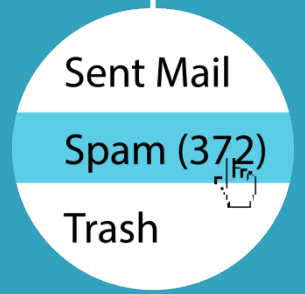
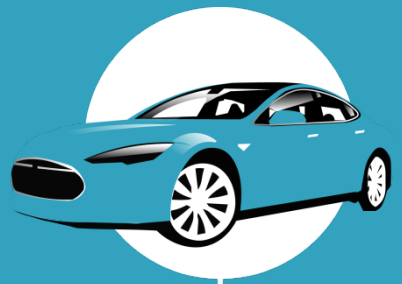
Module Overview

- What is Machine Learning?
- Machine Learning vs Traditional Development
- Types of Machine Learning
- Course Content
- Machine Learning and Data Science
- Python and Jupyter Notebook Demo

Is this email spam?

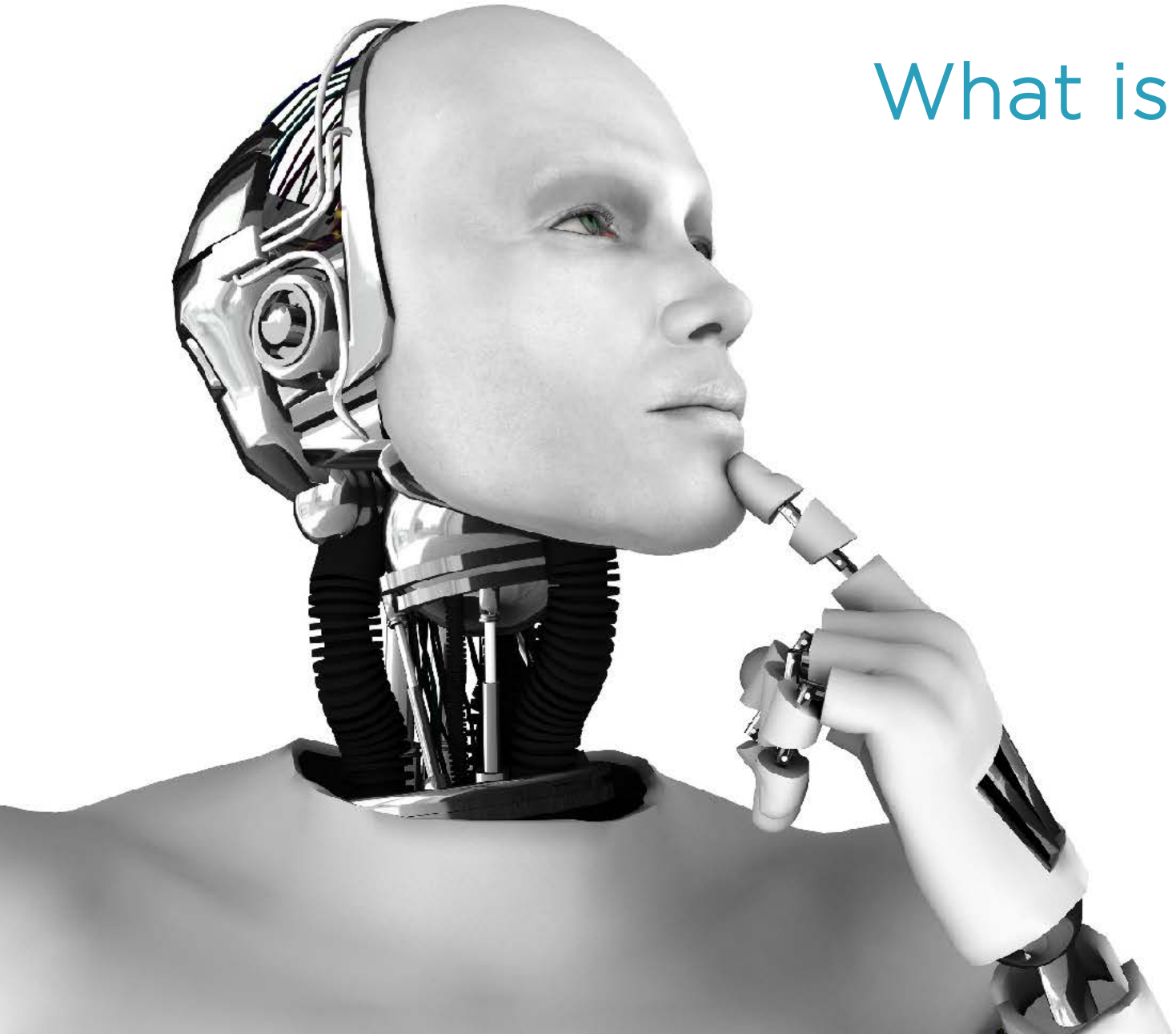




How can cars drive themselves?

What will people buy?

Machine Learning in Action



What is Machine Learning?

Machine Learning

Building a model from example inputs to make data-driven predictions vs. following strictly static program instructions.

Machine Learning

Building a **model from example inputs** to make data-driven predictions vs. following strictly static program instructions.

Machine Learning

Building a model from example inputs to make data-driven predictions vs. following strictly **static program instructions**.

Traditional Programming

Traditional Control Logic

If

Case

While

Until

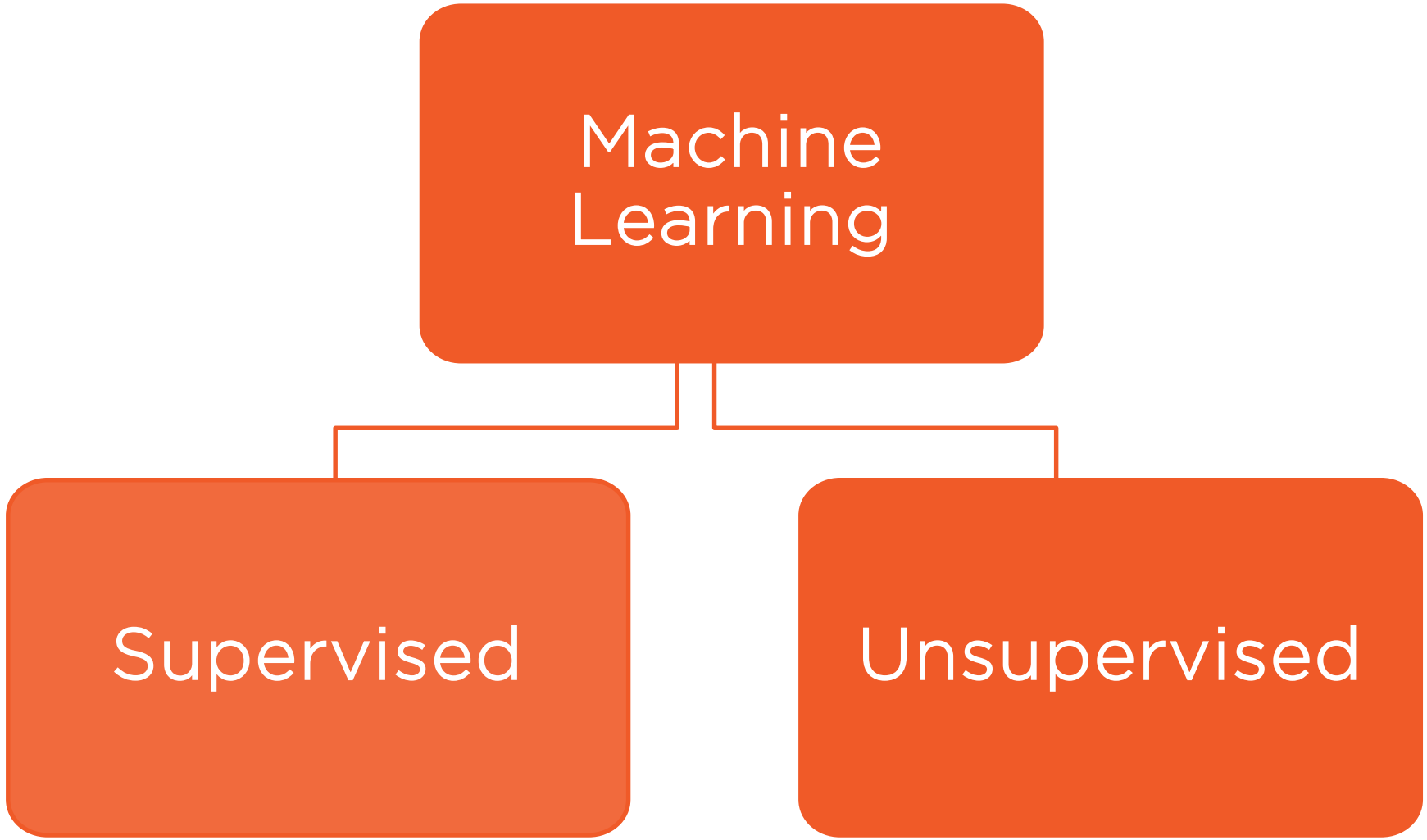
Machine Learning Logic

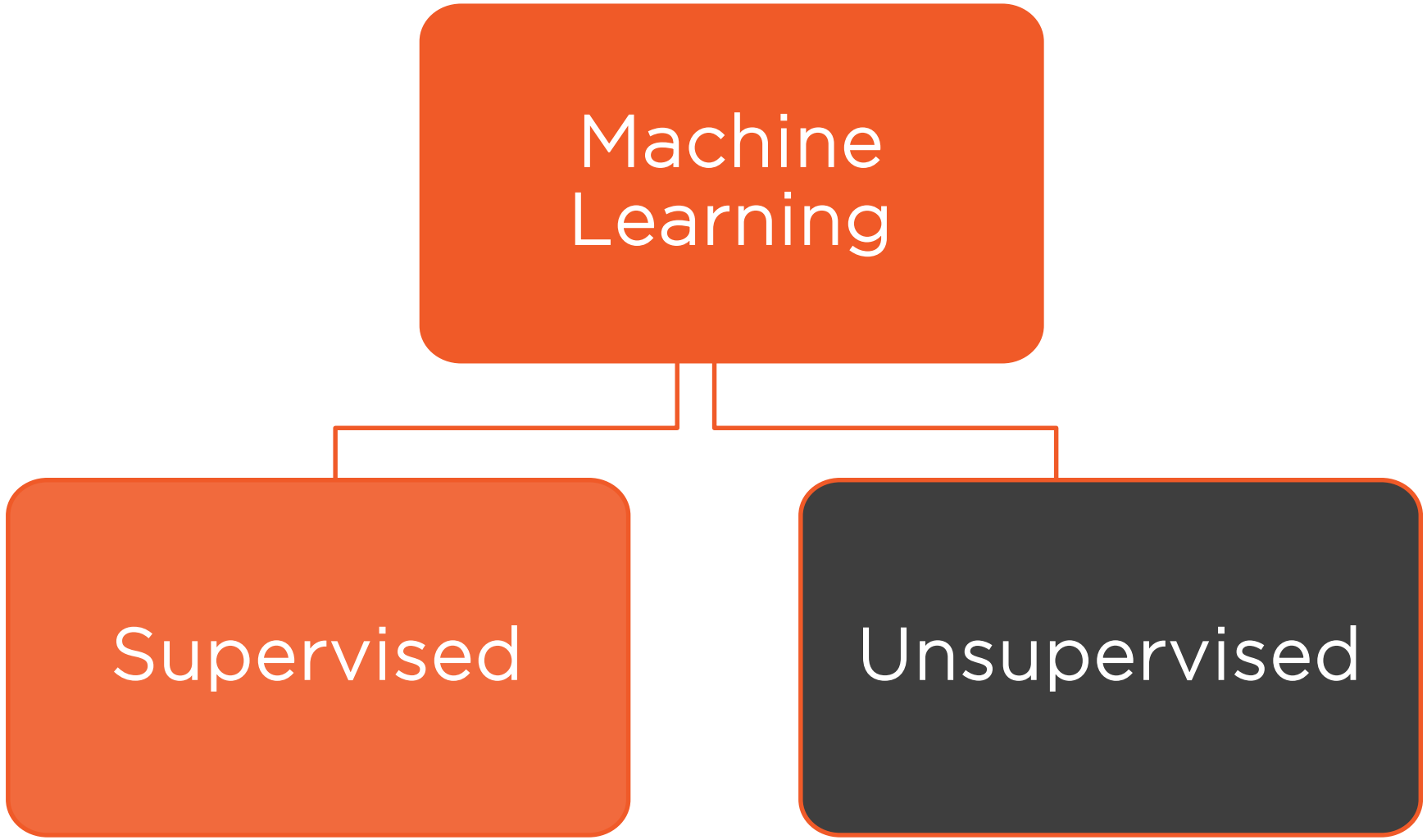
Data

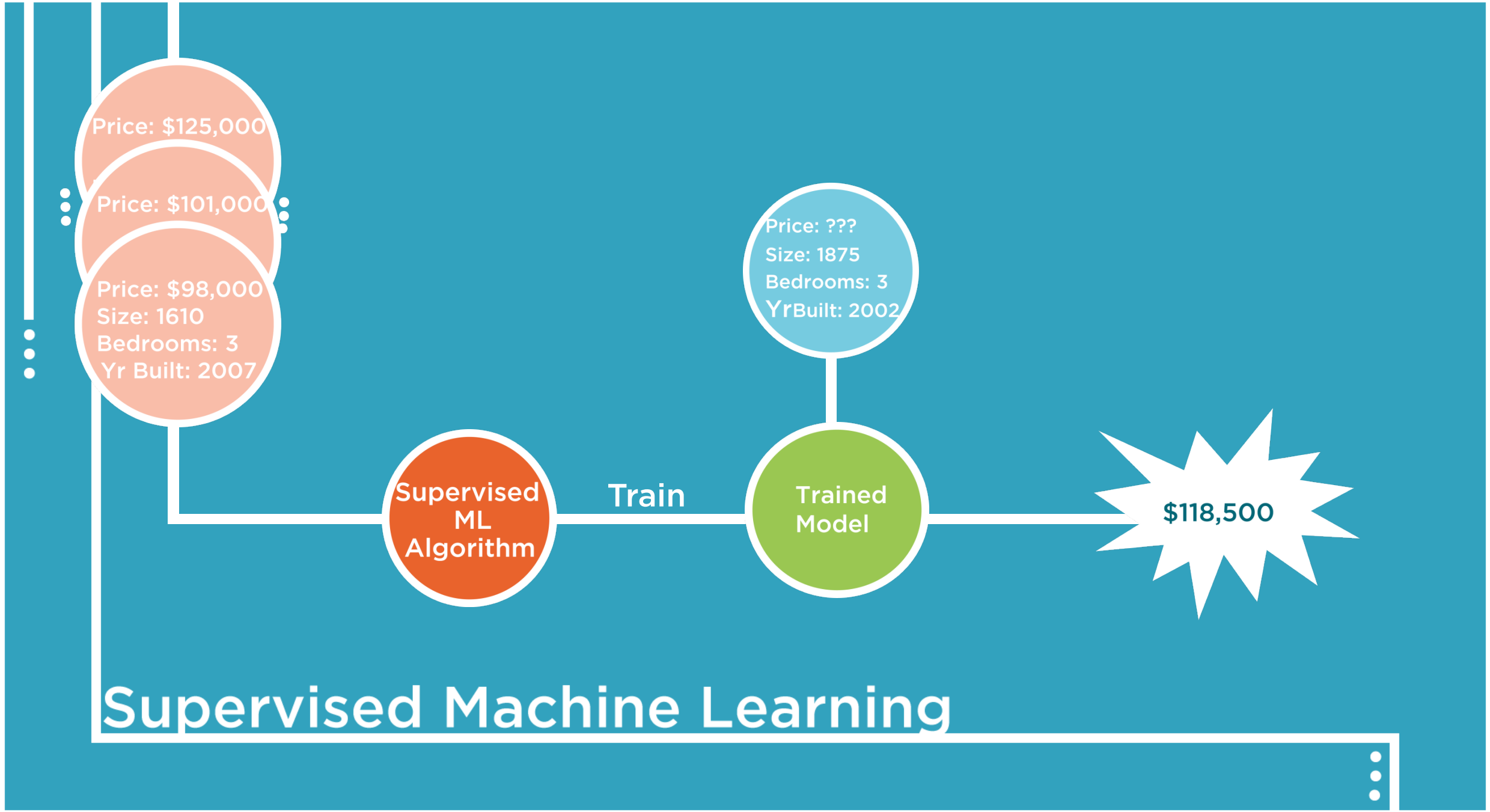
Algorithm

Data Analysis

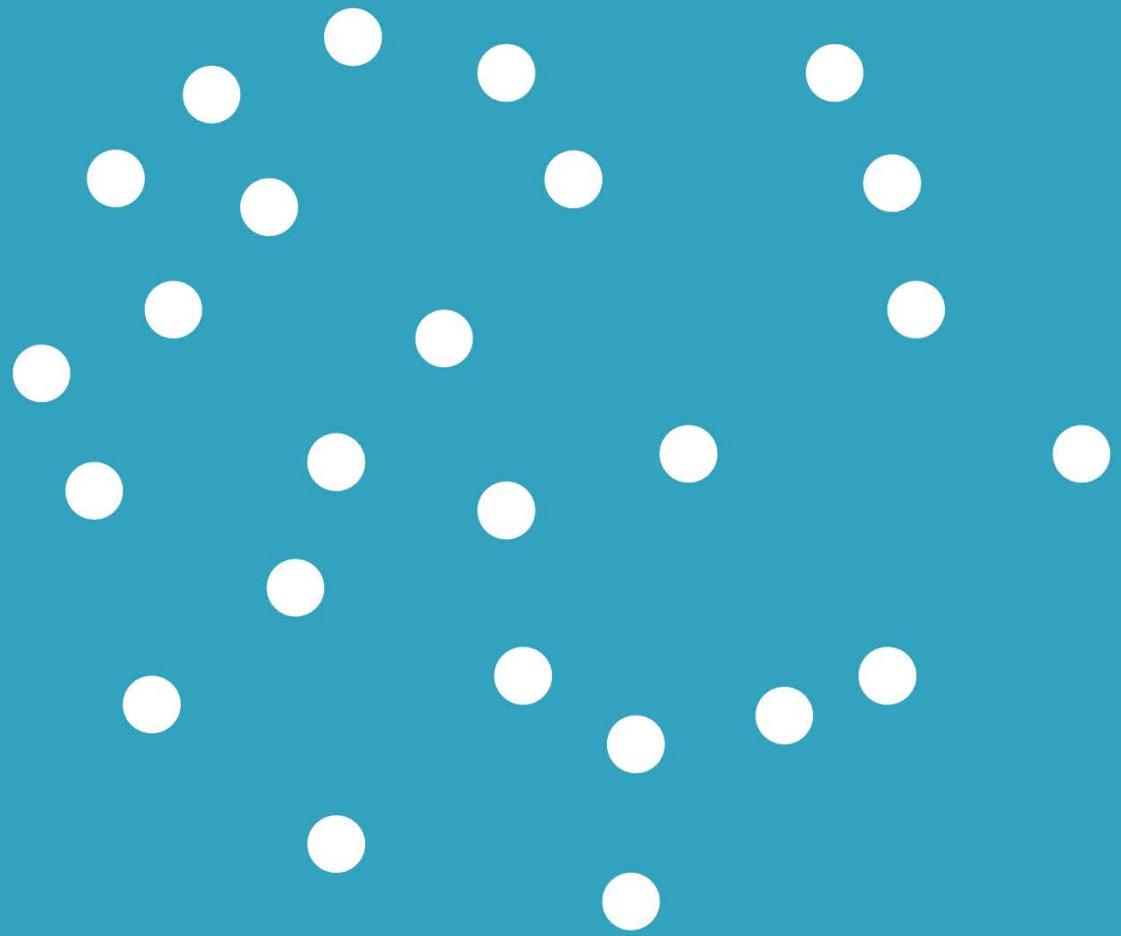
Model


Machine Learning
Supervised
Unsupervised

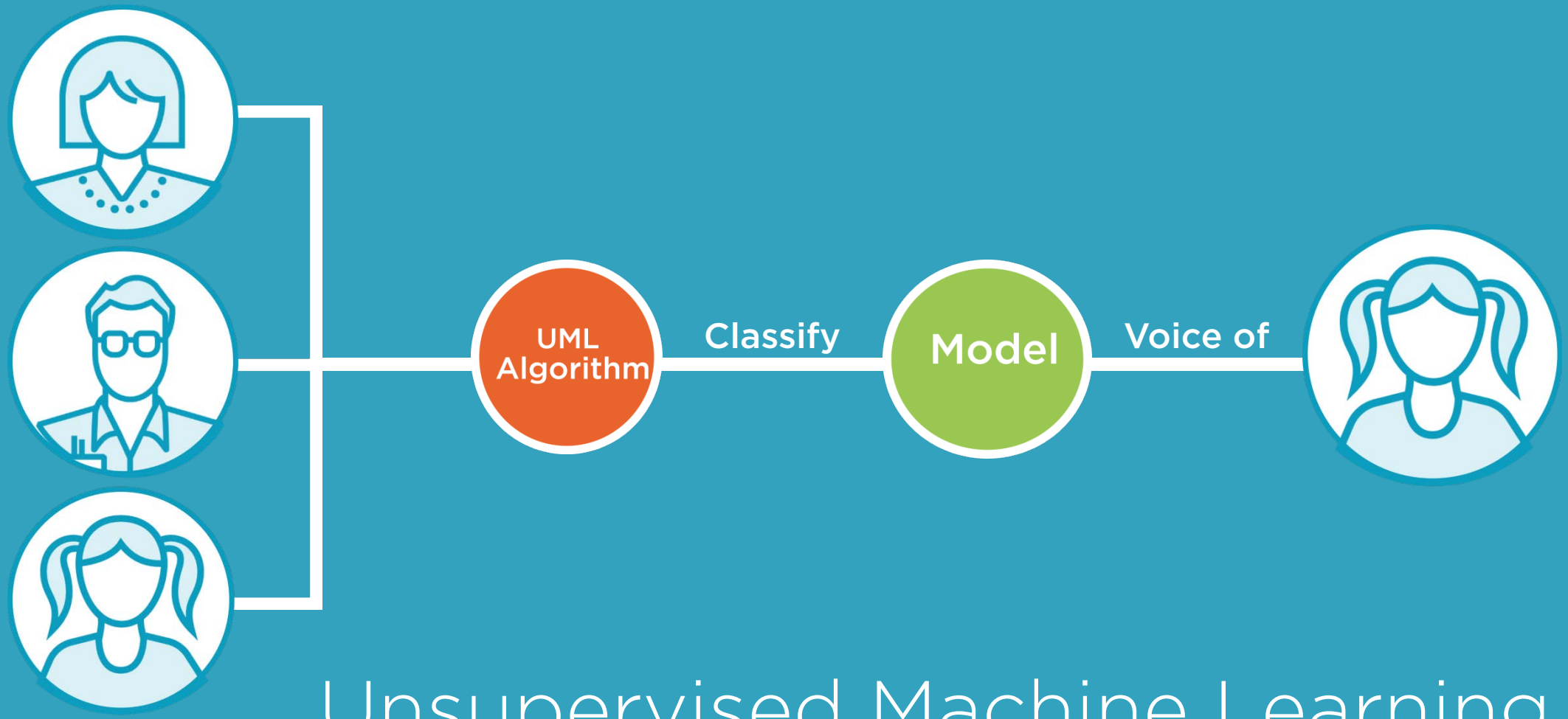

Machine Learning
Supervised
Unsupervised


Price: $125,000
Price: $101,000
Price: $98,000
Size: 1610
Bedrooms: 3
Yr Built: 2007
Price: ???
Size: 1875
Bedrooms: 3
YrBuilt: 2002
Supervised ML Algorithm
Train
Trained Model
$118,500
Supervised Machine Learning

Unsupervised Machine Learning



Unsupervised Machine Learning



Unsupervised Machine Learning

Machine Learning Technique Comparison

Supervised	Unsupervised
Value prediction	Identify clusters of like data
Needs training data containing value being predicted	Data does not contain cluster membership
Trained model predicts value in new data	Model provides access to data by cluster
Subject of this course	Not in this course

Course Overview

Machine Learning Workflow

Applying the Workflow Steps

Summary

Your Skills

Not Required

Experience in Python

Experience with Jupyter Notebook

Advanced statistics or math

Required

Software development experience

Experience with data in tables

Basic math and statistics skills

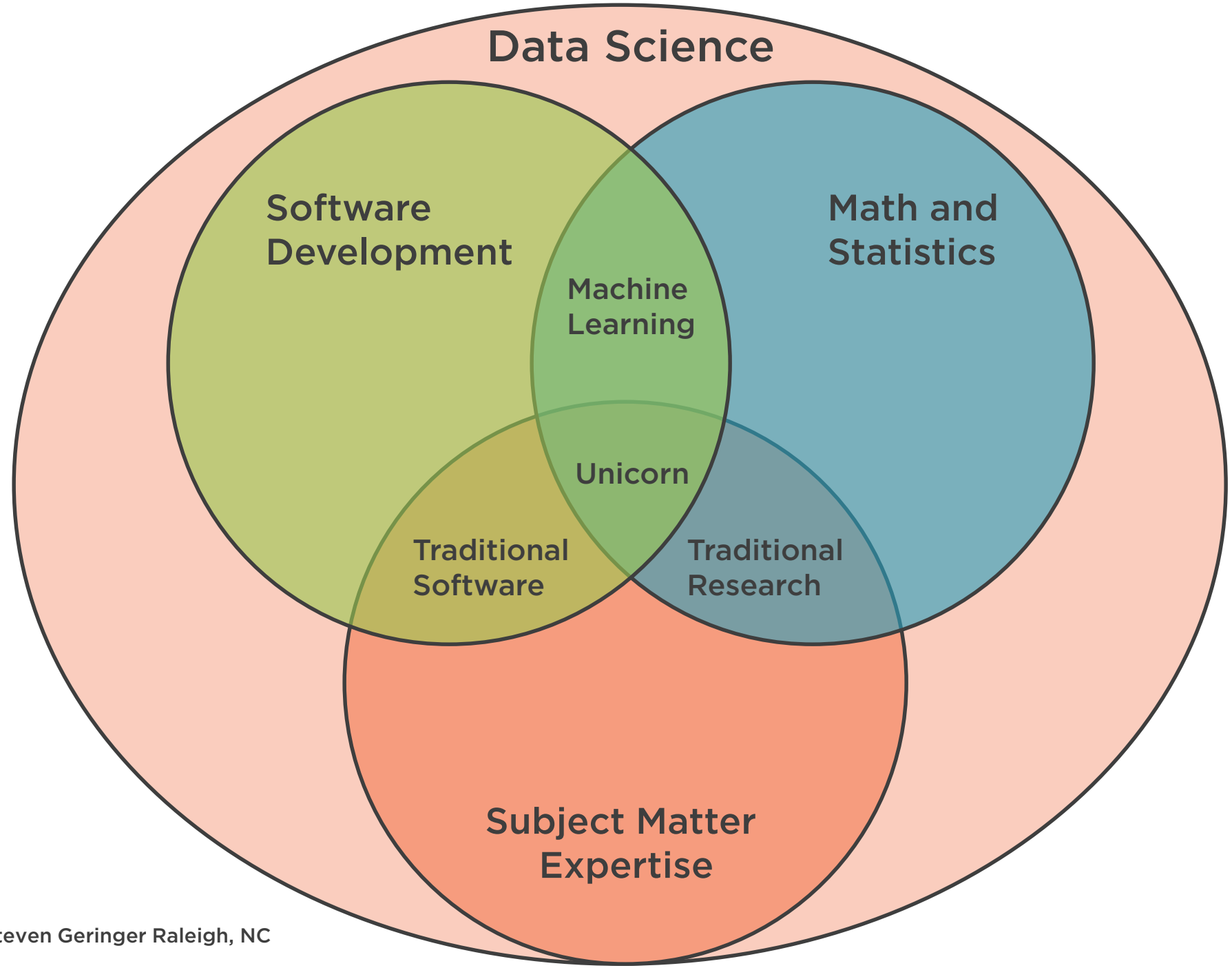
Passion to understand

Why This Course?

Add Machine Learning skills

Learn something new

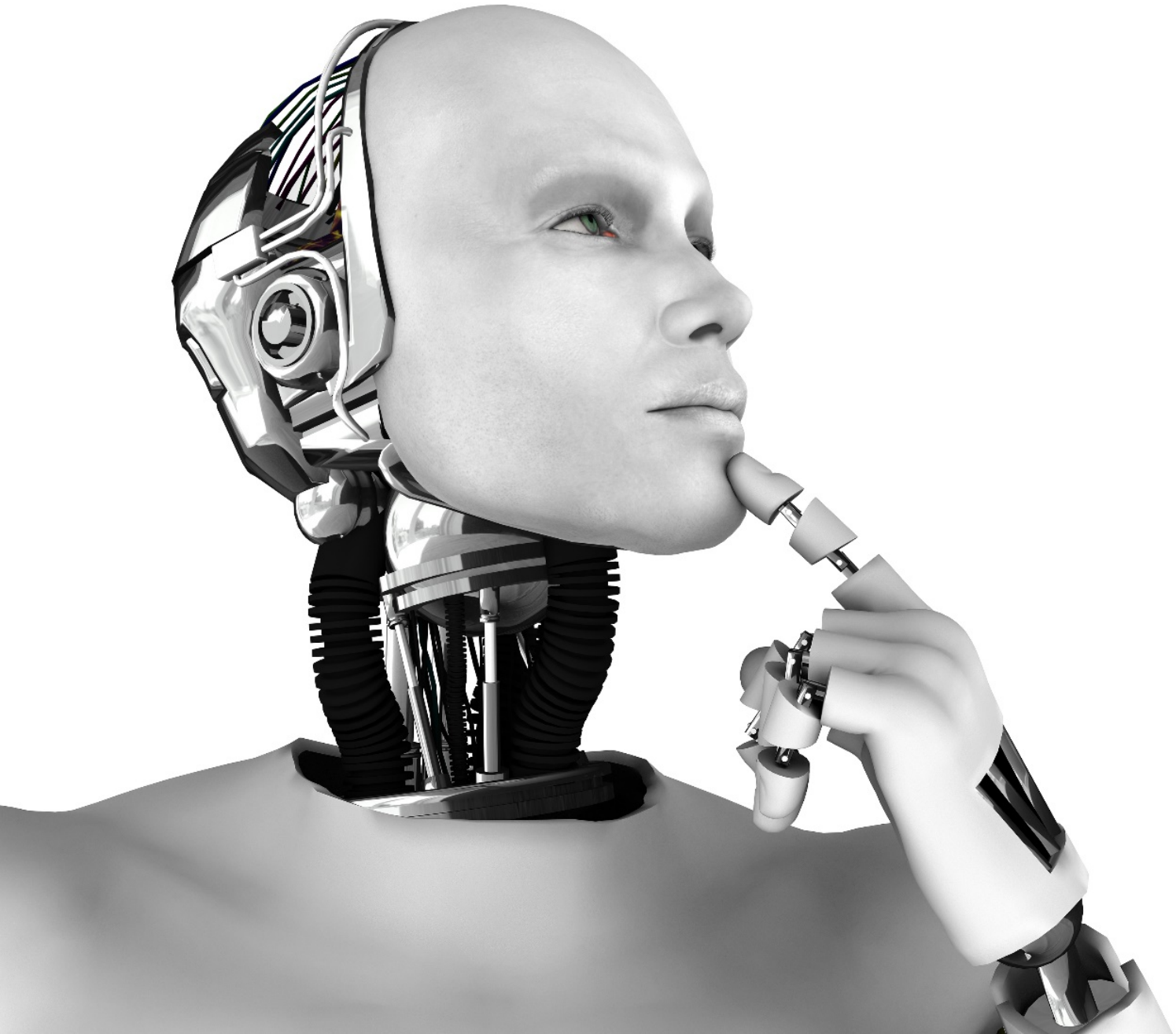
Learn about Data Science



A company’s success can be effected by Machine Learning

“Unicorn Data Scientists (upgraded from “sexy data scientists”) are hard to find and are paid more than $200,000 per year.”

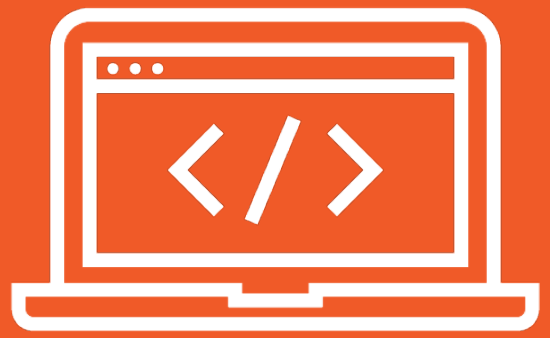
Gil Press. (2015). *Forbes*



Your next project?

Getting started with Python and Jupyter Notebook

Python



Easy to learn

Powerful, object-oriented

Elegant syntax, easy to read

Standard libraries for most common tasks

Python Versions

Python 2.7 and 3.x

- Both used
- Some incompatibilities

Python 3

- Future of Python
- Introduced in 2010

Python 2.7

- Last version of Python 2
- Static since 2012

Python 3.5 used in this course

Python Libraries For Machine Learning

numpy – scientific computing

pandas – data frames

matplotlib – 2D plotting

scikit-learn

- Algorithms
- Pre-processing
- Performance evaluation
- And more ...

Jupyter Notebook

Formerly IPython Notebook

Notebooks contain code and text

Perfect for iterable work like Machine Learning

Sharable

Supports multiple languages

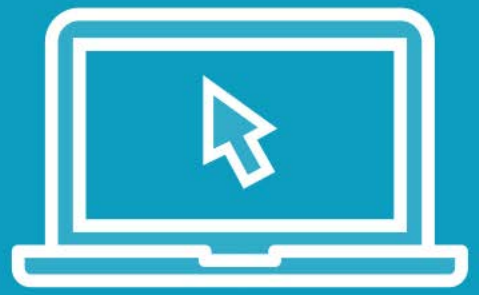
Installation

Anaconda Distribution

https://www.continuum.io/downloads

conda – package and environment manager

Demo



Jupyter Notebook

Python 3.5