

Task: Evaluation of AI Bot Conversation Quality

Objective: Your task is to create and execute a system that rates AI bot responses to lead inquiries on a scale of 1 to 10. This evaluation should assess how well the bot addresses lead questions and fulfills their needs, as well as whether the bot adheres to predefined instructions during conversations.

Bot Instructions:

A lead management bot connects with new leads who haven't had prior conversations, aiming to book them by suggesting available intro classes in groups of three and answering their questions. The bot's final goal is a successful booking. Once interest in booking any of the suggested classes is detected, the bot asks the lead to confirm the booking. In the case of a positive reply, the bot proceeds to book the lead and sends a success message to the lead.

The bot has specific instructions that should be followed. The bot should respond with the message "One of our representatives will be in touch with you shortly to assist you further" whenever any of the following is detected during a conversation:

- The lead wants to bring someone else (e.g., a friend, daughter, etc.).
- The lead asks about gift cards, payments, discounts, or costs.
- The lead wants to use 1Pass (onepass) credits.

Dataset: You have a dataset that includes conversations between leads and an AI bot. Each entry in the dataset consists of lead and bot messages, identified by numbers 2 (lead) and 1(bot) in the 'source' column, respectively, along with a lead_id and the message content.

Output: Jupiter Notebook with your codes and a CSV file as a result, where each row contains a lead ID and conversation score.