We use counterfactuals primarily to explain predictive models. This explanation requires a to define the prediction model we use in this thesis.

0.0.1 Practical Matters

Mention everything necessary to repeat this experiment: For instance, unbalanced data

0.0.2 Results

Show how this model is fine to use by reporting training and validation scores.

0.0.3 Discussion

Mention whatever is noticeable: 1. The models are extremely good for longer maximal sequence lengths of the BPIC dataset. The question is whether the model genuinely learns intrinsic features or just waits for trivial patterns like: How many padding events does the sequence have? What cyclical patterns are present? Does a particular event occur or not. Could be solved using attention and a visualisation method.