



# Breast Cancer Prediction using Logistic Regression

## Step 1: Importing the Data

Altair AI Studio Educational 2024.0.0 @ Acer\_Nitro-5

File Edit Process View Connections Settings Extensions Help

Repository

Import Data

Specify your data format

Header Row: 1

File Encoding: windows-1252

Use Quotes: ☒

Skip Comments: ☒

Trim Lines: ☒

Column Separator: Comma ','

Decimal Character: .

Multiline Text: ☐

|    | id       | diagnosis | radius_mean | texture_mean | perimeter_mean | area_mean | smoothness_mean | compactness_mean | concavity_mean | concave points_mean | symmetry_mean | fractal_dimension_mean | radius_worst |
|----|----------|-----------|-------------|--------------|----------------|-----------|-----------------|------------------|----------------|---------------------|---------------|------------------------|--------------|
| 1  | 842302   | M         | 17.99       | 10.38        | 122.80         | 1001      | 0.1184          | 0.2776           | 0.3001         | 0.1471              | 0.2419        | 0.07871                | 1.0          |
| 2  | 842517   | M         | 20.57       | 17.77        | 132.90         | 1326      | 0.08474         | 0.07864          | 0.0869         | 0.07017             | 0.1812        | 0.05667                | 0.5          |
| 3  | 84300903 | M         | 19.69       | 21.25        | 130            | 1203      | 0.1096          | 0.1599           | 0.1974         | 0.1279              | 0.2069        | 0.05999                | 0.7          |
| 4  | 84348301 | M         | 11.42       | 20.38        | 77.58          | 386.1     | 0.1425          | 0.2839           | 0.2414         | 0.1052              | 0.2597        | 0.09744                | 0.4          |
| 5  | 84358402 | M         | 20.29       | 14.34        | 135.1          | 1297      | 0.1003          | 0.1328           | 0.198          | 0.1043              | 0.1809        | 0.05883                | 0.3          |
| 6  | 843786   | M         | 12.45       | 15.7         | 82.57          | 477.1     | 0.1278          | 0.17             | 0.1578         | 0.08089             | 0.2087        | 0.07613                | 0.7          |
| 7  | 844359   | M         | 18.25       | 19.98        | 119.6          | 1040      | 0.09463         | 0.109            | 0.1127         | 0.074               | 0.1794        | 0.05742                | 0.4          |
| 8  | 84458202 | M         | 13.71       | 20.83        | 90.2           | 577.9     | 0.1189          | 0.1645           | 0.09306        | 0.05985             | 0.2196        | 0.07451                | 0.5          |
| 9  | 844981   | M         | 13          | 21.82        | 87.5           | 519.8     | 0.1273          | 0.1932           | 0.1859         | 0.09353             | 0.235         | 0.07389                | 0.3          |
| 10 | 84501001 | M         | 12.46       | 24.04        | 83.97          | 475.9     | 0.1186          | 0.2396           | 0.2273         | 0.08543             | 0.203         | 0.08243                | 0.2          |
| 11 | 845636   | M         | 16.02       | 23.24        | 102.7          | 797.8     | 0.08206         | 0.06669          | 0.03299        | 0.03323             | 0.1528        | 0.05697                | 0.3          |
| 12 | 84610002 | M         | 15.78       | 17.89        | 103.6          | 781       | 0.0971          | 0.1292           | 0.09954        | 0.06066             | 0.1842        | 0.06082                | 0.5          |
| 13 | 846226   | M         | 19.17       | 24.8         | 132.4          | 1123      | 0.0974          | 0.2458           | 0.2065         | 0.1118              | 0.2397        | 0.078                  | 0.9          |
| 14 | 846381   | M         | 15.85       | 23.95        | 103.7          | 782.7     | 0.08401         | 0.1002           | 0.09938        | 0.05364             | 0.1847        | 0.05338                | 0.4          |
| 15 | 84667401 | M         | 13.73       | 22.61        | 93.6           | 578.3     | 0.1131          | 0.2293           | 0.2128         | 0.08025             | 0.2069        | 0.07682                | 0.2          |
| 16 | 84799002 | M         | 14.54       | 27.54        | 96.73          | 658.8     | 0.1139          | 0.1565           | 0.1639         | 0.07364             | 0.2303        | 0.07077                | 0.3          |
| 17 | 848406   | M         | 20.13       | 94.74        | 684.50         | 0.099     | 0.072           | 0.074            | 0.053          |                     |               |                        |              |

no problems.

Previous Next Cancel

Altair AI Studio Educational 2024.0.0 @ Acer\_Nitro-5

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Interactive Analysis

Result History

ExampleSet (/LocalRepository/data/breast-cancer)

Open in: Turbo Prep Auto Model Interactive Analysis

Filter (569 / 569 examples): all

| Row No. | id       | diagnosis | radius_mean | texture_mean | perimeter_mean | area_mean | smoothness_mean | compactness_mean | concavity_mean | concave points_mean | symmetry_mean | fractal_dimension_mean | radius_worst |
|---------|----------|-----------|-------------|--------------|----------------|-----------|-----------------|------------------|----------------|---------------------|---------------|------------------------|--------------|
| 1       | 842302   | M         | 17.990      | 10.380       | 122.800        | 1001      | 0.118           | 0.278            | 0.300          | 0.147               | 0.241         | 0.07871                | 1.0          |
| 2       | 842517   | M         | 20.570      | 17.770       | 132.900        | 1326      | 0.085           | 0.079            | 0.087          | 0.070               | 0.181         | 0.05667                | 0.5          |
| 3       | 84300903 | M         | 19.690      | 21.250       | 130            | 1203      | 0.110           | 0.160            | 0.197          | 0.128               | 0.207         | 0.05999                | 0.7          |
| 4       | 84348301 | M         | 11.420      | 20.380       | 77.580         | 386.100   | 0.142           | 0.284            | 0.241          | 0.105               | 0.260         | 0.09744                | 0.4          |
| 5       | 84358402 | M         | 20.290      | 14.340       | 135.100        | 1297      | 0.100           | 0.133            | 0.198          | 0.104               | 0.181         | 0.05883                | 0.3          |
| 6       | 843786   | M         | 12.450      | 15.700       | 82.570         | 477.100   | 0.128           | 0.170            | 0.158          | 0.081               | 0.209         | 0.07613                | 0.7          |
| 7       | 844359   | M         | 18.250      | 19.980       | 119.600        | 1040      | 0.095           | 0.109            | 0.113          | 0.074               | 0.179         | 0.05742                | 0.4          |
| 8       | 84458202 | M         | 13.710      | 20.830       | 90.200         | 577.900   | 0.119           | 0.165            | 0.094          | 0.060               | 0.220         | 0.07451                | 0.5          |
| 9       | 844981   | M         | 13          | 21.820       | 87.500         | 519.800   | 0.127           | 0.193            | 0.186          | 0.094               | 0.235         | 0.07389                | 0.3          |
| 10      | 84501001 | M         | 12.460      | 24.040       | 83.970         | 475.900   | 0.119           | 0.240            | 0.227          | 0.085               | 0.203         | 0.08243                | 0.2          |
| 11      | 845636   | M         | 16.020      | 23.240       | 102.700        | 797.800   | 0.082           | 0.067            | 0.033          | 0.033               | 0.153         | 0.05697                | 0.3          |
| 12      | 84610002 | M         | 15.780      | 17.890       | 103.600        | 781       | 0.097           | 0.129            | 0.100          | 0.060               | 0.184         | 0.06082                | 0.5          |
| 13      | 846226   | M         | 19.170      | 24.800       | 132.400        | 1123      | 0.097           | 0.246            | 0.206          | 0.112               | 0.240         | 0.078                  | 0.9          |
| 14      | 846381   | M         | 15.850      | 23.950       | 103.700        | 782.700   | 0.084           | 0.100            | 0.099          | 0.054               | 0.185         | 0.05338                | 0.4          |
| 15      | 84667401 | M         | 13.730      | 22.610       | 93.600         | 578.300   | 0.113           | 0.229            | 0.213          | 0.080               | 0.207         | 0.07682                | 0.2          |
| 16      | 84799002 | M         | 14.540      | 27.540       | 96.730         | 658.800   | 0.114           | 0.160            | 0.164          | 0.074               | 0.230         | 0.07077                | 0.3          |
| 17      | 848406   | M         | 14.680      | 20.130       | 94.740         | 684.500   | 0.099           | 0.072            | 0.074          | 0.053               |               |                        |              |

ExampleSet (569 examples, 0 special attributes, 32 regular attributes)

Repository

Import Data

Training Resources (connected)

Samples

Community Samples (connected)

Local Repository (Local)

Connections

data

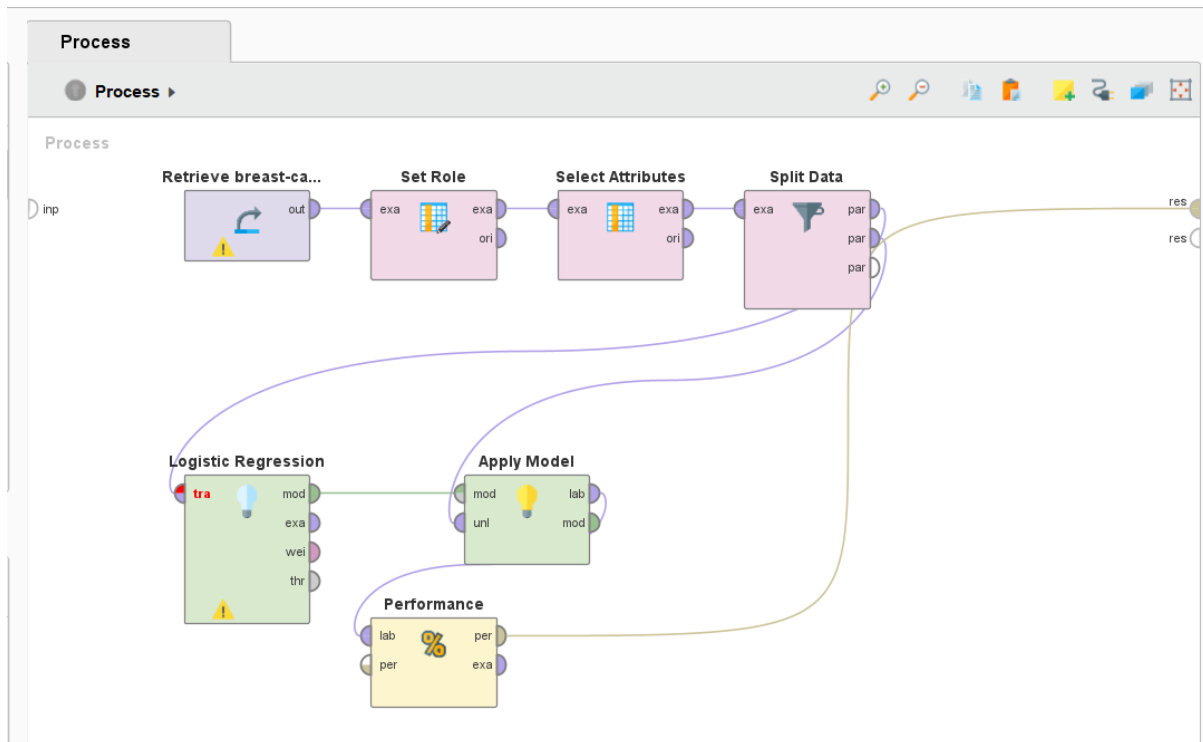
breast-cancer (9/30/24 12:53 PM - 152 KB)

processes

DB (Legacy)



## Step 2: Logistic Regression



Result History

PerformanceVector (Performance) x ExampleSet (/Local Repository/data/breast-cancer) x

Table View Plot View

accuracy: 91.81%

|              | true M | true B | class precision |
|--------------|--------|--------|-----------------|
| pred. M      | 57     | 7      | 89.06%          |
| pred. B      | 7      | 100    | 93.46%          |
| class recall | 89.06% | 93.46% |                 |

Criterion

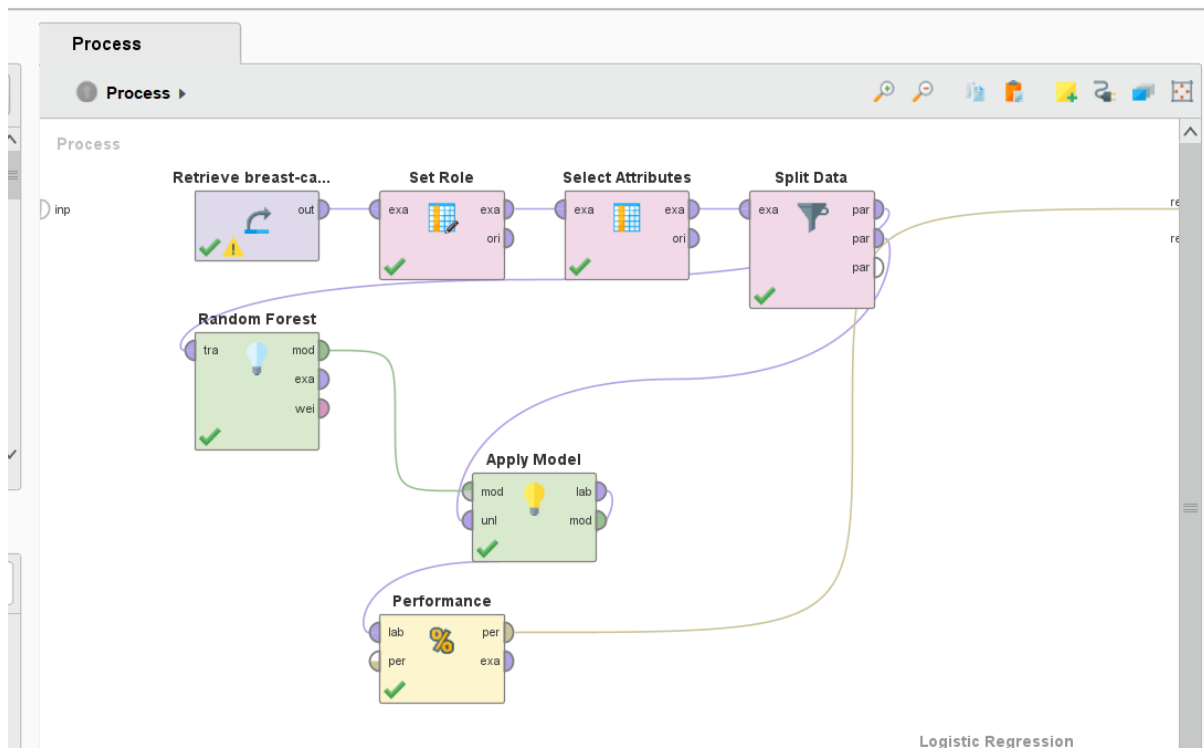
- accuracy
- precision
- recall
- AUC (optimistic)
- AUC
- AUC (pessimistic)

Description

Annotations



### Step 3: Random Forest



Result History

PerformanceVector (Performance) ExampleSet (/Local Repository/data/breast-cancer)

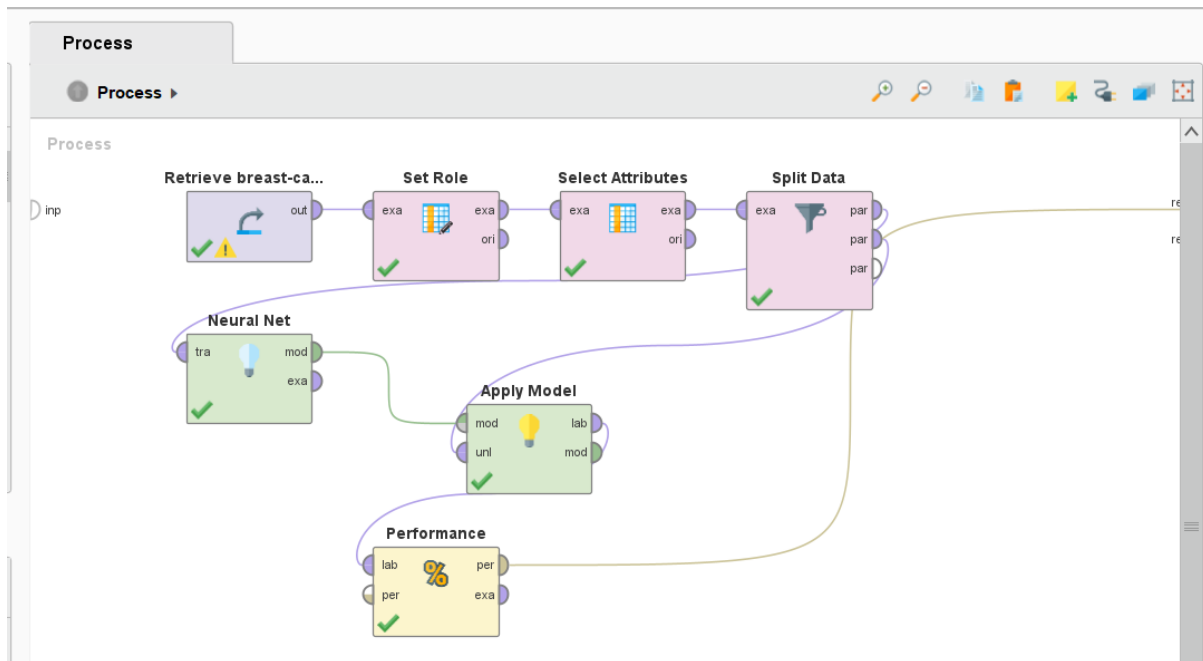
Table View Plot View

accuracy: 95.91%

|              | true M | true B | class precision |
|--------------|--------|--------|-----------------|
| pred. M      | 59     | 2      | 96.72%          |
| pred. B      | 5      | 105    | 95.45%          |
| class recall | 92.19% | 98.13% |                 |



## Step 4: Neural Net



Result History

PerformanceVector (Performance) ExampleSet (//Local Repository/data/breast-cancer)

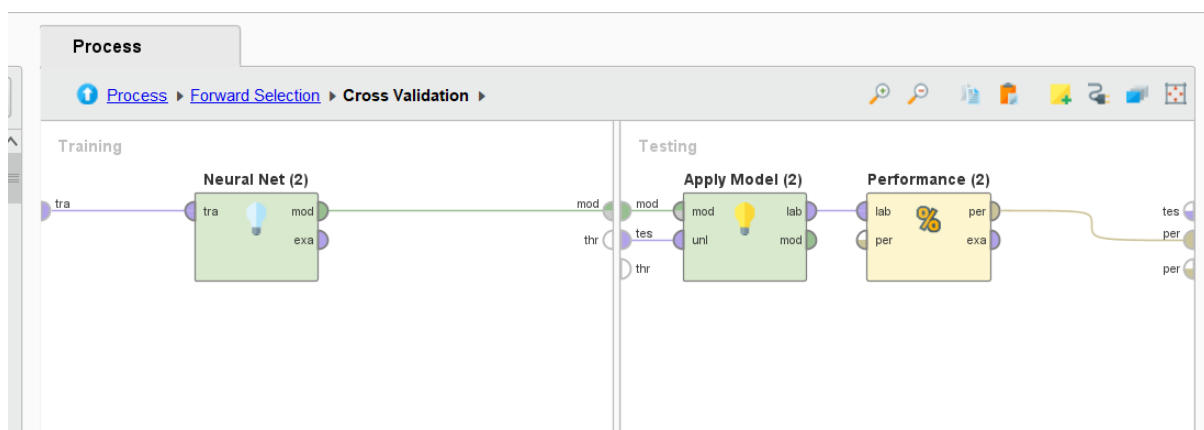
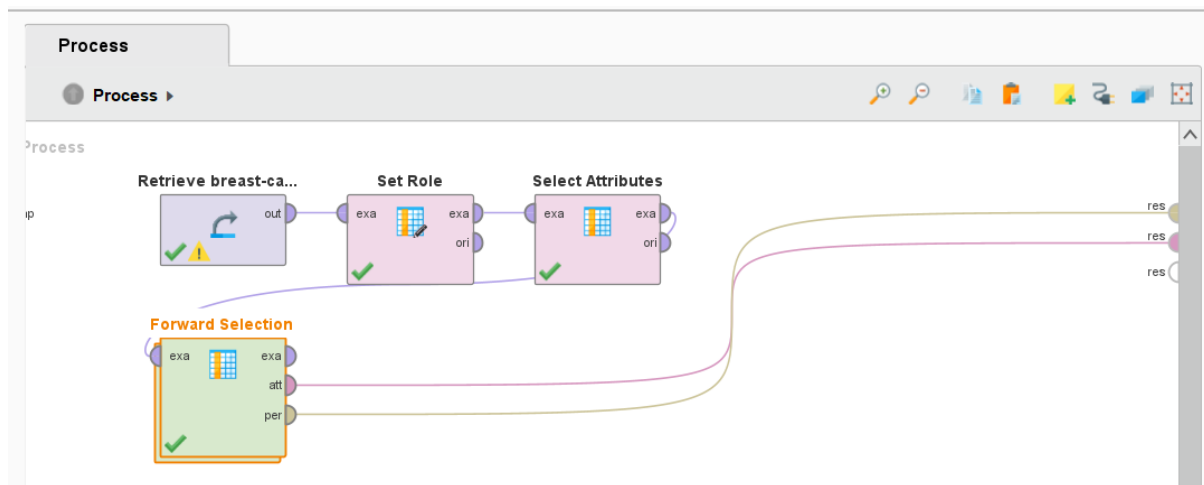
Table View Plot View

accuracy: 96.49%

|              | true M | true B  | class precision |
|--------------|--------|---------|-----------------|
| pred_M       | 58     | 0       | 100.00%         |
| pred_B       | 6      | 107     | 94.69%          |
| class recall | 90.62% | 100.00% |                 |



## Step 5: Forward Selection





Result History

AttributeWeights (Forward Selection) × PerformanceVector (Performance (2)) × ExampleSet (/Local Repository/data/breast-cancer) ×

Data

Weight Visualizations

Plot view

Annotations

| attribute              | weight ↓ |
|------------------------|----------|
| concavity_mean         | 1        |
| symmetry_se            | 1        |
| radius_worst           | 1        |
| texture_worst          | 1        |
| smoothness_worst       | 1        |
| radius_mean            | 0        |
| texture_mean           | 0        |
| perimeter_mean         | 0        |
| area_mean              | 0        |
| smoothness_mean        | 0        |
| compactness_mean       | 0        |
| concave points_mean    | 0        |
| symmetry_mean          | 0        |
| fractal_dimension_mean | 0        |
| radius_se              | 0        |
| texture_se             | 0        |
| perimeter_se           | 0        |
| area_se                | 0        |

Result History

AttributeWeights (Forward Selection) × PerformanceVector (Performance (2)) × ExampleSet (/Local Repository/data/breast-cancer) ×

Criterion

accuracy

precision

recall

AUC (optimistic)

AUC

AUC (pessimistic)

Description

Annotations

Table View Plot View

accuracy: 98.25% +/- 2.34% (micro average: 98.24%)

|              | true M | true B | class precision |
|--------------|--------|--------|-----------------|
| pred. M      | 206    | 4      | 98.10%          |
| pred. B      | 6      | 353    | 98.33%          |
| class recall | 97.17% | 98.88% |                 |