Predicting Health Risk: Blood Pressure Analysis & Web App Development

By Jugal

Problem statement:

- Hypertension affects affects 1 in 3 adults worldwide(WHO, 2023)
- Often underdiagnosed, yet a leading cause of heart disease and stroke(Known as “Silent killer”)
- Early detection using data can help target interventions
- Goal: Predict high-risk individuals using health data

Objective:

📊 Analyze health dataset for patterns related to blood pressure

🧠 Building a predictive ML model

🌐 Deploy it as a simple web application for user interaction

🧩 Incorporate feature engineering to boost model performance

Data Overview

Source: National Health and Nutrition Examination Survey(CDC, https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?Cycle=2017-2020)

Key features:

- Demographics: Age, Gender, Ethnicity
- Physical measurements: BMI, Waist-Hip Ratio, Blood Pressure
- Socioeconomic factors: Education level, Income
- Health behaviors: Smoking status, Physical activity

Target variable Hypertension status (BP ≥ 130/80 mmHg)

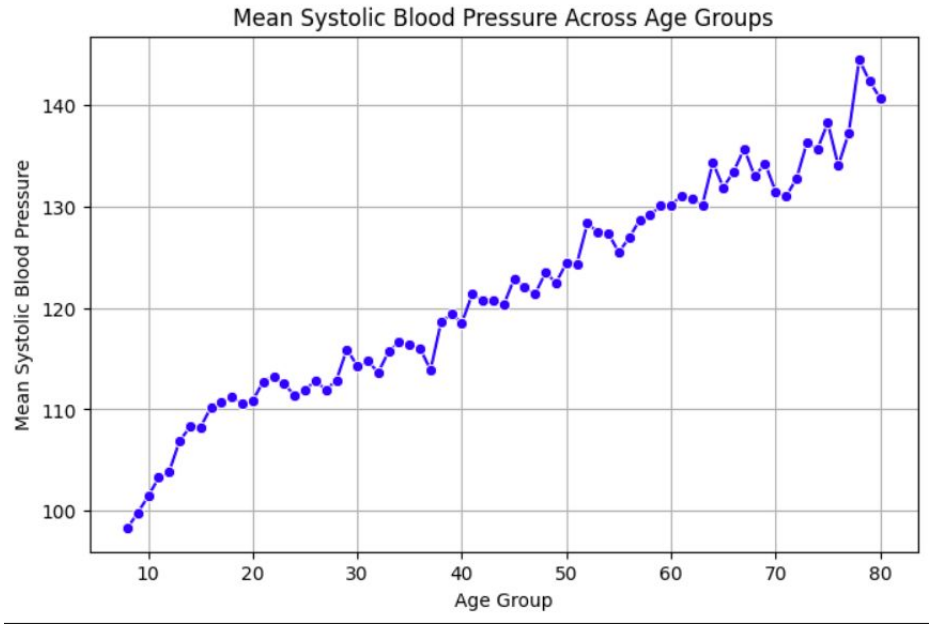
Data Insights

Key Findings from Exploratory Analysis

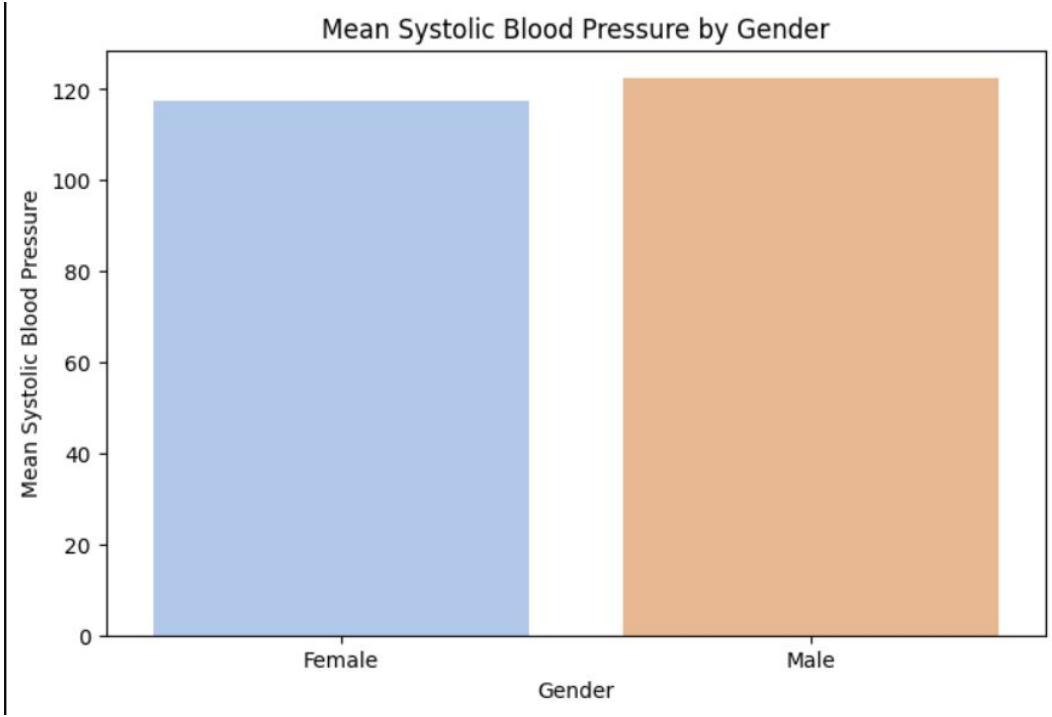
- Strong positive correlation between age and blood pressure
- Men show slightly higher average systolic pressure than women
- Significant variations across ethnic groups with highest prevalence in Black community
- Inverse relationship between education level and hypertension risk

Factors Affecting Blood Pressure:

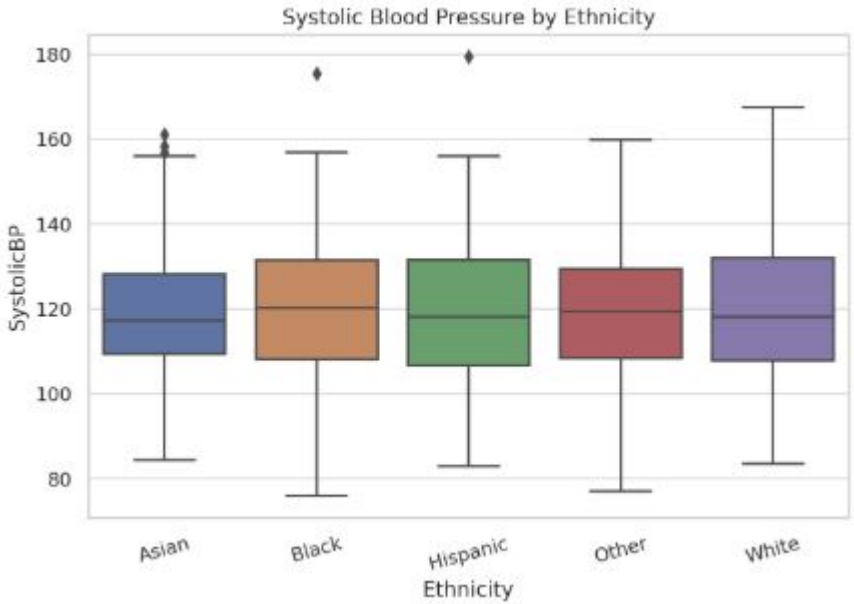
- **Age**: As people age, their blood pressure tends to rise.



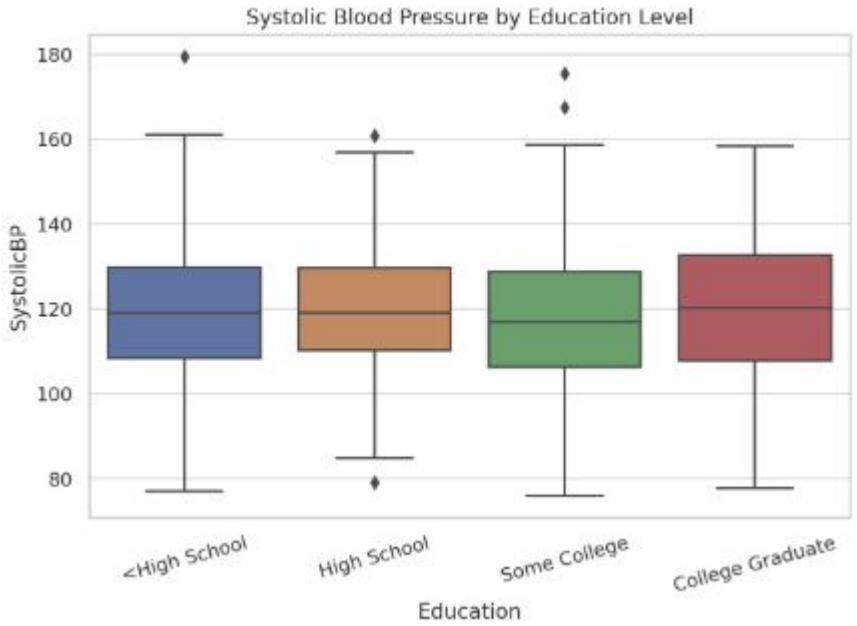
- **Gender: Men and women may experience different risk levels.**



- **Ethnicity: Certain ethnic groups may be more prone to hypertension.**



- **Education: Socio-economic status might also play a role.**



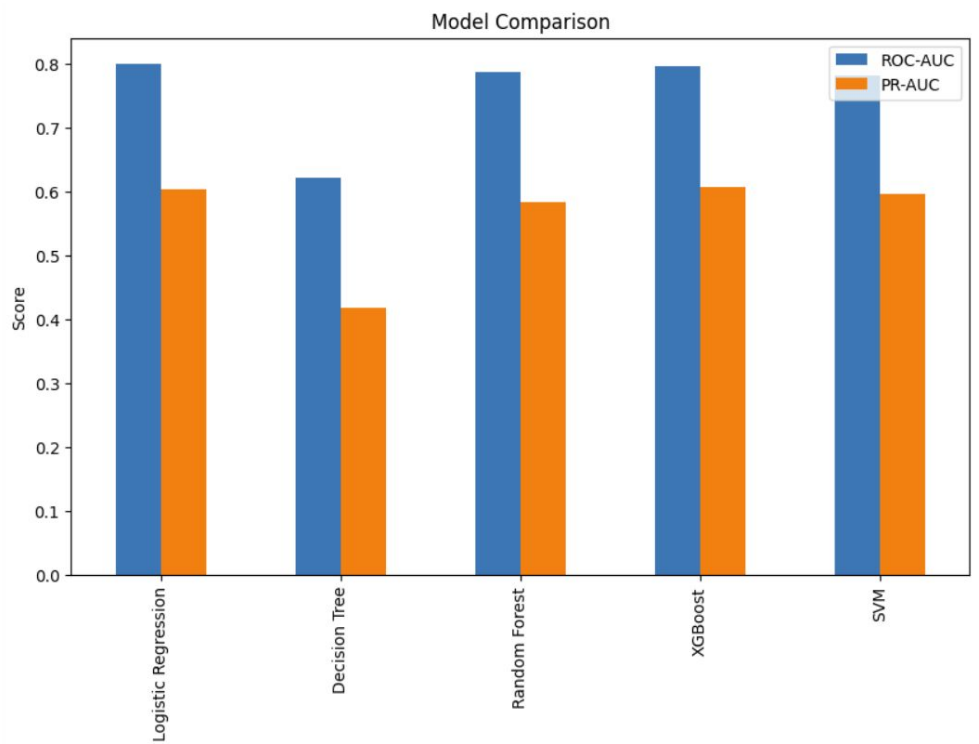
Feature Engineering

Enhancing Model Performance

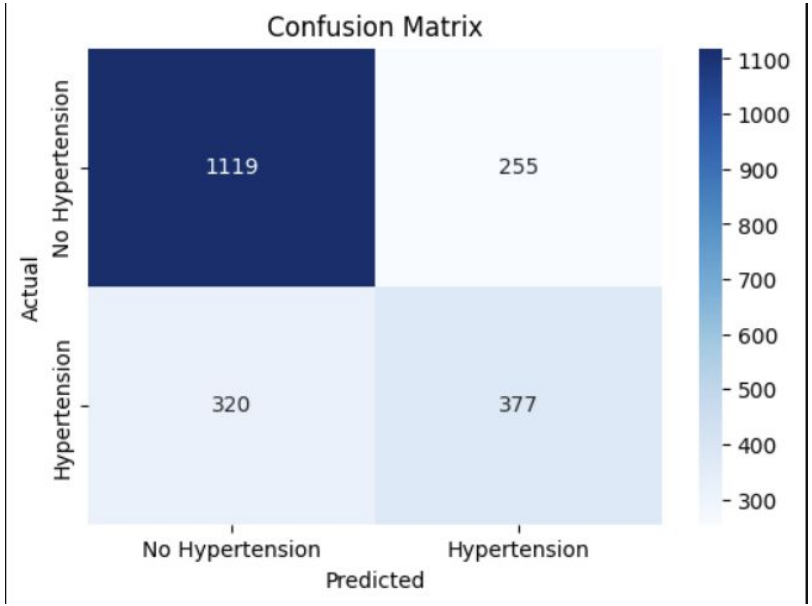
- Created BMI from height and weight measurements
- Calculated Waist-to-Hip Ratio as indicator of central obesity
- Developed interaction features:
    - BMI × Age: Captures compounding effect of weight and aging

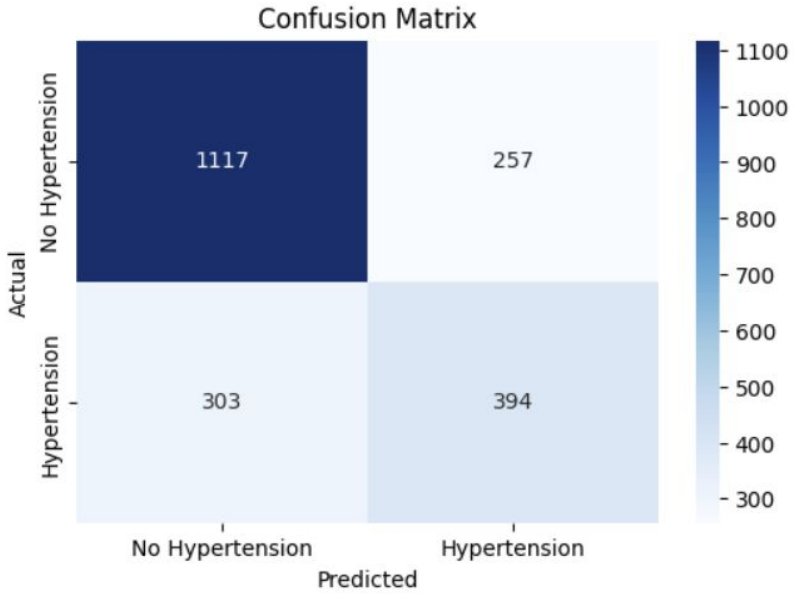
Machine learning model

1. Linear regression
2. Decision Tree
3. RandomForest
4. XGBoost
5. SVM



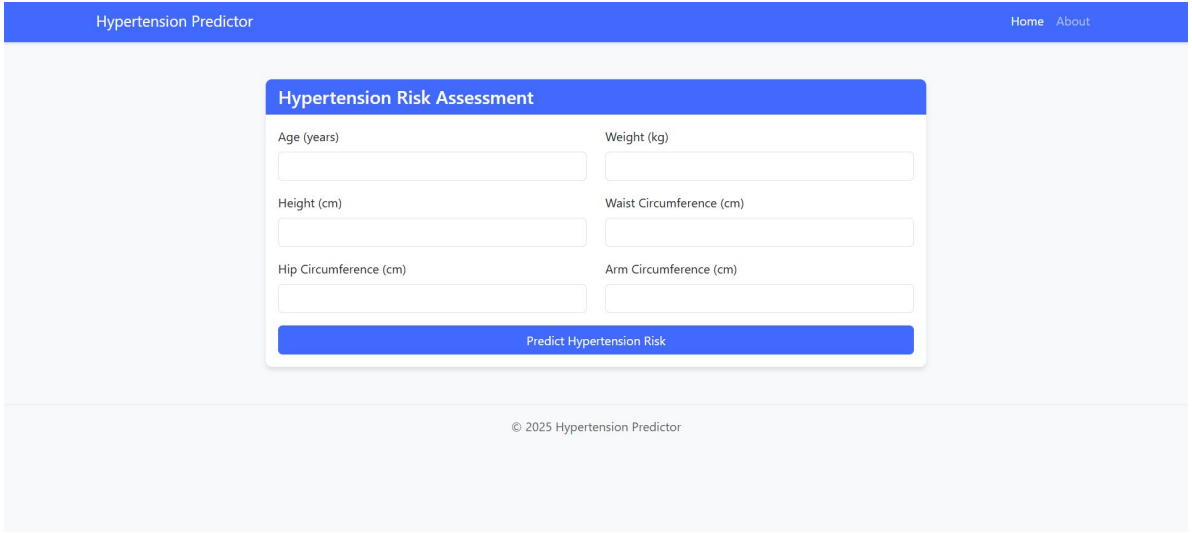
Hyperparameter tuning(For SVM)





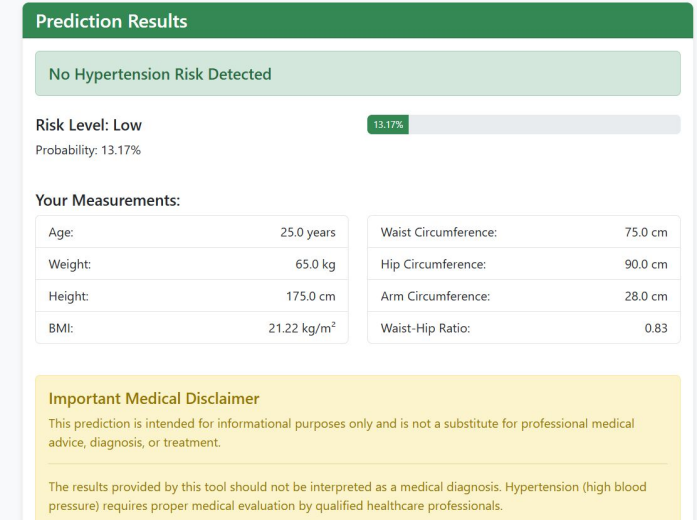
Web application

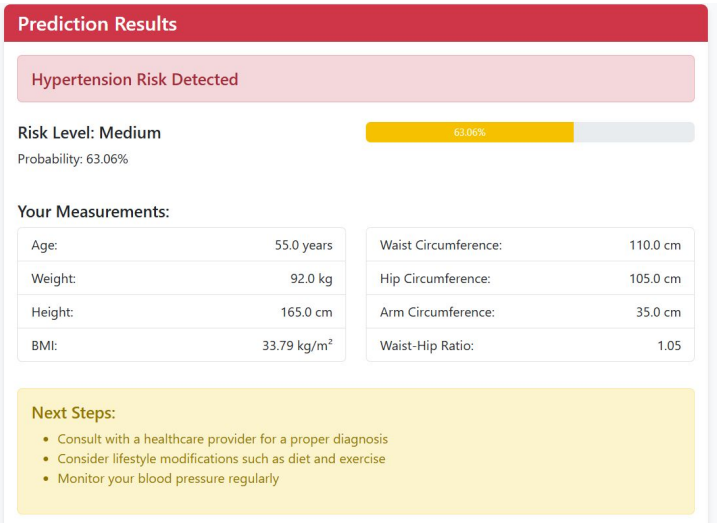
To make the prediction accessible, I developed a **Flask-based web app** that allows users to input their health data and get a risk prediction.



Key Features of the Web App

- **Input Form**: Users can input age, gender, height, weight, waist circumference, etc.
- **Prediction Results**: Users receive a prediction and risk level (Low & High).
- **Visual Feedback**: Displays a **risk level** with color-coded indicators (e.g., red for high risk, yellow for medium, green for low risk).





Live Demo

Challenges & Solutions

- **Challenges**:
  - Handling missing or incomplete data for some users.
  - Ensuring accurate model predictions across different demographics.
  - Deployment and ensuring the app is user-friendly.
- **Solutions**:
  - Used preprocessing techniques like **imputation** and **scaling**.
  - Cross-validation was used to ensure robustness and generalization.
  - Developed a Flask web app for user interaction(currently run locally)

Future Improvement

- **Potential Improvements**:
  - Cloud deployment on Render-Heroku for public access
  - Expanding model to include more factors, like lifestyle data (diet, exercise).
  - Expanding model to include additional health variables (Cholesterol, glucose)
  - Develop mobile application for wider accessibility
  - Integrating real-time health monitoring data (e.g., from wearables) to update predictions.

. **Impact**: I believe this project has the potential to help individuals assess their health risks easily and take preventive measures.

Q&A