Multimodal Data Analysis for Preventive Healthcare

Using machine learning to identify risks and personalize interventions

Presented by: Nader Abdullah

Introduction

Project Overview

- Data-driven preventive healthcare system
- Risk prediction for chronic conditions
- Personalized, understandable interventions

Why It's Important

- Chronic diseases cause 7 of 10 US deaths
- Early prevention could save $260B+ annually
- Bridges gap between clinical data and patient understanding

Annual US Healthcare Costs (in Billions)

Preventable Costs

26.0%

74.0%

Current Spending
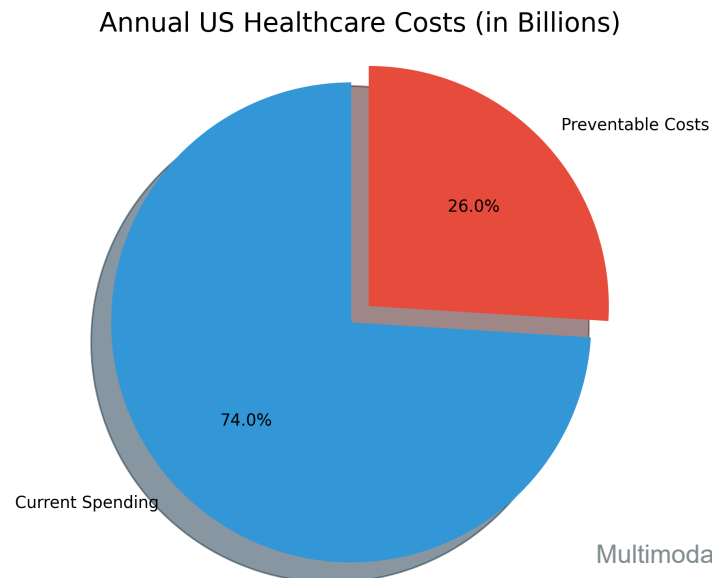
Method

Data Analysis Pipeline

- Synthetic EHR data from Synthea (patients, observations, conditions)
- Feature engineering focused on modifiable risk factors
- ML-based risk prediction with interpretable models
- Explanation generation with clinical context

We integrated GPT-4 to enhance the explanation engine, converting technical risk assessments into natural, empathetic language for improved patient communication.

Technical Implementation

System Architecture
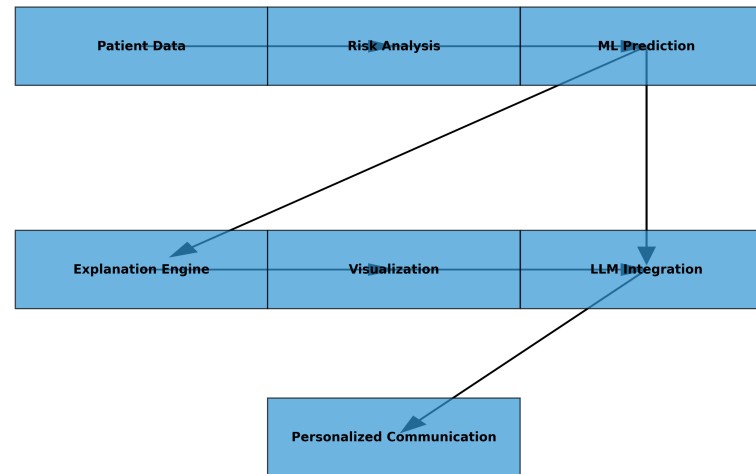
- Random Forest-based classification for risk prediction (95% accuracy)
- Feature importance analysis for risk factor identification
- Rule-based system for translating clinical values to recommendations
- Comprehensive visualizations for patient education

Results

Key Findings

- 95% accuracy for diabetes risk prediction
- 56% projected risk reduction with recommended interventions
- Successfully identified key modifiable risk factors

Example Outputs

5-Year Risk Projection
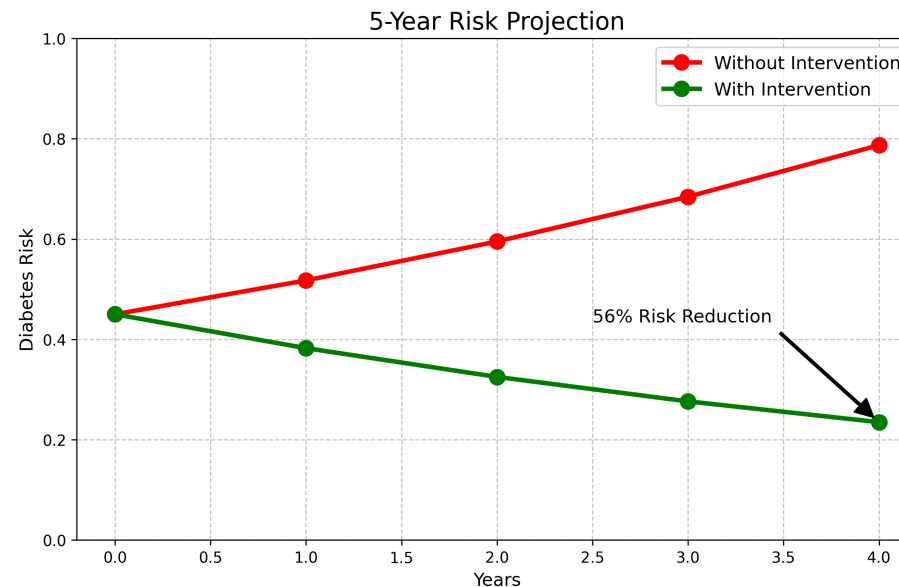
Implications

- Data-driven approach to preventive healthcare
- Improved patient understanding of health risks
- Support for both individual and population health management
- Practical integration pathway for AI in clinical workflows

While the core system functions effectively with rule-based explanations, our GPT-4 enhancement allows for natural, conversational explanations that can be tailored to different health literacy levels.

Example Personalized Communication

Dear Patient,

Thank you for your recent health check-up. I wanted to discuss your results and share some personalized recommendations to help you maintain optimal health.

Based on your recent measurements, we've identified that you may have an increased risk for diabetes. This isn't cause for alarm—many factors contribute to health risks, and taking small steps now can make a big difference for your future well-being.

I recommend focusing on:
• Incorporating more fruits and vegetables into your daily meals
• Taking a 15-20 minute walk after meals when possible
• Scheduling regular check-ups every 6 months
• Monitoring your blood pressure at home if you have a device available

Small, consistent changes often have the greatest impact on long-term health. Remember, these suggestions are preventive measures to help you maintain your health for years to come.

Do you have any questions? I'm here to support you on your health journey.

Wishing you the best of health,
Your Healthcare Team

Future Directions

Core System Improvements

- Enhanced explainability (SHAP values)
- More sophisticated risk models
- Integration with clinical workflows

Expansion Opportunities

- Additional preventable conditions
- Longitudinal analysis capabilities
- Mobile & wearable data integration
- Feedback loops for effectiveness

Future System Ecosystem
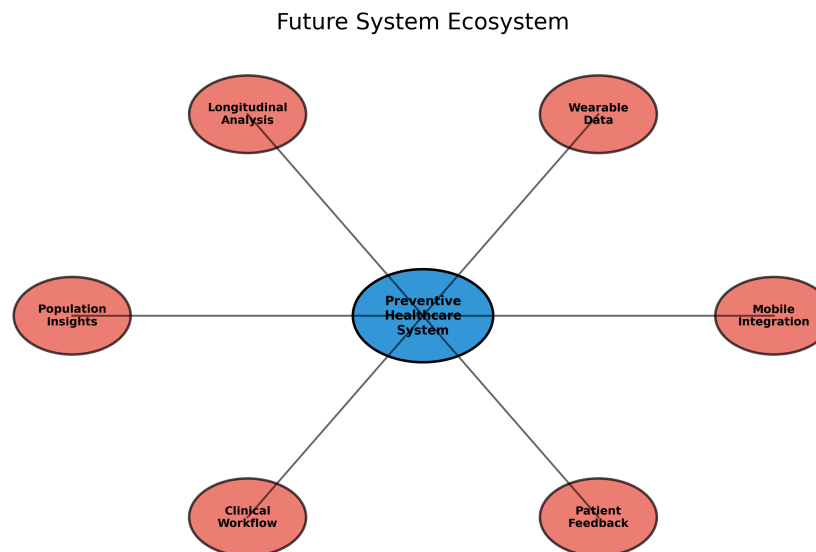
Conclusion

- Data analysis is key to identifying preventable health risks
- ML models can effectively predict chronic disease development
- Interpretable risk explanations improve medical decision-making
- Next steps focus on clinical validation and expanded coverage

While our core system provides effective risk assessment independently, the GPT-4 integration demonstrates how AI language models can enhance communication of complex health information.

Estimated Impact of AI-Driven Preventive Healthcare
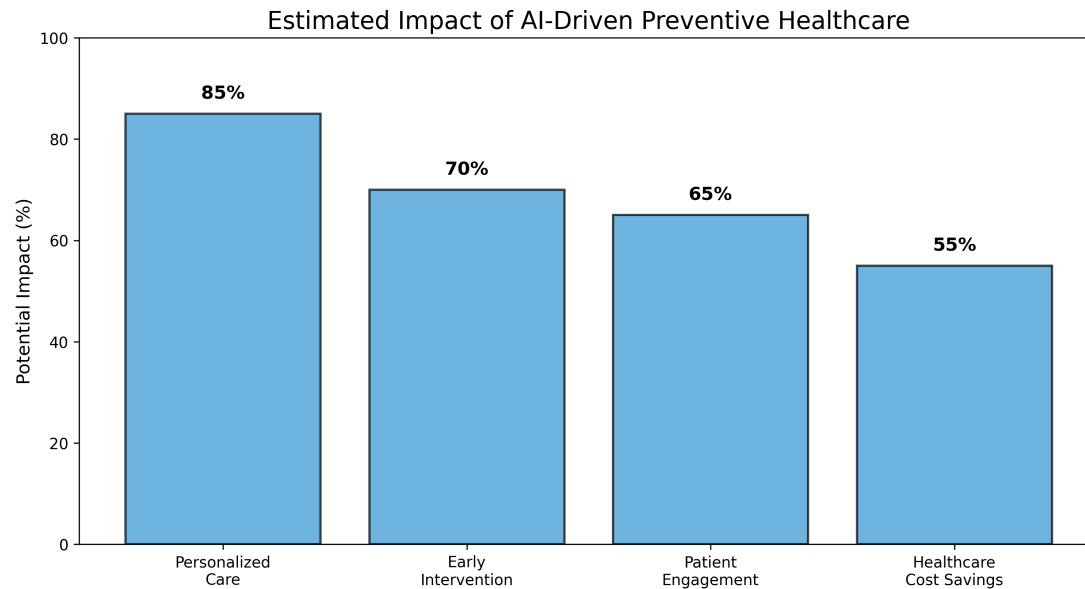
Thank you for your attention. Questions?