Report for
Application of Large Language Models
COS836
Project: **Tracing patterns and identifying divergences
in temporal data files with the ChatGPT Code
Interpreter.**

Jean Sette Siqueira Vianna

Up to Week 1

Abstract

Code Interpreter has proven to be a powerful tool for many IT professionals. due to its ability to interpret data coming from files of different extensions, plot graphics and even interpret them without the need for long code snippets. However, one of the functions of some of these professionals is also to compare these data with each other and verify the existence or not of patterns, and that is why this article has as main objective to test the capacity of this tool to do this with the input data and, in addition, check when this pattern is not obeyed.

Contents

Chapter 1

Week 1: Work Proposal

Nowadays much of the importance of the IT professional comes from its analytical capacity and its ability to extract information from large masses of data, but a difficulty that is common in this scenario is the impossibility of automating some processes in which it needs the human eye to determine certain guidelines. However, as a suggestion of these powerful AI's that, despite having been trained with immense data provided by great developers, were made accessible to most people for their own use, we can allocate some of these functions to them. Therefore, the objective of this work is to make use of the Chat-GPT Code Interpreter to automate a process of interpreting patterns in data files, and, in addition, to evaluate its dissension in classifying whether this pattern was obeyed or not.

An example of the use of this functionality would be to evaluate periodic extracts of sales of products from a given sector, and when, for example, there is a human error in entering the value of a SKU with the addition of a 0 (multiplying the total value of sales by 10), it would not take time to identify this error. Avoiding reaching the most external layers of analysis, such as the company's Director, for example. We know that in the job market it is impossible not to make mistakes of this type, but what makes the difference is the time it takes to execute and resolve them.

Another example would be an academic project that provides a service to a certain regulatory body in which data from different companies in the country is handled. Knowing that these companies send data annually, and this data is extracted from databases, it is up to the project members to verify that the data was sent correctly in order to respond in a timely manner so that the company can redo the recipe and subsequently send it.

The initial idea is to use the Chat-GPT Code Interpreter feature to carry out this work, we will also use some artificial data and create some curve interpolation algorithm to compare the results. It was thought to use this resource for the ease of file input at the same time that it analyzes and interprets the data. Your cache memory will be tested (if you can save information from old files) and your analytical capacity within and between files.

1.1   Work Report

- Done!
  - Papers read
    - Vaswani et al. (2017) is about...
  - Papers that were just skimmed
    - Vaswani et al. (2017) seems interesting...
  - Sites visited
    - https://powerbrainai.com/
  - Doubts and questions for the next class
    - List your questions and doubts
- New todos: New things you decided to do
  - Read another paper, like (ProducPlan, n.d.)
  - Create a program that reads a file and generate vectors from it

1.2   Keep KANBAN updated!

There is a file for your group to keep a Kanban updated in the end of this project.

1. Keep the kanban with things you think you must do, in priority order. You must use words as **Must**, **Should**, **Could**, **Won´t**, as in the MoSCoW method ProducPlan (n.d.).

Chapter 2

Week 2

2.1 To dos

You can make a todo list:
- ☐ Can you make an example of your application by hand, using Chat-GPT?
- ☐ Can you make mock screens?
- ☐ Fill the Software Project Canvas
- ☐ Not done yet
- ☑ Done
- ☒ Canceled Item

2.2 Software Project Canvas

At this point, you should have a good version of your Software Project Canvas (SPC), see Figure 2.1. Although it is in Portuguese, fill it in English (there will soon be an English version "soon").

Figure 2.1: Software Project Canvas, alternative 1 version, starts with WHY?.

2.3 Work Report

- Done!
  - Papers read
    - ⋄ Vaswani et al. (2017) is about...
  - Papers that were just skimmed
    - ⋄ Vaswani et al. (2017) seems interesting...
  - Sites visited
    - ⋄ https://powerbrainai.com/
  - Doubts and questions for the next class
    - ⋄ List your questions and doubts
- New todos: New things you decided to do
  - Read another paper, like (ProducPlan, n.d.)
  - Create a program that reads a file and generate vectors from it

2.4 Keep KANBAN updated!

There is a file for your group to keep a Kanban updated in the end of this project.

1. Keep the kanban with things you think you must do, in priority order. You must use words as **Must**, **Should**, **Could**, **Won´t**, as in the MoSCoW method ProducPlan (n.d.).

Chapter 3

Week 3...

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

3.1   Keep KANBAN updated!

There is a file for your group to keep a Kanban updated in the end of this project.

1. Keep the kanban with things you think you must do, in priority order. You must use words as **Must**, **Should**, **Could**, **Won´t**, as in the MoSCoW method ProducPlan (n.d.).

Chapter 4

More Instructions

You should use this model in Overleaf. However, you should use the option "from another project" for the document "padrao.sty", so the teacher can keep it updated.

The bibliography should be put in BiBTeX format in the file "references.bib" You can delete the following line after having your own bibliography

To use tables, if necessary, do not use cell borders and use `\toprule`, `\midrule` and `\bottomrule`. Table captions are above the table, as in Table 4.1, and figure cations are below the figure, as in Figure 4.1. Figures and tables must always be referenced in text.

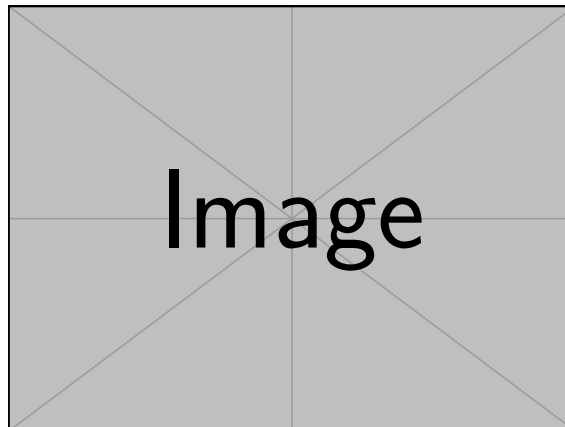Table 4.1: Sample table and caption

Title Cell	Title Cell 2
1	2
3	4

Figure 4.1: Placeholder Figure

Chapter 5

Bibliography

ProducPlan (n.d.). *What is MoSCoW Prioritization? | Overview of the MoSCoW Method.* https://www.productplan.com/glossary/moscow-prioritization/. (Accessed on 06/27/2023).

Vaswani, Ashish et al. (2017). *Attention Is All You Need.* Version v5. arXiv: 1706.03762 [cs.CL]. URL: https://arxiv.org/abs/1706.03762 (visited on 06/27/2023).

Chapter 6

Tools being used

You must put here a list with tools and tools' urls.

- Overleaf: https://www.overleaf.com/
- GitHub:

Chapter 7

Links to the Project

Here a list of links that should be used to evaluate the project. You must complete the urls.

- GitHub: https://github.com/
- Overleaf (edit access): https://www.overleaf.com/

Chapter 8

KanBan

8.1 Todo

> **Task without a name**
>
> **Status**: todo

> **WHY+WHAT (MUST): Describe scenario for application**
>
> **Status**: todo
> **Due date**: 4/7/2023
> **Assigned in**: 27/06/2023
> **From**: Geraldo Xexéo
> **To**: Group
> **Rationale**: You must say why your ideia is useful
> **Done in**: When it was done
>
> -
>
> Describe the business view and, as a user, how it will work. You must describe the current problems and opportunities and the future benefits.

> **HOW (MUST): Describe an overall idea for a technical solution**
>
> **Status**: todo
> **Due date**: 4/7/2023
> **Assigned in**: 27/06/2023
> **From**: Geraldo Xexéo
> **To**: Group
> **Rationale**: You must explain how it is feasible
> **Done in**: When it was done
>
> -
>
> Describe the business view and, as a user, how it will work.

HOW (MUST): Where is the data

Status: todo
Due date: 4/7/2023
Assigned in: 27/06/2023
From: Geraldo Xexéo
To: Group
Rationale: You must explain how it is feasible
Done in: When it was done

Describe how you will find the data.

HOW MUCH (MUST): Plan your expenses

Status: todo
Due date: 4/7/2023
Assigned in: 27/06/2023
From: Geraldo Xexéo
To: Group
Rationale: Teachers can try to find resources for the job
Done in: When it was done

Explain how much (even if zero) will you expend, say how

Software Project Canvas (SHOULD)

Status: todo
Due date: 11/7/2023
Assigned in: 27/06/2023
From: Geraldo Xexéo
To: Group
Rationale: Must know what you are doing
Done in: When it was done

Fill the Software Project Canvas

8.2 To Check

8.3 Checking

8.4 Done

8.5 Here only examples

Make coffe

Status: todo
Due date: Soon
Assigned in: July 18, 2023
From: Geraldo Xexéo

Please do the coffe strong

Drink a coffe

Status: doing
Due date: In 5 days
Assigned in: June, 1st 2021
From: The Boss
To: The Worker

Roast the beef

Status: tocheck
Assigned in: June, 1st 2021
From: The Boss
To: The Worker
Done in: July 18, 2023

Rare, plese

Drink another beer

Status: checking
Assigned in: June, 1st 2021
From: The Boss
To: The Worker
Done in: July 18, 2023

Please check if you have age to drink in your country. If not, drink orangejuice.

Drink a beer

Status: done
Assigned in: July 18, 2023
From: The Boss
To: The Worker
Rationale: Because I want
Done in: July 18, 2023

Task without a name

Task without a name

Status: done

arg 2

Status: arg 1
Due date: arg 3
Assigned in: arg 4
From: arg 5
To: arg 6
Rationale: arg 7
Done in: arg 8

arg 9