

Ajayeswar Reddy Peddyreddy

+1 (341)-732-8139 ♦ ajayeswarb2a@gmail.com ♦ [linkedin.com/in/ajayeswarreddy/](https://www.linkedin.com/in/ajayeswarreddy/) ♦ [Website](#)

A results-driven engineer with experience in the data team, committed to designing and developing scalable end-to-end data pipelines that drive operational efficiencies and enable data-driven decision-making. Passionate about leveraging emerging technologies and data management best practices to optimize data processing workflows and enhance business performance.

Education

Master of Science in Data Science – The University of San Francisco, USA *Jul 2022 – Jun 2023*

Courses: Data Acquisition, Relational Databases (SQL), Exploratory Data Analysis, Distributed computing and Data Systems (Apache Spark, MongoDB), Data Structures & Algorithms, Machine Learning, MLOps

B.E Electrical and Electronics – BITS PILANI, India *Jul 2016 – May 2020*

Professional Experience

AWS ML labs **San Francisco, USA**

Research intern *Oct 2022 – Present*

- Developing new algorithms and training pipelines to accelerate neural network training, optimizing model performance and accuracy.
- Implemented a sampling strategy for language models, resulting in faster convergence than traditional SGD and improved model efficiency.

GEP Worldwide **Hyderabad, India**
Software Engineer *Jul 2020 – Jun 2022*

Work involved handling end-to-end ML and engineering tasks for the product **OCR for invoices and credit memos**.

- Revamped the architecture of the system by migrating the API from **flask** to **FastAPI** and transitioned from synchronous to **asynchronous** processing, which resulted in a significant improvement in the reliability and cut down the error rate for processing complex invoices by more than **90%**.
- Successfully decoupled a monolithic API into a **microservices** architecture, resulting in seven independent services with improved scalability, fault tolerance, and agility.
- Developed a robust data factory pipeline in Azure (ADF) to efficiently transfer data across various environments including Dev, QC, UAT and Production, resulting in improved data accessibility and streamlined workflow.

Data Intern *Jul 2019 – Dec 2019*

- Created a middle layer for an OCR product, responsible for extracting specific data attributes from databases and comparing them with application outputs.
- Implemented matching algorithms that accurately identifies the best match from the database and provides relevant information, including a confidence score and row ID, resulting in improved accuracy and reliability of data extraction.

Projects (selected)

- **Job recommendation** – Developed an end-to-end distributed deep learning pipeline to match jobs with resumes. [GitHub](#)
 - Scraped and stored job postings from multiple sources to GCP (Google cloud), preprocessed data using PySpark on Databricks, and stored collections in Mongo atlas (**ELT pipeline**). Utilized Airflow for scheduled scraping.
 - Generated embeddings for job postings and resumes using Word2Vec and **Sentence Transformer models** (large and small).
 - Performed vector similarity search between job and resume embeddings using Pinecone achieving perfect accuracy.
- **Search Engine** – Implemented using Object Oriented hash table to retrieve matched documents from a corpus in constant time using tf-idf vectorization.
- **ML from scratch** – Implemented Decision Trees, Random Forest, Gradient Boosting, Adagrad, k-means clustering algorithms.

Technical Skills

Python (NumPy, Pandas, Matplotlib, Scikit-learn), Beautiful Soup, Fast API, Flask, Docker, Git, GitHub, Databases, SQL(Postgres), NoSQL (MongoDB), Apache Spark, Airflow, Cloud (Azure, GCP), Databricks, pipelines (ETL, ELT)