BOX OFFICE FORECAST

ZHANGTAO XUE

Abstract. With the development of deep learning, more and more applications use the knowledge of deep learning to solve practical problems. This paper discusses how to use depth to solve the problem of box office forecast. With the improvement of living standards, people have more time to improve their quality of life, such as watching movies. In the first half of 2017, China's film market showed new characteristics and trends. Under the dual promotion of policy guidance and market regulation, the development of the domestic film industry will enter the development stage driven by key high-quality content. As of June 30, 2017, the total box office revenue of Chinese mainland films was 27.177 billion yuan, with a year-on-year growth of 10.51. Therefore, it is necessary for the film industry to predict how much profit it may make if it makes a film. In this paper, deep learning neural network is used to predict the income of a movie.
Keywords: Deep learning, film, business income

Contents

Date: (None).
1991 *Mathematics Subject Classification.* Artificial Intelligence.
Key words and phrases. Deep learning, film, business income.

1. Introduction

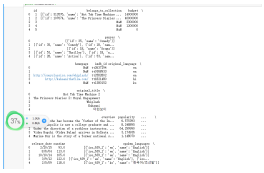
Deep learning has been studied in many research fields, such as image recognition, natural language processing and prediction system. In short, deep learning is a new field in machine learning. Its motivation is to build and simulate the neural network of human brain for analytical learning. It imitates the mechanism of human brain to interpret data.

In this study, we take out several key factors that affect the final business income, such as early investment, popularity, time and so on, and build a two-layer neural network to predict the film revenue. Then these key values are put into the neural network as feature input, and then the back propagation of BF neural network is used to update the weights, and finally the prediction model is obtained.

This paper aims at the existing methods to solve the practical problems of the film industry, so that the film industry can use the effective data to predict the possible revenue of making a film. The experimental results show that using neural network to solve the film revenue is a very desirable method.

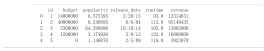
2. Related Work

First of all, our work can be divided into the following steps:
(1) Data download: the data source is directly downloaded and used from the kaggle project.



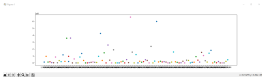
Flag.1

(2) Data cleaning: delete some unimportant or less influential data, such as poster_ path overview imdb_ ID, homepage, etc



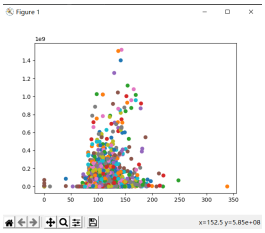
Flag.1

(3) Data analysis: To observe the impact of previous investment on the film, we use Matplotlib in Python to draw a picture, flag 1, we can see that the investment is directly proportional to the business income.



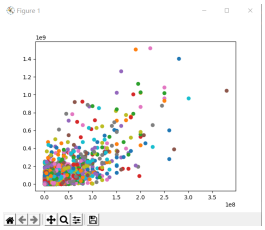
Flag.1

To observe the impact of previous investment on the film, we use Matplotlib in Python to draw a picture, flag 1, we can see that the investment is directly proportional to the business income.



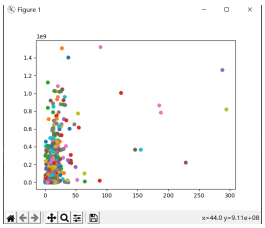
Flag.1

To observe the impact of previous investment on the film, we use Matplotlib in Python to draw a picture, flag 1, we can see that the investment is directly proportional to the business income.



Flag.1

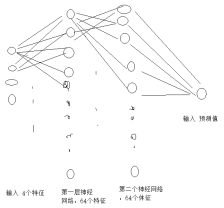
To observe the impact of previous investment on the film, we use Matplotlib in Python to draw a picture, flag 1, we can see that the investment is directly proportional to the business income.



The relationship between popularity and business income

3. Methodology

In this section, we will introduce our methodology framework. We use a three-layer neural network. There are 64 neurons in the first layer, 64 neurons in the second layer, and 1 neuron in the last layer, Finally, the cost function is selected as the square difference, and the design figure flag. 5



The relationship between popularity and business income

3.1 parameter learning

From the data analysis, we select several parameters which have the greatest impact on the business income as X. the weights of the first layer of neural network are W1 and bias value B1, the weight of the second layer of neural network is W2 and bias value B1, and the third layer of neural network W3 and bias value B1.
The first layer of neural network is: Z1 = relu (w1x + B1)
The second layer neural network is: Z2 = relu (w2z1 + B1)
The second layer neural network is: Z3 = relu (w3z2 + B1)
We get W1, W2 and W3 by neural network updating

4. Experiment

4.1 Dataset

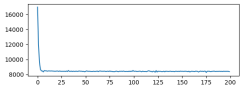
We downloaded data sets from the kaggle project, including test sets (including 3000 data sets) and training sets (1000 data sets). The dataset contains budget, genre, home page, IMDB_ ID, original language, original title, overview, popularity, poster path, production company, production country, release date, running time, spoken language, status, slogan, title, keywords, actors, staff, etc. Delete the weights that are not important or have little impact on the dataset, leaving only the feature values with large weight influence, such as budget, release date, running time, popularity, etc.

4.2 Hyper-parameter Settings

The training model of this neural network has three layers. The corresponding weights are W1, W2 and W3. The cost function is designed as mean square error, learning rate Le = 0.01, update times 200, each time 200 data.

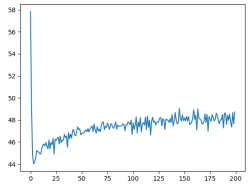
4.3 Overall Performance

Firstly, for the overall analysis, we forecast from these aspects:
(1) The update rate of cost function is the cost function of mean square error. We can clearly see that after 10 times of updating, the loss value of his cost function remains around 8000.



cost

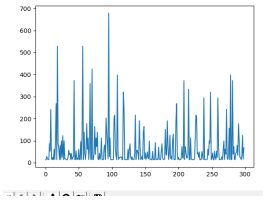
(2) The difference between the actual value and the predicted value, we can clearly see that the difference between the actual value and the final actual value is about 46.



cost

4.4 Forecast

We input the test data into the model and test the model to give the predicted value of movie evaluation.



cost

5. Conclusions

This paper combines deep learning with practice and application to solve the prediction value of profit demanded by the film industry. In our experiment, we compare each factor, compare it with the film profit factor, and make corresponding image display. And we use the neural network related knowledge design model structure, training, get a good result, from the final experimental results, the cost function is small to 8000, the square difference between the predicted value and the actual value is about 46, the effect is also good. In the future, we plan to use multi task learning to improve the method and improve the accuracy of prediction.

List of Todos

(A. 1) School of Computer Science,, Xi'an Shiyou University, Shaanxi 710065, China
Email address, A. 1: xxx@tulip.academy