Supplementary Algorithm S3 **Working of Long Short Term-Memory**

|  |
| --- |
| **Algorithm 3: Working of Long Short Term-Memory** |
| **Input:** A sequence of data points *x*, denoted as $x=\left[x\_{t-1, }x\_{t}, x\_{t+1}…\right]$ where $x\_{t}$represents the input at a specific time step "*t*" within the sequence.1. The forget gate determines whether the information from the previous timestamp should be retained or discarded, based on its relevance. Mathematically, forget gate can be represented as:

$$F\_{t}=σ(x\_{t}\*U\_{F}+f\_{t-1}\*θ\_{F})$$where $σ,x\_{t},f\_{t-1},θ\_{F}$ have the usual meanings and $U\_{F}$ is the weight associated with the input. This $F\_{t}$ is later multiplied with the cell state of the previous timestamp to determine whether the information is to be kept or not:$$S\_{t-1}\*F\_{t}=0 if F\_{t}=0$$$$S\_{t-1}\*F\_{t}=C\_{t-1} if F\_{t}=1$$1. The input gate enables the cell to learn and incorporate new information from the current input.

$$I\_{t}=σ\left(x\_{t}\*U\_{t}+f\_{t-1}\*w\_{t}\right)$$1. The updated information to be incorporated into the cell state is determined by the hidden state at the previous timestamp and the input at the current timestamp.

$$n\_{t}=tanh⁡(x\_{t}\*U\_{S}+f\_{t-1}\*w\_{c})$$The activation function used here is $tanh$, which limits the value of the new information between -1 and 1. The cell state is updated as given below:$$S\_{t}=F\_{t}\*S\_{t-1}+I\_{t}\*n\_{t}$$1. Ultimately, the cell transfers the updated information from the current timestamp to the subsequent timestamp using the output gate, mathematically represented as:

$$O\_{t}=σ(x\_{t}\*U\_{O}+f\_{t-1}\*w\_{O})$$1. The current hidden state is computed as:

$$f\_{t}=O\_{t}\*tanh⁡(S\_{t})$$ |