Loss Functions

Broadly speaking, loss functions can be grouped into two major categories concerning the types of problems we come across in the real world: **classification** and **regression**. In classification problems, our task is to predict the respective probabilities of all classes the problem is dealing with. On the other hand, when it comes to regression, our task is to predict the continuous value concerning a given set of independent features to the learning algorithm.

ASSUMPTIONS

- n/m — number of training samples
- i — i^{th} training sample in a data set
- y(i) — Actual value for the ith training sample
- y_hat(i) — Predicted value for the ith training sample

Classification Losses

TYPES OF CLASSIFICATION LOSSES

1. Binary Cross-Entropy Loss / Log Loss
2. Hinge Loss

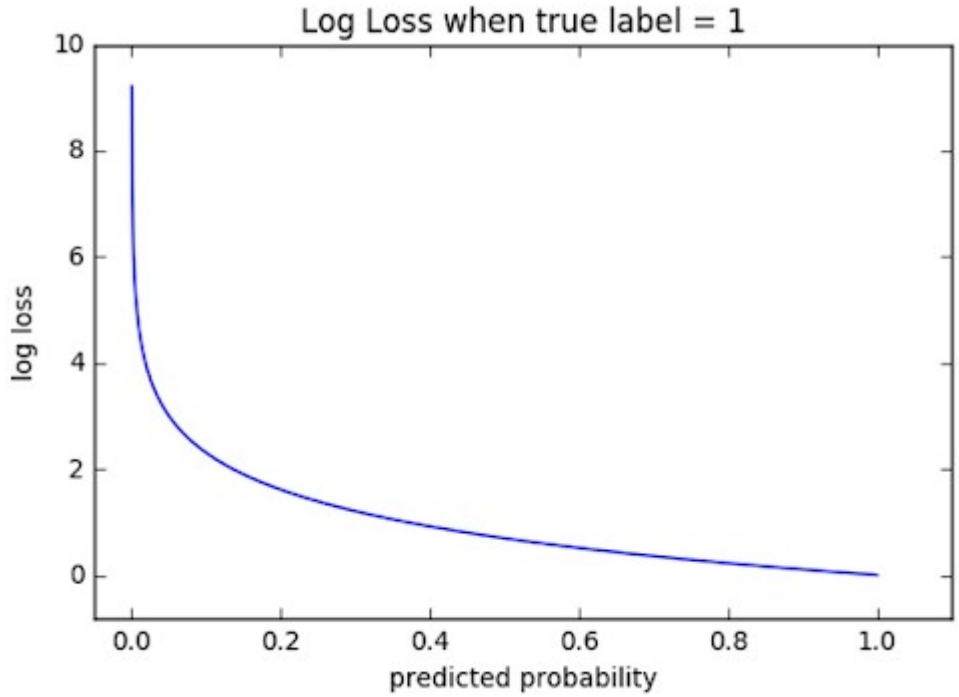
1. BINARY CROSS-ENTROPY LOSS / LOG LOSS

This is the most common loss function used in classification problems. The cross-entropy loss decreases as the predicted probability converges to the actual label. It measures the performance of a classification model whose predicted output is a probability value between `0` and `1`.
When the number of classes is `2`, it's *binary classification.*

$$L = -\frac{1}{m}\sum_{i=1}^{m}(y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i))$$

When the number of classes is more than `2`, it's *multi-class classification.*

$$L = -\frac{1}{m}\sum_{i=1}^{m} y_i \cdot \log(\hat{y}_i)$$

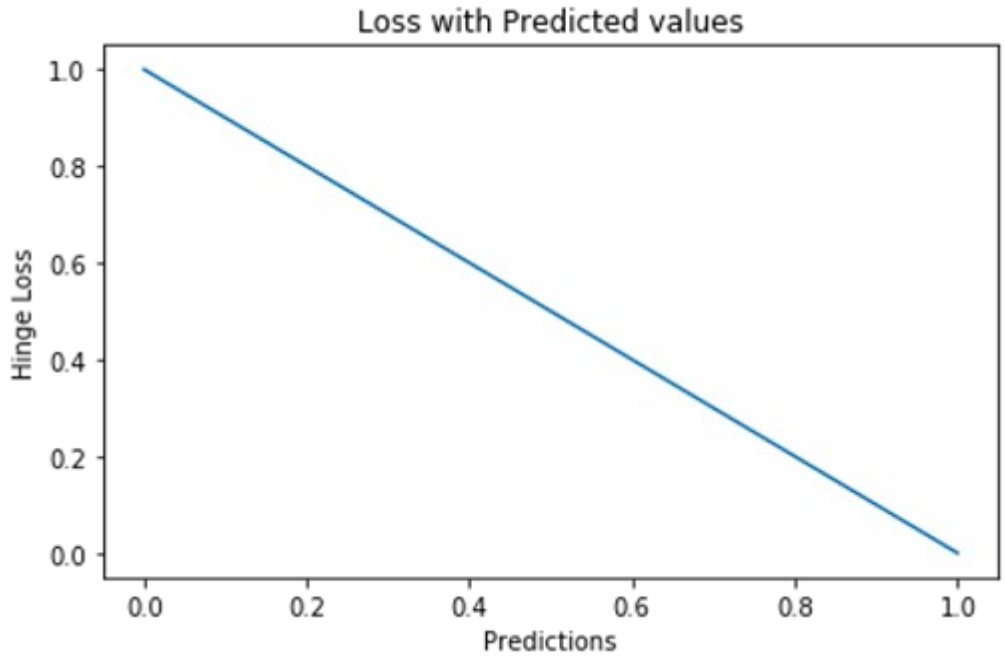


We derive the cross-entropy loss formula from the regular likelihood function, but with logarithms added in.

2. HINGE LOSS

The second most common loss function used for classification problems and an alternative to the cross-entropy loss function is hinge loss, primarily developed for support vector machine (SVM) model evaluation.

$$L = \max(0, 1 - y * f(x))$$



Hinge loss penalizes the wrong predictions and the right predictions that are not confident. It's primarily used with SVM classifiers with class labels as `-1` and `1`. Make sure you change your malignant class labels from `0` to `-1`.