COSC 4P82 Final Project: Comparing Standard GP with Island Model on Ionosphere Dataset

Grant Nike-6349302 Arin Yaldizciyan-6068431

1. Problem Statement

This project will compare the performance of a regular GP vs an Island Model GP on the Ionosphere dataset. This dataset is a binary classification based on which signals sent by antennae passed through the ionosphere or not. Good results are those which showed some structure in the ionosphere, which means the signals did not make it through. Those results which are labelled as bad are those where the signals could pass through the ionosphere easily. The goal is, given the data from the tests, to identify whether or not the ionosphere had structure in that area.

2. Experiment Description

2.1. Idea

We used two different GP models to evaluate the problem. The first is a standard GP, the second is an Island Model. The details of the island model will be described later on. The goal is to identify specifically which model achieves a higher fitness given the parameters we chose. Choosing parameters is an important process, so for this experiment we decided to find which parameters positively affected the results the most. We started by testing with two different parameter sets. These were used to identify what mutation and crossover rate would be more successful with the models. This would be used to standardize these values between the island model and generic GP. These values would stay the same as the test needs to be fair between both models. We decided to include an extra parameter set where the population size was increased. This was used to determine the effect of the population on the fitness, as well as the rate of growth between models. This means that there are three parameter sets used in total to compare between GP and Island model.

2.2. Island Model Description

The island model differs from a standard Genetic Program in a few important ways, the most important being its population. The population under island model is divided into separate "islands" that go through fitness evaluation, selection, crossover, and mutation independently. Every so often, the islands share some of their individuals with the other islands, the way they choose the individuals and how they are shared are parameters of the island model GP. Theoretically the separate islands can improve diversity in the overall population, and so they can provide improved results on some GP problems.

2.3. Constant GP Parameters

Tournament Size is only relevant to the standard GP model. While Migration size is only relevant to the Island model.

Table 1. Constant GP Parameters for All Experiments

GP Parameter	Parameter Value
# Generations	50
Tournament Size	3
Crossover Type	1-point
Training/Testing Split	0.7

Table 2. Constant Island Model Parameters for All Experiments

Island Model Parameter	Parameter Value
Migration Type	Ring Migration
Selection Type	Select Best Fit
Migration Size	5% of population
Replacement Type	Replace Emigrants

2.4. GP Parameters

These are the three parameter sets which are used to optimize and test the resulting fitness of the models. We will refer to them by Parameter Set 1 through 3.

Table 3. GP Parameter set 1

GP Parameter	Parameter Value
Population Size	50
Crossover Rate	0.5
Mutation Rate	0.2

Table 4. GP Parameter set 2

GP PARAMETER	PARAMETER VALUE
POPULATION SIZE	50
CROSSOVER RATE	0.7
MUTATION RATE	0.2

This final set of parameters is based on the performance of the first two. The only change made to this one is the population size.

Table 5. GP Parameter set 3

GP PARAMETER	PARAMETER VALUE
POPULATION SIZE	150
CROSSOVER RATE	0.7
MUTATION RATE	0.2

2.5. Training and Testing Selection

The Fisher-Yates shuffle algorithm was used to split the training and testing data. This algorithm is used to randomly create permutations of elements in a set. Once the dataset has been randomly shuffled, the data split parameter is used to divide the set into two arrays.

2.6. Fitness Formula

Since this is a binary classification, the fitness is measured by the comparison of the GP individual to the dataset training result. The function counts the number of correct classifications an individual makes against the entire training dataset.

2.7. GP Language

The language used constitutes the following operators: Addition,Subtraction, Multiplication, Division, A minimum function (This returns the smallest value of two given parameters), A maximum function (This returns the largest value of two given parameters), A negative operator (Simply negates the value, ex. 15.3 to -15.3), There is also an ephemeral constant, which is a random number between 0 and 10.

3. Results

Here we will discuss the various results of the parameters in different models. Note that for each test, the model was run 10 times, and the results of those runs were averaged to generate the data we are using for comparison. This is true for every graph or table in the results section.

3.1. Comparison of pure GP between parameters

The following graphs represent the results of the training on the pure GP model. It is important to note that the legend here represents the parameter sets being compared. Test 1 is in reference to parameter set 1, Test 2 is in reference to parameter set 2, and so on. It is clear that of the two initial

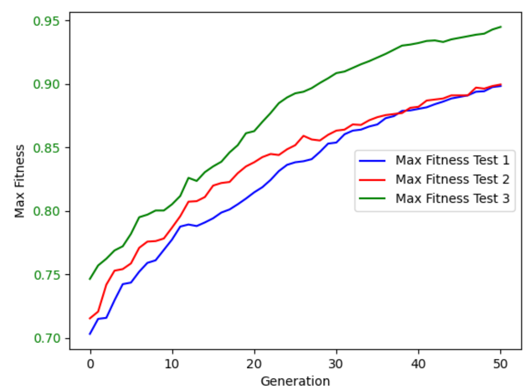

Figure 1. Max fitness per Generation on Pure GA

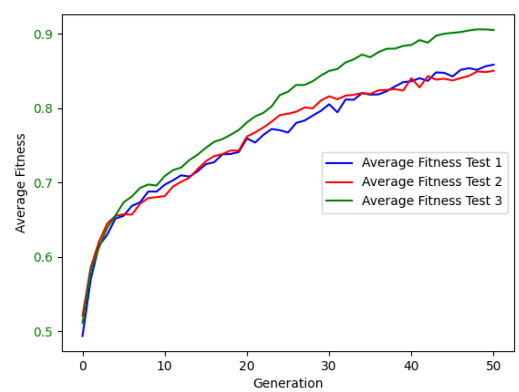

Figure 2. Average fitness per Generation on Pure GA

parameter sets, number 2 performs either better or equal to parameter set 1. Parameter set 3 should outperform both since its population size is three times greater, however its mutation and crossover rate are the same as parameter set 2.

3.2. Comparison of Island Model between parameters

The following graphs represent the result of the training on the Island Model. The legend follows the same rules as the previous set of graphs. They represent parameter sets 1 through 3. It is even more clear in the Island model that parameter set 2 outperforms parameter set 1. The same can be said about parameter 3 as was said in the pure GA. It

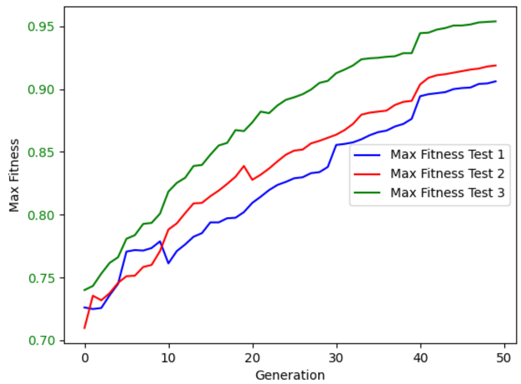

Figure 3. Max fitness per Generation on Island Model

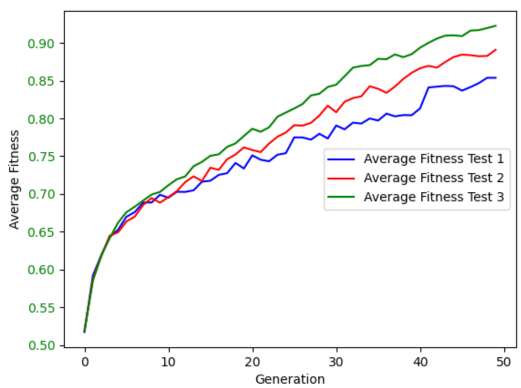

Figure 4. Average fitness per Generation on Island Model

is expected to out perform the rest of the results due to its higher population size.

3.3. Comparison between GP and Island Model

This comparison is the most important for our test. We will see, for each parameter set, which model performs the best. For this comparison, tables will be included which represent the average testing data as well as the graphs for the training data. Please note that blue is Pure GA and red is the Island model.

3.3.1. Parameter Set 1

The following table displays the average fitness on testing data for both models on Parameter set 1.

3.3.2. Parameter Set 2

The following table displays the average fitness on testing data for both models on Parameter set 2.

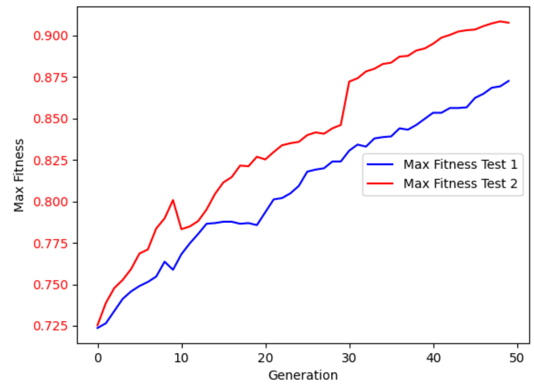

Figure 5. Max fitness per Generation, GP vs Island Model: Parameter set 1

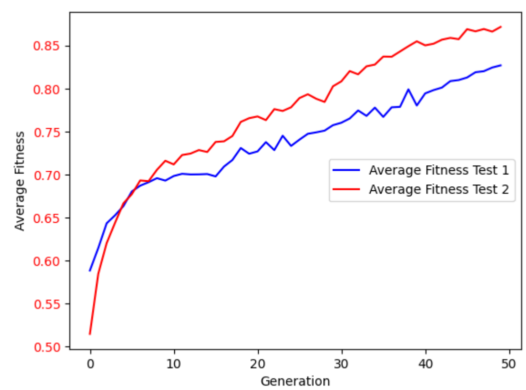

Figure 6. Average fitness per Generation, GP vs Island Model: Parameter set 1

Table 6. Average test fitness Parameter set 1

Model	Average Fitness
Pure GA	81.71%
Island Model	92.60%

Table 7. Average test fitness Parameter set 2

Model	Average Fitness
Pure GP	84.48%
Island Model	93.53%

3.3.3. Parameter set 3

This parameter set will perform better than parameter set 2, given that it is the same with a higher population count. The following table displays the average fitness on testing

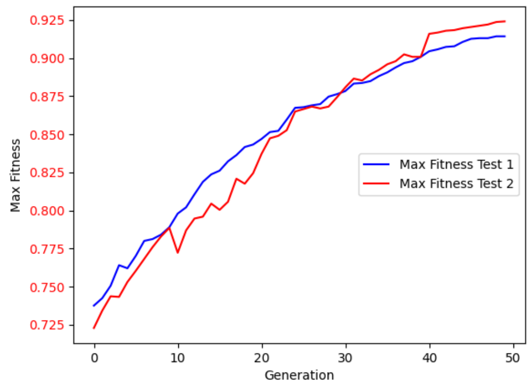

Figure 7. Max fitness per Generation, GP vs Island Model: Parameter set 2

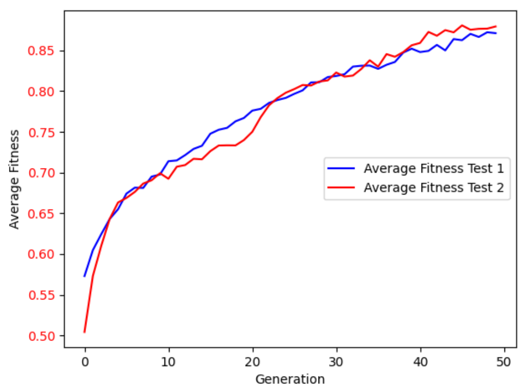

Figure 8. Average fitness per Generation, GP vs Island Model: Parameter set 2

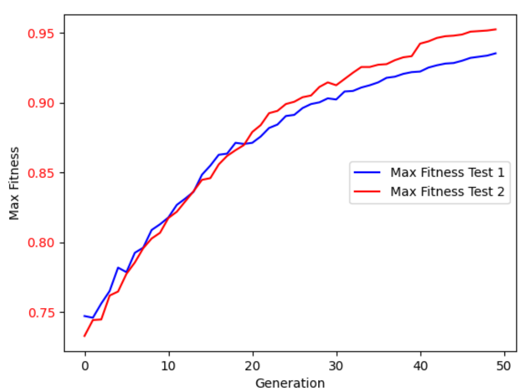

Figure 9. Max fitness per Generation, GA vs Island Model: Parameter set 3

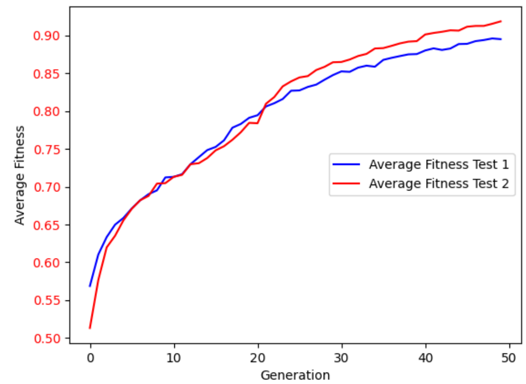

Figure 10. Average fitness per Generation, GA vs Island Model: Parameter set 3

data for both models on Parameter set 3.

Table 8. Average test fitness Parameter set 3

MODEL	AVERAGE FITNESS
PURE GP	88.28%
ISLAND MODEL	96.14%

3.4. Best population solutions

The following are two examples of high performing Individuals from the runs.

3.4.1. SOLUTION 1

add(ARG4, min(mul(min(3.7423677387445933, ARG3), sub(ARG8, ARG23)), add(ARG5, min(add(ARG2, sub(min(min(ARG17, min(neg(max(ARG17, ARG15)), min(ARG11, add(ARG30, min(ARG2, min(ARG15, min(ARG17, add(ARG9, min(ARG15, neg(ARG26))))))))), min(ARG15, min(ARG17, ARG4))), max(ARG15, ARG4))), add(ARG2, sub(min(min(ARG30, min(neg(max(ARG17, ARG15)), min(ARG11, add(ARG30, ARG15)))), sub(ARG18, ARG25)), max(neg(protectedDiv(ARG20, ARG20)), max(min(ARG15, neg(ARG26)), ARG22))))))))

3.4.2. SOLUTION 2

add(add(min(add(min(ARG5, ARG4), min(min(add(ARG4, add(ARG4, ARG25)), add(min(ARG4, ARG2), neg(protectedDiv(ARG17, ARG2)))), ARG2)), ARG4), add(min(ARG6, min(min(add(ARG13, add(add(ARG13, add(min(x, min(min(ARG24, ARG4), ARG4)), add(ARG5, add(ARG13, add(ARG6, ARG13))))), sub(ARG2, min(ARG26, ARG4)))), add(ARG4, neg(protectedDiv(ARG17, x)))), neg(min(ARG28, ARG4)))), *ARG25)), min(x, add(ARG13, add(min(add(add(min(ARG4, add(min(ARG4, ARG2), add(min(ARG16, ARG9), ARG17))), ARG9), add(min(add(protectedDiv(ARG17, ARG2), add(min(add(neg(protectedDiv(ARG17, ARG2)), neg(protectedDiv(ARG17, ARG2))), ARG4), add(max(ARG3, ARG4), add(add(min(ARG4, min(ARG26, ARG1)), ARG11), ARG25)))), ARG16), add(max(ARG4, ARG4), add(add(min(ARG4, ARG2), ARG11), ARG25)))), add(min(add(min(min(ARG6, ARG9), add(ARG4, neg(protectedDiv(ARG17, x)))), ARG9), min(ARG24, ARG4)), ARG4)), ARG4))))*

3.5. Result discussion

3.5.1. Comparison of Pure GP Parameters

The first set of tests shows that the increase of the crossover variable to 0.7 has a positive effect on the overall fitness for both the Maximum population graph, as well as the average. It is however quite close to when the crossover is set to 0.5. The population size increase obviously massively increases the fitness as there are simply many more individuals that are generated per generation.

3.6. Comparison of Island Model Parameters

The island model graphs show a significant increase in performance when crossover is set to 0.7 in comparison to crossover of 0.5. It is around an average difference of 3-6% better fitness over the 10 runs. It is also interesting to note that even though the population size increases by 3 times, the increase in fitness is not much greater in parameter set 3 than in parameter set 2. This is indicative that the island model seems to converge more quickly to a high fitness solution even with much lesser population.

3.7. Comparison of Pure GP to Island Model

Here we will discuss the results of the pure GP in comparison to the Island model. We will analyze specifically by the 3 parameter sets, and their impact on the performance of the population. This comparison is the most important for our results as we are looking to see if the island model performs any better with the same parameters as a pure GP. In particular we will discuss the performance implications of the results,and if Island model is worth the effort, specifically in terms of this dataset.

3.7.1. Parameter Set 1

We can see from the graphs for parameter set 1 that the island model GP consistently maintained a higher max fitness per generation during the training than the standard GP, and it also maintained a higher average fitness per generation than the standard GP after about 10 generations. Table 6 gives us the most important results, where we can see that the island model significantly outperformed the standard GP. The island model GP improved on the average testing fitness over 10 runs by about 11%.

3.7.2. Parameter Set 2

The graphs for parameter set 2 show a much more similar performance on training data between island model GP and standard GP, for both max fitness and average fitness per generation. It is not clear that either model outperformed the other from the graphs alone, although it does seem that island model had a small lead by generation 50 over standard. However, the results on the testing data, shown in table 7, clearly demonstrate that island model outperformed standard GP on the most important metric again, with a roughly 9% performance increase.

3.7.3. Parameter Set 3

The graphs for parameter set 3 show a very similar fitness growth, for both island model and standard GP, to parameter set 2. And the testing data in table 8 also gives a similar result to parameter set 2, with a performance increase of 8% between island model and standard GP.

4. Conclusion

The regular GP seems to perform well on this task regardless of the parameter changes. It is even quite good when compared to the island model. However the island model individually shows that it seems to converge more quickly even with more minimal population sizes. The most important results come from the comparison of island model to standard GP, which show that island model significantly outperformed standard GP that for all tested parameter sets. With the biggest performance difference shown in the comparison on parameter set 1, which had a crossover value of only 50%. This may be due to the ring migration of the island model making up for the lower crossover rate. The island model still did very well against the standard GP on the other two parameter sets, which had higher crossover rates. This may be due to the separate islands increasing the overall diversity, allowing the island model GP to avoid local solutions better than standard GP. Most importantly, we can see that with an increase in population size of 100 between parameter set 2 and 3, the difference in island model and standard GP performance only goes down from 9% to 8%. Overall, we can conclude from these experiments that island model outperforms standard GP on the problem of binary classification using the ionosphere data set. The island model is computationally more expensive than the regular GP, however it is possible to parallelize the computation which quite significantly improves the performance. Island model is actually more parallelizable than standard GP, which is already embarrassingly parallel, because each island can easily be assigned to a separate core, which only need to communicate to perform migration. Whereas standard GA can only be parallelized for fitness evaluation and mutation, for which island model can also be parallelized.

5. Future Improvements

Future experiments could improve on the results of this paper by comparing more parameter sets, between standard GP and island model, there are many combinations of parameters that this paper was unable to explore. Another improvement would be to compare island model to standard GP using more than one data set, and also more than one type of problem. The results of this paper can only conclusively tell us about this type of problem and this data set. The island model implementation could also be easily improved by enabling it to run with the ”scoop” parallel library for DEAP, this would significantly improve the runtime of the island model, and allow for more efficient experimentation, especially on larger data sets.

6. Running Instructions

The project folder contains a set of bash shell scripts which can be used to execute the models and plot the results automatically. The only modifications that should be made are the Python command in each shell script. As this is dependant on your system. If you choose to run the python scripts individually, the standard_GA.py will run the Pure GP solution while the island_model.py will run the island model. The outputs will be put into the output folders respectively and labelled by the run number. The plot_data.py program prints the MAX and Average plots for two data points, while the plot_data3point.py prints the graphs for three data points. The average testing results are printed in the average_output.txt file.

7. Bibliography

References

[1] https://www.researchgate.net/figure/Distributed-Island-Model-example_fig4_262176079

[2] https://archive.ics.uci.edu/ml/datasets/ionosphere

[3] https://deap.readthedocs.io/en/master/index.html