

Feature Film Features: Applying machine learning to movie genre identification



University of Colorado
Boulder

CSCI 5622 Group L: Grant Baker, John Dinkel, Derek Gorthy, Jeffrey Maierhofer, Luke Meszar

Project Overview

Goal

Apply machine learning techniques to the problem of identifying movie genres based off of promotional posters.

Justification

Machine learning has had success with image recognition and classification. An effective method of determining poster genre trends could lead to more sophisticated advertisement and marketing campaigns, allowing advertisers to better reach their target audiences.

Dataset

The dataset was obtained using IMDb and the IMDbPY Python package. We obtained a total of 37,928 movie titles, along with links to a thumbnail of the poster, IMDb rating, director, and top five actors.

Approach

We developed a total of four models:

1. Logistic Regression on a Histogram of Colors
2. Convolutional Neural Network
3. Convolutional Neural Network with Inception-v3 pre-trained feature extraction layers utilizing transfer learning
4. Convolutional Neural Network with Inception-v3 feature extraction and additional movie metadata (Actors and Director)

Histogram of Colors

As a baseline, we converted each image into a histogram of its RGB pixel values and fed this to a Logistic Regression algorithm. This yielded a testing accuracy of 24%.

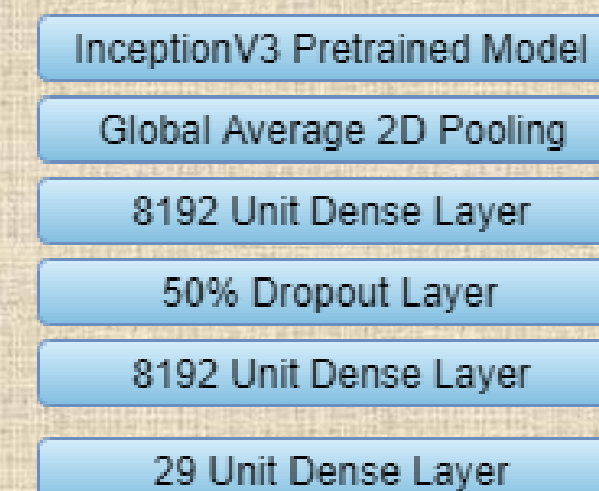
CNN

Our CNN model is described in the diagram. We achieved an f-score of 0.27.



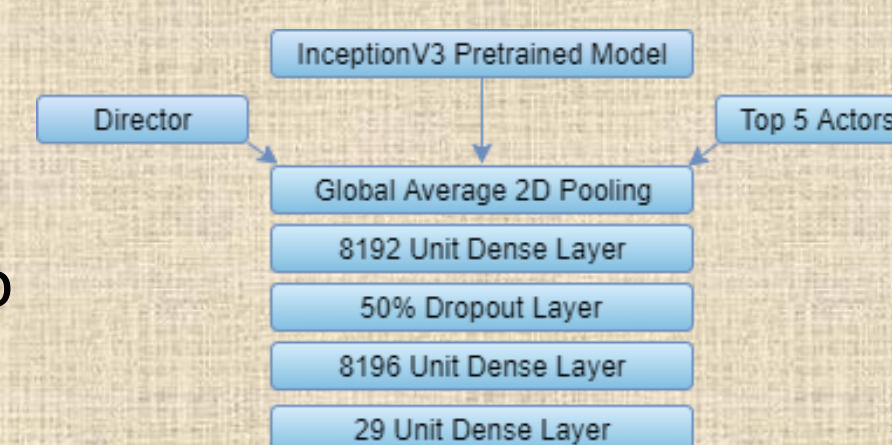
Inception-v3

Our pretrained CNN uses transfer learning with Google's Inception-v3 model as a feature extractor, training on the final layers. We achieved an f-score of 0.40.



Inception-v3 with Meta-Data

For our meta-data approach, we concatenated director and actor information to the Inception-v3 model. We have yet to finish training this model.



Human Baseline

In order to get a human baseline, we gave poster images to friends and asked them to guess three genres from the same list as our models. We averaged an f-score of 0.32. It is a small sample size of 12 people, but it is fairly clear that humans are not very good at this problem.

Conclusion

The best model is the Inception-v3 pre-trained model. This is unsurprising, as it utilizes transfer learning and a powerful feature extractor. We are hopeful that our meta-data model can perform even better.

We recognize that although our best model did not achieve a perfect f-score, due to the nature of the metric, f-scores of partially correct predictions are moderately low, so our model is able to partially solve the problem. Our current conclusion is that neural networks do not perfectly solve this problem.

Challenges

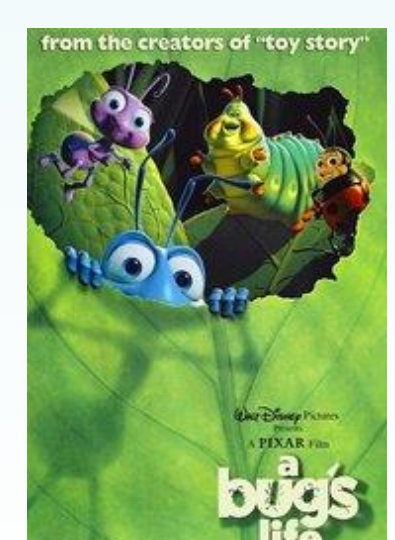
- Sparse classification → Sparsemax loss function
- Sparse classification → F-score accuracy metric
- A lot of movies are Dramas
- Training data is mostly unknown, atypical movies

Future Work

- A larger dataset containing higher resolution images might help with feature identification.
- Additional Meta-Data might also be useful. Examples include Title, Producers, Budget, etc.
- Other models architectures could lead to improved performance.
- Weighing classic movies of a genre higher could increase performance.

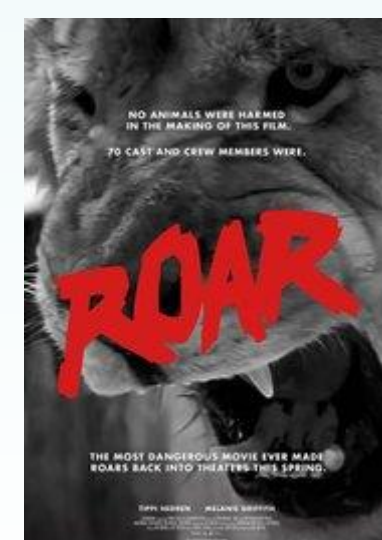
References

- Martins, Andre, and Ramon Astudillo. "From softmax to sparsemax: A sparse model of attention and multi-label classification." *International Conference on Machine Learning*. 2016.
- Ivasic-Kos, Marina, Miran Pobar, and Ivo Ipsic. "Automatic Movie Posters Classification into Genres." *ICT Innovations 2014*. Springer, Cham, 2015. 319-328.



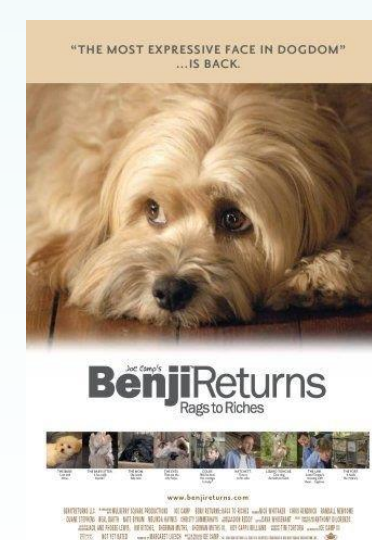
Guess
Drama
Animation
Comedy

Actual
Adventure
Animation
Comedy



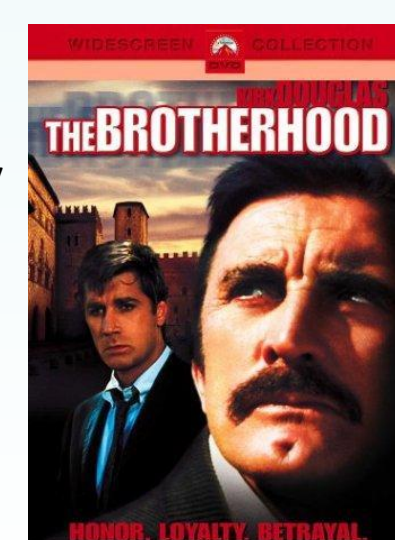
Guess
Drama
Comedy
Sport

Actual
Adventure
Horror
Thriller



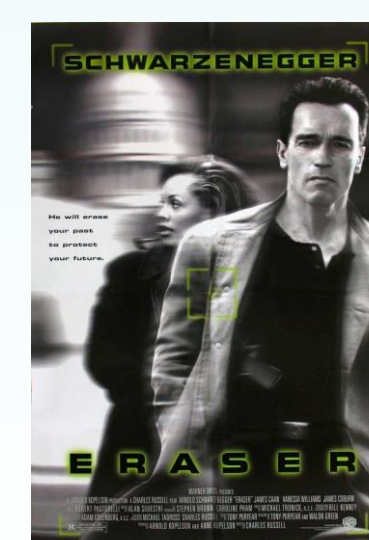
Guess
Drama
Documentary
Sport

Actual
Drama
Family



Guess
Drama
Comedy
Romance

Actual
Crime
Drama



Guess
Drama
Mystery
Thriller

Actual
Action
Drama
Mystery



Guess
Drama
Comedy
Horror

Actual
Action
Adventure
Fantasy