

Ajayeswar Reddy Peddyreddy

+1 3417328139 | ajayeswarb2a@gmail.com | <https://www.linkedin.com/in/ajayeswarreddy/> | [Blog](#)

Full stack data scientist with 3 years of test-driven development experience and a keen interest in building end-to-end scalable ML products and robust data pipelines to enable data-driven analytics and decision making.

Education

Master of Science in Data Science – The University of San Francisco, USA

Jul 2022 – Jun 2023

Courses: Machine Learning, Regression Analysis, Data structures & Algorithms, Statistical analysis, Exploratory Data Analysis and Data Visualization, Distributed computing, and Data Systems (Spark), Relational Databases), Data Acquisition, MLOps

B.E Electrical and electronics engineering (Minor in Finance) – BITS Pilani, India

Jul 2016 – May 2020

Courses: Neural Networks, Probability and Statistics, Linear Algebra, Differential Calculus, Object Oriented (Java), Operations Research

Professional Experience

Machine Learning Engineer intern

Oct 2022 – present

Amazon Machine Learning Solutions lab and USF Data Institute

San Francisco, CA

- Developed novel algorithms, optimization techniques, and training pipelines to expedite large-scale neural network training, while improving model accuracy and performance. Mentored by Amazon Machine Learning Solutions Lab and USF Data Institute.
- Enhanced Natural Language Processing (NLP) tasks by implementing a customized sampling strategy for language models using PyTorch data loader resulting in faster convergence rates than traditional Stochastic Gradient Descent and improved model efficiency.

Data Scientist

Jul 2020 – Jun 2022

GEP Worldwide

Hyderabad, India

As a Data Scientist, I worked on executing end-to-end Machine Learning and engineering tasks for the AI OCR product, specifically for invoices and credit memos.

- Developed and implemented a rule-based NLP pipeline for invoice processing, achieving up to 100% accuracy for select attributes and an overall accuracy range of 70-100%. Improved global usability and expanded market reach by expanding language support to 15 additional languages.
- Revamped the architecture of the system by migrating the API from Flask to FastAPI and transitioned from synchronous to asynchronous processing. Improved reliability and cut down the error rate for processing complex invoices by more than 90%.
- Successfully followed design patterns in software development and decoupled a monolithic API into a microservices architecture, resulting in seven independent services with improved scalability, fault tolerance, and agility.
- Engineered the product using Docker and Azure Kubernetes serving ~10000 requests per day and generating an estimated annual revenue of 3 million USD.

Data Scientist Intern

Jul 2019 – Dec 2019

GEP Worldwide

Hyderabad, India

- **Reverse Image Search:** Used pre-trained Res-net 50 architecture (fine-tuning computer vision model) to encode the images and used modified K-nearest neighbors to classify the images and recommend the appropriate item.
- Reduced false-positive rate from 25% to 5% using classification models and utilized YOLO for accurate object detection, to identify multiple objects from one image and recommend prediction accordingly.

Select Academic Projects

- **Job recommendation** – Developed an end-to-end distributed deep learning predictive modeling to match jobs with resumes. [GitHub](#)
 - Scraped and stored job postings (data sets) from multiple sources to Google Cloud bucket, preprocessed data using PySpark on Databricks, and stored collections in MongoDB. Utilized Airflow for scheduled and automation 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 score.
- **ML from scratch in Python**– Implemented Decision Trees, Random Forest, Adagrad, and k-means clustering algorithms.
- **Search Engine** – Implemented using Object Oriented hash table to retrieve matched documents from a corpus in constant time.
- **Mixed code sarcasm detection** – Implemented traditional Machine Learning algorithms (SVM, random forest) and bi-directional LSTM attention-based model to detect sarcasm in Hindi-English code-mixed tweets, achieving a 92% accuracy score. [GitHub](#)

Technical Expertise

Python (NumPy, Pandas, Matplotlib, Scikit-learn), TensorFlow, PyTorch, Big data (Apache Spark, Apache hive), Airflow, SQL, NoSQL (MongoDB), Databricks, pipelines (ETL, ELT), APIs (Fast API, Flask), Docker, Microsoft Azure, GCP, Version Control (Git, GitHub), testing frameworks (sonar cloud), experiments tracking (MLFlow), mathematics and statistical principles (t-test, F-test, ANOVA)