Dynamic Resume Evaluation a Comprehensive Approach to Part-Based Weightage Assignment and Score Generation

Sivaramakrishnan S
Department of Information Science and Engineering
New Horizon College of Engineering
Bengaluru, India
sivaramkrish.s@gmail.com

A Naveen
Department of Information Science and Engineering
New Horizon College of Engineering
Bengaluru, India
t18anaveen@gmail.com

Chethan A S
Department of Information Science and Engineering,
New Horizon College of Engineering
Bengaluru, India
chethansri1215@gmail.com

Hemsagar N M
Department of Information Science and Engineering,
New Horizon College of Engineering
Bengaluru, India
hemsagargowda33@gmail.com

Hrithik U
Department of Information Science and Engineering
New Horizon College of Engineering
Bengaluru, India
hrithik356hrithik@gmail.com

2024 International Conference on Recent Innovation in Smart and Sustainable Technology (ICRISST) | 979-8-3503-0914-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRISST59181.2024.10921820

Abstract: **Navigating the dynamic landscape of today's employment market to identify suitable opportunities can be challenging. The advent of AI has empowered candidates to be ranked for a job role processed by computers. While conventional methods can quantify certain elements of positions and applicants, the conversion of unstructured data from job descriptions and resumes often leads to the loss of crucial information. Recently, Large Language Models (LLMs) have demonstrated exceptional performance in text-based data domains within the AI field. Inspired by the prowess of LLMs in natural language understanding, this study exploits their capabilities to capture information previously lost during the conversion of unstructured data. In this innovative approach, traditional methods are surpassed by integrating web scraping techniques to extract valuable insights directly from resumes. By systematically gathering unstructured data from resumes, the LLMs are enriched with a diverse range of information, ensuring a more comprehensive understanding of candidates' skills and experiences. The acquired information is then utilized through a self-developed ranking mechanism tailored for the recruitment process, effectively ranking candidates based on their skills and capabilities.**

Keywords: **LLM, NLP, K-Means clustering, Rank scoring, Applicant tracking system, NIRF ranking.**

I. INTRODUCTION

While traditional approaches excel in quantifying specific facets of positions and candidates, the process of converting unstructured data from resumes and job descriptions into structured formats often leads to the omission of critical information. The advent of Large Language Models (LLMs) has brought about a paradigm shift in the field of AI, showcasing exceptional performance in domains reliant on text-based data. Motivated by the remarkable capabilities of LLMs, the study introduces an innovative Applicant Tracking System designed to leverage the natural language understanding skills inherent in these models.

This research paper provides a comprehensive exploration of the architecture, functionality, and real-world implications of "Talent-Insight." Through empirical evidence and case studies, the system's efficacy in streamlining the recruitment process, empowering HR professionals in decision-making, and ultimately enhancing the overall efficiency of talent acquisition strategies is illustrated. Embark on a journey through the transformative landscape of recruitment technology as it will unveil the potential of "Talent-Insight" in shaping the future of talent acquisition.

II. LITERATURE SURVEY

"JobRecoGPT a job recommendation system using Large Language Models (LLMs)".[1]. Comparing four approaches, it finds the LLM unguided approach outperforms others in both synthetic and real-world data experiments. The Hybrid method, combining LLM and deterministic approaches, offers practical benefits, emphasizing the importance of capturing unstructured elements for effective job recommendations.

"Resume Parser Using Natural Language Processing Techniques"[2] to streamline hiring processes. The model extracts and ranks relevant information from resumes based on company requirements, utilizing OCR and NLP. The proposed system includes a web portal for HR input, and future enhancements aim to parse resumes from diverse platforms and expand system capabilities.

"Applicant Tracking and Scoring System (ATSS) leveraging NLP and OCR."[3] to streamline resume screening. The web-based system, employing features like Named Entity Recognition and data visualization, aims to simplify the recruitment process for HR officers. The ATSS intends to enhance efficiency, transparency, and fairness in hiring practices, benefiting both employers and candidates.

"Predictive HR Candidate Ranking System using machine learning and AI."[4] to automate job applicant ranking based on qualifications. It addresses challenges in efficiently selecting qualified candidates and emphasizes systematic criteria and ranking policies. The system utilizes neural networks, fuzzy sets for transparency, and aims to revolutionize recruitment by streamlining processes and facilitating effective communication with top candidates.

"Attention Is All You Need"[5]. The Transformer is a new neural network architecture based on attention mechanisms, suitable for tasks like machine translation, text summarization, and question answering. It is a sequence-to-sequence model that uses attention mechanisms instead of recurrent neural networks (RNNs), which can be slow to train and difficult to parallelize. The Transformer has shown effectiveness in various tasks, such as the WMT 2014 English-to-German translation task. This significant advance in neural machine translation is expected to significantly impact the development of new and improved machine translation systems.

"Diversity based self-adaptive clusters using PSO clustering for crime data"[6]. The authors introduce a diversity-based self-adaptive PSO (DPSO) approach for clustering crime data, addressing the issue of inaccurate or misleading clusters. DPSO dynamically adjusts parameters based on data characteristics, aiming for better cluster quality and diversity [7]. Comparing DPSO with other algorithms, it improves clustering accuracy and diversity.

"Introduction to AI Technique and Analysis of Time Series Data Using Facebook Prophet Model."[8] [9]. The paper discusses the Facebook Prophet model, an AI technique designed to analyse and forecast time series data. It explores its limitations and real-world application, highlighting its potential for unlocking secrets within the data. The paper also highlights future research avenues.

"A Literature Review on Sentiment Analysis"[10]. They untangle the concept, exploring various approaches like lexicon-based and machine learning methods, and unveil its diverse applications across social media, product reviews, and beyond. The paper acknowledges the challenges inherent in sentiment analysis, but also paints a promising picture of future directions, leaving readers with a comprehensive understanding of this powerful tool for extracting opinions and insights from textual data.

III. SYSTEM DESIGN

Once the data is gathered, a large language model (LLM) processes the collected information, which likely involves natural language understanding to extract relevant details about the candidate's skills, experience, and qualifications [11][12][13][14][15].

The extracted data from both the resume and the web scraping are then passed to a ranking algorithm. This algorithm evaluates the information based on certain criteria to determine the relevance and quality of the candidate's profile. Fig 1 shows the flow chart of system design.

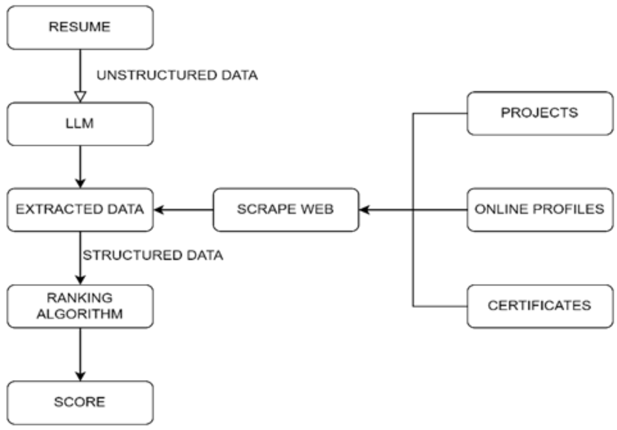

Fig. 1: System Design

Finally, the output of the ranking algorithm is a score that reflects the candidate's suitability for a position or a role. This score can be used to rank candidates, making the selection process more efficient for recruiters or hiring managers.

In summary, the system automates the assessment of candidate profiles by leveraging both traditional resume information and supplementary web data, analysed through advanced language models and ranked with a bespoke algorithm to produce a quantifiable score.

IV. METHODOLGY

Each resume can be split into two major parts:
Part A: Candidate's educational information.
Part B: Candidate's skills, projects, etc.

Each part is given different importance based on various factors. For instance, Part A might be given more consideration when the candidate is from a prestigious university, while Part B could be prioritized for candidates from lesser-known universities but with valuable skills.

In the given scenario, a clear pattern emerges, allowing each part of the resume to be assigned different weights according to the situation.

How can the weights be assigned?

Dynamic Resume Evaluation (DRE) Algorithm employs the NIRF ranking framework and the median salary of each university for assigning weights to different parts of the resume.

Explaining with example:

Taking the top 200 universities for engineering in the NIRF framework list for the past 3 years with their median salary and applying a k-means algorithm (the algorithm can use the elbow point method to determine the optimal group size).

Using the information from Fig. 2, DRE Algorithm finds the average salaries from each group as follows:

Average Salary for Cluster 1: 4,40,000
Average Salary for Cluster 2: 8,14,000
Average Salary for Cluster 3: 15,40,000
Average Salary for Cluster 4: 25,00,000

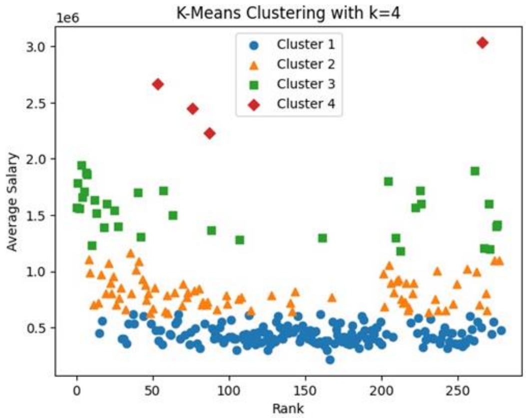

Fig. 2: K Means Clustering

Using the average salaries from each cluster, DRE Algorithm calculates the percentage variation in the median salary to assign weights to the parts of the resume. Calculating the starting weights involves finding the percentage increase between the highest median salary cluster and the lowest median salary cluster.

((25,00,000 – 4,40,000) / 25,00,000)

=0.82(82%) ≈ 80%

These are the starting weights for the highest median salary cluster. Resumes from Cluster 4 get 80% consideration for Part A and 20% consideration for Part B, as shown in Table I.

Table. I Weightage assignment for cluster 4

Cluster	Salary	Part A	Part B
4	25,00,000	80%	20%

Solving for cluster 3

α =current median Salary

β =previous median Salary

C_w =current cluster weight

P_w=previous cluster weight

$C_w = (\alpha/\beta) * P_w$

C_w = (15,40,000/25,00,000) *80 = 49.2 ≈ 50

Resumes from cluster 3 get 50% consideration for Part A and 50% consideration for Part B

As shown in below Table II

Table. II Weightage assignment for cluster 4 and 3

Cluster	Salary	Part A	Part B
4	25,00,000	80%	20%
3	15,40,000	50%	50%

Similarly solving for cluster 2 P_w=50, β=15,00,000, α=8,14,000

$C_w = (\alpha /\beta) * P_w$

C_w= (8,14,000/15,00,000) *50 = 27.1 ≈ 30%

Resumes from cluster 2 get 30% consideration for Part A and 70% consideration for Part B

As shown in below Table. III

Table. III Weightage assignment for cluster 4, 3 and 2

Cluster	Salary	Part A	Part B
4	25,10,000	80%	20%
3	15,40,000	50%	50%
2	8,14,000	30%	70%

Similarly solving for cluster 1 P_w=30, β=8,14,000, α=4,40,000

$C_w = (\alpha /\beta) * P_w$

C_w= (4,40,000/8,14,000) *30 = 16.2 ≈ 16%

Resumes from cluster 1 get 16% consideration for Part A and 84% consideration for Part B

As shown in below Table. IV

Table. IV Weightage assignment for cluster 4, 3, 2 and 1

Cluster	Salary	Part A	Part B
4	25,10,000	80%	20%
3	15,40,000	50%	50%
2	8,14,000	30%	70%
1	4,40,000	16%	84%

In evaluating Part B of the resume, the DRE Algorithm considers the candidate's competitive programming skills, specifically their average contests percentile on platforms such as LeetCode and Codeforces. Additionally, for project assessment, the algorithm examines GitHub activity, taking into account metrics like stars and forks. These criteria collectively contribute to generating an overall score for Part B, aiding in the assessment of the candidate's technical proficiency and project contributions.

Validating the results:

Validating the results involves leveraging LinkedIn to gather around 150 profiles from diverse colleges across India. These profiles encompass information on individuals' undergraduate employment and academic achievements. The DRE Algorithm, as outlined earlier, is then applied to calculate scores, with a focus on skills, projects, competitive programming, and college degree.

Table. V Recruitment Verses Score Validation

Cluster	Average Score for Part A	Top Recruiters
1	10	Wipro, Accenture, Delloite, ZOHO
2	19	Microsoft, deutsche bank, Goldman Sachs, Wells Fargo
3	43	Microsoft, amazon, Samsung, Mercedes Benz
4	75	Amazon

Table. V reveals a clear correlation between higher average scores and the recruitment preferences of leading tech product companies. Clusters with elevated average scores predominantly attract top product-based companies, while those with lower scores tend to align with service-based companies. This underscores the importance of academic performance and the subsequent professional choices made by individuals, showcasing a noteworthy pattern in the recruitment landscape.

V. ALGORITHM

Function weights_assign(ranks[], median_sal):
  clusters 🡐 k_means(ranks[],median_sal[],k)
  β 🡐 max(clusters[median_sal])
  sort(clusters[median_sal])
  N 🡐 len(clusters[median_sal])
  P_w 🡐 ((max(clusters[median_sal])-min(clusters[median_sal]))/max(clusters[median_sal])
  weight[0]ßP_w
  N 🡐 len(clusters[median_sal])
  For α 🡐 N-2 **to** 1 **do**:
    C_w ← (clusters [median_sal]$_\alpha$ /β)*P_w
    weight[α] ← C_w
    P_w 🡐C_w
    β 🡐 clusters[median_sal]$_\alpha$
  end for
return weight
end function

VI. CONCLUSION

After analysing the table VI provided, one can allocate specific weights to various sections of the resume. By utilizing web scraping techniques, it becomes possible to extract candidate profiles from online sources. Subsequently, these profiles can be assessed and assigned a score based on DRE Algorithm. This scoring methodology enables a more objective evaluation of potential candidates.

Table. VI Resume weightage of different clusters

Cluster	Salary	Part A	Part B
4	25,10,000	80%	20%
3	15,40,000	50%	50%
2	8,14,000	30%	70%
1	4,40,000	16%	84%

REFERENCES

[1] JobRecoGPT: Explainable job recommendations using LLMs: Preetam Ghosh, Vaishali Sadaphal, Intelligent Service Networks, Tata Research, Tata Consultancy Services.

[2] Resume Parser Using Natural Language Processing Techniques: Shubham Bhor et. al -IJRES ISSN (Online): 2320-9364, ISSN (Print): 2320-9356

[3] Applicant Tracking and Scoring System: Ankit Tiwari, Sagar Vaghela, Rahil Nagar, Mrunali Desai - IJRET p-ISSN: 2395-0072 © 2019, IRJET | Impact Factor value: 7.211 | ISO 9001:2008 Certified Jwenal.

[4] Predictive Human Resource Candidate Ranking System: Bhagyashree Anilkumar Koyande , Rucha Suhas Walke , Mithila Gajanan Jondhale , Deepali Shrikhande - IJRE ISSN (Online): 2581-5792

[5] Attention Is All You Need: (NIPS 2017), Long Beach, CA, USA.

[6] Patil, S., & Anandhi, R. J. (2020).Diversity based self-adaptive clusters using PSO clustering for crime data . IJIT, 12, 319–327

[7] Harika, Vangala Ramanuja, and S. Sivaramakrishnan. "Image Overlays on a video frame Using HOG algorithm." 2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE). IEEE, 2020.

[8] Sivaramakrishnan, S., et al. "Eye-Based Cursor Control and Eye Coding Using Hog Algorithm and Neural Network." Innovative Engineering with AI Applications (2023): 209-225.

[9] S. S, R. B, S. V and S. Premalatha, "Time Series Analysis of Impact of Covid-19 Using Facebook Prophet Model and Review of the Machine-Learning Algorithm," 2023 International Conference on Advanced & Global Engineering Challenges (AGEC), Surampalem, Kakinada, India, 2023, pp. 151-155, doi: 10.1109/AGEC57922.2023.00040.

[10] Vandana, C.P et. Al. (2020). A Literature Review on Sentiment Analysis. IJSRCSEIT, 3(6), (204-362).

[11] D. K V, J. S N, G. B, D. G. Chandu and A. Dinesh, "Security of Cyber Physical Systems," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 170-174, doi: 10.1109/ICACCS57279.2023.10112935.

[12] K. V. Divya and R. J. Anandhi, "Analysis of the Next Release Problem to Tackle Requirement Selection," (ICICCS, 2023, pp. 912-916.

[13] Aydın et. al "A method for analyzing suspect-filler similarity using convolutional neural networks." Physical Sciences and Engineering 64.2 (2022): 129-151.

[14] Devi, BS Kiruthika, et al. "An improved security framework in health care using hybrid computing." Malaysian Journal of Computer Science (2022): 50-61.

[15] Sivaramakrishnan, S., et al. "Real Time Mock Interview Evaluation using CNN." 2024 4th International Conference on Data Engineering and Communication Systems (ICDECS). IEEE, 2024.