* **Question 1**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Consider the following sentence: *The dog chased the cat down the road.* This sentence has two possible interpretations: that the dog moved further along the road to chase the cat, or that the dog chased the cat which is located at a lower part of the road. For the interpretation that the dog moved further along the road to chase the cat, draw the dependency graph. With this interpretation, which of the following tokens are the dependents of the head of the sentence? |  |  |  |
| |  |  | | --- | --- | | Answers: | The | |  | Correct  dog | |  | chased | |  | the | |  | Correct  cat | |  | Correct  down | |  | the | |  | road | |  | Correct  . (full stop) | |  | None of the above | |  |  |  |

* **Question 2**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following the previous question, which of the following tokens is the head of the token *down*? |  |  |  |
| |  |  | | --- | --- | | Answers: | The | |  | dog | |  | Correct  chased | |  | the | |  | cat | |  | down | |  | the | |  | road | |  | . (full stop) | |  | None of the above | |  |  |  |

* **Question 3**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Consider the same sentence as above. For the interpretation that the dog chased the cat which is located at a lower part of the road, draw the dependency graph. Which of the following dependency relations hold between the tokens *cat*and *down*? |  |  |  |
| |  |  | | --- | --- | | Answers: | det | |  | punct | |  | pobj | |  | Correct  nmod | |  | vmod | |  | sbj | |  | obj | |  | None of the above | |  |  |  |

* **Question 4**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Consider the sentence: *The dog chased the cat down the road.* Interpret this sentence as meaning that the dog moved further along the road to chase the cat. Draw the phrase structure tree for this interpretation. Identify the noun phrases in this tree. |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  The dog | |  | chased the cat | |  | down the road | |  | chased the cat down the road | |  | Correct  the cat | |  | the cat down the road | |  | Correct  the road | |  | None of the above | |  |  |  |

* **Question 5**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Consider the sentence: *The dog chased the cat down the road.* Interpret this sentence as meaning that the dog chased the cat which is located at a lower part of the road. Draw the phrase structure tree for this interpretation. Identify the noun phrases in this tree. |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  The dog | |  | chased the cat | |  | down the road | |  | chased the cat down the road | |  | Correct  the cat | |  | Correct  the cat down the road | |  | Correct  the road | |  | None of the above | |  |  |  |

* **Question 6**

1.5 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Consider the following sentence:  *The dog under the tree with the bone is happy.*  This sentence is ambiguous due to which type(s) of ambiguity? |  |  |  |
| |  |  | | --- | --- | | Answers: | Phonological | |  | Lexical | |  | Correct  Syntactic | |  | Correct  Semantic | |  | None of the above | |  |  |  |

* **Question 7**

2.66666 out of 4 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | For the input "They closed the deal with Pitt, Briggs & Co. at noon." a sentence splitter produces: "They closed the deal with Pitt, Briggs & Co." and "at noon." How will you try to improve the splitter so that it can better handle cases like this? |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  Create a dictionary of abbreviations. | |  | Write a rule specifying what tokens are disallowed before an end-of-sentence delimiter. | |  | Correct  Write a rule specifying features of tokens expected to follow an end-of-sentence delimiter. | |  | Correct  Train a model on data with examples similar to this text, that are annotated as only one sentence. | |  | None of the above |  |  |  | | --- | --- | | Response Feedback: | In this specific case, it does not help to write a rule specifying that the token "Co." is disallowed before an end-of-sentence delimiter because "Co." is a token that can possibly be the last token of other sentences, e.g., "He applied for a job at Pitt, Briggs & Co."  It will however help to create a dictionary of abbreviations with "Co." included so that the sentence splitter is aware that the full stop can simply be part of an abbreviation (and not necessarily used as an end-of-sentence delimiter). This should be combined with a rule that checks whether the token after the full stop is a possible beginning of a new sentence (in this case, the token is not capitalised hence it is unlikely to begin a new sentence). | |  |  |  |

* **Question 8**

Needs Marking

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | You were asked to write rules for extracting relations between companies that were acquired by other companies--in other words, purchased companies ([PURCHASED]) and their respective purchasers ([PURCHASER])--in text. You already know for a fact that Facebook bought Instagram.  You searched the Internet for text mentioning those entities and found the sentences below.   * + The Facebook $1 billion acquisition of Instagram was a watershed moment for Big Tech.   + Facebook acquired photo-sharing platform Instagram for $1 billion in 2012.   + Facebook bought Instagram in 2012 for $1bn (£760m).   + Instagram was acquired for $1 billion by Facebook.   + If a better analysis is used, it becomes clear that the Facebook acquisition of Instagram was illegal to begin with.   Write down a regex based on the token "acquisition". |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct  [PURCHASER] \* acquisition of [PURCHASED] | | Response Feedback: | [None Given] | |  |  |  |

* **Question 9**

Needs Marking

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Further to the previous question, write down a regex based on the token "acquired" used in the passive voice. |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct  [PURCHASED] was acquired \* by [PURCHASER] | | Response Feedback: | [None Given] | |  |  |  |

* **Question 10**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | The table linked below contains some headlines. Each headline has been labelled independently by two annotators who were asked to manually judge whether each headline is fake news or not. The labels provided by the two annotators are shown in the last two columns of the table.  [Headlines1.pdf](https://online.manchester.ac.uk/bbcswebdav/pid-12485509-dt-content-rid-68632673_1/xid-68632673_1)  We want to calculate inter-annotator agreement using the Kappa coefficient. What is the value of observed agreement? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.375 | | Answer range +/- | 0.01 (0.365 - 0.385) | |  |  |  |

* **Question 11**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous question, what is the value of expected agreement? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.5 | | Answer range +/- | 0 (0.5 - 0.5) | |  |  |  |

* **Question 12**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous questions, what is the value of the Kappa coefficient? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct -0.25 | | Answer range +/- | 0.01 (-0.26 - -0.24) |  |  |  | | --- | --- | | Response Feedback: | In this case, Kappa has a negative value. This can happen and can be interpreted as the annotators having "great disagreement" on the task. | |  |  |  |

* **Question 13**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | The table linked below contains some headlines. Shown under the Reference Label column are the corresponding labels given by a human annotator who was asked to judge whether each headline is fake or not. Meanwhile, shown under the Response Label column are the labels given by an automatic fake news detection tool.  [Headlines2.pdf](https://online.manchester.ac.uk/bbcswebdav/pid-12485509-dt-content-rid-68632755_1/xid-68632755_1)  Considering FAKE as the positive class, how many true positives were produced by the fake news detection tool? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 3 | | Answer range +/- | 0 (3 - 3) | |  |  |  |

* **Question 14**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous question, considering FAKE as the positive class, how many false positives were produced by the fake news detection tool? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 3 | | Answer range +/- | 0 (3 - 3) | |  |  |  |

* **Question 15**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous question, considering FAKE as the positive class, how many false negatives were produced by the fake news detection tool? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 1 | | Answer range +/- | 0 (1 - 1) | |  |  |  |

* **Question 16**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous question, what is the value of precision obtained by the fake news detection tool with respect to the FAKE class? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.5 | | Answer range +/- | 0 (0.5 - 0.5) | |  |  |  |

* **Question 17**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous question, what is the value of recall obtained by the fake news detection tool with respect to the FAKE class? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.75 | | Answer range +/- | 0 (0.75 - 0.75) | |  |  |  |

* **Question 18**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Following from the previous questions, what is the value of F-score obtained by the fake news detection tool with respect to the FAKE class? |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.6 | | Answer range +/- | 0 (0.6 - 0.6) | |  |  |  |

* **Question 19**

0 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | What describes the intuition of inverse document frequency (IDF)? |  |  |  |
| |  |  | | --- | --- | | Answers: | Frequent terms are more important than infrequent ones | |  | Correct  The fewer documents in which a term occurs, the higher its weight | |  | Correct  Terms that appear in many documents are down weighed | |  | Frequent terms are indicators of a document's meaning |  |  |  | | --- | --- | | Response Feedback: | IDF is used to decrease weights of common terms, i.e., less important ones, and increase weights of uncommon ones. Therefore, both option 2 and 3 are correct. | |  |  |  |

* **Question 20**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Given four documents after preprocessing as follows: Doc1: my cat meow all the time while my dog does n't bark . Doc2: can cat be friendly ?  Doc3: i do n't know why my dog always bark at squirrel .   Doc4: dog be sad when being home alone .  Using the collection and the formula in Week 3 lectures, calculate tf-idf(“bark”, Doc1). Please round off your answer to three decimal places. |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.091 | | Answer range +/- | 0.001 (0.090 - 0.092) |  |  |  | | --- | --- | | Response Feedback: | tf-idf(bark, Doc1) = log\_10(1+1) \* log\_10(4/2) = 0.091 | |  |  |  |

* **Question 21**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | With the same four documents in the previous question and the formula in Week 3 lectures, calculate tf-idf(“dog”, Doc3). Please round off your answer to three decimal places. |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct 0.038 | | Answer range +/- | 0.001 (0.037 - 0.039) |  |  |  | | --- | --- | | Response Feedback: | tf-idf(dog, Doc3) = log\_10(1+1) \* log\_10(4/3) = 0.038 | |  |  |  |

* **Question 22**

0 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | With the same four documents in the previous question, when we use a word-word matrix approach with a context window of 2, i.e., two words to the left and two words to the right of the current word, which of the following counts, namely count(current word, its context words), are correct? |  |  |  |
| |  |  | | --- | --- | | Answers: | count(cat,while) = 1 | |  | Correct  count(cat,can) = 1 | |  | count(dog,bark) = 2 | |  | Correct  count(dog,my) = 2 | |  | Correct  count(dog,sad) = 1 | |  |  |  |

* **Question 23**

2.0001 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Given the following term-document matrix, which of the following interpretation(s) is/are correct?  [term-document matrix](https://online.manchester.ac.uk/bbcswebdav/pid-12485509-dt-content-rid-69148261_1/xid-69148261_1)  NOTE: we assume that the matrix was created from a collection of 4 documents. |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  The word “good” can be represented as a vector of (114, 80, 62, 89) | |  | tf-idf("battle", “As You Like It”) < tf-idf("wit", “As You Like It”) | |  | Correct  The document “Henry V” can be represented as a vector of (13, 89, 4, 3) | |  | Correct  tf-idf("battle", “Julius Caesar”) > tf-idf("wit", “Julius Caesar”) | |  |  |  |

* **Question 24**

0 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Select disadvantages of using count-based vectors to represent word meanings: |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  Sparse vectors, hence difficult for machine learning | |  | Correct  Do not capture meanings of words in different contexts | |  | Cannot use count-based vectors to calculate the similarity of words | |  | When using count-based vectors, it's easy to debug your model | |  |  |  |

* **Question 25**

2 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | In a named entity recognition (NER) task, if we have to detect 10 NE categories using BIO scheme, how many classes will we have? |  |  |  |
| |  |  | | --- | --- | | Answers: | 30 classes | |  | 20 classes | |  | 31 classes | |  | Correct  21 classes | |  |  |  |

* **Question 26**

Needs Marking

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Given the following sentence “The British balloon is to be flown by Richard Branson, chairman of Virgin Atlantic Airways; Per Lindstrand, chairman of Lindstrand Balloons Ltd. of Oswestry, England, and an Irish balloonist, Rory McCarthy.” identify named entities (person names -- PER, organisations -- ORG, and locations - LOC) by putting each named entity within square brackets and its corresponding category as a subscript, e.g., ... [Johnson]\_PER ... |  |  |  |
| |  |  | | --- | --- | | Correct Answer: | Correct  The British balloon is to be flown by [Richard Branson]\_PER, chairman of [Virgin Atlantic Airways]\_ORG; [Per Lindstrand]\_PER, chairman of [Lindstrand Balloons Ltd.]\_ORG of [Oswestry]\_LOC, [England]\_LOC, and an Irish balloonist, [Rory McCarthy]\_PER. | | Response Feedback: | [None Given] | |  |  |  |

* **Question 27**

2 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Why is a global approach better than a local approach in NER? |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  A global approach uses the whole sentence/text, which gives context that helps inform each individual prediction. | |  | Correct  Since there are more Os then Bs and Is (in the case of BIO scheme), to avoid bias in a model, it's best to use a global approach to consider the whole sequence tag rather than considering them independently. | |  | Because a global approach can take into account POS features. | |  | Correct  Because a local approach can't completely address the context of the input sequence. | |  |  |  |

* **Question 28**

2 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Which sentences are correct? |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  CRF is a global approach to NER | |  | Using RNN to represent a sentence and apply a softmax layer on top of that is another global approach to NER | |  | Correct  Using BiLSTM to represent a sentence and apply CRF on top of that is also a global approach to NER | |  |  |  |

* **Question 29**

2 out of 2 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | What are disadvantages of using neural networks for NER? |  |  |  |
| |  |  | | --- | --- | | Answers: | We don’t need to do feature engineering | |  | Correct  Since neural networks implicitly create features for the input instances via hidden layers, it’s not easy to interpret the results | |  | Correct  To achieve good performance, we have to pre-trained vectors using big document collections | |  | All of the above | |  |  |  |

* **Question 30**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | What is true about Named Entity Linking (NEL)? |  |  |  |
| |  |  | | --- | --- | | Answers: | NEL is a trivial task since we only need to check entity mentions against an available knowledge base | |  | Correct  NEL is important in natural language understanding | |  | Correct  Basically, we can approach NEL by doing two steps: candidate generation and candidate ranking | |  | We can only use neural networks to address the candidate ranking step | |  |  |  |

* **Question 31**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | According to WordNet, which of the following sentences are correct? |  |  |  |
| |  |  | | --- | --- | | Answers: | 'Apple' is a hypernym of 'Edible Fruit' | |  | Correct  'Apple' is a hyponym of 'Edible Fruit' | |  | Correct  'Wheeled vehicle' is a hypernym of 'Bicycle'' | |  | 'Wheeled vehicle' is a hyponym of 'Bicycle'' | |  | Correct  'Leader' is an antonym of 'Follower' | |  |  |  |

* **Question 32**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Given the following information: Input sentence: Waves were hitting the steep **bank**.  **Senses for the word “bank”**: 1. sloping land (especially the slope beside a body of water) 2. a financial institution that accepts deposits and moves the money into lending activities 3. a building in which the business of banking transacted  **Context definitions:** wave – one of a series of ridges that moves across the surface of a liquid (especially across a large body of water) hit – hit against; come into sudden contact with steep – of a slope; set at a high angle  Using the Lesk algorithm with a similarity measure of counting the overlap words between two sets of sense definitions, what is the correct sense(s) of the word ‘bank’ in the input sentence? |  |  |  |
| |  |  | | --- | --- | | Answers: | Sense 3 | |  | Sense 2 | |  | Correct  Sense 1 | |  | All of the above | |  |  |  |

* **Question 33**

1.9998 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Which of the following tasks can be addressed by using sequence labelling models? |  |  |  |
| |  |  | | --- | --- | | Answers: | Correct  Part-Of-Speech Tagging | |  | Correct  Named Entity Recognition | |  | Correct  Semantic Role Labelling | |  | Syntactic Parsing | |  | Machine Translation | |  | All of the above | |  |  |  |

* **Question 34**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Overall, the success of deep learning in NLP can be attributed to |  |  |  |
| |  |  | | --- | --- | | Answers: | the fact that neural networks incorporate decades of linguistic theories in form of hand-crafted features | |  | Correct  its capability to learn and utilise expressive representations in unsupervised manner | |  | Correct  its capability to capture the variability of natural language | |  | Correct  the availability of large-scale natural language datasets, deep learning frameworks and specialised computational resources | |  |  |  |

* **Question 35**

3 out of 3 points

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | |
|  | Which of the following statements are true about Long Short Term Memory (LSTM) networks? |  |  |  |
| |  |  | | --- | --- | | Answers: | LSTM is a special case of Recurrent Neural Networks (RNN) that only focuses on learning short term dependencies. | |  | Correct  By using LSTM, we can avoid the vanishing gradients in RNN. | |  | Correct  LSTM cells have memory that can store previous timestep information, which helps the model incorporate context information. | |  | Correct  At each time step, LSTM learns how much information to be kept from the previous time step and how much information to be added at the current step. | |  | All of the above | |  |  |  |

Tuesday, 25 May 2021 11:23:32 o'clock BST