

Generate explanations for time-series classifier by ChatGPT

Xue Zhechang

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

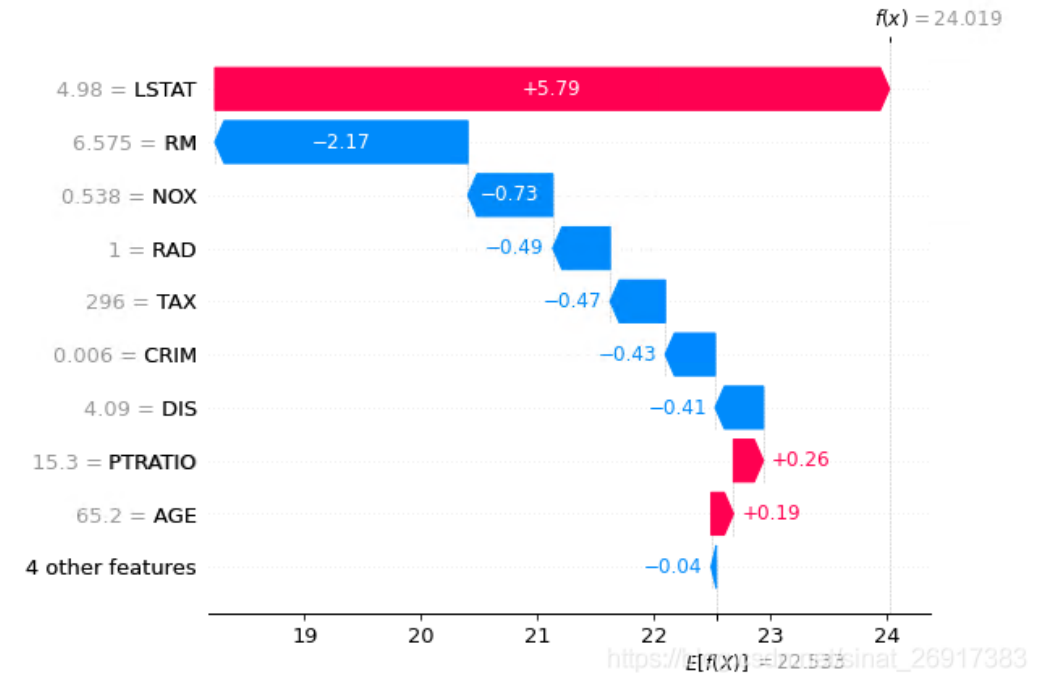


Outlines

- Motivation
- Ground knowledge
- Related Works
- Example of my solution
- Realization of my solution
- Planned Evaluation

Motivation

- Problem: Current method to explain AI is **not user-friendly** for **non-researchers**.
- Idea: Use **ChatGPT** to explain why the feature can contribute to the classification
- Benefit:
 - User-friendly and easy-accessible.
 - Link features to knowledge in other fields.
- Actions: Design a **prompt** to guide GPT give the answer.



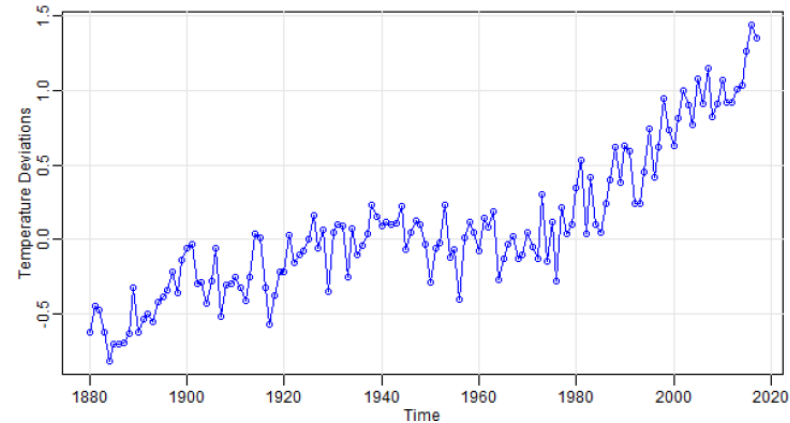
Why?

Ground knowledge

ChatGPT



Time series



Explanability



Related Works

■ Feature selection(tsfresh)

- Gašper et al. provided a new method to understand the importance of time series feature generated by tsfresh

■ Forecasting via Prompt

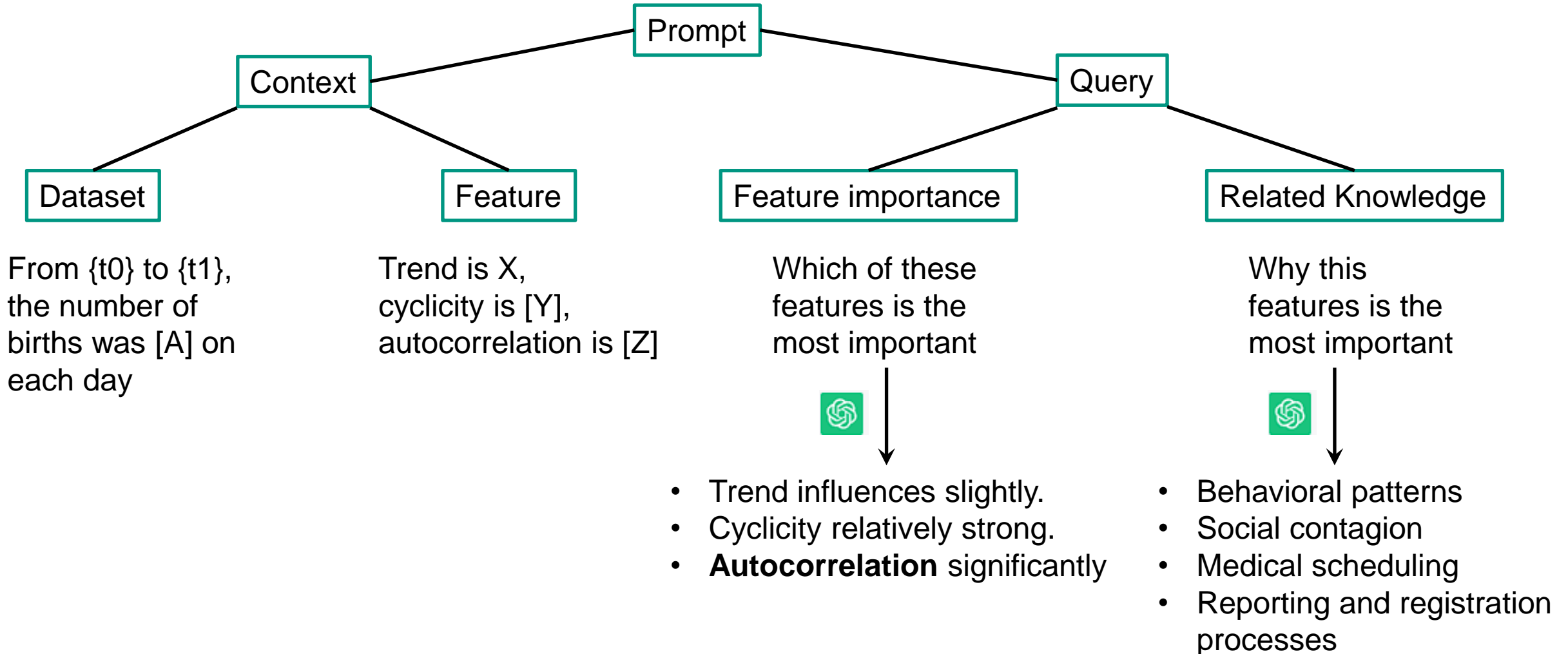
- Xue et al. used prompt-based method to forecast time-series data
- Set a benchmark to evaluate different methods and LLM

■ Causal Inference

- Stephanie et al. Proved that GPT can generate causal graphs in a high accuracy.

Example of my solution

Daily total female births in California



Realization of my solution

■ Feature selection(tsfresh)

■ Prompt design

■ Context

■ Dataset description

■ Features

■ Query

■ Study which key words improve answer stability

```
Try to classify
0.0000 0.0000 0.0000 1.0000 1.0000 1.0000
0.3214 0.0000 0.0000 1.0000 1.0000 0.0000
0.9286 1.0000 1.0000 1.0000 1.0000 1.0000
0.0779 0.0000 1.0000 0.0000 0.0000 0.0000
1.0000 1.0000 1.0000 0.0000 1.0000 0.0000
0.1429 0.0000 1.0000 0.0000 0.0000 0.0000
0.7786 0.0000 1.0000 0.0000 0.0000 0.0000
0.9286 1.0000 1.0000 1.0000 1.0000 0.0000
0.3571 1.0000 1.0000 1.0000 1.0000 0.0000
0.0714 0.0000 0.0000 1.0000 1.0000 1.0000
0.9407 0.0000 1.0000 1.0000 0.0000 1.0000
0.0250 0.0000 0.0000 1.0000 1.0000 1.0000
0.9587 0.0000 0.0000 1.0000 1.0000 1.0000
0.2321 0.0000 0.0000 1.0000 1.0000 0.0000
0.9407 0.0000 0.0000 1.0000 1.0000 1.0000
0.7184 0.0000 0.0000 1.0000 1.0000 1.0000
0.2500 0.0000 0.0000 1.0000 1.0000 1.0000
0.7321 1.0000 1.0000 0.0000 1.0000 0.0000
0.9444 0.0000 0.0000 0.0000 1.0000 0.0000
0.3036 1.0000 1.0000 1.0000 1.0000 1.0000
0.3214 0.0000 0.0000 1.0000 1.0000 0.0000
0.5385 1.0000 1.0000 1.0000 1.0000 1.0000
0.7184 0.0000 0.0000 1.0000 1.0000 0.0000
0.9444 1.0000 1.0000 1.0000 1.0000 1.0000
0.8008 0.0000 0.0000 0.0000 0.0000 0.0000
to Label 0.0 or 1.0, with the help of dataset given above. Don't show me the code. Give me the
label in format [label 1, label 2, ..., label 25]
```

```
[label 1, label 2, ..., label 25]
[1.0, 0.0, 1.0, 0.0, 0.0, 0.0, 0.0, 1.0, 1.0, 1.0, 0.0, 1.0, 0.0, 1.0, 0.0, 0.0, 1.0, 0.0, 1.0,
0.0]
```



■ Combine feature importance to domain knowledge

■ Get key words from answer and ask automatically.

■ Prove the reliability



Planned Evaluation

- Research object: A time-series dataset (From UCR)
 - With explanation generated by other methods
- Target:
 - Stable desired answer
 - Better prompt, avoiding incorrect answer or unwished format
 - Prove to be reliable
 - Compare GPT-selected importance and pre-set importance
 - Modify original data to strongly affect feature
 - Test consistency

Prompt engineering

1. Write clear and specific instructions

1. 使用分隔符
2. 获得指定输出 如JSON
3. 让LLM检查文本是否满足条件
4. 为模型提供几个问答的例子, 让 LLM 熟悉这种风格

2. Give the model time to think

1. 明确完成任务所需的步骤: Step1, Step2
2. 获取结构化的输出:

3. LLM的具体功能:

1. 总结, 推理 (态度, 情感, 主题), 转换, **扩展**

4. 迭代优化:

1. 更明确地聚焦于某个方面:
Bsp: The description is intended for furniture retailers, so should be technical in nature and focus on the materials the product is constructed from.

- Use the following format:
- Text: <text to summarize>
- Summary: <summary>
- Translation: <summary translation>
- Names: <list of names in French summary>
- Output JSON: <json with summary and num_names>

Dataset

Acuteinflammation & acutenephritis

■ Method: 15 times, 25 tests, decision tree vs ChatGPT

■ Decision tree: 100% correct on both datasets

■ ChatGPT: big range, not stable

■ Old_Prompt: There are A acuteinflammation, and each has B columns. Each column C is ...

Try to classify following... to Label 0.0 or 1.0, with the help of dataset given above.

Don't show me the code. Give me the label in format: [label 1, label 2, ..., label D].

New_Prompt: You are a data analyst, your job is to classify the given time series dataset.

There are 93 acutenephritis, and each has 6 columns.

The dataset will be given in format: [column 1, column 2, ..., column 6, label].

■ False rate:

Acuteinflammation: 20.0%, 16.0%, 20.0%, 20.0%, 20.0%, 36.0%, 20.0%, 28.0%, 16.0%, 20.0%, 16.0%, 16.0%, 20.0%, 28.0%, 20.0%

Acutenephritis: 40.0%, 40.0%, 12.0%, 40.0%, 8.0%, 36.0%, 36.0%, 28.0%, 44.0%, 40.0%, 32.0%, 36.0%, 8.0%, 24.0%, 0.0%

■ Average: 21.06%(35.2%), 28.26%(40.8%)

Dataset

Acuteinflammation & acutenephritis

1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	1.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	1.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0
1.0	1.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	1.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0
1.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	1.0	1.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0	1.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0
0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0
1.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0
0.0	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0

Repr2Seq

- 使用14.01.2019-12.01.2022的股票TS
- 使用了Encoder-decoder模型，来描述TS结构和关键信息
 - Encoder分为Sequence Processor+TS2Vec
 - SP: $q_i = \{x_{i_1} - x_{i_1}, \dots, x_{i_j} - x_{i_1}, \dots, x_{i_T} - x_{i_1}\}$
 - TS2Vec: 计算Loss
 - Decoder:
 - 用argmax operation收集了生成词的目录
- Repr2Seq相比Transformer和Seq2seq有着更高的得分 (BLEU)

Generative Agent

■ Memory and retrieval

- takes the agent's current situation as input and returns a subset of the memory stream
- $score = \alpha_{recency} \cdot recency + \alpha_{importance} \cdot importance + \alpha_{relevance} \cdot relevance.$

■ Reflection

- when the sum of the importance scores for the latest events exceeds a certain threshold
- Agents generate trees of reflections

■ Planning and Reacting

- Reacting and Updating Plans + Dialogue

Causal inference

- Can Large Language Models Build Causal Graphs
 - Using medical context
- Determining whether GPT-3 can signal the presence or absence of an edge between two variables in a directed acyclic graph from the medical context.
- Evaluating whether the use of certain language in **prompts** or linking verbs improves the classification accuracy of GPT-3.
 - Normally, a better prompt after iteration performs better
 - Both verb and noun
 - **Specificity lowers accuracy**