TRANG TRUONG

(650) - 267 - 1710 | trangtruong.071105@gmail.com | LinkedIn | Portfolio | GitHub

SUMMARY

Expertise in deep learning, supervised learning, and statistical modeling, using frameworks to build and optimize predictive models. Skilled in NLP (BERTopic, LLMs), developing data pipelines with Python and SQL, and applying critical thinking and problem solving to translate data into actionable business insights. Proven ability to deliver end-to-end ML solutions with strong performance and communicate results effectively to diverse stakeholders.

EDUCATION

Arizona State University ***Aug 2023 - Dec 2026***

Bachelor of Science, Business Data Analytics

- **GPA:** 3.86
- **Coursework:** Business Statistics, Data Structures and Algorithms, Machine Learning, Object-Oriented Programming, Software Engineering and System Design, Social Media Mining

SKILLS

- **Programming Languages:** Python (Pandas, NumPy, Matplotlib, Plotly, Scikit-learn), SQL, C++, HTML, CSS
- **Machine Learning & AI:** Supervised Learning (Regression, Classification, Tree Ensembles), Deep Learning (NLP, BERTopic, LDA, LLMs, PyTorch, TensorFlow, Keras), Generative AI
- **Database & Querying:** SQL, MySQL, NoSQL
- **Frameworks & Tools:** Jupyter Notebook, Visual Studio Code, Git, Docker
- **Cloud & Big Data Platforms:** AWS Lambda, Databricks, Spark
- **Data Visualization:** Matplotlib, Plotly, Excel, Tableau
- **Soft Skills:** Solution/ Technical Design, Problem Solving, Analytical and critical thinking, Communication, Teamwork, Management, Leadership

EXPERIENCE

BullyBlocker ***Jul 2025 - Present***

Undergraduate Researcher

- Optimized cyberbullying detection model by integrating multiple API-based LLMs, achieving a 10-13% F1-score improvement and enabling scoring of harmful conversations.
- Developed explainable NLP systems that identified key reasons behind flagged abuse, reducing false positives by 12% and enhancing trust in moderation outcomes.
- Designed and deployed LLM-powered mediation generation models, advancing ethical AI applications for safer and more constructive online interactions.

School of Mathematical and Statistical Sciences, ASU ***Oct 2024 - Jun 2025***

Office Assistant

- Automated scheduling workflows for students, cutting processing delays by 30% and improving transparency.
- Digitized math placement records, accelerating data access and reducing registration bottlenecks.

PROJECTS

Intelligent Education Chatbot LINK | PyTorch, NLP (BERTopic), LLM (LLaMA 3) ***Jun 2025 - Present***

- Developed and benchmarked time-series models (ARIMA, GARCH, RF, XGBoost, SVR, LSTM) on Tesla OHLCV data, achieving 0.0394 RMSE and 52.1% directional accuracy, surpassing baselines.
- Engineered key financial features (returns, volatility, momentum) to enhance model performance on non-stationary, fat-tailed data.
- Extended findings into a financial research paper, integrating market and sentiment signals for improved predictive accuracy.

Tesla Stock Price LINK | Time Series Forecasting, Feature Engineering, *ML/DL Models* ***Sep 2025 - Present***

- Built and benchmarked time-series models (ARIMA, GARCH, RF, XGBoost, SVR, LSTM) on Tesla OHLCV data, achieving lowest RMSE by 0.0394 and directional accuracy by 52.07%, outperforming naive baselines.
- Engineered financial features (returns, volatility, momentum) to improve model robustness on non-stationary, fat-tailed stock data.
- Extended work toward a financial research paper, incorporating external market and sentiment signals to further enhance predictive accuracy.

Sephora Product Reviews LINK | PyTorch, TensorFlow, NLP (BERTopic), LDA ***Mar 2025 - Aug 2025***

- Developed a production-grade data mining and NLP pipeline analyzing 160,000+ customer reviews using LDA and BERTopic, identifying key dissatisfaction drivers across product categories.
- Engineered a skin-type recommendation model that uses user reviews to map suitable skin types, enhancing product relevance and personalization.
- Applied SMOTE for class balancing, boosting low-rating classification precision by 28%, and proposed product enhancement strategies through customer user research analysis.

Loan Recovery System LINK | Random Forest, Risk Scoring, Behavioral Modeling ***Jun 2025***

- Built a Random Forest model leveraging financial and behavioral features to classify overdue loans by recovery risk and generate individualized risk scores for borrower segmentation.
- Enabled data-driven recovery strategies (legal action, settlements, automated follow-ups) by integrating predictive insights into borrower profiles for optimized decision-making.