Title: Instagram Reach Analysis - Travel

Abstract

This project aims to analyze the Instagram's media content specifically the content related to travelling vlogs and reels. It will focus on user engagement with the media content and analyze trends and patterns including the number of likes, types of comments, the post captions and the visual content. The problem being addressed here is the lack of systematic analysis of travel related content and its user engagement. Even with the vast amount of data being generated through these posts, there is no proper tool to analyze it and explore the hidden patterns in the posted media.

Instagram's Basic Display API can be utilized to collect the required data from travel vlogs and store it in proper format. **Streamlit, FastAPI** are used for frontend and backend in python. Using **Regex, TF-IDF, Word Embeddings, Clustering,** the data from the captions and hashtags will be classified into common subjects or topics. NLP techniques like **Topic modeling** will also be used for better accuracy. A **Linear Regression model** will be employed to measure the impact of the user metrics such as likes, shares, number of comments, and the total reach of the post. **Sentiment analysis** of the comments will reveal how the userbase percieves the posts, giving certain destinations or posts, positive or negative feedback. **T5 Transformer model** will be used as a pretrained model to summarize and analyze the post content, finding the user interaction patterns and major themes.

The project will help us learn more about the content creators, identify their reach and analyze the engagement of their userbase with thier content, and for Instagram to recognize the most trending content on their platform and find out the actual cause for it's trend.