B.E., V Semester Information Technology

Mini Project Abstract

On

Content Aggregator

Content Aggregator is any tool that gathers, collects and accumulates content pieces such as articles, social media posts, videos, images, updates and press releases from different types of media outlets and then displays all that information on one single page. However, most of this information is redundant, insignificant, and may not convey the intended meaning.

So we are developing a Web application which extracts and aggregates all the information of News, Science articles and History articles from a few popular websites and displays them in a summarized way, so that users won't struggle to surf all the different links and come to a conclusion. We can do this by Natural language processing. Broadly, there are two approaches to summarizing texts in NLP: extraction and abstraction. In our project we focus on extraction based summarization, a subset of words that represent the most important points is pulled from a piece of text and combined to make a summary. We use some of the python NLP and Web scraping libraries to design this application.

T. Manjunath Sai Vamshi
160119737161

P. Manoj Kumar
160119737162