

Shrikar Thodla

734-389-5134 • thodla.shrikar@gmail.com • San Francisco, CA • sthodla.github.io • github.com/sthodla • U.S. Citizen

Education:

C++ Nanodegree: Udacity

Nov 2020-Present

M.S. in Data Science: University of San Francisco

June 2020

- Scholarships: MS in Data Science Scholarship, Wicklow AI in Medicine Research Initiative Scholarship.
- Relevant Courses: Python, Machine Learning, SQL, Time Series Analysis, Distributed Computing, Data Acquisition, Data Ethics, Product Analytics, A/B Testing & Experimental Design, Data Structure & Algorithms.

B.S. in Informatics, Minor in Mathematics: University of Michigan

Dec 2017

- Co-Founder and Treasurer of Hoops For Hope Club.
- Relevant Courses: Linear Algebra, Coding Theory, Probability and Statistics, Data Mining, Computational Biology/Bioinformatics.

Work Experience:

AI Research Engineer at Riiid Labs!

Aug 2020-Present

- Increased model accuracy for knowledge tracing (transformer-based model) by 5% using data augmentation.
- Decreased model MAE (Mean Absolute Error) for the score prediction model by 11.4 (for reference this is score prediction for SAT exams which are out 1600) by using data augmentation and more data.

Data Science Intern at Retrace Labs

Oct 2019 - June 2020

- Standardized images that were accidentally rotated by training and deploying a convolutional neural network (CNN) that detected rotation with 99.7% accuracy (PyTorch).
- Decreased execution time of MongoDB queries by 5x in order to load images onto the virtual instance faster by batch querying and multiprocessing (Python).
- Led the migration of full-stack components from Google Cloud (GCP) to IBM by setting up virtual server instances and user access management rules.

Associate at Infosys Limited

Feb 2018 - Apr 2019

- Decreased the number of security vulnerabilities in production and non-production servers by 50% by leading the activities between the application and patching teams.
- Communicated updates on patching and database administration activities to stakeholders.

Projects:

Bengali Character Classification (Kaggle Competition) - [Code](#)

Jan 2020 - Mar 2020

- Top 10% in the competition out of 2059 teams (Bronze Medal).
- Predicted 3 different labels for 200k+ Bengali characters and achieved a 0.93 average recall score using CNNs (PyTorch).
- Implemented findings from research papers, like Cutmix and Mixup, to help improve model performance.

StreamHopper - [Code](#)

Mar 2020 - May 2020

- Developed and pitched a web application that recommends which streaming service users should choose based on their preferences to stave off subscription fatigue (Python).
- Data for all the shows and movies on each streaming platform was pulled from two APIs using Python and standardized to join the data.

ML Algorithm Implementations

Oct 2019 - Feb 2020

- Recreated implementations of popular machine learning algorithms in Python, such as Random Forests, Decision Trees, Naive Bayes, K-Means, TF-IDF, and multiple feature importance algorithms.

Technical Skills/Courses:

- Skills: Python (scikit-learn, NumPy, pandas, SciPy, spaCy, Flask), PyTorch, PostgreSQL, AWS, git, R (ggplot2), Spark (PySpark, SparkSQL, SparkML)