Shubham Shailesh Tamhane

shubhamtamhane2000@gmail.com linkedin.com/in/shubhamtamhane
github.com/shubhamtamhane shubhamtamhane.github.io

Education

University of Rochester **Aug 2022 – Dec 2023**
Master of Science in Data Science *Rochester, NY*

- GPA: 3.96/4. Recipient of 40% merit scholarship
- Secured 2nd position in the 2022 UR Biomedical Data Science Hackathon

University of Mumbai **Aug 2018 – May 2022**
Bachelor of Engineering in Information Technology *Mumbai, India*

- GPA: 3.73/4, CGPA: 8.95/10
- Hackathon winner and Guest speaker for creating video conferencing web application (Google Meet clone)

Relevant Courses: Time Series, Data mining, Statistics, NLP, Machine Learning, AI, Data Structures, Big Data, DBMS

Experience

Regeneron Pharmaceuticals **Jun 2023 – Dec 2023**
Data Science Intern *Tarrytown, NY*

- Implemented time series forecasting approach to predict customer demand of a complex **inventory management** problem employing multiple approaches including **statistical** and **deep learning** methods.
- Deployed a webapp built using **python-dash** that leverages **MLOps** workflow built on cloud-infrastructure to provide real-time up-to date data and forecasting predictions, customer analysis and model maintenance options to end users contributing significantly to **cost optimization**.
- Led the development of a **maintenance analysis** system, optimizing the upkeep of MFCs and related systems, which resulted in substantial monthly savings.

URMC - Center for Advanced Brain Imaging and Neurophysiology **Sept 2022 – Jun 2023**
Software Intern *Rochester, NY*

- Engineered a **verification system** to confirm over **10,000** dicom medical files within **60** seconds using pydicom.
- Developed a Flask web service with **JWT** authentication and caching, reducing large JSON file load times to under 10 seconds. Deployed the service using **Docker** for improved system efficiency.

Sciffer Analytics Pvt Ltd **Oct 2020 – Jan 2021**
Data Science Intern *Pune, India*

- Managed the development of image datasets using labelimg tool for **information extraction** from Google in 3 months empowering a computer vision model to recognize over 30 distinct objects.
- Employed the **YOLO v3** model to build a deep learning classifier model, attaining an accuracy rate of **80%**.

Technical Skills

Programming Languages: Python, R, C, C++, Java, Spark, SQL
Data Manipulation and Visualization: MySQL, MongoDB, Tableau, PowerBI, JMP, Excel
Framework and Libraries: Sklearn, OpenCV, Tensorflow, Keras, Pandas, Numpy, Ggplot2, Pytorch, MLFlow
Machine Learning Methods: Regression, Clustering, NLP, Computer Vision, Object Detection, Speech Recognition
Web Technologies: HTML5, CSS3, Django, Flask, Dash, Streamlit, Nodejs, JavaScript, Express
Cloud Tools and Project Management: AWS, Databricks, Docker, Git, Dataiku, Seeq, Jira, Confluence, Agile

Projects

Emotion Recognition Using Deep Convolutional Neural Networks | Publication Link **Apr 2022**

- Neural networks such as **ResNet50** and **VGG16** were used to identify the mood of the user based on facial expression.
- Applied **Haar Cascades** on the FER2013 dataset, followed by a custom deep convolutional neural network (DCNN) to achieve an accuracy of **83.9%** .
- **Tech Stack:** Python, OpenCV, Sklearn, Tensorflow, Keras, Youtube, Streamlit, Spyder

Predicting and Analysing the Viral Fragments of Songs | Project Link **Dec 2022**

- Implemented **dynamic time warping** for comparing extracted MFCC features in songs, leveraging a weighted SVM classifier to achieve a performance of 100% recall and 86% accuracy.
- **Tech Stack:** Python, Pandas, Matplotlib, Pytube, Apafy, Librosa, Imagehash, PIL, Sklearn, Jupyter Notebook

Dynamic QA generator for Research Papers | Project Link **May 2023**

- Fine-tuned a T5-base model to create a **Question-Answer** system that generates and answers questions from research papers, enhancing paper interpretation.
- Utilized the **QASPER** dataset to evaluate models, employing metrics such as **BLEU** score, **ROUGE**, and **QAeval**
- **Tech Stack:** Python, OpenAI, Pandas, Numpy, Huggingface, Simpletransformers, Pickle, Google Colab