1

Question: What does NER stand for in the context of Natural Language Processing?

A) Neural Entity Recognition

B) Named Entity Recognition

C) Natural Entity Recognition

D) Numeric Entity Recognition

Correct Answer: B) Named Entity Recognition

2

Question: Which university is associated with the author Arya Roy?

A) University of Pennsylvania

B) Harvard University

C) Carnegie Mellon University

D) Stanford University

Correct Answer: C) Carnegie Mellon University

3

Question: What is the primary focus of the paper discussed in the text?

A) The evolution of hardware for data processing

B) The history of computer-readable textual data

C) Learning methods employed for NER in the recent past

D) The development of new encoding systems for NER

Correct Answer: C) Learning methods employed for NER in the recent past

4

Question: Which of the following is NOT listed as a keyword in the abstract?

A) Sequence tagging

B) Deep learning

C) Conditional random fields

D) Semantic analysis

Correct Answer: D) Semantic analysis

5

Question: What general class of problems in NLP does NER belong to?

A) Semantic tagging

B) Sequence tagging

C) Syntactic parsing

D) Speech recognition

Correct Answer: B) Sequence tagging

6

Question: Which of the following is a task that is upstream or downstream to NER?

A) Sequence tagging

B) Entity linking

C) Word embedding

D) Deep learning architecture

Correct Answer: B) Entity linking

7

Question: What type of learning methods have produced state-of-the-art results for NER?

A) Linear learning methods

B) Deep learning models

C) Unsupervised learning methods

D) Rule-based learning methods

Correct Answer: B) Deep learning models

8

Question: Which model is known for being one of the most effective algorithms for NER?

A) Hidden Markov Models (HMM)

B) Support Vector Machines (SVM)

C) Conditional Random Fields (CRF)

D) Decision Trees

Correct Answer: C) Conditional Random Fields (CRF)

9

Question: What is the main advantage of using deep learning frameworks over linear techniques for NER?

A) They require less training data

B) They are easier to implement

C) They offer better performance across different domains and languages

D) They are based on simpler mathematical models

Correct Answer: C) They offer better performance across different domains and languages

10

Question: What is the purpose of using word embeddings in NER tasks?
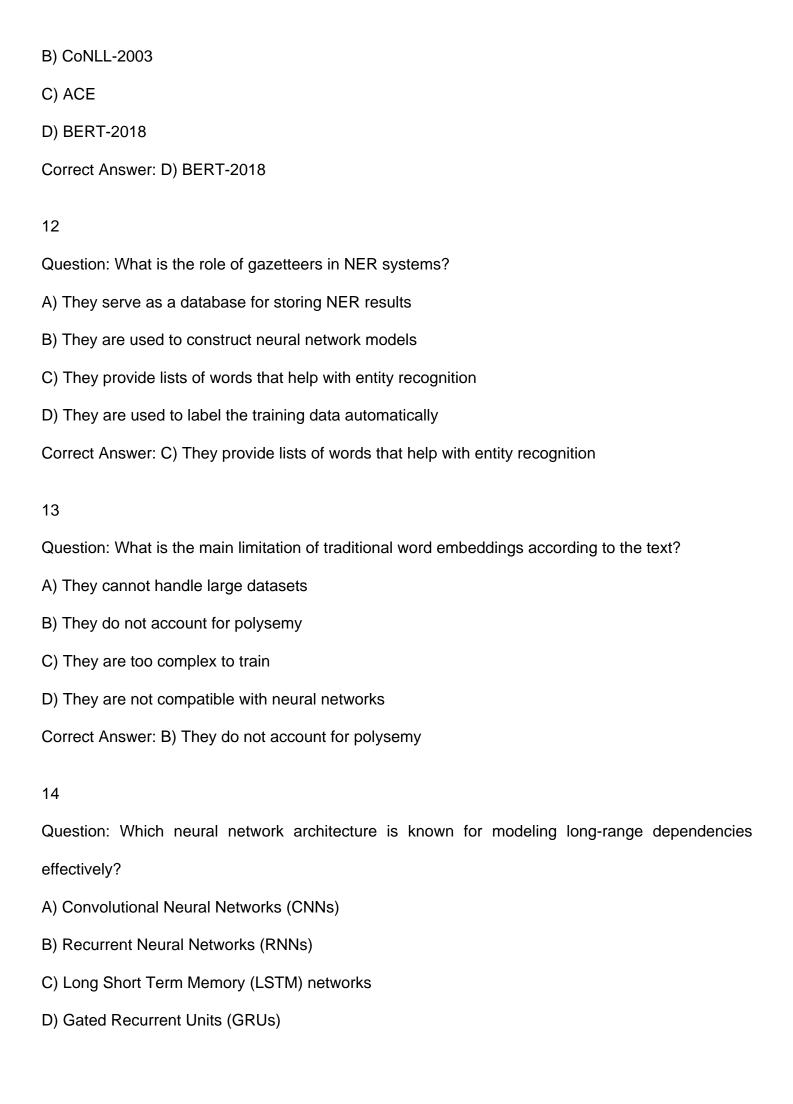
A) To reduce the size of the training dataset

B) To create a concise vector representation of words

C) To increase the speed of the NER system

D) To improve the graphical interface of the NER system

Correct Answer: B) To create a concise vector representation of words

11

Question: Which of the following is NOT a standard dataset used to benchmark NER performance?

A) CoNLL-2002

B) CoNLL-2003

C) ACE

D) BERT-2018

Correct Answer: D) BERT-2018

12

Question: What is the role of gazetteers in NER systems?

A) They serve as a database for storing NER results

B) They are used to construct neural network models

C) They provide lists of words that help with entity recognition

D) They are used to label the training data automatically

Correct Answer: C) They provide lists of words that help with entity recognition

13

Question: What is the main limitation of traditional word embeddings according to the text?

A) They cannot handle large datasets

B) They do not account for polysemy

C) They are too complex to train

D) They are not compatible with neural networks

Correct Answer: B) They do not account for polysemy

14

Question: Which neural network architecture is known for modeling long-range dependencies effectively?

A) Convolutional Neural Networks (CNNs)

B) Recurrent Neural Networks (RNNs)

C) Long Short Term Memory (LSTM) networks

D) Gated Recurrent Units (GRUs)

Correct Answer: C) Long Short Term Memory (LSTM) networks

15

Question: What is the main benefit of using a bi-directional LSTM for NER?

A) It reduces the computational complexity

B) It captures dependencies on both left and right context of a target word

C) It eliminates the need for feature engineering

D) It works well with small datasets

Correct Answer: B) It captures dependencies on both left and right context of a target word