Nirmal Amirthalingam

Blacksburg, VA | +1 (540) 824-8595 | nirmal@vt.edu | | | |

EDUCATION

Virginia Tech, Blacksburg, VA — May 2023

Master of Engineering in Computer Engineering — *GPA: 3.96/4*

Coursework: Advanced Machine Learning, Deep Learning, Computer Vision, Data Analytics, Information Visualization

Amrita Vishwa Vidyapeetham, Coimbatore, India — Jun 2020

Bachelor of Technology in Electronics and Communication Engineering — *GPA: 8.57/10*

Coursework: Pattern Recognition, Optimization Techniques, Image Processing, Probability and Random Processes

SKILLS

Programming: Python, SQL, C++, MATLAB

Libraries: PyTorch, TensorFlow, OpenCV, NumPy, SciPy, Keras, Scikit-learn, NLTK, spaCy, Pandas, Matplotlib, Seaborn

Tools/Frameworks: Docker, Git, GitLab, PostgreSQL, MongoDB, Spark, Kafka, Airflow, Tableau, D3.js, Flask

Cloud: AWS (SageMaker, Glue, Lambda, Athena, RDS, S3, EC2), Azure ML Studio

EXPERIENCE

Research Assistant — Aug 2022 - May 2023

Virginia Tech — *Blacksburg, VA*

- Implemented **YOLOv7** and **Faster R-CNN** to extract objects from Electronic Theses and Dissertations (ETD)
- Worked on self-supervised text-image (multimodal) pre-training using **transformers** for large-scale unlabeled corpus
- Parsed to a structured format after filtering detections, and developed read/write APIs for **PostgreSQL** database
- Integrated to a retrieval and search system that supports **50K ETDs** by providing **Dockerized** services with **CI/CD**

Machine Learning Intern — May 2022 - Aug 2022

AreaProbe — *Washington, DC*

- Developed an **object detection** and **tracking** model for pedestrians and vehicles based on **YOLOv5** in PyTorch
- Incorporated the model into production system for **35 RTSP cameras** in nearby parking lots and housing communities
- Integrated with a **MobileNet audio detection** pipeline with an SMS trigger to detect **gunshots** in real-time
- Performed **color recognition** on saved image crops for vehicles, and tracked object counts for the **surveillance** system
- Deployed the model using AWS **EC2** instances, and stored detected objects in **S3** bucket and structured data in **RDS**

PROJECTS

Loan Default Prediction — Aug 2022 - Nov 2022

- Implemented **gradient boosting** methods to assess credit risk and predict customer defaults from LendingClub dataset
- Performed **EDA** and data pre-processing to address class imbalance, and **PCA** to reduce feature dimensionality
- **XGBoost** classifier with Grid Search for hyperparameter tuning achieved an accuracy of **87 %** and AUC of **0.93**

Text-to-Image Generation using GANs — Mar 2022 - May 2022

- Experimented with Generative Adversarial Networks (**GAN**) for **text-to-image** synthesis in PyTorch
- Implemented a **deep fusion** model with text-guided image manipulation on the Caltech-UCSD Birds (**CUB**) dataset
- Reduced model parameters by a third, while achieving the same **FID** of 10.2 and **IS** of 5.9 as the baseline model

Hybrid Recommendation System — Jan 2022 - Mar 2022

- Designed a hybrid recommender system using user and item-based **collaborative filtering** to predict user ratings
- Used **Spark** to analyze and train model on **5.2M user ratings** with text reviews for restaurants from Yelp Open dataset
- Achieved **RMSE** score of **0.979** on the validation set using collaborative filtering with matrix factorization and SVD

Trending YouTube Video Analytics using Sentiment Analysis — Feb 2022 - May 2022

- Analyzed key performance indicators for trending videos in the US using **sentiment analysis** with NLTK
- Extended to trending videos in multiple countries by pre-processing raw JSON files using **AWS Lambda** and **Spark**
- Designed ETL pipeline using **Glue** to load analytical data into S3 bucket for creating visualizations in **QuickSight**