CS 335: Introduction to Large Language Models
Habib University

Activity Sheet 06

Name: _____ ID: _____

Question 01: Beam Search

Begin with the <START> token and use Beam Search Decoding (Beam size = 2) to generate the most probable sentence until all expanding branches reach either the End of Sequence (EOS) token or the max length (5 tokens). Note that the <START> and <EOS> tokens will be included in the token count.

The probability table for next words at each time step is as follows:

Current Sequence	Possible Next Words (with probability)
⟨START⟩	The (0.5), A (0.3), An (0.1), Silence (0.1)
"The"	traveler (0.6), scientist (0.2), stranger (0.1), warrior (0.1)
"A"	bird (0.5), dream (0.3), sound (0.1), shadow (0.1)
"The traveler"	explored (0.4), met (0.3), saw (0.2), heard (0.1)
"A bird"	sang (0.5), flew (0.3), disappeared (0.1), landed (0.1)
"The traveler explored"	<EOS> (0.5), the (0.3), vast (0.1), an (0.1)
"The traveler met"	a (0.4), an (0.3), wise (0.2), lost (0.1)
"The traveler saw"	a (0.4), something (0.3), nothing (0.2), ⟨EOS⟩ (0.1)
"The traveler heard"	a (0.5), whispers (0.3), music (0.1), nothing (0.1)
"The traveler explored the"	forest (0.4), city (0.3), desert (0.2), ocean (0.1)
"The traveler saw a"	mountain (0.4), bird (0.3), river (0.2), castle (0.1)
"The traveler met a"	guide (0.4), stranger (0.3), friend (0.2), ghost (0.1)
"A bird sang"	loudly (0.5), softly (0.3), beautifully (0.1), ⟨EOS⟩ (0.1)

Note: Take cumulative probabilities for each time step to atleast 3 decimal places.

$$\langle Start \rangle \quad The \quad traveller \quad explored \quad \langle EOS \rangle$$