Using Neural Networks to Predict Violent Crime Based on Movie Releases

Skylar Furey, Jimmy Kruse, Sivani Pillutla, Matthew Afsahi

Introduction

There has long been speculation regarding the effects of media viewership on human behavior. Since the introduction of visual art, researchers and the general public have been questioning the influence it has on society. In particular, great concern has been raised over whether viewing simulated violence on screen could instigate violent behavior among its viewers. Multiple studies have been done on the long term effects of violent video games on aggression in players of all ages (1) (2), to very mixed results. Studies also show there is sometimes a decrease in nearby violent crime immediately following a showing of a violent movie at a theater, but noting that there is no data to show that reduction continues. (3) Knowing this, can we accurately predict instances of crime based on violent movie releases?

Related Works

With the advent of new technologies, tools for crime detection and prevention have become more innovative and sophisticated. Machine learning as a field has made advances in predicting crimes before they occur, changing the way we view policing. A model built at Rutger's University used neural networks to predict criminal activity based on arrest and booking records. The models also used features from the environment, such as broken glass, disorder, and antisocial behavior to predict crime, as well as rehabilitation initiatives to steer offenders away from further crime. The model had an accuracy rate of 83.95%. (5)

Another study compares different approaches such as neural networks, machine learning, and computer vision for crime forecasting. The methods in this study included: KNN algorithm, with accuracies of 67% and 87%, decision trees, with accuracies of 60% and 87%, and naive Bayes models with accuracies of 66% and 87%. However, the study also highlighted how as the dataset becomes more complex, the accuracy also tends to drop. (6)

Machine learning is not only useful in policing, but is also used in social monitoring online. An article from the University of Chinese Academy of Sciences used a Multilevel Attention Residual Neural Network (MARN) to enhance online rumor prediction and track the spread of misinformation on social media platforms, specifically Weibo. The MARN used both text and visual data in the model. The system used in this study was able to classify and detect rumors and misinformation with high accuracy. (7)

The movie industry also utilizes machine learning to predict ratings, revenue, reception, and even recommendations. Many models and studies have been done exploring data from movie releases. Notably, a study utilizes an SVM to predict the success of a movie on a scale of 1-5 based on historical data with accuracies between 83.44% to 89.27%. The model identified budget, IMDb rating, and number of screens showing the film as the strongest predictors of box office success. (8)

Machine learning shows promising advancements in predicting in the social sphere, criminology, as well as cinema. Studies show there may be a significant correlation between viewership of simulated violence on screen to behaviors in real life. However, there seem to be significant gaps in research predicting if violent movies influence violent crime. We aim to predict local violent crime following the release of these movies.

Methodology

1. Data Aggregation and Feature Extraction

Our dataset comprises:

1. Movie Metadata: We extract attributes such as runtime, revenue, and release date. Let $x_{metadata}$ represent this feature vector.
2. Temporal Feature Extraction: Using the release date, we construct a feature x_{Month} that represents the month of release. This monthly alignment allows us to match each movie with the relevant crime data in each state.
3. Genre Selection: Each movie is associated with multiple genres. We select only the primary genre g_1, arguing that the first genre in the list most accurately represents the movie's dominant theme. Let $x_{genre}= g_1$, a categorical variable that we later encode using one-hot for model input.
4. Movie Ratings: We extract IMDb, Rotten Tomatoes, and Metacritic ratings, denoted as $r_{imdbRating}$, r_{RT_Rating}, $r_{Metacritic}$, from the dictionary attribute Ratings. Ratings offer insight into public sentiment, with lower ratings potentially correlating with societal discontent during the release period.

2. Image Processing and Feature Extraction from Posters

We hypothesize that visual elements of movie posters (color scheme, brightness, dominant colors like red for intensity) may correlate with the movie's tone and, indirectly, societal effects.

1. Poster Data: For each movie, we retrieve the poster image P and preprocess it using a pre-trained residual neural network (ResNet) based on the ImageNet model. The ResNet extracts high-level features f_{poster}=ResNet(P), where f_{poster} is a feature vector summarizing the poster content.

2. Feature Integration: These extracted features are appended to the dataset as a new attribute vector, $x_{Poster\_Features} = f_{poster}$.

3. Data Preprocessing and Scaling

To prepare the data for modeling, we concatenate all feature vectors into a single vector:

$$x = [x_{metadata}, x_{month}, r_{imdbRating}, r_{RT\_Rating}, r_{Metacritic}, x_{Poster\_Features}]$$

Each numeric feature is scaled to zero mean and unit variance using StandardScalar. This scaling ensures that features with large variances do not disproportionately influence the model.

4. Model Architecture: Convolutional Neural Network (CNN)

Our primary model is a CNN designed to integrate visual and non-visual features.

1. Feature Input Layers: The model initially processes non-image features $\hat{x}_{metadata}$, $\hat{x}_{month}$, $\hat{r}_{imdbRating}$, $\hat{r}_{RT\_Rating}$, $\hat{r}_{Metacritic}$ through dense layers to capture interactions among numeric, categorical, and rating features.
2. Feature Fusion: After several dense layers, we concatenate the outputs from the dense layers and ResNet output to form a joint feature representation. Let $c_{fusion}=[c_{dense},x_{Poster\_Features}]$, where c_{dense} is the outputs of the dense layers.
3. Output Layer: The final layer has a linear activation to predict violent crime count y_{pred} for each state-month. Our loss function is mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y_i} - y_i|$$

where $\hat{y}_i$ is the model's prediction and y_i is the actual violent crime count.

5. Evaluation and Hypothesis Testing

Our model's performance will be evaluated using cross-validation on a state-by-state basis as states have different ranges based on their populations and aptitude for violence. We hypothesize that movies with darker themes, lower ratings and other may correlate with a higher frequency of violent crimes. This hypothesis will be tested by analyzing the relationship between model errors and specific features to determine if certain movie attributes are more predictive of crime patterns.

Preliminary Results

Our preliminary results show some potential but indicate room for improvement. We began with a basic model, incorporating only the runtime and revenue numerical variables along with the poster image data. The model was trained on the entire dataset to maximize performance. After 10 epochs, we observed a MAE of 155.6 violent crimes for the state of Alaska. Given that Alaska's average monthly violent crime count from January 2020 through December 2023 is 1243, this MAE represents over 10% of the mean, which suggests considerable room for improvement. We expect better accuracy with additional input variables and effective feature selection.

5 Future Plans

Enhancing Feature Set and Variable Exploration

There is room to increase the number of variables included in our X set significantly. Currently the model uses only runtime and revenue, but there is more information still in the 'Budget', 'Rated', 'Ratings', 'Title', and 'Genre' fields. These will need to be appropriately preprocessed and added to the model inputs. Our model currently compares our predictor variables to crime by month by matching movies released in year-month 'x' to crime report numbers from the corresponding month. However, movie release dates can fall on any day in the month. Does a movie released on August 26th have an impact on crimes for the month of August? It is worth testing whether introducing a lag here can have beneficial results. That is, predicting crime rates in May based on movies from February/March/April. Another more interesting change would be the addition of movie summary information. Since we are relying on some sort of emotional impact from movies as a cause crime increases/decreases, it would be great if we could better capture the tone or themes of movies in our dataset. The online movies database offers short and long plot summaries. Acquiring these via the api should be quick, and with that extra variable we could perform sentiment analysis or keyword searching to produce rich features.

Neural Network Optimization and Interpretability

There is a significant amount of room to expand the number of layers and to experiment with the shapes of them. Along with this, we could also experiment with different types of activation functions to see how results differ. We can test changes in learning rate and consider early stopping to avoid overfitting. Trying different optimizers such as stochastic gradient descent as well may be useful. To do this we can try different layer structures and then organize hyperparameters into a gridsearch to find a more optimal model layout. Once we have come to a more optimized model, we can evaluate feature importance to see which features have the greatest impact. There are a number of methods to do this available including LIME for numeric inputs and LRP for image data, though this may be more complicated for our combined model.

Image Processing and Image Feature Extraction

In our current model we use the default pretrained model ResNet50 from tf.keras.applications. It is worth testing deeper pretrained ResNet models such as ResNet101 or ResNet152 to try to generate features accounting for more complex patterns. The current ResNet50 model generates feature vectors for each of our movies in a relatively short amount of time, so a larger but slower model would be viable. In addition to this we could try unfreezing some of the final layers of the ResNet model. In our current implementation the final classification layer is removed, but the remaining layers are all frozen by default in their current pretrained state. Unfreezing some of the last few layers would allow our ResNet model to adapt more specifically to our movie poster dataset. This could be useful because the pretrained ResNet models are trained on a number of everyday objects, scenes, people and animals but are not specifically trained on posters. Allowing the model to adapt to slightly different images could improve our combined models overall performance.

Extending Model Evaluation

Implementing K-fold cross validation and testing the model on a series of subsets may give a more robust view of model performance. We can test this with a number of different states or with states combined. This should lead to a more general measure of model performance. Importantly, we should compare our model to more simple models. We can compare the performance of our complex model to a linear regression model that just uses a subset of movie details. The task of predicting violent crime from movie posters/details is incredibly complex. We can spend a lot of time tinkering with our model, but to put this work in context we need to know how much better, if at all, this is than just a simple linear model. If not, we may need to reevaluate parts of our approach.

Other Things

Refining/preprocessing crime data. I do not understand the crime data very well but maybe we can better handle it somehow.
PCA is applicable to several areas here I think. I’m just not sure exactly where to put it.
One-hot encoding would be part of the preprocessing for genres.
I’m sure there are things I am leaving out here. Any ideas for neural network construction or any other part are welcome. We can put a timeline separately as below

or integrate it throughout the sections with peoples names. I'm not quite sure how sharing work and doing things in order exactly works. Happy to discuss further.

Order of Actions/Timeline

1. By November 11th - Expand our evaluation process so that we have more robust and usable metrics to use as we refine our model.
2. By November 15th - Integrate additional variables/variable transformations including additional details as well as movie summaries. Prioritize smaller achievable chunks, not super in-depth explorations.
3. By November 20th - Refine model and hyperparameters. Use gridsearch to explore neural net architectures and hyperparameter settings. Identify best model configuration.
4. By November 22nd - Interpretation of feature importance for numeric and image inputs. This will be essential for organizing final results and conclusions.
5. By November 25th - Final report. Perform final model evaluations. Ensure that model changes are documented and described over time to describe and justify model decisions. Produce conclusions and ideas/recommendations for future work.

References:

1) https://pmc.ncbi.nlm.nih.gov/articles/PMC6790614/
2) http://gip.uniovi.es/docume/pro_vs/vv_ado.pdf
3) https://econweb.ucsd.edu/~gdahl/papers/movies-and-violence.pdf
4) https://dl.acm.org/doi/pdf/10.1145/3325112.3328221
5) https://dl.acm.org/doi/pdf/10.1145/3325112.3328221
6) https://link.springer.com/content/pdf/10.1186/s42492-021-00075-z.pdf
7) https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2021.711221/full
8) https://ieeexplore.ieee.org/document/8281839