

Machine Learning

CS 412

Course Project

Onur Varol

Asst. Prof. of Computer Science

www.onurvarol.com | www.varollab.com

@onurvarol

Can we predict HW grades from ChatGPT interactions?

- Let's make a unique and challenging project
- We don't know if this is possible and we look at your approach to address this problem
- You all learn by grading other project teams and see their approaches.



ChatGPT

In Pandas, you can display variable (column) names in a DataFrame by accessing the column names or labels. You can do this in several ways:

1. Display Column Names:

To display the column names, you can access the `columns` attribute of the DataFrame. Here's how you can do it:

code. !whitespace-pre.hljs. language-python

462.12 x 188.8

Copy code

```
import pandas as pd

# Create a sample DataFrame
```

Upgrade to Plus to create your own GPT

[Report content](#) | [Terms of use](#) | [Privacy policy](#)

nts Console Sources Network Performance Memory Application Security Lighthouse Recorder 𐀀 Performance insights 𐀀

```
<div class="px-4 py-2 justify-center text-base md:gap-6 m-auto">
  <div class="flex flex-1 text-base mx-auto gap-3 md:px-5 lg:px-1 xl:px-5 md:max-w-3xl lg:max-w-[40rem] xl:max-w-[48rem] } group"> flex
    <div class="flex-shrink-0 flex flex-col relative items-end"> ... </div> flex
    <div class="relative flex w-full flex-col lg:w-[calc(100%-115px)] agent-turn"> flex
      <div class="font-semibold select-none">ChatGPT</div>
      <div class="flex-col gap-1 md:gap-3">
        <div class="flex flex-grow flex-col max-w-full"> flex
          <div data-message-author-role="assistant" data-message-id="6da913eb-4fad-43b9-9600-d033b710a1d3" class="min-h-[20px] text-message flex flex-col items-start gap-3 whitespac
e-pre-wrap break-words [.text-message+&]:mt-5 overflow-x-auto"> flex
            <div class="markdown prose w-full break-words dark:prose-invert dark">
              <p>... </p>
              <ol> flex
                <li>
                  ::before
                  <p>... </p>
                  <pre>
                    <div class="bg-black rounded-md">
                      <div class="flex items-center relative text-gray-200 bg-gray-800 dark:bg-token-surface-primary px-4 py-2 text-xs font-sans justify-between rounded-t-md"> ...
                        </div> flex
                      <div class="p-4 overflow-y-auto"> == $0
                        <code class="!whitespace-pre hljs language-python">... </code>
                      </div>
                    </div>
                  </pre>
                </li>
              </ol>
            </div>
          </div>
        </div>
      </div>
    </div>
  </div>
</pre>
<p>This will output:</p>
```

Format of the project

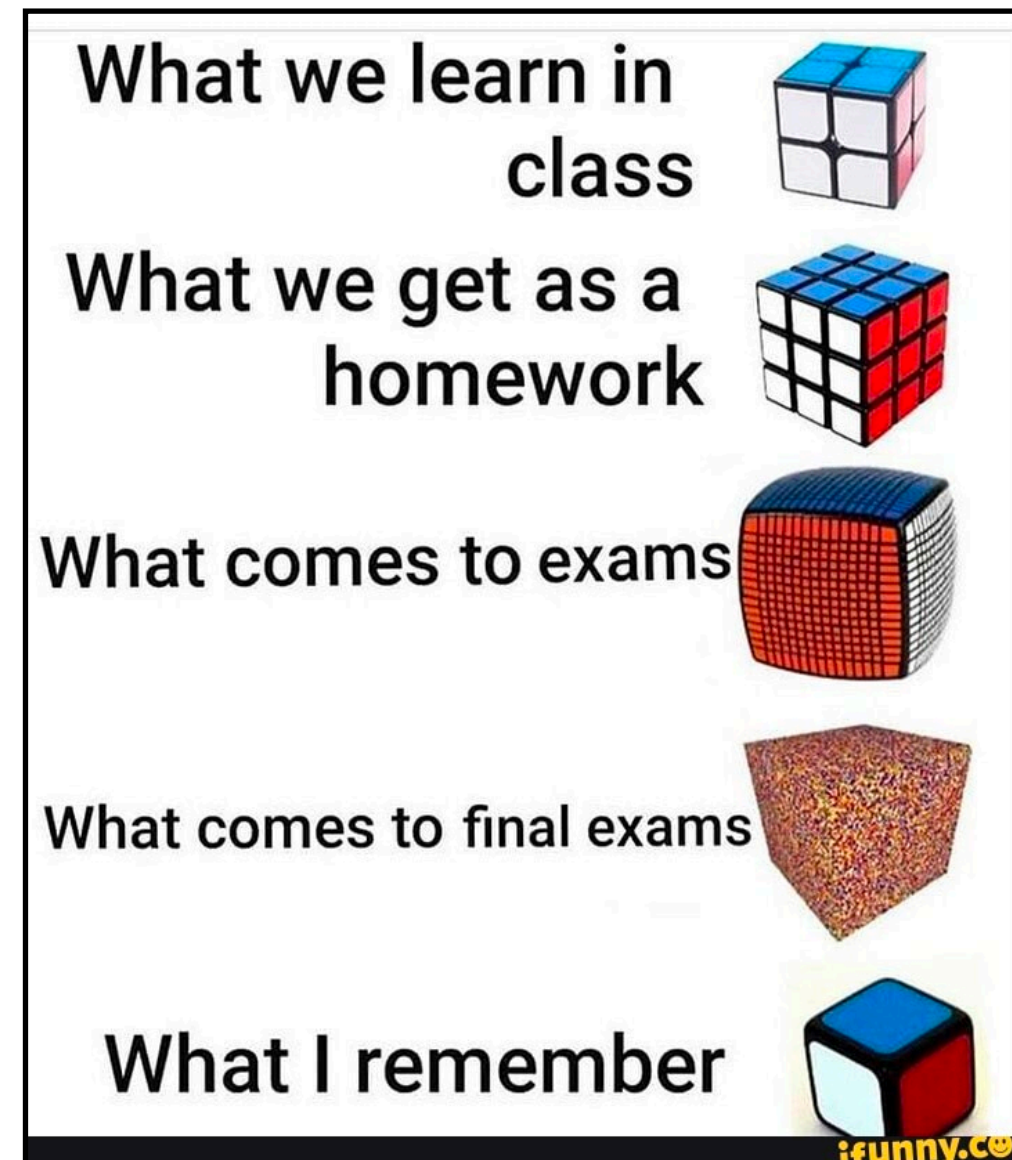
- You will receive the following materials
 - Collection of HTML files that contain ChatGPT histories (150-200 files)
 - CSV file that has a corresponding score of the student
- Output expected
 - Github repositories for all the material (code, scripts, etc.)
 - README file that explains your approach clearly and your findings and contribution of team members.

Structure of the README file

- Overview of the repository: Link different script and code pieces used and explain their function for the project
- Methodology: High-level explanation of things considered and solutions offered.
- Results: Experimental findings supported by figures, tables etc.
- Team contributions: List all team members by their names and how they contributed to the project

Challenges

- Extracting information from ChatGPT (feature engineering)
- Not all questions in the assignment solved using ChatGPT (missing data + imputation)
- Mapping prompt&answer to questions
- Dataset sizes
- Predicting scores (model building and evaluation)
- Clearly communicating your findings and methodologies used for the task



EVERY GROUP PROJECT



SAYS HE'S GOING
TO HELP, DOES
ONE THING, THEN
STOPS

THE BOSSY ONE
WHO **THINKS**
HE'S THE LEADER

THE QUIET ONE
WHO'S THE **REAL**
LEADER...BUT
GETS NO CREDIT
FOR IT

DOES NOTHING
UNTIL THE END

DOES 99% OF
THE WORK

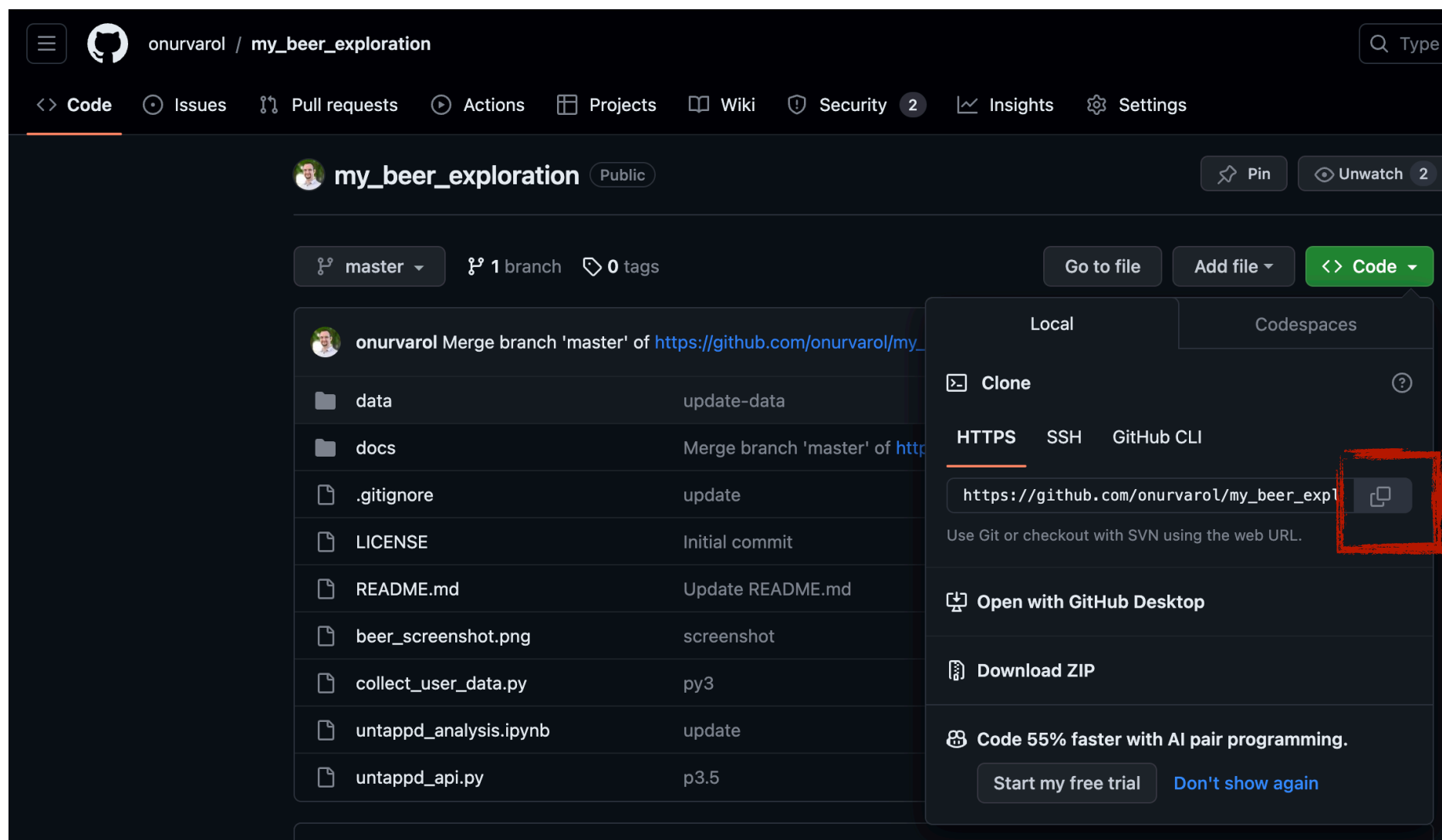
DISAPPEARS AT THE
BEGINNING AND
ONLY SHOWS UP
RIGHT AT THE END

Evaluation

- Code and material review (30%)
- Quality of the project ideas and results (40%)
- Peer-evaluation
 - Grading other projects (10%)
 - Grades received from peers (20%)

What is the submission format

- We will **only** ask for a **public** Github repository URL. Each team will submit “only one” repo.



- We will automate cloning Github repositories with a script. Mistakes in the URL format will cause you points!

Deadlines

- Since there are several hard deadline for grade submission, there will not be any late submission.
PLEASE PLAN ACCORDINGLY!
- **Deadline for submitting the project material: 19 Jan.**
- **Deadline for peer-evaluations: 22 Jan.**
- Deadline for Onur to submit grades: 25 Jan. at 3PM