

CSE 5693 Machine Learning HW4
Due Apr 9, 6:30pm
Canvas: HW4

1. Written assignment [pdf file]:
 - (a) 9.4
 - (b) 10.1
 - (c) 10.3
 - (d) 10.6
 - (e) From testIrisSelection in the programming assignment, compare the three selection strategies. Plot training and test set accuracy against number of generations and discuss your observations.
 - (f) From testIrisReplacement in the programming assignment, plot test set accuracy against replacement rate (r) and discuss/explain your observations.

2. Programming assignment: Genetic Algorithm
 - (a) allow variable-length individuals, each individual is a rule set similar to Section 9.3.
 - (b) Input parameters include:
 - population size (p)
 - replacement rate (r)
 - mutation rate (m)
 - stopping criterion (e.g. fitness threshold, number of generations)
 - selection strategy (fitness-proportional, tournament, rank)
 - (c) Test your implementation on:
 - i. Tennis dataset (same as HW2 and 3)
 - ii. Iris dataset (same as HW2 and 3)
 - (d) For each of the following experiments, provide a script/program/function (using parameter values you found are appropriate) for running the test:
 - i. testTennis: output the learned rules (in human-readable form similar to HW2), and accuracy on training and test sets.
 - ii. testIris: output the learned rules (in human-readable form similar to HW2), and accuracy on training and test sets.
 - iii. testIrisSelection: vary generation number, output generation number, training set accuracy, and test set accuracy for each of the three selection strategies
 - iv. testIrisReplacement: vary replacement rate r [.1 to .9, .1 increment], output replacement rate and test set accuracy for each of the three selection strategies.
 - (e) The same program should be able to handle the different data sets.
 - (f) Use C (GNU gcc), C++ (GNU g++), Java (Oracle Java), LISP (CLISP), or Python. If you don't have a preference, use Java since it's more portable.
 - (g) Your program should run on code01.fit.edu (linux) *without* non-standard packages/libraries (no additional installation of libraries/packages).
 - (h) Your submission will be evaluated on code01.fit.edu (linux).
 - (i) Submission:
 - i. README.txt: how to compile and run the tests on code01.fit.edu
 - ii. source code