A logistic regression model with L2 regularization was selected. Let's give a thorough mathematical analysis of the model and how it learns:

1. Model Information:

The model's binary classifier aims to categorize the data into two groups (0 or 1). The sigmoid function is used to convert the input features to probabilities, and a threshold of 0.5 is used to determine the final binary predictions.

2. The Sigmoid Function:

The weighted sum of the inputs is transformed into probabilities using the sigmoid function, which adds nonlinearity to the model. It is said that the sigmoid function is:

sigmoid(z) is equal to 1 / (1 + exp(-z)

3. Cost Function Combined with L2 Regularization:

The cost function for the model is the binary cross-entropy loss, which calculates the difference between the true binary labels' (y true) and predicted probabilities' (y pred) characteristics. The model also incorporates L2 regularization to avoid overfitting and enhance generalization. Given by: The cost function with L2 regularization.

cost is equal to:

- (y_true * log(y_pred) + (1 - y_true) * log (1 - y_pred)) + (lambda_reg / (2 * m)) * sum(w^2)

Gradient Descent with Regularization at L2:

With the aid of gradient descent and L2 regularization, the model learns from the data. The model calculates the gradients of the cost function with respect to the weights ('w') and the bias ('b') for each iteration. The weights are updated using the gradients, and they are biased in the direction that minimizes the cost function. The L2 regularized update equations are as follows:

```
dw = (dw / m) + ((lambda_reg / m) * w)
db = db / m
w = w - alpha * dw

b = b - alpha * db
```

5. In order to minimize the cost function with the specified learning rate and regularization strength during training, the model modifies its bias and weights. This process is repeated for a predetermined number of iterations (epochs). Finding the best values for "w" and "b" that produce precise predictions on the training data is the objective.