

Aufgabe 1 Logistische Regression

7 Punkte

Implementieren Sie logistische Regression. Benutzen Sie Ihre Implementierung, um einen besseren Spam¹-Klassifikator zu trainieren. Splitten Sie den Datensatz dazu erneut in Trainings- und Testdaten. Normalisieren Sie die Daten vor dem Trainieren.

Plotten Sie wie sich die Klassifikationsgenauigkeit über die Iterationen verändert. Geben Sie die Konfusionsmatrix aus. Wie könnte man das Modell anpassen, um eine der Klassen zu bevorzugen?

Aufgabe 2 Logische Funktionen

3 Punkte

Ein klassisches Problem in Machine Learning war es das Verhalten logischer Funktionen zu lernen. Wenden Sie dazu Ihre Implementierung aus Aufgabe 1 auf die logischen Operatoren *and*, *or* und *xor* an.

Dabei haben Sie für jede logische Funktion einen Datensatz mit zwei Merkmalen (x_1, x_2) und ein Label, das beispielsweise die Werte von $y = x_1 \wedge x_2$ annimmt.

¹<https://archive.ics.uci.edu/ml/datasets/spambase> → Data Folder bzw. Data Set Description