

Oliver K. Pan

 github.com/oliverkpan  [linkedin.com/in/oliverpan](https://www.linkedin.com/in/oliverpan)  oliverkpan.com  oliver.pan@mail.utoronto.ca

SKILLS

- **Languages & Frameworks:** Python, C, HTML, CSS, JavaScript (ReactJS), R, Git, Flask
- **Data Science:** pandas, numpy, scikit-learn, Keras, Tensorflow, nltk, SpaCy, PyTorch, Tableau
- **Big Data:** SQL (Oracle, MS SQL, MySQL), AWS, Apache Airflow, Spark

EXPERIENCE

Data Science Intern, CIBC Data Studio *May 2021 - August 2021*

- Customizes and designs **NLP pipeline**, using **SpaCy** for **Named Entity Recognition** and **Markov chains** for sentence generation, delivering results to various stakeholders.
- Optimizes end-to-end performance of web application by **70%**, using **ReactJs** to deploy machine learning model and **Python** to scrape websites, creating internal functions for business use.

Data Management Operations Intern, Ontario Teachers' Pension Plan *September 2020 - December 2020*

- Developed **ETL** process, pulling from large databases by writing **100+ SQL queries** and constructing **Python scripts** on 32 critical datasets capturing **100%** of data breaches during 2 months of production.
- **Automated 10+ Python scripts**, notifying business stakeholders by email notifications, which resulted in saving two hours of daily work for a team of eight members.

Data Science Intern, Royal Canin *June 2020 - August 2020*

- Filtered 300,000 consumer purchases using **pandas** and utilized machine learning models such as **XGBoost** to analyze feature importance, which resulted in company investment of \$30,000 into marketing activities.
- **Generated \$100,000 investment** for future promotions by developing **A/B tests** on four years of historical and **predicting customer churn/retention** using survival analysis, working with cross range of teams.
- Deployed **Streamlit** web application from a hyper-parameter optimized **CatBoost model**, to assist veterinarians in **multi-label classification** of urolith stones, achieving an accuracy of **90% F1 Score**.

PROJECTS

Menda | Hack Western 7 (Awarded Best Health Hack)

- Developed mental health management application which utilized **OpenCV** for emotion detection and **nltk** for sentiment analysis, connecting ML Model with **Flask** and storing backend data with **Firestore**.

Iowatch | Personal Project

- Scraped pricing data with **BeautifulSoup**, predicting house prices with an accuracy of **<0.05 normalized RMSE**, using techniques such as **Logistic Regression** on 80+ variables to deploy model with **Streamlit**.

EDUCATION

University of Toronto

B.Sc in Statistics & B.B.A in Finance with a Minor in Computer Science *September 2018 - May 2023*

- **Undergraduate Teaching Assistant** - Foundations of Business Management
- **Relevant Coursework:** Linear Algebra, Quantitative Methods, Probability, Computer Science (Python & C)