



Resume Analyzer- Resume Reveal

By:
Matan Cohen
Shira Shani
Eden Menahem





Motivation:

- Reduce time, bias and inconsistency in high-volume hiring through automated, objective resume screening.

Process Overview:

- Input: resume Dataset + job title.
- Output: seniority level (Junior, Mid, Senior).

Relevant NLP Tasks:

- Context extraction and seniority classification.

NLP Task:

Value Proposition:

- Faster, consistent skill matching and actionable feedback for candidates.

Potential Applications:

- Personalized job matching, enterprise talent management

Development Challenges:

- Diverse resume formats, nuanced context interpretation, seniority calibration, bias mitigation.

Training and test data:

Data Type:

Unstructured text from resumes and job titles.

Supervised Learning Approach:

All resumes are tagged with seniority levels corresponding to the job title.

Labels: Junior, Mid, Senior.

Data Sources:

Synthetic Data Combined with Real Dataset:

AI Resume Generation Using a Prerequisite Given Real Resume Examples

(For example, using the dataset "<https://www.kaggle.com/datasets/wahib04/multilabel-resume-dataset>")

Prompt example

ChatGPT ▾

↑ Share



Use the following real resumes as inspiration for tone, structure, and content. Do not copy or paraphrase directly. Instead, use them as reference points to guide the realism and design of your resumes created for tech jobs.

Create resumes for the following seniority levels: Junior, Mid, and Senior.

Make sure the resulting resumes remain unique and do not include information that identifies the source resume.

Then, follow the rest of the instructions:

Do not explicitly state the level of experience.

Subtly indicate this through job titles, responsibilities, projects, and education.

Change in style, tech stack, and professional background.



Data Usage for Evaluation:

- Loss Monitoring: Tracks Cross-Entropy Loss to evaluate convergence.
- Score: A scale from 0 to 10 to rate the level of fit for the position

Evaluation Metrics:

- Accuracy
- Confusion Matrix

Baseline Model & Comparisons

- Baseline Model: BERT – A deep learning model using bidirectional context for better text understanding.