MEE 3033: KEPINTARAN BUATAN / *ARTIFICIAL INTELLIGENCES*

SESI PENGAJIAN: SEM 2 (2021/2022)

KUMPULAN KULIAH MEE 3033 (A211): A

(*GROUP ASSINGMENT - CLASSIFICATION*)

RAPIDMINER

Should we play Golf?

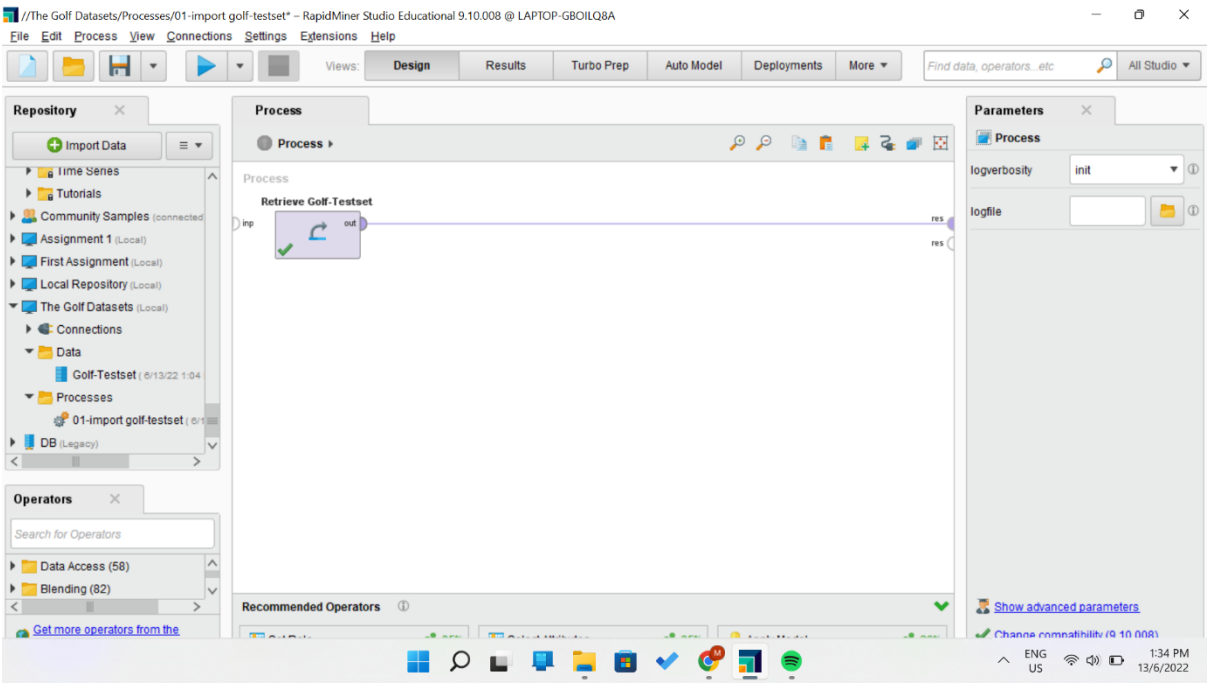
NAMA	NO MATRIC	NO. HP	PROGRAM
MOHD IZZUL IKHWAN BIN MOHD YUSOF	D20201095609	0197818481	AT20 - IJAZAH SARJANA MUDA PENDIDIKAN (TEKNOLOGI MAKLUMAT)
JAYATHEESH A/L ANPUSELVAM	D20201093973	0138200170	AT20 - IJAZAH SARJANA MUDA PENDIDIKAN (TEKNOLOGI MAKLUMAT)

NAMA PENSYARAH: PROFESOR MADYA DR. BAHBIBI BINTI RAHMATULLAH

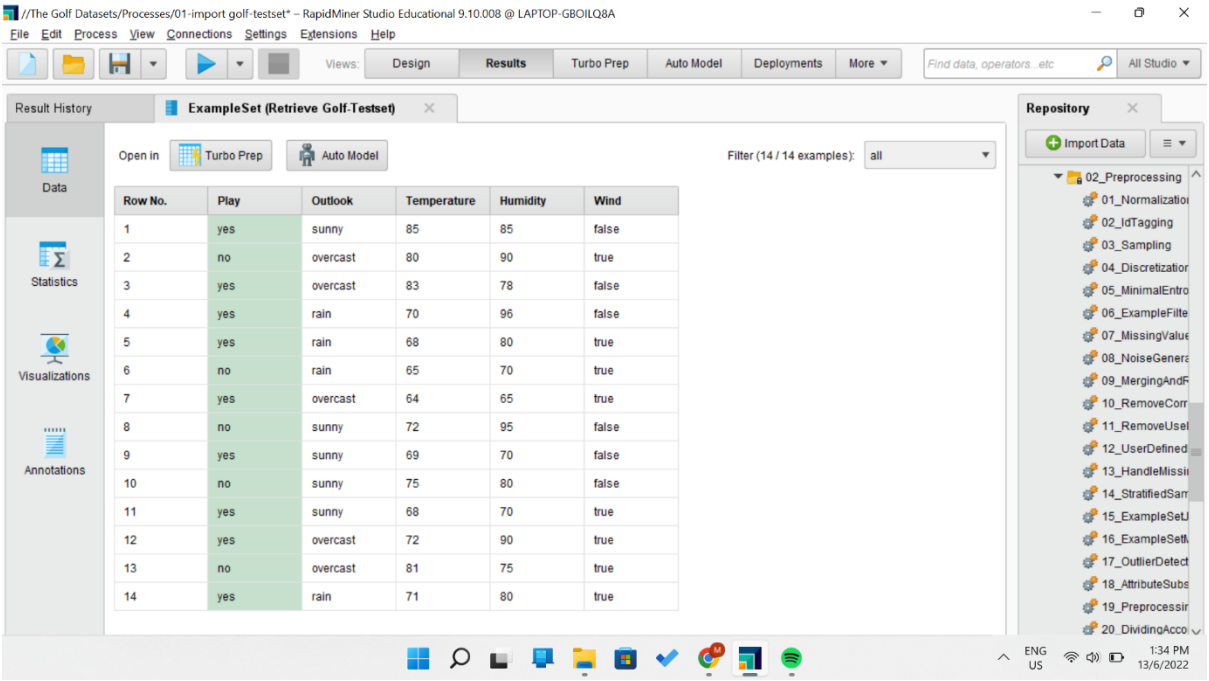
Should we play golf?

Golf data set contains attributes regarding the weather namely 'outlook', 'temperature', 'humidity' and 'wind'. Classify the Golf-Testset decide whether the game could be played or not by using decision tree for predicting the 'play' attribute.

0. Import Golf TestSet

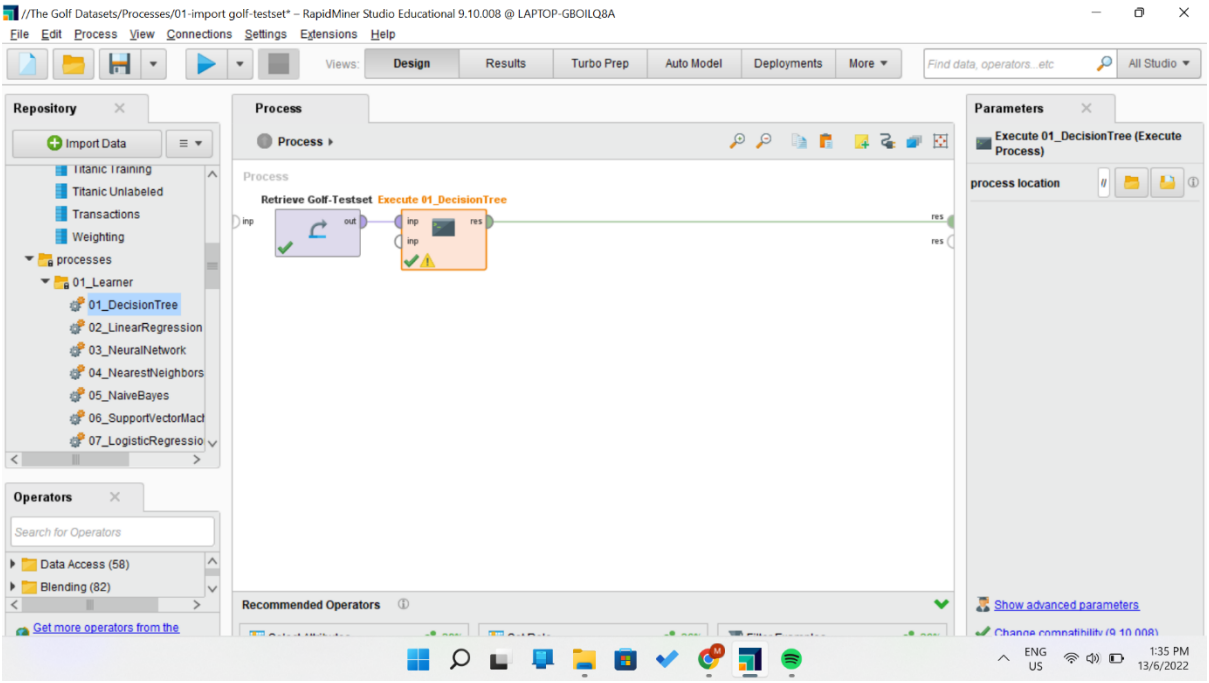


To import and store the Golf-Testset, retrieves the Golf-Testset from the Sample repository and drop it in the process and run to view the result.

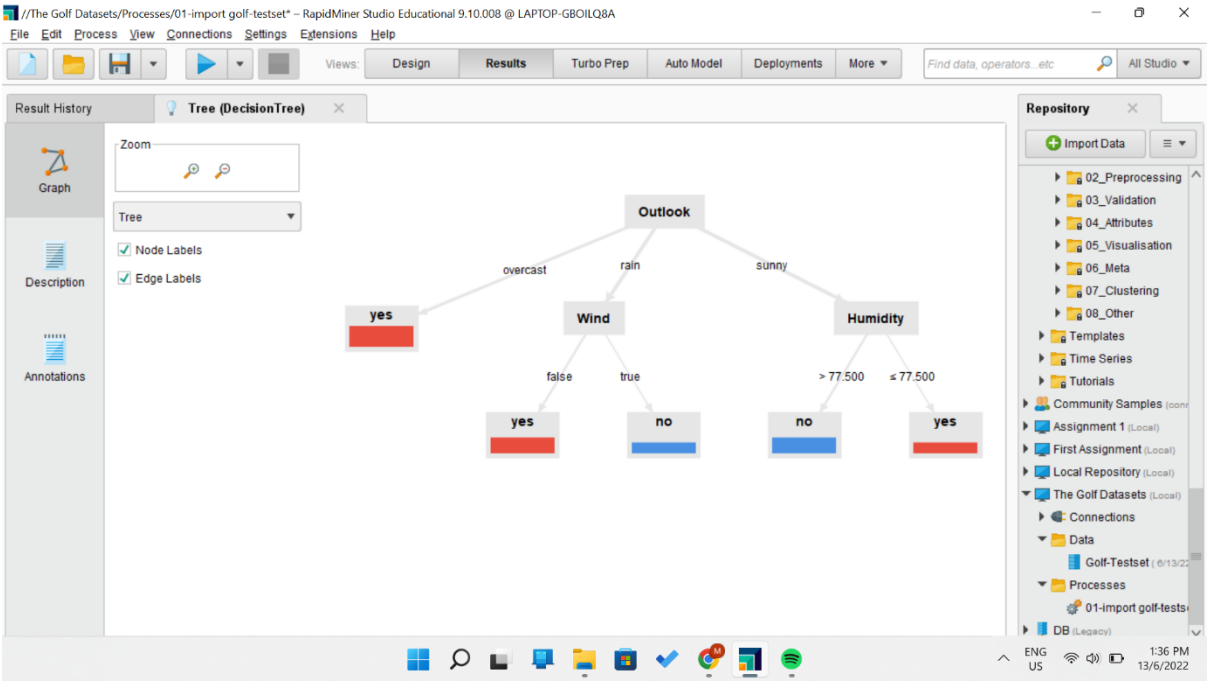


View result data retrieve Golf-Testset (play attribute is already labelled in label role)

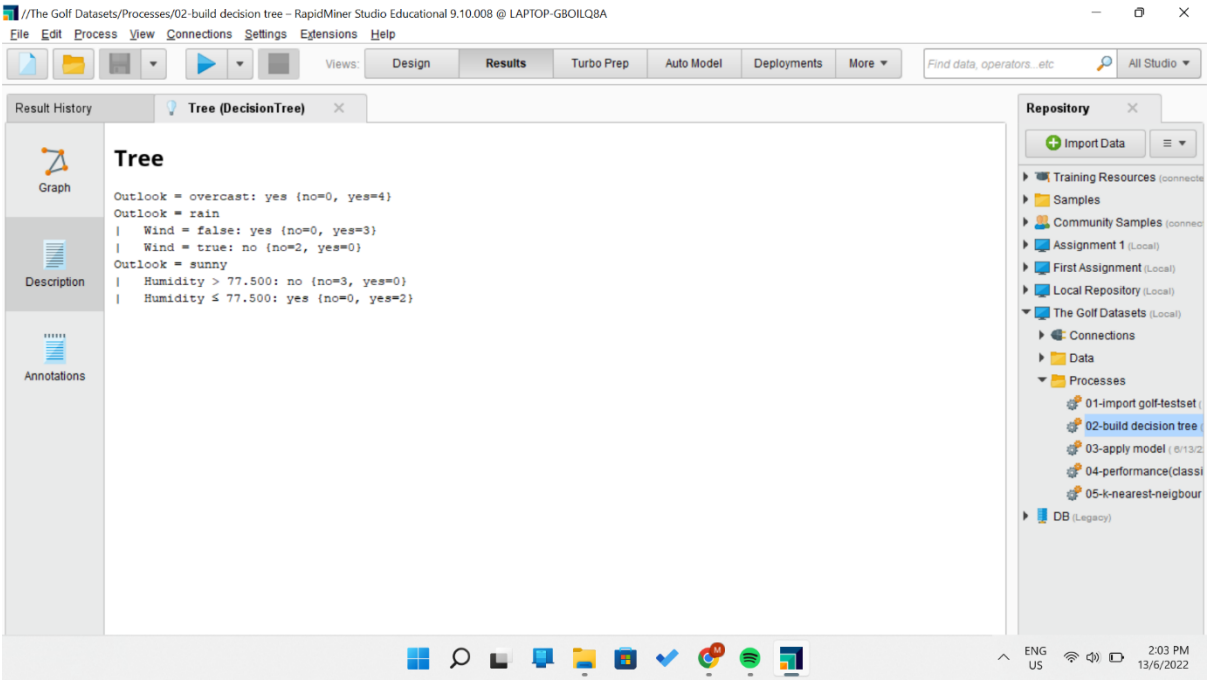
1. Decision Tree



Drop decision tree operator to the process and make passes to retrieves Golf-Testset. Now, the Decision tree operator receives the Golf data set from the Retrieve operator.

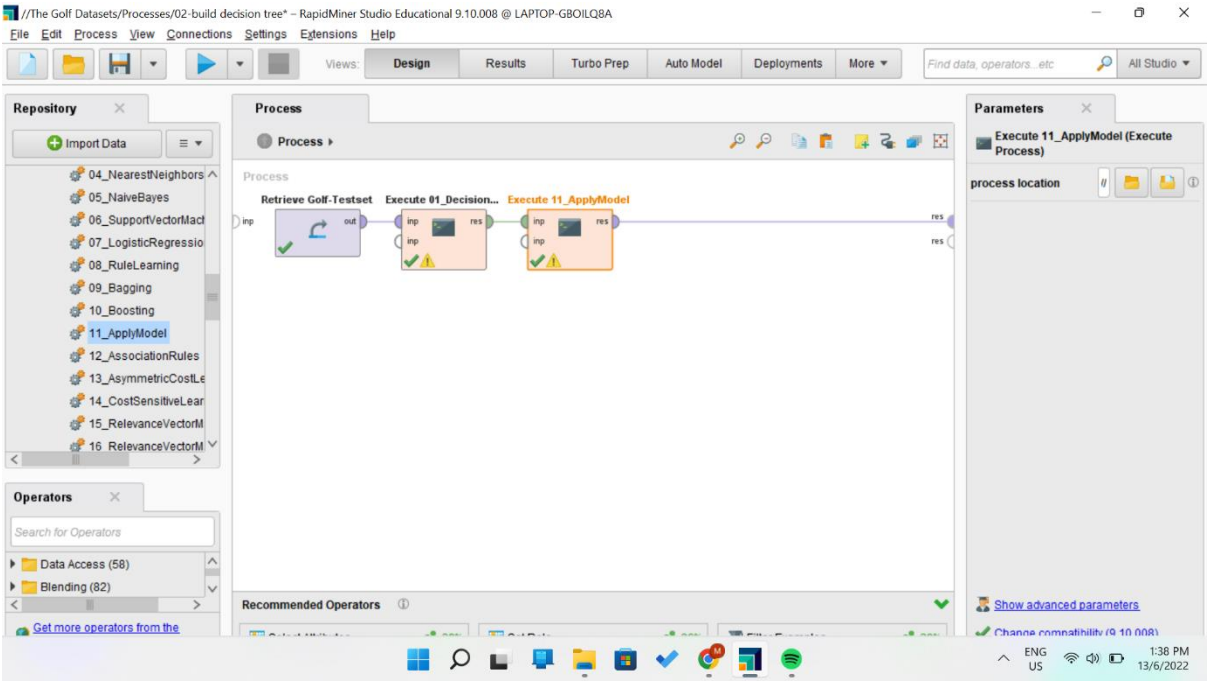


Click on the run button to view the Graph of Golf data set. This trains the decision tree model and takes to the results view, where can examine it graphically as well as in textual description. The tree shows that whenever the attribute 'outlook' has the value 'overcast', the attribute 'play' will have the value 'yes'. If the attribute 'outlook' has the value 'rain', then two outcomes are possible. If the attribute 'wind' has the value 'false', the 'play' attribute has the value 'yes'. If the 'wind' attribute has the value 'true', the attribute 'play' is 'no'.

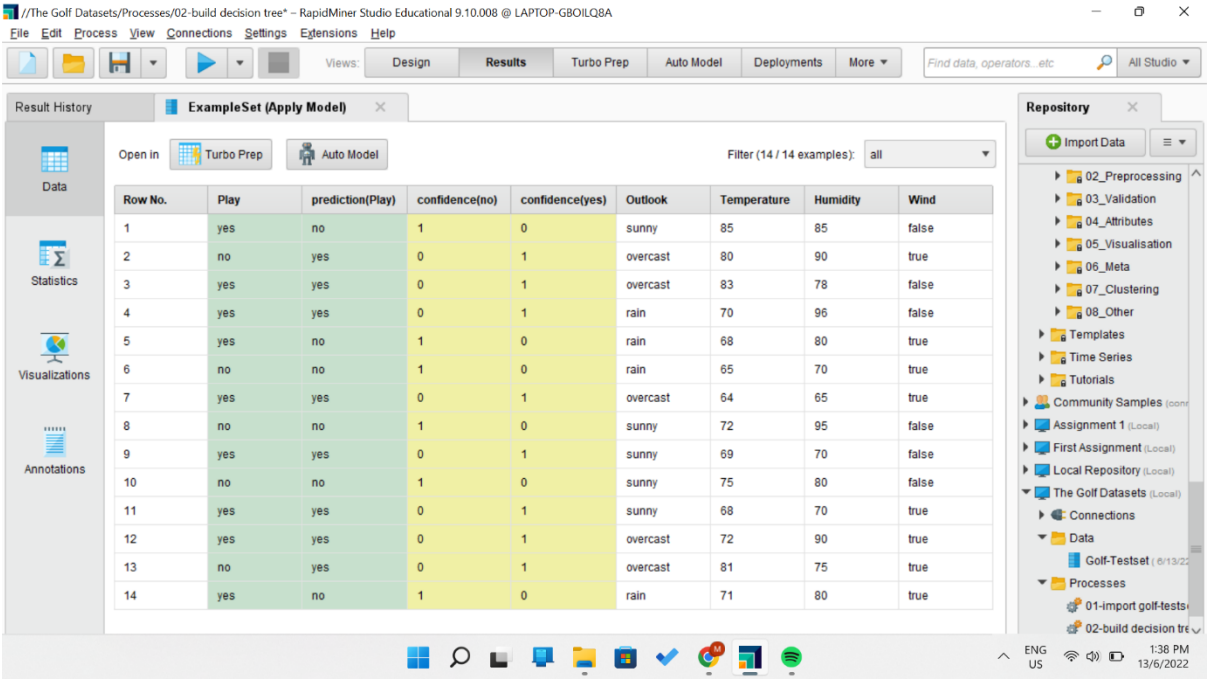


In the description, we can see if the attribute 'outlook' has the value 'sunny', there are again two possibilities. The attribute 'play' is 'yes' if the value of attribute 'humidity' is less than or equal to 77.5 and it is 'no' if 'humidity' is greater than 77.5.

2. Applying The Model



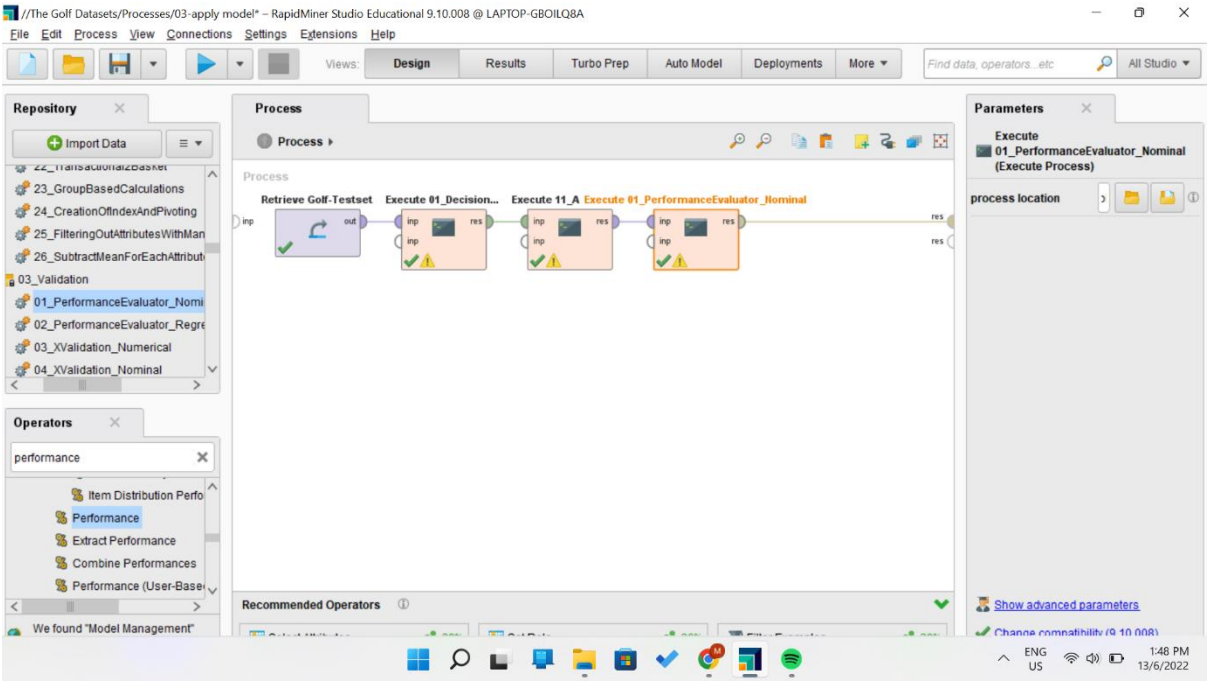
Drop ApplyModel operator and make passes to the decision tree to predict (play) and view the result.



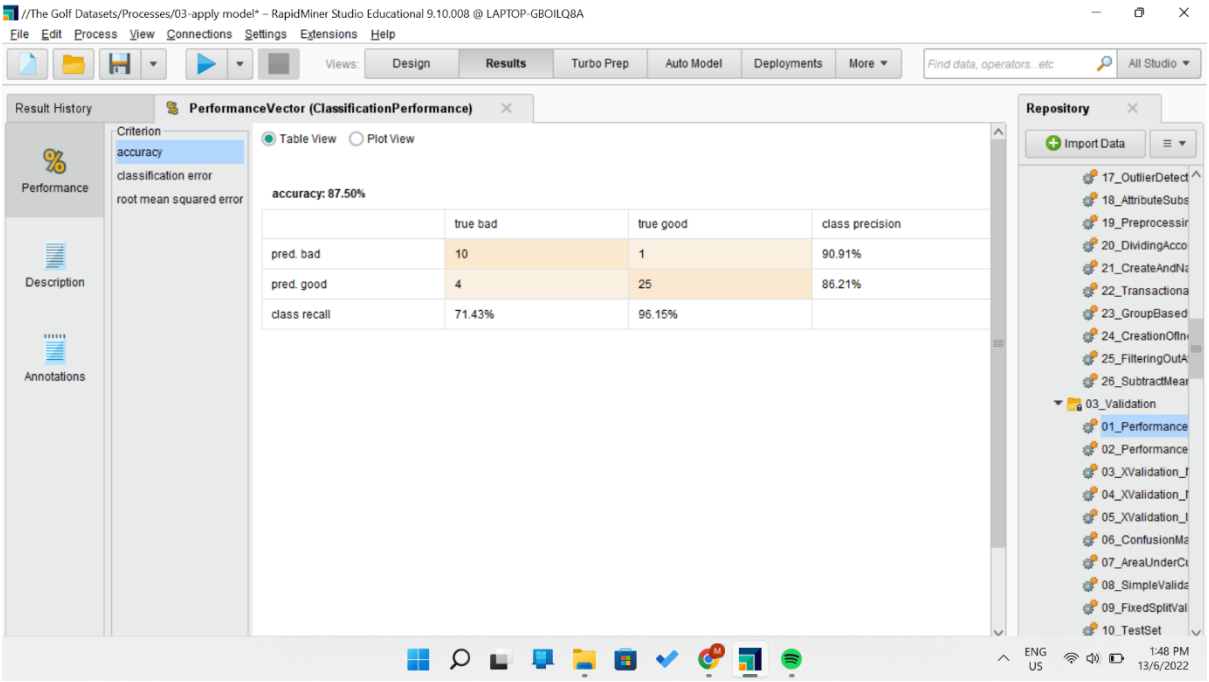
Row No.	Play	prediction(Play)	confidence(no)	confidence(yes)	Outlook	Temperature	Humidity	Wind
1	yes	no	1	0	sunny	85	85	false
2	no	yes	0	1	overcast	80	90	true
3	yes	yes	0	1	overcast	83	78	false
4	yes	yes	0	1	rain	70	96	false
5	yes	no	1	0	rain	68	80	true
6	no	no	1	0	rain	65	70	true
7	yes	yes	0	1	overcast	64	65	true
8	no	no	1	0	sunny	72	95	false
9	yes	yes	0	1	sunny	69	70	false
10	no	no	1	0	sunny	75	80	false
11	yes	yes	0	1	sunny	68	70	true
12	yes	yes	0	1	overcast	72	90	true
13	no	yes	0	1	overcast	81	75	true
14	yes	no	1	0	rain	71	80	true

After click on the run button, we can see play attributes as the original from the golf-testset itself and prediction for that play attributes. As we can see, five row doesn't match from the prediction. Row 1 the data set said we can play golf but the prediction is no. Row 2 said that we cannot play but the prediction is we can play the golf. In row 5, the prediction is no but the data said yes as we can play golf. In row 13, the prediction is yes and in the row 14 the prediction is no. After we applying the model the confidence (no and yes) collumn also shown in the table.

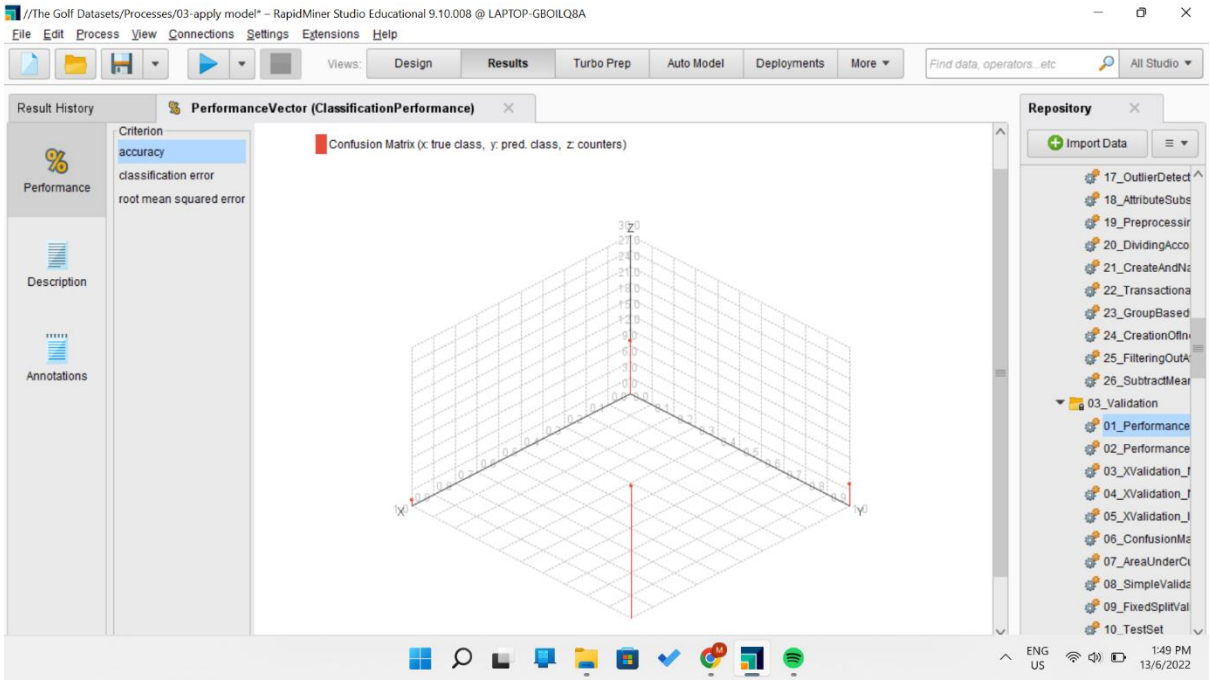
3. Performance (Classification) *[with decision tree]*



Drag the 'performance' in the operator and drop that performance operator examine the accuracy.

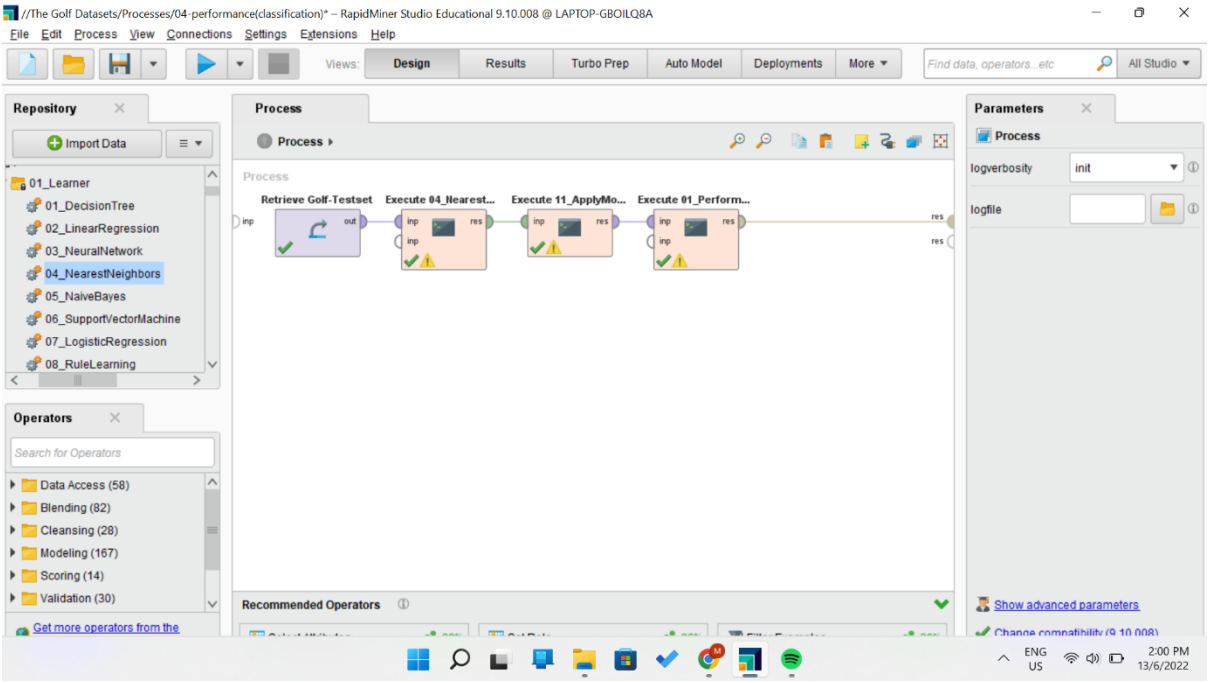


Click on the run button to view the results of accuracy. The result of accuracy shown 87.50%.

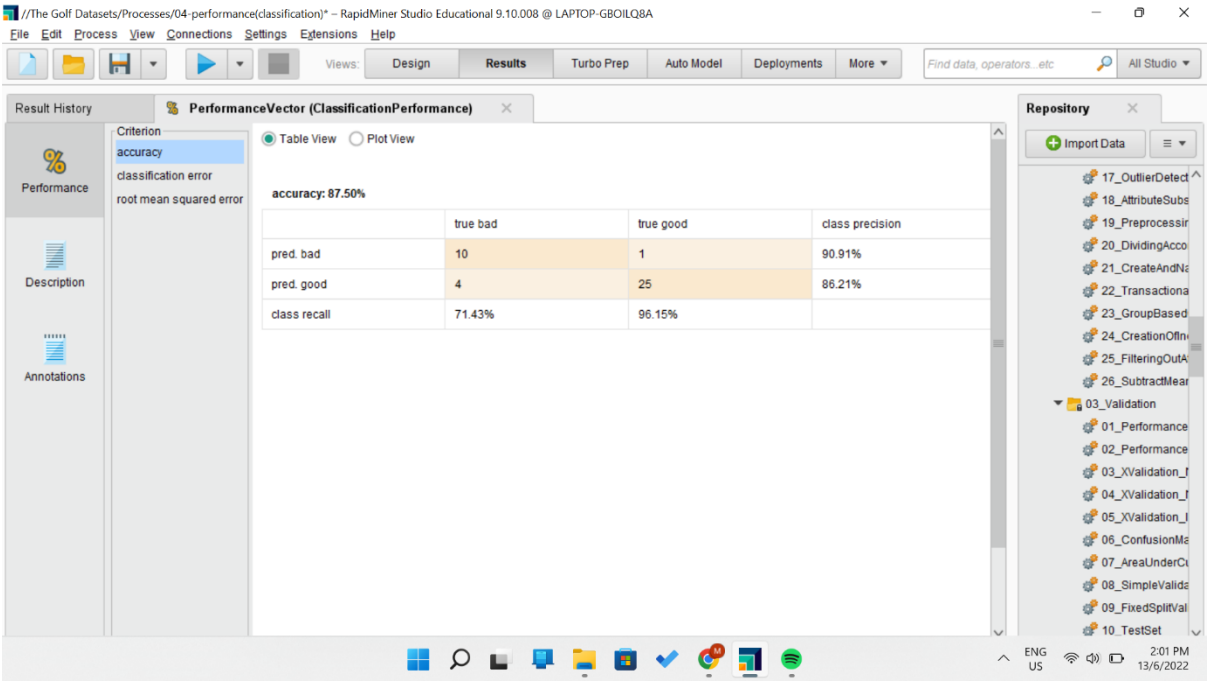


This is the confusion matrix for the accuracy of playing golf with the decision tree.

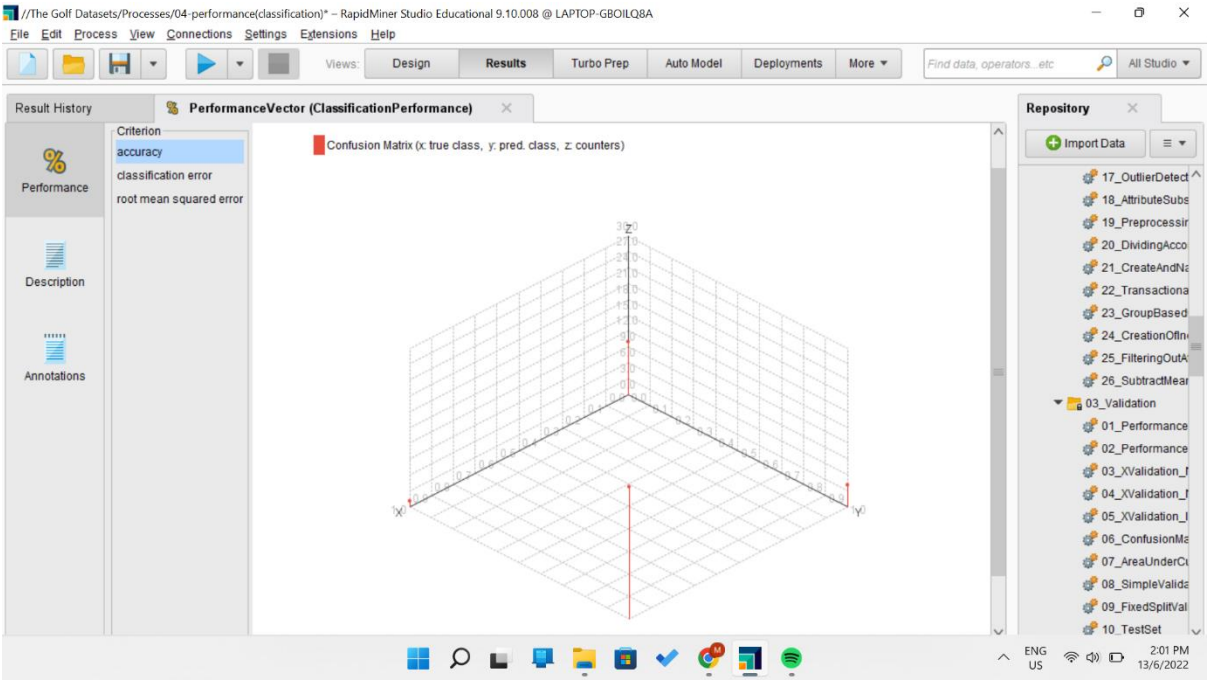
4. Performance (Classification) *[with k-nearest-neighbour]*



Replace the decision tree operator to the K-Nearest Neighbour (K-NN) classifier to find out does our accuracy will be change after that replacement.



The K-NN classifier does not had any change to our accuracy before [with decision tree]. The accuracy of playing golf after the replacement also 87.50%.



This is the confusion matrix for the accuracy of playing golf with the K-Nearest Neighbour (K-NN).