The following question will ask you about the below context-free grammar, where S is the start symbol.	
	-> NP V -> N A NP -> "small" "white" -> "cats" "trees" -> "climb" "run"
The following question will also ask you about the following four sentences. Sentence 1: Cats run. Sentence 2: Cats climb trees. Sentence 3: Small cats run. Sentence 4: Small white cats climb.	
✓	Of the four sentences above, *1/1 which sentences can be derived from the above context-free grammar?
0	Only Sentence 1 Only Sentence 1 and Sentence 2
0	Only Sentence 1 and Sentence 3 Only Sentence 1 and Sentence 4 Only Sentence 1, Sentence 2, and Sentence 3
0	Only Sentence 1, Sentence 2, and Sentence 4 Only Sentence 1, Sentence 3, and
	Sentence 4 All four sentences None of the four sentences
	Which of the following is not a *1/1
•	true statement? Attention mechanisms can be used to
0	determine which parts of an input sequence are most important to focus on. One-hot representations of words
	better represent word meaning than distributed representations of words. Transformers can be faster to train
0	than recurrent neural networks because they are more easily parallelized. A Naive Bayes Classifier assumes
0	that the order of words doesn't matter when determining how they should be classified.
✓	Why is "smoothing" useful when *1/1 applying Naive Bayes?
0	Smoothing allows Naive Bayes to turn a conditional probability of evidence given a category into a probability of a category given evidence.
0	Smoothing allows Naive Bayes to be less "naive" by not assuming that evidence is conditionally independent. Smoothing allows Naive Bayes to
	better handle cases where evidence has never appeared for a particular category. Smoothing allows Naive Bayes to
0	better handle cases where there are many categories to classify between, instead of just two.
✓	From the phrase "must be the *1/1 truth", how many word n-grams of length 2 can be extracted?
0	0 1
 (a) (b) (c) (d) (d) (d) (e) (e)<	3
	4
0	5 6
\bigcirc	15
0	17