

k-Means Clustering

1 Aufgaben

1.1 Testen des K-Means Clustering mit synthetischen Daten

Das Ziel dieser Aufgabe ist es, den K-Means Algorithmus mit verschiedenen synthetischen Datensätzen und verschiedenen Werten für K zu testen und grafisch darzustellen.

- (a) Erstellen Sie ein neues Python-Skript.
- (b) Machen Sie sich mit der *scikit-learn* Funktion *make_blobs* vertraut.
- (c) Erstellen Sie einen farbigen 2D-Plot eines beispielhaften Datensatzes.
- (d) Importieren Sie die Klasse *KMeans* und erstellen Sie ein Objekt davon. Führen Sie das Clustering mit verschiedenen Werten für K durch. Nutzen Sie hierfür die Methoden *.fit(X)* und *.predict(X)*.
- (e) Erweitern Sie den Plot um die Einteilung durch das K-Means Clustering.
- (f) Machen Sie sich mit den Funktionen *make_moons* und *make_circles* vertraut. Erstellen Sie mithilfe dieser Funktionen synthetische Datensätze und wiederholen Sie die vorigen Aufgaben.