# NLP lab mid-term presentation:
predicting IMDB movie score of an individual based on their review

Seyithan Dag, Universität Bonn

21.12.2017

# today’s talk

- sentiment analysis
  - what is sentiment and sentiment analysis
  - tasks, usages and applications
  - challenges
- topic / problem statement
  - IMDB
- problems to tackle
- roadmap
- conclusion

# sentiment analysis

## what is sentiment and sentiment analysis

- sentiment = feelings/emotions
  - attitudes, opinions
  - nothing factual, entirely subjective
  - generalization assumption: binary polarity

- sentiment analysis: using NLP and/or machine learning methods to extract, identify, or otherwise characterize the sentiment content of a text unit

# sentiment analysis

## tasks, usages and applications

is a product review positive or negative?

is a customer satisfied or dissatisfied?

how are the acute reactions of people to a certain advertisement campaign?

how are the acute reactions of people to political events?

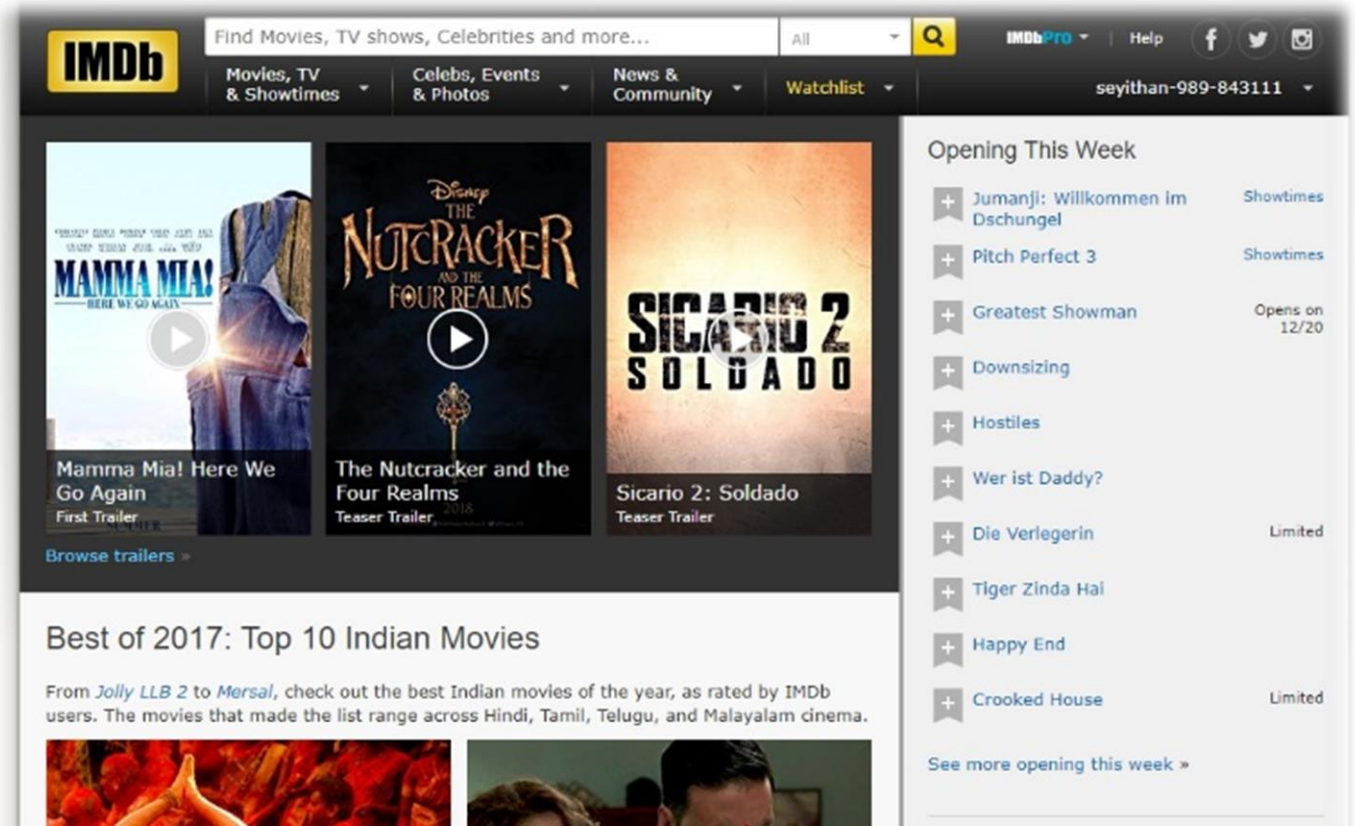
# sentiment analysis
## tasks, usages and applications

cross-domain applications

- sociology
- psychology
- political science
- business intelligence
- law

# topic / problem statement

predicting IMDB movie score of an individual based on their review



# topic / problem statement

predicting IMDB movie score of an individual based on their review

User Reviews

★★★★☆☆☆☆☆☆ **Is it a parody?**

14 December 2017 | by MoodyHank – See all my reviews

There are some well written and comprehensive reviews on here already so I won't waste your time in the details.

Essentially, the entire first half of The Last Jedi feels like a Star Wars parody. Instead of the odd moment of humour every now and again, every scene AND character is a comedian in TLJ. It's a crap gag a minute. And some of the great work done The Force Awakens to re-energise the franchise was just completely undone. You know that incredibly significant and poignant moment at the end of TFA as Rey hands the lightsabre to Luke? Just wait and see what happens next.

I've rated it 4 stars because the second half of the film does pick up but by then the damage is done. TLJ is a huge misstep in the franchise.

537 of 831 people found this review helpful. Was this review helpful to you? Yes No | Report this

# problems to tackle

- sentiments are subjective, i.e. no reference point of origin
- people express opinions in complex ways
- there might be irony or sarcasm involved
- text within the text within the text
- sudden change of expression or even topic
- particular case: gre words in reviews
- particular sad news: reviews and their scores do not always match

# problems to tackle

- particular sad news: reviews and their scores not always match

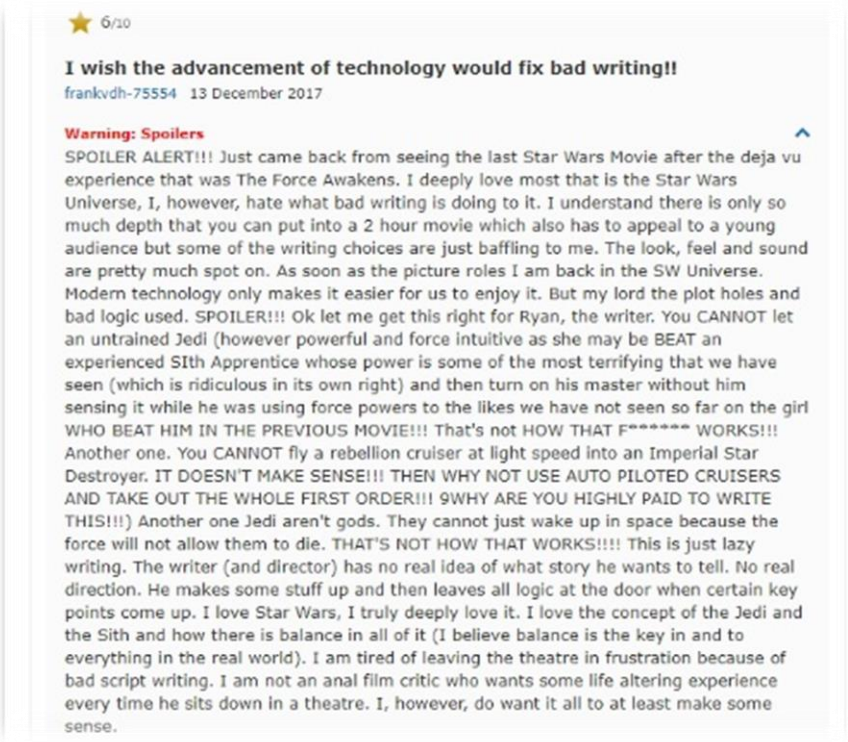
★ 5/10

**The Last Jedi is so bad on so many levels.**

btb90-387-617448 14 December 2017

**Warning: Spoilers**

I don't know where to start with this utter abominaton of a Star Wars movie. It's like Rian & Disney took everything that J.J. Abrams wrote and threw it out the window. There is no explaination who Ray and Snoke are and who were the Knights of Ren, Snoke is even killed off in the middle of the movie in a ridicilous way. The way they changed Luke Skywalker as a character is is just sad. All the hype where lukes blue lightsaber came from in The Force awakens was for nothing. There was not even a proper lightsaber battle in the end.



★ 6/10

**I wish the advancement of technology would fix bad writing!!**

frankvdh-75554 13 December 2017

**Warning: Spoilers**

SPOILER ALERT!!! Just came back from seeing the last Star Wars Movie after the deja vu experience that was The Force Awakens. I deeply love most that is the Star Wars Universe, I, however, hate what bad writing is doing to it. I understand there is only so much depth that you can put into a 2 hour movie which also has to appeal to a young audience but some of the writing choices are just baffling to me. The look, feel and sound are pretty much spot on. As soon as the picture roles I am back in the SW Universe. Modern technology only makes it easier for us to enjoy it. But my lord the plot holes and bad logic used. SPOILER!!! Ok let me get this right for Ryan, the writer. You CANNOT let an untrained Jedi (however powerful and force intuitive as she may be BEAT an experienced SIth Apprentice whose power is some of the most terrifying that we have seen (which is ridiculous in its own right) and then turn on his master without him sensing it while he was using force powers to the likes we have not seen so far on the girl WHO BEAT HIM IN THE PREVIOUS MOVIE!!! That's not HOW THAT F****** WORKS!!! Another one. You CANNOT fly a rebellion cruiser at light speed into an Imperial Star Destroyer. IT DOESN'T MAKE SENSE!!! THEN WHY NOT USE AUTO PILOTED CRUISERS AND TAKE OUT THE WHOLE FIRST ORDER!!! 9WHY ARE YOU HIGHLY PAID TO WRITE THIS!!!) Another one Jedi aren't gods. They cannot just wake up in space because the force will not allow them to die. THAT'S NOT HOW THAT WORKS!!!! This is just lazy writing. The writer (and director) has no real idea of what story he wants to tell. No real direction. He makes some stuff up and then leaves all logic at the door when certain key points come up. I love Star Wars, I truly deeply love it. I love the concept of the Jedi and the Sith and how there is balance in all of it (I believe balance is the key in and to everything in the real world). I am tired of leaving the theatre in frustration because of bad script writing. I am not an anal film critic who wants some life altering experience every time he sits down in a theatre. I, however, do want it all to at least make some sense.

# roadmap

reduction of dimension:

- instead of scores, polarity will be taken into consideration [0, 5] being negative, (5, 10] being positive

application of probabilistic graphical models to sentiment analysis

- celikyilmaz et. al. [1] uses probabilistic graphical models for analysing sentiments in tweets
- yates et. al. [2] utilizes bayesian inference to merge labelled sentences to analyse sentiment in user reviews

or sticking to only machine learning methods, or perhaps combining the two.

# any

questions?

# references


- [1] celikyilmaz, et.al., (2010). probabilistic model-based sentiment analysis of twitter messages. *IEEE*
- [2] yates, et. al., (2014). semi-supervised probabilistic sentiment analysis: merging labelled sentences with unlabelled reviews to identify sentiment. *proceedings of the american society for information science and technology*