Quiz 1:

Select the correct answer related to neural networks.

1. Deep learning always uses a composite of complicated functions to create simple non-linear functions
2. A fully recurrent network is the simplest neural network architecture because all nodes are connected to the same single nodes, and each node works as an input and output
3. A Jordan network is a closed-loop network in which the output is fed back to the input
4. During the training of networks under reinforcement learning, the network does not receive any feedback from the environment
5. A feedforward network with a single layer is insufficient to represent any functions

Quiz 2:

Select the correct answer related to neural networks.

1. Training neural networks is a non-convex optimization problem
2. Training learning weights during backup maximize the gradient descent
3. Maximizing weights involves adding an extra term to the cost function that penalizes the squared weights; keeping weights small unless they have big error derivatives
4. One way to use posterior uncertainty is to sample a set of values from the posterior, and then calculate the minimum value of their predictive distributions
5. We are unable to put error bars on the output of the work and uncertainty

Quiz 3:

Which of the following statements is true about LSTM networks?

1. LSTM networks have a fixed number of layers
2. LSTM networks are not well-suited for processing sequential data
3. LSTM networks use gates to selectively allow information to pass through
4. LSTM networks do not require any form of recurrent connections
5. None of the above

Quiz 4:

グラフ

自動的に生成された説明Consider the following RNN:

This specific type of architecture is appropriate when

Quiz 5:  
What’s the output shape of a bidirectional LSTM layer with 64 units?  
a) (None, 64)

b) (None, 128)

c) (128, None)

d) (128,1)

e) (1,128)

Quiz 6:

Which of the following best describes the vanishing gradient problem in RNNs?

1. When the gradients in the backward pass are too large and cause the weights to update too much
2. When the gradients in the backward pass are too small and cause the weights to update very little or not at all
3. When the hidden state of the RNN becomes too large and causes numerical overflow
4. When the hidden state of the RNN becomes too small and causes numerical underflow

Quiz 7:

Which of the following statements is true about the sigmoid function?

1. It maps any input value to a value between 0 and 1
2. It is a linear function
3. It can only be used as an activation function in neural networks
4. It has an unbounded range
5. None of the above

Quiz 8:  
Suppose we have an LSTM model with one input layer, one output layer, and two LSTM layers, each with 100 hidden units. The input data has a sequence length of 20, and each input data point has 50 features. What is the total number of parameters in the first LSTM layer?   
(NOTE: \*You may use Keras, TensorFlow, or PyTorch to answer this question)

Quiz 9:

Suppose you have 10,000 data points and set the number of training iterations per epoch to 500. What should be the batch size for this dataset? Please provide your answer as a numerical value.

Quiz 10:

This question asks you to use a Python notebook to answer.  
(NOTE: use Q10.ipynb, answer the below )

* You are building a neural network using the following layers:  
   A dense layer with 512 units with an input shape of max\_words.
* A ReLU activation function
* A dropout rate of 0.5
* Another dense layer with num\_classes units
* A final activation function with softmax

The network is compiled with categorical cross-entropy as the loss function, the Adam optimizer, and accuracy as an evaluation metric.  
  
**Hint**:

model.compile(loss='categorical\_crossentropy',

optimizer='adam',

metrics=['accuracy'])

model.summary()

What is the total number of parameters in this neural network?

Quiz 11  
Select all the hyperparameters that can be optimized in a neural network  
*Select all that apply.*  
a) Number of nodes

b) Number of layers

c) Dropout rate

d) Epoch

e) Functions (Tanh, ReLu, Sigmoid, etc.)

Quiz 12:

テキスト

自動的に生成された説明Here are the equations for the GRU and the LSTM:

From these, we can see that the Update Gate and Forget Gate in the LSTM play a role similar to \_\_\_\_\_\_\_ and \_\_\_\_\_\_ in the GRU. What should go in the the blanks?



Quiz 13:

Select the correct answer pertaining to LSTM and GRU:

1. LSTM's forget gate passes the previous hidden state through the tanh function
2. LSTM's output gate decides what the previous hidden state should be
3. GRU is an older generation of neural networks
4. GRU's forget and input gates function as the update gate in LSTM
5. Hybrid models can combine both layers of LSTM and GRU