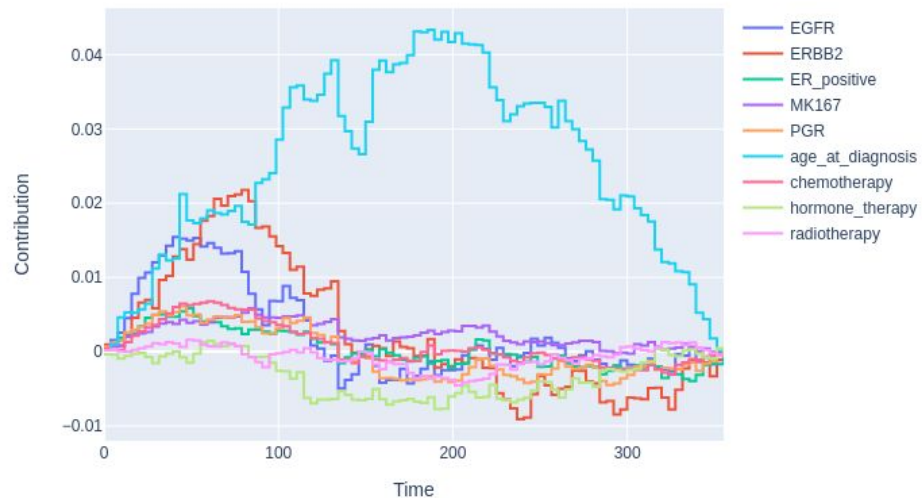


Time-dependent explanations of neural networks for survival analysis

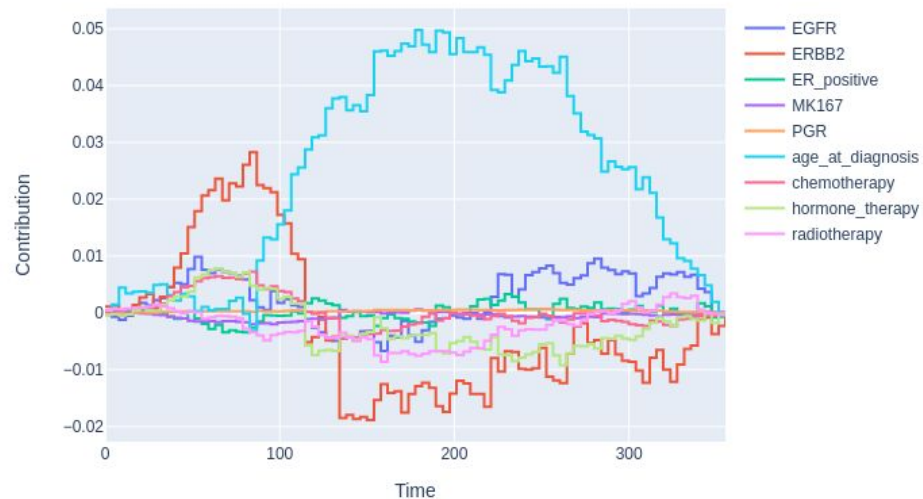
Jakub Bednarz
Kamil Grudzień
Krystian Sztenderski

Standardized explanations

SurvSHAP



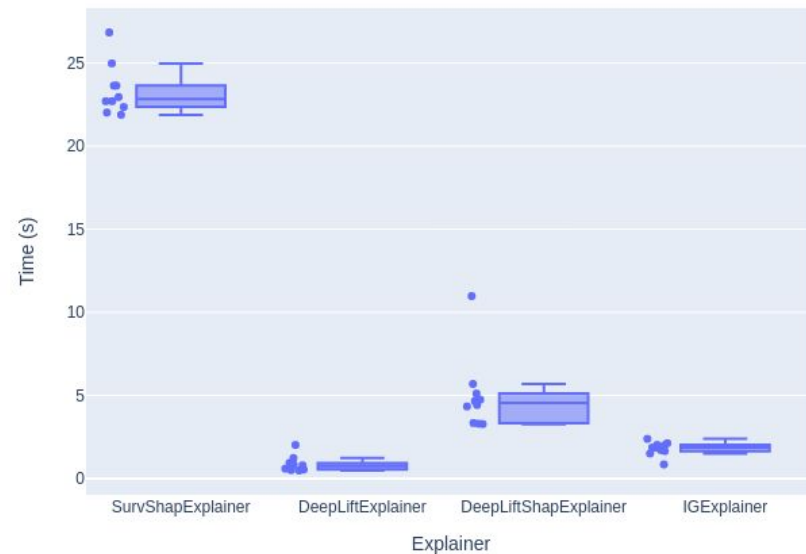
DeepLift



Execution time

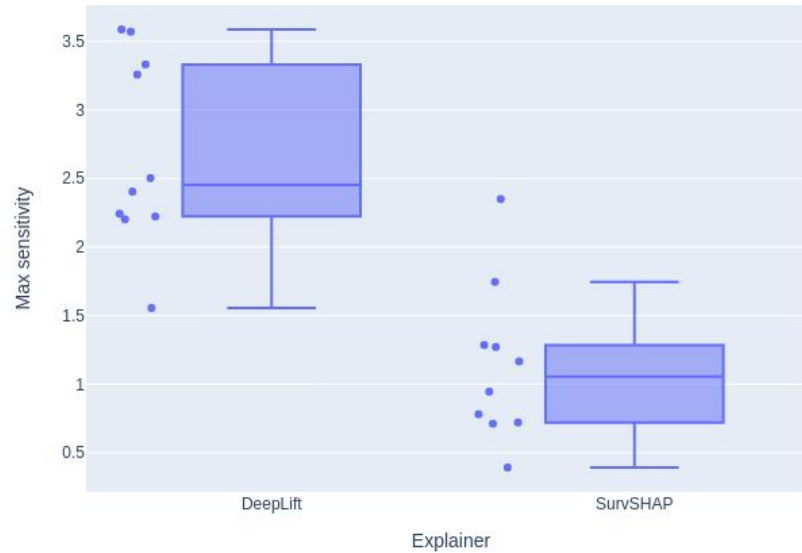
Method	Model	avg time (s)
SurvSHAP	Random Survival Forest	185.59
SurvSHAP	Cox Proportional Hazard	82.39
SurvSHAP	DeepHit	22.79
DeepLiftShap	DeepHit	2.91
Integrated Gradients	DeepHit	0.85
DeepLift	DeepHit	0.47

Execution time of explainer methods on a single sample for Metabric dataset



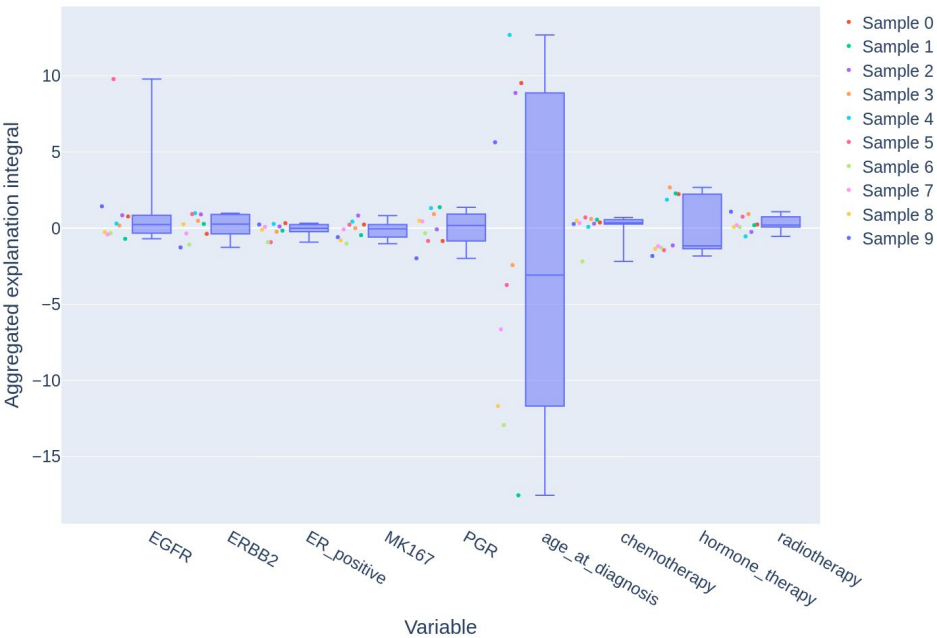
Max Sensitivity

Max sensitivity of explainers on Metabric dataset for 5 samples

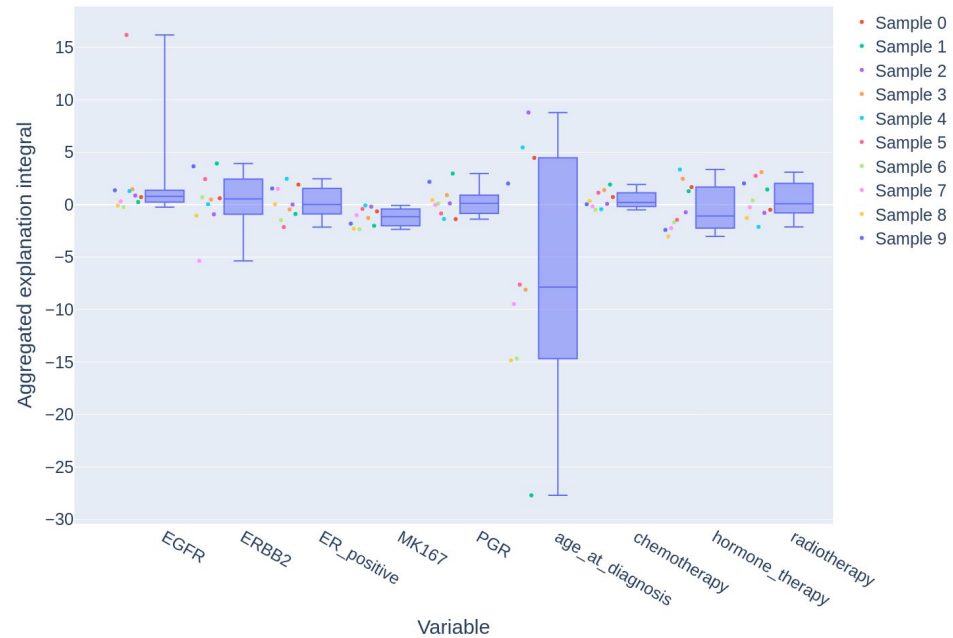


Integral aggregation

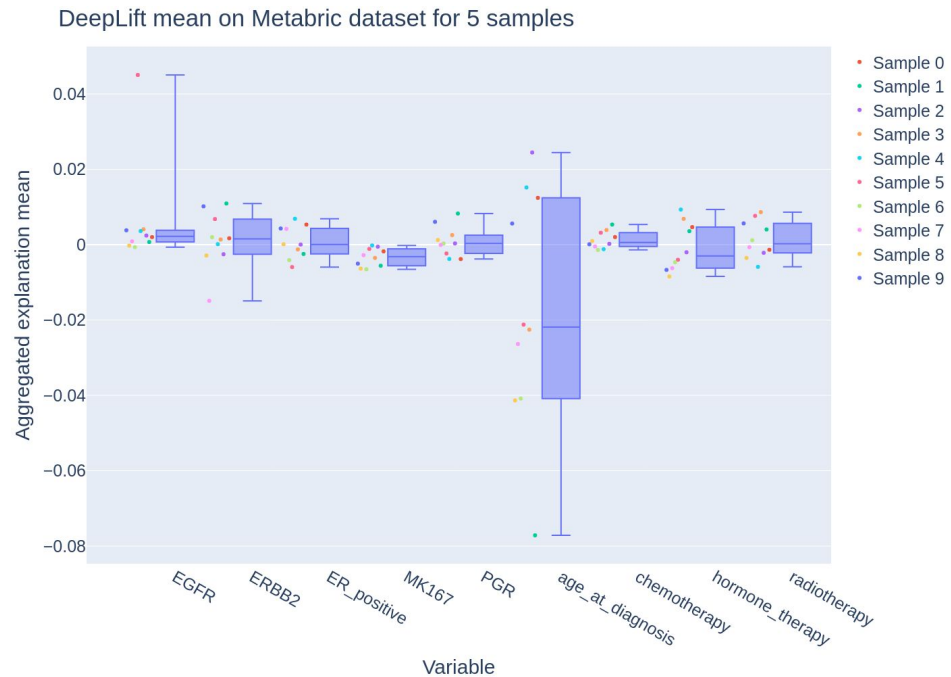
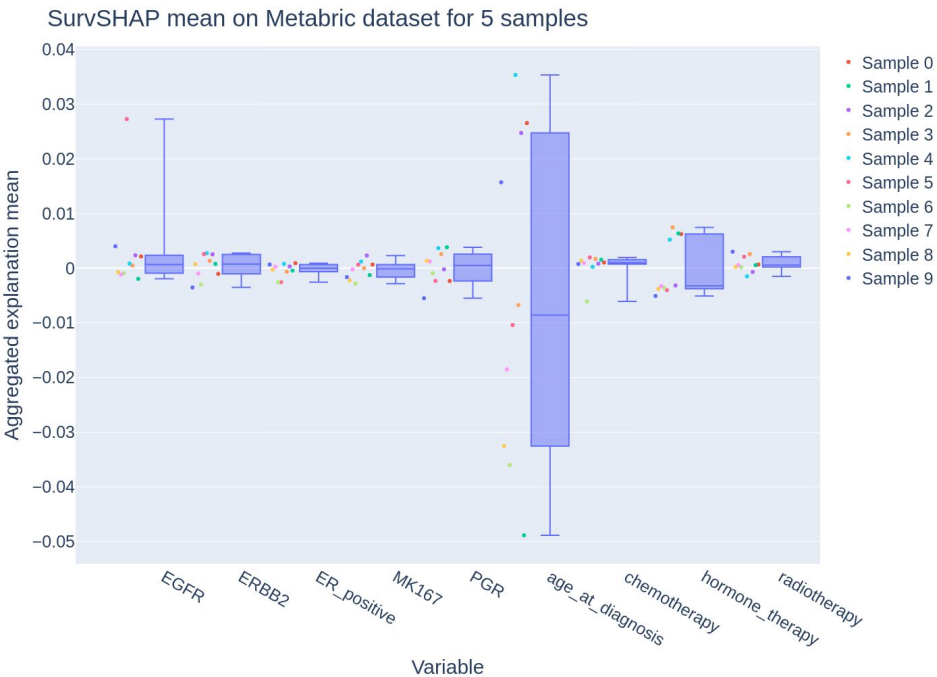
SurvSHAP integral on Metabric dataset for 5 samples



DeepLift integral on Metabric dataset for 5 samples

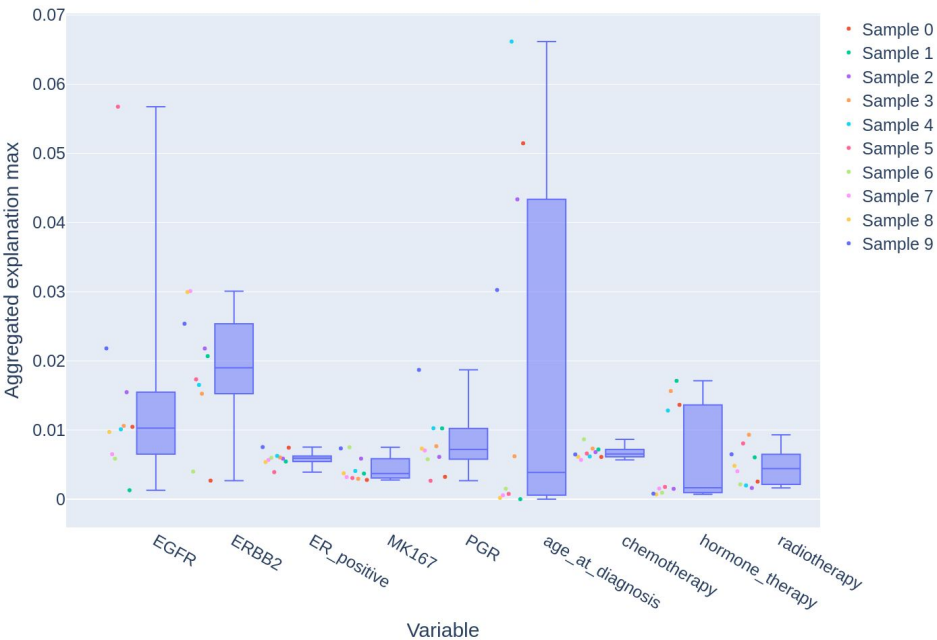


Mean aggregation

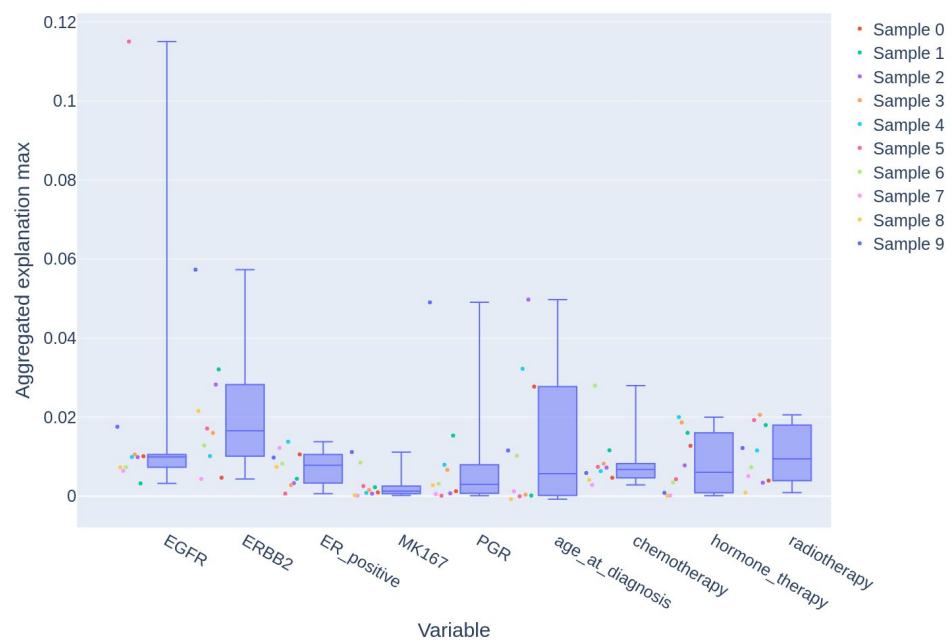


Max aggregation

SurvSHAP max on Metabric dataset for 5 samples

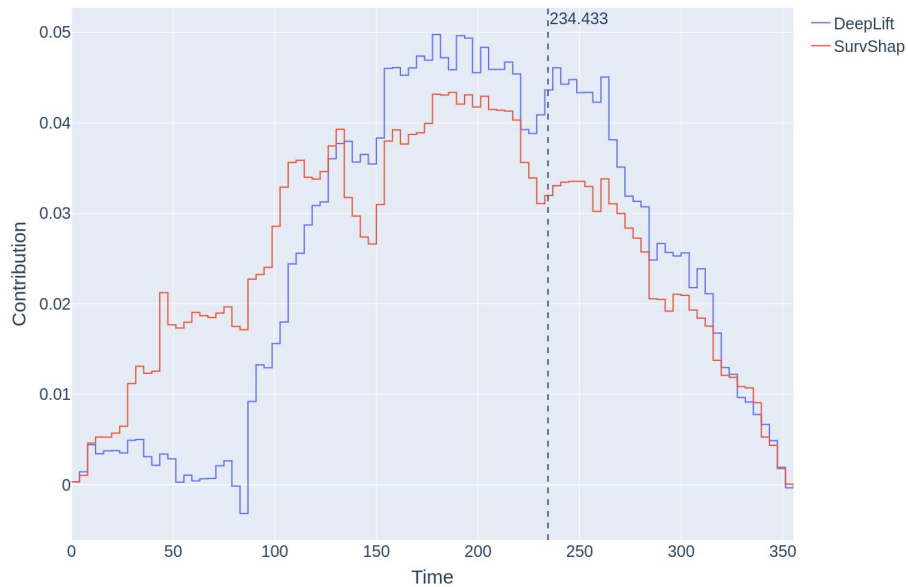


DeepLift max on Metabric dataset for 5 samples

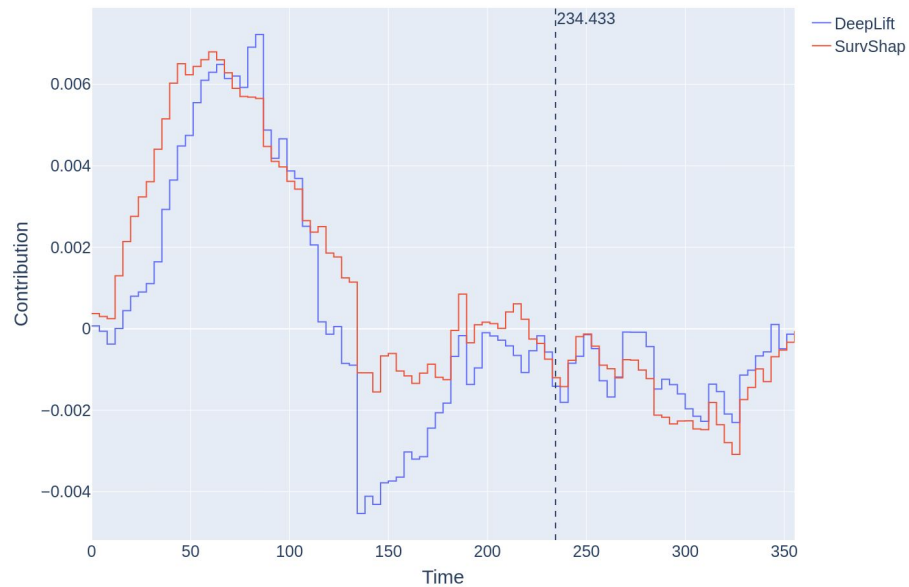


Explanation per feature

age_at_diagnosis

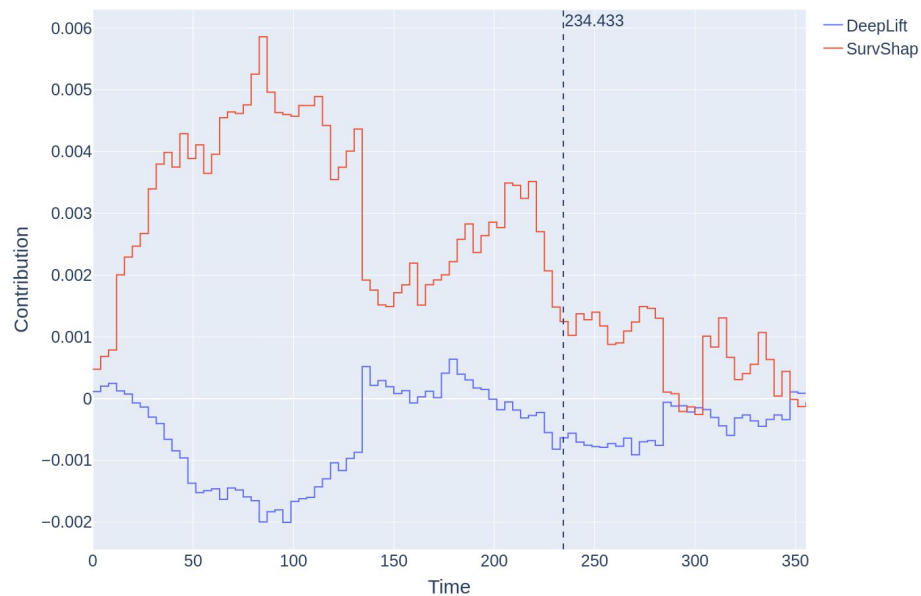


chemotherapy

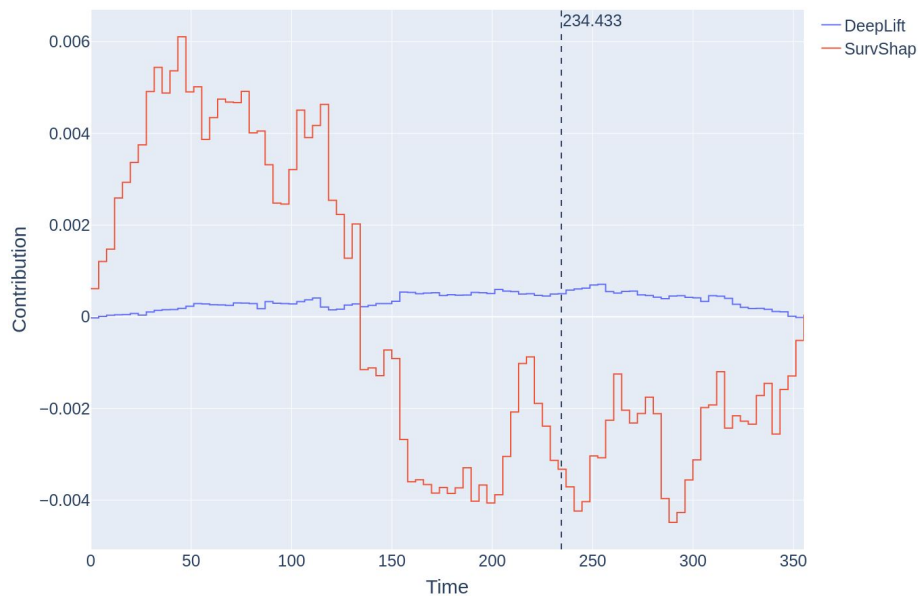


Explanation per feature

MK167



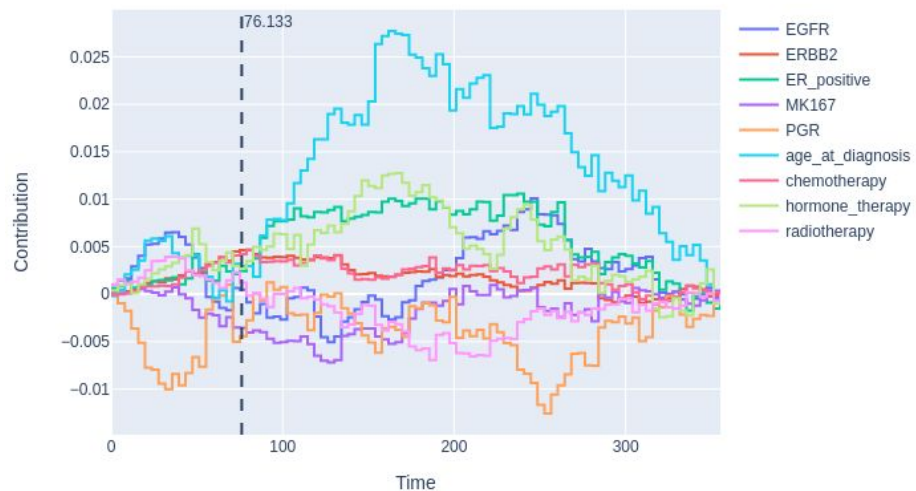
PGR



Baselines

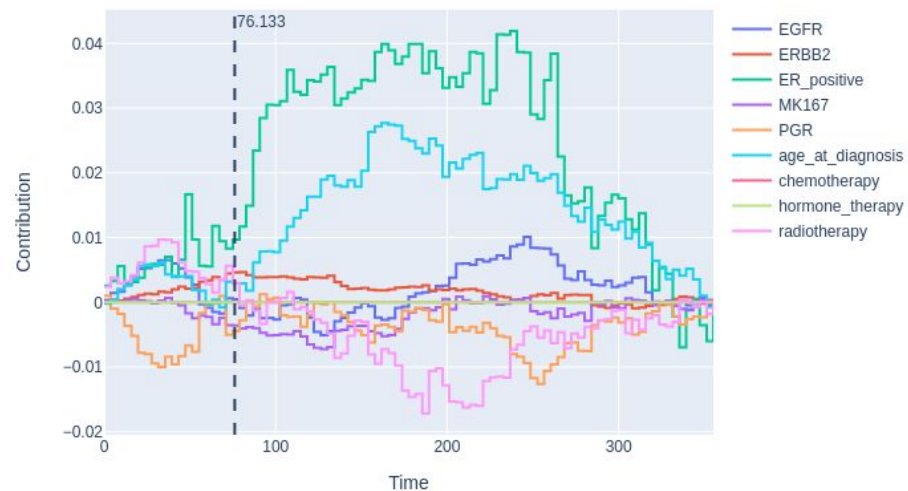
Mean baseline

DeepLiftExplainer DeepHitSingle



Zero baseline

DeepLiftExplainer DeepHitSingle



Takeaways

- Model specific explanations can leverage its internal architecture resulting in improved performance
- Explanation methods for single output models can successfully be adapted to survival analysis
- DeepLift achieves **similar results** to SurvSHAP
- DeepLift can be over **48 times faster** on average than SurvSHAP

Thanks for your attention