Train In Data

Precision, Recall and F-measure

Precision and Recall

- True Positive Rate (Recall or Sensitivity)

$$TP_{rate} = TP / (TP + FN)$$

- True Negative Rate

$$TN_{rate} = TN / (FP + TN)$$

- Positive Predictive Value (Precision)

$$PP_{value} = TP / (TP + FP)$$

- Negative predictive Value

$$NP_{value} = TN / (TN + FN)$$

Precision and Recall

- True Positive Rate (Recall or Sensitivity)

  $TP_{rate} = TP / (TP + FN)$

  Minority class

- True Negative Rate

  $TN_{rate} = TN / (FP + TN)$

  Majority class

- Positive Predictive Value (Precision)

  $PP_{value} = TP / (TP + FP)$

  Minority class

- Negative predictive Value

  $NP_{value} = TN / (TN + FN)$

  Majority class

Precision and Recall

- Both precision and recall vary between 0 and 1
- To select and tune machine learning models, our goal is to maximize both precision and recall.
- Both precision and recall depend on a probability threshold
  - Values vary depending on the threshold we use to determine the class output

F-measure

- The F1 score is a weighted harmonic mean of precision and recall.

$$F\text{-}measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

F-measure

- The F1 score is a weighted harmonic mean of precision and recall.
- The best score is 1.0 and the worst is 0.0.
- Optimizing this metric produces the best balance between precision and recall

Support

- Support is the number of actual occurrences of the class in the specified dataset.
- Indicates structural weaknesses in the reported scores. Highlights imbalanced datasets.
- Doesn’t change between models but instead diagnoses the evaluation process.

Recall

Target	Prediction
0	0
0	0
0	0
0	1
0	1
0	0
0	0
1	0
1	1
1	1

- Recall = TP / (TP + FN)
- Recall = 2 / (2+1) = 2 / 3 = 0.66

Precision

Target	Prediction
0	0
0	0
0	0
0	1
0	1
0	0
0	0
1	0
1	1
1	1

- Precision = TP / (TP + FP)
- Precision = 2 / (2+2) = 2 / 4 = 0.5

F-measure

Target	Prediction
0	0
0	0
0	0
0	1
0	1
0	0
0	0
1	0
1	1
1	1

- Recall = 2 / (2+1) = 2 / 3 = 0.66
- Precision = 2 / (2+2) = 2 / 4 = 0.5

$$F\text{-}measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$F\text{-}measure = \frac{2 \times 0.5 \times 0.66}{0.5 + 0.66} = 0.57$$

Support

Target	Prediction
0	0
0	0
0	0
0	1
0	1
0	0
0	0
1	0
1	1
1	1

- Majority class = 7
- Minority class = 3

Measures per class

Target	Prediction
0	0
0	0
0	0
0	1
0	1
0	0
0	0
1	0
1	1
1	1

- True Positive Rate (Recall or Sensitivity)

  TP rate = TP / (TP + FN)

  TP rate = 2 / (2 + 1) = 0.66

- True Negative Rate

  TN rate = TN / (FP + TN)

  TN rate = 5 / (2 + 5) = 0.71

- Positive Predictive Value (Precision)

  PP value = TP / (TP + FP)

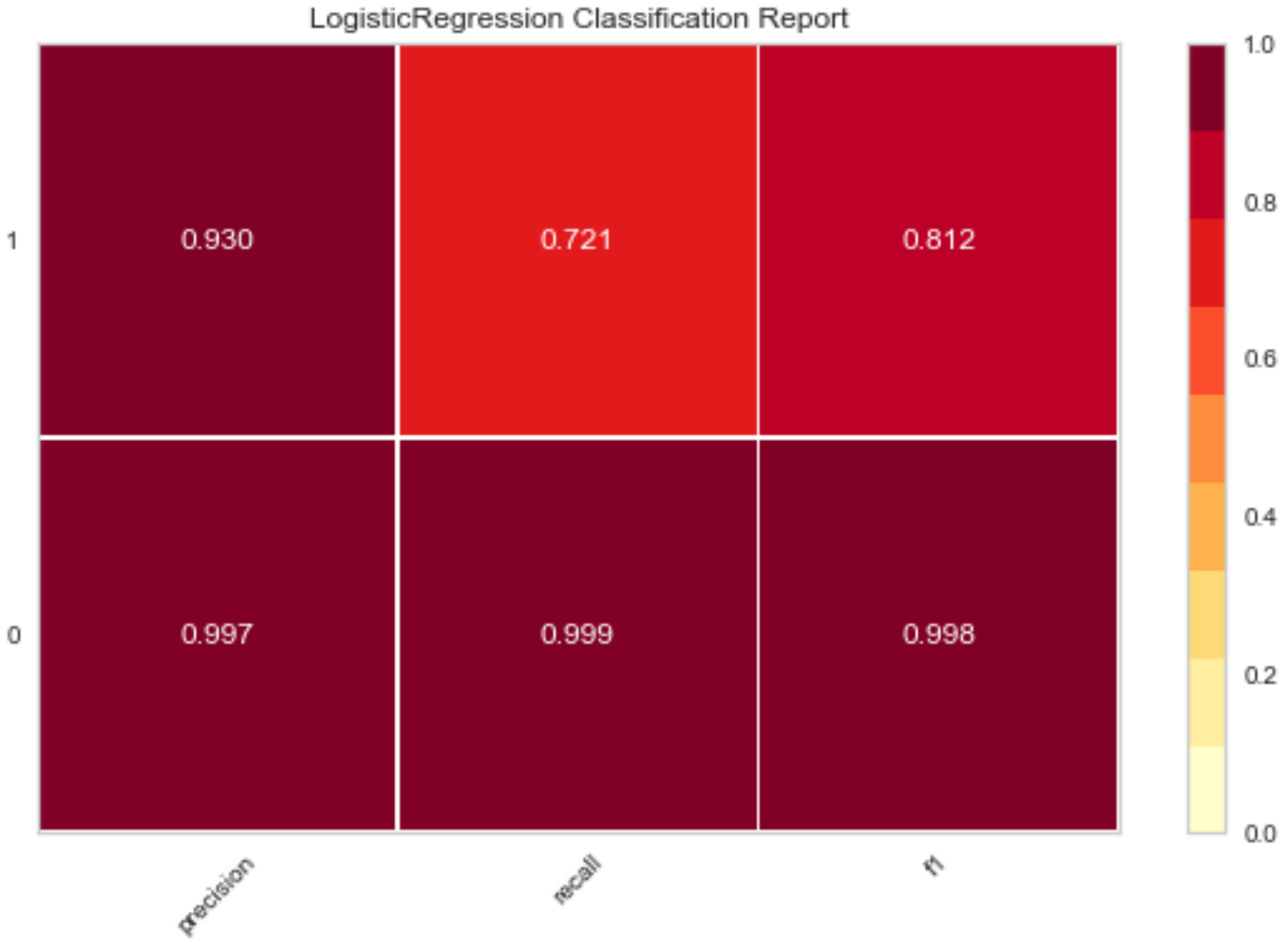
  PP value = 2 / (2 + 2) = 0.5

- Negative predictive Value

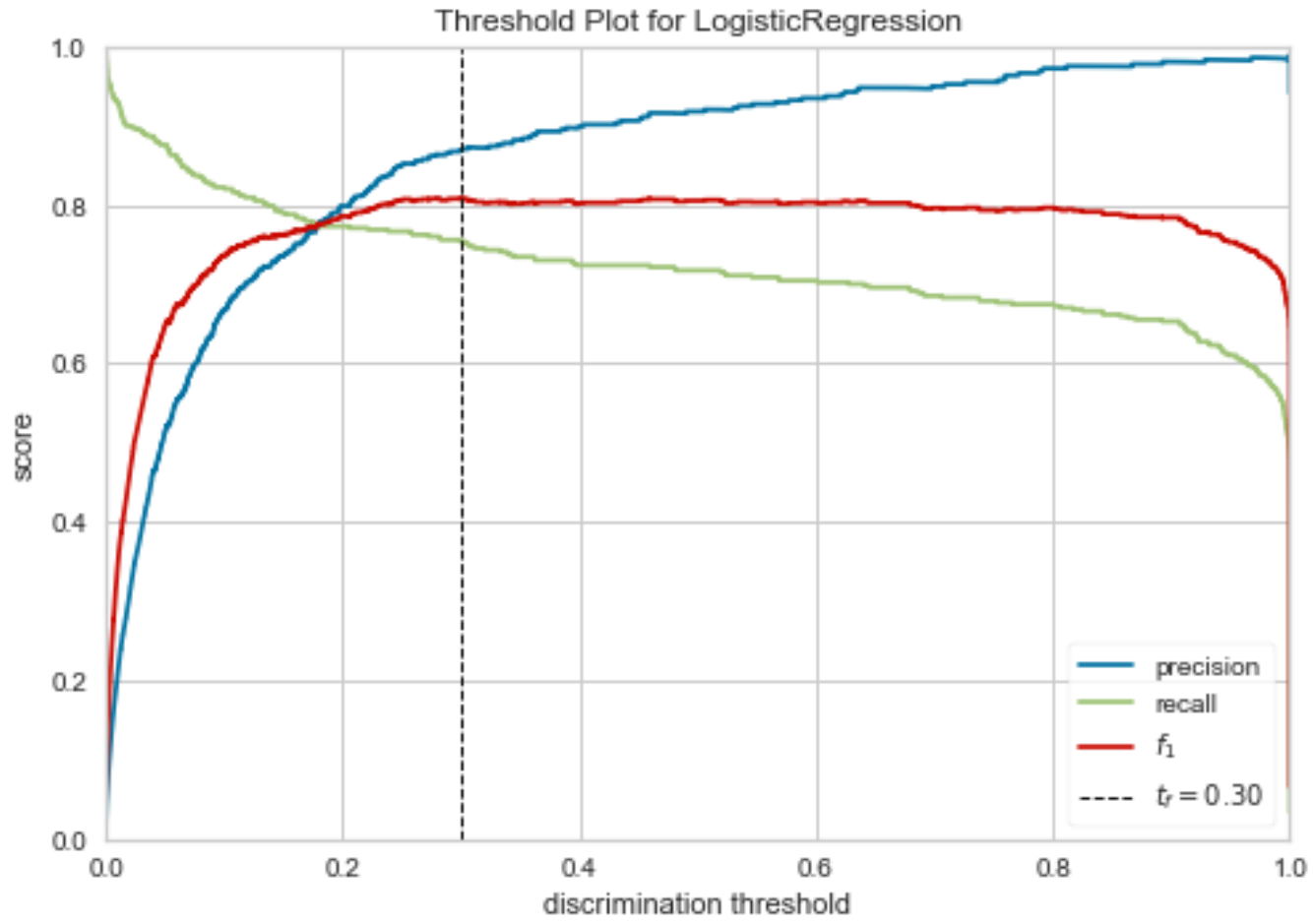
  NP value = TN / (TN + FN)

  NP value = 5 / (5 + 1) = 0.82

Breakdown


LogisticRegression Classification Report
1
0
0.930
0.721
0.812
0.997
0.999
0.998
precision
recall
f1
1.0
0.8
0.6
0.4
0.2
0.0

Optimising the threshold



Discrimination threshold: probability above which we classify a sample as positive

The optimal threshold is that at which F-measure is highest

Train In Data

THANK YOU

www.trainindata.com