Technical Report: Smart Screen AI

1. Resume Screening

1.1 Problem Statement

Manual resume screening is inefficient and error-prone, especially with high volumes of applications.

1.2 Solution Overview

Smart Screen AI offers two automated resume screening solutions:

Version 1: TF-IDF + Rule-Based Filtering

- **Input:** Job description and resumes (CSV)
- **Processing:**
  - **TF-IDF Vectorization:** Extracts keywords and computes relevance scores
  - **Rule-Based Filtering:** Applies custom rules (e.g., required skills, experience thresholds)
- **Output:**
  - List of resumes matching job criteria, ranked by relevance

Version 2: BGE Embeddings + Semantic Matching

- **Input:** Job description and resumes (CSV)
- **Processing:**
  - **BGE Embeddings (HuggingFace):** Generates semantic vectors for job and resumes
  - **Cosine Similarity:** Computes semantic relevance between job and each resume
  - **spaCy NER:** Named entity recognition for skill extraction
- **Output:**
  - List of resumes ranked by semantic fit and entity relevance

1.3 Workflow

1. **Upload Resumes and Job Description**
2. **Process Using Selected Version**
3. **Rank and Filter Candidates**
4. **Output:**
    - **Version 1:** CSV or JSON with ranked resumes (TF-IDF + rules)
    - **Version 2:** CSV or JSON with ranked resumes (BGE + NER)

1.4 Key Challenges & Solutions

Challenge	Solution
Unstructured resume formats	spaCy NER and sentence segmentation
TF-IDF ineffective for vague resumes	Integrated BGE embeddings for semantic scoring
Large resumes exceeding limits	Trimmed resume text to 4000 characters

---

2. Employee Sentiment Analysis

2.1 Problem Statement

Unaddressed employee dissatisfaction leads to increased attrition. Early detection and intervention are essential.

2.2 Solution Overview

A Flask-based API analyzes employee feedback, predicts attrition risk, and suggests engagement strategies.

2.3 Workflow

- **Input:**
  - Employee feedback (text)
- **Processing:**
  - **TextBlob:** Sentiment analysis to score feedback
  - **Feature Vector:** Combines sentiment and tenure data
  - **RandomForestClassifier:** Predicts attrition risk
  - **Rule-based Recommender:** Suggests engagement strategies
- **Output:**
  - **API Endpoint:**
    - **POST `/predict`**
    - **Request Body:**

```
{
  "feedback": "Employee feedback text here",
  "tenure": 2
}
```

    - **Response:**

```
{
  "sentiment": -0.5,
  "risk": "high",
  "recommendation": "Schedule a one-on-one meeting to discuss workload and recognition."
}
```

2.4 Key Challenges & Solutions

Challenge	Solution
Label encoding errors	Removed LabelEncoder; used raw sentiment
TextBlob sentiment often neutral	Added rule-based keyword detection
Deployment issues on Render	Added pre-deploy command for TextBlob corpora

---

3. Deployment

- **Resume Screening:**
    - Batch processing or API integration (endpoints to be defined as per implementation)
- **Sentiment Analysis:**
    - **API Endpoint:**
        - **Hosted on Render:**
            - **POST** `https://employee-feedback-ww3a.onrender.com/predict`
        - **Supports JSON input/output**

---

4. Conclusion

Smart Screen AI delivers:

- **Two robust resume screening solutions** for efficient candidate selection
- **Employee sentiment analysis API** for proactive attrition management