COMP0124 - Multi-Agent Artificial Intelligence Group Assignment (Group 04)

Kacy Chou
University College London
15007335
kacy.chou.15@ucl.ac.uk

Faizal Masrol
University College London
18099094
faizal.masrol.18@ucl.ac.uk

David van Rooij
University College London
18136657
david.rooij.18@ucl.ac.uk

1. INTRODUCTION

Since its emergence in circa 2009 [16], Real-Time Bidding (RTB) has grown as a method to buy and sell display advertising placements in real time over the internet. RTB uses the auction pricing mechanism where it allows placements, referred to as impressions, to be purchased via bids collected when they are made available. The process of RTB through the auction takes place between ad exchange that connects publishers and advertisers, that often happens within a fraction of a second [16]. The highest price bid by an advertiser would then win the impression followed by the relevant creative message would be displayed in the winning slot. Typically, RTB exchanges employ second-price sealed-bid auction model.

RTB leverages computer algorithms to allow advertisers to buy and sell ads impressions through auctions in real-time. Advertisers, often through demand-side platforms (DSP), purchase impressions that best suit their purpose or objective. If the advertiser wins the auction, the ads are instantly displayed on the publisher's website. The RTB ecosystem could be briefly visualised in Figure 1 [19].

One central challenge in RTB is the bid optimisation problem [8]. The goal is to find the best bid price for every impression or bid request, to maximise a Key Performance Indicator (KPI) e.g. the number of clicks of the display advertising given a budget constrained. One methodology that is widely employed to maximise a range of KPIs is estimating the click-through rate (CTR) by analysing historical behavioural profiles. This predicted CTR is used in deciding whether or not to bid on a certain ad request and, alternatively, used in determining the height of a bid. The CTR can be estimated based on features describing the context (e.g. slot height, website URL etc.) and the user (e.g. browser, city etc.). Note that the CTR can be expanded into conversion rates when additional data is available. Another challenge that emerges is the limited knowledge gain, historical data only provides insight into the winning bid and prevents developing a better understanding of bidding strategies of other agents in the system. This report aims to help advertisers develop an optimal bidding strategy to place their advertisements (ads) online in an RTB system in a single and multi-agent environment.

To maximise a KPI, two fundamental components that need to be addressed are (a) optimising a model to accurately predict the probability an impression is going to be clicked and (b) how an optimal bidding strategy can be formulated given the predicted probability. In this report, four different models for estimating the CTR are considered and four different types of bidding strategies are evaluated;

- **CTR estimation models:** *logistic regression, gradient boosting decision tree (XGBoost), and deep neural network.*
- **Bidding strategies:** *basic bidding strategies i.e. constant bidding and random bidding, linear bidding, non-linear bidding and multi-agent bidding strategies.*

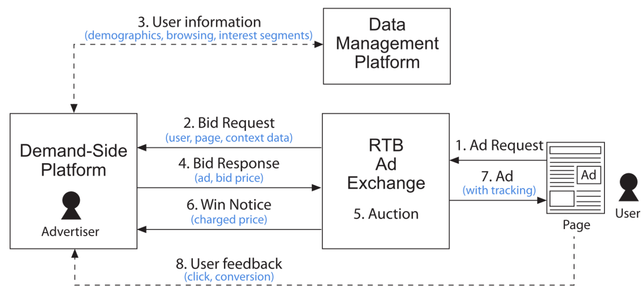

Figure 1: Illustration of a bidding engine in RTB display advertising.

Optimising bidding strategies for advertisers could potentially generate better returns and hence, that would be the main objective of this report.

The coming sections will begin with a review of related work on the field of RTB, followed by the approaches taken to address the fundamental components CTR estimations and optimise bidding strategies, analyses of the results and lastly the report will give its conclusion and potential future work.

2. RELATED WORK

Online display advertising is one of the major sources of revenue for most online publishers and the market size in [17] the United States alone has exceeded $40 billion in 2017 [12]. Traditionally, display advertisements were sold through pre-negotiated contracts between advertisers and publishers.

Nonetheless, the rise of RTB in 2009 has quickly dominated the market and disrupts online advertising by shifting the trend from 'ad-slot buying' to 'target audience buying'. RTB currently accounts for more than one-third of the display advertising market, which is predicted to exceed $65 billion by 2020 [4]. As a result, research on predicting CTR, and optimal bidding strategies, are highly popular among advertisers and DSPs [5, 6, 10].

For instance, Claudia et al. (2012) [10] proposed a bid-optimisation approach using supervised learning, which is shown to be effective in matching impressions to the relevant audience.

The CTR estimation often referred to as predicted CTR (pCTR) is a broadly researched subject in online display advertisement. For instance, linear or logistic regression models [1], boosted trees [14], deep neural networks [2, 7], as well as factorisation machines [9] are different models that were practised over pCTR estimations. One of the prevalent estimators are the regression models due to its efficiency and simplicity. Apart from that, although it's hard to get an interpretation, deep neural networks could be used to achieve good results. On the other hand, as feature importance could be generated, the boosted trees models can provide a direct interpretation of the model built.

In order to evaluate the performance of the pCTR models, the 'area under the curve' (AUC) measure is typically used amongst scholars. This evaluation metric is used to measure classification prediction accuracy [3]. AUC scores, which ranges between 0 and 1, is calculated from computing the area under the Receiver Operating Characteristic ('ROC') plot; which plots the true positive rate (TPR) against the false positive rate (FPR) of the predicted classes. Obtaining AUC ROC of between 70% to 80% has been reported in some other studies on display ads data[18].

Besides pCTR, bidding strategy is another fundamental component. Linear bidding strategies have proven to perform better than the basic bidding strategies [19]. Differently than the basic bidding; which takes no prior information to compute bids, the linear bidding is a strategy that utilises predicted CTR based on historical data. A linear function would then be applied to the pCTR to generate bid prices.

Although it has been studied less, the non-linear bidding strategy has shown to improve results compared to the linear bidding [19]. The bids generated from a non-linear bidding strategy is formed by considering the winning probability of an impression and its bid price. This implies the value of bids is also affected by other non-linear parameters beyond the pCTR learned from the data.

3. METHODOLOGY

This section provides an overview of the approaches utilised in answering the research questions. First, some preliminary data exploration is conducted to gain more insight into the data. Sequentially, the processing and feature engineering steps are described. Followed by different methodologies used for predicting CTR and how the performance of these predictions can be evaluated. Once these models have been established it is possible to turn these estimations into different bidding strategies that will be described in the last section. Again, the performance of the strategies can be evaluated using predefined KPI's in both a single agent and multi-agent setting.

3.1 Data Exploration

The data used in this analysis is an adapted version of the iPinYou dataset. The data available consisted of pre-split training, validation and testing sets. The datasets share the set of features related to impressions, except for the testing set does not contain the features 'bidprice', 'payprice' and 'click'.

Table 1: Overall Statistics for Each Provided Dataset

Data Set	Clicks	Imp	Spend	CTR	CPM	CPC
Train	1,793	2,430,981	189,984	7.38e-04	78.15	105.96
Val	202	303,925	23,777	6.65e-04	78.23	117.71
Test	n/a	303,375	n/a	n/a	n/a	n/a

3.1.1 *Input Features*

The dataset contains features describing the user and contextual information regarding the slot. These features include the day of the week, time of day, user's device types, ads exchange, details of the web page, formatting of slots and the creative message itself. All features in the datasets that are related to money (e.g., 'bidprice', 'payprice' and 'slotprice') are in the currency of CNY and the unit is Chinese fen x 1000; which refers to the cost-per-mille (CPM) pricing model.

3.1.2 *Display Advertising Metrics*

Through analysing historical data regarding the bid requests' that won ads auction an overall basic metrics can be computed as shown in Table 1. It can be observed that the CTR, CPM and CPC of train and validation sets show similar trends. Moreover, it is paramount to note the number of clicks and non-clicks are quite imbalanced; where the ratio of clicks ('click' = 1) to non-clicks ('click' = 0) is 1355:1. The challenge that arises given the presence of imbalance is addressed in Section 3.2 of this report.

3.1.3 *Statistics of User Feedback*

Figure 2 shows some relevant statistics of user feedback on all nine advertisers available in the training dataset. It shows the average value of CTR against the features the day of the week ('weekday'), time of day ('hour'), advertisers, publishers' slot sizes ('adsize') and ad exchanges. It can be observed that the average CTR peaks on Wednesdays and are the lowest on Sunday. While for the time of day, the CTRs fluctuates frequently and shows highest rates towards the end of the day. Amongst the nine advertisers, advertiser 2997 has a substantially higher average CTR. Lastly, smaller slot sizes get more clicks on average compared to larger slot sizes while different ad exchanges have relatively different average CTR within the dataset.

3.1.4 *Probability Distributions for Payprice*

Figure 3 shows the probability distribution function (PDF) and cumulative distribution function (CDF) graphs for the payprice found in training dataset. The CDF shows that almost all payprice is less than 100 CNY. For instance, 80% of the prices paid are less than 115 CNY. These charts could hence be used to guide the bidding strategy as described in this report.

3.2 Prepossessing and Feature Engineering

The first step of building a machine learning (ML) model to predict pCTR is to clean and extract the relevant features from the dataset. The available data is transformed into categorical features utilising a one-hot approach. In order to

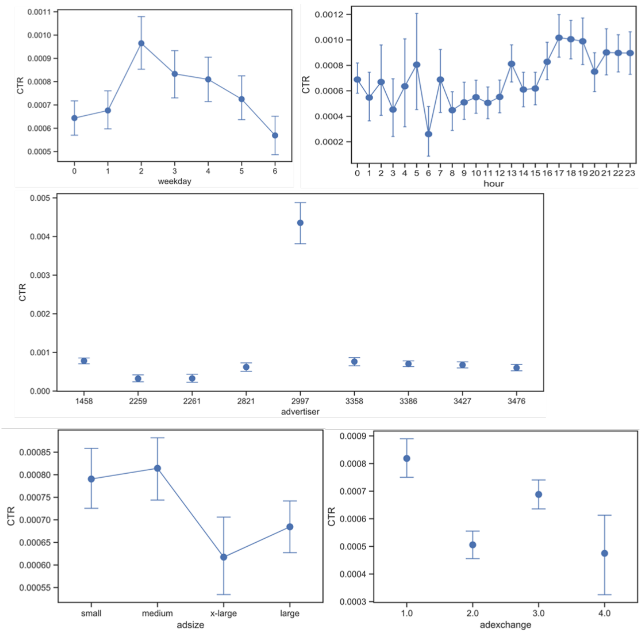

Figure 2: CTR distribution against selected features for training dataset.

reduce dimensionality, the entropy was calculated for all the features. The features with the highest entropy, 'userid', 'url', 'domain', 'slotid', and 'IP', were excluded from further analysis. Upon closer examination of the remaining features, it became apparent that 'adexchange' contained a substantial body of missing values. Several different approaches have been considered in how to address the missing values and it was ultimately decided that given the importance of this feature all missing values are assigned an additional class indicating 'not know' as opposed to allocating these to any of the existing classes. The same approach was used for addressing the missing values in the feature 'keypage'.

Additionally, some of the remaining data were transformed or used to derive new features. 'slotprice' was binned into bins $[1, 10]$, $[11, 50]$, $[51, 100]$ and $[101, +\infty)$ [20]. The features 'slot height' and 'slot width' were used to derive 4 categorical bins describing the ad size. These bins are constructed to have a similar number of observations per bin. Moreover, features such as 'os', 'apple', 'mobile device' and 'browser' were derived from the feature 'useragent'.

In addition, as aforementioned, the existence of the data imbalance needs to be addressed. An imbalanced dataset, indicates the two classes ('clicks and no clicks') are not equally represented. Especially rudimentary models are more accustomed to detect the majority class and result in a biased classification if the imbalanced dataset were not preprocessed. To reduce the imbalance two steps are used. First, an arbitrary subset of 100,000 items is selected of the non-click class. Second, the data is then balanced in such a fashion to reach a ratio of 10:1 by using the 'SMOTE' package. The used method combines oversampling the minority class ('no clicks') and under-sampling the majority class ('clicks') to achieve numerical stability. Additionally, Tomek [13] links were removed to increase the separability of the classes. Note that removing these links is most beneficial for models that rely on the geometric position of

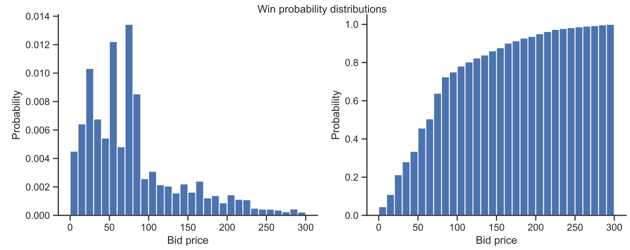

Figure 3: Probability Distributions for Payprice

features in an n-dimensional space. Therefore, Tomek links were only removed in the dataset used for training the Logistic Regression and the data used for the Neural Network and XGBoost is limited to the first balancing step. Note that typically further steps to calibrate performance evaluation after resampling data are needed, however, since ROC AUC scores are unaffected by calibration procedures since it is a measure of separability and not a probability this step is omitted.

3.3 CTR Estimation

To estimate pCTR as accurate as possible, three different machine learning approaches are considered of which the logistic regression (LR), XGBoost (XGB) and the deep neural network (NN). These models have also been used in other literature. Sequentially these models have been combined by using some rudimentary ensemble techniques. The models are trained to reduce log loss whilst the performance evaluation of the pCTR for each model is measured using AUC ROC. A semi-manual grid search was done on each model to tune the hyper-parameters of these models, in an attempt to maximise AUC ROC.

3.3.1 Logistic Regression

The LR is a supervised classification algorithm which returns a binary outcome($y(i) \in 0, 1$), i.e. 0- no click, 1-click. The logistic regression algorithm uses the sigmoid function to squash a linear equation with independent predictors into a probability, in other words, a predicted value in the range [0,1].

A logarithmic loss function is then used to calculate missclassifications of the binary outcome. The aim is to find the optimal weight vector theta using gradient descent to classify the training data to minimise the loss function and maximise the accuracy of the classification of clicks. In addition, L2 regularisation was used to minimise over-fitting. Some features were combined in the preprocessing step such as os and weekday to add higher dimensional features that are otherwise not evaluated by the LR.

3.3.2 Gradient boosting decision tree (XGBoost)

XGB is a tree-based model, which implements the gradient boosting decision tree algorithm. In short, the algorithm uses an ensemble technique, which produces a strong classifier in the form of a collection of weak learners (decision trees). Parameters to be tuned using the XGBoost algorithm include the learning rate and the number of iterations, and regularisation terms. L1 regularisation is used, since features have already been selected, hence encouraging sparsity is unnecessary. As a result, high lambda and low alpha would be the most effective in this case.

3.3.3 Deep Neural Network

Our final approach is to predict pCTR using a multi-layer neural network using the pyTorch module. The NN consisted of 4 hidden layers with 2048 nodes each, the RELU activation function is used to introduce non-linearity to the model. The 2-dimensional vector was obtained using the soft-max function, representing click and no-click class. The Adam optimizer was used, and a 0.5 dropout rate is used to minimise overfitting.

3.3.4 Ensemble

To reduce variability, two rudimentary ensemble techniques, average and weighted average, were used to improve and combined the three models discussed before.

3.4 Evaluation Bidding Strategies

In addition, the performance evaluation of the pCTR and the performance of the bidding strategy based on the pCTR can be evaluated through identifying a KPI. The main evaluation method used to evaluate the performance of the models and strategies built is based on the number of clicks. Alternatively, the click-through rate has been identified as secondary KPI.

3.4.1 Number of Clicks

Clicks can be defined as the action where a user has taken their cursor, placed it on the display ads image and clicked their mouse on that display ads as they surf around a particular website. The accumulated number of this action is defined as the number of clicks.

3.4.2 Click Through Rate

The click-through rate (CTR) is the probability of click per impression. It is crucial to optimise the accuracy of the pCTR, as the pCTR will be fed into the linear and non-linear bidding mechanisms to derive a bid price that maximises the KPI, in this case, the number of clicks.

3.5 Bidding Strategies

Note that bids are subjected to a budget constraint that needs to be enforced. Mind that the bidding environment is a second-price sealed-bid auction. Bids after the budget is depleted are not considered in determining the winner. Bidding strategies have been built to satisfy two winning criteria; which are determined:

(i) if $bid \geq payprice$ (Single-agent) and;

(ii) if $bid \geq payprice \& otherSubmittedBids$ (Multi-agent).

The performance of the bidding strategies hereafter was evaluated according to the identified KPI's.

3.5.1 Constant Bidding

Constant bidding is a bidding strategy where an agent bids a fixed price on all impressions. A constant bidding strategy can be used as a benchmark for other strategies. The base bid includes prices that range between 0 and 302 i.e. minimum and maximum(+10%) price paid found in the training set. For each impression, each base bid was compared to the pay price of each impression, and if the bid price is higher, while the budget has not been fully depleted, the impression was won. The budget of 6,250 CNY will be reduced by the payprice of the winning impression. The optimal constant bid was obtained by taking the base bid that achieved the maximum number of clicks across the validation set.

Table 2: Notations and Description

Notation	Description
$b()$	The bidding strategy
b_0	The base bid
θ	The predicted CTR
θ^*	The average CTR
λ	The Lagrangian multiplier
c	constant value
B	The campaign budget
T	The number of bid requests

3.5.2 Random Bidding

Two strategies for random bidding were conducted according to the two winning criteria. The first strategy is to find an upper and lower bound from random bid values chosen from a random bidding range. The optimal bounds were found using the same strategy used as mentioned in constant bidding i.e. satisfying winning criteria (i). Secondly, the strategy of considering 50 and 100 number of agents bidding on the same list impressions in the validation set was studied. This strategy was built on top of having the bid values randomly chosen from the same homogeneous random bid range as the other agents. The optimal bounds were also reported according to the winning criteria (ii).

3.5.3 Linear Bidding

The first approach to optimise bidding strategy is by implementing the linear bidding, with the objective to predict an optimal bid price. The linear bidding strategy can be characterised as the following equation:

$$b_{\text{linear}}(\theta) = b_0 \cdot \frac{\theta}{\theta^*} \quad (1)$$

where there exists a linear relationship between the bid price and pCTR, 'basebid' is a parameter to be tuned to maximise the number of clicks given the budget constraint and the pCTR is estimated by the machine learning models in Section 3.3.

3.5.4 Non-linear Bidding

To further explore a broader range of bidding models, non-linear models were experimented on. When budget-constraints and auction volumes are not considered, the truth-telling strategy [15] is the optimal strategy for second-price sealed auction. Truth-telling in auction referrers to the notion agents bid exactly their intrinsic value. Nonetheless, in a single campaign with a budget of 6,250 CNY and limited bid requests, the dominant strategies are not truth-telling. Our strategies to address this is to use several bidding strategies that incorporate budget constraints and constrained bid requests [15]. Zhang et al. (2014) [15] proposed two optimal real-time bidding (ORTB) in first-price auction, in the form of:

Optimal Real-time Bidding 1 (ORTB1)

$$b_{\text{ORTB1}}(\theta) = \sqrt{\frac{c}{\lambda}\theta + c^2} - c \quad (2)$$

m **Optimal Real-time Bidding 2 (ORTB2)**

$$b_{\mathrm{ORTB2}}(\theta) = c\left[\left(\frac{\theta + \sqrt{c^2\lambda^2 + \theta^2}}{c\lambda}\right)^{\frac{1}{3}} - \left(\frac{c\lambda}{\theta + \sqrt{c^2\lambda^2 + \theta^2}}\right)^{\frac{1}{3}}\right] \tag{3}$$

Additionally, the budget-constrained method was explored as a second-price bidding strategy [15]. It essentially is a function that maximises θ which is the pCTR across the auctions over the total number of bid requests within the given budget. The analytic form of the function is shown in equation (4).

Budget-Constrained (BC) Equation

$$b_{\text{Budget-constrained}}(\theta) = 2\,\theta^2 \sqrt[3]{\frac{Bc^2}{T}} \tag{4}$$

3.5.5 Multi-agent Bidding

In section 3.5.1 to 3.5.4, different bidding strategies are discussed with the assumption that the market data is stationary, where each impression is won when our predicted bid price is larger than the pay price in the validation set, i.e. the probability distribution is static. However, a different bidding strategy is required when multiple agents are involved in the auction. In this scenario, one must take into account both the auction environment and more importantly, the interactions among other bidders. Any changes in one bidder's strategy would impact other bidders' strategies. Note that agents are showing behaviour that is consistent according to mixed-strategy Nash equilibrium strategies. An agent has no incentive to deviate from their strategy given the other agents' strategies, as deviating will not lead to a payoff increasing outcome.

This can be framed as a non-cooperative game with incomplete information since there is no knowledge of the other bidders' values, nor about their strategies. Hence, a game-theoretic approach was taken, where multiple agents are simulated and agents' interactions are modelled. Since there is no way of knowing the bidding behaviour of other agents, simulating other agents in a competitive environment is an approach to cope with this problem in an attempt to estimate other agents' behaviour. All agents in the actual environment have a limited budget of 6,250 CNY, and the game may only reach a sub-optimal equilibrium. For instance, if we were risk-taking and bid aggressively at early auctions, we would run out of budget quickly and would result in a low matching efficiency of advertisers and consumers in a later stage of the auction. Therefore, multiple agents with the same budget constraint are created to simulate the environment and their interactions may provide us with valuable insights into how to bid in a multi-agent environment to maximise the number of clicks.

Combined Bidding Strategy

As a result, a script `multi_agent_simulator.py` is developed to simulate the non-stationary competitive environment, which will enable us to develop the optimal bidding strategy to maximise clicks against the pay price and the other agents. The script consolidated 37 different bidding strategies, including the constant and random bidding strategies from sections 3.5.1 and 3.5.2, linear bidding strategies from section 3.5.3 and the non-linear bidding from section 3.5.4; using pCTR estimations from the LR, XGB, and NN as inputs as discussed in Section 3.3. After all bids were collected in the environment, the winning criteria (ii) given a budget-constrained for each agent was used to determine the winners. Note that there is a discrepancy between the number of agents simulated and the number of agents in the actual environment (the other teams that are competed against). Therefore, the budget per simulated agent is corrected in such a fashion the total amount of money available both environments are the same. This environment was used to find the optimal bidding strategies given the presence of other agents in the system. It is also worth noting that the agents are simulated in such a fashion to resemble other agents to the best of our ability. A novel opportunistic strategy has also been proposed, which is to bid a constant price when pCTR is above a predefined threshold. An auction was then held among these 36 simulated agents to mimic how agents may interact in the environment.

Sequentially, the new optimised strategies are added to the environment following the assumption that the submitted bids have to compete against agents who have developed such strategies. Finally, this environment was then used to optimise the final strategy. This final strategy was used to determine the bids on the test set for criteria (ii).

Table 3: Constant Bidding Strategy: Evaluation of Optimal Bid from Validation Set

Strategy	Opt. Bid	Clicks	Spend	CTR	CPM	CPC
Const.	78	67	6,250	4.58e-04	42.71	93.28

Table 4: Random Bidding Strategies: Evaluation of Optimal Bounds from Validation Set

Strategy	Opt. Bounds	Clicks	Spend	CTR	CPM	CPC
Rand.	[69,99]	65	6,190	4.61e-04	43.89	95.14
Rand. 50 Agents	[267,297]	14	6,250	5.94e-04	265.70	446.44
Rand. 100 Agents	[269,299]	14	6,250	5.92e-04	264.09	446.44

4. RESULTS & DISCUSSION

4.1 Basic Bidding Strategies

4.1.1 Constant Bidding

As seen in Table 3, the optimal constant bid was found at 78 CNY on the validation set, with budget fully exhausted. The number of clicks found at this bid price is 67. Other performance metrics can also be observed in the same table. This result can be used as a baseline to compare more sophisticated models and strategies.

4.1.2 Random Bidding

Single agent. The results of optimal bidding bounds for both criteria can be seen in Table 4. To satisfy the first criteria, the strategy found an optimal bound of [69, 99]. This is in line with the optimal bid found in the constant bidding strategy. This setup achieves 65 number of clicks, with almost all of the budget spent.

Apart from having a similar number of clicks, the other

Table 5: Bidding Strategies: Model Evaluation

	Model	b_0/c	λ	Clicks	CTR	Spend	CPM	CPC
Linear	LR	3.60		149	0.0013	6,250	53.73	41.95
	XGB	213		164	0.0013	6,179	45.91	37.68
	NN	75		162	0.0018	6,204	69.01	38.30
ORTB1	LR	15	2.5e-05	151	0.0012	6,126	46.46	40.57
	XGB	80	1.0e-06	165	0.0011	6,230	43.37	37.76
	NN	20	1.0e-06	164	0.0017	6,169	62.66	37.62
ORTB2	LR	90	9.0e-05	151	0.0012	6,164	49.18	40.82
	XGB	80	1.0e-06	164	0.0011	6,250	42.64	38.11
	NN	50	2.0e-06	162	0.0016	6,210	61.19	38.33
BC	LR	15		151	0.0012	6,126	46.46	40.57
	XGB	1.37e+07		164	0.0012	6,043	45.56	36.84
	NN	4e+06		150	0.0017	6,250	69.39	41.67

performance-based metrics achieved by the random bidding strategy (for criteria (i)) shows similar results as the one previously achieved in constant bidding strategy.

Multi agent. Results from random bidding with multiple agents competing intuitively show that the presence of other agents drive the price at which bids are won upwards i.e. in order to win bids have to higher in the presence of other agents. From this simple simulation, it can be learned that there need to be more cultivated strategies adopted to build bidding strategies and to estimate CTRs. The results found here could also be used as the benchmark results in building better strategies.

4.2 pCTR Estimation

From the three pCTR estimators that were built, the XGB model was found to have the highest AUC ROC at 89.1%, followed by NN at 84.5% and LR at 80.1%. The AUC ROC of the ensemble model has been maxed at 86.9%. The superior performance of the XGBoost model can be attributed to its capability to handle sparse features, large datasets and support parallelisation. Additional steps can be taken to further strengthen and improve the pCTR models. Fist thing to remark is the sub-sampling approach favoured a very pragmatic approach. A more theoretical approach to randomly discarding data points such as using K-means to reduce sample size could improve performance. Moreover, some features with high entropy have been discarded and excluded from analysis. However, potentially valuable information might have been lost in the process. Applying bloom filters has been briefly considered but was ultimately deemed to be beyond the scope of this project. Alternatively, the ensemble methodology used in this report is rudimentary in nature and more sophisticated approaches such as stacking could strengthen the performance of the combined pCTR models.

4.3 Bidding Strategies

4.3.1 Linear Bidding

Having done the estimations for pCTR on the training set, the optimal bid prices were then computed using various strategies on the validation set, with the goal to achieve the highest number of clicks. The results for the linear bidding strategy across the three pCTR models used are shown in Table 5. Consistent with having the best accuracy to estimate pCTR, XGB manages to achieve the highest number of clicks of 164 with a base bid of 213. The budget ran out prematurely possibly because of the product of base bid and pCTR could still be high enough to win impressions. The other two models managed to achieve a similar number of clicks and CTR, however, they both have varying base bid values. The performance of XGB in the linear bidding strategy is also about 15-30% lower CPM and about 10% lower CPC compared to the other models.

4.3.2 Non-linear Bidding

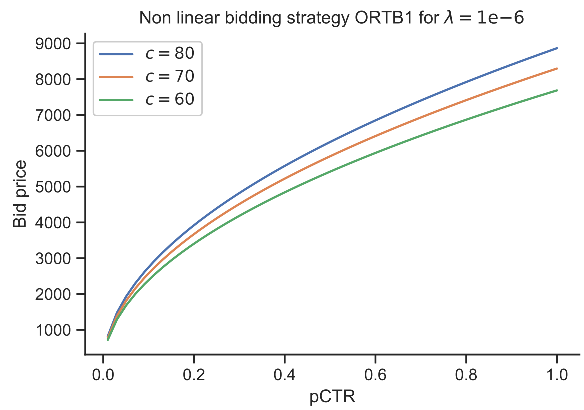

Figure 4: Relationship between Bid Prices and pCTR using XGB and ORTB1

Results from the three non-linear bidding strategies experimented on can be seen in Table 5. Amongst the different strategies, ORTB1 consistently performs the best as compared to the other two non-linear strategies across all pCTR estimators. However, the performance compared to others is only marginal. The highest number of clicks was reported by the ORTB1 strategy on the XGB estimator, at 165 clicks. Figure 4 shows the three best performing bidding strategies for ORTB1, from the predicted CTR; using the XGB estimator.

Comparison between the ranges of bid prices produced by different strategies can be seen in Figure 5. In this figure, the pCTR is estimated using the XGB model. One of the most prominent observations that can be made is the huge differences in the range of bid prices between the ORTB strategies and the linear and budget-constrained strategies implemented on the validation set. Remark that any bids higher than 300, in the first criteria setting, on the validation set has a 100 per cent probability of winning. Therefore the most relevant section is the body of bid prices where bid < 301 since all other extreme values will indeed all win the auction but only end up paying prices in the normal bid range. However, these extreme bids do provide some valuable insight into optimising a bid strategy when bidding against multiple agents.

4.3.3 Multi-agent Bidding

As discussed in Section 3.5.5, a game-theoretic approach was taken to emulate the competitive auction environment by simulating agents with different bidding strategies. The key challenge is that the number of clicks attained also depends on all the other agents in the module. All other agents attempt to come up with the best strategy to anticipate and counteract the bidding behaviour of other agents and any changes in one bidders' strategy will result in changes in

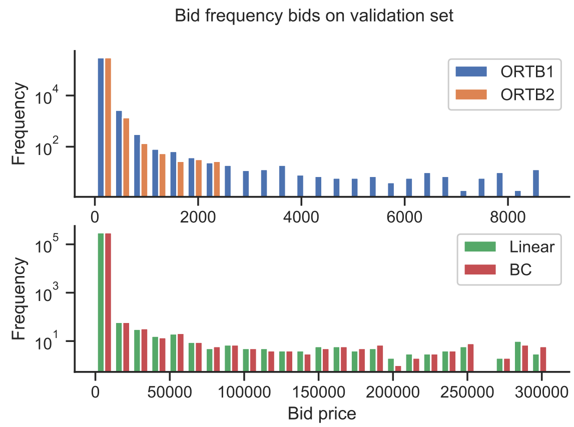

Figure 5: Bids Frequency on All Strategies using XGB estimator, on Validation Set

other bidders' strategies. This behaviour indicates the existence of a Nash equilibrium since the agents' strategy is the best strategy given the other agents' strategies.

It was found that auctions in a multi-agent scenario are extremely competitive. As often, increasing competition through increasing the number of people involved in the auction increases the price. If all teams were to spend all their budgets, this will result in the average pay price of at least 577 CNY. It should be noted that this is 7 times higher than the auction where only a single player auction. Therefore the knowledge of the bid distribution of the different strategies conveyed in Figure 5 can be used as an advantage i.e. since the bids need to be substantially higher in this competitive environment, it is the most likely a budget constrained or linear strategy will produce better results.

Moreover, traditional game theory is likely to be less relevant in the context of real-time bidding, as it is near to impossible to learn or model the distribution of bid prices. This is because each bidder comes up with a bid price according to his own algorithm to predict pCTR, and could have different objectives i.e. attempting to maximise different utility functions. In addition, unpredictable variables including irrationality, other strategic motives, campaign lifetime, and budget constraints would prevent bidders from bidding their private values [15]. Agents may also vary their bidding strategy at different stages of the campaign, such as prospecting [15] and re-targeting, which leads to great fluctuation of bidding activity of each agent in the limited time-frame of the campaign. Within the scope of this assignment, it can be argued both ways that agents either do or do not adjust their bidding strategy over time within this game setting. One could argue there is no sequential element in the game and prevents agents to adapt or develop strategies over time. On the other hand, one could argue that given the presence of a leader board it is to some extent possible to improve bidding strategies over time based on other agents strategies.

On a side note, one should note that since each bidder has its own valuation of the impression in each auction, given that each agent will have developed their own estimations for the CTR the auction is therefore not considered to be a common-value auction. Consequentially, this reduces the probability of encountering the winner's curse is minimal.

4.3.4 *Future work*

One of the glaring omission in this article is the fact that there is some historical knowledge of the price that is paid for bids and using a machine learning approach to predict the height of the price paid can be very valuable in improving the discussed bidding strategies subjected to budget constraints. Moreover, parameters used to optimise bids are determined in a quasi-scientific way i.e. there is a probability that these constants are not optimal constants and are potentially a local maximum or not even a local maximum at all. A gradient descent approach which is more rooted in science could improve the parameters used. However, this would include some very severe computational time. As alluded to before, there are some more sophisticated alternatives for the prepossessing available that given extra time could strengthen the foundation on which the ML models are built. Alternatively, the novel Distributed Coordinated Multi-Agent Bidding (DCMAB) solution proposed by Junqi Qi et al. (2018) [11] could be explored in the future. The DCMAB novel solution takes into account all advertisers' interactions to optimise their bidding strategies in a multi-agent environment. The approach outperforms the single-agent reinforcement learning approach, which echoes with the second-price winning criterion. Lastly, one could point out that the ensemble is mentioned in this report but no strategies were optimised for this estimated due to time constraints.

5. CONCLUSION

This report sought to explore the fundamentals of a Real-Time Bidding auction and how machine learning and game theory components can be used to maximise a certain predefined KPI given a budget constraint. The RTB typically employs a second-price sealed-bid auction model. In order to address the bid optimisation challenge that is present in RTB two fundamental components are addressed in this report: (a) finding the best model to predict CTR and (b) finding optimal bidding strategies. Two distinctive winning criteria are used to determine the auction' winners. The first criteria (i) resembles a single agent environment whereas the second (ii) resembles a multi-agent environment. To address the first fundamental component a Logistic regression, XGBoost, neural network and a rudimentary ensemble based on these models have been evaluated as predictive models for the CTR. The XGB model was found to have the highest AUC ROC at 89.1%, followed by NN at 84.5% and LR at 80.1%. Whilst the ensemble produced an AUC ROC of 86.9%. To address the second component for every model, a linear strategy, optimal real-time bidding 1, optimal real-time bidding 2 and a budget constrained strategy have been implemented and optimised to maximise the predefined number of clicks as KPI. Of all possible combinations, the XGB as the model to estimate the pCTR with ORTB1 as bidding strategy (given winning criteria (i)) yielded the maximum number of clicks on the validation set. However, compared to the other strategy's performance on the validation set is only marginally better. The predictive model with the best performing strategy was then applied to the test set which yielded 175 clicks, which at the time of writing, provided the 10$^{\text{th}}$ rank on the leader board. Sequentially, optimising a strategy for the multi-agent environment (criteria (ii)) was addressed. Since there is no information regarding other agents bidding strategies it was decided to optimise a strategy given an environment in which all the best performing models that have been identified throughout this report are actively partaking. The emulated environment consisted of 37 simulated agents of which some agents have been trained on a subset of this environment to get more realistic opponents. This approach was an attempt to best model other agents and develop a bidding strategy given the presence of other agents. The best performing strategy on the simulated environment was the budget constrained strategy whilst using the pCTR produced by XGBoost. This strategy was applied to the test set and submitted to the leaderboard. At the time of writing, the optimised strategy on a simulated environment enforcing winning criteria (ii) yielded 10 clicks resulting in the 8^{th} position on the leaderboard. In future work, several improvements can be made to strengthen the robustness of the analysis and mythologies presented.

6. GITHUB REPOSITORY

The codes for this project can be found at https://github.com/FaizalNafis/maai_bid_strategy.git

7. REFERENCES

[1] O. Chapelle, E. Manavoglu, and R. Rosales. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology (TIST) - Special Sections on Diversity and Discovery in Recommender Systems, Online Advertising and Regular Papers*, 51(4), January 2015.

[2] Q.-H. Chen, S.-M. Yu, Z.-X. Guo, and Y.-B. Jia. Estimating ads' click through rate with recurrent neural network. In *ITM Web of Conferences*, volume 7, pages 1–5, November 2016.

[3] K. chih Lee, B. Orten, A. Dasdan, and W. Li. Estimating conversion rate in display advertising from past performance data. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 768–776, August 2012.

[4] L. Fisher. Us programmatic ad spending forecast update, https://www.emarketer.com/content/us-programmatic-ad-spending-forecast-update-2018, 2018.

[5] A. Ghosh, B. I. Rubinstein, S. Vassilvitskii, and M. Zinkevich. Adaptive bidding for display advertising. In *Proceedings of the 18th international conference on World wide web*, pages 251–260, April 2004.

[6] K. J. Lang, B. Moseley, and S. Vassilvitskii. Handling forecast errors while bidding for display advertising. In *Proceedings of the 21st international conference on World Wide Web*, pages 371–380, April 2012.

[7] Q. Liu, F. Yu, S. Wu, , and L. Wang. A convolutional click prediction model. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, page 1743âĂŞ1746. ACM, October 2015.

[8] T. Maehara, A. Narita, J. Baba, and T. Kawabata. Optimal bidding strategy for brand advertising. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pages 424–432. IJCAI, July 2018.

[9] Z. Pan, E. Chen, Q. Liu, T. Xu, H. Ma, and H. Lin. Sparse factorization machines for click-through rate prediction. In *2016 IEEE 16th International Conference on Data Mining*. ACM, December 2016.

[10] C. Perlich, B. Dalessandro, R. Hook, O. Stitelman, T. Raeder, and F. Provost. Bid optimizing and inventory scoring in targeted online advertising. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 804–812, August 2012.

[11] J. Qi, C. Song, H. Li, K. Gai, J. Wang, and W. Zhang. Real-time bidding with multi-agent reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 2193–2201, October 2018.

[12] A. Sayedi. Real-time bidding in online display advertising. *Marketing Science*, 37(4):553–568, August 2018.

[13] I. Tomek. Two modifications of cnn. In *Systems, Man, and Cybernetics*, volume 6, pages 769–772. IEEE Transactions, November 1976.

[14] I. Trofimov, A. Kornetova, and V. Topinskiy. Using boosted trees for click-through rate prediction for sponsored search. In *Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy*. ADKDD, August 2012.

[15] J. Wang, W. Zhang, and S. Yuan. *Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting*. Now Publisher Inc, Hanover, 2017.

[16] S. Yuan, J. Wang, and X. Zhao. Real-time bidding for online advertising: Measurement and analysis. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*. ADKDD, August 2013.

[17] Y. Yuan, F. Wang, J. Li, and R. Qin. A survey on real time bidding advertising. In *IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI*, pages 418–423, November 2014.

[18] W. Zhang, T. Du, and J. Wang. Deep learning over multi-field categorical data: A case study on user response prediction. In *European Conference on Information Retrieval*, pages 45–57, March 2016.

[19] W. Zhang, S. Yuan, and J. Wang. Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1077–1086. KDD, August 2014.

[20] W. Zhang, S. Yuan, and J. Wang. Real-time bidding benchmarking with ipinyou dataset. *CoRR*, abs/1407.7073, 2014.