Binary GA pipeline is as follows:

1. Generate initial population. (P)
2. Apply crossover to population.
3. Mutate the population. (P)
4. Merge mutated population and initial population, which doubles population)
5. Get the indices of 1’s in parent’s chromosome.
6. Generate new train and test sets derived from the actual sets but with the columns with indices obtained from Step 5.
7. Calculate Euclidean distances -> run predictions -> assess the fitness of the parent and store it.
8. Get the fittest 50% of the population and go back to Step 2.

This pipeline loops until the fittest parent have met or surpassed the expected accuracy.

Initial population can be set via parameters. (Recommended sizes 100 (5-15 generations))

Previously crossovers were done between every pair starting from index 0 and 1.