

DATA MINING

Definition, Methoden und Tools

Data Mining bezeichnet die explorative Analyse von großen Datenmengen mittels Statistik und maschinellen Lernen um unbekannte Muster und Trends festzustellen. Während Data

Science den gesamten Datenverarbeitungsprozess abdeckt, fokussiert sich Data Mining als Aufgabe konkret auf das Identifizieren neuer Information in vorliegenden Daten.

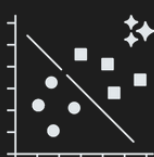
METHODEN

STATISTIK



Häufigkeiten, Relationen und Interaktionen erkennen.

CLUSTERING



Identifizieren von ähnlichen Einträgen und deren Relation.

ASSOZIATIONEN



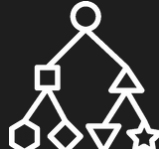
Quantifizieren von zusammen auftretenden Ereignissen.

OUTLIERS



Detektion von ungewöhnlichen Datensätzen.

KLASSIFIKATION



Die Vorhersage von Gruppenzugehörigkeiten.

REGRESSION



Die Vorhersage von numerischen Werten.

www.datadrivencompany.de

TOOLS

Coding



python



R



KNIME



RapidMiner



SPSS

Visualisierung



PowerBI



Tableau



Google Data Studio

Anderes



Excel



Wusstest du? Nur 25% seiner Zeit verbringt ein Data Scientist mit Datenanalysen wie Data Mining. Den Rest mit Datenakquise und Vorbereitung der Daten.