Q1(1)

Table

Description automatically generated

Q1(2)

small

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | μ(f1) | σ(f1) | μ(f2) | σ(f2) | μ(f3) | σ(f3) | μ(f4) | σ(f4) |
| 0 | -1.87 | 1.88 | -0.99 | 5.4 | 2.15 | 5.26 | -1.25 | 2.07 |
| 1 | 2.28 | 2.02 | 4.25 | 5.13 | 0.80 | 3.24 | -1.14 | 2.13 |
| all | 0.433 | 2.84 | 1.92 | 5.87 | 1.40 | 4.31 | -1.19 | 2.10 |

Q1(3)

The u(f1) and μ(f2) for class 0 is negative, for class 1 is positive.

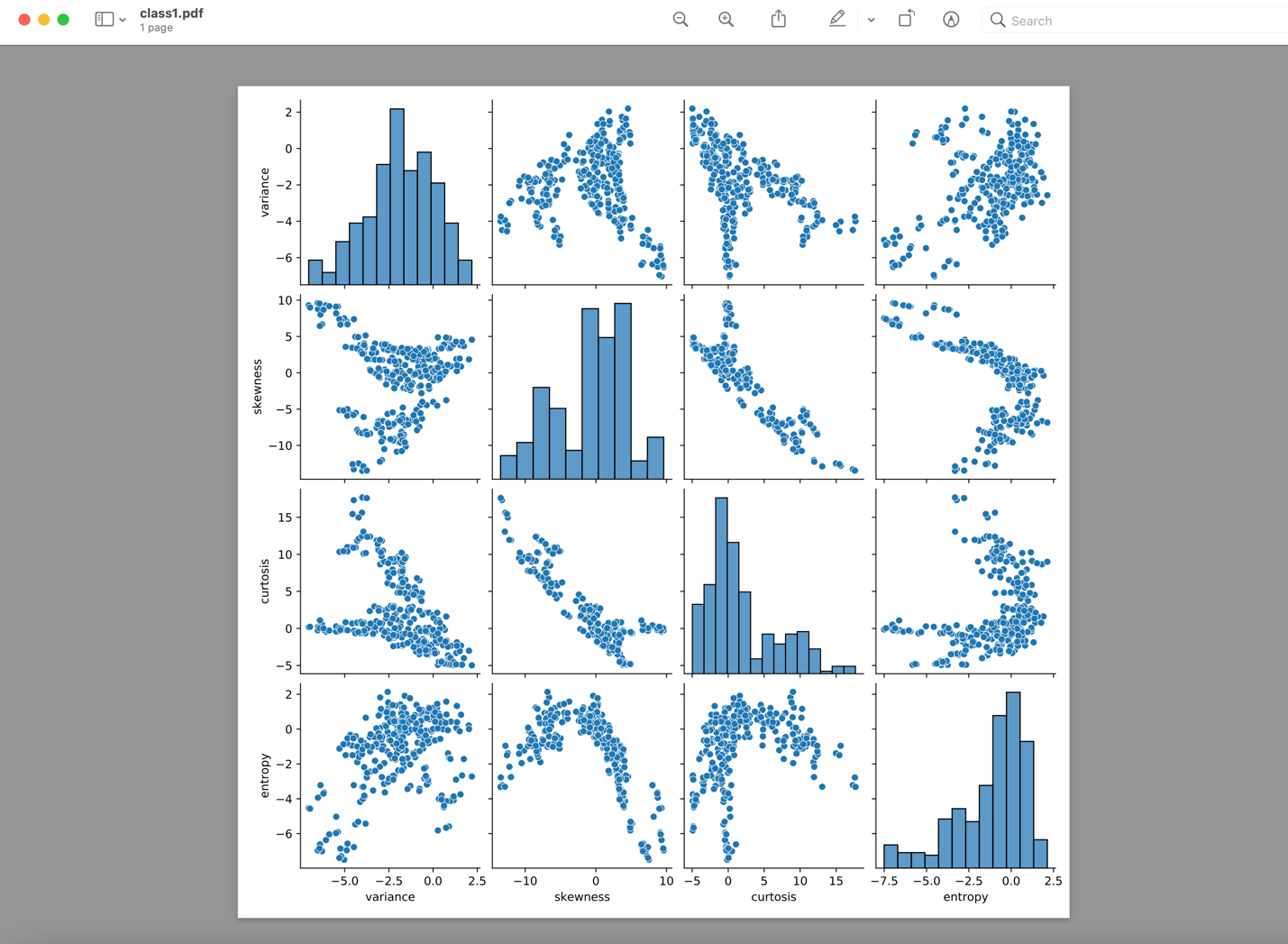
For σ(f2) , μ(f3) and σ(f3) the value for class 0 is larger than class 1.

For μ(f4) and σ(f4), the value for class 1 is larger than class 0.

Q2(1)

Chart, scatter chart

Description automatically generated



Diagram

Description automatically generated with medium confidence

Q2(5)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TP | FP | TN | FN | accuracy | TPR | TNR |
| 381 | 274 | 31 | 0 | 0.6 | 1 | 0.1 |

Text

Description automatically generated

Q2(6)

The accuracy for my simple classifier is 0.6. Since 0.6 >0.5, then, my accuracy is higher than 50%( coin flipping).

Q3(3)

Graphical user interface, chart

Description automatically generated

Text

Description automatically generated with medium confidence

All the accuracy are same as 1.

The optimal value is k = 3.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TP | FP | TN | FN | accuracy | TPR | TNR |
| 381 | 0 | 305 | 0 | 1 | 1 | 1 |

Q3(4)

The k-nn classifier is better than my simple classifier for any of the measures from the previous table except for TP and TPR.

Q3(5)

last 4 digits of my BU ID :6837

class label predicted for 6837 by knn =3 is class 0.

label predicted for 6837 by my simple classifier is good.

Icon

Description automatically generated

Q4(2)

The accuracy does not increase in any of the 4 cases because the accuracy was 1 before.

Q4(3)

The accuracy decreases the most after removing F1. Text

Description automatically generated

Q4(4)

Minimal reduction in accuracy after removal of F4.

Q5(2)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TP | FP | TN | FN | accuracy | TPR | TNR |
| 374 | 0 | 305 | 7 | ﻿0.99 | 0.98 | 1 |

Q5(3)

1. Yes, logistic regression is better than my simple classifier overall. How the TPR in simple classifier is better than logistic regression.

Q5(4)

Knn is better than logistic regression overall.

Q5(5)

last 4 digits of my BU ID :6837

class label predicted for 6837 by knn =3 is class 0.

class label predicted for 6837 by logistic regression

is class 0.

The predicted values are same as knn=3 and logistic regression

.

Icon

Description automatically generated with low confidence

Q6



Text

Description automatically generated

Q6(2)

The accuracy for all 4 features are used is 0.9898. However,

The accuracy for drop f4 is increased. It is 0.996.

Q6(3)

When F1 removed, contributed the most to loss of accuracy

Q6(4)

When F4 removed, contributed the least to loss of accuracy.

Q6(5)

Both Knn and logistic regression are same in relative significance of features.

F1 is the most important feature and F4 is least important.