



OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

FAKULTÄT FÜR
WIRTSCHAFTSWISSENSCHAFT

Big Data Anwendungen

Praxisübung - Bonusblatt 2 (Clustering)

Die Iris Daten sind die klassischen Daten zum Erlernen des Umgangs mit Clustering Algorithmen. Sie bestehen aus drei Unterarten der Irispflanze mit jeweils 50 Samples und vier Eigenschaften der Blume.

Aufgabe ist es mit Hilfe eines Clustering Algorithmus herauszufinden, wie am besten Cluster zu finden sind und diese zu visualisieren.

Download: <https://www.kaggle.com/uciml/iris>