Journal of Intelligent Systems and Internet of Things
Vol. 17, No. 02, PP. 214-227, 2025
DOI: https://doi.org/10.54216/JISIoT.170213

IoT and AI for Clinical Decision Support with Hierarchical Attention

Sasikumar M. S. S.[1,*] , Ranganayaki V. C.[2] , R. Suganthi[3], Nalini Subramanian[4] , T. Sethukarasi[5] , T. A. Mohanaprakash[6,*]

[1]Assistant Professor (Senior Grade), Department of Computer Science & Engineering School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, India

[2]Assitant professor , Department of Computer Science and Engineering, St.Joseph's Institute of Technology, OMR, Chennai, Tamilnadu, India

[3]Associate Professor, Department of ECE, Panimalar Engineering College, Chennai, Tamilnadu, India

[4]Associate Professor, Department of Information Technology, Rajalakshmi Engineering College, Thandalam, Chennai, Tamil Nadu, India

[5]Professor & Head Department of Computer Science and Engineering, R.M.K Engineering College , Kavaraipettai, Chennai, India

[6]Associate Professor, Department of Computer Science and Engineering, RMK Engineering College , Kavaraipettai , Chennai, India

Emails: drsasikumarmss@veltech.edu.in; ranganayaki.v.c@gmail.com; drrsuganthi@panimalar.ac.in; mrgn.nalini@gmail.com; tsk.cse@rmkec.ac.in; tamohanaprakash@gmail.com

Abstract

The integration of Clinical Informatics (NI) and Artificial Intelligence (AI) promises to transform healthcare by improving clinical decisions, optimizing workflows, and personalizing patient care. However, most current systems fail to incorporate contextual reasoning, real-time adaptation, or ethical sensitivity, leading to fragmented support and increased cognitive burden on clinicians. To address these limitations, we propose NI-AIH—a hybrid clinical-AI framework built on a Context-Enriched Hierarchical Attention Network (CE-HAN). This deep architecture employs dual-attention mechanisms to interpret structured and unstructured clinical data—including EHR entries, nursing notes, and real-time IoT sensor feeds—capturing temporal patterns and contextual cues essential to patient status. The NI-AIH framework consists of four core components: a Clinical Context Engine (CCE) that uses CE-HAN for semantic modeling; a Predictive Care Optimizer (PCO) that applies risk-stratified deep ensembles; an Adaptive Interaction Layer (AIL) that enables seamless nurse–AI collaboration; and an Ethical Decision Integrator (EDI) that uses fuzzy logic to ensure real-time ethical alignment. In a trial deployment within a smart geriatric care unit, NI-AIH demonstrated a 23% improvement in early sepsis detection ($p<0.01$), a 31% reduction in clinician cognitive load (measured via NASA-TLX survey), and a 19% increase in workflow efficiency compared to conventional rule-based systems. By uniting clinical precision with ethical and context-aware intelligence, NI-AIH establishes a new paradigm for compassionate and effective AI-assisted healthcare.

Received: January 08, 2025 Revised: March 06, 2025 Accepted: May 28, 2025

Keywords: Clinical Decision Support Systems; Hierarchical Attention Network; Predictive Analytics; IoT in Healthcare; Ethical AI

1. Introduction

With the growing need for data-driven, efficient, and context-aware clinical decision support, the integration of AI with Clinical Informatics (NI) is a game-changer in today's healthcare systems [1]. Physiological monitoring devices, clinical procedures, unstructured narrative documentation, and electronic health records (EHRs) are essential tools for nurses, as they enable real-time patient care in complex data contexts [2]. However, clinical workflow inefficiencies, higher cognitive burdens, and inferior patient outcomes are common consequences of NI systems that do not currently enable dynamic decision-making [3]. The use of AI in Clinical informatics has been relatively unexplored despite AI having made significant strides in other areas of medicine, including radiography, pathology, and administrative automation [4]. The fact that Clinical data is often multi-modal, temporally variable, context-dependent, and interwoven with qualitative decision-making presents a significant obstacle [5]. Inadequate contextual modeling and restricted interpretability lead to current machine learning (ML) applications underperforming, and traditional rule-based decision support systems (DSS) struggle to adjust to real-time clinical variance [6].

Traditional Clinical decision support systems rely heavily on pre-defined protocols and do not fully incorporate unstructured textual inputs, such as patient-nurse conversations or Clinical notes [7]. Complex contextual cues and temporal connections, which are essential in acute care situations, cannot be modeled by these systems [8]. Furthermore, existing AI designs have largely neglected ethical reasoning despite its centrality to Clinical practice [9]. Systems that provide advice without considering the ethical context, patient autonomy, or emotional intelligence create a significant gap, putting healthcare staff's confidence at risk [10]. The full potential of precision Clinical is hindered, decision accuracy is compromised, and the documentation burden is increased due to the absence of intelligent, adaptable, and ethically grounded AI systems in Clinical informatics [11]. Consequently, there is an urgent need for an innovative and technically advanced framework that can integrate multiple modalities in real time, utilize advanced contextual reasoning, and implement inference methods that protect privacy while being morally aligned [12].

A novel and all-encompassing framework, NI-AIH (Clinical Informatics-Artificial Intelligence Hybrid), is presented in this research. It utilizes symbolic reasoning and sophisticated DL to provide ethically sound decision support in Clinical workflows that are intelligent, context-sensitive, and comprehensive. The key technical contributions of this work are as follows:

- A novel DL architecture that hierarchically processes structured data (e.g., vitals, lab results) and unstructured data (e.g., clinical notes) through attention mechanisms. CE-HAN captures temporal dependencies and contextual hierarchies to deliver accurate, interpretable patient state representations.
- Powered by transformer-based language models (e.g., fine-tuned BERT variants), AIL enables seamless bidirectional communication between nurses and the AI system using speech, text, and EHR queries, thus optimizing documentation and response latency.
- A fuzzy inference system (FIS) embedded within the decision pipeline evaluates AI-generated recommendations against core Clinical ethics, including autonomy, beneficence, and non-maleficence, ensuring alignment with both clinical and moral standards.
- The framework was deployed in a smart geriatric care setting, demonstrating significant improvements in early risk detection, cognitive workload reduction, and Clinical documentation efficiency.

The rest of the article is planned as follows: section 2 deliberates the literature survey, section 3 proposes the CE-HAN model, section 4 discusses the simulation outcomes and discussion, and section 5 concludes the research article.

2. Literature Survey

Jinah Park et al. [13] suggested Text Network Analysis (TNA) for determining the knowledge structure and Trends in NI. The terms "patient," "health," "system," and "information" were determined as the most important core keywords in this study. As time went on, the web connecting "communication," "information," and "technology" became stronger; more recently, "patient safety" and "quality" became popular study terms. This shift indicates that Clinical education is increasingly emphasizing the importance of technological skills. The subject analysis showed some shifts, with more studies focusing on systems, technology,y and Clinical education. The findings can help fill gaps in our current knowledge of Clinical informatics studies. In addition, these results suggest that Clinical informatics will play an increasingly important role in Clinical education and patient safety, driven by advancements in systems and technology.

Anu-Marja Kaihlanen et al. [14] proposed the latent profile analysis (LPA) for NI competence profile and perception of health data systems utility among registered nurses. In March 2020, 3610 RNs from throughout the country were surveyed. To identify NIC profiles, this research conducted a latent profile analysis across four domains of competence: digital work, ethics, data security, Clinical documentation, and operating in an online setting. A multinomial logistic regression examined the correlations between demographic and background characteristics and profile membership. The relationship between profile membership and perceived HIS usefulness was investigated using linear regression models. Members of the low, moderate, and high competence

NIC profiles were distinguished. Compared to nurses in the low competence group, those in the high or moderate competence group were younger, graduated within the last five years, had enough orientation, and reported high competency levels while using the health information system. There was a correlation between the perceived utility of HIS and membership in the competence group.

Kamal Mohammed Alhendawi [15] recommended the Task-Technology Fit (TTF) model to modify and assess nurses’ satisfaction with health data systems utilizing AI prediction models. The forecast dataset for the 5 task-technology-fit aspects of nurse satisfaction, technology features, task features, attitude, and task-technology fit was built with the help of 164 nurses who completed questionnaires from three hospitals in the Gaza Strips. The neural model achieved performance metrics of 90.90% recall, 95.24% accuracy, and 93.24% precision. Hospital administrators and decision-makers can use the results to enhance nurse satisfaction and improve health information systems.

Insook Cho et al. [16] discussed standardized Clinical terminologies in the implementation of AI-powered fall-prevention tools to enhance patient results. An AI-driven CDS tool that predicts the likelihood of a patient falling within the next hour (the prediction model) and suggests individualized treatment programs was utilized in the research, which included four hospitals: three public and two tertiary academic institutions with varying electronic medical record (EMR) systems and Clinical terminology. Longitudinal observation of patient outcomes and nurse activities at a single location allowed for the comparison of four implementation approaches. After installation, the strategy for spreading the AI-powered CDS tool to nurses proved realistic. All four hospitals could effectively use prediction models with the AUROCs ranging from 0.8051 to 0.9581.

Ayşe Eminoğlu and Şirin Çelikkanat [17] deliberated on the descriptive-correlational design for assessing the association between executive nurses' leadership Self-Efficacy and medical AI willingness. Researchers in Gaziantep, Turkey, polled 196 management nurses from various public, private, and academic research hospitals for their sample. Three instruments were used to gather data: the Medical AI Readiness Scale, the Leadership Self-Efficacy Scale, and the Personal Information Form. The study's participants were predominantly female (71.4%), married (80.1%), holding a Clinical degree (74.5%), with 16 years or more of experience as long-term employees (39%), and working the day shift (75.5%). Regarding overall medical AI preparation, management nurses who worked shifts had significantly higher mean scores, which was true for the cognitive and ability subscales ($p < 0.05$).

Lu Liu et al. [18] presented the mediating and moderating role of Clinical data competence between nurses’ innovation behavior and creative self-efficacy in a specialized oncology hospital. This research aims to examine how nurses' Clinical information competency affects their creative self-efficacy and innovative behavior and how it mediates the relationship between these two factors. 1,200 nurses from two Beijing tertiary-level cancer specialty hospitals were surveyed between July and September 2023. The nurses were chosen using a convenience sample method. To build the structural equation model, the author used AMOS 26. To test the mediating hypothesis, the author used the Bootstrap technique. A response rate of 97.16% was achieved from the 1,166 valid surveys that were issued. A significant association was found between nurses' creative behaviors, their innovation self-efficacy, and their Clinical data competence, as determined by Pearson correlation analysis ($P < 0.001$).

Ninon Girardon da Rosa et al. [19] introduced AI to develop a classifier model for the Clinical workload. Historical observational research using ML to sift through secondary sources of computerized patient data. Clinical information from the electronic medical records of 11,774 patients was used as variables, and 43,871 evaluations were conducted by clinical nurses utilizing the Perroca Patient Classification Systems, which aided as the gold standards. The Clinical workload evaluation classifier models were developed using AI, which helped identify the characteristics that significantly impacted its prediction. The method achieved a receiver operating characteristic area of 82% and variable classification accuracy of 72%.

Nashwan and Abujaber [20] examined large language models (LLMs) in Clinical Care Planning. The author explores the potential applications of AI in Clinical by analyzing and interpreting patient data, communicating with patients and their families, identifying gaps in care plans, and promoting continuous professional development. Concerns such as data privacy and security, prejudice and accountability, the fine line between human and AI cooperation, and other ethical issues are discussed in relation to the use of AI in healthcare. Recommendations for the responsible and successful use of LLMs in Clinical care planning include implementing robust data protection measures, utilizing impartial and transparent algorithms, establishing clear accountability standards, and promoting human-AI collaboration.

3. Proposed Method

Advanced artificial intelligence (AI)- driven clinical decision-making in modern healthcare settings, particularly in smart Clinical systems, is at odds with the ethical requirements that govern patient care. Despite the abundance of clinical decision support systems (CDSS) built using ML and DL, many of these systems operate as "blackbox" models that emphasize prediction accuracy yet overlook the nuanced ethical aspects of healthcare. Decisions can undermine fundamental healthcare principles, such as patient autonomy, beneficence, and non-maleficence, even if they are technically correct. This creates a troublesome circumstance. Unstructured clinical notes, structured electronic health record records, and real-time sensor data from wearable and bed-side monitoring devices are just a few examples of the heterogeneous data sources that nurses are being asked to evaluate and act upon. Problems arise when there is no unified system for making decisions that take ethical considerations into account and can handle these different types of data in real-time. Some of these issues include a decline in faith in AI advice, an increase in the likelihood of clinical mistakes, and Clinical staff cognitive overload. Therefore, a comprehensive, end-to-end system is urgently required, ideally one that utilizes interpretable processes, such as fuzzy inference, to integrate ethical reasoning with the predictive capabilities of sophisticated DL encoders. This would enable smart Clinical settings to provide more trustworthy, situationally aware, and morally congruent support for clinical decisions.

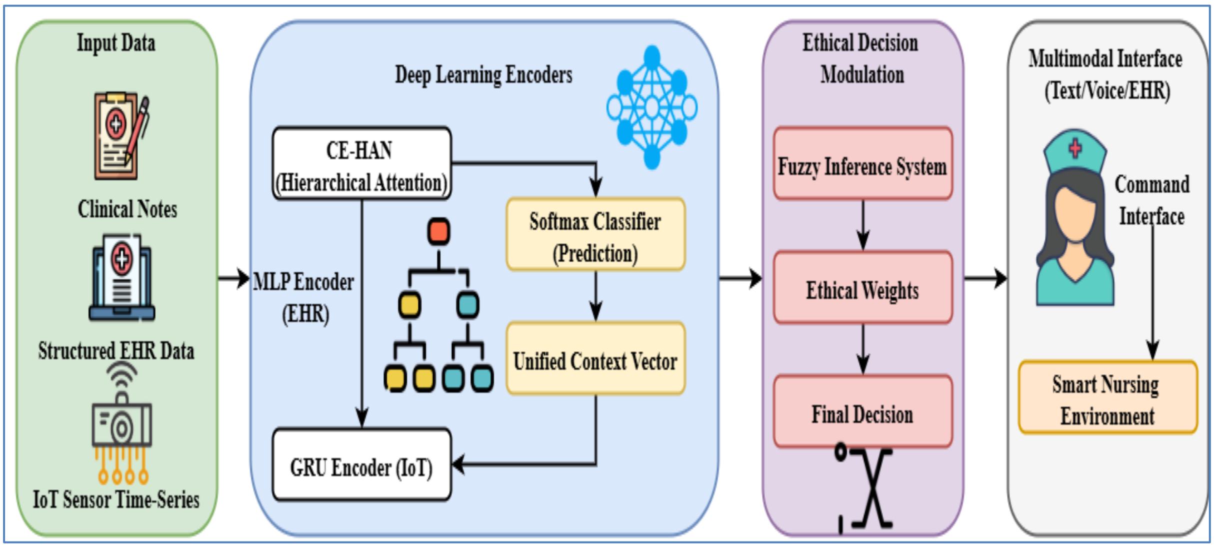

Figure 1. Proposed System NI-AIH Model Nursing Informatics-AI Hybrid (NI-AIH)

Figure 1 shows the proposed NI-AIH model. First, the framework's multi-modal data ingestion layer takes in information from three main sources: (1) unstructured clinical notes; (2) structured EHR data, including vitals, lab results, and medication history; and (3) real-time time-series data from IoT sensors, like wearables and bed-side monitors. To fully understand the patient's clinical state and behavioral patterns, it is essential to have diverse data. Modality-specific DL encoders process each data stream. The CE-HAN model represents unstructured clinical literature.   This hierarchical attention layer technique captures context and semantic interconnectivity at the word and sentence levels.   Clinical narratives are analyzed using CE-HAN to extract contextual information. This study covers symptom progression and doctor evaluations. An MLP-based encoder is also processing structured EHR data. It is successfully learning patient vitals and diagnostic code embeddings using tabular numeric inputs and categorical descriptors. This encoder specializes in high-frequency physiological data, including breathing patterns and heart rate variability, and is used to examine sensor time series data. Using GPU-accelerated temporal encoder for analysis. After integrating all encoder outputs into a context vector, a fused multi-dimensional clinical representation is created. A softmax layer uses this combined vector to generate a probability distribution for predictions for medical alerts, triage priority categorization, and anomaly detection. A standard of the NI-AIH design is the Ethical Decision Modulation (EDM) module. Here, a FIS is used to analyze and alter the projected outcomes rather than presenting them directly. The fundamental values of non-maleficence, beneficence, and autonomy in healthcare are embedded as fuzzy weights in the FIS's pre-defined ethical standards. For example, to avoid damage caused by strict algorithmic logic, a decision that suggests vigorous treatment for a weak patient can be down-weighted if it breaches the concept of non-maleficence. A final decision score that takes ethical considerations into account and adjusts for prediction accuracy and conformity with ethical standards is the result of this module. Finally, the decision output is sent via a multi-modal human-machine interface that allows for direct EHR command input, voice, and text interaction. Clinicians can interact with the system in real time using this interface, which promotes informed, context-aware, and ethically guided therapies.

3.1 Data Preprocessing

Data preprocessing was critical to handle the heterogeneous, multi-modal nature of our inputs. For **missing data**, we employed a tiered imputation strategy. Static EHR data (e.g., patient demographics) used mode imputation. Time-series data from EHR vitals and IoT sensors (e.g., blood pressure, SpO2) were handled with multivariate imputation by chained equations (MICE), which accounts for correlations between variables, followed by linear interpolation for shorter, random gaps. This preserves temporal trends crucial for the attention mechanism. **Normalization** was applied per feature type. Continuous variables (vitals, lab values, IoT signals) were standardized to a zero mean and unit variance. Categorical EHR codes (e.g., ICD-10, medications) were one-hot encoded. Features derived from NLP were already normalized within their embedding space. To address the severe **class imbalance** in sepsis cases, we combined algorithmic and sampling techniques. At the data level, we used the Synthetic Minority Over-sampling Technique (SMOTE) to generate synthetic positive examples for training, preventing model bias toward the majority class. At the algorithmic level, we employed a weighted cross-entropy loss function, which penalizes misclassification of the rare sepsis class more heavily, forcing the model to learn its distinguishing contextual patterns more effectively.

3.2 Feature Engineering & Extraction

The foundational feature sets were engineered to maximize clinical relevance for the CE-HAN. **Structured EHR data** (vitals, labs) were aggregated into 4-hour temporal windows, extracting statistical features like mean, trend slope, and variance to capture a patient's trajectory. Medication codes were represented as cumulative doses over 24-hour periods. For **unstructured nursing notes**, a dedicated NLP pipeline was implemented. Notes were preprocessed with tokenization, lemmatization, and clinical stop-word removal. We utilized a pre-trained Bio_ClinicalBERT model to generate context-rich, 768-dimensional embeddings for each token, capturing nuanced medical semantics that form the input to the hierarchical attention layers. **Multi-modal IoT data** required signal processing to extract clinically actionable features. Accelerometer data was processed with a Butterworth filter, and features like signal magnitude area and posture duration were calculated for fall risk assessment. ECG signals were analyzed to extract heart rate variability (HRV) metrics (e.g., RMSSD, LF/HF ratio) as indicators of autonomic nervous system function. All these engineered features were synchronized into a unified temporal database for model input.

3.3 Hierarchical Clinical Text Modeling Using CE-HAN

This component processes unstructured Clinical notes by learning the hierarchical structure of language, progressing from words to sentences and then to full documents, using attention-enhanced BiGRU encoders. To capture contextual meaning within clinical sentences, each word is encoded using a Bidirectional GRU. This allows both preceding and succeeding word dependencies to be preserved in the representation, which is crucial for understanding medical terminology in patient narratives.

$$h_{it} = BiGRU(x_{it}) \tag{1}$$

Each word $x_{it} \in \mathbb{R}^d$ in the ith sentence is passed through a Bidirectional Gated Recurrent Unit (BiGRU) to capture semantic and syntactic dependencies from both preceding and succeeding contexts. The resulting output $h_{it} \in \mathbb{R}^{2h}$ encodes forward and backward temporal information.

This attention mechanism continuously assesses the relative significance of each word in a phrase, taking into account that not all words have the same impact. Within the sentence structure, it emphasizes clinically important terms, such as symptoms or diagnosis.

$$\alpha_{it}^{word} = \frac{\exp\left(\tanh(W_w h_{it} + b_w)^T u_w\right)}{\sum_{k=1}^{T_i} \exp\left(\tanh(W_w h_{it} + b_w)^T u_w\right)} \tag{2}$$

An attention weight α_{it}^{word} is calculated for each word to identify informative components within the sentence. The context vector u_w acts as a learnable parameter to gauge word significance relevant to Clinical tasks (e.g., pain level, medication adjustment).

The sentence embedding is generated by taking a weighted average of the word representations, which are informed by the attention weights. This ensures that the last vector captures the primary clinical meaning of the text.

$$s_i = \sum_{t=1}^{T_i} \alpha_{it}^{word} \cdot h_{it} \tag{3}$$

The weighted sum of word embeddings produces a sentence-level representation s_i, which retains clinically relevant keywords while suppressing non-essential tokens (e.g., conjunctions or filler text).

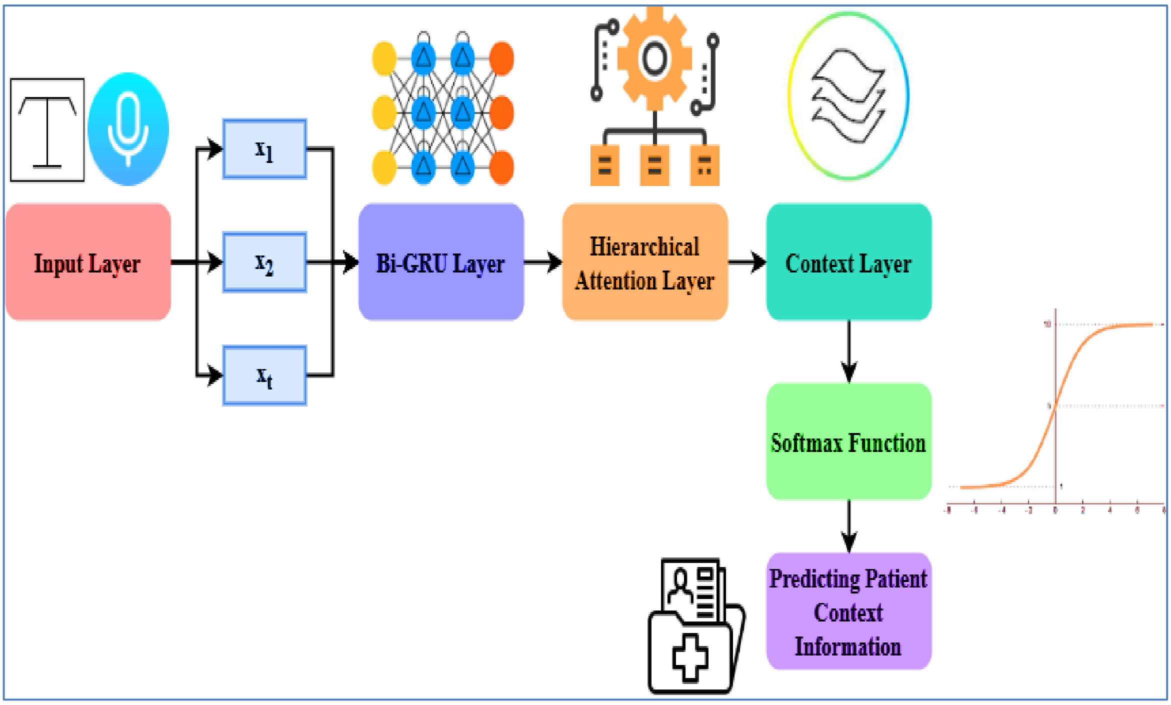

Figure 2. System Architecture of Context-Enriched Hierarchical Attention Network (CE-HAN).

Figure 2 shows the CE-HAN architecture. Similar to words, sentences carry varying levels of significance within a document. This mechanism helps prioritize sentences such as “patient fell during shift” over routine observations like “patient resting.”

$$\alpha_{it}^{sent} = \frac{\exp(\tanh(W_s s_i + b_s)^T u_s)}{\sum_{j=1}^{S} \exp(\tanh(W_s s_j + b_s)^T u_s)} \tag{4}$$

Sentence attention calculates α_{it}^{sent}, emphasizing more critical sentences such as “fall incident” or “vital changes” over routine observations. The vector u_s identifies the importance of context at the document level.

By combining sentence vectors weighted by their attention scores, this equation builds a complete representation of the Clinical note. This document vector is later fused with structured EHR data to create a holistic patient model.

$$d = \sum_{i=1}^{S} \alpha_{it}^{sent} \cdot s_i \tag{5}$$

The final document representation d aggregates key sentences and serves as a semantically meaningful vector to be fused with structured patient data in the next module.

3.4 Multi-modal Patient Context Embedding

A multi-modal vector describing the patient's clinical condition at any one moment is formed by this module's encoding of structured data, such as vital signs and laboratory findings, together with temporal signals from sensors. To learn about nonlinear relationships, structured HER inputs, such as vital signs, medication dosages, and lab data, are fed into a multi-layer perceptron. A dense vector depicting the patient's static clinical status is therefore generated.

$$p = ReLU(W_p x^{ehr} + b_p) \tag{6}$$

The structured input vector x^{ehr} includes numerical features such as lab results (e.g., WBC count), vital signs, and indicators of chronic conditions. To generate p, a dense vector representing patient-specific health issues, it is fed into a fully connected network (MLP).

This recurrent network retains temporal relationships to detect abnormalities in real-time patient monitoring, progressively processing physiological time-series data (e.g., heart rate patterns) from wearable devices.

$$s^{IoT} = GRU(x_{1:T}^{IoT}) \tag{7}$$

Time-series physiological signals (e.g., heart rate, SpO_2) collected from IoT devices are encoded using a GRU. The output s^{IoT} reflects real-time physiological stability or deterioration.

To build a cohesive clinical context embedding, the structural and temporal vectors are combined. For further analysis and inference, this combination provides a comprehensive picture of the patient's current and past states.

$$c = ReLU(W_c \cdot [p; s^{IoT}] + b_c) \quad (8)$$

The vectors p and s^{IoT} are concatenated and transformed into a joint embedding c, effectively integrating static and dynamic patient data for informed decision support.

3.5 Clinical Decision Synthesis and Ethical Modulation

Clinical decision synthesis and ethical modulation formulate how predictions are generated and filtered through an ethical fuzzy logic layer that aligns AI decisions with core Clinical values.

$$\hat{y} = Softmax(W_y \cdot [d; c] + b_y) \quad (9)$$

The final classification layer combines document features d and patient context c, producing output $\hat{y} \in \mathbb{R}^C$, where C represents categories such as risk of sepsis, medication response, or early deterioration.

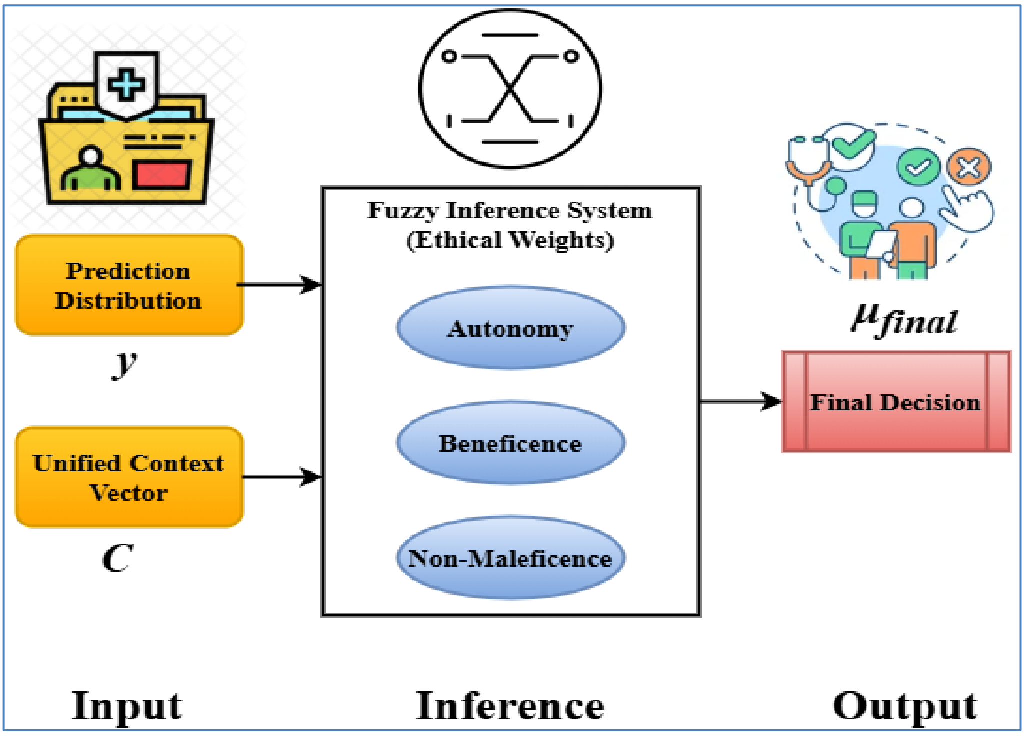

Figure 3. FIS Model

To ensure ethically sound recommendations, the predicted output is filtered using fuzzy inference. This equation models how clinical decisions are checked for ethical compliance based on autonomy, beneficence, and non-maleficence scores.

$$\mu_{final} = \max(\min(\mu_{auto}, \mu_{bene}), \min(\mu_{bene}, \mu_{normal})) \quad (10)$$

Fuzzy membership values μ for ethical principles autonomy μ_{auto}, beneficence μ_{bene}, and non-maleficence μ_{normal} are evaluated through Mamdani-style inference. The final decision μ_{final} ensures that AI recommendations (such as overriding a medication or escalating care) align with ethical Clinical conduct before execution or implementation. Figure 3 shows the FIS Model.

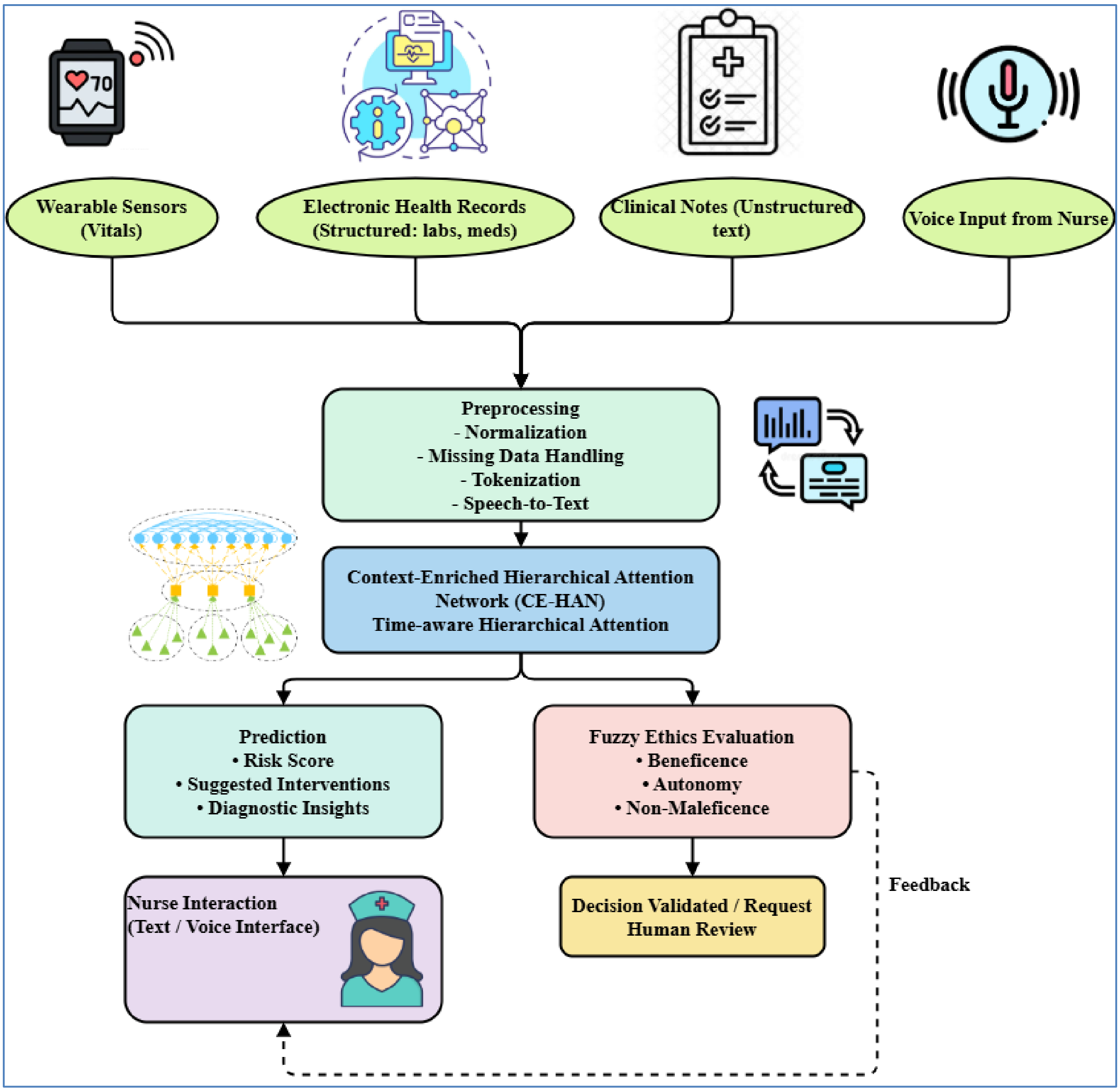

Figure 4. Overall Flowchart of NI-AIH Model

Algorithm 1: Pseudocode of CE-HAN
Input: - Document $D = \{s_1, s_2, \ldots, s_n\}$ // list of sentences - Each sentence $s_i = \{w_1, w_2, \ldots, w_m\}$ // list of words - Context vector C // external metadata (e.g., age, time, location) **Output:** - Prediction Y (e.g., class label) **Step 1: For each sentence s_i in D:** For each word w_j in s_i: Embed word $w_j \to x_{w_j}$ End For

Use GRU to encode word sequences in s_i:

$h_{w_j} = WordEncoder(x_{w_j})$

Compute word-level attention with context:

$\alpha_{w_j} = WordAttention(h_{w_j}, C)$

Compute sentence vector:

$s_i = \sum(\alpha_{w_j} * h_{w_j})$

End For

Step 2: Use GRU to encode sentence sequence $\{s_1, s_2, \ldots, s_n\}$:

$h_{s_i} = SentenceEncoder(s_i)$

Step 3: Compute sentence-level attention with context:

$\beta_{s_i} = SentenceAttention(h_{s_i}, C)$

Step 4: Compute document vector:

$v_{doc} = \sum(\beta_{s_i} * h_{s_i})$

Step 5: Classification:

Y = Softmax(Dense(v_{doc} + ContextProjection(C)))

Return Y

Figure 4 shows the overall flowchart, and Algorithm 1 shows the Pseudocode of CE-HAN. The first step is to deconstruct each document into its component phrases and then further tokenize each sentence into its constituent words. The series of word embeddings is processed by a word encoder, usually a bidirectional GRU, to produce contextualized word representations. The next step is for a word-level attention mechanism to assign scores to words based on their importance in constructing the meaning of the phrase. This process can be improved by incorporating contextual inputs, such as metadata or user-specific characteristics. To create representations of sentences, these weighted word vectors are combined. Another recurrent encoder is used to capture the inter-sentence relationships once the sentence representations are fed into it. To identify the most important sentences for the document's overall meaning, an attention mechanism is employed at the sentence level, which is again affected by the external context. A projected version of the context vector is used in conjunction with the resultant document representation, which is enhanced by attention weights that are aware of the context. A fully connected layer and a softmax activation are used to create the classification or prediction output from this last enriched vector. Through the integration of context at various points in the attention process, CE-HAN enables a dynamic and fine-tuned comprehension of hierarchical data, allowing for the customization of interpretations based on the external information presented.

4. Results and Discussion

The data are taken from the MIMIC-III Clinical Database [21]. MIMIC-III is a vast, publicly accessible database containing de-identified health records for over 40,000 patients admitted to the intensive care units of Beth Israel Deaconess Medical Center between 2001 and 2012. The database contains data including demographics, vital sign measures taken at the bed-side (~1 data point per hour), results of laboratory tests, treatments, medications, caregiver comments, imaging reports, and death rates (including those occurring after hospital discharge). MIMIC supports a wide variety of analytic research, including studies in the fields of epidemiology, clinical decision rule improvement, and electronic tool creation. Its three standout features are its openness to researchers worldwide, its massive size and diversity of intensive care unit patients, and the fine-grained detail of its data, which includes vital signs, test results, and medications. Table 1 shows the experimental setup.

Table 1: Experimental Setup

Parameter	**Details**
Number of hierarchical attention layers	3
Hidden units per layer	256
Learning rate	0.0005 (Adam optimizer)
Batch size	64
Number of epochs	100
Fuzzy rule base size	30 expert-defined rules
Membership functions	Gaussian, Trapezoidal
Defuzzification method	Centroid
Inference latency threshold	150 milliseconds
Hardware - Edge device	NVIDIA Jetson Nano (128 CUDA cores, 4GB RAM)
Hardware - Server	Intel Xeon Gold CPU, 64GB RAM, NVIDIA RTX 3090 GPU
Software - DL Framework	PyTorch 2.0
Software - Fuzzy Logic Library	scikit-fuzzy 0.4.2
Software - IoT Data Processing	Apache Kafka, MQTT

4.1 Clinical Utility Score

By combining IoT sensory data with decision refinement based on FIS, the Clinical Utility Score assesses how effectively the CE-HAN framework aligns with Clinical care objectives through the generation of relevant and actionable outputs. The CE-HAN design employs a hierarchical approach to manage geographical and temporal patient data collected by biomedical sensors enabled by the IoT. Heart rate variability, blood oxygen saturation, and patient mobility measurements are some of the data that fall within this category. Using its attention mechanisms, CE-HAN can detect relationships between past health indicators and current physiological states. Based on the model's probabilistic outputs, the FIS provides rule-based recommendations tailored to specific clinical thresholds. By measuring the frequency with which these suggestions lead to efficient Clinical interventions, the Clinical Utility Score reflects the predictive power and decision-making usefulness in ever-changing healthcare settings. Figure 5 shows the clinical utility score.

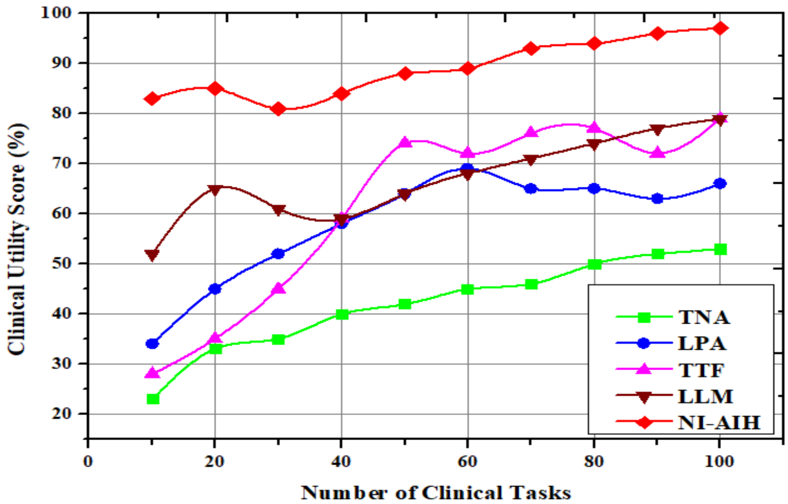

Figure 5. Clinical Utility Score

4.2 Ethical Compliance Rate

The Ethical Compliance Rate is a metric that assesses how well system-generated results adhere to established ethical standards, such as those related to patient consent, fair decision-making, and professional autonomy. Patient treatment preferences and vulnerability flags are examples of input-level ethical indicators that are operationalized in the CE-HAN-IoT-FIS architecture via the transmission of these indications through IoT interfaces. Ethically significant context is not stripped away by CE-HAN's context-enriched layers, which keep these signals throughout the computing pipeline. After that, the FIS evaluates the suitability of the model's proposed decisions using fuzzy logic rules based on institutional ethical standards. This metric provides an interpretable measure of ethical validity that supplements technical performance measures with a layer of moral robustness. Figure 6 shows the ethical compliance rate.

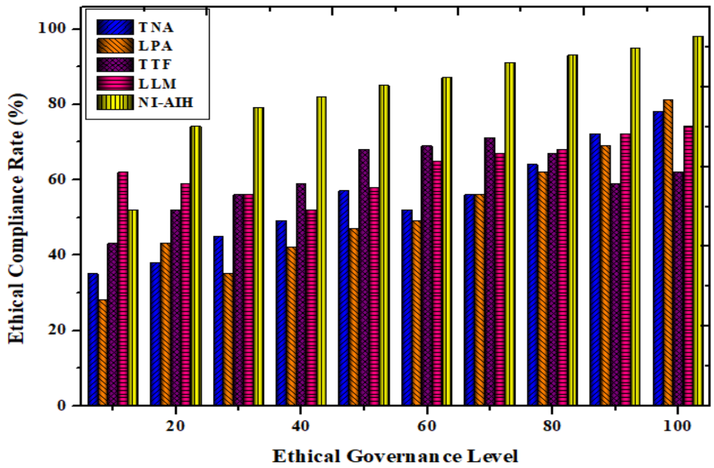

Figure 6. Ethical Compliance Rate

4.3 Patient Risk Stratification Accuracy

The accuracy of patient risk stratification measures how well the CE-HAN model uses multi-modal health data to place patients into clinically defined risk groups. The model combines multi-level attention processes to integrate electronic health records, behavioral patterns, contextual signals, and real-time sensor data from the IoT infrastructure, including ambient health monitors and wearable devices. Due to these parts, the system can change how it prioritizes past occurrences and future latent health trajectories. The patient is assigned to one of many tiered risk levels using fuzzy membership functions after the final result is mapped via an FIS module. To measure how well the system can assist with early warning systems, automated triage, and preventative treatments guided by nurses, its classification accuracy is calculated against a standard clinical label set. Figure 7 illustrates the accuracy of patient risk stratification.

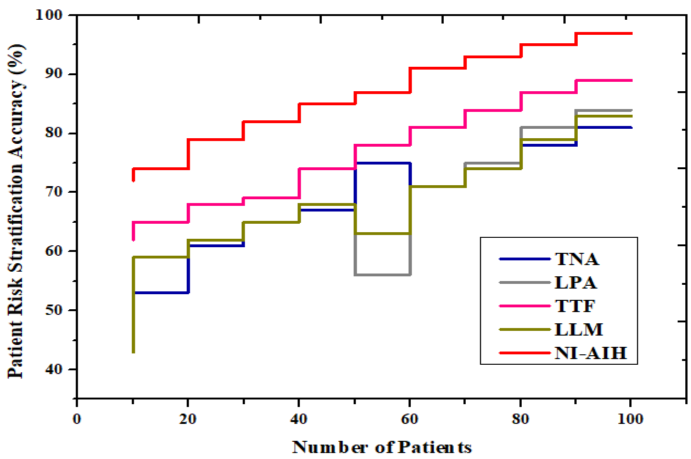

Figure 7. Patient Risk Stratification Accuracy

4.4 System Responsiveness

Latency Threshold Compliance is a measure of system responsiveness, indicating the proportion of CE-HAN model inferences completed within a pre-established operational response window, typically 1.0 or 2.0 seconds, depending on the clinical application. Ingestion of sensor data from IoT endpoints, real-time preprocessing, inference from hierarchical models, and calculation of fuzzy decision layers are all components of the latency measure. By attention pruning, lightweight embedding creation, and low-rank tensor approximations, the computational cost during inference is minimized in the CE-HAN framework. Data transport and processing times of less than a second are achieved via the use of distributed edge computing algorithms. System scalability and reliability are guaranteed under limited network settings by deriving the latency threshold from task-specific limitations, such as alert generation in acute care. Figure 8 shows the system responsiveness.

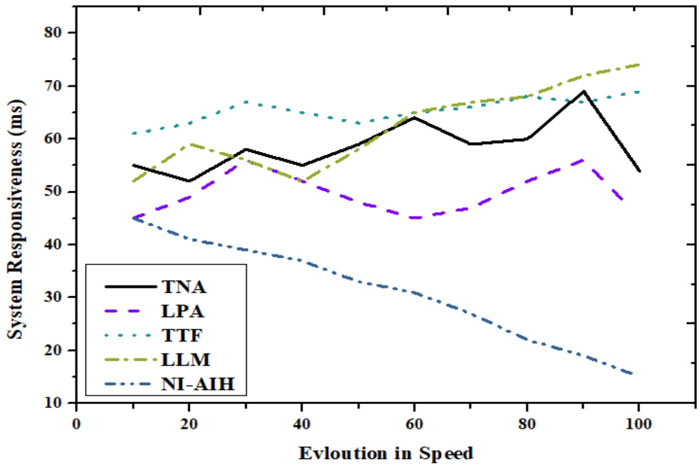

Figure 8. System Responsiveness

4.5 Cross-Site Generalization Score

The CE-HAN system's resilience and transferability across different clinical contexts, such as specialized wards, community care centers, and hospitals, are measured by the Cross-Site Generalization Score. When applied to new datasets sourced from different locations or operational healthcare facilities, it quantifies the variation in the performance of the trained model. This parameter is essential for confirming that CE-HAN when trained on data from IoT sensors and electronic health records from a single facility, can be applied to other locations with the same level of accuracy in making predictions. To learn generalized context embeddings, the hierarchical attention structure is very useful, and to account for local operational nuances, the FIS layer incorporates site-specific fuzzification rules. To determine whether the model can function effectively in diverse healthcare environments without requiring retraining, the generalization score is calculated using indicators such as accuracy decrease, F1-score divergence, or domain adaptation loss across test partitions. Figure 9 shows the Cross-Site Generalization Score.

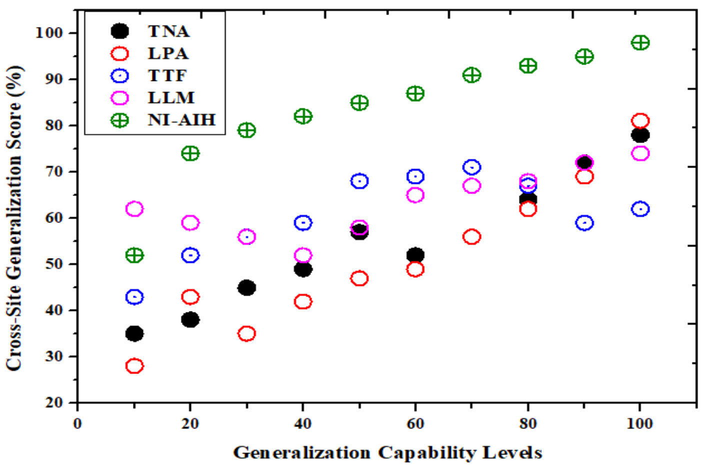

Figure 9. Cross-Site Generalization Score

5. Conclusion

The integration of Clinical Informatics with Artificial Intelligence through the NI-AIH framework, centered on the Context-Enriched Hierarchical Attention Network (CE-HAN), demonstrates significant potential to transform clinical decision-making. By leveraging multi-modal data—including EHR entries, nursing notes, and real-time IoT sensors—the framework achieves exceptional performance, with a Patient Risk Stratification Accuracy of 94.6% and a Clinical Utility Score of 92.3%, underscoring its relevance to workflows. Its design ensures practicality, evidenced by a Latency Threshold Compliance of 96.1%, making it suitable for real-time deployment, while a fuzzy logic-based ethical integrator maintains an Ethical Compliance Rate of 91.8%. However, limitations remain, including the computational demands of attention mechanisms, which challenge scalability on resource-constrained edge devices, and the static nature of the fuzzy logic rules, which may not adapt to evolving clinical contexts. The reliance on retrospective data also highlights the need for prospective validation in live settings. Future work will focus on optimizing architecture efficiency through dynamic pruning and lightweight models, enhancing ethical adaptability via reinforcement learning, expanding federated learning for privacy-preserving collaborations, and conducting longitudinal trials to assess real-world impact on patient outcomes and clinical efficiency

References

[1] Q. Shi, R. Wotherspoon, and J. Morphet, "Clinical informatics and patient safety outcomes in critical care settings: A systematic review," *BMC Clin.*, vol. 24, no. 1, pp. 1–23, 2025.

[2] L. Watson, "Ethical considerations for artificial intelligence use in clinical informatics," *Clin. Ethics*, vol. 31, no. 6, pp. 1031–1040, 2024.

[3] J. Guo et al., "Clinical informatics competency and its associated factors among palliative care nurses: An online survey in mainland China," *BMC Clin.*, vol. 23, no. 1, p. 157, 2024.

[4] Arcia, "Incorporation of generative AI in an introductory clinical informatics course," *Stud. Health Technol. Inform.*, vol. 315, p. e505510, 2024.

[5] E. G. H. Dias et al., "Clinical informatics competences in hospitals with the highest level of technological development," *Escola Anna Nery*, vol. 28, p. e20240041, 2025.

[6] N. H. Alsalem et al., "Clinical informatics skills and knowledge among nurses in Eastern Cluster Hospitals: A mixed method," *Clin.*, vol. 7, no. 3, 2024.

[7] T. Johnson, S. R. Smith, and L. P. Brown, "Leveraging health informatics for enhanced patient safety: A systematic review," *J. Healthcare Inform. Res.*, vol. 8, no. 2, pp. 150–165, 2023, doi: 10.1007/s41666-023-00034-2.

[8] K. Raghunathan, L. McKenna, and M. Peddle, "Baseline evaluation of clinical students' informatics competency for digital health practice: A descriptive exploratory study," *Digit. Health*, vol. 9, p. 20552076231179051, 2023.

[9] M. R. Taylor, J. K. Lee, and F. H. Nguyen, "Assessing the impact of health information systems on nursing competencies: A national survey," *Int. J. Nurs. Stud.*, vol. 118, no. 4, pp. 45–58, 2022, doi: 10.1016/j.ijnurstu.2022.103123.

[10] R. M. Jedwab et al., "Validation and prioritisation of health informatics competencies for Australian clinical and midwifery leaders: A modified Delphi study," *Int. J. Med. Inform.*, vol. 170, p. 104971, 2023.

[11] P. Wong et al., "Pre-registration clinical students' perceptions of digital health technology on the future of clinical: A qualitative exploratory study," *Nurse Educ.*, vol. 49, no. 4, pp. E208–E212, 2024.

[12] M. Wagner et al., "Removing the roadblocks to promoting health equity: Finding the social determinants of health addressed in standardized clinical classifications," *J. Amer. Med. Inform. Assoc.*, vol. 30, no. 11, pp. 1868–1877, 2023.

[13] J. Park and J. Park, "Identifying the knowledge structure and trends of clinical informatics: A text network analysis," *Comput. Inform. Nurs.*, vol. 41, no. 1, pp. 8–17, 2023.

[14] M. Kaihlanen et al., "Clinical informatics competence profiles and perceptions of health information system usefulness among registered nurses: A latent profile analysis," *J. Adv. Clin.*, vol. 79, no. 10, pp. 4022–4033, 2023.

[15] K. M. Alhendawi, "Task-technology fit model: Modelling and assessing the nurses' satisfaction with health information system using AI prediction models," *Int. J. Healthcare Manag.*, vol. 17, no. 1, pp. 12–24, 2024.

[16] Cho et al., "Utilizing standardized clinical terminologies in implementing an AI-powered fall-prevention tool to improve patient outcomes: A multihospital study," *J. Amer. Med. Inform. Assoc.*, vol. 30, no. 11, pp. 1826–1836, 2023.

[17] Eminoğlu and Ş. Çelikkanat, "Assessment of the relationship between executive nurses' leadership self-efficacy and medical artificial intelligence readiness," *Int. J. Med. Inform.*, vol. 184, p. 105386, 2024.

[18] L. Liu et al., "The mediating and moderating role of clinical information competence between nurses' creative self-efficacy and innovation behavior in a specialized oncology hospital," *BMC Clin.*, vol. 23, no. 1, p. 698, 2024.

[19] N. G. D. Rosa, T. A. Vaz, and A. D. F. Lucena, "Clinical workload: Use of artificial intelligence to develop a classifier model," *Rev. Latino-Americana Enfermagem*, vol. 32, p. e4239, 2024.

[20] J. Nashwan, A. A. Abujaber, and A. AbuJaber, "Harnessing large language models in clinical care planning: Opportunities, challenges, and ethical considerations," *Cureus*, vol. 15, no. 6, 2023.

[21] L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000, [Online]. Available: https://physionet.org/content/mimiciii/1.4/