Jasper Chow

Email: chow.jasper@gmail.com

Portfolio: https://jchow-ust.github.io

Telephone: (+44) 07570 891144

About Me

Analytically minded, detail-oriented, and a quick learner, I am capable of working both independently and in a collaborative setting. I have 7 years' experience in Python programming, and I frequently use Pytorch for machine learning and numpy/pandas for data manipulation. I am a native English speaker who is also fluent in Cantonese and Mandarin Chinese, all of which I have used during my work as a paralegal. I hope my experience with handling text, graph, and tabular data and my rigorous analytical skills can provide valuable, actionable insights.

Work Experience

Michelle Tsoi, Solicitors [Paralegal]   June 2020 - March 2022

● Liaising and communicating with counsel, clients, and solicitors, drafting and preparing legal documents, and assisting principal solicitor Ms. Michelle Tsoi in administrative duties such as keeping records of telephone conversations, managing inventory of office supplies, and performing simple statistical analysis on office expenses

Hong Kong University of Science & Technology [Undergraduate Helper]   June 2021 - Dec 2021

● Assisted Dr. Desmond Tsoi in preparing course materials for newly opened course, COMP2211 (Intro to Artificial Intelligence). Prepared example Jupyter notebook projects for students to complete and review materials of basic math/programming concepts for undergraduate students
● Assisted Dr. Cindy Li in teaching secondary school students introductory Python programming over the summer. Required clear communication of coding concepts to students unfamiliar with programming

Education

BSc Data Science and Technology & Computer Science   Sept 2018 - June 2022
Hong Kong University of Science and Technology (HKUST)   **First Class Honours (GPA: 3.82)**
          HKUST Scholarship for Continuing Undergraduate Students (HKD 10,000/year)   Sept 2020 - June 2022

Relevant Coursework

Main: Machine Learning, Database Management Systems, Cloud Computing & Big Data Technologies, Big Data Mining and Management, Design and Analysis of Algorithms, Statistical Inference, Regression Analysis
Electives: Stochastic Modeling, Operating Systems, Deep Learning Architectures, Computer Organization, Java Programming, Principles of Cybersecurity
Core: Probability, Statistics, Linear Algebra, Multivariable Calculus, Object-oriented Programming and Data Structures, Software Engineering

Skills

Technologies and programming languages
● 7 years' experience, proficient in **Python, SQL (Oracle SQL, MySQL), NoSQL (MongoDB), R, C++, C, Java, Ruby**
● **Databricks and Apache Spark**, **AWS EC2** for instance deployment and application testing
● **Git (Version control) and command line**
Statistical concepts and machine learning
● Machine learning frameworks and libraries: **Pytorch, Keras, TensorFlow**, **numpy, pandas, sklearn, matplotlib, seaborn, scipy**
● Probability, statistics, supervised statistical machine learning, **linear regression**, **decision trees, random forest, SVM**, bagging/boosting
● Deep learning architectures e.g. **CNN, RNN, GNNs, wide and deep models**
● Experience with computer vision, natural language processing (NLP), time series tasks
● Unsupervised learning: **k-means clustering, principal component/factor analysis,** association rule mining, singular value decomposition
Other
● Data scraping and cleaning of tabular, graph, and text data
● Data visualisation: **Tableau, PowerBI**, matplotlib/seaborn, Microsoft Excel, ggplot (R)
● Stakeholder management and communication

Projects (please view portfolio for more details)

Big2-RL (Capstone)   Supervisor: Dr. Qifeng Chen

Developed a Pytorch reinforcement learning framework for Big Two, a 4-player card game. We investigated and trialled a variety of approaches, such as Deep Monte Carlo. Despite not being fully trained, our reinforcement learning agent was capable of holding its own against both human players and the previous state-of-the-art benchmark, achieving a statistically significant average EV of over +1 unit/1000 games against both, thus demonstrating the effectiveness of Monte Carlo methods for multi-agent imperfect information tasks. Training was conducted on a distributed cluster of GPUs.

Pet photo popularity prediction

Developed model for predicting the popularity scores (0-100) of pets based on their pictures on the online pet adoption website Petfinder.my. Using metadata such as whether pets' faces are shown may influence how popular pet profiles are. Ultimately, the final model used both metadata and image features and managed to achieve an RMSE of 17.09 on the Kaggle private leaderboard, for which I was awarded **a bronze medal**.

Rating prediction with wide and deep model

Developed a model to predict user's ratings (from 1-5) of items given a sparse initial user-item rating matrix. After experimenting with model architecture, different embedding methods and hyperparameter tuning, the final model achieved a validation Root Mean Square Error of 1.014 which outperformed all other teams in the class.

Fine-grained Yelp review sentiment analysis

Used a logistic regression and LSTM ensemble in Pytorch to predict the number of stars (from 1-5) from a user's text review, achieving a score of 55% macro F1 score on an imbalanced test dataset. The ensemble's performance ranked among the top 10% in the class.