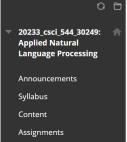
20233\_csci\_544\_30249: Applied Natural Language Processing

Assignments

Review Test Submission: Quiz 7



Tools

**USC Code of Ethics** 

Review	lest Submission:	Quiz 7
--------	------------------	--------

User	Kayvan Shah
Course	20233_csci_544_30249: Applied Natural Language Processing
Test	Quiz 7
Started	10/31/23 5:35 PM
Submitted	10/31/23 5:45 PM
Due Date	10/31/23 5:50 PM
Status	Completed
Attempt Score	100 out of 100 points
Time Elapsed	9 minutes out of 10 minutes
Results Displaye	ed All Answers, Submitted Answers, Correct Answers

10 out of 10 points

What are the foundational assumptions of the IBM Model 1 for machine translation?

Selected Answers: 😋 All possible alignments between source and target language sentences are equally likely.

Each word in the target language is generated independently from the source language words.

Answers:

Answers:

All possible alignments between source and target language sentences are equally likely. The translation probability of a word depends on its position in the sentence.

Each word in the target language is generated independently from the source language words. Contextual information is crucial for accurate translation.

Question 2 10 out of 10 points

In the context of phrase-based translation, what is a 'phrase table' primarily used for?

Selected Answer: 👩 Mapping phrases from the source language to corresponding phrases in the target language.

Storing common phrases to speed up the translation process.

Mapping phrases from the source language to corresponding phrases in the target language. Cataloging grammatical rules and exceptions for each language involved.

A lookup table for syntactic parsing rules.

Question 3 10 out of 10 points

Which of the following best describes a challenge associated with the EM algorithm?

Selected Answer: 👩 The EM algorithm can converge to different solutions depending on the initial parameter estimates.

The EM algorithm always converges to the global optimum of the likelihood function. Answers:

The EM algorithm requires a large amount of labeled data for accurate parameter estimation.

The EM algorithm can converge to different solutions depending on the initial parameter estimates.

The EM algorithm is computationally less intensive compared to other iterative algorithms.

**Question 4** 10 out of 10 points

Which of the following statements describe the limitations of IBM Model 1?

Selected Answers: 👩 It does not consider word context in translation.

It cannot handle non-literal translations effectively.

 $_{m{\lozenge}}$  It treats all possible alignments with equal probability.

Answers:

It does not consider word context in translation.

It cannot handle non-literal translations effectively. It is sensitive to the size of the training corpus.

It treats all possible alignments with equal probability.

**Question 5** 10 out of 10 points

In phrase-based translation models, what is the significance of 'language modeling'?

It is used to recognize speech in the source language. Answers: It is used for the phonetic translation of words. It predicts the likelihood of a sequence of words in the target language. It models the syntactic structure of the source language. Question 6 10 out of 10 points In phrase-based translation models, what is a critical step in extracting phrase pairs from a parallel corpus? Selected Answer: 👩 Identifying and aligning phrases based on statistical significance and frequency of occurrence. Applying part-of-speech tagging to identify noun and verb phrases separately. Using a bilingual dictionary to align phrases manually. Identifying and aligning phrases based on statistical significance and frequency of occurrence. Implementing syntactic parsing to establish grammatical relationships between words in a phrase. **Question 7** 10 out of 10 points In phase-based translation models, what role does 'translation probabilities' play? Selected Answer: 👩 Matching words or phrases in the source language with their equivalents in the target language. Aligning voice with text for better speech recognition. Answers: Matching words or phrases in the source language with their equivalents in the target language. Aligning the translation model with neural network parameters. Ensuring syntactical alignment in the source and target languages. **Question 8** 10 out of 10 points Which technique is often employed in phase-based translation models to handle the reordering of phrases? Selected Answer: o Distance-based reordering models N-gram modeling Answers: Beam search algorithms Convolutional neural networks Distance-based reordering models **Question 9** 10 out of 10 points What is the primary purpose of the Expectation step in the EM algorithm? Selected Answer: 👩 To estimate the probability distribution of the hidden variables given the observed data and current parameter estimates. Answers: To maximize the likelihood function based on current parameter estimates. To update the model parameters based on the expected complete data log-likelihood. 👩 To estimate the probability distribution of the hidden variables given the observed data and current parameter estimates. To compute the definitive classification of data points into distinct clusters. **Question 10** 10 out of 10 points What is a key feature that distinguishes IBM Model 2 from IBM Model 1 in machine translation? Selected Answer: 👩 IBM Model 2 adds a non-uniform positional alignment model to account for the relative positioning of words. IBM Model 2 introduces a fertility model to handle the dropping and adding words. Answers: IBM Model 2 adds a non-uniform positional alignment model to account for the relative positioning of words. IBM Model 2 incorporates a deep learning approach for context understanding. IBM Model 2 utilizes a grammar-based translation system. Saturday, January 20, 2024 11:42:06 PM PST

← OK

Selected Answer: 👩 It predicts the likelihood of a sequence of words in the target language.