1 Question: What is Named-entity Recognition (NER) also known as? A) Entity identification B) Entity extraction C) Entity analysis D) Entity recognition Correct Answer: A) Entity identification 2 Question: Which of the following is NOT a category that NER classifies entities into? A) Organizations B) Locations C) Colors D) Monetary values Correct Answer: C) Colors 3 Question: In which domain is NER extensively used for DNA and gene identification? A) Biomedical data B) Financial data C) Educational data D) Legal data Correct Answer: A) Biomedical data 4

Question: What is the main goal of the problem statement in the text?

- A) To extract named entities from movie trivia
- B) To build a general-purpose NER library

C) To optimize search engine algorithms
D) To perform machine translation
Correct Answer: A) To extract named entities from movie trivia
5
Question: Which algorithm is mentioned in the text as being suitable for sequence classification
problems?
A) Support Vector Machines
B) Decision Trees
C) Conditional Random Fields
D) K-Means Clustering
Correct Answer: C) Conditional Random Fields
6
6
Question: What is the output format of the data used from the MIT movie corpus?
A) .csv
B) .bio
C) .txt
D) .xml
Correct Answer: B) .bio
7
Question: Which Python library is mentioned as having a wrapper for the Stanford NER?
A) spaCy
B) NLTK
C) TensorFlow
D) PyTorch
Correct Answer: B) NLTK

Question: What is the main purpose of feature functions in CRF?

- A) To perform hyperparameter tuning
- B) To express characteristics of the sequence
- C) To reduce the dimensionality of the data
- D) To classify the text into different languages

Correct Answer: B) To express characteristics of the sequence

9

Question: What is the first step in the methodology and approach for solving the NER problem stated in the text?

- A) Data collection
- B) Model Building
- C) Data preprocessing
- D) Hyperparameter tuning

Correct Answer: A) Data collection

10

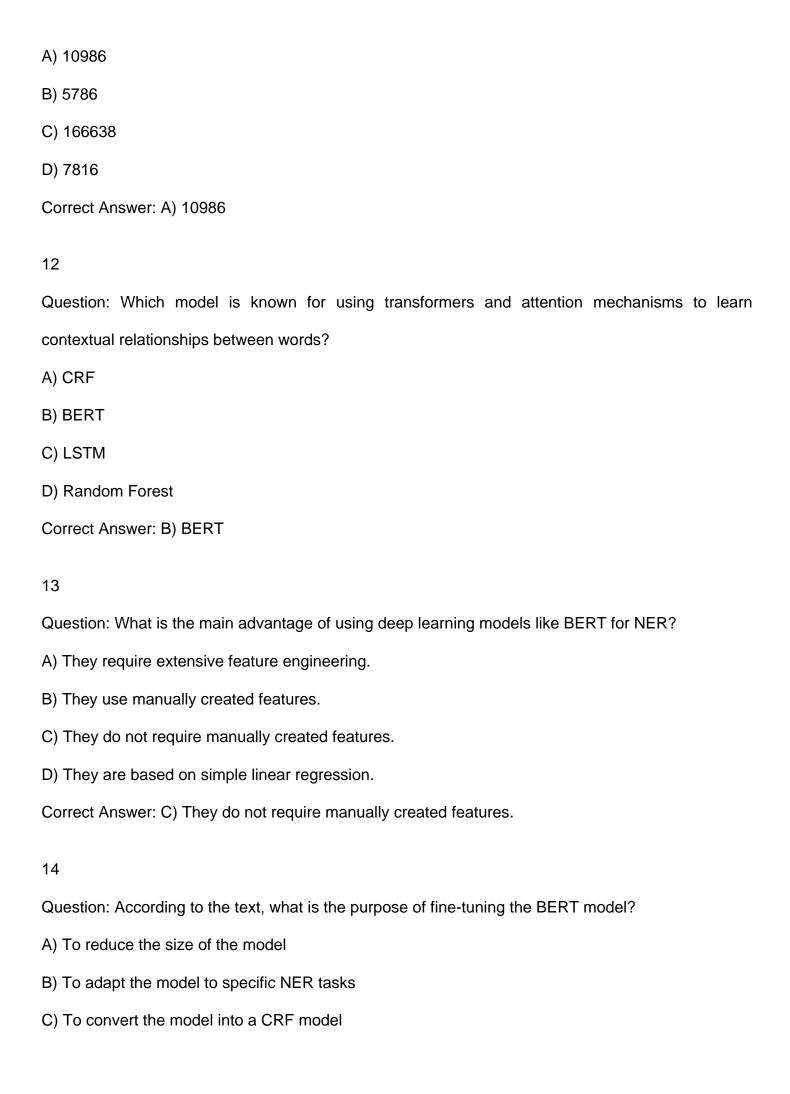
Question: Which of the following is NOT a step in solving the NER problem as per the text?

- A) Data understanding
- B) Feature mapping
- C) Data encryption
- D) Evaluating the model

Correct Answer: C) Data encryption

11

Question: What is the total number of unique words present in the training data set mentioned in the text?



D) To train the model on general NLP tasks

Correct Answer: B) To adapt the model to specific NER tasks

15

Question: What is the format of the data set after extraction as mentioned in the text?

- A) Two columns
- B) Three columns
- C) Four columns
- D) Five columns

Correct Answer: B) Three columns