

LLM Assignment-1 (Deadline- November 15 2025)

Provide suitable example, diagrams and code snippet wherever it is required.

1. Explain the encoder and decoder RNN model ?
2. Explain transformer model with encoder and decoder architectures ?
3. Write a program in core python to implement Byte Piece Encoder tokenization algorithm ?
4. Write a program in core python to implement Word Piece Encoder tokenization algorithm ?
5. Write a program in core python to implement Sentence Piece Encoder tokenization algorithm ?
6. Write a program in core python to implement Word2Vec embedding technique using CBOW model?
7. Write a program in core python to implement Word2Vec embedding technique using Skip-Gram model ?
9. Write a program in core python to implement GloVe embedding technique using co-occurrence matrix of words approach ?
10. Write a program in core python to implement attention mechanism using K, Q, V matrices ?

(All programs should be written in core python without using any third party libraries. First write the question , then logic for that question or program , then code with proper comments. In the last sample output).

Note - Submit this Assignment-1 in softcopy first with proper format. Google link will be shared on respective whatsapp group. Cover page will have SOA logo and headings.