

(<https://google.com/racialequity?authuser=0>)

Logistic Regression: Loss and Regularization

ited Time: 6 minutes

Loss function for Logistic Regression

The loss function for linear regression is squared loss. The loss function for logistic regression is **Log Loss**, which is defined as follows:

$$\text{Log Loss} = \sum_{(x,y) \in D} -y \log(y') - (1 - y) \log(1 - y')$$

where:

- $(x, y) \in D$ is the data set containing many labeled examples, which are (x, y) pairs.
- y is the label in a labeled example. Since this is logistic regression, every value of y must either be 0 or 1.
- y' is the predicted value (somewhere between 0 and 1), given the set of features in x .

Regularization in Logistic Regression

Regularization

(<https://developers.google.com/machine-learning/crash-course/regularization-for-simplicity/video-lecture?authuser=0>)

is extremely important in logistic regression modeling. Without regularization, the asymptotic nature of logistic regression would keep driving loss towards 0 in high dimensions. Consequently, most logistic regression models use one of the following two strategies to dampen model complexity:

- L_2 regularization.
- Early stopping, that is, limiting the number of training steps or the learning rate.

(We'll discuss a third strategy— L_1 regularization—in a later module

(<https://developers.google.com/machine-learning/crash-course/regularization-for-sparsity/video-lecture?authuser=0>)

.)

Imagine that you assign a unique id to each example, and map each id to its own feature. If you don't specify a regularization function, the model will become completely overfit. That's because the model would try to drive loss to zero on all examples and never get there, driving the weights for each indicator feature to $+\infty$ or $-\infty$. This can happen in high dimensional data with feature crosses, when there's a huge mass of rare crosses that happen only on one example each.

Fortunately, using L_2 or early stopping will prevent this problem.

ary

Logistic regression models generate probabilities.

Log Loss is the loss function for logistic regression.

Logistic regression is widely used by many practitioners.

arms

early stopping

(https://developers.google.com/machine-learning/glossary?authuser=0#early_stopping)

regularization

(https://developers.google.com/machine-learning/glossary?authuser=0#L1_regularization)

- log loss

(https://developers.google.com/machine-learning/glossary?authuser=0#Log_Loss)

- L_2 regularization

(https://developers.google.com/machine-learning/glossary?authuser=0#L2_regularization)

[Help Center](https://support.google.com/machinelearningeducation?authuser=0) (<https://support.google.com/machinelearningeducation?authuser=0>)

[Previous](#)

← Calculating a Probability

(<https://developers.google.com/machine-learning/crash-course/logistic-regression/calculating-a-probability?authuser=0>)

[Next](#)