SneakerBot: NLP-Powered Conversational Assistant for Sneaker Queries

# 1. Problem Definition

The sneaker market is booming, and users often need assistance filtering through brands, prices, sizes, and availability. The goal of this project is to develop an intelligent, NLP-based chatbot called SneakerBot to assist users in navigating sneaker-related queries such as price checks, availability, and recommendations.

This project integrates multiple NLP tasks including:  
- Intent detection  
- Sentiment analysis  
- Named Entity Recognition (NER)  
- Question answering  
- Summarization  
- Translation

# 2. Dataset Selection

- Primary Dataset: StockX Sneaker Price Dataset (CSV)  
 - Source: Kaggle StockX Data Contest  
 - Features: Brand, Shoe Name, Sale Price, Retail Price, Release Date

- Custom Intent Dataset: Manually created small dataset  
 - Includes labeled queries for intents like: greeting, filter\_size, check\_price, availability, recommend

# 3. Methodology and Implementation

3.1 Data Preprocessing:  
- Tokenization and lowercase transformation  
- TF-IDF Vectorization for intent classification

3.2 Model Selection:  
- Intent Detection: Logistic Regression trained on TF-IDF features  
- Sentiment Analysis: Pre-trained Transformer pipeline from HuggingFace  
- NER: spaCy's small English model (en\_core\_web\_sm)  
- QA System: Transformer-based QA pipeline with static context  
- Summarization: HuggingFace summarization pipeline  
- Translation: Googletrans for multilingual support

3.3 Training:  
- Model trained on small custom dataset  
- Evaluation using Accuracy, F1-score, Precision, Recall

3.4 Saving & Loading Models:  
- Used joblib to save intent\_model.pkl and vectorizer.pkl

# 4. Deployment

- Developed frontend using Streamlit  
- Chat interface accepts user input and displays responses  
- Integrates all NLP features dynamically based on intent  
- Hosted via Streamlit Community Cloud

# 5. Evaluation & Findings

Intent Detection Model Performance:  
- Accuracy: 100% on small test set  
- F1-score (macro): 1.0

Sentiment & NER: Working as expected in English inputs, capable of identifying brands, models, prices

Summarization & QA: Generates concise summaries and accurate answers from static sneaker context

Translation: Successfully translates input from Hindi, Tamil, Telugu into English

# 6. Conclusion

SneakerBot demonstrates the feasibility of combining rule-based and ML/NLP-based pipelines into a single intelligent assistant tailored for the sneaker domain. The multi-task setup allows for a fluid and context-aware user experience.

# 7. Future Enhancements

- Add real-time API integration with sneaker stores (e.g., StockX, GOAT)  
- Expand intent dataset with more samples  
- Use deep learning (LSTM/BERT) for more robust intent classification  
- Add memory-based chat context for multi-turn conversation support