

Senior Machine Learning Engineer Assignment

Time budget

8 hours of focused work (you may spend more, but we will judge as if you had one working day)

Submission guideline

Upload your solution to the provided Google Drive link.

Background

SumUp's Operations team manages a customer support (CS) chatbot that handles a high volume of inquiries. While the chatbot resolves many issues independently, there are situations where escalation to a human agent is necessary to ensure efficient and satisfactory support. For instance, if the bot cannot resolve a problem or if a user shows signs of frustration. Escalation should occur at the right moment - not too early, which wastes agent resources, and not too late, which risks harming customer satisfaction.

You are given a small corpus of **multi-turn chat transcripts** annotated with *is_escalation_needed* labels and the reasons why such escalations are granted. Each conversation consists of user and bot messages. Labels indicate whether an escalation is warranted at that turn.

Your task is to design and prototype a system that **predicts when to escalate** during an ongoing conversation.

Task

Design and prototype an ML/AI system which recognises the need of an escalation during an ongoing conversation.

During your solution, make sure to consider both potential product and technical requirements. Which technical challenges would you anticipate? Document your assumptions, any open questions, and the challenges you foresee when bringing your solution to production.

Apply good engineering practices throughout. Write clean, maintainable code with a strong focus on readability. You may use AI tools to assist you in creating your solution, but ensure you fully understand the output and feel confident presenting and explaining the code.

If your solution relies on AI services (e.g., LLMs), we recommend choosing a model that offers a free-tier quota (such as Google Gemini).

Data

You will receive labelled data (an example schema is shown below).

Example

JSON

```
{
  "conversation_id": "c001",
  "conversation_history": [
    {
      "role": "bot",
      "message": "Hi, how can I help you?"
    },
    {
      "role": "user",
      "message": "where is my money?"
    },
    {
      "role": "bot",
      "message": "Could you please give me more information about the issue of your money?"
    },
    {
      "role": "user",
      "message": "I already explained many times to you! I tried to contact support but no one answered. I need an agent now"
    }
  ],
  "is_escalation_needed": true,
  "reasoning": "The user has explained that they tried to contact Customer Support many times without success. They are frustrated and requesting agent support. Therefore, escalation is needed in this case."
}
```

[Full dataset.](#)

Deliverables

Please send us the code and a short readme how to execute it. Provide a simple CLI interface to simulate a conversation and test your solution. We do not assess the chatbot's detailed responses to user queries; our evaluation focuses solely on its escalation behavior.

Document your findings and assumptions in a way that is clear and easy for us to follow.

Evaluation

For us, it is important to understand not only the solution itself but also how you approached it. We will look at the clarity of your design decisions, the quality of your code, and how well you document your solution.