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

Python: Highly proficient with pandas, numpy, matplotlib, seaborn, scikit, scipy, etc.
Excel: Highly proficient with pivot tables, formulas, xlookup, conditional formatting, charts, data cleaning.
SQL: Experienced with query calls, views, joins, conditional statements, database management, etc.
Data Visualization: Familiar with telling stories using interactive dashboards in Tableau and PowerBI.
Statistical Methods: Highly proficient with supervised learning methods, time series, hypothesis testing, etc.
Cloud computing: Experienced with Google Cloud Platform, Microsoft Azure, Spark.

Work History

Data Developer & Analyst, OneDataTree

(01/2024 – Present, Toronto ON)

- 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%**.
- 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.
- Conducted financial predictive forecasting, classification, and ad hoc analytics **with Python**, boosting revenue **reporting accuracy by 15%** and supporting strategic budgeting decisions.

Data Science, Teck Resources

(08/2021 – 06/2022, Vancouver BC)
Machine Failure Classification

- 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.

Toronto AirBNB Popularity Insights Dashboard

Developed an **interactive Tableau dashboard** analyzing data from November 2022 to November 2023 for the Greater Toronto Area. Provided insights into popular neighborhoods, pricing trends, and accommodation types, enhancing decision-making and strategic planning.

Education

Honours B.Sc. in Statistics & Computational Cognitive Science

(University of Toronto, Toronto Ontario)
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.