Nimish Magre
(+1)617-817-0148 | nimishmagre@gmail.com | Boston, MA
LinkedIn | Kaggle | GitHub

EDUCATION

Northeastern University Boston, MA
Master of Science in Electrical and Computer Engineering August 2022
Concentration: Computer Vision, Machine Learning and Algorithms GPA: 3.9/4.00
Coursework: Deep Learning, Advanced Computer Vision, Fundamentals of Computer Engineering

Australian National University Canberra, Australia
Bachelor Of Engineering (Honors) July 2019
Concentration: Mechatronic Systems GPA: 5.75/7.00

TECHNICAL SKILLS

- **Programming Languages:** Python, C++, SQL
- **Python Libraries:** Numpy, Pandas, TensorFlow, Pytorch, OpenCV, Scikit-Learn, Matplotlib, SHAP, H2O, SciPy
- **Tools:** MATLAB, Git, Jupyter, MySQL, Anaconda, LaTeX, Visual-Studio, Gazebo, Docker, JIRA, Confluence, AWS-S3, AWS-EC2

INTERESTS

- Synthetic Data Generation, Model Explainability, Multi-modal learning, NLP

WORK EXPERIENCE

Squark AI (no-code AI SaaS and API for prediction/forecast models)
Python Developer (Data Science) **Dec 2022-Present**

- Responsible to build feature extraction and data curation pipeline.
- Presented model explainability features through SHAP (SHapley Additive exPlanations) summary graphs.
- Developed a weighted-averaging method to present Variable Importance for H2O based Stacked-Ensemble model.
- Integrated multi-thread processing to speed up the process for model explainability through SHAP.

Massachusetts General Hospital
Machine Learning Research Intern **Jan 2022-Jun 2022**

- Implemented a self-supervised denoising network to denoise MRI scans (helped in reducing the time required for patients to complete an MRI scan).
- Achieved average PSNR results at least 1.5dB higher compared to relevant single image based denoising methods for Additive White Gaussian Noise.
- Assessed the model with real world noisy dataset (PolyU) and attained an average PSNR value of 37.52 dB, similar in comparison to the best performing model at 37.55 dB.

AI Skunkworks (Northeastern University)
Graduate Student Researcher (Arxiv) **Dec 2021-Dec 2022**

- Designed a python-based pipeline to generate a synthetic dataset of 565,292 MNIST-style grayscale images representing 1,812 unique glyphs in styles of 1,355 Google-fonts for typography analysis.
- Published an introductory paper for the dataset on the *Arxiv* platform as a first author and trained a custom CNN-based classification model to identify non-Latin glyphs with ~97% accuracy.
- Inspected glyph identities and eye-tracking tools to effect real time mapping of type (fonts) to cognitive properties; generated “Panose” number identities used to characterize fonts for the entire dataset.

PROJECT EXPERIENCE

Northeastern University

Time Frequency Weighted Overlapping Group Shrinkage for Speech Denoising **Aug 2022-Dec 2022**
(GitHub)

- Adapted the l1-norm cost term to sparsify speech samples iteratively and introduced a mixed-norm non-separable penalty term to promote group sparsity and remove residual noise.
- Customized a time-frequency weight matrix to make the algorithm effective with impulsive noise.
- The project paper is currently under review for the IEEE-ICASSP conference.

Northeastern University

Siamese network for object-tracking (GitHub) **Jan 2021-Apr 2021**

- Reported a detailed literature review for the ResNet-50 based Siamese Box Adaptive Network (SiamBAN) for single object-tracking.
- Modified the template patch for track using correlation between initial, t-1 and search frames to detect occlusion and improve inference-time tracking performance when multiple object instances were present.
- Displayed differentiation of target object by model when partial occlusion and multiple instances of the object were present with qualitative testing on sample VOT-2018 video dataset.

KEY PERSONAL PROJECTS WITH GITHUB HYPERLINKS

- Image-based Sudoku Solver (C++): applied key image-processing and constraint problem solving techniques to solve a sudoku puzzle through the image of an unsolved puzzle.
- LSTM-based single instrument music generation: Represented guitar and violin samples as midi files to train the LSTM-based RNN model to predict after every 100 notes. For music generation, supplied a sequence of random 100 notes with a random offset value to generate 500 output notes.
- ML Research Paper Reviews: Comprehensive review of recent papers that focus on Advances in Computer Vision and techniques to induce sparsity with Big Data.
- ML-based Data Insights for NGO: Performed synthetic data generation, sentiment analysis on volunteer feedback and used data from a volunteer's visit log to generate scikit-learn based classification/regression models.

HONORS AND AWARDS

- ANU College of Engineering and Computer Science International Partnership Scholarship
- Bronze medal on Kaggle for typography dataset *(TMNIST, TMNIST_Glyphs)* and 2-year data hosting sponsorship by AWS