

Metric

Donnerstag, 18. November 2021 16:22

T: Richtig vorhergesagt

F: Falsch vorhergesagt

P: Bösartiger Tumor

N: Harmloser Tumor

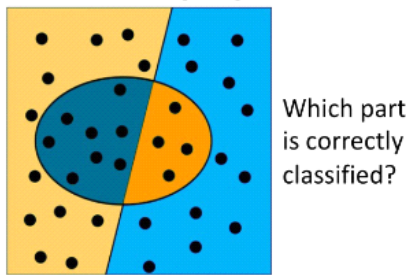
TP: Das Modell hat die Vorhersage richtig als bösartigen Tumor erkannt

TN: Das Modell hat die Vorhersage richtig als harmlosen Tumor erkannt

FP: Das Modell hat ein bösartigen Tumor vorhergesagt, obwohl es sich um einen Harmlosen handelt

FN: Das Modell hat harmlosen Tumor vorhergesagt, obwohl es sich um einen Bösartigen handelt

$$\text{Accuracy} = \frac{t_p + t_n}{t_p + f_p + t_n + f_n}$$



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

Gegeben ist ein 100 zeiliger Tumor-Datensatz.

Dieser Datensatz enthält Labels, die einen bösartigen Tumor (positiv) oder einen harmlosen Tumor (negativ) angeben, wovon 92 harmlos und 8 bösartig sind. Das hat zur Folge das der Datensatz unbalanciert ist (Accuracy wäre sinnvoll, wenn balanciert).

Dabei wurden vom trainierten Modell folgende Voraussagen entnommen:

Von den 92 Harmlosen wurden 91 richtig erkannt.

Von den 8 Bösartigen wurde nur 1 erkannt
 $1 / 8 = 0,125$ (Precision)

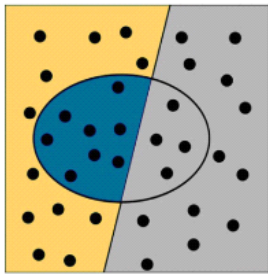
=>

Accuracy gibt an wie viel von den 100 Zeilen richtig vorhergesagt wurden.

$$\text{Accuracy} = (91 + 1) / (92 + 8) = 0,92$$

In dem Fall gibt die Accuracy einen guten Score an, obwohl es nicht der Fall ist. Von den bösartigen Tumoren wurden nur 12 % richtig erkannt, als ungefähr nur jeder 10te.

$$\text{Precision} = \frac{t_p}{t_p + f_p}$$



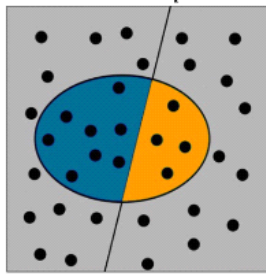
Which part of those predicted as positives are correct?

$$\frac{\# \text{ positives correct}}{\# \text{ labeled positive}}$$

Die Precision gibt an wie viel von den erkannten bös. Tumor tatsächlich solche sind.

Das bedeutet das bei dieser Metrik mehr Wert darauf gelegt wird, das von den gelabelten Daten möglichst viel richtig vorhergesagt wird. Dann wird allerdings FN (False Negative) nicht berücksichtigt, was in dem Kontext fatale Konsequenzen bewirken kann. In dem Fall bedeutet es, das es egal ist wenn ein bösartiger Tumor als harmlos eingestuft wird. Wenn beispielsweise die Precision 80 % beträgt, aber davon 2 vermeintlich harmlose Tumore dabei sind, dann stellt das ein Problem dar. Als Fazit kann die Precision problemlos verwendet werden, wenn FN Vorhersagen keine entscheidende Rolle spielen.

$$\text{Recall} = \frac{t_p}{t_p + f_n}$$



Which part of those you should have predicted did you predict?

$$\frac{\# \text{ positives correct}}{\# \text{ actually positive}}$$

Recall gibt an wie viel von den gesamten bös. Tumoren im Datensatz erkannt wurden.

Recall berücksichtigt FN, weshalb es in diesem Kontext eine bessere Metrik ist. Recall bewirkt, das möglichst viele bösartige Tumore richtig erkannt werden. Das hat leider zur Folge das harmlose Tumore eher als bösartig eingestuft werden. Dadurch kann es zwar nicht mehr zu FN Vorhersagen kommen, aber die Precision erreicht dadurch einen schlechteren Score.

=> Precision und Recall sind Gegensätze

=> Als Ausgleich F1 Score als Metrik verwenden