PROJECT
SYNOPSIS

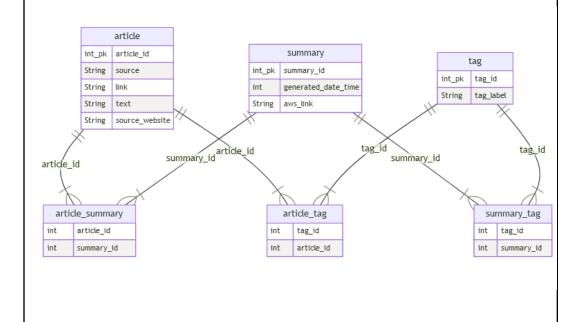| DEPARTMENT	Computer Science and Engineering	
TITLE OF THE PROJECT	News Articles Summary	
STUDENT NAMES/ USN/ PHONE/ MAIL ID	Vignesh N	1DS20CS130
	Mashhood Alam Mohammad	1DS20CS121
	Ridda Ishtiyaq	1DS20CS164
	Manoj M	1DS20CS119
MINI - PROJECT TIMELINE (Tentative Start date- End Date)	start date - 16/11/2022	
PROJECT GUIDE DETAILS	Prof. Ramya KM	
FIELD OF PROJECT	Data Science, Artificial Intelligence, machine learning, Natural Language processing, text classification, text summarization, journalism, news	
BACKGROUND OF PROJECT WITH REGARD TO THE DRAWBACK ASSOCIATED WITH EXISTING PRODUCT/PROJET	A lot of people get their news from a lot of different sources, people spend a lot of time visiting different websites, and reading the whole article. This process is very inefficient and time consuming and makes people not read news or take it from a single source.	
OBJECTIVE OF THE PROJECT	The objective of this project is to make an automation that creates a summary of news articles.	
PROJECT STATEMENT	Automation that creates and delivers news summary which can be read in 10mins everyday	
SUMMERY OF THE PROJECT	The project is based on the following steps: **Fetching articles from news websites:** News Websites have RSS feeds or API which help in retrieving articles, articles could also be found by web scraping websites. These articles are then stored in our database(mongo db) **Creating Summary:** The articles are fetched from the database, and according to topic they are grouped together and using the IBM Watson text summarizer they are summarized and written to a pdf	

Storing The Summary:
The PDF is then uploaded to aws s3 bucket and the title of the pdf and the link
are stored in the database

Sending the Summary:
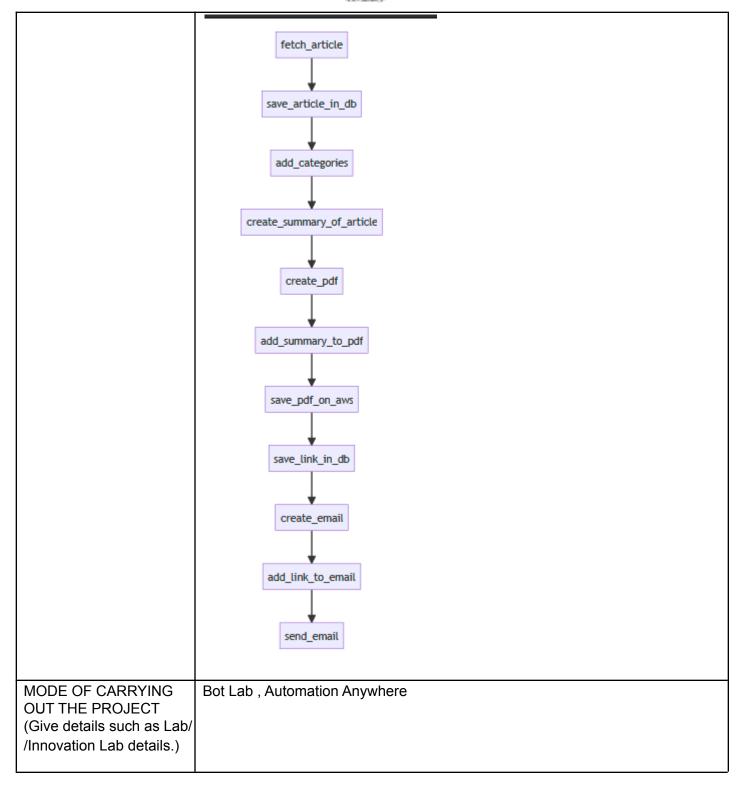The PDF is fetched from aws s3 and sent to users via auto generated email
every day

Database Design:

```
fetch_article
       │
       ▼
save_article_in_db
       │
       ▼
add_categories
       │
       ▼
create_summary_of_article
       │
       ▼
create_pdf
       │
       ▼
add_summary_to_pdf
       │
       ▼
save_pdf_on_aws
       │
       ▼
save_link_in_db
       │
       ▼
create_email
       │
       ▼
add_link_to_email
       │
       ▼
send_email
```

MODE OF CARRYING OUT THE PROJECT (Give details such as Lab/ /Innovation Lab details.)	Bot Lab , Automation Anywhere

INTENDED BENEFICIARIES OF THE PROJECT (industrial/commercial/R & D/social)	Social, People who traditionally visit news websites to get their news
ABSTRACT	Make an automation that fetches news from various websites, classifies it into categories and summarizes the contents of the article and create a pdf with the summary and links to the articles and save that pdf on the cloud