python barcelona

Introduction to Neural Nets with Pytorch

2023-11-11

Noé Casas

Speaker

Noé Casas

- PhD in artificial intelligence for machine translation.
- 15 year of experience as software dev, 2 years as data scientist, 5 years with NLP and deep learning.
- Solo founder of Langtern, a language learning app for teachers and students: https://langtern.com

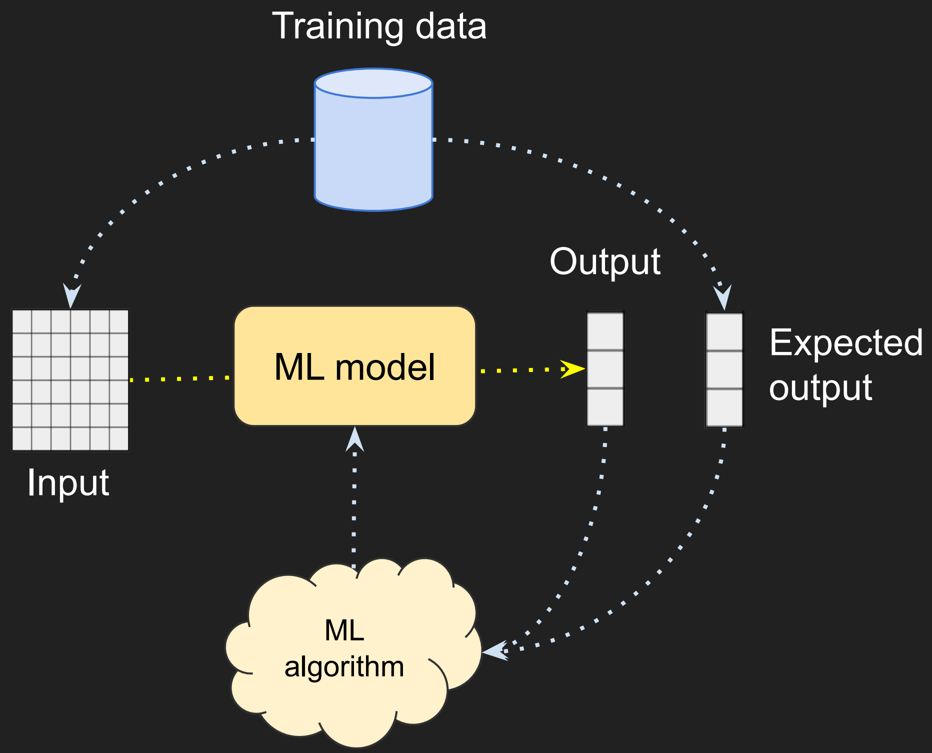
Contents

- Intro to supervised learning
- Intro to neural nets
- How PyTorch works
- Practical goal 1: linear regression
- Practical goal 2: logistic regression (binary classification)
- Practical goal 3: multiclass classification

Intro to Supervised Learning

What is supervised learning?

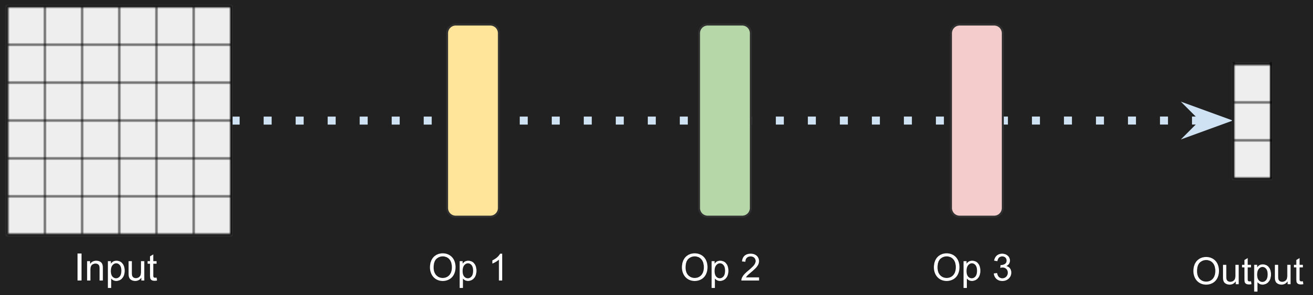
- Examples of input and expected output (training data) are used to train the ML model.
- The ML algorithm updates the model to minimize the difference between the model output and the expected output



Intro to Neural Nets

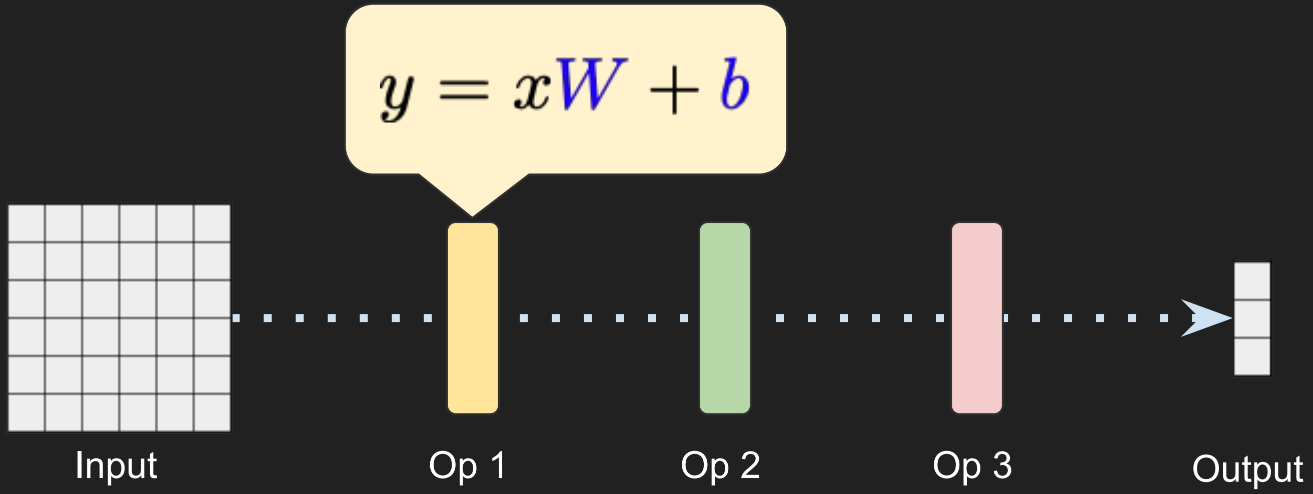
What are neural nets? (1/5)

A set of differentiable operations (layers) that receive some input and generate some output. Both input and output are usually matrices.



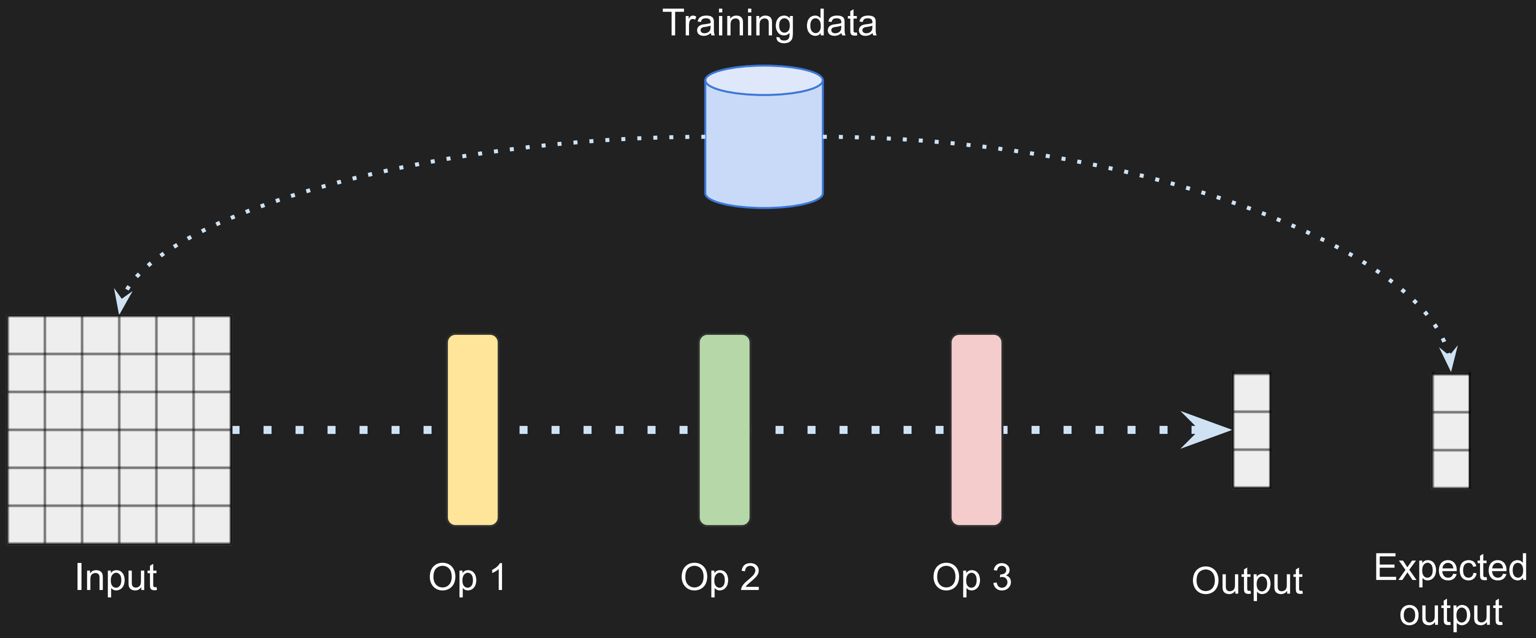
What are neural nets? (2/5)

Some of the layers have parameters whose values are not defined a priori. They are initialized to random (small) values and their value will be adjusted later.



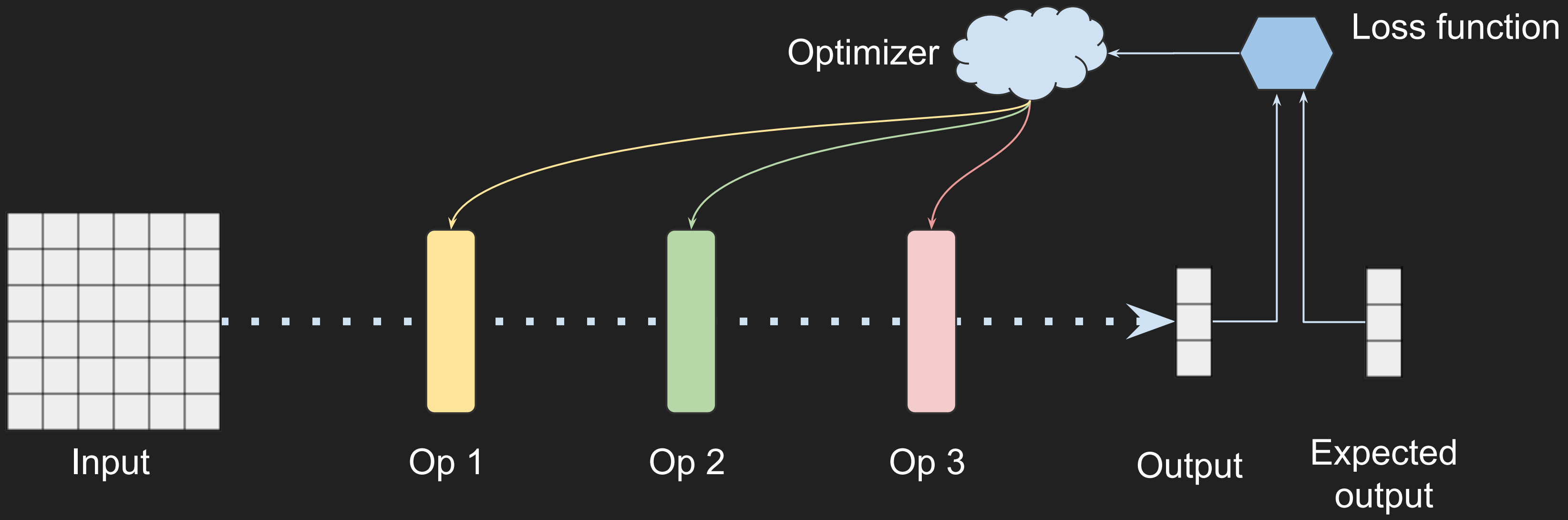
What are neural nets? (3/5)

Examples of input and expected output (training data) are used to train the net.



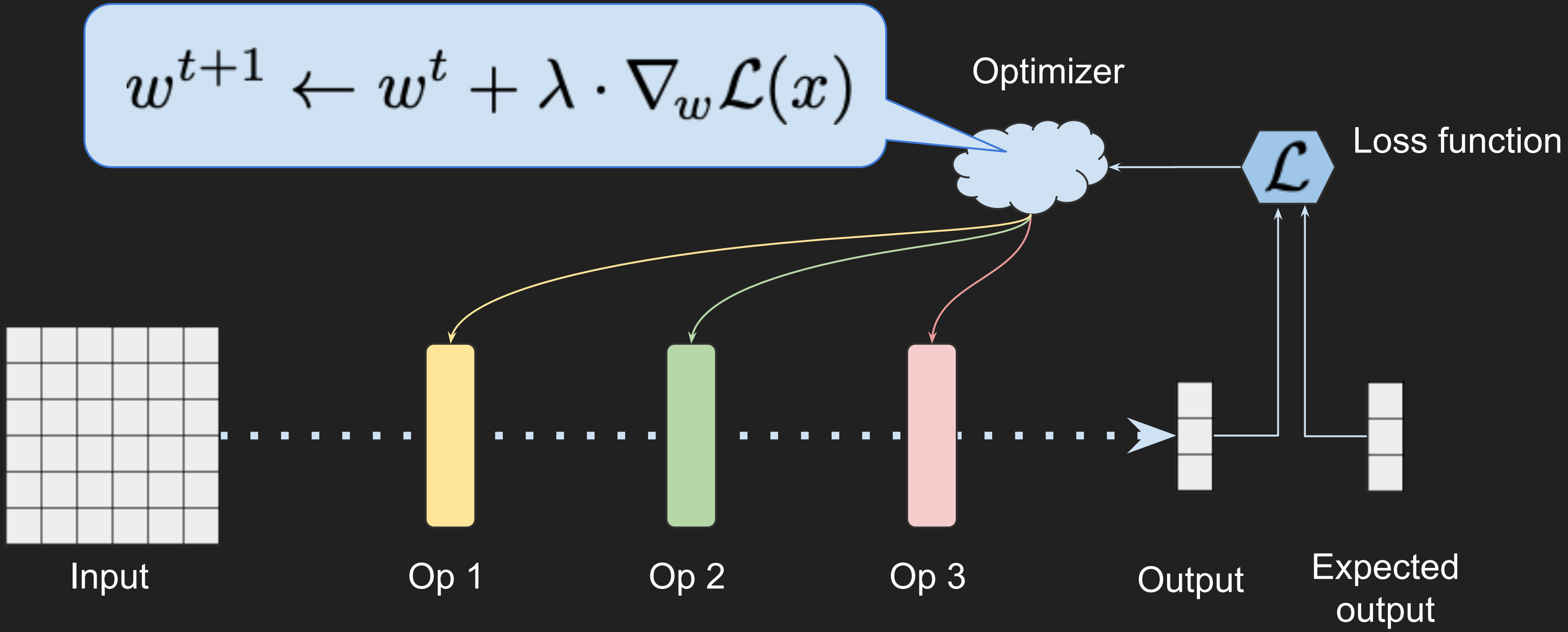
What are neural nets? (4/5)

The parameters are adjusted (trained) with an optimization process to minimize the difference (loss function) between the output and the expected output.



What are neural nets? (5/5)

To optimize the loss function, we use some variant of gradient descent.



$$w^{t+1} \leftarrow w^t + \lambda \cdot \nabla_w \mathcal{L}(x)$$

Hyperparameters

Elements of a neural network that cannot be trained. Examples:

- Structure of the network: how many layers, which type.
- Configuration of each layer (e.g. size W in a linear layer, activation function).
- Type of optimizer (e.g. SGD, Adam) and its configuration (e.g. learning rate).
- Training configuration (e.g. batch size, number of epochs).

Strategies to choose hyperparameter values: intuition, trial-error, grid search,...

Extra neural nets vocabulary

- Inference: to use a trained model (neural net) on some data to obtain the result.
- Deep learning = neural nets
- Tensor: a matrix, potentially multidimensional.
- Batch: group of samples used as input to the net as a single tensor.
- Epoch: total number of iterations needed to use all the training data.
- Graphical Processing Unit (GPU): device to compute fast.
- Back-propagation: algorithm to compute the gradient with respect to the params.

How PyTorch works

What is PyTorch?

- Deep learning Python library by Meta.
- “Imperative” programming model, like numpy.
- Automatic differentiation: it can compute the derivatives of its operations.
- Can use either CPU or GPU for both training and inference.

Practical goal 1:

Linear Regression

Practical goal 2:

logistic regression

(binary classification)

Practical goal 3:

Multiclass classification

Summary of the workshop

- Neural networks:
    - Are a set of differentiable layers with trainable params.
    - Are trained with gradient descent.
    - Have hyperparameters that cannot be trained, they are chosen e.g. with grid search.
- PyTorch:
    - Is like numpy (“imperative style”) but with automatic differentiation.
- Practical parts:
    - Linear regression: final linear layer + mean squared error loss.
    - Logistic regression (binary classif.): final sigmoid activation + binary cross-entropy loss.
    - Multiclass classification: final softmax activation + categorical cross-entropy loss.

Hope you enjoyed

contact@noecasas.com

Langtern

https://langtern.com