

BIG DATA

EMANUELE DELLA VALLE
18 MAGGIO 2018



Success story

The Big Data Behind Netflix's "House of Cards"

Emanuele Della Valle
DEIB - Politecnico di Milano
emanuele.dellavalle@polimi.it
<http://emanueledellavalle.org>

Netflix in numbers

- Netflix streaming service is said to account for **1/3 of peak time internet traffic** in the US.
- Last year, Netflix announced that it signed on **50 million** accounts worldwide.

BigData3

understanding viewing habits

- Netflix specialists wade through and capture important information from a variety of different analytics streams:
 - Personalization analytics
 - Messaging analytics
 - Content delivery analytics
 - Device analytics
 - ...

Back in 2006 ...

- Netflix was a **DVD delivery service**



The Netflix prize

- Central to the effectiveness of its DVD delivery service was its **recommendation engine**
- Netflix offered **\$1m** to the person that could come up with the **best algorithm for predicting how customers would rate a movie** based on their previous ratings.
- See <http://www.netflixprize.com/>

With four data points ...

- At the time, Netflix was mainly a DVD delivery service so programmers only had access to **four data points**:
 - Customer ID
 - Movie ID
 - Rating
 - The date the movie was watched

... Netflix predicts movie preferences

- The sophisticated winning approaches (see <http://www.netflixprize.com/leaderboard.html>) did not make it to production
- **A simple scalable** (based on Big Data technology) **approach made it**

Big Data

Nowadays, with more data points ...

- As streaming video became the primary focus, many **more data points** became available, giving insights into the customers
- These included...
 - Time of day something was watched
 - User gender and age (based on individual logins)
 - Time spent selecting movies
 - How often a movie or program was paused

... Netflix predicts "perfect situation"

- Netflix can now **analyze** how these factors impacted **viewers' enjoyment** (based on ratings given to movies).
- With no ads, Netflix's business model relies on customers who subscribe to their service for the long run. **The happier the customers are, the longer they will stay** subscribed to the service.
- Using these data points, Netflix developers built models to **predict the "perfect situation"** in which customers continuously receive programs they enjoy.

... creates "specific" category suggestions

- Another method of data analytics is movie tagging. When a viewer rates a movie highly, the Netflix algorithm suggests other movies they may enjoy based on matching tags
- The service has created **76,897 unique ways to describe types of movies.**
- These are called “alt-genres,” which is what leads to Netflix’s scarily specific category suggestions (e.g., “Comedies Featuring a Strong Female Lead”).

... bids on original programming

- When the series **House of Cards** began shopping around for a home, **Netflix** aggressively jumped on it, outbidding major cable networks with a massive **two-season order**.

BigData

Why the enthusiasm?

- **By analyzing its data**, Netflix saw that a large majority of its viewers enjoy programs :
 - *directed by David Fincher* (who directed Se7en, Fight Club and The Social Network)
 - *starring Kevin Spacey.*
- House of Cards was exactly that!



... promotes programs

- Netflix took a **data-driven approach to promote** House of Cards.
- For example, the company **modelled the show's cover image on the colours and styles for successful, similarly tagged programs**, to help draw new viewers in



Conclusion

- Netflix disrupted the TV industry using data analytics to **provide viewers with exactly the content they want.**

BigData

Credits

- The Big Data Behind Netflix's "House of Cards".
Hewlett Packard Enterprise Development
 - <https://www.hpematter.com/issue-no-8-winter-2016/big-data-behind-netflixs-house-cards>

Big Data

Thank you!
Any Question?

Emanuele Della Valle
DEIB - Politecnico di Milano
emanuele.dellavalle@polimi.it
<http://emanueledellavalle.org>