**Limitations**

**1. Education and Experience Extraction from Resume**

* **Current behavior:** The system does **not extract education or experience from the raw resume PDF**.
  + It relies on the applicant’s **MongoDB fields**: experience and education.
* **Limitation:** If these fields are missing, incomplete, or inconsistent, the system **cannot score the applicant accurately**.

**2. Parsing Job Eligibility**

* **Current behavior:** The system treats the **entire eligibility field** as one text block.
  + Both experience and education are compared to this **single text** using semantic embeddings.
* **Limitation:**
  + There’s **no separation of education vs experience**.
  + If the eligibility text includes multiple requirements (soft skills, certifications, experience, education), the similarity score may be **diluted**.
  + Hard numeric or categorical requirements are **not enforced**.

**3. Experience Scoring**

* **Current behavior:** Semantic similarity compares applicant’s experience text to the entire eligibility text.
* **Limitation:**
  + Cannot enforce numeric experience thresholds (e.g., 5 years required, 4 years candidate).
  + Candidate slightly below requirement may still get a **high similarity score**.

**4. Education Scoring**

* **Current behavior:** Same as experience — semantic similarity compares applicant’s education field to eligibility text.
* **Limitation:**
  + Cannot enforce exact degree requirement.
  + Candidate with a different but similar degree may get **partial credit**.
  + Transformer embeddings may overestimate similarity for loosely matching degrees.

**5. Fuzzy Matching vs Structured Comparison**

* **Observation:** System relies heavily on **SentenceTransformer embeddings** for all scoring (skills, experience, education).
* **Limitations:**
  + Semantic similarity is **fuzzy**, not precise.
  + No strict verification of requirements (years, degree, certifications).
  + Overlaps between unrelated text in eligibility may inflate scores accidentally.

**6. Skills Matching**

* **Current behavior:** Extracts resume text and compares with job skills & description.
* **Limitation:**
  + Only works if the resume is text-readable (PDF parsing may fail for images or complex layouts).
  + No handling of missing skills or synonyms beyond what embeddings capture.

**7. Translation Dependency**

* **Current behavior:** Translates all text to English.
* **Limitation:**
  + Accuracy depends on langdetect + GoogleTranslator.
  + Errors in translation may reduce similarity scores.

**✅ Summary Table**

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Current Behavior** | **Limitation** |
| Resume Education & Experience | Uses DB fields | Does not parse from PDF |
| Job Eligibility | Entire text used | No separation of education/experience; fuzzy matching |
| Experience Scoring | Semantic similarity | Cannot enforce numeric years requirement |
| Education Scoring | Semantic similarity | Cannot enforce exact degree requirement |
| Skills Matching | PDF text embeddings | Fails on non-text PDFs; synonyms/keywords may be missed |
| Translation | Auto-translate to English | Translation errors can reduce accuracy |

💡 **Bottom line:**  
The system **works as a fuzzy semantic ranking engine**, but **cannot enforce strict eligibility rules**. It relies on embeddings for both experience and education matching, which can **inflate scores for partially matching candidates**.