Predicting eBay Auction End-prices using Auction Dynamics
Boyang Qu
05/22/2019

Abstract

Along with the rapid development of the internet and e-commerce in today's world, online auctions have become one of the most popular modes of transactions. Having 94 million active members, both sellers and buyers, eBay is one of the largest online auction platforms. eBay auctions provide a large dataset, and it has been widely explored in different areas to generate different insights and applications. While most past research treats data from online auctions as cross-sectional and disregards the dynamics happening during an auction, this study focuses on the two dynamics in predicting the end-price: the dynamics of the current auction and other auctions. In this paper, I collected historical auction data from eBay and used multiclass classification approaches, which are machine learning algorithms, to predict the end-prices of auction items. I described the feature used, especially those capturing the dynamics of bidding history and recent auctions on similar items. I showed that by evaluating different classification methods, I was able to accurately predict the end-prices of eBay auctions, which can be useful for buyers and sellers to set prices and place bids in online marketplaces.

1. Data

eBay auctions are second-price auctions with proxy biddings, in which bidders submit a bid and the system automatically updates the bidding process to show bidding history and highest current price. The winner with the highest bid pays the second highest bid plus an increment set by the seller. eBay auctions have strict ending times, and data on closed auctions are available publicly on the website.

In this paper, a dataset containing 628 auctions ended in December 2012 and January 2013, including 136 auctions of Cartier wristwatch, 343 auctions of palm pilot M515 PDA, and 149 auctions of Xbox game console, is employed. The ending prices of these auctions extend over a range from $26 to $5400, which provides a wide range for exploration and thus a generalized prediction on ending prices. The average closing prices of auctions in the three categories are $884.55, $229.08, and $131.41 respectively. Furthermore, the average number of bids placed in an auction varies for different items. On average, 14.36 bids were placed in a Cartier wristwatch auction, 17.25 bids were placed in a palm pilot M515 PDA auction, and 18.87 bids were placed in an Xbox game console auction. The ratings of bidders in an auction have high standard deviations and are similar among auctions of the three items.

This dataset was obtained using a web crawler and is publicly available at http://www.modelingonlineauctions.com/datasets. For each auction in the dataset, auction id, a bidding history including the actual bid, the bidders, the rates of the bidders, and the bidding time, the opening bid, the ending-price, the item name were obtained from eBay website.

1.2 Bidding History

Figure 1. shows a scatterplot of the bidding history of a typical auction. The bids arrive at unevenly spaced time intervals and different time points during an auction. Figure 2. shows the scatterplot of all bids aggregated over all 628 auctions. The aggregated graph indicates that the bids arriving at the beginning and ending hours of the auctions are denser than the bids arriving at the time between. This is a typical feature of eBay auctions with a strict ending time. An auction is often found to contain three parts: an early part with some bidding activity, a middle part with very little bidding, and a final part with intense bidding (Shmueli et al).

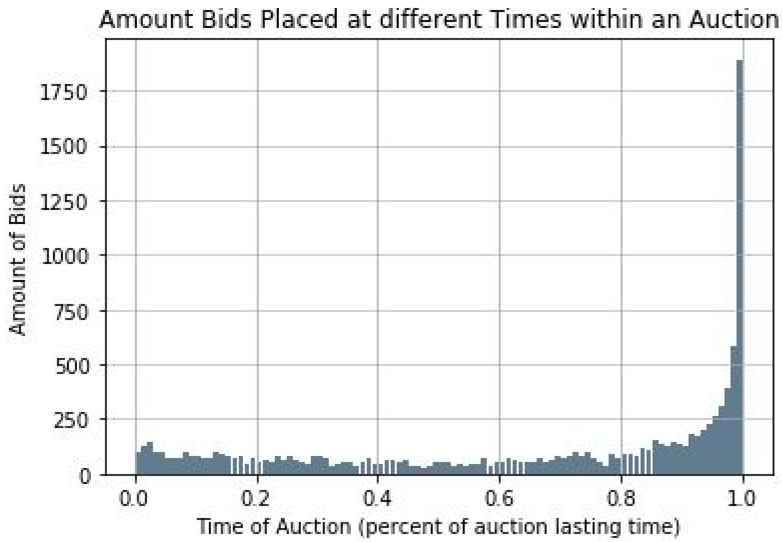

Figure 1. Amount of bids placed versus time of the bid during an auction for 632 auctions

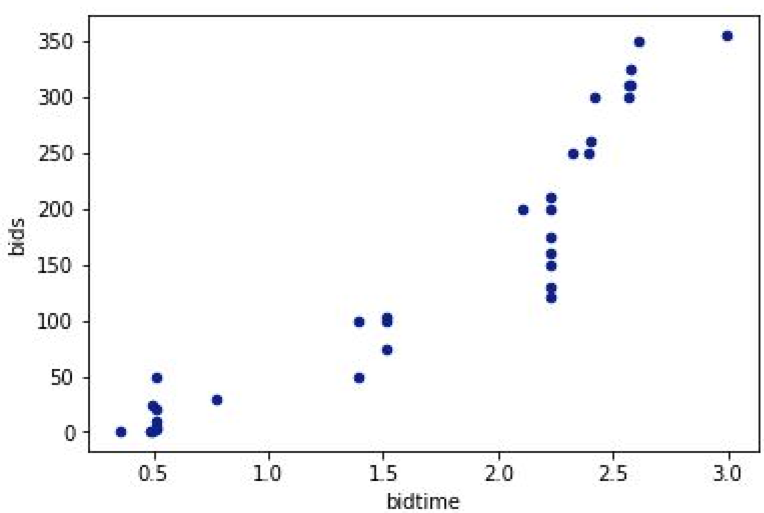

Figure 2. Bids placed in auction 1639453840 for Cartier wristwatch

1.2 Historical Auctions of similar Items

Figure 3. shows the end-price of auctions of items within each of the three categories. It is noticed that the end price of palm pilot M515 PDA and Xbox game console are relatively low and close to each other, and that of Cartier wristwatch is higher and has more oscillation. Overall, the end-prices of auctions within same category appears to be relevant.

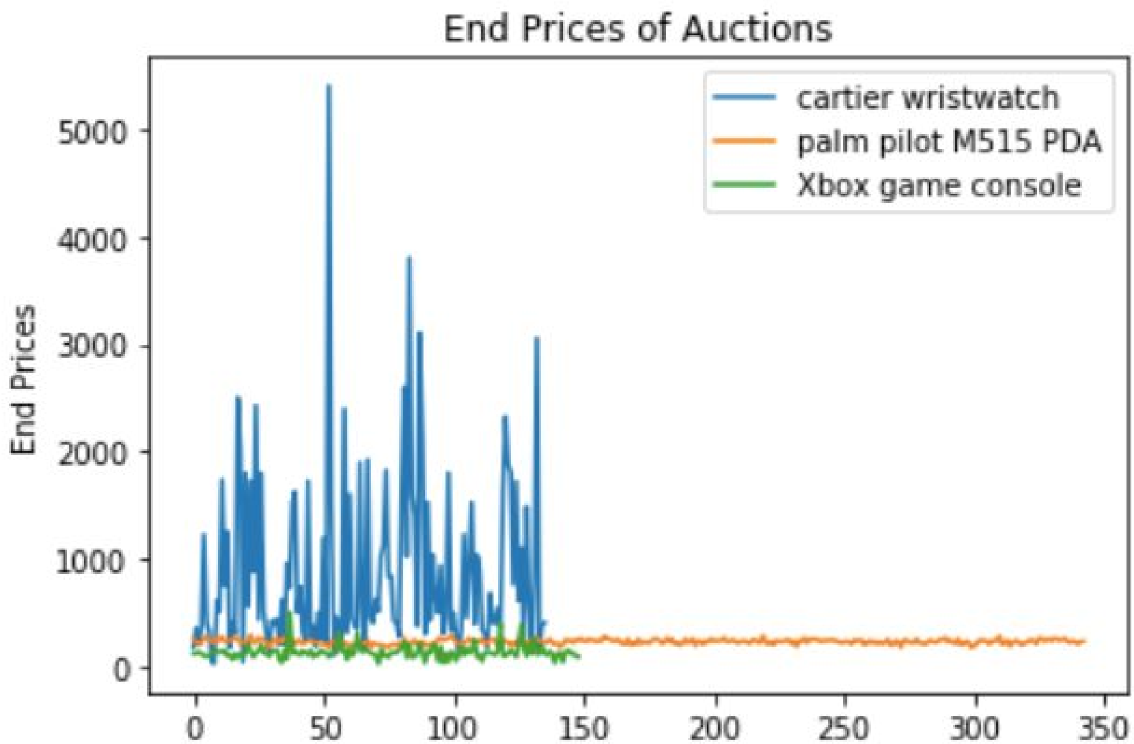

Figure 3. End-price of auctions of different items

2. Methods

2.1 Chebyshev Polynomials for Bidding History

In order to capture the dynamics in bidding history, functional regression is employed in this study. Functional regression analysis is similar to conventional regression analysis in that it treats a response variable as a predictor, but functional regression operates on functional objects, which can be a curve, a shape, or any object. In this study, Chebyshev polynomial is used to convert the bidding history containing the magnitude and time of bids into a continuous function. The continuous curve describes the price evolution happening during an auction and thus does a great job in capturing the dynamics bids of an auction.

Figure 3. shows a continuous curve recovered from the bidding history data using Chebyshev polynomial estimation (a 33-degree polynomial was used in this study), as well as the actual bidding history. The graph indicates that the Chebyshev polynomial estimation yields an almost perfect estimation for the bidding history, and thus is able to represent the dynamics greatly.

Furthermore, as noticed in the scatterplots of bidding history (Figure 1 and Figure 2), early and snapping bids, bids placed near the end of an auction, are common in eBay auctions, and they play an important role in the ending prices and bids of other bidders. I treated bids placed within the first half day and last hour of an auction as early and late bids. The presence of early and snapping bids are recorded as a binary variable, 1 for presence and 0 for absence.

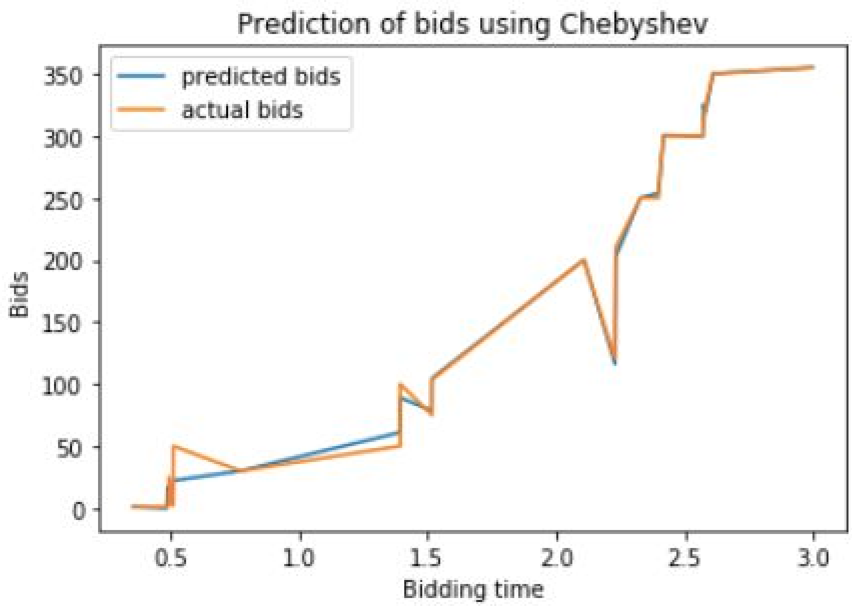

Figure 4. Continuous curve recovered from bidding history of auction 1639453840 for Cartier wristwatch using Chebyshev polynomial

2.2 Dynamics of similar-item auctions at the current time

While Chebyshev polynomial estimation perfectly captures the dynamics of the current auction, the dynamic of auctions of similar items also needs to be addressed. For each auction, the end-price of the last recently ended auction of a similar item is recorded, and the average end-price of all closed auctions on the similar items are also calculated. Since there are not many auctions for each item category and the prices of each item is fixed in the official provider websites or Amazon, a recently closed price and an average closing price are sufficient to capture the dynamics of other auctions.

2.3 Estimating the End-prices of eBay Auctions with Multiclass Classification

The end-prices are discretized into $34 intervals, which is approximately 10% of the average ending price of all auctions, $347.86. The goal is to predict the price within a 10% window of the average price. The end-price of each auction falls into one of these categories, and thus the price prediction problem can then be treated as a multiclass classification problem. The end-price output is a $34 range instead of a specific price as in conventional regression, and this has been proven to be more accurate than the regression approach (Ghani and Simmons). Decision tree classifier, supporter vector machine (SVM) classifier, and k-nearest neighbors (KNN) classifier are used to implement the multiclass classification approach.

3. Results

Auctions within the dataset are selected randomly for training and testing purposes. 125 auctions, 20% of all auctions are selected for testing purpose, and 507 auctions are selected for training the model.

The results in Table 1. shows that decision tree classifier is significantly better at predicting the end-price of auctions. The decision tree classifier indicates predicting the end-price of auctions using dynamics of the current auction and other auctions yield highly accurate results, considering the fact that there are only limited auctions in the dataset.

Classifier	Accuracy
Decision Tree	0.92
SVM	0.224
KNN	0.024

Table 1. Prediction Accuracy of different Classifiers

References

Shmueli, G., Russo, R. P., and Jank, W. (in press), “The Barista: A Model for Bid Arrivals in Online Auctions,” *Annals of Applied Statistics.*

Ghani, Tayid and Simmon, Hillery. “Predicting the End-Price of Online Auctions, ” *Accenture Technology Labs*.