**College Name:**

VIT Bhopal University

**Student Name:**

Shwetank Thakur (21MIM10003)

Sample - GEN AI PROJECT PHASE 1 SUBMISSION DOCUMENT

**Phase 1: Proposal & Idea Submission**

**1. Project Title:**

Named Entity Recognition using Transformer Models

**2. Domain:**

Generative AI | NLP | Information Extraction

**3. Problem Statement:**

Named Entity Recognition (NER) is a key NLP task involving detection and classification of entities like names, organizations, and locations. Accurate NER is essential for tasks such as document summarization, information retrieval, and content classification. This project aims to develop a high-performing NER system using transformer-based language models.

**4. Proposed Solution:**

The project leverages Hugging Face’s `transformers` library to develop an NER system using a pre-trained model such as `dslim/bert-base-NER`. The solution will include:  
- Input processing from users.  
- Named entity prediction using the selected transformer model.  
- Output visualization using text highlighting and labels.  
- Optional model fine-tuning for domain-specific use cases.

**5. Objectives:**

- Build a functional transformer-based NER system.  
- Evaluate prediction quality across different text samples.  
- Create an interactive interface for user testing and demonstration.

**6. Expected Outcome:**

- An accurate NER system based on pre-trained models.  
- Demonstration of model capabilities with labeled text outputs.  
- Optional: Evaluation against standard datasets or application-specific data.

**7. Tools & Technologies to be Used:**

- Python  
- Hugging Face Transformers  
- BERT-based NER models  
- Jupyter Notebook / Google Colab  
- Flask or Streamlit for frontend interface

**8. References:**

- Hugging Face Documentation  
- Research on BERT for NER  
- Related open-source NER projects and GitHub repositories