

Natural Language Processing (CS22N)

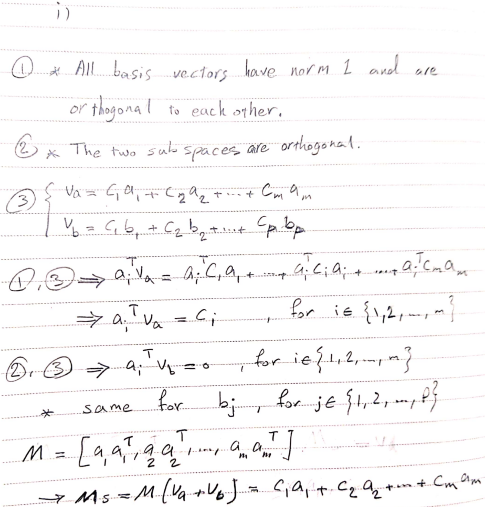
Student Name:

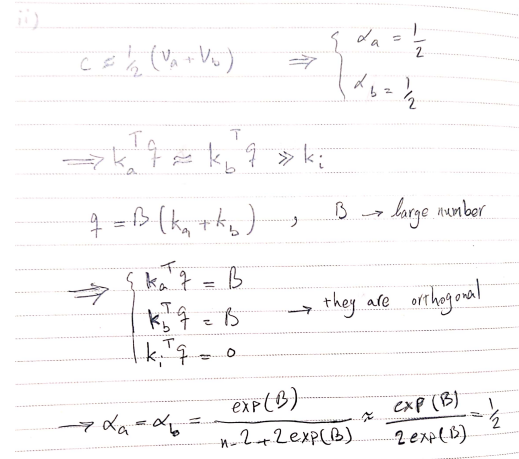
Babak Behkamkia

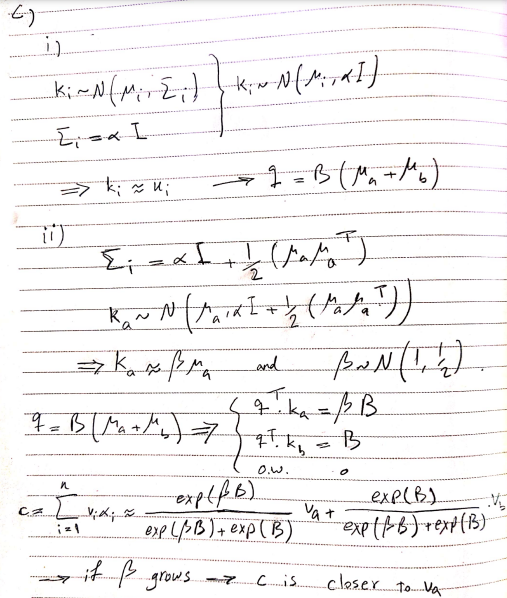
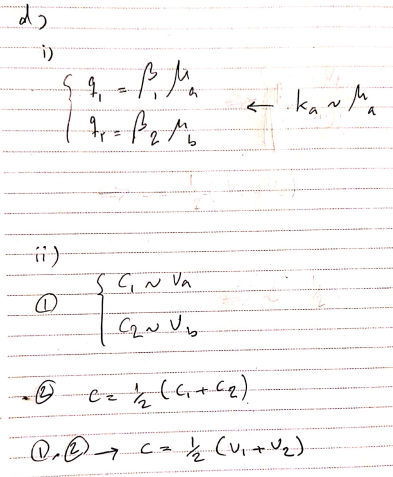
Instructor Name:

Dr. Sauleh Eetemadi

Spring 2023

1. **Attention exploration**
2. We use softmax to calculate the alpha values. Thus, these values will be a number between 0 and 1 and their sum would be 1. This shows these values are a probability distribution. Moreover, each alpha value assigned to a word in the sequence and shows the importance of that word in that sequence of words. Therefore, alpha scores are a categorical probability distribution.
3. ki q >> kj q where i, j ∈ {1, 2, …, n} and i!=j
4. In this case, the value of c would be very close to vi because ki q has the largest value and after applying the softmax function, the corresponding value would be 1 for it. Moreover, from the equation (1) we can understand that c = vi .
5. If the key value of a word is similar to q value of the chosen word, the product of those matrices will be large and after softmax, it will put almost all the weight on the corresponding alpha.
6. 



1. 
2. 
3. **Pretrained Transformer models and knowledge access**
4. Accuracy: 2%  
   London accuracy: 5%
5. Accuracy: 26.4%
7. Accuracy: 13%
8. **Considerations in pretrained knowledge**
9. Overall, pretrained models are better than non-pretrained models because they have trained on a large dataset previously. In other words, their weights are not random and they have learned some knowledge from the previous task.
10. 1. Trustworthiness: If an application generates unreal information like made up birthplaces, research papers, or websites, it will reduce the users trust on the application.
    2. Using wrong information: People may not notice the made up answer and use it in critical situations, which leads to a big problem.
11. Obviously, the model cannot determine the birthplace of a person that it never seen it before but with providing more information about that person for the model can help it in order to find similar individuals and make a prediction based on this information. For example, being angry most of the time may be exclusive to the people of a specific country.