Akshat Sanghvi

217-721-8155 | akshat72525@gmail.com | MyPortfolio | linkedin.com/akshatssanghvi | github.com/akshat3492

Experience

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

- Optimized supply chain efficiency by implementing advanced demand forecasting methods (Prophet, Croston) for 26k Service Parts across 11 Distribution Centers, leveraging parallel processing in Python to reduce training time from **18 hours to 1.5 hours**
- 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%**
- Enhanced forecast accuracy and improved failure rate estimates by **15%** for service parts inventory management using NLP's fuzzy matching techniques (Levenshtein distance, cosine similarity) to match new parts with existing ones
- 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 web application in **R** to streamline inter-team communication by automating the generation of Jira issue tickets. Integrated with Jira's **RestAPI** for seamless ticket creation and ensuring data accuracy through robust validation mechanisms

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

- Led research with Prof. Aravinda and an Indian NGO to encourage rural girls' education. Conducted statistical analysis (t-tests) on pre- and post-intervention surveys, indicating a noteworthy positive attitude shift **(approx. 36%)**, validating project 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 May 2022 – Aug 2022
Data Science intern *Bentonville, Arkansas*

- Led cross-functional collaboration to develop an end-to-end forecasting model for Cases per Trailer (CPT), potentially **saving $10M** and **reducing 4320 man-hours per year**
- Utilized SQL to extract CPT data from GCP BigQuery, followed by Python-based Exploratory Data Analysis for trend, pattern, and seasonality identification, along with data cleaning and feature engineering
- Leveraged 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

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 May 2023
Master of Science in Industrial Engineering *GPA: 3.6 / 4.00*

- **Relevant Coursework:** Deep Learning for Computer Vision, Cloud Computing, Analysis of Network Data. Statistics of Big Data, Algorithms of Data Analytics

University of Mumbai Jul 2021
Bachelor of Engineering in Mechanical Engineering *GPA: 3.89 / 4.00*

- **Relevant Coursework:** Object Oriented Programming, Operations Research, Applied Mathematics, Statistics & Probability, Database and Information Retrieval

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

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

Spotify – AI in Business | [Link] | *Python*

- Developed a content-based music recommendation system using KNN to suggest five similar music tracks to users
- Performed feature importance using the Random Forest on over 160k tracks containing 60+ features
- Predicted skip probability of a user with an 88% accuracy by hyperparameter tuning of LightGBM using Optuna