Mason Lee
Senior Data Engineer

M_lee86@u.pacific.edu ▪ www.linkedin.com/in/masonlee11 ▪ github.com/mason-lee19 ▪ San Jose, CA ▪ 408-838-8000

Summary

As a skilled and collaborative data engineer with professional experience in Python, C++ and data pipelines, I am committed to delivering high-quality solutions that drive data-based decision-making and improve business outcomes. With advanced knowledge of software tools and programming languages, I am confident in my ability to design and implement innovative data-driven solutions that add value to organizations.

Core Competencies

• Python	• Pandas	• C++	• SQL / NoSQL
• Agile Scrum	• Data Visualization	• Data Wrangling	• Data Analytics
• ETL Processes	• Tableau	• Version Control	• Communication

Experience

Western Digital, San Jose, CA

Senior Data Engineer **2024 – Present**

- Conducted exploratory data analysis (EDA) using Python libraries such as Pandas and Seaborn on data extracted and preprocessed from SQL and NoSQL databases, uncovering critical insights in increasing efficiency of the backend MFG process and failure analysis pipeline.
- Led the development of a key module for a data visualization platform, enabling visualization of customer and internal development failure data, resulting in improved failure analysis efficiency and faster issue resolution.
- Developed and optimized a machine learning algorithm to identify specific customer errors by analyzing and updating large datasets. Enhanced model accuracy by 20% through feature engineering and data preprocessing techniques, driving more reliable error detection.

Senior Software Engineer **2022 – 2024**

Software Engineer **2020 – 2022**

- Effectively managed and contributed to multiple version-controlled projects, leading to an improvement in code quality through proactive code reviews.
- Created and optimized Python scripts utilizing SQL to analyze large datasets, identifying and mitigating anomalies during manufacturing runs, reducing defects by 20%.
- Developed and maintained companywide C++ and Matlab test scripts for data collection and analysis, resulting in improved testing procedures and a 30% increase in data collection accuracy and efficiency.
- Took the lead in identifying root causes for encountered errors and performance issues, working collaboratively with cross-functional teams to implement effective solutions and achieve product goals on time.

Notable Projects

Market Headline Sentiment Analysis Dashboard, Personal **2024**

- Designed and implemented an ETL pipeline to extract daily news headlines, transform into relational data, and utilize a trained sentiment model for sentiment analysis, automating data processing and insight generation.
- Leveraged Google Cloud services to automate the ETL of headline data into a SQL server ensuring efficient and scalable data handling and storage.
- Created an interactive dashboard to visualize daily sentiment trends and track overall market sentiment, providing actionable insights and enhancing data-driven decision making.

Blackjack Strategy Optimization Using NEAT Algorithm, Personal **2023**

- Innovated a python-based AI for playing blackjack utilizing the NeuroEvolution of Augmenting Topologies (NEAT) algorithm to formulate strategies for playing blackjack.
- Conducted comprehensive data analysis including feature engineering, data visualization, and applied machine learning algorithms to build predictive models for outcomes.

Education, Certificates, and Professional Trainings

University of the Pacific, Stockton, CA

Bachelor of Science in Electrical Engineering