

Confusion matrix

Erarbeitet von
Dr. Katja Theune

Lernziele	1
Inhalt	2
Einstieg	2
Confusion matrix: Idee	2
Confusion matrix: Beispiel	3
Abschluss	5
Weiterführendes Material	6
Disclaimer	6

Lernziele

- Du kannst erläutern, was eine confusion matrix ist
- Du kannst die Einträge einer confusion matrix erklären
- Du kannst anhand von Beispielen eine confusion matrix erstellen

Inhalt

Einstieg

Wenn wir uns mit Prognosemodellen beschäftigen, fragen wir uns auch immer, wie gut oder genau unser Modell etwas vorhersagt und worin seine Stärken und Schwächen liegen. Häufig wollen wir auch verschiedene Prognosemodelle miteinander vergleichen. Wir brauchen also Maße, die uns bei der Beurteilung der Modelle helfen. Bei Klassifikationsproblemen bildet dafür die sogenannte confusion matrix eine gute Basis.

Confusion matrix: Idee

Wir befassen uns nun mit einem Klassifikationsproblem mit zwei Klassen. Das ist sehr intuitiv und leicht nachzuvollziehen. Kommen wir dafür zunächst wieder auf unsere Ausgangsfragestellung zurück: Wir wollen mit Daten von Studierenden ein Prognosemodell erstellen, das einen möglichen Studienabbruch vorhersagt. Nehmen wir an, wir haben so ein Modell aufgestellt und wollen nun prüfen, wie genau es Vorhersagen trifft. Häufig werden Beobachtungen in der für uns besonders relevanten Klasse als - englisch ausgesprochen- Positives bezeichnet, Beobachtungen in der anderen Klasse als Negatives. In unserem Fall wären z. B. Beobachtungen in der Klasse „Studienabbruch“ Positives und diejenigen in der Klasse „kein Studienabbruch“ Negatives. Diese Begriffe sind dabei nicht wertend gemeint.

Um unser Modell zu evaluieren, wollen wir die wahren Outputs, hier die Klassen, mit den vom Modell vorhergesagten Outputs vergleichen. Dafür lässt sich eine sogenannte confusion matrix oder auf deutsch auch Konfusionsmatrix oder Wahrheitsmatrix genannt, heranziehen. Die Zeilen geben die wahre Klasse einer Beobachtung an. Diese kennen wir ja aus den Daten. Die Spalten geben die vom Modell vorhergesagte Klasse an.

Wir können jetzt in die Zellen der Matrix die folgenden Konstrukte bzw. Werte eintragen:

Die True Positives, kurz TP. Sie geben die Anzahl an wahren Positives an, die auch als solche richtig prognostiziert wurden. Das wären bei uns z. B. wahre Studienabbrüche, die auch als solche vom Modell erkannt wurden.

Die True Negatives, kurz TN. Sie geben analog die Anzahl an wahren Negatives an, die auch als solche richtig prognostiziert wurden.

Die False Positives, kurz FP. Sie entsprechen der Anzahl an wahren Negatives, die fälschlicherweise als Positives prognostiziert wurden.

Und als letztes die False Negatives, kurz FN. Sie entsprechen analog der Anzahl an wahren Positives, die fälschlicherweise als Negatives prognostiziert wurden.

In unserem Fall wären das z. B. wahre Studienabbrüche, die aber als solche nicht erkannt wurden.

Klassifikation: Confusion matrix

True Positives (TP):

Anzahl an wahren Positives,
die auch also solche richtig prognostiziert wurden

True Negatives (TN):

Anzahl an wahren Negatives,
die auch als solche richtig prognostiziert wurden

False Positives (FP):

Anzahl an wahren Negatives,
die fälschlicherweise als Positives prognostiziert wurden

False Negatives (FN):

Anzahl an wahren Positives,
die fälschlicherweise als Negatives prognostiziert wurden

		Vorhergesagte Klasse	
Wahre Klasse		Positive	Negative
	Positive	TP	FN
	Negative	FP	TN

→ richtige Vorhersagen

→ falsche Vorhersagen

Die True Positives und die True Negatives sagen uns also, wo das Modell richtige Vorhersagen gemacht hat. Hier sollten also vergleichbar große Werte stehen. Die False Positives und False Negatives sagen, wo es falsche Vorhersagen gab, also wo sich das Modell geirrt hat.

Man kann solch eine confusion matrix natürlich auch für mehr als zwei Klassen aufstellen. Man hat dann eine sehr ähnliche, nur eben größere Matrix.

Confusion matrix: Beispiel

Confusion matrix: Beispiel

n=10 Beobachtungen

Beobachtung	wahre Klasse	vorhergesagte Klasse
1	Studienabbruch	kein Studienabbruch
2	kein Studienabbruch	kein Studienabbruch
3	Studienabbruch	Studienabbruch
4	Studienabbruch	Studienabbruch
5	kein Studienabbruch	kein Studienabbruch
6	kein Studienabbruch	kein Studienabbruch
7	kein Studienabbruch	Studienabbruch
8	Studienabbruch	Studienabbruch
9	kein Studienabbruch	kein Studienabbruch
10	kein Studienabbruch	Studienabbruch

←

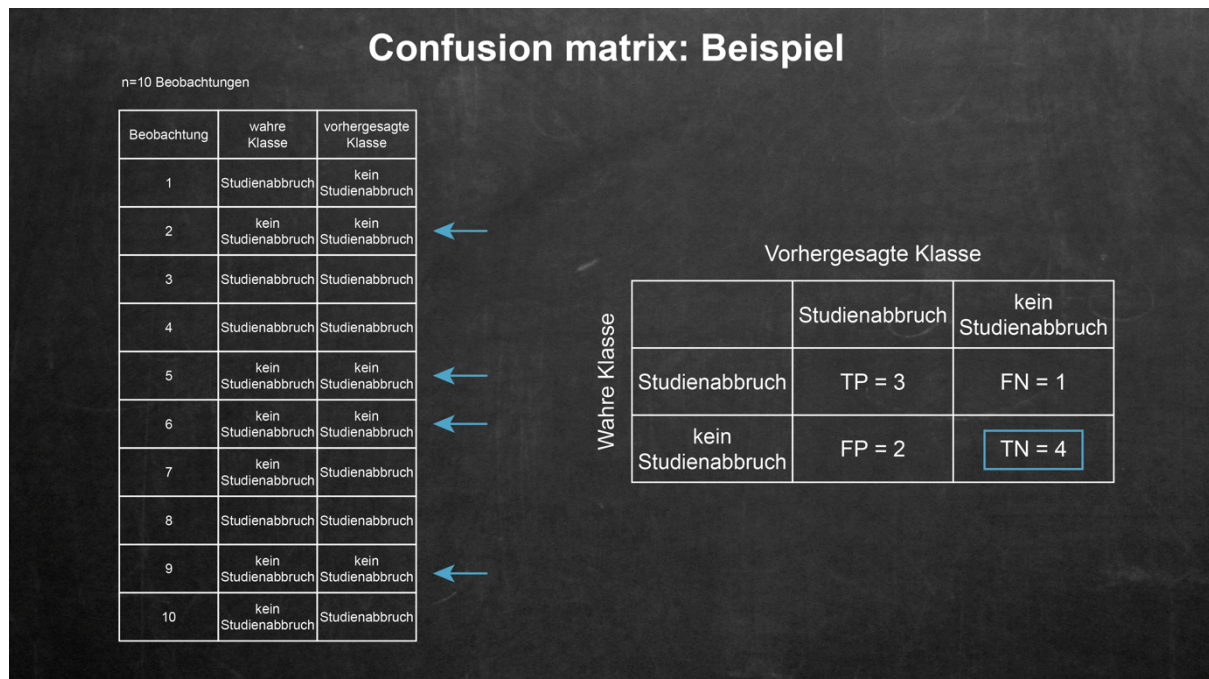
←

←

		Vorhergesagte Klasse	
		Studienabbruch	kein Studienabbruch
Wahre Klasse	Studienabbruch	TP = 3	FN = 1
	kein Studienabbruch	FP = 2	TN = 4

Schauen wir uns das einmal an einem einfachen Beispiel mit 10 Beobachtungen an. Für jede Beobachtung haben wir in dieser Tabelle einmal die wahre Klasse aus unseren Daten und einmal die vom Modell vorhergesagte Klasse abgetragen. Um jetzt zu den True Positives zu kommen, zählen wir alle Beobachtungen, deren wahre Klasse und auch die prognostizierte Klasse „Studienabbruch“ ist. Das sind folgende (siehe Grafik), also insgesamt drei.

Das gleiche machen wir für die True Negatives. Hier suchen wir alle Beobachtung mit der wahren und auch prognostizierten Klasse „kein Studienabbruch“. Das sind dann folgende (siehe Grafik), also vier.



Ähnlich ergeben sich die Werte für die False Negatives, hier ist es nur einer und für die False Positives, hier sind es zwei (siehe Grafiken).

Confusion matrix: Beispiel

n=10 Beobachtungen

Beobachtung	wahre Klasse	vorhergesagte Klasse
1	Studienabbruch	kein Studienabbruch
2	kein Studienabbruch	kein Studienabbruch
3	Studienabbruch	Studienabbruch
4	Studienabbruch	Studienabbruch
5	kein Studienabbruch	kein Studienabbruch
6	kein Studienabbruch	kein Studienabbruch
7	kein Studienabbruch	Studienabbruch
8	Studienabbruch	Studienabbruch
9	kein Studienabbruch	kein Studienabbruch
10	kein Studienabbruch	Studienabbruch

Vorhergesagte Klasse

	Studienabbruch	kein Studienabbruch
Studienabbruch	TP = 3	FN = 1
kein Studienabbruch	FP = 2	TN = 4

Wahre Klasse

Confusion matrix: Beispiel

n=10 Beobachtungen

Beobachtung	wahre Klasse	vorhergesagte Klasse
1	Studienabbruch	kein Studienabbruch
2	kein Studienabbruch	kein Studienabbruch
3	Studienabbruch	Studienabbruch
4	Studienabbruch	Studienabbruch
5	kein Studienabbruch	kein Studienabbruch
6	kein Studienabbruch	kein Studienabbruch
7	kein Studienabbruch	Studienabbruch
8	Studienabbruch	Studienabbruch
9	kein Studienabbruch	kein Studienabbruch
10	kein Studienabbruch	Studienabbruch

Vorhergesagte Klasse

	Studienabbruch	kein Studienabbruch
Studienabbruch	TP = 3	FN = 1
kein Studienabbruch	FP = 2	TN = 4

Wahre Klasse

Wir können diese Angaben nun verwenden, um zu schauen, wie gut unser Modell bestimmte Klassen vorhersagt, wo Stärken und Schwächen liegen und um sogenannte Evaluationsmetriken abzuleiten.

Abschluss

Wir haben nun die insbesondere für eine Klassifikation wichtige confusion matrix kennengelernt, welche die wahre und die vorhergesagte Klasse der Beobachtungen übersichtlich gegenüberstellt. Sie dient als Grundlage für verschiedenen Metriken zur Evaluation unserer Prognosemodelle.

Weiterführendes Material

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3. Auflage). Morgan Kaufmann.

James, G., Witten, D., Hastie, T., & Tibshirani, R., & Tylor, J. (2023). *An Introduction to Statistical Learning - with Applications in Python*. Springer.

Lantz, B. (2015). *Machine learning with R* (2. Auflage). Packt Publishing Ltd, Birmingham.

Disclaimer

Transkript zu dem Video „Prognosemodelle (Klassifikation und Regression): Confusion matrix“, Dr. Katja Theune.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.