Q II

II) Understand the working of a simple genetic algorithm involving operators of selection, cross-over, mutation. Apply these operators to an optimisation function such as max f(x)= x3-2x2+x within a range of (0,31).

(Refer David E. Goldberg material for basics of GA)

Maximising f is same as minimising 1/f = F

Minimising F,

import numpy as np

from scipy import optimize

# Function

def F(x):

    v = ((x\*\*3) -2\*(x\*\*2) + x)

    if v == 0:

        return 1

    return (1/((x\*\*3) -2\*(x\*\*2) + x))

# Genetic Optimisation

print(optimize.minimize(F, x0=0))

print(optimize.minimize(F, x0=0, method="L-BFGS-B"))

quit()

Output:

fun: array([-3782864.62824557])

hess\_inv: array([[-2.91991532e-20]])

jac: array([-1.51649125e+13])

message: 'Desired error not necessarily achieved due to precision loss.'

nfev: 192

nit: 2

njev: 60

status: 2

success: False

x: array([-2.6434979e-07])

fun: array([-1.2256126e+08])

hess\_inv: <1x1 LbfgsInvHessProduct with dtype=float64>

jac: array([6.65798886e+16])

message: b'ABNORMAL\_TERMINATION\_IN\_LNSRCH'

nfev: 136

nit: 2

status: 2

success: False

x: array([-8.15918494e-09])

Q III

III) Understand the working of Bucket Brigade Classifier[BBC] (discussed in Goldberg)

Q IV

IV) Explore the scope of applying the bucket brigade classifier in the context of decision tree induction. This question doesn't require the trace. You will have to discuss how the optimization, quicker convergence supported by a BBC can be useful for decision tree induction. (Not mandatory)

Q V

V) Understand confusion matrix and various performance measures associated with classification tasks such as accuracy, sensitivity, specificity, precision, recall, F1 score etc. Consider a sample binary classifier results (TP, FP, TN, FN) and compute various measures.

Confusion Matrix

A confusion matrix is a summary of predicted results on a classification problem

Terms,

1. Positive (P): outcome/observation is positive (Eg. it is a dog)
2. Negative (N): outcome/observation is not positive (Eg. it is not a dog)
3. True Positive (TP): prediction and observation are positive
4. False Positive (FP): prediction is positive, but true observation is negative
5. False Negative (FN): prediction is negative, but true observation is positive
6. True Negative (TN): prediction and observation are not positive

Formulas,

1. Recall = (Also called sensitivity)

High Recall means the recognized class is correct

Low Recall means the recognized class is wrong

1. Precision =

High Precision means if outcome is positive, in reality also it is positive

1. Specificity =

High Specificity means if outcome is negative, in reality also it is negative

1. Accuracy =
2. F1 Score =