The Machine Learning Reproducibility Checklist (v2.0, Apr.7 2020)

For all models and algorithms presented, check if you include: **Secs. "Defining ELENE" & "Learning with ELENE"**

- [x] A clear description of the mathematical setting, algorithm, and/or model.
- [x] A clear explanation of any assumptions.
- [x] An analysis of the complexity (time, space, sample size) of any algorithm.

For any theoretical claim, check if you include: **Sec. "Expressive Power"**

- [x] A clear statement of the claim.
- [x] A complete proof of the claim.

For all datasets used, check if you include: **Sec. "Experimental Results"**

- [x] The relevant statistics, such as number of examples. **Referenced**
- [x] The details of train / validation / test splits. **Defined in code**
- [x] An explanation of any data that were excluded, and all pre-processing step. **Referenced**
- [x] A link to a downloadable version of the dataset or simulation environment. **Standard PyG**
- [x] For new data collected, a complete description of the data collection process, such as instructions to annotators and methods for quality control. **Not applicable**

For all shared code related to this work, check if you include:

- [x] Specification of dependencies. **Defined in setup.sh**
- [x] Training code. **src/train_helper.py and train/*.py**
- [x] Evaluation code. **train/*.py — incl. metrics for each model**
- [x] (Pre-)trained model(s). **Not applicable**
- [x] README file includes table of results accompanied by precise command to run to produce those results. **See README.md and REPRODUCIBILITY.md**

For all reported experimental results, check if you include: **Sec. "Experimental Results"**

- [x] The range of hyper-parameters considered, method to select the best hyper-parameter configuration, and specification of all hyper-parameters used to generate results.
- [x] The exact number of training and evaluation runs.
- [x] A clear definition of the specific measure or statistics used to report results.
- [x] A description of results with central tendency (e.g. mean) & variation (e.g. error bars).
- [x] The average runtime for each result, or estimated energy cost. **We report mem. as a proxy in multi-user env.**
- [x] A description of the computing infrastructure used.

Reproduced from: *www.cs.mcgill.ca/~jpineau/ReproducibilityChecklist-v2.0.pdf*