JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST - 3 EXAMINATION- 2024

B.Tech-II Semester (CSE/IT/ECE/CE)

COURSE CODE(CREDITS): 22M11CI212 MAX. MARKS: 35

COURSE NAME: Deep Learning Techniques

COURSE INSTRUCTORS: HRI MAX. TIME: 2 hrs

---

***Note:*** *(a)All questions are compulsory.*

*(b)Marks are indicated against each question in square brackets.*

*(c) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems*

---

Q1. Compare Nestrov RMSProp, Adadelta and Adam gradient descent strategies [CO1][5]

Q2. Although nonlinearity can be achieved with the ReLU activation function, the key point is that the effects of max-pooling cannot be exactly replicated by strided convolutions either. At the very least, the two operations are not fully interchangeable. Why? [CO3][5]

Q3. Describe the architecture, forward propagation and backward propagation in an RNN. [CO4][4]

Q4. Describe the architecture of LSTM with a neat and clean diagram. Discuss the significance of all the gates used along with mathematical equations. Does the LSTM solve the issue of poor long-term memory in RNNs? [CO4][5]

Q5. Consider a document containing 100 words wherein the word cat appears 3 times. Now, assume we have 10 million documents and the word cat appears in one thousand of these. Now compute (a) TF, (b) IDF, and (c) TF-IDF [CO4][4]

Q6. Suppose a corpus contains the following two documents. [CO4][4]
The movie was very good.
The movie was not very good.
Compare BoW and N-Grams on the parameter "capturing contextual meaning".

Q7. Draw architecture of CBOW Word2vec model for a window size of 1 for the following document.
The movie was very good. [CO4][4]

Q8. Describe BERT architecture for word embedding. [CO4][4]