Weihao SUN

sweihao@student.ubc.com | *+1 2368662612*

EDUCATIONAL BACKGROUND

University of British Columbia   *Vancouver, Canada*
Bachelor of Science in Statistics | Minor in Data Science   *August 2019 – May 2024*
Academic Honors: UBC Faculty of Science Dean's Honor (Sep. 2020), Outstanding International Student Award (Aug. 2019)

WORK & RESEARCH EXPERIENCE

Triple Eagle Logistics | *Data Engineer Intern*   *May 2022 – August 2022*
- Developed, tested, and maintained an internally used automatic bill calculation system.
- Updated and maintained the company database; Analyzed the database structure; designed and modified the structure to cater to different use cases.
- Designed the Microsoft Power Automate pipeline to Use Excel Online to write Office Script to automatically fill in charging rates, amounts, and dates into the daily form.
- Matched and extracted essential information such as tracking numbers and prices from large, noisy data provided by clients for usage by other departments, which improved the working efficiency by over 60 %.

Physics Informed Neural Network for Battery Modeling Project / *Researcher*   *November 2021 – April 2022*
Supervisor: Maricela Best McKay (Ph.D. at University of British Columbia)
- Solved partial differential equations to improve the current battery models with Physics Informed Neural Network (PINN)
- Conducted literature survey on PINN, SPM, and Lithium batteries; Collected, cleansed data, prepared data for data visualization, and applied exploratory analysis
- Developed the Physics Informal Neural Network from scratch, fitted the model to existing data, and analyzed the PINN model

ACADEMIC PROJECT

Forecasting the Demands for Urgent Care Service   *July 2023 – Present*
- Applied and evaluated advanced machine learning models, including Neural Networks, Time Series models, etc., to predict local emergency department demands.
- Utilized the Hospital Triage Dataset on Kaggle as a starting point, and merged local climate and holiday data as additional feature creations.
- Kept identifying opportunities for adaptation and refinement based on future discussion and feedback from stakeholders.

Classification on Smog Ratings of Cars in the Year 2022   *January 2023 – April 2023*
- Performed early-stage raw data collection via public channels, including National Statistics Bureau, FRED, etc.
- Applied Naïve Bayesian analysis, KNN, and SVM strategies for model construction and autonomous testing.
- Applied classical statistical methods, including confidence and prediction intervals and their calibration to evaluate models.
- Adopted transformers including SVD and PCA for model improvements.

Animal Species Classification   *January 2022 – April 2022*
- Implemented and evaluated various machine learning models (k-nearest neighbor, decision tree, support vector machine, and logistic regression) to classify animal specifies based on physical features.
- Developed a reproducible pipeline for the classification problem, collaborated within a team using GitHub and maintained the Python package, integrated the package using Docker.

Amusement Park Management System (APMS)   *September 2021 – December 2021*
- Developed an amusement park management system from scratch using Django framework.
- Implemented an SQLite database modelling functions on services, employment management, and amusement facilities
- Designed the user interface of the application with the Bootstrap 4 template

EXTRACURRICULARS

PSUHacks 2021 | Computer Science & Statistics Online Programming Competition
- Collected and cleaned COVID-19 data; Visualized and analyzed the trends for the pandemic; Predicted the trend of the pandemic by using statistical methods.

VanHacks 2020 | Computer Science Online Programming Competition
- Designed website tools for local non-profit organizations in JavaScript; Used Google Map API to integrate the park information in Vancouver.
- This system can benefit both the non-profit organizations and the residents to find the best parks that satisfy their desires nearby.

UBC Data Science Club
- Participated in club meetings and workshops; Shared project ideas with other members; Participated in and analyzed data science competitions and projects on Kaggle.

LANGUAGE & TECHNICAL SKILLS

Programming Languages: **Proficient:** Python, R; **Familiar:** SQL, C, C++, Java, Julia, Racket
Frameworks: Git, Docker, Django, PyTorch
Natural Languages: Mandarin (Native), English (Native-like), French (Beginner)