

naive_resp

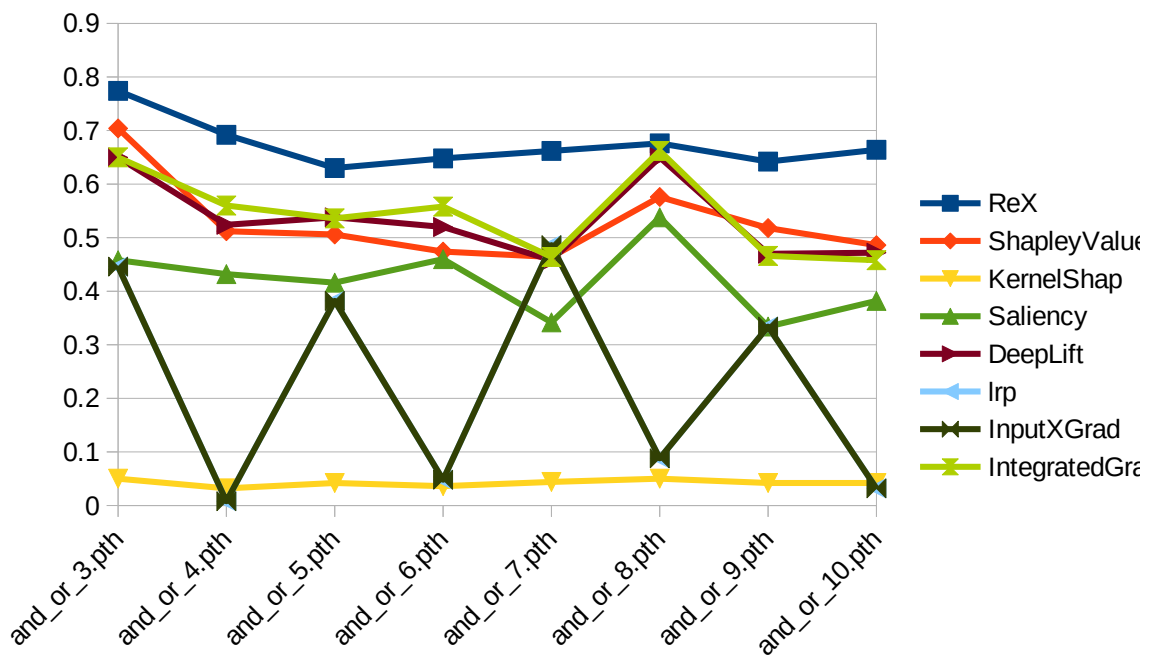
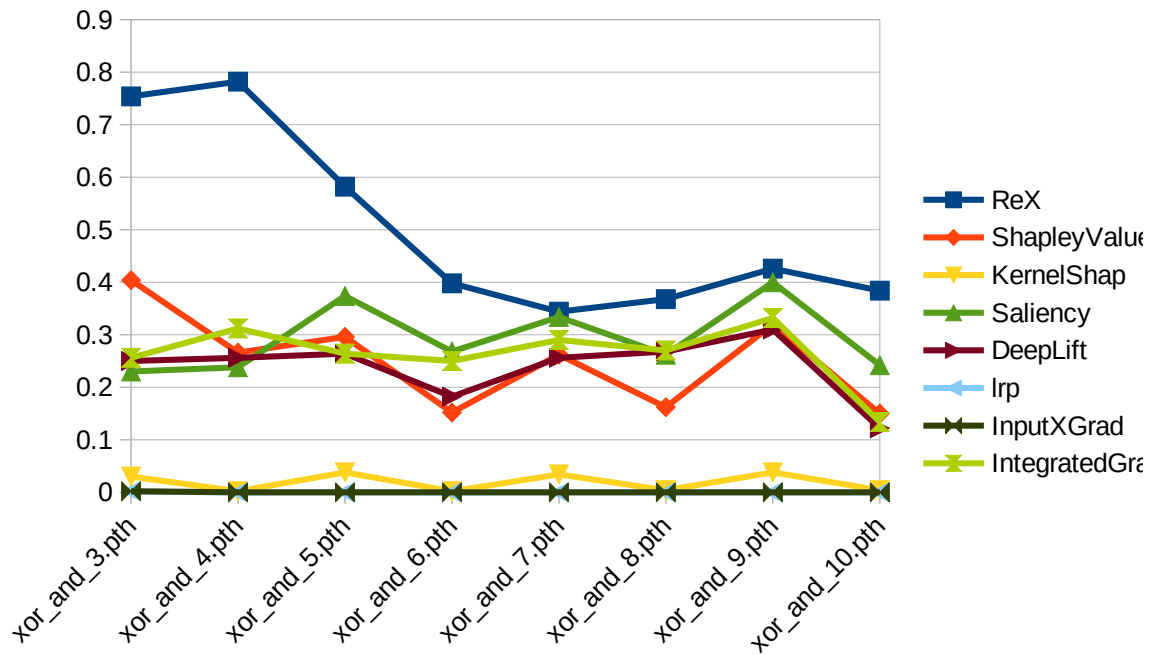
models	ReX	ShapleyValues	KernelShap	Saliency	DeepLift	Irp	InputXGrad
and.pth	0.992	0.88	0.068	0.842	0.958	0.136	0.136
or.pth	0.986	0.878	0.054	0.742	0.806	0.642	0.642
xor.pth	1	0.692	0.024	0.684	0.826	0.198	0.198
xor_and_xor.pth	0.638	0.536	0.006	0.774	0.734	0.314	0.314
xor_and_3.pth	0.754	0.404	0.03	0.23	0.25	0.002	0.002
xor_and_4.pth	0.782	0.266	0.002	0.238	0.256	0	0
xor_and_5.pth	0.582	0.296	0.038	0.374	0.264	0	0
xor_and_6.pth	0.398	0.152	0.002	0.268	0.182	0	0
xor_and_7.pth	0.344	0.262	0.034	0.334	0.256	0	0
xor_and_8.pth	0.368	0.162	0.004	0.262	0.268	0	0
xor_and_9.pth	0.426	0.32	0.038	0.4	0.31	0	0
xor_and_10.pth	0.384	0.15	0.004	0.242	0.122	0	0
and_or_3.pth	0.774	0.704	0.05	0.458	0.65	0.446	0.446
and_or_4.pth	0.692	0.512	0.032	0.432	0.524	0.008	0.008
and_or_5.pth	0.63	0.506	0.042	0.416	0.538	0.382	0.382
and_or_6.pth	0.648	0.474	0.036	0.46	0.52	0.048	0.048
and_or_7.pth	0.662	0.464	0.044	0.342	0.458	0.486	0.486
and_or_8.pth	0.676	0.576	0.05	0.538	0.65	0.088	0.088
and_or_9.pth	0.642	0.518	0.042	0.334	0.47	0.334	0.334
and_or_10.pth	0.664	0.486	0.042	0.382	0.472	0.032	0.032

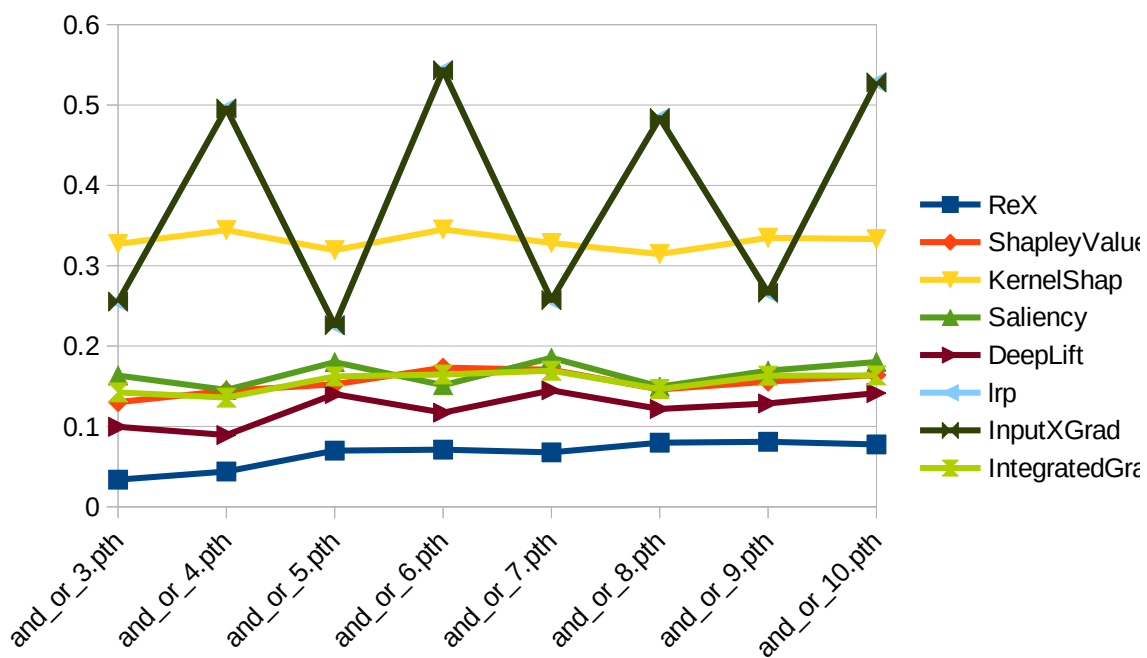
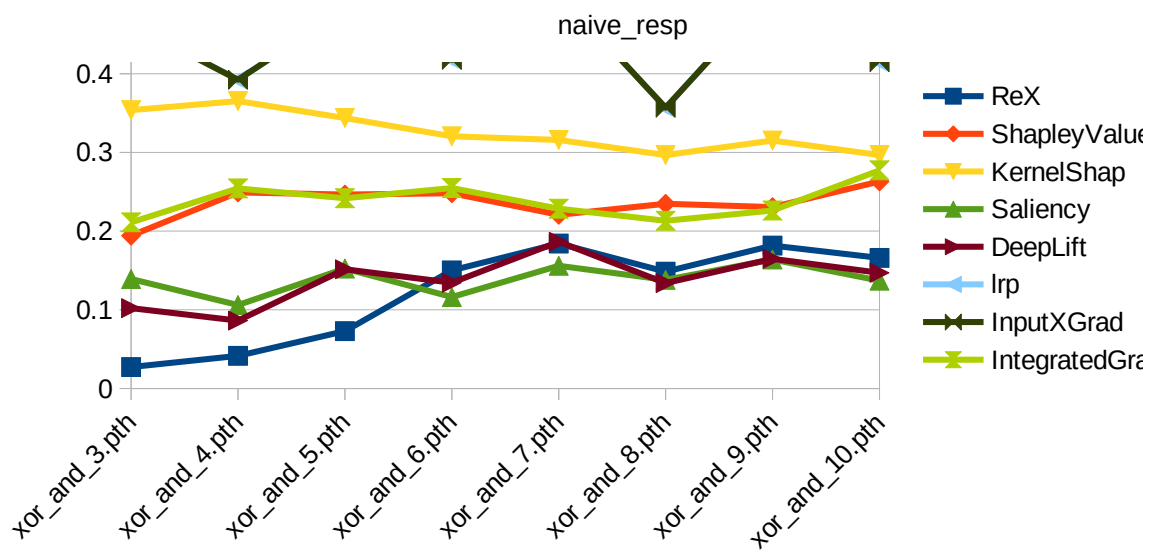
models	ReX	ShapleyValues	KernelShap	Saliency	DeepLift	Irp	InputXGrad
xor_and_3.pth	0.754	0.404	0.03	0.23	0.25	0.002	0.002
xor_and_4.pth	0.782	0.266	0.002	0.238	0.256	0	0
xor_and_5.pth	0.582	0.296	0.038	0.374	0.264	0	0
xor_and_6.pth	0.398	0.152	0.002	0.268	0.182	0	0
xor_and_7.pth	0.344	0.262	0.034	0.334	0.256	0	0
xor_and_8.pth	0.368	0.162	0.004	0.262	0.268	0	0
xor_and_9.pth	0.426	0.32	0.038	0.4	0.31	0	0
xor_and_10.pth	0.384	0.15	0.004	0.242	0.122	0	0

models	ReX	ShapleyValues	KernelShap	Saliency	DeepLift	Irp	InputXGrad
and.pth	0.992	0.88	0.068	0.842	0.958	0.136	0.136
or.pth	0.986	0.878	0.054	0.742	0.806	0.642	0.642
xor.pth	1	0.692	0.024	0.684	0.826	0.198	0.198
xor_and_xor.pth	0.638	0.536	0.006	0.774	0.734	0.314	0.314

models	ReX	ShapleyValues	KernelShap	Saliency	DeepLift	Irp	InputXGrad
and_or_3.pth	0.774	0.704	0.05	0.458	0.65	0.446	0.446
and_or_4.pth	0.692	0.512	0.032	0.432	0.524	0.008	0.008
and_or_5.pth	0.63	0.506	0.042	0.416	0.538	0.382	0.382
and_or_6.pth	0.648	0.474	0.036	0.46	0.52	0.048	0.048
and_or_7.pth	0.662	0.464	0.044	0.342	0.458	0.486	0.486
and_or_8.pth	0.676	0.576	0.05	0.538	0.65	0.088	0.088
and_or_9.pth	0.642	0.518	0.042	0.334	0.47	0.334	0.334
and_or_10.pth	0.664	0.486	0.042	0.382	0.472	0.032	0.032

naive_resp





naive_resp

IntegratedGrad	models	ReX	ShapleyValue	KernelShap	Saliency	
0.972	and.pth		0	0.08779535	0.351333063	0.103350216
0.79	or.pth		0	0.097243006	0.345345289	0.110038578
0.65	xor.pth		0	0.193114881	0.379917769	0.134277587
0.632	xor_and_xor.p	0.076010103	0.224653368	0.378633374	0.098156689	
0.256	xor_and_3.pth	0.02725465	0.194188537	0.353784172	0.139110004	
0.312	xor_and_4.pth	0.041509874	0.249231497	0.365252076	0.106025149	
0.264	xor_and_5.pth	0.072848822	0.24647888	0.343759386	0.15231993	
0.25	xor_and_6.pth	0.150044674	0.248163134	0.320507418	0.116236254	
0.29	xor_and_7.pth	0.184140848	0.220808695	0.315869117	0.155919072	
0.27	xor_and_8.pth	0.148457293	0.234789695	0.296722537	0.138039833	
0.332	xor_and_9.pth	0.181600093	0.230600154	0.315169067	0.163876201	
0.134	xor_and_10.pt	0.165938053	0.263131704	0.296760574	0.136953555	
0.65	and_or_3.pth	0.033728258	0.130193959	0.327255109	0.163356077	
0.56	and_or_4.pth	0.043895394	0.144309173	0.344373369	0.145458238	
0.536	and_or_5.pth	0.069893151	0.152362638	0.319394571	0.179962463	
0.558	and_or_6.pth	0.071056513	0.173137263	0.345240171	0.151248867	
0.464	and_or_7.pth	0.06773927	0.170568379	0.328410243	0.185426921	
0.662	and_or_8.pth	0.079789955	0.146006255	0.314529777	0.149488007	
0.466	and_or_9.pth	0.080885818	0.155661727	0.334800411	0.169343881	
0.458	and_or_10.pt	0.077632002	0.164000688	0.333183095	0.180153397	

IntegratedGrad	models	ReX	ShapleyValue	KernelShap	Saliency
0.256	xor_and_3.pth	0.02725465	0.194188537	0.353784172	0.139110004
0.312	xor_and_4.pth	0.041509874	0.249231497	0.365252076	0.106025149
0.264	xor_and_5.pth	0.072848822	0.24647888	0.343759386	0.15231993
0.25	xor_and_6.pth	0.150044674	0.248163134	0.320507418	0.116236254
0.29	xor_and_7.pth	0.184140848	0.220808695	0.315869117	0.155919072
0.27	xor_and_8.pth	0.148457293	0.234789695	0.296722537	0.138039833
0.332	xor_and_9.pth	0.181600093	0.230600154	0.315169067	0.163876201
0.134	xor_and_10.pt	0.165938053	0.263131704	0.296760574	0.136953555

IntegratedGrad	models	ReX	ShapleyValue	KernelShap	Saliency	
0.972	and.pth		0	0.08779535	0.351333063	0.103350216
0.79	or.pth		0	0.097243006	0.345345289	0.110038578
0.65	xor.pth		0	0.193114881	0.379917769	0.134277587
0.632	xor and xor.p	0.076010103	0.224653368	0.378633374	0.098156689	

IntegratedGrad	models	ReX	ShapleyValue	KernelShap	Saliency
0.65	and_or_3.pth	0.033728258	0.130193959	0.327255109	0.163356077
0.56	and_or_4.pth	0.043895394	0.144309173	0.344373369	0.145458238
0.536	and_or_5.pth	0.069893151	0.152362638	0.319394571	0.179962463
0.558	and_or_6.pth	0.071056513	0.173137263	0.345240171	0.151248867
0.464	and_or_7.pth	0.06773927	0.170568379	0.328410243	0.185426921
0.662	and_or_8.pth	0.079789955	0.146006255	0.314529777	0.149488007
0.466	and_or_9.pth	0.080885818	0.155661727	0.334800411	0.169343881
0.458	and_or_10.pt	0.077632002	0.164000688	0.333183095	0.180153397

naive_resp

es

ad

es

ad

naive_resp

es

ad

es

ad

DeepLift	lrp	InputXGrad	IntegratedGrad
0.065304782	0.546292308	0.546292318	0.08626769
0.06966697	0.165790678	0.165790672	0.108091832
0.060290205	0.304546289	0.304546289	0.18244577
0.082065052	0.349128203	0.349128193	0.166662404
0.102062563	0.462922112	0.462922111	0.211179122
0.086250404	0.392420988	0.392420982	0.254301249
0.151475887	0.481899485	0.481899487	0.241861658
0.134380469	0.4177682	0.417768204	0.254788221
0.186475479	0.503045582	0.503045585	0.228517893
0.133906814	0.356810358	0.356810357	0.213214525
0.164722698	0.50521786	0.505217859	0.226232602
0.146976679	0.41484253	0.414842532	0.277253611
0.099529414	0.255316309	0.255316311	0.142309814
0.089295945	0.495799587	0.495799581	0.135882388
0.140178767	0.225953518	0.225953516	0.162164457
0.117153644	0.543364465	0.543364451	0.164647861
0.144913134	0.257365154	0.257365159	0.169334458
0.121708332	0.483843114	0.483843107	0.146447151
0.128622724	0.266428044	0.266428043	0.163091295
0.141455482	0.528217793	0.52821779	0.163370782

DeepLift	lrp	InputXGrad	IntegratedGrad
0.102062563	0.462922112	0.462922111	0.211179122
0.086250404	0.392420988	0.392420982	0.254301249
0.151475887	0.481899485	0.481899487	0.241861658
0.134380469	0.4177682	0.417768204	0.254788221
0.186475479	0.503045582	0.503045585	0.228517893
0.133906814	0.356810358	0.356810357	0.213214525
0.164722698	0.50521786	0.505217859	0.226232602
0.146976679	0.41484253	0.414842532	0.277253611

DeepLift	lrp	InputXGrad	IntegratedGrad
0.065304782	0.546292308	0.546292318	0.08626769
0.06966697	0.165790678	0.165790672	0.108091832
0.060290205	0.304546289	0.304546289	0.18244577
0.082065052	0.349128203	0.349128193	0.166662404

DeepLift	lrp	InputXGrad	IntegratedGrad
0.099529414	0.255316309	0.255316311	0.142309814
0.089295945	0.495799587	0.495799581	0.135882388
0.140178767	0.225953518	0.225953516	0.162164457
0.117153644	0.543364465	0.543364451	0.164647861
0.144913134	0.257365154	0.257365159	0.169334458
0.121708332	0.483843114	0.483843107	0.146447151
0.128622724	0.266428044	0.266428043	0.163091295
0.141455482	0.528217793	0.52821779	0.163370782