Explainable Artificial Intelligence (XAI) in Management and Entrepreneurship: Exploring the Applications and Implications

ADEBAYO ROTIMI PHILIP
Department of Artificial Intelligence, Africa Centre of Excellence on Technology Enhanced Learning (ACETEL), Nigeria Open University (NOUN), Nigeria.

Abstract- Explainable AI (XAI) is a crucial aspect of AI model development. Success in XAI will skyrocket the use of AI in management and entrepreneurship, especially in crucial decision-making and resource allocations. This study evaluates how XAI enhances trust in decision-making, assesses the key applications of XAI in management and entrepreneurship, and considers the challenges and ethical considerations that arise in the adoption of XAI in management and entrepreneurship. Existing literature was explored and concrete conclusions were drawn from that literature. Findings show that AI solutions have proven to be effective in business operations etc., however, AI solutions present a huge challenge. High-accurate AI solutions are not explainable. This study shows that managers and entrepreneurs are likelier to trust AI solutions if they are interpretable. Besides, the literature reveals that guided XAI improves confidence in decision-making, particularly in strategic planning and resource allocations. XAI is found to reduce cognitive bias, promote transparency and accountability, and also address ethical concerns.

Indexed Terms- Explainable AI, entrepreneurship, management, SHAP, LIME, Post hoc AI, black box, white box, interpretability, explainability, transparency

I. INTRODUCTION

Due to the increase in economies of scale and inadequate financial resources, SMEs are gradually adopting the use of AI for their service operations. Government and international organizations have seen the need to promote the adoption of AI technologies by SMEs and business enterprises (Emil and Simon, 2021). For instance, the European Commission's 2021 proposal emphasizes the need to remove barriers for the adoption of AI technology by SMEs and the need for national governments to develop AI initiatives targeted at SMEs and users of AI systems (European Commission, 2021). In Europe, countries have established Digital Innovation Hubs, allowing businesses to access AI technologies, with the EU and member states investing €1.5 billion. Besides, Australian also released AI action in 2021 which incorporates the establishment of a National AI center that specifically address the barriers facing AI adoption (European Commission, 2021).

Despite the widespread awareness of the adoption of AI in management and entrepreneurship, a huge concern has emerged— the inherent opacity in their choice-making approaches (Shweta and Poonam, 2021). The AI solutions have been referred to as 'black box' which depicts that it's quite challenging to determine how AI models produce their outputs (Alejandro et al, 2020). In one of the most cited surveys of AI ethical principles, Jobin et al (2019) showed that transparency or interpretability was one of the most prevalent issues. Also, the European Commission's High-Level Expert Group Ethics Guidelines for Trustworthy AI incorporate the principle of explainability as one of the four AI ethical fundamental principles (European Commission, 2019).

Explainable AI (XAI) refers to a type of AI system designed to provide transparent, interpretable, and understandable decisions, ensuring trust, accountability, and usability in fields like management, healthcare, finance, and entrepreneurship (Shweta and Poonam, 2021). The quest to make AI systems interpretable has grown over a few years and this is because of the inherent duty to validate and evaluate AI-driven solutions and broaden its applications in management, entrepreneurship, healthcare, and even in sciences. The journey through this review encompasses an exploration of various methodologies, techniques, and frameworks that contribute to the development of more interpretable AI structures (Shweta and Poonam, 2021).

The research aims to investigate the significance of XAI in management and entrepreneurship, particularly in decision-making. The study considers the applications, and the challenges and ethical concerns associated with adopting XAI in management and entrepreneurship. The study considers three research questions which are highlighted below:

1. How does Explainable AI (XAI) enhance decision-making transparency and trust in management and entrepreneurship?
2. What are the key applications of XAI in managerial decision-making and entrepreneurial ventures, and how do they impact business performance?
3. What challenges and ethical considerations arise in the adoption of XAI within management and entrepreneurship, and how can they be addressed?

II. LITERATURE REVIEW

2.1. Introduction

Explainable AI (XAI) has gained attention among AI professionals and enthusiasts, including management and entrepreneurship. This is because AI models such as deep learning, which has produced more accurate results are not interpretable, hindering its total deployment in management processes and entrepreneurship. The main challenge in XAI is how do we make AI models interpretable without affecting their accuracy? As can be seen, there exists a gap between AI models particularly deep learning models and human understanding, which is currently referred to as the "black box" (Eschenbach, 2021). This literature review starts by examining the definition of XAI, discusses its concepts, and XAI techniques, after which we delve into XAI in management and entrepreneurship and lastly considers the theoretical framework such as Decision Theory (Simon, 1955), Technology Acceptance Model (TAM) (Davis, 1989), etc., for which this study is based.

2.2. Brief Overview of XAI: Definition, concept and techniques

The term explainable AI (XAI) is not a new buzzword, though could be viewed as a new name for a very old quest in science (Agarwal, C., Nguyen, 2020). XAI was coined by DARPA (Gunning and Aha, 2019) has gained popularity in the field of AI. As its name suggests, XAI is the ability of an AI model to provide a clear explanation for its actions and decisions (Miller T. H., 2017). The major goal of XAI is to enable humans to understand the behavior of the model and its underlying mechanisms in decision-making. Several efforts which explain the workings of AI models are primarily tailored to the researcher, rather than improving explanations to end-users. An ideal definition that captures explainability emphasizes the ability to explain the AI model's past actions, ongoing processes, and upcoming steps on which the processes are based (Gunning, 2019).

The definition of XAI isn't complete without considering interpretability, transparency, and explainability. These three terminologies, though are often used interchangeably, are quite different and represent different aspects in the definition of XAI. According to Agarwal and Nguyen (2020), transparency refers to the ability of humans to understand the operations of the entire AI model. Transparency is also considered at a level of individual components and a level of a particular training algorithm. Another way to see transparency is that the input, computation, and output admit an intuitive explanation (Lepri, 2018). An AI model is considered transparent if stakeholders can assess its decision-making process, identify biases and unfairness, and ensure it complies with legal and ethical standards (TechDispach, 2022).

Interpretability, on the other hand, is concerned with the ability to explain the internal workings of an AI model. Interpretability refers to the degree humans comprehend a given AI decision-making process (Lisboa, 2013). According to Burrell (2016), an AI model is said to be poorly interpretable if the users or the AI researchers do not have a concrete understanding of why a particular prediction or classification was arrived at by the model. For instance, linear regression and decision tree models are interpretable because we can comprehend how the models make predictions or decisions.

Explainability consider a broader aspect of XAI. It aims to answer the question of 'why'. It focuses on providing clear, understandable reasons for why a specific decision was made by the AI model. It is the process of communicating the rationale behind a model's prediction or output in a way that stakeholder and understand irrespective of the technicalities. In a nutshell, explainability relies on interpretability as a building block (TechDispach, 2022).

2.2.1. Concepts of XAI

To understand the concept of explainable AI, one needs to know that unlike conventional programming which follows strictly the algorithm written by the programmer, complex AI models make their own decisions building correlations and relationships among neurons with the input features to make decisions or predictions. In some cases, the input variables can be in thousands and the neurons in hundreds of thousands. These neurons build patterns and associations among themselves and with the input features on their own, making it difficult for even AI experts to ascertain and explain the underlying processes that birth those predictions (Peters, 2023). This resulting situation is described as a "black box", and this situation affects trust and confidence in AI systems.

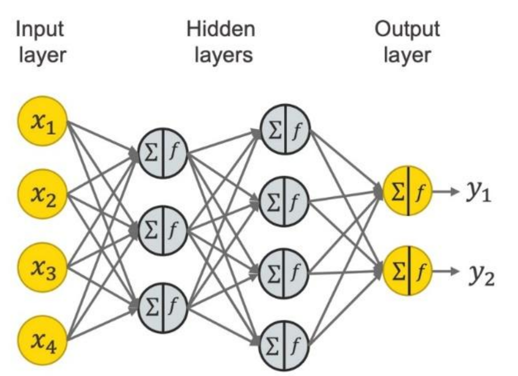

Figure 1 shows the architecture of the neural network (Melcher, 2021)

From Figure 1, the output represented by x_1, x_2, x_3, $and\ x_4$, can run to thousands of output in a complex system and the hidden layers can have thousands of layers with each layer containing hundreds to thousands of neurons. The neurons sometimes learn differently and each neuron feed its result to the succeeding neuron, which serves as the output of the succeeding neuron. The learning process continues until the final output is obtained.

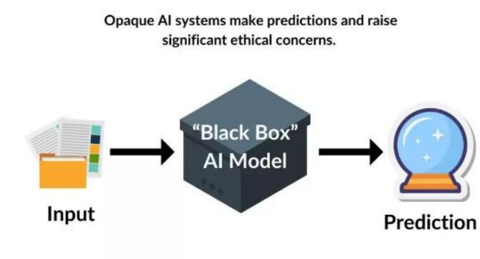

Figure 2 shows the opaque AI system (Neri, 2024)

The opacity of AI processes not only makes accepting AI solutions difficult but also has direct impacts on individuals since the processes involved in the AI model to arrive at its decision are unknown (Pearl, 2019). For instance, if an AI model is built to select the best candidates for an international job position, the AI model might choose people from certain demographics based on several reasons. The problem here is the recruiters might not be able to explain whether to make appropriate corrections or to integrate the AI solution with human reasoning for better decision-making, thus making it impossible to address the bias. Another situation might be in medical diagnosis, where the prediction of the AI results cannot be evaluated for bias.

From ongoing research about XAI, two major approaches to explainability have been put forward: self-explainable AI and post hoc explainable AI.

2.2.2. Self-explainable AI

Self-explainable AI also referred to as "white box" models provides its explanation during training (Xu, 2018). This model provides easy-to-understand algorithms that reveal how the data input affects the output or results (Thampi, 2002). Two common examples of the "white box" models are the decision trees and linear regression. In an email classification model, the decision tree algorithm could determine if the emails are spam or not spam. The decision tree algorithm is first trained to identify the features of spam and not spam emails. In making decisions about whether test data (emails) are spam or not spam, the algorithm divides the data recursively into binary partitions by calculating the entropy of the data. This process is recursive until it creates a tree-like structure. At each node, the tree selects the feature with the highest information gain to classify the emails.

However, more complex architecture like neural networks, which have provided incredible solutions such as generative AI, computer vision, and virtual assistance consist of multiple layers of interconnected neurons with each layer performing complex computations and passing its results to the next are virtually not self-interpretable (Lipton, 2018), so are referred to as "black box". Therefore the post hoc approach seems to be more appropriate in complex systems.

2.2.3. Post hoc AI Approach

The post hoc AI approach involves explaining how a complex AI model works after the model has been built or after the model makes decisions (Xu, 2018). The post hoc AI approach can be classified into two categories: local and global (Thampi, 2002). The global approach provides an overall understanding of the behavior and functioning of the model and its decision-making processes. It aims to capture general trends and patterns and provide broad insight into the model's behavior. An example of a global explanation is when AI experts try to interpret the reasons behind a model's decision-making by looking at the feature's importance (Breiman, 2001).

In music recommendation, an AI model might recommend some specific type of music to a user based on certain features displayed in the user selection of a type of music. If the user loves Jaz music, for instance, the AI model has hundreds of Jaz music to suggest to the user. However, how does it decide which one? It might consider the user’s listening history, genre preferences, and song metadata to make recommendations. Therefore, studying these features to decide how an AI model makes decisions is an example of a global explanation. Another example is "rule extraction" (Craven, 1996). AI experts and end-users can have an idea of how complex AI systems make decisions by looking at the rules. This was prominent during the symbolic AI era (first generation type of AI) that makes decisions based on rules. For instance, if the patient's age > 50 and blood pressure is high, then diagnose hypertension.

The local explanation focuses on specific results of decision-making by an AI model. Instead of trying to understand the general behavior of an AI model, the local explanation is more of why an AI model makes a particular decision. For instance, why did the AI model diagnose that Mr. A has cancer instead of trying to explain the general decision-making processes of an AI model? Two popular local explanation models are LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations) (Ribeiro, 2016; Lundberg, 2017).

The LIME model works by changing or manipulating the input data by creating a series of artificial data changing a part of the original attributes of the data, and observing how the output changes and the model behaves (Ribeiro, 2016). From the observation, the LIME creates an interpretable “surrogate model” to help understand how to make decisions (Ribeiro, 2016). These surrogate models are easier to understand, aiding users understanding. SHAP, on the other hand, is a method that works with the principle of game theory. It assigns values to each feature of the model and calculates the contribution or impact of each feature to the prediction of a specific instance (Lundberg, 2017). In the process, it considers the contributions of all possible features, which provides a unified measure of feature importance and helps explain the model’s decision at a local level. Consider an AI model which helps to predict the price of a house using SHAP. Each feature such as the number of rooms, distance to the golf course, size of land, etc., are assigned specific values. The values are varied to see the contribution and impacts of each feature on the predictions of the model. This way, the end-user would understand how the model makes predictions and the most significant features that impact the output.

There are other XAI models but most of them are coined from LIME and SHAP. Some of them are GraphLIME, Anchors, Layer-wise Relevance

Propagation (LRP), Deep Taylor Decomposition (DTD), Prediction Difference Analysis (PDA), Testing with Concept Activation Vectors (TCAV), etc. We will not explain what this XAI entails as this is beyond the scope of this study.

2.3. XAI in Management and entrepreneurship

Management is a multifaceted process that involves planning, organizing, leading, and controlling resources to achieve organizational goals, though definitions such as "planning, organizing, leading, and controlling to achieve results with people" can feel insufficient (Boddy, 2017; Koontz, 1961). As a broad human activity spanning various domains, it involves both a functional practice of managing and the role of managers who oversee operations (Drucker, 1954). In contrast, defining entrepreneurship is similarly complex, as it encompasses diverse perspectives related to opportunity pursuit, business creation, and profit-seeking across disciplines like economics, business, and psychology (Bennett, 2006; Shane & Venkataraman, 2000). The lack of a universally accepted definition of entrepreneurship reflects the discipline's interdisciplinary nature, leading to varying conceptualizations (Baker & Welter, 2017).

2.3.1. XAI in Management

The role of XAI in management cannot be over-emphasized. In management particularly in marketing, operations and strategic planning, XAI is highly essential as it aids to bridge the gap between machine learning decision-making models and managerial decision-making. In other words, managers cannot trust AI tools for critical decision-making especially in a dynamic business landscape because they do not understand how the AI tools got to the predictions. However, XAI is meant to provide explanations into the model decision-making processes, thus promoting trust and acceptance. Thus, XAI enhance trust and accountability use (Gilpin et al., 2018). Furthermore, by using XAI in performance management systems, organizations can improve employee evaluation and decision-making processes, reducing biases and increasing fairness. For instance, XAI can provide unbiased explanation why certain employees should be considered for promotion Molnar et al., 2020). XAI can help in supply chain optimization by providing insights and logical explanation behind inventory and logistics decisions, which is crucial for achieving operational efficiency (Briggs et al., 2021).

2.3.2. XAI in Entrepreneurship

For most entrepreneurs, understanding the market trends, customer preferences, and predicting future offerings are usually daunting tasks. XAI offers entrepreneurs valuable insights and provide explanation into market trends, customer preferences, and business performance, enabling entrepreneurs not only make informed decisions suggested by AI tools but also help them understand why those decisions are suggested. This enables entrepreneurs to either finetune this AI predictions to better suit the market trends or accept them or reject them. Besides, Entrepreneurs often operate in environments characterized by risk, ambiguity, and dynamic changes, which makes XAI a critical tool for enhancing decision-making and fostering innovation (González et al., 2020).

Furthermore, XAI can assist entrepreneurs in identifying investment opportunities and providing explanation on why the decisions should be pursued. XAI can also predict potential challenges, and mitigate risks by providing clear explanations for market trends or financial forecasts (Ribeiro et al., 2016). The transparency of XAI allows entrepreneurs to trust AI-driven insights, especially in areas like customer segmentation, pricing models, and demand forecasting, where understanding the "why" behind a decision is essential. Additionally, as startups often have limited resources, XAI can provide small businesses with the ability to compete with larger enterprises by offering them tools to improve decision-making processes without relying heavily on specialized expertise (Binns et al., 2020). In this way, XAI levels the playing field for entrepreneurs, enabling them to innovate, scale, and achieve sustainable growth while navigating the complexities of the entrepreneurial landscape.

III. RESEARCH METHODOLOGY

The research methodology outlines the approach used in research. In this research, we adopt the literature review method to obtain information from relevant articles and reports about XAI in management and entrepreneurship. The objective is to critically review, synthesize, and draw conclusions from existing research documents to understand the application and challenges of XAI in management and entrepreneurship.

3.1. Scope of the Literature Review

Time Frame: Research papers published in the last ten years (2013-2023) will be prioritized to ensure the review reflects the most up-to-date research.

Types of Sources: Peer-reviewed journal articles, conference papers, working papers, industry reports, and case studies

Areas of Focus: The study focuses on the application of XAI and the challenges in adopting XAI in management and entrepreneurship.

3.2. Search Strategy

A systematic approach is adopted to identify, collect, and select relevant literature. Relevant articles and reports are searched from popular databases such as Google Scholar, Scopus, IEEE Xplore, SpringerLink, JSTOR, and ResearchGate will be searched. Long-tailed and short-tailed keywords are searched for. Some of the examples are Explainable AI, XAI in management, Artificial Intelligence in management, AI in entrepreneurship, XAI in management and entrepreneurship, application of XAI in management, application of XAI in entrepreneurship, challenges of adopting XAI in management and entrepreneurship.

3.3. Inclusion and Exclusion Criteria

The inclusion criteria identify certain conditions considered in the study to determine which articles and reports to include in the study. Below are the inclusion criteria considered:

- Publications related to XAI and its application in business management and entrepreneurship.
- Research articles discussing frameworks, models, applications, and case studies of XAI in decision-making or entrepreneurship contexts.
- Peer-reviewed journals, conference papers, and industry reports within the last 10 years

Exclusion Criteria:

- Papers focused solely on theoretical AI models without practical applications.
- Articles not available in English or those lacking full-text access.
- Articles and reports older than 2015.

3.4. Data Extraction and Organization

The title, authors, publication year, journal/conference, theme, scope of study, type of research, sample size, research methods, findings, limitations and recommendations are extracted from the literature. This information is systematically included during and research findings and discussion are properly organized. The data extracted from the literature are categorized and discussed under three the research questions

3.4. Data Extraction and Organization

The title, authors, publication year, journal/conference, theme, scope of study, type of research, sample size, research methods, findings, limitations and recommendations are extracted from the literature. This information is systematically included during and research findings and discussion are properly organized. The data extracted from the literature are categorized and discussed under three the research questions

IV. DATA ANALYSIS

This chapter provides literature reviews and insight and draw conclusion on the three research questions:

1. How does Explainable AI (XAI) enhance decision-making transparency and trust in management and entrepreneurship?
2. What are the key applications of XAI in managerial decision-making and entrepreneurial ventures, and how do they impact business performance?
3. What challenges and ethical considerations arise in

the adoption of XAI within management and entrepreneurship, and how can they be addressed?

4.1. Research Question 1: XAI in Decision-Making in Management and Entrepreneurship

Decision-making is an important function of management and AI has been a vital tool. However, real-world AI models lack interpretability, causing limitation and reliability concerns in critical decision-making. According to Miller (2019), Ribeiro et al. (2016), and Lipton (2016), managers and CEOs are more likely to trust AI in making critical decisions in a dynamic business landscape if they understand the reasoning behind AI-generated predictions. Similarly, the work of Doshi-Velez and Kim (2017), which focuses on the impacts of XAI in managerial decision-making, shows that interpretability increases confidence in AI systems, particularly in strategic planning and resource allocations. The study done by Arrieta et al. (2020) to find out about the effects of XAI on cognitive biases suggests that XAI reduces cognitive bias by providing transparent insights, enabling managers to make decisions from an informed position with greater confidence.

The research by Olesja, et al (2024) compares four strategies to provide explanations by a decision support system, represented by the social agent Floka, which was designed to assist users in decision-making during uncertainty. 742 participants who make lottery decisions participated in the study. Two explanations that prioritize accurate explanations (transparent vs. guided) are compared with another two strategies that prioritize human-centered explanations (emotional vs. authoritarian). Findings show that a guided explanation strategy results in higher user reliance than a transparent strategy. The research shows that users trust guided explanation AI systems.

Nitin, R., Saurabh, C., & Jayesh, R. (2023) implemented XAI techniques to introduce transparency into financial AI systems. Various XAI methods, including rule-based systems, model-agnostic approaches, and interpretable machine learning models are used to determine their effectiveness in producing interpretable AI-driven financial solutions. Findings show XAI does not only promote transparency and accountability but also addresses ethical concerns, and promotes trust in AI models in the financial sector.

Overall, the literature reveals that managers and entrepreneurs are more likely to trust AI solutions if they are interpretable. Besides, the literature reveals that guided XAI improves confidence in decision-making, particularly in strategic planning and resource allocations, reduces cognitive bias, and promotes transparency and accountability. However, incorrect XAI has a negative impact on the trust and reliance of humans on AI solutions.

4.2. Research Question 2: Applications of XAI

4.2.1. XAI in Management

The application of XAI in management and entrepreneurship cannot be over-emphasized. XAI enhances decision-making, helps with risk assessment, and promotes operational efficiency by providing transparent and interpretable insights into AI solutions (Miller, 2019). In management, XAI helps managers understand AI recommendations, improving trust in the system. It also enhances accountability in strategic planning, human resource management, and performance optimization (Ribeiro et al., 2016). Decision making

- *XAI Promotes Operational Efficiency*

XAI contributes immensely to operational efficiency by enabling organizations to monitor, understand, and optimize AI processes. According to Gunning et al. (2019), when machine learning output becomes clearer and understood, humans can act on the AI insight faster, resulting in improved system responsiveness. In healthcare, transportation sector, and logistics, operational decision efficiency has been enhanced by 30% through the use of XAI (Gunning et al., 2019). Similarly, Chari et al. (2020) explored the impacts of XAI on operational efficiency in management and found that XAI improved collaboration between AI tools and managers, leading to better decision-making. Their findings justify that XAI does not only improve efficiency but also enhances understanding and collaboration among technical teams.

- *XAI Build Trust in AI Systems*

One of the challenges of AI models is trust. XAI helps managers to build trust in AI systems by making the systems transparent and understandable. Ribeiro et al. (2016) introduced the LIME XAI which promotes understanding of AI model individual predictions. Their findings show that managers are more likely to trust AI systems if they understand how they make predictions. Binns et al. (2018) carried out a significant study to ascertain the perception of managers and users of AI model decisions when the explanation is provided. They found that explanation improves user's confidence in the systems, especially in management, decision-making, and even in sensitive domains such as finance and criminal justice.

- *XAI in Decision-making*

Decision-making is another aspect that AI has been so useful, though managers still struggle to trust AI systems in significant decision-making. with XAI, managers will learn to trust AI systems over time. Lundberg and Lee's (2017) work on SHAP (Shapley Additive exPlanations), an XAI model, has been helpful in healthcare and finance. They use SHAP to make decisions for medical diagnosis and loan approval. Amann et al. (2020) reviewed the application of XAI in clinical decision-support systems and found that clinicians are more likely to rely on AI to support clinical decision-making if they understand how and why AI models make predictions and when its working aligns with clinical knowledge

- *Enhance Accountability in Strategic Planning*

XAI enhances accountability in strategic planning by providing auditable recommendations. Doshi-Velez and Kim (2017) highlight the importance of interpretability in ensuring that AI recommendations in critical policy and business decisions can be reviewed, justified, and corrected if necessary. They argue that explainability acts as a safeguard for aligning AI actions with human values and institutional goals. Besides, Watson et al. (2022) investigated the use of XAI in finance and found that XAI increases accountability by revealing strategic missteps. They found that XI improves post-decision audits by 25%.

- *XAI in Risk Management*

AI tools have been useful in predicting the risks involved in certain managerial decisions and suggesting appropriate solutions to reduce them. The research done by Ribeiro, Singh, and Guestrin (2016) proposes that SHAP and LIME models can help with risk predictions and explanations, framing the task as a submodular optimization problem. They demonstrated the flexibility of SHAP and LIME in risk management by using random forests and image classification AI models. Findings show that XAI helps managers assess operational risks more accurately.

Moreover, research has shown that XAI can enhance fraud detection and cybersecurity in financial management. Lluís, Anaya, and Jaume (2023) investigated several XAI techniques to equip risk managers with more XAI methods. They used a database of real universal-like policies to fit into a logistic regression model and several tree-based models. Then, they use SHAP to provide interpretable perspectives. Findings show that non-trivial ideas can emerge to improve paid-up risk management. The work of Niklas, et al (2020), uses XAI to determine the risk involved when credit is borrowed employing peer-to-peer lending platforms. The research employs the SHAP XAI model. An empirical analysis of 15,000 small and medium companies that ask for peer-to-peer lending credit reveals that both risky and not risky borrowers can be classified according to their financial characteristics, which is used to explain and under their credit scores.

- *XAI in Performance Optimization*

Performance Optimization involves refining the efficiency of systems to improve their speed, outputs, and scalability. It employs the scarce resources available and puts them into optimal use. AI-driven analytics play a crucial role in this domain, but without interpretability, their insights lack trust and reliability. The research done by Yogendr, et al (2024) delves into the integration of machine learning methodologies, with a specific focus on (XAI) models, within the domain of fiber optic surface plasmon resonance (SPR) sensors. They use XAI in a trained Gaussian Process Regression (GPR) model to gain insight into the operations of the AI model, thus providing insights into its decision-making process.

Furthermore, the study by Baker and Welter (2017) explores the impact of explainability on performance optimization in financial AI models, suggesting the need to balance interpretability with performance. The research investigated how to mitigate the tradeoffs between accuracy in performance optimization with interpretability. The research found that interpretability enhances trust and reliability in the AI model and processes, however, accuracy is likely to be affected by explainability. Lipton's (2018) research on interpretability on performance optimization enhances AI-driven performance metrics by providing managers with insights into the overall workings of the AI models.

4.2.2. XAI In entrepreneurship

In entrepreneurship, XAI provides explainable information about market predictions, investment risk analysis, and insights into customer behavior, enabling startups to make informed decisions (Lipton, 2016). Moreover, XAI helps startups and business managers comply with regulatory requirements by providing clear justifications for operational and financial decisions.

- *XAI in Market Predictions*

XAI is very useful in entrepreneurship for market prediction. It provides transparent forecasts, enhancing strategic planning and improving investor confidence. The study done by Tjoa and Guan (2020) reveals that explainable predictive AI models help startups study customer behavior, market trends, and market demands at little or no cost. Similarly, Guidotti et al. (2019) applied XAI models to time-series prediction and found that entrepreneurs benefit from the ability of AI models to provide explanations that can be understood by managers. These explanations are crucial for startups to improve planning and reduce risks.

- *XAI in Investment Risks Analysis*

Risk assessment is highly significant for startups and investors. XAI helps in assessing risks by making complex financial predictions and providing understandable explanations for those predictions. Barredo Arrieta et al. (2020) work demonstrated that XAI can help flag high-risk investments by pinpointing the factors that may enhance the failure of the investments. Similarly, Molnar (2022) emphasized the role of XAI in credit risk scoring, and startups to understand the rationale behind funding decisions.

- *Insights into customer behavior*

Entrepreneurs need to understand customers' behavior to develop engagement and value propositions. XAI provides insights into customers' behavior making it easier for startups to tailor their offerings. Binns et al. (2018) findings reveal that transparency in recommendation systems promotes customers' trust, sales conversion, and customer engagement, revealing that XAI systems are more effective in shaping customers' perceptions. Ribeiro et al. (2016) study finds something noteworthy about customers' responses to business campaigns. They found that the LIME framework helps entrepreneurs identify why customers respond to specific campaigns, enabling informed marketing campaigns.

4.3. Challenges of XAI

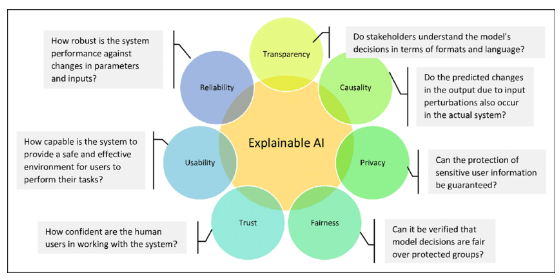

Figure 3 shows brief concerns about XAI in management

4.3.1. Tradeoffs between accuracy and interpretability

XAI has faced several challenges and one of the crucial dilemmas of XAI is choosing between white-box models which are easier to understand but less accurate and more dynamic accurate AI models termed black-box, which offer higher accuracy but are difficult to understand (Ding et al. 2022). Therefore, there is a concern about the trade-off between accuracy and interpretability in the context of AI models. According to Ding et al. (2022), complex AI models such as neural networks have been proven to improve accuracy but there are situations where complexity does not necessarily lead to improved performance. However, it has been generally accepted and empirically proven that accuracy reduces when interpretability becomes a priority.

4.3.2. Privacy and security considerations

Another crucial concern is choosing between privacy and security concerns and interpretability. The concern that diverse data sources are omnipresent already presents a huge concern about data privacy and security. When confronted with XAI, these concerns intensify, presenting the need to strike a balance between interpretability and security. The intrinsic nature of XAI which aims to make AI processes and decision-making more transparent presents a concern of exposing the data used to train the model.

Since AI models grapple with private data, it is data regulation and ethical to protect individual data (Ding et al. 2022). However, the XAI domain raises profound concerns regarding data protection during training and inference. This concern presents future research to ascertain how data could be protected during interpretability. Furthermore, research that could determine the lost privacy of XAI is of utmost importance (Ding et al. 2022). This will ascertain if AI interpretability is worth its investment in the first place.

4.3.3. Developing XAI evaluation metrics

This challenge presents a noteworthy concern in XAI. The simple question AI experts need to ask is: how do we evaluate the accuracy of XAI on the one hand, and how do we evaluate the effects of the accuracy of AI models while promoting interpretability on the other hand (Pawlicka et al. 2024)? Ding et al. (2022) argued that to establish a strong foundation of XAI, one or more evaluation metrics for XAI models are of utmost importance. Without these metrics, the push for XAI is weakened and lacks consistency (Ding et al., 2022).

The lack of uniform definitions for explainability, interpretability, and transparency in XAI poses a huge challenge to the development of XAI evaluation metrics. How do we develop evaluation metrics if interpretability means different things to different people? Therefore, a uniform definition of interpretability and explainability needs to be ascertained before talking about XAI metrics evaluation. At the moment, XAI performance is measured by two methods: objective methods which rely on analytical and mathematical evaluation, and human-centered evaluation which hinges on end-user observations.

4.3.4. Increasing complexity of AI models

In the early development of artificial neural networks, the network is just a few neurons in one layer and those are quite interpretable. However, to build AI systems that can solve real-world problems, a one-layer network cannot suffice. Generative AI for instance, which uniquely provides answers to most human queries and which can even answer sophisticated questions quite like humans has an extremely large architecture with thousands of layers, which makes it quite impossible to interpret.

V. FINDINGS AND DISCUSSION

XAI is an important aspect of AI that is evolving and it's quite an important field as it determines if AI solutions will be accepted for critical business operations and decision-making. From the exploration of literature about XAI in management and entrepreneurship: applications and challenges, findings established that AI solutions have been deployed into management and entrepreneurship and it has proven to be effective in business operations. Findings reveal that XAI plays a crucial role in enhancing decision-making, risk management, business performance, predictive analysis in entrepreneurship, and even resource allocations.

However, AI solutions present a huge challenge. High-accurate AI solutions are not explainable. This challenge is affecting the total adoption of AI solutions in management and entrepreneurship especially in crucial decision-making. This study shows that managers and entrepreneurs are more likely to trust AI solutions if they are interpretable. Besides, the literature reveals that guided XAI improves confidence in decision-making, particularly in strategic planning and resource allocations. XAI is found to reduce cognitive bias, promote transparency and accountability, and also address ethical concerns.

However, incorrect XAI has a negative impact on the trust and reliance of humans on AI solutions. The study revealed that businesses leveraging XAI in human resource management, financial analysis, and strategic planning experience increased efficiency and reduced cognitive biases.

This study also explores the challenges militating against XAI. Several challenges were identified in this study and a few of them include the trade-off between explainability and model accuracy, privacy and security concerns, lack of XAI evaluating metrics, and increased complexities of AI models. This study shows that despite hundreds of research in XAI, XAI still has a long way to go in achieving its aims. Future studies should focus on uniformly acceptable definitions of XAI, XAI evaluation metrics, XAI impacts on data security, and building more robust XAI models that can interpret complex AI systems like those of neural networks.

CONCLUSION

Explainable AI (XAI) is a crucial aspect of AI model development. Success in XAI will skyrocket the use of AI in management and entrepreneurship, especially in crucial decision-making and resource allocations. Presently, XAI is an evolving field though several successes have been achieved. Two approaches are used in XAI: the self-explainable approach and the Post hoc AI Approach. Examples of the Self-explainable approach model are decision trees and linear regression. The post hoc AI Approach requires that an interpretation model be built to explain already existing AI solutions and the two commonest post hoc models are LIME and SHAP.

XAI has been so useful in management and entrepreneurship by enhancing transparency, trust, and decision-making in AI-driven systems. This study covers three research questions which are how XAI enhances decision-making trust in management and entrepreneurship, key applications of XAI in management and entrepreneurship, and the challenges and ethical considerations that arise in the adoption of XAI in management and entrepreneurship.

The study shows that XAI plays a crucial role in enhancing decision-making, risk management, business performance, predictive analysis in entrepreneurship, and even resource allocations. However, AI solutions present a huge challenge. High-accurate AI solutions are not explainable. This challenge is affecting the total adoption of AI solutions in management and entrepreneurship especially in crucial decision-making. If XAI can finally be successful, the adoption of AI solutions will skyrocket beyond imagination.

6.1. Recommendation

1. More research needs to be done on XAI. XAI is an evolving aspect of AI and there are so many AI applications which are not understandable.
2. XAI should be taken into consideration when designing AI models. Presently, XAI is an afterthought and most AI models in place todays are not built with XAI in mind.
3. The tradeoff between interpretability and accuracy is a source of concern. There is a need to build AI systems which are highly accurate and fairly interpretable. This is highly mandatory for users to build trust in AI systems.
4. Organizations and startups should embrace the use of AI tools in resource allocation, predictive analysis, decision making and many more. The advantages of using AI undoubtedly outweighs the disadvantages.

6.2. Limitation and Future Research

Future research should focus on building AI systems which are explainable. XAI should be not an afterthought but should be a requirement when building AI systems. Besides, researchers need to do more to the tradeoff between accuracy and interpretability. Future research should focus on building systems and models which will negate this principle of AI: the more accurate a system is, the less interpretable it is.

ABBREVIATIONS

XAI – Explainable Artificial Intelligence

LIME – Local Interpretable Model-agnostic Explanations

SHAP – Shapley Additive Explanations

AUTHOR CONTRIBUTION

Adebayo Rotimi Philip is the sole author of the manuscript

REFERENCES

[1] C. Agarwal and A. Nguyen, “Explaining image classifiers by removing input features using generative models,” Lect. Notes Comput. Sci., vol. 12627, pp. 101–118, accepted for publication.

[2] B. A. Alejandro, N. Díaz-Rodríguez, and J. Del Ser, “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI,” Inf. Fusion, vol. 58, p. 82, accepted for publication.

[3] J. Amann, A. Blasimme, E. Vayena, D. Frey, and V. I. Madai, “Explainability for Artificial Intelligence in Healthcare: A multidisciplinary perspective,” BMC Med. Inform. Decis. Mak., vol. 20, no. 1, p. 310, accepted for publication.

[4] R. Anderson and M. Starnawska, “Research practices in entrepreneurship: Problems of definition, description, and meaning,” Int. J. Entrep. Innov., vol. 9, no. 4, pp. 221–230, accepted for publication.

[5] T. Baker and F. Welter, “Come on out of the ghetto, please! Building the future of entrepreneurship research,” Int. J. Entrepreneurial Behav. Res., vol. 23, no. 2, pp. 170–184, accepted for publication.

[6] A. Barredo Arrieta et al., “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” Inf. Fusion, vol. 58, pp. 82–115, accepted for publication.

[7] R. Bennett, “Business lecturers’ perceptions of the nature of entrepreneurship,” Int. J. Entrepreneurial Behav. Res., vol. 12, no. 3, pp. 165–188, accepted for publication.

[8] R. Binns, M. Veale, M. Van Kleek, and N. Shadbolt, “It’s reducing a human being to a percentage: Perceptions of justice in algorithmic decisions,” in Proc. CHI Conf. Human Factors Comput. Syst., pp. 377:1–377:14, accepted for publication.

[9] D. Boddy, Management – An Introduction, 7th ed. Harlow: Pearson Educ., accepted for publication.

[10] L. Breiman, “Random forests,” Mach. Learn., vol. 45, pp. 5–32, accepted for publication.

[11] J. Burrell, “How the machine ‘thinks’: Understanding opacity in machine learning algorithms,” Big Data Soc., vol. 3, no. 1, accepted for publication.

[12] S. Chari, A. Aggarwal, K. Chatterjee, and J. White, “Directions for responsible and explainable artificial intelligence,” in Proc. 1st ACM Conf. Fairness, Accountability, Transparency AI, pp. 123–132, accepted for publication.

[13] J. Chen et al., “The impact of explainable AI on decision-making,” Int. J. Bus. Anal., vol. 8, no. 3, pp. 15–30, accepted for publication.

[14] M. W. Craven, “Extracting comprehensible models from trained neural networks,” Univ. of Wisconsin-Madison, submitted for publication.

[15] J. Dastin, “AI systems face scrutiny over ethical concerns,” New York Times, 2018, submitted for publication.

[16] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint, arXiv:1810.04805, submitted for publication.

[17] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” arXiv preprint, arXiv:1702.08608, submitted for publication.

[18] P. F. Drucker, The Practice of Management, 1st Harper business ed., New York: HarperCollins, 1993.

[19] B. H. Emil and B. Simon, “Artificial intelligence and Internet of Things in SMEs,” J. Manuf. Syst., vol. 60, pp. 362, accepted for publication.

[20] W. J. Eschenbach, “Transparency and the black box problem: Why we do not trust AI,” Philos. Technol., vol. 34, pp. 1607–1622, accepted for publication.

[21] European Commission High-Level Expert Group on Artificial Intelligence, “Ethical guidelines for trustworthy AI,” submitted for publication.

[22] European Commission, "Fostering a European approach to artificial intelligence," COM(2021) 201 final, submitted for publication.

[23] European Commission, "Proposal for a regulation laying down harmonized rules on artificial intelligence," COM(2021) 206 final, submitted for publication.

[24] L. H. Gilpin et al., "Explaining explanations: An overview of interpretability of machine learning," in Proc. CHI Conf. Human Factors Comput. Syst., submitted for publication.

[25] R. Guidotti et al., "A survey of methods for explaining black box models," ACM Comput. Surv., vol. 51, no. 5, pp. 1–42, accepted for publication.

[26] D. Gunning and D. W. Aha, "DARPA's explainable artificial intelligence program," AI Mag., vol. 40, no. 2, pp. 44–58, accepted for publication.

[27] D. Gunning et al., "XAI—Explainable artificial intelligence," Sci. Robot., vol. 4, no. 37, eaay7120, accepted for publication.

[28] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," Nat. Mach. Intell., vol. 1, no. 9, pp. 389–391, accepted for publication.

[29] K. Matelyn et al., "The impact of imperfect XAI on human-AI decision-making," Human-Comput. Interact., submitted for publication.

[30] H. Koontz, "The management theory jungle," J. Acad. Manage., vol. 4, no. 4, pp. 174–188, accepted for publication.

[31] B. O. Lepri, "Fair, transparent, and accountable algorithmic decision-making processes," Philos. Technol., vol. 31, pp. 611–627, accepted for publication.

[32] Z. C. Lipton, "The mythos of model interpretability," Queue, vol. 16, no. 3, pp. 31–57, accepted for publication.

[33] P. J. G. Lisboa, "Interpretability in machine learning: Principles and practice," in Proc. Int. Workshop Fuzzy Logic Appl., pp. 15–21, accepted for publication.

[34] L. B. A. Lluís, D. B. Anaya, and J. B. C. Jaume, "Explainable AI for paid-up risk management in life insurance products," Finance Res. Lett., vol. 57, no. 2, 104242, accepted for publication.

[35] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Adv. Neural Inf. Process. Syst., vol. 30, pp. 4765–4774, accepted for publication.

[36] K. Melcher, "A friendly introduction to deep neural networks," https://www.knime.com/blog/a-friendly-introduction-to-deep-neural-networks, submitted for publication.

[37] T. Miller, "Explanation in artificial intelligence: Insights from the social sciences," Artif. Intell., vol. 267, pp. 1–38, accepted for publication.

[38] T. H. Miller, "Explainable AI: Beware of inmates running the asylum," arXiv preprint, arXiv:1712.00547, submitted for publication.

[39] C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2nd ed., submitted for publication.

[40] V. O. Neri, "Explainable AI made simple," https://spotintelligence.com/2024/01/15/explainable-ai/, submitted for publication.

[41] N. Buechner, P. Giudici, and D. Moneta, "Explainable AI in fintech risk management," Front. Artif. Intell., vol. 2, no. 5, submitted for publication.

[42] R. Nitin, C. Saurabh, and R. Jayesh, "Explainable artificial intelligence approaches in financial decision-making," SSRN, submitted for publication.

[43] L. Olesja et al., "Humans in XAI: Increased reliance in decision-making," Front. Behav. Health Ethics, vol. 3, 1377075, submitted for publication.

[44] J. Pearl, "The limitations of opaque learning machines," in J. Brockman (Ed.), Possible Minds: 25 Ways of Looking at AI, Penguin, pp. 13–19, submitted for publication.

[45] U. Peters, "Explainable AI lacks regulative reasons," AI Ethics, vol. 3, no. 3, pp. 963–974, accepted for publication.

[46] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in Proc. 22nd ACM SIGKDD Int.

Conf. Knowl. Discov. Data Min., pp. 1135–1144, accepted for publication.

[47] S. Shane and S. Venkataraman, "The promise of entrepreneurship as a field of research," Acad. Manage. Rev., vol. 25, no. 1, pp. 217–226, accepted for publication.

[48] S. Sharma and P. Chaturvedi, "Explainable AI and its importance," J. Manage. Entrep., vol. 10, no. 2, submitted for publication.

[49] TechDispatch, "Explainable AI," https://www.aepd.es/documento/techdispatch-xai.pdf, submitted for publication.

[50] V. Thampi, Interpretable AI: Building Explainable Machine Learning Systems, Simon & Schuster, submitted for publication.

[51] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 11, pp. 4793–4813, accepted for publication.

[52] A. Watson, L. Hassan, and R. Muldoon, "Explainable AI and the future of strategic decision-making," J. Bus. Strategy, vol. 43, no. 1, pp. 28–36, accepted for publication.

[53] H. S. Xu, "Deepobfuscation: Securing the structure of convolutional neural networks via knowledge distillation," arXiv preprint, arXiv:1806.10313, submitted for publication.

[54] S. D. Yogendr et al., "On the application of explainable AI in optimizing fiber optic SPR sensor performance," Opt. Fiber Technol., vol. 85, 103801, accepted for publication.