Sai Raghavendra Viravalli

(617)-379-3062
sairag1999@outlook.com
(He/Him)
Boston,US
Linkedin.com/sai
Github.com/sai

Education

Masters of Science in Information Systems, Northeastern University **Expected Dec 2023**

- Data Science Engineering Methods Tools | Data Management and Database Design | Big Data Systems and Intelligent Analytics| Advanced Data Science Architecture| Applied Machine Learning in Finance | Parallel Machine Learning AI

Bachelor of Technology in Computer Science Engineering, Vellore Institute of technology **Jul 2016 - June 2020**

Technical Skills

Data Science Frameworks:	Pandas, Keras, NumPy, Pytorch, Tensorflow, NLTK, sci-kit-learn, Streamlit
Languages:	C++,Python,Shell,Git,R, SQL, NestJS, NodeJS
Machine Learning:	Logistic Regression, Light GBM, Time Series forecasting, Natural Language Processing, Random Forests
Data Visualization/Analytics:	Data studio,Tableau, Microsoft Power BI
Cloud tools:	Big-Query, Airflow, Spark, Beam,GCP,AWS
Version Control Tools:	Git, Bitbucket, GitHub, Gitlab

Experience

Executive Analyst **Mar 2021 - Aug 2021**
Deloitte Touche Tohmatsu Ltd Bengaluru,India

- Designed and developed back end services for a private bank's net banking systems using **NestJs React** replacing previous version resulting in 25% faster transaction response time
- Part of the **Agile** process of the project and setup **Jira** boards accordingly with respect to business demands during the planning and design meetings with the Business clients and parties involved in the project
- Broadened client's Database systems by using **AWS EC2** services and queried using **PostgreSQL** to produce data resulting in better access of more cleaned data, cost savings, scalability, and ease of management
- Created and developed **Microservices** according to Business demands for several **APIs** using **gRPC** resulting in lowering client's costs by 10x times and 15% more efficient

Networking and Security Intern **May 2018 - Jul 2018**
Sify technologies Chennai,India

- Devised a chatbot using **Recurrent Neural Networks (RNN), Long Short Term Memory (LSTM) and Natural Language processing** in **python** for the networking department resulting in making their usage and saving time by 15 percent
- Built networking tools such as **ping, traceroute, ipconfig** amongst other tools in python and automated multiple **shell** commands to execute in a distributed way to save time as it includes large data

Academic Projects

Performance analysis of Plant-disease Detection over multiple GPUs | Deep learning,Python **Aug 2022 - Nov 2022**

- Implemented distributed parallel training of **CNN**s for plant disease detection tasks on the kaggle dataset resulting in comparison of performances on multiple **CNN** architectures over multiple **GPU**s. Utilized **Automatic Mixed Precision (AMP)** for speedup.
- Performed preprocessing of the entire dataset initially using the **multiprocessing** module and then compared the results with serial processing over mulitple **Nvidia Tesla v100 GPU**s.
- Evaluated performances of existing DNNs like **Mobilenetv2, VGG16, Resnet34** and ran the computation for each of the smaller mini-batches which were split in parallel using **Pytorch**

Image Captioning | Deep learning, Natural Language Processing,Python **May 2022 - Aug 2022**

- Generated captions for the images from the flicker 8k Database where the text will be summarized accordingly. Using the combination of **CNN** we extract features and use **RNN.**
- Utilized **LSTM and GRU** for the caption generation. Extended the model using Hyperparameter tuning and compared their results
- Calculated the **BLEU** scores for the captions produced according to the images and tested the accuracy across multiple models

TWITTER MARKETING CAMPAIGN ANALYSIS | Python,GCP,AWS, SQL,Docker **Jan2022 - May 2022**

- Based on the hashtag input and date range given by the user, the WebApp can analyze the relevant tweets, utilizing Huggingface's Sentiment Analysis and Named Entity Recognition to model-as-a-service deployed on **AWS ECR** as **Docker** containers on the **Serverless** framework and **AWS Lambda** to perform Sentiment analysis on tweets and derive entity names
- Locations from the named entities were used to find and display relevant news using an open-source News API along with real-time metrics on a **Google Data Studio** dashboard embedded in the WebApp.
- Created an **ETL** Pipeline Ingested data from **Twitter API** and **Google Pub/Sub** and associated Cloud functions to Store Raw data in **Google BigQuery** all packaged within **Airflow** running on **Google Cloud Composer**
- For Transform and Load in ETL, leveraged **Apache Beam** to set up an ad-hoc pipeline and an hourly pipeline ingesting raw data,cleaning and loading into a processed **BigQuery** dataset using **Google Cloud Dataflow** as the orchestration tool for Beam.