Classification Metrics (Part 2)

Dr. Kalidas Y., IIT Tirupati

64) key phrase… “Threshold on Score”

- Human to Decide
  - on which yi = 1 (positive) and
  - on which yj = 0 (negative)
- Get a Data Set
  - x1 = (1,1), y1 = 1
  - x2 = (1,2), y2 = 1
  - x3 = (1,3), y3 = 1
  - x4 = (2,1), y4 = 0
  - x5 = (3,1), y5 = 0
- Model Predictions
  - y1’ = 1
  - y2’ = 1
  - y3’ = 0
  - y4’ = 1
  - y5’ = 0

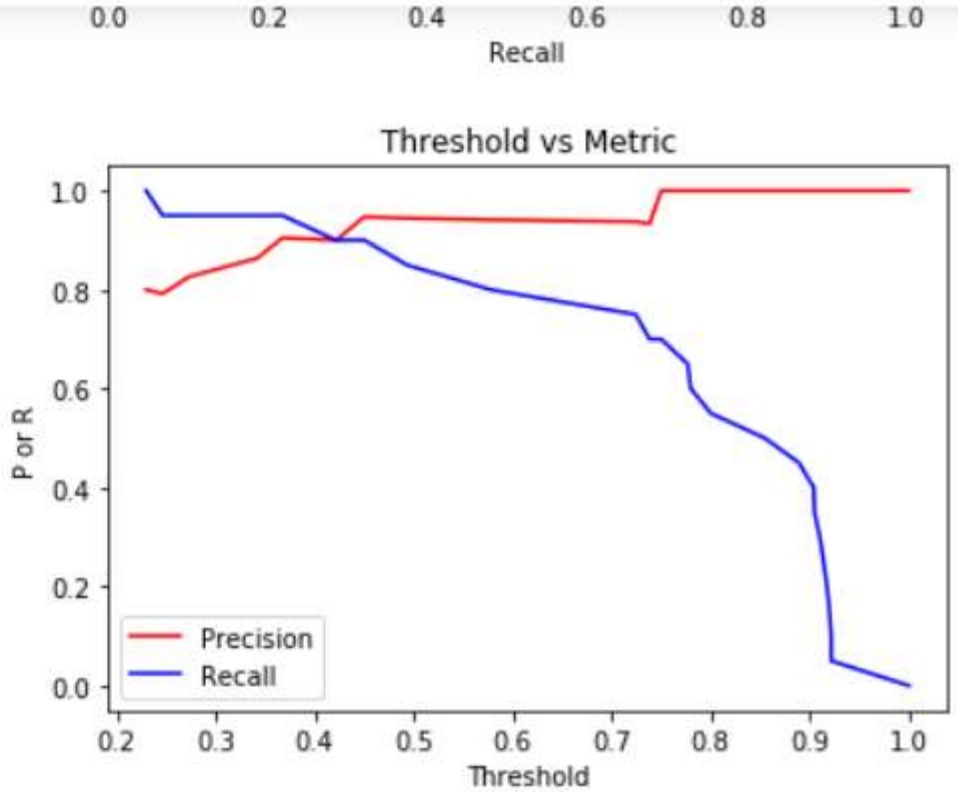
- Human to Decide
  - on which yi = 1 (positive) and
  - on which yj = 0 (negative)
- Get a Data Set
  - x1 = (1,1), y1 = 1
  - x2 = (1,2), y2 = 1
  - x3 = (1,3), y3 = 1
  - x4 = (2,1), y4 = 0
  - x5 = (3,1), y5 = 0
- Model Prediction Scores
  - y1_score = 0.9
  - y2_score = 0.7
  - y3_score = 0.3
  - y4_score = 0.8
  - y5_score = 0.4

Choose a Threshold

Above which are predicted +

below which are predicted -

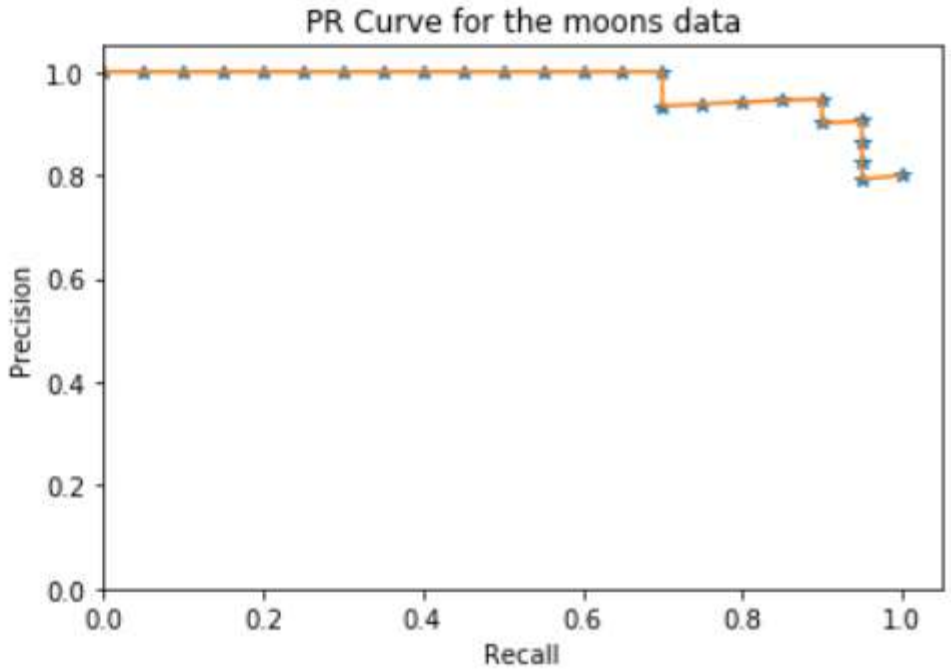
For example 0.55 ???



ROC Curve

65) key phrase… “PR Curve”

```
plt.legend(loc='best')
plt.show()
```



66) key phrase… “FPR or Fall out”

Google

false positive rate formula

All | Images | Videos | News | Maps | More | Settings | Tools

About 8,77,00,000 results (0.73 seconds)

False positive rate / Formula

$$\frac{FP}{N} = \frac{FP}{FP + TN}$$

FP = number of false positives

TN = number of true negatives

N = total number of negatives

From the web

The **false positive rate** is calculated as FP/FP+TN, where FP is the number of **false positives** and TN is the number of true negatives (FP+TN being the total number of negatives). It's the probability that a **false** alarm will be raised: that a **positive** result will be given when the true value is negative.

www.split.io › glossary › false-positive-rate

False Positive Rate | Split Glossary - Split Software

Fall out

67) key phrase… “AUC – Area Under Curve”

```
plt.ylim([0.0,1.05])
plt.show()
```

AUC – Area Under Curve

TPR is actually.. Recall