

Hongling Yang | Data Scientist

Address Temecula, CA 92593
Phone (915) 219-1391
E-mail hyang78227@gmail.com

LinkedIn www.linkedin.com/in/hongling-yang-909773227
Github <https://github.com/hyang78227>

PhD-level data scientist with over 10 years of analytical experience. Proficient in statistical modeling, machine learning, and big data tools. Passionate about translating complex data into actionable business insights.

SKILLS

Programming: Python, R, SAS, C++

Databases: MS SQL Server, MySQL, Postgres

Machine Learning: TensorFlow, Pytorch, Pyspark, Neural Networks,
Time Series, Clustering, SVM, Naive Bayes

Tools: Hadoop Ecosystem, Image Processing, Natural Language Processing, Recommendation Systems

Analytics: Data Wrangling, Data Visualization, Predictive & Forecasting Analytics

EDUCATION

PhD in Statistics 2005 - 2008
Arizona State University - Tempe, AZ

MS in Statistics 2003 - 2005
Texas - El Paso El Paso, TX

EXPERIENCE

Data Science Trainee 2023-2023
Sprinboard, San Francisco, CA

- **Fashion Product Image Classification** - [GitHub - hyang78227/capstone-project3](https://github.com/hyang78227/capstone-project3)
Achieved 97.5% accuracy with CNN via Transfer Learning; enhanced minority class representation and implemented Hyperband for optimization.
- **Google App Store Educational Apps Analysis** - [GitHub - hyang78227/CapstoneProjectTwo](https://github.com/hyang78227/CapstoneProjectTwo)
Predicted app ratings with up to 94% accuracy using multiple classifiers; addressed class imbalance and optimized using Hyperband.
- **Big Mountain Ski Resort Pricing Analysis** - [GitHub - hyang78227/DataScienceG](https://github.com/hyang78227/DataScienceG)
Applied regressions to increase ticket value by up to \$19; streamlined data pipeline and improved model accuracy to 96% with GridSearchCV.

Statistician 2016-2017
School Of Medicine, University Of California, San Diego, CA

- Spearheaded analyses in public health research focusing on HIV risk factors and interventions, utilizing multilevel modeling and structural equation modeling.

Statistical Consultant 2010- 2014
Texas Tech Health Center, El Paso , TX

- Guided medical research, mentored residents, and partnered with physicians in clinical trials.

Statistician & Lecturer 2008-- 2016
College of Engineering, University of Texas, El Paso, TX

- Conducted GIS research targeting pollution analysis and collaborated on various academic papers. Instructed courses in statistics and mathematics.

CERTIFICATION & LANGUAGES

SAS Advanced, Base, Clinical Trials Programmer for SAS 9
Fluent in Chinese and English