



Predicting the Likeability of Movies Using Movie Scripts

Mahsa Shafaei (mshafaei@uh.edu)

RiTUAL Lab

Motivation

- Predicting if a movie will be successful or not
- Save a huge amount of money in movie industry



Motivation



- Budget: \$275,000,000
- IMDB Rating: 6.4
- Worldwide Gross: \$260,002,115



- Budget: \$206,582,000
- IMDB Rating: 5.3
- Worldwide Gross: \$181,982,519

Success Metrics

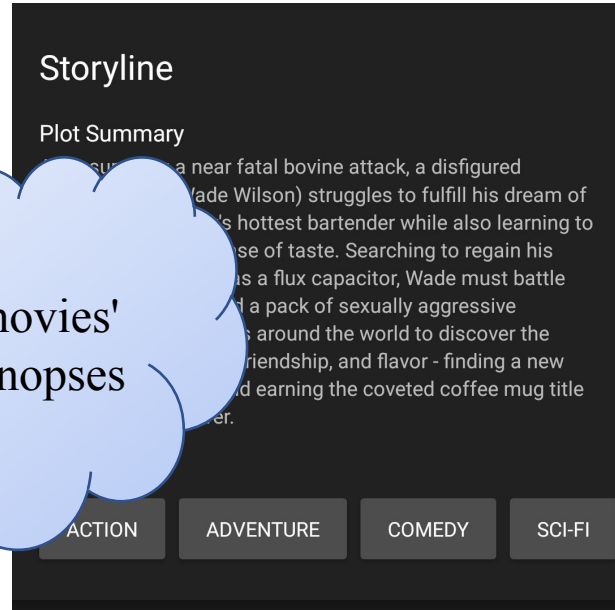
Reasons	IMDB Rating	Box-office Gross
Ticket price	✓	✗
Advertisements and competitive movies	✓	✗
Netflix and Amazon	✓	✗
Availability of data	✓	✗
Geographic popularity	✗	✓
Profitability	✗	✓

Related Work



used social
media like
Twitter and
YouTube

used movies'
plot synopses

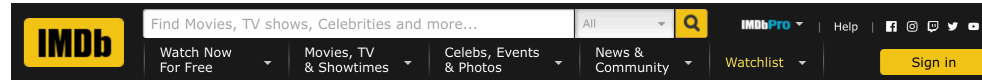


Extracted
features from
movies'
metadata

The IMDb logo, consisting of the letters 'IMDb' in a bold, black, sans-serif font, followed by a small yellow star icon inside a dark blue square.

Dataset

- Metadata (IMDB API)

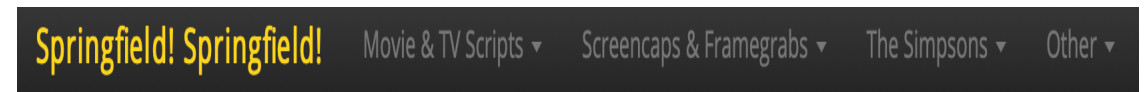


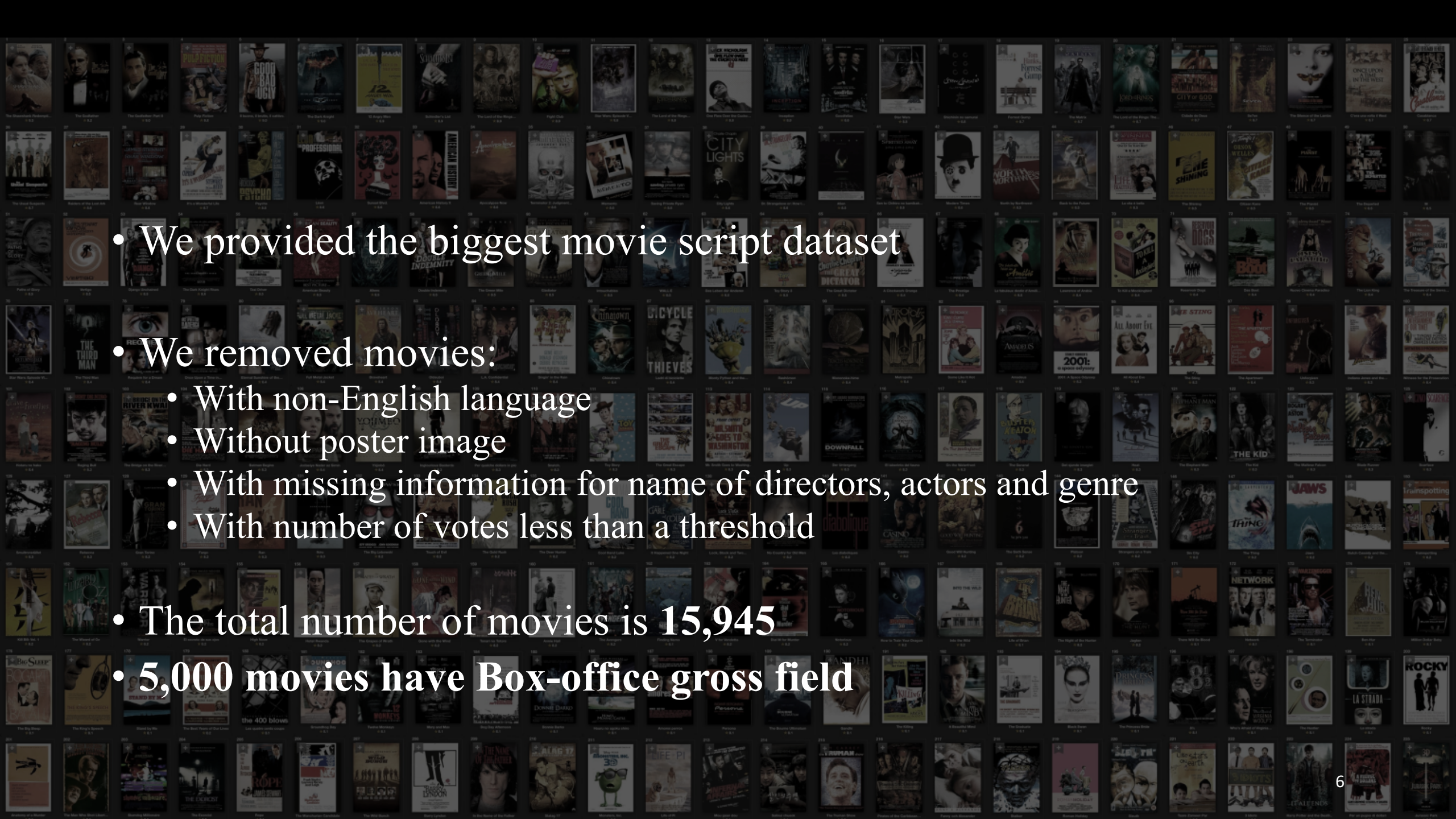
- Crawled websites

Box office revenue



Movie scripts



- 
- We provided the biggest movie script dataset
 - We removed movies:
 - With non-English language
 - Without poster image
 - With missing information for name of directors, actors and genre
 - With number of votes less than a threshold
 - The total number of movies is 15,945
 - 5,000 movies have Box-office gross field

Baseline

<https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS19/paper/viewPaper/18305>

- SVM Model
- Using handcraft features

Steps

- Preprocessing the input
- Choose your model:
 - Use a traditional machine learning model
 - Introduce new features
 - Use a Deep Neural Network
- Analyze the results

Köszönöm
Bedankt Meharbani
Blagodaram Tashakor
Blagodaria Tanemirt Matondo Tänan
Tak Cheers Manjuthe Arigato
Manana Toda Multumesc Sulpáy Welalin Dankegon Wado
Buznyg Multumesc Takk Blagodaram Zikomo Obrigada
Waybale Gracias Dakujem Saha Nuhun
Tack Paldies Salamat Shukria Dankewol Spasibu Mahalo
Soolong Obrigado Gracias Modupe Tack Akiba Cheers Talofa
Aabar Köszönöm Waybale Meharbani Dekuji Sobodi
Modupe Barkal Tenki Arigato Omol D koju
Aitäh Kiitos Barkal Maururu Efharisto Yekeniele Shukria Nuhun
Tak Saðol Manana Bedankt vinaka Hvala Shukran
Danki Skee Danke Ahsante Obrigado Thank You Dziakuju Mesi
Madlobt Akiba Grassie Toda Skee Spas Dankewol Paldies Danke
Arigato Supas Zikomo Miigwech
Kiitos Maururu Dziakuju Tanemirt Matondi Gracies Madlobt Marahaba
Menlau Gracies Nandri Tanmirt Danki Bayarilalaa Hvala Taiku
Toda Sobodi Nkosi Tashakor Tänan Grazie Talofa Murakoze
Merci Matondi Kinisou Ahsante Efharisto Sulpáy
Shukran Dakujem Gracies Miigwech Grassie Grazie Trugéré
Saðol Danke Merci Menlau Hvala
Dankegon Manjuthe Obrigada Dziekuje Takk Omol Shakkran
Gratzias Dziekuje Danki Merkzi Aabar Nkosi
Aitäh Mesi Grazie Waita Blagodaria D koju
Rahmat Matondo Merkzi Ngiyabonga
Tanmirt Wado Waita Cheers Kinisou Salamat Supas
Mahalo Misaotra Chokrane
Trugéré Marahaba Spasiba
Welalin

