

PROJECT 3

Milestone 1- Rojan Khatri

Topic

The project, 'Optimizing Community Engagement- Advanced Analytics of Facebook Group Posts', aims to deploy topic modeling techniques on Facebook group posts to dissect and understand the underlying community usage patterns. The project seeks to streamline communication processes, extract emerging trends, and craft innovative strategies from the group posts amongst community members to utilize group functionality and community engagement by analyzing these patterns.

Business Problem

The core challenges this project addresses center on the Facebook group's pivotal role as the primary communication channel for our community. Given the community's active growth—from 400 houses currently to a projected 700 upon completion—it becomes crucial to manage the increasing volume of posts and discussions efficiently. This initiative seeks to empower group admins and moderators with enhanced capabilities for managing, tracking, and investigating posts. Such support aims to streamline and improve the moderation process significantly. By proactively responding to the evolving needs and interests of the community, the project ensures that the group continues to serve as an invaluable resource, adeptly supporting both current and future residents.

Datasets

To tackle the outlined business problem, this project will leverage a dataset of approximately 3,500 posts from the community's Facebook group from March 2022 till the present. This dataset will focus on the textual content of the posts and associated engagement metrics like likes, comments, and shares, meticulously excluding any personal or directly identifiable information to uphold privacy standards.

Methods

The methodological framework of the project includes several key steps. Initially, Apify will be used to automate the collection of posts and relevant metadata from the Facebook group, with a strict focus on non-personal data. Subsequent data preprocessing will ensure the removal of potential personal identifiers and prepare the dataset for in-depth analysis. The project will then apply Natural Language Processing (NLP) techniques, specifically Latent Dirichlet Allocation (LDA), to identify the main discussion topics within the group's posts. An analysis of the evolution of these topics over time will further provide insights into changing community interests and engagement patterns.

Ethical Considerations

Ethical considerations form the backbone of this project. Prior to data collection, necessary permissions will be secured from the group's administrators and moderators, aligning the project with community guidelines and Facebook's data use policies. The project commits to the highest standards of privacy by ensuring data anonymization and conducting analysis on an aggregate level, thereby eliminating the use of personal or identifiable information.

Challenges/Issues

The project anticipates several challenges, one of the main ones being preserving privacy. Data extraction and analysis will necessitate advanced data handling and anonymization techniques to safeguard user privacy. Additionally, the project must navigate the unstructured nature of social media content, which requires filtering out irrelevant or low-quality data to maintain analysis integrity. The subjective nature of topic modeling results also poses a challenge, requiring careful validation to ensure that the identified topics accurately reflect community discussions. Finally, potential technical limitations of using *Apify* for data extraction, including compliance with Facebook's API limitations and ensuring reliable data collection, will be critical issues to address.

References:

Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the Social Sciences: Opportunities, Challenges, Ethical Considerations, and Practical Guidelines.

https://www.researchgate.net/publication/281621604_Facebook_as_a_Research_Tool_for_the_Social_Sciences

Jiang, J., Wilson, C., Wang, X., Sha, W., Huang, P., Dai, Y., & Zhao, B. Y. (2013). Understanding latent interactions in online social networks.

<https://people.cs.uchicago.edu/~ravenben/publications/pdf/latent-tweb13.pdf>