Data Science Bootcamp

<> ML crash course

Content designed and curated by DataRockie 2022

Content 🔥

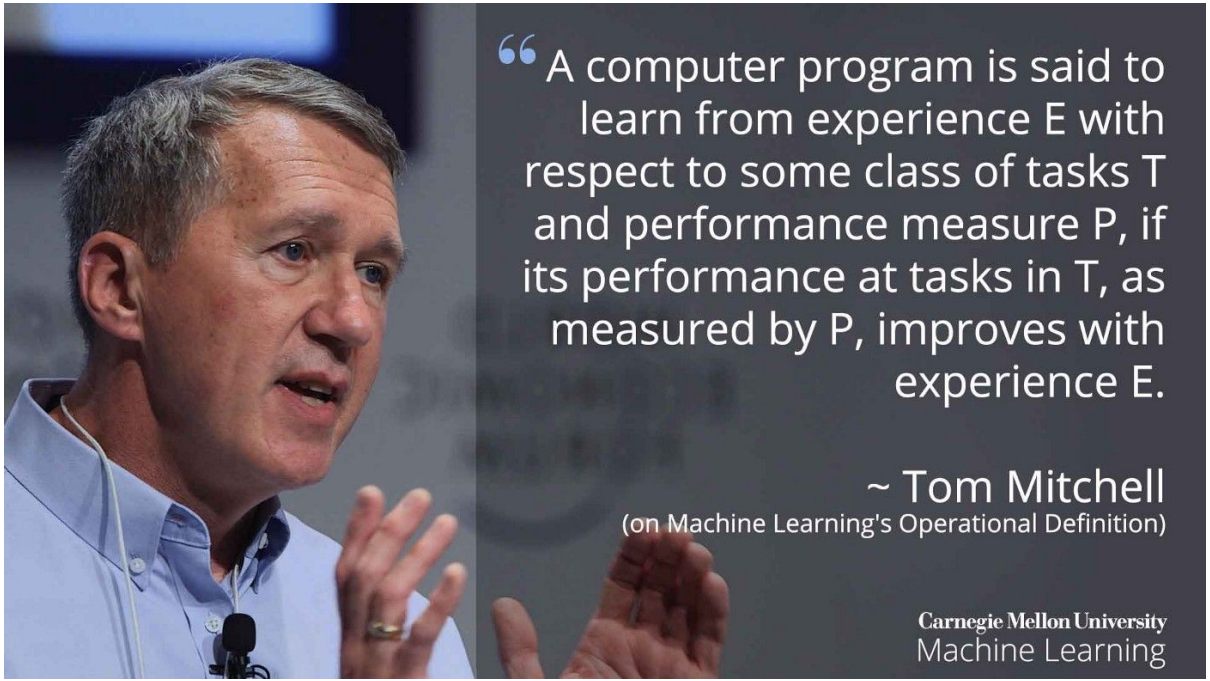
- What is ML
- Types of ML
- ML workflow
- R Caret
- Regression Model
- Classification Model

What is ML

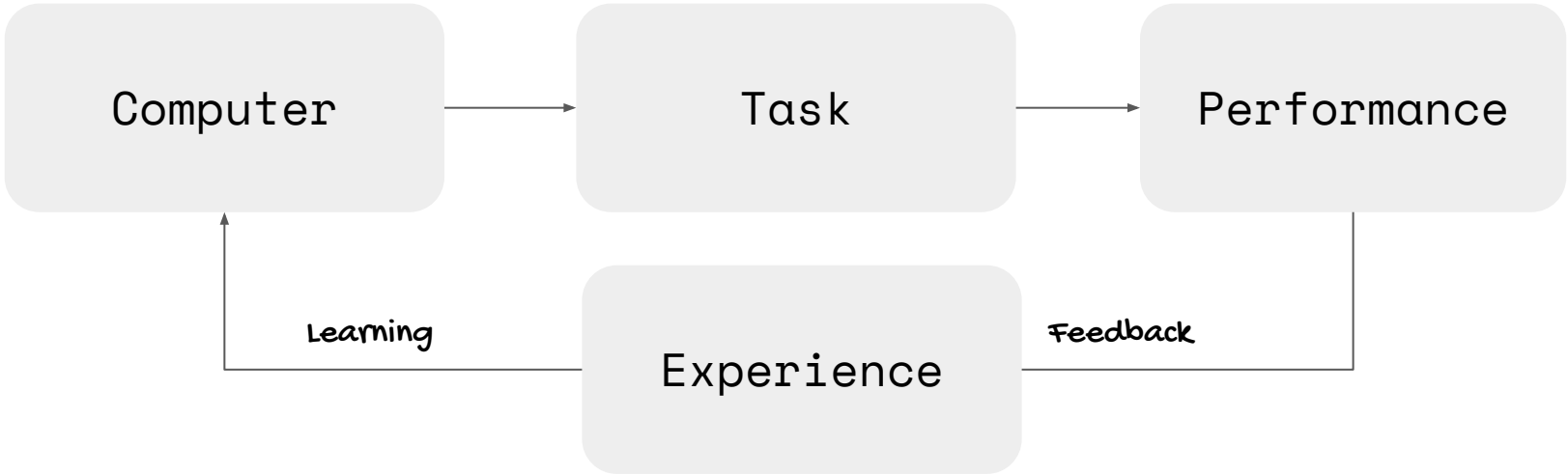
What is ML

Field of study that gives computers the **ability to learn** without being explicitly programmed

Arthur Samuel (1959)


A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.
~ Tom Mitchell
(on Machine Learning's Operational Definition)
Carnegie Mellon University
Machine Learning

The Idea



ML is everywhere

Give it a shot: Check out these 10 new camera upgrades on Pixel 7 and 7 Pro

Oct 06, 2022
6 min read

Here's a look at how camera upgrades to the Pixel 7 and Pixel 7 Pro pushes computational photography to new heights.

I **Isaac Reynolds**
Group Product Manager, Pixel Camera Software

A **Alexander Schiffhauer**
Group Product Manager, Pixel Camera Hardware & Research

Share

https://blog.google/products/pixel/pixel-7-camera/

We'll focus on Classical ML

We will build a lot of **predictive models** in this sprint

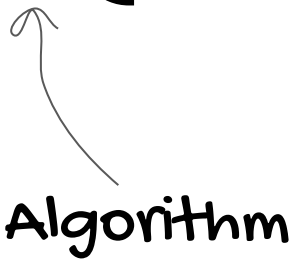
- Linear Regression
- Logistic Regression
- Regularized Regression
- Decision Tree
- KNN
- Simple Neural Nets

What is Model

Algorithm + Data = Model

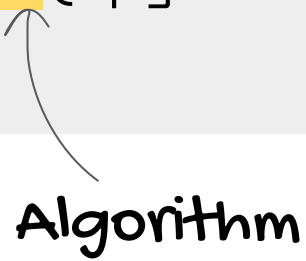
- Linear Regression
- Logistic Regression
- Decision Tree
- KNN

$$Model = f(Data)$$



R is A Great Example

```
model = lm(mpg ~ hp + wt + am , data = mtcars)
```



What does model look like?

Some models have equations

```
mpg = 30 + 0.5*hp + 0.34*wt + 1.5*am
```

Some models don't

```
IF (hp > 200)  AND (am = 0) AND (wt < 2) THEN mpg = 32.5
IF (hp <= 200) AND (am = 1) AND (wt < 2) THEN mpg = 39.2
ELSE mpg = 29.8
```


Types of ML

- Supervised Learning 🔥
- Unsupervised Learning
- Reinforcement Learning (RL)

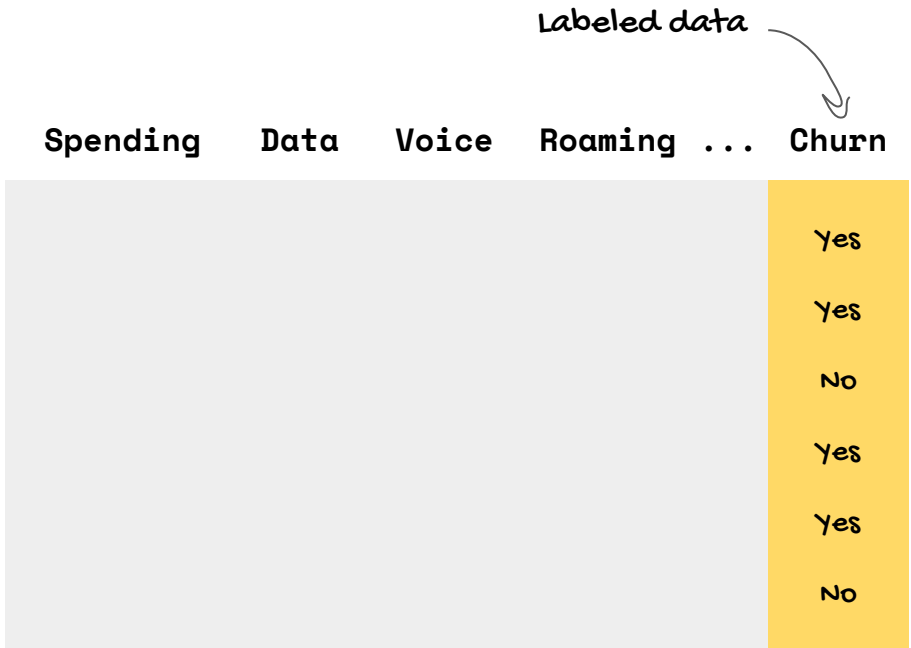
Supervised Learning

aka. predictive models/ analytics

Learn from **labeled data** to make a prediction

Supervised vs. Unsupervised Learning: What's the Difference? | IBM

What does labeled data mean?



Supervised Learning

aka. predictive models/ analytics

- Regression (numeric)
- Classification (category)

Supervised vs. Unsupervised Learning: What's the Difference? | IBM

Unsupervised Learning

e.g. how many customer segments do we have in our database

Learn from **unlabeled data** to find patterns and summarise data

Supervised vs. Unsupervised Learning: What's the Difference? | IBM

What does unlabeled data mean?

Spending	Data	Voice	Roaming	...

Let’s the computer find insights/ patterns for me

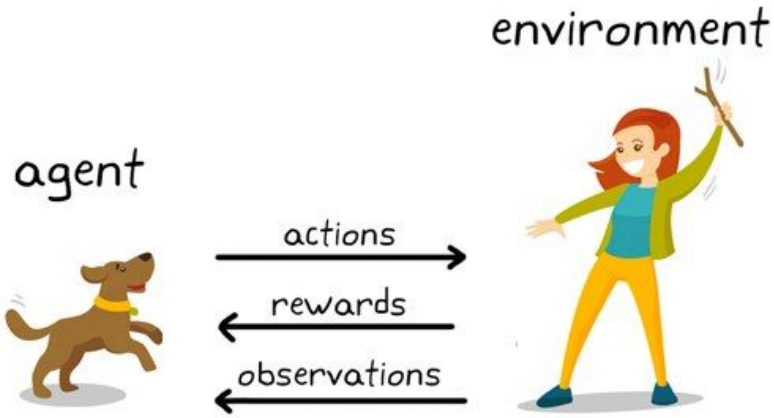
Unsupervised Learning

e.g. how many customer segments do we have in our database

- Clustering
- Association
- Dimension Reduction

Supervised vs. Unsupervised Learning: What's the Difference? | IBM

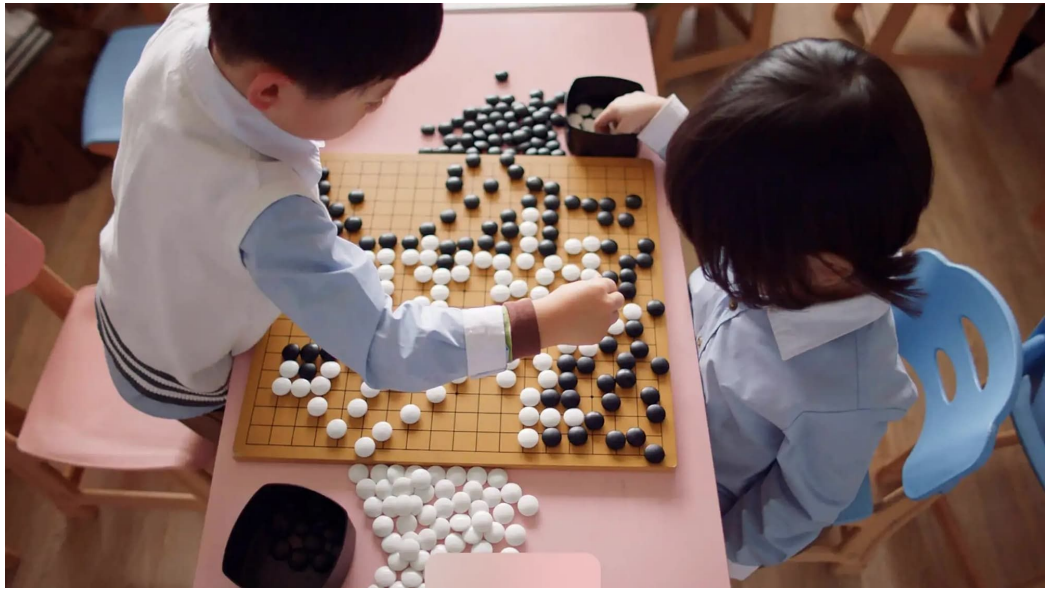
Reinforcement Learning



Agent learns to **maximize rewards**

Three Things to Know About Reinforcement Learning - KDnuggets

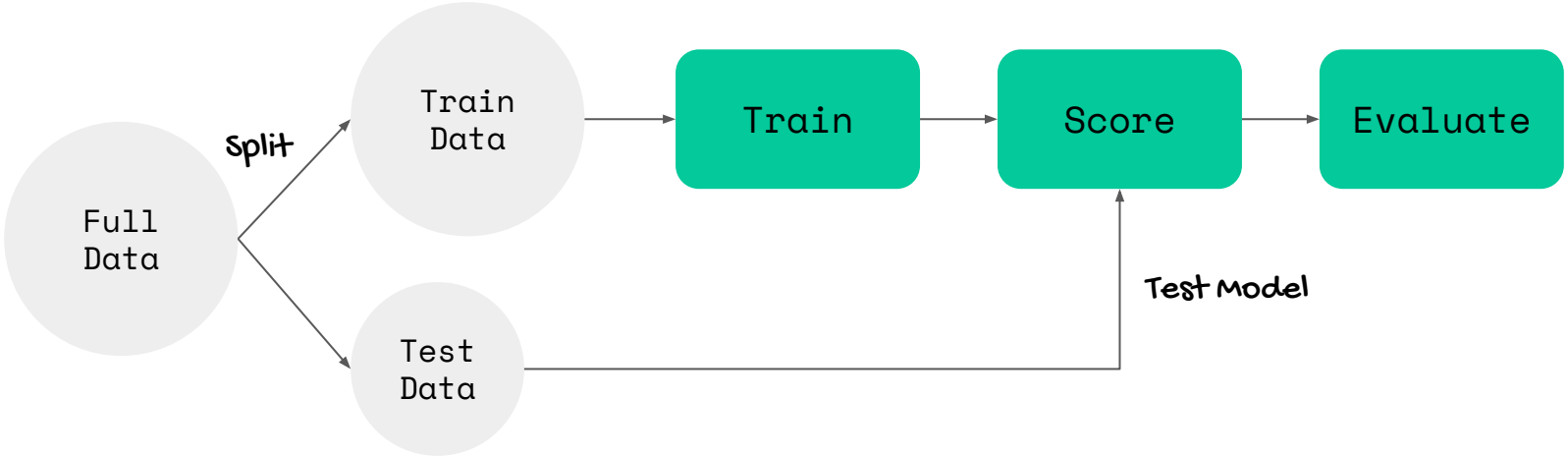
Alphago



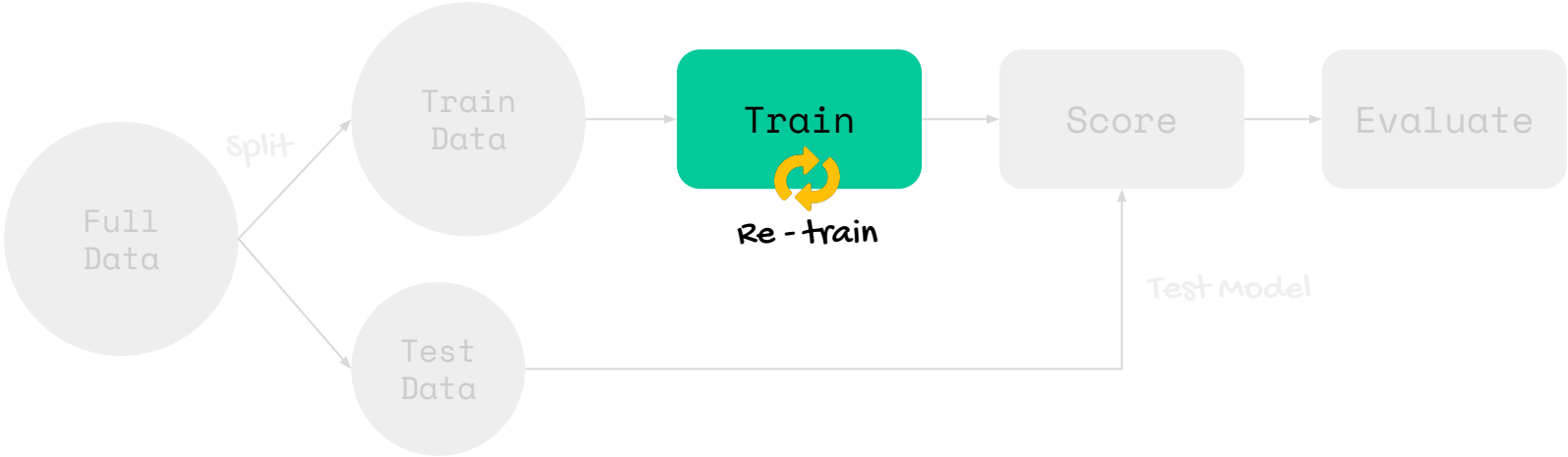
https://www.deepmind.com/research/highlighted-research/alphago

ML Workflow

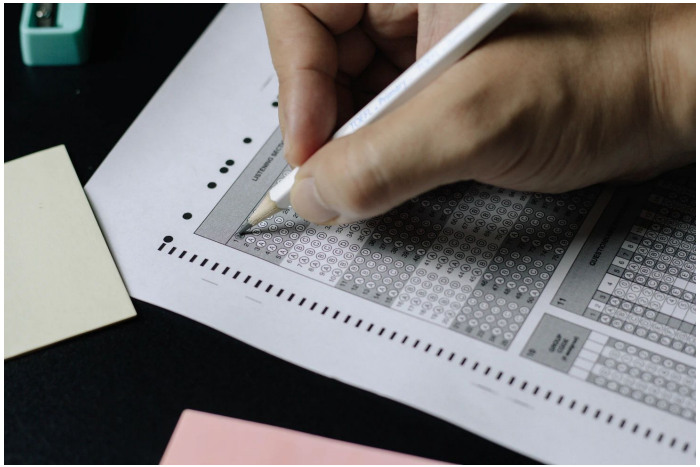
Basic ML Workflow



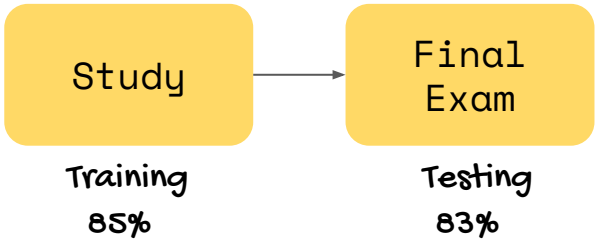
Basic ML Workflow



Testing Model



Testing model is a very important step in ML workflow



R Caret

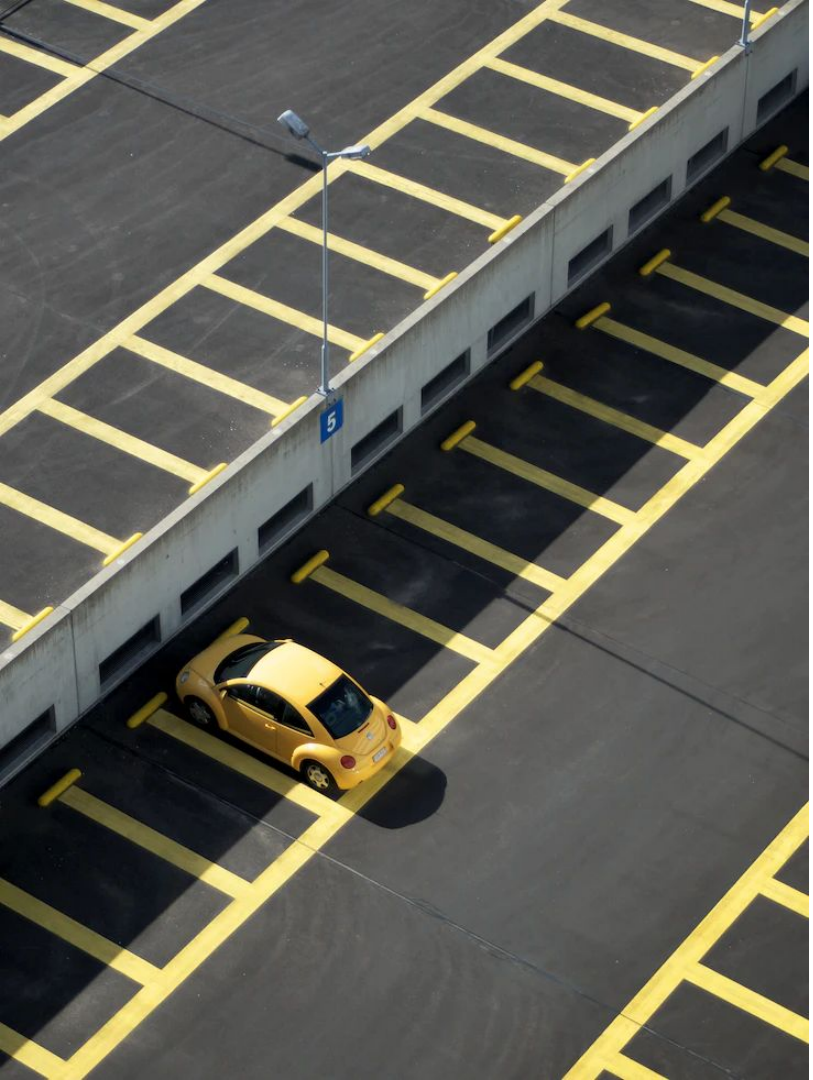
Classification **A**nd
Regression **T**ree

Max Kuhn - Applied Predictive Modeling

Install Package

```
# RStudio Cloud
install.packages("caret")
```


Regression Model



Build a simple model using **mtcars** dataset

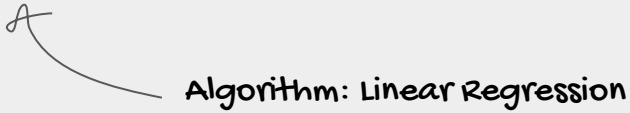
- Regression predicts mpg
- Classification predicts am

Train Test Split

```
set.seed(42)
n ← nrow(mtcars)
id ← sample(n, 0.8*n)

train_data ← mtcars[id, ]
test_data ← mtcars[-id, ]
```


Simple Regression

```
library(caret)

model ← train(mpg ~ hp,
              data = mtcars,
              method = “lm”)

print(model)
```



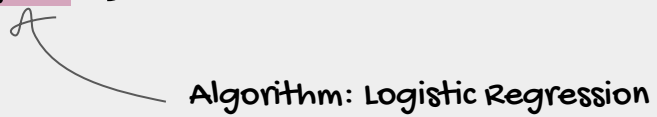
Classification Model

Train Test Split

```
# make sure y is factor
mtcars$am ← factor(mtcars$am)

# split data
set.seed(42)
n ← nrow(mtcars)
id ← sample(n, 0.8*n)

train_data ← mtcars[id, ]
test_data ← mtcars[-id, ]
```


Simple Classification

```
library(caret)

model ← train(am ~ mpg,
              data = mtcars,
              method = "glm")

print(model)
```



Key Takeaways

- ML algorithms learn from data
- Supervised learning learns from labeled data
- Unsupervised learning learns from unlabeled data
- Basic ML workflow
  - split data > train > score > evaluate
- We will use `caret` to build models in R

Data Science Bootcamp

<> ML crash course

Content designed and curated by DataRockie 2022