Log Loss

Problem with AUC-ROC

- Considers only the order of probability

Analytics Vidhya
Learn everything about analytics

Problem with AUC-ROC

- Considers only the order of probability
- Cannot be used for comparing two models

Comparing Models

ID	Actual	Predicted probabilities 1	Predicted probabilities 2
ID6	1	0.94	0.74
ID1	1	0.90	0.69
ID7	1	0.78	0.63
ID8	0	0.56	0.57
ID2	0	0.51	0.51
ID3	1	0.47	0.44
ID4	1	0.32	0.39
ID5	0	0.10	0.35

Comparing Models

ID	Actual	Predicted probabilities 1	Predicted probabilities 2	Predicted class
ID6	1	**0.94**	**0.74**	**1**
ID1	1	0.90	0.69	0
ID7	1	0.78	0.63	0
ID8	0	0.56	0.57	0
ID2	0	0.51	0.51	0
ID3	1	0.47	0.44	0
ID4	1	0.32	0.39	0
ID5	0	0.10	0.35	0

Comparing Models

ID	Actual	Predicted probabilities 1	Predicted probabilities 2	Predicted class
ID6	1	0.94	0.74	1
ID1	1	**0.90**	**0.69**	**0**
ID7	1	0.78	0.63	0
ID8	0	0.56	0.57	0
ID2	0	0.51	0.51	0
ID3	1	0.47	0.44	0
ID4	1	0.32	0.39	0
ID5	0	0.10	0.35	0

Log Loss

It is the negative average of the log of corrected predicted probabilities for each instance.

Corrected Probabilities

ID	Actual	Predicted probabilities
ID6	1	0.94
ID1	1	0.90
ID7	1	0.78
ID8	0	0.56
ID2	0	0.51
ID3	1	0.47
ID4	1	0.32
ID5	0	0.10

Corrected Probabilities

ID	Actual	Predicted probabilities	Corrected Probabilities
ID6	1	0.94	0.94
ID1	1	0.90	0.90
ID7	1	0.78	0.78
ID8	0	**0.56**	**0.44**
ID2	0	**0.51**	**0.49**
ID3	1	0.47	0.47
ID4	1	0.32	0.32
ID5	0	**0.10**	**0.90**

Log Loss

It is the negative average of the log of corrected predicted probabilities for each instance.

$$\log(\ p_i\ )$$

Corrected Probabilities

ID	Actual	Predicted probabilities	Corrected Probabilities	Log
ID6	1	0.94	0.94	-0.0268721464
ID1	1	0.90	0.90	-0.0457574906
ID7	1	0.78	0.78	-0.1079053973
ID8	0	0.56	0.44	-0.3565473235
ID2	0	0.51	0.49	-0.30980392
ID3	1	0.47	0.47	-0.3279021421
ID4	1	0.32	0.32	-0.4948500217
ID5	0	0.10	0.90	-0.0457574906

Log Loss

It is the negative average of the log of corrected predicted probabilities for each instance.

$$- \frac{1}{N} \sum_{i=1}^{N} (\log(p_i))$$

Analytics Vidhya
Learn everything about analytics

Corrected Probabilities

ID	Actual	Predicted probabilities	Corrected Probabilities	Log
ID6	1	0.94	0.94	-0.0268721464
ID1	1	0.90	0.90	-0.0457574906
ID7	1	0.78	0.78	-0.1079053973
ID8	0	0.56	0.44	-0.3565473235
ID2	0	0.51	0.49	-0.30980392
ID3	1	0.47	0.47	-0.3279021421
ID4	1	0.32	0.32	-0.4948500217
ID5	0	0.10	0.90	-0.0457574906

Log Loss =0.2144

Analytics Vidhya
Learn everything about analytics

Log Loss

$$\text{Log loss} = \frac{1}{N} \sum_{i=1}^{N} - \left(y_i * \log(p_i) + (1-y_i) * \log(1-p_i) \right)$$

Analytics Vidhya
Learn everything about analytics

Log Loss

$$\text{Log loss} = \frac{1}{N} \sum_{i=1}^{N} - (y_i * \log(p_i) + (1-y_i) * \log(1-p_i))$$

- p_i is probability of 1
- $1-p_i$ is probability of 0