Artificial Intelligence & Adversarial Machine Learning

(Assignment #3)

Gerry Vernon Dozier
Department of Computer Science & Software Engineering
Auburn University
gvdozier@auburn.edu

AUBURN UNIVERSITY
SAMUEL GINN COLLEGE OF ENGINEERING

Assignment #3

Assignment #3

1. **Given our HTML Malware Dataset & the Scikit-Learn Code:**
   - [30pts] Develop a SSGA and SEDA (Simple Estimation of Distribution Algorithm) for performing feature selection in an effort to fill out the following table:

Method	kNN	SVM_L	SVM_R	MLP
Baseline	best (avg)	best (avg)	best (avg)	best (avg)
SSGA	best (avg)	best (avg)	best (avg)	best (avg)
SEDA	best (avg)	best (avg)	best (avg)	best (avg)

   - Your SSGA should use the following control parameters:
     - Population Size = <discover the most effective>
     - Selection = Binary Tournament Selection
     - Crossover = Uniform Crossover (2-parent)
     - Mutation Rate = <discover the most effective>
     - Replacement = Replace the Worst

Assignment #3

Assignment #3 (1. Cont.)

- [30pts] Develop a SSGA and SEDA (Simple Estimation of Distribution Algorithm) for performing feature selection in an effort to fill out the following table:

Method	kNN	SVM_L	SVM_R	MLP
Baseline	best (avg)	best (avg)	best (avg)	best (avg)
SSGA	best (avg)	best (avg)	best (avg)	best (avg)
SEDA	best (avg)	best (avg)	best (avg)	best (avg)

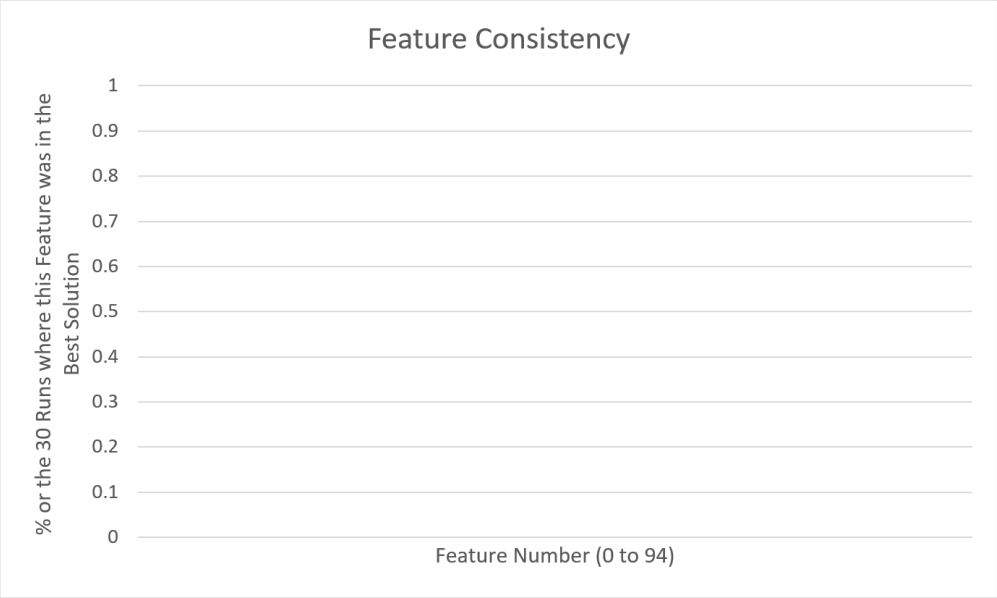
- Your SEDA should use the following control parameters:
  - Population Size = <discover the most effective>
  - Selection = Binary Tournament Selection
  - Crossover = Uniform Crossover((Population_Size/2)–Parent)
  - Mutation Rate = <discover the most effective>
  - Replacement = Replace the Worst
- [7pts] Discuss your results with your partner.

Assignment #3

Assignment #3

2. Given your results:

- [25pts] Provide a Feature Consistency Chart for all SSGA & SEDA Variants:



- [8pts] Discuss your results with your partner.

AUBURN UNIVERSITY
SAMUEL GINN COLLEGE OF ENGINEERING

Assignment #3

(Overview)

Assignment #3 (cont.)

- The points for this assignment will be broken down as follows:
  - Program/Demo 100pts
    - Problem #1 67pts
    - Problem #2 33pts

Have a Great Day!!!

AUBURN UNIVERSITY
SAMUEL GINN COLLEGE OF ENGINEERING