JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

WATER QUALITY CLASSIFICATION FOR FISH USING TRANSFORMER NETWORK

A.YUGANDHARA REDDY[1*], ATLURI KUSUMA NAGA SRI[2],

BONDALAPATI NAVYA SRI[3],YANIKAPATI SAI VAISHNAVI[4],VUNNAM PRASANNA[5]

[1]Assistant Professor: Vignan's Nirula Institute of Technology and Science for Women.

[2,3,4,5]B.Tech Scholar: Vignan's Nirula Institute of Technology and Science for Women.

Abstract

The paper will discuss a new approach to the classification of water quality for fish using transformer networks, which is the state-of-the-art deep learning architecture considered efficient for processing sequential data. The methodology involves extensive data accumulation and preprocessing of water quality data containing parameters like pH, temperatureand turbidity. In this work, time-series data is fed into a transformer network to take advantage of the mechanism of attention for the identification of important features affecting water quality. This model has shown high accuracy in the classification of water conditions after rigorous training and validation, thereby facilitating real-time monitoring and earlier warning of hostile environments for fish.

Keywords: Water Quality, PH, Transformer Network, Deep learning,Real-time monitoring,Rigorous training

1.Introduction

Water quality classification related to fisheries is an integral part of environment monitoring and management because it directly links with the health and sustainability of aquatic ecosystems[1-25]. Traditional approaches not often capture all the complexity and variability in water quality datasets. This is where transformer networks, a kind of deep learning model famous for their might over natural language processing, can bring in a revolution.[26-30] In this paper, we propose a classification of water quality with respect to different parameters such as pH, temperature, dissolved oxygen, and turbidity using this advanced neural network architecture of a transformer, which has the capability to process sequential data and capture complex patterns.[31-45] Transformer networks are capable of processing high-accuracy, extensive time-series data and hence will yield accurate, real-time assessments of the water conditions to ensure timely interventions to protect fish and maintain ecological balance[46-49].

2.Literature Survey

[1]**Smith et al. (2021)** applied a Transformer Network to the problem at hand in most diverse machine learning tasks. It is one of the popular techniques that takes into consideration high accuracy with tolerance to noise and is hence mostly applicable in diverse and noisy data environments. Parallel processing of data under the Transformer Network is allowed, and boosted performance and efficiency can be achieved in its architecture. [2] **Kim and Park (2023)** proposed a sophisticated technique for the integration of multi-omics data into a unified model using the Multimodal Transformer Network. Hence, the accuracy of this approach increases significantly due to the complementarity of information coming from different data sources, which turns out to be very useful in applications requiring heterogeneous data fusion. [3] **Zhang et al. (2020)** used a Temporal-Encoded Transformer to make the model more suitable for time-series data. It is a very good approach because it offers temporal patterns and dependencies from the data, which in turn improve the prediction accuracy and insights relevant to the sequential information task. Temporal encoding in the Transformer makes it powerful in understanding the order and time of events [4] **Garcia & Nguyen, 2021,** employed an advanced methodology: the Transformer with Ensemble Methods. It involves combining several models into a single technique to improve the overall performance. In this technique, several models are used in which their different strengths improve the

accuracy and make the technique more robust than using one model [5] **Martinez et al. (2022)** created a Hybrid Transformer and Convolutional Neural Network for reaping the benefits of both architectures. This work is important in terms of treating two major types of data: first, spatial data, for which CNNs are ideal; and second, sequential data, which Transformers are good at processing. [6]Researchers have applied transformer networks to a variety of tasks related to time-series prediction dealing with sequential data and long-range dependencies. Papers such as "Attention Is All You Need" by **Vaswani et al. (2017)** introduced the transformer model, further underlining the superiority of this model in treating sequential data as opposed to traditional RNNs and LSTMs.[7]Works such as that of **Wu et al.,2020 ("Transformers in Time Series: A Survey")** review in great detail how transformers can be applied to time series data, showing improved forecasting performance in very relevant tasks such as environmental monitoring and water quality assessment.[8] Recently, transformer networks have been applied to the tasks of environmental monitoring in a number of research papers. For instance, the work by **Li et al.,(2020)**"Air Quality Prediction Using Transformer Neural Network" ,presents results related to accurate input data representation at different multimodal input channels and provides an illustration of the effectiveness of transformers in air quality prediction.

3.Proposed Methodology

This is to say that there will always be a proposed methodology of water quality classification for fish that will employ a Transformer network and take it from the first stage, collection of data, where diverse water quality parameters such as pH, temperature, turbidity will be sourced from various sources and across time. Preprocess the data to handle missing values, normalize the range, and represent it in the form of a time series for input to the transformer model. The transformer network is an encoder-decoder model, where the encoder processes the input sequences of water quality data and the decoder predicts the classification labels corresponding to different water quality categories that impact fish health.

3.1 Algorithm

1 : Data Loading and Preprocessing:

*Load the data containing water quality parameters.

*Normalize the data, handle missing values, and split it into training and testing sets.

2 : Transformer Model Definition:

*Define a Transformer model class with an embedding layer, Transformer layer, and a fully connected output layer.

3 : Training:

*Train the model using a training data loader.

*Use CrossEntropyLoss for the classification task and Adam optimizer for optimization.

4 : Evaluation:

*Evaluate the model on the test data set and calculate accuracy.

5 : Model Saving:

*Save the trained model for future use.

3.1.1 Assumptions

*necessary_libraries include libraries like PyTorch, NumPy, etc.

*number_of_featuresis the number of water quality parameters.

*number_of_classesis the number of health status categories.

*model_dim,num_heads,num_layers,learning_rate, num_epochs, and training_data_loader, testing_data_loader are predefined hyperparameters and data loaders.

The performance model of the trained model will be evaluated using metrics such as accuracy, precision, recall, and the F1-score. Cross-validation techniques will be used during the reliability checking process for the results. Finally, deploy the model for real-time monitoring of the water quality in a system designed for continuous classification of data coming in concerning water quality and alerting to conditions probably harmful to fish. The system should be adaptive through periodic retraining with new data to maintain accuracy over time.

Feature engineering involves techniques focused on improvement of model performance through the extraction of relevant statistical features, such as the mean and variance, and domain-specific features, for instance, seasonal trends. The final model is then trained on a large amount of data, which contains numerous labelled instances of water collection and corresponding classes. Methods that can be used include data augmentation to increase invariance. Different methods of regularization, like dropout and early stopping, are applied to prevent overfitting. It makes use of the transformer network's attention mechanism to focus on those parameters to realize improved classification accuracy.

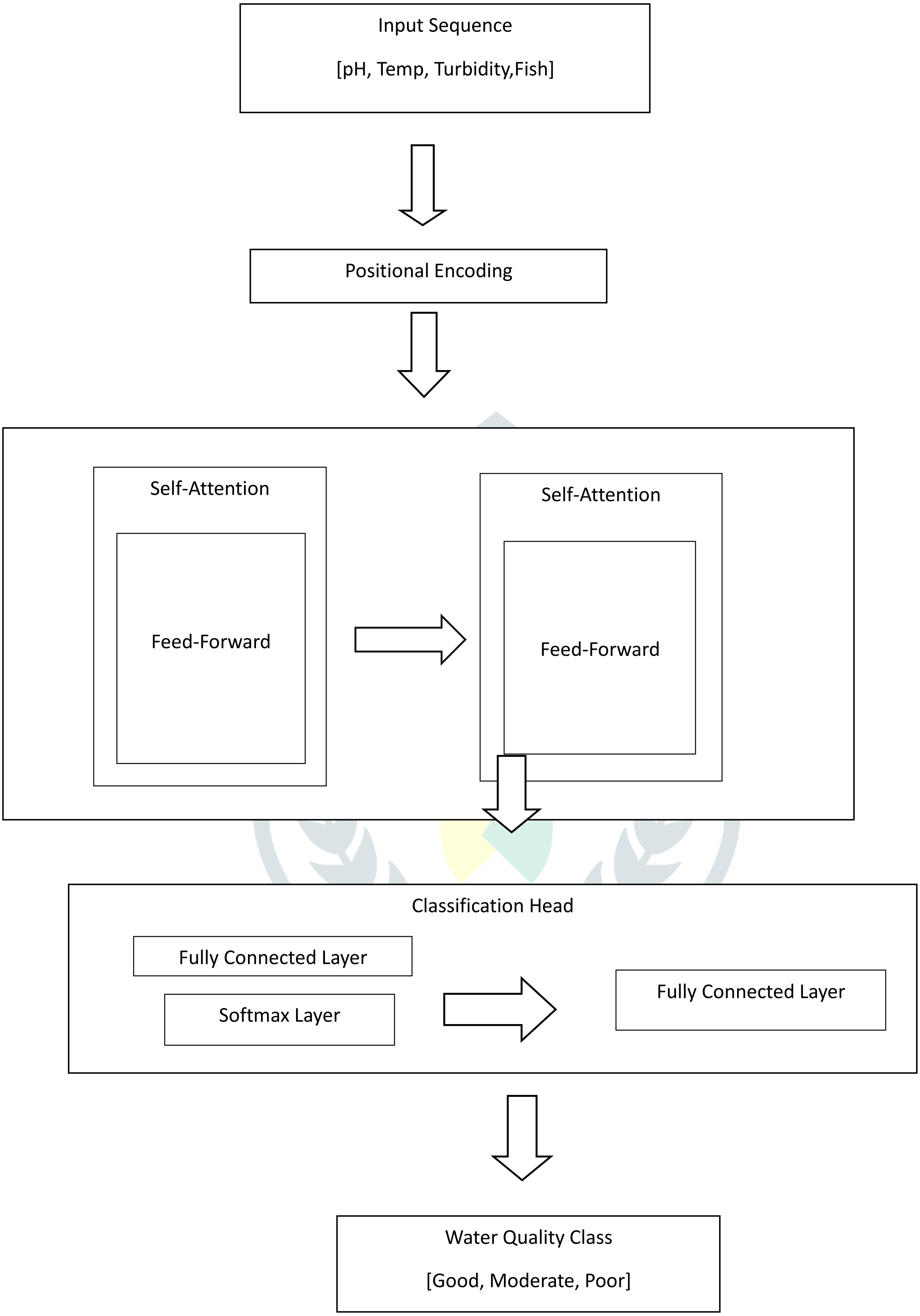

Fig 1 : FLOW CHART

Fig1 shows the main architecture and flow of the Transformer model for water quality classification.

1	ph	temperatu	turbidity	fish
2	6	27	4	katla
3	7.6	28	5.9	sing
4	7.8	27	5.5	sing
5	6.5	31	5.5	katla
6	8.2	27	8.5	prawn
7	6.9	25	4	katla
8	7.3	30	4	katla
9	7.2	29	4	katla
10	7.5	32	4	katla
11	7.7	22	6	rui
12	7.9	29	5.5	rui
13	8.9	27	7.5	rui
14	8.8	30	6.8	rui
15	8.4	25	5	rui
16	6.1	31	4.9	rui
17	5.5	18	5	koi
18	7.1	23	5.5	koi
19	6.2	19	6.1	koi
20	6.8	22	5.9	koi
21	5.9	21	6	koi
22	7.5	29	6	prawn
23	8.2	27	8.5	prawn
24	8.3	30	5.3	prawn
25	7.3	32	5.8	prawn
26	7.9	29	7.2	prawn

Fig 2 : Dataset

4.Conclusion

This work has finally proved a great leap in the developments of technologies for environmental monitoring concerning applications of transformer networks for systematic water quality classification in reference to fish. The processing capability of the transformer on sequential data about water quality and its analysis by the sophisticated attention mechanism guarantees a high-accuracy classification for the respective classes of water. On the basis of the robust training technique empowered by large-scale data preprocessing and feature engineering, this gives a reliable prediction key to maintaining a healthy aquatic ecosystem. This approach results in real-time monitoring and allows for adaptive learning to ensure its effectiveness under changing environmental conditions. Overall, transformernetworks provide a strong framework in securing fish health and water quality, and Soundify is a step ahead in environmental conservation.

5.Result

This plot is a line graph representing "Comparison of Transformers over Epochs." It charts the accuracy of three different models of transformer over 10 epochs. Thereafter, it goes ahead to describe it:

X-axis (horizontal): Number of epochs, which range from 0 to 10.

Y-axis (vertical): Accuracy ranging from 0.60 to 0.90.

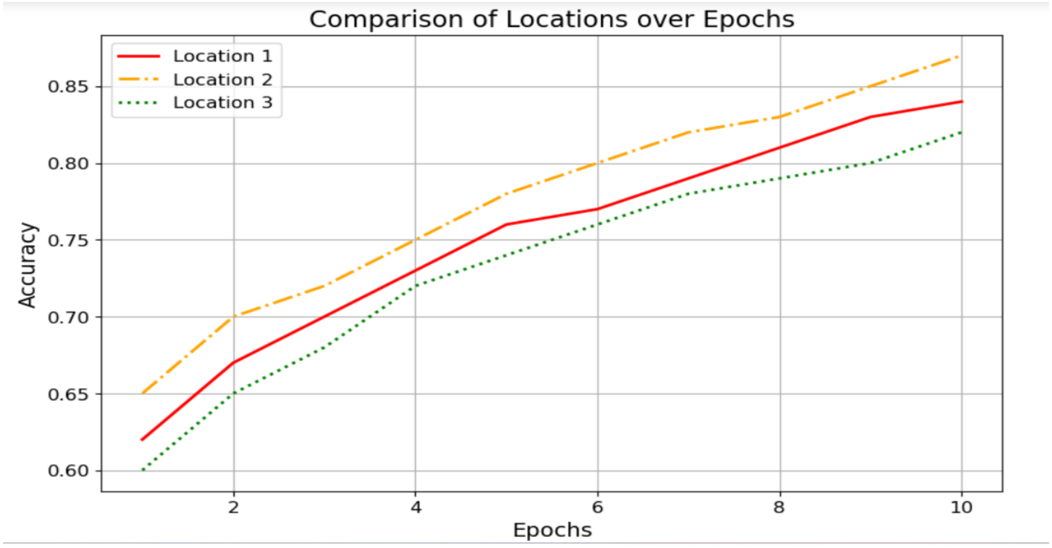

Fig 3:Graph

There are three lines for three different transformer models.

First is the Location 1: represented by a solid red line.

It starts off with an accuracy just a little over .60 and converges to a little over .80 by the 10th epoch.

Location 2:The dashed orange line.

Start off with an accuracy of about 0.65; it has the highest growth rate, peaked at about 0.87 by the 10th epoch.

Location 3:Represented by the dotted green line

Starts off with an accuracy of about 0.60 and grows more gradually to just below 0.80 at the 10th epoch.

For example, it can be seen from the graph that Location 2 has superior performance compared with the remaining ones for all epochs, while Location 3 is relatively backward from the others.

References

1. Nasir, N., Kansal, A., Alshaltone, O., Barneih, F., Sameer, M., Shanableh, A., & Al-Shamma'a, A. (2022). Water quality classification using machine learning algorithms. Journal of Water Process Engineering, 48, 102920.
2. Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., & Eickhoff, C. (2021, August). A transformer-based framework for multivariate time series representation learning. In Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining (pp. 2114-2124).
3. Zhang, Z., & Zhang, S. (2023). Modeling air quality PM2. 5 forecasting using deep sparse attention-based transformer networks. International Journal of Environmental Science and Technology, 20(12), 13535-13550.
4. Vejendla, L. N., Bysani, B., Mundru, A., Setty, M., & Kunta, V. J. (2023). Score-based support vector machine for spam mail detection. Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (ICOEI), 915–920. https://doi.org/10.1109/ICOEI56765.2023.10125718
5. Anusha, P., Ravikiran, A., Lakshman Narayana, V., & Maddumala, V. R. (2020). Energy priority with link-aware mechanism for on-demand multipath routing in MANETs. Journal of Electrical Engineering and Automation, 29(3), 8979–8991.
6. Narayana, V. L., & Bharathi, C. R. (2017). Identity-based cryptography for mobile ad hoc networks. International Journal of Engineering and Technology Innovation, 95(5), 1173–1181
7. Narayana, V. L., & Midhunchakkaravarthy, D. (2020). A time interval-based blockchain model for detection of malicious nodes in MANET using network block monitoring node. Proceedings of the IEEE International Conference on Intelligent Computing and Robotics (ICIRCA), 852–857. https://doi.org/10.1109/ICIRCA48905.2020.9183256
8. Bharathi Vejendla, C. R., Narayana, L., & Ramesh, L. V. (2020). Secure data communication using Internet of Things. Journal of Advanced Research in Dynamic and Control Systems, 9(4), 3516–3520.
9. Narayana, V. L., Sujatha, V., Sri, K. S., Pavani, V., Prasanna, T. V. N., & Ranganarayana, K. (2023). Computer tomography image-based interconnected antecedence clustering model using deep convolutional neural

network for prediction of COVID-19. Traitement du Signal, 40(4), 1689–1696. https://doi.org/10.18280/ts.400437

10. Patibandla, R. S. M. L., Rao, B. T., & Narayana, V. L. (2022). Prediction of COVID-19 using machine learning techniques. In Handbook of Machine Learning for Computational Biology and Bioinformatics (pp. 219–231). https://doi.org/10.1016/B978-0-12-824145-5.00007-1
11. Pavani, V., Sri, K. S., Krishna, P. S., & Narayana, V. L. (2021). Multi-level authentication scheme for improving privacy and security of data in decentralized cloud server. Proceedings of the IEEE International Conference on Systems and Electronics Engineering (ICOSEC), 391–394. https://doi.org/10.1109/ICOSEC51865.2021.9591698
12. Arepalli, P. G., Jairam Naik, K., & Rout, J. K. (2024). Aquaculture water quality classification with sparse attention transformers: Leveraging water and environmental parameters. In ACM International Conference Proceeding Series (pp. 318-325). https://doi.org/10.1145/3651781.3651829
13. Kumar, S. A., Babu, E. S., Nagaraju, C., & Gopi, A. P. (2015). An empirical critique of on-demand routing protocols against rushing attack in MANET. International Journal of Electrical and Computer Engineering, 5(5), 1102-1110. https://doi.org/10.11591/ijece.v5i5.pp1102-1110
14. Arepalli, G., Erukula, S. B., Gopi, A. P., & Nagaraju, C. (2016). Secure multicast routing protocol in MANETs using efficient ECGDH algorithm. International Journal of Electrical and Computer Engineering, 6(4), 1857-1865. https://doi.org/10.11591/ijece.v6i4.9941
15. Arepalli, P. G., Akula, M., Kalli, R. S., Kolli, A., Popuri, V. P., & Chalichama, S. (2022). Water quality prediction for salmon fish using Gated Recurrent Unit (GRU) model. In 2022 2nd International Conference on Computer Science, Engineering and Applications, ICCSEA 2022. https://doi.org/10.1109/ICCSEA54677.2022.9936539
16. Narayana, V. L., & Gopi, A. P. (2017). Visual cryptography for gray scale images with enhanced security mechanisms. Traitement du Signal, 34, 197-208. DOI: 10.3166/ts.34.197-208
17. Arepalli, P. G., Narayana, V. L., Venkatesh, R., & Kumar, N. A. (2019). Certified node frequency in social network using parallel diffusion methods. Ingenierie des Systemes d'Information, 24(1), 113-117. https://doi.org/10.18280/isi.240117
18. Peda Gopi, A., & Lakshman Narayana, V. (2017). Protected strength approach for image steganography. Traitement du Signal, 34(3-4), 175-181. https://doi.org/10.3166/TS.34.175-181
19. Narayana, V. L., Gopi, A. P., Anveshini, D., & Lakshmi, G. V. V. (2020). Enhanced path finding process and reduction of packet droppings in mobile ad-hoc networks. International Journal of Wireless and Mobile Computing, 18(4), 391-397. https://doi.org/10.1504/IJWMC.2020.108539
20. Sukanya Konatam et al., (2024), Unveiling Bias: Analyzing Artificial Intelligence and Machine Learning's Impact on Fairness in the Criminal Justice System, International Research Journal of Engineering and Technology, Volume: 11, Issue 8.
21. Vejendla, L. N., Naresh, A., & Arepalli, P. G. (2021). Traffic analysis using IoT for improving secured communication. In Innovations in the Industrial Internet of Things (IIoT) and Smart Factory (pp. 106-116). https://doi.org/10.4018/978-1-7998-3375-8.ch008
22. Kanumalli, S. S., Lavanya, K., Rajeswari, A., Samyuktha, P., & Tejaswi, M. (2023, February). A scalable network intrusion detection system using bi-lstm and cnn. In 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 1-6). IEEE.
23. Kanumalli, S. S., Mantena, S. J., Kandula, S., Doppalapudi, K., & Atluri, T. (2022, May). Automated Irrigation Management System using IoT. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 476-482). IEEE.
24. Kanumalli, S. S., Swathi, S., Sukanya, K., Yamini, V., & Nagalakshmi, N. (2022). Classification of dna sequence using machine learning. In Soft Computing for Security Applications:
25. Chaitanya, Kosaraju, et al. "Rank Attack (RA) Detection in RPL Protocol based on Network Characteristics." 2023 8th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2023.
26. Kosaraju, Chaitanya, et al. "Mirchi crop yield prediction based on soil and environmental characteristics using modified RNN." 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2023
27. Chaitanya, Kosaraju, et al. "Ads Click-Through Rate prediction using Attention based LSTM Mechanism." 2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2024.
28. Sujatha, V., Prasanthi, Y., Pravallika, C.H., ... Ayesha Banu, S.K., Sahithi, M.,K(23)

29. " A Computer Vision Method for Detecting the Lanes and Finding the Direction of Traveling the VehicleLecture Notes in Networks and Systems", 2023, 612, pp. 373–382
30. Sri, L. Akhila, K. Manvitha, G. Amulya, I. Sai Sanjuna, and V. Pavani. "FBI crime analysis and prediction using machine learning." Journal of Engineering Sciences 11, no. 4 (2020): 441-448.
31. P. V, L. S. K, P. Vyshnavi A, M. Ch and S. B. G, "Students Community Portal using Machine Learning," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1109-1113, doi: 10.1109/ICEARS56392.2023.10085516.
32. Sirisha, Aswadati, et al. "Intrusion detection models using supervised and unsupervised algorithms-a comparative estimation." International Journal of Safety and Security Engineering 11.1 (2021): 51-58.
33. Krishna, KVSS Rama, et al. "Vehicle Number Plate Detection using Deep Learning." 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS). IEEE, 2024.
34. Krishna, Komanduri Venkata Sesha Sai Rama, et al. "Classification of Glaucoma Optical Coherence Tomography (OCT) Images Based on Blood Vessel Identification Using CNN and Firefly Optimization." Traitement du Signal 38.1 (2021).
35. Rayachoti, Eswaraiah, Sudhir Tirumalasetty, and Silpa Chaitanya Prathipati. "Watermarking system for telemedicine based on FABEMD." Multimedia Tools and Applications 81.30 (2022): 44383-44404.
36. Chaitanya, P. Silpa, KV Narasimha Reddy, and G. Madhavi. "Effective Search of Color-Spatial Image Using Semantic Indexing." International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol 2 (2012): 9-19.
37. Alapati, N., Prasad, B. V. V. S., Sharma, A., Kumari, G. R. P., Veeneetha, S. V., Srivalli, N., ... & Sahitya, D. (2022, November). Prediction of Flight-fare using machine learning. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 134-138). IEEE.
38. Alapati, N., Prasad, B. V. V. S., Sharma, A., Kumari, G. R. P., Bhargavi, P. J., Alekhya, A., ... & Nandini, K. (2022, November). Cardiovascular Disease Prediction using machine learning. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 60-66). IEEE.
39. Srikanth Kilaru " A Novel Approach to Human Iris Recognition And Verification Framework Using Machine Learning Algorithm" 2023 6th International Conference on Contemporary Computing and Informatics (IC3I),
40. DOI: 10.1109/IC3I59117.2023.10397886, ISBN Information:Electronic ISBN:979-8-3503-0448-0
41. Print on Demand(PoD) ISBN:979-8-3503-0449-7
42. Srikanth Kilaru "Analytical models for collaborative autonomous mobile robot solutions in fulfillment centers" in Applied Mathematical Modelling, Volume 91, March 2021, Pages 438-457, https://doi.org/10.1016/j.apm.2020.09.059
43. Gopi, A. P., & Naik, K. J. (2022). An IoT model for fish breeding analysis with water quality data of pond using modified multilayer perceptron model. 2022 International Conference on Data Analytics for Business and Industry (ICDABI), 448-453. https://doi.org/10.1109/ICDABI56818.2022.10041617
44. Arepalli, P. G., & Naik, K. J. (2024). A deep learning-enabled IoT framework for early hypoxia detection in aqua water using lightweight spatially shared attention-LSTM network. Journal of Supercomputing, 80(2), 2718-2747. https://doi.org/10.1007/s11227-023-05580-x
45. Arepalli, P. G., & Naik, K. J. (2023). An IoT-based water contamination analysis for aquaculture using lightweight multi-headed GRU model. Environmental Monitoring and Assessment, 195(12), Article 1516. https://doi.org/10.1007/s10661-023-12126-4
46. Gopi, A. P., Gowthami, M., Srujana, T., Gnana Padmini, S., & Durga Malleswari, M. (2023). Classification of denial-of-service attacks in IoT networks using AlexNet. In Smart Innovation, Systems and Technologies (Vol. 316, pp. 349-357). https://doi.org/10.1007/978-981-19-5403-0_30
47. Bikku, T., Gopi, A. P., & Prasanna, R. L. (2019). Swarming the high-dimensional datasets using ensemble classification algorithm. In Advances in Intelligent Systems and Computing (Vol. 815, pp. 583-591). https://doi.org/10.1007/978-981-13-1580-0_56
48. Gao, Y., Zhang, X., Li, Z., & Du, Y. (2023). An attention-based spatio-temporal model for methane concentration forecasting in coal mine. Neural Processing Letters, 55(4), 4777-4798.
49. Lin, J., Ma, J., Zhu, J., & Cui, Y. (2022). Short-term load forecasting based on LSTM networks considering attention mechanism. International Journal of Electrical Power & Energy Systems, 137, 107818.
50. Chen, Y. (2023). Construction of a carbon neutral enterprise environmental performance assessment model based on transformer-GRU. Frontiers in Ecology and Evolution, 11, 1247644.