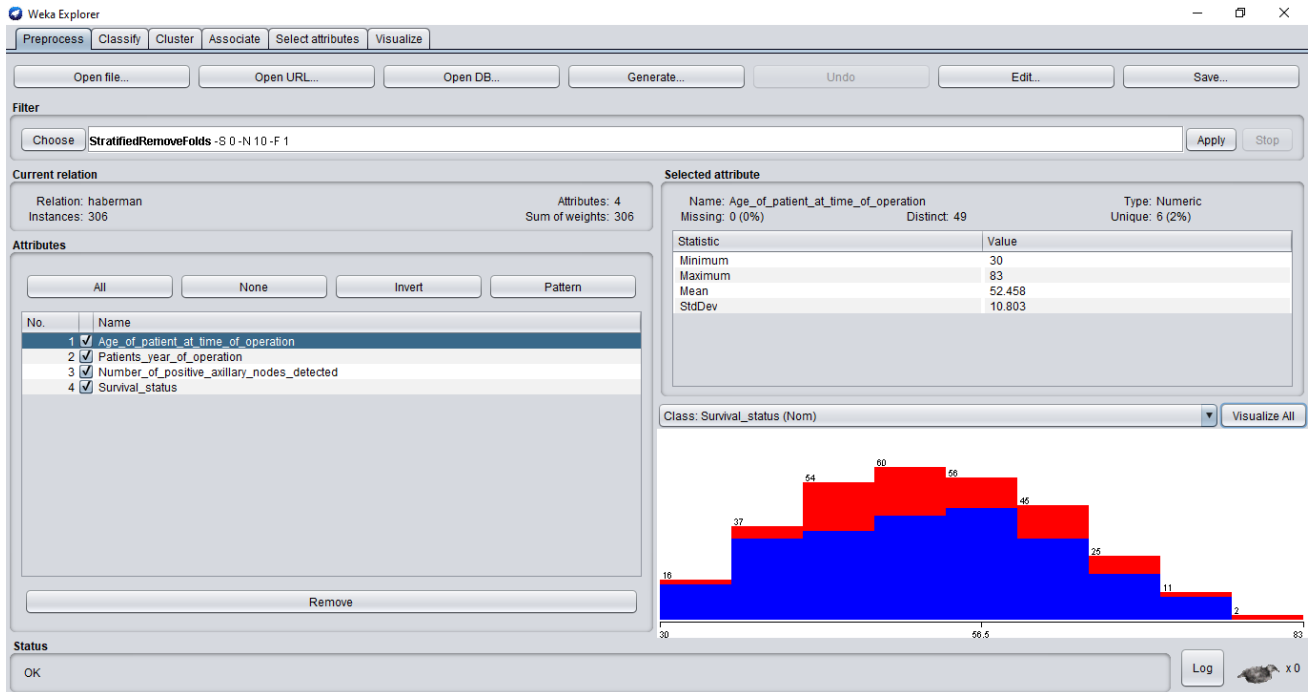
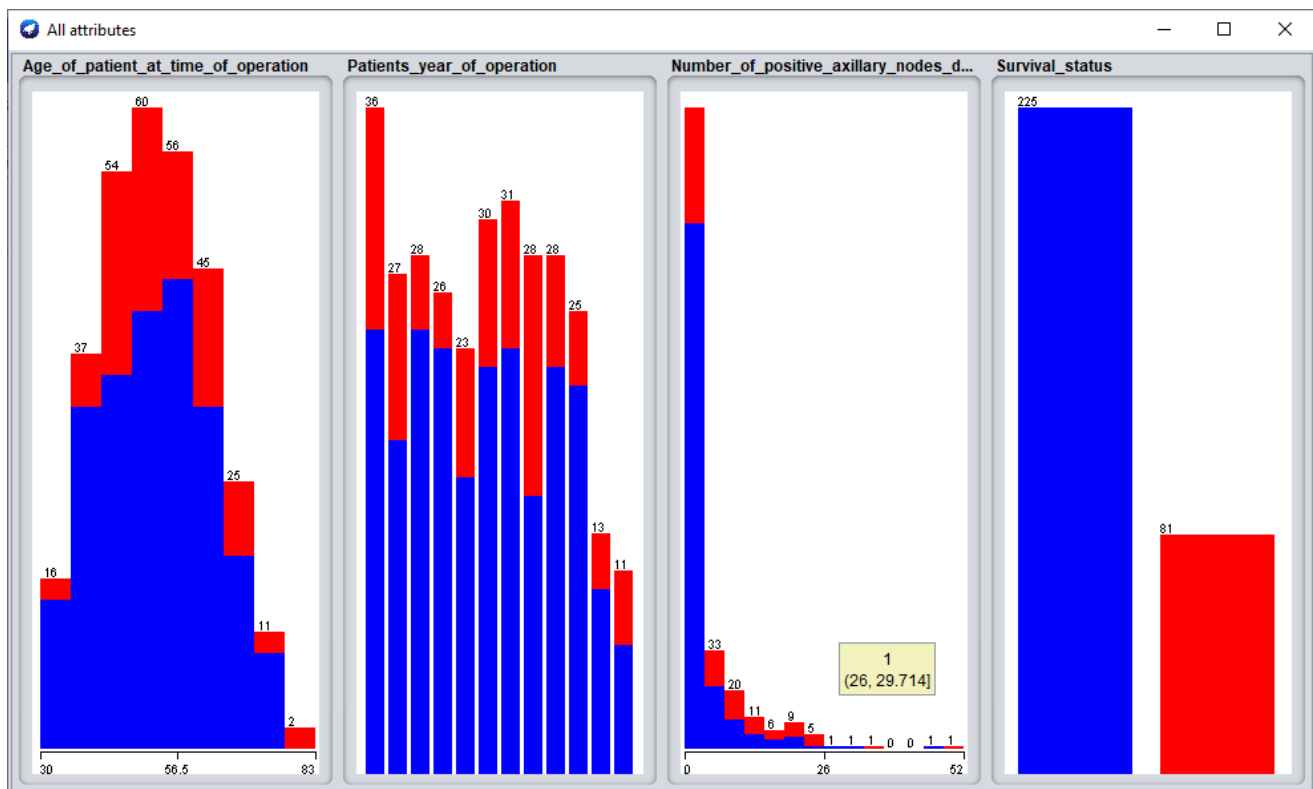


Telas WEKA –Haberman Dataset

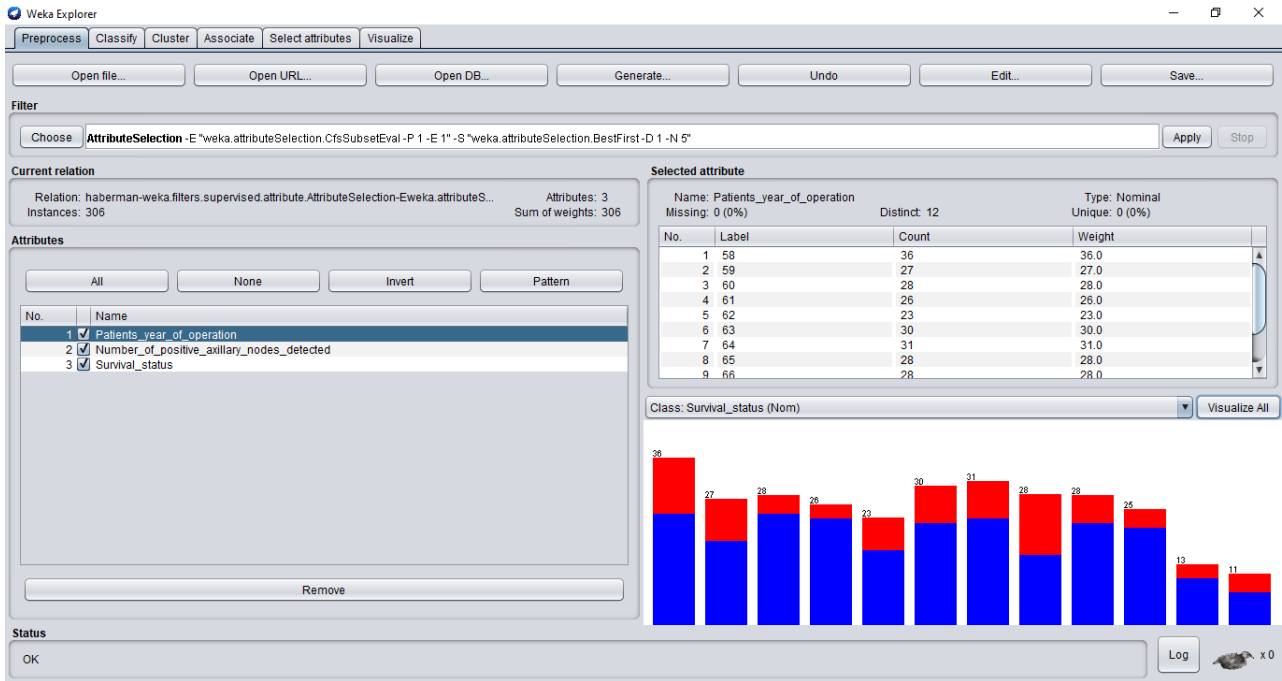
Tela 1 - Pré-processamento, sem a seleção de recursos:



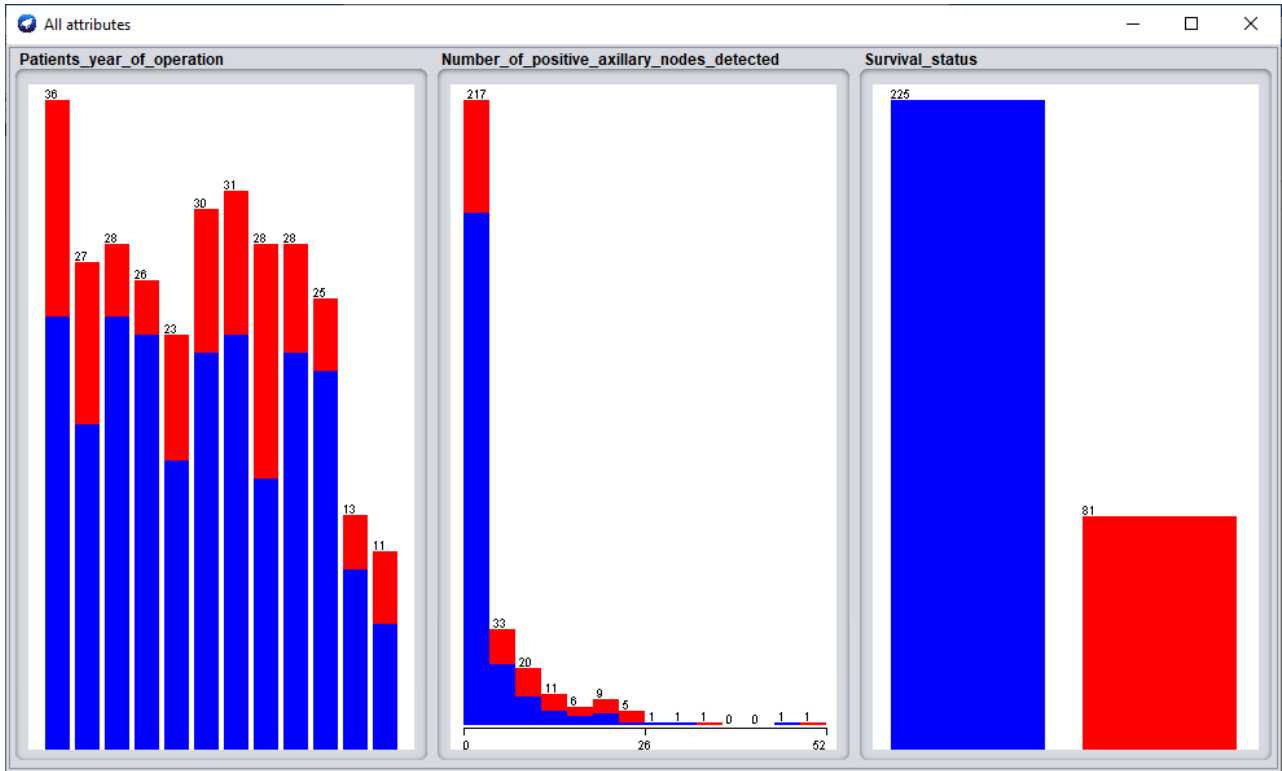
Tela 2 – Atributos, sem a SR:



Tela 3 - Pré-Processamento, Haberman com a seleção de recursos:



Tela – 4: Atributos mais importantes, após a seleção de recursos:



Tela 5 – KNN sem a seleção de Recursos (IBK_noSR.txt):

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'RandomForest' with parameters '-P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1'. The 'Test options' section shows 'Cross-validation' with 'Folds' set to 10. The 'Result list' on the left shows several models, with '20:03:31 - lazy.IBK' selected. The 'Classifier output' pane displays the following results:

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      209           68.3007 %
Incorrectly Classified Instances    97           31.6993 %
Kappa statistic                    0.1043
Mean absolute error                 0.3333
Root mean squared error             0.5633
Relative absolute error             85.4445 %
Root relative squared error         127.6833 %
Total Number of Instances          306

=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,836	0,741	0,758	0,836	0,785	0,107	0,577	0,785	1
	0,259	0,164	0,362	0,259	0,302	0,107	0,577	0,310	2

```

=== Confusion Matrix ===
 a b  <-- classified as
188 37 | a = 1
 60 21 | b = 2
```

The 'Status' bar at the bottom shows 'OK'.

Tela 6 - KNN com SR (IBK_SR.txt):

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'SGD' with parameters '-F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1'. The 'Test options' section shows 'Cross-validation' with 'Folds' set to 10. The 'Result list' on the left shows several models, with '20:02:48 - lazy.IBK' selected. The 'Classifier output' pane displays the following results:

```
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      220           71.8954 %
Incorrectly Classified Instances    86           28.1046 %
Kappa statistic                    0.1425
Mean absolute error                 0.3474
Root mean squared error             0.5221
Relative absolute error             89.0428 %
Root relative squared error         118.3345 %
Total Number of Instances          306

=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,898	0,778	0,762	0,898	0,824	0,155	0,522	0,768	1
	0,222	0,102	0,439	0,222	0,295	0,155	0,522	0,311	2

```

=== Confusion Matrix ===
 a b  <-- classified as
202 23 | a = 1
 63 18 | b = 2
```

The 'Status' bar at the bottom shows 'OK'.

Tela 7- Naive Bayes sem a SR (NB_noSR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:03:31 - lazy.IBk
- 20:08:52 - bayes.NaiveBayes
- 20:10:17 - trees.J48
- 20:13:54 - trees.RandomForest
- 20:15:10 - functions.MultilayerPerceptron
- 20:16:39 - functions.SGD

Classifier output

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	233	76.1438 %
Incorrectly Classified Instances	73	23.8562 %
Kappa statistic	0.2325	
Mean absolute error	0.3277	
Root mean squared error	0.4349	
Relative absolute error	84.0093 %	
Root relative squared error	98.5787 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0,951	0,765	0,775	0,951	0,854	0,275	0,654	0,824	1
	0,235	0,049	0,633	0,235	0,342	0,275	0,654	0,437	2
Weighted Avg.	0,761	0,576	0,738	0,761	0,719	0,275	0,654	0,721	

=== Confusion Matrix ===

a	b	<-- Classified as	
214	11	a = 1	
62	19	b = 2	

Status

OK Log x 0

Tela 8 - Naive Bayes com SR (NB_SR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose SGD -F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:02:48 - lazy.IBk
- 20:07:54 - bayes.NaiveBayes
- 20:11:37 - trees.J48
- 20:11:58 - trees.RandomForest
- 20:12:43 - functions.MultilayerPerceptron
- 20:13:23 - functions.SGD

Classifier output

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	230	75.1634 %
Incorrectly Classified Instances	76	24.8366 %
Kappa statistic	0.197	
Mean absolute error	0.3296	
Root mean squared error	0.4362	
Relative absolute error	84.4854 %	
Root relative squared error	98.8546 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0,947	0,790	0,769	0,947	0,849	0,236	0,659	0,818	1
	0,210	0,053	0,586	0,210	0,309	0,236	0,659	0,433	2
Weighted Avg.	0,752	0,595	0,721	0,752	0,706	0,236	0,659	0,716	

=== Confusion Matrix ===

a	b	<-- classified as	
213	12	a = 1	
64	17	b = 2	

Status

OK Log x 0

Tela 9 - J48 – sem SR (J48_noSR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **RandomForest** -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set

☒ Cross-validation Folds

☐ Percentage split %

(Nom) Survival_status

Result list (right-click for options)

- 20:03:31 - lazy.IBk
- 20:08:52 - bayes.NaiveBayes
- 20:10:17 - trees.J48
- 20:13:54 - trees.RandomForest
- 20:15:10 - functions.MultilayerPerceptron
- 20:16:39 - functions.SGD

Classifier output

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	223	72.8758 %
Incorrectly Classified Instances	83	27.1242 %
Kappa statistic	0.1526	
Mean absolute error	0.3572	
Root mean squared error	0.4624	
Relative absolute error	91.5564 %	
Root relative squared error	104.8123 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,916	0,790	0,763	0,916	0,832	0,172	0,564	0,750	1
	0,210	0,084	0,472	0,210	0,291	0,172	0,564	0,344	2
Weighted Avg.	0,729	0,603	0,686	0,729	0,689	0,172	0,564	0,642	

=== Confusion Matrix ===

```

a  b  <-- classified as
206 19 | a = 1
 64 17 | b = 2

```

Status

OK x 0

Tela 10 – J48 com a SR (J48_SR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **SGD** -F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set

☒ Cross-validation Folds

☐ Percentage split %

(Nom) Survival_status

Result list (right-click for options)

- 20:02:48 - lazy.IBk
- 20:07:54 - bayes.NaiveBayes
- 20:11:37 - trees.J48
- 20:11:58 - trees.RandomForest
- 20:12:43 - functions.MultilayerPerceptron
- 20:13:23 - functions.SGD

Classifier output

Number of Leaves : 14

Size of the tree : 17

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	223	72.8758 %
Incorrectly Classified Instances	83	27.1242 %
Kappa statistic	0.0914	
Mean absolute error	0.3734	
Root mean squared error	0.4487	
Relative absolute error	95.7186 %	
Root relative squared error	101.6975 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,947	0,877	0,750	0,947	0,837	0,120	0,553	0,758	1
	0,123	0,053	0,455	0,123	0,194	0,120	0,553	0,306	2
Weighted Avg.	0,729	0,659	0,672	0,729	0,667	0,120	0,553	0,638	

=== Confusion Matrix ===

```

a  b  <-- classified as
206 19 | a = 1
 64 17 | b = 2

```

Status

OK x 0

Tela 11 – Random Forest sem SR (RF_noSR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **RandomForest** -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set
☐ Supplied test set
☒ Cross-validation Folds
☐ Percentage split %

(Nom) Survival_status

Result list (right-click for options)

- 20:03:31 - lazy.IBk
- 20:08:52 - bayes.NaiveBayes
- 20:10:17 - trees.J48
- 20:13:54 - trees.RandomForest**
- 20:15:10 - functions.MultilayerPerceptron
- 20:16:39 - functions.SGD

Classifier output

Time taken to build model: 0.1 seconds

=== Stratified cross-validation ===
 === Summary ===

Correctly Classified Instances	211	68.9542 %
Incorrectly Classified Instances	95	31.0458 %
Kappa statistic	0.1522	
Mean absolute error	0.3379	
Root mean squared error	0.4711	
Relative absolute error	86.6255 %	
Root relative squared error	106.7699 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,822	0,679	0,771	0,822	0,796	0,154	0,646	0,832	1
	0,321	0,178	0,394	0,321	0,354	0,154	0,646	0,366	2
Weighted Avg.	0,690	0,546	0,671	0,690	0,679	0,154	0,646	0,709	

=== Confusion Matrix ===

a	b	-- classified as
185	40	a = 1
55	26	b = 2

Status

OK x 0

Tela 12 – Random Forest com SR (RF_SR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **SGD** -F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1

Test options

☐ Use training set
☐ Supplied test set
☒ Cross-validation Folds
☐ Percentage split %

(Nom) Survival_status

Result list (right-click for options)

- 20:02:48 - lazy.IBk
- 20:07:54 - bayes.NaiveBayes
- 20:11:37 - trees.J48
- 20:11:58 - trees.RandomForest**
- 20:12:43 - functions.MultilayerPerceptron
- 20:13:23 - functions.SGD

Classifier output

Time taken to build model: 0.08 seconds

=== Stratified cross-validation ===
 === Summary ===

Correctly Classified Instances	220	71.8954 %
Incorrectly Classified Instances	86	28.1046 %
Kappa statistic	0.1505	
Mean absolute error	0.3533	
Root mean squared error	0.4816	
Relative absolute error	90.5555 %	
Root relative squared error	109.1671 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0,893	0,765	0,764	0,893	0,824	0,162	0,563	0,776	1
	0,235	0,107	0,442	0,235	0,306	0,162	0,563	0,322	2
Weighted Avg.	0,719	0,591	0,679	0,719	0,687	0,162	0,563	0,656	

=== Confusion Matrix ===

a	b	-- classified as
201	24	a = 1
62	19	b = 2

Status

OK x 0

Tela 13 – MLP sem SR (MLP_noSR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:03:31 - lazy.IBk
- 20:08:52 - bayes.NaiveBayes
- 20:10:17 - trees.J48
- 20:13:54 - trees.RandomForest
- 20:15:10 - functions.MultilayerPerceptron
- 20:16:39 - functions.SGD

Classifier output

Time taken to build model: 0.32 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	212	69.281 %
Incorrectly Classified Instances	94	30.719 %
Kappa statistic	0.1576	
Mean absolute error	0.3366	
Root mean squared error	0.4705	
Relative absolute error	86.2786 %	
Root relative squared error	106.6486 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,627	0,679	0,772	0,827	0,798	0,159	0,649	0,830	1
	0,321	0,173	0,400	0,321	0,356	0,159	0,649	0,390	2

=== Confusion Matrix ===

a	b	<-- classified as
186	39	a = 1
55	26	b = 2

Status

OK Log x 0

Tela 14 – MLP com SR (MLP_SR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose SGD -F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:02:48 - lazy.IBk
- 20:07:54 - bayes.NaiveBayes
- 20:11:37 - trees.J48
- 20:11:58 - trees.RandomForest
- 20:12:43 - functions.MultilayerPerceptron
- 20:13:23 - functions.SGD

Classifier output

Time taken to build model: 0.28 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	220	71.8954 %
Incorrectly Classified Instances	86	28.1046 %
Kappa statistic	0.1177	
Mean absolute error	0.3589	
Root mean squared error	0.4497	
Relative absolute error	92.0114 %	
Root relative squared error	101.93 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,911	0,815	0,756	0,911	0,827	0,133	0,613	0,790	1
	0,185	0,089	0,429	0,185	0,259	0,133	0,613	0,362	2

=== Confusion Matrix ===

a	b	<-- classified as
205	20	a = 1
66	15	b = 2

Status

OK Log x 0

Tela 15 – SGD sem SR (SGD_noSR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:03:31 - lazy.IBk
- 20:08:52 - bayes.NaiveBayes
- 20:10:17 - trees.J48
- 20:13:54 - trees.RandomForest
- 20:15:10 - functions.MultilayerPerceptron
- 20:16:39 - functions.SGD

Classifier output

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	218	71.2418 %
Incorrectly Classified Instances	88	28.7582 %
Kappa statistic	0.0416	
Mean absolute error	0.2876	
Root mean squared error	0.5363	
Relative absolute error	73.7212 %	
Root relative squared error	121.5462 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,933	0,901	0,742	0,933	0,827	0,054	0,516	0,742	1
	0,099	0,067	0,348	0,099	0,154	0,054	0,516	0,273	2

=== Confusion Matrix ===

a	b	<-- classified as
210	15	a = 1
73	8	b = 2

Status

OK Log x0

Tela 16 – SGD com SR (SGD_SR.txt):

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose SGD -F 0 -L 0.01 -R 1.0E-4 -E 500 -C 0.001 -S 1

Test options

☐ Use training set

☐ Supplied test set Set...

☒ Cross-validation Folds 10

☐ Percentage split % 66

More options...

(Nom) Survival_status

Start Stop

Result list (right-click for options)

- 20:02:48 - lazy.IBk
- 20:07:54 - bayes.NaiveBayes
- 20:11:37 - trees.J48
- 20:11:58 - trees.RandomForest
- 20:12:43 - functions.MultilayerPerceptron
- 20:13:23 - functions.SGD

Classifier output

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	218	71.2418 %
Incorrectly Classified Instances	88	28.7582 %
Kappa statistic	0.0317	
Mean absolute error	0.2876	
Root mean squared error	0.5363	
Relative absolute error	73.7212 %	
Root relative squared error	121.5462 %	
Total Number of Instances	306	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0,938	0,914	0,740	0,938	0,827	0,042	0,512	0,740	1
	0,086	0,062	0,333	0,086	0,137	0,042	0,512	0,271	2

=== Confusion Matrix ===

a	b	<-- classified as
211	14	a = 1
74	7	b = 2

Status

OK Log x0