

# **Drug Safety Analysis and Recommendations**

## **Project Report:**

**FASNA-PP**

**29/09/24**

## 1. Project Overview

**Project Title:** Drug Safety Analysis

**Prepared by:** Fasna

**Date:** 29/09/24

**Objective:** To analyze drug usage patterns concerning various health indicators, including age, blood pressure, cholesterol levels, and sodium-to-potassium ratios, to make informed recommendations for enhancing drug safety.

## 2. Data Overview

- **Dataset Description:**
  - The dataset consists of patient demographics, including age, gender, blood pressure, cholesterol levels, sodium-to-potassium ratios (Na/K), and the drugs prescribed.
- **Key Variables:**
  - **Age Distribution:** Age of individuals segmented into various groups.
  - **Gender Balance:** Ratio of males to females in the dataset.
  - **Blood Pressure Levels:** Categories including high, normal, and low blood pressure.
  - **Cholesterol Levels:** Ranges categorized into high and normal.
  - **Sodium-to-Potassium Ratio:** A health indicator related to dietary habits.
  - **Drug Usage:** Frequency and types of drugs prescribed.

## 3. Insights

### 3.1 Age Distribution

- The majority of individuals fall within the 40-50 years age group, indicating a middle-aged population focus. This suggests a need for targeted marketing strategies and health interventions aimed at this demographic.

### 3.2 Gender Balance

- The dataset shows equal representation of males and females, ensuring that any gender-specific trends observed will be representative and free from bias.

### **3.3 Blood Pressure Distribution**

- A significant proportion of the population has high blood pressure, highlighting the need for targeted health interventions such as hypertension management programs.

### **3.4 Cholesterol Levels**

- The prevalence of high cholesterol levels indicates a heightened cardiovascular risk, necessitating cholesterol-lowering initiatives and preventive healthcare measures.

### **3.5 Sodium-to-Potassium Ratio (Na/K)**

- A lower Na/K ratio in the majority suggests healthier dietary habits, but higher ratios remain present, indicating a need for dietary interventions focused on sodium intake.

### **3.6 Drug Popularity**

- DrugY is the most prescribed drug, followed by DrugC and DrugX. The popularity of DrugY indicates its importance for treating prevalent conditions, while the lesser-used drugs may need further evaluation for effectiveness.

### **3.7 Age and Health Indicators**

- Older individuals tend to exhibit high blood pressure and high cholesterol, indicating age as a significant factor in these health issues.

## **4. Recommendations**

### **4.1 Targeted Health Campaigns**

- Design age-specific health awareness campaigns targeting middle-aged and older individuals prone to high blood pressure and cholesterol.

## 4.2 Lifestyle Changes

- Promote lifestyle modifications, such as low-sodium diets and exercise programs, to reduce cardiovascular risks.

## 4.3 Workforce and Demographic Planning

- Consider future workforce strategies, focusing on hiring younger employees to balance the age distribution, especially in the pharmacy sector.

## 4.4 Pharmaceutical Market Demand

- Increase production and marketing efforts for DrugY, while evaluating Drugs A and B for underutilization or the need for improvement.

## 4.5 Chronic Disease Management

- Develop targeted health interventions for individuals with high blood pressure and cholesterol, including tailored medication and lifestyle recommendations.

## 4.6 Gender-Specific Insights

- Utilize gender representation in the dataset to create inclusive health campaigns that address the specific needs of both genders.

## 4.7 Nutritional Guidelines

- Focus on non-age-related factors influencing Na/K levels, developing guidelines tailored to dietary habits.

## 5. Conclusion

The Decision Tree and Random Forest models (both untuned and tuned) performed exceptionally well, achieving an  $R^2$  of 0.989 and demonstrating very low error metrics. Given the Random Forest model's slightly more balanced performance and its tendency to avoid overfitting—common in decision trees—it is recommended as the best model for predictive performance. Its strong predictive capability ensures reliable generalization across training and testing datasets.

## 6. Appendices

- **Appendix A:**Data Summary and Statistics
- **Appendix B:**Model Evaluation Metrics
- **Appendix C:**Visualizations of Key Findings (Graphs, Charts)

**THANK YOU**