Iris Luo

+1 (672) 866-8119 iris0614ubc@gmail.com iris-luo iris0614 iris website

PROFESSIONAL SUMMARY

• Data Scientist with dual Master's degrees in Data Science and Economics and diverse internship experience.
• Strong proficiency in Python, R, SQL, Power BI, and Tableau, with additional expertise in PyTorch and TensorFlow.

EDUCATION

Master of Data Science (GPA: 4.23/4.33) 09/2023 – 06/2024
University of British Columbia *Vancouver, BC*

Public Economics (Exchange Program) (GPA: 3.7/4.0) 07/2021 – 08/2021
University of California Berkeley *California, United States*

Master of Science in Economics (GPA: 86.7/100) 09/2020 – 06/2023
Huazhong University of Science and Technology *Wuhan, China*

Bachelor of Science in Economics (GPA: 88.63/100) 09/2016 – 06/2020
Central China Normal University *Wuhan, China*

WORK EXPERIENCE

Illuminex AI, Data Scientist - UBC-MDS Capstone Project 04/2024 – 06/2024|*Waterloo,Canada*
• Developed an object detection pipeline using Faster R-CNN, YOLOv8, and RetinaNet, achieving a 60% mAP for bird strike mitigation at Chengdu Shuangliu Airport, surpassing the previous best of 49.5% mAP.
• Pioneered synthetic data augmentation with airplane images, enhancing model realism and boosting predictive accuracy in complex airport environments.
• Managed large-scale datasets on AWS, optimizing models for real-time, high-speed processing, and supporting robust validations. Awarded "Best Talk" in the Master of Data Science program for effectively communicating these complex data insights.

Beijing Academy of Social Sciences, Data Analyst Intern 10/2019 – 01/2020|*Beijing, China*
• Optimized MySQL database queries, reducing page load times by 15% and improving data retrieval speed by 20%, resulting in a smoother user experience across the platform.
• Conducted in-depth data analysis on over 20,000 records, uncovering critical insights. These findings directly contributed to enhancing educational programs. This led to a 12% improvement in student engagement metrics.
• Developed and presented detailed reports and dashboards to senior leadership, highlighting KPIs and trends. These insights facilitated data-driven decisions that contributed to a 10% increase in operational efficiency and strategic alignment across departments.

Deloitte, Audit and Assurance Department Intern 07/2019 – 08/2019|*Shenzhen, China*
• Performed financial analysis of Guangshen Railway Co. and Huajun Management Company for their yearly reports published to the public and the government. Utilized Python to analyze over 5,000 data points, identifying key metrics and potential risks, increasing the accuracy of financial assessments by 20%.
• Conducted a comprehensive data audit and drafted 15+ financial reports using Power BI, creating clear and accessible data visualizations. These visualizations improved report clarity and reduced the time required for senior stakeholders, including directors, executive managers, and company representatives, to review and understand the financial data by 30%.

Zhongshan Securities, Investment Banking Management Headquarters Intern 04/2019 – 06/2019|*Shenzhen, China*
• Conducted valuation and financial modeling, analyzing growth rates, profitability, and competitive advantages, achieving a 90% accuracy rate in valuation premiums. Employed time series models to predict stock prices and used Power BI to present results, providing stakeholders with actionable insights.
• Streamlined due diligence processes by 20%, conducting in-depth financial analysis and capital structure evaluations, and preparing detailed reports and presentations on private equity opportunities, leading to faster and more informed decision-making.

Agricultural Bank of China, Risk Management Department Intern 07/2018 – 08/2018|*Qingyuan, China*
• Applied machine learning techniques to analyze 806 auto installment loans, 1,325 home mortgage loans, and 2,021 large consumer installment loans, enabling more accurate customer ratings and risk assessments for the bank.
• Collaborated with the risk management team to integrate machine learning insights into existing financial models, resulting in a 10% improvement in the accuracy of risk predictions and customer assessments.

PROJECTS

Pyxplor 02/2024 – 03/2024

• Created a comprehensive Python package to automate and streamline the Exploratory Data Analysis (EDA) process, using technologies such as PyPI, Pytest, Seaborn, and Pandas. The package is tailored for various data types, including numeric, categorical, binary, and time series.

• Enhanced data interpretation through a suite of specialized plotting functions, reducing complexity and time investment in initial data analysis. Developed using tools like Poetry and Cookiecutter, this package is designed to be an essential tool for data scientists and analysts at all levels.

HomeScope Dashboard 03/2024 – 04/2024

• Led the design and deployment of the HomeScope dashboard using Dash, Plotly, and Altair for interactive data visualization, incorporating filters, correlative bar plots, and real-time geographic mapping.

• Deployed the dashboard on Plotly using dash-bootstrap-components, pandas, and pyarrow.

• Boosted user engagement by 30% by expanding the database and creating an intuitive interface for seamless housing data exploration, enabling city-by-city comparisons and more informed real estate decisions.

Churn Insights 06/2024 – 07/2024

• Engineered and processed over 6000 records of numerical, categorical, and binary data through ETL operations in PostgreSQL, followed by rigorous data cleaning and transformation to ensure high-quality input for modeling.

• Designed and implemented predictive models using advanced machine learning techniques, including Logistic Regression, Random Forest, and SVM, to forecast customer churn, achieving a test accuracy of 0.829.

• Leveraged a robust data science toolkit including scikit-learn for model development, Pandas for data manipulation, and Seaborn for exploratory data analysis and visualization, with insights effectively communicated through Tableau.

SKILLS

Programming & Scripting: Python (Pandas, NumPy, scikit-learn, Seaborn, Matplotlib, Altair), R, SQL, HTML/CSS

Data & Machine Learning: PyTorch, TensorFlow, Hugging Face, Neural Networks, Transformers, Feature Engineering, Hyperparameter Tuning, Cross-Validation, A/B Testing, Unit Testing, Data Modeling, Reproducibility

Tools & Technologies: Plotly, Dash, Shiny, Tableau, Power BI, NetSuite, Git, GitHub, Docker, Apache Airflow, AWS, Cloud Computing, MongoDB, PostgreSQL, MySQL