IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

MediDocker: An AI-Enabled Healthcare Intelligence Ecosystem Integrating Clinical Decision Support, Predictive Analytics, and Emergency Response

Sanika Ghume, Sneha Kanhurkar, Vedika Kulkarni, Shehnaz Siddique
Usha Mittal Institute of Technology, Santacruz, Mumbai, India
vedikak9299@gmail.com, snehakanhurkar12@gmail.com
sanika.mg22@gmail.com, work.shehnazsiddique@gmail.com

Abstract: *Patients keep their health information in places that do not talk to each other: handwritten prescriptions, PDF blood reports, appointment cards and paper bills, and in an emergency nobody can quickly see their blood group or allergies. This paper presents MediDocker, a web platform that keeps these records in one secure place and explains them in plain language. A vision-language model (Qwen2.5-VL, run locally) only copies what a prescription says; every medicine is then verified against RxNorm, an Indian brand dictionary and a hand-checked drug table, and its purpose and warnings come from verified sources, never from the model. Blood (CBC) reports are encrypted, read with a PDF text layer or PaddleOCR, checked by seven plausibility rules, and followed across reports with reference change values, a personal set point and Theil–Sen projection. A LightGBM model trained on 14,134 laboratory-confirmed NHANES participants screens for iron deficiency from the CBC alone and answers only when it is confident, otherwise advising a serum ferritin test. On 1,945 unseen participants from 2021 onwards it reached an AUROC of 0.836 (95% CI 0.816–0.857) and 94.2% accuracy on the 10.6% of cases it decides, with sensitivity 99.6% and specificity 99.3% at its two cut-offs. A record-aware assistant detects emergencies with rules before any language model is used, and an Ed25519-signed emergency QR code can be verified offline by anyone, without login.*

Keywords: personal health record, prescription verification, CBC trend analysis, iron-deficiency screening, LightGBM, NHANES, signed QR code, health assistant

I. INTRODUCTION

A visit to a doctor usually ends with two pieces of paper. One is a prescription written by hand, often in a hurry. The other, a few days later, is a laboratory report full of abbreviations and numbers that mean little to someone without medical training. Over the years these papers pile up in drawers and phone galleries, next to appointment slips and pharmacy bills. When the history is needed, for a new doctor or in an emergency, it is rarely complete and almost never in one place.

MediDocker is our attempt to fix this for a single patient. It is built around one strict rule: the AI may read and may phrase, but it may not decide a medical fact. Text is copied from images by a vision model and friendly wording is produced by a language model, while every medicine detail, every normal range and every risk score comes from a verified database, a fixed rule, or a model we trained and tested ourselves. Wherever a reading looks doubtful, the patient is asked to confirm it instead of the system guessing.

The paper makes five contributions. First, a prescription pipeline in which the vision model only transcribes and every medicine is verified before it is explained. Second, a CBC pipeline that validates each extracted value with seven

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

plausibility checks and follows every analyte across reports using reference change values. Third, an iron-deficiency screening model trained on laboratory-confirmed CDC NHANES data, calibrated, evaluated on later survey years, and designed to abstain when it is not confident. Fourth, a health assistant that detects emergencies with rules before any language model is consulted and then uses the patient's own records to suggest a specialist. Fifth, an emergency card signed with Ed25519 whose scan page verifies authenticity offline and offers one-tap help to a bystander.

Section II reviews related work, Section III presents the architecture, Section IV explains each component, Section V reports the evaluation, and Section VI discusses limitations before Section VII concludes.

II. RELATED WORK

Records, CBC prediction and chatbots. Musa et al. [1] built a cloud-based patient record system with rapid application development. Its focus is storing and organizing records, so AI-based document analysis and long-term tracking fall outside its scope. SympCare [2] predicts diseases from CBC data with machine learning behind a web interface; its dataset and clinical validation can be broadened further. Hussain [3] recommends specialists from symptoms using NLP and machine learning as a standalone system that does not draw on the patient's medical records. Secco and Nazemi [4] place long medical histories on a timeline in a modular dashboard designed mainly for clinicians, rather than for patient ownership of the data.

TABLE I: SUMMARY OF RELATED WORK

Ref.	Focus	Difference in scope from this work
[1]	Centralized web-based patient record system	AI document analysis and long-term tracking not covered
[2]	SympCare: disease prediction from CBC data	Dataset diversity and clinical validation can be extended
[3]	HealthBot: specialist recommendation from symptoms	Standalone; does not use medical records
[4]	Timeline dashboards for patient history	Clinician-oriented rather than patient-owned
[5]	Transformers for medication mentions in clinical notes	Typed notes; handwritten prescriptions not covered
[6]	Image-based QR codes with embedded visuals	General purpose; no medical use or security focus
[7]	Fuzzy logic and Random Forest for blood disease classes	Trend analysis not included
[8]	OCR and ML on CBC reports (Random Forest, about 91% F1)	One region and 12 features
[9]	QR-based emergency access to patient history	Security and AI analytics not the main focus
[10]	PaddleOCR and NLP for prescriptions with database validation	Not linked to records or emergency access

Medication notes, QR codes and blood reports. Guo et al. [5] fine-tuned transformer models to find medication mentions and medication changes in typed clinical notes, so handwritten prescriptions are outside their scope. Chen and Ti [6] show how images can be embedded in a QR code, a general-purpose design that is not aimed at medical use or data security. Ameen et al. [7] pair fuzzy logic with a Random Forest to classify hematological diseases from CBC values; trend analysis over time is not part of their approach. Bibi and Saeed [8] read CBC reports by OCR and apply machine learning; a Random Forest reached an F1-score of about 91% on data from one region and 12 features, which leaves room to extend the work to wider data.

Emergency access and prescriptions. Rahul et al. [9] use QR codes to give access to a patient's history in an emergency, with security and AI analytics outside their main focus. A system for doctors' prescriptions [10] digitizes them with PaddleOCR and NLP and validates them against a database, but it is not connected to health records or to emergency access. Table I summarizes these works. Taken together, these approaches address records, CBC prediction, chatbots, prescription OCR or emergency QR codes, each largely on its own, usually from a single report, and with

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

limited attention to securing emergency data. Combining them in one secure, patient-centric platform is the aim of this paper.

III. SYSTEM ARCHITECTURE

MediDocker consists of three programs that talk over HTTP, as drawn in Fig. 1, and a separate emergency page. The presentation layer is a single-page React application built with Vite. Its screens cover the dashboard with notifications, prescriptions, reports and the health summary, the calendar with doctor visits, expense tracking, the Medi health assistant, and the profile with the emergency QR code. The browser never talks to the database or to the AI service directly. Every call goes to the application layer, a Node.js and Express service, which checks a JSON Web Token first. The application layer owns the database and hosts authentication and profile handling, the record APIs for visits, calendar and expenses, the encrypted upload module, the chat engine and the emergency card service that signs profiles. Heavy processing is delegated over HTTP to the AI/ML service, a Python FastAPI application that contains the prescription pipeline, in which Qwen2.5-VL is served locally through Ollama, the CBC pipeline with PaddleOCR, and the blood analytics stage. Model training is done offline and kept apart from the serving code.

The data layer holds structured records in MongoDB Atlas (users, prescriptions, reports, analyses, expenses, doctor visits and calendar events) and the encrypted originals of reports and prescription images in a file store. Four external services are used for narrow jobs: the NLM RxNorm API for drug names, the openFDA API for label text, the Google Gemini API for the assistant's wording, and OpenStreetMap for nearby hospitals. Table II lists the technology stack.

IV. METHODOLOGY AND IMPLEMENTATION

A. Accounts, Profile and Data Protection

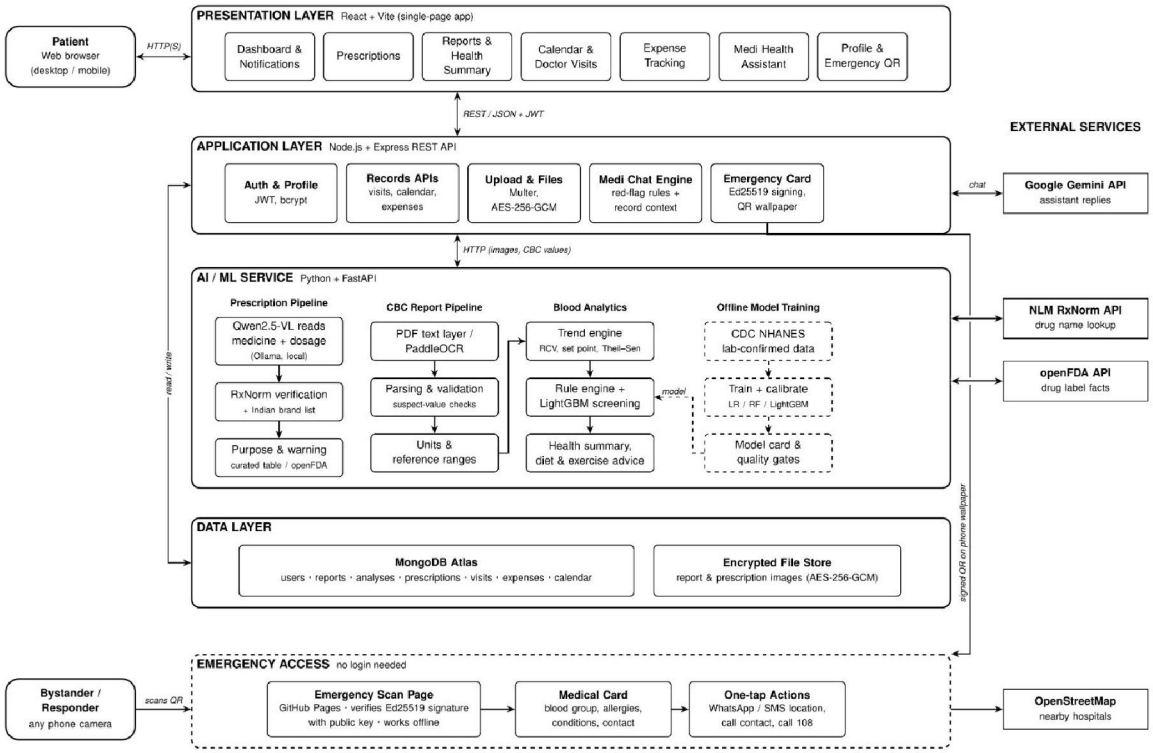


Fig. 1. System architecture of MediDocker. Dashed blocks are offline model training and the login-free emergency access zone.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

Passwords are stored only as bcrypt hashes (cost factor 10), and a successful login returns a JWT signed with HMAC-SHA256 and valid for seven days. An authentication middleware verifies the token on every private route, and every database query is filtered by the user identifier, so one account can never read another's documents. The onboarding profile holds age, gender, height, weight, a calculated BMI, blood group, allergies, conditions, pregnancy status and an emergency contact. These fields select the default reference ranges, give context to the assistant and fill the emergency card.

TABLE II: TECHNOLOGY STACK

Layer	Technology	Role
Frontend	React, Vite	Single-page app, charts, chat widget
Backend	Node.js, Express 5, JWT, bcrypt	REST APIs, authentication, encryption, signing
Database	MongoDB, Mongoose	Flexible documents for reports with different fields
AI service	Python, FastAPI	Prescription and report pipelines, model serving
Local AI	Ollama, Qwen2.5-VL, PaddleOCR	Reading images without leaving the deployment
Machine learning	scikit-learn, LightGBM, SHAP	Training, calibration, explanations
Cloud AI	Google Gemini	Assistant wording only
Emergency page	GitHub Pages, service worker, TweetNaCl	Offline signature verification

Report PDFs and prescription photographs are encrypted with AES-256-GCM before they reach the disk; the plain image is never written. Each file gets its own random key, and that key is wrapped with a master key (envelope encryption), so only the wrapped key is stored in the database and a stolen database alone cannot open any file. Decryption happens in memory only when the owner opens a file or starts an analysis, and keys are removed from every API response. Medical images are read by local models. RxNorm and openFDA receive only a medicine name, and Gemini receives a short summary of age, conditions and abnormal values, never a name, email or file.

B. Records, Calendar, Expenses and Notifications

The dashboard shows the report count, the latest report, the last five reports and trend lines for hemoglobin, white-cell count and platelets, and it is a read-only call that does not touch the AI service. The calendar stores appointments and health events, and a notification bell counts those falling in the next seven days: today is shown in red, tomorrow and later days in amber, with days counted by calendar date so that next month's events are included. The doctor-visit module keeps doctor and hospital names, phone numbers, reason and notes, and the assistant can save a visit there. Expenses are entered with an amount, category, reason and date and are shown in charts. These modules are ordinary create, read, update and delete features. Their value is that the summary and the assistant read from all of them, so the patient sees one picture instead of five screens.

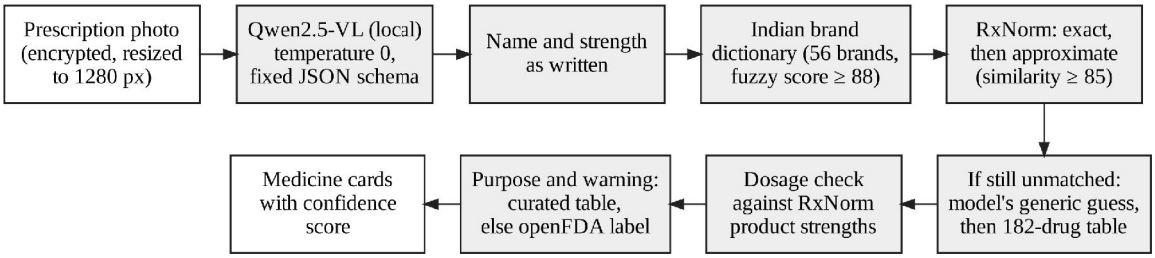


Fig. 2. Prescription simplification workflow.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, November 2026 Impact Factor: 8.2

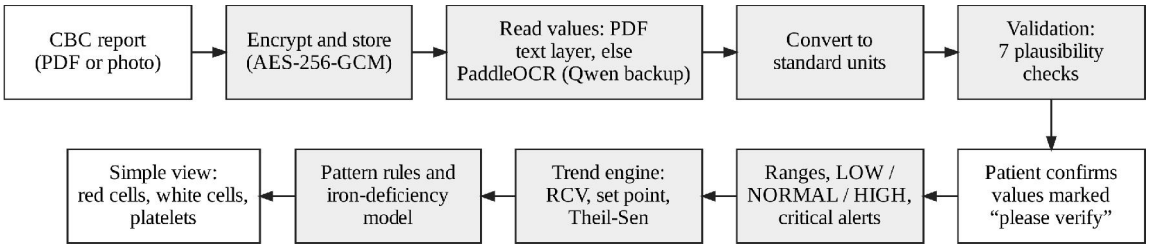


Fig. 3. CBC report workflow, from encrypted upload to the simple view.

C. Prescription Simplification

Fig. 2 shows the flow. The uploaded photograph is kept in memory (up to 20 MB), encrypted and saved, then resized to at most 1280 pixels so that the model runs quickly on ordinary hardware. The AI service sends it to Qwen2.5-VL, run locally through Ollama at temperature 0 with a fixed JSON schema. The model is told to copy each medicine name and strength exactly as written and to give a generic name only if it is sure. It is never allowed to write a purpose or advice.

Each name is then verified in a fixed order. A dictionary of 56 common Indian brands, such as "Dolo 650", is searched first with fuzzy matching (similarity of at least 88) and the resulting generic name is checked in RxNorm. This comes first because Indian brands are mostly absent from RxNorm, a United States database. Next, RxNorm is searched with the written name, exact match first and an approximate match only when the name similarity is at least 85%. The model's own generic guess is trusted least and is also checked in RxNorm. If RxNorm has nothing, a hand-checked table of 182 drugs is searched with a similarity of at least 90. The written strength is compared with the real strengths of that drug in RxNorm, and combination products are added up (625 mg is 500 mg plus 125 mg). Purpose and the main warning come from the drug table or from the openFDA label and are never generated by the model. A medicine that cannot be verified is shown as unverified. RxNorm answers are cached for 30 days, and only one vision job runs at a time so that a busy model makes the next request wait instead of failing.

Every medicine receives a rule-based confidence score. The base B is 90 when the medicine is verified in RxNorm, 75 when it is found only in the drug table and 30 when it is unverified. The score is

$$C = \min(100, \max(0, B - 100 - s - 15g + 10d - 10n))), \quad (1)$$

where s is the name similarity of an approximate RxNorm match (100 for an exact one), g is 1 if the name was only guessed by the model, d is 1 if the written strength matches a real product and n is 1 if no strength was written. For example, "Pantop 40" matches RxNorm at 94% similarity, giving $90 - 3 = 87$, and the strength of 40 mg is a real one, so $C = 97$. The confidence of a whole prescription is the average over its medicines, and the simplified list can be shared as a card through the Web Share API.

D. CBC Report Pipeline

The pipeline in Fig. 3 begins with an encrypted upload. A digital PDF is read directly from its text layer with pypdfium2, which gives the printed values exactly. A scanned PDF or photograph goes through PaddleOCR: text boxes with confidence below 0.5 are dropped, the table columns are located, and each value is read twice, the second time from a crop enlarged three times; if the two readings disagree the value is flagged. If OCR finds fewer than four of the eight core values, Qwen2.5-VL reads the report instead, and if that also fails the system returns what OCR found and asks the patient to check it.

All values are converted to one standard unit, for example g/L to g/dL and lakh/cumm to 109/L, so that reports from different laboratories can be compared. Seven validation checks then look for reading errors. Hemoglobin must lie between 2 and 25 g/dL, the five white-cell percentages must add up to about 100, and MCHC must equal hemoglobin divided by hematocrit times 100 within 12%, among others. Each failed check lowers the confidence of the value, and values below 0.6 are marked "please verify" so that the patient can correct them. The range printed on the report is used

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, November 2026 Impact Factor: 8.2

first; when it is missing, sex-specific defaults apply (for hemoglobin, 13.0–17.0 g/dL for men and 12.0–15.5 g/dL for women) and the screen says so.

Each value is labeled low, normal or high and graded mild, moderate or marked. Critical alerts are raised for hemoglobin below 7, platelets below 20, neutrophils below 0.5 or white cells above 50 (units as printed on the report). Finally, results are grouped into red blood cells, white blood cells and platelets, each shown as good, watch or see a doctor. No AI model decides whether a value is abnormal. Ranges, checks and alerts are fixed rules, so the same report always gives the same result.

E. Trend Engine and Pattern Rules

With two or more reports, every analyte is followed over time by three tools. First, the reference change value (RCV) separates a real change from normal variation of the test and of the person. For two serial results,

$$RCV(\%) = 1.96 \cdot \sqrt{2} \cdot \sqrt{\square CV\ 2 + CV\ 2}, \quad (2)$$

where CVA is the analytical variation of the laboratory and CVI the within-person biological variation, taken from the published EFLM biological variation database. For hemoglobin, CVI = 2.7% and CVA = 1.0% give an RCV of about 8%, so a fall from 13.0 to 12.5 g/dL (3.8%) is treated as noise while a fall from 13.0 to 11.8 g/dL (9.2%) is reported as real. Second, once three or more earlier reports exist, the median of the patient's own values becomes a personal set point. Someone whose hemoglobin is always near 15 is worth checking at 12.8, although 12.8 is inside the population range. Third, the direction of change is found with an ordinary least-squares line, and when there are three or more reports and a steady fall, a Theil–Sen estimate, the median of the slopes between every pair of reports,

$$\beta = \underset{i<j}{\text{median}} \frac{y_j - y_i}{t_j - t_i}, \quad (3)$$

projects when the value may leave the normal range, but only if that happens within 365 days. The median makes the estimate resistant to a single odd result.

Pattern rules then look at several values together, such as an iron-deficiency-type pattern of low hemoglobin with small red cells. A pattern earns 20 points for being met, 8 for each supporting value, points for supporting trends and 20 more if it repeats across three or more reports. The score is capped at 60 with one report and 69 with two, because a single report is weak evidence, and it is labeled strong (70 or more), moderate (45 or more) or weak. The Mentzer index, MCV divided by RBC, is also computed: a value below 13 points to thalassemia trait rather than iron deficiency, so the app lists it as a contradicting sign.

F. Iron-Deficiency Screening Model

A CBC alone cannot show iron stores, but its pattern carries information about them. We therefore trained a model to estimate the probability of iron deficiency from a routine CBC. The data come from eight cycles of the CDC National Health and Nutrition Examination Survey (NHANES), from 2005–2006 to 2021–2023, in which each participant has a full CBC and a serum ferritin result. The label is laboratory-confirmed: iron deficient means ferritin below 15 ng/mL (below 12 for children under five), and when C-reactive protein exceeds 5 mg/L the cut-off is raised to 70 ng/mL, because inflammation falsely raises ferritin. The model has 27 inputs: 17 CBC values (hemoglobin, hematocrit, RBC, MCV, MCH, MCHC, RDW, WBC, platelets, MPV, the five white-cell percentages, and neutrophil and lymphocyte counts), age, sex, pregnancy status and seven calculated indices, including the Mentzer index, an RDW index and the Green–King index. Missing values stay empty and are never filled with healthy values.

The split imitates real use, which is to learn from the past and predict the future. The 14,134 participants from the older cycles form the training data, split 80/20 into a fitting set and a separate calibration set. The test set holds the 1,945 participants from 2021 onwards, 478 of them iron deficient, and was not touched during training. Three algorithms were compared with five-fold cross-validation: logistic regression, a Random Forest of 400 trees and LightGBM. The LightGBM model builds 500 small trees in sequence, each correcting the errors of the previous ones, with a learning rate of 0.03, 15 leaves per tree, 80% of participants and 80% of features per tree, and a fixed random seed of 42. Its raw

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, November 2026 Impact Factor: 8.2

scores were then calibrated with isotonic regression on the calibration set, so that among people given a probability of 70%, about 70% are in fact deficient; the calibration slope on the test set was 1.048 (ideal 1).

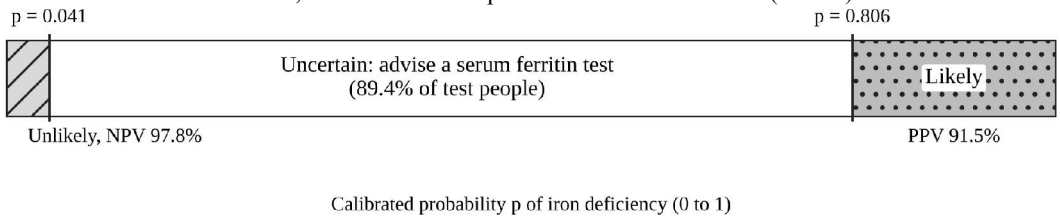


Fig. 4. Decision zones of the iron-deficiency model on the calibrated probability scale.

The model does not always answer. Two cut-offs divide the probability range (Fig. 4). A probability of at least 0.806 is reported as "likely", a value chosen so that at least 90% of such answers are correct. A probability below 0.041 is reported as "unlikely", chosen so that at least 97% of such answers are correct. Anything in between is reported as "uncertain", with advice to take a serum ferritin test. An answer is given only if the report has at least six CBC values and the age and sex of the person fall within the groups on which the model was tested. SHAP values show the three main reasons behind each result, such as a low MCV raising the risk. The same targets that fix the cut-offs act as release gates recorded in a model card: a model that cannot meet them is not shown to users. A vitamin B12 model that we also trained reached an AUROC of only 0.648 and is therefore hidden and marked experimental. The result is presented as a screening suggestion to discuss with a doctor and never as a diagnosis.

G. Medi Health Assistant

The user describes symptoms in a chat. The first step uses no AI. The message is matched against red-flag rules, written as regular expressions, for heart attack, stroke, breathing trouble, unconsciousness, severe bleeding, severe allergy, poisoning, pregnancy danger, head injury and meningitis (fever with a stiff neck), and Hinglish phrases such as "seene me dard" are included. The rules look at the words just before a match so that "I have no chest pain" does not raise an alarm. A match produces an emergency card with a call to 108, a call to the user's emergency contact and nearby hospitals; messages about self-harm get a separate support card with Tele-MANAS (14416) and 112. A test script of 16 cases, including negations, covers these rules.

If no red flag is found, the backend builds a short summary of the patient's records: age, gender, pregnancy, conditions, allergies, current medicines (from the last three prescriptions) and any blood value that is abnormal or changing (for example "hemoglobin 10.8, low, falling", from the last three reports). Gemini receives this summary with strict rules: never diagnose, never suggest medicines, choose from an Indian specialist list, and give an urgency of home, soon or now. It runs at temperature 0.3 and answers in JSON, so that the app can draw fixed fields. The reply is either a follow-up question with two to four tap options (at most two questions) or an answer card with the specialist, urgency, a reason in at most 20 words, a note from the patient's records, what to tell the doctor and, only for home urgency, simple home care. Buttons let the user book an appointment in the calendar, find nearby care on a map or save the visit. Chat history is not stored; only the actions the user chooses are saved. The same tiredness means something different when hemoglobin is falling, which is what separates this assistant from a generic chatbot such as [3].

H. Emergency Card and Offline Verification

From the profile the patient downloads a phone wallpaper with a QR code in the bottom-right corner. The card text M joins name, age, gender, blood group, height, weight, BMI, allergies, conditions, emergency contact and creation time with a separator. It deliberately holds no address and no reports. The server signs M with an Ed25519 private key that is kept outside the code repository:

$$\sigma = \mathrm{Sign}(sk, M), \text{ accept} \Longleftrightarrow \mathrm{Verify}(pk, M, \sigma) = 1. \quad (4)$$

The 64-byte signature is combined with a version byte and M and encoded in Base64url, and the QR holds the address of the scan page followed by this token after a # sign. Because the part of a URL after # is never sent to a web server,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

the medical data stay inside the QR and the scanner's browser. The scan page is a static site on GitHub Pages. It reads the token, verifies the signature with the public key written in the page (TweetNaCl) and shows a green "Verified by MediDocker" banner or a red "Fake or edited card" warning if even one character was changed; older cards show "Old card (not verified)". A service worker caches the page and the verification library after the first visit, and since verification

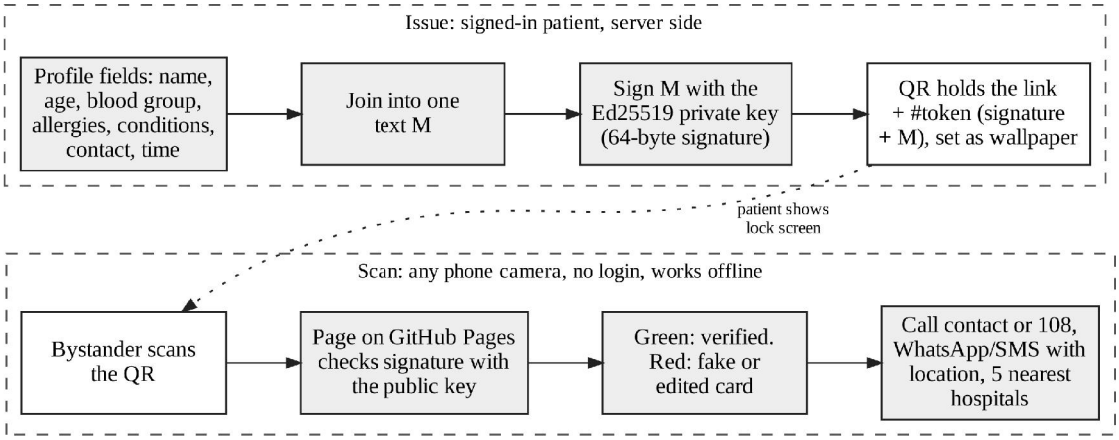


Fig. 5. Issue and scan flow of the signed emergency card.

TABLE III: FAILURE CASES AND SYSTEM BEHAVIOR

If this fails	The system does this
OCR finds fewer than 4 of 8 core values	Qwen2.5-VL reads the report instead
A value looks implausible	Confidence lowered, "please verify", patient can correct
Report has no printed range	Sex-specific default range, stated on screen
RxNorm has no match	Brand dictionary, then 182-drug table, else "unverified"
RxNorm or openFDA offline	30-day cache, otherwise "unverified"
Gemini busy or down	Retry after 800 ms, three backup models, then offline keyword match to 13 specialist rules
Emergency words in chat	Red-flag rules run without the language model
Fewer than 6 CBC values, or low confidence	No answer, or "uncertain: take a ferritin test"
Scanner phone is offline	Page and verifier load from the service-worker cache
Signing key is missing	Wallpaper download fails instead of issuing an unsigned card

TABLE IV: CROSS-VALIDATED AUROC OF THE CANDIDATE MODELS

Model	Main idea	Cross-validated AUROC
Logistic regression	Weighted sum of features, squeezed to 0–1	0.830
Random Forest	400 independent trees that vote	0.837
LightGBM (deployed)	500 trees, each fixing earlier errors	0.847

is pure arithmetic on the phone, no server or database is needed. This is why the card works in a basement or a place with poor coverage.

The page then offers one-tap help: a WhatsApp or SMS message to the emergency contact with the scan time, blood group and a map link to the current location, a call to the contact, a call to 108, and the five nearest hospitals with directions and phone numbers, searched within 5 km and widened to 15 km when fewer than three are found. Hospitals come from OpenStreetMap data through the Overpass API and are shown with Leaflet. If the profile changes, the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

patient downloads a new wallpaper, which carries newly signed data. If the private key were ever exposed, a new key pair would be generated, the public key in the page replaced and the wallpapers re-issued.

I. Fail-Safes

Every AI or online step has a fallback, and the safety-critical parts (emergency detection, QR verification and normal ranges) need no AI at all. Table III lists the main cases.

V. RESULTS AND DISCUSSION

A. Model Selection

The three algorithms were compared by five-fold cross-validated AUROC on the training data (Table IV). LightGBM ranked first with 0.847, ahead of the Random Forest (0.837) and logistic regression (0.830), and was chosen. The small gap between the best and the simplest model also suggests that the CBC signal is real and does not depend on one algorithm.

TABLE V: PERFORMANCE ON THE UNSEEN TEST SET (1,945 PEOPLE, 478 IRON DEFICIENT)

Measure	Value
AUROC (95% CI, 500 bootstraps)	0.836 (0.816–0.857)
Accuracy on decided cases	94.2%
Positive predictive value (“likely”)	91.5%
Negative predictive value (“unlikely”)	97.8%
Sensitivity at the low cut-off	99.6%
Specificity at the high cut-off	99.3%
Brier score	0.139
Calibration slope (ideal 1)	1.048
Coverage (share of people given an answer)	10.6%

B. Performance on Unseen Data

The deployed model was then scored once on the 1,945 held-out participants (Table V). The AUROC of 0.836 means that a randomly chosen deficient person receives a higher score than a randomly chosen healthy person 83.6% of the time (0.5 is a coin toss, 1.0 is perfect); its 95% confidence interval of 0.816–0.857 comes from 500 bootstrap resamples. It is close to the cross-validated 0.847, so performance did not collapse on later survey years. Among people who received a “likely” or “unlikely” answer, 94.2% were correct. A “likely” was right 91.5% of the time and an “unlikely” 97.8% of the time. The sensitivity of 99.6% at the low cut-off means that almost every truly deficient person is kept out of the “unlikely” group, and the specificity of 99.3% at the high cut-off means that almost nobody with normal iron stores is labeled “likely”. The Brier score is 0.139, where a model that always answered 50% would score 0.25.

C. Discussion

The coverage figure needs to be read with the accuracy. Only about one person in ten (10.6%) receives a confident answer, and the rest are sent for a ferritin test. This is a deliberate choice. A screening tool that says “you are fine” wrongly gives false comfort, so the model answers only when it is above the 90% and 97% targets and otherwise defers to the laboratory test that doctors use to confirm iron stores. In our data, simple rules such as low hemoglobin or small red cells caught only 32–42% of iron-deficient people, so even this limited coverage adds information that the rules cannot give.

Several other design choices explain the behavior of the platform. Restricting the vision model to transcription means that an unreadable word produces an “unverified” card rather than an invented explanation, which is the safer failure. Holding doubtful blood values back until the patient confirms them means a wrong digit costs a tap and does not

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, November 2026
Impact Factor: 8.2

corrupt a trend. Running emergency detection as rules means it works even when the language model is down. Hiding the vitamin B12 model, whose AUROC was 0.648, follows the same principle: only validated models are shown to patients.

VI. LIMITATIONS

NHANES describes a United States population, so cut-offs and prevalence may differ for Indian patients, and the test set from 2021 onwards is a time-based check, not an external one; validation on Indian laboratory data is still needed. The model covers one condition, uses CBC values only, and decides for a minority of people. The accuracy of the report and prescription readers has not yet been measured on a large labeled sample, and handwritten prescriptions in particular need a larger evaluation. On the engineering side, the token is kept in browser storage, email verification is implemented but not yet switched on, and a public deployment would need restricted origins, upload size limits and HTTPS. The emergency signature proves that a card is genuine but does not hide its content, so anyone who can scan the QR code can read the emergency fields. The platform has not yet been tested with clinicians.

VII. CONCLUSION AND FUTURE WORK

MediDocker keeps a patient's prescriptions, blood reports, visits, appointments and expenses in one encrypted place and explains them in plain language, while leaving every medical statement to a verified source, a fixed rule or a model that was trained and tested on laboratory-confirmed data. The iron-deficiency model reached an AUROC of 0.836 and 94.2% accuracy on the cases it decides in unseen data, the assistant detects emergencies before any language model is used, and the signed emergency card can be verified by anyone, offline. Future work includes validating the model on Indian patients, adding vitamin B12 and kidney screening once they are validated, doctor access, saved chat history, email verification and a cloud deployment with HTTPS.

ACKNOWLEDGMENT

The authors thank the Department of Computer Science and Technology, Usha Mittal Institute of Technology, for its support throughout this work.

REFERENCES

[1] M. O. Musa et al., "Design and development of a patient record management information system," IJES, 2024.
[2] V. V. Mahale, S. Wagh, R. Sable, and P. Pardeshi, "SympCare: A web-driven diagnostic framework for CBC-based disease prediction using machine learning," IEEE, 2025.
[3] S. M. Hussain, "HealthBot: A chatbot system for specialist recommendation based on symptoms input," IJCRT, 2023.
[4] C. A. Secco and K. Nazemi, "A modular visual analytics dashboard for patient health data," Information Visualization (SAGE), 2025.
[5] Y. Guo et al., "Detection of medication mentions and medication change events in clinical notes using transformer-based models," MEDINFO 2023, IOS Press, 2024.
[6] S.-K. Chen and Y.-W. Ti, "A design of multi-purpose image-based QR code," Symmetry (MDPI), 2021.
[7] S. Ameen, R. Balachandran, and T. Theodoridis, "Deriving hematological disease classes using fuzzy logic and expert knowledge: A comprehensive machine learning approach with CBC parameters," arXiv:2406.13015, 2024.
[8] M. Bibi and W. K. Saeed, "AI-enabled multi-disease diagnosis using low-cost CBC reports," Frontiers, 2026.
[9] Rahul et al., "A comprehensive approach to using QR code technology and digital health tracking to improve emergency medical response," IJERMT, 2024.
[10] "Intelligent system for scanning & processing doctor's prescriptions using OCR and NLP," JAAFR, 2025.