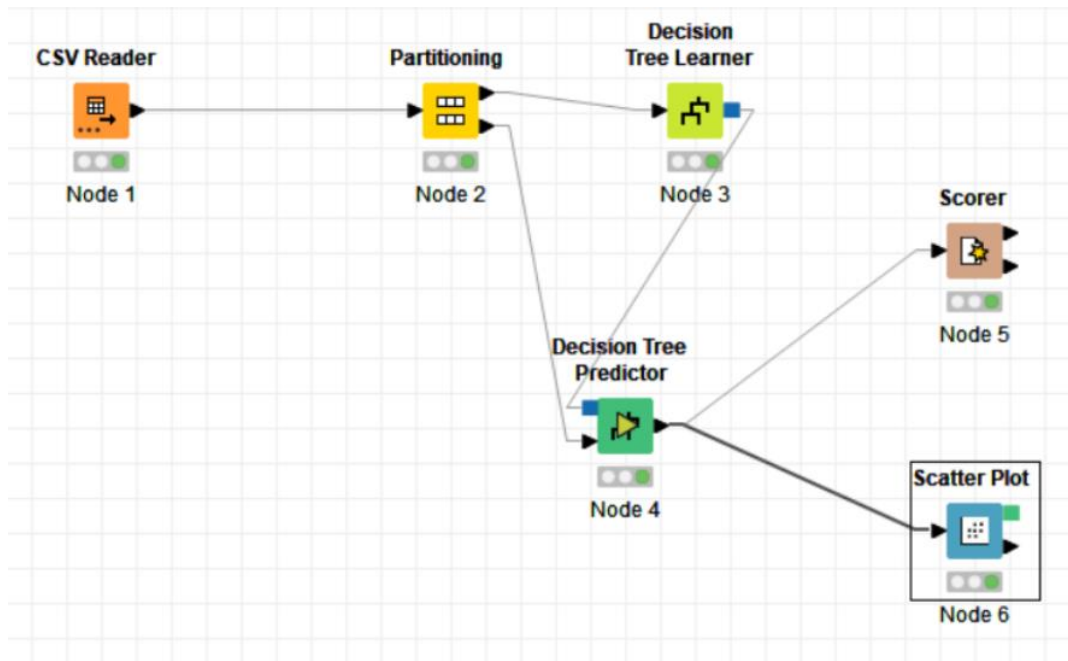


ILLUSTRATION OF 3 ALGORITHMS USING KNIME TOOL

DATASET: IRIS DATASET

1. CLASSIFICATION (DECISION TABLE) WORKFLOW



Accuracy statistics - 4:5 - Scorer

File Edit Hilite Navigation View

Table "default" - Rows: 4 Spec - Columns: 11 Properties Flow Variables

Row ID	I TruePo...	I FalsePo...	I TrueNe...	I FalseN...	D Recall	D Precision	D Sensitivity	D Specifcity	D F-meas...	D Accuracy	D Cohen'...
Iris-setosa	16	0	29	0	1	1	1	1	1	?	?
Iris-versicolor	15	0	28	2	0.882	1	0.882	1	0.938	?	?
Iris-virginica	12	2	31	0	1	0.857	1	0.939	0.923	?	?
Overall	?	?	?	?	?	?	?	?	?	0.956	0.933

Confusion matrix - 4:5 - Scorer

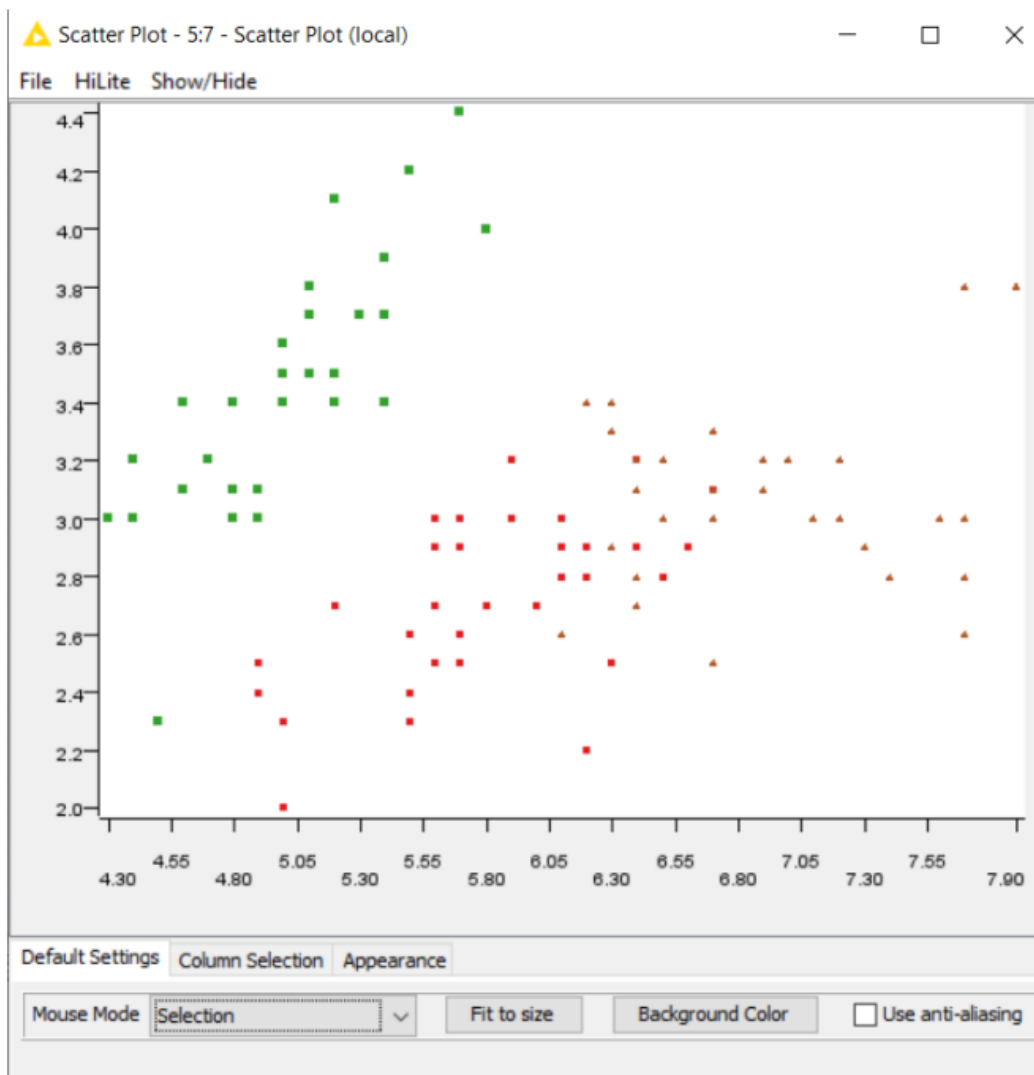
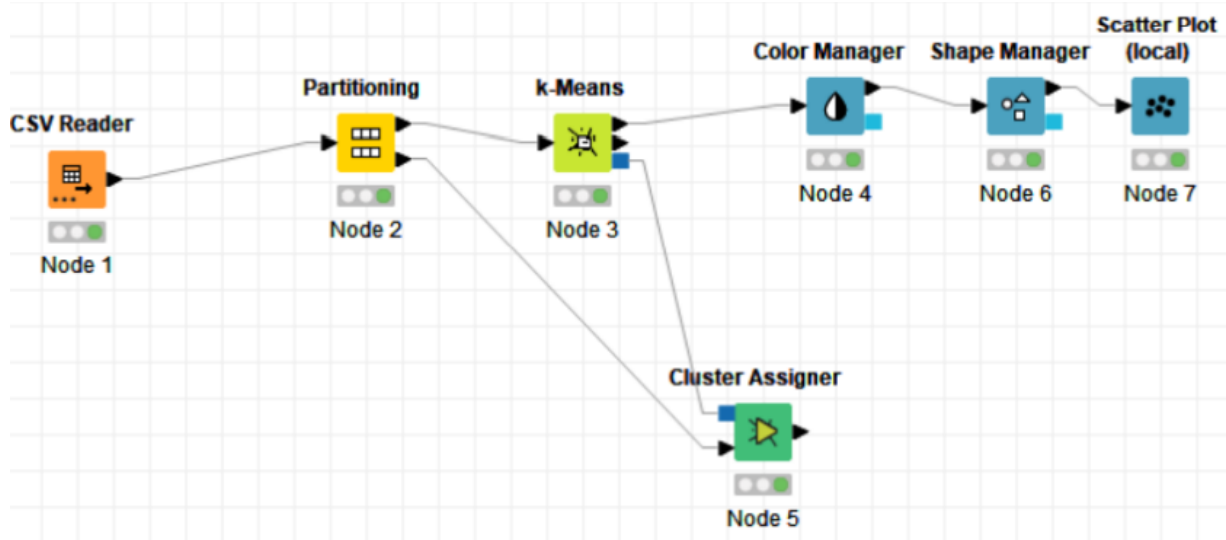
File Edit Hilite Navigation View

Table "spec_name" - Rows: 3 Spec - Columns: 3 Properties Flow Variables

Row ID	I Iris-set...	I Iris-ver...	I Iris-virg...
Iris-setosa	16	0	0
Iris-versicolor	0	15	2
Iris-virginica	0	0	12

The Accuracy statistics and Confusion Matrix can be used as performance metrics

2. CLUSTERING (K MEANS) WORKFLOW



Assigned Data - 5:5 - Cluster Assigner

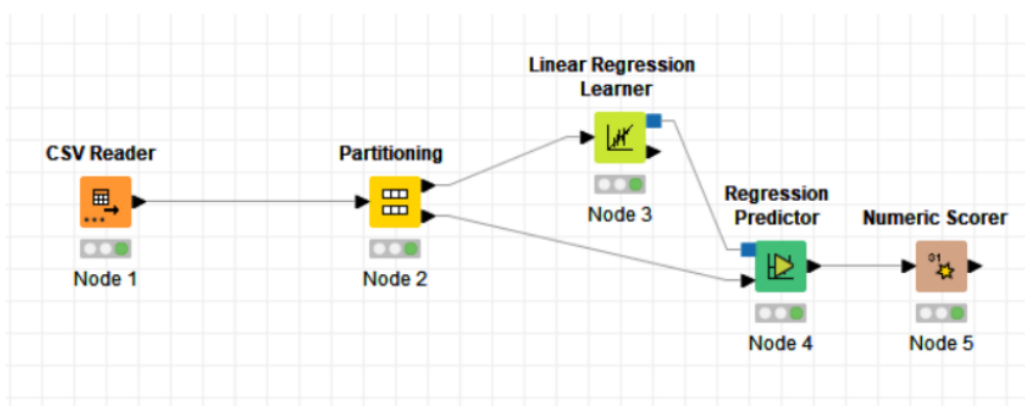
File Edit Hilite Navigation View

Table "default" - Rows: 45 Spec - Columns: 6 Properties Flow Variables

Row ID	D sepal le...	D sepal w...	D petal le...	D petal wi...	S class	S Cluster
Row5	5.4	3.9	1.7	0.4	Iris-setosa	cluster_0
Row8	4.4	2.9	1.4	0.2	Iris-setosa	cluster_0
Row17	5.1	3.5	1.4	0.3	Iris-setosa	cluster_0
Row18	5.7	3.8	1.7	0.3	Iris-setosa	cluster_0
Row22	4.6	3.6	1	0.2	Iris-setosa	cluster_0
Row23	5.1	3.3	1.7	0.5	Iris-setosa	cluster_0
Row25	5	3	1.6	0.2	Iris-setosa	cluster_0
Row29	4.7	3.2	1.6	0.2	Iris-setosa	cluster_0
Row35	5	3.2	1.2	0.2	Iris-setosa	cluster_0
Row36	5.5	3.5	1.3	0.2	Iris-setosa	cluster_0
Row39	5.1	3.4	1.5	0.2	Iris-setosa	cluster_0
Row45	4.8	3	1.4	0.3	Iris-setosa	cluster_0
Row47	4.6	3.2	1.4	0.2	Iris-setosa	cluster_0
Row49	5	3.3	1.4	0.2	Iris-setosa	cluster_0
Row55	5.7	2.8	4.5	1.3	Iris-versicolor	cluster_1
Row56	6.3	3.3	4.7	1.6	Iris-versicolor	cluster_1
Row62	6	2.2	4	1	Iris-versicolor	cluster_1
Row65	6.7	3.1	4.4	1.4	Iris-versicolor	cluster_1
Row66	5.6	3	4.5	1.5	Iris-versicolor	cluster_1
Row72	6.3	2.5	4.9	1.5	Iris-versicolor	cluster_1
Row75	6.6	3	4.4	1.4	Iris-versicolor	cluster_1
Row76	6.8	2.8	4.8	1.4	Iris-versicolor	cluster_1
Row78	6	2.9	4.5	1.5	Iris-versicolor	cluster_1
Row82	5.8	2.7	3.9	1.2	Iris-versicolor	cluster_1
Row84	5.4	3	4.5	1.5	Iris-versicolor	cluster_1
Row85	6	3.4	4.5	1.6	Iris-versicolor	cluster_1
Row87	6.3	2.3	4.4	1.3	Iris-versicolor	cluster_1
Row89	5.5	2.5	4	1.3	Iris-versicolor	cluster_1
Row92	5.8	2.6	4	1.2	Iris-versicolor	cluster_1
Row98	5.1	2.5	3	1.1	Iris-versicolor	cluster_1
Row99	5.7	2.8	4.1	1.3	Iris-versicolor	cluster_1
Row101	5.8	2.7	5.1	1.9	Iris-virginica	cluster_1
Row109	7.2	3.6	6.1	2.5	Iris-virginica	cluster_2
Row112	6.8	3	5.5	2.1	Iris-virginica	cluster_2
Row114	5.8	2.8	5.1	2.4	Iris-virginica	cluster_1
Row116	6.5	3	5.5	1.8	Iris-virginica	cluster_2
Row119	6	2.2	5	1.5	Iris-virginica	cluster_1
Row121	5.6	2.8	4.9	2	Iris-virginica	cluster_1

The scatter plot can be used as a performance metric

3. REGRESSION (LINEAR REGRESSION) WORKFLOW



ID	petal width
R^2	-1.3215412580169095
mean absolute error	2.56
mean squared error	7.694999999999999
root mean squared error	2.773986301336039
mean signed difference	-2.56
mean absolute percentage error	0.7108507897254764

The different types of errors calculated can be used as a performance metric

File

Statistics on Linear Regression

Variable	Coeff.	Std. Err.	t-value	P> t
sepal length	-0.1244	0.0755	-1.6489	0.1029
sepal width	-0.0169	0.0755	-0.2236	0.8236
petal length	0.2462	0.071	3.47	0.0008
petal width	0.58	0.1221	4.7487	8.22E-6
Intercept	1.1419	0.259	4.4091	3.02E-5

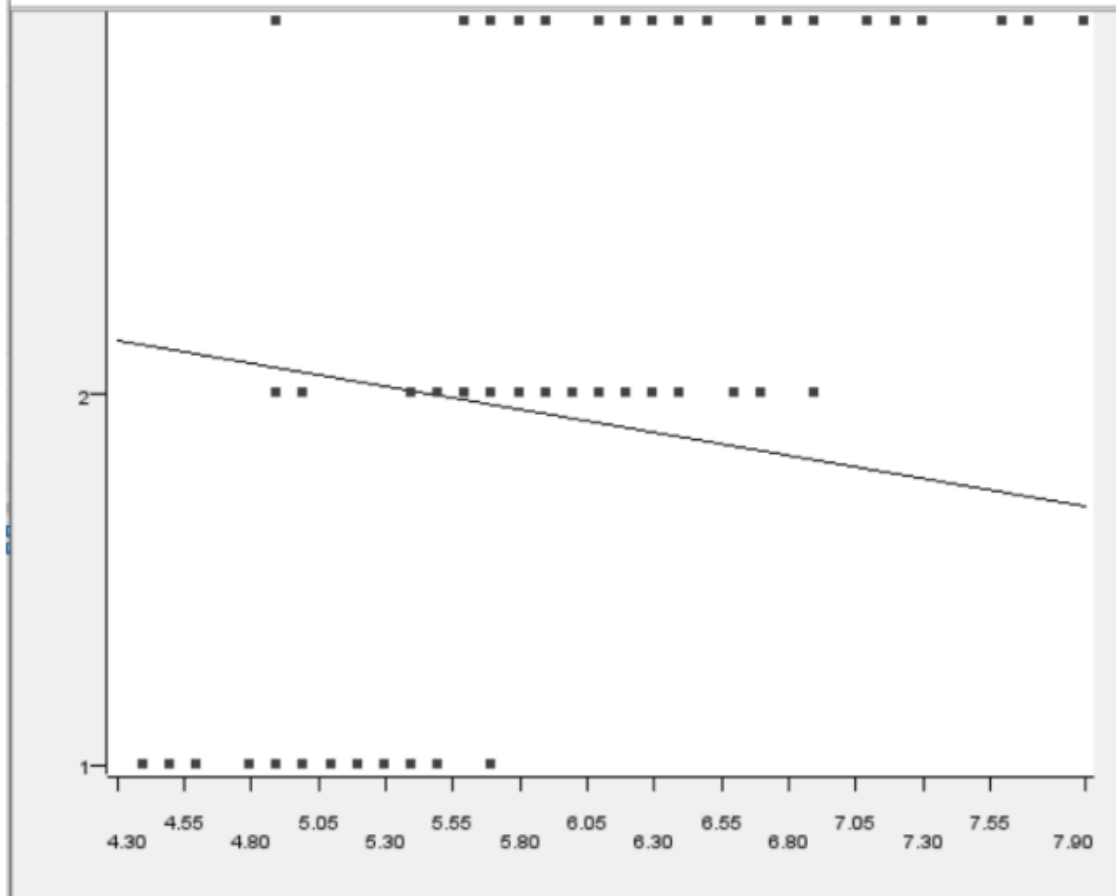
Multiple R-Squared: 0.9243

Adjusted R-Squared: 0.9207

Linear Regression Scatterplot View - 6:3 - Linear Regression Learner

— □ ×

File HiLite



Default Settings Column Selection Appearance

Mouse Mode Selection



Fit to size

Background Color

☐ Use anti-aliasing