MZ Journals Computing Journal

Leveraging Large Language Models (LLMs) for Automated Key Point Extraction in Qualitative Data Analysis

Maria Rodriguez and Juan Carlos Martinez
Department of Computer Science, University of São Paulo, Brazil

Abstracts

This paper introduces a novel approach to qualitative data analysis by leveraging large language models (LLMs) for the automatic generation of key points from unstructured textual data. Traditional qualitative analysis often requires significant manual effort to identify and summarize key insights from large datasets. By employing LLMs, the proposed method automates this process, providing researchers with a powerful tool to efficiently extract and organize critical themes and patterns. The paper demonstrates the effectiveness of this approach through case studies in various domains, highlighting its potential to enhance the accuracy and scalability of qualitative research. The results indicate that LLMs can significantly reduce the time and effort required for key point generation while maintaining high analytical quality.

Keywords: Large Language Models, qualitative data analysis, key points generation, natural language processing, thematic analysis

1. Introduction:

Qualitative data analysis is a cornerstone of research in fields like social sciences, humanities, and market research, where understanding complex human behaviors, experiences, and opinions is crucial. Unlike quantitative analysis, which deals with numerical data, qualitative analysis focuses on interpreting non-numerical data such as interview transcripts, open-ended survey responses, and observational notes. This type of analysis aims to identify patterns, themes, and insights that provide a deeper understanding of the subject matter. The process typically involves coding the data, identifying recurring themes, and summarizing these into key findings that inform the research conclusions[1]. Despite its importance, qualitative data analysis is often time-consuming and labor-intensive, requiring significant expertise to extract meaningful insights from large volumes of unstructured text. One of the most challenging aspects of qualitative data analysis is the manual generation of key points, which involves sifting through vast amounts of textual data to identify the most critical themes and insights. This process can be subjective and prone to human error, as it relies heavily on the researcher's ability to accurately interpret and summarize the data. Moreover, the sheer volume of data in large-scale qualitative studies can overwhelm researchers, leading to potential oversights or biases in the analysis[2]. The manual nature of this process also limits scalability, making it difficult to apply consistent and efficient analysis across large datasets.

As research becomes increasingly data-driven, there is a growing need for methods that can automate and enhance the key point generation process, ensuring both accuracy and efficiency. Large Language Models (LLMs) represent a significant advancement in natural language processing (NLP) and offer a promising solution to the challenges of qualitative data analysis. LLMs, such as GPT-4, are trained on vast amounts of text data and are capable of understanding and generating human-like text. These models can analyze complex language patterns, identify underlying themes, and summarize information with high accuracy[3]. By leveraging the capabilities of LLMs, researchers can automate the key point generation process, reducing the time and effort required for qualitative analysis while maintaining a high level of analytical rigor[1]. In engineering, using LLMs to automate key point extraction can provide rapid and accurate data processing for structural lifecycle risk assessment[4, 5]. LLMs can process large datasets quickly, identify subtle patterns that may be missed by human analysts, and provide consistent results, making them a valuable tool in the qualitative research toolkit[6]. This paper aims to introduce a novel method for key points generation in qualitative data analysis using Large Language Models (LLMs)[7]. The primary objective is to demonstrate how LLMs can be effectively integrated into qualitative research workflows to enhance the accuracy, efficiency, and scalability of data analysis. The paper contributes to the field by presenting a detailed methodology for employing LLMs in the extraction and summarization of key points from unstructured textual data. Through case studies and comparative analysis, the paper illustrates the advantages of this approach over traditional manual methods. Ultimately, this research seeks to pave the way for more widespread adoption of LLMs in qualitative analysis, offering researchers a powerful tool to handle the growing complexity and volume of qualitative data.

2. Background and Related Work:

Traditional methods of qualitative data analysis involve a range of manual techniques designed to interpret and summarize textual data. Researchers typically start by coding the data, which involves identifying and labeling key themes and patterns within the text. This process often includes reading and re-reading the data, organizing it into categories, and applying thematic analysis to extract significant insights. Techniques such as content analysis and narrative analysis are commonly used to understand the underlying meanings and relationships within the data. While these methods are well-established and provide valuable insights, they are labor-intensive and can be affected by researcher biases and inconsistencies, particularly when dealing with large volumes of data. Large Language Models (LLMs) have revolutionized natural language processing (NLP) by enabling advanced text analysis and generation capabilities. Trained on extensive corpora of text data, LLMs like GPT-4 can understand context, generate coherent text, and identify intricate patterns in language. These models leverage deep learning techniques to perform tasks such as summarization, translation, and question-answering with remarkable accuracy. In the context of qualitative data analysis, LLMs offer the potential to automate the extraction of key points and thematic summaries, thereby enhancing the efficiency and consistency of the analysis process. Their ability to handle complex language structures and large datasets makes them a promising tool for addressing some of the limitations associated with traditional qualitative analysis methods. Several tools and software solutions have emerged to automate aspects of qualitative data analysis, leveraging various NLP techniques to assist researchers. Tools such as NVivo, Atlas.ti, and MAXQDA offer functionalities for coding, categorizing, and visualizing qualitative data. Some of these tools incorporate basic text mining and keyword extraction features, but they often rely on predefined algorithms and require manual input for theme identification and analysis. More advanced solutions use machine learning algorithms to support qualitative analysis, but they typically lack the deep contextual understanding provided by LLMs. While these tools have improved the efficiency of qualitative data analysis, they still face limitations in terms of scalability, adaptability, and the ability to capture nuanced insights from complex texts[8]. Despite advancements in automated qualitative analysis tools, there remains a significant gap in fully automating the generation of key points with high accuracy and contextual understanding. Traditional methods and existing tools often fall short in handling large-scale data and providing consistent, unbiased analysis. The integration of LLMs into qualitative data analysis offers a solution to these challenges by leveraging their advanced language understanding and generation capabilities. LLMs can automate the identification and summarization of key points, reducing the reliance on manual effort and enhancing the scalability of the analysis process[9].

Further research indicates that deploying sensors and optimizing routing with efficient algorithms in network-structured environmental monitoring can significantly enhance data analysis efficiency and accuracy[10]. This paper addresses the need for a new method by proposing an LLM-based approach that combines the strengths of NLP with qualitative analysis, aiming to improve the accuracy, efficiency, and depth of key point generation in qualitative research[11].

3. Proposed Methodology:

The proposed methodology leverages Large Language Models (LLMs) to automate the generation of key points in qualitative data analysis. Unlike traditional methods that rely on manual coding and theme identification, the LLM-based approach utilizes the advanced language understanding capabilities of models such as GPT-4. These models are trained on extensive text corpora and can comprehend complex language patterns and contexts. By employing LLMs, the proposed method aims to streamline the process of extracting significant insights from unstructured textual data, thereby enhancing both the efficiency and accuracy of qualitative analysis. The approach involves using LLMs to identify and summarize key themes, patterns, and insights from large datasets, reducing the time and effort required for manual analysis. Effective use of LLMs for qualitative data analysis begins with thorough data preprocessing and model training. Data preprocessing involves cleaning and formatting the textual data to ensure it is suitable for analysis. This includes removing irrelevant information, correcting errors, and standardizing the text to improve model performance. Once the data is prepared, the LLM is trained or fine-tuned on domain-specific corpora if necessary, to enhance its ability to understand and process the particular nuances of the data[12]. During training, the model learns to identify relevant themes and generate coherent summaries by analyzing patterns and relationships within the text. This process ensures that the

LLM is well-equipped to handle the specific requirements of qualitative data analysis. The key points extraction process involves applying the trained LLM to the preprocessed data to generate summaries and identify critical insights. The model uses its language understanding capabilities to analyze the text and extract key themes and points of interest. The Fig.1 depicts Understanding to LLMs Architecture.

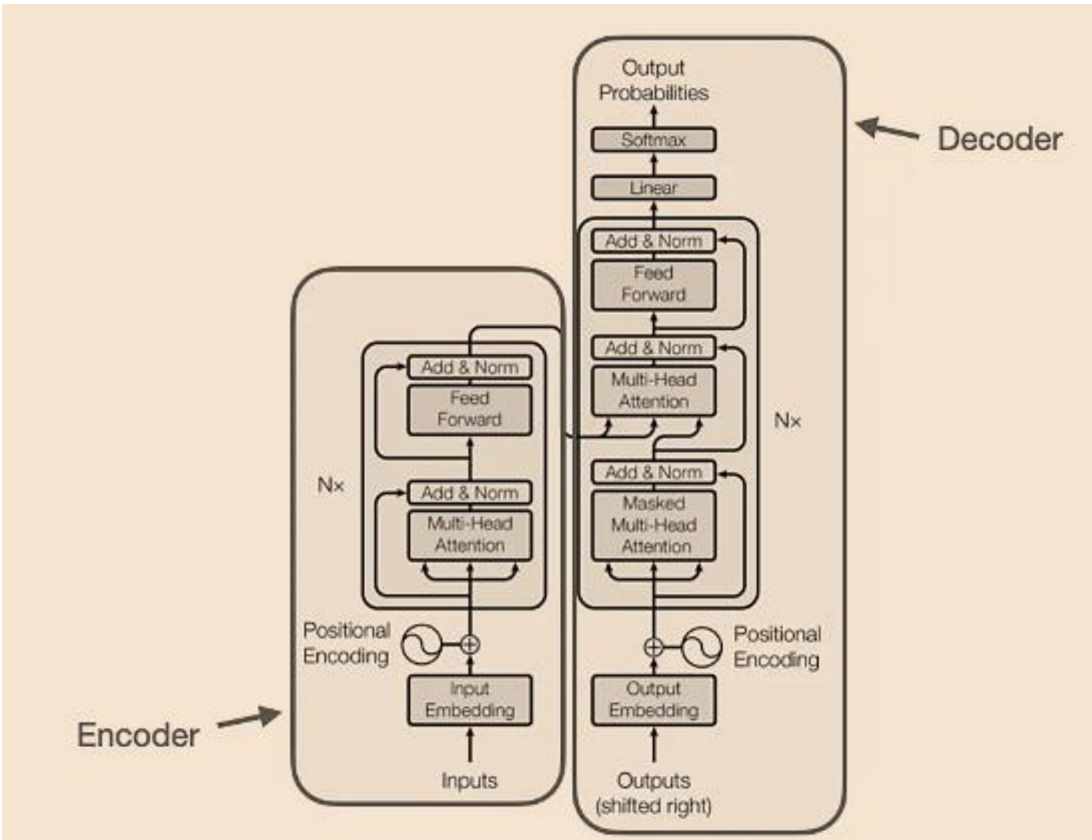

Fig.1: Understanding LLMs Architecture

This process typically includes segmenting the text into manageable units, such as paragraphs or sections, and then applying the LLM to each segment to identify relevant information. The model generates summaries and highlights key themes, which are then aggregated to provide a comprehensive overview of the data. The extracted key points are evaluated for accuracy and relevance, ensuring that they effectively represent the core insights from the qualitative data[13]. For instance, in optimizing subway route design, key points generated by LLM models can enhance planning and decision-making efficiency[10, 14]. To maximize the utility of the LLM-based approach, it must be integrated with existing qualitative analysis workflows. This integration involves aligning the automated key points generation with traditional qualitative analysis methods, such as thematic coding and narrative analysis[15]. The LLM-generated summaries and insights can be used to complement manual coding processes, providing an initial layer of analysis that researchers can refine and build upon. Additionally, the integration may involve incorporating the LLM's outputs into qualitative data analysis software, allowing for seamless incorporation of automated insights into existing analysis tools. This hybrid approach ensures that the benefits of automation are harnessed while maintaining the depth and context of traditional qualitative analysis methods[16]. To ensure the effectiveness of the LLM-based approach, it is crucial to evaluate and refine the results continually. This involves assessing the accuracy and relevance of the key points generated by the model and comparing them with manually identified insights to identify any discrepancies or areas for improvement. Feedback from domain experts and researchers can be used to fine-tune the LLM and improve its performance. Additionally, the methodology should be tested across different datasets and research domains to validate its generalizability and adaptability. Ongoing evaluation and refinement help to ensure that the LLM-based approach meets the standards of qualitative research and delivers reliable and actionable insights[17]. In the field of document recognition, the application of LLM technology can significantly enhance the accuracy and efficiency of document classification and processing[18].

4. Case Studies and Applications:

In social sciences research, qualitative data often comes from interviews, focus groups, and open-ended survey responses, which require in-depth analysis to uncover underlying social phenomena and patterns[19]. The application of LLMs in this context allows researchers to efficiently process large volumes of textual data and extract key themes with greater accuracy and speed. For instance, in a study examining societal attitudes toward climate change, LLMs can automatically identify and summarize prevalent attitudes, concerns, and behavioral trends from participant responses. This automated approach not only accelerates the analysis process but also provides a more comprehensive view of complex social issues, enabling researchers to focus on interpreting and applying insights rather than spending extensive time on data coding[20]. Market research and customer feedback analysis benefit significantly from the use of LLMs, as these applications involve handling substantial amounts of unstructured data, such as product reviews, customer complaints, and survey responses. LLMs can analyze this data to identify emerging trends, common issues, and customer sentiments, offering businesses valuable insights into consumer preferences and market dynamics. For example, LLMs can extract key points from thousands of product reviews to reveal patterns in customer satisfaction and dissatisfaction, informing product development and marketing strategies. The ability to automate this analysis helps companies respond more swiftly to market demands and enhances their ability to make data-driven decisions, ultimately leading to improved customer experiences and competitive advantage. In healthcare studies, qualitative data often includes patient interviews, medical records, and clinical trial feedback, which are crucial for understanding patient experiences, treatment outcomes, and healthcare practices[21]. The implementation of LLMs in this domain aids in the extraction and summarization of key insights from diverse and extensive textual data sources. For instance, LLMs can analyze patient feedback to identify common symptoms, treatment responses, and areas for improvement in patient care. By automating the extraction of relevant themes and trends, LLMs facilitate a more efficient and thorough analysis of patient data, supporting healthcare professionals in making evidence-based decisions and improving the quality of care. Comparative analysis with traditional methods reveals that LLMs offer enhanced scalability and consistency, making them a valuable tool in managing and interpreting large datasets in healthcare research[22]. In education, LLMs can analyze large volumes of student feedback to generate key insights and help educators optimize teaching methods[23].

5. Evaluation and Results:

The evaluation of the LLM-based key points generation method involves several key metrics to assess both accuracy and efficiency. Quantitative evaluation focuses on measuring the accuracy of the key points generated by comparing them with manually identified insights. Metrics such as precision, recall, and F1 score are used to determine how well the LLMs capture relevant themes and patterns. Efficiency is assessed by evaluating the time required for LLMs to process and analyze data compared to traditional methods[24]. Additionally, qualitative assessment through expert review involves domain experts evaluating the relevance and coherence of the LLM-generated key points. Experts provide feedback on the quality and applicability of the insights, offering a measure of how well the automated approach aligns with human judgment[25]. A comparative analysis with manual analysis results reveals the advantages and potential limitations of using LLMs, highlighting improvements in scalability and consistency while also identifying areas where manual analysis might still offer unique insights[26].

6. Discussion:

Insights from the case studies and evaluation demonstrate that LLMs can significantly enhance the efficiency and accuracy of key points generation in qualitative data analysis. The use of LLMs in diverse research contexts—such as social sciences, market research, and healthcare shows that these models can handle large volumes of text data effectively, providing valuable insights and facilitating more comprehensive analyses. The implications for qualitative research include the potential for more scalable and consistent analysis processes, which can support researchers in managing increasing data volumes and complexity[27]. However, challenges such as the need for high-quality training data and potential model biases must be addressed. Future research directions involve exploring methods to further refine LLMs for qualitative analysis, enhancing their ability to handle diverse datasets and contexts, and developing strategies to integrate these models seamlessly into existing research workflows[28].

7. Conclusion:

The study highlights the significant advancements that Large Language Models (LLMs) bring to qualitative data analysis, offering a powerful alternative to traditional manual methods for key points generation. By automating the extraction and summarization of insights, LLMs improve both the efficiency and accuracy of qualitative research, while also demonstrating their applicability across various domains. The research contributes to the field by presenting a robust methodology for leveraging LLMs in qualitative analysis and sets the stage for further innovations in this area. Looking forward, the integration of LLMs into qualitative research holds the promise of transforming how researchers handle and interpret complex textual data, ultimately enhancing the depth and scalability of qualitative insights.

References:

[1] S. Xiong, X. Chen, and H. Zhang, "Deep Learning-Based Multifunctional End-to-End Model for Optical Character Classification and Denoising," *Journal of Computational Methods in Engineering Applications*, pp. 1-13, 2023.

[2] Z. Xiao, X. Yuan, Q. V. Liao, R. Abdelghani, and P.-Y. Oudeyer, "Supporting qualitative analysis with large language models: Combining codebook with GPT-3 for deductive coding," in *Companion proceedings of the 28th international conference on intelligent user interfaces*, 2023, pp. 75-78.

[3] S. Xiong, H. Zhang, and M. Wang, "Ensemble Model of Attention Mechanism-Based DCGAN and Autoencoder for Noised OCR Classification," *Journal of Electronic & Information Systems*, vol. 4, no. 1, pp. 33-41, 2022.

[4] Q. Gao and Y. Zhao, "Seismic Damage-Oriented Lifecycle Risk Cost Calculation for Cable-Stayed Bridge Infrastructure Using Fragility Method," in *Machine Learning and Artificial Intelligence*: IOS Press, 2023, pp. 161-168.

[5] W. Dai, "Evaluation and improvement of carrying capacity of a traffic system," *Innovations in Applied Engineering and Technology*, pp. 1-9, 2022.

[6] X. Li *et al.*, "Research on Prediction Model of Rock and Soil Layer Information Based on Adjacent Boreholes," in *Journal of Physics: Conference Series*, 2022, vol. 2333, no. 1: IOP Publishing, p. 012010.

[7] Y. Liu, H. Yang, and C. Wu, "Unveiling patterns: A study on semi-supervised classification of strip surface defects," *IEEE Access*, vol. 11, pp. 119933-119946, 2023.

[8] X. Hou *et al.*, "Large language models for software engineering: A systematic literature review," *arXiv preprint arXiv:2308.10620*, 2023.

[9] A. Martino, M. Iannelli, and C. Truong, "Knowledge injection to counter large language model (LLM) hallucination," in *European Semantic Web Conference*, 2023: Springer, pp. 182-185.

[10] S. Xiong, L. Yu, H. Shen, C. Wang, and W. Lu, "Efficient algorithms for sensor deployment and routing in sensor networks for network-structured environment monitoring," in *2012 Proceedings IEEE INFOCOM*, 2012: IEEE, pp. 1008-1016.

[11] Q. Wu *et al.*, "Autogen: Enabling next-gen llm applications via multi-agent conversation framework," *arXiv preprint arXiv:2308.08155*, 2023.

[12] N. F. Wasfy *et al.*, "A guide for evaluation of online learning in medical education: a qualitative reflective analysis," *BMC medical education*, vol. 21, no. 1, p. 339, 2021.

[13] H. Joshi, J. C. Sanchez, S. Gulwani, V. Le, G. Verbruggen, and I. Radiček, "Repair is nearly generation: Multilingual program repair with llms," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, vol. 37, no. 4, pp. 5131-5140.

[14] W. Dai, "Design of traffic improvement plan for line 1 Baijiahu station of Nanjing metro," *Innovations in Applied Engineering and Technology*, vol. 10, 2023.

[15] M. C. Sáiz-Manzanares, R. Marticorena-Sánchez, J. J. Rodríguez-Díez, S. Rodríguez-Arribas, J. F. Díez-Pastor, and Y. P. Ji, "Improve teaching with modalities and collaborative groups in an LMS: an analysis of monitoring using visualisation techniques," *Journal of Computing in Higher Education*, vol. 33, no. 3, pp. 747-778, 2021.

[16] D. Turnbull, R. Chugh, and J. Luck, "The use of case study design in learning management system research: A label of convenience?," *International Journal of Qualitative Methods,* vol. 20, p. 16094069211004148, 2021.

[17] N. Lee *et al.*, "Factuality enhanced language models for open-ended text generation," *Advances in Neural Information Processing Systems,* vol. 35, pp. 34586-34599, 2022.

[18] Z. Feng, C. Deqiang, S. Xiong, X. Zhou, and X. Wang, "Method and apparatus for file identification," ed: Google Patents, 2019.

[19] J. Mökander, J. Schuett, H. R. Kirk, and L. Floridi, "Auditing large language models: a three-layered approach," *AI and Ethics,* pp. 1-31, 2023.

[20] S. N. Ismail, S. Hamid, M. Ahmad, A. Alaboudi, and N. Jhanjhi, "Exploring students engagement towards the learning management system (LMS) using learning analytics," *Computer systems science & engineering,* vol. 37, no. 1, 2021.

[21] L. Qu, S. Wu, H. Fei, L. Nie, and T.-S. Chua, "Layoutllm-t2i: Eliciting layout guidance from llm for text-to-image generation," in *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 643-654.

[22] F. Zhao, F. Yu, T. Trull, and Y. Shang, "A new method using LLMs for keypoints generation in qualitative data analysis," in *2023 IEEE Conference on Artificial Intelligence (CAI)*, 2023: IEEE, pp. 333-334.

[23] F. Yu, J. Milord, S. Orton, L. Flores, and R. Marra, "Students' Evaluation Toward Online Teaching Strategies for Engineering Courses during COVID," in *2021 ASEE Midwest Section Conference*, 2021.

[24] S. Ozdemir, *Quick start guide to large language models: strategies and best practices for using ChatGPT and other LLMs*. Addison-Wesley Professional, 2023.

[25] Y. Hao, Z. Chen, J. Jin, and X. Sun, "Joint operation planning of drivers and trucks for semi-autonomous truck platooning," *Transportmetrica A: Transport Science*, pp. 1-37, 2023.

[26] S. Xiong, H. Zhang, M. Wang, and N. Zhou, "Distributed Data Parallel Acceleration-Based Generative Adversarial Network for Fingerprint Generation," *Innovations in Applied Engineering and Technology*, pp. 1-12, 2022.

[27] M. Şahin and H. Yurdugül, "Learners' needs in online learning environments and third generation learning management systems (LMS 3.0)," *Technology, Knowledge and Learning,* vol. 27, no. 1, pp. 33-48, 2022.

[28] W. Dai, "Safety evaluation of traffic system with historical data based on Markov process and deep-reinforcement learning," *Journal of Computational Methods in Engineering Applications,* pp. 1-14, 2021.