This section dives into the generative models that we will explore in this thesis. They cover fundamentally different approaches to the counterfactual generation of sequences. The first approach acts as a simple baseline using existing instances in the dataset. Hence, the counterfactual candidates are not generated probablistically but chosen among the data itself. Next, we explore a deep generative approach. Here, we attempt to capture the latent state-space of a model and use this space to sample counterfactual candidates. Last, we explore a technique which does not require to optimise a differentiable objective function. Instead we use the validity measure as a fitness function and maximise the fitness of each counterfactual candidate.

0.0.1 Generative Model: Case-Based Approach

Case-based techniques leverage the information that the data provides. The idea is to find suitable candidates the fulfill the counterfactual criterion of precision. In this thesis we find the candidates by searching for alternative process instances that closely resemble the factual instance and still lead to the desired outcome. For this purpose, we use a distance function to compute the distance of every alternative to the factual instance and declare the top-K most-similar insances as counterfactual candidates. Inherently, this approach is restricted by the *representativeness* of the data. It is not possible to generate counterfactuals that have not been seen before. This method works for cases, in which the data holds enough information about the process. If this condition is not met, it is impossible to produce suitable candidates. We apply the validaity metric established in **??** on the subset of of instances which yield the desired outcome. The results of this approach are considered as a suitable baseline for the other methods explored in this thesis. Note, that all case-based approaches automaticall fulfill the criterion of being feasible, as they are drawn from data, that has been *seen* already. Hence, we expect their feasibility to always be higher than the counterfactuals generated by other methods.

0.0.2 Generative Model: Deep Generative Approach

The generative approach assumes, it is possible to capture a latent state z and use this state to generate suitable counterfactual candidates. We condition the generation procedure on the factual instance to generate counterfactuals that show sparse differences to the original sequence. The core idea is to sample randomly $x * p(z|x)$ to generate counterfactual candidates. We can sort each candidate by their *validity* and choose top-K contenders as valid couunterfactuals. There are a multitude of approaches to generate the

counterfactuals. However, we will limit our exploration to sequential Variational Autoencoders (VAEs) and sequential Generative Adversarial Models (GANs). Both technieques allow us to sample from a smooth latent space conditioned on the factual sequence. VAEs approximate $p(z|x)$ by trying to reconstruct the input using Monte-Carlo methods. GANs require a generator model and a distriminatior model. The generator model attempts to fool the discriminator model by generating, results that closely resemble true process instances. In contrast, the discriminator tries to distinguish generated instances from real instances.

0.0.3 Generative Model: Evolutionary Approach

Evolutionary approaches use ideas that resemble the process of evolution ^{CITE}. Here, we randomly generate candidates by modifying the factual sequence and evaluate their fitness based on a fitness function. Those, cadidates that are deemed as fit enough are subsequently modified to produce offspring. Afterwards, the procedure will repeat until a fitness convergence was reached. This procedure automatically maximises the fitness function. It differs from , because it does not require to use differiantible functions.