Supplementary Notes #05

Data Mining and Data Warehousing

Suggested Solution to Exercises on GA in Data Mining

a. How many bits do you need to encode one classification rule?

   10 bits

b. How many bits do you need to encode one set of classification rules?

   This depends on the number of rules you would like to encode.

   For 3 rules, it is 30 bits.

c. What is the set of rules encoded by the following chromosome?

   101001010110001010100100010110

   The first rule is encoded in: 1010010101

   If Income=L/H and Payment=AMEX and Not Frequent and Not late payment then Good

   The second rule is encoded in: 1000101010

   If Income=L and Payment=Cheque and Frequent and has late payment then Good

   The third rule is encoded in: 0100010110

   If Income=M and Payment=AMEX and Not Frequent and has late payment then Good

d. What is the fitness of the above chromosome?

   – None of the 10 records matches any rule, therefore, the prediction for all records are Bad

   – Hence the classification accuracy is 40%.