DAY 4: LOSS FUNCTIONS

TOWNIM FAISAL CHOWDHURY

ML DEVELOPER @BRAINSTATION23
RESEARCH ASSISTANT @NSU
GRADUATE TA @NSU

TOPICS

- Cross Entropy Loss
- Negative log likelihood loss
- Hinge loss
- Custom loss implementation: Focal loss

CROSS ENTROPY

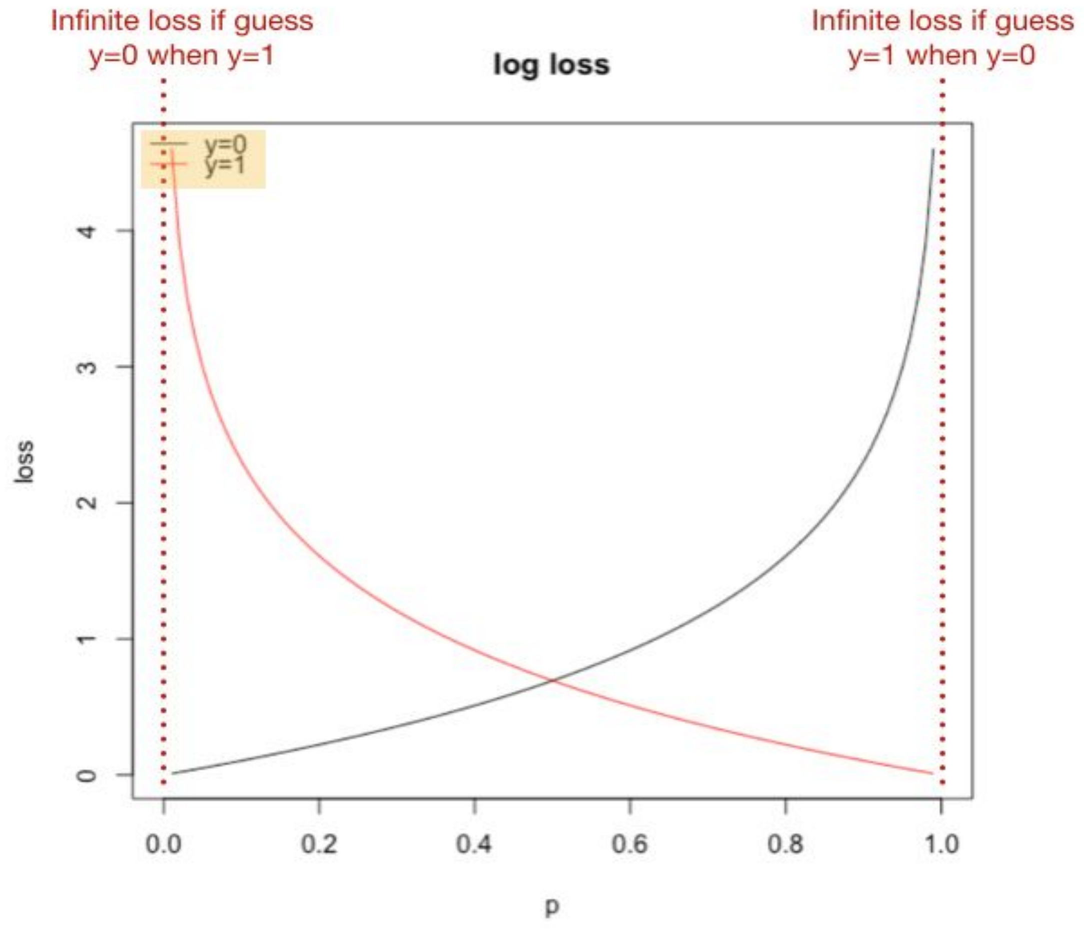
Cross-Entropy is used for classification problems. It minimizes the distance between the predicted and the actual probability distribution.

The Equation of Cross-Entropy loss:

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

So, it's used for Classification problems.

CROSS ENTROPY



Source: https://wandb.ai/site/tutorial/multilayer-perceptrons

NEGATIVE LOG LIKELIHOOD

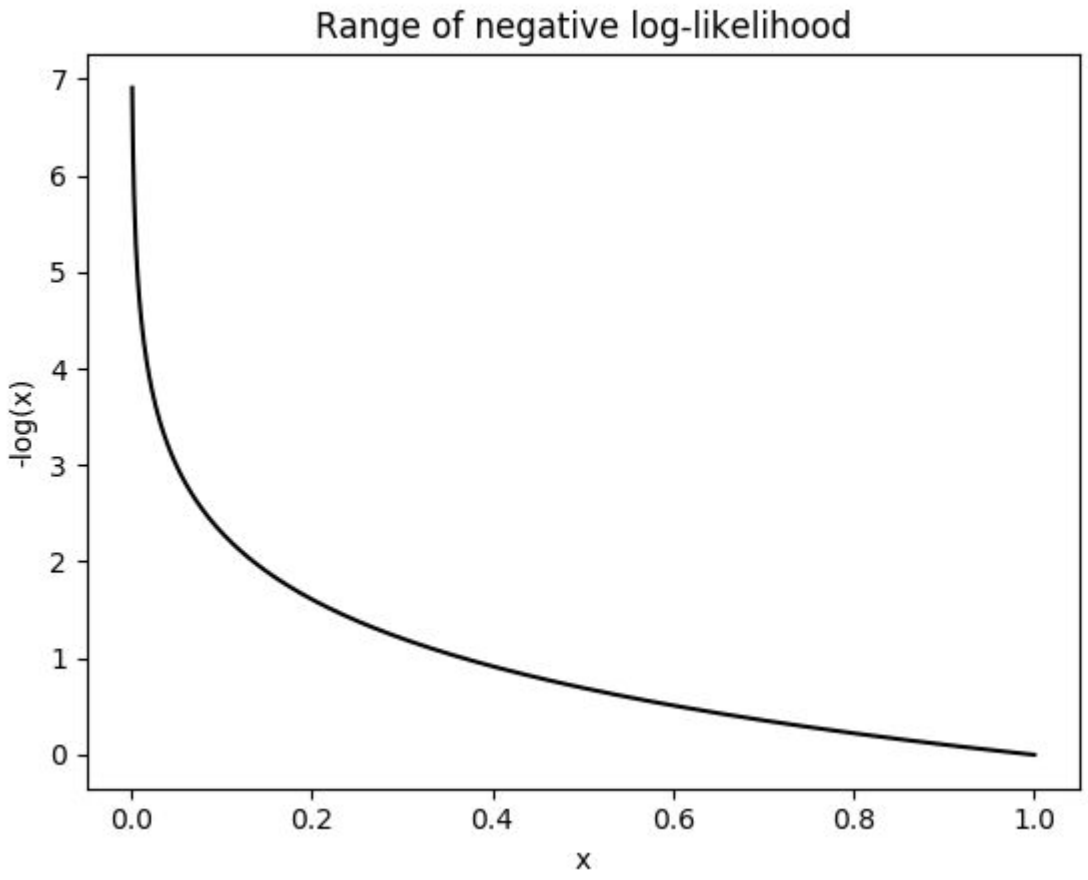
Negative Log Likelihood is also used for classification problems telling us how bad it's performing, the lower the better. Where most machine learning frameworks only have minimization optimizations, here we want to maximize the probability of choosing the correct category. We can **maximize by minimizing** the **negative** log likelihood.

The equation of the loss function: **L(y)=−log(y)**

$$\min -l(\theta)$$
$$w.r.t\ \theta$$

$$\min_{w.r.t\ \theta} \left(-\sum_{i=1}^{m} \left(y^{(i)} \times \log\left(h_\theta\left(x^{(i)}\right)\right) + \left(1 - y^{(i)}\right) \times \log\left(1 - h_\theta\left(x^{(i)}\right)\right)\right)\right)$$

NEGATIVE LOG LIKELIHOOD



Source: https://ljvmiranda921.github.io/notebook/2017/08/13/softmax-and-the-negative-log-likelihood/

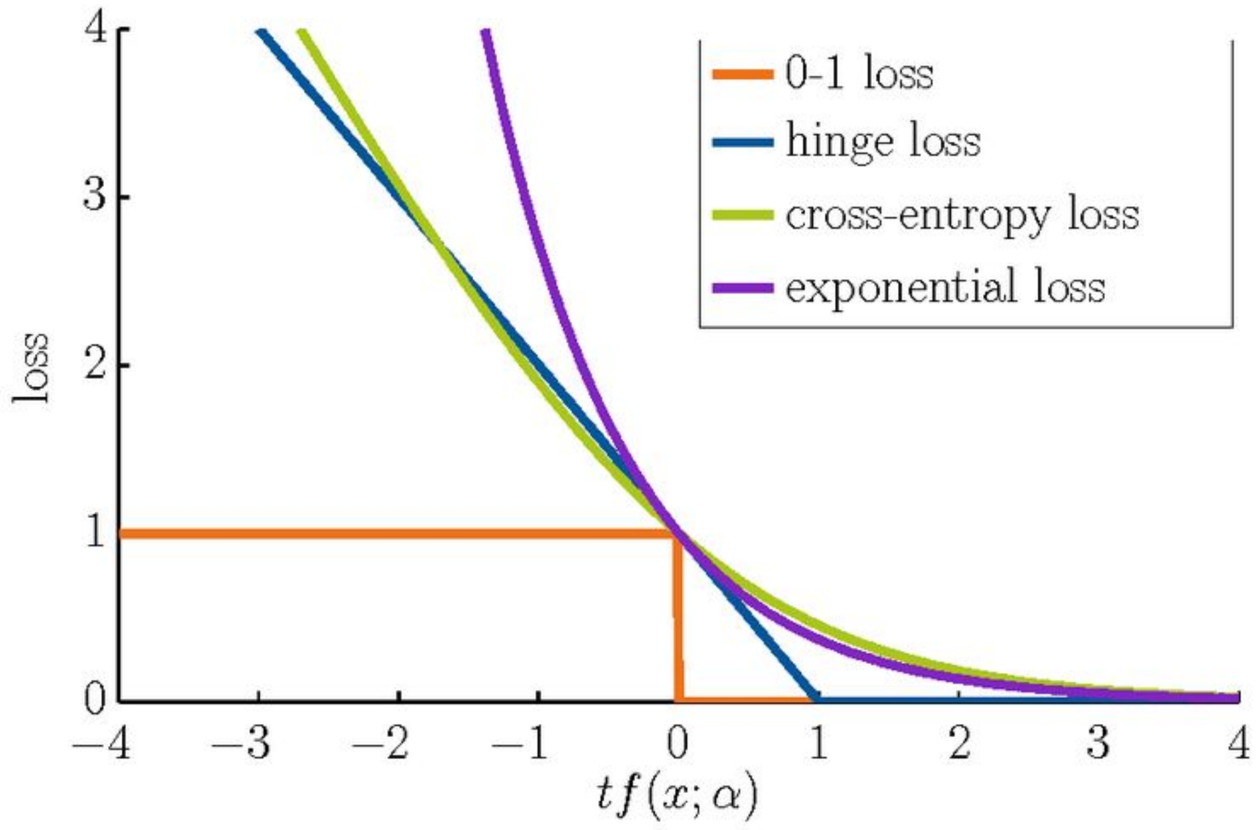
HINGE LOSS

The **hinge loss** is used for "maximum-margin" classification, most notably for support vector machines (SVMs). SVM uses a hinge loss, which conceptually **puts the emphasis on the boundary points**.

The formula for Hinge Loss:

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

HINGE LOSS



Source: https://www.researchgate.net/figure/Loss-functions-for-commonly-used-classifier-hinge-loss-SVM-cross-entropy-loss_fig5_236150927

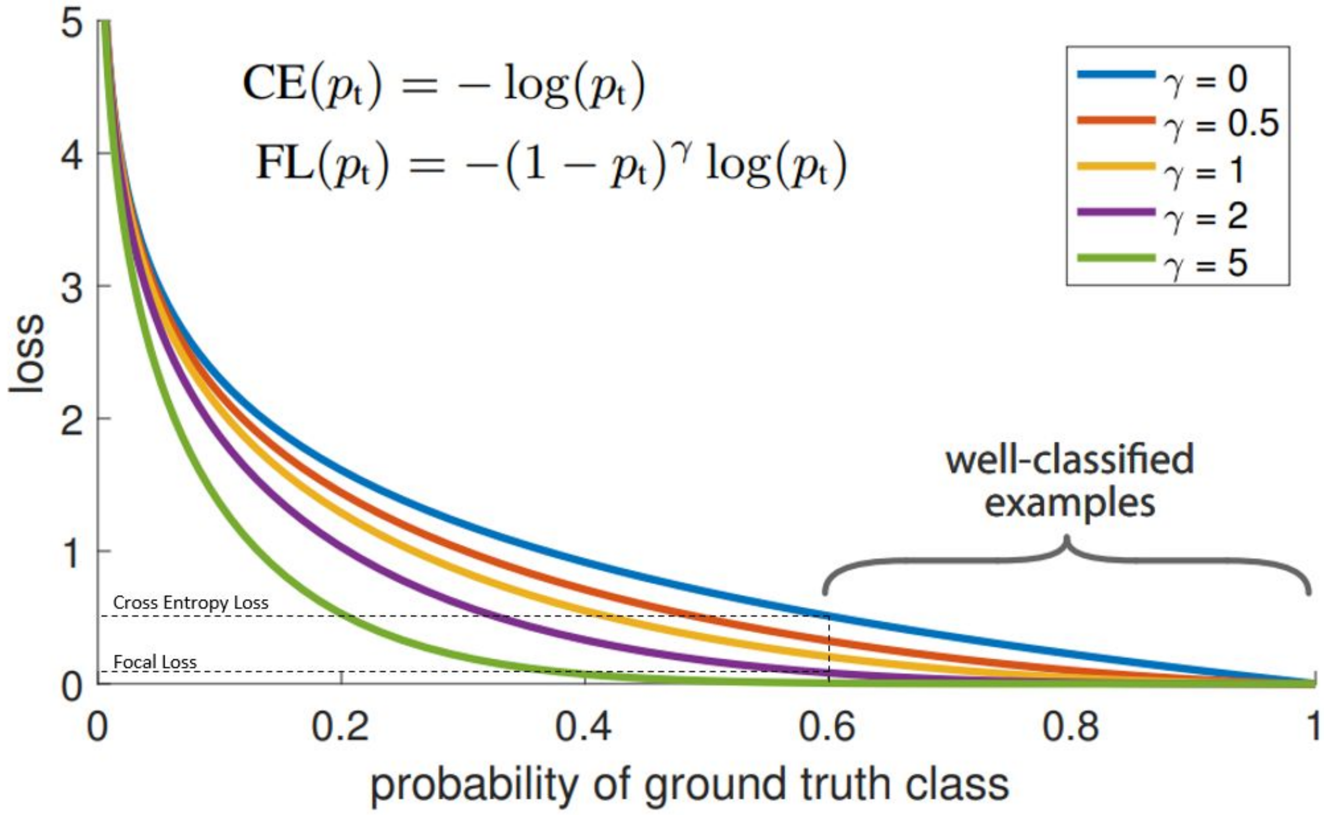
FOCAL LOSS

Focal loss is used to address the issue of the class imbalance problem. A modulation term applied to the Cross-Entropy loss function, make it efficient and easy to learn for hard examples which were prevailing in One-Shot Object Detectors.

Formula for Focal Loss

$$\mathrm{FL}(p_t) = -(1 - p_t)^{\gamma} \log(p_t)$$

FOCAL LOSS



Source: https://medium.com/swlh/focal-loss-what-why-and-how-df6735f26616

GET READY FOR CODE

Code: https://drive.google.com/file/d/1uBlc4XxfUbx9ANUZV_91FPLlLedUaCH3/view?usp=sharing

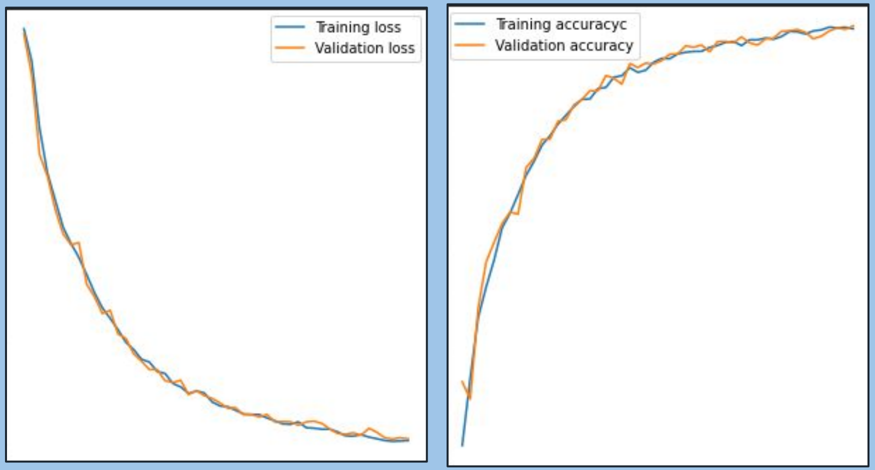
This Slide Link:
https://docs.google.com/presentation/d/1vvGifORiDHUU3RgR4ELVNdy-cya0VNfU/edit?usp=sharing&ouid=114298130813356779276&rtpof=true&sd=true

Visualizing 3D loss surface for cross-entropy loss: https://losslandscape.com/explorer

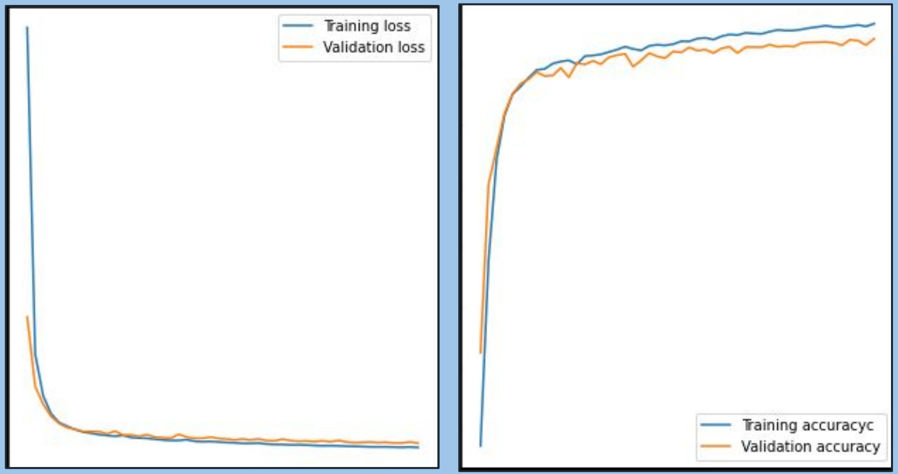
Other loss functions: https://pytorch.org/docs/stable/nn.html#loss-functions

Dataset: https://drive.google.com/file/d/1CLqe8c3EWdlt2zBxRSuCdEXwVHqYjYuO/view?usp=drivesdk

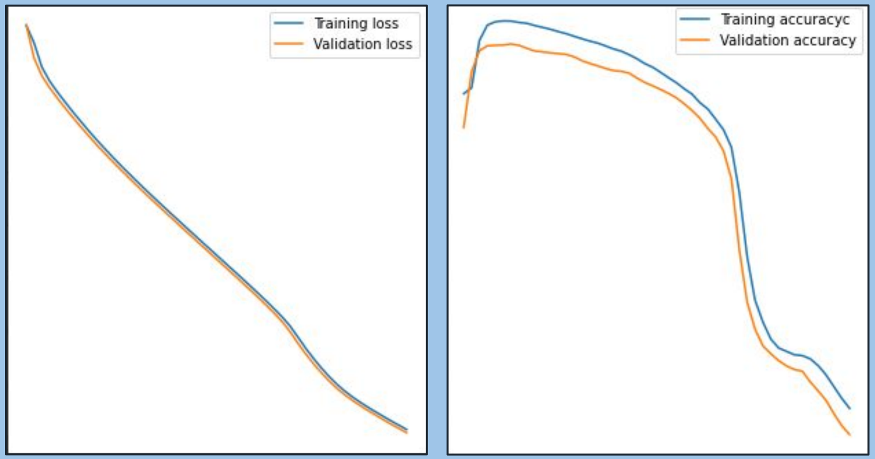
Impact of Loss



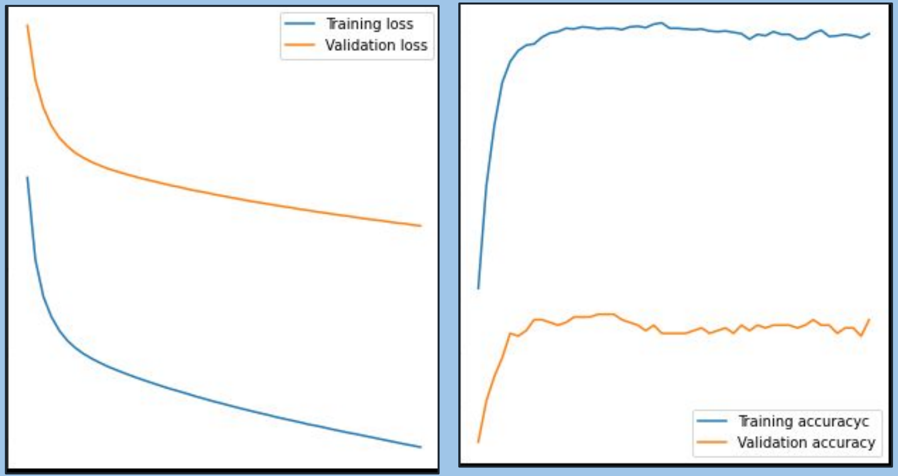
Cross Entropy Loss [Best val Acc: 91.70%]



Focal Loss [Best val Acc: 94.14%]



Hinge Loss [Best val Acc: 92.01%]



Negative Log Likelihood Loss [Best val Acc: 94.37%]

Useful Links

- https://wandb.ai/site/tutorial/multilayer-perceptrons
- https://ljvmiranda921.github.io/notebook/2017/08/13/softmax-and-the-negative-log-likelihood/
- https://medium.com/analytics-vidhya/understanding-loss-functions-hinge-loss-a0ff112b40a1
- https://medium.com/swlh/focal-loss-what-why-and-how-df6735f26616
- https://amaarora.github.io/2020/06/29/FocalLoss.html
- https://pytorch.org/docs/stable/nn.html#loss-functions
- http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture03.pd
- https://www.youtube.com/watch?v=h7iBpEHGVNc
- https://www.youtube.com/watch?v=bj1fh3BvqSU
- https://www.youtube.com/watch?v=CkiT0Muz62g

THANK YOU