Ke Xu

✉ kxu27@u.rochester.edu | 📞 585-358-9858 | ⭘ github.com/kefoto | ▣ linkedin.com/in/kxudata

Education

M.S. Data Science, Vanderbilt University	**May 2027**
B.S. Data Science — B.S. Business Analytics, University of Rochester	**May 2025**

- Cumulative GPA: 3.65/4.0

Experience

UR MEDICINE MOTION LAB ....

Data Science Intern   **January 2025 – May 2025**

- Engineered scalable **PostgreSQL** schema to normalize 3,000+ 3D motion and EMG datasets with 2,000+ features
- Automated **data ingestion pipelines** with **sqlalchemy** scripting, saving 10+ hours weekly
- Built high-accuracy **LSTM model** (92%) for real-time IMU motion sequence classification
- Enabled efficient querying and integration in research pipelines through **database management**

GOERGEN INSTITUTE OF DATA SCIENCE, UNIVERSITY OF ROCHESTER ....

Undergraduate Research Assistant   **September 2024 – May 2025**

- Deployed **multimodal AI** models for stock forecasting using sentiment-video-audio analysis
- Built **visual-audio processing pipeline** with **speech-to-text** and **speaker diarization**
- Implemented accelerated **PyTorch** and **HuggingFace** frameworks for emotion-based forecasting models
- Analyzed correlation between election precinct outcomes and visual OCR detection of yard signs under Professor Caliskan

ZTO EXPRESS INC. ....

Data Analytics Intern   **June 2024 – August 2024**

- Developed **NLP pipeline** to identify 15% of major clients from 34PB annual order data
- Leveraged **Spark** and **Hive** for data extraction from warehousing systems
- Refined client selection criteria and implemented new KPIs for secondary market classification and targeting

Additional Experience

MEMORIAL ART GALLERY ....

IT Assistant   **September 2024 – June 2025**

- Developed modular **Python scripts** for **API integration** and automated report generation
- Reduced report preparation time by 35% with structured documentation and accurate data visualization
- Built **React-Flask** web application with interactive calendar views and multi-location support
- Implemented **real-time** filtering, PDF generation, and emergency tracking features with IIS deployment on protected network

Skills

Technical Skills: **Machine Learning**, **Deep Learning**, **Data Science**, **Artificial Intelligence**, **Computer Vision**, **Natural Language Processing**, **Time-Series Analysis**, **Statistical Modeling**, **Data Mining**, **Big Data Analytics**
Programming: **Python**, **SQL**, **R**, **Java**, **JavaScript**, **C++**, **Bash Scripting**
ML/AI Frameworks: **PyTorch**, **TensorFlow**, **Scikit-Learn**, **HuggingFace Transformers**, **OpenCV**, **Pandas**, **NumPy**
Data Engineering: **Apache Spark**, **Hive**, **PostgreSQL**, **MySQL**, **Data Warehousing**, **Cloud Computing**
Development Tools: **Git**, **GitHub**, **CI/CD**, **REST APIs**, **React**, **Flask**
Languages: English (Professional), Mandarin (Native)

Awards

1st Prize Winner – AI Summit New York 2024 Hackathon,   **2024**

- Developed a digital assistant and dashboard for health and safety that analyzes vitals via a smartwatch to notify EMS during life-threatening crises with Google's Map, Gemini, and Telegram API.

Best Paper Award – Summer Institute of Computational Social Science (SICSS) 2024 Rochester,   **2024**

- Machine Learning-Based Analysis of Healthcare Institution Preferences in Turkey: Insights for Health Policy.