## horizontal line

Perceptron Loss Function (Class5)

08.01.2025

**─**

[Perceptron loss function](https://colab.research.google.com/drive/1dHjD4uVgRF7iPrAODYf9fT_h5S77O3P_#scrollTo=IFmdEztLBE1c)

[Problem with perceptron](https://colab.research.google.com/drive/1oWZSykMuXUOqvhEoZECPunFb-aoQMlV5)

# Problem with Perceptron Trick

1. There can be multiple random lines that classify the clusters but we can’t be able to find out the best fitted line .
2. We need to find convergence for each and every line , maybe there are thousands of points so for every point we need to check how many misclassified points by that particular line .
3. Hypothetical cases may be like that every time those points are considered that are correctly classified and the misclassified ones are left or vice versa. Thus an incorrect line will be the output.
4. Practically perceptron would not perform better than Logistic regression.
5. Would not classify non linear data.

# Perceptron Loss Function

This is used to quantify how much the line classified the points correctly .

This implies penalty , which line misclassified more , which one less. Now we need to find out the line with the minimum penalty.

Method is to simply take coordinates of misclassified points and put them in the equation of the line at present. What this gives is a penalty of that line .

By the help of SGD (Stochastic Gradient Descent) we would form it as an equation .

## *L(w1 , w2 ,b)* = 1/n

## *Where, L(yi , f(xi)) = max( 0 , - yi f(xi))*

## *and f(xi) = w1x1 +w2x2 + b*

# Explanation of loss Function

| Yi | Yi\_hat (or f(xi)) | Max( 0 , - yi f(xi)) |
| --- | --- | --- |
| 1 (+ve) | 1 (+ve) | Max (0 , -ve) = 0 |
| -1 | -1 | Max (0 , -ve) = 0 |
| 1 | -1 | Max (0 , +ve) = +ve no. |
| -1 | 1 | Max (0 , +ve) = +ve no. |

This clearly tells no loss while yi and yi\_hat are same and some non zero no. when they are different.

# Gradient Descent

For i in range (epochs):

W1 = w1 -

W2 = w2 -

b = b -

## Loss function differentiation

=

=

=

Therefore the above equations are :

W1 = w1 + \* yi \* x[i][0]

W2 = w2 + \* yi \* x[i][1]

b = b + \* yi

# More Loss Functions

| **Loss Function** | **Activation** | **Output** |
| --- | --- | --- |
| Hinge loss | step | Perceptron (binary classification) |
| Log loss (binary cross entropy) | sigmoid | Logistic regression (binary classification) |
| Categorical cross entropy | softmax | Softmax regression (multi class classification) gives prob. |
| mse | linear | Linear regression |