

7.3 Logistic Regression - Model Building

Tuesday, 08 November 2022 17:12

Summary	<ul style="list-style-type: none">Calculations: going from probabilities (real numbers) to 0 – 1 format.																																																										
<ul style="list-style-type: none">How do you go from probabilities to 0 – 1 format?	<h2>Regression: Calculations explained</h2> <ul style="list-style-type: none">Notice that this regression will give us the “forecasted” probability that the student will be placed ($Pr\{Y=1\}$).However, we want the value for our response variable (Y). And not the probability that Y will take on certain value.Towards that, we define a threshold probability – if the forecasted probability value is above the threshold, we say that $Y = 1$.On the other hand, if the forecasted probability is below the threshold, we can say that $Y = 0$ (the student won’t be placed).Now we have probabilities in hand, but we want to make the predictions in 0 – 1 format. To do that, we set a threshold on probabilities.																																																										
	<table><tr><td>Cutoff</td><td>0.5</td></tr></table>	Cutoff	0.5	So, we defined an arbitrary cut-off at 0.5																																																							
Cutoff	0.5																																																										
	<table><tr><th>Prob of Day-0 job</th><th>Predicted Y Classification</th></tr><tr><td>0.975166751</td><td>1</td></tr><tr><td>0.918065973</td><td>1</td></tr><tr><td>0.00457583</td><td>0</td></tr><tr><td>0.370317052</td><td>0</td></tr><tr><td>0.775330804</td><td>1</td></tr></table>	Prob of Day-0 job	Predicted Y Classification	0.975166751	1	0.918065973	1	0.00457583	0	0.370317052	0	0.775330804	1	<table><tr><td>If Prob of Day-0 job > 0.5</td></tr><tr><td> Predicted $Y = 1$</td></tr><tr><td>else</td></tr><tr><td> Predicted $Y = 0$</td></tr></table>		If Prob of Day-0 job > 0.5	Predicted $Y = 1$	else	Predicted $Y = 0$																																								
Prob of Day-0 job	Predicted Y Classification																																																										
0.975166751	1																																																										
0.918065973	1																																																										
0.00457583	0																																																										
0.370317052	0																																																										
0.775330804	1																																																										
If Prob of Day-0 job > 0.5																																																											
Predicted $Y = 1$																																																											
else																																																											
Predicted $Y = 0$																																																											
	<p>That lead us to a result that has some errors:</p> <table><tr><th>Observed Y</th><th>Predicted Y</th></tr><tr><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>1</td></tr></table>			Observed Y	Predicted Y	1	1	1	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	1
Observed Y	Predicted Y																																																										
1	1																																																										
1	1																																																										
0	0																																																										
0	0																																																										
1	1																																																										
0	0																																																										
0	0																																																										
0	0																																																										
1	0																																																										
1	0																																																										
1	0																																																										
1	1																																																										
1	1																																																										
1	1																																																										
0	0																																																										
0	0																																																										
0	1																																																										
0	0																																																										
0	0																																																										
0	0																																																										
1	1																																																										
0	0																																																										
0	0																																																										
0	0																																																										
1	1																																																										
0	0																																																										
0	1																																																										
	<ul style="list-style-type: none">This happened because we chose arbitrary value for the threshold.																																																										