ATNN: Adversarial Two-Tower Neural Network for New Item's Popularity Prediction in E-commerce

Shen Xin*, Zhao Li†, Pengcheng Zou†, Cheng Long*, Jie Zhang*, Jiajun Bu‡ and Jingren Zhou†
*Nanyang Technological University, Singapore
Email: {xin.shen, c.long, zhangj}@ntu.edu.sg
†Alibaba Group, China
Email: {lizhao.lz, xuanwei.zpc, jingren.zhou}@alibaba-inc.com
‡Zhejiang University, China
Email: bjj@zju.edu.cn

***Abstract*—The e-commerce era is witnessing a rapid increase in online shopping activities and rising new arrivals of items on e-commerce platforms every day. Identifying potential popular items accurately is of great importance in creating commercial value. Click-Through Rate (CTR) is a general indicator to evaluate item popularity. However, existing CTR prediction methods fail in new arrivals prediction because they are confronted with two main challenges: (1) sparse item features and missing item statistics; (2) high time complexity of computing for all pairs of users and items. To tackle the first challenge, we propose a novel Adversarial Two-tower Neural Network (ATNN) model for new arrivals CTR predictions by introducing an adversarial network to a two-tower network. We design a generator and a discriminator to better learn an item vector based on item profiles without item statistics. Moreover, we introduce a Deep & Cross Network (DCN) and a strategy of shared embeddings among multiple tasks to accurately predict CTR for new arrivals. To address the second challenge, we develop a strategy with an $\mathcal{O}(1)$ time complexity for a new item's popularity prediction. We first construct a user group and learn its mean user vector in a time-efficient manner. Then, an item's popularity score is calculated from its item vector along with the learned mean user vector, which leads to a huge decrease in computational overhead. This is the first work to introduce a generative adversarial network in an optimized two-tower neural network to efficiently learn new arrivals vectors and effectively make new arrivals predictions. We implement ATNN and examine it on a large-scale real-world dataset from one of the world's largest e-commerce platforms, "Tmall.com". Empirical results show that ATNN is strongly capable of learning item vectors from item profiles for CTR predictions. We also deploy ATNN on "Tmall.com". The online A/B test illustrates that ATNN is superior over human expert decisions in real-world applications. Furthermore, by introducing multi-task learning technology, we extend ATNN to another scenario, food delivery service, which is quite essential during the particular period of the COVID-19 pandemic. Deployment on one of the most popular food delivery platforms, "Ele.me", showcases that ATNN can be widely applied in various e-commerce services. Both offline and online experimental results demonstrate that ATNN can recognize attractive and welcoming new restaurants that have higher Value per Page View (VpPV) and generate more Gross Merchandise Volume (GMV).**

***Index Terms*—E-commerce, New arrivals prediction, Click-through rate prediction, Neural network, Generative adversarial network, Cold-start, Food delivery.**

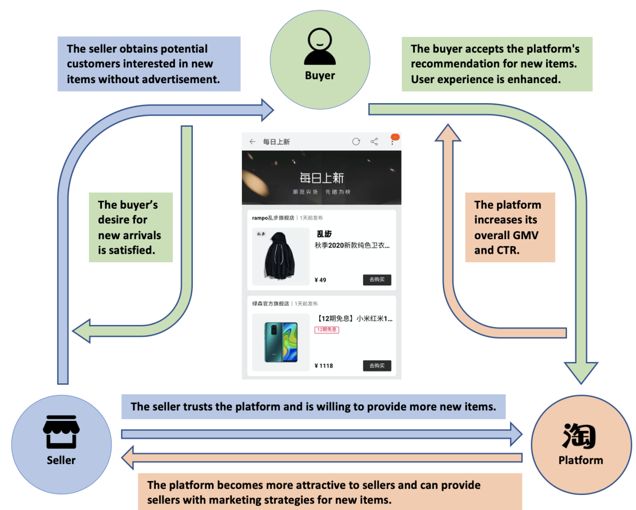

Fig. 1. A tripartite win-win mechanism based on new arrivals prediction

I. INTRODUCTION

With the rapid development of e-commerce, online shopping activities are surging. On a world-leading e-commerce platform, there are tens of millions of items released on the market every day. New arrivals prediction becomes a crucial task in the e-commerce industry. All online shopping participants, including sellers, buyers and platforms, benefit from accurate predictions of new arrivals. Personalized recommendations of new arrivals based on popularity estimates lead buyers to their preferred items, which satisfies buyers' desire for new arrivals and enhances user experience. Meanwhile, sellers are willing to provide more new items on this platform, because they benefit from increasing transactions of new arrivals. The e-commerce platform can also utilize the predicted popularity of new arrivals to implement various marketing strategies. Figure 1 shows a tripartite win-win mechanism based on new arrivals prediction.

Because there exists no general indicator to evaluate new arrivals' popularity, we adopt Click-Through Rate (CTR) as a key indicator to reflect item popularity. CTR Prediction is a regular task in e-commerce. Accurate CTR prediction leads buyers to their preferred items to improve user shopping experience, by reducing time for retrieving merchandise and discovering items that they are potentially interested in but unaware of. Items would be exposed to a group of potential buyers more effectively, and thus more transactions are generated. Under such circumstances, both sellers' sales volume and platform's Gross Merchandise Volume (GMV) would increase.

CTR prediction for new arrivals is a more challenging task than that for existing items since it suffers from a cold-start problem. We have no information about any user behaviors on new items, and thus the item statistics information is missing. Traditional methods focus on better utilization of item profiles: finding "nearest neighbors" based on contents for new arrivals, then making collaborative filtering predictions and recommendations. However, these methods do not exploit historical user-item interaction data when calculating item similarities. They suffer from very sparse features and missing interactive features.

To address the challenge of missing item statistics data, inspired by Generative Adversarial Networks (GANs) [1], we introduce a generator and a discriminator as an adversarial component to a Two-tower Neural Network (TNN) structure for better feature extraction. TNN consists of two neural networks. One tower maps user features to user embeddings. The other tower maps item features and historical user-item interactions to item embeddings. The generator extracts information from item profiles into vectors such that they are similar to the item vectors from the item tower's encoder. Besides, the discriminator aims to distinguish from two types of item vectors, i.e., maximizing the difference between item vectors obtained by the generator and the item encoder. The proposed Adversarial Two-tower Neural Network (ATNN) framework can be trained by two steps iteratively. The first step is to obtain both user and item vectors from full features for CTR predictions. The second step is to ensure that generated item vectors from item profiles are similar to normal item vectors. Inspired by transfer learning and multi-task learning, we let two item embedding layers share their embeddings. This strategy improves the feature extraction ability of the generator for CTR prediction and accelerates training progress.

To better obtain high-level features, we introduce Deep & Cross Network (DCN) [2] to the ATNN framework to enhance the ability to extract user and item features by learning explicit cross features of bounded-degree and traditional deep representations jointly.

Directly obtaining all new arrivals' popularity for a ranking task is confronted with a challenge of high time complexity. Typically, we need to construct samples of a Cartesian product of all new arrivals and all users to perform overall CTR prediction. However, it is not necessary to calculate the scores of all user-item pairs for determining the item popularity. To reduce the running time, we propose to learn and store a mean user vector of a user group before the prediction stage. We obtain one item's popularity score from its item vector and the stored mean user vector with only $\mathcal{O}(1)$ time complexity.

To examine ATNN, we design both offline and online experiments on a large-scale dataset from "Tmall.com". Offline experiments are divided into two parts. The first mainly aims to demonstrate the high generation ability of the adversarial component in ATNN. We test the performance of personalized CTR prediction compared with several popular methods being used in industry. The second evaluates the new arrivals prediction of ATNN by multiple business indicators, including Item Page View (IPV), adding to favorite, sales volume, etc. Furthermore, We deploy ATNN on one of the world-leading e-commerce platforms, "Tmall.com", to show its superiority of new arrivals prediction over human expert decisions.

Moreover, in the view of the COVID-19 pandemic, food delivery service becomes essential for daily life. The number of applications from new offline restaurants for joining food delivery platforms has increased unprecedentedly. We apply ATNN to this novel online-to-offline (O2O) scenario to choose attractive and welcoming restaurants by their popularity prediction. Based on current business needs, we introduce multi-task learning technology into ATNN to balance two evaluation metrics, Value per Page View (VpPV) and GMV. We also conduct both offline and online experiments. The experimental results show that ATNN successfully selects good restaurants that have higher VpPV and generate more GMV. In the meantime, we take corporate social responsibility to battle COVID-19.

We summarize the major contributions of this paper as follows:

1) We propose a novel Adversarial Two-tower Neural Network (ATNN) framework for new arrivals Click-Through Rate (CTR) predictions by enhancing the feature extraction capability. An adversarial network with a generator and a discriminator is developed to better learn item vectors from only item profiles without item statistics.
2) We design a Deep & Cross Network (DCN) and a strategy of shared embeddings among multiple tasks in the ATNN framework. The optimized ATNN model is utilized towards better CTR prediction for new arrivals.
3) We develop a high-efficiency method with $\mathcal{O}(1)$ time complexity for each new item's popularity prediction. This is achieved by computing one item's popularity score from its item vector and a pre-learned mean user vector of a user group.
4) We show both offline and online results of our experimental evaluation on a large-scale real-world dataset from "Tmall.com", which demonstrates the scalability of our method and its superiority in both extracting item vectors from sparse features and solving the task of new arrivals prediction.
5) We extend the proposed ATNN model to another popular scenario, food delivery service, by introducing multi-task technology. Empirical results on "Ele.me" show a kind solution to a cold-start problem in our food delivery platform and indicate extensive commercial values of the ATNN model.

The rest of the paper is organized as follows. In Section II, we give a brief review of related work. Section III presents our ATNN model. The experimental results are presented in Section IV. Further application studies are provided in Section V. Finally, we summarize the current work and suggest directions for future work in Section VI.

II. Related Works

In this section, we briefly introduce some related work about this paper in three aspects. First, we summarize the cold-start challenge in e-commerce and its current solutions. Second, we review popular methods for CTR prediction. Third, we revisit the development of GAN.

A. Cold-start problems in e-commerce

Cold-start problem is a traditional but very challenging task in recommender systems. There are two subproblems: recommendation for new users and new items.

There are some studies to solve the new user cold-start problem. An intuitive way is to let a new user get non-personalized recommendations built on top-seller rankings. Some models focus on exploiting side information, including user attributes and user social networks, to mitigate the cold-start problem. [3] defines a novel social collaborative filtering framework that generalizes standard item-based collaborative filtering to the cold-start recommendation setting. [4] proposes a principled approach based on user multi-relational factorization techniques. Feature mapping models are also used for the cold-start problem, such as [5], which normally learns a feature mapping between the side-information and one of the latent feature representations.

The new arrivals cold-start problem is more challenging. Unlike top-seller rankings recommendations, we cannot recommend "top-buyers" to new items. Meanwhile, we cannot construct social networks for new items without user-item interactive data. Classical content-based collaborative recommendation [6] and its multiple extensions, such as [7] and [8], are utilized to address the item cold-start problem. Recently, a variety of active learning strategies are proposed to solve this problem in the literature, such as [9] and [10]. However, active learning still relies on a few user-item interactions to train models. Under the real-world circumstances in industry, e-commerce markets require to predict overall CTR for new arrivals before they are released on the market. Active learning is incompetent in this case.

B. CTR predictions

CTR prediction is a crucial task in e-commerce, which guides users to click their favored items to increase transactions on the platform and commercial values. Traditional methods for this task is based on logistic regression, such as traditional Logistic Regression (LR) [11], Follow-The-Regularized-Leader (FTRL) [12], Large Scale Piece-wise Linear Model (LS-PLM) [13], etc. Because linear models are weak for extracting sophisticated feature interactions, various Factorization Machines (FMs) are proposed to improve CTR prediction by better learning interaction features [14].

Due to its powerful nonlinear modeling capabilities, DNNs are utilized for CTR prediction to automatically learn feature representations and high-order feature interactions recently. Two-tower structures are popular for CTR predictions. One item tower encodes item vectors from item features; the other user tower encodes user vectors from user features. Plenty of research works are related to this two-tower structure, such as [15], [16], etc. Deep & Cross Network (DCN) [2] introduces a novel cross network that is more efficient in learning certain bounded-degree feature interactions. Wide & Deep network [17] jointly trains wide linear models and deep neural networks to combine the benefits of memorization and generalization for recommender systems. Neural Factorization Machine (NFM) [18] is proposed for prediction under sparse settings, which seamlessly combines the linearity of FM in modeling second-order feature interactions and the non-linearity of neural network in modeling higher-order feature interactions. DeepFM [19] combines the power of factorization machines for recommendation and deep learning for feature learning in a new neural network architecture.

DNNs have been widely deployed in industrial applications. YouTube uses a deep candidate generation model with a separate deep ranking model for personalized recommendations [20]. Google introduces a large scale multi-objective ranking system for recommending videos to watch next [21]. Alibaba utilizes Deep Interest Network (DIN) to adaptively learn the representations of user interests from historical behaviors with respect to a certain advertisement by designing a local activation unit [22] [23]. However, these models need complete profiles of users and items, as well as user-item interactions, and they are not designed for solving cold-start problems.

C. Generative adversarial network

GANs are a class of machine learning systems invented by Goodfellow et al. in 2014 [1]. In a typical setting, GAN is composed of two components: one is a generative network G; the other is a discriminative network D. The generator G generates samples from its distribution. The discriminator D determines whether a sample is real data or generated data. The training objective of a generative network is to increase the error rate of the discriminative network. The discriminative network learns to distinguish candidates produced by the generator from the true data distribution. Both D and G are trained by a minimax game.

GANs are widely used in image recognition tasks. Conditional GANs are an extension of the GAN framework. For example, Stacked Generative Adversarial Networks (StackGANs) focus on improving the quality of the image conditioned on text descriptions [24]. The Deep Convolutional Generative Adversarial Networks (DCGAN) is used to improve GAN's performance by adding new constraints and modifying the CNN architecture [25]. GANs succeed in image super resolution [26], image to image translation [27], face aging [28] and video generation [29].

There are also some studies on GANs for information retrieval. For example, IRGAN [30] uses a generative model to learn to fit the relevance distribution over documents via the signals from a discriminative model and the discriminative model to exploit the unlabelled data selected by the generative model to achieve a better estimation for document ranking. Adversarial Personalized Ranking (APR) [31] enhances the pairwise ranking method Bayesian Personalized Ranking (BPR) by performing adversarial training. A general Minimax Game based method for selective Transfer Learning (MGTL) [32] is proposed for item recommendations by selecting high-quality source data to improve the transferring ability among different domains.

To predict popularity of new arrivals, we introduce a generative adversarial network in a two-tower network and utilize item vectors explicitly obtained from the two-tower network to enhance the generator's ability of feature extraction. We use item vectors obtained by the well-trained generator to accurately predict CTR and popularity for new arrivals.

III. Adversarial Two-tower Neural Network

In this section, we first describe the problem of new arrivals prediction. CTR is regarded as a crucial indicator of this problem. Then we present our Adversarial Two-tower Neural Network (ATNN) for ranking all new arrivals according to their popularity in detail.

A. Problem description

We aim to solve the cold-start problem for predicting popularity of new arrivals on e-commerce platforms. Since there is no common evaluation for new arrivals popularity, we use CTR prediction as a crucial task to solve this cold-start problem. Before releasing new items on the market, we utilize the model to make personalized recommendations according to CTR predictions. Accurate CTR predictions lead buyers to their preferred items, which satisfies buyers' consumption desire for new arrivals and enhances user experience. Meanwhile, sellers are willing to provide more new items on this platform, because they benefit from the increasing number of transactions for new arrivals.

Moreover, our main goal is to discover the potential popular items among new arrivals. Nevertheless, it is difficult for a model to evaluate the item popularity. We adopt that if one item has high CTR for a majority of buyers, it should be attractive with a high probability. Therefore, we extend our model to obtain the popularity of new arrivals among all users based on large-scale industrial data. This model enables to obtain the popularity of new items before releasing them on the market.

Specifically, we use item information of the first 30 days since they are released on the market as training data. We collect the statistics data of these items, including the number of Page Views (PV), the number of Unique Visitors (UV), series of user behaviors, such as clicking, adding to cart, adding to favorite and purchasing. We also obtain item profiles and user profiles as features of the training data. Item profiles

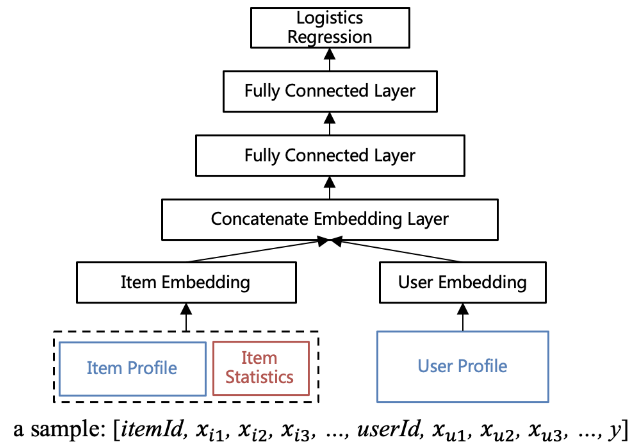

Fig. 2. A standard DNN model for pairwise user-item CTR prediction

contain seller information, product names, product images and categories. User profiles consist of personal data, such as user name, gender, location information, purchase preference, purchase power ratings, etc. New arrivals only have item profiles without item statistics information. Our goal is to rank all new arrivals according to their popularity among all users in the platform.

B. Two-tower neural network for ordinary pairwise user-item CTR prediction

DNNs are widely utilized for ordinary CTR prediction to automatically learn feature representations and high-order feature interactions. Item vectors obtained by DNNs can be used in various tasks, including popularity prediction for new arrivals.

Figure 2 shows a standard DNN model for pairwise user-item CTR prediction. It is a classical method that first concatenates an item embedding and a user embedding. We cannot obtain item vector nor user vector by this model.

To explicitly capture item vectors for new item's popularity prediction, we construct a two-tower neural network shown in Figure 3. The left tower extracts information from item profiles and item statistics to achieve item vectors; the right tower utilizes user profiles to get user vectors. We can explicitly capture item vector and user vector by this two-tower DNN. Obtained item vector and user vector can be utilized to train other models and solve other tasks related to pairwise CTR predictions. We use the obtained item vectors in our proposed adversarial network to train a generator, which will be described later.

We train the models by feeding each pair of one item and one user to the network, including the detailed interactions between every user-item pair. A sample of input is as follows:

$$[itemID, x_{i1}, x_{i2}, x_{i3}, ..., userID, x_{u1}, x_{u2}, x_{u3}, ..., y]$$

where $itemID$ and $userID$ are unique identifiers for each item and user, x_i and x_u indicate features of an item and a user, respectively, and y is the label for CTR prediction.

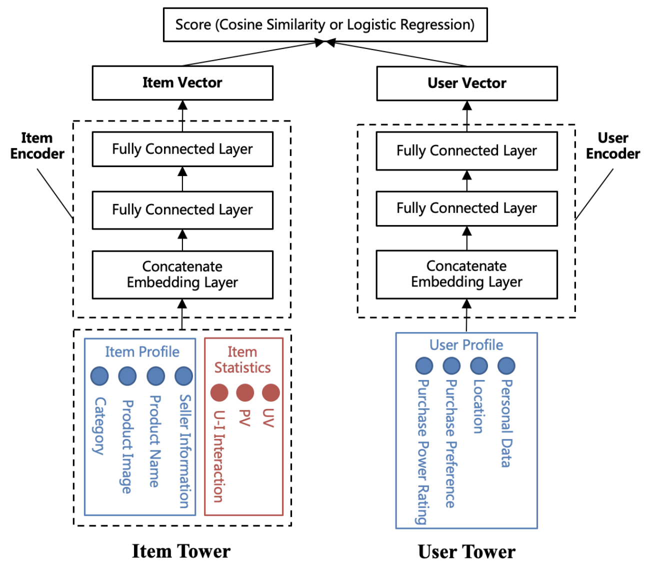

Fig. 3. A two-tower neural network for explicitly capturing item vector and user vector in pairwise user-item CTR prediction

C. Adversarial two-tower neural network framework for new arrivals CTR prediction

New arrivals CTR prediction differs from ordinary CTR prediction due to the lack of user-item interactions. For newly released items on the market, we usually have only a few user behaviors on them compared with the ordinary scenario. The data of clicking, adding to cart, adding to favorite and purchasing are too sparse to train a neural network. In addition, for new arrivals that have not been released on the market, there is no item statistics data. All classical methods face the challenge of missing item statistics information including PV, UV and series of user behaviors.

Motivated by the idea of GANs, we design an item generator and a discriminator for better learning item vectors from only item profiles. As we mentioned above, an original two-tower DNN model is able to achieve item vectors and user vectors, because there is an explicit layer behind item encoder and user encoder. We utilize the item vectors generated by the two-tower network to enhance the generator's ability of feature extraction. The quality of the generated item vector and user vector influences the precision of CTR prediction.

We propose to introduce an adversarial network in the two-tower structure for CTR predictions and call the resulting two-tower structure as *Adversarial Two-tower Neural Network* (ATNN) framework. The ATNN structure is shown in Figure 4. The left part is the adversarial component to learn to better extract item vectors using item profiles without any item statistics.

The generator is designed to generate item vectors from item profiles such that the generated item vectors are similar to item vectors generated by the item encoder, which is learned from both item profiles and item statistics. The discriminator is designed to distinguish generated item vectors generated by the item generator from item vectors generated by the item encoder. The generator and discriminator play a minimax game to improve their performance. The loss function of the discriminator is based on the similarity between two types of item vectors and is defined as $\mathcal{L}_s$.

In addition, we use all item vectors generated by both the generator and the encoder to train the CTR prediction model in the framework. The loss function of the original two-tower model for the CTR prediction task is defined as $\mathcal{L}_i$. The loss function of CTR prediction between the generated item vectors and user vectors is defined as $\mathcal{L}_g$.

$$\mathcal{L}_i = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i))$$

$$\mathcal{L}_g = -\frac{1}{N}\sum_{i=1}^{N}\left(y_i \log \hat{y}_i^{(g)} + (1-y_i)\log(1-\hat{y}_i^{(g)})\right)$$

where $y_i \in \{0,1\}$ is the label indicating whether the user clicks the item according to the records of user behaviors, $\hat{y} \in (0,1)$ is the label predicted based on the item vector and the user vector, $\hat{y}^{(g)} \in (0,1)$ is the label predicted based on the generated item vector and the user vector and N is the number of training samples.

We train this model by feeding each pair of one item and one user to the network, including the detailed interactions between every user-item pair. We iteratively optimize ATNN by reducing these three losses mentioned above. Both generation ability and CTR prediction quality can be enhanced.

Deep & Cross Network (DCN) [2] is utilized in all generators and encoders in the framework. DCN introduces a novel cross network that is more efficient in learning certain bounded-degree feature interactions without manual feature engineering. DCN explicitly applies feature crossing at each layer and adds negligible extra complexity to the DNN model. In e-commerce, there are plenty of high level features, e.g., item PV, seller PV and category PV. Manual feature engineering usually considers 2-level or 3-level features. Utilizing DCN can automatically obtain these bounded-degree features. In Section IV, we design a controlled experiment to test performance on both ordinary fully connected network and DCN in our model.

Furthermore, inspired by transfer learning and multi-task learning, we let two item embedding layers share their embeddings. The embedding layers map large-scale sparse features to low-rank vectors, requiring a large number of input samples to train a better model. Sharing features between embedding layers improves the generator component's ability for mapping item profiles to vectors.

We summarize the training steps for ATNN in Algorithm 1.

In each iteration, we first optimize ATNN by the following loss function:

$$\mathcal{L}_i(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y)$$

where $\boldsymbol{X}_i$ and $\boldsymbol{X}_u$ indicate features of one item and one user, respectively, $f_i(\boldsymbol{X}_i)$ and $f_u(\boldsymbol{X}_u)$ indicate item vector and

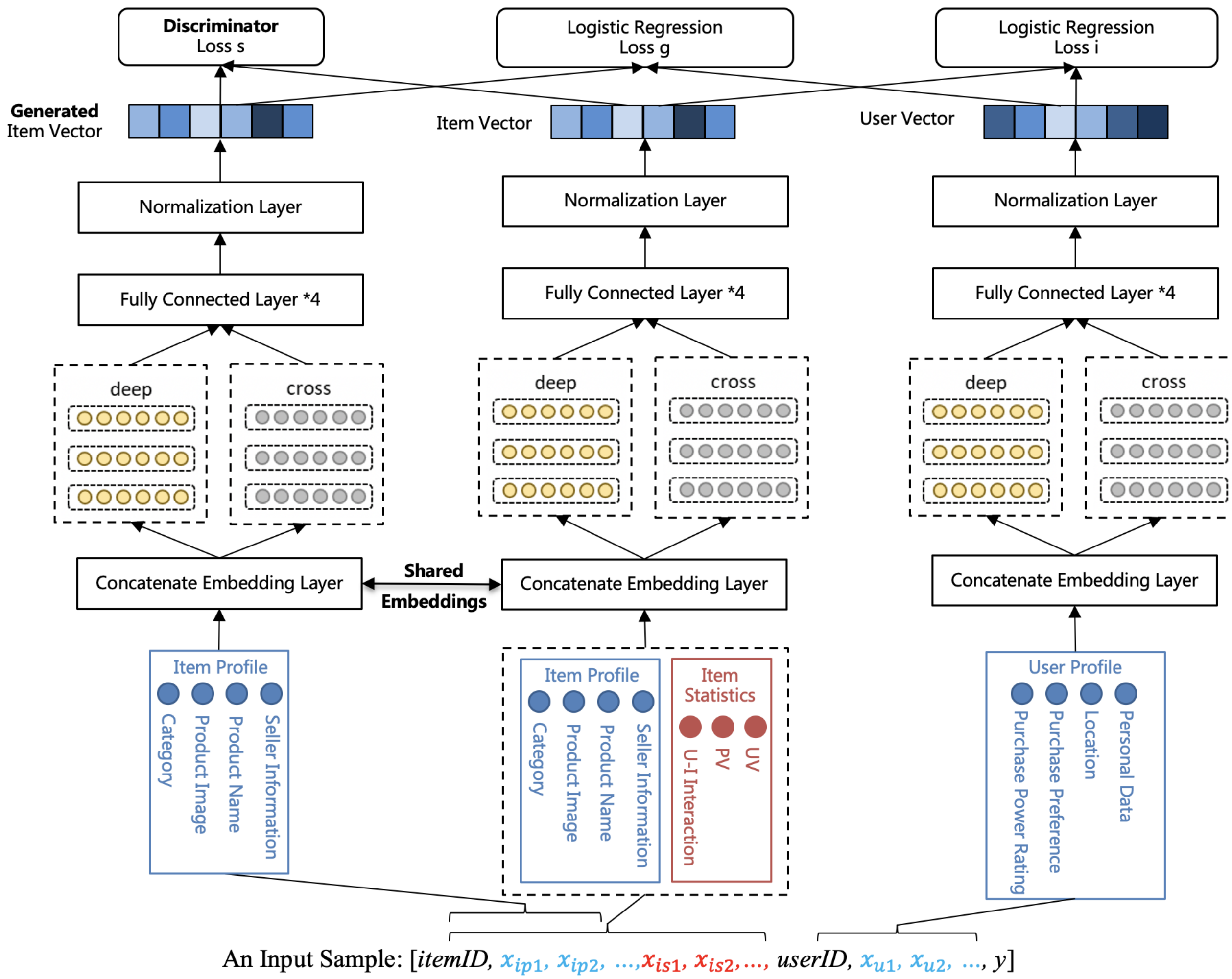

Fig. 4. Adversarial two-tower DNN model for new arrivals CTR prediction in E-commerce

user vector obtained by the item encoder and user encoder, respectively. $H(\cdot)$ function indicates CTR prediction score between one item and one user.

$\mathcal{L}_i$ uses logistic regression to evaluate the CTR predictions from both item profiles and item statistics information, according to given labels. In this step, we optimize CTR prediction by using item tower and user tower.

Then, we optimize ATNN by the following loss function:

$$\mathcal{L}_g(H(g(\boldsymbol{X}_{ip}), f_u(\boldsymbol{X}_u)), y) + \lambda\mathcal{L}_s(S(g(\boldsymbol{X}_{ip}), f_i(\boldsymbol{X}_i)))$$

where $\boldsymbol{X}_{ip}$ is the features of one item's profiles, $g(\boldsymbol{X}_{ip})$ is the generated item vector, λ is a weighting parameter to balance two losses and $S(\cdot)$ function indicates the similarity between a generated item vector and a normal item vector.

$\mathcal{L}_g$ uses logistic regression to evaluate the CTR predictions from only item profiles information, according to given labels. $\mathcal{L}_s$ uses the mean squared error as follows:

$$\mathcal{L}_s(X) = mean((1 - x_i)^2)$$

where $\mathcal{L}_s$ evaluates average similarity between generated item vectors and normal item vectors. In this step, we minimize the difference between generated item vector from the generator and item vector from the item encoder.

D. Large-scale new arrivals popularity prediction based on ATNN

We aim to discover potential popular new arrivals by ranking all items by their popularity. However, there is no general evaluation for rating items popularity. Based on a reasonable assumption that if one item has a high CTR for a majority of buyers, this product is very likely to be more attractive, we can utilize ATNN to estimate new arrivals popularity.

However, using a pairwise user-item CTR prediction model to achieve new arrivals popularity faces a critical challenge of high time complexity. In practice, for ranking new arrivals, we require to obtain all new arrivals popularity. In the prediction step, we need to construct a Cartesian product of all new arrivals and all users. Therefore, the time complexity of the prediction stage is $\mathcal{O}(N_U * N_{NA})$, where N_{NA} indicates the

Input: user features $\boldsymbol{X}_u$, item profile features $\boldsymbol{X}_{ip}$ and item statistics features $\boldsymbol{X}_{is}$

Output: generated item vector $\boldsymbol{X}_g$

$\boldsymbol{X}_i \leftarrow \mathrm{CONCAT}(\boldsymbol{X}_{is}, \boldsymbol{X}_{ip})$

for *number of training iterations* **do**

  Update D by minimizing loss: $\mathcal{L}_i(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y)$;

  Update G by minimizing loss: $\mathcal{L}_g(H(G(\boldsymbol{X}_{ip}), f_u(\boldsymbol{X}_u)), y) + \lambda\mathcal{L}_s(S(G(\boldsymbol{X}_{ip}), f_i(\boldsymbol{X}_i)))$;

  $\boldsymbol{X}_g \leftarrow G(\boldsymbol{X}_{ip})$;

end

Note:

$\mathcal{L}_i$: ctr prediction loss between item vector $G(\boldsymbol{X}_i)$ and user vector $f_u(\boldsymbol{X}_u)$

$\mathcal{L}_g$: ctr prediction loss between generated item vector $G(\boldsymbol{X}_{ip})$ and user vector $f_u(\boldsymbol{X}_u)$

$\mathcal{L}_s$: similarity loss between generated item vector $G(\boldsymbol{X}_{ip})$ and item vector $f_i(\boldsymbol{X}_i)$

Algorithm 1: ATNN Algorithm

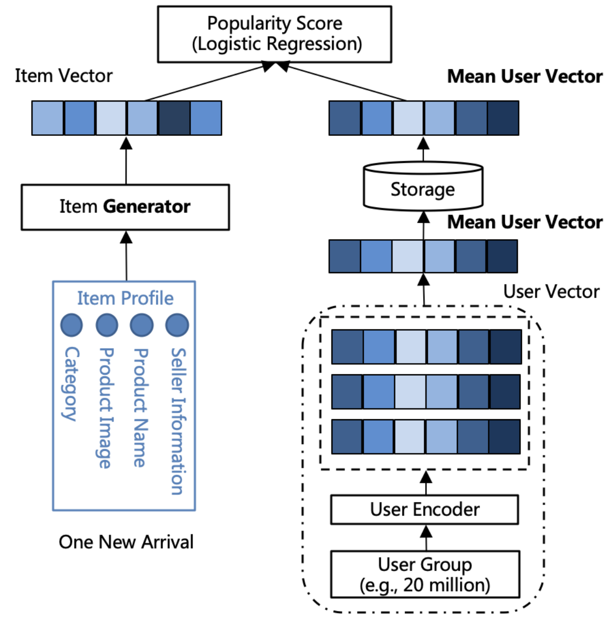

Fig. 5. ATNN for new arrivals popularity prediction

number of new arrivals. In one of the largest e-commerce platforms, there are hundreds of millions of existing users and millions of items released on the market every day. An $\mathcal{O}(N_U * N_{NA})$ algorithm does not work in a real-world system, especially in a world-leading e-commerce platform.

To rank new arrivals, it is unnecessary to obtain all pairs of user-item CTR predictions. Instead, we select the top 20 million active users who prefer new arrivals and regard them as a user group. We learn and store their mean user vector at the training stage. When predicting one item's popularity, we only need to utilize the stored mean user vector to make the prediction, which reduces the running time complexity for one item from $\mathcal{O}(N_U)$ to $\mathcal{O}(1)$. Figure 5 shows the efficient ATNN model for new arrivals popularity prediction.

IV. Experiments and Results

We implement and evaluate our algorithm on a large-scale dataset of a world-leading e-commerce company. In this section, we design two offline experiments and one online experiment. The first offline experiment aims to demonstrate the high generation ability of the adversarial component in ATNN. The second offline experiment evaluates the new arrivals popularity prediction of ATNN by multiple business indicators. The online experiment shows results of A/B test in a real-world e-commerce.

A. Offline experiments on item generation ability of ATNN

In this experiment, we aim to demonstrate the high generation ability of the adversarial component in ATNN compared with other popular methods.

1) Data description: We conduct the experiment on a real-world dataset. There are 23,107,452 items, 4 million users and 40 million user-item interactions in this large-scale dataset. User profiles include personal data, location, purchase preference, purchase power rating, etc. Item profiles contain seller information, product name, product image, category, etc. Item statistics include UV, PV, user behaviors among user-item interactions, etc. The numbers of raw features of user profiles, item profiles and item statistics are 19, 38 and 46, respectively. We apportion the data into training and test sets, with an 80-20 split.

2) Compared algorithms: To the best of our knowledge, no existing CTR prediction model utilizes an adversarial component to generate item vectors in the case of missing item statistics. Our baseline algorithms are as follows:

- Gradient boosting decision tree (GBDT) [33]
- Two-tower neural network with fully connected layers (TNN-FC)
- Two-tower neural network with deep & cross network layers (TNN-DCN)
- Adversarial two-tower neural network (ATNN)

GBDT is a popular machine learning method for regression and classification problems in industrial applications. It produces a prediction model in the form of an ensemble of weak prediction models. TNN-FC is our baseline method. TNN-DCN is tested to show whether DCN can improve the prediction.

3) Experimental settings: Categorical features cannot be fed into networks directly. We map these categorical features to fixed-length vectors according to their numbers of categories. For instance, user identifier, user occupation, user category preference, item category and item sub-category are pre-processed into 16, 8, 16, 6 and 16 dimensional vectors, respectively. We use the same network structure for obtaining generated item vector, item vector and user vector. The numbers of dimensions in DCN are 512, 256 and 128. The numbers

TABLE I
RESULTS OF OFFLINE EXPERIMENTS ON ITEM GENERATION ABILITY OF ATNN

Model	AUC for only item profiles (cold start scenario)	AUC for complete item features (ideal baseline)	Performance degradation due to missing item statistics
GBDT	0.6149	0.6590	-6.69%
TNN-FC	0.5934	0.6048	-1.88%
TNN-DCN	0.6860	0.7169	-4.31%
ATNN	**0.7121**	0.7124	**-0.04**%

of dimensions for fully connected layers after DCN are 256, 256, 256 and 128. Therefore, the numbers of dimensions of generated item vectors, item vectors and user vectors are all 128. Empirically, we set the weighting parameter λ as 0.1 in our experiments.

4) Evaluation metrics: We use Area Under the Curve (AUC) to evaluate CTR predictions. Higher AUC means better predictions. However, in this experiment, we aim to evaluate the feature extraction ability of the proposed ATNN. Better predictions indicate a more powerful item generation ability. Performance degradation due to missing item statistics is defined as:

$$\text{Performance degradation} = \frac{\text{AUC}_{\text{profile only}} - \text{AUC}_{\text{complete}}}{\text{AUC}_{\text{complete}}}$$

5) Experimental results and analysis: Table I summarizes the experimental results. The following conclusions can be drawn from the experimental results:

1) ATNN and TNN-DCN outperform GBDT and TNN-FC in the traditional CTR prediction from complete item features including both item profiles and item statistics. Introducing DCN as a part of encoders in a two-tower network can better extract bounded-degree features to improve CTR prediction.
2) Performance decreases significantly for all three compared algorithms in new arrivals CTR predictions, i.e., in the case of missing item statistics. Specifically, performance degradations are -6.69%, -1.88% and -4.31% for GBDT, TNN-FC and TNN-DCN, respectively.
3) The generator in ATNN has a strong ability to extract item vectors from only item profiles. There is nearly no performance degradation from using item vectors obtained by the original item encoder and item vectors obtained by the generator, for CTR predictions. This experimental result demonstrates the high generation ability of the adversarial component in ATNN.

B. Offline commercial value validations on new arrivals popularity prediction of ATNN

In this offline experiment, we introduce several key indicators of e-commerce to evaluate the new arrivals popularity prediction of ATNN and to validate its commercial value.

1) Data description and Experimental settings: Both the dataset and the experimental settings, including network settings and parameter settings, are the same as the previous experiment. However, differing from pairwise user-item CTR predictions, this experiment is designed to show the quality of new arrivals popularity prediction. We first obtain popularity scores for all new arrivals. Then we divide these new arrivals into five groups to observe their performance in the first 30 days starting from the release day. Since there is no general evaluation metrics for judging the item popularity, we introduce some important business indicators in e-commerce for evaluations.

2) Evaluation metrics: We utilize several business evaluation metrics to present the quality of new arrivals popularity predictions. Three indicators are listed as follows:

1) Item Page View (IPV): the number of requests for viewing an item webpage. An item with a larger IPV means that its item webpage has been watched for more times.
2) Add to Favorite (AtF): the number of user behaviors of adding the item to their favorite list. AtF explicitly shows how many users like this item.
3) Gross Merchandise Volume (GMV): the prevailing evaluation metric in online retail to indicate a total sales dollar value for merchandise.

The observation terms are set as 7 days, 14 days and 30 days. We calculate the average IPV, average AtF and average GMV for each item group. Ideally, the higher-ranking group should have larger IPV, AtF and GMV.

3) Experimental results and analysis: Table II summarizes the experimental results in terms of IPV, AtF and GMV for 7-day, 14-day and 30-day observation terms in five item groups, which are divided by their predicted popularity.

We group new arrivals by their popularity scores. The first group containing the top 20% predicted popular new arrivals achieves the best performance among all three evaluation metrics, i.e., IPV, AtF and GMV, for the first 7 days and 14-days on the market. All these business indicators present a positive correlation with the popular scores predicted by ATNN. This experiment result shows that the proposed ATNN succeeds in the popularity prediction of new arrivals in e-commerce.

C. Online experiments on new arrivals popularity prediction of ATNN

We deploy ATNN in an e-commerce platform of a world-leading Internet company. We design an online A/B test by selecting 300 thousand potential popular new arrivals among tens of millions of items to compare ATNN with manual selections by experts. The average time for the first five successful transactions is used to evaluate the popularity of new arrivals. A new item is more attractive if it has five successful transactions in a shorter time. The average time for the first five successful transactions reflects the overall quality of new arrival selections. Shorter average time means a better prediction for new arrivals.

TABLE II
RESULTS OF OFFLINE COMMERCIAL VALUE VALIDATIONS ON NEW ARRIVALS POPULARITY PREDICTION OF ATNN

Popularity Ranking (Top %)	7-day IPV	14-day IPV	30-day IPV	7-day AtF	14-day AtF	30-day AtF	7-day GMV	14-day GMV	30-day GMV
0-20	63.94	132.24	199.30	1.06	2.19	3.46	51.40	110.50	226.32
20-40	42.95	92.43	148.06	0.70	1.50	2.49	28.84	63.63	147.35
40-60	39.09	81.65	135.23	0.60	1.26	2.23	41.63	91.74	240.46
60-80	21.65	45.10	77.44	0.37	0.72	1.28	16.68	35.58	91.55
80-100	16.20	32.43	53.81	0.27	0.55	0.92	13.34	27.85	73.14
Average	36.77	76.77	122.77	0.60	1.24	2.08	30.38	65.86	155.76

TABLE III
RESULTS OF ONLINE A/B TEST

Expert selection	ATNN selection	Improvement
10.47 days	9.72 days	7.16%

The experimental results of the online A/B test are shown in Table III. The average time of the first five transactions of new arrivals selected by experts is 10.47 days. Outperforming the experts by 7.16%, our ATNN method not only improves the popularity prediction of new arrivals but also reduces human resources costs.

D. Deployments

ATNN has been deployed on Alibaba's online real-time data engine since August 2019. It is implemented in TensorFlow. The real-time data engine can obtain user behaviors, including clicking, adding to favorite, purchasing, etc. ATNN is capable of calculating new arrivals' popularity. Nowadays, two downstream applications benefit from ATNN's predictions. One is personalized search and recommendation, the other is smart selection of items for promotions.

V. FURTHER APPLICATIONS

We introduce multi-task learning technology and extend the ATNN model to a popular and essential scenario, food delivery service, to solve the cold-start problem of new restaurants. First, we describe the background and problem definition. Then, we propose the extended ATNN model for the food delivery scenario. Offline and online experimental settings and results are included at the end of this section.

A. Background and problem description

The rapid spread of the novel coronavirus, COVID-19, has raised alarm bells all over the world. To lower the risk of COVID-19 exposure and spread, a large number of restaurants and bars have suspended dine-in services and a lot of residents choose to stay home and keep social distancing. The demand for food delivery service has increased unprecedentedly.

Ele.me is an online-to-offline (O2O) catering and food delivery platform. As part of Alibaba Group's "New Retail", Ele.me had become the largest food delivery service in China with a 53.4% market share by 2019. [1]

[1]https://daxueconsulting.com/o2o-food-delivery-market-in-china/

During the battle period of COVID-19, tens of millions of restaurants have applied to Ele.me for providing delivery service on the platform. In order to try our best to ensure consumer safety and food quality, we must manually review the qualifications of each store, sign the contract and provide support services. It is impossible to review all applicants carefully in a short time. To create more commercial values, a model is required to assist us with selecting the most attractive and welcoming restaurants automatically, so that we can provide services to these merchants as soon as possible.

The problem turns out to be very similar to the new item's popularity prediction in e-commerce because two main challenges still exist: (1) sparse item features and missing item statistics; (2) high time complexity of computing for all pairs of users and items. Since all new applicants have not joined our platform, we only have the basic features of these restaurants, for instance, brands, locations, themes, cuisines, etc., without any retail data. In addition, calculation on each pair of one restaurant and one user is time-consuming and unnecessary for predicting the popularity of a new merchant, because we have tens of millions of new restaurants and hundreds of millions of users. Motivated by this, we attempt to modify our ATNN model to apply in the food delivery platform to solve the problem of new restaurant's popularity prediction.

Differing from choosing CTR prediction as a crucial task for new arrivals popularity prediction in e-commerce, in our real-world business scenario, Value per Page View (VpPV) and Gross Merchandise Volume (GMV) are keys to increasing business value. Unlike traditional online shopping, users tend to make quick decisions when they select food. The number of overall PVs in the food delivery platform should be much less than typical e-commerce. Current traffic allocation methods are difficult in creating more PVs. Based on the limited overall PVs, VpPV is crucial for evaluating a new restaurant. Besides, GMV is a general metric for measuring business value, which directly indicates the market size. Therefore, we change the training task from CTR prediction to VpPV prediction and GMV prediction.

B. Extended ATNN model for the food delivery scenario

As an O2O service, food delivery is location sensitive. Users often choose merchants close to them to keep food warm and fresh. We divide users into different groups based on their location information. Similar to Section III, we use mean

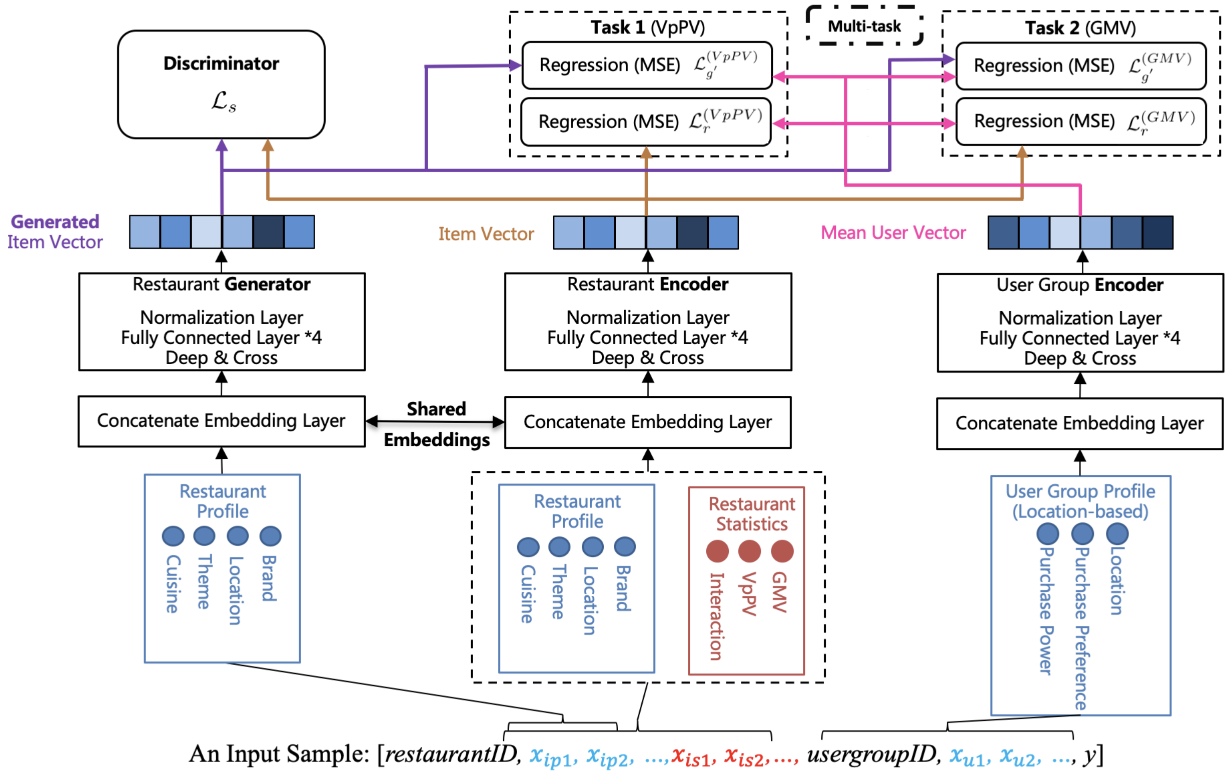

Fig. 6. Extended adversarial two-tower DNN model for new restaurant's popularity prediction in food delivery service

user features to replace single-user features to reduce time complexity.

To precisely predict new restaurant's popularity based on both VpPV and GMV aspects, we introduce multi-task learning technology in ATNN. Figure 6 shows the extended ATNN model to solve the cold-start challenge for new restaurants in a food delivery platform. By sharing networks, especially the restaurant representation among multiple tasks, our model can achieve better performance.

The loss functions for VpPV prediction task and GMV prediction task are as follows:

$$\mathcal{L}_r^{(GMV)} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{(GMV)} - \hat{y}_i^{(GMV)}\right)^2$$

$$\mathcal{L}_{g'}^{(GMV)} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{(GMV)} - \hat{y}_i^{(g')(GMV)}\right)^2$$

$$\mathcal{L}_r^{(VpPV)} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{(VpPV)} - \hat{y}_i^{(VpPV)}\right)^2$$

$$\mathcal{L}_{g'}^{(VpPV)} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{(VpPV)} - \hat{y}_i^{(g')(VpPV)}\right)^2$$

where $y_i^{(GMV)}$ is the GMV of $restaurant_i$, $\hat{y}_i^{(GMV)}$ is the GMV predicted based on the item vector and the mean user vector, $\hat{y}^{(g')(GMV)}$ is the GMV predicted based on the generated item vector and the mean user vector and N is the number of training samples.

In each iteration, we first optimize ATNN by the following loss function:

$$\mathcal{L}_r^{(GMV)}(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y^{(GMV)})$$
$$+\lambda_1 \cdot \mathcal{L}_r^{(VpPV)}(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y^{(VpPV)})$$

Then, we optimize ATNN by the following loss function:

$$\mathcal{L}_{g'}^{(GMV)}(H(g(\boldsymbol{X}_{ip}), f_u(\boldsymbol{X}_u)), y^{(GMV)})$$
$$+\lambda_1 \cdot \mathcal{L}_{g'}^{(VpPV)}(H(f_i(\boldsymbol{X}_{ip}), f_u(\boldsymbol{X}_u)), y^{(VpPV)})$$
$$+\lambda_2 \cdot \mathcal{L}_s(S(g(\boldsymbol{X}_{ip}), f_i(\boldsymbol{X}_i)))$$

where $\boldsymbol{X}_{ip}$ is the features of one restaurant's profiles, $g(\boldsymbol{X}_{ip})$ is the generated item vector, λ_1, λ_2 are weighting parameters to balance two losses and $S(\cdot)$ function indicates the similarity between a generated item vector and a normal item vector.

We summarize the training steps for extended ATNN with multi-task learning in Algorithm 2.

C. Experimental settings and results

We implement both offline and online experiments to test extended ATNN in a food delivery scenario. Offline experiments are designed to verify the effectiveness of our model compared with other popular DNN models. Online experiments are used for resolving real-world business to show that our model outperforms human expert decisions in creating business value.

Input: user group features $\boldsymbol{X}_u$, restaurant profile features $\boldsymbol{X}_{ip}$ and restaurant statistics features $\boldsymbol{X}_{is}$

Output: generated vector $\boldsymbol{X}_g$

$\boldsymbol{X}_i \leftarrow \mathrm{CONCAT}(\boldsymbol{X}_{is}, \boldsymbol{X}_{ip})$

for *number of training iterations* **do**

Update D by minimizing loss: $\mathcal{L}_r^{(GMV)}(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y^{(GMV)}) + \lambda_1 \cdot \mathcal{L}_r^{(VpPV)}(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y^{(VpPV)})$;

Update G by minimizing loss: $\mathcal{L}_{g'}^{(GMV)}(H(g(\boldsymbol{X}_{ip}), f_u(\boldsymbol{X}_u)), y^{(GMV)}) + \lambda_1 \cdot \mathcal{L}_{g'}^{(VpPV)}(H(f_i(\boldsymbol{X}_i), f_u(\boldsymbol{X}_u)), y^{(VpPV)}) + \lambda_2 \mathcal{L}_s(S(G(\boldsymbol{X}_{ip}), f_i(\boldsymbol{X}_i)))$;

$\boldsymbol{X}_g \leftarrow G(\boldsymbol{X}_{ip})$;

end

Algorithm 2: Extended ATNN Algorithm for New Restaurant's Popularity Prediction in Food Delivery

TABLE IV
RESULTS OF OFFLINE EXPERIMENTS FOR FOOD DELIVERY

Model	VpPV (MAE)	GMV (MAE)
TNN-DCN	0.077	1.445
ATNN	0.069	1.206
Improvement	10.4%	16.5%

1) Offline experiments: We utilize a dataset from "Ele.me" containing 1.2 million new sign-up restaurants. VpPV and GMV of the first 30-day upon sign-up are regarded as labels. We apportion the data into training and test sets with an 80-20 split. The features of restaurants include brands, locations, themes, cuisines, the number of similar restaurants nearby, overall VpPV, overall GMV, overall CTR, etc. The number of dimensions of the restaurants after data preprocessing is 211. λ_1 and λ_2 are set as 100 and 10. The network structure is the same as the previous ATNN model mentioned in Section IV. In the O2O scenario, there are more categorical features reducing the performance of GBDT. Previous experiments also show that GBDT and TNN-FC are significantly weaker than TNN-DCN. Therefore, we regard TNN-DCN as the compared algorithm and Mean Absolute Error (MAE) of VpPV and GMV as the evaluation metrics.

Table IV shows the results of offline experiments for food delivery. Compared with TNN-DCN, our ATNN model reduces the MAE of VpPV from 0.077 to 0.069 by 10.4% and the MAE of GMV from 1.445 to 1.206 by 16.5%. Introducing the adversarial network improves the accuracy of VpPV and GMV predictions for new restaurants significantly.

2) Online experiments: In traditional food delivery platform, we hire human experts to select competitive merchants for sign-up. Benefiting from the accurate forecasts by ATNN, we can filter tens of millions of new applications automatically. During the online experiment, we compare ATNN to human experts. Both ATNN and human experts select attractive and welcoming restaurants for sign-up in our business. Reviewing the first 30-day performance of newly recruited restaurants in this July, we regard actual VpPV and GMV as evaluation metrics.

TABLE V
RESULTS OF ONLINE EXPERIMENTS FOR FOOD DELIVERY

Source	VpPV	GMV
Human Experts	0.2656	191.23
ATNN	0.2872	219.33
Improvement	8.1%	14.7%

Table V shows the results of online experiments for food delivery. Compared with human experts, our ATNN model increases VpPV from 0.2656 to 0.2872 by 8.1% and GMV from 191.23 to 219.33 by 14.7%. Reviewing this A/B test, our ATNN model outperforms human experts on recruiting new restaurants in creating more business value.

VI. CONCLUSIONS

In this paper, we study a problem in e-commerce about new arrivals prediction, which is widely existed. Because there is no general indicator to evaluate new item popularity, we introduce Click Through Rate (CTR) to indicate item popularity. Existing CTR prediction models fail in new arrivals CTR prediction because of the sparse interactive features and missing item statistics. To tackle this challenge, we propose an Adversarial Two-tower Neural Network (ATNN) framework, which utilizes a novel adversarial component consisting of a generator and a discriminator. The adversarial component learns to better generate item vectors from only item profiles for CTR prediction. Besides, we introduce Deep & Cross Network (DCN) and shared embeddings in the ATNN framework. Experimental results show that ATNN has a powerful ability to generate item vectors from item profiles in CTR predictions. Another challenge in new arrivals prediction lies in huge overhead since pairwise computing between new arrivals and hundreds of million users is computationally costly. To address this problem, we construct a user group and store its mean user vector before the prediction stage. To predict a new item, we can obtain its popularity score from the item vector generated by the item generator using item profiles and the mean user vector stored in the system. This prediction step has only $\mathcal{O}(1)$ time complexity. Extensive offline and online experimental results on a large-scale e-commerce platform, "Tmall.com", validate the effectiveness and efficiency of our ATNN. In addition, by introducing multi-task learning, we extend ATNN to a food delivery platform, "Ele.me", to accurately predict VpPV and GMV of new restaurants.

There are several interesting problems to be investigated in our future work:

- It will be very interesting and valuable to extend our ATNN model to multiple online applications. Cold-start challenge exists widely in various online services using personalized recommendations or personalized information retrievals. ATNN utilizes an adversarial network to generate item vectors from only item profiles. Besides in e-commerce, this strategy can be applied to other scenarios, for example, movie recommendation.

- We can further group users by their preferences before making new arrivals predictions. Different groups have diverse preferences for different types of items. Introducing this information may further improve the accuracy of new arrivals predictions.

Acknowledgment

This work is partly conducted within the Delta-NTU Corporate Lab for Cyber-Physical Systems with funding support from Delta Electronics Inc. and the National Research Foundation (NRF) Singapore under the Corp Lab@University Scheme.

We would like to thank Alibaba Group for valuable datasets and platforms and Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies (AZFT) for technical support.

References

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[2] R. Wang, B. Fu, G. Fu, and M. Wang, "Deep & cross network for ad click predictions," in *Proceedings of the ADKDD'17*, 2017, pp. 1–7.

[3] S. Sedhain, S. Sanner, D. Braziunas, L. Xie, and J. Christensen, "Social collaborative filtering for cold-start recommendations," in *Proceedings of the 8th ACM Conference on Recommender systems*, 2014, pp. 345–348.

[4] A. Krohn-Grimberghe, L. Drumond, C. Freudenthaler, and L. Schmidt-Thieme, "Multi-relational matrix factorization using bayesian personalized ranking for social network data," in *Proceedings of the fifth ACM international conference on Web search and data mining*, 2012, pp. 173–182.

[5] S. Sedhain, A. K. Menon, S. Sanner, L. Xie, and D. Braziunas, "Low-rank linear cold-start recommendation from social data," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

[6] M. Balabanović and Y. Shoham, "Fab: content-based, collaborative recommendation," *Communications of the ACM*, vol. 40, no. 3, pp. 66–72, 1997.

[7] M. Nasery, M. Braunhofer, and F. Ricci, "Recommendations with optimal combination of feature-based and item-based preferences," in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, 2016, pp. 269–273.

[8] M. Saveski and A. Mantrach, "Item cold-start recommendations: learning local collective embeddings," in *Proceedings of the 8th ACM Conference on Recommender systems*, 2014, pp. 89–96.

[9] Y. Zhu, J. Lin, S. He, B. Wang, Z. Guan, H. Liu, and D. Cai, "Addressing the item cold-start problem by attribute-driven active learning," *IEEE Transactions on Knowledge and Data Engineering*, 2019.

[10] M. Aharon, O. Anava, N. Avigdor-Elgrabli, D. Drachsler-Cohen, S. Golan, and O. Somekh, "Excuseme: Asking users to help in item cold-start recommendations," in *Proceedings of the 9th ACM Conference on Recommender Systems*, 2015, pp. 83–90.

[11] M. Richardson, E. Dominowska, and R. Ragno, "Predicting clicks: estimating the click-through rate for new ads," in *Proceedings of the 16th international conference on World Wide Web*, 2007, pp. 521–530.

[12] 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 *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2013, pp. 1222–1230.

[13] K. Gai, X. Zhu, H. Li, K. Liu, and Z. Wang, "Learning piece-wise linear models from large scale data for ad click prediction," *arXiv preprint arXiv:1704.05194*, 2017.

[14] S. Rendle, "Factorization machines," in *2010 IEEE International Conference on Data Mining*. IEEE, 2010, pp. 995–1000.

[15] X. Yi, J. Yang, L. Hong, D. Z. Cheng, L. Heldt, A. Kumthekar, Z. Zhao, L. Wei, and E. Chi, "Sampling-bias-corrected neural modeling for large corpus item recommendations," in *Proceedings of the 13th ACM Conference on Recommender Systems*, 2019, pp. 269–277.

[16] M. Naumov, D. Mudigere, H.-J. M. Shi, J. Huang, N. Sundaraman, J. Park, X. Wang, U. Gupta, C.-J. Wu, A. G. Azzolini *et al.*, "Deep learning recommendation model for personalization and recommendation systems," *arXiv preprint arXiv:1906.00091*, 2019.

[17] H.-T. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir *et al.*, "Wide & deep learning for recommender systems," in *Proceedings of the 1st workshop on deep learning for recommender systems*, 2016, pp. 7–10.

[18] X. He and T.-S. Chua, "Neural factorization machines for sparse predictive analytics," in *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2017, pp. 355–364.

[19] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, "Deepfm: a factorization-machine based neural network for ctr prediction," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, 2017, pp. 1725–1731.

[20] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for youtube recommendations," in *Proceedings of the 10th ACM conference on recommender systems*, 2016, pp. 191–198.

[21] Z. Zhao, L. Hong, L. Wei, J. Chen, A. Nath, S. Andrews, A. Kumthekar, M. Sathiamoorthy, X. Yi, and E. Chi, "Recommending what video to watch next: a multitask ranking system," in *Proceedings of the 13th ACM Conference on Recommender Systems*, 2019, pp. 43–51.

[22] G. Zhou, X. Zhu, C. Song, Y. Fan, H. Zhu, X. Ma, Y. Yan, J. Jin, H. Li, and K. Gai, "Deep interest network for click-through rate prediction," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 1059–1068.

[23] G. Zhou, N. Mou, Y. Fan, Q. Pi, W. Bian, C. Zhou, X. Zhu, and K. Gai, "Deep interest evolution network for click-through rate prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 5941–5948.

[24] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. N. Metaxas, "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 5907–5915.

[25] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[26] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[27] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[28] Z. Zhang, Y. Song, and H. Qi, "Age progression/regression by conditional adversarial autoencoder," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5810–5818.

[29] S. Tulyakov, M.-Y. Liu, X. Yang, and J. Kautz, "Mocogan: Decomposing motion and content for video generation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1526–1535.

[30] J. Wang, L. Yu, W. Zhang, Y. Gong, Y. Xu, B. Wang, P. Zhang, and D. Zhang, "Irgan: A minimax game for unifying generative and discriminative information retrieval models," in *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2017, pp. 515–524.

[31] X. He, Z. He, X. Du, and T.-S. Chua, "Adversarial personalized ranking for recommendation," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018, pp. 355–364.

[32] B. Wang, M. Qiu, X. Wang, Y. Li, Y. Gong, X. Zeng, J. Huang, B. Zheng, D. Cai, and J. Zhou, "A minimax game for instance based selective transfer learning," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 34–43.

[33] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.