Oliver K. Pan

 oliverkpan   in oliverpan   🌐 oliverkpan.com   👤 oliver.pan@mail.utoronto.ca

SKILLS

Programming Languages: Python, SQL (Oracle, MS SQL), HTML, CSS, Git
Data Science: pandas, numpy, seaborn, matplotlib, plotly, Streamlit, Jupyter Notebook, FBProphet, Tableau
Machine Learning: scikit-learn, Keras, Tensorflow, OpenCV, nltk

WORK EXPERIENCE

Data Management Operations Intern, Ontario Teachers' Pension Plan   09/20 - 12/20
- Automated 5+ Python scripts, saving ~2 hours of daily work per week, notifying business stakeholders by email notifications, reducing company risk by 100%.
- Developed ETL process, writing 100+ SQL queries and building data quality checks using Datanomics.

Data Science Intern, Royal Canin/Mars Inc.   06/20 - 08/20
- Filtered 300,000 consumer purchases by building machine learning models such as XGBoost to provide actionable insights, which resulted in company investment of $30,000 into marketing activities.
- Developed A/B tests on four years of historical data to analyze impact of promotions on subscription service, successfully working with cross range of teams to generate $100,000 investment for future promotions.
- Assisted veterinarians in diagnosis of urolith stones, analyzing 60,000 submissions to deploy web application using Streamlit from a CatBoost model; conducted cross-validation for an accuracy of 90% F1 Score.

Data Analyst Intern, CardSwap Inc.   09/19 - 12/19
- 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.

EDUCATION

Bachelor of Science (BSc) - Statistics, University of Toronto   2018 - 2023
Bachelor of Business Administration (BBA) - Business Management, University of Toronto   2018 - 2023
- **Undergraduate Teaching Assistant**, Foundations of Business Management (MGTA05)

PROJECTS

Iowatch
- Scraped pricing data with BeautifulSoup, combining data sources to predict house prices with an accuracy of <0.05 normalized RMSE, using various regression techniques such as Logistic Regression and Booster Trees on 80+ variables, to deploy a model with Streamlit.

Menda
- Developed mental health management application which utilized OpenCV for emotion detection and nltk for sentiment analysis, resulting in personalized media recommendations for mental wellbeing.
- Presented the project at a Hackathon (Hack Western 7) and was awarded Best Health Hack, placing 1st out of 180+ participants.