

Figure 1 displays the results of running each algorithm on a set of factuals. The figure shows clear dominance of the evolutionary model all the models across all datasets.

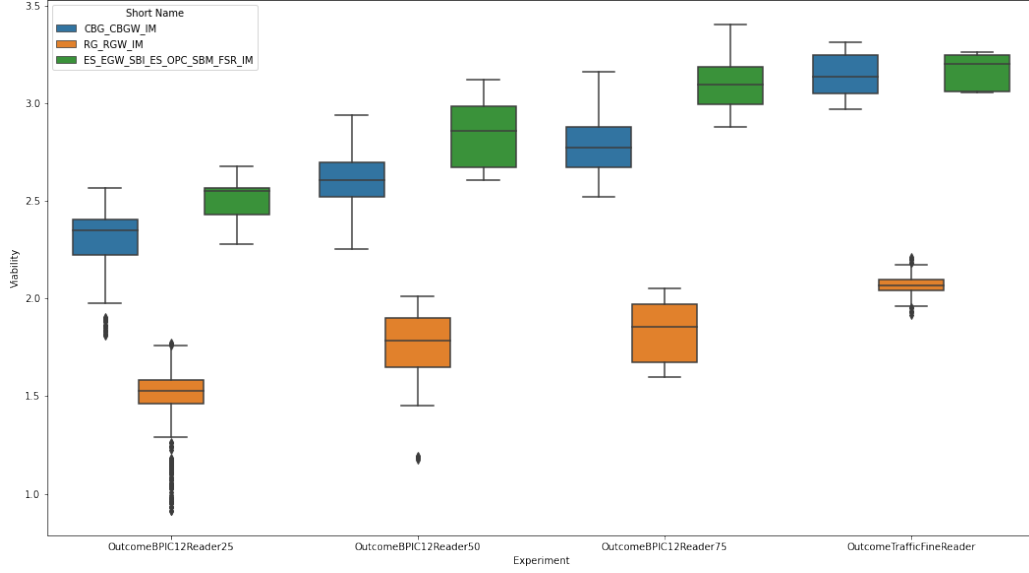


Figure 1: This figure shows boxplots of the viability of each models' generated counterfactual.

Here, [a specific model] returns an average viability of [some value]. This result is not surprising, as we optimize with the same metric used to evaluate the results.

A notable observation occurs in the processing time. Although all models require more time to produce their results, if we increase the maximum length of sequences, the evolutionary algorithms require substantially more time. ?? shows, the processing time is much higher than the baseline models. On average, [evo model] require [duration in seconds], while [other models] require [XX], [XX] and [XX], respectively.