Mughilan Muthupari

Data Scientist

mughilan.muthupari@gmail.com

240-779-6852

14115 Bear Creek Drive, Boyds, MD 20841

www.linkedin.com/in/mughil-pari/

www.github.com/MughilM

SKILLS

- **ML technologies:** PyTorch (incl. PyTorch Lightning), Tensorflow 2.x
- **Specializations and Interests:** NLP, image recognition, reinforcement learning, model interpretability
- **MLOps experience:** MLFlow, Weights and Biases, AWS Sagemaker, Apache Spark, Jenkins, Github Actions
- **Programming Languages:** Python, R, SQL, Java
- **Cloud Platforms:** AWS, Google Cloud Platform

EDUCATION

M.S / Data Science

Columbia University / New York, NY

Sep 2019 – Dec 2020

Classes Taken Include: Applied Deep Learning (*computer vision*), Data Science and Public Policy (*anomaly detection, regression*), Data Analysis and Visualization, etc.

Bloomberg Capstone Researcher (NLP)

- Improved sentence word structure through thematic fit by utilizing roles and words using **Tensorflow.**
- Measured effectiveness of various word embeddings (XLNet, FastText, Glove) and network structural changes on performance and evaluation.
- **Continued research after graduation resulted in paper published at EMNLP Blackbox and ready on arXiv:** https://arxiv.org/abs/2208.04749

B.S / Computer Science & Statistics

University of Maryland – Gemstone Honors Program / College Park, MD

Aug 2015 – Dec 2018

Team DIVA Research

- Created a novel climate visualization tool constructed with virtual reality in a team of multi-disciplinary students using **Python and Unreal Engine.**
- Awarded the James M. Wallace Outstanding Gemstone Thesis Award. Available for reading at UMD DRUM: https://doi.org/10.13016/wgrz-tatt

WORK EXPERIENCE

Data Scientist II

Battelle / Columbus, OH / Mar 2022 – Present

- Developing an image recognition model to examine possible biological pathways in rice and corn
- Analyzed employee data for workplace shocks as Model Lead
- Modeled evolving bioplumes using anomaly detection with graph neural nets and MLPs to detect changes in smoke concentration
- Implemented ML methods (clustering, RNNs) to detect patterns in human-generated random digits for an authentication system
- **Technologies:** Python, PyTorch, MLFlow, CNN, GNN, RNN

Staff Developer

FINRA / Rockville, MD / May 2019 – Mar 2022

- Worked with explainability techniques (LIME, Shap) for predictions on raw OCR'd PDFs and communicated to stakeholders on results as Team Lead
- Developed and maintained an advert ML classification system using CNNs and various transformer models to detect riskiness
- Prototyped abstractive document summarization using BERT
- Utilized CNNs to archive U5 filings to improve efficiency
- **FINRA Internship (Summer 2018)** – Predicted complaint types of disclosures with CNNs and LSTMs.
- **Technologies:** Python, PyTorch, Jenkins, Sagemaker, BERT, LongFormer, CNN, LIME, Shap

Software Intern

Moody's Investors Service / New York, NY / Jun 2020 – Aug 2020

- Optimized labels in a weak supervision NER task by generating refined labels, reducing labor costs by ~70%

Research Intern

NASA Center for Climate Simulations / Greenbelt, MD / Jun 2017 – Aug 2017

- Used the Advanced Data Analytics Platform (ADAPT) to examine historic daily temperature cycle and do other statistical calculations