

**Homework
in preparation of
Lecture 7
—
Feature Engineering**

Data Science

Prof. Dr. Markus Breunig

Photo by Pascal van de Vendel on Unsplash



Feature Engineering

Feature engineering is the process of using domain knowledge to extract features from raw data in order to improve the performance of machine learning models. We will be looking at the most important Feature Engineering techniques in the upcoming lecture.

In preparation, please recapitulate the following subjects covered in SVL und URL.

This homework assignment consists of the following parts:

- ◆ Recapitulate the material about Features (Categorical, Ordinal, Boolean, Numeric) from SVL (→ Chapter 2 – Types of Attributes)
- ◆ Recapitulate the material about k-means Clustering from URL (→ Chapter 0 – Introduction)
- ◆ Recapitulate the material about PCA from URL (→ Chapters 4 and 5)
- ◆ Complete the notebook "h07a..." - (expected time: **120min**)

