How can the F1 score be calculated in R, and can you provide an example of its implementation?

Authored by

stats writer

May 7, 2024

The F1 score is a commonly used metric for evaluating the performance of a classification model in machine learning. It takes into account both precision and recall, making it a more robust measure than accuracy alone. In R, the F1 score can be calculated by using the "f1_score" function from the "caret" package. This function takes in the predicted values and the actual values from a classification model and returns the F1 score. For example, if we have a model that predicts whether a patient has a disease or not, we can use the "f1_score" function to calculate the F1 score based on the model's predictions and the actual disease status of the patients. This allows us to assess the overall performance of the model and make adjustments if needed.

Calculate F1 Score in R (Including Example)

When using in machine learning, a common metric that we use to assess the quality of the model is the F1 Score.

This metric is calculated as:

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

where:

Precision: Correct positive predictions relative to total positive predictionsRecall: Correct positive predictions relative to total actual positives

For example, suppose we use a logistic regression model to predict whether or not 400 different college basketball players get drafted into the NBA.

The following confusion matrix summarizes the predictions made by the model:

		Predicted	
		Drafted = Yes	**Drafted = No**
Actual	**Drafted = Yes**	**120** (True Positive)	**40** (False Negative)
	Drafted = No	**70** (False positive)	**170** (True Negative)

Here is how to calculate the F1 score of the model:

Precision = True Positive / (True Positive + False Positive) = 120/ (120+70) = .63157

Recall = True Positive / (True Positive + False Negative) = 120 / (120+40) = .75

F1 Score = 2 * (.63157 * .75) / (.63157 + .75) = .6857

The following example shows how to calculate the F1 score for this exact model in R.

Example: Calculating F1 Score in R

The following code shows how to use the confusionMatrix() function from the caret package in R to calculate the F1 score (and other metrics) for a given

logistic regression model:

```
library(caret)

#define vectors of actual values and predicted values
actual <- factor(rep(c(1, 0), times=c(160, 240)))
pred <- factor(rep(c(1, 0, 1, 0), times=c(120, 40, 70, 170)))

#create confusion matrix and calculate metrics related to confusion matrix
confusionMatrix(pred, actual, mode = "everything", positive="1")
```

```
Reference
Prediction 0 1
0 170 40
1 70 120

Accuracy : 0.725
95% CI : (0.6784, 0.7682)
No Information Rate : 0.6
P-Value : 1.176e-07

Kappa : 0.4444

Mcnemar's Test P-Value : 0.005692
```

Sensitivity : 0.7500

Specificity : 0.7083

Pos Pred Value : 0.6316

Neg Pred Value : 0.8095

Precision : 0.6316

Recall : 0.7500

F1 : 0.6857

Prevalence : 0.4000

Detection Rate : 0.3000

Detection Prevalence : 0.4750

Balanced Accuracy : 0.7292

'Positive' Class : 1

We can see that the F1 score is 0.6857. This matches the value that we calculated earlier by hand.

If you use F1 score to compare several models, the model with the highest F1 score represents the model that is best able to classify observations into classes.

For example, if you fit another logistic regression model to the data and that model has an F1 score of 0.85, that model would be considered better since it has a higher F1 score.