YING (JULIA) JI

(615)-397-0143 yingji93@gmail.com Nashville, TN
www.linkedin.com/in/ying-ji-3a403483 yingji15.github.io

EDUCATION

Vanderbilt University, Nashville, TN Aug 2015 - May 2021 (Expected)
Ph.D. in Human Genetics GPA: 3.77
M.S. in Biostatistics GPA: 3.84
Coursework: Advanced Statistical Computing; Survival Analysis; Modern Regression Analysis; Contemporary Statistical Inference; Fundamentals of Probability; High Performance Computing

Tsinghua University, Beijing, China Sep 2011 - Jul 2015
B.S. in Biological Science GPA: 3.7 (90.5/100)

SKILLS

Languages: R, Python, SQL, C
Tools: Unix/Linux, Git, Rstudio, Jupyter, Anaconda, Pycharm
Data visualization: ggplot, matplotlib, Tableau
Statistical modeling: linear regression, logistic regression, SVMs, classification, decision trees

EXPERIENCE

Graduate Research Assistant May 2017 - Present
Vanderbilt University, Nashville, TN

Supervised classification of Schizophrenia risk genes using expression data

- Collected and constructed features from various datasets of 50-3000 dimensions via different sources to describe gene expression in diverse biological conditions
- Built machine learning models including logistic regression, SVM, random forest to classify risk gene from constructed features using skicit-learn (Python), randomForest, e1071, Caret (R)
- Analyzed important features in prediction using importance score and correlation matrix heatmaps to gain insights into disease-related gene expression changes
- Identified 10 times more candidate risk genes compared to well-supported known risk genes to inform more targeted downstream functional analyses

Predict patient-level genetic risk for breast cancer in East Asian populations

- Processed and filtered large quantities (millions of rows) of individual-level genotype data as features
- Improved variable selection and effect size estimation in linear model of individual-level breast cancer risk in East Asians by borrowing information from a higher-powered European study
- Significantly increased prediction auROC from 0.56 to 0.62 in East Asian population (P=3.6e-5), reduced the gap in risk prediction accuracy between European and East Asian populations

SELECTED PUBLICATION/PRESENTATION

Presentation: platform presentation at 2018 American Society of Human Genetics Annual Meeting (abstract scored in top 8% among more than 3000 participants)

Publication: Wang, Q., Chen, R., Cheng, F., Wei, Q., Ji, Y., et al.,(2019). A Bayesian framework that integrates multi-omics data and gene networks predicts risk genes from schizophrenia GWAS data. Nature neuroscience

OTHERS

Award: Graduate Leadership Development Institute (GLDI) travel award May 2019
Activity: Secretary of Human Genetics Graduate Student Association Aug 2017 - Jul 2018
Teaching: Mentored graduate rotation students on analysis using UNIX and R Jan - Feb 2019