

Heart Disease Prediction Using Logistic Regression

Project Overview

This project uses Logistic Regression to predict the likelihood of heart disease based on medical attributes from patient data. The model is trained and evaluated using a public dataset.

Step 1: Import Libraries

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LogisticRegression

from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
```

Step 2: Load Dataset

```
df = pd.read_csv('heart.csv')

df.head()
```

Step 3: Data Analysis and Visualization

```
df.isnull().sum()

df.describe()

sns.heatmap(df.corr(), annot=True)
```

Step 4: Preprocessing

Heart Disease Prediction Using Logistic Regression

```
X = df.drop('target', axis=1)
```

```
y = df['target']
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Step 5: Train Logistic Regression Model

```
model = LogisticRegression(max_iter=1000)
```

```
model.fit(X_train, y_train)
```

Step 6: Model Evaluation

```
y_pred = model.predict(X_test)
```

```
accuracy_score(y_test, y_pred)
```

```
confusion_matrix(y_test, y_pred)
```

```
classification_report(y_test, y_pred)
```

Step 7: Predict for New Patient

```
patient = [[63, 1, 3, 145, 233, 1, 0, 150, 0, 2.3, 0, 0, 1]]
```

```
prediction = model.predict(patient)
```

Conclusion

The model predicts heart disease with approximately 80-85% accuracy. Logistic Regression is effective for early detection and can be deployed in health monitoring applications.