

# FILM FINDER



Michael Feeley

Saving Private Ryan



Dunkirk



Tropic Thunder



# Film Finder

*Movie recommendation tool based on  
similar content.*

# THE DATA

# THE DATA

- Wikipedia movie plots from Kaggle



kaggle™

# THE DATA

→ Wikipedia movie plots from Kaggle

- ◆ ~30,000 movies
- ◆ Features:
  - Plot
  - Release Year
  - Director
  - Genre
  - Cast



kaggle™

# THE DATA

→ Wikipedia movie plots from Kaggle

- ◆ ~30,000 movies
- ◆ Features:
  - Plot
  - Release Year
  - Director
  - Genre
  - Cast



kaggle™

# THE DATA

- Wikipedia movie plots from Kaggle
  - ◆ ~30,000 movies
  - ◆ Features:
    - Plot
- Rich in information



kaggle™

# THE DATA

- Wikipedia movie plots from Kaggle
  - ◆ ~30,000 movies
  - ◆ Features:
    - Plot
- Rich in information
  - ◆ Lots of junk



kaggle™

# CLEANING THE TEXT

ewan mcgregor pierce brosnan kim cattrall olivia williams



# CLEANING THE TEXT

→ Strip down to lowercase letters



Ewan McGregor Pierce Brosnan Kim Cattrall Olivia Williams

# CLEANING THE TEXT

- Strip down to lowercase letters
- Remove low-value words



Ewan McGregor Pierce Brosnan Kim Cattrall Olivia Williams

# CLEANING THE TEXT

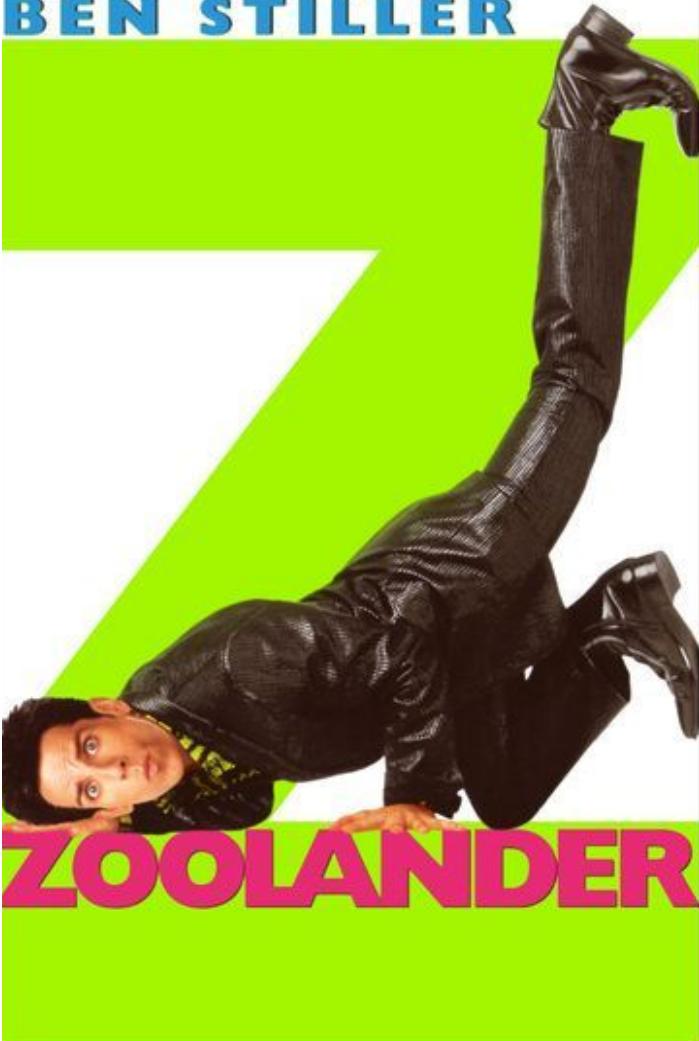
- Strip down to lowercase letters
- Remove low-value words
- Entity recognition



MODELING

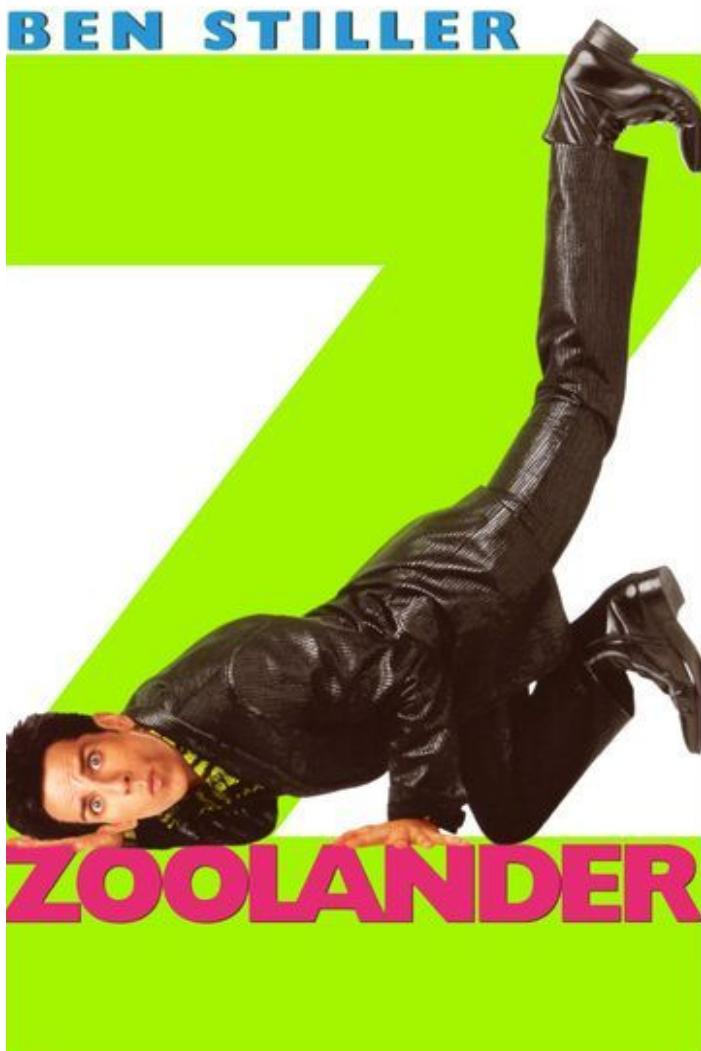
BEN STILLER

ZOOLANDER



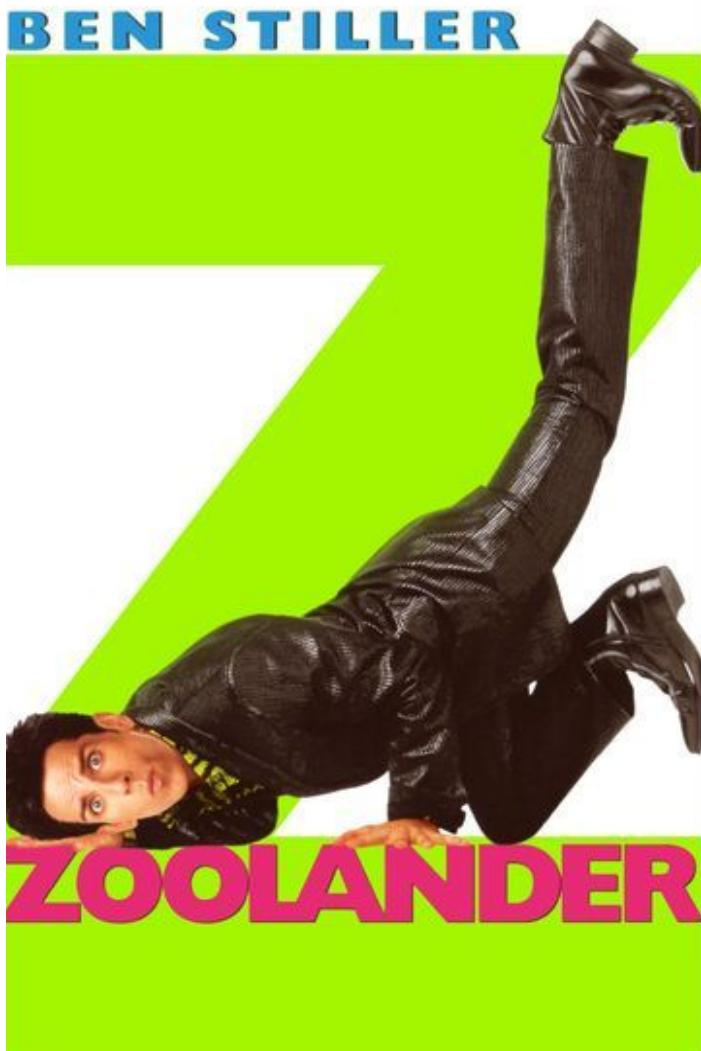
# MODELING

→ TfIdfVectorizer



# MODELING

- TfIdfVectorizer
- Shrink words into topic spaces  
(NMF)

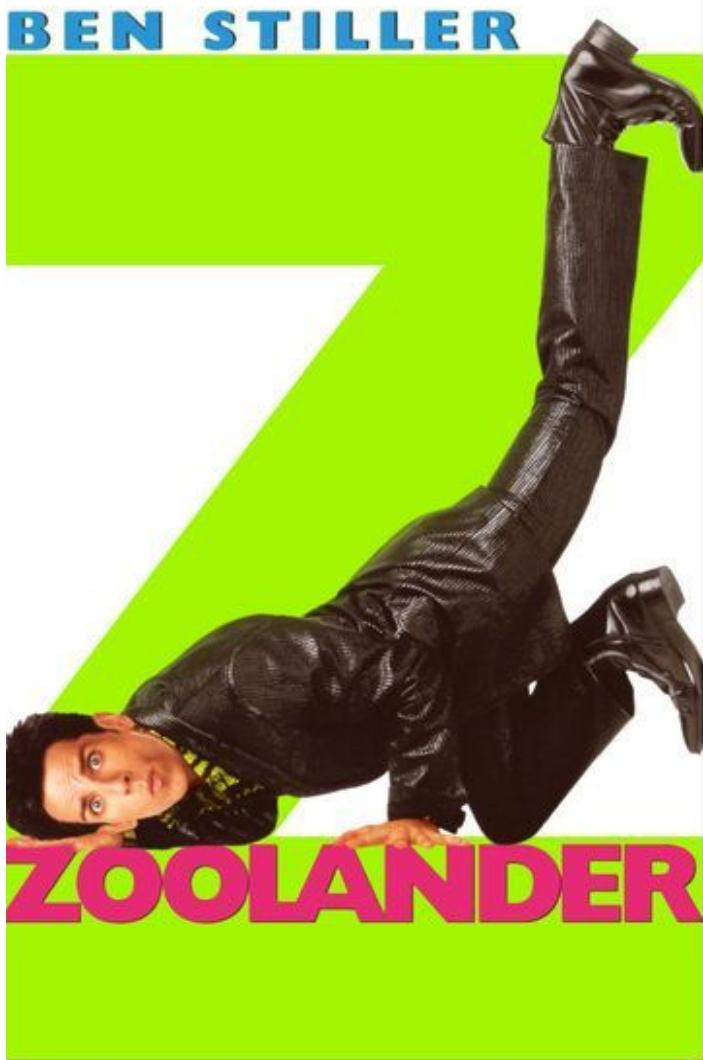


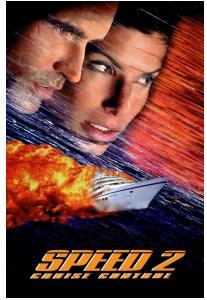
# 35 Topics

|               |              |             |                  |               |
|---------------|--------------|-------------|------------------|---------------|
| Relationships | Marriage     | War         | Police           | Village       |
| Medical       | Royalty      | Filmmaking  | School           | Money         |
| Sports        | Gangs        | Water       | Music            | Western       |
| Driving       | College      | Pregnancy   | Office Life      | Tale/Journey  |
| Writing       | Horse-Racing | Destruction | Train/Rail       | Investigation |
| Bollywood     | Combat       | Airplanes   | Criminal Justice | City Life     |
| Heist         | Death        | Animals     | Dancing          | Drug Crime    |

# MODELING

- TfIdfVectorizer
- Shrink words into topic spaces  
(NMF)
- Topic per movie
  - ◆ Each movie assigned a weight for each topic





Action



Action



Horror



Action



Horror



Thriller



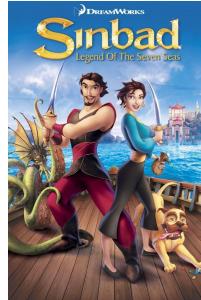
Action



Horror



Thriller



Animated



Action



Horror



Thriller

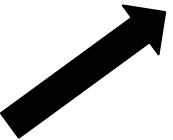


Animated



Drama

# Water



Action



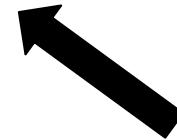
Horror



Thriller



Animated



Drama

# Filmmaking

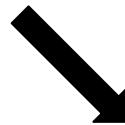
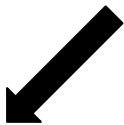
# Filmmaking



## Vellaripravinte Changathi (2011)

A look at the life of film-maker Augustine Joseph who committed suicide due to the debts incurred for an unreleased film. Later, his son acquired the prints of the film and got it released.

# Filmmaking



## Vellaripravinte Changathi (2011)

A look at the life of film-maker Augustine Joseph who committed suicide due to the debts incurred for an unreleased film. Later, his son acquired the prints of the film and got it released.

## Mere Dost Picture Abhi Baki Hai (2012)

Amar Joshi (Sunil Shetty) sells his business to become a filmmaker. As he tries to make his debut feature, he has to deal with unscrupulous financiers and producers and navigate through the complicated roads of Bollywood.

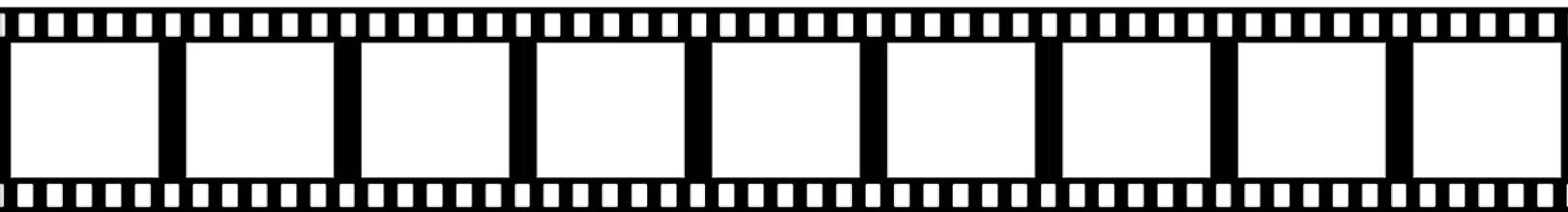
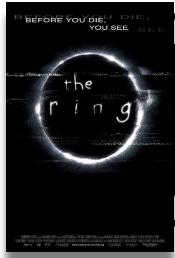
# RECOMMENDATION SYSTEM

# RECOMMENDATION SYSTEM

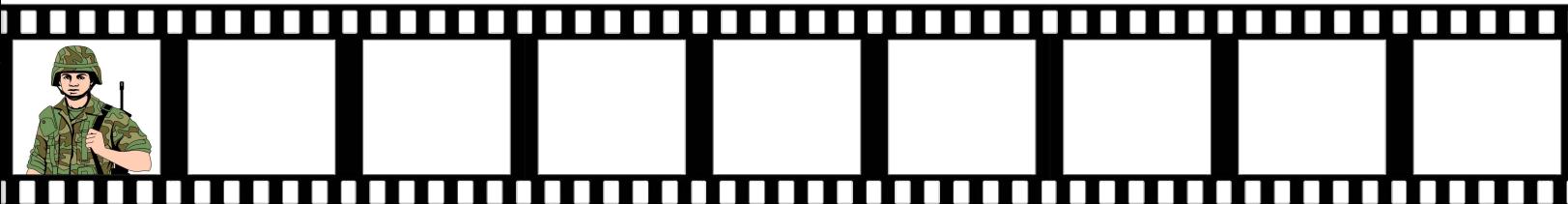
- Core → Movie Vector
  - ◆ Contains topic weights for movie

# RECOMMENDATION SYSTEM

- Core → Movie Vector
  - ◆ Contains topic weights for movie
- The Recommendation
  - ◆ Cosine distance between movie vectors



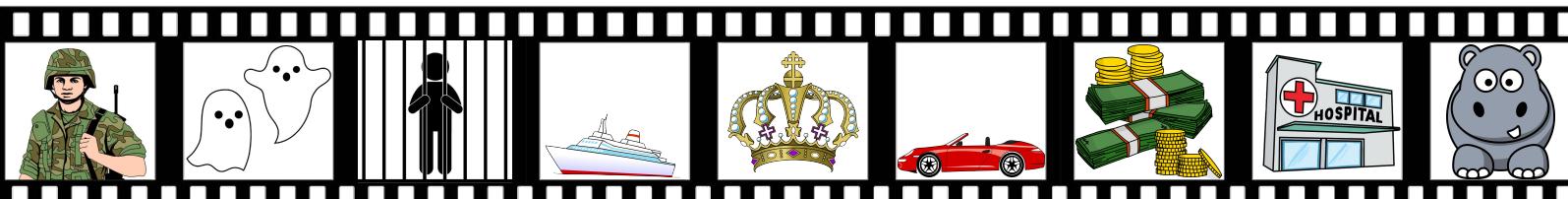
**The Ring**



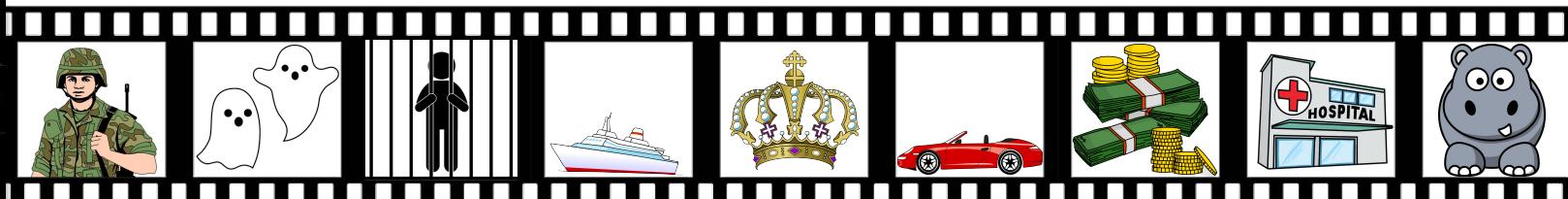
## The Ring



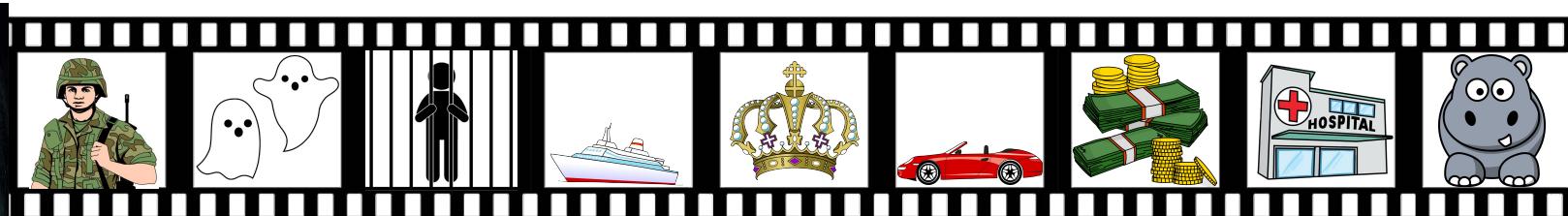
## The Ring



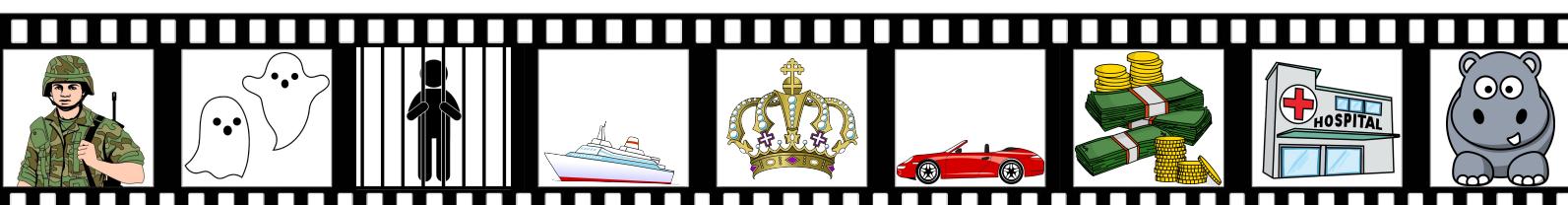
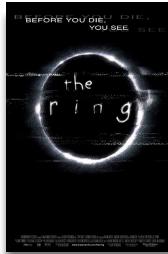
**The Ring**



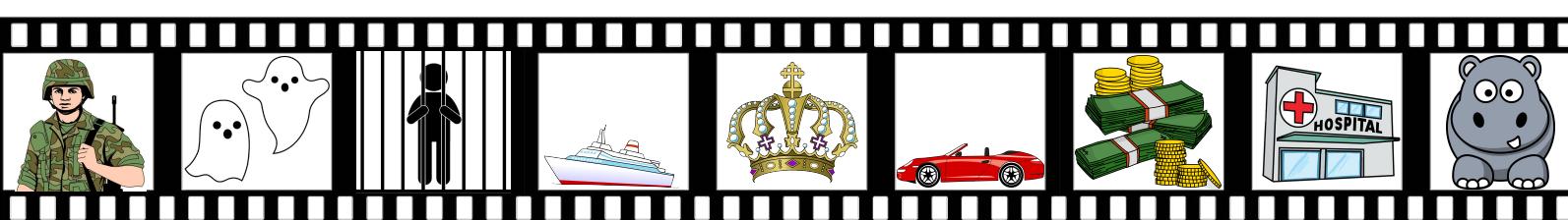
**The Ring**



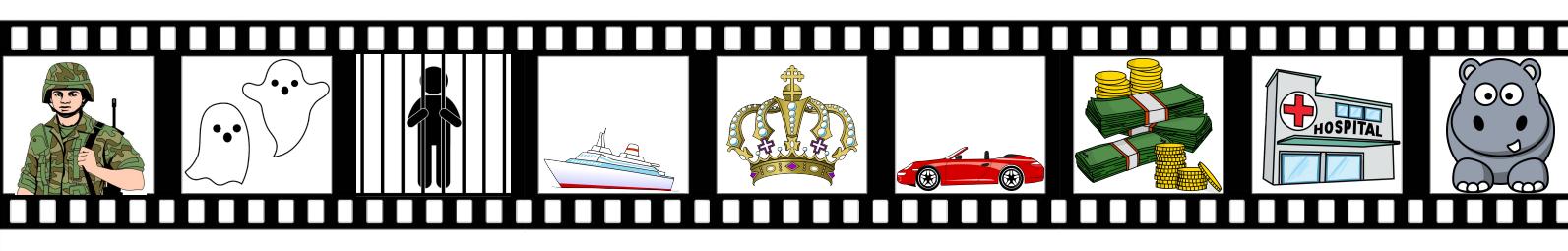
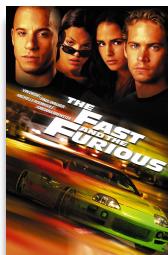
**The Grudge**



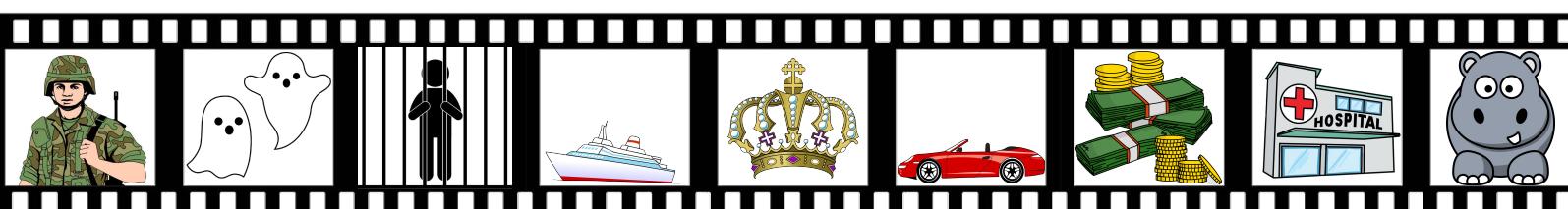
**The Ring**



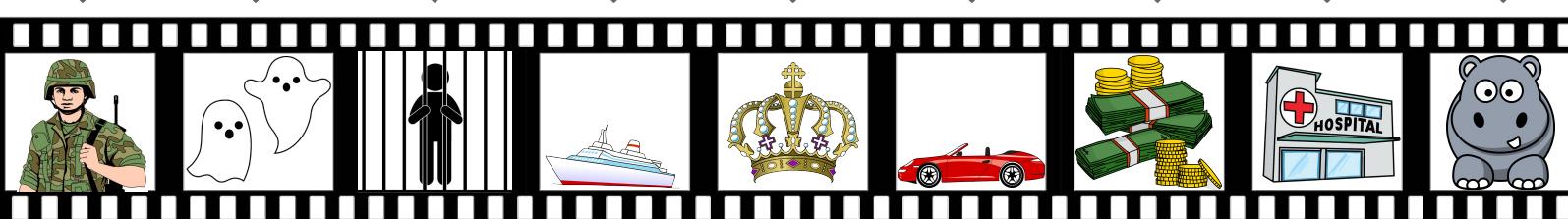
**The Grudge**



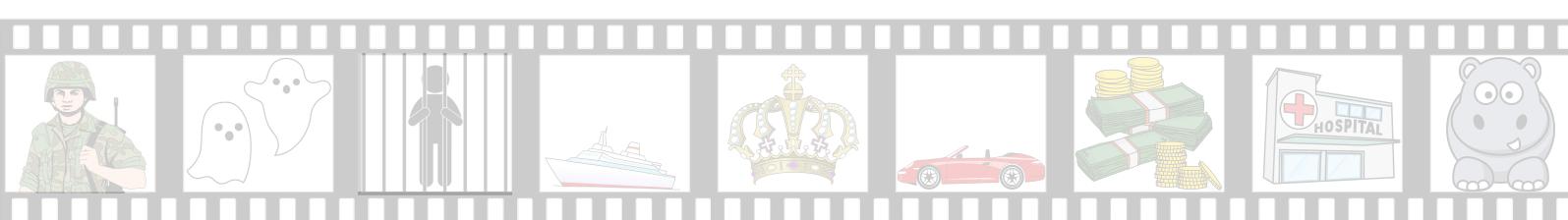
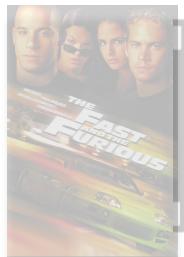
**Fast & Furious**

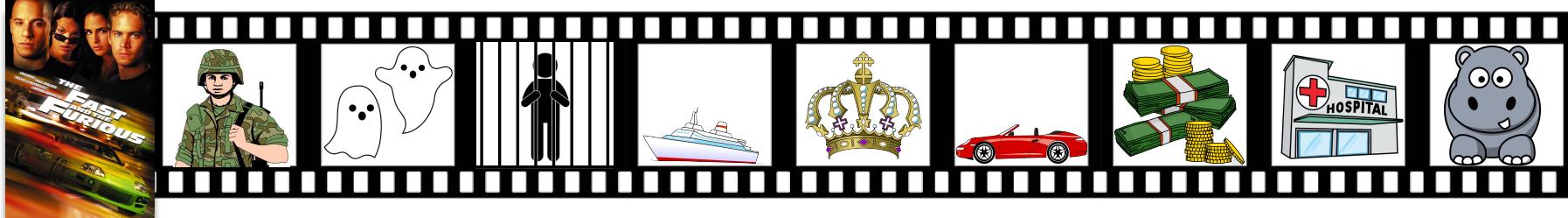
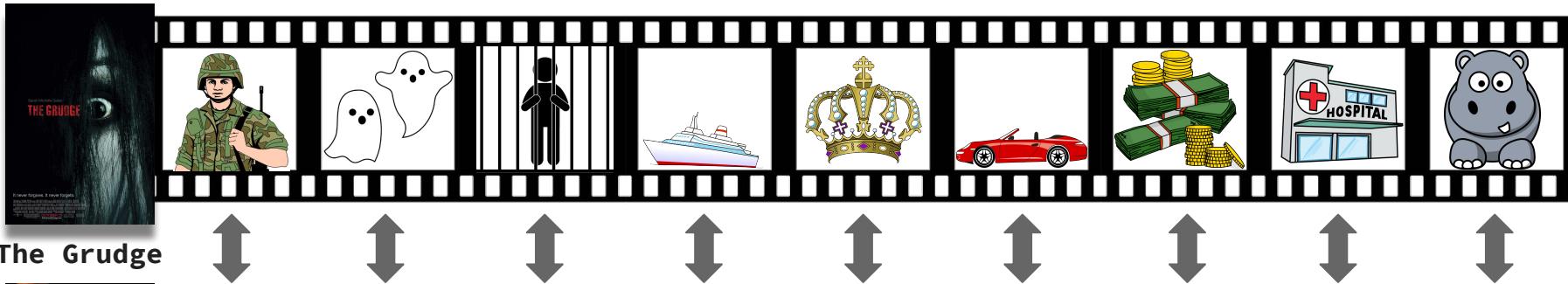
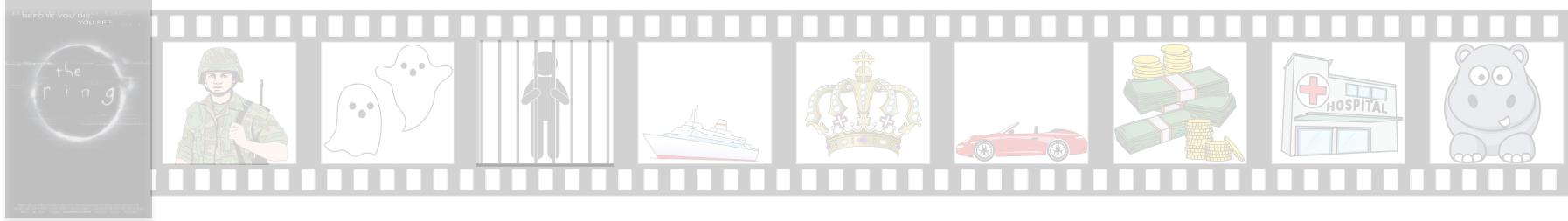


**The Ring**



**The Grudge**

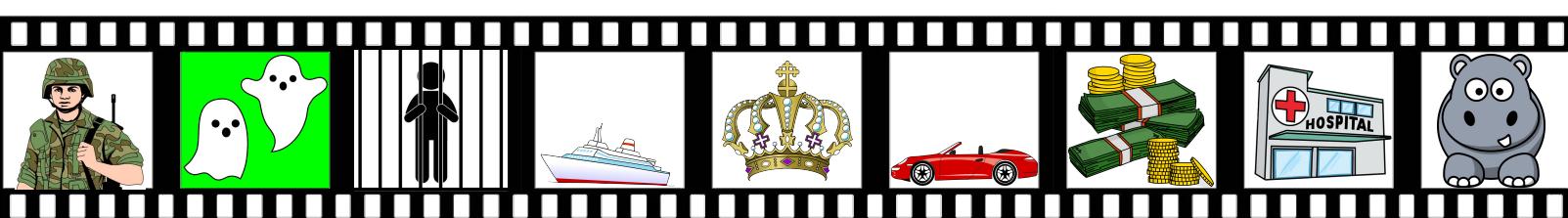




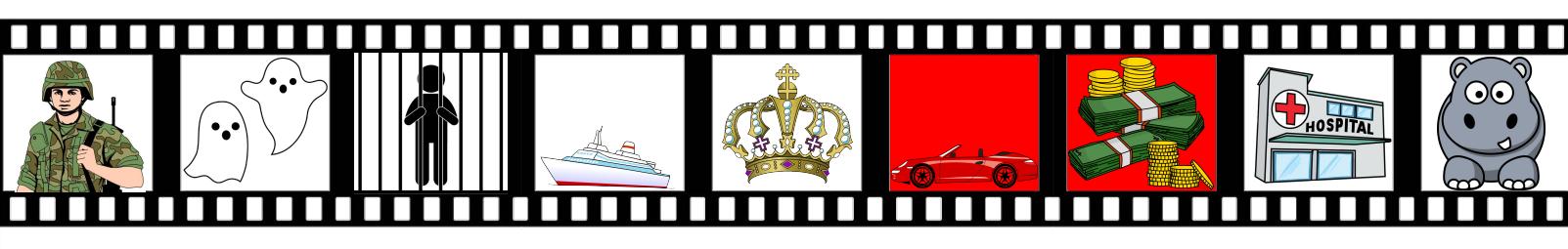
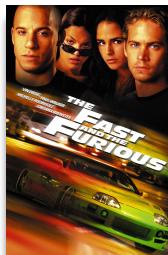
Fast & Furious



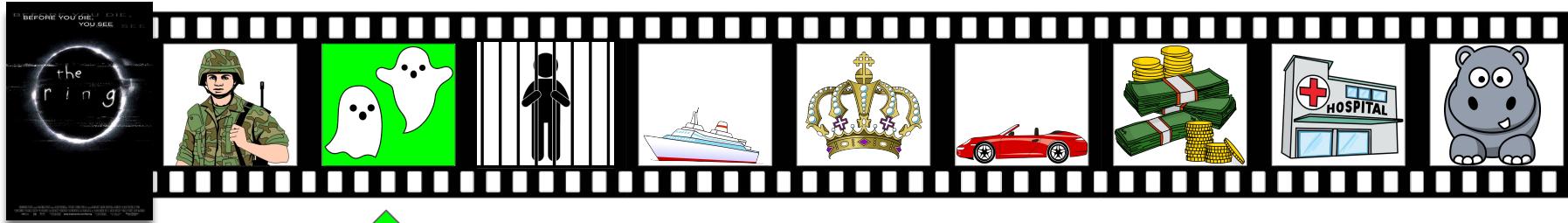
**The Ring**



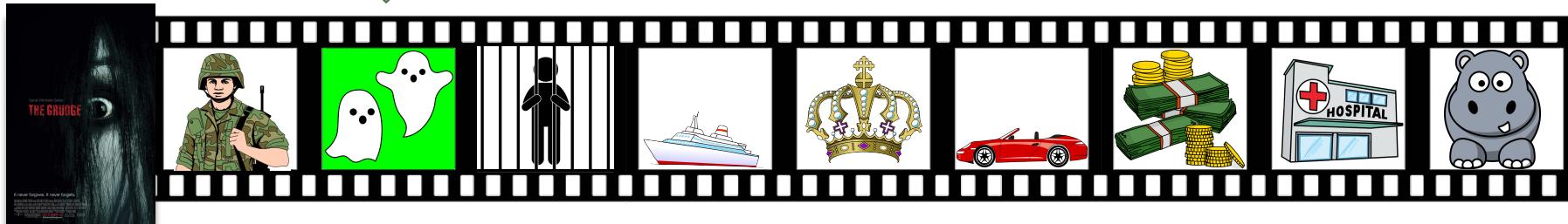
**The Grudge**



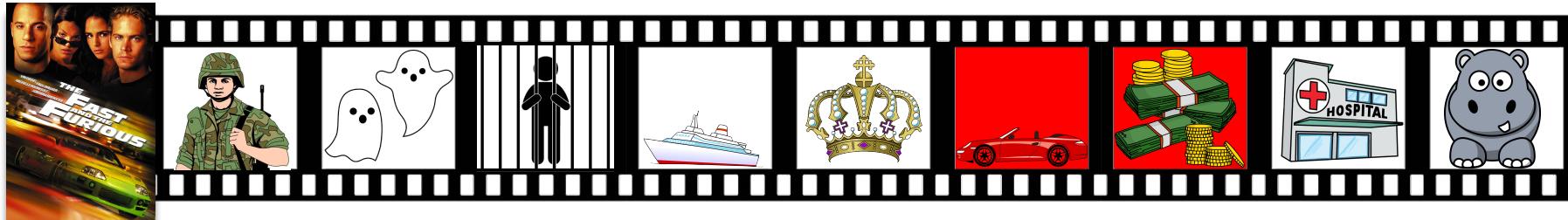
**Fast & Furious**



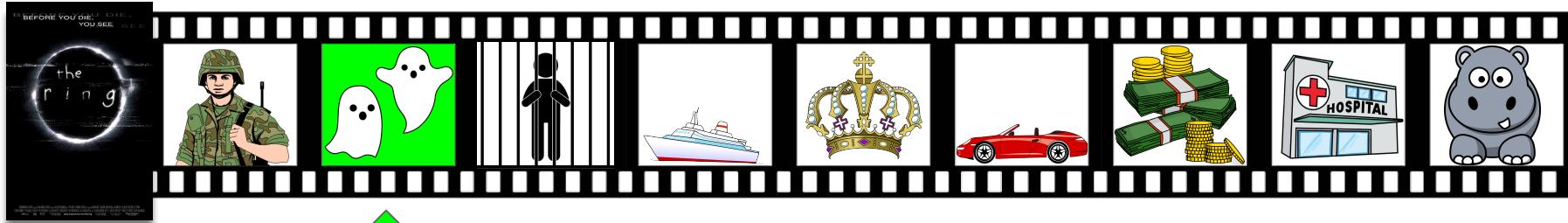
The Ring



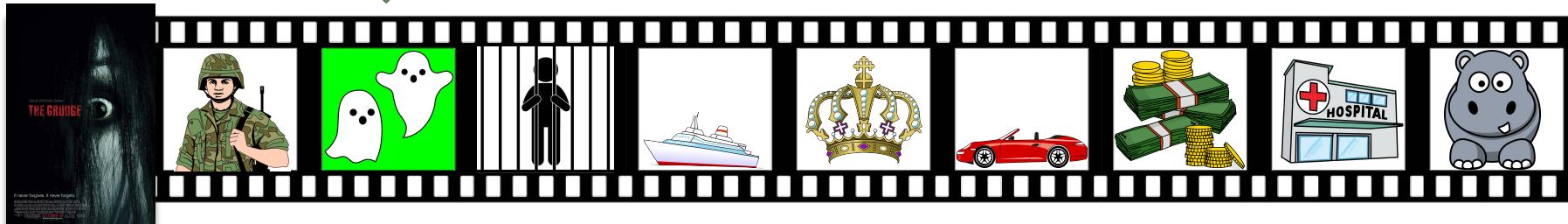
The Grudge



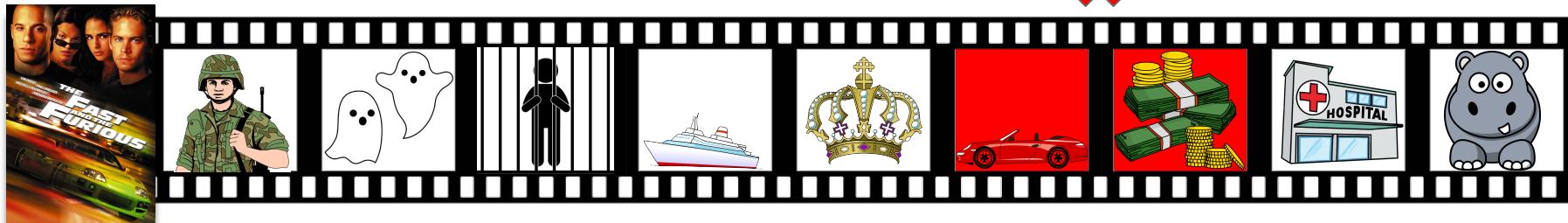
Fast & Furious



The Ring



The Grudge



Fast & Furious

# RECOMMENDATION SYSTEM

- Core → Movie Vector
  - ◆ Contains the weight of each topic for that movie
- The Recommendation
  - ◆ Cosine distance between movies
- Multiple inputs
  - ◆ Movie with best average ranking amongst movies searched

Let's try it out!

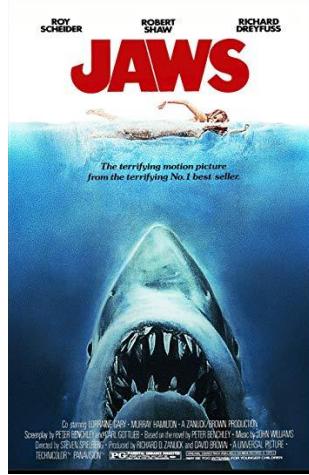
# Film Finder



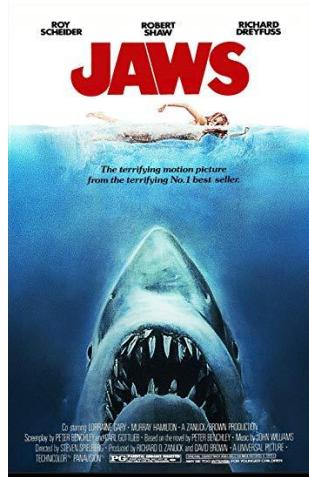
Search for one or more movies, separated by a comma and space (Get Out, The Matrix, Finding Nemo)

Movie plot data from Kaggle

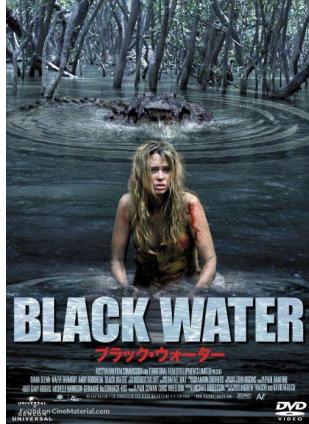
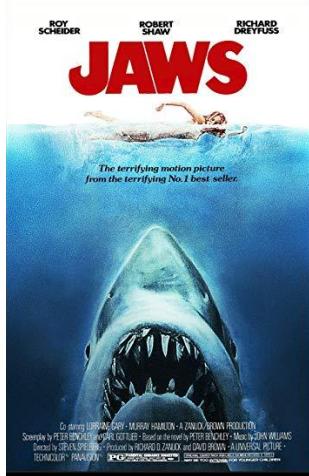
Created by Michael Feeley: [LinkedIn](#) | [GitHub](#)



When a young woman is killed by a shark while skinny-dipping near the New England tourist town of Amity Island, police chief Martin Brody (Roy Scheider) wants to close the beaches, but mayor Larry Vaughn (Murray Hamilton) overrules him, fearing that the loss of tourist revenue will cripple the town. Ichthyologist Matt Hooper (Richard Dreyfuss) and grizzled ship captain Quint (Robert Shaw) offer to help Brody capture the killer beast, and the trio engage in an epic battle of man vs. nature.

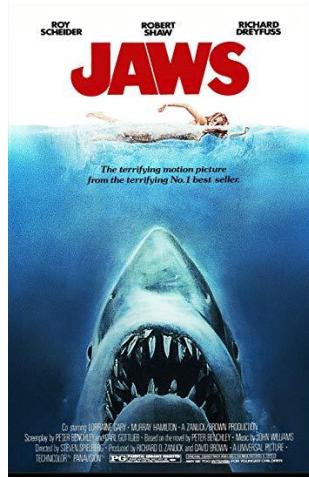


When a young woman is **killed by a shark** while skinny-dipping near the New England tourist town of Amity Island, police chief Martin Brody (Roy Scheider) wants to close the beaches, but mayor Larry Vaughn (Murray Hamilton) overrules him, fearing that the loss of tourist revenue will cripple the town. Ichthyologist Matt Hooper (Richard Dreyfuss) and **grizzled ship captain** Quint (Robert Shaw) offer to help Brody **capture the killer beast**, and the trio engage in an epic battle of **man vs. nature**.



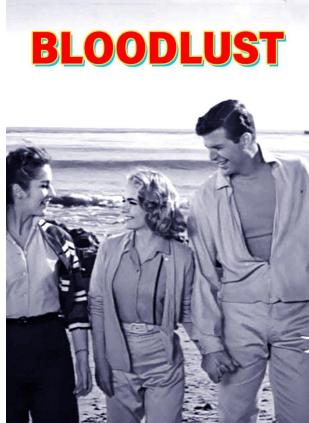
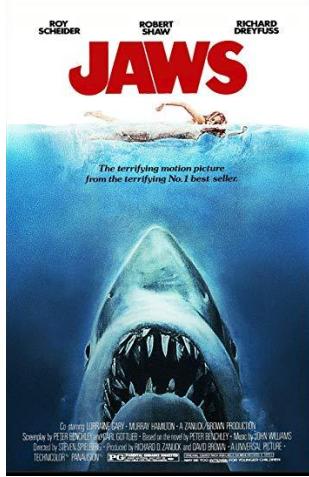
When a young woman is **killed by a shark** while skinny-dipping near the New England tourist town of Amity Island, police chief Martin Brody (Roy Scheider) wants to close the beaches, but mayor Larry Vaughn (Murray Hamilton) overrules him, fearing that the loss of tourist revenue will cripple the town. Ichthyologist Matt Hooper (Richard Dreyfuss) and **grizzled ship captain** Quint (Robert Shaw) offer to help Brody **capture the killer beast**, and the trio engage in an epic battle of **man vs. nature**.

A pregnant woman, her boyfriend and her sister take a **boat tour** of a mangrove swamp, where they are **terrorised by a killer crocodile**.



When a young woman is **killed by a shark** while skinny-dipping near the New England tourist town of Amity Island, police chief Martin Brody (Roy Scheider) wants to close the beaches, but mayor Larry Vaughn (Murray Hamilton) overrules him, fearing that the loss of tourist revenue will cripple the town. Ichthyologist Matt Hooper (Richard Dreyfuss) and **grizzled ship captain Quint** (Robert Shaw) offer to help Brody **capture the killer beast**, and the trio engage in an epic battle of **man vs. nature**.

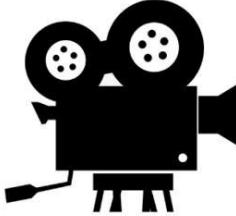
Spring break turns gory when an **underground tremor** releases hundreds of **prehistoric, carnivorous fish** into Lake Victoria, a popular waterside resort. Local cop Julie Forester (Elisabeth Shue) must join forces with a band of unlikely strangers -- though they are badly outnumbered -- to **destroy the ravenous creatures before everyone becomes fish food**.



When a young woman is **killed by a shark** while skinny-dipping near the New England tourist town of Amity Island, police chief Martin Brody (Roy Scheider) wants to close the beaches, but mayor Larry Vaughn (Murray Hamilton) overrules him, fearing that the loss of tourist revenue will cripple the town. Ichthyologist Matt Hooper (Richard Dreyfuss) and **grizzled ship captain** Quint (Robert Shaw) offer to help Brody **capture the killer beast**, and the trio engage in an epic battle of **man vs. nature**.

Two young couples **vacationing at sea** decide to visit a nearby **tropical island**. Here they encounter Dr Albert Balleau, a **demented doctor** who **kidnaps people** and then **hunts them for sport** on his property.

# Film Finder



Moneyball, s|

Movie plot data from Kaggle

Created by Michael Feeley: [LinkedIn](#) | [GitHub](#)

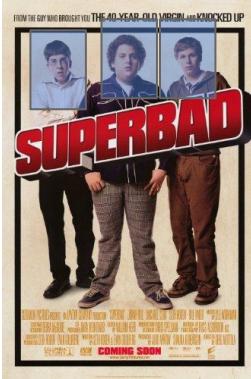


# Unfinished Business



R 2015 · Comedy · 1h 31m

Tired of playing second-fiddle to his obnoxious boss (Sienna Miller), businessman Dan Trunkman (Vince Vaughn) quits and forms his own competing mineral sales firm. He takes a retiring associate (Tom Wilkinson) and a dim-witted sales applicant (Dave Franco) with him. After a year and a half of struggle, a promising deal is in sight -- then promptly folds, thanks to Dan's ex-boss. Determined to save the deal, Dan and company fly to Germany, where a series of outrageous obstacles awaits.

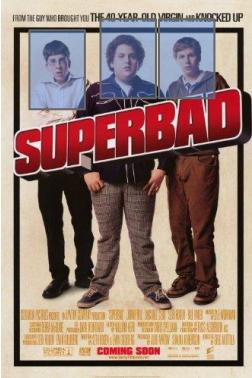


## Unfinished Business



R 2015 · Comedy · 1h 31m

Tired of playing second-fiddle to his obnoxious boss (Sienna Miller), businessman Dan Trunkman (Vince Vaughn) quits and forms his own competing mineral sales firm. He takes a retiring associate (Tom Wilkinson) and a dim-witted sales applicant (Dave Franco) with him. After a year and a half of struggle, a promising deal is in sight -- then promptly folds, thanks to Dan's ex-boss. Determined to save the deal, Dan and company fly to Germany, where a series of outrageous obstacles awaits.

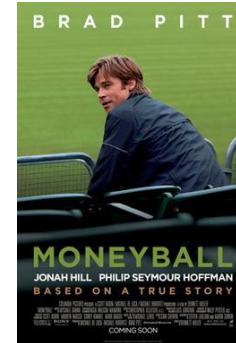


## Unfinished Business



R 2015 · Comedy · 1h 31m

Tired of playing second-fiddle to his obnoxious boss (Sienna Miller), businessman Dan Trunkman (Vince Vaughn) quits and forms his own competing mineral sales firm. He takes a retiring associate (Tom Wilkinson) and a dim-witted sales applicant (Dave Franco) with him. After a year and a half of struggle, a promising deal is in sight -- then promptly folds, thanks to Dan's ex-boss. Determined to save the deal, Dan and company fly to Germany, where a series of outrageous obstacles awaits.



# THANK YOU!



michael.f.feeley@gmail.com



mfeeley



mfeeley

# TOOLS USED



NLTK

