

David gegen Goliath

Für ein genaueres Bewertungssystem hat der Online-DVD-Verleih Netflix ein saftiges Preisgeld ausgeschrieben. Zwei junge Computerfreaks aus Graz mischen beim Kampf um die Million kräftig mit



Wenn sie durchhalten, winkt ihnen eine Million Dollar: Michael Jahrer (links) und Andreas Töschler setzen beim derzeit weltgrößten Informatikerwettbewerb ganz auf Nullen und Einsen.

Foto: Gudrun Profller; Montage: Lukas Friesenbichler

Es fehlte nicht viel, und Andreas Töschler und Michael Jahrer, zwei Telematik-Studenten aus Graz, hätten das Handtuch geworfen. Wie die Mittzwanziger es auch anpacken: Gegen das Tempo, das Forschungsabteilungen von Konzernen und Universitäten vorlegten, kamen die Studenten beim Netflix-Wettbewerb (www.netflix-prize.com) lange nicht an.

Dann aber spuckten ihre Computer, die in ihren Arbeitszimmern Tag und Nacht auf Hochtouren laufen, plötzlich immer exaktere Ergebnisse aus und kapultierten das Programmierduo namens BigChaos ins vorderste Spitzenfeld der täglich aktualisierten Zwischenrangliste. Nach einer Aufholjagd liefern sie die beiden jetzt sogar mit der Weltelite ein Duell um die Führung: Der bislang uneinholbar scheinenden Softwareabteilung des US-Konzerns AT&T sind sie seit Anfang Mai dicht auf den Fersen.

Und das, obwohl die Infrastruktur der Studenten nicht gerade Weltklasse ist: Anders als die Konkurrenz haben die privaten PCs der zwei Steirer die Resultate der Versuchsläufe viel langsamer vorliegen, sie können nicht so viele Rechenoperationen gleichzeitig durchführen.

Mit dem derzeitigen Platz 2 steht für die Studenten der Technischen Universität Graz plötzlich mehr auf dem Spiel als ihre Freizeit. Denn für die Verbesserung seiner Kundenempfehlungen lässt der größte amerikanische Online-DVD-Verleiher Netflix einiges springen: Immerhin eine Million Dollar warten auf denjenigen, der es schafft, die Treffsicherheit der Filmempfehlungen um ein Zehntel zu verbessern. Um diese halbwegs präzise vorhersagen zu können, be-

darf es komplexer Rechenoperationen, die Experten auf dem Gebiet der Künstlichen Intelligenz einiges abverlangen.

Die Informatiker müssen einen Algorithmus entwickeln, mit dem sich die individuellen DVD-Bewertungen der Kunden um zehn Prozent exakter vorhersagen lassen als mit der firmeneigenen Software. Netflix selbst war offenbar mit seinem Latein am Ende, wie sich die Treffsicherheit noch verbessern ließe. Also lagerte der Online-DVD-Verleih die Problemlösung einfach an die weltweite Forschergemeinde aus – die Auslagerung der Intelligenz an die Masse nennt man Crowdsourcing.

„Der Wettbewerb könnte noch vor Oktober entschieden sein“, zeigen sich die steirischen Million-Dollar-Jäger zuversichtlich – auch wenn so mancher Experte Zweifel hegt, ob die Zehn-Prozent-Grenze überhaupt erreichbar ist. „Diese Hürde ist gut durchdacht. Wahrscheinlich ist es machbar, sicher weiß das natürlich niemand“, meinen Töschler und Jahrer. Ein Gutteil des Weges liegt jedenfalls hinter ihnen, mittlerweile haben sie die Treffsicherheit bereits um 8,84 Prozent verbessert.

Das Kriterium, nach dem der prozentuelle Fortschritt berechnet wird, ist die durchschnittliche Abweichung der prognostizierten von den tatsächlichen Punktebewertungen der Kunden. Derzeit liegen die Vorhersagen der Grazer bei einem Fünf-Sterne-Raster im Durchschnitt deutlich weniger als einen Stern daneben (konkret: nur um 0,8673 Sterne).

Als das US-Unternehmen Netflix im Oktober 2006 den Wettbewerb ausgeschrieben und zu diesem Zweck ein riesiges anonymisiertes Datenpaket freigab,

ging weltweit ein Raunen durch die Informatikerreihen. Für die Experten des sogenannten „Maschinellen Lernens“, die es gewohnt sind, mit viel kleineren Datenmengen vorliebzunehmen, stellt der Wettbewerb einen seltenen Glücksfall dar: 100 Millionen Bewertungen, die 480.000 Netflix-Kunden zwischen 1998 und 2005 zu 18.000 Filmtiteln abgegeben haben, stehen ihrem Forscherdrang zur Verfügung. Ein Experimentierfeld, wie es die Programmierszene in dieser Fülle wohl noch nie vorgefunden hat.

Allerdings auch eines, bei dem Datenschützer den warnenden Zeigefinger hoben – und prompt kam es auch zu einem peinlichen Zwischenfall. Einigen Hackern gelang es, die anonymisierten Netflix-Kundenbewertungen mit den nicht anonymisierten User-Filmbewertungen der Internet Movie Database abzugleichen und sie so wieder konkreten Personen zuzuordnen.

„In unserem Forschungsgebiet“, sagt Töschler, „ist das zweifellos der größte Wettbewerb, den es je gegeben hat.“ Mehr als 30.000 Teams aus 170 Ländern machten sich bislang an die Arbeit, warfen ihre Rechner an und wühlten sich unermüdlich auf der Suche nach der perfekten Formel durch die Datenmassen. Welchem Erfolgsrezept sie ihren Durchbruch verdanken, wollen die beiden Steirer, die den Wettbewerb im Alleingang bestreiten, nicht verraten. Nur so viel: „Falls jemand das Ziel wirklich erreichen sollte, dann bestimmt nicht mit einem Einzelalgorithmus.“

Für das findige US-Unternehmen ist der Wettbewerb kein schlechtes Geschäft: Denn statt für teure Honorare eine

einzelne Softwarefirma anzuheuern, tüfteln und schufteten nun gleich zehntausende Forscherteams, um ein ausgeklügeltes Empfehlungssystem zu entwickeln, dem die Konkurrenz sicherlich nicht so bald das Wasser reichen kann. Während diese Experten jahrelang ihre Arbeit tun, ohne einen einzigen Cent Lohn zu verlangen, kann man sich bei Netflix entspannt zurücklehnen und sich zudem die Hände reiben über all den kostenlosen Werberummel, den das plötzlich entflammte Medieninteresse am Netflix Prize in Bewegung setzt.

Zwischendurch spendiert das Unternehmen zwar einen eher bescheidenen „Fortschrittspreis“ – der Gewinner musste natürlich seinen Algorithmus offenlegen, der umgehend bei Netflix zum Einsatz kommt –, ansonsten heißt es jetzt nur noch abwarten. Wenn eines der Teams das Plaziert erreicht, kassiert es einen schönen Batzen Geld – alle anderen Mitstreiter aber gehen leer aus.

Doch die Chancen stehen gut, dass auch die übrigen Teams am Ende nicht völlig durch die Finger schauen. Auch dem steirischen Studentenduo dämmert langsam, dass sie die Früchte ihres Netflix-Abenteuers selbst dann in bare Münze umsetzen könnten, wenn aus dem Millionen-Coup vielleicht doch nichts werden sollte. Deshalb haben die beiden Computerfreaks vor kurzem ihre zwei ehemaligen Schulkollegen Georg Profller und Michael Schrotter mit an Bord geholt. Die sollen aus dem unverhofften Geniestreich eine solide Firma machen. Die Palette an denkbaren Anwendungen für Maschinelles Lernen ist jedenfalls groß genug.

Roman Urbaner ■