Fanglu (Louie) Yang

416-829-9261 | louie.yang@mail.utoronto.ca | LinkedIn | Personal Website

Education

University of Toronto Sep 2023 – 2027
HBSc – Co-op Specialist in Statistics – Statistical Machine Learning and Data Science Stream cGPA: 3.70/4.00
Major in Computer Science, Minor in Economics

Projects

Car Evaluation Prediction ML Model Using XGBoost Classifier Feb 2025
Analyzed a dataset with **1000+** entries from the UCI Machine Learning Repository

- Achieved **98%** data completeness by handling missing values and applying label encoding via **Pandas and NumPy**.
- Trained an **XGBoost classifier** with **98.92% validation accuracy** using hyperparameter tuning via **GridSearchCV**.
- Visualized feature importance using **Matplotlib** via a bar chart highlighting top predictors for car acceptability.

Statistical Analysis on Research Project Investments by the MCU Mar 2024 – Apr 2024
Analyzed a 20-year (2004–2024) dataset from the Ministry of Colleges and Universities (STAA57 Group Project)

- Demonstrated data analysis skills through analyzing and visualizing investment data with plots, graphs, T-test and Var-test using **RStudio**.
- Conducted data cleaning and data wrangling on a dataset (from open.toronto.ca) with **5000+** entries.
- Applied statistical models including bootstrap, linear regression, cross-validation, and logistic regression with **81.7% accuracy**.

Planetze – A Carbon-Tracking Sustainability App (UI/UX) Dec 2024 – Jan 2025
Designed an Android app that tracks, reduces, and offsets carbon footprint while offering real-time data and personalized insights

- Implemented **SQL** for local storage and **Firebase** for **real-time database** operations such as registrations and authentication.
- Utilized **Java** for backend logistics, logistics, including data calculations and UI control, and performed **Mockito and JUnit tests**, increasing login reliability by **60%**.
- Developed the front-end in **Android Studio** using **XML** for an interactive **user experience**.

Experience

Best Overall Project in UofT SDSS (Students in Data Science and Statistics) Datathon Mar 2025
Completed a project of modeling, visualization and YouTube demo with business insights for transit **within 24 hours.**

- Achieved **over 99% accuracy** across all **9 datasets**, **150k+ entries** to predict delay of 3 different TTC transits using **Logistic Regression** and **XGBoost Regressor**.
- Designed an interactive **Power BI** dashboard featuring dynamic filters for station, route direction, date range, time and map-based location searches with scenario-based analysis of historical and predicted delay.

Research Assistant Apr 2021 – Nov 2021
Shanghai Institute of Technical Physics, Chinese Academy of Sciences

- Won **3rd Place** in the Shanghai Youth Science and Technology Innovation Competition **&** Shanghai Future Science Star by refining my **Piano Gesture Corrector project** using 3D printing and **Arduino**.
- 3D-modeled and printed **10+** prototypes with Autodesk & Gems, integrating an ultrasonic sensor and buzzer to provide real-time detection, improving gesture stability by **30%**.
- Designed a **flexible sensor** that recognizes real-time gestures via angle and curve detection, enhancing ergonomics for piano beginners by increasing detection accuracy by **55%**.

VP of Technology Aug 2024 – Present
BACSA (Biotechnology and Computer Science Association at University of Toronto)

- **Led** a team of **5** directors with weekly meetings and organized tech conferences and **python coding competitions**.
- Hosted and presented computer science workshops on **Python, Git**, etc.
- Managed club website development and updates using **HTML, CSS, and JavaScript**.

Skills

Languages: Python (Pandas, NumPy, Matplotlib, Tensorflow), C, Java, JavaScript, HTML/CSS, R, SQL
Tools: Git, VS Code, Office Suite, React, RStudio, PowerBI, Tableau, Microsoft Azure, Jupyter Notebook, OpenRefine, PostgreSQL, 3D-Modeling, Arduino, Data Analysis (Regression, Clustering, Cross-Validation, Hypothesis Testing, Bootstrap, K-Fold, Decision Tree, Random Forest, XGBoost, Sci-kit Learn, Keras).