

XUMING (MARCO) WANG

Jackson Ave 28-10, Jackson Park, Long Island City, 11101

Poforlio: <http://marcoxm.github.io>

Email : xwang423@fordham.edu

Mobile : +1-646-417-1681

EDUCATION

- **Fordham University** New York City, NY
Master of Science in Business Analytic; GPA: 3.80 Aug. 2018 – Dec. 2019(expected)
- **Guangdong University of Foreign Studies** Guangzhou, China
BS of Logistics; Aug. 2011 – July. 2015
- **Relevant Course:** Deep Learning, Data Mining, Machine Learning, Natural Language Processing and Big data Programming and Algorithms and Data Structure

EXPERIENCE

- **Fordham University** New York City, NY
Teaching & Research Assistant June 2019 - Present
 - **Research Assistant - Object Detection:** Research on Object detection and localization. Localized the damaged pixels in MRI picture, improving the diagnose result. Transform image data into pytorch gennerator, fine-tune the pretrain object detection(SSD and Faster RCNN)
 - **Teaching Assistant - Deep Learning:** Drafted assignments and final project. Collected and test image, text, audio and video data-set and generate baseline model performance for assignments with TF 2.0 and Keras. Built pipeline to automatically generated test data by Self-Attn GAN and VAE. Conducted Code-Review for students
- **Wright Star LLC** New York City, NY
Financial Data Analyst Oct 2019 - Present
 - **Portfolio Management:** Automated the data updating process for ETF and stock model using Python. Improved the overall efficiency by enabling the models to update the daily closing price of stocks automatically
 - **Quantitative Modeling:** Built a new financial model using Excel and Python, which classifies stocks based on multiple financial indicators, and provide insights to traders in finding valuable long-term stock investments
- **DHL-Sino Trans International Air Courier LTD** Dongguan, Guangdong, China
Operation Analyst Jan 2017 - Aug 2018
 - **Analytics Report:** Automated the data updating process for ETF by using Python. Improved the overall customer satisfaction score 15% by using root cause analysis, and delivered actionable solution plan.
 - **Workflows:** Dataduct an open source workflow framework to create and manage data pipelines leveraging reusables patterns to expedite courier productivity and customer satisfaction

PROJECTS

- **Wayfair Customer Return Purchase Prediction:** NYC, NY Apr. 2019 – May. 2019
 - Multitask prediction on customer conversion and convert revenue in Top 5 score, based on history data with 50,501 records and 181 features
 - Performed missing value imputation, feature engineering and feature selection through exploratory analysis
 - Reached 80% accuracy in classification prediction by stacking Catboost, Lightgbm and Xgboost models results
- **Patient Readmitted Prediction:** NYC, NY Apr. 2019 – May. 2019
 - Pre-processed data with 100000 records of diabetes patients by conducting EDA, data cleaning and implementing appropriate feature encoding with Pandas, Numpy and Scikit-learn libraries
 - Deployed regression models and classification models to fill missing values of important attributes
 - Written pipeline to conduct features selection (filter/ wrapper) and ensemble methods for classifiers including soft-margin SVM, Random Forest, Perceptron, Logistics Regression and Naive Bayes

PROGRAMMING SKILLS

- **Languages:** : Python, SQL, Scala, C++
- **Technologies:** : AWS((EC2, S3, EMR, sagemaker), Spark (Spark SQL, Spark ML), Git
- **Udacity Nano Degree:** Applied Neural Network in image classification and generation, time-series prediction, and model deployment
- **Kaggle NLP competition:** Jigsaw Unintended Bias in Toxicity Classification top 20% (534/2749)