



Unstatistik des Monats

Pressemitteilung

## Big Data knackt Ihre Psyche

„Zeige mir deine ‚Facebook-Likes‘ und ich sage dir, wer du bist.“ Wie ein Lauffeuer verbreitete sich im Dezember die Botschaft, dass ein derartiger Facebook-Algorithmus Donald Trump zum Wahlsieg verholfen haben soll. Viele renommierte Medien stellen diese Behauptung in Frage, denn bewiesen hat [der auslösende Artikel im Schweizer „Das Magazin“](#) den Zusammenhang keinesfalls. Doch warum wird nicht auch der Algorithmus selbst hinterfragt? Grund dazu gibt es, schließlich behauptet dessen Entwickler, dass er aufgrund von „Gefällt-mir“-Angaben, psychometrischer Tests und dem Facebook- Profil einer Person genaue Aussagen über ihre Persönlichkeit treffen könne. So könne, wie der Artikel im Magazin behauptet, mit einer Genauigkeit von 88% vorhergesagt werden, was für eine sexuelle Orientierung ein Mann hat. [„Facebook knackt Ihre Psyche“](#) verriet uns denn auch chip.de.

Dabei ist diese Zahl gar nicht die Genauigkeit der Prognose. Tatsächlich bedeutet das Ergebnis: Nimmt man je eine Person pro Gruppe, also einen homosexuellen und einen heterosexuellen Mann, so kann man diese mit einer Wahrscheinlichkeit von 88% ihren richtigen Gruppen zuordnen. Die Prozentzahl bemisst die so genannte „Area under the Curve“ und bezieht sich auf den Vergleich zwischen den Gruppen, nicht auf die Prognosegüte an sich.

Doch auch eine echte Wahrscheinlichkeitsaussage wäre mit Vorsicht zu genießen. Es reicht für unsere Zwecke aus, von einer groben Schätzung von rund 10% Homosexuellen in der Gesamtbevölkerung auszugehen. In einer Gruppe von 10.000 Personen wären dann 9.000 hetero- und 1.000 homosexuell. Wer alle Menschen als heterosexuell klassifizierte, überträfe den Algorithmus schon um 2 Prozentpunkte, läge aber bei den Homosexuellen sicher falsch. Ein etwas komplexerer Algorithmus, der in beiden Gruppen eine Korrektheit von 88% besäße, würde in der ersten Gruppe 7.920 Personen als hetero- und 1.080 fälschlicherweise als homosexuell einschätzen. In der zweiten Gruppe werden 880 Personen korrekt eingeschätzt, 120 falsch. Aufaddiert wird also für 1.960 Personen die Aussage getroffen, dass sie homosexuell sind. Davon sind aber tatsächlich nur 880 homosexuell, was zu einer Treffsicherheit von nur etwa 45% führt. Das ist eine ziemlich enttäuschende Leistung, die sich unmittelbar aus dem Satz von Bayes ergibt.

Dortmund/  
Essen/Berlin,  
20. Dezember 2016

### Algorithmus formalisiert Alltags-Klischees und liegt oft daneben

Bei Frauen, so wird berichtet, könne man die sexuelle Orientierung nur mit 75% Sicherheit vorhersagen. Mit der gleichen Analyse wie oben ergibt sich daraus, dass nur 25% von allen als homosexuell diagnostizierten Frauen tatsächlich homosexuell sind. Die meisten sind Fehldiagnosen.

Also, wenn Sie wissen, ob sie homosexuell sind, dann sagt Ihnen der Algorithmus nichts Neues. Wenn Sie sich nicht sicher sind und Sie die Diagnose „homosexuell“ erhalten, dann ist die Diagnose wahrscheinlich ein Irrtum. Wie wir schon zuvor bei der Vorhersage von Pankreaskrebs durch die Microsoft-Suchmaschine Bing berichteten: Viel Lärm um Big Data.

Selbst wenn alle Daten korrekt wären, ist der Algorithmus kaum mehr als eine Formalisierung von Alltags-Klischees. Wenn ein Mann freimütig erzählt, dass er sich brennend für Make-Up und Mode interessiert, müssen wir uns selbstkritisch fragen, wie gefeit wir vor unwillkürlichen Einschätzungen sind. So folgert auch der Algorithmus: Männer, die die Kosmetikmarke MAC mögen, sind mit höherer Wahrscheinlichkeit homosexuell. Eher sollten wir uns also darüber Sorgen machen, wie sehr Facebooks Algorithmus die in unseren Köpfen verankerten Vorurteile bestärkt.

-----

Ansprechpartnerin:

Katharina Schüller (STAT-UP),

Tel.: (089) 34077-447

*Mit der „Unstatistik des Monats“ hinterfragen der Berliner Psychologe Gerd Gigerenzer, der Dortmunder Statistiker Walter Krämer und RWI-Vizepräsident Thomas K. Bauer jeden Monat sowohl jüngst publizierte Zahlen als auch deren Interpretationen. In diesem Monat hat die Gastautorin Katharina Schüller, Geschäftsführerin und Gründerin von STAT-UP, die „Unstatistik“ verfasst. Alle „Unstatistiken“ finden Sie im Internet unter [www.unstatistik.de](http://www.unstatistik.de).*