ML22/23-15 Approve Prediction of Multisequence Learning

Vasanthakumar Ramesh Lakshmi

Introduction

Analyse the existing code and understand algorithm

Generate learning sequences and save to file

Generate testing sequences out of learning sequences (subsequences) and save to file

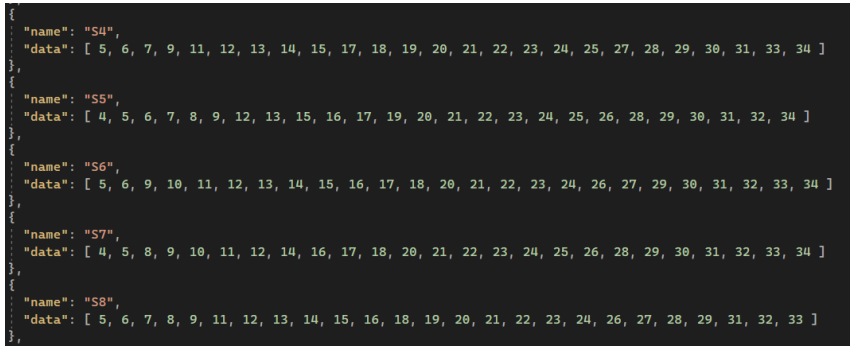
Learn the sequence, test the subsequence and

Calculate prediction accurracy

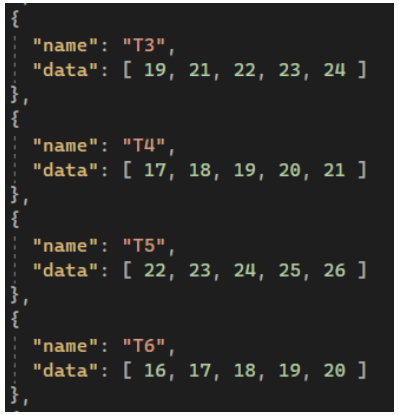
Approach: Sequence

- A particular order in which related things follow each other
- Sequence is the model of how we process and store the dataset.

```
public class Sequence
{
    public String name { get; set; }
    public int[] data { get; set; }
}
```



```
{
  "name": "S4",
  "data": [ 5, 6, 7, 9, 11, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 27, 28, 29, 30, 31, 33, 34 ]
},
{
  "name": "S5",
  "data": [ 4, 5, 6, 7, 8, 9, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 34 ]
},
{
  "name": "S6",
  "data": [ 5, 6, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 26, 27, 29, 30, 31, 32, 33, 34 ]
},
{
  "name": "S7",
  "data": [ 4, 5, 8, 9, 10, 11, 12, 14, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33, 34 ]
},
{
  "name": "S8",
  "data": [ 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 31, 32, 33 ]
},
```

```
{
  "name": "T3",
  "data": [ 19, 21, 22, 23, 24 ]
},
{
  "name": "T4",
  "data": [ 17, 18, 19, 20, 21 ]
},
{
  "name": "T5",
  "data": [ 22, 23, 24, 25, 26 ]
},
{
  "name": "T6",
  "data": [ 16, 17, 18, 19, 20 ]
},
```

Implementation: Methods

MultisequenceHelper

- HTM and Encoding Configs
- Read and Save data
- Get and Write to Log file

DatasetHelper

- Creating dataset and validating config
- Generating synthetic sequences

TestDatasetHelper

- Creating test dataset and validating config
- Creating subsequences from dataset

Implementation: Synthetic dataset and config

- Count – number of sequences to be created
- Size – length of each sequence (+3 is min. size)
- TestSize – length of each subsequence
- StartVal – start number of range
- EndVal – end number of range

```
public class ConfigOfSequence
{
    public int count { get; set; }
    public int size { get; set; }
    public int testSize { get; set; }
    public int startVal { get; set; }
    public int endVal { get; set; }

    public ConfigOfSequence(int Count, int Size,
        int TestSize, int StartVal, int EndVal)
    {
        this.count = Count;
        this.size = Size + 3;
        this.testSize = TestSize;
        this.startVal = StartVal;
        this.endVal = EndVal;
    }
}
```

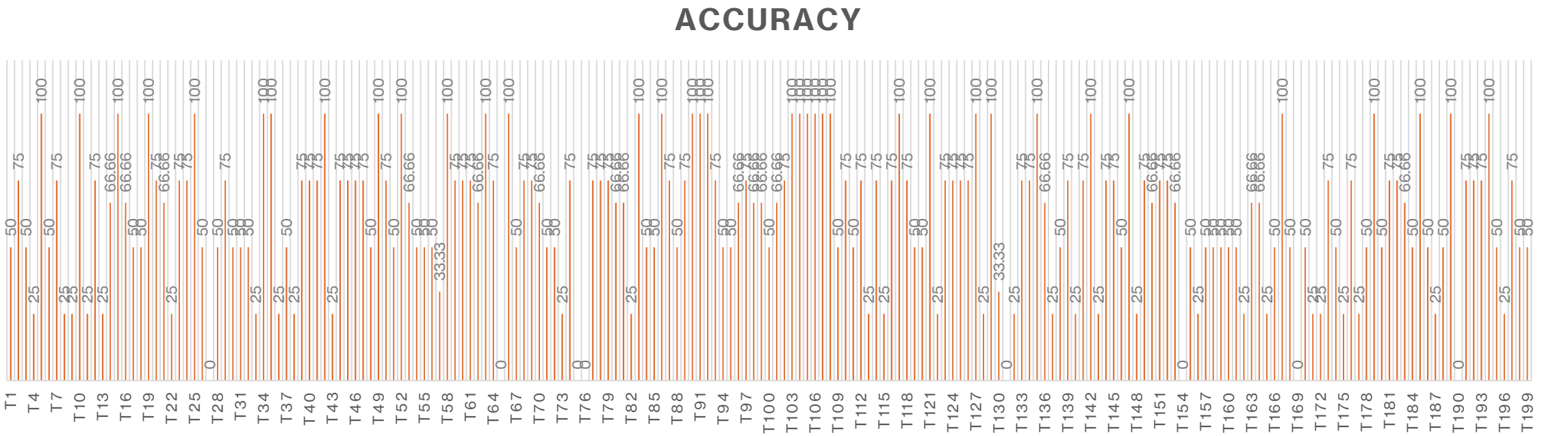
Implementation: Calculating Accuracy

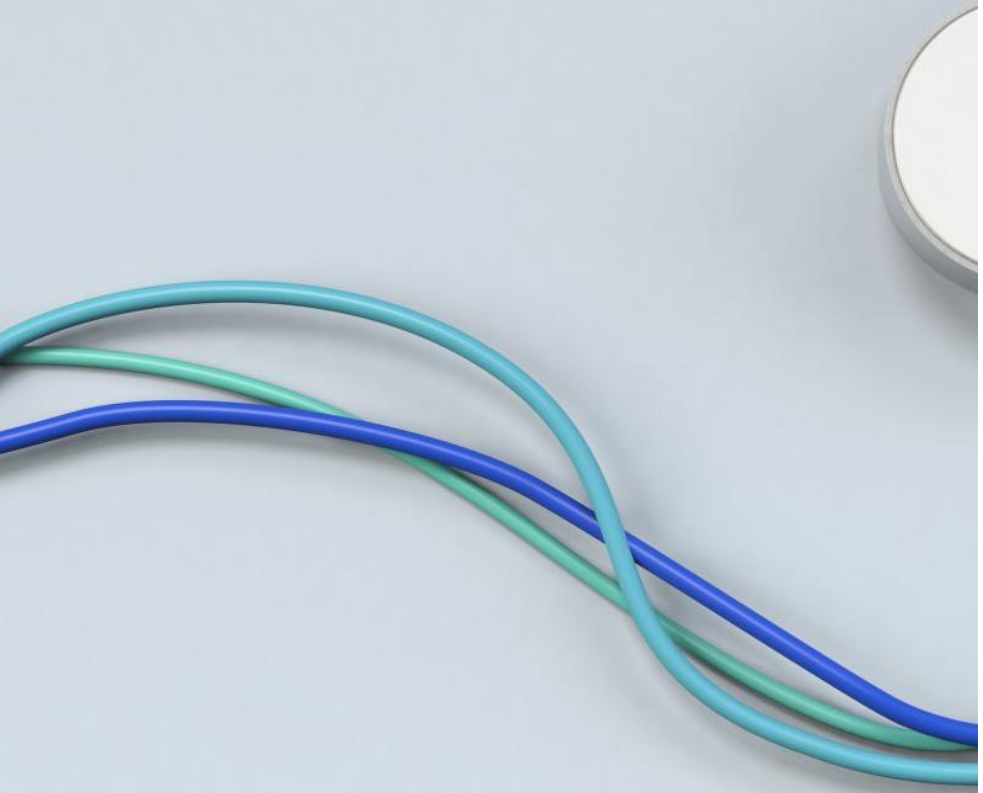
- Pseudocode

```
var matchCount = 0, prediction = 0
foreach( item in testSequence)
        predictedNextItem = predict(item)
        if(nextItem == predictedNextItem)
                matchCount++
        prediction++
        var accuracy = matchCount / prediction * 100
```

Results

- Ran the experiments 20times in sprint
- Learned over 1000 sequence of size 25 and tested with 200 sequence
- Each run has different dataset and test data
- Average run time was around 3.5hr
- 8runs 0% accuracy, 36runs 100% accuracy 58runs 75% accracy





Improvements

- Create a config json file for creating dataset
- Runs the experiment with different configuration in cloud

Thank you