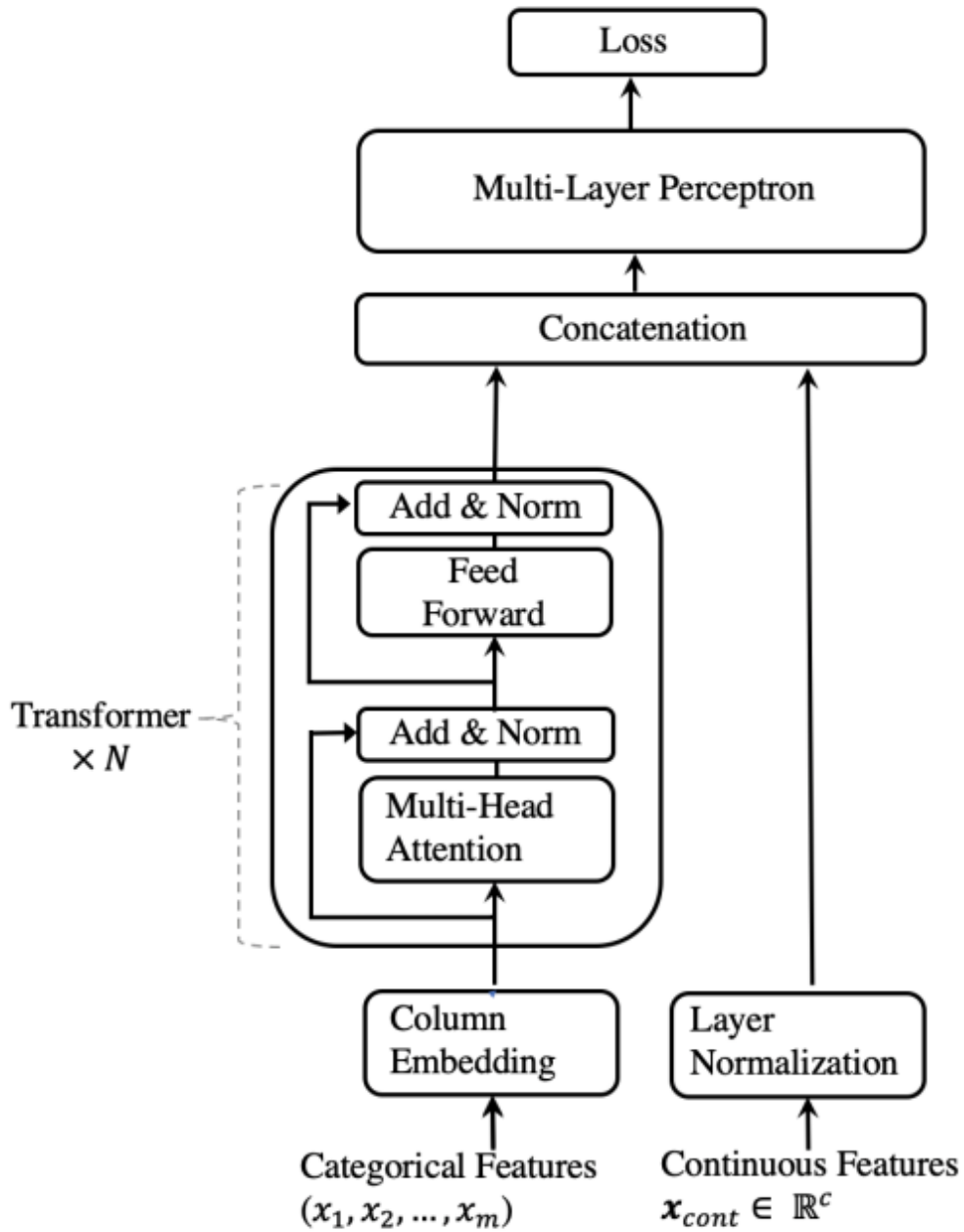
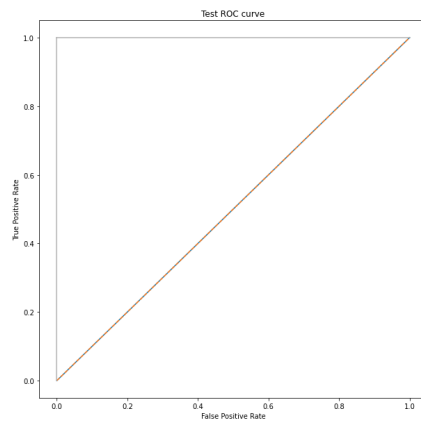
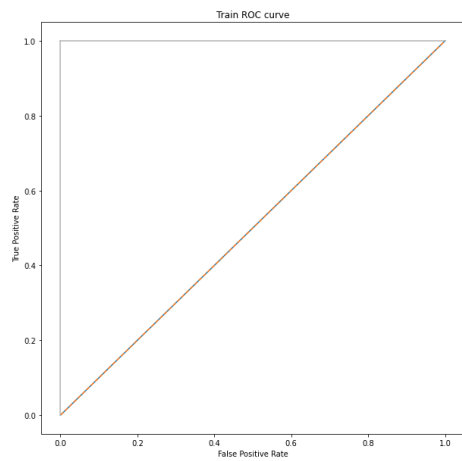


Architecture:

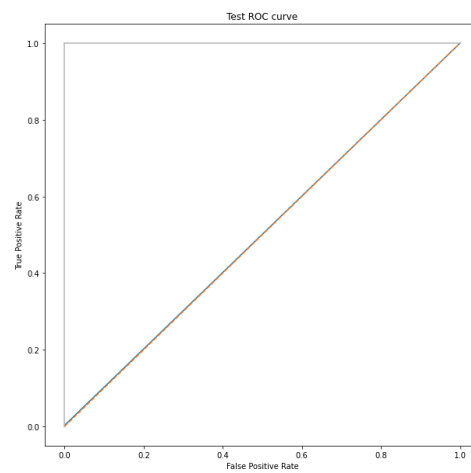
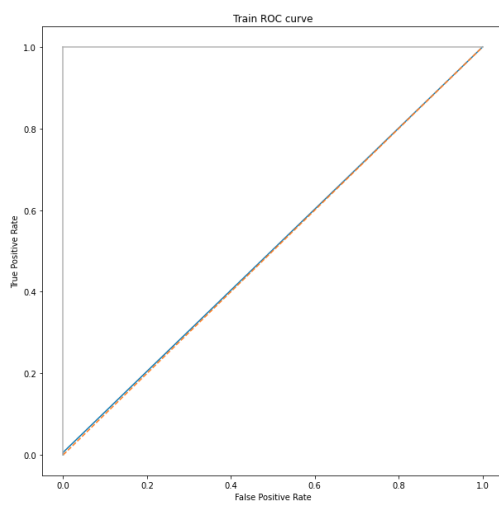


Datasets: <https://www.kaggle.com/shrutimechlearn/churn-modelling>

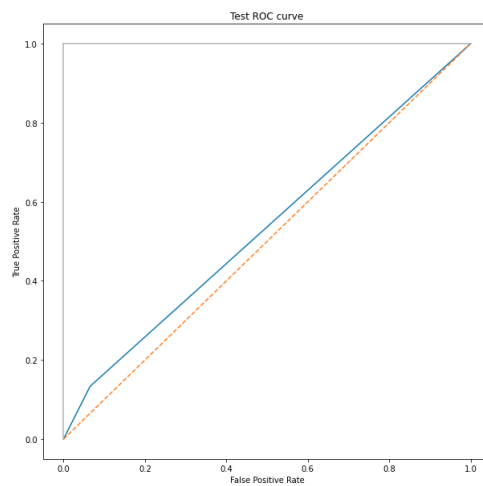
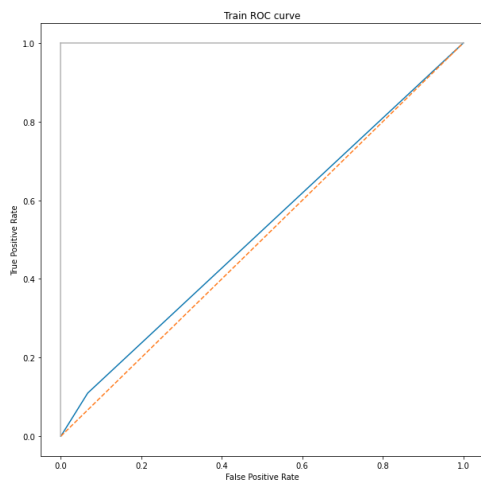
Heads = 1, depth = 6, attn_dropout=0.1, ff_dropout=0.1



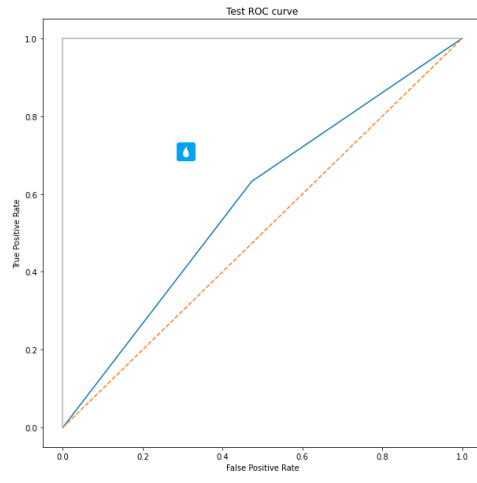
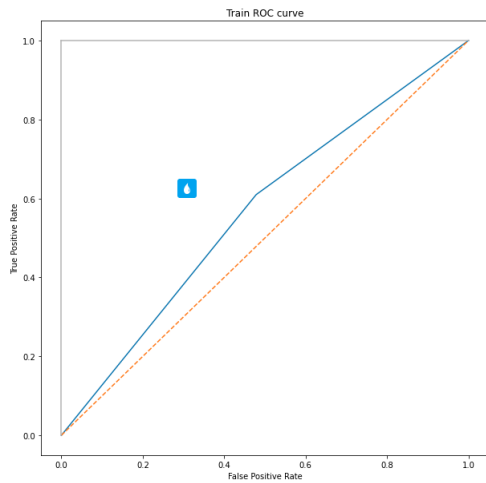
Heads = 2, depth = 6, attn_dropout=0.1, ff_dropout=0.1



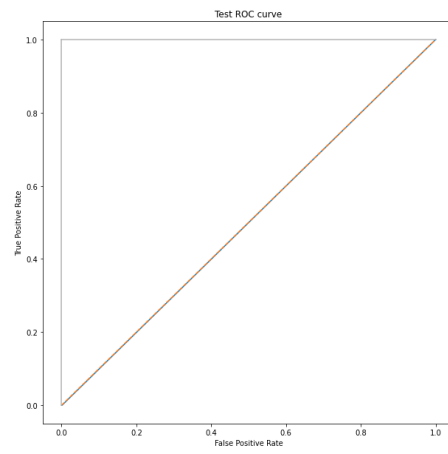
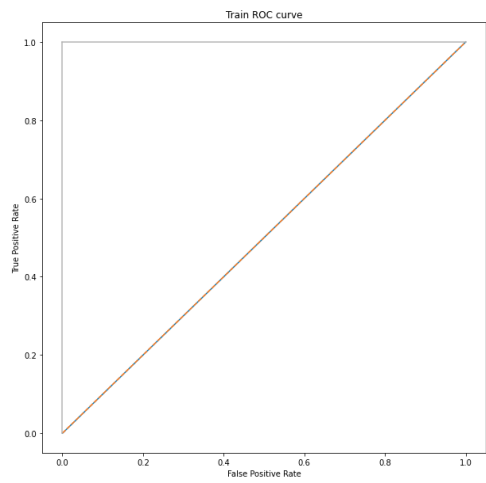
Heads = 4, depth = 6, attn_dropout=0.1, ff_dropout=0.1



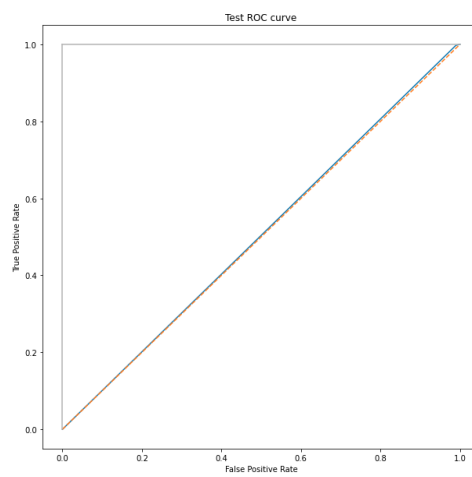
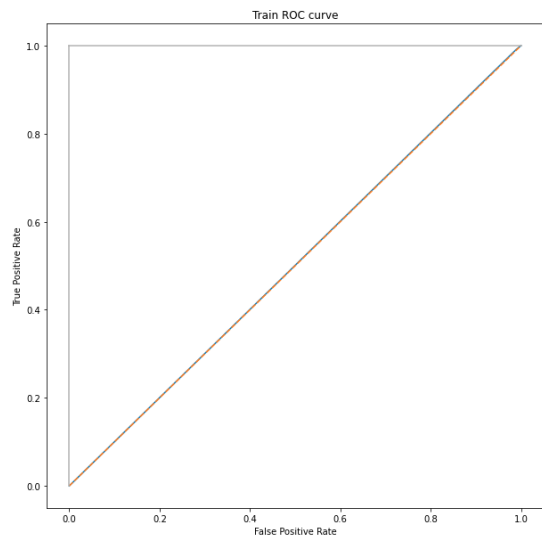
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1



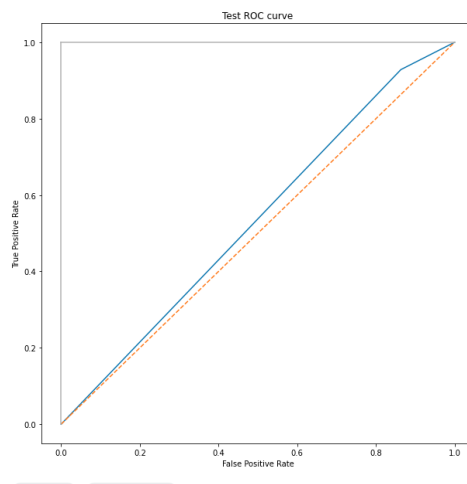
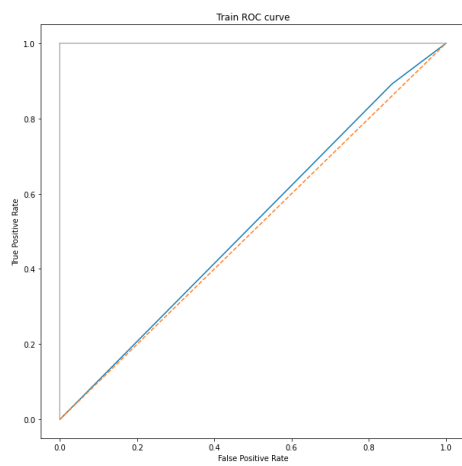
Heads = 16, depth = 6, attn_dropout=0.1, ff_dropout=0.1



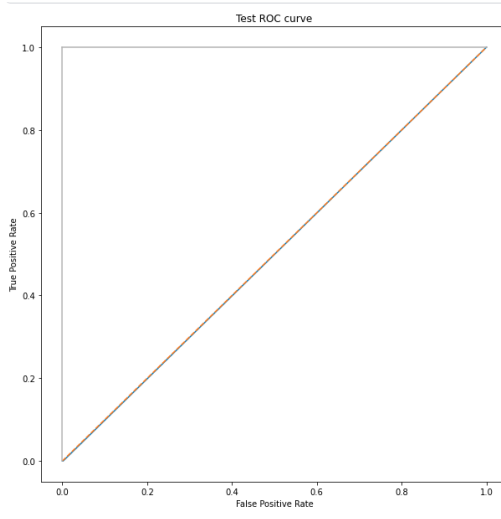
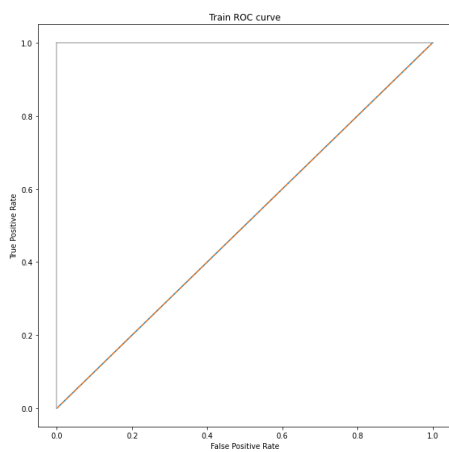
Heads = 32, depth = 6, attn_dropout=0.1, ff_dropout=0.1



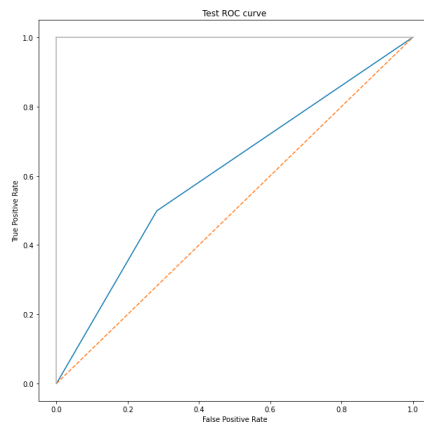
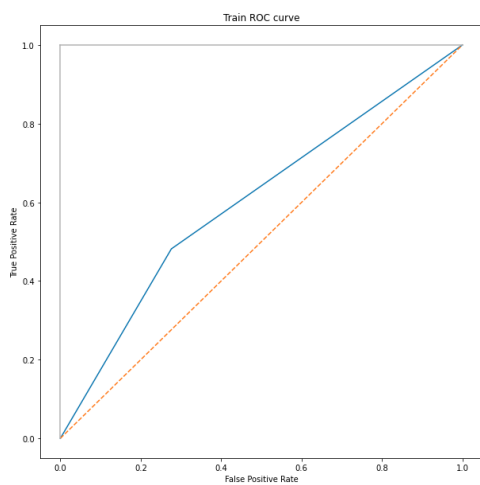
Heads = 8, depth = 6, attn_dropout=0.0, ff_dropout=0.



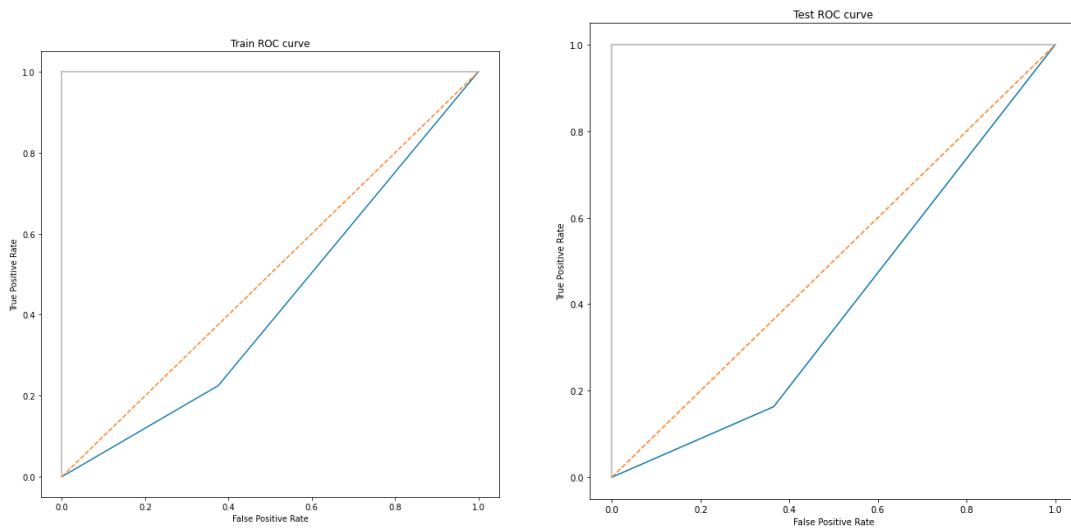
Heads = 8, depth = 6, attn_dropout=0.5, ff_dropout=0.5



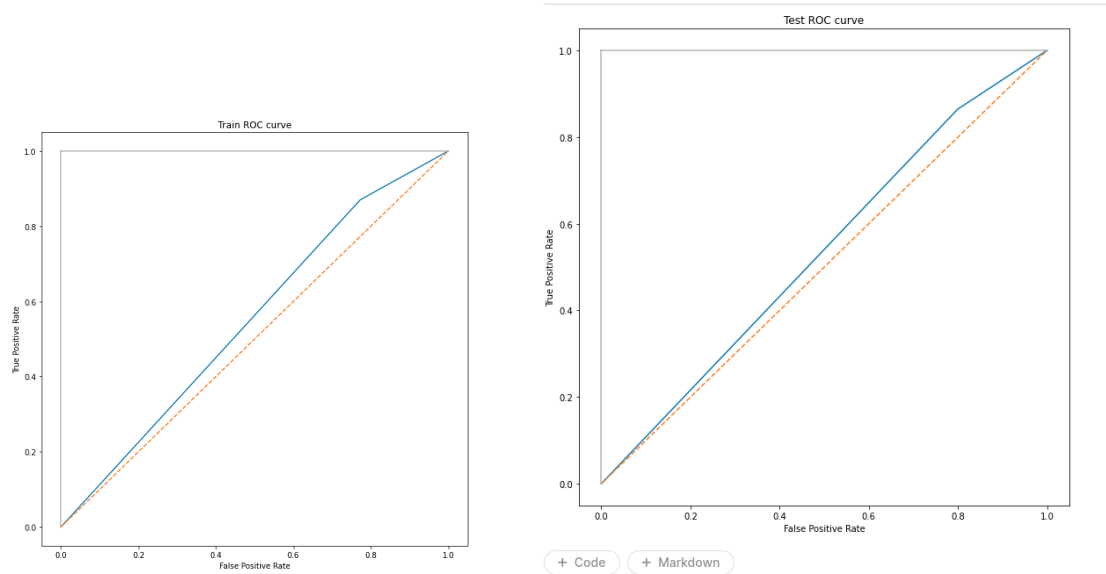
Heads = 8, depth = 2, attn_dropout=0.1, ff_dropout=0.1



Heads = 8, depth = 5, attn_dropout=0.1, ff_dropout=0.1

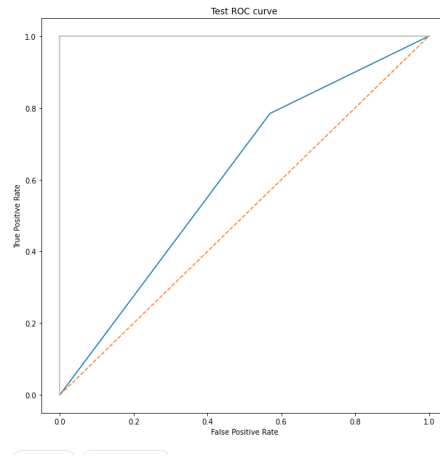
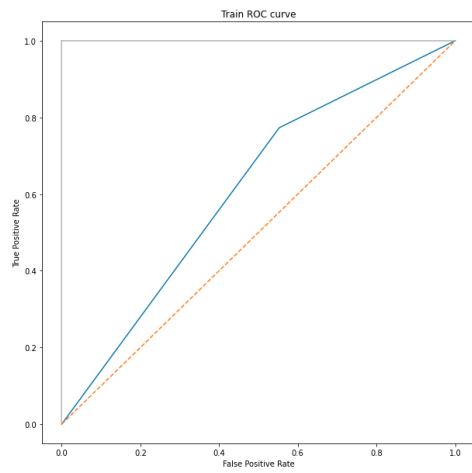


Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1,
mlp_hidden = (5,3)

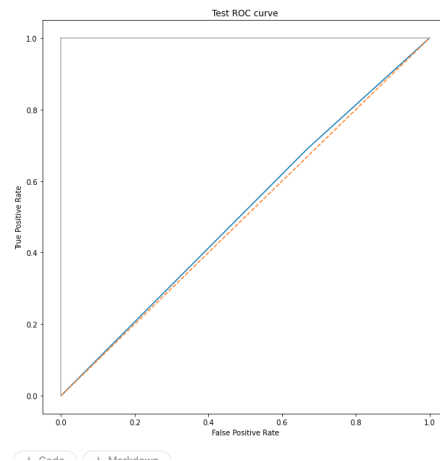
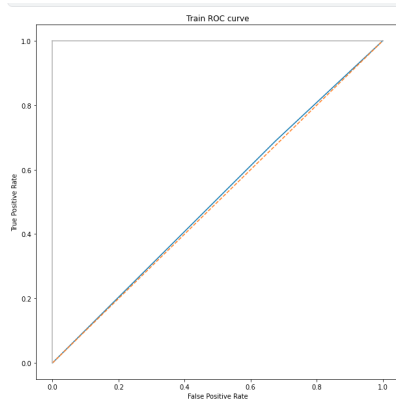


Dataset: <https://www.kaggle.com/lodetomasi1995/income-classification>

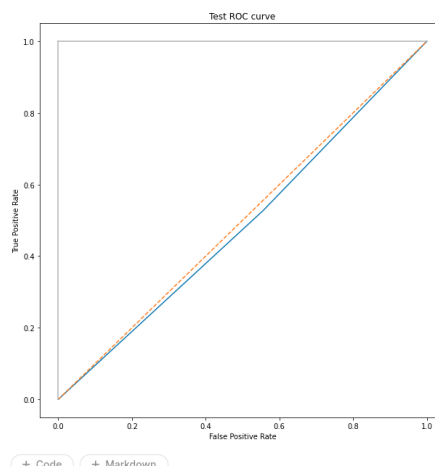
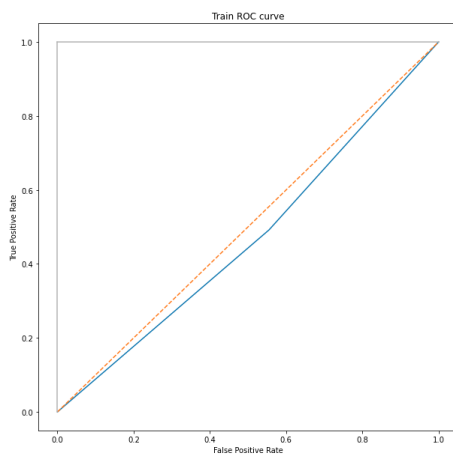
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1



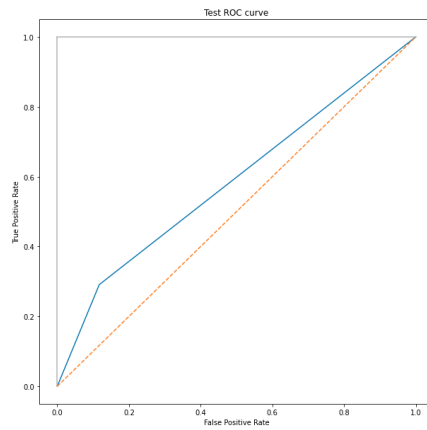
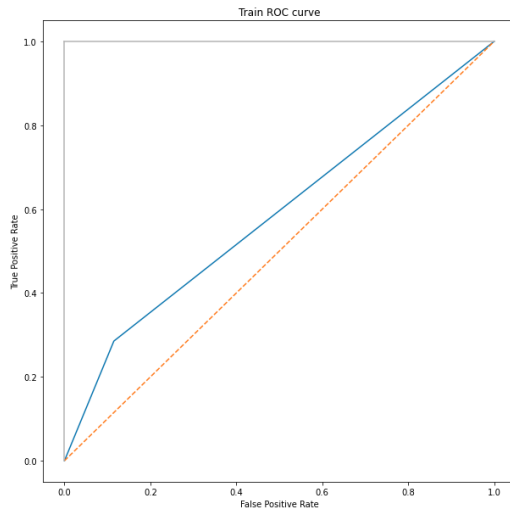
Heads = 1, depth = 6, attn_dropout=0.1, ff_dropout=0.1



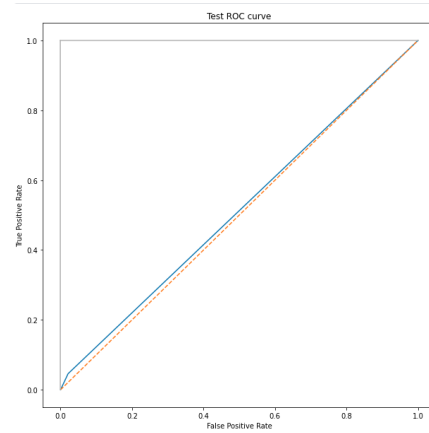
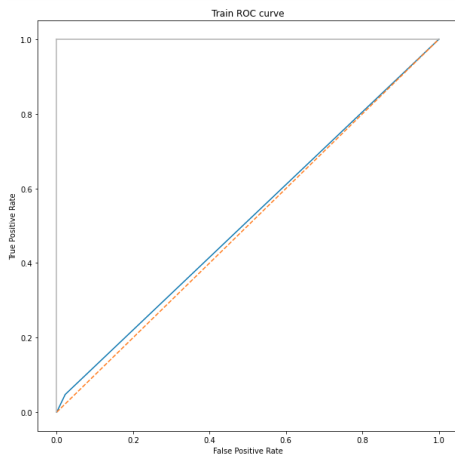
Heads = 2, depth = 6, attn_dropout=0.1, ff_dropout=0.1



Heads = 4, depth = 6, attn_dropout=0.1, ff_dropout=0.1



Heads = 8, depth = 6, attn_dropout=0.0, ff_dropout=0.



Churn_Modelling	Train_acc	Train_roc_auc	Test_acc	Test_roc_auc	
Heads = 1, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.796	0.5	0.7975	0.5	
Heads = 2, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.797	0.503	0.797	0.501	
Heads = 4, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.76525	0.52	0.7725	0.534	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.5395	0.57	0.5485	0.579	

Heads = 16, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.795375	0.4998352670 21382	0.7965	0.4993730407 52351	
Heads = 32, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.211	0.5009791605 08425	0.21	0.5037811060 79957	
Heads = 8, depth = 6, attn_dropout=0.0, ff_dropout=0.	0.2935	0.5158882648 5368	0.297	0.5325360888 57928	
Heads = 8, depth = 6, attn_dropout=0.5, ff_dropout=0.5	0.79525	0.4999846043 94522	0.796	0.4990595611 28527	
Heads = 8, depth = 6, attn_dropout=0.01, ff_dropout=0.01	0.796	0.5	0.7975	0.5	
Heads = 8, depth = 1, attn_dropout=0.1, ff_dropout=0.1	0.387125	0.5006450758 69544	0.386	0.4989898989 89899	
Heads = 8, depth = 2, attn_dropout=0.1, ff_dropout=0.1	0.6745	0.6027749039 31422	0.674	0.6086303649 52204	
Heads = 8, depth = 4, attn_dropout=0.1, ff_dropout=0.1	0.20675	0.4992209823 62794	0.2105	0.5022524091 48961	
Heads = 8, depth = 12, attn_dropout=0.1, ff_dropout=0.1	0.204	0.5	0.203	0.5003134796 23824	
Heads = 8, depth = 5, attn_dropout=0.1, ff_dropout=0.1	0.543	0.4247031727 26377	0.5395	0.3990363404 15651	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, mlp_hidden = (5,3)	0.3585	0.5487424869 44527	0.335	0.5324122450 55923	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, SELU	0.5085	0.4978199822 64262	0.5145	0.5067881883 97384	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, sigmoid	0.204	0.5	0.2025	0.5	

Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, Tanh	0.203625	0.4986251724 30781	0.2025	0.4981578234 4518	
Income_Evaluation					
Heads = 1, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.4119318181 81818	0.5068516449 25492	0.4179333640 41148	0.5107910802 99992	
Heads = 2, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.4561962530 71253	0.4683949146 50234	0.4644557039 76662	0.4856778439 12899	
Heads = 4, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.7412469287 46929	0.5851823542 98914	0.7379087977 89037	0.5868941550 51943	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1	0.5253762285 01228	0.608	0.5172731460 15661	0.6074944449 47292	
Heads = 8, depth = 6, attn_dropout=0.0, ff_dropout=0.	0.7547220515 97052	0.5126652272 48317	0.7504990019 96008	0.5124122760 60942	
Heads = 8, depth = 6, attn_dropout=0.5, ff_dropout=0.5	0.6391277641 27764	0.4853010704 14588	0.6328880700 13818	0.4864290986 60031	
Heads = 8, depth = 1, attn_dropout=0.1, ff_dropout=0.1	0.7478885135 13513	0.5049374575 18514	0.7432826654 38354	0.5057276003 57461	
Heads = 8, depth = 2, attn_dropout=0.1, ff_dropout=0.1	0.7291538697 7887	0.4966159616 77346	0.7228619683 7095	0.4960184645 74508	
Heads = 8, depth = 4, attn_dropout=0.1, ff_dropout=0.1	0.4601888820 63882	0.4256454703 67978	0.4647627821 28052	0.4369525420 66059	

Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, mlp_hidden = (5,3)	0.3566108722 35872	0.5403380353 1038	0.3592814371 25748	0.5333797157 3938	
Heads = 8, depth = 6, attn_dropout=0.1, ff_dropout=0.1, SELU	0.7154100122 85012	0.4984271198 08551	0.7124213112 23706	0.4980012008 85749	