Prediction Online Auction Price Using Functional Data Analysis

May Phyo Wai Lin, Khine Khine Oo
University of Computer Studies, Yangon
hikari.phyo@gmail.com, k2khine@gmail.com

Abstract

The goal of the proposed system is to derive models for forecasting the final price of ongoing auction. The forecasting task is important not only to the participants of an auction who compete against each other for the lowest price, but also to designers of bidder side. Forecasting price in online auctions is challenging from statistical point-of-view because traditional forecasting model do not apply. The reason for this are three typical feature of online auction data: (1) unequally spaced bid; (2) the limited time of an auction; (3) the dynamic of bidding change significantly over time. The key feature of our model is that it operates during the live –auction. Auction data typically arrive as a sequence of bids over a period of time. Taking a functional data analysis approach, the system acts the bid from a single auction as recognition from a continuous price.

Keywords: Functional Data Analysis (FDA), Dynamic Forecasting Model and predicts the price.

1. Introduction

Electronic commerce has created a lot of public interest in recent year. The system contains opening bid that is positively related to on-line auction price levels at the beginning of the auction, but its effect toward the end of the auction. The order in which the lots appear in an auction is negatively related to the current price level, with this relationship decreasing toward the end of the auction.

The online auction focuses on tools like summary statistics and more formal statistical method such as regression models. Predicting an auction's outcome is also important for the seller who may have the option to sell item before the auction is over. Online auction data have particular features that make traditional forecasting methods hard or even impossible to apply but the functional data analysis allows for an estimation of the price dynamics. Online auction data typically arrive as a sequence of bids placed over time.

Proposed system develops forecasting models to predict the final price of online auctions. Prediction the final price is interesting, especially to the participants of an auction who compete against each other for the lowest price.

Feedback mechanism is a popular feature of online auctions and can decrease the informational asymmetries between buyers and sellers.

2. Related Works

Some problem use Past Methodologies to create model (Wang, Jank and Shmueli (2004))[2] used 185 completed auctions of the item Palm M515 Personal Digital Assistant (PDA) with a duration of 7 days and collected all data about the bid histories. It is observed that most bids came either at the beginning or at the end of the auction. That focused on the development of a forecasting model in order to estimate the final price for specific items such as the Palm M515 for the same auction duration. Data is collected and put into context in order to make statistical analysis and subsequently to forecast the final price of the specific item.

(Ghani, 2005)[6]who suggest offering sellers an insurance that gurantees a minimum selleing price. Inorder to do so, it is important to correctly forecast the price, at least on average. While Ghani's method is static in nature, our dynamic forecasting approach could potentially allow more flexible fetures which would allow sellers to purchase an insurance either at the beginning of the auction, or during the live auction (coupled with a time-varying premium).

A further bidder whose entrance in the auction is at a late point of time less influences the velocity in the price enlargement. Bapna, Wolfgang and Shmueli, (2004)[3] also compared the US and European auctions and observed that US auctions have a 4 per cent higher price in the first half of the auction and subsequent price are equivalent. In addition, it is concluded that the main differences between US and European auctions occur in the middle of the auction. Although the US auctions have a bigger price at the start and the end of the auction duration, price increases in European auctions are faster in the middle part of the auction duration. An additional significant result obtained from the research is that the price level is negatively related to the auction duration even when the seller is low-rated.

3. Theory Background

3.1 Ascending Auction

The price is successively raised until only one bidder remains, and that bidder wins the object at the final price. This auction can be run by having the seller announce prices or by having the bidders call out prices themselves, or by having the prices rises continuously while bidders gradually quit the auction.

Ascending auctions provide a process of price discovery. Value is socially determined through the escalation of bids. Rarely does a bidder enter an auction with fixed values for the items being sold. Rather the bidders learn from each other's bidding, adjusting valuations throughout the process. This process is especially important when resale is a possibility or more generally when others have information relevant to assessing the item's value. This open competition gives ascending auctions a legitimacy that is not shared by other auctions.

3.2 Functional Data Analysis

Methodological and applied research related to the analysis of functional data is currently receiving a tremendous amount of interest in the statistics literature. Functional data analysis (FDA) is a tool set that, although based on the ideas of classical statistics, differs from it (and, in a sense, generalizes it), especially with respect to the type of data structures that it encompasses. While the underlying ideas for FDA have been around for a longer time, the surge in associated research can be attributed to the monographs. In FDA, the interest centers on a set of curves, shapes, images, or, more generally, a set of *functional objects*.

There are a number of recent studies devoted to the generalization of standard statistical methodology to the context of functional observations. For instance, develops a measure of centrality for a given functional observation within a group of curves. A principal component approach for a set of sparsely-sampled curves is developed. Other exploratory tools have been developed such as curve clustering and curve-classification. Classical statistical methods have also been generalized to functional canonical correlation analysis, functional ANOVA, functional regression, and functional generalized linear models. Differential equation models are fitted to data of functional form. This list is only a small part of the current methodological efforts in this emerging field.

Functional data analysis is a collection of techniques in statistics for the analysis of curves or functions. Most FDA techniques assume that the curves have been observed at all time points but in practice this is rarely the case. In some instances, curves may not be observed over all time periods.

In figure 1 describes the curves may only be observed over discrete intervals. Since we have many observations for each curve we first use a simple smoothing spline approach to generate a continuous smooth curve from our discrete, observation.

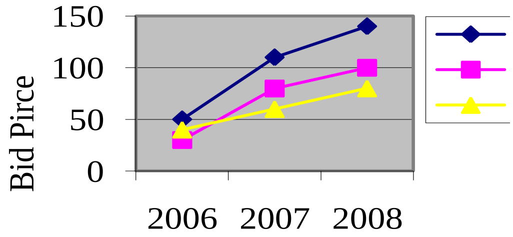

Figure 1 Estimation of Adoption of Item

3.3 Definition of Static and Dynamic Forecasting

The impact of proposed tax law changes on revenue collections is done initially with a *static* analysis. A change such as a tax rate reduction or an increase in the tax base would be applied to the existing baseline revenue forecast. For example, if a 10% reduction in individual income tax rates were to be enacted, the static approach would estimate that future tax collections would be reduced by 10% from the baseline forecast. Since this approach does not account for potential changes in economic growth or in taxpayer behavior that may be induced by tax law changes, static impact calculations may not estimate their full effects accurately.

Interest in *dynamic* methods increased during the 1990s. Dynamic forecasting goes beyond the static approach and attempts to predict changes in the economy brought about by changes in fiscal policies. In turn, the amount of revenue gained or lost from the subsequent changes in economic activity results in a “dynamic” estimate, which should better reflect the true impact of the policy change.

In order to estimate dynamic effects, it is necessary to develop a model that can produce statistical estimates of taxpayer responses to policy changes. The static estimates are put into the model, which is then used to calculate the additional impact as the economy responds to the changes. For example, an income tax rate reduction would generate added disposable household income. The dynamic model would forecast the level of added revenues resulting from this higher disposable income.

3.4 Data used in this Study

The data in this study are the bid histories of closed 3-years auctions for the Jade Log. The bids can arrive at unevenly spaced time intervals. While the number of incoming bids is sparse during some periods of the 3-year auction, it can be very dense other times such as at the very beginning of the auction and during the ending period.

In this work we focus on developing forecasting methods for auctions of a specific product. Having the capacity to forecast an auction for a specific product allows, e.g. the bidder, to focus her bidding-efforts on a select sub-sample from a potentially large population of auctions for the identical good.

Forecasting auctions for a specific product also allows for an exact measure of the forecasting-accuracy since the product value is relatively well-known. And finally, the lessons learned from this task can be used to derive more complex forecasting models for auctions of diverse analysis, functional ANOVA, functional regression, and functional generalized linear models. Differential equation models are fitted product types.

3.5 Forecasting Auction Dynamic

Firstly, the system must develop a model to estimate and forecast the auction dynamics.

$$D^{(m)}y_t = \alpha + \beta t$$

where,

$D^{(m)}y_t$ = the m^{th} times of the price-position y_t at time t.

α = bid price during auctioning

β = time experience about auction

$$\alpha = \bar{e} - \beta\,\bar{a}$$

$$\beta = \Sigma^{s}_{i=1} \left((a_i - \bar{a})(e - \bar{e}) \right) / \Sigma^{s}_{i=1} (a_i - \bar{a})^2$$

where,

$\bar{e}$ = average bid price for t time.

$\bar{a}$ = average time experience for time t.

3.6 Create the Vector of Price Dynamic

$$l(t) = y_{t-1} + y_{t-2} + \ldots + y_{t-q}$$

Where,

y_{t-q} = the first q lags of y_t

3.7 Create the Dynamic Forecasting Model

$$y_t = d(t) / l(t) * 100\%$$

where,

$d(t) = (D^{(1)} y_t, D^{(2)} y_t, \ldots D^{(p)} y_t)$ is the vector of price dynamics.

$D^{(p)}y_t$ = the vector of the first *p* derivatives of *y* at time *t*.

3.8 Forecast Accuracy

Due to figure 2, we measure forecast accuracy on the validation set using the mean-absolute-percentage-error (MAPE). We compute the MAPE in two different ways, once between the forecasted curve and the true functional curve

(MAPE1), and then between the forecasted curve and the actual current auction price (MAPE2).

Naturally, MAPE2 is higher than MAPE1, because it is harder to reach the second level of "truth" compared with the first level. MAPE1 is, at least on average, less than +5% for the entire prediction period (i.e. over the last day), implying that our model has a very high forecasting accuracy. MAPE2 is a bit larger magnitude due to the inevitable variation in fitting smoothing splines to the observed data. The widths of the confidence bounds underline the heterogeneity across all auctions in our data set.

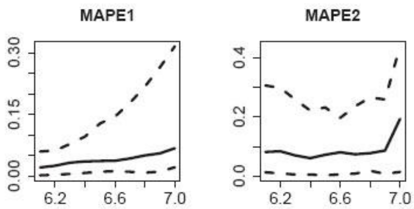

Figure 2 Mean Absolute Percentage Errors (MAPEs)

4 Implementation

Functional representation allows for measuring of process dynamics via derivatives. The calculation for the implementation of

$$b \qquad D^{(m)}y_t = \alpha + \beta t$$

When the user requested time stamp is 1-min we can calculate the value of t with 1. The value of t is depending on the user requested time stamp. The system is calculate the value of $D^{(m)}y_t$ after each 1 minute. After calculation each $D^{(m)}y_t$ he system will calculate total the last 2min sampled bid.

To calculate the value of d(t) by total calculation of $D^{(m)}y_t$. The calculation of d(t) is very important because it is the part of prediction online price.

After calculation total number of time at $D^{(m)}y_t$, next step is to calculate the price lap from $l(t) = y_{t-1} + y_{t-2} + \ldots + y_{t-q}$

Table 1 Sample data to calculate price dynamic model

Time experience (in sec)	Bid price
2	30
5	57
10	64
12	72
15	83
30	90
45	92
$\bar{a} = 17$	$\bar{e} = 69.71$

In table 1, these sample data are getting from ongoing online auction in specific time period. We can calculate by getting these data to create a dynamic model.

$$\beta = [(2-17)(30-69.71)+(5-17)(57-69.71)+\ldots+(45-17)(92-69.71)] / [(2-17)^2+(5-17)^2+\ldots+(45-17)^2]$$
$$= 1.17$$
$$\alpha = 69.71 - (1.17*17)$$
$$= 49.82$$
$$D^{(m)} y_t = 49.82 + 1.17t$$

Above the equation, $D^{(m)} y_t$ is the predicting price and t is the time (bidding time experience). So, $D^{(m)} y_t$ is a independent variable on time experience (t). If we substitute time experience in t, $D^{(m)} y_t$ value is obtained. The value of α and β are calculation for the derivative of bid price. We must compare predicting price with actual to get accuracy.

5 Predicitng Online Auction

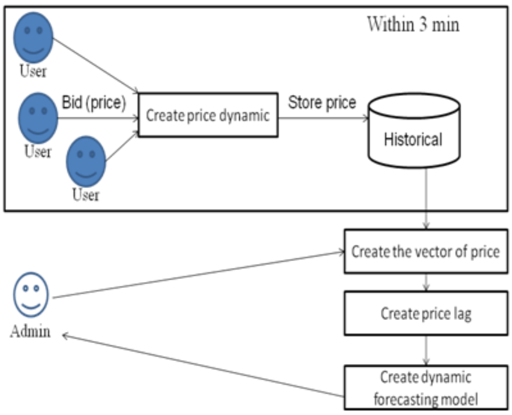

Figure 3 System Flow Diagram

System allow multiuser accessing nature to auction desire item. After bidding the system is automatically predict the bid price. After finishing all of predict calculation, try to compare the last price and test its accuracy.

In figure 3, multiuser is simultenously accessing the auction. After each 1 minute, the system is create the price dynamic model. The past time bid price are stored in the database as hitorical nature. The admin use can see the who is and how much bid price is live.

5.1 The overview of the proposed system

The prediction of continuous values can be modeled by statstical technique of regression. After the bid period time, the total calculation bid price is compared with the previous year bided price. It gives the result of comparisom remarks such as Good, Fair or Worst. And it also gives the result of curve from last year to this final bid year. You can really get the different type of different price form drop down list box in prediction session.

Some or more of the user are simultaneously log-in to this system for bid the various items. But the user is not the member who must register to be a member. After registeration, user as a member can access the application by its private user named and password. Specially all of members need to have a bank account. When a user log-in to the system, who can view various items's definition such as logId, Place where item is discovered, specification, floor price and its status such as Bidding or Not started.

In bidding process, the user cannot see with other bidding user information. But this page is described the floor price, current bid price, remanining time to be end, text box for submit price to be bid. After time up, it show 'Won' status to winner and 'Lose' status to loser.

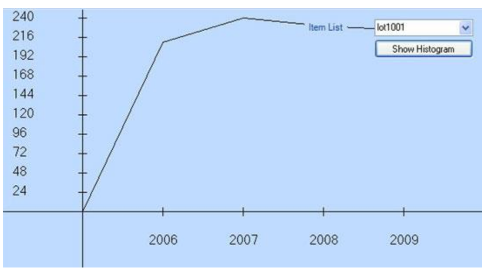

Figure 3 Historical Curves for Specific Item

In figure 3, First need to choose the item from the Item List box, and then tap the Show Histogram Box, the line graph of that item will be appear. In that line graph, show the year from the last three years and the price by the rate of accuracy. The nature of the curve will change by user requested item.

6. Conclusion

In this paper, we develop a dynamic forecasting model that operates during the live auction. Forecasting price in online auctions can have benefits to different auction parties. For instance, price forecasts can be used to dynamically score auctions for the same (or similar) item by their predicted price. Dynamic price soring can lead to a ranking of auctions with the highest expected price which, subsequently, can help bidder make decisions abut which auctions to prefer. On the other hand, auction forecasting can also be beneficial to the seller or the auctioner. Price forecasts can also be used by B2B that provide brokerage services to buyers or sellers.

The empiric is sufficient to conclude ascending auctions are superior to sealed-bid auctions. However, the case for ascending auctions is strong. The dynamic price discovery process of an ascending auction simply does a better job of answering the basic auction question: who should get the items and at what prices? Ascending auctions perform well on both efficiency and revenue rounds across a variety of settings. Two factors may favor sealed bidding – ex ante asymmetries and weak competition. In other cases, an ascending auction design is probably best. However, steps should be taken to limit the possibility of collusive outcomes in an ascending auction. This is accomplished by setting reserves, by imposing bid restrictions, and by limiting the information bidders receive during the auction.

Reference

[1] R. Bapna, P. Goes, and A. Gupta, "Analysis and design of business-to-consumer online auctions", *Management Science*, 49, 85-101, 2003.

[2] R. Bapna, W. Jank and G. Shmueli, "Consumer surplus in oline auctions", working paper, University of Connecticut, 2004.

[3] R. F. Easley and R. Tenorio, “Jump bidding strategies in internet auctions”, M*anagement Science*, 50(10), 1407-1419, October, 2004.

[4] Cuevas, M. Febrero and R. Fraiman, “Linear functional regression: The case of fixed design and functional response", *the Canadian Journal of Statistics*, 30, 285-300, 2002

[5] G. N. Allen, D. L. Burk and G. B. Davis, “Academic data collection in electronic environments: Defining acceptable use of internet resources”, *MIS Quarterly*, 30(3), 599-610, 12p, 1 chart (AN 21940319), Sep 2006.

[6] Ghani, R. (2005). Price prediction and insurance for online auctions. In the proceedings of the 11 th ACMSIGKDD International Conference on Knowledge Discover and Data Mining, Chicago, IL, 2005.

[7] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed.,New York: Wiley, 1984.