Guanyan Lin

1750 N Range Rd, IN, Bloomington
aaronarchiblad@gmail.com | 812-369-6456

EDUCATION

INDIANA UNIVERSITY BLOOMINGTON | MS in data Science
May 2018 | Bloomington, IN • Cum. GPA: 3.57/4.00

INDIANA UNIVERSITY BLOOMINGTON | BS in statistics | BA in Economics
Minor: Mathematcs, Business Consulting
May 2016 | Bloomington, IN • Cum. GPA: 3.34/4.00 | Major GPA: 3.74 / 4.00

WORK & RESEARCH EXPERIENCE

DANIEL MCDONALD'S LAB | Machine Learning Research Assistant
May 2017 – Present | Indiana Univeristy Bloomington, IN

- Read quantum chemistry papers and implement the most recent machine learning models and algorithms.
- Utilize programming languages, such as R and Python, to change the file format and clean chemical data.
- Use different molecule representations to transform input data such as BOB and Coulomb matrix.
- Implement kernel ridge regression and neural networks in R and Scikitlearn in serial and parallel.

INDIANA STATISTICAL CONSULTING CENTER | Statistical Anaylst Assistant
Sep 2015 - May 2016 | Indiana Univeristy Bloomington, IN

- Used strong analytical skill to assist graduate students and faculties by understanding their research projects and providing statistical solutions for research projects.
- Built statistical models such as generalized linear models and repeated measurement ANOVA to provide data driven evidences for research projects.
- Utilized programming software such as R, Excel, and Python to clean data and to run analysis on surveys and lab projects.
- Worked independently and collaboratively with 2 supervisors and 1-3 consulting associates to develop solutions for clients.

BUSINESS CONSULTING PROJECTS

MCD MACHINE, INC CASE | Data analyst
Aug 2014 - Dec 2014 | Bloomington, IN

- Interviewed the owner of the manufacturer to find out what he concerned and to understand bidding, ordering, manufacturing, shipping, and paying processes.
- Collected shipping logs, clock-in system, financial statements, and cost from the manufacturer and explore those data to find out how to reduce overhead cost.
- Collaborated with a team of three otherswit different academic backgrounds to provide a 45-minute presentation explaining our data based results and solution to not only layperson and professional business analyst.

KAGGLE
Aug 2016 - Dec 2016 | Bloomington, IN

- Ultilize R, Python, and postgreSQL to extract, transform, and clean data in different formats ranging from MB to GB.
- Implement different statistical models, such as generalized linear models, tuning the meta-parameter with cross validation criterion to find out the best model in the sense of minimizing the out of bag error based on mean squared error.
- Modify or search the models and algorithms from the literature in the industry, if the given loss function I have never seen.

TECHNICAL SKILLS

CORE COURSES
Econometric • Financial Economics • Managerial Economics
Machine Learning • Data Mining • High dimensional data analysis
Artificial Intelligence • Exploratory Data Analysis • Time Series Analysis

PROGRAMMING LANGUAGE
R • Python • PostgreSQL • LaTeX
C(basic) • JavaSE(basic)