

Aggregation of News Articles

Week	Heading	Task to be implemented
1	Web Interface	Making dynamic interface of website using HTML, CSS, Bootstrap, Java Script and JQuery.
2	Scrapping	Training of Web Scrapping.
3	Scrappy Code	Writing code for extracting information from news articles.
4	ASP.NET	Using Extracting information algorithm in application.
5	Crawling Module	Extracted URLs would be stored in Information Repository for efficient retrieval.
6	Information Submission Module	News based on specific search key or particular news portal would be retrieved.
7	Webpage Retrieval Module	Noisy information (like images) would be removed for useful information.
8	Term Frequency Calculator	Term frequency for checking usefulness of the content.
9	Data Collection Module	Data collection above the thresh hold calculated.
10	Summary Generation Module	Most related information is scanned according to user's interests.
11	NLTK	Natural Language Tool Kit Training.
12	Word Net	Implementation of WordNet lexical database for finding semantic relation between words using NLTK with ASP.
13	NLP ALGORITHM	Making of algorithm to create a new article from scanned news articles.
14	New Article generation Module	Using NLP algorithm and tools like NLTK WordNet new articles is created.
15	Testing	Code is tested using IDE Selenium.

	16	Documentation	After experimentation, It is documented and displayed to relevant Teacher.	