ABSTRACT The Internet of Things (IoT) in the healthcare market is propelled forward by the implementation of digital systems for monitoring and analysing health problems. IoT and smart devices can contribute to a highly smart environment. Smart medical devices interconnected with smartphone apps can collect medical and other required health data. “Data Fusion (DF)” refers to integrating data and knowledge from multiple sources. However, these techniques are also applied to other domains, including text processing. Using data from multiple distributed sources, the objective of DF in multisensory contexts is to reduce the chance of detection errors and increase their reliability. The objective is to increase scalability, performance efficiency, and identification. A medical device’s ability to scale up or down demonstrates its capacity to respond to environmental factors. A more scalable system performs as expected, with no interruptions, and makes the best available use of the resource management it has. To ensure that these tracking devices all work the same way, it is essential to form a specialised group to develop uniformity in areas such as communication channels, aggregation of data, and smart interfaces. The main contribution of this research is pre-processing, DF using the Improved Context-aware Data Fusion (ICDF) algorithm, feature extraction via Improved Principal Component Analysis (IPCA), feature selection through the Enhanced Recursive Feature Elimination (ERFE) algorithm, and a classifier using an ensemble-based Machine Learning (ML) model. The Improved Dynamic Bayesian Network (IDBN) is a good trade-off for tractability, becoming a tool for ICDF operations. The simulation results show that the proposed ICDF model achieves higher performance in terms of 97% accuracy, 96% precision, 97% recall, and 97% F1 score in the healthcare system.

INDEX TERMS Enhanced recursive feature elimination, ensemble-based machine learning model, health care data, improved dynamic Bayesian network, IoT.

I. INTRODUCTION
Health care is a crucial area where ubiquitous applications may be found. Pervasive computing takes place everywhere without users’ input [1], [2]. When it refers to ubiquitous systems, a collection of devices are linked together using wired and wireless technologies to create an invisible and Artificial Intelligent (AI) system that can operate with (or) without human involvement. These interconnected devices can work independently of one another. Each device is equipped with embedded chips that allow them to connect to a network of other devices, confirming that connection is constantly accessible [3], [4], [5]. Because of this, its users can use pervasive computing in different domains, such as healthcare, homecare, transportation, etc.

Universal healthcare’s primary goal is to employ pervasive computing technology to offer healthcare to humans everywhere, at any time. It alternates with the conventional healthcare system, which involves finding symptoms, communicating with a medical expert, reporting symptoms, and receiving treatment [6]. Pervasive healthcare, on the other hand, distributes healthcare to humans wherever they are and
at any time. It is based on using sensing and communication systems to frequently monitor a patient’s health by collecting information using sensors. This enables reliable health data to be sent to a doctor (or) medical expert, allowing for prompt diagnosis and treatment of a patient’s health issues everywhere [7], [8]. Significant advances in cloud-based communication, and sensing technologies, have led to the creation of intelligent portable and wearable devices like Personal Digital Assistants (PDA) (i.e., mobile phones and smartwatches) that have allowed the implementation of a wide range of ubiquitous medical system innovations.

IoT is a novel Information Communication Technology (ICT) method that sends and receives information using communication networks [9]. IoT means a connection of many things and their communication with one another using the Internet. The IoT plays a primary segment in the development as well as in the advancement of smart systems. IoT design is divided into five levels. Every layer has its own functions that, in the meantime, give services to its top (or) bottom layers. IoT is a platform which is built on a network of physical components, devices, vehicles, buildings, and so on; all of them are formed by electronic, software, and embedded sensor systems [10], [11]. Radio Frequency Identification (RFID) labels, sensors, and smart objects link physical devices using the internet. This research uses a Context-aware Data Fusion (CDF) method to accurately improve the data prediction process.

However, Such data are structured and maintained such that physicians may access them in order to provide improved treatment for their patients. In the existing work, CDF-EMLM is introduced for improving health data treatment. However, the classifier accuracy is reduced due to the irrelevant data in the database. In order to mitigate the problem, this research work focused on developing the ICDF and Efficient Feature Selection Algorithm (EFS) for improving the classification process for predicting the healthcare data. The main contribution of this paper is pre-processing, DF using the ICDF algorithm, Feature Extraction (FE) using IPCA, Feature Selection (FS) using the ERFE algorithm and classifier using Ensemble Machine Learning (EML) model.

The proposed method provides more accurate results using hands-on algorithms for the assumed dataset [12].

The main contribution of this research is pre-processing, DF using ICDF algorithm, feature extraction via Improved Principal Component Analysis (IPCA), FS using ERFE algorithm and classifier using EML model. The proposed method provides more accurate performance measures using practical algorithms for the assumed dataset [13].

In this paper, authors introduce an ICDF method that proactively determines the driving context and implements it to sequentially fuse sensor data from different models operating at distinct levels of the model. However, this work focuses on the challenge of classifying healthcare data perceptual and cognitive systems and assumes that the proposed ICDF-EMLM method for sensor fusion can be used in many settings, from object predictive models and monitoring to epidemiological research and clinical practice [14]. Finally, these data are learnt using the EMLM for performance checking. ICDF can improve the robustness of predicting healthcare data perception using a particular sensor-DF method. Our research is the first to investigate a context-aware DF method that can dynamically fine-tune when and how DF is carried out. EML methods are meta-algorithms that integrate many ML techniques into a single prediction model. When compared with a single model, this technique provides for higher predictive performance. Here, the Enhanced Neural Network (ENN), Modified Extreme Gradient Boost Classifier (MXGB), and Logistic Regression (LR) are combined to make an EML model for predicting healthcare data [15].

This paper identifies the following areas where further study is needed:

- The analysis and features of different types of real-world data and the functionality of distinct learning methodologies will assist us in defining the scope of our research project.
- To present a deep understanding of ML algorithms that can be used to improve the knowledge and capabilities of a data-driven application.
- The goal of this discussion is to study the viability of ML-based methods in a range of practical settings.
- This paper’s goal is to review the landscape of possible future research scenarios for smart data analysis and services within the scope of our ongoing investigation.

Section 5 describes the conclusion along with further enhancement.

II. LITERATURE REVIEW

Here, some of the recent techniques for IoT-based health care applications for improving the suggested model are reviewed. And this section motivates the improvement requirements of the prediction model in this field. There are probably millions of things associated with the internet. The higher-level ones can store essential data, process information efficiently, and take appropriate decisions because of their complex and sophisticated frameworks (e.g., smart devices). Contrarily, some of them have low-level models, very minimum storage, and severely restricted computing power (e.g., body sensors). The IoT is inherently complicated due to the interdependencies between these devices. Massive amounts of device-collected data can be collected, modelled, and sensible over in order to implement data knowledge.
extracting within the IoT set-up [16], [17]. This is referred to as IoT intelligence. The sensor-based DF, context-based modelling, and context-aware reasoning add further context awareness to the image. Space considerations, including time and location, are an added important feature of the IoT. For IoT processing, the location, routine, and timescale of data are critical factors to consider when extracting and analysing context-aware information from sensor data. Tracing these devices has become progressively more difficult as the number of IoT devices continues to grow. The temporal transformation is represented in Fig. 1.

By using a method proposed by [22], it was possible to determine the similarity measure for only the variables in the study. The technique offers a solution which can be applied with standard statistical methods and software either for discrete (or) continuous data. The common factors in most marketing applications are psychographic (or) demographic. Furthermore, factors to be merged are media viewing and product purchase. A scenario like this is one where the method is able to accurately predict the distribution of media usage and online ordering in combined application. That’s an object that is used to make marketing decisions. The fusion of discrete variables is essential in marketing applications. For this situation, the researchers devised a system for loosening the requirement of conditional independence. Researchers back up their method with information from extensive surveys of British consumers about what they buy and how they use the media.

Reference [23] suggested utilizing Deep Recurrent Neural Network (DRNN), a powerful DL method founded on sequential information, as a body sensor-based method for behaviour detection. They combine data from body sensors, including electrocardiography (ECG), accelerometers, and magnetometers. Kernel Principal Component Analysis (PCA) is used to improve the FE. After that, robust features are utilized for training an activity called Recurrent Neural Networks (RNN), which is subsequently utilized to recognize behaviours. Using three publicly available standard databases, the method was compared with traditional techniques. The experiments test demonstrates that the suggested method performs better than existing methods.

The suggested method by [24] was implemented utilizing complex event processing techniques that natively support a hierarchical processing method and concentrate on managing streaming data “on the fly,” which would be a significant necessity for storage-constrained IoT devices and time-critical application areas. Preliminary results show that the suggested technique facilitates fine-grained decision-making at various DF stages, improving overall performance as well as the reaction time of public healthcare services and therefore encouraging the use of IoT technology in the healthcare industry.

Reference [25] used sensor signal fusion and case-based reasoning to categorize physiological sensor signals. The suggested method was tested, utilizing sensor DF to identify people as Stressed (or) Relaxed. During the data collection phase, physiological sensor signals such as Heart Rate (HR), Finger Temperature (FT), Respiration Rate (RR), Carbon dioxide (CO2), and Oxygen Saturation (SpO2) are gathered. Sensor fusion is accomplished in two ways: (a) at the decision-level using features extracted using traditional methods, and (b) at the data-level using FE using Multivariate Multi-scale Entropy (MMSE). The categorization of the signals is done using Case-Based Reasoning (CBR). Compared to an expert in the subject, the developed method could correctly diagnose Stressed or Relaxed individuals 87.5% of the time. Consequently, it showed potential in...
the physiological area, and it may be feasible to apply the technique to other important healthcare methods in the future.

In a Fog Computing (FC) environment, [26] suggested a DF-supported Ensemble model for working with clinical information collected from BSNs. Daily activity information is gathered from multiple sensors and combined to produce high-quality activity information. The merged information is then sent into an Ensemble classification model to forecast early cardiac disease. Ensembles are hosted in a FC platform, as well as prediction calculations are distributed. The findings from the FC environment’s different nodes are integrated to generate a single result. A new Kernel Random Forest Ensemble (KFFE) is utilised for classification, providing considerably better outcomes than a Random Forest (RF).

A comprehensive research study backs up the solution’s applicability, and the findings are encouraging, as they obtained 98% accuracy whenever tree depth is equivalent to 15, the estimator’s number is 40, and the prediction work is based on 8 features.

Andò et al. [27] developed smart systems for categorising Activities of Daily Living (ADL) that rely on data from inertial sensors fixed in a user device. A unique multi-sensor DF technique is described, which combines data from an accelerometer as well as a gyroscope. Aside from alarm management, the data given by such a system may be used to track the progression of a user’s disease, as well as during rehabilitative efforts. Overall, user testing findings show that the adopted paradigm performs well regarding sensitivity and specificity while executing falls and ADL categorization activities. The sensitivity index for categories of falls and ADL studied in this article is 0.81, whereas the specificity index is 0.98.

Medjahed et al. [28] proposed a telemonitoring method based on a multi-modal platform with multiple sensors placed at home that allows access to a complete and carefully regulated universe of collections of data. It combines physiological as well as behavioural data from the aged, their acoustical surroundings, and environmental factors with medical expertise. Specific methods process and analyse every modality. To fuse different subsystem results, a DF technique depending on Fuzzy Logic (FL) with rule collection driven via clinical rules is utilised. By identifying multiple distress scenarios, this multimodal fusion improves the overall system’s dependability. In reality, this fusion technique accounts for temporary sensor failure and improves model reliability as well as robustness in the face of environmental disruptions (or) material limitations (Battery, Radio Frequency (RF) range). FL fusion approaches provide the telemonitoring platform much flexibility, particularly when merging modalities or introducing additional sensors. The suggested telemonitoring system would provide elderly individuals with continuous in-home health monitoring.

In the context of the Narrow Band-IoT (NB-IoT), [29] proposed a multi-source heterogeneous DF formed on perceptual semantics. To begin, examine the effects along with the core technologies of NB-IoT, which include physical as well as media access control layer fundamental methods.

Furthermore, to reduce data redundancy as well as increase network lifespan, they investigated centralized and dispersed modes of the NB-IoT network and suggested a multi-source heterogeneous DF formed on semantic perception to create a consistent system. Ultimately, to continue with DF and get the final fusion outcome, an enhanced D-S evidence theory is utilized. According to the study, the suggested method has a faster rate of convergence, is more stable, and is much better at judging the results of fusion in real-world situations.

Depending on the Dempster–Shafer (D–S) theory, as well as an Adaptive Weighted Fusion Method (AWFM), [30] introduced a Data Fusion-IoT. While DF analyses the dependability of every device in the network as well as any conflicts among devices. It is done whilst taking into account the information lifespan, the distance between sensors and entities, and the need for less computing. To describe uncertain information (or) measure the resemblance between two bodies of evidence, the suggested technique employs principles founded on the Basic Probability Assignment (BPA). Murphy provides a detailed study using benchmark data simulation and real datasets from a smart building test bed to test the effectiveness of the suggested technique in contrast to D–S. Regarding reliability, accuracy, and conflict management, DF-IoT performs better than the above approaches. The system’s accuracy was up to 99.18%, 99.18% on artificial datasets, 98.87%, 98.87% on actual datasets, 0.58%, and 0.58% on conflict.

Amrouche et al. [31] proposed an application-independent approach toward the segmentation of task executions in a semi-manual industrial assembly setup by exploiting the expressive features of the distribution-based gaze feature Nearest Neighbor Index (NNI) to build a Dynamic Activity Segmentation Algorithm (DASA). The proposed approach is enriched with a ML model acting as a feedback loop to classify segment qualities. The approach is evaluated in an alpine ski assembly scenario with real-world data, reaching an overall 91% detection accuracy.

To improve the classification accuracy, [32] presented the Recursive Feature Elimination with a Cross-Validation (RFECV) approach for Type-II diabetes prediction. Dealing with overfitting issues and improving accuracy without deleting unnecessary records is the primary challenge of this approach. To predict diabetes diagnosis, it used other preprocessing techniques and following classical Machine Learning (ML) algorithms: Logistic Regression (LR), Artificial Neural Network (ANN), Naïve Bayes (NB), Support Vector Machine (SVM), and Decision Tree (DT).

A personal healthcare system that is adaptable and scalable was presented by [33]. With the help of embedded sensors, the system continuously monitors the health of the wearer. The collected data is sent to the Raspberry Pi, which acts as a processor and processes and analyses the data. This data is stored in the cloud for reasons of scalability and adaptability. Digital health care is one of the most important applications of the IoT. The healthcare sector is given a new lease on
life chance to the IoT. With the help of IoT, physicians can confidently and swiftly use pertinent patient data to take appropriate action, which is one of the best methods. The quality of medical information and patient care will be greatly enhanced as a result of this development. A concrete platform for connecting all resources and raising standards of living is provided by the IoT. When a critical condition happens, the results of the analysis are sent automatically to the doctor.

From the above discussion, some methods have good prediction performance with the experimental results. However, those methods have some issues in the DF process. So this research focuses on improving the prediction rate of health care data using improved CDFT with an efficient FS model for reducing the problem in the fusion process. The details of this method are described in the below segment (Tab. 1).

| TABLE 1. Domains for future IoT study, quick background and suggestions. |
|--------------------------------------------------------------------------------------------------|
| **Domain of IoT**                          | **Quick Background and Suggestions**                                                                 |
| Addressing and Naming Devices              | In the IoT world, an issue that needs to be resolved is how to reach specific devices efficiently. It is proposed to create social IoT networks. |
| Context-Aware Computing                     | Context-aware computing is “the ability of an information system to detect and respond to changes in its immediate physical, social, and cultural context and to adjust its behaviour accordingly.” |
| Database Management                         | Big, dispersed, and diverse IoT data can be challenging to manage.                                        |
| Interoperability and Heterogeneity          | There has been a proliferation of IoT devices in production and development in recent years. The devices’ inability to properly communicate with one another is a major source of frustration for developers. This problem is being addressed by semantic solutions. |
| AI                                          | Designing self-organized, managed, adapted, and processed IoT devices necessitates ML.                 |
| Security and Privacy                        | When it comes to the IoT, security and privacy are two of the most crucial concerns. Sub-research areas for these problems include secure authorization and authentication, IoT data, and secure access. |

A DF method developed on DL was addressed by [34] for fusing the era of digital. The application of DL in this scenario is designed to assist adaptive fuzzy in learning the hidden mappings connected with these data. Applications of this framework include smart monitoring and user-friendly graphical analytics, and authors showed how to retrieve data from IoT devices linked to human objects and accurately tag them on the objects captured by a digital camera.

The use of DF to assess travel time in wireless sensor networks for Intelligent Transportation Systems is examined in this study [35], as well as its economic benefits and limitations. Specifically, the probability-based, AI-based, and evidence-theory-based methods were evaluated for this purpose. So, researchers can learn more about how to predict travel times if they look at both the pros and cons of DF techniques.

Propose a consortium blockchain-based Personal Health Records (PHR) and sharing scheme that is security-aware and privacy-preserving, as motivated by [36]. For Medica IoT (MIoT)’s PHR symmetric encryption storage, we use the Interplanetary File System (IPFS). Zero-knowledge verification can then provide proof for validating blockchain keyword index authentication. Modified attribute-based cryptographic primitives and custom smart contracts enable secure search, privacy-preserving, and personalized access control in MIoT scenarios. Security method analysis aims to illustrate that the designed protocols meet the desired design objectives.

This research [37] set out to develop a multi-sensor platform that would include a conductive fabric-based wearable device for monitoring plant growth and a system for assessing local microclimates (i.e., temperature and relative humidity). In a real-world scenario, the platform was tested on a tobacco plant. The findings were optimistic regarding the platform’s implications for tracking plant growth and its ability to measure temperature and relative humidity.

Combining Emotional Intelligence (EI) and a sensor network based on IoT devices [38] provided a human-centred method in the health sector. An IoT device is a Raspberry Pi wired with sensors and a camera. These monitors collect biometric data from the body and EI derived from facial expressions. For reliability, security, and tamper-proof data sharing and storage, the system is hosted on an Ethereum-permissioned blockchain. Physical Unclonable Functions (PUF) are used to verify the identities of connected devices in a network. PUF-based authentication is 330% faster than traditional methods, according to a comparative analysis. A low latency of 20 ms is achievable with this system. The suggested method uses smart contracts to control access based on roles and to help build scalable and harmonious digital health care platforms.

### A. LIMITATION OF RELATED WORKS
- Instead of manual monitoring, ML can be used to analyze ECG data. The process can proceed without human involvement, which contributes to reducing costs. IoT is highly vulnerable to cyberattacks, so security must be enhanced.
- Using smart devices and a microphone increase scalability even more. The standard IoT method doesn’t involve handling much data.
- ML-based Signal monitoring involves reducing errors and improving accuracy. Increased execution time is required when more FE is from the data.
- It is necessary to improve the adaptability as well as the scalability.
- For this method to effectively process data in the real-time system, an unsupervised model needs to be developed.
- It is required that wearable devices have their batteries replaced regularly. It is necessary to reduce the overall size of the data.

### III. PROPOSED METHOD
Using IoT devices, healthcare systems may gather patient data over a long period of time. This work focused on developing an ICDF and an efficient FS algorithm for improving the classification process for predicting healthcare data.
ML algorithms such as dimensionality reduction via feature engineering, dimensionality expansion via clustering, dimensionality reduction via association rule learning, and DL are discussed here. The term “context awareness” refers to a system’s (or) component’s ability to take in data about its immediate surroundings and adjust its actions accordingly. Automatic data collection and analysis to direct actions are at the heart of what’s known as “contextual” (or) “context-aware” computing. We suggest a guideline for incorporating context into ML models [39]. A conditional probability distribution is divided into two parts: context-free and context-sensitive. Put another way, context-aware systems are at least partially autonomous because they can adjust their offerings to meet the user’s requirements without any intervention on their part. Context-aware systems must first make context information accessible to the computer. EML methods increase complexity, which reduces the model’s interpretability and makes it challenging to draw important business insights. A lengthy period of time is required for the computation and design of EMLM [40].

In this article, the authors mainly focus on ML algorithms, which are commonly used in IoT-related domains like sensor networks and context-aware systems. In the following sections, we investigate Supervised Learning (SL), Unsupervised Learning (UL), and Reinforcement Learning (RL) techniques. We only discuss IoT-related and context-aware ML-related algorithms and techniques here because ML has such a broad range of applications [41]. The classifications of IoT-specific and general ML algorithms are depicted in Fig. 3.

The IoT-DF middleware is divided into four submodules: data receiving and computation, knowledge inference, user-related acquisition and adding new features, and service decomposition and performance [42]. The architecture of the system is depicted in Fig. 4. Using the smart library’s IoT and other sensing devices, the DF middleware’s receiving and processing module can collect and process data in real-time. In order to convert the raw data from the bottom layer into meaningful events, the knowledge reasoning module must first apply the rules to the data. Reference [43] and [44] Whenever an event is sent from the knowledge reasoning module, it is uniformly processed by the event description module into a simple form. In order to get the lower-level hardware devices to carry out the demand from the top layers, the event decomposition module is used to translate those requirements into commands the system can accept [45]. Fig. 5 depicts the system’s functional decomposition diagram, which represents the proposed methodology’s flow process step-by-step.

- **Data Fusion:** The data from IoT devices will be gathered and preprocessed to clarify the fusion processing. In this work, a Dual Filtering Method (DFM) for data preprocessing is introduced, which attempts to label the unlabelled attributes in the data gathered so that DF can be done accurately.
- **Improved Dynamic Bayesian Network:** IDBN is a good trade-off for tractability and becoming a tool for ICDF operations. Here, the inference problem is handled using the Hidden Markov Model (HMM) in the Deep Belief Network (DBN) model.
- **Improved Principal Component Analysis:** IPCA is used to pull out features and reduce the number of dimensions.
- **Feature Selection Process:** The FS process uses the ERFE method to eliminate irrelevant data in a dataset.
- **Ensemble-Based Machine Learning Model:** This data was learned using the EMLM model for performance checking. The ENN, MXGB, and LR are combined to make an ensemble mode for predicting healthcare data.
A detailed explanation of the proposed method is elaborated in the next part. Fig. 6. Shows the overall process of the suggested methodology.

FIGURE 6. Overall process of the proposed model.

A. PREPROCESSING
Data collection from physical devices is the initial stage in context acquisition. Next, to eliminate noise as well as other measurement outliers, measurement preprocessing is performed using filtering and estimating methods. To correct incorrectly labelled data, classification models that operate on clean data sets can be single classifiers or ensemble models are used after the filtering process. ‘Cleaned data sets’ refers to data elements that remain after cleaning (or) filtering. The preliminary filtering performed in this work removes noise from data instances before filters are applied. In this research work, a DFM is introduced to process the data. The Kalman Filter (KF) is a statistical method to measure a system’s state, while the Particle Filter (PF) is an unexpected way to measure a system’s moments.

B. KALMAN FILTER
The KF is a widely utilized Statistical State Estimation Technique (SSET) for fusing dynamic signal level information. A system’s state estimations are calculated using a recursively implemented prediction and update method, which believes that the present state of a system is dependent on the preceding time interval’s state. One of its many benefits is its computational efficiency. KF, for example, has frequently used combined accelerometer and gyroscope information to provide good estimation [46]; for example, KF could be used to identify postural sway during quiet standing (standing in one place while doing no other activity) while leaning on something.

C. PARTICLE FILTERING (PF)
Whenever target probability density moments cannot be calculated analytically, PF is a stochastic method to approximate them. Its idea is to create random numbers named particles from a readily sampled “importance” distribution. Next, for every particle, a weight is assigned that compensates for the difference between the goal and the significance probability. PFs are frequently employed in Bayesian contexts to determine the posterior density mean. They have the advantage of whole target distribution estimation without making any assumptions, making them especially helpful for nonlinear and non-Gaussian systems. PF may be used to check biomechanical conditions depending on the accelerometer and gyroscope information.

These two filtering steps make a significant contribution to noise-free, effectively labelled clean data for further analysis. Assuming each data instance x’ represents a multi-label set l, When dealing with Binary Classifications (BC), Let’s assume that C(+) represents positive labels and C(-) denotes negative labels in the set l’. Equation (1) and Equation (2) apply if pb(+)/pb(-) are positive and negative label probabilities in the set l’:

\[ pr(+) = \frac{C(+) + 1}{C(+) + C(-) + 2} \]  
\[ pr(-) = \frac{C(-) + 1}{C(+) + C(-) + 2} \]  

for a Laplace correction that is applied.

The margin between classes \(|pr(+) - pr(-)|\) is small if C(+) is very close to C(-). when items are mislabeled, this can happen in two ways. when labelling complex instances, the second case occurs. as a result, if \(|pr(+) - pr(-)|\) is small for an instance ‘x’, inference algorithms may not be able to integrate the instance and must be filtered. algorithm 1 is used in the proposed work, which is listed below. steps to 7 use \(|pr(+) - pr(-)|\) to perform preliminary filtering. noiseless data is returned from lines 14 TO 16 and correct labels from steps 9 to 13, while the second level of filtering is contained in steps 8 and 9.

D. IMPROVED CONTEXT-AWARE DATA FUSION (ICDF)
DBN is a good trade-off for tractability in this work, and it becomes a tool for DF operations. DBNs are used in this study to uncover the effects of context variables on environments without being bound by probability distributions. They use HMMs to discover the observable symptoms of an instance by dividing data into time slices that represent the states of the instance. The hidden variable ‘Vt’, abbreviated as DBN, is primarily used to infer the states of a known feature of interest. Sensory readings and their contexts are used to update the system. St=(S_t^1,...,S_t^n) is the set of sensor’s readings active in the ‘t’ time slice, and the contexts set is represented by C_t=(C_t^1,...,C_t^n) based on the application’s environment. The size of the Conditional Probability Table (CPT) and learning in the training phase are controlled by limiting the number of context variables.

Sensors and state transitions must be defined in DBN. Pb(S_t^1|V_t) represents how sensor information is affected by the system’s current state (or) sensor model, whereas Pb(V_t^1|V_{t-1}, C_m) represents the likelihood that a state variable has a specific value given its initial value and current context. The utilised DBN is a first-order Markov model and a given system state within the ‘t’ time slice; thus, ‘V_t’ can be defined
Algorithm 1: Pre-processing noisy dataset for noise reductions

```
Input: \( \hat{D} = \{(x_i, \hat{y}_i)\}_{i=1}^{N} \) a training data set with integrated labels; \( \{l_i\}_{i=1}^{N} \) multiple label sets of \( \hat{D} \); \( \delta \)-threshold
Output: \( \hat{D} \) the corrected data set

Step 1. A is a void set
Step 2. For \( i = 1 \) to \( N \), Do
Step 3. Account for the numbers of (+ve/-ve) label in \( l_i \), i.e., \( N(+) \) and \( N(-) \), correspondingly
Step 4. Analyze \( p(+) \) and \( p(-) \)
Step 5. If \( |p(+) - p(-)| \leq \delta \) Then
Step 6. The case ‘i’ is additional to the set A
Step 7. End For
Step 8. A PF to the Set \( \hat{D} \). ‘A’ and examples filtered out by the filter include a set ‘B.’
Step 9. \( \hat{D}_c = \hat{D} - A + B \)
Step 10. Build a classification model ‘f’ on the set \( \hat{D}_c \)
Step 11. For \( i = 1 \) to size of \( (A + B) \), Do
Step 12. The classifier of ‘f’ to relabel, for example, ‘i’ the set A+B
Step 13. End For
Step 14. Update the set A+B to \( \hat{A} + \hat{B} \) with corrected labels
Step 15. \( \hat{D} = \hat{D}_c + \hat{A} + \hat{B} \)
Step 16. Return \( \hat{D} \) as the corrected data set
Step 17. End For
Step 18. End For
Step 19. End
```

as the following Equation (3):

\[
\text{Bl} (v_i) = \text{Pb}(v_i | S_{1:t-1}, Cn_{1:t})
\]  

For the practical formulation of belief, a Bayes Filter-like procedure is applied using the Bayes rule, and it is possible to state Equation (4)

\[
\text{Bl} (v_i) = \text{Pb}(v_i | S_{1:t-1}, Cn_{1:t}) = \text{Pb}(v_i | S_{1:t-1}, S_t, Cn_{1:t}) \times \eta . \text{Pb}(S_t | v_i, S_{1:t-1}, Cn_{1:t}) \times \text{Pb}(v_i | S_{1:t-1}, Cn_{1:t})
\]

where “Normalizing Constant” is used. The sensor nodes in ‘S’ do not rely on \( C_n \) context variables in a Markov assumption; in a ‘S’ state variable, and assuming sensor measured values are mutually independent, parent node value ‘S’ can be computed using Equation (5)

\[
\text{Pb}(S_t | v_i, S_{1:t-1}, Cn_{1:t}) = \frac{\text{Pb}(S_t | v_i, Cn_{1:t})}{\text{Pb}(S_t | v_i)} = \prod_j \text{Pb}(s_j^t | v_i)
\]

where \( s_j^t \) is the specific value of the ‘i’ sensor in a ‘t’ time slice, Equation (6).

\[
\text{Pb}(v_i | S_{1:t-1}, Cn_{1:t}) = \sum_{v_{t-1}} \text{Pb}(v_i, v_{t-1} | S_{1:t-1}, Cn_{1:t}) \times \eta . \text{Pb}(v_{t-1} | S_{1:t-1}, Cn_{1:t}) = \alpha \sum_{v_{t-1}} \text{Pb}(v_i | v_{t-1}, S_{1:t-1}, Cn_{1:t}) \times . \text{Pb}(v_{t-1} | S_{1:t-1}, Cn_{1:t})
\]

where \( \alpha \)-normalizing constant. \( C_n \) could be sensibly neglected from the last term, as \( v_{t-1} \) does not rely on \( C_n \) following context when the next state ‘V’ is not considered. Therefore, using the Markov model stated in Equation (7):

\[
\text{Pb}(v_t | S_{1:t-1}, Cn_{1:t}) = \alpha \sum_{v_{t-1}} \text{Pb}(v_{t-1} | v_{t-1}, Cn_{1:t}) \times . \text{Pb}(v_{t-1} | S_{1:t-1}, Cn_{1:t-1}) = \alpha \sum_{v_{t-1}} \text{Pb}(v_{t-1} | v_{t-1}, Cn_{1:t}) \times . \text{Pb}(v_{t-1} | S_{1:t-1}, Cn_{1:t-1})
\]

belief is described with a recursive Equation (8)

\[
\text{Bl} (v_i) = \eta \sum_{v_{t-1}} \text{Pb}(s_j^t | v_i) \times \text{Pb}(v_{t-1} | v_{t-1}, Cn_{1:t})
\]

where \( \alpha \)-integrated with normalization constant ‘\( \alpha \)’.

The inference is executed via storing DBN in two slices where time, as well as space, updating the network’s beliefs, is not dependent on the length of the sequence. The Computational complexity is \( O(n + m) \), where ‘\( n \)’ represents sensors, ‘\( m \)’ represents possible \( v_t \) values, and the overall complexity of \( \text{Bl}(v_i) \) for all \( v_t \) is \( O(m^2 + m^2) \).

However, the inference issue in DBNs is identical to that in BNs, in which the desired quantity is the posterior marginal distribution of a collection of hidden variables with an order of observations (up to date of belief): \( P(X_0[t] \mid X_0[0], \ldots, X_0[t]) \) here \( X[t] = \{X_h[t], X_o[t]\} \) is a collection of time evolution variables where \( X_0[t] \) and \( X_h[t] \) specify observed variables as well as hidden, correspondingly. Time-series inference is classified as filtering (\( \tau = 1 \)), smoothing (\( \tau > 1 \)), (or) forecasting (\( \tau < 1 \)) based on the time frame of observation used in calculations.

Building a large static BN for required time section numbers and then using standard methods of inference for static BNs provides a direct method to imply probabilities in a DBN. Nevertheless, this necessitates knowing the end of a time interval in advance.

Furthermore, the data-processing complexity of this technique might necessitate a significant time (memory). As a result, DBN inference is often performed via recursive operators that revise DBN’s belief state as new observations become accessible. A message-passing mechanism for static BNs is identical in principle. So in this work, the Hidden Markov Model is used to mitigate this problem. The aim is to use messages specified on Markov’s cover of variables to D-separate the past from the future and use a forward-backwards technique to distribute the entire obviousness along DBN, as previously discussed.

E. HIDDEN MARKOV MODEL (HMM)

Each HMM consists of finite states (N) with its own probability distribution. Changes between states are controlled by a set of probabilities called “transition probabilities.” The
following processes must be completed to build up a word recognition model that is formed on HMMs: (a) Choose some states and observations, (b) choose an HMM topology, (c) choose training and samples, (d) train the system using training data, and (e) test it using testing data. Fig. 7 depicts the example of a 7-state; HMM that only allows transitions to the same state \( \rightarrow \) next state \( \rightarrow \) subsequent state [47]. Every state has transitions to the same state \( \rightarrow \) next state \( \rightarrow \) next state. During training and testing, the order in which the model’s states change is set by the FE described in the previous section.

![FIGURE 7. Hidden Markov Model (HMM) 7-State.](image)

HMMs are often designated by the letter ‘k’ and are specified by the following parameters set \( k=(p, A, B) \). Here \( (A, B, p) \) are parameters.

1) TRANSITION PROBABILITY (A) MATRIX

\[
A = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\tag{9}
\]

\[a_{mn} = P(S_n | S_m) ; m, n = 1, 2
\tag{10}
\]

where \( a_{mn} \) represents the probability that \( a_{mn} \) the current state provided by \( 'S_m' \). It is computed as a ratio of transitions that are expected from \( (S_m \rightarrow S_n) \) state of transitions predictable out of state \( 'S_m' \) (Equations (9) and (11)).

2) EMISSION PROBABILITIES (B) MATRIX

\[
B = \begin{bmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23}
\end{bmatrix}
\tag{12}
\]

\[b_{np} = b_n (p) = P(O_p | S_n) ; n = 1, 2; p = 1, 2, 3
\tag{13}
\]

Here \( b_n (p) \) is the likelihood that \( O_p \) present observation is given \( 'S_n' \) present state (Equation (12)). It is computed as a ratio of the expected number of times where \( O_p \) served with \( 'S_n' \) to the expected number of times in \( 'S_n' \) state (Equations (13) and (14)).

3) INITIAL STATES PROBABILITIES (\( \pi \))

\[
\pi = \begin{bmatrix}
\pi_1 \\
\pi_2
\end{bmatrix}
\tag{15}
\]

\[\pi = \{\pi_1 | \pi_1 = P(S_1 = i)\}
\tag{16}
\]

\[\pi_m = P(S_m) ; m = 1, 2
\tag{17}
\]

Due to the many advantages of the HMMs, the parameters \( k=(A, B, p) \) adjustment procedure is as (Equation (15) (16) and (18)).

**Step. 1** Set parameters \( \lambda = (A, B, \pi) \) arbitrarily: \( a_{mn} \) is set to 1/N, \( b_{np} \) is set to 1/M, \( \pi_m \) is set to 1/N.

**Step. 2** By the above equations, compute parameters \( \alpha_t(m), \beta_t(m), \xi_t(m, n), \gamma_t(m) \).

**Step. 3** Compute the model’s novel parameters \( \lambda^* = (A^*, B^*, \pi^*) \) as per values computed in the 2nd step:

**Step. 4**

\[
a_{mn}^* = \frac{\sum_{t=1}^{T} \xi_t(m, n)}{\sum_{t=1}^{T} \gamma_t(m)}
\tag{18}
\]

\[
b_{n}^*(p) = \frac{\alpha_t=O_p}{\sum_{i=1}^{T} \gamma_t(n)} \cdot \pi_m^*\gamma_1(m)
\]

**Step. 5** Compute \( P(V | \lambda^*) \) as probability \( P(V | \lambda^*) \) is rising, repeat 2 and 3 steps.

The parameters will describe a model that best matches the training observation sequences once they converge to particular values. Because the streams (DBN/HMM) are synchronised at each time slice, all observation sequences in the proposed ICDF model should work better than the current DF process.

**F. FEATURE EXTRACTION**

To conduct self-optimization, FE chooses important context information about the environment and internal state of the system. These features may be deduced from the DF module’s output and through a meta-analysis of the system’s internal activity. The objective of FS is to drive the self-configuring continuously component’s behaviour in response to context changes. For dynamic circumstances, static criteria such as frequency of reconfiguration (or) even the system’s overall goals are unsuitable. This level is split into two halves. The first performs measurements in the time (or) frequency domains. The signal (or) preliminary information required to determine the features used to make these measurements. Next, it gathers parameters from every measurement and combines them with data from the categorization issue. Fig. 8 depicts the FE procedure utilizing the suggested method.

![FIGURE 8. FE process.](image)

**G. IMPROVED PRINCIPAL COMPONENT ANALYSIS (IPCA)**

PCA is primarily utilized for exploratory data analysis and the creation of prediction models. It’s frequently used to show how genetic distance and relatedness among populations are shown. After normalising the raw data with eigenvalue decomposition of the data covariance (or) correlation; matrix (or) singular value decomposition of a data matrix, PCA is usually performed. Every attribute’s normalization entails subtracting all data values from all variable’s measured mean to arrive at a zero empirical mean (average); optionally, every variable’s variance is normalizing for making 1,
read Z-scores. Component scores, otherwise factor scores (transformed variable values relating to a specific data point), and loadings are commonly used to describe the outcomes of a PCA (the weight through which every original standardized variable has to be multiplied to obtain a component score). When component scores are normalized to unit variance, the loadings must be included in loadings (i.e., eigenvalues magnitude). When component scores are not standardized and (so include) data variance, loadings should be unit-scaled ("normalized"). All these weights are termed eigenvectors; they are the cosines of the orthogonal rotation of variables to their principal components (or) back [48].

Its function may be regarded as disclosing the internal structure of data in such a way that perfectly illustrates the data’s variation. Whenever a multivariate database is shown as a collection of coordinates in a high-dimensional data space, PCA can provide the viewer with a lower-dimensional image or projection of an object from its most relevant viewpoint. It is accomplished by reducing the dimension of the modified data by only utilizing the first few principal components. Whenever the data set contains outliers, nevertheless, PCA suffers primarily as a result of an investigation. In a setting with a massive amount of data, separating outliers is difficult. This problem is solved by constructing the Adaptive Gaussian Kernel Matrix (AGKM).

H. THE ADAPTIVE GAUSSIAN KERNEL MATRIX CONSTRUCTION

Assume a collection of s nodes $V = \{v_i, 1 \leq i \leq s\}$, each of which may interact with the central coordinator ‘$v_0$’ in a distributed setup. A Local Data Matrix (LDM) $P_i \in \mathbb{R}^{n_i \times d}$ with ‘$n_i$’ data points in ‘$d$’ dimension ($n_i > d$) exists on every node ‘$v_i$’. The Global Data Matrix (GDM) $P \in \mathbb{R}^{n \times d}$ is then formed by concatenating the LDM, that is $P^T = \left[ P_1^T, P_2^T, \ldots, P_s^T \right]$ and $n = \sum_{i=1}^{s} n_i$. The $i^{th}$ row of $P$ is denoted by ‘$p_i$’. Assume the data points are centred so that the mean is zero. That is, $\sum_{i=1}^{s} p_i = 0$. Uncentered data necessitates a rank-one change to methods, whose communication and computation expenses are conquered by the costs of other stages.

Assume a nonlinear transformation ‘$\phi(x)$’ from the original D-dimensional feature space to the M-dimensional feature space. Here, generally $M > D$. After that, every data point ‘$x_i$’, a point $\phi(x_i)$ is projected. Do conventional PCA in the new feature space. However, be aware that this might be both expensive and inefficient. However, make the calculation clearer by using kernel techniques. Basic notation and equations (19), (20), and (21) are described.

The benefit of using kernel techniques is that you don’t have to calculate manually ‘$\phi(x_i)$’. Directly generate the kernel matrix from the training data set $x_i$

$$k(x, y) = (x^T y)^d$$  \hspace{1cm} (19)

or $k(x, y) = (x^T y + c)^d$, \hspace{1cm} (20)

where $c > 0$ is constant, and the Gaussian kernel

$$k(x, y) = \exp(-\|x - y\|^2/2\sigma^2)$$ \hspace{1cm} (21)

$\sigma$ is an AGKM parameter.

Whilst $k=1$ and the centre is an r-dimensional subspace, PCA is a unique case. Top ‘$r$’ right singular vectors of ‘$P$’, defined as key components, span this optimum r-dimensional subspace, which may be determined via the Singular Value Decomposition (SVD).

I. FEATURE SELECTION USING ENHANCED RECURSIVE FEATURE ELIMINATION (ERFE)

Both approaches look for fewer input variables for predictive systems, and FS and dimensionality reduction techniques are linked. Differentiation is that FS selects which features to preserve (or) remove from a database, while dimensionality reduction produces new input features via projecting the information. Hence, dimensionality reduction is a form of FS instead of a substitute.

Furthermore, a smaller set of features may provide a better understanding of the system to be trained as well as a computation speedup. Other advantages might include cost-effectiveness, such as in biological applications where a smaller subset of attributes must be assessed to diagnose a disease with similar accuracy. The RFE technique uses ML’s generalization capabilities, making it ideal for minor sample issues.

Feature Selection Wrapper Technique (FSWT): FSWT removes redundant and weak features that have the most negligible impact on the training error while keeping the model independent and robust features to enhance the model’s generalization performance. It employs an iterative feature ranking process, an example of backward feature removal. This method begins by creating a model based on the complete collection of features and then ranking each feature by its relevance. The model is rebuilt, and the feature importance is recalculated after removing the least significant feature. To save the feature rating, assume ‘$T$’ is a sequence number. ‘$T_i$’ contains the top-ranked features upon which model refit and performance are accessible throughout every iterative round of backward feature reduction. A best-performing ‘$T_i$’ value is calculated, and top-performing features are matched to the final model. Algorithm 2 depicts the ERFE method procedure for removing unnecessary information.

**Algorithm 2 For Enhanced Recursive Feature Elimination**

**Input:** Dimensionality Reduced data

**Output:** Relevant features (discard the irrelevant data)

**Step 1.** Train the classification model by entire features with cross-validation

**Step 2.** Compute model performance

**Step 3.** Compute feature ranking

**Step 4.** For every subset $[T_i, i=0,1,2,3, \ldots, n]$ Do

**Step 5.** Maintain the $T_i$’s most vital features

**Step 6.** Update the Adaptive Learning (AL) function stage

**Step 7.** Re-compute the model’s performance

**Step 8.** Re-compute the ranking importance of every feature

**Step 9.** End

**Step 10.** Determine the optimal number of features
The mean has substituted some missing values in the specified features. Following the selection of the above features, the testing was performed using the entire dataset. For efficient learning, features are scaled at a unit variance. However, traditional RFE has some limitations, such as it does not concern the next state of the feature process practically. So, this work is mostly about making ERFE, a new method that changes the rules for removing features from each state.

1) ADAPTIVE LEARNING FUNCTION-BASED RFE

In this, the AL function $\phi$ is introduced for ordering the features as per their importance (high-value means highly vital). The primary difference between the new technique and the old RFE is that the original RFE has no regard for the future state, whereas ERFE will keep redundant (or) weak features that can be merged with other features. So, the suggested ERFE improves generalization accuracy, especially when there are few features.

$$|w_j| = \left| \sum_{i=1}^{n} \alpha_i y x_j \right| \quad (22)$$

where $x_j$ is the $j^{th}$ feature vector’s $i^{th}$ element. Following the calculation of Equation (22) for every one of the P features, they may be ranked in order of significance (high-value means more significant). A certain number of characteristics from the bottom of the sorted list are eliminated in the next phase. This procedure of removing a certain number of uninformative features is continued until the remaining features are empty. To figure out the best number of features, a performance measure must be made for each attribute trained on a certain number of features.

An AL function ‘$\phi$’ was used to combine the weights, which were specified as follows:

$$\phi(w_j, r_i) = w_j \frac{1}{\text{rank}(r_i)} \quad (23)$$

$$\text{rank}(r_i) = |r_i / r_i \geq r_i| \quad (24)$$

where $W_i$ is the RFE weight generated by Equation (23), and ‘$r_i$’ is the outcome of attribute Rank for the $i^{th}$ feature in Equation (24), the rank function was used to transform the actual ranks provided by findings into a rank-based. This implies that the feature with the Top Rank receives a weight of 1, the feature with the $2^{nd}$ Top rank receives a weight of 2, the feature with the $3^{rd}$ Top rank receives a weight of 3, and so on, up to P. The modification was carried out to avoid single attributes having a weight significantly more significant than the others, which could occur if these attributes had a considerably higher degree (interconnection with other attributes) than others.

IV. ENHANCED DEEP NEURAL NETWORK (EDNN)

DL is a successful approach for producing highly accurate predictions from complicated data sets. In this research work, you will design an improved NN model. Combining a fuzzy inference system with a DNN is the basis for this work. A Fuzzy Neural Network (FNN) with several hidden layers underlying this model. A FNN is a learning method that incorporates Neural Network (NN) methods to apply the attributes of FL systems [49]. In this approach, a DNN with n hidden layers is employed, as the value of ‘n’ may be fixed during training. Increasing the number of hidden layers in the network will improve the accuracy of the outcome, but increasing the number of hidden layers increases the system’s complexity and decreases training performance.

As a result, choosing hidden layer numbers is also part of the training procedure. Based on the objectives of an analyst, a DL framework may be designed to handle raw data such as images or 1-D signals. Without the requirement to label the data, the weights may be learnt iteratively in an unsupervised way. After these networks have been trained, every neuron’s activity may be evaluated for its response to ‘+ve’ and ‘-ve’ stimuli. A specific input may contain multiple labelled features. This validation process can also be used to fine-tune the neurons in the network by using backpropagation. To summarize the goal of utilising DNN, it should be noted that choosing the parameters of the provided data is a big challenge, and since many of these parameters are chosen based on user experiences or trial-and-error, an overfitting problem is introduced in the process. A computational method can be used to automate this procedure while minimising errors. As a result of this study, a Fuzzy Inference Method (FIM) for reducing this issue was developed. The neurons that respond can be fed into a FIM to help figure out which features are good.

- Integration WITH FIM

Complex non-linear situations can be modelled using if-then statements employing FIM algorithms. The benefit of structured rule-based algorithms is that they may be influenced by subjective data. This allows an analyst to give expert knowledge to the system, perhaps enhancing categorization findings (or) altering the system’s behaviour [50]. This feedback bias may also help autonomous systems learn faster while preserving stability.

FIS will further analyse the features derived from the data using DL, causing the system to replicate human reasoning while also providing a way of biasing the system with feedback from an analyst. Both feature vector inputs, as well
as the correctly labelled output, would be needed to train the system. The extension concept is used to expand DL models to include FL. Fuzzy aggregation, which accounts both for the granularity of the data collection and as a way to manage incomplete data from every modality, is enabled by the inclusion of FL. These processes also allow for a subjective assessment of every modality’s value. Linguistic statements given by the user are converted into symbolic notations that may be used to initiate NN. These may be seen in metrics such as the number of neurons per layer as well as the overall NN model’s number of layers. The fuzzy inference block is then driven by NN. Fuzzy rules, for example, can be modified to integrate specialized knowledge about the input data. This system has the advantage of using fuzzy rules to describe the behaviour of a fuzzy system. The FIM block’s input pattern is analysed by the learning algorithm block. Adjusted system weights are backpropagated through the system, allowing the system to adapt over time by automatically altering the behaviour of the predictors.

Assume the rules comprise three Takagi and Sugeno-style fuzzy If-Then rules.

• **Rule 1:** If \((a \text{ is } X_1, b \text{ is } Y_1, c \text{ is } Z_1)\) Then
  \[ f_{i1} = p_{1a} + q_{1b} + t_{1c} + r_1 \]

• **Rule 2:** If \((a \text{ is } X_2, b \text{ is } Y_2, c \text{ is } Z_2)\) Then
  \[ f_{i2} = p_{2a} + q_{2b} + t_{2c} + r_2 \]

• **Rule 3:** If \((a \text{ is } X_3, b \text{ is } Y_3, c \text{ is } Z_3)\) Then
  \[ f_{i3} = p_{3a} + q_{3b} + t_{3c} + r \]

The below expression defines the Fuzzy based NN mathematical operations, Layer I. An adaptive node with function node is included in Equations (25), (26) and (27).

\[
O_{1,i} = \mu_{A_1}(x), \text{ for } i = 1, 2 \tag{25}
\]

\[
O_{1,i} = \mu_{B_{i-2}}(y), \text{ for } i = 3, 4 \tag{26}
\]

\[
O_{1,i} = \mu_{C_{i-4}}(z), \text{ for } i = 5, 6 \tag{27}
\]

where \(\mu_{A_1}(x)\), \(\mu_{B_{i-2}}(y)\) and \(\mu_{C_{i-4}}(z)\) are any acceptable parameterized (MFs, \(O_1\)) is membership grade of a fuzzy set \(A = (A_1, A_2, B_1, B_2(\text{or})C_1, C_2)\), which shows the degree to which supplied input \((x,y,z)\) satisfies the quantifier. “Premise Parameters” is the name of this layer.

Furthermore, any suitably parameterized MF, such as a generalized “bell function”, can be used as the membership function for Equation (28):

\[
\mu_A(x) = \frac{1}{1 + |(x - c_i)/a_i|^b} \tag{28}
\]

where \(a_i, b, c_i\) are the set of parameters. The Bell function fluctuates as the values of parameters vary, revealing the different types of MF for fuzzy sets in the process. This type of membership function is also employed in the latest research.

Layer I : Each node in this layer is a fixed node, with an outcome equal to the sum of all incoming signals (Equation 29).

\[
O_{2,i} = \mu_{A_1}(x) \mu_{B_{i}}(y) \mu_{C_{i}}(z), i = 1, 2, 3 \tag{29}
\]

Every output node indicates a rule’s “Firing Strength”.

Layer II: The ‘N’ function of normalization has a fixed node labelled N Equation (30):

\[
O_{3,i} = \frac{w_i}{w_1 + w_2 + w_3}, i = 1, 2, 3 \tag{30}
\]

The results are generally known as “Normalized Firing Strengths”.

Layer IV: Adaptive nodes are included in Equation (31):

\[
O_{4,i} = \tilde{w}_f_i = \tilde{w}_i(p_1x + q_1y + t_3z + r_i) \tag{31}
\]

Because each node in this layer is a multiplication of the 3rd layer’s Normalized Firing Strength and the outcome of DNN, it’s termed “Consequent Parameters.”

Layer V : It contains a signal fixed node designated ‘S’ with a summing function that calculates the DNN network’s total output as the sum of all incoming signals Equation (32).

\[
\text{Overall output } O_{5,i} = \sum_i \tilde{w}_f_i = \frac{\sum_i w_f i}{\sum_i w_i} \tag{32}
\]

Finally, the enhanced DNN model provides the best results.

V. MODIFIED EXTREME GRADIENT BOOST (MXGB) CLASSIFIER

MXGB is ML model for classification and regression problems based on the Gradient Boosting Decision Tree (GBDT) [51]. The inner nodes of the regression tree contain attribute test values, whereas the leaf nodes with scores indicate a judgment. Fitting the model requires both the 1st-order and 2nd order derivatives of the prediction loss functions. This is because XGBoost uses an AL method with a 2nd order approximation.

To provide a clear process, this work first develops the additive tree boosting 2nd order approximation. The volume of data is referred to as ‘m’, while the number of features is ‘n’. The ‘raw prediction’ before the sigmoid function will be represented as ‘z_i’ as well as the probabilistic prediction will be \(\hat{y}_i = \sigma(z_i)\), in which the sigmoid function is represented by \(\sigma\). It’s crucial to remember that the notations differ, and the ‘\(\hat{y}_i\)’ n that their analysis is indicated as ‘z’. The true label is denoted by ‘\(y_i\)’, while the parameters for the two-loss functions are denoted by \((\alpha, \gamma)\) correspondingly. Gradients and Hessians’ expressions are recorded in a merged form independent of the value of ‘\(y_i\)’, as this simplifies programme implementation and facilitates vectorization in other programs.

In practice, the AL goal is Equation (33):

\[
\mathcal{L}(\theta) = \sum_{i=1}^{n} \left( y_i, z_i^{(0)}(x_i) + f_i(x_i) \right) + \Omega(f_i) \tag{33}
\]

\(t^{th}\) iteration of the training process is denoted by ‘\(t\)’. In the equation, noticed that the notations have been replaced. When
Equation (34) is worked with the 2nd-order Taylor expansion, the following results are found:

\[ L^{(t)} \approx \sum_{i=1}^{n} \left[ l(y_i, z_i^{(t-1)} + g f_i(x_i)) + \frac{1}{2} h_i(f_i(x_i))^2 + \Omega(f_i) \right] \]  \tag{34}

\[ \propto \sum_{i=1}^{n} \left[ g f_i(x_i) + \frac{1}{2} h_i(f_i(x_i))^2 \right] + \Omega(f_i) \]  \tag{35}

The last line is because \( l(y_i, z_i^{(t-1)}) \) term may be omitted from the learning goal because it does not affect the model fitting in the \( t \)-th iteration, Equation (35).

Hand-derived derivatives will be necessary because XGBoost does not enable automatic differentiation. Furthermore, the obtained expressions might be used in various ML applications. The activation for both loss functions is sigmoid, as the following fundamental characteristics of sigmoid will be applied consistently in derivatives Equations (36) (37) and (38):

\[ \frac{\partial \hat{y}}{\partial z} = \frac{\partial \sigma(z)}{\partial z} = \sigma(z)(1 - \sigma(z)) \]  \tag{36}

\[ = \frac{\hat{y}(1 - \hat{y})}{2} \]  \tag{37}

\[ \frac{\partial \sigma(z)}{\partial z} = \sigma(z)(1 - \sigma(z)) \]  \tag{38}

**A. REGULARIZATION-BASED ADAPTIVE FACTOR (RA)**

Next, methods are utilized to reduce overfitting in addition to the regularised objective. After every stage of tree boosting, the regularization-based AL function adds new weights by a factor of \( RA_{w_j} \). Adaptive reduces the effect of each tree and makes room for future trees to improve the model. This is similar to how a learning rate is used in optimization.

Calculate the optimum weight ‘\( RA_{w_j} \)’ of leaf \( j \) for a fixed structure \( q(x) \) by Equation (39)

\[ RA_{w_j} = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \]  \tag{39}

where ‘\( \epsilon \)’ is a factor of an estimate, this implies that there are approximately 5 chosen points. Every data point is weighted by the ‘\( h_i \)’ that indicates the weight; hence, EQU (34) may be rewritten as Equation (40)

\[ \sum_{i=1}^{n} h_i(f_i(x_i) - g_i/h_i)^2 + \Omega(f_i) \]  \tag{40}

Define \( I_j = \{ i | q(x_i) = j \} \) as an example, the set of leaves ‘\( j \)’. Then, by extending ‘\( \Omega \)’ as follows, it may be rewritten as Equations (41) and (42)

\[ L^{(t)} = \sum_{i=1}^{n} [g f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i)] + \Omega(f_i) \]  \tag{41}

\[ L^{(t)} = \sum_{i=1}^{n} [g f_i(x_i) + \frac{1}{2} h_i^2 f_i^2(x_i)] + \frac{1}{2} g_i^{2T} w_i \]  \tag{42}

With labels \( g_i/h_i \) and weights ‘\( h_i \)’, this is precisely a weighted loss. Finding candidate splits that meet the requirements is difficult in large datasets.

**VI. LOGISTIC REGRESSION (LR)**

The LR model used, which is a popular ML algorithm that is frequently used in real-world applications like data mining [52]. For example, in this study, the risk factors for provided data were examined, and the likelihood of occurrence was predicted depending on the risk factors. The most common use of logistic regression is for classification, mainly two-category problems (i.e., there are only two types of outcome, every representing one category), which may determine the possibility of each classification event occurring. The following is a representation of a LR model, Equation (43):

\[ prob(Y = 1) = \frac{e^z}{1 + e^z} \]  \tag{43}

where \( Y \) is a binary dependent variable (\( Y=1 \) If an event happens; \( Y=0 \) otherwise), ‘\( e \)’ is the foundation of natural logarithms, whereas ‘\( Z \)’ is \( Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \) with ‘\( \beta_0 \)’ constant, \( \beta_j \) coefficients and ‘\( X_j \)’ predictors for \( p \) predictors (\( j=1, 2, 3, \ldots, p \)).

**VII. RESULT AND DISCUSSION**

It discusses proposed authentication systems for healthcare applications. This proposed method was developed based on patient data and developed using MATLAB. This work uses two types of dataset to predict healthcare data. The first one is the Mobile HEALTH (M-HEALTH) dataset, and the second one uses a battery-less wearable sensor; researchers were able to recognise activity in healthy older individuals. The suggested methodology’s performance is evaluated using data sets.

With the help of IoT tracking devices, the authors could watch a patient getting long-term treatment for heart disease, their daily activities during treatment, and the effects of the treatment on the patient’s body, and then train based on the results. As a result to that, the trained dataset was around 7358 instances and 258 columns—the main attributes which were highly monitored as patient acceleration movement and angle (Tab.2).

**A. MHEALTH (MOBILE HEALTH) DATABASE**

The MHEALTH database (http://archive.ics.uci.edu/ml/datasets/mhealth+ dataset) contains recordings of physical movement and vital signs, while 10 participants with substantial variations were involved in various fitness activities. In order to monitor the movement of various parts of the body, sensors are attached to the chest, right wrist, and left ankle of a person. ECG readings from a sensor on the chest can also be used for essential heart monitoring, testing for various arrhythmias or examining the effects of exercise on ECG.

**B. RECOGNITION OF ACTIVITY IN HEALTHY OLDER PEOPLE USING A BATTERY-LESS WEARABLE SENSOR DATA SET**

The database comprises the motion data of 14 healthy older people aged 66 to 86 who conducted widely scripted tasks while wearing a battery-less, sternum-level wearable sensor.
TABLE 2. Summary of an experimental outcome.

| Room Setting | Actions | Class Labels |
|--------------|---------|--------------|
| Room_1       | Walking to a chair, sitting in it, getting out of it, walking to bed, laying down in bed, getting out of bed and making my way to the front door. | Sitting on a chair, Sitting on a chair, Lying on a bed, ambulating, which comprises standing and walking around the room. |

Owing to utilizing a passive sensor, the data is scarce and noisy. Participants were randomly assigned to one of two clinical rooms (S1, S2). Room setting S1 (Room_1) collects data using four RFID reader antennas (one on the ceiling and three on the walls), while room setting S2 (Room_2) collects motion data using three RFID reader antennas (2 on the ceiling and 1 on the wall level). The following activities were carried out: Walking to a chair, sitting in it, getting out of it, walking to bed, laying down in bed, Getting out of bed and making my way to the front door. As a result, the following class labels may be applied to each sensor observation: Sitting on a chair, Sitting on a chair, Lying on a bed, ambulating, which comprises standing and walking around the room.

Numerous criteria commonly used in binary classification were utilized in the tests to examine the effectiveness of the various approaches in predicting side effects. To compute multiple criteria, first measure the True Positive (TP), False Positive (FP), True Negative (TN), and False Negative Rates (FNR). The final performance measure was precision, defined as the ratio of relevant retrieved instances. The recall measure, defined as the ratio of relevant instances that have been retrieved, is the second performance measure. The metrics of accuracy and recall are both crucial in assessing the performance of a prediction technique, despite their frequently contradictory nature. As a result, these two metrics may be merged with equal weights to produce the F1-score, a single measure. The final performance measure was the accuracy of correctly predicted events to all predicted events.

The proportion of accurately obtained positive observations to the total number of observations is known as sensitivity/recall.

F1 - score = 2*(Recall * Precision) / (Recall + Precision) (47)

The following is how accuracy is computed for positives and negatives:

Accuracy = (TP + TN) / (TP + TN + FP + FN) (47)

TABLE 3. Comparison results of performance for proposed along with existing approaches.

| Metrics | DFA | CDFT | CDFT-HLCM | ICDFT-EMLM |
|---------|-----|------|-----------|------------|
| Accuracy | 86.059 | 90.950 | 94.400 | 97.9003 |
| Precision | 83.29 | 88.92 | 93.01 | 96.9710 |
| Recall | 88.19 | 90.86 | 94.67 | 97.7730 |
| F-score | 85.67 | 89.88 | 93.83 | 97.3704 |

Tab. 3 tabulates the performance comparison results for the proposed and existing methods for dataset 1. In this, the methods are Data Fusion Approaches (DFA), Context-Aware Data Fusion Technique (CDFT), Context-Aware Data Fusion Technique with Hybrid Learning Classifier Model (CDFT-HLCM) and Context-Aware Data Fusion Technique with Ensemble-based Machine Learning Model (CDFT-EMLM). The resulting graph identified the comparison performance among suggested and current techniques.

Fig.9 illustrates the precision comparison results of the proposed ICDFT-EMLM for the healthcare dataset. Based on the results, it was decided that the ICDFT-EMLM technique was very accurate. Fig. 10. illustrates that the results from the proposed ICDFT-EMLM have high recall results.

Fig.11. shows the F1-score comparison outcomes of the proposed ICDFT-EMLM for the healthcare data. From the results, it was concluded that the proposed ICDFT-EMLM technique was very accurate.
Fig. 9 illustrates the accuracy comparison results of the proposed ICDFT-EMLM model for the healthcare dataset. From the results, it is concluded that the proposed ICDFT-EMLM technique has high accuracy results.

From the above, Fig. 14 and Fig. 15 show accuracy and precision comparison outcomes for both datasets. And it is identified that the proposed method has higher results than the existing methods.

The above graph states the Accuracy comparison results of the proposed ICDFT-EMLM model for the healthcare data. The final metrics were obtained from the Performance Comparison Results for the Proposed and Existing Methods for Dataset-I and Dataset-II (Fig. 18).

Comparing the trained Dataset-I and Dataset-II evidence the results; HMM method improves the fusion process, which will increase the prediction performance. After that, the IPCA is utilized for FE as well as dimension reduction. The FS process uses the ERFE method to eliminate the irrelevant data in a dataset. Finally, this data will be learnt using the EMLM for performance checking. Here the ENN, MXGB

| Metrics | Dataset – I | Dataset - II |
|---------|-------------|--------------|
| Accuracy | 97.9003     | 95.800       |
| Precision| 96.9710     | 93.420       |
| Recall   | 97.7730     | 96.153       |
| F-score  | 97.3704     | 94.766       |

Tab. 4. tabulates the performance comparison results for both datasets. Through the table, it is proved that the introduced approach has more outstanding performance results for Dataset-I than Dataset-II, whereas Dataset-I has fewer.

Fig. 12 Precision analysis of the proposed ICDFT-EMLM healthcare data.

Fig. 13 shows the accuracy and precision comparison results of the proposed ICDFT-EMLM model for Dataset-I and Dataset-II.

Fig. 14 shows the recall measure for Dataset-I and Dataset-II.

Fig. 15 shows the F1-score for Dataset-I & Dataset-II.
and LR models are combined to make a ELM for predicting the healthcare data. Thus the results indicate that the proposed ICDFT-EMLM model improves the prediction performance of the healthcare data (Fig. 19).

As a result of comparing the performance of proposed approaches with the existing methods (Tab. 5).

The DFA method had 86.059% accuracy, a precision of 83.29%, and a recall of 88.19%. Also, with the F1-score, it was obtained as 85.67%. By using the same metrics, CDFT is 90.950% of accuracy, with a precision of 88.92%, and recall achieves 88.19%, resulting in an F1-score of 89.88%. Also, the 3rd parameter CDFT-HLCM shows 94.40% and a precision of 93.01%, with the Recall result as 94.67%, then also with F1-score achieves 93.83%. The proposed method, ICDFT-EMLM, with 97.900% Accuracy and Precision of 96.97%, verifies a 97.77% resultant for Recall, then an F1-score of 97.37%, which was the highest prediction among all these methods.

Tab. 6 has given the comparative results of II-datasets, Dataset-I shows the accuracy metrics of 97.90%, and comparatively, Dataset-II proves the accuracy rate of 95.80%. Dataset-I proves a comparative precision rate of 96.97%, while Dataset-II proves the precision rate of 93.42%. Then while performance comparison with Recall shows a percentage of 97.77% for Dataset-I and Dataset-II, it shows a recall rate of 96.15%. Finally, the performance analysis was made with the F1-score, which proves the final value of around 97.37% and then the dataset-2, which shows 94.76%. So these comparison proves that the proposed method shows the best practice analysis while comparing II-Datasets.

VIII. CONCLUSION AND FUTURE WORKS
This research work focused on developing the Improved Context-aware Data Fusion (ICDF) and efficient FS
algorithm for improving the classification process for predicting healthcare data. In order to accurately achieve data fusion, this study addresses a Dual Filtering Method for preprocessing the data, which attempts to label unlabeled data features. And then, the improved Dynamic Bayesian Network (DBN) is a good trade-off for tractability, becoming a tool for ICDF operations. Here the inference problem is handled using the Hidden Markov Model (HMM) in the Dynamic Bayesian Network (DBN) model. Thus proposed HMM method improves the fusion process, which increases the prediction performance. After that, the Improved Principal Component Analysis (IPCA) is utilized for feature extraction as well as dimension reduction. The feature selection method uses Enhanced Recursive Feature Elimination (ERFE) to eliminate the irrelevant data in a dataset. Finally, this data is learnt using the Ensemble-based Machine Learning Model (EMLM) for performance checking. Here the Enhanced Neural Network (ENN), Modified Extreme Gradient Boost Classifier (MXGB) and Logistic Regression (LR) model are combined to make a predictive model (Ensemble Model) for predicting the healthcare data. Thus the results indicate that the proposed ICDF-EMLM model improves the prediction performance of the healthcare data. The experimental results of two different datasets, Dataset-I, give the accuracy metrics of 97.90%, and comparatively, Dataset-II proves the accuracy rate of 93.85%. Dataset-I proves the comparative precision rate of 96.97%, while Dataset-II proves the precision rate of 93.42%. Then while performance comparison with Recall proves the percentage of 97.77% for Dataset-I, and Dataset-II shows a recall rate of 96.15%. Finally, the performance analysis was made with the F1-score, which proves the final value of around 97.37% and then the Dataset-II, which shows 94.76%. So these comparisons prove that the proposed method is the best practice analysis while comparing two different data sets.

Over reliance on technology, rising rates of medical mistakes and depression, shifting power dynamics, and other human factors all play a role. The Internet of Things has enabled remote monitoring and diagnosis of various health problems, the measurement of a wide range of health parameters, and the provision of testing facilities thanks to the application of these ideas in the healthcare industry. As a result, healthcare has shifted from focusing on hospitals to be more centred on the individual receiving treatment.

REFERENCES

[1] D. S. R. Krishnan, S. C. Gupta, and T. Choudhury, “An IoT based patient health monitoring system,” in Proc. Int. Conf. Adv. Comput. Commun. Eng. (ICACCE), Jun. 2018, pp. 1–7.

[2] F. Fioruiz, A. M. Rahmani, K. Manikodiya, M. Badaroglu, G. V. Merrett, P. Wong, and B. Farahani, “Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics,” Future Gener. Comput. Syst., vol. 78, pp. 583–586, Jan. 2018.

[3] A. Rghioui and A. Ounnad, “Challenges and opportunities of Internet of Things in healthcare,” Int. J. Electr. Comput. Eng., vol. 8, no. 5, p. 2753, Oct. 2018.

[4] J. T. Kelly, K. L. Campbell, E. Gong, and P. Scaffham, “The Internet of Things: Impact and implications for health care delivery,” J. Med. Internet Res., vol. 22, no. 11, Nov. 2020, Art. no. e201355.

[5] P. A. Laplante, M. Kassab, N. L. Laplante, and J. M. Voas, “Building caring healthcare systems in the Internet of Things,” IEEE Syst. J., vol. 12, no. 3, pp. 3030–3037, Sep. 2018.

[6] S. B. Baker, W. Xiang, and I. Atkinson, “Internet of Things for smart healthcare: Technologies, challenges, and opportunities,” IEEE Access, vol. 5, pp. 26552–26564, 2017.

[7] S. M. Kabene, C. Orchard, J. M. Howard, M. A. Soriano, and R. Leduc, “The importance of human resources management in health care: A global context,” Human Resour. Health, vol. 4, no. 1, pp. 1–17, Dec. 2006.

[8] R. V. Milani and C. J. Lavoie, “Health care 2020: Reengineering health care delivery to combat chronic disease,” Am. J. Med., vol. 128, no. 4, pp. 337–343, Apr. 2015.

[9] S. K. Dutta, C. Bonnet, A. Gyraud, R. R. P. da Costa, and K. Boudaoud, “Utilising Internet of Things for personalising healthcare in smart homes,” in Proc. 24th Wireless Pervasive Comput. Conf. (WOPC), Oct. 2015, pp. 164–169.

[10] V. Vippalalapalli and S. Ananthula, “Internet of Things (IoT) based smart health care system,” in Proc. Int. Conf. Signal Process., Commun., Power Embedded Syst. (SCOPES), Oct. 2016, pp. 1229–1233.

[11] F. Hu, D. Xie, and S. Shen, “On the application of the Internet of Things in the field of remote health monitoring,” in Proc. IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things Cyber, Phys. Social Comput., Aug. 2013, pp. 2053–2058.

[12] P. M. Kumar, L. U. Khan, and C. S. Hong, “Affirmative fusion process for improving wearable sensor data availability in artificial intelligence of medical things,” IEEE Sensors J., early access, Feb. 22, 2022, doi: 10.1109/JSEN.2022.3153410.

[13] F. Yang, Q. Wu, X. Hu, J. Ye, Y. Yang, H. Rao, R. Ma, and B. Hu, “Internet-of-Things-enabled data fusion method for sleep healthcare applications,” IEEE Internet Things J., vol. 8, no. 21, pp. 15982–15905, Nov. 2021, doi: 10.1109/JIOT.2021.3067905.

[14] S. Sobhan, S. Islam, M. Valero, H. Shahriar, and S. I. Ahamed, “Data analysis methods for health monitoring sensors: A survey,” in Proc. IEEE 45th Annu. Comput., Softw., Appl. Conf. (COMPSAC), Jul. 2021, pp. 669–676, doi: 10.1109/COMPSAC51774.2021.00097.

[15] H. Wu, T-T. Chang, S-C. Yeh, J. Wang, Y. M. Poole, and C. Shao, “Portable VR-based concussion diagnostics of mild traumatic brain injury,” in Proc. IEEE 3rd Eurasia Conf. IoT, Commun. Eng. (ECICE), Oct. 2021, pp. 21–25, doi: 10.1109/ECICES2819.2021.9645674.

[16] P. Ghamisi, B. Rasti, and R. Gloaeguen, “A novel composite kernel approach for multisensor remote sensing data fusion,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., Jul. 2019, pp. 2507–2510, doi: 10.1109/IGARSS.2019.8900136.

[17] B. Ando, S. Baglio, and V. Marletta, “A multisensor data fusion approach for the volcanic ash granulometry classification,” in Proc. IEEE Sensors, Nov. 2014, pp. 2050–2053, doi: 10.1109/ICS.2014.6985438.

[18] W. Gong, Y. Wang, M. Zhang, E. Mihankhah, H. Chen, and D. Wang, “A fast anomaly diagnosis approach based on modified CNN and multisensor data fusion,” IEEE Trans. Ind. Electron., vol. 69, no. 12, pp. 13636–13646, Dec. 2021, doi: 10.1109/TIE.2021.3155520.

[19] H. Zhang, E. Liu, R. Zhang, and Q. Miao, “RUL prediction and uncertainty management for multisensor system using an integrated data-level fusion and UPF approach,” IEEE Trans. Ind. Informat., vol. 17, no. 7, pp. 4692–4701, Jul. 2021, doi: 10.1109/TII.2020.3017194.

[20] N. E. I. Hamda, A. Hadjali, and M. Lagha, “An advanced weighted evidence combination method for multisensor data fusion in IoT,” in Proc. Int. Conf. Decis. Sci. Technol. (ICDST), Mar. 2022, pp. 810–815, doi: 10.1109/DASAS46582.2022.9765125.

[21] Z. U. Ahmed, M. G. Mortuza, M. J. Uddin, M. H. Kabir, M. Mahiuddin, and M. J. Hoque, “Internet of Things based patient health monitoring system using wearable biomedical device,” in Proc. Int. Conf. Innov. Eng. Technol. (ICICT), Dec. 2018, pp. 1–5.

[22] Z. Gilula, R. E. McCulloch, and P. E. Rossi, “A direct approach to data fusion,” J. Marketing Res., vol. 43, no. 1, pp. 73–83, Feb. 2006.

[23] M. Z. Uddin, M. M. Hassan, A. Alsanad, and C. Savaglio, “A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare,” Inf. Fusion, vol. 55, pp. 105–115, Mar. 2020.

[24] R. Dautov, S. Distefano, and R. Buyya, “Hierarchical data fusion for smart healthcare,” J. Big Data, vol. 6, no. 1, pp. 1–23, Dec. 2019.

[25] S. Begum, S. Barua, and M. Ahmed, “Physiological sensor signals classification for healthcare using sensor data fusion and case-based reasoning,” Sensors, vol. 14, no. 7, pp. 11770–11785, Jul. 2014.
[26] M. Mazammal, R. Talat, A. H. Sodhro, and S. Pirbhulal, “A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks,” *Inf. Fusion*, vol. 53, pp. 155–164, Jan. 2020.

[27] B. Ando, S. Baglio, C. O. Lombardo, and V. Marletta, “A multisensor data-fusion approach for ADL and Fall classification,” *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 1960–1967, Sep. 2016.

[28] H. Medjahed, D. Istrate, J. Bodady, J.-L. Baldinger, and B. Dorizzi, “A pervasive multi-sensor data fusion for smart home healthcare monitoring,” in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jun. 2011, pp. 1466–1473.

[29] Y. Liu, “Multi-source heterogeneous data fusion based on perceptual semantics in narrow-band Internet of Things,” *Pers. Ubiquitous Comput.*, vol. 23, nos. 3–4, pp. 413–420, Jul. 2019.

[30] S. Boukalaboul and D. Djennouri, “DPFOT: Data fusion for Internet of Things,” *J. Netw. Int. Manage.*, vol. 28, pp. 1136–1160, Mar. 2020.

[31] S. Anrouche, B. Göllan, A. Ferscha, and J. Heffberger, “Activity segmentation and identification based on eye gaze features,” in *Proc. 11th Pervasive Technol. Rel. Assistive Environ. Conf.*, Jun. 2018, pp. 75–82.

[32] P. Misra and A. S. Yadav, “Improving the classification accuracy using recursive feature elimination with cross-validation,” *Int. J. Emerg. Technol.*, vol. 11, no. 3, pp. 659–665, 2020.

[33] S. Saikh and V. Chitre, “Healthcare monitoring system using IoT,” in *Proc. Int. Conf. Trends Electron. Informat. (ICEI)*, May 2017, pp. 374–377.

[34] F.-H. Chiang, P.-Y. Kuo, H.-K. Shiu, and Y.-C. Tseng, “A framework for fusing video and wearable sensing data by deep learning,” in *Proc. 8th Int. Conf. Appl. Syst. Innov. (ICASI)*, Apr. 2022, pp. 134–139, doi: 10.1109/ICASI5125.2022.9774477.

[35] G. Guido, S. S. Haghshenas, A. Vitale, and V. Astarita, “Challenges and opportunities of using data fusion methods for travel time estimation,” in *Proc. 8th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, May 2022, pp. 587–592, doi: 10.1109/CoDIT55151.2022.9804014.

[36] Y. Wang, A. Zhang, F. Zhang, Y. Qu, and S. Yu, “Security-aware and privacy-preserving personal health record sharing using consortium blockchain,” *IEEE Internet Things J.*, vol. 9, no. 14, pp. 12014–12028, Jul. 2022, doi: 10.1109/JIOT.2021.3132780.

[37] J. Di Tocco, D. Lo Presti, C. Massaroni, S. Cinti, S. Cimini, L. De Gara, and E. Schena, “Wearable device for plant growth monitoring: A pilot study,” in *Proc. IEEE Int. Workshop Metrol. Ind. 4.0 IoT (MetroInd4.0&IoT)*, Jun. 2022, pp. 289–293, doi: 10.1109/MetroInd4.0&IoT54131.2022.9831678.

[38] K. P. Satamraju and M. Balakrishnan, “A secured healthcare model for sensor data sharing with integrated emotional intelligence,” *IEEE Sensors J.*, vol. 22, no. 16, pp. 16306–16313, Aug. 2022, doi: 10.1109/JSEN.2022.3189268.

[39] P. Singh and L. K. Singh, “Improved measurement accuracy in critical parameters of safety-critical systems with multisensor data fusion,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–8, 2021, doi: 10.1109/TIM.2021.3124852.

[40] A. Nastro, M. Ferrari, V. Ferrari, C. I. Mura, A. Labombarda, M. Viti, and S. D. Feste, “Noise reduction by data fusion in a multisensor system of replicated MEMS inclinometers,” in *Proc. IEEE Int. Symp. Inertial Sensors Syst. (INERTIAL)*, May 2022, pp. 1–4, doi: 10.1109/INERTIAL53425.2022.9787531.

[41] P. Ebel, A. Meraner, M. Schmitt, and X. X. Zhu, “Multisensor data fusion for cloud removal in global and all-season Sentinel-2 imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5866–5878, Jul. 2021, doi: 10.1109/TGRS.2020.3024744.

[42] X. Li and S. Xu, “Multi-sensor complex network data fusion under the condition of uncertainty of coupling occurrence probability,” *IEEE Sensors J.*, vol. 21, no. 22, pp. 24933–24940, Nov. 2021, doi: 10.1109/JSEN.2021.3061457.

[43] M. Webber and R. F. Rojas, “Human activity recognition with accelerometer and gyroscope: A data fusion approach,” *IEEE Sensors J.*, vol. 21, no. 15, pp. 16979–16989, Aug. 2021, doi: 10.1109/JSEN.2021.3079833.

[44] F. D’Apolito, C. Eliasch, C. Sulzbachner, and C. Mecklenbrauker, “A joint multiple hypothesis tracking and particle filter approach for aerial data fusion,” in *Proc. 25th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2022, pp. 1–7, doi: 10.23199/FUSION9751.2022.9841308.

[45] X. Qiang, R. Xue, and Y. Zhu, “Multi-sensor fusion approach with fault detection and isolation based on two-state probability ratio,” in *Proc. 6th Int. Conf. Wireless Commun., Signal Process. Netw. (WSPNNET)*, Mar. 2021, pp. 325–329, doi: 10.1109/WSPNNet51692.2021.9419448.

[46] J. Whittle and J. Schumann, “Automating the implementation of Kalman filter algorithms,” *ACM Trans. Math. Softw.*, vol. 30, no. 4, pp. 434–453, Dec. 2004.

[47] Z. Huang, “Medical inpatient journey modeling and clustering: A Bayesian hidden Markov model-based approach,” in *Proc. AMIA Annu. Symp.*, 2015, p. 649.

[48] L. Guo, P. Wu, S. Lou, J. Gao, and Y. Liu, “A multi-feature extraction technique based on principal component analysis for nonlinear dynamic process monitoring,” *J. Process Control*, vol. 85, pp. 159–172, Jan. 2020.

[49] S. C. Wang, “Artificial neural network,” in *Interdisciplinary Computing in Java Programming*, Boston, MA, USA: Springer, 2003, pp. 81–100.

[50] Y. P. Kondratenko, L. P. Klymenko, and E. Y. M. Al Zu’bi, “Structural optimization of fuzzy systems' rules base and aggregation models,” *Kybernetes*, vol. 42, no. 5, pp. 831–843, May 2013.

[51] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, and H. Cho, “Xgboost: Extreme gradient boosting,” *R Package Version* vol. 1, no. 4, pp. 1–4, Aug. 2015.

[52] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, “An introduction to logistic regression analysis and reporting,” *J. Educ. Res.*, vol. 96, no. 1, pp. 3–14, 2002.