Casey Soh

515-587-1724 | kachun.soh@gmail.com | https://github.com/sohkachun | linkedin.com/in/casey-soh-442098b9/

EDUCATION

Boston College, MA

MS in Applied Analytics | CGPA: 3.76   *Dec 2024*
- Coursework: Data Analytics, Computer Vision, ML Algorithms, Linear Algebra and Vector Calculus, NLP, Product Management, Big Data Econometrics, AI / ML Software Tools and Platforms

Drake University, IA

Bachelor of Science in International Business and Management | CGPA: 3.56   *Dec 2021*

SKILLS

- **Programming Languages:** Python (Intermediate), SQL (Intermediate), R (Intermediate)
- **Tools:** PowerBI, Tableau, Excel, PostgreSQL, Git, Docker, Azure, MySQL, Big Query
- **Soft Skills:** Team collaboration, Problem-solving, Public speaking, Time Management
- **Libraries/Frameworks:** PyTorch, TensorFlow, Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn
- **Data Science Skills:** Machine Learning, Neural Networks, NLP, Transformers, Computer Vision
- **Language:** English, Chinese, Cantonese, Malay

Projects

Sales Analysis and Forecasting (2023)
- Designed and implemented data aggregation systems using SQL to consolidate sales data from five disparate sources into a centralized database, ensuring data accuracy and integrity.
- Processed and manipulated over 10,000 records using SQL, addressing missing values, deduplication, and standardization. Developed and delivered 10 interactive Power BI dashboards to visualize key performance indicators (KPIs), sales trends, and market segmentation, providing actionable insights to stakeholders.

End-to-End Machine Learning Project: Student Performance Prediction (2024)
- Conducted EDA to identify key patterns and enhance feature selection for model accuracy.
- Developed and optimized a predictive model using XGBoost, evaluated with RMSE and R^2 to track performance.
- Developed modular code components with MLOps integration for workflows and exception handling.
- Deployed a Flask-based web application for real-time predictions, utilizing Docker for containerization.

Consumer Segmentation and Lifetime Value Prediction (2024)
- Conducted exploratory data analysis (EDA) and data cleaning to prepare transactional data for analysis.
- Performed data visualization using Matplotlib and Seaborn.
- Engineered key features such as recency, monetary value, and frequency to enhance clustering and predictive modeling.
- Performed cluster analysis and applied KNN clustering to segment consumers based on purchasing behaviors.
- Developed a regression model to predict consumer lifetime value (CLV), enabling data-driven customer retention strategies.

Sale Data Analysis Project(2023)
- Used Python to perform EDA on the client's store data, conducted time-series analysis on sales (ARIMA, Naive, and Seasonal Naive), and forecasted sales for the next 2 years.
- Researched and created a definition for shopper loyalty to classify shoppers as loyal customers and non-loyal customers
- Displayed public speaking skills by describing market trends and findings as well as presenting business recommendations and their potential impacts to an open panel

Work Experience

DLL, IA

Contract Administrator   *Feb 2022- Jun 2022*
- Created a project planner for a team of 5 members to improve efficiency by allocating tasks based on individual experience and skillset, ensuring the timely completion of time-sensitive deliverables.
- Demonstrated leadership by taking on manager's responsibilities during her 2-week leave, collaborating effectively with teammates to ensure priorities were addressed and tasks were completed on time.