



Die Welt der Algorithmen Relevanz und Risiko

Direktorenkonferenz der Landesmedienanstalten

Prof. Dr. Katharina A. Zweig

Algorithmus? Logarithmus?



WirtschaftsWoche [UNTERNEHMEN](#) [FINANZEN](#) [POLITIK](#) [ERFOLG](#) [TECHNOLOGIE](#)

[Banken](#) [Dienstleister](#) [Energie](#) [Industrie](#) [IT](#) [Handel](#) [Versicherer](#) [Mittelstand](#) [Auto](#)

DAX € 9.757,88 -0,41%	E-STDXX 60S 3.021,01 -0,78%	MDAX € 19.473,44 -0,02%	Dow Jones 16.449,18 -0,10%	Gold (USD) 1.128,50 +0,83%	EUR/USD 1,0890 +0,66%	■ Börsenkurse ■ € Indikatoren
-----------------------------------	--------------------------------	-------------------------------------	-------------------------------	-------------------------------	--------------------------	---

Die WirtschaftsWoche > Unternehmen > IT > Apple übernimmt Beats: Jimmy Iovine: Der Drei-Milliarden-Dollar-Mann

Thema [Apple >](#)

Apple übernimmt Beats

Jimmy Iovine: Der Drei-Milliarden-Dollar-Mann

30. Mai 2014



Er soll Apple fürs Musikbusiness fit machen: Plattenboss Jimmy Iovine.

Bild: REUTERS

0 Kommentare

Artikel teilen

- [Facebook](#)
- [Twitter](#)
- [LinkedIn](#)

Versenden Drucken Merken Startseite

Oder gar: **Algorithmus?**

„Menschen werden für Services zahlen (...). Aber sie werden nicht für den reinen Zugang zahlen.“ Daher dürfte Streaming nicht auf **persönlichen Logarithmen** und automatischen Empfehlungen basieren.



Was ist ein Algorithmus wirklich?

Ein Algorithmus ist ein Problemlöser

Problem



INPUT

**Der OUTPUT
der uns sagt,
wie Input
mit Output
zusammenhängt.**



OUTPUT

Input: By User:Bluemoose - Own work, [CC BY-SA 3.0](#)

Putput: By Yann (talk) - Own work, GFDL

Output: [CC BY-SA 3.0](#)

Ein Algorithmus ist...



...eine für jede **erfahrene Programmiererin** und jeden erfahrenen Programmierer **ausreichend detaillierte Lösungsvorschrift**, so dass bei **korrekter Implementierung** der Computer **für jede korrekte Inputmenge den korrekten Output** berechnet – in endlicher Zeit.



Beispiel

Problem: Sortieren



Sortieren 1: „Sortieren durch Einfügen“



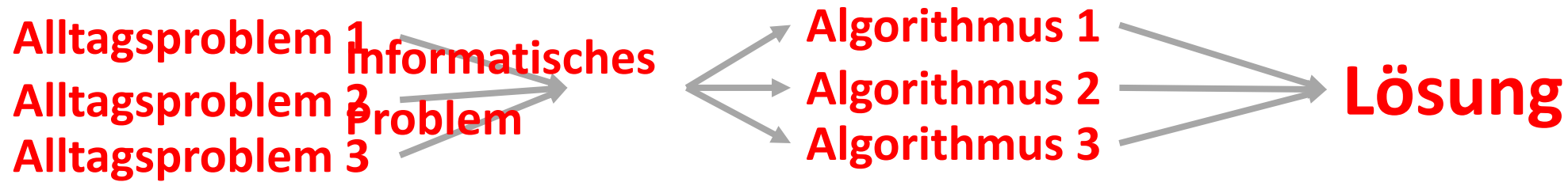
- Fange mit einem Buch an, stelle es ins Regal.
- Solange es noch Bücher gibt,
 - nimm das nächste,
 - geh am Regal entlang und sortiere es an der passenden Stelle ein.
- Alle Bücher, die schon im Regal stehen, sind in der richtigen, relativen Reihenfolge.
- Daher: wenn alle im Regal stehen, sind sie vollständig sortiert.

Sortieren 2: Aufsteigendes Sortieren



- Stelle alle Bücher irgendwie ins Regal.
- Gehe das Regal entlang – wenn dabei zwei Bücher in der falschen Reihenfolge nebeneinander stehen, vertausche sie. Tue dies bis zum Ende des Regals und gehe wieder zum Anfang.
- Laufe solange immer wieder am Regal entlang, bis im letzten Durchgang kein Tausch mehr nötig war.
- Wenn kein Tausch mehr nötig war, sind alle Bücher sortiert.

Problem-Algorithmus-Problem



- Ein informatisches Problem kann also meist durch mehrere Algorithmen gelöst werden.
- Jeder Algorithmus löst nur genau ein (informatisches) Problem.
- Im Sinne von „Alltagsproblemen“ löst derselbe Algorithmus sehr viele verschiedene Probleme:
 - Sortieren von Rechnungen nach Zahlfristende;
 - Suchmaschineneinträge sortieren nach Bewertung durch Suchmaschinenalgorithmus;
 - Produkte nach Bewertung, Preis oder Verkaufsrang anzeigen, etc.

Was kann schiefgehen?



Der Algorithmus löst das Problem nicht.





Komplexe Algorithmen

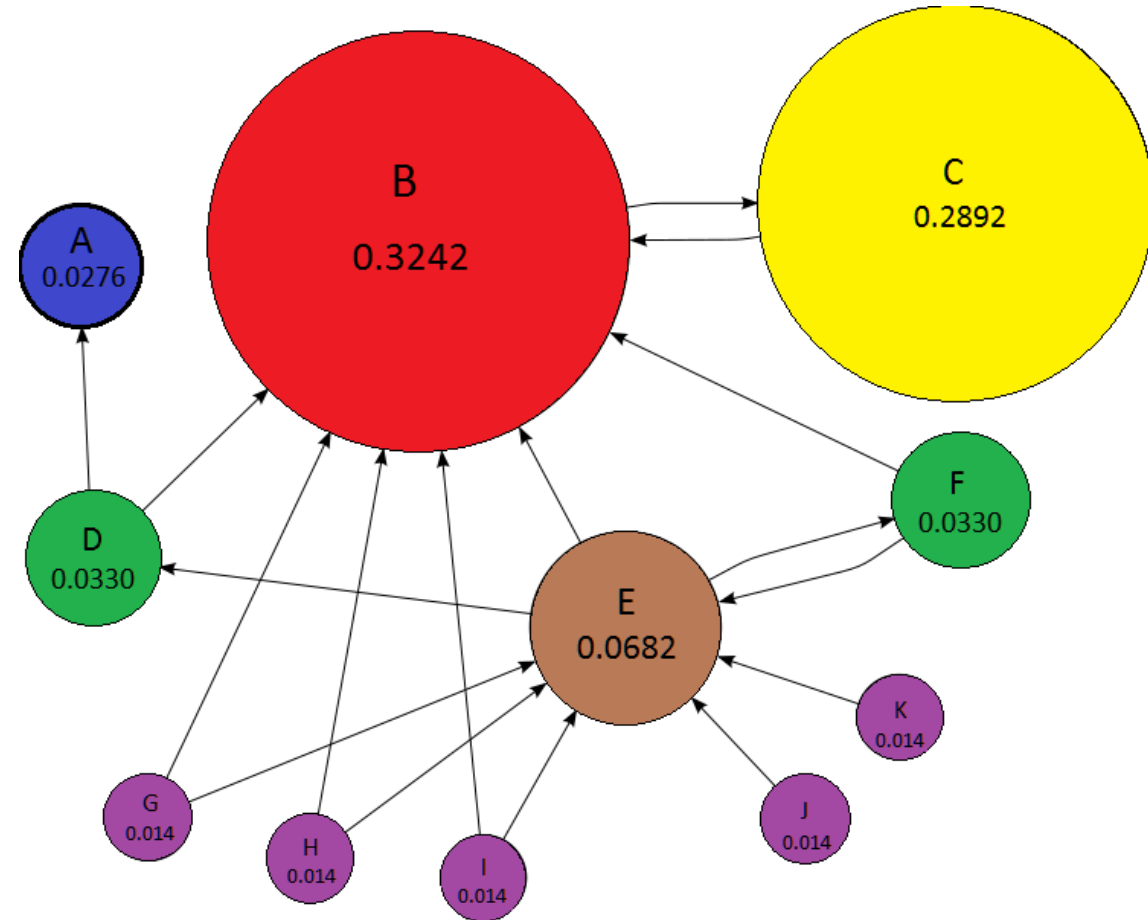
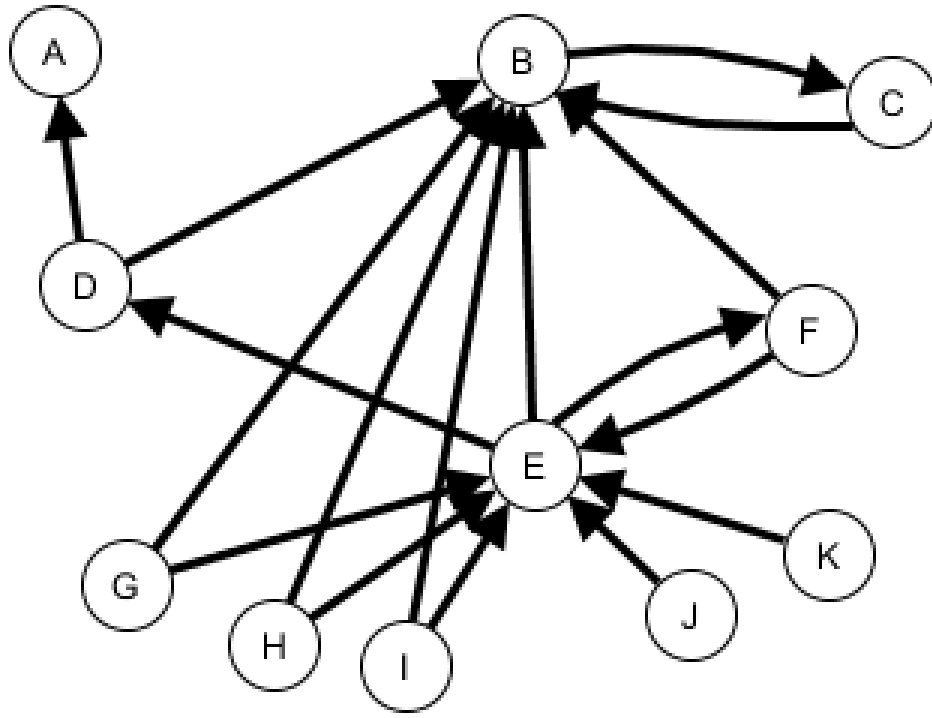
Beispiel: Suchmaschinenalgorithmen

Suchmaschinen 101



1. Filtern aus allen ihnen bekannten Webseiten diejenigen, deren Text mit den angegebenen Suchbegriffen zusammenhängen.
2. Sortieren diese anhand:
 - Der Vernetzungsstruktur der Seiten untereinander
 - Dem Clickverhalten anderer Nutzer und Nutzerinnen bezüglich derselben Suche
 - Bei Personalisierung: auch nach dem eigenen, bisherigen Suchverhalten

PageRank



Modellierungsannahme



- Gibt nur dann relevante Ergebnisse, wenn Webseiten
 - Links auf ähnliche Seiten wie ihre eigene setzen,
 - Links auf relevante, meinungsangebende Seiten setzen, und
 - ihre Links **unabhängig** voneinander setzen.
- Unter dieser Bedingung ist der Algorithmus neutral und gibt das kollektive Wissen der Welt nutzbringend weiter.
- Die Veröffentlichung des Algorithmus führte prompt zu Manipulationen seitens der Webseitenbetreiber.
 - Zu große Offenheit der Algorithmen ist manchmal **schädlich**.

Was kann schiefgehen?



Der Algorithmus löst das Problem nicht.

Die Modellierungsannahmen eines Algorithmus gelten nicht (mehr).



Lernende Algorithmen

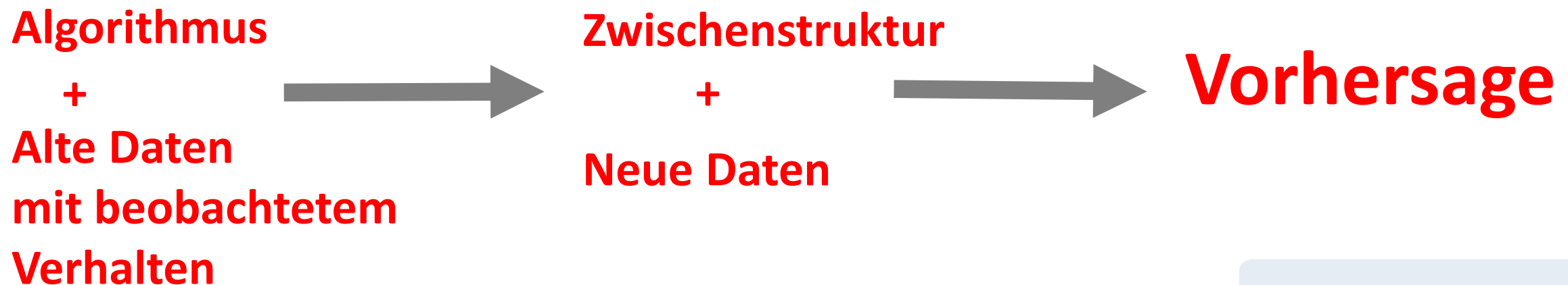
Lernende Algorithmen



Künstliche Intelligenz



- **Problem:** gegeben eine Menge von bekannten Daten, finde Muster, die auf neuen Daten vorhersagen, wie sich etwas oder jemand verhalten wird.
- Algorithmus baut – basierend auf bekannten Daten – eine Zwischenstruktur auf, die dann Vorhersagen für neue Daten generiert.
- Der Algorithmus wird „auf den Daten trainiert“.



Beispiel Suchmaschine (fiktiv)

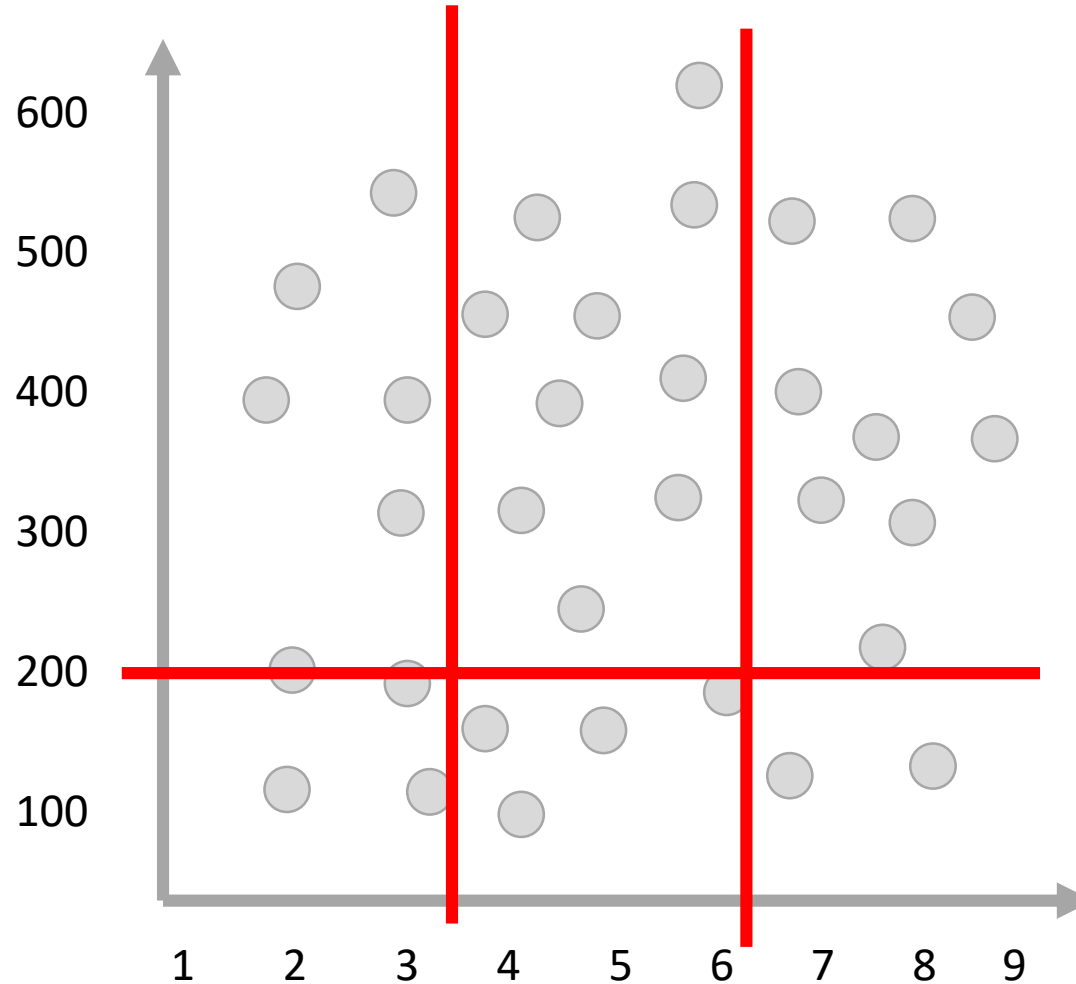


- Alte Daten beinhalten Suchanfrage, Klickverhalten des Nutzers.
- Weitere mögliche bekannte Informationen:
 - Quelle: SpiegelOnline vs. Blog
 - Aktualisierungsdatum
 - Textlänge
 - Anzahl der suchrelevanten Wörter
 - Position in der Web-Linkstruktur
 - ...
- Google nutzt nach eigenen Angaben 200 (!) Variablen.
- Das Verhalten von Interesse könnte sein:
 - Wahrscheinlichkeit, den Link anzuklicken,
 - Verweildauer,
 - Interaktion mit Werbung?

Beispiel mit 2 Variablen



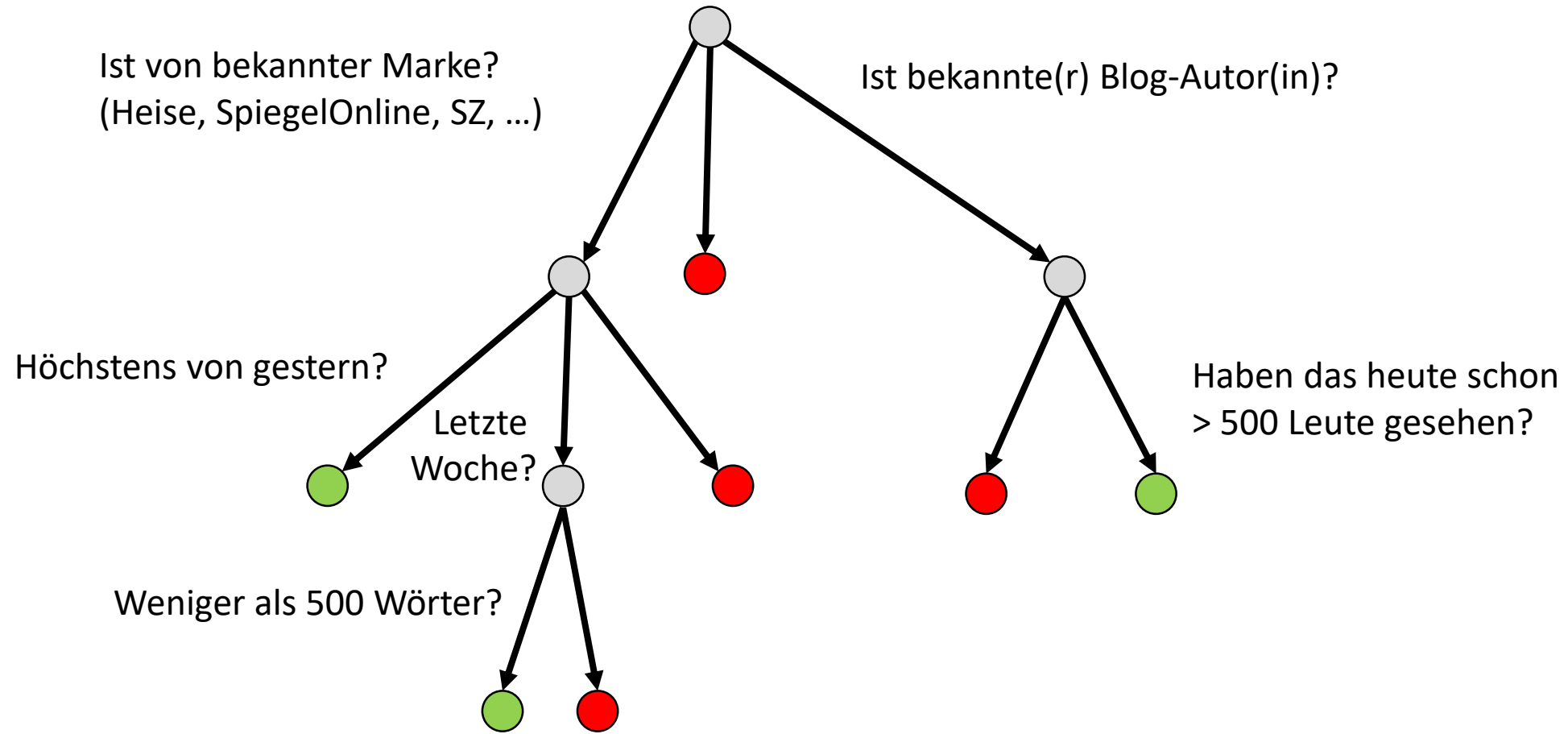
Länge des Textes [Worten]



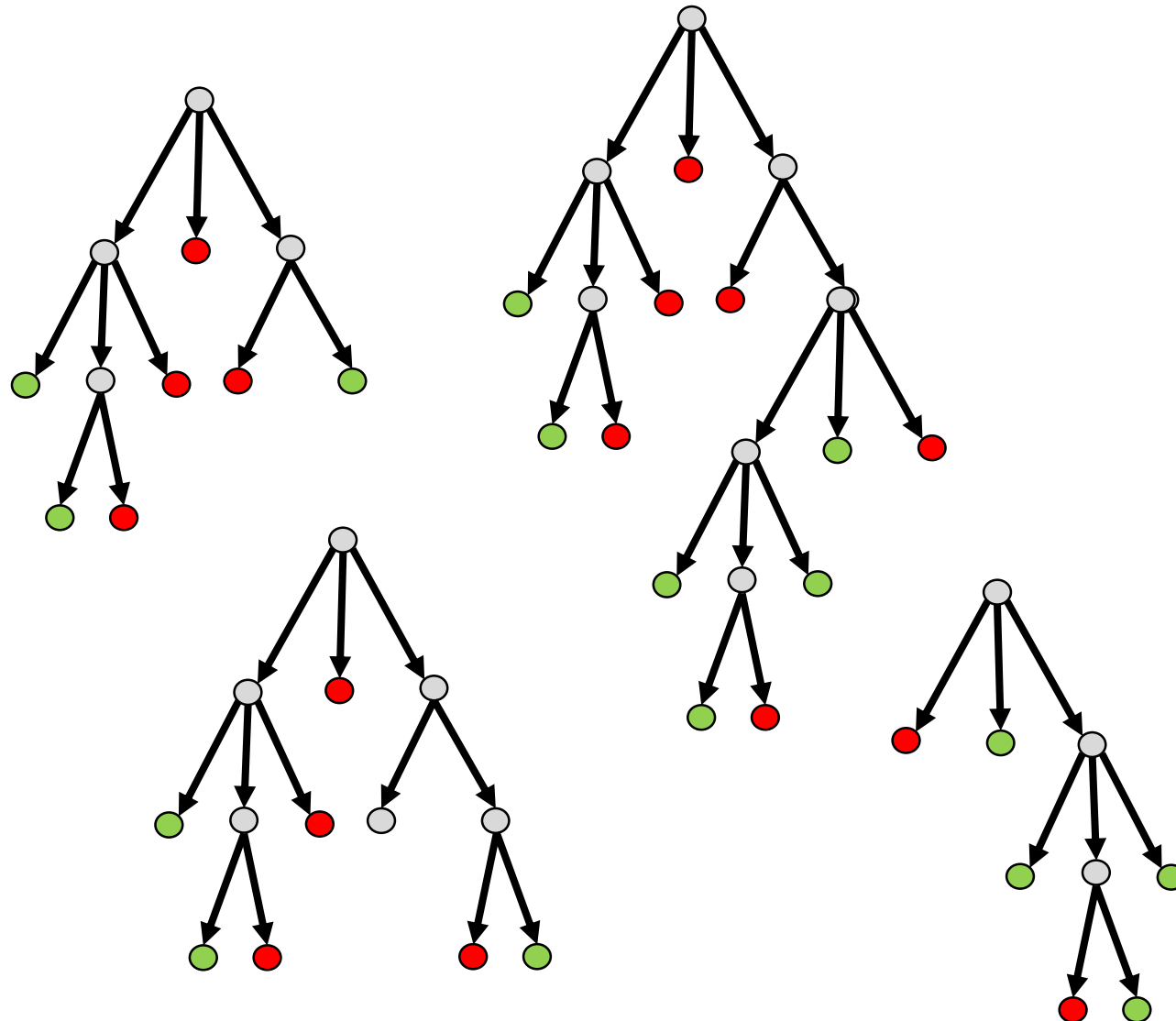
Aktualität

[Stunden seit Veröffentlichung]

Entscheidungsbaum



Entscheidungswald



- Aus einem Datensatz werde Hunderte dieser Bäume gelernt:
 - Unterschiedliche Variablenauswahl
 - Andere Reihenfolgen der Variablen
- Bei neuen Daten (Webseiten) „laufen“ diese durch alle Bäume, jeder Baum sagt „ja“ oder „nein“
- Es kommt eine Vorhersage heraus: 330 von 1000 Bäumen sagen: „Dieser Link wird geklickt werden“

Was kann schiefgehen?



Der Algorithmus löst das Problem nicht.

Die Modellierungsannahmen eines Algorithmus gelten nicht (mehr).

Die Daten sind nicht gut geeignet, um das Verhalten vorherzusagen oder sagen es nur für eine kleine Gruppe von Leuten gut vorher.



Relevanz von Algorithmen

Neutrale Empfehlungsalgorithmen?



- Empfehlungsalgorithmen filtern, lernen und sortieren.
- Sie machen Modellierungsannahmen, wählen Variablen aus und lernen nur aus einer Teilmenge von Daten.
- All diese Schritte können mehr oder weniger gut gelingen, keiner davon ist neutral im Sinne von „objektiv“.
- Viele Betreiber behaupten, sie würden **nur** nach Relevanz sortieren.

Relevanz – ein weites Feld



"A squirrel dying in front of your house may be **MORE RELEVANT TO YOUR INTERESTS** right now than people dying in Africa."

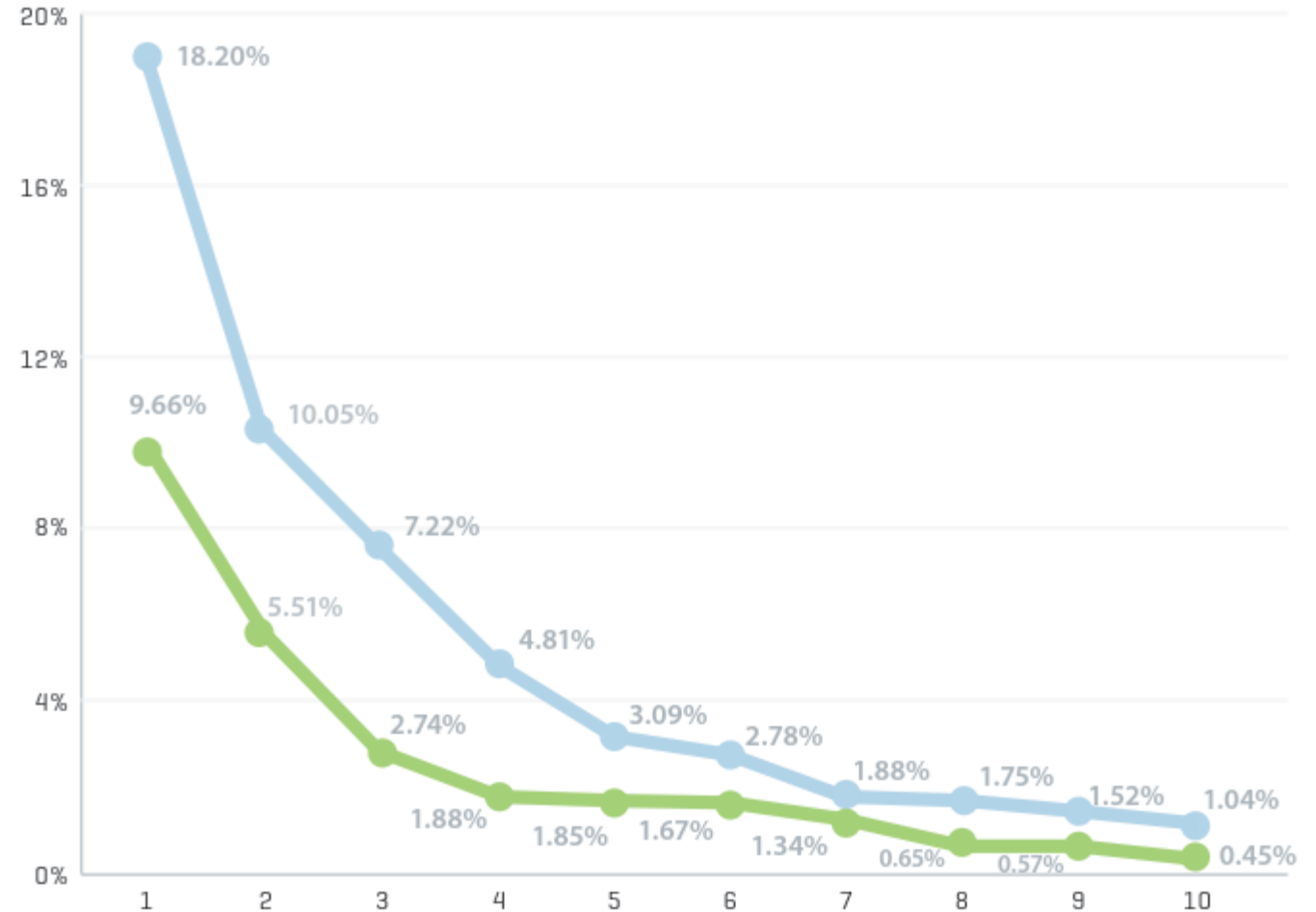
Mark Zuckerberg, CEO facebook,
nach David Kirkpatrick: „the facebook EFFECT“,
Simon & Schuster New York, New York, USA,
2010, S. 181



Vom Wert, der Erste zu sein

- Der erste bekommt ca. 18,2% (9,7%) aller User, der zweite nur noch 10,1% bzw. 5.51.
- 1. Platz fast doppelt so oft wie der 2., 2,5 mal so oft wie der dritte, fast fünfmal so oft wie der 4.

GOOGLE VS BING CLICK-THROUGH RATE



Paul Davison at Digital Relevance™: „A Tale of Two Studies: Establishing Google & Bing Click-Through Rates“, Study by Digital Relevance™ using client data from Jan-June 2011, available from <http://connect.relevance.com/a-tale-of-two-studies-establishing-google-bing-click-through-rates> or research@relevance.com; published 2013.

Sind wir beeinflussbar über Algorithmen?



- Suchergebnisreihenfolgen:
 - Manipulierte Suchreihenfolgen werden vom Nutzer nicht bemerkt und können die Tendenz eines unentschlossenen Wähler beeinflussen (Epstein & Robertson, 2015)
- Facebooks „Vote“ bzw. „Ich habe gewählt“-Button
 - Studie von Bond et al. über den Effekt auf das Wahlverhalten.
 - Effekt war klein, aber hochgerechnet ca. 60.000 mehr Wahlstimmen.

Epstein, R. & Robertson, R. E.: “The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections”, Proceedings of the National Academy of Science, 2015, E4512-E4521

Bond, R. M.; Fariss, C. J.; Jones, J. J.; Kramer, A. D. I.; Marlow, C.; Settle, J. E. & Fowler, J. H.: “A 61-million-person experiment in social influence and political mobilization”, Nature, 2012, 489, 295-298

Zuccherosconi



Wollte Zuckerberg der jüngste
Präsident der USA werden –

wer wäre rein technisch noch in
der Lage, als Gegenkandidat über
politische Themen zu punkten?





Algorithmen in einer demokratischen Gesellschaft

Beipackzettel für Algorithmen



Welches Problem „kuriert“ der Algorithmus?

Was ist das Einsatzgebiet des Algorithmus, was seine Modellannahmen?

Welche „Nebenwirkungen“ hat der Algorithmus?

Schlussformel



... zu Risiken und Nebenwirkungen der Digitalisierung befragen Sie bitte Ihren nächstgelegenen Data Scientist oder den deutschen Algorithmen TÜV.