SmartSales: An AI-Powered Telemarketing Coaching System in FinTech

Yuanfeng Song1,2, Xuefang Zhao2, Di Jiang2, Xiaoling Huang2, Weiwei Zhao2, Qian Xu2, Raymond Chi-Wing Wong1, Qiang Yang1,2

1The Hong Kong University of Science and Technology 2Al Group, WeBank Co., Ltd, China
{songyf,raywong,qyang}@cse.ust.hk

ABSTRACT
Telemarketing is a primary and mature method for enterprises to solicit prospective customers to buy products or services. However, training telesales representatives is always a pain point for enterprises since it is usually conducted manually and costs great effort and time. In this demonstration, we propose a telemarketing coaching system named SmartSales to help enterprises develop better salespeople. Powered by artificial intelligence (AI), SmartSales aims to accumulate the experienced sales pitch from customer-sales dialogues and use it to coach junior salespersons. To the best of our knowledge, this is the first practice of an AI telemarketing coaching system in the domain of Chinese FinTech in the literature. SmartSales has been successfully deployed in the WeBank’s telemarketing team. We expect that SmartSales will inspire more research on AI assistant systems.

CCS CONCEPTS
• Information systems → Expert systems;

KEYWORDS
AI-Powered coaching system, Telemarketing, FinTech

ACM Reference Format:

1 INTRODUCTION
Recent years have witnessed a huge rise in artificial intelligence (AI) coaching system in the sports [1, 10] and education [8] area. However, when it comes to telemarketing, the primary method for FinTech enterprises to solicit prospective customers to buy products, the telesales representatives training procedure is still a pain point since it is usually conducted manually and costs great effort. The reason is that the experience of senior salespersons is hard to accumulate.

2 SYSTEM OVERVIEW
SmartSales includes various critical components, and in this section, we will briefly introduce their details.

QA Mining & KnowledgeBase Construction As shown in Fig. 2, automated QA mining includes two steps: question identification and answer detection. For the question identification step, a classifier is trained to distinguish meaningful questions from the non-question from the dialogues, and then the questions are further filtered to ensure smoothness and quality. Candidate answers are prepared for the selected answers, and we consider two types of similarity features to detect the matched one, including lexical (e.g., the ratio of overlapping words and TF-IDF) and semantic (e.g., sentence embedding and topic distribution) [2]. Besides dialogue history, SmartSales also crawls knowledge from other sources such as financial websites and WeChat official accounts into its knowledge base.
Figure 2: The QA Mining and Recommendation

ASR Given the real-time acoustic input, the ASR module decodes its corresponding transcripts as follows:

$$w^* = \arg \max_w \log P_{LM}(w) + \lambda \log P_{AM}(a|w),$$

where $P_{LM}$ is a language model (LM), $P_{AM}$ is an acoustic model (AM), and $\lambda$ is a trade-off parameter. In SmartSales, we build the AM using the Kaldi "Chain" model [6, 7]. To further improve the performance, we also use an advanced L2RS mechanism [9] integrated with BERT [3] to do the N-best rescoring step.

User Intent Mining We manually identify 35 intents which belong to 4 categories, and then the intent mining problem is transformed into a classification one. We extract word level and character level features using Word2Vec [5], then feed them into a convolution neural network (CNN) to get the high-level representation $H_w^c$ and $H_c^i$, and finally train a classifier represented as:

$$y = \text{Softmax}([H_w^c; H_c^i]W^T),$$

where $y$ is the intent and $W^T$ is trainable weight matrices.

Response Recommendation As shown in Fig. 2, the response recommendation models employ a two-step approach, that is, retrieval and ranking. The essential words are extracted from the client’s dialogue, as query $q$, to retrieve the top-$n$ candidates, and then various features for candidates are obtained from pre-trained models. Eq. 3 is used to measure the similarity $S_i$ of the $i$-th feature between query and candidates, and a pairwise learning-to-rank model [4] is used to rerank these candidates. The overall correlation is calculated by a feedforward neural network using Eq. 4, and the model is trained with the loss function defined as Eq. 5. Finally, the top 3 candidates are displayed to the end-users.

$$S_i(q, a) = \cos(q, a) = \frac{\|q \cdot a\|}{\|q\| \cdot \|a\|}$$

$$\text{Sim}(q, a) = \sigma \left( \sum_i w_i S_i(q, a) \right),$$

$$\text{loss} = \max(0, 1 - \text{Sim}(q, a^*) + \text{Sim}(q, a^-)),$$

where $a^*$ and $a^-$ are two response candidates, and $a^*$ is more relevant than $a^-$ with respect to query $q$.

Automated Call Evaluation Automated call evaluation uses algorithms to analyze phone conversations. Insights gained can help identify the best sales leaders and enhance their effectiveness by timely feedback. SmartSales supports profanity detection and positive/negative comment detection with a coarse-to-fine retrieval and matching approach. It also supports call quality scoring.

Performance Analysis We build a corpus constructed from the real dialogue between telesales and customers in the Chinese Fin-tech domain. The original data came from 3-months WeChat official account logs and speech data with around 15K recordings. We first obtained the transcripts by the ASR module and then recruited two workers to annotate the intent and question-answer pair label. After

Figure 3: The System Performance of SmartSales

3 SYSTEM DEMONSTRATION

The functionalities are demonstrated with a screenshot in Fig. 4.

Real-time ASR Decoding After the call is made, the real-time decoding result by an ASR system will be displayed on the homepage.

User Intent Mining & Response Recommendation The ASR decoding results will be forwarded to the user intent mining and response recommendation modules, and the analysis results will be displayed on the webpage.

Call Evaluation & Scoring Typical functionalities include the sentiment analysis results of both the clients and the telesales.

Knowledge Base Maintenance The salesman can maintain a local knowledge base, which contains personalized question-answer pairs. This information will also be used to update the machine learning algorithms regularly.

User Management Different users have different management interfaces, depending on their role. The system administrators can create accounts for an institution and assign local administrators. The local administrators can access the user’s page in the administration console to view, create, and modify users. Sales users can modify their user profiles.

Data Analysis Report This functionality aims to summarize the statistics of the salespersons and the quality of calls by these salespersons. The summary report will be shown to the administrator.
REFERENCES