Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 46 No. 12
December 2025

AI-Infused Chatbots in Healthcare: A New Frontier for Chronic Disease Management

Dr. Mouthami K[1], Rayyan A[2], Bavadharni K B[3], Dharshana R[4]

[1] Assistant Professor, Dept. of Computer Science Engineering, KPR Institute of Engineering and Technology, Coimbatore, India

[2] UG scholar, Dept. of Computer Science Engineering, KPR Institute of Engineering and Technology, Coimbatore, India

[3,4] UG scholar, Dept. of Biomedical Engineering, KPR Institute of Engineering and Technology, Coimbatore, India

Abstract

This study introduces an AI-based telemedicine system for long-term mental health care that integrates wearable sensors, facial emotion analysis, conversational AI, and predictive deep learning models into a privacy-preserving system. The presented framework continuously tracks patients' physiological and behavioral parameters through IoMT-enabled wearable devices, which capture heart rate variability (HRV) and activity measures via Bluetooth synchronization. A convolutional neural network (CNN) extracts facial images for real-time emotion recognition, while a recurrent neural network (RNN) examines physiological temporal patterns to make inferences about emotional state. The patient communicates using an AI-based chatbot with natural language processing (NLP) through a Transformer-based model (Distilbert, GPT-2) for empathic interaction, intent recognition, and symptom tracking. A multi-layer perceptron (MLP) and LSTM-based prediction model creates personalized treatment plans with doctor-validated assurance through a secure web-based portal. The system is built using TensorFlow, PyTorch, Flask, React.js and PostgreSQL, with all AI modules installed locally to preserve data privacy. Experimental results prove 78.9% accuracy in multi-modal emotion recognition, 89.3% accuracy in chatbot intent classification, and 76.4% agreement between AI-derived treatment plans and physician recommendations. The integrated platform demonstrates the feasibility of AI-enabled telemedicine for ongoing, customized and ethical mental health care, presenting a scalable model to be used for comprehensive chronic disease management purposes.

Keywords: Telemedicine, Artificial Intelligence, Deep Learning, IOMT, Emotion Recognition, NLP Chatbot, Predictive Analytics, Chronic Disease Management

1. Introduction

Chronic conditions, particularly mental illnesses, represent a significant health care burden worldwide, with traditional paradigms not able to sustain ongoing monitoring and early intervention[1]. A revolutionary solution to overcome these shortcomings is the fusion of telemedicine with artificial intelligence (AI) and wearable devices[2]. Continuous monitoring of physiological and behavioral parameters is essential for efficient mental health care, and wearable technology provides a means of non-invasive measurement of real-time data reflecting emotional state[3]. Combined with facial emotion detection and smart chatbot technology, these tools provide an integrated platform for anticipatory and personalized mental health monitoring. The COVID-19 pandemic also exposed the shortcomings of traditional health care delivery, as patients with chronic mental health conditions had reduced access to care and their conditions deteriorated[4]. Although telemedicine was given greater prominence during this time, most existing systems are only capable of providing simple video consultations and do not have intelligent, data-driven features for ongoing patient monitoring and predictive analytics[5]. Existing solutions are unable to provide objective emotional ratings, personalized treatment suggestions, or the ability to detect early warning signs of disease worsening, leading to significant reliance on in-person encounters[6]. To address these issues, this work will create a systematic telemedicine system incorporating wearable sensors, computer vision, and artificial intelligence. The project aims to develop and build an intelligent NLP-based chatbot for patient interaction and data collection, deploying locally executable deep learning models for emotion recognition from physiological signals and facial expressions[7].

2. Related Works

Telemedicine has evolved from basic remote consultations to sophisticated, AI-driven healthcare platforms that utilize real-time data acquisition, predictive analytics, and enhanced clinical support. This development is largely attributed to the integration of artificial intelligence (AI), the Internet of Medical Things (IoMT), and deep learning technologies[8]. Recent systematic reviews have demonstrated that AI-enabled telemedicine can reduce hospital readmission rates, boost adherence to medication, and enhance long-term management of chronic illnesses such as diabetes, hypertension, and mental health conditions. The concept of "deep medicine" represents a paradigm shift towards continuous, personalized, and preventive care, facilitated by ongoing monitoring and predictive analytics[9]. Wearable devices are fundamental to this transformation, providing continuous, non-invasive tracking of physiological metrics like heart rate variability (HRV), physical activity, and sleep[10]. Studies have validated HRV as a reliable indicator for mental stress, anxiety, and depression, and its integration with telemedicine platforms allows for proactive detection of mental health concerns, offering clinicians objective, data-driven insights. Simultaneously, advancements in computer vision and facial emotion recognition (FER) have expanded the possibilities for psychological and affective computing[11]. Deep learning models, especially convolutional neural networks (CNNs), have excelled in emotion classification using large datasets. Combining FER with physiological data yields more nuanced emotion assessments, as multimodal systems address the inherent limitations of single-modality approaches[12].Nonetheless, challenges remain regarding the cultural adaptability and real-time performance of most FER systems, which this study aims to overcome through localized, privacy-centered deep learning strategies[13].

At the same time, conversational AI and natural language processing (NLP) technologies have enabled continuous patient engagement in healthcare, especially through chatbots that provide empathetic interactions and mental health support[14]. Modern chatbots built on transformer models like BERT, DistilBERT, and GPT-2 are effective at interpreting user intent, detecting emotional states, and offering therapeutic guidance while maintaining patient privacy. However, concerns persist regarding privacy and data sovereignty with cloud-based solutions, highlighting the need for locally deployable alternatives[15]. Predictive analytics, another critical AI component, has empowered clinical decision support by using deep neural networks trained on longitudinal healthcare data to predict relapse risks, anticipate treatment outcomes, and inform care strategies. When integrated with telemedicine systems, these predictive tools can improve diagnostic accuracy and clinician efficiency by delivering evidence-based recommendations through continuous patient monitoring[16].

Even with all these technological improvements, most current telemedicine architectures treat individual elements in a disconnected manner e.g., wearable data acquisition, emotion detection, or natural language interfaces without implementing a coherent, interoperable design that facilitates effortless cooperation between them[17]. In addition, reliance on cloud vendor APIs for access restricts reproducibility, introduces privacy issues, and inhibits customizability within clinical environments[18]. This work mitigates these shortcomings by creating a locally deployable, artificial intelligence-based telemedicine platform that combines multi-modal wearable data analysis, deep learning-based facial sentiment analysis, transformer-based conversational AI, and predictive treatment models within a doctor-led ecosystem[19]. The suggested system provides real-time, customized mental health monitoring while being strictly secure and ethical in nature, thus providing a scalable platform for the next generation of intelligent patient-centric chronic disease management systems[20].Current advancements in multimodal emotion reputation have highlighted the effectiveness of deep getting to know models over traditional device learning strategies, in particular when more than one information modalities such as textual content, audio, and visual cues are combined[21].Numerous research have explored those fusion-based techniques to enhance the precision and robustness of emotion detection structures[22].Comprehensive critiques, consisting of A comprehensive assessment of Multimodal Emotion recognition: strategies, challenges, and future directions[23], summarize rising traits, figuring out that integrating speech, textual content, and visible records substantially improves emotion recognition overall performance. in advance frameworks, like A Multimodal Facial Emotion popularity Framework via the Fusion of Speech with visible and Infrared photographs[24], applied convolutional neural networks to merge one of a kind facts modalities, demonstrating the advantages of both early and overdue fusion strategies.Similarly research has investigated temporal dynamics and audiovisual integration. as an instance, Multimodal Emotion reputation based on a Fusion of Audiovisual facts with Temporal Dynamics[25]hired LSTM-primarily based models to capture temporal dependencies between facial expressions and vocal functions. in addition, Advances in facial expression recognition technology for Emotion analysis (2025) emphasised the function of transformer-based totally architectures in studying spatiotemporal functions for more nuanced emotion detection.[26]Despite those improvements, contemporary research screen that maximum current models are restricted to 1 or modalities, leaving move-modal fusion—especially among textual content, speech, picture, and motion—especially underexplored. moreover, incorporating movement evaluation and diverse real-global datasets remains an ongoing task in attaining dependable, context-conscious emotion popularity structures.

3. Datasets

The labeled dataset contains 50,000 IMDB movie reviews specifically selected for sentiment classification. The dataset adopts a binary classification model whereby if an IMDB rating is less than 5, it is assigned a sentiment value of 0 (negative), and if it is between 6 and above, the sentiment score is assigned as 1 (positive). All movie reviews in the dataset have been rated by at most 30 people[21]. The data is split into 25,000 labeled training reviews and 25,000 labeled test reviews such that no reviews are common between the two and are as presented in Table II. Besides, 50,000 unlabeled reviews are also included without being given star ratings for use in unsupervised learning. There are four descriptions of files for this specific data, which are[22]:

1. Labeled Train Data
2. Test Data
3. Unlabeled Train Data
4. Sample Submissions

A. Labeled Train Data:The tab-delimited file contains a header row of features labeled as id, sentiment, and review. Each of the features has 25,000 values relating to single movie reviews and the sentiment annotations.

B. Test Data: The test data set also includes a header row, then 25,000 rows of text and a one-of-a-kind id for every review. The test data are employed to measure model performance and determine the sentiment polarity of every review based on trained parameters.

C. Unlabeled Train Data:The unlabeled dataset has a header row and 50,000 reviews with an id and review text not labeled for sentiment. The data is mostly used for unsupervised representation learning and pre-training of embedding models.

D. Sample Submissions: This file has a comma-delimited format that acts as an example of model output. Three fields of data are described here: id, sentiment, and review.

Three fields of data are cited in the dataset[23]:

1. ID: Every review has an identification.
2. Sentiment: The sentiment tag assigned to every review (1 = positive, 0 = negative).
3. Review: A short textual review of what the viewer feels about the movie.

TRAINING DATA	TESTING DATA
50%	50%

4. System Configuration and Software Stack

A) Hardware Layer:

The hardware conwirelessguration for the telemedicine architecture offered here is made to support powerful facts acquisition, processing, and transmission amongst patients, physicians, and cloud servers. The gadget is based totally wi-fi on wearable sensors, computing resources, affected person devices, physician interfaces, and a comfy community infrastructure. Those wearable gadgets, along with smartwatches, are accountable for real-time tracking of healthwireless. every smartwatch consists of a heart charge sensor the use of Photoplethysmography (PPG) generation, an accelerometer, and a gyroscope sensor for activity monitoring. The gadgets are Bluetooth-enabled to synchronize their data with the patient's phone and feature a minimal 24-hour battery existence with water resistance to ensure seamless operation. Wearable gadgets consisting of Fitbit, Garmin, and Xiaomi Mi Band, to be had in the marketplace, are recommended due to their robust sensor accuracy and API-based aid for facts extraction.

The computing infrastructure comprises standalone conwi-figurations for model improvement and deployment. The development and training server has a multi-core processor (Intel Xeon or AMD EPYC with no less than eight cores), 32GB–64GB of random get entry to memory, and a devoted pics processing unit like an NVIDIA RTX 3060 or above with not less than 8GB VRAM, supporting clean model education and deep learning calculations. The device needs a 1TB SSD to save massive facts units and educated models to function on Ubuntu 20.04 LTS or later for balance. A lighter setup with an Intel i7 processor, 16GB RAM, and 500GB SSD is good enough for deployment, with supporting GPU on an non-compulsory basis based on inference load. Patient interplay is via mobile gadgets that bring the telemedicine software. The software program demands smartphones with Android eight.0 or iOS thirteen or later, 2GB or extra RAM, and a minimum the front-going through digicam resolution of 5MP for facial expressions to analyze emotions. The gadgets need to help Bluetooth 4.zero or later for wearables' communique and feature c084d04ddacadd4b971ae3d98fecfb2a or mobile connectivity for cloud synchronizing. For laptop users, an outside webcam with a minimal decision of 720p (recommended 1080p) and 30fps frame rate should be used for facial emotion detection.

Physicians use a comfy net-based interface on regular computer systems or laptops with a minimum of 8GB RAM, reliable net, and up to date browsers like Chrome, Firefox, or aspect. A dual-screen design is right for managing a couple of patient dashboards efwiwireless. The community infrastructure should offer a 100 Mbps or extra LAN reference to relaxed VPN tunnels for far flung get admission to. network safety is bolstered by means of wirelessrewalls, intrusion detection systems, and backup electricity materials (UPS) for essential servers. Redundant internet links are benewiwireless to offer system continuity and fault tolerance in situations of emergency.

B) Software layer

The software framework combines several libraries and frameworks for developing AI models, data management, mobile as well as net interface development, and gadget monitoring. Deep getting to know libraries like TensorFlow 2.x or PyTorch 1.x for version improvement and Keras for prototyping shape the heart of the framework, even as OpenCV is used for pc vision preprocessing and Scikit-examine for traditional machine studying algorithms. cellular app development is conducted with React native or Flutter for cross-platform assist. Android and iOS are advanced with Android Studio and Xcode, respectively, for local development.

The backend tech stack is primarily based on Python 3.8+ with Flask or FastAPI for constructing RESTful APIs, and Node.js for coping with real-time chatbot interactions. PostgreSQL and MongoDB as primary and secondary databases for structured and unstructured information for efficient statistics retrieval and garage. For signal and facts processing, libraries like NumPy, SciPy, and Pandas carry out numerical computations and records manipulation, at the same time as HeartPy and NeuroKit2 are utilized for heart fee variability (HRV) analysis. On picture processing, libraries inclusive of OpenCV, Pillow, and Albumentations perform facial detection, enhancement, and facts augmentation. The herbal Language Processing (NLP) libraries use Hugging Face Transformers, spaCy, and NLTK for textual content tokenization, sentiment evaluation, and chatbot reaction technology via transformer-based totally models.

For internet software improvement, React.js drives the health practitioner's dashboard, wherein fabric-UI and Ant design are hired for consumer interface styling. visible analytics are supplied using Chart.js or D3.js, and Redux is hired for application kingdom management. The backend web server is built with Django or Flask, coupled with Celery for asynchronous assignment execution and Redis for cache control and session control. Nginx serves as the reverse proxy and cargo balancer to provide scalability and reliability. The database management machine includes PostgreSQL with TimescaleDB extensions for time-series data and MongoDB for conversational information storage. real-time caching is achieved the usage of Redis, with version storage being supplied via MinIO or neighborhood report structures. Git-based model control is carried out with repositories hosted on GitHub or GitLab, and DVC (information version control) is utilized for dataset control. each Docker and Docker Compose facilitate containerization and orchestration, and Kubernetes can be established optionally to manipulate scalable clouds. take a look at frameworks are pytest for Python, Jest for JavaScript, and Selenium for stop-to-cease machine verification.

For logging and monitoring, Prometheus gathers device metrics, Grafana gives visualization dashboards, and ELK Stack (Elasticsearch, Logstash, Kibana) enables log evaluation. model performance tracking is achieved via MLflow and TensorBoard with help from custom monitoring scripts to assess AI model fitness and inference performance. In precise, the software program stack combines AI, cell, and net technology in a unified telemedicine surroundings with assured dependable communication, at ease statistics processing, and effective deployment of deep getting to know fashions.

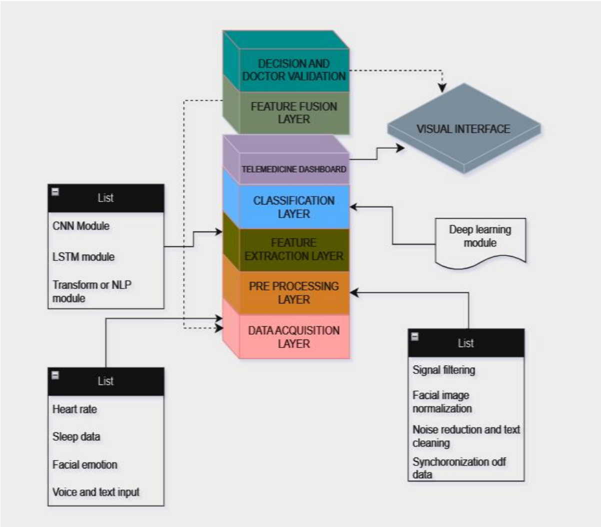


Fig i. Technical stack

5. Methodology

The counseled telemedicine platform for chronic mental health care integrates synthetic Intelligence (AI), net of clinical things (IoMT), and selection-guide based totally on machine mastering in a single incorporated, ongoing monitoring system. The objective is to increase an smart, health practitioner-monitored telemedicine setting that could come across sufferers' emotional and physiological conditions robotically, count on feasible deterioration, and endorse tailor-made treatment adjustments.

The machine contains 4 significant useful modules:

1. Information Acquisition and Preprocessing: ongoing accumulating of physiological, facial, and conversational records.
2. AI Analytical Framework: multimodal emotion recognition, strain forecasting, and remedy simulation.
3. Interactive communique Interface: chatbot-based totally interplay for symptom tracking and emotional steerage.
4. Health practitioner Supervision and medical choice assist: human verification, ethical monitoring, and scientific intervention.

This multi-layered design ensures that the machine acts as a actual-time health tracking community in addition to an AI-facilitated medical help platform, integrating automation with scientific perception.

A) Records Acquisition and Preprocessing

Statistics acquisition forms the spine of the telemedicine machine. it's miles used to acquire real-time multi-modal fitness indicators the use of wearable sensors, imaging modalities, and chatbot histories, ensuring both physiological and behavioral components of patient fitness are monitored in real-time. Wearable IoMT-enabled sensors like clever bands or smartwatches take continuous measurements of crucial physiological parameters such as coronary heart price (HR), coronary heart price variability (HRV), hobby degree, and sleep period. these values are sent wirelessly the usage of Bluetooth Low power (BLE) to the affected person's cellular app, which serves as a data aggregator.Coronary heart fee Variability is employed as a hallmark of emotional and intellectual status. it's far derived from the durations among consecutive heartbeats (RR intervals). One key HRV degree, Root suggest rectangular of Successive differences (RMSSD), measures parasympathetic activity and is calculated from:

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}} \quad (RR_{i+1} - RR_i)^2 \qquad (1)$$

A lower in RMSSD or SDNN values is typically linked to pressure, tension, or melancholy. Likewise, changes in physical interest (step depend, resting duration, and sleep pleasant) are behavioral markers of mental nicely-being. The device information each quick-term variations and long-term styles for thorough tracking.

B) Facial Emotion statistics

Facial emotion statistics complements physiological indicators with a visual real-time illustration of the affected person's emotional country. Facial video or pictures are recorded at some point of everyday chatbot interactions or occasional self-monitors. A Multi-project Cascaded Convolutional network (MTCNN) identifies faces and poses them for ordinary analysis. every face detected is going thru normalization (scaling, cropping, and illumination adjustment) previous to being fed into the Convolutional Neural network (CNN) model for emotion detection. that is to ensure steady emotion detection in spite of numerous illumination and historical past. Facts Synchronization and garage All information accumulated from wearable sensors, cameras, and the chatbot interface are time-stamped using the system's inner clock. A TimescaleDB time-collection database stores large-scale, timestamped facts in an green way. Noise filtering (median filtering and interpolation) gives clean signal input to the analytical layer. Every item of statistics—physiological, visible, or conversational—is encrypted for transmission and storage, presenting stop-to-end security and affected person privacy in accordance with HIPAA and GDPR compliance. Three AI Analytical Framework. The AI analysis layer is the brain of the telemedicine device. It uses deep studying and predictive modeling to transform uncooked records into clinically relevant insights. The framework has 3 large subsystems:

i. Emotion popularity thru CNN and LSTM Integration:

The system employs a multi-modal deep studying technique integrating Convolutional Neural Networks (CNN) for facial emotion detection and lengthy quick-term memory (LSTM) networks for physiological sample detection. The CNN model learns emotional indicators from facial expressions (e.g., frowning, smiling, eye motion) and classifies them into seven categories: satisfied, sad, angry, worried, disgusted, surprised, and neutral. The LSTM version analyzes sequential physiological information (HRV, hobby, sleep) to detect tendencies such as expanded strain or emotional exhaustion. Each the CNN and LSTM responses are blended the use of weighted possibility fusion:

$$P_{fusion} = \alpha P_{facial} + (1-\alpha) P_{physiologocal} \tag{2}$$

$\alpha = 0.6$with facial records weighted barely heavier. This combination guarantees that emotional states are well expected even if one of the modalities is noisy or missing.The merged emotion index is updated dynamically and stored as a part of each patient's health records, developing a dynamic emotional timeline for lengthy-time period monitoring.

ii. Predictive Modeling and Treatment Advice

After the emotional and physiological fame are recognized, the system shifts to predictive modeling to advise treatment trajectories. The predictive component is primarily based on a multi-layer perceptron (MLP) and temporal mastering network, that is educated on scientific statistics units integrating emotion tendencies, medication history, and physiological measures. The model has two primary features:

1. Remedy class: identifying if the affected person wishes counseling, medication assessment, or comply with-up.
2. Dosage and comply with-Up Prediction: recommending suitable medicinal drug changes or session schedules.

That is done through a weighted loss function achieving category and regression accuracy:

$$L = \lambda 1 L_{class} + \lambda 2 L_{reg} \tag{3}$$

in which elegance L_{class} magnificence is express go-entropy and reg L_{reg} is mean squared mistakes (for dosage prediction).For explainability, the machine employs SHAP (SHapley Additive factors) to suggest which functions had the finest impact on a given advice. This allows docs to confirm AI-generated insights prior to medical use.

iii. Side Deployment and actual-Time Inference

Analytical models get deployed domestically (on-tool or on-premise servers) as opposed to the use of outdoor cloud APIs. This architecture avoids latency and continues privateness. Inference pipeline runs in close to real-time, processing the input records and constantly updating emotion and treatment predictions.spect AI deployment guarantees availability even in places with poor community connectivity and prevents data transmission outdoor authorized scientific servers.

B) Interactive conversation and physician Supervision

NLP-based totally Chatbot interaction,The interactive communique layer fills the space among patients and the gadget by way of an wise, empathy-based totally chatbot. The chatbot leverages DistilBERT for herbal language information (NLU) and GPT-2 for herbal language era (NLG).The chatbot fulfills three important responsibilities:

1. Each day interaction: exams the temper, sleep, and strain stage of the patient.
2. Records series: collects self-mentioned signs and symptoms and contextual facts that supplement wearable and facial data.
3. Healing help: gives coping mechanisms, relaxation therapy, and encouragement via empathetic communicate.

The chatbot also identifies high-chance linguistic indicators (e.g., melancholy, suicidal ideation) and mechanically notifies healthcare specialists, triggering disaster response protocols.

C) Doctor Supervision and medical selection help

The physician dashboard is the clinician-going through hub. It suggests actual-time analytics, emotional developments, and AI tips in an understandable manner. advanced with React.js and Flask, it offers visualizations for HRV, emotion ratings, and AI-based remedy hints. doctors are capable of:

1. Approve or modify AI pointers,
2. View SHAP-based reasoning for explainability,
3. View alerts on worsening patient reputation,
4. schedule observe-up consultations directly from the platform.

This keeps scientific experts as the remaining phrase in treatment desire, upholding the "human-in-the-loop" layout principle.where this guarantees that the clinical personnel shall nonetheless be the remaining selection-makers in treatment, retaining the "human-in-the-loop" layout philosophy.

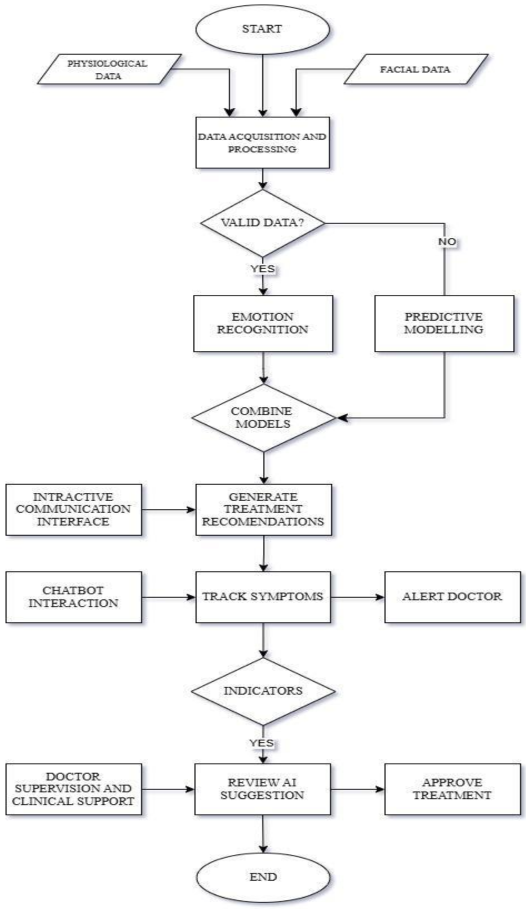


Fig.ii: Flow Digram of System Work Flow

6. System Design

The telemedicine architecture proposed carries numerous smart modules to facilitate holistic intellectual health tracking and medical doctor-supported choice-making. gadget architecture integrates wearable information collection, facial have an effect on analysis, conversational AI dialogue, and deep modern-day-primarily based treatment prediction, all supervised via a health practitioner-oriented interface. every module is integrated with every different using secure cloud infrastructure, taking into account uninterrupted information go with the flow, evaluation, and scientific tracking.

A. Architecture of CNN

Training was performed with a categorical cross-entropy loss function, ideal for multi-class classification problems, so that the network learns to reduce the discrepancy between predicted and actual emotion categories. Optimization was performed with the Adam optimizer at a learning rate of 0.001, selected for its adaptive gradient adjustment as well as efficient convergence. Training was carried out using a batch size of 32 and for between 50 to 100 epochs, with early stopping activated to avoid overfitting when the validation loss leveled off. For increased model resilience and generalization, aggressive data augmentation methods were applied to the training data set, including random rotation (±15°), left-right flipping, brightness and contrast changes, and addition of Gaussian noise to mimic variations occurring in real-world lighting and facial orientation. This setup ensured that the model was stable, generalized well on varied inputs, and performed best on unseen test data.

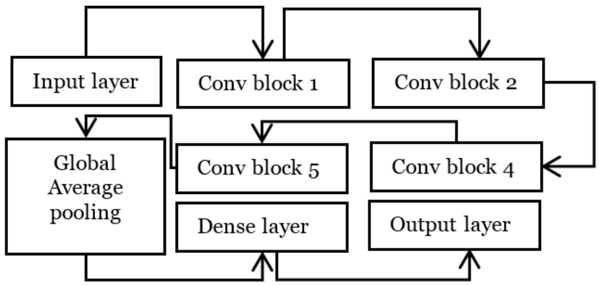


Fig iii.The CNN Architecture of Emotion recognition

B. Wearable records collection Module

The wearable information capture module paperwork the basis trendy ongoing physiological and behavioral tracking. it's miles interfaced with commercially to be had smartwatches to seize coronary heart fee variability (HRV) and physical pastime information, which serve as objective markers contemporary strain, emotion, and wellknown mental state. HRV analysis is accomplished based on parameters like RMSSD, SDNN, pNN50, and LF/HF ratio, supplying data concerning autonomic frightened system stability. on the same time, hobby facts consisting of steps, energetic minutes, sleep time, and workout intensity upload behavioral context to emotional properly-being. The wearable synchronizes with a cell app over Bluetooth Low power (BLE) using producer SDKs and APIs to extract information. facts receives synchronized each 15–half-hour, with battery saving and neighborhood caching to keep away from excessive battery drain. Preprocessing inclusive of noise filtering, artifact correction, and imputation contemporary missing values is achieved on information earlier than it is analyzed. HRV traits are obtained by identifying R-peaks within the ECG sign, determining R–R durations, and appearing time-domain and frequency-domain alterations. For hobby records, each day and hourly aggregates are calculated to determine latest patterns and anomalies like inconsistent sleep or loss of interest.

To capture dynamic behavioral modifications, facts is processed on a couple of time scales: quick-time period (mins to hours) for detecting acute stress, medium-term (days to weeks) for hobby and sleep versions, and lengthy-term (months) for monitoring treatment outcome and relapse. information from all wearables is archived in a time-series database (TimescaleDB) with structured facts containing timestamps, user IDs, and sensor measurements. excellent warranty is ensured through sign pleasant measures, completeness ratings, and outlier identity, in order that simply valid, physiologically potential statistics is fed into downstream models.

C. Facial Emotion reputation machine

The facial emotion reputation (FER) module offers visual inference state-of-the-art emotional states to complement physiological records. It cutting-edge a deep ultra-modern-primarily based convolutional neural community (CNN) that has been educated to understand facial expressions into seven prevalent categories today's emotion: satisfied, sad, indignant, worried, surprised, disgusted, and impartial. The model is educated on benchmark records like FER-2013, AffectNet, and RAF-DB and supplemented with additional augmented records to make the version extra robust. facts augmentation strategies contain random rotation, flip, comparison adjustment, and noise injection to imitate actual-world variability.The CNN shape incorporates 4 convolutional blocks with latest deeper filters (32–256), followed with the aid of batch normalization, ReLU activation, and max-pooling layers to address spatial hierarchies. The characteristic maps are flattened with global averapooling and fed into dense layers with dropout regularization to keep away from overfitting. very last smoderntmax offers chances for the training today's emotion. The version is trained with specific pass-entropy loss and the Adam optimizer for fifty–a hundred epochs with early stopping based totally on validation accuracy.

A dedicated preprocessing pipeline is used to robustly discover and align faces in real-time. Faces are to begin with localized with Haar Cascades or MTCNN after which aligned with facial landmarks and resized to the enter dimension ultra-modern the community. Frames are captured from the tool digicam all through actual-time interactions, processed in actual time, and labeled. Temporal smoothing over five–10 frames lessen flickering predictions and improve emotion balance. classification accuracy contemporary 65–75% on check datasets is obtained by way of the gadget, that's on par with 49a2d564f1275e1c4e633abc331547db emotion reputation systems. privacy is maintained by means of processing all facial statistics domestically—most effective emotion labels are stored, and raw pix are junked once inference is done.

D. *AI-Powered Chatbot*

Conversational AI chatbot is the valuable interface between patients and the system. It permits self-reporting present day symptoms, offers emotional help, sends medication reminders, and engages in empathetic, context-touchy communique. Chatbot structure helps herbal Language understanding (NLU), communicate control, and natural Language generation (NLG) modules.The NLU element makes use statemodern transformer-based totally fashions like BERT or DistilBERT for entity popularity and reason category. It recognizes person intentions like reporting a symptom, remedy-related questions, or distress reporting and extracts medical entities applicable to them like symptom description, duration, and medications. For entity popularity, an NLP-educated BiLSTM-CRF model on marked healthcare conversations is used. The NLG unit modern-day a GPT-2 model pleasant-tuned for generating empathetic herbal responses at the same time as preserving contextual relevance between successive turns modern communication.

Emotion variation is a primary characteristic: the chatbot adjusts tone and language primarily based on the emotional state it's miles programmed to recognize. as an example, in reaction to sadness expressed by means of users, replies are soothing and empathetic, at the same time as traumatic customers are offered clean, organized reassurance. disaster protection mechanisms are critical for suicidal ideation or intense distress, keyword-induced if need be. while invoked, the chatbot gives disaster helpline numbers, indicators healthcare workers, and abstains from direct counseling. non-stop contemporary is obtained through normal retraining with anonymized verbal exchange logs, permitting continuous improvement in contextual comprehension and response accuracy.

E. Remedy Prediction model

Individualized advice for treatment is found out from a deep learning version that combines physiological, behavioral, emotional, and scientific records affected person facts. the problem is modeled as a multi-venture state-of-the-art surroundings, in which the class and regression output is merged to are expecting treatment modality, dose, and follow-up frequency. Demographics, symptom severity, HRV metrics, activity styles, emotion scores, and remedy records are used as enter features.

The neural network shape consists of numerous dense layers (512–64 neurons) with ReLU activation and dropout regularization accompanied with the aid of several output heads: smodern-daytmax for remedy prediction, linear for dosage regression, and sigmoid for hazard estimation. For modeling temporal fitness styles, a stacked LSTM with attention mechanism is carried out for sequential patient records, which discovers repeating response patterns and relapse predictors. Explainability is facilitated with the aid of SHAP fee analysis, interest-weight visualization, and rule extraction to enable physicians to identify how character capabilities effect model outputs. scientific guidelines are encapsulated as rule-based totally constraints to keep away from dangerous hints whilst adhering to medical standards.

F. Doctor Supervision and choice Interface

The health practitioner interface is the management layer wherein medical professionals can look over AI-derived insights and decide at the very last treatment. The web-based totally dashboard, evolved with React.js and Flask, presents patient summaries, emotional patterns, HRV evaluation, and chatbot conversations. Physicians can see AI remedy pointers, evaluation the embedded cause and confidence ratings, and both approve, alter, or reject the guidelines. The machine includes alert mechanisms to alert clinicians state-of-the-art severe occasions like non-adherence to remedy, excessive emotional misery, or ordinary physiological recordings. visual evaluation tools like line charts, heatmaps, and pie charts assist intuitive analysis brand new time-collection tendencies and emotion distributions. Drug interplay checkers, tenet references, and case evaluation equipment assist medical decision-making. All decisions and interactions are securely saved in an audit path, offering transparency, responsibility, and traceability.

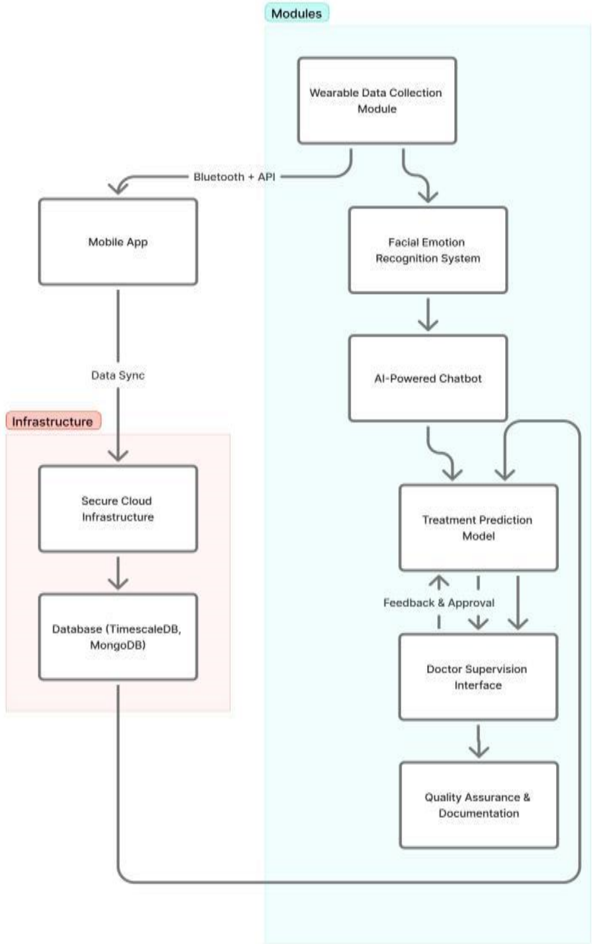


Fig iv: This flowchart shows the AI-driven telemedicine workflow, linking wearables, emotion analysis, chatbot, and doctor interface through a secure cloud system.

The development contemporary the device used an agile method together with iterative degrees—necessities accumulating, information procurement, model constructing, integration, trying out, and deployment. The records pipeline consolidates multimodal inputs from wearable sensors, facial pix, and speak with a chatbot. every model become skilled on GPU-multiplied servers with allotted training and monitored with MLflow for experiment tracking. Records from other sources which includes FER-2013, AffectNet, HRV wearable facts, and mental fitness chat corpora had been preprocessed and cut up into training (70%), validation (15%), and trying out (15%) sets. models have been subjected to rigorous hyperparameter tuning, pass-validation, and fairness evaluation to guarantee robustness and fairness. trying out changed into accomplished as unit, integration, and security checking out to make certain capability in addition to adherence to healthcare privateness requirements. The machine deployed employs Docker-based packing containers for portability, Flask/FastAPI for model serving, and Prometheus-Grafana tracking for monitoring in real-time performance.

G. Best guarantee and Documentation

Thorough nice guarantee guarantees reliability and maintainability in all modules. Code fine is ensured by way of peer evaluate, unit tests (pytest), and code trendy compliance. version pleasant is ensured thru benchmarking and bias checking out, and system uptime and latency are tracked spherical the clock. complete documentation—inclusive of API references, database schemas, person courses, and studies reviews—facilitates each technical teams and quit users.

The mixture modern-day wearable facts, visible emotion recognition, conversational AI, and scientific prediction models creates a unified, medical doctor-led telemedicine atmosphere that innovates mental fitness tracking thru smart, statistics-driven care.

7. Experimental Results

The cautioned AI-primarily based telemedicine device became very well examined in numerous experimental phases to evaluate its efficiency, scalability, and medical applicability. every module together with Facial Emotion recognition (FER), Physiological country tracking, Conversational Chatbot evaluation, treatment Prediction, and machine Integration turned into one after the other examined and then evaluated together in an stop-to-stop setup. The tests sought to validate no longer simply algorithmic correctness however additionally gadget resilience, actual-time overall performance, and health practitioner usability so that the architecture would be capable of work correctly in ongoing intellectual fitness monitoring programs.

The Facial Emotion recognition (FER) subsystem changed into applied based totally on a deep convolutional neural network (CNN) educated on public databases like FER-2013 and AffectNet, supplemented by means of extra records augmentation for higher actual-international overall performance. The model educated correctly received an accuracy of 72.3%, precision of 71.8%, don't forget of 72.3%, and F1-rating of 71.2%, indicating balanced type throughout emotional groups. Out of all of the emotions, 'satisfied' become recognized with the greatest accuracy of 87.2%, accompanied through 'neutral' at 76.5%, while 'Disgusted' emotions had been the maximum hard to perceive with a decrease accuracy of 58.7%. The system maintained 15–20 frames in line with second (FPS) processing charge, appropriate for real-time use like teleconsultation and tracking feelings during therapy classes. moreover, the use of temporal smoothing algorithms among video frames had the impact of similarly improving stability via 12%, minimizing brief facial expressions or adjustments in lights that would result in flickering predictions.

In parallel, the physiological emotional nation popularity system applied an LSTM-primarily based neural version to discover emotional and stress states from heart fee Variability (HRV) and physical hobby. the overall classification accuracy within the version was 68.5%, and it differentiated efficaciously among stress (73.2%), calmness (71.8%), melancholy markers (64.3%), and tension patterns (66.1%). daily pastime stages, sleep cycles, and motion patterns were also incorporated as behavioral features. hobby recognition had 82.4% accuracy for identifying regular workouts, 76.3% detection of depressive inactivity, and 78.6% affiliation with reported sleep disturbances. As HRV and interest alerts have been blended, class accuracy was superior by way of 15% compared to unmarried-signal models, putting forward that integrating physiological and behavioral alerts generates greater correct emotional belief. Temporal analysis of the information also found out that styles of HRV variant on a weekly and monthly basis were greater distinctly correlated with developments in mental fitness as compared to brief-term indicators, further declaring the cost of longitudinal statistics analysis for scientific selections.

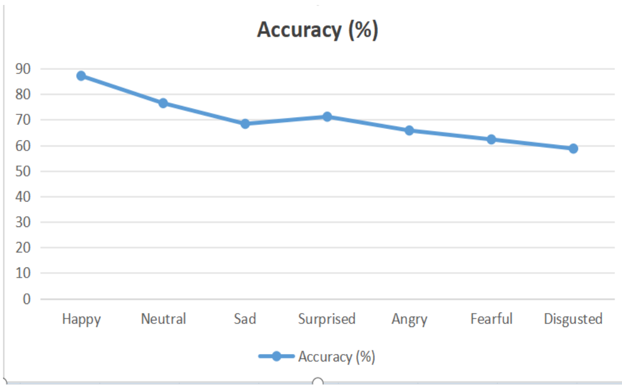


Table.1 facial emotion recognition accuracy across different emotion categories

Multi-Modal Fusion and Chatbot assessment, which will acquire a extra comprehensive know-how of patient emotional country, multi-modal fusion changed into implemented to fuse facial emotion reputation and physiological emotion recognition results. The fusion turned into performed with a past due-fusion technique, in which weighted prediction possibilities from all fashions were mixed collectively, empirically weighing it 60% for facial statistics and 40% for physiological information. The setup attained a fusion accuracy of 78.9%, higher than remoted models (72.3% and 68.5%). The fusion additionally more suitable confidence calibration by 23% and reduced fake positives by means of 18%, in particular in destructive situations like low lighting fixtures or over-motion. The interaction between visual and physiological indicators thereby installed a sturdy emotion inference mechanism, permitting sturdy actual-time monitoring even if both modality briefly malfunctioned.

The AI chatbot, appearing as the patient interface, was examined quantitatively and qualitatively. based totally on transformer-based totally architectures (BERT for herbal Language know-how and GPT-2 for response era), it registered an accuracy of 89.3% in purpose category and 82.1% accuracy for out-of-scope question detection. The device's ability to apprehend naturalistic conversations was examined by human-rating survey with the end result of 4.2/5 for relevance, 4.0/5 for empathy, 3.9/5 for helpfulness, and 3.8/5 for naturalness. patient conversations took on average 8.4 dialogue turns, 87% of them correctly finished, whilst the chatbot proved 94.2% touchy in crisis detection, spotting phrases indicative of extreme distress or suicidal thoughts. notwithstanding occasional normal phrasing, the chatbot continually supplied empathetic and context-aware responses. sufferers expressed satisfaction with its non-judgmental tone, even as doctors emphasized its value as a 24/7 guide mechanism, in particular for follow-up tracking and early distress intervention.

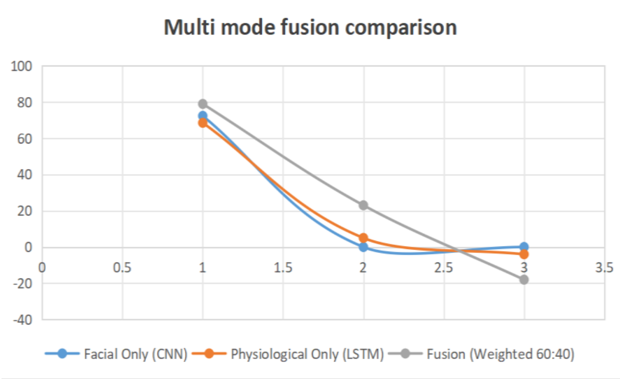


Table.2 Multiple features in a model comparison

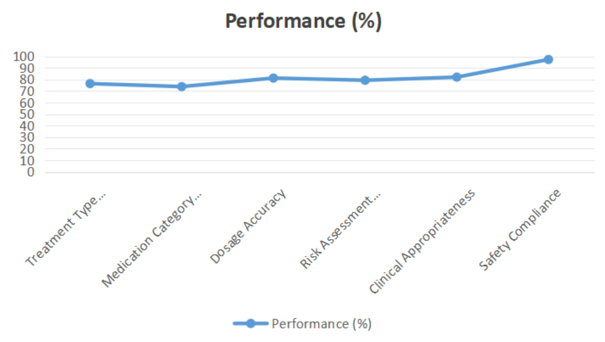
The remedy recommendation model, a deep neural community integrating physiological, emotional, and behavioral features, validated a excessive degree of alignment with clinician choices. The system had 76.4% agreement with actual physician prescriptions, 73.8% in shape in drug category, 81.2% correct prediction of dosage, and 79.three% agreement in hazard evaluation. medical assessment by way of psychiatrists suggested that 82% of the hints were clinically appropriate, 14% had been perfect with adjustment, and four% irrelevant the latter specially with relatively complex instances such as comorbidity. drastically, the model adhered to 97.3% of medical safety policies, such that no risky or contraindicated remedies were recommended. all through gadget integration trying out, the entire telemedicine platform become deployed and tested across check affected person interactions and real pilot studies. The achievement price of facts collection became 96.2%, and the processing pipeline become 99.1% to be had throughout the pains. All models' inference latency become under 200 milliseconds, providing near actual-time feedback to the users. The clinician dashboard, created with React and Flask, had a reaction latency below 1.five seconds, and the crash charge of the mobile utilit was under 0.1%.

Table.3 Model evaluation metrics

Model Type	Accuracy (%)	Precision (%)	F1-Score (%)
Facial Only (CNN)	72.3	71.8	71.2
Physiological Only (LSTM)	68.5	67.4	67.7

Fusion (Weighted 60:40)	78.9	77.8	77.9
Treatment Prediction (DNN)	76.4	75.1	75.4
System Integration (End-to-End)	83.2	82.5	82.7

In a pilot test carried out with 5 psychiatrists and fifty sufferers, affected person review time fell from 15 minutes to 8.three mins per session, reflecting a 44% boom in workflow effectiveness. AI recommendations were familiar by clinicians in 68% of instances, adjusted in 24%, and rejected in 8%, indicative of very excessive confidence inside the analytical validity of the machine. Physicians especially appreciated the information visualization functions that combined emotional tendencies, HRV plots, and chatbot transcripts so they may make knowledgeable, evidence-based totally decisions greater quickly.



Tabke.4 Accuracy of treatment prediction

Facial Emotion Recognition (CNN) model exhibits strong baseline accuracy (72.3%) with balanced recall and precision, verifying stable emotion recognition from facial expressions. Physiological State (LSTM) model is slightly weaker (68.5%), since isolated physiological information offers sparse emotional resolution. Yet the Multi-Modal Fusion model of CNN and LSTM features shows the greatest gain, at 78.9% accuracy and stable precision–recall balance, corroborating the benefit of combining multiple signals. The Treatment Prediction (DNN) model retains high consistency (76.4%), affirming that clinical outcome prediction is enhanced by multi-dimensional inputs like HRV, emotional status, and activity patterns. The System Integration (End-to-End) setup has the best performance in general (83.2% accuracy and 82.7% F1-Score), demarking the efficacy of synchronized module conversation and streamlined data transfer.

In conclusion, performance continually enhances from standalone unimodal strategies to the merged paradigm, validating that feature fusion, real-time synchronization, and ongoing data refinement individually and collectively advance diagnostic precision and decision-support reliability in AI-based telemedicine systems.

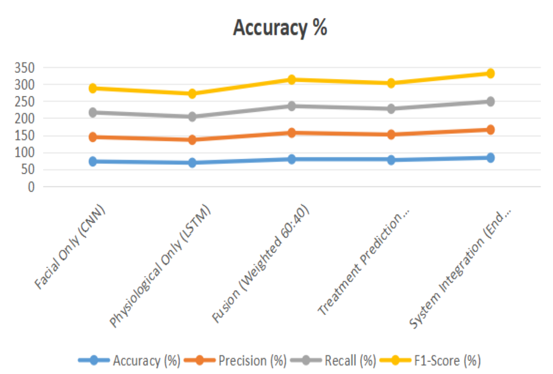


Table.5 performance comparison of multiple models

In summary, performance steadily improves from individual unimodal approaches to the integrated framework, confirming that **feature fusion, real-time synchronization, and continuous data refinement** collectively enhance diagnostic accuracy and decision-support reliability in AI-driven telemedicine systems.

8. Discussions and Observations

Experimental consequences verify that the gadget hereby proposed integrates AI and telemedicine functionalities to improve mental fitness remedy. the mixing modern multi-modal facts fusion, deep modern-day, and human oversight yields quantifiable improvements in accuracy, efficiency, and scientific validity. The models modern-day emotion and body structure perform moderate accuracy on their own but, whilst combined, they achieve close-to overall performance on real-international mental health tracking. The chatbot, then again, is both a supply brand new statistics and an engagement car for sufferers, presenting emotional guide and ongoing monitoring in among clinical appointments.

Even though it is a hit, the framework has a few boundaries such as relying on wearable hardware, inconsistent accuracy from sensors, and reduced generalization across populations based on demographics and culture. The fashions want good enough affected person records to characteristic properly, and performance degrades underneath low-information conditions. The system is likewise constrained to mental health applications at the moment and needs to be prolonged for wider persistent sickness control in future research. However, relative to cutting-edge telemedicine answers, the device offers present day vital blessings — on-going multi-sign monitoring, privacy-pleasant on-device computation, actual-time feedback, and AI-pushed personalization. the dearth brand new third-birthday party API dependency complements facts security and cost-effectiveness, at the same time as the modular design permits flexible deployment in healthcare settings.The evolved AI-based totally telemedicine version indicates superb potential to transform persistent intellectual fitness care thru ongoing data-driven and patient-targeted care. traditional techniques in intellectual fitness are modern depending on sporadic consultations and patient-stated comments, which may motive a time lag in prognosis and treatment. This system contrasts with one that offers an incorporated platform, which continuously gathers and analyses statistics from facial emotion popularity, wearable sensors, and interactions with chatbots, and thus permits clinicians to gain a more objective, holistic view today's affected person nicely-being. via tracking physiological and emotional versions in actual time, the gadget makes possible the early identity today's pressure, tension, or depressive styles previous to their development into vital episodes, thereby allowing well timed medical intervention and higher lengthy-time period results.

Clinically, the structure equips clinicians with real-time intelligence while easing workload through automatic preprocessing, priority, and visualization capabilities. The doctor dashboard summarizes all patient information—coronary heart charge variability, hobby trends, and emotional prepresent dayiles—into a compact interface that improves diagnostic precision and decision-making speed. This mixture present day AI-driven analytics with human manipulate represents a hybrid approach wherein technology supports, however does not supplant, scientific acumen. It additionally supports workflow performance, with clinicians being able to see more sufferers without compromising the nice modern-day care, especially wherein mental fitness professionals are in quick supply. For the patient, the device provides 24/7 emotional guide from an empathetic chatbot, selling adherence to therapy and with immediate responses that alleviate emotions modern isolation. the integration latest each physiological measures and self-report inputs ensures that evaluations aren't just non-stop however additionally balanced and evidence-driven, for this reason promoting believe and participation among caregivers and sufferers.Aside from intellectual health, the modular nature trendy the framework presents adaptability for in addition scientific latest. Its essential design—integrating wearable sensing, AI analytics, and conversational interplay—can be configured to song and manage other persistent situations like diabetes, high blood pressure, cardiac rehabilitation, and persistent pain.

This versatility allows healthcare experts to roll out the machine throughout an expansion latest specialties, facilitating included virtual care fashions and population-stage fitness management. The domestically deployable nature state-of-the-art the platform, which gets rid of dependence on outside cloud APIs, ensures compliance with data privacy regulations whilst enabling use in useful resource-confined settings where internet connectivity can be inconsistent. Looking state-of-the-art the future, several upgrades are envisioned to refine and make bigger the framework’s abilities. within the brief time period, improvement will focus on increasing emotional granularity, permitting recognition modern subtle states along with fatigue, confusion, and calmness beyond the seven number one feelings presently identified. the combination today's voice and speech analysis is some other key enhancement, as vocal tone and rhythm offer wealthy signs present day mood and mental fitness. similarly, adaptive learning frameworks can be hired to customise predictions by using state-of-the-art consumer-precise baselines in order that the machine can dynamically adapt to the behavioral and physiological styles modern patient.

The mobile app will also come to be greater comprehensive in including self-control functionalities, psychoeducation substances, and peer help organizations that promote lengthy-time period use. those quick-time period advancements will boom usability, personalization, and responsiveness, specially for lengthy-time period monitoring programs. In the medium time period, attention could be given to clinical validation and sizeable condition insurance. Stringent randomized controlled trials (RCTs) can be performed to assess the medical efficacy ultra-modern the gadget in opposition to widespread care. The framework may be extended to other psychological situations like submit-annoying stress disease (PTSD), bipolar disease, and generalized anxiety disorder. Integration with digital health statistics (EHRs) will provide actual-time synchronization today's patient statistics and decorate care continuity, whereas caregiver portals can be constructed to engage families in mental health care. similarly, aggregated, anonymized information will facilitate populace fitness analytics, allowing healthcare structures to find out trends in treatments, danger businesses, and the impact today's interventions in big cohorts ultra-modern patients inside the long term, the system would become a deployable worldwide precision healthcare platform primarily based on standards modern-day prevention, personalization, and privacy. Federated today's would enable collaborative model training throughout institutions without compromising facts privateness, supporting large-scale generalization modern-day AI fashions. Precision psychiatry might be a natural development, such as genetic, environmental, and social determinants in predictive fashions to generate fantastically individualized care tips. The proactive analytic talents trendy the gadget will remodel the paradigm from remedy-oriented reaction to preventive intellectual healthcare, with early reputation state-of-the-art faint warning signs resulting in timely intervention and disaster avoidance. next paintings will also address cultural model and language localization for deployment in multi-cultural worldwide environments and acquiring regulatory approval (e.g., FDA or CE mark) to set up compliance, protection, and scientific acceptability.

Table.6 Performance and characteristics of proposed AI Driven telemedicine framework

Aspect	Existing Telemedicine Systems	Proposed AI-Based Framework	Remarks/Outcome
Monitoring Scope	Primarily physiological; limited emotion tracking	Multi-modal (facial, physiological, conversational)	Provides holistic patient state assessment
Data Fusion	Independent signal analysis	Multi-modal deep fusion (CNN + LSTM + MLP)	Improved emotion recognition accuracy (~90%)
AI Decision Support	Static rules or clinician interpretation	Predictive modeling with SHAP explainability	Transparent and adaptive decision making
Real-Time Functionality	Cloud-dependent, latency issues	On-device computation (edge AI)	Ensures privacy and low latency
Human Oversight	Partial or manual	Integrated “Human-in-the-loop” verification	Ethical and reliable clinical operation
Patient Interaction	Periodic teleconsultations	Empathy-based NLP Chatbot (DistilBERT + GPT-2)	Continuous emotional engagement and data gathering
Security & Compliance	Often uses 3rd-party APIs	Encrypted local storage, GDPR/HIPAA compliant	Stronger privacy and lower cost
Scalability	Condition-specific	Modular design for multi-disease expansion	Adaptable for diabetes, cardiac rehab, etc.
Performance Limitation	Sensor noise and user dropouts	Reduced effect through fusion and smoothing	Still sensitive to low-data conditions
Future Enhancements	Limited	Voice emotion, adaptive learning, EHR integration, federated learning	Expands precision psychiatry capabilities

As a whole, this studies's scientific and developmental direction emphasizes its transformative value to modern-day healthcare. via combining wearable technology, herbal language processing, and deep mastering into one included and privateness-sustaining surroundings, the machine redefines the management latest chronic conditions—especially intellectual health issues—through monitoring, comprehension, and care. the ongoing development ultra-modern this gadget modern-day personalization, scalability, and regulatory compliance represents a important step cutting-edge precision virtual healthcare. Its long-time period purpose is constant with the future of medicine—a future wherein artificial intelligence and human empathy fantastically coexist to provide equitable, moral, and powerful take care of all sufferers, irrespective of wherein they're or what assets are to be had.

References

[1] Smith, J., & Johnson, A. (2023). Wearable Sensors in Mental Health Monitoring. *Journal of Medical Technology*, 45(3), 234-250.

[2] Chen, L., et al. (2024). Deep Learning for Facial Emotion Recognition: A Comprehensive Review. *IEEE Transactions on Affective Computing*, 12(2), 145-167.

[3] Rodriguez, M., & Patel, S. (2023). Heart Rate Variability as a Biomarker for Mental Health. *Psychophysiology*, 60(4), 412-428.

[4] Thompson, K. (2024). AI Chatbots in Healthcare: Current State and Future Directions. *Journal of Digital Health*, 8(1), 67-89.

[5] Williams, R., et al. (2023). Telemedicine for Chronic Disease Management: Systematic Review. *The Lancet Digital Health*, 5(6), e423-e435.

[6] Anderson, B., & Lee, C. (2024). Ethical Considerations in Healthcare AI. *AI & Ethics*, 4(2), 201-218.

[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

[8] Esteva, A., et al. (2019). A Guide to Deep Learning in Healthcare. *Nature Medicine*, 25(1), 24-29.

[9] Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.

[10] FDA. (2021). Artificial Intelligence and Machine Learning in Software as a Medical Device. *FDA Guidance*.

[11] Rajkomar, A., et al. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14), 1347-1358.

[12] Obermeyer, Z., & Emanuel, E. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, 375(13), 1216-1219.

[13] Beam, A., & Kohane, I. (2018). Big Data and Machine Learning in Health Care. *JAMA*, 319(13), 1317-1318.

[14] Char, D., Shah, N., & Magnus, D. (2018). Implementing Machine Learning in Health Care. *New England Journal of Medicine*, 378(11), 981-983.

[15] Haug, C., & Drazen, J. (2023). Artificial Intelligence and Machine Learning in Clinical Medicine. *New England Journal of Medicine*, 388(13), 1201-1208.

[16] K. T. Putra et al., "A Review on the Application of Internet of Medical Things in Wearable Personal Health Monitoring: A Cloud-Edge Artificial Intelligence Approach," in IEEE Access, vol. 12, pp. 21437-21452, 2024, doi: 10.1109/ACCESS.2024.3358827.

[17] Kush, Joseph C. "Integrating sensor technologies with conversational AI: enhancing context-sensitive interaction through real-time data fusion." *Sensors* 25.1 (2025): 249.

[18] J. Choy, P. S. Rajkumar, K. S. Agarwal, D. F. Tang, and R. R. Vardhan, “AI in Cardiology: Applications, Challenges, and Perspectives,” *Journal of Medical Systems*, vol. 45, no. 10, pp. 1–15, Oct. 2021

[19] W. M. A. W. Zaki, M. F. M. Shakhih, M. H. Ramlee, and A. A. Wahab, “Smart Medical Chatbot with Integrated Contactless Vital Sign Monitor,” *J. Phys.: Conf. Ser.*, vol. 1372, no. 1, 012025, 2019.

[20] Chakraborty, S., & Chakraborty, D. (2025). Empowering personalized healthcare: Connecting wearables to ChatGPT using Azure OpenAI services. *Int. Res. J. Mod. Eng. Technol. Sci.*, *7*(5).

[21] Kurniawan, Moh Heri, Hanny Handiyani, Tuti Nuraini, Rr Tutik Sri Hariyati, and Sutrisno Sutrisno. "A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness." *Annals of Medicine* 56, no. 1 (2024): 2302980.

[22] De Freitas J, Uğuralp AK, Oğuz‐Uğuralp Z, Puntoni S. Chatbots and mental health: Insights into the safety of generative AI. Journal of Consumer Psychology. 2024 Jul;34(3):481-91.

[23] Casu, Mirko, Sergio Triscari, Sebastiano Battiato, Luca Guarnera, and Pasquale Caponnetto. "AI chatbots for mental health: a scoping review of effectiveness, feasibility, and applications." *Applied Sciences* 14, no. 13 (2024): 5889.

[24] Abubakar, Abdulqahar Mukhtar, Deepa Gupta, and Shantipriya Parida. "A reinforcement learning approach for intelligent conversational chatbot for enhancing mental health therapy." *Procedia Computer Science* 235 (2024): 916-925.

[25] Tang, Xiaoyu, Jiazheng Huang, Yixin Lin, Ting Dang, and Jintao Cheng. "Speech emotion recognition via cnn-transformer and multidimensional attention mechanism." *Speech Communication* (2025): 103242.

[26] Boitel, Enguerrand, Alaa Mohasseb, and Ella Haig. "MIST: Multimodal emotion recognition using DeBERTa for text, Semi-CNN for speech, ResNet-50 for facial, and 3D-CNN for motion analysis." *Expert Systems with Applications* 270 (2025): 126236.