DERRICK YU

Los Angeles, CA • +1 626-996-2118 • derrick.usc.data.ft@gmail.com • LinkedIn

PROFILE SUMMARY

Skilled data engineer with a **Computer Science** background and hands-on experience across the data engineering lifecycle. Proficient in **Python, SQL, Spark, Scala, R, C++, and Java**. Expertise in data management (**SQL/NoSQL**), **statistical modeling**, **machine learning**, **NLP**, **and deep learning**. Proficient in **MSBI** (SSIS, SSAS, SSRS, SSMS), **Databricks**, **Tableau, Power BI** (DAX, Power Query, Power Apps), **SAS JMP**, **Adobe Analytics**, and **cloud platforms (AWS)**. Effective communicator able to translate insights into actionable recommendations.

EDUCATION

Master of Science in Business Analytics (STEM) 2022.09 – 2024.05
University of Southern California, Marshall School of Business Los Angeles, CA
Honors Bachelor of Science in Computer Science with Artificial Intelligence (First Class) 2017.09 – 2021.05
University of Nottingham Nottingham, UK

WORK EXPERIENCE

Marketing Data Analyst | Amtrak | Washington, DC 2023.05 – 2024.05

- Analyzed customer journeys using Adobe Analytics data, identifying key **usage patterns** that helped the UI/UX team streamline booking pathways.
- Developed **ETL** processes and **data pipeline** using MSBI (SSIS, SSAS) and Python to extract, clean, and transform large-scale datasets (**20M+ rows**), significantly streamlining data preparation and **increasing efficiency by 90%**.
- Designed **Month-Over-Month Analysis** Power BI dashboards with DAX to analyze marketing campaigns, conduct cohort analysis, and **monitor marketing KPIs** (eg. CTR, Open Rate, CLV, Conversion Rate).
- Segmented customers and analyzed spending patterns (including price sensitivity analysis) to uncover insights into top customer profiles and revenue drivers.

PROFESSIONAL EXPERIENCE

Game Player Retention Analysis 2023.06

- Developed a player retention model to predict 14-day retention based on in-game and out-of-game factors.
- Applied mutual information and ANOVA to identify crucial metrics like tutorial completion and playtime, as well as social engagement and purchase behavior.
- Implemented Logistic Regression and then enhanced prediction accuracy with Random Forest, achieving a 9% increase in the F1 score.

Movie Recommendation System Implementation 2023.03

- Developed a movie recommendation system combining machine learning and deep learning techniques, enhancing recommendation accuracy by 11%.
- Analyzed 1M+ movie rating dataset, applied Bayesian Personalized Ranking Matrix Factorization (BPRMF) to deliver tailored movie recommendations aligned with historical user preferences.
- Optimized the model with LightGCN using PyTorch, expertly configuring the model architecture to achieve significant performance gains.

Adobe Analytics Challenge 2022.10

- Analyzed over 1 billion Hilton booking data points using SQL & PySpark, transforming raw data into structured sessions and customer profiles for in-depth analysis.
- Utilized Adobe Customer Journey Analytics to evaluate user growth, engagement, and conversion across channels (bounce rate, session duration, etc.).
- Compared user booking behavior on Hilton platforms vs. third-party OTAs, driving increase in direct bookings.

Face Recognition System for Class Attendance Checking 2021.07

- Developed a face recognition webapp to streamline class attendance tracking for students, tutors, and courses
- Architected a MySQL database to store information, efficiently defining relationships between students, tutors, and courses using primary and foreign keys
- Implemented various access levels using role-based access control (RBAC) to ensure appropriate permissions and data security for different user types (students, tutors, and administrators)