Artifact Documentation:
ACSAC Submission #275, AE Submission #20

Can Large Language Models Provide Security & Privacy Advice? Measuring the Ability of LLMs to Refute Misconceptions

This document contains information about our artifacts which we make available via an Open Science Framework (OSF) page

Link - https://osf.io/xq37z/?view_only=00c450b4baef41979eae73f9c971e095

The artifacts reproduce the paper's main findings - experiments (E1, E2, E3, E4) which are detailed in Section 4 Evaluation.

This artifact document consists of two parts.

A) **File Directory** **pg 2**
   - Descriptions of files and folders made available

B) **Guideline to Reproduce Results** **pg 3-5**
   - Procedure to replicate results of experiments

We are available via the HOTCRP submission page to answer any questions if needed.

Thank you,
Authors of Submission #275, AE submission #20

A: File Directory

File	Description
reproduce_e1_e2_e3.py	Script to reproduce tables and figures for E1, E2, E3
reproduce_e4.py	Script to reproduce figures for E4
Dockerfile	To build docker image with required dependencies, and run artifacts
acsac275container_latest.tar	Pre-saved Docker container (built from Dockerfile)
requirements.txt	Python dependencies for scripts
dataset	
S&P-dataset.csv	Dataset of misconceptions
experiments	
Bard-responses-e1.csv	E1 experiment results, Bard (used in *reproduce_e1_e2_e3.py)*
Bard-responses-e2.csv	E2 experiment results, Bard (used in *reproduce_e1_e2_e3.py)*
Bard-responses-e3.csv	E3 experiment results, Bard (used in *reproduce_e1_e2_e3.py)*
Bard-responses-e4.csv	E4 experiment results, Bard (used in *reproduce_e4.py)*
ChatGPT-responses-e1.csv	E1 experiment results, ChatGPT (used in *reproduce_e1_e2_e3.py)*
ChatGPT-responses-e2.csv	E2 experiment results, ChatGPT (used in *reproduce_e1_e2_e3.py)*
ChatGPT-responses-e3.csv	E3 experiment results, ChatGPT (used in *reproduce_e1_e2_e3.py)*
ChatGPT-responses-e4.csv	E4 experiment results, ChatGPT (used in *reproduce_e4.py)*
plots	Empty directory where figures produced by scripts will be placed
supplemental-csv-files	
S&P-paraphrases.csv	Paraphrased S&P-dataset claims
supplemental-documentation	
Guideline for Manuscript Relevance.pdf	Guideline to filter relevant manuscripts
Guideline for Paraphrases.pdf	Guideline to decide if paraphrases are valid or not
Keyword for Manuscript Search.pdf	Keywords used in search for relevant manuscript
Prompt Template.pdf	Prompt template used to query LLM
labeling-guide.xlsx	Labeling guide used by annotators to label LLM responses

B: Guideline to Reproduce Results

<u>Expected Amount of Time to Reproduce Results : < 10 minutes</u>

Our artifacts reproduce results in Section 4. Given that findings are outlined in detail through tables and figures, our artifacts reproduce the 5 tables and 7 figures found in Section 4. We reproduce each figure twice - (1) the version found in the paper, and (2) a version annotated with percentages at the top of each bar.

1. Navigate to files in OSF repo (Linked here) [File size ~250 MB, download speeds may vary]
2. Download the zipped folder by clicking "*Download this folder*" (Icon pictured to the right) Download this folder
3. Unzip the file
4. Navigate to the unzipped folder in the terminal
   - e.g., `cd Downloads/osfstorage-archive`
5. Obtain the docker image: Choose (a) or (b)
   a) Produce a docker image from the dockerfile (Takes ~ 2 minutes)
   `docker build --no-cache -t acsac-275-container .`

Note: The dockerfile runs an ubuntu OS, installs python, relevant dependencies, and copies files from the working directory into the container

OR

   b) Load the image from the pre-saved container (Takes ~ 1 minute)
   `docker load -i acsac275container_latest.tar`

6. Run a container based on the image
   `docker run -v "$(pwd)"/plots:/app/plots -it acsac-275-container`

   Note: the -v flag syncs the local plots folder with the folder inside the container

7. To reproduce e1, e2, e3, type in
   `python3 reproduce_e1_e2_e3.py`

   Values for 5 tables will be reproduced (table # will be annotated in output)
   - Table 4,5 - Page 6 of paper
   - Table 6,7 - Page 7 of paper
   - Table 8 - Page 8 of paper

   Note: % produced are occasionally rounded and values are ensured to total to a 100 %

10 figures will be reproduced (in the local **plots** directory outside the container)

PDF Title	**Figure & Page**
e1-res-label-per-category-normalized-count-with-number.pdf	Fig 2, pg 6 (annotated w/ percentage)
e1-res-label-per-category-normalized-count.pdf	Fig 2, pg 6
e2-unique-label-types-per-category-normalized-count-with-number.pdf	Fig 3, pg 6 (annotated w/ percentage)
e2-unique-label-types-per-category-normalized-count.pdf	Fig 3, pg 6
e2-result-type-for-five-per-category-normalized-count-with-number.pdf	Fig 4, pg 7 (annotated w/ percentage)
e2-result-type-for-five-per-category-normalized-count.pdf	Fig 4, pg 7
e3-result-type-for-five-per-category-normalized-count-with-number.pdf	Fig 5, pg 7 (annotated w/ percentage)
e3-result-type-for-five-per-category-normalized-count.pdf	Fig 5, pg 7
e3-unique-label-types-per-category-normalized-count-with-number.pdf	Fig 6, pg 8 (annotated w/ percentage)
e3-unique-label-types-per-category-normalized-count.pdf	Fig 6, pg 8

8. To reproduce e4, type in
   `python3 reproduce_e4.py`

   4 figures will be reproduced (in the local **plots** directory outside the container)

PDF Title	**Figure & Page**
e4-valid-normalized-count-with-number.pdf	Fig 7, pg 8 (annotated w/ percentage)
e4-valid-normalized-count.pdf	Fig 7, pg 8
e4-relevance-normalized-count-with-number.pdf	Fig 8, pg 8 (annotated w/ percentage)
e4-relevance-normalized-count.pdf	Fig 8, pg 8

Note: Our scripts (`reproduce_e1_e2_e3.py` and `reproduce_e4.py`) parse corresponding csv files to print tables and produce figures.

If you wish to run locally instead of using a container, please install the required dependencies and run the scripts

1. `cd Downloads/osfstorage-archive`
2. `pip3 install -r requirements.txt`
3. `python3 reproduce_e1_e2_e3.py`
4. `python3 reproduce_e4.py`