RISET
JURNAL APLIKASI EKONOMI AKUNTANSI DAN BISNIS
https://ejournal.ibik.ac.id/index.php/riset E-ISSN : 2656-7113 P-ISSN : 2797-4057 DOI : doi.org/10.37641/riset.v4i2.166

FRAUD DETECTION AUTOMATION THROUGH DATA ANALYTICS AND ARTIFICIAL INTELLIGENCE

Wishmy Meinawa Ikhsan[1)], Elzami Haqie Ednoer[2)], Winanda Setyaning Kridantika[3)], Amrie Firmansyah[4)]

[1, 2,3,4)] Polytechnic of State Finance STAN

E-mail: *4132210045_wishmy@pknstan.ac.id* [1)], *elzamihaqie@gmail.com* [2)], *amrie@pknstan.ac.id* [4)]

ARTICLE INFO

FRAUD DETECTION AUTOMATION THROUGH DATA ANALYTICS AND ARTIFICIAL INTELLIGENCE

Submitted:
08 – July – 2022
Revised:
23 – August – 2022
Accepted:
27 – September – 2022

ABSTRACT

This study aims to review the use of data analytics and artificial intelligence in fraud detection to support internal audits. This study employs a qualitative method with a scoping review approach. The research data comprised 24 online journal articles indexed by Scopus and Sinta, which were used as the basis for scoping reviews. The stages carried out in this study consisted of identifying research questions, using keywords, selecting literature, mapping the results of research data, and compiling a summary of research results. This study concludes that the fraud detection model based on data analytics and artificial intelligence has a high accuracy value in improving audit quality. This study indicates that the Indonesian Financial and Development Supervisory Agency needs to increase the use of technology, including data analytics and artificial intelligence, to detect fraud optimally.

Keywords: Artificial Intelligence, Data Analytic, Fraud, Internal Audit

INTRODUCTION

The term industrial revolution is not something new anymore because the phenomenon has been introduced since the 18th century (Fonna, 2019). They started from the industrial revolution 1.0 with a focus on the revival of machines until now. Industrial revolution 4.0 is marked by the development of digital technology and the internet, increasing rapidly (Fonna, 2019). Areas that are trending in the industrial revolution 4.0 are the Internet of Things (IoT), Industrial Internet of Things (IioT), cyber-physical systems (CPS), cloud computing, and artificial intelligence (AI) (Fonna,

2019). These things focus on combining technology and the internet as an automation tool that supports the production process and activities carried out by humans.

The development of increasingly advanced technology certainly positively impacts various fields, such as industry, business, and finance. First, technology can automate business processes and encourage people to find innovations related to their business. Second, computerized accounting can provide predictions or projections regarding future business prospects for the benefit of stakeholders. Third, with the existence of a computerized system, it is no longer necessary to process the preparation of financial reports manually (Naqvi, 2020).

In addition to providing benefits, technological developments also provide new risks, such as viruses that attack the system, cybercrime, and illegal activity on computer networks and the internet. In addition, the risk that also needs attention is digital fraud, which can be interpreted as the use of web-based technology to commit fraud and misuse assets to obtain financial benefits. Fraud is generally defined as an activity that includes various behaviors in the form of fraud to make a profit (Button et al., 2009).

Fraud is one of the fraudulent acts that are of great concern to the international community. Fraud is a global problem that affects all organizations around the world. Fraud is often undetected and never reported, making it difficult to determine the full scope of global losses (ACFE, 2020). Indonesia ranks first in the country with the most fraud cases in Asia, with 36 fraud cases (ACFE, 2020). The number of cases decreased according to the ACFE study in 2022 to 23 cases (ACFE, 2022). Even though it is starting to improve in 2022, Indonesia is currently in the top 4 countries with the highest fraud cases in the Asia Pacific region, after Australia, China, and Malaysia (ACFE, 2022). This fraud case is detrimental to government organizations, the private sector, and the public, who have got the impact directly or indirectly. ACFE Indonesia has previously conducted a survey which proved that in 2019, losses due to fraud cases in Indonesia were around IDR 873 billion, of which acts of corruption caused almost 70% of these cases (ACFE Indonesia, 2020). It shows that fraud cases must be a concern to be mitigated by every organization.

The impact of fraud has affected performance in various fields and professions, one of which is auditing. Fraud causes the emergence of new areas that need the auditor's attention. In 2020, each person generated an average of 1.7 MB of data per second, increasing exponentially yearly. Data must be analyzed to ensure that fraud checking is done thoroughly. If the data are unchecked, IT developments can become a new opening for fraud because there are many hiding places for fraudsters. It is the background for the emergence of the need to use emerging technology to support auditor performance.

The extensive data population has resulted in the use of the sampling method in conventional audits is no longer effective. Without a thorough examination of each data, there is a risk that many critical areas go unnoticed. In addition, using conventional audit techniques also requires more time, resources, and costs to audit a more extensive scope. Thus, conventional audit techniques increase the risk of fraud detection failure and reduce audit effectiveness and efficiency. The audit sector is still slow in adopting emerging technologies to support audit activities (Satyawan et al., 2021). It is generally due to concerns over the high initial costs incurred to adopt it, even though this initial investment will be proportional to the long-term benefits obtained. In addition, the demand for the required competencies is a barrier for auditors to use new techniques.

Fraud techniques are developing rapidly, and auditors must be able to respond to them. It is not easy for auditors to keep up with the rapid development of information technology, especially with demands regarding specific technical capabilities in auditing these areas. However, it is still a must because adequate professional skills are the foundation of confidence in the opinion given by the auditor.

The demand for auditors to have high specific technical capabilities is reflected in Kovanen (2020), who finds there is an agreement in the views of auditors that in maintaining audit relevance, audit scope should not be determined only based on the limitations of existing capabilities. Hakami et al. (2020) found that the detection of fraud by auditors through conventional audit activities has a lower accuracy rate than the fraud detection model. It indicates the need for auditors to expand competence. Furthermore, Mui (2018) found data analysis ability as one of the skills needed by auditors in detecting fraud.

Fortress (2020) stated that big data technology, a new achievement in the world of information and technology, impacts various kinds of work, one of which is auditors. The use of big data analytics positively affects work performance and improves the audit quality carried out by auditors. Big data analytics will significantly add value to organizations by discovering hidden patterns in data (Handoko et al., 2020).

Big data is a good strategy for providing basic information, which is a significant advantage for its users. However, big data takes a long time, accuracy is not guaranteed, and costs are expensive. Given the availability of huge and complex amounts of data, big data processing will be optimal if it involves a computer system to analyze further the data they get to draw conclusions and take action. This system is called artificial intelligence (AI). AI has intelligence in thinking, has a broad knowledge base in a limited domain, and uses structured reasoning in making decisions to solve problems. It is considered the leading solution for various cases of auditor failure in detecting fraud.

AI can provide convenient automation and control and improve audit process efficiency. Muawanah et al. (2022) further observed how auditors perceive the ease with which AI can improve the audit process. It was concluded that the auditor considered the AI system to facilitate the audit process, although it is undeniable that there will also be direct threats to employees' work. In addition, in its implementation, auditors will undoubtedly face challenges in implementing AI to improve the audit process.

Specifically, internal audit has a significant role in fraud detection. The reporting and internal auditing media are the most effective means of detecting fraud (ACFE Indonesia, 2020). Reporting media originating from reports of employees of the company/organization is the tool with the most significant contribution to uncovering fraud. Furthermore, an internal audit is also very effective in the early detection of fraud. Internal audits can detect fraud better by utilizing data analytics and AI techniques. It is at the same time to maintain and increase the relevance of the internal auditor function.

This study aims to review the use of data analytics and AI in fraud detection to support internal audits. In addition, the risks and challenges that may be faced will also be discussed. This study differs from previous research because this study discusses data analytics support and artificial intelligence for internal audits. The previous studies mainly discussed data analytics and artificial intelligence separately (Albizri et al., 2019; Craja et al., 2020; West & Bhattacharya, 2016). They discussed audit coverage, whether internal audit, external audit, or both.

Internal audit is an independent and objective assurance and consulting activity designed to provide added value and improve the organization's operations (The Institute of Internal Auditors, 2022). Internal audit activities are carried out through a systematic and regular approach to evaluating and improving risk management, control, and governance effectiveness to achieve organizational goals (The Institute of Internal Auditors, 2022).

There is a shift in philosophy in viewing the role and function of internal audits in an organization in the current era. According to Zamzami et al. (2013), there has been a change in the orientation of the internal auditor profession from its initial role. As a watchdog around 1940 to a consultant role around 1970. With various developments, auditors play a role as strategic business partners for management and the board of directors (Zamzami et al., 2013).

Data analytics is the process of evaluating data to obtain conclusions to answer questions within a business (Richardson et al., 2020). Data analytics usually involves the technologies, systems, methodologies, databases, statistics, and applications used to analyze data (Richardson et al., 2020). The main point is that data analytics aims to turn raw data into knowledge to create value for its users.

Data analytics is often juxtaposed with big data, resulting in the term big data analytics. Big data is a complex data set with another term, namely 'the three Vs,' consisting of volume, velocity, and variety. Volume characteristics indicate the amount of big data that continues to grow and be produced every time. Velocity is the speed of growth of big data that cannot be controlled. So it is difficult to process it with conventional methods; variety means that all kinds of data are available in big data (Oracle.com, 2022). Big data analytics is the process of extracting useful information by analyzing various types of large data sets. Big data analytics is used to find hidden patterns, market trends, and consumer preferences to benefit a company's decision-making (Kurniawati, 2020).

AI is not new because its development started around the 1950s. This term was first coined by McCarthy et al. (1955) as science and engineering to make intelligent machines. The proposed definition of AI has developed along with the development of AI (McCarthy et al., 1955). However, artificial intelligence is developing a system capable of doing work that generally requires human thinking intelligence, such as visual perception, decision making, and language/communication understanding intelligence.

Now the development of AI is accelerating and indeed increasingly resembles human intelligence. According to Naqvi (2020), the factors that influence these developments include the exponential development of data that is so large and, of course, accompanied by storage capacity and data management. Furthermore, increased processing power increases data processing capabilities. Naqvi (2020) also stated that research to find new algorithms and approaches is increasingly widespread. It is also supported by the existence of a global network for researchers to share their research & accelerate the development of AI.

Data analytics is closely related to artificial intelligence. The tools used in data analytics are also used in artificial intelligence. In short, an artificial intelligence system can replace or perform data analysis work usually performed by data analysts to obtain information or implied knowledge of these data. Thus, artificial intelligence not only processes data as data analytics tools function but also analyzes the processing results. Naqvi (2020) recommended several tools, including robotic process automation (RPA),

expert systems, process mining, and machine learning to support audit automation. Data analytics and AI are kinds of leveraged technology in the continuous audit. They are real-time reporting and analysis in the form of an intelligent dashboard to show any anomalies, outliers, inconsistencies, and non-compliance earlier to provide feedback and feed-forward.

This research is expected to improve auditors' knowledge about various fraud detection methods that have the potential to be adopted according to organizational conditions. In addition to convincing auditors to immediately adopt new techniques, this study is also expected to increase the readiness of internal auditors to implement them. The potential benefits obtained are also expected to convince organizations to encourage and facilitate their internal auditors in implementing fraud detection systems based on data analytics and artificial intelligence. In addition, this research is expected to contribute to developing literature on data analytics and AI for internal audits.

METHOD

This study employs a qualitative method with a scoping review approach to discuss the implementation of data analytics and AI in fraud detection automation. Scoping review is a method that pays attention to the need for comprehensive literature, where each literature has a different level of depth of discussion (Arksey & Malley, 2005). This study took secondary data from 24 journal articles in online media indexed by Scopus and Sinta according to the theme of the discussion. The scoping review steps in this study are to identify research questions: "How is the implementation of data analytics and AI in fraud detection automation?". Next is to search the literature in the Scopus and Sinta indexed databases with the keywords: fraud detection with big data analytics and fraud detection with artificial intelligence. The next stage is to select the literature and map the data. The final stage of the scoping review of this research is compiling a summary of the results of data mapping. These stages used scoping review follows Pamungkas & Firmansyah (2021).

RESULTS

Implementing Data Analytics for Fraud Detection

Singh et al. (2019) started their research by determining the common fraud indicators in transactions. A total of 2,117 procurement transaction data at telecommunications companies in America became the object of testing in a model that utilizes binary logistic regression (BLR). However, BLR is strongly influenced by the cutoff value, where a high cutoff value tends to reduce the number of transactions marked as risky. In contrast, a low cutoff increases transactions marked as risky and false positives. False positives are usually caused by an algorithm of a program that states the existence of a symptom/signal/object that does not exist. The result is that the lowest cutoff value significantly produces the highest accuracy for marking fraud risk transactions of 68.6%. It should be noted that Singh et al. (2019) emphasized the importance of continuously updating the fraud indicator base in line with the development of fraud techniques.

Singh et al. (2019) also described the advantages offered by the built model. From the agency theory point of view, this model can reduce the costs required to create and execute contracts and monitor those contracts. It is a fundamental problem of agency theory where information asymmetry appears when shareholders entrust wealth management to the manager, so additional costs are needed to mitigate the risk. In addition to reducing incentives/encouragements to commit fraud, this model can also increase auditor efficiency and independence and improve overall audit quality. Moreover, this model can also be used to detect fraud in real-time, considering that the type of data used is also updated in real-time.

On a larger scale, Patil et al. (2018) performed big data analytics to detect fraud in real-time. This study uses logistic regression algorithms, decision trees, and random forest decision trees to process a large volume of banking data. The test results found that the fraud detection precision from logistic regression was 76%, the decision tree was 89%, and the random forest decision tree was 93%.

Rizki et al. (2017) examined how data processing applications detect fraud in companies in Indonesia. The companies studied were listed on the Indonesia Stock Exchange (IDX), which consisted of 100 non-fraud companies and 24 fraudulent companies. Significant indicators in detecting financial fraud are profitability and efficiency. Algorithm accuracy used Support Vector Machines (SVM) is 88.37%, while Artificial Neural Network (ANN) produces a better accuracy of 90.97%. However, there is no significant difference between the two algorithms, and the determination of the algorithm used is based on the data set, operating time, and ease of implementation.

Data analytics can also be used as a complement to other fraud detection techniques. Tarjo & Herawati (2015) found that the Beneish-M score can be combined with logistic regression data mining techniques to detect fraud. The results of this study of 70 companies in Indonesia show that the accuracy of fraud detection by this model is 77.1%. Tarjo & Herawati (2015) also found that depreciation, cost of sales, and discretionary accrual accounting policies can reduce the model's ability to detect fraud.

Another breakthrough in fraud detection is the use of text mining. Dong et al. (2018) examined financial and social media data to detect fraud. Social media is not directly related to organizational operations, but hidden insights can still be drawn from unstructured data on social media. The insights are extracted from signals such as sentiments, emotions, topics, lexical or related words, and social networks. The data obtained is, of course, very sufficient in number to be analyzed, considering that the use of social media is increasing rapidly. In addition, processing can be carried out in real-time, considering that the data can be recorded directly. The data is then processed with four algorithms (SVM, NN, decision tree/DT, LR) to determine which produces the best accuracy. The fraud detection accuracy rate among 128 companies ranged from 54.5% to 75.5%, with the highest accuracy held by SVM. In addition, the analysis of the language used in the management discussion and analysis section of the financial statements produces detection accuracy between 52.78% to 70.33%, with the best results given by the LR algorithm.

In comparison, fraud detection with financial ratios tested in this study has an accuracy ranging from 41.17% to 56.17%. However, the best results are produced from a combination using SVM, reaching an accuracy of 56.17% to 80%. This research also shows how this method can help auditors better assess the risk of material misstatements.

Goel et al. (2010) examined whether linguistic predictors detect fraud in finance. This study uses datasets from a sample of 126 companies that committed fraud from 1993 to 2006. The method used for testing is Natural Language Processing (NLP) tools to analyze, understand, and produce language. The Support Vector Machines (SVM) method is a machine learning method to study the characteristics of the tested dataset. This method can increase the prediction accuracy of the fraud detection model from the initial baseline with an accuracy of 56.75% to 89.51%. Interestingly, there have been systematic differences in communication and writing style of the annual report that contained fraud where financial figures did not accurately capture the narrative.

Goel & Uzuner (2016) further examined text analysis by examining whether sentiment affected fraud detection in the annual report. The methods used are Natural Language Processing (NLP) and Support Vector Machines (SVM) which are considered the most suitable for building the fraud sentiment classification model. This method identifies sentiment features to distinguish between fraudulent and honest Management Discussion and Analysis (MD&As) by measuring sentiment on polarity, subjectivity, and intensity dimensions. The classification model built can identify 81.84% accurately. This model can identify a more prominent use of negative and positive sentiment, a higher proportion of subjective expression than objective expression, and greater use of sentiment expression.

Data analytics can also detect fraud in other fields, not just financial statement fraud, one of which is international shipping. The risk of fraud often faced in this field is the manipulation of documents to avoid customs fees and smuggling. Usually, random audits are conducted to check the completeness of the documents. However, this technique cannot cover the entire sender, so the risk of passing the manipulated document is still large. Triepels et al. (2018) investigated fraud detection used the Bayesian Network to analyze trade and shipping route patterns. As a result, random audits with conventional methods showed a very low fraud detection precision of 9.99%, while the data-driven model reached a value between 34.6% and 60.7%. This detection model also has the potential to be applied to companies that have a similar business nature as retail companies that manage lots of inventory movement. With this model, auditors can reduce physical audits while increasing accuracy by involving a larger population.

Not just a model, data analytics for fraud handling has been implemented directly in the real world. Unmitigated, the giant company Alibaba uses big data as a basis for fraud risk management. Chen et al. (2015) reviewed how Alibaba uses CTU as a real-time fraud prevention monitoring system and a risk model in the form of a Readiness and Investment Navigator (RAIN) score. Alibaba's fraud risk method leads to a new generation of fraud risk monitoring and management based on big data processing and computing technology, as well as a real-time fraud prevention system.

The findings of testing the accuracy of data analytics models that show their benefits in detecting fraud seem to align with the auditor's perception. Syahputra & Afnan, (2020) surveyed 221 auditors working at Indonesia Audit Supreme Board and Financial and Development Supervisory Agency spread throughout Indonesia. This study examined the effect of big data and forensic audits on fraud detection. The test results suggest that big data and forensic audits are positively associated with fraud detection.

Application of Artificial Intelligence in Fraud Detection

Tang et al. (2018) developed an intelligent fraud detection model based on a fraud detection ontology. Ontology is a discussion of the most basic or deepest principles (Tang et al., 2018). In this study, the ontology was compiled based on financial statement data to produce financial variables that significantly affect detecting fraud (Tang et al., 2018) by combining the ontology knowledge base with the C4.5. An algorithm is used to serve the decision tree. This model was tested on 130 fraudulent companies and 130 non-fraudulent companies. As a result, the detection rate of fraudulent financial statements reached 86.67%, with a total accuracy of 80%, indicating this model's validity. However, it should emphasize that this model relies purely on financial reports' data without expert knowledge. Even though the auditor's field knowledge has its value which is not always reflected directly in the financial statements.

Financial variables from financial statements are also used in the research (Temponeras et al., 2019). If Tang et al. (2018) obtained the main variables for detecting fraud from the compilation of ontologies, Temponeras et al. (2019) determined the variables based on professional judgment. The machine learning algorithm used to detect fraud in this study is Denoising Diffusion Probabilistic Models (DDPM). It is then compared with Deep Dense Multilayer Perceptron Decision tree C4.5; k-NN type 3-NN; logistic regression; and minimal sequential optimization for support vector machine (SVM). In testing 164 companies in Greece, DDPM obtained the best results among other variables, with a fraud detection rate of 91.7% and a total accuracy of 93.7%. With this model, sound professional judgment from the auditor is needed to determine the proper characteristics of fraud as the basis for detection. Bao et al. (2020) used raw data instead of financial ratios because raw data can be used more flexibly and in more complex functional forms. With a data sample of 206,026 American companies, using the Normalized Discounted Cumulative Gain (NDCG)@k model can relatively avoid high direct and indirect costs for investigating cases suspected of fraud.

In addition to financial variables, non-financial variables were used together in the study of Jan (2018) on 40 fraudulent companies and 120 non-fraudulent companies in Taiwan. In the first stage, these essential variables are filtered by artificial neural networks (ANN) and SVM. Ten-fold cross-validation was carried out in the second stage to improve detection accuracy in 4 decision tree algorithms. From the eight models built, the combination of ANN and CART decision tree resulted in the highest fraud detection rate of 90.21%, with a total accuracy of 90.83%. This study also emphasized the need for internal control, audit systems, and corporate governance to avoid fraud and produce objective data for more accurate analysis. Internal auditors certainly play a significant role in ensuring the system's realization.

Mustika et al. (2021) examined fraud detection in retail financing companies in Indonesia. Departing from the pandemic conditions that limited audit space, remote audit activities became a solution that was intensively encouraged. However, limited direct interaction also increases the risk of document manipulation and deletion of relevant information. This study employs 46,536 transactions as objects and employs customer and finance profiles as predictors of fraud. Based on the test, the random forest algorithm became the method with the highest accuracy of 74.54%. When compared, the previously discussed studies have a total accuracy rate of more than 80% and fraud detection accuracy above 85%. Mustika et al. (2021) suggested further research to add appropriate predictor variables to improve accuracy. Thus, the

implication of conducting an audit is that auditors need to understand the characteristics of fraud in depth. Professional judgment in determining these variables will affect the accuracy of fraud detection.

AI is certainly not limited to automating the audit process itself but can be used in all operational activities of the organization. However, not a direct part of the internal audit process described above, AI in operational activities can also contribute indirectly to the audit process.

Several studies have used artificial intelligence to detect fraud used by management. Raj & Portia (2011) conducted a survey and evaluated AI-based fraud detection techniques used the Bayesian Belief Networks (BNN) and Artificial neural network (ANN) models. The results of this study are that the ANN and BNN models have a medium accuracy and are pretty expensive. Fraud detection rates for BNN and ANN are 74% and 77%, respectively.

Sadineni (2020) also conducted a similar study by utilizing the Kaggle data repository, which has 150,000 stored transactions. The research resulted in machine learning accuracy for detecting fraud, which was above 95%. While Alghofaili et al, (2020) developed methods for increasing computational speed, handling big data, and identifying unknown attack patterns. The deep learning technique used in this research is Long Short-Term Memory (LSTM) which can achieve 99.95% accuracy in less than one minute. At the same time, Thennakoon et al. (2019) evaluated a series of machine learning models to choose the optimal algorithm for the type of fraud. The data used in this study are files with 200 fraud case records, and transaction log files have 917,781 records with four different fraudulent transaction patterns. The results suggest a system for detecting fraud in real-time consisting of three central units: API module, fraud detection models, and data warehouse.

Exploring different variables, P. Singh & Singh (2015) built a fraud detection model by monitoring behavior and geographic location used the Bayesian Learning Approach. This model can assess the possibility of fraudulent activity by providing a risk score where a score greater than 0.5 will be considered fraud. Not only detection, but this model also applies advanced mechanisms to prevent fraud. When a possible fraud is detected, the system will authenticate the user by asking the user to input the correct verification code.

Vynokurova et al. (2020) examined Hybrid Machine Learning systems to overcome anomaly detection in user geolocation (GPS spoofing, Wi-Fi spoofing, location jumping). This hybrid system consists of two subsystems: an anomaly detection subsystem and an anomaly type interpretation subsystem, which uses a decision tree technique with supervised learning and unsupervised learning. This hybrid system produces a fraud detection accuracy rate of above 90%. The advantage of this hybrid system is higher data processing speed when given real-time data.

In the insurance sector, Dhieb et al. (2020) developed a Smart Insurance System based on Blockchain and AI (SISBAR), a fraud detection system used extreme gradient boosting (XGBoost), and Very Fast Decision Tree (VFDT) as dynamic fraud detection and classification model for insurance companies. Based on the test, it was concluded that the XGBoost method has the highest accuracy and prediction. Roy & George (2017) used more than 500 insurance claim data to detect fraud in insurance claims using machine learning. Machine learning techniques are used in the form of the decision tree, random forest, and Naive Bayes. Decision tree and random forest algorithm classification techniques perform better than naive Bayes.

Although all systems mentioned above are not used directly by the auditors, the information obtained from these systems can be used to evaluate performance and internal control carried out by management, for example, in detecting insurance claim fraud. If the system finds a high level of fraud, the auditor needs to audit the cause further, whether it comes from inadequate control, poor team member performance, or pure customer error. If the cause is found in the internal organization, the auditor must provide recommendations on the problem. In addition, these models can be used as a blueprint for developing new fraud detection models that can be used directly in audits. This development mainly explores new variables that have not been used in audit automation, such as behavior patterns, location, and geolocation.

Audit activities certainly do not stop when fraud is discovered. A series of audit work programs still need to be carried out to produce an audit report containing the auditor's opinion. In this case, Soeprajitno (2019) analyzed the potential of Artificial Intelligence in issuing auditor opinions. Machine learning Artificial Neural Network (ANN) tools were used in this study. Combining capabilities on big data, data mining, artificial neural networks, and final analysis through calculating total irregularities obtained using Fuzzy will help AI have the ability to detect fraud and issue audit opinions. Professional judgment, usually done manually, can be obtained from the ability of AI as the basis for issuing explanatory paragraphs.

Analysis of the Application of Data Analytics and Artificial Intelligence for Fraud Detection

The results of testing the fraud detection model with the highest accuracy in each article can be seen in the following table:

Table 1. The Summary Of Fraud Detection Model Test Results

Testing Accuracy Rate	Number of Models	Algorithm Used
90-99%	10	ANN[38], XGBoost+VFDT[15], ANN[32], ANN[35], SVM[45], LSTM[6], HMLS[47], RF[30], DDMP[2], ANN+CART[14]
80-89%	3	NLP+SVM[20], NLP+SVM[21], Ontology+C4.5[10]
70-79%	4	BNN+ANN[31], RF[5] SVM[18], Beneish M-Score+LR[27]
60-69%	2	BLR[26], BSN-NN[22]

Source: Data processed

Based on the accuracy table of the fraud detection model above, ten research results (52.63%) show that the accuracy of tools is above 90%, three research results (15.79%) show tool accuracy is in the 80%-89% range, four research results (21.05%)) shows tools are in the range of 70%-79%. Two research results (10.53%) show tools are in the range of 60%-69%. Data analytics and artificial intelligence accuracy in detecting fraud are very high. More than half of the samples used showed an accuracy rate above 90%, and the rest were within an average accuracy of 85.55%. It shows that data analytics and artificial intelligence can be very accurate in detecting fraud and successful in their application.

In addition, several exciting studies review the success of data analytics and artificial intelligence methods in identifying fraud. Syahputra & Afnan (2020) found that big data is positively associated with detecting fraud and conducting forensic

audits. Roy & George (2017) found that the decision tree and random forest algorithms have the best accuracy, precision, and recall performance. Furthermore, P. Singh & Singh (2015) used the Bayesian Learning Approach to detect fraud and perform authentication for security if anomalies are found in transactions. The fraud risk management model created by the Alibaba company uses big data, which was later examined by Chen et al. (2015) and demonstrated its ability to detect fraud in real-time. In contrast, Bao et al. (2020) used raw data processed with the NDCG@k algorithm considering flexibility.

All these studies offer an overview of how fraud detection can use various algorithms with various types of variables and input data. This condition shows flexibility in the design of fraud detection systems. Auditors can adjust the types of variables and input data according to organizational conditions. The data with adequate availability can be selected for analysis from the various types of data that can be used. In addition, the algorithm used can adjust the ability of the system builder because there is no rigid limit to only using specific algorithms.

The variable selection also plays an essential role in obtaining the best level of accuracy. Auditors can test different variable combination types to find the most ideal for the organization. In addition, since fraud techniques continue to develop in line with technology, auditors must be keen to see gaps that can become indicators of new fraud. Therefore, the variables used in the system also need to be updated regularly. It is where the auditor's professional judgment is needed to assess which areas need more attention as potential new variables.

The success and usefulness of implementing data analytics and artificial intelligence are beneficial as a reference and guide for internal auditors to be able to carry out supervisory duties. The rapid development of information has broadened the scope of organizational audits and is more diverse. In this case, the auditor cannot continue to rely on manual control techniques. Along with the development of information technology, fraud patterns continue to develop, so more advanced, automated, fast, accurate monitoring techniques are needed and can monitor an enormous scope in identifying fraud.

Based on the studies that have been carried out, it is known that the application of data analytics and artificial intelligence will benefit the world of internal audits. Implementing big data analytics and artificial intelligence can reduce the potential for audit delays due to its ability to detect fraud in real-time. Audit delay is the period for completing the audit from the date of the company's reporting until the audit report is issued (Widati & Septy, 2008).

Data analytics and artificial intelligence can detect fraud in real-time and are effectively used by organizations. Using these two tools can also reduce: manual fraud detection activities, increase the scope of detection, speed up the process, and have a better future prediction system. In addition, artificial intelligence can also assist auditors in reducing physical work, where previously auditors had to examine various documents with many pages. Artificial intelligence plays a role in automatically recognizing and processing all documents so that manual processes can be reduced and auditors have more time to analyze. This intelligent audit automation can also help improve audit quality because it has a high level of accuracy for detecting fraud and reduces costs for conducting audits.

Without reducing the benefits of artificial intelligence, it should still be noted that based on the analysis, not all artificial intelligence methods can detect fraud in real time

or are fully automated. Artificial intelligence has been able to conclude/information like humans, but in most models, the process is still semi-automated. The type of input data used can not all be captured directly by the system, or it still needs to be inputted first by the officer. Integrating the fraud detection system with the operating system that produces data as material for fraud analysis is necessary to automate the process. This integration allows the fraud detection system to analyze each new data entry during an activity/transaction. The development of this real-time system needs to be carried out continuously because the faster the early warning system works, the intervention can be carried out immediately so that losses can be minimized.

Previously, it was also mentioned how initial investment was an obstacle to adopting data analytics and artificial intelligence. Although it is very clear how many benefits will be obtained, there are not many studies that directly mention the number of costs required. This information is helpful for organizations to conduct a cost-benefit analysis. The existence of more explicit cost information can further convince organizations to facilitate the use of data analytics and artificial intelligence.

The many benefits derived from data analytics and artificial intelligence do not mean that these two tools are easy to implement. In addition to costs, auditors certainly need to consider the risks that must be mitigated so as not to harm the achievement of organizational goals. In addition, auditors face challenges in implementing data analytics and artificial intelligence to detect fraud optimally.

Kovanen (2020) identified several risks that accompany audit automation. First, there are technological risks in data errors, algorithm and system malfunctions such as producing incorrect output or system crashes. Second, the regulation on the use of data needs to be a concern, especially in the ethical use of data and the demand to maintain data security that concerns many parties. Another risk that needs to be considered is cyber attacks. Much data continuously processed by the system is sensitive data, such as financial and employee information. This condition further increases the vulnerability to cyber-attacks. This attack can be in the form of unauthorized access, extraction, deletion, and destruction of data, as well as other activities that result in disruption, operations, and losses, both financial and non-financial. In addition, detection accuracy also depends on the adequacy and authenticity of the processed data. Thus, internal auditors must ensure that the organization has adequate internal controls to produce quality data.

Apart from risks, there are challenges faced in using intelligent automation. The biggest challenge is the demand for auditor competence to keep pace with technological developments. The most significant demands are technical capability, regulatory understanding, and accountability for audit confidence and accuracy (Kovanen, 2020).

The transparency of this intelligent system is also a concern of its own. Audit automation makes it easier for auditors to audit activities that utilize artificial intelligence and data analytics and the organization's operational activities. The algorithms used may also be different from those used for audit automation. Auditors must also understand how this system works, requiring management's openness in conveying it.

High expectations from regulators and organizations as auditees are also challenging (Soeprajitno, 2019). Stakeholders have always expected the assurance level to reach 100%, although it is difficult to achieve. Previously, conducting conventional audits by sampling could cause missing critical areas to be examined. Nevertheless, now with the help of intelligent audit automation with high accuracy, the expected

standard for audit confidence levels is getting higher. There is no longer any excuse for auditors not to meet stakeholder expectations, so auditors must rush to master these new techniques for better audit quality.

CONCLUSION

This study concludes that various fraud detection models based on data analytics and artificial intelligence have a high accuracy value in improving audit quality. These models with various algorithms, variables, and input data can be a blueprint for developing a fraud detection system following the organization's characteristics.

This research is limited to the scoping review method on the literature that has reviewed and tested fraud detection based on data analytics and artificial intelligence. In this regard, there are still many things that need to be explored regarding techniques in data analytics and artificial intelligence, which also continue to develop in line with the development of information technology. Not all data analytics and artificial intelligence techniques are suitable for every characteristic of fraud. Further research is expected to be able to examine more deeply related to the techniques of applying data analytics and artificial intelligence to detect each type and characteristic of fraud.

This research indicates that the Indonesia Financial and Development Supervisory Agency, as the supervisory agency for the Government's Internal Supervisory Apparatus, is expected to encourage the urgency of using technology, including data analytics and artificial intelligence, to detect fraud optimally. In addition, the Indonesia Agency for the Assessment and Application of Technology needs to improve the competence of internal auditors and technology experts to cooperate in developing artificial intelligence for fraud detection.

REFERENCES

ACFE. (2020). *Report to The Nations on Occupational Fraud and Abuse: 2020 Global Fraud Study*. https://acfepublic.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf

ACFE. (2022). *Occupational Fraud 2022: A Report to The Nations*. https://acfepublic.s3.us-west-2.amazonaws.com/2022+Report+to+the+Nations.pdf

ACFE Indonesia. (2020). *Indonesian Fraud Survey 2019*. https://acfe-indonesia.or.id/wp-content/uploads/2021/02/SURVEI-FRAUD-INDONESIA-2019.pdf

Albizri, A., Appelbaum, D., & Rizzotto, N. (2019). Evaluation of Financial Statements Fraud Detection Research: A Multi-Disciplinary Analysis. *International Journal of Disclosure and Governance*, *16*(4), 206–241. https://doi.org/10.1057/s41310-019-00067-9

Alghofaili, Y., Albattah, A., & Rassam, M. A. (2020). A Financial Fraud Detection Model Based on LSTM Deep Learning Technique. *Journal of Applied Security Research*, *15*(4), 498–516. https://doi.org/10.1080/19361610.2020.1815491

Arksey, H., & Malley, L. O. (2005). Scoping Studies: Towards A Methodological Framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Bao, Y., Ke, Bi., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach. *Journal of Accounting Research*, *58*(1), 199–235. https://doi.org/10.1111/1475-679X.12292

Button, M., Lewis, C., & Tapley, J. (2009). *Fraud typologies and victims of fraud*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/118469/fraud-typologies.pdf

Chen, J., Tao, Y., Wang, H., & Chen, T. (2015). Big Data Based Fraud Risk Management at Alibaba. *Journal of Finance and Data Science*, *1*(1), 1–10. https://doi.org/10.1016/j.jfds.2015.03.001

Craja, P., Kim, A., & Lessmann, S. (2020). Deep Learning For Detecting Financial Statement Fraud. *Decision Support Systems*, *139*, 113421. https://doi.org/10.1016/j.dss.2020.113421

Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A Secure AI-Driven Architecture for Automated Insurance Systems: Fraud Detection and Risk Measurement. *IEEE Access*, *8*, 58546–58558. https://doi.org/10.1109/ACCESS.2020.2983300

Dong, W., Liao, S., & Zhang, Z. (2018). Leveraging Financial Social Media Data for Corporate Fraud Detection. *Journal of Management Information Systems*, *35*(2), 461–487. https://doi.org/10.1080/07421222.2018.1451954

Fonna, N. (2019). *Development of the Industrial Revolution 4.0 in Various Fields*. Bogor: Guepedia.

Fortress, H. (2020). *The Effect of Using Big Data Analytics on The Work of Auditors* [Universitas Katolik Parahyangan]. https://repository.unpar.ac.id/handle/123456789/11669

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

Goel, S., & Uzuner, O. (2016). Do Sentiments Matter in Fraud Detection? Estimating Semantic Orientation of Annual Reports. *Intelligent Systems in Accounting, Finance and Management*, *23*(3), 215–239. https://doi.org/10.1002/isaf.1392

Hakami, T. A., Rahmat, M. M., Yaacob, M. H., & Saleh, N. M. (2020). Fraud Detection Gap between Auditor and Fraud Detection Models: Evidence from Gulf Cooperation Council. *Asian Journal of Accounting and Governance*, *13*, 1–13. https://doi.org/10.17576/ajag-2020-13-01

Handoko, B. L., Mulyawan, A. N., Tanuwijaya, J., & Tanciady, F. (2020). Big Data in Auditing for Tthe Future of Data Driven Fraud Detection. *International Journal of Innovative Technology and Exploring Engineering*, *9*(3), 2902–2907. https://doi.org/10.35940/ijitee.b7568.019320

Jan, C. (2018). An Effective Financial Statements Fraud Detection Model for the Sustainable Development of Financial Markets: Evidence from Taiwan. *Sustainability*, *10*(2), 513. https://doi.org/10.3390/su10020513

Kovanen, A. (2020). *Risks of Intelligent Automation and Their Impact on Internal Audit* [Tampere University]. https://trepo.tuni.fi/handle/10024/121440

Kurniawati, G. N. (2020). *Big Data Analytics and Its Uses for Your Business Development*. DO Lab. https://www.dqlab.id/big-data-analytics-dan-kegunaannya-untuk-perkembangan-bisnis

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A Proposal for The Dartmouth Summer Research Project on Artificial Intelligence*. 1–13. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Muawanah, A., Adawiyah, D., Maisarah, I., Ali, M. R. A., & Widiastuti, N. P. E. (2022). Auditor Behavior Responding to the Emergence of Artificial Intelligence in the Audit Process. *Jurnal Publikasi Ekonomi Dan Akuntansi …*, *2*(1), 52–60. http://ejurnal.stie-trianandra.ac.id/index.php/jupea/article/view/152

Mui, G. Y. (2018). Defining Auditor Expertise in Fraud Detection. *Journal of Forensic and Investigative Accounting*, *10*(2), 168–186. https://s3.amazonaws.com/web.nacva.com/JFIA/Issues/JFIA-2018-No2-2.pdf

Mustika, N. I., Nenda, B., & Ramadhan, D. (2021). Machine Learning Algorithms in Fraud Detection: Case Study on Retail Consumer Financing Company. *Asia Pacific Fraud Journal*, *6*(2), 213–221. https://doi.org/10.21532/apfjournal.v6i2.216

Naqvi, A. (2020). *Artificial Intelligence for Audit, Forensic Accounting, and Valuation*. New Jersey: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119601906

Oracle.com. (2022). *What is Big Data?* Oracle Cloud Infrastructure (OCI). https://www.oracle.com/big-data/what-is-big-data/

Pamungkas, U. D., & Firmansyah, A. (2021). How is The Regulation of Cryptocurrency Ownership by Companies Based on Financial Accounting Standards? *Jurnal Ilmiah Akuntansi Kesatuan*, *9*(3), 489–510. https://doi.org/10.37641/jiakes.v9i3.895

Patil, S., Nemade, V., & Soni, P. K. (2018). Predictive Modelling for Credit Card Fraud Detection Using Data Analytics. *Procedia Computer Science*, *132*, 385–395. https://doi.org/10.1016/j.procs.2018.05.199

Raj, S. B. E., & Portia, A. A. (2011). Analysis on Credit Card Fraud Detection Methods. *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, 152–156. https://doi.org/10.1109/ICCCET.2011.5762457

Richardson, V., Terrell, K., & Teeter, R. (2020). *Data Analytics for Accounting* (2nd ed.). New York: McGraw-Hill. https://www.amazon.com/Data-Analytics-Accounting-Vernon-Richardson/dp/1260837831

Rizki, A. A., Surjandari, I., & Wayasti, R. A. (2017). Data Mining Application to Detect Financial Fraud in Indonesia's Public Companies. *2017 3rd International Conference on Science in Information Technology (ICSITech)*, 206–211. https://doi.org/10.1109/ICSITech.2017.8257111

Roy, R., & George, K. T. (2017). Detecting Insurance Claims Fraud using Machine Learning Techniques. *2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, 1–6. https://doi.org/10.1109/ICCPCT.2017.8074258

Sadineni, P. K. (2020). Detection of Fraudulent Transactions in Credit Card Using Machine Learning Algorithms. *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 659–660. https://doi.org/10.1109/I-SMAC49090.2020.9243545

Satyawan, M. D., Triani, N. N. A., Yanthi, M. D., Siregar, C. S., & Kusumaningsih, A. (2021). Accelerating the Role of Technology in Audit During Covid-19 Pandemic. *Jurnal Akuntansi Multiparadigma*, *12*(1), 186–206. https://doi.org/10.21776/ub.jamal.2021.12.1.11

Singh, N., Lai, K., Vejvar, M., & Cheng, T. C. E. (2019). Data-Driven Auditing: A Predictive Modeling Approach to Fraud Detection and Classification. *Journal of Corporate Accounting & Finance*, *30*(3), 64–82. https://doi.org/10.1002/jcaf.22389

Singh, P., & Singh, M. (2015). Fraud Detection by Monitoring Customer Behavior and Activities. *International Journal of Computer Applications*, *111*(11), 23–32. https://doi.org/10.5120/19584-1340

Soeprajitno, R. R. W. N. (2019). Artificial Intelligence (AI) Potencial Issue an Auditor's Opinion? *Jurnal Riset Akuntansi Dan Bisnis Airlangga*, *4*(1). https://doi.org/10.31093/jraba.v4i1.142

Syahputra, B. E., & Afnan, A. (2020). Fraud Detection: The Role of Big Data and Forensic Audit. *Jurnal ASET (Akuntansi Riset)*, *12*(2), 301–316. https://doi.org/doi.org/10.17509/jaset.v12i2.28939

Tang, X.-B., Liu, G.-C., Yang, J., & Wei, W. (2018). Knowledge-based Financial Statement Fraud Detection System: Based on an Ontology and a Decision Tree. *Knowledge Organization*, *45*(3), 205–219. https://doi.org/10.5771/0943-7444-2018-3-205

Tarjo, & Herawati, N. (2015). Application of Beneish M-Score Models and Data Mining to Detect Financial Fraud. *Procedia - Social and Behavioral Sciences*, *211*, 924–930. https://doi.org/10.1016/j.sbspro.2015.11.122

Temponeras, G. S., Alexandropoulos, S.-A. N., Kotsiantis, S. B., & Vrahatis, M. N. (2019). Financial Fraudulent Statements Detection through a Deep Dense Artificial Neural Network. *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1–5. https://doi.org/10.1109/IISA.2019.8900741

The Institute of Internal Auditors. (2022). *About Internal Audit*. The Institute of Internal Auditors. https://www.theiia.org/en/about-us/about-internal-audit/

Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., & Kuruwitaarachchi, N. (2019). Real-Time Credit Card Fraud Detection Using Machine Learning. *Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019*, 488–493. https://doi.org/10.1109/CONFLUENCE.2019.8776942

Triepels, R., Daniels, H., & Feelders, A. (2018). Data-Driven Fraud Detection in International Shipping. *Expert Systems with Applications*, *99*, 193–202. https://doi.org/10.1016/j.eswa.2018.01.007

Vynokurova, O., Peleshko, D., Bondarenko, O., Ilyasov, V., Serzhantov, V., & Peleshko, M. (2020). Hybrid Machine Learning System for Solving Fraud

Detection Tasks. *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, 1–5. https://doi.org/10.1109/DSMP47368.2020.9204244

West, J., & Bhattacharya, M. (2016). Intelligent Financial Fraud Detection: A Comprehensive Review. *Computers & Security*, *57*, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

Widati, L. W., & Septy, F. (2008). Factors Affecting the Time Span of Financial Statements to the Public (Empirical Study on LQ 45 Companies Listed on the Indonesia Stock Exchange). *Fokus Ekonomi (FE)*, *7*(3), 173–187.

Zamzami, F., Faiz, I. A., & Mukhlis. (2013). *Internal Audit: Concepts and Practices*. Gadjah Mada University Press. https://ugmpress.ugm.ac.id/userfiles/product/daftar_isi/Audit_Internal.pdf