ASSIGNMENT 9 - Exercise 15: Introduction to Machine Learning

Ragunath Gunasekaran

2020-11-01

Footnote

This is a Footnote test.
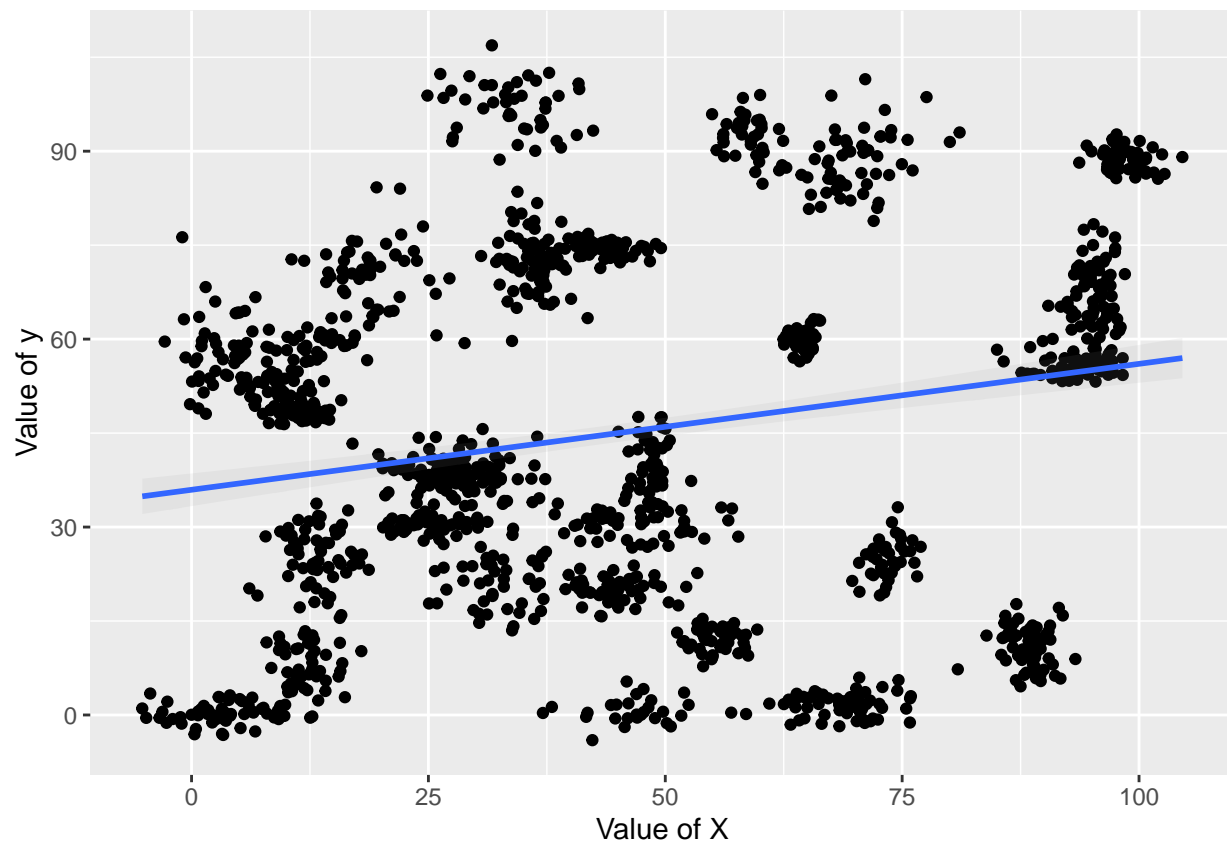
Citations
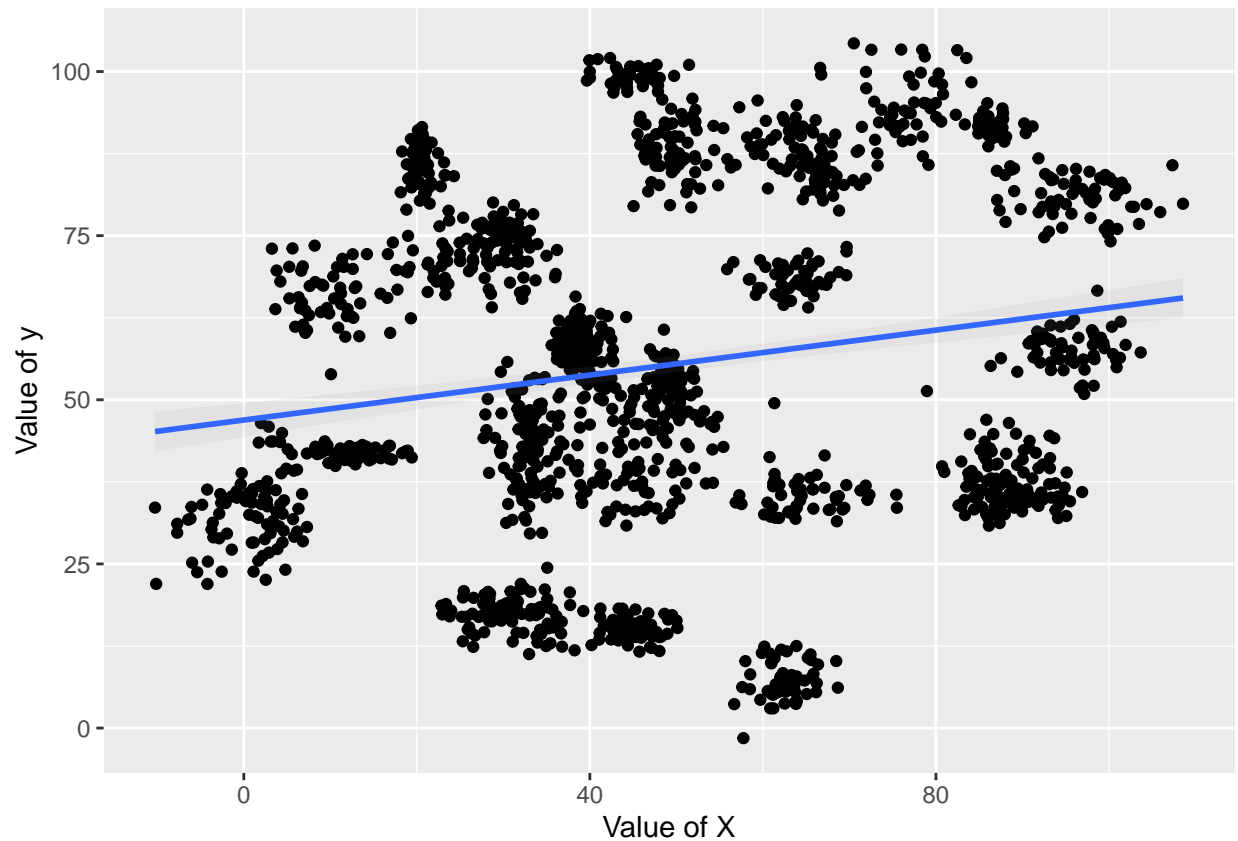
- R for Everyone
- Discovering Statistics Using R

a. Plot the data from each dataset using a scatter plot

Binary classifier Data - Scatter plot displays the relationship between x and y variables

```
## Warning: package 'ggplot2' was built under R version 4.0.2
```

Trinary classifier Data - Scatter plot displays the relationship between x and y variables

Fitting a model is when you use the input data to create a predictive model.Accuracy is simply the percentage of how often the model predicts the correct result. If the model always predicts the correct result, it is 100% accurate. If the model always predicts the incorrect result, it is 0% accurate.

Number of observations for Binary Dataset - Train, Test

```
## [1] 1048
```

```
## [1] 450
```

```
## [1] 1048
```

```
## [1] 450
```

Number of observations for Binary Dataset - Train, Test

```
## [1] 1097
```

```
## [1] 471
```

```
## [1] 1097
```

```
## [1] 471
```

b. Fiting a k nearest neighbors model for each dataset for k=3, k=5, k=10, k=15, k=20, and k=25. Compute the accuracy of the resulting models for each value of k. Plot the results in a graph where the x-axis is the different values of k and the y-axis is the accuracy of the model.

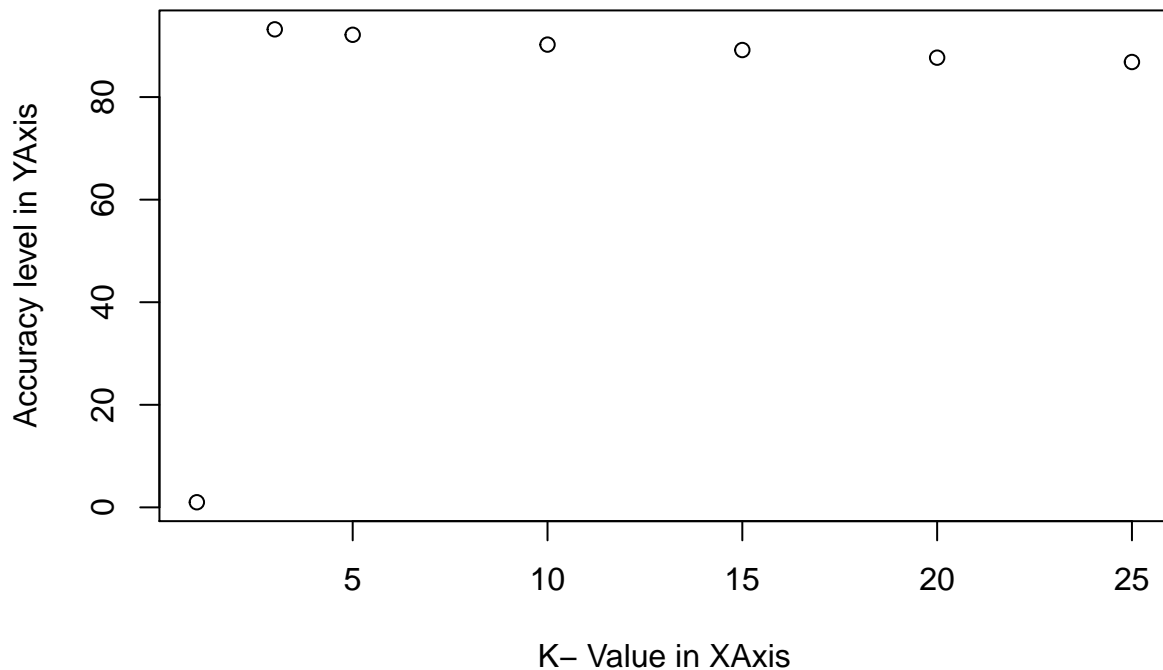Binary Dataset - k nearest neighbors model

```
## Warning: package 'class' was built under R version 4.0.3
```

```
## 3 = 97.55556 5 = 97.33333 10 = 97.55556 15 = 97.55556 20 = 96.66667 25 = 96.44444
```
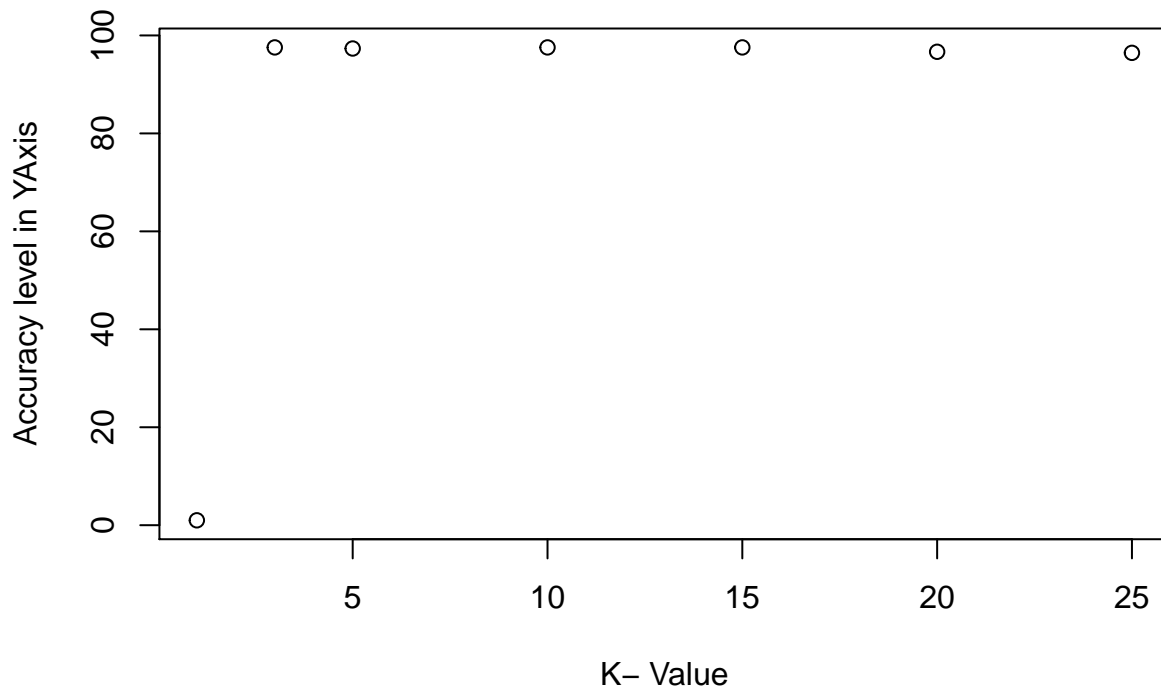
Trinary Dataset - k nearest neighbors model

```
## 3 = 93.20594 5 = 92.14437 10 = 90.23355 15 = 89.17197 20 = 87.68577 25 = 86.83652
```

c. Looking back at the plots of the data, do you think a linear classifier would work well on these datasets?

By looking at the plots, I don't think Linear classifier would work well on these datasets since not seeing the linear combination of the characteristics.

(As per Ref 5, Linear classifier achieves this by making a classification decision based on the value of a linear combination of the characteristics)

References

1. Bernard Marr. (2016). Supervised V Unsupervised Machine Learning – What's The Difference?
2. Bernard Marr. (2016). What Is The Difference Between Artificial Intelligence And Machine Learning?
3. Bernard Marr. (2016). What Is The Difference Between Deep Learning, Machine Learning and AI?
4. http://rstudio-pubs-static.s3.amazonaws.com/515845_ac9b3e6ee0b14ede809e2ce59dd43830.html
5. https://en.wikipedia.org/wiki/Linear_classifier