Phase 3 Expectations

// FLATIRON SCHOOL

Agenda

1. Phase 3 Overview
2. Week 1: Math & Classification
3. Week 2: More Models!
4. Gating / Assessments
5. What Will Phase 3 Feel Like?

Phase Overview

PHASE 1
Weeks 1 - 3

Data Engineering

Statistics

PHASE 2
Weeks 4 - 6

PHASE 3
Weeks 7 - 9

Machine Learning

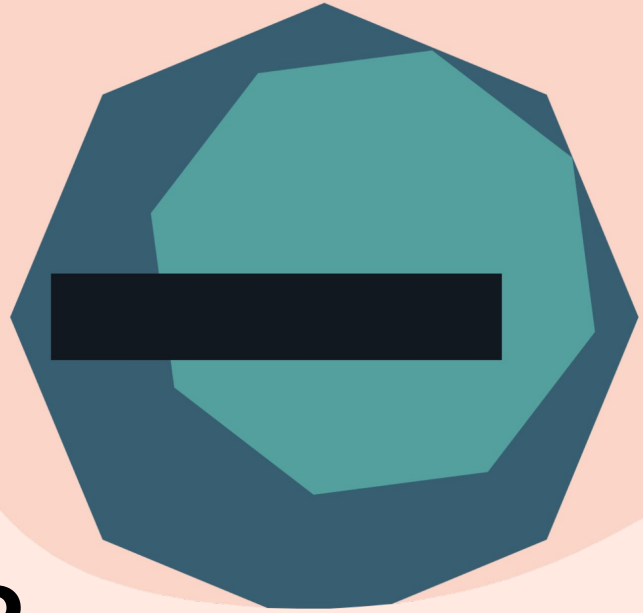
Advanced Topics

PHASE 4
Weeks 10 - 12

PHASE 5
Weeks 13 - 15

Capstone

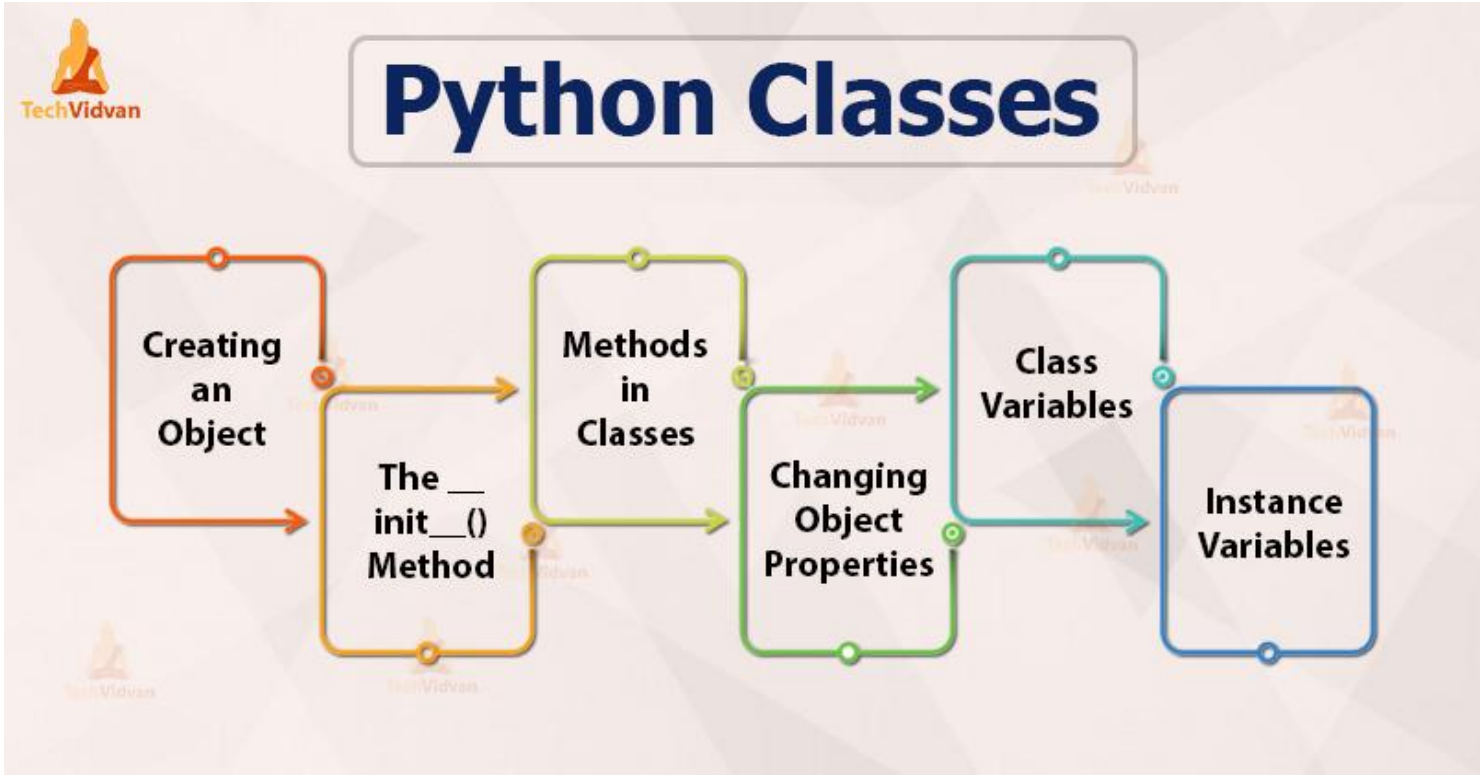
Week 1



More Linear Regression, Predictive Modelling, OOP, Little bit of Math

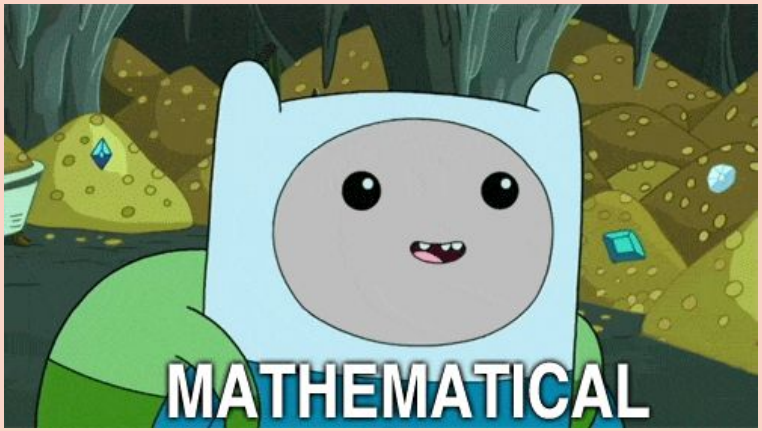
// FLATIRON SCHOOL

Object-Oriented Programming

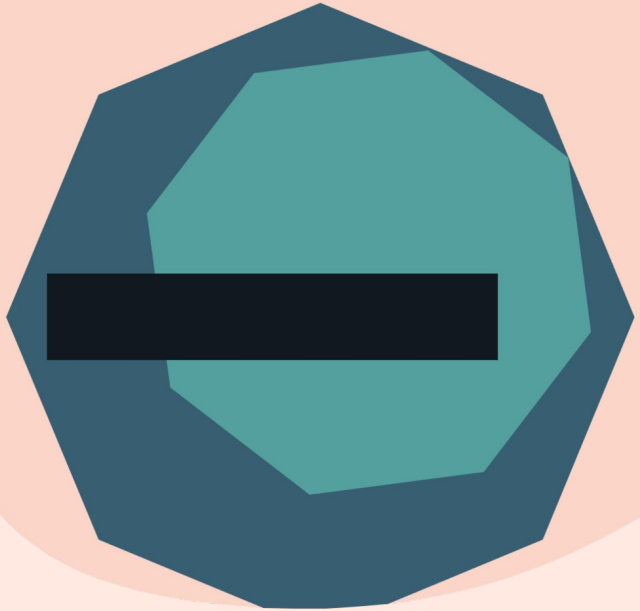


The Math Behind Data Science

- Linear Algebra
- Calculus, Cost Functions, Gradient Descent
- Focus on concepts and application!



Week 2

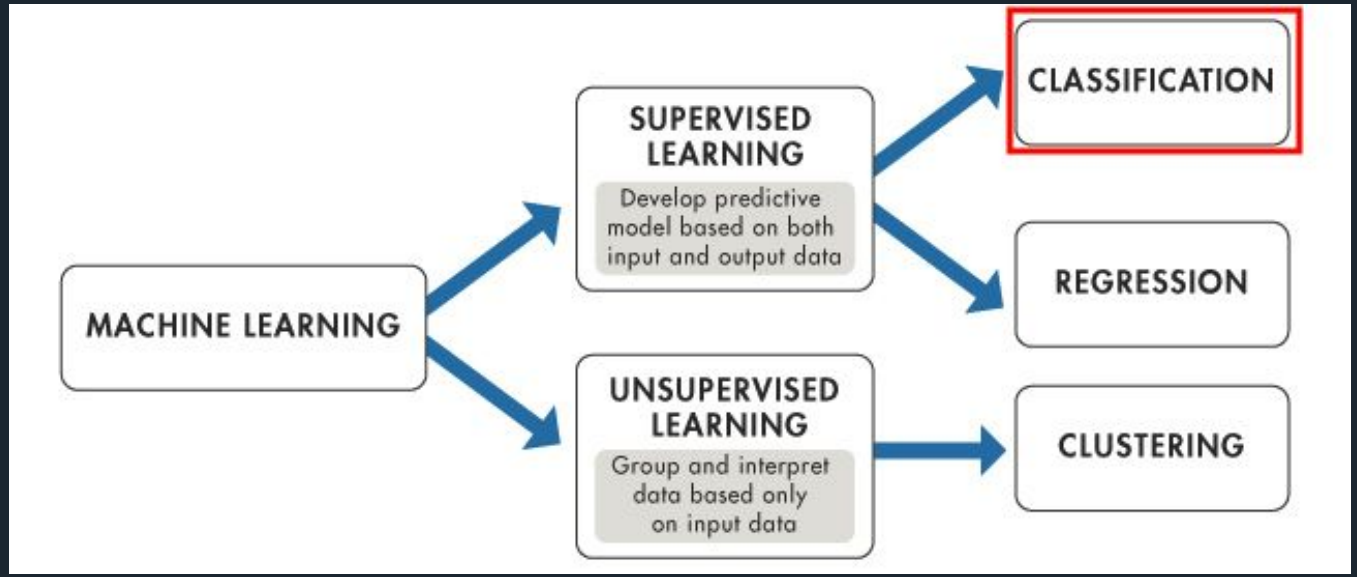


Logistic Regression, Model Tuning, Evaluation

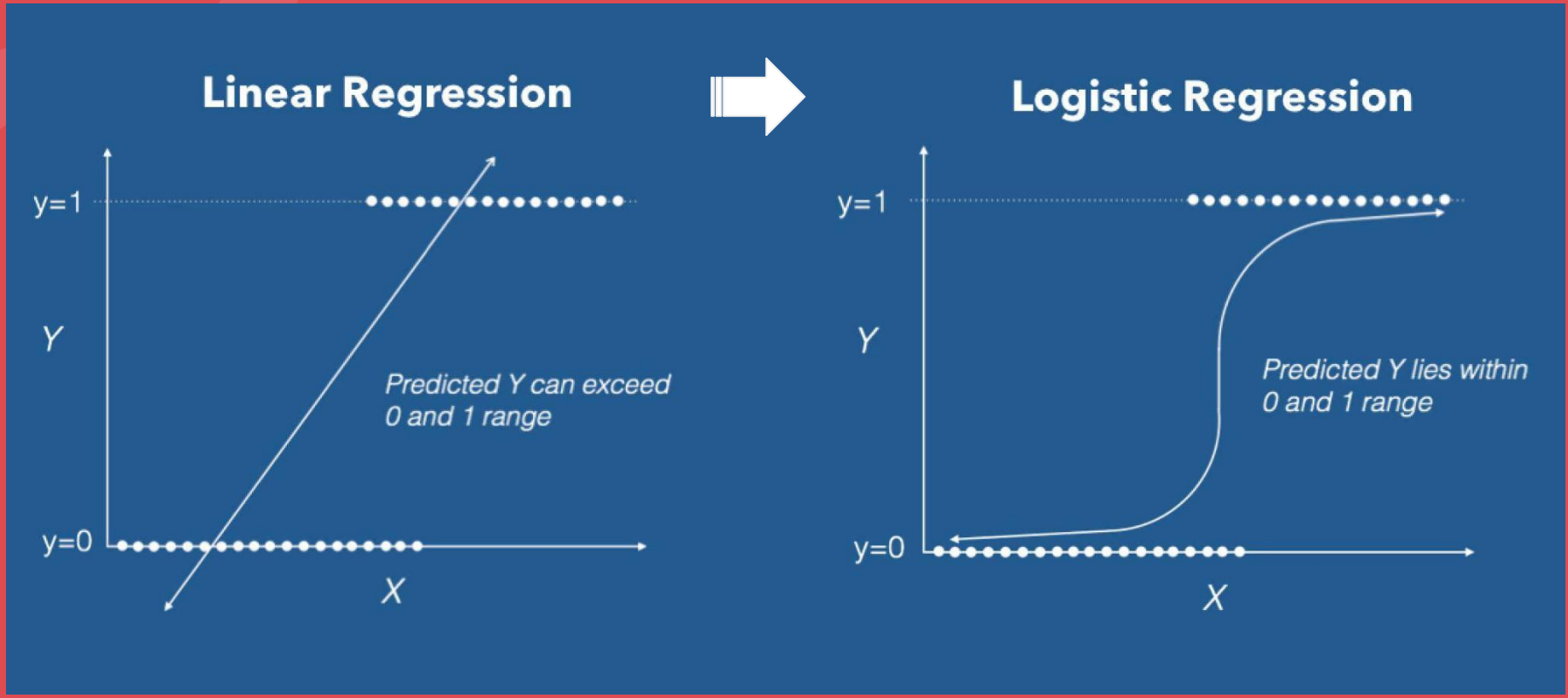
// FLATIRON SCHOOL

Classification

- Asking: Is it ____ or not?
- From continuous target to categorical target

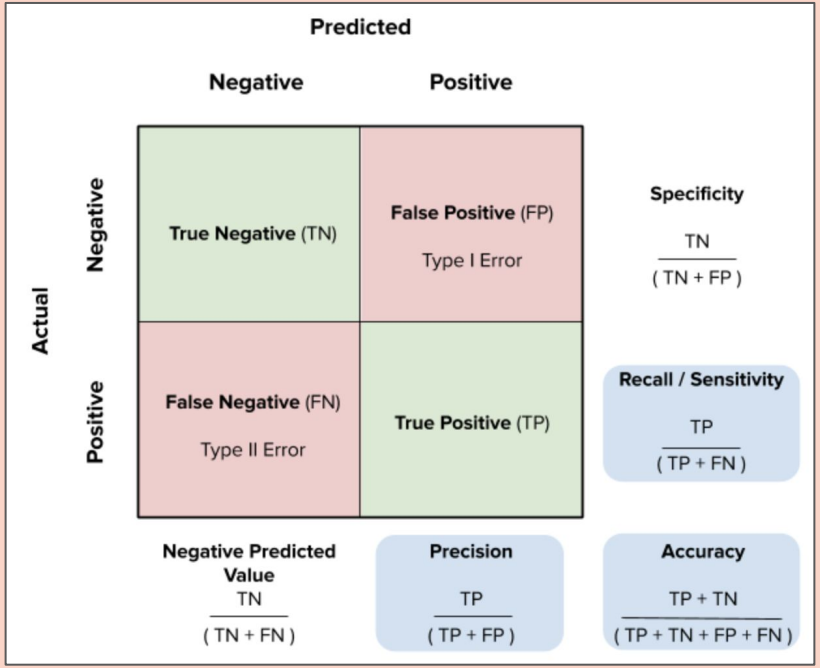


Logistic Regression



Classification Metrics

- New type of target - so, new metrics!



Student Progress and Assessments

Blog Post

Encouraged but not required

- Write a tutorial with code
- See next slide

Required Quizzes

1. Multiple Linear Regression
2. Data Ethics - Training on Bad Data
3. Decision Trees - Classification

Checkpoints & Code Challenge

1. Linear Regression - Wed 05/15
2. ML Fundamentals - Mon 05/20
3. Code Challenge - Thurs 05/23
    a. GD, Logreg, Metrics, Decision Tree

Phase 3 Project

Classification ML Project

- Predictive (and) Inferential modelling
- Datasets provided or find your own
- Iterative process

Phase 3 Blog Post

Write a tutorial (with a data set and code sample) on something that you think might be interesting to other people taking the course. It can be a topic we didn't cover at all, or can just go deeper into a topic that we did cover. Your tutorial can use the same tool/library from the previous blog post, or you can choose something new. Potential elements to include:

I. An introduction explaining why a data scientist would want to do what your tutorial does
II. A section explaining the data set you're working with (what are the features? if there is a target, what is it?)
III. A section explaining the libraries you're working with (ideally with links to documentation and/or tutorials for someone just getting started)
IV. Well-commented code
V. A conclusion section that recaps what you did and why

Feelings at the Start of Phase 3

Excited: Let's tackle Machine Learning!

Little Overwhelmed: Lots of new tools

Feelings at the End of Phase 3

Powerful: Modeling toolbelt

Capstone Stress Creep

Capstone Project

// FLATIRON SCHOOL

- Sneaking up!
- What can you do now?
  - Explore data!
    - UCI Machine Learning Datasets Repository
    - Kaggle Datasets
    - Awesome Datasets Repo on Github
    - New York City Open Data Portal
    - Inside AirBNB
    - Data is Plural
    - Image Classification
    - Rec Systems - Ecommerce