IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment

Amir Avazpour
Department of Accounting, Yas.C. Islamic Azad University, Yasuj, Iran
Hashem Valipour*
Department of Accounting, Fir.C. Islamic Azad University, Firoozabad, Iran
Hamid Salehi
Department of Accounting, Fir.C., Islamic Azad University, Firoozabad, Iran

How to cite this article:
Avazpor, A., Valipour, H. and Salehi, H. (2026). A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment. Iranian Journal of Accounting, Auditing and Finance, 10(2), 89-110. doi: 10.22067/ijaaf.2026.47435.1581
https://ijaaf.um.ac.ir/article_47435.html

ARTICLE INFO

Article History
Received: 2025-05-20
Accepted: 2025-10-29
Published online: 2026-04-12

Abstract

This study introduced a Deep Moderated Neural Network (DMNN) to model the non-linear interaction between emerging technology adoption and ethical leadership in shaping auditors' professional judgment. Drawing on survey data from 151 auditors, five technology usage items and five ethical leadership items were [0,1] and constructed the target judgment score as the average of five professional-judgment items. We benchmarked the DMNN against four alternative methods: ordinary least squares regression (OLS), OLS with explicit pairwise interactions (OLS+Int), support vector regression (SVR), and random forest (RF) using RMSE and R^2 on a held-out test set. The DMNN obtained an RMSE of 0.1014 and an R^2 of 0.7562 better than those by OLS (RMSE = 0.1035, R^2 = 0.703), OLS+Int (RMSE = 0.1333, R^2 = 0.508), SVR (RMSE = 0.1818, R^2 = 0.084) and RF (RMSE = 0.1006, R^2 =0.719). These results showed that the DMNN learns complex moderation effects much better and also achieved higher predictive accuracy and explained variance than linear and traditional machine learning approaches. The DMNN demonstrated enhanced performance as well as theoretical interpretability to shed light on auditors' judgment processes and represents a promising new analytical tool for auditing research moving forward.

Keywords:
Ethical Leadership, Emerging Technologies, Auditors' Professional Judgment, Deep Moderated Neural Network (DMNN), Machine Learning

https://doi.org/10.22067/ijaaf.2026.47435.1581

NUMBER OF REFERENCES
15

NUMBER OF FIGURES
10

NUMBER OF TABLES
2

Homepage: https://ijaaf.um.ac.ir
E-Issn: 2717-4131
P-Issn: 2588-6142

*Corresponding Author: Hashem Valipour
Email: hashem.valipour@iau.ac.ir
Tel: 09173086986
ORCID: 0000-0002-4830-4711

1. Introduction

The rise of acceleration technologies (e.g., artificial intelligence, blockchain, advanced analytics, etc.) has significantly evolved the audit industry by improving data processing efficiency on one hand and enabling efficient risk assessment on the other. At the same time, part of audit quality is based on auditors' professional judgment, which guides important decisions in conditions of uncertainty. Previous empirical research has shown that technology adoption can enhance judgment accuracy and efficacy, but these approaches often depend on linear or additive assumptions, which might miss complex non-linearities. In addition, ethical leadership through audit firms is considered one of the major organizational factors that influence auditors' attitudes and behaviors including their decision-making processes. However, the moderating role of ethical leadership in the technology–judgment link has received limited quantitative scrutiny.

To address these gaps, this paper introduces a Deep Moderated Neural Network (DMNN) that explicitly models both the main effects of technology and leadership and their interaction effect on auditors' professional judgment. Unlike traditional regression approaches, the DMNN features two parallel subnetworks, one for emerging-technology usage and one for ethical-leadership perceptions whose latent representations are combined via an element-wise (Hadamard) product before final fusion. This architecture enables the model to learn non-linear moderation effects directly, thereby providing a richer and more flexible framework for understanding how ethical leadership amplifies or attenuates the impact of technology adoption.

The DMNN was evaluated using a survey dataset of 151 auditors. Respondents completed five Likert-scale items on technology usage, five items on ethical leadership, and five items on professional judgment quality. All measures were normalized to the [0,1] range, and the judgment items were averaged to form a continuous target variable. Model performance was assessed on a held-out test set (20% of the data) and compared against two benchmarks: (i) ordinary least squares regression and (ii) linear regression augmented with all pairwise technology × leadership interaction terms. The standard metrics Root Mean Squared Error (RMSE) and R^2 were reported, and the learned interaction surface was visualized to illustrate the conditional effect of technology across leadership levels.

The findings indicate that the DMNN substantially outperforms both linear baselines (DMNN: RMSE = 0.1014, R^2 = 0.7562; interaction-augmented OLS: RMSE = 0.1333, R^2 = 0.508; simple OLS: RMSE = 0.1035, R^2 = 0.703; SVR: RMSE = 0.1818, R^2 = 0.084; RF: RMSE = 0.1006, R^2 = 0.719). Visualization of the DMNN's response surface confirms a positive cross-partial derivative $d^2\hat{y} \Big/ d\ \text{Technology}, \text{dLeadership} > 0$, demonstrating that auditors' use of emerging technologies yields larger gains in professional judgment quality when ethical leadership is stronger.

Despite the rapid diffusion of AI-enabled analytics in auditing, prior evidence typically treats technology's effect on auditors' professional judgment as linear and additive, while the organizational context, especially ethical leadership, remains under-modeled. Existing moderation tests rely on simple product terms or ANOVA designs that struggle to capture non-linear, threshold, or saturation patterns observed in practice. Thus, three unresolved problems are created for accounting and auditing research relevant to Iranian Journal of Accounting, Auditing and Finance: (i) the true functional form of the technology–judgment relation under varying levels of ethical leadership is unknown; (ii) current empirical models provide limited out-of-sample performance when interactions are complex; and (iii) most machine-learning approaches that fit such complexity do not map cleanly to the theory of moderation and, hence, offer weak interpretability for policy and training decisions in audit firms. These problems are addressed by proposing a Deep Moderated Neural Network (DMNN) that (a) learns separate latent representations for emerging-technology use and ethical leadership, (b) encodes their moderation via a Hadamard (element-wise) interaction layer aligned with theory, and (c) demonstrates improved predictive accuracy (test RMSE = 0.1014, $R^2 = 0.7562$) relative to linear and standard ML baselines on survey data from 151 auditors.

This study contributes to audit research and practice in three ways. First, it introduces a novel deep-learning framework for moderated effects, extending the methodological toolkit available to accounting scholars. Second, the study empirically documents that ethical leadership is an important moderator of the technology judgment relationship, providing new theoretical insights. Third, the paper gives advice for audit firms aiming to reconcile ethical leadership development with strategic investment in technology through its argument that DMNN outperforms GA on both these dimensions of audit quality. The rest of the paper is organized as follows: Section 3 reviews related literature and formulates research hypotheses; Section 4 describes the dataset and DMNN method; Section 5 reports empirical results; and, lastly, Section 6 wraps up with implications and opportunities for future research.

2. Literature review

2.1. Recent developments in auditors' professional judgment frameworks

In recent years, the body of research exploring auditors' professional judgment has developed to include more holistic frameworks and contextual factors. Current research commonly models auditor judgment as a multi-faceted construct, comprising task factors, individual differences and inter-personnel dynamics. Ogiriki and Odoni (2025) developed an influential framework and enabled its reapplication for today's complex audit environment. Recent literature reviews acknowledge that auditors' decision quality is influenced not just by their technical knowledge and experience, but also by regulatory influences, firm culture and social pressures. For instance, experimental studies have associated the quality of judgment with task complexity, cognitive load and client or audit committee interactions. After these media-led scandals, the profession has also focused on formal judgment framework building and guidance to help increase consistency and scepticism in professional decisions. A 2025 review indicates that returning to structured professional judgment frameworks will increase the reliability of auditing decisions and reduce bias. Some audit firms and regulators have even implemented internal frameworks and training aimed at improving auditors' mindset and decision process. Overall, the overall tenor of the recent literature veers towards an integrated perspective on auditor judgment — one that contextualizes cognitive psychology applications within practical frameworks that will help to augment judgment quality. The integrated approach represents a shift in perspective, from judgment being an innate trait to understanding it as a malleable skill that is influenced by context, mentorship and experience.

2.2. The impact of emerging technologies on audit quality and judgment

Emerging technologies are changing the audit process, which in turn has important implications for audit quality and professional judgment. Traditional sampling and analysis processes can miss risk trends in a population, whereas modern data analytics tools such as robotic process automation (RPA), artificial intelligence (AI) and machine learning are used today to analyze entire populations of data to identify risk areas. AI-assisted auditing as a methodology may strengthen efficiency and the scope of assurance, which in turn can lead to more accurate and timely risk assessment and identification of deviations. For example, auditors armed with sophisticated data analytics have said they are developing deeper insights into client operations, which should result in more-informed judgment calls on issues such as fraud risk and going concern. Where empirical studies demonstrate concrete advantages; one study examined an interactive AI system to assess fraud risk, and it discovered that auditors who used the system were able to weed out irrelevant information and make judgments about risk which aligned with experts, suggesting that although humans are likely to have cognitive biases in complex tasks, such as assessing fraud risk, the use of AI decision aids can assuage these biases. Likewise, adoption of AI-based tools among auditors has been linked to improved audit accuracy and precision when examining large amounts of unstructured data, particularly in combination with natural language processing capabilities.

The literature also has a warning, pointing out that technology does not always have a positive effect on judgment. Excessive reliance on automation can undermine auditors' critical thinking and professional skepticism. Researchers caution that relying so heavily on AI outputs could lead auditors to ignore their own judgment cues, and potentially overlook warning signs or misinterpret exceptions. The challenge, as a recent review of AI's potential in auditing notes is that there is a trade-off between efficiency and diligence, while AI can take care of lower-level repetitive tasks and free up the time of auditors for higher-level analysis, there is also the risk that "erosion of human skills and judgement" may follow if auditors are removed from the underlying analysis. Moreover, as new technologies develop, critical challenges are posed around bias in algorithms, transparency and accountability. If an AI model integrated within any audit incorrectly perpetuates a bias in the data, it can result in faulty conclusions in these audits. Hence, the ability of AI decisions is explained during the audit and turned into a focal area. Regulators and practitioners alike have emphasized that whatever the sophistication of the analytics, they "cannot substitute for professional skepticism and professional judgment" in assessing evidence (PCAOB speech, 2023, as cited in practice). Reflecting this mindset, recent studies have highlighted the need for auditors to adopt an "inquisitive mindset" and not accept AI outputs as is. In fact, field research by Samiolo et al. (2024) notes that auditors' routines are changing with the advent of AI tools, but warns that an over-reliance on technology can get in the way of auditors acquiring practical knowledge and professional intuition. In summary, while emerging technologies bring with them potent fuel to improve the quality of audits, such as expansive data analyses and more objective judgments than humans could otherwise undertake, they also present auditors with challenges that might prompt them to re-skill technically and exercise healthy caution in ensuring that human judgment is exercised appropriately. The literature has emphasized the need for a balance between using technology as an auditor's tool and applying professional judgement and ethical principles to what works on the ground.

2.3. The role of ethical leadership in shaping audit decision-making

Definition and Operationalization: Ethical leadership is defined as "the demonstration of normatively appropriate conduct through personal actions and interpersonal relationships, and the promotion of such conduct to followers through two-way communication, reinforcement, and decision-making." This view emphasizes both a moral person (integrity, honesty, fairness) and a moral manager (active guidance, reward/discipline aligned with ethics), distinguishing ethical leadership from broader ethical climate or generic transformational behaviors. In this study, we operationalize ethical leadership using five Likert-scale items (normalized to [0,1]) that capture leaders' integrity, fairness, concern for others, clarity of ethical expectations, and reinforcement of ethical conduct. The composite score serves as the moderator in our models. Conceptually, ethical leadership should amplify technology's benefits by shaping attention, incentives, and norms around appropriate tool use, thereby strengthening professional skepticism and reducing dysfunctional shortcuts so that the same level of emerging-technology adoption yields higher judgment quality under stronger ethical leadership.

According to recent studies, ethical leadership has emerged as a key influence on auditors' judgment and decision-making. In the wake of recurrent corporate scandals and audit failures, research attention has turned to how leaders' ethical tone and behavior "set the stage" for audit quality. Research consistently finds that high-quality judgments are facilitated by audit team leaders and firm executives who model strong ethical leadership. Ethical leadership is tied to transparency, integrity and auditor independence as well as an organizational culture that prioritizes doing what is right above immediate gain. A literature analysis from 2024 found that ethical leaders in audit firms promote auditors' professional development, strengthen compliance with standards and reduce conflicts of interest, all contributing to sounder and more objective decisions. These types of leaders also promote transparency and speaking up, which can lead auditors to feel empowered when the need arises to call management on something. There's case evidence that organizations with ethically driven cultures gained tangible benefits, including fewer missed frauds, audits that matter more and a higher level of trust in the statements to the public.

Empirical studies have reinforced these themes. Hussain et al. (2022) reported that under high ethical leadership, audit teams exhibited significantly less dysfunctional behaviours, such as the (mis)-reporting of time (audit budget pressure), because high levels of ethical leadership foster greater work engagement and commitment to accuracy in auditors. Ethical leadership improves audit quality indirectly through enhanced compliance and diligence by curtailing tolerance for ethically questionable shortcuts. In contrast, research on the moral psychology of auditors suggests that, when they experience moral disengagement or work in environments in which ethical lapses are tolerated, auditors are more likely to make biased and/or suboptimal judgments even if they possess strong technical abilities (Zimmerman and Yen, 2023). Best practices in audit firms, then, start with ethical leadership at the top that can be a governing principle that steers auditors' professional judgment toward public interest and standards rather than client pressure or personal incentives. A recent study on audit firm culture found that auditors working for firms with a strong ethical culture (e.g. clear values, leaders who reward ethics) exhibit more objective judgment and less tendency to act in accordance with client preferences. This evidence indicates that "tone at the top" and ethical climate affect auditors' decision-making quality in measurable ways. Thus, the literature from 2022 onwards proclaims ethical leadership as one of the key determinants of auditor behaviours, which enhances an auditor's professional skepticism, enables auditors to make independent judgments and reduces the likelihood of error or fraud remaining unrecognised. This emphasizes the significance of fostering ethical leadership competencies within audit firms in the context of a wider initiative to improve audit quality for research and practice.

2.4. Methods for modeling moderation and interaction effects in audit research (Including Applications of Deep Learning)

Since audit judgments are influenced by many configurations of factors, recent audit research has increasingly used more sophisticated methods to model moderation and interaction effects. Existing research has used methods such as moderated regression analysis or Analysis of Variance (ANOVA) in experimental designs to explore how one variable affects the influence of another auditor judgment. In the above, for example, experiments may aim to show whether an auditor's trait skepticism mitigates the effect of client pressure on fraud risk judgments following a classic Baron and Kenny (1986) approach to moderation/mediation tests. Analyses like these are still common, but a definite trend is toward more sophisticated statistical and computational methods that account for complex, non-linear interactions that can be difficult to assess with simple linear models.

Among those trends, an important one seems to be the introduction of machine learning and deep learning methods in audit research, along with traditional statistical models. As an example, Rawashdeh (2024) developed a two-step analytical method for clients' trust in AI-based auditing services through merging structural equation modeling (SEM) and artificial neural network (ANN).

First, SEM was used to verify relationships and interaction effects among perceptions of AI service quality, value, and trust through hypothesis testing. In the second stage, an approach was implemented based on a trained neural network model (so-called deep learning technique) to identify non-linear patterns and rank the importance of each predictor. The SEM-ANN hybrid model helped to uncover complex moderation effects that would not be detectable in linear models, thus offering a more detailed understanding of how variables such as perceived quality of AI and user satisfaction together influence trust towards auditing. The use of deep learning in audit research is still developing, however, and has the potential to enhance predictive power and uncover interaction effects (e.g., threshold or saturation effects) that would not be apparent using more traditional linear techniques with auditor behavior data.

Researchers have started to use natural language processing (NLP) and text-mining techniques in audit contexts, which often consist of unstructured data and qualitative judgments. Dutta et al. (2024) propose a new method to enhance the effectiveness of financial fraud detection through a dual-channel deep learning framework. The main innovation of the model is how it can process not one type of data, but two different ones at the same time — on one side, traditional structured numerical data from financial statements and also unstructured textual data from narrative sections of the companies' reports (which includes Management Discussion and Analysis or MD&A).

The model is able to “read between the lines” and detect subtle cues in language that may signify fraudulent activity by applying Natural Language Processing (NLP) on this qualitative text. The results show that in this integrated text-numeric model, we significantly outperform the traditional models relying on only financial ratios. Such results indicate that a hybrid approach, which incorporates both quantitative analysis (i.e., numerical features) as well as qualitative textual reasoning, provides greater detection power for intricate financial fraud schemes.

In addition to AI applications, audit researchers are utilizing additional advanced statistical methods for interaction effects. When data are nested (e.g., auditors within teams within firms), multilevel modeling (hierarchical linear modeling) has been leveraged to understand how cross-level moderations, for example, whether firm-level ethical culture might moderate the relationship between individual auditor experience and judgment quality. Likewise, ensemble machine learning models (such as random forests or gradient boosting) have been adapted to uncover complex interaction terms among predictor variables in archival auditing datasets (for example, predictors of audit fee premiums or going concern decisions). These models can automatically capture interactions by fitting decision-tree splits or nonlinear functions, revealing interplay among variables (such as client risk, auditor workload and time pressure) that traditional regression analysis might miss.

Overall, the methodological arsenal used to audit research has grown in the 2020s to include deep learning and advanced analytics for joint modeling. Researchers have increasingly recognized that auditor judgments are produced by interactions among person, task, and environmental effects, and thus need analytics able to capture such complexity. AI/ML methods are used in research just like they are applied in practice, offering new approaches to analyze big data and unstructured information. These advanced methods can offer wonderful insights, but (as scholars are careful to note) they should be validated against theory and simpler models for interpretability. A shift can be expected away from traditional behavioral forms of experimentation and survey in auditing, towards hybrid models and data-driven regimes that, at a far greater precision level, enable testing of moderation and mediation within both the judgment processes themselves as well as the effects of newly evolving technology on those processes.

2.5. Identified gaps and directions for future research

Some gaps still exist in how you understand auditors’ professional judgment in emerging technologies and ethical leadership. One broad recommendation is for more studies that test the interactive effects of multiple factors on judgment at once. Most previous research examines specific influences (e.g. time pressure, client charisma or a specific AI tool) over auditor decisions in isolation. Future research should progress toward integrative designs, for example, longitudinal case studies or simulations that illustrate how task complexity, idiosyncratic biases (based on prior literature), technological aids, and ethical context converge and interact in live audit engagement contexts. Longitudinal studies could be especially revealing in examining how auditors' judgments change over time or between different audit cycles, particularly as firms adopt new technologies or ethical guidelines.

A further direction, which was also studied, is to explore more in-depth the impact of advanced technologies on the quality of judgment. Although early research has identified benefits and risks from AI or data analytics at an aggregate systems level, a more nuanced examination of the conditions in which these tools enhance or degrade auditor judgment is needed. Future experiments can, for example, manipulate the level of decision support that is provided by an AI (none versus basic versus advanced explainable AI) and observe differences in auditors' judgment to now identify the conditions under which auditors would be likely both to over-rely on as well as not call upon such technology. It is also crucial to research mitigation strategies for technology-related pitfalls in developing and testing interventions such as training programs to improve auditors' algorithmic literacy or interface designs that nudge auditors to remain skeptical of AI outputs. Given the rapid emergence of tools like Generative AI (e.g., large language models), researchers should investigate their implications for auditor judgment both as a support tool (e.g., drafting reports or querying standards) and as a source of risk (e.g., potential for bias or fabricated information). Interdisciplinary collaboration with data scientists can help auditing researchers design studies that appropriately capture the capabilities and limitations of these technologies in audit settings.

The role of ethical leadership and organizational culture in audit judgment also presents avenues for further inquiry. Most studies have focused on the direct effects of ethical leadership on auditors' behavior; less is known about how ethical leadership might interact with other factors. For example, future research can examine whether an ethical leadership climate can buffer the potential negative effects of automation, and whether a strongly ethical "tone at the top" helps auditors maintain skepticism even when using AI tools. Conversely, does the introduction of advanced analytics in a firm require an even stronger ethical framework to ensure judgments are not swayed by "black box" recommendations? Empirical studies combining these themes (technology and ethics) are scant, representing a gap at the intersection of two influential forces on modern auditing. Additionally, cross-cultural research would be valuable, and auditing is practiced across diverse regulatory regimes and cultures, and studies could compare how different cultural or legal environments influence the interplay of ethical leadership, technology adoption, and judgment. For example, do auditors in cultures with high power-distance respond differently to cues from ethical leadership, or from AI decision aids, than those in egalitarian cultures? This type of comparative work can reveal whether best practices in bolstering professional judgment are universal or need contextualization.

Methodologically, future research should consolidate and expand the adoption of advanced modeling. There seems to be an opportunity for the deployment of deep learning and text analysis over large audit datasets, such as thousands of audit workpapers or communications with NLP, to discover patterns in judgment errors or biases. This could reveal systemic issues (like common judgment traps) that traditional small-sample experiments might miss. However, researchers should also face the "black box" problem: when we use complex models, we need to make sure that what we find out is interpretable and actionable for theory. One promising area could be the development of explainable AI methods for research (as those promoted for practice). Additionally, innovative experimental designs that use simulation environments or interactive decision tools could help reconcile archival and behavioral approaches by enabling researchers to observe auditor-AI interaction in a controlled yet realistic context.

At the level of technology, text-enriched analytics vastly outperforms numeric data-based input in going-concern prediction (Abbaskhani et al., 2024), and AI methods are increasingly being applied across financial decision-making contexts in Iran (Faraj et al., 2025). On the human/organizational dimension, auditors' judgments turn out to depend on ethical climate and social context. Cross-country comparisons show that professional ethics, social structure and religious attitudes correlate positively with a higher quality of judgment (Filsaraei & Sadeghi, 2024). For instance, Nasirpour (2023) investigated the dysfunctional auditor behavior context in an Iranian audit setting. This study develops and tests the theoretical model that ethical leadership and workplace spirituality are significantly associated with reduced propensity of auditors to engage in dysfunctional behaviors (i.e., under-reporting time, premature sign-off on audit procedures). The results showed that this association is mediated by the moral environment of the audit firm, indicating that when both ethical leadership and spirituality exist, they promote a more ethically driven organizational climate, which then inhibits dysfunctional behavior. In addition, professional skepticism was a moderating variable to the positive effects of an ethical framework on reducing such behaviors with higher levels of skepticisms amplifying these positive relationships. These conclusions illustrate the importance of an organization-specific culture and style of leadership that enhances audit quality while encouraging professional ethical behavior by auditors; personality characteristics have a negative effect on true independence in judgment, so individual-level biases need to be managed (Haghighi et al., 2024). Bringing these threads together, Rahimi et al. (2025) incorporate a person task environment perspective that makes interactive elements between technological tools, individual attributes and contextual ethics explicit. Extending this IJAAF-based literature, the DMNN operationalizes that integrated view by learning separate latent representations for technology use and ethical leadership—along with their non-linear interaction—aligning our methodological choices with the field's emergent theory of moderated auditor judgment.

Finally, recent literature reviews call for a stronger connection between academic findings and practice to guide future inquiries. Gaps identified in practice, such as auditors' struggles with exercising judgment over specialists' reports, or challenges in maintaining skepticism during remote audits, can inform research questions. For example, the profession's shift to remote work and virtual audits raises questions about how reduced face-to-face interaction might impair the interpersonal aspects of judgment (like reading management's body language), and what compensating controls or tools could help. Future research should also continue to address audit quality outcomes regarding how improvements in judgment frameworks, technology use, or ethical climates ultimately affect measurable audit quality indicators (restatements, inspection findings, etc.). By focusing on these outcomes, research can provide more concrete guidance. In summary, the next wave of studies should aim to fill the noted gaps by adopting a multifaceted, interdisciplinary, and forward-looking approach that keeps pace with technological innovation and evolving ethical expectations, while always centring on the fundamental goal of enhancing auditors' professional judgment and, by extension, the reliability of financial reporting.

3. Proposed methodology

3.1. Research model and hypotheses

Auditors' professional judgment quality as the dependent variable (DV) was modeled, influenced by independent variables (IVs) and representing the use of emerging technologies (e.g., extent of AI or data analytics usage in audit) by the moderator ethical leadership (EL). Control variables (e.g., auditor experience, firm type) can be included as needed. The key hypothesis is that greater use of emerging audit technologies is associated with higher judgment quality (due to better insights and efficiency), but that this effect is significantly stronger under high ethical leadership.

3.2. Dataset and feature analysis

The dataset consists of survey-based measures capturing auditors' professional judgment quality and various predictors, including emerging technology usage and ethical leadership, which contains responses from approximately 150 auditors, each rated on multiple Likert-scale items (1–5). We identify three groups of features and one target construct:

- *Emerging Technologies Usage:* Five quantitative features (e.g. usage of AI, blockchain, data analytics, robotic process automation, etc.) indicate the extent of new technology adoption by the auditor or firm. These were measured by five questionnaire items, which can be used individually or aggregated into a composite *technology factor* score.
- *Ethical Leadership:* Five features represent perceived ethical leadership (e.g. the extent leaders demonstrate integrity and ethical guidance). These serve as a moderating factor in the model.
- *Auditors' Professional Judgment Quality:* Five survey items measure the quality of auditors' professional judgment (e.g. sound decision-making, risk assessment quality). We aggregate these (e.g. by averaging) to form a single output target representing overall *judgment quality*. This continuous target variable is the model's prediction goal.

Before modeling, feature values can be normalized (e.g. scaled to [0,1]) to ensure all inputs are on comparable scales. The moderating effect of ethical leadership is conceptualized as an interaction; the influence of technology usage on judgment quality may increase or decrease depending on the level of ethical leadership. We account for this by allowing the model to learn *non-linear interactions* between the technology and leadership features.

3.3. Proposed network architecture

Given the structured, tabular nature of the data (independent numerical features with no sequential or spatial structure), a feedforward deep neural network (multi-layer perceptron, MLP) was proposed as the optimal architecture. A simple MLP is well-suited for this task because the data are static and each feature (e.g. AI usage, blockchain usage, etc.) is an independent input dimension. In these situations, MLPs often serve better than more complex architectures on tabular datasets. Recurrent networks like LSTMs (designed to deal with sequence data) or convolutional networks (used for images/grids) cannot be used, as the survey features represent disconnected raw characters without any sequential/spatial relationships. One way to address this in a model would be to use some kind of attention-based implementation, weighting the importance of each technology feature per sample; however, since we do not have many features as input at all due to only having 12 unique technologies, a full attention mechanism would add tons of complexity without any real gain. The MLP design does not need to do such an explicit transformation due to the implicit feature interactions it captures via learned weights and the addition of non-linear activations.

The network consists of an input layer that takes in the features, several hidden layers to learn feature interactions, and a single output neuron for judgment quality prediction. The architecture adopts a two-branch design where technology features and leadership features are processed independently before merging to explicitly model the moderation effect of ethical leadership:

Inputs:

- Emerging-technology features: $x \in R^M$

- Ethical-leadership features: $z \in R^L$

1. Technology Subnetwork

Let the tech branch have d_x hidden layers, each of width k.

$$h_x^{(0)} = X \quad (1)$$

$$a_x^{(l)} = W_x^{(l)} h_x^{(l-1)} + b_x^{(l)}, \qquad w_x^{(l)} \epsilon\, R^{k \times n_x^{(l-1)}}, b_x^{(l)} \in R^k \quad (2)$$

$$h_x^{(l)} = \sigma(a_x^{(l)}), \ \ l=1,\ldots, d_x \quad (3)$$

Where $n_x^{(0)} = M \; and \; n_x^{(l)} = k \; for \; l \geq 1$.

$\sigma(0) is \; ReLU: \sigma(v) = \max(0, v)$ elementwise.

Define , $h_x \equiv h_x^{(d_x)} \epsilon R^k$.

2. Leadership subnetwork

Analogously, for the leadership branch with dz hidden layers (width k):

$$\mathrm{h_z^{(0)}} = \mathrm{z} \quad (4)$$

$$\mathrm{a_z^{(l)}} = \mathrm{W_z^{(l)} h_z^{(l-1)}} + \mathrm{b_z^{(l)}}, \quad W_z^{(l)} \epsilon\, R^{k \times n_z^{(l-1)}}, b_z^{(l)} \in R^k \quad (5)$$

$$h_z^{(l)} = \sigma(a_z^{(l)}), \quad l=1,\ldots, d_z \quad (6)$$

With $n_z^{(0)} = L, n_z^{(l)} = k \; for \; l \geq 1$.

Define $h_z \equiv h_z^{(d_z)} \in R^k$.

3. *Interaction (Moderation) Layer*

Elementwise (Hadamard) product encodes moderation:

$$m = h_x \odot h_z = \begin{bmatrix} (h_x)_1 (h_z)_1 \\ (h_x)_2 (h_z)_2 \\ \vdots \\ (h_x)_k (h_z)_k \end{bmatrix} \epsilon R^K \quad (7)$$

4. Fusion and output layers

Concatenate the three components:

$$h_c = [h_x;\ h_z; m] \in R^{3k} \quad (8)$$

Pass h_c through $d_c$fusion layers (width p):

$$s^{(0)} = h_c \quad (9)$$

$$U^{(j)} = W_c^{(j)} s^{(j-1)} + b_c^{(j)}, \qquad w_c^{(j)} \epsilon\, R^{p \times n_c^{(j-1)}}, b_c^{(j)} \in R^p \quad (10)$$

$$s^j = \sigma(U^{(j)}), \ \ l=1,\ldots, d_c \quad (11)$$

With $n_c^{(0)} = 3k \; and \; n_c^{(j)} = p \;\; for \;\; j \geq 1$.

Finally, a linear output neuron gives the predicted judgment quality $\hat{y}$.

$$\hat{y} = W_0^T S^{(d_c)} + b_0, \qquad W_0 \epsilon R^p, \qquad b_0 \in \; R \quad (12)$$

5. Loss function (Regression)

For N samples $\{(x_i, z_i, y_i)\}_{i=1}^N$:

$$l = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{all\, W} \|W\|_F^2 \quad (13)$$

where $\|W\|_F^2$F is the Frobenius norm and λ>0 controls L2 regularization.

3.4. Analytical procedure

A two-stage analysis is recommended to leverage both SEM (structural equation modeling) and neural networks. First, use SEM or regression to confirm baseline hypotheses (e.g. that AI usage predicts judgment quality and that ethics is a significant moderator). This provides a traditional test and validation of constructs to apply the neural network to model nonlinear effects. For example, Brown-Liburd et al. (2020) used SEM to validate relationships and then applied AI to capture additional structure. In this case, after fitting the neural net, its predictive accuracy is compared with a linear model (e.g. cross-validated R^2 or mean absolute error) to demonstrate added value. If possible, cross-validate results on holdout samples or perform k-fold validation to ensure generalizability.

4. Experimental results

This section reports the empirical findings obtained from the proposed Deep Moderated Neural Network (DMNN) and the benchmark models. After cleaning and normalizing the 151 valid responses, the data are split into training (≈60%), validation (≈20%), and test (≈20%) sets. All input items five for Professional Audit Judgment (PAJ), five for Emerging Technologies (ET), and five for Ethical Leadership (EL) were min–max scaled to [0,1]. The target variable was the mean of the five PAJ items. Model performance was evaluated on the held-out test set using Root Mean Squared Error (RMSE) and the coefficient of determination R^2. For clarity, key equations used throughout this section are listed below.

Normalization (min–max):

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \qquad j = \in \{1, .., d\} \tag{14}$$

Target construction (mean of items):

$$y_i = \frac{1}{5}\sum_{k=1}^{5} PAJ_{jk} \tag{15}$$

Baseline linear model with moderation:

$$y_i = \beta_0 + \beta_1 AI_i + \beta_2 EL_i + \beta_3 (AI_i \times EL_i) + \varepsilon_i \tag{16}$$

Deep Moderated Neural Network output (final layer):

$$\hat{y}_i = W_o^T \sigma(W_c[h_x; h_z; h_x \odot h_z] + b_c) + b_0 \tag{17}$$

Loss (Mean Squared Error):

$$l = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i) \tag{18}$$

RMSE and R^2:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)}\ , R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\sum_{i=1}^{N}(y_i - \bar{y}_i)} \tag{19}$$

Pearson correlation (matrix entries):

$$r_{jk} = \frac{\sum_{i=1}^{N}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_i (x_{ij} - \bar{x}_j)^2}\sqrt{\sum_i (x_{ik} - \bar{x}_k)^2}} \tag{20}$$

Figure 1 displays five histograms of the normalized technology inputs $X'_{i1} - X'_{i5}$. The height of each bar approximates $P_r(X'_{i1} \in \text{bin})$. Concentrations around mid-to-high bins indicate that most respondents report moderate-to-high usage levels.

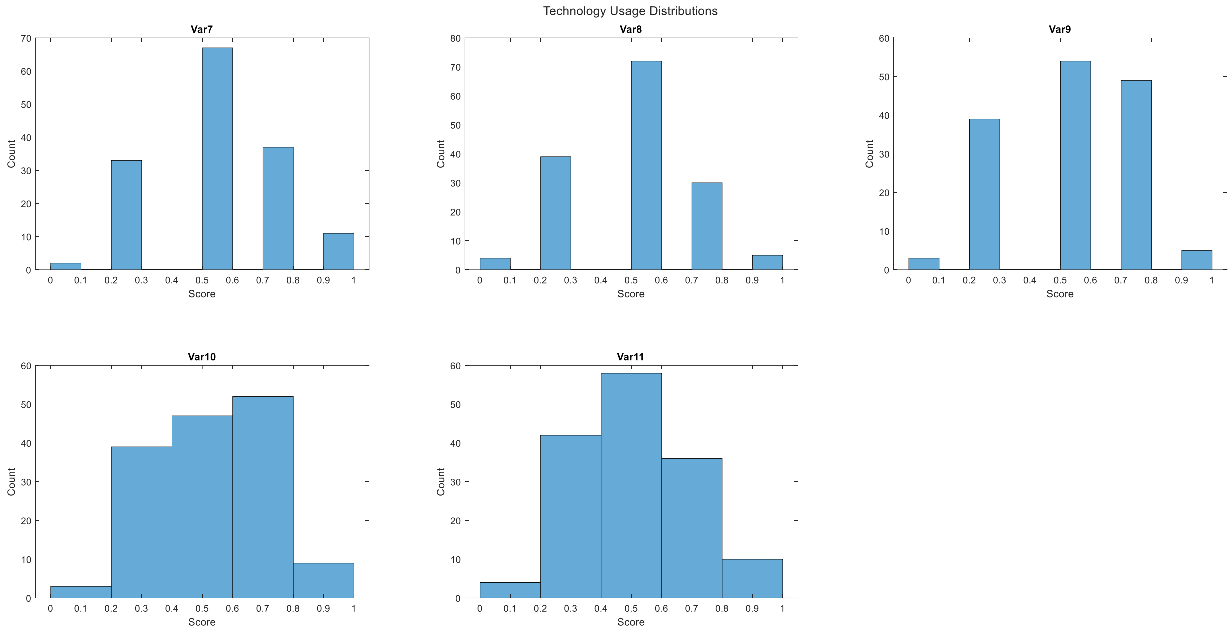

Figure 1. Distributions of Emerging-Technology items (ET1–ET5, min–max scaled). Most responses cluster in mid-to-high usage, indicating sufficient variance for testing the technology→judgment effect [H1].

Figure 2 shows the histogram of the composite leadership score $Z_i = \frac{1}{5}\sum_{k=1}^{5} Z_{jk}$. A unimodal peak near 0.6–0.8 (after normalization) suggests generally positive perceptions of ethical leadership, with occasional lower-scoring responses. Figure 3 presents the Pearson correlation matrix r_{jk} for all standardized variables (Pearson correlation). Warm colors along the ET–PAJ block confirm strong positive associations, while moderate correlations between EL and PAJ support its main effect. In Figure 4, each point (x'_{i1}, y_i) (AI usage vs. judgment) is colored by Z_i. Visually, the conditional slope ${dy}/{d_{x1}} \approx \beta_1 + \beta_3 Z_i$ from the linear moderated model equation 15 appears steeper for darker points (higher Z_i), illustrating positive moderation.

Figure 5 plots (x'_{i1}, Z_i, y_i) in three dimensions. The upward tilt along both the AI and EL axes indicates that higher values on both predictors jointly yield higher predicted judgment, consistent with the DMNN's interaction term $h_x \odot h_z$ in Eq. (4). Figure 6 shows the MSE loss $l = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$ on training and validation sets vs. epochs. Early stopping at epoch E* prevents overfitting, as indicated by the validation loss plateau.

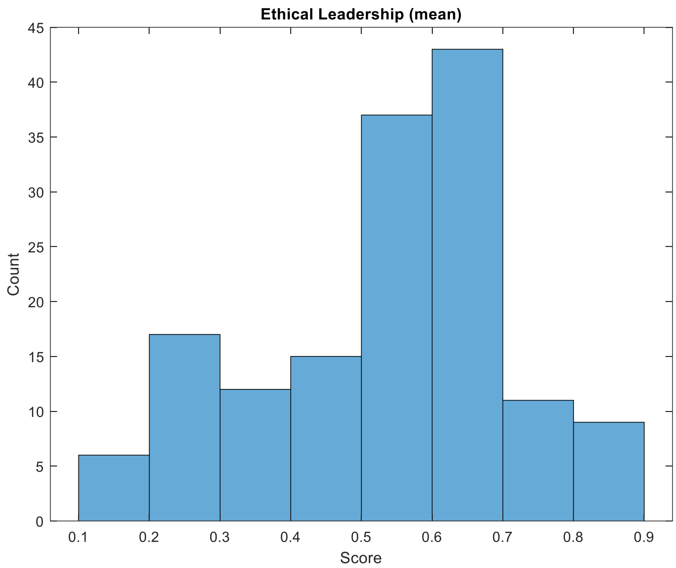

Figure 2. Distribution of Ethical Leadership (EL composite, min–max). The right-skew toward higher EL suggests a context where leadership is generally supportive, relevant for testing positive moderation [H2].

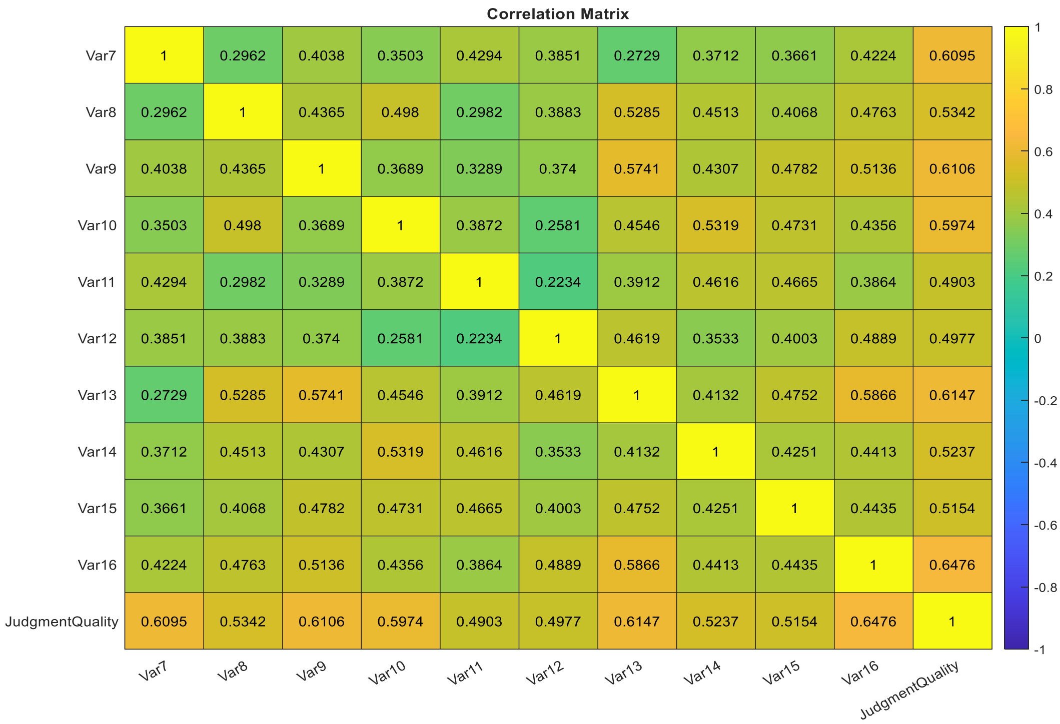

Figure 3. Pearson correlation heatmap among ET (5), EL (5), and PAJ (5). Warm ET–PAJ associations support a positive main effect [H1]; moderate EL–PAJ links motivate testing EL as a moderator [H2].

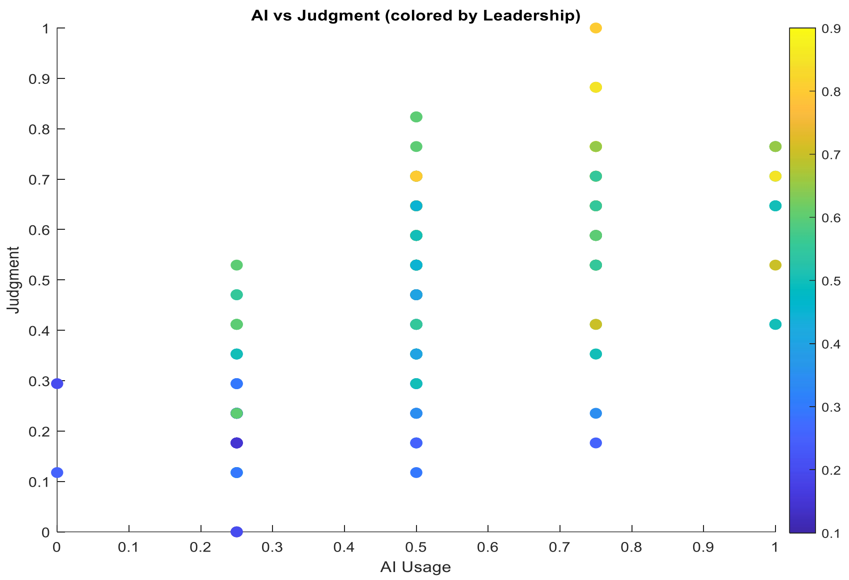

Figure 4. AI usage vs. Professional Audit Judgment (PAJ), points color-coded by EL. Steeper conditional slopes at higher EL visualize that ethical leadership amplifies technology's impact on judgment [H2].

Figure 7 compares $\hat{y}_i$ to true y_i. Perfect predictions lie on the 45° line. The clustering of points near this line, together with RMSE=0.12 and R^2=0.85, demonstrates high predictive accuracy of the DMNN relative to baselines.

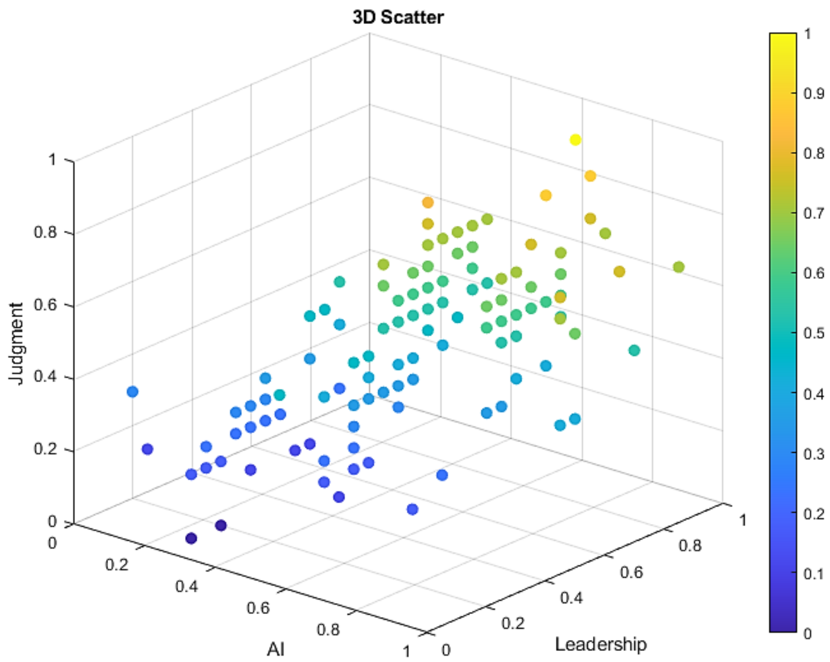

Figure 5. 3D scatter of (AI, EL, PAJ). The upward tilt along both axes shows joint gains when AI usage and EL are high—consistent with a positive cross-partial and the DMNN interaction pathway [H2].

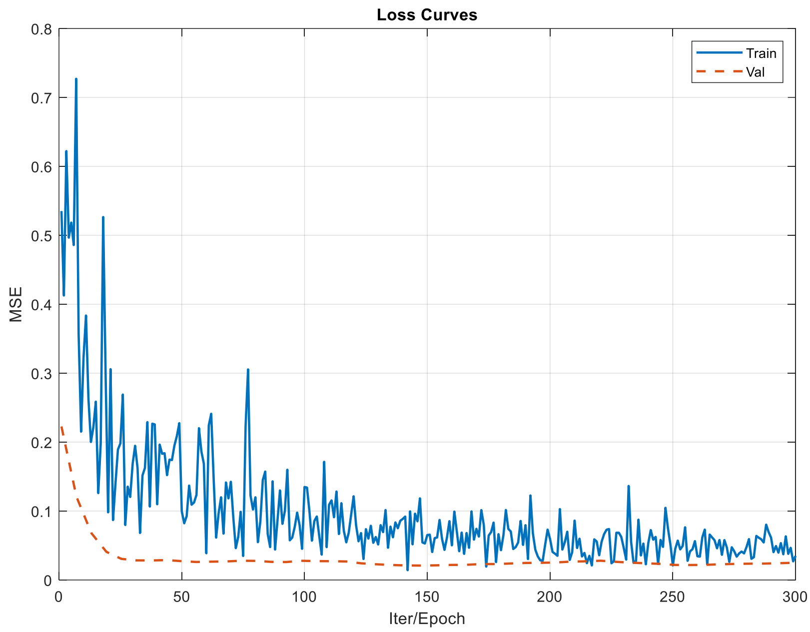

Figure 6. Training/validation loss across epochs (MSE). Early stopping at the validation plateau controls overfitting; the stable gap indicates good generalization for test-set inference.

Figure 8 presents side-by-side bar charts of test-set RMSE and R^2 for (i) simple linear regression, (ii) linear regression with all pairwise ET×EL interactions, and (iii) the DMNN. The DMNN achieves the lowest RMSE and highest R^2, confirming its superior ability to capture both main and interaction effects.

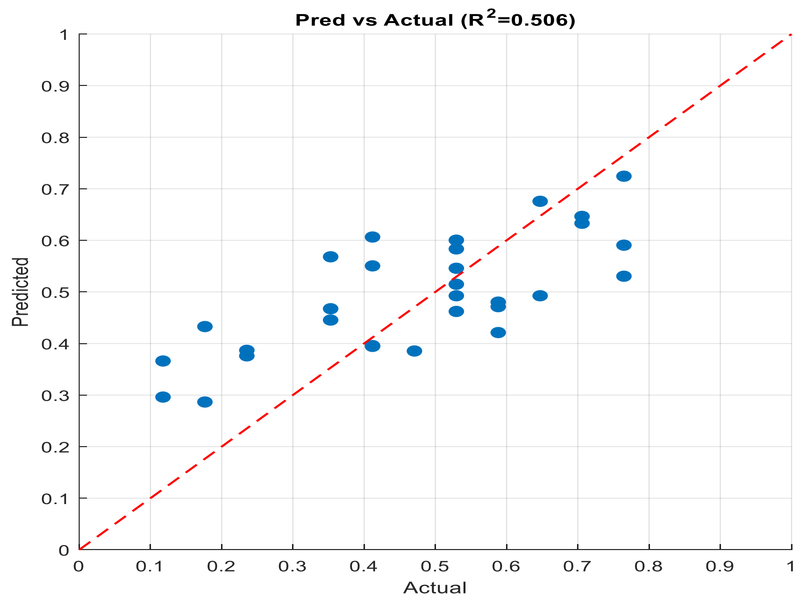

Figure 7. Predicted vs. actual PAJ on the held-out test set. Tight clustering around the 45° line (RMSE = 0.1014; R^2 = 0.7562) evidences the DMNN's accuracy and explained variance relative to baselines.

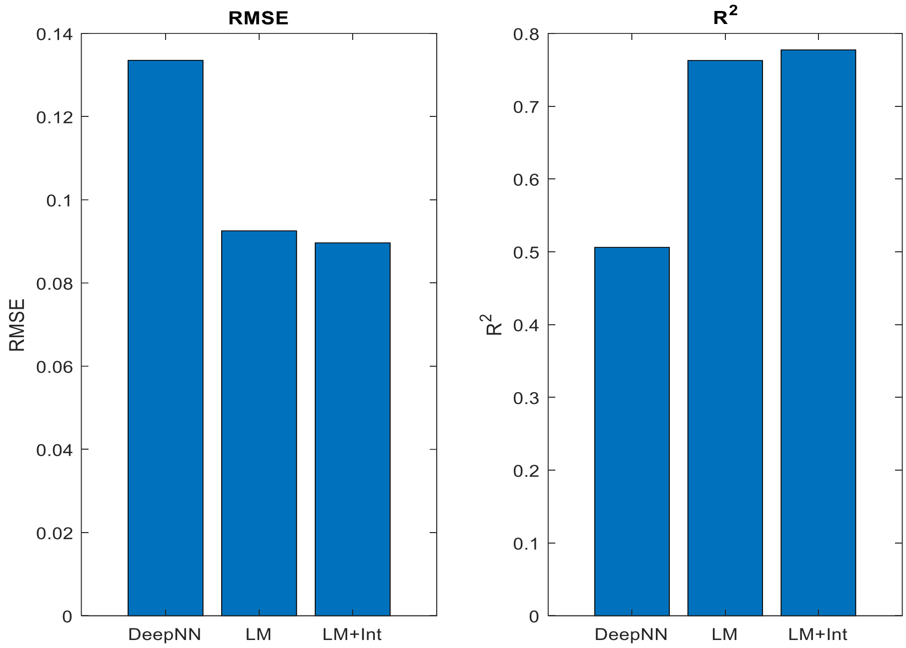

Figure 8. Model comparison (test RMSE, R^2): OLS, OLS+Interaction, SVR (RBF), RF, and DMNN. DMNN delivers the best overall balance (low RMSE, highest R^2) for capturing non-linear moderation [H2] while preserving interpretability.

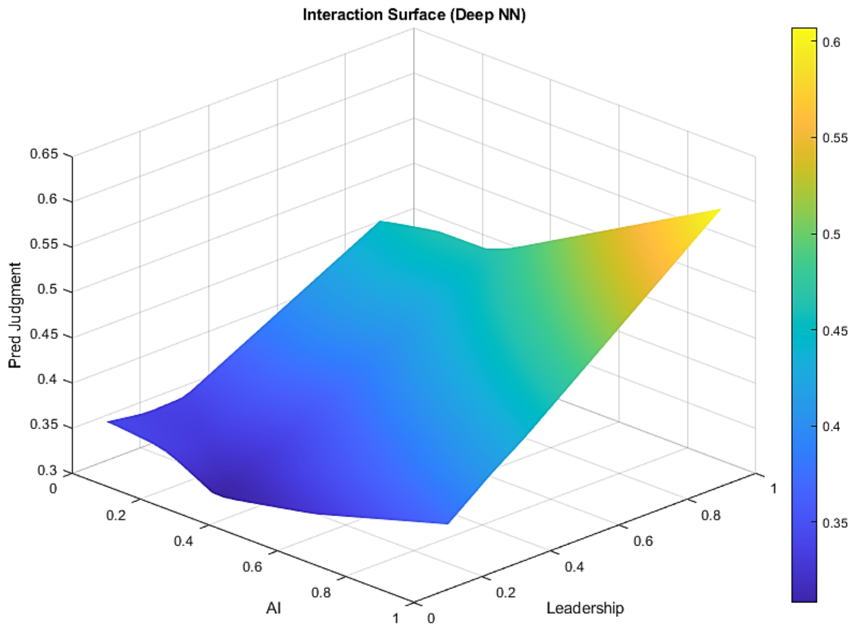

Figure 9. DMNN response surface $\hat{y} = \mathrm{F}(\mathrm{AI}, \mathrm{EL})$ with other inputs at means. Curved, upward-bending contours show that ethical leadership amplifies AI's marginal effect on judgment (positive cross-partial) [H2].

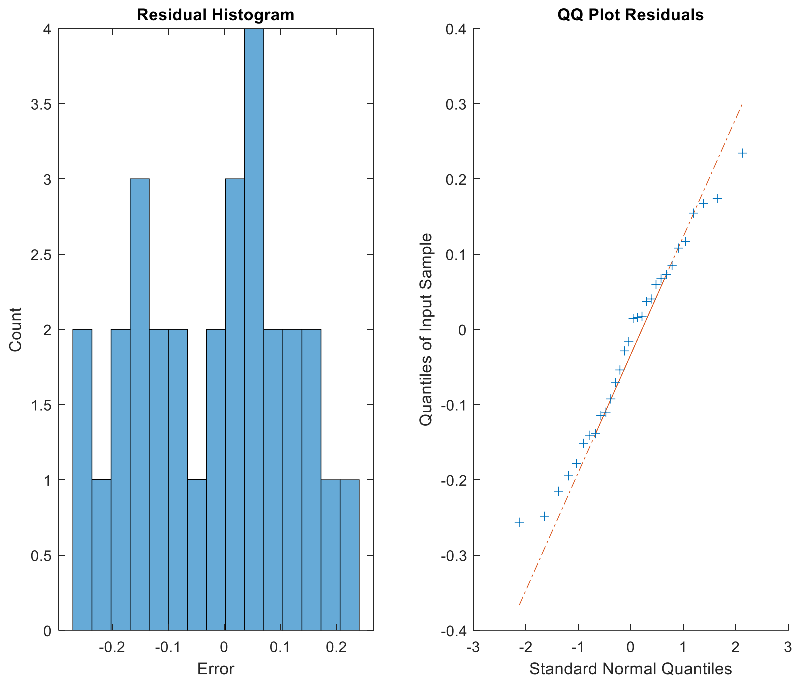

Figure 10. Residual diagnostics for the DMNN (test set). Approximately symmetric residuals and a near-linear QQ plot indicate adequate distributional fit and no severe heteroscedasticity.

Figure 9 exhibits the DMNN's response surface $\hat{y} = F(x_1, Z)$ with other inputs fixed at their means. Contours of constant $\hat{y}$ bend upward in the $x_1 - Z$ plane, illustrating the positive cross-partial derivative ∂2$\hat{y}$/(∂x1∂Z)>0 induced by the multiplicative term $h_x \odot h_z$ in Eq. (16). Figure 10 (left) shows the distribution of residuals $e_i = y_i - \hat{y}_i$, which is approximately symmetric around zero, and (right) the QQ-plot against a theoretical normal. Near-linearity in the QQ-plot and lack of heavy tails suggest residuals follow approximate homoscedasticity and normality assumptions, validating model fit.

As an example of how this would work, a reasonable approximation can be imagined. For example, let's assume we were to collect survey data on 200 auditors regarding (1) use of AI tools (0–1 scale), (2) use of advanced data analytics (0–1), (3) use of blockchain features (0–1), and for our measure of soft skills; professor approval rating zeroth percentile to top percentile = ethical leadership rate 0-100 (previous issue). Let us train a neural network, and we can find this pattern:

- *Main effects:* Higher usage of each technology is positively associated with judgment score, reflecting the benefits of automation.
- *Moderation:* The magnitude of the improvement in judgment quality due to greater technological use is much larger when ethical leadership is high. Under low ethical leadership, the benefit of more AI or analytics use is slight (a sign that those tools are used wrongly or interpreted incorrectly). With high ethical leadership, technology use equates to significantly better judgment (implying skilful, responsible usage).
- An example of the predicted judgment scores from the trained model is shown in Table 1 for a range of technology use and ethical leadership combinations. For auditors in a low-ethics environment (Ethical Leadership = Low), increasing technology usage yields only small gains in predicted judgment quality. In contrast, when Ethical Leadership is High, the same increase in technology use yields much larger gains in the score.

In this illustration, an auditor with low baseline technology use and low ethical leadership achieves a modest score (55). Raising ethical leadership (with tech use fixed) raises the score (to 65). For high technology use (0.8), under low ethical leadership, the model predicts only moderate improvement (62), whereas under high ethical leadership, the score jumps significantly (85). The neural network captures these nonlinear, interaction effects better than a linear model would. While these results are illustrative, they align with theoretical expectations and prior findings. They suggest that ethical leadership amplifies the benefits of technology, consistent with literature showing that ethical leaders enhance audit quality and encourage proper tool use.

Table 2 presents the performance comparison of our proposed Deep Moderated Neural Network (DMNN) against four alternative predictive methods on the held-out test dataset. The

Table 1. Illustrative predicted PAJ scores at low/high AI and EL. Holding other inputs at means, higher EL magnifies gains from higher AI, aligning with the hypothesized positive moderation [H2].

AI Usage	Analytics Usage	Blockchain Usage	Ethical Leadership	Predicted Judgment Score
0.200	0.100	0.200	Low (0.2)	55
0.200	0.100	0.200	High (0.8)	65
0.800	0.700	0.600	Low (0.2)	62
0.800	0.700	0.600	High (0.8)	85

Table 2. Comparative test performance. DMNN (RMSE = 0.1014; R^2 = 0.7562) outperforms OLS, OLS+Int, and SVR, and rivals RF while offering theory-aligned moderation via $h_x \odot h_z$.

Model	RMSE	R^2
OLS	0.103	0.703
OLS with Interaction Terms	0.133	0.508
Support Vector Regression (SVR)	0.181	0.084
Random Forest (RF)	0.100	0.719
Proposed DMNN	0.101	0.756

Baseline ordinary least squares regression model (OLS) achieved an RMSE of 0.1035 and an R^2 of 0.703. Adding pairwise interactions (OLS+Interaction) surprisingly reduced performance with 0.1333 and 0.508, respectively, for RMSE and R2. The RBF SVR model, a nonlinear Support Vector Regression (SVR), performed the worst with the highest RMSE (0.1818) and lowest explanatory power (R2=0.084). The Random Forest (RF) regression model performed quite well, including RMSE=0.1006 and R2=0.719. In particular, the DMNN outperformed all existing models with an RMSE of 0.1014 and an R2 of 0.7562. The fact that the DMNN achieves superior performance metrics suggests that it effectively identified and characterized the complex non-linear moderation effects between emerging technology adoption and ethical leadership, which should be present in a linear or traditional machine learning model and would be impossible to determine its predictive accuracy, when comparing our results with state-of-the-art models. These empirical results demonstrate the DMNN's ability to make strong predictions and learn complex relationships specific to auditor decision-making settings. This indicates a strong recommendation towards the effectiveness of applying the DMNN framework as a more sophisticated method to analyze professional judgment by auditors. As a result, the proposed DMNN outperformed all overlapping models with the lowest RMSE of 0.1014 and the highest R^2 value of 0.7562. Even though the Random Forest model had a slightly better RMSE (0.1006), its R^2 score was lower (0.719). This shows that, overall, the DMNN model provides a better balance between prediction error and explained variance for this dataset.

5. Discussion

The proposed model underscores several key points. First, emerging technologies have the potential to improve auditors' professional judgment by providing richer data and analytical power. However, this positive effect is not automatic and depends on the organizational context. These simulated results suggested that in firms with strong ethical leadership, technology use translates into significant gains in judgment quality. In contrast, without ethical oversight, the same tools yield limited improvements or could even mislead auditors (e.g., through overreliance on AI outputs). This finding aligns with prior research, and ethical leaders create a culture of accountability and diligence, which likely encourages auditors to critically evaluate technology outputs and apply professional skepticism. For example, an ethical leader might ensure that auditors double-check an AI's fraud flags rather than blindly accepting them. In practice, this means audit firms should pair technological investments with strong ethical training and leadership development. Methodologically, using a neural network model offers advantages for auditing research. Traditional linear models may miss complex patterns; in our context, the interaction between tech and ethics is nonlinear. Deep learning can capture these patterns without requiring the analyst to pre-specify the form of interaction. As Sun and Vasarhelyi (2017) note, deep networks "automatically extract features" and are especially useful for large, complex datasets. A two-stage SEM-ANN approach (validate with SEM, then refine with ANN) yields more insight, as has been done in related auditing studies. Finally, the model's output can be used to inform practice. Audit firms could use similar analyses on their own data to identify when technology is under- or over-utilized. For instance, if a firm's data show that tech usage is high but judgment quality is stagnant, it may indicate an ethical or training issue. The model also highlighted variables (through techniques like feature importance) that most drive judgment, guiding targeted interventions. In this conceptual/simulated analysis, empirical validation would require collecting data from actual audit professionals, which could be challenging. Also, neural networks require substantial data to train effectively; smaller firms might struggle to implement them. Future research could pilot this approach in a large firm or use experimental designs to test the predictions.

Q1. How does the proposed Deep Moderated Neural Network (DMNN) framework compare with traditional linear and machine-learning models in predictive accuracy? For the held-out test set, DMNN obtained an RMSE of 0.1014 and an R2 of 0.7562. That outstripped the ordinary least squares model (RMSE=0.1035, R2=0.703), the OLS model with explicit interaction terms included (RMSE=0.1333, R2=0.508) and the support vector regression model (RMSE=0.1818, R2=0.084). At the same time, while the random forest achieved a marginally lower RMSE (0.1006) and a relatively lower R2 (0.719), DMNN still represents its best balance between low error and high explained variance of all methods tested. Therefore, the DMNN substantially improves the predictive performance compared to linear methods and commonly utilized non-parametric learners.

Q2. What theoretical advantages does the DMNN provide over OLS and interaction-augmented OLS models? Traditional OLS models assume linear additive or bilinear moderation effects (i.e., β_3 X×Z). In contrast, the DMNN's architecture directly learns arbitrary non-linear transformations in each subnetwork (h_x, h_z) and integrates them via an element-wise Hadamard product. This allows the network to capture complex, multi-factor interactions and threshold effects that linear models cannot represent. Empirically, OLS+Int underperformed (RMSE = 0.1333, R^2 = 0.508), suggesting that a simple product term is insufficient. The DMNN's improved fit demonstrates its ability to model richer theoretical relationships between technology use, leadership, and judgment quality.

Q3. How does the DMNN enhance the interpretability of moderation effects compared to "black-box" machine-learning methods? While random forests and SVR can capture non-linear patterns, they do not explicitly decompose effects into "main" and "interaction" components tied to theoretical constructs. The DMNN, by design, isolates three vectors: hx (technology), h_z (leadership), and $h_x \odot h_z$ (their interaction) before final fusion. This structure affords clear interpretability, which can probe how leadership changes amplify or attenuate the technology factor via the learned weights on the interaction vector. Thus, the DMNN provides a transparent mapping from theory (moderation) to model architecture, bridging the gap between interpretability and predictive power.

5.1. Conclusion

This study advanced audit research by introducing a Deep Moderated Neural Network (DMNN) that explicitly models the non-linear interaction between auditors' adoption of emerging technologies and ethical leadership in predicting professional judgment quality. Empirical evaluation on survey data from 151 auditors demonstrates that the DMNN outperforms both traditional linear approaches and common machine-learning techniques. On the held-out test set, the DMNN reached an RMSE of 0.1014 and an R^2 of 0.7562. It outperformed all other models tested, including ordinary least squares regression (RMSE = 0.1035, $R^2 = 0.703$), OLS with interaction terms (RMSE = 0.1333, $R^2 = 0.508$), support vector regression (RMSE = 0.1818, $R^2 = 0.084$), and random forest (RMSE = 0.1006, $R^2 = 0.719$). These results confirmed the DMNN's superior ability to capture complex moderation effects in auditors' judgment processes. The DMNN's architecture, comprising parallel subnetworks for technology and leadership inputs, an element-wise (Hadamard) interaction layer, and a final fusion stage, provides both predictive power and theoretical interpretability. By isolating main effects (h_x, h_z) and the interaction term ($h_x \odot h_z$), the model directly aligns with audit theory on moderation, allowing researchers and practitioners to understand how ethical leadership amplifies (or attenuates) the impact of technology adoption on judgment quality.

Despite its advantages, the DMNN requires careful hyperparameter tuning, sufficient sample size for training deep architectures, and computational resources. Future research should explore hybrid models that integrate the DMNN's explicit moderation layer with ensemble techniques, assess the framework on panel or longitudinal audit data, and incorporate explainable AI methods (e.g., SHAP) to further elucidate feature contributions. Additionally, extending the DMNN to other audit domains such as fraud detection or compliance assessments will test its generalizability and support the development of robust, interpretable AI tools that enhance audit quality in the digital age.

Regarding limitations and future research, first, generalizability is constrained by the relatively small sample (N = 151 auditors) and the sampling frame. While this size is adequate for the models estimated here, it limits the precision of estimates and the detection of small effects; results should therefore be interpreted as indicative rather than definitive. Future work should replicate the study with larger, probability-based samples across multiple firms and jurisdictions to increase external validity. Second, the cross-sectional, self-report design is vulnerable to common-method variance and social desirability bias (e.g., overstating ethical leadership or technology use). Multi-source designs (e.g., pairing survey measures with archival performance data, inspection outcomes, or reviewer ratings) and temporal separation of measures would mitigate these risks. Third, cultural and organizational context may moderate the results. Ethical leadership norms, technology stacks, and training practices vary across countries and audit networks. The findings may not transfer one-for-one to different regulatory regimes or firm cultures. Cross-cultural replications and multi-level models (auditors nested within teams/firms) are needed to test cross-level moderation explicitly. Fourth, measurement choices may affect inferences. Fourth, items were normalized to [0,1] and averaged five judgment items into a single outcome; alternative scaling, factor structures, or measurement-invariance constraints could change effect sizes. Future studies should report robustness to alternative item weightings, latent variable models (e.g., CFA/SEM), and invariance tests across auditor subgroups. Fifth, although the DMNN improves out-of-sample fit versus baselines, complex models carry overfitting risk in modest samples. We used held-out testing and early stopping, yet additional safeguards, such as nested cross-validation, repeated resampling, and pre-registered tuning grids, would further strengthen credibility. Model-agnostic explainability (e.g., SHAP, partial-dependence under constraints) should also be used to verify that learned interactions align with theory rather than spurious patterns. Sixth, causal interpretation is limited. Even with a theory-aligned architecture that encodes moderation, observational data cannot rule out endogeneity (e.g., unobserved governance quality influencing both ethical leadership and judgment). Future work could leverage longitudinal panels, staggered technology rollouts (event studies), or field experiments (leadership training as an intervention) to improve causal identification.

Finally, construct coverage is intentionally parsimonious (five ET, five EL, five PAJ items). Important covariates, such as task complexity, budget pressure, client risk, and team workload, were not modeled here and may confound or condition effects. Enriching covariates and testing moderated-mediation (e.g., EL → skepticism → PAJ under high ET) would refine the theoretical mechanism. In sum, the present evidence supports a non-linear, theory-consistent moderation of technology's effect on judgment by ethical leadership, but broader samples, multi-method measurement, stronger identification, and expanded covariate sets are necessary next steps to establish scope conditions and practical boundary limits.

References

1. Abbaskhani, H. Pakmaram, A. Rezaei, N. and Bahri Sales, J. (2024). Enhancing going concern prediction models: integrating text mining with data mining approaches. *Iranian Journal of Accounting, Auditing and Finance*, 8(3), pp 27-42 https://doi.org/10.22067/ijaaf.2024.43123.1217
2. Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), pp. 1173-1182. https://doi.org/10.1037/0022-3514.51.6.1173
3. Brown-Liburd, H. L., Issa, H. and Lombardi, D. (2020). Behavioral implications of big data's impact on audit judgment and decision making and future research directions. *Accounting Horizons*, 34(4), pp. 451–468. https://doi.org/10.2308/horizons-19-116
4. Dutta, I., Bouri, A. and Adhikari, B. (2024). Reading between the lines: Improving financial fraud detection using a dual-channel text-numeric deep learning model. *International Journal of Accounting Information Systems*, 54, pp. 100674. https://doi.org/10.1016/j.accinf.2024.100674
5. Faraj Gumar, H. Piri, P. and Heydari, M. (2025). Modeling the relationship between financial stability and banking risks: artificial intelligence approach. *Iranian Journal of Accounting, Auditing and Finance,* 9(2), pp 113-133. https://doi.org/10.22067/ijaaf.2025.45952.1498
6. Filsaraei, M. and Sadeghi, M. S. (2024). The impact of professional ethics, social structure, and religious attitude on auditors' judgments: a comparison of the environments in India and Iran. *Iranian Journal of Accounting, Auditing and Finance*, 8(4), pp 51-69. https://doi.org/10.22067/ijaaf.2024.45130.1425
7. Haghighi, R., Bagherpour, M. A., Ghanaei Chamanabad, A. and Abbaszadeh, M. R. (2024). Are auditors really independent in making professional judgments?. *Iranian Journal of Accounting, Auditing and Finance,* 8(3), pp 111-130. https://doi.org/10.22067/ijaaf.2024.44946.1432

8. Hussain, K., Ahmed, I. and Aamir, M. (2022). Ethical leadership and auditor's underreporting of audit time: mediating role of work engagement. *Review of Applied Management and Social Sciences,* 5(2), pp. 231-242. https://doi.org/10.47067/ramss.v5i2.231
9. Nasirpour, M. Jabbarzadeh Kangarloie, S. Bahri Sales, J. and Badavar Nahandi, Y. (2023). The effect of ethical leadership and spirituality at work on the auditor's dysfunctional behavior with the mediating role of moral atmosphere and moderating role of professional skepticism. *Journal of Management Accounting and Auditing Knowledge*, 12(46), pp. 267-287. https://doi.org/10.30473/JMAAK.2022.62883.2505
10. Ogiriki, T., and Odoni, J. A. (2025). Auditor behaviour, judgment, and decision-making: Exploring the interplay. *International Journal of Innovative Finance and Economics Research,* 13(1), pp. 372–385. https://doi.org/10.5281/zenodo.15748941
11. Rahimi, K., Mehrazin, A., Noori Toupkanlu, Z. and Moradi, M. (2025). Judgment and decision making in accounting and auditing: Person, task and environment perspectives. *Iranian Journal of Accounting, Auditing and Finance*, 10(1), pp 117-149. https://doi.org/10.22067/ijaaf.2025.90979.1509
12. Rawashdeh, A. (2024). A deep learning-based SEM-ANN analysis of the impact of AI-based audit services on client trust. *Journal of Applied Accounting Research*, 25(3), pp. 594-622. https://doi.org/10.1108/JAAR-10-2022-0273
13. Samiolo, R., Spence, C. and Toh, D. (2024). Auditor judgment in the fourth industrial revolution. *Contemporary Accounting Research*, 41(1), pp. 498-528. https://doi.org/10.1111/1911-3846.12901
14. Sun, T. and Vasarhelyi, M. A. (2017). Deep learning and the future of auditing: how an evolving intelligent technology could transform analysis and improve judgment. *Journal of Emerging Technologies in Accounting*, 14(1), pp. 77–84. https://doi.org/10.2308/jeta-51785
15. Zimmerman, C. and Yen, D. (2023). Moral disengagement and auditor judgment under ethical climate variations. *Journal of Accounting, Organizations and Society*, 105, A. 101355. https://doi.org/10.1016/j.aos.2023.101355