

Most of the results are reasonable. Surprisingly, the models did not necessarily create counterfactuals that are 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 was expected, as well. We did not implement any safeguard option to keep certain attributes fixed. The values that were produced are more or less an indication of what the model deems as useful to change the outcome at a specific step in the process.

We are also not surprised, all models manage to capture the first few activities. These are mostly the same across all cases. If our models had not recognized these it’s usefulness could be questioned.

All models successfully manage to flip the outcome of the prediction model and surprisingly close to the factual compared to the model by Hsieh, Moreira, and Ouyang. However, we have to keep in mind that these observations tell us more about the model rather than the true process. More specifically, our model is capable of showing, 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 model primarily. Further improvement might show even more nuance in the models behaviour. We discuss some of them in the discussion chapter.