

# Andy (Xiangyu) Cui

[GitHub](#) | [LinkedIn](#) | [Portfolio](#)

[xiangyucui@outlook.com](mailto:xiangyucui@outlook.com) | (402)-853-3000

EDUCATION	Relevant Coursework	Location & Date
<b>Northeastern University</b> <i>M.S. in Artificial Intelligence of Khoury College</i>	📁	Boston, MA, USA Sept. 2021-Dec.2023
<b>University of Nebraska-Lincoln</b> <i>B.S. in Computer Science of Arts Science College</i>	📁	Lincoln, NE, USA Sept. 2016-May. 2020
<b>Dalian Neusoft University of Information</b> <i>B.S. in Electronic Information Engineering</i>	📁	Dalian, LN, CHN 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 RP, Servlet, Junit, Nginx, LaTeX, CAD design, Tableau

## WORK EXPERIENCE

<b>AlpaLifeBio Inc   Internship</b> <i>Software Engineer &amp; Data Analyst   Analysis of Potential Consumer Demand</i>	Woburn, MA, USA Dec. 2022-June. 2023
<ul style="list-style-type: none"><li>● Executed data collection from biomedical public databases using <b>Python</b> and <b>graphical web scraping tools</b>, 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.</li><li>● Applied <b>TF-IDF</b> techniques for data comparison and matching, effectively identifying potential clients with a match rate of up to 95%, significantly contributing to targeted marketing strategies.</li><li>● Consolidated existing datasets for integrated analysis and established a structured <b>SQL</b> 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.</li></ul>	
<b>Dutchgo LLC   Self-employed &amp; Co-founder</b> <i>Project Analyst   Analysis and Strategy Optimization</i>	Omaha, NE, USA Oct. 2020-Apr. 2022
<ul style="list-style-type: none"><li>● Analyzed customer demand for streaming media during the Covid-19 pandemic using advanced <b>SQL</b> queries and <b>Python</b>, then trained and tested a <b>Random Forest</b> model with <b>Python</b> to identify potential user profiles for conversion, achieving an accuracy of 98%.</li><li>● Determined customer markets and consumption levels based on selected categories, extracted and preprocessed monthly market data of customers using <b>SQL</b> and <b>Python</b>, and predicted revenue trends using a multiple regression model with an accuracy of 85%, providing recommendations for strategic business decisions.</li><li>● 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 <b>SQL</b>, communicated critical insights to the engineering team, and implemented funnel improvements based on findings.</li><li>● Created interactive dashboards with <b>Tableau</b> to monitor key investment metrics for stakeholders, including <b>ROI (Return on Investment)</b>, growth rate, and portfolio balance, and conducted comprehensive root cause analysis of any abnormal trends.</li></ul>	
<b>Chatchup (Zhuiguang) Information Technology Co. Ltd   Co-founder</b> <i>Project Analyst   Customized Analysis for Client Requirements</i>	Zhengzhou, Henan, CHN Jan. 2019-Sept. 2020
<ul style="list-style-type: none"><li>● Analyzed customer requirements to draft an <b>activity diagram</b> and outlined preliminary <b>UML</b> page functionality scenarios. Initiated the development of customized software (Web &amp; App) tailored to customer needs.</li><li>● Employed advanced <b>SQL queries</b> and <b>Python</b> to analyze over 50GB of market data on similar products, identifying customer usage patterns and business requirements.</li><li>● Conducted exploratory data analysis on numerical, categorical, and time-series data using <b>Matplotlib</b> and <b>Seaborn</b>. Constructed multivariate Time Series Clustering model, <b>DTW (Dynamic Time Warping)</b> with Hierarchical clustering, on 120 features to group customers, segment market and described the clustering centroid by <b>DBA (DTW Barycenter Averaging)</b>.</li><li>● Trained market demand models, including <b>Isolation Forest</b> and <b>Robust PCA (Principal Component Analysis)</b>, as well as <b>Linear Models</b> and <b>Random Forests</b>, to establish a system for identifying and responding to customers' potential requirements.</li><li>● Constructed dashboards with <b>Tableau</b> to depict customer usage activities and submitted analytical reports to clients.</li></ul>	