QIANYU CHENG

474 Warren St, Jersey City, NJ • 217-607-4217 • qc510@nyu.edu • Visa Status: F-1

EDUCATION

New York University **August 2016 - May 2018**
Master of Science in Data Science ***GPA: 3.76/4.00***
University of Illinois at Urbana-Champaign **August 2012 - May 2016**
Bachelor of Science in Actuarial Science and Bachelor of Science in Statistics **Statistic Major GPA: 3.74/4.00**
Related Courses: Deep Learning, Machine Learning, Big Data, Advanced Python, Natural Language Processing, Inference and Representation, Time Series

WORK EXPERIENCE

Northwestern Mutual **July 2018 – Now**
Associate Data Scientist **New York City, NY**
- Participating in Predictive Planning Team and studying the attrition rate of Financial Representatives
- Accessing database using SQL and built survival analysis and cox proportional-hazard rate model in R
- Building a collaborative filtering recommender based on Financial Representatives' sales history

NBCUniversal **May 2017 – August 2017**
Data Scientist Intern **New York City, NY**
- Collected and parsed movie Metadata from IMDb and Google API and used fuzzy logic to match movie titles
- Coded in PySpark to aggregate impression data stored in AWS and presented data visualization
- Applied K-means cluster model to identify light, medium, heavy movie viewer and PCA to visualize clusters
- Built a collaborative filtering recommender to find similar movie using Alternating Least Square model

NBCUniversal **September 2016 – December 2016**
Data Scientist Intern **New York City, NY**
- Contributed to the creation and maintenance of time series based forecasting models in SAS
- Contributed to the development and enhancement of forecasting methodologies
- Managed multiple priorities across a mix of ad-hoc and operational projects using PySpark
- Translated complex problems and solutions to all levels of the organizations

RESEARCH & PROJECT

Comparing Methods on Stanford Nature Language Inference **Fall 2017**
GitHub Repository: https://github.com/QianyuCheng/ds1011_final_project **New York City, NY**
- Implements models in PyTorch and ran the programs on GPU on HPU clusters
- Programmed and compared the performances of baseline model Continuous Bag of Words with Multi-Layer Perceptron, Attention-based Gated Recurrent Unit, as well as Decomposable Attention model.
- Tuned hyper-parameter and word embedding: GloVe, fastText, character n-gram to achieve better accuracy.

Improved Wasserstein Generative Adversarial Network With Group Normalization **Spring 2018**
GitHub Repository: https://github.com/QianyuCheng/wgan_with_group_normalization **New York City, NY**
- Translated Wasserstein Generative Adversarial Network Model from TensorFlow to PyTorch
- Implemented Group Normalization to improve WGAN performance and add ResNet structure
- Visualized generated figures and tested the Inception Score on CIFAR-10 dataset

Stochastic Process and Nested Simulation on Annuity Deferred **Spring 2016**
Research Assistant **Champaign, IL**
- Analyzed pricing, reserve and capital calculating process of equity-indexed annuities product in insurance industry
- Derived a generic iterative formula from both the reserve and capital processes to build a nested-stochastic model
- Simulated hedging procedure into the capital calculation and increased the efficiency by resource allocation
- Translated pricing algorithm into Matlab code and analyzed survey results from Society of Actuary

PROGRAMMING LANGUAGE

Python, R studio, PySpark, C++, Hadoop, Matlab