

Machine Learning

Tool Study : Comparative Analysis

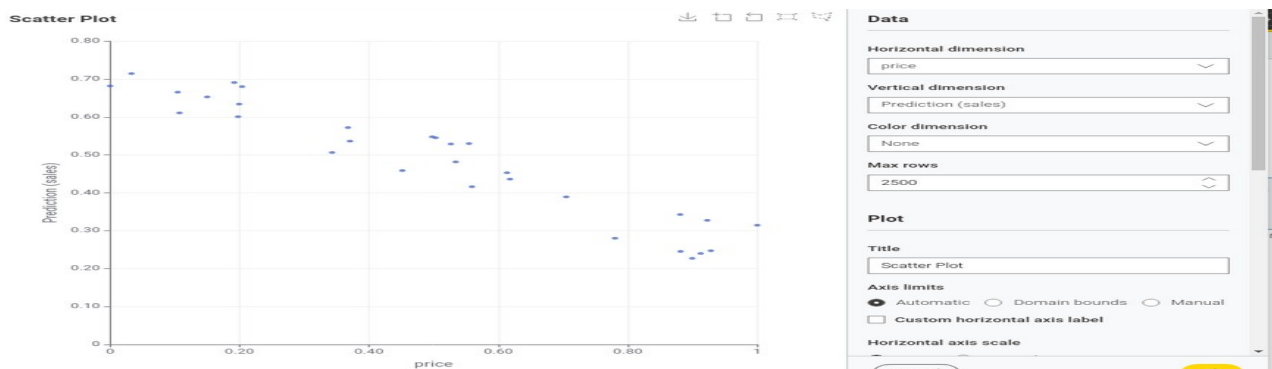
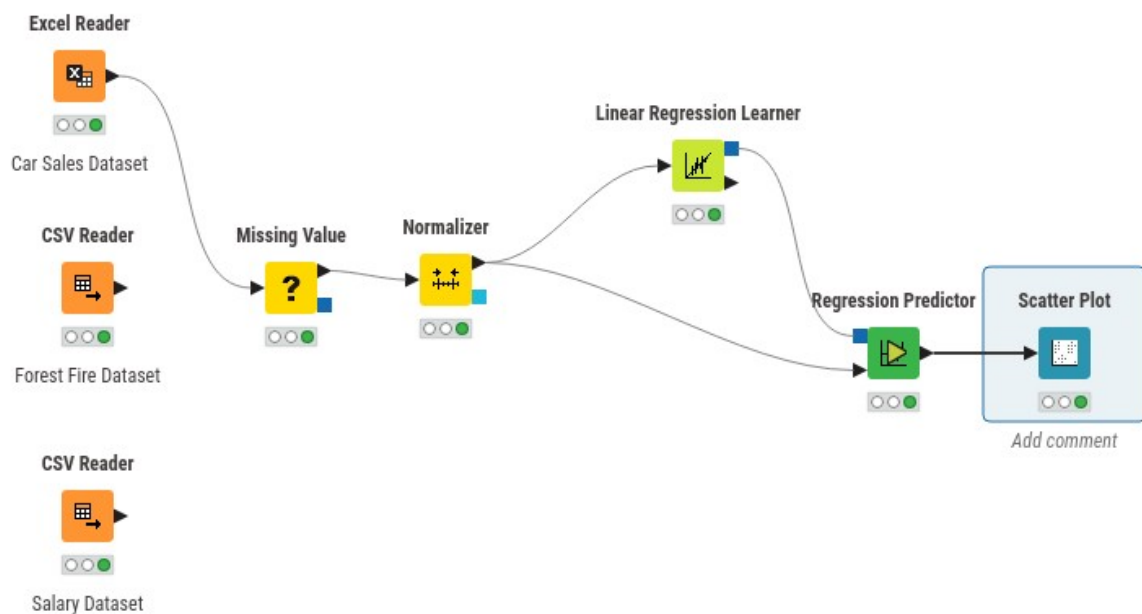
Name: Athithraja R

Reg.no: 2022503702

Tool: KNIME

Linear Regression:

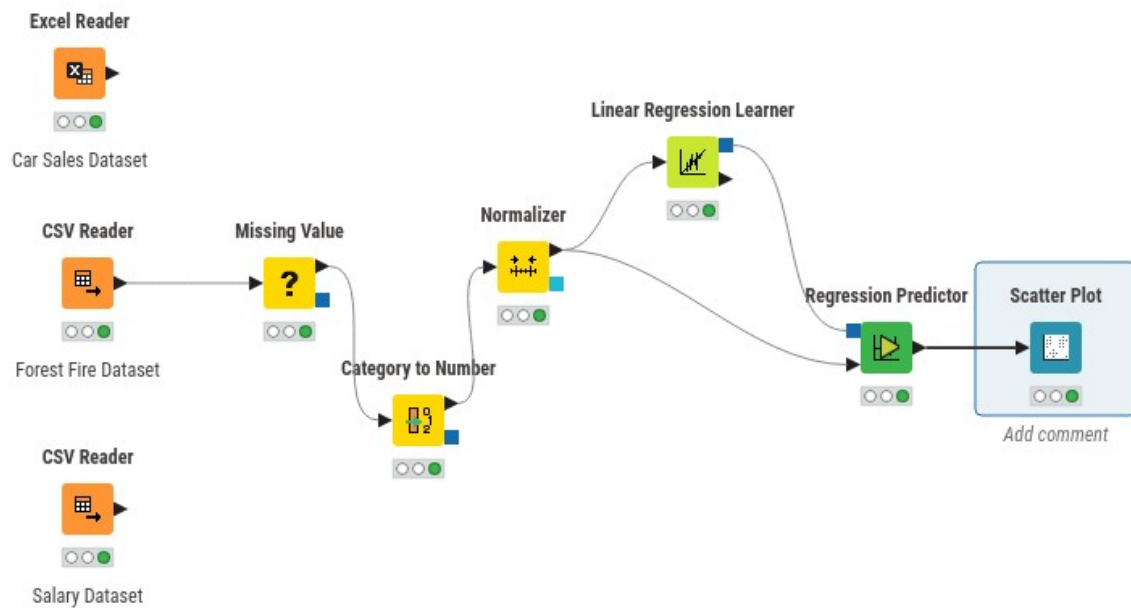
1) Car Sale Dataset



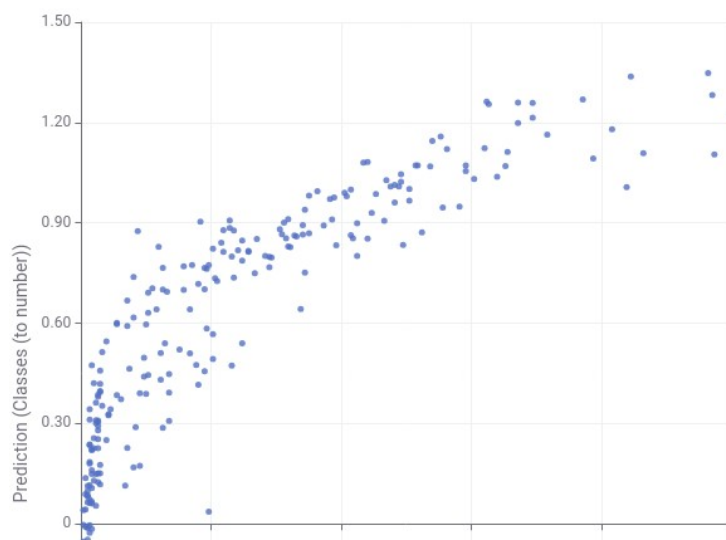
→ Only plot 2d or 3d

Linear Regression:

2) Forest Fire Dataset



Scatter Plot



Data

Horizontal dimension

FWI

Vertical dimension

Prediction (Classes (to number))

Color dimension

None

Max rows

2500

Plot

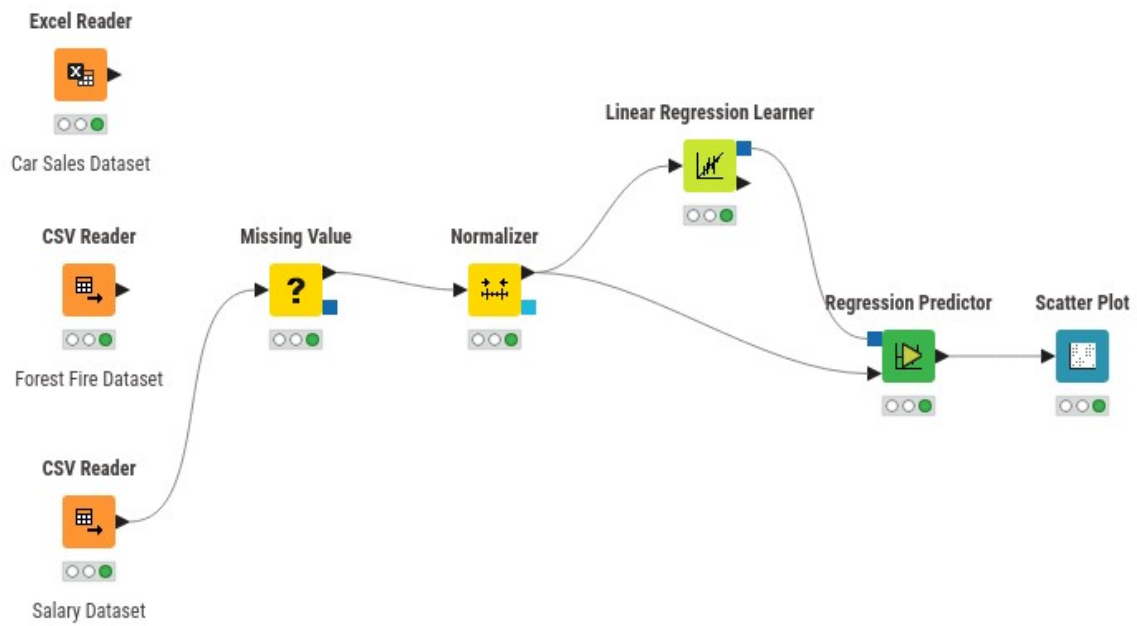
Title

Scatter Plot

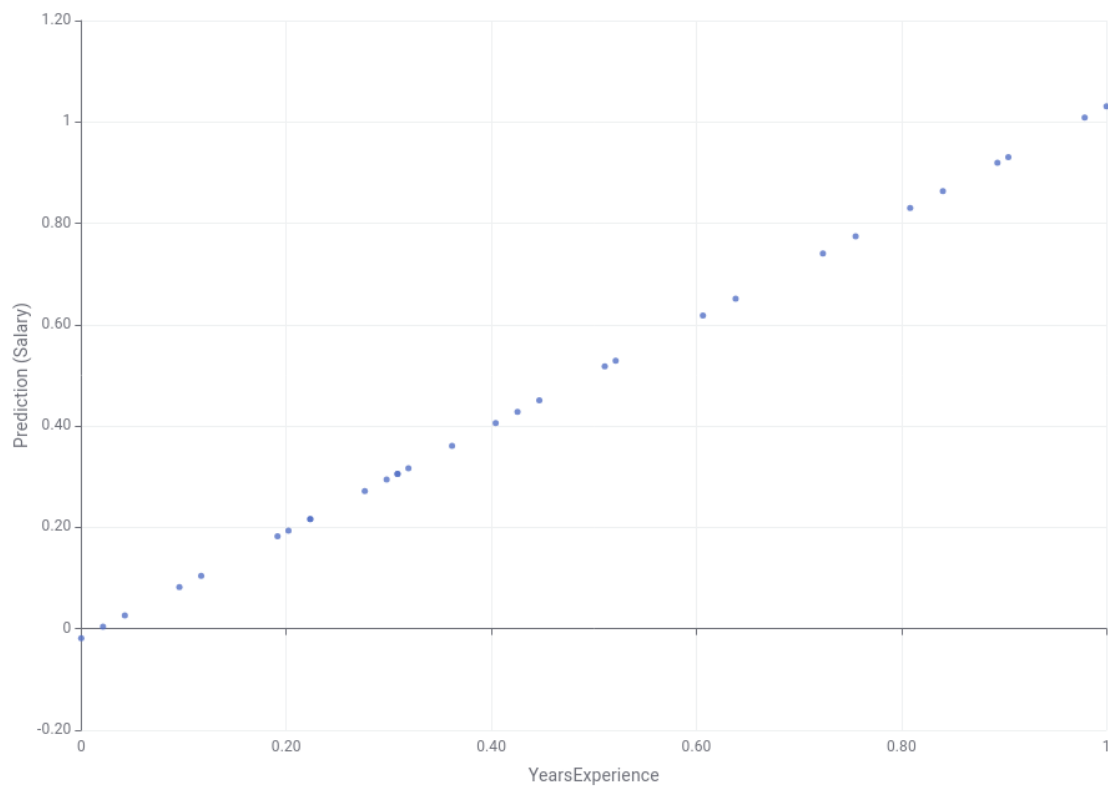
→ Only plot 2d or 3d

Linear Regression:

3) Salary Dataset

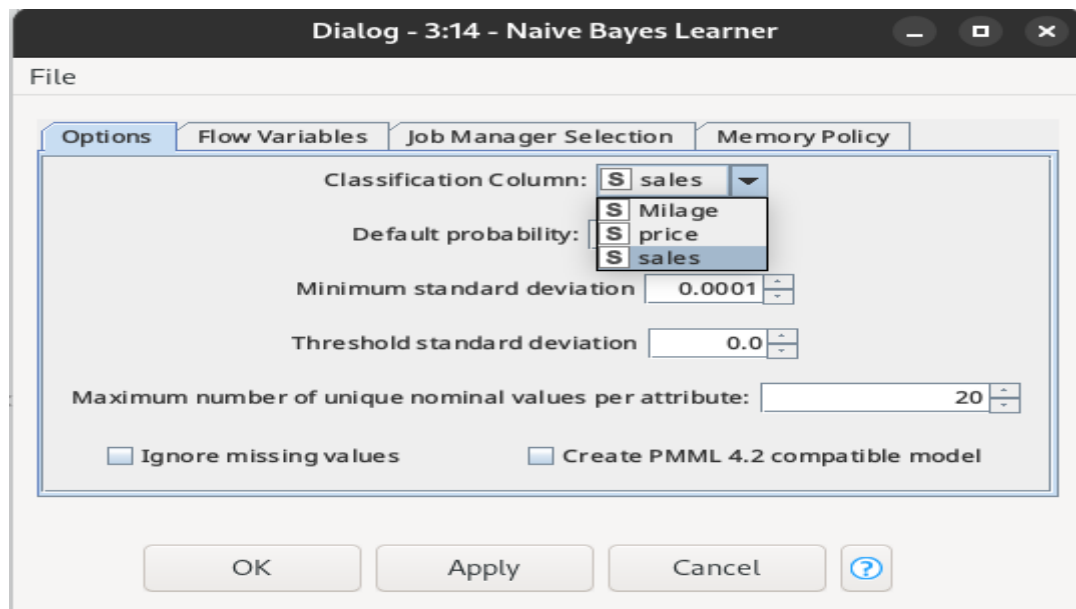
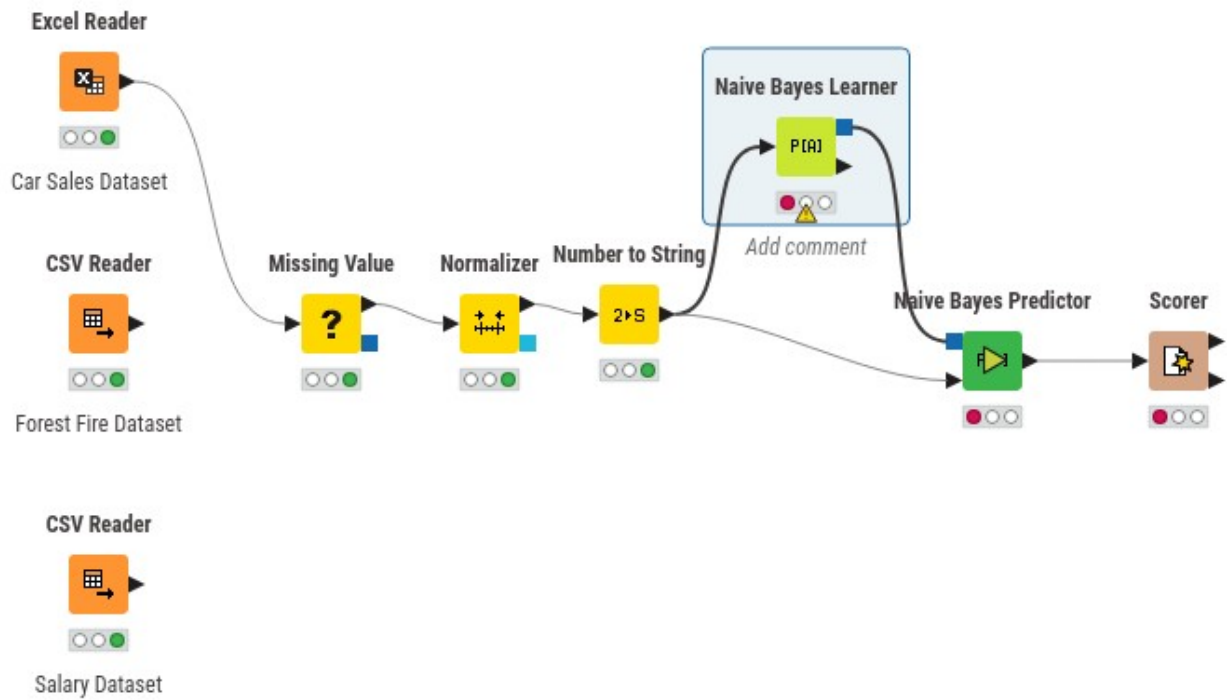


Scatter Plot



Naive Bayes

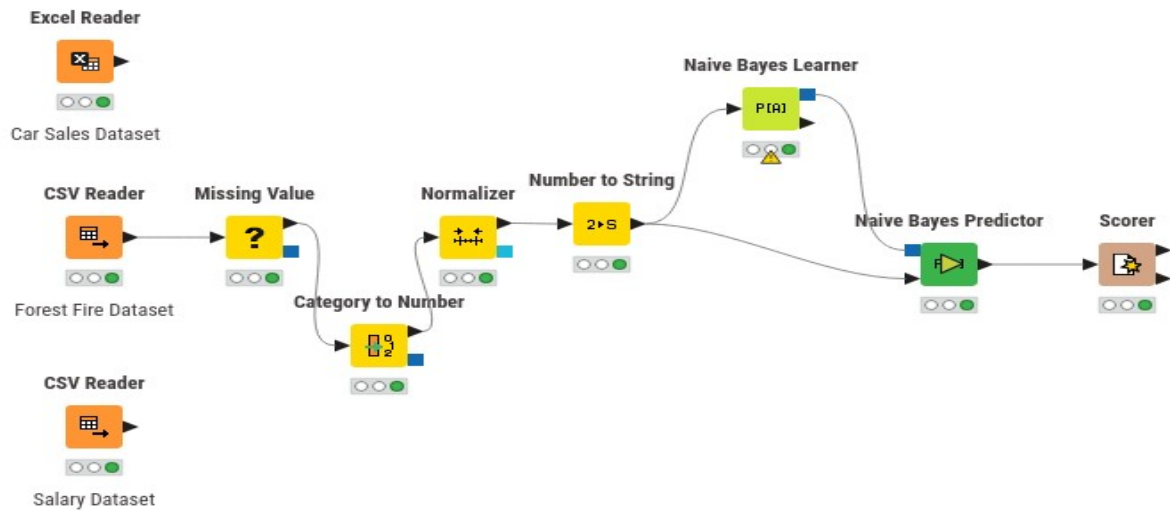
1) Car Sale Dataset



→ Only suits for classification based data

Naive Bayes

2) Forest Fire Dataset



► 1: Confusion matrix ► 2: Accuracy statistics Flow Variables

Rows: 3 | Columns: 11

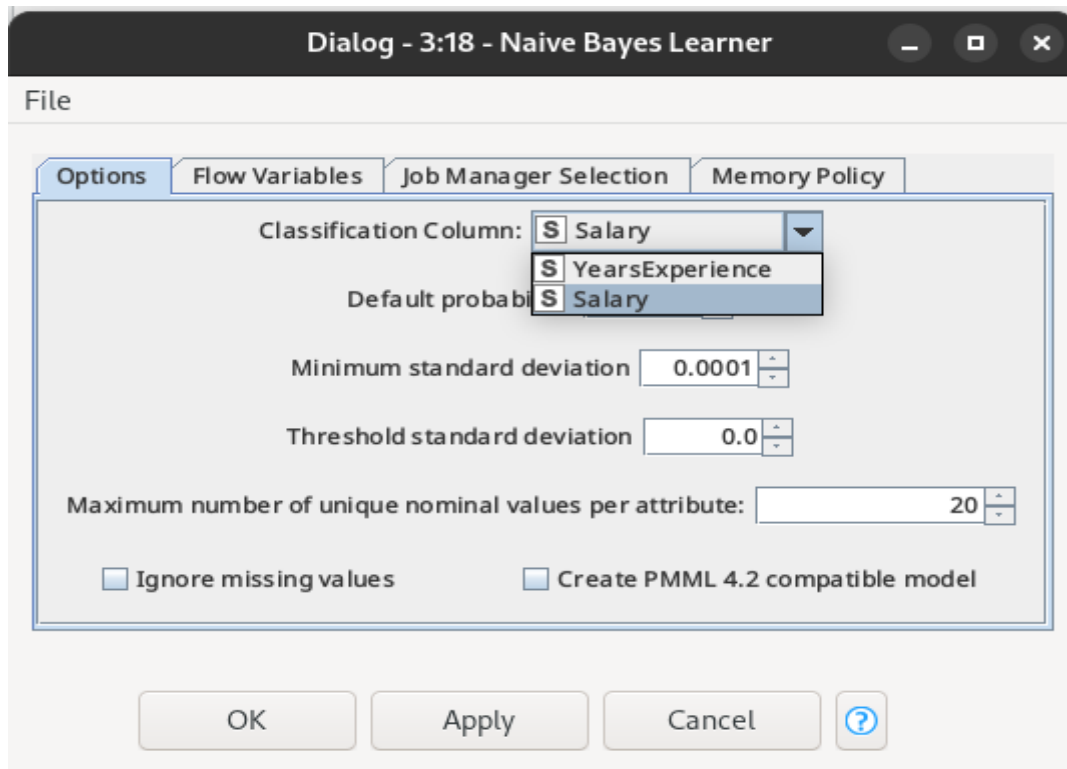
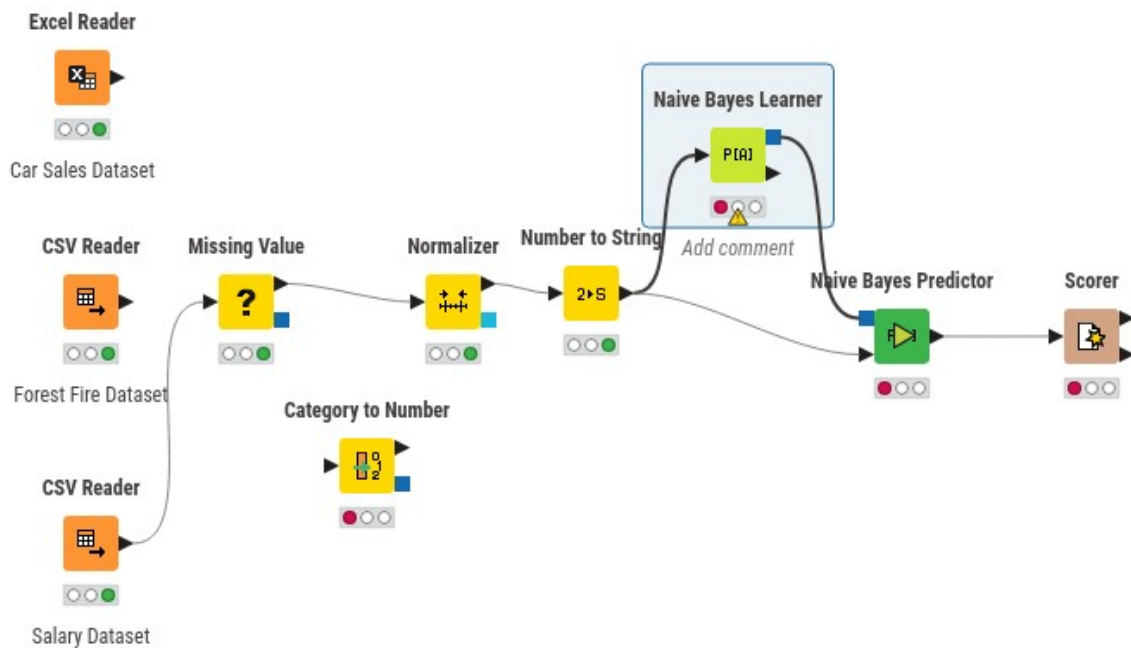
Table Statistics

<input type="checkbox"/>	#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNeg... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specificit... Number (dou...
<input type="checkbox"/>	1	0.0	106	0	137	0	1	1	1	1
<input type="checkbox"/>	2	1.0	137	0	106	0	1	1	1	1
<input type="checkbox"/>	3	Overall								

Confusion Matrix - 3:13 - Scorer				
File Hilite				
Classes (to ...	0.0	1.0		
0.0	106	0		
1.0	0	137		
Correct classified: 243			Wrong classified: 0	
Accuracy: 100%			Error: 0%	
Cohen's kappa (κ): 1%				

Naive Bayes

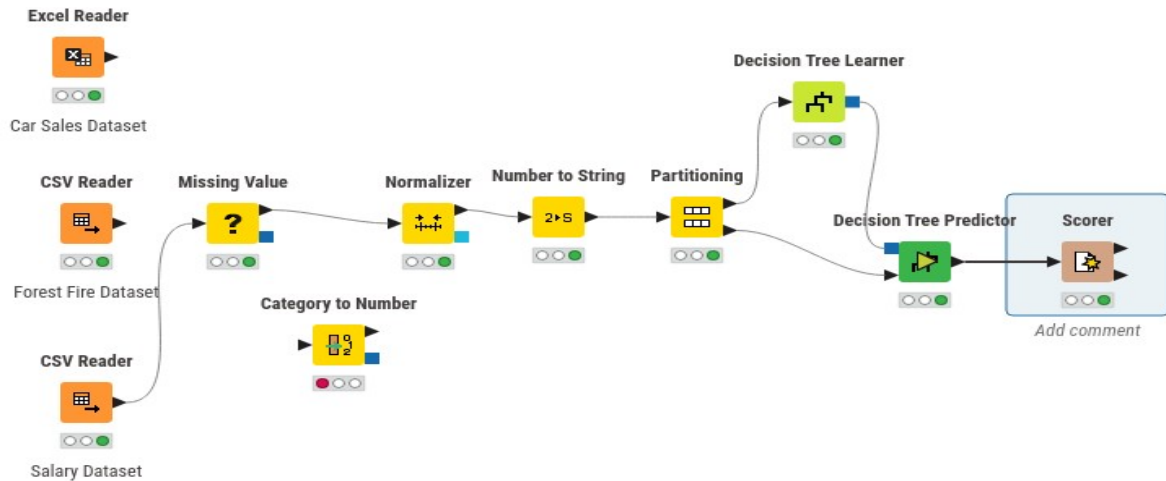
3) Salary Dataset



→ Only suits for classification based data

Decision Tree

1) Salary Dataset



► 1: Confusion matrix ► 2: Accuracy statistics ◀ Flow Variables

Rows: 31 | Columns: 11

Table Statistics

#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNeg... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specifi... Number (i...
1	0.0190	0	9	0	0	?	0	?	0
2	0.1000	0	0	9	0	?	?	?	1
3	0.0	0	0	9	0	?	?	?	1
4	0.0684	0	0	8	1	0	?	0	1
5	0.0255	0	0	8	1	0	?	0	1
6	0.2233	0	0	9	0	?	?	?	1
7	0.2648	0	0	9	0	?	?	?	1
8	0.1974	0	0	8	1	0	?	0	1
9	0.3155	0	0	9	0	?	?	?	1
10	0.2298	0	0	8	1	0	?	0	1

Confusion Matrix - 3:13 - Scorer

File	Hit				
Salary \ Pre...	0.0190408...	0.1000944...	0.0	0.0684384...	0.0255138...
0.0190408...	0	0	0	0	0
0.1000944...	0	0	0	0	0
0.0	0	0	0	0	0
0.0684384...	1	0	0	0	0
0.0255138...	1	0	0	0	0
0.2233758...	0	0	0	0	0
0.2648121...	0	0	0	0	0
0.1974249...	1	0	0	0	0
0.3155445...	0	0	0	0	0
0.2298369...	1	0	0	0	0
0.3010512...	0	0	0	0	0
0.2133593...	0	0	0	0	0
0.2270966...	0	0	0	0	0
0.2285613...	0	0	0	0	0
0.2761634...	0	0	0	0	0
0.3568036...	1	0	0	0	0
0.3342546...	1	0	0	0	0
0.5357547...	0	0	0	0	0
0.5153791...	0	0	0	0	0
0.6639381...	1	0	0	0	0
0.6379281...	0	0	0	0	0
0.7151193...	1	0	0	0	0
0.7508977...	0	0	0	0	0
0.8986652...	0	0	0	0	0
0.8469170...	0	0	0	0	0

Correct classified: 0

Wrong classified: 9

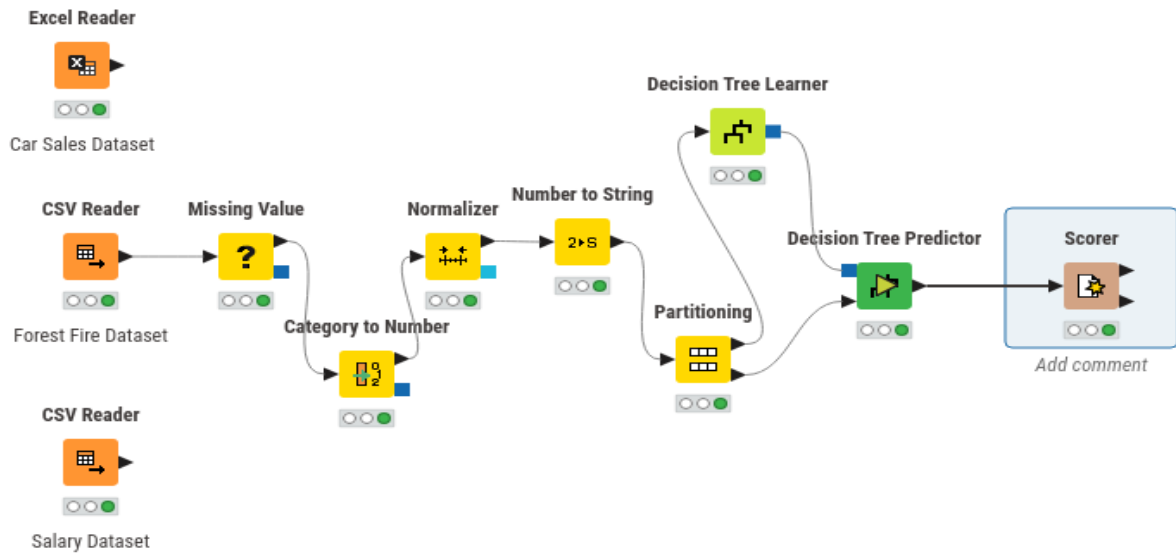
Accuracy: 0%

Error: 100%

Cohen's kappa (κ): 0%

Decision Tree

2) Forest Fire Dataset



► 1: Confusion matrix ► 2: Accuracy statistics Flow Variables

Rows: 3 | Columns: 11

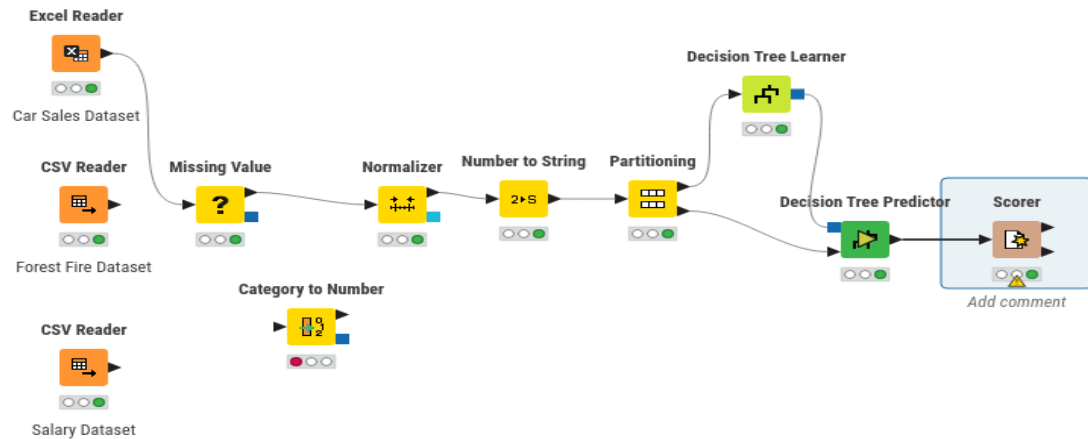
Table Statistics

<input type="checkbox"/>	#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNega... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specificit... Number (dou...
<input type="checkbox"/>	1	0.0	31	0	42	0	1	1	1	1
<input type="checkbox"/>	2	1.0	42	0	31	0	1	1	1	1
<input type="checkbox"/>	3	Overall								

File	Hit
Classes (to ...	0.0 1.0
0.0	31 0
1.0	0 42
Correct classified: 73	
Wrong classified: 0	
Accuracy: 100%	
Error: 0%	
Cohen's kappa (κ): 1%	

Decision Tree

3) Car Sale Dataset



► 1: Confusion matrix ► 2: Accuracy statistics ☑ Flow Variables

Rows: 28 | Columns: 11

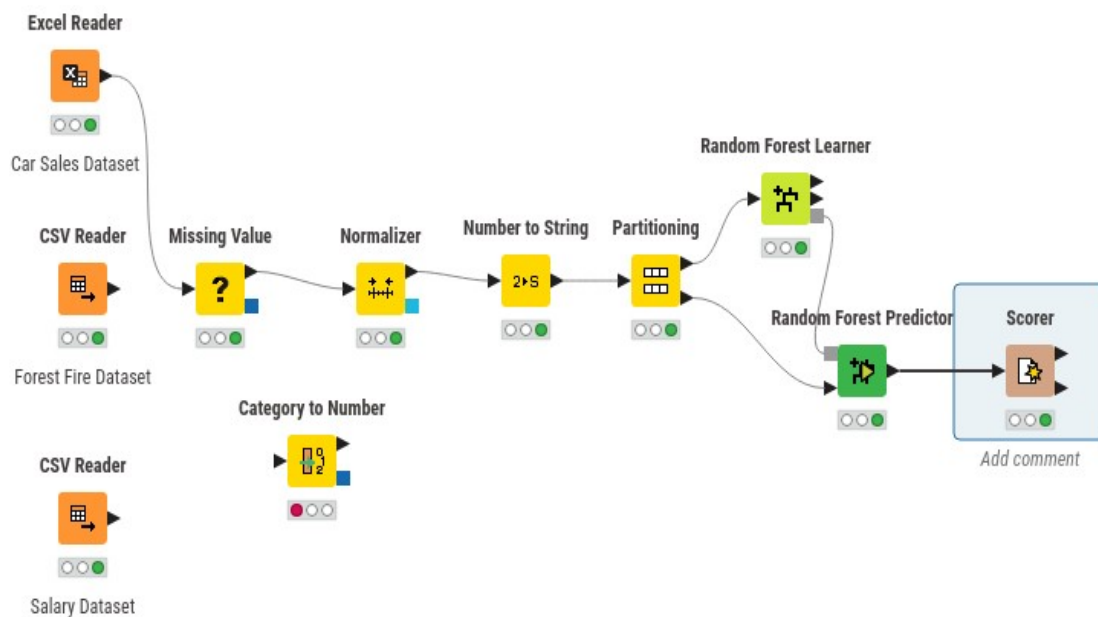
Table Statistics

<input type="checkbox"/>	#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNega... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specifi... Number (i
<input type="checkbox"/>	1	0.8842	0	1	3	0		0		0.75
<input type="checkbox"/>	2	0.1157	0	1	3	0		0		0.75
<input type="checkbox"/>	3	0.4210	0	1	3	0		0		0.75
<input type="checkbox"/>	4	0.2526	0	1	3	0		0		0.75
<input type="checkbox"/>	5	0.6526	0	0	4	0				1
<input type="checkbox"/>	6	0.3368	0	0	4	0				1
<input type="checkbox"/>	7	0.1684	0	0	3	1	0		0	1

Confusion Matrix - 3:13 - Scorer						
File Hilite						
⚠ There were missing values in the reference or in the prediction class column						
sales \ Pre...	0.8842105...	0.1157894...	0.4210526...	0.2526315...	0.6526315...	
0.8842105...	0	0	0	0	0	
0.1157894...	0	0	0	0	0	
0.4210526...	0	0	0	0	0	
0.2526315...	0	0	0	0	0	
0.6526315...	0	0	0	0	0	
0.3368421...	0	0	0	0	0	
0.1684210...	1	0	0	0	0	
0.0842105...	0	0	0	0	0	
0.5684210...	0	0	0	0	0	
0.6947368...	0	0	0	0	0	
0.9157894...	0	0	0	0	0	
0.4631578...	0	0	0	0	0	
0.1052631...	0	0	0	1	0	
0.6421052...	0	0	0	0	0	
0.0105263...	0	0	0	0	0	
0.6736842...	0	0	0	0	0	
0.0210526...	0	0	0	0	0	
0.8210526...	0	0	0	0	0	
0.0421052...	0	0	0	0	0	
0.7052631...	0	0	0	0	0	
0.9999999...	0	0	1	0	0	
0.8631578...	0	0	0	0	0	
0.7789473...	0	0	0	0	0	
0.4736842...	0	0	0	0	0	
0.7473684...	0	0	0	0	0	
Correct classified: 0 Wrong classified: 4						
Accuracy: 0% Error: 100%						
Cohen's kappa (κ): 0%						

Random Forest Algorithm

1) Car Sale Dataset



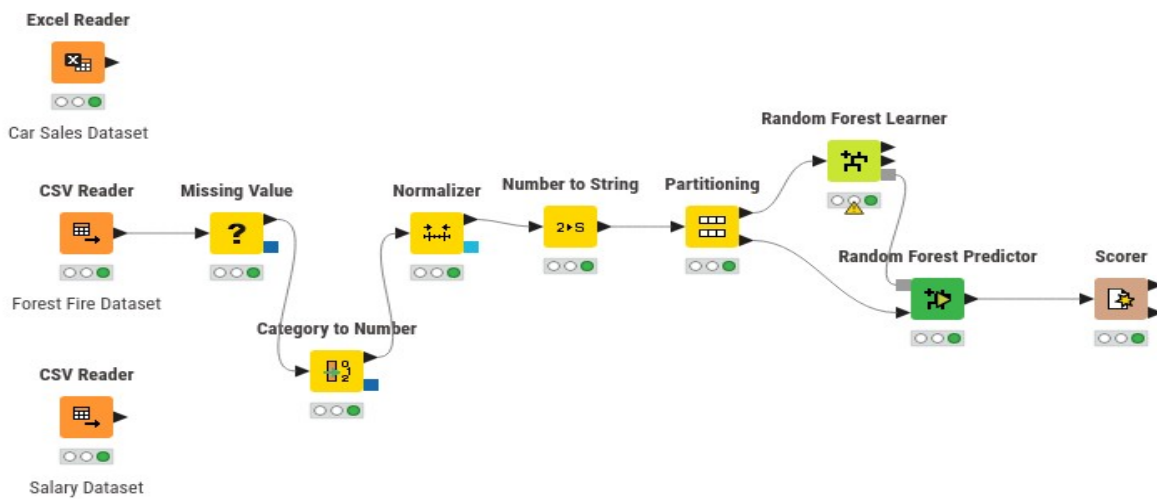
File	Hillite	0.1684210...	0.1052631...	0.1157894...	0.8210526...	0.4736842...
sales \ Pre...	0.1684210...	0	0	0	0	0
0.1052631...	0	0	0	0	0	0
0.1157894...	0	0	0	0	0	0
0.8210526...	0	0	0	0	0	0
0.4736842...	0	0	0	0	0	0
0.8842105...	1	0	0	0	1	0
0.6526315...	0	0	0	0	0	0
0.3368421...	0	0	0	0	0	0
0.0842105...	0	0	0	0	0	0
0.5684210...	0	0	0	0	0	0
0.6947368...	0	0	0	0	0	0
0.9157894...	0	0	0	0	0	0
0.4631578...	0	0	0	0	0	0
0.6421052...	0	0	0	0	0	0
0.0105263...	0	0	0	0	0	0
0.6736842...	1	0	0	0	0	0
0.0210526...	0	0	0	0	0	0
0.0421052...	0	0	0	0	0	0
0.4210526...	0	0	0	0	0	0
0.7052631...	1	0	0	0	0	0
0.9999999...	0	0	0	0	0	1
0.8631578...	1	0	0	0	0	0
0.7789473...	1	0	0	0	0	0
0.7473684...	0	0	0	0	0	0
0.1473684...	0	0	0	0	0	0

Correct classified: 0 Wrong classified: 9
 Accuracy: 0% Error: 100%
 Cohen's kappa (κ): 0%

→ Only suits for classification based data

Random Forest Algorithm

2) Forest Fire Dataset



► 1: Confusion matrix ► 2: Accuracy statistics ⚙ Flow Variables

Rows: 3 | Columns: 11

Table Statistics

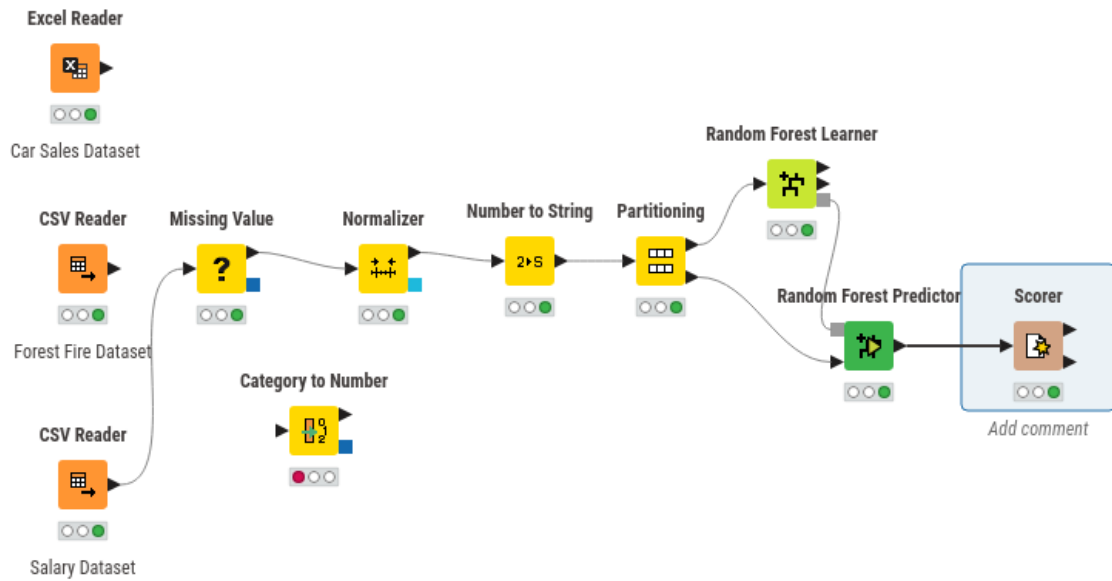
#	RowID	TruePosit... Number (inte...	FalsePosi... Number (inte...	TrueNega... Number (inte...	FalseNeg... Number (inte...	Recall Number (dou...	Precision Number (dou...	Sensitivity Number (dou...	Specificity Number (dou...
1	0.0	30	17	21	5	0.857	0.638	0.857	0.553
2	1.0	21	5	30	17	0.553	0.808	0.553	0.857
3	Overall	?	?	?	?	?	?	?	?

Classes (to ...	0.0	1.0
0.0	30	5
1.0	17	21

Correct classified: 51	Wrong classified: 22
Accuracy: 69.863%	Error: 30.137%
Cohen's kappa (κ): 0.404%	

Random Forest Algorithm

3) Salary Dataset



Confusion Matrix - 3:13 - Scorer					
File	Highlight				
Salary \ Pre...	0.2133593...	0.8847625...	0.0190408...	0.1000944...	0.0
0.2133593...	0	0	0	0	0
0.8847625...	0	0	0	0	0
0.0190408...	0	0	0	0	0
0.1000944...	0	0	0	0	0
0.0	0	0	0	0	0
0.0684384...	0	0	0	0	0
0.0255138...	0	0	0	0	0
0.2233758...	0	0	0	0	0
0.2648121...	0	0	0	0	0
0.1974249...	0	0	0	0	0
0.3155445...	0	0	0	0	0
0.2298369...	0	1	0	0	0
0.3010512...	0	0	0	0	0
0.2270966...	1	0	0	0	0
0.2285613...	0	1	0	0	0
0.2761634...	0	0	0	0	0
0.3568036...	0	0	0	0	0
0.3342546...	0	1	0	0	0
0.5357547...	0	0	0	0	0
0.5153791...	0	0	0	0	0
0.6639381...	0	0	0	0	0
0.6379281...	0	1	0	0	0
0.7151193...	0	0	0	0	0
0.7508977...	0	1	0	0	0
0.8986652...	0	1	0	0	0
Correct classified: 0					
Wrong classified: 9					
Accuracy: 0%					
Error: 100%					
Cohen's kappa (κ): 0%					

→ Only suits for classification based data