AI - Assignment 1

Q1.

State Representation: A boolean list with size equal to the number of subsets. The bit represents whether each subset was taken or not. A '1' means that subset was taken and '0' means the subset was not taken.

Fitness function(state) = {

100 + fraction of untaken subsets * 100 : if the state covers the universe

100 – number of elements not covered : if the state does not cover the universe

}

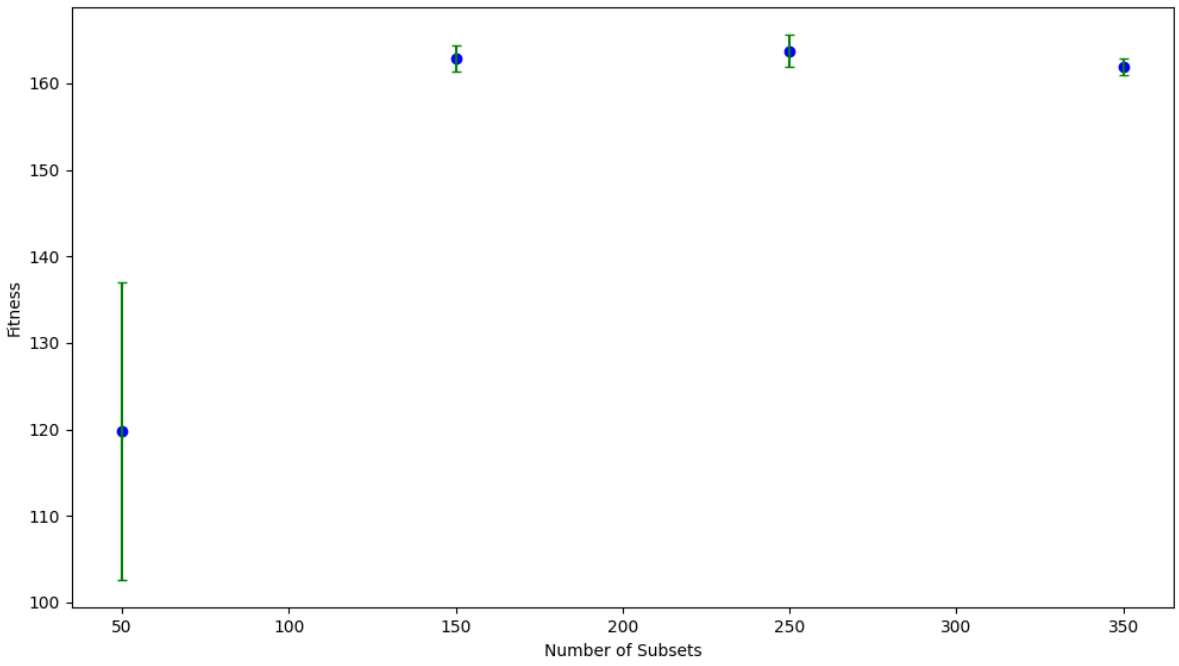
The idea behind this function is that we will prioritize covering the universe first, and then we need to minimize the number of states.

Mutation Rate: 0.15

Elitism: I have implemented elitism by keeping the top-scoring parent from the previous generation.

(a)

- Below is the plot of mean fitness and standard deviation of the genetic algorithm over 50 generations.



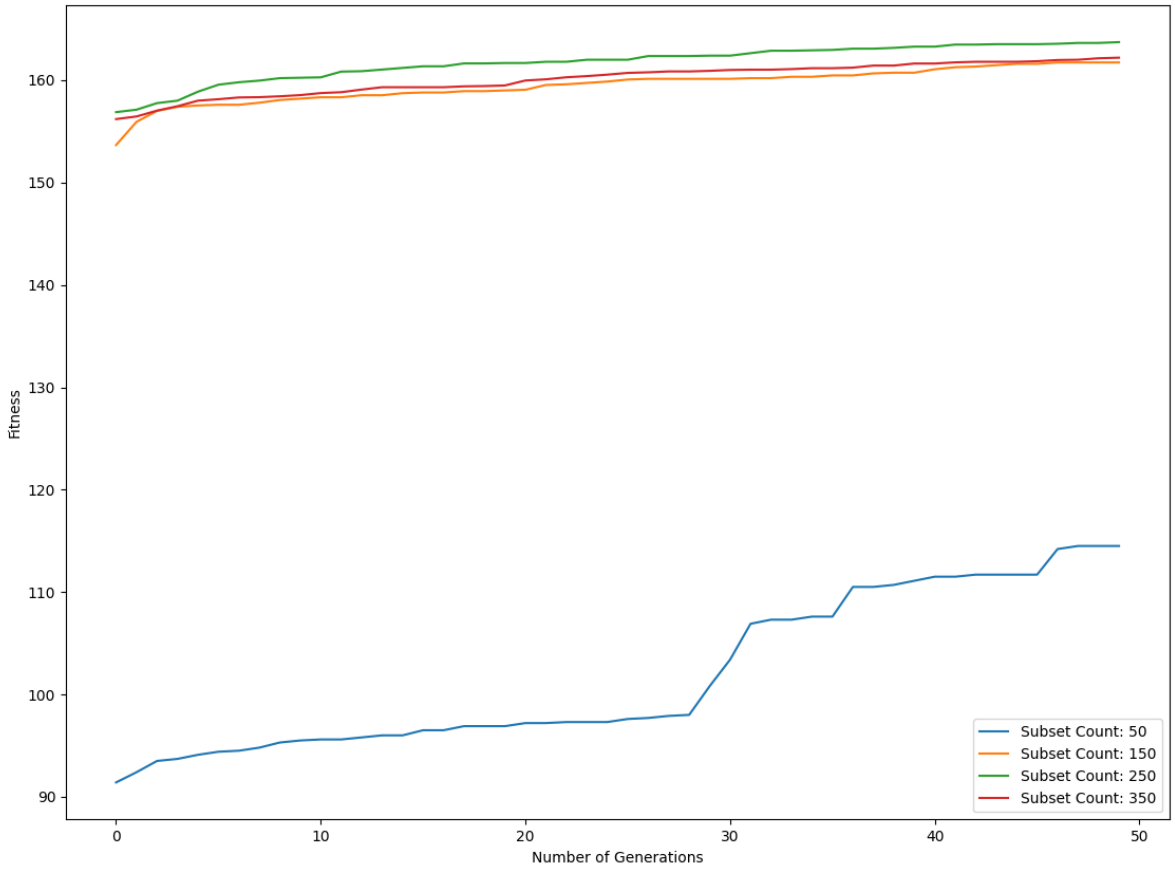
The blue points represent the mean fitness. The green lines represent the standard deviation.

The fitness initially increases with number of subsets and then gradually decreases. It increases as for a small number of subsets, it was difficult to find a subset that covered all the elements.

Also, the fraction of untaken subsets of small number of subsets will be lower as compared to large subset sizes.

This is also true for very large subset sizes. The fraction of untaken subsets will be less.

- Below is the plot of how the best fitness value changes over 50 generations

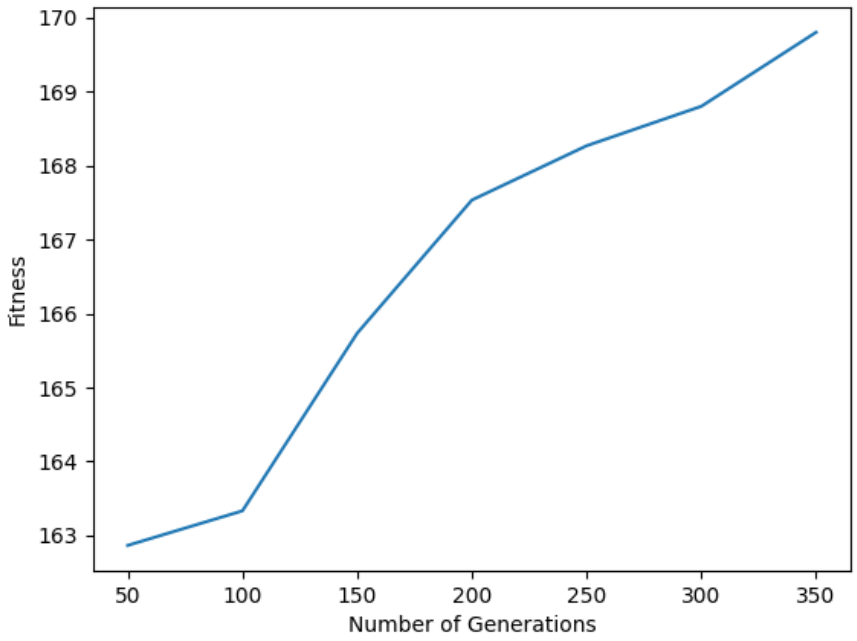


The fitness is non-decreasing with number of generations because in my implementation of the genetic algorithm, we always keep the most fit individual in the next generation.

For high subset counts, the initial fitness is high as it is likely that the whole universe will be covered initially as on average we will select many subsets initially.

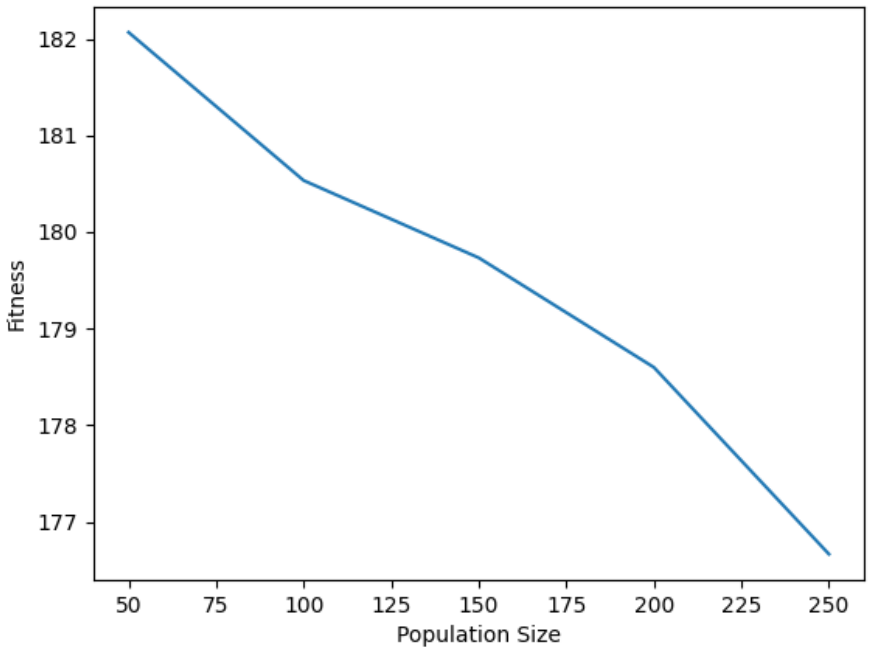
(b) Observations and improvements:

- Number of generations: The fitness increases as we increase the number of generations. This is expected as by reproducing more, we will explore more and find better individuals.



- Initial Population: Having an individual = (1, 1, 1, …, 1) increases the average fitness from 162.66 to 163.33 in the case where population size = 50 and it runs for 50 generations This could be as this individual would be responsible to add 1s in the next generation, resulting in covering of all the sets.

- Population Size: Increasing the population size results in decreased fitness. This could be as with a large population, the less fit individuals will be passed on to the next generations and could reproduce



- Mutation Rate: The fitness decreases as we increase the mutation rate. This occurs as if we mutate a lot, the children lose the good characteristics gained from the previous generations.