

IIAS: An Intelligent Insurance Assessment System through Online Real-time Conversation Analysis

Mengdi Zhou*, Shuang Peng*, Minghui Yang, Nan Li
Hongbin Wang, Li Qiao, Haitao Mi, Zujie Wen, Teng Xu, Lei Liu
Ant Group, Hangzhou, China
{jacquelyn.zmd, jianfeng.ps, minghui.yhm, ligan.ln, hongbin.whb, li.qiaol,
haitao.mi, zujie.wzj, harvey.xt, leihu.ll}@antgroup.com

Abstract

With the development of Chinese medical insurance industry, the amount of claim cases is growing rapidly. Ultimately, more claims necessarily indicate that the insurance company has to spend much time assessing claims and decides how much compensation the claimant should receive, which is a highly professional process that involves many complex operations. Therefore, the insurance assessor's role is essential. However, for the junior assessor often lacking in practical experience, it is not easy to quickly handle such an online procedure. In order to alleviate assessors' cognitive workload, we propose an Intelligent Insurance Assessment System (IIAS) that helps effectively collect claimant information through online real-time conversation analysis. With the assistance of IIAS, the average time cost of the insurance assessment procedure is reduced from 55 minutes to 35 minutes.

1 Introduction

With the development of Chinese medical insurance industry, the assessor's role becomes essential in handling growing claim cases. The traditional insurance assessment procedure is carried out offline. Due to the coronavirus (COVID-19) pandemic, the previous procedure has to be interrupted temporarily, and thus many claimants can not get their insurance compensation in time. To help these claimants, many insurance companies shift the offline procedure to the online platform in which insurance assessors can communicate with claimants via video connection. In this way, claimants can submit their applications and get compensation entirely online.

However, limited by the complexity of the insurance assessment procedure and the instability of the online environment, simply providing a video connection for insurance assessors and claimants is still insufficient. We summarize three primary problems for the current online procedure: (1) Instability. The online video connection may sometimes be unstable, and the assessor may not distinguish the claimant's

words. (2) Complexity. The insurance assessment procedure is very complex, and not all assessors are sufficiently experienced. The junior assessors, in particular, need much assistance to complete their work expertly, and even some senior assessors still have room for improvement. (3) Time-consuming. Recording the key information and writing the insurance report is tedious for most assessors, and junior assessors may even forget some important information during the inquiry.

In this paper, we propose IIAS: Intelligent Insurance Assessment System, to promote the work efficiency of the insurance assessor by keyword display and report filling during the insurance assessment procedure.

The contributions of this work are as follows.

1. We develop a real-world industrial intelligent system to shift the previous insurance assessment procedure online.
2. We utilize recently advanced NLP technologies [Devlin *et al.*, 2019; Cui *et al.*, 2019; Mani *et al.*, 2020; Ouyang *et al.*, 2020] to address the problem of online insurance assessment. Our methods significantly improve the assessor's work efficiency, where the average time cost of the procedure is reduced from 55 minutes to 35 minutes.
3. We deploy the system in the real world. Until now, it has served thousands of online insurance claim cases, which delivers potential value to the insurance industry.

2 System Overview

Figure 1 demonstrates the architecture of IIAS with key components such as streaming automatic speech recognition (ASR), dialogue segmentation, keyword extraction, keyword filtering, and downstream applications. For better understanding, we discuss the technical details for each key component in the following subsections.

2.1 Dialogue Segmentation

Dialogue segmentation aims at detecting the topic of the real-time conversation text. As the insurance assessment procedure is conducted according to the content that needs to be filled in the insurance report, different parts of the dialogue between the assessor and claimant are focused on different topics. These topics can be seen as different questions that

*Equal contributions.

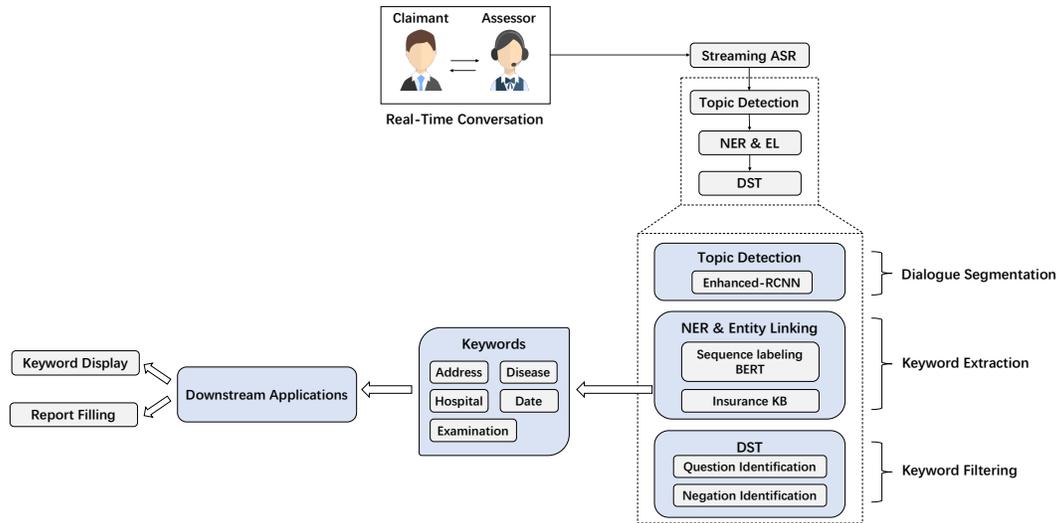


Figure 1: The Architecture of Intelligent Insurance Assessment System.

the claimant needs to answer. Here we use the sentence similarity learning method to map the assessor utterance into the corresponding topic.

We first order the one-to-many mapping relationship between the topic and standard questions that most assessors inquire in the insurance assessment procedure. We select the representative utterance from the assessor as the standard question. Since the ways of inquiry among assessors are different, we use the Enhanced-RCNN model [Peng *et al.*, 2020] to calculate the similarity score between the utterance and selected standard questions, and get the topic of the utterance if the score is greater than the threshold (the default value is 0.5) or follow the topic of the previous utterance if the score is less than the threshold.

After the dialogue segmentation, the whole dialogue is segmented into many parts that belong to different topics¹ so that the keywords extracted in the following process can be linked with these topics in real time. We evaluate the performance of dialogue segmentation by 200 online cases with human annotation. The accuracy of segmentation reaches 90%.

2.2 Keyword Extraction

In the insurance assessment scenario, the keyword contains five types, and their values are not within a limit number. Therefore, any string fragment of assessor or claimant utterance may become a keyword. Traditional slot filling methods with classification models in the task-oriented dialogue system cannot predict unseen slot values in a pre-defined value list and are not appropriate to our scenario. To solve the problem, we convert the keyword extraction as a sequence labeling task and use the BERT-based NER model [Devlin *et al.*, 2019].

We first pre-annotate the keyword type in the utterance as *Addr*, *Hos*, *Dis*, *Date*, *Exam* or *Other* (non-named entity) and follow the official tutorial in NER². For fine-tuning,

¹Every segmented part only belongs to one single topic.

²<https://github.com/google-research/bert>

we feed the final hidden representation for each token into a classification layer over the NER label set [Devlin *et al.*, 2019]. After the NER process, we get the raw entities extracted from the utterances. As the word accuracy of ASR is about 85%, some of the derived entities may be misrecognized. In order to alleviate the negative effect of the ASR module, we use the EL model [Le and Titov, 2018] to recognize and disambiguate the raw entities to the insurance KB that contains the normalized address, hospital, and disease information for getting all possible keywords.

We evaluate the NER/EL model’s performance on the human-annotated testing set. The experimental results show that the average F1-score of the model has reached 82%.

2.3 Keyword Filtering

During the insurance assessment procedure, the assessor always asks the claimant many detailed questions. As a result, a part of entities extracted by the previous keyword extraction process may come from the assessor’s question text and actually do not need to be displayed or suggested if the claimant’s following response has negative semantics. To filter the irrelevant keywords extracted by the previous process and increase the accuracy of keyword display and report filling, we utilize the Dialogue State Tracking (DST) method for keyword filtration. Like the application in the task-oriented dialogue system, the DST method [Ouyang *et al.*, 2020] tracks the state of the keyword (addition or deletion) and contains two major modules: question identification and negation identification. The assessor’s conversion text is first fed to the question identification module. Suppose the text is identified to be a question. In that case, the extracted keyword from it will be reserved temporarily and only displayed (or recommended) once they are confirmed by the claimant (decided by the negation identification module).

We convert question and negation identifications as the binary-classification task and utilize the LSTM-based text classification model to solve it [Hochreiter and Schmidhuber, 1997; Liu *et al.*, 2017; Yang *et al.*, 2020]. From the perfor-

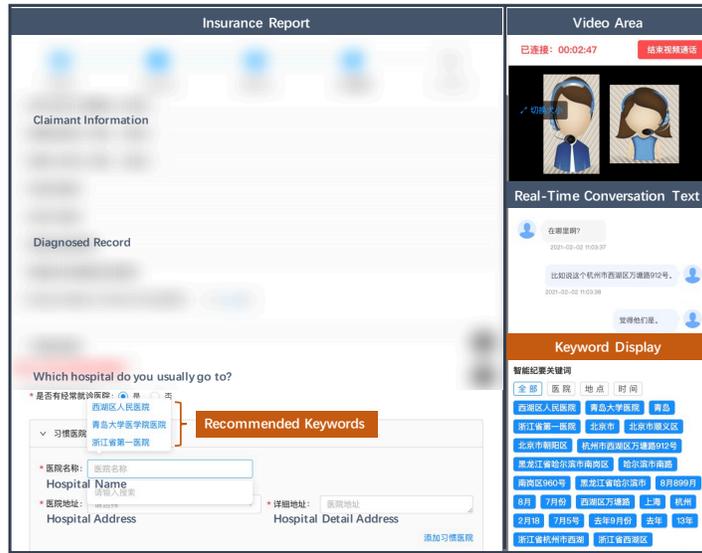


Figure 2: The dashboard of IAS. The left part is the insurance report that the assessor needs to fill and the right part displays the real-time conversation text (upper right) and the extracted keyword (bottom right).

mance on the human-annotated testing set, the question identification’s accuracy is 78%, and the negation identification’s accuracy is 92%.

2.4 Applications

Based on the previous three steps, the system gets filtered keywords extracted from the real-time conversion text, and these keywords are classified into different topics simultaneously.

We identify two types of generic applications for derived keywords: keyword display and report filling. The keyword display aims at displaying information extracted from the real-time conversation text. Furthermore, the report filling aims at recommending related contents when the assessor fills in the specific part of the insurance report. For ensuring a good user experience, we sort extracted keywords by the reversed order of utterances in the dialogue and limit the maximal number of recommended contents as 5.

3 Use Cases

Figure 2 shows a typical use case of IAS. In the figure, we present the dashboard of IAS. During the insurance assessment procedure, the dashboard’s upper right part demonstrates the real-time conversation text between the claimant and assessor, and the bottom right part displays the keywords extracted and filtered. The assessor fills in the insurance report on the left part of the dashboard, and the system will suggest related content when the assessor fill in the different form. Here, we demonstrate the system via a short video ³.

For the insurance assessor, keyword display and report filling significantly alleviate the workload of remembering all the claimant information during the inquiry procedure. From the latest 289 claimant cases, the overall adoption ratio of

recommended content achieves 20%. To further demonstrate IAS’s overall contribution to the assessor’s work, we present some online performance in the following section.

4 Online Performance

Our IAS has been deployed in the real insurance assessment scenario for three months and has served 300 claim cases a day on average. Based on statistical results, the average time cost of the insurance assessment procedure is reduced from 55 minutes to 35 minutes, and the overall human resources cost is saved 30% compared with previous offline insurance assessment. Moreover, with the assistance of our IAS, the insurance company provides more job opportunities for junior assessors. This indicates that IAS has effectively lowered the bar of insurance assessment jobs.

5 Conclusions and Future Work

We have presented IAS, a system that combines different NLP technologies in a novel way for providing value, especially to the junior insurance assessor.

Although there has been considerable work done for each of the independent modules/models, this systematic way of combining them delivers potential value to the insurance industry.

Acknowledgements

The authors would like to thank Jianzeng Liang, Chaojin Zhou, Weiwei Fan, and other members of Ant Group for helpful discussions and comments. We would also like to thank reviewers for their valuable comments.

References

[Cui *et al.*, 2019] Yiming Cui, Wanxiang Che, Ting Liu, Bing Qin, Ziqing Yang, Shijin Wang, and Guoping Hu.

³<https://youtu.be/3vmdmqBtJvw>

Pre-training with whole word masking for chinese bert.
arXiv preprint arXiv:1906.08101, 2019.

- [Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [Hochreiter and Schmidhuber, 1997] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- [Le and Titov, 2018] Phong Le and Ivan Titov. Improving entity linking by modeling latent relations between mentions. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1595–1604, Melbourne, Australia, July 2018. Association for Computational Linguistics.
- [Liu *et al.*, 2017] Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. Adversarial multi-task learning for text classification. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1–10, Vancouver, Canada, July 2017. Association for Computational Linguistics.
- [Mani *et al.*, 2020] Anirudh Mani, Shruti Palaskar, Nimshi Venkat Meripo, Sandeep Konam, and Florian Metze. Asr error correction and domain adaptation using machine translation. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6344–6348. IEEE, 2020.
- [Ouyang *et al.*, 2020] Yawen Ouyang, Moxin Chen, Xinyu Dai, Yinggong Zhao, Shujian Huang, and Jiajun Chen. Dialogue state tracking with explicit slot connection modeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 34–40, Online, July 2020. Association for Computational Linguistics.
- [Peng *et al.*, 2020] Shuang Peng, Hengbin Cui, Niantao Xie, Sujian Li, Jiaying Zhang, and Xiaolong Li. Enhanced-rnn: An efficient method for learning sentence similarity. In *Proceedings of The Web Conference 2020*, pages 2500–2506, 2020.
- [Yang *et al.*, 2020] Minghui Yang, Hengbin Cui, Shaosheng Cao, Yafang Wang, and Xiaolong Li. Servicegroup: A human-machine cooperation solution for group chat customer service. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2189–2192, 2020.