

CSC 411: Machine Learning in Action

Challenge: Movie Rating and Genre Prediction

Sanja Fidler

University of Toronto

Jan 22, 2015

Hands-On Experience

- A good (and fun!) way to learn the concepts we are talking about is to implement them and try them on real data.
- We prepared movie data:
<http://www.cs.toronto.edu/~fidler/teaching/2015/slides/CSC411/movies.zip>
- Try to predict **ratings** and **genres**! Report performance on piazza
- Questions:
 - ▶ What would be good features?
 - ▶ How does each method perform?
 - ▶ What happens if you use less vs more training data? Which method did better?
 - ▶ What happens if your features are low or high dimensional? Which method did better?
 - ▶ How did you choose your hyper-parameters?
- **Note: This is not an assignment!**

Putting Machine Learning to Practice

- Good advice by Andrew Ng for practical ML:

<http://cs229.stanford.edu/materials/ML-advice.pdf>

- In video:

<https://www.youtube.com/watch?v=TxJe4xeDI7g&feature=relmfu>

- Pedro Domingo's paper:

<https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>

- The data has:
 - ▶ **Train:** 700 movies
 - ▶ **Test:** 300 movies
- You may want to split the training set into train and validation
- Do **not** use test data for training or parameter tuning
- Each movie has lots of available meta-data, e.g.:
 - ▶ Cast
 - ▶ Director(s)
 - ▶ Writer(s)
 - ▶ Year of release
 - ▶ Storyline (short description of the movie)
 - ▶ Plot (longer description of the movie)
 - ▶ Box-office information (try not using this for rating prediction)
 - ▶ Keywords (try not using this for genre prediction)
- Report performance on test data

Kung Fu Panda

- This movie will be released on Jan 29 (next Friday):



IMDb Find Movies, TV shows, Celebrities and more... All

Movies, TV & Showtimes | Celebs, Events & Photos | News & Community | Watchlist

THE WEIGHT IS OVER
ANIMATED BY
KUNG FU PANDA 3

Kung Fu Panda 3 (2016) 154

PG | 95 min | Animation, Action, Adventure | 29 January 2016 (USA)

Not yet released

(voting begins after release) Metascore: **72/100**
Reviews: | 1 critic | 4 from Metacritic.com

Continuing his "legendary adventures of awesomeness", Po must face two hugely epic, but different threats: one supernatural and the other a little closer to his home.

Directors: Alessandro Carloni, Jennifer Yuh
Writers: Jonathan Aibel, Glenn Berger
Stars: Jack Black, Angelina Jolie, Dustin Hoffman | See full cast and crew »

<http://www.cs.toronto.edu/~fidler/teaching/2015/slides/CSC411/panda.mat>

Whose method can best guess what the rating will be?

- We will revisit this question in the end of Feb (when enough votes come in)
- Don't cheat: by e.g., going to watch the movie and giving it a score!

Kung Fu Panda

- This movie will be released on Jan 29 (next Friday):



IMDb Find Movies, TV shows, Celebrities and more... All

Movies, TV & Showtimes | Celebs, Events & Photos | News & Community | Watchlist

Kung Fu Panda 3 (2016) 154

PG | 95 min | Animation, Action, Adventure | 29 January 2016 (USA)

Not yet released

(voting begins after release) Metascore: **72/100**

Reviews: | 1 critic | 4 from Metacritic.com

Continuing his "legendary adventures of awesomeness", Po must face two hugely epic, but different threats: one supernatural and the other a little closer to his home.

Directors: Alessandro Carloni, Jennifer Yuh
Writers: Jonathan Aibel, Glenn Berger
Stars: Jack Black, Angelina Jolie, Dustin Hoffman | See full cast and crew »

<http://www.cs.toronto.edu/~fidler/teaching/2015/slides/CSC411/panda.mat>

Can you accurately predict its genres?