

Exercise on Lift charts

- We are given a test set with the real labels

X	Y	Z	Class
7	8	45	Yes
30	8	40	No
13	23	21	No
47	43	34	No
37	10	29	Yes
19	49	31	No
20	13	8	Yes
33	44	16	Yes
47	12	41	No
49	21	3	Yes

- Our model provides the following predictions and associated confidences
- Plot the corresponding Lift chart

X	Y	Z	Class	Predicted Class	Confidence of Prediction
7	8	45	Yes	No	0.5618863452
30	8	40	No	Yes	0.6701976614
13	23	21	No	No	0.6816996196
47	43	34	No	No	0.8795983369
37	10	29	Yes	Yes	0.8245785853
19	49	31	No	Yes	0.8194210517
20	13	8	Yes	Yes	0.5079911998
33	44	16	Yes	No	0.5736005213
47	12	41	No	No	0.8702045378
49	21	3	Yes	Yes	0.7856012356

- Step 1: focus on relevant information

Class	Predicted Class	Confidence of Prediction
Yes	No	0.5618863452
No	Yes	0.6701976614
No	No	0.6816996196
No	No	0.8795983369
Yes	Yes	0.8245785853
No	Yes	0.8194210517
Yes	Yes	0.5079911998
Yes	No	0.5736005213
No	No	0.8702045378
Yes	Yes	0.7856012356

- Step 2: from prediction and confidence, derive the score
 - Score = probability of having positive

Class	Predicted Class	Confidence of Prediction	Score
Yes	No	0.5618863452	0.4381136548
No	Yes	0.6701976614	0.6701976614
No	No	0.6816996196	0.3183003804
No	No	0.8795983369	0.1204016631
Yes	Yes	0.8245785853	0.8245785853
No	Yes	0.8194210517	0.8194210517
Yes	Yes	0.5079911998	0.5079911998
Yes	No	0.5736005213	0.4263994787
No	No	0.8702045378	0.1297954622
Yes	Yes	0.7856012356	0.7856012356

- Sort records according to score
 - Descending order (most likely positives first)

Sorted Class	Sorted Scores
Yes	0.8245785853
No	0.8194210517
Yes	0.7856012356
No	0.6701976614
Yes	0.5079911998
Yes	0.4381136548
Yes	0.4263994787
No	0.3183003804
No	0.1297954622
No	0.1204016631

- Evaluate, for each possible threshold, how many true positives we capture

Sorted Class	Sorted Scores	True Positives
		0
Yes	0.8245785853	
No	0.8194210517	
Yes	0.7856012356	
No	0.6701976614	
Yes	0.5079911998	
Yes	0.4381136548	
Yes	0.4263994787	
No	0.3183003804	
No	0.1297954622	
No	0.1204016631	

- Evaluate, for each possible threshold, how many true positives we capture

Sorted Class	Sorted Scores	True Positives
		0
Yes	0.8245785853	1
No	0.8194210517	1
Yes	0.7856012356	2
No	0.6701976614	2
Yes	0.5079911998	3
Yes	0.4381136548	4
Yes	0.4263994787	5
No	0.3183003804	5
No	0.1297954622	5
No	0.1204016631	5

- Plot

Sorted Class	Sorted Scores	True Positives
		0
Yes	0.8245785853	1
No	0.8194210517	1
Yes	0.7856012356	2
No	0.6701976614	2
Yes	0.5079911998	3
Yes	0.4381136548	4
Yes	0.4263994787	5
No	0.3183003804	5
No	0.1297954622	5
No	0.1204016631	5

