Research on the Marketing Transformation of Insurance Industry Under Generative Artificial Intelligence Technology

Xiaoying Lin[1a]* and Wei Ruan[2b]

*[a]linxy06@cuc.edu.cn , *[b]ruanwei_c@cbirc.gov.cn

[1] Shanghai Sanda University, Shanghai, 201209, China
[2] Wei Ruan, Shanghai Regulatory Bureau of State Administration for Financial Regulation,Shanghai,200127,China

Abstract. With the continuous development of artificial intelligence (AI) And natural language processing (NLP) technologies, Generative AI (AIGC) has become one of the most prominent technologies in the insurance industry. The application of AIGC technology can provide deep technological empowerment for the insurance industry in the marketing field: improving the production efficiency of insurance marketing content, improving the recommendation efficiency of insurance products, improving customer satisfaction and loyalty, and improving the service level of insurance marketing personnel; In response to the shortcomings of current AIGC large-scale language models, the paper proposes two innovative technical architecture solutions: combining large language models with knowledge graphs and vector databases to provide more knowledge for reasoning and understanding; The main issues faced by AIGC technology in the field of insurance marketing include copyright issues related to marketing content, potential privacy violations of users using data, data security issues, and the professionalism of data used in large language training models.

Key words:AIGC, insurance marketing,NLP technologies,knowledge graphs, vector databases.

1.Introduction

Marketing is an organizational function and a process of creating, communicating, and conveying customer value, as well as managing customer relationships in a way that benefits both the organization and stakeholders. [1] In the traditional insurance industry, offline marketing by agents plays a dominant role in the insurance marketing channel structure. Insurance agents mainly seek customers through product presentations and introductions from friends and family, understand customer needs, and recommend suitable insurance products to customers, providing them with various insurance services. In the wave of digital transformation, the traditional insurance industry's extensive "mass marketing" marketing strategy has become unsustainable. According to the report of the China Banking and Insurance Regulatory Commission, as of June 30, 2022, there were 5. 707 million sales personnel registered in the insurance intermediary supervision information system of insurance companies nationwide, compared with 9. 712 million on June 30, 2020, a decrease of more than 40 % in just two years. [2] The insurance industry urgently needs to integrate insurance technology with

ICPDI 2023, September 01-03, Chongqing, People's Republic of China

DOI 10.4108/eai.1-9-2023.2338781

marketing to achieve technological transformation in insurance marketing. With the continuous contraction of the agent scale, this transformation aims to enhance agent marketing services and efficiency, improve the production efficiency of marketing materials in self-operated modes, and elevate customer experience standards in after-sales services.

In recent years, artificial intelligence technology has been vigorously developed, and Large Language Models (LLMs) have been able to handle various Natural Language Processing (NLP) tasks. Based on large models such as BERT and knowledge bases, decision based artificial intelligence has been widely used in customer service, recommendation, and other fields.With the rapid growth of model scale, LLM, especially AIGC represented by ChatGPT last year, has the potential to solve many complex practical tasks, and can quickly generate various types and styles of content based on specific dialogue instructions, including pure text, graphics, animations, short videos, etc. In addition, it supports multiple rounds of dialogue, and better understands the intent as the conversation progresses, resulting in more accurate content.

Therefore, AIGC has become one of the most high-profile technologies in the insurance industry. The application of AIGC technology can provide the insurance industry with deep technological empowerment in the field of marketing, reducing marketing costs and improving marketing efficiency, helping the insurance industry shift from extensive growth to high-quality development.

2. The role of AIGC in insurance marketing

2.1 Improve the efficiency of insurance marketing content production

AIGC can help insurance companies improve the production efficiency of marketing materials. Some insurance companies use AIGC technology to quickly generate creative marketing copy, such as advertising language, email, text messages, WeChat public number articles, etc., to promote insurance products and services. For example, the domestic internet insurance company Zhong' an Insurance disclosed in the AIGC & ChatGPT Insurance Industry Application White Paper that the company has tried to use AIGC technology to generate various types of marketing materials. The company used to take about three days to produce a poster image for insurance products, but now it can produce multiple sets of different styles in less than three hours. [3] Compared to manually generated marketing materials, AIGC can not only build a more efficient content production line and bring in massive content supply, but also generate and optimize different content for different target audiences to achieve better and more effective reach and conversion results.

AIGC technology can be used by insurance companies to intelligently generate creative copywriting, as well as be integrated with different AI tools to automatically generate marketing images and videos for various scenarios, thus better presenting insurance products. For example, in advertising production practice, AI painting tools such as MidJourney and Stable Diffusion have been widely used in the production of print and video advertising. With the help of AIGC technology, models, set-up, filming, and other advertising processes can all be saved.

2.2 Improving the recommendation efficiency of insurance products

It is vital to possess the mindset of customers during the process of insurance marketing. User thinking means to take users as the core and understand their unique psychological needs accurately to provide differentiated products and services, making user focus the key. Personalized recommendation is one of the major applications of AIGC technology in insurance marketing. The basic principle is to predict user needs or interests based on their historical behavior and interest preferences through big data analysis and deep learning, and then present relevant product or service recommendations.

With the continuous development and popularity of social media, insurance companies have begun to utilize social media to promote insurance products. AIGC can interpret the network-wide public sentiment data in real-time to help brand owners grasp social media trends faster, as well as collect and analyze customer data in real time to interpret customers' behavior and preferences on social media, so as to optimize marketing strategies, generate corresponding insurance product recommendations, and improve promotional effectiveness on social media. Additionally, insurance companies need to consider the needs of customers from different regions, languages, and cultures when conducting cross-regional and cross-cultural marketing. AIGC technology supports multi-language analysis and recommendation to generate insurance product recommendations for different regions and languages, thus increasing product coverage and marketing effectiveness.

When recommending insurance products, in addition to considering users' needs and preferences, it is also necessary to consider the emotional state of the customer. The sentiment analysis technology in AIGC can help insurance companies better understand users'emotions and reactions. When users post comments or make posts, emotion analysis can determine whether their sentiment is positive or negative. This information can be used to respond to user needs more effectively and improve the efficiency of product promotion and customer satisfaction. Overall, by analyzing customer data and generating personalized insurance product recommendations, AIGC enhances marketing effectiveness and market share.

2.3 Improve the efficiency of insurance consulting services

Intelligent customer service is another application of AIGC technology, which is based on the principle of using artificial intelligence technology to simulate human chat interaction and achieve 24 / 7 online customer service. Intelligent customer service can effectively reduce marketing costs and improve customer satisfaction and loyalty.

In insurance marketing practice, insurance companies use AIGC technology to build intelligent insurance product consultation robots, providing customers with fast and convenient insurance product consultation services through natural language interaction. The robot can answer various questions about insurance products, including insurance types, insurance periods, insurance terms, insurance premiums, etc., providing customers with 24 / 7 online consulting services. When customers raise questions or requirements, AIGC can intelligently analyze the questions raised by users. Intelligent robots automatically generate corresponding solutions and service plans through deep learning algorithms and massive data analysis, meeting customer needs in the shortest possible time, thereby improving customer satisfaction and loyalty.

2.4 Improve the level of marketing services provided by agents

In 2022, the Chinese insurance industry achieved a total original premium income of 4. 7 trillion yuan, a year-on-year increase of 4. 58 %. [4] Facing such a large and complex market, how to accurately capture the needs of customers and provide insurance products that meet their needs is a problem that insurance agents need to address. Using AIGC technology, insurance companies can build intelligent insurance sales assistance robots for sales personnel, providing personalized insurance sales advice and support through natural language interaction. Robots can analyze customer needs and situations, provide sales personnel with the best sales strategies and solutions, and improve sales efficiency and customer satisfaction.

3. AIGC's technical solution for empowering insurance industry marketing

Productivity tools are key focus for implementing large models. However, solely relying on the training dataset of large models for domain-specific applications, such as insurance marketing, often falls short of achieving production-ready, instant-use outcomes. This article proposes two technical solutions: one is to supplement the training set with vertical / private domain content and integrate the knowledge graph, which is the LLM + Knowledge Graph + Prompt (LKP) architecture; The second is to introduce vector databases to provide long-term and short-term memory for large models, integrating domain knowledge bases, which is the LLM + VectorDB + Prompt (LVP) architecture. They will be discussed separately.

3.1 The realization mode of LLM + Knowledge Graph + Prompt (LKP) architecture

The knowledge graph is an artificial intelligence technology based primarily on graph database storage and the use of graph algorithms to retrieve and calculate data. It was initially proposed by Google in 2012 as an enhancement to its search results, which is different from traditional word segmentation matching. Instead, it aims to deeply explore the relationship between entities to provide users with more accurate search results and better user experience. Large Language Modeling (LLM) performs well in natural language processing tasks such as multiple rounds of dialogue and question-answering, but there are still many issues such as lack of up-to-date factual evidence. Specifically, LLMs memorize facts and knowledge contained in the training corpus. [5]The improvement of training effect by integrating Knowledge Graph technology (KG) in different stages is expected to solve the above problems. The LLM also assists knowledge graphs in having powerful abilities for unstructured data processing and natural language understanding, thereby facilitating more efficient knowledge extraction and expression from knowledge graphs.

3.1.1 Using LLM to enhance the knowledge graph

The greatest advantage of the knowledge map is the knowledge representation of the association between entities in complex relational networks, which can be used in tasks such as question answering, recommendation, and web search.However, traditional methods of constructing knowledge graphs often overlook non-structured textual data, leading to redundancy or gaps in information. In order to solve these issues, researchers have considered using LLM to enhance the construction of knowledge graphs, allowing them to fully consider unstructured data such as

text, thereby more accurately identifying the relationship between entities and ultimately improving performance on downstream tasks. The use of LLM to enhance knowledge maps includes aspects such as knowledge map embedding, knowledge map completion, knowledge map to text generation, and knowledge map question and answer. [6](p11)

Among them, knowledge graph embedding mainly utilizes the strong advantage of LLM to understand text information, which can promote more accurate graph embedding. The completion of the knowledge graph mainly uses LLM to mine implicit relationships between entities, which can promote the construction of a more complete knowledge graph; The generation of knowledge graph to text (KG to text) has always been limited by traditional technologies, so it is difficult to generate high-quality text that can accurately describe the input knowledge graph information. However, with the help of LLM, the usability of knowledge graphs can be significantly improved in more realistic natural language generation scenarios, including story creation and knowledge-based dialogue. [6](p16) Knowledge mapping Q & A using LLM can better combine the inference advantages of existing knowledge maps to answer questions in unknown fields.

3.1.2 LLM and knowledge mapping synergy

LLM can be used to understand the needs expressed by users' natural language, and the knowledge map can serve as a knowledge base providing the basis for knowledge. By combining the two in training, LLM and Knowledge Graph can create powerful language models for knowledge representation and reasoning. LLM is usually based on multiple rounds of iterative training with massive text corpora. Both text corpus and knowledge graphs contain a wealth of knowledge. But their knowledge structures are inconsistent. Generally speaking, knowledge in text corpora is unstructured and implicit, while knowledge in knowledge graphs is pre-processed and structured and explicit. [6](P18) To effectively collaborate between LLM and the knowledge graph, the first thing to address is aligning the knowledge in the text corpus and the knowledge graph.

In order to leverage the advantages of LLM and knowledge graphs simultaneously, researchers also use LLM and knowledge graphs to perform reasoning tasks for multiple applications. In question answering tasks, LLM can be first utilized to understand the unstructured text presented by the user and convert it into appropriate prompts to guide the inference of the knowledge graph. Relevant results are combined with the preceding prompts to be further analyzed through LLM.[6](p19)This establishes a bridge between unstructured text and structured data, aligning relevant knowledge, and making the process of reasoning and analysis more efficient and accurate.

3.1.3 Combination of Prompt Embedding and Knowledge Graph

Generally speaking, compared to human-driven approaches, knowledge graphs can be characterized by their ability to extract and refine data into finer and smaller granular dimensions. Combining the refined and understood small granular data from the knowledge map with the previously manually processed large chunks of data, we can use LLM to better understand the knowledge representation that requires comprehensive global complex networks / spanning multiple nodes for queries.

For example, we use LLM to query the knowledge representation of topic X. Without the combination of knowledge maps, prompt words (Prompt) can only be used to ask questions and queries in conjunction with the information on the topic X itself, which is relatively thin and insufficient, and the results are generally not ideal. However, if the prompt (Prompt) is embedded with a knowledge graph, the Prompt can perform two or more jump queries on the topic X entity from the knowledge graph. At this point, the information that comes from other nodes and involves multiple layers of entities related to the topic X can be included in Prompt for context learning due to the natural technical advantages of the knowledge graph, thereby overcoming the aforementioned problems and achieving better knowledge representation.

Although the LLM + Knowledge Graph + Prompt (LKP) architecture has received attention from some scholars and researchers, its training, research and development costs are high, and manufacturers rarely use this mode in practical applications.

3.2 Implementation method of LLM + VectorDB + Prompt (LVP) architecture

In the era of big models, some big model vendors such as OpenAI and Google Bard have chosen the LVP technology route of providing long-term and short-term memory for big models using vector databases. In the LVP architecture, an external memory is supplemented by a vector database as a large model. L acts as the computational unit to provide logical analysis and natural language interface capabilities, while V serves as the storage unit to provide stable, accurate, high-capacity, and highly scalable knowledge. At the same time, the content accumulated in the business can be directly transformed into assets at the knowledge base level, forming a closed-loop knowledge storage and output. On the basis of the former two, providing specific business-oriented adaptation capabilities. Taking the recommendation and consultation of insurance products as an example:

Build a product vector database (offline): first of all, the documents in the product document repository need to be semantically segmented and vectorized using a large language model to build a fragment-level vector database. In order to comprehensively consider the content of the entire document, it is necessary to summarize the entire document in the product document library and use a large language model to vectorize the summary language to construct a text-level vector database.

Product Advisory Q & A (Online): Vectorize the user question text using a large language model, and then use semantic query functionality to retrieve highly relevant product text fragments (Top-k) in a vector database based on the question vector. Finally, combine the user question and related document fragments to construct a Prompt as a prompt input for the large language model to generate question answers.For example, let's take the sale of assistant insurance agents:

Building a vector database (offline): Based on the product document library, build a fragment-level vector database and a text-level vector database.

Search for similar customer needs (online): summarize and vectorize user descriptions using a large language model, and then retrieve highly relevant similar customer needs (Top-k) from a vector database based on semantic query capabilities.

Reasoning user needs and revising marketing strategies (online): Using large language models to summarize user descriptions (multiple rounds of dialogue), vectorize the obtained product

descriptions related to users, and then retrieve relevant product text fragments in the vector database based on semantic query functions. Finally, based on the relevant product text fragments and the similar customer needs obtained in the previous step, compare and analyze and revise marketing strategies.

Generate marketing recommendations (online): The similar customer needs and user product portfolio obtained in the previous two steps are used to construct Prompt as prompt inputs for the large language model, which then generates relevant marketing recommendations based on reasoning.

4. Challenges in AIGC Implementation in Insurance Marketing

4.1 Copyright attribution of marketing materials

Advertising ideas and content such as ad copy, image and video ads generated by AIGC are works produced by AI. Due to the similarity between the source of demand and the training dataset, there is currently no clear legal regulation on the copyright of these works. Therefore, when insurance companies use AIGC to produce marketing materials, there may be potential commercial risks.

Based on current legislation status, Japan, Australia, UK and USA did not provide the civil subject qualification for AI, so its output does not have copyright. [7] This means that other subjects can freely use and distribute the creative advertising content generated by the AIGC system. Moreover, AIGC needs to be trained with vast amounts of data and material in the process of generating marketing ideas. Using other copyrighted materials may lead to disputes and legal disputes. [8] If an insurance company invests a lot of resources in developing the AIGC system to obtain commercial marketing ideas, it will undoubtedly face significant risks and increase other compliance and risk costs.

4.2 Data privacy and security concerns

AIGC is facing not only the copyright issue of marketing materials, but also issues related to the potential violation of user privacy and data security in the insurance marketing field.

AIGC needs a lot of data during its training process, and it is likely to contain personal information and sensitive data of users. For example, in utilizing AIGC technology to make personalized recommendations for insurance customers, in-depth analysis of the user's historical behavior and interests is required. These data sources are diverse, and even training on public datasets may contain personal identity information. Therefore, the collection and use of user data must comply with relevant laws and regulations and rules. The purpose and scope of user data must be clearly communicated to users and their explicit consent must be obtained. In addition, enterprises also need to adopt technological means for management to minimize the risk of such information leakage.

4.3 Scarcity of professional insurance data

The data sources of AI models include the following categories: public datasets, public websites, own data, crowdsourced data, synthetic data, and more. [9] Currently, the data used by AIGC in the large-scale language training model technology is mainly derived from public

datasets and public websites. However, if it needs to be applied in the context of insurance marketing, such as assisting agents in analyzing customer preferences or responding to customer requests for product consultation and recommendations, it requires more professional capabilities to ensure the professionalism and accuracy of relevant responses. [10]Therefore, how to ensure the professionalism of training data in the field of insurance marketing with AIGC technology is also a major challenge.

5.Conclusion

In the field of insurance marketing, AIGC technology can help insurance companies rapidly generate marketing creative text to increase the productivity of insurance marketing content. AIGC technology can predict users 'needs or interest points based on users' historical behaviors and interest preferences through big data analysis and deep learning, thereby improving the recommendation efficiency of insurance products; AIGC technology can provide intelligent customer service, effectively reducing marketing costs, and enhancing customer satisfaction and loyalty. AIGC technology can also provide personalized insurance sales advice and support to salespersons to enhance their service quality.

In analyzing how AIGC technology can support insurance marketing, this paper addresses the deficiencies of large language models, and proposes a method that combines large language models with knowledge graphs. Innovatively, two technical solutions are proposed: one is to supplement vertical / private domain content to the training set and integrate knowledge graphs, which is the LLM + Knowledge Graph + Prompt (LKP) architecture; Second is the introduction of a vector database to provide long and short term memories for large models and integrated with domain knowledge repositories, known as the LLM + VectorDB + Prompt (LVP) architecture. This technique achieves great results for a variety of NLP tasks such as answering open domain questions, generating automated summaries and dialogues, and more.

In the last chapter of this thesis, a forecast was presented to the potential problems AIGC technology will encounter in the field of insurance marketing, including copyright issues in the marketing content produced by AIGC technology, data security issues related to the potential infringement of user privacy, and professional data usage in large-scale language training models.

References

[1]Gregory T. Gundlach, William L. Wilkie.(2009), The American Marketing Association’s New Definition of Marketing: Perspective and Commentary on the 2007 Revision. Journal of Public Policy & Marketing.,Vol.28:
259 ,https://www3.nd.edu/~wwilkie/pdf/2009MKTGDefinationGundlachAndWilkieJPPM.pdf
[2]China’s State Administration of Financial Regulation .(2022),Announcement of the General Office of the China Banking and Insurance Regulatory Commission on the Practice Registration of Insurance Company Sales Practitioners in the First Half of 2022.http://www.cbirc.gov.cn/cn/view/pages/ItemDetail.html?docId=1071608&itemId=925&generaltype=0

[3]Zhong An Insurance: White Paper on the Application of AIGC in the Insurance Industry, https://finance.ifeng.com/c/8QSA2oCmQRR
[4] China's State Administration of Financial Regulation.(2023)Table of Original Insurance Premium Income by Region in China as of December 2022.http://www.cbirc.gov.cn/cn/view/pages/ItemDetail.html?docId=1093184&itemId=954&generaltype=0
[5] F. Petroni, T. Rocktaschel, S. Riedel, P. Lewis, A. Bakhtin, Y. Wu, ¨ and A. Miller, "Language models as knowledge bases?" in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019, pp. 2463– 2473.
[6]Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, Xindong Wu.(2021)Unifying Large Language Models and Knowledge Graphs: A Roadmap. Journal of Latex Class Files.14:1-29 https://doi.org/10.48550/arXiv.2306.08302
[7] Xin Yuan Media Integration ,Does artificial intelligence generated content (AIGC) have copyright, https://rmh.pdnews.cn/Pc/ArtInfoApi/article?id=36261019
[8]Duan Chunlin,2023, The Impact and Challenge of AIGC on Advertising Marketing Communication, https://baijiahao.baidu.com/s?id=1768311597816221757&wfr=spider&for=pc
[9] Intelligent Governance Dialogue ChatGPT: The Integration of Knowledge Graph and LLM and Its Future Development, https://www.sohu.com/a/681878093_121124777
[10]Zhong An Insurance: White Paper on the Application of AIGC in the Insurance Industry, https://finance.ifeng.com/c/8QSA2oCmQRR