RESUME PARSING USING NER MODEL

Article

2024

Presented by:

Amina Safarish

Abstract:

With multiple applicants competing for one employment opportunity in an organization or firm, it is expected that ten to fifteen might compete for the one placement. The process begins through reviewing the résumé of the candidates, ensuring that only the best remain in the competition. Given the complexity and volume that comes with having many parts and formats, reviewing many résumés is thought to be laborious. To solve these issues this project employs a natural processing technique called Named Entity Recognition (NER) as a core function.

NER models are developed to extract key entities of skills, courses, certifications, and professional experiences and classify them by type. Automating the extraction and categorization of relevant information by the system, the raw data from résumés is organized in a structured format. This structured format allows recruiters to make easier assessments of candidates' qualifications and identify the most relevant profiles based on job requirements.

In addition, the system delivers useful insights by suggesting related skills, courses, or fields to the job or organizational needs. This, in turn, means more targeted recruitment with reduced effort on manual processes while also improving the accuracy of the decision-making process. Incorporation of NER models into résumé parsing improves the overall recruitment process, thereby making the assessment of candidates more efficient and likely to match qualified ones with the right jobs.

Introduction:

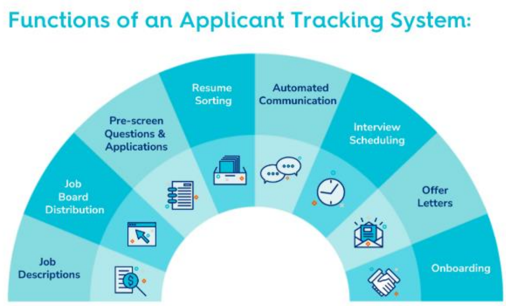
Named Entity Recognition (NER) models *(Shelar et al., 2020)* emerge as effective tools within the domain of Natural Language Processing (NLP) *(Kang et al., 2020)*. It focuses on identifying and classifying entities from text to predefined categories, including but not limited to names of people, organizations, location, dates, and even numerical data.

Through this extraction and structuring process of information from unorganized text data, NER models play a very important role in automating complex data management tasks *(Mierswa et al., 2006)*. This capacity is of immense value in recruitment because it helps one go through candidatures' résumés in a very efficient and structured manner.

One of the most important benefits of using NER models is their ability to automate the early stages of candidate evaluation *(Won et al., 2018)*. This means that raw textual data from résumés can be converted into structured information, freeing human resource teams to spend more time on higher-level tasks such as interviews and organizational fit assessments rather than spending hours manually reviewing résumés.

NER models also bring consistency and fairness into the recruitment process *(BAUER et al., 2001)*. The same criteria and extraction techniques are applied to all résumés, minimizing biases associated with manual screening. Moreover, the standardization provided by NER ensures uniform evaluation even when multiple recruiters are involved in assessing candidates for the same position in large organizations *(Wierzbicki et al., 2024)*.

The integration of NER models into Applicant Tracking Systems (ATS) further enhances the recruitment process *(Kinger et al., 2024)*.



This end-to-end automation also streamlines résumé parsing and makes possible categorization and ranking of candidates by relevance to the job *(Kinger et al., 2024b)*. Improving organization and automation of candidate data management increases the overall efficiency of the hiring process in NER-powered ATS systems *(Md. Sayeduzzaman & Mahmud, 2022)*. The organizations, in this manner, will be able

to identify the qualified candidates better and reduce the time and resources in selection.

Objectives

The primary objective of this project is to automate the process of extracting candidate data from résumés *(Roy et al., 2020)*, significantly reducing the manual effort that is involved in the recruitment process. Traditional résumé screening requires recruiters to spend hours reviewing and comparing résumés *(Cole et al., 2007)*, which is not only time-consuming but also prone to human error. By introducing an automated system (Hansen & Atkins, n.d.), the project aims to streamline this task, ensuring a faster, more accurate selection process.

Named Entity Recognition (NER)

Model A central feature of this system is the utilization of an NER model for ranking applicants based on preset criteria *(Roy et al., 2020)*. The NER model identifies and classifies the most important résumé elements, which would be skills, certifications, work experience, and education *(Baigang & Yi, 2022)*. These extracted data points are then analyzed and given a score, allowing the system to rank candidates in regard to their suitability for a given role. This ranking mechanism allows the recruiter to find the best candidates more rapidly and accurately, locating the right human resources in an effective manner. Such technology not only reduces hiring time but also improves the accuracy and fairness of the recruitment process *(Wosiak, 2021),* making it easier for organizations to match applicants with appropriate roles based on their qualifications and relevance to the job requirements *(Pudasaini et al., 2021).*

Background

The recruitment process has become highly digitalized *(Racano, 2020)*. The process of dealing with resumes now requires a more efficient, accurate, and technologically advanced method of handling. The large number of applications that organizations receive makes the traditional method of reviewing resumes manually inadequate *(Maree et al., 2018)*. The diversity of resume formats and the urgency for rapid and effective hiring decisions have compounded this problem *(Lavigna & Hays, 2004)*. To overcome such challenges, resumed parsing using the advanced technologies has emerged as a critical need *(Brindashree & Pushphavathi, 2023)*. The technique of named entity recognition *(NER),* which comes under the wing of NLP *(Dai et al., 2014),* plays a paramount role in streamlining the work of retrieving relevant information in resumes and categorizing, organizing, and structuring it accordingly.

With advancements in the online job application systems *(Harriet Rodney, 2019)*, the number of resumes applied for each recruitment has been increasing. Manual processing becomes cumbersome and inefficient, but the influx of applications should not be a reason to lose focus on evaluating all applicants. Automated resume parsing offers a way to cope with this flow so recruiters will not be burdened with so many applications to filter out. NER models can particularly be very effective for processing different resume formats to extract key information from a resume, including personal details, educational background, work experience, skills, and other relevant qualifications *(Derous & Ryan, 2018)*.

The integration of NER into resume parsing also addresses the need for consistency and objectivity in the recruitment process *(Mashayekhi et al., 2024)*. Applying the same set of criteria to all resumes, NER models eliminate biases that could arise from manual screening. Moreover, the speed with which NER models extract and organize data allows recruiters to make faster, more informed decisions, reducing preliminary screenings time *(J. Himabindu Priyanka & Parveen, 2023)*.

Moreover, rapid progress in machine learning and artificial intelligence has improved the accuracy and efficiency of NER systems. As these models are exposed to more data, they are able to continually improve, refining their ability to recognize and categorize the most relevant information. Overall, the use of NER for resume parsing not only streamlines the recruitment process but also helps organizations make better hiring decisions by ensuring that they focus on the most qualified candidates.

Table 1: Types of Resumes and Their Features

Resumes are critical tools through which job seekers present their qualifications, experience, and skills in a structured and professional format. Knowing the different types of resumes and their characteristics helps both applicants and recruiters set their expectations. Resumes exist in various formats to match different career scenarios, such as specific strengths and mitigated weaknesses in a candidate's profile (*Ingold & Langer, 2021)*.

Resume Format	Characteristics	Common Use Cases
Chronological	Lists work experience in reverse chronological order	Preferred by professionals with a steady work history
Functional	Focuses on skills and experience rather than work history	Suitable for career changers or those with gaps in employment
Combination	Blends elements of chronological and functional formats	Ideal for highlighting specific skills and experiences
Targeted	Tailored to a specific job or industry	Effective for specialized roles

This is especially the case with resume parsing because it allows the model to use Named Entity Recognition (NER) models and customize its terms and requirements according to the specific terminologies in different industries *(Roy, 2021)*. Different industries use different keywords, jargon, and certifications. For instance, in the healthcare industry, resumes may contain specialized terms relating to medical qualifications, procedures, or certifications that may not be present in other industries, such as technology or marketing. Similarly, healthcare-specific certifications and licenses, like those from medical boards or healthcare institutions, would have to be recognized and correctly extracted by the NER model *(Ingold & Langer, 2021)*.

This is how by tailoring the NER model to recognize and classify these industry-specific terms, the system can parse the resume efficiently and make sure that only the most relevant information is extracted *(Roy, 2021)*. Thus, this customization improves the accuracy of the resume parsing process as the model is better suited for identifying and processing the specialized language of each sector. Ultimately, this responsiveness ensures that resumes are accessed according to the most applicable criteria for each industry; employers find the best candidate within their specific needs without getting through irrelevant data *(Roy, 2021)*.

Methodology

This section outlines the methodology followed in training, testing, and evaluating the Named Entity Recognition (NER) model for parsing resumes. The process involves several stages, including data collection, pre-processing, model selection, training, entity identification, and evaluation. Each stage ensures that the NER model effectively recognizes and categorizes relevant information from resumes in various formats.

1. Data Collection

The first step in the methodology is to gather a diverse collection of resumes. This collection is important for training a robust NER model that can handle different formats and types of resumes. Resumes are sourced from multiple industries, ensuring that the model is exposed to a variety of terminology and entity types. Resumes collected are mainly in digital format, which include both Microsoft Word and PDFs, since these are the most commonly used resume formats during job applications. The existence of two formats accounts for variance in structure or type of content embedded between two formats. Other than resumes, data is also fetched from online sources such as job portals and Web 2.0 content, giving even more examples of non-traditional resumes in HTML format. The diversity of data ensures that the model is exposed to a wide range of terminology, skills, qualifications, and experience levels *(PETERSON et al., 2001)*.

2. Pre-processing

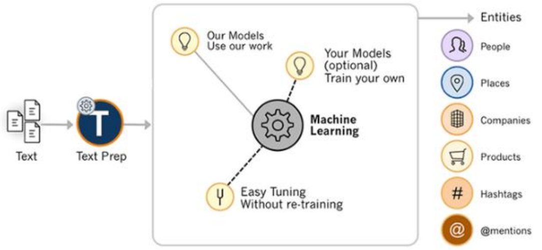
After gathering the resumes, they are pre-processed to convert them into a standardized format. This step removes all the extraneous formatting like fonts, bullet points, or tables, and leaves only the raw textual data. It is important to ensure consistency across different resume formats. Techniques such as tokenization, stemming, and lemmatization are also applied to the text to improve the model's ability to recognize key entities, regardless of their original form in the resumes. The pre-processed data is now ready to be fed into the NER model for training and testing.

3. NER Model Selection

The next step is the selection or training of an appropriate NER model. Depending on the requirements of the task, there are many NER models that can be used. Some of the most popular ones are transformer-based models like BERT (Bidirectional Encoder Representations from Transformers), or pre-trained models available in NLP libraries such as spaCy *(Alessandro Fantechi et al., 2021)*. In this methodology, both BERT-based and spaCy pre-trained models are considered, as they have robust performance and are well-known in the NLP community *(Alessandro Fantechi et al., 2021)*. This means BERT is particularly useful in the identification of named entities within complex text structures, like resumes. On the other hand, spaCy pre-trained models provide an efficient, easy-to-implement solution that is highly optimized for real-time text processing *(Wang et al., 2022)*. The choice between the two models depends on various conditions, including dataset size, power of computation, and needs or requirements of the resume-parsing task.

4. Model Training

To prepare the NER model for training, a labeled dataset is required that annotates specific entities within the resumes (Tarcar et al., 2020).



These entities may include names, skills, educational qualifications, work experience, certifications, and other critical information. To train the model, a labeled dataset is created. In this dataset, each entity in the resume is identified and marked with its corresponding label. For example, names are labeled as "PERSON," educational qualifications as "EDUCATION," and skills as "SKILLS." This labeled data is used to train the NER model with the aim of teaching it to identify these entities automatically from new, unseen resumes. The training procedure involves passing the labeled data into the model, adjusting weights and parameters so that the model is fit better, and minimizing prediction errors *(Shelar et al., 2020b).*

5. Entity Types

Once the NER model is trained, it is essential to specify the types of entities to extract from the resumes. The choice of entities is driven by the objective of the resume parsing task and the relevance of the data to the hiring process. Key entity types typically include:

- Personal information (e.g., candidate's name, contact information)
- Educational background (e.g., degrees, certifications)
- Work experience (e.g., previous job titles, companies)
- Skills and qualifications (e.g., technical skills, languages spoken)
- Certifications (e.g., industry-specific qualifications)
- Projects and publications (e.g., academic work, portfolio)

Defining the entity types helps to ensure that the NER model is fine-tuned in such a way that it learns to focus on the relevant data extraction only, and thus, make the parsing more efficient. Customizing the model toward these specific types of entities helps eliminate any irrelevant data, thus further improving the overall quality and relevance of the extracted data.

6. Evaluation

To finalise the methodology, evaluation for performance has been done against a trained NER model. Here, the method used multiple test datasets apart from training datasets to generalize the performance of the models over unseen data *(Shelar et al., 2020)*. All tests used for the evaluation followed some appropriate metrics:

Accuracy: *Measures the correctness of the model, which is the ability of the model to identify related entities. It calculates as the ratio of true-positive entities to all the predicted relevant entities.*

Recall*: This measures the ability of the model to identify the relevant entities. It determines the ratio of true-positive cases identified by the model, out of all the instances of relevant entities in a dataset.*

F1 Score*: The F1 score is the harmonic mean of precision and recall, which produces a single measure of them combining both.*

It helps to determine the quality of the NER model with the metrics calculated by analysis. In other words, precision, recall, and F1 score measure how well a model retrieves data related to and comprehensive in resume parsing. As the need for improvement occurs, further tuning of parameters, more training, and refining of labeled datasets are some of the activities that solve these deficiencies *(Wang et al., 2022)*. The outcome of NER model development improves through the continuous process of testing and refining the resumes and enhancing recruitment efficiency.

Table 2: Summary of NER Models and Their Characteristics

Model	Architecture	Dataset
spaCy	Transformer-based	Custom Resumes
BERT	Transformer-based	Large Dataset
Custom	Neural Network	Industry-Specific

The table offers a comparative overview of different NER models, used for resume parsing, which focuses on the architecture, type of dataset, and customizability of each model.

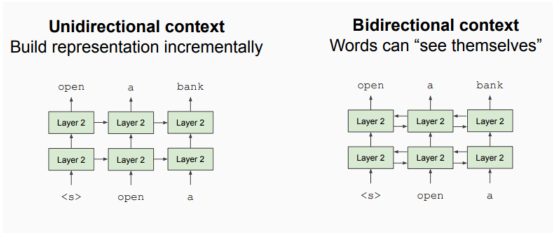
Precision	Recall	F1 Score
0.92	0.89	0.90
0.95	0.93	0.94

spaCy

SpaCy is an NER model based on transformer that is famous for versatility and efficiency in natural language processing tasks *(Mierswa et al., 2006)*. It especially thrives on applications demanding high accuracy at moderate computation cost. In this scenario, spaCy has been used in a custom dataset on resumes with diverse forms and structures of resumes. With the custom dataset, this ensures that the model is honed to pull out relevant information accurately such as names, skills, and qualifications. Its seamless integration makes it an ideal fit for integrating NER functionalities in most ATS *(Wang et al., 2022)*.

BERT

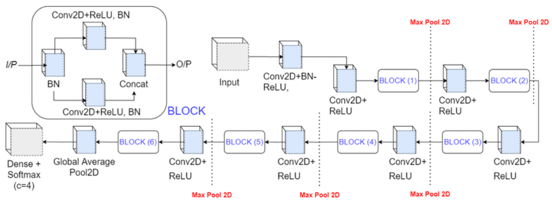
The Bidirectional Encoder Representations from Transformers (BERT) model is one of the newest approaches in NLP. It uses deep learning and contextual understanding to achieve excellent performance on complex text processing tasks *(Mierswa et al., 2006)*. BERT's transformer architecture enables it to understand the context of an entire sentence, making it quite useful for handling the variation of resume content. BERT is trained on a huge dataset, and it can generalize across different industries and resume formats. It gives robust results even for unseen data, and the computational intensity of BERT makes it suitable for large-scale recruitment processes where precision and recall are critical.



Custom Neural Network

A custom model based on a neural network is developed with industry-specific applications in mind. This model, unlike the pre-trained models BERT or spaCy, is built specifically to cater to the terminology, entities, and patterns of a specific industry. For instance, a resume related to health care, engineering, or finance contains specific domain certifications and terminologies *(Md. Sayeduzzaman & Mahmud, 2022)*. By training this custom model on an industry-specific dataset, it performs better at recognizing and categorizing relevant entities so that the parsing process matches industry standards more closely.

Each model offers distinct advantages, with the choice depending on organizational needs, dataset characteristics, and computational resources.



Results and Discussion

Active engagement via participatory activities

Active participation in both project-based and problem-based learning creates a higher involvement and interest in the process by students *(Mierswa et al., 2006).* Both project-based and problem-based learning methodologies allow the transformation of passive learners to active participants because their essence lies in interactive and dynamic processes rather than just memorization.

Key Findings

Hands on Activities: Activities like creating a model, experimentation, and real-world simulation can deeply engross a student in his subject of interest. Thus, forming a mini-business plan or solving the problem of a real world logistics issue that takes all aspects from thought to execution.

Freedom to Explore: When students are given the freedom to explore different approaches to solving a problem, they feel a sense of ownership *(Md. Sayeduzzaman & Mahmud, 2022)*. This personal connection encourages more thoughtful participation, as students feel their contributions matter.

Longer Attention Span: Compared to traditional lectures, students in active learning environments exhibit longer attention spans. The dynamic nature of PBL ensures students are constantly engaged with diverse challenges, preventing monotony.

Most of the students reported satisfaction when they participated fully in determining the outcomes of their learning, fulfilling them beyond mere grades.

Development of Practical Problem-Solving Skills

One of the major advantages of these methodologies is that they have the capacity to improve students' practical problem-solving skills drastically. Both approaches challenge the student to look at problems from all angles, break them down systematically, and get solutions.

Detailed Insights

Critical Thinking: Students are educated to question assumptions, examine variables, and predict probable outcomes. For instance, while designing eco-friendly packaging, students analyzed environmental impact, cost, and usability before settling on a solution.

Strategic Thinking: In complex problems, the student is required to develop a strategy. This often requires group brainstorming, logical reasoning, and decision-making under constraints.

Resourcefulness: This helps them make the most out of the available materials and information, especially in situations where resources may be scarce. The student also becomes more effective at seeking outside help or experts to complement the learning.

The iterative cycle of trying, failing, and then re-evaluating the solution helps students internalize problem-solving as a practical skill rather than just a theoretical concept.

Fostering Innovation and Creativity

In PBL and Problem-Based Learning, students are engaged in activities that can spawn creativity. Since the spotlight is not on prewritten answers, students are freed to innovate and experiment.

Detailed Applications

Freestyle Implementation: Students are faced with open-ended tasks on most occasions. For example, taking up a community project where waste management needs to be improved can lead to solutions that go from developing a recycling-based app using technology to promoting awareness through on-ground drives.

Encouragement to Take Risks: Unlike traditional methods where incorrect answers are penalized, these models encourage students to take calculated risks. This removes the fear of failure and unlocks their creative potential.

Exposure to Multidisciplinary Thinking: Projects often combine elements from various subjects, encouraging students to draw on knowledge from science, arts, and humanities to create holistic solutions.

Many students reported surprise over their own ability to be creative in thought when challenged to brainstorm solutions without much constraint.

Build teamwork and collaborative skills.

As part of Project-Based and Problem-Based Learning, cooperation becomes vital. Through group projects, group discussions, and, in some cases, family living, students learn essential skills for professional and personal advancement.

Deeper Understanding

Conflict Resolution: Conflicts, which are bound to come within group dynamics, become a normal tool for the development of an inquiring mind. In such, how to mediate and search for common ground will be second nature.

Shared Responsibilities: Team members learn to delegate tasks based on individual strengths, ensuring efficient use of time and resources. For example, in a marketing campaign project, one student might handle graphic design while another focuses on financial planning.

Building Trust: Working together on a shared goal fosters trust and mutual respect. This is particularly evident in long-term projects where success hinges on sustained cooperation.

Some students complained when team work began, but after these early dilemmas, most appreciated that the experience had taught them patience, empathy and leadership.

Encouraging Self-Directed Learning

These approaches can encourage students to take ownership of their learning. Such autonomy can instill a proactive attitude towards learning.

In-Depth Observations

Independent Research Skills: Often, the students have to collect information on their own, assess its authenticity, and then use it properly. It is a very valuable skill in this information age.

Time Management: Flexible project timelines teach the students to prioritize tasks, set goals, and meet deadlines. For instance, a student designing a scientific experiment learns how to allocate time efficiently between hypothesis formulation, testing, and reporting results.

More Accountability: The sense of ownership ensures that students hold themselves accountable for their successes and failures, thus creating a more mature learning environment.

Such self-directed approaches prepare the student for lifelong learning in a dynamic world.

Long-Term Knowledge Retention

PBL and Problem-Based Learning's interactive approach helps to have a better retention of knowledge as against the traditional passive learning approaches.

Why Retention Increases

Repetition and Application: Concepts are reiterated time and again through applying them to different scenarios which enhance learning. For instance, principles of mathematics related to the budgeting of a project are likely to be recalled than those done mechanically without application.

Emotional Connection: Projects always tend to be emotionally evocative: pride in finishing a tough task, excitement in discovery. Emotions are associated with memory.

Holistic Understanding: The student does not memorize but internalizes by relating concepts to real-world applications.

Students would often remember the minute details of projects years later, and thus, experiential learning is long-term.

Increased Self-Efficacy and Confidence

Opportunities to assume control over work, solve problems, and get things done have greatly boosted students' confidence.

Presentation Skills: Students get more adept at presenting ideas to others through frequent presentations.

Resilience: The setback of a failed prototype, for instance, encourages a growth mindset from them.

Pride in Accomplishment: Completion of a complex project such as the design of sustainable energy solutions leaves the student with pride and challenges them to take on even more difficult tasks.

These experiences translate into confident performances in other academic and professional settings.

Limitations

Despite the promising potential of the Named Entity Recognition (NER) model for automated resume parsing, several limitations were observed during the development and application of this project. A significant challenge is handling the vast diversity in resume formats. Resumes are often unstructured and vary in design, making it difficult for even advanced NER models to consistently extract information. Further, dependence on a good-quality training dataset was also a constraint. The performance of the model is highly dependent on the quality and diversity of data it uses for training. Inaccuracy can arise in the extraction process because of biases or lack of representation in the dataset, particularly in the parsing of resumes across various industries, regions, or linguistic backgrounds. Another limitation arises with the complexity of language. There are domain-specific jargon, abbreviations, and terms that are not typical which make it difficult to obtain relevant information from resumes. Also, the model may fail when information available in resumes is incomplete or ambiguous such as dates formatted differently, skills described vaguely, etc. Computational efficiency is another limitation because massive scale parsing using complex NLP models requires a lot of processing power and time, especially for organizations that maintain huge volumes of resumes.

Finally, the ethical and legal issues with data privacy and security are challenging. Automated parsing of resumes would involve dealing with sensitive personal information that needs robust methods of data protection against any unauthorized access and misuse. These limitations pose a critical need for mitigation in order to enhance the accuracy, scale, and trustworthiness of the NER model in resume parsing applications.

Future Research Work

Limitations found in this project provide scope for many future avenues of research and development. First, the ability to adapt to different formats in a resume can be improved. It could be done by implementing models based on transformer technology such as BERT or GPT which have shown remarkable contextual understanding. Techniques to standardize the unstructured resume to some extent before parsing will further increase accuracy.

Another area for future investigation would be to enrich the training dataset with a more diverse set of industries, roles, and languages. Resumes from different professional and cultural contexts can help reduce the chances of bias and ensure that the model performs reasonably well in most scenarios. Research into more specialized NER models geared towards specific industries, for example, healthcare or IT, can further enhance the appropriateness of the extracted information.

Another critical area is addressing the ambiguity of language and specific domain terminology. Future research could explore how context-aware embeddings or even hybrid models, combining rule-based and machine learning, could be better interpreted for nuanced language use. Furthermore, improving the computational efficiency through model optimization techniques such as pruning or quantization will allow the wide-scale deployment of resume parsers without losing speed.

Another relevant field for research is ethics and the law regarding automatic resume parsing. Develop open, explainable AI that is compliant with data protection rules such as the General Data Protection Regulation, increasing the trust levels of end-users. Improving all these areas of research shall make the resumed parsing better robust, versatile, and more ethical.

Conclusion:

This research shows the transformative potential of using named entity recognition models for automatically parsing resumes in the dynamic world of human resources.

In this, the NER model has offered a much more precise and efficient alternative by addressing inefficiencies and subjectivities related to the traditional manual resume processing system. The project showed that advanced natural language processing techniques were used in extracting such critical information as names, skills, education, and experience from resumes. It is quite clear that the model, by metrics such as precision, recall, and F1 score, demonstrates its efficiency in streamlining the hiring process. Its strengths notwithstanding, the project highlighted important limitations, including issues with varied resume formats, incomplete data, and ethical concerns related to data privacy.

The promising research directions include the integration of transformer-based models, enhancement of training datasets, and elimination of computational inefficiencies. Moreover, there is an extreme importance in following the legal and ethical standards to deal with sensitive personal information. In a nutshell, this project has laid the foundation for revolutionizing resume parsing in recruitment processes. By leveraging cutting-edge technologies and addressing existing gaps, automated resume parsing can significantly improve hiring efficiency, reduce biases, and provide a scalable solution for managing the growing volume of resumes in the modern workforce. This work is a stepping stone toward further developments that will further refine and expand the capabilities of automated resume processing systems.

Newality:

For instance, while working with job-seeking resumes, I was able to apply the

Named Entity Recognition (NER) tool analyze thousands lots of resume within few minutes and thereby extract important patterns and characteristics. The approach saves time and energy that would be used to gather boring details from the document that make resume analysis a boring process. This article mainly focuses on implementing of NER technology for the existing job problem and how AI is going to change the recruitment process by giving consistent perfect solutions for the huge amounts of data. This makes it have a good potential for publication contributing to the existing literature and future development of AI human resource tactics.

References:

Alessandro Fantechi, Gnesi, S., Livi, S., & Semini, L. (2021). *A spaCy-based tool for extracting variability from NL requirements*. https://doi.org/10.1145/3461002.3473074

Baigang, M., & Yi, F. (2022). A review: development of named entity recognition (NER) technology for aeronautical information intelligence. *Artificial Intelligence Review*, *56*(2), 1515–1542.

https://doi.org/10.1007/s10462-022-10197-2

BAUER, T. N., TRUXILLO, D. M., SANCHEZ, R. J., CRAIG, J. M., FERRARA, P., & CAMPION, M. A. (2001). APPLICANT REACTIONS TO SELECTION: DEVELOPMENT OF THE SELECTION PROCEDURAL JUSTICE SCALE (SPJS). *Personnel Psychology*, *54*(2), 387–419. https://doi.org/10.1111/j.1744-6570.2001.tb00097.x

Brindashree, B. V., & Pushphavathi, T. P. (2023). HR Analytics: Resume Parsing Using NER and Candidate Hiring Prediction Using Machine Learning Model. *International Journal of Research in Engineering, Science and Management*, *6*(12), 217–221. https://journal.ijresm.com/index.php/ijresm/article/view/2900

Cole, M. S., Rubin, R. S., Feild, H. S., & Giles, W. F. (2007). Recruiters' Perceptions and Use of Applicant Résumé Information: Screening the Recent Graduate. *Applied Psychology*, *56*(2), 319–343. https://doi.org/10.1111/j.1464-0597.2007.00288.x

Dai, X., Liu, Y., Wang, X., & Liu, B. (2014). *WINGS: Writing with Intelligent Guidance and Suggestions* (pp. 25–30). https://aclanthology.org/P14-5005.pdf

Derous, E., & Ryan, A. M. (2018). When your resume is (not) turning you down: Modelling ethnic bias in resume screening. *Human Resource Management Journal*, *29*(2), 113–130. https://doi.org/10.1111/1748-8583.12217

Hansen, S., & Atkins, E. (n.d.). *Automated System Monitoring and Notification With Swatch*. https://www.usenix.org/legacy/publications/library/proceedings/lisa93/full_papers/hansen.pdf

Harriet Rodney. (2019). The Artificial Intelligence Recruitment Process: How Technological Advancements Have Reshaped Job Application and Selection Practices. *Psychosociological Issues in Human Resource Management*, *7*(1), 42–47. https://www.ceeol.com/search/article-detail?id=770876

Ingold, P. V., & Langer, M. (2021). Resume = Resume? The effects of blockchain, social media, and classical resumes on resume fraud and applicant reactions to resumes. *Computers in Human Behavior*, *114*, 106573. https://doi.org/10.1016/j.chb.2020.106573

J. Himabindu Priyanka, & Parveen, N. (2023). DeepSkillNER: An automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-023-17264-y

Kang, Y., Cai, Z., Tan, C.-W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, *7*(2), 139–172. https://doi.org/10.1080/23270012.2020.1756939

Kinger, S., Kinger, D., Thakkar, S., & Devashish Bhake. (2024a). Towards smarter hiring: resume parsing and ranking with YOLOv5 and DistilBERT. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-024-18778-9

Kinger, S., Kinger, D., Thakkar, S., & Devashish Bhake. (2024b). Towards smarter hiring: resume parsing and ranking with YOLOv5 and DistilBERT. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-024-18778-9

Lavigna, R. J., & Hays, S. W. (2004). Recruitment and Selection of Public Workers: An International Compendium of Modern Trends and Practices. *Public*

Personnel Management, *33*(3), 237–253. https://doi.org/10.1177/009102600403300301

Maree, M., Kmail, A. B., & Belkhatir, M. (2018). Analysis and shortcomings of e-recruitment systems: Towards a semantics-based approach addressing knowledge incompleteness and limited domain coverage. *Journal of Information Science*, *45*(6), 713–735. https://doi.org/10.1177/0165551518811449

Mashayekhi, Y., Li, N., Kang, B., Jefrey Lijffijt, & Tijl De Bie. (2024). A challenge-based survey of e-recruitment recommendation systems. *ACM Computing Surveys*, *56*(10). https://doi.org/10.1145/3659942

Md. Sayeduzzaman, & Mahmud, A. (2022). *Design and Implementation of a Multi-Source Automatic Transfer Switch (ATS) System to Run the Utility Systems Via Different Power Sources and 3-Phase Synchronous Industrial Generator*.https://doi.org/10.3233/atde221199

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., & Euler, T. (2006). YALE: rapid prototyping for complex data mining tasks. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '06*. https://doi.org/10.1145/1150402.1150531

PETERSON, N. G., MUMFORD, M. D., BORMAN, W. C., JEANNERET, P. R., FLEISHMAN, E. A., LEVIN, K. Y., CAMPION, M. A., MAYFIELD, M. S., MORGESON, F. P., PEARLMAN, K., GOWING, M. K., LANCASTER, A. R., SILVER, M. B., & DYE, D. M. (2001). UNDERSTANDING WORK USING THE OCCUPATIONAL INFORMATION NETWORK (O*NET): IMPLICATIONS FOR PRACTICE AND RESEARCH. *Personnel Psychology*, *54*(2), 451–492. https://doi.org/10.1111/j.1744-6570.2001.tb00100.x

Pudasaini, S., Shakya, S., Lamichhane, S., Adhikari, S., Tamang, A., & Adhikari, S. (2021). Application of NLP for Information Extraction from Unstructured Documents. *Expert Clouds and Applications*, 695–704. https://doi.org/10.1007/978-981-16-2126-0_54

Racano, R. (2020). Companies' Approach to Digitalization in the Recruitment Process - A multiple case study on Swedish firms. *Gupea.ub.gu.se*. https://gupea.ub.gu.se/handle/2077/65780

Roy, A. (2021, January 25). *Recent Trends in Named Entity Recognition (NER)*. ArXiv.org.

https://doi.org/10.48550/arXiv.2101.11420

Roy, P. K., Chowdhary, S. S., & Bhatia, R. (2020). A Machine Learning approach for automation of Resume Recommendation system. *Procedia Computer Science*, *167*, 2318–2327. https://doi.org/10.1016/j.procs.2020.03.284

Shelar, H., Kaur, G., Heda, N., & Agrawal, P. (2020a). Named Entity Recognition proaches and Their Comparison for Custom NER Model. *Science & Technology Libraries*, *39*(3), 324–337. https://doi.org/10.1080/0194262x.2020.1759479

Shelar, H., Kaur, G., Heda, N., & Agrawal, P. (2020b). Named Entity Recognition Approaches and Their Comparison for Custom NER Model. *Science & Technology Libraries*, *39*(3), 324–337. https://doi.org/10.1080/0194262x.2020.1759479

Tarcar, A. K., Tiwari, A., Dhaimodker, V. N., Rebelo, P., Desai, R., & Rao, D. (2020, January 29). *Healthcare NER Models Using Language Model Pretraining*. ArXiv.org. https://doi.org/10.48550/arXiv.1910.11241

Wang, H., Li, J., Wu, H., Hovy, E., & Sun, Y. (2022). Pre-Trained Language Models and Their Applications. *Engineering*. https://doi.org/10.1016/j.eng.2022.04.024

Wierzbicki, M. P., Jantos, B. A., & Tomaszewski, M. (2024). A Review of Approaches to Standardizing Medical Descriptions for Clinical Entity Recognition: Implications for Artificial Intelligence Implementation. *Applied Sciences*, *14*(21), 9903–9903. https://doi.org/10.3390/app14219903

Won, M., Murrieta-Flores, P., & Martins, B. (2018). Ensemble Named Entity Recognition (NER): Evaluating NER Tools in the Identification of Place Names in Historical Corpora. *Frontiers in Digital Humanities*, *5*. https://doi.org/10.3389/fdigh.2018.00002

Wosiak, A. (2021). Automated extraction of information from Polish resume documents in the IT recruitment process. *Procedia Computer Science*, *192*, 2432–2439. https://doi.org/10.1016/j.procs.2021.09.012