ZIQIAN DING

📱 (+86) 183-8109-2137· ✉ ziqian.ding@smail.nju.edu.cn· 🏠 www.ziqianding.com

EDUCATION

Nanjing University <div align="right">Sep. 2019 - Jul. 2023 (Expected)</div>
B.S. in Computer Science and Technology <div align="right">*Jiangsu, China*</div>

- **GPA**: 4.60/5.0 (92.0/100) | **Rank**: 3/189
- **Relevant Coursework**: Econometrics, Microeconomics, Precision Marketing Analytics, Algorithm Design & Analysis, Big Data Processing, Data Structures, Machine Learning, Applied Time Series, Linear Algebra, Probability & Statistics
- **Honors**: National Scholarship (Top 0.2% nationwide), Ruli Scholarship (Top 5%), Wang Laoji Scholarship (Top 5%), Outstanding Student of Nanjing University (Top 5%)

RESEARCH EXPERIENCE

Harvard Business School <div align="right">Feb. 2022 - Present</div>
Research Associate <div align="right">*Boston, MA*</div>

- Implemented a ResNet-based deep learning model to categorize the visual emotions of images using PyTorch and achieved 69.1% classification accuracy for 8 emotion categories on the FI dataset; fine-tuned a BERT model to detect the Big 5 Personality Traits of Airbnb hosts from their listing descriptions.
- Developed text mining algorithms to estimate speed, volume and circuitousness of texts with Python and MATLAB; extracted, cleaned and analyzed 1 TB of consumer purchase data from a leading shopping platform in Hong Kong.
- Built an automated tool in Python to scrap 30,000+ Airbnb listings using Scrapy and Airbnb's unofficial API.

UNC Kenan-Flagler Business School <div align="right">Oct. 2021 - Mar. 2022</div>
Research Assistant <div align="right">*Chapel Hill, NC*</div>

- Conducted data extraction and cleaning on 800 GB of demographic, account and transaction data for 2M+ Bank of China customers; performed exploratory analysis using Pandas and Seaborn.
- Analyzed data employing econometric, statistical models and machine learning tools, including PSM (Propensity Score Matching), DiD (Differences-in-Differences) and DML (Double Machine Learning).

WORKING PAPER

- "COVID-19 Telemedicine Expansion: Hidden Consequences on Doctor Churn.", with Vinayak Deshpande and Yuqian Xu. 2022. Soon to be submitted to *Manufacturing & Service Operations Management*.

PROJECTS

Big Data Processing on Basketball Game Logs <div align="right">Jun. 2022 - Jul. 2022</div>

- Cleaned and explored 3 TB of basketball game log data with Hive (a distributed data warehouse system).
- Constructed a distributed KNN model to predict future game results using Hadoop MapReduce (a distributed computing framework) and achieved a prediction accuracy of 61.2%.
- Summarized preferences and tactics of players and teams based on data analysis and machine learning predictions.

Instacart Market Basket Analysis <div align="right">May. 2022 - Jun. 2022</div>

- Cleaned and performed feature engineering on past purchase data for 200K+ Instacart customers from Kaggle.
- Implemented a LightGBM model to predict which previously purchased products a user will buy in the next order using R, and attained 62.8% accuracy and 0.478 F1-Score with the baseline model.
- Improved model performance by about 17% with hyper-parameter tuning and grid search cross-validation.

SKILLS

Programming Languages & Tools	C++, C, Java, Python, R, MATLAB, Stata, SQL, Markdown and LaTeX
Frameworks & Packages	Pandas, NumPy, PyTorch, Scrapy and Hadoop