Learning to Bid the Interest Rate in Online Unsecured Personal Loans

Dong Jun Jee*
PFC Technologies
Seoul, South Korea
dongjun@pfct.co.kr

Seung Jung Jin*
PFC Technologies
Seoul, South Korea
seungjung@pfct.co.kr

Ji Hoon Yoo
PFC Technologies
Seoul, South Korea
jihun@pfct.co.kr

Byunggyu Ahn
PFC Technologies
Seoul, South Korea
byungkyu@pfct.co.kr

ABSTRACT

The unsecured personal loan (UPL) market is a multi-billion dollar market where numerous financial institutions compete. Due to the development of online banking, loan applicants start to compare numerous loan products. They aim for high loan limits and low interest rates. Since loan applicants have a desired loan amount, institutions instead focus on adjusting interest rates. Despite the importance of determining optimal interest strategies, institutions have traditionally relied on heuristic methods by human experts to set interest rates. This is done by adding a target return on assets (ROA) to the applicant's expected default probability predicted by a credit scoring system (CSS) such as the FICO score.

We conceptualize the UPL market dynamics as a repeated auction scenario, where loan applicants (akin to sellers) seek the lowest interest rates, while financial institutions (akin to bidders) aim to maximize profits through higher interest rates. To the best of our knowledge, this is the first time anyone has approached the UPL market through the viewpoint of a repeated auction. While there are several research done in learning to bid in repeated auctions, those works cannot be directly applied to the UPL market due to the lack of any feedback about other bidders' strategies and the need to satisfy the bidder's target loan volume and profit variance. We present an algorithm named AutoInterest, which is a modification of the dual gradient descent algorithm. In addition, we provide a framework to evaluate interest rate bidding strategies on a benchmark dataset and the credit bureau dataset of actual loan applicants in South Korea. We evaluate AutoInterest on this framework and show higher cumulative profit compared to other common online algorithms and the current fixed strategy used by real institutions.

*Both authors contributed equally to this research.

KDD '24, August 25–29, 2024, Barcelona, Spain

ACM ISBN 979-8-4007-0490-1/24/08
https://doi.org/10.1145/3637528.3671584

CCS CONCEPTS

• **Applied computing → Online banking**; *Online auctions*; • **Computing methodologies → Sequential decision making**.

KEYWORDS

Unsecured Personal Loan, Interest Rate Strategy, Loan Comparison, Online Learning, Budget Constraint, Variance Constraint

ACM Reference Format:
Dong Jun Jee, Seung Jung Jin, Ji Hoon Yoo, and Byunggyu Ahn. 2024. Learning to Bid the Interest Rate in Online Unsecured Personal Loans . In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24), August 25–29, 2024, Barcelona, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3637528.3671584

1 INTRODUCTION

In recent years, the growth of online banking has made unsecured personal loans (UPL) much more accessible, leading to a significant expansion of the market. Various financial institutions, including banks, credit unions, and savings banks, actively engage as lenders within the UPL market. According to the Federal Reserve Bank of New York, the UPL market grew from \$104 billion in 2017 to \$232 billion in 2023 [40]. However, finding the right balance in loan products to appeal to both loan applicants and institutions remains a complex challenge. Online banking and loan comparison services have empowered applicants to make more informed decisions, increasingly seeking higher loan amounts and lower interest rates. Meanwhile, institutions are focused on maximizing the utility of their capital by tweaking loan conditions. Due to constraints such as government regulations, institutional risk management policies, and applicants' needs, there is more room for adjustment in interest rates than in adjusting loan amounts [21]. Nevertheless, the study of optimizing interest rate strategies has been minimal, primarily relying on risk based pricing (RBP), which is slow and highly dependent on expert-driven modifications. [16, 39].

Adopting optimization algorithms has improved efficiency in areas such as inventory management [37], dynamic pricing [46], marketing [12] and online advertisement [9]. In this paper, we present two ideas to optimize the interest rate strategy: first, modeling the UPL market as a variant of the first price auction; and second, employing optimization techniques to design interest rate strategies.

In the UPL market, loan applicants seek the lowest interest rates, while financial institutions (akin to bidders) try to maximize profits by increasing the offered interest rate. Financial institutions compete against each other for each loan applicant. Therefore, the UPL market can be viewed as a repeated auction where the lowest bidder wins and receives the interest as the reward, with each round of the repeated auction corresponding to a single loan applicant. To the best of our knowledge, we are the first to study the UPL market as a kind of auction problem. However, key differences distinguish the first price auction from the UPL market modelled as auction. In first price auctions, the winning bid is disclosed to all bidders. In contrast, lenders are only informed if an applicant chooses them, without any insight into the bids of competitors or the selected institution. Furthermore, there exist constraints for lenders in the loan market: limited budget for loans and the variance in the profit of their loan portfolio.

In this paper, we present AutoInterest, an adaptation of the online gradient descent algorithm with dual variables that aims to maximize the cumulative reward (profit). AutoInterest combines ideas from multi-armed bandit algorithms to incorporate the feedback structure of the UPL auction, as well as ideas from the gradient descent algorithm to consider the loan budget and profit variance constraints. Unlike the RBP strategy, where interest rates are largely fixed for every applicant, AutoInterest updates its interest rates after each applicant to maximize the cumulative profit of the institution.

We test AutoInterest on our open-sourced framework[1] on the Default of Credit Cards (Credit Default) dataset [47] and the NICE credit bureau (CB) data of South Korea. The CB dataset is a real-life dataset of loan applicants from January to March of 2023, with RBP strategies developed by human experts from PFCT, an online lending company. This serves as a production-level backtest of AutoInterest's performance on the actual UPL market. Through this framework, we show that our algorithm outperforms the RBP strategy and other common online learning algorithms.

Our contributions of this work are:

- We model the UPL market as a variant of the first price auction for the first time.
- We study optimal interest rate strategies as a constrained optimization problem with budget and variance constraints, departing from the traditional RBP strategy.
- We develop an algorithm named AutoInterest to solve the constrained optimization problem online, combining ideas of multi-armed bandit and dual gradient descent.
- We develop an offline evaluation tool using real-world finance data to evaluate various interest rate strategies.
- We show that our learning algorithm surpasses the current interest rate strategies used in financial institutions and other learning algorithms on the benchmark dataset and on real-life CB data from South Korea.

2 RELATED WORKS

The multi-armed bandit is a classic problem that finds a balance between exploration and exploitation to maximize the cumulative reward. In the simplest case, we select one of the k possible actions every round and observe a reward sampled from the stochastic distribution of the chosen action. As the algorithm can only observe the reward of chosen actions, it has to explore various arms while trying to exploit its observed reward. Learning algorithms for multi-armed bandit include upper confidence bound algorithm (UCB) [4], its variations [3, 26, 32], ϵ-greedy [42] and gradient-based algorithms [36]. Several variations of the multi-armed bandit are related to learning the interest rate strategy. The bandit with knapsack [5, 13] problem addresses scenarios with limited budgets, where the available budget is consumed every round based on the actions selected. In contextual bandit [49], we observe contextual information for each round before choosing actions.

Learning in repeated auctions has been extensively studied from both the perspectives of sellers and buyers in various types of auctions. In the first price auction, the highest bidder wins and pays its own bid; in the second price auction, the highest bidder wins but pays the second highest bid. Sellers set a reserve price where sellers sell only if the winning bid is greater than the reserve price. Sellers learn how to set reserve prices during repeated auctions [2, 14, 23]. From a bidder's perspective, the bidder learns a bidding strategy to maximize its rewards. For the second price auction, it is known that the optimal bidding strategy is bidding the truthful value [38] regardless of the strategy of other bidders. Therefore, in a second price auction, when the bidder does not know the true value, the bidder learns the true value online [45]. Furthermore, the bidder observes additional information about the second highest bid when the bidder wins. In [7], the authors considered cross-learning in the first price auction with binary feedback where the bidders do not gain information other than whether the bidders won. In [28], the authors considered the first price auction with partial feedback where the bidders observe the highest bid when they lose.

Auctions have mainly been studied in the context of online advertisements [10, 25, 27, 44]. Most online advertisements are sold through real-time auctions on platforms such as Google [19, 22] where advertisements with the highest bid win. There are many areas of research on online advertisement auctions. A pay-per-click auction studied in [20, 24, 41] is standard in online advertisement auctions; bidders pay by the number of clicks multiplied by the bid. In the auto-bidding model setting, the bidder tries to maximize values such as the number of conversions under the budget and ROI constraints. As the target for auto-bidding agents is value-maximizing, the agent might overspend. Therefore, the additional ROI constraint where the ratio between spending and reward must be higher than a fixed number is commonly considered. Various papers study the theoretical properties in auto-bidding settings such as efficiency, robust design, and price of anarchy [1, 6, 18, 33]. In the online setting, an online learning algorithm called pacing is introduced for repeated second auction with budget constraint [8, 10]. These algorithms utilize the property of a second price auction where information about the highest bid of other bidders can be observed when the bid is won. In [25, 27], the authors design an online learning algorithm for a second price auction with return on spending (ROS) constraints. On the first price auction, [44] devised an online learning algorithm for the first price auction with a budget under full feedback and partial feedback.

[1] https://github.com/seungjungjin-pf/AutoInterest

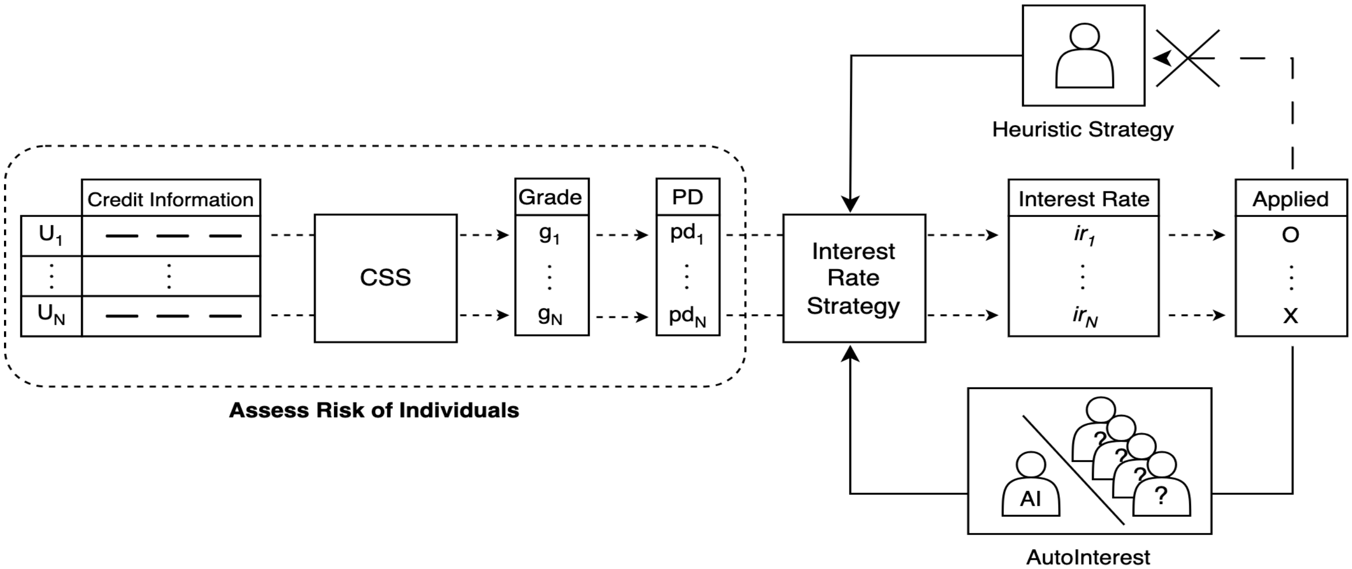

Figure 1: The standard process of providing loans by financial institutions. Institutions estimate the probability of default (PD) based on the credit grade assessed by the credit scoring system (CSS). Traditionally, human experts establish a heuristic-based interest rate strategy with almost no feedback on loan applications. In contrast, AutoInterest operates on a repeated auction, competing with unknown bids, and incorporates partial feedback from every interaction into its strategy. U_i : loan applicant, g_i: credit grade range in [1, 10], pd_i: PD corresponding to g_i, ir_i: loan interest rate

3 PROBLEM FORMULATION

3.1 Risk Based Pricing

Risk based pricing (RBP) is currently the most prevalent interest rate strategy in the UPL market [21]. This strategy calculates interest rates by summing the loan applicant's probability of default (PD), operational and procurement costs, and the desired return on assets (ROA) [16, 39]. Institutions use credit scoring systems either built internally or provided by credit bureaus, such as FICO in the U.S. and NICE in South Korea [34, 35] to evaluate the applicants' credit risk factors. Loan applicants are then typically divided into grades ranging from 1 to 10, where each grade has an expected probability of default. Operational and procurement costs consist of costs such as having employees to operate the loans, and costs of financing the loan. On top of other expenses, human experts set the target ROA and also make heuristic adjustments to account for other factors such as economic outlook and additional information regarding the applicants [48]. It should be noted that new RBP on revised ROA and heuristic adjustments are updated on an irregular basis.

3.2 The Interest Rate as Auction

For loan applicants, financial institutions offer interest rates based on the applicants' credit grades evaluated via the institutions' credit scoring systems (CSS). The applicant then chooses the institution that provides the lowest interest rate. Since we are focusing on the optimization of the interest rates, we assume that every institution offers the same loan limit equal to the applicants' desired loan amount. The applicant will pay back the loan amount along with the given interest rate. Consequently, this process can be viewed as a variant of the first price auction with a unique reward structure where the lowest bidder, as opposed to the highest, wins. When an applicant defaults on their loan, the institution incurs losses equal to the loan amount. Therefore, the expected reward can be computed as the loan amount multiplied by the difference between the interest rate and the default rate.

Most institutions maintain a consistent interest rate strategy over short periods, with any alterations being relatively minor [16, 39]. Therefore, we assume that the minimum competing bids for each grade are independent and identically distributed (i.i.d), given grade. Additionally, we assume that the minimum competing bids are independent of the loan amount, given grade. In this context, contrary to the advertisement auction scenario where bids and values are presumed independent [30], the minimum competing bid is directly linked to the grade. Moreover, due to every institution's shared objective of predicting applicant defaults using their own CSS, the outputs of these systems lead to a correlation between the bids and grades of institutions. We approach this problem as a multi-item auction with grades serving as distinct *items* but, due to shared budget and variance constraints across grades, we cannot treat each grade as a distinct *auction*. In typical auctions, bidders receive either full or partial feedback on other bidders' bids upon losing. However, we only receive information on whether the applicant selected us, which makes the problem particularly challenging.

3.3 Model

Each round $t = 1, \ldots, T$ corresponds to a single loan applicant. Let a_t be the credit grade of CSS for the applicant t and p_i is the expected default rate of grade i with N being the total number of grades. The observed loan amount of the applicant t is v_t, b_{t,a_t} represents the

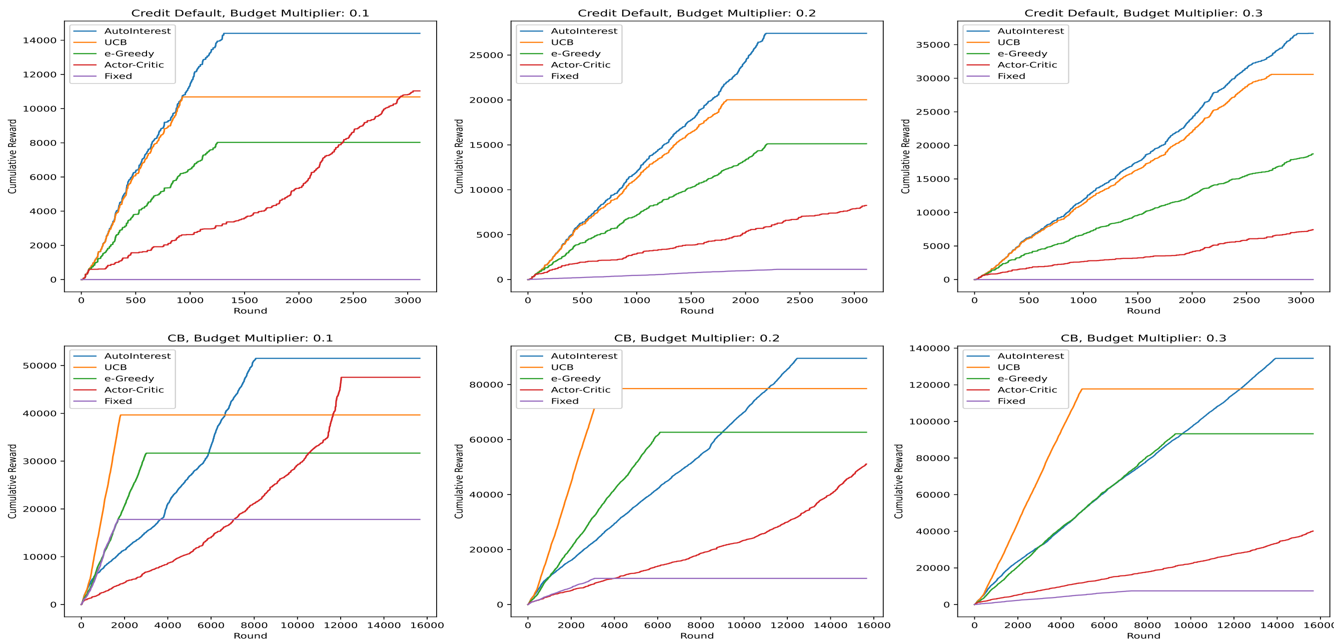

Figure 2: The cumulative reward for algorithms on budget multiples of 0.1, 0.2, and 0.3. The top half is run on the Credit Default dataset, and the Bottom Half is run on the CB dataset.

interest rate offered by the bidder and d_{t,a_t} represents the minimum interest rate offered by other institutions for the applicant t. $\bar{b}$ represents the maximum possible interest rate that can be offered. Let the minimum competing interest rate $d_{t,i}$ be i.i.d sampled from the distribution F_i given that applicant t's credit grade is i. The binary variable $1\{b_{t,a_t} \leq d_{t,a_t}\}$ represents whether the bidder is selected by applicant t or not. For simplicity of notation, we define $G_i = 1 - F_i$ as the probability of being selected. Note that G_i is unknown to the bidder. Given grade, whether the applicant t defaults or not is given by the random variable $X_t | a_t = i \sim Ber(p_i)$, and by taking conditional expectation given grade, we get $E(X_t | a_t) = p_{a_t}$. W.L.O.G, we assume operational costs and procurement costs are zero. Otherwise, we can add operational and procurement costs to the default rate. The conditional expected reward is then given by $1\{b_{t,a_t} \leq d_{t,a_t}\} \cdot v_t \cdot (b_{t,a_t} - p_{a_t})$.

The institution has a budget B and let $\rho = \frac{B}{T}$ be the budget ratio. The amount of budget consumed for applicant t is given by $1\{b_{t,a_t} \leq d_{t,a_t}\} \cdot v_t$. Note that the amount of budget we consume is not related to b_{t,a_t}, as b_{t,a_t} only affects whether we win the bid. In addition, the institution needs to maintain the variance of loss resulting from defaults in its portfolio. Let σ^2 be the maximum variance the institution can tolerate and $\nu = \frac{\sigma^2}{T}$. As $X_t \sim Ber(p_i)$, the variance of X_t is given by $p_i \cdot (1 - p_i)$. Since the institution gives out loan amount v_t to applicant t, the variance of loss due to default is given by $v_t^2 \cdot (p_i) \cdot (1 - p_i)$. Note whether the applicant t defaults or not is independent of the other applicants. Therefore, the total variance of the portfolio can be determined as $\sum_{t=1}^{T} 1\{b_{t,a_t} \leq d_{t,a_t}\} \cdot v_t^2 \cdot (p_i) \cdot (1 - p_i)$.

3.4 Optimization Problem

We consider a constrained optimization on the expectation of the budget and variance constraints.

$$
\begin{aligned}
\max_{b} \ & E_{d,v,a} \sum_{i=1}^{N} \sum_{t=1}^{T} 1\{a_t = i\} \cdot 1\{b_{t,i} \leq d_{t,i}\} \cdot v_t \cdot (b_{t,i} - p_i) \\
\textbf{s.t} \ & E_{d,v,a} \sum_{i=1}^{N} \sum_{t=1}^{T} 1\{a_t = i\} \cdot 1\{b_{t,i} \leq d_{t,i}\} \cdot v_t \leq \rho T \\
& E_{d,v,a} \sum_{i=1}^{N} \sum_{t=1}^{T} 1\{a_t = i\} \cdot 1\{b_{t,i} \leq d_{t,i}\} \cdot v_t^2 \cdot p_i \cdot (1 - p_i) \leq \nu T
\end{aligned}
$$

The budget constraint is a hard constraint that must be satisfied and cannot be violated. However, the variance constraint may be hard or soft, depending on business needs.

4 ALGORITHM

4.1 Initialization

In most online learning algorithms, each action must be explored at least once. In our setting, the number of actions (possible bids) is in a similar order to the number of applicants. Therefore, blindly trying each action once during the exploration stage will consume most of the budget. We reduce the possible action space by only considering the bids greater than the estimated default rate. In real life, the possible interest rate is discrete with a step size of d, where d is commonly equal to 0.01%. Therefore, we define the bid space B_i for each grade i as $[p_i, p_i + d, p_i + 2d, ..., \bar{b}]$.

In addition, when we choose an action, we know the result of other actions. For instance, if an applicant selected our institution when the bid is 0.1, we know that the applicant would choose us when we bid any interest rates lower than 0.1. Similarly, if the applicant did not select us, any bid higher than 0.1 will not get selected. Therefore, we choose the middle point of bids that have yet to be selected, reducing the initial exploration to $\log K$ instead of K

Algorithm 1 Initialization

INPUT: Count array C; Selected number array R; Bid Space B
OUTPUT: C; R
IF $\exists\, b$ in B such that $C[z] == 0$:
  Let z_{min}, z_{max} be the minimum and maximum of such z
  Bid $z_{mid} = \frac{z_{max}-z_{min}}{2}$
  If selected :
    For $z <= z_{mid}$:
      $C[z] = C[z] + 1$
      $R[z] = R[z] + 1$
  else :
    For $z >= z_{mid}$:
      $C[z] = C[z] + 1$

4.2 AutoInterest

To solve the constrained optimization problem with budget and variance constraints, we utilize a primal-dual gradient descent approach [11], which has been utilized for learning to bid in the advertisement platforms [44]. The Lagrangian dual L of constrained optimization with dual variables λ_1, λ_2 is given by

$$\begin{aligned}
E_{d,v,a} &\sum_{i=1}^{N}\sum_{t=1}^{T} 1\{a_t = i\} \cdot 1\{b_{t,i} \leq d_{t,i}\} \cdot (v_t \cdot (b_{t,i} - p_i) \\
&- \lambda_1 \cdot v_t - \lambda_2 \cdot v_t^2 \cdot p_i \cdot (1-p_i)) + 1\{a_t = i\} \cdot (\lambda_1 \cdot \rho + \lambda_2 \cdot \nu) \\
&= E_{v,a} \sum_{i=1}^{N}\sum_{t=1}^{T} 1\{a_t = i\} \cdot G_i(b_{t,i}) \cdot (v_t \cdot (b_{t,i} - p_i) \\
&- \lambda_1 \cdot v_t - \lambda_2 \cdot v_t^2 \cdot p_i \cdot (1-p_i)) + 1\{a_t = i\} \cdot (\lambda_1 \cdot \rho + \lambda_2 \cdot \nu)
\end{aligned}$$

At each time step, before bidding, we observe credit grade a_t, its corresponding probability of default p_{a_t}, and loan amount v_t. To maximize the Lagrangian given dual variables λ_1, λ_2, we need to bid:

$$b_{t,a_t} = \arg\max_b G_{a_t}(b) \cdot (v_t \cdot (b - p_{a_t}) - \lambda_1 \cdot v_t - \lambda_2 \cdot v_t^2 \cdot p_{a_t} \cdot (1 - p_{a_t}))$$

Given b_{t,a_t}, the gradient of the Lagrangian respect to λ_1 is $\rho - v_t \cdot G_{a_t}(b_{t,a_t})$ and the gradient of the Lagrangian respect to λ_2 is $\nu - G_{a_t}(b_{t,a_t}) \cdot (1 - p_{a_t}) \cdot p_{a_t}$. We update dual variables using gradient descent.

As the bid space B_i is discrete, we can treat G_i discrete and represent G_i with an array of length B_i. Let C_i be an array of the number of times a bid has been made for grade i and t_i as the sum of elements of C_i. R_i represents an array of the frequency of selection when this interest rate is offered to applicants for grade i. For simplicity of the notation, for each element b_i in B_i, the value of the corresponding element in array G_i, C_i, R_i is represented as $G_i[b_i], C_i[b_i], R_i[b_i]$, where b_i indexes the position of the element within array B_i.

Contrary to the bidding algorithm in [44], we do not gain any knowledge about other institutions' bids. Therefore, we need to estimate $G_i[b_i]$ from the previous history of selections to compute the arg max. We estimate $G_i[b_{t,i}]$ using the previous history of bidding and the response of the loan applicants to the bids. The estimate of $\tilde{G}_i[b_{t,i}] := R_i[b_{t,i}]/C_i[b_{t,i}]$

In [17], the authors present the SGD-UCB algorithm, which combines UCB with stochastic gradient descent(SGD) for constrained optimization with bandit feedback. Motivated by this, we use a UCB-like algorithm to promote the exploration of b_i while estimating $G_i[b_i]$. For the interest rate, as the bid increases, the reward increases if selected. Therefore, exploring a high interest rate has a higher worth than exploring a low interest rate. To promote exploration of higher interest rates, instead of using the confidence of UCB $\sqrt{\frac{2\log t_i}{C_i[b_{t,i}]}}$, we use adjusted confidence of $\sqrt{\frac{2\log t_i \delta[b_{t,i}]}{C_i[b_{t,i}]}}$, where δ is the relative ranking of the bid in the bid space.

In [7], updating the reward for every input context is studied, suggesting cross-learning. While we cannot use cross-learning directly because other institutions' bids correlate with the context, we partially adapt this idea by updating the reward for multiple bids at once using the property of the first price auction. After offering $b_{t,i}$ to the applicants, we observe whether R_i, C_i is updated similarly to the initialization process shown in section 4.1.

A pseudo-code for AutoInterest that combines these ideas is shown in Algorithm 2.

5 OFFLINE EXPERIMENT

To evaluate the effectiveness of the learning algorithm, we developed a framework that provides an auction setting where the bidder with the lowest interest rate wins. We test five types of bidding strategies: AutoInterest, UCB[4], ϵ-greedy[43], Actor-Critic algorithm[31] and RBP, on other generated RBP strategies. We consider RBP strategies fixed during the whole time period because the update frequency of RBPs is irregular in real life, and the period considered is relatively short. Fixed RBP is the main baseline to be compared as it is the most prevalent algorithm used in the UPL market. Considering each bid as possible arms, we compare AutoInterest with 3 common multi-armed bandit algorithms. In this setting, every institution has equal loan budgets, and the budget is subtracted only when a given institution's bid wins the auction. We test on various loan budgets, expressed as a fractional multiple of the total sum of applied loan amounts shown in 1. For example, a multiple of $\frac{1}{5}$ indicates the institution being able to fund $\frac{1}{5}$ of the whole UPL market. The reward on each bid is calculated by multiplying the loan amount by the difference between the bid and the expected default rate of the applicant. For example, with a loan application \$100, an expected default rate of 5%, a bid of 5.2% will yield a reward of \$0.2. We show experimental results of various bidding strategies against 10 RBP strategies on the Credit Default dataset and also backtest on the CB dataset. We open-sourced the framework to allow others to test the learning algorithms on their internal financial datasets. All numbers for rewards shown in the following figures and tables are in the units of 10,000 Korean Won, which is around \$8 US Dollars.

Table 1: Average Rank of the algorithm with various other institution's strategy with varying budget multiplier

Algorithm	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$
Credit Default						
AutoInterest	1.21	1.11	1.07	1.05	1.03	1.27
UCB	2.56	2.26	2.19	2.10	2.05	1.77
ϵ-greedy	3.79	3.5	3.22	3.07	3.04	2.96
Actor-Critic	2.45	3.15	3.52	3.82	3.96	4.12
Fixed RBP	4.97	4.99	4.99	4.96	4.93	4.88
CB						
AutoInterest	1.29	1.31	1.34	1.33	1.29	1.24
UCB	2.87	2.95	2.92	2.93	2.75	2.57
ϵ-greedy	3.69	3.63	3.65	3.6	3.35	2.93
Actor-Critic	2.15	2.11	2.09	2.14	2.61	3.26
Fixed RBP	5	5	5	5	5	5

Algorithm 2 AutoInterest

INPUT: p_i - the expected default rate of grade i; total budget $B = \rho T$; total variance tolerable $\sigma^2 = \nu T$; max interest rate biddable $\bar{b}$; step size of possible bid d; gradient update step for budget ϵ_1; gradient update step for variance ϵ_2
Let $B_i = [p_i, p_i + d, p_i + 2d, ..., \bar{b}]$ for each grade i; $C_i[b_i] = 0$, $R_i[b_i] = 0$ for $b_i \in B_i$; $\lambda_1 = 0$; $\lambda_2 = 0$;
For $t = 1, ..., T$:
  The bidder observe the credit grade a_t and the loan amount v_t
  t_i be sum of elements of C_i
  If $B < v_t$: continue;
  If $\exists\, z$ in C_{a_t} such that $z == 0$:
    Run Initialization with C_{a_t}, R_{a_t} and B_{a_t}
  else :
    Estimate probability of winning :
    $\tilde{G}_{a_t}[b_{a_t}] = R_{a_t}[b_{a_t}]/C_{a_t}[b_{a_t}]$
    Calculate Confidence :
    $\delta[b_{a_t}] = \sum_{b'_{a_t} \in B_{a_t}} 1\{b'_{a_t} < b_{a_t}\}/len(B_{a_t})$
    $Conf[b_{a_t}] = \sqrt{\frac{2\delta[b_{a_t}] \cdot \log t_{a_t}}{C_{a_t}[b_{a_t}]}}$
    Find the interest rate to bid:
    $b_{t,a_t} \in \arg\max_{B_{a_t}} \tilde{G}_{a_t}[b_{a_t}] \cdot (v_t \cdot (b_{a_t} - p_{a_t})$
    $-\lambda_1 \cdot v_t - \lambda_2 \cdot v_t^2 \cdot p_{a_t} \cdot (1 - p_{a_t}) + v_t \cdot Conf[b_{t,a_t}])$
    Update dual variables:
    $\lambda_1 = \lambda_1 - \epsilon_1(\rho - v_t \cdot \tilde{G}_{a_t}[b_{t,a_t}]);$
    $\lambda_2 = \lambda_2 - \epsilon_2(\nu - \tilde{G}_{a_t}[b_{t,a_t}] \cdot v_t^2 \cdot p_{a_t} \cdot (1 - p_{a_t}))$
    $\lambda_1 = \max(\lambda_1, 0); \lambda_2 = \max(\lambda_2, 0)$
    Update the estimated probability of selection:
    If selected :
      For all $b < b_{t,a_t}$:
        $C_{a_t}[b] = C_{a_t}[b] + 1,$
        $R_{a_t}[b] = R_{a_t}[b] + 1$
      $B = B - v_t$
    else :
      For all $b \geq b_t$:
        $C_{a_t}[b] = C_{a_t}[b] + 1$

5.1 Credit Default Dataset

The Credit Default dataset has 30,000 instances and 23 features aimed at studying credit card default payments in Taiwan. We trained two binary classifiers to predict default using random forest [29] and xgboost [15]. Loan applicants are divided into ten grades based on the output of the models, with each grade associated with an expected probability of default. We set aside 20% of the dataset as the test dataset, and of those, we restrict to grades that have an expected default of less than 20%, yielding about 3.3k samples. This hard upper limit to the interest rate adheres to regulations where interest rates above 20%

5.2 CB Data

In South Korea, there are multiple loan comparison services where users can get interest rates from different institutions in one sitting. During the instance of a loan inquiry from an applicant, financial institutions receive the applicant's credit data from the Korean credit bureau, NICE. The CB dataset is comprised of 300k applicants, and of those, about 48k have actually received a loan from 2023.01~2023.03. Five CSS systems developed by a South Korean online lending company, PFC Technologies and NICE CB Score, were chosen as possible credit scoring systems. Based on these 6 CSS models, 50 experts from PFC Technologies were divided into 16 teams, resulting in 914 distinct fixed RBP strategies. We compare AutoInterest against ten randomly selected fixed RBP strategies every run and report the results. Similar to the Credit Default setting, we create our test dataset by setting aside the test set and then restricting again on grades with less than 20% of the expected default rate - yielding a set of around 16k samples.

6 RESULT

We would want to point out that the maximum reward an institution can achieve depends on the strategy of other institutions. For example, if there exists an unrealistic institution with unlimited loan volume that bids 0 for all applicants, the best achievable profit is 0 - there exists no strategy that can compete with 0% interest.

Table 2: Mean & STD of Reward

Algorithm	Mean	Std
Credit Default		
AutoInterest	**13628.40**	2469.46
UCB	11167.02	2139.77
ϵ-greedy	10019.33	904.04
Actor-Critic	12078.91	1763.92
Fixed RBP	1826.33	2125.94
CB		
AutoInterest	**67428.39**	7070.63
UCB	35317.44	4457.41
ϵ-greedy	37572.89	2722.61
Actor-Critic	41827.69	1605.35
Fixed RBP	5207.51	4682.38

Similarly, if all other institutions bid high interest rates, the maximum reward the institution can achieve would be higher. Therefore, it is meaningless to compare the rewards of algorithms when the strategies of other institutions change. The importance is the relative position of the given algorithm compared to other algorithms in the same setting. We show two tables where table 1 shows the relative ranking of algorithms over the wide range of strategies of competing institutions, and table 2 shows the average reward of algorithms when the strategies of other institutions are fixed. For this section, we only consider the budget constraint instead and ignore the variance constraint. This is because budget is the main constraint in real life. In section 7.3, we show the case where we consider both variance and budget constraints.

6.1 Ranking of Algorithms

Each bidding algorithm was tested against the same ten randomly generated fixed RBP strategies. The ranking of each algorithm was then calculated by looking at the resulting profit. This experiment was repeated 100 times over six different budget multipliers, and the results are shown in Table 1 - the top half is tested on the Credit Default dataset, and the bottom half is on the CB dataset. We also show the graph of the cumulative reward for various budget multiples for one specific instance in Fig 2. While all online learning algorithms, including AutoInterest, are able to outrank the traditional fixed RBP strategy on both datasets, on the Credit Default dataset, AutoInterest ranks the highest for multipliers from $\frac{1}{10}$ up to $\frac{1}{5}$. Since AutoInterest considers the constraints at every step, the relative performance of AutoInterest degrades as the budget becomes a non-limiting factor. However, even the $\frac{1}{5}$ multiplier for a single institution is unrealistic - it would represent an institution having $\frac{1}{5}$ of the multi-billion dollar UPL market. A similar trend of AutoInterest's ranking going down with higher multipliers is also shown in the CB dataset; however, in this case, it is not enough for UCB to outrank AutoInterest. The reasoning for AutoInterest's dominance on the CB dataset stems from the fact that since there are more applicants, there is more time for the algorithm to catch up on its sub-optimal bidding during the earlier stages of exploration.

6.2 Reward of Algorithms

To test the reward of AutoInterest, we fixed the other institutions' strategies and instead shuffled the dataset and compared the mean and std of the reward against other algorithms. The result is shown in Table 2. In both the Credit Default and CB datasets, AutoInterest has the highest cumulative reward, and the fixed RBP is significantly weaker than other learning algorithms. Furthermore, the mean reward for AutoInterest, as compared to UCB's mean reward in the CB dataset, is much higher than that of Credit Default.

7 ADDITIONAL EXPERIMENTS

7.1 Bids of AutoInterest

The bids for grade 5,6 for both Credit Default and CB data can be seen in Fig 3. We only show grade 6 here since the default rate for grade 5,6 is relatively high; there's more room for variation in the bid values among the competing institutions. For the low multiplier of $\frac{1}{8}$, the algorithm can be more selective on its bid, resulting in higher overall bid values. One interesting behavior of AutoInterest can be seen on the budget multiple of $\frac{1}{8}$, where the bid seems to be oscillating back and forth between a high and low bid value. AutoInterest tries to maintain a constant rate of consumption. As the constant rate is smaller in lower budget settings, the algorithm changes its bid more drastically for each applicant t to balance reward maximization and budget use.

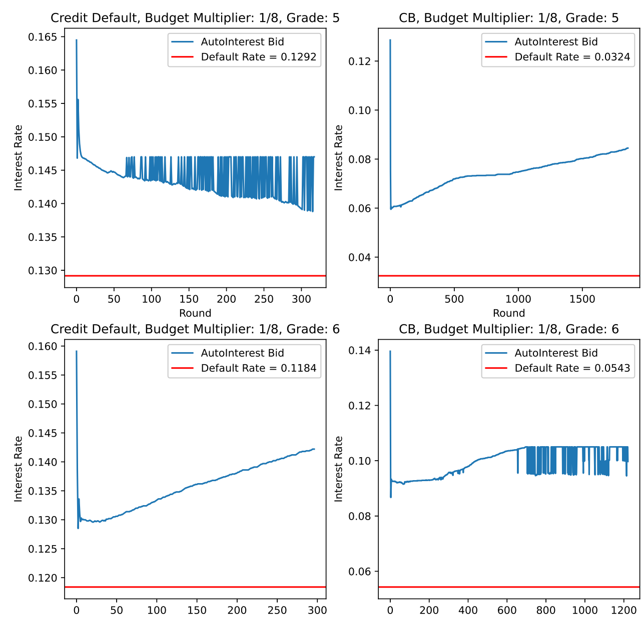

Figure 3: Bid values of AutoInterest for grade 6. After a short period of initial exploration on approximating the appropriate bid level, AutoInterest constantly tunes the interest rate over time

7.2 Without budget constraint

Figure 4 shows the effects of the learning rate on the budget constraint. Since the possible step size of the bid in the experiment is 0.001, ϵ_1 is chosen such that λ_1 is of a similar magnitude to the step size. While the cumulative reward is significantly lower in the case where the budget constraint is completely ignored, the rewards for ϵ_1 values ranging from 10^{-6} to 10^{-4} are similar. This shows that the performance of AutoInterest does not depend on ϵ_1.

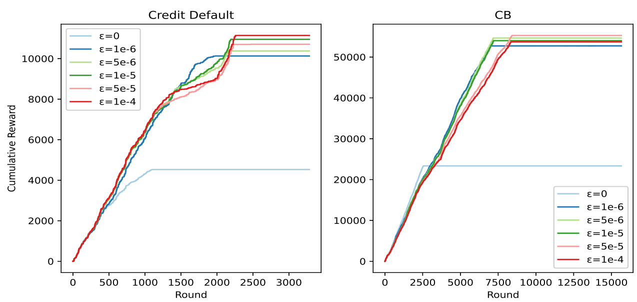

Figure 4: Effect of the learning rate of the budget constraint on cumulative reward

7.3 Variance Constraint

We consider the case where both the budget and variance constraints are considered hard limits. The variance constraint acts to add to the predictability of the algorithm's profit. While the variance constraint doesn't improve the cumulative reward as much when compared to the budget constraint, we can still see that non-zero values of ϵ_2 increase the cumulative reward.

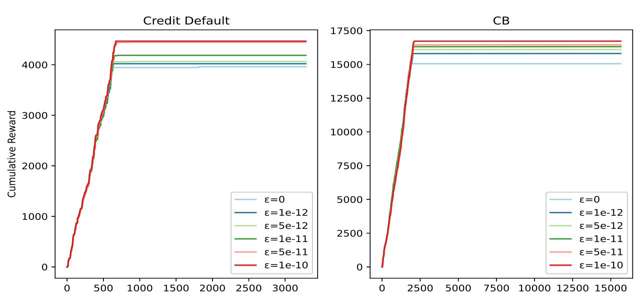

Figure 5: Effect of the learning rate of the variance constraint on cumulative reward

7.4 AutoInterest vs AutoInterest

We experiment the scenario where two financial institutions use AutoInterest. As shown in Figure 6 both institutions receive almost same amount of rewards.

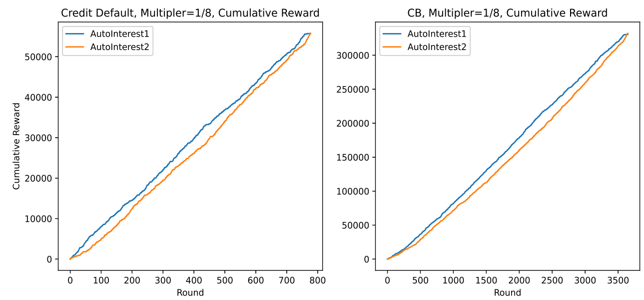

Figure 6: AutoInterest vs AutoInterest

8 CONCLUSIONS & LIMITATIONS

We have shown that the the UPL market can benefit from revisioning the interest rate strategies. We modeled the UPL market as an auction problem, marking a novel departure from an industry perspective. We have presented AutoInterest, an online learning algorithm that finds the optimal strategy for any opposing strategy. Furthermore, we provide a comprehensive framework for offline testing of the algorithm's efficacy. Our empirical analysis, conducted on both the widely-used Credit Default dataset against simulated RBP strategies and on the CB dataset featuring bidding strategies devised by human experts, demonstrates the effectiveness of AutoInterest.

Fixed RBP strategies have shortcomings due to human experts' reliance on heuristic decision-making. This approach is inherently suboptimal, as it fails to account for the complex dynamics of the UPL market. AutoInterest considers both the default rate of loan applicants and the competing institutions' strategy to optimize the bidding strategy. As a result, AutoInterest consistently outperforms RBP strategies in terms of profitability, a critical metric within the UPL market landscape. This underscores the potential of adopting optimization algorithms to enhance the market performance of financial institutions. Moreover, we compare AutoInterest with other online algorithms UCB and the ϵ-greedy algorithm. We demonstrated that AutoInterest outperforms other online algorithms under limited budgets. This setting reflects real-world dynamics where no single financial institution dominates the entire UPL market, highlighting the relevance and effectiveness of AutoInterest in practical settings.

For future work, we hope to analyze various properties of the UPL market further through the lens of the auction beyond finding the bidding strategy for profit maximization. For example, we can explore concepts such as the price of anarchy, the equilibrium, and robust design to gain a more comprehensive understanding of the UPL market dynamics. In addition, we could theoretically find the upper bound of cumulative regret of AutoInterest to study its properties. Furthermore, given AutoInterest's reliance on CSS to calculate expected default rates, it would be valuable to investigate the extent to which CSS performance impacts our algorithm's effectiveness.

REFERENCES

[1] Yeganeh Alimohammadi, Aranyak Mehta, and Andres Perlroth. 2023. Incentive Compatibility in the Auto-bidding World. *arXiv preprint arXiv:2301.13414* (2023).
[2] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2013. Learning prices for repeated auctions with strategic buyers. *Advances in Neural Information Processing Systems* 26 (2013).
[3] Jean-Yves Audibert, Sébastien Bubeck, et al. 2009. Minimax Policies for Adversarial and Stochastic Bandits.. In *COLT*, Vol. 7. 217–226.

[4] Peter Auer. 2003. Using confidence bounds for exploitation-exploration trade-offs. *J. Mach. Learn. Res.* 3, Nov (mar 2003), 397–422.
[5] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. 2018. Bandits with knapsacks. *Journal of the ACM (JACM)* 65, 3 (2018), 1–55.
[6] Santiago Balseiro, Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. 2021. Robust auction design in the auto-bidding world. *Advances in Neural Information Processing Systems* 34 (2021), 17777–17788.
[7] Santiago Balseiro, Negin Golrezaei, Mohammad Mahdian, Vahab Mirrokni, and Jon Schneider. 2022. Contextual Bandits with Cross-Learning. *Mathematics of Operations Research* (2022).
[8] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. 2020. Dual mirror descent for online allocation problems. In *International Conference on Machine Learning*. PMLR, 613–628.
[9] Santiago R Balseiro, Yuan Deng, Jieming Mao, Vahab S Mirrokni, and Song Zuo. 2021. The landscape of auto-bidding auctions: Value versus utility maximization. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. 132–133.
[10] Santiago R Balseiro and Yonatan Gur. 2019. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science* 65, 9 (2019), 3952–3968.
[11] Dimitri P Bertsekas. 2014. *Constrained optimization and Lagrange multiplier methods*. Academic press.
[12] Colin Campbell, Sean Sands, Carla Ferraro, Hsiu-Yuan Jody Tsao, and Alexis Mavrommatis. 2020. From data to action: How marketers can leverage AI. *Business horizons* 63, 2 (2020), 227–243.
[13] Matteo Castiglioni, Andrea Celli, and Christian Kroer. 2022. Online learning with knapsacks: the best of both worlds. In *International Conference on Machine Learning*. PMLR, 2767–2783.
[14] Nicolo Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. 2014. Regret minimization for reserve prices in second-price auctions. *IEEE Transactions on Information Theory* 61, 1 (2014), 549–564.
[15] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) *(KDD '16)*. ACM, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785
[16] Domenico Curcio and Igor Gianfrancesco. 2009. Bank loans pricing and Basel II: a multi-period risk-adjusted methodology under the new regulatory constraints. *Banks and Bank Systems* 4, 4 (2009), 56–66.
[17] Yuan Deng, Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. 2023. Multi-channel Autobidding with Budget and ROI Constraints. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*, Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (Eds.). PMLR, 7617–7644. https://proceedings.mlr.press/v202/deng23c.html
[18] Yuan Deng, Jieming Mao, Vahab Mirrokni, Hanrui Zhang, and Song Zuo. 2022. Efficiency of the first-price auction in the autobidding world. *arXiv preprint arXiv:2208.10650* (2022).
[19] Stylianos Despotakis, Ramamoorthi Ravi, and Amin Sayedi. 2021. First-price auctions in online display advertising. *Journal of Marketing Research* 58, 5 (2021), 888–907.
[20] Nikhil R Devanur and Sham M Kakade. 2009. The price of truthfulness for pay-per-click auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*. 99–106.
[21] Wendy Edelberg. 2006. Risk-based pricing of interest rates for consumer loans. *Journal of monetary Economics* 53, 8 (2006), 2283–2298.
[22] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American economic review* 97, 1 (2007), 242–259.
[23] Zhe Feng, Sébastien Lahaie, Jon Schneider, and Jinchao Ye. 2020. Reserve price optimization for first price auctions. *arXiv preprint arXiv:2006.06519* (2020).
[24] Zhe Feng, Christopher Liaw, and Zixin Zhou. 2023. Improved online learning algorithms for CTR prediction in ad auctions. In *International Conference on Machine Learning*. PMLR, 9921–9937.
[25] Zhe Feng, Swati Padmanabhan, and Di Wang. 2023. Online Bidding Algorithms for Return-on-Spend Constrained Advertisers. In *Proceedings of the ACM Web Conference 2023*. 3550–3560.
[26] Aurélien Garivier and Olivier Cappé. 2011. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *Proceedings of the 24th annual conference on learning theory*. JMLR Workshop and Conference Proceedings, 359–376.
[27] Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. 2021. Bidding and pricing in budget and roi constrained markets. *arXiv preprint arXiv:2107.07725* 8, 8.1 (2021), 3.
[28] Yanjun Han, Zhengyuan Zhou, and Tsachy Weissman. 2020. Optimal no-regret learning in repeated first-price auctions. *arXiv preprint arXiv:2003.09795* (2020).
[29] Tin Kam Ho. 1995. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1. IEEE, 278–282.
[30] Krishnamurthy Iyer, Ramesh Johari, and Mukund Sundararajan. 2014. Mean field equilibria of dynamic auctions with learning. *Management Science* 60, 12 (2014), 2949–2970.
[31] Volodymyr Kuleshov and Doina Precup. 2014. Algorithms for multi-armed bandit problems. *arXiv preprint arXiv:1402.6028* (2014).
[32] Tor Lattimore. 2015. Optimally confident UCB: Improved regret for finite-armed bandits. *arXiv preprint arXiv:1507.07880* (2015).
[33] Christopher Liaw, Aranyak Mehta, and Andres Perlroth. 2022. Efficiency of non-truthful auctions under auto-bidding. *arXiv preprint arXiv:2207.03630* (2022).
[34] Francisco Louzada, Anderson Ara, and Guilherme B Fernandes. 2016. Classification methods applied to credit scoring: Systematic review and overall comparison. *Surveys in Operations Research and Management Science* 21, 2 (2016), 117–134.
[35] Anton Markov, Zinaida Seleznyova, and Victor Lapshin. 2022. Credit scoring methods: Latest trends and points to consider. *The Journal of Finance and Data Science* (2022).
[36] Jincheng Mei, Zixin Zhong, Bo Dai, Alekh Agarwal, Csaba Szepesvari, and Dale Schuurmans. 2023. Stochastic Gradient Succeeds for Bandits. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*, Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (Eds.). PMLR, 24325–24360. https://proceedings.mlr.press/v202/mei23a.html
[37] Grzegorz Michalski. 2009. Inventory management optimization as part of operational risk management. *Economic computation and economic cybernetics studies and research* (2009), 213–222.
[38] Paul Robert Milgrom. 2004. *Putting auction theory to work*. Cambridge University Press.
[39] Carola Müller, Ragnar Juelsrud, and Henrik Andersen. 2021. Risk-based pricing in competitive lending markets. *Available at SSRN 4120047* (2021).
[40] Ambika Nair and Eldar Beiseitov. 2023. *The Role of Fintech in Unsecured Consumer Lending to Low-And Moderate-Income Individuals*. Technical Report. Federal Reserve Bank of New York.
[41] Alessandro Nuara, Francesco Trovò, Nicola Gatti, and Marcello Restelli. 2022. Online joint bid/daily budget optimization of internet advertising campaigns. *Artificial Intelligence* 305 (2022), 103663.
[42] Aleksandrs Slivkins. 2019. Introduction to Multi-Armed Bandits. *Foundations and Trends® in Machine Learning* 12, 1-2 (2019), 1–286. https://doi.org/10.1561/2200000068
[43] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.
[44] Qian Wang, Zongjun Yang, Xiaotie Deng, and Yuqing Kong. 2023. Learning to Bid in Repeated First-Price Auctions with Budgets. In *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 202)*, Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (Eds.). PMLR, 36494–36513. https://proceedings.mlr.press/v202/wang23ao.html
[45] Jonathan Weed, Vianney Perchet, and Philippe Rigollet. 2016. Online learning in repeated auctions. In *Conference on Learning Theory*. PMLR, 1562–1583.
[46] Robert M Weiss and Ajay K Mehrotra. 2001. Online dynamic pricing: Efficiency, equity and the future of e-commerce. *Va. JL & Tech.* 6 (2001), 1.
[47] I-Cheng Yeh and Che hui Lien. 2009. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Syst. Appl.* 36 (2009), 2473–2480. https://api.semanticscholar.org/CorpusID:15696161
[48] Cristiano Zazzara and Iftekhar Hasan. 2006. Pricing risky bank loans in the new Basel II environment. *Bank of Finland Research Discussion Paper* 3 (2006).
[49] Li Zhou. 2015. A survey on contextual multi-armed bandits. *arXiv preprint arXiv:1508.03326* (2015).

A EXPERIMENT DETAILS

A.1 Credit Default CSS

The Credit Dataset was split into 72% train, 8% valid and 20% test set. The train and valid sets were used to train the model and to create the bins for grades. The test set was solely used to evaluate AutoInterest on figures 1 through 5. Two models using random forest classifier from sklearn and xgboost classifer from the xgboost library were trained with default parameters. Columns x12 through x23 were fit by sklearn's Standard Scaler on the train dataset. All train, valid and test were then transformed with the scaler. Each competing institution was given a 50-50 chance of using the xgb model or the rf model. To meet regulatory requirements regarding interest rates in South Korea, grades with a default probability of 20% on either of the models are omitted. The default probability on each grade is shown in table 3.

Table 3: Default Probability and Count of CSS

	RF Model		XGB Model	
Grade	Default Probability	Count	Default Probability	Count
1	0.0493	609	0.0617	540
2	0.0897	621	0.0716	654
3	0.0827	717	0.0969	626
4	0.1063	470	0.1152	552
5	0.1429	664	0.1292	640
6	0.1429	682	0.1184	626
7	0.2379	449	0.2057	590
8	0.2570	545	0.2713	557
9	0.4229	592	0.4492	587
10	0.6900	651	0.7225	628

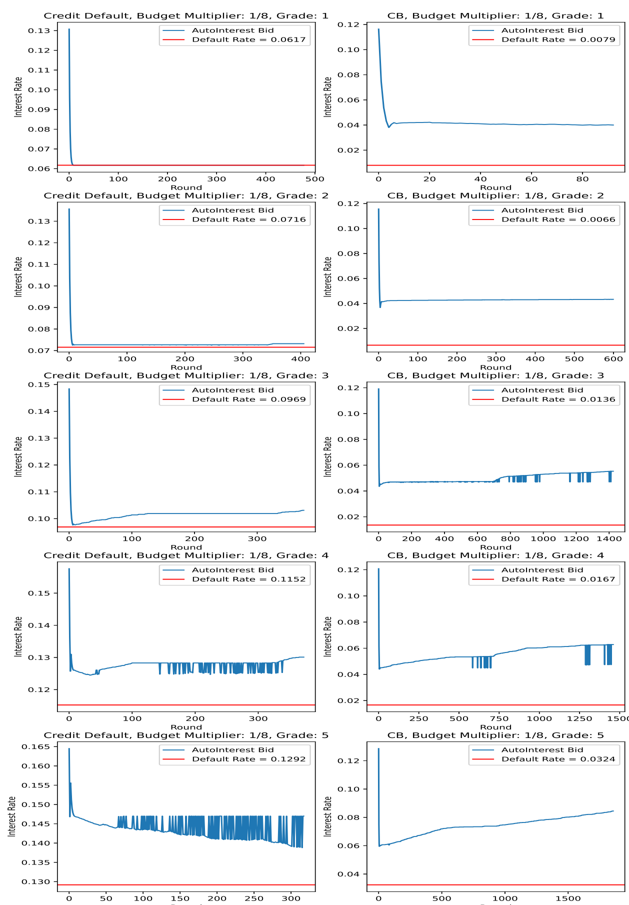

Figure 7: Bids of AutoInterest on all grades run on Credit and CB datasets.

A.2 Risk Based Pricing

Risk based pricing, X_t is sampled from a beta distribution with parameters $\alpha = 0.1$ and beta $\beta = 0.90$ which centers the beta distribution with mean $\mu = 0.1$. The values for the interest rate, b_{t,a_t} is then calculated as $b_{t,a_t} = (1+p_{a_t}) \cdot X_t$, where a_t is the credit grade, p_{a_t} is the default rate at a given grade, and $X_t \sim Beta(\alpha, \beta)$.

A.3 UCB and E-Greedy

Both UCB and ϵ-greedy algorithms were initialized with initialization shown in 4.1. The epsilon value of ϵ-greedy was 0.5.

A.4 Cumulative Reward of Algorithms

Cumulative reward in figure 2 shows single run test results on budget multiples of 0.1, 0.2, and 0.3 on both the Credit Default dataset and CB dataset. The epsilon value for the budget constraint is $2 \cdot 10^{-6}$.

A.5 Ranking of Algorithms

The experiment result shown in table 1 was done on budget multipliers of $\frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}$. The variance multiplier is not used here. The budget is the total sum of loan amounts multiplied by the given multiplier, where each loan amount, v_t is sampled from a normal distribution of $\mu = 2700$ and $\sigma = 500$ which resembles the loan amount values from the CB dataset. Each multiplier is tested on 100 different seeds on the random generator. At the beginning of each test, the ten competing institutions generate interest rates on grades 1 through 6 with the RBP strategy, and each institution is given a budget. On every round, every bid is according to their strategy, and the winning institution receives a reward of $(b_{t,a_t} - p_{a_t}) \cdot v_t$. The winning institution subtracts the loan amount from the remaining budget when the budget runs out.

A.6 Reward of Algorithms

The experiment result of table 2 shows the reward and std of each algorithm in a fixed setting - the RBP strategy of the institutions is fixed with a budget multiplier of $\frac{1}{8}$ and the variance constraint is not used. The changing part of the experiment is the order of the number of applicants; with each iteration, the order of the applicants is shuffled.

A.7 Bids of Interest

The bid values shown in figure 3 are in the setting of budget multiplier $\frac{1}{8}$ on grade 6 for the Credit Default and CB datasets. The bid values for which AutoInterest has made a bid, whether it was a win or not, are shown in the figure.

A.8 AutoIntereset vs. AutoInterest

The result shown in figure 6 are in the setting of budget multiplier of $\frac{1}{8}$.

A.9 Without Budget Constraint

The mean and the std of the cumulative reward in the setting of the budget multiple of $\frac{1}{8}$ is shown in figure 4.

A.10 Variance Constraint

The effects of the variance constraint shown in figure 5 have the budget multiplier at $\frac{1}{8}$ and variance multiplier at $\frac{1}{16}$. The variance multiplier was purposefully chosen to be lower than the budget multiplier such that the variance multiplier would become the bottleneck for every institution. In this setting, the total allowable variance is calculated as a ratio of the total variance. The total variance is calculated as $\sum_T (v_t^2 \cdot p_{a_t} \cdot (1 - p_{a_t}))$.

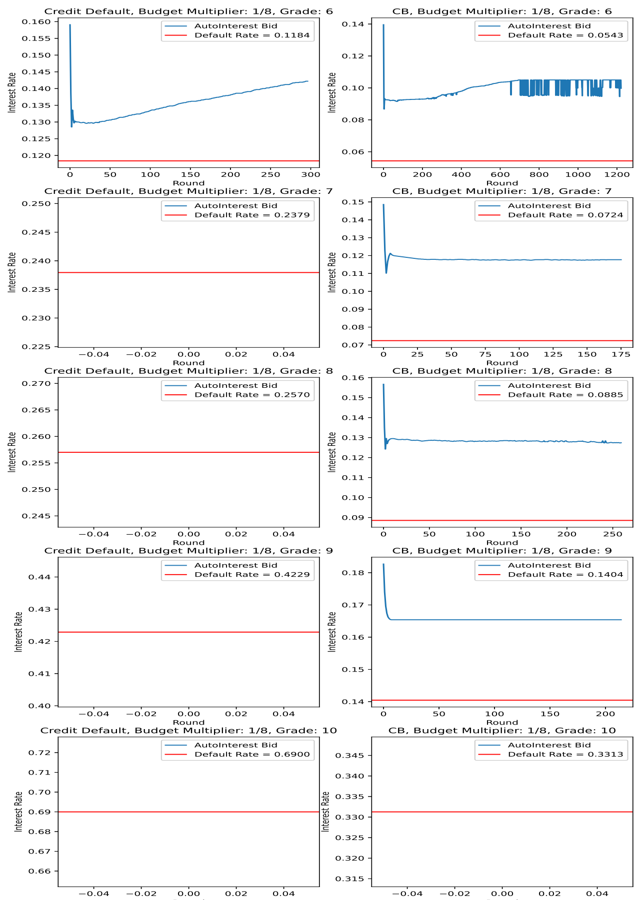

Figure 8: Bids of AutoInterest on all grades run on Credit and CB datasets.

A.11 Bid of Grade

We include graph of bids of other grade not included in section 7.1