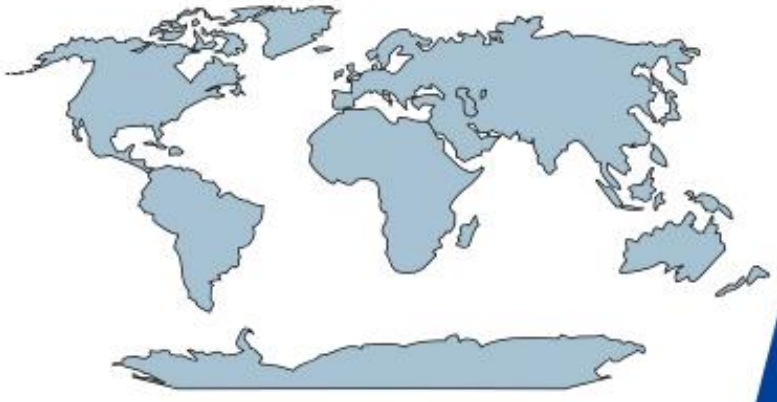
ISSN 2322 - 0899

INTERNATIONAL JOURNAL OF RESEARCH IN MANAGEMENT & SOCIAL SCIENCE



Volume 9, Issue 2 (I)
April - June 2021

GRADIENT PROBLEM 46 – 50

Durga Laxman Ursal and Mithilesh Chauhan

IMPROVING EMOTIONAL QUOTIENT OF HUMANOIDS WITH THE USE OF THERMOLOGY AND THERMOGRAPHY 51 – 57

Hemant V Karekar and Mithilesh Chauhan

IMPACT OF SOCIAL MEDIA ON TODAY'S YOUTH 58 – 63

Ketki Deepak Kunchikar and Rajendra Patil

CHANGING DYNAMIC OF COVID ERA: NEW NORMAL IN SOCIETY AND INDUSTRY BANKING SECTOR AND IT'S DYNAMICS 64 – 65

Bichu Sarah Naveed

ELECTRICITY POWER OUTAGE ISSUE FACED BY PEOPLE RESIDING IN NALLASOPARA 66 – 69

Maithili Anil Dhumal

ENTREPRENEURSHIP & 'TRANSITION OF MANAGEMENT AND MANAGEMENT STRATEGIES' OF STARTUPS : A CASE STUDY OF PROFESSIONALS REVIEW REGARDING STARTUP EXPERIENCE & PROBLEMS 70 – 75

Harshit Khanna

A STUDY ON INVESTMENT PATTERN AMONG THE EMPLOYEES OF WESTERN RAILWAY WITH RESPECT TO KANDIVALI (EAST) E.M.U. CAR SHED 76 – 79

Harisa Tuscano and Rajendra Patil

CHANGING DYNAMICS IN THE INDIAN THEATRE INDUSTRY DURING COVID-19 AND THE PATH AHEAD 80 – 83

Jennifer Jagose

SECURING BYOD [BRING YOUR OWN DEVICE]: A ORGANIZATIONAL SECURITY CHALLENGE 84 – 87

Rakesh Suresh Dalvi

STUDY ON DISASTER MANAGEMENT AND TO ACCESS THE AWARENESS OF PEOPLE LIVING IN KOTTAYAM DISTRICT, KERALA 88 – 92

Athil Kalloor

A STUDY ON TWITTER DURING COVID 93 – 97

Smit Shah

ETHEREUM BLOCKCHAIN BASED VOTING SYSTEM 98 – 100

Pragati Vilas Nivate and Mithilesh Chauhan

STUDY ON STUDENTS' PREFERENCE BETWEEN ONLINE AND CLASSROOM LEARNING 101 – 104

Tanisha A. Nikam

International Journal of Research in Management & Social Science
Volume 9, Issue 2 (I) April - June 2021
ISSN 2322 - 0899

GRADIENT PROBLEM

Durga Laxman Ursal[1] and Mithilesh Chauhan(Mentor)[2]
Student[1], Department of Information Technology, Vikas College of Arts, Science & Commerce
Assistant Professor[2], Department of Information Technology, Bunts Sangha's S.M. Shetty College of Science, Commerce and Management Studies, Powai

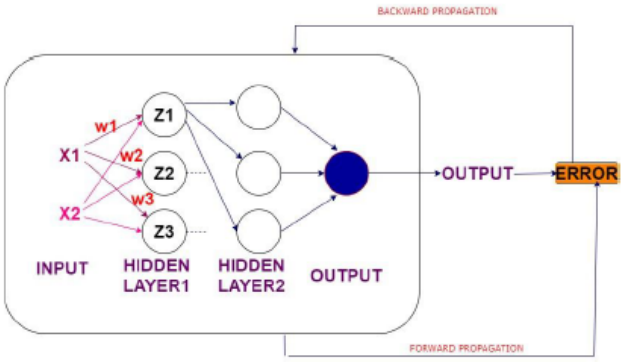
ABSTRACT
This research paper highlights the gradient problem which occurs during the training of the neural network. During the training, the loss is calculated till the desired output is obtained. With the help of loss, we get the magnitude and direction to change the weights. This weight change is done by the chain rule where we multiply the derivatives. This works well for the shallow network but as the number of layers increases, the value of these derivatives becomes vanishingly small causing vanishing gradient problem or very large causing exploding gradient problem.

Keywords: Backpropagation, gradient descent, Learning rate, activation function, loss function

INTRODUCTION
A Neural network consists of interconnected weights, biases, and activation functions. The sum of the product of input and weights becomes the net input which is then given as an input to the activation function. The output of the activation function becomes the actual output of the neuron. The group of neurons is various layers form a neural network where each neural output participates in the decision making.

To make correct decisions, it is required to train the neural network. The above process continues for layers till the last layer arrives. At this point, we check whether the output obtained from the last layer matches the target output. If it does not, we need to adjust the weights. This adjustment of weights is done by a method called Backpropagation.

When the actual output and expected output don't match, a loss is calculated by a function called as loss/cost function.



In Backpropagation, to change the output from the output nodes, there is a need to update the weights connecting the output layer. Another way to do this is to change the nodes of the previous layer which is the activation output. As we cannot directly change the activation output as it has been influenced by the weights and activation outputs of the previous layer, we need to make modifications in the weights of the previous layer, this process continues till we reach the first layer where we cannot change the actual inputs but we can change the weights. We are moving from right to left for a slight change in the output value.

The rate at which we change the weights is called the **learning rate**. Learning rate is the most important hyperparameter in training the neural network. A high learning rate may converge too quickly to a suboptimal solution and may skip the necessary values. However, a low Learning rate may lead to a slow training of the network and might cause the network to get stuck. So choosing an appropriate learning rate is of most importance.

CONTENT
The Gradient is a numeric calculation that allows us to adjust our parameters in such a way that the desired output is obtained with minimum deviation.