Does Dynamics Happening during an auction play an Important Role in Predicting the Ending Price?

Boyang Qu
06/12/2019

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. eBay, one of the largest online auction platforms, has been widely explored in different areas to generate different insights and applications. However, most past research treats data from online auctions as cross-sectional and disregards the dynamics happening during an auction. This study focuses on two dynamics in predicting the ending price: the dynamics of bidding in a given auction and the dynamics of other auctions. This paper then explores that does dynamics happening during an ongoing auction play an important role in predicting the ending price? This paper addresses this question by using a multi-class classification model with parameters capturing these two dynamics as inputs, and this model predicts the ending prices within price intervals. The dynamics of bidding history in a given auction is extracted using Chebyshev polynomial estimation, and dynamics of other auctions is represented using ending prices of previously ended auctions. The results show that a decision tree classifier is able to predict the ending price with high accuracy. Since the input variables only represent the auction dynamics, the high accuracy proves that dynamics happening during an ongoing auction are important in predicting the ending price.

Keywords: eBay auctions, predicting model, auction dynamics.

1. Introduction

Online auctions have become increasingly popular in recent years, and eBay, one of the largest online auction platforms, provides a large dataset for empirical research in this area. With millions of items dispersed across thousands of categories on eBay's auction platform, 135.5 million registered users, and 95 million active users, eBay is popular worldwide. Research on eBay auction ties closely to everyday life.

Many research on eBay auctions or other online auctions has been published in economics, marking, and information systems literature in the past. However, I find that most past research ignores one of the most important features of online auctions: the dynamics involved in the bidding process of an ongoing auction. Though many papers ignore the evolving dynamics and treat the data from online auctions as cross-sectional, it is useful to consult their approaches. For instance, Ghani and Simmons (2004) evaluate the possibility and accuracy of using three machine learning algorithms, simple regressions, multi-class classification, and multiple binary classification approach, to predict the ending price of given auctions with features known prior to the start of the auctions. They find that the discrete price prediction (predict $5 ending price interval) using multi-class classification approaches outperforms simple continuous price prediction using the regression approach.

Although many papers ignore the dynamics in the bidding process, which lead to a great loss of information, there are some recent papers step forward and address the dynamics in online auctions. I focus on two major auction dynamics mentioned in past literature: the dynamics of bidding in the current auction and that of closing prices of recent auctions on similar items. To the best of my knowledge, past research has only focused on one dynamics and has not examined the importance of these two dynamics in predicting the ending price as a whole. This paper can contribute to this missing area. The goal of this research is to explore that do the two dynamics happening during an ongoing auction together play an important role in predicting the ending price.

Past research employs different approaches in addressing the two dynamics, and these approaches inspire the model used in this paper. Wang et al. (2012) develop a forecasting system addressing the dynamics of bidding in a current auction. They recover the bidding history of each auction to a functional object, which is a curve that models the price velocity and acceleration, and then use functional object regression as the core of their forecasting model predicting the ending price of an ongoing auction. This model significantly outperforms standard forecasting methods. Li et al. (2014) employ a different approach using attribute construction, which transforms the one-to-many relationship between auction and bids to a one-to-one relationship taking in the consideration of the bidding times and magnitudes. The data transformation process enables traditional machine learning algorithms and statistical analysis methods to be used to predict the closing prices, and they conclude that incorporating the dynamics improves the accuracy of prediction. Lim et al (2008) compare the accuracy of artificial neural network and the grey system approach in predicting closing prices of auctions, and they show that including data capturing the dynamics and closing prices of ongoing auctions of similar items can yield higher accuracy rate. Schapire et al. (2002) propose a boosting-based conditional density estimation approach to estimate the ending price based on past auction results and information on the current auction. This approach is useful in estimating the ending price of an auction if auctions of multiple interacting goods are held simultaneously.

To explore the importance of the two dynamics happening during an ongoing auction in predicting the ending price, in this article I develop a forecasting model using parameters capturing both dynamics and use an eBay auction dataset on Cartier wristwatch, Palm pilot M515 PDA, and Xbox game console. First, inspired by functional object regression in Wang et al (2012), I utilize Chebyshev polynomial estimation in recovering the bidding history of an auction into a curve, and the coefficients in the polynomial estimation are used to represent the dynamics, such as acceleration and timing space, in bidding history. Some other parameters including number of bids, number of bidders, and number of early and snapping bids are also incorporated to further capture the bidding history dynamics. Then to capture the dynamics in other auctions of similar items, ending prices of last recently ended auction of a similar item, as well as average ending price of last ten recently ended auctions of similar items, are used. Lastly, with the variables mentioned to capture the two dynamics, discrete price prediction following the idea of Ghani and Simmons (2014) is employed. Several classifiers including decision tree, support vector machine, and k-nearest neighbors, are used for multi-class classification discretizing the ending prices into price intervals. The accuracy of each classifier in predicting the ending prices is explored.

The results show that the decision tree classifier is the most accurate (0.90 for Palm pilot M515 PDA, 0.78 for Xbox game console, and 0.66 for Cartier wristwatch) in predicting the ending price interval of an auction within 10% of the actual ending price. Since only parameters representing the dynamics of bidding history in a given auction and the dynamics of previously ended auctions of similar items are used, the high accuracy proves the importance of the two dynamics together in predicting the ending price of an auction.

The appeal of the forecasting model in this research is that it captures both dynamics of bidding history in a given auction and that of other auctions, which, to the best of my knowledge, no previous paper has done. Furthermore, with more information about an auction and other ongoing auctions, this model can be further developed into a dynamic forecasting model using both the dynamic information and static information of an auction known in the beginning of an auction to predict the end-price of an ongoing auction, instead of an ended auction. This future direction can be enormously useful for eBay sellers and users.

2. Theory

Based on previous literature, parameters capturing two dynamics are used in the model: dynamics of bidding history of an ongoing auction and dynamics of other recently ended auctions of similar items.

First, in order to capture the dynamics in bidding history, functional regression is employed in this study. This idea follows the functional object regression model in Wang et al (2012), while I choose a different approach to recover the functional object from the bidding history of a given auction. Chebyshev polynomial is used to convert the bidding history containing the magnitude and timing of bids into a continuous function in this study. The estimated curve using Chebyshev polynomial describes the price evolution happening during an auction and is able to capture the dynamics of bidding in auctions. 8-degree Chebyshev polynomial estimation is used, which hopefully does not lead to overfitting problem and captures the dynamics well:

$$y_i = c_0 + c_1x_i + c_2x_i^2 + c_3x_i^3 + c_4x_i^4 + c_5x_i^5 + c_6x_i^6 + c_7x_i^7 + c_8x_i^8,$$

where y_i is an actual bid and x_i is the corresponding bidding time. The coefficients of the estimated curve c_0, c_1, …, c_8 are used as independent variables. Additionally, variables including number of bids (n_{bid}), number of bidders (n_{bidder}), and number of early and snapping bids (n_{early} and n_{late}) are also incorporated to further capture the bidding history dynamics.

The next step is to generate parameters capturing the dynamics in other auctions of similar items. The ending prices of a last recently ended auction of a similar item p_{i-1}, as well as average ending price of last ten recently ended auctions of similar items p_{ave} are used as independent variables.

The final ending price prediction step uses discrete price prediction based on the idea in Ghani and Simmons (2014). They claim that discrete price prediction is much more accurate than continuous price prediction. For discrete price prediction, the ending prices are discretized into price intervals less than 10% of the average ending prices, and multiclass-classification approach is used. Using independent input variables c_0, c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, n_{bid}, n_{bidder}, n_{early}, n_{late}, p_{i-1}, and p_{ave} and a dependent variable of actual ending price p_i, the accuracies of three classifiers, decision tree, support vector machine, and k-nearest neighbors, are examined.

3. Data

This paper uses an eBay auction dataset on Cartier wristwatch, Palm pilot M515 PDA, and Xbox game console. The dataset was created by Wolfgang Jank, Galit Shmueli, and Shanshan Wang using a web crawler for their book (2010) and it is publicly available at http://www.modelingonlineauctions.com/datasets. Observations for a given auction in the dataset includes auction id (the unique identifier of an auction), opening bid, ending-price, item name, and bidding history including the actual bid, the bidder, the eBay feedback ratings of the bidder, and the bidding time (number of days since the start of the auction) were available. According to Jank and Shmueli's book using this dataset, the auctions within each item category were ordered by their ending time (2010).

The eBay dataset contains 628 auctions ended between December 2012 and January 2013. Within the 628 auctions, there are 136 7-day auctions of Cartier wristwatch, 343 7-day auctions of palm pilot M515 PDA, and 149 7-day auctions of Xbox game console. The ending prices of these auctions extend over a range from $26 to $5400, which provides a wide price range for exploration. Endings prices of Cartier wristwatch is most unstable with oscillation over a great range, and ending prices of M515 PDA is the most stable with low standard deviation (Fig.1 and Table 1). By using this dataset, I am able to examine the model on a broad price range and diverse auction categories, and thus can reach a generalized conclusion.

This paper seeks to explore the importance of dynamics of bidding history and of auctions of similar items in predicting the ending price, and thus the bidding history variables are employed to represent the bidding history dynamics, and the ending prices of previously ended auctions of same-category items are employed to represent the dynamics of other auctions. In the dataset, the bidder rate variable is missed in many bidders (with value zero), and the opening prices of many auctions are set to 0.1 or 1, which are conventional values in eBay auctions irrelevant to the item price, and thus these two variables are not used in the model in this paper.

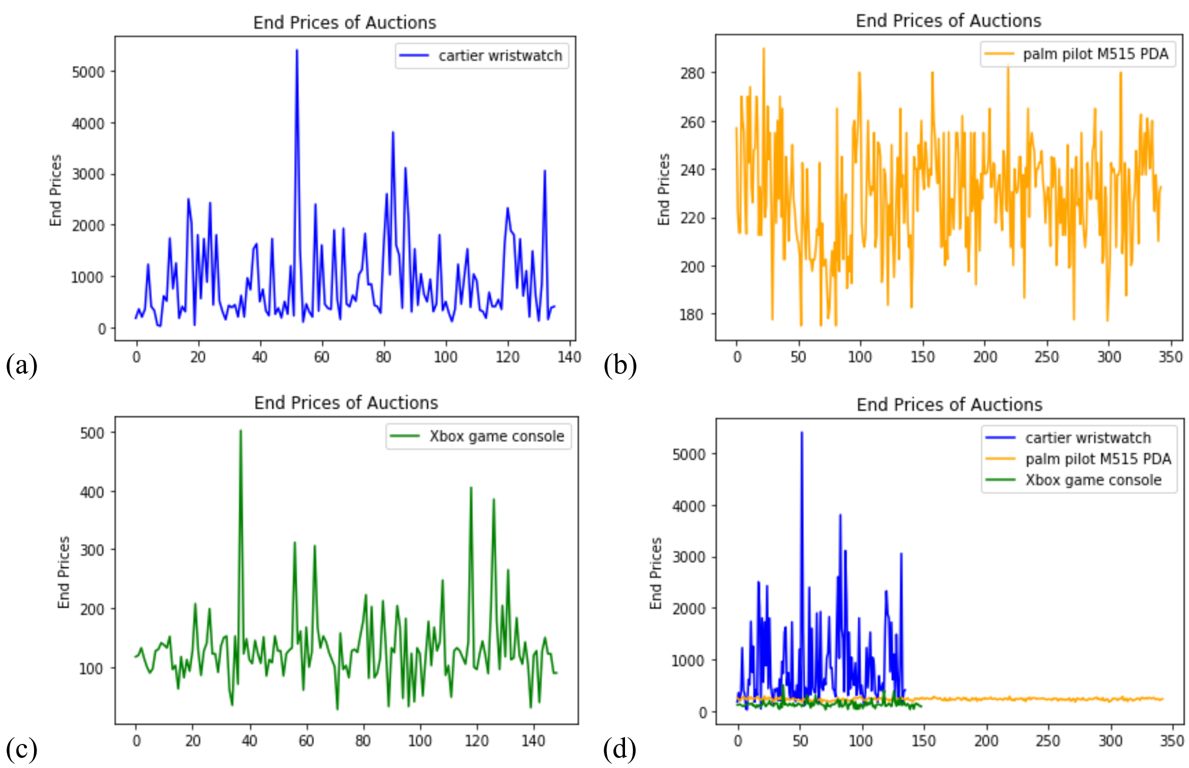

Figure 1. Ending prices of eBay auctions. (a) Cartier Wristwatch auction. (b) Palm pilot M515 PDA auctions. (c) Xbox game console auctions. (d) All auctions.

Table 1. Mean, Standard Deviation, and Variance of ending prices of Auctions within each category, in dollars.

	Mean	**Standard Deviation**	**Variance**

Cartier Wristwatch	884.5574	832.4350	692948.0658
Palm Pilot M515 PDA	229.0836	21.9660	482.5052
Xbox Game Console	131.4140	63.5348	4036.6696
All	347.8601	481.3986	231744.6432

3.1. Bidding History

In the bidding history of a given item, there are three important aspects: the number of bids placed, the number of bidders, and timing and acceleration of bids. According to Table 1, the number of bids placed and bidders in each auction vary a lot both among auctions within each category and among auction categories. The diversity of eBay auctions represents diverse bidding dynamics.

Bids in an auction arrive in unevenly spaced time intervals and different time points, as shown in Figure 2. Moreover, the scatterplot of bids in a single auction suggests that there are more bids arrive early and late in an auction, an aggregate bidding history over 628 auctions in Figure 3 further confirms this phenomenon. Bids arriving during the beginning and ending hours of the auctions, approximately first and last 0.02 of auction lasting time, 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, 2007). The number of bids and bidders, the number of early and late bids in the first and last three hours, together with parameters capturing the acceleration and timing in bidding history, are then used in the model to capture the dynamics of bidding history.

Table 2. Mean, Standard Deviation, and Variance of bids placed and bidders in a given auction in each category.

	Number of Bids			**Number of Bidders**		
	Average	**Standard Deviation**	**Variance**	**Average**	**Standard Deviation**	**Variance**
Cartier Wristwatch	14.4	8.9	79.8	6.8	3.8	14.8
Palm Pilot M515 PDA	17.2	11.1	123.8	8.8	5.0	25.4
Xbox Game Console	18.9	11.5	132.9	8.3	13.75	14.1
All	17.0	10.9	118.5	8.2	20.9	4.6

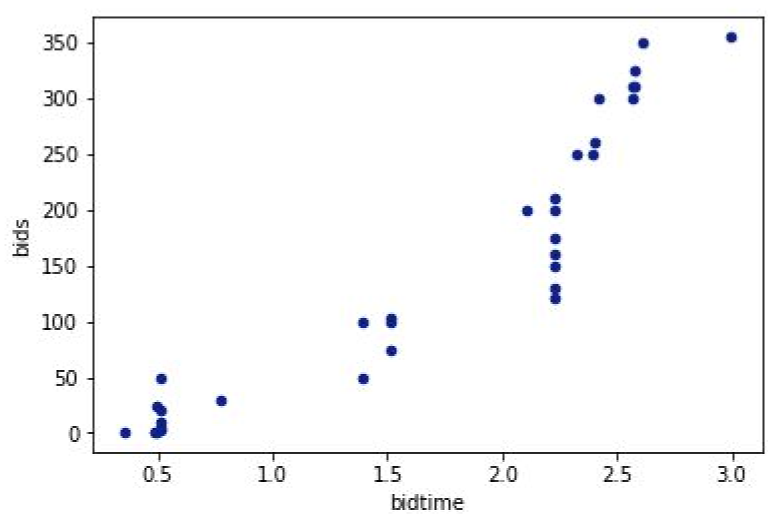

Figure 2. Bids placed versus bidding time in auction 1639453840 for Cartier wristwatch.

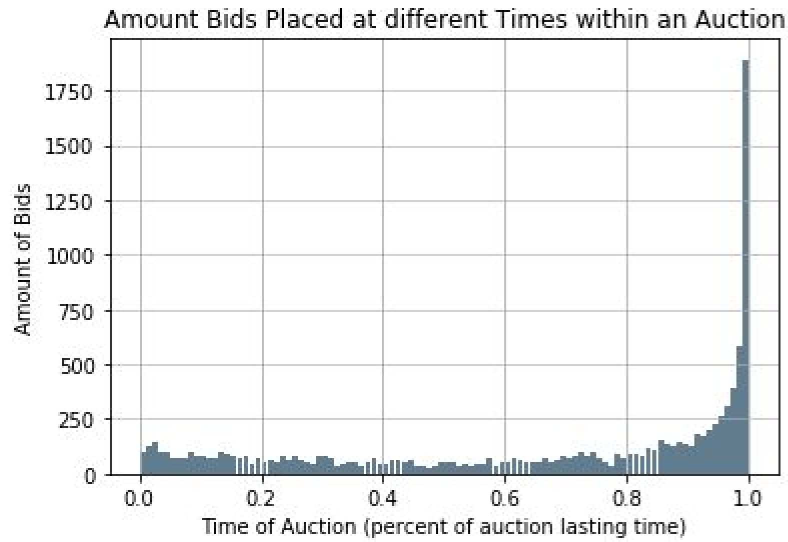

Figure 3. Amount of bids placed versus bidding time for 632 auctions.

4. Methods and Results

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

4.1. Capturing Bidding Dynamics

Using Chebyshev polynomial estimation, a continuous curve is recovered from the bidding history. An 8-degree Chebyshev polynomial is used in this study, which is able to capture bidding dynamics while likely does not lead to overfitting problem. Figure 4. shows a continuous curve recovered from the bidding history data using Chebyshev polynomial estimation, as well as the actual bidding history. The graph indicates that the Chebyshev polynomial estimation yields a good estimation for the bidding history, and thus is able to capture the dynamics greatly.

Coefficients from the Chebyshev polynomial estimation, number of early and snapping bids within the first and last three hours, and number of bidders for a given auction is used to represent the bidding dynamics.

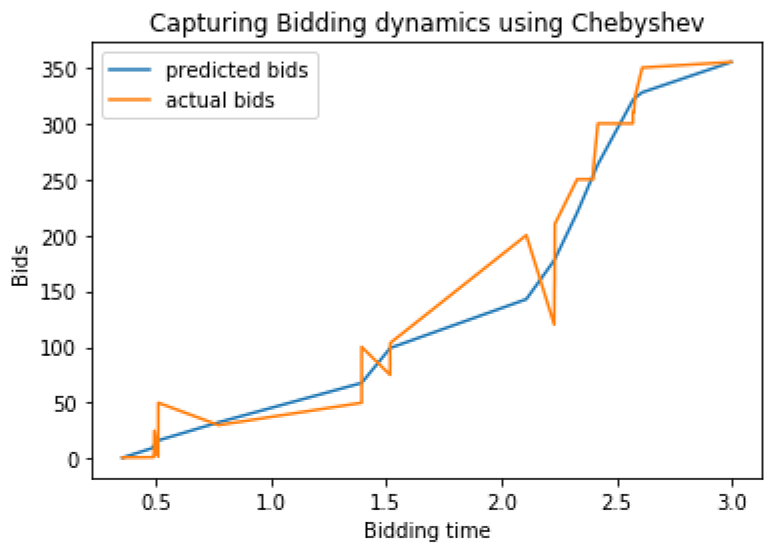

Figure 4. Continuous curve recovered from bidding history of auction 1639453840 for Cartier wristwatch using Chebyshev polynomial

4.2. Capturing dynamics of similar-item auctions

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 last ten recently closed auctions on the similar items is also calculated. While the closing prices of auctions within each category oscillate much, the oscillation are more likely to be associated with the used state of the items. Since the prices of each item, unlike unique art pieces, are provided 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.

4.3. Multiclass Classification

The end-prices are discretized into price intervals, and three different classifiers, decision tree classifier, supporter vector machine (SVM) classifier, and k-nearest neighbors (KNN) classifier, are used to implement multiclass classification that predicts the ending price interval. The price intervals are in $35, $80, $20, and $10 for all auctions, Cartier Wristwatch auctions, Palm pilot M515 PDA auctions, and Xbox game console auctions respectively, which is less than 10% of the average prices of auctions within each category. The goal is to predict the price within a 10% window of the average price.

The results in Table 1. shows that decision tree classifier is significantly better at predicting the ending prices of auctions in all categories than SVM and KNN classifiers. KNN classifier seeks to find similar neighbors, and the fact that it does not perform well implies that there are no significant differences in bidding history and previously ended auctions' prices that can divide auctions into groups associated with each ending price interval. However, there should be differences that can separate auctions into groups that are associated with the ending prices, but not as many groups as the price interval groups. The SVM classifier seeks to identify a hyperplane that separates auctions into classes, but parameters in auction predictions are very close or even overlay, and thus the line is hard to find. The decision tree, which is most accurate, classifies auctions using multiple steps according to different criteria associated with the auction parameters, and the multi-step process works best in this paper. The accuracy of predicting the ending price of Palm pilot M515 PDA using decision tree classifier is the highest, and that of predicting the ending price of Cartier wristwatch is the lowest. Recall that the ending prices of Cartier wristwatch is most unstable and that of Palm pilot M515 PDA is most stable. A stable ending price suggests that ending prices of previously ended auctions of similar items can be better used to predict the ending price of a given auction.

Considering that there are only limited auctions in the dataset, the decision tree classifier is able to predict the ending price interval of auctions using dynamics of the current auction and other auctions with relatively highly accuracy.

Table 3. Average Accuracy of different Classifiers from 10-trials.

	Decision Tree	**SVM**	**KNN**
Cartier Wristwatch	0.66	0.15	0.02
Palm Pilot M515 PDA	0.90	0.08	0.514
Xbox Game Console	0.78	0.14	0.32
All	0.27	0.09	0.17

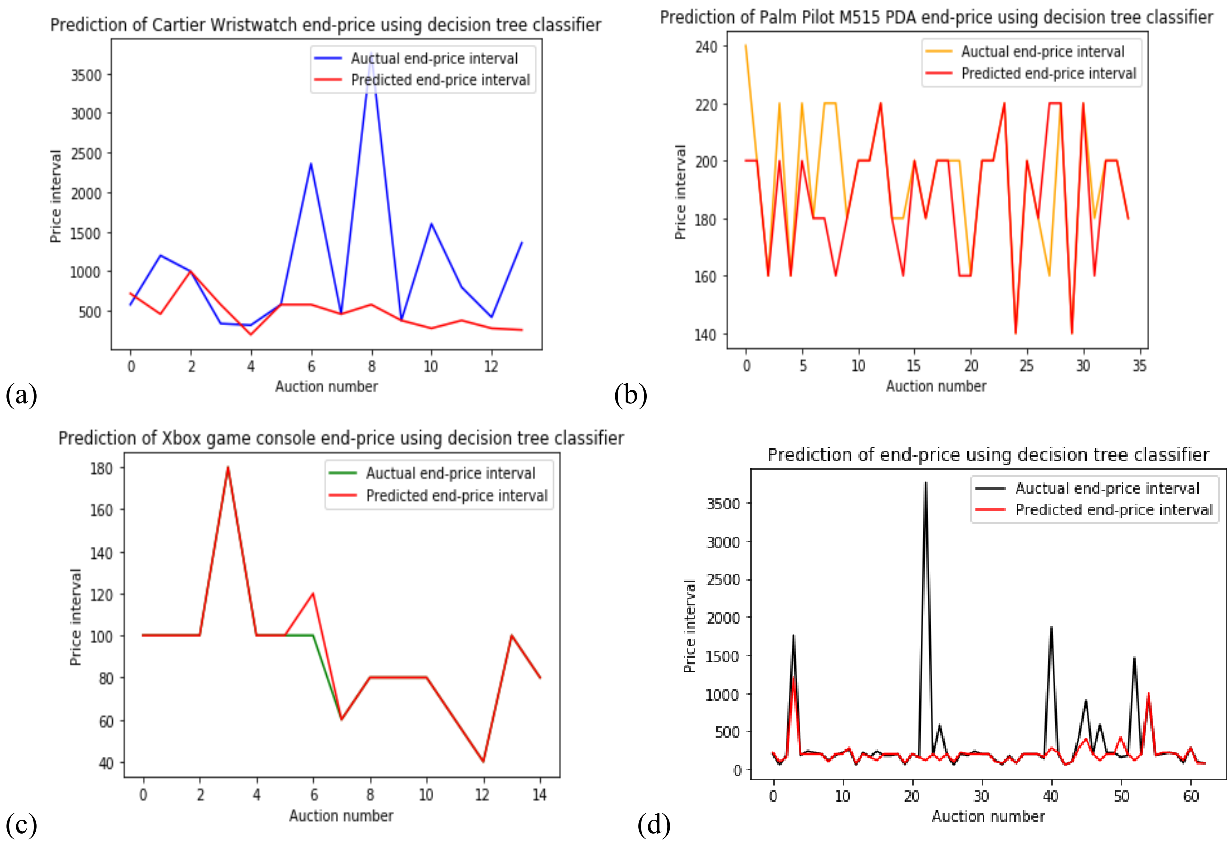

Figure 5. Actual Ending Price versus Predicted Ending Price using Decision Tree Classifiers. (a) Cartier Wristwatch ending prices. (b) Palm Pilot M515 PDA ending prices. (c) Xbox game console ending prices. (d) All auctions ending prices.

5. Conclusion

From above, I find that using parameter from Chebyshev polynomial estimation to capture dynamics in bidding history of a given auction and using ending price of recently ended auctions of similar items to capture the dynamics of other auctions as input, a decision tree classifier can predict the ending price of an auction with relatively high accuracy. Because only parameters representing auction dynamics are included, the result convincingly suggests that dynamics in an auction is important in predicting the ending prices. Furthermore, the lower accuracies using KNN and SVM classifiers indicate that there is no huge difference in bidding dynamics that grouped auctions into similar neighbors or generate a hyperplane separating auctions to predict the ending prices, but instead, minor differences in auction dynamics can be used to predict the ending price using a decision tree classifier. The accuracy differences among auction categories also indicate that ending price of auctions for items with a drastically oscillated price range as Cartier wristwatch can be harder to predict, while those for items with a relatively high oscillated price range as Xbox game console and Palm pilot M515 PDA can produce relatively high accuracy.

There is one important limitation in this study for capturing dynamics in other auctions of similar items. Using only the ending prices of a recently ended auctions of similar item may be insufficient to capture dynamics in other auctions. For any given auction, there are also other ongoing auctions of similar items. To better capture the dynamics of these ongoing auctions, which do not have ending prices, bidding dynamics of them can also be incorporated together with the ending prices of recently ended auction of similar items. However, since it is only known that auctions in the dataset are ordered by ending time and the exact time of each bid is not known, the progress of other ongoing auctions (time point within other auctions) cannot be decided using the dataset, and thus this cannot be implemented in this study. With dataset containing more information, this can be implemented.

The model in this paper can be further developed into a dynamic forecasting model that is able to predict the ending price of an ongoing auction dynamically. Using dataset with more information, both the dynamic information and static information of an auction can be used as input variables in the dynamic forecasting model, and the predicted ending price of an ongoing auction is hopefully relatively accurate, which can be enormously beneficial for eBay users and sellers.

References

1. Ghani, R. and Simmons, H. (2004). Predicting the End-Price of Online Auctions. In the Proceedings of the International Workshop on Data Mining and Adaptive Modelling Methods for Economics and Management

2. Lim, D., Anthony, P., Ho, C.M. (2008). Agent for Predicting Online Auction Closing Price in a Simulated Auction Environment. PRICAI 2008: Trends in Artificial Intelligence, vol 5351. Springer, Berlin, Heidelberg

3. Wang, S., Jank, W., and Shmueli, G. (2012). Explaining and Forecasting Online Auction Prices and Their Dynamics Using Functional Data Analysis. Journal of Business & Economic Statistics, 26: 144-160.

4. Li, X., Liu, L., Wu, L., and Zhang, Z. (2005). Predicting the final prices of online auction items. Expert Systems with Applications, 31: 542-550.

5. Schapire, R., Stone, P., McAllester, D., and Csirik, J.A. (2002). Modeling Auction Price Uncertainty Using Boosting-based Conditional Density Estimation. Machine Learning: Proceedings of the Nineteenth International Conference, 2002.

6. Shmueli, G., Russo, R. P., and Jank, W. (2007), “The Barista: A Model for Bid Arrivals in Online Auctions,” *Annals of Applied Stati*stics.

7. Wolfgang Jank and Galt Shmueli (2010). “Modeling Online Auctions, ” *Wiley and Sons*.