Data Science & AI

Machine Learning

Practical Implementation

Lecture No.- 02

By- Krish Naik Sir

Recap of Previous Lecture

Topic

Topic

Topic

Topic

Topic

Machine Learning Model

Topics to be Covered

- Topic ROC Curve
- Topic
- Topic Threshold
- Topic
- Topic

Area Under the Curve OF Reciever Operating Characteristics (ROC - AUC)

Confusion Matrix

Threshold = 0.5 $\quad \geq 0.5 \Rightarrow 1$ $\quad < 0.5 \Rightarrow 0$

	Actual 1	Actual 0
Predicted 1	TP	FP
Predicted 0	FN	TN

[0, 0.2, 0.4, 0.6, 0.8] True Positive Rate

y	$\hat{y}$	$\hat{y}(0)$	$\hat{y}(0.2)$	$\hat{y}(0.4)$
1	0.8	1	1	
→ 0	0.96	1	1	
1	0.4	1	1	
1	0.3	1	1	
→ 0	0.2	1	0	
1	0.7	1	1	

$$TPR = \frac{TP}{TP+FN}$$

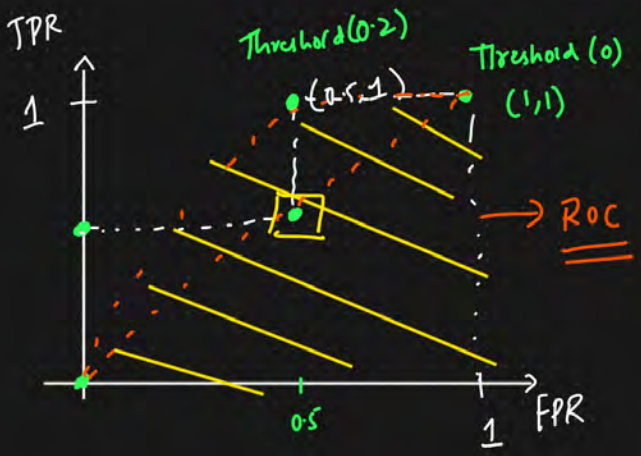
$$FPR = \frac{FP}{FP+TN}$$ False Positive Rate.

$\hat{y}(0)$: $TPR = \frac{4}{4+0} = 1$

$\hat{y}(0)$: $FPR = \frac{2}{2+0} = 1$

$\hat{y}(0.2)$: $TPR = \frac{4}{4+0} = 1$

$\hat{y}(0.2)$: $FPR = \frac{1}{1+1} = 0.5$



Threshold ↓

(0.4)

Area of the curve ↑↑.

Question:
What does a point on the ROC curve represent?

(A) The trade-off between precision and recall at a certain threshold.
(B) The trade-off between true positive rate and false positive rate at a certain threshold. ✓
(C) The trade-off between accuracy and error rate at a certain threshold.
(D) The overall accuracy of the model at a certain threshold.

Machine Learning

① Linear Regression → Loss, Gradient Descent, Weight Updation

↓

Logistic Regression

② Naive Bayes, SVM, SVR, KNN → Lazy Learner

④ Optimizers, Forward Prop, Backward Prop, Optimizers, Epoch, Iteration, Parameters

⑤ PCA, Curse of Dimensionality, KMeans, Hierarchy, DBScan

⑥ R^2, Adjusted R^2, CROSS VALIDATION, Hyperparameter Tuning, Confusion Matrix, Precision, Recall, F Beta Score, ROC-AUC Curve

You are building a spam classifier using Naive Bayes. You have the following data from your training set:

60% of the emails are spam.
40% of the emails are not spam (ham).
In the spam emails, 30% of them contain the word "offer".
In the ham emails, 5% of them contain the word "offer".
If you receive an email that contains the word "offer", what is the probability that this email is spam? Choose the closest answer.

DPP BAYE's Theorm

(A) 30%
(B) 45%
(C) 78%
(D) 85%

P(Yes|( )) =

P("offer")=P("offer"|Spam)×P(Spam)+P("offer"|Ham)×P(Ham)

2 mins Summary

- Topic
- Topic
- Topic
- Topic
- Topic

THANK - YOU