National University of Computer and Emerging Sciences

School of Computing

Spring 2024

Islamabad Campus

AI3003 ANN

Tuesday, May 21, 2024

Course Instructor

Mr. Hassan Raza

Final Exam

Total Time: 3 Hours

Total Marks: 145

Student Name	Roll No	Section	Signature
	Instructions		
 Return the question paper with 	the answer sheet after attemp	ting.	
• Must fill the answers(A,B,C			at the end of questions.
Question I Check the correct box (or boxes	for MCQs. Overwriting will	result in a zero	
(a) (1 Mark) Which of the fol	lowing is an application of NI	N (Neural Netw	ork)?
O Data validation			
Sales forecasting	-		
Risk managemer			
All of the mention			
(b) (1 Mark) What are dendri	tes?		
fibers of nerves	ma		
nuclear projection other name for n			
none of the men			
(c) (1 Mark) The fundamenta			
brain	t time of not work to		
O nucleus			
() neuron			
o axon	_		
(d) (1 Mark) What is shape o	f dendrites like		
○ oval			
○ round			
(V) "Tree			
rectangular			
(e) (1 Mark) Function of den	drites is?		
○ receptors			
○ both receptor &	transmitter		

	one of the mentioned
(f)	(1 Mark) What is purpose of Axon?
	○ transmitter
	○ transmission
	one of the mentioned
(g)	(1 Mark) On what parameters can change in weight vector depend?
	O learning parameters
	O'input vector
	○ learning signal
	all of the mentioned
(h)	(1 Mark) What is generalization?
	the ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the test set.
	The ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the training set.
	can be either way
	one of the mentioned
(i)	(1 Mark) Back propagation is a learning technique that adjusts weights in the neural network by propagating weight changes.
	O Forward from source to sink
	Backward from sink to source
	O Forward from source to hidden nodes
	Backward from sink to hidden nodes
(j)	(1 Mark) Identify the following activation function:
	$f(V) = Z + (1/1 + \exp(-x * V + Y)) Z, X, Y \text{ are parameters.}$
	Step function 7 + 1 + ePo
	C Kamp function
	Sigmoid function
	○ Gaussian function ②
(k)	(1 Mark) A neuron with 3 inputs has the weight vector [0.2 -0.1 0.1]T and a bias $\theta = 0$. If the input vector
	is $X = [0.2 \ 0.4 \ 0.2]T$ then the total input to the neuron is:
	\bigcirc 0.20 \bigcirc \bigcirc 2.
	♦ 0.02
4.	○ -1.0
(1)	(1 Mark) Why is the XOR problem exceptionally interesting to neural network researchers?
	Because it can be expressed in a way that allows you to use a neural network
	Because it is complex binary operation that cannot be solved using neural networks
	 Because it can be solved by a single layer perceptron Because it is the simplest linearly inseparable problem that exists.
/ N	
(m)	(1 Mark) What is back propagation?
	O It is another name given to the curvy function in the perceptron O It is the temperature of arrow healt through the naturally to adjust the inputs.
	It is the transmission of error back through the network to adjust the inputs It is the transmission of error back through the network to allow weights to be adjusted so that the
	network can learn.

	0	None of the mentioned
(n)		A perceptron adds up all the weighted inputs it receives, and if it exceeds a certain value, it outputs wise it just outputs a 0.
	0	True
	0,	Palse
	Ø	Sometimes – it can also output intermediate values as well
	~	Can't say
(o)	(1 Mark)	What is an Artificial Neural Network (ANN)?
	Θ	A computational model inspired by the human brain b) A machine learning algorithm used for image processing c) A statistical analysis technique for data clustering d) A programming language for neural network implementation
(p)	(1 Mark)	What is the basic building block of an Artificial Neural Network?
	0	Neuron b) Activation function c) Gradient descent d) Loss function
(q)	(1 Mark)	Which of the following activation functions is commonly used in ANNs?
	\circ	ReLU (Rectified Linear Unit)
	_	Sigmoid
		Tanh (Hyperbolic Tangent)
	_	All of the above
(r)		What is the purpose of the activation function in an ANN? It determines the output of a neuron
	-	It introduces non-linearity to the network
	-	It enables the network to learn complex patterns
		All of the above
(s)	(1 Mark)	What is the function of the input layer in an ANN?
. ,		It receives input data and passes it to the hidden layers
	0	It performs mathematical computations on the input data
	0	It stores the trained weights and biases of the network
	\sim	None of the above
(t)) Which layer of an ANN is responsible for making predictions or producing the final output?
		Input layer
	_	Hidden layer
	4	Output layer All layers contribute equally
<i>(</i>)		What is the purpose of the backpropagation algorithm in ANN training?
(u)		To update the weights and biases based on the prediction error
		To initialize the weights and biases of the network
	0	To determine the number of hidden layers and neurons
	Ŏ	None of the above
(v)	(1 Mark) Which of the following is a common loss function used in ANNs for binary classification?
		Mean Absolute Error (MAE)
	C	Mean Squared Error (MSE)
	0	Binary Cross-Entropy
	\circ	Categorical Cross-Entropy
(w)	(1 Mark	What is the purpose of the forward pass in ANN training?
	(To compute the predicted output based on the current weights and biases

To adjust the majohts and bioses using andiant descent
 To adjust the weights and biases using gradient descent To identify misclassified samples and update the model
None of the above
· · · · · · · · · · · · · · · · · · ·
(x) (1 Mark) What is the primary goal of training an ANN?
To minimize the prediction error on the training data
To maximize the number of neurons in the hidden layers
To achieve 100% accuracy on the test data
O None of the above
(y) (1 Mark) Which of the following is a common optimization algorithm used in ANN training?
○ Gradient Descent
Stochastic Gradient Descent (SGD)
Adam
All of the above
Question II(25 Marks)
(a) (1 Mark) What is the purpose of regularization in ANN training?
To prevent overfitting by adding a penalty term to the loss function
To increase the model's capacity for learning complex patterns
O No speed up the training process by adjusting the learning rate
O None of the above
(b) (1 Mark) What is the vanishing gradient problem in ANNs?
When the gradients become extremely small during backpropagation
When the gradients become extremely large during backpropagation
When the weights and biases are initialized randomly
O None of the above
(c) (1 Mark) Which type of ANN architecture is used for processing sequential data?
© Recurrent Neural Network (RNN)
O Convolutional Neural Network (CNN)
Multilayer Perceptron (MLP)
Radial Basis Function Network (RBFN)
(d) (1 Mark) What is the purpose of dropout regularization in ANN training?
To randomly disable neurons during training to prevent overfitting
O To increase the learning rate for faster convergence
To add additional layers to the network for increased capacity
O None of the above
(e) (1 Mark) Which of the following is an advantage of using ANNs for pattern recognition?
Ability to learn from large amounts of data
Robustness to noise and variations in input
Scalability to handle complex tasks
All of the above
(f) (1 Mark) What is the purpose of cross-validation in ANN training?
To evaluate the generalization performance of the model
To split the data into training and test sets
O None of the above

(g) (1 Mark) Which type of ANN architecture is commonly used for image classification tasks?
Convolutional Neural Network (CNN)
Recurrent Neural Network (RNN) Recurrent Neural Network (RREN)
Radial Basis Function Network (RBFN) Makilawa Basasatan (MLP)
Multilayer Perceptron (MLP)
(h) (1 Mark) What is the purpose of weight initialization in ANN training?
To set the initial values of the weights and biases in the network
 To adjust the learning rate during training To compute the gradient of the loss function
None of the above
(i) (1 Mark) Which activation function is commonly used in the output layer for binary classification in ANNs?
Sigmoid
ReLU (Rectified Linear Unit)
Tanh (Hyperbolic Tangent)
Softmax
(j) (1 Mark) What is the purpose of learning rate scheduling in ANN training?
To adjust the learning rate during training for better convergence
To increase the number of epochs for longer training
O To shuffle the training data between epochs
None of the above
(k) (1 Mark) Which of the following techniques can be used to prevent overfitting in ANN training?
O Dropout regularization
○ L1 and L2 regularization
○ Early stopping
All of the above
(l) (1 Mark) What is the purpose of the bias term in an ANN?
To provide a threshold for neuron activation
To add an additional feature to the input data
 To prevent overfitting by adjusting the learning rate
None of the above
(m) (1 Mark) What is the purpose of momentum in the optimization algorithm used for ANN training?
To accelerate the convergence of the algorithm
O To prevent overfitting by regularizing the model
O To adjust the learning rate during training
None of the above
(n) (1 Mark) What is the purpose of a validation set in ANN training?
O To tune the hyperparameters of the model
To evaluate the model's performance during training
O News of the above
None of the above
(o) (1 Mark) Which type of ANN architecture is commonly used for natural language processing tasks?
Recurrent Neural Network (RNN) Convolutional Neural Network (CNN)
Convolutional Neural Network (CNN) Multilayer Perceptron (MLP)
Radial Basis Function Network (RBFN)
C Radial Dasis I microll Letwork (KD11)

	(1 Mark) What is the purpose of mini-batch training in ANN training? To update the model's weights and biases after processing a subset of the training data To reduce the computational complexity of the training process To increase the learning rate for faster convergence None of the above
(q)	(1 Mark) What is the main advantage of using deep neural networks compared to shallow neural networks? Ability to learn hierarchical representations of data Faster convergence during training Lower computational complexity None of the above
(r)	(1 Mark) Which technique is used to initialize the weights of a deep neural network layer by layer?
	Xavier/Glorot initialization
	O Particular of the control of the c
	 Random initialization None of the above
(0)	
(8)	(1 Mark) Which activation function is commonly used in the hidden layers of a neural network to introduce non-linearity?
	○ Sigmoid
	ReLU (Rectified Linear Unit)
	○ Tanh (Hyperbolic Tangent)
	○ Linear
(t)	(1 Mark) What is the purpose of the activation function in a neural network?
	It determines the number of layers in the network
	It normalizes the input data
	Of the natural section of the network
()	It controls the learning rate of the network (1.Mark) Which appeals a transfer and for hardling accuration data such as natural language.
(11)	(1 Mark) Which neural network architecture is used for handling sequential data, such as natural language processing or time series analysis?
	Feedforward Neural Network (FNN)
	O Convolutional Neural Network (CNN)
	Recurrent Neural Network (RNN)
	O Radial Basis Function Network (RBFN)
(v)	(1 Mark) Which neural network architecture is commonly used for image classification tasks?
	O Feedforward Neural Network (FNN)
	Convolutional Neural Network (CNN)
	Recurrent Neural Network (RNN)Radial Basis Function Network (RBFN)
(111)	
(w)	(1 Mark) Which algorithm is used for updating the weights in a neural network during the training process? Backpropagation
	O Gradient Descent
	Stochastic Gradient Descent (SGD)
	All of the above
(x)	(I Mark) What is the purpose of the bias term in a neural network?
	Old It controls the learning rate of the network
	It adds flexibility to the decision boundaries of the network

It introduces non-linearity in the network	
 It allows shifting the activation function 	
(y) (1 Mark) Which algorithm is used for updating the weights in a neural network with a single training example at a time?	
Backpropagation	
○ Gradient Descent	
Stochastic Gradient Descent (SGD)	
Mini-batch Gradient Descent	
Question III)
(a) (1 Mark) Which technique is used for preventing overfitting in a neural network by randomly dropping ou neurons during training?	
() Dropout	
Batch Normalization	
○ L1 Regularization	
○ L2 Regularization	
(b) (1 Mark) What is the purpose of the loss function in a neural network?	
It measures the accuracy of predictions	
lt measures the complexity of the model	
It quantifies the difference between predicted and actual values	
 It controls the learning rate of the network 	
(c) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the previous	S
weight update?	
Backpropagation through time (BPTT)	
Momentum Propagation (MProp)	
Levenberg-Marquardt Algorithm	
Quickprop	ie
(d) (1 Mark) Which neural network architecture is used for handling both sequential and spatial data, such a video processing or 3D image analysis?	
○ Feedforward Neural Network (FNN)	
Oconvolutional Neural Network (CNN)	
Recurrent Neural Network (RNN)	
Long Short-Term Memory (LSTM) Network	
(e) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the directic of steepest descent?	n
O Backpropagation	
Gradient Descent	
Momentum Gradient Descent	
Newton's Method	
(f) (1 Mark) What is the purpose of the learning rate in a neural network?	
It controls the speed of convergence during training	
It determines the number of hidden layers in the network	
It introduces non-linearity in the network	
It allows shifting the activation function	
(g) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient?)TI

Gradient Descent Adam Optimization Adaboost (h) (1 Mark) Which state of the art neural network architecture is used for handling variable-length sequential data, such as text generation or machine translation? Feedforward Neural Network (RNN) Corrolutional Neural Network (RNN) Recurrent Neural Network (RNN) Transformer Network (i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L2 Regularization L2 Regularization Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (RNN) Corrolutional Neural Network (RNN) Recursive Neural Network (RNN) Recursive Neural Network (RRFN) Dropout Batch Normalization L2 Regularization L1 Regularization Recursive Neural Network (RNN) Radial Basis Function Network (RBFN) Batch Normalization L2 Regularization R1 (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization R1 (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Corrolutional Neural Network (FNN)		
Adahoost (h) (1 Mark) Which state of the art neural network architecture is used for handling variable-length sequential data, such as text generation or machine translation? Feedforward Neural Network (FNN) Convolutional Neural Network (RNN) Recurrent Neural Network (RNN) Transformer Network (i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L1 Regularization L2 Regularization (j) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Baskpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (RSPN) Radial Basis Function Network (RBNN) Radial Basis Function Network (RBNN) Batch Normalization Dropout Batch Normalization Dropout Batch Normalization Dropout Batch Normalization Radial Basis Function seed on the weights? Dropout Batch Normalization Dropout Batch Normalization Radial Basis Function Seed or updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Batch Pormalization Radial Basis Function Seed on the weights? Dropout Batch Normalization Radial Basis Function Seed on the weights? Dropout Batch Normalization Radial Basis Function Network (RBN) Radial Basis Function Seed on the weights in a neural network by considering the direction of the negative gradient and the squared history? Batch Normalization Radial Rad		O Backpropagation
Adaboost (h) (1 Mark) Which state of the art neural network architecture is used for handling variable-length sequential data, such as text generation or machine translation? Feedforward Neural Network (CNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Transformer Network (i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L2 Regularization L2 Regularization C3 (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (RNN) Recursive Neural Network (RNN) Resilient Descent Dropout Batch Normalization L2 Regularization RMSPror (n) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (KNN) Convolutional Neural Network (KNN) Convolutional Neural Network (KNN) Convolutional Neural Network (KNN)		
th) (1 Mark) Which state of the art neural network architecture is used for handling variable-length sequential data, such as text generation or machine translation? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (CNN) ○ Recurrent Neural Network (RNN) ○ Transformer Network (I) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? ○ Dropout ○ Batch Normalization ○ L1 Regularization ○ L2 Regularization (I) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? ○ Backpropagation through time (BPTT) ○ Resilient Propagation (RProp) ○ Levenberg-Marquardt Algorithm ○ RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (RNN) ○ Recursive Neural Network (RNN) ○ Radial Basis Function Network (RNN) ○ Radial Basis Function Network (RNN) ○ Dropout ○ Batch Normalization ○ L2 Regularization (m) (1 Mark) Which technique is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? ○ Backpropagation ○ Gradient Descent ○ Conjugate Gradient ○ RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (FNN)		•
data, such as text generation or machine translation? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Transformer Network (i) (I Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout		· ·
Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Transformer Network (i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L1 Regularization L1 Regularization L2 Regularization Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (RNN) Redula Basis Function Network (RNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)	(h)	
Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Transformer Network (i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L1 Regularization L2 Regularization L2 Regularization L3 Response at the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (RNN) Recursive Neural Network (RENN) Readial Basis Function Network (RBFN) (l) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)		The production of the contract
Recurrent Neural Network (RNN) Transformer Network (i) (I Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dopout Batch Normalization LI Regularization LI Regularization LI Regularization (i) (I Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (I Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (RNN) Convolutional Neural Network (RNN) Radial Basis Function Network (RBFN) (I) (I Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (I Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (I Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)		
(i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L2 Regularization L2 Regularization Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Recursive Neural Network (RNN) Radial Basis Function Network (RBFN) (I) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization T2 Regularization (I) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)		
(i) (1 Mark) Which technique is used for normalizing the input data in a neural network to ensure similar scales across different features? Dropout Batch Normalization L1 Regularization L2 Regularization L2 Regularization U2 Regularization U3 (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (RNN) Radial Basis Function Network (RBFN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)		- Parameter and the state of th
across different features? Dropout		plant and the second se
Dropout Batch Normalization L1 Regularization L2 Regularization L2 Regularization L3 (i) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (ReNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization T2 Regularization RMSPror (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)	(i)	
Batch Normalization L1 Regularization L2 Regularization L3 Regularization L3 Regularization L3 Regularization (i) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization M (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		_
OL1 Regularization ○L2 Regularization ○L2 Regularization ○L2 Regularization ○L1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? ○Backpropagation through time (BPTT) ○Resilient Propagation (RProp) ○Levenberg-Marquardt Algorithm ○RMSPrp ○KMSPrp ○KMSPP		
O L2 Regularization (j) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? ○ Backpropagation through time (BPTT) ○ Resilient Propagation (RProp) ○ Levenberg-Marquardt Algorithm ○ RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (RNN) ○ Radial Basis Function Network (RPNN) ○ Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? ○ Dropout ○ Batch Normalization ○ L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? ○ Backpropagation ○ Gradient Descent ○ Conjugate Gradient ○ RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (FNN)		· ·
(i) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (FNN)		* *
of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (ReNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (l) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)	(J)	
Resilient Propagation (RProp) Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
Levenberg-Marquardt Algorithm RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (l) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
RMSPrp (k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (l) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~
(k) (1 Mark) Which neural network architecture is used for handling both sequential and hierarchical data, such as natural language parsing or speech recognition? Feedforward Neural Network (FNN)		
as natural language parsing or speech recognition? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (I) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization 12 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)	(k)	•
Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (I) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization 12 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)	(**/	
Recursive Neural Network (ReNN) Radial Basis Function Network (RBFN) (I) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
Radial Basis Function Network (RBFN) (I) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		Oconvolutional Neural Network (CNN)
(1) (1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term to the loss function based on the weights? Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		© Recursive Neural Network (ReNN)
to the loss function based on the weights? Dropout Batch Normalization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		Radial Basis Function Network (RBFN)
Dropout Batch Normalization L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)	(1)	(1 Mark) Which technique is used for preventing overfitting in a neural network by adding a penalty term
 □ Batch Normalization □ L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? □ Backpropagation □ Gradient Descent □ Conjugate Gradient □ RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? □ Feedforward Neural Network (FNN) □ Convolutional Neural Network (CNN) 		to the loss function based on the weights?
L2 Regularization (m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
(m) (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		=
of the negative gradient and the squared history? Backpropagation Gradient Descent Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		CL2 Regularization
O Backpropagation O Gradient Descent O Conjugate Gradient O RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? O Feedforward Neural Network (FNN) O Convolutional Neural Network (CNN)	(m)	
 ○ Gradient Descent ○ Conjugate Gradient ○ RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (CNN) 		Barryy
Conjugate Gradient RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN)		
 RMSPror (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) 		
 (n) (1 Mark) Which neural network architecture is used for handling both sequential and non-sequential data, such as sentiment analysis or document classification? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) 		
such as sentiment analysis or document classification? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (CNN)		
Feedforward Neural Network (FNN)Convolutional Neural Network (CNN)	(n)	
Oconvolutional Neural Network (CNN)		
() Recurrent Neural Network (RININ)		Recurrent Neural Network (RNN)
Transformer Network		

serial No 055118

Date



Ple	ease Tick (() Car	npus:	CFD		KH		PWR	Seme	ster:	SP	SU FA 20
	•				FINAL	EXAM	ANSW	ÆR BO	OK			
Co	urse Code &	Title:		1114	icia	1	Ne	v sel	-	Ne	+wo~	/ h &
Rol	l No:		_ Sect	tion:	·····		udeni Çnatu:		"	Nost		Date:
con	ial No. of tinuation et if attached	ú:			C	otal No of Extra ets Us)		vigilato ignatu		· ·	
			W	(f	HIS ANS	SWER BO)(e) K (c(e))	Traine i	AGE 148)		
			ON NOT	OPEN T						TL INST	RUCTED	·
						Insti	4					÷
	Please ensure there may a c				thresho	d is fre	e of an	y matei	rial clas	sified a	s 'useful	in the paper or else
,		essary,	as neith	er the								Make assumptions ur queries or provide
	Fit in all your mark questio								tra she	et if rec	quired. If	you do so, clearly
	Use only your your own cald							d by yo	ur teach	ier/exa	miner). I	f you do not have
	Use only perr Any part of pa									anent ir	ık pen w	ill be considered.
6. E	Ensure that yo	ou do no	ot have	any ele	ctronic	gadget	(like m	obile of	ione. sn	nart wa	itch. ear	buds etc.) with you.
7.)uestio:	n Paper	along v								ed) to the invigilator
	Q./Part No.	Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Q-7	Q-8	Q-9	Q-10	
	Total Marks											Total Marks
	Obtained Marks	2	2-2	-26	ĮÙ	10	10	-17	08			The second secon
	CLO NO.											Total Marks
			J		F		· ·		t			Obtained

Examiner/Course Teacher

/Part No.	Page No . 02	Rough Work
STATE OF THE STATE		
garja pa gengada a feed maren võudarelistus) marendik (
25/3/Similardo-Equificação		and geographic photos and an initial state of the management geographic contribution of the contribution o
SCON PARTICULAR VALIGURAS ANNO UNAS ANNO UNICO		ang granggangganggang an an at think
gen a un no massa famour a anno ana de rado britán		in a second seco
Misser Malaceteraissen values anno anno		and the proposation of monotonical stability of the stabi
		THE PROPERTY OF THE PROPERTY O
Clare I Complete Coloredo SE ESCOCIO PER FOCUENCIO.		A STATE OF THE STA
etember destatement of of all limited to deal and other and all the consti-		
CALO LLAMANACIONOS ATÚMICA ARUMANA		
5504V-0004V-0004V-0004V-0000		The state of the s
To the Control of the		The state of the s
[60]-60[control of the control of t		
ann tha ann an t-mailting de ann an ann ann an an		Annual control of the
		The state of the s
van de se met districte avant district de communication de districtiva de		A CONTRACTOR OF THE PROPERTY O
undri un houselour Harristel (948 brokken) (966 brokken) Ab		A LINEAR CANADA
meetikatiinissä kätiinin kun keen on tekkikiiniste		
Ann and Street S		
vigras por nasopon mesesa i seridam polonicinali el		
но-ынованения сечало сечало		
THE CANCELLAND CONTRACTOR OF CONTRACTOR CONT		
The behavior day and instrument memory memory and the		The second secon
mpfed derdomhile-friedrich das der Albeiterbessen von		
and all animomens continues of the animomen of		
PHICAGOGORIANTARIA		

t No. Page	No.03 Rough	Woi
		7710
		HIPPOWERSHAM
		di Vindo de di
	- 2 (he(x ^a))-y ^a)	BEET NOOTHEE HAVE ALL
		20000000000000000000000000000000000000
and the same and t	Lander Comment of the	www.yorwyorwyy
	(O) + O) (3)	erc vettimine e p. 11.
		adrate/ade/access
	(8, 10, 27-19) 27-200	
ta turnak jajan usenja ja j	29-4) + (8+29-1)+	
	40, 3) F(6, -1)]	00thris anne
		mossiumaasium monesiuma
		10000000000000000000000000000000000000
The second secon	All Committee and the committee of the c	menosanan.

Part No.	Page No . 04	Rough Wor
	dJ = 2 / 0 + 30 - 4) x3 + (0 , + 20, -1) 2	A CONTRACTOR OF THE PROPERTY O
		The state of the s
	+ (0,+40,-3)+4+(0,-1)x0)	
		A Superior Control of the Control of
A STATE OF THE PARTY OF THE PAR	- 36, +90, -19 + 200+40, -4	The state of the s
	2 +40, +160, -12	ACT LANGE OF THE PARTY OF THE P
		A CONTRACTOR OF THE PROPERTY O
	- 4.5 Bo + 14/5 B, -13 -> (B)	777777700 m 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
		The state of the s
· ·	Man Q and B	Annual Communication of the Co
1	4.56, 414.50, -13 = 0	THE PROPERTY OF THE PROPERTY O
	$\frac{4.56. + 14.50 13 = 0}{2.60. + 4.50 4.5 = 0}$	TO THE PROPERTY OF THE PROPERT
	Mushiply A with =2.25	**
The state of the s	<u> </u>	
WERRITE CONTROL TO THE PROPERTY OF THE PROPERT	- 4. CB, - 10.13 CB, +10.126=0	
		and the second s
	4.3750, -935-166E0	
And the state of t	0. = 0.66	
	and the second s	

rt No.	Page No . 05	Rough Wor
		N F F (1980) (1981) - To Andrews (1980) (198
45 M Million had windows by January and water some water specific process of the control of the		CONTROL OF THE STATE OF THE STA
PP (2 MMC) A STATE of The College of	<u> </u>	
PP-PS-CCCC CONTROL OF STATE OF		
	<u> </u>	
		1730 kinkiskal 1831 - 1832 aradisani kabbaki kilakis abbandansa yansar saratan sara -
	<u> </u>	
	10 - 0.13 A 4.6((C)	
in a construction of the c	us conference and control of the con	
SHAN IS Some Pallary and Angel Pallary and Annual Three Andreas Andrea		
		Transference (Company)
		And the state of t
		NOW NOW THE RESIDENCE AND ADMINISTRATION OF THE PROPERTY OF TH
		To the second se
Policinia		A CONTRACTOR OF THE CONTRACTOR
	the second secon	

/Part No.	Page No . 06	Rough Work
-		
	has 6(0.0) + 2	e e transporte de la companya del companya del companya de la companya del la companya de la com
	2 2	
		a de la companio del companio de la companio del companio de la companio del companio de la companio de la companio de la companio del companio de la companio del compa
	162) = ((0 10101)+2 · · ·	
	<u> </u>	
	<u> </u>	
gardys at the state of the section o		
n, to 17 fermines of majories before 17 maj to 18 fermines and a superproduction and the		The state of the s
	= 20·18 18	
ANAMATAN INTO THE REPORT OF STREET ST		
en e		na i norma (ili) - Panaminor ami reliment i normi sa araa pari Ili) - Panaminor ami reliment i normi sa araa pari
	- 247424	
enemicke volg Francosomore sprant og sprant visualenska		
rojectorny i indian narodoka kommunika sumana kamanan narodoka kommunika sumana kamanan kamanan kamanan kamana		
ordering an execution of the second		
Service of the Service Service of Children's Fore a surgery. Superior manifestations	and the second of the second o	
on extra and a second entertainment of the second entertai	one and the contract of the co	
		can wante grant institution on the Abbertal Mills (Methods institution)
		·

÷		
Part No.	Page No . 07	Rough Work
	<u> </u>	
	<u> </u>	
	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
	and the second s	

·		
/Part No.	Page No . 08	Rough Work
wind this billion of contract from the billion of	(3) = 10 (2.02020) + ((g.,20202)+2	
William Comment of Market Comm	54.935273	
A CONTRACTOR OF THE CONTRACTOR		
The second secon	4 (h) -10 (3.020303) +6 (3.030303)+2 2 112 · 44 009 181	
		BOTH THE STATE OF
	L(C) = (0 (4.040404) +6 (6.040+7-4)+	
	= 189.49 1819	
TO THE RESIDENCE AND ASSESSED ASSESSED.		
The CALL TOURISM CONTROL OF A C		
III PORTINI IN TOTO PORTINI PO		
	the state of the s	
	<u> </u>	
		Manuscript Communication Commu
2000 erroratii iii isaanaa anta anta anta anta anta anta ant	- 64.64	
2042/YAA-MISSIONYASIONIESEN PALAMPINIS EPIRENDO ()		A Part of the Control
COMO REPORTED DE PROSENCIA DE TRANSPORTED DE TRANSP		
**************************************	- 129.92	
	$M \in \mathbb{C} \longrightarrow M = M = M = M = M = M = M = M = M = M$	
20000000000000000000000000000000000000	$MSE = \frac{1}{5} + \frac{1}{5} + \frac{3}{5} + \frac{3}{5} + \frac{2}{5} + \frac{1}{5} + \frac{2}{5} + \frac{3}{5} + \frac{1}{5} + \frac{2}{5} $	Control of the Contro
	4(-(469) + (-170.63	
gyndags frysiaigd fewyddydd ardiada dd dibrenn o mae'r y gyndigol ardiadad a dibrenn o mae'r y gyndigol add a dibrenn o fel y gyndigol a dibrenn o fel y gy	1307-66	
wice a treatment or common or treatment or transport or the control of the contro		Management of the state of the

t No.	Page No . 09	Rough Work
		annequementaurmeinssensiellensses i Sidd Million (1900) alli
	S MSE OF MODELL	
man sa	Less than that	The state of the s
are consecutive of the consecuti	Madel 2, We'll chose	
		gengengengen grann naven som en state en statistisk klade klade til 1800 for til
		paga pang pang paga sa 2 pang manananan na di di mbalifi d ^{a s} a 12 (^{an} ^{an} an 14
	La company of the second of th	
		The state of the s
		18 September 18 Se
The state of the s		
	<u> </u>	
The management of the state of	hypothesis = np. At (x, theta)	
	en e	
The land of the la	and the second of the second o	
gen general (g g g) and a g g g g g g g g g g g g g g g g g g	Light No Syn Correr	
i (sprks)st meerik si		Planting the state of the state
	<u> </u>	

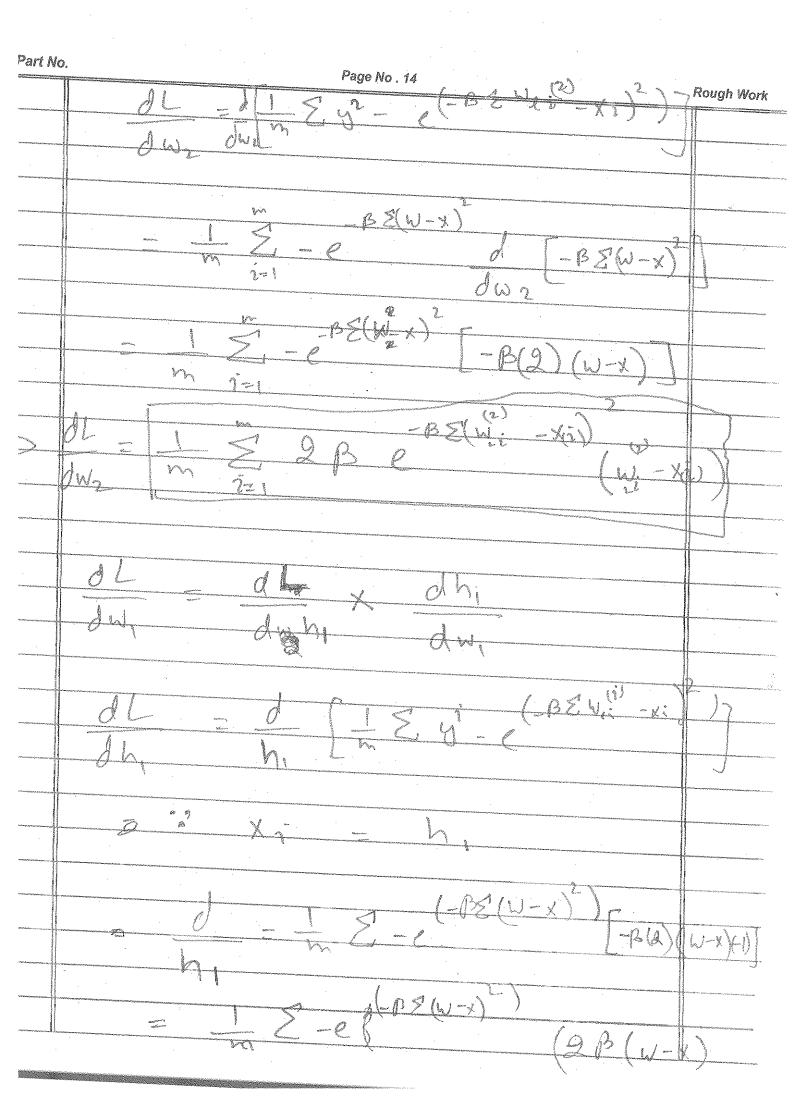
% /5% A ?	Ph 63	Fine contains
Q/Part No.	Page No . 10	Rough Worl
		Antonio de Santo de S
	175 Assurant Mar = Oo+G, X	
Control to the contro		
	45 - 4 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2	
	and the state of t	
To all the second secon		The state of the s
The second section of the section		A Carlotte and A Carl
	1 = 1 5 1 5° - 6 7 6 3) x 1	
To the state of th	The state of the s	
to the state of th		
	07 / 2 / 1 2 / 1 0 9)	
many light of the control of the con	Total to the second	
The state of the s		Section 1
West Constitution of the C	= 2= 3 (9° : 0, -8, x)(-1)	
		The state of the s
The state of the s	7/1====================================	
A TANAN CARICAGAN AND AND AND AND AND AND AND AND AND A		I de all trans
		es que mentrenamentans a su es estante

'art No.	Page No . 11	Rough Work
	in (Viji)	and the contract of the contra
** 11,11*** \$\frac{1}{2}\$ 11,11** \$1		
	and the second s	
	Clef efficient forward (1, w, bas):	
	Lyfothesis VI Legg - XII	
	Ly Pathesis - 2 - DP SUN (hy Pathesis Manie)	
And Andreas and An	hydethesis-22 Meddin & bedg/2 hydeth	
	1/2 W = hy Po Hosis - 3 * (- f) * 1623 4	han Ald heri
	11 x = hybothesis 3 x (+2) x + 10 2 4 m (2 4 m) (2 4 m)	
	- Liwa (hapathesis a Janda 1)	
\$6.00 mm 1 m		A CANADA
		and the second section of the
methodore and development of the control of the con		ng et kaladalana karalijan kenda indo ominahnya kamane kaladasa kara
Manager (Applications of the Control		
The second secon		To you was a state of the state

.

art No.	Page No . 12	Rough Won
A CONTRACTOR OF THE PROPERTY O		
Sub-rich Strategy	Class Descriminator I manable :	All a property of the contract
The second secon	- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	
	- Carlotte C	AND
the second continuous productions of the second		
	- Self. Jones - Nn. Spacestill	
	100 1 100 100 100 100 100 100 100 100 1	
WILLIAM OF THE PARTY OF THE PAR	- LANGE LANGE TO THE TOTAL PROPERTY OF THE PARTY OF THE P	
		or my of particular and particular a
(Action of Children on Children	- Milinipología (N)	The state of the s
	- Land Defendance of the second secon	
The state of the s		
The state of the s		
	- Minder Andrew College Andrew	
THE REAL PROPERTY OF THE PERSON OF THE PERSO		
a control of the cont	TOTAL SUB- DOWN (1)	
2	Construction BCELOSCO	Special region and the state of the special state o
T-Ville in a principal in the second	OPENING CONTRACTORS ()	and from the first the first of
Without I Have been seen and the seen and th	offinited = SCO (NHO Peremetras)	
	date lecture to recholision date more (partition date production) and continue to a co	
	Pack 7 10	
		neconnect this commence to what place the statement and another

art No.	Page No . 13	Rough Wor
- Andrews		The state of the s
1 3 M (F - 2 M) - 1 3 M	- if the ing place in detailed	
70 - 10 - 10 - 10 - 10 - 10 - 10 - 10 -	out = neto(ing. Plateuro)	
AND AND LOAD TO THE PROPERTY OF THE PROPERTY O	Los Alexander Carlo Della	
The second secon	LASS: LES WWARD ()	A Carrier de Laboratorio de Laboratorio de Provincio de Provincio de Laboratorio
Annual Comment of the	<u> </u>	
	and the second s	
PROPERTY IN CONTRACT IN CONTRA		
	(Q VIII)	CANAGOS SANGOS S
A CONTRACTOR OF THE PROPERTY O		TO SETTLE FOR Submitted of the sale and the
		And distributed a summarism on the content of the polytopical poly
	Loss + 1 - (4 ² - 4.)	
PO 17 TO 17	overall 1055 = L-15(13"-13")2-	
	=> Assuming there are no 1	очення под 15 година на Адганова III на поста на принципа на принципа на принципа на принципа на принципа на п Принципа на принципа на пр
· · · · · · · · · · · · · · · · · · ·	activation functions involved	
TOTAL CONTRACTOR OF THE PROPERTY OF THE PROPER		
The state of the s		
And an information regarding the control of the con		r s statement fra silvent en
STATE OF THE PROPERTY OF THE P		II (ch-)



'Part No. Page No . 15 Rough Work

VPart No.	Page No . 16	
		Rough Work
manife of decomposition of the State of the		
on the control of the		Control of the contro
		A THE STREET OF
		A PROPERTY OF THE PROPERTY OF
		The state of the s
The state of the s		The second science of
The state of the s		Control of the Contro
7		The state of the s
Million of Stilling College Co		and the state of t
		Americal de la company de la c
April 14.2 provided by the supplemental state of the supplement of		and the second s
D-SMCDrawpp square		
TOTAL PROPERTY OF THE PROPERTY		The second secon
And the second of the second o		And the state of t
State Company of the State of t		Marine and American September 1985 (September 2) (Septembe
		Considerate Springer Springer Springer (Springer Springer
AND AND SOLD TO SOLD SOLD SOLD SOLD SOLD SOLD SOLD SOL		management of the section of the sec
		an ang distinct of the support of th
The second secon		· ·
		CONTINUED IN CONTINUED AND ADMINISTRATIVE AND ADMIN
The second secon		Sellengerica Elizarologorica committà del protecti New dissiporica del committà del protection del committa
and it is equivalently writing the following of the production of the Collection of Topics.		The state of the s
approxy communication for the first of the f		and making the section of the property of the section of the secti
		er Shift was the General Statement of particular of the Shift was the Shift of the
The state of the s		- Consequence - Production
And the state of t		
- Charles (American St.) Special Charles (Charles Charles) and the Special Charles (Charles Charles Charles) and the Special Charles (Charles Charles Charles) and the Special Charles (Charles Charles Charles) and the Special Charles (Charles Charles Charles Charles (Charles Charles Charles Charles Charles (Charles Charles Charles Charles Charles (Charles Charles		in and the state of the state o

(0)	(1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the momentum term?
	○ Backpropagation
	O Gradient Descent
	Stochastic Gradient Descent (SGD)
	Momentum-based Gradient Descent
(p)	(1 Mark) What is the purpose of the momentum term in a neural network?
	It controls the speed of convergence during training
	It introduces non-linearity in the network
	It allows shifting the activation function
	It helps accelerate the convergence and overcome local minima
(q)	(1 Mark) Which technique is used for preventing overfitting in a neural network by randomly selecting a subset of the training examples for each iteration?
	O Dropout
	Batch Normalization
	 ◯ L1-Regularization ⊘ Mini-batch Gradient Descent
	•
(T)	(1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient, momentum and directionless history?
	Backpropagation
	Gradient Descent
	Adam Optimization
	○ Adaboost
(s)	(1 Mark) What is the purpose of the early stopping technique in a neural network?
	Of the prevents the network from overfitting the training data
	On the Heavy chiffing the notivetien function.
	 It allows shifting the activation function It controls the learning rate of the network
4.1	
(1)	(4 % 4 1) 177 1 1 1 4 1 4 1 4 1 4 1 4 1 4 1 4 1 4
(.)	(1 Mark) Which neural network architecture is used for handling both sequential and spatial data, such as video processing or 3D image analysis?
(*)	video processing or 3D image analysis? ○ Feedforward Neural Network (FNN)
(*)	video processing or 3D image analysis? ○ Feedforward Neural Network (FNN) ○ Convolutional Neural Network (CNN)
(*)	video processing or 3D image analysis? O Feedforward Neural Network (FNN) O Convolutional Neural Network (CNN) Recurrent Neural Network (RNN)
	video processing or 3D image analysis? O Feedforward Neural Network (FNN) O Convolutional Neural Network (CNN) O Recurrent Neural Network (RNN) O Long Short-Term Memory (LSTM) Network
	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Tong Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update?
	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Tong Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT)
	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Long Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp)
	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Tong Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm
(u)	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Long Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm Momentum Optimization
(u)	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Long Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm Momentum Optimization (1 Mark) What is the purpose of the dropout technique in a neural network?
(u)	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Tong Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm Momentum Optimization (1 Mark) What is the purpose of the dropout technique in a neural network? It prevents the network from overfitting the training data
(u)	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Long Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm Momentum Optimization (1 Mark) What is the purpose of the dropout technique in a neural network? It prevents the network from overfitting the training data It speeds up the convergence of the network
(u)	video processing or 3D image analysis? Feedforward Neural Network (FNN) Convolutional Neural Network (CNN) Recurrent Neural Network (RNN) Tong Short-Term Memory (LSTM) Network (1 Mark) Which algorithm is used for updating the weights in a neural network by considering the direction of the negative gradient and the magnitude of the previous weight update? Backpropagation through time (BPTT) Resilient Propagation (RProp) Levenberg-Marquardt Algorithm Momentum Optimization (1 Mark) What is the purpose of the dropout technique in a neural network? It prevents the network from overfitting the training data

(w)	(1 Mark) Which algorithm is used for updating the weights in a neural network by consider of the negative gradient and quadratic derivative? Backpropagation Gradient Descent Conjugate Gradient	ering the directi	ion
()	 RMS Prop(1 Mark) Choose the correct option describing the features of Artificial neural network		
(X)			•
	 It is essentially machine learning algorithm. It is useful when solving the problems for which the data set is very large. 	Ô	Q
	They are able to extract features without input from the programmer.		ii palaanna,
	4. These are systems modeled on the human brain and nervous system		
	Choose the most appropriate answer from the options given below:	٨	
	All the statements are correct.	Q.	To Tamong
	Only 2 and 3 are correct.		
	Only 1 and 4 are correct.		
	All the statements are not correct.		
(y)	(1 Mark) Which Boolean operation on two variables can be represented by a single percent	ption layer?	s Mill
-	1. X1 AND X2		W.K.
	2. X1 OR X2		
	3. X1 NOR X2		m, O
	4. X1 XOR X2	n O	0 700
	Choose the most appropriate answer from the options given below:		yee,
	O 1 and 2 Only	.0 6	0
	2 and 3 Only) 4.	t
	① 1, 2 and 3 Only	10	
	(7.15.1.) Yes first the following apprintment of the following apprintment		
(z)	(5 Marks) In a feed forward neural network with the following specifications: Input layer has 4 neurons, hidden layer has 3 neurons and output layer has 2 neurons	•	ioid
	activation function for given input values [0.5, 0.8, 0.2, 0.6] as well as the initial weights for	or the connection	ons.
	W1: [0.1, 0.3, 0.5, 0.2]		
	W2: [0.2, 0.4, 0.6, 0.2] Input layer to hidden layer weights		
10	W2: [0.2, 0.4, 0.6, 0.2] Input layer to hidden layer weights W3: [0.3, 0.5, 0.7, 0.2] W4: [0.4, 0.1, 0.3] W5: [0.5, 0.2, 0.4] Hidden layer to output layer weights		l of 1
0631	W4: [0.4, 0.1, 0.3]		No.
£ * /	W5: [0.5, 0.2, 0.4] Hidden layer to output layer weights	XY	3M \
	What is the output of the output layer when the given input values are passed through neurathe answer to two decimal places:	H Hetwork: No	Dill
	$\bigcirc [0.62, 0.68]$ $\bigcirc [0.72, 0.78]$ $\bigcirc [0.82, 0.88]$		
	O [0.82, 0.88]	MXI	
	\bigcirc [0.92, 0.98]	3	0.521
~			
0.62	0.66 0.69) 3x2 0.51		0.718
	1,63	_ \	
and the second s	100	war and a second	
	143	124	
	Page 10 - test Ux	<u> </u>	
	1000 000	w- 1	

Question 1	Answer	Question 2	Answer	Question 3	Answer
a		a	0	a	a_
b	- OA	b 1-7	Δ	b	c
C garage	Came	c L	a	c	a ×
d /	C	d	a	d	d
e &	6	е :	d	e	b 1/2
f X	a	f townsend	<u>a</u>	f	2
g		g 1	A	g	6-
h ,		h	<u>a</u>	h	de
i		i 1	a	i	Ь
j	Con	j de marieman de m	- Q	j	d
k /	Con-	k \	à	k	San January Control of the Control o
1	1	1	α	1	
m	Com	m į	- Q	m	C7
n 💃	<u> </u>	n ×	6	n	d X
0	- a	0	a	O	
p	14	р	\	p	1
q /	11	q [a	q	d
r /	1	r hammen	<u>a</u>	r	C
s /	<u> </u>	S James S	0	S	19:
t /		L Immediately	<u>C</u>	t	0
u /	1 a	u L	Commen	u	d
v	Com	V	, ,	v	10.
w	Δ	W \		W	d
x	(C	x ×	6	х	191
у	10	у	C	у	C.
	1)	<u>l (29</u>	<u>D</u>	Z	G.
	no market and the second secon	Page 1	ĺ		(26)

Consider the problem of predicting how well a student does in his/her second semester of university, given how well they did in their first semester. Specifically, let x be equal to the number of "A" grades (including A-, A, and A+ grades) that a student receives in their first semester of college. We would like to predict the value of y, which we define as the number of "A" grades they get in their second semester. Refer to the following training set of a small sample of different students' performances. Here each row is one training example. Recall that in linear regression, our hypothesis is $h_{\theta}(x) = \theta_0 + \theta_1 x$, and we use m to denote the number of training examples.

×	У
3	4
2	1
4	3
0	1

Figure 1: Example Dataset

- (a) (7 Marks) What values of θ_0, θ_1 should we use. [Hint: Remember our goal is to find the parameters that minimize following cost function. $J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left(h_{\theta}(x^{(i)}) y^{(i)} \right)^2$. You may use simple Calculus rules to arrive at the solution]
- (b) (2 Marks) How many A's a student will get in his second semester, if he had 5 A's in his first semester.
- (c) (1 Mark) How many A's a student will get in his second semester, if he had 1 A's in his first semester.

$$J_{\theta} = \frac{1}{2m} \sum_{i=1}^{m} (y^{(i)} - h_{\theta}(x^{(i)}))^{2}$$

Here m is the number of training examples.

(a) (5 Marks) Now your goal is to write a function to compute the cost function for a given dataset.

```
def cost_function(self, X,Y, theta):

///

Computes the Cost function for given
input data (X) and labels (Y).

Input:

X: can be either a single n X d-dimensional
vector or n X d dimensional matrix of inputs
theta: must d X l-dimensional vector for
representing vectors
Y: Must be n X l-dimensional label vector

Return:
Returns the cost of hypothesis with input parameters

///
```

(b) (5 Marks) Write the equation of derivative of given cost function with respect to θ_1 and θ_0

A WYY	 (111 Marke)	
Omestion VI	 (TO TANDET SON)	

Consider the problem of predicting how well a student does in his/her second semester of university, given how well they did in their first semester. Specifically, let x be equal to the number of "A" grades (including A-, A, and A+ grades) that a student receives in their first semester of college. We would like to predict the value of y, which we define as the number of "A" grades they get in their second semester. Refer to the following training set of a small sample of different students' performances. Here each row is one training example. Recall that in linear regression, our hypothesis is $h_{\theta}(x) = \theta_0 + \theta_1 x$, and we use m to denote the number of training examples.

x	У	
3	4	
2	1	
4	3	
0	1	

Figure 2: Example Dataset

- (a) (7 Marks) What values of θ_0 , θ_1 should we use. [Hint: Remember our goal is to find the parameters that minimize following cost function. $J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left(h_{\theta}(x^{(i)}) y^{(i)} \right)^2$. You may use simple Calculus rules to arrive at the solution]
- (b) (2 Marks) How many A's a student will get in his second semester, if he had 5 A's in his first semester?
- (c) (1 Mark) How many A's a student will get in his second semester, if he had 1 A's in his first semester?

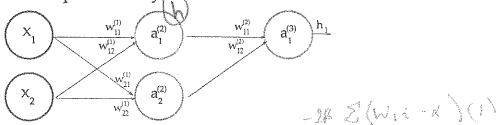
	X	Υ
0	0.000000	20.000000
4	1.010101	27.080910
2	2.020202	36.202428
3	3.030303	47.364555
4	4.040404	60.567289

Figure 3: Example dataset.

- (a) (3 Marks) A linear regression run resulted in the model: y = 6x + 2. Determine the error in each estimation and the mean sum of square of errors (MSE) (in other words what will be the cost function).
- (b) (4 Marks) Another linear regression run resulted in the model: $y = 10 * x^2 + 6x + 2$. Determine the error in each estimation and the mean sum of square of errors (MSE) (in other words what will be the cost function).
- (c) (3 Marks) What model among the both of the above will you use? Now use the selected model to estimate y for x = 15.

 network to RBF neural network we simply update the definition of affine functions $a_1^{(2)}, a_2^{(2)}$, and $a_1^{(3)}$ with their corresponding RBF definitions.

Example of 3 Layer Neural Network



For instance, in our RBF $a_1^{(2)}$ will have following definition:

$$a_1^{(2)} = \exp(-\beta \sum_i (W_{1i}^{(2)} - X_i)^2)$$

instead of $a_1^{(2)} = \sum_i (W_{1i}^{(2)} * X_i)$. $a_2^{(2)}$ and $a_1^{(3)}$ also have similar forms. Here β is a hyperparameter and remains same for the neurons of a layer. Also note that there are no biases involved. For this current RBF neural network we will be using L_2 loss function that is, $L_i = (y^{(i)} - h_i^{(i)})^2$

(a) (5 Marks) Write code for forward propagation,

```
def affine_forward_RBF(x, w, beta):

"""

Computes the forward pass for an RBF (fully-connected) layer.

The input x has shape (N, d_1, ..., d_k) where x[i] is the ith input.

Inputs:

x - Input data, of shape (N, d_1, ..., d_k)

w - Weights, of shape (D, M)

beta - a scalar for the current layer

Returns a tuple of:

- out: output, of shape (N, M)

- cache: (x, w, beta)

"""

(D, M)

(D, M)
```

(b) (10 Marks) Now given this definition of RBF networks, your goal is to derive derivatives for $W^{(2)}$ and $W^{(1)}$.

Code the following network in pytorch. The fully connected neural network predicts whether an RGB image of size 32x32 is real or fake.

- There are 3 hidden layers
- All layers except the last one have ReLU() as activations.
- Use nn.Sequential to build the network
- Choose an appropriate loss function and an optimizer.
- Do a forward pass and a backward pass.
- The class name should be Discriminator and the object of the class should be netD=Discriminator()
- The dataloader is a tuple containing an image and a label.

1001

