In the following, we discuss the results in three aspects: (1) the quality in terms of 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, and (3) their implications in terms of the general utility of our solution.

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

Based on the results, we have seen towards the latter experiments that we can confidently say the models are capable of generating viable counterfactuals. In fact, 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 mostly 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 it its applicability to real world scenarios is depends on the viability of the model itself. But regardless of the prediction model's performance, we can clearly gain an understanding about its internal reasoning pattern.

The viability measure we proposed shows that structural difference can help us to better understand when and where we have to 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 that were produced are a clear indication of 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. Obviously, 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 greatest contribution. Furthermore, our models are capable of generating solutions that are not currently present within the data. This is often an oversight for case-based solutions, as they obviously are heav-

ily biased towards the data input. Second, they can also fail to deliver the necessary structural nuance when it comes to understanding sequences.