

*Generalisability List:*  
Average Test AUCs For Each Architecture  
Where Trainset  $\neq$  Testset, i.e. Out of Training Dist.

