

## TEXT SUMMARIZER - GEN AI PROJECT PHASE 2 SUBMISSION DOCUMENT

### Phase 3: Final Report and Submission

#### 1. Project Title:

AI-Based Text Summarization Using Pre-trained Transformers

#### 2. Summary of Work Done:

##### **Phase 1 – Proposal and Idea Submission:**

In this phase, we identified the problem statement and proposed the idea of developing a Text Summarization system using Generative AI.

The objectives were clearly defined:

- Understand the working of generative models for summarization.
- Use pre-trained models to generate concise summaries.
- Create an interactive interface for users to test the model.

##### **Phase 2 – Execution and Demonstration:**

In the second phase, we implemented the summarization system using Python, HuggingFace Transformers, and ipywidgets in Google Colab. The following tasks were completed:

- Built an interactive Colab interface using ipywidgets.
- Loaded the facebook/bart-large-cnn model for summarizing input text.
- Developed a button-based workflow where users input text and view the generated summary.
- Tested the model with various text inputs to ensure concise and contextually correct summaries.
- Documented complete code and sample outputs for submission.

College Name: VIT Bhopal University

Student Name: Komal Tripathi

### 3. GitHub Repository Link:

You can access the complete codebase, README instructions, and other resources at the following GitHub link:

<https://github.com/KomalTripathi/Text-Summarizer>

### 4. Testing Phase:

#### 4.1 Testing Strategy:

The system was tested across a variety of text inputs to ensure robustness and performance.

Testing focused on:

- **Input Handling:** Verified the system could handle large paragraphs, blogs, and random content.
- **Summary Quality:** Checked for relevance, grammar, and preservation of important points.
- **Edge Case Testing:** Handled extremely short inputs, blank entries, or irrelevant content gracefully.

#### 4.2 Types of Testing Conducted:

- **Unit Testing:**  
Each module (like text preprocessing, summarization function, and UI elements) was tested independently to ensure individual correctness.
- **Integration Testing:**  
Integration of the HuggingFace summarizer with the widget-based interface was tested to ensure smooth functionality.
- **User Testing:**  
Test users interacted with the interface and provided feedback regarding the summary quality and ease of use.
- **Performance Testing:**  
Summarization speed was tested for large and small inputs to ensure optimal performance without noticeable delay.

#### 4.3 Results:

College Name: VIT Bhopal University

Student Name: Komal Tripathi

- **Accuracy:**  
The system consistently generated concise, contextually relevant summaries.
- **Response Time:**  
The application responded quickly even for longer paragraphs (~500 words).
- **Edge Cases:**  
The summarizer returned short, coherent outputs even for small or random text, showing robustness.

## 5. Future Work:

Even though the system is functional, there are areas for improvement:

- **Fine-tuning the Summarization Model:**  
Training the model on a specific domain (e.g., educational content, news articles) for domain-adapted summaries.
- **Advanced Summarization:**  
Adding options for extractive vs. abstractive summarization based on user choice.
- **Real-time Summarization for Live Data:**  
Integrating APIs that pull live news/blog data and generate real-time summaries.
- **Multi-language Summarization:**  
Extending support to summarization in languages like Hindi, Spanish, etc., using multilingual models.
- **User Feedback Integration:**  
Allowing users to rate or correct summaries, thereby training a self-improving model.

## 6. Conclusion:

This project successfully demonstrates the potential of **Generative AI** in solving real-world NLP problems such as text summarization.

Through all three phases, the project evolved from concept design to functional implementation and testing.

The system proves how transformer-based models like BART can be efficiently applied for generating meaningful, coherent summaries, significantly benefiting applications in education, content writing, journalism, and research fields.