# Andy (Xiangyu) Cui

xiangvucui@outlook.com | (402)-853-3000 | Linkedin | Portfolio

#### **EDUCATION**

**Northeastern University** Boston, MA M.S. in Artificial Intelligence of Khoury College Dec 2023 University of Nebraska-Lincoln B.S. in Computer Science of Arts Science College

Lincoln, NE May 2020

### PROFESSIONAL EXPERIENCE

King 7 Club Corp Senior Software Engineer Jan. 2025-Present Los Angeles, CA

- Developed and deployed a responsive website using React and Node.js, enhanced with CSS 5 animations and hover effects for superior UI/UX, and implemented a local storage-based language toggle to efficiently switch between English and Chinese, reducing API latency and speeding up page loads. Ensured a consistent mobile interface with React's responsive design, maintaining uniform content and UI across all platforms.
- Integrated Google Analytics Reporting API to track website traffic and generate detailed reports, providing insights that helped design interactive dashboards for monitoring metrics such as page views and URL clicks, thereby facilitating data-driven strategies to boost social media engagement on platforms like TikTok, Red Book, and YouTube.
- Optimized global website performance by deploying through AWS, ensuring top access speeds with services including CloudFront for content delivery, S3 for storage, Route 53 for intelligent DNS routing, and Global Accelerator for optimized connectivity. Used **JSON** for data structuring to support scalable backend development and consistent data management.
- Automated tracking and reporting of web-driven social media engagement, increasing operational efficiency by 80%, and organized common modules like navigation bars and footers into separate directories for better scalability and maintainability through modular design. Managed version control with Git and led Agile development processes, ensuring effective project execution and timely delivery.

**CAC Auto Group LLC** Feb. 2024-Dec. 2024 Data Engineer Southborough, MA

- Developed and maintained a predictive pricing system for vehicles on CarGurus using AWS serverless architecture, optimizing market compatibility and operational efficiency. Utilized AWS services including S3, Lambda, DynamoDB, SNS, CloudWatch, and Kinesis for real-time data processing, Python and AWS CloudFormation for infrastructure.
- Implemented real-time monitoring with AWS Kinesis to track market data fluctuations, enabling automated detection and response to price deviations. Improved the accuracy of vehicle pricing on the company's website, resulting in an 80% boost in daily operational efficiency and over 50% improvement in real-time price adjustments compared to industry standards.
- Designed and deployed a CRM system on Docker using WordPress, ensuring the system's scalability and simplifying ongoing maintenance; Customized the CRM with essential features such as source tracking and lead status monitoring to enhance sales processes and improve visibility, saving significant time in monitoring sales conversion rates and successfully recovered 80% of potential sales leads, resulting in a notable increase in overall sales performance.

### **PROJECTS**

# Job Recommendation System Design

- Developed a user interface for job searching using Axure RP 10; Applied content-based filtering using TF-IDF and cosine similarity, achieving 82% precision in matching user skills to job descriptions; Conducted collaborative filtering in Python with implicit user feedback, improving recommendation diversity by 18% via matrix factorization.
- Leveraged deepseek API to dynamically adjust recommendations based on real-time user feedback; Reduced cold-start bias by 30% through RL-driven exploration of niche roles.

## **Stock Price Prediction with Deep Learning**

Oct 2024

- Collected the historical stock price and other financial assets data on the company of interest; Conducted data preprocessing by applying min-max scaling in **Sklearn** to normalize stock price values, ensuring consistency across the dataset.
- Implemented LSTM, GRU, and Transformer models in PyTorch, optimizing hyperparameters (e.g., number of layers, optimization methods) through grid search, increasing model accuracy by 20%; Visualized opening and closing price trends to assess model performance in Python.

## Auto QA Chat Agent for Customer via NLP

- Collected and cleaned large-scale Amazon customer Q&A datasets, ensuring high-quality data for effective AI model training; Developed a conversational AI agent in PyTorch leveraging state-of-the-art NLP models, including BERT and GPT2, fine-tuned for question answering (QA) tasks.
- Optimized hyperparameters for BERT and GPT-2 to improve contextual understanding and response generation; Conducted 10-fold cross-validation for BERT and 5-fold for GPT-2, achieving average BLEU scores of 0.9 for BERT and 0.8 for GPT-2.
- Integrated continuous learning capabilities to adapt to evolving customer queries, improving long-term operational efficiency; Successfully automated 70% of routine customer inquiries, significantly reducing operational costs.