measures to mitigate these concerns.

The mentioned articles present successful implementations of automated recruitment processes in various industries. These include web applications using NLP and ML algorithms for resume screening, ranking candidates, and providing suggestions for improvement. Such systems have the potential to reduce hiring costs and improve talent acquisition.

The future of automated recruitment processes holds potential for advancements in technology, including the use of more advanced algorithms, increased self-learning capabilities, and the incorporation of new data inputs. Organizations should be prepared to adapt to these changes by staying updated on emerging technologies, conducting further research, and ensuring compliance with evolving ethical and legal standards.

Overall, implementing automated recruitment systems can offer significant benefits, but it is essential to address the associated challenges and consider ethical and legal considerations. Continued research and development are needed to refine techniques, mitigate biases, and enhance the effectiveness of these systems in the recruitment process.

Chapter 4

Analysis and Design

In SEG Automotive Portugal's case all the recruitment process is performed manually, with no usage of automation's during its duration. So, the HR collaborator does the entire procedure manually. In a given month an average of 170 curriculum are received for an average of 3 vacancies available. In Figure 4.1, the process for curricula evaluation is described using Business Process Model and Notation (BPMN).

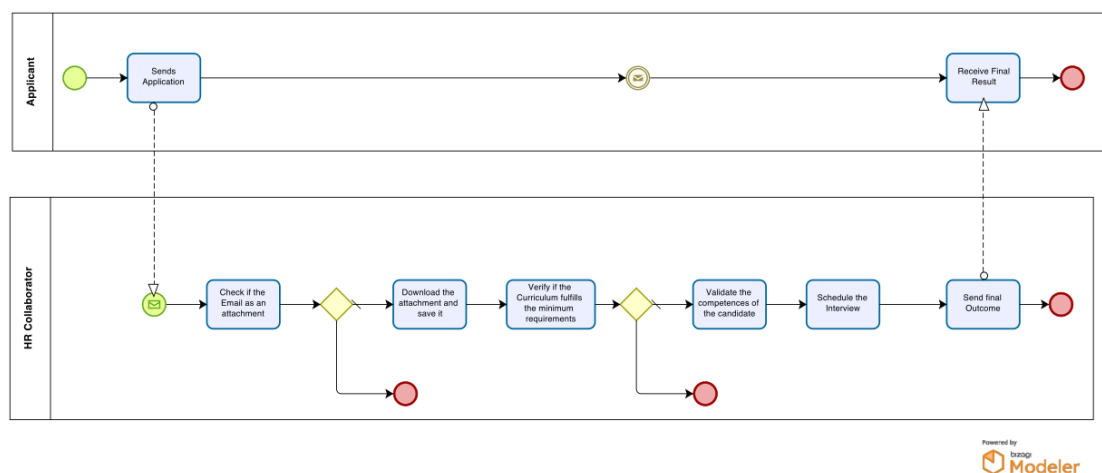


Figure 4.1: Curricula Evaluation Process As-Is

In the Figure 4.1 we can see that the curriculum's are analysed in a first step in a fast way to verify if it contains the essential requirements to the vacancy that the applicant is applying. Due to the large amount of curriculum's received there is no time to verify old applicants or even spontaneous applications that where received previously. Also a document with all information's about the applicants is filled after the first triage.

At this stage the HR collaborator will take around 30 seconds for each curricula in order to check the basic requirements. In the case that the curricula does not have the requirements that are needed the curricula is discarded and the process stops for that applicant. On the other hand, in the case that the minimum requirements are fulfilled a deeper analysis is made to check the total requirements to classify the curriculum and to fill the document with all applicants. In this case, after classification, an interview is schedule to know the applicant and get more information about him/her.

Saving the attachment will take about 30 seconds, spending around 85 minutes per month

in this task.

At this moment, only the preliminary evaluation is made in a regular basis, and it takes around 85 minutes per month, for the detailed analysis it would take around 2-3 minutes but at the moment due to lack of time it is made during the interviews, and the entirety of the recruitment process takes up to 1.5 to 2 months.

4.1 Robotic Process Automation

Robotic Process Automation (RPA) is a technology that uses software robots or "bots" to automate repetitive, rule-based tasks within business processes. RPA bots mimic human interactions with computer systems and perform tasks such as data entry, data extraction, form filling, data validation, and more. There are several technologies to use such as UiPath, Automation Anywhere, Blue Prism and Power Automate.

4.2 Requirements

After analysing and understanding the process as it is made, a list of requirements has been made.

4.2.1 Functional Requirements

The functional requirements that have been defined for the process were:

- Download the curricula from the email and save in the correct folder.
- Read the curricula in the format that is available (PDF or Word).
- Read the requirements for the vacancy.
- Extract the relevant information from the curricula.
- Compare the curricula and the vacancy.
- Give a classification to the curricula based on the comparison.
- Send automated emails for applicants.
- Notify HR team with the results from the comparison.

4.2.2 Non Functional Requirements

In addition to the functional requirements, non functional requirements were defined, and they are grouped by: adaptability, maintenance, parameterization, technologies and security.

Adaptability

- The algorithm must be able to compare documents in different formats, and languages.
- The algorithm must be able to extract information from documents, of the same type, with different structures.

- Curricula are created in several formats and templates, so it is complicated to predict the format it is sent.

Maintenance

- The process must be maintained for new requirements in the classification process as well as new formats of curricula.

Parameterization

- The algorithm must be parameterized with the requirements for the vacancies to be dynamical and to automatically change with the open vacancy.
- The confidence values for the NPL algorithm must be defined and parameterized.

Technologies

- The algorithm must be developed to be used in an RPA software, such as UiPath, PowerAutomate, BluePrism, and others.
- The documents must be saved in SEG Automotive Portugal SharePoint.
- Notifications must be sent to the HR Recruitment Team.

Security

- All data related to the candidate should be protected under the RGPD.
- The algorithm should only use the data required for the evaluation process.

4.3 Process To-Be

In this section we are going to present the process To-be, this means the process as it is going to be executed after the automation is totally implemented.

The general overview for the process To-Be is presented in the Figure 4.2, here it is represented both the process made by the applicant and the SEG Automotive Portugal Ida automation. As seen in the Figure 4.2, the application is sent by the applicant and received by the automation.

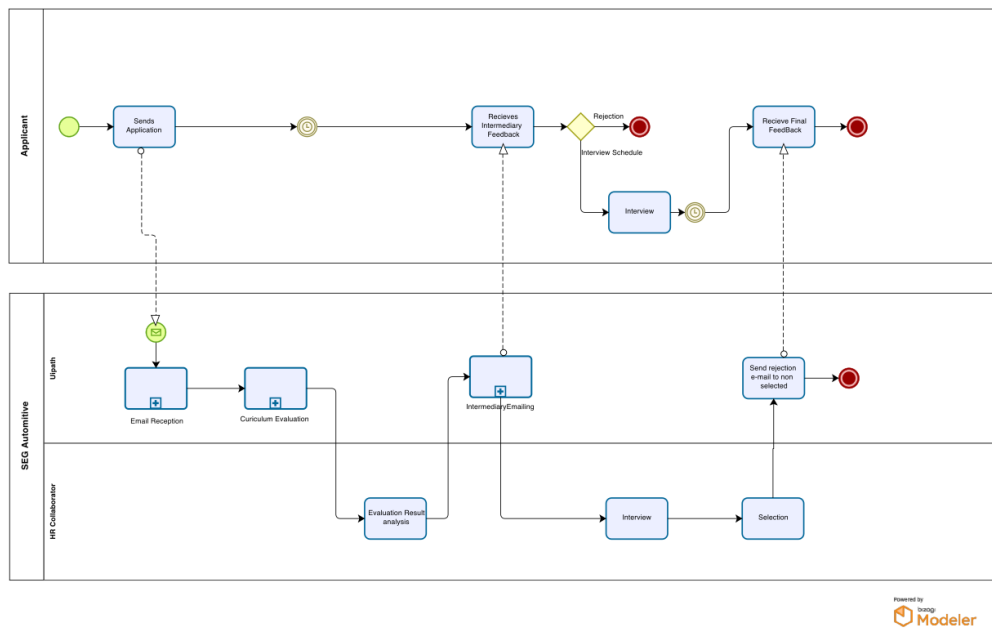


Figure 4.2: Curricula Evaluation Process To-Be - general overview

At this point the sub-process represented in the Figure 4.3 is represented, and the curriculum is saved into the folder corresponding to the open vacancy.

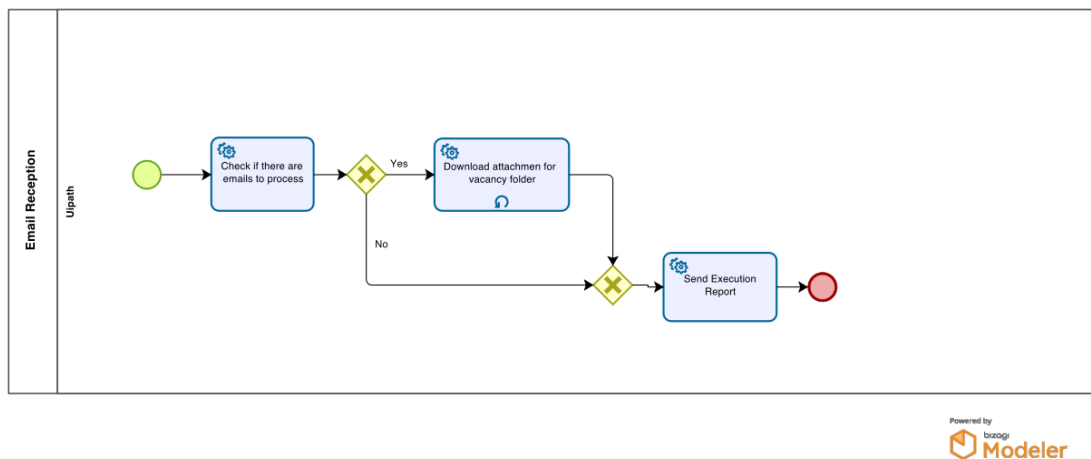


Figure 4.3: Sub-process Email Reception

For each Curricula received the automation will be validated and evaluated, as shown in Figure 4.4. It is at this point that the CurEval algorithm is activated and the NLP is used to Evaluate the curriculum in question as shown in Figure 4.5.

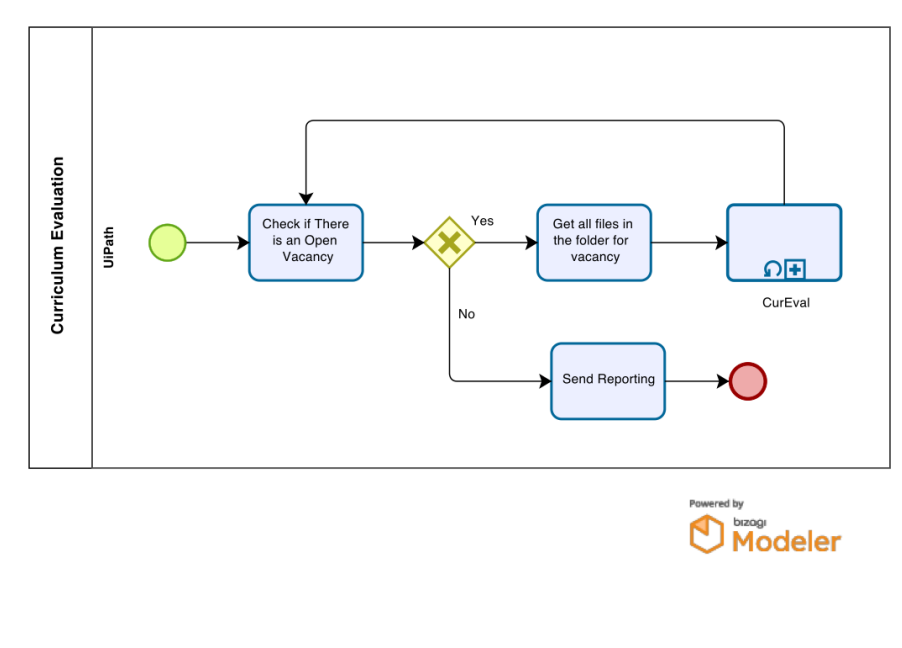


Figure 4.4: Sub-process Curriculum Evaluation

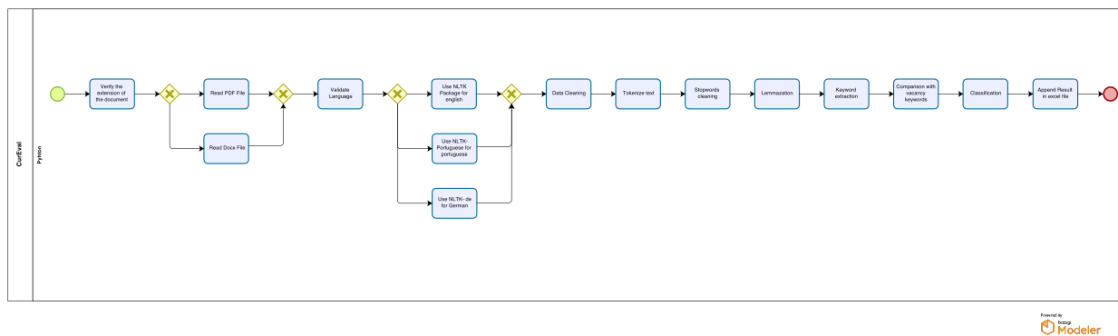


Figure 4.5: Algorithm CurEval

After the curriculum is evaluated by the automatism, the results are analyzed by the Human Resources Collaborator and the interview dates are scheduled and added to the control file, in order to send an intermediary email for the applicants with the results at that point in the process, meaning, scheduling interviews and rejection emails for the applicants that do not fulfill the minimum requirements for the vacancy, as described in Figure 4.6.

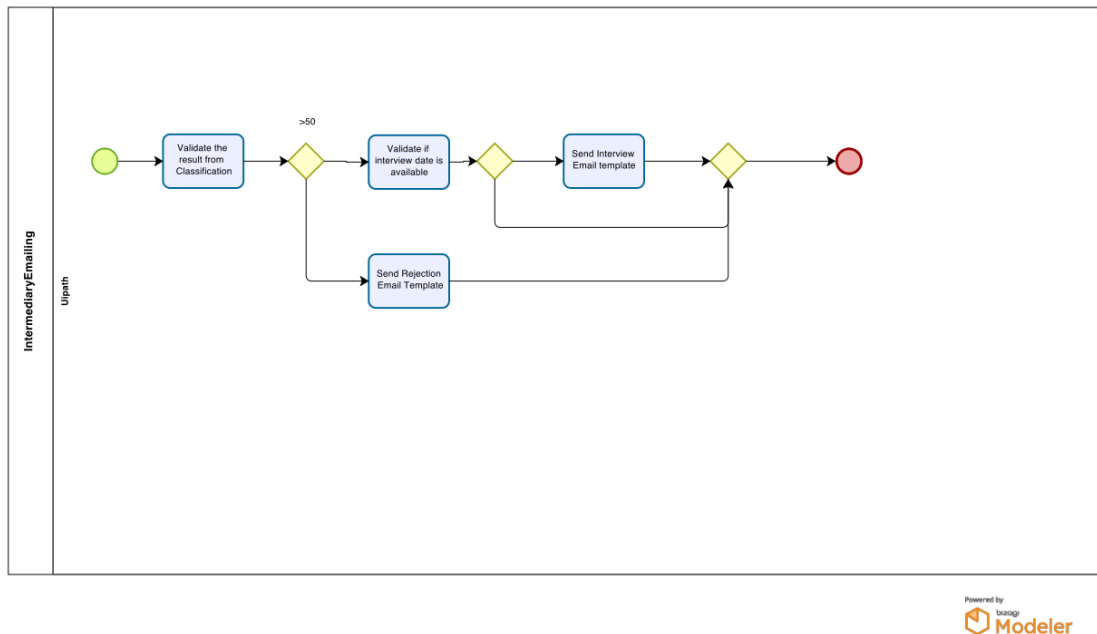


Figure 4.6: Sub-process Intermediary Emailing

After this automated part the interviews are made and the selection of the best candidate is made, as shown in Figure 4.2. Then the automation closes the process by sending the final emails with the conclusion of the process.

4.3.1 Proof of concept

The algorithm to be developed and shown in Figure 4.5 is going to be the final delivery to the company, since the objective of this thesis is to develop a POC to evaluate the potential of the algorithm and the possibility of expanding it to the entirety of the use cases available in the curriculum evaluation sub-process.

As a POC, an algorithm using the NLTK package is going to be used. This way we can compare the results from the three algorithm in order to verify if they all return the same results or if not which one in better.

The POC for email sending was developed with the HR collaborator using Power-Automate and SharePoint lists to facilitate the sending of emails before all automation is concluded.

The rejection email template developed to send for the applicant is the following:

Bom dia X,

Agradecemos toda a disponibilidade e interesse demonstrados na nossa oportunidade para Y, contudo para esta oportunidade optamos por não dar continuidade à tua candidatura.

Continuaremos com o teu currículo na nossa base de dados para eventuais oportunidades onde o mesmo se possa enquadrar.

Até lá, convidamos-te a acompanhar as nossas ofertas através da consulta do nosso site: <https://www.seg-automotive.com/career-application-platform>

Votos de muito sucesso!

Cumprimentos / Mit freundlichen Grüßen / Best regards,

And, the proposal email is the following:

Olá X,

É com agrado que envio em anexo a nossa proposta para a posição de Y, no nosso departamento de x.

Alguma dúvida ou questão, não hesites em colocar-nos.

Obrigada.

Cumprimentos / Mit freundlichen Grüßen / Best regards,

Three examples of vacancy announcements are presented in Appendix A.

4.4 Summary

Based on the information gathered, it is clear that the current recruitment process at SEG Automotive Portugal is heavily reliant on manual work, which can be time-consuming and inefficient. The average number of received curriculum for the available vacancies is quite high, which increases the workload for the HR collaborator. The current process involves a preliminary evaluation of the curricula to check if they meet the minimum requirements, followed by a more detailed analysis of the relevant curricula.

To optimize and streamline this process, a list of functional and non-functional requirements has been defined. The functional requirements include downloading and reading curricula in different formats, extracting relevant information, comparing curricula with vacancy requirements, classifying curricula based on the comparison, sending automated emails to applicants, and notifying the HR team of the results. On the other hand, the non-functional requirements include adaptability, maintenance, parameterization, technologies, and security.

To meet these requirements, an algorithm or system could be developed that automates the recruitment process. The POC algorithm use Natural Language Processing (NLP) Harsha et al. 2022 to extract relevant information from the curricula and comparing it with the job requirements where developed. After the POC is approved, the final algorithm must be able to handle different formats and templates of curricula and languages. It must also be parameterized to adapt to changes in the requirements for different vacancies.

Furthermore, the algorithm must be developed in a way that ensures the security and privacy of candidate data, complying with the RGPD regulations. The automated process should only use data necessary for the evaluation process.

In conclusion, automating the recruitment process can greatly improve the efficiency of the HR department at SEG Automotive Portugal, Unipessoal Lda, while also reducing the workload of the HR collaborator. Developing an algorithm that meets the functional and non-functional requirements listed above can provide a reliable and efficient way to screen, evaluate, and classify applicants' curriculum.

Chapter 5

Value analysis

Value analysis is defined as a method of assessing the worth of goods or services measuring their cost against their perceived value. It is a process that allows businesses to make decisions regarding the most cost-effective and financially beneficial products or services. So, value analysis looks at the balance between cost and performance.

Performing a value analysis implies looking at a product from different perspectives, e.g. quality, cost and usability, to determine which elements are more important to the customer and which ones can be reduced or eliminated to optimize savings. It also involves analyzing the current processes used to manufacture goods and services and considering other, more cost-effective methods of production.

In the following section, the value for the dissertation proposal will be analysed.

5.1 The concept of value

According to *Cambridge English Dictionary* n.d., value is:

1. "The amount of money that can be received for something.
2. The importance or worth of something for someone.
3. How useful or important something is."

In *Oxford dictionary* n.d., value is defined as:

1. "How useful or important something is.
2. How much something is worth compared with its price.
3. The quality of being useful or important."

From the definitions found in both dictionaries we can conclude that the definition of value can be considered as a subjective idea since the value depends on several aspects that may be influenced by personal ideas and perspectives.

In the context of this dissertation the value that an Automated Curriculum Classification is the ability to improve the accuracy, efficiency and objectivity of the recruitment process. AI systems use advanced algorithms and natural language processing to analyze resumes Harsha et al. 2022, making them less prone to human error and more accurate in identifying relevant skills and experiences. Due to the fact that an AI pre-screens the curricula this process becomes faster and enables recruiters to process the documents faster and more efficiently, leaving more free time to perform other important tasks.

In Kleinberg et al. 2020, the value brought by Automated systems is analysed in terms of

reducing bias. So the key benefits that can be added in the case of the AI Curriculum Evaluation are:

- **Cost Savings** - Automated systems can significantly reduce the time and cost of the recruitment process when done manually.
- **Improved Accuracy** - The usage of AI systems to analyze curriculum's makes them less prone to human errors and more accurate in identifying relevant skills and experiences.
- **Reduced Bias** - The Automated systems are design in an way to reduce the unconscious bias, making them more impartial and for this reason, the risk of discrimination is reduced.
- **Enhanced Data Analysis** - Using automated systems can also provide valuable insights into the strengths and weaknesses of each candidate, this will help recruiters to make well-informed hiring decisions.

5.1.1 Value For Human Resources Department

In the case of this project, the customers of the built algorithm will be the Internal Human Resources from SEG Automotive Portugal, Unipessoal Lda. In the present, the process of analysis and classification curriculum's is made manually for each vacancy and there are always some positions open. There is only one person to perform this task in the HR department, and it is a lengthy process since there are always many applicants, and many curriculum's to evaluate.

Also, automated systems can provide data-driven insights that help HR personnel to make better informed decisions.

5.1.2 Value for the Company

For SEG Automotive Portugal, Unipessoal Lda, automatizing this process will save time and decrease the workload of the resource, since the automated process will screen the curriculum's in a faster velocity in a short time, reducing the costs associated to the hiring process. Also the accuracy of the automated process will assure that human error will be eliminated and that the most qualified candidates are selected.

The automated process also improves the efficiency of the process, since it can screen and sort through several volumes of curriculum's and applications, helping to identify candidates for a given position. Speeding the hiring position.

Also as value gained by the automated systems is the reduction of unconscious bias in the hiring process by standardizing the selection criteria and removing any subjective factors.

Automating the curriculum selection process can help companies identify the best candidates for a position more efficiently and effectively, leading to a more streamlined and effective hiring process.

5.1.3 Value for the Candidates

From the candidates point of view there are also several benefits to take into account when the curriculum selection process is automatized.

When companies use automated systems to select applicants to a given vacancy, it can help to eliminate unconscious bias that may influence hiring decisions. This implies that

candidates that might otherwise have been overlooked due to their gender, or other factors unrelated to their qualifications may now have a better chance of being selected for an interview or further consideration.

When an algorithm for curriculum screening is coded it is design to evaluate based on a standardized set of criteria, which can be particularly beneficial for candidates who may not have connections or other advantages that could give them an edge in the hiring process, since it evaluates the curriculum's fairly and objectively.

When the process is automated, companies can evaluate a large number of resumes more quickly than they would be able to do manually. So, for candidates, this means that they receive the response for their application sooner, which can be particularly important if they are in need of a job urgently.

It also can be assumed that the automated analysis tools can provide feedback to the applicants, which can help them to identify areas of improvement in their curriculum's. This can also help to improve the overall candidate experience and enhance the company's reputation as a fair and supportive employer.

So, resume analysis can provide several potential benefits for candidates, including increased objectivity, faster screening, and improved chances of being selected for further consideration.

5.2 Opportunity Identification and Analysis

The recruitment process lasts a very long period of time in SEG Automotive (according to HR Department, around 1.5 to 2 months), and the amount of curriculum's received is very high (average 170 per month). Since all process is manual, there is room for improvement since the automation of part of the process will decrease the time that it takes for the process to finish.

This will allow the reduction of costs with the recruitment and save time and resources for recruiters and hiring managers, as well as improve the overall quality of the hiring process by ensuring that the candidates selected for consideration are a good match for the job requirements.

Also, automated curriculum analysis can help to reduce bias and improve diversity in hiring, as it can be programmed to prioritize certain qualifications and experience over others, rather than relying solely on subjective judgments.

However, it is important to also consider the potential drawbacks of relying too heavily on automated curriculum analysis, such as the risk of overlooking talented candidates who may not fit neatly into predefined categories or failing to account for relevant soft skills and personal attributes that may be important for success in a particular role.

Overall, the opportunity that can arise from automated curriculum analysis lies in its ability to streamline the hiring process and improve the quality and diversity of candidate pools, while also recognizing the limitations of this approach and supplementing it with other methods of candidate evaluation as needed.

5.3 Comparison between automated solutions

In the website *Datalabor - DADOS ESTATÍSTICOS E INFORMAÇÃO JURÍDICA NAS ÁREAS DO TRABALHO, EMPREGO E PROTEÇÃO SOCIAL* n.d. the statistics related to work are compiled. Searching the website for the values for last year, 2022, we encounter that there where 11431 employment vacancies, and 475802 applications in total, the values

where obtained from the metadata of Instituto do Emprego e Formação Profissional (IEFP), assuming this values and averaging one gets that for a vacancy there where around 42 applications for each job. This value can be smaller than reality since the statistics only assume the values registered on the IEFP website, neglecting candidates that are not registered in the IEFP.

Assuming the values given by the HR department of SEG automotive, that is an average 170 application per month for 3 open vacancies, one gets the average value of 57 applications, that is a value superior to the one from IEFP.

In the Chapter 2, some commercial applications were described, in Table 5.1, we present a comparison between three commercial solutions and the algorithm that is going to be developed CurEval.

Table 5.1: Comparison between Comercial software and CurEval

	Lever	Workable	JazzHR	CurEval
Cost	Payed	Payed	Payed	Free
Scalable	yes	yes	yes	yes
Reporting	yes	yes	yes(in highest plan)	yes
Bias	yes	no	N/A	yes
Peer Pressure	no	no	N/A	yes
Ranking Curriculum's	no	no	no	yes
Automated Responses	yes	yes	yes	yes
AI	yes	yes	yes	yes
On prem	no	no	no	yes
UiPath	no	no	no	yes

In the case of Lever no information about pricing was found but there is the information that it only has a free trial for 7 days. While in the case of JazzHR several payed plans are available depending on the size of the company as well as the functionalities needed.

The main requirement that was given by SEG Automotive is that the solution implemented must be free and that can be implemented in Uipath, or other Robotic Process Automation Software available. So, the solution CurEval will be developed in Python that can be easily called in an Uipath process automation. Uipath has an AI environment that can be used in a payed version but in a first approach it is not going to be used due to its value.

5.4 Quality Function Deployment

The Quality Function Deployment, or simply QFD, is a structured approach to product design and development that originated in Japan in late 1960's, as mentioned by Akao 1990 and Hauser and Clausing 1988. It is a process that involves translating the voice of the customer into specific product characteristics and design requirements.

Itenables organizations to prioritize customer needs, determine the product features that are most important to them, and develop a product that meets them. It is a comprehensive process that considers not only customer needs, but also technical requirements, business objectives, and other factors that are critical to a product's success. The QFD process typically involves the following steps:

1. Identifying customer needs and expectations.
2. Determining how the product will meet those needs and expectations.

3. Identifying key technical requirements for the product.
4. Prioritizing product features and characteristics based on customer needs, technical requirements, and business objectives.
5. Translating those product features into specific design requirements.
6. Continuously evaluating the product throughout the development process to ensure that it meets customer needs and expectations.

QFD has been widely used in product development, engineering, and manufacturing. It has been particularly popular in the automotive industry, where it has been used to improve the quality and reliability of vehicles. QFD has also been applied to other industries, such as software development, healthcare, and service industries. In chapter 4, the requirements from the customer were summarized, and these are the ones assumed in the QFD analysis. In Figure 5.1, the algorithm to be developed is compared with the requirements from the client as well as three competitors. As is visible the main strength that CurEval has is that it is going to be developed in an open Source language, so the costs are very low. Also as UiPath is a language that is already used in SEG Automotive Portugal, Unipessoal Lda, no added cost will be spent. Also All process will run in in-house servers and computers, so no data will be saved on cloud, as solicited by the company. Since all other software's available are payed, this makes the future CurEval as the best solution based on the requirements made by the customer.

Chapter 6

Experiments, Evaluation and Discussion

The problem and objectives of this dissertation are described in Chapter 1, where it is possible to highlight productivity problems, errors, problems, and concerns. When all of these matters of business are well defined, leads to another problem that is to successfully implement an algorithm avoiding the risks that could lead to project failure. It is then determined as objectives the study and implementation of Curriculum Analysis using Python.

6.1 Research Hypothesis

The use of automated curriculum screening techniques can improve the efficiency of the process of screening and selecting candidates for job vacancies Harsha et al. 2022.

Screening and selecting curricula for job openings can be a time-consuming and labor-intensive process, especially when there are a large number of applicants for a single vacancy (Chiavenato 2010). In addition, manual screening can be prone to bias and human error, which can result in the selection of unsuitable candidates or the exclusion of qualified candidates Several 2009. In this sense, the hypothesis proposes that the application of data analysis and automated classification techniques can help to improve the efficiency and accuracy of this process.

Based on this hypothesis, a research will be developed to test the effectiveness of the automated analysis and classification of CVs, comparing it with the traditional manual screening process. Research could include collecting candidate and resume data, applying automated analysis and ranking techniques, comparing results with manual screening, and evaluating the effectiveness and efficiency of each method. The research could also include the analysis of other factors, such as the accuracy in selecting suitable candidates, the speed of the process and the perception of recruiters and candidates in relation to each method.

6.2 Indicators and Information Sources

When analyzing and classifying CVs, some indicators and sources of information that can be considered include: Indicators:

1. Number of resumes received.
2. Number of qualified candidates for the vacancy.
3. Candidate response rate.

4. Average time for screening and selection of candidates.
5. Accuracy in selecting suitable candidates.
6. Satisfaction of candidates with the selection process.
7. Satisfaction of recruiters with the effectiveness and efficiency of the process.

Information sources:

1. Resumes sent by candidates.
2. Candidates' personal data, such as age, gender, education level, professional experience, among others.
3. Interviews with selected candidates.
4. Feedback from recruiters and candidates on the selection process.
5. Performance data of selected candidates after hiring.
6. Market data, such as unemployment rate, industry average salary, among others.

Some of these indicators can be obtained directly from the curricula submitted by candidates, while others may require the collection of additional information, such as conducting interviews with selected candidates or monitoring the performance of hired candidates. Additionally, it is important to remember that market data may be obtained from secondary sources such as market research reports, government data or industry statistics.

With the analysis of these indicators and sources of information, it is possible to evaluate the effectiveness and efficiency of the process of analysis and classification of resumes, and to carry out continuous improvements to increase the accuracy and effectiveness of the selection of candidates.

6.3 Methodology

The evaluation methodology for analyzing and classifying CVs can follow the following steps:

1. **Define the objectives of the evaluation:** It is important to define the objectives of the evaluation, including what you hope to achieve with the automated analysis and classification of resumes, such as improving the efficiency of the process, selecting qualified candidates and reducing bias.
2. **Select Performance Indicators:** Based on the evaluation objectives, select the performance indicators that will be used to measure the effectiveness and efficiency of the resume review and ranking process. Some examples of indicators might include the number of curricula received, the candidate response rate, the average time to screen and select candidates, the accuracy in selecting suitable candidates, and the satisfaction of candidates and recruiters with the process.
3. **Collect Data:** It is necessary to collect data from the candidates and the selection process to evaluate the selected performance indicators. Data can be collected from resumes submitted by candidates, interviews with selected candidates, feedback from recruiters and candidates on the selection process, and performance data of hired candidates after hiring.

4. **Apply Data Analysis and Automated Ranking:** It is important to apply data analysis and automated curricula ranking to a candidate data-set to compare the performance of the automated process to the traditional manual screening process. This can be done using software specialized in data analysis and CV classification.
5. **Compare the results:** With the data collected and the data analysis performed, compare the results of the automated process with the traditional manual process. Analyze the performance indicators and identify the differences, strengths and weaknesses of each method.
6. **Evaluate Results:** Evaluate the results of the benchmarking and identify necessary improvements to the automated process to increase the effectiveness and efficiency of candidate selection.
7. **Apply improvements:** Based on the assessment results, apply improvements to the automated curricula review and ranking process. Repeat the evaluation process to assess the effectiveness of the implemented improvements.

This methodology can help to continuously evaluate and improve the process of reviewing and ranking resumes, increasing accuracy and effectiveness in selecting suitable candidates for job openings.

6.4 Sample Data

In the evaluation of the CurEval algorithm, which utilizes NLP techniques to search for keywords, a meticulous approach was taken to test its performance for an Accounts Receivable vacancy. The target position for evaluation was carefully selected, and a manual analysis was conducted to extract the required data from the vacancy description. This data encompassed crucial qualifications, skills, and other pertinent criteria specific to the position.

The extracted information from the vacancy description was then used to create a CSV file that served as the basis for evaluating the curricula. The CurEval algorithm utilized NLP to search for keywords within the curricula, aiming to identify candidates who possessed the essential qualifications and skills outlined in the vacancy requirements.

By employing NLP techniques, the algorithm could analyze the textual content of the curricula and compare it against the extracted keywords. This allowed for a more precise and automated evaluation process, as the algorithm could identify candidates whose curricula matched the desired qualifications for the Accounts Receivable vacancy.

The utilization of NLP in the CurEval algorithm not only saved time but also enhanced the accuracy of the evaluation. By automating the search for keywords, the algorithm reduced the reliance on manual analysis, thereby increasing efficiency. Moreover, the algorithm's ability to identify relevant keywords within the curricula helped to identify suitable candidates more effectively, streamlining the selection process.

Table 6.1, presents the keywords that were identified during the analysis and used as the test requirements for the vacancy. These keywords encompassed various aspects, such as economics, finance, management, proficiency in Microsoft Windows, English communication skills, customer service, problem-solving ability, teamwork, a bachelor's degree, and accounting expertise. These keywords formed the basis for evaluating the suitability of candidates through the algorithm.

Table 6.1: Keywords used for the open vacancy to be used as an example

Keywords
Economics
Finance
Management
Microsoft
Windows
English
Communication
Costumer
Problem
Team
Team Environment
Bachelor
Accounting

In order to test the algorithms effectively while adhering to data privacy regulations, such as the RGPD (General Data Protection Regulation), the decision was made to acquire online curriculum templates rather than using actual applicant resumes. These templates, available in the appendix B, were chosen as examples that were not tied to specific individuals. This approach ensured compliance with data protection regulations while providing representative examples of resumes for testing purposes.

To assess the presence of any biases within the algorithms, it was crucial to introduce subtle variations into the curriculum templates. These differences were deliberately incorporated to validate whether the algorithms exhibited any discriminatory behavior based on factors like race, nationality, or other protected characteristics. By employing slightly modified versions of the templates, it became possible to evaluate the algorithm's fairness and impartiality in candidate selection, focusing on the shared requirements while varying the personal attributes.

Through this meticulous testing approach, which involved the analysis of the vacancy, the use of representative curriculum templates, and the introduction of subtle discrepancies, the algorithms could be assessed for potential biases and their ability to fairly evaluate candidates based on the specified requirements.

6.5 Applicant data - Synthetic data

To conduct a comprehensive evaluation of bias in the algorithm's coding, a meticulous approach was taken. Synthetic data was generated using the curriculum templates provided in Appendix B. Each template was used to create three distinct curriculum vitae, resulting in a diverse dataset for analysis. The variations in the generated curricula primarily focused on personal data, including applicant names, addresses, telephone numbers, email addresses, universities attended, and dates of work experience. By keeping all other aspects, such as skills, education, and professional experience, constant across the variations, the investigation specifically targeted potential bias arising from personal attributes.

This approach allowed for a systematic examination of how the algorithm handled personal information within the curricula. By varying personal attributes while maintaining consistency in other areas, it became possible to isolate the impact of such attributes on the algorithm's evaluation and decision-making processes.

Through this analysis, it was possible to assess whether the algorithm exhibited any biased

behavior or disparate treatment based on personal attributes such as names, addresses, or educational backgrounds. By carefully controlling the variables and generating a diverse set of curricula, the investigation aimed to uncover any potential biases encoded in the algorithm.

The thorough evaluation of bias in the algorithm's coding is crucial to ensure fairness and mitigate any discriminatory effects in the recruitment process. By proactively identifying and addressing biases, organizations can strive for equal opportunities and promote diversity and inclusion in their hiring practices.

Table 6.2 presents the synthetic data that was meticulously crafted for the curricula. It encompasses a wide range of information, such as application codes, the specific template employed, applicant details, and educational and professional background. Notably, the dataset deliberately incorporates examples that vary in gender and nationality, as these are prominent factors associated with bias in hiring processes. By introducing such diversity into the synthetic data, the algorithm's ability to make unbiased and fair decisions across different demographic groups could be thoroughly scrutinized.

Table 6.2: Synthetic Data Generated for the curricula

Applications Code	Template used	Applicant Name	Address	Telephone	Email	University	Date	Work Experience	Date
AP001	Template A	Sara Teixeira	Rua ABC Porto	91000000	Sou.eu22@qualquemail.com	FEUL	2020	XYZ, Coimbra	2020-2021
AP002	Template A	Rui Silva	Rua das Flores Lisboa	92100000	xpto2@qualquemail.com	FEUp	2015	Chi, Lisboa	2015-2021
AP003	Template A	Abdul Mohamed	Rua das Flores Lisboa	92100001	xpto1@qualquemail.com	FEOP	2015	AIL, Lisboa	2015-2021
AP004	Template B	Helena Cardoso	Rua ABC Porto	92999999	Aqui.vai@gmail.co	UA	2015	Acv, Braga	2014/2015
AP005	Template B	Bruno Mcneal	Rua das mimosas Porto	934999999	Mais.um.mail@gmail.com	FEUP	2017	Mxv, Bragança	2018
AP006	Template B	Safaa Dibb	Avenida da Liberdade Viana do Castelo	935999999	safaaaaaaa@gmail.com	FEUA	2016	Mxv, Bragança	2017/2018
AP007	Template C	Octávio Filipe	Avenida 25 de Abril Bragança	915579475	oioi@gmail.com	ISEP	2000	AEIOY, Bragança	2001
AP008	Template C	Teresa Almeida	Travessa de Cima Mangualde	9985115215	escreve.para.mim@gmail.com	FEUP	2006	Info1, Mangualde	2008
AP009	Template C	Xing Ming	Travessa da tristeza Peso da Régua	9564620054	xing.mim@gmail.com	University of Tokyo	2010	Ming, Tokyo	2010

To gauge the algorithm's performance, the generated curricula vitae were subjected to evaluation by the Human Resources department. These synthetic applications were treated as real candidates, and the assessments provided by human evaluators were collected. This step aimed to establish a benchmark against which the algorithm's outcomes could be compared. By contrasting the algorithm's results with the human evaluations, it became possible to identify any disparities or biases in the algorithm's decision-making process.

The comprehensive analysis of the algorithm's performance using synthetic data served multiple objectives. Firstly, it aimed to validate the presence or absence of biases in the coding, with a particular focus on personal attributes such as gender and nationality. This evaluation highlighted areas where the algorithm may exhibit unfair treatment or discrimination. Secondly, it facilitated an assessment of the algorithm's consistency and reliability in evaluating candidates' qualifications, irrespective of their personal characteristics. By comparing the algorithm's outputs with human evaluations, insights could be gained into its overall effectiveness and fairness.

The meticulous examination conducted using synthetic data aimed to address the issue of bias in automated curriculum selection. By delving into the nuances of personal attributes and their potential impact on the algorithm's decisions, the research aimed to promote transparency, accountability, and fairness in the recruitment process. The findings would contribute to the development of more inclusive and unbiased algorithms, fostering a level playing field where candidates are evaluated solely based on their qualifications and merits.

6.6 Results Comparison

As mentioned, to evaluate the accuracy and performance of the developed algorithm, the generated curricula were meticulously reviewed by a Human Resources collaborator with expertise in the field. This collaboration aimed to provide a reliable benchmark for assessing the algorithm's effectiveness. Additionally, the results obtained from one or two paid websites were compared to gain further insights into the algorithm's performance.

The evaluation of the paid application Recruitee highlighted its limitation in automating the comparison between curricula and job opening requirements. In contrast, the CurEval algorithm developed in this research aims to address this limitation by leveraging Natural Language Processing techniques. The CurEval algorithm can analyze the content of curricula, extract relevant information, and compare it against the desired skills and qualifications for the job. This automated approach eliminates the need for manual analysis and significantly reduces the time and effort required by the HR collaborator.

Unlike Recruitee (Figure 6.1), which only displays the listed skills in the curricula, the CurEval algorithm can go beyond surface-level keyword matching. It can understand the context and semantic meaning of the information provided in the curricula, allowing for a more comprehensive evaluation. By using NLP techniques, such as text cleaning, word embedding, and semantic similarity measures, the algorithm can identify the relevance and compatibility of the curricula with the job requirements.

The automation provided by the CurEval algorithm brings several advantages. Firstly, it reduces the HR collaborator's workload by automating the initial screening and classification process. Instead of manually analyzing each curriculum, the algorithm can quickly assess the compatibility and rank them based on their relevance. This significantly speeds up the screening process, allowing the HR collaborator to focus on higher-value tasks.

Secondly, the CurEval algorithm can improve the accuracy and consistency of the screening process. It can analyze a large volume of curricula consistently, ensuring that each application is evaluated based on the same criteria. This reduces the risk of human error and bias that may occur in manual evaluations.

Furthermore, the CurEval algorithm provides a standardized and objective approach to screening curricula. It applies predefined criteria and algorithms to evaluate and compare the qualifications of applicants, ensuring fairness and transparency in the selection process. This helps mitigate potential biases that may arise from subjective manual evaluations.

Overall, the CurEval algorithm developed in this research offers significant improvements over the limited functionality of Recruitee. It automates the comparison between curricula and job opening requirements, reduces manual effort, improves efficiency and accuracy, and provides a standardized and unbiased evaluation process.

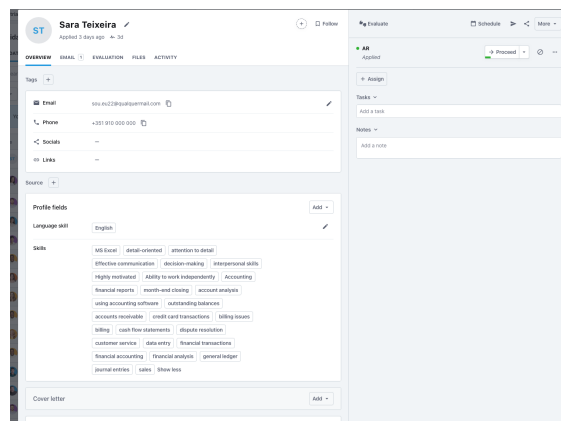


Figure 6.1: Screenshot from one of the synthetic data analysed in Recruitee

The evaluation of the Workable application (Figure 6.2) highlighted its lack of automated candidate evaluation, similar to Recruitee. Despite providing a summary of the information presented in the curricula, the application did not offer automated analysis or comparison against job requirements. The generated summary aimed to assist the HR collaborator by condensing the information for easier review. However, it still required manual analysis of each application, which did not effectively reduce the workload or improve the efficiency of the selection process.

Compared to the limited capabilities of Workable, the CurEval algorithm provides a significant advantage in automating the evaluation of candidates. It reduces manual effort, improves efficiency, and offers a more thorough analysis of curricula based on the desired qualifications. By automating the evaluation process, the CurEval algorithm empowers the HR collaborator to focus on higher-value tasks, such as interviewing and assessing the short-listed candidates, ultimately enhancing the overall recruitment process.

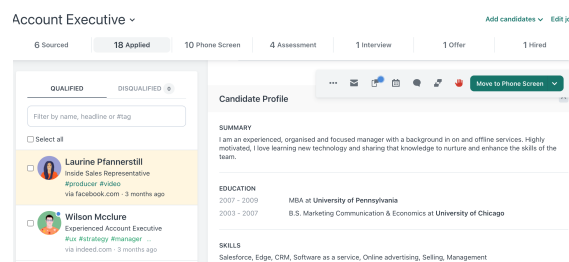


Figure 6.2: Screenshot from one of the synthetic data analysed in Workable

Unfortunately, platforms such as iCIMS, SmartRecruiters, Jobvite, and Lever could not be tested due to the unavailability of free trial options. Hence, a comprehensive comparison of results with these platforms could not be conducted.

Overall, the results indicated that the developed algorithm relied heavily on manual assessment by the Human Resources collaborator. While the algorithm facilitated the analysis of curricula by structuring the data and providing a systematic approach, the final evaluation and classification were still dependent on human judgment. The comparison with paid websites highlighted the limitations of these platforms in providing automated assessment and pre-selection capabilities, as they primarily focused on presenting information rather than performing in-depth analysis or classification.

The detailed examination of the synthetic data using Recruitee and Workable allowed for a

deeper understanding of the platforms' user interfaces and their limited impact on reducing the manual workload of the HR collaborator.

In order to evaluate the relevance and performance of the CurEval algorithm, a comparison was made between the scores assigned by the algorithm and the evaluations conducted by the HR collaborator. The objective was to assess how well the algorithm identified suitable curricula for the given vacancy and to determine the level of consistency between the algorithm's assessments and the human evaluations.

Table 6.3: Comparison between CurEval and HR Collaborator's results

Curriculum code	CurEval	HR Collaborator
AP001	5	9
AP002	5	9
AP003	5	9
AP004	6	11
AP005	6	11
AP006	6	11
AP007	3	5
AP008	3	5
AP009	3	5

Table 6.3 presents the scores given by CurEval and the HR collaborator for each curriculum. The sample data consisted of curricula with unique codes (AP001 to AP009), each having different personal data as described in Table 6.2 and based on the templates from appendix B. CurEval assigned scores to the curricula based on the presence and significance of specific keywords related to the job requirements, while the HR collaborator evaluated them based on her experience and expertise in the field.

The results indicate that the CurEval algorithm successfully identified incompatible curricula for the given vacancy. Specifically, curricula with codes AP007 to AP009 received lower scores from both the algorithm and the HR collaborator. These curricula only partially matched the desired soft skill keywords, suggesting that the algorithm effectively recognized and screened out curricula that did not meet the requirements of the vacancy. This outcome demonstrates the algorithm's capability to save time and effort in the screening process by eliminating unsuitable candidates.

The results provide evidence that there is no apparent bias in the evaluations conducted by either the algorithm or the HR collaborator. The HR collaborator assigned consistent scores to similar templates, even when the personal data was altered. Similarly, the CurEval algorithm assigned identical scores to curricula with similar content, regardless of variations in personal data. These findings indicate that both the algorithm and the HR collaborator's evaluations were fair and not influenced by factors such as gender or nationality. It is important to note that the absence of bias is a crucial aspect in ensuring a fair and equitable evaluation process.

There is a noticeable difference in the scores given by CurEval and the HR collaborator. This disparity can be attributed to the inherent nature of the algorithm itself. CurEval relies on predefined keywords and does not account for synonymous terms or contextual variations. Consequently, if a specific keyword is missing from a curriculum, the algorithm may assign a lower score compared to the HR collaborator. For instance, the HR collaborator might consider curricula with a background in finance and accounting as similar, whereas the algorithm may assign lower scores if certain finance-related keywords are absent. It is important to recognize that this score discrepancy is not indicative of a flaw or bias in the

algorithm but rather a consequence of its design and limitations.

To maintain confidence in the algorithm's performance and ensure its ongoing effectiveness, regular evaluations and comparisons should be conducted. Continuously comparing the algorithm's results with the assessments of the HR collaborator allows for identification of any discrepancies or areas for improvement. This iterative process enables fine-tuning of the algorithm's performance, ensuring its alignment with human evaluations and enhancing its accuracy and reliability over time.

In conclusion, the comparison between the CurEval algorithm and the HR collaborator's evaluations provides valuable insights into its relevance and performance. The algorithm demonstrates its ability to effectively filter out incompatible curricula, thereby streamlining the screening process. Moreover, the absence of bias in both the algorithm and the HR collaborator's evaluations ensures fair treatment of applicants. However, it is important to acknowledge the limitations of the algorithm, such as its reliance on predefined keywords and potential discrepancies in scores. Regular evaluations and improvements will help address these limitations and enhance the algorithm's effectiveness, ensuring its continued alignment with human assessments.

Chapter 7

Conclusions

This dissertation began by explaining the Curriculum Analysis concept and the ethical questions that arise from manual processing. It is possible to understand that the automation of the process will help to overcome these issues and can be used to help businesses pursue the best results in many aspects. It is clear the benefits that bring not only to SEG Automotive Portugal, Unipessoal Lda but also to employees and applicants since it provides faster results with a lower error rate compared with the work done by humans.

As a first step to understand the importance of this process and the steps performed during it, so a small theoretical research was made, using HR literature available.

After understanding the theoretical background, the state of the art was analysed and understood in the present context. There are several commercial solutions that can be used, in the free trial to verify the results but, since the company intended to have a free POC that uses data analysis knowledge and also NLP, the algorithm is going to be made from scratch, as also a prevention of the existence of bias in the data. Recent developments in AI, specifically in the application of NLP and ML, offer promising solutions for automating curriculum analysis. Proposed systems and models, such as the "Resumé Screening" web application by Harsha et al. (2022), CNN-based screening by Mridha et al. (2021b), and ML/NLP techniques in the system proposed by Kinge et al. (2022), demonstrate the potential to streamline the screening process, improve efficiency, and provide valuable insights for candidate selection. Moreover, the importance of addressing bias and ethical considerations in AI-driven recruitment processes is evident in the research, with studies advocating for transparency, monitoring of metrics, diversity training, and human involvement in decision-making. Future directions focus on leveraging advanced technologies, such as game-based interfaces and standardized CV formats, to enhance automated analysis and ensure fairness in hiring practices.

Implementing an automated recruitment system brings various benefits to organizations. Firstly, it enables efficient screening of resumes and job applications, allowing recruiters to process a large volume of candidates in a shorter time frame. This efficiency leads to time savings and allows recruiters to focus on more strategic aspects of the hiring process.

Automation also helps in identifying the most qualified candidates by using effective techniques and algorithms for extracting and analyzing information from resumes and job applications. Natural Language Processing (NLP) and Machine Learning (ML) algorithms can be applied to analyze resumes, extract relevant skills and qualifications, and compare them with job requirements. This automated analysis improves the accuracy of candidate assessment and reduces the chances of overlooking potential talent.

However, there are ethical and legal considerations that organizations must address when using automation in recruitment processes. One key concern is the potential for biases and discrimination. Algorithms used in automated systems may inadvertently reflect biases

present in the data they are trained on, leading to unfair treatment of certain demographic groups. It is crucial for organizations to ensure that their algorithms and models are developed and trained using diverse and representative data, and to regularly monitor and audit their systems for bias.

Transparency is another important ethical consideration. Candidates have the right to understand how their information is being processed and evaluated during the automated recruitment process. Organizations should strive to provide clear explanations of the criteria and algorithms used in the selection process to maintain transparency and trust.

From a legal perspective, organizations need to comply with data protection and privacy regulations when handling candidate data. They must obtain informed consent from applicants, ensure secure storage and transmission of data, and adhere to applicable laws and regulations, such as the General Data Protection Regulation (RGPD). So, while implementing an automated recruitment system offers benefits such as efficiency and improved candidate assessment, organizations must navigate ethical challenges, such as addressing biases, ensuring transparency, and complying with legal requirements. By proactively addressing these considerations, organizations can leverage the advantages of automation while maintaining fairness, transparency, and compliance in their recruitment processes.

The automated system can analyze the content of curricula at a much faster rate compared to manual screening. It can swiftly scan and extract key information such as educational qualifications, work experience, skills, and other relevant attributes. By leveraging NLP algorithms, the system can identify and categorize this information accurately and efficiently. Furthermore, the system can be trained using machine learning algorithms to recognize patterns and prioritize specific qualifications or keywords based on the job requirements. This enables the system to quickly filter out irrelevant curricula and prioritize those that closely match the desired criteria. As a result, the screening process becomes streamlined and expeditious.

The time saved through automated screening allows HR professionals to dedicate their efforts to more strategic aspects of the recruitment process, such as candidate engagement, conducting interviews, and assessing cultural fit. It optimizes the allocation of human resources, leading to a more efficient and effective recruitment workflow.

The manual screening of curricula can introduce ethical concerns and biases that may unfairly impact the selection process. Human evaluators may unintentionally favor or discriminate against certain candidates based on personal biases, leading to potential issues of fairness, diversity, and inclusion.

By leveraging NLP techniques, the automated system can help address these ethical questions. Firstly, the system follows predefined criteria and objective rules consistently when evaluating curricula. It treats all applicants equally, assessing their qualifications based solely on relevant skills, experience, and educational background. This reduces the potential for bias introduced by human evaluators and promotes a fair and unbiased selection process.

Secondly, NLP techniques enable the system to analyze curricula in a standardized manner. The system applies the same evaluation criteria to all applicants, regardless of their personal characteristics, such as gender, race, or background. This mitigates the risk of unconscious bias and ensures that candidates are assessed solely on their qualifications, aligning with principles of equal opportunity and fairness.

Additionally, the use of machine learning algorithms allows for ongoing evaluation and improvement of the automated system. By regularly monitoring and analyzing the system's performance, it becomes possible to identify and address any potential biases or inconsistencies that may emerge over time. This commitment to continuous improvement ensures the system remains fair, transparent, and accountable.

Verifying the results obtained by the CurEval algorithm and the payed solutions have different purposes, the first screens the curricula to find if it has the required keywords, while the second presents only all keywords that are available on the curricula without comparing to the vacancy. So, in this case the CurEval algorithm allows the reduction of time in the pre-screening.

The CurEval algorithm proved the potential for the application of NLP in the pre-screening or Curricula, reducing the duration of the entire recruitment process, fulfilling one of the objectives of this thesis that was studying the potential for the usage of NLP techniques into this process, as well as the potential to expand it to other processes.

An automated solution was developed in PowerAutomate to develop a solution to send the feedback to applicants, and it was tested and in running in the HR collaborator profile.

In conclusion, the application of NLP techniques in automating curriculum screening addresses the research question by significantly reducing the time spent on evaluation while simultaneously minimizing ethical concerns. It empowers HR professionals to focus on higher-level tasks, streamlines the recruitment process, and fosters a fair and inclusive environment by mitigating biases and promoting transparency.

There is further work to be done to improve in the CurEval POC, such as the following:

- **Expansion of Keywords** - The algorithm can be enhanced by incorporating a more extensive list of relevant keywords and synonyms. This would improve the algorithm's ability to identify and classify curricula that align with the vacancy requirements.
- **Machine Learning** - Integrating machine learning techniques can enhance the algorithm's performance and adaptability. By training the algorithm on a larger dataset of classified curricula, it can learn patterns and make more accurate predictions. Machine learning can also help in identifying biases and reducing their impact on the evaluation process.
- **Real-Time Data Integration** - Instead of relying solely on predefined templates, the algorithm can be adapted to process real-time data from applicants' actual curricula. This would require integrating the algorithm into the recruitment process to analyze and evaluate incoming curricula in real-time, providing immediate feedback and reducing manual effort.
- **Multi-Lingual Support** - Expanding the algorithm's capabilities to evaluate curricula written in multiple languages would broaden its applicability and make it suitable for global recruitment processes.
- **Ethical Considerations** - Continuously evaluating the ethical implications of the algorithm is essential. Ensuring compliance with data protection regulations, addressing biases, and maintaining transparency and accountability in the algorithm's decision-making process are ongoing concerns.

By addressing these areas, the CurEval algorithm can evolve into a more sophisticated and accurate tool for automating the screening and evaluation of curricula, improving efficiency, and enhancing the fairness of the recruitment process.

Bibliography

- Akao, Yoji (1990). *Quality Function Deployment (c): Integrating Customer Requirements into Product Design*. Ed. by Elsevier Editora Lda.
- Alsaif, Suleiman Ali et al. (2022). "NLP-Based Bi-Directional Recommendation System: Towards Recommending Jobs to Job Seekers and Resumes to Recruiters". In: *Big Data and Cognitive Computing* 6.4. issn: 2504-2289. doi: 10.3390/bdcc6040147. url: <https://www.mdpi.com/2504-2289/6/4/147>.
- Alsharef, Ahmad et al. (2023). "Exploring the Efficiency of Text-Similarity Measures in Automated Resume Screening for Recruitment". In: *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 36–42.
- El-Ariss, Wafic, Naseem Daher, and Imad H. Elhaji (2021). "To Resume or Not to Resume: A Behavior Tree Extension". In: *2021 American Control Conference (ACC)*, pp. 770–776. doi: 10.23919/ACC50511.2021.9482969.
- Aronson, Elliot, Timothy D. Wilson, and Robin M. Akert (2012). *Social Psychology*. Ed. by Pearson.
- Banaji Mahzarin R. and Greenwald, Anthony G (2013). *Blindspot: hidden biases of good people*. Ed. by Random House Publishing Group.
- Bondielli, Alessandro and Francesco Marcelloni (2021). "On the use of summarization and transformer architectures for profiling résumés". In: *Expert Systems with Applications* 184, p. 115521. issn: 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2021.115521>. url: <https://www.sciencedirect.com/science/article/pii/S0957417421009301>.
- Burdick, Annie (2021). *Unconscious Bias - Everything You Need to Know About Our Hidden Prejudices*. Ed. by Publicações Dom Quixote.
- Camara, Pedro B., Paulo Balreira Guerra, and Joaquim Vicente Rodrigues (1997). *Humanator Recursos Humanos e Sucesso Empresarial*. Ed. by Publicações Dom Quixote.
- Cambridge English Dictionary* (n.d.). <https://dictionary.cambridge.org/dictionary/english/>. Accessed: 2023-01-29.
- Cardoso, Adelino Alves (2005). *Recrutamento & Seleção de Pessoal*. Ed. by Editora RH.
- Chiavenato, Idalberto (2010). *Gestão de Pessoas*. Ed. by Productivity Press.
- Consul, Nikita et al. (2021). "Mitigating Unconscious Bias in Recruitment and Hiring". In: *Journal of the American College of Radiology* 18.6. Focus on Private Practice, pp. 769–773. issn: 1546-1440. doi: <https://doi.org/10.1016/j.jacr.2021.04.006>. url: <https://www.sciencedirect.com/science/article/pii/S1546144021003227>.
- Cooper, Harris, Larry V. Hedges, and Jeffrey C. Valentine (2009). *The Handbook of Research Synthesis and Meta-Analysis*. Ed. by Russell Sage Foundation Publications.
- Datalabor - DADOS ESTATÍSTICOS E INFORMAÇÃO JURÍDICA NAS ÁREAS DO TRABALHO, EMPREGO E PROTEÇÃO SOCIAL* (n.d.). <https://datalabor.pt/>. Accessed: 2023-02-19.
- Fuller, Pamela, Mark Murphy, and Anne Chow (2020). *The Leader's Guide to Unconscious Bias*. Ed. by Simon & Schuster.

- Harsha, Tumula Mani et al. (2022). "Automated Resume Screener using Natural Language Processing(NLP)". In: *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1772–1777. doi: 10.1109/ICOEI53556.2022.9777194.
- Harvard Business Review* (n.d.). <https://hbr.org/>. Accessed: 2023-02-05.
- Hauser, J. R. and D. Clausing (1988). "The house of quality". In: *Harvard Business Review* 63, pp. 63–73.
- iCIMS* (n.d.). <https://www.icims.com/>. Accessed: 2023-02-07.
- JazzHR* (n.d.). <https://www.jazzhr.com/>. Accessed: 2023-02-07.
- Jobvite* (n.d.). <https://www.jobvite.com/>. Accessed: 2023-02-07.
- Kahneman, Daniel (2011). *Thinking, Fast and Slow*. Ed. by Straus Farrar and Giroux.
- Kinge, Bhushan et al. (2022). "Mitigating Unconscious Bias in Recruitment and Hiring." In: *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*.
- Kleinberg, J. et al. (2020). "Bias in hiring algorithms: Evidence from field experiments". In: *Proceedings of the National Academy of Sciences*. Vol. 117. 23, pp. 13343–13348. doi: 10.1145/3351095.3372828.
- Köchling, Alina and Marius Claus Wehner (2020). "Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development". In: *Business Research* 13.3, pp. 795–848. doi: 10.1007/s40685-020-00134-w. url: <https://doi.org/10.1007/s40685-020-00134-w>.
- Lever* (n.d.). <https://www.lever.co/>. Accessed: 2023-02-07.
- Martins, Dora and Rita Moreira Cruz (2019). *Gestao do Talento em Organizações da Península Iberica*. Ed. by Editora RH.
- Mridha, M.F. et al. (2021). "A Machine Learning Approach for Screening Individual's Job Profile Using Convolutional Neural Network". In: *2021 International Conference on Science & Contemporary Technologies (ICSCT)*, pp. 1–6. doi: 10.1109/ICSCT53883.2021.9642652.
- Myers, David G. (2009). *Social Psychology*. Ed. by McGraw-Hill.
- Oxford dictionary* (n.d.). <https://www.oxfordlearnersdictionaries.com/>. Accessed: 2023-01-29.
- Ozimek, Adam (2020). "The Future of Remote Work". In: *Harvard Business Review*, p. 7. doi: 10.2139/ssrn.3638597.
- Peppers, Ken et al. (2007). "A Design Science Research Methodology for Information Systems Research". In: *Journal of Management Information Systems*, pp. 45–77. doi: 10.2753/MIS0742-1222240302.
- Pimpalkar, Amit et al. (2023). "Job Applications Selection and Identification: Study of Resumes with Natural Language Processing and Machine Learning". In: *2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1–5. doi: 10.1109/SCEECS57921.2023.10063010.
- Raghavan, Manish et al. (2020). "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. FAT* '20. Barcelona, Spain: Association for Computing Machinery, pp. 469–481. isbn: 9781450369367. doi: 10.1145/3351095.3372828. url: <https://doi.org/10.1145/3351095.3372828>.
- RGPD* (n.d.). https://commission.europa.eu/law/law-topic/data-protection_en. Accessed: 2023-06-25.
- RGPD in all EU languages* (n.d.). <https://eur-lex.europa.eu/eli/reg/2016/679/oj>. Accessed: 2023-06-25.

- Roxburgh, E. and K. Hansen (2015). "Bias in Recruitment and Selection". In: *Centre for Labour, Employment and Work* 0, pp. 1–3. url: https://www.wgtn.ac.nz/__data/assets/pdf_file/0008/1816604/bias-in-recruitment-and-selection.pdf.
- Sanyal, Satyaki et al. (Mar. 2017). "Resume Parser with Natural Language Processing". PhD thesis. doi: 10.13140/RG.2.2.11709.05607.
- Several (2009). *Recruitment and Selection Revised Edition*. Ed. by Rev. ed. Amsterdam : Routledge.
- Sharot, Tali (2011). *The optimism bias: a tour of the irrationally positive brain*. Ed. by Knopf Doubleday Publishing Group.
- SmartRecruiters (n.d.). <https://www.smartrecruiters.com/>. Accessed: 2023-02-07.
- Society for Human Resource Management (n.d.). <https://www.shrm.org/>. Accessed: 2023-01-29.
- Society for Human Resource Management (n.d.). https://www.shrm.org/hr-today/news/hr-news/pages/cms_017619.aspx. Accessed: 2023-02-05.
- Soleimani, Melika et al. (2021). "Cognitive biases in developing biased Artificial Intelligence recruitment system". In: *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 9. doi: 10.24251/HICSS.2021.620.
- Spears, T. and W. Schmader (2021). "Bias in Recruitment and Selection". In: *Physician Leadership Journal* 0, pp. 1–3. url: <https://www-test.physicianleaders.org/articles/unconscious-bias-healthcare>.
- Sudha, G. et al. (2021). "Personality Prediction Through CV Analysis using Machine Learning Algorithms for Automated E-Recruitment Process". In: *2021 4th International Conference on Computing and Communications Technologies (IC CCT)*, pp. 617–622. doi: 10.1109/IC CCT53315.2021.9711787.
- Workable (n.d.). <https://www.workable.com/>. Accessed: 2023-02-07.

Appendix A

Appendix A



For more than a century, we have been shaping progress in our product sector as technology leader – whether it's starter motors, generators, start/stop systems or mild hybridization. Innovation, development expertise and the highest quality standards make us a reliable partner for automotive manufacturers – at 16 locations worldwide. Our products are not only efficient and durable, but also provide high functional flexibility and the right answers for global product platforms and more stringent CO₂ requirements.

Accounts Payable Analyst

Join us to Work for shaping the future of mobility. And yours.

You are there when we work together to shape the future of mobility! Become part of our interdisciplinary, international team at the location in Vila do Conde, Portugal. SEG Automotive Portugal, Unipessoal Lda. is the SEG group's Shared Service Center.

How your future looks like with us

- Responsible for process, tracking, and recording expenses and payments in an accurate, efficient, and timely manner;
- Assure all issues related to invoicing, account analysis, and reconciliations in the organization, producing the necessary reports;
- Ensure the implementation of Detective Controls and preventive measures;
- Ensure all audit and compliance obligations related to the function;
- Manage all relations and communications with vendors;
- Analyze the existing processes to find opportunities for improvement to be implemented.

What you bring with you for a future with us

- Bachelor or Post-Graduate in Economics, Management, or Accounting;
- The domain of MS Office tools;
- Fluent in English (written and spoken) – mandatory;
- Soft skills: Excellent Communication Skills (Customer Service Skills) - Structured approach; Problem-solving skills; Good organizational and prioritization skills; Team player, Good presentation skills.

What we will offer you in the future

- The opportunity to work with a lot of innovation, the scope for design, and responsibility. This is supported by attractive additional benefits, such as flexible working hours and a hybrid working model.

Please send us your CV and tell us why do you want to join SEG Automotive Portugal

Human Resources: sgpt.recrutamento@seg-automotive.com

Visit our website: www.seg-automotive.com

Reference code SGPT-Vc-FIN-2023-09



For more than a century, we have been shaping progress in our product sector as technology leader – whether it's starter motors, generators, start/stop systems or mild hybridization. Innovation, development expertise and the highest quality standards make us a reliable partner for automotive manufacturers – at 16 locations worldwide. Our products are not only efficient and durable, but also provide high functional flexibility and the right answers for global product platforms and more stringent CO₂ requirements.

Professional Internship | Purchasing

Join our **Great Place to Work** for shaping the future of mobility. And yours.

You are there when we work together to shape the future of mobility! Become part of our interdisciplinary, international team at the location in Vila do Conde, Portugal. SEG Automotive Portugal, Unipessoal Lda. is the SEG group's Shared Service Center.

How your future looks like with us

- Create purchase orders according to the defined processes and procedures;
- Contact internal business partners to clarify doubts and correct errors;
- Support the development of reports and analyses related to ongoing processes and activities, providing information to support decision-making processes;
- Support in creating and updating data in ERP;
- Support the development of documentation for concept implementation for the PUS department;
- Support procurement activities for different regions for both Direct and Indirect Purchasing

What you bring with you for a future with us

- Degree in Economics or similar area;
- Good command of English language spoken & written (essential), German or Spanish language is a plus;
- Soft skills: Excellent Communication Skills (Customer Service Skills) - Structured approach; Problem solving skills; Good organizational and prioritization skills; Team player, Good presentation skills;
- Eligible for IEFP internship.

What we will offer you in the future

- The opportunity to work in an **great place to work**, with a lot of innovation, the scope for design, and responsibility. This is supported by attractive additional benefits, such as flexible working hours and a hybrid working model.

Please send us your CV and tell us why do you want to join SEG Automotive Portugal

Human Resources: sgptrecrutamento@seg-automotive.com

Visit our website: www.seg-automotive.com

Reference code SGPT-Vc-PUS-2023-13



For more than a century, we have been shaping progress in our product sector as technology leader – whether it's starter motors, generators, start/stop systems or mild hybridization. Innovation, development expertise and the highest quality standards make us a reliable partner for automotive manufacturers – at 16 locations worldwide. Our products are not only efficient and durable, but also provide high functional flexibility and the right answers for global product platforms and more stringent CO₂ requirements.

Professional Internship | Accounts Receivable

Join our **Great Place to Work** for shaping the future of mobility. And yours.

You are there when we work together to shape the future of mobility! Become part of our interdisciplinary, international team at the location in Vila do Conde, Portugal. SEG Automotive Portugal, Unipessoal Lda. is the SEG group's Shared Service Center.

How your future looks like with us

- Development and reporting of relevant KPIs (Key Performance Indicators) related to Order to Cash performance.
- As part of the Accounts Receivables team be responsible for the following tasks:
- Extract from SAP Customer accounts status (open and due items) and prepare the weekly and monthly trade receivable outstanding KPI.
- Collect and prepare data related to customer deductions.
- Prepare and track with team Trade receivables aging KPI.
- Manage Accounts Receivables mailbox.
- Support other areas with the following tasks:
- Extract from ERP demands and/or deliveries and prepare sales prognosis reports for sales and other areas.
- Monitor and report EDI unprocessed iDocs.
- Support in the continuous improvement, namely with the roll-out of automation solutions.
- Support Sales Business Frames (SBF) preparation.

What you bring with you for a future with us

- Degree in Economics, Finance, Management or similar areas;
- No more than 1 year of working experience;
- Domain of Microsoft operating systems;
- Good command of English language spoken & written;
- Very good IT-skills;
- Soft skills: Excellent Communication Skills (Customer Service Skills) - Structured approach; Problem solving skills; Good organizational and prioritization skills; Team player, Good presentation skills.

What we will offer you in the future

- The opportunity to work in a **great place to work**, with a lot of innovation, the scope for design, and responsibility. This is supported by attractive additional benefits, such as flexible working hours and a hybrid working model.

Please send us your CV and tell us why do you want to join SEG Automotive Portugal

Human Resources: sgpt.recrutamento@seg-automotive.com

Visit our website: www.seg-automotive.com

Reference code SGPT-Vc-SCC-2023-16

Appendix B

Appendix B

The first Template for the vacancy is:

Your Name
Address
City, State, ZIP
Phone Number
Email Address

Objective:

Highly motivated and detail-oriented professional with a strong background in accounts receivable seeking a challenging position to contribute my skills and expertise in ensuring accurate and timely processing of financial transactions.

Education:

- Bachelor of Science in Accounting
University Name, Year

Skills:

- Proficient in accounts receivable processes and procedures
- Strong knowledge of financial accounting principles and practices
- Excellent attention to detail and accuracy in data entry and record-keeping
- Proficient in using accounting software and MS Excel for financial analysis
- Effective communication and interpersonal skills for client interactions and dispute resolution
- Ability to work independently and collaboratively in a team environment

Experience:

Accounts Receivable Specialist
Company Name, City, State
Dates

- Processed incoming payments, including checks, credit card transactions, and electronic transfers, accurately applying them to customer accounts
- Generated and distributed customer invoices and statements on a regular basis
- Investigated and resolved any discrepancies or billing issues in a timely manner

- Reconciled accounts receivable records with the general ledger and performed regular account analysis
- Communicated with customers to collect outstanding balances and resolve payment disputes
- Prepared various financial reports, including aging reports and cash flow statements, to support management decision-making
- Collaborated with the sales and customer service teams to ensure accurate and timely billing and collection processes
- Assisted in the month-end closing process, including preparing journal entries and reconciling accounts

References:

Available upon request

The second template is:

Your Name
Address
City, State, ZIP
Phone Number
Email Address

Objective:

Results-driven and detail-oriented professional with a strong background in accounts receivable and a solid understanding of economic principles. Seeking a challenging position in accounts receivable to apply my expertise in financial analysis and contribute to efficient financial operations.

Education:

- Bachelor of Science in Economics
University Name, Year

Skills:

- Proficient in accounts receivable processes and procedures
- Strong knowledge of economic theories, principles, and analysis
- Excellent attention to detail and accuracy in financial record-keeping
- Proficient in using accounting software and MS Excel for financial analysis
- Understanding of economic indicators and their impact on financial operations
- Effective communication and interpersonal skills for client interactions and dispute resolution
- Ability to work independently and collaboratively in a team environment

Experience:

Accounts Receivable Specialist
Company Name, City, State
Dates

- Processed incoming payments, including checks, credit card transactions, and electronic transfers, accurately applying them to customer accounts
- Conducted financial analysis to identify trends, risks, and opportunities related to accounts receivable
- Collaborated with cross-functional teams to optimize financial operations and ensure timely collections
- Investigated and resolved any discrepancies or billing issues in a timely manner
- Prepared regular reports on accounts receivable performance and presented findings to management
- Communicated with customers to collect outstanding balances and resolve payment disputes
- Assisted in the development and implementation of credit and collection policies
- Participated in the month-end closing process, including preparing journal entries and reconciling accounts

Relevant Coursework:

- Macroeconomics
- Microeconomics
- Financial Economics
- Econometrics
- Money and Banking

References:

Available upon request

And the third template is:

Your Name
Address
City, State, ZIP
Phone Number
Email Address

Objective:

Highly skilled and analytical professional with expertise in data mining and machine learning. Seeking a challenging position in Data Mining to leverage my advanced statistical knowledge, programming skills, and problem-solving abilities to extract actionable insights from complex datasets.

Education:

- Master of Science in Data Science
University Name, Year

Skills:

- Proficient in data mining techniques, including data cleaning, preprocessing,

and feature selection

- Strong knowledge of statistical analysis and predictive modeling
- Experience with machine learning algorithms and data visualization tools
- Proficient in programming languages such as Python and R for data manipulation and analysis
- Familiarity with database systems and SQL for querying and managing large datasets
- Strong problem-solving skills and attention to detail in identifying patterns and trends in data
- Excellent communication and presentation skills for effectively conveying complex findings to non-technical stakeholders

Experience:

Data Mining Analyst

Company Name, City, State

Dates

- Conducted data mining and analysis on large datasets to identify patterns and trends
- Developed and implemented data cleaning and preprocessing techniques to improve data quality
- Applied statistical analysis and predictive modeling to extract insights and make data-driven recommendations
- Collaborated with cross-functional teams to define business objectives and design data mining strategies
- Created and maintained data pipelines for efficient data extraction, transformation, and loading
- Developed machine learning models for predictive analytics and anomaly detection
- Prepared and presented reports and visualizations to communicate findings and recommendations
- Stayed up-to-date with the latest data mining techniques and tools to drive continuous improvement

Projects:

- List any relevant data mining projects you have worked on, highlighting the techniques used and the outcomes achieved

Technical Skills:

- Programming: Python, R, SQL
- Data Mining Tools: scikit-learn, TensorFlow, PyTorch, RapidMiner, Weka
- Statistical Analysis: Regression, Clustering, Classification, Time Series Analysis
- Data Visualization: Matplotlib, Seaborn, Tableau

References:

Available upon request