Mushrifah Hasan

mushrifahhasan@gmail.com | +919920536165 | Mumbai, India
Linkedin | Website

SUMMARY

Motivated and detail-oriented Data Scientist with 3+ years of experience in machine learning, predictive & data analytics. Enthusiastic about Deep Learning, collaborative, positive attitude and always exploring and learning.

PROFESSIONAL EXPERIENCE

General Mills | *Mumbai, India* ***Nov 2023 - Present***

Data Scientist

- Optimized and automated ML-based regression model for demand prediction and price elasticity using **GCP, Python and Airflow**, reducing process time by 90%, enabling seamless scalability to new markets with minimal manual intervention, and enhancing the model by incorporating a component to handle holiday effects.
- Analyzing a retailer's household-level data and developing ML models using **Databricks, Python, and PySpark** to increase household penetration through targeted digital coupon campaigns, driving engagement and sales.
- Developing an **embeddings-based ML model** to analyze and identify key competitive clusters down to the UPC level, enhancing competitive intelligence and decision-making.
- Collaborate with cross-functional teams to deliver ad-hoc decision analytics and actionable insights tailored to stakeholder needs, enabling informed commercial investments and driving revenue growth.

Mobicule Technologies Pvt. Ltd | *Mumbai, India* ***Sept 2021 - Nov 2023***

Data Scientist ***Oct 2022 - Nov 2023***

- Improved the debt collection efficiency and recovery of assets by identifying patterns in payment behavior and predicting which customers are most likely to default using **classification and clustering-based ML algorithms.**
- Developed an AI-driven debt collection bot using **Rasa, Python, Text-to-Speech, and Speech-to-Text**, reducing operational costs and efficiently capturing key collection status information.
- Increased the response time and user interaction, by developing a POC chatbot to answer aggregate-based queries on a private database (i.e. text to SQL) using **OpenAI API, Python, Langchain and Rasa**.
- Developed end-to-end data pipeline for analytics dashboard with **Python, Airflow, and Superset**, resulting in reducing dashboard downtime and increasing stakeholder engagement.
- Partnered closely with different teams to develop AI-driven solutions for the mCollect debt-collection platform, implementing a POC for automated campaign rules to personalize communication timing and channels.

Machine Learning Intern ***Sept 2021 - Sept 2022***

- Worked closely with teams across functions to design and develop dashboards using **SQL, Superset & Grafana** that track debt collection (digital campaigns & physical) metrics KPIs across hierarchy levels.
- Developed a clustering-based ML model to classify customers and implemented a rule-based algorithm to automate the allocation of achievable targets to debt collectors based on location and predicted customer type.

TECHNICAL SKILLS

- **Languages/Tools**: Python, SQL, Google Cloud, Apache Superset, Docker, Git, Tableau, Grafana, Pyspark, R
- **Libraries/Frameworks**: Pandas, NumPy, Matplotlib, Scikit-Learn, MLflow, Flask, Rasa, Kubeflow, Airflow, FastApi, Streamlit, TensorFlow, Spacy, BeautifulSoup, PyTorch, FastAI, hugging face libraries

PROJECTS

Depression Detection Based on Sentiment Analysis in Social Media Using Deep Learning **Sept 2022**

- Implemented a two-step depression detection system with data scrapped from Twitter using deep learning language modeling in **Tensorflow,** and deployed the model as a **web application** with **Flask**.

Stress Detection in Tomato Plants with Thermal Images Using Deep Learning **Dec 2020**

- Implemented a non-invasive multi-modal analysis technique for stress detection in tomato plants with thermal images, trained a CNN-based image classification model using **Pytorch**, and deployed it using **Flask**.

PUBLICATIONS

- Data-driven Depression Detection System for Textual Data on Twitter using Deep Learning, IEEE, 2022
- Application of Deep Learning Coupled with Thermal Imaging in Detecting Water Stress in Plants, Book: Design of Intelligent Applications using Machine Learning and Deep Learning Techniques, 2021
- Image Processing based Application of Thermal Imaging for Monitoring Stress Detection in Tomato Plants,IEEE,2019

EDUCATION

Sardar Patel Institute of Technology, Mumbai ***Dec 2020 - Sept 2022***

MTech, Computer Engineering (CGPA: 9.85/10)

Computer Vision Nanodegree, Udacity ***Oct 2019 - Feb 2020***

University of Mumbai, Mumbai ***Jul 2016 - Nov 2020***

B.Tech, Computer Engineering (CGPA: 8.12/10)