

XUMING (MARCO) WANG

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

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

Email : xmwmarco@gmail.com

Mobile : 646-417-1681

EDUCATION

- **Fordham University** New York City, NY
Master of Science in Business Analytic; GPA: 3.80 Aug. 2018 – Dec. 2019
- **Guangdong University of Foreign Studies** Guangzhou, China
*Bachelors Degree in Logistics(**Supply Chain Management**);* Aug. 2011 – July. 2015
- **Relevant Course:** Deep Learning(**Pytorch, Tensorflow**), Data Mining(**Scikit-Learn**), Machine Learning, Bigdata Programming and Big data Analytic(**Hadoop, Hive, Pig, Spark and NoSQL**)

EXPERIENCE

- **Fordham University** New York City, NY
Teaching & Research Assistant June 2019 - Present
 - **Teaching Assistant - Deep Learning:** Directed the gathering of massive data, assessing data validity and synthesizing data into large analytical datasets to support assignments and the final project. Provided technical leadership and mentoring students for best practices on deep learning
 - **Research Assistant - Object Detection in NYU Langone Medical Center Project:** Achieved **20%** reduction in RMSE, by building and training machine learning algorithms for object detection and localization. Localized the damaged pixels in MRI pictures, improving the MRI image quality and diagnose result. Deployed the **End to End solution** independently, fine-tuned the detection models, SSD and Denosing AE model in school server, Implemented distributed computing and Flask shortens model inference time 70%
- **Wright Star LLC** New York City, NY
Financial Data Analyst Oct 2019 - Present
 - **Quantitative Modeling:** Built risk-based financial models and advanced proficiency in Excel, which classifies stocks based on multiple financial indicators, and provide insights to traders in finding valuable long-term stock investments
 - **Portfolio Management:** By adding new features in Stock Price Update Script, achieved auto-checked missing records and updated. Improved the overall efficiency of 30% while maintaining data integrity
- **DHL-Sino Trans International Air Courier LTD** Dongguan, Guangdong, China
Operation Analyst Jan 2017 - Aug 2018
 - **Analytics Report:** Updated data by using Python and Excel. Identified root causes and developed potential solutions, Improved the overall customer satisfaction score by 15% by using root cause analysis, and delivered an actionable solution plan.
 - **Project Management:** Communicated with sales, other technical teams, and senior management to collect requirements, describe real problems and solving strategy in operation efficiency improvement project applied insights to the development, execution, and improvement of action plans.

PROJECTS

- **Million Song Dataset Analysis:** New York City, NY May. 2019 – Jul. 2019
 - Developed the ETL Pipeline copying datasets from S3 buckets, data processing using Spark and writing to S3 buckets using efficient partitioning and parquet formatting on EMR Hadoop Cluster.
 - Fast-tracking the data lake build-out using (serverless) AWS Lambda and cataloging tables with AWS Glue Crawler.
 - Transforming data from various sources into a star schema optimized for visualization
- **Patient Readmitted Prediction:** New York City, NY Oct. 2018 – Dec. 2018
 - Reached 0.83 F1 score. Built Pre-processed data pipeline with over 100000 records of diabetes patients, achieving EDA, cleaning and implementing appropriate feature encoding.
 - Integrated EDA, preprocessing, data cleaning and modeling parts into presentation, illustrated it in the final section

PROGRAMMING SKILLS

- **Programming Languages:** : Python, SQL, Scala, C++
- **Technologies:** AWS(Redshift, EMR, Kinesis, RDS and sagemaker), Spark(Spark SQL, Spark ML), SPSS
- **Udacity Nano Degree:** Deep Learning & Data Engineering (data modeling, data analysis and model deployment)
- **Kaggle ASHRAE - Great Energy Predictor III:** Energy Consumption Prediction top 13% (455/3614)