Article

LLMs for Insurance: Opportunities, Challenges and Concerns

Chunling Cao [1,*], Xiuling Yao [1], Wenying Gong [1], Yutong Li [2], Yang Pan [3], Yuechen Yang [1] and Chaoran Zhao [1]

1 Jilin University; yaoxl@jlu.edu.cn (X.Y.); gongwy@jlu.edu.cn (W.G.); yyc23@mails.jlu.edu.cn (Y.Y.); zhaocr23@mails.jlu.edu.cn (C.Z.)
2 Northeast Electric Power University; liyutong241@gmail.com
3 Jilin Provincial Institute of Education; panyang202310@163.com
* Correspondence: caocl@jlu.edu.cn

Abstract: Large Language Models (LLMs) have revolutionized the financial services sector by enhancing data processing, decision-making, and customer interaction. Particularly in the insurance industry, LLMs facilitate significant advancements by automating complex processes and personalizing customer engagements, which increases efficiency and satisfaction. This paper explores the integration of LLMs within the insurance sector, highlighting their capabilities in sentiment analysis, risk assessment, and tailored service provision. However, deploying these models presents substantial challenges concerning data privacy, security, and the ethical implications of automated decision-making. Ensuring the fairness and transparency of AI-driven processes is imperative to address potential biases and maintain consumer trust. The paper also discusses robust risk management strategies essential for implementing LLMs in sensitive environments, focusing on continuous monitoring and the need for regular updates to security practices and compliance with data protection laws. The insurance industry can leverage LLMs to improve operational efficiencies and enhance customer service and risk management practices, positioning themselves at the forefront of technological innovation in the financial sector.

Keywords: large language model; insurance; financial service

---

1. Introduction

Large Language Models (LLMs) represent a significant advancement in artificial intelligence, especially within the financial services sector. These models, powered by deep learning and extensive training on diverse text data, excel in understanding and generating human language, which is pivotal for numerous Natural Language Processing (NLP) tasks [1–3]. Their capabilities extend from simple text generation to complex tasks like sentiment analysis and automated decision-making.

As the financial services industry continues to embrace technological advancements, LLMs are becoming increasingly integral in fostering innovation, particularly within the insurance sector [4–6]. These powerful models streamline complex processes by automating tasks that traditionally require human intelligence, such as processing claims and personalizing customer interactions. This automation significantly enhances operational efficiency and boosts customer satisfaction by delivering faster and more accurate services. For example, the integration of sentiment analysis and advanced risk assessment algorithms allows insurance companies to tailor their services with unprecedented precision [7–10]. By analyzing customer sentiment and behavior patterns, these technologies enable insurers to design insurance packages that are better suited to individual needs and risk profiles.

However, implementing LLMs in the insurance industry is fraught with challenges, primarily due to the sensitive nature of the data involved and the high standards required for accuracy and fairness. The insurance sector deals with large volumes of personal and financial information that must be handled with the utmost care to maintain privacy and security. Ensuring the integrity and confidentiality of this data while utilizing LLMs poses significant challenges [11,12]. Furthermore, there is a critical need to develop LLMs that make unbiased decisions. Since these models learn from historical data, there is a risk that they might perpetuate existing biases unless carefully managed. This necessitates robust mechanisms to monitor and refine the decision-making processes of LLMs continuously. Moreover, maintaining transparency in AI-driven decisions is vital to building and retaining trust among consumers. Insurers must provide clear explanations for decisions made by AI systems [13], especially when these decisions have significant implications for policyholders.

Given these complexities, the potential for LLMs in insurance also raises several concerns. As these models learn from vast datasets, there is a risk of perpetuating existing biases or creating new ethical dilemmas, such as discrimination in premium calculations or claim approvals [14–16]. The future applications of LLMs will need to address these issues carefully, ensuring that technological advancements enhance the industry without compromising ethical standards and customer trust. As the financial services industry continues to embrace technological advancements, LLMs are becoming increasingly integral in fostering innovation, particularly within the insurance sector [17]. These powerful models streamline complex processes by automating tasks that traditionally require human intelligence, such as processing claims and personalizing customer interactions. This automation significantly enhances operational efficiency and boosts customer satisfaction by delivering faster and more accurate services. For example, integrating sentiment analysis and advanced risk assessment algorithms allows insurance companies to precisely tailor their services. By analyzing customer sentiment and behavior patterns, these technologies enable insurers to design packages better suited to individual needs and risk profiles.

However, implementing LLMs in the insurance industry is fraught with challenges, primarily due to the sensitive nature of the data involved and the high standards required for accuracy and fairness [18,19]. The insurance sector deals with large volumes of personal and financial information that must be handled with the utmost care to maintain privacy and security [20–22]. Ensuring the integrity and confidentiality of this data while utilizing LLMs poses significant challenges. Furthermore, there is a critical need to develop LLMs that make unbiased decisions. Since these models learn from historical data, there is a risk that they might perpetuate existing biases unless carefully managed. This necessitates robust mechanisms to continuously monitor and refine the decision-making processes of LLMs. Moreover, maintaining transparency in AI-driven decisions is vital to building and retaining trust among consumers. Insurers must provide clear explanations for decisions made by AI systems, especially when these decisions have significant implications for policyholders.

2. Development of Large Language Models

Large Language Models are models trained with deep learning techniques to understand and generate human language [23,24]. The core function of a language model is to predict the next word or continue a text sequence, serving as the backbone for NLP tasks [25]. These NLP tasks range from text classification and language translation to sentiment analysis, and they rely on sophisticated language models to process and interpret extensive textual data.

Early language models relied on statistical methods like n-grams to predict the next word based on word sequence frequency. Though effective on limited vocabulary tasks, these models fell short in long-range dependencies and complex sentence structures. Then, the evolution of machine learning, especially with neural networks like Recurrent Neural Networks (RNNs) and later the Transformer architecture, greatly enhanced the ability to handle long sequences [26–28]. With a self-attention mechanism, transformers can process all sequence elements simultaneously. This is necessary for effectively capturing long-range dependencies needed for understanding and generating complex texts [29].

Current LLMs have demonstrated exceptional performance in semantic understanding, sentiment analysis, dialogue systems, and text generation [30,31]. Their success highlights their critical role in advancing AI technology, as they enhance human-computer interaction and automate information processing [32–34]. LLMs represent a significant leap in NLP, driving advancements in modern AI and paving the way for broader AI adoption and future applications.

The foundational technology of LLMs is the Transformer architecture [35,36]. This architecture revolutionized NLP by replacing recurrent layers with a series of attention mechanisms that allow the model to weigh the importance of different words in a sentence, regardless of their positional distance. Transformers consist of multiple layers of self-attention and feed-forward neural networks, enabling them to process exceptionally long data sequences and capture complex dependencies within them [37,38]. To build LLMs from the Transformer base, the architecture is scaled up by increasing the number of layers, attention heads, and parameters, allowing the model to learn from a vast corpus of textual data and capture nuances that smaller models might miss. This scalability is crucial as it enhances the model's ability to perform a wide range of linguistic tasks, from real-time translation to generating coherent text over extended passages. Distributed computing further facilitates the transition to LLMs by providing the necessary computational power to process large volumes of data across thousands of GPUs or TPUs. This not only accelerates the training process but also allows for larger batches of data, critical for model convergence on such scales. Increasing training volumes improves generalization, minimizes overfitting, and enhances the model's performance across a broader range of tasks, capturing more detailed and varied linguistic patterns. These technological advancements significantly set LLMs apart from their predecessors, pushing the boundaries of what is achievable in artificial intelligence and natural language processing.

The training of LLMs typically follows a structured approach. First, the model undergoes pre-training on a vast corpus of textual data to grasp the basic rules and patterns of language [39–41]. During this stage, the model learns to predict vocabulary, construct sentence structures, and understand language usage across various contexts by analyzing thousands of text samples [42–44]. Further, to enhance performance for specific applications, the model undergoes fine-tuning on particular downstream tasks such as text summarization, question answering, or sentiment analysis [45–48]. This fine-tuning phase enables the model to adapt to the specific demands of individual tasks, improving accuracy and effectiveness in targeted areas. LLMs boast fundamental capabilities like text generation, comprehension, and transformation, supporting a wide array of linguistic tasks. These capabilities make them powerful and versatile tools for applications, including machine translation, content summarization, automated question answering, and dialogue systems [49–52].

As technology evolves, the scope of LLM applications has vastly expanded beyond basic text handling. Modern LLMs not only manage text generation and comprehension but also engage in more complex cognitive tasks such as sentiment analysis and automatic summarization [53–55]. These functions enable the model to detect emotional tendencies in text or extract key information from lengthy documents. Moreover, LLMs are developing capabilities in logical reasoning and common-sense judgment, mimicking human thought processes to some extent and handling intricate reasoning tasks [56–58]. The latest developments include multimodal tasks like joint image and text generation and comprehension, enabling models to process and generate visuals related to textual information. Interactive learning marks another advancement for LLMs, with models continuously adjusting and optimizing their performance and outputs based on user feedback. This adaptability enhances the model's practicality and impact as an AI application in various fields.

Currently, LLM-related studies has seen significant breakthroughs, including OpenAI's GPT series, Google's BERT and T5, and Gemini [59–62]. These models have garnered widespread attention in academia and have proven highly effective in commercial applications, such as customer service, content recommendation, and automated editing. For example, Amazon [63] uses conversational bots powered by LLMs to enhance user interaction and improve customer service. Character.AI [64,65] leverages advanced LLMs to create engaging chatbots that provide human-like interactions for a unique conversational experience. These examples demonstrate that LLMs drive innovation across diverse industries and have become crucial in transforming modern applications.

3. Application of LLMs in the Fintech Industry

The financial services industry is a cornerstone of the modern economic system, encompassing key areas such as banking, insurance, investment management, payment systems, and financial consulting [66–68]. While these sectors vary in their service content and operational models, their core functions revolve around managing money and assets, facilitating capital flow, risk management, and delivering investment returns. These industries share a heavy reliance on large-scale data processing capabilities and advanced information technologies to handle complex financial transactions and vast amounts of customer data [69–72]. An effective customer service system is critical to their success, directly impacting customer satisfaction and the market reputation of businesses. Given the financial operations involved, risk management is also crucial, requiring precise risk assessment and management strategies to prevent potential financial losses and legal risks. The financial services industry faces common challenges in enhancing the efficiency of transactions and data processing, reducing operational costs, while also ensuring the security and accuracy of services. The complexity and scale of operations in this industry often involve cross-border transactions and multi-currency handling, making high-frequency transactions a norm. Moreover, these sectors must comply with strict regulatory and compliance requirements, including anti-money laundering regulations, data protection laws, and other financial regulatory standards. Ensuring data security is another major challenge, as the financial services industry is a high-risk target for cyber-attacks [73–75]. Any data breach or security flaw can lead to significant financial losses and damage to reputation. Therefore, improving customer experience, optimizing risk management, and enhancing data security are pressing issues that these industries need to address.

Facing these challenges, LLMs can leverage their advanced text comprehension and generation capabilities to help the financial services industry address multiple common issues. For instance, in banking, LLMs can automate customer inquiry services by understanding and responding to customer queries, enhancing service efficiency and customer satisfaction [76,77]. In the insurance sector, LLMs assist in automating claim processing by quickly and accurately analyzing claim requests and related documents, significantly speeding up the process and improving accuracy. In payment systems, LLMs provide smarter fraud detection by analyzing transaction patterns and behavior to effectively identify potential fraudulent activities and reduce losses. These applications not only boost operational efficiency but also enhance customer experience and help businesses cut costs and risks.

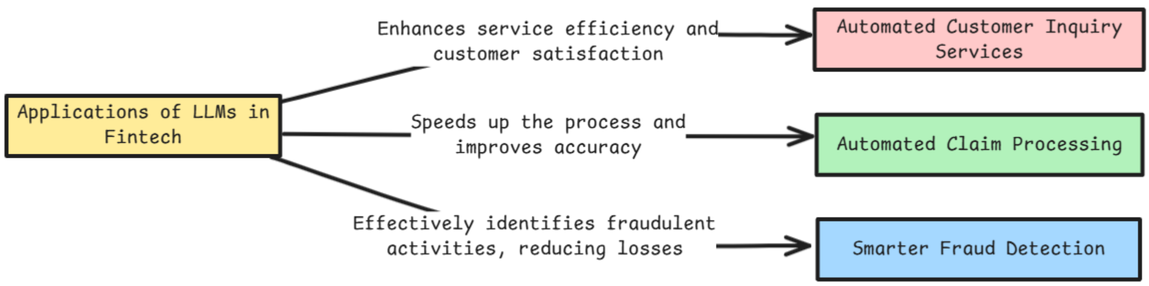

Figure 1. Caption

As LLM technology matures, its ability to understand complex texts, generate high-quality responses, and adapt to various financial environments will significantly improve. This means LLMs will be better equipped to handle complex queries and transactions, meet customer needs more effectively, and comply with increasingly strict regulatory demands:

1. LLMs greatly enhance service efficiency and quality in automated customer inquiry and support. They can understand specific customer questions and generate precise, personalized responses [78,79], boosting customer satisfaction and loyalty.
2. LLMs play an essential role in risk assessment and compliance monitoring. Analyzing vast amounts of transaction data and identifying patterns helps detect potential risks and fraud early [80,81]. This enables financial institutions to take preventive measures, reduce losses, and ensure regulatory compliance.
3. LLMs can automatically process and analyze large volumes of complex financial documents and transaction data, reducing the human workload and improving processing speed and

accuracy [82,83]. This capability is especially critical for financial institutions dealing with high-frequency transactions and cross-border activities.

4. LLMs drive financial innovation and new product development. By analyzing market trends and consumer behavior data, LLMs help financial institutions design and promote new financial products and services, enabling them to respond more effectively to changing market demands [84].

4. Application of LLMs in the Insurance Industry

The insurance industry forms the backbone of the financial services sector, crucial for managing risk and providing economic security [85–88]. The industry broadly splits into life insurance and non-life insurance categories. Life insurance, including health policies, primarily aims to offer financial support to individuals and their families in the event of death or health issues [89]. Non-life insurance covers property, auto, travel, and liability insurance, helping individuals and businesses mitigate financial losses from physical damage, accidents, and natural disasters [90]. By offering these products, the insurance sector helps transfer risk and ensures necessary financial support and recovery capabilities in the face of uncertainty and potential economic losses, thus supporting the stability and growth of society and the economy.

Meanwhile, the insurance industry, known for its protective role and extensive risk assessment, carries unique complexity and demands high predictiveness [91,92]. Beyond the usual challenges of the financial services sector—like data processing and customer service—the insurance industry faces additional hurdles. These include maintaining accurate risk assessments, detecting insurance fraud, and meeting the growing demand for customized insurance policies.

As the potential of LLMs is tapped, advanced models like ChatGPT are reshaping various aspects of the insurance industry [93–95]. These models significantly boost operational efficiency, customer interaction quality, and internal process management by utilizing technologies like reasoning, sentiment analysis, and text classification. LLMs particularly impact customer service, claims processing, and risk management.

In customer service, LLMs like ChatGPT handle a wide range of interactions in real-time, offering instant answers to common questions, assisting with policy details, and guiding customers through claims processes—all in a natural and engaging conversational style [96–99]. This not only raises customer satisfaction but also enhances service efficiency. By analyzing customer data, LLMs provide personalized interactions that improve service quality.

In claims processing, LLMs equipped with tabular learning capabilities expedite the process by automating initial assessments and document handling, employing text mining and knowledge extraction capabilities [100,101]. This automation shortens processing times and reduces human errors, making the entire claims process more efficient and customer-friendly.

Moreover, LLMs show great promise in risk management. LLMs effectively identify potential risks and fraud by analyzing extensive historical data and behavioral patterns and integrating time series analysis and anomaly detection technologies [102–105]. This helps insurance companies optimize product design and pricing strategies, better meeting market demands and customer expectations.

The insurance industry currently faces numerous challenges, including delayed customer service responses, complexity in the claims process, and inaccuracies in risk assessment. LLMs' advanced technology offers effective solutions to these issues. With their powerful text processing and data analysis capabilities, LLMs significantly enhance the quality and efficiency of insurance services.

1. **Improving Customer Service Response Times:** Traditional customer service often suffers from lengthy response times and low efficiency. With LLM's advanced text understanding and generation capabilities, real-time, automated customer responses become possible. LLMs can instantly answer customer queries and provide detailed information about insurance policies, using their deep understanding of natural language [106–109]. This immediate responsiveness

greatly enhances customer satisfaction, optimizes the customer experience, reduces waiting times, and boosts overall service efficiency.

2. **Streamlining the Claims Process:** The claims handling process is typically complex and time-consuming, requiring substantial manual document processing. LLMs utilize their natural language processing capabilities to automate the processing and review of claim documents, identify key information, and accelerate the document workflow [110–113]. This automation reduces the workload on human staff and increases processing speed and accuracy, making the claims process more efficient and customer-friendly.
3. **Enhancing the Accuracy of Risk Assessment:** Traditional risk assessment methods often rely on outdated data and limited analytical capabilities, which can lead to inaccurate risk predictions. LLMs, with their pattern recognition and data analysis capabilities, analyze vast amounts of historical data and behavioral patterns. Through precise data analysis and trend prediction, they enhance the accuracy of risk assessments. More accurate risk assessments help insurance companies optimize their insurance products and pricing strategies, reduce unnecessary losses, and offer more reasonable insurance services to customers, thereby increasing customer trust and satisfaction.

5. Opportunities for LLM in the Insurance Industry

5.1. Establishing Dialogue Agents to Enhance Customer Interaction and Personalized Services

In the insurance industry, the application of LLMs offers new possibilities for customer service and personalized interactions. As shown in Figure 2, using advanced multi-turn dialogue technology, LLMs can process customer inquiries and feedback in real-time, providing quick and accurate responses. For example, by deploying LLM-based dialogue agents, these systems can continuously track the context of conversations, better understand customer needs, and provide relevant solutions. This capability ensures faster responses and more personalized and precise customer service.

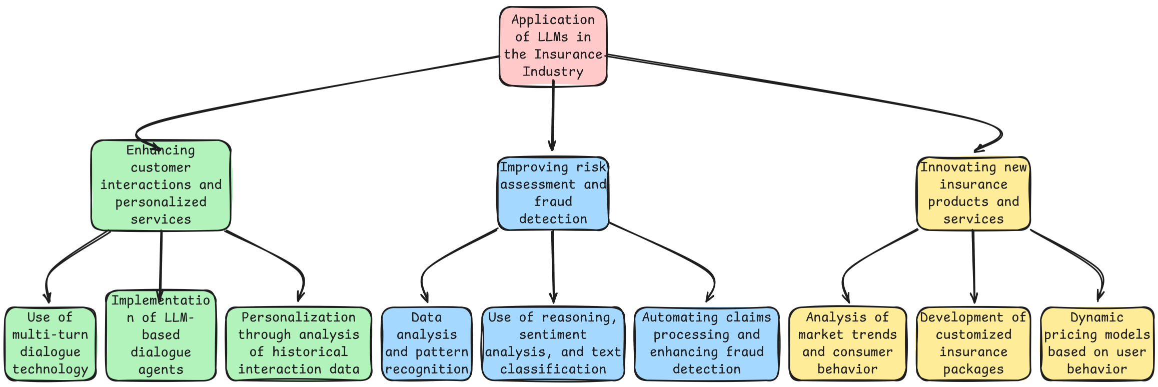

Figure 2. LLMs' application in insurance.

Specifically, LLMs personalize service by analyzing customers' historical interaction data, inferring potential needs from past inquiries, and proactively providing relevant information. This not only improves customer satisfaction, but also significantly improves efficiency by reducing repeated requests and misunderstandings. For example, if a customer frequently asks about car insurance details, the LLM could proactively offer the latest deals and coverage options in the next conversation, further enhancing customer experience and loyalty.

In the insurance industry, LLMs can revolutionize customer service and interaction through the use of Personalized Conversational Information Retrieval (CIR) [114,115] technology. This technology allows LLMs to tailor responses and recommendations by analyzing users' historical conversations and behavior patterns. This involves learning from vast amounts of dialogue data to recognize preferences and needs, allowing LLMs to provide more accurate and relevant information in future interactions.

For example, suppose a customer previously expressed interest or queries about a specific type of insurance. In that case, the LLM can proactively offer detailed information and the latest deals on that insurance type in subsequent conversations, achieving highly personalized customer service.

Moreover, this analysis and learning from historical data not only enhance the relevance of information but also significantly improve service efficiency. By automatically identifying user needs and behavior trends, LLMs can anticipate potential questions and needs, thus providing solutions before the customer even asks. This predictive service approach reduces customer wait times and greatly increases customer satisfaction and loyalty. For instance, by analyzing a customer's history of conversations about insurance claims, an LLM can immediately provide updates on claim status upon the customer's next contact, reducing uncertainty and enhancing the transparency of the overall service.

5.2. Improving Risk Assessment and Fraud Detection

LLMs enhance fraud detection and risk assessment in the insurance industry by leveraging their mathematical, textual, and symbolic reasoning capabilities. Each type of reasoning plays a crucial role in interpreting data and making informed decisions:

1. **Mathematical Reasoning:** Mathematical reasoning enables LLMs to handle complex calculations and statistical analyses required for risk assessment [116–118]. For instance, LLMs can evaluate the probability of certain events based on historical data, such as the likelihood of a fraudulent claim based on past claim patterns and financial behaviors. This type of reasoning helps insurers quantify risk levels and make decisions based on statistical evidence, thereby improving the precision of risk assessments.
2. **Textual Reasoning:** Textual reasoning involves understanding and interpreting written content within context [119–122]. In the realm of insurance, this ability allows LLMs to analyze the textual data from claims or customer interactions to identify signs of possible fraud. For example, discrepancies in incident reports or claims that deviate from typical patterns can be flagged for further investigation. Textual reasoning helps LLMs understand the nuances of language used in claims, spotting inconsistencies and anomalies that could indicate fraudulent activities.
3. **Symbolic Reasoning:** Symbolic reasoning refers to the ability of LLMs to manipulate and reason with symbols, typically used in logical deduction and problem-solving scenarios [123–125]. In insurance, symbolic reasoning can be applied to automate and refine the decision-making processes. For example, by defining certain rules and conditions for what constitutes a high-risk claim, LLMs can apply these criteria systematically, checking claims against established risk indicators. This type of reasoning ensures that decisions are consistent and based on defined insurance policy parameters, reducing human error and bias in fraud detection and risk assessment processes.

These advancements not only speed up the processing of claims and assessments but also ensure a higher degree of accuracy and reliability in detecting fraudulent activities and assessing risks accurately.

Current reasoning models like o1 and table-based models such as TableGPT2 [126,127] significantly enhance fraud detection capabilities in industries like insurance by adeptly processing and interpreting structured data prevalent in claims data and financial records. LLMs with tabular processing capabilities excel in analyzing tabular data, identifying patterns, anomalies, or inconsistencies that suggest fraudulent activities by examining datasets for irregular claims patterns such as unusually high claims from specific regions [128–131]. Meanwhile, models like o1 apply advanced mathematical and symbolic reasoning to correlate diverse data points—such as the frequency of claims by a single policyholder or typical payout amounts—thereby constructing comprehensive risk profiles that aid in preemptive fraud detection. These models automate and accelerate the fraud detection process, improving the accuracy of fraud identification and reducing financial losses and operational costs for insurance companies, increasing trust and ensuring a fair and transparent claims process.

5.3. Innovating New Insurance Products and Services Through LLM Insights

Leveraging the deep learning and pattern recognition capabilities of LLMs, insurance companies can develop innovative products and services that meet the evolving market demands. The foundational technology of LLMs allows them to process and analyze vast, complex data, identifying potential patterns and trends crucial for understanding market dynamics and consumer behavior.

5.3.1. Emotional Analysis for Consumer Insight

Emotion analysis and empathetic computing are key technologies within LLMs, enabling the models to extract and interpret consumers' emotional states from text data [55,132]. Emotion analysis assesses positive or negative sentiments, while empathetic computing delves deeper into the motives and depths of these emotions. This deep emotional understanding allows LLMs to predict consumer preferences for specific insurance products [133–135], such as comprehensive health insurance, sought by those expressing a strong need for security.

For example, by analyzing customer feedback on insurance claim processes, LLMs can identify aspects that satisfy or dissatisfy customers, such as processing speed, communication efficiency, or transparency in compensation amounts. This information is crucial for insurance companies as it directly guides product adjustments and service improvements to better meet customer needs and enhance loyalty.

The application of this technology extends beyond collecting and analyzing historical data; it is vital for real-time monitoring and adaptation to market changes. The combination of emotion analysis and empathetic computing provides insurance companies with a powerful tool that allows them to adjust their marketing strategies and product designs in real time, responding to changes in consumer emotions and market trends. This flexibility and adaptability are indispensable in the competitive modern insurance market, especially today when consumer demands are increasingly personalized and diverse.

5.3.2. Synthetic Data for Market Simulation

LLM's synthetic data technology also uses advanced algorithms to generate virtual yet realistic market data that can simulate future market conditions and consumer behaviors. This capability is beneficial in the early stages of the development of new products when accurate market data are insufficient for decision making [136–139]. Synthetic data allows insurance companies to test the performance of their products in the market and identify potential risks and issues in a controlled environment before the products hit the market.

Generating synthetic data involves extracting statistical features from existing data and creating new data instances based on these features. This preserves the core characteristics and patterns of the original data while introducing a range of possible variations in the simulated datasets, thus improving the comprehensiveness and depth of testing. For example, by simulating consumer behavior during economic downturns or market upheavals, insurance companies can better understand how their products perform under complex conditions, ensuring robust product design.

The application of data synthesis provides strong support for the development, development, and marketing strategies of research for insurance products. By predicting product performance under various market conditions, companies can optimize risk management strategies, adjust product features to meet consumer needs, or identify and address potential market failures. This increases the success rate of the products and significantly reduces the economic losses that could result from products not meeting market demands. Therefore, LLM's synthetic data technology is indispensable for modern insurance companies facing rapidly changing and highly competitive market environments.

5.3.3. Time Series Analysis for Market Trends

LLMs' time series analysis capability is crucial for simulating market trends and consumer behavior cycles [6,140–142]. LLMs analyze historical market data to detect periodic changes and identify key factors driving these fluctuations, providing precise data support for pricing and positioning insurance products. Time-series analysis involves statistically examining data collected at specific intervals. LLMs leverage this capability to forecast future market conditions based on past patterns, enabling insurance companies to anticipate changes in demand and market dynamics [143]. This foresight helps them adjust strategies proactively, aligning their offerings with market needs. LLMs solve this problem effectively by processing and analyzing large datasets over time, a task traditional analytical tools often struggle with. Integrating LLMs with time series analysis improves prediction accuracy and accelerates data processing, empowering insurance companies to make faster, data-driven decisions. This capability is essential to maintain a competitive edge in a fast-paced market where predicting consumer behavior and trends determines business success.

6. Security, Risks, and Ethical Concerns

6.1. Data Privacy and Security Challenges with LLMs

Integrating LLMs into the insurance industry raises significant concerns about data privacy and security. LLMs need access to vast amounts of personal and sensitive data to train and operate effectively. This data often includes personal identifiers, financial history, and other private information that, if mishandled, could lead to serious breaches. The challenge is to implement robust encryption and access controls that protect data integrity and confidentiality while still allowing the models to function effectively. Moreover, compliance with regulations like GDPR [144,145] or HIPAA [146,147] is crucial for ensuring data is handled and protected properly.

6.2. Ethical Implications of Automated Decision-Making in Insurance

The use of LLMs for automated decision-making in insurance raises several ethical concerns. Algorithmic decisions, such as determining insurance premiums or claims outcomes, can significantly impact individuals' lives. If not properly overseen, these decisions may lead to unfair treatment or discrimination due to biases in the underlying data or algorithms. Ensuring transparency in decision-making and providing clear explanations is essential to maintain trust and fairness. Moreover, there must be a mechanism for human review and recourse, allowing individuals to contest decisions, especially in cases where the automated system may have made an error.

6.3. Risk Management Strategies for Deploying LLMs in Sensitive Environments

Deploying LLMs in sensitive environments like the insurance sector requires careful risk management. Thorough risk assessments are essential before implementation to identify vulnerabilities and develop mitigation strategies. Continuous monitoring of LLM activities and their decisions helps detect and address issues promptly. Insurance companies must also keep LLM systems updated with the latest security practices and data protection laws to prevent exploitation. Multi-layered security measures, combining physical and software-based defenses, help ensure system resilience against attacks. Finally, regular audits and updates of LLM training datasets are necessary to maintain effectiveness, reduce bias, and prevent operational failures or reputational damage.

References

1. Chang, Y.; Wang, X.; Wang, J.; Wu, Y.; Yang, L.; Zhu, K.; Chen, H.; Yi, X.; Wang, C.; Wang, Y.; others. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology* **2024**, *15*, 1–45.
2. Minaee, S.; Mikolov, T.; Nikzad, N.; Chenaghlu, M.; Socher, R.; Amatriain, X.; Gao, J. Large language models: A survey. *arXiv preprint arXiv:2402.06196* **2024**.
3. Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Wang, H. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997* **2023**.
4. Li, Y.; Wang, S.; Ding, H.; Chen, H. Large language models in finance: A survey. Proceedings of the fourth ACM international conference on AI in finance, 2023, pp. 374–382.

5. Nie, Y.; Kong, Y.; Dong, X.; Mulvey, J.M.; Poor, H.V.; Wen, Q.; Zohren, S. A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges. *arXiv preprint arXiv:2406.11903* **2024**.
6. Hadi, M.U.; Al Tashi, Q.; Shah, A.; Qureshi, R.; Muneer, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Wu, J.; others. Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints* **2024**.
7. Xing, F.Z.; Cambria, E.; Welsch, R.E. Natural language based financial forecasting: a survey. *Artificial Intelligence Review* **2018**, *50*, 49–73.
8. Wu, S.; Irsoy, O.; Lu, S.; Dabravolski, V.; Dredze, M.; Gehrmann, S.; Kambadur, P.; Rosenberg, D.; Mann, G. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564* **2023**.
9. Zhang, J.; Xie, H.; Zhang, X.; Liu, K. Enhancing Risk Assessment in Transformers with Loss-at-Risk Functions. *arXiv preprint arXiv:2411.02558* **2024**.
10. Zhao, H.; Liu, Z.; Wu, Z.; Li, Y.; Yang, T.; Shu, P.; Xu, S.; Dai, H.; Zhao, L.; Mai, G.; others. Revolutionizing finance with llms: An overview of applications and insights. *arXiv preprint arXiv:2401.11641* **2024**.
11. Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günnemann, S.; Hüllermeier, E.; others. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* **2023**, *103*, 102274.
12. Yang, J.; Jin, H.; Tang, R.; Han, X.; Feng, Q.; Jiang, H.; Zhong, S.; Yin, B.; Hu, X. Harnessing the power of llms in practice: A survey on chatgpt and beyond. *ACM Transactions on Knowledge Discovery from Data* **2024**, *18*, 1–32.
13. Dan, H.C.; Lu, B.; Li, M. Evaluation of asphalt pavement texture using multiview stereo reconstruction based on deep learning. *Construction and Building Materials* **2024**, *412*, 134837.
14. Ozbayoglu, A.M.; Gudelek, M.U.; Sezer, O.B. Deep learning for financial applications: A survey. *Applied soft computing* **2020**, *93*, 106384.
15. Kearney, C.; Liu, S. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis* **2014**, *33*, 171–185.
16. Naveed, H.; Khan, A.U.; Qiu, S.; Saqib, M.; Anwar, S.; Usman, M.; Akhtar, N.; Barnes, N.; Mian, A. A comprehensive overview of large language models. *arXiv preprint arXiv:2307.06435* **2023**.
17. Girotra, K.; Meincke, L.; Terwiesch, C.; Ulrich, K.T. Ideas are dimes a dozen: Large language models for idea generation in innovation. *Available at SSRN 4526071* **2023**.
18. Zhang, J.; Bao, K.; Zhang, Y.; Wang, W.; Feng, F.; He, X. Is chatgpt fair for recommendation? evaluating fairness in large language model recommendation. Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 993–999.
19. Wang, P.; Li, L.; Chen, L.; Cai, Z.; Zhu, D.; Lin, B.; Cao, Y.; Liu, Q.; Liu, T.; Sui, Z. Large language models are not fair evaluators. *arXiv preprint arXiv:2305.17926* **2023**.
20. Yao, Y.; Duan, J.; Xu, K.; Cai, Y.; Sun, Z.; Zhang, Y. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. *High-Confidence Computing* **2024**, p. 100211.
21. Pan, X.; Zhang, M.; Ji, S.; Yang, M. Privacy risks of general-purpose language models. 2020 IEEE Symposium on Security and Privacy (SP). IEEE, 2020, pp. 1314–1331.
22. Liu, X.; Xie, L.; Wang, Y.; Zou, J.; Xiong, J.; Ying, Z.; Vasilakos, A.V. Privacy and security issues in deep learning: A survey. *IEEE Access* **2020**, *9*, 4566–4593.
23. Hadi, M.U.; Qureshi, R.; Shah, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Wu, J.; Mirjalili, S.; others. A survey on large language models: Applications, challenges, limitations, and practical usage. *Authorea Preprints* **2023**.
24. Zhao, W.X.; Zhou, K.; Li, J.; Tang, T.; Wang, X.; Hou, Y.; Min, Y.; Zhang, B.; Zhang, J.; Dong, Z.; others. A survey of large language models. *arXiv preprint arXiv:2303.18223* **2023**.
25. Gillioz, A.; Casas, J.; Mugellini, E.; Abou Khaled, O. Overview of the Transformer-based Models for NLP Tasks. 2020 15th Conference on computer science and information systems (FedCSIS). IEEE, 2020, pp. 179–183.
26. Vaswani, A. Attention is all you need. *Advances in Neural Information Processing Systems* **2017**.
27. Kalyan, K.S.; Rajasekharan, A.; Sangeetha, S. Ammus: A survey of transformer-based pretrained models in natural language processing. *arXiv preprint arXiv:2108.05542* **2021**.
28. Ansar, W.; Goswami, S.; Chakrabarti, A. A Survey on Transformers in NLP with Focus on Efficiency. *arXiv preprint arXiv:2406.16893* **2024**.

29. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence* **2013**, *35*, 1798–1828.
30. Bai, G.; Liu, J.; Bu, X.; He, Y.; Liu, J.; Zhou, Z.; Lin, Z.; Su, W.; Ge, T.; Zheng, B.; others. Mt-bench-101: A fine-grained benchmark for evaluating large language models in multi-turn dialogues. *arXiv preprint arXiv:2402.14762* **2024**.
31. OpenAI. New and Improved Embedding Model. https://openai.com/index/new-and-improved-embedding-model/, 2024.
32. Yuan, Y.; Huang, Y.; Ma, Y.; Li, X.; Li, Z.; Shi, Y.; Zhou, H. Rhyme-aware Chinese lyric generator based on GPT. Fourth International Conference on Advanced Algorithms and Neural Networks (AANN 2024). SPIE, 2024, Vol. 13416, pp. 667–672.
33. Li, S.; He, Y.; Guo, H.; Bu, X.; Bai, G.; Liu, J.; Liu, J.; Qu, X.; Li, Y.; Ouyang, W.; others. GraphReader: Building Graph-based Agent to Enhance Long-Context Abilities of Large Language Models. *arXiv preprint arXiv:2406.14550* **2024**.
34. Gao, M.; Wei, Y.; Li, Z.; Huang, B.; Zheng, C.; Mulati, A. A survey of machine learning algorithms for defective steel plates classification. International Conference on Computing, Control and Industrial Engineering. Springer, 2024, pp. 467–476.
35. Liu, Y.; Han, T.; Ma, S.; Zhang, J.; Yang, Y.; Tian, J.; He, H.; Li, A.; He, M.; Liu, Z.; others. Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology* **2023**, p. 100017.
36. Mahowald, K.; Ivanova, A.A.; Blank, I.A.; Kanwisher, N.; Tenenbaum, J.B.; Fedorenko, E. Dissociating language and thought in large language models. *Trends in Cognitive Sciences* **2024**.
37. Kaddour, J.; Harris, J.; Mozes, M.; Bradley, H.; Raileanu, R.; McHardy, R. Challenges and applications of large language models. *arXiv preprint arXiv:2307.10169* **2023**.
38. Wang, Z.; Yan, H.; Wei, C.; Wang, J.; Bo, S.; Xiao, M. Research on autonomous driving decision-making strategies based deep reinforcement learning. *arXiv preprint arXiv:2408.03084* **2024**.
39. Min, B.; Ross, H.; Sulem, E.; Veyseh, A.P.B.; Nguyen, T.H.; Sainz, O.; Agirre, E.; Heintz, I.; Roth, D. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys* **2023**, *56*, 1–40.
40. Gao, Z.; Wu, Y.; Bu, X.; Yu, T.; Yuan, J.; Jia, Y. Learning a robust representation via a deep network on symmetric positive definite manifolds. *Pattern Recognition* **2019**, *92*, 1–12.
41. Han, Z.; Gao, C.; Liu, J.; Zhang, J.; Zhang, S.Q. Parameter-efficient fine-tuning for large models: A comprehensive survey. *arXiv preprint arXiv:2403.14608* **2024**.
42. Zhang, S.; Dong, L.; Li, X.; Zhang, S.; Sun, X.; Wang, S.; Li, J.; Hu, R.; Zhang, T.; Wu, F.; others. Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792* **2023**.
43. Ziegler, D.M.; Stiennon, N.; Wu, J.; Brown, T.B.; Radford, A.; Amodei, D.; Christiano, P.; Irving, G. Fine-tuning language models from human preferences. *arXiv preprint arXiv:1909.08593* **2019**.
44. Feng, W.; Bu, X.; Zhang, C.; Li, X. Beyond bounding box: Multimodal knowledge learning for object detection. *arXiv preprint arXiv:2205.04072* **2022**.
45. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; others. Language models of protein sequences at the scale of evolution enable accurate structure prediction. *BioRxiv* **2022**, *2022*, 500902.
46. Creswell, A.; Shanahan, M. Faithful reasoning using large language models. *arXiv preprint arXiv:2208.14271* **2022**.
47. Zhang, J.; Wang, X.; Jin, Y.; Chen, C.; Zhang, X.; Liu, K. Prototypical Reward Network for Data-Efficient RLHF. *arXiv preprint arXiv:2406.06606* **2024**.
48. Petroni, F.; Rocktäschel, T.; Lewis, P.; Bakhtin, A.; Wu, Y.; Miller, A.H.; Riedel, S. Language models as knowledge bases? *arXiv preprint arXiv:1909.01066* **2019**.
49. Brants, T.; Popat, A.; Xu, P.; Och, F.J.; Dean, J. Large language models in machine translation. Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), 2007, pp. 858–867.
50. Kocmi, T.; Federmann, C. Large language models are state-of-the-art evaluators of translation quality. *arXiv preprint arXiv:2302.14520* **2023**.

51. Mo, F.; Mao, K.; Zhu, Y.; Wu, Y.; Huang, K.; Nie, J.Y. ConvGQR: generative query reformulation for conversational search. *arXiv preprint arXiv:2305.15645* **2023**.
52. Bu, X.; Wu, Y.; Gao, Z.; Jia, Y. Deep convolutional network with locality and sparsity constraints for texture classification. *Pattern Recognition* **2019**, *91*, 34–46.
53. Xiao, L.; Chen, X.; Shan, X. Enhancing large language models with evolutionary fine-tuning for news summary generation. *Journal of Intelligent & Fuzzy Systems*, pp. 1–13.
54. Zhang, W.; Deng, Y.; Liu, B.; Pan, S.J.; Bing, L. Sentiment analysis in the era of large language models: A reality check. *arXiv preprint arXiv:2305.15005* **2023**.
55. Mao, R.; Liu, Q.; He, K.; Li, W.; Cambria, E. The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection. *IEEE transactions on affective computing* **2022**, *14*, 1743–1753.
56. Chu, Z.; Chen, J.; Chen, Q.; Yu, W.; He, T.; Wang, H.; Peng, W.; Liu, M.; Qin, B.; Liu, T. A survey of chain of thought reasoning: Advances, frontiers and future. *arXiv preprint arXiv:2309.15402* **2023**.
57. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Xia, F.; Chi, E.; Le, Q.V.; Zhou, D.; others. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems* **2022**, *35*, 24824–24837.
58. Self-Reflection, S. Self-Contrast: Better Reflection Through Inconsistent Solving Perspectives.
59. Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; Almeida, D.; Altenschmidt, J.; Altman, S.; Anadkat, S.; others. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* **2023**.
60. Lin, J.; Nogueira, R.; Yates, A. *Pretrained transformers for text ranking: Bert and beyond*; Springer Nature, 2022.
61. Team, G.; Anil, R.; Borgeaud, S.; Alayrac, J.B.; Yu, J.; Soricut, R.; Schalkwyk, J.; Dai, A.M.; Hauth, A.; Millican, K.; others. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805* **2023**.
62. Team, G.; Mesnard, T.; Hardin, C.; Dadashi, R.; Bhupatiraju, S.; Pathak, S.; Sifre, L.; Rivière, M.; Kale, M.S.; Love, J.; others. Gemma: Open models based on gemini research and technology. *arXiv preprint arXiv:2403.08295* **2024**.
63. Weise, K. Amazon Has New Chatbot For Shoppers. *The New York Times* **2024**, pp. B1–B1.
64. Syaqih, K.N.; Qaisara, A.D.; Lail, M. Character. Ai: Analisis Penyalahgunaan Chatbot Artificial Intelligence Untuk Kepuasan Seksual Pada Remaja Perempuan Daerah Jakarta. *ADAPTASI: Jurnal Sosial Humaniora Dan Keagamaan* **2024**, *1*, 35–44.
65. Character.AI. https://character.ai/. Accessed: 2023-10-15.
66. Ahmed, S.; Alshater, M.M.; El Ammari, A.; Hammami, H. Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance* **2022**, *61*, 101646.
67. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance* **2021**, *32*, 100577.
68. Xu, Z.; Guo, L.; Zhou, S.; Song, R.; Niu, K. Enterprise supply chain risk management and decision support driven by large language models. *Applied Science and Engineering Journal for Advanced Research* **2024**, *3*, 1–7.
69. Kumar, S.; Sharma, D.; Rao, S.; Lim, W.M.; Mangla, S.K. Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research. *Annals of Operations Research* **2022**, pp. 1–44.
70. Xu, W.; Xiao, J.; Chen, J. Leveraging Large Language Models to Enhance Personalized Recommendations in E-commerce. *arXiv preprint arXiv:2410.12829* **2024**.
71. Khodadadi, A.; Ghandiparsi, S.; Chuah, C.N. A Natural Language Processing and deep learning based model for automated vehicle diagnostics using free-text customer service reports. *Machine Learning with Applications* **2022**, *10*, 100424.
72. Mo, F.; Mao, K.; Zhao, Z.; Qian, H.; Chen, H.; Cheng, Y.; Li, X.; Zhu, Y.; Dou, Z.; Nie, J.Y. A survey of conversational search. *arXiv preprint arXiv:2410.15576* **2024**.
73. Xie, Q.; Han, W.; Zhang, X.; Lai, Y.; Peng, M.; Lopez-Lira, A.; Huang, J. Pixiu: A large language model, instruction data and evaluation benchmark for finance. *arXiv preprint arXiv:2306.05443* **2023**.
74. Abu-Salih, B.; Alweshah, M.; Alazab, M.; Al-Okaily, M.; Alahmari, M.; Al-Habashneh, M.; Al-Sharaeh, S. Natural language inference model for customer advocacy detection in online customer engagement. *Machine Learning* **2024**, *113*, 2249–2275.

75. Xiang, C.; Zhang, J.; Li, F.; Fei, H.; Ji, D. A semantic and syntactic enhanced neural model for financial sentiment analysis. *Information Processing & Management* **2022**, *59*, 102943.
76. Doerr, S.; Gambacorta, L.; Garralda, J.M.S. Big data and machine learning in central banking. *BIS Working Papers* **2021**.
77. Petropoulos, A.; Siakoulis, V.; Stavroulakis, E.; Vlachogiannakis, N.E. Predicting bank insolvencies using machine learning techniques. *International Journal of Forecasting* **2020**, *36*, 1092–1113.
78. Wang, J.; Mo, F.; Ma, W.; Sun, P.; Zhang, M.; Nie, J.Y. A User-Centric Multi-Intent Benchmark for Evaluating Large Language Models. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 2024, pp. 3588–3612.
79. Mo, F.; Nie, J.Y.; Huang, K.; Mao, K.; Zhu, Y.; Li, P.; Liu, Y. Learning to relate to previous turns in conversational search. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023, pp. 1722–1732.
80. Yuan, T.; He, Z.; Dong, L.; Wang, Y.; Zhao, R.; Xia, T.; Xu, L.; Zhou, B.; Li, F.; Zhang, Z.; others. R-judge: Benchmarking safety risk awareness for llm agents. *arXiv preprint arXiv:2401.10019* **2024**.
81. Tang, X.; Jin, Q.; Zhu, K.; Yuan, T.; Zhang, Y.; Zhou, W.; Qu, M.; Zhao, Y.; Tang, J.; Zhang, Z.; others. Prioritizing safeguarding over autonomy: Risks of llm agents for science. *arXiv preprint arXiv:2402.04247* **2024**.
82. Wang, Y.; Gangi Reddy, R.; Mujahid, Z.M.; Arora, A.; Rubashevskii, A.; Geng, J.; Afzal, O.M.; Pan, L.; Borenstein, N.; Pillai, A.; others. Factcheck-gpt: End-to-end fine-grained document-level fact-checking and correction of llm output. *arXiv e-prints* **2023**, pp. arXiv–2311.
83. Biswas, A.; Talukdar, W. Robustness of Structured Data Extraction from In-Plane Rotated Documents Using Multi-Modal Large Language Models (LLM). *Journal of Artificial Intelligence Research* **2024**.
84. Qiu, L.; Singh, P.V.; Srinivasan, K. How Much Should We Trust LLM Results for Marketing Research? *Available at SSRN 4526072* **2023**.
85. Dercon, S. Risk, insurance, and poverty: a review. *Insurance against poverty* **2005**, pp. 9–37.
86. Levy, H.; Meltzer, D. The impact of health insurance on health. *Annu. Rev. Public Health* **2008**, *29*, 399–409.
87. Lewis, F.D. Dependents and the demand for life insurance. *The American economic review* **1989**, *79*, 452–467.
88. Zelizer, V.A.R. *Morals and markets: The development of life insurance in the United States*; Columbia University Press, 2017.
89. Gerber, H.U. *Life insurance mathematics*; Springer Science & Business Media, 2013.
90. Straub, E.; of Actuaries (Zürich), S.A. *Non-life insurance mathematics*; Springer, 1988.
91. Council, N.R.; others. *Science and Judgment in Risk Assessment*; National Academy Press, 1994.
92. Zio, E. The future of risk assessment. *Reliability Engineering & System Safety* **2018**, *177*, 176–190.
93. Dimri, A.; Paul, A.; Girish, D.; Lee, P.; Afra, S.; Jakubowski, A. A multi-input multi-label claims channeling system using insurance-based language models. *Expert Systems with Applications* **2022**, *202*, 117166.
94. Kang, I.; Woensel, W.V.; Seneviratne, O. Using Large Language Models for Generating Smart Contracts for Health Insurance from Textual Policies. In *AI for Health Equity and Fairness: Leveraging AI to Address Social Determinants of Health*; Springer, 2024; pp. 129–146.
95. Balona, C. ActuaryGPT: Applications of large language models to insurance and actuarial work. *Available at SSRN 4543652* **2023**.
96. Biswas, S. Using chatgpt for insurance: Current and prospective roles. *Available at SSRN 4405394* **2023**.
97. Mo, F.; Qu, C.; Mao, K.; Zhu, T.; Su, Z.; Huang, K.; Nie, J.Y. History-Aware Conversational Dense Retrieval. *arXiv preprint arXiv:2401.16659* **2024**.
98. Chen, H.; Mo, F.; Wang, Y.; Chen, C.; Nie, J.Y.; Wang, C.; Cui, J. A customized text sanitization mechanism with differential privacy. *arXiv preprint arXiv:2207.01193* **2022**.
99. Mao, K.; Qian, H.; Mo, F.; Dou, Z.; Liu, B.; Cheng, X.; Cao, Z. Learning denoised and interpretable session representation for conversational search. Proceedings of the ACM Web Conference 2023, 2023, pp. 3193–3202.
100. Hofert, M. Assessing ChatGPT's proficiency in quantitative risk management. *Risks* **2023**, *11*, 166.
101. Troxler, A.; Schelldorfer, J. Actuarial applications of natural language processing using transformers: Case studies for using text features in an actuarial context. *British Actuarial Journal* **2024**, *29*, e4.
102. Yu, X.; Chen, Z.; Ling, Y.; Dong, S.; Liu, Z.; Lu, Y. Temporal Data Meets LLM–Explainable Financial Time Series Forecasting. *arXiv preprint arXiv:2306.11025* **2023**.

103. Aracena, C.; Rodríguez, N.; Rocco, V.; Dunstan, J. Pre-trained language models in Spanish for health insurance coverage. Proceedings of the 5th Clinical Natural Language Processing Workshop, 2023, pp. 433–438.
104. Wu, J.; Yang, S.; Zhan, R.; Yuan, Y.; Wong, D.F.; Chao, L.S. A survey on llm-gernerated text detection: Necessity, methods, and future directions. *arXiv preprint arXiv:2310.14724* **2023**.
105. Jin, M.; Wang, S.; Ma, L.; Chu, Z.; Zhang, J.Y.; Shi, X.; Chen, P.Y.; Liang, Y.; Li, Y.F.; Pan, S.; others. Time-llm: Time series forecasting by reprogramming large language models. *arXiv preprint arXiv:2310.01728* **2023**.
106. Cao, H.; Zhang, Z.; Li, X.; Wu, C.; Zhang, H.; Zhang, W. Mitigating Knowledge Conflicts in Language Model-Driven Question Answering. *arXiv preprint arXiv:2411.11344* **2024**.
107. Mo, F.; Yi, B.; Mao, K.; Qu, C.; Huang, K.; Nie, J.Y. Convsdg: Session data generation for conversational search. Companion Proceedings of the ACM on Web Conference 2024, 2024, pp. 1634–1642.
108. Gao, H.; Wang, H.; Feng, Z.; Fu, M.; Ma, C.; Pan, H.; Xu, B.; Li, N. A novel texture extraction method for the sedimentary structures' classification of petroleum imaging logging. Pattern Recognition: 7th Chinese Conference, CCPR 2016, Chengdu, China, November 5-7, 2016, Proceedings, Part II 7. Springer, 2016, pp. 161–172.
109. Huang, K.; Mo, F.; Li, H.; Li, Y.; Zhang, Y.; Yi, W.; Mao, Y.; Liu, J.; Xu, Y.; Xu, J.; others. A Survey on Large Language Models with Multilingualism: Recent Advances and New Frontiers. *arXiv preprint arXiv:2405.10936* **2024**.
110. Tao, T.; Wang, X.; Mei, Q.; Zhai, C. Language model information retrieval with document expansion. Proceedings of the Human Language Technology Conference of the NAACL, Main Conference, 2006, pp. 407–414.
111. Tirumala, K.; Simig, D.; Aghajanyan, A.; Morcos, A. D4: Improving llm pretraining via document de-duplication and diversification. *Advances in Neural Information Processing Systems* **2023**, *36*, 53983–53995.
112. Li, X.; Cao, H.; Zhang, Z.; Hu, J.; Jin, Y.; Zhao, Z. Artistic Neural Style Transfer Algorithms with Activation Smoothing. *arXiv preprint arXiv:2411.08014* **2024**.
113. Yasunaga, M.; Leskovec, J.; Liang, P. Linkbert: Pretraining language models with document links. *arXiv preprint arXiv:2203.15827* **2022**.
114. Yu, S.; Liu, Z.; Xiong, C.; Feng, T.; Liu, Z. Few-shot conversational dense retrieval. Proceedings of the 44th International ACM SIGIR Conference on research and development in information retrieval, 2021, pp. 829–838.
115. Wu, Y.; Ma, X.; Yang, D. Personalized response generation via generative split memory network. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021, pp. 1956–1970.
116. Zhang, W.; Shen, Y.; Ma, Y.; Cheng, X.; Tan, Z.; Nong, Q.; Lu, W. Multi-view reasoning: Consistent contrastive learning for math word problem. *arXiv preprint arXiv:2210.11694* **2022**.
117. Wu, Y.; Liu, J.; Bu, X.; Liu, J.; Zhou, Z.; Zhang, Y.; Zhang, C.; Bai, Z.; Chen, H.; Ge, T.; others. Conceptmath: A bilingual concept-wise benchmark for measuring mathematical reasoning of large language models. *arXiv preprint arXiv:2402.14660* **2024**.
118. Zhang, W.; Shen, Y.; Hou, G.; Wang, K.; Lu, W. Specialized Mathematical Solving by a Step-By-Step Expression Chain Generation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **2024**.
119. Zhang, J.; Wang, X.; Ren, W.; Jiang, L.; Wang, D.; Liu, K. RATT: AThought Structure for Coherent and Correct LLMReasoning. *arXiv preprint arXiv:2406.02746* **2024**.
120. Huang, J.; Chang, K.C.C. Towards reasoning in large language models: A survey. *arXiv preprint arXiv:2212.10403* **2022**.
121. Zhang, J.; Mo, F.; Wang, X.; Liu, K. Thought space explorer: Navigating and expanding thought space for large language model reasoning. *arXiv preprint arXiv:2410.24155* **2024**.
122. Huang, X.; Liu, W.; Chen, X.; Wang, X.; Wang, H.; Lian, D.; Wang, Y.; Tang, R.; Chen, E. Understanding the planning of LLM agents: A survey. *arXiv preprint arXiv:2402.02716* **2024**.
123. Lu, P.; Gong, R.; Jiang, S.; Qiu, L.; Huang, S.; Liang, X.; Zhu, S.C. Inter-GPS: Interpretable geometry problem solving with formal language and symbolic reasoning. *arXiv preprint arXiv:2105.04165* **2021**.
124. Besold, T.R.; d'Avila Garcez, A.; Bader, S.; Bowman, H.; Domingos, P.; Hitzler, P.; Kühnberger, K.U.; Lamb, L.C.; Lima, P.M.V.; de Penning, L.; others. Neural-symbolic learning and reasoning: A survey and interpretation 1. In *Neuro-Symbolic Artificial Intelligence: The State of the Art*; IOS press, 2021; pp. 1–51.

125. Pan, L.; Albalak, A.; Wang, X.; Wang, W.Y. Logic-lm: Empowering large language models with symbolic solvers for faithful logical reasoning. *arXiv preprint arXiv:2305.12295* **2023**.
126. Su, A.; Wang, A.; Ye, C.; Zhou, C.; Zhang, G.; Zhu, G.; Wang, H.; Xu, H.; Chen, H.; Li, H.; others. TableGPT2: A Large Multimodal Model with Tabular Data Integration. *arXiv preprint arXiv:2411.02059* **2024**.
127. Yin, P.; Neubig, G.; Yih, W.t.; Riedel, S. TaBERT: Pretraining for joint understanding of textual and tabular data. *arXiv preprint arXiv:2005.08314* **2020**.
128. Huang, X.; Khetan, A.; Cvitkovic, M.; Karnin, Z. Tabtransformer: Tabular data modeling using contextual embeddings. *arXiv preprint arXiv:2012.06678* **2020**.
129. Zhang, X.; Zhang, J.; Rekabdar, B.; Zhou, Y.; Wang, P.; Liu, K. Dynamic and Adaptive Feature Generation with LLM. *arXiv preprint arXiv:2406.03505* **2024**.
130. Lu, P.; Peng, B.; Cheng, H.; Galley, M.; Chang, K.W.; Wu, Y.N.; Zhu, S.C.; Gao, J. Chameleon: Plug-and-play compositional reasoning with large language models. *Advances in Neural Information Processing Systems* **2024**, *36*.
131. Zhang, X.; Zhang, J.; Mo, F.; Chen, Y.; Liu, K. TIFG: Text-Informed Feature Generation with Large Language Models. *arXiv preprint arXiv:2406.11177* **2024**.
132. Hou, G.; Shen, Y.; Zhang, W.; Xue, W.; Lu, W. Enhancing Emotion Recognition in Conversation via Multi-view Feature Alignment and Memorization. Findings of the Association for Computational Linguistics: EMNLP 2023, 2023, pp. 12651–12663.
133. Dan, H.C.; Huang, Z.; Lu, B.; Li, M. Image-driven prediction system: Automatic extraction of aggregate gradation of pavement core samples integrating deep learning and interactive image processing framework. *Construction and Building Materials* **2024**, *453*, 139056.
134. Tao, C.; Fan, X.; Yang, Y. Harnessing LLMs for API Interactions: A Framework for Classification and Synthetic Data Generation. *arXiv preprint arXiv:2409.11703* **2024**.
135. Liu, J.; Zhou, Z.; Liu, J.; Bu, X.; Yang, C.; Zhong, H.S.; Ouyang, W. Iterative Length-Regularized Direct Preference Optimization: A Case Study on Improving 7B Language Models to GPT-4 Level. *arXiv preprint arXiv:2406.11817* **2024**.
136. Borisov, V.; Seßler, K.; Leemann, T.; Pawelczyk, M.; Kasneci, G. Language models are realistic tabular data generators. *arXiv preprint arXiv:2210.06280* **2022**.
137. Zhang, W.; Shen, Y.; Lu, W.; Zhuang, Y. Data-copilot: Bridging billions of data and humans with autonomous workflow. *arXiv preprint arXiv:2306.07209* **2023**.
138. Jones, E.; Palangi, H.; Simões, C.; Chandrasekaran, V.; Mukherjee, S.; Mitra, A.; Awadallah, A.; Kamar, E. Teaching language models to hallucinate less with synthetic tasks. *arXiv preprint arXiv:2310.06827* **2023**.
139. Peng, C.; Yang, X.; Chen, A.; Smith, K.E.; PourNejatian, N.; Costa, A.B.; Martin, C.; Flores, M.G.; Zhang, Y.; Magoc, T.; others. A study of generative large language model for medical research and healthcare. *NPJ digital medicine* **2023**, *6*, 210.
140. Yang, H.; Liu, X.Y.; Wang, C.D. Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031* **2023**.
141. Fan, X.; Tao, C.; Zhao, J. Advanced Stock Price Prediction with xLSTM-Based Models: Improving Long-Term Forecasting. *Preprints* **2024**.
142. Xi, Z.; Chen, W.; Guo, X.; He, W.; Ding, Y.; Hong, B.; Zhang, M.; Wang, J.; Jin, S.; Zhou, E.; others. The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864* **2023**.
143. Chang, C.; Peng, W.C.; Chen, T.F. Llm4ts: Two-stage fine-tuning for time-series forecasting with pre-trained llms. *arXiv preprint arXiv:2308.08469* **2023**.
144. Voigt, P.; Von dem Bussche, A. The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing* **2017**, *10*, 10–5555.
145. Sirur, S.; Nurse, J.R.; Webb, H. Are we there yet? Understanding the challenges faced in complying with the General Data Protection Regulation (GDPR). Proceedings of the 2nd international workshop on multimedia privacy and security, 2018, pp. 88–95.
146. Gostin, L.O.; Levit, L.A.; Nass, S.J. Beyond the HIPAA privacy rule: enhancing privacy, improving health through research **2009**.
147. Benitez, K.; Malin, B. Evaluating re-identification risks with respect to the HIPAA privacy rule. *Journal of the American Medical Informatics Association* **2010**, *17*, 169–177.