



Evaluating Classification Models Part 2

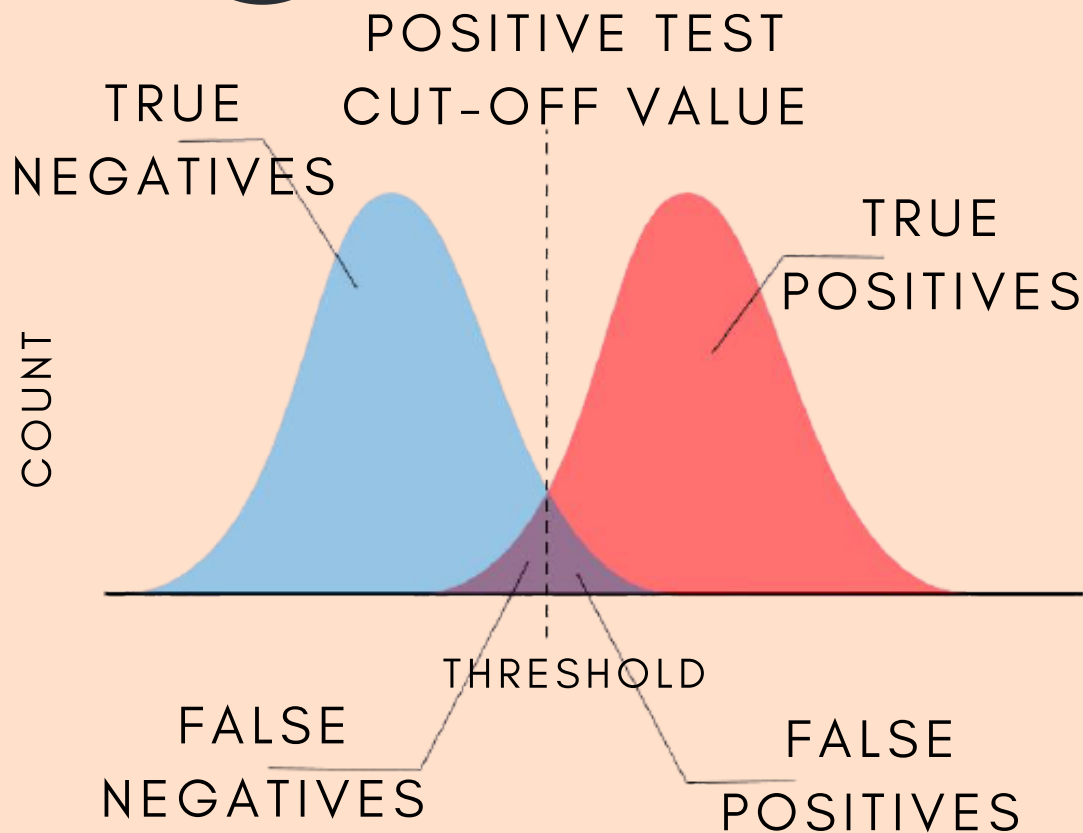
FIXING FUNDAMENTALS
BY ROBOFIED



WWW.ROBOFIED.COM

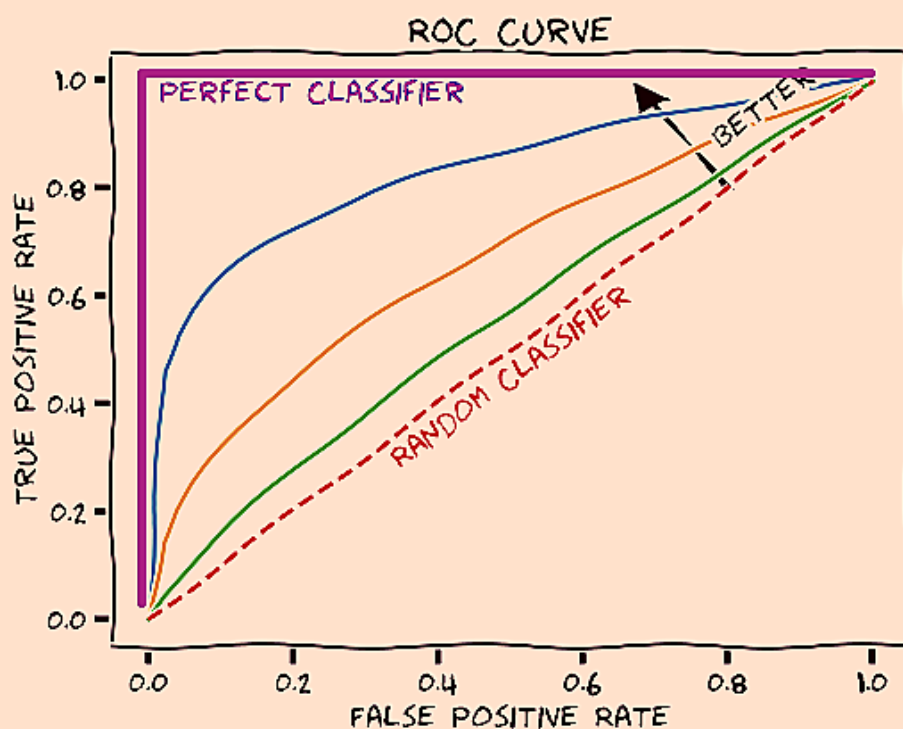
ROC & AUC

01



ROC curve is TPR(Recall) plotted against FPR (measures among the truly negative cases, how much percentage of them are actually false positive) and bigger the area under the curve the better the separation between different groups. AUC is the area under the ROC curve and it is a good measure on two things:

- How well the model can separate the two classes (positive and negative).
- How accurate is the model in identifying different categories (like whether it identify group A to positive correctly)?

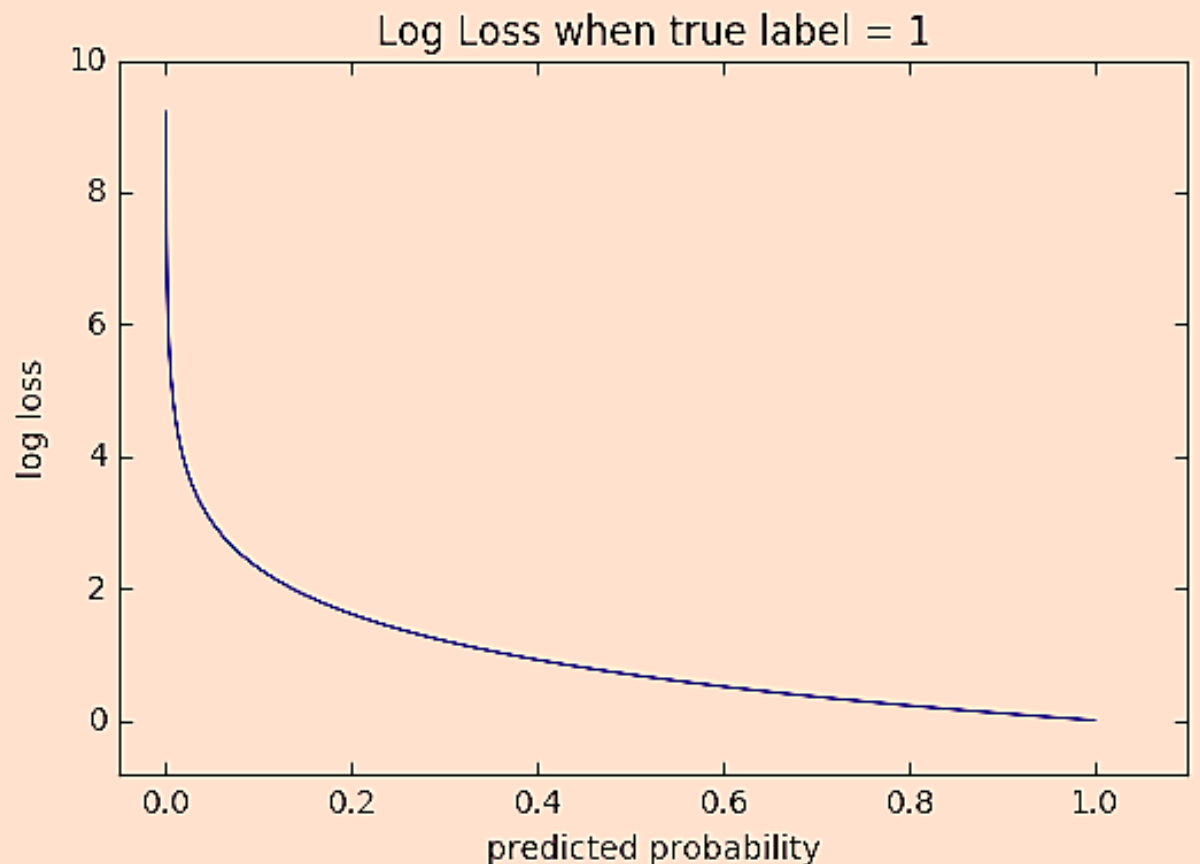


Code

```
● ● ●  
  
#Calculate AUC  
from sklearn.metrics import roc_curve, auc  
fpr, tpr, treshold = roc_curve(y_test, y_score_new)  
roc_auc = auc(fpr, tpr)  
#auc is 0.96
```


Binary Cross-Entropy/Log Loss

Log loss function penalizes wrongly classified cases and also lack of confidence in correctly classified cases. Same as ROC-AUC, it does not only consider the classification accuracy or results but also considers the probability assigned to each of the cases from the model. The better the model, the lower the log loss value.



$$\text{Log - Loss} = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))$$

Code

```
  
#Calculate log loss  
from sklearn.metrics import log_loss  
log_loss(y_test,y_score)  
#log loss is 0.32241485496779987
```

05

Tips from Team at Robofied

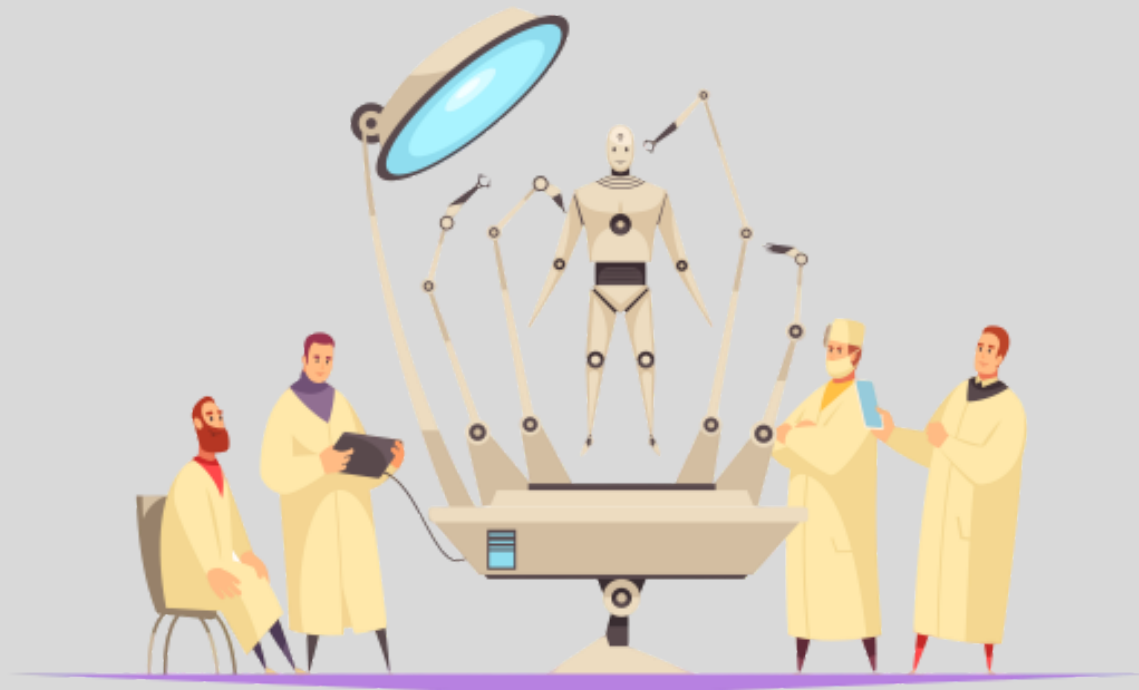


In conclusion, different metrics serve different purposes and can be applied accordingly to different problems. However it is advised to use ROC-AUC or log loss generally because they consider the prediction probability of the prediction for each classes. Even though both models predict correctly for the case, 90% confidence prediction model is definitely more preferred than the 51% confidence prediction model.

Other Resources

06

1. Machine Learning Mastery Blogs by Jason Brownlee
2. CS229 Machine Learning @Stanford by Andrew Ng
3. The Elements of Statistical Learning Book by Trevor Hastie, Robert Tibshirani, Jerome H. Friedman





At **Robofied**,
we aspire to democratize **AI**
Education

Join our channels:

 www.robofied.com

 <https://instagram.com/robofied>

 <https://twitter.com/Robofied>