AntProphet: an Intention Mining System behind Alipay's Intelligent Customer Service Bot

Cen Chen , **Xiaolu Zhang** , **Sheng Ju** , **Chilin Fu** , **Caizhi Tang** , **Jun Zhou** and **Xiaolong Li**

Ant Financial Services Group

{chencen.cc; yueyin.zxl; jusheng.js; chilin.fcl; caizhi.tcz; jun.zhoujun; xl.li}@antfin.com

Abstract

We create an intention mining system, named AntProphet, for Alipay[1]'s intelligent customer service bot, to alleviate the burden of customer service. Whenever users have any questions, AntProphet is the first stop to help users to answer their questions. Our system gathers users' profile and their historical behavioral trajectories, together with contextual information to predict users' intention, i.e., the potential questions that users want to resolve. AntProphet takes care of more than 90% of the customer service demands in the Alipay APP and resolves most of the users' problems on the spot, thus significantly reduces the burden of manpower. With the help of it, the overall satisfaction rate of our customer service bot exceeds 85%.

1 Introduction

Customer service is one of the pillars for a company's success, as it is highly related to customers' satisfaction and affects how the company is viewed by the public. Traditional customer service, which is made through either manual phone calls or online chats, is resource-intensive and time-consuming. And the traditional ways of services sometimes are hard to reach due to the availability issue, especially during peak periods. [Koole, 2004] reported that customers' patience for waiting usually only lasts for 1 minute and consequently may cause abandoned calls. Such common issues will lead to deteriorated or unsatisfied service experiences. Thus, it is critical for a company to design an intelligent customer service bot to help customers in need.

Witnessing the advances in data mining and deep learning, many intelligent customer service bots are introduced in online platforms/APPs to revolutionize the customer service, such as Microsoft SuperAgent [Cui *et al.*, 2017], Amazon Alexa [2], and AliMe [Li *et al.*, 2017]. Such bots have become extremely important to save both labor and time costs, where common and repetitive questions can be automatically addressed 24/7. However, often times users may be unable to

generate a precise question or correctly describe the complicated situations encountered. Or sometimes users themselves are not sure what to ask. In such cases, customers most likely are still going to find a customer service agent and explain the situation explicitly. What we do differently is that we are seeking to understand user questions/intent without their explanation. One observation is that we find a large number of user questions can be inferred from their historical behaviors. For example, a user who is facing payment issues on a platform may have tried to make the payment for a few times.

In this paper, we propose an intention mining system, named AntProphet, to facilitate our service bot by mining user's intention based on his/her profile and historical behaviors and providing help (in forms of a list of recommended questions/answers or easily accessible tools) even before they ask anything. Whenever users meet any problem, they will be first routed to AntProphet as their primary choice of consultation once entering the service bot. As a result, our system takes care of more than 90% of the customer service demands in Alipay, i.e., the majority of user questions are automatically handled by our system.
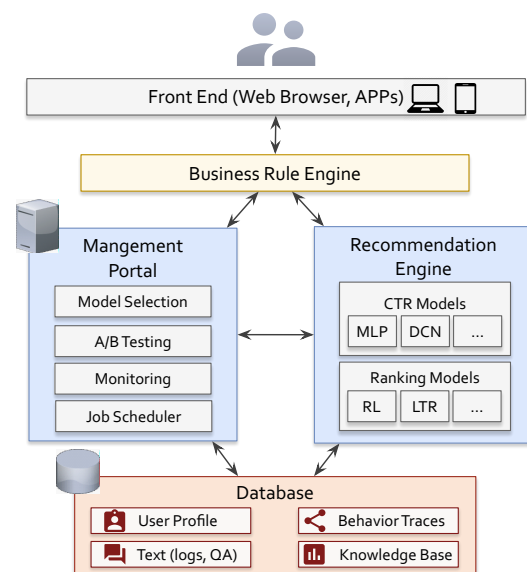

Figure 1: System Architecture Overview.

---

[1]The leading third-party mobile payment APP in China.

[2]https://developer.amazon.com/alexa

Figure 2: An Illustration of APP Interface.

2 System Architecture

The overall architecture of AntProphet is illustrated in Figure 1. The front end interacts with the business rule engine whenever a user requests for customer service. Subsequently, the business rule engine requests the recommendation engine for a list of recommended items(i.e., the questions that a user faced or a list of recommended tools for easy accessing), combining with the business rules to produce the final recommendations for users. More specifically, the recommendation engine is responsible for model training, interacting with the database for data processing and making real-time recommendations. Meanwhile, the management portal supports a variety of management tasks across the whole pipeline, such as model selection, testing, monitoring, and scheduling.

2.1 Front End

In Figure 2, we give an illustration of AntProphet user interface. AntProphet entries are embedded in the Alipay APP. Once a user clicks any customer service button in the APP, he/she will be immediately routed to AntProphet's main page, which consists of three main gadgets:

- Recommended questions in Figure 2(a) shows a list of most likely questions that relate to the problems a user may face. Answer details will be popped up when the user clicks questions that he/she is interested in.

- Recommended tools in Figure 2(b) presents a list of tools that highly likely to help the user to resolve the problems faced.

- Interactive recommendation gadget in Figure 2(c) provides a gamified question recommendation interface for a user to interact, where questions are recommended considering user clicked intent tags on a real-time basis.

Whenever users enter any customer service page, real-time requests are sent to the engine for recommendations.

2.2 Question Recommendation

The main component for our system is the recommendation engine. To perform top-N recommendation, two types of models are supported by the engine: Click-Through-Rate (CTR) prediction models and ranking models, which are periodically trained and updated in the engine for requesting.

Feature Extractions

Three types of features are considered: user profile (e.g., user's age, gender and etc.), user historical behavior traces (such as click, purchase and etc.), question-related information (e.g., texts, question-level statistics and etc.). Note that the latter two types of features are collected at different time granularities. We find user's historical behavior traces plays the most important role in predicting intention. This is intuitive as the problems users face in the APP are usually strongly associated with their recent behaviors.

Base CTR Models

CTR models aim to predict the likelihood of an item/question being clicked given the user query and related features. Based on the abovementioned features, we use CTR models to recommend questions in the knowledge base. This task is a large-scale problem in the industry, due to the large data size and feature dimension after discretization. Neural Network (NN) based models are chosen due to their efficiency in learning high-degree feature interactions without feature handcrafting. More specifically, for textual features, we use CNN [Kim, 2014] to get a feature representation. For non-text data such as user profile, we use a fully connected layer to learn a feature representation, while sequential behavior traces are embedded with LSTM. Three types of feature representations are concatenated to feed into two fully connected layers to make final predictions. Other NN variants, such as Deep&Cross [Wang *et al.*, 2017] and DeepFM [Guo *et al.*, 2017], are implemented to better model feature interactions.

Question Reranking

We then use ranking based models to perform the reranking of questions to optimize the presentation of questions, based on three types of ranking signals: CTR scores, question popularity, and question diversity. We build a set of base CTR models beforehand and use their model results as our CTR score signals. Aside from Learning-To-Rank, we introduce Reinforcement Learning(RL) based model for our top-N recommendation task. We propose to view the problem as a sequential decision-making process, where each step is to find a new question to add to the already selected questions, to better capture the long-term effects of each recommendation in the list and optimize a mixture of long-term accumulated rewards, such as click and diversity. We adapt Deep Deterministic Policy Gradients [Lillicrap *et al.*, 2015] for our RL model. Details of our proposed RL based approach can be found in [Chen *et al.*, 2019], omitted due to the limited space.

In all, AntProphet is designed to be a closed-loop platform that supports automatic data collection, feature extraction, model training, deployment, and management pipeline, thus making it easy to be applied for any new scenarios.

3 Conclusion

In this work, we present AntProphet, an intention mining system for Alipay's intelligent customer bot. We show the whole system and demonstrate how it works. Interested readers can try the Alipay App directly to test the system. Meanwhile, we have recorded a video to demonstrate our system.

References

[Chen *et al.*, 2019] Cen Chen, Chilin Fu, Xu Hu, Xiaolu Zhang, Jun Zhou, Xiaolong Li, and Forrest Sheng Bao. Reinforcement learning for user intent prediction in customer service bots. In *42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2019.

[Cui *et al.*, 2017] Lei Cui, Furu Wei, Shaohan Huang, Chuanqi Tan, Chaoqun Duan, and Ming Zhou. Superagent: A customer service chatbot for e-commerce websites. In *ACL 2017, System Demo*, pages 97–102, 2017.

[Guo *et al.*, 2017] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. Deepfm: a factorization-machine based neural network for ctr prediction. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 1725–1731. AAAI Press, 2017.

[Kim, 2014] Yoon Kim. Convolutional neural networks for sentence classification. In *EMNLP*, pages 1746–1751, 2014.

[Koole, 2004] Ger Koole. Performance analysis and optimization in customer contact centers. In *Quantitative Evaluation of Systems, 2004. QEST 2004. Proceedings. First International Conference on the*, pages 2–5. IEEE, 2004.

[Li *et al.*, 2017] Feng-Lin Li, Minghui Qiu, Haiqing Chen, and et al. AliMe Assist : An Intelligent Assistant for Creating an Innovative E-commerce Experience. In *CIKM*, 2017.

[Lillicrap *et al.*, 2015] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *CoRR*, abs/1509.02971, 2015.

[Wang *et al.*, 2017] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. Deep & cross network for ad click predictions. In *Proceedings of the ADKDD'17*, page 12. ACM, 2017.