Yifan Zhang

Tucson, AZ 85719 (607)-319-1395 yifanzhang0705@gmail.com

Research Interest
- Machine Learning
- Statistical Application
- Bioinformatics

Education
08/21 – Present
University of Arizona, Tucson
- Ph.D. Student
- Academic Highlights: Single Cell RNA Sequence

08/18 - 05/20
Cornell University, Ithaca
- Master of Professional Studies in Statistics
- Academic Highlights: Machine Learning, Tableau, Python, R, SQL, SAS, Big data management: Hadoop, Hive, MapReduce

09/14 - 06/18
University of California, Santa Barbara, Santa Barbara
• Bachelor of Science in Mathematical Sciences
• Minor in Statistical Science & Spanish
• Academic Highlights: Python, R, Regression analysis, SAS, Stochastic process and Markov chain

Related Projects/ Research/ Employment
05/16 - 06/16, University of California Santa Barbara, Goleta
- Project using R for regression analysis.
- Forecasting if donation amounts can be inferred from factors such as the contributor's income, education level and political party affiliation.
- Leading the group to interpret data trends and implement analytical models such as linear regression and probabilistic model.
- Formulating analyses, insights, and recommendations based on the built models.
- Visualizing models to track the performance of the prediction.

11/19 - 12/19, Cornell University, Ithaca
- Project using SQL and SAS to perform the data analysis of school revenues.
- Developing and implementing databases, collecting big amount of data.
- Acquiring data from databases by writing SQL commands.
- Testing the results using different commands to ensure the quality of the product.

01/20 - 06/20, Cornell University, Ithaca
- Research project collaborating with the customer Gravity Investments using R to forecast quarterly stock returns of Russell 3000 in financial market.
- Communicating with the company executives to prioritize business and data objectives.
- Collecting the big finance data set from Bloomberg Terminal and the Economic Research Database of the Federal Reserve Bank of St. Louis, ranging from 2005-2020.
- Filtering, cleaning and transforming the data set.
- Designing risk-based analysis plan and implementing analytical models such as random forest and gradient boosting.
- Visualizing the models to compare the results. Improving the quality of the final model.
- Integrating the final result with the financial service of Gravity Investments to help extract insights and make decisions for the clients.

04/20 - 06/20, Cornell University, Ithaca
- Project using python and Tableau to solve problems for the game League of Legends (LOL).
- Using Tableau to visualize the performance of each game.
- Identifying different quality issues such as variable correlations and cross-validation for the big data set.
- Implementing models such as logistic regression and support vector machine to find out factors that are most efficient in predicting the outcome of the games.
- Writing the project report including both detailed explanation of analytical methods applied, and the solution preferred.

05/20 - 06/20, Cornell University, Ithaca
- Project using SQL, Hive, MapReduce, Pig to perform data analysis of a big data set.
- Creating the databases and collecting data from the information of delayed flights.
- Acquiring and analyzing the data by applying statistical techniques such as finding the correlation.
- Identifying and interpreting trends among features such as type, location, etc.
- Integrating the result with Python to filter the information for the better quality of the final result.