



Macht der Algorithmen

Akademietage des Landkreises Biberach
1.2.2017

Prof. Dr. Katharina A. Zweig

„Aufträge“ an die Technikerin



- Beeinflussen uns Algorithmen bei den Wahlen? (mit Dank an Prof. Klein)
- Ist die Komplexität durch die Technologisierung so groß, dass Politik als ordnungssichernde Macht ihre Grundlage verliert, weil keine Ordnungssicherheit mehr möglich ist? (Danke an den Frager im Anschluss an den Vortrag von Prof. Anter)
- Und: kann man Macht vielleicht doch besitzen, im Gegensatz zu dem, was Hannah Arendt sagte? (Mit Dank an Prof. Anter)

Haben Sie auch gezittert?



Vor der Wahl: „Google manipuliert für Hillary“



- SourceFed behauptete, negative Suchvervollständigungen würden von Google bei Ms Clinton unterdrückt
- Psychologe Prof. Dr. Epstein sah massive Manipulation bei Google und betonte wiederholt, dass bis zu 20% der unentschlossenen Wähler durch Suchmaschinen manipuliert werden können.

Nach der Wahl:

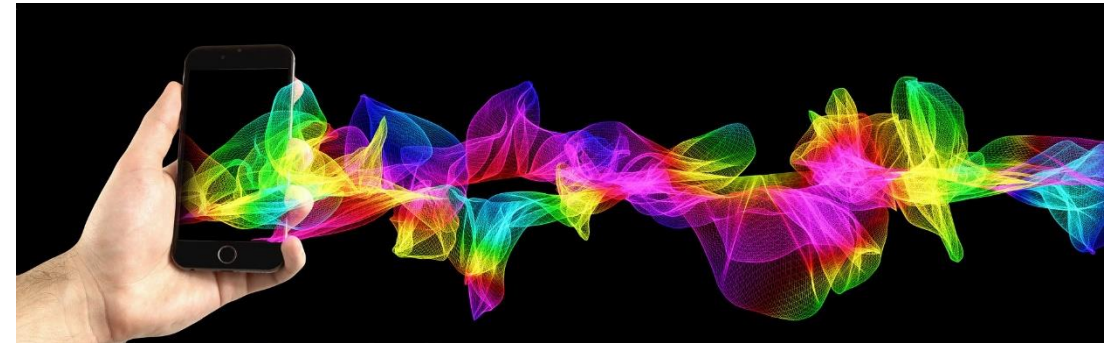


- Ach nee.
- Doch nicht.
- Dann war's Facebook:
- Filter bubble und Echokammern werden algorithmisch erzeugt, so dass wir in einer postfaktischen Welt leben.

Ich möchte Sie gerne besser kennenlernen...



- Haben Sie einen aktiven Facebook-Account?
 - 1 = Ja
 - 2 = Nein
- Haben Sie das Gefühl, dass Sie von Algorithmen manipuliert werden?
 - 1 = Ja
 - 2 = Nein
- Glauben Sie, dass Computer „lernen“ können?
 - 1 = Ja
 - 2 = Nein





Teil I:

Das kleine ABC der Informatik

Teil I: 50 Minuten plus 10 Minuten Diskussion

Dann 20 Minuten Pause

Teil II: 50 Minuten plus 20 Minuten Diskussion

Das kleine ABC der Informatik



Wann gefährden

Algorithmen,

Big Data und

Computerintelligenz

unsere Demokratie?



A wie Algorithmus

Ein Algorithmus ist eine genaue Handlungsanweisung,
die ein allgemeines Problem löst

Mathematisches Problem



INPUT

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



OUTPUT

Ist das Unkraut oder kann das weg?



- Unkrautjäten bedarf genauer Instruktionen!
- Dann wird aber jeder Garten zu demselben Ergebnis führen:
 - Unkrautfreiheit!
- Der Garten ist der **Input** (die **Eingabe**) für das Problem.
- Der **unkrautfreie** Garten ist der **Output** (die **Ausgabe**) der Prozedur.

Weitere Beispiele für mathematische Probleme



- Gegeben zwei Zahlen, nenne die größte dazwischenliegende Zahl, die eine Quadratzahl ist. Gibt es keine, gib -1 aus.
- Gegeben ein Pokerblatt, gib die Wahrscheinlichkeit dafür aus, mit dieser Hand zu gewinnen.
- Die „Algorithmen“ sind hier einfache, mathematische Formeln.
- Ein Problem, aber kein mathematisches:
 - Gegeben eine Soße, gib das Rezept aus, wie man sie schmackhaft machen kann.
 - „Schmackhaftigkeit“ ist nicht quantifizierbar und es wäre unklar, welche Aromen gemischt werden müssen, um diesen Geschmack zu erreichen
- Ab jetzt Probleme mit Algorithmen, die nicht einfach eine mathematische Formel sind.



1. Beispiel: Navigation



Navigation

Gegeben das Kartenmaterial und weitere Daten, berechne die kürzeste Route zwischen Start und Ziel.

Das **Problem** sagt nicht, wie man die Lösung **findet**.



Input: Straßen, Länge, Staus, ...
Start und Ziel



Output: optimale Route

Ein Algorithmus ist...



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

Eine **Implementierung** ist die Übersetzung der für den Menschen verständlichen Handlungsanweisung in eine Programmiersprache.



2. Beispiel: Sortieren

Problem: Sortieren



Sortieren 1: „Sortieren durch Einfügen“

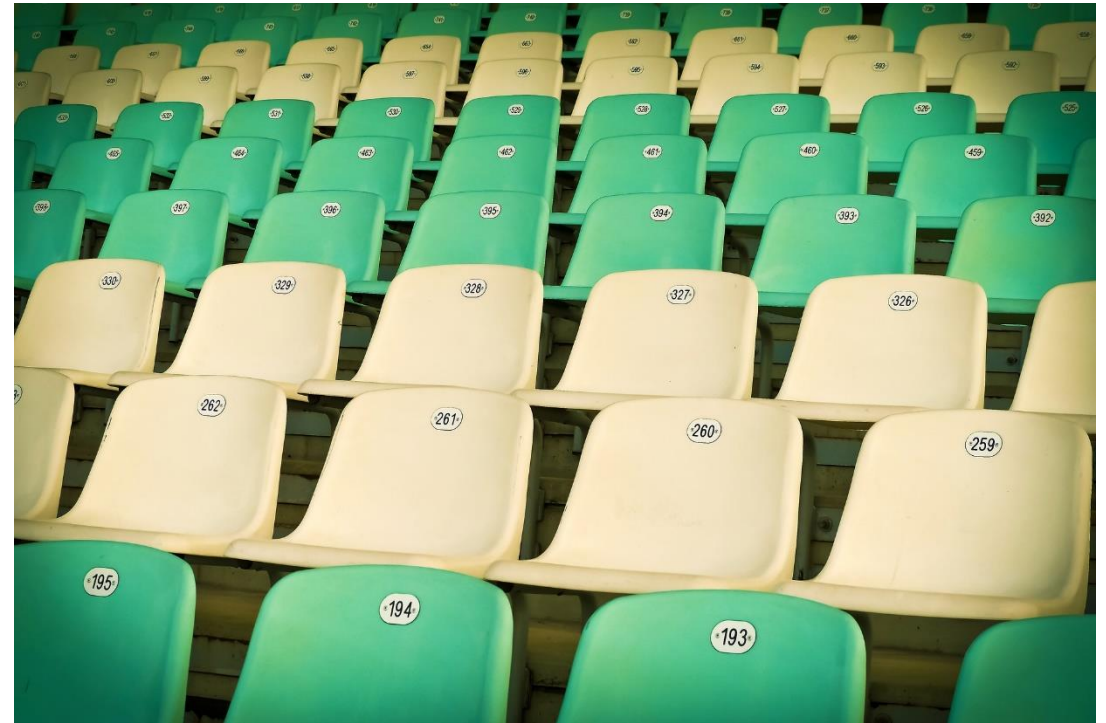


- Hier: Karten mit Zahlen darauf.
- Möglicher Algorithmus, um Zahlen und Besitzer(in) pro Reihe zu sortieren:
 - Alle stehen auf.
 - Eine oder Einer setzt sich hin.
 - Solange noch jemand steht,
 - Geht eine davon an der Reihe entlang und identifiziert die Lücke. Falls rechts davon jemand sitzt, müssen diese Personen alle eins nach rechts weiterrücken.

Sortieren 2: Aufsteigendes Sortieren



- Alle bleiben sitzen.
- Alle überprüfen immer wieder, ob sie ihre Karte mit ihrem Nachbarn tauschen müssen.
 - Wenn ja, dann bitte tauschen.
 - Nur die Zahlen tauschen, nicht die Sitzplätze!
 - Jeder hat immer nur eine Karte in der Hand.
- Wenn kein Tausch mehr nötig ist, sind alle Karten sortiert.



Alle Sortierprobleme auf einen Schlag



- Gegeben eine Menge von Objekten oder Subjekten...
- ... und ein Sortierkriterium, das für je zwei von diesen besagt, welches nach links, welches nach rechts sortiert werden muss,...
- ... kann jeder beliebige Sortieralgorithmus die korrekte Lösung berechnen.

- Eine Interpretation der Ergebnisse (dies sind die relevantesten Nachrichten, die wichtigsten Freunde, die kaufenswertesten Produkte) liefert er **nicht**.



Und wie funktioniert Google?

Suchmaschinen 101

1. Filtern aus allen ihnen bekannten Webseiten diejenigen, deren Text mit den angegebenen Suchbegriffen zusammenhängen.
2. Sie bewerten diese Seiten anhand der Vernetzungsstruktur der Seiten untereinander



Idee hinter dem Algorithmus



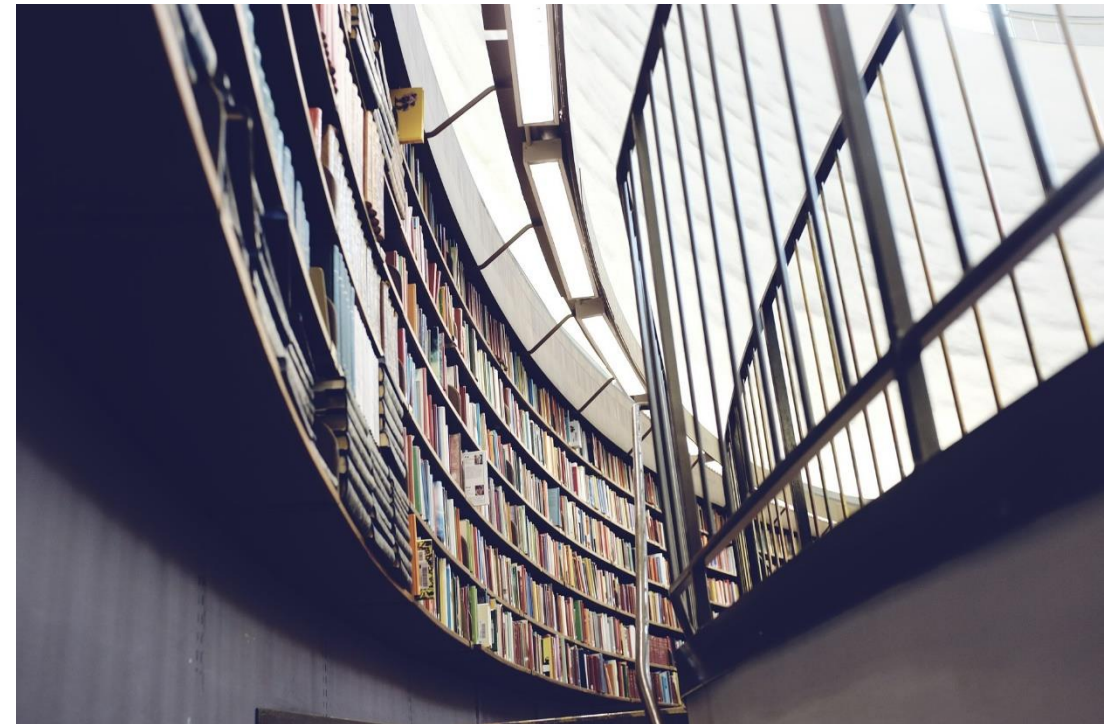
Ein Modell menschlichen Verhaltens: der Random Surfer

- Ein Surfer klickt auf eine Webseite
- Folgt einem der Links auf der Webseite zufällig
- Von Zeit zu Zeit springt er auf eine völlig neue Webseite
 - Modelliert externes Wissen (z.B. Werbung, bekannte Seiten)

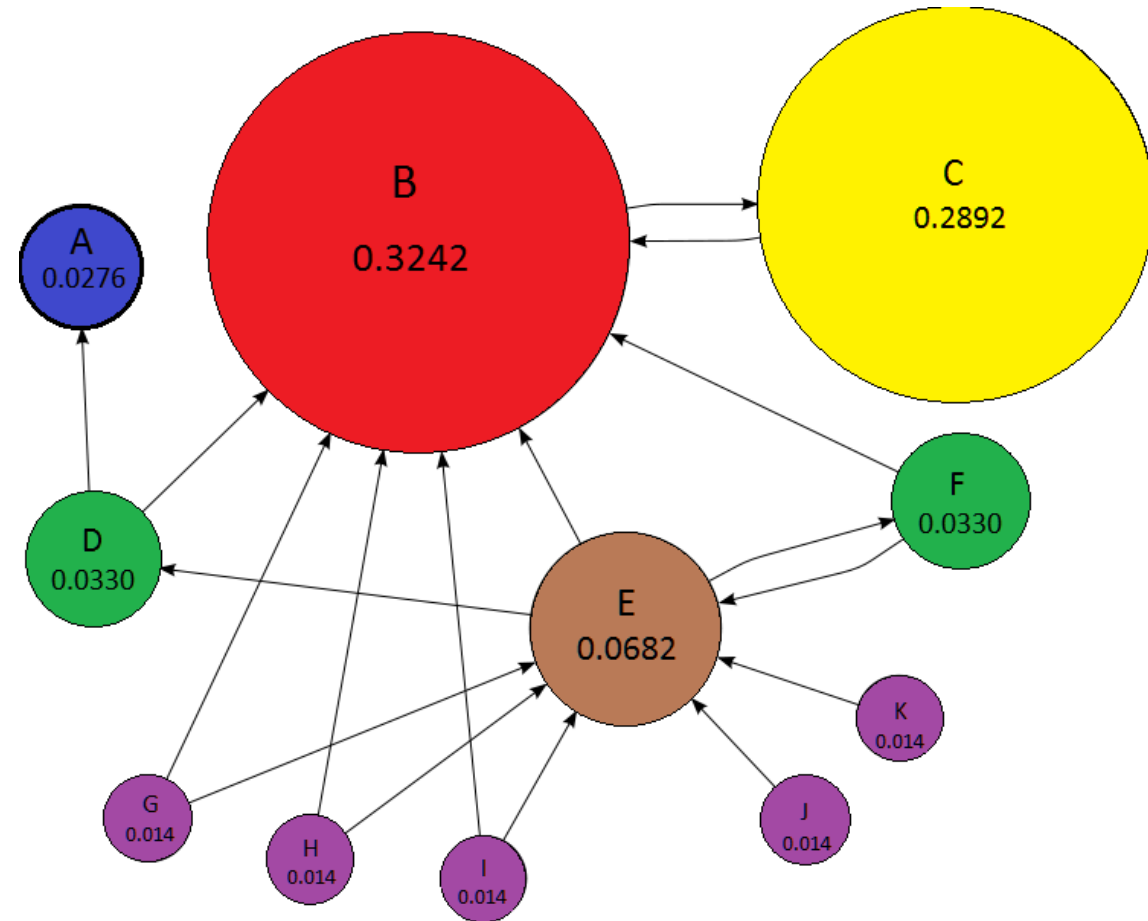
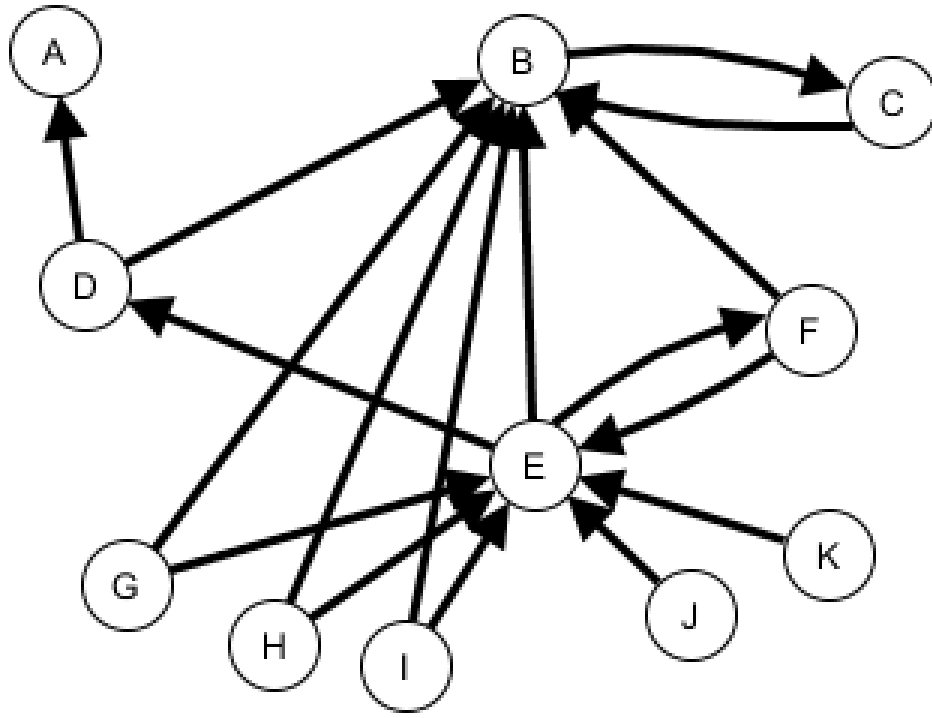
Analogie Buchreferenzen



- Bücher verweisen auf andere Bücher
- Bei eigenen Recherchen findet man immer wieder dieselben Person
 - Z.B. Hannah Arendt zum Thema Macht!
- Wir können davon ausgehen, dass sie wichtig für das Thema ist.
- Unterschied: Bücher können nur auf schon erschienene Bücher verweisen, Webseiten unterliegen keinen zeitlichen Einschränkungen.



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.

„Harmlose“ Algorithmen



- Die meisten Algorithmen laufen unbemerkt:
 - Suchen, Speichern, Filtern, Sortieren
 - Logistik, Scheduling
 - Bilderkennung in Industrierobotern
 - Verwaltungssoftware
 - Bürosoftware
 - ...
- Aber lernende Algorithmen können mehr. Sie können Antworten finden auf die Fragen:
 - Was mögen Kunden, die diese Bücher mögen?
 - Welche Webseite suchen SIE? (Personalisierung)
- Richtig mächtig werden sie mit „Big Data“.



B wie Big Data

Was ist die relevanteste Nachricht zur Anfrage: „Donald Trump“?



Wie Erdogan und die EU sich von Anfang an belogen
DIE WELT - vor 13 Stunden
Schon der Beginn war stürmisch, voll emotionaler Hochspannung. Als die Regierungschefs der EU Mitte Dezember 2004 beschlossen, ...



Donald Trump – In Rekordgeschwindigkeit als Präsident unbeliebt
DIE WELT - vor 5 Stunden
Die Empörung über **Donald Trumps** jüngstes Dekret zum Einreiseverbot ist groß. Die Proteste finden auf der Straße statt, und im Netz: ...
Widerstand gegen Donald Trump: Ab wann es für den US ...
Ausführlich - **FOCUS Online** - vor 7 Stunden
Alle ansehen



Neuer Richter für Supreme Court - Die Liebe der Abtreibungsgegner ...
Süddeutsche.de - vor 1 Stunde
US-Präsident Trump wird einen Richter an den Supreme Court schicken, der ... Auch hier bricht **Donald Trump** mit seinen Vorgängern.

Big Data



- Wie kann Relevanz modelliert und „quantifiziert“ werden?
- Big Data Methoden nutzen, z.B.:
 - Sprache der Anfrage, Niveau der Anfrage, Wörter, Wortkombination
 - Tageszeit und geographische Informationen, Gerätetyp
 - Ihre bisherigen Suchanfragen und Ihr persönliches Klickverhalten
 - Welche Seiten wurden angeklickt, wie lange betrachtet, kam die Nutzerin wieder zurück zu den Ergebnissen?
 - Metadaten der Nachrichten/Medien: wann erstellt, durch wen, wo publiziert, Verschlagwortung, Wahl der Wörter
 - Verhalten anderer Nutzer, „ liken “ auf sozialen Netzwerken, Interaktion mit Beiträgen

Big Data



Ganz allgemein:

- Große Datenmengen,
- die schnell verarbeitet werden müssen.
- Außerhalb ihres spezifischen Zwecks genutzt
- Daher im Einzelnen vermutlich fehlerbehaftet
- Dank großer Masse und wenig individualisiertem Verhalten statistisch nutzbar





C wie Computerintelligenz

Lernende Algorithmen



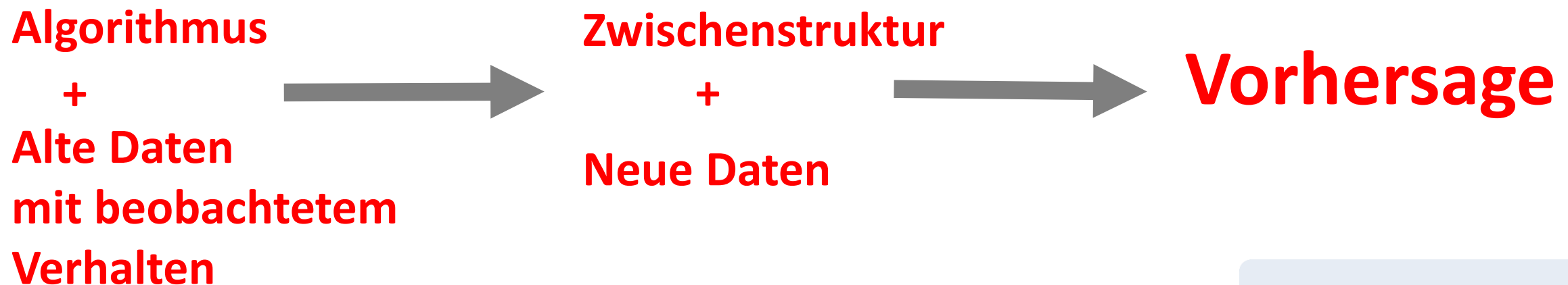
Die Buchhändlerin Ihres Vertrauens



Computerintelligenz



- **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)

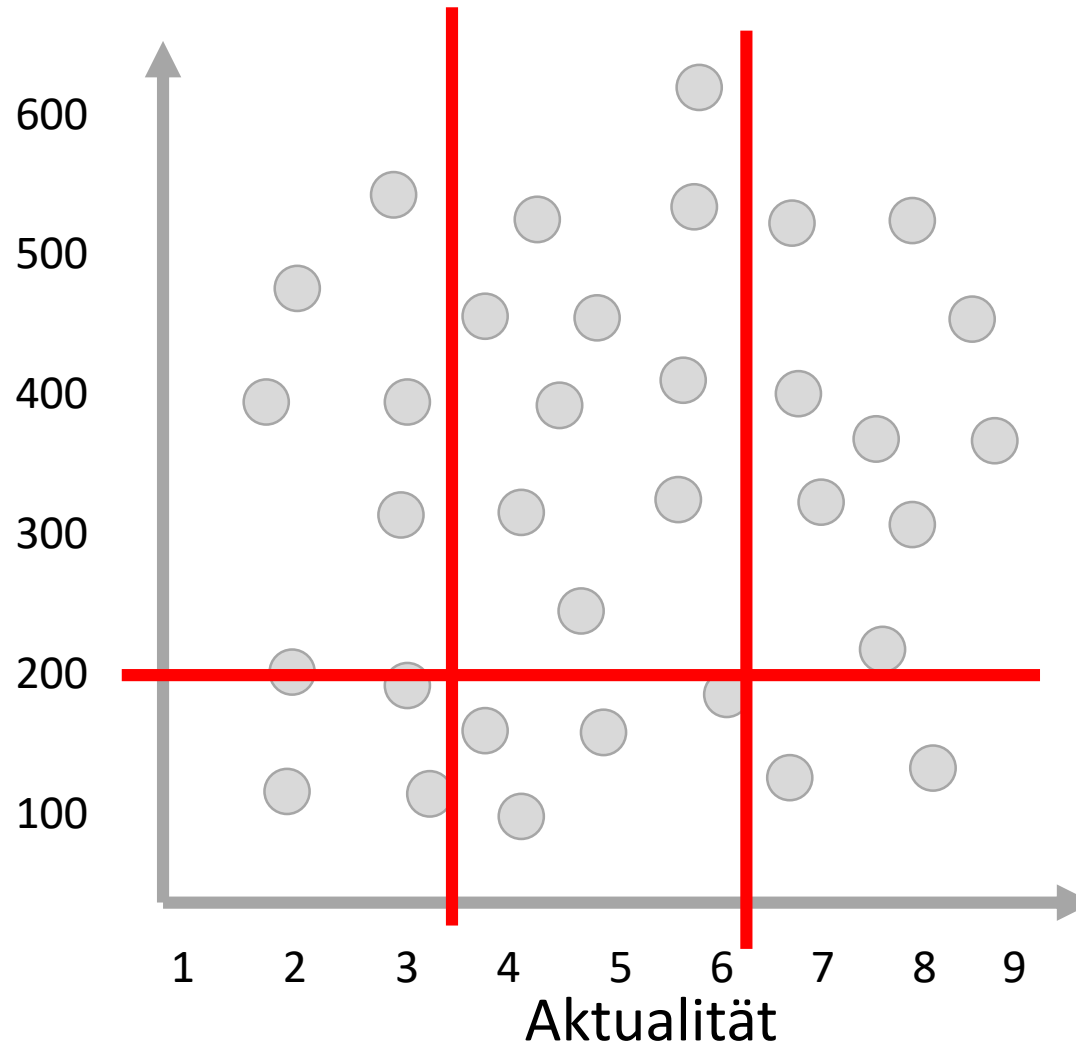


- Alte Daten beinhalten Suchanfrage, Klickverhalten des Nutzers.
- Weiterhin: Big Data wie oben genannt.
- Google nutzt nach eigenen Angaben 200 (!) Variablen.
- Das Verhalten von Interesse, das die **Relevanz** suggeriert, könnte sein:
 - Wahrscheinlichkeit, den Link anzuklicken,
 - Verweildauer,
 - Interaktion mit Werbung?

Beispiel mit 2 Variablen



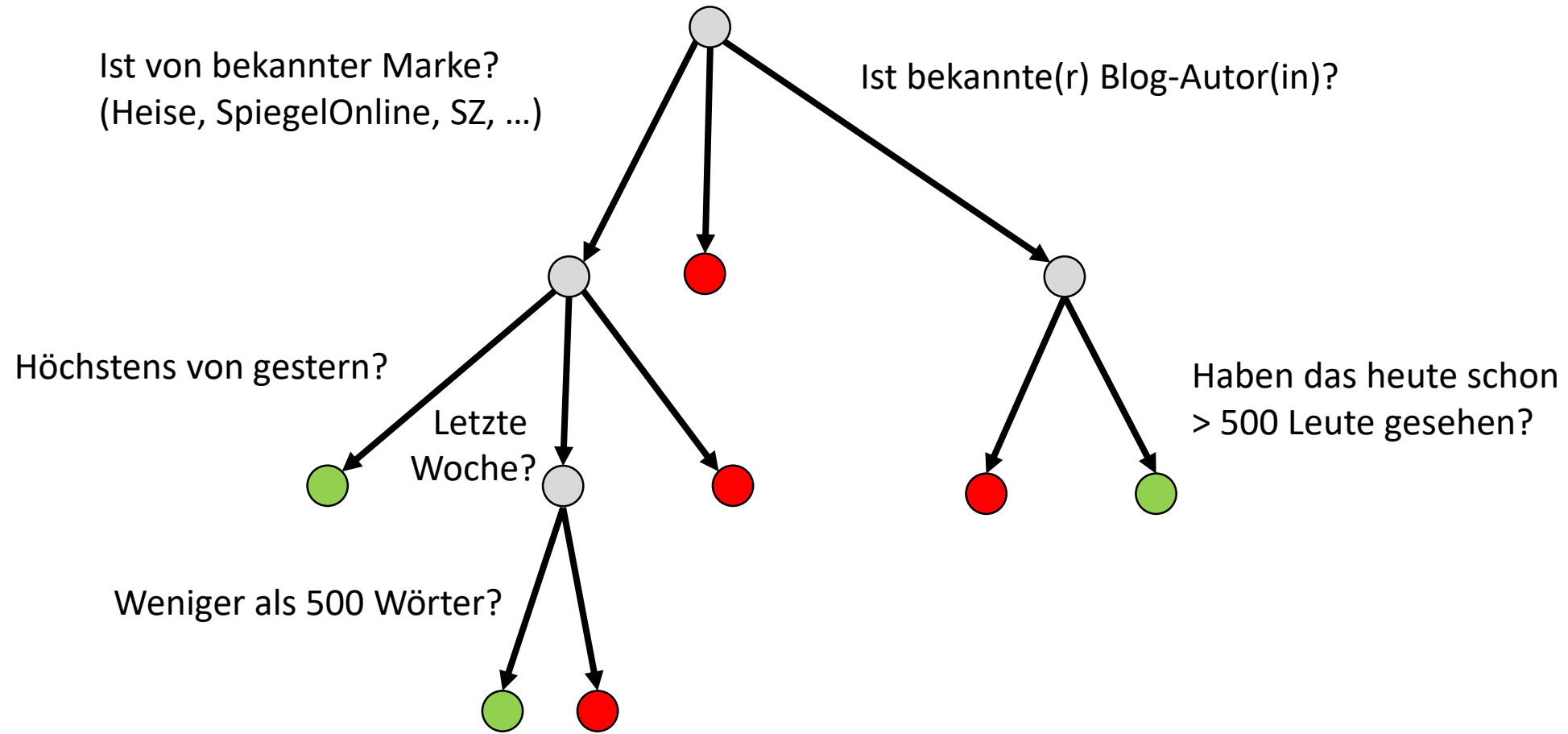
Länge des Textes [Worten]



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“

Lernende Algorithmen



- Sehr mächtiges Werkzeug,
- das aus viel mehr Daten lernen kann,
- als eine menschliche Expertin (wegen ihrer kurzen Lebensspanne).
- Sie betrachten nur die Daten, die ihnen zur Verfügung stehen -> das kann für mehr Objektivität sorgen!
- Sie haben keine „Tagesform“.

Wo funktioniert das nicht?

- Wenn es keine Ähnlichkeit zwischen den angeklickten Inhalten gibt.
- Beispiel: Die Bilder, die ich von Pixabay lade, sind thematisch nicht miteinander verwandt, sondern maximal unterschiedlich.
- Hier kann der Algorithmus vom „Klickmuster“ nichts lernen.



Ähnliche Bilder





Zusammenfassung Algorithmen



Algorithmen...

- ...können aus nur einer Formel bestehen,
- aber meistens sind sie eine komplexe Mischung aus Handlungsanweisungen und Berechnungen,
- die genau eine Lösung für das Problem basierend auf den Eingabedaten berechnen.
- Es sind eingefrorene Handlungsanweisungen, von Menschen erdacht, um mathematische Probleme durch Computer zu lösen.
- Das Ergebnis **muss interpretiert** werden.

```
var simple = type.slice(
  forward = type.slice(
    ofType = what === "of-type";
  return first === 1 && last === 0 ?
  // Shortcut for :nth-*(n)
  function( elem ) {
    return !elem.parentNode;
  } :
  function( elem, context, xml ) {
    var cache, outerCache, node, diff, nodeIndex, start,
        dir = simple !== forward ? "nextSibling" :
            "previousSibling",
        parent = elem.parentNode,
        name = ofType && elem.nodeName.toLowerCase(),
        useCache = !xml && !ofType;
    if ( parent ) {
      ...
    }
  }
);
```



Ich mache auch eine Vorhersage...



Ich mache in 15
Minuten weiter.
Daher...

...13.5 Minuten
PAUSE





Teil II:

Algorithmen in einer Demokratie

Chance oder Risiko?

Haben Sie auch gezittert?





Relevanz von Algorithmen

Gibt es neutrale Empfehlungsalgorithmen in Suchmaschinen, Newsfeed Aggregatoren und sozialen Netzwerken?



- Viele Betreiber behaupten, sie würden Nachrichten **nur** nach Relevanz sortieren.
- Aber: 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“.

Spielkampsche Regel



**Alle Algorithmen sind objektiv
Bis auf die von Menschen gemachten!**

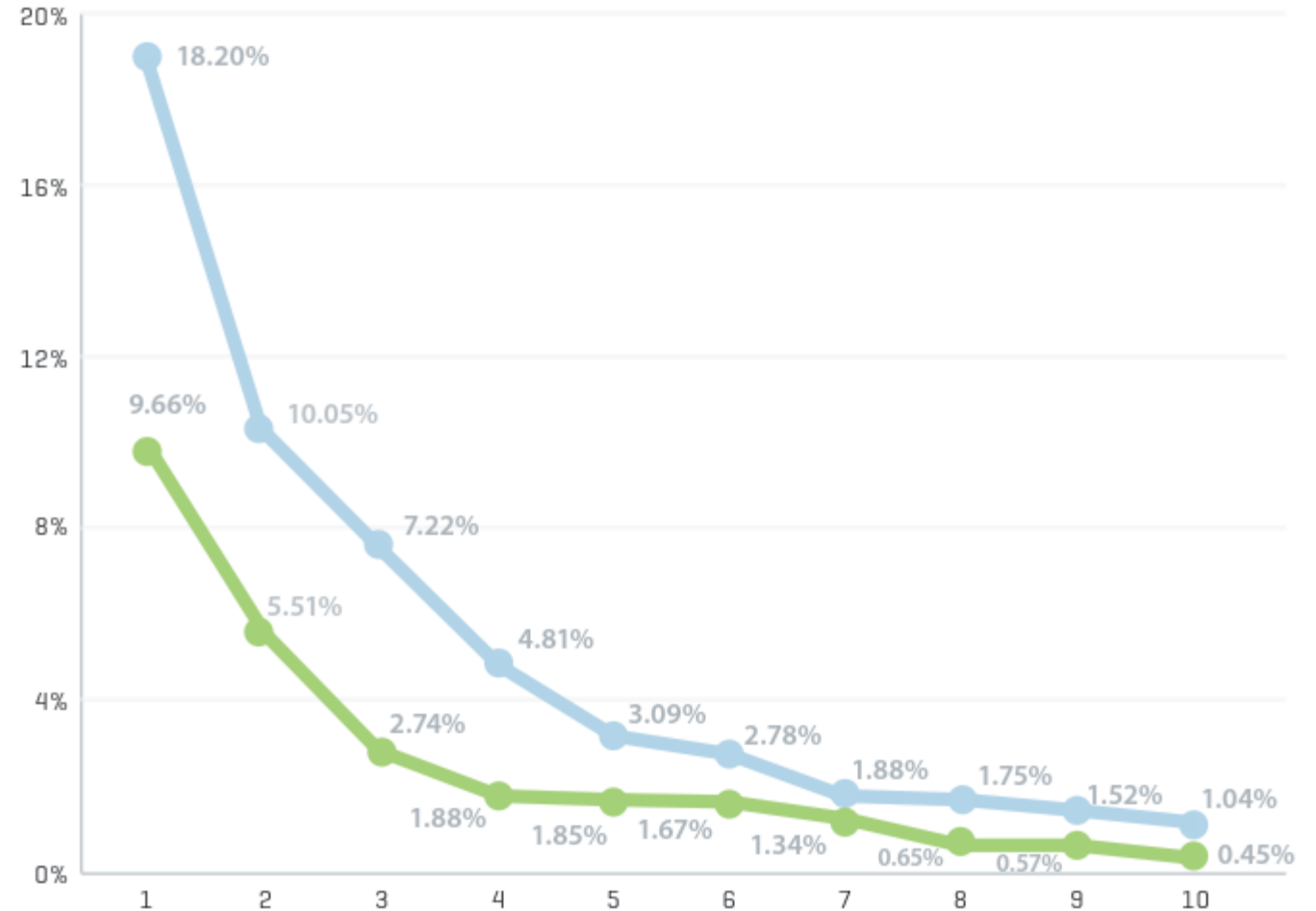


Menschliche Wahrnehmung

Am Beispiel von Suchmaschinen, sozialen Netzwerken und News-Aggregatoren

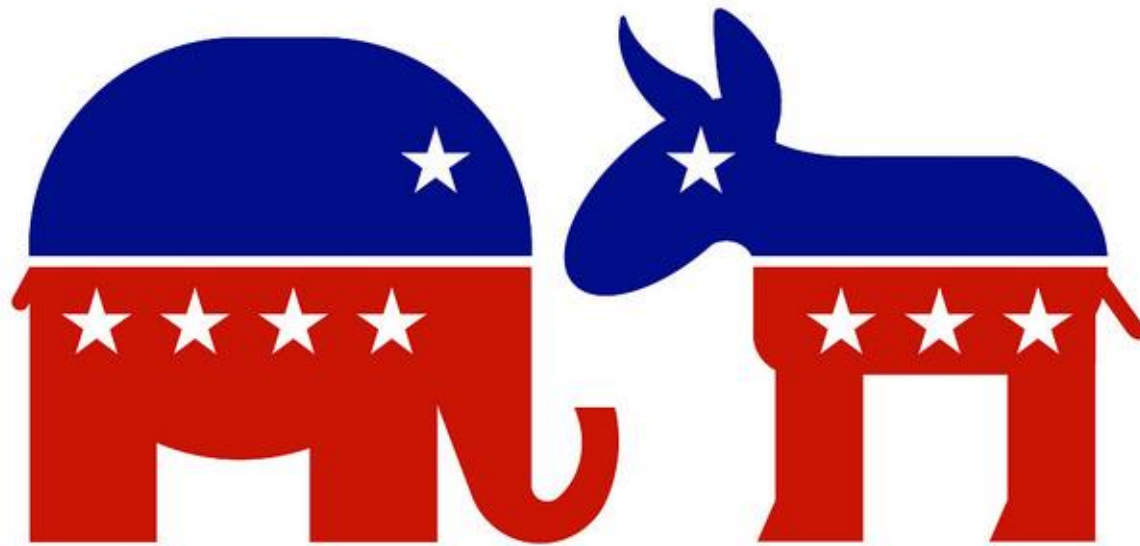
Vom Wert, der Erste zu sein

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.

Bevorzugt Google Demokraten?



DiederHunt

Studie von Trielli, Mussenden und Diakopolous¹:

Unter 16 Präsidentschaftskandidaten (USA) gab es bei Demokraten unter den ersten 10 Suchergebnissen 7 positive Berichte, bei Republikanern nur 5,9.

1 <http://algorithmwatch.org/warum-die-google-suchergebnisse-in-den-usa-die-demokraten-bevorzugen/>

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

Oder waren es ...



- BuzzFeed schreibt, dass ein deutlicher Anteil der Nachrichten auf hyperkonservativen Webseiten falsch oder eine Mischung aus falschen und korrekten Nachrichten ist.
- Ein Teil dieser „Hyperpartisan“-Webseiten stammt aus....
- ...Mazedonien!

Motivation



„Die meisten Texte auf diesen Seiten sind zusammengeschrieben oder einfach abgeschrieben von rechtslastigen Webseiten der USA. Die Mazedonier sahen irgendwo eine Story, versetzten sie mit einer aufgeheizten Überschrift, und veröffentlichten es dann schnell auf ihrer Webseite. Dann teilen sie die Nachricht auf Facebook und veranlassen damit Leute, auf ihre Seite zu kommen. Je mehr Leute sich verführen lassen, desto mehr Geld verdienen sie durch Werbung auf ihrer Webseite.“

Craig Silverman und Lawrence Alexander: „**How Teens In The Balkans Are Duping Trump Supporters With Fake News**“, BuzzFeed Nov. 4th, downloaded on the 27th of November, 2016
https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.wvzZe7D5#.he3BEIYV

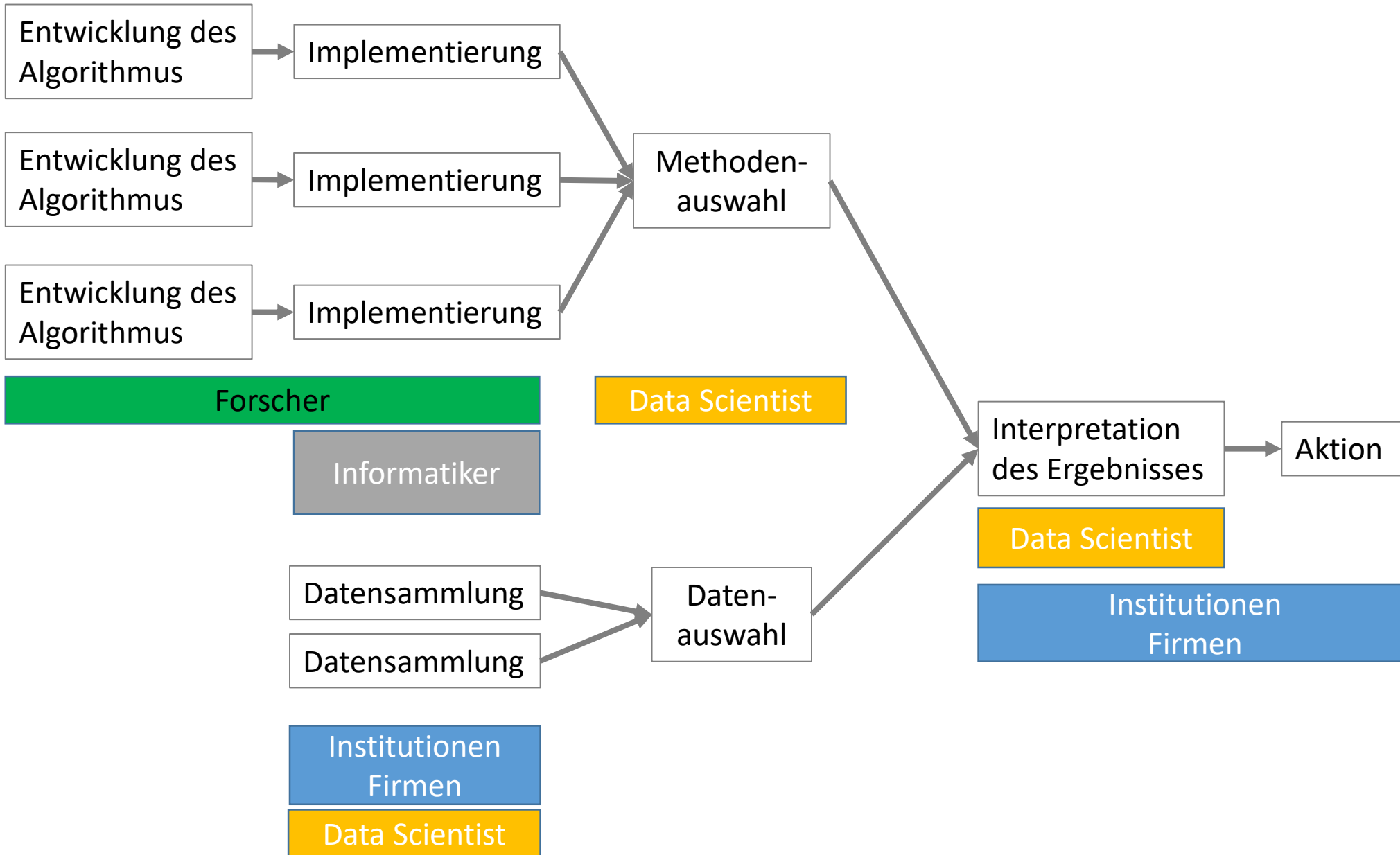


... und was bedeutet das für die nächste Bundestagswahl?



Algorithmen in einer demokratischen Gesellschaft

Verkettete Verantwortlichkeiten



Wer überwacht die Auswirkungen auf die Gesellschaft?

Medien?
Gesellschaft?
Politik?
Institutionen?
Firmen?
Recht?



Quis custodiet ipsos algorithmos

Der „Automated Decision Making“-TÜV vulgo: „Algorithmen TÜV“

Gründung von „Algorithm Watch“



ALGORITHM
WATCH



Lorena Jaume-Palasi, Mitarbeiterin im iRights.Lab



Lorenz Matzat, Datenjournalist der 1. Stunde, Gründer von lokaler.de, Grimme-Preis-Träger



Matthias Spielkamp, Gründer von iRights.info, ebenfalls Grimme-Preis-Träger, Vorstandsmitglied von Reporter ohne Grenzen.



Prof. Dr. K.A. Zweig, Junior Fellow der Gesellschaft für Informatik, Digitaler Kopf 2014, TU Kaiserslautern



Notwendige Eigenschaften

- Unabhängige Prüfstelle mit Siegelvergabe
- Möglichst auch mit Forschungsauftrag
- Identifikation der **kleinstmöglichen Menge** an zu überprüfenden Algorithmen
 - Die meisten Algorithmen sind harmlos;
 - Produkthaftung ermöglicht, dass andere, z.B. Versicherungen, Interesse an korrekten Algorithmen haben;
 - Wettbewerb ermöglicht, dass andere ‚neutralere‘ Algorithmen anbieten.
 - **Kein weiteres Innovationshemmnis!**
- **Non-Profit**

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.

Kontakt Daten

Prof. Dr. Katharina A. Zweig

TU Kaiserslautern

Gottlieb-Daimler-Str. 48

67663 Kaiserslautern

zweig@cs.uni-kl.de

Algorithmwatch.org





Algorithmen, Grundrechte und Diskriminierung

Wie sagt man die Rückfallrate eines Verbrechens voraus?



Predictive Policing



Wir haben schon
auf Sie gewartet!



Vorhersagen,
wann und wo
Straftaten
wahrscheinlich
sind.

Predictive Policing



Ein **Algorithmus**
hat mir geflüstert,
dass Du **fast** ein Krimineller bist.
Dann komm mal mit!



Aber auch: Vorhersagen,
ob ein Individuum
straffällig werden könnte!

Beispiel USA:

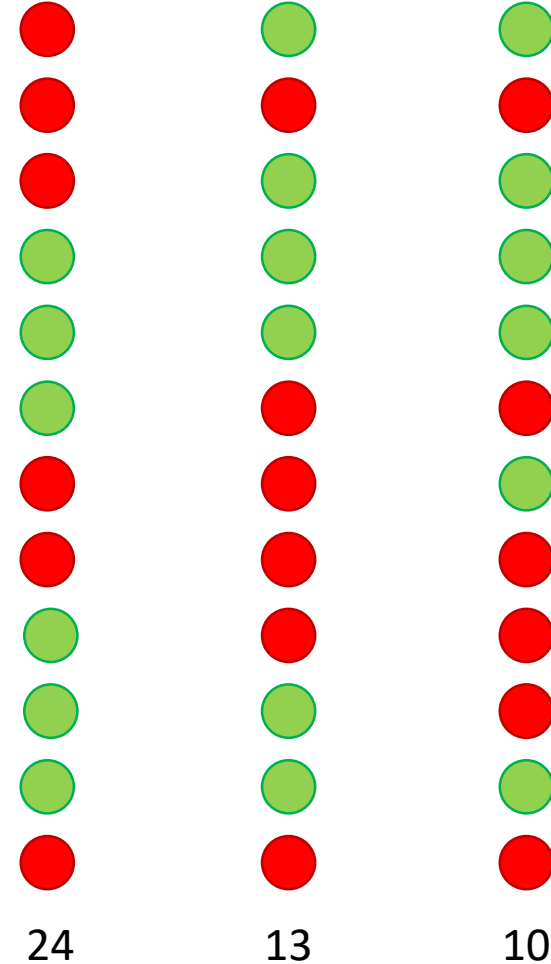
- 1) Oregon
- 2) Andere Bundesstaaten



Wieso benutzt eine Gesellschaft solche Algorithmen?

Evaluation von Vorhersagen

- Der Algorithmus bewertet jeweils, wie viele wirklich rückfällig gewordene Personen möglichst weit oben stehen. (Auf „alten“ Daten)
- Der Algorithmus, der das maximiert, wird für weitere Daten genommen.



Rot soll nach oben sortiert werden.

Oregon Recidivism Rate Algorithm

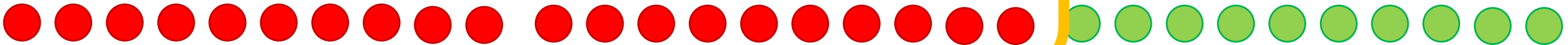
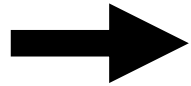


- Das oben genannte Qualitätsmaß dieses Algorithmus: 72 von 100 Paaren werden korrekt sortiert.
- Der in Oregon benutzte Algorithmus hat also, gegeben einen „Rückfaller“ und einen „Nichtrückfaller“, eine Chance von ca. 1:3 den Rückfaller höher zu gewichten als den Nichtrückfaller.
- Nur 25% aller so gemachten Prognosen sind falsch!
 - Das klingt doch ganz gut, oder?
- So werden aber keine Urteile gefällt!
- Problem: die Klassen sind ungleich verteilt!
 - 1000 Delinquenten
 - Ca. 2000 werden rückfällig

Optimale Sortierung



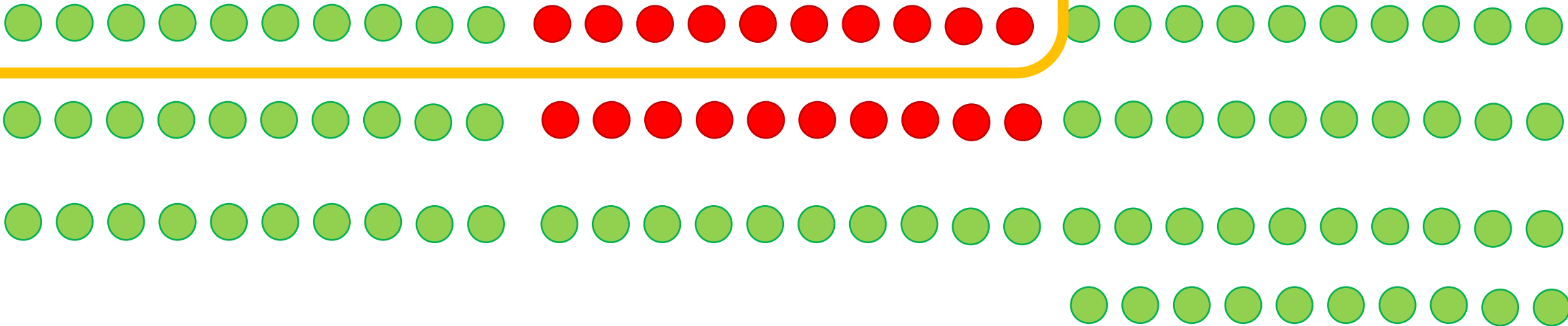
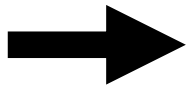
Erwartete 20% „Rückfällige“



Mögliche Sortierung eines Algorithmus mit dieser „Güte“ (75/100 Paaren)



Erwartete 20% „Rückfällige“



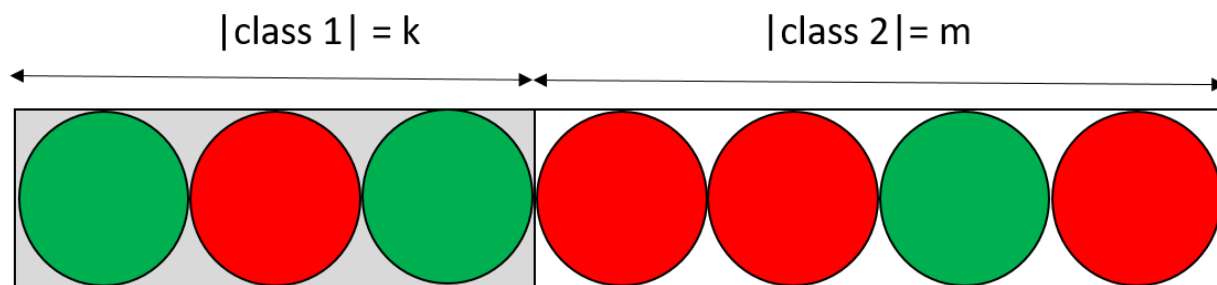
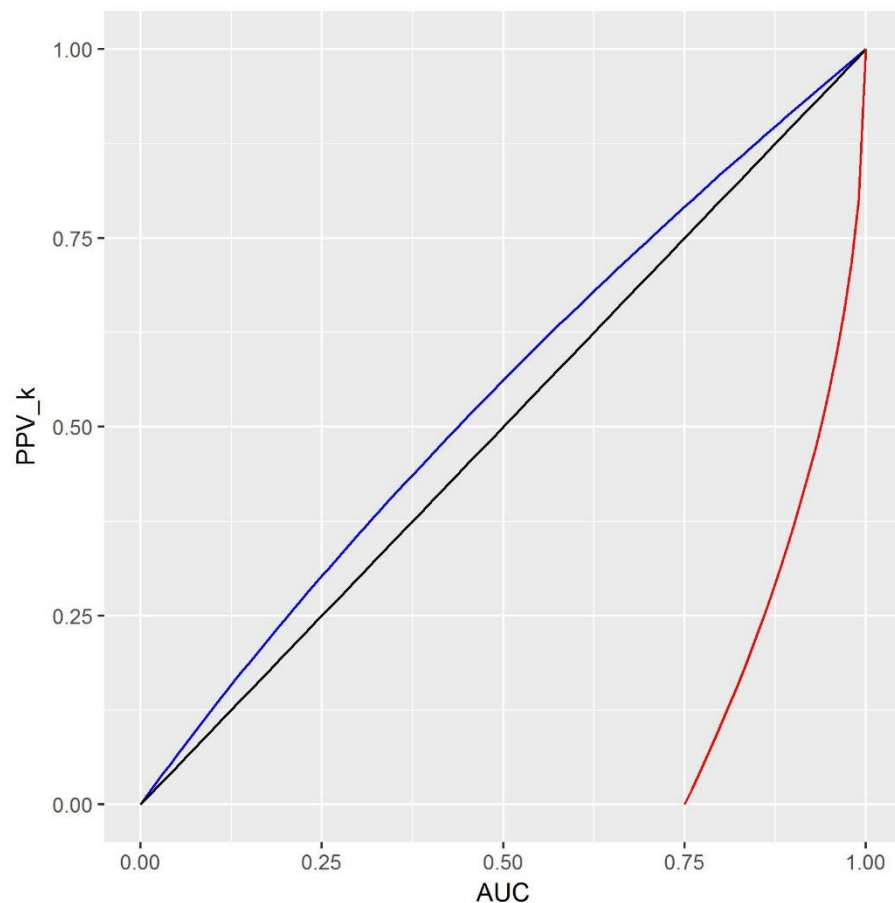
Problem: Unbalancierte Klassen



- Bei optimaler Sortierung: die ersten 200 rot – keine Fehlentscheidung.
- Jetzt: nur die Hälfte!
- Damit **50% Fehlentscheidungen**
- **Es könnten sogar 100% sein.**



Qualitätsmaße



- Diagramm zeigt, wie sehr das „klassische Qualitätsmaß“ (schwarze Diagonale) vom besser verständlichen Maß abweicht.
- Blaue, obere Linie: Höchstwert gegeben das klassische Maß
- Rote, untere Linie: Minimaler Wert gegeben das klassische Maß.

Das ist wie...



„Kaufen Sie diesen wunderbaren Wagen. TÜV? Brauchen Sie nicht! Und sehen Sie nur, die unglaublich gut erhaltenen Sommerreifen. Das ist noch Qualität!“



Rückfallvorhersagealgorithmus ist rassistisch (Propublica)



- In einer Studie von Propublica (anderer Algorithmus) war die Quote noch schlechter:
 - Nur 20% der (vorhergesagten) Gewalttäter begingen eine Straftat
 - Bei allen möglichen Straftaten war die Vorhersage etwas besser als ein Münzwurf.
 - Bei schwarzen Mitbürgern war die Vorhersage immer zu pessimistisch;
 - Bei weißen zu optimistisch.
- Northpoint Software ist eine Firma, der Algorithmus ist unbekannt.
- Rasse ist an sich keine Variable des Algorithmus...

Neueste Entwicklung

- Der Algorithmus ist unfair, aber nicht rassistisch.
- Beide Kriterien können mathematisch gesehen (!) nicht gleichzeitig erfüllt werden!
- Welche Art von Gerechtigkeit wollen wir?



Zweig'sche Regel



Algorithmen der künstlichen Intelligenz werden da eingesetzt, wo es **keine einfachen Regeln** gibt.

Sie suchen **Muster** in hoch-verrauschten Datensätzen.

Die Muster sind daher grundsätzlich **statistischer Natur**.

Versuchen fast immer, eine **kleine Gruppe** von Menschen zu identifizieren (Problem der **Unbalanciertheit**)

Wenn es **einfache Regeln zur Entscheidungsfindung gäbe, würden wir sie schon kennen**.

Daher machen sie immer Fehler.

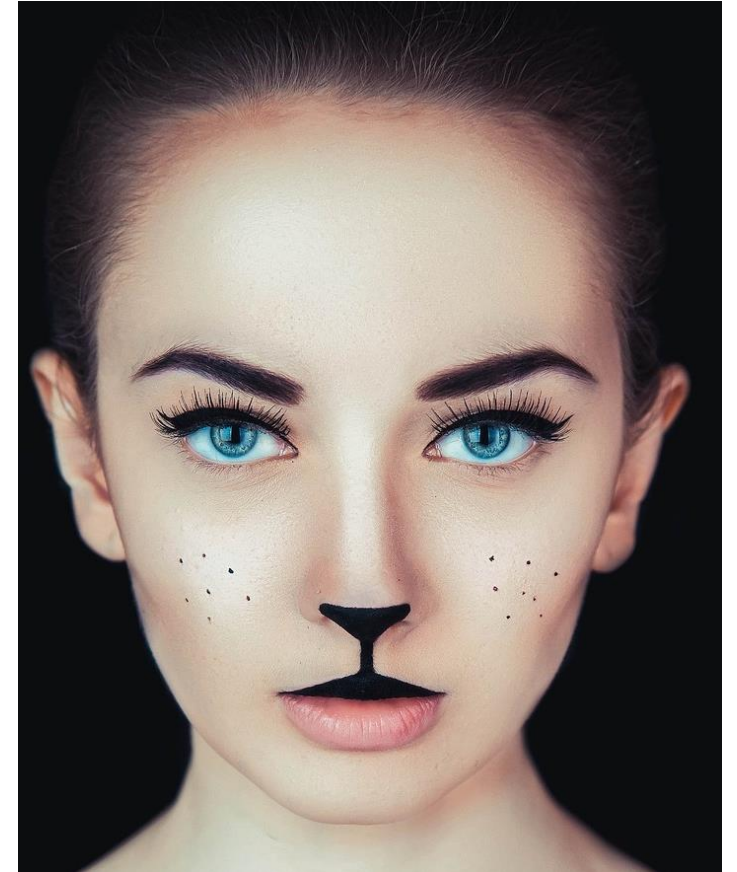


Statistische Vorhersagen über Menschen

Was bedeutet das eigentlich?

Zu 70% ein Krimineller....

- Wenn dieser Mensch eine Katze wäre und 7 Leben hätte, würde er in 5 davon wieder rückfällig werden...
- Nein!
- **Algorithmische Sippenhaftung**
 - Von 100 Personen, die „genau so sind wie dieser Mensch“, werden 70 wieder rückfällig;
 - Mitgefangen, mitgehungen;
 - In einer dem Delinquenten (der Delinquentin) völlig unbekanntem, algorithmisch bestimmten „Sippe“.



Probleme



- Aufmerksamkeitsökonomie der Richter und Richterinnen.
- „Best practice“ erfordert Nutzung der Software.
- Eine Nichtbeachtung der Empfehlung und gleichzeitige Fehleinschätzung wirkt viel schwerer als eine Beachtung der Empfehlung.
- Grundlegende Modellierung und Datenqualität kann schlecht sein.
- Der ins Gefängnis geschickte Delinquent **kann die Vorhersage prinzipiell nicht entkräften!**
 - Dies gilt auch für: Kreditvergaben, Bildungsangebote, Jobs, Personen, die von Drohnen erschossen werden oder als Terrorist eingesperrt werden, ...

Fazit: Lernende Algorithmen und gesellschaftlich relevante Entscheidungen



- Lernende Algorithmen bieten tatsächlich die Chance, in vielen Fällen, in denen der Mensch keine Regelmäßigkeiten mehr finden kann,
- objektiv und neutral diese Regeln zu finden
- und für Entscheidungen (Entscheidungsvorbereitungen) aufzubereiten.
- Sie (müssen) Fehler machen.
- Wenn wir gesellschaftlich nicht diskutieren, was dabei unsere Prioritäten sind und welche Qualitätskontrollen notwendig sind, verlieren wir Macht an kleine Gruppen von Algorithmenbesitzern.
- Dazu kommen handwerkliche Fehler.

Terroristenidentifikation SKYNET



TOP SECRET//COMINT//REL TO USA, FVEY

We've been experimenting with several error metrics on both small and large test sets

Training Data	Classifier	Features	100k Test Selectors		55M Test Selectors	
			False Alarm Rate at 50% Miss Rate	Mean Reciprocal Rank	Tasked Selectors in Top 500	Tasked Selectors in Top 100
None	Random	None	50%	1/23k (simulated)	0.64 (active/Pak)	0.13 (active/Pak)
Known Couriers	Centroid	All	20%	1/18k		
		Outgoing	43%	1/27k		
+ Anchory Selectors	Random Forest		0.18%	1/9.9	5	1
			0.008%	1/14	21	6

Random Forest trained on Known Couriers + Anchory Selectors:

- 0.008% false alarm rate at 50% miss rate
- 46x improvement over random performance when evaluating its tasked precision at 100

Windows
Wechseln
aktivieren

TOP SECRET//COMINT//REL TO USA, FVEY

<https://theintercept.com/document/2015/05/08/skynet-courier/>

<https://theintercept.com/2015/05/08/u-s-government-designated-prominent-al-jazeera-journalist-al-qaeda-member-put-watch-list/>

Top-“Kurier“ der Terroristen laut Algorithmus ist...



TOP SECRET//COMINT//REL TO USA, FVEY

The highest scoring selector that traveled to Peshawar and Lahore is PROB AHMED Z Aidan

The image shows a map of Pakistan with a purple path connecting several locations: Quetta, Miram Shah, Wana, Faisalabad, and Lahore. A profile window for 'PROB AHMED MUWAFAK ZAIDAN' is overlaid on the right. The profile includes a photo of a man in a suit and a list of attributes: 'TIDE Person Number: [redacted]', 'MEMBER OF AL QAEDA', 'MEMBER OF MUJAHID BROTHERHOOD', and 'WORKS FOR AL JAZEERA'.

Miles
Kms 100 200 300

Windows
Wahlseil

Algorithmen und Teilhabe



- Personalisierte, lernende Software als Assistenz erlaubt z.B. mehr Teilhabe für Ältere (Ambient Assisted Living), für Ausländer im deutschen Bildungssystem (z.B. Einblenden von Begriffen in Muttersprache) und Analphabeten.
- Chance für „seltene“ Gruppen nach menschlichen Maßstäben: z.B. Lernbehinderungen, seltene Krankheiten, Hochbegabung.
- Könnte deutlich objektiver sein als menschliche Entscheiderinnen und Entscheider (wenn sie gut gemacht sind, ausreichend unverzerrte Daten vorliegen und sie konstant überwacht werden).

Algorithmen und Teilhabe



- Klassifiziert uns nach uns unbekanntem Leitlinien trainiert mit unbekanntem Daten, mit nicht ausgehandelten Qualitätsansprüchen (keine gleichberechtigte Teilhabe)
- Könnte uns alle überfordern im Bereich lebenslanges Lernen: wann welche Weiterbildung machen? Trifft besonders weniger gut Ausgebildete.
- Hinterlässt besonders reichhaltige Datenspuren zur Persönlichkeit.

Algorithmen und Teilhabe



- Nicht alle werden den Algorithmen gleichmäßig unterworfen sein. Je individueller (je reicher, gebildeter, eingebetteter), desto weniger, denn Vorhersage ist nur da möglich, wo es viele gibt, und nur da nötig, wo jemand nicht schon bekannt ist.
- Gefahr der Granularisierung der Gesellschaft in Kleinstgruppen (ins. Versicherungen, aber auch Transportkosten?)
- Einmal Ausgeschlossene können sich nicht mehr beweisen: Kredit, Gefängnis, Jobinterview- und vergabe, Studienplatzvergabe, ...

Der Raser



- Gestern wurde ein YouTuber zu zwei Jahren und neun Monaten Haft verurteilt.
- Auf einer Raserfahrt überfuhr er einen betrunkenen Fußgänger.
- Die Helmkamera war nicht an.
- Mit seinem YouTube-Kanal hat er 80.000 Abonnenten, und verdiente im ersten Halbjahr ca. 2.200 Euro.
- Wie funktioniert das eigentlich?
- Und was hat das mit der nächsten Bundestagswahl zu tun?

