

Explainable model for time-series via ChatGPT

Xue Zhechang

CHAIR FOR PERVASIVE COMPUTING SYSTEMS, INSTITUTE OF TELEMATICS, DEPARTMENT OF COMPUTER SCIENCE



Outlines

- Paper review
- Baseline
- 2-Class classification
- Generate data trend by ChatGPT
- Time Table

Paper Review

Temporal Data Meets LLM

- 实验模型: few-shot/zero-shot GPT-4
 - Instruction-based fine-tuning LLaMa
- 实验对象: NASDAQ-100
- 实验结果: 效果比ARMA-GARCH, gradient-boosting tree model好, GPT-4 > LLaMa
- 评价标准: Binary precision: 看是Up还是Down
Bin precision: 看是不是在这个bin里
MSE: 测量与实际值误差
- 实验方法: 周预测: 把数据集根据周切成12份, (D5+, U5+)
月预测: (D10+, U10+)
让GPT-4自己总结公司profile, 输入相关的新闻 (google), 让他自己写meta-summary

	Weekly			Monthly (Every 4 Weeks)		
	Binary Precision	Bin Precision	MSE	Binary Precision	Bin Precision	MSE
Most-Frequent Historical Bin	50.7%	16.4%	43.5	51.4%	17.2%	155.1
ARMA-GARCH	52.4%	11.1%	22.1	50.5%	6.2%	90.1
Gradient Boosting Tree Model	60.8%	26.4%	24.3	56.4%	17.7%	85.6
GPT-4 Zero-Shot	64.5%	31.2%	20.5	64.8%	26.0%	60.1
GPT-4 Few-Shot	65.8%	32.7%	20.6	65.3%	26.5%	58.2
GPT-4 Few-Shot w/ COT	66.5%	35.2%	18.7	69.5%	28.6%	50.4
Open LLaMA (13B) Fine-Tuned	62.2%	26.5%	23.3	60.1%	22.6%	63.3
Open LLaMA (13B) Fine-Tuned w/ COT	64.7%	30.7%	21.0	62.2%	24.4%	63.5

Table 1: Performance comparison between the baseline models and LLMs for stock price weekly/monthly forecasting.

Baseline

- 300 train, 700 test, 10 classes
- Imputed with 0
- XGBoost*: 46.57% (XGBoost doesn't really need to impute)
- RandomForest: 54.29%
- DecisionTree: 32.86%

Baseline

- 120 train, 280 test, 4 classes
- Imputed with 0
- XGBoost*: 68.93% (XGBoost doesn't really need to impute)
- RandomForest: 79.29%
- DecisionTree: 57.50%

	Decision Tree	Gradient Boosting	Random Forest**
Average	51.9%	54%	40.2%
Highest	90% (5)	100% (1)	60%(2)
Lowest	20% (2)	25% (1)	20%(1)

Baseline

- 60 train, 140 test, 2 classes (1&2)
- Imputed with 0
- XGBoost*: 75.71% (XGBoost doesn't really need to impute)
- RandomForest: 92.86%
- DecisionTree: 77.86%

2-class classification

Based on **10** features extracted from Gradient boosting tree

2 classes & each **20** train samples & **10** test samples

	3 & 6	3 & 4	6 & 8	2 & 9	4 & 5	2 & 4
Average 72.93%	63.8 %	67.7%	81.0%	96.7%	71.9%	56.5%
Highest	100% (10)	100% (7)	100% (4)	100% (39)	100% (10)	100% (1)
Lowest	35% (2)	35% (1)	50% (4)	15% (1)	40% (3)	35% (2)

Generate data trend by ChatGPT

Based on **10** features extracted from Gradient boosting tree

输入源数据，让GPT先想象数据的曲线图，然后分析曲线图得到对趋势的描述和特征。



分析所提供的数据集，我们可以观察到以下趋势和特征：



1. **总体模式：**

- 数据呈现周期性的模式，在正负值之间波动。
- 存在相对稳定值的时期，以及急剧变化的时期。

2. **振荡频率：**

- 振荡频率似乎在数据集的后半部分增加。
- 在初始部分，振荡相对较慢，在中间部分，变得更加频繁。

3. **振幅：**

- 振荡的幅度不同。一些振荡的幅度比其他的大。
- 随着数据集末尾的临近，振荡的幅度总体上在增加。

4. **方向性变化：**

- 数据的整体趋势存在明显的变化。最初，数据呈现正趋势，然后变为负趋势，后来再次变为正趋势。

5. **稳定期：**


- 存在数据相对稳定的时期，特别是在数据的中间部分。

6. **不对称性：**

- 振荡的正负阶段并不对称。

7. **异常值：**

- 存在一些数据出现急剧变化的情况（异常值），超出了当前趋势。

 Regenerate

Generate data trend by ChatGPT

Based on **10** features extracted from Gradient boosting tree

告诉GPT背景，让他猜测是什么动作



上面的数据集是持有运动检测器（佩戴于手上）的人，做某个动作所检测到的坐标轴X值，请猜测这个动作



通过分析提供的数据集，很难直接确定具体的动作，因为数据只提供了在某个坐标轴上的数值变化而没有具体的标签或上下文。不过，我们可以从数据的波动和变化中提出一些建议。

1. 周期性波动：

- 数据呈现了一些明显的周期性波动，这可能与某种循环性动作有关。

2. 方向性变化：

- 在数据集的不同部分，坐标轴的方向发生了变化，可能表示在执行不同动作时手的方向发生了改变。

3. 振幅的变化：

- 振幅的变化可能反映了动作的幅度或力度的改变。

4. 稳定期：

- 存在一些时期，数据相对稳定，这可能对应于某些静止或相对不活动的状态。

5. 不对称性：

- 振荡的正负阶段并不对称，这可能与动作的不对称性有关。

基于以上观察，该动作可能是一种循环性动作，涉及到手的方向变化、幅度变化，可能包含某种不对称性。然而，确切的动作还需要更多的上下文信息或标签，例如特定动作的名称或描述，以便更准确地推断。如果您有更多的信息可提供，我将尽力提供更具体的猜测。

Generate data trend by ChatGPT

Based on **10** features extracted from Gradient boosting tree

给GPT限定范围，让他猜测



请猜测这个数据集是哪个动作：A，捡东西；B，左右摇手；C，向左移动；D，向右移动



根据对数据集的观察和分析，我会倾向于猜测这个数据集可能与动作 B，即左右摇手有关。

理由如下：

1. **周期性波动：** 数据呈现了一些明显的周期性波动，这与左右摇手的动作可能相关。
2. **方向性变化：** 在数据集的不同部分，坐标轴的方向发生了变化，这可能反映了手在左右移动的变化。
3. **振幅的变化：** 振幅的变化可能反映了手的左右移动幅度的改变。
4. **不对称性：** 振荡的正负阶段并不对称，这可能与手在左右摇摆时的不对称性有关。

请注意，这只是基于观察到的模式的猜测，准确性可能受到数据集本身的特定性以及采集设备和方法的影响。如有可能，最好通过提供标签或更多上下文信息来验证或精确描述这些数据的含义。

Time Table

10/6-10/17: 找出最好的prompt

10/18-10/24: 基于最好的prompt生成解释

10/25-10/31: 人工评测

11/1 - 11/15: 写论文

周日 10月 1日(十七)	周一 2 (十八)	周二 3 (十九)	周三 4 (二十)	周四 5 (廿一)	周五 6 (廿二)	周六 7 (廿三)
					找出最好的prompt	
8 (寒露)	9 (廿五)	10 (廿六)	11 (廿七)	12 (廿八)	13 (廿九)	14 (三十)
找出最好的prompt						
15 (九月)	16 (初二)	17 (初三)	18 (初四)	19 (初五)	20 (初六)	21 (初七)
找出最好的prompt			基于最好的prompt生成解释			
22 (初八)	23 (初九)	24 (霜降)	25 (十一)	26 (十二)	27 (十三)	28 (十四)
基于最好的prompt生成解释			人工评测			
29 (十五)	30 (十六)	31 (十七)	11月 1日(十八)	2 (十九)	3 (二十)	4 (廿一)
人工评测			写论文			