All of Machine Learning

A summary under eternal construction

Moritz Gück
github.com/MoritzGuck/All_of_ML-under_construction
last changes: March 12, 2020

Abstract

This is a reference for machine learning approaches and methods. The topics range from basic statistics to complex machine learning models and explanation methods. For each method and model, the underlying formulas (objective functions, prediction functions, etc.) are given, as well as code snippets from the respective python libraries. This reference should give data scientists a catalogue to find methods for their problem, refresh their knowledge and give references for further reading. If you find errors or unclear explanations in this text, please file an issue under
`github.com/MoritzGuck/All_of_ML-under_construction` .

Contents

1 Classification methods

1.1 Support Vector Machines (SVM)

2 Classification methods

2.1 Support Vector Machines (SVM)

3 Kernel methods

Kernel methods help in using linear decision boundaries on non-linearly separable data by morphing the feature space. They can also help in disentangling data for clustering. Kernels can incorporate domain knowledge into your model.

3.1 Introduction to Kernels

Kernels can be seen in two ways: As similarity measures between data-points or transformations of the data-points into a higher dimensional space. For two points x and y a Kernel is given by:

$$K(x,y) = \sum_{i=1}^{n} h_i(x)h_i(y) = \langle h(x), h(y) \rangle, \tag{1}$$

where $h(x)$ is a transformation-function and $\langle \cdot \rangle$ is the inner product.

3.2 Kernel SVM

4 Interpretability Methods

Complex machine learning algorithms (e.g. NNs) are hard/impossible to interpret. Interpretability methods help with debugging, trust and taking appropriate action on the results.

4.1 Local interpretability methods

Explain, why your model made this/these exact decisions.

4.1.1 Shapley values

Find attributes that determine the deviation of your output from the *expected value*. **How:** Calculate how much each feature pushes the prediction away from the expected value, by shuffling through all combinations of features having the sample-value or expected value respectively.

SHapley Additive exPlanations (SHAP) SHAP calculates Shapley values

4.1.2 Local interpretable model-agnostic explanations (LIME)

! $\rightarrow$ LIME interpretaions are not always consistent.

4.1.3 Example-based explanations

Influence Functions What would happen to the model parameters, if you would up-weight an instance? (model is function of training data.)

Index