Stephanie Nicole Garibay Lim All Modules Help ▼ You are currently viewing Archive 2022 Site **∷** Ŋ Dashboard / My Modules / ECS763U/ECS763P - Natural Language Processing - 2022/23 - Semester 1 / Week 11 - (Unit 9) Semantics II: Meaning Representations with Vectors / Unit 9 Quiz: Semantics 2: Meaning Representation with Vectors (2% of grade) ECS763U/ECS763P - Natural Language Processing -2022/23 - Semester 1 **Quiz navigation ₹ Started on** Friday, 23 December 2022, 2:23 PM **State** Finished 命 **Completed on** Friday, 23 December 2022, 2:48 PM **Time taken** 24 mins 43 secs **Grade 9.67** out of 10.00 (**97**%) Show one page at a time \bigcirc Question **1** Finish review Correct Mark 1.00 out of The purpose of using a weighting technique like TF-IDF is to minimise the effect of unimportant words which appear in _____. Select one: a. many documents b. few documents The correct answer is: many documents Question **2** Correct Mark 1.00 out of 1.00 Word2Vec is a technique like term-context matrix building which uses defined contexts of words as the basis for its vectors. Select one: True ● False The correct answer is 'False'. Question **3** Correct Mark 1.00 out of 1.00 $\ensuremath{\mathbb{F}}$ Flag question See the term-context matrix for two target words `princess' and `queen' below. Select the correct depiction of the two vectors representing the two words in 2-dimensional space from the two plots. powerful woman 27 16 princess 35 30 queen queen princess 'woman' 16 'powerful' b) princess queen 'woman' 'powerful' Select one: a. a b. b The correct answer is: a Question 4 Correct Mark 1.00 out of Which of the following structure preserving maps have been experimented with between a syntactic system and vector space semantics system? Select one or more: ✓ a. Lambek calculus -> vector spaces. pregroup grammar -> vector spaces. c. Domain of individuals -> vector spaces. d. Properties -> vector spaces. ✓ e. CCG -> vector spaces. f. Referents -> vector spaces. The correct answers are: pregroup grammar -> vector spaces., Lambek calculus -> vector spaces., CCG -> vector spaces. Question **5** Partially correct Mark 0.67 out of 1.00 Select all of the statements true of TF-IDF. Select one or more: ☑ a. The IDF part of TF-IDF is a function of document frequency of a term divided by the number of documents. × ☑ b. TF-IDF gives more importance to words that occur in few documents. c. TF-IDF is always guaranteed to give more realistic similarity scores. ☑ d. The IDF part of TF-IDF is a function of the number of documents divided by the document frequency of a term. ☑ e. TF-IDF uses smoothing to avoid getting the logarithm of 0. ☐ f. TF-IDF gives more importance to words that occur in many documents. The correct answers are: TF-IDF uses smoothing to avoid getting the logarithm of 0., The IDF part of TF-IDF is a function of the number of documents divided by the document frequency of a term., TF-IDF gives more importance to words that occur in few documents. Question **6** Correct Mark 1.00 out of 1.00 Select all of the following statements which are consistent with the distributional hypothesis of language meaning. Select one or more: ☑ a. You shall know a word by the company it keeps. b. The meaning of a sentence is determined by its denotation according to a formal model of the world. Words that appear in the same contexts can be assumed to be synonymous. Meaning is determined by a logical form attached to each word or constituent. e. The syntax of a language determines the meaning of sentences. ☑ f. A word's meaning can be approximated by the distribution of the contexts it occurs in. The correct answers are: You shall know a word by the company it keeps., A word's meaning can be approximated by the distribution of the contexts it occurs in., Words that appear in the same contexts can be assumed to be synonymous. Question **7** Correct Mark 1.00 out of You have a corpus of 50 documents. See the table below for the document frequency for a list of terms in the corpus. What is the logarithm of the inverse document frequency (IDF) (not the full TF-IDF calculation) of the term `Han', using a log base of 10, to 3 decimal places? DFt (document frequency of term) 30 Luke 22 Leia 18 Han N = 5045 soldier 24 trooper ewok 20 50 and Select one: a. 2.778 b. 0.444 c. -0.444 d. 0.360 e. -1.123 The correct answer is: 0.444 Question 8 Correct Mark 1.00 out of 1.00 $\begin{tabular}{ll} \Bbb Flag question \end{tabular}$ Consider the table of document frequencies for the terms in the table across a corpus of 50 documents shown below, in addition to the termdocument matrix shown for three of the 50 documents in the corpus. Using a log base of 10 and using add-one smoothing to term frequencies, calculate the TF-IDF score of the term `soldier' in the document `Star Wars Episode VI: Return of the Jedi'. Give your answer to 2 Decimal Places. DFt (document frequency of term) 30 Luke Leia 22 18 Han N = 5045 soldier 24 trooper ewok 20 50 and 50 it TF*t*,*d* (frequency of term *t* in document *d*) Star Wars Star Wars Episode Star Wars Episode VI: Return of the V: The Empire Episode IV: A Strikes Back Jedi New Hope Luke 39 43 50 38 Leia 30 36 40 43 Han 15 soldier 80 102 94 98 89 78 trooper 0 0 30 ewok 1009 897 988 and it 754 659 680 Answer: 0.09 The correct answer is: 0.09 Question **9** Correct Mark 1.00 out of 1.00 Consider the table of document frequencies for the terms in the table across a corpus of 50 documents shown below, in addition to the termdocument matrix shown for three of the 50 documents in the corpus. Using a log base of 10 and using add-one smoothing to term frequencies, calculate the TF-IDF score of the term `ewok' in the document `Star Wars Episode VI: Return of the Jedi'. Give your answer to 2 Decimal Places. DFt (document frequency of term) Luke 30 22 Leia 18 Han N = 50soldier 45 24 trooper ewok 20 50 and it 50 TF*t*,*d* (frequency of term *t* in document *d*) Star Wars Episode Star Wars Star Wars Episode V: The Empire VI: Return of the Episode IV: A Strikes Back New Hope Jedi Luke 43 39 50 38 30 36 Leia 40 43 15 Han soldier 102 94 80 98 89 78 trooper 0 0 30 ewok 1009 897 988 and it 754 659 680 Answer: 0.59 The correct answer is: 0.59 Question **10** Correct Mark 1.00 out of Consider the table of document frequencies for the terms in the table across a corpus of 50 documents shown below, in addition to the termdocument matrix shown for three of the 50 documents in the corpus. Using a log base of 10 and using add-one smoothing to term frequencies, calculate the TF-IDF score of the term `Han' in the document `Star Wars Episode VI: Return of the Jedi'. Give your answer to 2 Decimal Places. DFt (document frequency of term) Luke 30 22 Leia 18 Han N = 5045 soldier 24 trooper 20 ewok 50 and 50 it TF*t*,*d* (frequency of term *t* in document *d*) Star Wars Star Wars Episode Star Wars Episode VI: Return of the V: The Empire Episode IV: A Strikes Back New Hope Jedi Luke 39 43 50 Leia 38 30 36 43 40 15 Han soldier 80 102 94 98 89 78 trooper 0 0 30 ewok 1009 897 988 and it 754 680 659 Answer: 0.53 The correct answer is: 0.53 Finish review Programming Assignment 2 (worth 40% of ◀ Lecture live Zoom link for remote overall mark): Vector Space Semantics for Jump to... Similarity between Eastenders Characters participants week 11 **Student Life** Library Archives Student email Library Landing Page Archive My QMUL Library Website 2021/22 Find It! Use It! Reference It! Queen Mary Students' Union 2020/21 Student Enquiry Centre Library Search 2019/20 QMplus for students Subject Guides 2018/19 Cite Them Right 2017/18 Careers Academic Skills

 Ω