Kai Lu

bravo6k@gmail.com | (312)447-1396 | linkedin.com/in/kai-lu-bravo6k | Mishawaka, IN, 46545

Education

University of Illinois Urbana-Champaign Aug. 2016 – May. 2018
Master of Science in Statistics

Shanghai University of Finance and Economics Sep. 2012 – Jun. 2016
Bachelor of Arts in Statistics

Experience

Aunalytics. INC *Data Scientist, South Bend* Jul. 2018 – Jul. 2024

- **Natural Language to SQL RAG System** Mar. 2020 – Jul. 2024
  - **Overview:** Developed an NL2SQL RAG system for **10+** local banks and a supermarket, reducing data retrieval time by **70%** and significantly improving data accessibility for **50+** employees.
  - **Responsibilities:** Led the entire machine learning system design, including requirements gathering, problem definition, model development, and deployment.
  - Implemented a two-stage NL2SQL translation model: an attention-based classification model (**variable class numbers**), integrated GPT-4 for SQL generation, achieving an execution accuracy of **85%** on public spider dataset.
  - Enhanced training data quality and diversity by implementing a SQL2NL model using DeepSpeed and QLoRA for multi-GPU fine-tuning of Llama 3.
  - Engineered an interactive error analysis dashboard, streamlining model diagnostics and **reducing error analysis time by 90%**, thereby accelerating the iterative improvement cycle of the NL2SQL model.
  - Designed a comprehensive CI/CD pipeline for the machine learning system, including DVC data version control, automated deployment using GitHub action, rollback mechanisms and backup model management.
- **Multi-Purpose RAG Chatbot** May. 2024 – Jul. 2024
  - **Overview**: Developed a multi-tier system chatbot leveraging GPTs, capable of handling conversations, complex data retrieval tasks like SQL generation, dashboard creation and report/dashboard retrieval.
  - Led the design of the chatbot architecture, focusing on efficient task routing, prompt engineering.
- **Customer Attrition Assessment for Banks** Jun. 2019 – Oct. 2019
  - **Overview**: Developed a customer churn prediction system for **10+** local banks, enabling proactive retention strategies with the potential to retain up to **$20 million in annual deposits** across all participating banks.
  - **Responsibilities:** Scoped and led the project, including problem definition, data architecture design and implementation of the machine learning model.
  - Conducted extensive analysis on historical churn behavior patterns to define optimal churn criteria. Implemented a monthly snapshot approach to construct a historical dataset with a 0.85% churn rate.
  - Developed an XGBoost model achieving an AUC-ROC of 0.79 and utilized SHAP for model interpretation, enabling bank stakeholders to identify key factors influencing customer churn.
  - Achieved a precision of 5.43% within the highest decile of churn scores, and 13.3% precision for the top 100 high-risk customers identified, demonstrating strong predictive capability for prioritizing retention efforts.
- **Node Congestion Prediction** Jan. 2019 – Mar. 2019
  - **Overview**: Improved an existing network congestion prediction system, reframing from time series forecasting to a classification approach for more accurate abnormal peak detection.
  - Increased the system's recall by 4 times, dramatically improving its ability to detect network congestion events.

Publication

Kaiwen Dong, **Kai Lu**, Xin Xia, David Cieslak, Nitesh Chawla, "An Optimized NL2SQL System for Enterprise Data Mart." *In the proceedings of Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track, 2021.*

Skills

Python(Pandas, Matplotlib, Scikit-learn, Pytorch, Keras, Tensorflow), SQL, Dashboard, Spark, Docker, Git, Linux