

(15th March, Week 3)

1 Comparing Models - (Which one is better?)

We can compare models with different blunt metrics like accuracy etc... we would like to rule out the possibility of tainted results by *chance*. We do this concretely by incorporating t-tests to provide a statistical measure against a null hypothesis.

So considering models M_1 and M_2 . That is two models. We perform a 10-fold cross validation, where respectively each iteration result populates accuracy samples A_{M_1} and A_{M_2} . Hence $|A_{M_1}| = |A_{M_2}| = 10$. We then perform an independant sample t-test (null hypothesis is that A_{M_1} differs not to A_{M_2}).

NOTE: we need to go back to the notes here and copy down all the manual formulas for calculating those t-test values etc, didnt bother since I would just use sci py package and just perform a significance test that way...

1.1 Alterantive measures

And keep in mind that **tree models are not always the best**. How does the model perform when certain features are not present? Does the model still predict well? Which features are critical etc... The no free lunch theorem.