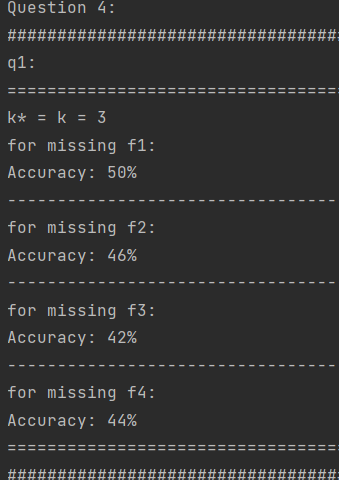
**Assignment 7 banknote**

# **Question 4:**

## **1.take your best value k ∗ . For each of the four features f1, . . . , f4, generate new Xtest and Xtrain and drop that feature from both Xtrain and Xtest. Train your classifier on the ”truncated” Xtrain and predict labels on Xtest using just 3 remaining features. You will repeat this for 4 cases: (1) just f1 is missing, (2) just f2 missing, (3) just f3 missing and (4) just f4 is missing. Compute the accuracy for each of these scenarious.**

| Case | missing f1 | missing f2 | missing f3 | missing f4 |
| --- | --- | --- | --- | --- |
| Accuracy | 50% | 46% | 42% | 44% |



## **2. did accuracy increase in any of the 4 cases compared with accuracy when all 4 features are used?**

Answer:

No, the accuracy decrease in any of the 4 cases compared to that when all features are used.

## **3. which feature, when removed, contributed the most to loss of accuracy?**

Answer:

When the feature f3 is removed, the accuracy is loss the most.

## **4. which feature, when removed, contributed the least to loss of accuracy?**

Answer:

When the feature f1 is removed, the accuracy is loss the least.

# 

# 

# 

# **Question 5:**

## **1. Use 50/50 split to generate new Xtrain and Xtest. Train your logistic regression classifier on Xtrain and compute its accuracy for Xtest**

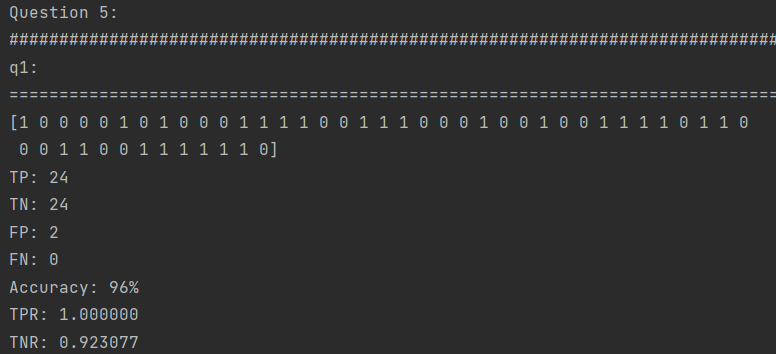
| Accuracy |
| --- |
| 96% |

## 

## 

## **2. summarize your performance measures in the table**

| TP | FP | TN | FN | Accuracy | TPR | TNR |
| --- | --- | --- | --- | --- | --- | --- |
| 24 | 2 | 24 | 0 | 96% | 1 | 0.923077 |



## **3. is your logistic regression better than your simple classifier for any of the measures from the previous table?**

Answer:

Yes, when in my simple classifier, the accuracy is 80%, but in my logistic regression, the accuracy is 96%, which is higher than the one of the simple classifier

## **4. is your logistic regression better than your k-NN classifier (using the best k ∗ ) for any of the measures from the previous table?**

Answer:

Yes, when in my k-NN classifier, the accuracy is about 90% when using the best k\* =3, but in my logistic regression, the accuracy is 96%, which is higher than the one of the k-NN classifier

## **5. consider a bill x that contains the last 4 digits of your BUID as feature values. What is the class label predicted for this bill x by logistic regression? Is it the same label as predicted by k-NN?**

Answer:

The label predicted by logistic regression is 0, which is the same one as predicted by k-NN

# 

# 

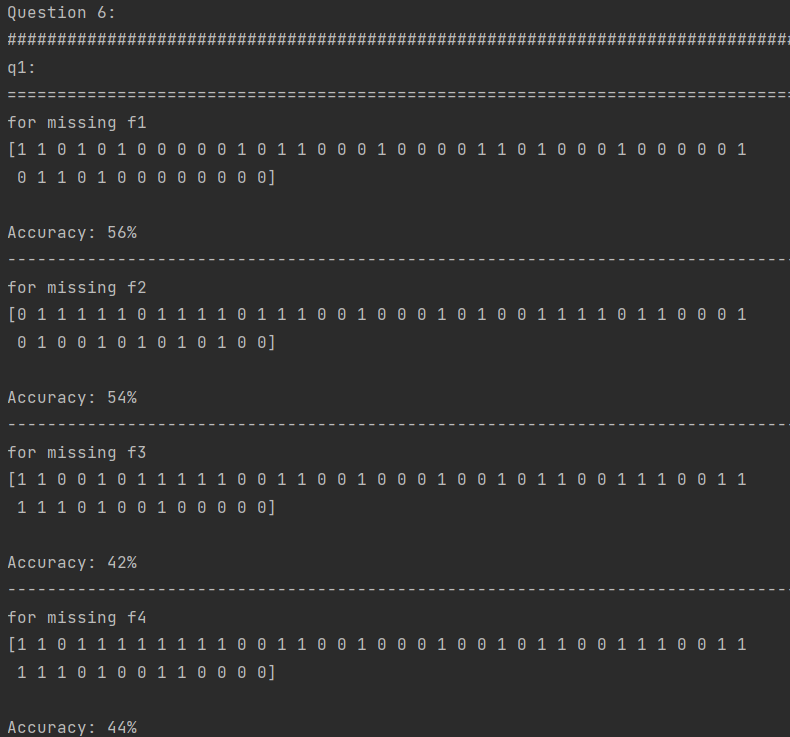
# 

# 

# **Question 6:**

## **1. For each of the four features f1, . . . , f4, generate new Xtrain and Xtest and drop that feature from both Xtrain and Xtest. Train your logistic regression classifier on the ”truncated” Xtrain and predict labels on ”truncated” Xtest using just 3 remaining features. You will repeat this for 4 cases: (1) just f1 is missing, (2) just f2 missing, (3) just f3 missing and (4) just f4 is missing. Compute the accuracy for each of these scenarious.**

| Case | missing f1 | missing f2 | missing f3 | missing f4 |
| --- | --- | --- | --- | --- |
| Accuracy | 56% | 52% | 42% | 44% |



## **2. did accuracy increase in any of the 4 cases compared with accuracy when all 4 features are used?**

Answer:

No, the accuracy decrease in any of the 4 cases compared to that when all features are used.

## **3. which feature, when removed, contributed the most to loss of accuracy?**

Answer:

When the feature f3 is removed, the accuracy is loss the most.

## **4. which feature, when removed, contributed the least to loss of accuracy?**

Answer:

When the feature f1 is removed, the accuracy is loss the least.

## **5. is relative significance of features the same as you obtained using k-NN?**

Answer:

Yes