



Machine Learning mit SVM

VON: CEVIN SIEPMANN, HENDRIK WEISGERBER, NIKLAS SCHÄFER,
DAVIT MELKONYAN, MOAYAD YAGHI

Gliederung

- ▶ SVM
- ▶ Lineare Regression
- ▶ (Aktien)
- ▶ Programm Vorstellung
- ▶ Quiz

Allgemeines

- ▶ Was ist Machine learning
- ▶ Methoden des Machine learnings
 - ▶ 1. Association Rule Learning
 - ▶ 2. Artificial Neural Networks
 - ▶ 3. Deep Learning
 - ▶ 4. Inductive Logic Programming
 - ▶ 5. Support Vector Machines
 - ▶ 6. Clustering
 - ▶ 7. Bayesian Networks
 - ▶ 8. Reinforcement Learning
 - ▶ 9. Representation Learning
 - ▶ 10. Similarity und Metric Learning

Support Vector Machine

- ▶ Allgemeines
- ▶ Mathematisches Grundprinzip
- ▶ Kernel

Allgemeines

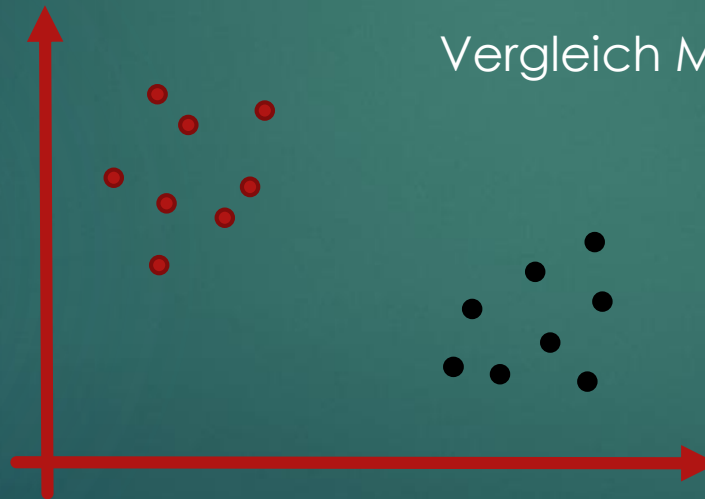
- ▶ Ab 1990 von Wladimir Naumowitsch Wapnik
- ▶ Zwei Klassen eines Graphen bestmöglich unterteilen
- ▶ Sog. Hyperplane in den Graphen einsetzen



<https://datascience.columbia.edu/vladimir-vapnik>

Y=Felldichte

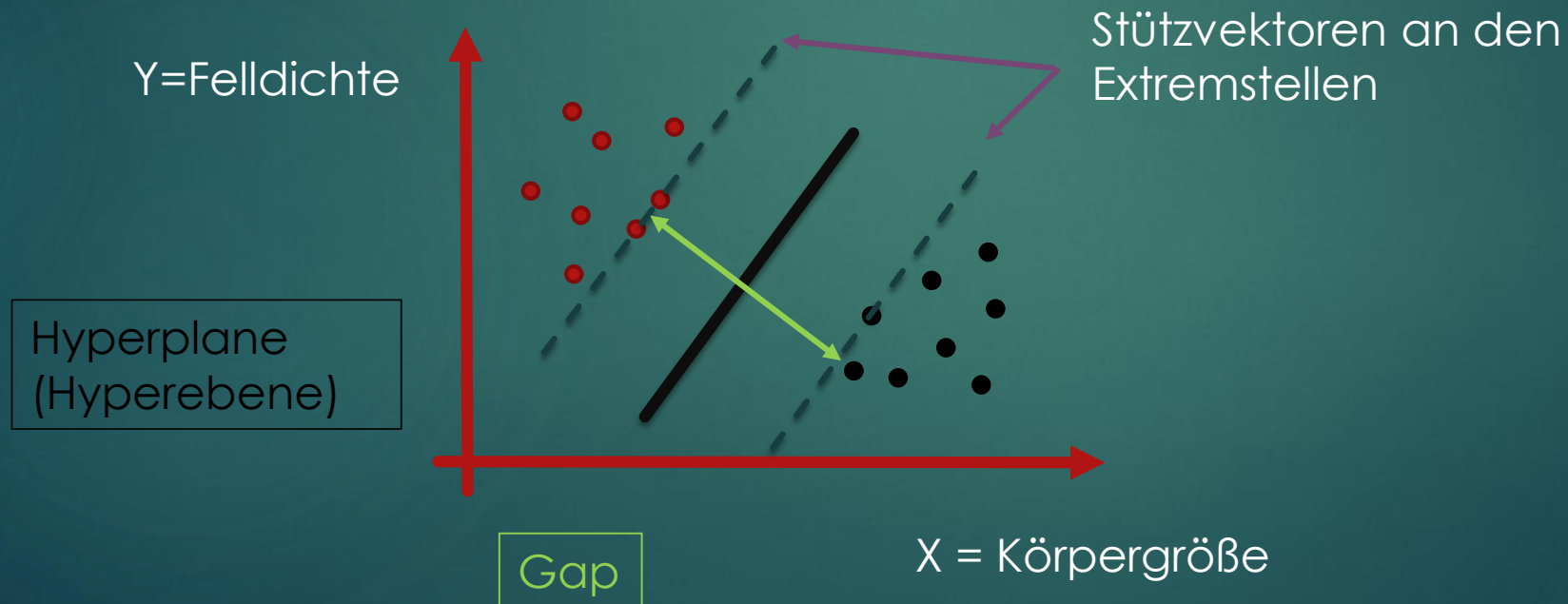
Vergleich Mensch zu Hund



X = Körpergröße

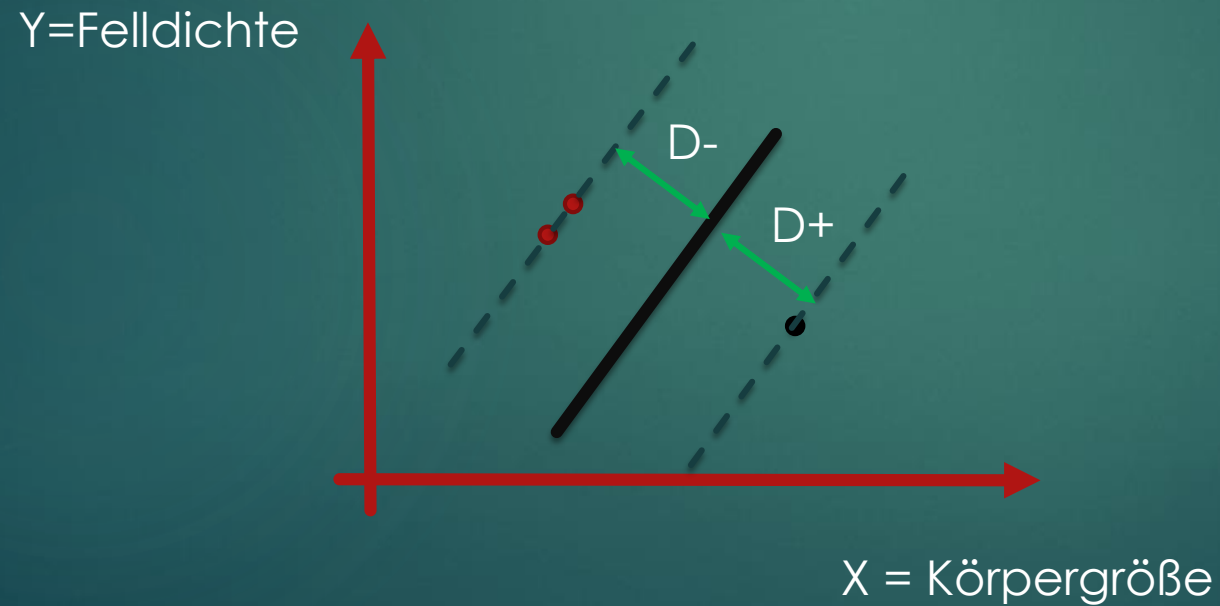
Mathematische Grundlagen

Linear Support Vector Machine (LSVM)

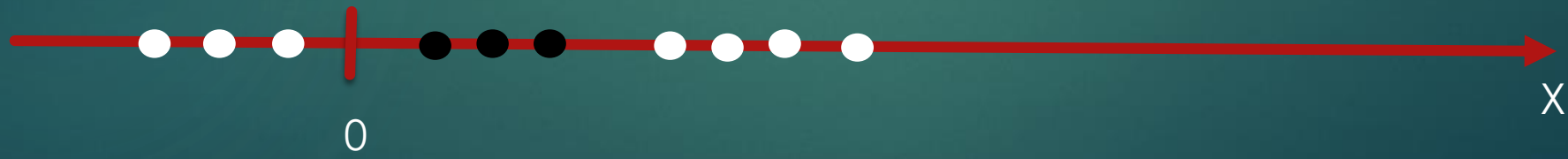


Mathematische Grundlagen

Linear Support Vector Machine (LSVM)

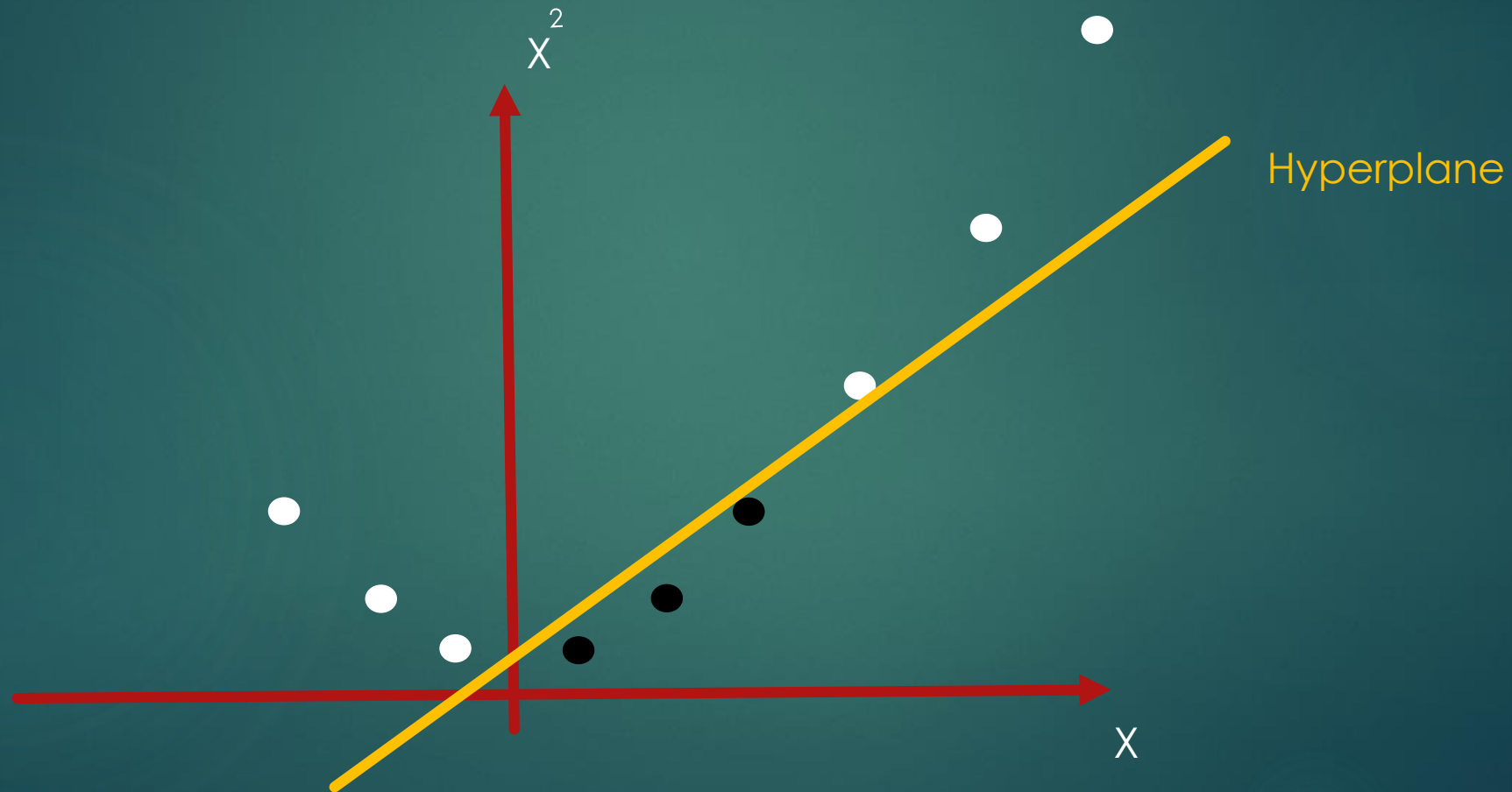


Ist das möglich ?



Mathematische Grundlagen

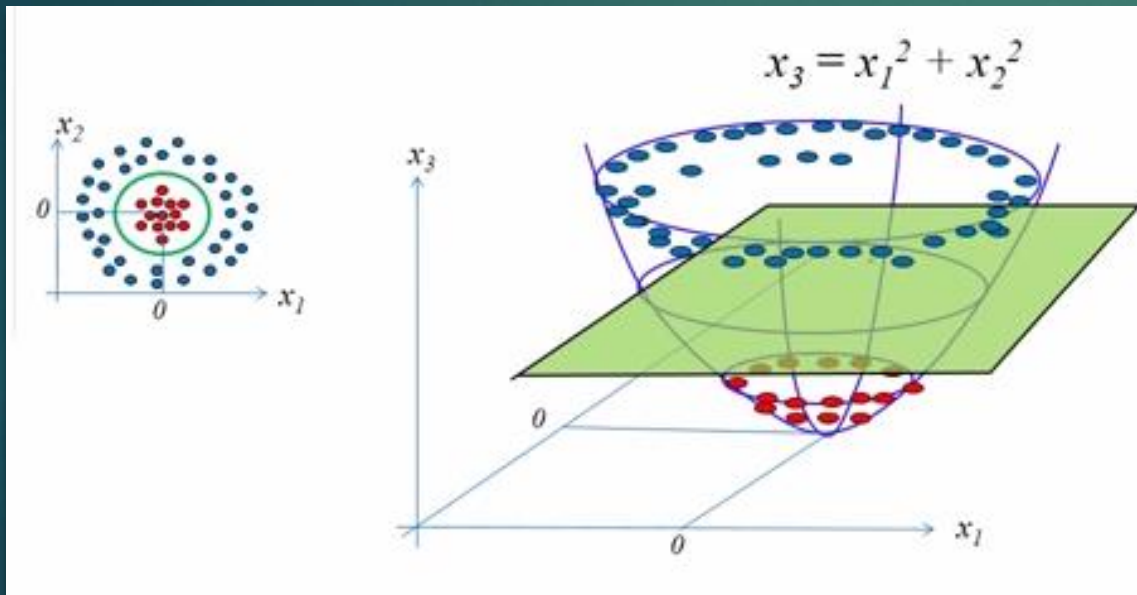
Nicht lineare Support Vector Machine



Kernels

2-Dimensional zu 3-Dimensional

- Möglich auch durch sogenannte Kernels-Funktion



<https://www.bogotobogo.com/python/scikit-learn/images/svm2/Non-linear-boundaries3.png>

Kernel

Problem ist :

- ▶ Muss für jede Rechnung angepasst werden

Positiv ist:

- ▶ Leicht zu lernen für Neuronales Netz
- ▶ Effektiv in hohen Dimensionen

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^p$$

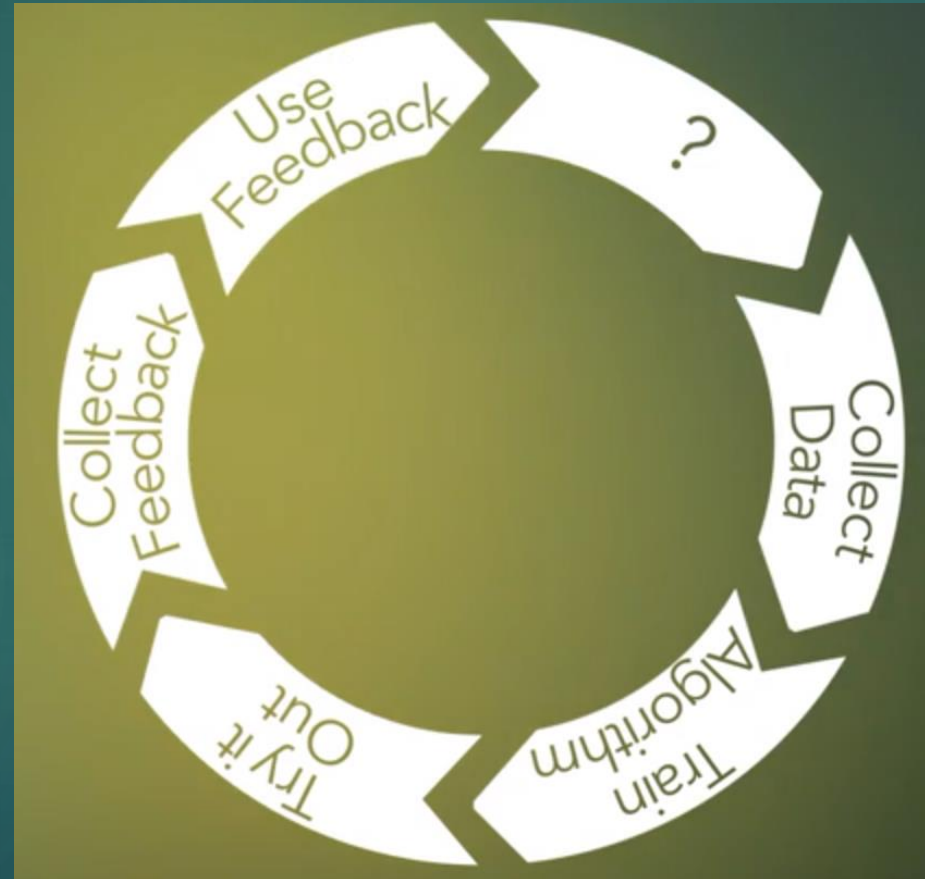
$$K(\mathbf{x}, \mathbf{y}) = e^{-\|\mathbf{x} - \mathbf{y}\|^2 / 2\sigma^2}$$

$$K(\mathbf{x}, \mathbf{y}) = \tanh(\kappa \mathbf{x} \cdot \mathbf{y} - \delta)$$

Lineare Regression

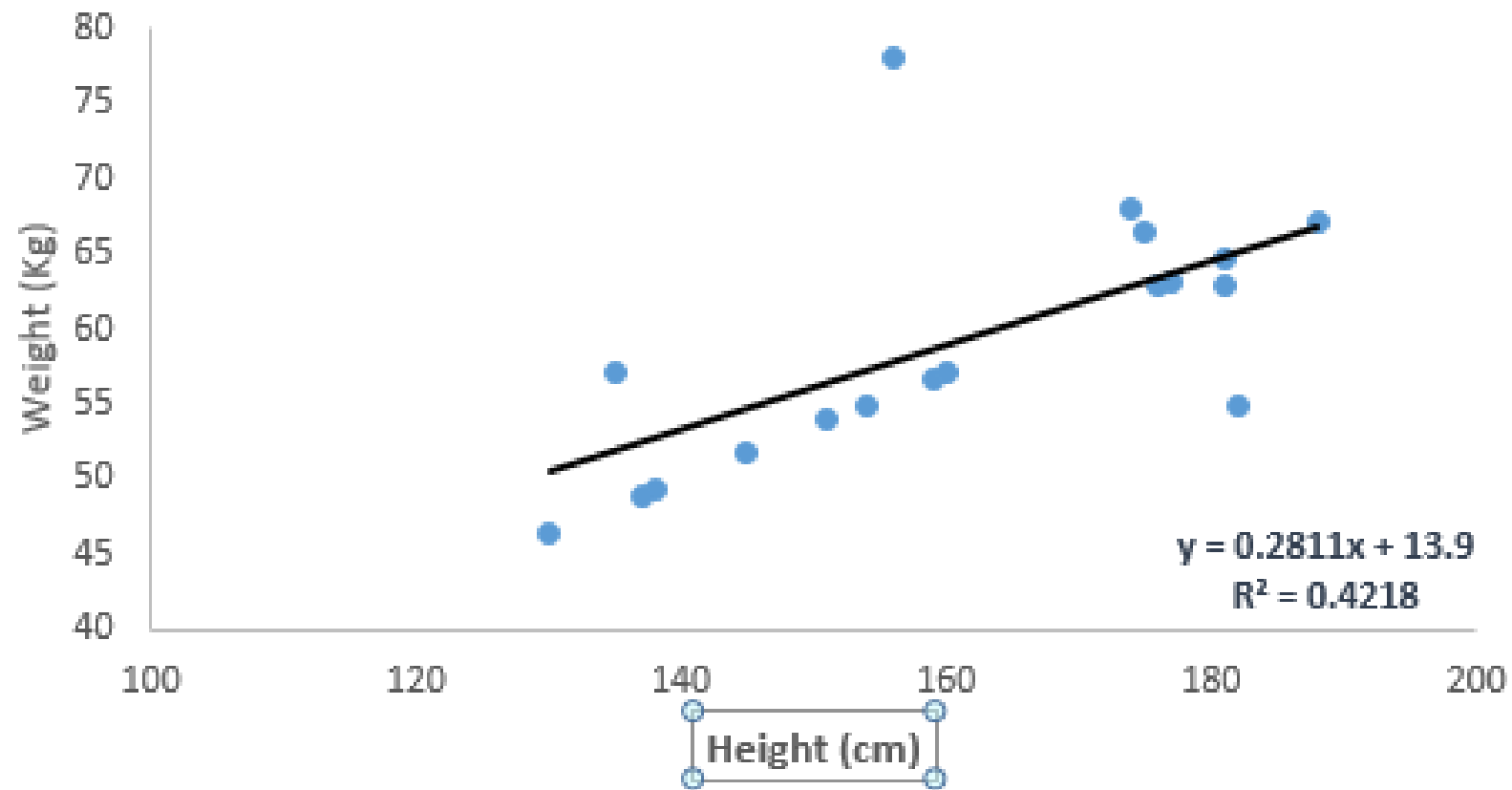
- ▶ Lebenszyklus Maschinelles lernen
- ▶ Beispiel der Methode „Lineare Regression“
- ▶ Fehler und Bias
- ▶ Loss function
- ▶ R^2

Maschinelles lernen Lebenszyklus



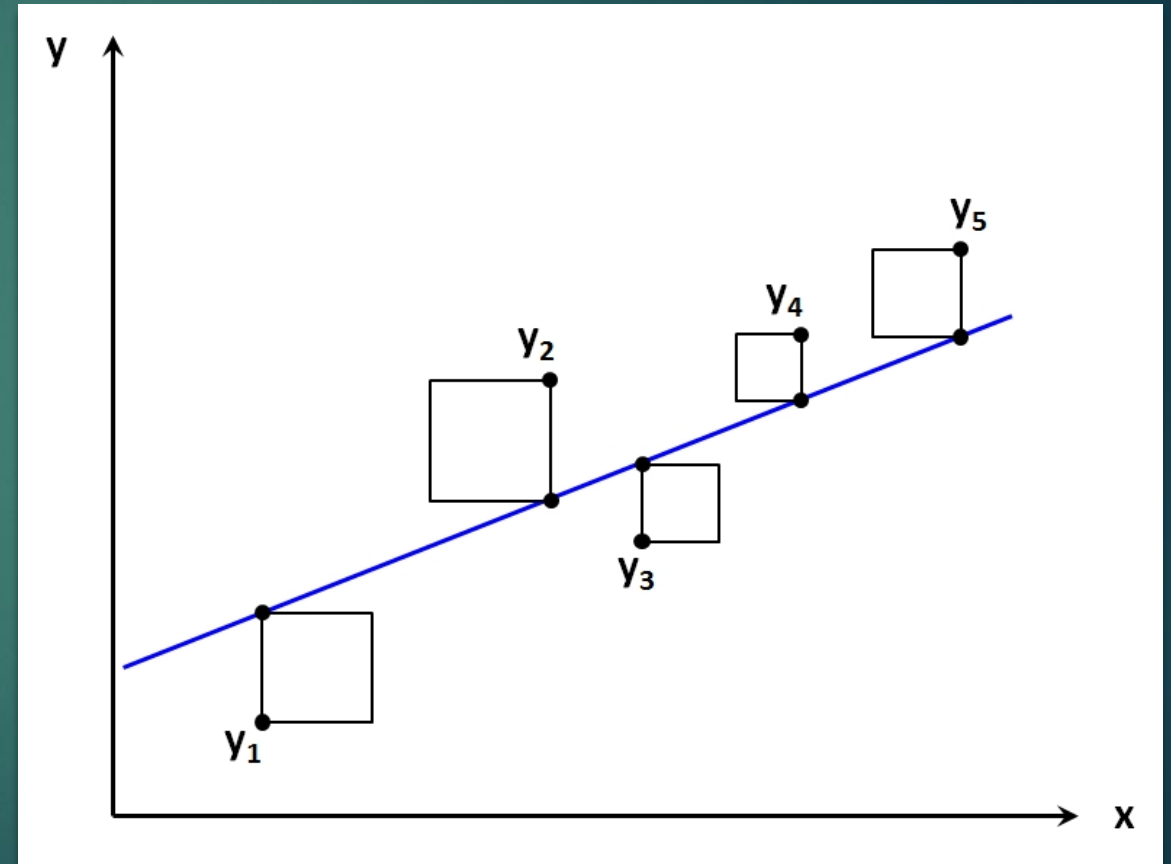
Beispiel

Relation B/w Weight & Height



Fehlerquellen und Optimierung

- ▶ Was sind Fehler?
- ▶ Warum quadriert man diese?
- ▶ Was ist der Bias?
- ▶ Hold out Data?



Loss function

- ▶ Mean Squared Error Loss (MSE)
 - ▶ $L = \frac{1}{N} \sum_{i=1}^n (y(i) - (mx(i) + b))^2$
- ▶ Mean Squared Logarithmic Error Loss (MSLE)
 - ▶ $L = \frac{1}{n} \sum_{i=1}^n (\log(y(i) + 1) - \log(\hat{y}(i) + 1))^2$
- ▶ Mean Absolute Error (MAE)
 - ▶ $L = \frac{1}{n} \sum_{i=1}^n ||y(i) - \hat{y}(i)||$

R²

- ▶ Erklärung der Varianz der abhängigen Variable
- ▶ Varianz: Empfindlichkeit auf Schwankung
- ▶ Zwischen 0 und 1

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\text{erklärte Variation}}{\text{Gesamtvariation}}$$

oder

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\text{unerklärte Variation}}{\text{Gesamtvariation}}$$

- ▶ Kritikpunkt: zu hohe Werte für R²

Korrigierte R^2

- ▶ Das korrigierte R^2

$$R^2_{\text{kor}} = 1 - (1 - R^2) \cdot \frac{n - 1}{n - p - 1}$$

- ▶ Je höher das korrigierte R^2 desto besser
- ▶ Kann negativ werden
- ▶ Guter Wert für R^2 ?

Programmvorstellung

$$y = mx + b$$

$$b = \bar{y} - m\bar{x}$$

$$m = \frac{\bar{X} \cdot \bar{Y} - \overline{XY}}{(\bar{X})^2 - \overline{x^2}}$$

Quiz

Link zum Quiz:

▶ joinmyquiz.com

Game Code:

Quellen

- ▶ <https://blog.nxtgn.de/maschinelles-lernen-verfahren>
- ▶ <https://data-science-blog.com/blog/2017/12/20/maschinelles-lernen-klassifikation-vs-regression/>
- ▶ <https://www.educba.com/loss-functions-in-machine-learning/>
- ▶ <https://www.crashkurs-statistik.de/einfache-lineare-regression/>
- ▶ https://medium.com/@lachlanmiller_52885/understanding-and-calculating-the-cost-function-for-linear-regression-39b8a3519fcb
- ▶ <https://pythonprogramming.net/machine-learning-tutorials/>