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, LDA, 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:** Computer Applications and Information Technology, Math for Business Analysis, Business Statistics, Information Systems Analytics, AI Business, Principles of Programming with C++

SKILLS

- **Programming Languages:** Python (Pandas, NumPy, Matplotlib, Plotly, Scikit-learn), SQL, C++
- **Machine Learning & AI:** Supervised Learning (Regression, Classification), Deep Learning (NLP, BERTopic, LDA, LLMs, PyTorch, TensorFlow, Keras)
- **Database & Querying:** SQL, MySQL, NoSQL
- **Frameworks & Tools:** Jupyter Notebook, Visual Studio Code, Git
- **Data Visualization:** Matplotlib, Plotly, Excel, Tableau

EXPERIENCE

BullyBlocker ***Jul 2025 – Present***

Undergraduate Researcher

- Developed transformer-based NLP models for cyberbullying detection, achieving 10-13% improvement in F1-score over baseline classifiers on noisy social media text.
- Fine-tuned large language models (LLMs) and optimized preprocessing pipelines, reducing false positives by 12% and boosting moderation precision in real-world datasets.
- Contributed to ethical AI research in online safety, supporting publication-quality results and advancing responsible automated abuse detection methods.

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

Office Assistant

- Automated scheduling workflows for 200+ students, reducing advisor processing delays by 30% and enhancing operational transparency.
- Digitized and maintained math placement test records, improving access speed and minimizing administrative bottlenecks during registration periods.

PROJECTS

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

- Developed two NLP-based models for personalized learning: one collecting user preferences to recommend courses, and another interpreting direct commands for fast, intent-driven suggestions.
- Building a course-sequencing recommender system that suggests the next course based on user progression, leveraging behavioral patterns and course content relationships.
- Currently integrating LLaMA3 to power natural conversations and deliver adaptive, context-aware course recommendations in real time.

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 = 0.0394 and directional accuracy = 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, uncovering top 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 | Python, Random Forest, Risk Scoring, Behavioral Modeling ***Jun 2025***

- Developed a Random Forest model using key financial and behavioral features to accurately classify overdue loans by recovery risk.
- Created individualized risk scores to segment borrowers into actionable categories for optimized collection strategies (legal action, settlements, automated follow-ups).
- Integrated predictive insights with borrower profiles to enable targeted, data-driven recovery planning and decision-making.