

0. Introduction

[FinGPT](#) is an open-source financial large language model project focused on real-world financial NLP tasks. In this project, I tried to finetune three LLMs: Llama3.1-8B, Qwen2.5-7B-Instruct, DeepSeek-R1-Distill-Llama-8B using data from [FinGPT/fingpt-forecaster-dow30-202305-202405 · Datasets at Hugging Face](#). I will show some details of my finetuning and some meaningful observations.

To get the basic information on the FinGPT project, you can refer to [this blog](#).

1. Experimental Setup

I performed finetuning and inference on one Nvidia RTX 5090 GPU with CUDA 13.0.

I downloaded the three base models from [modelscope](#).

The training and test datasets are both from [FinGPT/fingpt-forecaster-dow30-202305-202405 · Datasets at Hugging Face](#).

2. Finetuning and Inference

I used [llamafactory](#) for my finetuning: It is a very efficient tool for finetuning base models.

Some hyperparameters of interest:

lr : 5*10**(-5); num_of_epochs: 4; lora_rank: 16; lora_alpha: 32; cutoff_length: 2560; batch_size: 2, gradient_accumulation_steps: 8

I used the 'prompt' values in the dataset [FinGPT/fingpt-forecaster-dow30-202305-202405 · Datasets at Hugging Face](#) as 'instruction' and 'answer' values as 'output' to form my training data in the llamafactory-accessible format.

```
[
  {
    "instruction": "[INST]<<SYS>>\nYou are a seasoned stock market analyst. Your task is to list the posi",
    "output": "[Positive Developments]:\n1. Market Outperformance: Despite the recent fall, American Express",
  },
  {
    "instruction": "[INST]<<SYS>>\nYou are a seasoned stock market analyst. Your task is to list the posi",
    "output": "[Positive Developments]:\n1. American Express's stock portfolio value increased significantl",
  },
  {
```

I used [vllm](#) for local inference.

3. Observations

Llama3.1–8B

Llama3.1–8B is an uninstructed model, so I expect it to have poor performances.

The outcome confirms my expectation: It seems hard to let this model follow the prompts, and thus it is difficult to parse out meaningful information.

Little improvement is witnessed in the finetuned version, due to the limited amounts of training data. Here I attach a typical output of this model:

```
[Headline]: American Express' (NYSE:AXP) Dividend Will Be Increased To $0.70
[Summary]: The board of American Express Company ( NYSE:AXP ) has announced that it will be paying its dividend of $0.70 on the...

[Headline]: American Express' (NYSE:AXP) Dividend Will Be Increased To $0.70
[Summary]: The board of American Express Company ( NYSE:AXP ) has announced that it will be paying its dividend of $0.70 on the...

[Headline]: American Express' (NYSE:AXP) Dividend Will Be Increased To $0.70
[Summary]: The board of American Express Company ( NYSE:AXP ) has announced that it will be paying its dividend of $0.70 on the...

[Headline]: American Express
None
```

We can see that the model is copying or repeating parts of the prompt (headlines, summaries) and it is not producing the structured [Positive Developments], [Potential Concerns], [Prediction & Analysis] format, which is a typical behavior of non-instructed models.

Qwen2.5–7B-Instruct

This instructed model produces much more meaningful results.

```

{
  "valid_count": 297,
  "bin_acc": 0.4276094276094276,
  "mse": 6.617845117845118,
  "pros_rouge_scores": {
    "rouge1": 0.39941974130564806,
    "rouge2": 0.12541717080316309,
    "rougeL": 0.25409868531325047
  },
  "cons_rouge_scores": {
    "rouge1": 0.34585829615427527,
    "rouge2": 0.08957099403172535,
    "rougeL": 0.2148732567389812
  },
  "anal_rouge_scores": {
    "rouge1": 0.3880227554173664,
    "rouge2": 0.08961055821567994,
    "rougeL": 0.1879528461350897
  },
  "avg_time": 3.590280412832896
}

```

We can see that 297/300 model outputs are parsed successfully. Here is a typical base model output:

```

[Positive Developments]:
1. Dividend Hike: American Express announced a significant 16.7% increase in its dividend, indicating financial strength and a commitment to rewarding shareholders.
2. Innovation and Market Position: As a leading player in the financial services sector, American Express continues to innovate and maintain a strong market position, which could drive future growth.

[Potential Concerns]:
1. Market Underperformance: American Express's stock underperformed relative to its competitors, suggesting that broader market or sector-specific issues may be affecting its performance.
2. Economic Uncertainty: The overall market conditions and economic uncertainty could pose risks to the company's revenue and profitability.

[Prediction & Analysis]:
Prediction: The stock price of American Express (AXP) is likely to remain stable or experience minor fluctuations over the next week.

Analysis:
- Positive Factors: The dividend hike and the company's strong market position provide a solid foundation for investor confidence. However, the recent underperformance suggests that market sentiment might still be cautious.
- Negative Factors: Economic uncertainties and market-wide underperformance could continue to weigh on the stock price. Additionally, the high price-to-earnings ratio (PE) of 16.3032 indicates that the stock is relatively expensive, which could deter some investors.
- Technical Indicators: Given the recent price movement, the stock might consolidate around current levels or experience minor adjustments. The lack of significant news or events in the near term could lead to a sideways trend.

Overall, while there are positive developments that support the stock, the current market environment and recent underperformance suggest a cautious outlook for the coming week.

```

Now compare that with the performance of our finetuned version:

```
{
  "valid_count": 297,
  "bin_acc": 0.4444444444444444,
  "mse": 7.341750841750842,
  "pros_rouge_scores": {
    "rouge1": 0.4551213048479248,
    "rouge2": 0.18145296750531867,
    "rougeL": 0.287095131766227
  },
  "cons_rouge_scores": {
    "rouge1": 0.42959388683400396,
    "rouge2": 0.1547375922206421,
    "rougeL": 0.27206717850675793
  },
  "anal_rouge_scores": {
    "rouge1": 0.44306560722765,
    "rouge2": 0.14005517999155312,
    "rougeL": 0.22622359768424058
  },
  "avg_time": 4.201723875999451
}
```

We can see that the binary prediction accuracy is moderately improved, but a larger MSE value is obtained. From another aspect, we witness obvious improvements on rouge values. Such behavior is also expected: our target is the entire contexts of analysis, so although the finetuned model matches the ground-truth wording better (all ROUGE scores up), it does not strengthen its quantitative prediction ability very much. A good alternative approach would be to add more weights on the quantitative prediction section. Here is a typical finetuned model output:

```
[Positive Developments]:
1. American Express has been recognized for its strong performance over the past five years, delivering a stellar return of 124%. This indicates in
vestor confidence in the company.
2. The company has announced an increase in its dividend payout to $0.70, which could attract income-focused investors and potentially boost the st
ock price.
3. The company's gross margin stands at 0.6334, indicating efficient cost management and profitability.

[Potential Concerns]:
1. The company's total debt to equity ratio is quite high at 6.355, suggesting that the company is heavily leveraged, which could pose risks in cas
e of a downturn.
2. The company's current ratio is less than 1, indicating potential liquidity issues.
3. The company's long-term debt to total asset ratio is also relatively high at 0.1833, which might raise concerns about the company's ability to m
eet its long-term obligations.

[Prediction & Analysis]:
Prediction: Up by 0-1%
Analysis: Despite some concerns regarding the company's high debt levels and potential liquidity issues, the overall outlook for American Express r
emains positive. The company's strong historical performance, coupled with the recent announcement of a dividend increase, suggests that the compan
y is in a good position to attract and retain investors. Furthermore, the company's gross margin indicates effective cost management and profitabil
ity, which bodes well for future earnings growth.

However, the company's high debt levels do warrant caution, especially given the current economic environment where interest rates are expected to
rise. If the company cannot effectively manage its debt levels, it could face challenges in meeting its financial obligations, which could negative
ly impact the stock price.

Given these factors, I predict that American Express's stock price may experience a slight upward movement of around 0-1% in the upcoming week. How
ever, investors should closely monitor the company's financial health and macroeconomic conditions, particularly interest rate trends, as they coul
d significantly impact the company's performance and stock price.
```

DeepSeek-R1-Distill-Llama-8B

The instructed deepseek base model performs a little worse than qwen2.5-7B base model:


```
{  
  "valid_count": 145,  
  "bin_acc": 0.4,  
  "mse": 13.355172413793104,  
  "pros_rouge_scores": {  
    "rouge1": 0.400652000648513,  
    "rouge2": 0.10934604571946498,  
    "rougeL": 0.23478521523208915  
  },  
  "cons_rouge_scores": {  
    "rouge1": 0.3126972005035444,  
    "rouge2": 0.06644256545352357,  
    "rougeL": 0.17054937678117588  
  },  
  "anal_rouge_scores": {  
    "rouge1": 0.25815301312022326,  
    "rouge2": 0.04454926172966924,  
    "rougeL": 0.14225181940635678  
  },  
  "avg_time": 8.579050907293956  
}
```

An obvious improvement is witnessed in the finetuned version. This time both the prediction accuracy and rouge values are improved, which makes great sense considering the moderate base model capability:

```
{
  "valid_count": 297,
  "bin_acc": 0.531986531986532,
  "mse": 8.412457912457912,
  "pros_rouge_scores": {
    "rouge1": 0.44607514455794633,
    "rouge2": 0.17634100378381995,
    "rougeL": 0.2836017688847098
  },
  "cons_rouge_scores": {
    "rouge1": 0.4230040157142374,
    "rouge2": 0.15398771959070515,
    "rougeL": 0.26160950612427303
  },
  "anal_rouge_scores": {
    "rouge1": 0.4421233689837768,
    "rouge2": 0.14366381105631867,
    "rougeL": 0.22771010277125608
  },
  "avg_time": 9.664793166319528
}
```

4. Conclusion

In this project I carried out finetuning on three base models: Llama3.1-8B, Qwen2.5-7B-Instruct, DeepSeek-R1-Distill-Llama-8B using [FinGPT/fingpt-forecaster-dow30-202305-202405 · Datasets at Hugging Face](#). The results shown above have demonstrated the sound effects of my finetuning. Further improvements can be made through adding more weights to the quantitative prediction.