IIITV

End-Semester Examination (Autumn 2022)

Student's name : Urmila Rathore
Roll No: 2019511164
Time Duration : 3 hours
Total Mark : 65

Degree Program : B.Tech (CS) (VII Sem)
Course Title : Natural Language Processing
Course Code : IT 437
Date of Examination : 14 November, 2022.

Instructions:

1. This exam consists of 4 questions. Before answering, ensure that all the questions are present.

Q.1 Answer the following questions:

(a) What does the term "exploding gradients" refer to? What circumstances lead to "exploding gradients?"

(b) Why are "vanishing gradients" a more common problem in basic RNNs compared to feed forward networks?

(c) Name two methods for alleviating the vanishing gradient problem and explain why they work.

(d) Consider a recurrent neural network with a single hidden unit and a sigmoid activation, $h_m = \sigma(\theta h_{m-1} + x_m)$. Prove that if $|\theta| < 1$, then the gradient $\frac{\partial h_{m+k}}{\partial h_m}$ goes to zero as $k \to \infty$.

10 marks

Q.2 Consider the function:

$$f(x, y, z, v) = \exp\left(y^{-x} - \log(z)\right) + (y - z)^2 \cdot \log(v)$$

where log is the natural logarithm.

1. Draw the computational graph for f.
2. Evaluate the computational graph from (1) at the point $(0, 1, 2, 3)$. Fill in all intermediate values in your drawing.
3. Apply backpropagation to the computational graph from (1) evaluated at $(0, 1, 2, 3)$, making use of the values obtained from forward propagation in (2). Fill in all intermediate values.

Compute $\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial v}\right]$ with symbolic differentiation. How many computations are repeated when using this method for calculating ∇f ?

20 marks

try

Q.3 Desc... det ls following architectures:

100

(a) Skip-gram

Q.4 Consider the following training data

10 marks

⟨ s ⟩ am Sam ⟨/s⟩
⟨ s ⟩ Sam I am ⟨/s⟩
⟨ s ⟩ Sam I like ⟨/s⟩
⟨ s ⟩ Sam I do like ⟨/s⟩
⟨ s ⟩ do I like Sam ⟨/s⟩

Assume that we use a bigram language model based on the above training data.

1. What is the most probable next word predicted by the model for the following word sequences?

   (a) ⟨ s ⟩ Sam ...
   (b) ⟨ s ⟩ Sam I do ...
   (c) ⟨ s ⟩ Sam I am Sam ...
   (d) ⟨s⟩ do I like ...

12 marks

2. Which of the following sentences is better, i.e., gets a higher probability with this model?

   (a) ⟨s⟩ Sam I do I like ⟨/s⟩
   (b) ⟨s⟩ Sam I am ⟨/s⟩
   (c) ⟨s⟩ I do like Sam I am ⟨/s⟩

9 marks

3. Consider again the same training data and the same bigram model. *Compute the perplexity* of
   ⟨ s ⟩ I do like Sam

4 marks