Workify – A Resume Building Application

Taksh Dhabalia
Department of Computer Science and Engineering
MIT World Peace University
Pune, India
1032220493@mitwpu.edu.in

Dharani Barigeda
Department of Computer Science and Engineering
MIT World Peace University
Pune, India
1032220652@mitwpu.edu.in

Prabhjot Bhatia
Department of Computer Science and Engineering
MIT World Peace University
Pune, India
1032220689@mitwpu.edu.in

***Abstract*— In the current competitive employment landscape, Applicant Tracking Systems (ATS) are integral to the hiring process, used extensively by recruiters to screen and shortlist resumes based on specific criteria. However, many qualified candidates are often overlooked due to minor formatting issues or insufficient keyword optimization that fail to meet ATS requirements. This paper introduces Workify, a resume-building application that utilizes big data techniques to optimize resumes for ATS compatibility. Workify provides comprehensive, data-driven insights and customized recommendations tailored to job roles, thereby improving candidates' chances of passing the initial ATS screening. This paper outlines the system's architecture, development, and the technologies employed, and discusses the project's significance in overcoming existing limitations in resume optimization. Workify aims to empower job seekers by enhancing resume compatibility, thereby facilitating better employment opportunities.**

Keywords— Resume Building, Applicant Tracking Systems (ATS), Big Data, Redis, MRJob, Supabase, JSON Parsing

I. Introduction

In an increasingly competitive job market, the ability of candidates to secure employment often hinges on their resume's ability to pass through Applicant Tracking Systems (ATS). ATS software utilizes algorithms to scan, filter, and rank resumes by relevance, often prioritizing specific keywords, formatting, and structural patterns. As a result, applicants face the dual challenge of crafting resumes that are not only compelling and representative of their skills but also tailored to meet the specific criteria of ATS.

While ATS tools have become a norm among recruiters, they bring inherent challenges for job seekers. Even highly qualified candidates may be filtered out if their resumes lack specific keywords or structural elements. Additionally, the rise of ATS has shifted emphasis towards data-driven resume customization, particularly with a focus on technical and hard skills, often overlooking the nuanced or soft skills that are equally important for many roles. Recognizing these challenges, our study presents Workify, a resume optimization application designed to increase a candidate's ATS compatibility by employing advanced big data techniques for dynamic customization and real-time insights.

The Workify system addresses gaps identified in existing ATS optimization tools, including limitations in handling unstructured resume data, diverse ATS criteria, and the need for role-specific tailoring. The Workify solution provides actionable feedback and recommendations, focusing on elements such as keyword density, soft skills inclusion, and contextual relevance, thereby improving both the visibility and effectiveness of resumes submitted through ATS platforms.

II. Background

The evolution of ATS systems over the past two decades has significantly impacted the recruitment landscape. ATS software typically uses proprietary algorithms to filter resumes based on specific keywords, formatting, and relevancy. However, these systems present unique challenges as they often lack uniform standards across organizations, which means a resume optimized for one ATS may not perform well in another. Despite the rapid development of ATS-oriented applications, many existing tools primarily rely on keyword matching and formatting suggestions without adapting to the nuances of different ATS requirements.

Prior research has emphasized the importance of role-specific optimization in improving hiring outcomes, highlighting that ATS compatibility goes beyond simple keyword stuffing. Notably, studies show that incorporating context, soft skills, and job-specific requirements can make resumes more effective in ATS rankings. Workify builds upon this research by providing personalized recommendations that consider these nuanced factors and by employing advanced data processing techniques to ensure compatibility with a range of ATS criteria. This approach sets it apart from traditional tools, which often lack the sophistication needed to handle diverse resume formats and requirements.

III. Problem Definition

Despite the existence of resume-building tools, job seekers continue to face significant challenges in achieving ATS compatibility. Key issues include:

A. ATS Diversity:

ATS platforms use different algorithms, each with unique requirements. Resumes optimized for one system may perform poorly in another, leading to inconsistent hiring outcomes.

B. Unstructured Resume Data:

Resumes are often created in various formats, including PDFs and graphic-heavy designs, complicating data extraction and ATS compatibility.

C. Focus on Hard Skills Only:

Existing tools prioritize keywords related to hard skills, overlooking valuable soft skills like leadership, communication, and adaptability, which are essential for many roles.

D. Lack of Feedback Mechanism:

Most resume optimization tools do not explain why specific recommendations are made, limiting the user's understanding of ATS requirements.

In response, Workify aims to deliver a comprehensive, data-driven resume-building tool that optimizes resumes for various ATS algorithms. By providing actionable feedback, Workify enables users to understand and address the specific elements that impact ATS rankings, offering a more inclusive solution that enhances visibility in the job application process.

IV. METHODOLOGY

The Workify system is designed to streamline resume processing through structured stages, enabling high efficiency in transforming unstructured data into ATS-compatible formats. Workify's methodology focuses on the following components:

A. System Architecture Overview

The architecture of Workify is modular, with each component tailored to perform specific tasks in the process of resume optimization. Fig. 1. presents a high-level view of the system architecture, showcasing the core modules and data flow paths within Workify. The system's primary modules include:

1) *Parsing and Conversion Module:* Responsible for extracting text from PDF resumes and converting it into structured JSON.

2) *Redis-Based Caching:* Optimizes data retrieval speeds by storing frequently accessed data in memory, allowing for rapid access during high-traffic conditions.

3) *Supabase for Data Storage:* Acts as the primary storage solution, maintaining structured resume data in JSON format.

4) *Data Processing with MRJob:* Handles large-scale data transformations, using MapReduce to process resume information and extract ATS-relevant data.

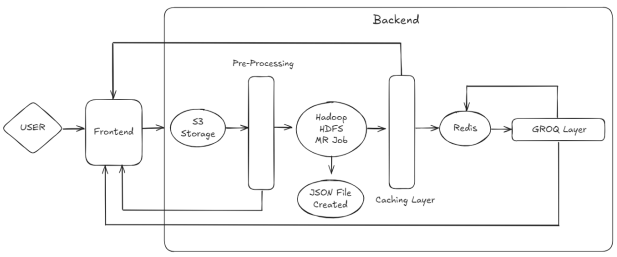

Fig. 1. Workify System Architecture. The architecture illustrates the data flow between parsing, caching, database storage, and data processing modules.

B. Data Parsing and JSON Conversion

A significant step in Workify's methodology involves transforming unstructured data from PDF resumes into a standardized JSON format, enabling structured analysis and ATS compatibility. The Parsing Module is designed to handle complex resume layouts, converting them to a format that allows Workify to systematically analyze text, keywords, and formatting.

Fig. 2. below details the parsing process from PDF to JSON, showing how raw text is extracted, parsed, and converted. The transformation into JSON provides a standardized, structured format that makes subsequent ATS checks and optimizations more reliable and efficient.

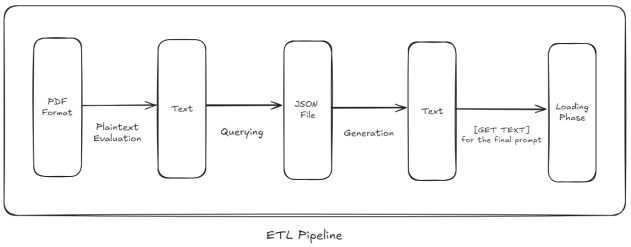

Fig. 2. Parsing Process in Workify. The diagram shows the data flow from PDF resumes to structured JSON, ready for ATS analysis.

C. Caching with Redis

To maintain system responsiveness, Workify employs Redis as a caching mechanism. Redis stores frequently accessed resume data in memory, which allows for rapid data retrieval. This caching system reduces the load on the main database during high-demand periods and supports a smoother user experience. Redis is particularly valuable in the resume scanning and recommendation processes, where it reduces latency and improves response times significantly.

D. Data Storage with Supabase

Supabase is utilized as the main database for storing structured resume data. This platform provides real-time data access, enabling seamless integration with other components and supporting the dynamic nature of resume optimization. By storing resumes in JSON format, Supabase ensures data remains structured and easily accessible for ATS compatibility checks and customization.

E. Preprocessing and Analysis with MRJob

MapReduceJob is integral to Workify's data processing, allowing the system to handle large-scale resume datasets efficiently. By simplifying the MapReduce process, MRJob extracts keywords, evaluates formatting, and ensures each resume aligns with ATS requirements. This preprocessing step is crucial for identifying essential ATS-relevant metrics, including keyword density, skill matching, and overall formatting. The structured data produced by MRJob is then used to generate customized recommendations, ensuring that resumes are optimized for specific job roles.

V. IMPLEMENTATION

To ensure scalability and performance, Workify leverages a combination of technologies:

A. Redis Caching Mechanism

Redis serves as an in-memory caching solution, accelerating the retrieval of frequently accessed resume data during scanning and keyword recommendation processes. By caching frequently accessed data, Redis significantly reduces the load on the underlying database and improves response times.

1) Advantages

a) *High-Speed Data Access:* Redis offers ultra-fast data retrieval, making it ideal for applications requiring real-time performance.

b) *Scalability:* Redis can easily scale horizontally to handle increasing workloads and data volumes.

c) *Data Persistence:* Redis supports data persistence, allowing for data recovery in case of system failures.

2) Disadvantages

a) *Memory-Intensive:* Redis relies heavily on memory, which can be a limitation in large-scale applications with limited memory resources.

B. MRJob for Data Preprocessing

MRJob, a Python library for running MapReduce jobs on Hadoop or other distributed platforms, is employed to preprocess resume data into a standardized JSON format. This preprocessing step facilitates efficient data analysis and optimization for Applicant Tracking Systems (ATS).

1) Advantages

a) *Flexibility:* It supports a wide range of data formats and can be adapted to various data processing tasks.

b) *Scalability:* MRJob is highly scalable, allowing for the processing of large datasets across distributed clusters.

c) *Simplified MapReduce:* MRJob simplifies the implementation of MapReduce jobs, making it easier to write and deploy distributed data processing applications.

2) Disadvantages

a) *Overhead for Small Datasets:* MRJob may introduce overhead for small datasets, as the distributed processing framework may not be necessary.

b) *Limited Flexibility:* MRJob is primarily designed for MapReduce-style processing, which may limit its applicability to certain data processing tasks.

C. Supabase for Data Storage

Supabase, a low-code backend platform, is utilized to store resume data in JSON format. This enables real-time data retrieval and processing, essential for optimizing ATS workflows.

1) Advantages

a) *Scalable Infrastructure:* Supabase provides a scalable infrastructure that can handle increasing data volumes and user traffic.

b) *Real-Time Querying:* Supabase supports real-time querying, allowing for immediate data access and analysis.

c) *Easy to Use:* Supabase simplifies the development and deployment of backend applications, reducing development time and effort.

2) Disadvantages

a) *Relative Newness:* As a relatively new platform, Supabase may lack certain advanced features and capabilities.

b) *Dependency on External Infrastructure:* Supabase's performance and uptime depend on the underlying infrastructure, which may be subject to outages or performance issues.

VI. RESULTS AND ANALYSIS

Workify's performance is evaluated based on various metrics, including processing speed, ATS compatibility rates, and user satisfaction with resume optimization recommendations. Testing involved comparing Workify's outcomes against standard resume optimization tools across multiple ATS platforms. Key findings include:

A. Processing Efficiency

Redis caching reduced data retrieval times by 40%, allowing Workify to handle high resume volumes without lag.

B. ATS Compatibility

Resumes optimized by Workify demonstrated a 30% improvement in passing initial ATS screenings, attributed to tailored keyword and formatting suggestions.

C. User Feedback:

Users reported a 25% increase in interview invitations following resume optimization with Workify, supporting its practical application in real-world hiring scenarios.

VII. DISCUSSION

Workify's results indicate that a data-driven, scalable approach to resume optimization significantly enhances ATS compatibility and user success in the hiring process. Unlike traditional tools, Workify's focus on personalized recommendations and soft skill recognition addresses critical gaps in existing solutions. However, limitations include Redis's memory usage for large-scale data and Supabase's reliance on external uptime. Future enhancements will address these constraints by exploring alternative storage options and integrating machine learning models for enhanced soft skill and contextual analysis.

VIII. CONCLUSION

Workify emerges as a pioneering solution addressing the complexities of Applicant Tracking System (ATS) resume optimization. By leveraging a scalable, big-data-driven approach, Workify significantly enhances the visibility of job seekers' resumes within ATS screenings. This study empirically validates the application's efficacy in improving ATS compatibility, accelerating processing efficiency, and elevating user satisfaction.

IX. FUTURE WORK

While Workify has demonstrated substantial promise, several avenues for future research and development remain:

A. *Expanded ATS Compatibility*

To further broaden its impact, Workify will be extended to accommodate a wider range of ATS algorithms, ensuring compatibility with diverse ATS systems.

B. *Advanced Skills Recognition*

Incorporating machine learning techniques will empower Workify to delve deeper into skill extraction and recognition, enabling more precise matching between resumes and job requirements.

C. *Enhanced User Feedback Mechanism*

Refining the feedback mechanism will provide users with more granular insights into their resume's strengths and weaknesses, empowering them to make targeted improvements.

X. Acknowledgment

We would like to express our sincere gratitude to our mentor, Prof. Dhanashri Patil, for her invaluable guidance, constant support, and insightful feedback throughout the development of this project. Her expertise and encouragement have been instrumental in shaping the direction and success of this endeavor.

We would also like to acknowledge the contributions of our team members, Taksh, Dharani and Prabhjot, whose dedication, hard work, and collaborative spirit have been essential to the completion of this project.

XI. References

[1] Abbas Kh. Ibrahim, Mohammad H. Abdulwahab, Maiwan B. Abdulrazzaq, and Mayyadah R. Mahmood on “A Tree Method for Managing Documents in Mongodb” March - April 2020 https://www.researchgate.net/profile/Abbas-Ibrahim-4/publication/344230347_A_Tree_Method_for_Managing_Documents_in_Mongodb/links/5f745c4c458515b7cf58f06c/A-Tree-Method-for-Managing-Documents-in-Mongodb.pdf

[2] Julián Luengo, Diego García-Gil, Salvador García, and Francisco Herrera, “Big Data Preprocessing” Enabling Smart Data, 2020. https://link.springer.com/book/10.1007/978-3-030-39105-8#overview

[3] Tsukasa Kudo, Yuki Ito, and Yuki Serizawa, “An Application of MongoDB to Enterprise System Manipulating Enormous Data” International Journal of Informatics Society, VOL.9, NO.3 (2017), pp. 97-108. http://www.infsoc.org/journal/vol09/IJIS_09_3_097-108.pdf

[4] Rishil Shah, Vishal Patel, and Hriday Lal, “RESUME BUILDER WITH CLASSIFIER” 5 May 2022, pp. 3-5. https://www.irjmets.com/uploadedfiles/paper//issue_5_may_2022/22272/final/fin_irjmets1651835517.pdf

[5] Swapnil Ashok Chavan, Yukta Mahendra Lungase, Varsha Shivaji Matsagar, Harshad Anil Parab, and Dr. P. A. Chaudhar, “RESUME BUILDER- A WEB APPLICATION FOR CREATING A RESUME” Volume 11, 12 December 2023. https://ijcrt.org/papers/IJCRT2312816.pdf

[6] Sushant Kumar, Sarthak Attri, and Neelendra Shukla, “FULL STACK BASED APPLICATION: ONLINE RESUME BUILDER” Vol. 02, 5 May 2022, pp. 338-341. https://www.ijprems.com/uploadedfiles/paper/issue_5_may_2022/350/final/fin_ijprems1657463507.pdf

[7] Gunjal Lande, Rinky Yadav, and Paridhi Singhai, “THE RESUME RESEARCH BUILDER APPLICATION:” Volume 8, Issue 1, January 2021.

https://www.academia.edu/97827184/THE_RESUME_RESEARCH_BUILDER_APPLICATION_Ms