

Natural Language Processing (NLP) Project

BY

- Rahaf Alqahtani
- Dimah Albunayyih

Problem Description

With the spread of social media and the ease of spreading news, there is a lot of fake news, which may cause many problems. So by using Natural language processing (NLP), classification, clustering and different techniques to provide an information that help to know whether the news is true or not.

Data Description

The data that will be used in this project has been taken from kaggle website (<https://www.kaggle.com>). It includes data such as: (Text, Subject, Date, etc.) for each news. It includes more than 44K columns and 4 features, below is the description of each one:

Features	Description
Title	The news title
Text	News content
Subject	What is the subject of this news? Is it a sports ,politics etc
Date	The date the news was published

Tools

- Technologies: Python, Jupyter notebook.
- Libraries: NumPy, Pandas, Matplotlib, Seaborn, Sklearn, NLTK and gensim.