

Netflix Turned to Computer Science Professor for Million Dollar Competition

September 21, 2009

Daniel Kane

Computer science professor and data mining expert Charles Elkan contributed to the structure of the Netflix Prize.

Netflix today awarded a \$1 million to BellKor's Pragmatic Chaos -the grand prize winner of the Netflix Prize and the team that created the most accurate movie recommendation system. This is the final stage in the Netflix Prize, a worldwide data mining competition for which University of California, San Diego computer science professor Charles Elkan has served as a contest designer, consultant and judge for the past three years.

Netflix offered a \$1 million prize to the person or team who could develop a system that is more than 10 percent more accurate than the in-house Netflix system for predicting what movies customers will like based on their movie preferences. Charles Elkan, the UC San Diego Jacobs School of Engineering professor, runs an annual data mining competition for students and postdoctoral researchers, sponsored by FICO. This expertise positioned Elkan to make a series of contributions to the structure of the Netflix Prize competition.

First, Elkan advocated for an online leaderboard to allow both spectators and contestants to monitor the competition from the Web.

Elkan also supported the idea that the competition switch into a final month-long phase as soon as one team reached the 10 percent improvement mark. "If you want a competition to promote the growth of knowledge, you let people know the task is feasible and see if anyone else can reach the goal even better," explained Elkan.

"The Netflix competition has raised awareness among the public and various business communities of the availability and usefulness of data mining," said Elkan.

The Netflix recommendation challenge is a special case of "collaborative filtering" and is similar to the problem that Amazon.com must solve in order to recommend books, movies and other products to individual customers based on what else they have looked at and purchased.

"Almost any business that is providing products or services has the issue of getting the right products and services in front of the people who want them. Traditional marketing is a shotgun approach. Focusing just on blockbusters is another option. But with data mining and the Internet, you can do much more individualized targeting. There are still a lot of opportunities for further research in this area," said Elkan.

For shopping Web sites, Elkan explained, past purchases can be predictive of future purchases, but it's not that simple. "People generally only buy one digital camera. But if you looked at a specific digital camera, you might buy some other product. That is one of the questions researchers and businesses alike are working on using various data mining techniques," said Elkan.

People have been doing research on these kinds of problems for more than 10 years, but the data sets have not been as big and difficult as the Netflix data set. The availability of the Netflix data set has enabled researchers to compare and extend methods, and has brought new ideas to the arena, according to Elkan.

"Attacking the problem in an open minded data driven way is advantageous. It's possible your purchasing behavior regarding digital cameras is predictive of your vacation purchasing behavior, or maybe it's the other way around," said Elkan who will be giving an invited talk on recommender systems at a computational advertising workshop organized by the National Institute of Statistical Sciences on November 4, 2009.

<http://www.niss.org/event/niss-affiliates-workshop-computational-advertising>

Media Contact: Daniel Kane, 858-534-3262 or dbkane@ucsd.edu

