From Machine Learning to Deep Learning: Advances in Techniques to Practical Applications

Dr. Jayeshkumar Madhubhai Patel

Associate Professor – PG Department [M.C.A.], Ganpat University, Gujarat, INDIA.

Abstract

The complexity and increasing frequency of cyber threats necessitate the development of advanced detection systems capable of identifying both known and emerging attacks. In this study, we present a hybrid Network Intrusion Detection System (NIDS) based on anomaly detection, which integrates several machine learning and deep learning algorithms, including XGBoost, Random Forest, Graph Neural Networks (GNN), Long Short-Term Memory (LSTM), and autoencoders. The study investigates the realworld application areas in which ML and DL techniques are implemented. Moreover, the study highlights the latest trends and possible future avenues for research and development (R&D), which consist of developing hybrid models, generative AI, and incorporating ML and DL with the latest technologies. The study aims to provide a comprehensive view on ML and DL technologies, which can serve as a reference guide for researchers, industry professionals, practitioners, and policy makers.

Keywords: Artificial Intelligence, Deep Learning, Engineering, Systematic Literature Review, Neural Networks, Machine learning algorithm, Deep neural network architectures, Cybersecurity, Network intrusion detection system and Ensemble learning

1. Introduction

Developing machines with the ability to 'think' has been a long-running aspiration of inventors throughout history. The popular idea of replicating intelligent human behavior arranged as processes in machines [1] has fueled researchers' imaginations. In the present time, artificial intelligence (AI) is a thriving and rapidly changing field with various applications in society and economics, such as understanding speech or images, textual analysis, and in supporting an actively growing research body. Machine learning (ML), a part of AI, is a multidisciplinary field spanning computer science, statistics, and data science that addresses the need for computers to improve automatically through experience and by the use of data. ML is advancing evidence-based decision-making in the fields of healthcare, education, national security, finance, economics, manufacturing, and marketing, specifically by implementing various approaches to teach computers to achieve tasks. However, conventional ML techniques cannot efficiently process raw data and require mindful engineering and great expertise. In the real world, every piece of data may be influenced by different factors of variations, thus requiring humans to factor in those variations and decide whether to incorporate them or not. Overcoming such flaws, deep learning (DL) has recently emerged as a promising approach in ML, currently dominating the majority of the works in the field of ML.

While it may appear as a seemingly new concept, the idea of DL can be traced back to the 1940s and subsequently underwent roughly three waves of development with the most recent current revival beginning in 2006. During the first wave between 1940 and 1960, DL was known as cybernetics, then it gained popularity again in the 1980s-1990s as connectionism. Fundamental methods such as radial basis function networks and multilayer perceptrons were employed in 2014 to solve the problem of designing mobile adaptive tracking controllers. These two neural networks were found suitable for decision-making and control. Later, pointed out a few reasons why DL rose to prominence in the twenty-first century, including

the surge of "big data" with quality labels, improvements in regularization techniques, development of near-perfect optimization algorithms, creation of niche software platforms that can enable the integration of architectures, and advancements in parallel computing power and multi-core, multi-threaded execution. In fact, big data became a huge issue for conventional ML algorithms along with the increasing size of the network, whereby the performance of old algorithms either became overloaded or deteriorated. The enhanced performance of DL can be attributed to its ability to utilize two or more levels of non-linear feature transformation of the given data[2].

Deep learning allows computational models with multiple layers to gradually extract higher-level features from the raw input. The "deep" in DL, therefore, denotes a high credit assignment path (CAP) depth, which has been assigned a value of 2 by most researchers. Deep learning enables computers to learn complex concepts by forming them out of simple ones. Adequately explained that, "Deep learning is a particular kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones"[9].

2. Related Work

Deep learning (DL) is considered an evolution of machine learning (ML) that incorporates algorithms to learn from data to accomplish some tasks without being explicitly programmed. Both ML and DL are a subset of artificial intelligence (AI). ML powers a wide range of automated functions in various businesses, from data security services hunting down malware to finance specialists looking for trade warnings. It has also a wide variety of applications in modern society, such as: developing intelligent personal assistants for finding helpful information, recommender systems that can suggest relevant items to the users, machine translation to provide the most accurate translation of any text in any language [3], and predicting the class of object in an image. The way machines can learn new techniques becomes interesting whenever deep learning techniques are employed. The effectiveness of traditional machine-learning approaches is comparatively lower than DL techniques, considering that they require a large volume of data to provide significant results. For a long time, designing a feature extractor for machine learning systems demanded hand-crafted features to simplify the learning process. However, such feature extraction techniques need human expertise and significant domain understanding [5].

Deep learning allows machines to learn from their mistakes and comprehend the world as a hierarchy of concepts. In this learning process, the machines learn from data using a general-purpose learning algorithm, thus needing less human expertise to describe all the knowledge that the machine requires expressly. The models of DL employ a layered network architecture, known as an artificial neural network (ANN), which is modelled after the human brain's analogous networks. The embedding of layers results in a significantly more efficient learning experience than traditional machine learning models. The ability of deep learning to achieve high-level features from a massive amount of input data, referred to as feature engineering, distinguishes it from machine learning. As a result, deep learning is gaining popularity with innovative applications in natural language processing (NLP), computer vision, and predictive modelling[4].

3. Materials and Methods

This review utilized an integrative literature method to analyze almost all available deep learning modelling approaches, as well as their current progress, applications, advantages, and challenges. Throughout this method, relevant and reliable papers were selected, collected, filtered, carefully evaluated and analyzed. The database was found using credible websites, e.g. Scopus and refereed journals from reputable publishers such as Nature, Elsevier, Taylor & Francis, Springer, Wiley, ACS, Inderscience, MDPI, Frontiers, and Sage. Relevant keywords such as "Deep learning", "Deep learning architecture", "Deep learning modelling", "Advantages of deep learning", "Challenges of deep learning", "Future of deep learning", and each deep learning model such as "Vector space model", "Convolutional neural network", "Recurrent neural network" and so on, were used to find out publications related to

the present work. Through the Scopus database, 186,154 papers published within the last five years were identified. The references and bibliographies of the aforementioned publications were sifted and compiled in order to locate more relevant papers. The following criteria were used to thoroughly scan and categorize the abstract, introduction, and conclusion from selected papers:

(i.) Preliminary consideration was given to only peer-reviewed articles from reputable publishers and websites
(ii.) Researchers who are actively engaged in the relevant research field were chosen and collected
(iii.) The selected publications were evaluated for their balance between modern studies and prior research
(iv.) Referring to websites that employ the aforementioned keywords commercially
(v.) The most recent and cutting-edge algorithms relevant to the present work were emphasised
(vi.) Some publications of relevance that were cited in recent studies were rigorously retrieved as the original source of the studies.

The above criteria assisted in selecting papers that are more relevant. Throughout the entire review process of available relevant papers, several questions were raised. To answer these questions, some other references were sourced and examined for further clarification and improvement. Papers for the present study were chosen based on a set of inclusion and exclusion criteria[6].

4. Proposed model

The proposed model integrates multiple machine learning and deep learning techniques through a hybrid ensemble framework to enhance intrusion detection accuracy and robustness. First, the raw network trafc data is preprocessed by handling missing values, encoding categorical variables, normalizing numerical features, and applying the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. Five individual models are trained separately: XGBoost, Random Forest, Graph Neural Network (GNN), Long Short-Term Memory (LSTM), and Autoencoder. XGBoost and Random Forest are employed to capture non-linear feature interactions and provide diverse perspectives from ensemble-based learners. GNN is designed to exploit graph-structured relationships between entities in network trafc, such as IP addresses and ports, using message-passing mechanisms across graph layers. Te LSTM model captures temporal dependencies within sequential network trafc flows, learning patterns over time that indicate potential intrusions.

Meanwhile, the Autoencoder operates in an unsupervised manner to learn compressed representations of normal behavior, flagging anomalies based on reconstruction errors. After training the individual models, their outputs whether probability scores or anomaly scores are combined using a weighted soft-voting ensemble strategy, where weights are assigned based on each model's validation performance.

This integration leverages the complementary strengths of traditional machine learning and deep learning architectures, resulting in a robust and generalized detection framework.

5. Discussion

The study aimed to discuss and identify ML and DL's current and potential developments, emphasizing their utilization across different sectors. AI, particularly ML and DL, are gaining immense popularity and actively changing various industries, including business, healthcare, fnance, and many others. In this research, we demonstrated how these techniques are used, how they perform in practice, and their merits and demerits. According to the frst objective, the study identified four significant data types that ML and DL techniques use: structured, unstructured, semi-structured, and time series data. The data were collected from different sources, such as healthcare, e-commerce, marketing transportation, finance, manufacturing, retail, agriculture, education, and entertainment industries, and processed using several ML and DL techniques. For the study's second objective, a comparison table followed by a comprehensive discussion was designed. The literature holistically evaluates the differences between ML and DL regarding data types, algorithms they use, feature engineering, data volume, model complexity, training times, interpretability, computational requirements, and application areas[7].

This study further identified the specific tasks that ML and DL can handle in various application domains. Table 7 discusses the specific techniques for the tasks in detail. For example, in healthcare, ML is used to predict patient outcomes, disease classification, diagnostic processes, and drug discovery, while DL is used to analyze medical images, genomics, and personalized medicine. Similarly, in computer vision, ML is used for image recognition, object detection and classification, and handwriting recognition, and DL is used for facial recognition, real-time video processing, and emotion detection. In other words, ML and DL have vast application domains with huge, specialized tasks[8].

6. Conclusions

The study provides a holistic view of machine learning and deep learning algorithms, their techniques, data types, application domains, and their potential use in future. According to the research objectives, the study comprehensively discussed the different types of machine learning and deep learning techniques implemented in various domains to solve real-world problems. The study identified that the potential for machine learning or deep learning can be assessed based on data and the problem-solving capability of algorithms. Each technique has specifc merits and demerits; using specifc methods depends on the amount or type of data available, the computing power available, and the type of issue being faced

7. Authors' Biography

Dr. Jayeshkumar Madhubhai Patel (b.1978) joined the faculty of the Department of Computer Science, Kadi Sarva VishwaVidyalaya, India in 2003 and presently working as an Associate Professor – PG Department [M.C.A.], Acharya Motibhai Patel Institute of Computer Studies, Ganpat University. He received hir M.S. (2002) and Ph.D (2006) in Computer Science from the Hemchandrachaya North Gujarat University, Patan. He has 92+ publications in books, book chapters, journals, and in the proceedings of national and international conferences. His CAREER AWARD FOR YOUNG TEACHER (CAYT) have got from All India Council for Technical Education (AICTE) , Government of India, New Delhi. He is author of four books. He is supervising work of a few doctoral research students.

8. References

1 Deep Learning: Concepts, Architectures, Workflow, Applications and Future Directions - Jayeshkumar Madhubhai Patel - IJFMR Volume 5, Issue 6, November-December 2023. DOI 10.36948/ijfmr.2023.v05i06.11497

2 Assortment of Machine Learning model on Gujarati Web Page Sets. CD Patel, J Patel. GRADIVA REVIEW JOURNAL 8 (7), 1081-1087, 2022. 2022

3 Natural computing algorithms–a survey; B Chawda, J Patel; International Journal of Emerging Technology and Advanced Engineering 6 (6)

4 Stock Market Portfolio Management: A Walk-through BV Chawda, JM Patel International Journal on Recent and Innovation Trends in Computing

5 Bharat V. Chawda, Jayeshkumar Madhubhai Patel,"Investigating Performance of Various Natural Computing Algorithms", International Journal of Intelligent Systems and Applications(IJISA), Vol.9, No.1, pp.46-59, 2017. DOI:10.5815/ijisa.2017.01.05

6 Strategic Innovations and Future Directions in Deep Learning -Based Intrusion Detection Models - Jayeshkumar Madhubhai Patel - IJFMR Volume 7, Issue 6, November-December 2025. DOI https://doi.org/10.36948/ijfmr.2025.v07i06.65104

7 Nested Learning: The Illusion of Deep Learning Algorithms - Jayeshkumar Madhubhai Patel - International Journal of Engineering and Applied Sciences (IJEAS) Volume 12, Issue 11, November 2025. DOI : https://dx.doi.org/10.31873/IJEAS.12.11.02

8 A Conceptual Framework for Deep Learning Algorithms and Their Applications - Jayeshkumar Madhubhai Patel - IJFMR Volume 6, Issue 6, November-December 2024. DOI https://doi.org/10.36948/ijfmr.2024.v06i06.33737

9 Chowdhary, K.; Chowdhary, K. Natural language processing. In Fundamentals of Artifcial Intelligence; Springer: New Delhi, India,2020