

Lucy Xu

lucyxu521@gmail.com | Chicago, IL | Open to Relocation | 949-332-0094 | linkedin.com/in/luxu66

EDUCATION

Northwestern University , McCormick School of Engineering <i>Master of Science in Machine Learning and Data Science, Minor in Engineering Management</i>	December 2025
Certificate: AWS Cloud Solutions Architect, Google Data Analytics Professional Certificate, JavaScript Algorithms & Data Structure	GPA:3.7/4.0

University of California, San Diego <i>Bachelor of Science in Mathematics and Economics, Minor in Accounting</i>	June 2022
	GPA: 3.7/4.0

PROFESSIONAL EXPERIENCE

PACCAR	Lewisville, TX
<i>Data Analyst Intern</i>	June 2025 – September 2025

- Orchestrated ETL pipelines within Azure DevOps to ingest and process 350K+ ECU fault logs, improving data accessibility and reducing per-vehicle startup time by 22 percent through streamlined data flow and monitoring automation
- Designed time-series forecasting models to identify early fault signals and detect anomalous system behaviors, reducing diagnostic time by 78% and driving data-driven maintenance decisions that improved operational reliability

TransUnion	Chicago, IL
<i>Data Science & Analytics, Industrial Practicum</i>	September 2024 – June 2025

- Created an end-to-end Artificial Intelligence (AI) pipeline to optimize CNN architectures & reduced model risk by 22%
- Leveraged soft-ordering 1D CNN & convolutional autoencoders for feature extraction & learning on large-scale datasets
- Automated PyTorch-based CNN model pipelines and converted tabular data into multi-channel inputs, cutting model development cycle by 43% and reducing decision latency in fraud detection, accelerating incident response across risk teams

PricewaterhouseCoopers, LLP	Los Angeles, CA
<i>Financial Assurance Associate, Full-Time</i>	October 2022 – July 2023

- Awarded PwC's Data-Driven Achievement Award for creating ML pipelines that automated audit exception detection. Applied Isolation Forests, VAEs, and NLP to uncover risk signals, raising detection precision by 25% and reducing close cycles by 38%
- Analyzed 10M+ financial transactions (Python, Tableau) to identify spending anomalies and reconciliation gaps. Engineered features and applied Z-score detection, reducing review workload by 40% & enhancing audit accuracy under PCAOB standards

Monster Energy	Corona, CA
<i>Financial Data Analyst, Full-Time</i>	April 2024 – August 2024

- Implemented A/B testing and regression models in Python (Pandas, SciPy) to test pricing hypotheses and quantify variance drivers, reducing forecast error by 12% and improving model generalization across product lines
- Built risk management models (logistic regression, decision trees) to streamline amortization workflow, reducing time by 23%

Astra Solutions, Inc.	San Diego, CA
<i>Data Science Intern</i>	June 2021 - September 2021

- Engineered REST API pipelines to automate competitor pricing extraction, improving accessibility and refining benchmarks
- Leveraged unsupervised clustering and XGBoost models in Looker to segment high-value users and detect retention drivers, driving targeted engagement initiatives that increased campaign ROI by 32%
- Developed regression and ensemble models with cross-validation and hyperparameter tuning, boosting ROC-AUC by 12% and reducing portfolio risk by \$50K through improved rental pricing predictions

PROJECT EXPERIENCE

Topic-Aware Tweet Sentiment Analysis (RAG & Fine-Tuned LLM)

- Designed a multi-agent RAG pipeline integrating Lang Graph orchestration, Pinecone retrieval, GPT-3.5 summarization, and a LoRA fine-tuned SmoLM2-360M classifier, boosting classification accuracy from 0.319 to 0.8675 on 1M+ tweets
- Developed a real-time trend analysis (<2s latency) via web app, boosting overall model performance by 15-20% over baseline

TECHNICAL SKILLS

Programming Languages: Python (Pandas, TensorFlow, NumPy, PySpark), SQL, Java, Hive, R, Agile, HTML, SAS, Julia, Shell

Technical Skills: Machine Learning, Deep Learning, Generative AI, Data Mining, Databases Management, A/B Testing, Advanced Analysis, Data Visualization, Quantitative Analysis, Big Data Technologies, Project Management, Risk Management, Statistics

Tools & Platforms: Tableau, Power BI, Snowflake, Spark, Dask, Git, Alteryx, Azure DevOps, Salesforce, Oracle, Jira, Looker