Build a Better Netflix, Win a Million Dollars?

Lester Mackey

2012 USA Science and Engineering Festival
• Rents & streams movies and TV shows
• 100,000 movie titles
• 26 million customers

Recommends “Movies You’ll ❤️”
Recommending Movies You’ll ♥️

Rate what you’ve seen to discover suggestions for you

- Mrs. Doubtfire
  - Haven't Seen It

- Iron Man 2
  - Hated it!

- Food, Inc.
  - Loved it!
Recommending Movies You’ll ❤️
Inception

2010  PG-13  148 minutes

Dom Cobb earns a tidy sum infiltrating the dreams of corporate titans to steal their most closely held secrets.

Starring: Leonardo DiCaprio, Joseph Gordon-Levitt
Director: Christopher Nolan
Genre: Sci-Fi & Fantasy
Availability: DVD and Blu-ray

Our best guess for LESTER: 4.4

Recommended based on your interest in: *Batman Begins*, *The Matrix* and *Memento*
How This Works

I don’t like this movie.

Now I’m unhappy!

Top Secret Computer Program: Cinematch

Your Predicted Rating: 4 stars
Back at Netflix

How can we improve Cinematch?

Let’s have a contest!

What should the prize be?

How about $1 million?
The Netflix Prize

October 2, 2006

- Contest open to the world
- 100 million movie ratings released to public
- **Goal**: Create computer program to predict ratings
  - **$1 Million** Grand Prize for beating Cinematch accuracy by 10%
  - **$50,000** Progress Prize for the team with the best predictions each year

5,100 teams from 186 countries entered
Dinosaur Planet

David Lin

David Weiss

Lester Mackey

Team Dinosaur Planet
The Ratings

• Training Set
  – What computer programs use to learn customer preferences
  – Each entry: 100,500,000 ratings in total
  – 480,000 customers and 18,000 movies

July 5, 1999
# The Ratings: A Closer Look

## Highest Rated Movies

<table>
<thead>
<tr>
<th>Movie</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Shawshank Redemption</td>
</tr>
<tr>
<td>Lord of the Rings: The Return of the King</td>
</tr>
<tr>
<td>Raiders of the Lost Ark</td>
</tr>
<tr>
<td>Lord of the Rings: The Two Towers</td>
</tr>
<tr>
<td>Finding Nemo</td>
</tr>
<tr>
<td>The Green Mile</td>
</tr>
</tbody>
</table>

## Most Divisive Movies

<table>
<thead>
<tr>
<th>Movie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fahrenheit 9/11</td>
</tr>
<tr>
<td>Napoleon Dynamite</td>
</tr>
<tr>
<td>Pearl Harbor</td>
</tr>
<tr>
<td>Miss Congeniality</td>
</tr>
<tr>
<td>Lost in Translation</td>
</tr>
<tr>
<td>The Royal Tenenbaums</td>
</tr>
</tbody>
</table>
How the Contest Worked

• Quiz Set & Test Set
  – Used to evaluate accuracy of computer programs
  – Each entry:

• Each team predicts Quiz Set and Test Set ratings once per day
• Netflix displays Quiz score on public Leaderboard
<table>
<thead>
<tr>
<th>Team Name</th>
<th>Best Score</th>
<th>% Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Grand Prize candidates yet</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Prize - RMSE &lt;= 0.8563</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Thought Gang</td>
<td>0.9413</td>
<td>1.06</td>
</tr>
<tr>
<td>Progress Prize 2007 - RMSE &lt;= 0.9419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wxyzconsulting.com</td>
<td>0.9430</td>
<td>0.88</td>
</tr>
<tr>
<td>Sparkling_Destiny</td>
<td>0.9488</td>
<td>0.27</td>
</tr>
<tr>
<td>Cinematch score on quiz subset - RMSE = 0.9514</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline0</td>
<td>0.9525</td>
<td>-0.12</td>
</tr>
<tr>
<td>CodeMonkey</td>
<td>0.9571</td>
<td>-0.60</td>
</tr>
<tr>
<td>Bjornson</td>
<td>0.9648</td>
<td>-1.41</td>
</tr>
<tr>
<td>jsnell</td>
<td>0.9670</td>
<td>-1.64</td>
</tr>
</tbody>
</table>
How the Contest Worked

• Quiz Set & Test Set
  – Used to evaluate accuracy of computer programs
  – Each entry:

• Each team predicts Quiz Set and Test Set ratings once per day

• Netflix displays Quiz score on public Leaderboard

• Test score is hidden – but best Test score wins!
  – 10% improvement ➔ $1 Million Grand Prize
  – Most improvement in 1 year ➔ $50,000 Progress Prize

Sept. 9, 2006
How to Predict Ratings

Nearest Neighbor Rule: **4-5% better** than Cinematic Match

- Find customer with the most similar ratings
- Use her rating as best guess for new guy’s rating
How to Predict Ratings

Nearest Neighbor Rule: 4-5% better than Cinematch

- Find movie with the most similar ratings
- Use its rating as best guess for new guy’s rating
Map of Movie Neighbors

Produced by Yifan Hu, joint work with Emden Gansner, Stephen Kobourov & Chris Volinsky, Data from Yehuda Koren, Copyright AT&T
Map of Movie Neighbors

Produced by Yifan Hu, joint work with Emden Gansner, Stephen Kobourov & Chris Volinsky, Data from Yehuda Koren, Copyright AT&T
The First Progress Prize

One day before the deadline...

<table>
<thead>
<tr>
<th>Rank</th>
<th>Team Name</th>
<th>Best Score</th>
<th>% Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BellKor</td>
<td>0.8728</td>
<td>8.26</td>
</tr>
<tr>
<td>2</td>
<td>Gravity</td>
<td>0.8750</td>
<td>8.03</td>
</tr>
<tr>
<td>3</td>
<td>Dinosaur Planet</td>
<td>0.8753</td>
<td>8.00</td>
</tr>
<tr>
<td>4</td>
<td>ML@UTorontoA</td>
<td>0.8787</td>
<td>7.64</td>
</tr>
<tr>
<td>5</td>
<td>Arek Paterek</td>
<td>0.8789</td>
<td>7.62</td>
</tr>
<tr>
<td>6</td>
<td>basho</td>
<td>0.8805</td>
<td>7.45</td>
</tr>
<tr>
<td>7</td>
<td>NIPS Reject</td>
<td>0.8808</td>
<td>7.42</td>
</tr>
<tr>
<td>8</td>
<td>Ensemble Experts</td>
<td>0.8841</td>
<td>7.07</td>
</tr>
</tbody>
</table>

Grand Prize - RMSE <= 0.8563

Progress Prize 2007 - RMSE: 0.9419
The Power of Teamwork

- First Progress Prize
  - joins forces with for 8.38% improvement
  - BellKor improves to 8.43%, wins $50,000

- Second Progress Prize
  - BigChaos joins BellKor for 9.44% and $50,000

- Grand Prize Team (GPT) founded by +
  - Anyone could join
  - The more you improve the GPT score, the bigger your share of the $1 million Grand Prize
The Last Call

- Remaining teams had 30 days to respond
- Over the next 24 days, **30 individuals** from **11 countries** combined forces to challenge BPC
The Ensemble

- Grand Prize Team
  - Gravity
    - Gabor Takacs, Istvan Pilaszy, Bottyan Nemeth, Domonkos Tikk
  - Dinosaur Planet
    - David Lin, Lester Mackey, David Weiss
  - Joe Sill
  - Ces Bertino
  - Dan Nabutovsky
  - William Roberts
  - Wojtek Kulik
  - Willem Mestrom
  - David Purdy

- Vandelay Industries!
  - Greg McAlpin
  - Bill Bame
  - Bo Yang
  - Chris Hefele
  - Jeff Howbert
  - Xiang Liang
  - Larry Ya Luo
  - Aron Miller
  - Steve Pagliarulo
  - Opera Solutions
    - Bruce Deng, Peng Zhou, Priyanka Rastog, Arvind Gangadha, Jacob Spoelstra
  - Craig Carmichael
  - Mike Linacre
  - Edward de Grijs
  - Clive Gifford
  - Feeds2
    - Nicholas Ampazis, George Tsagas

Learn more at http://the-ensemble.com/
The Road to the Grand Prize

• Next to Last Day
  – The Ensemble submits
  – 10.09% improvement on Quiz Set

• Final Day, 6:18pm
  – BellKor’s Pragmatic Chaos responds
  – 10.09% improvement on Quiz Set

• Final Day, 6:38pm
  – The Ensemble makes its final submission
  – 10.10% improvement on Quiz Set

• Final Day, 6:42pm: Contest closes
The Other Road to the Grand Prize

• Next to Last Day
  – The Ensemble submits
  – 10.05% improvement on Test Set

• Final Day, 6:18pm
  – BellKor’s Pragmatic Chaos responds
  – 10.06% improvement on Test Set

• Final Day, 6:38pm
  – The Ensemble makes its final submission
  – 10.06% improvement on Test Set

• Tie breaker: Time of submission
And then there were two...

Teams shown from first appearance in top 20.

Dinosaur Planet
Gravity
WDPU
Newman!
David Purdy
Craig Carmichael
Edward de Grijs
xiangliang
Mike Linacre
Aron Miller
Steve Pagliarulo (ESP)
Dace
Chef Ele
OfADifferentKind
Vandelay
Industries!
clueless
Grand Prize Team
William J.J. Roberts
photek
Green Circle
Expensive Lunch
blednotik
The Ensemble
The End

Pragmatic Theory
BellKor in BigChaos
BellKor
Big Chaos
1st Progress Prize
BellKor Paper Published
2nd Progress Prize
BellKor Paper Published
2007 2008 2009
8.43 9.44 10.05

Copyright AT&T