Tri Le

triple14@seas.upenn.edu | https://www.linkedin.com/in/tlcsun

Education

University of Pennsylvania — Philadelphia, Pennsylvania
Master of Science in Engineering, Data Science — *2025 – 2027*

Experience

Undergraduate Research Assistant
California State University, Northridge — *Northridge, California*

- Developed and evaluated machine learning Models: Applied various ML and DL models, including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbor (KNN), and Multi-Layer Perceptron (MLP), achieving a 97 % accuracy with MLP and 95 % accuracy with SVM for Parkinson's Disease detection based on voice signal features.
- Enhanced model performance through advanced techniques: Utilized the Synthetic Minority Over-sampling Technique (SMOTE), Feature Selection, and hyperparameter tuning (GridSearchCV) to improve model performance significantly.
- Processed and analyzed 195 voice recordings from 31 patients, and achieved optimal results with a 70:30 train/test split using the UCI Machine Learning Repository dataset. Implemented data preprocessing and visualization techniques to ensure high-quality input for model training and evaluation.

Projects

MLB Stasher | *Python, scikit-learn, TensorFlow, PostgreSQL*

* Collected and processed high-dimensional MLB data from various sources, such as MLB Statcast and advanced tracking systems, using Python and Pandas, handling up to 5GB of data per game
* Implemented machine learning models, including regression models for predicting player statistics (e.g., home runs, batting average) with 92 % accuracy and clustering algorithms for player classification (e.g., identifying different hitting styles), using scikit-learn
* Utilized TensorFlow for pitch classification and biomechanics analysis via player pose estimation; designed a PostgreSQL database to manage resulting datasets (5,000+ concurrent queries)

NBA Grabber | *Python, Flask, React, Git*

* Led end-to-end development of a web application (5,000+ players, 30 teams) for in-depth NBA performance analysis.
* Engineered advanced NBA statistical measures (EPM, RAPM, LEBRON) using BBall Index data across 10+ seasons and Developed a k-Means Clustering model to analyze NBA offensive evolution from 1,000+ games, uncovering key performance trends
* Architected a Flask/SQLAlchemy data pipeline integrating the NBA API for under 500ms latency in delivering up-to-the-minute statistics

UtilMe | *Javascript, React Native, Google Firebase/Cloud, Google Map API, Git*

* Built a user-friendly cross-platform mobile application to connect homeowners with nearby home service providers such as plumbers and electricians, utilizing React Native.
* Implemented Google Maps API integration, enabling users to visualize and access a comprehensive database of local service providers, enhancing the application's search and discovery capabilities.
* Integrated the app with Google Firebase/Cloud Messaging to provide multiple features including user authentication, account management, and real-time communication.

Technical Skills

Languages: Java, Python, SQL (Postgres), JavaScript, HTML/CSS, R
Frameworks: React, Node.js, Flask, FastAPI
Developer Tools: Git, Docker, VS Code, Visual Studio, AWS
Libraries: pandas, NumPy, Matplotlib, PyTorch, MLFlow, Scikit-Learn, TensorFlow