Team Members:

Hazim Bukhari

Amena AlSamah

Rashed AlHilal

NLP Project Proposal

Business Problem:

The Saudi Arabian government is thinking of creating a suicide hotline and it wants to know what are the most frequents words/patterns in Social media posts by suicidal victims to identify them and provide the needed help to them before they harm themselves.

Data:

Data is collected from r/SuicideWatch, r/depression and downloaded from Kaggle, it contains Text and the outcome of the writer, Suicide or non-Suicide.

Number of observation: ~230k

Tools:

Tools: Python, Jupyter Notebook.

Libraries: Pandas, SeaBorn, Numpy, Sci-Kit Learn, Keras, Tensorflow.