Quiz Questions

1. How does Latent Semantic Analysis allow for dimensionality reduction?

a) It randomly projects using normalized Gaussians onto a low dimensional subspace

b) It uses Canonical Cluster Analysis to provide a supervised technique for reducing dimensions

**c) Principal Component Analysis allows us to take the largest singular value/vector combinations and project**

d) We take the largest vectors by norm in our SVD

2. How should we interpret the basis vectors in LSA?

**a) They represent underlying topics**

b) They represent underlying sentences

c) They represent underlying words

d) They represent underlying paragraphs

3. What techniques in machine learning do we use in Word2Vec?

a) Linear Regression

**b) Neural Networks**

c) Ridge Regression

d) Boosting

4. How does Word2Vec create the word embedding?

**a) It learns a matrix and treats it as a lookup table**

b) It utilizes neural nets to optimize efficiency via PCA

c) It passes words through a recurrent neural network

5. What theorem do we invoke in order to perform PCA in our document-term matrix?

**a) Eckart-Young Theorem**

b) Tikhonov Theorem

c) Low-Rank Approximation Theorem

d) Latent Dirichlet Analysis

6. Word2Vec is a supervised algorithm?

**a) True**

b) False

7. What do we train our neural network on in our Word2Vec model?

**a) Context words**

b) Skip-gram words

c) Continuous Bag of Words

8. What matrix do we perform PCA on?

a) Neural weight matrix

b) Term-paragraph matrix

**c) Document-term matrix**

9. What are some of the drawbacks in LSA?

a) The basis vectors are not necessarily interpretable

b) They have huge variance

c) Vectors for individual words cannot be easily combined to form a vector for a large block of text

10. Is LSA a supervised algorithm?

a) True

**b) False**