**Topic: Recurrent Neural Network (RNN)**

**Instructions**

Please share your answers filled inline in the word document. Submit Python code and R code files wherever applicable.

Please ensure you update all the details:

**Name:** Nithin Dsouza

**Batch Id:** 05012021-10AM

**Topic: Recurrent Neural Network.**

1. **Business Problem**
   1. **Objective**
   2. **Constraints (if any)**
2. **Work on each feature of the dataset to create a data dictionary as displayed in the below image:**



**2.1 Make a table as shown above and provide information about the features such as its Data type and its relevance to the model building, if not relevant provide reasons and provide description of the feature.**

**Using Python code perform:**

1. **Data Pre-processing**

**3.1 Data Cleaning, Feature Engineering, etc.**

**3.2 Outlier Imputation if applicable**

1. **Model Building**
   1. **Build the Recurrent Neural Network**
   2. **Train and Test the data**
   3. **Briefly explain the model output in the documentation**
2. **Share the benefits/impact of the solution - how or in what way the business (client) gets benefit from the solution provided**
3. **Use Tensorflow for this assignment depending on your system configuration either Tensorflowgpu or Tensorflowcpu version.**

# Note:

**The assignment should be submitted in the following format:**

* **Python code**
* **Code Modularization should be maintained**
* **Documentation of the model building (elaborating on steps mentioned above)**

**Problem Statement:-**

1. Here is the time series data [110,125,133,146,158,172,187,196,210].

Build RNN/LSTM model to predict the next 10 digits.

**Business Objective:** Maximize the accuracy of the model and recognize a data’s

sequential characteristics

**Business Constraint:** Minimize the Errors

2. Write down the multiple applications of RNN.

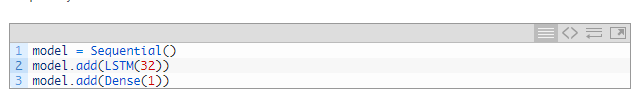
**Ans:**

* Prediction problems
* Language Modelling and Generating Text
* Machine Translation
* Speech Recognition
* Generating Image Descriptions
* Video Tagging
* Text Summarization
* Call Center Analysis
* Face detection, OCR Applications as Image Recognition
* Other applications like Music composition

3. How to do select the inputs for a LSTM/RNN models. Explain in the terms of timesteps, samples and feature.

Ans: The LSTM input layer is specified by the “input shape” argument on the first hidden layer of the network.

For example, below is an example of a network with one hidden LSTM layer and one Dense output layer.



In this example, the LSTM () layer must specify the shape of the input.

The input to every LSTM layer must be three-dimensional.

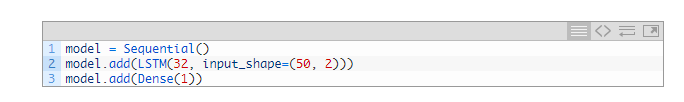
The three dimensions of this input are:

* **Samples**. One sequence is one sample. A batch is comprised of one or more samples.
* **Time Steps**. One-time step is one point of observation in the sample.
* **Features**. One feature is one observation at a time step.

This means that the input layer expects a 3D array of data when fitting the model and when making predictions, even if specific dimensions of the array contain a single value, e.g. one sample or one feature.

When defining the input layer of your LSTM network, the network assumes you have 1 or more samples and requires that you specify the number of time steps and the number of features. You can do this by specifying a tuple to the “*input shape*” argument.

For example, the model below defines an input layer that expects 1 or more samples, 50-time steps, and 2 features.



4. What are the disadvantages of MLP when dealing with sequence data.

Ans: Disadvantages of MLP include too many parameters because it is fully connected. Parameter number = width x depth x height. Each node is connected to another in a very dense web — resulting in redundancy and inefficiency.