Pattern Recognition

Fundamentals

Objectives

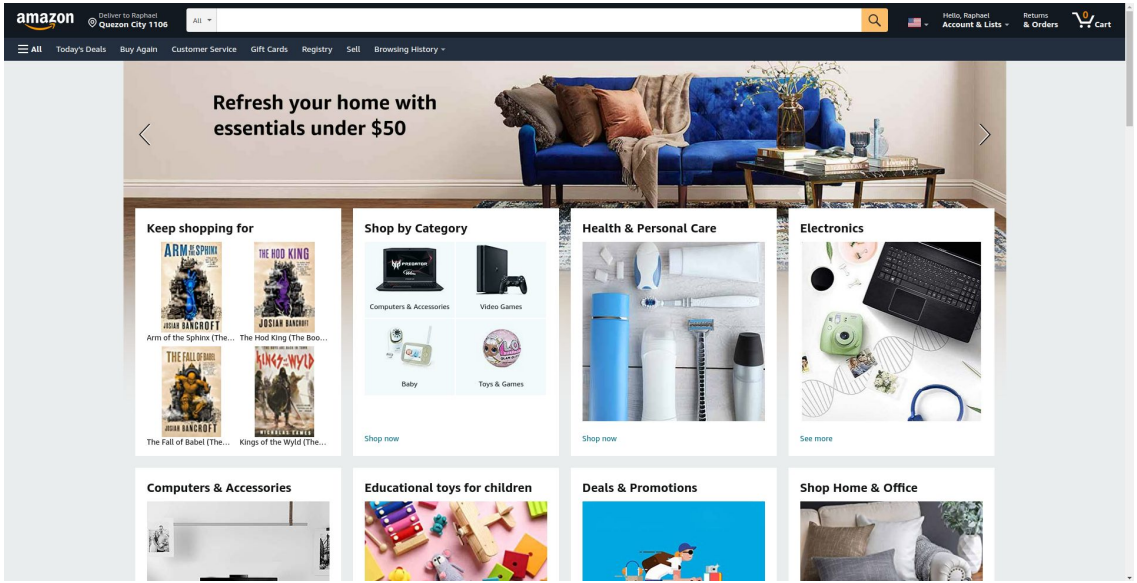
- Overview of Pattern Recognition
- Programming vs Machine Learning
- Definition of Terms
- Setup and Tools

Pattern Recognition

- Method of analysis that applies machine learning to recognize patterns / regularities in data
- **Data:** Any form of input that can be objectively defined that characterizes an entity
- **Pattern:** a set of objective rules that govern how data is treated

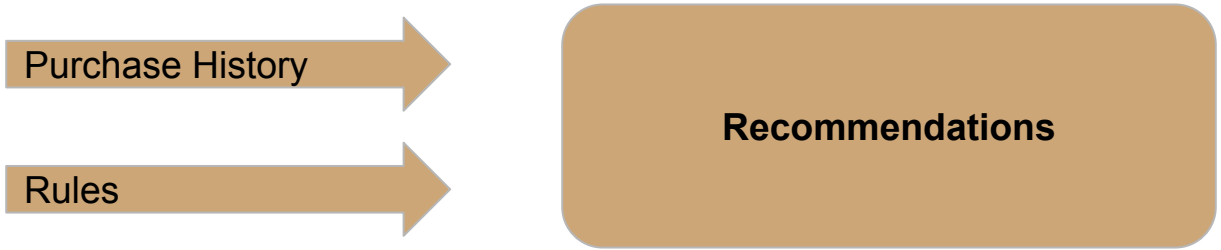
A Simple Case Study: Product Recommendations

- Your site needs to provide a set of recommended products based on past purchases



A Simple Case Study: Product Recommendations

- System:

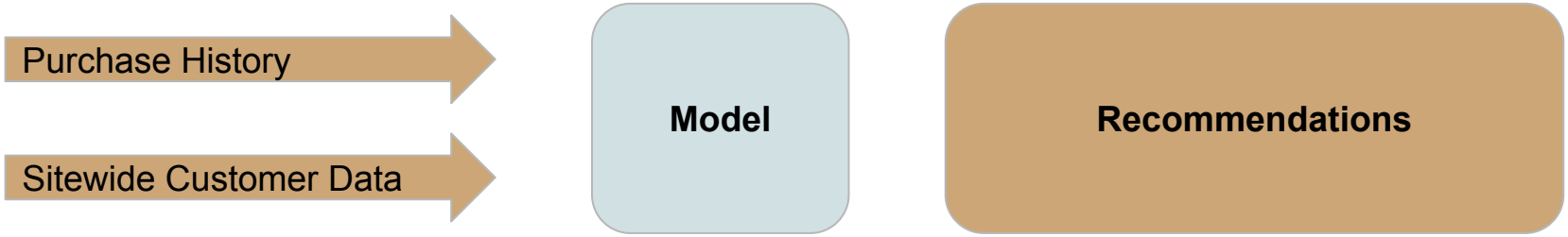


A Simple Case Study: Product Recommendations

- Classical Programming:
  - Requires the programmer to explicitly set rules
  - Customer recommendations can be unpredictable and variates from customer to customer
  - Each rule adds more time to the process

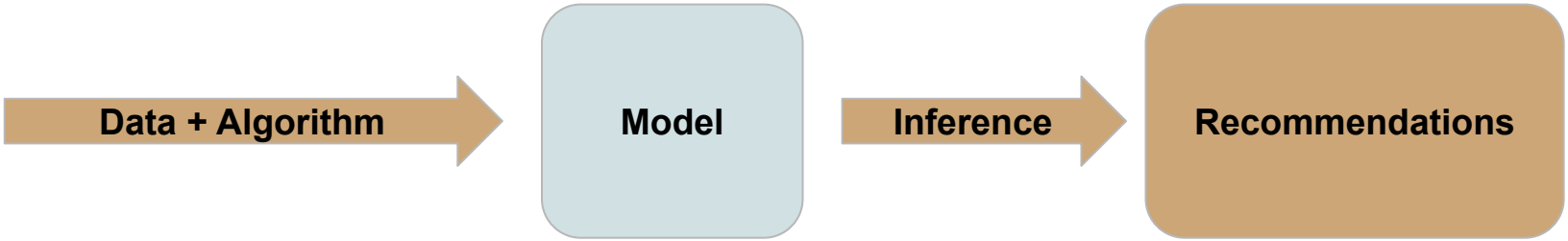
A Simple Case Study: Product Recommendations

- Machine Learning:
  - Derives rules from the data itself
  - Adapts to the complexities in customer profiles
  - Trains models ahead of time, saving time towards prediction
- The output of a Machine Learning algorithm is a model



What is a Model?

- A trained algorithm that:
  - Use to identify patterns / rules in your data
  - Does not require explicitly setting such rules (since it is already derived)
- Inference: Application of a model to perform a prediction



Simple Algorithm

$$f(x) = w_1x_1 + w_2x_2 \ldots + w_nx_n$$

Features and **Weights**

Simple Algorithm

$$f(x) = w_1x_1 + w_2x_2 \ldots + w_nx_n$$

Features: Objective characteristics of your data

x_1: Is this product a hat? Yes = 1

Simple Algorithm

$$f(x) = w_1x_1 + w_2x_2 \ldots + w_nx_n$$

Weights: How much does the feature affect the prediction?

w_1: Hats make up 80% of the purchases this customer has made

Simple Algorithm

$$f(x) = (0.8 * 1) + w_2x_2 \ldots + w_nx_n$$

Feature x_2: Is this item from Brand Z

Weight w_2: 2 / 8 items this person bought in the past were Brand Z (0.25)

Simple Algorithm

$$f(x) = (0.8 * 1) + (0.25 * 1)$$
$$f(x) = 1.05$$

Output: If > 1, then recommend the product

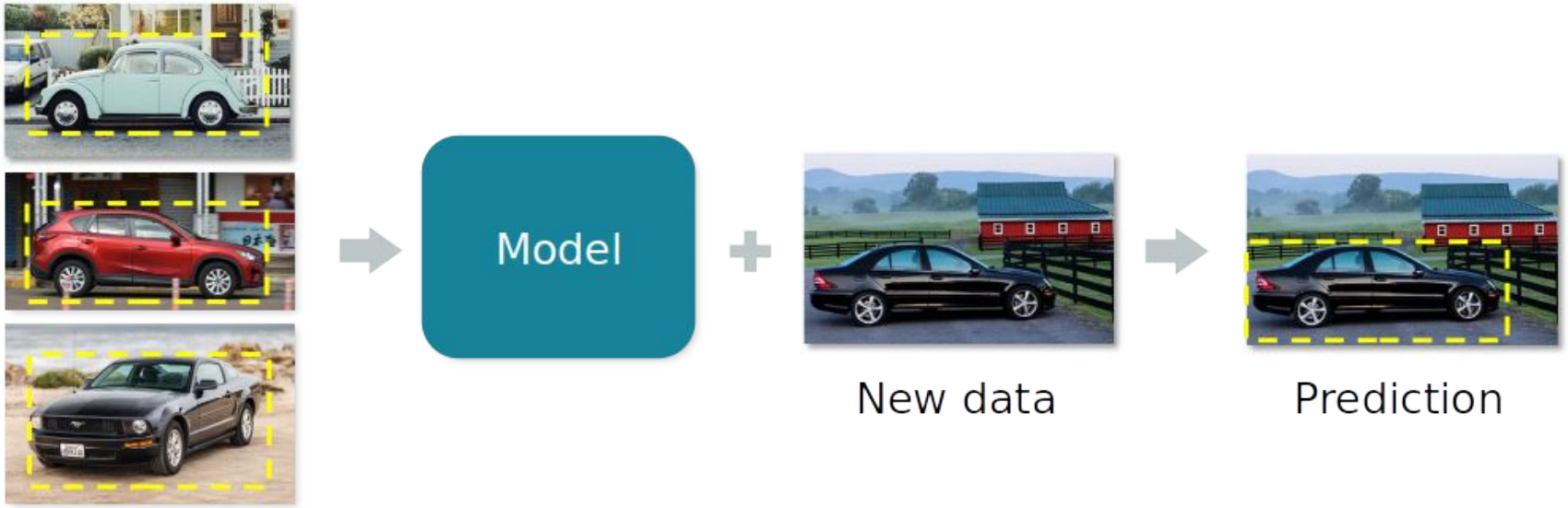
Common Terminologies

- **Feature Vector:** An array of numbers representing an entity
- **Dimensionality:** The size of the feature vector
- **Algorithm:** The algorithm (machine learning algorithm) used to create a model
- **Model:** The output of the algorithm representing a solution
- **Weights:** Numerical coefficient of dimensionality to come up with a solution (read model)

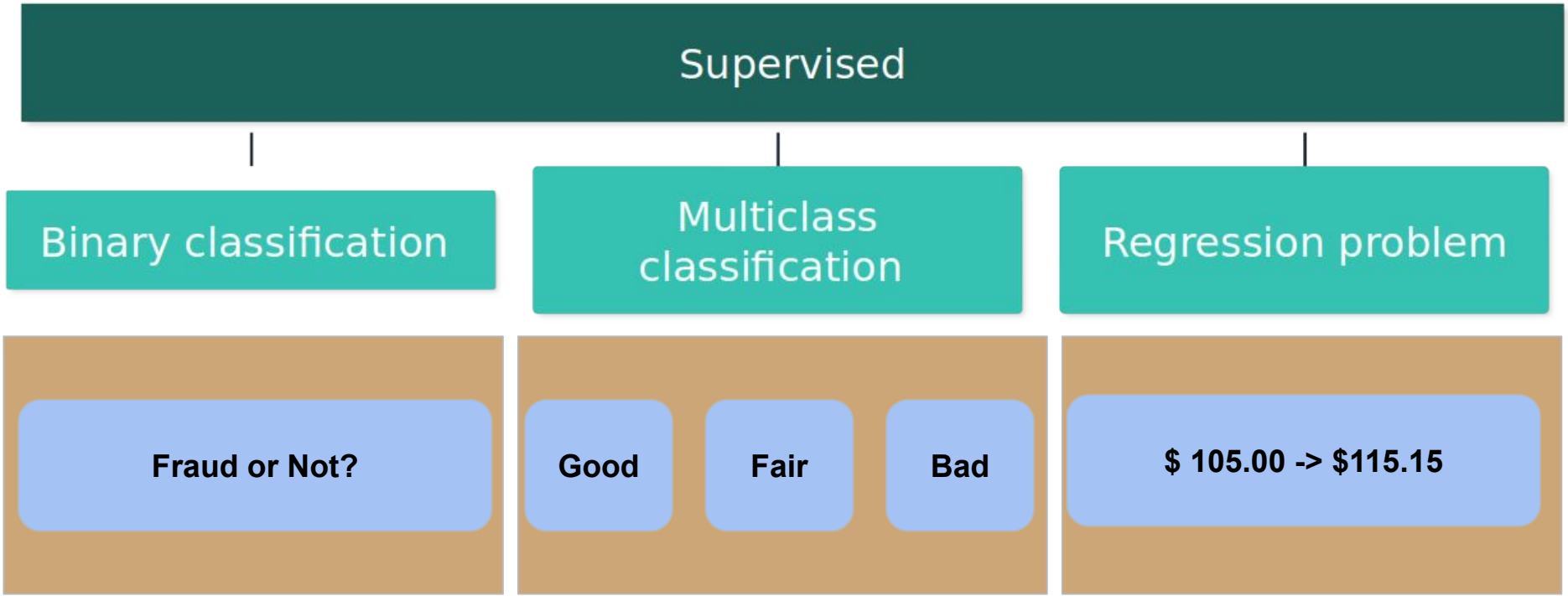
Categories of Machine Learning Algorithms

Supervised

- Learn by identifying patterns in data that’s already been labeled



Supervised


Supervised
Binary classification
Multiclass classification
Regression problem
Fraud or Not?
Good
Fair
Bad
$ 105.00 -> $115.15

Mathematical Representation

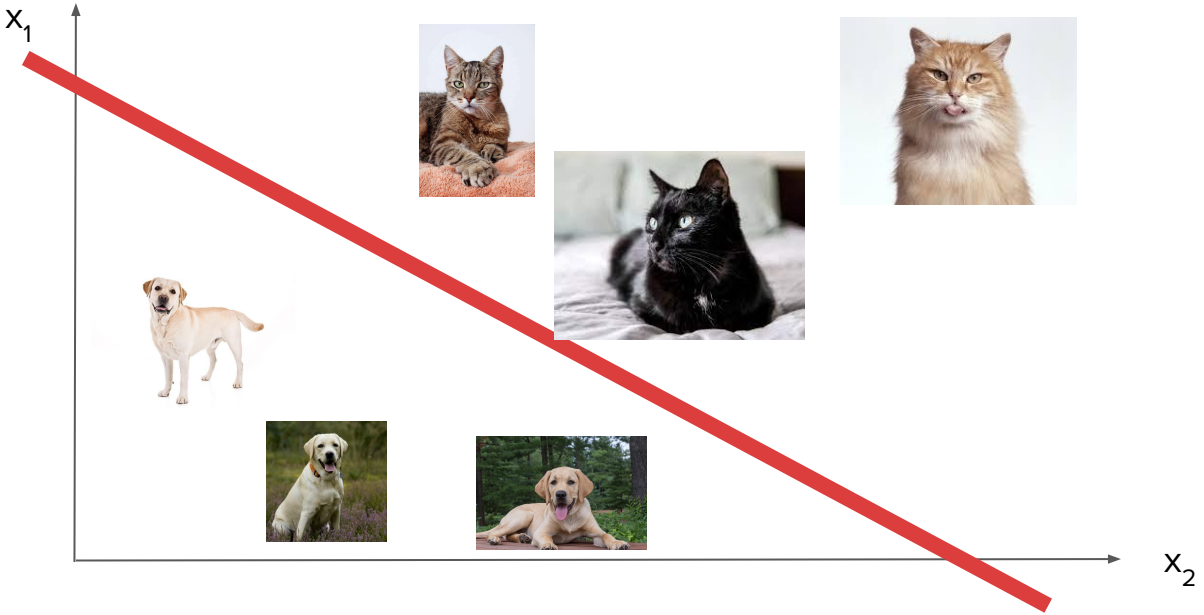
$$f(x) = mx + b$$

$$f(x) = y$$

$$y = \{ 1, 0 \}$$

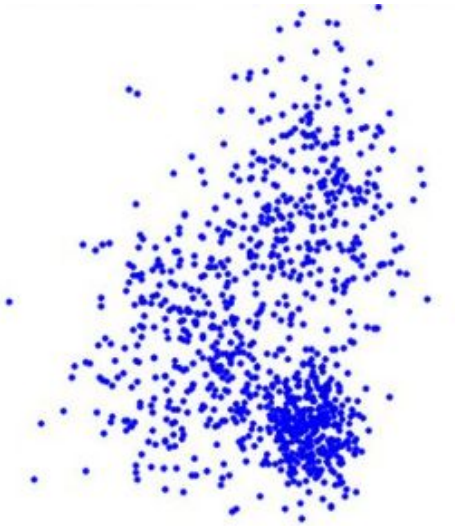
$$f(x) = w_1x_1 + w_2x_2 + w_3x_3 \dots w_nx_n$$

Mathematical Representation

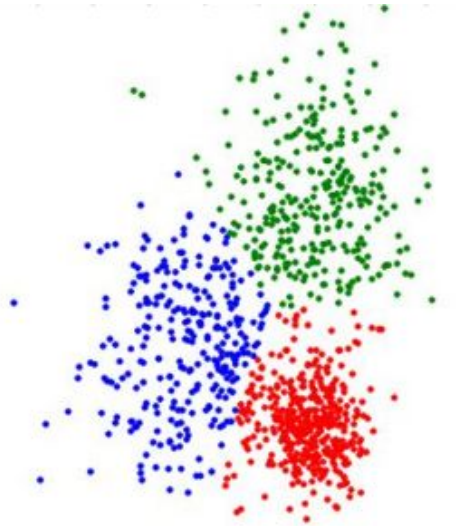


Unsupervised

- The machine has to uncover and **create the labels** itself

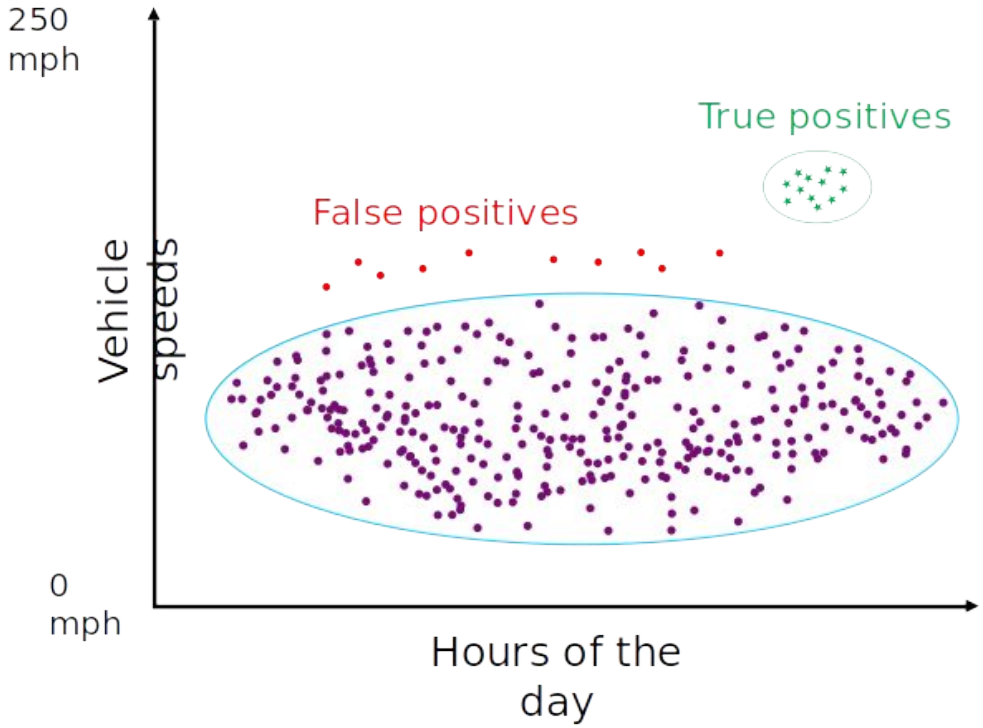


Original data



Clustered data

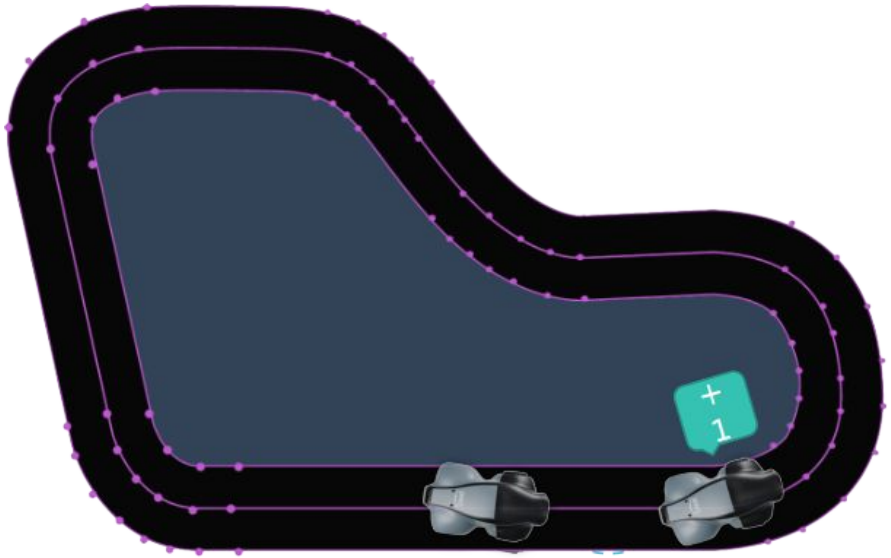
Unsupervised



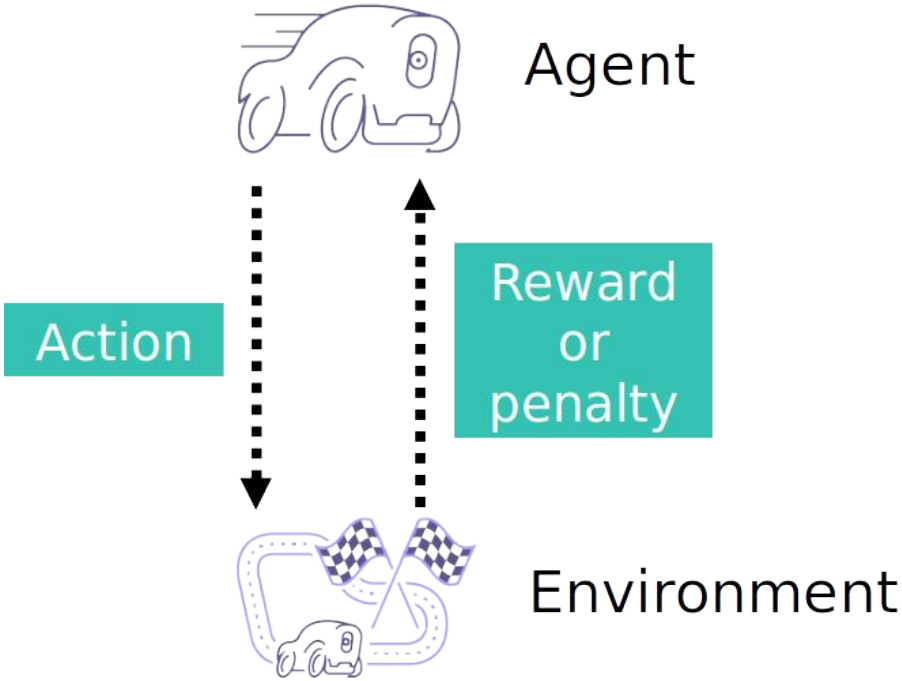
Are these anomalies simple outliers or indicators of hardware failure?

Reinforcement Learning

- Learning through trial and error. Best when the desired outcome is known but the exact path to achieving it isn’t.

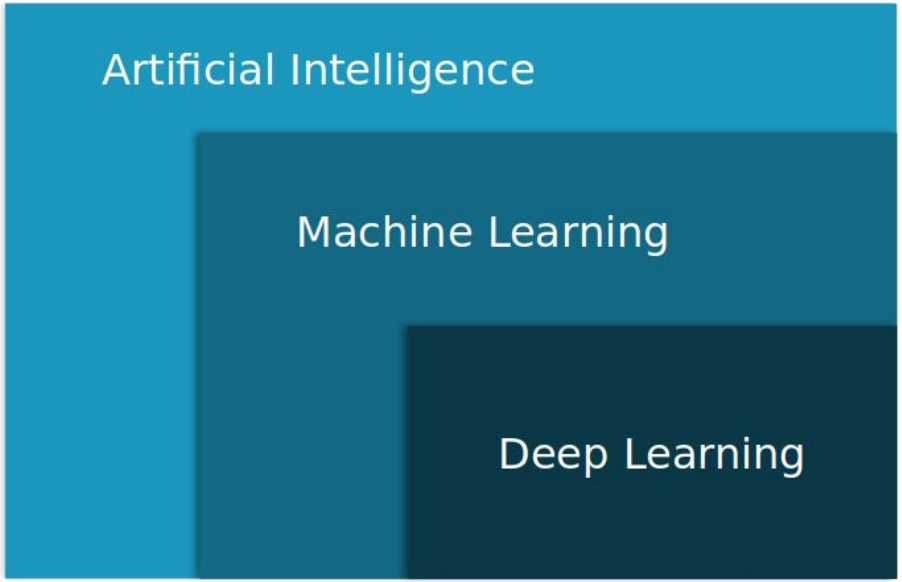


Reinforcement Learning

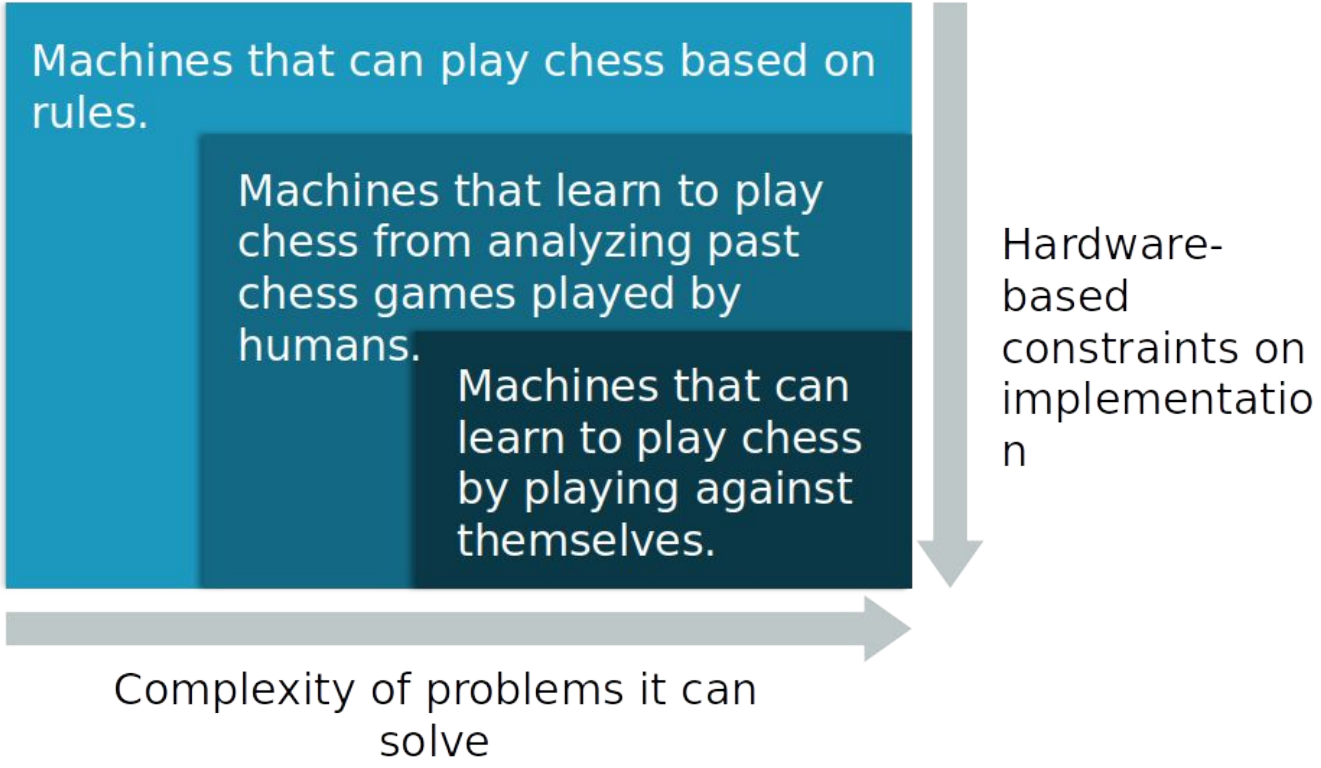


Deep Learning

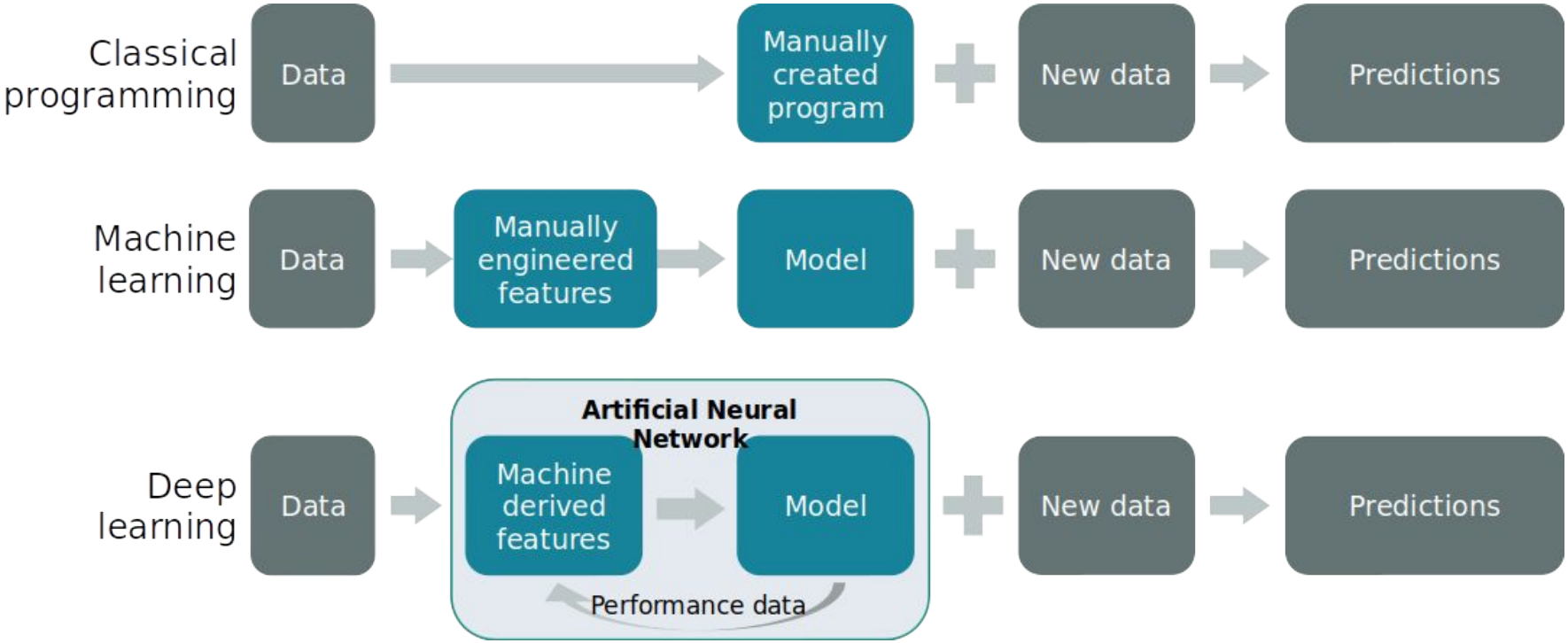
- Sub-category of machine learning



Deep Learning

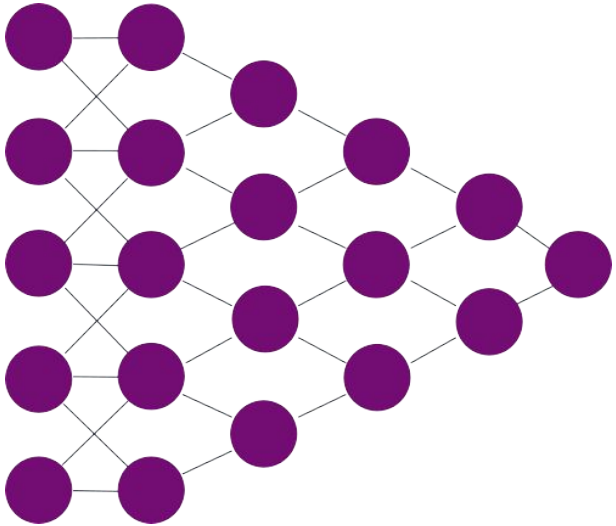


Deep Learning

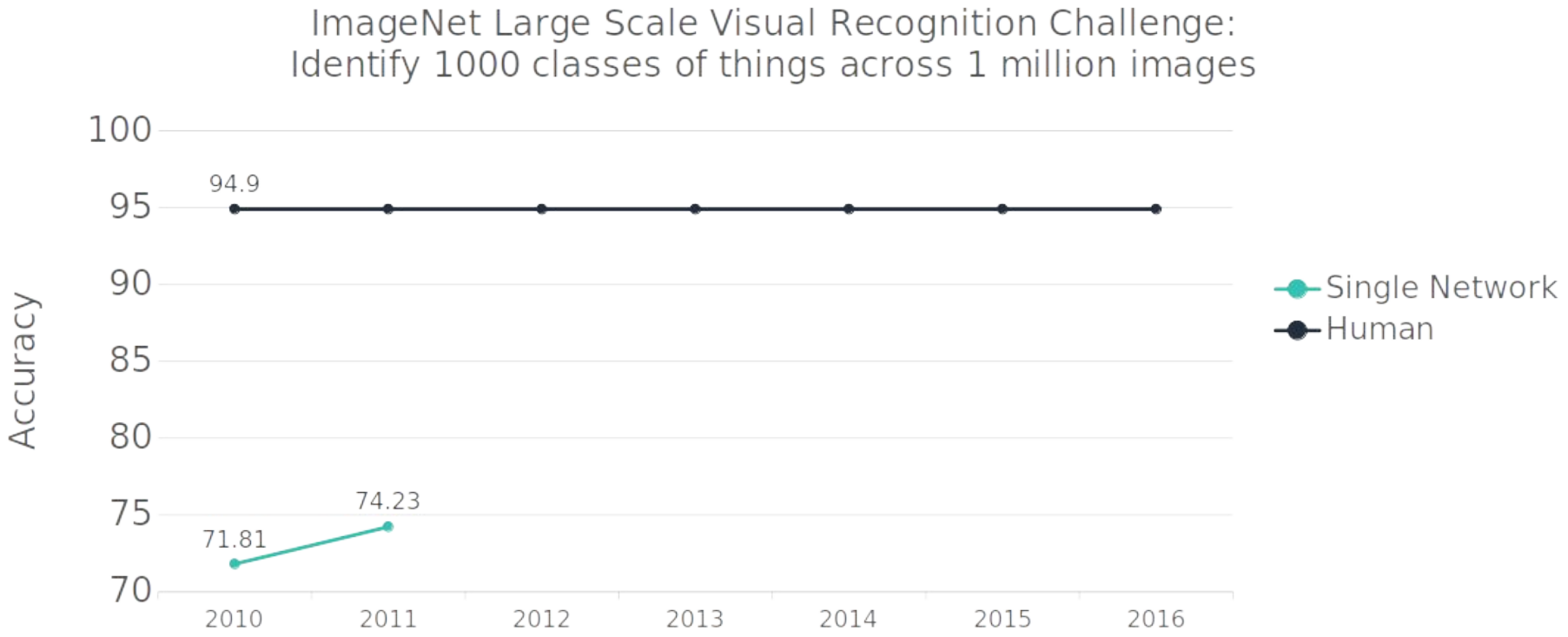


Deep Learning

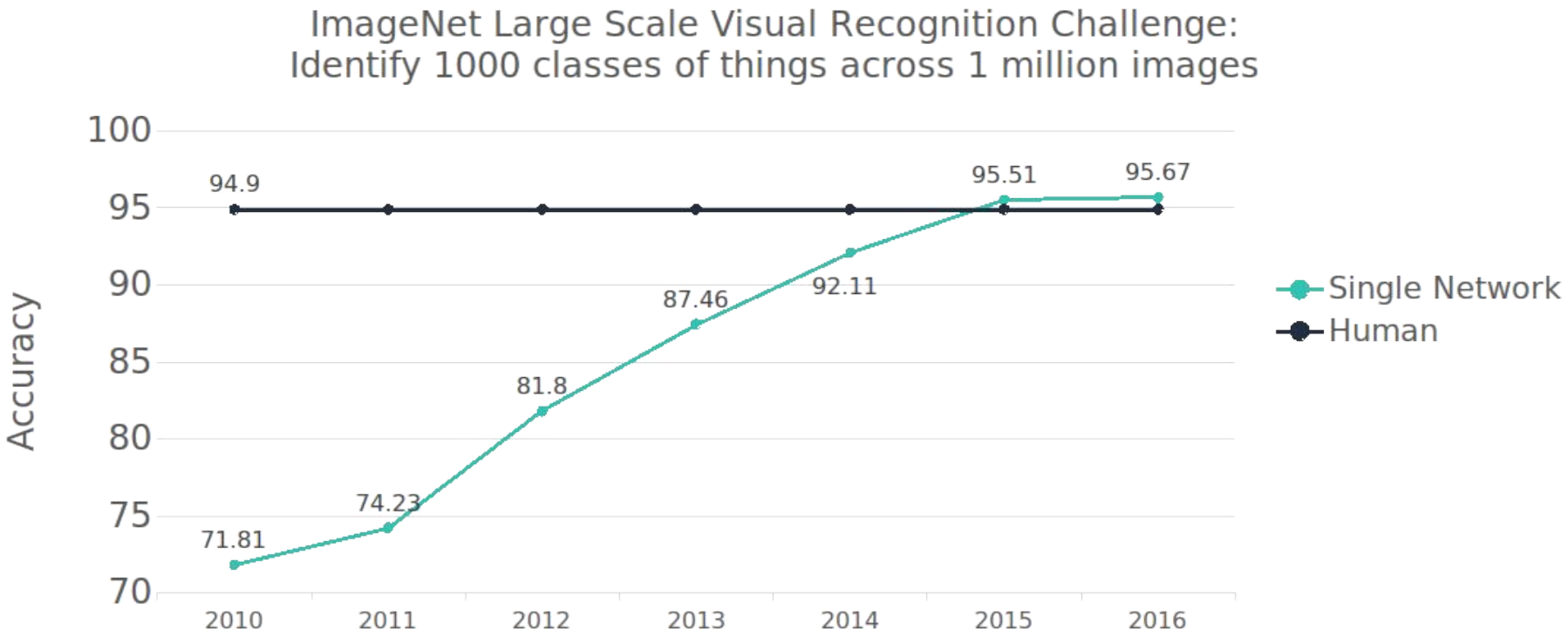
- Artificial Neural Network
  - Model that can be trained on raw data
  - Feature extraction is automatically performed



Deep Learning

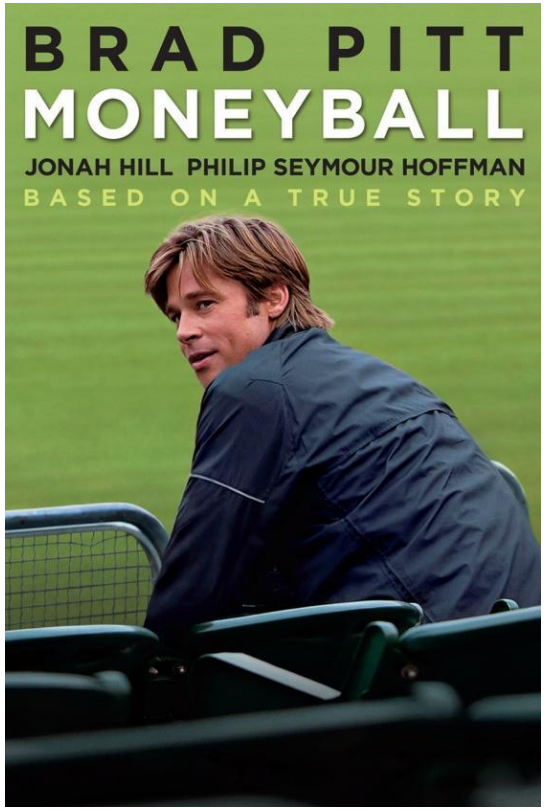


Deep Learning



Demo

Inspiring Shows



Wizards of the modern Day

Two Primary Objectives

Represent anything as a multi-dimensional vector

Develop a Model based on data

Tools of the Trade

Setting up Python and Jupyter

1. Download Python for your Operating System (https://www.python.org/downloads/)
2. Use **pip** to install jupyter
3. Run Jupyter notebook

Checking your Python Version

```
python --version
```

Creating an Environment (Windows)

```
py -m venv env
```

Creating an Environment (Mac / Linux)

```
python -m venv env
```

Allowing Execution (Windows)

```
Set-ExecutionPolicy Unrestricted -Scope Process
```

Activating an Environment (Windows)

```
env\Scripts\activate.ps1
```

Activating an Environment (Linux / Mac)

```
source env/bin/activate
```

Installing Jupyter

```
pip install jupyter
```

Running Jupyter

```
jupyter notebook .
```