



NLP Project



News Category Classification

Group -

Team Members



Amelia Morencia Irena

2602115122

Carissa Fabellina

2602071120

Karen Regina Susanto

2602073403

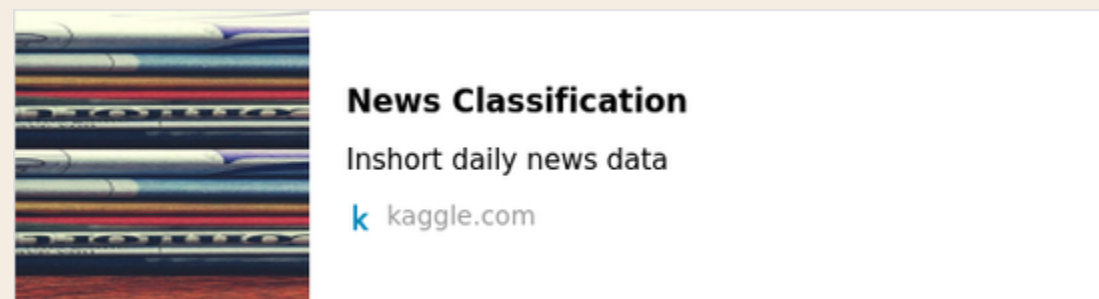
INTRODUCTION

Pada era digital ini kita semakin dipermudah dalam mengakses informasi melalui teknologi secara digital. Dengan banyaknya artikel berita yang dipublikasikan mempersulit proses organisasi / sortir terhadap konten-konten tersebut berdasarkan kategorinya. Maka dari itu untuk project Natural Language Processing ini kami mengimplementasikan **News Category Classification** dengan tujuan untuk meningkatkan efisiensi dan mempercepat proses organisasi konten berita berdasarkan kategorinya masing-masing.

DATASET USED

Untuk Project News Category Classification ini kami menggunakan dataset “**News Classification**” yang dibuat oleh **Kishan Yadav** melalui kaggle yang dapat diakses melalui link berikut

<https://www.kaggle.com/datasets/kishanyadav/inshort-news/data>



DATASET USED

	news_headline	news_article	news_category
0	50-year-old prof	DeepMind's AI s	technology
1	Microsoft Team	Microsoft Team	technology
2	Hope US won't e	China, in respor	technology
3	Global smartph	The global smar	technology
4	EU hoping Bider	The European U	technology
5	Parliamentary p	The members of	technology
6	Google, Facebo	Google CEO Sur	technology
7	Facebook could	Facebook could	technology
8	Canada plans di	Canada's Finan	technology
9	Twitter's 'Audio	Twitter's "Audio	technology
10	US FCC Chairm	US Federal Com	technology
11	Tokyo Stock Exc	Tokyo Stock Exc	technology
12	Bitcoin hits new	Cryptocurrency	technology
13	Ant Group's IPO	Jack Ma-owned	technology
14	Italy fines Apple	Italy's antitrust	technology
15	Theme park with	Universal Studic	technology
16	General Motors	General Motors	technology
17	Facebook buys	Facebook on Mc	technology
18	Ex-Tesla factory	A former Tesla f	technology
19	Global 5G subsc	Swedish telecor	technology
20	Bitcoin is gold 2	Cryptocurrency	technology
21	Trump loses 13	President Donal	technology
22	Biden urged by	US President-el	technology
23	Samsung may d	South Korean te	technology
24	Instagram 'Live	Instagram on Tu	technology
25	Anushka does sl	Anushka Sharm	sports

Dataset memiliki 3 kolom

- news_headline
- news_article
- news_category

Dataset memiliki 7 kategori berita

- technology
- sports
- politics
- entertainment
- world
- automobile
- science

Total data adalah **9051**



IMPLEMENTATION



Hasil implementasi project News Category Classification kami dapat diakses melalui link Google Colaboratory sebagai berikut

https://colab.research.google.com/drive/190yFNj0F8q_16hPatDk6rczj8Q0RVPXP?usp=sharing



Google Colab

 google.com

 Thank You 

Group - News Category Classification