Exploitation of Natural Language Processing for Financial Audits

MSc Research Project
MSc. Artificial Intelligence for Business

Joshua Komolafe
Student ID: x22124926

School of Computing
National College of Ireland

Supervisor: Victor Del Rosal

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Joshua Komolafe
Student ID:	x22124926
Programme:	MSc. Artificial Intelligence for Business **Year:** 2023
Module:	MSc. Research Practicum
Supervisor:	Victor Del Rosal
Submission Due Date:	31/01/2024
Project Title:	Exploitation of Natural Language Processing for Financial Audits
Word Count:	4247 **Page Count** 15

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……………………………………………………………………………………

Date: 31/01/2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Exploitation of Natural Language Processing for Financial Audits

Joshua Komolafe
x22124926

Abstract

Financial audits play a crucial role in ensuring reliable corporate financial reporting. However, manually reviewing increasing volumes of textual data pose challenges to Auditors in thoroughly and consistently examining all relevant information. This research explores leveraging Natural Language Processing (NLP) to enhance the audit process. Reviewed literature identifies that NLP shows promise in areas like risk assessment, fraud detection, and workpaper review. However, most prior studies focus on limited datasets and lack integration into real audits. To address this, an NLP solution called SwiftAuditAI is designed using a large language model. The bot analyses financial statements and responds to audit questions to evaluate its ability to support audits. A comparison finds that SwiftAuditAI completes verification and extraction tasks on average 7 times faster than a human Auditor. For a sample of 20 questions, the bot's responses achieved 95% contextual accuracy compared to a human Auditor based on cosine similarity scoring of BERT embeddings.

1 Introduction

The information technology industry has always been through disruptions by the invention of new technologies and these technologies often influence other fields, professions, or industries. The current trending technology is the Artificial Intelligence (AI) technology with a whole lot of excitement and expectation around it. As a result of the promises and expectations around AI, several types of businesses and industries have already explored or are considering exploring AI and are altering their strategies, products, and procedures to better have a competitive advantage. One important area where AI is being explored is the Auditing profession. Financial audits often entail examining a company's financial records and statements to provide a reasonable assurance that they accurately represent the true and fair financial position (Fisher, Garnsey and Hughes, 2016).

However, the volume and complexity of textual data to be reviewed have increased enormously, making manual review time-consuming, inconsistent, and potentially less effective (Sifa et al., 2018). While Computer Aided Auditing Techniques (CAATs) has been leveraged by Auditors in certain areas of audit, there is a substantial need for technologies such as Natural Language Processing (NLP) that can assist Auditors to review bulky relevant financial and non-financial documents with potential speed and accuracy. NLP provides new opportunities to exploit text analytics to improve financial statement audits. NLP is a field of AI that gives computers the ability to understand, interpret, and derive meaning from human language. NLP is considered a cognitive technique that comes from AI, according to Osika (2021). The power of NLP is supported by extremely intricate computer techniques that employ machine learning and statistical modeling. NLP techniques such as semantic analysis, topic modeling, and document classification have been used to extract deeper meaning from textual data and detect patterns indicative of high risk (Hajek, Olej & Myskova, 2014; Fissette et al., 2017).

Researchers have explored various NLP methods to support different aspects of financial auditing. A key application is analysing the textual content of financial reports and other corporate disclosures to identify linguistic cues that may indicate financial misreporting or fraud risk (Goel et al., 2010; Cecchini et al., 2010). Other studies have focused on using NLP to improve the efficiency and consistency of manual review procedures in areas such as audit workpaper review (Arnold et al., 2012). According to Issa, Sun and Vasarhelyi (2016), the Big 4 audit firms: Deloitte, EY, PwC and KPMG have already started leveraging AI in their audit and assurance services. However, smaller audit firms and internal audit departments might not have the resources and capacity of these big firms. Therefore, an easily accessible NLP solution that cuts across board is required.

Overall, NLP shows strong potential to enhance financial auditing through automated analysis of unstructured textual data. However, most existing studies have focused on experimental applications using limited datasets. Additional research is needed to transition NLP techniques into practical tools that can be integrated into real-world audit workflows. Challenges include adapting NLP algorithms to handle noisy and complex textual data, integrating NLP with Auditors' domain knowledge, and providing appropriate transparency and validation of NLP systems (Vasarhelyi et al., 2015). Addressing these challenges can help unlock the benefits of NLP for improved audit quality, efficiency, and risk assessment. This research aims to provide guidance to a practical use of NLP in the auditing profession with little or no coding taking into consideration that Auditors often don't have coding experience. This research's main contribution is a novel and less resources intensive exploitation of NLP for a speedy and accurate extraction of information from financial documents. To scientifically ascertain that NLP would make the Auditors job easier, a competition between AI and a human is carried out to ascertain the speed and accuracy with which tasks are performed with a bot called SwiftAuditAI[1] designed and hosted on Poe[2] powered by Claude-2-100k, a Large Language Model (LLM) that is owned by Anthropic.

The exploitation of NLP in related works is discussed in section 2 of this research. The research methodology used is discussed in section 3. Section 4 discusses the design components for the NLP solution. Section 5 entails the discussion of the implementation of this research. Section 6 presents and discusses the evaluation results. The research conclusion and discussion of future works are contained in Section 7.

[1] https://poe.com/SwiftAuditAI
[2] https://poe.com

2 Related Work

NLP has been explored on how it can improve the auditing profession. Rudžionis et al. (2022) investigated the implementation of NLP-based algorithms to identify irregular operations using comments left by Accountants. Using cosine similarity which measures semantic similarities between texts, the researchers implemented content analysis to analyse over 500 million financial records of Dutch companies. Following the analysis of comments and financial data, it was discovered that 0.3% of operations are likely suspicious. This provided a lead that using NLP can immensely cut the number of probable suspicious operations (Rudžionis et al., 2022). The strengths of the study are the use of a practical application, the use of real-world datasets and content analysis approach, while its limitations are generalizability and lack of comparison, also about 40% of the documents have no Accountants' comments.

Fisher, Garnsey and Hughes (2016) studied NLP in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research. The methodology the researchers used entails a systematic approach to choices of literature. The authors initially scan the bibliographies of four literature reviews on NLP in accounting, auditing, and finance. They then perform keyword searches in databases such as the ACM Digital Library, ProQuest, and EBSCO Host. The selected literature is assessed and categorized based on its relevance to NLP combined with AI and ML. The research also provides a chronological overview of the evolution of NLP research in the field. The result of the research identifies two major uses of NLP to be classification and prediction. Additionally, the study discusses frequently observed applications and data sources, as well as readability studies in the field. The strengths of the study are a comprehensive literature review, identification of research gaps and the use of a systematic methodology for literature selection. However, the weakness of the study includes lack of primary research or empirical analysis. It solely focuses on reviewing and synthesizing existing literature. Found limitation is the bias in literature selection: The study's reliance on existing literature reviews and keyword searches may introduce bias and overlook relevant studies not included in the selected sources.

A research paper by Zemankova (2019) on AI in audit and accounting provides a comprehensive analysis of current trends, opportunities, and threats. It highlights the significance of using AI in accounting and the audit process. The paper identifies seven essential audit tasks that benefit from AI implementation, with a focus on risk assessment. It confirms the use of genetic algorithms, fuzzy systems, neural networks, and hybrid systems as the most used technologies. The synthesis of expert systems and neural networks is found to be the most successful combination. The paper also summarizes the latest AI tools and innovations developed by the Big 4 companies, mainly for audit planning, benchmarking, and document analysis. However, the paper lacks a detailed explanation of the research methodology used and does not address potential limitations or ethical implications. Overall, it provides valuable insights into the current state of AI in audit and accounting but could benefit from more specific examples and a clearer research methodology.

Li and Liu (2019) proposed an intelligent NLP-based audit plan knowledge discovery system (APKDS) that collects and extracts important contents from audit brainstorming discussions. It addresses the challenge of retrieving valuable knowledge from spoken dialogs in audit engagements. The proposed system integrates NLP technologies and aims to provide decision support for future audit plan engagements. However, the research lacks detailed explanations of the NLP modules and empirical evidence to validate the system's effectiveness. It also does not discuss potential limitations. The results highlight the potential benefits of the proposed system in enhancing brainstorming session effectiveness and providing informational support to auditors. The research methodology includes a literature review and the proposal of a conceptual framework but lacks details on data collection and analysis methods. Sample characteristics are also not provided.

Jain et al. (2019) wrote a research paper focused on the development of an AI platform called AppZen for finance teams to audit employee expenses. The authors highlight the challenges in processing and analysing different types of transaction documents, such as receipts, invoices, contracts, and purchase orders. They propose the use of NLP and a lightweight semantic layer to address these challenges. The paper emphasizes the effectiveness of their approach in detecting unauthorized expenses, with a high accuracy rate (>85%), leading to significant cost and time savings for organizations. However, the research has limitations as it is specific to expense auditing and the AppZen platform. The scalability and applicability to other domains are not discussed. The research methodology involves data collection, NLP techniques, statistical models, and semantic analysis. Data was collected by annotating receipts and utilizing DBpedia for unauthorized expense categories. NLP and statistical models were used for entity extraction and semantic analysis. The research lacks a comprehensive evaluation and does not discuss potential limitations or drawbacks. Overall, the paper provides valuable insights into using NLP for financial transaction document understanding, but further research is needed to address the limitations and evaluate the approach more comprehensively.

Mayer et al. (2020) carried out a research titled 'Towards Natural Language Processing: An Accounting Case Study'. The research paper addresses a relevant and practical use case of applying NLP for data extraction from physical leasing contracts. The research study is based on a case study approach, involving discussions with an Auditor and a Manager from a chemical company to evaluate the design guidelines. The paper follows the Design Science Research (DSR) approach, which is a rigorous methodology for developing and evaluating artifacts. The paper provides design guidelines for implementing NLP in the context of the International Financial Reporting Standard (IFRS) 16, which can be valuable for organizations dealing with leasing contracts. However, the paper does not provide a detailed discussion of the limitations and challenges of implementing NLP for data extraction from physical leasing contracts as well as not present any empirical evaluation or validation of the proposed design guidelines. Limitations of the study is that it is based on a single case study of a leading international technology group, which may limit the generalizability of the findings. The result of the study is the presentation of design guidelines for implementing NLP in the context of IFRS 16 fulfilment, including steps such as evaluating machine readability, clustering documents, and designing an intuitive user interface.

Research by Theron (2020) focused on identifying financial risk through NLP of company annual reports. The strength lies in addressing a practical problem and developing a pipeline to convert unstructured reports into a novel corpus. Machine learning techniques, including logistic regression and deep learning models, are used to analyse the textual data. Based on accessible wordlists, a bag of words and word embedding techniques were used and enhanced with linguistic qualities such as tone, uncertainty, and casualness. Weaknesses include a lack of detailed explanation of the specific techniques used and the absence of discussion on potential biases and limitations. The research is limited to South African banks and relies on failed financial events for labelling. The impact of external factors is not considered. The results provide a baseline for further research, but specific findings are not mentioned. The research methodology involves feasibility testing, data collection, labelling, and analysis using various techniques. Both traditional and modern models are employed, along with feature selection and validation processes.

Gao and Han (2021) investigated the effects of AI on the goals of auditing financial statements. They contend that AI technologies can improve the guarantee of the dependability of accounting data. The report, however, lacks a thorough literature assessment and empirical evidence to back up its findings. There is no discussion of potential problems or ethical implications, and the approach is not fully stated. The study focuses on the theoretical aspects of AI in auditing rather than concrete conclusions or discoveries. While the topic is interesting, the paper has flaws in the literature review, methodology, and empirical findings.

Research conducted by Karmańska (2022) focused on the benefits of applying AI in the audit profession. The study utilized a questionnaire administered through an online survey, with a sample size of 206 auditing and accounting practitioners and students. The data collected were analysed using a principal axis factor analysis with Promax rotation to assess the underlying structure of the questionnaire. The findings revealed that AI adoption increases audit efficiency, enhances client communication and service, and can automate time-consuming and routine tasks. The research also identified three factors that accounted for 62.223% variance in the data. However, a limitation of the study is its focus solely on respondents from Poland, which limits the generalizability of the findings. Overall, the research methodology employed was appropriate for the research objective, and the study contributes to the understanding of the advantages of AI adoption in the audit profession.

Xiao (2022) carried out research that focused on the applied strategies of business financial audit in the age of AI. It highlights the challenges faced by Auditors due to the rapid development of high and new technologies. The study emphasizes the need for Auditors to adapt proactively to the changing audit environment. However, the study lacks specific details about the research methodology used. The study does not provide specific results of the research and while the research addresses an important topic, the lack of methodological details and limited information on results and limitations weaken the study's credibility.

Liu, Yen and Wu (2022) investigated the associations between the sentiment perceived in key audit matters (KAMs) and current and future firm performance. The researchers used a two-year sample of 1,606 firm-year observations, including manually labelled sentiment data in 2017 and sentiment data extracted using the BERT model in 2018. Positive associations were found between KAM sentiment and current and next-year firm performances, measured by Tobin's Q, ROA, and ROE. However, the evidence of the positive association between KAM sentiment and current firm market performance was weaker in 2017 than in 2018. The study supports the use of the BERT deep learning approach for textual mining. Limitations include the focus on Taiwanese listed firms and the exclusion of other factors influencing firm performance. The study contributes to the literature by demonstrating the informational value of KAM sentiment and suggests implications for investors, auditors, and regulators. The research methodology involves manual labelling of KAMs and the use of regression analysis.

Ugur et al. (2022) carried out research that aimed to develop a financial intelligence system using NLP techniques to analyse online news channels. It successfully addresses various NLP issues and solves them using traditional machine learning techniques and modern deep learning architectures. The research utilizes a dataset obtained from a reliable source, the national official news agency, and the news datasets have been labelled by financial experts. However, it lacks detailed explanations of the specific NLP techniques used and their implementation. The research also does not discuss potential limitations or challenges faced during system development. It focuses on analysing news channels filtered by specific keywords, limiting its scope. The research does not consider other sources of information, such as financial reports. It lacks a comprehensive evaluation of the system's performance and effectiveness. The research successfully develops a financial intelligence system that can understand and analyse news channels, providing potential opportunities and threats for companies. It demonstrates the application of various NLP techniques, but the research methodology lacks detailed explanations of data preprocessing, feature extraction, model creation, and evaluation.

Siciński (2023) evaluated the effectiveness and usability of the ChatGPT application in assessing the financial condition and bankruptcy risk of entities. The study utilizes analysis and synthesis, critical analysis of literature, and an experiment with a NLP application. The findings indicate that the ChatGPT tool has extensive usability and can conduct interactions like human communication. However, the research is limited to simple algorithms and entities with widely available financial statements. The paper lacks specific details about the sample size and experimental design. The research highlights the development potential of NLP technology in management and its potential role as a digital managerial advisor. The ChatGPT language model shows a higher level of training on general data compared to specific fields. The research methodology includes analysis and synthesis, critical analysis of literature, and an experiment with the ChatGPT application. The paper emphasizes the need for an implementation strategy for NLP technology in management. Overall, the research contributes to filling the research gap in the realm of bankruptcy early warning systems and highlights the potential of NLP technology in supporting financial analysis.

Goel (2009) carried out a research study to investigate the qualitative contents of annual reports using NLP techniques to detect fraud. The study concluded that to detect fraud it is essential to examine the qualitative contents of the reports as they tend to provide more leads than the quantitative or financial contents. Goel et al. (2010) further improved on this research with another one with results showing that the use of linguistic features can help to detect fraud. The result was feasible to improve prediction accuracy from 56.75 percent using an earlier proposed NLP-based method and a "bag of words" principle to 89.51 percent by using verbally inspired features drove by informed logic and integrating domain-related insight.

3 Research Methodology

The research methodology from data gathering and data preprocessing to data transformation, modelling and results and evaluation is illustrated in Figure 1 below.

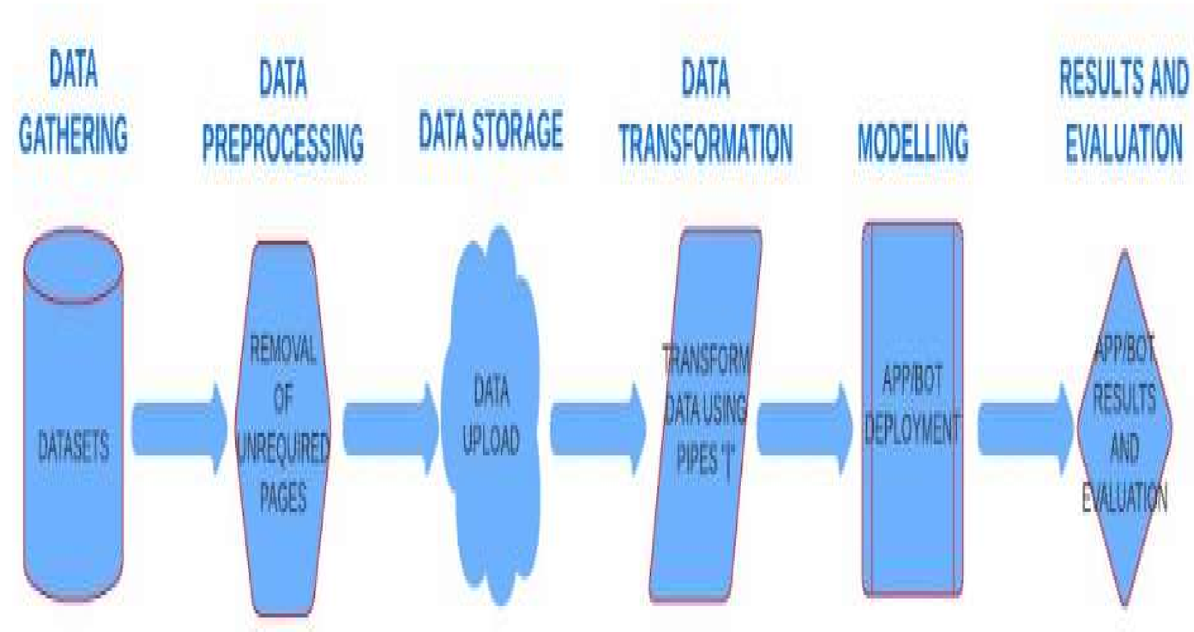

Figure 1. Research Methodology

The initial step entails gathering the required datasets consisting of Dataset1 Suzano Consolidated Financial Statements, Dataset2 International Standard on Auditing (Ireland) 700, Dataset3 International Financial Reporting Standards in your Pocket.

The second step is the process of removing unwanted contents from the dataset. The dataset to be audited was stripped of the Auditors report in other to make the AI solution independent of it.

The third step involves uploading the datasets. The designed bot, SwiftAuditAI which leverages Claude-2-100k LLM on the Poe platform gets the unstructured datasets excluding the dataset to be audited uploaded to its' knowledge base.

For the fourth step, to make financial figures well-arranged and understandable for the prompt engine during the AI audit process, the bot was prompted to arrange multiple financial figures copied from financial statements and pasted into it by it using the symbol "|" (pipe) (on both side of separations) to separate and delineate columns in other to present the financial figures and information in a standard format. In financial statements or tables, this symbol is commonly used to create a clear visual separation between different data columns.

In the fifth step which is the modelling stage, for SwiftAuditAI, the bot behaviour is determined by the given prompts. With Claude-2-100k as its model, the knowledge base comprises of the relevant datasets that the bot will access to inform its responses. The bot will retrieve relevant sections from the knowledge base based on the user message. The custom temperature of the bot which controls the creativity of the bot's responses is set at 0.0/1. This is to ensure that the bot doesn't give abstract responses. The degree of randomness in token selection is controlled by temperature. Higher temperatures can provide more varied or unexpected results; lower temperatures are ideal for prompts that anticipate a true or proper response. The token with the highest probability is always chosen when the temperature is zero.

The sixth step is the Results and Evaluation stage which involves evaluating the performance of the Claude-2-100k model and the human performance using speed and accuracy.

4 Design Specification

The SwiftAuditAI design architecture leverages the Claude-2-100k LLM and the Poe chat user interface as shown in Figure 2 below. The components of Claude-2-100k model entails encoder, decoder, attention mechanism, transformer architecture, pre-trained transformer model. The model also has several features that allows it to be more versatile and powerful (Reever, 2023; Wiggers, 2023).

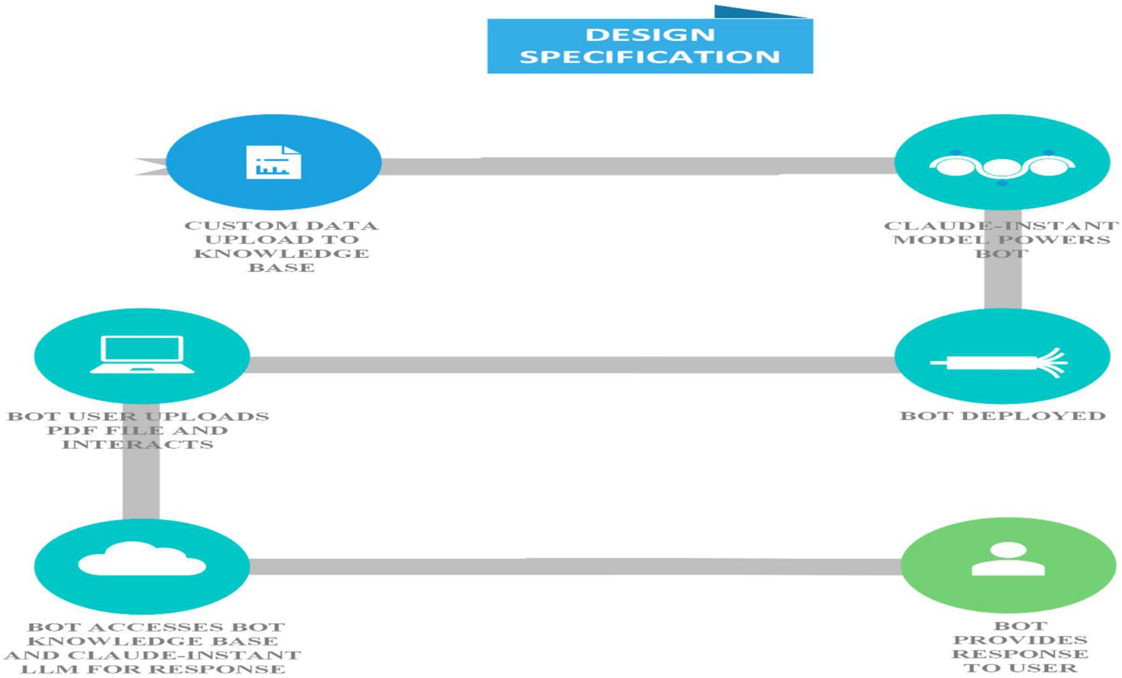

Figure 2. Design Specification

5 Implementation

The SwiftAuditAI was implemented as a bot on the Poe chat platform leveraging the Claude-2-100k LLM. The user interface is the Poe chat user interface shown in Figure 3 below which allows users to upload a PDF containing financial statements for interaction. Based on given questions, answers are generated from the knowledge base of SwiftAuditAI where relevant data has been based as well as Claude-2-100k LLM. The temperature of the bot responses is set at zero to ensure no abstract responses. The bot is given a profile display name and information, custom picture shown in Figure 4 below, custom prompt welcome message. The hardware used has 16GB RAM, a 1TB hard drive, and 12th Gen Intel(R) Core (TM) i7-1255U 1.70 GHz.

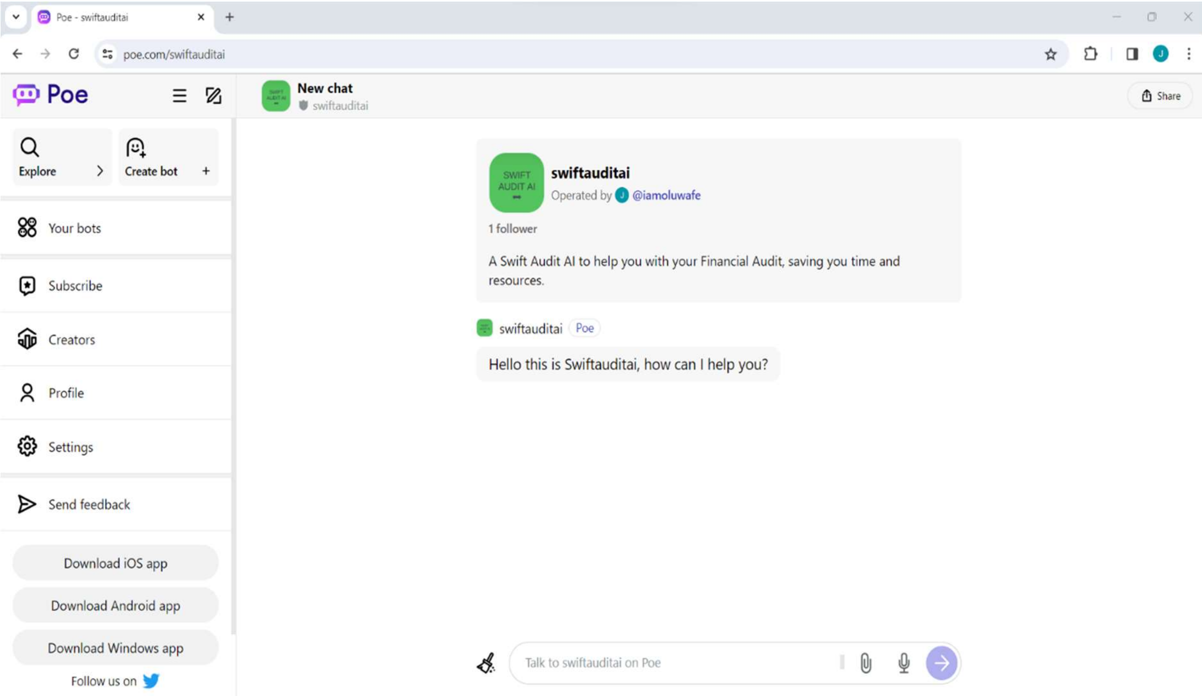

Figure 3. SwiftAuditAI User Interface

Figure 4. SwiftAuditAI Logo

6 Evaluation

The aim of this research is to compare the speed and accuracy of AI and humans in carrying out financial audits. Through the leverage of NLP, the AI bot was able to analyse the unstructured dataset containing the financial information as well as the dataset in the knowledge base thereby being capable of competing with a human Auditor.

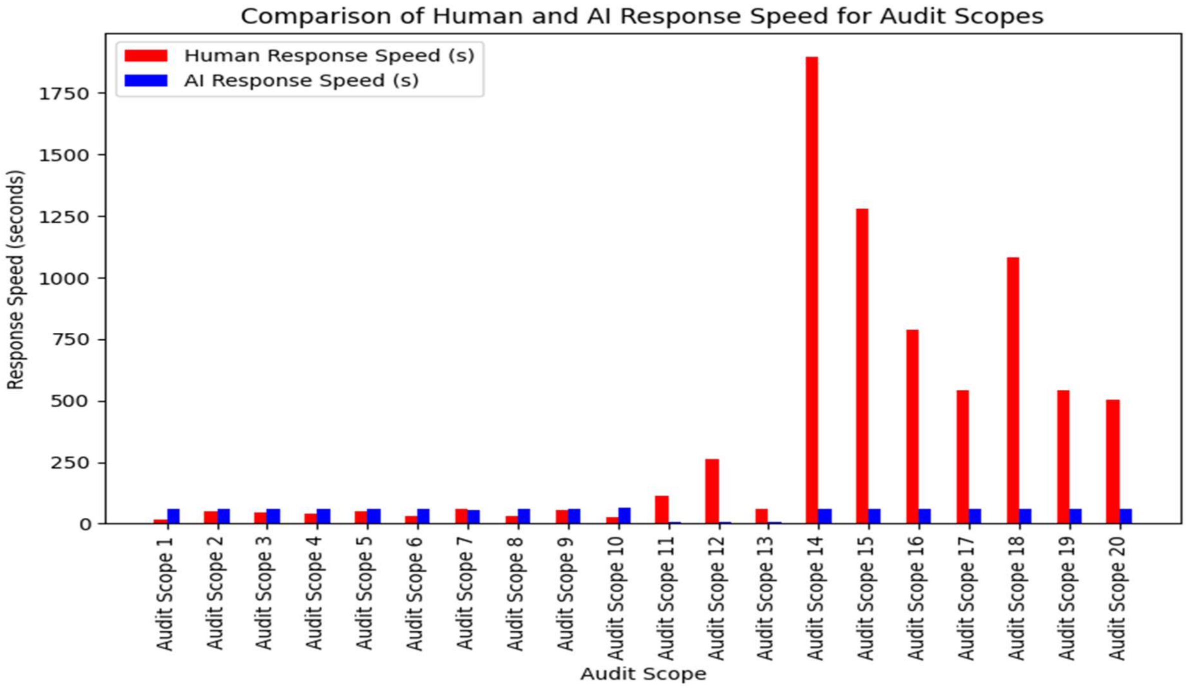

Figure 5. Human vs AI Comparison by Speed.

A comparison of the Human and AI speed is shown in Figure 5 above. Across twenty (20) audit scope questions, the speed at which SwiftAuditAI and a human Auditor complete a task were examined using a timer.

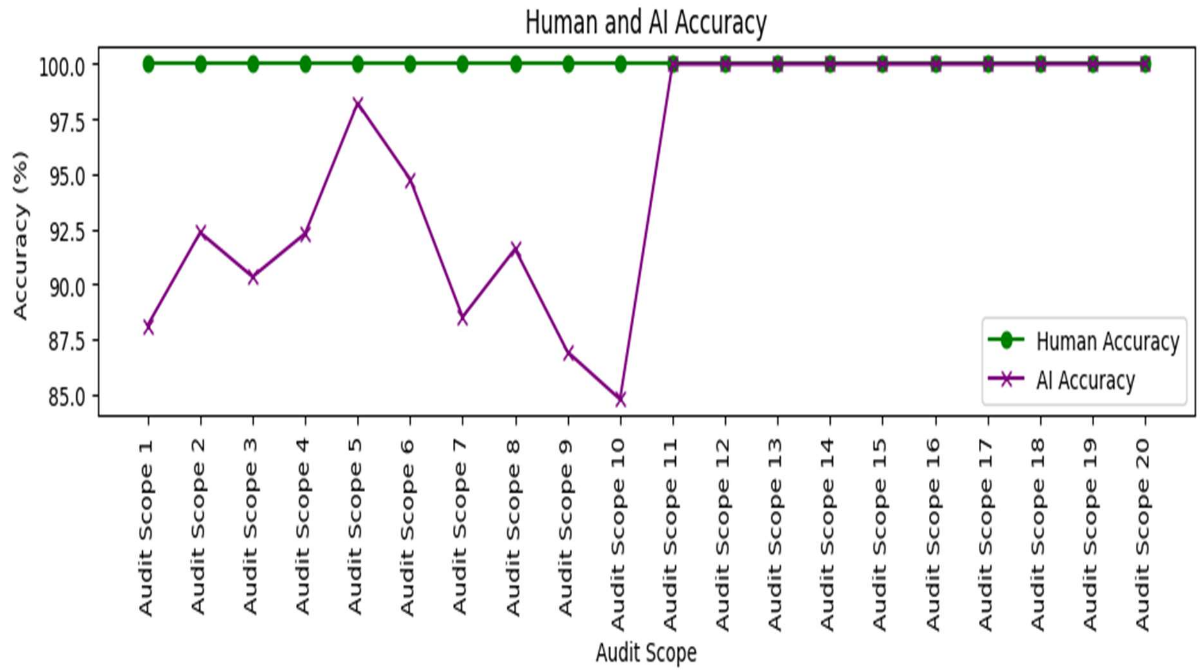

Figure 6. Human vs AI Comparison by Accuracy.

A comparison of the human and AI accuracy is shown in Figure 6 above. Across twenty (20) audit scope questions, the contextual accuracy of SwiftAuditAI is compared to a professional human Auditor and a code is written in Python which uses the BERT model to calculate the contextual accuracy between human and AI responses. The code imports the necessary libraries such as ipywidgets, IPython display, transformers, torch, and sklearn.metrics.pairwise. The bert_encode function takes in text, a pre-trained BERT model, and a tokenizer as input. It returns the mean of the last hidden state of the BERT model after encoding the input text. The code then loads the pre-trained BERT model and tokenizer using the BertModel from_pretrained and BertTokenizer from_pretrained functions respectively. The calculate_bert_contextual_accuracy function takes in a button as input. It retrieves the human and AI responses from the human_response_widget and ai_response_widget respectively. It goes ahead to encode the human and AI responses using the bert_encode function and calculates the cosine similarity between the two embeddings using the cosine_similarity function from sklearn.metrics.pairwise. Finally, it displays the cosine similarity score as a measure of accuracy in percentage (McGregor, 2020; Muller, 2022 & Kumari, 2023).

6.1 Discussion

The experiment is a good practical attempt to provide Auditors with a swift route to leveraging AI for their auditing processes with little or no coding skills and this experiment aimed to leverage on the weakness of previous research works. The NLP solution was experimented based on speed and accuracy with AI having an average speed of 51.5 seconds and human Auditor having an average speed of 373.75 seconds, showing AI is about 7 times faster than human auditor. Furthermore, the human speed was faster in quickly locating where qualitative terms are in the dataset, however the slow pace of AI in this regard is because AI takes its time to read the entire content with an average of 60 seconds. When it gets to calculations and verifications of financial figures, AI beats human Auditors with significant speed.

In the aspect of accuracy, the contextual accuracy of AI response to human Auditor response was compared and results reveal that AI has an average accuracy of 95%. AI was 100% accurate in financial figures and mathematical responses but had the lowest of 85% in text-based responses. This is because human Auditor responses were based on a particular context of the questions while AI goes deeper to provide more insights with related figures and contexts related to these questions in the PDF.

7 Conclusion and Future Work

The aim of this research was to provide a practical guidance for Auditors to leverage NLP for financial audits. This research proposes a framework that uses Claude-instant LLM for the auditing process. Results show that NLP can help Auditors to carry out their responsibilities in little or no time if the motivation is for speed and accuracy. This research can potentially assist Auditors and researchers. The research can be improved upon by developing an audit specific intelligent bot.

References

Arnold, V., Bedard, J.C., Phillips, J.R. and Sutton, S.G. (2012). The Impact of Tagging Qualitative Financial Information on Investor Decision Making: Implications for XBRL. *International Journal of Accounting Information Systems, 13(1), pp.2-20.*

Cecchini, M., Aytug, H., Koehler, G.J. and Pathak, P. (2010). Detecting Management Fraud in Public Companies. *Management Science, 56(7), pp.1146–1160. doi: https://doi.org/10.1287/mnsc.1100.1174.*

Dataset1 (Suzano Consolidated Financial Statements) – Source: https://s201.q4cdn.com/761980458/files/doc_news/2023/02/4Q22/2022-12-DFP-EN.pdf

Dataset2 (International Standard on Auditing (Ireland) 700 Updated October 2023) – Source: https://iaasa.ie/wp-content/uploads/2023/10/ISA_700_UpdatedOct2023.pdf

Dataset3 (International Financial Reporting in your Pocket by Deloitte) – Source: https://www.iasplus.com/en/publications/global/ifrs-in-your-pocket/2023/at_download/file/IFRS%20in%20your%20pocket%20-%202023.pdf

Fisher, I.E., Garnsey, M.R. and Hughes, M.E. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research. *Intelligent Systems in Accounting, Finance and Management, [online] 23(3), pp.157–214. doi: https://doi.org/10.1002/isaf.1386*

Fissette, M. (2017). Text Mining to Detect Indications of Fraud in Annual Reports Worldwide. [online] Available at: https://ris.utwente.nl/ws/portalfiles/portal/20665286/thesis_M_Fissette.pdf [Accessed 11 Dec. 2023].

Gao, Y. and Han, L. (2021). Implications of Artificial Intelligence on the Objectives of Auditing Financial Statements and Ways to Achieve Them. *Microprocessors and Microsystems, p.104036. doi: https://doi.org/10.1016/j.micpro.2021.104036*

Goel, S. (2009). Qualitative Information in Annual Reports & the Detection of Corporate Fraud: A Natural Language Processing Perspective. *[online] ERIC. Available at: https://eric.ed.gov/?id=ED513187 [Accessed 09 Dec. 2023].*

Goel, S., Gangolly, J., Faerman, S.R. and Uzuner, O. (2010). Can Linguistic Predictors Detect Fraudulent Financial Filings? *Journal of Emerging Technologies in Accounting, [online] 7(1), pp.25–46. doi: https://doi.org/10.2308/jeta.2010.7.1.25.*

Hajek, P., Olej, V. and Myskova, R. (2014). Forecasting Corporate Financial Performance using Sentiment in Annual Reports for Stakeholders' Decision-Making. *Technological and Economic Development of Economy, 20(4), pp.721–738. doi: https://doi.org/10.3846/20294913.2014.979456.*

Issa, H., Sun, T. and Vasarhelyi, M.A. (2016). Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. *Journal of Emerging Technologies in Accounting, 13(2), pp.1–20. doi: https://doi.org/10.2308/jeta-10511*

Jain, P., Verma, K., Gaikwad, A. and Gadde, P. (2019). Understanding Financial Transaction Documents using Natural Language Processing. *doi: https://doi.org/10.1145/3360901.3364439.*

Karmańska, A. (2022). Artificial Intelligence in Audit. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, [online] 66(4), pp.87–99. doi: https://doi.org/10.15611/pn.2022.4.06.*

Kumari, K. (2023). Building Language Models: A Step-by-Step BERT Implementation Guide. *[online] Analytics Vidhya. Available at: https://www.analyticsvidhya.com/blog/2023/06/step-by-step-bert-implementation-guide/.*

Li, Q. and Liu, J. (2019). Development of an Intelligent NLP-Based Audit Plan Knowledge Discovery System. *Journal of Emerging Technologies in Accounting, 17(1), pp.89–97. doi: https://doi.org/10.2308/jeta-52665.*

Liu, W.-P., Yen, M.-F. and Wu, T.-Y. (2022). Report Users' Perceived Sentiments of Key Audit Matters and Firm Performance: Evidence from a Deep Learning-Based Natural Language Processing Approach. *Journal of Information Systems. doi: https://doi.org/10.2308/isys-2020-061.*

Mayer, J., Stritzel, O., Eßwein, M. and Quick, R. (2020). Towards Natural Language Processing: An Accounting Case Study Practitioner Paper. *[online] p.1. Available at: https://www.rcw.wi.tu-darmstadt.de/media/bwl4/forschung_9/kompetenzzentrum_1/20200927_ICIS_1383_NLP_track_23_final_version_letter_size.pdf.*

McGregor, M. (2020). Google BERT NLP Machine Learning Tutorial. *[online] Available at: https://www.freecodecamp.org/news/google-bert-nlp-machine-learning-tutorial/.*

Muller, B. (2022). BERT 101 - State Of the Art NLP Model Explained. *[online] huggingface.co. Available at: https://huggingface.co/blog/bert-101.*

Osika, G. (2021). Dilemmas of Social Life Algorithmization - Technological Proof of Equity. *Scientific Papers of Silesian University of Technology Organization and Management Series, 2021(151). doi: https://doi.org/10.29119/1641-3466.2021.151.36.*

Reever, L. (2023). 3 Free Ways to Access Claude — Anthropic's AI Model (Outperformed GPT3.5). *[online] AI Bytes. Available at: https://medium.com/ai-bytes/free-access-claude-ai-5eedadf8f162 [Accessed 01 Dec. 2023].*

Siciński, J. (2023). Early warning systems against bankruptcy risk and NLP: can ChatGPT predict corporate distress? *2023(170), pp.439–455. doi: https://doi.org/10.29119/1641-3466.2023.170.28*

Sifa, R., Ladi, A., Pielka, M., Ramamurthy, R., Hillebrand, L., Kirsch, B., Biesner, D., Stenzel, R., Bell, T. and Lübbering, M.; et al. (2019). Towards Automated Auditing with Machine Learning. *doi: https://doi.org/10.1145/3342558.3345421.*

Theron, J.L. (2020). Identifying Financial Risk through Natural Language Processing of Company Annual Reports. *repository.up.ac.za. [online] Available at: http://hdl.handle.net/2263/83181 [Accessed 12 Dec. 2023*].

Ugur, O., Kalay, T., Demirel O. and Yıldırım, S. (2022). Leveraging the power of Natural Language Processing for Financial Intelligence System. *doi: https://doi.org/10.1109/iisec56263.2022.9998106*

Vasarhelyi, M.A., Kogan, A. and Tuttle, B.M. (2015). Big Data in Accounting: An Overview. *Accounting Horizons, [online] 29(2), pp.381–396. doi: https://doi.org/10.2308/acch-51071.*

Wiggers, K. (2023). Anthropic launches improved version of its entry-level LLM. *[online] TechCrunch. Available at: https://techcrunch.com/2023/08/09/anthropic-launches-improved-version-of-its-entry-level-llm/ [Accessed 14 Dec. 2023].*

Xiao, W. (2022). Research on applied strategies of business financial audit in the age of artificial intelligence. *[online] IEEE Xplore. doi: https://doi.org/10.1109/CIS58238.2022.00098.*

Zemankova, A. (2019). Artificial Intelligence in Audit and Accounting: Development, Current Trends, Opportunities and Threats - Literature Review. *[online] IEEE Xplore. doi: https://doi.org/10.1109/ICCAIRO4*