IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Application of Mathematics in Machine Learning

Manisha Anand Gund
Mathematics
Scholar in JJTU, JhunJhunu
Rajasthan, India

Dr. Swati Desai
Professor in Mathematics
JJTU, JhunJhunu
Rajasthan, India

ABSTRACT

In today's era of big data, machine learning is the modern technology. Machine learning is nothing but the application of algorithms to solve the real-life problems. Mathematics provides useful tools for data representation, matrix multiplication, optimization, and decision-making in machine learning algorithms. This paper highlights the importance of these mathematical tools, which include linear algebra, calculus, and Probability theory.

The important pillar of a machine learning algorithm is data. Linear algebra is useful for representing data in matrix form systematically and for reducing the dimension of the given data, making it easier to handle large datasets. Since the data is in matrix form, various operations can be performed using matrix operations. (Linear Transformations). Moreover, in the optimization part of a machine learning algorithm, calculus plays a crucial role. Probability theory is very much useful in the decision-making step of machine learning algorithms.

So, mathematics acts as a building block for modern technology, machine learning. With the help of Mathematics, we can better understand the working behavior of Machine Learning. Knowledge of Mathematics helps us choose the appropriate or correct algorithm for a given data set and also to improve the accuracy of the machine learning algorithms.

Keywords—Machine learning; Linear Algebra; Calculus; Probability Theory

I. INTRODUCTION

In machine learning, some concepts from mathematics are really very important or helpful in the area of machine learning, deep learning, artificial intelligence, and so on.

What is machine learning?

Machine learning is one of the most important applications of artificial intelligence, due to which computers are enabled to learn automatically from given data and improve from past experience. As we learn from our past experience, we then take decisions based on that experience in the present or future. Similar to this in machine learning, a machine has the ability to make decisions in the present or future based on past experience.

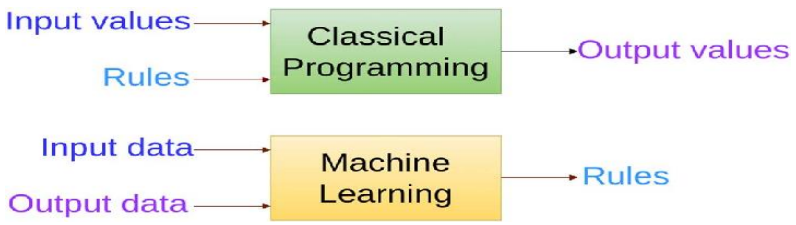

Fig. 1 Block diagram of traditional programming and Machine Learning

In traditional programming, we fed the program and input data to the computer and used to get output. However, machine learning is different from traditional programming. In machine learning, we provide input and output to the computer, and we will try to model our program.

Machine learning is divided into three different learning types depending upon their different learning processes and objectives, namely supervised, unsupervised, and reinforcement learning.

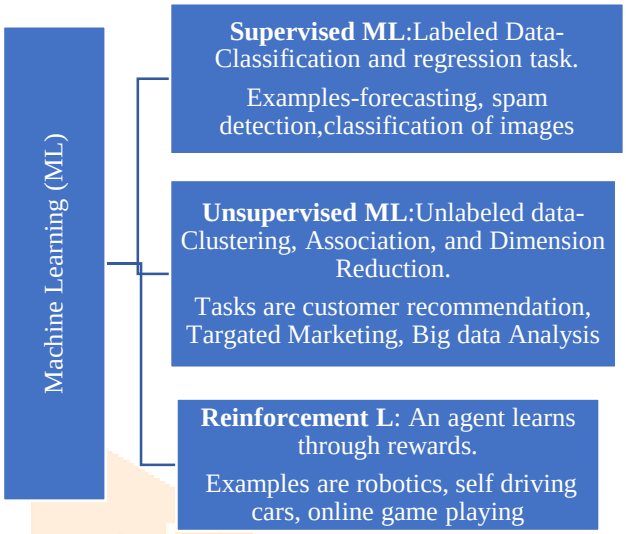

Fig. 2 Types of Machine Learning

Following given are different Machine Learning Algorithms

1. Linear Regression Algorithm
2. Logistic Regression Algorithm
3. Support Vector Machine
4. Support Vector Regression
5. Decision Tree
6. Random Forest Tree
7. K-Nearest Neighbour
8. Naïve Bayes
9. Neural Networks
10. Principal Component Analysis
11. Gaussian Mixture Models

Why Mathematics?

Linear algebra, matrix multiplication, basic probability theory, and calculus are roots of most of the algorithms from machine learning.

So, to understand the behavior of any machine learning tool, theoretically one must have sufficient knowledge of the above-listed topics.

OBJECTIVE

To explore how the mathematical concepts from linear algebra, calculus, and probability theory form the building blocks of machine learning algorithms by helping with data handling, optimization, and decision-making.

II. LITERARURE REVIEW

1. This paper is focused on how mathematical foundations are important in machine learning algorithms. Machine learning has four mathematical pillars, namely statistics, calculus, linear algebra, and probability. Mathematics is essential to discover the relations between parameters and decision variables. In machine learning, statistics is useful for the analysis of data, hypothesis testing, and inference. Bayes theorem is useful in probabilistic algorithms. Gradient descent is the key to minimizing the loss function.

Optimization techniques rely on multivariable calculus in deep learning. Yes, the conclusion of the above paper is that mathematical concepts are essential to construct accurate machine learning models.[1]

2. Various problems and tasks are in the world that are not easy to solve by computers or programming. So, by teaching computers how to learn and, through practice, improve every time to get better solutions. The paper gives an idea about mathematical concepts used to build machine learning models and their contribution to optimizing predictions. Linear algebra is required to handle large-sized data and to perform matrix operations. Calculus is essential in optimization techniques through gradient descent. Probability is a must in modelling uncertainty in the data. Statistics helps in the analysis of data, regression, and hypothesis testing. Also, an illustration of prediction by linear regression is given. Thus, the paper throws light on the role of mathematics in various machine learning techniques.[2]

3. Extensive coverage of tools for machine learning is done. The paper discusses various tools for machine learning, such as scikit-learn, TensorFlow, PyTorch, etc. Classification of tools is done depending on their applications, features, and functions. In paper supervised, unsupervised, and reinforcement learnings are explained through algorithms regression, classification, and clustering. An important part of the paper is that it highlights some industry use examples, such as Amazon Go, Google Microscope (detection of cancer cells), Tesla Autopilot (self-driving car), and Netflix and Spotify recommended systems.[3]

4. The paper gives an explanation of basic concepts used in machine learning and deep learning. The author discussed supervised, unsupervised, semi-supervised, and reinforcement learning. Outline of working of algorithms K-NN (K-Nearest Neighbour), DT (Decision Tree), and linear regression are given. The author discusses the application of machine learning in healthcare, social media, and aerospace.

In healthcare, machine learning is used for robotic surgery, medical diagnosis, and virtual nursing.

In social media, machine learning is helpful in content recommendations and advertisements.

In aerospace, machine learning is useful for pilot training and predictive maintenance.[4]

5.P. Manisha Kumar (2020) explains about detection of fake reviews using various algorithms including KNN, Naïve Bayes, Random Forest, Decision Tree, SVM and Logistic Regression of Machine Learning.

By reading fake review, customers declining the purchase of products. The system on Machine Learning built by author removes the fake reviews and junk from the reviews.

Out of six algorithms used, SVM classifier gives best accuracy.[5]

III. MATHEMATICAL CONCEPTS BEHIND MACHINE LEARNING

3.1 LINEAR ALGEBRA

Linear algebra is a branch of mathematics where we study vector space and matrices.

Vector is used to represent an n-dimensional point in mathematics. In machine learning, a vector is used to represent one data point with n different features.

A matrix is a rectangular arrangement of numbers in mathematics. Order of Mathematics is denoted by mxn, where m is the number of rows and n is the number of columns. In ML, a matrix is used to represent data or to handle data. Each row represents a data point with n number of features, and each column represents one feature of all points. Matrix A of order mxn represents data with m inputs, and each input has n features.

Linear transformation (L.T.) in linear algebra is in the form of a matrix and is very much useful in graphics and animation. By applying L.T. on a given object, we can change the shape of that object in a plane or in a space. Also, we can see from one viewing angle to another viewing angle. Using geometric transformation, we can rotate and scale the object within a space. L.T. can be used to project data from one space to another space, which is required if data is not linearly separable.

Eigenpairs is a term for eigenvalues and eigenvectors together. In most of the learning algorithms, the complexity depends on the dimension of input variables. Therefore, for the reduction of memory, we are interested in reducing the dimensionality of the input variables, preserving most of the information. We have two methods for reduction of dimensionality, namely feature selection and feature extraction.

In feature selection, our aim is to find k of the d dimensions that gives us most of the information, and we can eliminate the remaining d - k dimensions. In feature extraction, our aim is to find a new input set of k dimensions that is a combination of the original d dimensions. The best and widely used methods are Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Eigenvalues and eigenvectors are very much crucial in ML, as dimensional reduction algorithms like Principal Component Analysis and Linear Discriminant Analysis are totally based on them (eigenvalues and eigenvectors) only.

Application in Machine Learning:

1. Many times, data is represented in matrix form.
2. Matrix operations are used to train neural networks in deep learning.
3. Eigenvalues and Eigenvectors are used for dimensionality reduction in Principal Component Analysis.

3.2 PROBABILITY and STATISTICS

Statistics plays a vital role in technological advancements, particularly when working with data. As data forms the base, statistical concepts are essential for understanding, interpreting, and drawing conclusions from it. Consequently, statistics is a crucial component of machine learning.

Statistics is fundamentally the science of data collection, analysis, interpretation, presentation, and organization. It serves as a powerful tool for extracting insights from information.

Statistics plays a pivotal role in machine learning, enabling us to understand, develop, validate, and refine models used to analyze vast datasets. As machine learning relies on uncovering meaningful patterns and predictions within extensive data, statistics

provides the essential tools to navigate and make sense of this information. The steps involved in the formation of a machine learning algorithm are as follows:

Understanding Data → Selecting and Creating a Model → Validation of model

Preparation of Data:

Accurate data preparation is very effective in machine learning algorithm. Very often the given data is imperfect, unclear, or may be incomplete. Such data cannot be fed directly to the machine. So, we make use of statistical techniques to refine, analyse, and visual it.

-Handling missing entries (Here most of the time, we use central tendency)
-Use correlations to uncover relationships among features and the relationship between input features and output i.e., target variable.
-Removing redundant features that are dependent on others.
-Normalize data into standard normal variate. (Z-Variable)

By systematic application of above statistical methods, we can transform raw data into clean and organized format, which enables effective analysis and modelling of the data.

Selecting and Creating model:

Based on statistical insights of data, one can select and create a machine learning model. This model leverages statistical principles to derive a mathematical expression that represents the best relationship between input variables (features) and output variables (targets). Some useful statistical techniques to build models are regression, probability distribution, and multivariate statistical techniques.

For the building of different machine learning models of classification, the basic and central concept is the concept of probability theory from statistics. In many cases, we have to deal with uncertainties of the output due to lack of some information or due to other reasons. To overcome the issue of uncertainty, we make use of probability theory in various algorithms from classification, clustering, and reinforcement learning. In addition, Bayesian networks are powerful tools for visually representing dependencies of random variables within themselves. So, we can conclude that important factors of machine learning algorithms are Bayes theorem and probability distribution.

Validation of the model:

A very important and necessary step of each algorithm because the authenticity of the algorithm is calculated. To calculate it, techniques like cross-validation or R^2 score and mean squared error are used.

Application in Machine Learning: 1. In feature analysis, to discover significant variables that affect the outcome most.
2. Used in Bayesian classifiers (binary classifiers).
3. Use of Gaussian distribution in the uncertainty (Gaussian mixture models are used in clustering).
4. To estimate performance of model by using statistical significance test.

3.3 CALCULUS

A branch of mathematics where the study of continuous change with respect to time is done is called Calculus. We study, integration to calculate the total amount of quantity that accumulates over a time period. To study the change of the rate of a function with respect to time or independent variables, “derivative” from calculus is used. So, it is useful in the optimization of solutions given by machine learning algorithms.

Specifically:

The gradient descent algorithm is used to minimize the error between predicted and actual values of the target variable to increase the accuracy of the model. To optimize the distance between data points and clusters, we use distance formulas in various algorithms of clustering. It confirms effective grouping and segmentation.

A framework is provided by calculus to study the limitations and understanding bias-variance trade-offs in machine learning models. When the model is too simple and loses its ability to capture the complexity of data points, high bias occurs, and high variance is due to overfitting of the model on training data. The balance of both (bias and variance) is necessary to achieve accuracy and generalization of the model. Using calculus, we are able to understand the interaction between bias and variance to have balance in both.

Application in Machine Learning: 1. Minimization of loss function to improve accuracy of model.
2. Optimization techniques are used to maximize margin in Support vector machine.
3. Using gradient descent neural networks are trained (backpropagation in deep learning).

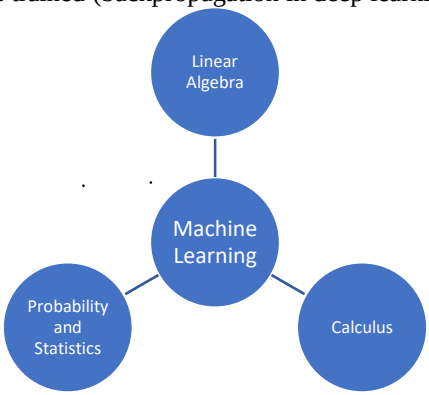

Fig.3 Mathematical concepts used in Machine Learning

IV. DIFFERENT ML TOOLS WITH MATHEMATICAL CONCEPTS

Behind each step of machine learning algorithms, mathematics is used. Thus, different machine learning algorithms are underpinned with different mathematical concepts.
Below is the complete breakdown of the mathematical dependencies for various machine learning algorithms.

Machine Learning Algorithm	Useful Mathematical Concept	Mathematical Equation	Area of Application
LR (Linear Regression Algorithm)	1.Matrix operation from Linear Algebra is used to get parameters of model. 2.For optimization gradient descent is used to minimize loss function.	Minimize loss function $J(\theta) = \frac{1}{2m}\sum_{1}^{m}(\widehat{y_i} - y_i)^2$ Updating by Gradient descent $\theta = \theta - \alpha\nabla J(\theta)$ And Matrix multiplication to update the prediction $\widehat{y_i} = x.\theta$, $\widehat{y_i}$ is predicted output, y_i is actual value, x is vector of inputs, θ is vector of parameters.	1.forecast of stock market 2. prediction of house rent 3.Trends based on historical data.
SVR (Support Vector Regression)	1.Linear Algebra is used to define decision boundary i.e. to calculate hyperplane. 2. Prediction of continuous value of target is done instead of classification	ε −insensitive Loss function $L_\varepsilon(y, \hat{y}) = max(0,\ \|y - \hat{y}\| - \varepsilon)$	1.Forecast of sales prediction, 2.Sentiment analysis, 3.Predictive maintenance.
SVM (Support Vector Machine)	1.Linear Algebra is used to define decision boundary and maximum margin between classes. 2.Optimization is used for maximum margin.	Hyperplane is defined by the equation $w^T x + b = 0$ Optimization is done for $min_w \frac{1}{2}\|w\|^2$ (Lagrange's multiplier method is used)	1.Face detection, 2. spam detection, 3.Disease diagnosis.
Machine Learning Algorithm	Useful Mathematical Concept	Mathematical Equation	Area of Application
DT (Decision Tree)	1.By using information theory calculation of information gain and entropy is done 2.Optimization	Entropy H(S) $= -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$ $p(x_i)$ is probability of outcome x_i Information Gain = Entropy H(S) $- \sum_{a\epsilon A} \frac{\|S_a\|}{\|S\|} * Entropy\ H(S_a)$, A is Attribute	1.Customer segmentation and 2. fraud detection
K-NN (K-Nearest Neighbor)	1.To calculate distance between two data points, distance metrics are used. 2. Prediction is done on proximity.	Euclidean distance formula $d(x, y) = \sqrt{(x_i - y_i)^2}$ Maximum vote for classification and mean in regression.	1.Recommendations to consumer 2. Medical diagnosis 3. Fraud detection
NB (Naïve Bayes)	Bayes theorem is used to calculate probability	$P(H\|E) = \frac{P(H) * P(E\|H)}{P(E)}$ P(H|E)= Posterior probability of H 'given the evidence; P(H)= Prior Probability; P(E|H)= Likelihood of the evidence E' if the Hypothesis 'H' is true; P(E)=Priori probability that the evidence itself is true	1. Speech recognition 2. Text classification 3. Medical diagnosis
GMM (Gaussian Mixture Model)	Use Gaussian distribution to model data For parameter estimation iterative optimization technique is used.	Formula for Gaussian distribution μ is mean $f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	1. image processing 2.clustering 3.portfolio analysis.

		From Byjus.com	
Logistic Regression	Matrix operations are used to get prediction. Gradient Descent is used for optimization of weights For classification Probability is used in logistic function	Hypothesis is given by $h_\theta(x) = g(\theta^T x)$ Sigmoid function is given by $g(z) = \frac{1}{1+e^{-z}}$ $J(\theta) = \frac{1}{m}\sum_{i=1}^{m}[-y^{(i)}\log\left(h_\theta(x^{(i)})\right) - (1-y^{(i)})\log(1-h_\theta(x^{(i)}))]$	1, Binary classification
NN (Neural Networks)	Using Matrix operations updating of weights is done in layers. Backpropagation to estimate updated weights.	Forward propagation $a^i = \sigma(W^i a^{i-1} + b^i)$ Backpropagation (GD) = $W^{i+1} = W^i - \eta\nabla W$ W is weight.	1.Natural Language processing and 2. image recognition.
PCA (Principal Component Analysis)	To compute eigenvalues and eigenvectors. Variance maximization across principal component.	Covariance Matrix = $\frac{1}{m}X^T X$ Decomposition of eigenvalues = $(X^T X)v = \boldsymbol{\lambda v}$ λ is eigen value and v is eigen vector.	1.Feature selection and 2.compression of image

V. APLLICATIONS OF MACHINE LEARNING

Machine Learning has plenty of applications in various field, bring about innovations and transforming industry in new domains. Following are some important applications of Machine Learning in different areas.

1. Healthcare: Mostly deep learning and neural networks techniques are broadly applied to disease diagnosis. ML techniques have been applied to discover similar patterns in images getting from CT scans, X-Rays, MRIs for diagnosis of diseases like cardiovascular disease, cancer, Alzheimer's.[6], [7] Also, ML algorithms predict autism spectrum disorder,[8] skin lesion detection,[9] depression,[10] , Predictive diseases [11].

2.Retail and E-commerce: ML algorithms are useful for building recommendation system which recommend products to consumer depending on their purchase history, and behavior. Helpful in prediction of creating requirements to improve supply chain optimization. Do analysis of feedback and review of customers to improve services. Virtual assistant and chatbots sufficiently interact and support to consumers.[5], [12], [13].

3. Financial: Fraud detection is done by identifying fake transaction. Chatbots and virtual assistant smartly handle the queries of consumers. ML techniques decide the creditworthiness depending on transaction pattern and credit history of loan borrower.[14], [15], [16]

4. Sports analysis: Predict performance of player, injury risks, outcome of matches depending on historical data. This information is used by players to decide the strategies of game.[17]

5. Agriculture: Based on quality of soil, recommendations of fertilizers, planting stratifies, and improving crop productivity is done. Also, crop yield prediction is done.

6. Cybersecurity: ML techniques are useful in spam filtering and secure system from malicious software.
7. Application of ML in Quantum computing**.**[18]

VI. CONCLUSION

From the topic that we discussed above, conclude that mathematics has a vital role in building and implementing the algorithms of machine learning. Also, we can say that mathematics and computer science together gave a powerful tool named "machine learning", which has many applications in real-world life all over the world.

REFERENCES

[1] K. Singh, "The Role of Mathematics in Artificial Intelligence and Machine Learning," *Int. J. Res. Publ. Semin.*, vol. 14, no. 5, pp. 186–197, 2023, doi: 10.36676/jrps.v14.i5.1434.

[2] Rajendra Kumar Vilas Thorat, "Role of Mathematics in Data Science - Machine learning," *Int. J. Sci. Res. Mod. Sci. Technol.*, vol. 3, no. 3, pp. 18–21, 2024, doi: 10.59828/ijsrmst.v3i3.191.

[3] M. R. Kishore, "A Review on Machine Learning Tools and Techniques," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 4270–4283, 2022, doi: 10.22214/ijraset.2022.44888.

[4] S. Reema Sree, S. B. Vyshnavi, and N. Jayapandian, "Real-World Application of Machine Learning and Deep Learning," *Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019*, no. Icssit, pp. 1069–1073, 2019, doi: 10.1109/ICSSIT46314.2019.8987844.

[5] P. M. Kumar, S. S. Harrsha, K. Abhiram, M. Kavitha, and M. Kalyani, "Role of Machine Learning in Fake Review Detection," *6th Int. Conf. Electron. Commun. Aerosp. Technol. ICECA 2022 - Proc.*, no. Iceca, pp. 1026–1030, 2022, doi: 10.1109/ICECA55336.2022.10009174.

[6] S. C. Mana, G. Kalaiarasi, R. Yogitha, L. S. Helen, and R. Senthamil Selvi, "Application of Machine Learning in Healthcare: An Analysis," *3rd Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2022 - Proc.*, no. Icesc, pp. 1611–1615, 2022, doi: 10.1109/ICESC54411.2022.9885296.

[7] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.

[8] S. Karim, N. Akter, M. J. A. Patwary, and M. R. Islam, "A Review on Predicting Autism Spectrum Disorder(ASD) meltdown using Machine Learning Algorithms," *2021 5th Int. Conf. Electr. Eng. Inf. Commun. Technol. ICEEICT 2021*, pp. 1–6, 2021, doi: 10.1109/ICEEICT53905.2021.9667827.

[9] R. Yadav and A. Bhat, "A Survey on Skin Lesion Detection and Classification using Machine Learning," *2nd Int. Conf. Artif. Intell. Mach. Learn. Appl. Healthc. Internet Things, AIMLA 2024*, pp. 1–5, 2024, doi: 10.1109/AIMLA59606.2024.10531571.

[10] A. M. Putri, K. Wijaya, O. A. Salomo, A. A. Santoso Gunawan, and Anderies, "A Review Paper: Accuracy of Machine Learning for Depression Detection in Social Media," *Proceeding - IEEE Int. Conf. Commun. Networks Satell. COMNETSAT 2022*, pp. 39–45, 2022, doi: 10.1109/COMNETSAT56033.2022.9994553.

[11] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey," *J. Electr. Syst. Inf. Technol.*, vol. 10, no. 1, 2023, doi: 10.1186/s43067-023-00108-y.

[12] A. M. Khedr and S. R. S, "Enhancing supply chain management with deep learning and machine learning techniques: A review," *J. Open Innov. Technol. Mark. Complex.*, vol. 10, no. 4, p. 100379, 2024, doi: 10.1016/j.joitmc.2024.100379.

[13] H. Soltani and P. Bhandari, "The Use of Machine Learning in Supply Chain Management, A Systematic Review," no. October 2024, pp. 21–49, 2023, doi: 10.46254/an13.20230529.

[14] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," *IEEE Access*, vol. 11, no. December 2022, pp. 3034–3043, 2023, doi: 10.1109/ACCESS.2022.3232287.

[15] E. Pan, "Machine Learning in Financial Transaction Fraud Detection and Prevention," *Trans. Econ. Bus. Manag. Res.*, vol. 5, pp. 243–249, 2024, doi: 10.62051/16r3aa10.

[16] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, "Financial fraud detection through the application of machine learning techniques: a literature review," *Humanit. Soc. Sci. Commun.*, vol. 11, no. 1, pp. 1–22, 2024, doi: 10.1057/s41599-024-03606-0.

[17] R. P. Bunker and F. Thabtah, "A machine learning framework for sport result prediction," *Appl. Comput. Informatics*, vol. 15, no. 1, pp. 27–33, 2019, doi: 10.1016/j.aci.2017.09.005.

[18] S. B. Ramezani, A. Sommers, H. K. Manchukonda, S. Rahimi, and A. Amirlatifi, "Machine Learning Algorithms in Quantum Computing: A Survey," *Proc. Int. Jt. Conf. Neural Networks*, no. 2, 2020, doi: 10.1109/IJCNN48605.2020.9207714.

[19] V. Nasteski, "An overview of the supervised machine learning methods," *Horizons.B*, vol. 4, no. December, pp. 51–62, 2017, doi: 10.20544/horizons.b.04.1.17.p05.

[20] R. Nath Mohalder, M. Alam Hossain, and N. Hossain, "Classifying the Supervised Machine Learning and Comparing the Performances of the Algorithms," *Int. J. Adv. Res.*, vol. 12, no. 01, pp. 422–438, 2024, doi: 10.21474/ijar01/18138.

[21] V. Gupta, V. K. Mishra, P. Singhal, and A. Kumar, "An Overview of Supervised Machine Learning Algorithm," *Proc. 2022 11th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2022*, no. December, pp. 87–92, 2022, doi: 10.1109/SMART55829.2022.10047618.