Learning more about LSTMs

## Task 1: Further Reading

**Read and Summarize:** Read the this article: [Understanding LSTM Networks -- colah's blog](https://colah.github.io/posts/2015-08-Understanding-LSTMs/)

Summarize the key concepts, including the structure of an LSTM cell and its components (forget gate, input gate, output gate).

**Dataset Preparation**: Choose a time-series dataset (e.g., stock prices, weather data) or a text corpus (e.g., a collection of tweets, a literary work).

**Implement an LSTM Model**: Using a machine learning framework (e.g., TensorFlow, PyTorch), implement an LSTM model to either predict the next value in your time-series dataset or generate text in your chosen corpus.

**Experiment and Analyze:** Train your model and perform experiments by varying the hyperparameters (e.g., number of LSTM layers, hidden units, learning rate). Document the impact of these changes on the performance of your model.

## Task 2: MCQ Refresher

1. What does LSTM stand for?

a) Long Short-Term Memory

b) Linear Sequence to Memory

c) Long Sequence Time Model

d) Learning Sequential Temporal Matrix

2. Which problem in sequence modeling are LSTMs particularly designed to overcome?

a) Overfitting

b) Vanishing Gradient

c) Exploding Gradient

d) Bias-Variance Tradeoff

3. What is unique about the structure of an LSTM compared to a basic RNN?

a) It has multiple hidden layers.

b) It uses a feedforward network approach.

c) It includes gates to regulate information flow.

d) It has a simpler structure.

4. Which of the following is not a component of an LSTM cell?

a) Forget Gate

b) Input Gate

c) Output Gate

d) Relay Gate

5. LSTMs are commonly used in which type of data?

a) Static Images

b) Unordered Categorical Data

c) Sequential Data

d) Independent Data Points

6. In an LSTM cell, what does the forget gate do?

a) Determines what new information to add to the cell state

b) Decides what information to remove from the cell state

c) Regulates the output of the cell to the next layer

d) Connects the current cell with the previous cell in the sequence

7. Which activation function is commonly used in the gates of an LSTM?

a) Sigmoid

b) ReLU

c) Tanh

d) Linear

8. Which of the following tasks is LSTM least suitable for?

a) Time-series forecasting

b) Image classification

c) Text generation

d) Speech recognition

9. What is the primary advantage of LSTMs over traditional RNNs?

a) Faster computation

b) Ability to learn long-term dependencies

c) More straightforward architecture

d) Lower memory usage

10. In LSTMs, which component creates a vector of new candidate values that could be added to the state?

a) Forget Gate

b) Output Gate

c) Input Gate

d) State Gate

Answers: 1-a, 2-b, 3-c, 4-d, 5-c, 6-b, 7-a, 8-b, 9-b, 10-c.