

JORDON CHEN

jordonochen@gmail.com | (778)-251-6396 | Vancouver, BC | github.com/jochenncn | linkedin.com/in/jordon-chen

EDUCATION

BSc, University of British Columbia (Vancouver) – MBIM & Data Science 2018 - 2023

- Dean's Honours List for all eligible semesters, (2018, 2019, 2020)
- *Key Courses*: Data Science Research in Microbiology (95%), Statistical Modelling for Data Science (86%), Visualization for Data Science (86%), Applied Machine Learning (83%)

KEY SKILLS

- **Dashboard Creation** - (Looker, Redash, Tableau), **ML Algorithms** - (random forest, XGBoost, regression), **Languages** - (MySQL, Python, R, Java), **Software** - (Git, Jira, GCP, Microsoft Suite), **Generative AI** - (ChatGPT, Gemini, DeepSeek)

RESEARCH & PROJECTS

Spatial region of habitat drives fish gut microbial diversity, composition, and functionality Jan 2023

- Wrote reproducible **QIIME 2** workflows to clean, filter, and export sequencing data for downstream analyses
- Performed alpha and beta diversity analyses using statistical tests such as **ANOVA**, **Wilcoxon**, & **PERMANOVA** in **R stats** package
- Conducted taxonomic and functional analyses utilizing R packages: *phyloseq*, *microbiome*, & *ggpicrust2*
- Created and managed **Github** repository to share and compile R scripts and figures

Predicting NBA Player Points per Game based on Player Characteristics Jan 2022

- Worked in team of 4 to analyze a dataset for an association between a NBA player's years of experience, height, and division on points per game
- Read, cleaned, and conducted **exploratory data analysis (EDA)** on dataset using R packages including *dplyr*, *tidyr*, and *GGally*
- Utilized functions in *glmnet* & *caret* to conduct cross-validation alongside **linear models** to assess performance

WORK & RELEVANT EXPERIENCE

Beatdapp – Uncovering Music Streaming & Royalty Fraud **Vancouver, BC**
Data Analyst Aug 2023 - Present

- Performed **EDA** on distributor datasets of 30,000,000+ data points, using **Python** and **SQL** to aggregate, group, and select for downstream analyses
- Designed internal and external **Google Looker** and **Redash** dashboards for music distributors, providing comprehensive overviews of data and aiding in outlier detection
- Improved detection of music streaming fraud using **Analytic Hierarchy Process (AHP)**, optimizing subsequent labeling
- Provided validation on **XGBoost** predictions in feedback loop to refine model accuracy, precision, and recall
- Supported construction of **Neo4j** and **NetworkX** network graphs that uncovered streaming botnets of over 100,000+ nodes/fraudulent payees
- Assisted in the creation of **string-matching** features that increased the identification of fraud upwards of **30%**

UBC Institute for the Oceans and Fisheries – Species Distribution Modelling **Vancouver, BC**
Research Intern Jan 2023 - August 2024

- Extracted data from marine species databases, cleaning and tidying data using R for downstream analyses
- Conducted **PCA** and **ENFA** tests on ecogeographical variables and species presence data
- Created and optimized **ML** models including *random forests*, *boosted regression trees*, & *ensembles* to predict future species distribution

UBC Department of Microbiology & Immunology – Foundations of Microbiology **Vancouver, BC**
Teaching Assistant Jan 2023 - May 2023

- Cooperated with teaching team to deliver course syllabus to class of over 200 students, fostering environment of open communication and interest in discovery
- Implemented weekly tutorials under the supervision of the professor, presenting and building on topics and ensuring learning goals were met
- Monitored discussion board, patiently answering questions and leading students towards conclusions