A Patient-Centric Dialysis Decision Support System: From Lab Report Parsing to Diet-Aware Short-Horizon Electrolyte Projection

Abstract

Patients with chronic kidney disease (CKD) face day-to-day decisions impacted by fluid intake, electrolyte loads, and dialysis scheduling. We present a practical, end-to-end system that: (1) parses Kidney Function Test (KFT) reports to extract key labs; (2) quantifies nutrients and fluids from free-text diet logs; (3) projects short-horizon changes in serum sodium and potassium from planned intake; and (4) integrates a pre-trained neural predictor for dialysis frequency. The system combines robust PDF parsing, canonicalized nutrition lookup, physiologically grounded fluid–electrolyte kinetics, and a Keras model for dialysis cadence estimation. A demonstration scenario yields a −2.0 mEq/L sodium dilution and +0.33 mEq/L potassium rise at 1 hour after 1.47 L of intake, illustrating actionable 'what-if' insights. This paper details the architecture, algorithms, and demonstration results, and outlines limitations and future work.

1. Introduction

Fluid and electrolyte management is central to CKD and dialysis care. We propose an end-to-end assistive system that operationalizes this workflow: parse KFT lab reports, translate diet logs into nutrient/fluid loads, predict near-term sodium/potassium changes from intake kinetics, and estimate dialysis cadence using a pre-trained model. The design emphasizes practical robustness, interpretability of intermediate features, and seamless interactivity.

2. System Overview

KFT parsing extracts Creatinine, Urea, Sodium, Potassium, and Phosphorus from PDFs using geometry-aware heuristics. Diet-to-features canonicalizes free-text items to nutrition tables and computes per-item Water, Na, K, P, Protein with drink-specific sodium concentration and absorption half-lives. A short-horizon kinetics model predicts dilutional ΔNa and ECF-linked ΔK. A pre-trained Keras model estimates dialysis frequency. A Streamlit UI orchestrates the pipeline and logs sessions.

3. Architecture

Key modules: (a) KFT parser (pdfplumber-based) for robust extraction; (b) Food lookup with unit parsing, canonicalized columns, fuzzy matching, beverage inference, and category-specific Na concentration; (c) Fluid kinetics with TBW mixing for sodium and ECF appearance plus insulin shift for potassium; (d) Optional Keras predictor and scaler; (e) UI and CSV history for longitudinal tracking.

4. Methods

KFT PDF Parsing: words are grouped into lines; 'Result' column bounds are detected; synonym dictionaries map heterogeneous labels; plausibility filters remove ranges and implausible values. Diet Featureization: quantities unified to grams; match priority exact→startswith→contains→fuzzy; per-100g scaling computes nutrients; beverage detection drives Na concentration bounds and absorption t½. Fluid–Electrolyte Projection (~1h): TBW = weight×TBWfraction; ECF ≈ 0.2×weight. Per-item absorption fraction uses first-order kinetics; sodium mixes with TBW under a ±2 mEq/L/h cap; potassium appears in ECF then shifts with $\tau \approx 1.2$ h. Dialysis Predictor: a pre-trained Keras model consumes standardized clinical and lab features to estimate sessions/week.

5. Results

Demonstration scenario (baseline Na=138 mEq/L, K=4.5 mEq/L; 70 kg): intake of 1100 ml Apple Juice, 300 g Watermelon, and 1 cup Soy milk. Computed totals: 1.468 L water, 135.8 mg Na, 1735 mg K, 186.8 mg P, 9.79 g protein. 1-hour projection: ΔNa = −2.0 → Na_new = 136.0 mEq/L; ΔK = +0.33 → K_new =

4.83 mEq/L. Magnitudes align with expected physiology for hypotonic/sweet beverages and mixed solids.

6. Discussion

Robustness stems from geometry+text fallbacks in KFT parsing and multi-strategy food matching. Interpretability is provided by per-item fields (na_conc_mmol_per_L, absorption t½). The system supports ‘what-if’ counseling before dialysis and integrates a model for cadence context.

7. Limitations

No labeled benchmark for KFT parsing is bundled; nutrition matching may misclassify composite dishes; kinetics omit renal handling, medications, and acid–base effects; the dialysis model’s training data and metrics are not included.

8. Future Work

Add labeled PDFs and metrics, expand nutrition coverage and recipe decomposition, extend kinetics beyond 1 h with personalization and uncertainty bounds, and open a documented training pipeline for the dialysis model.

9. Ethics and Safety

Decision support only; not a substitute for clinical care. Protect PHI in logs; add alerting guardrails for hyperkalemia or hyponatremia risk.

10. Reproducibility

Python 3; key libraries: pdfplumber, pandas, joblib, tensorflow (for inference). Data: CSVs under app/data/. Run the Streamlit app (app_ui.py) to orchestrate; process_history.py builds feature CSVs.

11. Conclusion

We connect lab parsing, diet understanding, physiological projections, and dialysis frequency inference in one workflow. The demo shows immediate, interpretable insights for clinic and patient use.

Appendix A: Demonstration Output

Water_ml: 1468.53; Sodium_new: 136.0 (Δ –2.0); Potassium_new: 4.83 (Δ +0.33). Baseline rule-of-thumb: if Potassium>6.0 or Creatinine>10.0, flag for urgent dialysis consideration.