Generalisability List:
Average Test AUCs For Each Model
Where Trainset ≠ Testset, i.e. Out of Trianing Dist.