

Capstone 2 Project Proposal

Project Title: Sports article Summarization

Problem Statement:

With growing digital media and long rails of articles publishing every day - Text summarization comes to rescue for generating concise and meaningful summary of text from multiple text resources such as books, news articles, blog posts, research papers, emails and tweets. The goal of this capstone is to employ Automatic Text Summarization NLP mechanism to provide time saving and useful synopsis to the user on tennis news articles dataset.

Potential Clients:

- Social Media marketing
- Financial Research

Dataset:

Sports articles



Features:

There are three features in the sports article dataset

1. article_id
2. article_text
3. source

The most important feature of these is 'article_text' which contains the text of articles.

Outcome:

Generate a single summary for all articles in the dataset.

Capstone 2 Project Proposal

Deliverables:

- Prediction model
- Report
- Visualization