Overview

This project aims to analyze the effectiveness and prescription patterns of various drugs based on patient demographics and health metrics. The dataset, named `drug200.csv`, contains information on 200 patients, including their age, sex, blood pressure, cholesterol levels, sodium to potassium ratio, and the drug prescribed to them.

Dataset Description

The dataset comprises several columns, each representing different attributes related to the patients and the drugs prescribed:

- `Age`: The age of the patient in years.
- `Sex`: The sex of the patient (`F` for female, `M` for male).
- `BP`: The blood pressure level of the patient (`HIGH`, `LOW`, or `NORMAL`).
- `Cholesterol`: The cholesterol level of the patient (`HIGH` or `NORMAL`).
- `Na_to_K`: The sodium to potassium ratio in the patient's blood.
- `Drug`: The name of the drug prescribed to the patient (`DrugY`, `drugC`, `drugX`, and others).

Methodology

The project was conducted through the following steps:

Data Preprocessing: Cleaned the data by checking for null values, duplicate values and encoding categorical variables.

Exploratory Data Analysis (EDA): Analyzed the distribution of patient characteristics and prescriptions to identify patterns and trends.

Statistical Analysis: Using statistical tests to understand the significance of relationships between patient characteristics and drug prescriptions.

Predictive Modeling: Developed a model to predict the drug prescription based on patient characteristics.

Insight Generation: Drawing insights from the analysis to inform prescription practices.

Outcomes:

- Understood the factors influencing the prescription of different drugs.
- Analyzed the relationship between patient characteristics (age, sex, BP, cholesterol, Na_to_K ratio) and the drug prescribed.
- Developed a model to predict the drug prescription based on patient characteristics.
- Created a web application using streamlit with the help of model developed to predict the drug prescription based on patient characteristics.