Wangjing (Jing) Ke, PhDc, MS, MPH

Tel: 520-360-8048 | **Email:** kwangjing@gmail.com | **Website:** http://wangjingke.com
LinkedIn: https://www.linkedin.com/in/wangjingke | **GitHub:** https://github.com/wangjingke

PROFILE

Data scientist with domain knowledge on healthcare and insurance data, and hands-on experience in statistical analysis, machine learning, and data engineering. Lead cross-functional teams and collaborate with business partners to extend the internal strategic business process, re-engineer the workflow and generate tangible business and healthcare values.

SKILLS

Programming: R, SAS, Python, Scala, SQL, Hadoop, Hive, Spark MLlib, BI tools
Statistics: study design, A/B testing, predictive modeling, time series
Machine learning: random forests, classification, clustering, regression models, neural network

WORK EXPERIENCE

Molina Healthcare, Inc *Long Beach, CA*
Data Mining Specialist *Mar 2017 – Present*

- Led the advanced analytics team in an enterprise initiative to avoid $9M interest payment and government penalty:
  - Trained stochastic gradient descent model on streaming data to identify providers with configuration issues, reducing medical claim denial errors and cutting claim processing time.
  - Detected abnormal patterns in member and provider activities for fraud investigations.
  - Created tier mechanisms over 1M healthcare providers using kernel SVM, offering automatic recommendations on doctor referrals, yielding an estimated $8M first year saving.
- Developed statistical methodology to maximize the efficacy of healthcare intervention programs:
  - Built hybrid Bayesian and time series models to predict thousands of healthcare quality measures for the coming year within 2% margin, cutting the turnaround time by 95%.
  - Trained random forest and regression models to design personalized health campaigns, improving the enterprise-wide quality compliance by 3%, increasing member satisfaction with no extra cost.
- Queried >600M data lines and identified 39K TX victims at risk of evacuation hours before hurricane Harvey's arrival. Predicted and requested essential medical supplies for 200 rescued members based on their healthcare history.

University of Southern California *Los Angeles, CA*
Data Scientist *2013 – Mar 2017*

- First author and co-author on multiple academic publications.
- Managed >20 TB data, and provided 40+ researchers with analytic insights and presentation of statistical results.
- Designed and built machine learning models, using principal component analysis, clustering and kNN algorithms to scan 40 million DNA mutations from 250K patients, and identified 71 heart disease related genes.

EDUCATION

University of Southern California *Los Angeles, CA*
Master of Science – Biostatistics (GPA: 3.86/4.0) *2016*
Doctor of Philosophy (ABD) – Epidemiology *2013 – 2015*

University of Arizona *Tucson, AZ*
Master of Public Health – Epidemiology (GPA: 3.92/4.0) *2013*
Bachelor of Science – Biology *2011*

- **Awards/Honors:** Graduated *Summa Cum Laude*, *Phi Beta Kappa* inducted

ADDITIONAL INFORMATION

Certificates: SAS Certified Base Programmer, SAS Certified Advanced Programmer
Interests: skiing, analyzing basketball games, and making drinks (certified by the Tucson Bartending Academy)