Web Economic Individual Report

Weisi Han
University College London
School of Computer Science
uczlwha@ucl.ac.uk

1 INTRODUCTION

Technological advancements and and high data availability in display advertising have opened a wealth of options for advertisers to precisely target campaigns. It provides a more streamlined and transparent way to buy ad inventory for each impression, allowing advertisers to bid based on individual impression's worth. The recent emergence of real-time bidding (RTB) has made the advertising work more efficiently for buyers. In an RTB auction, advertisers bid against each other to win the ad space. Therefore, how to develop an better bidding strategy to accurately compute the bid value for each impression is of importance in this performance-driven business.

In order to address the problem, this report proposed an online bidding strategy which contains two individual components: a Click-Through-Rate (CTR) estimation model based on Factorisation Machines (FM) and an optimised bidding price estimator based on PRUD algorithm. The former learns the features of each impression in iPinYou advertising dataset and generates the predicted ad-click probability (pCTR), while the latter predicts the winning price and combines it with the pCTR to compute the bidding value. The results show that FM CTR prediction model exhibits a better performance than other linear predictors and the PRUD bidding price calculator also outperforms other strategies with regard to both click numbers and effective Cost-Per-Click (eCPC). Additionally, a series of data exploration, feature analysis and model evaluation will be also discussed in this report. All source code related to this report has been uploaded to [15].

2 LITERATURE REVIEW

There are lots of techniques related to RTB based display advertising [1]. Focusing on the Display-Side Platform (DSP), how to maximise the user response rate (eg. the CTR and Conversion Rate (CVR)) and improve the advertisement performance should be treated as the main goal. Therefore, a number of machine learning models to predict CTR and CVR have been proposed in recent years. Among them, Logistic Regression ([3], [4], [5]) is one of the simplest but effective tools, however, it is difficult for a LR model to analyse high dimensional features. Approaches based on Neural Networks (NN) include [6] and [7], although NN is able to deal with more complex feature interactions, it is difficult to obtain a global optimisation. Factorisation Machines, which relies on feature engineering and matrix design could be seen as an excellent approach applied in CTR estimation ([8], [9]). It allows us to implement pairwise combination of features in an efficient way both in terms of time and space complexity.

To optimise the bidding price prediction, two state-of-the-art bidding strategies have been discussed in [10] and [11]. The first paper applies a linear function to compute the bid value, whereas the second one (ORTB) establishes non-linear equations to calculate the bid price. However, both of them only take a single parameter pCTR into the bidding value computation. In this report, we would like to incorporate a new indicator to enhance the overall bidding performance.

3 APPROACH AND RESULTS

3.1 Data Exploration

3.1.1 *Data Format*.

The whole dataset contains a training set, a validation set and a test set with the data number proportion about 8:1:1. The training dataset comprises a set of processed iPinYou DSP bidding logs. Each row represents a successful impression which could be categorised into the following four groups:
- **Scenario Feature**: click time (day & week), browser and operating system, etc.
- **User Features**: user IP, region, city, user segments, etc.
- **Ad Feature**: Ad exchange id, advertiser id, Ad slot information (slot height, width, formats), etc.
- **Price Feature**: bid price, floor price, market price, etc.

Each piece of impression in the training set is associated with click situation, (whether it is clicked or not) Except the price and slot shape information, most of the features are categorial variables which will be further processed. The feature processing part will be described in detail in the group report.

3.1.2 *Basic Statistics of the Dataset*.

In order to analyse the basic statistical information of the training set, a number of indicators which could be used to evaluate the bidding performance will be summarised in Table1. This includes the number of clicks, total cost, Click-Through-Rate(CTR), Cost-Per-Millie(CPM), wining impressions, and effective Cost-Per-Click(eCPC) for each advertiser. The overall situation has also been computed and appended at the bottom of the table.

The **CTR** measures the ratio of total impressions to clicks in display advertising.

$$CTR = \frac{Total\ number\ of\ clicks}{Total\ number\ of\ impressions} \tag{1}$$

The average **CPM** refers to the cost for thousand impressions.

$$CPM = \frac{Total\ cost \times 1000}{Total\ number\ of\ impressions} \tag{2}$$

The **eCPC** reflects the effective cost per click in online advertising, which can be calculated as:

$$eCPC = \frac{Total\ cost}{Total\ number\ of\ clicks} \tag{3}$$

As can be seen from Table 1, advertiser 2997, which belongs to the mobile e-commerce app industry, generates an abnormally

Table 1: Basic Statistics of the Training Set

Adv.	Imps	Clicks	Cost	CTR	CPM	eCPC
1458	492353	385	33969	0.0782%	68.99	88.23
2259	133673	43	12428	0.0322%	92.97	289.03
2261	110122	36	9874	0.0327%	89.66	274.27
2821	211366	131	18828	0.0620%	89.08	143.73
2997	49829	217	3129	0.4355%	62.80	14.42
3358	264956	202	22447	0.0762%	84.72	111.12
3386	455041	320	34932	0.0703%	76.77	109.16
3427	402806	272	30459	0.0675%	75.62	111.98
3476	310835	187	23919	0.0602%	76.95	127.91
Total	2430981	1793	2430985	0.0737%	78.15	105.95

high ratio of CTR. This can be attributes to its mobile environment. Apart from that, most of the advertisers in our dataset has a relatively high CTR, (the average CTR for a display ad across all industries in real life is just around 0.35%). Advertiser 1458 has the highest number of ad clicks, which is followed by advertiser 3386 and 3427. Advertiser 3386 (international e-commerce) has allocated great amount of money to bid for impressions whereas for advertiser 2261 (Telecommunication), they did not spend much money in wining the impressions and thus, the generated clicks is very low. The CPM for the nine advertisers ranges from around 69 to 90 fen, however, their corresponding eCPC varies a lot. The lowest eCPC belongs to advertiser 2997 (14.42 fen), whereas advertiser 2259 has an extremely high eCPC, which is about 195 higher than the overall situation (105.95 fen). Therefore, it is necessary for advertiser 2259 to identify their potential customers to better target their advertising.

3.1.3 *User Feedback and Bidding Behaviour*.

In this section, we choose advertiser 3358 and 3386 and compare their user feedback (represented by CTR) together with bidding behaviour (represented by market price) distributions against different features. The CTR and the market price are plotted in one graph with two y-axis for joint observation and better comparison.

The first four plots in Figure1 depict the time impact on CTR and market price of advertiser 3358 and 3386. Focusing on the first subplot, we can find that advertiser 3358 receives the lowest CTR on Wednesday and the highest on Thursday. However, the market price rises to the peak on Wednesday but drop to the bottom a day after. Therefore, the high CTR combined with the low pay price will result in a remarkable eCPC and the return-on-investment (ROI) for advertiser 3358 on Thursday should be relatively high. For advertiser 3386 (second subplot), both of the CTR and market price followed the same overall trend over a week. It only exhibits a low market price on Monday. If we observe the trend of daily fluctuation, the two indicators for both advertisers follow the similar overall tendency. In general, the bid competitiveness for both advertisers is higher in the afternoon and evening than in the morning. The market price for advertiser 3386 is lower than 60 fen from 1 to 6 am as the CTR at those times is expected to be low, whereas the price is at least 70 fen for advertiser 3358 throughout the whole day.

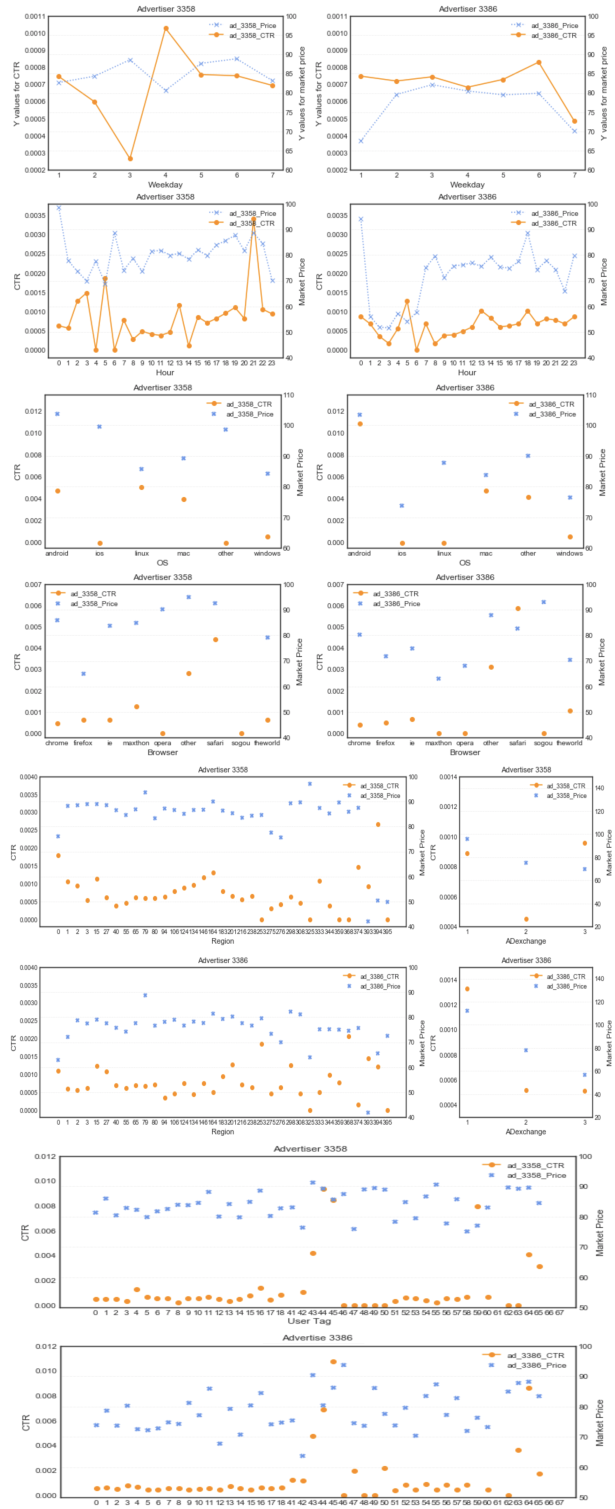

Figure 1: CTR and Market Price Distribution against Time, User Agent, Region, Ad Exchange, and User Tag for Ad3358 and Ad3386

When it comes to the impact caused by browser and operating system, there is a greater possibility for users to click an ad from advertiser 3358 on linux operating system and its market price is very low. Therefore, bidding the impressions which are shown on Linux operating system can be treated as a cost effective choice. Nevertheless, bidding an impression on ios platform is unprofitable as the market price is around 100 fen, with almost 0 CTR. In terms of browsers, both advertisers could obtain a high CTR with a low pay price spending when they advertise on Safari.

Regions and Ad exchanges also influences the CTR and market price heavily. Specifically, the market price for both advertisers in region 393 is just over 40 fen, but the CTR for them are similar with others. In terms of Ad exchange, they facilitates buying and selling process of media advertising. It is obvious that for both advertisers, the market price on Ad exchange 1 is higher than Ad exchange 2 or 3, which shows a strong competitiveness. In addition, the CTR and market price of Ad exchange 1 for advertiser 3358 is lower than that for advertiser 3386. However, those for Ad exchange 3 experiences an opposite situation.

The last two plots in Figure1 illustrates the distributions against different user tags. We could find that the market price for two advertisers are both volatile across different user tags, which varies between 60 and 100 fen. The majority of CTR are below 0.2%, however, advertiser 3358 has a noticeably high CTR on user tag 44, 45, and 59. Furthermore, the CTR from advertiser 3386 exceed 1% on user tag 45, which is extremely high for display ads.

3.1.4 ***eCPC and Further Analysis.***

As each individual plot in the above figure connect both CTR as well as market price, it could also give us an insight to analyse the eCPC for both advertisers across different features. A lower eCPC indicates a lower actual price the advertiser pay for each click, which also reflects the effectiveness of the bidding strategy and a high ROI. Therefore, if we take advertiser 3358 as an example, it is better for them to launch their bidding strategy on safari with mac operating system on Thursday. If possible, they could also launch their advertising strategy to a specific user segmentation (tag 59) to improve their ROI.

3.2 Bidding Strategy

The bidding strategy is composed of two parts, CTR estimation to predict the click possibility of each impression, and the bidding function to generate appropriate bid price according to the pCTR and other bid related features.

3.2.1 ***CTR Estimation.***

Model: Factorisation Machines.

FMs have been widely applied in the field of advertisement click prediction and collaborative recommendation systems. One of the reasons that we apply FM into our bidding strategy involves its ability to deal with feature-rich datasets.

Considering the following predictive model for a simple multivariate linear/logistic regression:

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i \tag{4}$$

Table 2: Training Feature: useragent, after OneHot Encoding

...	linux	mac	win	chrome	firefox	ie	opera	...
...	0	0	1	0	0	1	0	...
...	0	0	1	1	0	0	0	...
...	0	1	0	0	0	0	0	...
...	0	0	1	0	0	1	0	...

where y is the label, indicating click or not, x_i are eigenvectors independent from each other and w_0 and w_i are weights for the features. This equation learns the effect caused by individual variables. However, based on our dataset exploration, we found the CTR, market price and eCPC are more likely to correlate with some combined features. From this, we can assume if we integrate some individual feature with others, users' click behaviour can be better determined. For example, if we combine advertiser with id 1458 (the e-commerce advertiser) and users with tag 38 (inmarket/clothing, shoes and bag), we could find a strong positive correlation between these pair-wised features and click.

This is one benefit of using FM, which explores the feature interactions and allows us to build a model based on the feature combinations in the dataset. The equation above can be therefore rewritten as:

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} w_{ij} x_i x_j \tag{5}$$

The last part of the equation takes the interactions between pairs of features into account.

Another big advantage of FMs related to its remarkable performance when dealing with sparse training data. As is described in our group report, the way we adopt to process most of our categorial variables is through one-hot encoding. After converting the categorial features into dummy variables, our dataset will be represented by a huge amount of features. Table 2 shows part of the processed features in our training set. It is obvious that even when we purely focus on useragent category, only two of the transformed features are filled non-zero values. Expending to the entire feature matrix, each impression could possess 127 feature dimensions and over a hundred of columns are equal to zero. Therefore, data sparsity is an unavoidable challenge in our CTR prediction, which can be misleading to modelling algorithms and influence prediction accuracy. In order to avoid expensive computation during model training, FM adopt Matrix Factorisation strategy which extracts the most important latent features from the existing ones and presents almost the same correlation between the target and predictors by using a lower-dimension dense matrix. Since it deals with the feature interactions as dot product of low dimensional vectors, Equation (5) could be improved to:

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} < v_i, v_j > x_i x_j \tag{6}$$

Where $<v_i, v_j>$ represents the inner product of the k-dimensional latent vectors associated with each variable (i.e. it models the feature interaction).

Table 3: Main Training Parameters

Parameter	Value	Remark
order	3	3-order feature interaction
rank	10	k in the matrix factorisation
optimizer	Adam	choose from Adam or FTRL
learning rate	0.001	α
init_std	0.001	standard deviation for initialization

Table 4: CTR Estimation Performance

Model	AUC	Accuracy	Precision	Recall	F1 Score
LR	0.794	99.93%	20.00%	0.50%	0.97%
FM	0.876	99.93%	46.94%	11.39%	18.33%

Model Construction and Parameter Tunning.
There are numbers of ways to implement FM [12], I would like to use ***tffm*** [13], which is an implementation of an arbitrary order Factorisation Machine. Theoretically, as FM is capable to automatically identify significant latent features, and it works with numerical data and categorial represented as binary integers, all of our pre-processed 127 features could be directly fitted into the model. However, sparse matrix computation is not cost efficient in terms of time and space complexity. Thus, these training features will be further transformed by using Scipy's Compressed Sparse Row (CSR) algorithm.

In order to obtain the best prediction performance, we have to decide the optimal model parameters. Our parameter tunning process will be based on the model's accuracy as well as AUC. We first set parameter *order* from 2 to 9. Interestingly, both of the accuracy and AUC of the model drops after reaching a peak of order 3, which indicates for our sparse dataset, a second or third order FM model suffices, as it may not be able to provide enough information to estimate more complex interactions. Similarly, we initially thought the higher *rank* (factorisation parameter k) is, the more latent features the model could learn. However, after running with a large number of iterations, we set rank-k as 10, since it provides the best performance. We also determined other parameters in the similar way and the optimal set of parameters are shown in Table 3. As is discussed in the group report, different down sample ratio has been tried to balance the positive and negative samples in the training set and finally the ratio is set as 0.1 for the best FM model performance.

Evaluation and Result.
As the prediction of CTR can be seen as a logistic regression task, percentage correct (Accuracy) is often a good statistic to measure the performance of the CTR estimator. In addition, the ROC curve and the Area Under the Curve (AUC) are also common tools used with binary classifiers, thus, they are joint considered in our case. Other statistics, like the confusion matrix and the precision, recall and F1 score will be listed and briefly interpreted in the following part as well.

The experimental results for CTR estimation with LR and FM are displayed in Table 4. We could find that although both of the models have generated the same accuracy, this does not automatically mean that the quality of their predictions are equal. Due to the huge imbalanced characteristic of positive and negative labels, the accuracy scores are biased. In fact, in our specific CTR prediction case, the recall is very important, which represents the ratio ***tp/(tp+fn)***, where *tp* is the number of true positive predictions and *fn* is is the number of false negatives. For LR, it only has a 0.5% recall rate, which means it is very difficult for a simple LR classifier to determine all click=1 samples. However, the recall score for FM is over 11%, much better than LR. Similarly, the precision describes the ability of the classifier to not label a sample as positive if it is negative. FM also has a higher precision than LR. When we combine these two indicators, F1 score tells us that FM is still better than LR.

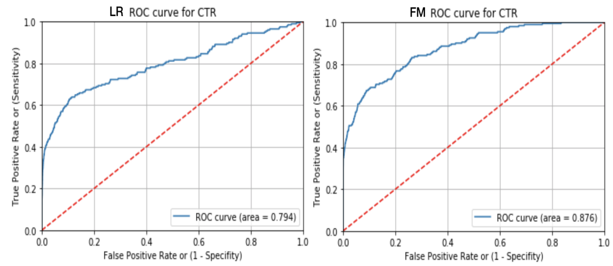

Figure 2: ROC and AUC Plot for LR (left) and FM (right)

Another evaluation metric for our CTR predictor is the ROC and AUC since they are insensitive to the class distribution. Figure 2 shows the ROC curve and AUC for LR and FM, where the true positive rates are plotted against false positive rates. The dotted red line represents the ROC curve of a purely random classifier and a good classifier aims for the upper left corner of the chart. Obviously, the ROC of the FM is in accordance with this requirement. With an AUC of 0.876, FM displays a promising overall performance.

3.2.2 ***Bidding Strategy Optimisation.***
PRUD Algorithm.
After calculating the pCTR for each impression, we need to construct an additional winning price predictor to complete the whole bidding process. In the group report, we compared three types of bidding strategies, a constant bidding strategy, a random one and a linear bidding strategy. Although the linear formed strategy has already generated better performance than the previous two, it merely relies on a single parameter, pCTR. However, once the the accuracy of the pCTR estimator could not be guaranteed, the predicted bidding price will be affected as well. Therefore, apart from considering the pCTR, I would like to integrate an additional estimator, the predicted Winning Price (pWP) into my bidding strategy to enhance the performance of the bidding function.

Inspired by Lin, et.al [14], I choose to use PRUD algorithm to make prediction. This algorithm combines the powers of two estimators mentioned above. It takes the features of each impression as input and output three variables, the pCTR, the pWP and a threshold of the bidding efficiency cutoff (ρ_{cut}).

Based on what has been discussed in Section 3.2.1, we have completed the training process of calculating the pCTR and thus we only need to focus on the winning price prediction. Considering the feature sparsity, we decide to apply a FM Regressor to predict the advertisers' pay price. After these two process, a bidding efficiency value (ρ_{cut}) will be computed for each impression.

$$\rho(\vec{x}) = \frac{pCTR(\vec{x})}{pWP(\vec{x})} \tag{7}$$

This equation presents the correlation between the bidding efficiency, *pCTR* and *pWP*. A higher bidding efficiency can be observed when an impression is predicted to have a strong click possibility but with a low market price, which will motivate the advertiser to bid for an impression. However, when there is an ad with low predicted click-through rate and low market price as well, it is also worth investing, as even if it ends up with no click-through, the loss is acceptable. Similarly, by combining the *pPW* and *pCTR* simultaneously, biddings with extremely high predicted price but low *pCTR* will be excluded from our consideration, which could alleviate the risk of blindly investment. Intuitively, the higher ρ, the higher value a bid has, and the stronger inclination we would allocate our budget on it.

After calculating the bidding efficiency for each impression, the next step is to find a proper cutoff, ρ, to determine whether the value of this impression can be treated as high or low. The idea is once the efficiency value of one impression is higher than the cutoff, it is of high value and we would therefore bid on it. Under our specific context, the best threshold value can be determined by tunning the bid efficiency value with the validation set as the validation set and the test set express the same impression distributions. The best cutoff value would be the lowest ρ which generates the maximum click number under the predefined budget 6250 fen. When an impression possesses a higher ρ than this threshold, we would bid on it with the predicted price plus a small *lift-up* (RTB is worked on a second-price auction model), if not, we would just skip this impression. This cutoff value will be finally used in the test set and it is expected to produce the same level of optimal result.

Evaluation and Result.
In this part, we analyse the bidding strategy performance after combing two kinds of pCTR estimation models with different bidding algorithms. The result is show in table 5 We will use the LR + Linear bidding strategy as a baseline to make comparisons with other bidding strategies. This strategy can obtain 142 clicks. However, if we combine the linear bidding strategy with FM CTR estimator, 27 more click can be achieved, which is the highest click number (169) among three strategies. The last strategy highlighted in this report combines FM and PRUD algorithm. The click number it achieves lies between the previous two strategies.

In general, both biding strategies combined with FM model has won more clicks than the baseline. For the PRUD algorithm, although we initially thought it could in some manner compensate the defect of the CTR predictor and improve the overall click number by incorporating an additional winning price estimator, it does not offer stronger outcome to the other model. This is because we could not guarantee that both the CTP and WP prediction are always be 100% precise for all cases.

However, in terms of real online display advertising, focusing on the generated clicks is far from enough. One of other good proxies for evaluation is the effective cost per click. It can be seen that PRUD exhibits a lower eCPC than other two cases even though it sacrifices a certain number of clicks. This is reasonable as the objective of the algorithm we developed starting from filtering out the impressions with low bidding efficiency, while the bidding efficiency is subject to the ratio of pCTR (related to click number) and winning price (related to cost). As a result, this value positively corresponds with the eCPC and the majority of impressions with higher eCPC will be retained. For other metrics like CTR and CPM, PRUD also shows superiority upon other strategies.

Table 5: Bidding Strategy Performance

Model	Click	CTR	Spend	CPM	eCPC
FM+Linear	169	0.001439	6240	53.197	37.202
LR+Linear	142	0.001086	6189	47.321%	43.585
FM+PRUD	159	0.001381	6249	45.871	34.738

PRUD Enhancement.
As the main goal for this project is to maximise the click number on the test set, I decide to further enhance the PRUD bidding algorithm performance. To achieve that, I plot the predicted CTR distributions of click 1 impressions and click 0 impressions based on the validation set. For click = 0, the vast majority of predicted click probability is concentrated around zero, which decays as it approaches to pCTR = 1. The click=1's distribution is not as good. Although there is a proportion of pCTR concentrated at one, more are spread across the x-axis. In spite of this, we still find that the minimum pCTR of all click=1 impressions is $6.18 * 10^{-7}$, which means other impressions with a pCTR lower than this value can be treated as 0-click impressions. Bidding on these impressions is a waste of money and thus, we set the bid price of all low-pCTR impressions as 0. Since this part will be examined on the test set, detailed explanations are put in the Group report. The result shows that, by allocating limited budget on higher-value impressions, the overall CTR could be optimised and the eCPC has declined as well.

4 CONCLUSION AND RETROSPECTIVE

In this report, we first described the problem in the current RTB context. After conducting a basic exploration of the iPinYou impression dataset, we analysed the basic statistical information together with user feedback and bidding behaviour. In order to predict the CTR, we applied an FM model, which has been proved to be an effective solution to deal with high-dimensional sparse data. The PRUD bidding strategy also showed an outstanding bidding performance. By combining the pCTR and WP predictor, it outputs the most optimal bidding price, which not only contributes to a higher click number, but also a lower eCPC.

Further potential direction for this coursework could include analysing the pCTR and bid price for each advertiser under different budget constraints. As different advertisers belong to different industrial categories, their target audiences could also vary from each other. If we focus on specific advertisers, we could work out more pertinent strategy in terms of target clients and markets.

Table 6: Task Allocation

Tasks	Weisi	Qiuru	Boyang
Data Exploration	X	X	X
Feature Extraction	X		
Down Sampling			X
Constant Bidding	X		
Random Bidding		X	
CTR Estimation Model	X	X	X
Logistic Regression	X	X	X
Factorisation Machine	X		
XGBoost		X	
LGBM			X
DNN			X
Bidding Strategies	X	X	X
Linear	X	X	X
ORTB		X	
PRUD	X		
Combined Strategy	X	X	X
Multiagentt		X	

5 GROUP ROLES

The team is composed of Weisi Han, Qiuru Dai, and Boyang Liu. Although the whole project has been divided into several tasks and each of us has to develop our individual model, we discuss everything and share our code with each other in order to realise our common goal: IMPROVE CLICKS!!! Table 6 summarised each individual work in detail.

REFERENCES

[1] Hairen Liao, Lingxiao Peng, Zhenchuan Liu, and Xuehua Shen. 2014. iPinYou Global RTB Bidding Algorithm Competition Dataset. Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining - ADKDD14 (2014). DOI:http://dx.doi.org/10.1145/2648584.2648590

[2] Xinran He et al. 2014. Practical Lessons from Predicting Clicks on Ads at Facebook. Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining - ADKDD14 (2014). DOI:http://dx.doi.org/10.1145/2648584.2648589

[3] M. Richardson, E. Dominowska, and R. Ragno. Predicting clicks: estimating the click-through rate for new ads. In WWW, pages 521âĂŞ530. ACM, 2007.

[4] O. Chapelle, E. Manavoglu, and R. Rosales. Simple and scalable response prediction for display advertising. ACM Transactions on Intelligent Systems and Technology (TIST), 5(4):61, 2015.

[5] H. B. McMahan, G. Holt, D. Sculley, M. Young, D. Ebner, J. Grady, L. Nie, T. Phillips, E. Davydov, D. Golovin, et al. Ad click prediction: a view from the trenches. In SIGKDD, pages 1222âĂŞ1230. ACM, 2013.

[6] A. I. Baqapuri and I. Trofimov. Using neural networks for click prediction of sponsored search. arXiv preprint arXiv:1412.6601, 2014.

[7] Q. Liu, F. Yu, S. Wu, and L. Wang. A convolutional click prediction model. In CIKM, pages 1743âĂŞ1746. ACM, 2015.

[8] A.-P. Ta. Factorization machines with follow-the-regularized-leader for ctr prediction in display advertising. In Big Data, pages 2889âĂŞ2891. IEEE, 2015.

[9] R. J. Oentaryo, E.-P. Lim, J.-W. Low, D. Lo, and M. Finegold. Predict- ing response in mobile advertising with hierarchical importance-aware factorization machine. In WSDM, pages 123âĂŞ132. ACM, 2014.

[10] C. Perlich, B. Dalessandro, R. Hook, O. Stitelman, T. Raeder, and F. Provost. Bid optimizing and inventory scoring in targeted online advertising. In KDD, 2012.

[11] W. Zhang, S. Yuan, and J. Wang. Optimal real-time bidding for display advertising. In KDD, 2014.

[12] Steffen Rendle. 2012. Factorization Machines with libFM. ACM Transactions on Intelligent Systems and Technology 3, 3 (January 2012), 1âĂŞ22. DOI:http://dx.doi.org/10.1145/2168752.2168771

[13] Geffy. 2018. geffy/tffm. (March 2018). Retrieved March 10, 2018 from https://github.com/geffy/tffm

[14] Chi-Chun Lin, Kun-Ta Chuang, Wush Chi-Hsuan Wu, and Ming-Syan Chen. 2016. Combining Powers of Two Predictors in Optimizing Real-Time Bidding Strategy under Constrained Budget. Proceedings of the 25th ACM International on Conference on Information and Knowledge Management - CIKM 16 (2016). DOI:http://dx.doi.org/10.1145/2983323.2983656

[15] https://github.com/BlaBlaPer/WebEconomics/tree/master/Weisi