The background features abstract, overlapping green geometric shapes in various shades, including light lime green, medium green, and dark forest green. These shapes are primarily located on the left and right sides of the slide, framing the central text.

Learning Perturbations to Explain Time Series Predictions

Joseph Enguehard

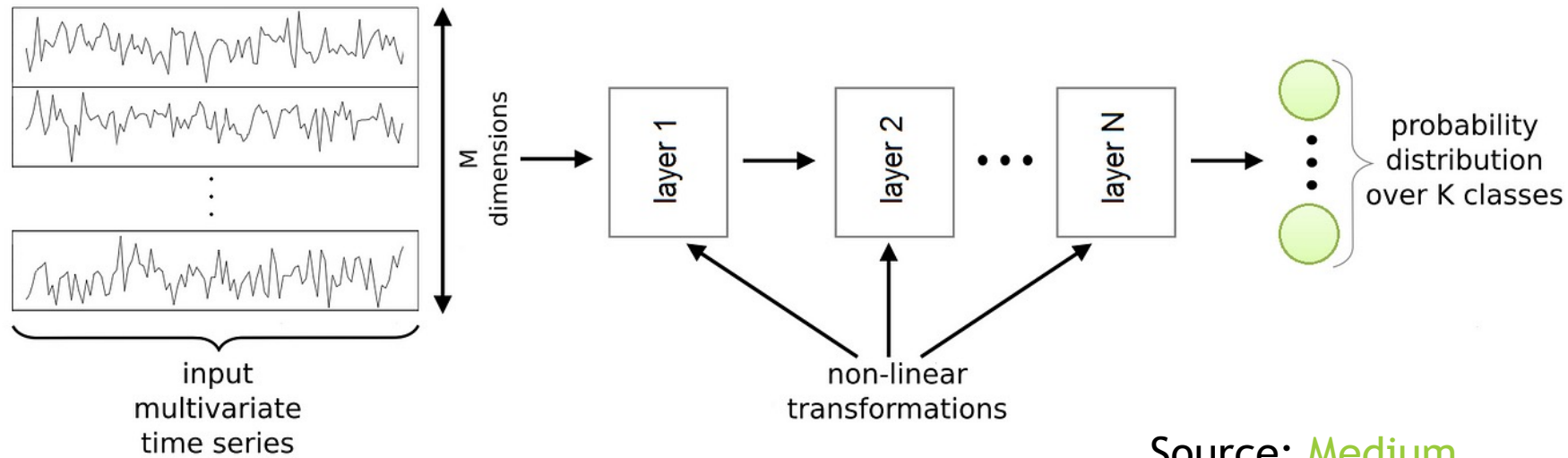
Learning Perturbations to Explain Time Series Predictions



Learning Perturbations to Explain Time Series Predictions

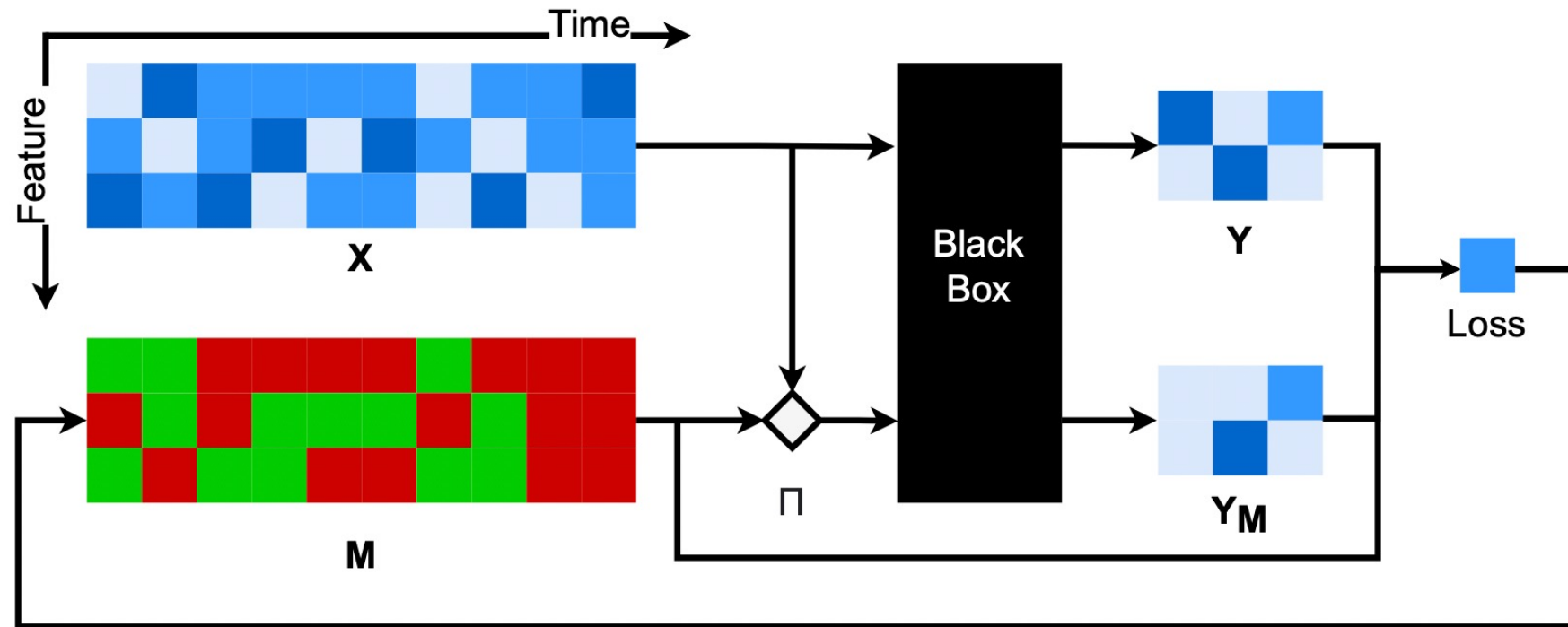


Learning Perturbations to Explain Time Series Predictions



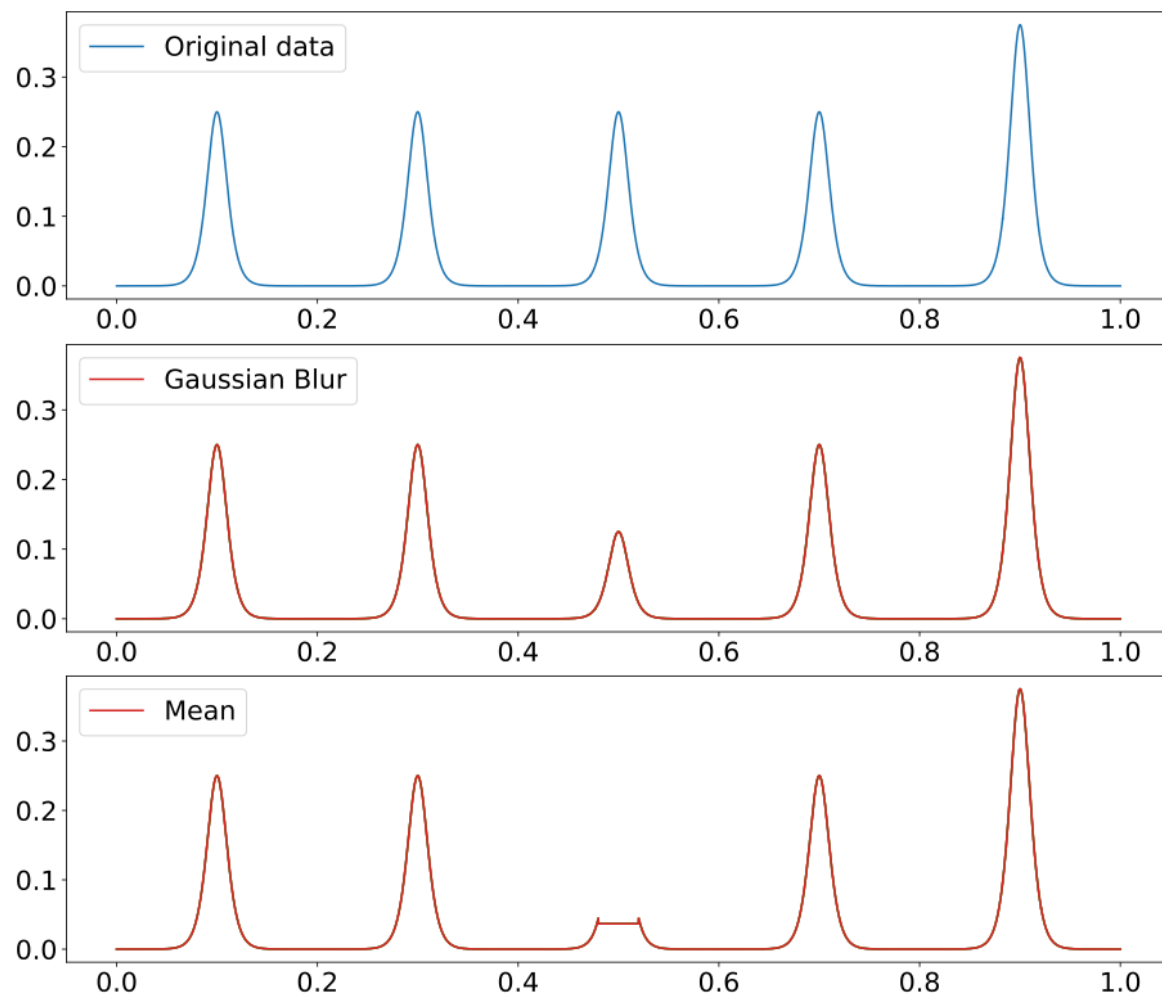
Source: [Medium](#)

Learning Perturbations to Explain Time Series Predictions

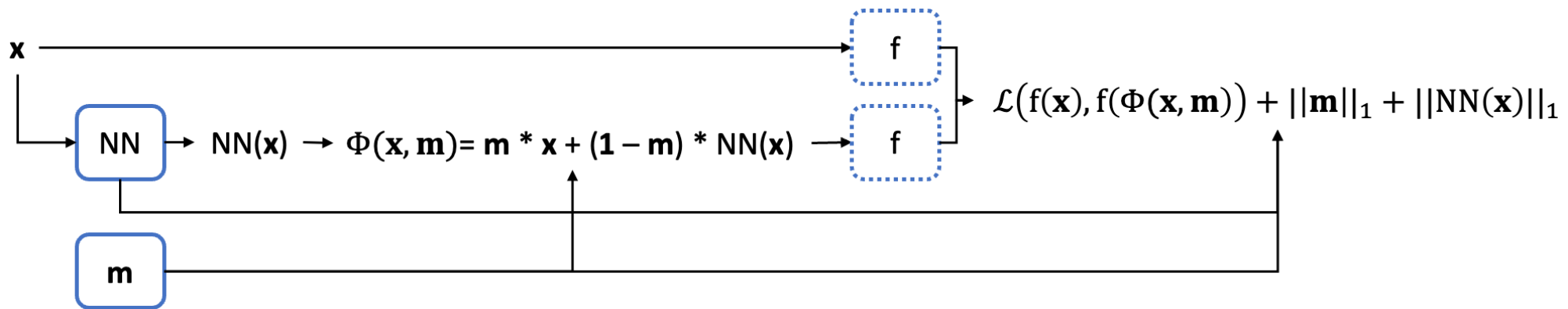


Source: [Crabbé et. al.](#)

Learning Perturbations to Explain Time Series Predictions



Learning Perturbations to Explain Time Series Predictions



Learning Perturbations to Explain Time Series Predictions

$$\operatorname{argmin}_{m, \Phi \in NN} \mathcal{L}(f(x), f(\Phi(x, m))) + \|m\|_1 + \|NN(x)\|_1$$

$$\Phi(x, m) = m \times x + (1 - m) \times NN(x)$$

Learning Perturbations to Explain Time Series Predictions

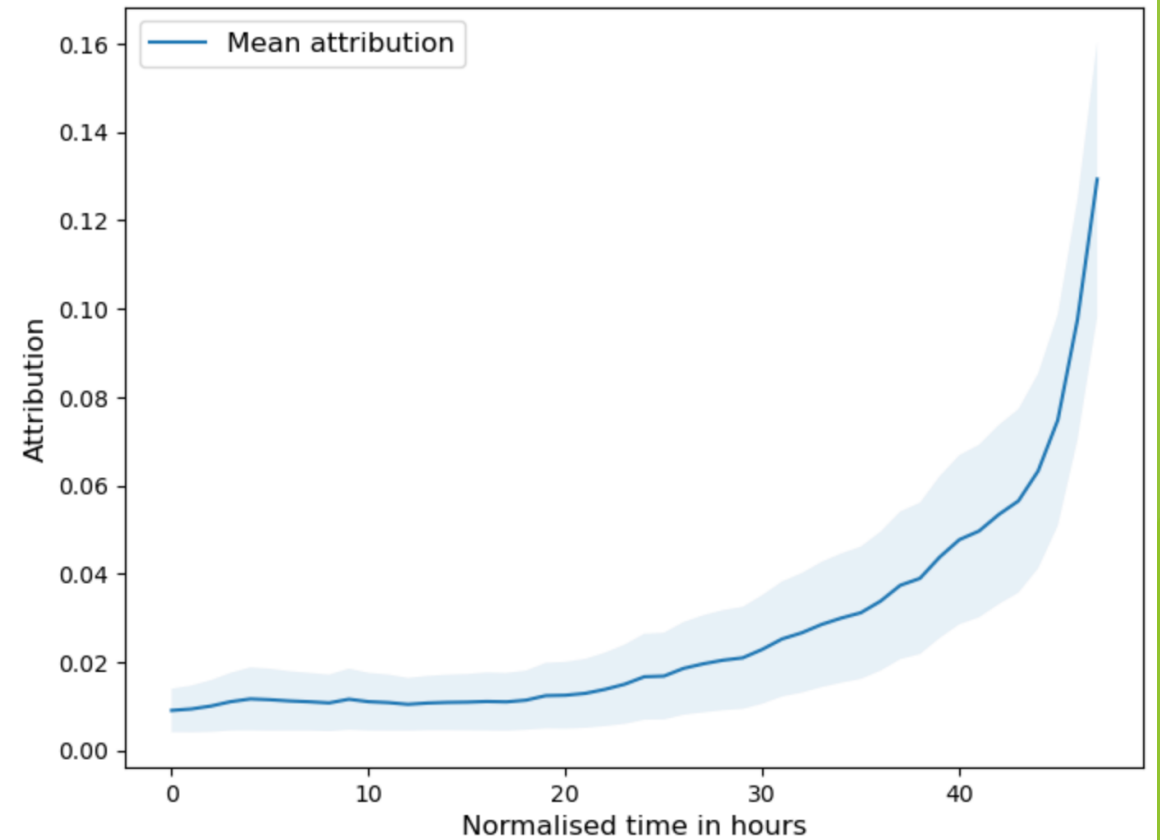
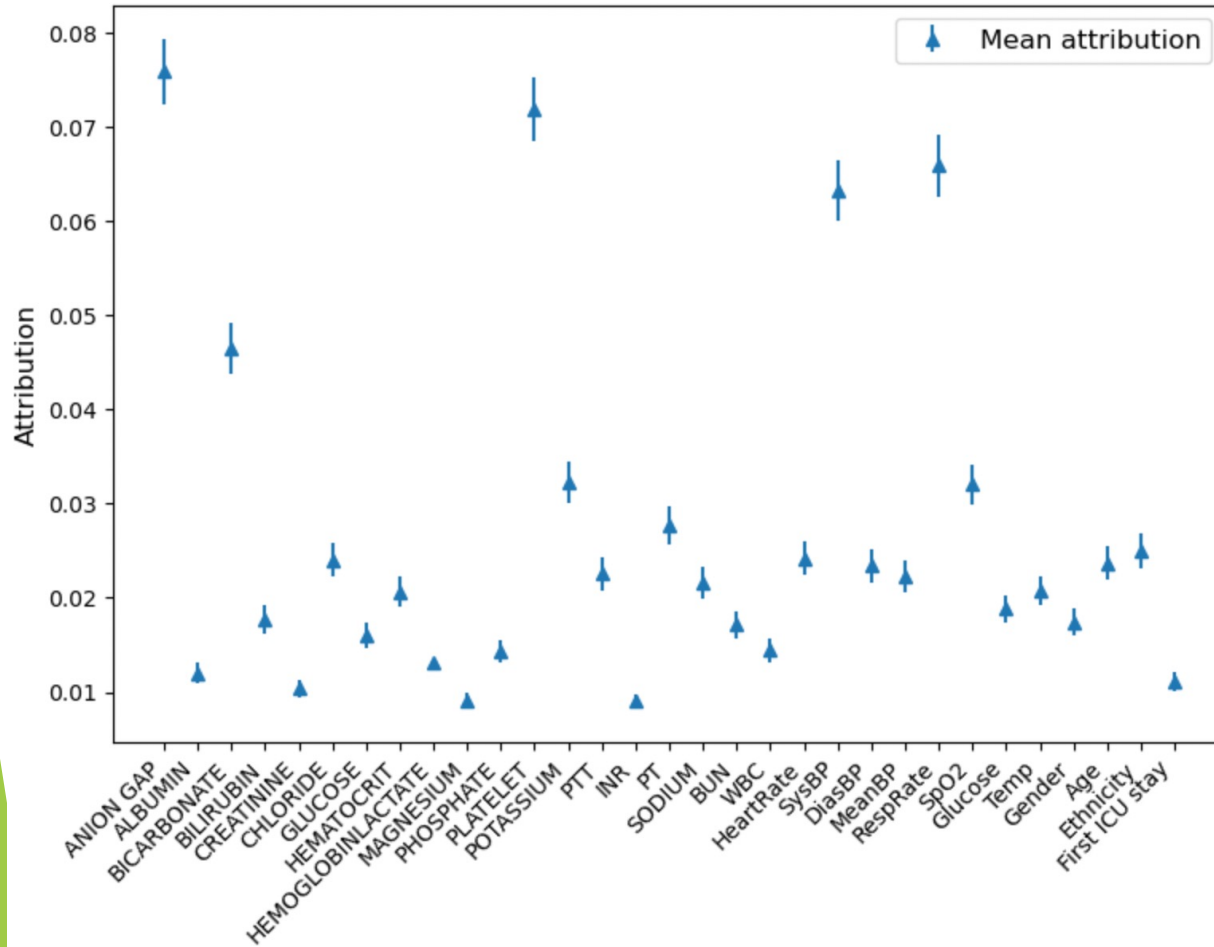
► HMM Dataset

Method	AUP \uparrow	AUR \uparrow	I \uparrow	S \downarrow
DeepLift	0.920 (0.019)	0.454 (0.011)	359 (9.55)	145 (0.949)
DynaMask	0.711 (0.020)	0.763 (0.026)	954 (50.0)	45.4 (0.781)
IG	0.918 (0.019)	0.454 (0.011)	359 (11.6)	146 (0.871)
GradientShap	0.849 (0.030)	0.414 (0.015)	335 (14.8)	138 (2.44)
Fit	0.421 (0.013)	0.549 (0.017)	436 (22.7)	164 (2.79)
Lime	0.932 (0.017)	0.438 (0.008)	347 (8.46)	143 (1.47)
Occlusion	0.866 (0.032)	0.393 (0.006)	322 (14.6)	137 (1.90)
Aug Occlusion	0.755 (0.043)	0.388 (0.025)	364 (9.02)	165 (1.42)
Retain	0.645 (0.088)	0.334 (0.013)	206 (21.2)	138 (5.85)
Ours	0.885 (0.030)	0.781 (0.013)	1536 (79.0)	34.1 (3,70)

► In-Hospital Mortality

Method	Acc \downarrow	Comp \uparrow	CE \uparrow	Suff \downarrow
DeepLift	0.972 (0.003)	-1.19E-3 (0.007)	0.125 (0.014)	-6.92E-3 (0.006)
DynaMask	0.975 (0.002)	-1.27E-3 (0.004)	0.106 (0.009)	6.57E-3 (0.012)
IG	0.972 (0.003)	1.24E-4 (0.007)	0.127 (0.015)	-7.61E-3 (0.006)
GradientShap	0.968 (0.006)	-6.28E-3 (0.004)	0.128 (0.017)	6.61E-4 (0.005)
Lime	0.983 (0.003)	-5.22E-3 (0.004)	0.093 (0.008)	-2.23E-3 (0.019)
Occlusion	0.971 (0.003)	-4.03E-3 (0.003)	0.122 (0.008)	-4.97E-3 (0.008)
Aug Occlusion	0.972 (0.003)	-6.88E-4 (0.004)	0.121 (0.009)	-4.62E-3 (0.011)
Retain	0.971 (0.003)	-8.01E-3 (0.006)	0.0123 (0.009)	4.90E-4 (0.007)
Ours	0.943 (0.008)	1.09E-1 (0.023)	0.318 (0.057)	-6.94E-2 (0.006)

Learning Perturbations to Explain Time Series Predictions



Any Questions? Please use the Chat !