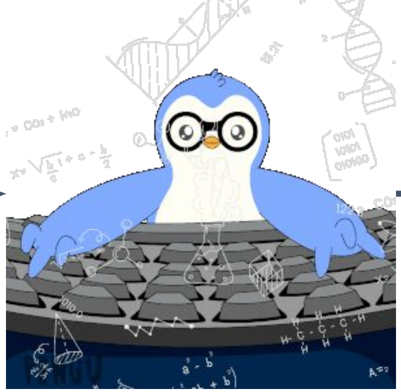
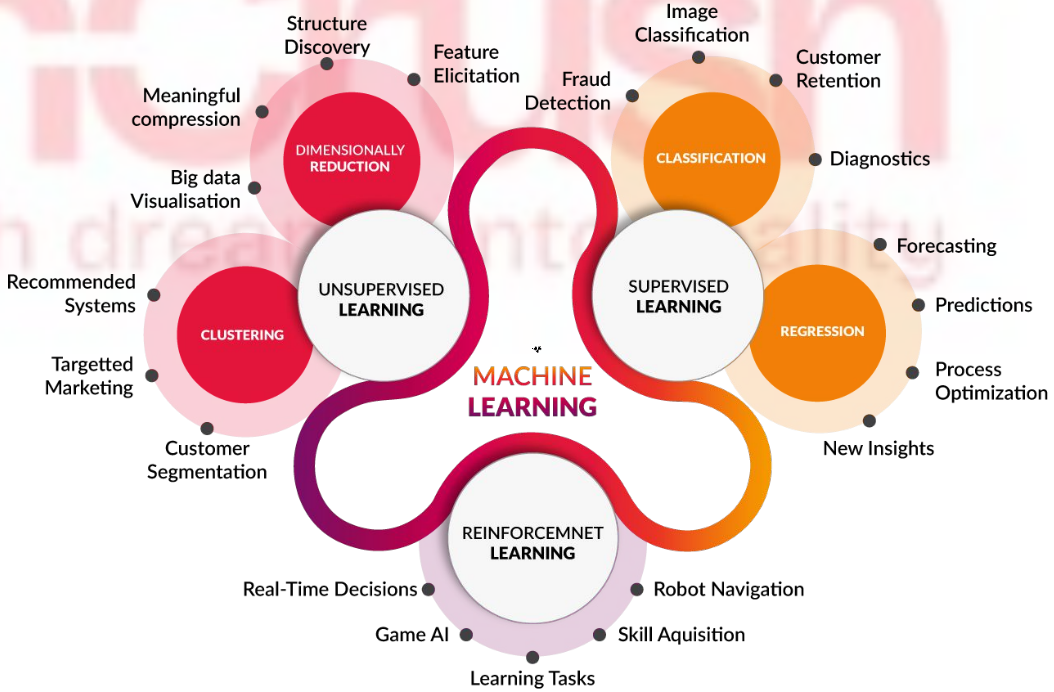
TECHCRUSH ARTIFICIAL INTELLIGENCE BOOTCAMP

Facilitator: Hammed Obasekore
August 20th & 26th, 2025

Recap



Programming

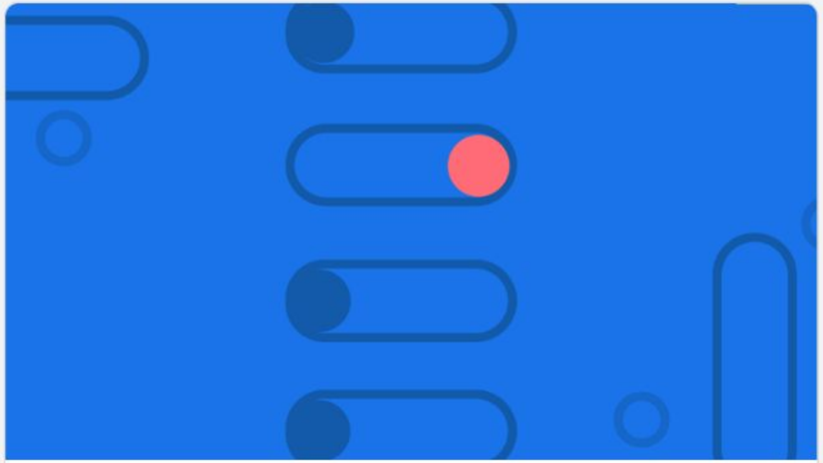


Linear Regression

An introduction to linear regression, covering linear models, loss, gradient descent, and hyperparameter tuning.

Logistic Regression

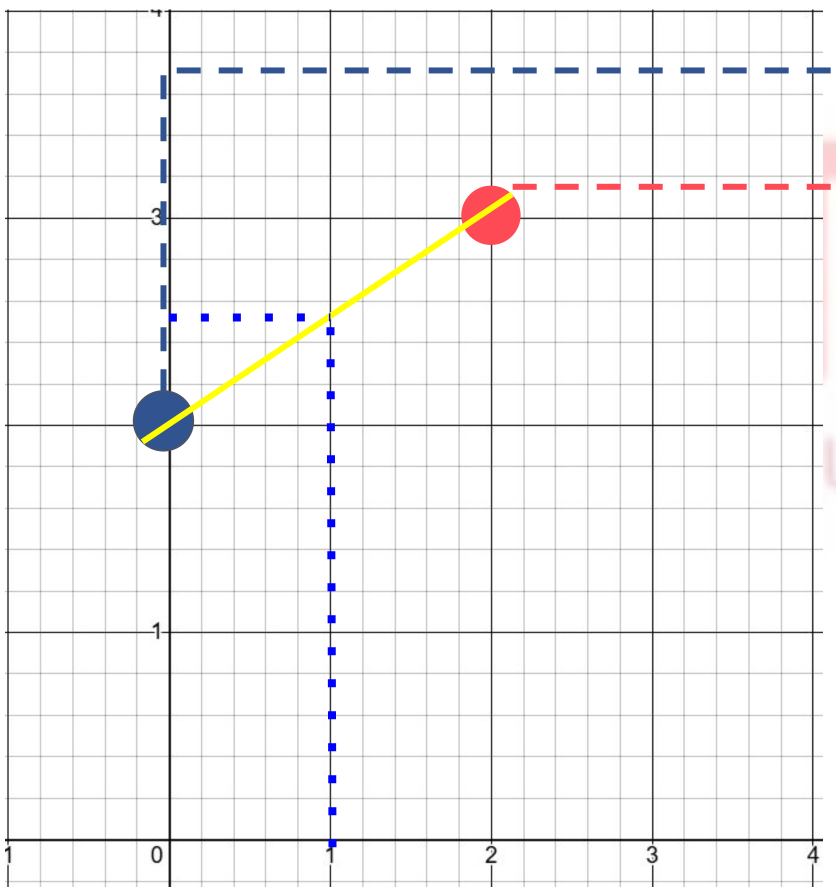
An introduction to logistic regression, where ML models are designed to predict the probability of a given outcome.



Classification

An introduction to binary classification models, covering thresholding, confusion matrices, and metrics like accuracy, precision, recall, and AUC.

Graphing Online



Data

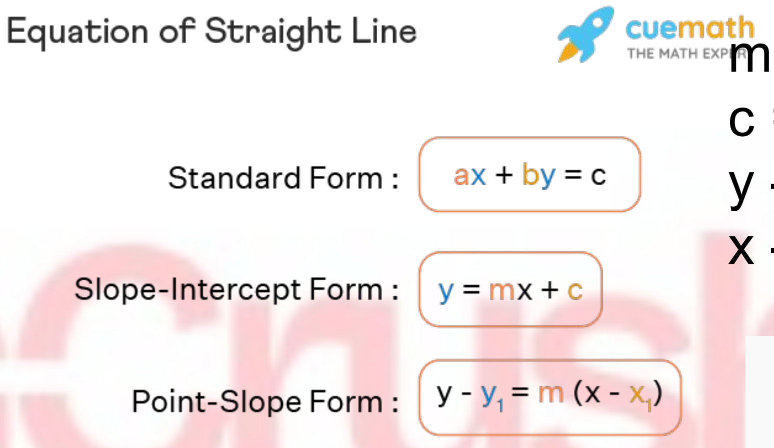
x	y
0	2
2	3
1	?

when x = 1

y = 0.5*x + 2

y = 0.5*1 + 2

y = 2.5

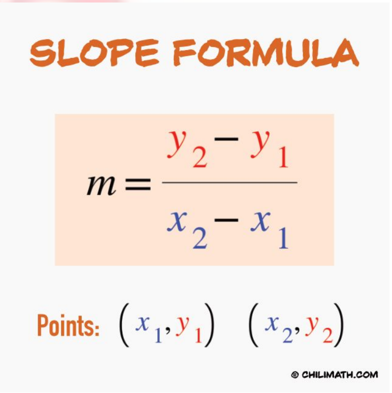


m ⇒ slope

c ⇒ y - intercept

y - axis

x - axis



Analytically

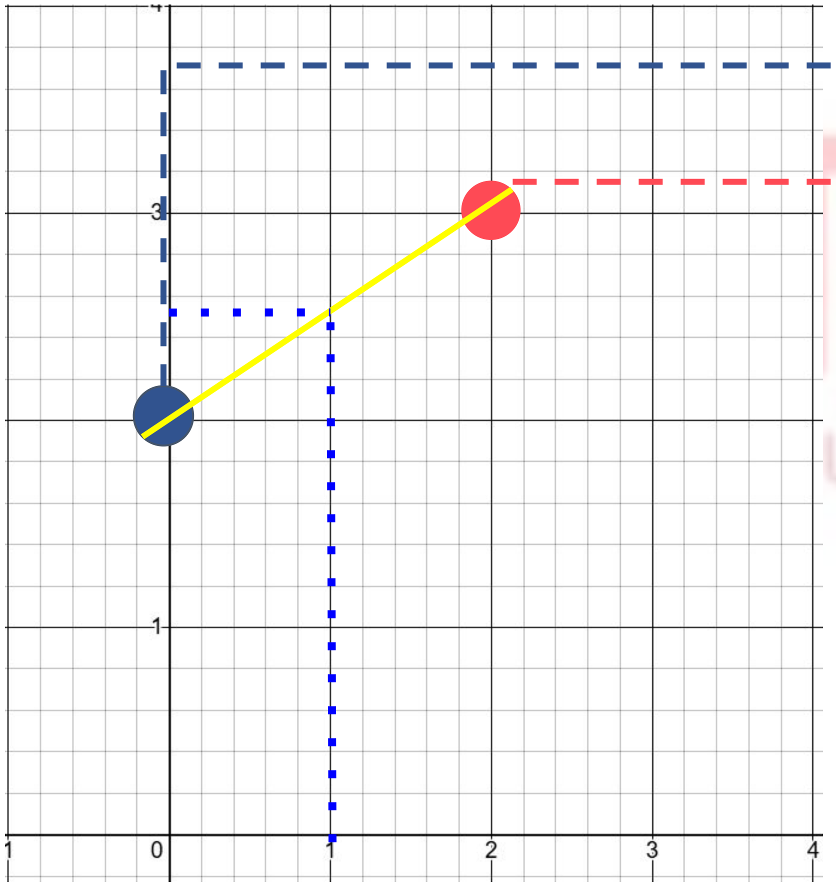
m = (3-2)/(2-0) = 0.5

$y - y_1 = m*(x - x_1)$

y - 2 = 0.5 (x - 0)

y = 0.5*x + 2

Graphing Online



Data

x	y
0	2
2	3
1	?

The model

$y = 0.5*x + 2$

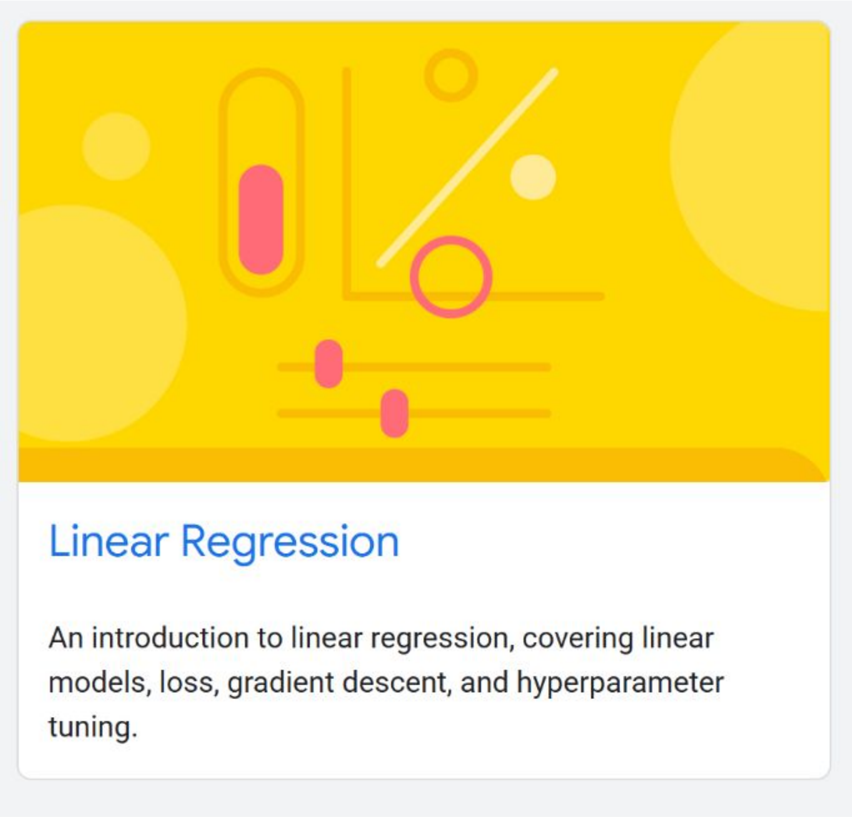
when $x = 100$
$y = 0.5*x + 2$
$y = 0.5*100 + 2$
$y = 52$

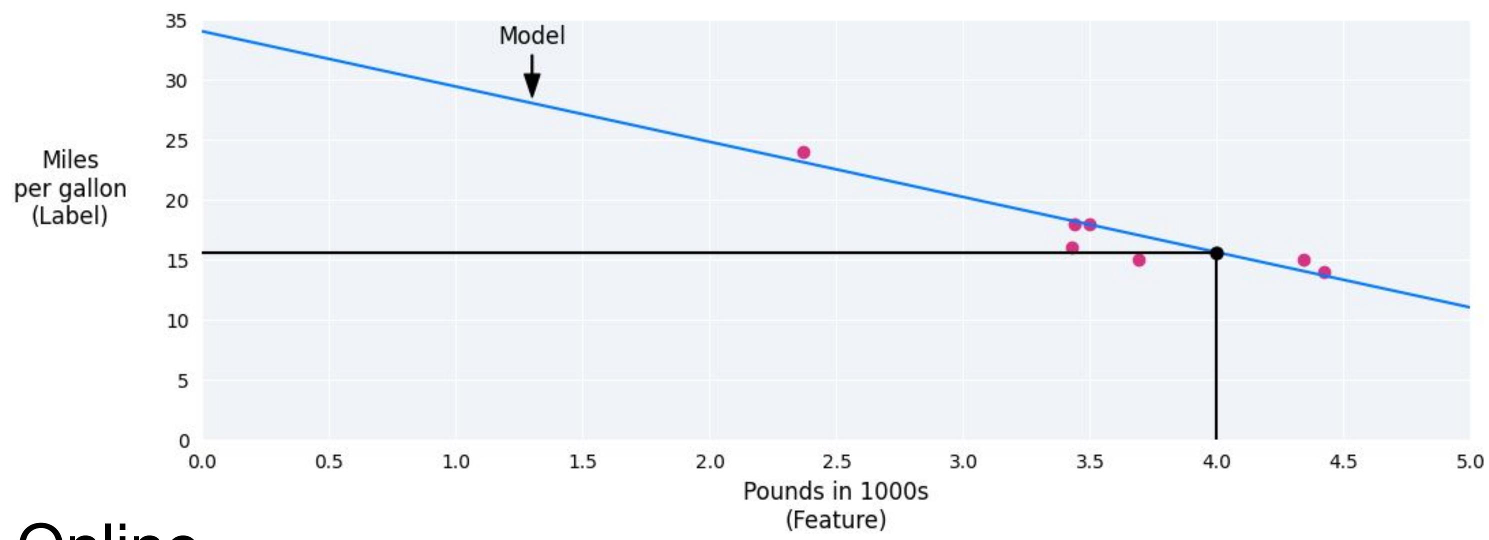
In real world scenario,
$y \Rightarrow$ fuel price, $x \Rightarrow$ quantity of fuel
$y \Rightarrow$ house rent, $x \Rightarrow$ number of rooms

$$y' = b + w_1x_1$$

Prediction Bias Weight Feature value

Calculated from training





[Linear regression]() [Graphing]() Online

Linear Regression

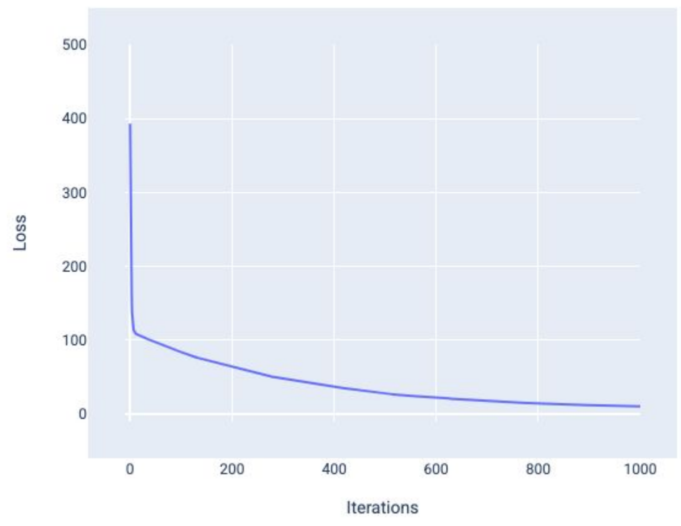
An introduction to linear regression, covering linear models, loss, gradient descent, and hyperparameter tuning.

Linear regression

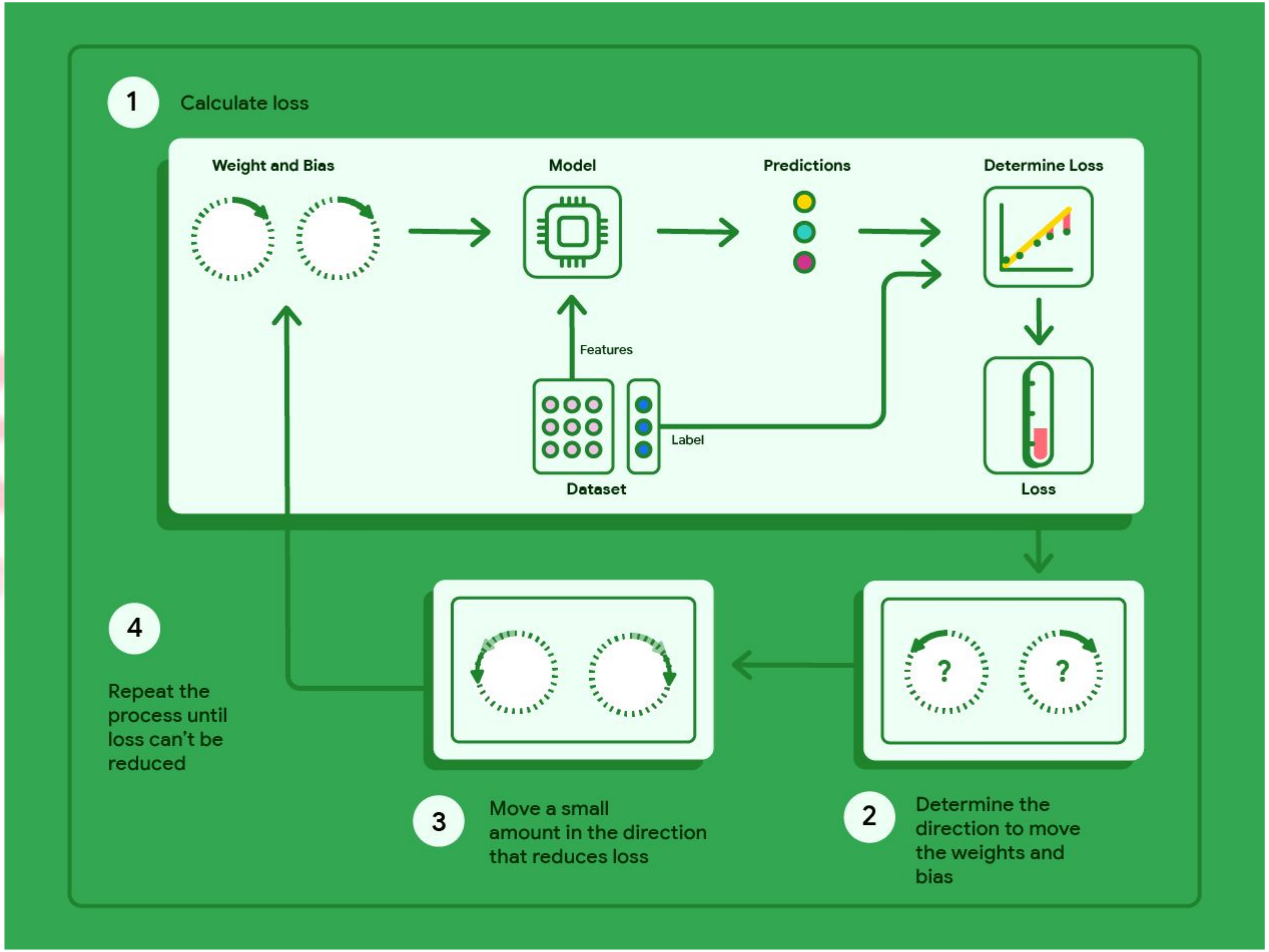
- Loss
- Gradient descent
- Hyperparameters

Loss

Loss type	Definition	Equation
L_1 loss	The sum of the absolute values of the difference between the predicted values and the actual values.	$\sum \lvert actual\ value - predicted\ value \rvert$
Mean absolute error (MAE)	The average of L_1 losses across a set of *N* examples.	$\frac{1}{N} \sum \lvert actual\ value - predicted\ value \rvert$
L_2 loss	The sum of the squared difference between the predicted values and the actual values.	$\sum (actual\ value - predicted\ value)^2$
Mean squared error (MSE)	The average of L_2 losses across a set of *N* examples.	$\frac{1}{N} \sum (actual\ value - predicted\ value)^2$



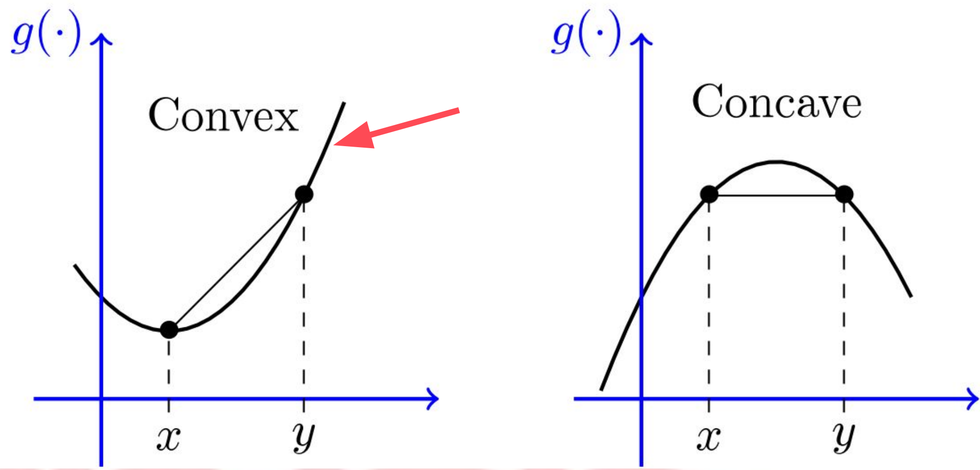
Gradient descent



Gradient descent

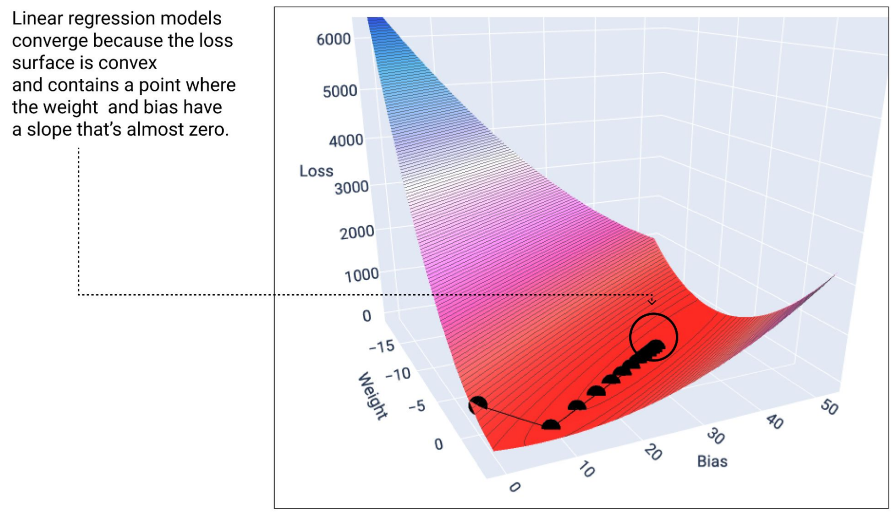
Convex functions

The loss functions for linear models always produce a convex surface



Convergence

A linear model converges when it's found the minimum loss



Disclaimer: This training material belongs to techcrush and shouldn't be shared

Gradient descent

1. function ⇒ f(x)
2. loss
   a. (MSE) ⇒ $(f(x) - y)^2$
3. gradient ⇒ derivative (loss)
4. update ⇒
   a. $\theta_{new} = \theta_{old} - (\propto * \text{gradient})$
   b. θ ⇒weight or bias

$$f_{w,b}(x) = (w * x) + b$$

$$y' = b + w_1 x_1$$

Prediction Bias Weight Feature value

Calculated from training

Derivative

$$\frac{\partial}{\partial w}\left(((w x + b) - y)^2\right) = 2x(b + w x - y)$$

derivative of (((w*x) + b) - (y))^2 with respect to w

Derivative

$$\frac{\partial}{\partial b}\left(((w x + b) - y)^2\right) = 2(b + w x - y)$$

derivative of (((w*x) + b) - (y))^2 with respect to b

wolframalpha

Hyperparameters

parameters are the variables, like the weights and bias, that are part of the model itself.

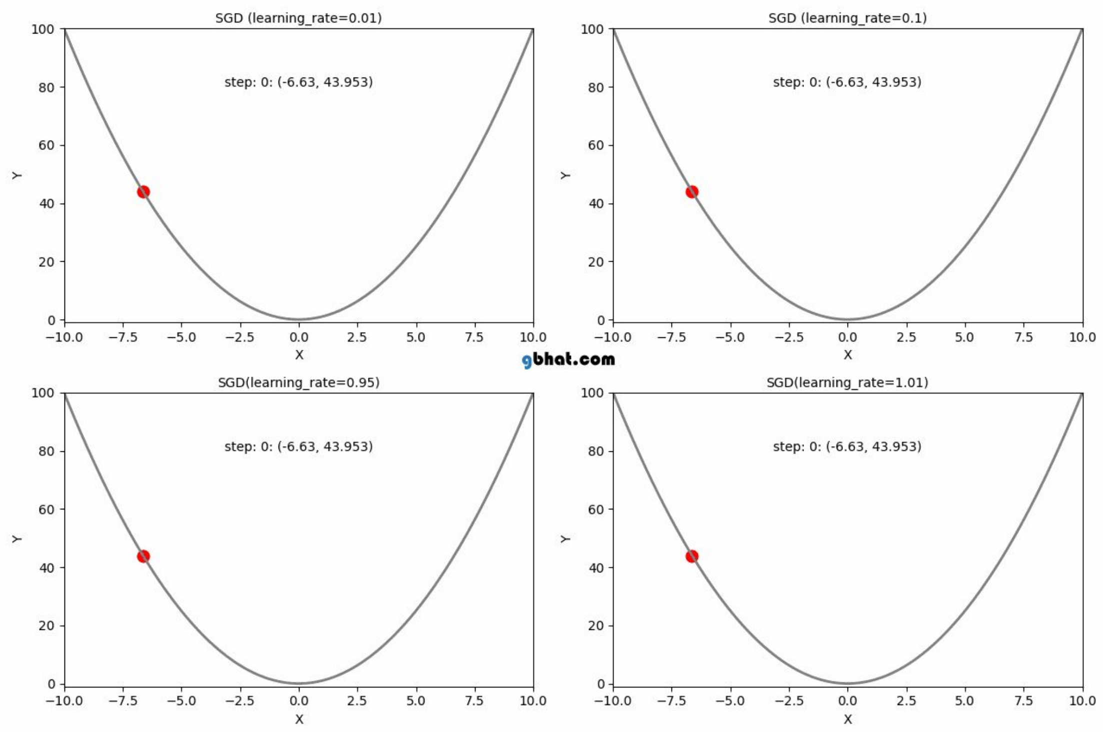
hyperparameters are values that you control

Learning rate (∝) is a floating point number you set that influences how quickly the model converges.

- too low: model can take a long time to converge
- too high: the model never converges

The goal is to pick a learning rate that's not too high nor too low so that the model converges quickly.

Usually between 0 - 1



Hyperparameters

hyperparameters are values that you control

Batch size refers to the number of examples the model processes before updating its weights and bias.

- **Stochastic gradient descent (SGD)**: Stochastic gradient descent uses only a single example (a batch size of one) per iteration.
- **Mini-batch stochastic gradient descent (mini-batch SGD):** The model chooses the examples included in each batch at random, averages their gradients, and then updates the weights and bias once per iteration.

Determining batch size

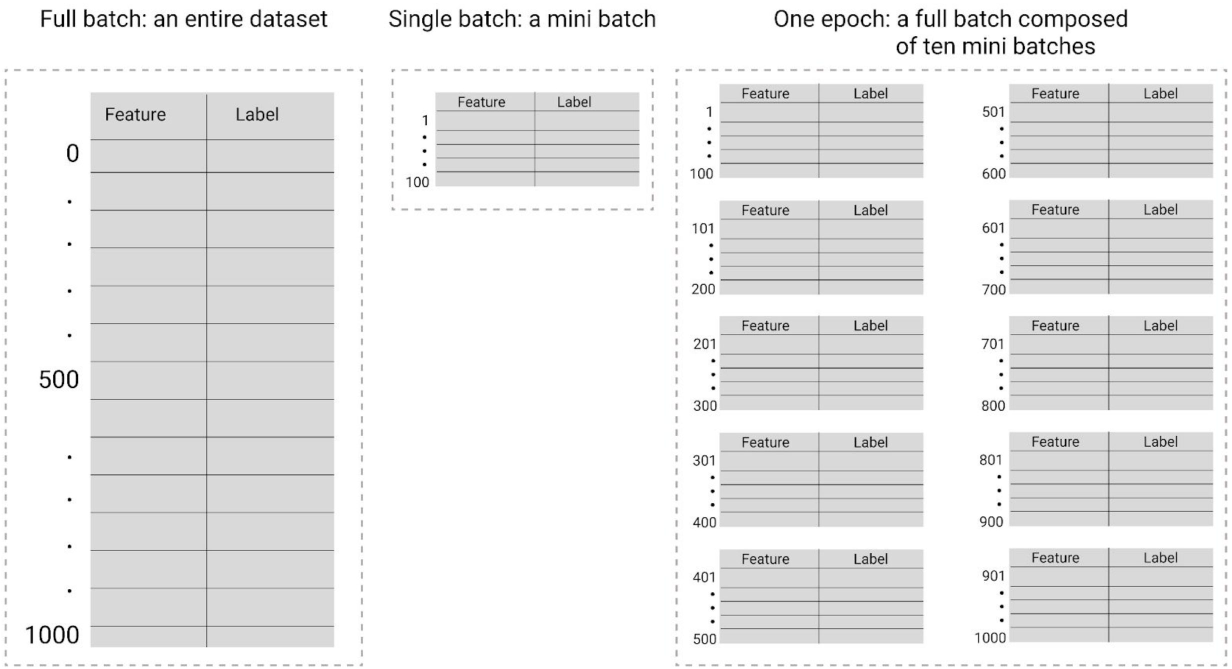
- Usually even number
- dataset and the available compute resources

Hyperparameters

hyperparameters are values that you control

Epochs means that the model has processed every example in the training set once.

For example, given a training set with 1,000 examples and a mini-batch size of 100 examples, it will take the model 10 iterations to complete one epoch.



Python

- **Interpreted language**
- **Syntax & Semantic**
  - **White-Space and Indentation**
- **Comments**
- **Variables & Data Types**
  - **Variable Name**
  - **Case-Sensitivity**
  - Single or Double quotes
  - Data-mismatch
  - Scope (Global & Local)
- **Array-like**
  - **Strings**
  - **Lists**
  - **Tuples**
  - sets
  - **Dictionaries**
- **Operator**
- **Condition**
- Loops
- Functions
- Importing and Installing libraries

Visualize Python Execution

W3School - Python

Notebook commands

- **!   - shell commands**
  - !pip install pandas
- **%   - Line magics**
  - %pip install pandas
- **%%   - Cell magics**
  %%html
  <h1>Hello TechCrush AI</h1>

- Installing Packages: Using PIP
  - pip install numpy
  - pip install pandas
  - pip install -U scikit-learn
  - pip install matplotlib
- Importing Packages
  - import math
  - import pandas as pd
  - from matplotlib import pyplot as plt
  - from sklearn.linear_model import LinearRegression

Questions?