

**START OF QUIZ**

**Student ID:**

**84574284, Cheng, Yushun**

## Question 1

Topic: Lecture 6

Source: Lecture 6

What are two significant shortcoming of the Word2Vec model? (1)

## Question 2

Topic: Lecture 7

Source: Lecture 7

What is the difference between coherence and cohesion? (1)

### Question 3

Topic: Lecture 5

Source: Lecture 5

Generally speaking, why are we not interested in negative PMI? (1)

## Question 4

Topic: Lecture 6

Source: Lecture 6

Think back to week 1 of this block when we were doing word sense disambiguation. Do you think there would be benefits or disadvantages to disambiguating all words before running word2vec? Explain. (2)

## Question 5

Topic: Lecture 8

Source: Lecture 8

Do you think we could use word embeddings for coreference resolution? What kind of assumptions would we be making, and why do you think it might still be a very difficult task?  
(2)

## Question 6

Topic: Lecture 7

Source: Lecture 7

We took a look at 2 different ways of implementing the TextTiling algorithm - one with vector overlap, and one with BERT. Can you think of how we might modify the algorithm further to strengthen up its weaknesses? (No is not a valid answer.) (2)

## Question 7

Topic: Lecture 8

Source: Lecture 8

What is an anaphor? (1)



## Question 8

Topic: Lecture 5

Source: Lecture 5

Which is likely to have the lowest PMI? A rare word and a frequent word that appear together frequently, or two frequent words that appear together frequently? (1)

## Question 9

Topic: Long

Source: Lecture 7

Identify the shifts in the following discourse (show your work): Jonathan Harker was a solicitor from England. He was sent to Transylvania to meet with the mysterious Count Dracula. Dracula wanted to buy property in London. That's where all the wealthiest nobles lived. Dracula had other plans, too, but Harker didn't know that. (3)

**END OF QUIZ**