Jack Li

Data Analyst | Data Scientist
Jack.li.2000@outlook.com | www.linkedin.com/in/jack-li-2000 | Jack-Li-2000.github.io/JackLiPortfolio/

Skills

Data Science: Regression, Time Series, A/B Testing, Clustering, Feature Engineering, Statistics, Machine Learning.
Data Management: MSSQL, MySQL, BigQuery, Postgres, NoSQL, Excel.
Data Visualization: Tableau, PowerBI, Matplotlib, Seaborn.
Python: Pandas, Numpy, Scikit, Scipy, Tensorflow, Pytorch.
Big Data: GCP, Azure, AWS, Spark.
Collaboration: Confluence, Jira, Git

Experience

Data Developer Analyst, OneDataTree
(04/2025 – Present, Remote)

- Built and maintained end-to-end data ingestion and dashboard automation pipeline using **Python, SQL, and Tableau**, **slashing daily reporting time commitment** by **75%** and **cutting report-generation errors by 10%**.
- Conducted financial **predictive forecasting**, leveraging **Time Series** and **Regression** analytics **with Python**, boosting revenue **reporting accuracy by 15%** and supporting strategic budgeting decisions.
- Developed and maintained **Classification** and **Customer Lifetime Value** model in **Python** to **identify high risk churn** profiles, enabling **targeted retention campaigns**, reducing churn rate by **7%.**
- Developed and deployed an **Entity Flow Analysis** application tracking revenue, balance, and expenditure for **50+ customer entities.**

Reconciliation Data Analyst, OneDataTree
(01/2024 – 04/2025, Remote)

- Implemented cross-database validation report to maintain data accuracy and reliability between internal and client records with **SQL, Python, and Excel,** significantly **reducing discrepancies by 43%.**
- Designed and delivered weekly KPI dashboards, providing real-time visibility into revenue, churn, and product usage, enabling leadership to identify and act on emerging trends within 24 hours.

Data Scientist, Teck Resources
(08/2021 – 06/2022, Vancouver BC)

- Performed data cleaning, conveyed statistical inferences, and conducted exploratory analysis with **Python and SQL**, resulting in a **22% reduction in daily reporting time commitment**.
- Retrieved relevant features for classification using ElasticNet feature selection, **improving model precision and reducing overfitting,** leading to **16% increase in detection rate**.
- Predicted equipment lifecycle trends to inform replace vs. maintain decisions using regression models, **resulting in a 14% reduction** in maintenance costs and a **10% increase in equipment uptime**.

Projects

Wordnet Similarity Recognition Analysis
Conducted WordNet similarity recognition analysis to explore NLP data structure similarities with human semantic relationships, performing **hypothesis and significance testing in Python**. Leveraged data from the Semantic Priming Project to assess human reaction time and accuracy in response to stimuli. Analyzed correlation between WordNet path similarity and human word-pair recognition performance and concluded NLP concept structure mimics our semantic relationships.

Education

Honors B.Sc. in Statistics & Computational Cognitive Science
(University of Toronto, Toronto ON)
Practical knowledge in data regression models, predictive data techniques, classifiers, time series analysis, dimensionality reduction, data visualization, and extracting key insights from complex data sets. In-depth understanding of data structures and programming.