Basic Details of the Team and Problem Statement

Ministry/Organization Name/Student Innovation:

Ministry of Information & Broadcasting

PS Code: NS1230

Problem Statement Title: An AI/ML-based solution that generates short summaries of press releases for easy and more targeted dissemination of information to the people. Such TL, DR (Too Long, Didn't Read) software may be integrated into an app-based solution.

Team Name: Sudocode

Team Leader Name: Vedansh Arun

Institute Code (AISHE):

Institute Name: Netaji Subhas University of Technology

Theme Name: Smart Automation
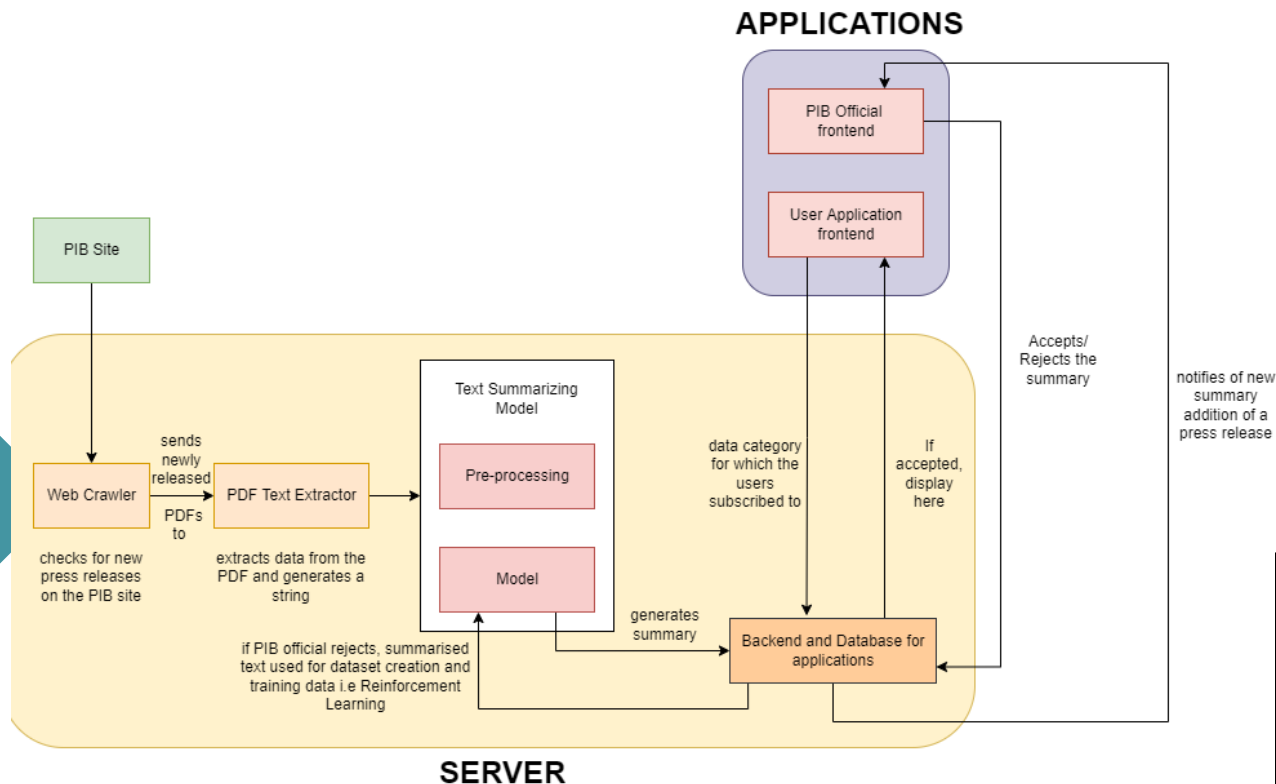
SMART INDIA HACKATHON 2022

Idea/Approach Details

Describe your idea/Solution/Prototype here:

➤ The idea of the app is to consider all the press releases from PIB's site using web crawler, extract the text from the pdf, pre-process the data and run the ML model.

➤ The summarised text is then sent to apps' backend which sent a notification to PIB officials to vet and edit the summarised text in 50-60 words. In case, the PIB officials reject the summary, we use that data to train the model via supervised learning.

➤ The user gets notified about the summary generated based on the ministries he/she has chosen to subscribe to.

➤ AI/ML Model Implementation: Extract text -> Data Preprocessing -> Extraction Summarization Algorithm -> Output Summary

Describe your Technology stack here:

➤ Flutter has been used to build the frontend of the app

➤ Python for the implementation of AI- based model and backend of the application.

Idea/Approach Details

Describe your Use Cases here

> ➢ People who fall in the age gap of 18-40 years who spend a lot of time on the mobile and don't follow any other traditional channel or read a newspaper.

Describe your Dependencies / Show stopper here

> ➢ The vetting done by the officials of PIB has an external stakeholder as PIB. The delay in the update of the summary i.e. vetted by the officials. Such time delays due to slow updates by PIB officials is a dependency.

Stretch Goals

> ➢ We plan to generate infographics from text using our ML model.
> ➢ At this satge, we have used extraction as the best possible model to generate summary in 50-60 words, however we are open to exploring other algorithms which may give better accuracy.

Competitive Advantages

> ➢ The vetting power given to PIB officials makes the summary more authentic.
> ➢ Other products that exist in the market do not provide infographics along with text summary, presently.
> ➢ Other summarizer apps that exist are not as authentic and user-friendly as our application.

Team Member Details

Team Leader Name: Vedansh Arun

Branch (Btech)   Stream: EE   Year (II):

Team Member 1 Name: Maanas Talvar

Branch (Btech)   Stream : COE   Year (III):

Team Member 2 Name: Rinav Gupta

Branch (Btech)   Stream: CSAI   Year (II):

Team Member 3 Name: Pratyush Gupta

Branch (Btech)   Stream: CSAI   Year (II):

Team Member 4 Name: Manvi Gupta

Branch (Btech)   Stream: EE   Year (II):

Team Member 5 Name: Priya Agrawal

Branch (Btech)   Stream: CSAI   Year (II):

Team Mentor 1 Name: Type Your Name Here

Category (Academic/Industry):   Expertise (AI/ML/Blockchain etc):   Domain Experience (in years):

Team Mentor 2 Name: Type Your Name Here

Category (Academic/Industry):   Expertise (AI/ML/Blockchain etc):   Domain Experience (in years):