Secure AI Powered Customer Service Chatbot for E-commerce

Contributors

- Harshit Pant (CS20BTECH11021)
- Mahin Bansal (CS20BTECH11034)
- Satpute Tukaram Aniket (CS20BTECH11056)

Overview

- The project has three models that were trained on synthetic datasets.
- Github Repository
- The models are:
  - Intent Classification Model
  - Named Entity Recognition Model
  - AI Security Model [Login Anomaly Detection] Using Isolation Forest
- To train the models on a fresh clone it is advised to run the `backend/chatbot/train.ipynb` notebook.

NLP Model Architecture

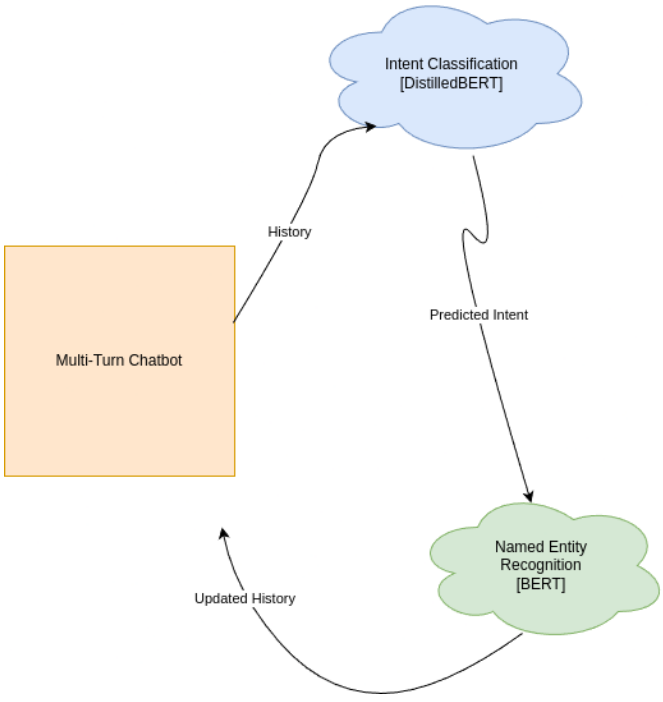

Figure 1: NLP Model Architecture

- The model aims to emulate a customer service chatbot for an e-commerce platform.
- The model is a multi-turn chatbot that can handle multiple intents and entities.

- It maintains a history of the conversation to provide context to the user and uses the following input pattern to predict next intent and entities.

`[INT] <previous intent> [BOT] <bot response if not first query> [USR] <user query>`

- The architecture splits the model into two parts:
  - Intent Classification Model
  - Named Entity Recognition Model
  - The models are trained on synthetic datasets.
- The synthetic dataset covers multiple intents such as:
  - track_order
    - Entities: order_id
  - cancel_order
    - Entities: order_id, reason, affirmation(for confirmation)
  - list_orders:
    - Entities: count(count of items from last order), end_date, start_date

Intent Classification Model

- The intent classifier model is a fine-tuned DistilBERT model and is trained on the dataset present in `chatbot/data/bert_input.csv`.

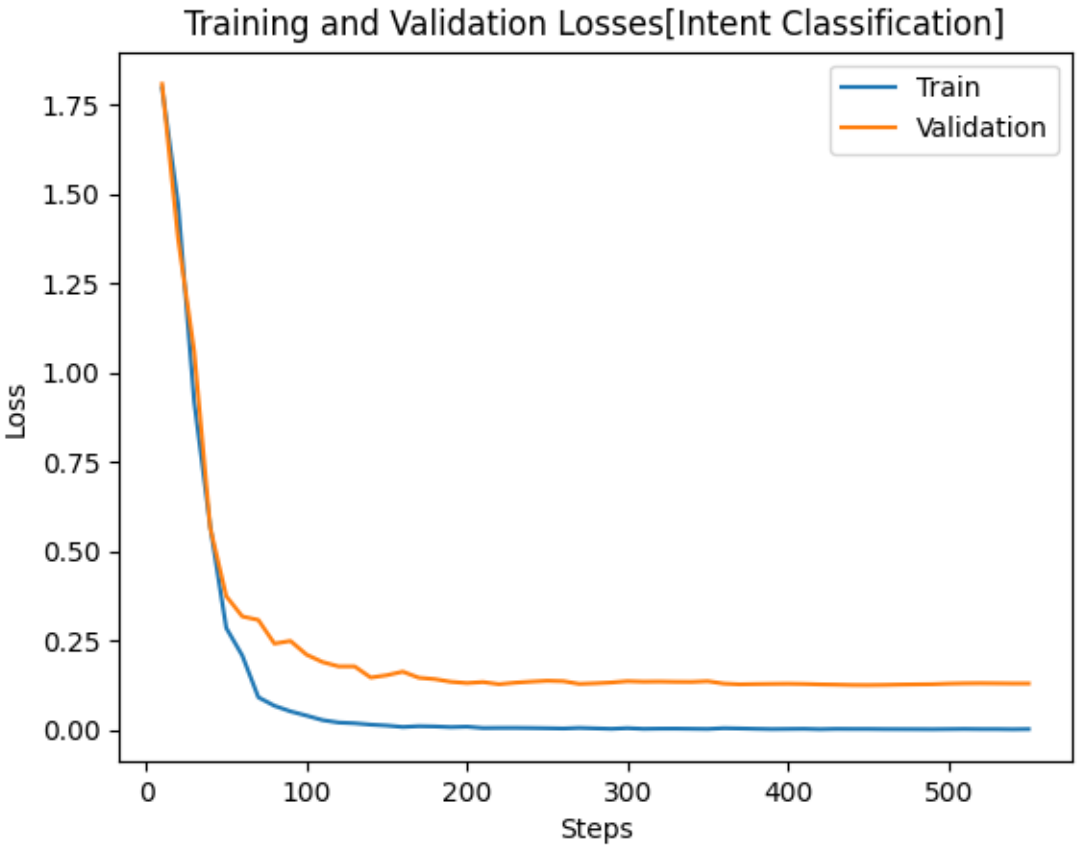

Figure 2: Loss

Named Entity Recognition Model

- The model is a fine-tuned BERT model and is trained on the dataset present in `chatbot/data/ner_data.csv`.
- The data is labelled in the IOB format which is the typical format for NER tasks.

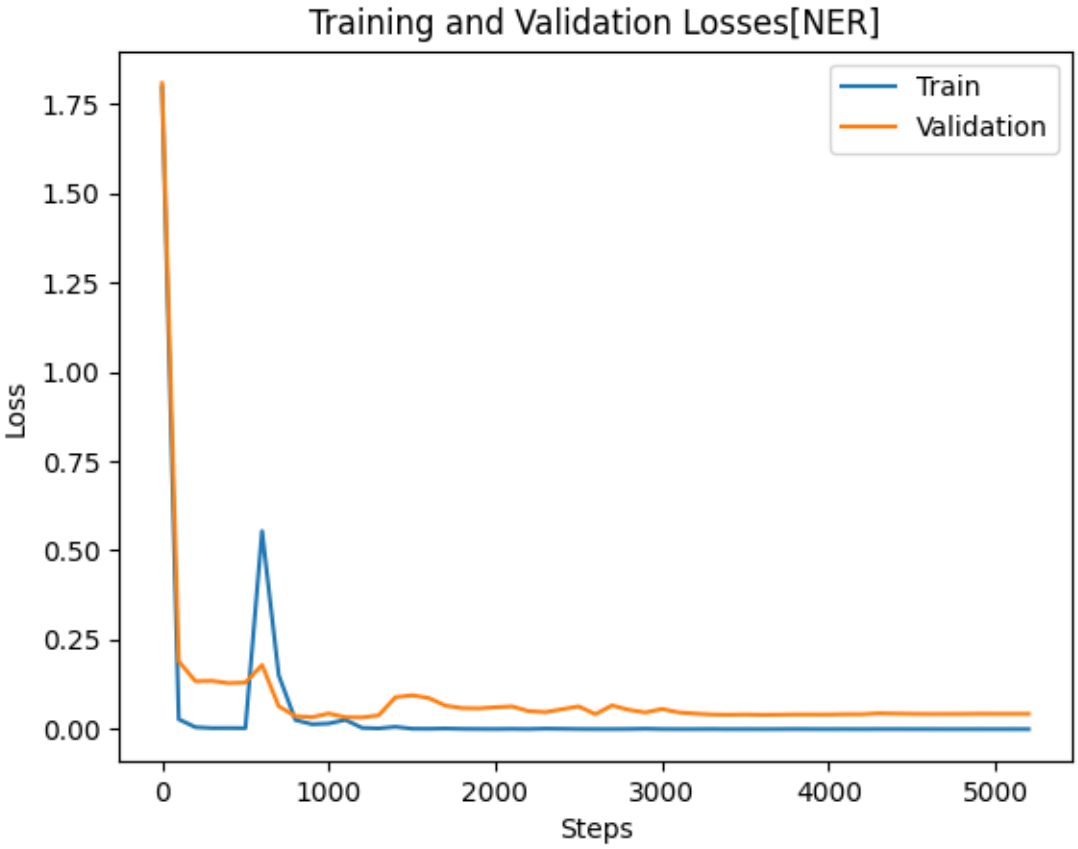

Figure 3: Loss

AI Security Models

- The AI Security Models used in this project are:
- Anomaly Detection in User Login Data:
- The model uses additional feature engineered vectors such as the number of failed login attempts in a certain time frame which is exponentially weighted.
- The model uses the Isolation Forest algorithm to detect anomalies in the login data.
- Anomalous IP addresses and user IDs are flagged as potential security threats.
- The model is present in `backend/AIModels/loginModel/final_model.py`.
- The model is trained on the dataset generated from the log files which are directly generated from the login screen of the e-commerce platform.
- The data is logged into the `backend/logs/login.json` file.