

NEURO-SYMBOLIC AI FOR AUTONOMOUS ETHICAL DECISION-MAKING IN BRIDGING HUMAN VALUES WITH MACHINE INTELLIGENCE

Balasubramani Murugesan

Director of Engineering, Digit7, United States of America

ABSTRACT

This research explores the potential of neuro-symbolic AI in bridging the gap between human values and the machine intelligence. Neuro-symbolic AI combines the neural networks with symbolic reasoning to deliver an ethical reasoning capability in artificial intelligence systems. Through the integration of the neural networks and symbolic reasoning, neuro-symbolic AI seeks to enhance the learning, reasoning and decision-making capabilities of the AI system. This study examines the theoretical underpinnings, applications, and ethical considerations encompassed by neuro-symbolic AI in various domains such as academia, healthcare and business. Through the utilization of the qualitative research methodology, the study analyzes the existing literature to address the main themes and potential challenges. The findings underline the encouraging potential of neuro-symbolic AI in the identification of the limitations of traditional AI approaches and promote human-centered AI development. The findings from the research underline that neuro-symbolic AI can contribute to the explainability, transparency, and fairness of AI. However, there are still open challenges on scalability, interpretability, and cross-cultural adaptability. Future research should be devoted to overcome the limitations above and foster further interdisciplinary collaboration on ethical development and deployment associated to neuro-symbolic AI systems. This paper focuses on the exploration of the innovative methods to integrate fairness into AI models, leveraging Explainable AI (XAI) tools such as SHAP to mitigate biases during training. The framework autonomously maintains fairness through diversified data handling, with applications in income prediction, credit risk, and recidivism. This paper analyzes the neuro-symbolic AI approaches for ethical reasoning, highlighting their integration of deep learning and symbolic logic to enhance transparency, explainability, and decision-making. Applications in healthcare, finance, and education are explored, with a focus on aligning AI with human values, ensuring ethical, trustworthy outcomes across diverse domains.

KEYWORDS: *Neuro-Symbolic AI, Ethical Decision-Making, Human Values Integration, Artificial Neural Networks (ANN), Machine Intelligence, Symbolic Reasoning, Explainability in AI.*

Article History

Received: 22 Nov 2024 | Revised: 26 Nov 2024 | Accepted: 12 Dec 2024

1. INTRODUCTION

1.1 Background

In the present landscape of rapid digitalisation, Neuro-symbolic AI is expanding its area as a combination of neural and symbolic AI architectures. It is widely recognised as neuro-symbolic artificial intelligence which integrates the strengths of statistical AI such as machine learning with the abilities of human-centric symbolic knowledge. The traditional neural

networks work with pattern recognition with the goal of constructing artificial general intelligence that assists in reasoning, learning and modelling cognition [1]. This approach leverages the opportunities of the neural networks in pattern recognition and the strengths of the symbolic AI in the decision-making process. The AI techniques are integrated significantly into the critical decision-making process which expands the way for making ethical decisions which align with human values. This integration helps to identify the current AI systems, especially in the areas that require ethical decision-making that aligns with human values. This research focuses on examining the neuro-symbolic AI for autonomous ethical decisions for integrating human values with machine intelligence.

1.2 Aim and Objectives

Aims

The research aims to examine the neuro-symbolic AI framework that integrates into autonomous ethical decision-making leading to bridging the gap between human values and machine intelligence.

Objectives

- To analyse neuro symbolic AI approach for making ethical reasoning capabilities
- To explore potential applications of the neuro symbolic AI in different domains
- To assess the influence of this neuro-symbolic AI in the integration of human values with machine intelligence

1.3 Research Question

- *RQ 1:* What is the neuro-symbolic approach for designing ethical reasoning abilities?
- *RQ 1:* What are the general applications of neuro-symbolic AI in the different domains such as academic, athletic and business?
- *RQ 1:* What are the influences of this neuro-symbolic AI in the integration of human values with machine intelligence?

1.4 Research Rationale

The rationale of this research lies within the identification of the emerging requirements for AI technologies which assist in making decisions in consistent manner by aligning it with the ethical standards. The present AI system sometimes fails to explain the decisions and resonates with human values which expands the way of trust issues along with significant ethical concerns [2]. This inadequacy in AI capacity poses substantial risks, particularly in the critical domain where the after-effects of unethical decisions lead to potential loss and severe deprivation. In addition to this, machine cognition techniques embrace more complicated computations such as the utilisation of the knowledge of the environment to provide assistance to analogy, reasoning and long-term planning. The incorporation of neuro-symbolic AI assists these fields by offering effective solutions that integrate the strengths of neural networks along with symbolic reasoning.

1.5 Problem Statement

Current system of the artificial intelligence grapples more with the creation of ethical decisions and aligns it with human values due to its data-driven nature. It created a significant requirement for an assurance that those systems, deploying AI technologies in autonomous decision-making, are able to make decisions resonating with human values. The issues posed

by AI are increasingly difficult, especially in the area of data protection and matters of personal and private information, transparency in ethics, and algorithm biases. The ongoing AI approaches within the decision-making processes became inappropriate for the incorporation of the ethical reasoning and explain their decisions in a transparent manner[3]. This research tries to shed light on the issues in the present AI framework and highlights the importance of neuro-symbolic AI to combine human values with machine intelligence.

1.6 Significance of the Research

The capability to create a culture for ethical decision-making makes neuro-symbolic AI more acceptable in the present market landscape [4]. The importance of this research lies in the identification power it holds over critical challenges in the field of AI, with a focus on ethical decision-making and value alignment. Through the development of the neuro-symbolic AI framework, this research focuses on the enhancement of the transparency, trustworthiness and ethical behaviour of autonomous decision-making. This research holds the potential for the enhancement of the capabilities in ethical decision-making through neuro-symbolic AI systems across distinct domains. The findings of this research can pose extensive implications for the various applications including academic, business and athletic fields. This ultimately contributes to the wider range of goals that ensure the benefits of AI technologies for the overall society.

2. LITERATURE REVIEW

2.1 Introduction

The current realm of the advanced artificial intelligence (AI) and machine language offers a huge impact on the human values within various sectors such as sports, academics or business. Hence, neuro-symbolic computing has been an interesting area of research for a long time focusing on bringing together robust learning in neural networks. These published journals include ‘reasoning and explainability’ by offering symbolic representations for neural models. In this project, the literature review segment will go through a few existing articles to accumulate perceptions of other researchers on the subject.

2.2 Neuro Symbolic AI Approach for Making Ethical Reasoning Capabilities

Neuro-symbolic artificial intelligence is recognised as the subfield of artificial intelligence which integrates neural with the symbiotic approaches. The neural refers to the “Artificial Neural Networks” (AMNNs) that are considered as a sub-symbolic approach which includes deep learning leading to significant development in the present decade. The study by Sheth *et al.* (2023) focuses on the exploration of the growing paradigm of neurosymbolic AI which integrated neural networks with symbolic reasoning for creating a more flexible AI system. Humans connect with their environment through perception which highlights the transformation of sensory inputs into systems and cognition which assists in depicting the symbols of the knowledge for planning [5]. Through combining the symbolic approaches with the neural networks, the neuro-symbolic AI can able to improve algorithm-based abilities such as analogy, abstraction and reasoning and application-based capacities such as explainable decision-making. The symbolic reasoning allows the developers to use better auditing of the AI-based decision-making which is essential for the applications in healthcare, sports and autonomous driving. The integration of deep learning and symbolic reasoning for the enhancement of AI applications heavily influences the activities of the construction industry. The deep learning model required a larger labelled dataset which is difficult to obtain within the construction process. The deep learning model lacks transparency which makes it more difficult to understand the decision-making process. In this field, the neuro symbolic computation assists in

combining deep learning technologies with symbolic reasoning leading to the creation of more accurate and holistic AI systems [6]. Symbolic reasoning assists in the enhancement of the interpretability of the AI models which make the decisions more transparent and clarity. Neurosymbolic computing fosters the computer vision model which is used for the expedition of safety and quality within the construction site. The research by Wan *et al.* (2024) delivers a holistic review of the advancements and future directions in neurosymbolic AI. In the present digital settings, the unsustainability in the computational along with the energy requirements of the current AI systems is driven by the larger neural networks. The ongoing AI systems face a shortage of explainability and comprehensiveness which limits its applicability in the critical fields. The “*Neuro Symbolic AI Paradigm*” (NSAI) integrates symbolic reasoning, neural networks and probabilistic methodologies which helps to enhance the interpretability and effectiveness of AI [7]. The NSAI represents the potential in the areas where AI systems are required to collaborate with humans through leveraging cognitive abilities and symbolic reasoning. The integration of symbolic reasoning in the AI system allows the NSAI system to act as an expertise in conducting complicated reasoning tasks. It further helps to make it more compatible with applications requiring higher levels of cognitive functions.

2.3 General Applications of Neuro-Symbolic AI in the Different Domains such as Academic, Athletic and Business

The neuro-symbolic AI assist in making more powerful reasoning along with a streamlined learning system by integrating the strengths of the symbolic reasoning techniques and deep-learning-based methods. The neuro-linguistic programme, which is designed depending on the neuro-symbolic AI assists in the identification of the area, reason and process for the creation of miscommunication in different fields. The neuro-symbolic AI paradigm is commonly used in the healthcare sector, especially for medical diagnosis and drug redesigning. This AI system assists in explaining the procedure of the link prediction in the knowledge graphs [8]. The successful combination of the user behaviour data with symbolic and precise rules helps the NSAI system improve the transparency and accuracy of the delivered recommendations. The neuro-symbolic AI aims to be more comprehensive than neural approaches that assist in acquiring a higher level of explainability which is difficult to maintain. The combination of symbolic reasoning with the neural network seeks to navigate the strengths of both approaches for overcoming individual issues. The study by Bhuyan *et al.* (2024) covers a broader range of applications for the neuro-symbolic which includes the answers to the questions, robotics and computer vision [9]. However, the main challenge in this regard is scaling the neuro-symbolic systems for handling complicated and larger datasets. The neuro-symbolic AI-based framework for the “*Sports Ball Concerning Toddlers*” presents an innovative technique for the identification of sports balls in the area that includes toddlers using a neuro-symbolic AI framework[10]. This framework involves the neural networks for visual recognition with symbolic reasoning for the enhancement of clarity and accuracy. This hybrid approach leverages the strengths of both methods for the improvement of the identification of the overall process. This framework was tested in different scenarios, which shows high accuracy and vigorousness. The successful integration of symbolic reasoning allows the systems to handle inconclusive cases more efficiently than neural-based approaches.

2.4 Significance of Neuro-Symbolic AI with Machine Intelligence on the Integration of Human Values

The integration of ‘Artificial Intelligence (AI)’ with ‘the Internet of ‘Things (IoT)’, known as ‘the Artificial Intelligence of Things (AIoT)’, enhances the processing and analysis capabilities of various devices. It aims to reduce the cost, time, and labour of human beings by linking with the complexity of machine learning (ML). By addressing the challenges and limitations of this process the ‘neuro-symbolic paradigm’ acts as a saviour. Neuro-symbolic models represent knowledge

in a structured manner, allowing for better interpretability [11]. The goal of the neuro symbolic AI is to provide reasoning similar to the human brain along with the capability of learning, representation, and decision-making.

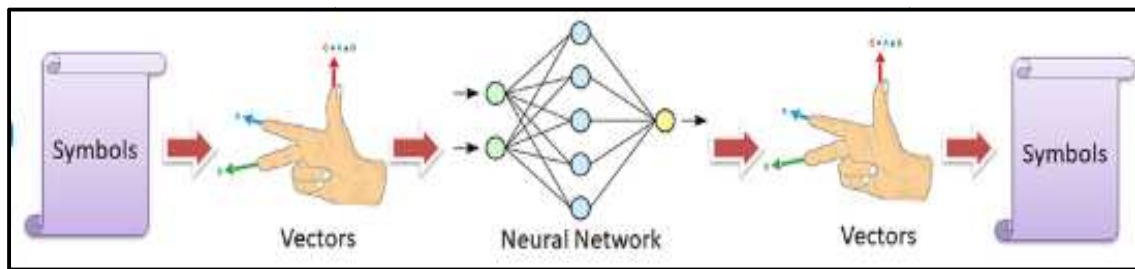


Figure 1: Process of Neuro-Symbolic AI through Vector Representation [11]

Symbolic reasoning allows for the explicit representation of knowledge using rules, logic, and symbols, while connectionist learning such as neural networks excels at handling complex patterns in data. By combining these approaches, neuro-symbolic AI bridges the gap between symbolic and statistical AI, potentially leading to more robust and interpretable models. They learn from data, leveraging neural networks or other machine learning techniques. Symbolic reasoning enables logical deduction and inference. Integration of both learning and reasoning approaches can easily facilitate the decision-making processes [12]. In the case of academic research people can rely on neuro symbolic AI to improve question and answer systems by combining deep learning with logical reasoning. In the case of image recognition tasks, Neurosymbolic models can perform deep learning. It has applications in robotics which can be used in business as well as the sports industry, where reasoning about the environment and making decisions are crucial.

Several articles suggest that AI along with neuro symbolic systems is creating milestones in healthcare services by aiding medical diagnosis and treatment planning with efficiency. It is applied in medical Imaging and AI assists in interpreting X-rays, MRIs, and CT scans, for improving accuracy and speed. AI models are used for predictive disease progression, patient outcomes, and potential complications, especially in cancer treatment. The process of drug discovery becomes smooth by analysing vast datasets and identifying potential candidates. It is helpful for personalized medication as AI tailors treatments based on individual patient characteristics. The neuro-symbolic AI can be used to integrate and interpret huge datasets with patient information including their diagnosis and treatment histories. Regarding safety concerns, the article [13] explains that the overuse of traditional generative AI makes significantly vulnerable to the personal information of individuals.

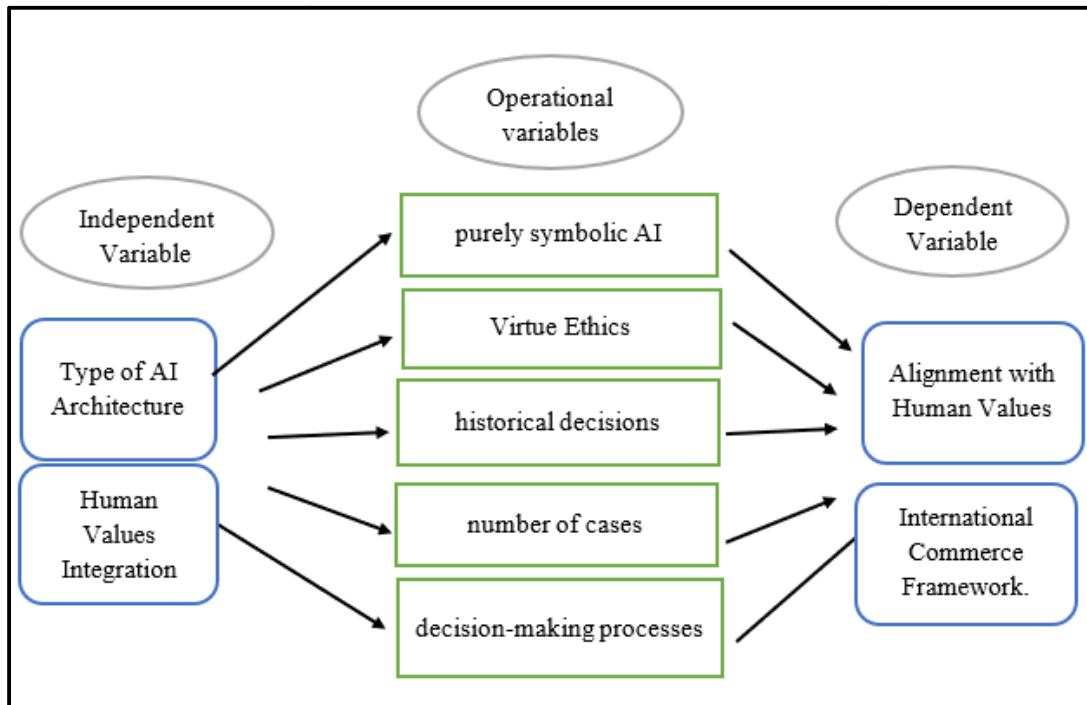
2.5 Theoretical Underpinning

The study about neuro symbolic AI for autonomous decision-making in gap-filling within humans with machine intelligence involves several theoretical perspectives. The theory of Cognitive architecture refers to both a theory about the structure of the human mind and a computational instantiation [14]. This theory is used in artificial intelligence (AI) and computational cognitive science to serve as blueprints for intelligent agents. A cognitive architecture proposes fixed structures that constitute a mind, whether in natural or artificial systems. ACT-R, SOAR, and other architectures contribute to AI programs and advance our understanding of cognition. Dual-process theory involves the idea of both automatic and controlled processes [15]. This model explains persuasion based on intense scrutiny or superficial thinking in social psychology and cognitive psychology by working on memory. The ethical series of artificial intelligence such as deontology, utilitarianism, and virtue ethics compose a foundation of AI used in decision-making techniques.

2.6 Literature Gap

There are a number of research papers on Neurosymbolic Artificial Intelligence for advanced AI applications nevertheless those papers have loopholes to mitigate. For example, while the article [13] provides the foundational principles of neurosymbolic AI are discussed, there is a need for more in-depth exploration of scalability challenges. In the case of integration with emerging technologies, this article could benefit from discussing how neurosymbolic AI can effectively integrate with emerging technologies like quantum computing or advanced robotics. Article [3] lacks information about longitudinal studies to track the development process of neuro-symbolic AI systems. The area of data security solutions lacks emphasis on the article [5] which overlooks the limitations of traditional AI models. The framework of [10] focuses on sports ball identification but lacks a user-centric approach. There is a gap in exploring how the techniques used in this specific framework can be transferred or adapted to other domains or age groups.

2.7 Conceptual Framework



(Source: Self-Developed)

Figure 2: Conceptual Framework.

2.8 Chapter Summary

The integration of Artificial Intelligence (AI) aims to reduce the cost, time, and labour required by linking with the complexity of machine learning (ML). The neuro-symbolic paradigm acts as a saviour by addressing challenges and limitations in AIoT. Neuro-symbolic models represent knowledge in a structured manner, allowing for better interpretability. The goal of neuro-symbolic AI is to provide reasoning similar to the human brain, combining learning, representation, and decision-making. Symbolic reasoning allows for explicit knowledge representation using rules, logic, and symbols. Connectionist learning (such as neural networks) excels at handling complex patterns in data. Cognitive architecture such as ACT-R refers to both a theory about the structure of the human mind and a computational instantiation. This section highlights that the area of data security solutions deserves more emphasis. In summary, neuro-symbolic AI holds immense potential, but addressing these gaps will contribute to its robustness and effectiveness.

3: METHODOLOGY

3.1 Introduction

The methodology section outlines the specific process, techniques, and approaches used in the report to conduct the study properly. It serves as a complete framework for the research philosophy, design, approaches, data collection procedures, and data analysis methods. It explains the process of neurosymbolic AI research, and understanding the interplay between neural networks and symbolic knowledge-based approaches is essential. Research design describes the overall structure of the research and states that whether it is qualitative, quantitative, or mixed-methods. The data collection and analysis part describe the statistical methods, qualitative analysis techniques, or other approaches used to interpret the collected data. Neurosymbolic AI studies must address ethical aspects regarding the study to ensure confidentiality and obscurity and address potential risks. Researchers may discuss several measures in the methodology to ensure the validity and reliability of the study.

3.2 Research Philosophy

A research philosophy includes the assumption, knowledge and nature of the study. In this study, different paradigms of Neurosymbolic AI for implementing those within human values are mentioned. Neurosymbolic AI aims to integrate the strengths of both neural networks and symbolic reasoning. While neural networks excel at large-scale pattern recognition called machine perception, symbolic knowledge-based approaches are better suited for complex computations such as reasoning, analogy, and long-term planning called machine cognition. This involves interpretivism to emphasize the perspectives of researchers on this topic. It will explore the opinions of the developers, users, and administrators about the significance of AI systems with machine learning.

3.3 Research Design

The research design of a study refers to the overall discussion about the strategies to answer the research questions [16]. To collect and analyse the data about the performance of the actions of neurosymbolic AI in the different industries, the secondary qualitative method is used. It will help to provide an overview of the This data will be collected from browsing history, search records, social network activities, and feedback on applications of AI systems from different scholarly articles. From the existing literature, the perception of people regarding the merits and challenges faced by them can be reviewed.

3.4 Research Approach

The research approach implicitly uses the methods of data collection and data analysis by providing the differences between quantitative and qualitative methods. It can be categorized into three types such as deductive, inductive, and abductive approaches. The research regarding the complete framework of neurosymbolic AI for autonomous ethical decision-making techniques for simultaneous management in various sectors. This research includes the deductive approach to investigate the efficacy of the integration of neurosymbolic artificial intelligence over traditional methods for automated ethical decision-making. By using the deductive method this project will provide the opinion of previously working individuals through data collection and analysis.

3.5 Data Collection

For this study about the attribution's neuro-symbolic AI qualitative method is used to capture a comprehensive understanding of user behavior and developers' perceptions. Qualitative data will be collected via a secondary method by gathering large-scale datasets from scientific journals and official websites of applied marketing segments. Additionally, qualitative metrics related to recommendation performance, such as accuracy, precision, recall, and user engagement metrics, will be collected to evaluate the effectiveness of the proposed neuro symbolical approach of AI. Data will be collected through platform databases including previously published articles, ensuring a robust representation of user activities and system performance. The data collection will be performed through peer review of previously published scholarly articles, and journals accessed through Google Scholar and PubMed.

3.6 Data Analysis

The obtained information through secondary qualitative data collection will be accumulated and analyzed based on thematic analysis. depending on the objectives and research questions of the study a few themes will be framed. These themes will be described from the information gathered from review of the existing literatures. Thematic analysis is a qualitative resrch method that helps to identify and interpret patterns or themes of the topic from textual information. During this process it has to make sure that selected and designed themes are accurate and useful.

3.7 Reliability and Validity

The reliability and validity aspects of "Neuro-Symbolic AI for Autonomous Ethical Decision-Making in Bridging Human Values with Machine Intelligence." This paper explores the integration of Artificial Intelligence of Things (AIoT) with neuro-symbolic approaches, aiming to enhance reliability and ethical decision-making. In this context, reliability requires the consistency and stability of the secondary data sources used in the thematic analysis. Along with this to maintain the transparency in gathered and provided data dependability must be fostered on previous researches to derive themes. The validity of this research is concerned about the accuracy of secondary data collection as the credibility of thematic analysis depends on this. The ethical decision making require strength for the support of applications of neuro symbolical AI in various fields such as academic, business and sports.

3.8 Ethical Considerations

Ethical Considerations involve the principles that guide the research design and practices to protect from potential harm and confidentiality. In this study, ethical considerations ensure the protection of participants' rights, privacy and data integrity. It will be obtained through informed consent from all participants involved in the study including the users' data collected from AI platforms and different websites. Clear and transparent information will be provided regarding the purpose of the study, data collection methods, and potential risks and benefits. Privacy and confidentiality will be strictly maintained throughout the research process.

3.9 Chapter Summary

In Chapter 3, methodology outlined comprehensive, covering various aspects such as research philosophy, design, approach, data collection, and analysis methods. By adopting an interpretivist philosophy and a deductive research approach, the study effectively integrates existing knowledge and perspectives related to neuro-symbolic AI. The use of secondary qualitative data collection and thematic analysis allows for a thorough exploration of the interaction between

neural networks and symbolic reasoning. This approach highlights the strengths of both paradigms in ethical decision-making. Ethical considerations are meticulously addressed, safeguarding participants' rights, privacy, and data integrity.

4: FINDINGS/ ANALYSIS/ DISCUSSION

4.1 Findings

As per Wagner, et al. 2024 [17], attaining intuitive and human-like explanations, it is crucial to represent the knowledge level of abstraction, aiding for in-depth exploration as required- similar to how a child repeatedly asks "why" to comprehend something. The statistical explanations and visualizations are great for debugging, and logic-based explanations provide precise, descriptions of the abstract regarding the knowledge base that can be utilized for reasoning and offer the clarity for further activities. The representation of the symbolic aspects is advised as the method to mitigate the underlying between human and artificial intelligence (AI). It symbolizes the scrutinization and the navigation of the activities associated with communication and explainability. Nevertheless, the utilization of the insufficient because the natural level of the languages is frequently informal and ambiguous. The paper advocates for the utilization of formal symbolic languages meant for logical intervention to streamline the communication channel between humans and AI systems.

Considering the neurosymbolic cycle, it has been illuminated that it is a mechanism where the knowledge base is enacted from a deeper network through interactive querying and the continuous inclination of the proactive values that are being attained. The cycle aids the human-in-the-loop approach by offering a shared platform for facilitating the communication channel and integration of the system network. The symbolic networks extracted from the networks at a proper and precise level of abstraction aid the consolidation of the knowledge base, evaluations of the comparative aspects, and targeted revisions of the neural model for networking activities. Illumination is the implementation of the neural network querying at various levels for alignment purposes. Scholarly articles discuss the extent to which neural networks in Logic Tensor Networks (LTN) gain profound information through the assignment of the grounding activities [18]. Within the LTN, the outputs and the internal aspects of the concepts reciprocate are configured over certain logical predicates, allowing elaborations in association with the other predicates. A potential user querying the neural networks or symbolic knowledge, directly intervening with the abstract of the presentation and their operations. querying with bounded variables, one can analyze how certain features influence outcomes. For instance, is querying whether variables under specific risk conditions produce similar outcomes across different groups, ensuring fairness. This is formally expressed as a logical query that mandates consistent and non-discriminatory outcomes across groups with similar risk profiles. The findings emphasize the significance of comprehending the undermined facets and mechanisms in the AI mechanism. A method designed for querying pertains to the prevalence of conceptual relationships, such as the conjunctions of the logical literals, which is critical for the inclination of the understanding of the human aspects of the AI system configurations. The approach presents in the research paper proved to be relevant, with an approximation of the procedures for attaining the proper inclination between humans and AI.

Scholarly, articles refer to grounding as the process of association with the abstract symbols and the conceptual aspects within the algorithms in the AI domain to real-world datasets and expertise. Contrary to the traditional symbols grounding, which frequently depends upon linguistic symbolizations, that approach emphasizes the grounding technique within the Interactive AI mechanisms. These elaborations reciprocate the symbolic aspects of acting as the bridge gap between the deep learning parameters and the reasoning of the humans that comprehend the understanding [19]. The

explanations illustrated within certain models deliver the process within the AI systems to communicate its decision-making approaches or reasoning alignment with human user bases. These elaborations are crucial for the assurance of the actions demonstrated by AI with human values and are understandable to both the scenarios showcasing experts and non-experts.

Lastly, the semantic representation of the spaces is considered as the multi-dimensional periphery where the concepts associated with the relational patterns are demonstrated. This space is configured upon the norms of the symbolic logic and the patterns pertained by the neural networks that understand both aspects. The findings proactively illustrate the requirement for the clear definitions of the terminologies to effectively incline the AI systems with the understanding rate of humans. It assures that the AI decisions are convenient, interpretable, and relevant in terms of real-world scenarios.

4.2 Analysis

Theme 1: Analysis of the Neuro symbolic AI Approaches for Making Ethical Reasoning Capabilities

Research over the neuro symbolic AI has been profound and determined the interests recently, that drives by deep integration of the advancement in learning approaches. This reliability has been evaluated in a concentrated aligned symbolic complexity, that involves the deduction of the logical aspects and the learning of the rules, utilization of the techniques associated with the deep learning phenomenon. These modulated efforts have correctly aided the researchers and demonstrated their capabilities and the limitations regarding deep learning in maintaining the complicated nature of the tasks assigned symbolically. Another prominent area of the research is the integration of knowledge bases that align symbolic measures and the expressive ideas allocated in metadata into deep learning channels. This proactive approach utilizes well-configured knowledge for the background of the knowledge bases such as the graphs required for the integration of knowledge [20]. It illuminates the performance of the system operators, specifically in certain tasks that undermine the contextual shreds of evidence, such as zero-shot learning and conversational agents. This integration improves accuracy, bolsters generalizability, and ensures safety in neural control systems. The critical determination regarding the neuro-symbolic approaches underscores the explainability of AI. By incorporating the background knowledge demonstrated by humans, the researchers planned to offer more optimistic and transparent explanations of AI behaviors. The prevalent neuro-symbolic AI research is focused on capitalizing the deep learning approaches to enforce the problems associated with symbolic approaches, that streamline the AI within the well-configured knowledge bases and enhance the explainability [21]. The multidisciplinary prospects aid in generating the more robust, that are further generalized and enhance the transparency within the AI system framework, addressing some of the inherent limitations of the traditional deep learning practices.

Theme 2: Evaluating the Potential Implications of the Neuro Symbolic AI in Diversified Domains

- **Healthcare:** Neuro-symbolic AI revolutionizes healthcare by streamlining diagnostic accuracy and the planning for treatment. Through the integration of symbolic reasoning with deep learning approaches, the incorporation of AI systems provides the proper interpretation of the complicated dates associated with medical datasets [22]. This proactive approach paved the way for the openness and reliability of AI tools, enhancing the outcomes of the patient and the trustworthiness of AI-generated solutions for healthcare.

- **Finance:** In the finance sector, neuro symbolic approach and the AI-driven approaches can enhance the assessment of the risks and the fraudulent detection. The combination of the patterns associated with the deep learning approaches the identification of the logical reasoning of the allowance and the analysis of the betterment of the complicated financial datasets and the regulatory norms [23]. The approaches which comprised several phenomena in the areas of decision-making and the provide strategies for investment purposes, ultimately lead to extensive explainable aspects of the financial goals.

Within the education domain, the neuro-symbolic AI can further personalize the experiences associated with the learning approaches through the comprehension of the requirements provided by the students and the analysis of the data sets and certain symbolic reasoning. From the research paper, it is evident that AI can offer well-configured content for education and aids in the adaptation of the diversified learning phenomenon that offers explanations for its profound recommendations.

Theme 3: Understanding the Influence of Neuro-Symbolic AI in the Integration of Human Values with Machine Intelligence

Articles narrate that AI performs a significant role in amalgamating the values of humans within the intelligence machine learning approaches by mitigating the underlying gap between the raw evidence and the human understandable reasoning approaches. Contrary to previous models practiced within AI, which frequently operate as the black box, neuro-symbolic AI amalgamates with the deep learning potential that aids in identifying the symbolic logical overview and the allowance of system integration within the decision-making procedures undertaking the values of humans more precisely. The implementation of the AI mechanisms aids in offering the elaboration of the activities in a way that delivers the reasoning concisely of the humans [24]. It crafts the activities more transparent and accountable. In crafting the ethical decision-making approaches and understanding the scenarios, neuro-symbolic AI can utilize logical stances to assure the consequences regarding the alignment with the norms of the society and the value models, providing a rational judgment that humans can comprehend and ensure trustworthiness. The ability to incorporate symbolic knowledge, such as ethical rules and cultural norms, directly into the AI's reasoning process ensures that the system can adapt to different contexts and respect diverse human values. This makes neuro-symbolic AI particularly valuable in sensitive applications like healthcare, law, and social governance, where adherence to ethical principles is crucial. Neuro-symbolic AI aids a more rapid and continuous amalgamation of the integration of the values into intelligence provided by the machines encouraging the trustworthiness and assuring systems integration of AI models that are consistent with the ethical approaches of the humans and societal expectations.

4.3 Discussion

Certain methods were utilized to capitalize on the information from the Explainable AI (XAI) to preserve the AI models from capitalizing on the advancement of unprecedented biases. The integration of the knowledge bases and the extraction of the ideas from the deep neural networks into the segment of the Logic Tensor Network (LTN) platform and implementing the well-crafted fairness of the aspects provide several constraints during the training programs. The approaches proactively incorporate the fairness of the facts into the AI system configurations that involve biases utilizing XAI tools such as the SHAP. The mechanism's potential innovative measures underlie the application of the transparent constraints to perform the programs associated with training approaches. The primary approaches regarding the innovative measures underlie the constraints associated with the process, specifically addressing the complexities such as gender

biases. For instance, fairness is considered by the assurance of the assessment of the risk factors for the male (RMi) and female (RFi) segmentation in the risk assessment credits models that are equitable. The framework is trained to treat both genders equitably by the logical justification of the contrasting characters that maintain the equivalence of the groups [25]. The process associated with fairness is entirely operated with an autonomous mechanism, assuring that both the datasheets are preserved within a class segment, that involves a diversified range of the gender, groups that are properly treated upon the decision-making process for the model. The scrutinization operated upon the real-world information for income prediction, the risk with the credit assessment, and the recidivism demonstrate that these neuro-symbolic aspects attained openness while managing the top-tier classification of the applications [26]. The discussion section meticulously highlights the potential of the methods to implement fairness into the AI models in a flexible agnostic manner.

5: CONCLUSION

5.1 Conclusion

From the above context, it can be concluded that neuro-symbolic AI depicts a prominent advancement toward the inclination of AI with the values demonstrated by humans and ethical conduct. It further entails the implications of AI in several fields of study with meticulous justifications regarding its usability and applicability. The amalgamation of neurosymbolic prospects with AI underscores both positivity and limitations. The implication of human values and thinking in machine-based learning mechanisms aids in streamlining daily tasks. The research undermines the overall effectiveness of the neuro-symbolic procedures in offering intuitive, explanations of human nature and streamlining the interpretability of the AI configurations. The ambiance towards knowledge-based neural networking approaches mitigates the complexities between the reasoning adopted and its practical incorporation making AI more interpretable and accountable. The findings illuminate the transitional potential of the neuro-symbolic approaches and AI implementation over diversified domains. It enhances the diagnostic and the patient care over trustworthiness. Additionally, integrating human values into the AI systems ensures that decisions lie with the societal norms and the ethical practices. Neuro symbiotic AI offers a promising framework for developing AI systems that are technically advanced and ethically sound aligning with human values.

5.2 Linking with Objective

The analysis of the neuro-symbolic AI emphasizes its capability to streamline ethical reasoning through the integration of the symbolic aspects of logic with the deep learning aspects. The proactive approaches ensure that AI-generated systems can further provide openness, making it understandable for explorations and alignment with the values of humans. Through testing and refining the neuro-symbolic AI systems within practical applications such as the academic field, athletes, health, sports, businesses and various others allow future researchers to conduct in-depth research in this segment. The delivery of encouragement for creating collaboration between the researchers, sociologists, ethicists and psychologists. This further provides assistance to making interdisciplinary collaboration that ensures the ethical integration of AI systems.

5.3 Future Scope

The future scope of this research lies in the enhancement of interpretability and dynamic learning with adaption, cross-cultural frameworks, real-world applications and streamlined decision-making.

- The development of methods for making the decision-making process for the neuro-symbolic AI system more interpretable and transparent. This further assists in understanding the procedure of ethical decisions by ensuring that they resonate with human values.
- Further research can focus on the construction of systems that assist in dynamic learning and the adaption of emerging ethical guidelines coupled with societal norms. These regulations include continuous learning from human feedback and altered ethical standards.
- The integration of contextual awareness into the neuro-symbolic systems can expand the way for future scope. This ensures that the decisions are made by considering the specific circumstances and variations in the ethical standards.
- The development of frameworks which can able to handle ethical decision-making across distinct cultures. This required understanding and integration of diverse human values into the AI systems.
- The application of neuro-symbolic AI in practical scenarios such as vehicles, healthcare, academics, sports, athletes, law enforcement and businesses have to be focused on. The testing and refinement of these systems within the practical application can deliver valuable insights into the applicability of neuro neuro-symbolic AI framework.

5.4 Recommendation

Analysing the overall research findings, the recommendation can be provided in the following manner:

- **Interdisciplinary collaboration:** It is essential to encourage the interconnectivity between AI researchers, sociologists and ethicists so that a well-organised and ethical framework for the AI system can be designed [27]. This approach assists in the identification of complicated ethical issues.
- **User-centric design:** Organisations and businesses have to focus on designing a user-centric AI system that caters to the diverse range of preferences, requirements and values of the users [28]. Through the inclusion of end-users in the improvement procedure can ensure that the AI system resonates with the ethical expectations of the users.
- **Vigorous ethical guidelines:** The establishment of precise ethical guidelines for the development as well as deployment of the neuro-symbolic AI systems. These guidelines have to be regularly updated to showcase the technological advancements and altered societal values.
- **Ethical auditing:** It is recommended to incorporate the techniques for continuous ethical auditing and monitoring of the AI techniques. It will further assist in the identification and recognition of the ethical issues that arise during the operation and application of this AI system [29].

Through focusing on these crucial areas, researchers can stimulate the field of neuro-symbolic AI. It further ensures that the autonomous systems assist significantly in making ethical decisions that resonate with human values.

FUNDING DECLARATION

This research was conducted without collecting any specific grant from any commercial, public or other funding agencies. The authors received no financial support for the research, authorship, and/or publication of this article. All the resources utilized for the development and conduction of the research were self-financed by the authors or their affiliated institutions. Any incidental charges were covered through institutional along with personal support. The authors declare no financial conflicts of interest related to this research.

REFERENCES

1. A. E. Williams, "A model for artificial general intelligence," in *Proc. Int. Conf. Artif. Gen. Intell.*, Cham, Switzerland, Jul. 2020, pp. 357-369. Available: https://link.springer.com/chapter/10.1007/978-3-030-52152-3_38.
2. J. B. Telkamp and M. H. Anderson, "The implications of diverse human moral foundations for assessing the ethicality of artificial intelligence," *J. Bus. Ethics*, vol. 178, no. 4, pp. 961-976, 2022. Available: <https://link.springer.com/article/10.1007/s10551-022-05057-6>.
3. W. Rodgers, J. M. Murray, A. Stefanidis, W. Y. Degbey, and S. Y. Tarba, "An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes," *Hum. Resour. Manag. Rev.*, vol. 33, no. 1, p. 100925, 2023. Available: <https://www.sciencedirect.com/science/article/pii/S1053482222000432>.
4. E. Barnes and J. Hutson, "Natural Language Processing and Neurosymbolic AI: The Role of Neural Networks with Knowledge-Guided Symbolic Approaches," *J. Artif. Intell. Robot.*, vol. 2, no. 1, 2024. Available: <https://digitalcommons.lindenwood.edu/faculty-research-papers/603/>.
5. A. Sheth, K. Roy, and M. Gaur, "Neurosymbolic artificial intelligence (why, what, and how)," *IEEE Intell. Syst.*, vol. 38, no. 3, pp. 56-62, 2023. Available: <https://ieeexplore.ieee.org/abstract/document/10148662/>.
6. X. Luo, H. Li, and S. Lee, "Bridging the gap: Neuro-Symbolic Computing for advanced AI applications in construction," *Front. Eng. Manag.*, vol. 10, no. 4, pp. 727-735, 2023. Available: <https://link.springer.com/article/10.1007/s42524-023-0266-0>.
7. Z. Wan, C. K. Liu, H. Yang, C. Li, H. You, Y. Fu, C. Wan, T. Krishna, Y. Lin, and A. Raychowdhury, "Towards cognitive AI systems: a survey and prospective on neuro-symbolic AI," *arXiv preprint arXiv:2401.01040*, 2024. Available: <https://arxiv.org/abs/2401.01040>.
8. D. Bouneffouf and C. C. Aggarwal, "Survey on applications of neurosymbolic artificial intelligence," *arXiv preprint arXiv:2209.12618*, 2022. Available: https://www.researchgate.net/profile/Djallel-Bouneffouf/publication/364382002_Toward_The_Applications_of_Neurosymbolic_Artificial_Intelligence/links/635061aa8d4484154a18d98f/Toward-The-Applications-of-Neurosymbolic-Artificial-Intelligence.pdf.
9. B. P. Bhuyan, A. Ramdane-Cherif, R. Tomar, and T. P. Singh, "Neuro-symbolic artificial intelligence: a survey," *Neural Comput. Appl.*, pp. 1-36, 2024. Available: <https://link.springer.com/article/10.1007/s00521-024-09960-z>.
10. A. Mangal, T. Vyas, R. Gupta, S. Tanwar, and H. Shahinzadeh, "Neurosymbolic AI-based Framework For Sports Ball Identification Concerning Toddlers," in *Proc. 2024 20th CSI Int. Symp. Artif. Intell. Signal Process. (AISP)*, Feb. 2024, pp. 1-6. Available: <https://ieeexplore.ieee.org/abstract/document/10475201/>.

11. B. P. Bhuyan, A. Ramdane-Cherif, R. Tomar, and T. P. Singh, "Neuro-symbolic artificial intelligence: a survey," *Neural Computing and Applications*, pp. 1-36, 2024. Available: <https://link.springer.com/article/10.1007/s00521-024-09960-z>.
12. Y. R. Shrestha, V. Krishna, and G. von Krogh, "Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges," *Journal of Business Research*, vol. 123, pp. 588-603, 2021. Available: <https://www.sciencedirect.com/science/article/pii/S0148296320306512>.
13. A. Sheth, V. Pallagani, and K. Roy, "Neurosymbolic AI for enhancing instructability in generative AI," *arXiv preprint arXiv:2407.18722*, 2024. Available: <https://arxiv.org/abs/2407.18722>.
14. F. Grassiotto and P. D. P. Costa, "CogToM: A Cognitive Architecture Implementation of the Theory of Mind," in *ICAART* (2), 2021, pp. 546-553. Available: <https://www.scitepress.org/PublishedPapers/2021/101942/101942.pdf>.
15. R. Augusto, "Two kinds of process or two kinds of processing? Disambiguating dual-process theories," *Review of Philosophy and Psychology*, vol. 15, no. 1, pp. 277-298, 2024. Available: <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.805386/full>.
16. J. L. Johnson, D. Adkins, and S. Chauvin, "A review of the quality indicators of rigor in qualitative research," *American Journal of Pharmaceutical Education*, vol. 84, no. 1, p. 7120, 2020. Available: <https://www.sciencedirect.com/science/article/pii/S0002945923016066>.
17. B. Wagner and A. D. A. Garcez, "A Neurosymbolic Approach to AI Alignment," *Neurosymbolic AI*, 2024. Available: <https://neurosymbolic-ai-journal.com/system/files/nai-paper-729.pdf>.
18. M. Gong, H. Yang, S. Zhang, and J. Zhao, "Challenges and opportunities of climate change for Chinese agriculture," *Natl. Sci. Rev.*, vol. 9, no. 6, p. nwac035, 2022. Available: <https://academic.oup.com/nsr/article/9/6/nwac035/6542460>.
19. I. Abdullah, A. Javed, K. M. Malik, and G. Malik, "DeepInfusion: A dynamic infusion based-neuro-symbolic AI model for segmentation of intracranial aneurysms," *Neurocomputing*, vol. 551, p. 126510, 2023. Available: <https://www.sciencedirect.com/science/article/pii/S0925231223006331>.
20. A. D. A. Garcez, M. Gori, L. C. Lamb, L. Serafini, M. Spranger, and S. N. Tran, "Neural-symbolic computing: An effective methodology for principled integration of machine learning and reasoning," *arXiv preprint arXiv:1905.06088*, 2019. Available: <https://arxiv.org/abs/1905.06088>.
21. R. Agrawal and N. Pandey, "Neuro-symbolic AI: Revolutionizing advanced reasoning in AI assistants," *Int. J. Res. Comput. Appl. Inf. Technol. (IJRCAIT)*, vol. 7, no. 1, pp. 30-40, 2024. Available: https://ijrcat.com/index.php/home/article/view/IJRCAIT_07_01_004.
22. J. M. Ptaschunder, "Artificial intelligence in the healthcare sector," *Scientia Moralitas - Int. J. Multidiscip. Res.*, vol. 4, no. 2, pp. 1-14, 2019. Available: <https://www.ceeol.com/search/article-detail?id=820666>.
23. C. Maple et al., "The AI revolution: Opportunities and challenges for the finance sector," *arXiv preprint arXiv:2308.16538*, 2023. Available: <https://www.ceeol.com/search/article-detail?id=828906>.

24. A. Sheth and K. Roy, "Neurosymbolic value-inspired artificial intelligence (Why, what, and how)," *IEEE Intell. Syst.*, vol. 39, no. 1, pp. 5-11, 2024. Available: <https://ieeexplore.ieee.org/abstract/document/10453391/>.
25. A. Oltramari, J. Francis, C. Henson, K. Ma, and R. Wickramarachchi, "Neuro-symbolic architectures for context understanding," in *Knowledge Graphs for eXplainable Artificial Intelligence: Foundations, Applications and Challenges*, IOS Press, 2020, pp. 143-160. Available: <https://ebooks.iospress.nl/doi/10.3233/SSW200016>.
26. R. Agrawal and N. Pandey, "Neuro-symbolic AI: Revolutionizing advanced reasoning in AI assistants," *Int. J. Res. Comput. Appl. Inf. Technol. (IJRCAIT)*, vol. 7, no. 1, pp. 30-40, 2024. Available: https://ijrcait.com/index.php/home/article/view/IJRCAIT_07_01_004.
27. R. Moirano, M. A. Sánchez, and L. Štěpánek, "Creative interdisciplinary collaboration: A systematic literature review," *Think. Skills Creat.*, vol. 35, p. 100626, 2020. Available: <https://www.sciencedirect.com/science/article/pii/S1871187119302196>.
28. L. Bu, C. H. Chen, K. K. Ng, P. Zheng, G. Dong, and H. Liu, "A user-centric design approach for smart product-service systems using virtual reality: A case study," *J. Clean. Prod.*, vol. 280, p. 124413, 2021. Available: <https://www.sciencedirect.com/science/article/pii/S0959652620344589>.
29. I. Munoko, H. L. Brown-Liburd, and M. Vasarhelyi, "The ethical implications of using artificial intelligence in auditing," *J. Bus. Ethics*, vol. 167, no. 2, pp. 209-234, 2020. Available: <https://link.springer.com/article/10.1007/s10551-019-04407-1>.
30. C. S. Stafie, I. G. Sufaru, C. M. Ghiciuc, I. I. Stafie, E. C. Sufaru, S. M. Solomon, and M. Hancianu, "Exploring the intersection of artificial intelligence and clinical healthcare: a multidisciplinary review," *Diagnostics*, vol. 13, no. 12, p. 1995, 2023. Available: <https://www.mdpi.com/2075-4418/13/12/1995>.