

In the following, we discuss the results in three aspects:

1. The quality in terms of the viability of the counterfactual sequences generated by our models.
2. Their quality compared to two baseline approaches and the state-of-the-art DICE4EL approach.
3. Their implications in terms of the general utility of our solution.

Our first two experiments show that we can optimise towards viability successfully. We defined four criteria for the viability of counterfactuals (similarity, sparsity, feasibility, and delta in likelihood) and showed that a model optimising towards those criteria can return superior results. Furthermore, we created models capable of optimising complicated operationalisations of these criteria without the limitation of a function with a clearly defined gradient.

Based on the results, we have seen in the latter experiments that we can confidently say the models can generate viable counterfactuals. Compared to other methods in the literature, we show that our counterfactuals attempt to be closer to the factual we desire to understand. We have to note that these counterfactuals are primarily a reflection of the underlying prediction model. One might argue that this does not translate to a real-world scenario. However, a model never truly does. If our framework attempts to explain how a prediction model behaves, then its applicability to real-world scenarios depends on that model’s viability. But regardless of the prediction model’s performance, we can clearly gain an understanding of its internal reasoning pattern.

The viability measure we proposed shows that structural difference can help us better understand when and where we must apply counterfactual changes. Other approaches often seem to overlook the importance of the sequence structure. However, the *CBI-RWS-OPC-SBM-FSR* model shows that it may be reasonable to incorporate structural differences in our viability measures. Especially, if we talk about sequences and processes. The gaps within the counterfactuals our models produced clearly indicate that. If a model attempts to align sequences, it becomes much easier to compare them side-by-side.

In contrast to the closest alternative approach by Hsieh, Moreira, and Ouyang, we show that we can create these counterfactuals without incorporating domain-specific knowledge such as an understanding of milestone patterns. Domain knowledge can always help us create better solutions.

However, we do not always have access to them. We believe that showing it is possible to create viable counterfactuals without domain-specific knowledge is our most significant contribution. Furthermore, our models can generate solutions not currently present within the data. Case-based solutions often overlook this aspect, as they are heavily biased towards the data input. Second, they can fail to deliver the necessary structural nuance when understanding sequences.