Report

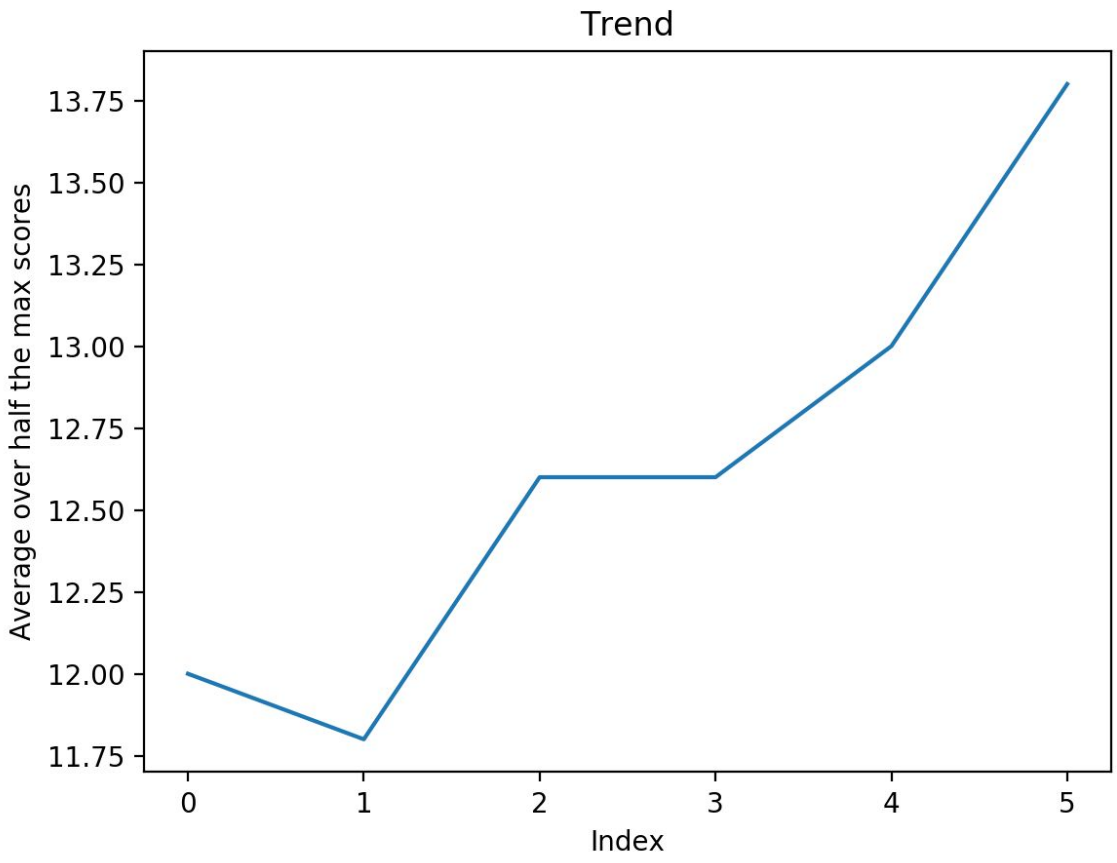
Genetic Algorithm Project

Pratyusha Karanam & Prerit Pathak
Data Science in BioInformatics

University of Florida

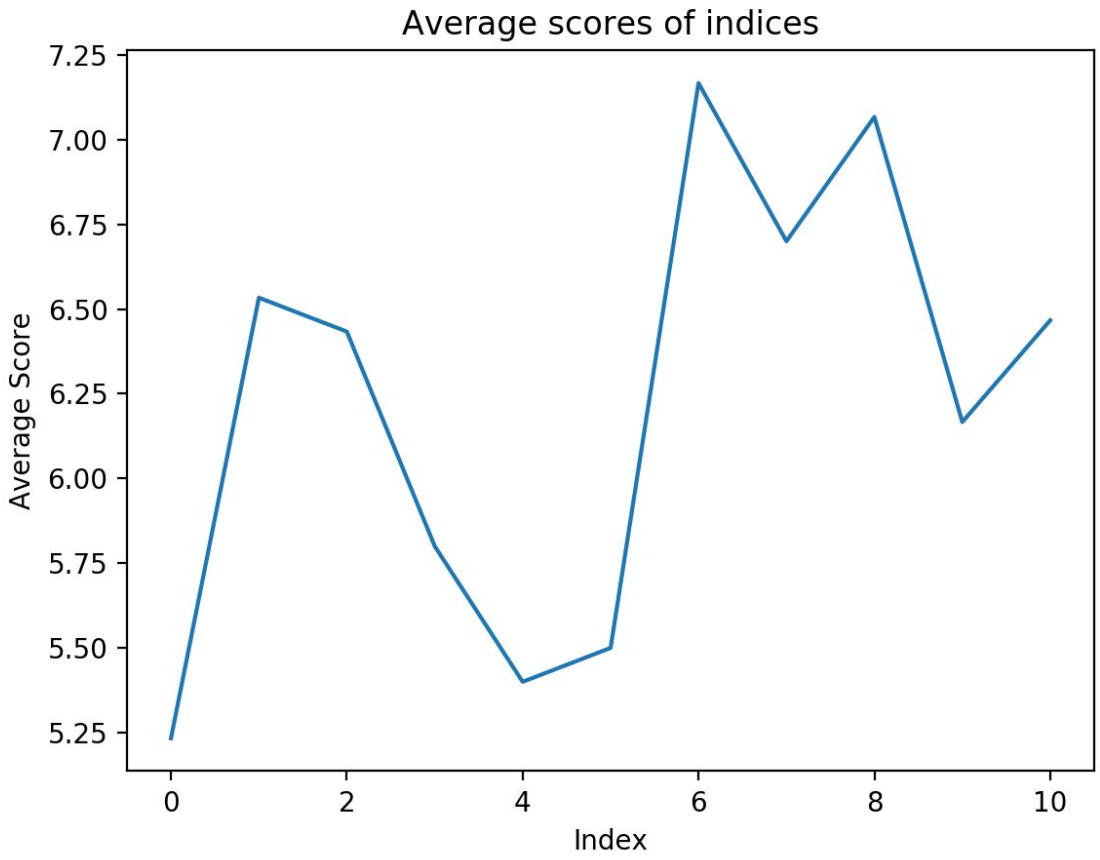
The number of sequences generated is specified as a part of the command while trying the script. When this number is set to be 30, the following graphs are generated:

TREND ANALYSIS



The first graph is built to analyse the trend. This plot shows the value of average of half the max scores at a particular index. The indices are represented on the x axis whereas the average over half the max scores value is represented on the y axis. We observe that this is overall a rising graph.

AVERAGE ANALYSIS



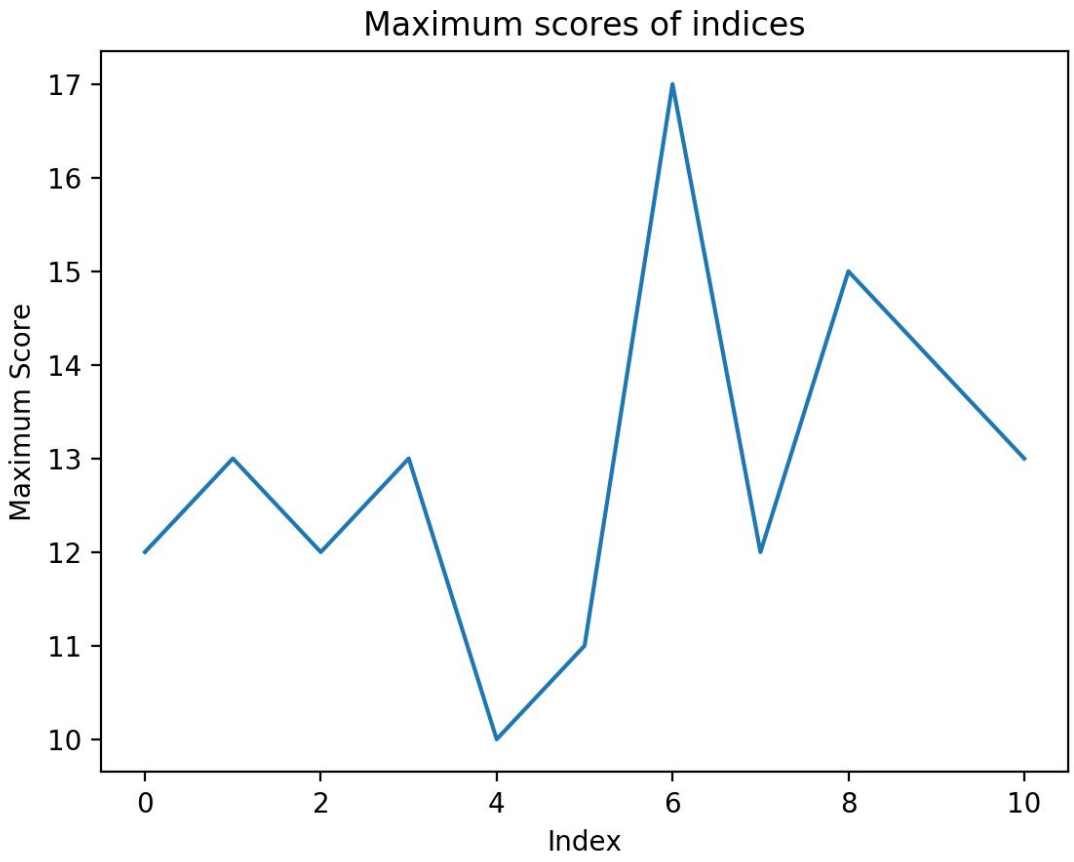
The second graph shows the average score of indices. The indices are represented on the x axis and the average score is represented on the y axis. This graph can be referred to in order to get a basic idea about the average results of our project.

COMPARISON OF MAXIMUM SCORE PER INDEX

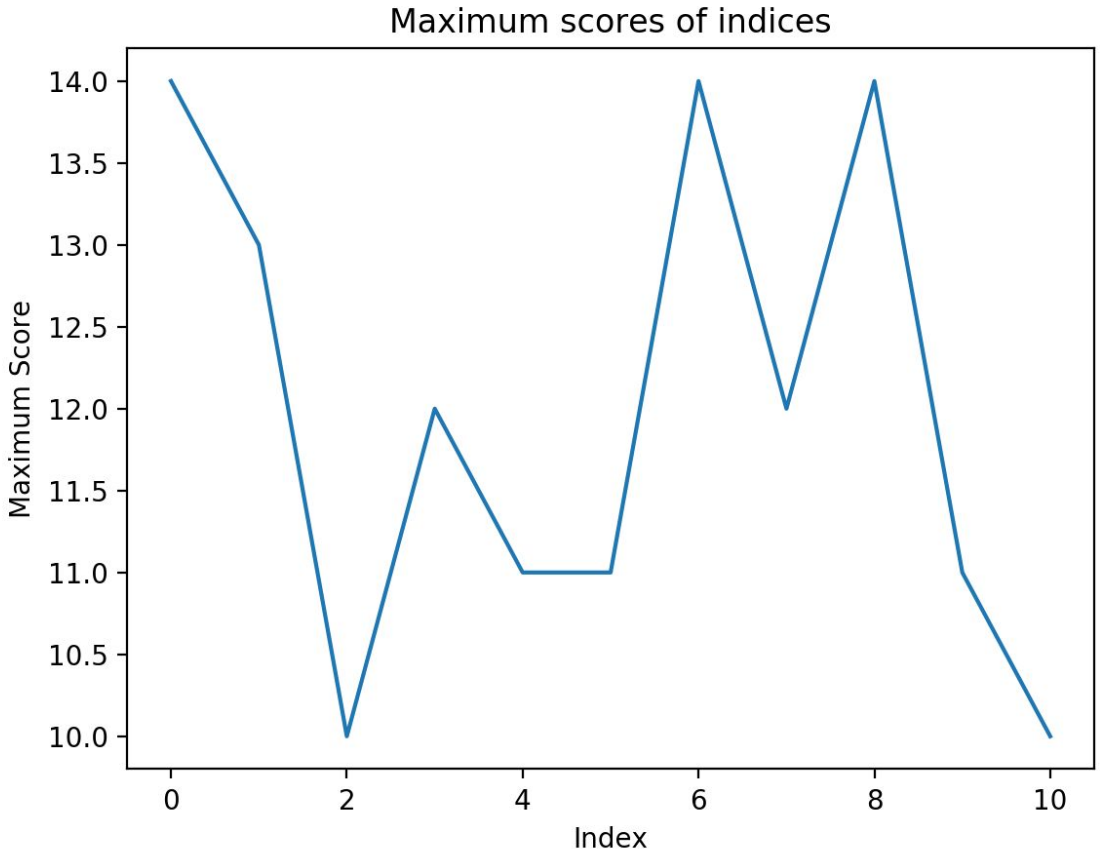
In this part of the visualisation, we generated graphs for various N values; where N = number of sequences generated.

Each graph represents the indices on the x axis and their corresponding max scores on the y axis.

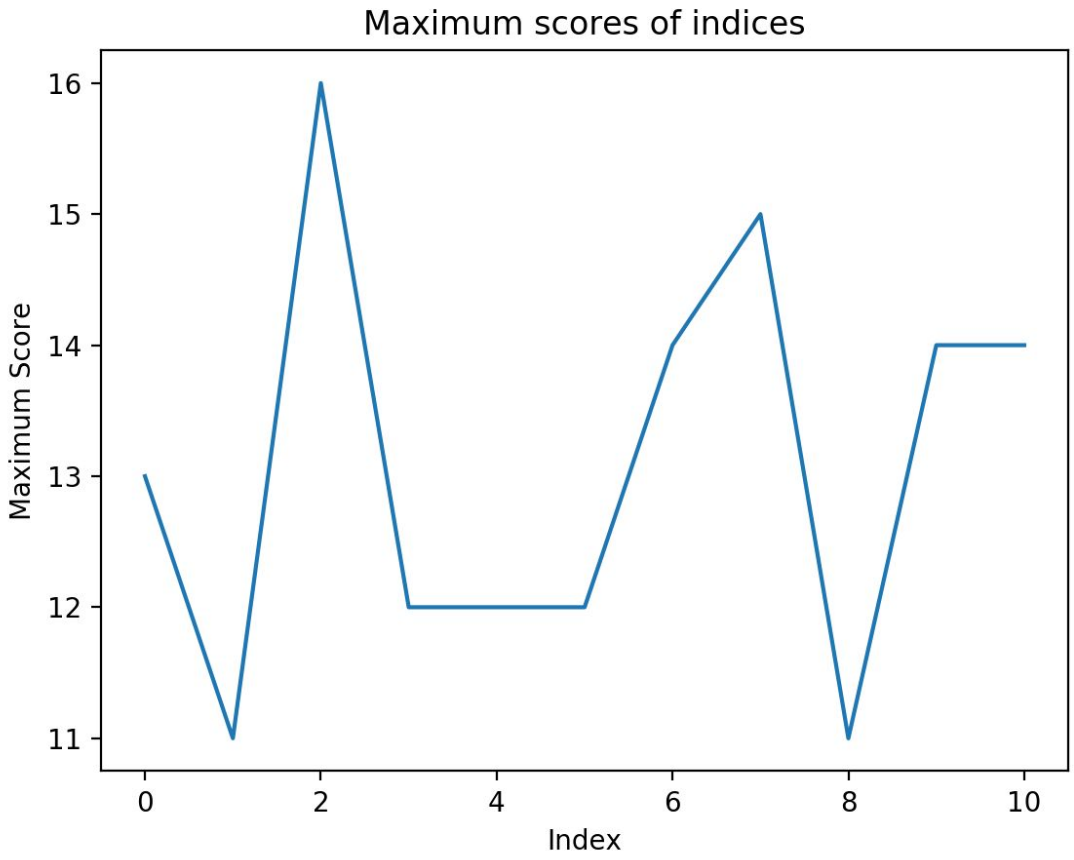
N = 30 & Mutation Probability = 0.01



N = 30 & Mutation Probability = 0.5



N = 50 & Mutation Probability = 0.01



Hence, we can observe the drastic variation in the graphs when we change the value of N.

The following is a screenshot showing the results generated via the terminal:

```
preritpathak Data-Science-Project2-master-3
$ python3 DSproj2.py "sample_input_pdb_test(only H)" 30 10 .01
30 10 0.01
generation 0
[12, 10, 9, 8, 8, 7, 7, 7, 7, 7, 7, 6, 6, 6, 5, 5, 4, 4, 4, 4, 4, 3, 3, 3, 3, 2, 2, 2, 1, 1]
generation  1
[13, 12, 11, 11, 9, 9, 8, 8, 8, 8, 7, 7, 6, 6, 6, 6, 6, 6, 6, 6, 5, 5, 5, 5, 4, 3, 3, 3, 2, 2]
generation  2
[12, 11, 11, 10, 9, 9, 9, 8, 8, 7, 7, 7, 7, 7, 6, 6, 6, 6, 6, 5, 5, 5, 5, 4, 4, 4, 4, 3, 3, 3, 2]
generation  3
[13, 12, 12, 10, 8, 8, 8, 8, 7, 7, 7, 6, 6, 5, 5, 5, 5, 5, 4, 4, 4, 4, 4, 4, 3, 3, 2, 2, 2, 1]
generation  4
[10, 10, 10, 9, 9, 9, 8, 8, 8, 7, 6, 6, 6, 5, 4, 4, 4, 4, 4, 4, 4, 4, 4, 3, 3, 3, 2, 2, 2, 2, 2]
generation  5
[11, 9, 8, 8, 8, 8, 7, 7, 6, 6, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4, 3, 3, 3, 2, 1]
generation  6
[17, 14, 11, 11, 11, 11, 10, 10, 8, 8, 8, 7, 7, 7, 6, 6, 6, 6, 6, 6, 5, 5, 4, 4, 4, 4, 4, 4, 3, 3, 3]
generation  7
[12, 11, 11, 9, 9, 9, 9, 9, 9, 8, 8, 8, 8, 8, 7, 7, 6, 6, 6, 5, 5, 4, 4, 4, 4, 4, 4, 4, 3, 2, 2]
generation  8
[15, 15, 14, 13, 10, 10, 9, 9, 8, 8, 8, 7, 7, 7, 6, 6, 6, 6, 6, 6, 5, 5, 5, 4, 4, 4, 4, 3, 3, 3, 2]
generation  9
[14, 11, 10, 10, 9, 9, 8, 8, 8, 8, 7, 7, 6, 6, 6, 6, 6, 6, 5, 5, 5, 4, 4, 4, 4, 3, 3, 3, 3, 2, 1]
generation  10
[13, 11, 11, 10, 10, 9, 9, 9, 9, 9, 8, 7, 7, 7, 6, 6, 6, 5, 5, 5, 5, 4, 4, 4, 4, 4, 3, 3, 3, 2, 2]
Max scores are: [12, 13, 12, 13, 10, 11, 17, 12, 15, 14, 13]
```