

# Machine Learning

Pawel Wocjan

University of Central Florida

Fall 2020

# An Iterative Approach

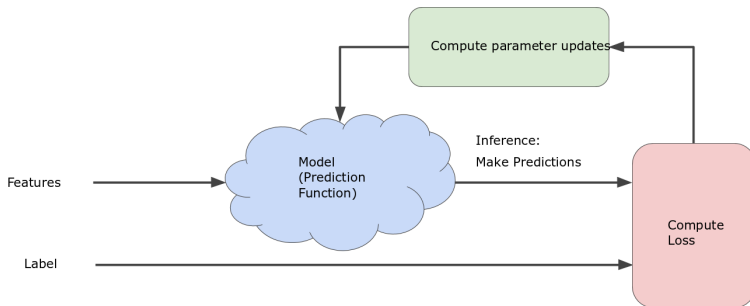
- ▶ We now examine how a machine learning model reduces loss iteratively.
- ▶ Iterative learning is reminiscent of the “Hot and Cold” kid’s game for finding a hidden object.
- ▶ The “hidden object” is the best possible model.
- ▶ We start with a wild guess *the value of  $w_1$  is 0* and wait for the system to tell us what the loss is.

Then, we try another guess *The value of  $w_1$  is 0.5* and see what the loss is etc.

- ▶ If we adapt the weights cleverly, we will usually be decreasing the loss.
- ▶ The real challenge is to find the best possible model as efficiently as possible.

# An Iterative Approach

- The figure below depicts the iterative trial-and-error process that machine learning algorithms use to train a model:



# An Iterative Approach

- ▶ Iterative strategies are prevalent in machine learning, primarily because they scale so well to large data sets.
- ▶ The model takes features  $x_1, \dots, x_n$  as input and returns one prediction  $\hat{y}$  as output.
- ▶ To simplify the discussion, consider the simplest linear regression model that takes only one feature as input:

$$\hat{y} = b + w_1 x_1$$

# An Iterative Approach

- ▶ What initial values should we set for  $b$  and  $w_1$ ?
- ▶ For linear regression problems, it turns out that the starting values aren't important.
- ▶ We could pick random values, but we'll just take the following trivial values instead:

$$b = 0, \quad w_1 = 0$$

# An Iterative Approach

- ▶ Let us examine what happens inside the “Compute parameter updates” part of the diagram.
- ▶ It is here that the ML system examines the value of the loss function and generates new values for  $b$  and  $w_1$ .
- ▶ For now, just assume that this mysterious box devises new values and then the ML system re-evaluates all those features against all those labels, yielding a new value for the loss function, which yields new parameter values.
- ▶ The learning continues iterating until the ML system discovers the model parameters with the lowest possible loss.
- ▶ Usually, we iterate until the loss stops changing or at least changes extremely slowly. When that happens, we say that the model has **converged**.

# Key Point

- ▶ A Machine Learning model is trained by starting with an initial guess for the weights and the bias and iteratively adjusting those guesses until finding the weights and the bias with the lowest possible (or sufficiently low) loss.

# Key Terms

- ▶ convergence
- ▶ loss
- ▶ training



# Gradient Descent

- ▶ The iterative approach diagram contained a green hand-wavy box entitled “Compute parameter updates.” We’ll now discuss the underlying algorithms in detail.
- ▶ Suppose we had the time and the computing resources to calculate the loss for all possible values of  $w_1$ .