AI-Powered Resume Screening System And Text Summarization

Kunal Sharma –(22bcs058)
Ketan Sharma –(22bcs053)
Karan Kant   –(22bcs051)
Priyansh   –(22bcs091)

ABSTRACT:

Our AI-Powered Resume Analysis and Text Summarization platform combines state-of-the-art deep learning and natural language processing (NLP) to automate recruiting and document processing. The resume analysis feature forecasts an applicant's professional field—e.g., Software Engineering, Data Science, Finance, or Marketing—based on their experience, skills, and credentials. This is done by mapping textual resume information into numerical forms with the help of word embeddings such as Word2Vec, GloVe, or BERT, which enable the AI model to grasp semantic word relationships. For instance, words such as "Python" and "Machine Learning" will be identified as extremely relevant to Data Science, whereas "Financial Analysis" and "Risk Management" will be linked to Finance. The deep learning classifier, which has been trained on a massive dataset of resumes with labels, subsequently predicts the most appropriate professional field for each candidate.

Automated screening of resumes by this approach removes human prejudice and increases the fairness of recruitment, so candidates are chosen according to their competence and not arbitrary considerations. Unlike conventional resume shortlisting, which is susceptible to unconscious gender, ethnic, or personal biases, this AI system solely bases its decision on qualifications and experience. Automating the resume categorization eliminates the time and effort that recruiters have to invest in manually sifting through thousands of resumes, freeing them up to focus on interviewing top candidates rather than manually sorting resumes. Also, the AI platform maintains consistency and high precision in candidate categorization, enhancing the overall effectiveness of the recruitment process.

The feature of text summarization supports the resume analysis feature by enabling users to feed any document, paragraph, or report and produce a brief, contextually correct summary. This feature is driven by cutting-edge NLP models, such as BERT, T5, and BART, which are accessed via the Hugging Face API. The system provides support for extractive summarization, where it chooses and scores the most salient sentences of the text, and abstractive summarization, where it creates a fully new summary in a more human-like and natural way. Such functionality is very helpful to experts in HR, business, and research, as they tend to deal with great volumes of text in an efficient manner. For example, a hiring manager can employ the summarization feature to extract important information from a long job posting in a matter of seconds, while a researcher can summarize a lengthy report to understand its key findings within seconds. Its capability to create summaries of variable length makes the tool versatile and easily accommodative to varying user requirements. By streamlining resume filtering and document summarization, this AI-based system minimizes human workload substantially, improves decision-making, and brings equity and efficiency to the recruitment process. In addition, utilizing machine learning models that are trained on various datasets, the system reduces bias to ensure equal opportunities to all applicants without compromising text processing quality in terms of precision and reliability.

INTRODUCTION:

In the fast-paced employment market of today, hiring managers are confronted with the daunting task of sifting through job applications to identify top talent. Sifting through hundreds of resumes manually not only takes time but is also subject to inconsistency, and unconscious bias. As a result, top talent may be lost, while subpar candidates may get through.

Therefore, we created an Artificial Intelligence powered resume screening system that screens candidates according to their skill. Using sophisticated Natural Language Processing and machine learning, the system accurately sorts resumes into the appropriate fields such as Softwares Engineering, Data Science, Marketing, and Finance. This kind of automation allows the recruiter to devote attention to interacting with top candidates without cusing bias or accelerting hiring. This improves not just effectiveness but also equitable and inclusive recruitment practices.

But smart automation is hiring only necessary. Specialists in fields ranging from law and money to medicine and the academy wrestle with sorted-through amounts of written information. Reading reports, briefs, or research studies by hand is a slowdown on energy and decision-making pace.

To address this, we also developed an AI-powered text summarization tool that summarizes long papers into crisp, precise abstracts automatically. Users can specify the summary length, in case they want a concise summary or a more detailed digest. Powered by state-of-the-art NLP models such as BERT, T5, and BART, the platform is capable of carrying out both extractive and abstractive summarization—either extracting key sentences or rewording in a natural, human-like fashion.

Together, these tools create a powerful, end to end solution that combines resume screening and document summarization. Organization can streamline their hiring processes and professionals can make faster, more informed decisions—all while saving time, reducing manual efort, and

ensuring fairnes. By bridging human insight with AI automation, we're helping businesses and individuals work smarter, not harder.

LITERATURE REVIEW:

Prior to creating our AI-based system, we thoroughly researched the work of others in terms of resume screening and summarizing text. Knowing the classic methods as well as current AI solutions enabled us to create a superior, more efficient tool.

- **Traditional Methods:**
  The traditional recruitment process is still heavily reliant on manual resume screening carried out by human resources. This traditional process has three inherent drawbacks that affect hiring efficiency and quality. For one, the process is very time-inefficient, because processing every application manually is a time-consuming endeavor requiring extensive personnel time, especially for positions getting high numbers of applicants. Second, human assessors necessarily bring subjectivity by way of innate cognitive biases, which in turn may unwittingly affect candidate evaluation. Third, assessment criteria tend to differ across reviewers, resulting in inconsistent application of hiring criteria among candidates.
  To counter these challenges, Applicant Tracking Systems (ATS) emerged as an automated option, using mostly keyword comparison algorithms. Although these systems have enhanced processing rates and lowered initial screening time, they have a number of serious drawbacks. The most significant problem is their excessive dependence on precise keyword matching, which often leads to the rejection of otherwise suitable candidates whose CVs employ different wordings to express the same qualifications. Furthermore, these systems do not have the capacity to understand semantic relationships within terms and therefore cannot perceive that

variously phrased content can embody the same competencies. In addition, they showcase rigidity in reading diverse resume structures and wording conventions, frequently dismissing highly qualified candidates based on the differences of formatting instead of qualitative lacks in qualifications

- **AI-based and Machine Learning Methods:**
  The provided technology by natural language processing (NLP) and machine learning has transformed the screening process of resume, and the paradigm is now from basic keyword matching to intelligent candidate assessment. Current AI systems now scan resumes by truly comprehending the contents, assessing skills, experience, and qualifications in light of job specifications. Instead of merely scanning for predetermined buzzwords, these systems are able to understand that "Python development" and "built applications using Python" are the same skill. Top companies are using advanced pre-trained language models such as BERT and GPT, which have shown impressive ability to understand complex information in resumes, such as interpreting diverse phrasing, understanding contextual relationships among skills, and even inferring unmentioned abilities. Following this tried-and-tested method, our system also uses these sophisticated AI techniques to provide more accurate, equitable, and holistic resume analysis - going beyond the limitations of conventional screening processes to identify more genuinely qualified candidates.

- **Text Summarization Techniques**
  These days, hiring managers may simply not have the time to read through long resumes. Text summarization comes in handy here by

extracting the most important and relevant details, and enabling faster and more effective decision-making. Apart from recruitment, this method is also extensively utilized by book readers, students reading study guides, researchers reading documents, and business professionals in need of processing large amounts of text. Through the compression of massive amounts of data into succinct, easy-to-consume abridgments, text summarization makes productivity higher, time savings, and increased understanding possible in multiple fields.

.

OBJECTIVE:

Our solution tackles two key issues in contemporary recruitment: effective candidate filtering and impartial resume assessment. Through the use of AI-driven automation, we turn the traditionally time consuming recruitment process into a smart, data-driven process. Here's how our solution adds value:

- The system itself examines and classifies resumes automatically according to candidates' true competencies and not merely keywords. With the power of advanced NLP methods, it picks up and understands skills, experience, and qualifications in context - understanding that "led a dev team" and "managed software developers" are equivalent leadership experience.
- At the heart of it, advanced machine learning algorithms transform resumes and job descriptions into similar vector representations. This facilitates semantic similarity-based correct matching rather than shallow keyword similarity overlap, allowing candidates to be judged on their actual qualifications
- By standardizing the evaluation criteria and stripping away identifiable demographic data at processing, the system reduces unconscious human biases that typically characterise conventional hiring. It

concentrates solely on quantifiable abilities and experiences critical to the job.

- The embedded summarization feature condenses long resumes into brief professional summaries, underscoring key qualifications and career development. This provides the hiring managers with the critical information they require without having to sift through full documents.
- We also created easy-to-implement through simple, intuitive web interface in which recruiters can upload resumes in multiple formats and obtain result for which position this candidate is best suited. And also view condensed professional profiles.

METHODOLOGY:

In order to attain these goals we undertook the following steps that involves:

- **Data Gathering**: We began by collecting a dataset of resumes from Kaggle.
- **Data Preparation**: We pulled out the resumes to organized forms through data cleaning and the use of NLP techniques such as vectorization, stopword removal, and Named Entity Recognition to extract the most crucial information.
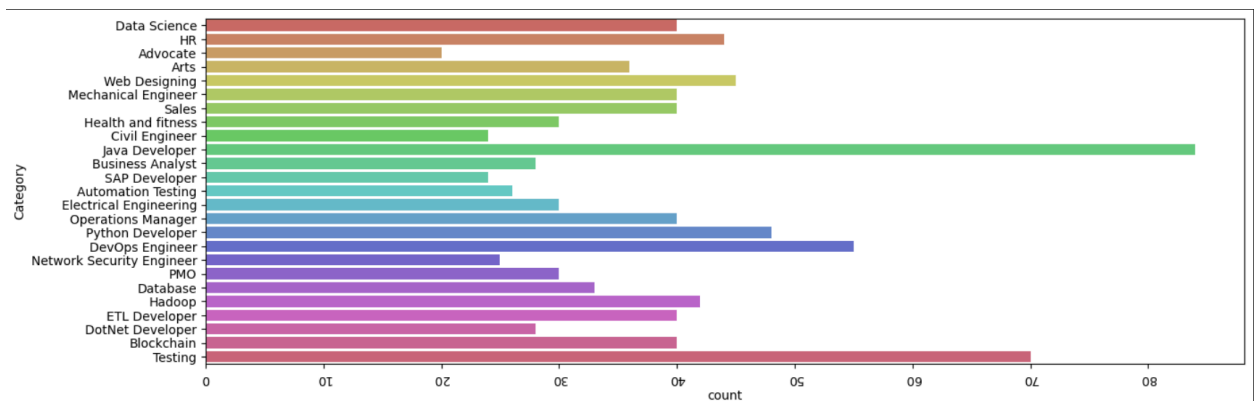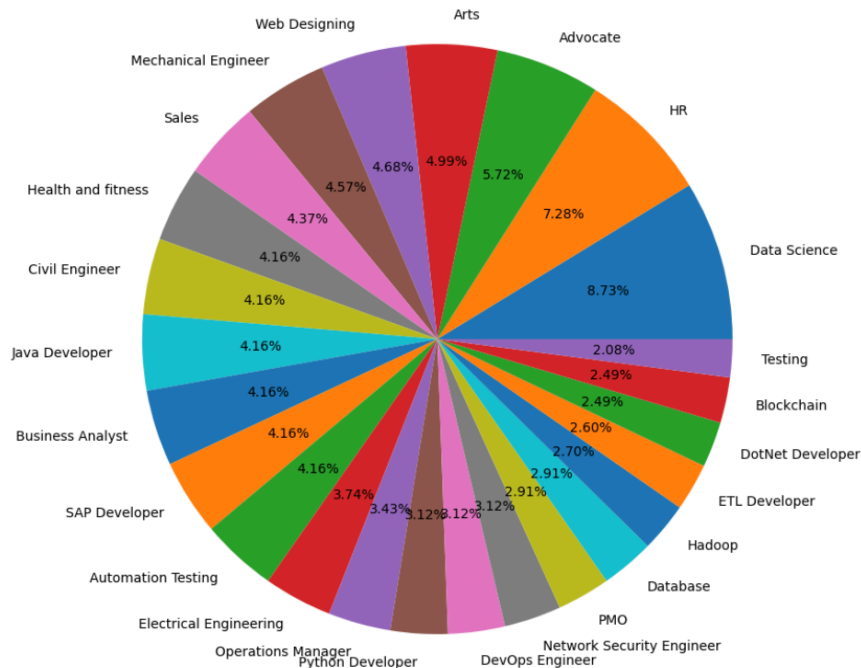
- **Resume Screening Model**: We are using several classification algorithms to train and test our model, which include K-Nearest Neighbors (KNN), Naive Bayes, Support Vector Machine (SVM), and Random Forest. Each of them provides us with different accuracy values: KNN gets 98%, Naive Bayes 96%, SVM 99%, and Random Forest also 99%.

- **Implementation & User Interface**: We created a friendly web application based on HTML, CSS, JavaScript, and Flask that automates the hiring process. The site enables recruiters to conveniently upload resumes of candidates and enter paragraphs of descriptive text. The system later reads through and summarizes the input content, pulling out important information to enhance effective decision-making.

- **Text Summarization**:To integrate automatic text summarization into our system, we employed Hugging Face's Transformers API, which provides access to industry-leading pre-trained models that are specifically fine-tuned for summarization tasks. The integration begins

with selecting an appropriate model architecture such as BART or T5 which are designed to carry out abstractive summarization by extracting contextual relationships in the text and generating coherent, shortened versions that preserve the most critical information. They have been pre-trained on massive corpora (e.g., science articles, news stories) and will generalize extremely well to other domains, so they are suited to processing resumes, reports, or other business documents.

CONCLUSION:

This AI-powered recruitment solution is a revolutionary leap in recruitment processes by addressing two of the most persistent issues of talent identification in the current era: the inefficiency of manual resume screening and the sheer volume of information recruiters must sort through. Traditional recruitment has been a process of resume reading for hours, gut feeling, keyword spotting, and format-driven judgments—all prone to human weakness, inconsistency, and prejudice. Our platform disrupts this old tradition by utilizing cutting-edge natural language processing (NLP) algorithms and transformer-based artificial intelligence models, such as those provided by Hugging Face, to turn hiring into a far more objective, accurate, and efficient process.

The resume screening module enhances keyword matching by validating the context, pertinence, and compatibility of a candidate's skills and experiences so that it can effectively categorize applicants into professional groups like Software Engineering, Data Science, or Marketing. Contextual classification ensures that the applicants are being mapped to positions not just because they have identified a trending skill, but because their credentials also fit the position's requirements. It highly reduces the burden

of recruiters, releasing hours on end from being wasted on monotonous, manual screening. While the text summarization functionality condenses long resumes or documents into concise, consistent summaries containing only the key points—education, work history, relevant tools, and achievements—so recruiters can glance and compare candidates at the speed of light.

Together, these two robust abilities simplify the hiring process, cutting hiring cycles by as much as 70–80% while enhancing fairness and minimizing bias.
By eliminating the most likely noise to corrupt human judgment and focusing only on information pertaining to the issue at hand, the system allows for equitable hiring choices—giving each of the qualified candidates an equal chance to get noticed. In fact, this does not mean doing away with human interaction but allowing hiring managers to make value-based, decisive decisions instead of administrative hurdles. Recruiters now have time to engage meaningfully with candidates, interview more efficiently, and collaborate on hiring strategy, and organizations have data-driven, consistent, and scalable evaluation. Basically, this AI-hiring solution doesn't just make hiring easier—it improves the whole hiring process. It turns what used to be a time-consuming, inconsistent process into a fast, smart system that matches talent with possibility more effectively than ever before. By combining machine efficiency with human smarts, the platform provides a future-facing talent acquisition solution that is faster, fairer, and brighter by design.

FUTURE SCOPE:
- In the future, we can improve these models and large transformer models like GPT-4 or LLaMA for better contextual understanding.
- We might design an explainable AI system to provide insights as to why a candidate was selected or not.

- We might also add support to these NLP models for covering many languages so as to conduct global hiring from various geographical areas.
- We can also optimize abstractive summarization to produce human-like summaries with higher clarity.

REFERENCES:

Research Paper Link:

- https://arxiv.org/abs/2307.08624?utm_source=chatgpt.com
- https://doi.org/10.1108/SHR-04-2019-0024
- https://dl.acm.org/doi/10.1145/3457607

Dataset Link:

- https://www.kaggle.com/datasets/gauravduttakiit/resume-dataset?resource=download

Documentation Reference of Algorithm:

- https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
- https://scikit-learn.org/stable/api/sklearn.naive_bayes.html#module-sklearn.naive_bayes
- https://scikit-learn.org/stable/api/sklearn.svm.html#module-sklearn.svm