

Decision Support and Profit Prediction for Online Auction Sellers

Chia-Hui Chang
Computer Science & Information Engineering
National Central University
Jhongli, Taiwan
chia@csie.ncu.edu.tw

Jun-Hong Lin
Computer Science & Information Engineering
National Central University
Jhongli, Taiwan
csieflyman@gmail.com

ABSTRACT

Online auction has become a very popular e-commerce transaction type. The immense business opportunities attract a lot of individuals as well as online stores. With more sellers engaged in, the competition between sellers is more intense. For sellers, how to maximize their profit by proper auction setting becomes the critical success factor in online auction market. In this paper, we provide a selling recommendation service which can predict the expected profit before listing and, based on the expected profit, recommend the seller whether to use current auction setting or not. We collect data from five kinds of digital camera from eBay and apply machine learning algorithm to predict sold probability and end-price. In order to get genuine sold probability and end-price prediction (even for unsold items), we apply probability calibration and sample selection bias correction when building the prediction models. To decide whether to list a commodity or not, we apply cost-sensitive analysis to decide whether to use current auction setting. We compare the profits using three different approaches: probability-based, end-price based, and our expected-profit based recommendation service. The experiment result shows that our recommendation service based on expected profit gives higher earnings and probability is a key factor that maintains the profit gain when ultra cost incurs for unsold items due to stocking.

Keywords

Online auction, profit prediction, sample selection bias, probability calibration, expected profit

1. INTRODUCTION

Online auction has become a very popular e-commerce transaction type. With very low entering barrier, online auction has attracted a lot of individuals as well as small online stores. Though sellers enjoy the enormous business opportunities on online auction, they also faced a variety

of competitors on the same marketplace. How to maximize the profit has become the critical success factor in online auction.

One of the key questions that sellers encounter is how to list their commodities. For online auctions, sellers need to decide many auction settings like starting bid price, reserve price, duration time, and whether to use buy-it-now or advertising option, etc. Some sellers give a high starting price in order to get a high revenue, while some sellers lower the price and purchase advertising to increase sold probability. In fact, in the pursuit of maximized profit, the seller has to make a tradeoff between sold probability and revenue. How to find an auction setting that could maximize sellers' profit is a challenging problem. However, it is not easy to determine the best auction setting. Thus, we turn to an easier question: **Given an auction setting, should the current auction setting be used for the given item?** Furthermore, if there exists a service that could predict the expected profit, then we might apply such services to determine the best auction setting a commodity should use for a particular seller.

There are several researches on end-price (or closing price) prediction for online auction [1][3][6]. Ghani and Simmons apply three models, including regression, multi-class classification and multiple binary classification tasks, to predict auction end-price and show that the last model with neural network outperforms other approaches [1]. Heijst et al. incorporate textual information contained in the item description and ensemble decision trees using boosting algorithms for prediction [3]. They show that item description is more important predictor than sellers' feedback score and items' pictures. Wang et al. develop a dynamic forecasting model based on functional data analysis which can predict the end-price of an "in-progress" auction [6]. Such a service is beneficial for bidders to skip auctions items with high end-price and focus on others with potentially low price. However, for the decision support of commodity listing, dynamic forecasting is not necessary since sellers could not change the auction setting when an auction begins. Thus, Ghani and Simmons' methods are more suitable in this context.

A potential problem in Ghani and Simmons' work is that they use historical "sold items" for end-price prediction, which violates the assumption of most machine learning algorithms that examples for training are randomly drawn from the same distribution of as the test set. The problem, known as **sample selection bias** in econometrics is studied and formalized in Zadrozny's paper in [7]. In this paper, we consider our case as feature bias and apply the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

U'09 June 28, 2009, Paris, France.

Copyright 2009 ACM 978-1-60558-675-5 ...\$5.00.

cost-proportionate rejection sampling method in [7] to correct our biased training examples which contains only “sold items”. We use starting price and sellers’ feedback score as the sample selection probability. We show in experiments that the predicted end-price is closer to the actual end-price.

On the other hand, end-price alone is not enough for auction sellers to predict their net revenue. As described above, end-price prediction and sold probability are two things that could not be maximized at the same time. Thus, we argue in this paper that sold probability should also be considered in sellers’ decision process and a decision based on both sold probability and end-price should give higher profit than considering them alone. To this end, an accurate probability for each class membership is necessary. Although most supervised classifiers output ranking scores for examples, they need to be calibrated into probability. In this paper, we use Platt’s parametric approach to map SVM scores into well-calibrated sold probability [4]. We show in experiments that we can get higher profit gain with the support from calibrated probability than probability without calibration.

Finally, although some researches suggest the use of predicted end-price for selling strategy, their papers did not address the problem of how to apply predicted end-price in decision making but only focus on precision of end-price prediction. For this problem, we proposed the use of cost-sensitive decision making to resolve whether to list a commodity or not. We compare three different decision criteria respectively: one depends on sold probability), another depends on end-price, and the other depends on both of them. We use profit gain over average profit of similar items sold by similar sellers as our measure and show that CS (cost-sensitive) approach provides highest profit gain than the first two approaches.

The rest of the paper is organized as follows: In Section 2, we describe the data and features used by our learning algorithms. Section 3 discusses the criteria to support the decision whether to use current auction setting. Section 4 and 5 details the procedure for sample bias correction and sold probability calibration, respectively. We report our experimental evaluations to prove our argument in Section 6. Finally, in Section 7, we summarize the contribution of this paper and suggest the directions for future work.

2. DATA

We write a crawler to collect auction data from eBays in the category of digital cameras over a period of two months in March-April 2006. We select five models (A530, SD600, SD550, S2, A620) of digital camera with maximum transactions. There are a total of 4,852 records extracted for our research. Table 1 shows the statistical information of the data set. Note that we exclude old commodities from our data set because they are highly affected by their descriptions and pictures which is hard to predict by our collected data.

We preprocess the collected data to extract for each auction item 72 features which are classified into 3 classes as used in [1]:

- **Seller Features:** feedback score, negative feedback, positive feedback, IsPowerSeller, HasAboutMePage, HasEBayStore, ActivePeriod, number of products listed by the seller, etc.
- **Item Features:** memory size, warranty, bundled kits

Table 1: Statistics of the collected data set

Model	No. of Items	Sale Ratio	Average end-price	Standard Deviation
A530	881	0.63	177	34
SD600	831	0.64	296	38
SD550	807	0.44	321	44
S2	1021	0.73	361	62
A620	1312	0.61	269	47

Table 2: Comparison of auction listing and BuyItNow listing

Model	Auction Listing			BuyItNow Listing		
	No. Items	Sale Ratio	End-Price	No. Items	Sale Ratio	End-Price
A530	427	0.92	172.12	454	0.35	189.46
SD600	394	0.96	292.75	437	0.34	304.06
SD550	322	0.76	313.31	485	0.22	338.07
S2	538	0.96	339.98	483	0.46	409.95
A620	626	0.94	258.30	686	0.30	301.06

(bag, battery, memory, tripod, lens, memory, reader, etc.).

- **Auction Features:** auction listing or BuyItNow listing, starting price, BuyItNow price, shipping cost, start date, end date, auction duration, pictures, presence of reserve price, payment method, listing upgrade features (bold, subtitle, etc.) some predefined words in title or subtitle (no reserve, fast ship, etc.).

Generally speaking, sellers have two ways to sell items on eBay - Auction listing and BuyItNow listing. For auction listing, sellers give a starting price and then the buyers bid for the item until the auction duration is over. For BuyItNow listing, sellers give a direct price (BuyItNow price). Once any buyer is willing to pay the price, the bid is over. As shown in Table 2, except the SD550, almost all items were sold out (Sale Ratio: 0.92~0.96) for auction listing. For BuyItNow listing, the sale ratio is much lower (0.22~0.35) and the average end-price is higher since most sellers are professional sellers and the sale is usually associated with more kits as we will see later.

Table 3 shows the average starting price and listing cost for sold and unsold items, respectively. Listing cost refers to money paid to eBay, which depends on the auction setting. As we can see, unsold items have higher starting price and listing cost than those of sold items. This can be explained as auction sellers set higher starting bids in order to get higher end-price, but the auction fails due to the same reason. In this situation, auction sellers also incur additional lost due to higher starting price. As for BuyItNow listing, items with higher BuyItNow price seem to scare many buyers away. Thus, sellers need to spend more listing cost to promote their commodities (e.g. via advertisement), showing the intense competition. This analysis demonstrates that we need to consider both sold probability and end-price in order to get high profit.

For items using auction listing, there is no BuyItNow price; while for items using BuyItNow listing, the starting price is typically null. If we mix these two types of items, the trained model will be greatly affected by the null values in the data set. Meanwhile, the end-price for items using

Table 3: Starting price (SP) and listing cost (LC) comparison

Model	SP/ LC	Auction Listing		BuyItNow Listing	
		Sold	Unsold	Sold	Unsold
A530	SP	55.28	101.13	189.56	211.59
	LC	1.27	1.81	7.04	3.91
SD600	SP	67.50	232.73	304.29	354.25
	LC	1.37	2.91	6.01	4.88
SD550	SP	102.89	323.35	338.07	373.50
	LC	1.93	2.27	6.11	4.26
S2	SP	87.07	104.23	410.15	425.63
	LC	1.62	1.84	6.82	4.94
A620	SP	55.10	88.23	301.22	303.33
	LC	1.41	2.30	6.11	3.71

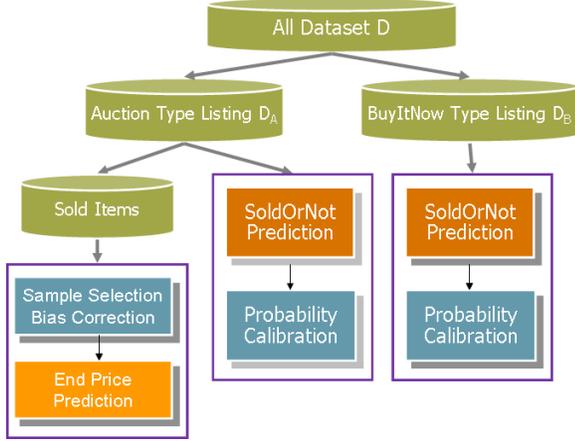


Figure 1: Two separate tasks: sold probability estimation and end-price predication.

BuyItNow listing is exactly the BuyItNow price. There will be no need to predict the end-price. Thus, we separate the data into auction listing D_A and BuyItNow listing D_B . For BuyItNow listing D_B , we need to estimate and calibrate the sold probability. For auction listing data D_A , we also need to build an end-price prediction model and solve the sample selection bias problem as well as the sold probability estimation model (see Figure 1).

3. TO SELL OR NOT TO SELL?

Upon deciding whether to use a given auction setting, the major problem is how the predicted end-price and sold probability be used for judgment and how to evaluate various approaches. Intuitive criteria like sold probability could be used to maximize sale amount by suggesting sell when the sold probability is greater than 0.5. End-price, however, could not be easily modeled since every seller has his own profit expectation or requirement.

To be more specific, given the predicted end-price, the seller would use an auction setting (i.e. decide to sell) if the profit satisfies the sellers' expectation. Since we do not know the seller's cost and his requirement of profit, the judgment is hard to formalize. Thus, instead of comparing to one's absolute requirement, we suggest a comparison to the end-price of other similar items sold by similar sellers. In a way, this judgment denotes whether the current auction setting

Table 4: Profit gain matrix

Profit Gain	SOLD	UNSOLD
SUGGEST SELL	$Profit(x) - AvgP(x)$	$-lc(x) - uc$
SUGGEST NOT TO SELL	$AvgP(x) - Profit(x)$	$lc(x) + uc$

has higher revenue than average auction setting. Thus, criteria based on end-price can be formalized as suggesting sell if the predicted end-price of item x is greater than the average end-price of similar items sold by similar sellers¹.

However, approaches based on sold probability alone or end-price alone are not enough. Both tips should be used for the decision support. On the other hand, a measurement to evaluate whether sale amount based or end-price based performs better is necessary. Again, we compare the profit of an item with the average profit of other similar items² sold by similar sellers³. Let $y(x)$ and $lc(x)$ be the end-price and listing cost for item x , respectively. The profit that item x generates is defined as $Profit(x) = y(x) - lc(x)$ and the average profit of this item can be calculated from other similar items by equation 1.

$$AvgP(x) = \frac{1}{n_x} \sum_{x' \sim x} Profit(x') \quad (1)$$

where n_x is the number of items similar to x and sold by similar sellers. Thus, the gain of adopting this auction setting against others could be defined as $Profit(x) - AvgP(x)$ if item x is sold; or if it is not sold, the seller would incur the loss of listing cost and other cost due to stock or value decreasing. We denote such a cost by uc . From the other aspect, we can save the seller from less earning of $AvgP(x) - Profit(x)$ or $lc(x) + uc$ by suggesting not to use current auction setting. Table 4 shows the profit matrix we use to measure the gain of current auction setting for item x against others. To evaluate the performance of a decision criterion, we sum up all $M_x(i_x, j_x)$ for every x in the testing set, which is the gain of predicting class i_x when the true class is j_x .

Such a measurement implies that minimization of classification error in general approaches do not work since each outcome of the four cases has different profit gain. For this kind of problem, cost-sensitive decision making is a good approach that could change decision boundary to help us choose the class with maximum profit gain. Suppose we can estimate the probability for each class j , $P(c = j|x)$, Bayes' risk theory suggests that we choose class i with the maximum expected profit gain as follows.

$$\arg \max_i \sum_j P(c = j|x) M_x(i, j) \quad (2)$$

¹Readers might wonder why not using reserve price for seller's cost? The reason is that reserve price is optional, thus could not be adopted here.

²Note that items of the same model have different values when they come with different kits. In this paper, we assume items with the same kits amount have the same value and are similar.

³Different sellers have different profit expectation. To distinguish such sellers, we divide sellers into four types by their feedback score: range 0~100, 100~1000, 1000~5000, and above 5000.

The idea is similar to Zadrozny and Elkan’s work on KDD-Cup98 data set [8], where they apply direct cost-sensitive decision making to send mail only when the expected donation is greater than mailing cost. In this case, we suggest the use of current auction setting only when the expected profit gain is greater than listing cost plus ultra cost.

In summary, we have three approaches of decision criteria:

- **Sold Probability Based Approach:** When the predicted sold probability is greater than 50%, i.e. $P(c = 1|x) > P(c = -1|x)$, we suggest the seller list the auction. Thus, this approach depends only on sold probability.
- **End-price Based Approach:** When the predicted end-price is higher than the average end-price of similar items, i.e. $y(x) > Avgy(x)$, we suggest the seller list the auction. Thus, this approach depends only on end-price.
- **Expected Profit Based Approach:** When the expected profit gain for suggesting sell is greater than zero, i.e.

$$\frac{P(c = 1|x)[Profit(x) - AvgP(x)]}{P(c = -1|x)[lc(x) + uc]} > \quad (3)$$

we suggest the seller use the current auction setting. Thus, this approach combines both sold probability and end-price.

4. END-PRICE PREDICTION

There are two types of auction end-price predication: static and dynamic. The former uses machine learning algorithms and is more useful for auction bidders while the later applies functional data analysis and is important to auction sellers. To support auction sellers in commodity listing, we only need static end-price predication as auction setting must be determined before the auction begins.

4.1 Multiple Binary Classification

In this paper, we use multiple binary classifiers to predict the end-price for auction listing items [1]. The idea is to sort auction items by their prices and divide data into two classes at various prices. For each price $\$Y_i$, we train a binary classifier which judges whether the end-price is greater than $\$Y_i$ or not. The output for a test example x upon this binary classifier is then a three-tuple (" $>\$Y_i$ ", true/false, θ) showing that x is greater than $\$Y_i$ (true; or false otherwise) with confidence θ . Finally, the end-price is determined by the largest $\$Y_i$ (and its larger neighbor $\$Y_{i+1}$) which predicts x as true. For example, if the outputs of three binary classifiers $\$50$, $\$45$, and $\$40$ are (" $>\$50$ ", -1, 0.8), (" $>\$45$ ", 1, 0.85), (" $>\$40$ ", 1, 0.9), we determine the end-price locates between segment 50 and 45 and take the mean value 47.5 as the predicted end-price. As discussed in [1], each binary classifier can use the full data set as training data, which is important for data sets with few transaction items.

4.2 Sample Selection Bias Correction

As described above, since we can only use sold items as the training data for end-price prediction, the selection of examples for training is biased. We can interpret this situation as many high end-price items are not selected from training set due to their lower sold probability. In other words, the

negative correlation between sold probability and end-price will cause the prediction price lower than its real value.

Sample selection bias problem has been studied in econometrics and statistics. Heckman [2] proposed a two-step procedure to solve the problem in 1979. However, the method is limited to linear regression model. In recent years, there are some researches from machine learning field [5, 7] which focus on discrete predication, i.e. classification problem. There are four kinds of sample selection bias [7]:

- **Complete Independent:** Selection variable s is independent of attribute x and label y , that is, $P(s = 1|x, y) = P(s = 1)$, then the selection is not biased.
- **Feature Bias:** Selection variable s is dependent of attribute x and independent of label y with attribute x , that is, $P(s = 1|x, y) = P(s = 1|x)$.
- **Class Bias:** Selection variable s is dependent of label y and is independent of attribute x with label y , that is, $P(s = 1|x, y) = P(s = 1|y)$.
- **Complete Bias:** Selection variable s is dependent of both attribute x and label y , that is, $P(y|s = 1) \neq P(y)$ and $P(x|y, s = 1) \neq P(x|y)$.

In our work, whether an item is sold can be taken as a selection variable s , label y is the discretized end-price segment, and property x means the features (including seller features, item features, and auction features). Since the data used for end-price prediction are all sold items ($s=1$), we only have data sets for $(x, y, s=1)$ and $(x, s=0)$ but not $(x, y, s=0)$, which is the most complicated complete bias case (s is dependent of both x and y). In theory, we cannot predict under complete bias. Thus, we need further assumption to solve the complete bias problem.

We assume that items with the same features have the same end-price, and therefore the same label, no matter it is sold or not, that is, $P(y|x) = P(y|x, s = 1)$. With this assumption, the problem can be simplified to feature bias problem as shown in equation (4).

$$\begin{aligned} P(s = 1|x, y) &= \frac{P(x, y|s=1)P(s=1)}{P(x, y)} \\ &= P(x|s = 1)P(y|x, s = 1) \times \frac{P(s=1)}{P(x, y)} \\ &= P(x|s = 1)P(y|x) \times \frac{P(s=1)}{P(x, y)} \\ &= \frac{P(s=1|x)P(x)}{P(s=1)} \times \frac{P(x, y)}{P(x)} \times \frac{P(s=1)}{P(x, y)} = P(s = 1|x) \end{aligned} \quad (4)$$

Comparing to complete bias problem, feature bias calibration is solvable and there exists methods for bias correction. Since we model the end-price predication problem by multiple binary classifiers, we can apply Zadrozny’s reweighting method [7] to calibrate sample selection bias. Assume that we know the selection probabilities $P(s = 1|x)$ and they are greater than zero for all x . The idea of Zadrozny’s reweighting method is to calibrate data distribution by equation (5):

$$\hat{D}(x, y, s) \equiv \frac{P(s = 1)}{P(s = 1|x)} \times D(x, y, s) \quad (5)$$

We can consider each record (x, y, s) of data distribution $\hat{D}(x, y, s)$ as a record $D(x, y, s)$ of distribution D with weight $P(s = 1)/P(s = 1|x)$. Zadrozny proved that, after calibration by the above formula, the expected loss function $l(h(x), y)$ of a classifier h learned from feature bias data set (\hat{D}) is equal to the result for non-bias data, that is,

$$E_{x, y \sim \hat{D}}[l(h(x), y)|s = 1] = E_{x, y \sim D}[l(h(x), y)] \quad (6)$$

Algorithm 1 Modified Costing

Input: Learner A , Training set D , iteration t

```

for  $i = 1$  to  $t$  do
   $S_i = \emptyset$ 
  for  $j = 1$  to  $m$  do
    sample  $x$  from  $D$  and  $u$  from  $U(0, 1)$ 
    if  $u < P(s = 1)/P(s = 1|x)$  then  $S_i = S_i \cup \{x\}$ 
  endfor
  for each price  $Y$  do  $h_i^Y \equiv A(S_i)$ 
endfor
for each price  $Y$  do  $h^Y \equiv \text{VotingOf}(h_1^Y, h_2^Y, \dots, h_t^Y)$ 

```

Figure 2: Modified costing algorithm

However, to obtain a sample from a weighted distribution is not completely straightforward. Zadrozny recommends the use of Costing method for feature bias correction 0. Costing is an ensemble learning algorithm which aggregates multiple base classifiers learned from each sampled data set using rejection sampling. The algorithm was originally designed to solve problems with non-uniform misclassification costs, where straightforward sampling (with or without replacement) does not work well since samples are not drawn independently from \hat{D} . Rejection sampling on the other hand ensures that the sampled examples are distributed independently according to \hat{D} .

Algorithm 1 shows the modified costing algorithm with selection ratio $P(s = 1)/P(s = 1|x)$ as a weight for each example. We repeat t iterations to produce t sample sets S_i ($i = 1$ to t) and use them to train t binary classifiers h_i^Y ($i = 1$ to t) for each price Y . The outputs of the aggregated h^Y for each price are then used to predict the end-price for testing examples. As shown in Figure 1, this reject sampling procedures are conducted first to correct data distribution before the application of multiple binary classification tasks. While averaging over multiple learners gives better result both in theoretical and empirical views, the overall computation time is also reduced since the size of S_i is smaller compared with the whole training set D .

5. SOLD PROBABILITY CALIBRATION

To estimation the sold probability for an item, we can use any binary classification to train a model for predicting whether an item would be sold ($c = 1$) or not ($c = -1$). The output of this model is then mapped into posterior probability. In this paper, we use SVM (Support Vector Machine) as our classifier. SVM is a very popular classification algorithm in recent years, and it is used to solve a variety of classification problems. However, SVM produces an uncalibrated value that is not a probability. To yield well-calibrated probability, we adopt Platt Scaling (1999) to transfer a value $f(x_i)$ into probability $P(c = 1|x_i)$. Platt Scaling assumes a parametric model of a sigmoid mapping with two parameters α and β (equation 7).

$$p_i \equiv P(c = 1|x_i) = \frac{1}{1 + \exp(\alpha f(x_i) + \beta)} \quad (7)$$

The parameters can be fit using maximum likelihood estimation from a training set (f_i, t_i) , where t_i are target probabilities defined as $t_i = \frac{c_i + 1}{2}$ (i.e. 1 for $c_i = 1$ and 0 for $c_i = -1$). By minimizing the negative log likelihood of the training data, which is a cross-entropy error function

Table 5: Classification accuracies for SoldOrNot prediction

MODEL	AUCTION LISTING	BUYITNOW
A530	91.9	71.5
SD600	96.1	76.8
SD550	92.9	84.8
S2	94.5	71.4
A620	92.1	76.5
AVERAGE	93.5	76.2

(Kullback-Liebler divergence between f_i and t_i),

$$\arg \min_i \left\{ - \sum_{x_i \in D} t_i \log p_i + (1 - t_i) \log(1 - p_i) \right\} \quad (8)$$

we have an unconstrained optimization function which can be solved using gradient descent algorithm.

To avoid overfitting to a small number of examples, Platt suggests using non-binary targets (equation 9) instead of regularization for parameter space (α, β) ,

$$t_+ = \frac{N_+ + 1}{N_+ + 2}, t_- = \frac{1}{N_- + 2} \quad (9)$$

where N_+ and N_- are numbers of positive and negative examples, respectively.

6. EXPERIMENTS

We use the data described in Section 2 for the following experiments. For each camera model, we randomly select 70% data as training data and the remaining 30% as testing data. The process repeats five times to have 5 sets of training and testing data. For each experiment, we take the average result of the five data sets. The experiments can be divided into three parts: The first part reports the performance on item sold prediction, and the relationship between sold probability and auction features. The second part shows the accuracy of end-price prediction by using multiple binary classifiers, and the effects of sample selection bias correction. The third part compares performance of three approaches: probability-based, end-price based, and expected profit based approaches.

6.1 Accuracy of Sold Probability Estimation

We use SVM as our classifier for predicting whether an item is sold or not. The accuracy of prediction whether an item will be sold is about 93.5% and 76.2% for auction listing and BuyItNow listing, respectively (Table 5). The accuracy for BuyItNow is lower than the corresponding items in auction listing. We found that some items even with the same auction setting have different outcome. This is because some sellers have multiple items to sell. Some might be sold while others might not. Thus, the accuracy for BuyItNow is much lower than Auction listing.

To analyze the relationship between sold probability and auction settings, we apply Platt scaling to transfer the SVM output into probability and divide the data into ten bins according to their sold probability. We then calculate the average starting price and kits amount of the commodities in each bin. We also examine the average feedback score of sellers for those items to see if seller's feedback is also relevant to sold probability.



Figure 3: Negative correlation between difference of starting price and sold probability

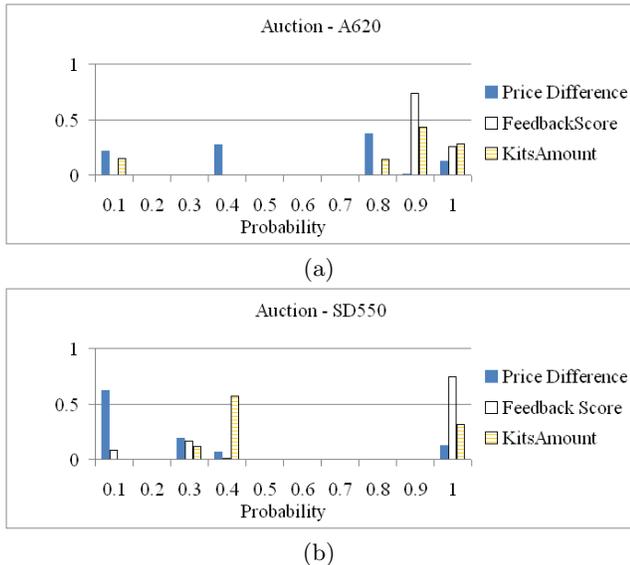


Figure 4: Difference in starting price, feedback score and kits amount vs sold probability for Auction data with high starting price

Recall from Table 3 that sold items have lower starting price. Thus, we expect to see that items with higher starting price should have lower probability. Figure 3(a) shows the result between sold probability and difference of starting price for A620 and SD550 under the BuyItNow category⁴. This product shows typical negative correlation between the price difference and sold probability: if the BuyItNow price is greater than the average price of similar items, the items would have very low sold probability.

As for auction listing, the negative correlation between

⁴Due to space limitation, we are not able to show all figures for all products.

starting price and sold probability is not so obvious as shown in Figure 3(b). Thus, we further examine other factors. Since most items are sold for auction listing, we extract data with starting price greater than average end-price from the auction listing to analyze the relationship between sold probability and other relevant factors. Figure 4 shows two such factors as well as difference of starting price: the solid bars represent the difference in starting price, the white bars represent seller’s average feedback score, while the strip bars represent the average amount of kits associated with items (all of them are normalized between 0~1). As the figure shows, feedback score and kits amount has positive correlation with sold probability, while difference in starting price shows negative correlation with sold probability. Although the relationship is not very definite for A620, the trend is apparent for SD600 as shown in the figure.

From the analysis above, we confirm our argument that starting price could strongly affect sold probability, especially in BuyItNow category. Thus, when sellers use high starting price or BuyItNow price approach to ensure higher end-price, they should also consider sold probability as well. For auction listing, there are several factors that affecting the sold probability, the influence strength varies for different product. Note that, probability calibration is necessary; otherwise the probability will be either 0 or 1 and cannot indicate the relationship between probability and listing setting.

6.2 End-price Prediction

To use multiple binary classifiers for end-price prediction, we need to decide the reference value \$Y\$ for each binary classifier. We exclude the highest and lowest end-prices, and use 10% window of the average end-price as the interval size, i.e. \$Y\$ (starting from the lowest end-price to the highest end-price) is increased by 10% of the average end-price for each binary classifier. The number of intervals for each model is shown in the second column (denoted by “No. of Int”) in Table 6. Again, SVM is used as the base classifier to predict whether the end-price is greater than the reference value \$Y\$.

Table 6: Accuracies of end-price prediction for sold items without or with bias correction

MODEL	No. of Int	BEFORE BIAS CORRECTION			AFTER BIAS CORRECTION		
		TGT	± 1	± 2	TGT	± 1	± 2
A530	17	48.7	83.6	95.0	47.2	84.0	94.8
SD600	29	59.8	95.4	99.6	58.5	95.5	99.5
SD550	31	46.3	87.4	98.9	42.8	92.7	99.5
S2	33	52.9	93.9	98.3	54.9	94.6	98.7
A620	25	46.9	89.1	98.0	48.2	90.1	98.7

For the collected digital camera data, the major factors affecting end-price are the kits amount and the memory size. Since there are a variety of kits can be associated, the price varies greatly even for the same digital camera item, which causes the difficulty for prediction. Table 6 shows the result of multiple binary classifiers for sold items. Though the accuracy is not high (51%) for the target interval, the accuracy approximates 90% within ± 1 interval range.

As described above, due to sample selection bias problem, the predicted model (based on training data with only sold items) will underestimate the end-price of items with low sold probability. With the assumption that items with the same features will also have the same end-price, we apply the modified costing algorithm to correct the feature bias problem. Hence, we also need to ensure that $p(x, s = 1) > 0$, that is, the selected data contains all feature spaces of the data and with enough examples in each feature space. As starting price and feedback score are the main factors affecting the sold probability, we use these two features to predict $p(s = 1|x)$ for each example. For each of the five training sets (of each digital camera model), we calculate the weight (i.e. $p(s = 1)/p(s = 1|x)$) for every example and use rejection sampling to generate 10 sample sets ($t=10$). Each sample set has about 250 examples. We then apply SVM to train classification models for each $\$Y$, and use the vote of the 10 predictions as the final result for each binary classifier with judgment whether the end-price is greater than $\$Y$.

To see how the correction of feature bias calibrates the prediction of unsold items, we extract unsold items with starting price greater than average end-price of similar items and compare the predicted end-price before and after calibration in Table 7. The result shows that the predicted end-price based on sold items without costing correction is lower than the average starting price of these items (i.e. underestimate), while the predicted end-price after calibration is closer to the average starting price. Similarly, we compare the predicted price before and after costing calibration for linear regression. Contrast to the result for multiple binary classifiers, only SD550 presents visible change by the correction. We suspect it is because SD550 has more unsold items, thus has more impact for linear regression approach based on mean square errors. As shown in Table 6, the accuracy before and after calibration is about the same. Thus, the calibration for sample selection bias does not only maintain the accuracy for sold items but also has better prediction for unsold items.

6.3 Profit Gain Comparison

In order to validate our argument that end-price alone does not guarantee high profit due to low sold probability, we use profit gain over other sellers to compare the three ap-

Table 7: Increase of predicted end-price for unsold items via sample bias correction

Model	Average Starting Price	Multiple Binary Classifier		Linear Regression	
		-Corr.	+Corr.	-Corr.	+Corr.
A530	190.77	181.02	180.17	190.85	192.83
SD600	389.09	310.80	328.1	300.27	300.04
SD550	366.46	341.60	355.22	293.03	323.52
S2	289.20	273.2	293.0	264.33	264.41
A620	296.21	294.0	304.5	259.21	261.38

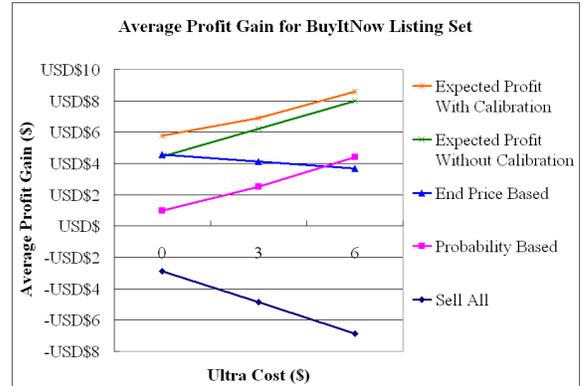


Figure 5: Average profit gain for BuyItNow listing.

proaches described in Section 3. Figure 5 and Figure 6 show the average profit gain of the three approaches for BuyItNow listing and auction listing, respectively. We also use “Sell All” approach as the baseline for comparison, which suggest selling all the time.

For BuyItNow listing set, the end-price based approach uses the BuyItNow price as end-price and suggests sale if the BuyItNow price is greater than average BuyItNow price of similar items sold by similar sellers; while profit based approach uses sold probability either with or without sold probability calibration. Figure 5 shows that the baseline approach (“Sell All”) makes a loss of \$2.882 than average even at $uc=\$0$ since the sale ratio is only 22%~46%. With the help of sold probability prediction, we can obtain an average profit gain of \$0.967 at $uc=\$0$. As uc increases, probability based approach gains even more profit than average since we have ultra cost if the commodities are not sold. On the other hand, end-price based approach although makes a high gain of \$4.578 than average at $uc=\$0$, the profit gain decreases to \$3.769 as uc increases to \$6 due to the decrease in sold probability. Finally, approaches based on expected profit (either with or without probability calibration) obtain the highest profit gain. Even when extra cost uc increases, the average profit gain increases as well, which means expected profit based approaches avoid the loss of unsold items by balancing the expected profit. As seen from Figure 5, probability calibration improves profit gain by making more accurate estimation in sold probability.

Figure 6 shows a similar result in auction listing, though the change with ultra cost is not as obvious: the base line approach has a profit gain of \$0.025 at $uc=\$0$, while probability based approach even makes a loss of \$0.081 at $uc=\$0$. Thus, if most items could be sold, sellers could simply use

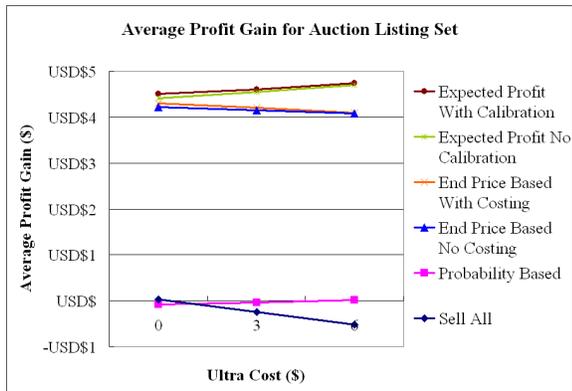


Figure 6: Average profit gain for auction listing.

Table 8: Increase of profit gain by bias correction in auction set

Model	End-Price Based	Expected Profit Based	
		No Calibration	With Calibration
SD550	+0.759	+0.409	+0.689
AVG	+0.089	+0.023	+0.1

“sell all” approach. However, sale amount based approach does not deteriorate like the base line approach as ultra cost increases.

Similar to BuyItNow listing, end-price based approaches (either with or without correction by costing) show a negative correlation with ultra cost, while only profit based approach maintains an increasing profit as ultra cost increases for auction listing. Figure 6 also shows that end-price based approach with feature bias correction performs better than without correction, though the difference is not obvious. The reason why the profit gain by costing is not obvious can be attributed to high sale ratio for auction listing data set. Thus, we can observe more profit gain in SD550 (i.e. the lowest sale ratio model), showing that sample selection bias correction is more important in categories with a lot of unsold items in Table 8.

7. CONCLUSIONS AND FUTURE WORK

A critical question for online auction sellers is to find an auction setting that could maximize profit for their commodities. Although several researches are proposed for end-price prediction, how such information could support sellers’ decision has not been fully explored. In this paper, instead of enumerating all possible auction settings, we provide a selling recommendation service which suggests the use of current auction setting based on comparing average profit of other similar items sold by similar sellers. In a way, this approach denotes whether the current auction setting is better than average auction setting.

We apply machine learning algorithms for end-price prediction and sold probability estimation. For end-price prediction, since we can only use sold items to train the classifiers, the model could under-estimate the end-price for unsold items. As shown in the experiments, approaches that do not involve the consideration of sold probability (e.g. end-price based and the baseline approaches) have deteriorating profit as ultra cost increases. This proves our argument that

sold probability should also be considered for profit maximization. In conclusion, the approach based on both probability and end-price have the best performance among all other approaches.

Although the service described in this paper can support sellers in whether to use the current auction setting or not, a more aggressive goal will be to find the optimal auction setting that could maximize the profit of an item. Since feedback score is an important factor affecting sold probability, simply imitating other sellers’ auction setting may not have the same return. Meanwhile, analyzing description text to provide guidance for writing commodity description is also an interesting research direction. In practice, it is not realistic to build prediction models for every kind of commodity. Transfer learning technique is a promising way to predict “similar” commodity.

8. ACKNOWLEDGEMENT

This work is sponsored by National Science Council, Taiwan under grant NSC-97-2221-E-008-088.

9. REFERENCES

- [1] R. Ghani and H. Simmons. Predicting the end-price of online auctions. In *Proceedings of the International Workshop on Data Mining and Adaptive Modelling Methods for Economics and Management*, Pisa, Italy, 2004.
- [2] J. Heckman. Sample selection bias as a specification error. *Econometrica*, (47):153–161, 1979.
- [3] D. Heijst, R. Potharst, and M. Wezel. A support system for predicting ebay end prices. Technical report, Econometric Institute Report, 2006.
- [4] J. Platt. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. *Advances in Large Margin Classifiers*, pages 61–74, 1999.
- [5] A. Smith and C. Elkan. A bayesian network framework for reject inference. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 286–295, 2004.
- [6] S. Wang, W. Jank, and G. Shmueli. Forecasting ebay’s online auction price using functional data analysis. *Journal of Business and Economic Statistics*, 2006.
- [7] B. Zadrozny. Learning and evaluating classifiers under sample selection bias. In *Proceedings of the 21th International Conference on Machine Learning*, 2004.
- [8] B. Zadrozny and C. Elkan. Learning and making decisions when costs and probabilities are both unknown. In *Proceedings of the Seventh ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pages 204–213, 2001.