Heart Health Prediction

Adia Ainsworth

Table of contents

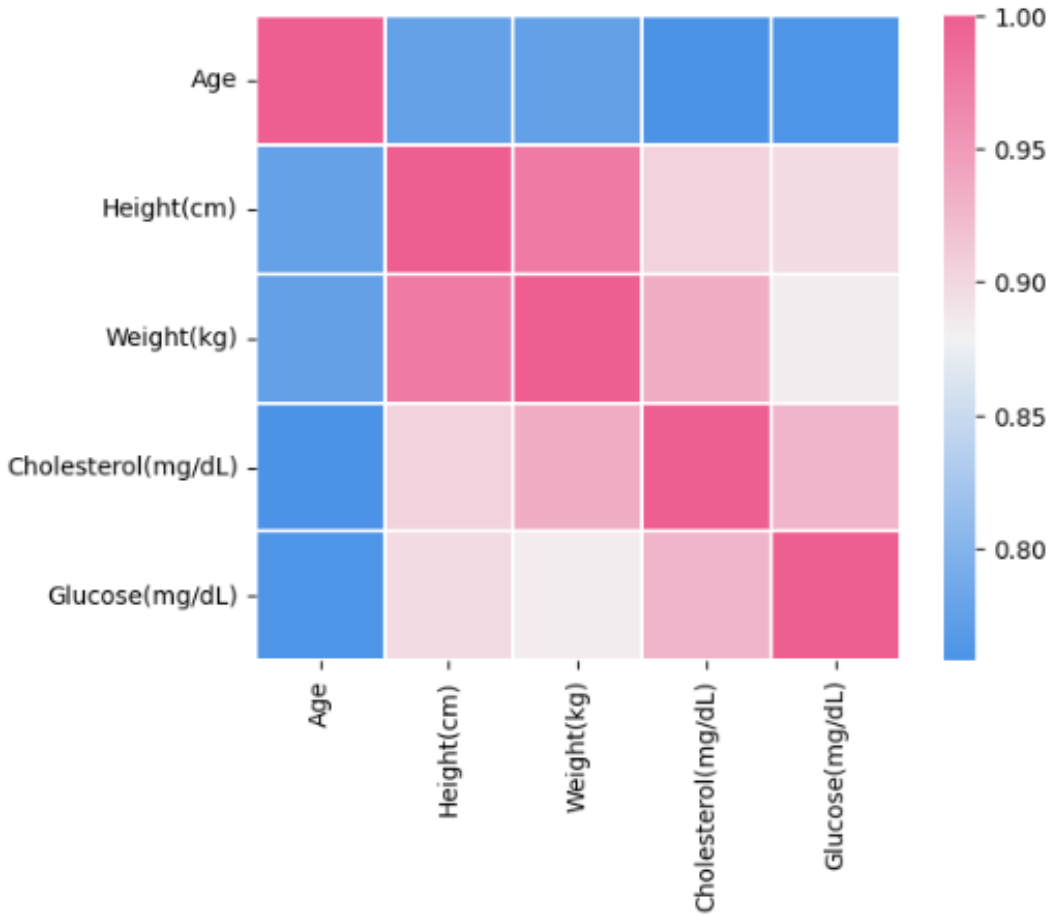
Naive Bayes Classifier

- Bayes' Theorem: $P(B|A) = \frac{P(B|A)P(A)}{P(B)}$
- Real world applications
- Naive Bayes
  - Independent features
  - Equal contribution

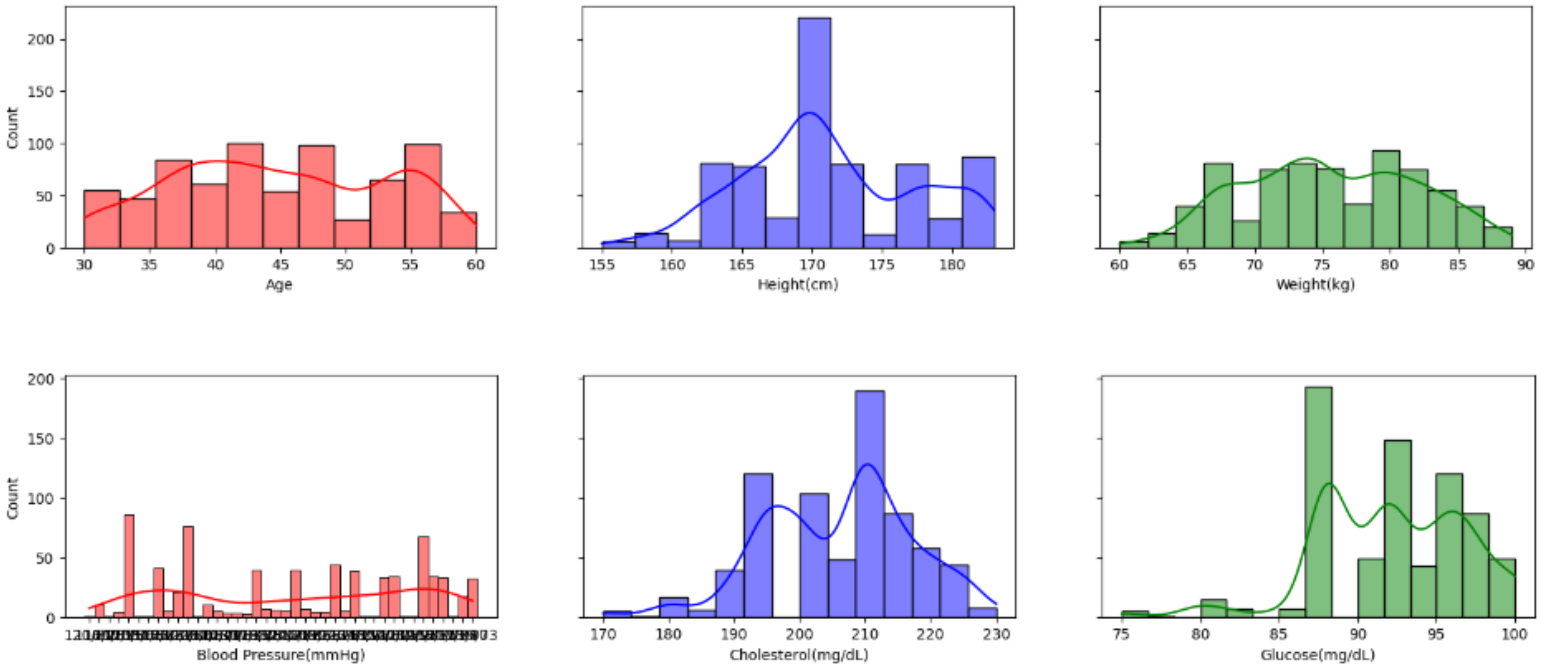
Implementation

- Naive Bayes classifier for diagnosis prediction
- Heart health data set
- Features
  - ID, Name, Age
  - Height(cm), Weight(kg)
  - Blood pressure(mmHg), Cholesterol(mg/dL), Glucose(mg/dL)
  - Exercise(hours/week)
  - Heart Attack

Implementation



Implementation



Implementation

- Gaussian model
- $\begin{bmatrix} 57 & 0 \\ 1 & 87 \end{bmatrix}$
- 0.9943

Implementation

	catGlucose	catChol	Heart Attack
0	1	1	0
1	0	0	0
2	2	2	1
3	1	0	0
4	2	1	1
5	0	0	0
6	2	2	1
7	1	1	0
8	2	1	1
9	1	1	0

Implementation

- Categorical features model
- $\begin{bmatrix} 55 & 2 \\ 4 & 84 \end{bmatrix}$
- 0.9655

Summary

- Naive Bayes Classifier
  - Gaussian model
  - Categorical features model
- Next steps
  - Larger dataset
  - More complex models

References

- “Bayes Classifier and Naive Bayes Tutorial (Using the Mnist Dataset)” Lazy Programmer, lazyprogrammer.me/bayes-classifier-and-naive-bayes-tutorial-using/.
- “Naïve Bayes Classifier.” H2O 3.46.0.1 Documentation, docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/naive-bayes.html.
- Kavlakoglu, Eda. “Classifying Data Using Multinomial Naive Bayes Algorithm.” IBM Developer, developer.ibm.com/tutorials/awb-classifying-data-multinomial-naive-bayes-algorithm/.
- https://www.youtube.com/watch?v=3I8oX3OUL6I
- https://www.kaggle.com/datasets/mahad049/heart-health-stats-dataset/data