

Tutorial Business Analytics

Exercise 8

Exercise 8.1

Create a 3-Fold Cross-validation for the specified instances

- Partition the data set into complementary subsets
- Decide which subsets will be used for training and which for testing

Additionally, create a *stratified* 3-Fold Cross-validation for specified instances

Instance	Class
1	+
2	+
3	+
4	+
5	-
6	+
7	+
8	-
9	-
10	-
11	+
12	+
13	+
14	-
15	-

Exercise 8.2

True Class	Predicted Class
0	0
0	1
1	1
1	0
0	0
1	0
0	0
1	1
0	1
1	0

Calculate Recall, False Alarm Rate, Specificity and Accuracy.

Exercise 8.3

The table below contains 3 classifier's **accuracy values**. Evaluate whether the results obtained by the new classifiers 1 and 2 are significantly different from the baseline classifier 0 (two-sided test, significance level 5%).

Classifier 0	Classifier 1	Δ	Classifier 2	Δ
0.67	0.98	-0.31	0.67	0.00
0.63	0.91	-0.28	0.69	-0.06
0.95	0.93	0.02	0.90	0.05
0.75	0.86	-0.11	0.91	-0.16
0.75	0.95	-0.20	0.86	-0.11
0.79	0.85	-0.06	0.75	0.04
0.79	0.90	-0.11	0.68	0.11
0.82	0.87	-0.05	0.83	-0.01
0.83	0.98	-0.15	0.82	0.01
0.78	0.91	-0.13	0.80	-0.02

$$t = \frac{\bar{d}}{s_d / \sqrt{k}} \sim t_{k-1}$$

$$\bar{d} = \frac{1}{k} \sum_i d_i$$

$$s_d = \sqrt{\frac{1}{k-1} \sum_i (d_i - \bar{d})^2}$$

Exercise 8.4

Use the given result of an evaluation (Cutoff = 0.87) to construct:

- a gain curve (10% steps)
- a lift curve
- an ROC curve

Remember: A Cutoff value of 0.87 means, we will classify an instance as positive until its probability falls under 0.87

Number	Probability	Class
1	0.991	+
2	0.977	+
3	0.973	+
4	0.945	+
5	0.918	+
6	0.915	-
7	0.906	+
8	0.889	-
9	0.873	+
10	0.871	+
11	0.869	-
12	0.866	-
13	0.862	+
14	0.852	-
15	0.837	+
16	0.831	-
17	0.829	-
18	0.811	-
19	0.787	-
20	0.779	-

Annex

t-table

q	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807