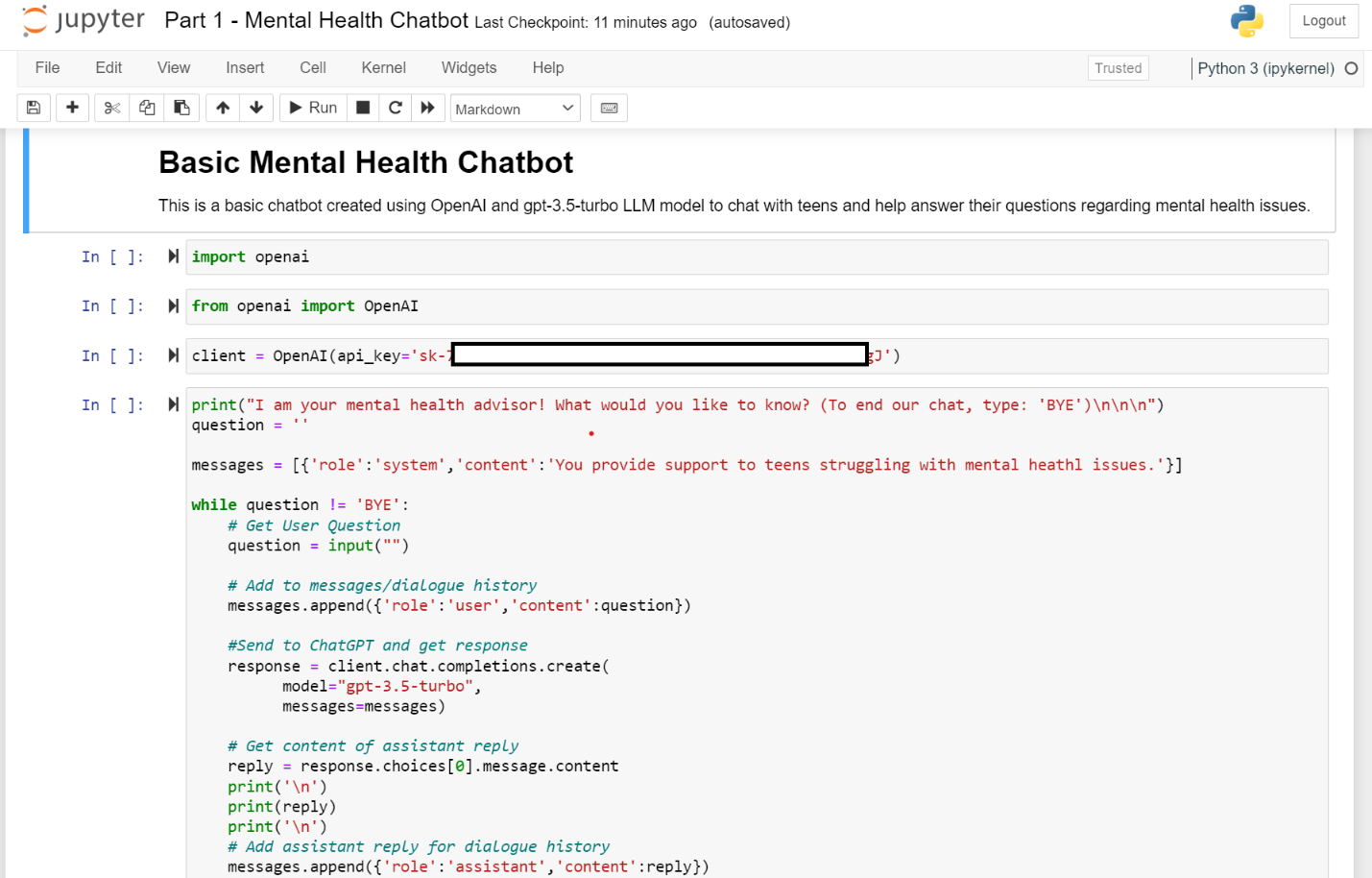
# Mental Health Chatbot and AI Assistant to support mental health advisors

The researchers decided to build an AI assistant to support the teen talk app mental health advisors. This AI assistant will be able to scan the incoming messages from the teens who need support and detect their intent. The virtual assistant would then be able to provide helpful tips and prompts to the teen talk advisor regarding how to respond based on the intent identification. The algorithm would also be able to identify serious situations that need escalation such an intent to self-harm and suicide and help the teen talk advisor regarding how to deal with the issue and escalate the call and request support from the supervisor. This AI assistant can chat directly with the teens, however, there are several risks to AI such as bias, hallucination, security & privacy risks and legal liability issues due to which we decided to limit the functionality to a virtual assistant that can support the health advisors rather than having the chatbot directly interact with the teens.

We created the project in the following steps:

### Part 1 - Learn how to use LLMs and create a basic mental health chatbot to support teens struggling with mental health issues using Open AI and “gpt-3.5-turbo” LLM model



#### Examples of conversation using this chatbot:

A close-up of a letter

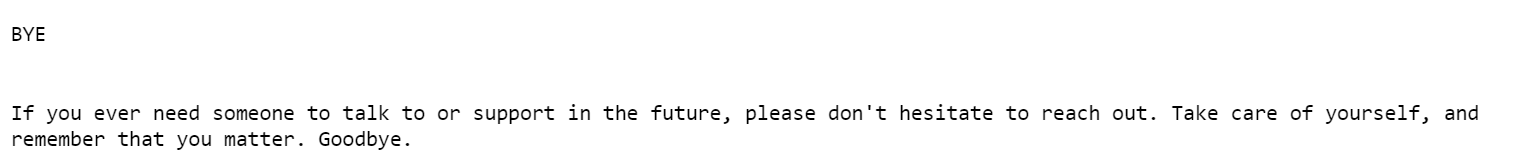
Description automatically generated

A close-up of a document

Description automatically generated

A close-up of a message

Description automatically generated



Note – This is a basic chatbot leveraging Chat GPT LLM. It is not advisable to use this chatbot for actual live use as the GPT model may be biased and hallucinate. The model needs to be supplemented with the organization’s resource files for knowledge base and to accurately respond to user’s questions. Work in Progress.

### Part 2 - Create the intent Classification Model

Researchers used the following data file and code from Kaggle.com –

kaggle/input/mental-health-conversational-data/intents.json

Researchers reused the code from Kaggle.com to build the machine learning model for Teen Talk App.

This dataset is a set of past conversations related to mental health. It consists of tags, patterns and responses. Patterns are the messages that are sent by the users who need mental health support, and the responses are the replies sent by the health care advisor to answer the messages (patterns) received from the users. Tags are the “intents” behind the message sent by the users. For e.g. if the user sends a message such as - I am lonely or sad or empty etc., then the tag (intent) would be classified as “sad”

A close-up of a message

Description automatically generated

**Steps**

1. Import the “intent.json” file downloaded from Kaggle.com and create a pandas data frame in Python Jupyter notebook.

A screenshot of a computer

Description automatically generated

1. Create a dictionary using the data frame

A screenshot of a computer program

Description automatically generated

1. Data Visualization

The data consists of the following tags (intents) which include greetings (good morning, afternoon, bye etc.), emotional state (depressed, suicide, anxious etc.) and mental health facts.

Facts are intentions where the users ask specific question about mental health such as – “What is mental health?”, “What is depression?” etc.

A white background with red text

Description automatically generated

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer screen

Description automatically generated

1. Create a Machine Learning model that can be trained using this dataset and predict users’ intent for future chats. The text data was vectorized using the TF-IDF vectorization method.

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

A graph with different colored lines

Description automatically generated

1. Deploy the model and teach it how to generate responses for each intent that the ML model identifies.

A screenshot of a computer program

Description automatically generated

This is a basic code that identifies how to respond to a user when the intent is determined to be greeting, farewell, question or suicide. The code can be extended to include responses for all the intents in the model.