Scientifica - Was Daten verraten

„Schere, Stein, Papier: Schlagen Sie die Statistik-Maschine!“ lautet die Aufforderung der Statistiker an die Besucherinnen und Besucher der diesjährigen Scientifica. Nicolai Meinshausen (NM) und Martin Mächler (MM) vom Seminar für Statistik haben die „Statistik-Maschine“ entwickelt, gegen die Besucherinnen und Besucher am Stand „*Schere, Stein, Papier“* spielen können. Im folgenden Interview erklären sie weshalb sie dieses Spiel gewählt haben und wie sie es mit statistischen Methoden schaffen, das menschliche Spielverhalten vorauszusagen.

Was war Ihre Motivation an der Scientifica mit einem eigenen Stand dabei zu sein?

NM: Das diesjährige Motto “Was Daten verraten“ hat uns motiviert. Wir hatten das Gefühl, dass wir als Statistiker nicht aussen vor sein sollten, denn das Motto fasst sehr gut zusammen, was die Statistik ausmacht.

MM: Das Wort „verraten“ ist dabei ausschlaggebend. Es geht nicht darum Daten zu organisieren, sondern aus ihnen Schlüsse über die Wirklichkeit zu ziehen.

Warum haben Sie das *Spiel Schere, Stein, Papier* ausgewählt?

NM: Es sollte etwas sein, das die Leute schnell verstehen, das man leicht erklären kann und das gleichzeitig die Tür offen lässt, um mit den Leuten ins Gespräch zu kommen, zum Beispiel darüber wie man aus den Daten lernen und Vorhersagen machen kann.

Bei diesem Spiel setzten Sie voraus, dass eine Person nicht zufällig agiert, sondern dass sie nach einem bestimmten Muster vorgeht. Woher wissen Sie das?

NM: Es wurde schon seit langem beobachtet, dass es Menschen generell sehr schwer fällt zufällige Sequenzen zu generieren.

MM: Wenn eine Person aufgefordert wird sich vorzustellen hundert Mal eine Münze zu werfen und das jeweilige Ergebnis zu notieren, ohne dass sie die Münze tatsächlich geworfen hat, dann kann jeder Experte sofort den Unterschied zu einem echten Münzwurf erkennen.

Das heisst Menschen können nicht zufällig denken?

MM: Ja, Menschen haben das Gefühl, der Zufall ist anders als er ist. Zum Beispiel beim Münzwurf hat der Mensch das Gefühl, Zufall heisst möglichst viel abzuwechseln. Und zwar häufiger als der tatsächliche Zufall abwechseln würde.

Das heisst wenn Sie eine zufällige Sequenz notieren müssten, könnten Sie das besser, weil Sie den Zufall besser verstehen?

MM: Das glaube ich nicht. Wenn wir zufällig sein wollen, nehmen wir den Computer.

Für die Vorhersage der Spielzüge bei *Schere, Stein, Papier* brauchen Sie ein Datenmodell, dass das Verhalten des menschlichen Spielers voraussagen kann. Wie haben Sie dieses generiert?

NM: Wir haben verschiedene Strategien programmiert und jede macht etwas sehr Einfaches: Eine Strategie geht davon aus, dass wenn ein Spieler gewonnen hat, wiederholt er denselben Spielzug nochmals, wenn er verliert macht er hingegen etwas anderes. Bei einer weiteren Strategie wiederholt die Spielerin den Spielzug nie, sondern wechselt immer ab. So kann das Modell für jeden Menschen, der gegen die Maschine spielt vorhersagen welche Strategien er gerade verwendet.

MM: Die Person wird nicht immer gleich spielen sondern passt ihre Strategie dem an was gerade passiert. Das heisst, sie ändert ihre Strategie dauernd. Und jede Person spielt anders.

NM: Die Maschine analysiert wie die Spielerin spielt und stellt sich dann darauf ein. Wir haben unsere Doktoranden und Doktorandinnen damit spielen lassen und das System so angepasst, dass es gut funktioniert.

Wie lange hat diese Anpassung gedauert?

NM: Wir haben ein paar tausend Spielzüge von etwa zehn bis fünfzehn Doktorandinnen und Doktoranden analysiert. Damit haben wir das System optimal angepasst. Aber das System stellt sich zusätzlich noch auf jeden neuen Spieler ein, der gerade spielt.

Wie wird am Stand gespielt?

NM: Die Leute spielen eine Sequenz von 25 Spielen auf einem Tablet. An einem zweiten Bildschirm wird jeweils der nächste Spielzug vorausgesagt und angezeigt, ohne das die spielende Person diesen sieht, aber die anderen Standbesucher schon.

Verwenden Sie die Methoden für die Analyse des Spielverhaltens auch in der Forschung?

NM: Es geht bei unserer Forschung prinzipiell darum, gute Vorhersagen zu ermöglichen. Beispiele von Fragestellungen sind: Wie reagieren Zellen auf genetische Veränderungen? Wie sicher können wir den Klimawandel vorhersagen? Ein verbindender Aspekt in moderner Datenanalyse ist dabei, dass die wenigen guten und hilfreichen Daten oft in einem grossen Berg von eher nutzlosen Daten verborgen sind: Wir suchen eine sprichwörtliche Nadel im Heuhafen.