Supplementary materials

Selection of Learning Rate: We compare the average 4-fold cross validation accuracy of TabNet and DeepTable trained by two different learning rates, 0.01 and 0.1. The performance values on both the training and the validation sets as shown in Table 1 indicate that TabNet with learning rate of 0.01 overfits on the training data after few epochs. However, it obtains good fit on the data using a higher learning rate 0.1. DeepTable in both cases obtains about the same accuracy values using two learning rates, however, it achieves higher performance both on the validation and the training sets using the learning rate of 0.01. Therefore, we choose 0.1 for TabNet and 0.01 for DeepTable as their learning rates in our evaluations.

Table 1: 4-fold cross validation accuracy of DeepTable and TabNet measured on the validation and the training sets (validation set%, training set%).

Model	0.1	0.01
DeepTable	(76.42, 76.57)	(76.51, 78.02)
TabNet	(75.09, 78.78)	(73.18, 85.74)