

Relationship between Data Mining and Audit Quality: Evidence from Big 4 Audit Firms in Nigeria

Ogiriki, Tonye (PhD) and Odogu, Terry Keme Zuode

Accounting Department, Niger Delta University, Amassoma, Bayelsa State, Nigeria.

Accounting Department, University of Africa, Toru-Orua, Bayelsa State, Nigeria

Abstract: *This study was necessitated by the need to eliminate and minimize sampling error and fraud, to improve audit quality in the public sector of Nigeria. Consequently, a descriptive survey design was adopted to investigate the relationship between data mining and audit quality. A sample of 220 practicing auditors were drawn from the Big4 audit firms operating in Nigeria, in line with Krejcie & Morgan's Sample Size Calculator of 1970. The reliability of the research instrument (questionnaire) was determined by Cronbach Alpha test, and questionnaire responses were coded, converted and analysed with Karl Pearson's Moment Correlation at 0.05 level of significance. The results revealed that text and web mining can effectively discover trends, patterns, errors, misrepresentations and fraud in accounting records and reports. This study therefore, concludes that data mining and audit quality are positively related, and that data mining can improve audit quality and reduce the age long audit expectation gap. On this premise, this study recommends that contemporary auditors should acquaint themselves with data mining skills, and audit firms and shareholders should apply and insist on the use of data mining in external audit.*

Keywords: *Data mining, Text mining, web mining, audit quality, Big 4 audit firms, inspired confidence theory.*

INTRODUCTION

The auditing profession has started witnessing new methodologies such as data mining. This is attributable to the high volume and complex-dynamic nature of transactions and data in the third industrial era, and the inadequacies and shortcomings of the traditional sampling approach that interrogates and impairs the reliability and quality of the auditor's report. The extensive deployment of information technology by large contemporary firms and the inappropriate application of sampling has severally occasioned audit failures and performance gap, and has regrettably widened the audit expectation gap between (audited) financial statement users and auditors. Consequently, scholars like Ecaterina (2007), Efiang, Inyang & Joshua (2016), Ogungbade, Adekoya & Olugbodi (2021) argued that auditing in contemporary time requires a modern audit approach and or model such as data mining, to improve

audit quality to regain and increase the confidence of the public. Actually, auditing in the digital age requires a modern alternative approach that can efficiently and holistically process and analyze large, complex and dynamic data, to reduce audit risk to an acceptable level.

Data mining is a powerful model that seamlessly analyzes, classifies and summarizes large volumes of complex data, to discover important events, facts and hidden patterns. This study is thus, motivated by the need for a better data analytical approach that would facilitate the detection and prevention of fraud and errors to improve audit quality, because audit quality is determined by the auditor's capacity to protect the interests of financial statement users by detecting and revealing fraud, errors and information asymmetry between managers and users of financial statements

(Baah & Fogarty, 2018; Elliott, Fanning, & Peecher (2020). Evidence from extant literature reveal that Big 4 audit firms have started deploying data mining in the auditing profession. This calls for an empirical study on the four international audit firms in Nigeria, to ascertain the relationship between audit quality and relevant data mining dynamics like text mining and web mining. The outcome of this study would benefit auditors, shareholders and financial statement users, as it would suggest a better alternative to sampling, and restore the seemingly lost confidence the public once reposed on the auditor.

1.1 Statement of Problem: The relevance of annual audit has been severally interrogated, particularly by external financial statement users and stakeholders, due to the increasing spate of fraud in the public and private sectors of nations. This continuously erodes the confidence of the public on audited financial statements, and seems to have plummeted the once exalted position of

Objectives of the Study:

The specific objectives of this study are:

1. To find out whether text mining can uncover all trends, patterns and errors in accounting records and reports.

the external auditor in the society. This was the major concern in the works of Ecaterina (2007), Efiog, Inyang & Joshua (2016), Ogungbade, Adekoya & Olugbodi (2021). The question is, while the auditor is busy standing on professional definitions and provisions to absolve himself from fraud discovery and prevention, is there no better alternative data analytical model that can reduce sampling errors to bridge the ever increasing audit expectation gap in the auditing profession? This is because the traditional audit approach of sampling is error bound and limits the auditor's chances of discovering fraud and error, and most (unintentional) audit failures are attributable to sampling errors, which occur when samples used in auditing do not absolutely represent large volume of data to achieve or guarantee accuracy. This study is therefore, driven by the necessity for a data mining-based remedy to the escalating incidents of fraud that daily erodes the integrity, credibility and reliability of auditors' report in Nigeria.

2. To find out whether web mining can reveal all trends, patterns and errors in accounting records and reports.

Literature Review

Conceptual Framework and Hypotheses Formation

Meaning of Data Mining:

This is a kind of artificial intelligence technique called computer aided/assisted audit technique (CAAT). Accordingly, data mining can be described as a computerized technique that extracts hidden trends and patterns from massive database. It is a systematic process of discerning meaningful new correlations, patterns and trends by examining and scrutinizing large volumes of data stored in repositories, using statistical, mathematical and pattern recognition techniques and technologies. Data mining as it were, can be defined as the application of

a set of digitalized mathematical techniques, tools and strategies designed to identify and uncover meaningful trends and patterns in big-complex databases. Remarkably, data mining can be alternatively called knowledge discovery, data warehouse and data mart, and is the extraction of hidden predictive and descriptive information from massive quantitative, qualitative, textual and multimedia databanks (Fayyad, Piatetsky-Shapiro & Smyth, 1996; Venkatadri & Reddy, 2011).

The Historical Development of Data Mining:

Historically, data mining began with the need for retrospective and static data collection in the 1960s, which later resulted to the need

for retrospective and dynamic data access in the 1980s

(Fayyad, et al., 1996; Venkatadri & Reddy, 2011). This further necessitated the need for retrospective and dynamic data warehousing for decision support in the 1990s, which eventually informed the need for prospective and proactive information delivery called data mining in the 21st century (Fayyad, et al., 1996; Venkatadri & Reddy, 2011). It must be mentioned however that, the thought of data mining in auditing in contemporary time was partly necessitated by the introduction of the specific requirement to verify the appropriateness of journal entries as part of the auditor's responsibilities in relation to fraud in early 2005, by ISA 240.

Audit Quality: Audit quality, according to Rostami, Samadi, Omrani, Margavi, Asadzadeh & Nazari (2011), is the chance that there is no material omissions,

Significance of Data Mining in Auditing and Accounting Research:

Data mining is regarded by the Institute of Internal Auditors (IIA) as one of the four research primacies in accounting research (Koh & Low, 2004). This is because data mining, according to Pienaar (2012) seamlessly evaluates a variety of patterns, such as association, sequence, path, classification, grouping, and forecasting, to discover past unknown patterns and facts that can be used to predict future behaviour. The Chartered Global Management Accountants (2013) observed that, due to its

Audit Quality and Data Mining:

The definitions of data mining suggest that data mining can detect and prevent fraud and errors in auditing, and can indeed, improve the quality of audit report. This is evident in the work of Wang & Yang (2009) that the use of data mining in auditing facilitates the detection of fraud. This implies that data mining offers auditors the opportunity and privilege to discover organized frauds and misrepresentations in clients' financial

Relevant Dimensions of Data Mining:

Although, data mining is broad in scope and widely practiced, the dimensions of interest

misrepresentation or misstatements in financial statements, and that the financial statements and related documents have been examined, verified and certified to have been prepared according to generally accepted accounting principles, and are therefore reliable, by an external auditor. This implies that a high quality audit report is expected to be materially free from sampling errors and oversight, and assure users of financial statements that the accounting standards and methods applied in the preparation of financial statements of an audited firm are generally acceptable, and the accounting records are relevant, complete, comparable, understandable and faithfully represent the true and fair view of what transpired in the accounting period. The question however is, how reliable is an audit report premised on an error-prone sampling outcome?

importance, about 50% of managers and directors in the private sector rank data mining among the top ten priorities. Congruently, the American Institute of Certified Public Accountant (AICPA) remarked that data mining is one of the top ten trending paradigms of business analytics and decision support for the future (Amanni & Fadlalla, 2017). This thus, is the rationale for a study on the relationship between data mining and audit quality.

transactions and records, as it has the capacity to uncover and interpret hidden patterns to predict the future (Rostami, et al., 2011). It further explains that data mining particularly, text mining and web mining can guide contemporary auditors to make valid and reliable opinion about the financial position and going concern of firms, as it has the capacity to predict the future.

that constitute the independent variables of this study are:

a) Text Mining:

This is a dimension of data mining that transforms unstructured text into a structured format to identify meaningful patterns and new insights. It is the process of extracting structured meaningful numeric data and indices from unstructured scripts and memos (Botham, 2004). Text mining enables auditors to analyse words or cluster of words, for decoding, decryption, interpretation and untangling of hidden connections and affiliations between and amongst variables of interest in fraud examination and investigation (Allahyari,

b) Web Mining:

This is another aspect of data mining that extract and process data from the World Wide Web to discover trends and patterns. Web mining entails the application of automated techniques to extract both structured and unstructured data from web pages, server logs and link structures. It is the scanning and mining of text, graphs and pictures from web pages to ascertain trends and patterns in electronic transactions and documents by auditors. Application areas in auditing include: web personalization, multimedia e-commerce, e-commerce-consumer behaviour. Obviously, web mining

Theoretical Framework

The problem, objectives and variables of this study are further explained by the following theories:

a) The Theory of Inspired Confidence:

This theory proposes that auditors must exercise due professional diligence in their audit assignment to meet the anticipation of a rational outsider, as outsider parties cannot monitor material misstatements, misrepresentation and or fraud in financial statements and reports (Limperg, 1932). This theory therefore, advocates that audit report should be of high quality and reliable in order to satisfy the expectations and gain the confidence of the public. This theory is inspired by the anxiety and skepticism of external stakeholders over the true and

Poiriyeh, Assefi & Safaei, 2017). This suggests that, unlike traditional sampling, text mining can facilitate the discovery of fraud and errors, and would likely improve audit quality and contract the expectation gap between the auditor and financial statement users. However, the first hypothesis of this study is stated in the null form as:

H₀₁: *Text mining cannot uncover all trends, patterns and errors in accounting records and reports.*

enables auditors search, extract and obtain profile and transition trends, patterns and behaviour of e-customers (Zhang & Segall, 2008). It can therefore be argued that web mining can expedite the discovery of fraud and errors, as well as improve audit quality to contract the age long audit expectation gap between the auditor and financial statement users. However, the second hypothesis of this study is stated in the null form as:

H₀₂: *Web mining cannot reveal all trends, patterns and errors in accounting records and reports.*

fairness of the financial records and report prepared by managers of firms they have interest in. The inspired confidence theory therefore, requires auditors to painstakingly examine the books of accounts in a way that will not impair the expectation of external stakeholder. Relatively, this theory subtly requires external auditors to deploy appropriate data mining techniques such as text mining and web mining to facilitate the discovery of trends, patterns, misstatements and misrepresentations for high audit quality.

b) Lending Credibility Theory:

This theory proposes that high quality audit report increases stakeholders' confidence in management's stewardship and credibility (Ecatarina, 2007). It argues that the primary purpose of an audit is to increase the reliability and credibility of financial statements and reports (Okpala, 2015). The lending credibility theory posits that the competitive hedge that endears an auditor to his employer is the quality and reliability of his findings and report. It argues that the confidence of financial statement users is guaranteed by the auditor's professional credibility expressed in his report. Thus, this theory emphasizes high quality audit which is

a function of good data mining skills and practice. Relatively, this theory proposes that data mining dynamics such as text mining and web mining increase audit quality, and therefore, can increase the credibility of financial statements, and satisfy the expectation of financial statement users. The lending credibility theory, though closely related to the inspired confidence theory, aptly and better explains the problem and offers direction to the achievement of the objectives of this study. Consequently, this study is underpinned by the lending credibility theory.

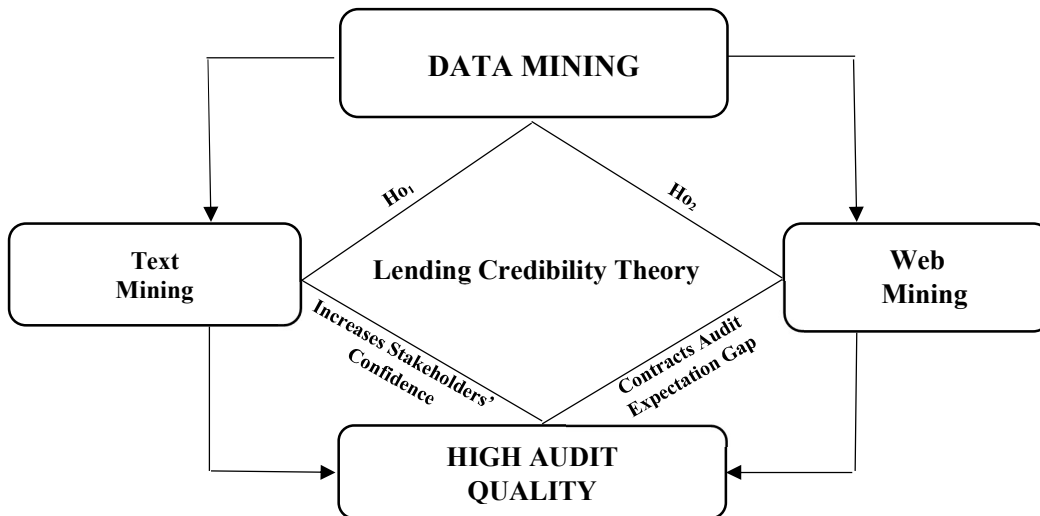


Fig. 1: Conceptual and Theoretical Framework (2022)

Empirical Review:

A search for related literature on data mining and audit quality suggests that data mining, particularly in accounting research is relatively new. This was revealed by the comparatively few empirical studies done in the area. Consequently, this study reviewed only four studies that are germane to its problem, objectives and questions. Firstly, Kirkos, et al. (2007) investigated the effectiveness of data mining classification techniques such as, decision trees, neural networks and bayesian belief networks in detecting firms that issue fraudulent financial statements. The study used ratios derived from financial statements of 76 Greek

manufacturing firms, and found that data mining techniques have advanced classification and prediction capabilities and could facilitate the detection of management fraud. Secondly, a study carried out by Olasami (2013) on computer aided audit techniques and fraud revealed that, with pivotal tables in spreadsheets, auditors can review, analyze and simplify complex data to decode and untangle unknown and unseen facts and information. The study found that auditors can undertake simple data mining with Windows, Linux, and Macintosh operating systems in desktop, notebook and server computers to detect hidden

patterns and fraud (with Excel Spreadsheets). Thirdly, (Amanni & Fadlalla, 2017) explored the applications of data mining techniques using a multi-dimensional framework, namely: retrospective - descriptive, retrospective-prescriptive, prospective-prescriptive, and prospective-predictive data mining application. The framework of the study revealed that data mining is positively related to assurance services, compliance testing, fraud detection, business health and forensic accounting in a significant manner. Remarkably, the findings of Kirkos, et al. (2007), Olasami (2013) and (Amanni & Fadlalla, 2017) suggest that data mining can reveal hidden patterns and trends in accounting records and is positively related to audit quality.

However, the findings of Ezenwafor & Udukeke (2019) do not absolutely or explicitly confirm the above assertion. Ezenwafor & Udukeke (2019) examined the utilization of data mining and anonymous communication

techniques for fraud detection in large scale business organisations in Delta State. They adopted a descriptive survey research design on a population of 260 accounting staff. Using simple random sampling, they obtained primary data from a sample of 160 accountants through a 4-point rating scale questionnaire. The study analyzed data using mean, standard deviation, t-test and analysis of variance to ascertain the homogeneity of the research variables, and to test the hypotheses of the study at 0.05 level of significance. The study found that accountants and auditors in large scale business firms in Delta State of Nigeria are yet to utilize data mining for fraud detection in their practice. Although, this study did not explicitly state the relationship between data mining and fraud detection, it could be assumed and argued that the authors knew its potency but only wanted to know the extent it is applied in Nigeria.

Methodology

Research Design, Data Source, Sample Size and Instrument:

This study deployed survey research design to collect 5-point scale questionnaire responses from a sample of 220 randomly chosen from a population of 516 staff of the Big4 audit firms in Nigeria. The population was ascertained by dividing the total reported staff strength of Big4 audit firms in Nigeria by

the different services they provide (audit and assurance, accounting services, tax planning and preparation, and business development and valuation). The sample size was ascertained from Krejcie & Morgan's Sample Size Determination Table of 1970 at 95% confidence level and 5% margin of error:

Table 1: Population of the Study

S/N	Name of Big 4 Audit Firm	Staff Strength
1	PricewaterhouseCoopers (PwC)	1,000
2	Deloitte Nigeria	600
3	Klynveld Peat Manick Geerdeler (KPMG)	1,053
4	Ernst & Young	410
Gross Population		3,063
Adjusted Population (3,063 / 4)		516

Source: Authors' Compilation from Firms' Website (2022)

Table 2: Sample of the Study

S/N	Confidence Level = 95%; Margin of Error = 5%	
	Population Size	Calculated Sample Size
1	320	175
2	340	181
3	360	186
4	380	191
5	400	196
6	420	201
7	440	205
8	460	210
9	480	214
10	500	217 = 220 <i>Approximately</i>

Source: Extract from Krejcie & Morgan’s Sample Size Calculator (2022)

Validity and Reliability Tests:

The questionnaire of this study was reviewed and vetted by a Fellow Chartered Accountant in the Bursary Unit and an Associate professor in the Accounting Department of university of Africa, Toru-Orua Bayelsa State of Nigeria, for *content validity*, to ensure that the questionnaire items were in line with the

questions and objectives of the study. This study further applied *Cronbach Alpha Statistics* to ascertain the *reliability* of the research instrument, and obtained outcomes which implies that the research instrument is reliable, as shown in Table 3:

Table 3: Reliability Statistics of Research Instrument

S/N	Variables	Alpha	No of Items
1	Audit Quality	0.73	4
2	Data Mining	0.81	6

Source: Authors’ Computation (2022)

Model Specification and Data Analysis:

This study tested and measured one dependent variable and two independent variables. The dependent variable of this study is *Audit Quality*, while the independent as:

variables designed to explain the dependent variable are *Text Mining* and *Web Mining*. This is econometrically specified

$$AUQ = \beta_0 + \beta_1TEM + \beta_2WEM + \epsilon \dots\dots\dots (1)$$

Where:

TEM= Text Mining;

WEM=Web Mining;

AUQ = Audit Quality;

β_0 = Intercept of the Regression;

β_1, β_2 = Coefficients of the Regression, and

ϵ = Error Term capturing other explanatory variables not explicitly included in the model.

This study used a parametric test and analysis based on the nature of data obtained from the questionnaire. Specifically, the study employed Karl Pearson's Product Moment

Correlation and Regression analysis, at 0.05 level of significance (corresponding to 95% level of confidence).

Results and Discussion of Findings

Presentation of Data and Results

This study administered 220 copies of questionnaire on staff of the Big4 audit firms in Nigeria. It retrieved 202 copies, which represents a good (acceptable) proportion,

92% of the total copies of questionnaire administered. The demographic details of questionnaire respondents from the Bayelsa State civil service are presented in Table 4.

Table 4: Demographic characteristics of the Respondents

Demographic Details	Responses	Percentages
Gender:		
▪ Male	132	65%
▪ Female	70	35%
Total	202	100%
Firm:		
▪ PwC (1,000/3,053) = 0.3275 x 202	66	33%
▪ Deloitte Nigeria (600/3,053) = 0.1965 x 202	40	20%
▪ KPMG (1,053/3,053) = 0.3449 x 202	70	34%
▪ Ernst & Young (400/3,053) = 0.1310 x 202	26	13%
Total	202	100%
Educational Qualification:		
▪ HND/B.Sc.	112	55%
▪ MBA/M.Sc.	86	43%
▪ PhD.	4	7%
Total	202	100%
Professional Qualification:		
▪ FCA	48	24%
▪ FCCA	12	6%
▪ ACA	114	56%
▪ ACCA	28	14%
▪ CAN	0	0%
Total	202	100%
Experience:		
▪ 1 to Years	48	24%
▪ 11 to 20 Years	134	66%
▪ 21 Years and Above	20	10%
Total	202	100%

Source: Authors' Field Survey (2022)

Table 4 shows the demographic characteristics of respondents comprising of gender, firm, educational qualification, professional qualification and experience of the respondents. Table 4 also indicates that majority of the respondent are males (65%). It also reveals that majority of the respondent holds HND/B.Sc. certificates (55%) with the

least being PhD holders (7%). Moreover, in professional qualification, 56% of the respondents are Associate Chartered Accountants (ACAs), and majority of respondents (66%) have work experience of 11 to 20 years. The opinions of the 202 respondents are presented and analysed in the following tables:

Table 5: Responses to Questionnaire Items 1, 2, 3 & 4 for Audit Quality

Questionnaire Items	Q ₁		Q ₂		Q ₃		Q ₄	
	Resp	%	Resp	%	Resp	%	Resp	%
Strongly Agree	48	24	84	41	50	25	94	47
Agree	100	49	82	41	90	44	72	35
Undecided	0	0	0	0	4	2	2	1
Disagree	34	17	8	4	42	21	14	7
Strongly Disagree	20	10	28	14	16	8	20	10
Total	202	100	202	100	202	100	202	100

Source: Authors' Computation (2022)

NB:

Q₁: Data mining holistically examines and scrutinizes large-complex quantitative and qualitative data, and can improve audit quality.

Q₂: Data mining eliminates sampling errors and makes better audit reports.

Q₃: The use of data mining increases investors' confidence on audit report.

Q₄: Audit quality is positively related to data mining.

Responses to research questionnaire items 1, 2, 3 and 4 all point to the fact that the dependent variable (audit quality) is related to data mining from which the independent variables are derived from. Specifically, 73%,

82%, 69% and 82% of respondents to Q₁, Q₂, Q₃ and Q₄ agreed at varying degrees that data mining improves audit quality and report, and is positively related to audit quality.

Table 6: Responses to Questionnaire Items 5, 6 & 7 for Text Mining

Questionnaire Items	Q ₅		Q ₆		Q ₇	
	Resp	%	Resp	%	Resp	%
Strongly Agree	100	50	46	23	73	36
Agree	36	17	106	52	71	35
Undecided	6	3	6	3	6	3
Disagree	30	15	14	7	22	11
Strongly Disagree	30	15	30	15	30	15
Total	202	100	202	100	202	100

Source: Authors' Computation (2022)

NB:

Q₅: The use of text mining in statutory audit facilitates the discovery of errors in source documents and subsidiary books.

Q₆: The deployment of text mining in external audit enhances audit trail and explicitly reveals hidden and organized fraud in

financial transactions and accounting records.

Q₇: Text mining is capable of increasing the confidence of financial statement users on audit report.

Table 6 reveals that 67% (50% and 17%) of the total respondents accept the claim of question item 5 that the use of text mining in statutory audit facilitates the discovery of errors in source documents and subsidiary books. 75% (23% and 52%) of respondents also accepted the proposition of questionnaire item 6 that, text mining

enhances audit trail and reveals hidden and organized fraud in financial transactions and accounting records. Again, 71% (35% and 36%) of respondents to questionnaire item 7 submitted that text mining is capable of increasing the confidence of financial statement users on audit report.

Table 7: Responses to Questionnaire Items 8, 9 & 10 for Web Mining

Questionnaire Items	Q ₈		Q ₉		Q ₁₀	
	Resp	%	Resp	%	Resp	%
Strongly Agree	50	25	36	18	43	21
Agree	92	45	104	51	98	49
Undecided	0	0	8	4	4	2
Disagree	20	10	20	10	19	9
Strongly Disagree	40	20	34	17	38	19
Total	202	100	202	100	202	100

Source: Author’s Computation (2022)

NB:

Q₈: *Web mining can reveal e-transaction patterns and trends, and facilitates audit trail and evidence gathering.*

Q₉: *The use of web mining in auditing significantly reduces errors and fraud, and improves audit quality.*

Q₁₀: *The application of web mining can increase financial statement users’ confidence and contract the audit expectation gap.*

Responses to questionnaire item 8 in Table 7 reveals that 70% (25% and 45%) of respondents concurred that web mining facilitates audit trail and evidence gathering, especially in electronic transaction. Moreover, 18% and 52% (69%) of respondents believed that the use of web mining in auditing can significantly errors and fraud, and improves audit quality. Above all, 21% and 49% (70%) of respondents submitted that the application of web mining can increase financial statement users’

confidence and contract the audit expectation gap between auditors and the public.

Responses to the 10 questionnaire items are analysed according to the dependent and independent variables, and the research hypotheses of this study. Questionnaire items 1 to 4 were designed to generate data for the dependent variable (audit quality), while questionnaire items 5, 6, 7 and 8, 9, 10 were designed to generate data for the two independent variables (text mining and web mining).

Test of Hypotheses and Discussion of Findings

The two null hypotheses of this study were tested systematically. First, respondents’ opinions were coded (Strongly Agree = 5; Agree = 4; Undecided = 3; Disagree = 2; Strongly Disagree = 1) and converted to ordinal data for each questionnaire item retrieved from respondents. Secondly, the ordinal data derived from the coded 5-Point

Likert scale opinions of respondents were tested with E-view software for analyses. Finally, the E-view output was analyzed and used to reject or accept the two hypotheses, based on a decision rule that says: *Reject Ho, if P value is < 0.05; Accept Ho, if P value > 0.05.*

Table 8: Multivariate Group Data Set for Hypotheses 1 and 2

Opinion	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
S /Agree	50	50	94	84	32	28	100	46	50	32	566
Agree	98	88	70	82	110	80	34	104	90	102	858
Undecided	0	4	2	0	4	6	6	6	0	0	28
Disagree	34	42	14	8	14	80	30	14	20	20	276
S/Disagree	20	18	42	26	40	8	32	32	40	34	292
Total	202	202	202	202	202	202	202	202	202	202	2,020

Source: Authors’ Computation (2022)

Table 8 is the outcome of coded and converted data from respondents of the big4 audit firms in Nigeria. This was turned into

the E – View Software for multivariate regression test, to test the two hypotheses of this study.

Table 9: Multivariate Regression (Test) Result for Hypotheses 1 and 2

Variables	Coefficients	Std. Errors	t-statistics	Prob.
C	0.90932	0.05761	2.126743	0.0093
TEM	0.128736	1.680482	-0.076607	0.0113
WEM	1.881819	2.715749	0.692928	0.0483
R ² = 0.788258 , F-statistic 19.85449 , Durbin-Watson stat = 2.04654 ,				Prob. = 0.0093

Source: Author's Computation from E-View Version 10 (2022)

Table 9 is a multivariate regression result of the dependent variable, Audit Quality (AUQ) regressed against the independent variables namely: Text Mining (TEM) and Web Mining (WEM). Generally, the Probability (P) value of the constant (C) in Table 9 is 0.0093, which is less than the predetermined alpha value of

0.05 (P = 0.0093 < 0.05), suggesting that, there is a positive significant relationship between data mining and audit quality. This is in line with the decision rule of the study, and is specifically applied in the rejection and or acceptance of the two null hypotheses of the study.

Test of Hypothesis 1 (H₀₁)

The multivariate regression output in Table 9 on the first independent variable, Text Mining (TEM) reveals a probability (P) value of 0.0113, which is less than the predetermined alpha (significant) value of 0.05 (P = 0.0113

< 0.05). This implies a rejection of the first null hypothesis (H₀₁) of the study for the alternate, meaning that: *Text mining can uncover all trends, patterns and errors in accounting records and reports.*

Test of Hypothesis 2 (H₀₂)

The multivariate regression result in Table 9 with respect to the second independent variable, Web Mining (WEM) discloses a probability (P) value of 0.0483, which is below the chosen significant (alpha) value of 0.05 (P = 0.0483 < 0.05). This again,

requires a rejection of the second null hypothesis (H₀₂) of the study for the alternate, meaning that: *web mining reveals trends, Web mining can reveal all trends, patterns and errors in accounting records and reports.*

Discussion on Findings

The coefficient of determination (R²) in the multivariate regression Table (9) indicates that the model is a good fit because the explanatory factors roughly explained 79 % of the link between data mining techniques and audit quality. This implies that, all things being equal, other factors not described by the result, the model accounted for around 21% of the variability in audit quality. This further implies that the regression output meets a priori theoretical expectations.

there is a statistically significant relationship between Text mining (TEM) and Audit quality (AUQ). This is supported by the coefficient in the multiple regression model in Table 9 which indicates that a unit increase in the use of text mining in external audit in the Nigerian public sector would results to a positive increase in audit quality by 0.128736. This implies that the deployment of the data mining can improve audit quality and report in the Nigerian public sector.

Remarkably, the decision on the first hypothesis of this study clearly shows that,

More so, the decision on the second hypothesis of this study reveals that, there is a strong positive link between web mining (WEM) and Audit quality (AUQ). This is informed by the coefficient in the multiple regression model in Table 9 which shows that a unit increase in the deployment of web mining in statutory audit in the public sector of Nigeria would result to a significant improvement in the quality of audited report by 1.881819 points. This therefore implies that, the deployment of data mining can significantly improve audit quality in the Nigerian public sector. The findings and position of this study is in agreement with the empirical findings of

CONCLUSION AND RECOMMENDATIONS

This study concludes from its findings that, the application of data mining in auditing facilitates the discovery of errors, misstatements, misrepresentations and fraud, and would therefore, improve audit quality and secure the confidence of the public. On this premise, this study recommends that:

Kirkos, et al. (2007) who discovered that data mining techniques have advanced classification and prediction capabilities and can facilitate the detection of fraud. The findings of this study is also in line with the empirical position of Olasami (2013) which says that data mining with Windows, Linux, notebook and server computers can help auditors detect hidden patterns and fraud. The findings of this study further confirms the empirical opinion of Fadlalla, et al. (2017) which says that, data mining is positively related to assurance services, compliance testing, fraud detection, business health and forensic accounting in a significant manner.

- Auditors should acquire data mining skills for efficiency and relevance in their career and practice;
- Audit firms should increasingly use data mining to increase and regain the confidence of their clients and the public;
- Shareholders should recommend and insist on the use of data mining in external auditing.

REFERENCES

Allahyari, M., Poiriyeh, S., Assefi, M. & Safaei, S. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. Retrieved on May 12th, 2022 from: <https://www.researchgate.net>>3183.

...

Amanni, F.A. & Fadlalla, A.M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, 24, 32-58.

Baah, G. K., & Fogarty, T. J. (2018). What auditors think about audit quality - A new perspective on an old issue. *Journal of Managerial Issues*, 30(4), 404–504.

Chartered Global Management Accountants (2013), Report from insight into impact: Unlocking opportunities in big data. Retrieved on June 8th, 2022 from: https://www.cgma.org/resources/Reports?Down/unlocking_the_opportunities_in_big_data.pdf

Ecaterina, V. (2007). The theories of audit expectations and the expectations gap. Retrieved on January 22nd, 2022 from: <https://doi.org/https://www.grin.com/document/116911>.

Effiong, E.J. Inyang, I.O. & Joshua, U. (2016). Effectiveness of the mechanisms of fraud prevention and detection in Nigeria. *Advances in Social Sciences Research Journal*, 3(3), 206-218.

- Elliott, W. B., Fanning, K., & Peecher, M. E. (2020). Do investors value higher financial reporting quality, and can expanded audit reports unlock this value? *The Accounting Review*, 95(2), 141–165.
- Ezenwafor, J. I. & Udukeke, O. F. (2019). Utilization of data mining and anonymous communication techniques for fraud detection in large scale business organisations in Delta State. *British Journal of Education*, 7(11), 74-86.
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases, AAAI Press/ MIT Press, Massachusetts Institute of Technology.
- Kirkos, E., Spathis, C. & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32, 995–1003.
- Koh, H. C., & Low, C. K. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19(3), 462–476.
- Krejcie, R.V. & Morgan, D.W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30, 607 – 610.
- Limperg, T. (1932). *Theory of Inspired Confidence*. University of Amsterdam.
- Ogungbade, O.I, Adekoya, A.C. & Olugbodi, D.I. (2021). Audit quality and financial reporting quality of deposit money banks listed on the Nigerian stock exchange, *Journal of Accounting, Finance and Auditing Studies* 7(1), 77-98.
- Okpala, K. E. (2015). Stamatis generic model and audit quality in Nigeria. *Journal of Social Sciences Research*, 1(8), 94–101.
- Olasanmi, O. O. (2013). Computer aided audit techniques and fraud. *Research Journal of Finance and Accounting*, 4 (5), 67-79.
- Rostami, K.H., Samadi, S., Omrani, H., Margavi, A.K., Asadzadeh, H. & Nazari, H. (2011). Data mining and application in accounting and auditing. *Journal of Education and Vocational Research*, 2(6), 211-215.
- Venkatadri, M . & Reddy, L.C. (2011). A review on data mining from past to the future. *International Journal of Computer Applications*, 15 (7), 1-4.
- Wang, J. & Yang, J.G.S. (2009). Data mining techniques for auditing attest function and fraud detection. *Journal of Forensic and Investigative Accounting*, 1 (1), 1-24.
- Zhang, Q. & Segall, R. (2008). Web mining: A survey of current research, techniques and software. *International Journal of Information, Technology and Decision Making*, 7(4), 683-720