**Mar Athanasius College of Engineering**

**Department of Computer Applications**

**Seminar Abstract**

**Topic: Long Short-Term Memory (LSTM)**

**References: Das, S., Partha, S. B., & Imtiaz Hasan, K. N. (2020). Sentence Generation using LSTM Based Deep Learning. 2020 IEEE Region 10 Symposium (TENSYMP)**

**Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN Architecture for Human Activity Recognition. IEEE Access, 8, 56855–56866**

**Hsieh, S.-T., & Lin, C.-L. (2020). Fall Detection Algorithm Based on MPU6050 and Long-Term Short-Term Memory network. 2020 International Automatic Control Conference (CACS)**

Long Short-Term Memory is an advanced version of recurrent neural network (RNN) architecture that was designed to model chronological sequences and their long-range dependencies more precisely than conventional RNNs. The major highlights include the interior design of a basic LSTM cell, the variations brought into the LSTM architecture, and a few applications of LSTMs that are highly in demand such as fall detection, and human activity recognition.

LSTM has been so designed that the vanishing gradient problem is almost completely removed, while the training model is left unaltered. Long-time lags in certain problems are bridged using LSTMs which also handle noise, distributed representations, and continuous values. LSTMs provide us with a large range of parameters such as learning rates, and input and output biases. Hence, no need for fine adjustments. The complexity to update each weight is reduced to O(1) with LSTMs, similar to that of Back Propagation Through Time (BPTT), which is an advantage.

The LSTM contains special units called memory blocks in the recurrent hidden layer. The memory blocks contain memory cells with self-connections storing the temporal state of the network in addition to special multiplicative units called gates to control the flow of information. Each memory block in the original architecture contained an input gate and an output gate. The input gate controls the flow of input activations into the memory cell. The output gate controls the output flow of cell activations into the rest of the network. Later, the forget gate was added to the memory block. This addressed a weakness of LSTM models preventing them from processing continuous input streams. The forget gate scales the internal state of the cell before adding it as input to the cell through the self-recurrent connection of the cell, therefore adaptively forgetting or resetting the cell’s memory. In addition, the modern LSTM architecture contains peephole connections from its internal cells to the gates in the same cell to learn the precise timing of the outputs.

Submitted by Faculty Guide

Name: Anlin Albert Prof. Sonia Abraham

Reg No: MAC20MCA-2003 Assistant Professor

S4 MCA MCA Dept, MACE