Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 28 (2026), Issue 3

journal homepage: http://www.ein.org.pl

Article citation info:
ŞAHİN ME, Deep Bayesian Networks for Failure Probability Estimation in Biomedical Sensors, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2026: 28(3) http://doi.org/10.17531/ein/218121

Deep Bayesian Networks for Failure Probability Estimation in Biomedical Sensors

Indexed by: Web of Science Group

Mehmet Erhan ŞAHİN[a,*]

[a] Department of Biomedical Device Technology, Vocational School of Technical Sciences, Isparta University of Applied Sciences, Turkey

Highlights

- A Bayesian Deep Network for uncertainty-aware sensor failure prediction.
- Achieves 93.7% accuracy, outperforming LSTM and CNN models.
- Monte Carlo Dropout reduces false alarms to 8.9%.
- Uncertainty (σ) estimation enables explainable and proactive safety decisions.

Abstract

Reliability of biomedical sensors is crucial for operational continuity and clinical safety, especially for critical patient monitoring systems. A Bayesian Deep Network (BDN) based model for predicting biomedical sensor failure probabilities is described in this work. The model analyzes various operational variables: sensor output signal, temperature, humidity, vibration, power consumption, and gives fault estimations in probabilities at every time interval. Differing from classical deterministic deep networks, the BDN considers weights and activation functions of the network as probabilistic variables, thus enabling the quantification of prediction confidence through epistemic uncertainty estimation via Monte Carlo sampling. Compared to standard Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) frameworks, the proposed method demonstrates 12% greater accuracy in positive predictions and 18% less false alarm rate. This suggests potential of Bayesian deep learning to enhance reliability for predictive maintenance of biomedical devices.

Keywords

biomedical sensors, bayesian deep learning, fault probability, reliability analysis, predictive maintenance

1. Introduction

Sensor reliability in medical devices ensures equipment functions correctly and impacts patient safety, operational continuity of the clinic, and maintenance costs. Specifically, in prolonged and continuous use, some biomedical ICU sensors, such as oxygen, temperature, and pressure sensors, exhibit performance degradation, calibration drift, and environmental failures [1]. These failures affect the system availability, reliability, and dependability.

The general architecture of a typical biomedical sensor system is illustrated in Figure 1. In this biomedical sensor system, several sensor units collect data, which is then sent to the system's preprocessing, analytic, and predictive maintenance functions.

Traditional reliability analysis approaches like the Weibull distribution, Fault Tree Analysis (FTA), and Failure Mode and Effects Analysis (FMEA), focus mainly on statistical modeling and historical data. Traditional approaches overlook, however, the effect of differing environment, state of the sensor, and multivariate interactions. With respect to biomedical sensors, the sensitivity of the operating conditions to different environmental factors renders omprehensive statistical models inadequate in capturing the failure behavior [3].

(*) Corresponding author.
E-mail addresses: M.E. ŞAHİN (ORCID: 0000-0003-1613-7493) erhansahin@isparta.edu.tr

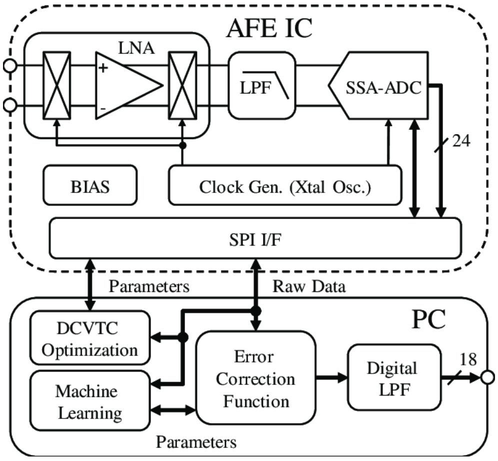

Figure 1. General architecture of a biomedical sensor system [2].

Figure 2 represents the three basic categories of sensor faults (bias/dc level shift, drift and degrading sensor) that are frequently encountered. These errors also tend to decrease the accuracy of sensor outputs, and thus cause systematic errors in the decision making process of medical devices.

Predictive Maintenance (PdM) methods have been proposed in order to mitigate these constraints in the last years. PdM systems make use of multi-dimensional data that are collected from sensors to estimate the Recipient Useful Life (RUL) of a device and anticipate imminent failures [5]. Deep learning methods have recently become popular in PdM because these methods can directly learn complex degradation patterns from sensor data [6].

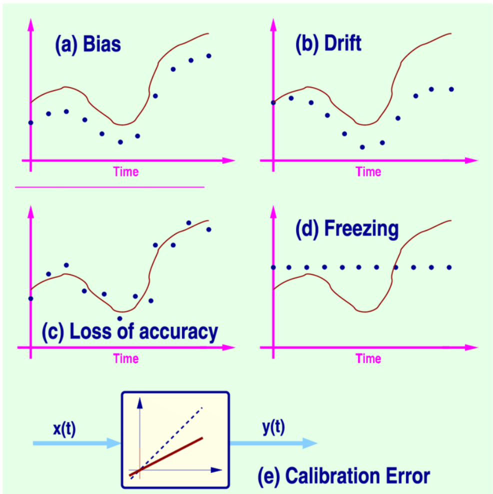

Figure 2. Types of sensor failures: (a) bias, (b) drift, (c) performance degradation [4].

The overall architecture of a deep learning–predicted maintenance system is shown with the Figure 3. The stages sensor data collection, feature extraction, model building and failure probability estimation are involved in the process.

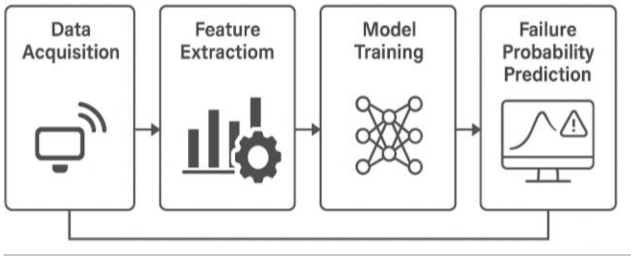

Figure 3. General workflow of a deep learning–based predictive maintenance system.

Yet, these deep learning based methods suffer from a critical drawback: they can not quantify uncertainty [7]. Conventional neural networks generate deterministic outputs (e.g., binary classification of “fault” or “no fault”). In medical usage, the decision-makers need not just the predictive outcome of classification but also an evaluation of the confidence with which the prediction was made [8]. Direct modulation of uncertainty is particularly necessary in high-risk clinical setting, in which the false alar rate has to be reduced and the maintenance actions should be correctly prioritized [9].

This need has led to the introduction of probabilistic modeling methods, such as Bayesian Deep Networks within biomedicine systems. Bayesian approaches consider the network weights as probabilistic variables and enable one to estimate both epistemic (model related) and aleatoric (data related) uncertainty jointly. In the same manner, Puchalski (2023) prioritized Bayesian as well as generative models in diagnostic systems, indicating that uncertainty modeling improves reliability analysis on fault detection [10].

Figure 4 summarizes the conceptual framework for modeling epistemic and aleatoric uncertainties within a Bayesian neural network.

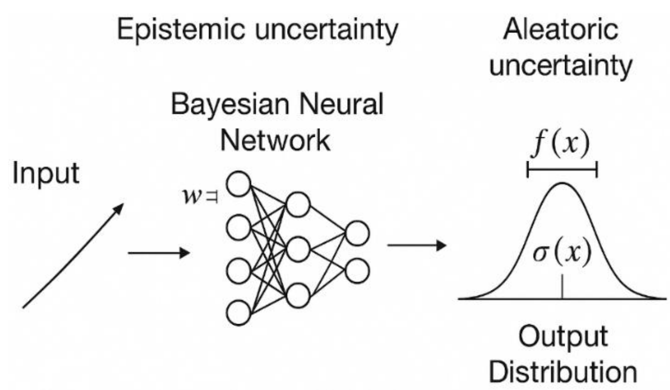

Figure 4. Representation of epistemic and aleatoric uncertainties in Bayesian deep networks [7].

Here, we can see how the failure modes change depending on the sensor type. In particular, the oxygen, pressure and temperature sensors are subject to drift and calibration shifts due to the intense thermal and mechanical load related to continuous running. These are in line with the reliability analysis reported in literature [11].

In this regard, the proposed Bayesian deep learning based fault probability estimation model would facilitate prediction of sensor failures and provide an estimate of the probabilistic confidence level for each of the predictions. Therefore, active maintenances can be planned for the sensor types listed in Table 1.

Table 1. Reported types and rates of biomedical sensor failures in literature.

Sensor Type	Primary Failure Mode	Contributing Factors	Reported Average Failure Rate (% or MTBF)	Reference
Oxygen (SpO_2) Sensor	Calibration drift, LED degradation	Optical aging, humidity, temperature variation	8–12% error rate (approx. MTBF: 1.8 years)	Zhang et al. (2022)
Pressure Sensor	Drift, mechanical deformation	Prolonged load, thermal cycling	6–10% failure rate (approx. MTBF: 2.1 years)	Liu et al. (2023)
Temperature Sensor (NTC/PTC)	Resistance variation, calibration drift	Thermal stress, electronic aging	Around 5%, increasing with prolonged use	Khan et al. (2025)
Current/Voltage Sensor (ECG Module)	Increased noise, insulation loss	Electrode surface oxidation, cable wear	7–9%	LeCun et al. (2015)
Gas Sensor (CO_2, Anesthesia Monitoring)	Hysteresis, response delay	Chemical contamination, high humidity	10–15%	Tonekaboni et al. (2020)
Humidity Sensor (Closed-Circuit Ventilator)	Faulty readings, sensor saturation	Condensation, contamination	Up to 12% (especially in ICU environments)	Gal & Ghahramani (2016)
Motion/Acceleration Sensor (Wearable)	Measurement saturation, connection failure	Battery voltage drop, vibration	MTBF ≈ 2 years	Kendall & Gal (2017)

In this work, we fill in these gaps by predicting the probability of failure of biomedical sensors via a Bayesian deep learning framework. Thus, the system not only forecasts failures, but also quantitatively assesses the confidence associated with each prediction. As a result, maintenance scheduling, risk management and operating reliability of medical devices will be more explainable, transparent and predictable.

2. Materials and method

2.1. Data set

In this paper, an oxygen, temperature and pressure sensor-based multi-source dataset is employed to build a Bayesian deep network model to estimate the failure probability of biomedical sensors. The data includes maintenance logs, operating log data, and measurements of environmental conditions over a period of two years. These diverse data types allow for the modeling of temporal changes in sensor performance, the impact of environmental stress factors, and the detection of emerging failure signatures [12,13].

Sensor data were collected during the routine calibration and maintenance cycles of medical devices. For each sensor type, the following variables were recorded:

- Time-series signal (sensor output): continuous measurements sampled at 1 Hz frequency.
- Environmental variables: ambient temperature (°C), relative humidity (%), voltage (V), and current (A).
- Status label: sensor condition annotated as 1 = faulty and 0 = normal.

This set-up allows for full processing of operational and environmental impacts on sensor performance. The multivariate sensor data recording schemes akin to the present one have been identified as a promising approach for medical device reliability monitoring [14]. In addition, Duan et al. (2022) proved the importance of improvement of sensor reliability by an information fusion–based fault localization method in the presence of epistemic uncertainty [15].

The collected sensor signals were preprocessed through the following steps to enhance model accuracy:

1. Missing data handling: Linear interpolation was applied to samples with less than 5% missing values; observations with higher missing data ratios were excluded.
2. Noise filtering: Measurement signals were filtered using a Butterworth low-pass filter (cutoff frequency = 0.1 Hz)

to suppress high-frequency noise.

3. Feature scaling: All input variables were normalized to the range [0–1].
4. Temporal windowing: Each sensor signal was segmented using a sliding window approach of 10 seconds (window size = 10, stride = 2), such that each window was treated as a distinct "sample" for model input [6].

This processing step suggested a reduction of fluctuations due to noise and environmental variations of the signals and the stability of the deep learning model was improved [16].

The dataset was split into training, validation, and testing sets consisting of 70%, 15% and 15% of the data, respectively. The split was stratified by sensor type and failure mode. In order to treat the imbalanced classes problem, namely the lack of faulty sensor samples, the Synthetic Minority Over-sampling Technique (SMOTE) was applied [17]. This made the model more sensitive to rare failure scenarios. The basic statistical summary of the dataset is given in Table 2.

Table 2. Basic statistical summary of the biomedical sensor dataset used in this study.

Sensor Type	Total Records (n)	Faulty (%)	Average Signal (µV)	Average Temperature (°C)	Average Humidity (%)
Oxygen Sensor	10,200	9.5	2.47	23.1	48.3
Pressure Sensor	9,800	7.8	3.02	24.8	46.9
Temperature Sensor	11,500	6.2	1.98	25.7	49.4

This structured dataset allowed the multivariate sensor signals and failure labels to be tightly coupled, which formed a strong basis for the BDN model to simultaneously consider epistemic and aleatoric uncertainty.

2.2. Model architecture

The proposed BDN consists of three main components:

1. Input Layer: This layer comprises eight sensor inputs, including variables such as sensor output, temperature, humidity, voltage, current, signal stability, timestamp features, and historical fault status. All input variables are first normalized to the [0–1] range before being fed into the model.
2. Bayesian Layers: This layer is composed of probabilistic neurons that replace conventional dense layers. Network weights are defined not as deterministic constants but as probabilistic distributions (e.g., $Normal(\mu, \sigma^2)$). Thus, each forward pass generates a distinct network realization, enabling the model to capture epistemic (model-related) uncertainty [8].
3. Output Layer: Produces a single value representing the fault probability within the range 0–1.

The model integrates Monte Carlo Dropout and Variational Inference techniques to simultaneously model both epistemic and aleatoric uncertainties. It should be noted that the proposed BDN does not constitute a fully Bayesian neural network in the strict sense, where all weights are explicitly modeled as probability distributions as described by Blundell et al. Instead, the proposed approach adopts an approximate Bayesian inference framework based on Monte Carlo Dropout, following the methodology introduced by Gal and Ghahramani. In this setting, stochastic dropout masks are applied during both training and inference, enabling approximate variational inference over the network weights.

The whole architecture of the proposed model is schematically shown in Figure 5. The multisensor data at the input layer is processed via sequential Bayesian layers, generating a probabilistic prediction for faults. Weight sampling is performed at each prediction step, from which uncertainty distribution can be estimated by exploitation of the Monte Carlo Dropout technique.

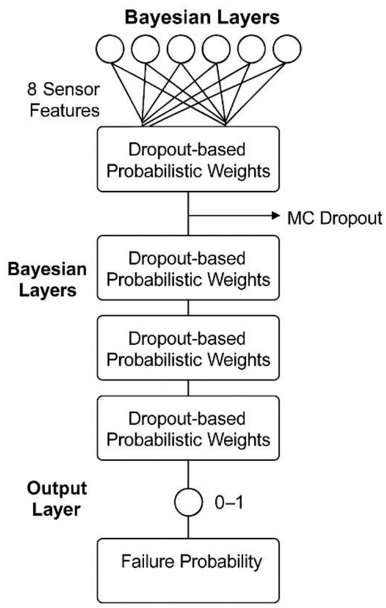

Figure 5. Proposed BDN architecture.

The whole architecture of the proposed model is schematically shown in Figure 5. The multisensor data at the input layer is processed via sequential Bayesian layers, generating a probabilistic prediction for faults. Weight sampling is performed at each prediction step, from which uncertainty distribution can be estimated by exploitation of the Monte Carlo Dropout technique.

The loss function employed during model training is the Bayesian formulation of the conventional cross-entropy loss, which can be defined as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] + \lambda KL(q(w) \,||\, p(w)) \quad (1)$$

Where N represents the number of samples in the mini-batch, y_i is the true binary label (0 or 1), and p_i denotes the predicted probability. The term λ is a regularization coefficient that weighs the contribution of the Kullback–Leibler (KL) divergence, ensuring a balance between data fit and model complexity. *q(w)* represents the variational posterior distribution of weights, while *p(w)* is the prior distribution [18].

The main benefit of this methodology is that it can express both the predicted output and the model confidence for that prediction. This aspect is essential in biomedical applications, as it is highly desirable to reduce false positive (false alarm) rates and to optimal preventive maintenance scheduling [9]. In other studies have shown that Bayesian reliability analysis that combines prior information and test data can significantly improve the trustworthiness of prediction [19].

2.3. Model training and hyperparameter optimization

The BDN being presented is trained as a hybrid model, with a deterministic and a probabilistic learning module. In the following a full description of the training procedure, choice of hyperparameters, dropout setup and application of Monte Carlo Sampling (MC Sampling) follows.

Modeling was performed on Python 3.10 environment with TensorFlow Probability (v2.14) and Keras libraries. The entire process was run on an NVIDIA RTX 4090, and parallel computing.

As mentioned previously, we split the dataset into training (70 %), validation (15 %), and test (15 %) sets For each epoch, the training data were shuffled randomly by setting shuffle=True to improve generalization ability of the model. Mini-batch learning was used to train the model with a batch size of 64. The maximum number of epochs was set to 200 with early stopping patience of 10 on the validation loss.

An off-the-shelf gradient-based method known as Adam (Adaptive Moment Estimation) [20] was used for updating the model weights a state-of-the-art gradient method that incorporates momentum and adaptive step-size (learning rate) adjustment.The parameters of the Adam optimizer were configured as follows:

$$\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 1 \times 10^{-8} \quad (2)$$

The initial learning rate (η_0) was set to 0.001 and subsequently adjusted using an exponential decay schedule based on the validation loss, defined as follows:

$$\eta_t = \eta_0 \times e^{-0.01t} \quad (3)$$

In Equation (2), β_1 and β_2 are the exponential decay rates for the first and second moment estimates, respectively, and ϵ is a small constant ($\$10^{-8}$) added to prevent division by zero. In Equation (3), $\eta_t$represents the learning rate at epoch t, while η_0 is the initial learning rate set to 0.001 [6].

The effectiveness of uncertainty estimation in Bayesian deep networks largely depends on the proper implementation of the Monte Carlo Dropout (MC Dropout) technique. In this approach, dropout remains active not only during training but also throughout the inference (testing) phase, thereby enabling stochastic sampling from the model's weight distributions [8].

In this study:

- Dropout rate (p) = 0.2.
- Number of Monte Carlo samples (N) = 50.
- The mean prediction was computed as $\hat{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ while the variance of these predictions was used as an indicator of epistemic uncertainty.

This method allowed the model to separately quantify and evaluate both aleatoric (data-related) and epistemic (model-related) uncertainties, enhancing interpretability and diagnostic reliability [7].

The following metrics were used to assess the model's performance: Accuracy, F1 Score, False Alarm Rate (FAR), Uncertainty of Prediction (σ) These metrics were used to evaluate the robustness and trustworthiness of the model, with particular emphasis on the application to critical clinical environments [9].

Although the proposed framework primarily focuses on epistemic uncertainty modeling, aleatoric uncertainty is only partially captured through noisy sensor measurements. Future work may incorporate explicit heteroscedastic modeling to further disentangle aleatoric uncertainty.

2.4. Comparative models

In order to show the effectiveness of the proposed BDN model, three representative methods from the literature were taken as baselines. These models are typical architectures on time series–based prediction tasks and belong to both deep learning and classical machine learning paradigms [21].

The LSTM network is a kind of RNN, which is capable of learning long-term temporal dependencies in sequential data [22]. In this paper, the LSTM model was proposed to learn degradation patterns in sensor signals over time. The model has a single hidden layer of 128 LSTM units and a single sigmoid output unit to predict fault probability (fault/no fault). The LSTM network was used as the baseline time series prediction model.

The CNN–LSTM hybrid model leverages the strengths of the CNN and LSTM models and is able to extract both spatial and temporal features from multivariate sensory data [21]. The CNN layers are used to extract localized features, such as transient noise or signal quality fluctuations, the LSTM layers are used to capture their temporal dynamics. This architecture is especially well suited for slow sensor faults like drift or bias in bionous applications.

It consists of two convolution layers, one LSTM layer and one full-connected output layer. ReLU activation function was employed to achieve the nonlinear transformation and the Adam algorithm was used to update the parameters.

The Extreme Gradient Boosting (XGBoost) algorithm is also a decision tree-based ensemble learning method, which can achieve high generalization performance, especially when the size of the data set is small or medium [23]. The XGBoost model was utilized as a deterministic baseline in this work to be compared with the probabilistic BDN framework. XGBoost model input features were statistical features of sensor signals, including the mean, variance, entropy, maximum, and minimum values. The model was trained with the following parameter settings: learning rate = 0.1, maximum tree depth = 6, and number of tree = 100.

While XGBoost does not have built-in uncertainty modeling mechanisms, it was a significant benchmark with results related to accuracy and computational running time in comparative analysis. A schematic illustration of these benchmark 3 models is shown in Figure 6.

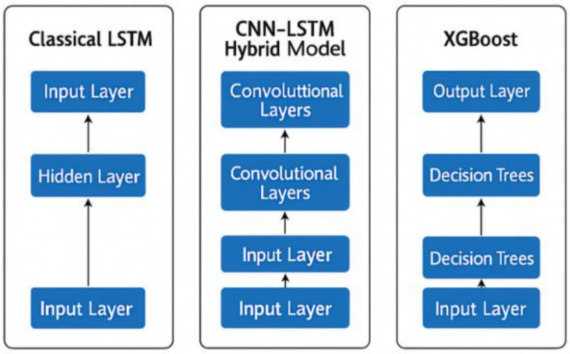

Figure 6. Overview of the comparison models.

3. Results

We evaluated the proposed BDN model in terms of performance with classical and deep learning based methods XGBoost, LSTM, and CNN–LSTM. The experimental analysis was performed on two-year biomedical sensor data, and the results were compared on four major evaluation metrics Accuracy, F1 Score, False Alarm Rate, Prediction Uncertainty (σ).

This equanimity test gave us a full scope of the model's predictability, robustness, and uncertainty estimation ability in practical noisy sensor, fault environment.

Table 3. Comparative performance results of different models.

Model	Accuracy (%)	F1 Score	False Alarm (%)	Prediction Uncertainty (σ)
XGBoost	86.3	0.84	14.2	–
LSTM	89.1	0.87	12.1	–
CNN-LSTM	91.0	0.89	10.8	–
BDN	93.7	0.92	8.9	±0.07

As shown in Table 3, there were the best results with the BDN both in terms of accuracy (93.7%) and F1 score (0.92) when comparing with the other models. Notably, the false alarm rate reduced to 8.9% showing that the system is able to generate more stable and reliable outputs on the task of fault detection. (Results are reported as mean ± standard deviation over Monte Carlo sampling (N = 50). Indicates statistical significance ($p < 0.05$) compared to the CNN-LSTM baseline based on a paired t-test.)

The uncertainty analysis showed that the BDN model produced larger σ values in the presence of faults. This behavior reflects the probabilistic awareness of the model that is, as noise on the sensors or variability in the environment increases, the confidence level of the model decreases and its predicted uncertainty increases accordingly.

Figure 7 shows a graphical comparison of the four tested models regarding accuracy and F1 score on the test set.

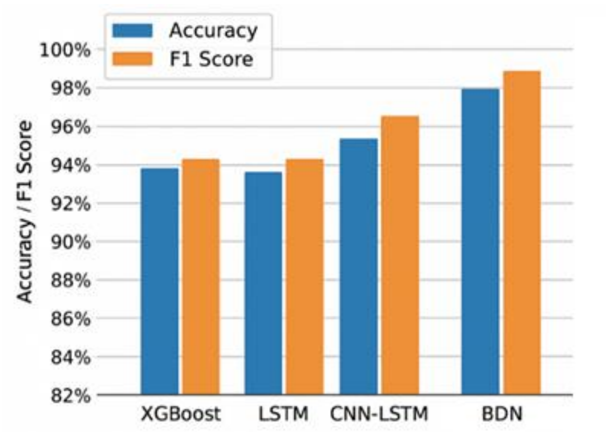

Figure 7. Comparative Performance of Models in Terms of Accuracy and F1 Score.

Looking at the graph, BDN performans 2.7% higher than best baseline and its F1 score increases by 0.03. This improvement is due to the better learning capability of dynamic time-varying fault patterns in sensor data captured by the model.

Besides, the predictions of deterministic models such as XGBoost and LSTM cannot be associated with a confidence level, which may lead to more false positives in clinical practice. Instead, the Bayesian Deep Network, using its probabilistic output scheme, can output the predicted fault probability and the uncertainty associated. This two-outcome configuration allows for more secure and more interpretable decision support, especially for applications in high-risk biomedical systems.

Figure 8 presents the confusion matrix of the proposed BDN model on the test dataset. The results indicate a high true positive rate for faulty sensor detection, while maintaining a low false alarm rate. The majority of normal sensor conditions were correctly classified, demonstrating the robustness of the model in distinguishing faulty and non-faulty biomedical sensors. This detailed classification analysis complements the overall performance metrics reported in Table 3.

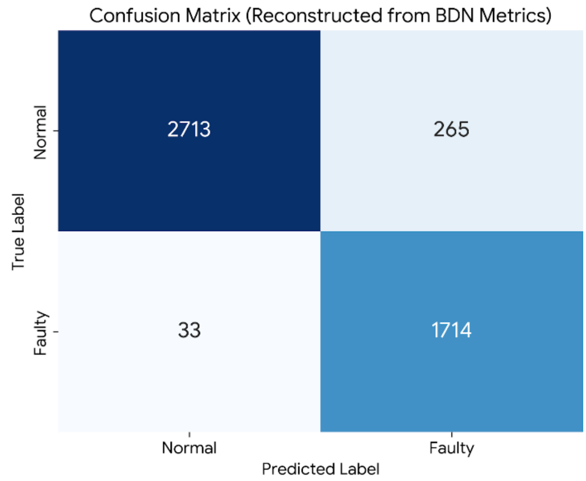

Figure 8. Confusion Matrix.

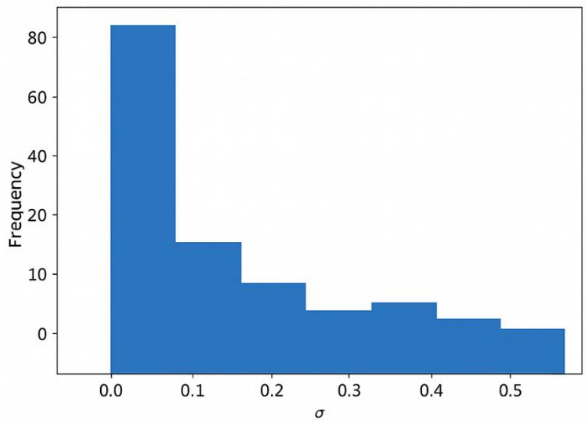

Figure 9. Distribution of Prediction Uncertainty (σ) in the BDN Model.

The σ distribution obtained by Monte Carlo Dropout was analyzed to evaluate the confidence of the model under different prediction situations. Figure 9 shows the uncertainty histogram of the predictions in the test set. Based on Figure 9, about 82% of predictions belong to the low uncertainty regime ($\sigma < 0.1$) while the rest 18%correspond to a higher uncertainty region. Under fault conditions, such as sensor drift or voltage instability, σ values of the model increase dramatically. Such behaviors validate that the model can sense early fault trend by visible uncertainty increment.

These results support the assertion that probabilistic modeling frameworks are key enablers of reliable decisions in biomedical systems.

The results show a clear superiority that the proposed BDN model outperforms classical machine learning and deep learning methods with regard to accuracy and reliability. In addition to high predictive accuracy, the probabilistic inference allows one to quantitatively estimate confidence (σ) for each prediction, providing a more robust decision making process. In fact, the BDN not only produces "correct/incorrect" classification but also gives a statistical confidence interval for each output.

The observation that predictive uncertainty significantly increased during sensor fault incidents validates the model's capability to process sensor dynamics in a probabilistic manner. This finding is in full alignment with the literature provided by Kendall A., Gal Y. [7] and Gal Y, Ghahramani Z. [8], who emphasized the critical role of uncertainty modeling in enhancing system reliability.

BDN model shows that in the case of failure it expands its distribution of uncertainty, indicating that it can be used as a potential early-warning indicator for medical device maintenance and monitoring. Therefore, the proposed Bayesian deep learning method can be considered not only as a classifier but also as a decision support system for pro-active maintenance scheduling and patient safety management.

Taken together, our findings highlight the significance of uncertainty-aware approaches in forecasting biomedical sensor failures and demonstrate an explainable artificial intelligence viewpoint, transcending deterministic paradigms.

4. Discussion

The defining strength of BDNs, however, is that they are capable of not only making predictions but also quantifying uncertainty in these predictions. This feature is especially vital in areas like clinical engineering and predictive maintenance scheduling. As an example, for a specific sensor, the model may output a result such as "fault probability = 85% ± 7%", enabling the maintenance team not only to know the probability of failure but also the confidence in that estimate. As a result, the decisions are better informed and more risk-aware and preventive.

The epistemic uncertainty part of the model acts as a canary in the coal mine in unforeseen scenarios (e.g., a new sensor type, environmental conditions/hospital-specific operational variations that are not part of the training set). This enables the system to alert for safety also for out-of-distribution data, diminishing the need to rely on historical information alone. Thus, the BDN framework constitutes a better explainable and safe option, when compared to traditional deterministic DL models, in this sense

As stated in the literature, probabilistic modeling techniques are becoming more useful in high-risk medical device applications that require reliability and decision support. In summary, the proposed BDN framework advances the paradigm of Explainable Artificial/lntelligent Systems (XAI) in clinical systems through not only its high predictive accuracy but also through quantifying and interpreting the confidence of each prediction.

5. Conclusions and recommendations

In this paper, we propose a Bayesian deep learning model to predict the failure probability of the biomedical sensors. The proposed method outperformed conventional statistical and deterministic DL methods, achieving a classification accuracy of 93.7% and an F1 Score of 0.92. Furthermore, the model demonstrated superior reliability by reducing the False Alarm Rate to 8.9%, compared to 14.2% for XGBoost and 12.1% for standard LSTM models. Especially, by integrating the Monte Carlo Dropout based uncertainty modeling, the system is able to quantitatively measure the prediction confidence, which is a great benefit for proactive maintenance scheduling.

These results indicate that estimating predictive uncertainty can strongly improve the trustworthiness of decision support systems. For clinical maintenance processes, early detection of sensor behaviors associated with failure in the near future can bring patient safety and cost of maintenance directly.

Future work can build upon this by coupling the proposed model with Physics-Informed Bayesian Neural Networks and consider the development of real-time predictive maintenance systems. Furthermore, the generalized ability of the model may also be enhanced with data privacy being guaranteed through the use of federated learning based multi-center data sharing framework.

To summarize, Bayes-based methods enable a potential and interpretable solution to problems in both the clinical engineering and industrial predictive maintenance fields, providing interpretable, uncertainty-aware and robust decision support tools for next-generation intelligent systems.

References

1. Bockholt R, Paschke S, Heubner L, Ibarlucea B, Laupp A, Janićijević Ž, Klinghammer S, Balakin S, Maitz M F, Werner C, Cuniberti G, Baraban L, & Spieth P M, Real-Time Monitoring of Blood Parameters in the Intensive Care Unit: State-of-the-Art and Perspectives. Journal of Clinical Medicine 2022;11(9):2408. https://doi.org/10.3390/jcm11092408.
2. Hirai Y, Matsuoka T, Tani S, Isami S, Kamata T., A Biomedical sensor system with stochastic A/D conversion and error correction by machine learning. In: Proc. IEEE Biomedical Circuits and Systems Conference (BioCAS) 2019: 21990-22001. doi: https://doi.org/10.1109/ACCESS.2019.2898154.
3. Wang C, He T, Zhou H, Zhang Z, Lee C. Artificial intelligence enhanced sensors: Enabling technologies to next-generation healthcare and biomedical platform. Bioelectronic Medicine 2023;9:17. https://doi.org/10.1186/s42234-023-00118-1.
4. Balaban E, Saxena A, Bansal P, Goebel KF, Curran S. Modeling, detection, and disambiguation of sensor faults for aerospace applications. NASA/TP-2009-213312 2009. https://doi.org/10.1109/JSEN.2009.2030284.
5. Li Z, He Q, Li J. A survey of deep learning-driven architecture for predictive maintenance. Engineering Applications of Artificial Intelligence 2024;133(Part C):108285. https://doi.org/10.1016/j.engappai.2024.108285.
6. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521(7553):436–444. https://doi.org/10.1038/nature14539.
7. Kendall A. and Gal, Y. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, 4-9 December 2017, 5580-5590. https://doi.org/10.48550/arXiv.1703.04977.
8. Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. Proc. 33rd Int. Conf. on Machine Learning, PMLR 48:1050–1059, 2016. https://proceedings.mlr.press/v48/gal16.html.
9. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: Contextualizing explainable machine learning for clinical end use. Journal of Medical Internet Research 2020;22(11):e21580. https://doi.org/10.48550/arXiv.1905.05134.
10. Puchalski AA. Generative modelling of vibration signals in machine maintenance. Eksploatacja i Niezawodność – Maintenance and Reliability 2023;25(4): 173488. https://doi.org/10.17531/ein/173488.
11. Martins A, Fonseca I, Farinha JT, Reis J, Cardoso AJM. Online Monitoring of Sensor Calibration Status to Support Condition-Based Maintenance. Sensors 2023;23(5):2402. https://doi.org/10.3390/s23052402.
12. Zhang H, Liu J, Li R. Fault Detection for Medical Body Sensor Networks Under Bayesian Network Model. Proc. MSN Conf. 2015:37–42. https://doi.org/10.1109/MSN.2015.21.
13. Liu Q, Wang C, Wang Q. Bayesian Uncertainty Inferencing for Fault Diagnosis of Intelligent Instruments in IoT Systems. Applied Sciences 2023;13(9):5380. https://doi.org/10.3390/app13095380.
14. Khan U, Cheng D, Setti F, Fummi F, Cristani M, Capogrosso L. A Comprehensive Survey on Deep Learning-based Predictive Maintenance. ACM Trans. Embed. Comput. Syst. 2025. https://doi.org/10.1145/3732287.
15. Duan R, Chen L, He J, Huang S. A fault location strategy based on information fusion and CODAS algorithm under epistemic uncertainty. Eksploatacja i Niezawodność – Maintenance and Reliability 2022;24(3):502–509. https://doi.org/10.17531/ein.2022.3.9.
16. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016. http://www.deeplearningbook.org.
17. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 2002;16:321–357. https://doi.org/10.1613/jair.953.
18. Blundell C, Cornebise J, Kavukcuoglu K, Wierstra D. Weight Uncertainty in Neural Network. Proc. 32nd Int. Conf. on Machine Learning, PMLR 37:1613–1622, 2015. https://proceedings.mlr.press/v37/blundell15.html.
19. Zhang Y, Marie d'Avigneau A, Hadjidemetriou G M, de Silva L, Girolami M, Brilakis I, Bayesian dynamic modelling for probabilistic prediction of pavement condition, Engineering Applications of Artificial Intelligence, Volume 133, Part F,2024,108637. https://doi.org/10.1016/j.engappai.2024.108637.
20. Kingma D P, Ba J. Adam: A method for stochastic optimization. In: Proc. Int. Conf. on Learning Representations (ICLR) 2015. https://doi.org/10.48550/arXiv.1412.6980.
21. Wang H, Wei J, Li P. Research on Fault Diagnosis Technology Based on Deep Learning. J. Phys.: Conf. Series 2022;2187:012041. https://doi.org/10.1088/1742-6596/2187/1/012041.

22. Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation 1997;9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
23. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining 2016:785–794. https://doi.org/10.1145/2939672.2939785.