ALBERT YU

📞 647-679-2872 ✉ changmiao.yu@mail.utoronto.ca 💼 LinkedIn ⓞ Github 🌐 Personal Website

Summary

Passionate data enthusiast with strong skills in programming languages, databases, and data transformation, eager to advance my career in the data engineering field. Highly skilled in modern data science techniques, machine learning, deep learning, and API development.

Education

Wilfrid Laurier University   **Waterloo, ON**
Master of Science in Management Analytics (MMA) Candidate   *2022 – 2023*

University of Toronto   **Toronto, ON**
Bachelor of Science in Mathematics and Statistics   *2017 - 2022*

Technical Skills

Languages: Python, SQL, R, C, HTML, CSS, JavaScript
Tools: Tableau, Power BI, Microsoft Azure, Git/Github
Certifications: Data Manipulation with Pandas, Data Analysis with SQL, Azure Fundamentals
Github: https://github.com/AlbertYu97

Experience

Ceridian Dayforce   **January 2023 - Present**
Data Scientist Intern   *Toronto, ON*

- Developed a highly scalable machine learning model with Azure function and verified results using Postman
- Designed and implemented time series analysis to forecast KPIs with Facebook Prophet, improving the accuracy by 30%
- Automated data cleaning, cross-validation and hyperparameter tuning to reduce model-building time by 80%
- Served as scrum master for the Workforce Management team, leading daily development scrum meetings and supporting team members with workload management

Stars Technologies Services (PokerStars)   **May 2022 – August 2022**
Data Analyst Intern   *Toronto, ON*

- Led the development of a machine learning model with Python to predict tournament duration used by millions of PokerStars customers
- Managed data extraction using SQL queries in a large-scale database environment to access, transform, and clean data for modelling; Selected and created suitable features for the model based on customer behaviour
- Conducted XGBoost and Random Forest to the model and utilized Shapley value for interpretability
- Performed hyperparameter tuning using grid search to improve model performance and achieved over 90% accuracy

Stars Technologies Services (PokerStars)   **May 2021 – August 2021**
Junior Poker Delivery Analyst Intern   *Toronto, ON*

- Executed test cases across different OS systems and verified actual results to ensure application logs are free of errors and database tables contain valid data as a member of the Quality Assurance team
- Performed systematic tests of the software and identified over 50 requirement gaps and functional bugs
- Communicated and cooperated with developers, designers, product management and other business owners in all stages of the bug detection and debugging lifecycle to maintain server stability
- Performed ad-hoc analysis on reproducing potential bugs on test/released servers and provided clarification and guidance to support code repair, reducing issue resolution time by 20%

Projects

Twitter Data Pipeline with Airflow and AWS S3   **January 2023**

- Developed an end-to-end data engineering pipeline using Apache Airflow, Python, and Tweepy to extract, transform, and store Twitter data from the Twitter API, focusing on user timelines and relevant tweet metrics.
- Gained hands-on experience with AWS services (EC2 and S3), data orchestration using Airflow, and data manipulation using pandas, while enhancing my proficiency in Python and data engineering best practices.

Deep Learning for Fraud Detection   **November 2022**

- Utilized deep learning techniques with Scikit-Learn, TensorFlow, and Keras to detect fraudulent transactions in a large, imbalanced dataset
- Improved sensitivity and overall accuracy by experimenting with various deep learning architectures, hyperparameter tuning, and data preprocessing strategies