

CSD402 - Project Proposal

Pranav Dahiya (1710110249)

Motivation

Today there are numerous digital means to access news. On one hand, there are the tried and tested RSS readers, which allow you to manually curate your news sources so that the news the user sees is not dictated by some algorithm. However, with this method, the user is bombarded with news articles. Even with a few news sources selected (3 – 5), one can expect to receive 100s of news articles every day, which is not ideal. On the other hand, there are news readers like Google News or InShorts, which aim to simplify this experience for the user. However, in this case, the user loses all control. Google News simply gives the user news that is popular in your geographical area and there is no way of ensuring that the publishing organization is credible. For example, since I live in an area with mostly conservative people who are supporters of BJP, that's the audience that the news articles that I see when I open my Google News feed. Finally, there is also social media, but with that the user has even less control. Therefore, there is a gap here, for a news reader that gives the user control, but still organizes incoming articles in a way that is not overwhelming.

Idea

I propose to build an advanced RSS news reader where the user can select which news sources are deemed credible and useful, but incoming articles are organized in a better way. I plan to use NLP to categorize news articles based on events, and then organize articles in each category based on an importance score. When the user opens the reader, the app will allow him/her to select events, and then show articles for that particular event in order of importance.

Features

Front-End Features:

- An event browser
- An article browser (inside each event)
- An article viewer
- A (RSS) feed selector

Back-End features:

- A database of news feeds and news articles (Updates regularly with new articles and categories)
- An API to service requests from clients (Every time a client requests articles, the database is queried, and the requested categories/articles are returned)