CONG (MARK) MU

108 W 39th St Apt 4, Baltimore, MD 21210 | 410-710-7091 | cmu2@jhu.edu | https://congm.github.io

EDUCATION

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

Master of Science in Engineering | **Major: Applied Mathematics and Statistics**

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

Bachelor of Science | **Major: Statistics** | **Minor: Finance**

Dissertation: Rank-based Integrated Feature Selection Methods | Advisor: Caixia Li

RESEARCH EXPERIENCE

Statistical Models for Large Networks | Johns Hopkins University

- Built network models that could be scaled to analyze large networks; estimated and simulated network formation models using high performance computing; develop R package with research objectives such as identifying the community structure, influence of nodes on core-periphery
- Key words: **Hierarchical Exponential-Family Random Graph Models, (Generalized) Random Dot Product Graph, SBM with Covariates, Variational generalized EM algorithms, Minorize-Maximization, Parallel Computation**

Automatic Tools for Dash Cam Video | Johns Hopkins University

- Developed automatic tools for analyzing and annotating video stream with relevant information such as timing, speed, traffic, accidents, objects and etc
- Key words: **Structural Similarity Index, Earth Mover's Distance, Oriented FAST and Rotated BRIEF, Image Hashing, Deep Neural Networks**

Constructing Affinity Matrix for Spectral Clustering | Johns Hopkins University

- Built a framework on constructing affinity matrix for spectral clustering to apply it to a more general problem; developed corresponding theoretical justification on different setting
- Key words: **Low-Rank Subspace Clustering, Sparse Subspace Clustering, Spectral Curvature Clustering, Profile Likelihood, Model-based 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, Causal Inference**

Text Mining and Information Extraction | Johns Hopkins University

- Collaborated with different teams to mine the large-scale text data, speculated gender based on names; extracted information from large-scale data sets; crawled online data
- Key words: **Natural Language Processing, Parallel Computation, Regular Expression, Crawler**

PROFESSIONAL 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 (statnet, mclust, dplyr, ggplot2, shiny, Rmpi), Python (NumPy, Pandas, scikit-learn), SQL, Matlab, C/C++
TensorFlow, OpenCV, Data Mining, Data Visualization, Machine Learning, Model Development