

Artificial Intelligence

Albert-Ludwigs-Universität Freiburg



**UNI
FREIBURG**

Thorsten Schmidt

Abteilung für Mathematische Stochastik

www.stochastik.uni-freiburg.de

thorsten.schmidt@stochastik.uni-freiburg.de

SS 2017

Literature (incomplete, but growing):

- I. Goodfellow, Y. Bengio und A. Courville (2016). **Deep Learning**. <http://www.deeplearningbook.org>. MIT Press
- D. Barber (2012). **Bayesian Reasoning and Machine Learning**. Cambridge University Press
- R. S. Sutton und A. G. Barto (1998). **Reinforcement Learning : An Introduction**. MIT Press
- G. James u. a. (2014). **An Introduction to Statistical Learning: With Applications in R**. Springer Publishing Company, Incorporated. ISBN: 1461471370, 9781461471370
- T. Hastie, R. Tibshirani und J. Friedman (2009). **The Elements of Statistical Learning**. Springer Series in Statistics. Springer New York Inc. URL: <https://statweb.stanford.edu/~tibs/ElemStatLearn/>
- K. P. Murphy (2012). **Machine Learning: A Probabilistic Perspective**. MIT Press

Our goal today

Get to know Bagging, Bumping, Boosting and Random Forest!

Main content has been presented on blackboard.

See also chapters 8,10,14 and 15 in The Elements of Statistical Learning.

Diskreter Ada-Boost für Klassifizierung in 2 Klassen

Die zwei Klassen seien kodiert als -1 und 1 .

Algorithmus:

1. Initialisiere Gewichte $w_i = \frac{1}{N}$
2. Für $n = 1, \dots, M$
 - 2.1. Passe Klassifizierer $G_m(x)$ an die Trainingsdaten mit Gewichten w_i an und setze
 - i) $\text{err}_m = \frac{\sum_{i=1}^N w_i 1_{\{y_i \neq G_m(x_i)\}}}{\sum_{i=1}^N w_i}$
 - ii) $\alpha_m = \log\left(\frac{1 - \text{err}_m}{\text{err}_m}\right)$
 - iii) $w_i \leftarrow w_i \exp(\alpha_m \cdot 1_{\{y_i \neq G_m(x_i)\}})$
3. $G_{\text{AdaBoost}}(x) = \text{sign}\left[\sum_{m=1}^M \alpha_m G_m(x)\right]$

Algorithmus für Random Forest (Regression)

1. Für $b = 1, \dots, B$
 - a) Ziehe Bootstrap-sample Z^{*b} aus den Trainingsdaten Z
 - b) Lass einen Baum wachsen:
 - i) Wähle zufällig m aus p Prädiktoren aus
 - ii) Wähle den besten Split in 2 Teilgruppen unter diesen Prädiktoren aus
 - iii) Splitte dort
2. Der Random Forest besteht also aus B Regressions-Bäumen $\{T_b\}_{b=1, \dots, B}$
3. Sage vorher

$$\underbrace{\hat{f}_{r_f}^B(x) = \frac{1}{B} \sum_{b=1}^B T_b(x)}_{\text{Regression}}$$

$m = p \Rightarrow$ Random Forest entspricht Bagging.