InSight Bulletin: A Multidisciplinary Interlink International Research Journal
Peer Reviewed International, Open Access Journal.
ISSN : 3065-7857 / Website: https://ibrj.us / Volume-3, Issue-1 / January - 2026

Original Article

A Comparative Review- From Artificial Intelligence to Deep Neural Networks: Evolution, Architecture, Capabilities, and Applications

Archana Milind Tank
Ram Krishna More, Arts, Commerce and Science College (Autonomous), Akurdi, Pune

Manuscript ID:
IBMIIRJ -2026-030112

Submitted: 07 Dec. 2025

Revised: 11 Dec. 2025

Accepted: 06 Jan. 2026

Published: 31 Jan. 2026

ISSN: 3065-7857

Volume-3

Issue-1

Pp. 57-63

January 2026

Correspondence Address:
Archana Milind Tank
Ram Krishna More, Arts, Commerce and Science College (Autonomous), Akurdi, Pune
Email: archanamtank@gmail.com

OPEN ACCESS

Quick Response Code:

Web. https://ibrj.us

DOI: 10.5281/zenodo.18956695

DOI Link:
https://doi.org/10.5281/zenodo.18956695

Abstract

In moment's digital and tech- driven world, terms like AI (Artificial Intelligence), ML (Machine Learning), DL (Deep Learning). These buzzwords are frequently used interchangeably, creating confusion about their true meaning. While they partake some parallels, each field also has its own unique characteristics. AI serves as the Broadview, encompassing conception, while ML learns retired patterns and connections from your data and focuses on developing algorithms that can help you prognosticate what will be next. DL is a technical subset of ML that uses Deep Neural Networks (DNNs) — networks with multiple retired layers to dissect data. The depth of the network allows it to automatically learn complex features from raw data, bypassing the need for mortal- engineered point birth. This paper presents a relative study of Artificial Intelligence (AI), Machine learning (ML) and Deep Learning (DL)on crucial generalities like elaboration infrastructures, capabilities, and operations.

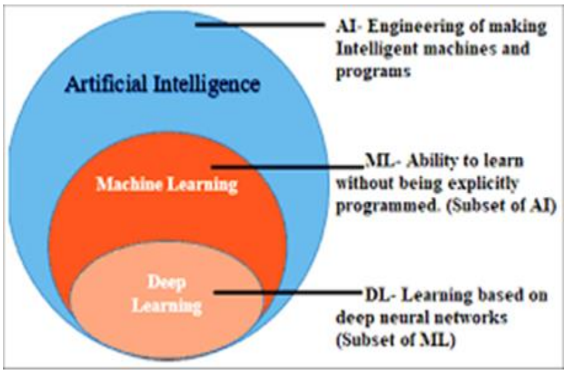
Keywords: *Artificial Intelligence, Machine Learning, Deep leaning, Deep neural network, Supervised learning, Unsupervised learning, Reinforcement learning, Artificial neural network, Feed forward Neural Networks, Convolution neural networks, Recurrent Neural Networks, Activation Function, Generative Adversarial Networks, Multilayer Perceptrons.*

Introduction

Artificial Intelligence (AI) has emerged as a transformative discipline aimed at enabling machines to perform tasks that typically require human intelligence, such as reasoning, learning, perception, and decision-making. Over the past few decades, AI has evolved from rule-based symbolic systems to data-driven approaches that leverage statistical learning and optimization techniques. As computational power and data availability have increased, AI-based systems have demonstrated significant success across diverse domains, including healthcare, finance, transportation, and natural language processing (Russell & Norvig, 2021). Machine Learning (ML) represents a crucial subset of AI that focuses on developing algorithms capable of learning patterns and relationships directly from data without being explicitly programmed. ML models improve their predictive performance by identifying statistical regularities in historical data and applying this learned knowledge to unseen scenarios. Common learning paradigms within ML include supervised learning, unsupervised learning, and reinforcement learning, each suited to different problem settings and data characteristics (Mitchell, 1997). Deep Learning (DL), a specialized branch of machine learning, employs deep neural networks composed of multiple hidden layers to automatically extract hierarchical representations from raw data. Unlike traditional ML techniques that rely heavily on manual feature engineering, deep learning models can learn complex, non-linear feature representations directly from large-scale datasets. This capability has enabled significant advancements in image recognition, speech processing, natural language understanding, and generative modeling (Good fellow, Bengio, & Courville, 2016).

With the rapid development of deep neural network architectures and training techniques, generative AI systems have further expanded the scope of intelligent applications by producing realistic images, text, and audio content. Understanding the conceptual distinctions, architectural differences, and application domains of AI, ML, and DL is therefore essential for researchers and practitioners. This paper presents a comparative analysis of Artificial Intelligence, Machine Learning, and Deep Learning, emphasizing their evolution, core architectures, capabilities, and practical applications.

How to cite this article:
Tank, A. M. (2026). A Comparative Review- From Artificial Intelligence to Deep Neural Networks: Evolution, Architecture, Capabilities, and Applications. InSight Bulletin: A Multidisciplinary Interlink International Research Journal, 3(1), 57–63. https://doi.org/10.5281/zenodo.18956695



Overview of Artificial Intelligence

Artificial Intelligence (AI) is a multidisciplinary field of computer science concerned with the design and development of systems capable of exhibiting intelligent behavior. Such systems are designed to perceive their environment, analyze information, and take actions that maximize the likelihood of achieving predefined goals. AI seeks to replicate key aspects of human cognition, including learning, reasoning, problem-solving, perception, and language understanding, through computational models and algorithms (Russell & Norvig, 2021). At its core, AI integrates data processing with decision-making mechanisms to enable machines to operate autonomously in complex and dynamic environments. Modern AI systems rely on large volumes of data and advanced computational techniques to identify patterns, infer relationships, and generate predictions. These capabilities allow AI-driven solutions to perform tasks such as speech recognition, computer vision, automated planning, and intelligent recommendation systems, which were traditionally considered exclusive to human intelligence (Nilsson, 2014). Learning constitutes one of the fundamental components of artificial intelligence, enabling systems to improve their performance over time through experience. Learning mechanisms in AI are commonly categorized into supervised, unsupervised and reinforcement learning paradigms. Through these approaches, intelligent agents adapt their behavior based on feedback obtained from data or interactions with the environment, leading to progressively optimized outcomes (Poole &Mackworth, 2017). Another essential aspect of AI is reasoning, which allows systems to draw logical conclusions and make informed decisions. Reasoning techniques in AI include deductive, inductive, and abductive inference, each supporting different forms of knowledge representation and decision-making. By combining reasoning with heuristic search strategies, AI systems can efficiently explore large problem spaces to identify optimal or near-optimal solutions (Russell & Norvig, 2021). Perception and natural language understanding further extend the capabilities of AI by enabling machines to interpret sensory inputs such as images, audio, and textual data. Through advances in computer vision and natural language processing, AI systems can recognize objects, understand spoken or written language, and interact with humans in an intuitive manner. These capabilities form the foundation of intelligent assistants, autonomous systems, and human–computer interaction technologies (Good fellow et al., 2016).

Machine Learning

Machine Learning (ML) is a core subfield of Artificial Intelligence that focuses on the development of algorithms capable of learning from data and improving performance over time without explicit programming. Instead of relying on predefined rules, ML systems identify patterns, trends, and relationships within data to make predictions or decisions when exposed to new and unseen inputs. This data-driven approach has made machine learning a foundational technology in modern intelligent systems (Mitchell, 1997). Machine learning models are typically trained using historical datasets, where statistical and computational techniques are employed to minimize prediction error and enhance generalization. The effectiveness of an ML model depends on several factors, including the quality and quantity of data, feature representation, algorithm selection, and model evaluation strategies. ML techniques are widely applied in tasks such as classification, regression, clustering, and anomaly detection across domains including healthcare, finance, cyber security, and recommendation systems (Bishop, 2006).

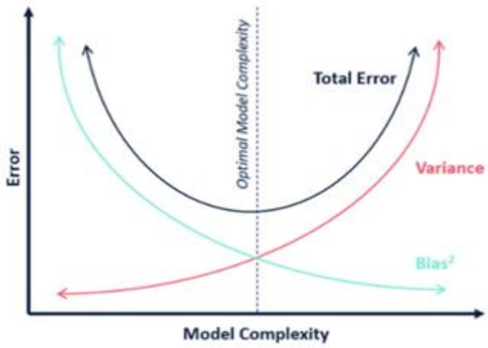
Types of Machine Learning

Machine learning approaches are broadly categorized into supervised, unsupervised, and reinforcement learning, based on the nature of the data and the learning process involved.

- **Supervised learning** involves training models on labeled datasets, where each input is associated with a known output. The learning objective is to map input features to correct outputs, enabling tasks such as classification and regression. Common supervised learning algorithms include linear regression, decision trees, support vector machines, and k-nearest neighbors (Hastie, Tibshirani, & Friedman, 2009).
- **Unsupervised learning** operates on unlabeled data, allowing models to discover hidden structures or patterns without explicit guidance. Techniques such as clustering and dimensionality reduction help in organizing data into meaningful groups or reducing feature complexity. Algorithms like k-means clustering and principal component analysis are frequently used in exploratory data analysis and knowledge discovery (Bishop, 2006).
- **Reinforcement learning** is a learning paradigm in which an agent interacts with an environment and learns optimal behavior through trial and error. The agent receives feedback in the form of rewards or penalties, guiding it toward strategies that maximize cumulative rewards over time. Reinforcement learning has been successfully applied in robotics, game playing, autonomous systems, and control optimization problems (Sutton & Barto, 2018).

Model Generalization: Under fitting and Over fitting

A critical challenge in machine learning is achieving a balance between model complexity and generalization capability. Under fitting occurs when a model is too simple to capture underlying data patterns, leading to high bias and poor performance on both training and unseen data. In contrast, over fitting arises when a model becomes excessively complex and learns noise or irrelevant patterns from the training data, resulting in high variance and reduced predictive accuracy on new data (Hastie et al., 2009).



To address these issues, techniques such as feature engineering, cross-validation, and regularization are commonly employed. Regularization methods, including L1 (LASSO) and L2 (Ridge) regularization, introduce penalty terms into the loss function to constrain model complexity and improve generalization. By controlling the magnitude of model parameters, regularization helps prevent over fitting while maintaining predictive performance (Bishop, 2006)

Deep Learning

Deep Learning (DL) is an advanced subset of machine learning that focuses on learning hierarchical representations of data through deep neural network architectures. Unlike traditional machine learning approaches that rely heavily on handcrafted features, deep learning models automatically extract relevant features directly from raw data by leveraging multiple layers of nonlinear transformations. This capability has enabled significant improvements in handling complex and high-dimensional data such as images, audio signals, text, and video (Good fellow et al., 2016). The fundamental building block of deep learning is the artificial neural network, which is inspired by the biological structure of the human brain. Neural networks consist of interconnected processing units, known as neurons, organized into input, hidden, and output layers. Each neuron performs a weighted sum of inputs followed by a nonlinear activation function, allowing the network to model complex and non-linear relationships within data. As the number of hidden layers increases, the network gains the ability to learn increasingly abstract and high-level features, which is why such models are referred to as "deep" networks (Le Cun, Bengio, & Hinton, 2015). Training deep neural networks involves an iterative optimization process that minimizes prediction error using algorithms such as gradient descent combined with back propagation. During forward propagation, input data is passed through successive layers to generate predictions, while back propagation adjusts the network's weights by propagating the error backward through the layers. This learning process enables deep learning models to continuously improve their performance as they are exposed to larger and more diverse datasets (Rumelhart, Hinton, & Williams, 1986).

Deep learning has demonstrated remarkable success across a wide range of real-world applications. In computer vision, deep convolutional neural networks achieve state-of-the-art performance in image classification, object detection, and facial recognition tasks. In natural language processing and speech recognition, deep learning models enable machine translation, sentiment analysis, and voice-controlled intelligent systems. More recently, deep generative models have expanded the scope of artificial intelligence by producing realistic synthetic data, including images, text, and audio, further enhancing creativity and automation in intelligent systems (Le Cun et al., 2015; Good fellow et al., 2016).

Components of a Neural Network and Activation Functions

Artificial neural networks form the foundational architecture of deep learning systems and are designed to model complex relationships within data. A neural network is composed of multiple interconnected layers, each responsible for a specific stage of information processing. The primary layers include the input layer, one or more hidden layers, and the output layer, which collectively enable the transformation of raw data into meaningful predictions or classifications (Good fellow et al., 2016). The input layer serves as the entry point of data into the network, where each neuron represents a feature of the input dataset. This layer does not perform computation but transmits the input values to subsequent layers. The hidden layers perform the core computational tasks by applying weighted transformations and nonlinear activation functions to extract relevant features and learn internal representations of the data. The depth and structure of hidden layers determine the model's ability to capture complex and abstract patterns. The output layer produces the final result of the network, such as a predicted value or class label, based on the task being performed (Le Cun et al., 2015). Activation functions play a critical role in neural networks by introducing non-linearity into the model, allowing it to learn complex and non-linear relationships. Without activation functions, neural networks would be limited to modeling linear transformations, significantly reducing their expressive power. Activation functions operate by transforming the weighted sum of inputs and biases of a neuron before passing the output to the next layer (Bishop, 2006).

Several activation functions are commonly used in neural network architectures. The linear activation function is primarily applied in regression tasks where continuous output values are required. The sigmoid function maps input values to a range between 0 and 1, making it suitable for binary classification problems and probabilistic interpretations. However, sigmoid functions may suffer from vanishing gradient issues in deep networks. The hyperbolic tangent (tanh) function addresses this limitation to some extent by producing outputs in the range of −1 to 1, enabling better gradient flow during training (Goodfellow et al., 2016). The Rectified Linear Unit (ReLU) has become the most widely used activation function in deep learning due to its computational efficiency and ability to mitigate vanishing gradient problems. ReLU outputs zero for negative inputs and a linear response for positive inputs, enabling faster convergence during training. Variants such as Leaky ReLU and Parametric ReLU have been proposed to address limitations associated with inactive neurons. For multiclass classification tasks, the softmax activation function is commonly employed in the output layer, as it converts raw output scores into a probability distribution across multiple classes (Le Cun et al., 2015).

Working of Neural Networks: Forward and Back propagation

The operation of a neural network is primarily governed by two fundamental processes: forward propagation and back propagation. Together, these processes enable the network to generate predictions and iteratively improve its performance by adjusting internal parameters during training.

Forward propagation refers to the process by which input data flows through the network from the input layer to the output layer. At each layer, neurons compute a weighted sum of their inputs, add a bias term, and apply an activation function to produce an output. These outputs are then passed to the next layer, allowing information to propagate through the network. The final layer generates the network's prediction, which may represent a continuous value in regression tasks or class probabilities in classification problems. The quality of the prediction is evaluated using a loss function that quantifies the difference between the predicted output and the actual target value (Good fellow et al., 2016).

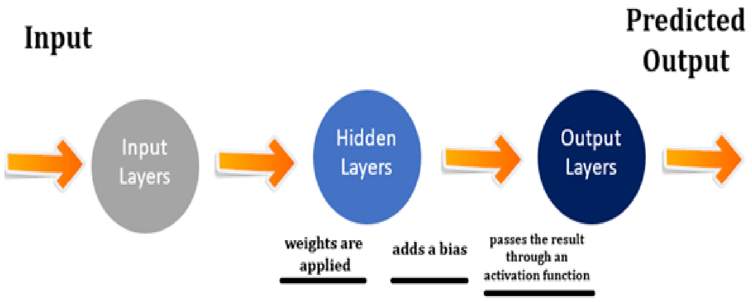
Back propagation is the learning mechanism that enables neural networks to minimize prediction error by systematically updating model parameters. After computing the loss during forward propagation, the error is propagated backward through the network using the chain rule of calculus. This process calculates the gradient of the loss function with respect to each weight and bias in the network. Optimization algorithms, such as gradient descent and its variants, use these gradients to adjust the parameters in a direction that reduces the overall error. Through repeated iterations of forward and backward propagation, the network gradually learns optimal representations of the data (Rumelhart et al., 1986). The effectiveness of neural network training depends on several factors, including learning rate selection, weight initialization, network depth, and the choice of activation and loss functions. Proper tuning of these components helps ensure stable convergence and prevents issues such as vanishing or exploding gradients. When trained on sufficiently large and representative datasets, neural networks can achieve high levels of accuracy and generalization across a wide range of complex tasks (Le Cun et al., 2015).

Types of Neural Networks in Deep Learning

Deep learning employs a variety of neural network architectures, each designed to address specific types of data and problem domains. These architectures differ in structure, information flow, and learning mechanisms, enabling efficient modeling of complex patterns in high-dimensional datasets.

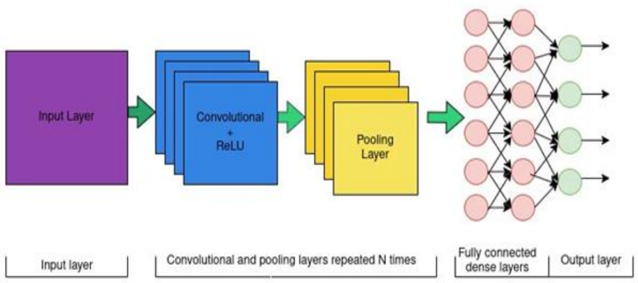
Feed forward Neural Networks (FNNs)

Feed forward neural networks represent the simplest form of neural network architecture, in which information flows in a unidirectional manner from the input layer to the output layer through one or more hidden layers. These networks do not contain cycles or feedback connections, making them suitable for tasks such as classification and regression. Feed forward networks typically rely on nonlinear activation functions and loss functions such as mean squared error or cross-entropy to optimize performance (Bishop, 2006).



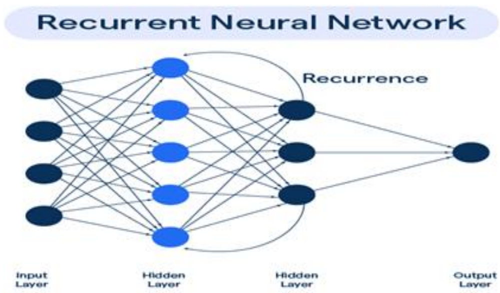
Convolution Neural Networks (CNNs)

Convolution neural networks are specialized architectures designed to process grid-like data structures, particularly images and videos. CNNs utilize convolution layers with learnable filters to automatically extract spatial features such as edges, textures, and shapes from input data. Pooling layers are often incorporated to reduce dimensionality and enhance spatial invariance, while fully connected layers integrate extracted features for final prediction. CNNs have achieved remarkable success in applications including image classification, object detection, medical imaging, and facial recognition (LeCun et al., 1998; LeCun et al., 2015).



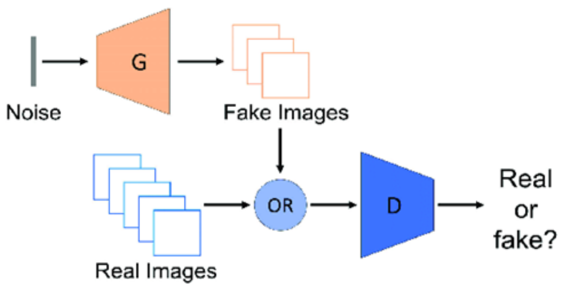
Recurrent Neural Networks (RNNs)

Recurrent neural networks are designed to model sequential and temporal data by incorporating feedback connections that allow information from previous time steps to influence current outputs. This internal memory mechanism enables RNNs to capture temporal dependencies in data sequences such as text, speech, and time-series signals. Variants such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) address limitations related to vanishing gradients, making them effective for natural language processing, speech recognition, and sequence prediction tasks (Hochreiter & Schmidhuber, 1997).



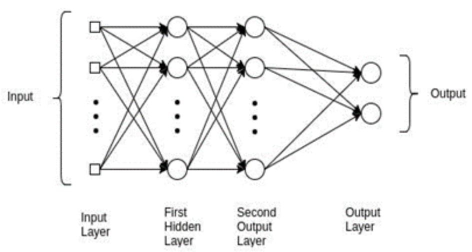
Generative Adversarial Networks (GANs)

Generative Adversarial Networks consist of two competing neural networks: a generator and a discriminator. The generator attempts to produce synthetic data that resembles real data, while the discriminator evaluates whether the generated samples are authentic or artificial. Through this adversarial training process, GANs learn to generate highly realistic data samples. GANs are widely used in image synthesis, data augmentation, style transfer, and creative content generation (Goodfellow et al., 2014).



Multilayer Perceptron (MLPs)

Multilayer perceptron are fully connected feed forward networks consisting of multiple hidden layers and nonlinear activation functions. MLPs are capable of modeling complex relationships between inputs and outputs and are commonly used in tasks such as pattern recognition, classification, and regression. While they lack the spatial and temporal modeling capabilities of CNNs and RNNs, MLPs remain effective for structured and tabular data (Hastie et al., 2009).



Applications of Neural Networks and Deep Learning

Neural networks and deep learning techniques have become integral to modern intelligent systems due to their ability to automatically learn complex representations from large and diverse datasets. Their data-driven nature and high predictive accuracy have enabled widespread adoption across multiple domains, transforming traditional approaches to problem-solving and decision-making. In the field of computer vision, deep learning models—particularly convolution neural networks—are extensively used for image classification, object detection, facial recognition, and medical image analysis. These models automatically extract hierarchical visual features, enabling accurate interpretation of complex visual data. Applications such as autonomous driving, surveillance systems and diagnostic imaging rely heavily on deep neural networks to achieve robust and real-time visual perception (LeCun et al., 2015). Natural language processing (NLP) is another prominent application area where deep learning has significantly advanced machine understanding of human language. Recurrent neural networks and transformer-based architectures enable tasks such as machine translation, sentiment analysis, speech recognition, and text summarization. Deep learning-powered language models facilitate more natural human–computer interaction in applications including virtual assistants, chat bots, and automated content generation systems (Goldberg, 2017).

In healthcare, deep learning techniques support disease diagnosis, medical image interpretation, drug discovery, and personalized treatment planning. Neural networks are employed to analyze medical records, genomic data, and radiological images, assisting clinicians in early disease detection and improved decision-making. These applications contribute to enhanced diagnostic accuracy and improved patient outcomes (Esteva et al., 2017). The financial sector utilizes neural networks for fraud detection, credit risk assessment, algorithmic trading, and customer behavior analysis. Deep learning models excel at identifying subtle patterns and anomalies within large transactional datasets, enabling organizations to detect fraudulent activities and optimize financial decision-making processes in real time (Good fellow et al., 2016). In autonomous systems and robotics, deep learning plays a critical role in perception, navigation, and control. Neural networks process sensory inputs such as camera images, lieder data, and sensor signals to enable autonomous vehicles and robotic systems to interact safely and efficiently with their environment. Reinforcement learning combined with deep neural networks has further enhanced decision-making capabilities in dynamic and uncertain environments (Sutton &Barto, 2018). Additionally, generative models, including Generative Adversarial Networks, have expanded the creative potential of artificial intelligence by enabling the synthesis of realistic images, videos, music, and text. These models are widely applied in data augmentation, content creation, virtual reality, and simulation-based training, demonstrating the growing influence of deep learning beyond traditional analytical tasks (Goodfellow et al., 2014).

Conclusion

This paper has presented a comparative review of Artificial Intelligence, Machine Learning, and Deep Learning, highlighting their conceptual foundations, architectural differences, and practical capabilities. Artificial Intelligence represents the broader goal of enabling machines to exhibit intelligent behavior, while Machine Learning provides data-driven methods that allow systems to learn from experience and improve performance over time. Deep Learning, as an advanced subset of machine learning, further enhances these capabilities by leveraging deep neural network architectures to model complex and high-dimensional data. The analysis demonstrates that while traditional AI approaches rely on rule-based reasoning and symbolic representations, machine learning and deep learning techniques emphasize statistical learning and pattern recognition. Machine learning methods are effective for structured data and moderately complex tasks, whereas deep learning models excel in handling large-scale, unstructured data such as images, speech, and natural language. The hierarchical feature learning capability of deep neural networks significantly reduces the dependence on manual feature engineering and enables superior performance in complex real-world applications. Furthermore, the rapid advancement of neural network architectures, training algorithms, and computational resources has expanded the scope of intelligent systems across diverse domains, including computer vision, natural language processing, healthcare, finance, and autonomous systems. Despite their success, challenges related to data requirements, computational cost, model interpretability, and ethical considerations remain important areas for future research.

In conclusion, understanding the evolution and interrelationship of Artificial Intelligence, Machine Learning, and Deep Learning is essential for researchers and practitioners aiming to develop robust and scalable intelligent systems. Continued research and innovation in these fields will play a crucial role in shaping the future of intelligent technologies and their responsible integration into society.

Comparative Analysis of Artificial Intelligence, Machine Learning, and Deep Learning

Aspect	Artificial Intelligence (AI)	Machine Learning (ML)	Deep Learning (DL)
Definition	A broad interdisciplinary field focused on designing systems capable of performing tasks that require human-like intelligence, such as reasoning, decision-making, and perception.	A subset of AI that enables systems to learn patterns and relationships from data to improve performance without explicit rule-based programming.	A specialized subset of machine learning that utilizes deep neural network architectures to learn hierarchical representations from large-scale data.
Primary Objective	To simulate intelligent behavior and cognitive functions in machines.	To develop predictive models by learning from historical data.	To automatically extract complex features and representations from raw, high-dimensional data.
Learning Approach	May involve rule-based systems, symbolic reasoning, or learning-	Relies on statistical learning techniques and optimization	Uses multi-layered neural networks trained through back

	based methods.	algorithms.	propagation and gradient-based optimization.
Data Dependency	Can function with limited or structured data in rule-based systems.	Requires labeled or unlabeled structured data for effective learning.	Requires large volumes of data, often unstructured, to achieve high accuracy.
Model Complexity	Varies widely depending on the approach used.	Moderate complexity, dependent on algorithm and feature engineering.	High complexity due to deep architectures and large parameter spaces.
Feature Engineering	Often manual and knowledge-driven.	Largely manual, requiring domain expertise.	Mostly automatic, with features learned directly from data.
Computational Requirements	Low to moderate, depending on implementation.	Moderate, requiring efficient computation for training models.	Very high, often requiring GPUs or specialized hardware.
Interpretability	Generally high in rule-based systems.	Moderate, depending on the model used.	Relatively low due to the black-box nature of deep networks.
Typical Applications	Expert systems, planning, game playing, and intelligent agents.	Classification, regression, clustering, recommendation systems.	Image and speech recognition, natural language processing, generative modeling, autonomous systems.

Acknowledgment
The authors sincerely acknowledge the support and guidance provided by the Department of Arts, Commerce and Science College (Autonomous), Akurdi, Pune, for facilitating this research. We are grateful to the faculty members for their valuable suggestions and encouragement. We also thank the scholars whose work contributed to this study and our family members for their constant support.

Financial support and sponsorship
Nil.

Conflicts of interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

References
1. Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
2. Esteva, A., Kuprel, B., Novoa, R. A., KO, J., Swetter, S. M., Blau, H. M., &Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056
3. Goldberg, Y. (2017). Neural network methods for natural language processing. Morgan & Claypool.
4. Good fellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
5. Good fellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., &Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672–2680.
6. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd Ed.). Springer.
7. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
8. Le Cun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, *521*(7553), 436–444. https://doi.org/10.1038/nature14539
9. Le Cun, Y., Bottom, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791
10. Mitchell, T. M. (1997). Machine learning. McGraw-Hill.
11. Nilsson, N. J. (2014). Principles of artificial intelligence. Morgan Kaufmann.
12. Poole, D. L., &Mackworth, A. K. (2017). Artificial intelligence: Foundations of computational agents (2nd Ed.). Cambridge University Press.
13. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0
14. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th Ed.). Pearson.
15. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd Ed.). MIT Press.