3. Confidence Level for Favourite Algorithm Outperforming Decision Tree

We are comparing two algorithms using

10-fold cross-validation repeated 10 times (paired t-test is appropriate for this setup, but the table only shows the results of the 10 final folds). The comparison involves the mean difference in error rates.

U(^d) between the two algorithms over the k = 10 folds.

Totals:

di = E^(Fav) − E^(DT) = **- 0.0109**

(di −U^(d))^2 = **2.299 × 10^(-4)**

- Number of folds: k = 10.
- Mean difference: U^d = **Sum(**d(i) / k**)** = −0.0109 / 10 = −0.00109.
    - This negative mean difference suggests the Decision Tree (DT) algorithm has a lower average error rate (a higher accuracy) than the Favourite Algorithm (Fav), since d(i) = E^(Fav) − E^(D(T)).

The standard deviation of the differences is:

S(d) = **Root(**1 / (10 - 1) * 2.299 * 10^(-4)**)**

= **Root(**2.554 * 10^(-5)**)**

= 0.005054

The t-statistic for comparing U^d to 0 (where Ud is the true difference) is:

T = U^(d) - 0 / S(d) * **Root(**k**)**

= -0.00109 / (0.005054 / **Root(**10**)**)

= -0.00109 / 0.00160

= -0.681

We have k − 1 = 9 degrees of freedom. We are looking for the confidence level that the Favourite Algorithm will outperform the Decision Tree. Outperforming means E(Fav<E(DT, so the difference U(d = E(Fav)−E(DT) must be negative (U(d<0).

The t-statistic is t = −0.681. Since U^(d) is already negative, we need to find the probability of a t-score being this low or lower (one-tailed test). We look up the absolute value

|t| = 0.681 in the t-table (Table 4) for

N = 9 degrees of freedom.

The smallest value for v = 9 in the table is t = 0.985 at a 65% confidence level (two-sided).

- For a 65% two-sided confidence interval, the t-value is t65,9 = 0.985.
- This corresponds to a one-tailed probability of 21 − 0.65 = 0.175 in the right tail, and 1 − 0.175 = 0.825 in the left tail (confidence of being worse).

Since our t-statistic (0.985) is bigger than 0.825 but smaller than 1.383, we can be **at least 80% confident** (but not quite 90%) that your favorite algorithm will outperform the decision tree induction algorithm in this situation.