Tanvi Nikhil Shroff

(213) 272-4022 • tnshroff16@gmail.com • linkedin.com/in/tanvi-shroff • tshroff16.github.io/portfolio-tanvishroff/

Graduate student in Computer Science from the University of Rochester with 1 year of industry experience. Passionate about creating solutions to solve complex business problems. Adept in fostering cross-functional collaboration and driving project success through communication and strong technical expertise in data science.

EDUCATION:

University of Rochester, Rochester, New York **August 2022 – May 2024**
Master of Science, Computer Science (Data Science) GPA: 3.83 / 4.0
Coursework: Computational Statistics, Data Mining, Time Series Analysis, Statistical Machine Learning

Savitribai Phule Pune University, Pune, India **August 2018 – May 2022**
Bachelor of Technology, Information Technology GPA: 3.80 / 4.0
Courses: Machine Learning, Advanced Machine Learning, Business Intelligence
Awarded the **'Cummins – Excellence Award'** as the Best Outgoing Student Award for 2021-2022

SKILLS:

Languages:	Python, R, SQL, Java, C++, PowerQuery
Technologies:	Tableau, Git, JIRA, Databricks, MLflow, Jupyter Notebook, PowerPoint, Excel
Libraries and Frameworks:	PyTorch, PySpark, Tensorflow, NumPy, Pandas, Matplotlib, NLTK, scikit-learn
Statistical Methods:	Hypothesis Testing, ANOVA, ROC curve
Machine Learning Algorithms:	Regression, Random Forest, Gradient Boosting, XG Boost, Clustering, Forecasting

WORK EXPERIENCE:

Data Analyst Intern, *Paychex* **September 2023 – December 2023**
- Discovered opportunities for revenue **optimization** and identified **trends** by analyzing customer behavior data.
- Built a **random forest classification** model to predict sales for Paychex's payroll solution with a recall of 90%.
- **Optimized** complex **SQL** queries by eliminating view dependencies and streamlining joins reducing runtime by 50%.
- **Automated** data insertion from Excel to MySQL Server using Python and streamlined data integration process using **Git**.
- **Collaborated** with Sales, Marketing, and Data Engineering teams to ensure seamless data integration and data accuracy.

Data Warehouse Intern, *University of Rochester* **June 2023 – August 2023**
- Developed an **XGBoost** classification model to predict the academic performance of students with an F1 score of 87%.
- Developed a **Tableau** dashboard to monitor multiple model performances ensuring comprehensive oversight of model efficacy.
- Fostered strong **collaboration** with **cross-functional** teams to implement the predictive model, ensuring smooth integration with Cognos, Denodo, and Oracle SQL systems.

Data Science Intern, *Women in Data Science* **May 2021 – July 2021**
- Analyzed consumer complaints for electronic devices by global companies using **Advanced Excel** and **SQL**.
- Designed an automated web scraper for extracting >1MM customer reviews, performed EDA using **Python**, and visualized sentiments using **Tableau**.
- Implemented **topic modeling** to identify common complaints and provided solutions, reducing complaints by 10%.

Founder, *The Lavender Whisk* **June 2020 – June 2022**
- Managed advertising, inventory planning, and finances as the founder of a bakery startup.
- Analyzed product-level **time-series** data and social media trends to design profitable bakery products using **Excel**.

RESEARCH & ACADEMIC PROJECTS:

Justified Representation in Multi-Winner Approval Voting, *University of Rochester*
- Employed **algorithmic design** to ensure computational effectiveness in implementing Justified Representation.
- Conducted empirical studies using **inferential statistics** and mathematical theories to explore multi-winner approval voting.

CITI Bikes Forecasting Model, *University of Rochester*
- Developed an end-to-end **pipeline** for a bike-sharing system using **Databricks**, **PySpark**, and real-time **structured streaming**.
- Built an **FB prophet forecasting** model to predict hourly net bike change for a station.
- Developed a real-time monitoring application utilizing **Git** and **MLflow** for model tracking and registry.

English to Indian Sign Language, *Savitribai Phule Pune University*
- Led a team of 4 to generate a 3D avatar with hand signs using speech-to-text API and phrase re-ordering.
- Leveraged **NLTK** python library to implement **tokenization** and **lemmatization** for text analytics.
- Accelerated the inference latency by 500%, by caching the SiGML files containing HamNoSys notations.

PUBLICATIONS:

- Shroff, Tanvi, *et al.* "Literature Review on Machine Translation Systems for Sign Language Generation." International Conference on Data Management, Analytics & Innovation. Springer, 2023