

Most of the results are reasonable. Surprisingly, the models did not necessarily create counterfactuals much shorter than their factual counterparts. In fact, most of Hsieh, Moreira, and Ouyang’s counterfactuals are shorter in length. This characteristic can be an advantage for use cases, such as medicine. The fluctuations in the loan amount were expected, as well. We did not implement any safeguard option to keep certain attributes fixed. The values our generative models produce are more or less an indication of what the prediction model deems as a useful change to turn over the outcome at a specific step in the process.

We are also not surprised that all models manage to capture the first few activities. These are mostly the same across all cases. If our generative models had not recognised these, one could question their utility.

All models successfully flip the outcome of the prediction model and are surprisingly close to the factual compared to the model by Hsieh, Moreira, and Ouyang. However, we must remember that these observations tell us more about the model than the true process. More specifically, our model can show which events and attributes have to be present at a specific point within the process.

All in all, we claim that the generator model can teach us more about the prediction model primarily. Further improvement might show even more nuance in the model’s behaviour. We discuss some of them in the discussion chapter.