A Real-World Human-Machine Interaction Platform in Insurance Industry

Wei Tan, Chia-Hao Chang[†], Yang Mo, Lian-Xin Jiang, Gen Li, Xiao-Long Hou
Chu Chen, Yu-Sheng Huang, Meng-Yuan Huang, Jian-Ping Shen
AI Department
Ping An Life Insurance of China, Ltd.
{tanwei818,zhangjiahao206[†],moyang853,jianglianxin769,ligen947,houxiaolong430,
chenchu870,huangyusheng112,huangmengyuan334,shenjianping324}@pingan.com.cn
Correspondence[†]:strategist922@gmail.com

Abstract

In the insurance industry, lots of effort is putting into helping the customer to solve their problems that occurred during and after purchasing cycle and helping telemarketers to practice selling skills. Chat bots and assistant bots are widely used in these business scenarios, but building a bot application from scratch is expensive. In this paper, a human-machine interaction platform specially designed for intelligent bot applications in insurance industry that combined the technologies of Question Answering (QA), task-oriented dialogue and chit-chat was proposed and we demonstrate the architecture design of this platform, key technologies and the scenario of applications in real-world insurance industry. It has been supporting many intelligent bot applications of insurance industry already, such as Intelligent Coach Bot (ICB) which helps telemarketers to practice their selling skills, Intelligent Customer Service Bot (ICSB) which provides after-sales services and Insurance Advisor Bot (IAB) which helps customer to purchase the most suitable insurance product. Currently, these bot applications serve millions of users per day and are able to solve 80% of the online problems.

KEYWORDS

Human-Machine Interaction Schema, Dialogue Tree, Text Similarity, Text Matching, Chit Chat Engine, Anaphora Resolution, Intention/Slot Driven, Attention Mechanism

1   INTRODUCTION

All authors contributed equally to this manuscript.

With the development of deep learning technology, speech recognition and natural language understanding (NLU) have developed rapidly and the way of human-computer interaction has also changed from GUI to CUI (Conversation User Interface). The traditional GUI depends on the keyboard, mouse, or touch screen. But CUI carries out input and output through dialogue, breaking through the limitation of contact and getting closer to the way of communication between people. CUI is a tremendous change in the way of interaction with a variety of AI techniques. We have implemented a human-computer interaction platform that uses CUI as a starting point to construct a conversational input and output interface and a smart brain. Based on the insurance scenario, our platform has been widely and solidly applied in the fields of salesman training, intelligent customer service, and recommendation of products.

In salesman training scenario, we rely on this platform to build intelligent coach robot that can simulate customers to do human-machine dialogue with the telesales man. It helps them learn sales and handle customer objections, improving their sales skills. In the intelligent customer service scenario, we built a customer service robot under the technical support of the platform which can answer users' questions at anytime and anywhere and guide users to apply the request for business, such as the policy loan. In the scenario of insurance products recommendation, we implemented the insurance advisor bot that could find customers' demands from the human-computer dialogue and recommend insurance products that our customers might need. All of these bots were built on the human-computer interaction technology platform. They share basic technologies provided by the platform, such as intent recognition, semantic understanding, dialogue management, and a unified knowledge center. The main contributions of this paper include: (1) We design a set of human-computer interaction application solutions and based on these to develop the coach bot, the intelligent customer service bot and intelligent insurance advisor bot. (2) We propose the human-machine interaction schema which uses dialogue tree of intentional organization as the core. (3) We propose and implement a calculation method of text similarity that integrated various of technical methods which used for text matching. (4) We propose and implement the intent detection, recognition of chit chat based on the classification technology. (5) We propose and implement a recommendation technical solution based on dialogue.
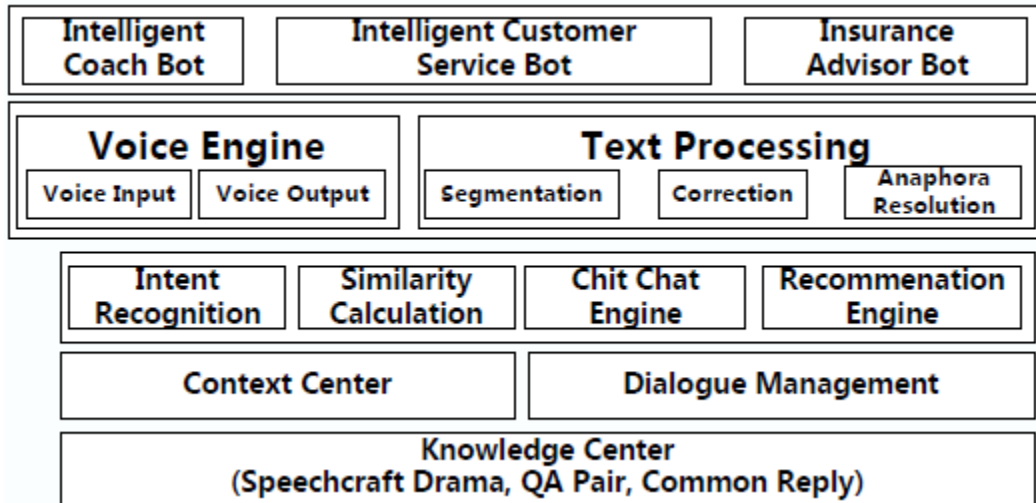
2　SYSTEM OVERVIEW

Figure 1: The Overall System Architecture

In Figure 1, it shows the overall architecture of the human-computer interaction technology platform. The first layer is the input and output layer. The system supports the interactive form of speech and text. The input speech needs to be recognized by the ASR engine, and the text needs to be synthesized into the voice output by the TTS engine. Text processing module will pre-process the text input or the text comes from the ASR engine. The pre-processing step includes word segmentation, error correction and anaphora resolution. Using uniform pre-processing module can ensure the consistency of the results. By tuning this module in the system, it will be applied to all downstream modules. The second layer is the basic technical support layer. It includes intent detection, text similarity calculation, the recommendation engine and chit chat engine. The third layer is the context center and dialogue management module used to track the whole human-machine dialogue and manage slot information. The fourth layer is the knowledge center, which includes the speech craft drama, QA pairs, common replies and so forth. The entire human-computer interaction technology platform can support a variety of business scenarios, such as intelligent coaching, intelligent customer service, and insurance products recommendation, which is demonstrated on top of Figure 1. The three bots share basic components of the platform, like the Voice Engine and Text Processing modules. The Intelligent Coach Bot uses Intent Recognition and Similarity Calculation modules for NLU (Natural Language Understanding), and Chit Chat Engine is used to generate reasonable replies when a user has said some task-unrelatable sentences. For Intelligent Customer Service Bot, QA is especially important, so it integrates Similarity

Calculation including semantic similarity and literal similarity as its sub-module. Both Intelligent Coach Bot and Intelligent customer Service Bot use Content Center and Dialogue Management for context understanding and dialogue state tracking [14]. As for Insurance Advisor Bot, Recommendation Engine and Content Center is the core module.

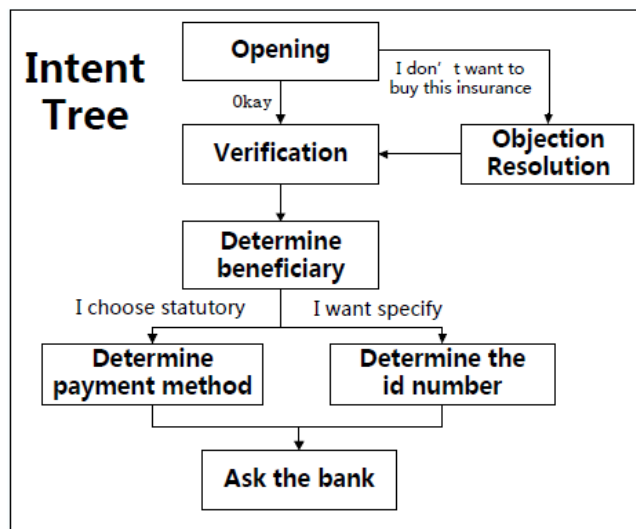## 3 SYSTEM FEATURES

3.1 Dialogue Management

Figure 2: Dialogue Management

In the man-machine dialogue, we need to identify the user's intention or extract the slot value from the conversations. The basic framework for our implementation shows in Figure 2. In salesman training scenario, the system was configured with a lot of dialog dramas in advance. The purpose is to train salesman to improve sales skills by using the system to simulate as a customer and perform human-computer dialogue. The salesman needs to sell the insurance product to the customers according to the pre-defined intention of the script. Each intention has a standard speech and allows the salesman to have a certain degree of free play, but cannot deviate from the main meaning. The customer a machine simulated will give an ordinary response or an objection during the selling process. For example, "I don't want to buy your insurance" is the customer's objection, and the salesman needs to resolve the objection before entering the next

round of dialogue. In order to give maximum flexibility for agents, we organize drama by using intent dialog tree, one drama corresponding to an insurance product is organized as one dialog tree. We use the quadruples *(intent_name, objection, finished_intents, prattle_times)* to save the information of dialogue status and store in the Redis cache system. In a conversation, *intent_name* is used to track the salesman's intent node's position in the last round of dialogue tree, *objection* indicates whether the machine raises the objection or not in the last dialogue round, *finished_intents* is used to track the intents of the entire conversation. Owing to the limitation on the number of consecutive chats, *prattle_times* is used to record the number of chats.

3.2 Intent Detection

In the human-machine dialogue process, the system needs to identify the intention of the agent or the intention in the customer's dialogue. This is a text classification task [1], we use the FastText tool. An important reason why we chose FastText is its fast inference speed so that we can make a robot dialogue system has a good interactive experience. When training, we try to join the n-gram feature, after the addition of this feature the size of the model will reach nearly 1GB, while F1 score of our model is only slightly improved by 0.3%. In our use case, we do training for each dialog drama with an individual model. In order to make memory consumption at a reasonable level, we did not use n-gram feature choice in the final model.

Model	FastText	FastText+2-gram	BERT
Avg. of F1	0.9664	0.9694	0.97245
Avg. Time of Prediction	48.46 μs / sentence	59.38 μs / sentence	213333 μs / sentence
Model Size	8.9M	986M	1.1G

Table 1: Benchmark Results

In Table 1, it shows the performance benchmark results with the FastText model, FastText+2-gram features model and BERT model [13], by using the intent classification dataset. The dataset contains 36 categories, and each category contains 1536 sentences. This dataset corresponds to a

drama, and every category is an intent like "opening", "verification". The corpora are collected by many sophisticated salesmen, they divide the insurance sale into several key processes and each of them corresponds to an intent, e.g. the intent "verification" means verify the identity of a customer. We use 80% of the dataset for training and other 20% of the dataset as the test set. We get one F1 score for every category and compute the average F1 over all categories as show in Table 1.

3.3 Semantic Calculation

In the intent-based dialogue, customers will raise objections or questions. The questions need to be matched with the corpus in the knowledge base. The system will match the customer's objections or questions with the collection set of customer objections in the knowledge base or in our FAQ set one by one. In our system, we used a combination of literal similarity, and Siamese CBOW [2].

3.3.1 Literal Similarity Algorithm

Based on the literal similarity of edit distance, the Levenshtein Distance algorithm is used to calculate the minimum number of steps to transform sentence A into sentence B by adding, deleting, and replacing operations. After finding the edit distance of Levenshtein, the similarity of the two sentences is calculated by the following formula: ***Similarity = (Max(x, y) - Levenshtein) / Max(x, y)***, where *x*, *y* are the length of the sentence *A* and *B*.
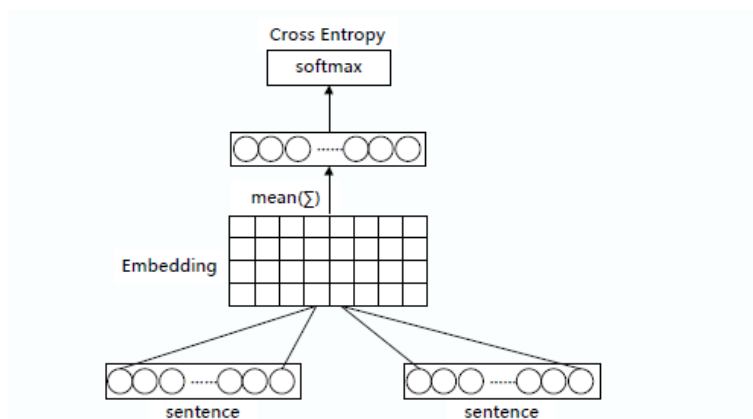

Figure 3: CBOW Optimization

3.3.2 CBOW Optimization

We have optimized the semantic representation of the word vector and adjusted the word vector at the sentence level. The model structure we use was shown in Figure 3. For each sentence, we select a positive case and several negative cases, and calculate the sentence vector according to the formula ***SentenceVec(A)=1/n∑Vec(wi)***. The cosine similarity of the selected sentence and the positive and negative examples is then calculated and normalized using the SoftMax function and uses cross entropy as the loss function. The embedding layer uses a pre-trained Glove word vector and adjusts the word vector weight by adding the sentence similarity as considerations.

3.4 Chit Chat Recognition and Reply

We pre-defined some common Chit Chat response categories and answer templates. The template defines some frequent Chit Chats. For a Chit Chat conversation that cannot be classified as a pre-defined Chit Chat category, we generate a chat response by using the seq2seq model [4] [15] [3].

3.5 Dialog-Based Recommendation

We proposed a dialogue recommendation model show as in Figure 4 to effectively recommend insurance products to users. The dialogue model includes three sub-networks: the graph recommendation sub-network, the semantic recommendation sub-network, the behavior recommendation sub-network and the output layer is a weighted sum of the scores of networks.

The input of the graph recommendation sub-network is the entity embedding in the query. The intelligent insurance advisor bot has established a knowledge graph for insurance approbation and insurance indemnity, with about 7,000 nodes. Based on the knowledge graph, system can learn the entity embedding which including insurance products, diseases, and people. The graph recommendation sub-network employs the attention mechanism. The point multiplication is used to expose the similarity between the insurance product and the entity, which is efficient and performs well. The input of the semantic recommendation sub-network is the word embedding in the query. The word embedding is pre-trained. The insurance, disease and other professional terms are very helpful for the model. Since the conversations are generally short texts, using the CNN

model is more efficient, and the semantic recommendation sub-network uses one-layer convolution-pooling CNN. The input of the behavior recommendation sub-network is the item embedding of the insurance products in the query. Using the user click behavior data in intelligent insurance advisor bot mobile application, item embedding can be learned by item2vector model [5]. The behavior recommendation sub-network employs the attention mechanism and the attention value is calculated by using point multiplication between clicked item and target item. Increasing the behavior recommendation sub-network AUC can increase by 1.2%.
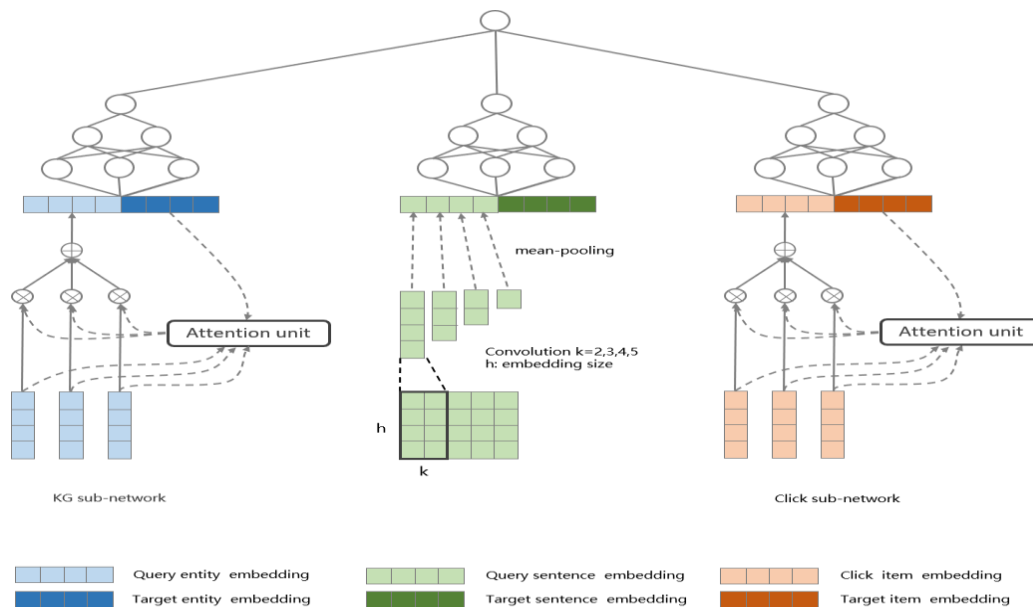

Figure 4: Dialog-Based Recommendation

4 RELATED WORK

Dialog Management. This is a core module of human-computer interaction system, which solving the problem of dialogue state tracking [6] and dialogue logic processing. Here we use the dialog tree to manage user intent and state.

Intent Recognition. In the human-machine dialogue system, intent recognition is crucial. Traditional intent recognition methods typically use a template-based approach. With the development of neural networks, intent recognition based on CNN[7] and RNN[8] methods have emerged. We used the FastText method achieving an accuracy of 86%.

Semantic Computation. It is possible to compute the semantic representation of the query and then retrieve the knowledge base through the semantic matching technology to obtain the best matching question and the corresponding answer. The popular methods of semantic matching include literal matching and semantic matching (semantic matching models include DSSM[9], CLSM[10], Deep Math[11], Siamese). In order to improve the coverage of matching, we use the literal similarity and Siamese CBOW model [2] and set the corresponding weight for each method's results.

Chit chatting. Commonly used methods for open domain chit chatting include IR-based [12] and generation models. We use the IR-based methods on the priority, and use the generative methods when the answer can't be found with IR-based methods, as insurance field is financial related, we need to give a rigorous answer.

Dialogue Based Recommendations. In the human-computer interaction scenario, we recommend products, tasks, and news to users based on user behavior and conversation content to improve user experience and conversion rate in human-machine applications. We constructed a dialogue recommendation model consisting of three sub-networks (semantic recommendation sub-network, map recommendation sub-network, behavior recommendation sub-network).

5 CONCLUSION

The human-machine interaction platform proposed in this paper has been supporting many intelligent bot applications of insurance industry already, such as Intelligent Coach Bot (ICB) which helps telemarketers to practice their selling skills, Intelligent Customer Service Bot (ICSB) which provides after-sales services, and Insurance Advisor Bot (IAB) which helps customer to purchase the most suitable insurance product. Currently, these bot applications serve millions of users per day and are able to solve 80% of the online problems. In the future, the improvements of the platform include intent recognition and semantic computing based on BERT, end-to-end tasks modeling, scene-based sequence labeling and shopping guide based on reinforcement learning.

REFERENCES

[1] Sun, Y., Wong, A. K., & Kamel, M. S. (2009). CLASSIFICATION OF IMBALANCED DATA: A REVIEW. International Journal of Pattern Recognition and Artificial Intelligence, 23(04), 687-719.

[2] Tom.Kenter, et al, 2016, Siamese CBOW: Optimizing Word Embeddings for Sentence Representations

[3] Luong, T., Pham, H., & Manning, C. D. (2015). Effective Approaches to Attention based Neural Machine Translation. empirical methods in natural language processing,, 1412-1421.

[4] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. neural information processing systems,, 3104-3112.

[5] Barkan, O., & Koenigstein, N. (2016). ITEM2VEC: Neural item embedding for collaborative filtering. international workshop on machine learning for signal processing,, 1-6.

[6] Mahew Henderson. 2015. Machine Learning for Dialog State Tracking: A Review. In MLSLP'15.

[7] Hashemi H B., Asiaee A, Kraft R, Query intent detection using convolutional neural networks, International Conference on Web Search and Data Mining, Workshop on Query Understanding , 2016.

[8] Bhargava A, Celikyilmaz A, Hakkanitur D, et al. Easy Contextual Intent Prediction and Slot Detection, IEEE International Conference on Acoustics. IEEE, 2013: 8337-8341.

[9] Po-Sen Huang, et al., 2013, Learning Deep Structured Semantic Models for Web Search using Clickthrough Data

[10] Yelong Shen, et al, 2014, A Latent Semantic Model with Convolutional-Pooling Structure for Information Retrieval

[11] Zhengdong Lu & Hang Li, 2013, A Deep Architecture for Matching Short Texts

[12] Zhao Yan, Nan Duan, Jun-Wei Bao, Peng Chen, Ming Zhou, Zhoujun Li, and Jianshe Zhou. 2016. DocChat: An Information Retrieval Approach for Chatbot Engines Using Unstructured Documents. ACL'16

[13] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[14] Julien Perez, Fei Liu. 2013. Dialog state tracking, a machine reading approach using Memory Network. arXiv preprint arXiv:1606.04052

[15] Jiatao Gu, Zhengdong Lu, Hang Li, Victor O.K. Li. 2016. Incorporating Copying Mechanism in Sequence-to-Sequence Learning. arXiv preprint arXiv:1603.06393