

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

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

Abstract

We consider a publicly available dataset gathered by the INVALSI institute, which is a public research centre for the assessment of the Italian education system. The 2016-2017 dataset is part of a national examination program conducted in Italy with the aim of “carrying out periodic and systematic checks on knowledge and skills of students’”, as declared in the official documentation of the INVALSI statistical service.

The documentation (in Italian) is available at: <https://invalsi-serviziostatistico.cineca.it>.

We focus on data related to 8th grade students from schools in the city of Bologna: more specifically we consider those questions related to the comprehension of the Italian language. The resulting dataset comprises a total of 8126 observations (students), belonging to 84 educational institutions.

Availability

- ☒ Data **are** publicly available.

Publicly available data

- ☒ Data are publicly available by request, following the process described here:

The considered dataset must be requested following the procedure described at the link (in Italian):

<https://invalsi-serviziostatistico.cineca.it>

Please select “Dati” on the left menu and then select the link “Istruzioni e moduli per la richiesta dei dati” for a detailed description of the request procedure.

Description

File format(s)

- ☒ CSV or other plain text.
- ☒ Software-specific binary format (.Rda, Python pickle, etc.): sav

Data dictionary

- ☒ Available at the following URL: <https://invalsi-serviziostatistico.cineca.it>

Part 2: Code

Abstract

The uploaded code contains the main relevant functions for the simulation study and the data analysis. Most of the code is in the form of an R package (`HNRMI_0.0.4.tar.gz`). A few additional functions are provided (`miscellanea.R`, `miscellanea.cpp`, and `MASTER_FILE_SIMULATION_STUDY.R`).

Description

Code format(s)

- ☒ Script files
 - ☒ R
 - ☒ Other: C++
- ☒ Package
 - ☒ R
 - ☒ C++

Supporting software requirements

Version of primary software used R version 4.1.2 (2021-11-01) - “Bird Hippie”

Libraries and dependencies used by the code Package: `HNRMI` (0.0.4), `tidyverse` (1.3.1), `gridExtra` (2.3), `mcclust` (1.0), `mcclust.ext` (1.0), `bridgesampling` (1.1.2).

Additional dependencies are required for the installation of the `HNRMI` package, as described in the `DESCRIPTION` file of the package.

Parallelization used

- ☒ No parallel code used

License

- ☒ MIT License (default)

Part 3: Reproducibility workflow

Scope

The provided workflow reproduces:

- ☒ The computational method(s) presented in the paper (i.e., code is provided that implements the method(s))

The `MASTER_FILE_SIMULATION_STUDY.R` allows the reproduction of the simulation study. Different scenarios of the simulation study can be obtained by changing the settings at the beginning of the script.

Workflow

Format(s)

- ☒ Single master code file

Instructions

The `MASTER_FILE_SIMULATION_STUDY.R` should be execute. Please first install all the necessary packages, including the uploaded `HNRMI` library.

Expected run-time

Approximate time needed to reproduce the analyses on a standard desktop machine:

☒ 1-8 hours