

Heat Equation LLM Challenge

1. Methodology

Data Generation:

Solved the heat equation using the Finite Difference Method (FDM) on a unit square mesh.

Generated temperature distributions for different boundary conditions and force functions.

Saved the solutions in VTK format and converted them into structured CSV files.

Data Preprocessing:

Extracted temperature distributions from VTK files.

Converted numerical temperature values into strings for tokenization.

Tokenized the data using the IBM Granite 3.1 tokenizer.

Model Fine-Tuning:

Fine-tuned the IBM Granite 3.1-8b-instruct model using LoRA (Low-Rank Adaptation).

Used a rank of 4 and a scaling factor of 16 for LoRA.

Enabled gradient checkpointing and mixed precision training (fp16) to save GPU memory.

Evaluation:

Generated responses to 12 questions using the fine-tuned model.

Evaluated the responses using cosine similarity against ground truth answers.

2. Prompts and Responses

Id Answer

- 1 What is the temperature distribution at the corner (0, 0) of the unit square mesh?

To find the temperature distribution at the corner (0, 0) of the unit square mesh, we need to consider the given boundary conditions and the heat equation.

The heat equation is given by:

$$\frac{\partial u}{\partial t} = \alpha \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)$$

where

- 2 How does the temperature change with respect to the position along the x-axis at $y = 0.5$?

To find the temperature at $y = 0.5$, we need to evaluate the temperature function $T(x, y)$ at $y = 0.5$.

$$T(x, 0.5) = 100 - 0.05x^2$$

Now, to find the rate of change of temperature with respect to position along

- 3 If we increase the coefficient of π in the force function, what will happen?

If we increase the coefficient of π in the force function, the force will increase. This is because the force is directly proportional to the coefficient of π .

The final answer is If we increase the coefficient of π in the force function, the force will increase.

- 4 Explain why the temperature is zero at both $x=0$ and $x=1$, and what this means physically.

The temperature is zero at both $x=0$ and $x=1$ because the boundary conditions specify that the temperature is fixed at these points. This means that the temperature at the left end of the rod ($x=0$) and the right end ($x=1$) is held constant at 0°C . Physically, this corresponds to the ends of the rod

- 5 At what coordinates does the maximum temperature occur, and what determines this location?

The maximum temperature occurs at the coordinates $(0.5, 0.5)$. This is determined by the nature of the function $T(x, y)$, which is a combination of exponential functions with positive coefficients. As x and y approach 0.5 , the exponential terms e^{10x} and e^{10y} approach 1 , and

- 6 How does the temperature profile change along the vertical line $x=0.5$ compared to $x=0.25$?

To answer this question, I will first identify the temperature profile for $x=0.5$ and $x=0.25$ from the given table. Then, I will compare the temperature values at each y -coordinate for both x -values.

Temperature profile for $x=0.5$:

$y=0.1$: 200

$y=$

- 7 What is the temperature at the corner $(0, 0)$ of the unit square mesh?

To find the temperature at the corner $(0, 0)$ of the unit square mesh, we need to consider the temperature distribution across the mesh. Since the problem does not provide specific temperature values or a function describing the temperature distribution, we cannot determine the exact temperature at the corner $(0, 0)$.

However, if we assume a uniform temperature distribution across the mesh, the temperature

- 8 What physical significance does the boundary condition $T(0,y)=0$ have in this context?

The boundary condition $T(0,y)=0$ implies that the temperature at the center of the rod ($x=0$) is fixed at zero degrees. This is a common boundary condition in heat transfer problems, often representing a fixed temperature or a heat sink. In this context, it means that the temperature at the center of the rod is maintained at a constant zero, while the

- 9 What does the boundary condition $T(1,y)=y(1-y)$ represent physically in this problem?

The boundary condition $T(1,y)=y(1-y)$ represents the temperature distribution at the boundary of the plate, where $y=1$ corresponds to the right edge of the plate. The term $y(1-y)$ describes a parabolic temperature distribution along the right edge, with the maximum temperature at $y=0.5$ (the midpoint of

10 What can you infer about the decay rate of temperature?

The decay rate of temperature can be inferred to be exponential. This is because the temperature at any given time is a function of the initial temperature and the decay rate, which is a constant. The formula for exponential decay is $T(t) = T_0 * e^{(-kt)}$, where T_0 is the initial temperature, t is time, and k is the decay constant. The decay constant k is a measure of how quickly the

11 Comment on the physical interpretation of why the spatial pattern remains unchanged while only the amplitude decreases with time?

The physical interpretation of this phenomenon can be understood by considering the wave equation and its solutions. In the context of wave propagation, the wave equation describes how a disturbance (or wave) propagates through a medium. The solutions to the wave equation are typically sinusoidal functions, which represent standing waves or traveling waves.

In the case of a

12 What is the effect of alpha on the decay rate of heat dissipation?

The effect of alpha on the decay rate of heat dissipation is that it determines the rate at which the heat dissipation decreases over time. Alpha is a parameter that characterizes the rate of decay in exponential decay processes. In the context of heat dissipation, a higher alpha value indicates a faster decay rate, meaning that the heat dissipation decreases more rapidly over time. Con

3. Validation Approach

Dataset Splitting:

Used the generated dataset for training and validation.

Split the dataset into training and validation sets (80% training, 20% validation).

Evaluation Metric:

Used cosine similarity to compare the model's responses with the ground truth answers.

The score reflects how closely the model's responses match the expected answers.

Fine-Tuning Validation:

Monitored the training loss and validation loss during fine-tuning.

Ensured the model was not overfitting by using early stopping.

4. Analysis of Results

Model Performance:

The fine-tuned model achieved a cosine similarity score of 0.49938.

The score indicates moderate similarity between the model's responses and the ground truth.

Strengths:

The model provided physically meaningful interpretations of the heat equation solutions.

Responses were concise and relevant to the questions.

Weaknesses:

Some responses lacked depth or specific details (e.g., mathematical derivations).

The model struggled with questions requiring complex reasoning or calculations.

Improvements:

Use a larger and more diverse dataset for fine-tuning.

Experiment with different hyperparameters (e.g., learning rate, batch size).

Incorporate additional numerical methods (e.g., FEM, FVM) for data generation.

Conclusion

The fine-tuned IBM Granite 3.1-8b-instruct model provided meaningful interpretations of the heat equation solutions. While the model performed well on straightforward questions, there is room for improvement in handling complex queries. Future work will focus on enhancing the dataset and fine-tuning process to improve the model's accuracy and depth of responses.