

TANZIL AHMED

☎ (408) 780-5984 ✉ tannahmed03@gmail.com [in linkedin.com/in/tanzilahmed/](https://www.linkedin.com/in/tanzilahmed/) github.com/tanzilahmed0

Education

San Jose State University

Aug 2022 – May 2026

Bachelor of Science, Data Science

San Jose, CA

Relevant Coursework: Data Structures and Algorithms, Data Visualization, Linear Algebra, Machine Learning, Processing Big Data, Applied Probability and Statistics — **GPA: 3.63**

Professional Experience

Machine Intelligence and Complex Systems Lab

January 2025 - Present

Machine Learning Research Assistant

San Jose, CA

- Engineered **agentic RAG pipelines** using **LangChain**, **LlamaIndex**, and **LLaMa-3b**, enhancing multi-step retrieval and reasoning capabilities.
- Implemented a custom RAG pipeline from scratch using **PyTorch**, **LangChain**, and **LlamaIndex**, optimizing retrieval and response generation.
- Developed a document indexing pipeline with **BM25**, **FAISS**, **Pinecone**, and **Weaviate**, improving search efficiency for unstructured text.

Projects

NBA Game Predictor

Python | Pandas | Scikit-learn

- Developed and implemented a **predictive analytics model** to analyze and predict the winning team of NBA games using **Python**, **Pandas**, and **Scikit-learn** with **64% accuracy**.
- Utilized advanced data preprocessing techniques including **sorting**, **rolling windows**, and **groupby operations** to prepare data for analysis.
- Scraped and parsed box score data using **Playwright for browser automation** and **BeautifulSoup for HTML parsing**, ensuring comprehensive and accurate data collection.

NFL EloTracker

Python | Pandas | Scikit-learn

- Conducted statistical analysis of **50+ years of NFL data** to identify key performance patterns and quarterback impact on team success, utilizing **pandas** and **sk-learn**.
- Engineered **8+ predictive features** from raw game data, including ELO differentials and QB performance metrics, and applied **PCA to retain 95% variance** while reducing dimensionality.
- Developed predictive models using **Random Forest** and **Logistic Regression** to forecast NFL game outcomes, achieving **82% accuracy** through feature engineering and hyperparameter optimization.
- Created interactive data visualizations using **Plotly** and **Seaborn** to analyze historical team performance trends across **2,000+ games**, enabling insights into dynasty formation and playoff success factors.

Music Mind

Python | Scikit-learn | TensorFlow

- Developed and implemented a music genre classification system using ensemble methods (**SVM**, **Random Forest**, **XGBoost**), achieving **85%+ accuracy** through hyperparameter optimization and cross-validation techniques.
- Engineered **50+ audio features** using signal processing libraries (librosa), extracting MFCCs, spectral characteristics, and temporal features from **8,000+ audio samples** to improve model performance.
- Designed and trained a **Convolutional Neural Network (CNN)** using **TensorFlow** for spectrogram image analysis, incorporating data augmentation techniques and achieving **78% validation accuracy** across 10 music genres.

Technical Skills

Languages: Python, Java, SQL, HTML5, CSS, JavaScript, C++

Developer Tools: Visual Studio Code, IntelliJ, Git, Github

Libraries/Frameworks: React, Scikit-learn, PyTorch, NumPy, TensorFlow, Pandas, Plotly, Seaborn, OpenCV, Next.js