A Predictive Model for Video Lectures Classification

Priscylla Silva
Institute of Computing
Federal University of Alagoas
Alagoas, Brasil
pmss@ic.ufal.br

Roberth Pinheiro
Institute of Computing
Federal University of Alagoas
Alagoas, Brasil
rraraujop@gmail.com

Evandro Costa
Institute of Computing
Federal University of Alagoas
Alagoas, Brasil
evandro@ic.ufal.br

ABSTRACT

In the educational context, it is important to provide students with learning resources, such as tutorials, video lectures, and educational games to help their learning process, especially when they are not in the school and have difficulties or doubts. In these situations, recommendation systems may be used to suggest learning resources for students, avoiding, for instance, the task of making the manual process of searching and selecting resources. Most generic recommendation systems for video lectures use viewing history of the user to make recommendations of videos, which are in according of the user interests. In the educational context, other factors must be considered, the video should not only be of interest to the student, as it will be used as a learning resource for the main purpose of helping the student to learn a particular subject or clarify doubts. Hence, in this paper we evaluated three classifiers and propose a predictive model to classify video lectures according to their quality. We applied machine-learning algorithms on a set of video lectures by classified students according to some quality requirements. We conducted an experiment and preliminary results indicate good quality of the selected prediction model.

1. INTRODUCTION

The search for relevant information on the Web is a well known problem that has been addressed by several studies in the literature. Recommendation systems have been considered one important way to address this problem. In the educational context, recommendation systems are used to recommend learning resources for students, saving the students the time by manual process of searching and selecting resources. One of the most used resources by students are the video lectures. Websites like Youtube[1] and Vimeo[2] have many videos within various themes, including video lectures of various topics.

[1]http://www.youtube.com/
[2]http://vimeo.com/

The amount of videos on the Internet is growing at an explosive rate [1], making it harder the search for good and appropriate video lectures. When students have to learn a subject or they are in doubt, they generally perform the following steps: 1. Search on a website of videos using keywords from the subject; 2. Choose one of the first videos ranked to watch; 3. If the video is not good enough for them, then they stop watching it and try another video. It may occur students selecting several bad videos followed, until they find a video that they classifies as good and watch the full video to learn what they need. This happens because many students have no way to predict the quality of video lecture selected.

Most generic recommendation systems for videos uses viewing history from previous users to suggest videos [3]. In the educational context, other factors must be considered, the video should not only be of interest to students, as it will be used as a learning resource for the main purpose of helping students to learn a particular subject or clarify doubts. In this paper we evaluated three classifiers to classify video lectures according to their quality. The classifiers Navie Bayes, SVM and C4.5 were used. The classifier with better performance compared to others was selected.

2. EXPERIMENT

The purpose of the experiment was to evaluate the predictive ability from classifiers towards video evaluation. For perform the experiment the Weka software was used [2].

To perform the experiment 120 video lectures from *YouTube* were collected in the mathematic domain. The video lectures belong to the following topics: logarithms, Cartesian plane, set theory, polynomial functions, geometric progression and matrices. A total of 15 undergraduate students volunteered to evaluate videos. Each video lecture was assessed by 5 students who had attended courses that have mathematical knowledge as a prerequisite, such as calculus and linear algebra. Students evaluated the video lectures by applying grades from 0 to 10, median grade from this evaliation was used to generate the overall assessment of video lecture. Students were instructed to review the video lectures on the following criteria: clarity, teaching method, depth in the proposed issue, audio quality and image, teacher's didactic, among others. A video lecture is composed of the following attributes: title size, description size, duration, date of publication, view count, like count, dislike count and comment count. The attributes were normalized and then discretized.

All attributes have been discretized using histogram analysis. Each video has a class called "evaluation" that can take the labels: inadequate, bad, average, good, excellent.

The experiment was performed using 10-fold stratified cross-validation. This procedure divides the sample into k mutually exclusive parts (folds), for each step, $k-1$ folds are used for training and the induced hypothesis is tested on the remaining fold. In order to get statistically meaningful results, the number of iterations used was 10. In case of 10-fold cross-validation this means 100 calls of one classifier with training data and tested against test data. The current experiment performs 10 runs of 10-fold stratified cross-validation on the dataset using Navie Bayes, SVM and C4.5 scheme, this means 300 calls. The experiment consists in to confirm if the video lectures were automatically labeled correctly (in the sense of assigning a evaluation).

2.1 Evaluation Metrics

Given an algorithm A and a set of instances denominated T, assume that T is divided into k partitions. In the case of 10-fold cross-validation, $k = 10$. For each partition i, the hypothesis h_i is induced and the error denoted by $err(h_i)$, where $i = \{1, 2, ..., k\}$ is calculated. The mean, variance and standard deviation for all partitions are calculated using the following formulas: i) $mean(A) = mean(A, T) = \frac{1}{k}\sum_{i=1}^{k} err(h_i)$; ii) $var(A) = var(A, T) = \frac{1}{k}\left[\frac{1}{k-1}\sum_{i=1}^{k}(err(h_i) - mean(A, T))^2\right]$; iii) $sd(A) = sd(A, T) = \sqrt{var(A, T)}$.

When comparing two inductors in the same domain T, the standard deviation can be seen as a picture of the robustness of the algorithm: if the errors (calculated on different test sets) derived from induced hypotheses using different training sets are very different from one experiment to another, this indicates that the inductor is not robust to changes in the training set, coming from the same distribution. To compare two machine learning algorithms and decide which one is better (with confidence level of 80%), just take the general case to determine whether the difference between two algorithms (A_i and A_j) is significant or not, assuming a distribution normal. For this, the mean and standard deviation combinations are calculated according to the following equations: i) $mean(A_i - A_j) = mean(A_i) - mean(A_j)$; ii) $sd(A_i - A_j) = \sqrt{\frac{sd(A_i)^2 + sd(A_j)^2}{2}}$; iii) $ad(A_i - A_j) = \frac{mean(A_i - A_j)}{sd(A_i - A_j)}$. The absolute difference (ad) is given in standard deviations.

If $ad(A_i - A_j) > 0$ then A_j overcomes A_i and if $ad(A_i - A_j) >= 1.29$ then A_j overcomes A_i with 80% degree of confidence.

If $ad(A_i - A_j) <= 0$ then A_i overcomes A_j and if $ad(A_i - A_j) <= -1.29$ then A_i overcomes A_j with 80% degree of confidence.

3. RESULTS AND DISCUSSION

In the results of the experiment, the classifier that showed the best performance was the SVM. Table 1 shows the results of the comparative analysis of used classifiers.

In the comparison between Navie Bayes and SVM, where A_i = Navie Bayes and A_j = SVM, we have $ad(A_i - A_j) > 0$ and $ad(A_i - A_j) > 1.29$, therefore the SVM outperforms Navie Bayes with confidence level of 80%. In the comparison between SVM and C4.5, where A_i = SVM and A_j = C4.5, we have $ad(A_i - A_j) < 0$, therefore the SVM outperforms C4.5, but does not overcome the level of confidence of 80%, because $ad(A_i - A_j) > -1.29$. In the comparison between C4.5 and Navie Bayes, where A_i = C4.5 and A_j = Navie Bayes, we have $ad(A_i - A_j) > 0$, therefore the Navie Bayes outperforms C4.5, but does not overcome the level of confidence of 80%, because $ad(A_i - A_j) < 1.29$.

Table 1: The comparative analysis of used classifiers.

A_i	Navie Bayes	SVM	C4.5
A_j	SVM	C4.5	Navie Bayes
mean	0,08	-20,09	20,09
sd	0,05	56,96	56,96
ad	1,41	-0,35	0,35

Although we have achieved good results with our experiments, we verified three treats to validity of our work: i) The small number of volunteers (15) for evaluate the video lectures; ii) The limited domain and limited dataset; iii) The limited number of attributes - in our work we used only nine attibutes. We pretend to perform new experiments increasing the number of attributes, such as: analysis of the subtitles, audio and image quality, type of lesson (theoretical or problem solving), resources used in the video lecture (blackboard, slides or pen and paper), among others.

4. CONCLUSION AND FUTURE WORK

In this work we present an analysis of classifiers to predict the quality of video lectures. We conducted experiments with the classifiers: Navie Bayes, SVM and C4.5. The classifier that showed the best performance was the SVM, that was selected as a predictive model. We conducted an experiment and preliminary results indicate good quality of the SVM as prediction model. In our future work, we will conduct experiments with more users and videos. The analysis performed in this paper is part of an initial work to build a predictive model to determine the quality of video lectures. We plan to improve the prediction model with other factors such as context, viewing history, audio quality and relationships between user can be used to provide better results. In the future, the authors plan to integrate this predictive model in a recommendation system of video lectures.

5. REFERENCES

[1] S. Boll. Multitube–where web 2.0 and multimedia could meet. *IEEE MultiMedia*, 14(1):9–13, Jan. 2007.

[2] R. R. Bouckaert, E. Frank, M. A. Hall, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. WEKA–experiences with a java open-source project. *Journal of Machine Learning Research*, 11:2533–2541, 2010.

[3] X. Zhao, H. Luan, J. Cai, J. Yuan, X. Chen, and Z. Li. Personalized video recommendation based on viewing history with the study on youtube. In *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, ICIMCS '12, pages 161–165, New York, NY, USA, 2012. ACM.