**Short Answer:**

**1.** The window size of a skipgram/CBOW model is the context window around the target word which is used for predicting either the target word given the context word or vice versa. Therefore, the window size controls the context for target word. You may increase it when you are looking at documents with a large sequence of strings such as a legal report or analyst report. You can decrease if when documents have a short sequence of strings such as tweets or text messages since the words to analyze are more limited.

**2.**

a. Julie loves me and Linda likes me

b. Julie likes me and Linda loves me

**4.** option B

**5.**

a. Observed States: Falak, studies, USC

b. Hidden States: PERSON, NON\_NAMED\_ENTITY, PLACE

c. The transition matrix shows the probability of transitioning from one hidden state to another

d. The emission matrix provides the probability of emission of the observed state given the hidden state.

**6.** BBC news articles would suffer from the exploding gradients as they have a large sequence of strings which cannot be handled by RNN. Due to the chain rule application while calculating the error gradients, the domination of the multiplicative term increases over time and due to that the gradient has the tendency to explode or vanish. If the largest eigenvalue is less than 1, the gradient will vanish. If the largest eigenvalue is more than 1, the gradient explodes.

**True/False:**

**B.** False. After dimensionality reduction using PCA, the number of reduced dimensions is usually far less and are not correlated with each other

**C.** True. As this will reduce the false positives.

**D.** False. Semantically similar words like canine and dog have low similarity scores

**E.** False. Glove and word2vec embeddings are non-contextual. Eg. The embedding of the word ‘sack’ would remain the same regardless of whether the document was ‘sack of rice’ or ‘sack the employee’

**G.** True. In variable length encoding, it is important to specify that more bytes exist in the encoding, due to which continuation bytes are used to indicate a large sequence of bytes. These bytes start with 10.