Electronic Commerce Research and Applications 21 (2017) 38–49

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Predicting long-term product ratings based on few early ratings and user base analysis

Sushant S. Khopkar *, Alexander G. Nikolaev

Department of Industrial and Systems Engineering, State University of New York at Buffalo, Buffalo, NY 14260, United States

ARTICLE INFO

Article history:
Received 2 February 2016
Received in revised form 7 December 2016
Accepted 21 December 2016
Available online 28 December 2016

Keywords:
Machine learning
Bayesian network modeling
Product rating prediction

ABSTRACT

Product reviews and ratings are major quality indicators in online shopping systems. In making the decision about carrying a product as part of their inventory, an online seller often pays attention to the product's rating. However, when the observed average is based on a small number of individual user-submitted ratings, the decision-maker may not feel as confident about the product, even when the average is high. The long-term average rating predictions can help online retailers to identify products to promote on their websites as "top picks". The paper proposes a Bayesian Network model to predict the long-term average product ratings based on a (limited) number of early submitted ratings. Performance of the proposed model is compared with the performance of five other prediction methods/models; a Linear Regression model, two variations of a Running Average predictor, an Ordered Logistic Regression model and a Confirmatory Factor Analysis model. Each model's performance is evaluated using the "MovieLens" dataset (GroupLensResearch, 2012). The training is done on 56,590 data points, the ratings submitted for 1155 movies, and the prediction results are reported for 495 movies. It is demonstrated that the proposed Bayesian Network and the Linear Regression models are particularly effective in making accurate predictions around the time of product introduction when the information about the prospective, future ratings is especially valuable. The Bayesian Network model performs better in the very early stage of feedback collection, and in the later stages, as more feedback is received, the Linear Regression model emerges as the best predictor. The strengths and weaknesses of all the assessed prediction methods are discussed.

1. Introduction

With the spread and growth of online businesses such as amazon.com and ebay.com, online shopping has become a major form of shopping over the past decade. One of the primary reasons behind the preference for online shopping is that people can buy products without physically visiting stores, which reduces the efforts required for shopping, saves time and allows for quick research about product quality. While shopping online, customers cannot evaluate a product in person, "by the feel". Hence, they often resort to online peer reviews: the textual feedback and reported product ratings people share on retailers' web pages and forum threads. Easy access to reviews and ratings is not limited to products; through this mechanism consumers help each other decide what movies to watch, which restaurants and attractions to visit, and so on. A statistical study (digitalvisitor.com, 2012) revealed that after price considerations, consumers pay the highest attention to the available reviews and ratings in making their buying decisions. Also, 24% of users access online reviews before paying for a service delivered offline (Zhu and Zhang, 2010).

The availability of accurate tools for predicting long-term average product ratings will benefit online retailers to decide which products to promote on their websites as "top picks". It can also help them to forecast demand of a product based on its predicted rating and manage their inventory accordingly. This paper addresses the problem of predicting long-term average ratings of products, with the intent to better inform online retailers of product quality based on early, limited user feedback. In doing so, it relies on available quantitative information about multiple products rated by the retailer's user base. Note that people's buying decisions may be affected not only by numerical product ratings, but also by textual feedback/comments.

These long-term average product ratings can also help consumers in their buying decisions. For any product with its rating information available, two quantities can be observed: (1) the average rating the product has received so far, and (2) the number of peers who have contributed to this rating. If a particular online product has been rated by many users, then the decision-maker is

* Corresponding author.
E-mail addresses: skhopkar@buffalo.edu (S.S. Khopkar), anikolae@buffalo.edu (A. G. Nikolaev).

http://dx.doi.org/10.1016/j.elerap.2016.12.002
1567-4223/

likely to trust the information, i.e., rely on it in making her decisions. On the other hand, if a product has been rated by only a few users, then the decision-maker may not feel as confident. For example, when a customer observes an average rating of 4.5 out of 5 for a product while observing that this product has been rated by 47 people, then she feels confident about the quality of the product. She thinks that the "real quality" of the product indeed evaluates to 4.5 out of 5 (on her personal scale, whatever that might be), and hence, she will be comfortable buying this product if its price is in her budget. However, if the same person observes a 5.0 out of 5 rating for another product, while the product has been rated by only two people, then she might believe the rating is subjective and not sufficiently trustworthy (see Fig. 1). Note that the described chain of thoughts can be justified by simple statistics. The more ratings are averaged, the more likely the average is to be close to the true mean of the population of ratings, i.e., the combined opinion of all the potential voters. Hence, long-term average rating prediction is also beneficial to a buyer. The phenomenon of information cascading (Bikhchandani et al., 1992) might also be responsible for this "feeling confident" about a rating when a large of number ratings have already been submitted. However, no such assumption is necessary for the prediction methods considered in this paper. Note that the knowledge of the long-term average product rating may not be as useful to a consumer as it is to a retailer. Indeed, an *average* rating does not capture individual buyer's interests/preferences, while it helps establish an overall interest in the product from all the buyers. As an example, consider the current top rated movie on IBDb website (www.imdb.com), "The Shawshank Redemption", having the average rating of 9.3/10 (rated by 1,734,646 users). This movie comes under the genre of drama. This means that if a person does not prefer the drama genre, then this high average rating might not indicate how she in particular would be impressed by this movie. On the other hand, for a product such as vacuum cleaner, where its performance or dependability is more important than other subjective attributes such as color and design, the determination of long-term average ratings would be more useful to consumers. Meanwhile, in any case, the average product rating is useful to a product distributor to help predict further sales. In summary, the knowledge of the long-term average rating of a product is useful to online retailers, and *may be* useful to consumers.

The challenge of having to use limited information about a product while making judgment of its quality can be addressed by predicting its long-term average rating based on the information about the user base of a specific website. This paper proposes a Bayesian Network model to address the question "Can one predict the long-term average rating for a product (i.e., its rating when a large number of people will have voted on it) that has currently received only a few ratings, using the information of the past ratings given to other products by the same user base?" It also compares the proposed model with five other prediction methods/models.

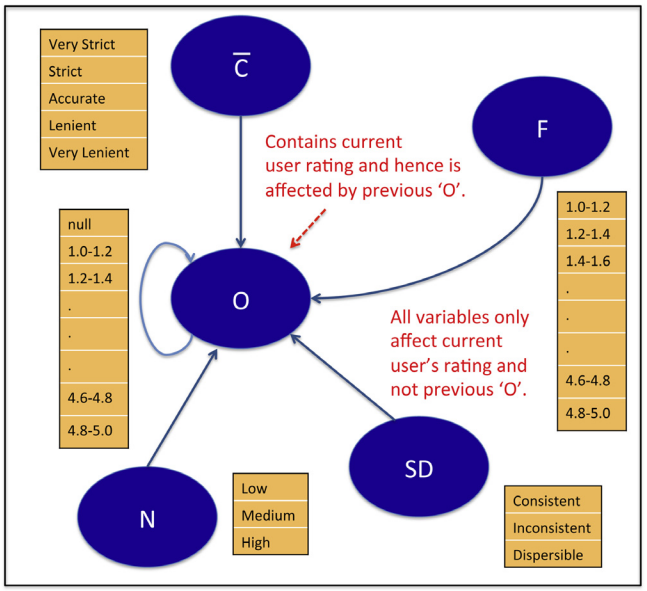

Fig. 2. The predictive bayesian network model variables.

2. Related work

Few prior research efforts pursued similar objectives. Among those that did is a recent effort of forecasting product adoption that relies on learning the user interaction network characteristics in building projected product diffusion curves based on the observed diffusion trajectories of other products (Trusov et al., 2013). Dover et al. (2012) study the differences in adoption rates of diffusion processes to make inferences about degree distributions of underlying networks, which in turn help to predict product diffusions. The proposed Bayesian Network model in the present paper, however, do not require any knowledge about the users' social network. A prominent research direction relevant to the presented research is recommender system development, which has now become an integral part of many e-commerce sites including amazon.com and ebay.com (Schafer et al., 1999). The relevance comes

Fig. 1. The number of ratings combined with the average rating of a product determines users' perspectives towards it.