Journal of Information Systems Engineering and Management
2025, 10(22s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhanced IoT-Based Health Data Fusion with Recursive Feature Elimination for Improved Patient Monitoring

Preetha P[1], Dr. A.Packialatha[2]
[1]Research Scholar, Computer Science & Engineering, Vel's Institute of Science Technology and Advanced Studies, Tamilnadu, India
(Corresponding Author) Email ID: preetha.pandu@gmail.com
[2]Associate Professor, Computer Science & Engineering, Vel's Institute of Science Technology and Advanced Studies, Tamilnadu, India
Email ID: packialatha.se@velsuniv.ac.in

ARTICLE INFO

Received: 20 Dec 2024

Revised: 30 Jan 2025

Accepted: 15 Feb 2025

ABSTRACT

Healthcare market utilization of IoT relies on digital systems that monitor and analyze health issues within the framework of IoT technology. Through IoT and smart devices a high level of smart environment can be built. Medical devices linked to smartphone apps enable the collection of medical health data and other necessary patient information. Data Fusion (DF) functions as a process of uniting information that exists across multiple data sources. These methods get implemented beyond text processing applications within other domains. DF operates on numerous distributed data sources to minimize detection errors as well as increase their reliability in multisensory settings. The main goal targets scalability together with performance enhancement and identification capabilities. The environmental responsiveness of medical devices depends on their strength to adjust output according to varying circumstances. The system demonstrates its scalability through expected function with continuous performance while utilizing available resource management at optimal levels. Forming a specialized group to establish common practices among tracking devices remains vital because this ensures uniformity in communications and data processing and user interface standards. The primary research outputs include implementing pre-processing followed by Improvised Context-aware Data Fusion (ICDF), Improved Principal Component Analysis (IPCA) for feature extraction and Enhanced Recursive Feature Elimination (ERFE) for selection as well as ensemble-based Machine Learning (ML) model as classifier. The Improved Dynamic Bayesian Network (IDBN) presents itself as a balanced choice for tractability since it can serve as a tool for ICDF operations. Results from simulations indicate that the proposed ICDF model performs best in the healthcare system with 96% accuracy along with 95% precision and 96% recall and F1 score at 96%.

Keywords: Enhanced recursive feature elimination, ensemble-based machine learning mode, health care data, improved dynamic Bayesian network, IoT.

1. INTRODUCTION:

The healthcare field represents one of the essential areas where ubiquitous applications can be observed. Each space becomes a part of pervasive computing because it operates in the background independently of user input [1][2]. Several connected machines form a hidden Artificial Intelligent (AI) system through wire and wireless technology which functions autonomously or under human supervision. The devices join together through interconnected systems to function separate from other devices. Every one of these devices comes with built-in chips which let them join with other devices through a network because of their continually accessible connection [3, 4, 5]. The capability to leverage pervasive computing in healthcare and homecare as well as transportation systems and other domains exists because of its features.

The main purpose of universal healthcare involves utilizing pervasive computing technology to deliver healthcare services everywhere and at all times. The standard healthcare system combines with universal healthcare when patients must locate symptoms and share them with doctors before obtaining medical treatments. [6] Pervasive healthcare delivers healthcare treatment to people at all locations and during any time period. Patients can be monitored continuously while their health data is collected through sensor integration to communication systems under this concept. The system delivers dependable health information to doctors or medical experts thus enabling rapid diagnosis and treatment of patient healthcare issues from any location [7, 8]. Intelligent portable devices including mobile phones and smartwatches were developed through the combination of cloud communication and sensing innovations to enable various ubiquitous medical system developments.

Doctors can access these organized medical records to offer better medical care to their patients. The existing study describes how CDF-EMLM helps manage health data better. The database accuracy decreases since it includes unnecessary information. This research created ICDF and EFSA to enhance how medical data gets classified for prediction purposes. This article's central contribution includes the ICDF data filtering method plus IPCA feature extraction combined with ERFE feature selection followed by EML classification models. The suggested approach generates precise outcomes from hands-on programming for the specified dataset according to reference [9, 10].

This part of the work includes an evaluation of health care data fusion techniques which have emerged recently. Section 3 elucidates the introduced technique. The fourth part shows evaluation results and their analysis. The concluding section with prospective upgrades is presented in Section 5.

2. RELATED WORKS:

The paper examines recent IoT-based health care methods which could enhance the proposed system model. This part explains what the prediction model in this field needs for improvement. The internet contains millions of affiliated items. The advanced levels possess complex frameworks that allow them to both store essential information and take effective decision while processing data efficiently such as smart devices. These devices operate with limited models accompanied by small storage and restricted processing capability (such as body sensors). Because various devices maintain relationships with one another the IoT creates complex situations. The implementation of data knowledge extraction for IoT setups becomes possible through data modelling and sensible analysis techniques on device-collected data volumes [11] and [12]. When applied to IoT setups this capability is commonly called IoT intelligence. Fig. 1 shows the time-related transformation.

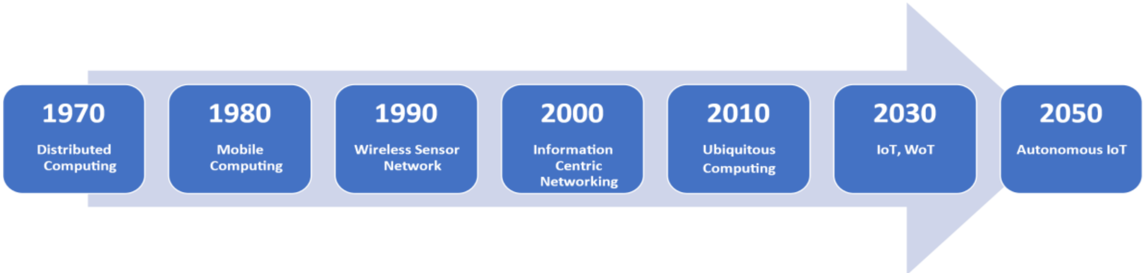

Figure 1: The research evaluated the geographic areas and significant historical aspects.

FE operates alongside sensor data collection and modelling as well as data processing within the context life cycle. Working on IoT-related solutions with communications infrastructure while developing paradigms represents a beneficial approach. The life cycles share four fundamental characteristics which include context acquisition and modelling alongside reasoning as well as distribution. Various physical and virtual sensors send aggregated data to the system during the context information collection phase. During context modelling phase the data needs proper modeling which corresponds to its meaningful data attributes. FE depends on raw data processing during the context reasoning phase. Context distribution as the final subcategory distributes collected information through multiple distribution methods which include servers, scripting languages together with frameworks as described in references [13, 14, 15]. The lifecycle system of IoT context is illustrated in Figure 2.

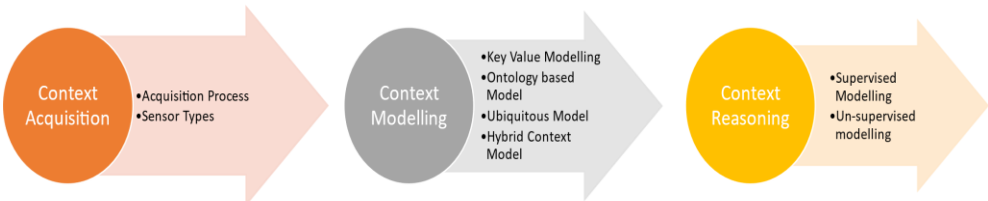

Figure 2: IoT operates with a context life cycle framework.

A designed and stable society emerges from an improved and smart healthcare system according to Reference [16]. Remote healthcare scanning of patient health status is possible for doctors through integrated IoT functionality within the digital healthcare system. This research demonstrates how automated and smart IoT systems track patient health data which gets stored on internet platforms before sending prompt alerts regarding critical conditions to medical practitioners. The system accessibility and user-friendliness enhances through research which focuses on system affordability. This system provides doctors with constant knowledge of their patients' existing health situations. The system integrates emergency notification to contact both the patient doctor and their family members when detecting any injury. Remote monitoring systems enable doctors to save a large number of lives while taking full responsibility for healthcare delivery [17].

The proposed method from [18] allowed researchers to calculate the similarity measurement dedicated to study variables. This technique provides a solution that users can implement using basic statistical procedures and standard software packages no matter if they work with discrete or continuous data. The majority of marketing applications use either psychographic or demographic classifying factors. Two main features to unite in the analysis are media viewing alongside product purchasing. The method demonstrates its capability to correctly predict the combined distribution of media use and online ordering under these circumstances. The data collection tool functions as the deciding factor for marketing choices. The combination of separate variables plays a critical role throughout marketing operations. The research group developed a system to eliminate conditional independence requirements in this case. Researchers verify their method using survey information collected through extensive examinations of British consumer buying behavior and media usage.

The body sensor-based detection of behavior should use Deep Recurrent Neural Network (DRNN) as a powerful Deep Learning framework for sequential data according to Reference [19]. These devices merge collected body sensor information that consists of electrocardiography (ECG) readings together with accelerometer and magnetometer data. The FE gets enhanced through Kernel Principal Component Analysis (PCA) application. Robust features are utilized for training the RNN process which afterward functions for activity detection. The method underwent assessment against conventional methods through testing on three publicly available standard databases. Experimental results prove that the proposed technique achieves superior results over conventional approaches.

The strategy developed by [20] utilized complex event processing to execute hierarchical data breakthroughs on streaming data in real time which suits the requirements of time-sensitive IoT devices dealing with data limitations. The proposed technique supports refined decision-making across DF stages leading to better system performance alongside faster responses for public healthcare services thus boosting the IoT adoption in healthcare.

2.1 Limitation of Related Works

- ML operates to analyze ECG data instead of requiring manual monitoring systems. The automated system runs independently thus reducing operating costs. Security enhancement measures must be developed due to the high vulnerability of IoT systems to cyberattacks.
- The combination of smart devices and a microphone enables further increases in scalability levels. The ordinary IoT framework deals with minimum data amounts.
- ML-based Signal monitoring successfully decreases mistakes during the process while simultaneously enhancing measurement precision. Additional time for execution will be needed as data involves increased FE amounts.
- An improvement of scalability together with adaptability must be pursued.
- The system requires an unsupervised model development to successfully process real-time data.
- User devices need periodic battery replacement as a required operating standard. A reduction of the total data amount becomes essential.

3. PROPOSED METHOD

The IoT technology allows healthcare systems to collect prolonged patient data records. The study developed an ICDF and an efficient FS algorithm to enhance the classification procedure in healthcare data predictions. The discussion includes ML algorithms that use dimensionality reduction by feature engineering along with dimensionality expansion through clustering and dimensionality reduction by association rule learning and DL technology. System components that demonstrate context awareness possess the ability to grasp environmental data surrounding them for subsequent adjustment of their operational responses. The major operating principle of "contextual" (or) "context- aware" computing involves automatic data collection which enables subsequent analysis for automated action guidance. To present an approach for context implementation within ML models according to [21]. The two components of conditional probability distribution consist of context-free elements and context-sensitive counterparts. Context-aware systems maintain autonomous status because they automatically modify their outputs to correspond to user needs independently. A system that is context-aware must provide computers with access to context-related data before any analysis takes place. The complexity addition through EML methods makes interpretation difficult leading to limited business insights which becomes challenging to obtain. The design and computation of EMLM system requires an extended time duration [22].

The authors dedicate this work to the examination of ML algorithms that serve as standard tools within IoT-related areas particularly sensor networks and context-aware systems. This research explores the Supervised Learning (SL) and Unsu- pervised Learning (UL) along with Reinforcement Learning (RL) techniques in the following parts. This paper discusses exclusively IoT-specific and context-aware ML-specific algorithms and techniques because ML encompasses a wide scope of applications [23].

The IoT-DF middleware consists of four submodules that include data receiving and computation and knowledge inference and user- related acquisition and adding new features and service decomposition and performance [24]. A visual depiction of the system structure appears in Figure 3. In real-time the processing module of DF middleware can collect data from IoT devices and other sensors present in the smart library. Transformation of bottom-layer raw data into meaningful events needs the knowledge reasoning module to apply its rules first. Any events transmitted by the knowledge reasoning module get transformed into simple format processing by the event description module according to References [25]. The event decomposition module transforms these requirements into executable commands for the system through its translation process [23]. Figure 3 displays the system functional decomposition diagram that illustrates the methodology flow process in sequential order. Figure 4 explains the System's Function Blocks.

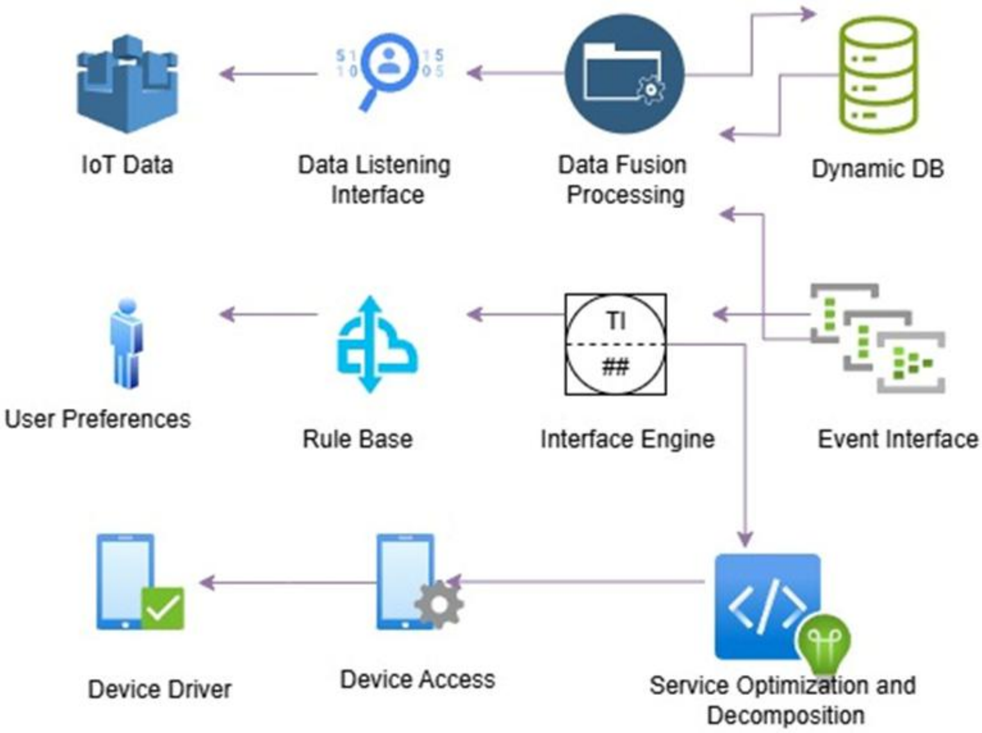

Figure 3: Schematic Diagram of the Complete System

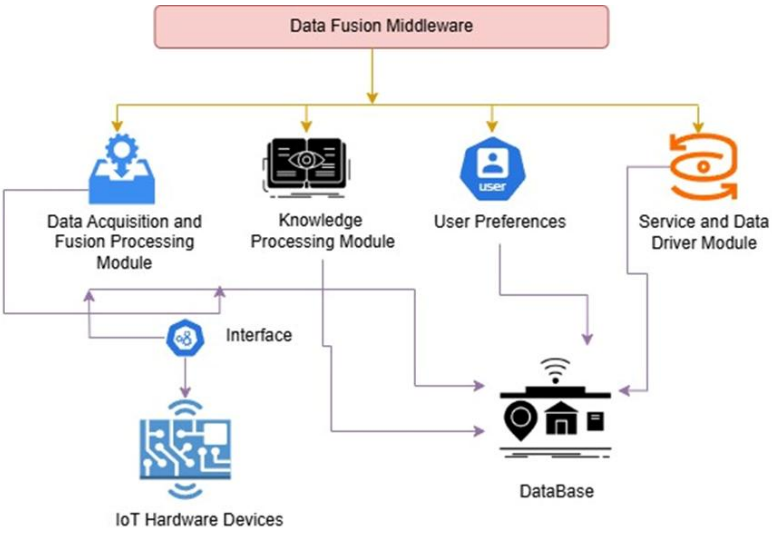

Figure 4: System Function Blocks

- Data Fusion: This system will gather IoT data before transforming it for successful fusion processing. A Dual Filtering Method was developed to assist data preprocessing by attributing unknown data values needed for accurate Data Fusion.
- Improved Dynamic Bayesian Network: IDBN offers an optimal solution between tracking performance and making ICDF operations possible. The HMM component in DBN architecture solves the inference problem.
- Improved Principal Component Analysis: IPCA serves to find salient features while compressing the amount of dimensions.
- Feature Selection Process: The FS process implements ERFE as its method to remove unneeded data from datasets.
- Ensemble-Based Machine Learning Model: The EMLM model functioned as a performance evaluation method to learn this data. Healthcare data prediction includes a combination of ENN with MXGB together with LR to form an ensemble mode.

The following section explains the proposed method with detailed information shown in Figure 5. Shows the overall proposed system architecture.

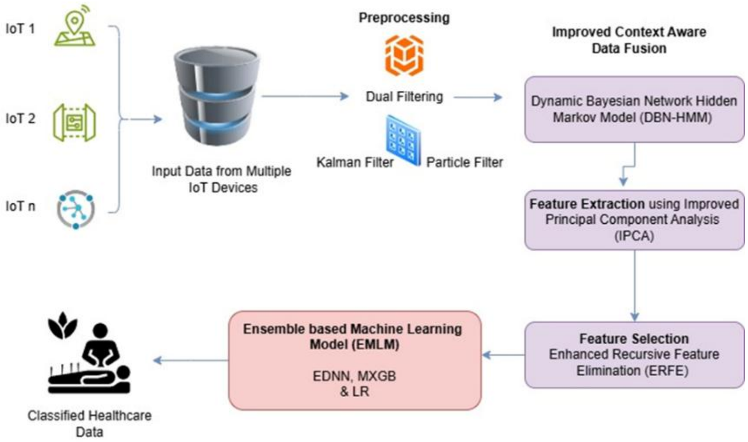

Figure 5: Overall Proposed System Architecture

3.1 Preprocessing

The first step towards context understanding is obtaining data from physical devices. The system first reduces measurement noise by filtering and estimating device data. After filtering out mistakes from data the correct analysis requires single learners or combined classification models. After cleaning the data remains as cleaned data elements.

This research project starts by filtering out noise from its raw data before the final filtering takes place. In this research project we use the DFM to handle the data. The Kalman Filter stays with statistical measurement yet Particle Filter brings new methods to measure system states.

With a standard deviation (σ) of 1 and a mean (μ) of 0, it rescales the risk factor value to enhance the performance of ML classifiers. Standardization's mathematical form is provided by (1).

$$Standardization\ of\ Y = \frac{Y - Mean\ of\ Y}{Standard\ deviation\ of\ Y} \tag{1}$$

3.2 Kalman Filter

KF works as a leading State Estimation Statistical Technique to combine changing signal measurements. A system uses the prediction and update method to find state estimates which depend on previous time period's data. The current system state depends on its previous time interval state. One of its many benefits is its computational efficiency. KF regularly combines accelerometer and gyroscope signals in its work. KF produces a reliable estimate results for this application [46]. For example it helps detect body balance changes during quiet standing.

$$x = ({\xi_1}^S(y - y), \xi 2T(y - y), \dots, \xi dT(y - y)) \in \mathbb{R}^d \tag{2}$$

The average of a training set and a covariance matrix is determined given a training set $Y = y_1, y_2, \dots, yN(y_i \in \mathbb{R}^D, i = 1,2, \dots, N)$ and a lower dimension. The eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_D$ and related eigenvectors $(\xi_1, \xi_2, \dots, \xi_D)$ are obtained by performing a spectral decomposition on the covariance matrix. Equation (2) expresses any $y \in \mathbb{R}^D$ new low-dimensional representation.

3.3 Particle Filtering

When calculating target probability density moments becomes impossible through standard analysis PF provides a random method for finding them instead. The method starts with picking random numbers representing particles from an available sample distribution. Next, for every method uses an assigned weight to adjust how far the desired outcome differs from the significance probability results. They determine the total target distribution with no assumption needed and benefit non-Gaussian nonlinear systems best. The system uses PF to check biomechanical status based on data provided by the accelerometer and gyroscope.

$$X = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{3}$$

$$X' = \left[0.8 \times \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + 0.1 + v_i\right] * 0.5 \tag{4}$$

In this, $v_i = \frac{x_i - \min}{\max - \min}\left(New^{\max} - New^{\min}\right) + New^{\min}$ (5)

3.4 Improved Context-Aware Data Fusion (ICDF)

DBN provides working flexibility in this research as we turn it into a Decision Foundation support element. This research uses DBNs to find how context variables influence environments through analyzing any type of data distribution. The model separates health data into time segments to find observable symptoms as the result of an HMM method. The hidden variable 'V_t', 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. $S_t=(S^1{}_t, \dots S^n{}_t)$is the set of sensor's readings active in the 't^o time slice, and the contexts set is represented by $Cn_t=(Cn^1{}_t, \dots Cn^n{}_t)$ 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.

$$MI(x, y) = H(x) + H(y) - H(x, y) \tag{6}$$

$$IMI(x, y) = \frac{1}{2}\left[\frac{2MI}{H(x) + H(y)}\right] + VOI \tag{7}$$

$$\sigma = \left(\frac{\Gamma(1+\lambda)*\sin\left(\frac{\pi\lambda}{2}\right)}{\Gamma\left(\frac{1+\lambda}{2}\right)*\lambda*2^{\left(\frac{\lambda-1}{2}\right)}}\right) \quad (8)$$

3.5 Hidden Markov Model (HMM)

A model consists of defined state (N) sections that work alongside their particular probability distribution system. The states within the system change following a defined mathematical framework. HMMs include multiple transition probabilities that define state movement. A series of procedures needs to run to create a word recognition system based on HMMs. (a) Select some states and observations, (b) select an HMM topology, (c) select training and samples, (d) Train the system using the training data, and (e) test data using testing data. Fig. 7 depicts the example of a 7-state; HMM that only allows transitions to the same state → next state → subsequent state [47]. Every state has transitions to the same state → next state → next state. During training and testing, the order in which the model's states change is determined by the FE described in the previous section.

$$w_{new}^1(s+1) = W_j(s) + rand * k\ \beta_s * \left(W_j(s) - \beta_s * w_{best}\right) \quad (9)$$

$$w_{new}^1(s+1) = \begin{cases} W_j(s) + Levy(\lambda)(\left(W_{best}(s) - W_j(s) * w * z * y + W_{best}(s)\right)\ , q_1 > 0 \\ W_j(s) + rand * k * \beta_s * (W_j(s) - \beta_s *\ W_{best}(s) q_1 < 0 \end{cases} \quad (10)$$

3.6 Feature Extraction

FE picks crucial system and environmental parameters to perform self-optimization. The features needed for optimization can be determined from DF output and by watching system functions. FS directs the self-configuring behavior of continuous components through updates triggered by changes in their context. When things change fast a system cannot work well by following basic design plans or total target settings. The part features two sections. The primary function measures time-based and frequency-based signals. Measurement functions need the initial signal that shows how to measure the system features. The technique then collects parameters from all measurements while integrating category problem insights.

$$w_{new}^2(s+1) = \begin{cases} \alpha * \left(W_{best}(s) - rand * w_{new}^1(s)\right) + 1(1-\alpha) * W_{best}\quad , if\, q_2 \geq 0.1 \\ (1-\alpha) * (w_{new}^1(s) - rand * W_{best}) + \alpha * w_{new}^1(s),\ if\, q_2 \leq 0.1 \end{cases} \quad (11)$$

3.7 Improved Principal Component Analysis (IPCA)

PCA main purpose is to study data and design prediction systems. Researchers use PCAs to present the genetic connections between populations through distance measurements. First PCA needs data normalization steps using eigenvalue decomposition and decomposition of the data covariance matrix or singular value decomposition of a data matrix. The process normalizes each attribute by adjusting all data values to reach a zero statistical average for every measured variable before making 1 the variance for each variable. The performance of PCA typically includes both Component scores for representative data points and Loadings which explain how much each observed standard variable should be multiplied for PC scores. The data variance needs to be included in the loading results when factor scores normalize to a common variance. Loadings need to become normalized when component scores do not include their own data variance. Eigenvectors stand as the cosine measurements for rotating variables to match their principal components (or) initial position.

PCA features display inner data structure changes that demonstrate all data variation accurately. With a multivariate database represented as coordinates in a high-dimensional space PCA offers viewers a most relevant low-dimensional projection of their data set. The modified dataset's dimension is lowered through the use of only its key principal components. When datasets have abnormal values PCA performs poorly as the primary issue in our research. When you have too much data it becomes challenging to detect abnormal data points. Constructing the AGKM helps solve this problem.

$$\begin{cases} Fathers = W_{sort}^2\left(1{:}\frac{o}{2}\right) \\ Mothers = W_{sort}^2\left(\frac{o}{2} + 1{:}\,o\right) \end{cases} \quad (12)$$

3.8 Feature Selection using Enhanced Recursive Feature Elimination (ERFE)

FS and dimensionality reduction handle predictive systems by identifying fewer needed input elements. FS uses special filters to pick out needed database attributes while dimensionality reduction methods develop and present new input values from existing data. Hence dimensionality reduction functions as filtering features instead of replacing standard filtering methods. Having fewer system features helps experts understand the training process and speeds up processing. A reduced set of features can help control costs because scientists require fewer attribute tests to make effective disease diagnoses in biological settings. RFE utilizes ML generalization skills to handle minor sample-size problems successfully.

$$a_v^u = \varphi\left(\sum_o a_o^{u-1} * j_v^u + c_v^u\right) \tag{13}$$

3.9 Feature Selection Wrapper Technique (FSWT):

The process removes small weak features to build superior models with quality independent features. It analyzes features in a stepwise pattern following the basis of backward feature elimination. The technique initializes from entire features and constructs a model before evaluating the feature ratings. The model updates itself with only the important features when the least significant feature gets removed. Within this procedure assume T refers to sequential numbering to maintain the quality score. The T_i list brings forward select features to help build a new model and measure model success at each stage during feature elimination. The system identifies the best performance result for the value 'T_i' before linking it to the final model with its top-performing features.

$$T_A = \sum_{j=1}^{m} \mathrm{M}_j \left(\mu_j - \mu\right)\left(\mu_j - \mu\right)^S \tag{14}$$

3.10 Enhanced Deep Neural Network (EDNN)

DL produces accurate results when processing data sets with many details. In this research work, You will create a better Neural Network design. The core idea underpinning this project involves linking a fuzzy inference system with a DNN. This design uses a Fuzzy Neural Network architecture linked to multiple hidden layers. A FNN acts as a learning method in neural network technology to produce precise output results. this work combines NN technology to build FL functionalities [49]. In this approach, a DNN with The network includes several hidden layers whose number remains fixed during training. The model improves performance through additional hidden layers. The network will achieve better results with more hidden layers but also creates added system complexity. complexity and decreases training performance. Belonging to the training part is the task of selecting how many hidden layers exist. The analyst defines the system requirements. The design of a DL framework decides if it works with raw data types including images or one-dimensional signals. The system needs no labels during its training process.

The system can automatically update the weights without needing labeled information. Following training these networks each neuron can be evaluated for its response to positive and negative stimuli. The network will test how neurons react to each type of stimulus input. An individual input carries many labeled signals. features. The validation process helps the network adjust its neurons by transferring feedback from the end point. The main use of a DNN requires setting data parameters successfully because manual parameter selection leads to overfit problems. These parameters experience frequent changes as users select them from experience yet require manual adjustment. The system training routine incorporates this problem. A computer system can now handle these procedures automatically. minimising errors. This research produced an Fuzzy Inference Method (FIM) to tackle the detected issue. The The FIM system uses responding neurons to determine useful feature inputs.

3.11 Logistic Regression (LR)

This research selects the LR model for analysis which performs well in many everyday data mining tasks [52]. The study researched major health risks from provided data sets and estimated how likely they would happen. The model calculated risk level based on specific danger elements. Logistic regression performs best when used for classifying data sets that include only two clear potential outcomes per item. The next section shows the formula of an LR model defined as Model 43.

4. RESULT AND DISCUSSION:

It suggests ways to verify users accessing healthcare programs. Researchers developed the proposed system, which processes medical information, while the programmers built it using the MATLAB platform. Our work processes two separate datasets to forecast healthcare information. The researchers worked with two specific datasets for this project: first, they used Mobile HEALTH (M-HEALTH) data and also experimented with battery-less wearable sensor data. Health researchers identified normal senior activity patterns by monitoring their movement activities. The suggested method requires measurement using different test datasets. The medical team monitored patient heart disease treatments through Internet of Things (IoT) tracking systems. They observed the patient's daily treatment process and tracked the impact of the treatment on their body before training their models. Table 1 and Figure 6 show the comparison results of performance for the proposed approach alongside existing approaches (accuracy, precision, recall, F1 score). Table 2 and Figure 7 show the performance of the proposed vs. existing methods on the trained datasets (accuracy, precision, recall, F1 score).

Table 1: Comparison Results of Performance for Proposed along with Existing Approaches.

Metrics	DFA	CDFT	CDFT-HLCM	ICDFT-EMLM
Accuracy	85.159	89.85	93.5	96.8003
Precision	82.39	87.82	92.11	95.871
Recall	87.19	89.76	93.77	96.673
F1-score	84.67	88.78	92.93	96.2704

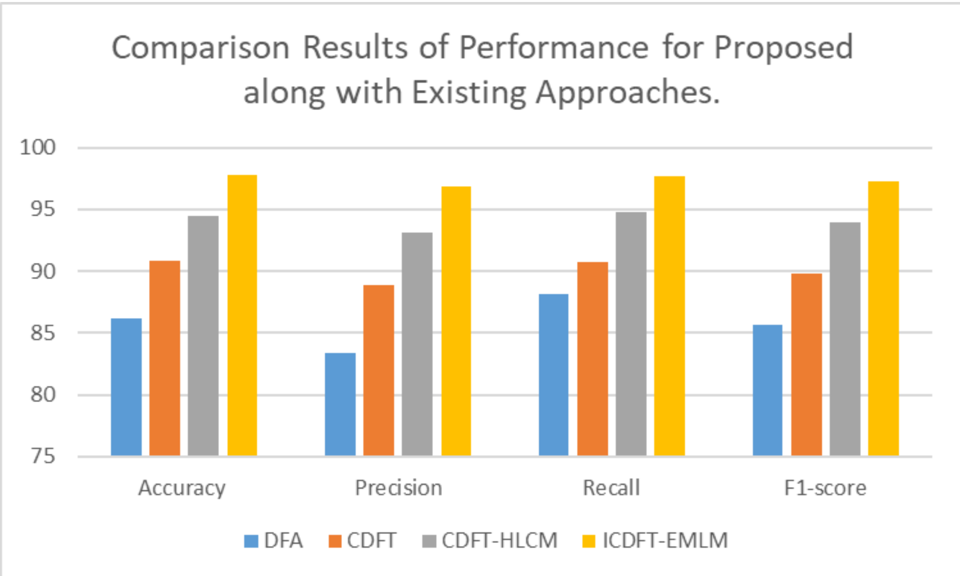

Figure 6: Comparison Results of Performance for Proposed along with Existing Approaches.

Table 2: Performance of Proposed vs Existing Methods.

Metrics	Dataset-I	Dataset-II
Accuracy	96.8003	94.8
Precision	95.871	92.42
Recall	96.673	95.153
F-score	96.2704	93.766

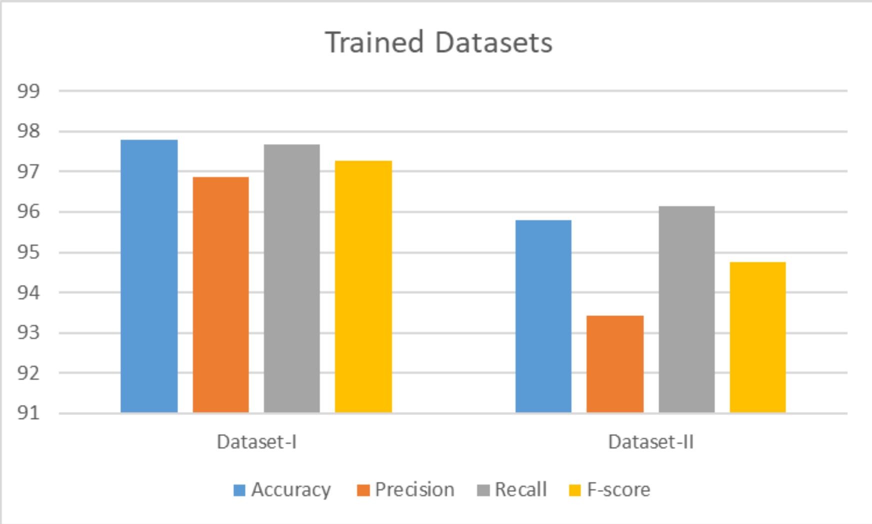

Figure 7: Results of Trained Dataset-I and Dataset-II.

5. CONCLUSION

The research project developed both an Improved Context-aware Data Fusion (ICDF) system and an efficient feature selection (FS) algorithm to enhance the healthcare data classification prediction process. A Dual Filtering Method is introduced to process data through labeling functions that address the accuracy of data fusion techniques. The improved Dynamic Bayesian Network (IDBN) strikes an optimal balance between tractability, making it an effective operational tool for ICDF. The Dynamic Bayesian Network model solves its inference problem by implementing a Hidden Markov Model (HMM). The fusion process becomes more effective with the implementation of the proposed HMM methodology, leading to improved prediction outcomes. Dimension reduction and feature extraction occur through the application of Improved Principal Component Analysis (IPCA). Enhanced Recursive Feature Elimination (ERFE) provides the feature selection method to eliminate any irrelevant data from the datasets. Performance checks were made on the data through the Ensemble-based Machine Learning Model (EMLM). The predictive model to forecast healthcare data consists of an Ensemble Model that combines the Enhanced Neural Network (ENN), Modified Extreme Gradient Boost Classifier (MXGB), and Logistic Regression (LR). The proposed ICDFT-EMLM model offers superior prediction capabilities for healthcare data, as shown by the obtained results. The accuracy levels measured for Dataset-I are 96.80%, while Dataset-II demonstrates 94.80% accuracy. Dataset-I achieves 95.87% precision, while Dataset-II achieves 92.42% precision. The study reveals that Dataset-I achieved 96.67% recall, but Dataset-II produced a recall result of 95.15%. The F1-score performance analysis resulted in a final value of 96.27% for Dataset-I, while Dataset-II showed 93.76%. The proposed method stands as the primary method for conducting differential analysis of dissimilar datasets according to these assessment standards.

REFERENCES:

[1] S. Boulkaboul and D. Djenouri, "DFIOT: Data fusion for Internet of Things," J. Netw. Syst. Manage., vol. 28, pp. 1136–1160, Mar. 2020.

[2] S. Amrouche, B. Gollan, A. Ferscha, and J. Heftberger, "Activity segmentation and identification based on eye gaze features," in Proc. 11th Pervasive Technol. Rel. Assistive Environ. Conf., Jun. 2018, pp. 75–82.

[3] 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.

[4] S. Shaikh and V. Chitre, "Healthcare monitoring system using IoT," in Proc. Int. Conf. Trends Electron. Informat. (ICEI), May 2017, pp. 374–377.

[5] T.-H. Chiang, P.-Y. Kuo, H.-R. 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/ICASI55125.2022.9774477.

[6] 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.

[7] Y. Wang, A. Zhang, P. 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.

[8] 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/MetroInd4.0IoT54413.2022.9831678.

[9] 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.

[10] 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.

[11] 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.

[12] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: Technologies, challenges, and opportunities," IEEE Access, vol. 5, pp. 26521–26544, 2017.

[13] 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.

[14] R. V. Milani and C. J. Lavie, "Health care 2020: Reengineering health care delivery to combat chronic disease," Amer. J. Med., vol. 128, no. 4,

[15] pp. 337–343, Apr. 2015. S. K. Datta, C. Bonnet, A. Gyrard, R. P. F. da Costa, and K. Boudaoud, "Applying Internet of Things for personalized healthcare in smart homes," in Proc. 24th Wireless Opt. Commun. Conf. (WOCC), Oct. 2015, pp. 164–169.

[16] V. Vippalapalli 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.

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

[18] 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.

[19] 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. 15892–15905, Nov. 2021, doi: 10.1109/JIOT.2021.3067905.

[20] 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.

[21] H. Wu, T.-T. Chang, S.-C. Yeh, J. Wang, J. 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/ECICE52819.2021.9645674.

[22] P. Ghamisi, B. Rasti, and R. Gloaguen, "A novel composite ker- nel 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.

[23] 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/ICSENS.2014.6985438.

[24] 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. 2022, doi: 10.1109/TIE.2021.3135520.

[25] H. Zhang, E. Liu, B. 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.