Discussion Summary

Michael Wacey

For Doctor Walid Hassan

SEAS 8520: Deep Learning and NLP

August 17, 2024

George Washington University

# Request

Write a summary of the class discussions, and provide a reflection on what you learned from the topics discussed in class (2~3 pages)

Below is a copy of the discussion topics, overview, summary, and reflection.

# Discussion Topics

## Discussion 1

Discuss the factors that make Artificial Neural Networks (ANNs) superior in certain applications, while also exploring scenarios where logistic regression might be a more appropriate choice. Limit to 300 words.

## Discussion 2

Explore how gradient descent influences the learning and performance of neural networks. How does adjusting the learning rate affect model accuracy and training speed? Consider the challenges of choosing an appropriate learning rate and the strategies that might be employed to overcome these challenges. Limit to 300 words.

## Discussion 3

Discuss the challenges and strategies in designing neural network architectures that are complex enough to capture underlying patterns in data without overfitting. Explore the role of techniques such as dropout, early stopping, and regularization in achieving this balance. Limit to 300 words.

## Discussion 4

Discuss the basic components of CNNs and their roles as well as some challenges in training CNNs and strategies to overcome them. Give a real-world application of CNNs and its impact on that field.

## Discussion 5

Please read the first 2-3 pages of the paper listed under lecture 6 (An Empirical Study of Using Large Language Models for Unit Test Generation.... can be downloaded from https://arxiv.org/abs/2302.06527). What is the paper about? can you comment on the structure of the write-up of the paper (not the content)?

## Discussion 6

Transformers have fundamentally changed the landscape of NLP. In your own words, reflect on the core mechanisms that make Transformer models unique and effective. Compare and contrast with other models we have seen in class.

## Discussion 7

Retrieval Augmented Generation (RAG) is an innovative approach that combines the power of retrieval mechanisms with language generation models to enhance the quality and relevance of generated text. After conducting your research on RAG, discuss in your own words some of its use, benefits and drawbacks. To get credit for this discussion please cite your sources.

# Overview

The discussion questions were aligned to but extended the material being covered in class. It gave the students an opportunity to extend their understanding. I found this interesting and very useful. Oftentimes each student would find a subtly or even extremely different way of addressing the question. This was very important learning for me. I sometimes get focused on the idea that there is only one way to do something.

# Summary

As this is our third term in the program, it made sense that the first discussion would jump right into the material. In prior classes, we had many machine learning techniques but had not directly discussed Artificial Neural Networks (ANN). This discussion was a great opportunity to tie what we had learned previously to the material that we were going to cover in this class.

In discussion two we started to delve into the details of how neural networks work. How they are trained using gradient decent and a learning rate was the specific question. For me, this was a good chance to think through the implications of decisions such as learning rate. Generally, I find it useful to visualize how something works. In responding to this discussion question I spent a fair amount of time thinking through how gradient descent works and what the learning rate does.

For the next discussion, we brought it back to important topics that we had covered with techniques other than neural networks. Overfitting is a common problem that I expect is fairly easy to fall into. Since it makes your accuracy look good, it is important to be aware of overfitting and be prepared to address it. For this discussion, we covered some previous methods such as L1 and L2 regularization but also some new methods such as Dropout. It was nice to put everything in perspective.

Now that we had a good understanding of Neural Networks, discussion four was the start at looking at ways to architect NN based solutions. In this case we discussed the pros and cons of Convolutional Neural Networks (CNN). The architectures of neural networks are heavily dependent on solving simultaneous equations. It seems to me that I am just starting to understand how all the parts fit together. As part of answering this discussion question, I read several papers about CNNs. While I could understand most of the material, I am quite sure that I am not yet at the level where I could create it.

Discussion five took us in a very different direction. Here we were to read a paper for its structure, not content. Having read many papers in the past for content, this gave me a different view on the paper at hand. I looked for how the authors connected the parts into a coherent whole. I thought about the rational for each part of the paper. I wondered if the parts were missing anything. It was a good experience.

Finally, we get to discussion six. This is what I really wanted to know when we started this class. I had read but not understood the “Attention Is All You Need” paper. Now I was able to read it and really begin to understand it. For all other discussions, I read and appreciated other student’s responses. This discussion was the one where I really wanted to hear the other student’s perspectives.

Discussion seven was a good opportunity to do some research on a topic and try to put it in perspective. I found this very interesting. In doing my research I found both the theoretical underpinnings of RAG and the practical applications. It was interesting to see the promise spelled out in the theory. Then the practical papers went in two directions. One was a product view that was careful to state the benefits and the limitations. The second was a marketing view that said RAG can solve every problem ever encountered. This was very illuminating to me.

# Reflection

I found the discussions to be beneficial in this class for several reasons. First, they reinforced the material that we were covering in class. Second, they gave me an opportunity to see how other people address the same problem. This was very useful for broadening my approach to problem solving. Overall, a great learning experience.

Some specific things that I learned while responding to the discussion questions were that there is rarely one answer, the use of multiple sources is critical to understanding a topic, and that I rarely know as much as I think I do on a topic. Each of these are discussed in more detail below.

As I prepared my answer to each question, I was rarely sure that I had the “right” answer. Generally, I felt good about my answers, but I always felt there was something lacking. As I read posts from other students, I would see some of the things I was missing. It seems to me that these items will stay with me.

For most of the discussions I had to do research beyond the lecture slides and book. This was either because the question stated it or because I wanted to round out my knowledge. While I have experienced it before, in these cases I often found both subtle and substantial differences. It seems to me that the areas of AI and Machine Learning are growing so quickly that this is inevitable.

The more you know the less you understand. I think that is a rule of thumb. It certainly has been for me in this course. I heard about the “Attention Is All You Need” paper at the beginning of this program in January. So, I read it. But only when I read it as part of this course did I really know how little of it I still understand. I have so much more to learn.

In conclusion, learning is a journey and not a destination. This course and these discussions are an important part of my learning journey. I am so happy to be on this journey and I thank you for being one of my guides.