Classification: Accuracy

Estimated Time: 6 minutes

Accuracy is one metric for evaluating classification models. Informally, **accuracy** is the fraction of predictions our model got right. Formally, accuracy has the following definition:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where *TP* = True Positives, *TN* = True Negatives, *FP* = False Positives, and *FN* = False Negatives.

Let's try calculating accuracy for the following model that classified 100 tumors as malignant (https://wikipedia.org/wiki/Malignancy) (the positive class) or benign (https://wikipedia.org/wiki/Benign_tumor) (the negative class):

True Positive (TP):

- Reality: Malignant
- ML model predicted: Malignant
- **Number of TP results: 1**

False Positive (FP):

- Reality: Benign
- ML model predicted: Malignant
- **Number of FP results: 1**

False Negative (FN):

- Reality: Malignant
- ML model predicted: Benign
- **Number of FN results: 8**

True Negative (TN):

- Reality: Benign
- ML model predicted: Benign
- **Number of TN results: 90**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{1 + 90}{1 + 90 + 1 + 8} = 0.91$$

Accuracy comes out to 0.91, or 91% (91 correct predictions out of 100 total examples). That means our tumor classifier is doing a great job of identifying malignancies, right?

Actually, let's do a closer analysis of positives and negatives to gain more insight into our model's performance.

Of the 100 tumor examples, 91 are benign (90 TNs and 1 FP) and 9 are malignant (1 TP and 8 FNs).

Of the 91 benign tumors, the model correctly identifies 90 as benign. That's good. However, of the 9 malignant tumors, the model only correctly identifies 1 as malignant—a terrible outcome, as 8 out of 9 malignancies go undiagnosed!

While 91% accuracy may seem good at first glance, another tumor-classifier model that always predicts benign would achieve the exact same accuracy (91/100 correct predictions) on our examples. In other words, our model is no better than one that has zero predictive ability to distinguish malignant tumors from benign tumors.

Accuracy alone doesn't tell the full story when you're working with a **class-imbalanced data set**, like this one, where there is a significant disparity between the number of positive and negative labels.

In the next section, we'll look at two better metrics for evaluating class-imbalanced problems: precision and recall.

Key Terms

- accuracy (/machine-learning/glossary#accuracy)
- class-imbalanced data set (/machine-learning/glossary#class_imbalanced_data_set)

Help Center (https://support.google.com/machinelearningeducation)

Previous
← True vs. False; Positive vs. Negative (/machine-learning/crash-course/classification/true-false-positive-negative)

Next
Precision and Recall (/machine-learning/crash-course/classification/precision-and-recall) →

Except as otherwise noted, the content of this page is licensed under the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the Apache 2.0 License (https://www.apache.org/licenses/LICENSE-2.0). For details, see the Google Developers Site Policies (https://developers.google.com/site-policies). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-02-10 UTC.