

Question 1

1.

The architecture of my sentiment classifier is as below:

```
SentimentClassifier(  
    (token_embeddings): Embedding(38482, 300, padding_idx=0)  
    (dropout): Dropout(p=0.5, inplace=False)  
    (linear): Linear(in_features=300, out_features=1, bias=True)  
    (transformer_encoder): TransformerEncoder(  
        (layers): ModuleList(  
            (0-1): 2 x TransformerEncoderLayer(  
                (self_attn): MultiheadAttention(  
                    (out_proj): NonDynamicallyQuantizableLinear(in_features=300, out_features=300,  
bias=True)  
                )  
                (linear1): Linear(in_features=300, out_features=512, bias=True)  
                (dropout): Dropout(p=0.5, inplace=False)  
                (linear2): Linear(in_features=512, out_features=300, bias=True)  
                (norm1): LayerNorm((300,), eps=1e-05, elementwise_affine=True)  
                (norm2): LayerNorm((300,), eps=1e-05, elementwise_affine=True)  
                (dropout1): Dropout(p=0.5, inplace=False)  
                (dropout2): Dropout(p=0.5, inplace=False)  
            )  
        )  
    )  
)
```

2.

The accuracy on test set is 0.8658 after 50 epochs.

3.

The accuracy on test set of TF-IDF, Word2vec and Transformer are 0.9055, 0.8714 and 0.8658. Obviously, Transformer is less effective than the others if we only focus on accuracy. The below is advantages and limitations of these models:

| Model | Advantages | Limitations |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TF-IDF | <ul style="list-style-type: none">① Simple and straight: easy to understand and achieve② Effective on computation: low computation cost③ Efficient on performance: its accuracy is the best than the others in this case although transformer could be better according to experience | <ul style="list-style-type: none">① Mainly focus on frequency of words and ignore the relationship between words② Unable to understand different meanings of the same word |
| Word2vec | <ul style="list-style-type: none">① It can capture the semantic information of words② Can better handle words that never met | <ul style="list-style-type: none">① Unable to understand different meanings of the same word② Need more computation compared to TF-IDF |
| Transformer | <ul style="list-style-type: none">① Able to understand different meanings of the same word and the relationship between words② Transformer can handle parallel calculations of sequence data and improve calculation efficiency. | <ul style="list-style-type: none">① Need more computation: it's more complex than the others② Too complex to adjust parameters: it has lots of parameters, it's more difficult to get best parameters because it takes a long time to get result③ Suitable for long paragraphs or complex articles: because movie reviews are all relatively short, it always doesn't need long-distance dependencies and complex semantic relationships to get sentiments. We can always get audiences' sentiments from some words from movie reviews. That's the main reason why Transformer's accuracy is lower the others. |

Question 2

1. The final score of my model is 0.71431. My model is based on DistilBERT to solve the classification problem because it's more effective on computation and good on performance. I used pretrained model and fine-tuning the model using the train data to classify the text in test data. I don't use the labels "horror", "Science Fiction" and so on. I just make the label parameter as 4 to train and make predictions. The following are some parameters: max_length=300, batch_size=64, lr=3e-5, and epochs = 20.

2. Training: The model is trained using the AdamW optimizer with the learning rate set to $5e-5$. The model goes through all the training data in each epoch, calculates the loss between the model's output and the actual label, and uses this loss to update the model's weights. The learning rate scheduler updates the learning rate in each step.

Testing: I have tried some parameters for better performance like `max_length = 256`, `batch_size = 18`, `lr=5e-5`, `epochs = 10` and so on. The best performance is the parameters above (Question 1).

3. I have tried BERT and XLNet. I get a score from XLNet with 3 epochs and the result is 0.69897, which is lower than the result of DistilBERT. I didn't get a result from BERT because it needs too much computation and my Google colab just run out of memory.

4. I didn't get a result from BERT because it needs lots of computation. Although BERT and XLNet are all big model, BERT is more complex and need more time. Although they may be better than DistilBERT, the cost is huge. So, I just use DistilBERT to train and make a prediction, and it's easier to adjust parameters.

Question 3

Through LLM

Question:

Is nuclear power plant eco-friendly?

Answer:

- No, nuclear power plants produce radioactive waste that can harm the environment and human health.
- The process of generating electricity through nuclear reactions also releases greenhouse gases.
- Nuclear power plants require large amounts of water for cooling, which can impact local ecosystems.
- Accidents at nuclear power plants, such as Chernobyl and Fukushima, have had severe environmental and health consequences.

Answer the question below in text using about 50 words, your answer should be in bullet points.

Question:

How to stay safe during severe weather?

Answer:

- Stay informed through local news and weather alerts
- Create a family emergency plan
- Stock up on essential supplies
- Stay indoors and away from windows
- Avoid driving during severe weather
- Keep your phone charged and have a backup power source.

Through RAG

Query: Is nuclear power plant eco-friendly?

Top-3 documents (similarity scores):

data/gov/documents/05/G00-05-2684553 0.4045

data/gov/documents/01/G00-01-1806077 0.3790

data/gov/documents/00/G00-00-0799717 0.3509

The Nuclear Science Division at Lawrence Berkeley National Laboratory (LBNL) is a leading research center in the field of nuclear science, with a wide range of programs and activities including nuclear theory, relativistic nuclear collisions, heavy ion

The U.S. Department of Energy Office of Nuclear Energy, Science and Technology has an online portal called FRAME (Federal Register Application Management Environment) that provides information on nuclear energy plant optimization, including research areas, notices

The T-2 Nuclear Information Service is a website run by Group T-16 of the Theoretical Division of the Los Alamos National Laboratory, which provides nuclear data and information. The site has educational materials, datasets, computer

Given the following context, answer the question below in text using about 50 words. Your answer should be in bullet points.

Context:

Document: data/gov/documents/05/G00-05-2684553

Summary: The Nuclear Science Division at Lawrence Berkeley National Laboratory (LBNL) is a leading research center in the field of nuclear science, with a wide range of programs and activities including nuclear theory, relativistic nuclear collisions, heavy ion

Document: data/gov/documents/01/G00-01-1806077

Summary: The U.S. Department of Energy Office of Nuclear Energy, Science and Technology has an online portal called FRAME (Federal Register Application Management Environment) that provides information on nuclear energy plant optimization, including research areas, notices

Document: data/gov/documents/00/G00-00-0799717

Summary: The T-2 Nuclear Information Service is a website run by Group T-16 of the Theoretical Division of the Los Alamos National Laboratory, which provides nuclear data and information. The site has educational materials, datasets, computer

Question:

Is nuclear power plant eco-friendly?

Answer:

- Yes, nuclear power plants are considered to be one of the cleanest forms of energy production.
- They produce no greenhouse gas emissions during operation, which can help reduce carbon footprint and combat climate change.
- Nuclear

Query: How to stay safe during severe weather?

Top-3 documents (similarity scores):

data/gov/documents/04/G00-04-0688776 0.1400

data/gov/documents/01/G00-01-0938351 0.1384

data/gov/documents/03/G00-03-2373928 0.0699

The National Severe Storms Laboratory's Bibliographic Database (NBD) contains over 1000 references on hazardous winter weather from 1882-1998. The database is searchable and

The National Severe Storms Laboratory (NSSL) is a NOAA research laboratory that investigates all aspects of severe weather. Based in Norman, Oklahoma with staff across the US, NSSL partners with the National Weather Service to improve

The California Integrated Waste Management Board is responsible for protecting public health and safety and the environment through waste management. The board works with local governments, industries, and the public to reduce waste disposal and ensure safe landfills.

Given the following context, answer the question below in text using about 50 words. Your answer should be in bullet points.

Context:

Document: data/gov/documents/04/G00-04-0688776

Summary: The National Severe Storms Laboratory's Bibliographic Database (NBD) contains over 1000 references on hazardous winter weather from 1882-1998. The database is searchable and

Document: data/gov/documents/01/G00-01-0938351

Summary: The National Severe Storms Laboratory (NSSL) is a NOAA research laboratory that investigates all aspects of severe weather. Based in Norman, Oklahoma with staff across the US, NSSL partners with the National Weather Service to improve

Document: data/gov/documents/03/G00-03-2373928

Summary: The California Integrated Waste Management Board is responsible for protecting public health and safety and the environment through waste management. The board works with local governments, industries, and the public to reduce waste disposal and ensure safe landfills.

Question:

How to stay safe during severe weather?

Answer:

- Stay informed about weather conditions through local news and weather radios
- Have a plan that includes evacuation routes and emergency contact numbers
- Stock up on essential supplies such as food, water, and medications
- Stay

Answers:

1. For question 1, they provide opposite answers for the same questions. Directly through LLM, it gave negative answer and explained some reasons in different aspects. Through RAG, it gave positive answers and explained why. In question 2, they give almost the same answers and introduce some aspects.

Main Differences:

- Directly through LLM produce more coherent answers and more consistent with topic.
- RAG provided answers according to the summaries of three most relative documents, while directly through LLM doesn't.
- The answers from RAG may be incomplete while directly through LLM is.

2.

The answers are follows:

| Name | Advantages | Disadvantages |
|----------------------|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| Standard IR System | 1. Can be highly precise in returning relevant documents. 2. Less computation (restricted to relevant documents) | 1. Users must find the information that they want in returned documents 2. Hard to understand the context of queries |
| Directly through LLM | 1. Can answer any kinds of questions | 1. Need lots of computations |

| | | |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | <ul style="list-style-type: none"> 2. Able to understand context of queries 3. Can generate answers with high relevance | <ul style="list-style-type: none"> 2. Need lots of relative data to make the answers accurate |
| RAG | <ul style="list-style-type: none"> 1. Can make the answers more relevant according to provide relevant documents 2. Increase the reliability of answers through providing relevant documents 3. Reduce biases of LLM | <ul style="list-style-type: none"> 1. It's more complex to setup and optimize 2. It's difficult to decide the length of summary, it should not be too long or too short 3. Relative documents could be conflict with the truth, which can produce wrong answers 4. Need even more computation |