UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Engineering, Built Environment and IT
Department of Computer Science

COS 314

Assignment 3 - Machine Learning

Due 20 May 2024

Question (25 Marks)

Constraints

1. For this assignment you may only use C++ or Java. Furthermore, you may not use external libraries (built in ones are fine). You may be expected to demo your submission (a schedule will be issued).
2. This assignment involves performing classification by implementing the following:
   a) an Artificial Neural Network,
   b) a GP Classification Algorithm.

   The dataset is provided. You are allowed to pre-process the dataset.

Model Details

The models are specified as follows

1. **Artificial Neural Network** (10 marks)
   For this task you must build a neural network model as follows:
   - Have at least 1-hidden layer.
   - The weights must be optimized using back-propagation.
   - Select an activation function for the output layer.(You will need to motivate this choice)
   - You will need to determine a good learning rate. (You will need to motivate this choice)
   - You will need to determine a good stopping condition.(You will need to motivate this choice)
2. **Genetic Programming Classification Algorithms**(10 marks)
   - The GP classification algorithm should evolve arithmetic classifiers.
   - Population size 100
   - Number of generations 50
   - Other parameters are your decision.
   - Care must be taken in choosing tree size as individuals can grow exponentially.

Submission

For each program you are expected to submit the code for two programs.

- Each program must clearly output the training and testing results.
  - For the neural network, the code must display the error after each epoch.
  - For GP the code must display the training accuracy at each evolution. Show the testing accuracy at the end of the run.
- A report in PDF format containing your models description and results should also be submitted. Including the following metrics **Accuracy**, **Specificity**, **Sensitivity** and **F-measure**.(5 marks)
- Pre-processing of the data if any must be reported.
- In order for your results to be replicated a seed value must be used. The seed value needs to reported as this will be used in the demo.
- Results should be reported in a table illustrating the performance of the two models.
- You will be asked questions about your models during the demo.