Appendix A. SIMULATION PARAMETER

Table A.1. Simulation setup for three-tank dataset. Each parameter describes the mode of operation, where 1 in the case of q_i means full pumping power. In the case of valves kv_i, 1 corresponds to fully open and 0 to the closed valve.

state	q1	q3	kv1	kv2	kv3
$Q1$	0.1	0	0	0	0
$V12$	0	0	0.1	0	0
$V23$	0	0	0	0.1	0
$V3$	0	0	0	0	0.1
$V12_{faulty}$	0	0	0.01	0	0
$V23_{faulty}$	0	0	0	0	0
$V3_{faulty}$	0	0	0	0	0.05
$Q1_{faulty}$	1	0	0	0	0

Appendix B. REPRODUCABILITY OF CATVAE MODELS

All experiments were conducted using Python 3.9 on a 64-bit Linux VM equipped with an 8GB Nvidia V100 GPU and a CPU featuring 10 cores. The necessary software package prerequisites are detailed within the repository of the provided code.

In order to establish the replicability of our assessment concerning Discret2Di, the subsequent hyperparameters were employed.

Table B.1. CatVAE hyperparameters used for Siemens SmA discretization. The model is trained for 400 epochs. *beta* and temperature τ were changed from 0.1 to 1 in 0.1 steps.

Parameter	Values
kernel size	12
batch size	128
temperature τ	0.5
β	0.1
early stopping patience	40
training / validation samples	29,740 / 7,436
encoder block (our model)	(39, 512, 512, 4)
linear layer (encoder - latent)	(4, 20)
decoder block 2 (our model)	(20, 512, 512, 8)
linear layer (decoder - μ)	(8, 39)
linear layer (decoder - σ)	(8, 39)

Table B.2. CatVAE hyperparameters used for BeRfiPl dataset (**?**) discretization. The model is trained for 400 epochs. *beta* and temperature τ were changed from 0.1 to 1 in 0.1 steps.

Parameter	Values
kernel size	12
batch size	32
temperature τ	0.5
β	0.1
early stopping patience	40
training / validation samples	3,385 / 847
encoder block (our model)	(154, 512, 512, 4)
linear layer (encoder - latent)	(4, 20)
decoder block 2 (our model)	(20, 512, 512, 8)
linear layer (decoder - μ)	(8, 154)
linear layer (decoder - σ)	(8, 154)

Table B.3. CatVAE hyperparameters used for SWaT dataset discretization. The model is trained for 400 epochs. *beta* and temperature τ were changed from 0.1 to 1 in 0.1 steps.

Parameter	Values
kernel size	12
batch size	128
temperature τ	0.5
β	0.1
early stopping patience	40
training / validation samples	48,000 / 12,000
encoder block (our model)	(4, 512, 512, 8)
linear layer (encoder - latent)	(8, 20)
decoder block 2 (our model)	(20, 512, 512, 16)
linear layer (decoder - μ)	(16, 4)
linear layer (decoder - σ)	(16, 4)

Table B.4. CatVAE hyperparameters used for Three-Tank Dataset discretization. The model is trained for 400 epochs. *beta* and temperature τ were changed from 0.1 to 1 in 0.1 steps.

Parameter	Values
kernel size	12
batch size	64
temperature τ	0.3
β	0.2
early stopping patience	40
training / validation samples	3,500 / 1500
encoder block (our model)	(3, 256, 256, 8)
linear layer (encoder - latent)	(8, 10)
decoder block 2 (our model)	(10, 256, 256, 16)
linear layer (decoder - μ)	(16, 3)
linear layer (decoder - σ)	(16, 3)

Table B.5. GMM hyperparameters employed for the process of discretization delineated across the used datasets (library: sklearn).

Parameter	Values
random state	0
covariance type	"full"
SWaT dataset: n components	5
SmA dataset: n components	4
BeRfiPl dataset: n components	7

Table B.6. Hyperparameters used for association rule learning by FP-Growth (library: mlxtend).

Parameter	Values
threshold metric (association learning)	"confidence"
min threshold (association learning)	0.005
threshold likelihood	-50