	Recurrent Neural Networks Graded Quiz • 30 min			Due Feb 3, 3:59 PM CS1
		✓ Congratulations! You passed! TO PASS 80% or higher Keep Learni	grade 100%	
		TO PASS 60% Of Higher		
		Recurrent Neural Networks LATEST SUBMISSION GRADE		
		100%		
		1. Suppose your training examples are sentences (sequences of words). Which of the following refers to the j^{th} word in the i^{th} training example?	le 1/1 point	
		$\bigcirc x^{< i > (j)}$ $\bigcirc x^{(j) < i >}$		
		$igcup x^{< j > (i)}$		
		\checkmark Correct We index into the i^{th} row first to get the i^{th} training example (represented by parentheses), then the j^{th} column to get the j^{th} word (represented by the brackets).		
		column to get the y word (represented by the brackets).		
<	Recurrent Neural Networks Graded Quiz • 30 min	y y v		Due Feb 3, 3:59 PM CS
		$ \uparrow \qquad \uparrow \qquad$		
		$a^{<0>} \longrightarrow \begin{array}{ c c c c c c c c c c c c c c c c c c c$		
		$\uparrow \qquad \uparrow \qquad \uparrow$		
		$x^{<1>}$ $x^{<2>}$ $x^{<3>}$ $x^{}$		
		This specific type of architecture is appropriate when:		
		$igcup_{T_x} < T_y$ $igcup_{T_x} > T_y$		
		$igcup_x = 1$ $igcup_x$ Correct		
		It is appropriate when every input should be matched to an output.		
		3. To which of these tasks would you apply a many-to-one RNN architecture? (Check all that apply).	1/1 point	
		\hat{y}		
<	Recurrent Neural Networks Graded Quiz • 30 min	a<□>→ → → · · · · →		Due Feb 3, 3:59 PM CS
		a · · · · →		
		$\chi^{<1>}\chi^{<2>}$ $\chi^{}$		
		Speech recognition (input an audio clip and output a transcript)		
		Sentiment classification (input a piece of text and output a 0/1 to denote positive or negative sentiment)		
		Correct Correct!		
		 Image classification (input an image and output a label) ✓ Gender recognition from speech (input an audio clip and output a label indicating the speaker's gender) 		
		✓ Correct Correct!		
		4. You are training this RNN language model.	1 / 1 point	
<	Recurrent Neural Networks Graded Quiz • 30 min			Due Feb 3, 3:59 PM CS
		$a^{<0>}$		
		$\uparrow \qquad \uparrow \qquad \qquad \downarrow \qquad \qquad \uparrow \qquad \qquad \downarrow \qquad \qquad \uparrow \qquad \qquad \downarrow \qquad \qquad $		
		$0 y^{<1>} y^{}$		
		At the t^{th} time step, what is the RNN doing? Choose the best answer.		
		C Estimating $P(y^{<1>}, y^{<2>}, \dots, y^{< t-1>})$		
		Estimating $P(y^{< t>})$ Estimating $P(y^{< t>} \mid y^{< 1>}, y^{< 2>}, \dots, y^{< t-1>})$		
		Correct $P(y^{< t>} \mid y^{< 1>}, y^{< 2>}, \dots, y^{< t>})$		
		Yes, in a language model we try to predict the next step based on the knowledge of all prior steps.		
		5. You have finished training a language model RNN and are using it to sample random sentences, as follows:	1 / 1 point	
		$\hat{y}^{<1>} \qquad \hat{y}^{<2>} \qquad \hat{y}^{<3>} \qquad \qquad \hat{y}^{}$		
<	Recurrent Neural Networks Graded Quiz • 30 min			Due Feb 3, 3:59 PM CS
		A \ A \ \ A \ \ .		
		$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		
		What are you doing at each time step t ?		
		What are you doing at each time step t ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then		
		What are you doing at each time step t ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then		
		What are you doing at each time step t ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step.		
		What are you doing at each time step t ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass this selected word to the next time-step.		
		What are you doing at each time step t ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as $\hat{y}^{< t>}$. (ii) Then pass this selected word to the next time-step.		
		 What are you doing at each time step t? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<t></t>}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<t></t>}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<t></t>}. (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<t></t>}. (ii) Then pass this selected word to the next time-step. 		
		 What are you doing at each time step t? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<t></t>}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<t></t>}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<t></t>}. (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<t></t>}. (ii) Then pass this selected word to the next time-step. ✓ Correct Yes! 6. You are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number"). Which of these is the most likely cause of this problem? 		
		 What are you doing at each time step t? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ct (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ct (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ct (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ct (iii) Then pass this selected word to the next time-step. ✓ Correct		
	Recurrent Neural Networks Graded Quiz • 30 min	 What are you doing at each time step t? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-(t)}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-(t)}. (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-(t)}. (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-(t)}. (ii) Then pass this selected word to the next time-step. ✓ Correct Yes! 6. You are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number") Which of these is the most likely cause of this problem? ✓ Vanishing gradient problem. © Exploding gradient problem. © Exploding gradient problem. ○ ReLU activation function g(.) used to compute g(z), where z is too large. 		Due Feb 3, 3:59 PM CS
		 What are you doing at each time step <i>f</i>? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<□}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<□}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{<□}, (ii) Then pass this selected word to the next time-step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{<□}, (ii) Then pass this selected word to the next time-step. ✓ Correct Yes! Vou are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number") Which of these is the most likely cause of this problem? Vanishing gradient problem. Exploding gradient problem. ReLU activation function g(.) used to compute g(z), where z is too large. Sigmoid activation function g(.) used to compute g(z), where z is too large. Sigmoid activation function g(.) used to compute g(z), where z is too large. 		Due Feb 3, 3:59 PM CS
		 What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-CD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-CD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-CD}, (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-CD}, (ii) Then pass this selected word to the next time-step. ✓ correct	1/1 point	Due Feb 3, 3:59 PM CST
		 What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-cD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-cD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{-cD}, (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{-cD}, (ii) Then pass this selected word to the next time-step. Correct Vest Vou are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number") Which of these is the most likely cause of this problem? Vanishing gradient problem. Exploding gradient problem. ReLU activation function g(,) used to compute g(z), where z is too large. Sigmoid activation function g(,) used to compute g(z), where z is too large. Sigmoid activation function g(,) used to compute g(z), where z is too large. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a^{-cD}, What is the dimension of Ω_u at each time step? 1 	1/1 point	Due Feb 3, 3:59 PM CST
		 What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{<d< sup="">, (ii) Then pass the ground-truth word from the training set to the next time-step.</d<>} (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{<d< sup="">, (ii) Then pass the ground-truth word from the training set to the next time-step.</d<>} (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{<d< sup="">, (ii) Then pass this selected word to the next time-step.</d<>} (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{<d< sup="">, (ii) Then pass this selected word to the next time-step.</d<>} ✓ Correct	1/1 point	Due Feb 3, 3:59 PM CST
		 What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{CD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{CD}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{CD}, (ii) Then pass this selected word to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{CD}, (ii) Then pass this selected word to the next time-step. ✓ Correct	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
		 What are you doing at each time step t? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{CD→}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{CD→}, (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ^{CD→}, (ii) Then pass this selected word to the next time-step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ^{CD→}, (iii) Then pass this selected word to the next time-step. ✓ Correct	1/1 point	Due Feb 3, 3:59 PM CS
		What are you doing at each time step ℓ? ○ (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c+2} · (ii) Then pass the ground-truth word from the training set to the next time-step. ○ (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c+2} · (ii) Then pass the ground-truth word from the training set to the next time-step. ○ (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c+2} · (ii) Then pass this selected word to the next time-step. ● (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c+2} · (ii) Then pass this selected word to the next time-step. ✓ Correct Vest 6. You are training an RNN, and find that your weights and activations are all taking on the value of NaN (*Not a Number*) Which of these is the most likely cause of this problem? ○ Vanishing gradient problem. ● Exploding gradient problem. ● Exploding gradient problem. ○ RetU activation function g(.) used to compute g(z), where z is too large. ○ Sigmoid activation function g(.) used to compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary and are using an LSTM. 8. Here're the update equations for the GRU.	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
		What are you doing at each time step ℓ? (i) (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c-t-} , (ii) Then pass the ground-truth word from the training set to the next time-step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (ii) Then pass the ground-truth word from the training set to the next time-step. (iii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c-t-} , (ii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word to the next time-step. (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word to the next time-step. (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (ii) Then pass this selected word to the next time-step. (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c-t-} , (iii) Then pass this selected word for the next time-step. (iii) Then pass this selected word for the sample sample sample sample sampl	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
		What are you doing at each time step ℓ? ○ (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c+2} · (ii) Then pass the ground-truth word from the training set to the next time-step. ○ (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c+2} · (ii) Then pass the ground-truth word from the training set to the next time-step. ○ (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c+2} · (ii) Then pass this selected word to the next time-step. ● (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c+2} · (ii) Then pass this selected word to the next time-step. ✓ Correct Vest 6. You are training an RNN, and find that your weights and activations are all taking on the value of NaN (*Not a Number*) Which of these is the most likely cause of this problem? ○ Vanishing gradient problem. ● Exploding gradient problem. ● Exploding gradient problem. ○ RetU activation function g(.) used to compute g(z), where z is too large. ○ Sigmoid activation function g(.) used to compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary and are using an LSTM with 100-dimensional activations a compute g(z), where z is too large. 7. Suppose you are training a LSTM. You have a 10000 word vocabulary and are using an LSTM. 8. Here're the update equations for the GRU.	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Graded Quiz • 30 min	What are you doing at each time step <i>t</i> ? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as <i>ŷ</i> ^{-cto} . (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as <i>ŷ</i> ^{-cto} . (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as <i>ŷ</i> ^{-cto} . (ii) Then pass this selected word to the next time-step. (ii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as <i>ŷ</i> ^{-cto} . (ii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as <i>ŷ</i> ^{-cto} . (iii) Then pass the selected word time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word fo	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
		What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ĝ ^{CED} , (ii) Then pass the ground-druth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ĝ ^{CED} , (ii) Then pass the ground-druth word from the training set to the next time-step. (i) Use the probabilities down to the next time-step. (ii) Use the probabilities output by the RNN to trandomly sample a chosen word for that time-step as ĝ ^{CED} , (iii) Then pass this selected word to the next time-step. ✓ cerrect Yest! 6. You are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number") Which of these is the most likely cause of this problem? ✓ vanishing gradient problem. © Exploding gradient problem. © Exploding gradient problem. © ReLU activation function gi.) used to compute gizi, where z is too large. 7. Suppose you are training a LSTM, You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a cfD., What is the dimension of Γ _w at each time step? 1 © 100 300 10000 ✓ cerrect CGRU c ^{CCD} = tanh(W _c (Γ _r * c ^{CC-1D} , x ^{CCD}) + b _u) Γ _w = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _w = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u) Γ _r = σ(W _w (c ^{CC-1D} , x ^{CCD}) + b _u)	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step £7 (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c, +1, -1, -1, -1} (ii) Then pass the ground-truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c, +1, -1} (ii) Then pass this ground-truth word from the training set to the next time-step. (ii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as ŷ ^{c, +1} (ii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to probe the pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c, +1} (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c, +1} (iii) Then pass this selected word to the next time-step. (iv) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as ŷ ^{c, +1} (iii) Then pass this selected word to the next time-step. (iv) Then pass the most like and the next time-step. (iv) Then pass time and find that your weights and activations are all taking on the value of NaN ('Not a Number') Which of these is the most like years of this problem? (iv) Variability and the time-step as ŷ ^{c, +1} (iii) Then the value of NaN ('Not a Number') Which of these is the most like years of this problem? (iv) Variability and the value of NaN ('Not a Number') Which of these is the most like years of this problem? (iv) Variability and the value of NaN ('Not a Number') Which of these is the most like years of this problem? (iv) Variability and the value of NaN ('Not a Number') Which of these is the most limit and the value of NaN ('Not a Number') Which of these is the most limit and the problem? (iv) Variability and the problem? (iv) Va	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step £? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^(ct) , (ii) Then pass the ground-routh word from the training set to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^(ct) , (ii) Then pass this productive toword to the next time-step. (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^(ct) , (ii) Then pass this selected word to the next time-step. (ii) (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step as g ^(ct) , (iii) Then pass this selected word to the next time-step. (iii) Then pass this selected word time-step as g ^(ct) , (iii) Then pass this selected word time-step as g ^(ct) , (iii) Then pass this selected word time-step as g ^(ct) , (iii) Then pass this selected word time-step as g ^(ct) , (iii) Then pass this selected word time-step as g ⁽	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step £? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{r.to.} , (ii) Then past the ground truth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{r.to.} , (ii) Then past the ground-truth word from the training set to the next time-step. (ii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{r.to.} , (ii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{r.to.} , (ii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{r.to.} , (ii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{r.to.} , (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{r.to.} , (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{r.to.} , (iii) Then past this selected word to the next time-step as g ^{r.to.} , (iii) Then past this selected word to the next time-step as g ^{r.to.} , (iii) Then past this selected word to the next time-step. (iii) The past time selected word to the next time-step as g ^{r.to.} , (iii) Then past time step? (iv) Carrect vertically a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a g ^{r.to.} , (white is the dimension of Γ ₀ , at each time step? (iv) 10000 and 10000 and 10000 are taking a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations a "Co. 10000 are taking a LST	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step f7 (i) (but the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{x+x} . (ii) Then past the ground truth word from the training set to the next time step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (ii) Then past the ground struth word from the training set to the next time-step. (ii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{x+x} . (ii) Then past this selected word to the next time step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{x+x} . (iii) Then past this selected word to the next time-step without find that your weights and activations are all taking on the value of NaN ('Not a Number') Which of these is the most likely cause of this problem. (iii) Seplocing gradient problem. (iv) Sep	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step (?) (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{-(*)} , (i) Then pass the ground-truth word from the training set for the next time-step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (ii) Then pass the ground-truth word for the training set to the next time-step. (ii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (ii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (iii) Then pass this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-(*)} , (iii) Then pass this selected word to the next time-step. (iii) Variability and the selected word to the next time-step as g ^{-(*)} , (iii) Then pass this selected word to the next time step? (iii) Sigmoid activation function g() used to compute g(t), where z is too large. (iiii) Sigmoid activation function g() used to compute g(t), where z is too large. (iv) Sigmoid activation function g() used to compute g(t), where z is too large. (iv) Sigmoid activation function g() used to compute g(t), where z is too large. (iv) Sigmoid activation function g() used to compute g(t), where z is too large. (iv) Sigmoid activation function g() used to compute g(t), where z is too large. (iv) Sigmoid activation function g(t) used to comput	1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step f? (i) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{-1/2} , (ii) Then past the ground druth word from the training set to the next time-step. (i) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (ii) Then past the ground druth word from the training set to the next time-step. (ii) Use the probabilities output by the RNN to pick the highest probability word for that time-step as g ^{-1/2} , (ii) Then past the ground druth word from the training set to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (ii) Then past the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (iii) Then past this selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (iii) Then past the selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (iii) Then past the selected word to the next time-step. (iii) Use the probabilities output by the RNN to randomly sample a chosen word for that time-step as g ^{-1/2} , (iii) Then past time step. (iv) Variation gradient problem. (iii) RNN that the selected word to the the time step. (iv) Variation gradient problem. (iv) RNN that time step as g ^{-1/2} , (iii) Then past time step as gradient samples and statistic time. (iv) Variation gradient problem. (iv) RNN that time step as g ^{-1/2} , (iii) Then past time step as g ^{-1/2} , (iii) Then past time step as g ^{-1/2} , (iii) Then past time step as a variation of the gradient samples back through that timestep without much decay. (iv) All	1/1 point 1/1 point 1/1 point	
	Recurrent Neural Networks	Vihal are you doing at each time step <i>t</i> ? (I) Use the propositiones output by the titon to pick the highest proposition you for that time-step as <i>y</i> ⁻¹⁰⁰ , (ii) Then past the ground-huth word from the training set to the next time-step. (I) Use the propositiones output the titon the training set to the next time-step. (I) Use the propositiones output by the RINN to pick the highest probability word for that time-step as <i>y</i> ⁻¹⁰⁰ , (ii) Then past the ground-huth word from the training set to the next time-step. (I) Use the probabilities output by the RINN to pick the highest probability word for that time-step as <i>y</i> ⁻¹⁰⁰ , (iii) Then past this set set word to the next time-step. (I) Use the probabilities output by the RINN to randomly sample a drosen word for that time-step as <i>y</i> ⁻¹⁰⁰ , (iii) Then past this set set output to the next time-step. (I) Use the probabilities output by the RINN to randomly sample a drosen word for that time-step as <i>y</i> ⁻¹⁰⁰ , (iii) Then past this set set of the problem? (I) Use the probabilities output by the RINN to randomly sample a drosen word for that time-step as <i>y</i> ⁻¹⁰⁰ , (iii) Then past this set of the problem? (I) Variation of the set of the rest time by the problem? (I) Variation of the set of the most timely cause of this problem? (I) Variation of the set of the most timely cause of the problem? (I) Variation of the set of the most timely cause of the problem? (I) Suppose you are training a LSTM. You have a 10000 word variability, and are using an LSTM with 100-dimensional activations at "1," what is the dimension equal to the number of hidden units in the LSTM. (I) Suppose you are training a LSTM. You have a 10000 word variability, and are using an LSTM with 100-dimensional activations at "1," what is the dimension of T ₁ , at each time step? (I) I the past training a LSTM, you have a 10000 word variability and are using an LSTM with 100-dimensional activations as "1," what is the dimension equal to the number of hidden units in the LSTM. (I) Carrier t	1/1 point 1/1 point 1/1 point	
	Recurrent Neural Networks	What are you doing at each time step t ? Of the the probabilities contain by the filt No to juck the highest probability word for that time-step as g^{***} . (i) Then pass the ground-truth word from the training stat to the next time-step. Of the the probabilities countly the filt No to juck the highest probability word for that time-step as g^{***} ?. (ii) Then pass the ground-truth word from the training sat to the next time-step. Of use the probabilities countly the filt No juck the highest probability word for that time-step as g^{***} ?. (ii) Then pass the selected word to the next time-step. Of use the probabilities countly the filt No juck the filt of pass the selected word to the next time-step as g^{***} ?. (iii) Then pass the selected word to the next time-step. Violate the probabilities countly the filt No juck the filt No juck the filt of the pass the selected word to the next time-step. Violate the selected word to the next time step. C you are training an RNN, and find that your weights and activations are all taking on the value of NaN ("Not a Number") which in these as the most filely cause of this problem? Variating gradient problem. Disposition of the selected problem. Disposition of the selected problem. Signoid activation finction gij used to compute gij, where a is too large. Suppose you are training a LSTM. You have a 10000 word vocabulary, and are using an LSTM with 100-dimensional activations at f^* . What is the dimension of f^* , at each time step? In a GW ₁ (c^* -to- f^* -y, c^* -to- f^* -to-	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	What are you doing at each time step f ? (I) Use the probabilities output by the RNN to pick the highest probability word for that time-step as $g^{+(2)}$, (ii) Then past the ground-study word from the training est to the next time-step. (I) Use the probabilities country to the RNN to randomy senips a chosen word for that time-step as $g^{+(2)}$, (ii) Then pasts the ground-study word from the training set to the next time-step. (I) Use the probabilities country the RNN to randomy senips a chosen word for that time-step as $g^{+(2)}$, (ii) Then pasts this salested word to the next time-step. (I) Use the probabilities country the RNN to randomy seniple a chosen word for that time-step as $g^{+(2)}$, (ii) Then pasts this salested word is the next time-step. (II) You see training an RNN, and find that your neights and activations are all taking on the value of fickly (in) take the rest view. (IV) A salesting gooders problem: (IV) A salesting gooders problem: (IV) A salesting gooders problem: (IV) Exploring gooders	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	What are you doing at each time step f ? (ii) Use the probabilities caugus by the Ribit to pick the highest probability word for that time-day as $g^{*(G)}$, (iii) Then pass the ground such word from the training set to the next time step. (iii) Use the probabilities caucus by the Ribit to analysis to the next time-day. (iii) Use the probabilities caucus by the Ribit to analysis to the next time-day. (iii) Use the probabilities caucus by the Ribit to analysis to the next time-day. (iii) Use the probabilities caucus by the Ribit to analysis to the next time-day. (iii) Use the probabilities caucus by the Ribit to analysis and activations are all taking on the value of filtre ("To take pass this selected word to the next time-day. (iii) Use the probabilities caucus by the Ribit to analysis and activations are all taking on the value of filtre ("Not a filtrember") which of these to the most likely cause of the problem? (iv) Varioting gradient problem. (iv) Expedient gradient problem. (iv) Expedient gradient problem. (iv) Expedient gradient problem. (iv) Expedient gradient gradient. (iv) Expedient gradient gra	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	(i) (1) Use the provious diese couput by the BNN to pack the highest probability word for that time-atop as g ^{1/2} , (iii) Then pass the ground study word form the saving acts to be next mere step. (i) (1) Use the provious direct word from the saving acts to be next mere step. (ii) (1) Use the productions couput by the BNN to condenly again access were for that time-atop as g ^{1/2} , (iii) Then pass the ground-study word from the saving act to the next time-step. (ii) (1) Use the productions coupled by the BNN to condenly sample a chosen word for that time-atop as g ^{1/2} , (iii) Then pass the selected word to the rest time-step. (iii) (1) Use the productions coupled by the BNN to condenly sample a chosen word for that time-atop as g ^{1/2} , (iii) Then pass the selected word to the next time step. (iv) (1) Use the productions coupled by the BNN to condenly sample a chosen word for that time-atop as g ^{1/2} , (iii) Then pass the selected word to the next time step. (iv) (1) Use the production coupled by the BNN to condenly sample a chosen word for that time-atop as g ^{1/2} , (iii) Then pass the production of the selected word to the next time step. (iv) (1) Various gradeer problem. (iv) Various gradeer problem. (iv) Lead active form function (g) used to compute gg3 where a to too large. (iv) Suppose you are training a LSTM. You have a 10000 word locability, and are using an LSTM with 100-dimensional activities and the selection of the selection	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	What are you doing at each time stop D (i) (I) Use the probabilities coupled by the BIN to pick the highest probability word for that time also as giffer, (ii) Then pass the ground-stank word from the staining at the bankst in-deby. (ii) Use the probabilities coupled by the BIN to pack the highest probability word for that time also as giffer, (iii) Then pass the ground-stank word from the signing at the bankst fine date). (ii) Use the probabilities coupled by the BIN Love downly passed by the state of that time also as giffer, (iii) Then pass the selected word to be constituted date. (iii) Use the probabilities coupled by the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then pass the selected word to the next time-also. (iii) Use the probabilities coupled by the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then pass the selected word to the next time-also. (iii) Use the probabilities coupled by the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then the BIN Love downly sample a chosen word for that time-also as giffer, (iii) Then the BIN Love downly sample as chosen as giffer, (iii) Then the BIN Love downly sample as chosen as giffer, (iii) Then the BIN Love downly sample as the second secon	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	(i) (1) Use the probabilities computely the BRN to pick the inglest probability word for that time-dep as g ^{1,0+} , (6) Then pass the ground-drift word form the tearing set to the rest time-dep. (ii) (1) Use the probabilities couply by the BRN to pack the inglest probability word for that time-dep as g ^{1,0+} , (6) Then pass the ground-drift word from the tearing set to the rest time-dep. (iii) (1) Use the probabilities couply by the BRN to randomy surple as chosen word for that time-dep as g ^{1,0+} , (6) Then pass the ground-drift word from the time-dep and g ^{1,0+} , (6) Then pass the ground-drift word from the time-dep and g ^{1,0+} , (6) Then pass the selected word to the rest time-dep. (iii) (1) Use the probabilities couply by the BRN to rendomly sample a Chosen word for that time-dep as g ^{1,0+} , (6) Then pass the selected word to the rest time-dep. (iii) (1) Use the probabilities couply by the BRN to rendomly sample a Chosen word for that time-dep as g ^{1,0+} , (6) Then pass the selected word to the rest time-dep. (iii) (1) Versioning groof problem. (iv) Use the probabilities couply by the BRN to rendomly sample a Chosen word for that time-dep as g ^{1,0+} , (6) Then pass the selected word to the rest time-dep. (iv) Use the probabilities couply by the BRN to probability word for that time-dep as g ^{1,0+} , (6) Then pass the selected word to the rest time-dep. (iv) Versioning groof problem. (iv) Use the probabilities couply by the BRN to probability word for that time-dep as g ^{1,0+} , (6) Then pass the selected word to the probability word for the time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the probabilities couply by the BRN to rest time-dep. (iv) Use the pr	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	What are you doing at each time step in (ii) the the proposition output by the infoll to pols the highest probability word for that time-step as g ⁺⁽⁰⁾ , (ii) then probability and for that time-step as g ⁺⁽⁰⁾ , (iii) then pass the ground-studh word from the staining at the fine rection rection change on the fine state ground-studh word from the staining at the fine rection rection ground the pass the ground-studh word from the staining at the rection rection in the state of the time-step as g ⁺⁽⁰⁾ , (iii) then pass the decision word to be most time stay. (iii) the the production about of the fine rection rection, sample is chosen word for that time-step as g ⁺⁽⁰⁾ , (iii) then pass the selected word to the most time-step. (iv) the time production of the fine state of the problem? (iv) the time state of the state of the problem? (iv) to an extended word to the most time-step. (iv) to an extended word to the most time-step. (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the state of the problem? (iv) to select state of the stat	1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	(i) Use the probabilists output by the RNH to pick the highest probability word for that since date as gifts (ii) the possible growd such word for firm the destinate at the restination of the state	1/1 point. 1/1 point.	Due Feb 3, 3:59 PM CST
	Recurrent Neural Networks Graded Quiz + 30 min	 What are you doing at each time stop to 10 on the photo process this word for that time atop at giffs, (iii) Then pass the provide form for the time are given to the rest time atop at giffs, (iii) Then pass the provide form for the time are given to the rest time atop. If the time pass shall be controlled for the time atop as giffs, (iii) Then pass the provide form for the time atop as giffs, (iii) Then pass the provide form for the time atop. If the time pass shall be controlled for the time atop. If the time pass shall be controlled for the time atop. If the time pass shall be controlled for the time atop. If the time are gift the time atop at gift time. If the time are timened an interface of the problems. The timened pass of the time atop at gift time atop at gift time. If the timened pass are problems. If the timened problems. If the timened problems. If the timened problems. If the timened problems are timened passed problems. If the timened passed problems. If the timened problems are timened passed problems. If the timened passed problems. If the timened problems are timened passed problems. If the timened problems are training a LETIA. You have a 1600 word forecholders, and are using an LETIA with 100 dimensional activations are training a LETIA. You have a 1600 word forecholders, and are using an LETIA with 100 dimensional activations are training a LETIA. You have a 1600 word forecholders, and are using an LETIA with 100 dimensional activations are training a LETIA with 100 dimensional activations are training a LETIA with 100 dimensional activation are training at the passed representation are training at LETIA. If the area is training a LETIA with 100 dimensional activation are using a training a LETIA with 100 dimensional activation are using a controlled p	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CS
	Recurrent Neural Networks Graded Quiz + 30 min	What are you comp at each time stop (!) (i) the the probabilisms output by the BBBI to put the highest probability and for that time-step, any ****, (ii) Then put the time ground stop and the stop time time the stop time time the stop time time time time time time time time	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CST
	Recurrent Neural Networks Graded Quiz + 30 min	What are your during at each time step (7) (9) to see the presentations exposed by the North Spice in the register of the section of the se	1/1 point 1/1 point 1/1 point	Due Feb 3, 3:59 PM CS