A Feature-Based Neural Network Movie Selection Approach

N. Karunanithi
2Q-265, Bellcore
445 South Street
Morristown, NJ 07960
karun@faline.bellcore.com

Joshua Alspector
2P-396, Bellcore
445 South Street
Morristown, NJ 07960
josh@bellcore.com

Abstract

The huge amount of information available in the currently evolving information infrastructure at any one time can easily overwhelm end-users. One way to tackle this situation is to use an "information filtering agent" which can select information according to the interest and/or need of an end-user. However, at present no information filtering agents exist for the evolving multimedia information infrastructure. In this study, we demonstrate the use of a feature-based neural network filtering agent for the video-on-demand application. We evaluate the neural network approach for 10 voluntary subjects who provided ratings for the movies. Our preliminary results suggest that the feature-based selection can be a useful tool to recommend movies according to the taste of the user and that the feature-based approach can be as effective as a movie rating expert. We also compare our feature-based approach with a clique-based approach.

1 Introduction

In recent years, computer-network-based information services have gained wide acceptance both within commercial and non-commercial sectors. The information content in such services is mostly textual. However, the currently evolving U. S. National Information Infrastructure (NII) is expected to support not only a variety of text-based information services but also various multimedia (hypertext, audio and video-based) information services. Some of the potential application domains in which the information infrastructure is likely to have impact are banking at home, access to electronic libraries, distance learning and laboratories, delivery of news and entertainment on demand, electronic shopping malls, law enforcement and security alertness, legal services, national health care and weather services, and telecommuting. Thus, the NII has the potential to change the way we work, communicate, travel, and generally access information.

The huge amount of information available in the information infrastructure at any one time can easily overwhelm end-users. Even within existing computer-network-based information services, providing information that is of interest to a particular end-user is not an easy task. For example, filtering relevant information in the "Internet" is not easy because a single message may be sent over a set of mailers, a message may consists of a "thread" (a sequence of "replies" to the original mail), or the header may not reflect what the actual content means. This situation is likely to worsen in the future multimedia information infrastructure unless the end-user has the ability to filter information based on what is relevant to him/her.

Several useful text-based tools exist for navigational purposes[5] on the "Internet". An example of a common-to-all user interface for the Internet is "Mosaic". Mosaic is a hyper-text based easy-to-use X-Window interface built on top of various Internet navigational, browsing tools such as "Gopher", WAIS, World Wide Web, Archie, FTP service etc. Even with such a common interface, these navigational aids require network support, and "active participation" of the end-user. However, at present no equivalent filtering systems exist for the evolving multimedia information infrastructure.

In this study, we evaluate the use of a feature-based neural network movie recommendation system for the evolving video-on-demand service. The initial study is based on a feature database of 1548 movies. We evaluate the approach for 10 voluntary subjects who provided ratings for the movies. Our preliminary results suggest that the feature-based selection can be a useful tool to recommend movies according to the taste of the user and that the feature-based approach can be as effective as a movie rating expert. We also compare our feature-based approach with a clique-based approach.

2 An Experiment on Movie Selection

In this study, we evaluate two fundamentally different approaches: a *clique-based* approach[1] and a *feature-based* approach. In the *clique-based* approach, movies are recommended using the ratings of a set of users who might have similar taste. In the *feature-based* approach, first a model is built using a set of important features of the movies that a user has seen and rated, and then that model is used to predict the ratings for movies that the user wants to see. It should be noted that a hybrid approach can also be developed by combining these two approaches. We present more details and a preliminary evaluation of these approaches in subsequent sections.

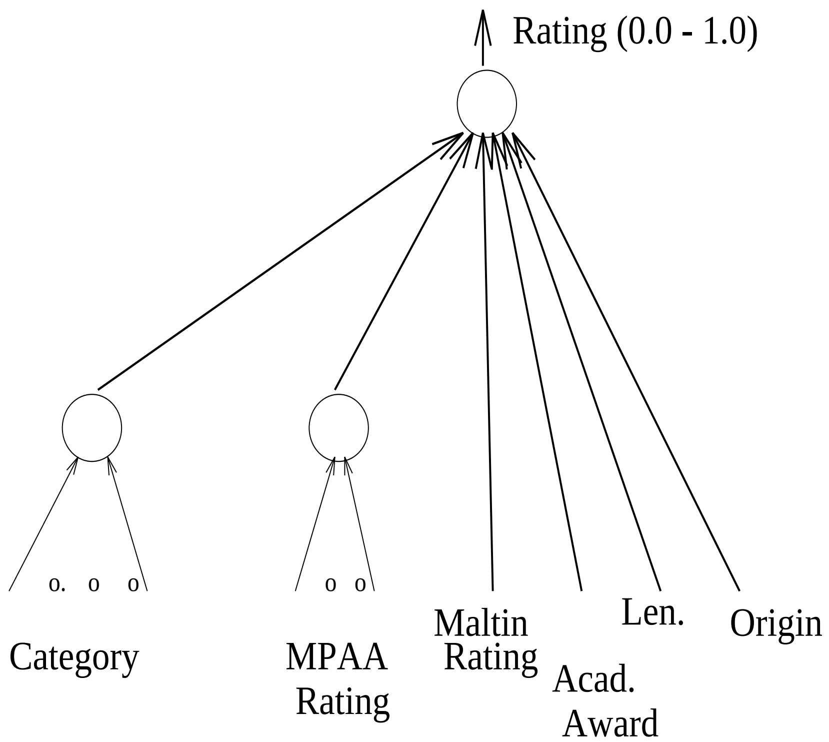

Figure 1: The Neural network architecture used for the feature-based approach

[1]The clique-based approach is a variant of the approach suggested by Will Hill of Bellcore.

2.1 A Feature-Based Approach

The feature-based approach is based on the notion that the features of the movies can be useful in recommending movies. Some of the features that can be used to recommend movies include MPAA ratings, expert critic ratings, movie category, name of the director, leading actor/actresses, and awards received. For example, a particular user may have a strong inclination to see only movies that are rated G, acted by Ben Kingsley, and belonging to the category Comedy or Drama. Thus, the feature-based approach exploits the bias of a user towards a set of important features of the movies.

The algorithm for the feature-based approach is as follows:

1. Extract relevant features from the movies that the user has rated.
2. Build a model for the user by associating selected features (as inputs) and the ratings (as output).
3. Estimate ratings for the movie that the target user wants to see by considering its features as new input to the model.

To implement this algorithm, one must select a proper set of features and a correct model to predict ratings. We already suggested several useful features for rating movies. However, in this study, we use only the six features that are shown in Figure 1.

The rationale for selecting a small set of features is that we wanted to evaluate whether we can build a reasonable model with as few features as possible. For step 2, we do not make any specific recommendation as to what modeling approach might be superior. The choice for modeling, for example, may include a linear or nonlinear regression model, an expert system, a neural network, or other approaches. In this study, we illustrate the use of the neural net.

The architecture of the neural network model is shown in Figure 1. The network is a simplified form of a modular network in that the inputs for features Category and MPAA Ratings are first fed to separate hidden units rather than directly into the network. The actual encodings used for these features are illustrated in Table 1.

2.2 A Clique-Based Approach

The clique-based approach is based on the hypothesis that the average rating of a *clique* of users is the best indicator of an individual's future rating. A set of users form a clique if their movie ratings are similar. A member of the clique for whom we want to predict a future rating is considered a *target user*. In order to identify a clique for a target user, we define two parameters: C_{min}, the *correlation threshold* and S_{min}, the *size threshold*. The parameter C_{min} defines the minimum correlation required for a user to become a member of the clique of a target user. Pearson's correlation

Feature	Type of Encoding	No. of Input Units
Category	1-of-N unary	25
MPAA Ratings	1-of-N unary	6
Maltin Ratings	Real value between 0.0 and 1.0	1
Academy Award	Real value between 0.0 and 1.0	1
Length (minutes)	Real value normalized by the mean	1
Origin	Real value between 0.0 and 1.0	1

Table 1: Encoding used for features

coefficient is computed between users using vectors of ratings for the movies that both users have seen and rated. Thus, the parameter C_{min} imposes a minimum limit on similarity of ratings required for a user to be a member of the clique of a target user. The second parameter, S_{min}, defines a lower limit on the number of movies that a user must have seen and rated in common in order to be a member of the clique. The parameter S_{min} is used to restrict users who have not seen the same movies that the target user has seen from entering the clique. Thus, a clique is defined for a target user by specifying suitable values for C_{min} and S_{min}.

Thus, the movie recommendation algorithm for the clique-based approach involves the following steps.

1. Initialize parameters C_{min} and S_{min}.
2. Identify a clique for the target user.
3. Estimate ratings for the movie that the target user wants to see by considering the ratings by the members of the clique who have already rated that movie.

In our preliminary implementation, we use a constant value of 10 for S_{min} and a positive variable value for C_{min} such that the number of users in a clique is held constant at 40. And in step 3 we use a simple arithmetic mean of the ratings of the members of the clique. Thus, the average ratings of the clique becomes the predicted ratings for the target user. The rationale behind computing the average is that a target user's taste may not be very different from that of his/her clique.

2.3 Data Collection

The feature database for the experiment was populated from the Microsoft Cinemania[2] CD-ROM for 1548 movies in Will Hill's survey. The features

[2] Cinemania is a registered trademark of Microsoft Inc.

collected include MPAA ratings, Maltin ratings, Category, Academy award, Length of the movie in minutes, the country of Origin, Director, leading Actor/Actresses, and a short review. As pointed out earlier, only the first six features are used in this study. The test-case users are 242 Internet subscribers who volunteered to rate the movies that he/she had seen. Each movie was rated on a scale of 0 to 10, with 0 being the worst, and 10 the best. The number of movies rated by the individual user varied from 0 to a maximum of 456 with the average being 177. For the purpose of evaluation, we selected 10 users who have rated approximately 350 or more movies as our *target users*. For convenience we label them as S3, S21, S39, S41, S46, S77, S111, S124, S145, and S178.

2.4 Results

The experiment for this study was conducted as follows: For each target user, we split the data set into a *training set* and a *test set*. The training set was used to build the model while the test set was used to validate the model. (In the clique-based approach, the training set was used to identify the clique for the target user.) The training set had 90% of the ratings and the test set had the remaining 10% of the data. In order to make a fair comparison, we cross-validated our model by splitting the data into 10 different mutually exclusive training and test sets. Thus, for each target user there are 10 experiments.

Target User Ids	No. of Movies Rated	Correlation Vs.		
		Maltin Ratings	Neural Net Ratings	Clique Averages
S3	364	0.23	0.29	0.51
S21	397	0.10	0.28	0.43
S39	436	0.43	0.45	0.67
S41	354	0.40	0.53	0.52
S46	456	0.38	0.48	0.57
S77	348	0.13	0.27	0.50
S111	374	0.58	0.58	0.64
S124	420	0.43	0.48	0.65
S145	367	0.41	0.39	0.64
S178	402	0.48	0.48	0.72

Table 2: A summary of ratings by different approaches.

Table 2 shows the results of our study. The results are presented in terms of the correlation coefficient between the actual ratings by the target users and the ratings by different movie recommendation approaches. For the neural network approach, the ratings represent the average over 25 experiments for each training set. The third column in Table 2 represents the correlation between target users and Maltin ratings. The fourth column represents the correlation between neural network predictions and actual ratings. The fifth column represents the correlation between ratings by the clique-based approach and target users. A higher correlation in Table 2 implies that the predicted ratings are close to the actual ratings. A comparison of results in column three and four suggests that the feature-based neural network ratings are better than Maltin (a movie rating expert) ratings for the target users S3, S21, S39, S41, S46, S77, and S124. For the remaining three users, Maltin ratings seems to be marginally better than the neural network ratings. The average difference between actual ratings and the neural network ratings for the target users is around 1.2 (on a scale of 10) with a standard deviation of 0.95. Thus, the neural network based movie recommendation approach may be used as an alternative to Maltin ratings. A comparison of the fifth column against other columns suggests that the clique-based approach is better than other two approaches for all target users except S41. Thus, for some users, the clique-baseed approach is a better movie recommendation system than our preliminary neural network implementation.

3 Conclusion

We developed a feature-based neural network rating system and compared against an expert and a clique-based approach. Our preliminary results, based only on a few important features, suggest that the feature-based neural network approach can be used as an alternative to an expert. The clique-based approach, on the other hand, may be used if there is a clique for the target user. The clique-based approach cannot be used for new movies and for users without a clique. So, the feature-based approach may provide an advantage if a new movie (or product) needs to be selectively targeted for the customers.

Acknowledgement

We are indebted to Will Hill of Bellcore for allowing us to use his movie list and ratings by Internet users. We are also thankful to Thomas Judd of Bellcore for populating the database with features from the Microsoft Cinemania CD-ROM.

References

[1] Dent, L., Boticario, J., McDermott, J., Mitchell, T. M., and Zabowski, D., "A Personal Learning Apprentice", In *Proc. of the National Conference on Artificial Intelligence (AAAI-92)*, August, 1992.

[2] Dumais, S. T., Furnas, G. W. and Landauer, T. K., "Using Latent Semantic Analysis to Improve Access to Textual Information", *Proc. of the 1988 ACM SIGCHI Conference*, pp. 281-295.

[3] Gifford, W. and Turock, D. "The Electronic Receptionist A Knowledge Based Approach to Personal Communications", *Proc. of the 1992 Computer Human Interactions Conference*, Monterey, CA, Sep. 1992.

[4] Harman, D. and Candela, G., "Retrieving Records from a Gigabyte of Text on a Minicomputer Using Statistical Ranking", *Journal of the ASIS*, Vol. 41, pp. 581-589, 1990.

[5] Obraczka, K., Danzig, P. and Li, S., "Internet Resource Discovery Services", *IEEE Computer*, Vol. 26, No. 9, pp. 8-22, Sep. 1993.

[6] Loeb, S., "Architecting Personalized Delivery of Multimedia Information", *Communications of the ACM*, Vol. 35, No. 12, pp. 39-48, Dec. 1992.