VIVA

----

1. What is the significance of Cohere, LangChain, and Pydantic?

- Cohere: Provides NLP APIs including embeddings, classification, and summarization via LLMs. Used for model inference.

- LangChain: Framework to build applications powered by LLMs. It connects LLMs with tools like memory, APIs, and agents.

- Pydantic: Used for data validation and parsing using Python type hints. Useful in defining schemas and ensuring structured output from LLMs.

----

2. Explain the techniques used in word embeddings.

- Traditional: One-hot encoding, Bag of Words (BoW), TF-IDF.

- Neural: Word2Vec (CBOW, Skip-gram), GloVe (co-occurrence matrix), FastText (subword n-grams), BERT (contextual embeddings).

----

3. Why do we use word embeddings?/ Need for word embeddings.

- To reduce dimensionality and sparsity.

- To capture semantic and syntactic meanings.

- To enable words with similar meanings to have similar representations.

----

4. Discuss the real-world applications of LLMs and their limitations.

- Applications: Chatbots, summarization, code generation, sentiment analysis, content creation.

```
- Limitations: Bias in training data, hallucination, high
computational cost, lack of real-time awareness.

-----------------------------------------------------------------

5. Which model is used for summarization?

- Pre-trained summarization model from Hugging Face
(facebook/bart-large-cnn)

-----------------------------------------------------------------

6. Explain the BART model in detail.

- BART (Bidirectional and Auto-Regressive Transformer) combines
BERT (encoding) and GPT (decoding).

- It is a sequence-to-sequence model used for text generation,
summarization, translation, etc.

- Trained by corrupting text and learning to reconstruct it.

-----------------------------------------------------------------

7. What is sentiment analysis and its applications?

- It is the process of identifying sentiment (positive, negative,
neutral) from text.

- Applications: Customer feedback, brand monitoring, political
analysis, market research.

-----------------------------------------------------------------

8. Discuss and explain the significance of the parameter
perplexity in t-SNE.



- hyperparameter that defines the effective number of neighbors.

- Controls the balance between local(less perplexity) vs.
global(more perplexity) structure.

- Should be less than the number of data points; typical range:
5–50.

-----------------------------------------------------------------
```

9. Describe the algorithm (step-by-step, in words) for building an IPC chatbot.

a. Download the Indian Penal Code document.

b. Preprocess and split the document into retrievable chunks.

c. Use embeddings to store the chunks in a vector store.

d. Accept user queries.

e. Retrieve relevant sections using similarity search.

f. Use LLM (via LangChain) to answer based on the retrieved context.

----

10. Discuss PCA and t-SNE.

- PCA: Linear, preserves global variance, faster, used for large datasets.

- t-SNE: Non-linear, preserves local relationships, ideal for visualizing word clusters in small data.

- Used to visualize high-dimensional word embeddings in 2D/3D.

----

11. What are the uses of prompt engineering?

- To control LLM outputs by carefully designing the input prompts.

- Used in chatbots, summarization, translation, data extraction, and few-shot learning.
