Ying (Julia) Ji

Linkedin: ying-ji | Github: yingji15| Blog: yingji
Email: yingji93@gmail.com| Cell: 615-397-0143| Nashville,TN

EDUCATION

- **Vanderbilt University** — Nashville, TN
  Ph.D. in Human Genetics, M.S. in Biostatistics; GPA: 3.87 — *2015/08 – present*
  - **Relevant coursework::** Advanced Statistical Computing; Applied Survival Analysis; Modern Regression Analysis; Contemporary; Statistical Inference; Fundamentals of Probability; High performance computing

- **Tsinghua University** — Beijing, China
  B.S. in Biological Science; GPA: 3.7 (90.5/100) — *2011/09 – 2015/07*

EXPERIENCE

- **Vanderbilt University** — Nashville, TN
  Graduate Research Assistant — *May 2016 - Present*
  - : **Risk gene prediction using machine learning methods**
    * **Data Collection**: collected high-dimensional, both dense and sparse gene expression profiles as features from various sources
    * **Data Integration**: combined multiple sources of data through data-concatenation, data transformation and ensemble learning depend on specific questions
    * **Machine learning**: built random forest model to quantify disease risk of genes
    * **Visualization**: utilized dimension reduction techniques and cluster methods to visualize high dimensional data gene expression datasets and identify patterns

  Predict genetic risk for breast cancer in East Asian population
    * **Feature Engineering**: developed approaches to transform data from various population sources to same scale for further prediction models
    * **Statistical Analysis**: conducted meta-analysis to predict genetic risk in East Asian by leveraging European sample results and increased prediction auROC from 0.56 to 0.62

PROJECTS

- **Predict patient readmission from high dimensional hospitalization Data**: conducted feature engineering to reduced feature scale by 10 times of sparse, messy hospital billing data with 10,000 subjects, improved computation efficiency and prediction accuracy

SKILLS

- : **Environment**: Unix/Linux
- : **Coding Languages**: R Python
- : **Data science toolkits**: dplyr, numpy, scipy, scikit-learn, pandas
- : **Analytics & visualization tools**: ggplot2, matplotlib, Tableau

OTHER

- **Awards**: 2019 Graduate Leadership Development Institute (GLDI) travel award
- **Presentations**: 2018 American Society of Human Genetics Annual Meeting platform talk (top 8% abstracts submitted)
- **Activities**: 2017-2018 Secretary of Human Genetics Graduate Student Association
- **Teaching**: Mentored graduate rotation student perform analysis using UNIX and R