Pakistan Social Sciences Review
June 2018, Vol. 2, No. 1 [295-304]
P-ISSN 2664-0422
O-ISSN 2664-0430

Pakistan Social Sciences Review
www.pssr.org.pk

RESEARCH PAPER

Sentiment-based Movie Recommender System using Deep Learning

Naveed Jhamat[1] Ghulam Mustafa[2] Zeeshan Arshad[3]

1. Assistant Professor, Department of Information Technology, The University of the Punjab, Gujranwala Campus, Punjab, Pakistan
2. Assistant Professor, Department of Information Technology, The University of the Punjab, Gujranwala Campus, Punjab, Pakistan
3. Lecturer, Department of Information Technology, The University of the Punjab, Gujranwala Campus, Punjab, Pakistan

PAPER INFO

Received:
February 15, 2018
Accepted:
June 24, 2018
Online:
June 30, 2018

Keywords:
Artificial Neural Networks, Emotions, Movies Recommendations Recommender Systems

Corresponding Author:
gmustafa@pugc.edu.pk

ABSTRACT

The emergence of overloaded internet content poses many new challenges for users and content providers alike. To limit the amount of content viewed, such as videos, music, or other products of content providers such as Netflix or Amazon, the recommender systems are used to guide the user via the available material. These services collect knowledge about the customer and strive to deliver personalized experience. Many sophisticated programs have an approach to contents, but often neglect to take into consideration the nature of user sentiments. This leads to the creation of a sentiment-as-input paradigm incorporating current studies on the recognition of human sentiment and emotionally classifying material within the movie domain. Multiple models of the learning are tested and an ANN is selected due to its outstanding performance. The results of this analysis show that the user sentiment should be used to suggest tailor-made content rather than random content.

Introduction

New technology offered a mass of information and diverse forms of content to virtually everyone connected to the internet. Internet academic search engines such as Google Scholars, music streaming platforms such as Spotify or firms that offer on-demand video services such as Amazon are only some examples of services that have arisen since the turn of the last century. Since the customer gets a vast amount of new knowledge and content introduced almost half of the time. This leads to many problems for consumers such as locating the right information or handling the increasing amount of data.In order to find out whether a consumer actually needs to turn this knowledge into the actual content, expert systems are built. These programs are known as recommendation systems (Burke, 2002) and are meant to assist consumers in discovering the content they want. It aims at directing users by seeking the right information and details. They are also designed to resolve the fact that hundreds of different items will result in a search query. This greatly increases the Internet surfing experience and usability (Leiner et al., 2009).

Many of today's recommended systems are focused on the user's consumption and search experience. The user history can be looked at and linked to another users' history. Related users are clustered and the contents accessed by one user to the other users within the community are indicated. But psychological studies found that, because of many influencing factors, a man's emotional state would change frequently. It may be argued that these advanced programs will give each customer precise and tailored recommendations. In order to make better decisions, a recommendation system can also contain a user's current emotional state. To build a smoother user interface and greater user loyalty, a recommendation should be a natural fit. The background of movies was chosen to illustrate case of how a system of emotions recommenders would appear.

Here we discuss the philosophy behind the classification of human mood (user knowledge). The functional knowledge aspect is explained in depth. Until looking at the classification of movie knowledge, the principle behind movie classification (catalogue knowledge) is presented. Emotions are strong about someone or something. Overall, the sentiment was observed to have either a positive effect or a negative influence. Both of these two dimensions has also been grouped into a 10-point list reflecting a particular mood. Each individual consumer scores each object from 1 (very or not) to 5 (extremely) when assessing the mood of a person using PANAS (Staw and Cummings, 1996). The answers given are paired with a score that reflects the current sentiment of the consumer. As this measured score excludes quite a lot of insights it should be addressed whether all the individual responses are to be taken into account.

Most recommender systems do not really have a function that captures these sentiment changes and suppresses the user's interest about new topics. Although the user model is often connected to an individual account of the most traditional solutions, the framework is only suitable for the single user. For example, if a user just watches science-fiction (sci-fi) films, the suggested film bubble is full of films such as “Inception” and “Avengers”. If this 'personalized' recommendation method is currently used by a consumer with a preference for romantic drama films, he will not be pleased with the recommendations offered. Also, if this user continued looking for this genre for “The Notebook” or other films, this would conflict with the initial user's current model.

There is little trust in their recommendations in the current methods. Before the correct type of content is defined and the like-minded community is sought, it takes some time. Although there might be a number of other factors, other than the content, in a particular case, there are reasons to think a movie is not suitable. Very limited information about how effective those guidelines are based on these programs is available. It is hard to argue which of these strategies has the better performance. The author concludes that the better solution would be to use a hybrid of the two alternatives to the recommendation system, taking into consideration the length of time it would take to locate a comparable party. The customer would then be able to make a more educated decision which would make the solution more successful.

In this paper, we're going to look at how to create a video recommendation system based on the user's emotional state. We would also analyse whether it is feasible to merge various recommended frameworks in order to make high-quality recommendations. The paper would then describe the research challenge that this paper will tackle and then address the various research questions to help solve it. Research problems can then be addressed in a straightforward manner to illustrate how this issue can be approached in the most transparent way possible. In the following sections, we will discuss some of the key research problems which have not yet been thoroughly examined in this area of study. The questions are as follows: 1. How does a recommendation system that recommends movies using traditional machine learning algorithms work against the sentiment-based movie Recommender system that uses the artificial neural networks? 2. How do we compare a sentiment-based system's success with the performance of a system that use traditional machine learning recommendations?

Literature Review

The need for recommendation systems is generated by growing amounts of movie contents. The purpose of this section is to illustrate the key principles and methodologies used in the most common recommendation systems.

A recommender system is a framework that provides consumers with individualized recommendations. Depending on different requirements, there are several distinct types of recommendation systems. In the succeeding paragraphs, the most familiar types are explained. By learning about the various consumer expectations and then predicting their particular needs, the goal is to achieve these individualized outputs. What sort of method is picked and adapted to the condition depends on the context. The types or techniques of recommender systems vary in the manner in which and how a particular recommendation is provided.

In content based recommender systems, research has centered on recommending products with textual content associated with them. By learning something about a customer, such as demographic statistics, one may make a better customized suggestion. For example, provided knowledge about the movie genre, and knowing that a user liked "Ratatouille" and "Cars" one might assume an Animation predilection and might thus recommend "The Loin King". The principal benefit of content oriented guidance is that data sparsity is alleviated. One downside, though, is that many consumers and objects in large-scale real-world applications frequently lack content and background knowledge.

In comparison to content-based recommended systems, collaborative recommended systems are one of the most widely used systems. They can be binary, time-based, model-based, or memory-based (Burke, 2002). Movie recommendations on Netflix are a typical example for this recommendation method. How systems are currently marked and matched against each other is also kept confidential by the various service providers. The framework detects correlations between users on the basis of their scores and then provides new recommendations on the basis of inter-user comparisons.

The baseline of certain types of systems is generated by demographic recommenders. These characteristics can include age, gender and other personal characteristics. The derived model of all composite personal characteristics is then matched to a manually generated catalogue of consumer stereo types. For example, when the user first uses a recommendation method, the user's model does not yet work. Although the preferences of the users are uncertain, only the demographic details of the current users can be used as the basis of suggestions. Long-term estimation of the user with scores is not mandatory since the system knows the user over time. The major challenge is that each user can define this function, which is mostly achieved using various techniques of user satisfaction. These programs aim to quantify each object's usefulness with respect to the user. There is basic information about the interaction between the interests of a customer and a particular object in a knowledge based recommender system.

In their article, (Wakil et al., 2015) proposed a recommendation framework that incorporates a user's emotion. They attempt to solve the problem and how to relate the required film material in a valid user model. It aims to understand the user-percept of a movie using multiple forms of ranking systems comparable to a content-based filtering scheme. The more people who offer a ranking, the more ideas the user would have and the more exact the groups are, so that a dialogue. The user's role on the receiving end switches and creates a dialogue. The system invites the user to select from multiple colors three times, where each color is emotionally attributed. If two of the chosen colors are of the same emotional state, this emotional state is assigned to the individual. The author believes that the framework of suggestion is a core aspect of the online experience of users on the website of film streaming.

Material and Methods

Data Collection

To choose potential outputs, the IMdB movie database with its extensive marking was used. As the primary grouping mark, it was agreed to use the genre. A simple review has shown that 24 distinct genres exist. These genres vary from adventure to history, with more than one genre labelled in most movies. Many brands are only marginally capable of explaining the entire plot of a movie. The aim of the study is to obtain adequate data to generate substantial findings. The time available is the principal constraint for the volume of data obtained. As the data model is largely processed, a much smaller amount is taken into account. There is one significant benefit dependent on the nature of the recommendation process.I got much of the answers through sharing structures with social media, movie forums and machine learning forums and a variety of online channels. I collected data for about 8 weeks for the recommender system. A total of 1000 responses received from participants to the sentiment-based recommender.

Recommender Model

Supervised learning is a computer science concept that attempts to relate an input to the output by looking at the input/output pairs that actually exist. In this sense, it is possible to characterize the accuracy of an algorithm as the extent to which the algorithm can predict the result of a given dataset. We would use five different algorithms to test the recommender system. The algorithms are chosen according to the type of knowledge and on their own strengths and weaknesses. The algorithms are then used for a test data set, the results are then compared to see how well they performs. On the basis of training data sets, the five algorithms are then checked to see if their precision, recall and F1-measure and Accuracy can be calculated using a variety of data models to produce best performance.

Generally speaking, Accuracy can be defined as the percentage of cases in which the algorithm was able to predict a movie set's correct rating against a test dataset. For the moment, all consumers who have not been able to locate a movie are combined into one category. This will have a negative effect on the data model as it can aim to recommend a film for unique and even contradictory sentiments when no movie is currently available so far. Following algorithms are used in our sentiment-based recommender system.

First algorithm is known as Nearest Neighbor (KNN) algorithm(Bishop, 2006). The subset of adjacent instances is interpolated by the k-Nearest Neighbor (KNN) algorithm. The distance between the multiple subsets is determined. The kNN is a conventional way to solve complex problems of grouping. Second algorithm used in our experiments is known as Naive Bayes (NBC)(Bishop, 2006). A classifier Naive Bayes (NBC) takes a more mathematical approach and its underlying probability equations characterize it. It can be used in tracking learning scenarios and the outcome of a test instance can be correctly estimated. For my data model, I selected a special Gaussian distribution as many problems in the real-world classification have been shown to be solved (John and Langley, 1995).

The third algorithm that is used is known as Adaboost(Freund and Schapire,1996). A Adaboost is a sub-concept of anensemble learning technique that can be described as a mix of several decision trees. A Adaboostconsists of multiple trainable decision-trees and aims to overcome general problems of classification.

To make the method thorough, the established Adaboost uses a couple of parameters that need to be recorded. I have to specify in the Adaboost how many iterations I would like to have in the Adaboost. The fourth algorithm that is used in our experiments is known as Artificial Neural Network (ANN).ANN is attempting to overcome the problems by changing the weights of the connections. The algorithm was practiced at a batch size of five for 1000 epochs. To assess whether a sensor is shot or not, a ReLU (Rectified Linear Unit) activation feature was used. In order to represent the 22 dimensions of feedback into one of the available genres, four hidden strata with 100 perceptrons were selected.

The last algorithm is C4.5(Quinlan, 2014). In most popular packages the popularity of classification tree models derives from their strong adoption, simple analysis and availability of acceptable estimation routines. The decision tree algorithm C4.5 is one of the most commonly used in the business and engineering of data mining. When the target variable is discrete, a classification tree can be created. In general, the construction of the C4.5 model decision tree consists of the three phases that follow. The first phase is the branch growth. The tree growth is also dependent on the data gains rate of the target variable section. The processing of discrete variables is the second step. The third step has been pruning. The pre-selection pruning approach is chosen in order to avoid over-fitting of the model.

Results and Discussion

The five models are tested in using four different measures. The numbers are meant to explain or refute the performance of different models. The tests attempt to achieve a better understanding of the proposing systems' results. The observations from graph are presented and all results are summarized.

Table1
Comparison of Different Algorithms for Sentiment-based recommender systems using four difference measures

Methodology	Accuracy	F1-Score	Precision	Recall
kNN	0.5628	0.5767	0.6089	0.5477
Naïve Bayes	0.7037	0.7124	0.7171	0.7078
Adaboost	0.6581	0.6708	0.6566	0.6855
C4.5	0.5754	0.5892	0.5924	0.5860
ANN	0.7692	0.7692	0.7692	0.7692

On the test data, k-Nearest Neighbor's algorithm returned the 0.5628 accuracy. As seen in Table 1, kNN was given an F1-score of 0.5767. The Precision and Recall measure of this approach is 0.6089 and 0.5477, respectively. On the other hand, the Accuracy and other measures are used for Navie Bayes (NB) algorithm. The NB is given a 0.7171 precision ranking above the kNN. The Accracy is 0.7037 which is ahead of the previous algorithm kNN and similarly other measures are better than kNN. The Recall of NB method is 0.7078 which is better than kNN. The F1-score of the NB is 0.7124 which is better than previous algorithm.

By splitting the data collection into a training and test set, the accuracy is measured. The training set (consisting of multiple decision trees) is used to construct the Adaboost. The test set is then used to verify the algorithm. 0.6581Accuracy of the Adaboost was measured, as can be seen in Table 1.The Accuracy of Adaboost is better than kNN but below than NB. Similarly, F1-score which is 0.6708 of Adaboost is better than kNN and less than NB. The Precision and Recall of the Adaboost algorithm are 0.6566 and0.6855 respectively better than kNN but lower than NB.

The Accuracy of the given test data and labels are specified and measured usingC4.5. The C4.5 earned a 0.5924 Precision score that is below the previous methods provided. However, the Precision of C4.5 is slightly better than kNN. The Accuracy of the C4.5 is also better than all previous methods. Lastly, The ANN achieved a 0.7692 accuracy which is better than all previous algorithms. Similarly, on other measures ANN outperformed all other models on all measure. The Precision, Recall and F1-score of ANN is 0.7692,0.7692 and 0.7692 respectively. ANN performed better than all other models on all measures.

The Sentiment-based Recommendation System using ANN provides high success rates of movies recommendations which demonstrate that the recommendations are interesting and useful to the consumer. The system was developed to demonstrate its ability to screen the right movie. The study found that the sentiment-base recommendation system functions better than other recommender system with traditional algorithms.

It also provides recommendations of high importance more effectively by substantially reducing error. The ANN model performs better than other models and it is difficult to equate it with other recommender models. The recommender based on the sentiment and Artificial Neural Network (ANN) is better than the recommendation framework based on recommender system with traditional algorithms. That is because the Sentiment-based method can capture the correct sentiment and provide an acceptable recommendation for all genre.

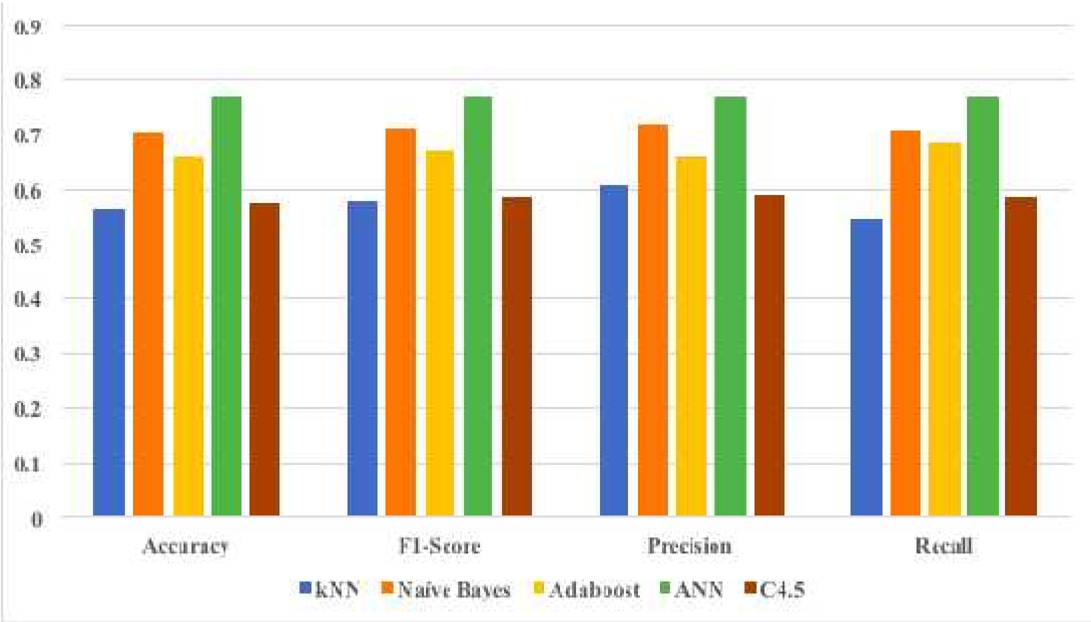

Figure 1: The performance of recommendation models using different measures.

The new user dilemma will emerge if a new user attempt to request a recommendation. Since the user uses the system for the first time, he has no details about the user's interests. This dilemma is "solved" over time as the machine gets to know the user, but no information is known on the length of this process until the user has scored appropriate movies to have good quality feedback. The time taken to solve the new user issue can be easily addressed for the Sentiment-based recommender System. For a consumer to be comfortable with the system it takes between few minutes. The framework then generates movie recommendations for the user. Sentiment-based recommendation programs are based on a range of highly personal knowledge. A new form of filter bubble based on the user sentiment can be generated by the system.

Conclusion

In order to bridge the information gap, a sentiment-based recommendation system was built to consider what the consumer needs? The artificial neural network (ANN)model is effective throughout and outclass the other models. The process of incorporating these approaches into the current movie recommendation mechanisms and creating a smoother experience for consumers. The Sentiments added to the recommender process enhanced the performance of the ANN based recommender system. The additional knowledge added to the system also help to address the new user problem effectively.

Recommendations

By comparing various machine learning approaches, the sentiment-based recommendation method using deep neural network was chosen. By adding more movies to the database and getting more feedback from users, the sentiment-based system could be enhanced. Just one or two genres are used in the new system. This may be updated to make more than one genre possible. In both directions, depth and width, database capacity can be increased to make the system more scalable. The approach involves work on optimizing movie recommender systems performance with more than one movie per recommendation. In order to align the effect of each method with other approaches (such as collaborative or content-based systems). It would be promising to apply a hybrid approach for recommendations, especially in overcoming the problem of the new user or in scaling up the project.

References

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, *12*(4), 331-370.

Freund, Y., &Schapire, R. E. (1996, July). Experiments with a new boosting algorithm. In *icml* (Vol. 96, pp. 148-156

John, G. H., & Langley, P. (2013). Estimating continuous distributions in Bayesian classifiers. *arXiv preprint arXiv:1302.4964*.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... & Wolff, S. S. (1997). The past and future history of the Internet. *Communications of the ACM*, *40*(2), 102-108.

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Staw, B. M., & Larry, L. Cummings (1996). *Research in organizational behavior: An annual series of analytical essays and critical reviews. Greenwich Conn. and London: JAI Press*.

Wakil, K., Bakhtyar, R., Ali, K., &Alaadin, K. (2015). Improving web movie recommender system based on emotions. *International Journal of Advanced Computer Science and Applications*, *6*(2), 218-226.