Abstract

Within the field of Process Mining, deep recurrent networks (such as LSTM) have been used to predict the next state or the outcome of a multivariate sequence. However, these models tend to be complex and are difficult for users to understand of the underlying process model. Counterfactuals answer "what-if" questions, which are used to understand the reasoning behind the predicted outcome of a process. Current methods to generate counterfactual explanations do not take the structural characteristics of multivariate discrete sequences into account. In this work we propose a framework that uses evolutionary methods to generate counterfactuals, while incorporating criteria that ensure their viability. Our results show that it is possible to generate counterfactuals that are viable and automatically align with the factual. The generated counterfactuals outperform baseline methods in viability and yield comparable results compared to other methods in the literature.