In this lecture, we'll be discussing the story of Netflix and how their recommendation system is worth a million dollars.

Through this example, we'll introduce the method of clustering.

Netflix is an online DVD rental and streaming video service.

Customers can receive movie rentals by mail, and they can also watch selected movies and TV shows online.

Netflix has more than 40 million subscribers worldwide and has an annual revenue of $3.6 billion.

A key aspect of the company is being able to offer customers accurate movie recommendations based on a customer's own preferences and viewing history.

From 2006 through 2009, Netflix ran a contest asking the public to submit algorithms to predict user ratings for movies.

This algorithm would be useful for Netflix when making recommendations to users.

Netflix provided a training data set of about 100 million user ratings and a test data set of about three million user ratings.

They offered a grand prize of $1 million to the team who could beat Netflix's current algorithm, called Cinematch, by more than 10% measured in terms of root mean squared error.

Netflix believed that their recommendation system was so valuable that it was worth a million dollars to improve.

The contest had a few rules.

One with that if the grand prize was not yet reached, progress prizes of $50,000 per year would be awarded for the best results so far, as long as it was at least a 1% improvement over the previous year.

Another rule was that teams must submit their code and a description of the algorithm to be awarded any prizes.

And lastly, if a team met the 10% improvement goal, a last call would be issued, and 30 days would remain for all teams to submit their best algorithm.

So what happened?

The contest went live on October 2, 2006.

By October 8, only six days later, a team submitted an algorithm that beat Cinematch.

A week later, on October 15, there were three teams already submitting algorithms beating Cinematch.

One of these solutions beat Cinematch by more than 1%, already qualifying for a progress prize.

The contest was hugely popular all over the world.

By June, 2007, over 20,000 teams had registered from over 150 countries.

The 2007 progress prize went to a team called BellKor, with an 8.43% improvement over Cinematch.

The following year, several teams from across the world joined forces to improve the accuracy even further.

In 2008, the progress prize again went to team BellKor.

But this time, the team included members from the team BigChaos in addition to the original members of BellKor.

This was the last progress prize because another 1% improvement would reach the grand prize goal of 10%.

On June 26, 2009, the team BellKor's Pragmatic Chaos, composed of members from three different original teams, submitted a 10.05% improvement over Cinematch, signaling the last call for the contest.

Other teams had 30 days to submit algorithms before the contest closed.

These 30 days were filled with intense competition and even more progress.

But before revealing what happened, let's investigate how we could try to predict user ratings.

In the next video, we'll discuss how recommendation systems generally work.