Detecting the insider’s threat with Long Short Term Memory (LSTM) neural networks

Eduardo Lopez
McMaster University
Hamilton, Ontario
lopeze1@mcmaster.ca

Kamran Sartipi
East Carolina University
Greenville, NC
sartipik6@ecu.edu

ABSTRACT

The experiments described in this study aim to demonstrate how the use of deep learning, and more specifically Long Short Term Memory (LSTM) neural networks, can be used for an effective detection of the insider’s threat.

2 BACKGROUND AND RELATED WORK

The experiments described in this study are anchored in computational and mathematical models. In this section we explain the theoretical framework used and how it is applied towards the detection of an insider’s threat.

2.1 Neural networks foundations

Neural networks are computational models used for classification and prediction. On the most basic form, a neural network takes a set of inputs, or predictors, performs some calculations on it and provides an output or outputs as the result of the process. A simple neural network with one node takes the inputs and generates the output based on weights that it assigns to each of the inputs.

\[ \hat{y} = \sigma(w \cdot x + b) \]

In the case of Figure 1, there are 3 inputs, each denoted with \( x_i \). The neural network assigns the weights \( w \) and sums the resulting numbers. The results of this operation are then entered into an activation function, which is used to introduce non-linearity into the output. There are three typical activation functions found in the literature. They are depicted in Figure 2.

If the activation function we use is the sigmoid (see Figure 2), the function calculation would be as follows:

\[ \hat{y} = \sigma(w \cdot x + b) \]

where \( \hat{y} \) is the predicted value, \( w \) is the weight assigned to the input \( x \) and \( b \) is the bias. The \( \sigma \) function is defined as: