Jade Cheng

984-484-0239 | haojing.cheng2025@gmail.com | haojingcheng-jade.github.io/jade-cheng | NC

Educations

Duke University Durham, NC
Master of Science in Statistical Science *Aug. 2025 – May 2027*

Communication University of China Beijing, China
Bachelor of Engineering in Data Science *Sep. 2021 – June 2025*

- Major GPA: 3.80/4.00 Top 10%
- Data Structure, SQL, Python, Probability, Machine Learning, Data Visualization

Experience

Software Center of the Industrial and Commercial Bank of China June 2023 – Aug 2023
Data Analyst Intern *Beijing, China*

- Analyzed 12 months of customer card activation, transaction, and app binding data for users aged 18-22 using **SQL**, identifying a 130.97% surge in new activations during the back-to-school season.
- Segmented user binding data to identify significant adoption gaps: 70.47% of target users unbound on JD.com and 56.13% on Meituan, providing data-driven insights to inform partnership and marketing strategies.
- Designed and executed a multi-variant **A/B test** for a card-binding campaign; used SQL for real-time monitoring of **CTR** and conversion rates to optimize channel selection and push timing.
- **Collaborated with cross-functional teams** (legal, algorithms, operations) to implement the campaign, resulting in an 18% increase in overall conversion, 25% lift on JD.com, 22% on Meituan, and a 15.6% CTR.

Projects

Behavior Analysis and Visualization for China Mobile | *SQL, PowerBI* June 2024 – July 2024

- Deliver actionable intelligence: Established and optimized a **MySQL** database to ensure integrity and efficient querying of large-scale mobile signal data. Integrated user distribution patterns from China Mobile across major tourist attractions during peak holiday seasons.
- Transform data into insights: Leveraged **PowerBI** to design an interactive **DASHBOARD** highlighting real-time metrics such as visitor density, attraction clustering, and demographic profiles. Enabled data-driven decision making for optimizing resource allocation and enhancing visitor management strategies.

Neural News Recommendation System with LLM Explanation | *Prompt, ML* Dec 2024 – May 2025

- Processed and cleaned the MIND dataset using **Pandas** and **NumPy**, handling missing values, normalizing features, and tokenizing text to create a dataset for model training and evaluation.
- Conducted comparative experiments on **DEEP LEARNING** models (DKN, NRMS, NAML), performing parameter tuning and evaluation to identify NAML as the optimal algorithm.
- Integrated the **DeepSeek API** and engineered **PROMPTS** to generate natural-language explanations for news recommendations, including news summaries, relevance to user history, and potential interest areas.

Technical Skills

Programming Languages: Python, SQL, C/C++, R
Analysis & Statistical Techniques: Statistical Modeling, Machine Learning, A/B Testing
Data Tools: Tableau, Power BI, Excel
Libraries: Pandas, NumPy, Matplotlib
Additional: Teamwork, Communication, Time-management