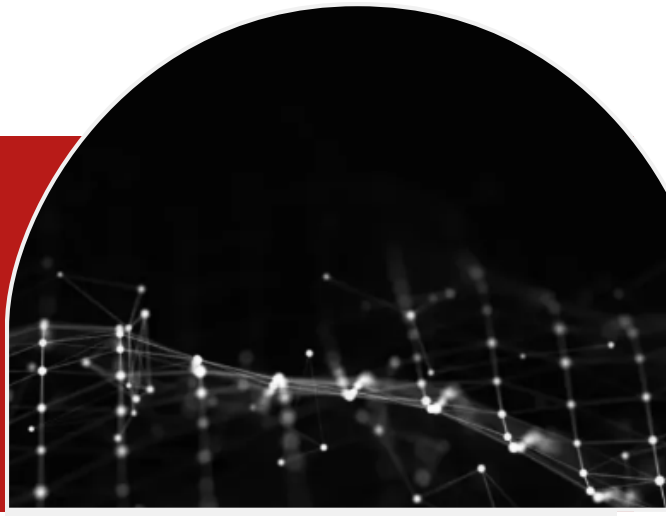
YEARLING AI

SCALABLE RESUME PROCESSING FOR ENTERPRISES

Smarter Decisions.
Faster Hiring.
Stronger Teams.



INTRODUCTION

Resume processing in bulk is a tedious, time consuming and difficult task for Human Resource people. Medium and Large enterprises receive thousands of resumes every month for a variety of job openings. Furthermore, job search portals like jobs.com, Monster.com and countless others also need to process resumes at scale. In this project, Yearling AI has implemented a scalable resume processing solution for enterprises that automates resume processing and mining data from them in bulk, and thus helps reduce the time and manual labor of extracting information from resumes. The solution developed in this project is a critical component of office process automation for enterprises and human resource companies.

CUSTOMER STORY

This solution was developed based on requirements from two customers via Yearling AI's consulting partner on Google Cloud Platform. Currently, it's being demonstrated to both clients, with a software license agreement in progress. The customers aim to eliminate manual resume processing for hundreds of thousands of resumes. Key requirements include:

- Parsing resumes and storing key terms for easy searchability.
- Extracting sections like Education, Experience, Skills, and Contact Info.
- Identifying named entities (e.g., universities, companies) for quick lookup.

OUR SOLUTION

We developed an end-to-end machine learning pipeline for resume processing with the following steps:

1. **Text Extraction:** Detect and extract text from PDF or DOC resume files.
2. **Embedding Creation:** Generate text embeddings for sentences/paragraphs using NLP-based vectorization.
3. **Clustering:** Use unsupervised learning to group text into sections like Education, Experience, and Contact Info.
4. **Named Entity Recognition:** Identify entities such as organizations and locations within each section.
5. **Data Storage:** Store original text, extracted information, and embedding vectors in a datastore.
6. **Search Indexing:** Build global and intra-resume search indexes to facilitate efficient keyword and phrase searches.

KEY TECHNOLOGY & FRAMEWORK

Technology

- Optical Character Recognition (OCR) for Text Detection
- OCR for Text Recognition
- Natural Language Processing
- Clustering

Framework

- Pytorch-Lightning
- Huggingface Transformers
- Jupyter Notebooks
- Pandas
- Numpy
- FastAPI
- Google Cloud Storage
- Google Kubernetes Engine

CUSTOMER BENEFITS

This solution is a critical component of office process automation for enterprises and human resource companies. It can reduce human engagement in the tedious process of extracting useful information from potentially hundreds of resumes daily. This leads to significantly more efficient office operations and saves valuable time for the Human Resource personnel.

ABOUT YEARLING AI

We build AI that works. At YearlingAI, we bring deep technical expertise to solve complex problems with machine learning, natural language processing, and generative AI. From intelligent automation to custom LLM agents, we design, build, and deploy solutions that drive results. As a Google Cloud partner, we specialize in cloud-native development—but also support AWS, Azure, and hybrid environments. Whether you're a growing startup or a global team, we deliver practical AI solutions that scale with your needs.