

Guidelines

- This assessment must be completed individually. Do not seek assistance from others or use unauthorised resources.
- Ensure that all work submitted is your own.
- Solutions generated using ChatGPT or similar AI tools will automatically be rejected.

Prerequisites:

- This test requires you to have an OpenAI API key, which you get from your OpenAI account. If you already have an account, please skip the account setup steps and proceed directly to the assessment. Otherwise, follow the instructions below to set up your account.

To enrol for an OpenAI account, follow these steps:

- **Step 1:** Go to OpenAI's website.
<https://platform.openai.com/signup>
- **Step 2:** You can sign up using your email address, Google account, or Microsoft account.
- **Step 3:** Complete your profile by providing the necessary details.
- **Step 4:** Once your account is set up, you can navigate to the API section to get started with using OpenAI's services.
- **Credit and Usage**

New users typically receive **\$5 worth of free** credits upon signing up. You can check your usage limits by navigating to the left-hand side menu, then selecting Settings > Billing.

Please note, if you already have an account associated with the same email on the same device, once the credit limit is exhausted, it cannot be recovered. The system checks the device, and if it has already availed the free account limit, you will not be able to receive the free credits again on the same device, even with different email IDs.

Assessment:

About the dataset:

In this assessment, you will work on sentiment analysis of Reddit comments **using the OpenAI API**. The data consists of comments collected from eight different subreddits, with 50 comments from each subreddit, totaling 400 comments.

Note:

- Please use **Python exclusively** for this assessment. We want to see your prowess in it.
- Ensure that you use a model version of 3.5 or beyond to utilise the system and prompt features provided by OpenAI.

Process:

- Download and store the CSV file enclosed in the google drive folder.
- The dataset has columns for subreddit, post, comment, and a placeholder for the sentiment score.

Sentiment Scoring Method:

The sentiment scoring method should assign a score to each comment on a continuous scale from -1 to 1, where:

- -1 indicates the most negative sentiment
- 0 indicates a neutral sentiment
- 1 indicates the most positive sentiment

The approach depends on your understanding of AI, but we expect you to do some exploratory data analysis before starting the AI process. Create neat functions that iterate through each row of the dataset, produce a score, and store it in the “Score” column which is currently blank and given as a placeholder.

Submission:

At the end, the sentiment scores should be added to the original dataset. Store the results in a CSV file and submit your file along with the code file.

Please ensure that the code file is clear and easy to understand. Additionally, providing a text explanation for each step you take and why you take it will help us understand your familiarity with OpenAI.

Best of Luck! 😊