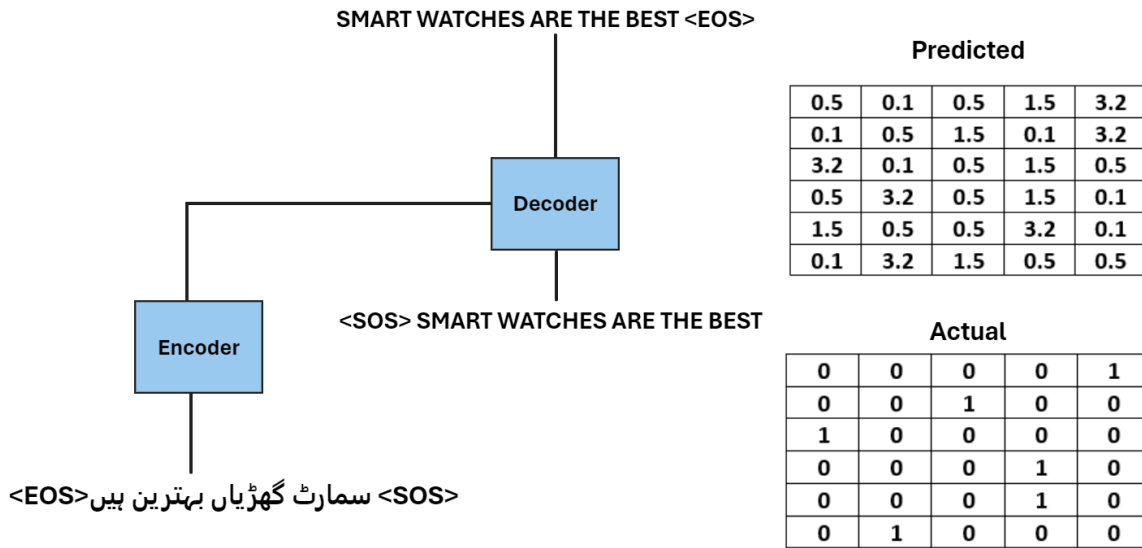


Q: Suppose we have a Transformer based Seq2Seq Model for Translation given below. You are required to perform following given tasks. [2+5 = 7 Marks]



- How many words we have in our vocabulary?
- Compute loss value using cross entropy loss function.