

# 1. ÜBUNGSBLATT

**Aufgabe 1.1**

(3 Punkt)

Ein binärer Klassifizierer  $K_1$  liefert auf den Validierungsdaten die nachfolgende Confusion Matrix.

		Ground Truth	
		Klasse A	Klasse B
Predicted	Klasse A	100	8
	Klasse B	2	5

Hierbei sei die Klasse A die positive Klasse und die Klasse B die negative Klasse.

Beantworten Sie die folgenden Fragen:

- (a) Wie viele Daten gehören tatsächlich zur Klasse A und wie viele zur Klasse B?
- (b) Berechnen Sie Precision, Recall, Accuracy und den  $F_1$ -Score
- (c) Ist eine dieser Metriken ausreichend um die Qualität des Klassifizierers zu beschreiben?  
Wenn ja: Welche Metrik beschreibt die Qualität des Klassifizierers am besten? Falls nicht:  
Was ist das Problem?
- (d) Angenommen wir haben einen zweiten Klassifizierer  $K_2$  der über eine Gleichverteilung zufällig entscheidet ob ein Datenpunkt zur Klasse A oder B gehört. Welche Accuracy hat  $K_2$ ?
- (e) Wie können Sie ganz einfach einen Klassifizierer entwickeln, der eine Precision von 100% erreicht?
- (f) Was würden Sie empfehlen, um  $K_1$  zu verbessern?

Hier finden Sie weitere Informationen zu gängigen Metriken<sup>3</sup>.

**Aufgabe 1.2**

(4 Punkte)

Erklären Sie die folgenden Klassifizierer in Ihren eigenen Worten (jeweils ca. 5-10 Sätze) mit Visualisierungen und ggf. Berechnungsformeln:

- (a) k-Nearest Neighbor
- (b) Decision Tree
- (c) Support Vector Machine (SVM)
- (d) Perceptron

<sup>3</sup>[https://en.wikipedia.org/wiki/Confusion\\_matrix](https://en.wikipedia.org/wiki/Confusion_matrix)

**Aufgabe 1.3 (Perzeptron)**

(6 Punkte)

Bearbeiten Sie das Notebook `Perzeptron_Aufgaben.ipynb`.

**Aufgabe 1.4**

(3 Punkte)

Beantworten Sie die folgenden Fragen:

- a) Beschreiben Sie die Begriffe Underfitting und Overfitting.
- b) Erklären Sie das Gradient Descent Verfahren.
- c) Geben Sie die Bayes'sche Regel an und erklären Sie die Bedeutung der einzelnen Terme.

**Bonusaufgabe 1.5**

(6 Punkte)

In dieser Aufgabe soll der k-Nearest-Neighbor Algorithmus implementiert werden. Sehen Sie sich dazu die Aufgaben im Jupyter-Notebook an und bearbeiten Sie diese.