



Machine Learning Übung 1: kNN

Allgemein:

Implementieren Sie einen Klassifizierer mittels k-Nearest-Neighbor Algorithmus. Sie können dafür Java oder C# als Sprache auswählen und dürfen keine Machine Learning Frameworks verwenden. Messen Sie jeweils die Zeit für jeweils 1000, 10.000 oder 100.000 Klassifizierungen (also ohne Einlesen der Daten, Lernen etc.). Teilen Sie die Daten jeweils in Lern- und Test Daten. Verwenden Sie dafür ein beliebiges Validation Verfahren, erstellen Sie eine Confusion Matrix und berechnen Sie die Accuracy für Ihren Algorithmus.

Achten Sie bei der Implementierung, dass Sie konkret dieses Beispiel lösen. Sie müssen keine allgemeine, wiederverwendbare Lösung implementieren, die auch für andere Daten verwendet werden kann.

Abzugeben ist ein einseitiger Bericht im PDF Format mit den gemessenen Zeiten und den Confusion Matrizen. Beim Prüfungsgespräch ist auch der Code zu erklären.

Variante 1:

Daten zu den Iris-Blumen aus <https://archive.ics.uci.edu/ml/datasets/Iris>.

Variante 2:

Daten zur Weinqualität aus <https://archive.ics.uci.edu/ml/datasets/Wine+Quality>, Rot-, oder Weißwein

Zusatzaufgabe (Optional):

Erstellen Sie zunächst ein Framework für den kNN Algorithmus und implementieren Sie Klassen, die auch für andere Daten zum Einsatz kommen können.