



Problem Statement:

Personalized Health Prediction System

Welcome to the Personalized Health Prediction System, a cutting-edge solution designed to provide personalized wellness recommendations. This system leverages synthetic datasets and advanced machine learning techniques to offer actionable insights for optimizing your health. This presentation will cover the problem statement, data and models, training methodologies, system functionality, and future enhancements.

Developed by 1st Years -

Bilir Goyari - 2412053

Gourav Kumar Singh - 2412065

Vadlakonda Vivek - 2412076

Problem Statement

Problem

The system addresses the challenge of providing personalized wellness recommendations. Each individual has unique health profiles and needs. A one-size-fits-all approach to wellness often falls short, leading to ineffective or even harmful advice.



Solution

The system employs advanced ML to analyze various health-related features. The system generates personalized wellness plans, including recommendations for nutrition, exercise, and lifestyle adjustments tailored to individual needs.

Libraries Used



Data & ML

pandas, numpy, scikit-learn (RandomForest, Pipeline, metrics)



Visualization

matplotlib, seaborn



Utils

datetime, random, warnings

The system harnesses the power of several key libraries. **pandas** and **numpy** are essential for data manipulation and numerical computations, while **scikit-learn** provides robust machine learning algorithms and evaluation metrics. **matplotlib** and **seaborn** are used for generating insightful visualizations. Utility libraries such as **datetime**, **random**, and **warnings** enhance the system's functionality and reliability.

Datasets

Fitness Tracker Dataset

Includes steps, workout type/duration, and activity level to measure physical activity patterns.

Sleep Health Dataset

Includes duration, quality, wake-up time, and efficiency to assess sleep habits and patterns.

Lifestyle Dataset

Captures stress, hydration, BMI, heart rate, and blood pressure to understand overall lifestyle factors.

Nutrition Dataset

Encompasses calories, macros (protein/carbs/fats), and meal timing to analyze dietary habits.



Input → Output Flow



Input Features

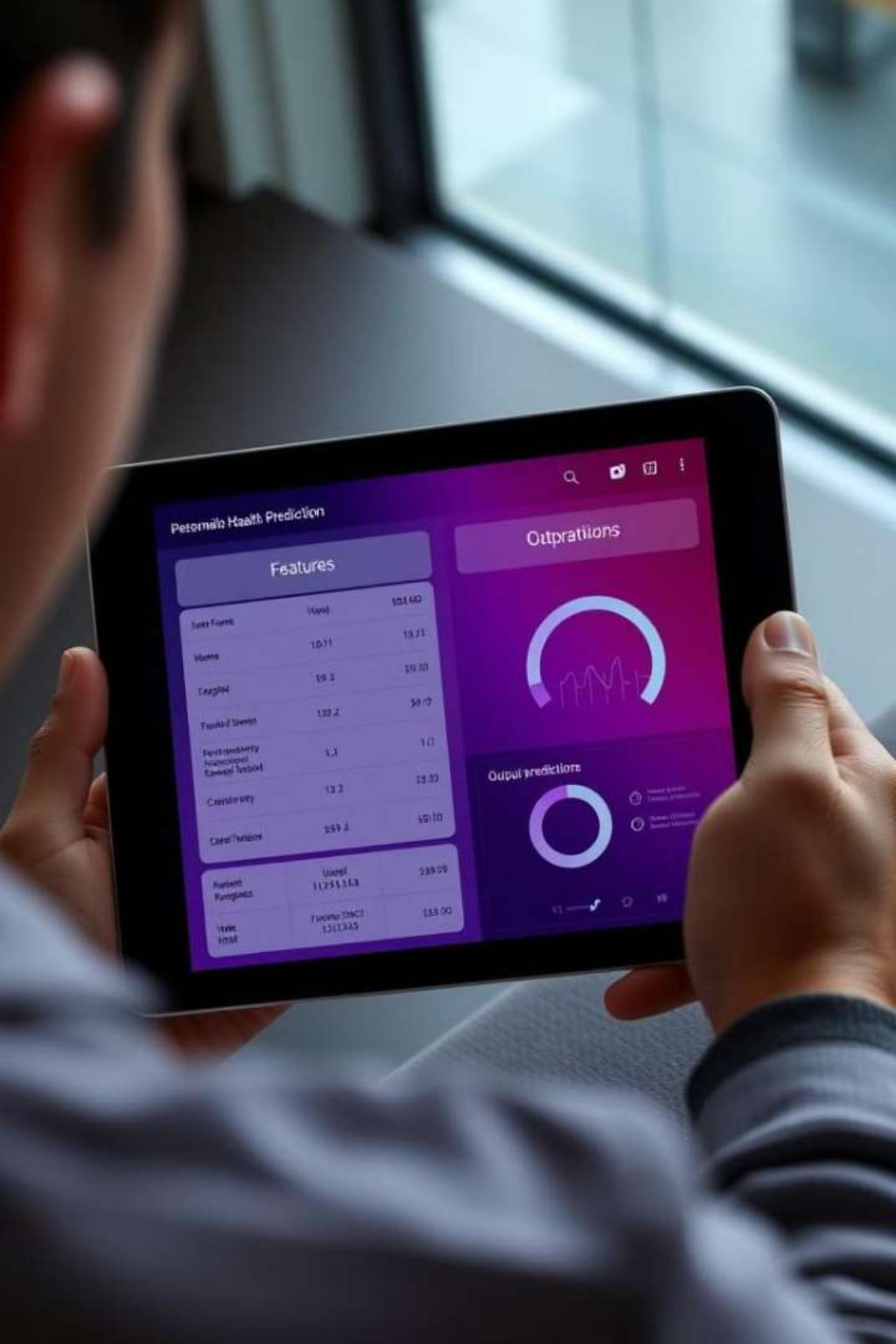
Age, BMI, Sleep Duration, Workout Type, Stress Level, Daily Steps, Hydration, Resting Heart Rate, etc. (20+ features)



Output Predictions

Regression: Calories, sleep hours, workout duration

Classification: Supplements ("Vitamin D, Omega-3"), meal frequency ("3 meals/day")



Training & Testing

1

Preprocessing

- Numerical: Standard scaling (StandardScaler)
- Categorical: One-hot encoding (OneHotEncoder)
- Time features: Converted to minutes (e.g., "06:30" → 390)

2

Model

Multi-Output Random Forest (for regression + classification)

Hyperparameters: `n_estimators=150`, `max_depth=20`

3

Evaluation

Regression: RMSE (e.g., ± 150 kcal), MAE

Classification: Accuracy (>85%), F1-score

The training and testing phase involves rigorous preprocessing to prepare the data. Numerical features are standardized, categorical features are one-hot encoded, and time features are converted to minutes. A Multi-Output Random Forest model is then trained and evaluated using appropriate metrics.

How It Works

Data Generation

Simulates realistic health profiles with correlations (e.g., high BMI → lower activity).

Visualization

Compares user data against training distributions (histograms/bar plots).



User Input

Accepts manual entry or random samples (age, sleep, workouts, etc.).

Prediction

Preprocesses input → ML model → Generates personalized plan.

Future Scope

Real Data Integration

Fitbit/Apple Health API for live tracking.

Chatbot

Gemini/LLM for wellness Q&A (like NeuraDoctor).

Expanded Metrics

Blood tests, genetic data for precision health.

The future scope of the system includes several exciting enhancements. Integrating real-time data from devices like Fitbit and Apple Health will provide more accurate and dynamic recommendations. Implementing a chatbot powered by Gemini/LLM will offer interactive wellness Q&A. Additionally, incorporating expanded metrics such as blood tests and genetic data will enable even more precise and personalized health insights.

