SYED DANISH AHMED

+1 (412) 708 5459 | syeddanish@cmu.edu | linkedin.com/in/syeddanishahmed/ | github.com/SyedDanishAhmed

Education

CARNEGIE MELLON UNIVERSITY — Pittsburgh, PA
Master of Information Systems Management, BIDA – Data Science, **QPA: 3.89** — Dec 2020
Relevant Courses: Machine Learning, Artificial Intelligence, Unstructured Data Analytics, Statistics, Econometrics, Business Analytics

S.G.S.I.T.S — Indore, India
Bachelor of Engineering in Information Technology, GPA: 3.81/4.00 (First Class with Distinction) — Jun 2015
Relevant Courses: Data Structures, Algorithms, Database Management, Data Mining, Linear Algebra, Distributed Systems

Skills

- **Machine Learning:** Regression, Classification, Segmentation, Dimensionality Reduction, Gradient Boosting, Deep Learning - DNN, CNN, RNN
- **Functional:** Natural Language Processing, Anomaly Detection, Time Series Forecasting, Hypothesis Testing, A/B Testing, Agile - Scrum
- **Languages:** Python, Java, R, SQL, Microsoft SCOPE
- **Business Intelligence and Visualization Tools:** R Shiny, Tableau, Power BI, Excel VBA
- **Data Engineering & Deployment:** IBM MQ, Flask, Hadoop, Spark, AWS, MongoDB, SSIS, Azure - Blobs, Data Factory, ML, Cosmos, PowerShell

Academic Experience

CARNEGIE MELLON UNIVERSITY — Aug 2019 – Present

- Developed a model for Musculoskeletal Disorder physiotherapy exercise recognition using sensor data from accelerometers
  - Generated features by windowing with overlap and used Discrete Cosine Transform (DCT) for feature transformation and selection
  - Implemented a **1D Time Distributed CNN – LSTM** architecture using **Keras and Tensorflow** & achieved AUC of 97% over seven exercises
- Implemented a **Multilayer Perceptron (MLP)** classifier for text-based Emotion Classification using **Word2vec word embedding**
  - The model was able to identify nine different emotions, like joyful, terrified, jealous, etc. with an accuracy of 77%
- Implemented a **Gradient Boosting – LightGBM** classifier with **PCA** on Vesta's real-world **high-dimensional** e-commerce transactions dataset
  - The model can improve the efficacy of fraudulent transaction alerts with an AUC of 94%
- Built a predictive model using **scikit-learn** to guide profitable loan investments using the historical dataset for loans issued on LendingClub
  - Built classification models to predict loan default probability and regression models to predict the expected return on each loan
- Developed a semi-supervised **autoencoder model** in **Keras** to recognize fraudulent credit card transactions using reconstruction error
- Built a Landmark Classification model by **Transfer Learning** using a pre-trained **ResNet18** CNN in **PyTorch** with 85% accuracy over ten classes

Work Experience

PricewaterhouseCoopers (PwC) Digital Transformation and Innovation Center at CMU — Pittsburgh, PA
Data Science Research Intern — Jun 2020 – August 2020

- Developing a Machine Learning framework for **fraud detection** on transactional data using unsupervised **Anomaly Detection** techniques

MU SIGMA: Data Analytics provider for 140 Fortune 500 Companies — Bengaluru, India
Senior Decision Scientist – Data Analytics — Oct 2015 – May 2019

- Developed **R Shiny** application to visualize high-dimensional data using **Hierarchical Voronoi Tessellations** & **Sammon Projection**
- Developed pipelines for automating the periodic dashboard back-end refresh process using Microsoft Azure framework
  - Automated data transfer from Azure **Cosmos DB** to **Azure Blob storage** and Azure **SQL** using **Azure Data Factory (ADF)** and **PowerShell**
- Developed and **deployed a real-time microservice** for auto-releasing incorrectly flagged financial transactions
  - Financial Messages (SWIFT) are picked from an IBM Message Queue and reviewed using **Cosine Similarity** matching algorithm
- Devised and implemented a **time-series Anomaly Detection** framework for detecting Bot traffic and the associated timeframe
  - Deployed the pipeline on **Azure ML** for automating model refresh and **Power BI** report generation
- Developed an **Excel VBA dashboard** that generates reports for comparing adoption in Paid Search Remarketing across different timeframes
- Optimized the performance of a Power BI dashboard by creating a **Star Schema Data Model** in SQL Server and **indexing** the required fields
- Created a **Power BI dashboard** for tracking usage of Excel VBA and Power BI dashboards to purge the dashboards not being consumed
  - Extracted usage information of the Power BI dashboards from **Blobs** generated through **Azure SQL Database Auditing Logs**
  - For Excel dashboards, created a **VBA** code that records the username and timestamp when the dashboard is opened and stores in a DB
- Automated the annotation process for a Customer Support Chat-Bot using **Topic Modeling** algorithm - **Latent Dirichlet Allocation (LDA)**
- Implemented **Named Entity Recognition (NER)** model using **Bidirectional LSTM** & **ELMo embeddings** and deployed as a service using **Flask**
- Implemented **ARIMA time-series forecasting** to predict the baseline Spend dollar value for Online Advertisers with MAPE less than 9%

Awards & Recognition

- Won the 3rd prize in Deloitte Case Challenge 2019 organized at CMU Heinz College — September 2019
- Dean's List for outstanding academic performance — June 2020
- Received multiple Mu Sigma Spot Awards for showcasing quality and timeliness in project execution — 2016, 2017, 2018