

Note: Doing means folds testing for 20 * 10-fold cross-validation in which the average accuracy for a single 10-fold cross validation acts as a sample													
Runs	Tuned Decision Trees	Default Decision Trees	Difference in accuracy	Sample mean	Mean squared Distance	Degrees of Freedom	SQRT DF + 1	Sum of mean squared Distance	Sample std deviation	Calculated Z-statistic	p-value	Z-value from Table (one tailed test)	
trial 1	94.2	93	1.2	-0.07	1.6129	19	4.472135955	9.502	0.7071812096	-0.4426722778	0.05	1.645	
trial 2	94.5	93.9	0.6		0.4489								
trial 3	94.1	94.5	-0.4		0.1089								
trial 4	93.6	94.2	-0.6		0.2809								
trial 5	94.7	94.2	0.5		0.3249								
trial 6	93.6	94.2	-0.6		0.2809								
trial 7	94.8	93.6	1.2		1.6129								
trial 8	94.5	94.5	0		0.0049								
trial 9	94.8	94.1	0.7		0.5929								
trial 10	93.6	93.9	-0.3		0.0529								
trial 11	93.3	94.2	-0.9		0.6889								
trial 12	93.6	94.5	-0.9		0.6889								
trial 13	93.6	94.5	-0.9		0.6889								
trial 14	94.2	94.7	-0.5		0.1849								
trial 15	93.9	93.4	0.5		0.3249								
trial 16	93.9	94.2	-0.3		0.0529								
trial 17	94.4	94.5	-0.1		0.0009								
trial 18	94.2	93.9	0.3		0.1369								
trial 19	93.6	94.8	-1.2		1.2769								
trial 20	94.2	93.9	0.3		0.1369								

Since, $-0.4426722778 < 1.645$, we cannot reject the null hypotheses that the tuned decision tree model will not perform better than the default decision tree model when using features extracted from Link-Target Identification