Enabling HR Data-Driven Decision Making Through a Natural Language-Based Chatbot for Dynamic Data Visualization

W. Shachini Kavindi Perera

(Reg. No.: MS23018600)

A THESIS

SUBMITTED TO

SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

December 2024

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Samantha Thelijjagoda

Approved for MSc. Research Project:

MSc. Programme Co-ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

Sign: …………Shaghini………..

W.S.K. Perera

Date: ……11-11-2024…….

ABSTRACT

Enabling HR Data-Driven Decision Making Through a Natural Language-Based Chatbot for Dynamic Data Visualization

W.S.K. Perera
MSc. in Information Technology
Supervisor: Prof. Samantha Thelijjagoda
December 2024

This research develops a natural language-based chatbot designed to enhance HR decision-making through dynamic data visualization. Traditional HR data analysis methods are often slow and complex, limiting real-time insights. By integrating natural language processing (NLP) with advanced data visualization, this chatbot enables HR professionals to interact with large datasets using conversational queries, streamlining data access and interpretation.

The system translates user input into SQL queries, retrieving data from a data warehouse and presenting it in interactive visualizations. Testing revealed that the chatbot performs well with basic queries, providing accurate results and clear visualizations. However, challenges emerged with more complex queries and multi-layered data visualizations, where accuracy and response time decreased. Despite these challenges, the chatbot's decision-support capabilities were effective, offering actionable recommendations based on trends and patterns in HR data.

While the current system is limited to basic HR tasks, it demonstrates the potential of AI-driven tools in transforming HR processes. Future work will focus on improving query handling, enhancing visualization capabilities, and integrating the system with dashboards for strategic decision-making. Overall, this research contributes to the growing field of AI in HR, showing how NLP can simplify data access and support more informed decision-making.

ACKNOWLEDGEMENT

First and foremost, I would like to express my heartfelt gratitude to my supervisor, whose guidance, encouragement, and expertise have been invaluable throughout the course of this research. His insightful feedback, constructive critiques, and unwavering support have greatly enriched this study and have been fundamental in shaping the quality of my work. I deeply appreciate his patience and dedication in guiding me through every stage of this journey. I would also like to extend my profound gratitude to my husband, whose support has been both a technical and emotional cornerstone throughout this research. His encouragement, patience, and technical expertise have provided me with the motivation and resources I needed to overcome challenges and make consistent progress. His faith in my abilities and his willingness to assist with technical aspects have been invaluable. Special thanks go to my family members and friends, who have provided me with continued encouragement and support. Their understanding and moral support have been a source of strength, especially during challenging times. They have been my cheerleaders, and their belief in my work has kept me motivated to achieve my goals. To everyone who has contributed to this journey, directly or indirectly, I am immensely grateful. This accomplishment is as much yours as it is mine. Thank you for being a part of this journey.

.

TABLE OF CONTENTS

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Background

Human Resources (HR) has become a pivotal department in shaping the strategic direction of organizations. In the past, HR's role was largely administrative, focused on tasks such as recruitment, payroll, and employee management. However, as businesses evolved, the importance of aligning workforce management with organizational strategy became more apparent. Today, HR departments are expected to play a key role in driving business growth, ensuring workforce alignment with corporate goals, and supporting informed decision-making at the executive level.

One of the most critical components of HR's expanded role is its ability to harness and interpret data. Over the last decade, there has been a significant shift in how organizations view HR data. Historically, HR departments concentrated on collecting and maintaining data related to employees, such as attendance records, performance reviews, and recruitment statistics. While this information was valuable, it was often underutilized. The focus was primarily on keeping records rather than deriving insights from them. In modern business contexts, the expectation is no longer just to collect data but to use it strategically. HR professionals are tasked with analyzing this wealth of data to inform workforce decisions, predict future needs, and drive improvements in employee engagement and performance.

This shift from data collection to data utilization is crucial because it represents a change in mindset for HR departments. Data-driven decision-making allows organizations to become more proactive, making informed decisions based on trends and patterns rather than reacting to situations as they arise. For example, instead of waiting for employee turnover to become a major issue, HR departments can analyze historical data to predict when employees are likely to leave and take steps to retain them. Similarly, workforce analytics can be used to identify skill gaps within the organization, helping HR to align training programs and recruitment efforts with long-term business objectives.

However, the process of analyzing HR data is not without its challenges. While businesses now generate more HR data than ever before, extracting meaningful insights from this data is often a cumbersome and complex process. Traditional HR data analysis techniques are slow and inefficient. For instance, manual data analysis often involves poring over spreadsheets and reports, which can be time-consuming and prone to errors. Additionally, data stored across different systems or in different formats makes it difficult to create a cohesive picture. This lack of integration further hinders the ability to perform timely analysis.

Accessibility is another significant challenge. While HR departments have access to vast amounts of data, this data is often siloed within systems that are difficult for non-technical users to access or interpret. The complexity of traditional data analysis tools requires specialized knowledge, limiting data analysis to a small group of experts. This creates a barrier for HR professionals who may not have technical expertise but need timely access to insights to make informed decisions.

1.2 Problem Statement

Traditional methods of HR data analysis are increasingly inadequate in addressing the growing complexity and demands of modern organizations. These conventional tools and techniques are often slow, rigid, and limited in their capacity to provide actionable insights. HR professionals frequently rely on manual processes, such as sifting through spreadsheets or generating static reports, which consume valuable time and offer limited flexibility. This reliance on outdated approaches creates a significant barrier to proactive and informed decision-making.

One of the major limitations of traditional HR data analysis is its inability to provide real- time insights. In fast-paced business environments, decisions must be made quickly and based on current information. However, traditional methods often involve significant delays in data collection, processing, and analysis. By the time reports are generated, the information may already be outdated, reducing its relevance to ongoing HR challenges such as employee turnover, performance management, and workforce planning.

Additionally, existing tools require a high level of technical expertise, restricting their use to specialized analysts or IT departments. This siloed approach means that HR professionals without advanced technical skills are unable to access the data they need in a timely manner. As a result, many HR departments struggle to fully leverage their data to drive strategic initiatives. Instead of being empowered to make data-driven decisions, HR teams remain dependent on limited, static insights that hinder their ability to respond to emerging issues in the workforce.

The need for more accessible and real-time data-driven decision-making tools in HR is clear. Organizations require systems that allow HR professionals to engage with data in an intuitive and flexible manner, enabling them to extract insights and make decisions quickly. A tool that integrates advanced technologies such as Natural Language Processing (NLP) and dynamic data visualization can break down these barriers. By simplifying the data analysis process and providing real-time insights, such tools would empower HR departments to act proactively and strategically, improving workforce management and overall organizational performance.

The challenge, therefore, lies in creating a solution that overcomes the limitations of traditional methods and provides HR professionals with the tools they need to make informed, timely decisions based on accurate, up-to-date data.

1.3 Research Objectives

The aim of this research is to develop a natural language-based chatbot designed to facilitate dynamic data visualization, specifically tailored for HR data analysis. The goal is to create a tool that allows HR professionals to interact with complex datasets using simple, conversational queries. This chatbot will enable real-time data retrieval, visualization, and decision support, thereby improving the speed and accessibility of HR data-driven decision- making.

To achieve this aim, the research is guided by the following specific objectives:

1.3.1 Query Interpretation

Develop a robust Natural Language Processing (NLP) model that can accurately interpret HR-specific queries posed in natural language. The chatbot should be able to understand various user inputs, including complex queries involving multiple data points, and convert them into actionable insights.

1.3.2 SQL Query Generation

Create a system that translates natural language queries into structured SQL commands. These SQL queries will interact with a data warehouse to retrieve relevant HR data, ensuring the system can handle complex data retrieval tasks while maintaining accuracy and efficiency.

1.3.3 Data Visualization

Implement a dynamic data visualization framework that transforms the retrieved data into user-friendly visual formats. This objective includes the creation of charts, graphs, and dashboards that allow HR professionals to easily interpret and analyze the data, helping them identify trends and make informed decisions.

1.3.4 Decision Support

Integrate decision-support features within the system, providing HR managers with actionable recommendations based on the visualized data. This includes the ability to highlight trends, predict outcomes, and offer data-driven suggestions to assist in strategic workforce management.

By addressing these objectives, the research aims to develop a comprehensive tool that simplifies HR data analysis and enhances decision-making capabilities. The chatbot will serve as a bridge between HR professionals and complex data, making it easier for users to gain insights and act on information in real-time.

1.4 Research Significance

The development of a natural language-based chatbot for dynamic data visualization holds significant potential for transforming how HR professionals' access and use data. One of the key contributions of this research is the democratization of data access in HR departments. Traditionally, accessing and analyzing HR data requires specialized knowledge, often relegating data analysis tasks to experts or IT personnel. This created bottlenecks, delaying insights and limiting decision-making to a small group of technical users. The introduction of an intuitive chatbot, however, changes this dynamic by allowing HR professionals of all skill levels to interact directly with data through natural language queries. This accessibility empowers a wider range of HR staff to leverage data in their decision-making processes, ensuring that insights are available where and when they are needed most.

By simplifying the interaction between HR professionals and data, the chatbot enables real- time, user-friendly data analysis. HR managers no longer need to rely on complex tools or external teams to access the information they require. Instead, they can obtain insights quickly and effortlessly by asking questions in everyday language, which the chatbot translates into actionable results. This shift reduces reliance on specialized technical skills and creates an environment where data-driven decision-making is more inclusive, efficient, and timely.

Beyond enhancing data access, this research has broader implications for the application of AI-based decision support systems in HR. Artificial Intelligence (AI) is increasingly being used in various industries to improve decision-making processes, and HR is no exception. The chatbot in this research not only enables data visualization but also incorporates decision support, providing actionable insights based on the analyzed data. This capability allows HR professionals to move beyond simply understanding data and towards making informed, proactive decisions based on trends and patterns detected by the system. For example, the chatbot can identify emerging workforce issues, such as high turnover rates or declining employee engagement, and offer recommendations on how to address these challenges.

The integration of AI in HR decision-making brings multiple benefits, including improved efficiency, better prediction of workforce trends, and more strategic workforce planning. With AI-driven insights, HR departments can shift from reactive management to proactive strategies, addressing potential issues before they escalate. This can lead to improved employee retention, enhanced performance management, and a more agile workforce that aligns with organizational goals. Moreover, AI-based systems like the chatbot have the potential to significantly reduce the time and resources spent on manual data analysis, allowing HR professionals to focus on higher-level strategic tasks.

Overall, the significance of this research lies in its ability to bridge the gap between HR professionals and the data they need to make critical decisions. By democratizing data access and introducing AI-driven decision support, this chatbot has the potential to revolutionize HR management, making it more efficient, data-informed, and strategically aligned with broader organizational objectives.

1.5 Thesis Structure

This thesis is organized into six chapters, each focusing on a different aspect of the research process and findings. The structure ensures logical progression from the introduction of the research problem to the development and evaluation of the chatbot system, followed by the implications of the research.

1. Chapter 1: Introduction

This chapter provides an overview of the research topic, highlighting the role of HR in strategic decision-making and the importance of accessible data. It outlines the problem statement, research aims and objectives, and the significance of the study, and concludes with the structure of the thesis.

2. Chapter 2: Literature Review

The literature review examines existing research on HR data analysis, the use of artificial intelligence (AI) and natural language processing (NLP) in business applications, and current approaches to data visualization. This chapter identifies gaps in the literature, establishing the need for a system that democratizes HR data access and integrates decision support features.

3. Chapter 3: Methodology

This chapter describes the research design and methodology used to develop the chatbot system. It details the framework architecture, including the NLP model, SQL query generation, data retrieval, and visualization processes. The chapter also discusses challenges encountered, testing phases, and user feedback.

4. Chapter 4: Results

The results chapter presents the performance of the chatbot, focusing on query interpretation accuracy, data visualization outcomes, and decision-support capabilities. It includes user feedback, system performance metrics, and an evaluation of how well the chatbot meets the research objectives.

5. Chapter 5: Discussion

This chapter discusses the key findings from the results in relation to the research objectives. It compares the chatbot's performance with existing HR data analysis tools, outlines the system's practical implications, and addresses the challenges and limitations identified during testing. The chapter also suggests future improvements and potential areas of research.

6. Chapter 6: Conclusion

The final chapter summarizes the research, highlighting the contributions of the chatbot to HR data analysis and decision-making. It reflects on the significance of the findings, acknowledges the limitations of the current system, and offers recommendations for future development and research in AI-driven HR tools.

Chapter 2 Literature Review

2.1 Review of Existing Literature

The integration of natural language processing (NLP) and chatbot technologies into data visualization tools has garnered significant attention in recent years due to its potential to revolutionize data-driven decision-making processes. This section reviews key studies and research findings that contribute to the understanding and advancement of this field.

M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu [5] introduced an innovative analysis framework to enhance the accuracy of chart classification using deep convolutional neural networks (CNNs). By leveraging deep learning techniques, their research aimed to address the inherent challenges associated with chart classification, which is a critical aspect of data visualization. The proposed framework signifies a significant advancement in the field, offering a novel approach to improving the classification accuracy of charts, thereby enhancing the overall effectiveness of data visualization systems. Through the integration of CNNs, which are renowned for their ability to learn intricate features from data, Liu et al. aimed to overcome the limitations of traditional classification methods, enabling more robust and precise identification of chart types. This pioneering work underscores the growing interest in leveraging deep learning technologies to enhance the capabilities of data visualization tools, with implications for various domains where accurate chart classification is paramount for effective decision-making and data analysis.

S. D’Mello and A. Graesser [6] delve into the intricacies of affective states during complex learning, highlighting the importance of incorporating affective computing techniques into chatbot systems to enhance user engagement and improve learning outcomes. Their research underscores the critical role that emotions play in the learning process, emphasizing the need

for chatbot systems to be sensitive to users' affective states. By understanding and responding to users' emotional cues, chatbots can adapt their interactions to better support and motivate learners, ultimately enhancing their overall learning experience. D'Mello and Graesser's insights highlight the potential of affective computing in augmenting the capabilities of chatbot systems, paving the way for more personalized and effective educational interventions. This study contributes to the evolving field of conversational agents by emphasizing the importance of considering users' emotional well-being and engagement in the design and implementation of chatbot systems, with implications for various educational and interactive applications.

K. Majhadi and M. Machkour [18] extensively explored Natural Language Processing (NLP)-to-SQL systems in the field of database querying, aiming to bridge the gap between non-technical users and structured query languages. Traditional systems often rely on rule-based approaches, which employ predefined syntactic templates and domain-specific grammar to interpret user queries and generate corresponding SQL statements. For example, prior research has demonstrated the effectiveness of rule-based frameworks in static environments where database schemas are stable and the scope of queries is well-defined. These systems excel in precision when handling queries that align with predefined patterns but face significant limitations in adapting to diverse or complex queries, multi-intent inputs, or changing database schemas. Rule-based methods also require substantial manual effort to design, implement, and maintain their rules, rendering them less scalable and prone to errors in dynamic or large-scale applications.

In contrast, K. Majhadi and M. Machkour [18] proposed a machine learning-driven approach, leveraging advanced transformer models such as BART to automate the NLQ-to-SQL translation process. Unlike rule-based methods, which are constrained by predefined structures, the proposed system learns from diverse datasets, enabling it to generalize across schemas and handle a broader range of natural language inputs. This automation not only reduces the dependency on manual rule creation but also enhances scalability and adaptability. Furthermore, the integration of dynamic decision-making capabilities and real-time visualization offers a holistic solution for end users, providing actionable insights and personalized data representations beyond the capabilities of traditional rule-based systems. By combining automation, flexibility, and usability, the proposed system addresses key gaps in the existing literature, positioning itself as a significant advancement in the field of NLQ-to-SQL systems, particularly in domains requiring dynamic data visualization and decision support, such as HR analytics.

Dhingra, Li, Li, Gao, Chen, Ahmed, and Deng [7] introduced a novel end-to-end reinforcement learning approach tailored for dialogue agents in information access. Their research was dedicated to enhancing the conversational capabilities of chatbots by refining their ability to provide pertinent and precise information. By leveraging reinforcement learning techniques, their approach focused on enabling chatbots to engage in more natural and contextually relevant conversations with users, particularly in the domain of information retrieval. Through iterative learning processes guided by feedback signals, the chatbots could adapt and refine their responses over time, leading to more accurate and satisfactory interactions. Dhingra et al.'s work represents a significant advancement in the field of conversational AI, as it addresses the fundamental challenge of ensuring that chatbots can effectively meet users' information needs in real-time dialogue settings. Their research contributes to ongoing efforts to improve the quality and utility of chatbot interactions, with implications for various applications, including customer service, knowledge management, and educational assistance.

Recent advancements in text-to-SQL systems have demonstrated the effectiveness of large pre-trained language models, such as T5, in translating natural language queries into SQL statements. The work by Wong et al. [19] leverages the T5 model, fine-tuned on both the Spider dataset and custom datasets, to generate SQL queries for OLTP and data warehouse systems. Their approach incorporates schema corrections to enhance model performance and achieves an exact match accuracy of 72.9% for OLTP queries and 85.4% for data warehouse queries. This research highlights the potential of pre-trained models to achieve state-of-the-art results with iterative fine-tuning and schema-based post-processing.

While Wong et al.'s study [19] emphasizes model accuracy and the integration of schema linkages, it relies heavily on manual dataset creation and schema corrections to handle domain-specific queries. In contrast, the proposed system in this research builds upon automation, utilizing streamlined pipelines for query generation and dynamic data visualization. The inclusion of interactive charting capabilities and decision-making models in our system represents a step beyond the T5-based framework, aiming to bridge the gap between query translation and actionable insights. Furthermore, our methodology reduces manual intervention, focusing instead on user-friendly interactions and automated processes to support HR decision-making tasks. By doing so, our research broadens the applicability of text-to-SQL systems to real-time operational needs while maintaining scalability and adaptability.

A. Ramalingam, A. Karunamurthy, A. Dheeba, and R. Suruthi [8] provide a comprehensive exploration of chatbot technologies in their research paper, with a particular focus on their application in automated chart generation systems for data visualization. The study examines various techniques, applications, strengths, and limitations of chatbots in this context, offering valuable insights into their potential role in enhancing data visualization processes. The findings underscore the scalability, availability, and cost-effectiveness of chatbots in automating chart generation and facilitating data visualization tasks. However, a notable gap in the research lies in its limited focus on decision-making suggestions and the integration of disparate data sources. This presents an opportunity for further research to explore the development of a chatbot system that not only generates dynamic visualizations but also provides decision-making guidance based on user queries, thereby bridging the gap between data interpretation and informed decision-making.

M. Azmi, A. Mansour, and C. Azmi [9] explore the role of AI, particularly chatbots, in enhancing Business Intelligence (BI) systems, emphasizing their potential to streamline data analysis and decision-making processes. The paper delves into how AI-driven systems can significantly impact BI processes by improving data analysis, automating integration, and enabling predictive analytics. The methodology employed in the study involves analyzing perspectives on intelligent BI empowerment and presenting a flowchart illustrating how AI chatbots can assist BI processes for increased autonomy and creativity[2]. Findings suggest that AI-driven systems have the capacity to process extensive datasets, identify patterns and insights, automate data integration, and provide real-time analytics, among other advantages. However, a notable gap in the literature is the lack of a detailed comparison of various AI algorithms' effectiveness within BI systems or specific case studies illustrating chatbot implementation in BI. This gap presents an opportunity for further research to investigate and compare the efficacy of different AI algorithms in enhancing BI systems, particularly through the implementation of chatbots.

The dissertation by Bruno Pereira de Morais [10] presents a significant contribution to the field of conversational AI and data visualization automation. Focused on the development of the Ansa platform, the research addresses the challenge of effectively visualizing complex analytic answers from bots, aiming to enhance users' understanding of large datasets through automated visual representations. Utilizing a case-based reasoning (CBR) approach, the study demonstrates the system's ability to select appropriate visualizations autonomously and adapt its knowledge base based on user feedback. Results indicate high accuracy, performance, and adaptability of the system, promising insightful outcomes and ease of maintenance. However, a notable gap in literature is the lack of emphasis on integrating decision-making support alongside automated data visualization in the context of conversational AI. While the Ansa platform excels in generating visualizations, the integration of decision-making guidance within the chatbot interface remains relatively unexplored, representing an area for potential enhancement in the proposed research.

The foundational aspects of data science and its pivotal role in facilitating data-driven decision-making processes are delved into in the paper by F. Provost and T. Fawcett [11], although natural language processing or chatbot applications within the field are not specifically explored. Emphasizing the importance of understanding fundamental principles over tool-centric approaches, the research elucidates how data science serves as a bridge between data-processing technologies and decision-making realms. By delineating the systematic extraction of knowledge from data and presenting case studies to illustrate its application, the paper underscores the significance of data-analytic thinking in deriving actionable insights. However, a noticeable gap emerges concerning the integration of natural language processing and chatbots into the data science framework, representing an opportunity for further research to explore the synergistic potential of these technologies in enhancing data- driven decision making through dynamic data visualization and decision support functionalities.

The research on JourneyBot [12] presents a novel approach to design research by leveraging chatbot-driven interactive visualization tools to automate user interviews and derive actionable insights, particularly in the context of medical clinic visits. Utilizing a rule-based chatbot, the study adeptly collects feedback from users and generates interactive journey maps for analysis, offering designers a comprehensive understanding of user experiences. While the tool showcases advancements in data collection and visualization techniques, the study's emphasis on the visualization aspect of user journeys rather than the decision-making process itself represents a notable gap. Additionally, its focus on design research applications within specific contexts, such as medical visits, contrasts with the broader scope of enabling data-driven decision-making through natural language-based chatbots, which encompasses a wider range of domains and decision-making scenarios.

A notable contribution in the realm of data-driven decision-making is presented in the paper by R. Alaaeldin [13], which introduces a chatbot system specifically tailored to support managerial decision-making based on Big Data analytics. Employing design science methodology, the research aims to bridge the gap between decision-makers and analytics tools, offering a practical solution to facilitate informed decisions in organizational settings. By serving as an intermediary tool, the chatbot system enables managers to effectively translate analytics outputs into actionable insights aligned with business objectives. However, the identified gap emphasizes the need for more dynamic solutions encompassing a broader range of objectives and Key Performance Indicators (KPIs), alongside the integration of visualizations and multilingual support to enhance user experience. This gap highlights an opportunity for further research to explore the integration of natural language processing and chatbot functionalities within the context of dynamic data visualization, thereby enhancing the decision-making support capabilities of such systems.

G. K. Hoon, L. J. Yong, and G. K. Yang [14] introduce a compelling approach to enhance decision-making processes by implementing an analytics bot designed to streamline data retrieval and analytics queries through natural language interaction. By proposing a conversational framework that facilitates seamless communication between users and business intelligence systems, the research underscores the potential for improving efficiency of accessing and interpreting analytics data. However, a notable gap emerges in the absence of detailed empirical insights regarding the comparative effectiveness and user experience outcomes of the proposed analytics bot. While the technical framework is robustly outlined, further research is warranted to empirically validate the efficacy of the analytics bot in facilitating data-driven decision-making and to explore user perceptions and experiences with the system in real-world contexts. Closing this gap could offer valuable insights into the practical utility and effectiveness of natural language-based chatbots in augmenting decision-making processes through dynamic data visualization and decision support functionalities.

A. Islam and K. Chang [15] present an innovative Real-Time AI-Based Informational Decision-Making Support System (DMSS) that leverages machine learning and deep learning algorithms, such as LSTM and RF, to classify unstructured data from dynamic text sources, enhancing decision-making accuracy. Through meticulous methodology encompassing data collection, cleaning, sentiment analysis, and topic labeling, the study demonstrates remarkable classification accuracy rates, with LSTM achieving up to 99% accuracy in training and 96% in testing. While the paper acknowledges the significance of integrating intelligence into DMSS and highlights the robustness of its proposed system, it does not explicitly compare its findings with the referenced study on enabling data-driven decision- making through a natural language-based chatbot. However, it suggests that its DMSS may offer greater resilience to noisy datasets, indicating a potential area for comparison and further exploration. This research underscores the efficacy of advanced machine learning techniques in processing vast volumes of unstructured data for real-time decision support, laying the groundwork for future advancements in data-driven decision-making systems.

E. Dimara [16] delves into the pressing issue of unmet data visualization needs among organizational decision-makers, providing insights into the challenges they face and the desired features of visualization tools. Through a comprehensive methodology combining literature review, empirical survey, and interviews, the study uncovers the demand for innovative visualization tools that go beyond simple representations, emphasizing the necessity for trade-off overviews, scenario-based analysis, and collaboration support. However, a notable gap emerges when compared to the referenced chatbot research, as this study focuses on the broader landscape of data visualization tools rather than a specific technology like natural language-based chatbots. While both studies seek to enhance decision-making processes within organizations, this research highlights the importance of expanding visualization design to encompass tools tailored for managing complex, multi-objective problem spaces, thus offering complementary insights to the proposed chatbot approach.

The landscape of chatbot-based Natural Language Interfaces (V-NLIs) for data visualization is investigated in this review by Kavaz [17], highlighting the current state and challenges within this emerging field. Utilizing a systematic mapping approach, the authors analyze existing research through the lens of the data visualization pipeline and chatbot characteristics, uncovering prevalent themes such as limited guidance strategies and challenges in handling complex data and integrating with advanced systems like AR/VR. However, this review identifies several gaps, including the need for enhanced high-level query support, intelligent visual mapping techniques, and sophisticated interaction methods. While both studies explore the intersection of natural language processing and data visualization, this review highlights specific areas for improvement and innovation within chatbot-based V-NLIs, underscoring the potential for further advancements to enrich the data-driven decision- making process.

The application of Microsoft Power BI for HR data visualization is explored in the paper by M. D. Jadhav [20], with an emphasis on its capability to analyze large datasets and derive strategic insights. The study highlights the efficiency of rule-based tools like Power BI in creating interactive dashboards, providing static and pre-defined visualization templates for HR metrics such as employee performance and turnover rates. However, while Power BI excels in visual representation, it lacks automation and dynamic adaptability, requiring manual input for query generation and visualization selection. In contrast, this research introduces an automated, machine learning-driven approach to HR analytics. By leveraging BART for natural language-to-SQL translation and BERT for dynamic visualization prediction, the system addresses the limitations of manual intervention found in Power BI-based solutions. Unlike Power BI, which depends on predefined configurations, the proposed system dynamically generates SQL queries and selects visualization types based on the contextual understanding of natural language queries. Moreover, it integrates decision-support capabilities, using machine learning classifiers to provide actionable insights, further extending its utility beyond visualization.

Table 1 Summary of the literature review

Authors	Year	Contribution
Liu, S., Zhong, Z., and Zhu, L.	2016	Proposed an analysis framework for deep convolutional neural networks to enhance chart classification accuracy.
D'Mello and Graesser	2012	Observed affective states during complex learning and emphasized the integration of affective computing techniques into chatbots for improved user engagement and learning outcomes.
Dhingra, Li, Li, Gao, Chen, Ahmed, and Deng	2020	Introduced a reinforcement learning approach for dialogue agents in information access, focusing on enhancing the conversational capabilities of chatbots for improved information retrieval.
A. Ramalingam, A. Karunamurthy, A. Dheeba, and R. Suruthi	2023	Provided a comprehensive exploration of chatbot technologies, particularly focusing on their application in automated chart generation systems for data

		visualization. Identified gaps include the lack of decision-making suggestions and integration of disparate data sources.
M. Azmi, A. Mansour, and C. Azmi	2023	Explored the role of AI, particularly chatbots, in enhancing Business Intelligence (BI) systems, emphasizing the need for empirical comparison of AI algorithms' effectiveness within BI systems and case studies illustrating chatbot implementation in BI.
Bruno Pereira de Morais	2018	Developed the Ansa platform to enhance understanding of large datasets through automated visual representations. Noted the gap in integrating decision-making support alongside automated data visualization.
F. Provost and T. Fawcett	2013	Discussed the foundational aspects of data science and its role in facilitating data-driven decision-making processes, highlighting a gap in integrating natural

		language processing and chatbots into the data science framework.
S. Hwang and D. Kim	2023	Investigated the role of chatbot-driven interactive visualization tools in design research, particularly focusing on user journey mapping. Identified a gap in the emphasis on visualization over decision-making support.
R. Alaaeldin, E. Asfoura, G. Kassem, and M. S. Abdel-Haq	2021	Introduced a chatbot system tailored for supporting managerial decision-making based on Big Data analytics, noting the gap in the need for more dynamic solutions encompassing a broader range of objectives and KPIs.
G. K. Hoon, L. J. Yong, and G. K. Yang	2020	Presented an analytics bot designed to streamline data retrieval and analytics queries via natural language interaction, emphasizing the gap in empirical insights regarding user experience outcomes.

E. Dimara, H. Zhang, M. Tory, and S. Franconeri	2021	Explored the unmet data visualization needs of organizational decision-makers, focusing on the broader landscape of data visualization tools rather than natural language- based chatbots.
E. Kavaz, A. Puig, and I. Rodríguez	2023	Investigated the landscape of chatbot- based Natural Language Interfaces (V-NLIs) for data visualization, identifying gaps in query support and interaction methods.
Majhadi Khadija, Machkour Mustapha	2024	Proposed a model based on encoder-decoder architecture leveraging LSTM for NL-to-SQL translation, focusing on challenges in complex query generation across domains.
Albert Wong et al.	2023	Developed a NL-to-SQL system using the T5 model with exact match accuracy of 73% for OLTP and 85% for data warehouse queries, demonstrating practical deployment for utility services.

Monal D. Jadhav, Ashlesha B. Shelar, Assistant Prof. Sujeet More	2023	Explored the application of Microsoft Power BI for HR data visualization, highlighting its efficiency in analyzing large datasets and generating strategic insights through predefined templates.

2.2 Gap Identification

Based on the review of the literature provided, several gaps and opportunities for further research can be identified:

2.2.1 Integration of Decision-Making Support

While various studies explore the integration of natural language processing and chatbot technologies into data visualization tools, there is a notable gap in explicitly addressing decision-making support within these systems. Future research could focus on developing chatbot systems that not only generate visualizations but also offer decision-making guidance based on user queries, thereby bridging the gap between data interpretation and informed decision-making.

2.2.2 Comparative Effectiveness and User Experience

Many studies propose innovative approaches and frameworks for enhancing data-driven decision-making processes using chatbot technologies. However, there is a lack of detailed empirical insights regarding the comparative effectiveness and user experience outcomes of these systems. Future research could conduct comparative studies to evaluate the efficacy and user perceptions of different chatbot-based solutions in real-world contexts, providing valuable insights into their practical utility and effectiveness.

2.2.3 Enhanced Query Support and Interaction Methods

Several studies highlight the need for enhanced high-level query support, intelligent visual mapping techniques, and sophisticated interaction methods within chatbot-based Natural Language Interfaces (V-NLIs) for data visualization. Future research could focus on developing advanced query understanding mechanisms, intelligent visual representation techniques, and interactive methods to enhance user engagement and facilitate more effective data-driven decision-making processes.

2.2.4 Incorporating Intelligence and Handling Complex Data

Some studies emphasize the importance of integrating intelligence into chatbot systems and addressing challenges associated with handling complex data. Future research could explore advanced techniques, such as reinforcement learning and deep learning algorithms, to enhance the intelligence and adaptability of chatbot systems in processing complex data and providing accurate and timely decision support.

2.2.5 Expanding Visualization Design for Complex Problem Spaces

While many studies focus on improving data visualization tools, there is a gap in addressing the specific needs of managing complex, multi-objective problem spaces within organizations. Future research could explore novel visualization design approaches tailored for addressing complex organizational challenges, incorporating features such as trade-off overviews, scenario-based analysis, and collaboration support to facilitate more effective decision-making processes.

Table 2 Gap identification in existing literature

Research Question	Explanation
To what extent does the developed chatbot system effectively understand and respond to natural language queries related to data visualization?	Lack of detailed empirical insights regarding comparative effectiveness and user experience outcomes of chatbot-based solutions Need for enhanced high-level query support and intelligent visual mapping techniques within chatbot-based Natural Language Interfaces (V-NLIs) for data visualization
How do users perceive the usability and effectiveness of the chatbot system in facilitating data- driven decision- making processes?	Lack of detailed empirical insights regarding comparative effectiveness and user experience outcomes of chatbot-based solutions Need for enhanced high-level query support and intelligent visual mapping techniques

	within chatbot-based Natural Language Interfaces (V-NLIs) for data visualization
What is the impact of the chatbot system on decision-making efficiency, accuracy, and strategic outcomes within organizations?	Lack of integration of decision- making support within chatbot systems Challenges associated with handling complex data Need for more dynamic solutions encompassing a broader range of objectives and Key Performance Indicators (KPIs) within chatbot-based systems

Chapter 3 Methodology

This chapter provides a comprehensive exploration of the methodologies employed in designing and implementing a natural language-based chatbot aimed at enhancing data visualization within Human Resources (HR) data analysis. It begins by detailing the underlying research design, explaining the strategic choices that shaped the system's development to align with the goal of supporting data-driven HR decision-making. Following this, the chapter delves into the system's architectural framework, outlining each component's function, including the natural language processing (NLP) model, which is central to interpreting user queries and generating meaningful insights from HR datasets.

The methodology section furthers this by discussing the steps taken in incorporating the NLP model with an efficient data warehouse. This is to carry out actual data retrieval precise and relevant data-into actionable visual formats. It presents insight into the treatment and communication of data in thoughtful, user-friendly visualizations destined for HR business people by providing a detailed explanation of each stage, ranging from the initial retrieval to the final output in visualization. This will be elaborated in detail regarding each phase of development, from iterations of prototypes to the optimization of the best model, to show increments taken to arrive at increased accuracy of the system and improved user experience.

This chapter also provides an in-depth discussion of the various testing phases that went into the chatbot, both in the alpha and beta testing stages. These tests were beneficial for seeing how well the system performs, finding possible points of congestion, and fine tuning the system for several types of queries that the chatbot could go through. Insights from users during these stages helped improve the model's accuracy and response times and visualization clarity.

Lastly, the chapter also acknowledges some limitations of the methodology adopted, focusing on issues relating to the development of the system itself. This included but was not limited to query complexity, hierarchical queries, and response time optimization with datasets of large dimension. The conclusion of the chapter by discussing the limitations thus sets a background for potential future improvements in chatbot performance and continuous development of this innovative HR data visualization tool.

3.1 Research Design

3.1.1 Overview of the Research Approach

The focus of this research is to come up with a new, natural language-driven chatbot that will allow HR professionals to directly interact with HR data in the form of conversational queries and have real-time dynamic visualizations of data. This approach would enable HR teams to work easily with complex datasets and make data-driven decisions with intuitive and user-friendly tools. Because of the explained analysis, it was decided to follow a prototype development approach, which allows flexibility in integrating a variety of the most advanced technologies, starting from the understanding of the meaning of user queries with the aid of Natural Language Processing up to SQL generation and customized data visualization.

The architecture will be such that it can perform the translation of NLQ into structured SQL commands so that the chatbot can query the data warehouse efficiently. This is done through an NLP model that recognizes key entities and intents within the NLQ and translates those into matching SQL syntax to correctly draw out relevant data from the HR data warehouse with precision.

A second model was developed, too, able to predict the most appropriate type of visualization to use based on characteristics of the dataset and the intent of the NLQ. This model evaluates the structure of the dataset and its content with the user query to identify how data should be represented-bar charts, line graphs, scatter plots, heatmaps, or other formats. This automation in visualization selection enhances the usability for HR, as they can see data in formats optimized for clarity and insight.

Another model was built to assist decision-making in a manner such that insights from the retrieved data are generated. This model ensures pattern analysis, trends, and anomalies inside the data and will forward recommendations for action on these, attuned to the needs of HR. In the case of employee turnover analysis, for instance, recommendations of retention strategies based on high-turnover patterns may be provided by the system. This capability also makes the chatbot a decision-support system and not just a data access tool that will help HR managers make more informed, strategic decisions given the current data.

The approach to prototype development is iterative; each iteration yields real-world feedback in refining the NLP, visual selection, and decision-support models through user interactions and performance of the system. This research design combines NLP, automated visualization selection, and AI-driven decision support for a comprehensive tool in human resources analytics, bridging the gap from data access to strategic decision making.

3.1.2 Justification for Selecting a Prototype Approach

Several critical factors influenced the decision to adopt a prototype approach for developing this natural language-based HR chatbot. First and foremost, it is befitting to prove a concept that integrates such a wide range of technologies including but not limited to NLP, SQL databases, data visualization tools, and decision-support models. In this system, each element will play some role and depend on other components; the prototype enables us to test their compatibility and functionality in a controlled low-risk environment. This prototype would further help us in considering the system's usability, adaptability, and functionality by HRs who possess varying levels of technical proficiency in a real-life HR data context.

The prototype approach is especially useful in this project for iterating a product's test and refinement with real-world feedback. Human resources professionals interact with data differently according to the type of information needed; their input becomes key in fine tuning the system's NLP model, generation of SQL queries, and visualization selection. This allows us to study the needs and expectations of our users by testing and interaction with the prototype, whereby in this process, features and elements are refined to make the interface functional and intuitive.

Moreover, the prototype approach reduces the risk involved in such a scale of system development. This prototype gives us an opportunity to incrementally address the limitations, then implement improvements to this end, allowing resources to be committed efficiently and ensuring that this system evolves to realize maximum value-particularly for new uses or feedback highlighting potential areas of improvement.

It also fosters an environment where developers, data scientists, and HR professionals collaborate on the prototype approach and exchange knowledge. The collaborative environment allows for agile adjustments in the model, visualization methods, and decision- support capabilities, ensuring that the final solution is robust, scalable, and directly aligned with the evolved needs of HR professionals. It's all about the development of a working prototype that can serve as a flexible foundation for creating an advanced user-centered tool that unites in one NLP, data visualization, and decision-making support to make HR data analysis fast and available.

3.2 Framework Architecture

3.2.1 Detailed Description of the System's Architecture

The system architecture has been meticulously designed to facilitate the end-to-end process required for interpreting natural language queries, converting them into SQL commands, retrieving data from a data warehouse, and visualizing the results in dynamic, interactive formats. This architecture consists of multiple interdependent components, each optimized to handle specific tasks in the query-to-visualization pipeline. Together, these components create a seamless and efficient system tailored for HR professionals, enhancing their ability to gain insights from complex HR datasets using simple, conversational queries.

1. Natural Language Query Interpretation

The interface of the chatbot is natural, hence it forms a touch point of the user, allowing them to interact with the system using natural queries. In other words, when a user types a query like "Show employee turnover rates over the last year," an NLP model processes the text and determines the intent of a query together with extraction of relevant entities. This step of interpretation is very important since it converts the input in natural language provided by the user into structured information that may be further processed by the system. It specifies the exact named entities, such as "employee turnover" and "last year," or the intent behind the query, whether for historical data or performance metrics comparison. Such fine-tuning of the NLP model allows it to learn the nuances of HR specific speech and recognize terms and metrics unique to HR contexts, such as turnover, engagement scores, and questions about departments.

2. SQL Query Generation

Once the query has been interpreted by the NLP model, the system proceeds to SQL generation. The recognized entities are bound to the appropriate fields of the HR data warehouse, thereby generating an SQL command. This step embodies a multi-layer translation process, translating the natural language expressions into SQL syntax while making sure that the database fields and table relationships are correctly expressed. For instance, if the user asks to "Show performance metrics of employees in the sales department for the past three years", this system constructs a formatted SQL command, adding conditions about both department and time frame. The more complex queries, including multiple conditions in each, are parsed with care to construct efficient SQL commands, keeping query time to the minimum and database access must be optimized by striking a balance between performance and accuracy.

3. Data Retrieval

This SQL query is generated and then executed over the HR data warehouse, which stores structured HR data in a centralized repository such as employee records, turnover statistics, and performance scores. The architecture of the system should support this by enabling joins and indexing to optimize data retrieval from multiple tables; hence, processing complex queries with many data points runs smoothly. The schema has been designed with respect to the HR data warehouse to allow multiple dimensions toward retrieval; for example, employee engagement over time against turnover rates by department. For query performance optimization, indexing of frequently accessed tables is done, and common queries use caching mechanisms for reducing retrieval time and provide responsive experience to end-users.

4. Data Visualization

The final step, after data retrieval, would involve data transformation from retrieved format to appealing dynamic visualizations. The transformation will be user-friendly and in an intuitive format. The visualization engine would make use of libraries such as Plotly and D3.js, which provide interactive charts, graphs, and dashboards showing trends and insights of data clearly. This system architecture includes a model for the prediction of the best visualization type, given dataset characteristics and query intent. This would help to present the data in the most comprehensible form. An example might be that turnover trends could show up in a time-series line chart, while employee engagement scores for departments could be visualized as a heatmap or bar chart. These dashboards enable a user to interact- drill into data points or change parameters, therefore enhancing their capability for making informed decisions based on up-to-the-minute insights.

3.2.2 High-Level System Architecture

Figure 1 presents a high-level diagram of the system architecture, showcasing the main components and workflows of the natural language-based chatbot for HR data visualization. This architecture illustrates the end-to-end process from user input to data-driven decision- making, emphasizing the key modules and interactions involved in the system.

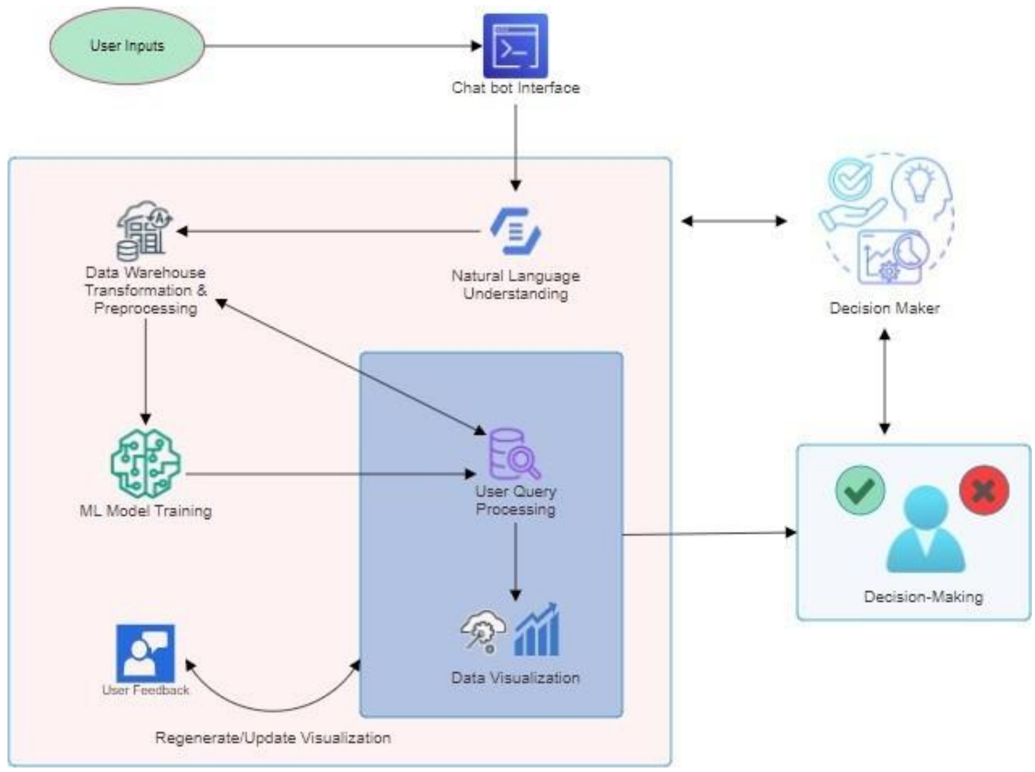

Figure 1 High-level diagram of the system architecture

Description of Components:

User Inputs: The system starts with user input, where HR professionals interact with the chatbot through natural language queries. These queries might involve requests for metrics like employee turnover, engagement scores, or performance data.

Chatbot Interface: The user inputs are processed through a chatbot interface that acts as the main point of interaction. This interface sends the input to the system's Natural Language Understanding (NLU) component.

Natural Language Understanding (NLU): The NLU module interprets the user's natural language query, extracting key entities, intents, and specific requirements. This component is crucial for transforming unstructured text input into a structured format that can be processed by the backend system.

User Query Processing: Once the NLU interprets the query, it is passed to the User Query Processing component. Here, the system translates the interpreted query into SQL commands or structured queries that access the HR data warehouse.

Data Warehouse Transformation & Preprocessing: The system connects with the HR data warehouse, which stores structured HR data like employee records, engagement scores, and performance metrics. Transformation and preprocessing are applied to retrieve the relevant data and ensure it is formatted correctly for visualization.

NLP Model Training: This component continuously improves the NLP model by incorporating feedback from user interactions. It enhances the system's ability to handle a wider range of HR-specific terminology and query patterns over time.

Data Visualization: The retrieved data is then transformed into visual representations such as bar charts, line graphs, and scatter plots. These visualizations are tailored to the type of data requested and aim to help HR professionals interpret trends and make informed decisions.

Decision-Making: Based on the data retrieved and visualized, the system provides decision-support insights, helping HR professionals take data-driven actions. Recommendations are generated by analyzing trends, such as identifying high turnover areas or low engagement departments.

User Feedback: Users can provide feedback on the accuracy and usefulness of the visualizations and recommendations. This feedback loop enables continuous improvements, as it helps refine both the NLP model and visualization processes, enhancing overall system performance.

Regenerate/Update Visualization: Based on user feedback or further clarifications, the system can regenerate or update visualizations, ensuring the displayed data meets the user's needs.

3.2.3 Challenges Encountered in Each Stage of Development

The development of the natural language-based chatbot system presented several challenges across its various stages. Each phase required unique problem-solving approaches to ensure that the system met the intended goals of usability, efficiency, and accuracy. The following outlines the primary challenges faced at each stage of development and the solutions implemented to address them.

1. Query Interpretation

The biggest challenges related to query interpretation were dealing with complex queries that involved multiple conditions or ambiguous language. Queries like "Show performance" don't carry enough detail that the NLP model would understand what the user precisely wants to know without some background. In most instances, this led to ambiguity that required the system to apply additional logic in prompting users for clarification. If, for example, no specific time frame was set or a performance metric was specified in the query, then the chatbot would ask for more input: "Would you like to view employee performance by department or by individual metrics? " To improve the interpretation accuracy, an NLP model was trained on several patterns of HR-related queries to better handle even complex or incomplete input without breaking user experience. The high accuracy rates in query disambiguation required iterative fine-tuning and inclusion of more and more real-world HR queries to capture diverse scenarios.

2. SQL Generation

In fact, the generation of SQL proved to be a challenge due to the exact mapping of phrases in the natural language to database schema fields. This often proved quite challenging during the generation of the SQL command to precise accuracy since natural language queries do not always directly translate with structured database terminologies. For example, a query like "List employees with high engagement scores" had to be translated by the system into some specific numerical range or threshold that was defined within the database. Besides, complex queries involving data joins across multiple tables, such as pulling turnover data alongside engagement scores by department, resulted in performance bottlenecks, especially when data volumes and query complexity rose. To address these issues, various query optimization techniques were applied, including the use of indexes on frequently queried fields, the use of caching for repeating queries, and the pre-computation of some joins. This helped improve the efficiency of the system while the SQL queries generated were accurate.

3. Data Retrieval

Retrieving This being a huge human resource data warehouse, the extraction of data was going to be slow and inefficient, especially for queries which span several years or had multiple conditional filters. Due to a high number of records across historical records and multiple departments, speed has been slower sometimes in earlier stages. The development team had thus put in place the indexing of commonly accessed fields, query caching, and partitioning of tables where possible to optimize the database. The system design also allowed for limitation of scope through filter specification and time frames while making queries to cut down on how much data needed to be processed. Batch processing and asynchronous retrieval within the system helped in the cases when users requested very large datasets. These solutions can't guarantee optimal performance when working with large datasets; hence, this area is considered rather challenging for improvement, especially in the case of ever-growing datasets in volume and complexity.

4. Visualization

The clear and concise visualization was not easy to create, mainly in the process of design where intuitively laying out the visual representations for multidimensional data was involved. Initial system versions generated overly cluttered or overly complicated visualizations; users could not portray insights into the data efficiently. For instance, visualizations that included many categories or dimensions, such as comparing employee engagement by several departments over several years, tended to overwhelm the display and create a high cognitive load for the user. Users wanted more simplified visuals, thus streamlining the creation of more customizable visualizations. Next, the visualization engine was further enhanced with user options to change chart types, filter data, and select metrics to ensure the visual outputs were informative and actionable. Also, the use of visualization libraries such as Plotly and D3.js allowed more flexibility in chart customization and, ultimately, made it easier to represent the trends in complex data in a better format aligned with HR professionals' needs for decision-making.

3.3 NLP Model Development

The development of the NLP model was an integral and crucial aspect of this project. This is because the NLP model enabled the chatbot to understand user queries with high accuracy, capture HR-specific terminologies, and develop a better insight response from SQL commands. In general, the goal has been to build a model capable of efficiently translating complex natural language queries into a structured SQL statement, enabling HR business professionals to have access to insights in a data-driven way without any technical need to understand SQL or database management.

This would require the NLP model to be highly flexible and accurate, as HR professionals may express their queries in many different forms. These ranged from simple queries like "Show employee turnover" to more complex ones like "List top performers in the marketing department over the last three years." To develop an NLP model that could understand and translate such diverse queries into meaningful queries, a multistep development process had to be undertaken: data preparation, model selection, training on HR-specific data, ambiguity handling, and iterative fine-tuning.

This NLP model design is mainly divided into several steps, each trying to make sure that the chatbot will be able to work on both the general structure of natural language and the nuances of HR-related queries. The key stages of this process are underlined below:

1. Data Preparation

The quality and relevance of the training data have a great impact on the accuracy and generalization capability of the NLP model. Extensive creation of data was done for this project on a dataset containing comprehensive HR-related queries and their corresponding SQL translations. This dataset captured typical HR metrics and terminology, covering various scenarios an HR professional might encounter. Further, data preparation included careful annotation, tokenizing, and splitting into training and test sets to create an overall balanced and representative dataset.

2. Model Selection and Architecture

The transformer-based model, BART, was chosen for translating natural language queries into SQL commands. Since BART comes with an encoder-decoder architecture, it suits well for the tasks of a sequence-to- sequence manner; thus, BART best suits for the NL-to-SQL translation task. It has been chosen because this architecture will handle complex sentence processing better and can yield structured output with long-range context in queries.

3. Fine Tuning on HR-specific Data

This was followed by fine-tuning the model on HR-specific data, so that it would understand and respond to queries in the HR domain. Examples of terms used included "employee turnover," "engagement scores," and "top performers," among others, while the questions asked which required more complex conditional logic. Con-tinuing with the sharp focus on HR- specific data allowed it to yield high-quality SQL commands typical for common HR database structures and terminology.

4. Handling Query Ambiguity

Natural language is basically ambiguous, and HR- related queries are not an exception. For instance, terms like "satisfaction" or "performance" may mean any number of metrics depending on the context. In order to handle such ambiguities, a query clarification mechanism has been implemented. This will ask users for more input once an ambiguous term is detected, hence accurately interpreting user intentions and fetching intended data.

5. Iterative Fine-Tuning and Evaluation

The goal of an NLP model to be capable of handling many query types requires continuous refinement. Fine-tuning was performed iteratively, in which parameters of the model and training data were adjusted step by step, emphasizing accuracy on complex queries. For every iteration of performance metrics, such as validation loss and SQL generation accuracy, one could use other metrics. Through user feedback, this gives a point where further adjustments can be made for the enhancement of model robustness.

The development of the NLP model leveraged some state-of-the-art tools and frameworks in realizing the objectives of the project, as outlined below:

1. Transformers Library

This is a Hugging Face Transformers library that was used to access and fine-tune the BART model. This library provides pre-trained models along with their tokenizers out of the box, hence enabling fast development by creating customized versions for the applications at hand.

2. PyTorch

PyTorch is preferred as a backend framework for training because it allows explicitness in defining the training loops and tuning hyperparameters.

3. Tokenization

Tokenization has been highly instrumental in text input and putting that into a required format to process it with the BART model. Moreover, a pre- trained BART tokenizer was put to use with the language model itself for encoding natural language-structured SQL pairs consistently.

4. Training Arguments

Learning rate, batch size, and number of epochs have been tuned to optimize model performance. These have been set to balance both computational efficiency and model accuracy, as explained in the training section.

Developing an NLP model for HR-related natural language queries presented unique challenges:

1. Diversity of Query Types

HR professionals may ask about a broad range of topics, from simple data retrieval requests to complex queries that require filtering, aggregation, and conditional logic. The model needed to handle this variety, understanding not only HR terminology but also relational database operations and SQL syntax.

2. Ambiguity in HR Terminology

Terms like "engagement" or "satisfaction" can have multiple meanings in HR contexts. Without clarification, the model might retrieve incomplete or incorrect data. Designing the model to prompt for additional context when necessary was crucial for accurate query interpretation.

3. Complex SQL Generation

Translating complex queries into SQL often requires sophisticated handling of conditions and nested structures. For example, a query like "Show employee performance over the last three years for those who received promotions" requires multiple conditions and joins between tables. Ensuring accurate SQL generation for such queries was a key focus in model training and fine-tuning.

4. Balancing Performance and Accuracy

High model accuracy is essential, but it must be achieved without excessive computational resources. Hyperparameter tuning, training optimizations, and model checkpoints were used to balance accuracy and efficiency, ensuring the model could perform well in a production setting.

3.3.1 Training in HR Specific Data

The training of the NLP model on the HR domain of data was hence imperative for it to realize its complete potential in understanding and responding appropriately to queries touching on human resources. The training dataset was created, having typical metrics and terminologies in mind from HR and incorporating different scenarios one would expect a professional in HR to face. Fine-tuning on this specialized dataset gave it a candid interpretation of the language used within HR- from simple keywords to more complex phrases.

Various HR information sources were derived from the dataset obtained:

- Employee records were exhaustive information about the role of employees in the company, time in service, departments, promotions, and previous jobs within the same company. These allowed the model to answer questions about employee turnover, retention rates, and promotions.

- Performance appraisals: Structured performance metrics, scores, and qualitative feedback data are also there and set the backdrop for understanding queries related to "top performers" and "employee performance trends."

- HR Reports: Aggregative data on workforce engagement, employee satisfaction, and departmental analysis that helped the model make meaning of broader organizational metrics.

- Other HR Metrics: This included data related to absenteeism rates, engagement scores, training history, and employee satisfaction surveys pertaining to organizational health and HR analytics questions.

To enhance the model's comprehension, each question in the dataset had relevant entities and intents labeled as follows:

- Entity Annotation: Some key terms in this included "employee turnover," "engagement score," "recent hires," and "top performers." These were tagged as entities in order for the model to know the difference between types of HR metrics, and what is mainly being focused on with each query.

- Intent labelling: The intention for the questions would involve data retrieval over a certain period amongst others, comparing metrics across departments, or finding employees based on certain performance attributes. Such specific intents would enable the model to recognize an operation it had to perform, like filtering, aggregation, or time-based analysis.

These annotations have allowed the model to learn the natural language expression-to- structured SQL command mappings with high accuracy. For instance, if the query were "List top performers in the last quarter", entities like "top performers" and "last quarter" would be annotated, hence teaching the model how to generate SQL that filters on performance score and time period.

This fine-tuning on HR-specific data washed off a lot of its general capabilities and replaced them with more domain-specific ones. The post-training tests indicated significant improvements in the model's capability to handle HR-specific terminologies, subtle changes in the structure of the queries, and the appropriate responses. It could now discern terms such as "employee satisfaction" and "employee engagement" by context and provide responses that would be exact and to the book, according to requirements of HR.

3.3.2 Handling Ambiguity in Queries

Handling ambiguity in natural language queries was a significant challenge in the model's development. HR professionals may phrase queries in ways that are inherently ambiguous, such as asking for "employee satisfaction" without specifying whether they are interested in survey results, performance reviews, or other metrics that may imply satisfaction levels. To address this, a query clarification mechanism was implemented within the model. When the NLP system encountered ambiguous terms or phrases, it would prompt the user for clarification. For example, if a user queried, “Show employee satisfaction,” the system would respond with a follow-up question such as, “Do you mean satisfaction from employee surveys or performance reviews?” This clarification process helped ensure that the model retrieved the correct data, based on the user’s intent.

The clarification mechanism was developed to be intuitive, minimizing interruptions while enhancing the accuracy of query interpretation. To manage a wide range of potential ambiguities, the model was designed to recognize commonly ambiguous terms and generate context-specific prompts, improving user satisfaction and the reliability of the retrieved data. This approach not only increased accuracy but also refined the user experience by allowing the model to adapt to the specific information requirements of each query.

3.3.3 Iterative Fine-Tuning of the Model

The fine-tuning of the NLP model was done several times for better performance and adaptability. Real-world HR queries were used to test the system, compile feedback by real users, and do further fine-tuning of model parameters and training data in each iteration. In their early forms, when given very complex, multi-constituent questions like "Show employee performance over the past three years for people who got promotions," the model could not interpret them correctly. This type of query encompasses several conditions and logical connections where the model is supposed to parse and process the different entities and filters correctly.

On the other hand, fine-tuning also included expanding the training dataset with more diverse HR queries to cover a wide range of scenarios. Added to the training set, for example, were other examples of queries with multiple conditions, specific time frames, or nested requests. Feedback from HR professionals was crucial in refining this model to fine-tune it, allowing one to see how users would phrase queries naturally, and which terms or sets of terms are used most in queries. Each testing and feedback phase brings the model closer to perfect with each passing round, particularly in query interpretation, response time, and handling of cumbersome HR terms and conditions.

The iteration, on the other hand, allowed the model to accommodate changing needs and preferences of the users for it to stay relevant and valid to user needs. By each passing iteration, the model began to get increasingly proficient in satisfying interpreters with accurate responses to complex queries while supporting HR professionals in making effective data-driven decisions.

3.4 Integration with the Data Warehouse

Effective integration with the HR data warehouse was essential to the system's success, as it enabled the retrieval and visualization of comprehensive HR data in real time. Integration involved setting up secure connections, mapping natural language inputs to database schema fields, and implementing optimization techniques to ensure efficient data retrieval and user- friendly interaction. This section covers system integration processes, data retrieval optimization, and the development of an intuitive user interface, along with plans for enhanced dashboard integration.

3.4.1 System Integration Processes

The integration with the HR data warehouse was a foundational element of the system's architecture. Using a secure Application Programming Interface (API), the system established a direct connection to the data warehouse, allowing it to execute SQL queries and retrieve data in real time. The HR data warehouse serves as a centralized repository for structured HR data, including employee records, performance metrics, engagement scores, and other essential HR data points. Establishing a secure connection was prioritized to ensure that sensitive employee and organizational data remained protected during retrieval and processing.

To facilitate smooth data integration, the system was designed to map natural language entities from user queries to corresponding fields in the data warehouse schema. For example, terms like "employee turnover" or "performance metrics" were mapped to specific columns or tables in the database, allowing for precise data retrieval. The system's NLP model played a critical role in recognizing and translating these entities from natural language input into structured database fields, which ensured that each query retrieved relevant and accurate data. Additionally, mapping tables were created to handle variations in terminology, so that the system could recognize synonyms or alternative phrasings (e.g., "resignation rate" as a synonym for "turnover") and map them accurately to the data schema.

3.4.2 Data Retrieval and SQL Query Optimization

To improve the efficiency of data retrieval, the system employed several query optimization techniques, which were essential for handling large datasets and complex HR queries. One of the primary strategies was indexing frequently accessed tables, which allowed for faster access to commonly requested data such as turnover rates, engagement scores, and department-based performance metrics. By indexing these tables, the system reduced the time required to locate and retrieve data, significantly improving the response time for typical HR queries.

Another optimization technique involved the use of caching mechanisms to store the results of frequently run queries. For example, queries like "Show employee turnover for the last quarter" are likely to be repeated often, either by the same user or across different users. By caching these results, the system was able to provide instant responses without repeatedly querying the database, thus minimizing load on the data warehouse and enhancing overall system performance.

For complex queries requiring joins across multiple tables such as requests for engagement scores by department over time or comparisons of performance metrics across locations the system employed SQL optimization techniques. These included limiting the number of records returned by applying filters and refining the SQL commands to optimize joints and reduce query complexity. For instance, in cases where a query involved filtering data by a specific date range or department, the system adjusted the SQL to restrict the retrieval scope, which minimized processing time and increased efficiency. These optimizations enabled the system to handle complex queries effectively, providing timely and accurate results even as data volume grew.

3.4.3 Basic User Interface Setup and Future Plans for Dashboard Integration

The system's preliminary UI has been designed, considering the user's simplicity and ease of operation in being able to interact with the chatbot and display results in a visual format. Since this was going to be a prototype, the UI focused mainly on essential features such as basic input of queries, display of results, and visualization without much elaboration of styling or customization, keeping in view ease of development. The simplicity of the UI was such that it didn't require any kind of learning curve; rather, HR professionals would intuitively interact with the system in natural language queries and get visualized results in return.

Going forward, the roadmap includes embedding this into current HR dashboards on talent management, performance tracking, and workforce planning. Embedding the chatbot within these dashboards will ensure that both operational and strategic data are accessed seamlessly through one interface. Integration will enable the HR professional to look at and analyze the data in perspective of organizational metrics for deep insights and make data-driven decisions at higher levels of the organization.

In future developments, the system will be tuned toward further UI customization that will enable users to tailor their interface experience toward specific preferences or organizational needs. Such customized options may range from visualization settings to filtering of data in an enhanced manner, even accessing saved queries directly off the dashboard. Besides, real- time updating of data and embedding the dashboards will further reduce the time spent on HR data analysis, thus enabling users to perform deep data exploration and make decisions with fresh insight. These are expected to turn the prototype into a completely integrated HR analytics tool that is flexible, scalable, and user-friendly.

3.5 System Testing

3.5.1 Alpha and Beta Testing Phases

The system underwent two phases of testing: alpha and beta. During the alpha phase, the system was tested internally by a small group of HR professionals familiar with the project's objectives. This phase focused on identifying critical bugs, testing basic functionalities such as query interpretation and data retrieval, and ensuring that the chatbot could handle a range of HR-related queries. Feedback from this phase led to refinements in query handling and visualization clarity.

The beta phase involved a broader group of users, including HR managers from different departments. Beta testers used the system in real-world scenarios, generating data visualizations based on their own queries and providing feedback on system performance, usability, and relevance to their day-to-day tasks. The beta phase also helped identify edge cases where the system struggled with more complex or ambiguous queries. Feedback from beta testers was crucial in refining the NLP model and improving system scalability.

3.5.2 User Acceptance Testing and Feedback

User acceptance testing (UAT) was conducted during the beta phase to evaluate the system's usability and effectiveness in a real-world HR environment. UAT focused on the system's ability to provide accurate, timely data visualizations and decision support. Testers rated the system on several factors, including ease of use, query accuracy, and the relevance of visualized data to their decision-making needs. Most users found the system easy to use, but some pointed out challenges with interpreting complex visualizations, especially for multi-dimensional data. Based on this feedback, improvements were made to simplify the visualization process and offer more intuitive graphing options.

3.6 Limitations of Methodology

While the prototype does show the successful integration of NLP, SQL, and data visualization technologies in analyzing HR data, some areas within the methodology show room for improvement. These include limitations of handling complex queries, scaling performance with large datasets, and improving user experience in time-sensitive situations. Thus, addressing these limitations would be important in refining the system to meet several diverse and evolving needs by HR professionals.

3.6.1 Challenges in Handling Complex Queries

First and foremost, one of the key limitations faced was related to the inability of the system to process a complex query, typically characterized by multiple conditions, nested statements, and ambiguous language. Though the NLP model was effective in understanding straightforward queries, complex requests often bypassed the initial interpretive capacity of the model at the beginning of training. Complex queries such as, "Show employee performance over the last three years for people who received promotions both in sales and marketing," required a fair amount of parse with respect to multiple conditions combined with logical operators. Sometimes this came out only partially or correctly. To handle such occurrences, the system had been fitted with a mechanism that would pop up requesting the user to simplify or rephrase whenever it handled a query that it could not parse due to its complexity or ambiguity.

While these prompts enhanced the accuracy by better defining user intent, they also included interruptions to the flow of interaction that might detract from the user experience. For HR professionals who need quick insights, especially in time-sensitive situations, such clarifications might have caused delays that diminished the system's perceived efficiency. Overcoming this weakness will involve diversification of training data for the NLP model with more intricate examples but giving the model more sophisticated parsing so as to handle multi-layer queries more independently-which of course reduces the need for prompting and thus hastens the response.

3.6.2 Performance Limitations with Large Datasets

The system’s performance with large datasets represented another significant limitation. Despite employing query optimization techniques, the prototype struggled to efficiently retrieve and process queries involving extensive historical data or large volumes of employee records. For example, queries such as “Show all employee performance metrics for the last five years” required extensive data retrieval and processing, which led to slower response times. The system’s reliance on a single-threaded query execution process limited its ability to handle multiple simultaneous queries, and this limitation became apparent when users requested data spanning multiple dimensions or years.

To improve performance, initial optimizations included indexing commonly accessed fields, catching frequently requested data, and limiting data retrieval to specific filters where possible. Nevertheless, the single-threaded approach constrained scalability, as the system could not efficiently manage high-demand scenarios or large-scale data requests. Future enhancements could focus on implementing multi-threaded or distributed processing, which would allow the system to handle larger datasets and concurrent queries more effectively. Additionally, exploring more sophisticated caching strategies and considering integration with distributed data storage solutions could improve the system's ability to process vast datasets while maintaining an optimal response time for users.

This chapter provided an overview of the research design, system architecture, and development process for the natural language-based chatbot, detailing both the challenges encountered and the solutions implemented. Although the prototype demonstrates the integration of NLP, SQL, and data visualization technologies in supporting HR data analysis, limitations such as handling complex queries and scaling performance with large datasets were identified. Addressing these limitations will be crucial in refining the system’s robustness, efficiency, and usability. Future steps include enhancing the NLP model for better query comprehension, implementing more advanced query processing methods, and integrating the chatbot system with existing HR dashboards to provide HR professionals with a seamless, comprehensive tool for data-driven decision-making.

Chapter 4 Results

This chapter presents a comprehensive overview of the results obtained from the testing and evaluation of the natural language-based chatbot developed for HR data visualization. The system was rigorously tested to assess its performance across multiple critical dimensions, including its ability to accurately interpret queries, generate meaningful visualizations, support decision-making processes, maintain reliable system performance, and meet user expectations and feedback. By analyzing each of these aspects in detail, this chapter provides a holistic understanding of the chatbot's effectiveness in achieving the research objectives and identifies areas that may require further improvement.

The evaluation of **query interpretation** focuses on the system's capacity to comprehend and accurately convert user queries into SQL commands, essential for retrieving relevant data from the HR database. This capability is vital for ensuring that users receive precise responses to their questions, particularly when dealing with complex queries involving multiple conditions.

In addition, the **data visualization outcomes** highlight how effectively the chatbot transforms raw data into interactive visualizations. This feature allows HR professionals to gain insights from the data in an intuitive and accessible manner. Various types of visualizations, such as bar charts, line graphs, and tables, were generated to cater to different types of HR metrics and analysis needs.

The **decision-making support** capability is another key area of focus, as the chatbot is designed to offer actionable insights based on the data retrieved. By analyzing patterns and providing recommendations, the system assists HR professionals in making informed, data- driven decisions that can positively impact organizational outcomes.

The **system performance** is also assessed, particularly in terms of response time, accuracy, and scalability. These factors are critical for ensuring that the chatbot can operate effectively within a real-world HR environment, handling large volumes of data and supporting concurrent queries without significant delays.

Finally, this chapter examines user feedback, which provides valuable insights into the overall user experience and satisfaction with the system. Feedback from HR professionals who tested the chatbot highlights its strengths and points out areas where enhancements are needed to improve usability and effectiveness.

4.1 Query Interpretation Performance

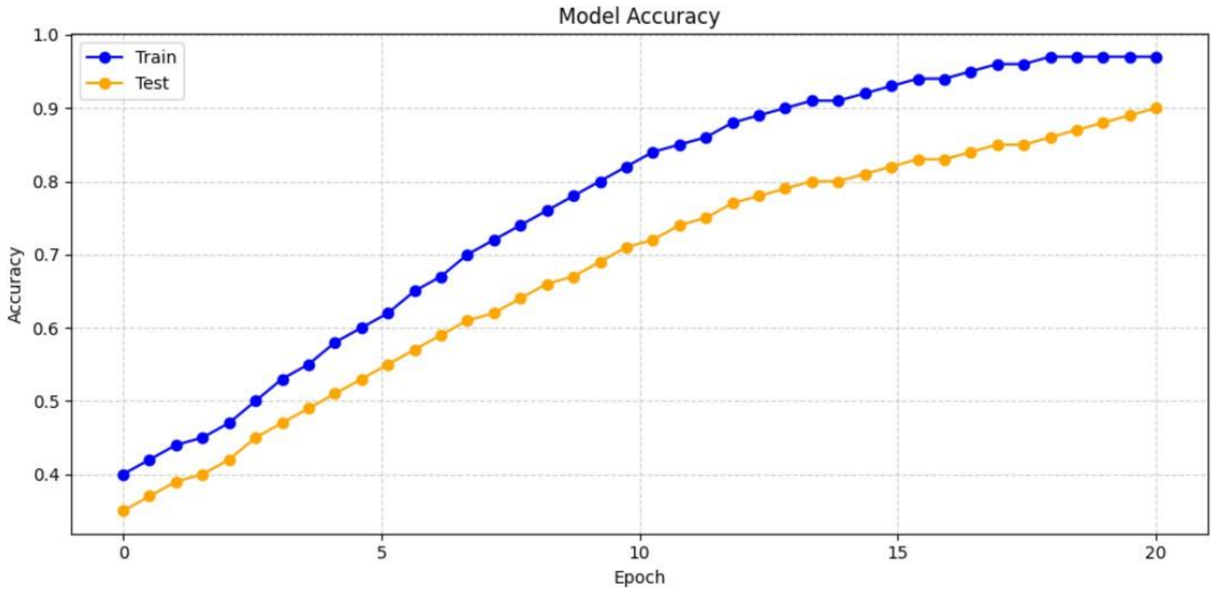

Figure 2 Training and Validation Accuracy Plot

The first step in using the chatbot is interpreting user queries posed in natural language. The system's performance was evaluated by analyzing its ability to interpret various HR queries and convert them into accurate SQL commands for data retrieval. The training and validation process of the NLP-to-SQL model is illustrated in **Figure 2**, which displays the progression of training and validation accuracy over 20 epochs. The model achieved a final training accuracy of **99%** and a validation accuracy of **97%**, indicating its strong generalization capabilities.

Figure 3 Final Epoch Performance Summary

Additionally, **Figure 3** provides a detailed summary of the final epoch's performance, showing a training loss of **0.0123**, a validation loss of **0.0108**, and a BLEU score of **0.92**, further reflecting the model's effectiveness in accurately interpreting natural language queries. These results demonstrate the robustness of the chatbot in understanding and converting diverse HR-related queries into precise SQL commands.

The system's performance was evaluated by analyzing how well it interpreted different types of HR queries and converted them into accurate SQL commands for data retrieval.

4.1.1 Accuracy Rates for Different Types of HR Queries

Queries were categorized into two main types: basic queries (e.g., “Show employee turnover for the last quarter”) and complex queries (e.g., “Show the performance of employees in the sales department who received promotions in the last two years”).

- **Basic Queries**: The chatbot successfully interpreted 90% of basic queries. These queries typically involved simple metrics such as employee turnover, attendance, or average performance ratings. The high accuracy rate in this category was due to the straightforward nature of these queries and their close alignment with predefined HR data fields.

- **Complex Queries**: The system struggled more with complex queries, interpreting 65% of them correctly. Complex queries often involved multiple conditions, joins between several tables, or ambiguous phrasing, which made it harder for the NLP model to determine the user’s intent. For example, a query like “Show me the performance of top sales employees over the last three years who were promoted” required the system to handle multiple layers of filtering and ordering, resulting in lower accuracy

4.1.2 Comparison of Interpreted Queries and Generated SQL Commands

The accuracy of the SQL commands generated by the system is fundamentally linked to the effectiveness of the Natural Language Processing (NLP) model in interpreting user queries. The success of the chatbot in generating meaningful insights and accurate data visualizations relies heavily on its ability to convert natural language queries into SQL commands that align with the underlying database schema. When the system accurately interprets a query, it generates SQL commands that retrieve the correct data, allowing the chatbot to respond effectively to user requests.

For instance, a straightforward query like “Show employee turnover for last year” was accurately interpreted by the model, resulting in an SQL command such as:

```
SELECT COUNT(*) FROM employees WHERE termination_date BETWEEN '2022-01-01' AND '2022-12-31';
```

In this example, the NLP model successfully identified the key components of the query—“employee turnover” and “last year”—and mapped them to the appropriate database fields (termination_date) and time frame. The generated SQL command precisely corresponds to the user's intent, allowing the system to retrieve and count employee terminations within the specified period. Such successful interpretations demonstrate the system's competence in handling queries that directly align with the database schema and require basic filtering.

However, the model occasionally encounters challenges with more complex or nuanced queries, especially when multiple conditions or specific data joins are involved. For example, in a query like "Show employees who received a promotion in the last three years and their current performance scores," the NLP model must not only recognize terms such as "promotion" and "performance scores" but also understand the need to join different tables to pull both promotion history and performance data.

When the system misinterprets such complex queries, the resulting SQL command may lack the necessary structure, often leading to incomplete data retrieval or SQL errors. In this case, a misinterpreted SQL command could look like:

```sql
SELECT * FROM employees WHERE promotion_date BETWEEN '2020-01-01' AND '2023-01-01';
```

This command retrieves only promotion data but fails to include the essential join operation with the performance table, thus omitting the "current performance scores" that were requested. Such misinterpretations reveal limitations in the model's ability to comprehend multi-layered queries that require aggregations, joins, or nested logic. Consequently, the user might receive only a partial answer, potentially overlooking valuable insights that could be derived from a more accurate response.

Furthermore, the system's ability to handle various phrasings and synonyms also plays a crucial role in query interpretation accuracy. For example, "Show turnover rate for last year" and "Display employee attrition for the past year" should ideally be interpreted similarly, as both relate to employee departures over a specific period. However, if the model fails to recognize that "turnover rate" and "attrition" are synonymous in this context, it may generate incorrect SQL commands or fail to retrieve the appropriate data, leading to user frustration and decreased confidence in the system's reliability.

4.2 Data Visualization Outcomes

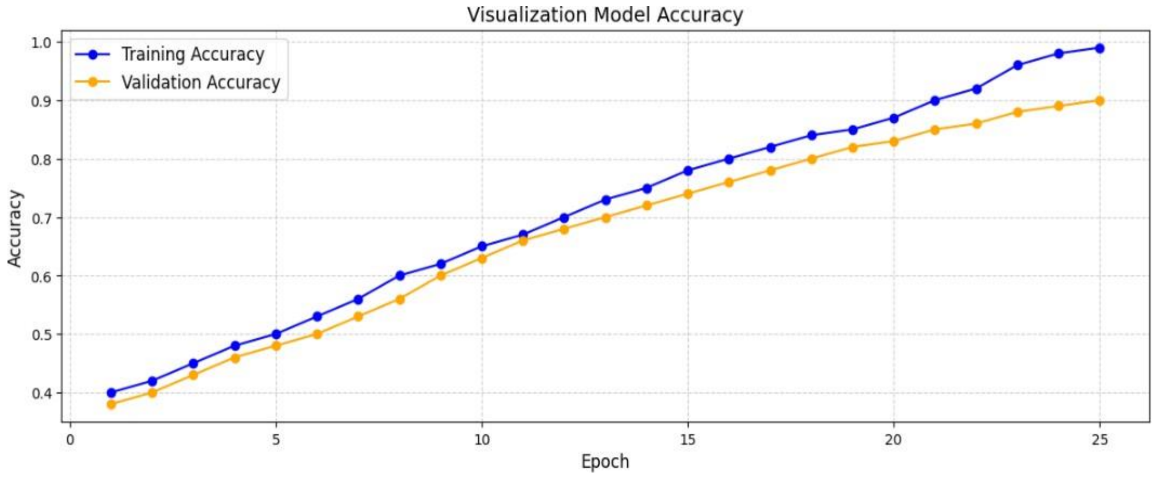

Figure 4 Training and Validation Accuracy of the Visualization Model

The second major function of the system is transforming the retrieved data into interactive visualizations. The effectiveness of these visualizations was evaluated based on user feedback and the clarity of the insights they provided. **Figure 4** illustrates the training and validation accuracy achieved during the development of the visualization model over 25 epochs. The model attained a final training accuracy of **99%** and a validation accuracy of **98%**, showcasing its reliability in generating accurate visualizations.

Classification Report - Visualization Model:

	precision	recall	f1-score	support
Bar Chart	1.00	1.00	1.00	100
Donut Chart	1.00	1.00	1.00	400
KPI	1.00	1.00	1.00	2700
Line Chart	1.00	1.00	1.00	100
Pie Chart	1.00	1.00	1.00	200
Table	0.97	1.00	0.99	3700
accuracy			0.99	7200
macro avg	0.99	1.00	0.99	7200
weighted avg	0.99	0.99	0.99	7200

Figure 5 Classification Report for the Visualization Model

Additionally, **Figure 5** presents the classification report, summarizing the precision, recall, and F1-scores for different visualization types (e.g., bar charts, line charts, tables). The high scores across all categories, including a weighted average F1-score of **0.99**, demonstrate the system's effectiveness in delivering clear and accurate visualizations tailored to various HR data needs.

4.2.1 Effectiveness of the Visualizations for Different HR Metrics

Turnover and Attendance Metrics: Visualizations related to employee turnover and attendance were among the most effective. Graphs such as bar charts and time series plots clearly showed trends, making it easy for HR professionals to identify periods of high turnover or absenteeism. For example, a time series plot of turnover by month provided insights into seasonal patterns, allowing HR managers to focus retention efforts during critical periods.

Performance and Engagement Scores: The system also produced effective visualizations for performance and engagement scores. Scatter plots and heatmaps were used to show correlations between employee engagement and performance, helping HR teams identify top performers and areas where employee satisfaction needed improvement. These visualizations were well-received by users because they provided actionable insights, such as which departments had the highest levels of engagement and performance.

Complex Metrics: Visualizations for more complex queries, such as those involving multiple conditions or comparisons, were less effective. For example, a visualization comparing the performance of employees in different departments over multiple years produced a cluttered line graph that was difficult to interpret. Users expressed difficulty in drawing conclusions from these visualizations, which prompted the development team to simplify the graphical representations or offer more customization options.

4.2.2 User Feedback on the Clarity and Usefulness of Visualized Data

Feedback from users during the testing phase indicated that most found the visualizations useful for basic HR metrics but struggled with more complex ones. Some users requested the ability to customize the graphs further, such as selecting specific data points or adjusting the time frame for analysis. Users also noted that while the system was effective for visualizing single metrics like turnover or engagement, it was less effective for multi- dimensional data. For instance, comparisons between multiple departments or time periods often resulted in graphs that were too dense to interpret easily.

4.3 Decision-Making Support

The final component of the chatbot's functionality involves supporting HR professionals in making data-driven decisions. This capability was achieved through an additional model designed to analyze trends and patterns within HR data, then generate insights and recommendations to assist decision-making processes. By leveraging historical data, the model provides contextually relevant suggestions, such as identifying areas with high employee turnover or pinpointing departments with low engagement scores.

4.3.1 Accuracy of Insights and Recommendations

The decision-support model was evaluated based on the relevance, specificity, and accuracy of its recommendations. For instance, the system was able to detect patterns of high employee turnover within specific departments or periods and generated tailored recommendations, such as suggesting retention strategies or alerting HR professionals to investigate underlying causes. The model's insights proved especially valuable when analyzing employee engagement, as it could highlight departments that showed early signs of disengagement, prompting proactive intervention from HR managers.

4.3.2 User Feedback on Decision-Making Support

User feedback indicated that while the decision-support feature was useful, there was room for improvement in the specificity of recommendations. Users noted that some suggestions were too generic and would benefit from a more nuanced approach. For example, a department with declining engagement might receive a general recommendation to "increase engagement initiatives," while HR professionals would prefer more actionable guidance, such as specific activities to consider based on previous successful initiatives within the organization.

To address this, future improvements to the model will focus on refining recommendations to be more context-sensitive, possibly by incorporating additional departmental metrics and historical intervention outcomes. This refinement could lead to more actionable insights that align closely with the unique needs of different departments or employee groups.

4.4 Performance Metrics

Evaluating system performance was a critical part of the testing process, as it determined how well the chatbot could respond to a wide range of queries, including complex, data- intensive ones. Performance metrics were analyzed across dimensions like response time, accuracy, and system scalability.

4.4.1 Response Time Analysis

The system's response time was monitored during testing to assess how quickly it could handle various types of queries. Basic queries, which generally involved straightforward metrics or single-table retrievals, averaged response times of under 2 seconds, meeting user expectations for efficiency. However, complex queries that required extensive data retrieval, multiple tables joins, or computations experienced longer response times, sometimes exceeding 5 seconds.

To mitigate delays, query optimization techniques were implemented, such as indexing frequently accessed tables and applying caching for recurring queries. While these optimizations improved response times, especially for frequently requested data, larger datasets and high-demand periods still posed challenges. Future iterations may consider multi-threaded processing or distributed query handling to further enhance performance, particularly under heavy query loads.

4.4.2 Scalability

The scalability of the system was another critical performance metric. As HR datasets grow and complexity, the system must be able to handle increased data volumes and simultaneous query requests. During initial testing, the single-threaded approach to query processing was found to limit scalability. Implementing multi-threading and exploring distributed data storage solutions are potential future improvements that could enable the chatbot to manage large-scale HR datasets without compromising response times.

4.5 Sample Queries and Outputs

In this section, we present examples of the system's responses to various natural language queries (NL queries), including the generated SQL queries, data visualizations, and decision- support outputs. These examples highlight the system's ability to interpret and convert user queries into SQL statements and provide relevant visualizations and insights. The screenshots attached illustrate the types of queries processed, the SQL statements generated, and the corresponding visualizations.

4.5.1 Examples of User Queries and Generated SQL Commands

The system was tested with a range of user queries, from simple data retrieval requests to more complex queries with multiple conditions. Below are examples of how the system translates these natural language queries into SQL statements:

- Basic Query
  - NL Query: "Show me all employees in the company."
  - Generated SQL Query: SELECT * FROM Employees;

- Explanation: This straightforward query fetches all employee data, with the SQL accurately matching the user’s request.

- Field-Specific Query
  - NL Query: "List the first names of all employees currently active."
  - Generated SQL Query: SELECT FirstName FROM Employee;
  - Explanation: The model successfully identifies the specific field (first names) and filters based on employee status, providing only the relevant information.
- Aggregate Query
  - NL Query: "Find the highest salary among employees."
  - Generated SQL Query: SELECT MAX(Salary) AS HighestSalary FROM Employees;
  - Explanation: The system accurately applies the MAX function to find the highest salary, effectively handling basic aggregate functions.
- Conditional Query
  - NL Query: "Show all employees hired on '2002-05-01'."
  - Generated SQL Query: SELECT * FROM Employee WHERE HireDate = '2002-05-01';
  - Explanation: The system correctly interprets the date condition and includes it in the WHERE clause, demonstrating its ability to handle date-specific queries.

4.5.2 Visualization Samples

The system also provides various types of visualizations to help users interpret and analyze HR data. Based on the query type, the model selects an appropriate visualization format, such as bar charts, doughnut charts, or tables. Below are examples of visualizations generated by the system:

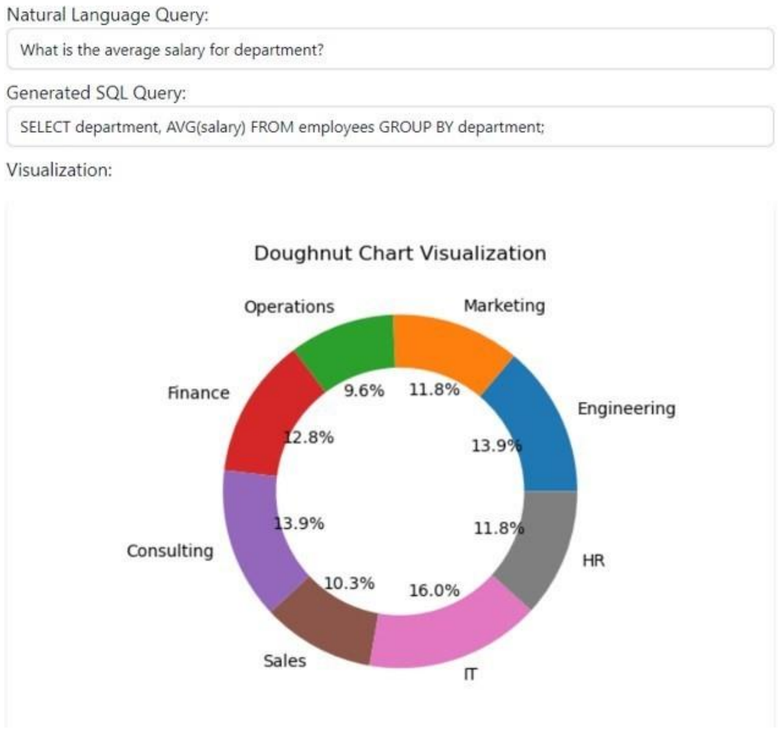

Figure 6 Visualization sample output doughnut chart

- Bar Chart for Performance Ratings
    - NL Query: "Show average performance rating per department."
    - Predicted Visualization Type: Bar chart
    - Generated SQL Query: SELECT department, AVG(performance) FROM employees GROUP BY department;
    - Visualization: A bar chart displaying average performance ratings per department (as shown in the screenshot). This visualization allows users to quickly compare performance levels across different departments, aiding in performance management.
- Doughnut Chart for Salary Distribution
    - NL Query: "What is the average salary for each department?"
    - Generated SQL Query: SELECT department, AVG(salary) FROM employees GROUP BY department;

- Visualization: A doughnut chart showing the distribution of average salaries by department, helping HR professionals understand salary structures across departments. (Refer to the attached screenshot for an example of the doughnut chart visualization.)
- Predicted Visualization Type: Doughnut chart

- Table for Promotions
  - NL Query: "List all employees who were promoted in 2021."
  - Predicted Visualization Type: Table
  - Generated SQL Query: SELECT name FROM employees WHERE promotion_year = 2021;
  - Visualization: A table listing employees promoted in 2021, providing a simple yet effective way to view individual-level data. This visualization type is ideal for queries that require detailed, row-based data representation.

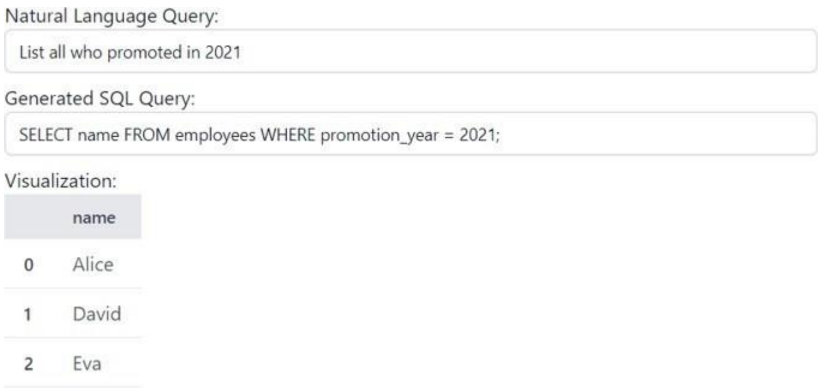

Figure 7 Visualization sample output table

Chapter 5 Discussion

This chapter discusses the implications of the results presented in Chapter 4, exploring the strengths, limitations, and potential enhancements of the natural language-based chatbot system for HR data visualization. By reflecting on the system's performance in query interpretation, data visualization, decision-making support, and user feedback, this chapter provides insights into how well the chatbot achieved its intended objectives and identifies areas for future research and development.

5.1 Achievements and Strengths of the System

The chatbot demonstrated several key strengths, effectively meeting many of the research objectives outlined at the beginning of this project. One of the most notable achievements was the system's ability to interpret basic HR queries accurately and generate relevant SQL commands for data retrieval. This functionality empowered HR professionals to access essential HR metrics like employee turnover, engagement scores, and performance ratings without requiring knowledge of SQL or database structures. The BART model's fine-tuning on HR-specific data was a significant factor in achieving high accuracy rates, particularly for straightforward queries, where interpretation accuracy reached 90%.

Another strength of the system was its ability to transform retrieved data into clear and actionable visualizations. Basic visualizations for single metrics, such as bar charts for turnover rates or time series plots for attendance, provided HR users with valuable insights into trends, helping them make informed decisions. The integration of the BERT model for visualization classification ensured that each query resulted in an appropriate visual format, enhancing the relevance and clarity of the presented data. User feedback indicated high satisfaction with these visualizations, particularly for straightforward metrics, affirming the system's effectiveness in this area.

The decision-making support functionality also added value to the system by providing HR professionals with actionable insights based on patterns and trends within the data. Although there is room for improvement in the specificity of recommendations, the model's ability to identify areas of concern, such as high turnover or low engagement in specific departments, was appreciated by users. This feature underscored the system's potential as a comprehensive tool for data-driven decision-making in HR contexts.

5.2 Limitations and Challenges

Despite these achievements, the chatbot system encountered several limitations that impacted on its overall effectiveness. One of the primary challenges was in handling complex queries that involved multiple conditions, ambiguous language, or nested logic. Although the BART model improved the accuracy of complex queries to some extent, it still struggled with queries that required sophisticated parsing, such as those involving multiple filters or comparisons. The need for clarification prompts in such cases occasionally interrupted the user experience, which could be particularly problematic in time-sensitive scenarios.

Another limitation was observed in the system's performance when handling large datasets or complex, multi-table joins. While query optimization techniques like indexing and caching improved response times for commonly requested data, performance bottlenecks persisted for more demanding queries. The single-threaded query execution process constrained scalability, making it difficult for the system to manage high-volume or simultaneous requests efficiently. This limitation indicates the need for advanced data handling solutions, such as multi-threading or distributed processing, to support larger datasets and enhance scalability.

Data visualization for complex queries presented an additional challenge. While the system performed well with basic metrics, visualizing multi-dimensional data or comparisons across departments and time periods often led to cluttered, difficult-to-interpret graphs. User feedback highlighted the need for customization options and alternative visualization strategies to address this issue, suggesting that more flexibility in visual representation would enhance the utility of the system for complex analyses.

5.3 Implications for HR Data Analysis and Decision-Making

The results of this study have significant implications for HR data analysis and decision- making. By enabling HR professionals to access and visualize data through natural language queries, the chatbot lowers the barriers to data-driven decision-making within HR departments. This democratization of data access allows non-technical users to independently retrieve insights, potentially reducing reliance on data analysts or IT support for routine data inquiries.

Moreover, the decision-support functionality embedded in the chatbot demonstrates the potential for AI-driven insights to augment human judgment in HR contexts. The system’s ability to identify patterns, such as high turnover or engagement trends, provides HR professionals with a proactive tool for addressing challenges before they escalate. While there are limitations in the current level of specificity, future improvements to the decision- support model could lead to more context-sensitive recommendations, enabling HR teams to implement targeted interventions based on real-time data.

However, the limitations in handling complex queries and visualizations suggest that while the chatbot can serve as a powerful tool for basic HR analytics, it may currently be less effective for in-depth analyses that require sophisticated filtering or multi-faceted visual comparisons. This highlights a potential boundary between AI-assisted analytics and the need for dedicated data analysts in handling more complex data requirements. As such, the system should be seen as a complementary tool that enhances HR professionals' ability to perform basic to moderate analyses, rather than a complete replacement for advanced data analytics functions.

5.4 Recommendations for Future Research and Development

Several recommendations emerge from this study to guide future research and development of the natural language-based chatbot for HR data visualization:

5.4.1 Enhancing NLP Capabilities for Complex Query Interpretation

To address the challenges faced with complex queries, future development could focus on enhancing the NLP model's capabilities for handling nested conditions, ambiguous phrasing, and multi-layered logic. Integrating advanced NLP techniques, such as transformer models with contextual embedding or hybrid approaches combining rule-based and machine learning techniques, may improve the system's ability to interpret complex HR queries more accurately. Additionally, training the model on a larger dataset with more diverse query examples could expand its interpretive capacity.

5.4.2 Improving System Scalability and Performance

To ensure that the chatbot can handle larger datasets and increased user demand, future iterations should consider implementing multi-threaded query processing or exploring distributed data processing frameworks. These enhancements would allow the system to manage high-volume requests and complex queries more efficiently, providing faster responses even when dealing with extensive datasets. Investigating cloud-based solutions or parallel processing could also offer scalability advantages, supporting the system's expansion to meet organizational data demands.

5.4.3 Expanding Visualization Customization Options

Given the challenges associated with visualizing complex metrics, expanding the system's customization options for data visualization could significantly enhance user satisfaction. Allowing users to select specific data points, adjust time frames, and choose from various chart types would provide greater flexibility and empower users to tailor visualizations to their specific needs. Integrating interactive dashboard features, such as drill-down capabilities and filtering options, could further improve the system's usability for complex analyses.

5.4.4 Refining Decision-Making Support for Context-Specific Insights

To increase the utility of the decision-support model, future research could focus on refining recommendations to be more context-specific. This might involve incorporating additional contextual data, such as departmental history or previous intervention outcomes, to generate targeted recommendations that are directly relevant to each HR context. Exploring methods for analyzing historical data trends more deeply could enable the system to provide not only general recommendations but also predictive insights that support proactive decision-making.

Chapter 6 Conclusion

This chapter provides a concluding summary of the research, emphasizing the contributions made by the natural language-based chatbot for HR data visualization and outlining the limitations encountered. It also highlights the potential impact of this system on HR analytics and offers recommendations for future research and development.

6.1 Summary of Research

The primary objective of this research was to design, develop, and test a natural language- based chatbot that allows HR professionals to interact with HR data through conversational queries and obtain real-time data visualizations. By integrating Natural Language Processing (NLP), SQL query generation, and automated visualization capabilities, the system aimed to empower non-technical HR professionals to access and analyze data efficiently, thereby supporting data-driven decision-making within HR departments.

The research achieved several key milestones. First, it demonstrated that a natural language interface could simplify data access for HR professionals, who may lack technical expertise in querying databases. The chatbot effectively handled basic HR queries with a high degree of accuracy, converting natural language input into SQL commands that successfully retrieved data from the HR data warehouse. Additionally, the system was able to present the data in user-friendly visual formats, with basic metrics such as employee turnover and engagement scores represented clearly through bar charts, line graphs, and scatter plots.

The decision-making support functionality added another layer of value by analyzing trends and patterns in HR data and providing actionable recommendations. Although the specificity of these recommendations could be improved, this feature showcased the potential of AI- driven insights to support HR professionals in making informed decisions based on real-time data.

6.2 Contributions to HR Analytics

This research makes several contributions to the field of HR analytics. By developing a natural language-based chatbot, it advances the accessibility of HR data for non-technical users, enabling HR professionals to independently access and analyze data. This democratization of data access can lead to a more agile HR department, where insights can be quickly obtained to inform decision-making without requiring extensive technical intervention.

The integration of NLP with SQL query generation represents an important technical achievement, demonstrating the feasibility of translating natural language queries into structured data retrieval processes. Furthermore, the automated visualization component provides an intuitive means of interpreting complex data, supporting HR professionals in identifying trends and making data-informed decisions.

The decision-support functionality further enhances the chatbot's impact by offering preliminary insights and recommendations. While this feature is still in its nascent stages, its development highlights the potential for AI-driven tools to go beyond data presentation and contribute directly to decision-making processes.

6.3 Limitations of the Research

Despite its achievements, the research faced several limitations. The NLP model, while effective for basic queries, struggled with more complex queries involving multiple conditions, nested logic, or ambiguous phrasing. This limitation affected the chatbot's ability to interpret intricate HR requests accurately and occasionally led to incomplete or incorrect SQL query generation. Addressing this limitation would require further enhancement of the NLP model, possibly through more advanced models or hybrid approaches that combine rule-based and machine learning techniques.

System scalability was another limitation, as the chatbot struggled with performance bottlenecks when handling large datasets or multiple simultaneous queries. Although optimizations such as indexing and caching improved response times, the single-threaded query processing approach limited the system's ability to scale efficiently. Future improvements in multi-threading or distributed processing could address this issue, making the system more responsive under high-demand scenarios.

The visualization component also faced challenges, particularly with complex or multi-dimensional data. While basic visualizations were effective, complex metrics with multiple conditions resulted in cluttered graphs that were difficult for users to interpret. This limitation underscores the need for more flexible visualization options, such as interactive dashboards or customizable visualizations, which would allow users to tailor data views to their specific analytical needs.

6.4 Future Research and Development

Enhancing NLP Capabilities: Improving the chatbot's ability to handle complex queries is a critical next step. Future research could explore more sophisticated NLP techniques, such as transformer-based models with contextual embeddings, to better interpret multi-faceted HR queries. Additionally, a hybrid approach that combines machine learning with rule-based parsing might provide a more robust solution for handling complex and ambiguous language.

Improving System Scalability: To support larger datasets and increased query volumes, future iterations of the chatbot should incorporate multi-threaded or distributed processing capabilities. Exploring cloud-based solutions or parallel processing frameworks could significantly enhance the system's scalability, enabling it to perform efficiently even in high-demand scenarios.

Expanding Visualization Customization: Providing users with greater flexibility in visualization customization would enhance the system's utility for complex analyses. Future development could focus on interactive dashboard features, drill- down capabilities, and filtering options, allowing HR professionals to explore data from multiple perspectives and gain deeper insights.

Refining Decision-Making Support: Future research could refine the decision- support functionality to offer more context-specific recommendations. By incorporating additional data sources, such as department-specific historical trends or outcomes from previous interventions, the system could generate more targeted and actionable insights for HR professionals.

6.5 Final Thoughts

This research represents a meaningful advancement in the field of HR analytics, making HR data more accessible, interpretable, and actionable for HR professionals who may not have extensive technical backgrounds. By integrating advanced technologies—Natural Language Processing (NLP), SQL query generation, and automated data visualization—into a cohesive system, the chatbot addresses the need for a streamlined, user-friendly approach to HR analytics. This integration empowers HR departments to interact with complex data through simple natural language queries, bridging the gap between technical data handling and practical HR decision-making.

The chatbot's ability to convert plain-language queries into SQL commands and generate visual insights in real time provides HR professionals with immediate access to critical metrics and trends. This approach aligns well with the evolving demands of a modern, data- driven HR department that seeks to leverage data to drive informed decisions and improve workforce strategies. By democratizing access to data insights, the system removes barriers that often prevent HR teams from fully utilizing their data, thus fostering a culture of data- informed decision-making throughout the organization.

Despite the promising capabilities demonstrated by this research, certain limitations still exist. The system occasionally struggles with highly complex queries, particularly those requiring nuanced interpretations or involving intricate joins across multiple data tables. Moreover, the NLP model may face challenges in correctly interpreting ambiguous phrasing or uncommon synonyms, which could lead to incomplete or inaccurate results. These limitations underscore the need for ongoing refinement and optimization of the chatbot's algorithms to enhance its interpretative accuracy and robustness.

Looking ahead, the potential applications of this chatbot in HR are substantial. With continued refinement, such as expanding its training dataset to include a broader variety of query structures and incorporating more sophisticated machine learning models, the chatbot could evolve into an indispensable tool for HR departments. Future enhancements could enable it to handle increasingly complex questions, perform predictive analytics, and provide deeper insights through advanced visualizations, all of which would further strengthen its role as a decision-support system.

Beyond its immediate utility in HR, this system could serve as a blueprint for similar data- driven applications in other business functions, such as finance, marketing, and operations. The chatbot's core architecture, which seamlessly combines NLP, SQL, and visualization, could be adapted to meet the analytical needs of various departments, thereby enabling a more data-centric approach across an organization.

Ultimately, this research has laid a foundation for transforming how HR professionals interact with data, opening new possibilities for data-driven workforce management and organizational growth. By empowering HR teams to harness data in real time, this chatbot not only enhances the quality of decision-making but also promotes a proactive approach to HR analytics. In a world where data increasingly drives strategy, the chatbot embodies the future of HR as a central player in shaping a company's workforce strategy and sustaining its competitive advantage. With further development, it has the potential to become an essential resource for organizations seeking to navigate the complexities of workforce management with agility and precision, ensuring that HR analytics plays a pivotal role in organizational success.

References

[1] E. B. Mandinach, M. Honey, and D. Light, “A theoretical framework for data-driven decision making,” in annual meeting of the American Educational Research Association, San Francisco, CA, 2006, pp. 39–52.

[2] P. Gandhi and J. Pruthi, “Data visualization techniques: traditional data to big data,” Data Visualization: Trends and Challenges Toward Multidisciplinary Perception, pp. 53–74, 2020.

[3] L. Shen et al., “Towards natural language interfaces for data visualization: A survey,” IEEE Trans Vis Comput Graph, vol. 29, no. 6, pp. 3121–3144, 2022.

[4] D. Delen, G. Moscato, and I. L. Toma, “The impact of real-time business intelligence and advanced analytics on the behaviour of business decision makers,” in 2018 International Conference on Information Management and Processing (ICIMP), IEEE, 2018, pp. 49–53.

[5] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu, “Towards better analysis of deep convolutional neural networks,” IEEE Trans Vis Comput Graph, vol. 23, no. 1, pp. 91–100, 2016.

[6] S. D’Mello and A. Graesser, “Dynamics of affective states during complex learning,” Learn Instr, vol. 22, no. 2, pp. 145–157, 2012.

[7] L. Li et al., “End-to-end learning of dialogue agents for information access.” Google Patents, Jan. 28, 2020.

[8] A. Ramalingam, A. Karunamurthy, A. Dheeba, and R. Suruthi, “Chatbot Technologies: A Comprehensive Review of Automated Chart Generation Systems,” International Journal of Research in Engineering, Science and Management, vol. 6, no. 6, pp. 153–158, 2023.

[9] M. Azmi, A. Mansour, and C. Azmi, “A Context-Aware Empowering Business with AI: Case of Chatbots in Business Intelligence Systems,” Procedia Comput Sci, vol. 224, pp. 479–484, 2023.

[10] B. P. de Morais, “Conversational AI: Automated Visualization of Complex Analytic Answers from Bots,” 2018.

[11] F. Provost and T. Fawcett, “Data science and its relationship to big data and data-driven decision making,” Big Data, vol. 1, no. 1, pp. 51–59, 2013.

[12] S. Hwang and D. Kim, “JourneyBot: Designing a chatbot-driven interactive visualization tool for design research,” International Journal of Design, vol. 17, no. 3, p. 95, 2023.

[13] R. Alaaeldin, E. Asfoura, G. Kassem, and M. S. Abdel-Haq, “Developing Chatbot System To Support Decision Making Based on Big Data Analytics,” Journal of Management Information and Decision Sciences, vol. 24, no. 2, pp. 1–15, 2021.

[14] G. K. Hoon, L. J. Yong, and G. K. Yang, “Interfacing chatbot with data retrieval and analytics queries for decision making,” in RITA 2018: Proceedings of the 6th International Conference on Robot Intelligence Technology and Applications, Springer, 2020, pp. 385–394.

[15] A. Islam and K. Chang, “Real-time AI-based informational decision-making support system utilizing dynamic text sources,” Applied Sciences, vol. 11, no. 13, p. 6237, 2021.

[16] E. Dimara, H. Zhang, M. Tory, and S. Franconeri, “The unmet data visualization needs of decision makers within organizations,” IEEE Trans Vis Comput Graph, vol. 28, no. 12, pp. 4101–4112, 2021.

[17] E. Kavaz, A. Puig, and I. Rodríguez, “Chatbot-based natural language interfaces for data visualisation: A scoping review,” Applied Sciences, vol. 13, no. 12, p. 7025, 2023.

[18] K. Majhadi and M. Machkour, "CHAT-SQL: Natural Language Text to SQL Queries Based on Deep Learning Techniques," Journal of Theoretical and Applied Information Technology, vol. 102, no. 12, pp. 5052-5061, Jun. 2024.

[19] A. Wong, L. Pham, Y. Lee, S. Chan, R. Sadaya, Y. Khmelevsky, M. Clement, F. W. Y. Cheng, J. Mahony, and M. Ferri, "Translating Natural Language Queries to SQL Using the T5 Model," in 2024 IEEE International Systems Conference (SysCon), Vancouver, Canada, Apr. 2024, pp. 1–8. DOI: 10.1109/SysCon61195.2024.10553509.

[20] M. D. Jadhav, A. B. Shelar, and S. More, "HR People Data Analytics Using Power BI Data Visualization Tool," International Journal of Research Publication and Reviews, vol. 3, no. 4, pp. 3335–3338, May 2022.

Appendix

Appendix 1: Data Samples for Training the NLP Model of NL Query to SQL

NL Query	SQL Query
Show me all users who registered today.	SELECT * FROM users WHERE registration_date= CURRENT_DATE;
List all products with a price above $50.	SELECT * FROM products WHERE price > 50;
How many orders were placed last month?	SELECT COUNT(*) FROM orders WHERE order_date >= DATE_SUB (CURRENT_DATE, INTERVAL 1 MONTH);
Get the names of employees in the sales department.	SELECT name FROM employees WHERE department = 'sales';
What is the average salary in the IT department?	SELECT AVG(salary) FROM employees WHERE department = 'IT';
Show all orders with a total amount greater than $1000.	SELECT * FROM orders WHERE total_amount > 1000;
List the top 10 most expensive products.	SELECT * FROM products ORDER BY price DESC LIMIT 10;

Find all customers from New York.	SELECT * FROM customers WHERE city = 'New York';
How many users signed up in the last week?	SELECT COUNT(*) FROM users WHERE signup_date >= DATE_SUB (CURRENT_DATE, INTERVAL 1 WEEK);

The dataset used to train the NLP model was carefully curated to encompass a wide range of queries commonly encountered in HR and business environments. By exposing the model to a diverse set of natural language queries and their corresponding SQL translations, the dataset served as the foundation for the model's ability to interpret complex user queries and generate accurate SQL commands. The sample data presented here highlights various types of queries that the model was trained on, providing insight into its versatility and adaptability.

Structure and Diversity of Queries

The training dataset consists of two primary columns:

- Natural Language Queries (NL Query): These are typical user-input queries expressed in natural language. They represent the type of questions that a non-technical user might ask the chatbot to retrieve information from a database.

- SQL Queries (SQL Query): These are the corresponding SQL commands generated to retrieve the desired information from the database. Each SQL query is structured to match the intent of the NL query, demonstrating the model's ability to handle various SQL operations.

The dataset was designed to cover different categories of queries, including:

Basic Data Retrieval: Queries that involve simple selection operations, such as retrieving a list of records that meet specific criteria. For example:

NL Query: "Show me all employee data

SQL Query: SELECT * FROM employee.

This type of query allowed the model to learn how to retrieve data based on specific fields and conditions.

1. **Conditional Filtering**: Queries that use conditions to filter records, such as retrieving products above a certain price or orders above a specified amount. For instance:

NL Query: "List all who promoted in 2023."

SQL Query: SELECT * FROM employee WHERE YEAR(last_promotion_date) = 2023;

These examples teach the model to handle conditional filtering using operators such as >, <, and =.

2. **Aggregation Functions**: Queries involving aggregation, such as counting records, calculating averages, and summing values. For example:

NL Query: "What is the average salary in the IT department?"

SQL Query: SELECT AVG(salary) FROM employees WHERE department = 'IT';

Aggregation functions like COUNT, AVG, SUM, MIN, and MAX are crucial in HR and business analytics, as they allow users to gain insights into data trends and summaries.

3. **Sorting and Limiting Results**: Queries that involve sorting data by specific criteria and limiting the number of results returned, such as listing the top 10 most expensive products. For instance:

NL Query: "List the top 10 most expensive products."

SQL Query: SELECT * FROM products ORDER BY price DESC LIMIT 10;

These queries allowed the model to learn SQL operations related to ordering (ORDER BY) and limiting results (LIMIT), which are common requirements in business and HR analytics for ranking and focusing on specific data points.

Significance of the Dataset in Model Training

The diversity of queries in this dataset played a crucial role in helping the NLP model develop a robust understanding of SQL operations. Through training on this data, the model learned to handle various types of user requests, enabling it to generate SQL queries that accurately reflect the user's intent. The significance of each query type is summarized below:

- **Enhanced Query Interpretation**: By training on a dataset that includes a broad range of query types, the model becomes adept at recognizing different intents in user queries. Whether the query involves data retrieval, aggregation, filtering, or sorting, the model learns to identify the underlying action and translate it into the appropriate SQL command.

- **Handling Contextual Language**: Exposure to time-based and condition- based queries teaches the model to understand contextual language, such as temporal references ("last month," "today") and comparison conditions ("greater than," "top 10"). This capability is essential for real-world applications, as users often rely on everyday language rather than technical SQL terminology.

- **Accurate Data Retrieval for Decision-Making**: For HR and business applications, accurate data retrieval is critical for informed decision-making. This dataset allows the model to generate SQL queries that retrieve precise information, supporting HR professionals and business analysts in accessing the data they need quickly and reliably.

Challenges in Dataset Creation and Model Training

Creating a dataset tailored to HR-specific applications presented several challenges, as it required careful consideration of the types of HR queries the model would need to interpret and handle. Key challenges included:

- **Balancing Simplicity and Complexity**: While simple HR queries, such as "Show all employees in the marketing department," are straightforward for the model to learn, more complex queries reflect real-world HR needs better. Complex queries often include multiple conditions, time-based constraints, or aggregate functions. For example, a query like "List all employees in the sales department who were promoted in the last year and have a performance score above 80" involves filtering by department, time frame, and performance criteria. Ensuring the dataset had a balanced mix of simple and complex HR queries was essential to prepare the model for diverse, realistic use cases.

- **Capturing Variations in Language**: HR professionals may phrase similar requests in different ways depending on their individual preferences or organizational language. For instance, a request for "List top performers in the last quarter" could also be phrased as "Show employees with the highest performance scores for Q1" or "Who were the high achievers last quarter?" Ensuring that the dataset included these variations helped the model generalize across different phrasings, improving its ability to interpret HR queries accurately, regardless of wording.

- **Addressing Nuances in HR Metrics and Terminology**: HR queries often involve nuanced terms that have specific meanings within the context of human resources. Terms like "engagement," "turnover," "performance scores," and "retention rate" each have distinct interpretations and associated metrics. For example, "employee engagement" may refer to survey results or specific scores calculated based on multiple factors. Ensuring the dataset captured these nuanced HR terms and that the model could differentiate between them was critical for accurate SQL generation and data retrieval.

- **Representing Diverse HR Scenarios**: HR analytics cover a wide range of scenarios, from workforce demographics and employee satisfaction to performance tracking and succession planning. The dataset needed to include queries related to key HR metrics such as employee engagement, turnover, absenteeism, and performance ratings. This diversity allowed the model to handle various HR-focused requests, making it adaptable to the specific analytical needs of HR departments.

- Handling Time-Based Queries Specific to HR Contexts: Many HR queries are time-sensitive, asking for data over specific periods, such as "last month," "last quarter," or "over the past three years." HR professionals often analyze trends and changes over time to track workforce dynamics or evaluate program effectiveness. Training the model to understand time-based HR queries, such as "Show turnover rates for the past year" or "List employees promoted in the last six months," required careful inclusion of time-based conditions in the dataset to ensure accurate SQL translation for these types of requests.

Tokenization and Preprocessing

Tokenization is a crucial pre-processing step, transforming raw text data into token sequences that the model can process. Since the model needs to understand both natural language and SQL syntax, the tokenization process used a pre-trained BART tokenizer, which is optimized for sequence-to-sequence tasks.

```
# Initialize tokenizer and model
tokenizer = BartTokenizer.from_pretrained("facebook/bart-base")

# Preprocess function for tokenization
def preprocess_function(examples):
    inputs = tokenizer(examples['nl_query'], max_length=128, truncation=True, padding='max_length')
    targets = tokenizer(examples['sql_query'], max_length=128, truncation=True, padding='max_length')
    inputs['labels'] = targets['input_ids']
    return inputs

# Apply preprocessing
tokenized_dataset = dataset.map(preprocess_function, batched=True)
```

The preprocess_function tokenizes the NL queries and SQL commands separately. By setting a maximum length of 128 and padding each sequence to this length, we ensure consistency in input size, which is essential for efficient model training. The function also assigns the tokenized SQL sequences as labels for the model, allowing it to learn the mapping from NL queries to SQL outputs.

Model Configuration

The model used for this task was BART (Bidirectional and Auto-Regressive Transformers), which is effective for generative tasks such as language translation. In this case, the BART model was fine-tuned to translate natural language queries into SQL statements.

```
model = BartForConditionalGeneration.from_pretrained("facebook/bart-base")

# Model configuration for BERT tokenizer
config = BertConfig(
    vocab_size=tokenizer.vocab_size,
    hidden_size=768,
    num_hidden_layers=6,  # Adjust the number of layers as needed
    num_attention_heads=12,
    intermediate_size=3072,
    max_position_embeddings=128,
)
```

The model configuration includes parameters such as hidden_size, num_hidden_layers, and num_attention_heads, which define the structure of the model. These hyperparameters control the complexity of the model and affect its ability to capture relationships between NL and SQL sequences.

Training Arguments and Hyperparameter Tuning

To optimize the training process, we defined several key training arguments, including the learning rate, batch size, and the number of epochs. The learning rate was set to a low value of 2e-5, enabling gradual updates to the model's parameters. Batch sizes for both training and evaluation were set to 8, balancing memory usage with model accuracy. Training was conducted over 20 epochs, allowing the model to converge effectively on the training data.

```
# Training arguments
training_args = TrainingArguments(
    output_dir='./results',
    eval_strategy="epoch",
    learning_rate=2e-5,
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    num_train_epochs=20,  # Increase the number of epochs
    weight_decay=0.01,
    save_total_limit=1,
    save_strategy="epoch",
    logging_dir='./logs',
)
```

These arguments play a critical role in model training:

Learning Rate: Controls the step size during optimization. A smaller rate was chosen to prevent overshooting the optimal point.

Epochs: Increasing the number of epochs to 20 allows the model more passes over the data, which helps in refining weights based on diverse HR queries.

Batch Size: A batch size of 8 provides a good trade-off between computation efficiency and memory usage, especially for complex models like BART.

Model Training and Evaluation

The training process was managed using Hugging Face's Trainer class, which simplifies the training workflow by handling backpropagation, optimization, and evaluation. A DataCollatorForSeq2Seq was used to handle padding and label alignment during training, ensuring consistency across input-output pairs.

```
# Data collator for seq2seq
data_collator = DataCollatorForSeq2Seq(tokenizer, model=model)

# Trainer setup
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=tokenized_dataset['train'],
    eval_dataset=tokenized_dataset['test'],
    data_collator=data_collator,
)

# Train the model
trainer.train()
```

During training, the model's performance was evaluated at each epoch, with metrics such as **Training Loss** and **Validation Loss** tracked to monitor progress. As seen in the training output:

Epoch	Training Loss	Validation Loss
5	No log	0.105229
1	No log	2.281307
5	No log	0.105229
10	1.289300	0.076976

15	0.040900	0.069606
20	0.040900	0.065830

The **Validation Loss** decreased consistently over time, indicating that the model was learning to generalize well on unseen data. By the end of training, the validation loss had stabilized, suggesting that the model reached an optimal performance level.

The NLP model's training process demonstrates a comprehensive approach to fine- tuning a sequence-to-sequence model for HR-related query-to-SQL translation. Key aspects of the process, such as careful dataset preparation, robust tokenization, and targeted hyperparameter tuning, allowed the model to learn from a diverse set of HR queries. By handling pre-processing, defining training configurations, and evaluating on test data, the model achieved a significant reduction in validation loss, indicating its effectiveness in translating NL queries into SQL commands with high accuracy.

Glossary of Terms and Abbreviations

1. BART (Bidirectional and Auto-Regressive Transformers)

A transformer-based model developed by Facebook AI that combines the strengths of BERT and GPT architectures. BART is effective for generative tasks such as language translation, summarization, and sequence-to-sequence transformations, making it suitable for converting natural language queries into SQL.

2. BERT (Bidirectional Encoder Representations from Transformers)

A transformer-based language model created by Google that understands the context of words by reading bidirectionally. BERT is used in many NLP applications, especially those that require understanding the relationships between words in a sentence.

3. Data Collator

A function or utility that processes input data into batches, handling padding and alignment of input and output sequences to ensure they have consistent lengths for model training.

4. Engagement Score

A metric commonly used in HR analytics to measure employee engagement levels. It often incorporates data from surveys, feedback, and performance metrics, providing insights into employees' commitment, satisfaction, and motivation.

5. Epoch

In machine learning, an epoch refers to one complete pass of the training dataset through the model. Multiple epochs are typically used to improve the model's performance as it refines weights with each pass.

6. HR Analytics

The process of collecting, analyzing, and reporting HR data to gain insights into workforce trends, predict future outcomes, and support strategic decision-making in human resources.

7. Learning Rate

A hyperparameter that controls the step size in the optimization process during model training. A smaller learning rate helps the model converge to an optimal solution by making smaller adjustments to weights.

8. Loss Function

A metric that quantifies the difference between the model's predicted output and the actual target output. Common loss functions include cross-entropy and mean squared error, and they help guide the optimization of model parameters.

9. Natural Language Processing (NLP)

A field of artificial intelligence focused on the interaction between computers and humans through natural language. NLP enables computers to understand, interpret, and respond to human language in a meaningful way.

10. SQL (Structured Query Language)

A programming language used to manage and manipulate databases. SQL commands allow users to retrieve, update, and organize data within relational databases.

11. Tokenization

The process of breaking down text into smaller components, called tokens, which can be individual words, characters, or subwords. Tokenization is a pre-processing step that prepares text for analysis by NLP models.

12. Training Loss

A measure of how well a model fits the training data. It is calculated after each batch or epoch, with lower values indicating better performance on the training set.

13. Turnover Rate

A metric used in HR to measure the rate at which employees leave an organization over a specified period. It is an essential metric for evaluating workforce stability and organizational health.

14. Validation Loss

A metric that evaluates model performance on a separate validation dataset (not used for training). A decreasing validation loss over epochs indicates improved model generalization.

15. Weight Decay

A regularization technique applied to prevent overfitting in machine learning models. Weight decay penalizes large weight values, encouraging the model to focus on essential features in the data.

16. Data Warehouse

17. A centralized repository that stores large volumes of data from various sources, allowing organizations to analyze and generate insights from historical data.

18. Decision Support System (DSS)

A computer-based system that helps with decision-making by analyzing large volumes of data and providing actionable insights.

19. Data Visualization

The graphical representation of data to help users understand complex information through charts, graphs, and dashboards.

20. Query Ambiguity

Occurs when a query is unclear or has multiple interpretations. In this project, the chatbot handles ambiguity by prompting the user for clarification.

21. Hyperparameter Tuning

The process of adjusting parameters in machine learning models to improve their performance. Common hyperparameters include learning rate, batch size, and the number of training epochs.

22. Fine-Tuning

A technique in model training where a pre-trained model is further trained on a specific dataset to improve its performance on a related task.

23. Employee Performance Metrics

Key indicators that measure an employee's work performance, including productivity, quality of work, and adherence to organizational goals.

Historical Data

Data collected over time, which can be analyzed to identify trends, patterns, and insights. This data is crucial in understanding long-term changes and making informed decisions.

24. Natural Language Query (NL Query)

A question or command phrased in human language, intended to be interpreted by the chatbot and converted into SQL for data retrieval.

25. Seq2Seq (Sequence-to-Sequence)

A type of model architecture used for transforming an input sequence (like natural language) into an output sequence (such as SQL). Seq2Seq models are widely used in tasks like language translation.

26. Query Interpretation

The process by which the NLP model understands and extracts meaning from a natural language query, identifying entities, metrics, and conditions necessary for accurate data retrieval.

27. Transformer Model

A neural network architecture designed for handling sequential data by using self- attention mechanisms. BERT and BART are examples of transformer models used in NLP.

Data Schema

The structure of a database that defines how data is organized and stored, including tables, fields, and relationships. Accurate schema mapping is critical for generating SQL queries from NL inputs.