International Journal of Multidisciplinary Studies and Innovative Technologies

e-ISSN: 2602-4888
dergipark.org.tr/en/pub/ijmsit
Research Article

2022, X (X), pp-pp
DOI: 10.36287/ijmsit.xxxx
Received: xx.xx.2022; Accepted: xx.xx.2022

AI Tabanlı Kan Testi Analiz Chatbot'u: PDF Tabanlı Tıbbi Verilerin Türkçe Yorumlanması

Sidal Deniz Bingöl [1] ve Damla Nur Alper [2]

[1]İzmir Bakırçay Üniversitesi,Mühendislik Fakültesi, Bilgisayar Mühendisliği, İzmir, Türkiye (210601009@bakircay.edu.tr) (ORCID: 0009-0004-0666-0920)*
[2] İzmir Bakırçay Üniversitesi,Mühendislik Fakültesi, Bilgisayar Mühendisliği, İzmir, Türkiye (220601017@bakircay.edu.tr) (ORCID:0009-0002-0210-4826)

Türkçe Özet – Bu çalışma, e-Nabız üzerinden alınan kan tahlili PDF dosyalarını analiz eden, doğal dilde açıklamalar sunan Türkçe bir yapay zekâ tabanlı chatbot sistemi geliştirmeyi amaçlamaktadır. Geliştirilen sistem, kullanıcıların yüklediği PDF dosyalarını PyMuPDF kütüphanesiyle ayrıştırarak JSON formatında veri elde eder. Ardından bu veriler, referans aralıklarına göre sınıflandırılır ve Ollama platformu üzerinde çalışan Mistral dil modeliyle analiz edilir. Sistemin ön yüzü React ile, arka planı ise FastAPI kullanılarak geliştirilmiştir. Chatbot, yalnızca mevcut test verileriyle çalışmakta ve eksiksiz/eksik sonuçlara göre sade ve kişiselleştirilmiş öneriler sunmaktadır. Benzerlerinden farklı olarak sistem, doğrudan sayısal tıbbi verilere dayalı yorumlama yapar; semptom temelli değildir. Türkçe olması, dosya tabanlı çalışması ve yorumların yönlendirilmiş olması projeyi özgün kılmaktadır.

Anahtar Kelimeler – Kan Testi, Yapay Zeka, Chatbot, PDF Analizi, Türkçe Sağlık Uygulaması

Atıf Alper, D. N., Bingöl, S. D. (2025). AI Tabanlı Kan Testi Analiz Chatbot'u: PDF Tabanlı Tıbbi Verilerin Türkçe Yorumlanması. International Journal of Multidisciplinary Studies and Innovative Technologies , x(x): xx-xx.

AI-Based Blood Test Analysis Chatbot: Interpretation of Medical PDF Reports in Turkish

Extended Abstract

Research Problem/Questions – How can users interpret blood test results taken from Turkish e-health systems (e.g., e-Nabız) in a meaningful and personalized way without medical knowledge?

Short Literature Review – While AI-based health chatbots like Ada Health or Infermedica rely on symptoms entered by the user, few tools exist to automatically analyze structured medical documents like blood test PDFs. Most such systems are English-based and lack support for Turkish-language medical data.

Methodology – Blood test PDFs are parsed using PyMuPDF and regex patterns. Extracted data is transformed into structured JSON and labeled by comparing with reference ranges. Relevant information is retrieved from a knowledge base and passed to the Mistral model via a prompt template. All layers communicate over FastAPI with a React-based frontend.

Results and Conclusions – The system successfully provides simple, understandable, and accurate explanations for blood test results in Turkish. Users receive real-time, context-based suggestions. The modular design enables easy expansion to other test types. It improves accessibility of medical data for non-expert users in Türkiye.

Keywords – Blood Test, AI Chatbot, Medical NLP, RAG, PDF Parsing

Citation: Alper, D. N., Bingöl, S. D. (2025). AI-Based Blood Test Analysis Chatbot: Interpretation of Medical PDF Reports in Turkish. International Journal of Multidisciplinary Studies and Innovative Technologies , x(x): xx-xx.

I. INTRODUCTION

This study aims to develop a Turkish-language AI chatbot system that analyzes blood test PDF reports obtained from the e-Nabız platform and explains the results in understandable language. Unlike symptom-based systems such as Ada Health or Buoy Health, this chatbot directly processes quantitative medical data. The need for a localized, document-based interpretation tool is evident, particularly in regions like Türkiye where e-Nabız is widely used. The purpose of this

work is to provide users with personalized, accurate, and reliable interpretations of their health data using advanced AI methods.

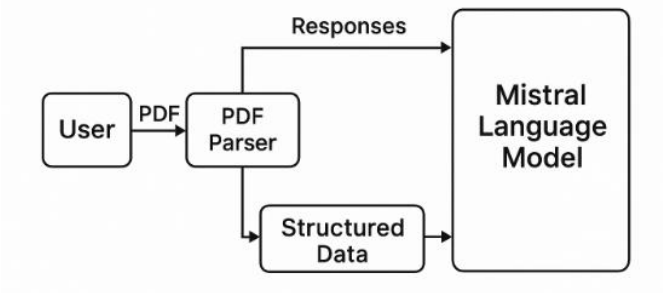
II. Materials and Method

The system architecture is composed of several key components designed to parse, analyze, and present medical data extracted from blood test reports in PDF format. The frontend of the application is built using **React.js**, enabling users to upload their lab reports in a user-friendly environment. Once uploaded, the PDF is processed by a backend service built with **FastAPI**, which utilizes **PyMuPDF** and regular expressions to extract relevant medical data from unstructured text.

The extracted values are then mapped to a **structured JSON format**. These values are programmatically compared to standard reference intervals to identify abnormalities (e.g., high or low test results). A separate medical knowledge base, including predefined explanations and recommendations, is maintained in a JSON file and used for reasoning.

For natural language responses, the system uses the **Mistral language model**, integrated via the **Ollama** framework. A customized prompt template is constructed dynamically by merging user-specific lab data with predefined medical insights. The model then generates context-aware, human-readable answers. Only the values present in the uploaded report are used—no assumptions are made for missing tests.

This modular pipeline ensures accuracy, flexibility, and easy expansion to other medical test types.



An example of the table is given below.

Table 1. Example of a table

Test Name	Measured Value	Reference Range	Interpretation
Glucose	110 mg/dL	70–99 mg/dL	High – May indicate early diabetes or stress.
Hemoglobin	12.5 g/dL	12–16 g/dL	Normal.
MCV	74 fL	80–100 fL	Low – Possible iron deficiency.
ALT	45 U/L	7–35 U/L	Slightly High – May indicate mild liver stress.

III. Results

The developed system successfully interprets blood test results from e-Nabız PDFs and produces customized, concise explanations for each out-of-range value. For example, questions like 'What does low MCV mean?' are answered with context-specific insights about iron deficiency. Users receive health-related guidance in Turkish, making medical data more accessible

IV. Discussion

This chatbot differs from existing systems by focusing on medical documents rather than symptoms. The Turkish language support and PDF-based data extraction make it a unique and localized solution. Its modular architecture allows for scalability and integration with other test types in the future.

V. Conclusion

The project presents an innovative AI-driven approach to interpreting blood test results for non-expert users. It simplifies complex medical data, improves patient engagement, and can potentially reduce reliance on repetitive medical consultations for basic lab interpretation. Further improvements could include multi-test comparison and integration with wearable health devices

Acknowledgment

This study was conducted as part of an undergraduate computer engineering project course. The authors thank their academic advisor for their guidance and feedback

Authors'Contributions

The authors' contributions to the paper are equal.

Statement of Conflicts of Interest

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The authors declare that this study complies with Research and Publication Ethics

References

[1] T.C. Sağlık Bakanlığı, 'Kan Tahlili Değerleri ve Açıklamaları'.
[2] E-Nabız Kişisel Sağlık Sistemi, https://enabiz.gov.tr/

[3] Lab Tests Online TR, 'Kan Testleri Rehberi', https://labtestsonline.org.tr/
[4] Mayo Clinic, 'Complete Blood Count (CBC) Testi', https://www.mayoclinic.org/
[5] PyMuPDF Documentation, https://pymupdf.readthedocs.io/
[6] React Documentation, https://react.dev/
[7] FastAPI Documentation, https://fastapi.tiangolo.com/
[8] Hugging Face, https://huggingface.co/
[9] Ollama, https://ollama.com/
[10] GitHub, PDF Parsing Projects, https://github.com/
[11] Stack Overflow, React Developer Forum, https://stackoverflow.com/
[12] ChatGPT, Technical Support for Prompt Engineering