Artificial intelligence for smart bidding

Pratyush Shandilya
Data Scientist, Amplify Analytix Ltd, India

Pratyush Shandilya is a Data Scientist at Amplify Analytix. He has developed bespoke end-to-end data science solutions for clients in diverse industries. His expertise is in deep learning, reinforcement learning and natural language processing. Pratyush holds a MSc from the Carnegie Mellon University.

Amplify Analytix Ltd, India
E-mail: pratyush.shandilya@amplifyanalytix.com

Laura Murphy
Chief Executive Officer, Amplify Analytix, The Netherlands

Laura Murphy is the co-founder and Chief Executive Officer of Amplify Analytix. A former head of business transformation in international markets, Laura specialises in helping organisations embrace the benefits of data science to improve commercial performance.

Amplify Analytix BV, De Cuserstraat 93, 1081CN Amsterdam, The Netherlands
E-mail: laura.murphy@amplifyanalytix.com

Fernando Perales
Head of the Research Lab and Strategic Partnerships, JOT Internet Media, Spain

Fernando Perales is Head of the Research Lab and Strategic Partnerships at JOT Internet Media, where he is responsible for the coordination of innovation projects, from developing in-house technology to participating in pan-European collaborations to test cutting-edge Big Data applications.

JOT Internet Media, General Ramirez de Madrid 8, 28020 Madrid, Spain
E-mail: fernando.perales@jot-im.com

Abstract
Growth in industry digitisation in recent years has resulted in consumers, manufacturers and service providers succumbing to the allure of the online platform, leading to the erosion of the long-standing supremacy of traditional businesses. With increasing competition among manufacturers both at local and global scope, it becomes crucial for them to make intelligent digital marketing decisions which could help get ahead of their competitors. Artificial intelligence (AI) has proven to be an effective medium for supporting decision-making in digital marketing. This paper discusses the application of contextual multi-armed bandits to optimise the bidding strategy used by digital advertisers in main search platforms. The highly scalable algorithm, apart from suggesting a winning strategy in an advertising auction, also enables clients to improve return on investment (ROI) using digital advertising.

KEYWORDS: ad exchange, advertiser, contextual multi-armed bandits, digital marketing, publisher, reinforcement learning, AI (artificial intelligence), data-driven decision making
INTRODUCTION
During the last couple of years, the COVID-19 pandemic has been the driving force behind a substantial shift in the digital marketing landscape. The traditional retail market is gradually being displaced by e-commerce platforms at a rate that is 2–5 times quicker than it was in the pre-pandemic period. Over-the-top (OTT) platforms are gradually replacing movie theatres in consumers’ lives and so on.¹ As a result of this tendency, digital channels will continue to hold the title of the most heavily funded advertising medium globally in 2021. Digital promotion garnered almost 59 per cent of worldwide advertising spending, compared to television’s less than 25 per cent.²

According to Gartner’s annual CMO survey,³ organisations’ marketing budgets are at their lowest level as a proportion of sales despite the unexpected expansion of digital media, creating the perfect storm. Budget constraints in organisations are one major detrimental effect of COVID-19, the war in Ukraine, the unfolding energy crisis and other macroeconomic factors, which also took a toll on the marketing departments within organisations. This calls for reassessing and revamping the marketing strategies.

To stay relevant and make the most of the marketing budget, it is important to react swiftly when user interests and trends are changing and the digital marketing field is so competitive. This often involves considerable resources and analysis of what has previously worked and not worked. Trends these days, however, are so diverse as a result of the unpredictable macroeconomic environment that it is necessary to take on a cautious approach. As a result, automation and data analytics are used increasingly, but only after carefully selecting the most effective strategies.

The paper proposes an application of artificial intelligence (AI) to develop digital marketing strategies that result in better return on investment (ROI), overcoming the shortcomings created by marketing budget constraints. The proposed approach proved to be effective in common problems in digital marketing. Investing the right amount in ad allocation scenarios like paid search or sponsored products led to serving the right advertisements to the right audience at the right time.

BACKGROUND AND RELATED WORK
In the past couple of years, companies have been focusing more on digital advertisements for promotions of their products and services. Since digital advertisements are one of the cheapest forms of marketing, the competition among advertisers has increased to a great extent, such that every advertiser wishes to have an edge over the competitors, aiming to improve the reach of their products or services to the target audience.

Advertisement landscape
The two main entities in the advertising world are advertisers and publishers (from here on referred to as platform or publisher platform). An advertiser is any organisation that aims to market their products or services on a publisher platform, which is an online platform that publishes ads; it may be an e-commerce platform, an OTT platform, a search engine, a streaming platform, etc. Broadly, there are two types of channels by which an advertiser can approach a publisher platform:

1. Direct channel: Advertiser works directly with the publisher platform to purchase advertising space on a website or in a traditional publication. The ad revenue generated via this channel is fixed, as the terms and rates for publishing the ad on the platform are decided beforehand;

2. Programmatic channel: This channel supports automated buying and selling of online advertising. Advertisers take part in an auction where they bid for
Artificial intelligence for smart bidding

Although direct advertisement supports fixed ad revenue, it becomes difficult to engage a lot of advertisers in the channel, which might result in a specific type of ad not getting published on the platform. On the other hand, an open auction scenario, supported by a programmatic channel, provides a fair chance for every advertiser taking part in it to publish their ad.

Figure 1 shows different components involved in a programmatic channel. It involves multiple indirect partners (supply side partner, ad exchange, demand side partner), facilitating the ad request sent by a user to be processed and finally getting an ad published on the platform. The steps involved for the same are as follows:

1. While a user or consumer browses content on the publisher platform, an ad request is sent to the publisher’s ad server;
2. If there is available ad space on the publisher platform, the request is routed to the supply side platform (SSP);
3. SSP then further sends the request to an ad exchange;
4. The ad exchange hosts an open auction by creating a real-time bidding environment. Other ad exchanges and demand side platforms (DSP) take part in the open auction. DSP is further connected with advertisers, acting as a means for advertisers to take part in the auction. The winning bid by one of them decides which ad is to be published;
5. Instructions for ad display are then passed to the audience’s session and an ad is retrieved to be published from the advertiser.

Ad allocation strategy

Advertisers, seeking an opportunity to publish their ad on the platform, participate in the auction by placing a bid amount for the available ad inventory, either directly or via DSP. This is a blind auction; an advertiser places a bid amount for the available ad slot without any information about who their competition is, how many competitors they have or how much bid amount is placed by the competition.

Every real-time auction scenario is different. For example, in some auction scenarios, the advertiser that bids the highest wins the auction, whereas a reserve price (also called floor price) is set for some auctions such that the bidder closest to that reserve price wins the auction. In every auction scenario, however, irrespective of the winning criteria used, the advertisers try to decide an optimum bid amount that could improve their chances of winning the auction. Setting an extremely low or remarkably high bid value can result in losing the auction and/or wasting the investment. Like any other auction, two scenarios can happen with the bidder: (1) they lose the auction to the competition and do not need
to make any payment; (2) they win the auction and are charged some price to display the ad. The price charged for the advertisers varies based on the type of auction and the rules set by the ad exchange.4,5

Every advertiser taking part in the auction aims to maximise the surplus value (profit made or loss incurred based on the outcome of the auction) that they can obtain by participating in the auction. The surplus value depends on the number of users watching the ad and clicking on it; in other words, the surplus amount is dependent on the allocated ad and the context. On the other hand, winning the auction is completely dependent on the bid value placed by the advertisers, irrespective of the ad and the context. These call for the development of intelligent bidding strategies to optimise key performance indicators (KPIs), having the optimum bid value for the right advertisement at the appropriate time to maximise the benefits for the advertisers.6

Reinforcement learning
In the past couple of decades, reinforcement learning has become a bleeding-edge topic in AI and is now considered to be the go-to approach because of its potential to transform most businesses. As explained by Russell and Norvig,7 reinforcement learning is similar to natural learning processes (NLP) where a supervisor is not available and the learning is made purely based on trial and error — that is, it is a continuous learning algorithm that learns from mistakes. It is a field of machine learning (ML) in which an agent interacts with an environment with the goal of maximising cumulative reward through time.8

The way these reinforcement learning models work is analogous to a toddler making its way around the house and discovering new things. While doing this, the toddler discovers some fun things to do and seeks to repeat those activities again. On the other hand, sometimes it leads to unpleasant experiences like bumping their head or touching a hot iron. Over time, the child learns the activities it enjoys and avoids the unpleasant activities. In the world of reinforcement learning, the toddler is an agent, the action taken by the child is the action taken by the agent, the area where the child moves around is the environment, and the pleasant or unpleasant experience is the reward. A reinforcement learning agent takes an action, based on which the environment provides a positive or negative reward, and by learning from this, the agent takes the next action. The goal of the agent is to learn from the positive or negative feedback provided by the environment and take action to maximise the total reward.

The agent has two options to take next actions based on the rewards received. Either it can choose to take the actions that provide the maximum reward every time, trying to exploit the available action space to gain immediate payoffs (exploitation), or explore the action space with the aim of achieving long-term benefits (exploration). There is always a trade-off between exploration and exploitation in reinforcement learning. Exploitation is a greedy approach which makes use of partial information to receive rewards which might result in missing better rewards by not exploring unknown actions. On the other hand, exploration provides better information about the action space, but if the agent is exploring all the time, it might end up wasting a lot of time on futile actions. It is important for an agent to maintain a balance between exploration and exploitation while deciding the next best action.

REINFORCEMENT LEARNING AND ML
In traditional ML paradigms, algorithms are trained to learn patterns on a batch of historical training data. These batches come with labels or numerically valued outcomes. Once the desired outcome has
been modelled reasonably well, the models are deployed into business processes. With the passage of time, new business patterns can evolve, and a previously trained model may no longer be relevant. This problem is referred to as model drift. Reinforcement learning as an ML paradigm offers an alternative. It involves training an AI agent that learns to take optimal decisions based on its interactions with its environment and the reward signals it obtains. In the beginning, this training does use historical data, but the design of the agent allows it to be explorative even after it has been deployed. Every now and then, the agent takes an action that cannot be considered ‘optimal’ based on its prior experience. This allows the agent to explore alternative action choices, even as business patterns change. This continues, keeping the AI agent relevant with time.9

Reinforcement learning models, being online, even after deploying the model keep on learning and adapting the changes with every new prediction request, incessantly trying to improve the reward. Moreover, on the fly learning capability provided by live feedback further helps in adjusting the data abnormality. With the ever-changing advertising environment based on emerging trends, reinforcement learning appears to be the most logical choice.

**MABs AND CMAB**

The name ‘multi-armed bandit’ (MAB) came from a practical scenario of a gambler trying their luck with slot machines in a casino. Imagine that a casino has different slot machines, and the goal is to figure out which one can provide the best pay-out without losing too much money. Trying every machine once and picking up the machine that provides the best outcome every time after that is not the most effective strategy. The selected machine might provide a lucky outcome at first attempt but can be a sub-optimal option. Instead, the better approach is to go back to the bad performing machines to collect more information about them, which can help in more precise decisions afterwards, and finally making decisions on the sequence of slot machines to interact with.

MAB is a classic reinforcement learning example that aims to solve the exploration/exploitation dilemma faced by an agent using a sequential decision-making approach.10 At each step, the agent chooses one of several possible actions (arms) and gets a reward for the choice made from the environment. The main goal of the agent is to maximise the cumulative reward by using a combination of different exploring arms and exploiting arms that resulted in high rewards for the previous actions. In other words, at each step, the agent takes action on the environment based on policy to move from state to, upon getting a reward from the environment. The agent tries to improve the policy at every step to maximise the sum of rewards.

Contextual MAB (CMAB) is an extension to MAB, with additional information on ‘context’ associated with each arm. These contexts help in learning the policy function to take the optimal sequence of actions leading to maximum cumulative reward.

**CMAB FOR BID OPTIMISATION**

The main problem faced by advertisers is the issue of budget management for digital ads. With the digital advertisement scenario being a blind auction, there is no visibility on likelihood of winning the real-time bidding, so it becomes crucial to make intelligent decisions while bidding for the available ad inventory slot to ensure maximum ROI for the bid amount placed. If an advertiser has a monthly budget of, and if they bid too high and pay too much per click, their Artificial intelligence for smart bidding budget may run out and their ads will stop appearing; in turn,
they will miss out on potential impressions, clicks and conversions. The budget is typically set at a campaign level and affects all ad groups in that ad campaign. It constrains optimal advertising in two ways:

1. **Temporal** — across time, especially if a fixed advertising budget is set for a month, or even a lifetime, or when advertisers bid multiple times in a day;

2. **Cross-sectional** — across different features of advertisements on the publisher platform (like keywords for e-commerce platform, content on OTT platform, etc.) in each campaign at a fixed point in time.

The model needs to be built keeping budget constraints under consideration and hence, it is used as one of the contexts in the CMAB to decide policy function in order to take optimal actions. Apart from keeping budget constraints under consideration, other key features used as context include details of the ad group (ad group ID name, serving status, ad group bid, etc.), information of ad campaign (campaign targeting type, campaign type, campaign bidding strategy, bid optimiser, placement adjustment, etc.) and temporal features based on the frequency with which the advertiser participates in the real-time auction.

To achieve the goal of improving the ROI via the delivered ad, it is important to understand user behaviour towards the delivered ad.

For precise identification of user behaviour, the following performance metrics were considered:

1. **Number of impressions**: Once the ad request made by the user is completed and the ad is displayed, it is then called an impression. The number of impressions delivered for the desired period of the ad campaign gives the idea of success in the real-time bidding scenario;

2. **Cost of clicks**: The total cost of the clicks by users on the ad accumulated for the desired period of the campaign provides a direct idea of performance of the previous action taken on the bid;

3. **Discounted attributed sales**: Provides an idea of the future sales value based on the historic performance of the campaign.

As shown in Figure 2, we trained two agents: (1) **bid recommender** to develop an optimal bidding strategy; and (2) **dynamic predictor** to analyse the estimated performance for the decided bid. Upon receiving the input features to develop the context of the CMAB models, a decision is made, based on the performance of the previously recommended bid values, on whether the reinforcement learning models need to be trained on the bulk data. Once the context parameters and the performance parameters are collected, the agents work together to recommend an optimum bid value for the ad and using that dollar amount, the advertiser takes part in the real-time auction. The performance parameters obtained based on the suggested bid value are then added to the input data to the CMAB model, thereby completing the feedback loop. The feedback loop acts by considering a large number of features, which could be easily overlooked while deciding the bid amount manually, making the AI-based model much more efficient.

The two agents are dependent on each other, such that the output of one agent drives the other. Upon receiving context features and performance metrics, the bid recommender provides a bid value (next action) which could lead to maximising the cumulative reward based on the context features. The output from the bid recommender then triggers the dynamic predictor agent which analyses the historic performance and estimates the performance of the ad in the form of discounted sales value and future clicks and conversions,
to justify the action taken by the previous model. Finally, the dual agent engine decides the suggested bid value, which is then applied in the blind auction.

**Bid recommender**

The main goal of the bid recommender is to suggest bid values for the ad inventory slots to an advertiser by updating policy function based on context features, keeping the budget constraint under consideration. The idea is to predict bid instance 2, given that it has already computed the bid for values sequentially, keeping the last bid value under consideration but identifying the bid recommendations for all the bidding instances every time.

If an advertiser plans to participate in real-time auction in a period, at each time, the bid recommender produces bids for each bidding instance. To do this, it predicts a bid for each one of them sequentially (not simultaneously): first, it computes the bid for instance 1, then it computes the bid for instance 1, then it computes the bid for instance 3, knowing bid value for instances 1 and 2, and so on.

The approach is formulated using a sequential neural network, where the batch...
size is 1 and the sequence length is. With batch size of 1, the model architecture does not depend on the sequence length, and the addition or removal of instances does not affect the number of layers, output size of model, etc.

**Dynamic predictor**

Dynamic predictor acts as the performance monitor for the dual agent model. It uses such critical targets as impressions, clicks, cost, and conversion. To use these criteria as the reward for bid recommender, ‘discounted conversion’ value is calculated. A problem that arises here is that these critics have vastly different scales. To bring all the metrics to a comparable level, we used z-scaling such that each metric has a mean of 0 and standard deviation of 1, before calculating the discounted conversion. It gives the idea of future conversion rate for the ad campaign, which helps in estimating reward to predict the optimum bid amount, as the main goal is to maximise conversions and hence increase ROI.

The discounted reward based on the estimation of future conversions is calculated as:

\[ r = \epsilon_\text{d} + \sum_{d_i} \Upsilon^{d_i - d_0}\left(\epsilon_\text{d_i} - \epsilon_\text{d_0}\right) \]

where \( r \) is discounted reward, \( \epsilon \) is the number of conversions, \( \Upsilon \) is the discount factor, \( d \) is a list of days in the future for which discounted reward is calculated and \( i \) is the index for the list.

In the model, discounted reward was calculated for days; the discounted reward is denoted as:

\[ r = \epsilon + \Upsilon^6(\epsilon_{n+7} - \epsilon_{n}) + \Upsilon^{13}(\epsilon_{n+14} - \epsilon_{n+7}) + \Upsilon^{29}(\epsilon_{n+30} - \epsilon_{n+14}) \]

The output from the dynamic predictor helps in identifying the sequence of the bids which could lead to maximum cumulative reward, suggesting higher ROI.

**IMPLEMENTATION OF BID OPTIMISATION IN REAL LIFE**

**Amazon-sponsored product optimisation**

The digital revolution has had a direct impact on the retail industry, leading to a decline in sales in physical stores, even resulting in many of them shutting down; the global pandemic of COVID-19 has accelerated the process to a great extent, leading to increase in online shopping. One of the forefathers leading to such a revolution is Amazon.

The significant impact of Amazon on the transition from traditional brick-and-mortar shops to e-commerce is called the ‘Amazon effect’. It is usually associated with the variety of products on the platform, facilities such as same-day deliveries, 30-day exchange policy, etc., which attract consumers to move towards online shopping experience. This change in consumer behaviour has led to an increase in competition among manufacturers on the platform.

Amazon-sponsored products are the cost-per-click ads on the platform to promote product listings on it. With the increasing competition, every manufacturer aims for their product to be featured among sponsored products. To select sponsored products, Amazon takes help from ad exchange to create a real-time bidding environment for these advertisers to take part in. The bidding takes place on a daily basis. Manufacturers (advertisers) choose to select the frequency of bids to be placed in a day. The bid price in this scenario is placed on the search keywords, which means advertisers have the option to bid for multiple keywords in a day.

The decision on the required keywords for which a bid needs to be placed is made by competitors’ analysis. The list of search terms on which competitors gain conversions is identified and is considered to be the list of target keywords to bid on. These keywords, however, are quite dynamic in nature and are expected to change over time. Using these keywords as a feature for the model would necessitate updating the model architecture.
every time. As an alternative, we built a model architecture that is independent of the number of keywords by ensuring that:

1. The number of keywords is entered as a ‘batch’ dimension, instead of a ‘feature’ dimension. This means that predictions on one keyword are treated independently from predictions on the other keywords;
2. The keywords are represented as continuous vectors in space (embeddings). The model outputs a vector in that exact space, and we find the K-nearest neighbours to that prediction.

Upon identification of target keywords, the optimal bid value for each keyword is identified using the dual-agent model, to win the real-time auction and to ensure a high number of conversions in the future.

Jeunen et al.12 from Amazon developed a simulation environment ‘AuctionGym’ that enables researchers and professionals to use bandit learning for bidding strategy in an open auction. The open-source environment simulates advertising problems and closely resembles the real-life auction scenario at Amazon. The simulated environment presents an impression opportunity to the bidders, on which they could decide a bid price and the ad to be shown. The environment then decides the winning ad and its price, which is then shown as a sponsored product, possibly leading to conversion. The internal system in the AuctionGym environment consists of Bernoulli process to simulate whether the allocation decision leads to conversion, enabling an end-to-end auction scenario.

We took help from the open-source AuctionGym simulation environment to fine-tune our reinforcement learning models and strategies for deciding the target keywords, which could possibly lead to better conversion. Figure 3 shows the comparison between historical discounted conversions and the discounted conversions on applying the CMAB-based models (bid suggested by the agent). The discounted conversion value is the calculated reward value from z-scaled critic parameters used to select the optimum bid and to show the

Figure 3: Comparison between discounted conversion calculated from z-scaled critic targets based on the historic bidding strategy and bids suggested CMAB model
estimate of the performance based on the bid price set. As time progresses, the agent learns the bidding strategy better and starts outperforming manually suggested bids for the keywords. For a completely trained model, RL agent’s bid suggestion provided approximately 30 per cent better-discounted conversions than a human agent (a manual process).

**Google Ad budget optimisation**

JOT Internet Media, a Spanish company specialized in performance based digital marketing, helps its partners to reach new audiences relevant to their business, increasing and monetising their web traffic. JOT has a long history of experimenting with data science to identify data-driven ways to improve the service it provides to its clients, and to save data-crunching time for its digital marketing experts, freeing up more of their time for the creative part of the job.

In 2021, JOT Media teamed up with Amplify Analytix, a data science for business company, aiming to devise an intelligent bidding strategy for Google Paid Search. Just like Amazon sponsored products, Google uses real-time auctions to select advertisers for Google Ad Search, giving them the opportunity to attract more traffic towards their websites. The main goal of this partnership was to improve JOT's detection of non-trivial temporal search patterns, with a view to identifying the most relevant keywords and estimation of optimal bidding on the suggested sequence of keywords, meeting the budget constraint for the advertisers. To achieve the desired goals, CMAB agents were used.

Since paid searches are depended on a large number of factors, including target audience, search category, platform, device, browser, ad space, match type, temporal features in the interest of the day, month and season), the ad type, the content of the ad, campaign properties, geographic features etc., making a context feature set for CMAB agent to learn policy function from and to further achieve a better cumulative reward. Based on the learned policy function, the CMAB agents took action on decreasing or increasing the bid value as compared to the previously placed dollar amount, thereby suggesting an optimal bid, such that it helps in winning the real-time auction and can improve the number of clicks and click-through rate (CTR) for the website as well.

With the solution implemented in a live environment, the CMAB application to Google Search bid optimisation resulted in 15 per cent higher CTR, 7 per cent estimated increase in ROI and 38 per cent of digital expert time saved. Further, the time taken to onboard new digital specialists joining the team was reduced by approximately 12 months.

Our partnership with JOT Internet Media was enabled by the European Data Incubator, an Innovation Action project co-funded by the European Union to facilitate sustainable business incubation around big data. The application of the bandit technique that helped in improving both clicks and CTR, thereby better serving JOT’s clients — succeeded in developing the award-winning implementation of the algorithm using CMAB technique, to optimise Google Ads spending. Figure 4 shows a comparison between CTR observed by an advertiser using our (Amplify) bidding approach and the approach developed by a competitor. The exponential increase in the CTR depicts the learning curve of the CMB-based agent.

**INVESTING IN REINFORCEMENT LEARNING PAYS OFF**

After observing the positive impact on click-through rate (CTR) for the advertisers participating in the real-time auction for the Google Paid Search by application CMAB agent, Fernando Perales, Head of the Research Lab, and Strategic Partnerships at JOT says:
'After analysing the results, it is clear that the model helps to improve marketing campaign performance. In particular, its ability to digitise and replicate the optimisation processes carried out by human accounts, resulting in a significant reduction in time consumed by analyses and delivering a data-driven decision support system. This was made possible because the team sought to understand end-user roles, pain points and built a model that added value to daily work. We are very interested in continuing the collaboration to develop these predictions one or even two steps further. The gains are definitely worth it.'

Based on the success stories seen with Amazon sponsored product optimisation and Google Ad budget optimisation by the implementation of reinforcement learning, Laura Murphy, CEO of Amplify Analytix, says:

‘Over the years, we keep improving the recipe for success: a business problem or opportunity that is clearly articulated and understood by all stakeholders, a “translation” of the business problem into terms that can be handled by data science, a top-notch data science model built by the best data scientists, and a deep understanding of how the output or insights will be used by the business users in question are all crucial ingredients. Our best work happens when we create a multi-disciplinary, cross-organisation team, as we did with JOT, made up of those who need and those who build the insights, to enable us to change established ways of working and deliver improved performance. We encourage every team to get on the journey, experiment, learn fast and embrace occasional failure as part of the learning. We are convinced that the possible benefits, for now, and in the future, far outweigh any teething problems on the way to unlocking the value of your data.’

**CONCLUSION**

The drift of consumers towards digital media for shopping, entertainment, education, etc. resulted in development of mountains of data that could serve in-depth data analysis for depicting complex consumer behaviour. Meanwhile for companies around the world recovering from the impact of COVID-19 and currently dealing with increasing
inflation prices following the energy crisis, limited budget is allocated for marketing. In such a scenario, implementation of an intelligent bidding strategy by extracting meaningful information from vast amounts of data as efficiently and cost-effectively as possible helps in utilising the limited budget in the best possible way.

The solution lies in the application of CMAB technique to inform decision-makers to improve customer service and save data-crunching time for digital marketing experts. The reinforcement learning-based approach not only provides a more optimal bidding strategy than manual decision, but also can accommodate a huge set of constraints and contexts, making it a highly scalable approach.

References
10. Deshmukh et al., ref. 8 above.
12. Jeunen, O. et al., ref. 5 above.