PRANEETH SEEMAKURTHY +91 9502579913

Senior Machine Learning Engineer, Quantiphi praneethkrishna27@hotmail.com

About

I am an experienced ML Engineer with 5 years of professional experience in developing machine learning solutions. I have a strong foundation in both NLP and Computer Vision and have worked on several end-to-end projects that involve designing the architecture, implementing, and deploying ML models. I have expertise in utilizing various GCP services to develop solutions for clients across different domains.I am passionate about developing innovative solutions using cutting-edge technology and always strive to stay up-to-date with the latest advancements in the field of machine learning.

Skill Set

Technical Abilities: Python, Ml/DL Modeling

AI/ML Skills: Ensembling techniques, generalized linear models, Boosting Techniques, MLPs, CNNs, RNNs, Encoder-Decoder, Transformers, Attention Mechanisms, Large Language Models, Retrieval Augmented Generators

Tools and Frameworks: Git, Pytorch, Docker

Cloud Services : GCP

Work Experience

Organization – Quantiphi Analytics Pvt Ltd: April 2022 - Present

Sr Machine Learning Engineer: I currently lead an ML Team that is responsible for delivering multiple workloads in the GCP R&D Practice. We work on creating MVPs using GCP services within short timelines. The MVPs are designed to meet the requirements coming from the Sales team and showcase our capabilities to potential clients.I specialize in developing NLP (Natural Language Processing) and CV (Computer Vision) applications

Project: **Document Processing**: working on creating a web application that can extract tabular data from PDFs that contains Tables in multiple formats

- utilizing DocAI-OCR for data extraction, custom trained form parser and LLMs for table identification, row & column information extraction
- built post processing script for extraction of relevant information and arrangement into structured format
- API creation for interacting with front end and information display

Project: **Traffic Violator Detection**: developed an AI-powered web application PoC for a Government body in New York to automate the process of detecting Traffic Noise Violators from Video Data

- created pipelines for both noise classification and violator vehicle identification using vehicle tracking and detection by using transfer-learning methodology
- developed post processing script that recognises vehicle number plates for further processing
- with the help of this automation, the time and resources required to manually monitor all the videos is drastically reduced which led to an overall increase in the revenue via fine collection

Project: **Vehicle Damage Detection**: developed an AI-powered web application for an automobile insurance company that can detect different types of damages by analyzing uploaded images of defective vehicles.

- employed transfer learning using Yolo V5 to identify multiple body parts of a vehicle, and utilized Visual Inspection AI (VIAI) to detect any damages in an uploaded image.
- Implemented the backend logic in Python and utilized several GCP services for storage and deployment

Project: **Document Insights**: developed an NLP based web application for an investment management client that is capable of analyzing multiple large document PDFs for Information retrieval

- leveraged the Doc AI Warehouse service to enable document search from an extensive corpus of documents.
- created a Custom RAG (Retrieval Augmented Generator) pipeline using PaLM APIs (text-bison model) to extract most relevant content for a Query passed from the front end which is further utilized by the downstream tasks such as Question Answering and Summarization

- This helped end-users to extract the relevant information more precisely without the need to read the whole document

Project: ArtWork Data Extraction:
developed a demo that leverages NLP and computer vision to identify printing errors in artworks. The demo involved extracting data from various entities in the artworks such as logos and nutritional information, and comparing it against the product label repository (PLR) using GCP services.

- designed end-to-end pipelines that enable the uploading of artwork PDFs, extraction of data, and comparison against the ground truth.
- developed a text extraction script using Python and utilized GCP services such as Vision AI and Doc AI to facilitate this process. The script is capable of understanding 16 different languages.
- devised the success metric for this project in terms of reducing the time required for a single artwork verification from one hour to just 10 minutes.

Organization – Gramener Technologies Pvt Ltd: March 2021 - March 2022

Project: SCAI Document Classification:

- Developed a digital twin for a pharma client for clinical trials document classification with more than 70 different artifact classes using NLP based techniques to improve the document archival process
- Performed data quality checks by defining quality metrics followed by preprocessing and text feature generation
- Worked extensively on developing classification models using ML/DL algorithms using sklearn and Hugging-face transformers libraries respectively and demonstrated the model explainability and interpretability using Lime library visually
- Mentored junior developers and interns across teams, participated in pair programming sessions, which helped them in better contribution in their projects

Project: NLP Reusable components

- Developed reusable, customizable NLP pipelines for training deep neural networks using transfer learning techniques for text classification
- this pipeline has been used across several teams within the Organization and reports indicated reduced model training time by upto 20% and improved the prediction performance in terms of accuracy and f1-score

Project: Semantic Search Engine.

- Worked on designing and building an NLP Based Search System POC that finds a relevant document in Elasticsearch DB based on a input query
- used BOW, TF-IDF, W2V, USE pre-trained algorithms for creating high dimensional word/sentence vectors
- used ElasticSearch Indexes for storing the data vocabulary as Inverted Indices

Organization – Tata Consultancy Services: Aug 2018 - March 2021

Systems Engineer
Project: Aviation Digital Hub, (Leaders in Aircraft Engine Manufacturing):

- created airplane sensor data archive by building data pipelines in python that pulls sensor data from airplane components in near real-time, parses and then ingests into PostgreSQL in Azure cloud
- built generic python scripts that parses data from various unstructured file formats using beautiful soup
- received client appreciation for building and scaling up the data pipelines that seamlessly consumes data from multiple airline fleets

Achievements and Awards

Awarded Certificate of Appreciation in recognition for outstanding contribution to TCS-R2 Data Labs Engagement in the year 2018-19.

Education

B-Tech, Civil Engineering, KL University Vijayawada – 2015: CGPA- 7.77 || HSC – 86.8% || SSC 86.6%