**TEAM NAME:**

Egyptian Rackers

**TEAM MEMBERS:**

1. Medhat Elhady
2. Abdelrahman Attia
3. Abdelrahman Rezkallah

**PROJECT NAME:**

Arabic Sentiment Analysis

**PROBLEM:**

the need to address the challenge of understanding and analyzing the sentiment or opinion expressed in large volumes of text data. Prior to the advent of sentiment analysis, extracting subjective information from textual sources in an automated and efficient manner was a complex and time-consuming task.

Businesses find that it’s hard to effectively understanding customer sentiment, monitoring brand reputation, or gaining insights from public opinion.

There is no support for arabic language so current models can’t understand text written in arabic.

**DATASET:**

The dataset combines reviews from hotels, books, movies, products and a few airlines. It has three classes (Mixed, Negative and Positive). Most were mapped from reviewers' ratings with 3 being mixed, above 3 positive and below 3 negative. Each row has a label and text separated by a tab (tsv). Text (reviews) were cleaned by removing Arabic diacritics and non-Arabic characters. The dataset has no duplicate reviews.

**SOLUTION:**

The solution is a web application that is designed to receive comments, feedback, or reviews and classify them as either positive or negative based on their content.

**ML APPROACH:**

Describe the types of model you are considering, the type of feature engineering you are planning to perform, and what kind of data pipeline you're planning to build if any.

Describe the simplest solution you can think of (that could be your baseline).

**LEARNING GOAL:**

1. Learn how to load a pre-trained BERT model from TensorFlow Hub
2. Learn how to build your own model by combining with a classifier
3. Learn how to train a BERT model by fine-tuning
4. Learn how to save your trained model and use it
5. Learn how to evaluate a text classification model
6. Create a streamlit application that call the endpoint

**REFERENCES:**

* [distilbert | Kaggle](https://www.kaggle.com/models/jeongukjae/distilbert/frameworks/tensorFlow2/variations/multi-cased-l-6-h-768-a-12)
* [asl-ml-immersion.git](https://github.com/GoogleCloudPlatform/asl-ml-immersion.git)
* [tensorflow.org/](https://www.tensorflow.org/)

**TOOLS:**

* [**github.com/medhat-elhady/Arabic-Sentiment-Analysis.git**](https://github.com/medhat-elhady/Arabic-Sentiment-Analysis.git)
* [**notebooks.googleusercontent.com/lab/tree/arabic-char-classifier**](https://52ea28689268c9ac-dot-us-central1.notebooks.googleusercontent.com/lab/tree/arabic-char-classifier)