News Summarization using LLM

In this work, we generate news summary using Large Language Model and test whether the generated summary is effective in comparison to original one.

Dataset: We use BBC new dataset as obtained from Kaggle (https://www.kaggle.com/datasets/pariza/bbc-news-summary/). It contains news articles from different categories namely business, entertainment, politics, sport and tech. This dataset also contains summary. 10 news from business category have been used.

Model: We use GPT-3.5 Turbo model for applying Large Language Model on the dataset for summary generation. The settings used are as follows.

model="gpt-3.5-turbo",
messages=messages,
temperature=0.5,
max_tokens=1000,
top_p=1.0,
frequency_penalty=0.0,
presence_penalty=0.0

Results: The summaries generated have been compared with original summaries using ROGUE score.

Summary for News at Serial Number	rouge1	rouge2	rougeL	rougeLsum
1	0.03	0.0	0.03	0.03
2	0.0375	0.0	0.0375	0.0375
3	0.035	0.0	0.035	0.035
4	0.0425	0.0	0.0425	0.0425
5	0.04	0.0	0.04	0.04
6	0.0475	0.0	0.0475	0.0475
7	0.055	0.0	0.055	0.055
8	0.055	0.0	0.055	0.055
9	0.065	0.0	0.065	0.065
10	0.06	0.0	0.06	0.06