D 71	DATAT	\sim .	.	
Dav/L	RNN	()1117.	Attempt	review
ouy , i		Quiz.	riccinp	1011011

Started on Thursday, 12 June 2025, 4:53 PM	Time taken Tim		
Completed on Time taken 5 5 mins 9 secs Marks 7.00/15.00 Grade 46.67 out of 100.00 In a standard RNN, the hidden state h _t is updated as: (a) a. h _t =tanh(W x _t + U h _{t-1} + b) (b. h _t =σ(W x _t + b) (c. h _t = tanh(W x _t + b) (d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer (b. Calculates gradients (c. Stores long-term memory) (d. Stores hidden layers) Question 3 Complete	Time taken 1 mins 19 secs 1 mins 9 secs 100.00 Guestion 1		
Time taken 5 mins 9 secs Marks 7.00/15.00 Grade 46.67 out of 100.00 In a standard RNN, the hidden state h _t is updated as: In a standard RNN, the hidden state h _t	Time taken 5 mins 9 secs 100 mins 7 secs 100 mins 9 secs 100 mins 9 secs 100 mins 9 secs 100 mins 9 secs 100 mins		
Marks 7.00/15.00 Grade 46.67 out of 100.00 Question 1 Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =\(\text{o}(W x_t + b)\) c. c. h _t = tanh(W x _t + b) d. h _t = \(\text{ReLU}(x_t)\) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Marks 7.00/15.00		
Question 1 Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =σ(W x _t + b) c. h _t = tanh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. C. Stores long-term memory d. Stores hidden layers Question 3 Complete Question 3 Complete	Complete Mark 1.00 out of 1.00		
Question 1 Complete Mark 1:00 out of 1:00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =σ(W x _t + b) c. h _t = σ(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0:00 out of 1:00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Question 1 Complete Mark 1:00 out of 1:00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =o(W x _t + b) c. h _t = tanh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0:00 out of 1:00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0:00 out of 1:00 Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers		
Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =\(\sigma(W x_t + b)\) c. h _t = tanh(W x_t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =o(W x _t + b) c. h _t = anh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculate gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	Grade	46.67 Out of 100.00
Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =\(\sigma(W x_t + b)\) c. h _t = tanh(W x_t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Complete Mark 1.00 out of 1.00 In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =o(W x _t + b) c. h _t = anh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculate gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	Question 1	
In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =\(\sigma(W x_t + b)\) c. h _t = tanh(W x _t + b) d. h _t = ReLU(x _t) question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers question 3 Complete	In a standard RNN, the hidden state h _t is updated as: a. h _t =tanh(W x _t + U h _{t-1} + b) b. h _t =c(W x _t + b) c. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C ₁ ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
■ a. $h_t = \tanh(W x_t + U h_{t-1} + b)$ b. $h_t = \sigma(W x_t + b)$ c. $h_t = \tanh(W x_t + b)$ d. $h_t = ReLU(x_t)$ Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C_t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	a. h ₁ =tanh(W x _t + U h _{t-1} + b) b. h ₁ =o(W x _t + b) c. h ₁ = tanh(W x _t + b) d. h ₁ = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
	a. h ₁ =tanh(W x _t + U h _{t-1} + b) b. h ₁ =o(W x _t + b) c. h ₁ = tanh(W x _t + b) d. h ₁ = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
	a. h ₁ =tanh(W x _t + U h _{t-1} + b) b. h ₁ =o(W x _t + b) c. h ₁ = tanh(W x _t + b) d. h ₁ = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
$\begin{array}{ll} b. & h_t = \alpha(Wx_t + b) \\ & c. & h_t = \tanh(Wx_t + b) \\ & d. & h_t = \text{ReLU}(x_t) \\ \\ \hline \textbf{Question 2} \\ \hline \textbf{Complete} \\ \text{Mark 0.00 out of 1.00} \\ \\ \hline \textbf{In an LSTM cell, what is the function of the cell state C}_t? \\ & a. & \text{Acts as the output layer} \\ & b. & \text{Calculates gradients} \\ & c. & \text{Stores long-term memory} \\ \hline \textbf{0} & d. & \text{Stores hidden layers} \\ \hline \\ \hline \textbf{Question 3} \\ \hline \textbf{Complete} \\ \\ \hline \end{array}$	b. h _t =σ(W x _t + b) c. h _t = tanh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	In a standard RNN, t	he hidden state h _t is updated as:
$\begin{array}{ll} b. & h_t = \alpha(W x_t + b) \\ & c. & h_t = \tanh(W x_t + b) \\ & d. & h_t = \text{ReLU}(x_t) \\ \\ \hline \textbf{Question 2} \\ \hline \textbf{Complete} \\ \text{Mark 0.00 out of 1.00} \\ \\ \hline \textbf{In an LSTM cell, what is the function of the cell state C}_t? \\ & a. & \text{Acts as the output layer} \\ & b. & \text{Calculates gradients} \\ & c. & \text{Stores long-term memory} \\ \hline \textbf{0} & d. & \text{Stores hidden layers} \\ \hline \\ \hline \textbf{Question 3} \\ \hline \textbf{Complete} \\ \hline \end{array}$	b. h _t =σ(W x _t + b) c. h _t = tanh(W x _t + b) d. h _t = ReLU(x _t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
 C. h_t = tanh(W x_t + b) d. h_t = ReLU(x_t) Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Question 3 Complete	Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
 Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory ⑥ d. Stores hidden layers Question 3 Complete	Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	\bigcirc b. $h_t = \sigma(W x_t +$	b)
Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Question 2 Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	\bigcirc c. $h_t = tanh(W$	$x_t + b$)
Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	\bigcirc d. $h_t = ReLU(x_t)$	
Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	Question 2	
In an LSTM cell, what is the function of the cell state C _t ? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	In an LSTM cell, what is the function of the cell state Ct? a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	Complete	
 a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	Mark 0.00 out of 1.00	
 a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
 a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	a. Acts as the output layer b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	L LCTM II L	
 b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	 b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate 	In an LSTM cell, what	: is the function of the cell state C_t ?
 b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete	 b. Calculates gradients c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate 	Ω a Acts as the o	output laver
 c. Stores long-term memory ● d. Stores hidden layers Question 3 Complete	 c. Stores long-term memory d. Stores hidden layers Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate 		
Question 3 Complete	Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
Question 3 Complete	Question 3 Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate	d. Stores hidde	en layers
Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
Complete	Complete Mark 0.00 out of 1.00 In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
	In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
Mark 0.00 out of 1.00	In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward? a. Forget gate b. Memory gate		
	a. Forget gateb. Memory gate	Mark 0.00 out of 1.00	
	a. Forget gateb. Memory gate		
In an LSTM cell, which gate controls how much of the previous hidden state should be carried forward?	a. Forget gateb. Memory gate	In an LSTM cell, whic	h gate controls how much of the previous hidden state should be carried forward?
	b. Memory gate		
a. Forget gate		a. Forget gate	
b. Memory gate		b. Memory gat	e
	c. Output gate		
○ c. Output gate			
	d. Input gate	_ a. input gate	
c. Output gated. Input gate	○ d. Input gate		
a. Forget gateb. Memory gate		a. Forget gateb. Memory gatc. Output gate	e
		 d. Input gate 	
	○ d. Input gate		
	○ d. Input gate		

D 71	DATAT	\circ .	A	
1)av/1	KNN	()1117:	Aftemr	ot review
_ ~,		~ ~ · · · ·		

Question 4 Complete Mark 100 gut of 100
Mark 1.00 out of 1.00
In sequence-to-sequence models, what is the role of the encoder?
○ a. Translate output sequence
O b. Predict next token
c. Encode input sequence into a fixed representation
Od. Update output vocabulary
Question 5 Complete
Mark 1.00 out of 1.00
What does teacher forcing refer to during RNN training?
a. Using the model's own output as input
b. Feeding the ground truth output at time t-1 to predict time t
C. Pre-training the encoder before decoder
d. Resetting hidden states between batches
Question 6 Complete Mark 1.00 out of 1.00
What is gradient clipping in the context of training RNNs?
a. Limiting updates to only the final layer
O b. Reducing batch size to avoid overfitting
 c. Restricting the magnitude of gradients to prevent exploding gradients
Od. Applying dropout to avoid vanishing gradients
Question 7 Complete
Mark 0.00 out of 1.00
What is the main reason RNNs struggle with learning long-term dependencies?
a. Gradient explosion
○ b. Vanishing gradients
c. Lack of activation functions

D 71	DATAT	O . A.,	
Day/L	KINN	Quiz: Attempt review	
_ ~,		_ ~	

Question 8 Complete Mark 1.00 out of 1.00 What is the primary advantage of using bidirectional RNNs? a. Reduced computation time b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout d. Beam search
Mark 1.00 out of 1.00 What is the primary advantage of using bidirectional RNNs? a. Reduced computation time b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
What is the primary advantage of using bidirectional RNNs? a. Reduced computation time b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
 a. Reduced computation time b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
 a. Reduced computation time b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
 b. Works with images c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
 c. Access to both past and future context d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
 d. Replaces the need for attention mechanisms Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
Question 9 Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
Complete Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
Mark 0.00 out of 1.00 What technique is commonly used during inference in seq2seq models to improve generation quality? a. Adam optimizer b. Batch normalization c. Dropout
a. Adam optimizerb. Batch normalizationc. Dropout
a. Adam optimizerb. Batch normalizationc. Dropout
a. Adam optimizerb. Batch normalizationc. Dropout
b. Batch normalizationc. Dropout
○ c. Dropout
○ d. Beam search
Question 10 Complete
Mark 0.00 out of 1.00
Which loss function is most commonly used in training sequence-to-sequence models with RNNs for classification?
a. Hinge Loss
a. Hinge Lossb. Categorical Crossentropy
a. Hinge Lossb. Categorical Crossentropyc. Mean Squared Error
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete Mark 0.00 out of 1.00
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete Mark 0.00 out of 1.00 Which mechanism allows RNN-based models to focus on specific parts of the input during decoding?
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete Mark 0.00 out of 1.00 Which mechanism allows RNN-based models to focus on specific parts of the input during decoding? a. Dropout
 a. Hinge Loss b. Categorical Crossentropy c. Mean Squared Error d. Binary Crossentropy Question 11 Complete Mark 0.00 out of 1.00 Which mechanism allows RNN-based models to focus on specific parts of the input during decoding? a. Dropout b. Beam search

Day71	RNN	Quiz:	Attemp	ot review
-------	-----	-------	--------	-----------

uestion 12
omplete
lark 0.00 out of 1.00
Which of the following statements about GRU is incorrect?
a. GRU has fewer parameters than LSTM
b. GRU combines the forget and input gates into a single update gate
c. GRU has a separate memory cell c_t like LSTM
d. GRU is generally faster to train than LSTM
uestion 13
omplete
lark 0.00 out of 1.00
Which one is not a typical application of RNNs?
a. Machine translation
b. Speech recognition
c. Object detection
d. Sentiment analysis
uestion 14
uestion 14 omplete
omplete
omplete lark 1.00 out of 1.00
omplete
Omplete Slark 1.00 out of 1.00 Which RNN variant is specifically designed to solve the vanishing gradient problem?
omplete lark 1.00 out of 1.00
Omplete Slark 1.00 out of 1.00 Which RNN variant is specifically designed to solve the vanishing gradient problem?
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN westion 15 complete lark 1.00 out of 1.00 Why are RNNs not inherently parallelizable across time steps?
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN uestion 15 complete lark 1.00 out of 1.00 Why are RNNs not inherently parallelizable across time steps? a. Each output depends on previous output
omplete lark 1.00 out of 1.00 Which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN uestion 15 omplete lark 1.00 out of 1.00 Why are RNNs not inherently parallelizable across time steps? a. Each output depends on previous output b. Due to weight sharing
which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN westion 15 complete lank 1.00 out of 1.00 why are RNNs not inherently parallelizable across time steps? a. Each output depends on previous output b. Due to weight sharing c. They use convolutional filters
omplete lark 1.00 out of 1.00 Which RNN variant is specifically designed to solve the vanishing gradient problem? a. LSTM b. GRU c. Vanilla RNN d. Bidirectional RNN uestion 15 omplete lark 1.00 out of 1.00 Why are RNNs not inherently parallelizable across time steps? a. Each output depends on previous output b. Due to weight sharing