Arnav Raut

Pune, Maharashtra | arnavraut390@gmail.com | GitHub | LinkedIn

EDUCATION

Bachelor of Technology in Computer Science Engineering **CGPA - 8.47** | August '23 - June '27
MIT-ADT University | Loni Kalbhor, Pune

SKILLS

Programming Languages – Python, C++, JavaScript, HTML, CSS, R
AI/ML Frameworks – PyTorch, TensorFlow, Hugging Face Transformers, scikit-learn
Data Science & Tools – Numpy, Pandas, Matplotlib, Seaborn, Selenium, Streamlit, Tableau, Power BI

PROJECTS

Stock Price Prediction using Statistical Model (ARIMA) & Deep Learning Model (LSTM) July '25

- Built and comprehensively compared two independent stock forecasting models **ARIMA**, a statistical model, and **LSTM**, an advanced RNN-based deep learning architecture using historical Tesla (TSLA) stock data sourced via **Yahoo Finance API**.
- Designed a robust and scalable ARIMA pipeline with automatic retraining on each new ticker input, employing fixed (1,1,1) parameters for ARIMA(p,d,q) to maintain temporal stability and ensure quick adaptability across diverse financial datasets.
- Developed and optimized a high-performance LSTM sequential neural network capable of capturing long-term temporal dependencies in stock trends, leveraging its superior ability to model intricate and complex nonlinear financial data patterns.
- Conducted comprehensive and rigorous performance evaluation using standard metrics like **RMSE, MSE, and MAPE,** where the LSTM model achieved a **significantly lower RMSE**, establishing its edge in predictive accuracy over ARIMA.
- Created and deployed interactive, **Streamlit dashboards** for both ARIMA and LSTM, enabling users to interactively visualize **live forecasts**, compare model outputs, and analyze predictive trends in real time for actionable insights.

Sentiment Analysis Using Twitter Dataset June '25

- Developed an end-to-end and automated sentiment classification system leveraging **Logistic Regression** on a standard and public Kaggle Twitter dataset to predict tweet polarity across positive, neutral, and negative categories systematically.
- Implemented an extensive text preprocessing pipeline incorporating URL and mention removal, tokenization, POS tagging, and **WordNet lemmatization**, while selectively retaining hashtags and sentiment-heavy stopwords for contextual integrity.
- Applied and specifically optimized the **TF-IDF Vectorization** technique using a range of unigrams and bigrams across 5,000 feature dimensions to effectively capture linguistic patterns crucial for nuanced sentiment recognition in short-form text.
- Trained the supervised Logistic Regression model by fine-tuning regularization and iteration hyperparameters, achieving approximately **70% test accuracy**, which was then validated through a **classification report** and **confusion matrix** analysis.
- Designed and deployed an interactive **Streamlit** web interface to display predicted sentiment labels along with confidence probabilities and comprehensive bar chart visualizations, offering a highly practical and interpretable user experience.

Game Recommendation Engine using BERT, MIT-ADT University March '25 - May '25

- Built a content-based game recommender system using **BERT**, which was trained on features including title, user reviews, description, genres, tags, and publishers, these features were included in a Kaggle dataset which was used for the project.
- Also fetched game metadata like descriptions, tags, genres, and release dates from **RAWG API**, a community for gamers having a database of all the information related to games, and merged this data with the main dataset taken from Kaggle.
- Pre-processed and combined multiple text columns into one aggregated feature per game for embedding generation.
- Generated semantic similarity scores using **cosine similarity** on BERT vectors to recommend contextually similar games, and created a deployable site for demo using **Streamlit** for users to input a game and receive similar recommendations.
- Evaluated the model with hit ratio, had 50 users in our campus use the engine and accumulated the hit scores, **60% hit ratio**.

ACHIEVEMENTS

Winner - WePitch 1.0 Startup Pitch Event, **MIT-ADT University** April '25

- Proposed a gig-based student marketplace, led ideation and revenue model strategy - involving commission-based structure.

POSITION OF RESPONSIBILITY

Technical Team Member at Microsoft Learn Student Chapter (MLSC), MIT-ADT University April '25 - Present