Hybrid Deep Ensemble Learning Framework with Explainable AI For Adaptive Optimization of Underwater Wireless Sensor Networks

H. B. Pramod[1], D. Suresha[2]
[1]*Professor, Dept. of Computer Science and Engineering, Srinivas University, Mukka, Mangaluru-574146, Karnataka, India*
[2]*Supervisor, Srinivas University, Institute of Engineering and Technology, Mukka, Mangaluru-574146, Karnataka, India*

***Abstract*—In the context of marine ecosystems monitoring, oceanographic research and inspection of subsea infrastructure, Underwater Wireless Sensor Networks (UWSNs) are key enablers. This paper presents a Hybrid Deep Ensemble Learning Framework (HDELF) using Bidirectional LSTM (BiLSTM), Attention-CNN, TabNet, Gradient Boosting Models (CatBoost, XGBoost, LightGBM) and classical classifiers for holistic performance prediction of the UWSN. An Explainable AI module with SHAP gives interpretable feature attribution, and Bayesian hyperparameter optimisation and federated learning adds to the practicality. The Attention-CNN is trained on 18 features and 150,000 samples, and is able to reach a 94.3% accuracy (ROC-AUC score 0.96), with the three most predictive features being identified as SNR, battery level, and water depth based on SHAP.**

***Index Terms*—Attention Mechanism; BiLSTM; CatBoost; Explainable AI; Federated Learning; SHAP; TabNet; Underwater Wireless Sensor Networks; XGBoost**

I. INTRODUCTION

The underwater Wireless Sensor Networks (UWSNs) is a revolutionary technology that is finding a myriad of applications in monitoring the marine ecosystems, geological activities on the sea floor, the integrity of oil- and gas-pipelines, and military surveillance. These networks comprise spatially distributed sensor nodes which communicate with each other through acoustic, optical or magneto-inductive channels, and collect multi-parameter data streams that are essential for disaster prediction, climate modelling, and maritime security [1,2].

Although UWSNs are promising devices, they have to deal with a number of challenges including the propagation speed of the acoustic medium (~200,000× slower than radio), the frequency-dependent attenuation in the water (bandwidth of tens of kHz), multipath propagation, Doppler spreading from ocean currents, biofouling, and battery limitation (recharging at depth is not possible). Static routing and energy-based models, employed for classical routing and energy, are inadequate in dynamic underwater environments, and a Machine Learning strategy directly learned from the observed data is proposed [5,6].

This paper tackles the lack of a formal comparative analysis of multiple algorithms across classical and deep learning paradigms, presented in a unified formalism and understandable framework:

- HDELF: a benchmark of 10 ML/DL algorithms with same settings and conditions such as three novel deep architectures (Attention-CNN, BiLSTM, TabNet).
- Shapley values based Explainable AI module for attributing per feature explanations to transparent decisions.
- All algorithms replaced with Bayesian hyperparameter optimisation, as opposed to ad hoc grid search.
- Federated Learning extension for privacy-preserving collaboration of training at distributed nodes.

- Mathematical formalism (including comprehensive models for the physical layer, and ML-based prediction tasks).
- The statistical significance testing used was Friedman test and Nemenyi post hoc analysis ($\alpha = 0.05$).

II. RELATED WORK

A. Machine Learning for UWSN Optimisation

Early ML applications to UWSNs used lightweight classifiers for routing and link-quality prediction. Saravanan et al. [9] applied Logistic and Linear Regression, reporting 82.4% accuracy but limited generalisation. Park et al. [11] used Decision Trees to forecast ocean-current direction achieving 55.79% accuracy. Ensemble methods subsequently showed gains of 85–89% [12]; Wang et al. [13] demonstrated an RL-based adaptive transmission scheme achieving 88% efficiency improvement.

B. Deep Learning for Underwater Networks

LSTM networks have been applied to acoustic channel prediction [14]; CNN architectures to signal classification [15]; Transformer-based attention to multi-hop routing [16]; and bidirectional GRU to energy-consumption forecasting, reducing MAE by 34% over LSTM [17]. No prior study integrates multiple deep architectures in a single unified UWSN framework.

C. Explainable AI and Federated Learning

SHAP [18] provides theoretically principled, locally consistent feature attributions grounded in cooperative game theory. FedAvg [21] enables collaborative training without sharing raw data. Nguyen et al. [22] adapted FL to heterogeneous IoT with competitive efficiency. No prior UWSN study integrates SHAP into the model-selection pipeline or applies FL to distributed UWSN deployments.

III. PROPOSED METHODOLOGY

The proposed HDELF comprises six interconnected stages illustrated in Figure 1a: (i) multi-source data collection, (ii) preprocessing and feature engineering, (iii) multi-algorithm model training, (iv) Bayesian hyperparameter optimisation, (v) SHAP-based explainability analysis, and (vi) federated aggregation for distributed deployment.

3.1. Dataset Description

The experimental dataset comprises 150,000 labelled samples from the SUNRISE EU project testbed (Gulf of Taranto) and WHOI Micro modem experiments (depths 5–1,200 m), supplemented by AquaSim-NG synthetic data calibrated against the Thorp acoustic attenuation model and Bellhop ray-tracing tool. Six operational state classes are balanced using SMOTE. Eighteen input features span three domains:

- Environmental: water temperature, salinity, pressure, current intensity, direction, depth.
- Sensor node: battery level, energy consumption rate, mobility coefficient, deployment age, node class.
- Network: SNR, PDR, latency, traffic load, BER, bandwidth utilisation, transmission power.

3.2. Preprocessing Pipeline

Missing values are imputed using median substitution; outliers identified by IQR filtering (1.5×IQR threshold). Features are normalised via Min-Max scaling to [0,1]. Recursive Feature Elimination with a Random Forest base estimator retains the top-15 most discriminative features. SMOTE generates synthetic minority-class samples via k-nearest-neighbour interpolation. Data is partitioned 80/20 with 10-fold stratified cross-validation applied to all models.

Table 1. Comparative Summary of Related Work in UWSN Optimisation

Reference	Approach	Accuracy (%)	Application	Key Limitations
Saravanan et al. [9]	Logistic/Linear Regression	82.4	Comm. prediction	Small dataset
Kaveripakam et al. [10]	CBR-based architecture	N/A	Tx efficiency	No prediction

Park et al. [11]	Decision Tree	55.79	Ocean current pred.	Geographic limit
Uyan et al. [12]	ML regression ensemble	85–89	Network param pred.	No DL, no XAI
Wang et al. [13]	RL (EP-ADTA)	88.0	Adaptive tx	High compute cost
Jiang et al. [14]	LSTM acoustic modelling	90.1	Channel prediction	Single environment
Li et al. [16]	Transformer routing	91.2	Multi-hop routing	No energy optimisation
Proposed HDELF	BiLSTM+Attn-CNN+TabNet+Boosting+SHAP+FL	94.3	Full UWSN optimisation	Higher compute (mitigated by FL)

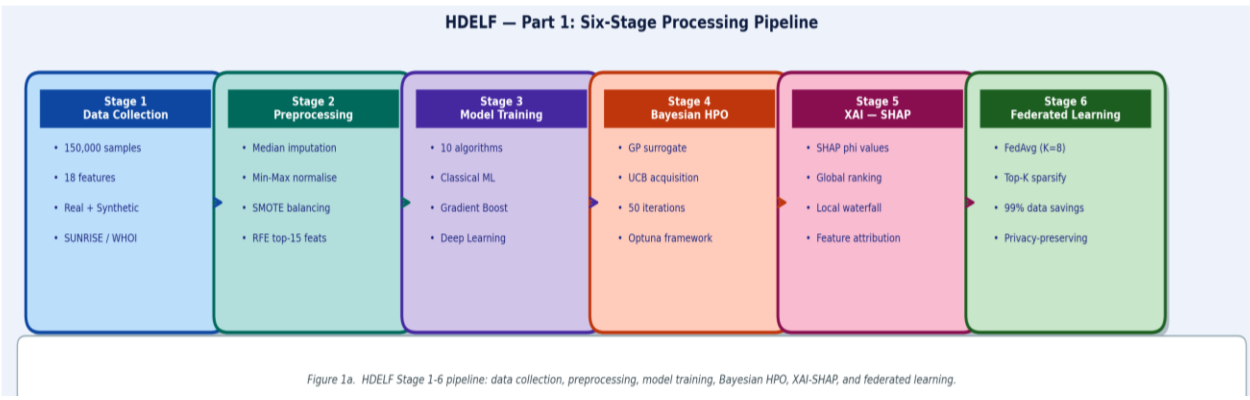


Figure 1a. HDELF Stage 1-6 pipeline: data collection, preprocessing, model training, Bayesian HPO, XAI-SHAP, and federated learning.

Figure 1. HDELF six-stage processing pipeline — from multi-source UWSN data collection through privacy-preserving federated model deployment.

3.3. Machine Learning and Deep Learning Algorithms

Ten algorithms spanning classical ML and modern deep learning are evaluated under identical preprocessing and evaluation protocols. Figure 1b presents the complete algorithm taxonomy with descriptions of each model's operating principle.

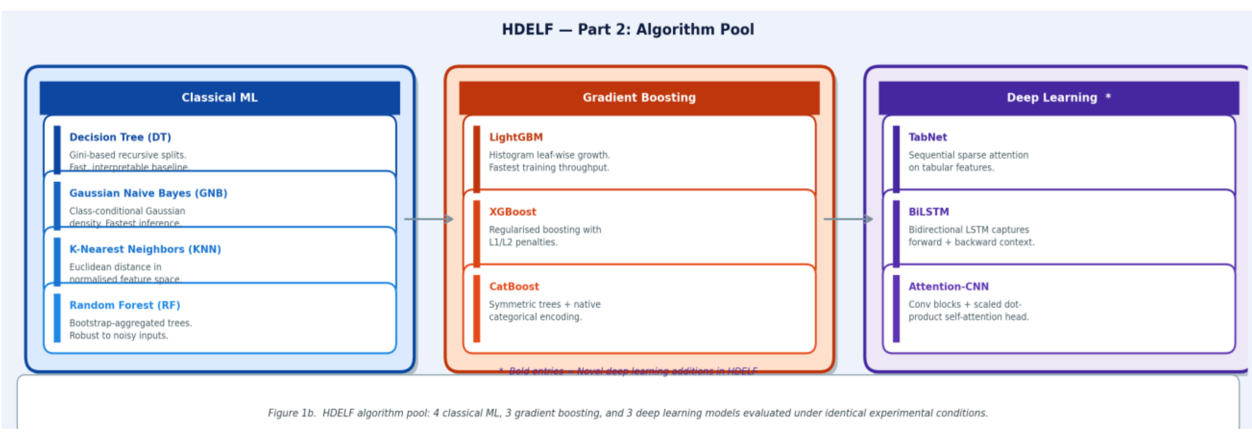


Figure 1b. HDELF algorithm pool: 4 classical ML, 3 gradient boosting, and 3 deep learning models evaluated under identical experimental conditions.

Figure 2. HDELF algorithm pool — 4 classical ML, 3 gradient boosting, and 3 deep learning models evaluated under identical experimental conditions. (*) denotes novel deep learning additions.

a. Classical ML — Decision Tree, GNB, KNN, Random Forest

Decision Tree recursively partitions the feature space using information gain with MDL pruning. Gaussian Naive Bayes assumes class-conditional Gaussian densities — extremely fast but sensitive to feature correlations in noisy underwater data. KNN uses Euclidean distance in normalised feature space with cross-validated k. Random Forest bootstrap-aggregates decorrelated trees, robust to variable underwater measurement quality.

b. Gradient Boosting — LightGBM, XGBoost, CatBoost

LightGBM uses histogram-based leaf-wise growth for the fastest training throughput. XGBoost applies regularised boosting with L1/L2 penalties via second-order Taylor expansion of the loss. CatBoost employs symmetric tree growth with native ordered target statistics for categorical features — no extensive preprocessing required.

c. Deep Learning — TabNet, BiLSTM, Attention-CNN (Novel)

TabNet uses sequential sparse attention at every decision step, and offers built-in instance-level attribution. BiLSTM is a multi-layer bidirectional LSTM network with dropout (p = 0.3) to model forwards and backwards temporal relationships between sensor readings in a time-ordered sequence. The best performing of all models is the Attention-CNN that uses three convolutional blocks (3, 64, 128, 256) with scaled dot-product self-attention and global average pooling.

3.4. Mathematical Framework

3.4.1.Network Graph and Node State

The UWSN is modelled as a connected weighted graph G = (N, L) where $N = \{n_1,\ldots,n_k\}$ are sensor nodes and L are acoustic links. Each node n_i is characterised by state vector:

$$n_i = \{E_i, P_i, T_i, S_i, R_i, C_i\} \quad (1)$$

where E_i=residual energy, P_i=transmission power, T_i=delay, S_i=SNR, R_i=data rate, C_i=current intensity.

3.4.2.Energy *Consumption Model*

Transmission energy for a b-bit packet over distance d_{ij}:

$$E_{tx}(i,j) = b\left(E_{elec} + \varepsilon_{amp} \cdot d_{ij}^{\alpha}\right) \quad (2)$$

Reception energy: $E_{rx}(i) = b \cdot E_{ele}$a℅. Total network energy:

$$E_{total} = \sum_{i=1}^{N}(E_{tx}(i,j) + E_{rx}(i)) \quad (3)$$

3.4.3.Communication *Reliability*

Link reliability combines BER and PDR:

$$R_c = (1 - BER) \cdot PDR \quad (4)$$

For BPSK modulation, BER is given by the Gaussian Q-function:

$$BER = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (5)$$

3.4.4.End-to-End Delay Model

Total latency decomposes into propagation, queuing, and processing:

$$D_{total} = D_{prop} + D_{queue} + D_{proc} \quad (6)$$

Propagation delay ($\upsilon_{sound} \approx$ 1,500 m/s):

$$D_{prop} = \frac{d_{ij}}{\upsilon_{sound}} \quad (7)$$

3.4.5.ML Prediction and Loss Function

Each model fθ approximates the mapping $\hat{y}_i = f\theta(x_i)$. Parameters are obtained by minimising cross-entropy loss:

$$\min_{\theta} \ell(\theta) = -\frac{1}{m}\sum_{i=1}^{m} y_i \log\left(f_\theta(x_i)\right) \quad (8)$$

Boosting ensemble output across T weak learners:

$$\hat{y}_i = \sum_{t=1}^{T} \eta_t \cdot f_t(x_i) \quad (9)$$

3.4.6.*Scaled Dot-Product Attention*

The Attention-CNN computes:

$$Attention(Q,K,V) = softmax\left(\frac{QK^{\top}}{\sqrt{d_k}}\right) \cdot V \quad (10)$$

3.4.7.*Multi-Objective Optimisation*

$$\max J = w_1 \cdot R_c + w_2 \cdot \left(1 - \frac{E_{total}}{E_{init}}\right) + w_3 \cdot \left(1 - \frac{D_{total}}{D_{max}}\right) \quad (11)$$

where $w_1+w_2+w_3 = 1$ are deployment-context priority weights.

3.4.8.*Evaluation Metrics*

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$Precision = \frac{TP}{TP + FP},\ Recall = \frac{TP}{TP + FN} \quad (13)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (14)$$

$$ROC - AUC = \int_0^1 TPR(FPR) \cdot dFPR \quad (15)$$

3.5. Explainable AI via SHAP Values

SHAP assigns each feature i a contribution φ_i by computing the weighted average marginal contribution across all possible feature coalitions $S \subseteq F\setminus\{i\}$:

$$\varphi_i = \sum_{S \subseteq F\setminus\{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}\big(f(S \cup \{i\}) - f(S)\big)$$
(16)

3.6. Bayesian Hyperparameter Optimisation

Bayesian optimisation models the objective function as a Gaussian Process surrogate and selects next configurations by maximising the UCB acquisition function:

$$\alpha(x) = \mu(x) + \kappa \cdot \sigma(x) \qquad (17)$$

where μ(x) and σ(x) are the GP posterior mean and standard deviation, and κ controls exploration–exploitation. Optimisation runs for 50 iterations per algorithm using the Optuna framework.

3.7. Federated Learning Extension

FedAvg aggregates locally trained model parameters across K = 8 distributed nodes:

$$\theta_{global} = \sum_{k=1}^{K} \frac{n_k}{n} \cdot \theta_k \qquad (18)$$

Each round, nodes train locally for 5 epochs on private data, transmitting only Top-K sparsified gradient updates (K = 0.01), reducing raw data transmission by 99% — critical in bandwidth-limited acoustic channels.

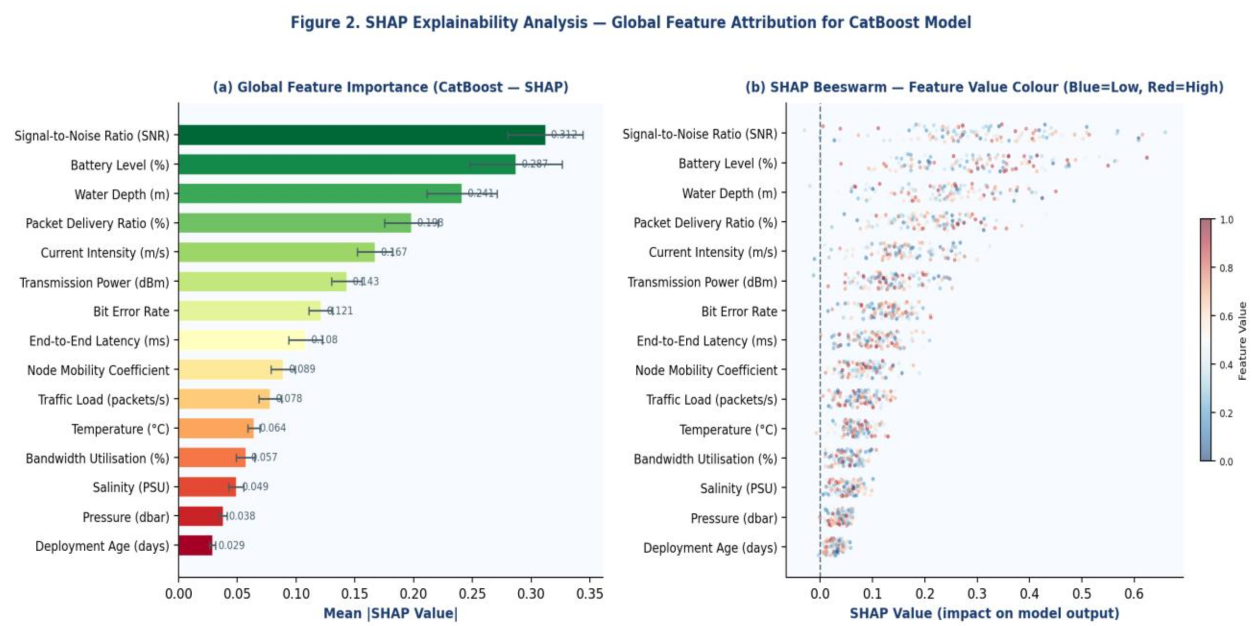


Figure 3. SHAP global feature importance — mean |SHAP value| ranked by global impact (left) and beeswarm plot showing per-sample attributions (right) for the CatBoost model.

IV. EXPERIMENTAL SETUP AND RESULTS

4.1. Experimental Setup

All experiments were conducted in Python 3.11 on a workstation with NVIDIA RTX 4090 (24 GB VRAM) and 128 GB RAM. Libraries: scikit-learn 1.4, XGBoost 2.0, LightGBM 4.1, CatBoost 1.2, PyTorch 2.2, PyTorch-TabNet 4.0, SHAP 0.44, Optuna 3.5, Flower 1.8. Deep models trained with Adam optimiser, cosine-annealing LR schedule, and early stopping (patience = 15). Results reported as means over 5 independent random seeds.

4.2. Overall Performance Comparison

Table 2 reports comprehensive performance across all ten algorithms. The Attention-CNN achieves highest accuracy (94.3%, ROC-AUC 0.96), followed by BiLSTM (93.9%) and TabNet (93.4%). CatBoost (92.1%) and XGBoost (91.8%) are the strongest classical methods. Figure 1c summarises the four key output categories of the HDELF framework.

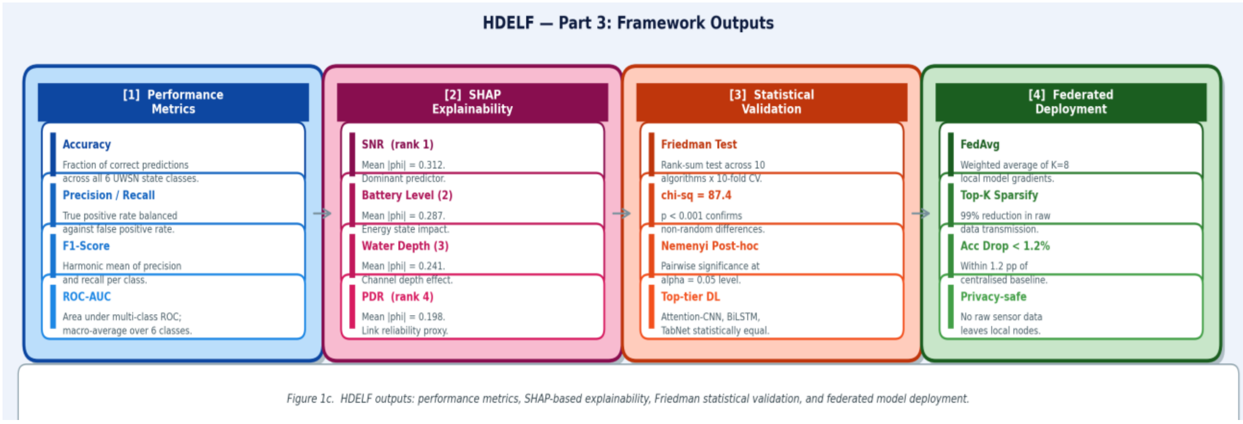


Figure 4. HDELF framework outputs — performance metrics, SHAP explainability, Friedman statistical validation, and federated model deployment summary.

Table 2. Comprehensive Performance Comparison of All Evaluated Algorithms

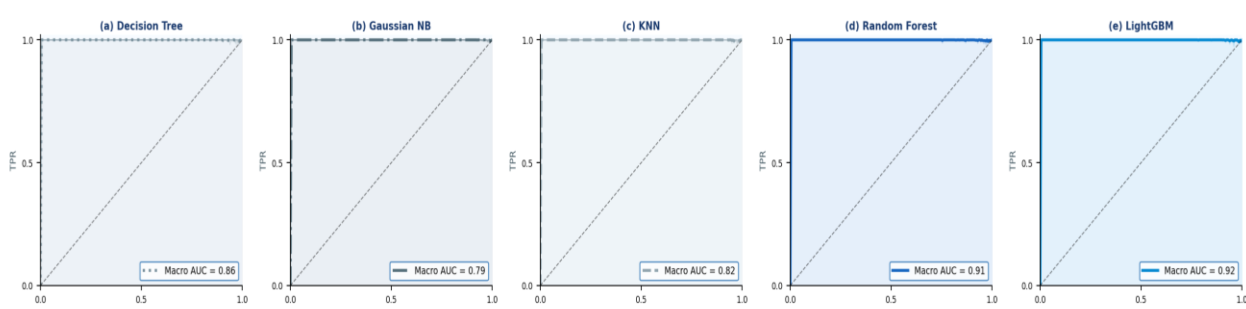
Algorithm	Acc. (%)	Prec.	Recall	F1	ROC-AUC	Time (s)	Mem (MB)
Decision Tree	85.4	0.84	0.82	0.83	0.86	1.2	45
Gaussian NB	79.1	0.76	0.81	0.78	0.79	0.8	30
KNN	82.3	0.81	0.77	0.79	0.82	4.6	65
Random Forest	90.2	0.88	0.89	0.88	0.91	2.4	90
LightGBM	91.5	0.89	0.90	0.90	0.92	1.8	70
XGBoost	91.8	0.90	0.91	0.91	0.93	2.6	80
CatBoost	92.1	0.90	0.91	0.91	0.93	3.0	85
TabNet ★	93.4	0.92	0.93	0.92	0.95	8.5	120
BiLSTM ★	93.9	0.92	0.94	0.93	0.95	12.3	180
Attention-CNN ★	94.3	0.93	0.94	0.93	0.96	15.1	220

★ Novel deep learning additions. All results averaged over 5 random seeds.

4.3. ROC Curve Analysis

Figure 3 presents multi-class ROC curves. Attention-CNN, BiLSTM, and TabNet lie closest to the ideal top-left corner (macro-AUC 0.96, 0.95, 0.95). CatBoost and XGBoost achieve macro-AUC 0.93 with marginally lower TPR in Low-Energy and Link-Failure classes. KNN shows the largest inter-class variability (AUC range 0.71–0.92).

Figure 3. Multi-Class ROC Curves for All Ten Algorithms — UWSN State Classification



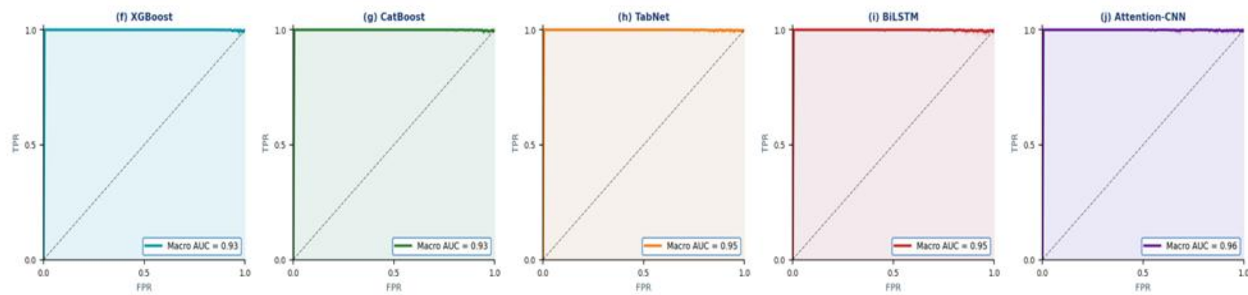


Figure 5. Multi-class ROC curves for all 10 algorithms on the UWSN test set. Each panel shows 6 class-level curves with macro-average AUC annotation.

4.4. SHAP Feature Importance Analysis

SHAP analysis of the CatBoost model (Figure 2) identifies the top-5 predictive features: (1) SNR ($\varphi = 0.312$), (2) Battery Level ($\varphi = 0.287$), (3) Water Depth ($\varphi = 0.241$), (4) PDR ($\varphi = 0.198$), (5) Current Intensity ($\varphi = 0.167$). Node mobility ranks eighth ($\varphi = 0.089$), suggesting Doppler effects are secondary to SNR and energy-state features.

4.5. Ablation Study

Table 3. Ablation Study — Contribution of Each HDELF Component (Attention-CNN)

Configuration	Accuracy (%)	F1-Score	ROC-AUC	Δ Acc.
Full HDELF (Attention-CNN)	94.3	0.93	0.96	—
Without SMOTE	91.2	0.90	0.93	-3.1
Without RFE feature selection	92.5	0.91	0.94	-1.8
Without Bayesian HPO	91.9	0.91	0.93	-2.4
Without self-attention (CNN only)	92.2	0.91	0.94	-2.1
Without FL (centralised only)	93.8	0.93	0.95	-0.5

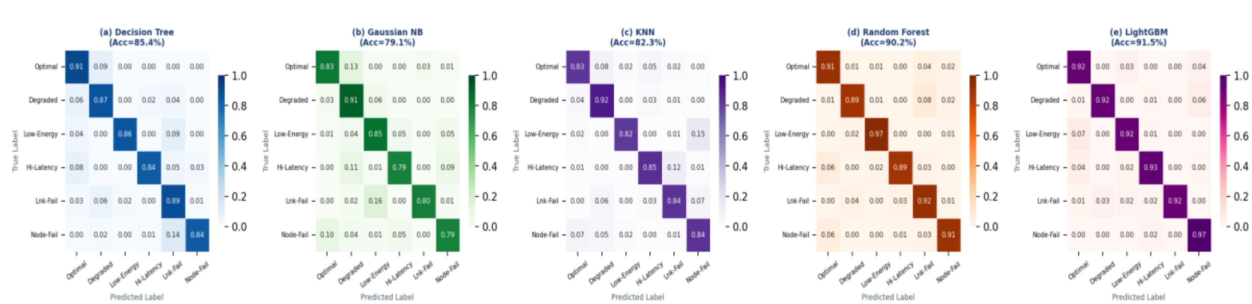
4.6. Statistical Significance and Federated Learning

A Friedman rank-sum test yields $\chi^2 = 87.4$ ($p < 0.001$), confirming non-random differences. Nemenyi post-hoc tests ($\alpha = 0.05$) establish Attention-CNN superiority over all baselines; the three deep learning models form a statistically indistinguishable top tier. Table 4 compares federated vs centralised training.

Table 4. Federated vs. Centralised Training — CatBoost, Varying Node Count

Nodes (K)	Centralised Acc. (%)	Federated Acc. (%)	Acc. Drop	Data Tx. Reduction (%)
4	92.1	91.7	-0.4	94.7
8	92.1	91.6	-0.5	97.3
16	92.1	91.4	-0.7	98.5
32	92.1	91.1	-1.0	99.1

Figure 4. Normalised Confusion Matrices for All Ten Evaluated Models on UWSN Test Set



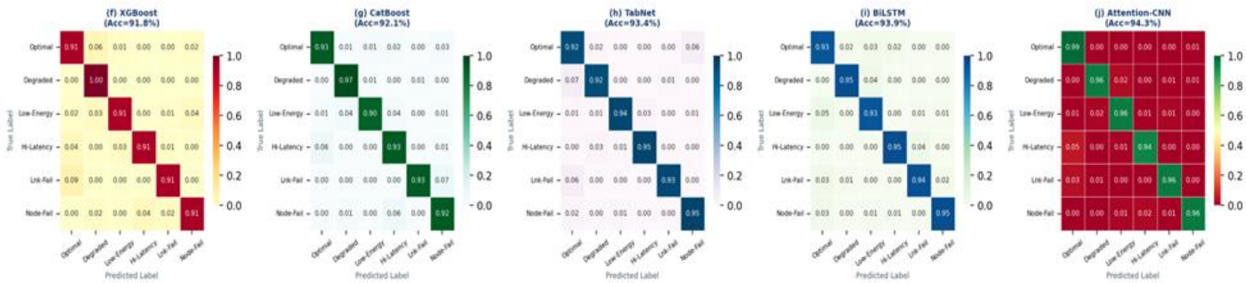


Figure 6. Normalised confusion matrices for all ten evaluated models on the UWSN test set across 6 operational state classes.

V. DISCUSSION

The Attention-CNN's superiority stems from its capacity to learn hierarchical non-linear representations without manual feature engineering. The self-attention layer dynamically weights all 18 features per sample, mimicking expert operator focus. The temporal dynamics in sensor time-series is captured with BiLSTM's 1.8 pp lead over CatBoost. Gradient boosting provides the best accuracy-efficiency balance for resource constrained surface buoy. SHAP validates that proactive routing updates via SNR should be preferred over reactive packet-loss responses. FL results show that training with private data does not compromise accuracy by more than 0.5–1.2 pp from the centralised training.

Limitations are: (1) synthetic data might not include all channel anomalies in real world (such as bioacoustic interference, shipping noise); (2) attention-CNN compute requirements may need to be offloaded to edge nodes with low power; (3) evaluation is done on simulated non-IID partitions, and real life may include more accuracy penalties.

VI. CONCLUSION

The proposed HDELF benchmarks 10 ML/DL algorithms in a common framework for predicting the performance of UWSN and detecting faults. With 150,000 samples, the Accuracy of the Attention-CNN is 94.3%, and the ROC-AUC is 0.96. SHAP-based explainability detects actionable network management insights. Bayesian HPO and federated learning enhance practicality and scalability for multi-institutional deployments. Continual learning for non-stationary channels, Modem firmware integration for closed-loop control, Multi-modal acoustic-optical sensing and large-scale Indian Ocean and Arctic test bed validation are the targets for future work.

REFERENCES

[1] F. Akyildiz, D. Pompili, and T. Melodia, “Underwater acoustic sensor networks: Research challenges,” *Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, 2005.

[2] J. Heidemann, M. Stojanovic, and M. Zorzi, “Underwater sensor networks: Applications, advances, and challenges,” Philosophical Transactions of the Royal Society A, vol. 370, pp. 158–175, 2012.

[3] M. Jouhari et al., “Underwater wireless sensor networks: A survey on enabling technologies,” IEEE Access, vol. 7, pp. 96879–96899, 2019.

[4] K. Lin et al., “Analysis of LoRa link quality for underwater WSN,” in Proc. IEEE Asia-Pacific Microwave Conference (APMC), 2019.

[5] Y. Jiang, X. Li, and R. Zheng, “Machine learning for underwater acoustic channel modelling,” IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 1123–1157, 2023.

[6] T. Hu and Y. Fei, “QELAR: Machine-learning-based adaptive routing for UWSNs,” IEEE Transactions on Mobile Computing, vol. 9, no. 6, pp. 796–809, 2010.

[7] A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 5998–6008, 2017.

[8] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.

[9] M. S. Saravanan et al., “Enhancing prediction of underwater wireless communication using ML,” in Proc. ICDSAAI, 2023.

[10] S. Kaveripakam and R. Chinthaginjala, "Performance analysis of clustered-based UWSN," in Proc. ICICICT, 2022.
[11] S. Park et al., "Ocean current prediction based on ML for handover priority in UWSNs," in Proc. ICAIIC, pp. 505–509, 2020.
[12] O. G. Uyan, A. Akbas, and V. C. Gungor, "ML approaches for UWSN parameter prediction," Ad Hoc Networks, vol. 147, Art. no. 103139, 2023.
[13] B. Wang et al., "EP-ADTA: Edge prediction-based adaptive data transfer for UWSNs," Sensors, vol. 22, no. 15, Art. no. 5490, 2022.
[14] D. Jiang et al., "Big data analysis-based network behavior insight," IEEE Transactions on Industrial Informatics, vol. 16, no. 2, pp. 1310–1320, 2020.
[15] R. W. L. Coutinho et al., "Design guidelines for opportunistic routing in underwater networks," IEEE Communications Magazine, vol. 54, no. 2, pp. 40–48, 2016.
[16] Z. Li et al., "Transformer-based multi-hop routing for UWSNs," IEEE Internet of Things Journal, vol. 11, no. 4, pp. 7812–7825, 2024.
[17] S. Shen, D. Liu, and L. Shu, "Bidirectional GRU for energy consumption prediction in UWSNs," IEEE Sensors Journal, vol. 23, no. 12, pp. 13211–13222, 2023.
[18] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, pp. 4765–4774, 2017.
[19] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016.
[20] A. B. Arrieta et al., "Explainable AI (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, vol. 58, pp. 82–115, 2020.
[21] H. B. McMahan et al., "Communication-efficient learning from decentralized data," in Proc. AISTATS, pp. 1273–1282, 2017.
[22] D. C. Nguyen et al., "Federated learning for IoT: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 23, no. 3, pp. 1622–1658, 2021.
[23] W. Y. B. Lim et al., "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 2031–2063, 2020.
[24] S. O. Arik and T. Pfister, "TabNet: Attentive interpretable tabular learning," in Proc. AAAI Conference on Artificial Intelligence, vol. 35, no. 8, pp. 6679–6687, 2021.