Wenxin (Grace) Fang

wenxinf2@illinois.edu | 812-964-9039 | 210 S 4th St, Champaign, IL 61820

EDUCATION

University of Illinois at Urbana-Champaign(UIUC) | Champaign, IL — Aug 2019–May 2021(expected)
Master of Science in Statistics — GPA: 3.89/4.00

- Coursework: Data Structures & Algorithms, Database Systems, Machine Learning

Indiana University - Bloomington | Bloomington, IN — Jan 2016–June 2019
Bachelor of Science in Information Systems, Minor in Mathematics — GPA: 3.61/4.00

- Coursework: Data Programming, Application Development, Management Information Systems, Technology Infrastructure

TECHNICAL SKILLS

- **Programming**: Python, SQL, C++, C#, Java, VBA, R, HTML, CSS
- **Tools & Systems:** NoSQL, MongoDB, Neo4j, MapReduce, Hive, Jupyter Notebook, Linux

PROFESSIONAL EXPERIENCE

FinTell Financial Services — **Beijing, China**
Big Data Intern (HiveSQL) — *June 2018–Aug 2018*

- Built purchasing power profiles for over 900 million loan borrowers based on the online shopping apps downloads and usage time to predict default risk
- Wrote HiveSQL queries on Hadoop to classify over 1,700 shopping apps from multiple mobile app databases into three big categories and nine subcategories based on the quality and price of products sold
- Created a shopping app table in Hive that integrates the number of downloads and usage in each shopping app category along with a calculated self-defined Purchasing Power Score (PPS) for given IMEI number

My Sister's Closet (NGO) — **Bloomington, IN**
Software Management Intern (VBA) — *Feb 2018–May 2018*

- Developed a VBA program to detect and remove duplicate records of donor contact information from multiple data sources
- Built an Access database to track volunteer activity and increased efficiency of volunteer data management

PROJECTS

Walmart Weather-Sensitive Products Sales Prediction (R, Python) | Champaign, IL — *Fall 2019*

- Extracted and cleaned over four million weather and sales data to predict sales of 111 weather–sensitive Walmart products
- Conducted exploratory data analysis using ggplot, performed feature engineering by extracting date information and built Stepwise Regression with backward and AIC criterion in R
- Built ensemble models (Random Forest, CatBoost) in Python that achieved an RMSLE of 0.11744

Machine Learning Data Modeling | Champaign, IL — *Fall 2019*

- House Sales of King County Analysis: Built supervised methods (Random Forest, Boosting, Regression, GLM) by tuning hyperparameters with cross-validation to predict the housing sales in King County
- Mnist Handwritten Digit Analysis: Built 1NN to predict the handwritten digit that decreased the test error to 5%
- Credit Card Fraud Detection Analysis: Preprocessed imbalanced data using smote method and built Logistic Regression model to detect credit card fraud, resulting ROC-AUC Score of 90% on the test data

**Database Design Project (SQL*)* | Bloomington, IN — *Spring 2018*

- Constructed a database for a music truck business by designing E-R Diagram and implementing it in the Oracle database
- Conducted SQL queries to extract and display useful information to customers, employees, and managers, resulting in the improved customer shopping experience and efficiency of business operation

SAP ERP Simulation Competition | Bloomington, IN — *Spring 2018*

- Led a team of four to make strategic decisions for a dairy company and implemented a concentration strategy targeting on high margin products on SAP system, achieving the 2nd highest company's equity value in the final competition
- Designed a data visualization dashboard in SAP Lumira to analyze transactional data, identify the pattern of customer preference in different regions and adjust forecast amount based on product demand accordingly