

# ANÁLISIS DE DATOS MASIVOS

## CASO DE ESTUDIO: EL DESAFÍO NETFLIX

---

Blanca Vázquez

4 de noviembre de 2024

## Sistema de recomendación



Desafío: construir un algoritmo de filtrado colaborativo para predecir los *ratings* de los usuarios

## Contexto

- En el 2006, Netflix lanzó un desafío de un millón de dólares al equipo que lograra mejorar el rendimiento predictivo de *Cinematch*.
- En ese año, Netflix contaba con 7.5 millones de usuarios y 1,500 películas.
- Problema: la matriz de *ratings* estaba mayormente vacía.

## Cinematch

- El objetivo de *Cinematch* era rellenar toda la matriz de *ratings*.
- Posteriormente, generaba una recomendación, tomando las puntuaciones más altas.
- El desafío consistía llenar todas las predicciones para usuarios y películas.

## Cinematch

- *Cinematch* tenía un RSME de 0.8525 sobre el conjunto de prueba.
- El equipo ganador tenía que mejorar el rendimiento de *Cinematch* en un 10 % sobre el conjunto de prueba.

## Datos disponibles

- Netflix proporcionó un conjunto de entrenamiento de 100,480,507 *ratings* provenientes de 480,189 usuarios para 17,770 películas.
- Forma de los datos:  
`usuario, película, fecha, rating`

## Datos disponibles

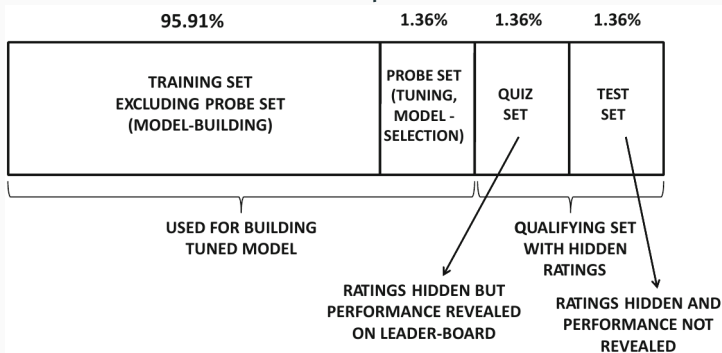


Imagen tomada de Aggarwal, 2016

## Participantes

- Al inicio de la competencia 3 equipos se inscribieron, para el 2007 habían más de 20,00 equipos de más de 150 países.
- En el 2009, el ganador fue *BellKor's Pragmatic Chaos* los cuales habían mejorado en un 10.09 % el RSME sobre el conjunto de prueba.





## Netfilx Prize

[Home](#) | [Rules](#) | [Leaderboard](#) | [Update](#)

### Netflix Prize: Forum

Forum for discussion about the Netflix Prize and dataset.

Announcement

Congratulations to team "BellKor's Pragmatic Chaos" for being [awarded the \\$1M Grand Prize](#) on September 21, 2009. This Forum is now read-only.

2009-09-18 16:58:04

#1

**prizemaster**  
Administrator  
From: Netflix HQ  
Registered: 2006-06-29  
Posts: 181  
[Website](#)

It is our great honor to announce the \$1M Grand Prize winner of the Netflix Prize contest as team [BellKor's Pragmatic Chaos](#) for their verified submission on July 26, 2009 at 18:18:28 UTC, achieving the winning RMSE of 0.8567 on the test subset. This represents a 10.06% improvement over Cinematch's score on the test subset at the start of the contest. We congratulate the team of Bob Bell, Martin Chabbert, Michael Jahrer, Yehuda Koren, Martin Pott, Andreas TÄtscher and Chris Volinsky for their superb work advancing and integrating many significant techniques to achieve this result.

The Prize was awarded in a ceremony in New York City on September 21st, 2009. We will post a video on this forum of the presentation the team delivered about their Prize algorithm. In accord with the [Rules](#) the winning team has prepared a system description consisting of three papers, which we both make public below.

Team BellKor's Pragmatic Chaos edged out team [The Ensemble](#) with the winning submission coming just 24 minutes before the conclusion of the nearly three-year-long contest. Historically the [Leaderboard](#) has only reported team scores on the quiz subset. The Prize is awarded based on teams' test subset score. Now that the contest is closed we will be updating the Leaderboard to report team scores on both the test and quiz subsets.

To everyone who participated in the Netflix Prize: You've made this a truly remarkable contest and you've brought great innovation to the field. We applaud you for your contributions and we hope you've enjoyed the journey. The Netflix Prize contest is now closed.

We will soon be launching a new contest, Netflix Prize 2. Stay tuned for more [details](#).

The winning team's papers submitted to the judges can be found below. These papers build on, and require familiarity with, work published in the [2008 Progress Prize](#).

Y. Koren, ["The BellKor Solution to the Netflix Grand Prize"](#) (2009).

A. TÄtscher, M. Jahrer, R. Bell, ["The BioChaos Solution to the Netflix Grand Prize"](#) (2009).

M. Pott, M. Chabbert, ["The Pragmatic Theory solution to the Netflix Grand Prize"](#) (2009).

Offline

[FAQ](#) | [Forum](#) | [Netflix Home](#)

© 1997-2009 Netflix, Inc. All rights reserved.

## Equipo: BellKor's Pragmatic Chaos

