

Problem Statement

MOTIVATION

Implicit Health Cues

During home hospitalization, the patient is not under constant supervision by medical staff and may experience a deterioration in critical indicators (such as blood sugar, blood pressure, or oxygen saturation) or symptoms (such as pain or shortness of breath) without realizing how severe the situation is. The lack of continuous monitoring can lead either to unnecessary hospitalization or, conversely, to neglect of a medical emergency.

Example

Mr. Levi, 73, receiving home hospitalization for chronic heart failure, reported via the app a 3 kg weight gain over three days, shortness of breath at rest, and significant ankle swelling. The system identified a critical cardiac deterioration, integrated sensor data (elevated weight and blood pressure), and immediately alerted the medical team for an urgent home visit with a recommendation to adjust diuretics or consider ER referral.

Application

Without an integrative tool, patients and staff rely on sporadic visits and phone calls that don't provide a full picture. A dedicated app enables trend detection, automatic alerts, and continuous care—crucial for the safety and success of home hospitalization.

NLP TASK

Challenge:

Interpreting patients' free-form reports in everyday language (text and/or speech) and combining them with physiological measurements (blood pressure, glucose, weight, etc.) to detect changes in health status

Input:

- Free-form transcription (written or spoken) from the patient.
- Quantitative data from connected home monitoring devices (blood pressure monitor, blood glucose, weight scale, oxygen saturation, etc.).

Output – Structured Version:

- Classification of the patient's condition into one of the following categories:
- Deterioration /Improvement/ Stable/Normal
- Triggering an alert when needed: ER referral, treatment adjustment, urgent home visit.
- Real-time recommendations to physician/nurse via a dashboard.

Example

Textual input: "I gained 2 kg in 48 hours, and I'm having trouble breathing at rest."

Quantitative input: Sharp weight increase; blood pressure 165/105.

Output: Classified as "cardiac deterioration"; alert sent to physician with recommendation for urgent home visit or adjustment of diuretics

Training and Test Data

"Home Hospitalization Symptom Reports Dataset"

- A database of free-form transcriptions from home-hospitalized patients, labeled by status: deterioration / improvement / stable
- Used to train models for NER (Named Entity Recognition) to identify symptoms, and for text classification to assess patient condition

"Home Monitoring Sensor Dataset" Data from connected home sensors:

- Daily blood pressure, oxygen saturation, heart rate, weight, activity level
- Additional measurements as needed (temperature, blood glucose, urine output)
- Used for system testing and validation in continuous monitoring scenarios

Data Integration:

- Integration of free-form transcriptions with real-time sensor readings
- Training and validation of the Virtual Care Assistant on deterioration and improvement scenarios
- Supports tasks such as symptom extraction, triage prediction, and realtime recommendations

LLM Integration:

- Use of a Large Language Model (LLM) to interpret patient-reported descriptions
- Extract, normalize, and classify symptoms using natural language processing (NLP)
- Enhance triage decisions by combining chronological text with sensor data
- Generate real-time, personalized recommendations for the patient and medical team via a dashboard

Evaluation

Closed-World Metrics Classification Metrics (Symptom Extraction / Triage)

- Precision, Recall, F1
- Measure the accuracy of identifying known symptoms or classifying patients into risk levels

Baseline

- Dictionary-based / keyword-matching approach for symptom extraction
- Static rules (thresholds) applied to sensor values



Open-World Metrics

Text Generation & Recommendation Quality

- ROUGE / BLEU or LLM-based evaluation for clarity and accuracy of recommendations
- Human rating (clinician-rated helpfulness)

Clinical Impact & Outcomes

- Reduction in unnecessary ER visits
- Shorter time-to-intervention

Patient & Clinician Satisfaction

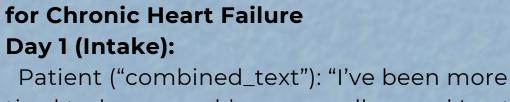
SUS or NPS surveys to assess system usability and trust

Baseline

- Template text messages without contextual adaptation
- Manual recommendations based on static clinical rules



PROMPT 1



Scenario - Mr. Levi, 73, Home Monitoring

tired today, my ankles are swollen, and I get out of breath walking from the bedroom to the living room."

Vitals: BP 145/95 mmHg, SpO₂ 92 %, Pulse 84 bpm

System: Classified as "Critical Deterioration"; alert sent to CHF care team.

Day 2 (No Change):

Patient ("combined_text"): "I still weigh 71 kg, my ankle swelling is the same, and breathing feels no easier when I lie flat."

Vitals: BP 147/97 mmHg, SpO₂ 91 %, Pulse 85 bpm

System: Flagged as "No Improvement"; continue current diuretic dose and monitor closely.



PROMPT 2

Scenario – Ms. Cohen, 65, Home Treatment for Left-Leg Cellulitis Day 1 (Intake):

Patient ("combined_text"): "My left lower leg is red and swollen. I feel a mild burning sensation but I'm not feverish."

Vitals: Temp 37.1 °C, BP 125/80 mmHg, Pulse 78 bpm

System: "Routine Monitoring" – continue home antibiotics and leg elevation.

Day 2 (Mild Improvement):

Patient ("combined_text"): "The redness is fading, and the swelling is less pronounced. I still feel a little tenderness but no new pain or fever."

Vitals: Temp 36.8 °C, BP 120/78 mmHg, Pulse 75 bpm

System: "Stable Trend" – maintain current antibiotic regimen and daily assessment.



PROMPT 3

Scenario – Mr. Ben-Ami, 58, Home Monitoring for COPD Day 1 (Intake):

Patient ("combined_text"): "I've been coughing more than usual, and when I lie down I feel short of breath quickly."

Vitals: SpO₂ 91 %, BP 130/85 mmHg, Pulse 88 bpm

System: "Watch – Early Instability" flagged; start daily remote check-in protocol.

Day 2 (Deterioration):

Patient ("combined_text"): "Today even sitting up I'm very short of breath, and I'm exhausted after just a few steps."

Vitals: SpO₂ 88 %, BP 140/90 mmHg, Pulse 96 bpm

System: "Critical Deterioration" flagged; urgent home visit triggered and inhaler therapy adjusted.

Data

In this project, we synthetically generated a dataset using GPT-based prompting to simulate realistic home hospitalization cases.

We crafted structured prompts to produce diverse, natural-language caregiver reports that included:

Symptom descriptions in everyday language

Vital signs embedded in the text (HR, RR, BP, Temp)

A status label (0 = No change, 1 = Improvement, 2 = Deterioration)

Each sample was returned in JSON format with both free text and structured metadata. This dataset was essential for training and evaluating our models on both classification and multimodal fusion tasks.



Sample GPT Prompt Used for Dataset Creation

```
diagnoses = [
"Hypertension", "COPD", "CHF", "Pneumonia", "Cellulitis", "UTI",
 "DEHYDRATION", "FEVER", "Bronchiectasis", "DIARRHEA", "influenza",
 "gastroenteritis", "Skin infections", "heart failure", "hyperemesis"
genders = ["Male", "Female"]
condition_changes = ["no change", "improvement", "deterioration"]
def generate_prompt(patient_id, age, gender, diagnosis, change, description_length, tone):
return f"""
You are simulating progress notes from a patient at home. Write a JSON object describing their experience over two days.
Fields to include:
- "patient_id": "{patient_id}"
- "age": {age}
- "gender": "{gender}"
- "diagnosis": "{diagnosis}"
- "change": "{change}"
- "narratives": {{
"Day 1": a short first-person message written by the patient (about {description_length}). It should describe how they feel in simple, everyday language, matching their
diagnosis. Use natural, human phrasing like someone texting or journaling. At the end of the message, include their vitals: HR, BP, Temp (°C), RR.
"Day 2": a similar first-person message, with different symptoms or experiences. Include new vitals at the end.
- "reasoning": A brief, objective explanation of how the patient's change is inferred from their complaints and vital signs. Avoid emotional or vague wording like "better" or
"worse".
Return valid JSON only. No markdown, no extra commentary. Only the JSON object.
```

