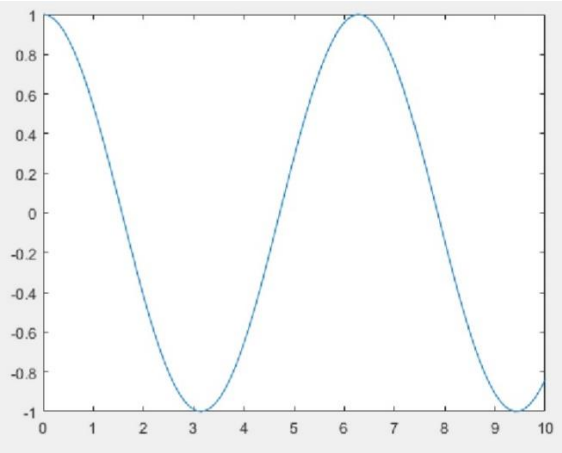
Gunjan K Khut
UID–115813083

Homework 2

Answer 1:

We considered a *'cosine of x'* 1-D function for this case. The 100 evenly space function inputs were varied between (0,10). The overlap area was varied from 1 to 34 with a weight matrix of 35.

The function plot as shown below:



1. **Main Function:**
   - ✓ Distributed the inputs of (0,10) over 100 evenly spaced points
   - ✓ Calculated the output using the function '*cos (x)*'
   - ✓ Used ***initialize*** function to initialize Weights and overlap area.
   - ✓ Assigned 70 random values from **X** to training data and **30** to test data.
   - ✓ Then 'Run' the ***train*** and ***test*** function to plot the values.
2. **Initialize Function:**
   - ✓ Defined the input vector as evenly distributed between the maximum and the minimum of the inputs. This was used to index the inputs to associate them with the weight matrix.
   - ✓ Defined the binary Lookup table to know how many and which associate cells are linked for the given inputs.
3. **Train Function:**
   - ✓ Defined variable indexing, it categorizes the training data so that it can then use the lookup table to find out how many associative cells are linked with the inputs.
   - ✓ Calculated the output by adding the values of all the associated weights.
   - ✓ Calculated the error and updated the weights
   - ✓ Calculated the simultaneous mean absolute error
   - ✓ Return the updated values of weights along with the number of iterations performed.
4. **Test Function:**
   - ✓ Defined similar as ***train*** function, used the testing data instead of the training data.
   - ✓ Calculated the accuracy and returned the number of iterations performed

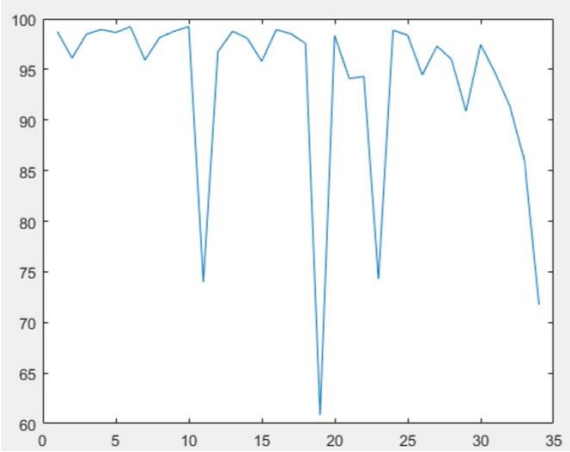
Gunjan K Khut
UID–115813083

Now the function was executed from an overlap area of 1 to an overlap area of 34. The following graphs are obtained:

1. **Accuracy for the discrete function:**
   - We observe that as the overlap area increases the accuracy decreases. This can be explained by the concept of generalization. The more the number of overlaps or more the generalization the difficult it would be to have a higher accuracy as a change in a single weight is affected across all the other weights.
   - The worst accuracy was 60.8522 for the overlapping area 19 of 35 weights and the best accuracy obtained at 99.2264 at 10 overlapping weights.

   Accuracy = [98.7367 96.1160 98.4927 98.9312 98.6461 99.2160 95.9199 98.1481 98.7692 99.2264 74 96.7481 98.7573 98.0913 95.7987 98.9259 98.5203 97.5658 60.8522 98.3462 94.0890 94.3079 74.2656 98.8645 98.3771 94.4426 97.2998 95.9987 90.8668 97.4602 94.6252 91.3684 86 71.7503]

   The plot as shown below:

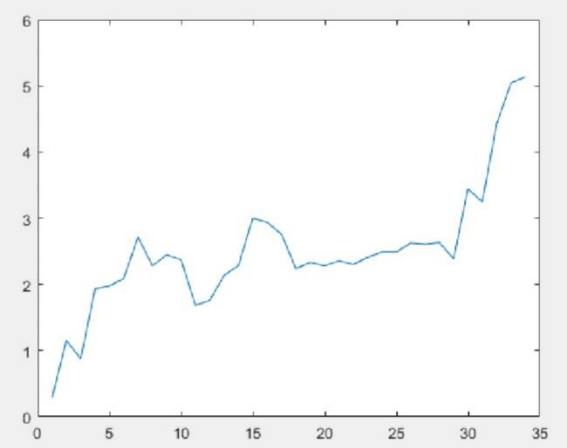


2. **Time to convergence for the discrete function:**
   From the graph, as accuracy of the function decreases, the time for convergence increases.
   - The maximum time taken was 5.1344 sec at the overlap area of 34.
   - The minimum time taken was 0.2834 sec when the overlap was 1.
   - The minimum time was justified by the fact that, as the overlap was 1, i.e. similar to one to one mapping and therefore minimum time.

   $\mathbf{T_{discrete}}$: [0.2834 1.1501 0.8768 1.9326 1.9764 2.0886 2.7154 2.2816 2.4472 2.3701 1.6813 1.7589 2.1373 2.2853 2.9989 2.9381 2.7557 2.2384 2.3318 2.2805 2.3571 2.2998 2.4079 2.4887 2.4886 2.6237 2.6058 2.6317 2.3840 3.4401 3.2451 4.4243 5.0430 5.1344]

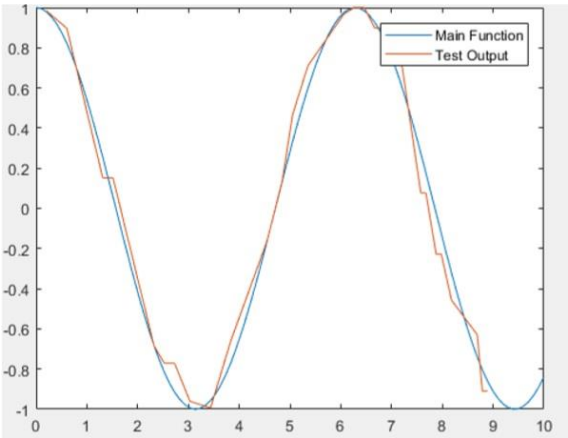
Gunjan K Khut
UID–115813083

The plot as shown below:

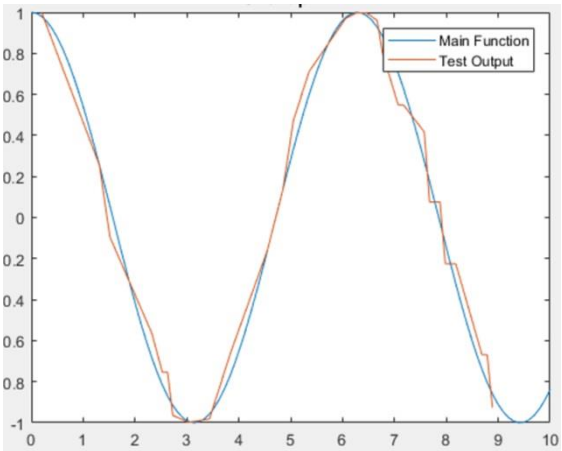


- The plots for the discrete function, the training data and the test data as shown:

The plot for the overlapping value of 19 or the case of maximum accuracy:



The plot for the overlapping value of 2 or in the case of the quickest convergence:

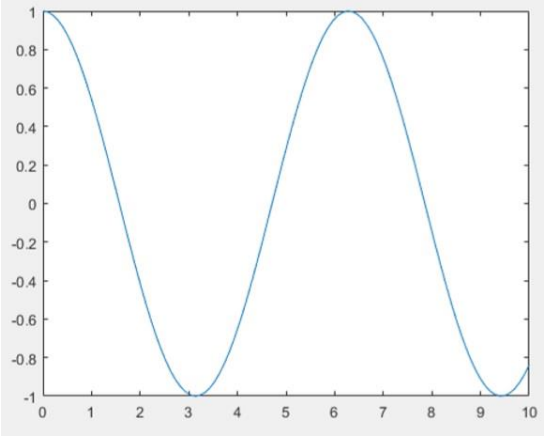


Gunjan K Khut
UID–115813083

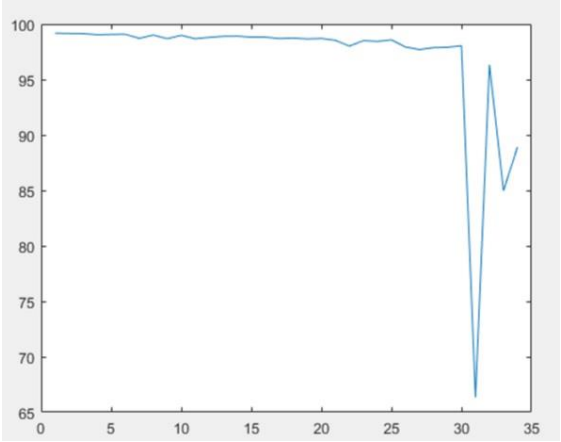
Answer 2:

The same code used, just the continuous values of *cos(x)* were considered instead the discretized values.

The Plot of input function i.e. *cos(x)*:



The Plot of accuracy against overlapping space:



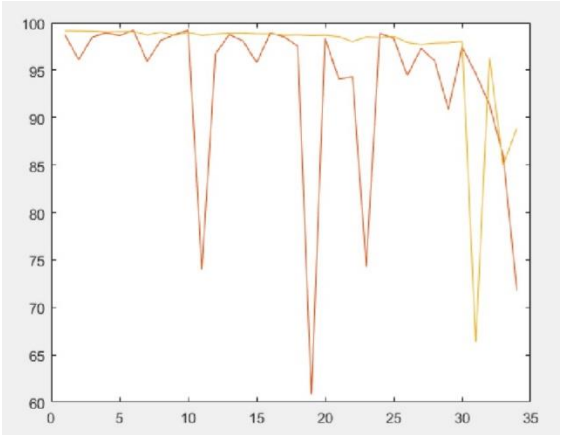
1. **Accuracy for the continuous function:**
    - The results obtained were similar to the previous case. Again, the accuracy was decreasing with the increase in generalization or increase in overlapping area. The maximum accuracy was 99.1740 when the overlapping area was 1 and the minimum accuracy was 66.3411 at overlapping area of 31.

      Accuracy= [99.1740 99.1535 99.1366 99.0373 99.0662 99.0837 98.7136 99.0117 98.6817 98.9848 98.6795 98.7994 98.9010 98.9112 98.8219 98.8227 98.6956 98.7386 98.6574 98.7009 98.5348 98.0162 98.5085 98.4395 98.5701 97.9353 97.7035 97.8771 97.9084 98.0485 66.3411 96.3204 84.9688 88.9222]

- When we compared accuracies for both discrete and continuous functions, we saw that the accuracies for both were almost the same at the beginning, but towards the end or with an increase in the overlap area, the accuracies decease randomly.

- To generalize, we say that the accuracy of a continuous function is better than that of the discrete function, but it decreases in both cases towards the end.
  Discrete function = Red Continuous function= yellow

The Plot between accuracy vs overlapping area:

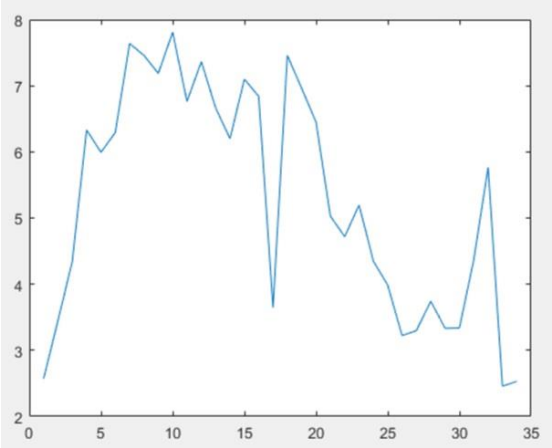


2. Time for Convergence for the continuous function:

- The time taken for convergence was varying between 3 to 8 seconds.
- The maximum time taken was 7.6388 seconds at the overlap space of 7 and minimum was 2.5732 seconds at the overlap space of 1.

$\mathbf{T_{continous}}$ = [2.5732 3.4468 4.3424 6.3318 5.9932 6.2927 7.6388 7.4569 7.1859 7.8085 6.7617 7.3635 6.6637 6.1999 7.0990 6.8395 3.6472 7.4571 6.9580 6.4494 5.0279 4.7179 5.1973 4.3467 3.9857 3.2227 3.2991 3.7414 3.3322 3.3380 4.3681 5.7640 2.4583 2.5790]

The plot as shown below:

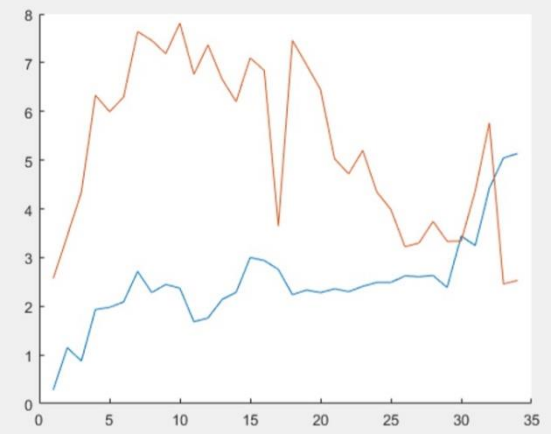


Gunjan K Khut
UID–115813083

Comparing $T_{discrete}$ and $T_{continous}$ overweight space:

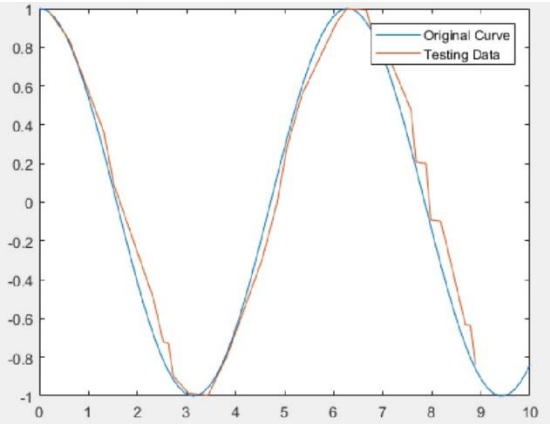
- On Comparison, we saw that the time taken to converge was generally slightly larger for continuous than the discrete. May be because of a greater number of input points in the continuous function.
  $T_{discrete}$: Blue $T_{continous}$ = Red

The Plot Comparing $T_{discrete}$ and $T_{continous}$ overweight space:



- At this numcell value, there was maximum accuracy and least time for convergence.

The Plot for continuous function and test data (when the overlap – one):



Gunjan K Khut
UID–115813083

Answer 3:

- Recurrent connections depend upon the idea that it receives two inputs, the current and the present past. There is information in the sequence of the past output which is analogous of having memory in humans. This sequential information is preserved in the networks hidden state and manages to span many times as it cascades forward to affect the processing of each new example. To avoid the problems of vanishing and Exploding gradients, Long Short-Term Memory Units were proposed.
- In this method, the information is stored outside the normal flow of the recurrent network in gated cells. These analog gates receive, pass or block information based on their own set of weights. These are also adjusted via the recurrent networks learning the process.
- To implement this concept in the original code, we would need to correct the output by error function and an additional function which depends on the output. This will make the function depend on itself or its previous outputs.

References:

[1] A Beginner's Guide to Recurrent Networks and LSTMs by DL4J. (https://deeplearning4j.org/lstm.html)