AJITH KUMAR ETHIRAJULU

44 Heath St., Buffalo, NY 14214 | 716-256-4115 | aethiraj@buffalo.edu
LinkedIn: www.linkedin.com/in/aje2907 | GitHub: github.com/aje2907 | Portfolio: aje2907.github.io

EDUCATION

University at Buffalo, The State University of New York   *Buffalo, NY*
Master of Science – Data Science   *Expected Graduation - January 2024*

EXPERIENCE

Fractal Analytics   *Bengaluru, KA, India*
Data Engineer   *July 2021 – April 2022*

- Developed two Python-based Azure Functions based on customer requirement to seamlessly receive and process streaming data from a logistics firm.
- Leveraged Azure Functions to efficiently capture and transform real-time data streams originating from Azure Event Hub. Implemented data storage in Azure Cosmos DB utilizing MongoDB API for optimized retrieval and analysis.

Infosys Limited   *Chennai, TN, India*
Senior Systems Engineer   *April 2018 – July 2021*

- Managed and maintained four trade store applications hosted on Linux, performed several DML operations on databases working with Oracle, ensuring data integrity, database backup, and cleanup.
- Leveraged various monitoring and scheduling tools including Autosys, Geneos, and ServiceNow to track system performance, schedule jobs, and ensure service level agreements (SLAs) were met by applying Agile methodology.
- Designed and implemented a procedural SQL script to automate multiple manual tasks, reducing time taken to resolve critical data compatibility issues by 25 minutes.

PROJECTS

Research Project - Predicting 3D Protein Structures for Huntington's Disease leveraging AlphaFold
Under the guidance of Professor Dr. Murali Ramanathan, University at Buffalo   *May 2023 – September 2023*

- Explored the innovative AlphaFold 3D structure prediction model, an artificial intelligence tool created by Google's DeepMind, and assessed its potential in the realm of healthcare, pharmaceutical science and drug development.
- Investigated the effects of mutation in the huntingtin protein as the root cause of Huntington's disease and compared the predicted and existing structures of two protein structures yielded by established techniques.
- Orchestrated a series of 40 experiments using AlphaFold to predict the structures of HTT protein. Analyzed the distorted protein regions having less than 50 pLDDT confidence scores, while employing the Multimer model to forecast potential protein aggregation tendencies.

Time Series Forecasting on Retail Store Sales in R   *February 2023 – May 2023*

- Led an empirical study forecasting retail sales applying time series methods like ETS, ARIMA, and Prophet in R, on a dataset of 3 million records from 54 stores, identifying Prophet as the best-performing model.
- Interpreted sales components, trends, and seasonality, conducting rigorous model training and testing. Attained best Root Mean Squared Error of 474.02 and Mean Absolute Error of 269.9.

Skin Melanoma Prediction   *September 2022 – December 2022*

- Leveraged Convolutional Neural Networks to build a binary classification model, trained and validated with a dataset containing about 33000 images to estimate the probability of a given image being skin cancer (melanoma).
- The highest accuracy achieved was 94.1% and 94.6% AUC with the help of EfficientNetB4 transfer learning model. Deployed a web app using Python Flask and HTML to predict skin melanoma on the uploaded images.

Credit Defaulter Analysis and Prediction   *September 2022 – December 2022*

- Integrated SQLite and Python to store, perform exploratory data analysis, and feature extraction on the vast data containing 120 features and 1.2 million records.
- Built binary classification machine learning model using Support Vector Machines, RandomForest, and Tensorflow (Keras). The highest accuracy obtained was 91% by the optimized Tensorflow model.

SKILLS

Programming:	Python, R, SQL, Machine Learning, Deep Learning, Reinforcement Learning
Databases:	MySQL, SQLite, Postgres SQL, Data Warehousing
Data Engineering:	Pyspark, ETL, HDFS
Cloud:	Azure Data Factory, Azure Databricks, Azure Data lakes
Visualization:	Tableau, Power BI, Seaborn, Matplotlib