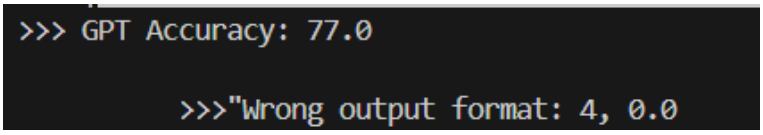
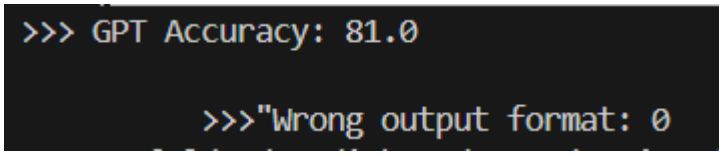
MLDS GENERATIVE AI – ASSIGNMENT 2

Prompt Engineering Accuracy Results:

GPT 3.5 without prompt engineering



GPT 3.5 with prompt Engineering:



OPRO:

To check if the algorithm was working as intended, I ran OPRO for 3 iterations with just 10 training examples. The optimized prompt and training accuracy score are given below:

OPRO Prompt 1:

Best prompt found with accuracy: 0.250

Prompt: To effectively tackle any math word problem, follow these systematic problem-solving steps:

1. Begin by carefully reading the problem to identify the key information provided.

2. Determine the type of mathematical operation needed (addition, subtraction, multiplication, division) based on the relationships between the quantities mentioned.

3. Break down the problem into smaller, more manageable parts.

4. Solve each part step by step, showing all your work and calculations.

5. Combine the individual solutions to find the final answer accurately.

Report your final answer with 4 hashtags at the end.

Then I ran the OPRO algorithm for 10 training iterations with 30 training examples provided. There is a significant increase in the training accuracy score after making this change:

OPRO Prompt 2:

Best prompt found with accuracy: 0.806

Prompt: To effectively conquer any math word problem and derive the correct solution, adhere to this structured step-by-step strategy:

1. Begin by thoroughly understanding the problem statement to grasp the given information and identify the unknown quantity awaiting calculation.

2. Break down the problem into smaller, manageable components based on the required mathematical operations (addition, subtraction, multiplication, division).

3. Determine the suitable mathematical operation needed for each part based on the context provided in the problem scenario.

4. Formulate precise equations or expressions to accurately represent the relationships between the quantities involved.

5. Progressively solve each component step by step, ensuring accurate calculations are performed at every stage.

6. Verify your solution by checking that it satisfies all conditions specified within the problem statement.

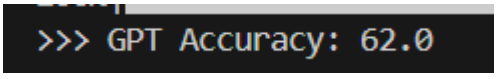
7. Present your final answer with confidence and clarity, followed by 4 hashtags for easy identification.

Follow these structured steps diligently to navigate through diverse math word problems successfully and reach the correct solution. Let's now proceed to solve the math word problem step by step.

OPRO Evaluation:

Finally, I tested the optimized OPRO prompt using the holdout test set that was used to evaluate the earlier prompt engineering techniques.

The final evaluation accuracy is:



We see a reduction in the accuracy score compared to other methods. However, the number of training examples recommended in the OPRO paper is 50, and training should be run for around 30 iterations ideally. I was not able to do this due to extended runtime (almost 9h). If these changes are made, the OPRO algorithm converge to the most optimal prompt which will deliver a higher accuracy on the test set.