CONG (MARK) MU

108 W 39th St Apt 4, Baltimore, MD 21210 | 410-710-7091 | cmu2@jhu.edu

EDUCATION

Johns Hopkins University (Baltimore, MD) 08/2017 –

Master of Science in Engineering | **Major: Applied Mathematics and Statistics** | **GPA: 3.9/4.0**

Sun Yat-Sen University (Guangzhou, China) 08/2013 – 06/2017

Bachelor of Science | **Major: Statistics** | **Minor: Finance** | **GPA: 3.8/4.0**

RESEARCH EXPERIENCE

Network Analysis | Johns Hopkins University

- Built network models to analyze large social and economic networks; estimated and simulated network formation models using high performance computing
- Key words: **Hierarchical Exponential-Family Random Graph Models, Variational generalized EM algorithms, Bootstrap Standard Error, Parallel Computation**

Image Similarity | Johns Hopkins University

- Constructed framework to estimate lap time given dash cam video of auto racing; identified unusual frames in dash cam video; annotated video with relevant information
- Key words: **Structural Similarity Index, Earth Mover's Distance, Oriented FAST and Rotated BRIEF, Image Hashing, Deep Neural Networks**

Subspace Clustering | Johns Hopkins University

- Examined different subspace clustering methods and created algorithms to choose optimal methods based on given data
- Key words: **Low-Rank Subspace Clustering, Sparse Subspace Clustering, Spectral Curvature Clustering, Locally Linear Manifold Clustering, Random Dot Product Graph**

Therapy Functional Measures | Johns Hopkins University & Johns Hopkins Hospital

- Identified patterns in patient functional trajectories; measured causal effect of different physical therapy dosage regimes on patient functional status; constructed features and built model to predict AMPAC score to optimize physical therapy in the hospital
- Key words: **Linear Mixed-Effect Model, ARIMA**

Text Mining | Johns Hopkins University

- Collaborated with different teams to mine the large-scale text data, speculated gender based on names and explored characteristic distribution across gender
- Key words: **Natural Language Processing, Parallel Computation**

Information Extraction | Johns Hopkins University

- Extracted information from large-scale data sets and reconstructed data; crawled online data
- Key words: **Regular Expression, Crawler, Data Wrangling**

WORK EXPERIENCE

Analyst Intern | **GF Fund Management** (Guangzhou, China) 11/2016 – 04/2017

- Selected features to build market emotional indicators and developed model to predict market, achieved 92% accuracy (**XGBoost, Random Forest, Logistic Regression, Lasso**)
- Mined key business data and constructed data reporting system; analyzed and visualized product and user data to provide decision support (**R Markdown, R Shiny, R ggplot2**)

Data Science Intern | **Research Center of Statistical Science** (Guangzhou, China) 02/2016 – 10/2016

- Classified users to optimize delivery of advertisements and constructed program recommendation system; predicted whether user will be secondary loans to explore potential customers and evaluate risk in advance (**Collaborative Filtering, Spectral Clustering**)
- Presented in 9th China-R Conference and Regional Data Science Conference on how to use R to interact and share ideas by using Shiny in R to make an interactive interface rapidly (**R Shiny**)

TEACHING EXPERIENCE

Teaching Assistant | Johns Hopkins University

- Data Mining (Spring 2018)

RELEVANT COURSES

Johns Hopkins University 08/2017 –

- Probability Theory
- Statistical Theory
- High-Dimensional Approximation, Probability and Statistical Learning
- Statistical Pattern Recognition Theory & Methods
- Statistical Machine Learning
- Data Mining
- Mathematical Image Analysis
- Applied Statistics and Data Analysis

Sun Yat-Sen University 08/2013 – 06/2017

- Real Analysis
- Mathematical Statistics
- Applied Regression Analysis
- Time Series Analysis
- Nonparametric Statistics
- Modern Multivariate Statistics: Data Mining
- Complex Data Analysis (Case Study)
- Mathematical Experiments and Mathematical Software

SKILLS

R (dplyr, ggplot2, shiny, Rmpi), Python (NumPy, Pandas, scikit-learn), SQL, Matlab, C/C++

TensorFlow, OpenCV, Data Mining, Data Visualization, Machine Learning, Model Development