Shubham Shailesh Tamhane

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

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

University of Mumbai **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.

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**
- Leveraged **JMP** to conduct a comprehensive analysis and forecast of protein demands, enhancing operational efficiency and culminating in a weekly time saving of **8** hours.
- Devised and implemented a robust **inventory analysis** system, strategically preventing a quarterly material wastage worth **$25,000,** thereby contributing to cost optimization.
- Spearheaded the development of a **maintenance analysis** system, optimizing the upkeep of MFCs & related systems, which resulted in a substantial monthly savings of **$10,000**.

URMC – Department of Neuroscience, Software Intern Sep*t 2022-Jun 2023*, **Rochester, NY**
- Engineered a verification system using pydicom library to confirm over **10,000** anonymized 'dcm' medical files within **60** seconds.
- 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.

Exposys Data Labs, Data science intern *May 2021-Jun 2021***, India**
- Implemented clustering algorithms for customer segmentation & knowledge mining achieving 85% accuracy.

Sciffer Analytics Pvt Ltd, Data science intern *Oct 2020-Jan 2021***, India**
- Managed the development of image datasets using 'labelimg' tool for information extraction from Google in **3** months. This empowered 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%.

Department of Information Technology, RAIT, Software Intern *Jun 2020-Jul 2020* **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, Hadoop
- **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:** Time series, ARIMA, CNN, SVM, Transfer-Learning, Computer Vision
- **Web Technologies:** HTML5, CSS3, Django, Flask, Dash, Streamlit, Nodejs, JavaScript, Express
- **Cloud Tools and Project Management:** Databricks, Dataiku, Seeq, Jira**,** Confluence, Agile

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.
- **Forecasting Bike Inventory for Citibike** Link
  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.
- **Dynamic QA generator for Research Papers** Link
  Fine-tuned a **T5-base** model to create a **QA** system that generates and answers questions from research papers, enhancing paper interpretation.