



Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

CAS Practical Machine Learning Introduction

Project: K-Means with Python

Prof. Dr. Jürgen Vogel (juergen.vogel@bfh.ch)

K-Means

K-Means via scikit-learn

- ▶ documentation
 - ▶ <http://scikit-learn.org/stable/modules/clustering.html#k-means>
 - ▶ <http://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>
- ▶ datasets
 - ▶ Iris
 - ▶ K = 3 (try different K too)
 - ▶ http://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_iris.html
 - ▶ handwritten digits
 - ▶ K = 10 (try different K too)
 - ▶ http://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_digits.html
 - ▶ alternatively: MNIST dataset (more samples and better resolution)
 - ▶ <https://www.kaggle.com/c/digit-recognizer>
- ▶ evaluation metrics
 - ▶ <http://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics.cluster>