
Einführung

WS 2023-24

DI Emil Marinov | DI David Bechtold

Übersicht

1.	Überblick Data Science	3 – 8
2.	Grundbegriffe Statistik	9 – 16

Überblick

Überblick

Was ist Data Science?

- Prozess der Gewinnung neuer Erkenntnisse unter Verwendung von datenbasierten Methoden
- Unterstützung von Geschäftsentscheidungen in Unternehmen mit dem Ziel der Steigerung von Effizienz, Umsatz und Produktivität

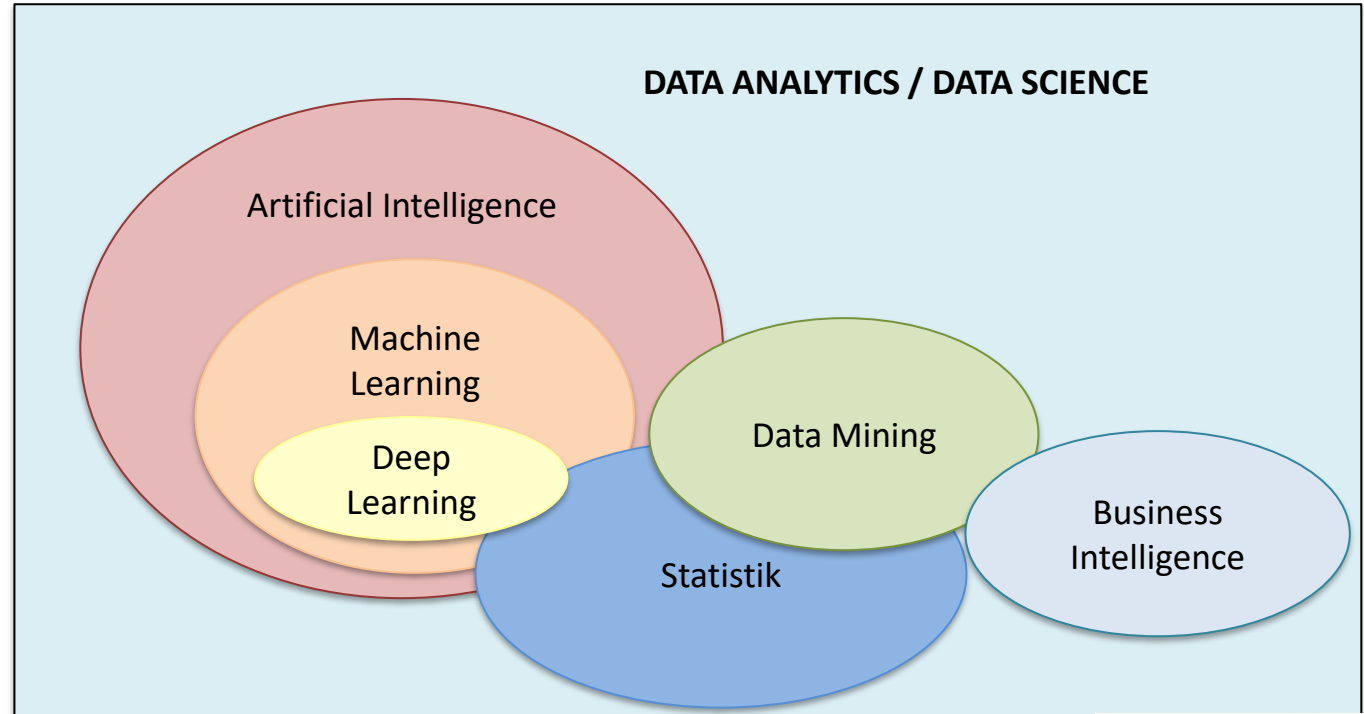


„Führende Unternehmen in allen Branchen nutzen Daten und Analysen, um sich einen Wettbewerbsvorteil zu erschaffen“

Gartner

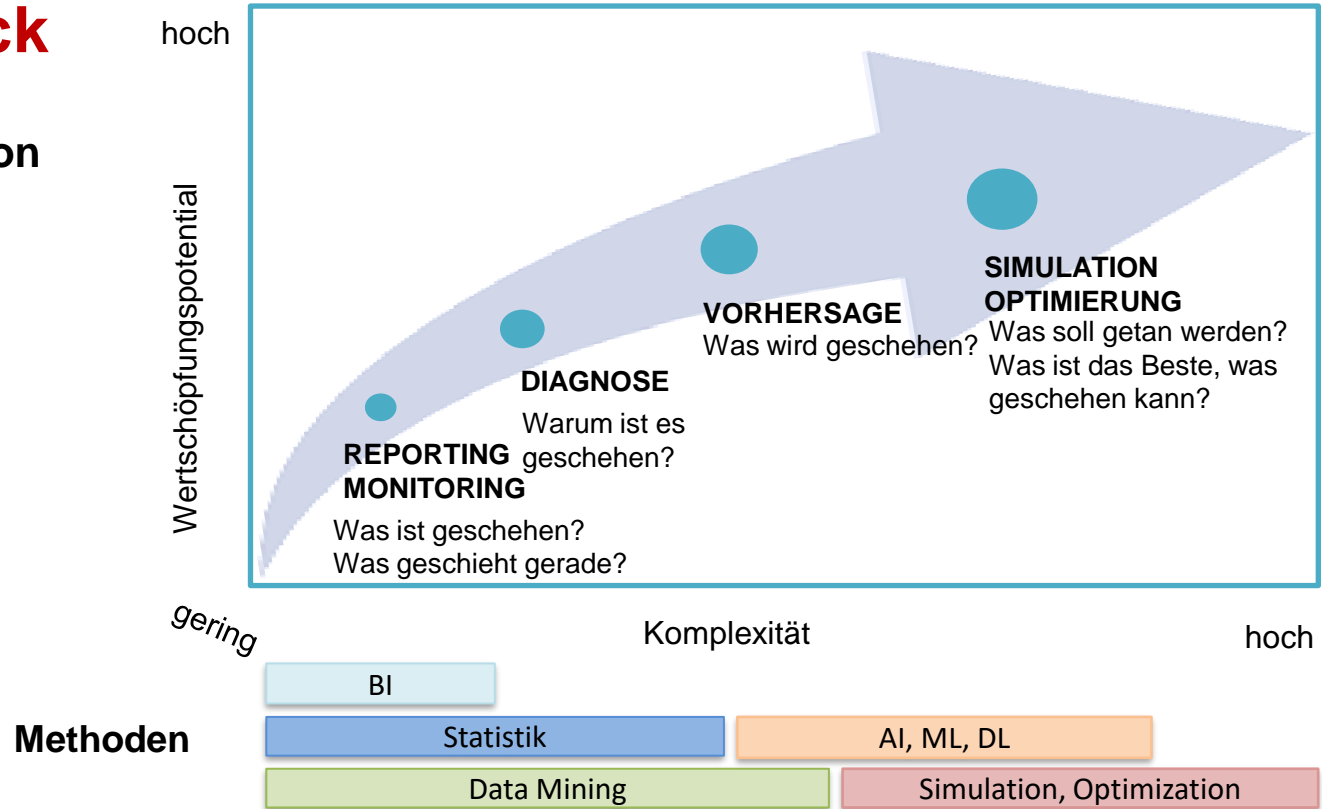
Überblick

**Welche
Methoden
werden
eingesetzt?**



Überblick

Spektrum von Analysen



Überblick

Data Analytics bei ...

Überblick

Teilgebiete der Statistik

1. Beschreibende (deskriptive) Statistik:

- > Beschreibung der erhobenen Daten mit Tabellen, Grafiken, Kennzahlen
- > Es liegen vollständige Daten über eine Gesamtheit vor.

2. Wahrscheinlichkeitsrechnung:

- > Bewertung von zufälligen Ereignissen
- > Grundlage für schließende Statistik

3. Schließende (induktive) Statistik:

- > Es liegen nur Daten einer ausgewählten Stichprobe der Gesamtheit vor.
- > Überprüfung von Hypothesen für die Grundgesamtheit.

Grundbegriffe der Statistik

Grundbegriffe

- **Daten:** Sammlung von Datenobjekten und deren Attributen
- Ein **Attribut** ist eine Eigenschaft oder ein **Merkmal** eines Objekts.
- Eine Sammlung von Attributen beschreibt ein **Objekt (Merkmalsträger)**.
Ein Objekt wird auch als Datensatz, Punkt, Fall, Beispiel, Entität oder Instanz bezeichnet.
- **Größe** = Anzahl der Objekte
- **Dimensionalität** = Anzahl der Attribute

Merkmale / Attribute

FS	Percentage Setup	CRL Factor	Utilization Factor
1	0.050	1.500	0.750
2	0.050	1.500	0.850
3	0.050	1.500	0.950
4	0.050	3	0.750
5	0.050	3	0.850
6	0.050	3	0.950
7	0.100	1.500	0.750
8	0.100	1.500	0.850
9	0.100	1.500	0.950
10	0.100	3	0.750
11	0.100	3	0.850
12	0.100	3	0.950

Objekte

Grundbegriffe

Merkmale / Attribute

= Eigenschaften von Merkmalsträgern / Objekten

Beispiele:

Merkmalsträger	Merkmal
Angestellte einer Firma	Jahreseinkommen
Studenten	Lerndauer pro Woche
Unternehmen	Aktienkurs

Grundbegriffe

Merkmalsausprägungen

= mögliche Werte eines Merkmals

Beispiele:

Merkmal	Merkmalsausprägungen
Geschlecht:	männlich, weiblich
Lerndauer pro Woche	5 h, 10 h, 17 h, ...
Zufriedenheit mit Einkommen:	Sehr zufrieden, wenig zufrieden, nicht zufrieden

Grundbegriffe

Merkmaltypen

Unterscheidung nach **Art** (quantitativ / qualitativ) und **Anzahl** (diskret / stetig) der Ausprägungen

- **qualitativ:** Ausprägungen beschreiben die Qualität, Angabe durch Namen
- **quantitativ:** Ausprägungen geben Intensität wieder, Angabe durch Zahlen
- **diskret:** endlich oder abzählbar unendlich viele Ausprägungen
- **stetig:** alle Werte eines Intervalls sind mögliche Ausprägungen

Grundbegriffe

Merkmalstypen

Beispiele:

Merkmal	Merkmalstyp
Qualitätsklasse eines Produkts	qualitativ, diskret
Bearbeitungszeit eines Bauteils	quantitativ, stetig
Anzahl der Mitarbeiter	quantitativ, diskret

Grundbegriffe

Messskalen

- **nominal:** Werte sind Namen
- **ordinal:** Ordnungsbeziehung, Abstand nicht sinnvoll
- **metrisch / numerisch:** Ordnung und Abstand sinnvoll
 - > **intervallbasiert:** Verhältnis nicht sinnvoll, Nullpunkt willkürlich
 - > **verhältnisbasiert:** Verhältnis sinnvoll, natürlicher Nullpunkt

Skala	Beispiele
nominal	Maschinentyp, Losgrößenregel
ordinal	Qualitätsklasse, Vorgangsnummer
metrisch – intervallbasiert	Liefertermin, Umgebungstemperatur
metrisch – verhältnisbasiert	Länge, Durchlaufzeit

Kontrollfragen

Thema	<i>Statistische Grundbegriffe</i>
Fragen	<ol style="list-style-type: none">1. Geben Sie für folgende Merkmale mögliche Merkmalsträger und Ausprägungen an: Lieblingssportart, Baujahr, Haushaltseinkommen, Liefertreue2. Welche Merkmale sind diskret, welche stetig? Warmwasserverbrauch, Ausschussstücke pro Tag, Bierkonsum in Flaschen, Bierkonsum in Liter3. Ordnen Sie die folgenden Merkmale den Merkmalstyp (diskret – stetig) und das Skalenniveau (nominal – ordinal – metrisch) zu: Zigarettenkonsum, Bauteilgewicht, Wohnfläche, Verkaufspreis