

# Oliver K. Pan

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

## EDUCATION

---

### University of Toronto

B.Sc in Statistics & B.B.A in Finance with a Minor in Computer Science

*September 2018 - May 2023*

- **Relevant Coursework:** Linear Algebra, Probability, Quantitative Methods in Economics, Calculus of Several Variables, Computer Science (Python & C), Derivative Markets

## 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

*October 2020 - December 2020*

- Developed **ETL** process for Quantitative Research project, writing **100+ SQL queries** and building Python scripts on 32 critical datasets to detect breaches in data extraction, capturing 100% of critical incidents 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 built 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 data to analyze impact of promotions on subscription service, working with cross range of teams.
- Deployed web application using **Streamlit** from a **CatBoost model**, conducting cross-validation for an accuracy of **90% F1 Score** to assist veterinarians in the diagnosis of urolith stones for thousands of pets.

### Data Analyst Intern, CardSwap Inc.

*September 2019 - December 2019*

- Increased transaction processing speed by 25%, successfully identifying and analyzing the data of over 500+ customers, which resulted in business transactions of over \$4000 daily.

## 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, resulting in personalized media recommendations for mental wellbeing.

## SKILLS

---

- **Languages & Frameworks:** Python, HTML, CSS, JavaScript (ReactJS), C, Git, Flask
- **Data Science:** pandas, numpy, seaborn, matplotlib, plotly, Streamlit, Jupyter Notebook, Tableau
- **Big Data:** SQL (Oracle, MS SQL, MySQL), AWS, Apache Airflow
- **Machine Learning:** scikit-learn, Keras, Tensorflow, OpenCV, nlTK, SpaCy, FBProphet, PyTorch