

# Task

## Business Intelligence

### Description

Apply Text Analysis on Restaurant\_Reviews dataset to predict if the customer will like or dislike the restaurant based on his review.

The dataset contains 2221 rows  $\times$  2 columns (Review, Liked). Use this dataset to build a prediction model that should accurately classify the customers' reviews if they will like or dislike the restaurant.

- Apply the most appropriate preprocessing steps (Tokenization, stemming, lemmatization, etc.)
- Apply Feature Extraction. (Bag of words, word embedding)
- Apply the model (select the classifier that is suitable for this data).
- Evaluate the selected model (Accuracy, F1-score, Precision, Recall).
- Use Python and needed libraries like (nltk, pandas, sklearn).

The [Dataset file](#) is attached to this file.

### Teams:

- Each team **must** consists of **at least 5** members and **maximum 6**.
- The team members of the **project** will be **the same team members** of the **task**.

### Deadline & Delivery:

- The task delivery will be in **week 6** that started on 16<sup>th</sup> of March 2024.
- All the team members should deliver and **discuss** the task with the assigned TA.
- Each team should come **before** the discussion **at least 15 minutes**.
- There will be oral questions about the labs.

The registration form's link is:

<https://forms.gle/rviCMGJwbwvdTEBX9>

The **deadline** of the registration is **Saturday 9/3/2024 at 11:59 PM**.