Andy (Xiangyu) Cui

GitHub | LinkedIn | Portfolio
xiangyucui@outlook.com | (402)-853-3000

EDUCATION

Northeastern University — Boston, MA, USA
M.S. in Artificial Intelligence of Khoury College — Sept. 2021-Dec.2023

University of Nebraska-Lincoln — Lincoln, NE, USA
B.S. in Computer Science of Arts Science College — Sept. 2016-May. 2020

Dalian Neusoft University of Information — Dalian, LN, CHN
B.S. in Electronic Information Engineering — Sept. 2013-Sept. 2015

SKILLS

- Programing languages: Java, Python, JavaScript, HTML, CSS, C/C++, Assembly, VB
- Database: MySQL, PostgreSQL, MongoDB, Redis, SQLite, Hive
- Machine Learning: TF-IDF, Naïve Bayes, GPT-2, Bert, CNN, Transformer
- Framework/Libraries: Spring Boot, React.js, Node.js, jQuery
- Version Control & CI/CD: GitHub, TFS, Azure DevOps, Jenkins, Jira
- Others: AWS, Docker, Maven, Tomcat, Axure, Servlet, Junit, Nginx, LaTeX, CAD design, Tableau

WORK EXPERENCE

AlpaLifeBio Inc | Internship — Woburn, MA, USA
Software Engineer & Data Analyst | Analysis of Potential Consumer Demand — Dec. 2022-June. 2023

- Executed data collection from biomedical public databases using **Python** and **graphical web scraping tools**, enhancing the dataset for advanced analysis. Additionally, processed and analyzed over 500,000 data entries daily, utilizing graphical data representations to streamline reporting and significantly boost workflow efficiency.
- Applied **TF-IDF** techniques for data comparison and matching, effectively identifying potential clients with a match rate of up to 95%, significantly contributing to targeted marketing strategies.
- Consolidated existing datasets for integrated analysis and established a structured **SQL** database, implementing tag processing for improved search and retrieval operations. This approach has saved 80% of the time for future data searches and significantly enhanced work efficiency, thereby saving costs for the company.

Dutchgo LLC | Self-employed & Co-founder — Omaha, NE, USA
Team Leader & Project Analyst | Analysis and Strategy Optimization — Oct. 2020-Apr. 2022

- Analyzed customer demand for streaming media during the Covid-19 pandemic using advanced **SQL** queries and **Python**, then trained and tested a **Random Forest** model with **Python** to identify potential user profiles for conversion, achieving an accuracy of 98%.
- Determined customer markets and consumption levels based on selected categories, extracted and preprocessed monthly market data of customers using **SQL** and **Python**, and predicted revenue trends using a multiple regression model with an accuracy of 85%, providing recommendations for strategic business decisions.
- Built a secondary sales server pipeline based on selected categories and customer markets, analyzed key churn points in the conversion funnel using funnel analysis in **SQL**, communicated critical insights to the engineering team, and implemented funnel improvements based on findings.
- Created interactive dashboards with **Tableau** to monitor key investment metrics for stakeholders, including **ROI (Return on Investment)**, growth rate, and portfolio balance, and conducted comprehensive root cause analysis of any abnormal trends.

Chasing Light(Zhuiguang) Information Technology Co. Ltd | Co-founder — ZhengZhou, Henan, CHN
Team Leader & Project Analyst | Customized Analysis for Client Requirements — Jan. 2019-Sept. 2020

- Analyzed customer requirements to draft an **activity diagram** and outlined preliminary **UML** page functionality scenarios. Initiated the development of customized software (Web & App) tailored to customer needs.
- Employed advanced **SQL queries** and **Python** to analyze over 50GB of market data on similar products, identifying customer usage patterns and business requirements.
- Conducted exploratory data analysis on numerical, categorical, and time-series data using **Matplotlib** and **Seaborn**. Constructed multivariance Time Series Clustering model, DTW (Dynamic Time Warping) with Hierarchical clustering, on 120 features to group customers, segment market and described the clustering centroid by DBA (DTW Barycenter Averaging).
- Trained market demand models, including **Isolation Forest** and **Robust PCA** (Principal Component Analysis), as well as **Linear Models** and **Random Forests**, to establish a system for identifying and responding to customers' potential requirements.
- Constructed dashboards with **Tableau** to depict customer usage activities and submitted analytical reports to clients.