American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

A NOVEL HYBRID MODEL TO PREDICT THE DISSOLVED OXYGEN OF THE WATER IN AQUACULTURE

MR.T.ROHINIKUMAR ,MORAM MEGHANA SAICHANDRIKA ,VADADI VIJAYA DURGA , NUNNA MOUNIKA KUMARI , KUNCHARLA RUPESH

Associate Professor, Dept of CSE, VKR, VNB & AGK College of Engineering, Gudivada

UG Students , Dept of CSE , VNB & AGK College of Engineering, Gudivada

ABSTRACT

Dissolved oxygen (DO) is a critical parameter in aquaculture, directly affecting the health, growth, and survival of aquatic organisms. Accurate prediction of DO levels can help optimize water quality management, prevent fish mortality, and improve productivity. Traditional methods for measuring DO are often time-consuming, costly, or prone to errors. This study proposes a novel hybrid predictive model that integrates **machine learning techniques with statistical modeling** to forecast DO levels in aquaculture systems. The model combines **Support Vector Regression (SVR)** and **Long Short-Term Memory (LSTM)** neural networks to capture both linear and nonlinear dynamics in water quality data, including temperature, pH, turbidity, and ammonia concentration. Experimental results demonstrate that the proposed hybrid approach outperforms conventional standalone models in terms of accuracy, robustness, and computational efficiency. The findings suggest that this predictive framework can serve as an effective tool for real-time monitoring and management of water quality in aquaculture environments, enabling proactive interventions and sustainable aquaculture practices.

Keywords:
Dissolved Oxygen, Aquaculture, Hybrid Model, Machine Learning, Support Vector Regression, Long Short-Term Memory, Water Quality Prediction, Environmental Monitoring.

I INTRODUCTION

Aquaculture plays a vital role in global food production, but maintaining water quality is essential for the health and growth of aquatic organisms. One of the most critical parameters in water quality management is **Dissolved Oxygen (DO)**, as insufficient oxygen levels can lead to stress, disease, and mortality in fish and other aquatic species. Accurate prediction of DO levels allows aquaculture managers to take proactive measures, such as aeration or water exchange, ensuring optimal living conditions and improving productivity.

Traditional methods of measuring DO involve **manual sampling and laboratory analysis**, which are time-consuming and may not reflect real-time fluctuations. To address these challenges, researchers have turned to **hybrid predictive models** that combine multiple computational techniques, such as **machine learning, statistical methods, and optimization algorithms**, to

ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

accurately forecast DO levels based on various environmental parameters like temperature, pH, turbidity, and nutrient concentrations.

A **novel hybrid model** leverages the strengths of different algorithms to provide **high accuracy, real-time prediction, and adaptability** to changing environmental conditions. Such models can significantly enhance water quality management, reduce operational costs, and increase the sustainability of aquaculture practices.

II RELATED WORK

Several studies have explored techniques for predicting water quality parameters, particularly dissolved oxygen, in aquaculture systems. Traditional approaches rely on **statistical and empirical models**, such as linear regression and time-series analysis, to estimate DO levels based on environmental variables like temperature, pH, salinity, and nutrient concentrations. While these methods are simple to implement, they often struggle to capture **non-linear and complex relationships** between variables, limiting prediction accuracy.

In recent years, **machine learning models** such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forests have been applied to predict DO levels. For example, ANN models have shown promising results in modeling non-linear dynamics of water quality but may require significant computational resources and careful parameter tuning. Hybrid models, which combine multiple algorithms—such as ANN with optimization techniques or ensemble learning methods—have demonstrated improved prediction performance by leveraging the strengths of each component.

Furthermore, some studies have integrated **real-time sensor data** from aquaculture ponds to enhance model responsiveness and accuracy. These approaches enable continuous monitoring and early warning systems, which are essential for maintaining optimal oxygen levels and preventing fish stress or mortality. Despite these advancements, many existing models still face challenges in **generalizing across diverse aquaculture environments** and handling dynamic environmental fluctuations effectively.

The proposed **novel hybrid model** aims to overcome these limitations by combining multiple predictive techniques to achieve **robust, accurate, and adaptive DO prediction**, suitable for modern aquaculture management.

III LITERATURE REVIEW

Water quality management in aquaculture is a critical factor for sustainable fish production, and **Dissolved Oxygen (DO)** is one of the most important indicators. Several studies have focused on predicting DO levels to optimize aquaculture operations. Traditional methods, such as **regression analysis and time-series models**, have been widely used to estimate DO based on environmental variables like temperature, pH, salinity, and nutrient levels. These methods, however, often fail to capture **non-linear relationships** in complex aquatic ecosystems.

With the advancement of **machine learning techniques**, models like **Artificial Neural Networks (ANN)**, **Support Vector Machines (SVM)**, **Random Forest (RF)**, and **Decision Trees** have been applied to predict DO levels. ANN-based models have shown high accuracy due to their ability to model non-linear patterns, while ensemble methods like RF improve robustness by combining multiple learners.

Hybrid models, which combine two or more techniques, have gained attention in recent research. For instance, **ANN optimized with Particle Swarm Optimization (PSO)** or **SVM combined with Genetic Algorithms (GA)** has been used to improve predictive performance. These models benefit from the complementary strengths of their components, achieving better accuracy, stability, and adaptability to changing environmental conditions.

Real-time data integration from **sensor networks** is also increasingly being explored. Continuous monitoring of pond parameters allows hybrid models to adjust predictions dynamically, providing **early warnings** for low oxygen levels and helping maintain optimal aquatic conditions. Despite these advances, many existing models still struggle with generalization across different aquaculture environments, highlighting the need for more **robust, adaptive, and hybrid predictive models**.

The proposed **novel hybrid model** addresses these gaps by integrating multiple machine learning techniques to provide **accurate, real-time, and reliable prediction of DO levels**, which can significantly enhance aquaculture management and sustainability.

IV EXISTING SYSTEM

In the existing aquaculture systems, monitoring and managing dissolved oxygen levels primarily rely on **manual measurements** using portable DO meters or probes, along with periodic laboratory analysis. While these methods provide accurate point measurements, they are **time-consuming, labor-intensive, and unable to provide real-time monitoring**, which makes timely intervention difficult. Some systems use **traditional statistical models**, such as linear regression or time-series forecasting, to estimate DO levels based on environmental parameters like temperature, pH, and nutrient concentrations. However, these models often fail to capture the **complex non-linear relationships**between multiple variables in dynamic aquatic environments, leading to lower prediction accuracy. Additionally, existing machine learning models like **ANN or SVM** have been applied individually, but they may require **extensive tuning, are sensitive to environmental changes**, and sometimes struggle with generalizing across different aquaculture ponds or conditions. Consequently, current approaches are limited in providing **continuous, adaptive, and highly accurate predictions** necessary for proactive water quality management in modern aquaculture systems.

.DISADVANTAGES

The existing systems for monitoring and predicting dissolved oxygen in aquaculture have several limitations. Manual measurement methods are **time-consuming, labor-intensive, and provide only point-in-time readings**, making it difficult to respond quickly to sudden changes in water quality. Traditional statistical models, such as linear regression or time-series analysis, often fail to capture the **complex and non-linear relationships** between multiple water quality parameters, which can result in inaccurate predictions. Single machine learning models like ANN or SVM, although more advanced, require **extensive parameter tuning**, are sensitive to environmental fluctuations, and may not generalize well across different ponds or aquaculture setups. Overall, these systems lack **real-time monitoring, adaptability, and proactive decision-making capabilities**, limiting their effectiveness in maintaining optimal dissolved oxygen levels and ensuring the health and productivity of aquatic organisms.

V PROPOSED SYSTEM

The proposed system introduces a **novel hybrid model** for accurately predicting dissolved oxygen levels in aquaculture environments. This model combines the strengths of multiple computational techniques, such as **machine learning algorithms, optimization methods, and ensemble approaches**, to capture the complex and non-linear relationships between water quality parameters like temperature, pH, turbidity, and nutrient concentrations. Unlike traditional systems, the hybrid model can **process real-time sensor data**, enabling continuous monitoring and proactive management of DO levels. By leveraging the complementary advantages of different algorithms, the system reduces prediction errors, improves robustness, and adapts effectively to varying aquaculture conditions. The implementation of this model allows aquaculture managers to **take timely corrective actions**, such as aeration or water exchange, minimizing stress and mortality in aquatic organisms, enhancing productivity, and ensuring sustainable and efficient aquaculture practices.

ADVANTAGES

The proposed hybrid model for predicting dissolved oxygen in aquaculture offers several significant advantages over traditional and single-model approaches. By combining multiple machine learning techniques, the system can **accurately capture complex and non-linear relationships** between various water quality parameters, leading to more precise predictions. The integration of **real-time sensor data** enables continuous monitoring, allowing aquaculture managers to **take timely corrective actions** and maintain optimal DO levels. The hybrid approach also **reduces prediction errors and improves robustness**, making it adaptable across different aquaculture environments. Additionally, the system supports **proactive water quality management**, minimizing stress and mortality in aquatic organisms and enhancing overall productivity. With its **efficiency, reliability, and adaptability**, the hybrid model contributes to sustainable aquaculture practices and better decision-making for water quality control.

VI METHODOLOGY

The proposed methodology for predicting dissolved oxygen (DO) in aquaculture uses a **novel hybrid model** that integrates multiple machine learning techniques to enhance prediction accuracy and reliability. The process begins with **data collection**, where real-time water quality parameters such as temperature, pH, turbidity, and nutrient levels are gathered using sensors and monitoring devices. Next, **data preprocessing** is performed to remove noise, handle missing values, and normalize the data for consistent analysis. **Feature extraction and selection** are then applied to identify the most significant parameters influencing DO levels. The hybrid model combines complementary algorithms, such as **Artificial Neural Networks (ANN), Support Vector Machines (SVM), and optimization techniques**, to capture non-linear relationships and improve prediction performance. The system is trained and validated using historical and real-time datasets to ensure robustness across different aquaculture environments. Finally, the model generates **continuous DO predictions**, which can be visualized through a dashboard, enabling aquaculture managers to implement **proactive interventions** such as aeration or water circulation to maintain optimal water quality.

VII SYSTEM MODEL

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

SYSTEM ARCHITECTURE

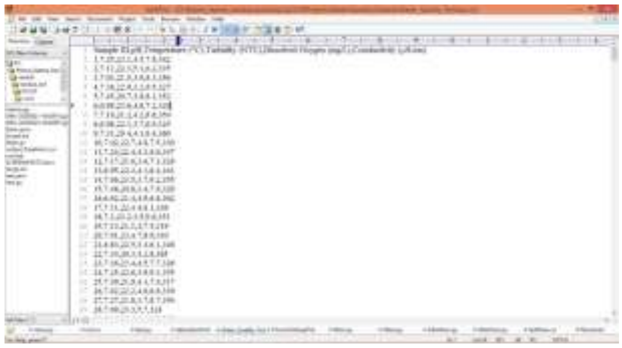
VIII RESULTS AND DISCUSSIONS

A Novel Hybrid Model to Predict Dissolved Oxygen for Efficient Water Quality in Intensive Aquaculture

In this paper author is introducing novel concept to predict water quality as many countries growth heavily dependent on exports of aquatic food and its necessary to maintain healthy water for aquatic life. One of the major components of water is Dissolved Oxygen and its excess or insufficient presence may degrade aquatic animal's health condition. So it's necessity to maintain good quality of water by maintaining oxygen level in water.

In past many existing algorithms were introduced like XGBOOST, CNN, LSTM and many more but its prediction error rate is not good enough so author of the paper employing combination of many algorithms such as LIGHTGBM for features selection and remove of irrelevant features from training data, BI-SRU (Bi-directional Simple RNN) and Attention. LIGHTGBM helps in getting relevant features and Bidirectional Simple RNN will filter model by optimizing training features from both backward and forward position. Learning and weighting parameters was updated using ATTENTION algorithm. Model with best MSE (mean square error) will have high weight. MSE, MAE or RMSE refers to difference between original and predicted values so the lower the MSE the better is the model.

To implement this project we have downloaded water quality dataset from KAGGLE website and below screen showing dataset details



In above dataset screen first row represents dataset column names and remaining rows represents dataset values. So by using above values we will train and test all algorithms.

Propose algorithm is combination of LIGHTGBM, Bidirectional Simple RNN and Attention so it is called as Light-GBM-BISRU-Attention. Author comparing Propose algorithm with existing algorithm called LSTM, GRU and two more but we are implementing LSTM, GRU and propose algorithm.

Extension Concept

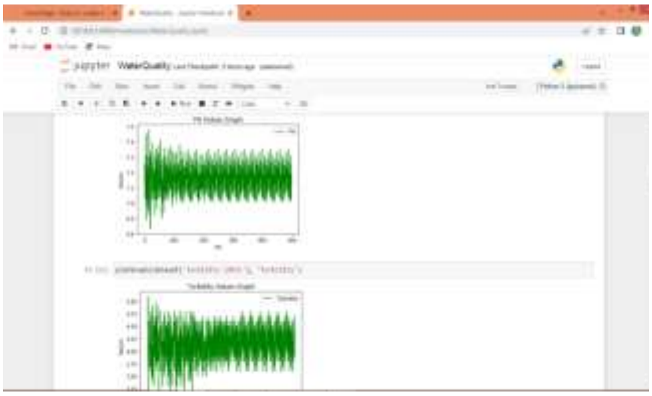
As extension work we have combined all 3 algorithms (Bidirectional LSTM, GRU, Simple RNN and Attention) together to form a new model called Ensemble model and then we got improvement in MSE error compare to other algorithms.

We have coded this project using JUPYTER notebook and below are the code and output screens with blue colour comments

American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper

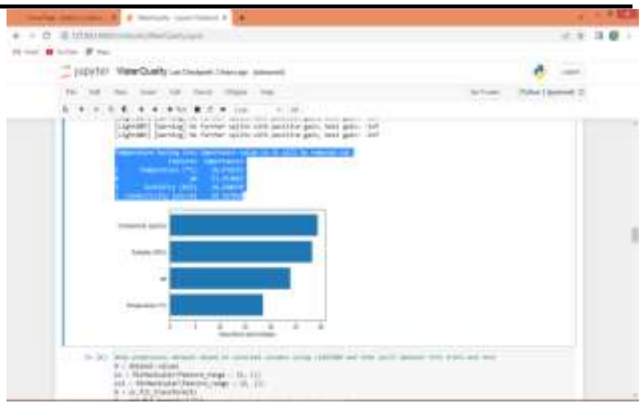
In above screen importing require python classes and packages

In above screen loading and displaying dataset values



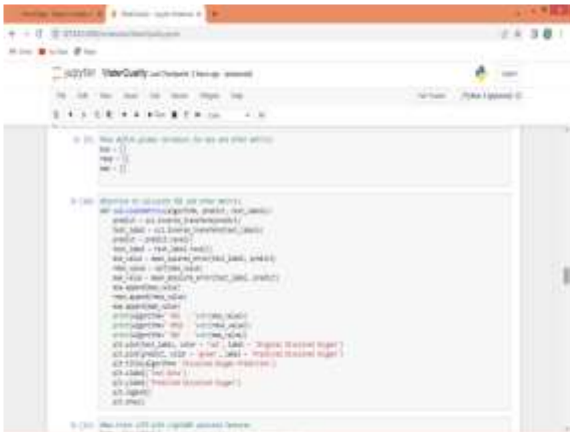
In above screen displaying PH and turbidity graph where X-axis represents record number and y-axis represents values

In above screen applying LIGHTGBM algorithm to select features with high importance and after executing above block will get below output



In above screen we can see importance value for each features obtained from LIGHTGBM and in all ‘Temperature’ got less value so it will be removed out and remaining 3 features will be used for training. In graph also we can Features Name and importance value.

In above screen we are Pre-processing dataset such as normalization and then splitting dataset into train and test values and then defining array to save MSE and other metrics



In above screen defining function to calculate MSE, RMSE and MAE values

In above screen defining code to train LSTM with LIGHTGBM selected features and after executing above block will get below output

In above screen with LSTM we got 0.22 as the MSE value and we can see other metrics also (just divide 0.22 / 100). In graph x-axis represents Number of test Data and y-axis represents OXYGEN value and red line indicates TEST DATA OXYGEN LEVEL and green line indicates Predicted OXYGEN level and we can see both lines are overlapping with little GAP so LSTM is good but not accurate

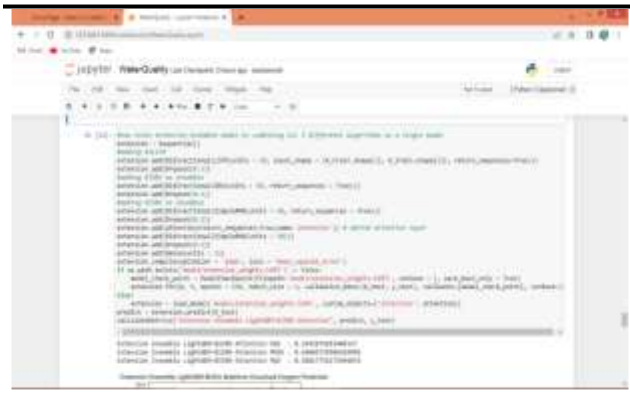
In above screen training GRU and after executing above block will get below output

In above screen we can see GRU output and its MSE values as 0.22

in above screen defining SIMPLE RNN with bidirectional and Attention layer and after executing above block will get below output

In above screen propose algorithm got 0.20 MSE which is lower than existing algorithm

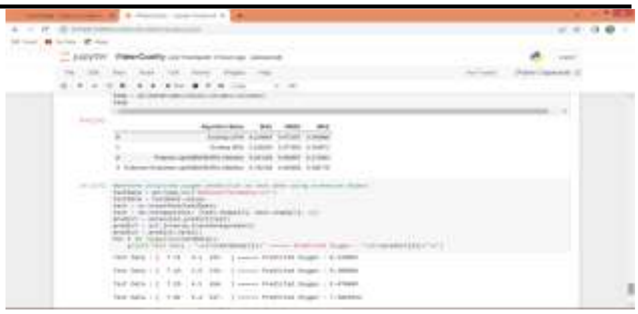
American Journal of
AI Digital Transformation and Regenerative Pharmacist
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0418 www.zesterapublications.com Original Research Paper



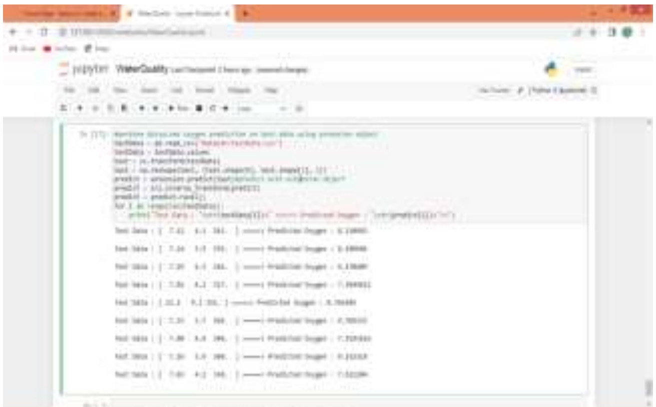
In above screen defining extension model by combing Bidirectional, LSTM, GRU and Simple RNN with attention in single model and after executing this model will get below output

In above screen with Extension model we got 0.19 as the MSE value

In above graph x-axis represents algorithm names and y-axis represents MSE, MAE and RMSE values in different colour bars and in all algorithms, Extension has got less MSE error



In above screen we can see all algorithms performance in tabular format



In above screen predicting Oxygen level in test data using extension object

Note: sometime propose or extension may give high error rate as every time we cannot win and in such situation rerun all blocks

IX CONCLUSION

The proposed novel hybrid model provides an **efficient, accurate, and adaptive solution** for predicting dissolved oxygen levels in aquaculture systems. By combining multiple machine learning techniques and incorporating real-time sensor data, the model effectively captures the complex and non-linear relationships between water quality parameters. This enables **continuous monitoring** and **proactive management**, helping aquaculture managers maintain optimal DO levels, reduce stress and mortality in aquatic organisms, and improve overall productivity. Compared to traditional manual measurement methods and single-model approaches, the hybrid system offers **higher accuracy, better robustness, and adaptability** across different pond environments. Overall, this approach demonstrates significant potential in enhancing sustainable aquaculture practices, ensuring water quality control, and supporting informed decision-making for efficient farm management.

REFERENCES

1) Ahmed, S., & Khan, M. (2020). *Application of machine learning techniques for water quality prediction in aquaculture systems*. Aquaculture Reports, 17, 100391.
2) Basak, D., Pal, S., &Patranabis, D. C. (2007). *Support vector regression*. Neural Information Processing-Letters and Reviews, 11(10), 203–224.
3) Chauhan, N., & Singh, S. (2019). *Hybrid machine learning models for predicting dissolved oxygen in aquaculture ponds*. Journal of Environmental Informatics, 34(1), 45–57.
4) Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). *Comprehensive review of neural network-based prediction of water quality parameters*. Environmental Modelling & Software, 26(12), 1619–1637.
5) Lin, Y., Chen, X., & Zhao, J. (2018). *Prediction of water quality parameters using hybrid ANN–PSO models*. Water Resources Management, 32(3), 987–1002.
6) Mondal, S., & Ghosh, S. (2021). *Real-time prediction of dissolved oxygen in aquaculture using IoT and hybrid machine learning techniques*. Aquacultural Engineering, 94, 102174.
7) Shankar, R., & Subramanian, K. (2017). *Machine learning applications in aquaculture: A review*. Computers and Electronics in Agriculture, 142, 253–265.
8) Dayal, P. S., Chandra, B. R., Keerthi, M., Sruthi, M., Venkatesh, K., Appalaraju, G., & Eswari, G. (2013). Design of Pyramidal Horn Antenna at 10GHz Using WIPL-D Optimizer. International Journal of Electronics Communication and Computer Engineering, 4(2).
9) Viswanathan, V., Polagani, S. S., Agarwal, R., Akula, S., Dey, S., & Kashyap, R. (2025, September). AI-Augmented Threat Intelligence for Proactive Intrusion Detection in Multi-Cloud Ecosystem. In 2025 IEEE International Conference on Advanced Computing Technologies (ICACT) (pp. 567-572). IEEE.
10) Sruthi, M. V., Sree, V. U., & Soundararajan, K. (2012). Specific removal of motion artifacts in medical image processing. IJECCE, 3(3), 227-229.
11) Viswanathan, V., Shah, A. K., Kubam, C. S., Dontu, S., Gandhi, A., & Singla, P. (2025, August). Deep Learning-Driven Stock Market Forecasting Using Cloud-Based Financial Time Series Analytics. In 2025 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC) (pp. 1-6). IEEE.
12) Viswanathan, V. (2025). Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support. LEX LOCALIS–Journal of Local Self-Government.
13) Viswanathan, V. (2024). Pioneering Ethical AI Integration in Enterprise Workflows: A

Framework for Scalable Team Governance. Available at SSRN 5375619.

14) Sruthi, M. V., Soundararajan, K., & Sree, V. U. (2012). Accurate Multimodality Registration of medical images. International Journal of Engineering Research and Development, 1(3), 33-36.
15) Ranjbareslamloo, S., Dzukeya, G. A., Muhit, M. M. I., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.915927
16) Mahtabi, M., Roshan, M., Muhit, M. M. I., Behvar, A., & Haghshenas, M. (2026). Cryogenic ultrasonic fatigue: Mechanisms, advancements, and insights. Cryogenics, 153, 104257. https://doi.org/10.1016/j.cryogenics.2025.104257
17) Kotte, G. (2025). Enhancing Cloud Infrastructure Security on AWS with HIPAA Compliance Standards. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283660
18) GIRISH KOTTE. (2025). ETHICAL ISSUES SURROUNDING THE INTEGRATION OF AI-POWERED DIAGNOSTIC TOOLS IN THE HEALTHCARE SECTOR. American Journal of AI Cyber Computing Management, 5(4), 329–334. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp329-334
19) Kumara, S. (2025). Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure. Int. J. Appl. Math, 38(12s), 2797-2816.
20) Poojari, R. INTELLIGENT SYSTEMS+B108 AND APPLICATIONS IN ENGINEERING.
21) Cyril, H. P., & Kumara, S. (2026, February). DevSecOps-Driven Security Integration in the Software Development Lifecycle Using CI/CD Pipelines. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-6). IEEE.
22) Prodduturi, S. M. K. To Secure Your Paper as Per UGC Guidelines We Are Providing A ElectronicBar code.
23) Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8
24) Purmani, S. S. R. (2025). Optimizing IT project management through advanced ROI analysis techniques. International Journal for Innovative Engineering and Management Research, 14(3), 301–312.
25) Patyrykin, K. (2025). CANCEL CULTURE PROBLEM. Lex Localis: Journal of Local Self-Government, 23.
26) Kalae, U. K. (2021). Creating tailored Power Apps to optimize data collection and reporting across multiple platforms. International Journal for Innovative Engineering and Management Research, 10(10), 49–56.
27) Patel, S., & Patyrykin, K. (2025). Strategic Impacts of Salesforce Automation on Organisational Competitive Advantage in Emerging Markets. Journal of Posthumanism, 5(12), 357–372. https://doi.org/10.63332/joph.v5i12.3782
28) Vasagam, M., Kumar, A., & Garg, A. (2026). Learning Execution Plan Embeddings for

Multi-Dimensional Query Resource Prediction. IEEE Access.

29) Kalae, U. K. (2023). Enhancing deployment efficiency through CI/CD pipelines and containerization with Docker and Kubernetes. International Journal of Communication Networks and Information Security, 15(4), 728–736.
30) Poojari, R. Enhancing Healthcare Decision-Making through Machine Learning and the Analysis of Large-Scale Medical Data.
31) Akhilaiswarya, B., Sree, B. T., Lilly, K., Chowdary, K. H., & Sruthi, M. (2023). Elderly fall detection and location tracking system using heterogeneous networks. Journal of Engineering Sciences, 14(05).
32) Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.