ARABIC NEWS ARTICLES

PROJECT PROPOSAL

Proposed by:
Alanoud Alosaimi, Mashael Alfehaid, Rana Alzahrani

INTRODUCTION

In this project, we have a large collection of Arabic news articles which is a great resource that contains natural language text from various newspapers and that can be used in different tasks such as Text Classification and Word Embedding. Our goal is to analyze a series of these articles to identify the domain that each article belongs to.

DATA DESCRIPTION

Over 30K Arabic newspaper articles have been gathered from several online Arabic newspapers.

TOOLS

Here the tools that will be initially uses in this project:

- Software platform: Jupyter Notebook
- Programming language: Python
- Includes libraries:
    - Pandas
    - Numpy
    - Seaborn
    - Matplotlib
    - Sklearn
    - NLTK
    - re
    - Scipy