EightFoldAI Finals

Team Trial

September 2024

1 Data Pre-Processing and Interpreting

Firstly our goal was to convert the unstructured pdf data into a structured one, for that we used **Llama-70B** model, hosted on **Groq Cloud**. It was divided into three steps

1.1 Structuring Resumes

For each job the person has worked in we parsed the following

- Role name
- Work description
- company level (For eg. Startup, Large corporation, Medium Sized)
- Start Date
- End Date

We were able to extract meaningful insights form this data such as **Work experience**, and cross checks for **Fraud Detection**

We also parsed the following related to education:

- University
- Start date
- End date
- Domain/Degree
- Level (Masters, Bachelors etc)

These proved useful in determining the influence of the person.

Finally we parsed the following related to skills and interests:

- Projects
- Interests

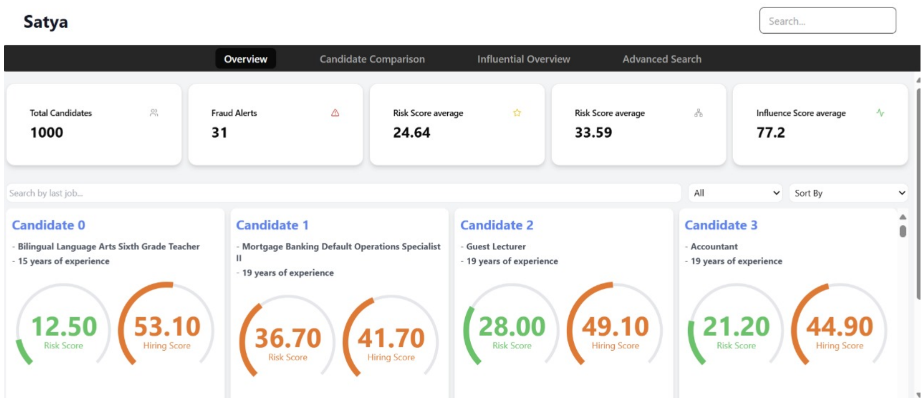


Figure 1: Overview

- Skills
  - Name
  - Level
  - Explanation

For setting the skill level, we have put benchmarks

- Identify specific skills from the work descriptions, education, and projects.
- For each skill, assign a skill level (on a scale from 1 to 5):
  - 1: Basic familiarity – The person has been exposed to the skill, but usage has been limited to beginner tasks or minimal exposure in their role.
  - 2: Some practical experience – The person has used the skill in a few practical situations but still lacks deep understanding or significant work with it.
  - 3: Solid working proficiency – The person can work with the skill regularly and efficiently with minimal supervision, demonstrating a good command of the skill.
  - 4: Strong proficiency with extensive practical experience – The person uses the skill extensively, handles complex tasks related to it, and may mentor others.
  - 5: Expert level – The person has mastered the skill, having led major projects or teams where the skill was central. They are seen as a go-to resource for this skill.
  - If someone mentions some skills in skill section, but there is no backing, give it 0.

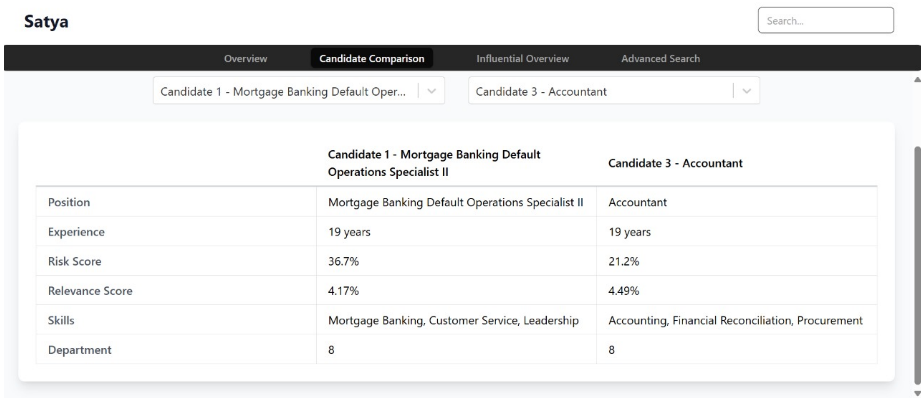
1.2 Structuring Recommendations

Now as we have extracted resume, it is the time to process recommendation letters, for this our inputs were LORs and structured resumes.
We focused on getting these three outputs.

- **Relevance**: The alignment between skills, roles and accomplishments along with supporting details. Ranked from 1 to 10.
- **Fraud Chance Level**: Contradictory data, improper language, obvious signs of exaggeration, no concrete examples to support claims etc. Ranked from 1 to 5.
- **Verified Skills**: Skills from the resume which are also present in the recommendation along with supporting details. It is a common problem HR faces and something applicants try to take advantage of. Our model verifies skills by cross checking between various texts like job histories from the CV, skills in the CV, LORs received from the recommenders. Our model also assigns scores to the skills of the applicants by considering supporting details and such similarities.

The Benchmarks for these were:

- Relevance Scores (1 to 10)
  - 1-3 (Low): The recommendation is barely relevant to the resume, with little or no connection between the skills, experience, or roles mentioned.
  - 4-6 (Moderate): The recommendation mentions some skills or experiences from the resume but lacks substantial detail or alignment with the key areas of expertise.
  - 7-9 (High): The recommendation covers many of the key skills, roles, or accomplishments from the resume, with strong alignment and relevant details.
  - 10 (Exceptional): The recommendation is highly relevant, fully aligning with all key skills, experiences, and roles from the resume, providing strong support for the candidate's qualifications.
- Fraud Chance Level (1-5)
  - 1 (Low): The recommendation seems genuine and well-supported by the resume. No informal language, statistically improbable data, or red flags.
  - 2 (Moderate): Some minor inconsistencies or exaggerated claims are present, but the recommendation remains largely credible.
  - 3 (Concerning): Multiple claims seem exaggerated or unsupported, with a lack of concrete examples or evidence from the resume.



	Candidate 1 - Mortgage Banking Default Operations Specialist II	Candidate 3 - Accountant
Position	Mortgage Banking Default Operations Specialist II	Accountant
Experience	19 years	19 years
Risk Score	36.7%	21.2%
Relevance Score	4.17%	4.49%
Skills	Mortgage Banking, Customer Service, Leadership	Accounting, Financial Reconciliation, Procurement
Department	8	8

Figure 2: Candidate Comparison

  - 4 (High): Several aspects of the recommendation are unrealistic, with improper language or statistically improbable claims.
  - 5 (Very High): The majority of the recommendation seems fraudulent, with obvious signs of exaggeration, improper language, or contradictory data.

Points we have covered:-

- Informal Language: Check for the use of casual or informal language inappropriate for a formal recommendation. If found, increase the fraud chance level by 1.
- Statistically Improbable Claims: Analyze whether the claims made in the recommendation seem unrealistic given the candidate's domain, experience level, or position. If such claims are present, set the fraud chance level to 5 and provide reasoning.

1.3 Evaluation of Influence in the industry

Now we focus on the influence and credibility of the person in the industry and domain.

We have decided it on the basis of three parts, summing them gives the value of the person.

- Job Role Level (Score: 1-5)
  - Executive Leadership (5 points): Roles like CEO, CFO, COO, CTO, President, Vice President.
  - Senior Management (4 points): Roles like Director, Senior Manager, Head of Department.

- Middle Management (3 points): Roles like Manager, Team Lead, Supervisor.
- Professional Staff (2 points): Roles like Senior Engineer, Senior Analyst, Specialist.
- Entry-Level Positions (1 point): Roles like Junior Engineer, Assistant, Associate.

- Company Size and Influence (Score: 1-3)
  - 3: Large Multinational Corporations or Highly Influential Organizations (3 points): Fortune 500 companies, major global brands, top-tier consultancies.
  - 2: Medium-Sized Companies or Well-Known Organizations (2 points): Regional leaders, established mid-sized companies, recognized NGOs.
  - 1: Small Companies or Lesser-Known Organizations (1 point): Startups, local businesses, small nonprofits.
- Experience and Tenure (Score: 0-2)
  - Extensive Experience (2 points): Over 10 years in their field or role.
  - Moderate Experience (1 point): 3 to 10 years in their field or role.
  - Limited Experience (0 points): Less than 3 years in their field or role.

Calculation:
- Add up the points from each category to get the total **Influence Factor Score out of 10**.

We also mentioned the number of years of experience. If end date is given as 'Present' (or similar meanings) then take "September 2024" as the present date. If not able to find the working experience store NaN

2 Key findings and Performance Metrics

2.1 Risk Score

Used the following
I. Fraud detection score by analyzing recommendation letter, which we have covered in above part
II. We have created a code to detect cycles using Depth First Search, among the peers. For cycles ¡ 7 persons the risk score is increased.
III. If the recommendations given by a person have high Fraud rates too then the risk score of that person increases.

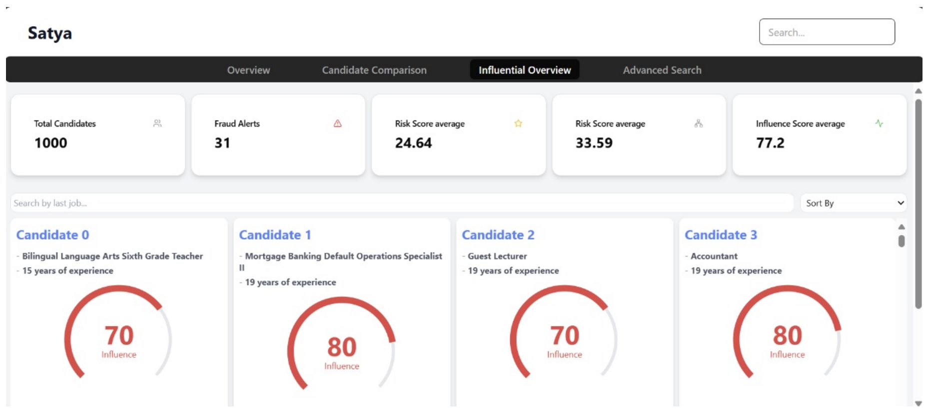


Figure 3: Enter Caption

Calculations Firstly assume we have i recommendation letter, each with fraud score : x1, x2 . . . xn We have calculated

$$\frac{\sum (x_i - 1)/4}{n}$$

.

Secondly if someone gives recommendation which are fraud, then that person 100-risk score decreases by 10 percent for fraud score of 4 and 20 percent for fraud score of 5.

if there are cycles:
Then for cycle of 3 : 10 percent is reduced
For cycle of 4 : 8 percent is reduced
For cycle of 5 $\rightarrow$ 6 percent reduced
For cycle of 6 $\rightarrow$ 4 percent reduced

In this way we have calculated the risk score

2.2 Hiring Score

It is defined by:
We have taken recommender's id, checked their safety score = 100-risk score
Then we have allocated weights proportional to the square of the safety score of that recommender to the relevance score for that recommender's ID.
And then we have used these weights, to get the weighted mean of relevance scores, let's call it x.

Hiring score =

$$x \cdot [\text{influence}(i)]^{1.5} \cdot \left(1 - \frac{\text{risk of that person}}{100}\right)^2$$

2.3 Advanced Search

Now we are going to implement advanced search for getting better results.

So for each ID, we have taken the job history of the person, and combined the text, and created an **embedding** using **paraphrase-MiniLM-L12-v2** model, this is the **semantic part**, we get results by comparing **cosine similarities**.

For keyword based search, we have implemented **bm25 algorithm**, for making it a hybrid search.

2.4 Front end and Back end

Front end is done using react, tailwind CSS
Back end is done using flask

3 Bias and Fairness : A Discussion

- We have considered company influence of 2/3 for recognized NGOs but some NGOs may have a large social impact given the new social media world. This can cause them to have a large influence.
- A CEO in a small company may not necessarily have larger influence then an associate in a very large corporation
- For fairness we have considered weights of 3 to company size and 5 to positional hierarchy. These weights may be easily changed as more data flows in.

4 Scalability, Optimization and Cost Analysis

- For the embedding part, we could use ONNX runtime, which would make the embedding creation part more faster.
- For vector search, we could use vector db such as milvus, pinecone. They use algorithms such as HNSW for a faster search, in O(logn) complexity
- For LLM part, we have taken a mid sized LLM (70B), by using LORA, it's size could be decreased, and made faster and cheaper. Still Llama 3 - 70b is cheap, costing 90 cents for 1 millions token.

- For Backend hosting, we could use aws lambda function, which could scale up very fast, and scale down as well.
- For Frontend, we could use AWS or netlify.