

Understanding Large Language Models: Final Project

Benjamin Tremblay
bptremblay@wm.edu

1 Introduction

Throughout the semester, we have learned how pretrained, then fine-tuned Large Language Models can be used to accurately predict an expected response with varying degrees of success. For the final project, I wanted to answer the question: how well does fine-tuning a model go when given in-context knowledge and meant to act in a specific way? Using a dataset that pulls from a call center, I made a LLM model that predicts the third response by the call center agent, given the full conversation (before and after the masked statement). This challenges the AI's ability to think in-context, and outside of coding. For the code, I used my code in <https://github.com/bentrem38/Agent-Reponse-model> as a starting point, as fine-tuning Hugging Face models all follow a certain pattern, allowing more efficient use of time.

My motivation for choosing this goal stems from the AI projects I have already worked on this year. While I have seen a fine-tuned models' capability in code completion, I have yet to personally try out a fine-tuned model for Natural Language Processing outside of coding. I want to see how well a model can predict a human's response to conversation when given a specific context. The capabilities required to do well on predicting the agent's response is a good understanding of customer support service interactions, how this specific call center handles customers, and an ability to handle test cases not seen before. While working on the project, I used an ipynb file to not redundantly rerun code, but the `agent_response.py` file is the final working version of the code. The project can be found in the following GitHub repository: <https://github.com/bentrem38/Agent-Reponse-model>

2 Implementation

Model Description: I decided to use Google/Flan-t5-base as my model to fine-tune. While Flan-t5-base has a huge 250 million parameters, this is actually a lot smaller than many other pre-trained models, allowing for an easier time with computation and making it easier to train. It is also pretrained on numerous NLP tasks, giving it a good basis to build off. It runs off t5 architecture, and is a seq2seq transformer (encoder-decoder). It runs on float32, and changing to binary16 or bfloat16 instructions was unnecessary, as the computation times were manageable. I had no trouble computing the model locally on my desk-top PC, but perhaps running it on a laptop would have been harder, leading me to running binary16 instead.

Data Preparation: Using HuggingFace, I found and used the `llama2_transcripts_healthcare_callcenter` dataset. Numbering 1000 entries, it has written scripts of phone conversations of a healthcare call center. In order to be used for training a model, I flattened the conversation, masked the third response from the agent at the call center, replacing it with `<MASK>`, and took the masked response at a `target_block` for the AI to predict. 800 examples were used for the training, while 100 were used for both validation and testing.

Before Data Preparation:

Original Script:	<p><s>[INST] Classify the following call transcript:</p> <p>Agent 3: Thank you for calling HealthHarbor, my name is Agent 3. How can I assist you today?</p> <p>Customer: Hi Agent 3, my name is Emma Johnson. I've been experiencing some symptoms lately and I wanted to seek medical advice or get a symptom assessment.</p> <p>Agent 3: I'm sorry to hear that, Emma. I'll do my best to help you. Can you please describe the symptoms you've been experiencing?</p> <p>Customer: Sure. I've been having a persistent headache for the past few days, and it's been accompanied by dizziness and occasional nausea. I'm not sure what could be causing it.</p> <p>Agent 3: I understand your concern, Emma. Headaches can have various causes. Have you experienced any recent changes in your lifestyle or any other symptoms besides the headache, dizziness, and nausea?</p> <p>Customer: No major lifestyle changes, but I have noticed that my vision seems a bit blurry at times. And I've been feeling more fatigued than usual.</p> <p>Agent 3: Thank you for sharing that information, Emma. Blurry vision and fatigue can also be related to your symptoms. It's important to consider all these factors for a proper assessment. Based on your symptoms, I would recommend consulting with a healthcare professional. They will be able to conduct a thorough examination and provide a more accurate diagnosis.</p> <p>Customer: I was hoping to get some advice before scheduling a doctor's appointment. Is there anything I can do to relieve these symptoms in the meantime?</p> <p>Agent 3: While I'm not a doctor, I can offer some general suggestions. You could try applying a cold or warm compress to your forehead to see if it helps with the headache. It's also important to stay hydrated and get enough rest. However, it's crucial to understand that these measures may not address the underlying cause of your symptoms. A medical professional will be able to provide a more targeted approach.</p> <p>Customer: I understand. I'll make sure to schedule an appointment with a doctor as soon as possible. Thank you for your advice.</p> <p>Agent 3: You're welcome, Emma. It's always better to be safe and have a professional evaluation. Is there anything else I can assist you with today?</p> <p>Customer: No, that's all for now. Thank you for your help, Agent 3.</p> <p>Agent 3: You're welcome, Emma. I hope you feel better soon. If you have any more questions or concerns, don't hesitate to reach out. Take care!</p> <p>Customer: Thank you, Agent 3. Have a great day!</p> <p>Agent 3: You too, Emma. Goodbye! [/INST] Medical Advice or Symptom Assessment </s></p>
-------------------------	---

After Data Processing:

<p>Processed Script:</p>	<p><s>[INST] Classify the following call transcript: Agent 3: Thank you for calling HealthHarbor, my name is Agent 3. How can I assist you today? Customer: Hi Agent 3, my name is Emma Johnson. I've been experiencing some symptoms lately and I wanted to seek medical advice or get a symptom assessment. Agent 3: I'm sorry to hear that, Emma. I'll do my best to help you. Can you please describe the symptoms you've been experiencing? Customer: Sure. I've been having a persistent headache for the past few days, and it's been accompanied by dizziness and occasional nausea. I'm not sure what could be causing it. Agent 3: <MASK> Customer: No major lifestyle changes, but I have noticed that my vision seems a bit blurry at times. And I've been feeling more fatigued than usual. Agent 3: Thank you for sharing that information, Emma. Blurry vision and fatigue can also be related to your symptoms. It's important to consider all these factors for a proper assessment. Based on your symptoms, I would recommend consulting with a healthcare professional. They will be able to conduct a thorough examination and provide a more accurate diagnosis. Customer: I was hoping to get some advice before scheduling a doctor's appointment. Is there anything I can do to relieve these symptoms in the meantime? Agent 3: While I'm not a doctor, I can offer some general suggestions. You could try applying a cold or warm compress to your forehead to see if it helps with the headache. It's also important to stay hydrated and get enough rest. However, it's crucial to understand that these measures may not address the underlying cause of your symptoms. A medical professional will be able to provide a more targeted approach. Customer: I understand. I'll make sure to schedule an appointment with a doctor as soon as possible. Thank you for your advice. Agent 3: You're welcome, Emma. It's always better to be safe and have a professional evaluation. Is there anything else I can assist you with today? Customer: No, that's all for now. Thank you for your help, Agent 3. Agent 3: You're welcome, Emma. I hope you feel better soon. If you have any more questions or concerns, don't hesitate to reach out. Take care! Customer: Thank you, Agent 3. Have a great day! Agent 3: You too, Emma. Goodbye! [/INST] Medical Advice or Symptom Assessment </s></p>
<p>Target_block:</p>	<p>I understand your concern, Emma. Headaches can have various causes. Have you experienced any recent changes in your lifestyle or any other symptoms besides the headache, dizziness, and nausea?</p>

Evaluation Method: I will be using Recall-Oriented Understudy for Gisting Evaluation (ROUGE) as my main evaluation method for this project. Exact matches are too strict, and Bleu does not pick up on similar syntax, while ROUGE takes into consideration the similarity of the text.

Obtaining a Baseline: To have a standard for comparison, I input the masked test data into the original not fine-tuned model. Unsurprisingly, the untrained model was very bad at predicting the masked statement, having an overall score of just a little over 5%. Although trained for NLP tasks, without in-context knowledge, flan-t5-base is not sure what to do with the given input.

Baseline Example:

Processed Script	Target_block	Untrained Model prediction
<s>[INST] Classify the following call transcript: Agent 4: Thank you for calling HealthHarbor. This is Agent 4. How can I assist you today? Customer: Hi, Agent 4. My name is John Smith, and I have a few questions regarding my insurance and billing. Agent 4: Hello, Mr. Smith. I'll be happy to help you with that. Could you please provide me with your patient ID or any other identifying information? Customer: Sure, my patient ID is 123456789. Agent 4: <MASK> Customer: I recently received a bill from HealthHarbor, and I wanted to verify the charges. There seems to be an error in the amount billed. Agent 4: I apologize for any inconvenience, Mr. Smith. Let me take a look at your account and investigate this for you. Can you please provide me with the details of the charges you are questioning? Customer: It's regarding a surgery I had last month. The bill shows a charge for a procedure that I don't recall having. I want to make sure that this is accurate. Agent 4: I understand your concern, Mr. Smith. Let me pull up your account and review the details. While I do that, could you please provide me with the date of the surgery and any other relevant information? Customer: The surgery was on April 15th, and I had it in the HealthHarbor main facility. The procedure mentioned on the bill is a knee replacement, but I actually had a hip replacement. Agent 4: Thank you for providing that information, Mr. Smith. I see the discrepancy now. It seems like there was an error in coding. I will make a note of it and have our billing department correct the charges on your account. You will receive an updated bill reflecting the accurate charges. Customer: Thank you, Agent 4. I appreciate your help in resolving this matter. Is there anything else I need to do or provide? Agent 4: You're welcome, Mr. Smith. I'm glad I could assist you. You don't need to do anything else at this point. You will receive an updated bill within the next few weeks. Is there anything else I can assist you with today? Customer: No, that was my main concern. Thank you for your	Thank you, Mr. Smith. How can I assist you today?	Customer: Hello

assistance, Agent 4. Agent 4: It was my pleasure assisting you, Mr. Smith. If you have any further questions or concerns, feel free to reach out to us. Have a great day! Customer: Thank you, Agent 4. You too. Goodbye. Agent 4: Goodbye, Mr. Smith. Take care! [/INST] Insurance and Billing Inquiries </s>		
--	--	--

In this example, the untrained model's prediction is nonsensical, and has no understanding of the context of the dialogue, leading to a confusing prediction. Hopefully, with in-context training, the model predictions will improve significantly

Untrained model ROUGE Scores:

Rouge1:	0.0663
Rouge2	0.0168
RougeL	0.0593
RougeLsum	0.0591
Average ROUGE Score:	0.0503776024659935

3. Fine-tuning: When fine-tuning the model, I first tried to run the full model locally to see if I needed to change my approach to a parameter efficient fine-tuning method, but it ended up running fine. The whole run time of the training was under half an hour, as with a small dataset, it does not take for the fine-tuning process to finish. I chose eval_loss as the way to determine when to stop fine-tuning the model, as this allows for an optimized fitting to the data, stopping whenever it starts overfitting after two epochs. After deliberating, I chose for the model to run for 10 epochs, or until eval_loss stopped it early, as this allowed for it to continue fitting the data until it was just right. While I did not use any checkpoints during the training process, I did save the completed model to my local drive, as this saves time from reloading your data if ever interrupted. My model stopped early after 5 epochs, hitting a minimum of validation loss, and pulling epoch that has the lowest validation loss

Epoch	Training Loss	Validation Loss
1	0.094700	0.111950
2	0.082600	0.110143
3	0.079200	0.109998
4	0.068600	0.114833
5	0.065000	0.115739

Model Generation: To run the model’s outputs efficiently, I broke up the test cases into batches, and ran them batch by batch. This stopped any memory errors from popping up, which I have struggled to fix before while generating model outputs.

Fine-tuning Results: After fine-tuning the model, the model underwent an enormous positive change in its ROUGE evaluation.

Trained model ROUGE Scores:

Rouge1	0.633
Rouge2	0.4708
RougeL	0.5879
RougeLsum	0.5891
Average ROUGE Score	0.5701928157335951

Impressively, the model overall does a great job of keeping on task and addressing the concerns of the caller. Many of the test case results are similar to the target_block

Processed Script	Target_block (agent response)	Trained Model prediction:	Untrained Model Prediction
<p><s>[INST] Classify the following call transcript: Agent 3: Thank you for calling HealthHarbor, my name is Agent 3. How can I assist you today? Customer: Hi, Agent 3. My name is Emily Johnson. I've been experiencing some symptoms lately and I'm not sure if I should be concerned. Agent 3: Hello, Emily. I'm here to help. Can you please let me know what symptoms you've been experiencing? Customer: Sure. I've been having frequent headaches and occasional dizziness. It's been happening for the past week or so. Agent 3: <MASK> Customer: No, not really. I haven't experienced any nausea or vomiting. It's mainly just the headaches and dizziness. Agent 3: Okay, thank you for sharing that. Have you had any recent changes in your lifestyle or any significant stressors? Customer: Not really. My lifestyle has been pretty consistent, and I haven't been under any excessive stress lately. Agent 3: Alright. In</p>	<p>I understand how concerning that can be. Have you noticed any other symptoms accompanying the headaches and dizziness?</p>	<p>I understand your concern, Emily. Headaches and dizziness can be quite bothersome. Have you experienced any other symptoms along with these?</p>	<p>Customer: Hello,</p>

<p>this case, it's always a good idea to consult with a healthcare professional. However, based on the symptoms you've described, it is possible that your headaches and dizziness could be related to tension or stress. It's important to manage stress levels and try relaxation techniques such as deep breathing exercises or meditation. If the symptoms persist or worsen, I recommend scheduling an appointment with your primary care physician for a thorough evaluation. Customer: That makes sense. I'll definitely try those relaxation techniques and monitor my symptoms. If they continue, I'll make an appointment with my doctor. Agent 3: That's a wise decision, Emily. It's always better to be safe and get professional advice. Is there anything else I can assist you with? Customer: No, that was all. Thank you for your help, Agent 3. I appreciate your time and advice. Agent 3: You're welcome, Emily. I'm glad I could assist you. If you have any further questions or concerns, don't hesitate to reach out. Take care and have a great day! Customer: Thank you, Agent 3. You too. Goodbye. Agent 3: Goodbye, Emily. Have a wonderful day! [/INST] Medical Advice or Symptom Assessment</p>			
--	--	--	--

In this example, although not getting an exact match, the syntax and similarity between the agent response and the trained model prediction are undeniable. I would call this a success for the fine-tuned model. On the other hand, the untrained model did not even give a coherent response, with most of its responses similar to this result, further showing how big of an impact the training had on the model.

Comparison to untrained model in previous example:

Processed Script	Target_block	Untrained_model prediction	Trained_model prediction
<p><s>[INST] Classify the following call transcript: Agent 4: Thank you for calling HealthHarbor. This is Agent 4. How can I assist you today? Customer: Hi, Agent 4. My name is John Smith, and I have a few questions regarding my insurance and billing. Agent 4: Hello, Mr. Smith. I'll be happy to help you with that. Could you please provide me with your patient ID or any other identifying information? Customer: Sure, my patient ID is 123456789. Agent 4:</p>	<p>Thank you, Mr. Smith. How can I assist you today?</p>	<p>Customer: Hello</p>	<p>Thank you, Mr. Smith. How can I assist you with your insurance and billing inquiries today?</p>

<p><MASK> Customer: I recently received a bill from HealthHarbor, and I wanted to verify the charges. There seems to be an error in the amount billed. Agent 4: I apologize for any inconvenience, Mr. Smith. Let me take a look at your account and investigate this for you. Can you please provide me with the details of the charges you are questioning? Customer: It's regarding a surgery I had last month. The bill shows a charge for a procedure that I don't recall having. I want to make sure that this is accurate. Agent 4: I understand your concern, Mr. Smith. Let me pull up your account and review the details. While I do that, could you please provide me with the date of the surgery and any other relevant information? Customer: The surgery was on April 15th, and I had it in the HealthHarbor main facility. The procedure mentioned on the bill is a knee replacement, but I actually had a hip replacement. Agent 4: Thank you for providing that information, Mr. Smith. I see the discrepancy now. It seems like there was an error in coding. I will make a note of it and have our billing department correct the charges on your account. You will receive an updated bill reflecting the accurate charges. Customer: Thank you, Agent 4. I appreciate your help in resolving this matter. Is there anything else I need to do or provide? Agent 4: You're welcome, Mr. Smith. I'm glad I could assist you. You don't need to do anything else at this point. You will receive an updated bill within the next few weeks. Is there anything else I can assist you with today? Customer: No, that was my main concern. Thank you for your assistance, Agent 4. Agent 4: It was my pleasure assisting you, Mr. Smith. If you have any further questions or concerns, feel free to reach out to us. Have a great day! Customer: Thank you, Agent 4. You too. Goodbye. Agent 4: Goodbye, Mr. Smith. Take care! [/INST]</p> <p>Insurance and Billing Inquiries</p> <p></s></p>			
---	--	--	--

The fine-tuned model prediction is almost an exact match to the target_block, a far cry from the two word response of the untrained model. Overall, the model's fine-tuning was a major success, showing a huge improvement in predicting the

target_block. Further, showing that AI models made for Natural Language Processing are not only useful for coding, but can also replicate human responses in conversation, leaving the door open for many possibilities in the future.