



Question Answering

Automatic Labeling

Evaluate Automatic Labeling

✔ **Video:** Evaluating label extraction
1 min

✔ **Video:** Precision and recall and F1 score
4 min

✔ **Video:** Evaluating on multiple disease categories
3 min

📅 **Lab:** Lecture notebook: Preparing Input for Text Classification
1h

Quiz: Information Extraction with NLP

📋 **Practice Quiz:** Quiz: Information Extraction with NLP
10 questions

Assignment: Natural Language Entity Extraction



Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE

100%

Quiz: Information Extraction with NLP

Quiz: Information Extraction with NLP

TOTAL POINTS 10

1. Which of the following is not true about BERT's inner word representations?

1 / 1 point

Try again



Each unique word has exactly one vector representation



The representation of a word depends on the words around it



Words which are similar in meaning are typically close as vector



None of the above



Correct

Explanation: Unlike typical word vectors, BERT uses contextualized word vectors. Therefore, since a given word's vector depends on the other vectors around it, it in general can correspond to representations. This affords BERT more flexibility, which contributes to its power.

2. True or False: the start and end vectors are fixed throughout training

1 / 1 point



True



False



Correct

Explanation: This is false. The start and end vectors, which we dot product with our word vectors to get start and end scores, are in fact learned as well during training. They are fixed at test time, however.

3. Which of the following is a difference between BERT and LSTM models?

1 / 1 point



BERT can be trained on multiple languages, while LSTMs cannot



BERT is trained using backpropagation while LSTMs are not



BERT takes entire sequences as input, while LSTM models process words one by one



BERT uses regular word vectors, while LSTMs use contextualized word vectors



Correct

Explanation: A major difference between BERT and LSTMs is that BERT process an entire sequence of input, while LSTMs only in words one by one. This enables greater parallelization and results in better training among a variety of tasks.

4. Given the following word vectors and start and end vectors, determine the start and end of the sequence of interest.

1 / 1 point

The	0.1	-0.1	S	E
BRCA1	0.25	0.05		
gene	-0.3	-0.4		
is	-0.2	0.25		
associated	-0.5	0.01		
			1	0
			0	1