



PROBLEM STATEMENT

Examine trends in the entertainment industry and attempt to predict Academy Award favorites based on information such as: release date, plot summary, genre, production company, budget, opening WE results, and history of Oscars among a movie's cast & crew.

What's the big deal with the Oscars?

The Academy Awards mean much more than a celebration of talent:

- ♦ ~\$149 M in revenue for ABC the exclusive home of the Oscars since 1997
- \diamond ~\$130 M in economic boost that LA sees from hosting the Oscars
- ~\$100 M+ is spent each year on awards season lobbying
- ♦ 20% is the boost in pay that a Best Actor/Actress can expect on their next film
- \$150-225k goes to every nominee in party favors
- \Diamond .



Big deal - big datasets

LOTS of data engineering

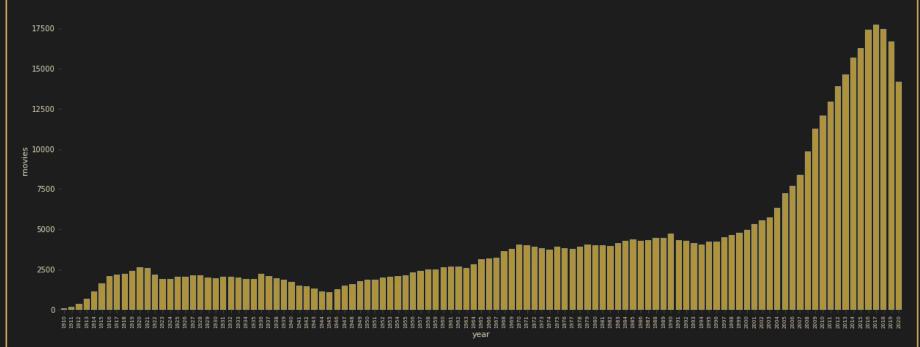
Oscar details scraped for movies from 1990-2020

Details scraped for movies from 1990-2020: budget, opening WE, release date, language, etc.

Base data downloads: <u>basic</u> title, name, principals, and ratings data (1900-2020), in tables linked by unique keys

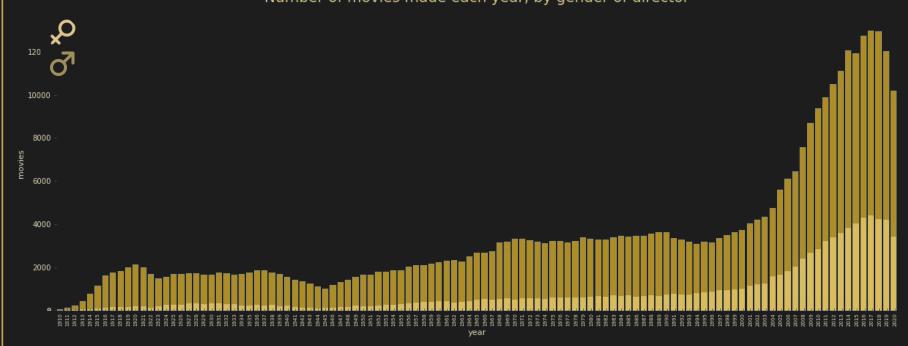
The movies, in numbers

Number of movies produced worldwide, by year

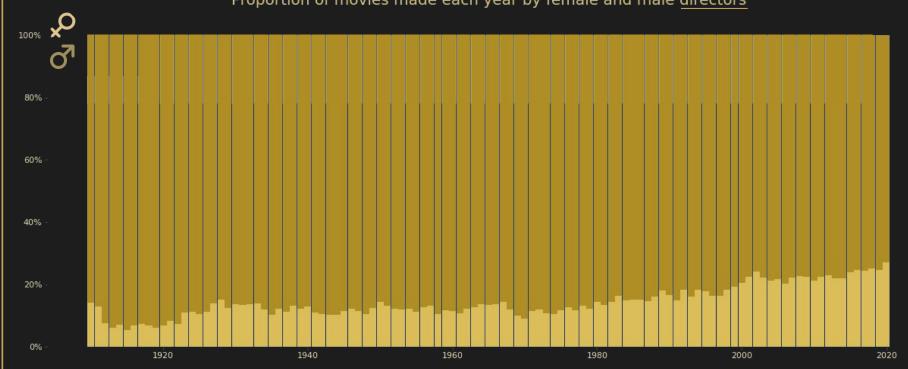


The movies, in numbers

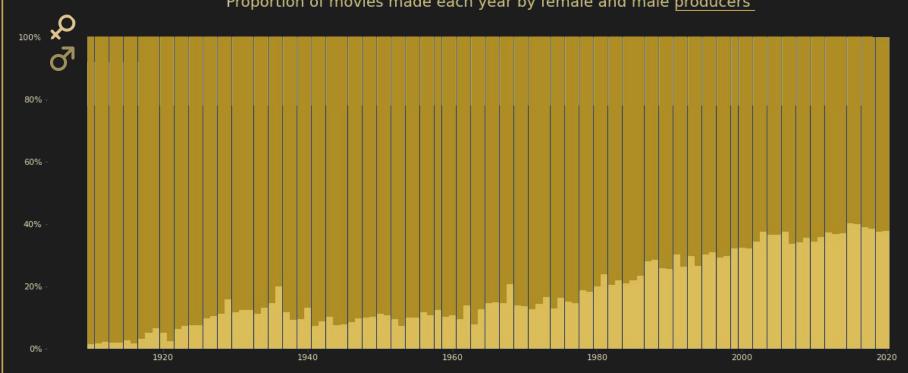
Number of movies made each year, by gender of director



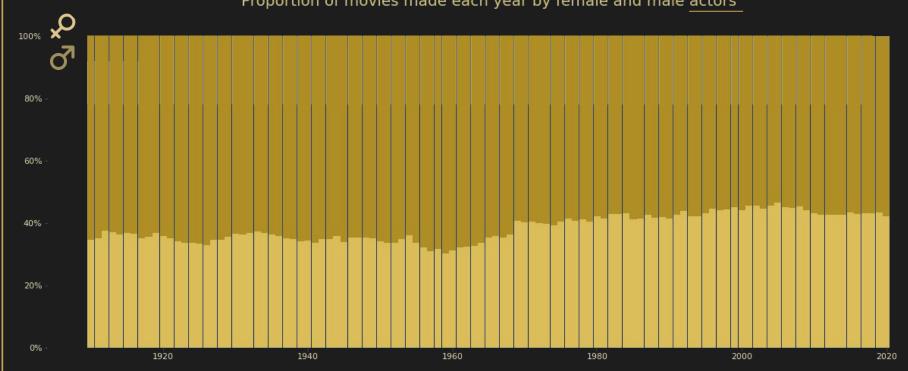
Proportion of movies made each year by female and male directors



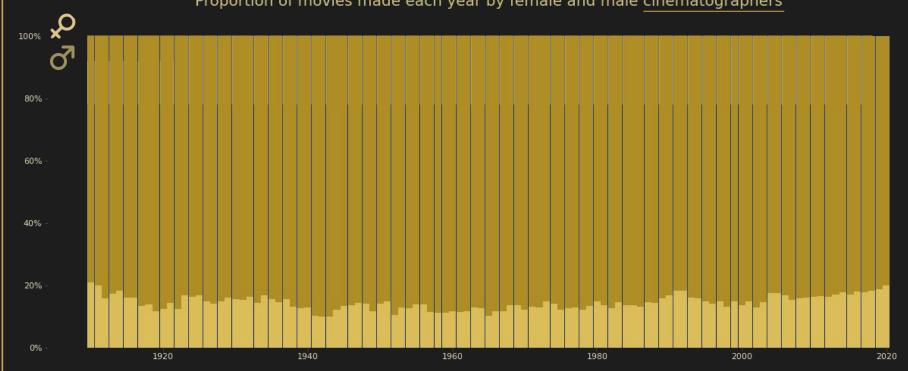
Proportion of movies made each year by female and male producers



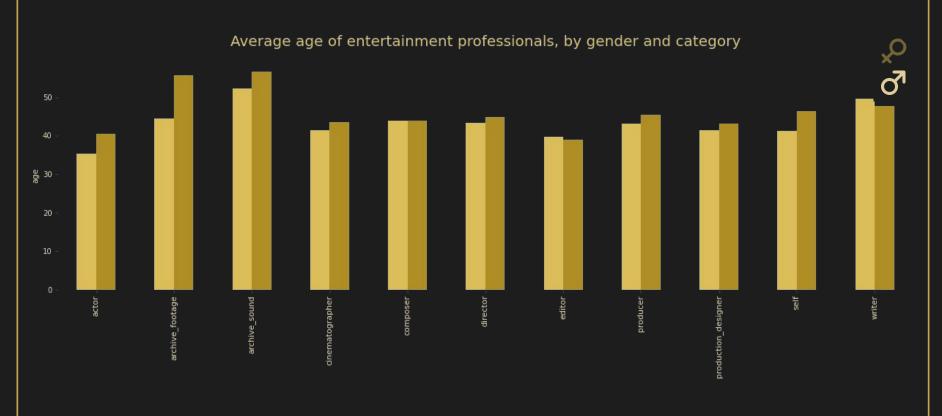
Proportion of movies made each year by female and male actors



Proportion of movies made each year by female and male cinematographers



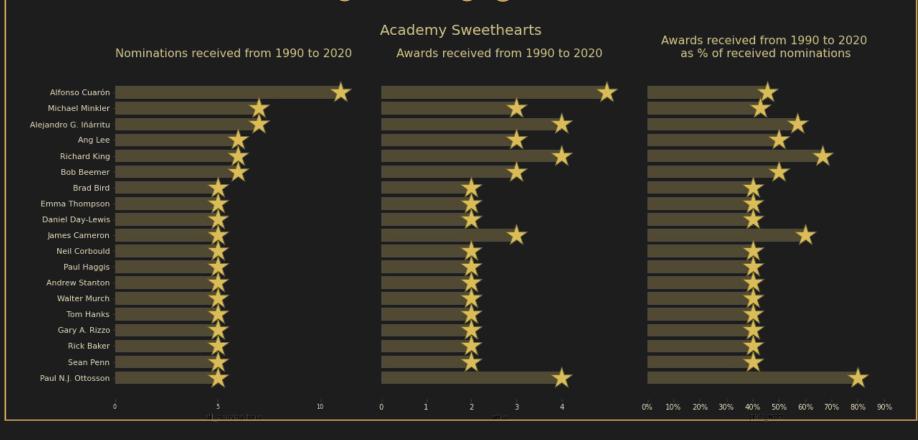
Age differences



Back to the Oscars



The Academy is very generous to some...

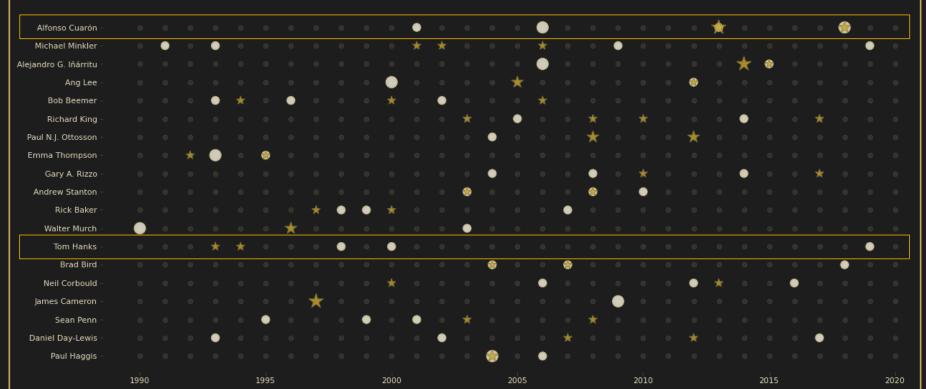


Nomination/-s

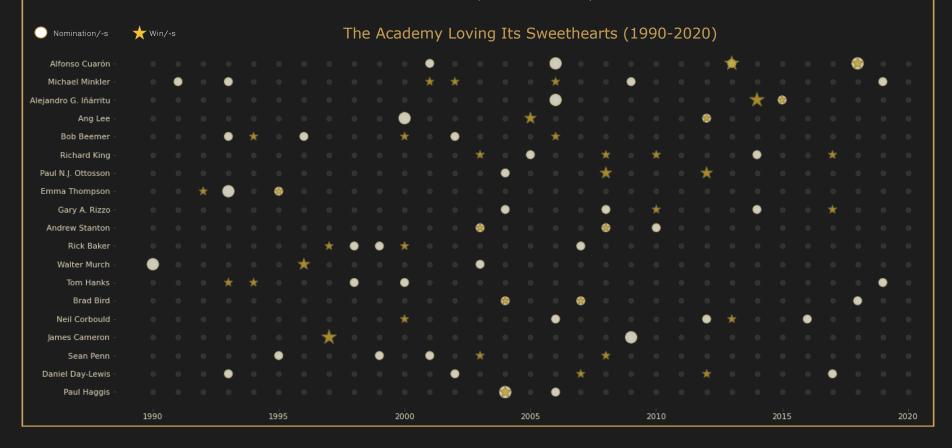
★ Win/-s

ACADEMY LOVE

Academy loving its sweethearts (1910-2020)



ACADEMY LOVE



The Academy is <u>not</u> generous to all

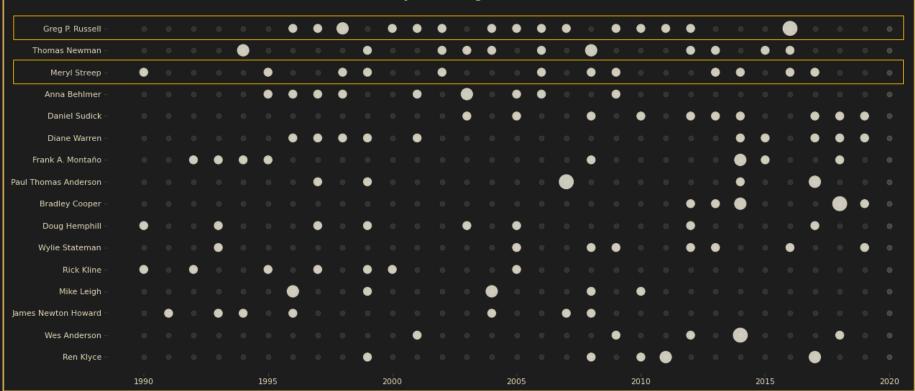


Nomination/-s

★ Win/-s

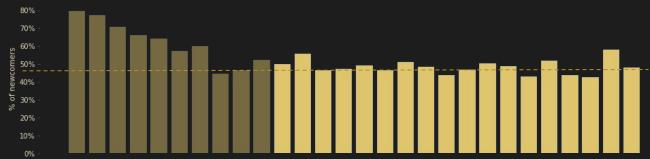
ACADEMY LOVE... of snubbing

Academy snubbing its sweetsnubs (1910-2020)

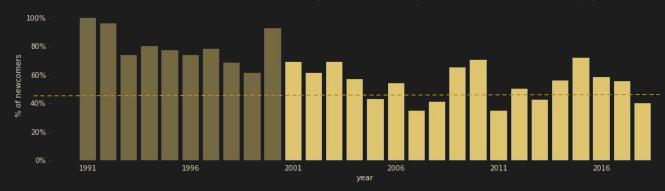


Can newcomers break in?

Number # of Oscar newcomers nominated by the Academy as % of all nominees, by year (1990-2020)



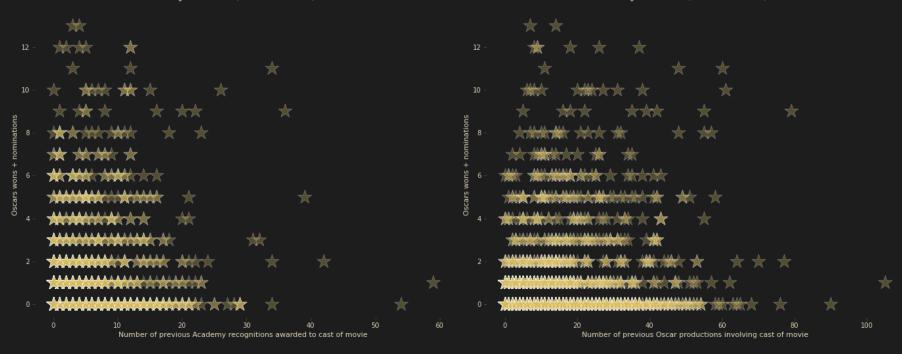
Number # of Oscar newcomers awarded by the Academy as % of all winners, by year (1990-2020)



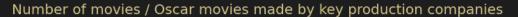
Once a winner, always a winner?

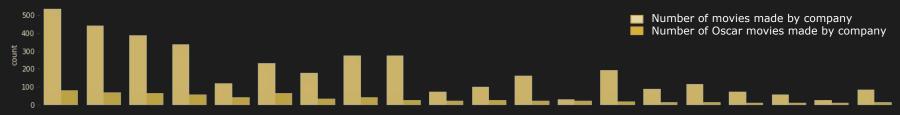
Oscar wins & nominations vs Oscar-density of cast and crew, by movie (1995-2020)

Oscar wins & nominations vs Oscar-density of cast and crew, by movie (1995-2020)

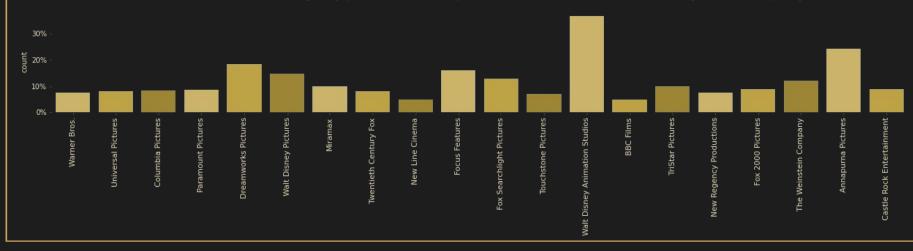


Oscar luck by studio



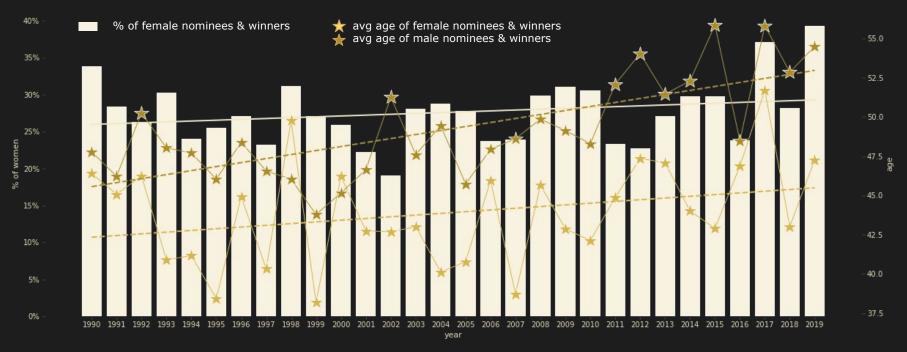


Oscar movies made by key production companies as % of all movies made by each company



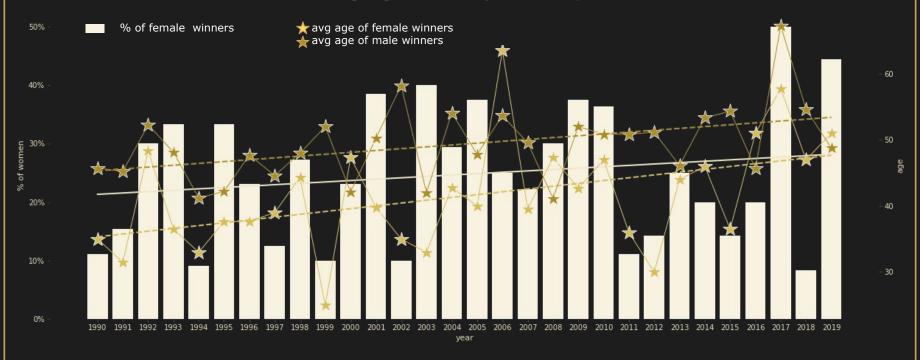
Academy's support of diversity

Average age of Academy Award nominees and recipients



Academy's support of diversity

Average age of Academy Award recipients





The magnitude and complexity of the task

175,693

movies made between 1990 and 2019 1,197

movies recognized by the Academy during this period

0.7%

237

movies which received Academy Awards

 \bigcirc .]%

Four questions - four predictors

Which movies are most likely to receive a statuette?

AWARD

Which movies have the best chances of winning THE award?

BEST PIC



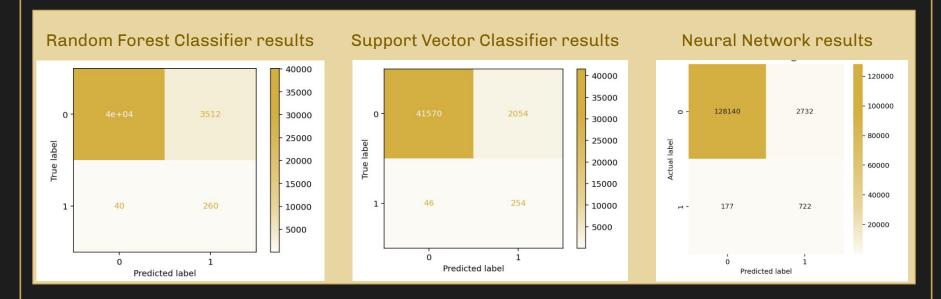
NOMINATION

Which movies will receive an Oscar nomination, in any category?

BEST PIC NOMINATION

Which movies will the Academy nominate for Best Motion Picture?

Predicting movies to be <u>nominated</u> for an Oscar



- ★ ACCURACY: 0.92-0.97 on train data, 0.92-0.95 on test data, 0.92-0.99 in 5-fold cross-validation
- * RECALL: 0.87-1.0 on train data, 0.79-0.87 on test data. 0.63-0.83 in 5-fold cross-validation
- ★ PRECISION: 0.19-0.93 on train data, 0.07-0.18 on test data, 0.07-0.53 in 5-fold cross-validation

Models' top picks for Oscar <u>nominations</u>*

TITLE	GENRES	RUNTIME
Sound of Metal	Drama, Music	120
Tenet	Action, Sci-Fi, Thriller	150
Onward	Adventure, Animation, Comedy	102
Hillbilly Elegy	Drama	116
Never Rarely Sometimes Always	Drama	101
The Croods: A New Age	Adventure, Animation, Comedy	95
Wolfwalkers	Adventure, Animation, Family	103
News of the World	Action, Adventure, Drama	118
Emma.	Comedy, Drama, Romance	124
The Call of the Wild	Adventure, Drama, Family	100

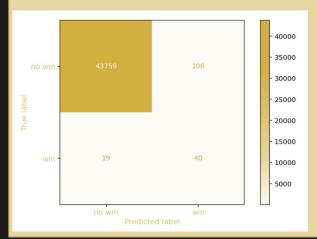
TITLE	GENRES	RUNTIME
I Care a Lot	Comedy, Crime, Thriller	118
Soul	Adventure, Animation, Comedy	100
The Way I See It	Documentary	100
Over the Moon	Adventure, Animation, Comedy	95
The Eight Hundred	Action, Drama, History	149
The Trial of the Chicago 7	Drama, History, Thriller	129
Pieces of a Woman	Drama	126
Trolls World Tour	Adventure, Animation, Comedy	90
Worth	Biography, Drama, History	118
The Last Thing He Wanted	Crime, Drama, Thriller	115

Color indicates movies with nominations as of March 15th 2021

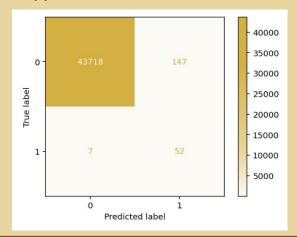
^{**} based on hard voting with an ensemble classifier using estimators listed on previous page

Predicting movies to <u>receive</u> an Oscar

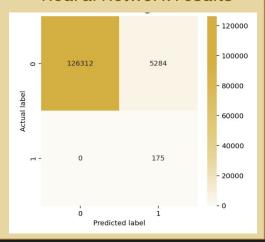
Random Forest Classifier results



Support Vector Classifier results



Neural Network results



- ★ ACCURACY: 0.90-0.96 on train data, 0.96-0.99 on test data, 0.98-1.0 in 5-fold cross-validation
- * RECALL: 1.0 on train data, 0.78-1.0 on test data, 0.52-0.88 in 5-fold cross-validation
- ★ PRECISION: 0.03-0.53 on train data, 0.03-0.26 on test data, 0.06-0.30 in 5-fold cross-validation

Top picks for Oscar wins*

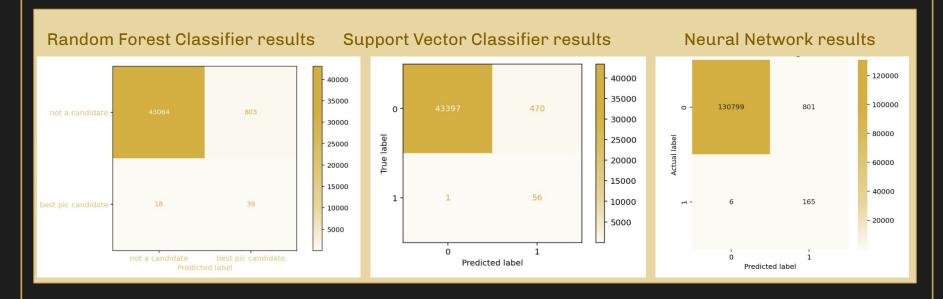
TITLE	GENRES	RUNTIME
Hillbilly Elegy	Drama	116
Tenet	Action, Sci-Fi, Thriller	150
The Invisible Man	Drama, Horror, Mystery	124
Sound of Metal	Drama, Music	120
Onward	Adventure, Animation, Comedy	102
Mank	Biography, Comedy, Drama	131
l Care a Lot	Comedy, Crime, Thriller	118
Sonic the Hedgehog	Action, Adventure, Comedy	99
Birds of Prey: And the Fantabulous Emancipation of One Harley Quinn	Action, Adventure, Comedy	109

TITLE	GENRES	RUNTIME
Dolittle	Adventure, Comedy, Family	101
The Call of the Wild	Adventure, Drama, Family	100
Shikara	Drama, Romance	118
The Rhythm Section	Action, Drama, Mystery	109
Pinocchio	Drama, Fantasy	125
News of the World	Action, Adventure, Drama	118
Bad Boys for Life	Action, Comedy, Crime	124
Bloodshot	Action, Drama, Sci-Fi	109
Last Call	Biography, Drama	110
Underwater	Action, Horror, Sci-Fi	95
Timmy Failure: Mistakes Were Made	Adventure, Comedy ,Drama	99

Color indicates movies with nominations as of March 15th 2021

^{&#}x27;* based on hard voting with an ensemble classifier using estimators listed on previous page

Predicting movies to be <u>nominated</u> for <u>Best Picture</u>



- ★ ACCURACY: 0.99-1.0 on train data, 0.99 on test data, 0.98 in 5-fold cross-validation
- * RECALL: 0.98-1.0 on train data, 0.72-0.95 on test data, 0.51-0.93 in 5-fold cross-validation
- ★ PRECISION: 0.14-1.0 on train data, 0.14-0.30 on test data, 0.07-0.36 in 5-fold cross-validation

Top picks for <u>Best Picture nominations</u>*

TITLE	GENRES	RUNTIME
Hillbilly Elegy	Drama	116
Tenet	Action, Sci-Fi,Thriller	150
News of the World	Action, Adventure, Drama	118
The Call of the Wild	Adventure, Drama, Family	100
The Comeback Trail	Comedy, Crime	104
Borat Subsequent Moviefilm	Comedy	95
The Croods: A New Age	Adventure, Animation, Comedy	95
Let Them All Talk	Comedy, Drama	113
I Care a Lot	Comedy, Crime, Thriller	118
Dolittle	Adventure, Comedy, Family	101

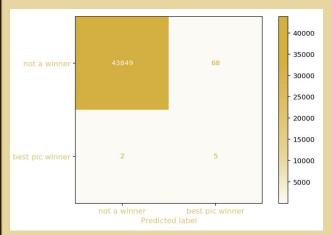
TITLE	GENRES	RUNTIME
Wendy	Drama, Fantasy	111
The Witches	Adventure, Comedy, Family	106
Sound of Metal	Drama,Music	120
Bad Boys for Life	Action, Comedy, Crime	124
John Williams: Live in Vienna	Music	128
Mank	Biography, Comedy, Drama	131
The Father	Drama	97
Minari	Drama	115
The Invisible Man	Drama, Horror, Mystery	124
Soul	Adventure, Animation, Comedy	100

Color indicates movies with Best Picture nominations as of March 15th 2021

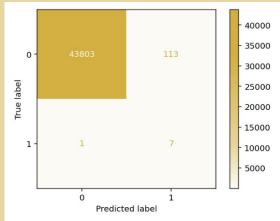
^{&#}x27;* based on hard voting with an ensemble classifier using estimators listed on previous page

Predicting movies to win Best Picture

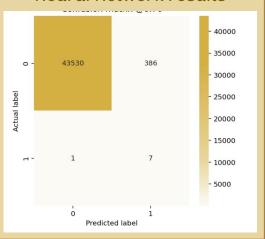
Random Forest Classifier results



Support Vector Classifier results



Neural Network results



- ★ ACCURACY: 0.57-0.99 on train data, 0.99-1.0 on test data. 1.0 in 5-fold cross-validation
- * RECALL: 1.0 on train data, 0-0.88 on test data. 0.0 in 5-fold cross-validation
- ★ PRECISION: 1.0 on train data 0.0-0.06 on test data. 0.0 in 5-fold cross-validation

Top picks for <u>Best Picture wins*</u>

TITLE	GENRES	RUNTIME
Hillbilly Elegy	Drama	116
I Care a Lot	Comedy, Crime, Thriller	118
News of the World	Action, Adventure, Drama	118
Borat Subsequent Moviefilm	Comedy	95
Wolfwalkers	Adventure, Animation, Family	103
The Croods: A New Age	Adventure, Animation, Comedy	95
I Am Better Because of You	Drama, Romance	150
Tenet	Action, Sci-Fi, Thriller	150
The Comeback Trail	Comedy, Crime	104
Hillbilly Elegy	Drama	116

TITLE	GENRES	RUNTIME
The Call of the Wild	Adventure, Drama, Family	100
The Invisible Man	Drama, Horror, Mystery	124
Pieces of a Woman	Drama	126
Greyhound	Action, Drama, History	91
Let Them All Talk	Comedy, Drama	113
The Prom	Comedy, Drama, Musical	130
Dolittle	Adventure, Comedy, Family	101
The Glorias	Biography, Drama, History	139
Sound of Metal	Drama, Music	120
The Father	Drama	97

Color indicates movies with Best Picture nominations as of March 15th 2021

^{**} based on hard voting with an ensemble classifier using estimators listed on previous page

— THANKS!—

What questions can I answer?



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution