Shubham Shailesh Tamhane

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

EDUCATION

University of Rochester **Rochester, NY**
Master of Science: Data Science *Aug 2022-Dec 2023*
- GPA: 3.95/4. Recipient of 40% merit scholarship
- Secured 2nd position in the 2022 UR Biomedical Data Science Hackathon

Ramrao Adik Institute of Technology **Mumbai, India**
Bachelor of Engineering: Information Technology *Aug 2018-May 2022*
- GPA: 3.73/4, CGPA: 8.95/10,
- Hackathon winner for creating video conferencing web application. Was invited next year to give guest lecture

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

EXPERIENCE

Regeneron Pharmaceuticals, Data Science Co-op *Jun 2023-Present*, **Tarrytown, NY**
- Developed a **demand forecasting** algorithm in **JMP**, which was used to predict protein demands from a central repository, resulting in a time saving of **8** hours per week.
- Instituted an **inventory management** system that effectively prevented **$25,000** worth of material from being wasted each quarter.
- Formulated a **predictive maintenance** system to ensure proper upkeep of MFCs and related systems saving **$10000**/month

URMC – Department of Neuroscience, Software Intern *Sept 2022-June 2023*, **Rochester, NY**
- Implemented a system using the "**pydicom**" library that verifies if over **10,000** medical files in "**dcm**" format, anonymized by a specific function, are present in the output, all within **60** seconds.
- Developed a web service in **Flask** having **JWT** authentication in combination with caching features to reduce loading time of large JSON files to under **10 seconds** and deployed on **Docker**

Exposys Data Labs, Data science intern *May 2021-June 2021***, India**
- Performed knowledge mining and segmented data by implementing **k-means clustering** with **85%** accuracy

Sciffer Analytics Pvt Ltd, Data science intern *Oct 2020-Jan 2021***, India**
- Developed and annotated image datasets by extracting information from Google, using tools such as "labelimg". This process enabled the training of a machine learning model that successfully identified more than **30** objects within a span of **3** months.
- Utilizing the **YOLO** v3 model, a deep learning classifier model was built, achieving an accuracy rate of **80%.**

Department of Information Technology, RAIT, Intern *Jun 20-Jul 20 & Dec 19-Jan 20*, **India**
- Engineered a multi-user video communication application utilizing **Express** and **Node.js**.
- Implemented competitive programming practices, resulting in a **10-50%** performance optimization in C/C++.

SKILLS

- **Programming Languages :** Python, R, C, C++, Java, Spark
- **Data Manipulation and Visualization:** Oracle SQL, MySQL, MongoDB, Tableau, PowerBI, Dataiku,Seeq, JMP
- **Framework and Libraries :** Sklearn, OpenCV, Tensorflow, Keras, Pandas, Numpy, ggplot2, pytorch
- **Machine Learning Methods :** Time series, ARIMA, CNN, SVM, Transfer-Learning, Computer Vision
- **Web Technologies :** HTML5, CSS3, Django, Flask, Nodejs, , JavaScript, Express

PROJECTS & PUBLICATIONS

- Tamhane, Shubham, et al. "**Emotion Recognition Using Deep Convolutional Neural Networks**." SSRN doi.org/10.2139/ssrn.4091264
  A **deep convolutional neural network (DCNN)** was created and used to identify the mood of the user based on his facial expression. Accuracy of over **83.9%** was achieved.
- **Dynamic QA generator for Research Papers**
  Fine-tuned a T5-base model to create a QA system that generates and answers questions from research papers, enhancing paper interpretation
- **Predicting and Analysing the Viral Fragments of Songs**
  Utilized **dynamic time warping** to compare extracted **MFCC** features from songs, and applied a weighted **SVM** classifier. Achieved 100% recall and 86% accuracy due to presence of data imbalance.