

NLP News Company

Abstract:

A News company want to know what news topics people are interested in and prefer, which helps the company attracting largest number of people.

Design:

This project is one of the T5 Data Science Boot Camp requirements. the data provided by [Kaggle](#).

Data Description:

The data will be used in this project is provided by [Kaggle](#). it contains of 32000 rows and 2 columns.

Columns contains:

1. headline: contains headlines from different news sites
2. clickbait: which 1 represents it is clickbait and 0 represents non-clickbait

What we gonna use:

1. Topic Modeling: To find the topic of the text and create new column topic
2. Nlb Classification: depending on the topics from topic modeling, classify it to know what type of news are clicked the most and predict what is the type of news depending on headline news

Tools:

- Numpy and Pandas for data manipulation
- Matplotlib and Seaborn and WordCloud for visualise data
- Sklearn
- NLTK
- Keras

Communication:

The slides are provided in [presntation](#) , and proposal provided at [ReadMe](#) file of the project in gitgub.