Basislerner

Autor und Herausgeber

Alexander Jung*, Department of Computer Science, Aalto University, Espoo, Finland (ORCID 0000-0001-7538-0990)
Editors: Konstantina Olioumtsevits (ORCID 0000-0003-2697-8659), Ekkehard Schnoor (ORCID 0000-0001-8252-1499), Aalto University, Espoo, Finland
*Corresponding author: alex.jung@aalto.fi

Zusammenfassung

Ein Basislerner ist die Methode des maschinellen Lernens, die ein Ensemble-Verfahren mehrfach ausführt: ein Algorithmus, der einen Trainingsdatensatz einliest und eine gelernte Hypothese zurückgibt. Das Ensemble-Verfahren kombiniert die gelernten Hypothesen seiner Basislerner zu einer einzigen Hypothese. Die wichtigsten Ansätze unterscheiden sich darin, wie sie die Trainingsdatensätze erstellen und wie sie die Ergebnisse kombinieren. Bagging trainiert seine Basislerner parallel auf mittels Bootstrap gezogenen Stichproben, während Boosting sie nacheinander trainiert, jeweils auf einem Trainingsdatensatz, der durch die früheren Ergebnisse modifiziert wird. Stacking behandelt ihre Vorhersagen stattdessen als eine Merkmalstransformation und lernt eine Aggregationsregel auf den transformierten Merkmalen.

Definition

Ein Ensemble, das die morgige Höchsttemperatur in Krems an der Donau vorhersagt, mittelt die Vorhersagen eines linearen Modells, eines Entscheidungsbaums und eines künstlichen neuronalen Netzes, die jeweils für sich aus demselben Trainingsdatensatz gelernt wurden. Jedes der drei ist ein Basislerner: eine Methode des maschinellen Lernens, die innerhalb eines Ensembles eingesetzt wird. Formal handelt es sich um einen Algorithmus $\mathcal{A}$, der einen Trainingsdatensatz $\mathcal{D}^{(\text{train})}$ einliest und eine gelernte Hypothese $\widehat{h} = \mathcal{A}(\mathcal{D}^{(\text{train})})$ erzeugt. Ein Ensemble-Verfahren führt mehrere Basislerner aus und kombiniert deren gelernte Hypothesen zu einer geeigneteren Hypothese.

Die wichtigsten Ansätze unterscheiden sich darin, wie sie die Trainingsdatensätze konstruieren und die Ergebnisse kombinieren. Bagging trainiert Basislerner parallel auf mittels Bootstrap erzeugten Stichproben (Breiman, 1996). Boosting trainiert sie sequenziell und nutzt frühere Ergebnisse, um den Trainingsdatensatz für den nächsten Basislerner zu modifizieren. Stacking verwendet ihre Vorhersagen als eine Merkmalstransformation und lernt eine Aggregationsregel aus den transformierten Merkmalen (Hastie et al., 2009, Kap. 8; Wolpert, 1992).

Siehe auch: Ensemble, Bagging, Stacking, Boosting.

Querverweise

Literatur

1. Breiman (1996). Bagging predictors. Mach. Learn., 24(2), pp. 123–140.
2. Wolpert (1992). Stacked generalization. Neural Netw., 5(2), pp. 241–259.
3. Hastie et al. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed., Springer Science+Business Media, New York, NY, USA.