Author Contributions Checklist Form

This form documents the artifacts associated with the article (i.e., the data and code supporting the computational findings) and describes how to reproduce the findings.

Part 1: Data

☐ This paper **does not** involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Abstract

4.1 million continuous ratings of 100 jokes from 73421 users: collected between April 1999-May 2003.

Availability

☒ Data **are** publicly available
☐ Data **cannot be made** publicly available

If the data are publicly available, see the *Publicly available data* section. Otherwise, see the *Non-publicly available data* section, below.

Publicly available data

☒ Data are available online at: https://goldberg.berkeley.edu/jester-data/
☐ Data are available as part of the paper's supplementary material.
☐ Data are publicly available by request, following the process described here:

☐ Data are or will be made available through some other mechanism, described here:

Non-publicly available data

Discussion of lack of publicly available data:

Description

File format(s)

☒ CSV or other plain text: jester-1.csv,jester-2.csv
☐ Software-specific binary format (.Rda, Python pickle, etc.):
☐ Standardized binary format (e.g., netCDF, HDF5, etc.):
☐ Other (described here):

Data dictionary

☒ Provided by the authors in the following file(s): codes/Real_data
☐ Data file(s) is (are) self-describiing (e.g., netCDF files)
☐ Available at the following URL:

Additional information (optional)

Part 2: Code

Abstract

The codes are provided to reproduce the methodology of key figures and tables in this paper.

Description

Code format(s)

☒ Script files
- ☒ R ☐ Python ☐ Matlab
- ☐ Other:

☐ Package
- ☐ R ☐ Python ☐ MATLAB toolbox
- ☐ Other:

☐ Reproducible report
- ☐ R Markdown ☐ Jupyter notebook
- ☐ Other:

☐ Shell script

☐ Other (described here):

Supporting software requirements

Version of primary software used

R 3.6.1, Rstudio 1.2.5001

Libraries and dependencies used by the code

None.

Supporting system/hardware requirements (optional)

Parallelization used

☒ No parallel code used
☐ Multi-core parallelization on a single machine/node
Number of cores used:
☐ Multi-machine/multi-node parallelization
Number of nodes and cores used:

License

☒ MIT License (default)
☐ BSD
☐ GPL v3.0
☐ Creative Commons
☐ Other (described here):

Additional information (optional)

Part 3: Reproducibility workflow

Scope

The provided workflow reproduces:
☐ Any numbers provided in text in the paper
☒ The computational method(s) presented in the paper (i.e., code is provided that implements the method(s))
☐ All tables and figures in the paper
☐ Selected tables and figures in the paper, as explained and justified here:

Workflow details

Location

The workflow is available:
☐ As part of the paper's supplementary material
☒ In this Git repository: https://github.com/qooyqpqy123/Ranking_Inference
☐ Other:

Format(s)

☐ Single master code file
☐ Wrapper (shell) script(s)
☐ Self-contained R Markdown file, Jupyter notebook, or other literate programming approach
☐ Text file (e.g., a readme-style file) that documents workflow
☐ Makefile
☒ Other (more detail in 'Instructions' below)

Instructions

There are two folders named "Real-Data" and "Simulation".
In folder "Real-Data, we provide the original dataset "jester-data-1.zip" and ``jester-data-2.zip"
The codes named Table_4.R, Table5-6.R reproduces the methodology used in Table 4 and Table 5-6.

In ths folder “Simulation”, we provide codes to reproduce methodologies used in our simulation results. In specific, Figure1-2.R, Figure_3.R, Figure_4.R and Table 1-3.R reproduce the methodogy used in Figure 1-4, Table 1-3, respectively.

Expected run-time

Approximate time needed to reproduce the analyses on a standard desktop machine:
□ <1 minute
□ 1-10 minutes
□ 10-60 minutes
□ 1-8 hours
☒ >8 hours
□ Not feasible to run on a desktop machine, as described here:

Additional documentation (optional)

Notes (optional)