<div align="center">

July 2022 Term
CSE 471 – Machine Learning
Class Test# 4 Time: 20 minutes

</div>

Name:_____ Student No._____

<div align="center">

No extra sheet shall be provided.

</div>

1. Write down the expression for BLEU, explaining each term. Explain how it penalizes shorter predicted sequences and promotes longer n-gram precision. **(10 marks.)**

2. During training of a Transformer model, for the input "Hi There", check the information below, where symbols have their usual meanings:

Embedding of "Hi", $X_1 = [1\ 2\ 3\ 4]^T$ Embedding of "There", $X_2 = [5\ 6\ 7\ 8]^T$	$W^Q = \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{matrix}$ $W^K = \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$ $W^V = \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$

Calculate the Attention Map providing detailed steps. Then fill up the below Attention Map Table, where each row represents a token's attention to all tokens. **(10 marks.)**

<div align="center">

Attention Map

	Hi	There
Hi		
There		

</div>