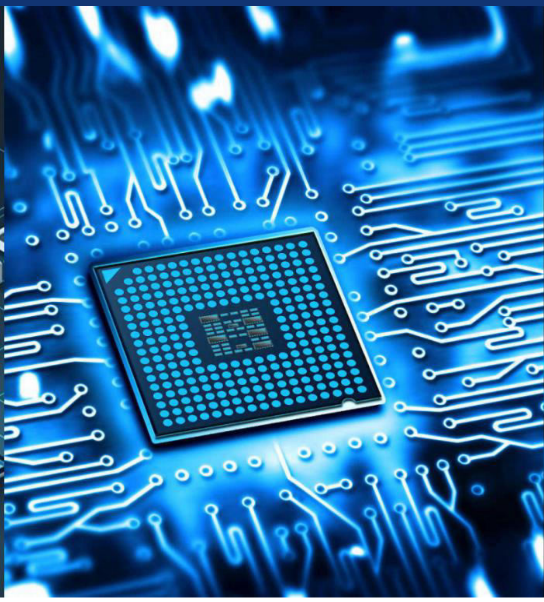
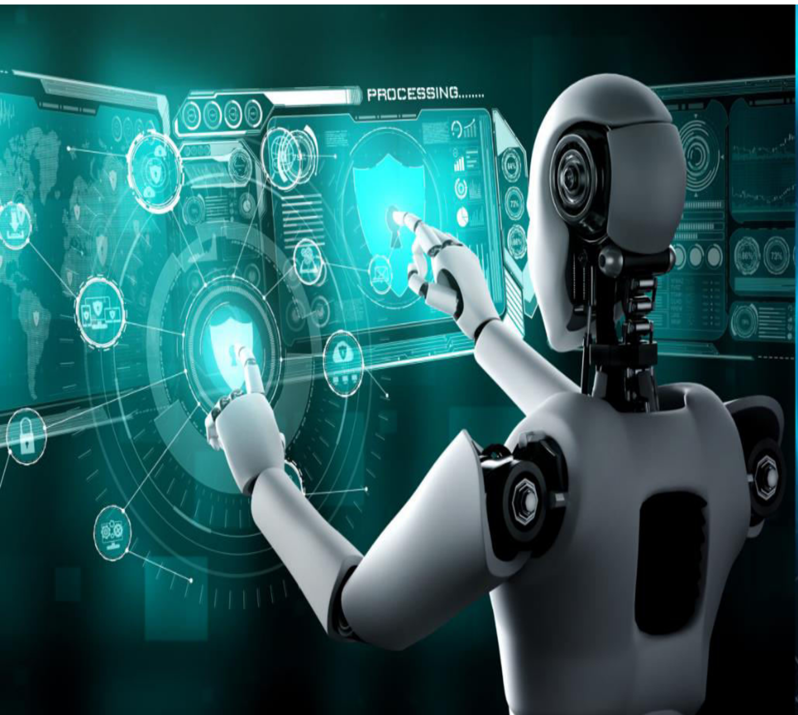
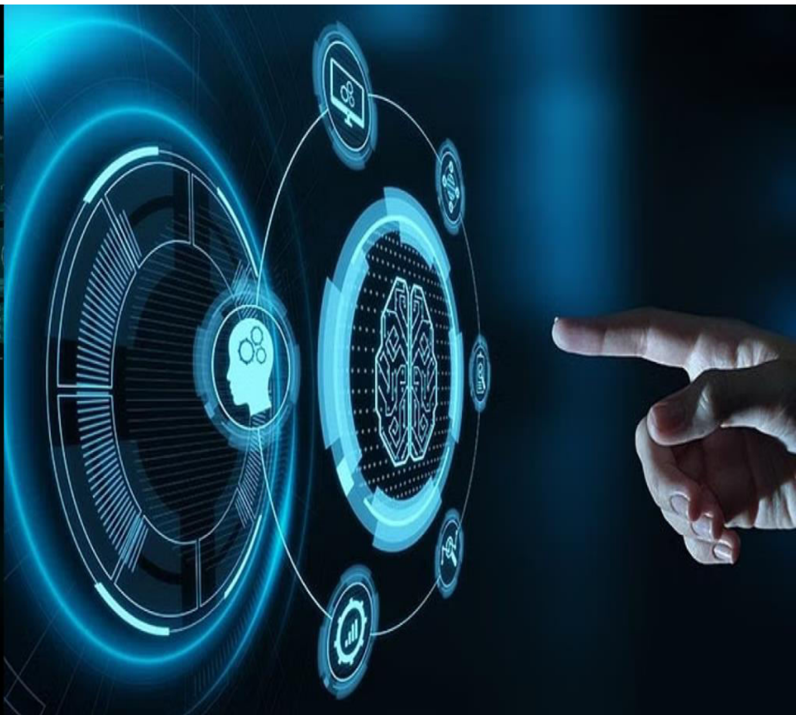
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 14, Issue 3, March 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Machine Learning for Sustainable Water Resource Management: A Comprehensive Review

Savita[1], Renu Shukla[1], ManMohan Singh Rawat[2]

Gurukula Kangri (Deemed to be University), Haridwar, Uttarakhand, India[1]

Uttarakahnd State Council for Science & Technology, Dehradun, India[2]

ABSTRACT: Recently, machine learning (ML) techniques have gained significant attention for their ability to analyze large datasets, identify hidden patterns, and support intelligent decision-making. In this context, modern strategies for managing water resources based on machine learning offer an effective, accurate, and sustainable solution. This paper provides a comprehensive review of ML algorithms and their applications in water resource management. In this paper, authors discuss in detail the theoretical background of ML algorithms, including methodology and performance metrics. The application section details various uses of ML in water resource management, including water demand forecasting, smart irrigation, leak detection, and water quality prediction. Furthermore, the paper discusses the current challenges, research gaps, and future prospects of ML-driven water conservation systems.

KEYWORDS: Water, Machine Learning, Natural resource, Water conservation, Survey, Environment

I. INTRODUCTION

Water is considered the foundation of life. Ancient civilizations always developed around rivers and water sources. Currently, the pressure on water resources has reached unprecedented levels. According to the United Nations, nearly half the world's population will be affected by water scarcity in the coming decades (UNESCO, 2023). This problem is even more serious in an agricultural country like India, where excessive dependence on rainfall, excessive groundwater exploitation, and pollution contribute to the water crisis (Vanham et al., 2021).

With technological advancements, innovations are also being seen in the field of water resource management. ML and artificial intelligence (AI), in particular, have played a key role in making water management data-driven and predictive. Water is a vital natural resource that supports life, economic development, and ecosystem sustainability. However, increasing population, urbanization, climate variability, and overexploitation of water resources have intensified water scarcity worldwide. According to global water assessments, freshwater demand is expected to rise significantly in the coming decades, particularly in developing countries (Vanham et al., 2021). Conventional water management and conservation practices rely heavily on historical data and static models, which often fail to adapt to rapidly changing environmental and usage conditions. ML offers a data-driven alternative by enabling adaptive, predictive, and intelligent water management systems capable of addressing complex and nonlinear relationships.

Water management aims to ensure the sustainable use and management of water resources. Major challenges include groundwater depletion, water losses due to leakage, inefficient irrigation, and deterioration of water quality. Climate change has further aggravated these challenges by altering rainfall patterns and increasing the frequency of extreme weather events. These issues highlight the need for advanced, adaptive, and intelligent conservation strategies. ML techniques can play a crucial role by providing accurate predictions, real-time monitoring, and optimized control mechanisms.

The remaining parts of the paper are organised the following way: Section 2 provides a concept and importance of water resource management. Section 3 details the theoretical background, methodology of ML, and performance metrics. Section 4 wrapped up the applications of ML in water resource management. Section 5 presents challenges and research gaps. Section 6 concluded the article.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table 1: Comprehensive List of Abbreviations used in the Study

Sno.		Abbreviations	Sno.		Abbreviations
1.	Artificial Neural Networks	ANNs	2.	K-Nearest Neighbors	KNN
3.	Support Vector Machine	SVM	4.	Deep Q-Network	DQN
5.	Decision Tree	DT	6.	Principal Component Analysis	PCA
7.	Naïve Bayes	NB	8.	Temporal Difference	TD
9.	Gradient Boosting	GBoost	10.	Support Vector Regression	SVR
11.	Gaussian Mixture Models	GMM	12.	Regression Tree	RT
13.	Deep Deterministic Policy Gradient	DDPG	14.	Random Forest	RF
15.	eXtreme Gradient Boosting	XGBoost	16.	Adaptive Boosting	AdaBoost
17.	Linear Discriminant Analysis	LDA	18.	Non-negative Matrix Factorization	NMF

III. CONCEPT AND IMPORTANCE OF WATER RESOURCE MANAGEMENT

Water management means the judicious use, conservation, and reuse of water resources. Its goal is to preserve water resources for future generations while meeting current needs. Water conservation refers to the balanced, judicious, and sustainable use of water resources, ensuring adequate water availability for future generations while meeting current needs. This includes preventing water wastage, efficiently managing available water, and conserving and recharging natural water sources. In the modern context, water management is not limited to water conservation; it is a holistic management process that encompasses aspects such as water availability, quality, distribution, use efficiency, and recycling. The concept of water management is based on the principles of sustainable development, where environmental, social, and economic balances are prioritized.

- **Environmental Importance**

Water conservation plays a vital role in preserving natural ecosystems. Rivers, lakes, wetlands, and groundwater systems are essential for biodiversity. Overexploitation of water leads to degradation of these ecosystems, while effective water conservation can maintain ecological balance.

- **Social Importance**

Water is a fundamental necessity for human life. Access to safe drinking water is directly linked to health, sanitation, and quality of life. Water management ensures equitable and equitable access to water in rural and urban areas, thereby reducing water-related social conflicts.

- **Economic Importance**

Key economic sectors such as agriculture, industry, and energy production are water-dependent. Water scarcity adversely impacts food production, industrial development, and employment. Water conservation supports economic stability and long-term development by increasing water use efficiency.

- **Importance in the Context of Climate Change**

Climate change is increasing uncertainty in rainfall patterns, droughts, and floods. Water conservation strategies, such as rainwater harvesting and groundwater recharge, help mitigate the adverse effects of climate change.

- **Importance for Future Generations**

Water resources are limited, and their excessive use impacts the rights of future generations. Water conservation ensures intergenerational equity, ensuring that future generations can access adequate and safe water.

Water management is not only a technical necessity; more than that, it is a social and moral responsibility. Given the current global water crisis, understanding the concept and importance of water management is crucial. Effective water management not only promotes environmental sustainability but also provides a strong foundation for social well-being and economic development.

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403075

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. MODERN WATER MANAGEMENT STRATEGIES: MACHINE LEARNING

India has a rich traditional water management system. Examples include ponds, wells, stepwells, johads, and the ahar-pain system. These systems were developed to suit local geographical and climatic conditions. However, in modern times, the neglect of these systems is exacerbating the water crisis. At present, due to water crisis this topic is again being considered important and is being considered seriously again. Technology plays a vital role in modern water conservation. Technologies such as remote sensing, geographic information systems (GIS), the Internet of Things (IoT), and big data analytics have made water management more accurate and efficient. ML is at the core of all these technologies. In this section we explain the basis of ML and their performance metrics.

3.1 ML: Theoretical Background

ML is a branch of AI in which computer systems learn from data to recognize patterns and make predictions. ML focuses on enabling systems to learn from data without explicit programming. ML generally categorizes into supervised, unsupervised, semi-supervised, and reinforcement learning. Supervised learning techniques such as LR, SVM, DT, and RF are widely used for prediction and classification tasks in water management. Unsupervised learning methods, including K-means clustering and anomaly detection algorithms, are applied for pattern discovery and leak detection. Deep learning models such as ANN and LSTM networks are particularly effective for time-series forecasting. ML is a branch of AI in which computer systems learn from data to recognize patterns and make predictions.

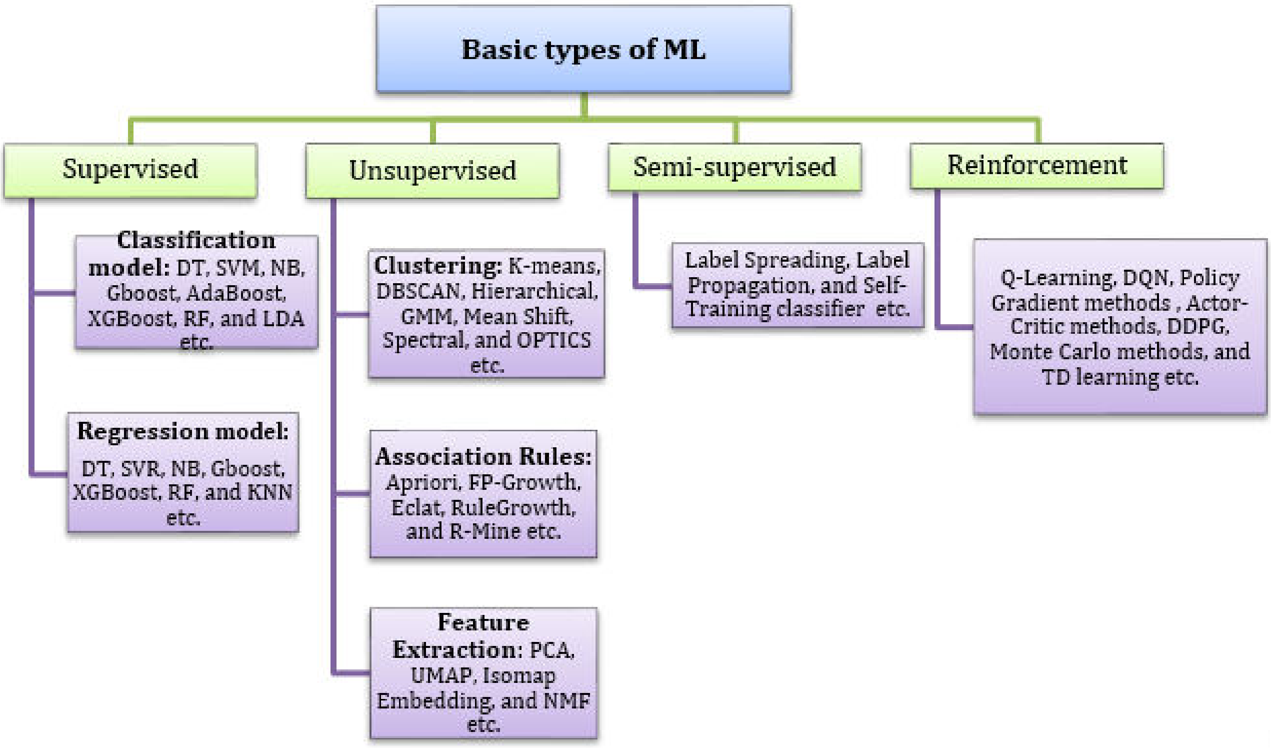


Figure 1: Overview of the Main Types of Machine Learning models

Its detailed methodology can be broadly divided into four categories and their specific steps (Fong 2010):

• **Supervised Learning:**

This method uses labeled data, where the input data is already associated with the corresponding output. It works through classification (e.g., flood prediction) and regression (e.g., accurate water demand estimation). A classification model is used for nominal dependent variables, and a regression model is used for continuous dependent variables. Classification models are DT (Breiman et al., 1984; Quinlan, 1986), SVM (Cortes and Vapnik, 1995; Steinwart and Christmann, 2008), NB (Domingos and Pazzani, 1997), AdaBoost (Freund and Schapire, 1995), GBoost (Friedman,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2001; Hastie et al., 2009), XGBoost (Chen and Guestrin, 2016), RF (Breiman, 2001), and LDA (Friedman, 1989). Regression models incorporate the regression forms of the classification algorithms mentioned earlier, namely DT (Breiman et al., 1984), SVR (Drucker et al., 1997), NB regressors (Frank et al., 2000), AdaBoost regressors (Freund and Schapire, 1995), GBoost regressors (Friedman, 2001; Hastie et al., 2009), XGBoost regressors (Chen and Guestrin, 2016), RF regressors (Breiman, 2001), and KNN regressors (Cover and Hart, 1967).

• Unsupervised Learning:

It works on unlabeled data and discovers hidden patterns or groups in it without human intervention. Its main technique is clustering, which organizes data points with similar characteristics into groups (e.g., identifying pollution sources) (Hsu et al., 2019). It uses similarity metrics to maximize similarity within groups and to clarify differences between different groups. Some basic clustering algorithms are K-means (Kodinariya & Makwana, 2013), DBSCAN (Kumar & Reddy, 2016), Hierarchical (Zhao & Karypis, 2002), GMM (Liu et al., 2010), Mean Shift (Carreira-Perpiñán, 2015), Spectral (von Luxburg, 2007), OPTICS (Deng et al., 2014), agglomerative (Müllner, 2011), density-based (Bhattacharjee & Mitra, 2020), etc. Another technique is association, which is based on finding relationships and associations between variables in the dataset (Kotsiantis & Kanellopoulos, 2006). Some important algorithms used in water management are Apriori (Al-Maolegi & Arkok, 2014), FP-Growth (Said, 2009), Eclat (Girotra et al., 2013), CAR-SPAN (Mooney & Roddick, 2013), FPMax (Miani & Junior, 2018), RuleGrowth (Fournier-Viger et al., 2011), R-Mine (Liu et al., 1999), Tertius (Baher & Lobo, 2012), etc.

• Semi-supervised Learning:

It is a hybrid of supervised and unsupervised learning (Chapelle et al., 2006). It uses a small amount of labeled data and a large amount of unlabeled data for training, which reduces costs and improves model understanding. Label Spreading (Zhu & Goldberg, 2009), Label Propagation (Zhu & Ghahramani, 2002), Self-Training classifier (Triguero et al., 2015), etc. are some well-known semi-supervised learning algorithms.

• Reinforcement Learning:

It is based on the principle of reward and punishment. In this, an 'agent' interacts with its environment and learns to make decisions through trial and error (learning from mistakes) to maximize profit (Kaelbling et al., 1996). It is used in complex tasks such as reservoir operations. Q-Learning (Van Hasselt et al., 2016), Deep Q-network (DQN) (François-Lavet et al., 2016), Policy Gradient methods (Sutton et al., 1999), Actor-Critic methods (Kumar et al., 2023), Proximal policy optimization (PPO) (Schulman et al., 2017), Deep deterministic policy gradient (DDPG) (Yoo et al., 2020), Monte Carlo methods (Lazaric et al., 2007), Temporal difference (TD) learning (Taylor et al., 2006), Asynchronous advantage actor–critic (A3C) (Babaeizadeh et al., 2017), Soft actor–critic (SAC) (Haarnoja et al., 2019), Twin delayed deep deterministic policy gradient (TD3) (Kim et al., 2020), etc. are Reinforcement Learning algorithms.

3.2 Methodology of ML

According to sources, the methodology of ML is based on the process of learning from data and improving itself based on experiences. The effectiveness of ML-based water conservation systems largely depends on the availability and quality of data. Common data sources include in-situ sensors, smart water meters, meteorological stations, satellite imagery, and water quality monitoring systems. Data preprocessing steps such as noise removal, missing data handling, normalization, and feature selection are essential to enhance model performance and reliability. Integration of heterogeneous data sources remains a key research challenge.

There are five main steps for creating a classification model (Han et al., 2012):

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403075

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

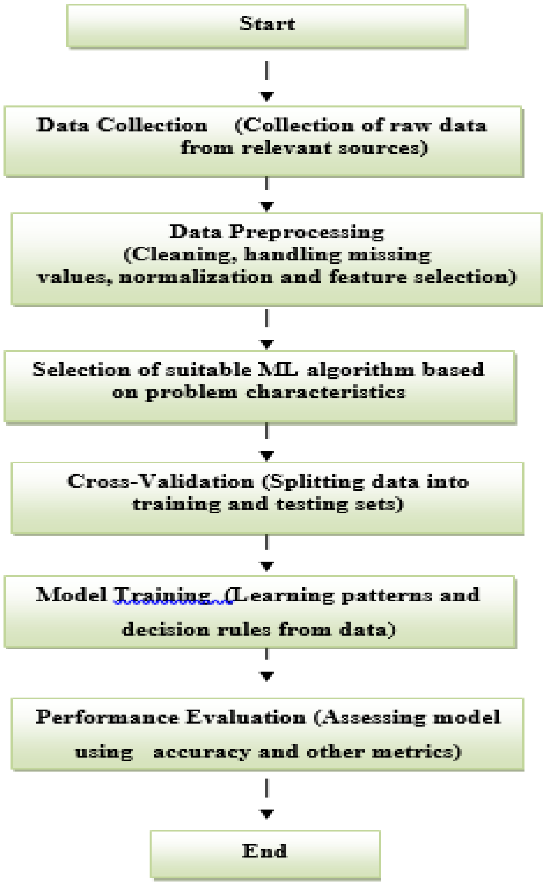


Figure 2: five main steps for creating a classification model

In short, ML is a set of algorithms and statistical models for analyzing data, learning from it, and making complex decisions, reducing human intervention in areas such as water resource management.

3.3 Performance metrics

Table 2 represents summary of commonly used evaluation metrics.

Table 2: Performance metrics ((Ghobadi & Kang, 2022a), (Ghobadi & Kang, 2022b))

	Metric.	Formula's	Description
Error-based metrics	Mean normalized bias error	$MNBE = \frac{1}{x}\sum_{i=1}^{x}\frac{p_i - v_i}{v_i}$	Average normalized bias of the predictions.
	Mean percentage error	$MPE = \frac{100}{x}\sum_{i=1}^{x}\frac{p_i - v_i}{v_i}$	Mean signed percentage deviation.
	Mean absolute error	$MAE = \frac{1}{x}\sum_{i=1}^{x}\lvert p_i - v_i\rvert$	Average magnitude of prediction errors.
	Mean absolute percentage error	$MAPE = \frac{100}{x}\sum_{i=1}^{x}\left\lvert\frac{p_i - v_i}{v_i}\right\rvert$	Average absolute error expressed in percentage terms.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

	Relative absolute error	$RAE = \frac{\sum_{i=1}^{x}\lvert p_i - v_i\rvert}{\sum_{i=1}^{x}\lvert p_i - \overline{v_i}\rvert}$	Compares model error with a mean-based baseline.
	Weighted mean absolute percentage error	$WMAPE = \frac{\sum_{i=1}^{x}\lvert p_i - v_i\rvert}{\sum_{i=1}^{x}\lvert v_i\rvert} X100$	MAPE with magnitude-based weighting.
	Normalized mean absolute error	$NMAE = \frac{MAE}{v_{max} - v_{min}}$	Scale-independent form of MAE.
Squared-error metrics	Mean squared error	$MSE = \frac{1}{x}\sum_{i=1}^{x}(p_i - v_i)^2$	Penalizes large deviations more strongly.
	Root mean square error	$RMSE = \sqrt{MSE}$	Represents typical error magnitude in original units.
	Normalized root mean square	$\text{NRMSE} = \frac{RMSE}{v_{max} - v_{min}}$	Allows comparison across different scales.
Variability and goodness-of-fit	Error Coefficient of variation	$CV = \sigma / \bar{v}$	Measures relative dispersion of observations.
	Coefficient of determination	$R^2 = 1 - \frac{\sum(p_i - v_i)^2}{\sum(p_i - \overline{v_i})^2}$	Fraction of variance explained by the model.
	Willmott's index agreement	$d = 1 - \frac{\sum(p_i - v_i)^2}{\sum(\lvert p_i - \overline{v_i}\rvert + \lvert v_i - \overline{v_i}\rvert)^2}$	Measures agreement relative to potential error.
	Legates–McCabe's	$LM = 1 - \frac{\sum\lvert p_i - v_i\rvert}{\sum\lvert v_i - \overline{v_i}\rvert}$	Absolute-error-based goodness-of-fit measure.
	Kling–Gupta efficiency	$KGE = 1 - \sqrt{(r-1)^2 + (a-1)^2 + (b-1)^2}$	Assesses model efficiency by considering accuracy, precision, and consistency components.
Information criterion	Akaike information criterion	$AIC = 2k - 2\ln(l)$	Evaluates model performance while considering model complexity. k = number of parameters.
Probabilistic and interval metrics	Continuous ranked probability score	$CRPS = \int_{-\infty}^{\infty}(F(p_i) - H(p_i - v_i))^2 dp_i$	Quantifies the quadratic difference between the forecasted and empirical cumulative distribution functions (CDF).
	Average width of the prediction intervals	$AWPI = \frac{1}{x}\sum_{I=1}^{x} p_i^u - p_i^l$	Mean interval width at a given confidence level. The upper and lower bounds of the 95% prediction interval are denoted by u and l, respectively.
	Prediction interval coverage	$PICP = \frac{1}{x}\sum_{I=1}^{x} c_{i,}$ $c_i = \begin{cases} 1, if\ v_i \in [p_i^l, p_i^u] \\ 0, if\ v_i \nexists [p_i^l, p_i^u] \end{cases}$	The proportion of instances in a holdout set for which the prediction interval successfully captures the actual value.
	Prediction interval normalized average width	$PINAW = \frac{1}{x}\sum_{I=1}^{x}(p_i^u - p_i^l)/R$	This metric quantifies the width or extent of the prediction interval. The range of variation of the observed value (R) is used to determine the width of the interval.
Notation Used	p_i= predicted value, v_i= observed value, $\overline{v_i}$ =mean of observed values x = number of samples. r =Correlation coefficient, $a = \frac{\sigma_p}{\sigma_v}$ (variability ratio), $b = \frac{\mu_p}{\mu_v}$ (bias ratio), $l = likelihood$, σ =representing the standard deviation.		

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. APPLICATIONS OF ML IN WATER RESOURCE MANAGEMENT

This section discussed different ML applications in water management.

- **Machine Learning in Water Demand Forecasting**

Accurate water demand forecasting is a crucial component of water management. Machine learning models analyze historical consumption data, population growth, industrial expansion, and weather data to predict future water demand (Ahmed et al., 2020).

- **Smart Irrigation Systems**

The agricultural sector is the world's largest consumer of water. Machine learning-based smart irrigation systems control irrigation based on soil moisture, temperature, crop type, and weather forecasts. This saves water and increases crop productivity (Doe & Smith, 2021).

- **Water Quality Monitoring**

Water quality monitoring is essential to ensure the availability of clean water. Machine learning models assess water quality by analyzing pH, turbidity, dissolved oxygen, and pollutant levels (Zhang et al., 2020).

- **Machine Learning for Water Demand Forecasting**

Accurate water demand forecasting is critical for efficient planning and management of water supply systems. Machine learning models such as RF, SVR, and LSTM networks have demonstrated superior performance compared to traditional statistical methods. These models can capture nonlinear relationships between water demand and influencing factors such as population growth, climate variables, seasonal trends, and consumer behavior. Short-term and long-term demand forecasts support proactive decision-making and resource allocation.

- **Leak Detection and Non-Revenue Water Reduction** (Kumar et al., 2020)

Leakage in water distribution networks leads to significant water losses, commonly referred to as non-revenue water. ML-based leak detection techniques utilize pressure, flow, and acoustic data to identify anomalies indicative of leaks. Algorithms such as Isolation Forest, SVM, and deep autoencoders have shown promising results in detecting leaks with high accuracy. Early detection enables timely maintenance and reduces operational costs.

- **Smart Irrigation and Agricultural Water Management**

Agriculture accounts for the largest share of global freshwater consumption. Machine learning–based smart irrigation systems optimize water usage by predicting soil moisture, crop water requirements, and weather conditions. DTs, RFs, and ANNs are commonly employed to determine optimal irrigation schedules. These systems not only conserve water but also enhance crop yield and energy efficiency.

- **Water Quality Monitoring and Prediction**

Maintaining water quality is essential for human health and ecosystem stability. Machine learning models are widely used to predict water quality parameters such as pH, turbidity, dissolved oxygen, nutrients, and chlorophyll-a concentration. Deep learning models, particularly LSTM and CNN, have shown high accuracy in time-series and spatial water quality prediction. Such models support early warning systems for pollution and eutrophication.

- **Integration of ML with IoT and Smart Cities**

The integration of machine learning with IoT technologies enables real-time monitoring, analysis, and control of water systems. Smart meters, sensors, and automated valves generate continuous data streams that can be processed using ML algorithms. In smart city environments, ML-driven decision-support systems enhance operational efficiency, reduce water losses, and improve service reliability (Gupta et al., 2021).

V. CHALLENGES AND RESEARCH GAPS

Despite significant advancements, several challenges hinder the large-scale adoption of ML-based water conservation systems. These include data scarcity, lack of standardized datasets, model interpretability issues, high implementation costs, and cybersecurity concerns. Addressing these challenges requires interdisciplinary research and collaboration between engineers, data scientists, and policymakers. High initial cost

In the future, machine learning may be integrated with digital twins, blockchain, and advanced sensor technologies. Data-driven decision-making will play a crucial role in policymaking. In general terms, data availability is a major challenge for ML approaches in this field. The quality of the data determines the predictions of the ML algorithm.

VI. CONCLUSION

Machine learning has emerged as a transformative technology for water conservation. By enabling intelligent prediction, monitoring, and optimization, ML-based systems offer effective solutions to contemporary water challenges. Their integration with IoT and smart infrastructure will play a crucial role in achieving sustainable water management and global development goals. This research demonstrates that machine learning-based water conservation strategies can play a crucial role in addressing the water crisis. Widespread adoption of these technologies is essential for achieving the Sustainable Development Goals. Future research directions include the development of hybrid AI models, integration of satellite remote sensing data, edge computing–based ML frameworks, and explainable artificial intelligence (XAI) for transparent decision-making. The incorporation of climate change projections into ML models will further enhance the resilience of water management systems.

Funding: This research was funded by Uttarakhand Council for Science and Technology (UCOST), Dehradun, India, grant number UCS&T/R&D-06/23- 24/24531.

Informed Consent Statement: Not applicable.
Conflicts of Interest: The authors declare no conflicts of interest.

REFERENCES

1. Ahmed, M. S., et al. (2020). Machine learning methods for water demand forecasting. *Water Resources Management, 34*(2), 1–15.
2. Al-Maolegi, M., & Arkok, B. (2014). An improved Apriori algorithm for association rules. *International Journal of Natural Language Computing, 3*, 21–29.
3. Babaeizadeh, M., Frosio, I., Tyree, S., Clemons, J., & Kautz, J. (2017). *Reinforcement learning through asynchronous advantage actor-critic on a GPU*. arXiv. https://arxiv.org/abs/1611.06256
4. Baher, S., & Lobo, L. M. (2012). A comparative study of association rule algorithms for course recommender system in e-learning. *International Journal of Computer Applications, 39*, 48–52.
5. Bhattacharjee, P., & Mitra, P. (2020). A survey of density based clustering algorithms. *Frontiers of Computer Science, 15*, 151308.
6. Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
7. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Inc, Belmont California.
8. Carreira-Perpiñán, M. Á. (2015). *A review of mean-shift algorithms for clustering*. arXiv.
9. Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi-supervised learning*. MIT Press.
10. Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. New York, NY, USA.
11. Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.
12. Cover, T.M., Hart, P.E., 1967. Nearest neighbor pattern classification. IEEE Trans. Inf. Theor. 13, 21–27.
13. Deng, Z., Hu, Y., Zhu, M., Huang, X., & Du, B. (2014). A scalable and fast OPTICS for clustering trajectory big data. *Cluster Computing, 18*, 549–562.
14. Doe, J., & Smith, A. (2021). Smart irrigation systems using IoT and AI. *IEEE Access, 8*, 12345–12358
15. Domingos, P., Pazzani, M., 1997. On the optimality of the simple Bayesian classifer under zero-one loss. Mach. Learn. 29, 103–137.
16. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. In: Mozer, M.C., Jordan, M.I., Petsche, T. (Eds.), Advances in Neural Information Processing Systems 9. MIT Press, Cambridge, MA, pp. 155–161.
17. Fong, A. (2010). Welcome message from the editor-in-chief. *Journal of Advances in Information Technology, 1*(1).

18. Fournier-Viger, P., Nkambou, R., & Tseng, V. S. M. (2011). RuleGrowth: Mining sequential rules common to several sequences by pattern-growth. In *Proceedings of the ACM Symposium on Applied Computing* (pp. 956–961).
19. François-Lavet, V., Fonteneau, R., & Ernst, D. (2016). *How to discount deep reinforcement learning: Towards new dynamic strategies*. arXiv. https://arxiv.org/abs/1512.02011
20. Frank, E., Trigg, L., Holmes, G., Witten, I.H., 2000. Naive bayes for regression. Mach. Learn. 41, 5–25.
21. Freund, Y., Schapire, R.E., 1995. A Decision-Theoretic Generalization of On-Line Learning and Application to Boosting. Lecture Notes in Computer Science. Springer berlin Heidelberg, pp. 23–37.
22. Friedman, J.H., 1989. Regularized discriminant analysis. J. Am. Stat. Assoc. 84, 165–175.
23. Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.
24. Ghobadi, F., & Kang, D. (2022a). Improving long-term streamflow prediction in a poorly gauged basin using geo-spatiotemporal mesoscale data and attention-based deep learning: A comparative study. *Journal of Hydrology, 615*, 128542. https://doi.org/10.1016/j.jhydrol.2022.128542.
25. Ghobadi, F., & Kang, D. (2022b). Multi-step ahead probabilistic forecasting of daily streamflow using Bayesian deep learning: A multiple case study. *Water, 14*(22), 3672. https://doi.org/10.3390/w14223672
26. Girotra, M., Nagpal, K., Minocha, S., & Sharma, N. (2013). Comparative survey on association rule mining algorithms. *International Journal of Computer Applications, 84*, 18–22.
27. Gupta, A., et al. (2021). AI-based smart water management systems. *Sustainable Cities and Society, 67*.
28. Haarnoja, T., Zhou, A., Hartikainen, K., Tucker, G., Ha, S., Tan, J., Kumar, V., Zhu, H., Gupta, A., & Abbeel, P. (2019). *Soft actor-critic algorithms and applications*. arXiv. https://arxiv.org/abs/1812.05905
29. Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
30. Hastie, T., Tibshirani, R., Friedman, J.H., 2009. Chapter 10. Boosting and additive tress. The Elements of Statistical Learning, second ed. Springer, New York, pp. 337–384.
31. Hsu, K., Levine, S., & Finn, C. (2019). Unsupervised learning via meta-learning. arXiv. https://arxiv.org/abs/1906.05426
32. Kim, M., Han, D. K., Park, J. H., & Kim, J. S. (2020). Motion planning of robot manipulators for a smoother path using a twin delayed deep deterministic policy gradient with hindsight experience replay. *Applied Sciences, 10*(2), 575.
33. Kodinariya, T., & Makwana, P. (2013). Review on determining of cluster in K-means clustering. *International Journal of Advanced Research in Computer Science and Management Studies, 1*, 90–95.
34. Kotsiantis, S., & Kanellopoulos, D. (2006). Association rules mining: A recent overview. *GESTS International Transactions on Computer Science and Engineering, 32*, 71–82.
35. Kumar, H., Koppel, A., & Ribeiro, A. (2023). On the sample complexity of actor-critic method for reinforcement learning with function approximation. *Machine Learning, 112*, 2433–2467.
36. Kumar, K. M., & Reddy, A. R. M. (2016). A fast DBSCAN clustering algorithm by accelerating neighbor searching using groups method. *Pattern Recognition, 58*, 39–48.
37. Kumar, R., et al. (2020). Leak detection in water distribution systems using machine learning. *Journal of Hydroinformatics, 22*(3), 456–470.
38. Lazaric, A., Restelli, M., & Bonarini, A. (2007). Reinforcement learning in continuous action spaces through sequential Monte Carlo methods. In *Advances in Neural Information Processing Systems*.
39. Liu, B., Hsu, W., & Ma, Y. (1999). Mining association rules with multiple minimum supports. In *Proceedings of the Knowledge Discovery and Data Mining Conference* (pp. 337–341). San Diego, CA, USA.
40. Liu, J., Cai, D., & He, X. (2010). Gaussian mixture model with local consistency. *Proceedings of the AAAI Conference on Artificial Intelligence, 24*, 512–517.
41. Miani, R. G. L., & Junior, E. R. H. (2018). Eliminating redundant and irrelevant association rules in large knowledge bases. In *Proceedings of the 20th International Conference on Enterprise Information Systems* (pp. 17–28). Funchal, Madeira, Portugal.
42. Mooney, C. H., & Roddick, J. F. (2013). Sequential pattern mining: Approaches and algorithms. *ACM Computing Surveys, 45*, 1–39.
43. Müllner, D. (2011). *Modern hierarchical, agglomerative clustering algorithms*. arXiv.
44. Said, A. M. (2009). A comparative study of FP-growth variations. *International Journal of Computer Science and Network Security, 9*, 266–272.
45. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). *Proximal policy optimization algorithms*. arXiv. https://arxiv.org/abs/1707.06347

46. Steinwart, I., Christmann, A., 2008. Support Vector Machines. Springer.
47. Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. In *Advances in Neural Information Processing Systems* (pp. 1057–1063). MIT Press.
48. Tang, C. Y., Liu, C. H., Chen, W. K., & You, S. D. (2020). Implementing action mask in proximal policy optimization (PPO) algorithm. *ICT Express, 6*(3), 200–203.
49. Taylor, M. E., Whiteson, S., & Stone, P. (2006). Comparing evolutionary and temporal difference methods in a reinforcement learning domain. In *Proceedings of the Genetic and Evolutionary Computation Conference* (pp. 1321–1328).
50. Triguero, I., García, S., & Herrera, F. (2015). Self-labeled techniques for semi-supervised learning: Taxonomy, software, and empirical study. *Knowledge and Information Systems, 42*(2), 245–284. https://doi.org/10.1007/s10115-013-0706-y
51. UNESCO (2023). World Water Development Report. Paris: UNESCO Publishing.
52. Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double Q-learning. *Proceedings of the AAAI Conference on Artificial Intelligence, 30*(1), 2094–2100.
53. Vanham, D., Alfieri, L., Flörke, M., Grimaldi, S., Lorini, V., De Roo, A., & Feyen, L. (2021). The number of people exposed to water stress in relation to how much water is reserved for the environment: A global modelling study. *The Lancet Planetary Health, 5*(11), e766–e774. https://doi.org/10.1016/S2542-5196(21)00234-5.
54. Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing, 17*, 395–416.
55. Yoo, H., Kim, B., Kim, J. W., & Lee, J. H. (2020). Reinforcement learning based optimal control of batch processes using Monte-Carlo deep deterministic policy gradient with phase segmentation. *Computers & Chemical Engineering, 144*, 107133.
56. Zhang, S., et al. (2020). Deep learning for water quality prediction. *Environmental Monitoring and Assessment, 192*(5), 1–18.
57. Zhao, Y., & Karypis, G. (2002). Evaluation of hierarchical clustering algorithms for document datasets. In *Proceedings of the Eleventh International Conference on Information and Knowledge Management (CIKM '02)* (pp. 515–524). New York, NY, USA.
58. Zhu, X., & Ghahramani, Z. (2002). *Learning from labeled and unlabeled data with label propagation* (Technical Report CMU-CALD-02-107). Carnegie Mellon University, School of Computer Science.
59. Zhu, X., & Goldberg, A. B. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool Publishers.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING