Towards reproducing and replicating results

You have chosen 1 paper. The goal now is to (1) reproduce some results of the original paper; (2) define a protocol to make a variation of the original study, including what and how to vary, the hypothesis and expected results

Read the paper

A mandatory step is to read and analyze the paper.
In particular, you first need to understand the context and scope.

Start by examining: the abstract, introduction, and conclusion to gain an overview of the study's objectives, hypothesis, methods, and key findings. Look at where this paper fits within the broader landscape.

Review related work and prior research that the authors cite. If references are somehow missing or you feel that other works could be studied, try to find them using a search engine or specialized engine like Google scholar https://scholar.google.fr/
To do so, use key terms of the paper and take a careful look at the targeted problem.
This helps in understanding any assumptions, common methodologies, or datasets that may have influenced the study's design.

Questions you can address to check your understanding:

- What is the main research question or hypothesis of the paper?
    - Why is this question significant in its field?
- How does it address a gap in existing knowledge?
- How does this study fit within the broader landscape?
- What prior work does it build upon or differ from?
- What are the primary objectives of the study?
    - Are there secondary or exploratory questions?

Identify key components for Reproducibility and Replicability

A second step is to have a "reproducibility" perspective on the paper

Data: Check if the datasets are available and accessible. Note any specific pre-processing steps, transformations, or selection criteria applied.

Methods and Code: Pay attention to the methods section for details on experimental setup, algorithms, statistical analyses, or software versions used. Look for code repositories linked to the paper, if provided. Make sure to record any dependencies or environmental requirements.

Results and Metrics: Review how results are measured and reported. Identify key metrics, thresholds, and statistical tests used to support conclusions, as replicating these results will be essential to confirm findings.

Some papers include a reproducibility statement or checklist, often found in the appendix or supplementary material, outlining the reproducibility strategy.
Review the paper's data and code sharing policies. Note if any key components are missing (e.g., proprietary code or inaccessible datasets), as this could limit your ability to reproduce the study.

Questions you can address to check your understanding:

- What datasets, materials, or inputs are required for this study?
    - Are these resources openly available, and if not, how might this affect reproducibility?
- What specific methods and algorithms are used?
    - Are the algorithms standard or customized, and are detailed parameter settings provided?
- What tools, libraries, or software environments are needed to replicate the study?
    - Are the dependencies clearly documented, including versions and configurations?

Papers aiming for reproducibility should clearly document their methodologies, including parameters, experimental settings, and analysis procedures.
Identify any ambiguities or vague descriptions. If a certain method or parameter setting is unclear, highlight it, as it might impact your ability to reproduce the study accurately.

As you read, make note of potential variability factors like random seeds, data randomness, and subjective decisions in data processing or analysis. These can influence reproducibility and may be critical in any future variation of the study.

Reproduce

The goal of reproduction is to confirm the original study's results as closely as possible by following the same methodology, using the same data, and applying the same analyses. This step involves verifying that the results can be obtained under the same conditions, often using the original code and datasets if they are available.

This effort is also crucial prior to replication.

Here's how to approach this:

Prepare and Reuse Code/Data

Ensure you have access to the original datasets, either through a repository link or by contacting the authors. Confirm that the data format and contents match the descriptions in the paper.

Check if any pre-processing steps were applied to the data. Some studies involve significant transformations, filtering, or selection criteria that need to be replicated exactly.

Code Setup and Environment:

Locate the original code provided by the authors, usually in an attached repository or supplement. Set up the same environment as specified in the study (e.g., programming languages, libraries, or software versions).

If the study does not provide a detailed setup, carefully record your chosen versions and configurations, as minor differences can affect the results. Pay attention to any dependencies, such as third-party libraries or software packages. If the versions aren't specified, try using commonly referenced ones based on the study's date or language version.

Use Docker and automated scripts to have a reproducible environment

For replicability reasons (see next section), make note of environmental settings, random seeds, or hardware requirements that may influence the output, especially if working with machine learning models or simulations.

Analyze and Document Reproduced Results

Compare Reproduced Results to Original Findings:

Evaluate the reproduced results against those reported in the paper. Ideally, results should match closely, though some minor variations can occur due to differences in hardware or software environments.

Pay close attention to key metrics and statistical results, as these are often the basis for confirming the study's main claims.

Document Any Deviations or Observations:

Note any discrepancies between your reproduced results and those reported in the study. Record whether these deviations seem random or if they could be linked to environmental factors, missing parameters, or other variations.

Describe any challenges you faced during reproduction, such as unclear code, undocumented parameters, or data inconsistencies, and suggest ways future work might overcome these issues.

Draw Conclusions: Is It Reproducible?

Based on your reproduction process, summarize whether you confirm the original study's results. Reflect on any difficulties encountered and the degree to which your results align with the original findings.

If you achieved closely matching results, conclude that the study is likely reproducible under the same conditions. If significant variations or inconsistencies arose, consider the factors that could explain these differences and whether they suggest limitations in the study's reproducibility.

Replicate

The goal of replication is to test the study's robustness and generalizability by introducing controlled variations and observing whether the findings still hold. Unlike reproduction, which aims to exactly replicate conditions, replication assesses the study's reliability under different scenarios. Your goal here is to define a protocol to make a variation of the original study, including what and how to vary, the hypothesis and expected results.

The protocol will be validated next week!

Identify Variability Factors and Threats to Replicability

Examine Key Sources of Variability:

- **Data Sensitivity**: Check if the original findings might depend heavily on specific data properties (e.g., dataset size, distribution, or feature types). Consider introducing different datasets to see if the results generalize.
- **Methodological Adjustments**: Identify steps in the methodology that might introduce variability, such as model parameters, feature selection techniques, or random seeds.
- **Environmental Factors**: Note if there are any specific hardware, software, or library dependencies. For example, hardware-based variations might impact computational performance or training results for machine learning studies.

List of Potential Threats to Replicability:

Look for aspects that may pose a risk to replicability, such as:

- **Randomness**: Ensure that any random seed is appropriately set or varied to observe robustness against different random states.

- **Sampling Bias**: Some studies may rely on specific data samples; try varying sample groups to test consistency.
- **Parameter Sensitivity**: Certain parameter values (e.g., regularization strength or learning rate in machine learning models) may affect outcomes; consider varying these to see if results change.

Document all possible variations:

- identify variability factors
    - for each variability factor, define possible values (eg True/False)
    - explain why you think it is worth considering
- identify logical constraints across variability factors

Design the Variation/Deviation Protocol

Hypothesis for the Variation

Formulate a hypothesis around each variation. For instance:

- “The original results are robust to parameter variations in model X settings.”
- Hypotheses should state both expected results and any limits on deviation that would suggest fragility in the original findings.