



# TestHorse

Certified IT practice exam authority

---

Accurate study guides, High passing rate!  
Testhorse provides update free of charge in one year!



<http://www.testhorse.com>

**Exam** : **E20-065**

**Title** : Advanced Analytics  
Specialist Exam for Data  
Scientists

**Version** : Demo

1. Given an input vector of features, a Random Forests model performs a classification task and ends in a tie.

How does the model handle this outcome?

- A. The model will be rebuilt
- B. A winner is chosen at random
- C. The tree that caused the tie is discarded
- D. One more tree is added to the forest

**Answer: B**

2. Which HDFS feature protects against user errors causing accidental loss of data?

- A. Encryption
- B. Replication
- C. Namenode federation
- D. Snapshots

**Answer: B**

3. What process must address acoustic ambiguity in NLP?

- A. Part-of-speech tagging
- B. Word sense disambiguation
- C. Speech recognition
- D. Discourse

**Answer: C**

4. A data engineer is asked to process several large datasets using MapReduce. Upon initial inspection the engineer realizes that there are complex interdependencies between the datasets.

Why is this a problem?

- A. MapReduce works best on unstructured data
- B. There is no problem; MapReduce accommodates all the data
- C. MapReduce can only parse one file at a time.
- D. MapReduce is not ideal when the processing of one dataset depends on another.

**Answer: D**

5. What is a characteristic of stop words?

- A. Used in term frequency analysis
- B. Include words such as "a", "an", and "the"
- C. Meaningful words requiring a parser to stop and examine them
- D. Don't occur often in text

**Answer: B**