



Glossar

Aktivierungsfunktion: Bestimmt, ob ein Neuron aktiviert ist und ein Signal aussendet.

AUC (Area Under the Curve): Metrik zur Bewertung der Leistung eines Klassifikators anhand der Fläche unter der ROC-Kurve.

Backpropagation: Algorithmus, der die Gewichte der Neuronen in Abhängigkeit von dem erzeugten Fehler des Gesamtmodells anpasst.

Bagging: Ensemble-Methode, bei der mehrere Modelle parallel trainiert werden und ihre Ergebnisse kombiniert werden, um die Vorhersagegenauigkeit zu verbessern. Ein bekanntes Beispiel ist der Random Forest-Algorithmus.

Boosting: Ensemble-Methode, bei der Modelle sequenziell trainiert werden, wobei jedes Modell versucht, die Fehler des vorherigen Modells zu korrigieren. Beispiele sind Adaptive Boosting und Gradient Boosting.

Deep Learning: Teilgebiet des Machine Learning, das sich auf die Anwendung neuronaler Netze mit vielen verdeckten Schichten ("Hidden Layers") konzentriert. Ermöglicht die Darstellung komplexer Zusammenhänge in Daten.

Entscheidungsbaum: Modell, das eine baumartige Struktur verwendet, um Entscheidungen zu treffen, indem es Daten anhand von Entscheidungsregeln in verschiedene Kategorien aufteilt.

Feature: Input-Variable, die zur Vorhersage der Zielgröße verwendet wird.

Gradient Boosting: Ensemble-Methode, die Entscheidungsbäume sequentiell trainiert, wobei jeder Baum die Fehler des vorherigen Baums korrigiert.

Grid-Search: Verfahren zur Optimierung der Hyperparameter eines Modells durch systematisches Ausprobieren verschiedener Parameterkombinationen.

Klassifikation: Modell, das Daten in diskrete Kategorien einteilt, z. B. "Regentag" oder "kein Regentag".

Kreuzvalidierung: Verfahren zur Evaluation der Vorhersagegenauigkeit eines Modells durch Aufteilung des Datensatzes in mehrere Teilmengen (z.B. k-Fold-Kreuzvalidierung).

Lernrate: Parameter, der steuert, wie stark die Gewichte des Neuronalen Netzes bei jeder Iteration angepasst werden.

Machine Learning: Entwicklung von Algorithmen und Modellen, die selbstständig Muster in Daten finden und daraus lernen. Es wird eine allgemeine Regel aus den Daten abgeleitet, ohne explizite Anweisung durch den Programmierer.

Modell: Stellt den Zusammenhang zwischen Features und der Zielgröße her.

Modell-Evaluation: Verfahren, um die Vorhersagegenauigkeit eines Modells zu bestimmen.

Neuronale Netze: Bestehen aus miteinander verbundenen Knoten ("Neuronen") in Schichten, die Eingabedaten verarbeiten und Informationen weitergeben. Sie zeichnen sich durch ihre Fähigkeit aus, komplexe, nichtlineare Beziehungen zu erkennen.

Overfitting: Das Modell lernt die Trainingsdaten "auswendig" und kann neue Daten nicht gut vorhersagen.

Propagierungsfunktion: Funktion, die innerhalb eines Neurons alle Eingabesignale verdichtet.

Glossar - Teil 3

Random Forest: Ensemble-Methode, die mehrere Entscheidungsbäume parallel trainiert und deren Vorhersagen kombiniert.

Regression: Modell, das einen kontinuierlichen Wert vorhersagt, z. B. die Temperatur.

Reinforcement Learning: "Bestärkendes Lernen" - Algorithmen lernen durch Interaktion mit einer Umgebung und erhalten Belohnungen für erfolgreiche Aktionen.

ReLU (Rectified Linear Unit): Aktivierungsfunktion, die für positive Eingabewerte den Wert selbst zurückgibt und für negative Eingabewerte Null

ROC-Kurve (Receiver Operating Characteristic Curve): Grafische Darstellung der Sensitivität und Spezifität eines Klassifikators bei verschiedenen Schwellenwerten.

Supervised Learning: "Überwachtes Lernen" - Der Algorithmus lernt aus Daten, bei denen das korrekte Ergebnis bereits bekannt ist.

Testdaten: Datensatz, der verwendet wird, um die Performance des Modells auf unbekannten Daten zu evaluieren.

Trainingsdaten: Datensatz, der zum Trainieren des Modells verwendet wird.

Unsupervised Learning: "Unüberwachtes Lernen" - Der Algorithmus lernt aus Daten, bei denen kein vorgegebenes Ergebnis existiert, z.B. bei der Clusteranalyse oder Anomalieerkennung.

Vanishing Gradient Problem: Problem in Neuronalen Netzen, bei dem die Gradienten bei der Backpropagation in den frühen Schichten immer kleiner werden, was das Lernen des Modells behindert.

XGBoost (Extreme Gradient Boosting): Erweiterung von Gradient Boosting mit zusätzlichen Features zur Verbesserung der Leistung und Geschwindigkeit.