Enhancing Online Auction Transaction Likelihood: 
A Comprehensive Data Mining Approach

Lei Chen, Jianghan University, Hubei, China 
Yanbin Tu, Robert Morris University, Pittsburgh, USA

ABSTRACT

This article compares four data mining models (discriminant analysis, logistic regression, decision tree, and multilayer neural networks) for online auction transaction predictions. It aims to choose the best model in terms of prediction accuracy and to identify determinants significant for auction transactions. By using datasets from eBay, the authors find that the best data mining model for auction transactions is multilayer neural networks. Logistic regression and decision tree models can be used to identify determinants significant for auction transaction such as seller’s feedback profile, listing picture, listing files size, return policies, and others. By adjusting these listing options, sellers could increase the auction transaction likelihood. This study will help sellers improve their auction listings by constructing effective selling strategies so that they can enhance the likelihood of online auction transactions. All these efforts will help improve their online auction performances and finally lead to a more efficient electronic marketplace.

KEYWORDS
Auction Outcome Prediction, Auction Price, Auction Success, Data Mining, eBay, Online Auction, Selling Strategies

INTRODUCTION

Online auctions are attracting more sellers and buyers for their low search costs, high interaction potential, and world-wide accessibility (Bakos 1997). Sellers and buyers can achieve significant cost savings and enhance their surplus through online auction transactions (Eric van and Peter 1998). Online auctions become an important component of electronic marketplaces. According to the second quarter financial report in 2015, eBay, the Internet auction house leader, had a record high of 157 million active users, over a 200% increase compared to the second quarter of 2005. In the same quarter, eBay had new listings totaling a record of 800 million at any given time, which is 40% higher than that in the second quarter of 2005.

With fast growing electronic marketplaces, there are still various challenges for sellers to succeed and grow. For example, individual sellers may find that it is increasingly difficult for them to survive in an electronic marketplace when competing among businesses (Cohen 2002). Moreover, compared to experienced sellers, new sellers suffer many disadvantages in terms of auction performance and outcome. While there are abundant studies on buyers’ behavior or bidding strategies (Bapna et al. 2004, Chen, Chen and Song 2002, and Livingston 2003), estimating the success chance of auction listings and constructing effective selling strategies have not been fully studied. Houser and Wooders (2006),
Dholakia (2005), Ba and Pavlou (2002), and Ba, Ow and Wood (2008) investigate the role of seller reputation in online auctions. Anderson et al. (2007), Becherer and Halstead (2004), Lucking-Reiley et al. (2007), Wolk and Spann (2008), and Joo, Mazumdar and Raj (2012) evaluate the effectiveness of different selling options and explore seller’s choices in online auctions. However, these studies mainly focus on individual selling options and use traditional data analytical tools such as structure equation models and linear regression models.

As more data becomes available online, data mining becomes an important tool to explore the uncovered knowledge and intelligence underlying a dataset (Hand, Mannila and Smyth 2001, Tan, Steinbach and Kumar 2005, Ali 2012a, Ali 2012b). Data mining techniques can be used to discover hidden patterns for business intelligence, minimize the impact of human biases on data analysis, and provide robust analysis without distribution assumptions (Fayyad, Pietetsky-Shapiro and Smyth 1996, Al-Janabi 2017, Al-Janabi and Esraa Alwan 2017, Al-Janabi et al. 2018). Applications of data mining have emerged in different business areas. Roussinov and Zhao (2003) use web mining techniques to detect similarity relationships. Bapna et al. (2004) use K-means clustering to classify bidders’ behavior. In particular, data mining techniques are widely used in customer relationship management (CRM) and other marketing fields (Albert, Goes, and Gupta 2004, Padmanabhan and Tuzhilin 2003, Eunju, Wooju and Yillbyung 2003, Shaw et al. 2001, and YongSeog and Street 2004). For the applications of data mining in online auction, Tu (2009) introduces an application of Web-based data mining in selling strategies for online auctions and illustrates the process of using spider to collect online data from eBay, and the application of classification and regression tree (CART) in constructing effective selling strategies. However, these studies do not apply multiple data mining tools to address compressively auction transaction and the associated selling strategies. There are a variety of data mining models that can be used to estimate auction transactions and to build selling strategies.

In this study, using real auction data from eBay, we compare four data mining techniques for predicting auction transactions. Auction transactions are defined as a deal achieved by sellers and buyers. We will identify the best data mining technique for auction transaction predictions. We will also discuss the most significant determinants for auction transactions. This study contributes to the literature by designing a practical method that helps sellers list their auctions in an effective way. Rather than applying one data mining technique, we combine four as a combo to boost auction listing success.

This study has significant implications to practitioners. By using multiple appropriate data mining techniques, sellers can measure the effectiveness of their proposed auction listings. The determinants identified could help sellers improve auction listings if needed so that the chance of auction transactions can be significantly increased. All of these efforts will help increase the efficiency of the online auction marketplace.

**BRIEF INTRODUCTION TO DATA MINING MODELS**

There are various analysis classifications in data mining literature. We choose four data mining models in this study: 1) Discriminant Analysis (DA), 2) Logistic Regression (LR), 3) Decision Tree (DT), and 4) Multilayer Neural Networks (MNN). These four represent both parametric and non-parametric methods. For example, logistic regression, discriminant analysis and neural networks are parametric, and CART model is non-parametric. We could include more non-parametric methods in the combo. However, the entire algorithm might need more data to support, and it might take a long time to implement. As a practical method for business, the simplest, most efficient, and effective data mining method is highly expected by eBay sellers. As an auction listing can have one of the two possible outcomes: transaction (success) or non-transaction (non-success) after the listing time ends, these four techniques are appropriate tools for predicting this kind of binary category variable with reasonable data demand and fast speed.
www.igi-global.com/article/xbrl-taxonomy-estimating-effects-greenhouse/53840?camid=4v1a