We want to train the logistic regression model for classification problem. You are provided with a text passage for training purpose.

Input	Actual Label
What truly sets this book apart is the depth of emotion it evokes. I laughed, I cried, and I felt my heart race with anticipation during the most gripping moments. The themes explored in this story are both timely and timeless, touching on the complexities of human nature, the power of friendship, and the triumph of hope in the face of adversity. I cannot recommend this book enough. It is a masterpiece of modern literature that deserves a place on every bookshelf. Whether you're a seasoned reader or just looking for a captivating story to dive into, this book will not disappoint. Prepare to be transported on an unforgettable journey that will stay with you for a lifetime.	**1**

Each training observation would be represented by the 5 crafted features shown in the following table.

Features	
1	log of Word count
2	number of punctuations (period, comma, apostrophe, quotation, question, exclamation, colon etc.)
3	number of positive words
4	number of negative words
5	ratio of capitalized words (words starting with capital letter) to total words

Dictionary for positive and negative words is given below

Positive dictionary	Negative dictionary
Good, unforgettable, masterpiece, depth, laughed, timeless, captivating, happy, triumph, friendship, modern, lifetime, gripping, enjoy, proud.	Not, cannot, disappoint, sad, cried, hopeless, adversity, waste, weird, complexities, anger, seasoned, anticipation, bad, rude,

Your task is to update the weights once, using stochastic gradient descent algorithm. Assume all initial weights set to 0.2 and learning rate α=0.5. Also compute the Loss function: binary cross entropy loss.

Solution:

Features	X_1	X_2	X_3	X_4	X_5
Values	**2.07**	**14**	**11**	**8**	**0.07**

$$P(1|X) = \text{Sigmoid}\,(W_1X_1 + W_2X_2 + W_3X_3 + W_4X_4 + W_5X_5 + b)$$

$$= \text{Sigmoid}\,[(0.2)(2.07) + (0.2)(14) + (0.2)(11) + (0.2)(8) + (0.2)(0.07) + 0.2]$$

$$= \text{Sigmoid}\,(7.228)$$

$$= 0.99$$

P(0|X) = 1 – P(1|X) = 1 – 0.99 = 0.01

Predicted Class = 1

$$L_{CE}(\hat{y}, y) = -[y \log \sigma(w \cdot x + b) + (1 - y) \log (1 - \sigma(w \cdot x + b))]$$
$$= -[\log \sigma(w \cdot x + b)]$$

= -log (0.99) = 0.01

Gradient Descent:

$W_1 = W_1 - \alpha (\hat{y} - y)X_1 = 0.2 - 0.5 (0.99 - 1)2.07 = 0.21$

$W_2 = W_2 - \alpha (\hat{y} - y)X_2 = 0.2 - 0.5 (0.99 - 1)14 = 0.27$

$W_3 = W_3 - \alpha (\hat{y} - y)X_3 = 0.2 - 0.5 (0.99 - 1)11 = 0.25$

$W_4 = W_4 - \alpha (\hat{y} - y)X_4 = 0.2 - 0.5 (0.99 - 1)8 = 0.24$

$W_5 = W_5 - \alpha (\hat{y} - y)X_5 = 0.2 - 0.5 (0.99 - 1)0.07 = 0.2003$

$b = b - \alpha (\hat{y} - y) = 0.2 - 0.5 (0.99 - 1) = 0.205$