**Mandsaur University, Mandsaur**

**Department of Computer Science & Engineering**

### **Subject: Deep Learning**

**Assignment 4**

**Q1. (Sequence-to-Sequence Translation using RNNs)**  
You are tasked with building a simple **English-to-Hindi translation model** using an **Encoder–Decoder RNN architecture**.

**Tasks:**  
a) Design the architecture using the following:

* Encoder: 2-layer LSTM
* Decoder: 2-layer LSTM with attention mechanism
* Embedding size = 128, Hidden units = 256
* Vocabulary size (English: 5000, Hindi: 6000)

b) Write down the input/output sequence dimensions for:

* Input sentence: 8 English words
* Output sentence: 10 Hindi words

c) Explain how using **Bidirectional RNNs** in the encoder can improve translation quality.

**OR**

**(Handling Long-Term Dependencies with LSTMs vs Vanilla RNNs)**  
You are training a model to **predict the next word** in long text sequences.

**Tasks:**  
a) Using a **vanilla RNN** with tanh activation, explain why the model struggles when sequences are longer than 50 words.  
b) Show, with equations, how the **forget gate** in LSTMs solves this problem.  
c) Implement a small PyTorch/Keras code snippet for a single-layer LSTM to process input sequences of length 100 with hidden size = 128.