TINGJI TAO

548-989-1782 tgg.tao@mail.utoronto.ca

EDUCATION

University of Toronto *May 2021 – Apr 2025 (Expected)*
Honours Bachelor of Science Toronto, ON

- **Program:** Mathematics and Statistics
- **Cumulative GPA:** 2.7/4.0
- **Relevant Courses:** Regression Analysis, Number Theory, Time Series Analysis, Analysis of Big Data, Statistical Methods for Machine Learning, Categorical Data Analysis

SKILLS

Technical Skills: **Programming Languages:** Python, R, SQL
Tools: Git *(GitHub)*, Jupyter Notebook, R Studio, Numpy, Torch

PROFESSIONAL EXPERIENCES

Huawei Technologies *May 2024 - Sept 2024*
Machine Learning Intern Toronto, Canada

- Participated the development and deployment of a customer service chatbot using GPT-based models, achieving a 70% resolution rate for common inquiries.
- Collaborated with the customer service team to continuously improve the chatbot's performance based on user feedback.
- Conducted extensive hyperparameter tuning and cross-validation to achieve optimal model performance, improved the accuracy of credit risk predictions by 25%

The BRIDGE Lab, University of Toronto *Sep 2022 - Dec 2022; May 2023 - Present*
Data Scientist in Management Scarborough, Canada

- Spearheaded the design and implementation of data research projects, primarily using Python in a Jupyter Notebook environment, resulting in a 30% increase in task efficiency.
- Conducted rigorous data cleaning, analyzing on 100k+ raw data, scrutinizing and rectifying 10k+ lines of potential errors.
- Leveraged advanced analytical techniques, including Time Series Analysis and Clustering, to design and implement predictive models. Used cross-validation to evaluate performance and select the best-fit model for further data predictions.

ACADEMIC PROJECTS

ICR - Identifying Age-Related Conditions *May 2023 - Jul 2023*
Personal Kaggle Project Scarborough, ON

- Preprocessed a dataset with 124 features, employing meticulous techniques like imputation, feature scaling, and generation of interaction terms using polynomial features to enhance data quality and readiness for model development.
- Architected an Ensemble Learning model, amalgamating base learners including Random Forest Classifier, Support Vector Classifier (SVC), and Logistic Regression.
- Evaluated and fine-tuned the model's hyperparameters, capitalizing on accuracy and log-loss metrics, achieving an 82% accuracy and a log-loss reduction of 0.26.

Comparative Performance Analysis of BERT-based Models for Binary Classification Tasks *Apr 2023 - May 2023*
Course Project, Supervised by Prof. Qiang Sun Scarborough, ON

- Evaluated various BERT-based models, including BERT, RoBERTa, XLM, and ALBERT, across diverse NLP tasks,empha-sizing their capabilities and potential in sentiment analysis, grammatical error detection, and paraphrase identification.
- Analyzed performance outcomes across SST, CoLA, and MRPC datasets, highlighting the paramountcy of model selection.

Registrar's Office Data Analysis *Oct 2022 - Dec 2022*
Course Project, Supervised by Prof. Andre Cire Scarborough, ON

- Processed and refined a robust dataset of over 280,000 student records; implemented data cleansing techniques that bolstered the quality assurance rate to 98%, fortifying the foundation for subsequent intricate analyses and model formulation.
- Proposed a risk model by evaluating the Information Value (IV) of over 30 features, culminating in the design of a nuanced scorecard to segment students into five distinct risk categories, enabling the Registrar's Office to anticipate with an 80%precision rate a student's likelihood to accept the offer and proceed to registration.
- Orchestrated a comprehensive Exploratory Data Analysis (EDA) on admitted students, pinpointing pivotal factors influencing offer acceptance and transition to registration.