Akshat Sanghvi

217-721-8155 akshat72525@gmail.com MyPortfolio LinkedIn GitHub

Experience

TESLA Fremont, California
Data Scientist *Nov 2023 – Present*

- Implemented **machine learning forecasting models** (Prophet, Croston) to optimize supply chain efficiency for 26k Service Parts across 11 Distribution Centers, using parallel processing in **Python** to reduce training time from **18 hours to 1.5 hours**
- Engineered a solution using **GitHub Actions** to automate multiple tasks, including forecast generation, exception management, and consumption file creation, reducing manual workload by **80%** and ensuring consistent file preparation for Supply Chain teams weekly
- Spearheaded a team of three in developing an automation pipeline to send demand forecasts for 14,000 service parts with intermittent demand to suppliers, eliminating manual intervention and reducing the workload of demand planners **by at least 50%**
- Constructed an **ETL** pipeline to upload 40+ local CSV, Excel, and Parquet files into **MySQL** database using Python, ensuring updated data for cross-functional teams and achieving **100% data automation**
- Developed a **Streamlit** web app for reviewing time series plots, improving forecast analysis by **60%** and enabling real-time feedback, enhancing collaboration between analytics and supply chain teams

GIES BUSINESS SCHOOL Champaign, Illinois
Data Science Research Associate *Aug 2023 – Nov 2023*

- Led research with Prof. Aravinda and an Indian NGO to promote rural girls' education. Analyzed pre- and post-intervention surveys using **t-tests**, showing a **36% increase** in positive attitudes toward education, demonstrating the project's success.
- Utilized Spark to analyze 1M+ credit consumers' data with 200+ attributes. Built and deployed a credit default classification model on **AWS (SageMaker)** using Random Forest and Decision Tree models in **PySpark MLlib**, achieving **60% accuracy**

WALMART Bentonville, Arkansas
Data Science intern *May 2022 – Aug 2022*

- Coordinated cross-functional collaboration to develop an end-to-end forecasting model for Cases per Trailer (CPT), potentially **saving $1M** and **reducing 4320 man-hours per year**
- Extracted over **five million rows** of CPT data from GCP BigQuery using SQL, followed by Python-based Exploratory Data Analysis for identifying trends, patterns, and seasonality, along with data cleaning and feature engineering
- Applied ARIMA, XGBoost, and Markov Chain based forecasting techniques, complemented by rolling cross-validation and back testing, to attain a robust **94% accuracy for non-seasonal forecasts** and an **88% accuracy for seasonal forecasts**
- Designed performance metrics KPIs to track and monitor the continuous improvement of model performance via a **Tableau** dashboard, ensuring effective communication and reporting to non-technical stakeholders

Projects

Image Captioning using Transformer | [Link] | *Python, PyTorch*

- Applied image rotation as a pretext task to train a ResNet18 encoder, achieving robust feature representation for image captioning
- Built a Transformer decoder with custom positional encoding and attention layers, trained on 30,000 images to generate descriptive captions

Generative Adversarial Networks in PyTorch | [Link] | *Python*

- Coded LSGAN and DCGAN architectures from scratch using PyTorch to generate high-quality synthetic images
- Experimented with different network architectures and hyperparameters to enhance GAN training stability and improve image quality

Land Cover and Crop Type Segmentation | [Link] | *Python, TensorFlow*

- Used TensorFlow to create pixel-level labels based on crop-type maps from Cropland Layer images provided by USDA
- Incorporated UNet to segment Corn, Soybeans from other crops on RapidEye Satellite image with 85% Pixel accuracy data

Technical Skills

Languages: Python, R, C/C++, Java, MATLAB, SAS, SQL, NoSQL, MongoDB, Neo4j
Data Tools: NumPy, Pandas, SciPy, Tableau, PowerBI, PyTorch, TensorFlow, Gurobi, SciKit-Learn, Microsoft Excel
DevOps Tools: AWS (S3, SageMaker), CI/CD pipelines, Git, GCP BigQuery, Docker, Kubernetes, Apache Spark, Hadoop
Concepts: Machine Learning, A/B Testing, Neural Networks, Computer Vision, Exploratory Data Analysis

Education

University of Illinois, Urbana-Champaign Champaign, Illinois
Master of Science in Industrial Engineering *May 2023*

University of Mumbai Mumbai, India
Bachelor of Engineering in Mechanical Engineering *July 2021*