Reproducibility Report for ACM SIGMOD 2021 Paper: "Explaining Black-Box Algorithms Using Probabilistic Contrastive Counterfactuals"

ANGELOS CHRISTOS ANADIOTIS, École Polytechnique, France

Most of the results of the paper were fully reproduced. There were only a few differences in the generated numbers, but they did not affect the conclusions of the paper. The reproducibility process had been mostly straightforward, with only minor configuration knobs that had to be addressed.

1 INTRODUCTION

The paper "Explaining Black-Box Algorithms Using Probabilistic Contrastive Counterfactuals" by Sainyam Galhotra (University of Massachusetts Amherst), Romila Pradhan (University of California San Diego), and Babak Salimi (University of California San Diego) was successfully reproduced [1].

2 SUBMISSION

The authors submitted a zip file that includes all the material required to reproduce the paper results. The scripts work almost out of the box, with the only exception being that the output files are not copied to the output directory pointed out in the README file. Still, the files were easily accessible by simply going through the scripts and checking the output directories there. Every figure was being generated by a different script, and there was a script to generate all of them together. This granularity helped a lot in running the experiments. Setting up the Python environment was either straightforward or simply impossible; however, that was not a problem of the paper, but of Python version and package discontinuity. Therefore, the authors are not to be blamed for that.

Summary of the submission contents:

- web link with code and scripts at: https://sainyamgalhotra.github.io/lewis.zip
- detailed readme file inside the compressed file
- data generators
- data sources inside the compressed file

3 HARDWARE AND SOFTWARE ENVIRONMENT

Provide information about the environment used in the paper and in the reproducibility. A table like Table 1 might be a good idea. Please feel free to add any rows you see fit.

Table 1. Hardware & Software environment

	Paper	Repro Review
CPU	Intel E5-2698 v3	Intel 5218
cores	64	32
GHz	2.3	2.3
RAM	256GB	192GB

4 REPRODUCIBILITY EVALUATION

4.1 Process

The reproducibility process was relatively straightforward. The zipped file that the authors provided includes all the data required to run the experiments, as well as a separate script for every figure included in the paper. Setting the Python environment had been problematic in a set of servers, but this is because of Python version problems. As soon as the Python environment had been properly set up, then the experiments were running without any problems. The only issue was the output files were not written in the output directory set by the authors, and some tuning was required in the scripts. Nothing big, though, and it was not necessary to contact the authors.

4.2 Results

The key findings of the paper were reproduced. Some figures were not exactly the same, but they did not affect the arguments made by the paper.

REFERENCES

[1] Sainyam Galhotra, Romila Pradhan, and Babak Salimi. 2021. Explaining Black-Box Algorithms Using Probabilistic Contrastive Counterfactuals. In *SIGMOD '21: International Conference on Management of Data, Virtual Event, China, June 20-25, 2021*. ACM, 577–590. https://doi.org/10.1145/3448016.3458455