SWETHA SRIKARI MAGANTI

Somerville, Massachusetts 02143 | (857) 500-9633 | mswethasrikari@gmail.com
swethasrikari.github.io | linkedin.com/in/swetha-maganti

I hold a Master's degree in Machine Learning and I'm looking for Data Scientist / ML Engineer roles to utilize my skills and create an impact. I worked on collaborative and independent projects covering various domains including health care, climate, social media and language models leveraging Python and open-source ML tools.

SKILLS

- Course work: Machine learning | Artificial Intelligence | Data Science | Data analysis | Representation learning | Statistical modeling | Natural Language Processing | Computer Vision | Data Structures and Algorithms | Probability and Statistics
- Data sets: Structured | Unstructured
- Programming Languages: Python | Matlab | R | LaTeX | HTML
- Database: MySQL | SQLite
- Machine learning / Deep learning libraries: Pandas | Numpy | Scikit-learn | NLTK | Tensorflow | Pytorch | SpaCy | Scikit-survival
- Data Visualization: Tableau | Matplotlib | Plotly | Seaborn | Shap
- Other: Git | Linux | Jupyter notebook

WORK EXPERIENCE

Machine learning Intern | Perceiv AI, Montreal | 01/2021 – 07/2021

- Worked with clinical data extracted using ICD 9 and ICD 10 codes.
- Performed data manipulation, feature engineering and prepared data pipelines for training
- Applied resampling techniques to tackle the problem of class imbalance
- Explored ensemble models and a variety of machine learning models
- Built Framingham and European SCORE risk models for a 10-year risk assessment of cardiovascular disease.

(*Python, Pandas, Numpy, Scikit-learn, Git, Linux, Jupyter notebook, Scikit-survival, Matplotlib, Plotly, Shap, Docker*)

EDUCATION

M.Sc, Computer Science | Mila, University of Montreal, Montreal | 09/2019 - 11/2021

B.E, Electronics and Instrumentation | Madras Institute of Technology, Anna University, Chennai | 08/2015 - 05/2019

CERTIFICATION

Tableau 2022 A-Z: Hands-On Tableau Training for Data Science | 2022
SQL for Data Science | 2022
R Programming, Data Science specialization | 2022
Regression models, Data Science specialization | 2022
Statistical Inference, Data Science specialization | 2022

PROJECTS

Image generation using diffusion model

Images are generated from complete noise using a diffusion (generative) model. The forward diffusion process involves gradually corrupting the original images with noise at each timestep. The images are

then recovered by training a model to predict the noise added at each timestep in the reverse diffusion process.

(*Python, Pytorch, Diffusion model, Matplotlib, Git, Jupyter notebook, Linux*)

Low Resource Machine Translation

To overcome the lack of a large aligned dataset, source-side monolingual data was exploited using a self-learning technique similar to back translation, which strengthens the encoder and boosts the performance of the transformer by generating synthetic data.

(*Python, Tensorflow, Numpy, Pandas, Scikit-learn, Git, NLTK, Jupyter notebook, Linux*)

Solar Irradiance Prediction

Implemented a 3-D convolutional neural network to nowcast the solar irradiance from GOES satellite imagery.

(*Python, Tensorflow, Numpy, Pandas, Scikit-learn, Git, Jupyter notebook, Linux*)

User Profiling in Social Media

Trained a model to construct a user profile i.e., infer age, demographics, and Big 5 personalities by leveraging textual, visual, and relational cues from social media.

(*Python, Pandas, Numpy, Scikit-learn, Jupyter notebook, Linux*)

Multi-class text classification

Performed comparative analysis of various language models, hyper-parameter configurations, and training regimes to boost the accuracy of classification.

(*Python, Pandas, Numpy, Scikit-learn, NLTK, SpaCy, Matplotlib, Git, Linux, Jupyter notebook*)

OTHER ACTIVITIES

- Adopted a tree - Cambridge Urban Forest Friends
- Volunteered at Computer Society of Madras Institute of Technology
- National Service Scheme volunteer, Anna University