Peiyao Wang

(984) 215-8186 | https://peiyaow.github.io/ | 1351 Mason Farm Rd
peiyao@email.unc.edu | https://github.com/peiyaow | Chapel Hill, NC 27514

EDUCATION

University of North Carolina, Chapel Hill, NC — Aug. 2015 - July 2020
Master of Science, Statistics and Operations Research
PhD, Statistics

Fudan University, Shanghai, China — Sept. 2011 - May 2015
Bachelor of Science, Pure and Applied Mathematics

COURSEWORK

Statistical Inference, Machine Learning, Deep Learning, Algorithm Analysis, Big Data & NoSQL, Linear Programming, Convex Optimization, Time Series Analysis, Survival Analysis

WORK EXPERIENCE

Amazon.com, Inc., Seattle, WA — May 2019 - Aug. 2019
Applied Scientist Intern, Customer Service Machine Learning Team
Build **intent models** to generate personalized order suggestions from customer click-stream and profile features. Predict **scores** and analyze Top-K mean average accuracy **ranking** results. Solution incorporates a pairwise ranking loss into a binary classifier.

PROJECTS

High Dimensional Heterogeneous Factor Regression Model — Nov. 2018 - Present
Develop methods and theory utilizing **approximate factor models** for a regression task with heterogeneous subgroups. Identify groupwise factors by **PCA**. Perform high dimensional inference to achieve optimal rates on model estimators.

Locally Weighted Penalized Regression with Applications to Alzheimer's Disease — Aug. 2017 - Mar. 2018
Train a locally weighted penalized regression model to predict ADAS-cog clinical scores from ˜200 brain image features. Local weights extracted from **random forests** and **ordinal logistic regression**. Training data include ˜900 subjects in the ADNI study.

Large-Scale Classification on Forest Cover Type via L_2-Regularized Support Vector Machine — Feb. 2017 - May 2017
Classify forest cover type in northern Colorado with 50000+ training samples and 54 cartographic variables. Implement random **coordinate descent** and proximal stochastic dual coordinate ascent algorithms to solve for the **large-scale** quadratic programming problem.

Study on Association of HIV Infection Rate with Contraception Methods — Mar. 2016 - May 2016
Assess the effect of three different hormonal contraceptives on the risk of HIV infection. Solution includes Kaplan-Meier **survival** curves comparison and **log-rank tests**. Identify marital status as a risk factor by fitting a **Cox proportional hazard** model.

PUBLICATIONS

Wang, P., Liu, Y., & Shen, D. (2018). Flexible Locally Weighted Penalized Regression With Applications on Prediction of Alzheimer's Disease Neuroimaging Initiative's Clinical Scores. *IEEE transactions on medical imaging*, 38(6), 1398-1408.
Li, J., **Wang, P.**, Li, Q., Zhang, K., & Liu, Y. Nonparametric Prediction Distribution for Regression with Heterogeneous Data. *Submitted.*

Wang, P., Xiao, W., Wang, S., Bhardwaj, V., & Sun, J. Combining Classification and Ranking for Personalized Customer Order Suggestions. *In Preparation.*
Wang, P., Li, Q., Liu, Y., & Shen, D. High Dimensional Heterogeneous Factor Regression Model. *In Preparation.*

TECHNICAL SKILLS

Data Querying & Wrangling: SQL, Hadoop, Spark, pandas, dplyr
Data Visualization: matplotlib, seaborn, Jupyter Notebook, Plot.ly, ggplot2
Machine/Deep Learning: Scikit-learn, lightgbm, xgboost, TensorFlow, Keras
Miscellaneous: Linux, macOS, git