# BITS Pilani presentation


**BITS** Pilani
Pilani Campus

Tanmay Tulsidas Verlekar
CSIS


BITS Pilani
Pilani Campus

# Applied Machine Learning SE ZG568 / SS ZG568

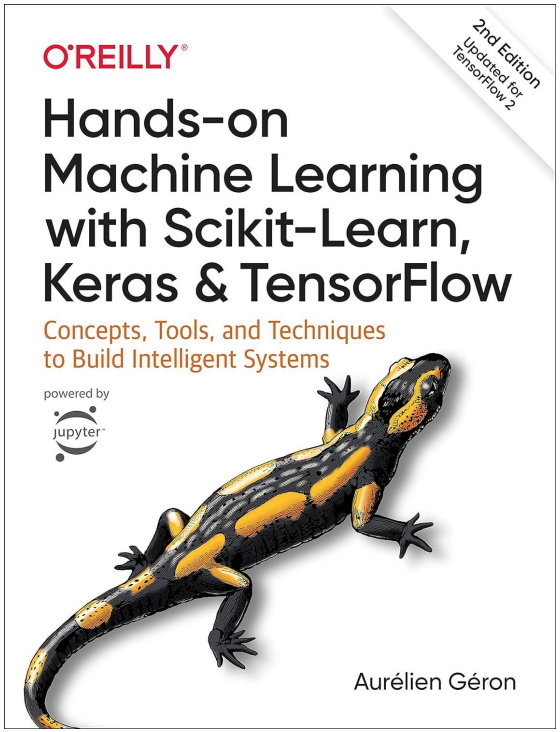
# Lecture No.1

# Course content

- The Fundamentals of Machine Learning
  Covers the machine learning landscape, what an end-to-end machine learning project looks like, and a variety of algorithms including classification methods, trees, and clustering

- Neural Networks and Deep Learning
  Introducing then training artificial neural networks, and then covering multiple types including CNNs, RNNs, GANs, and reinforcement learning

# Textbook




Codes: https://github.com/ageron/handson-ml2

# Evaluation Scheme

Legend: EC = Evaluation Component; AN = After Noon Session; FN = Fore Noon Session

| No | Name | Type | Duration | Weight | Day, Date, Session, Time |
|---|---|---|---|---|---|
| EC-1 | Quiz-I | Online | - | 5% | September 1-10, 2024 |
| | Programming Assignment I | Offline | - | 12% | October 10-20, 2024 |
| | Programming Assignment II | Offline | - | 13% | November 1-10, 2024 |
| EC-2 | Mid-Semester Test | Closed Book | 1.5 hours | 30% | Saturday, 21/09/2024 (AN) |
| EC-3 | Comprehensive Exam | Open Book | 2.5 hours | 40% | Saturday, 30/11/2024 (AN) |

# What Is Machine Learning?

Here is a slightly more general definition:

> [Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.
>
> —Arthur Samuel, *1959*

And a more engineering-oriented one:

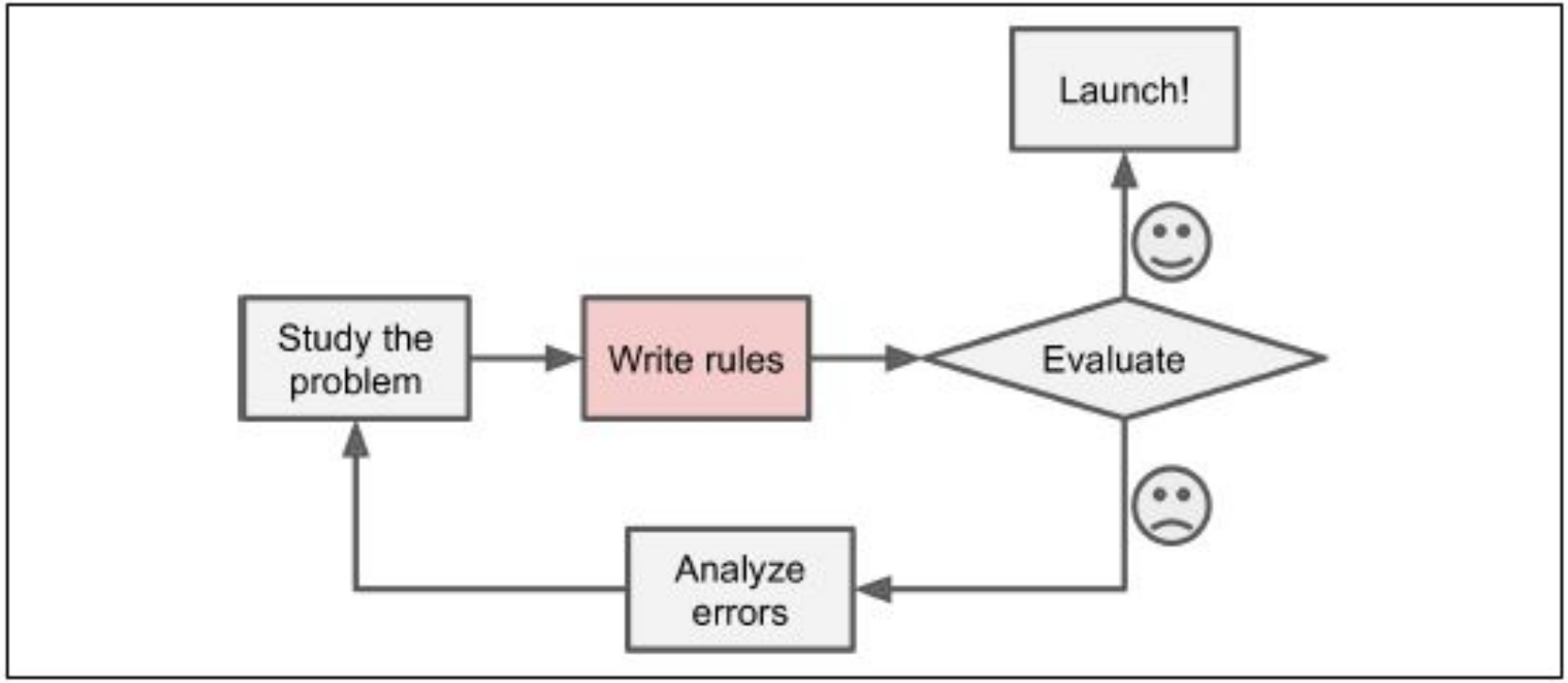
> A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.
>
> —Tom Mitchell, *1997*

# Face recognition

Owner of the phone

Friends

# The traditional approach



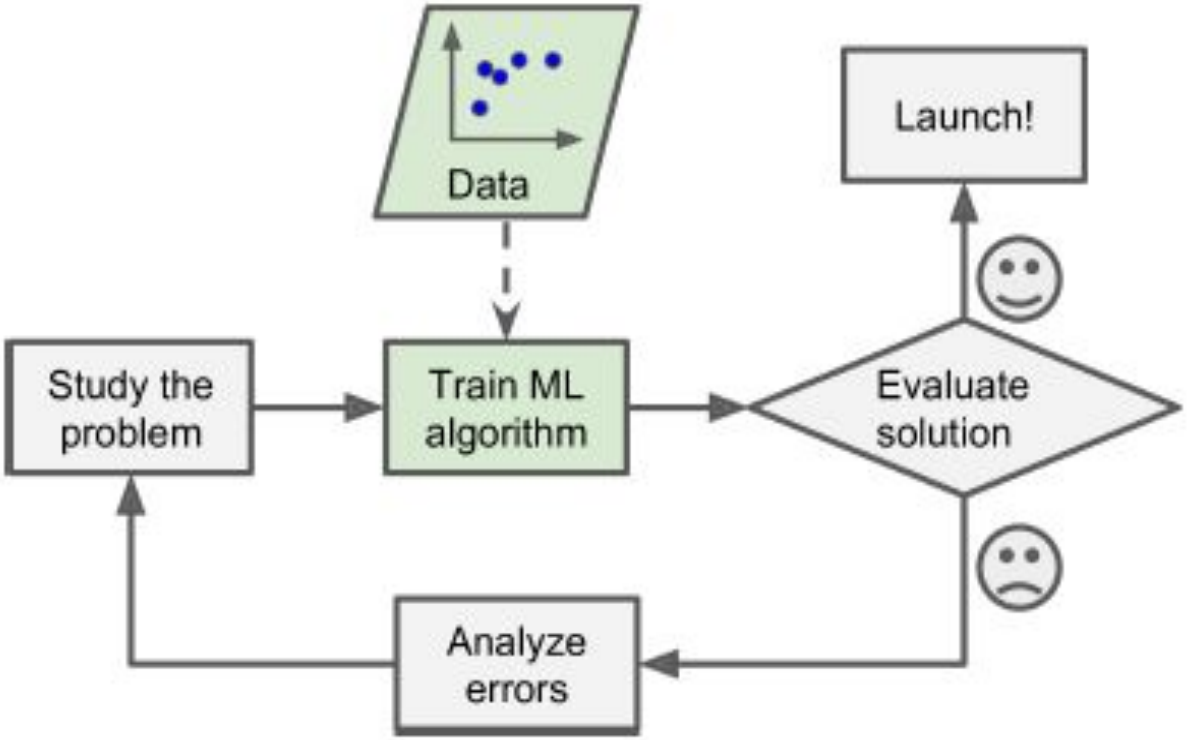
# Accuracy

| | | Predicted Class | |
|---|---|---|---|
| | | No | Yes |
| Observed Class | No | TN | FP |
| | Yes | FN | TP |

| | |
|---|---|
| TN | True Negative |
| FP | False Positive |
| FN | False Negative |
| TP | True Positive |

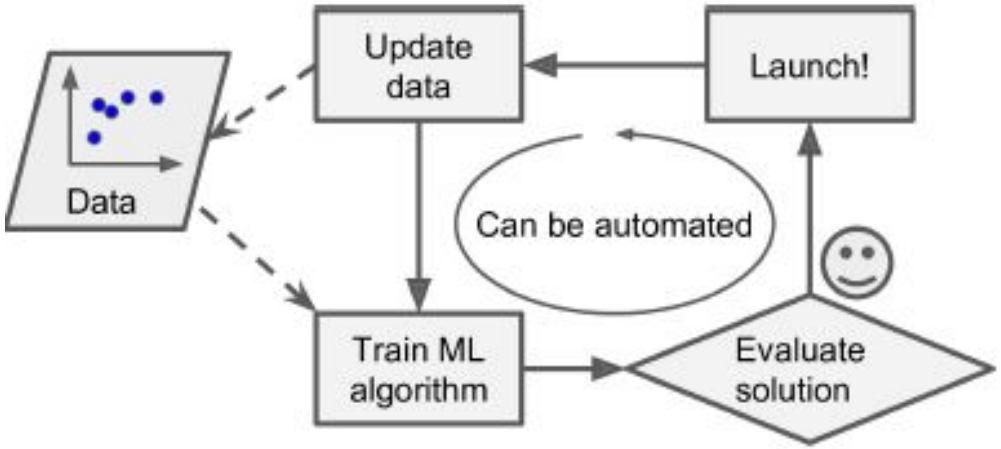
$$\text{Accuracy} = (TN+TP)/(TN+FP+FN+TP)$$

# Machine Learning approach



# Why Use Machine Learning?

- Training an algorithm with data can involve a shorter, easier to maintain program, and also be more accurate
- Can tackle super complex problems
- Easier to automate retraining for constantly changing environments
- May offer insights about the problems ML solves



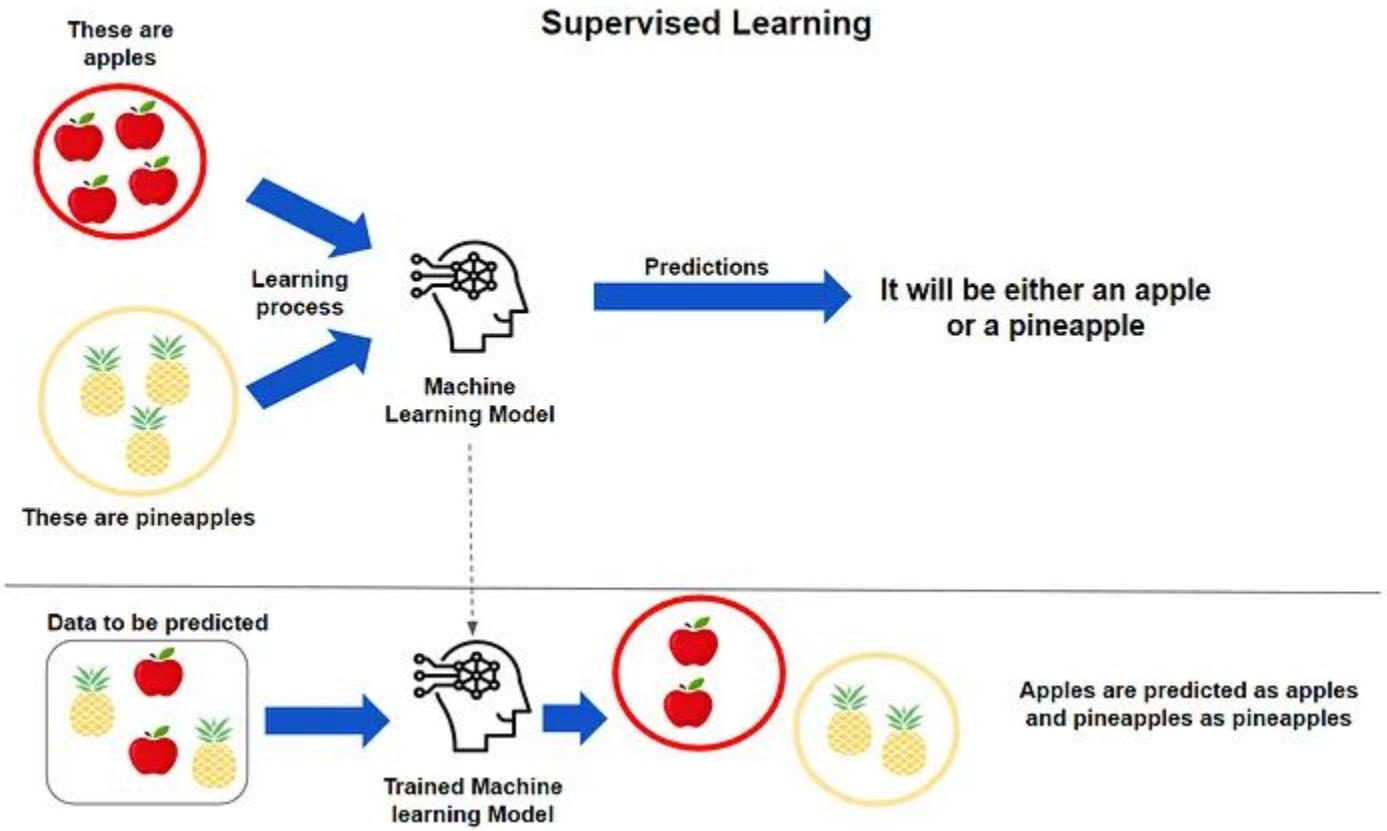

One-Attack 5

# Types of Machine Learning Systems

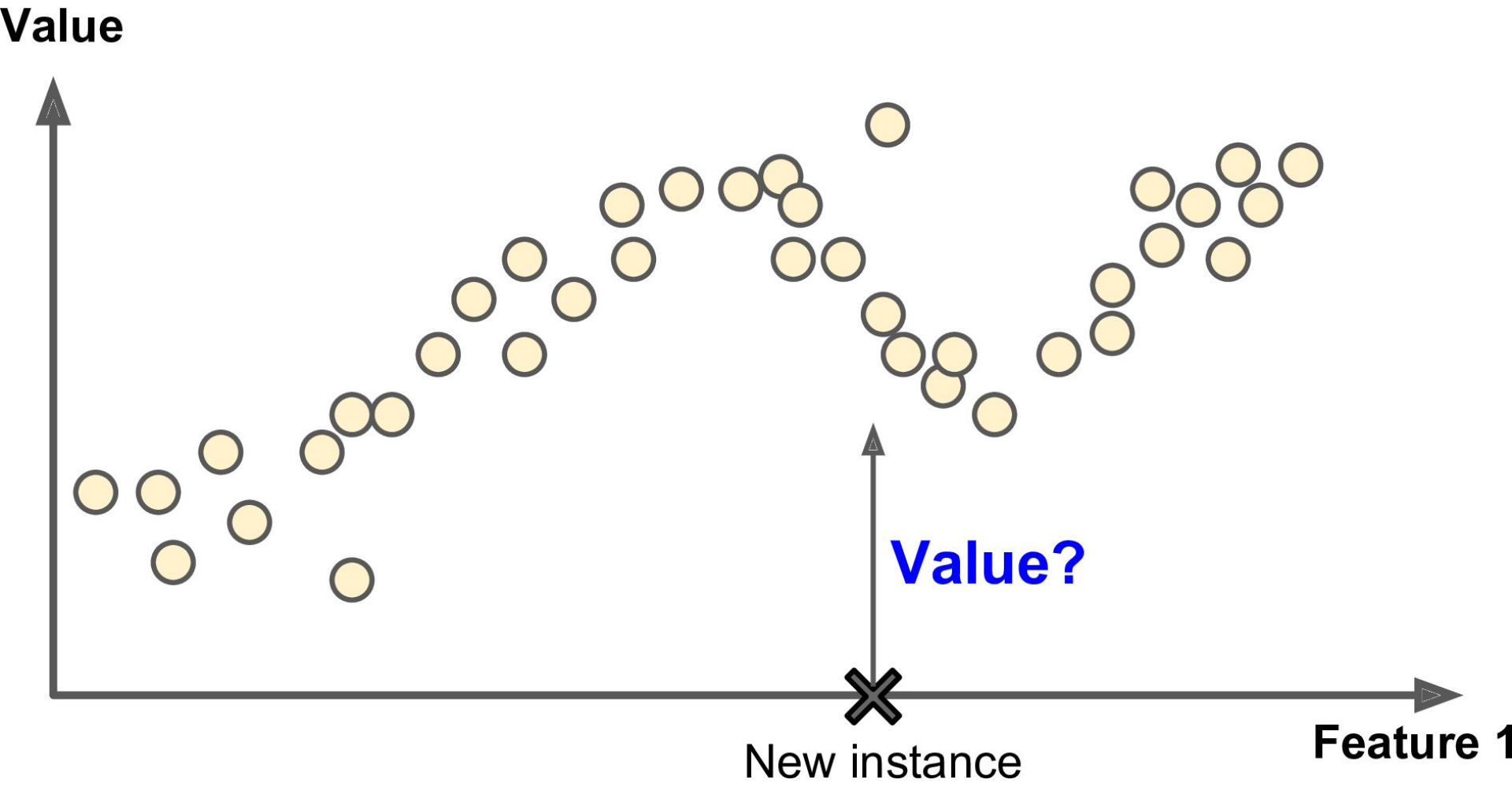


- Supervised learning - labeled data, e.g. classification and regression.
- Unsupervised learning - unlabeled pattern recognition, such as clustering or some anomaly detection
- Semi-supervised learning - solving supervised learning problems without labeling all of the data
- Reinforcement learning - training an agent to choose actions to maximize its numeric reward metric
- Batch and Online Learning
- Instance-Based Versus Model-Based Learning

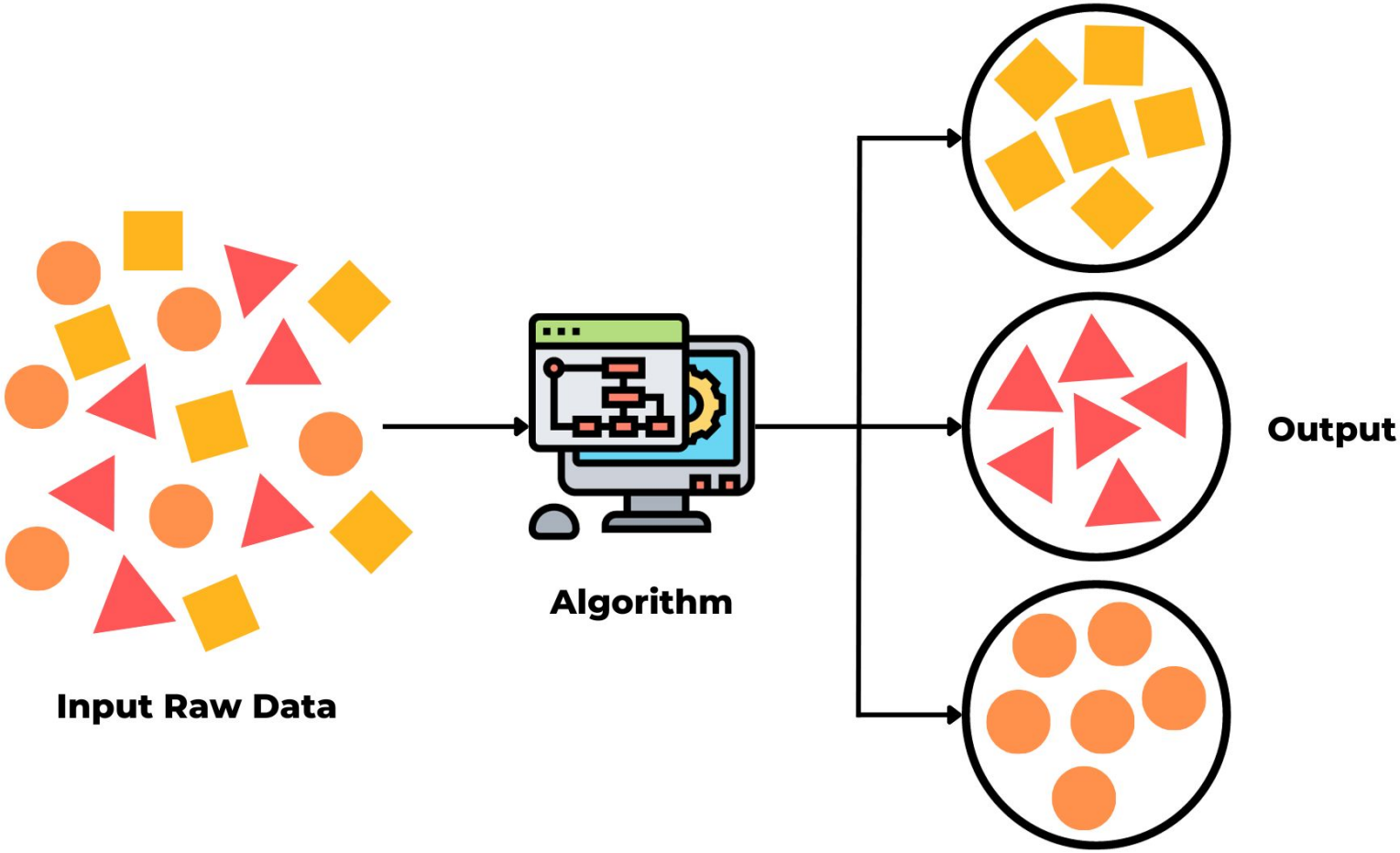
# Classification



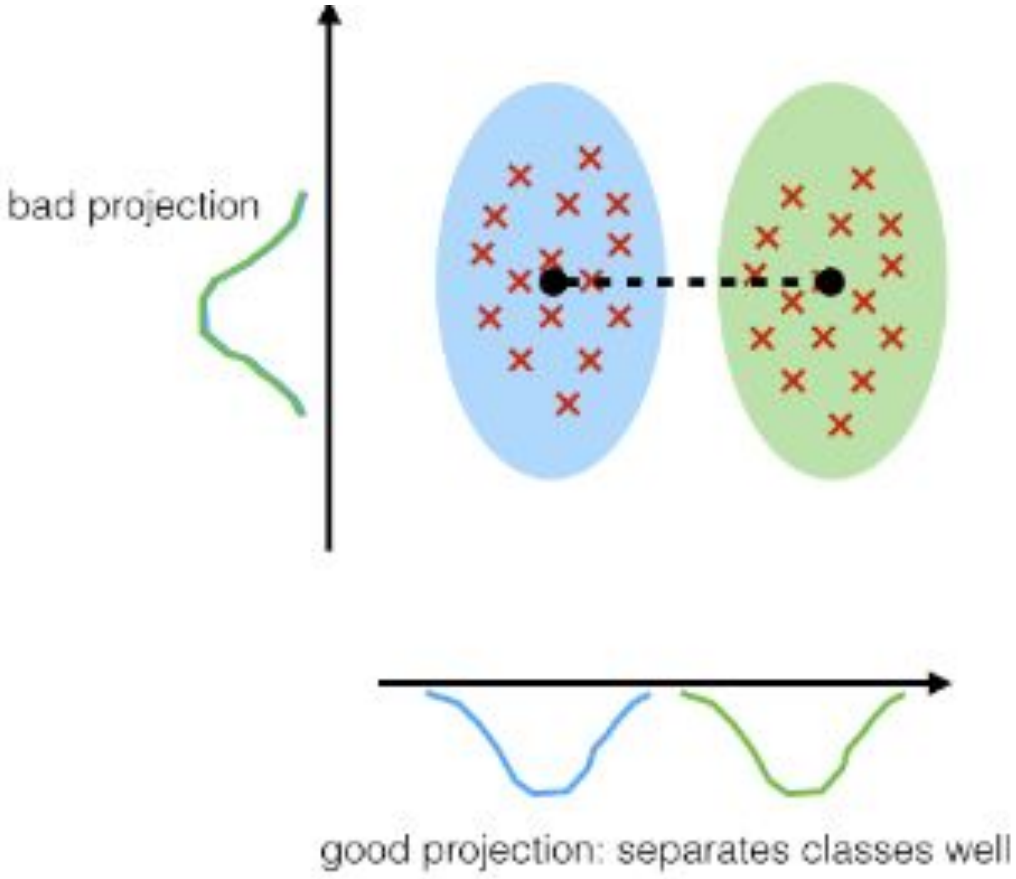
# Regression



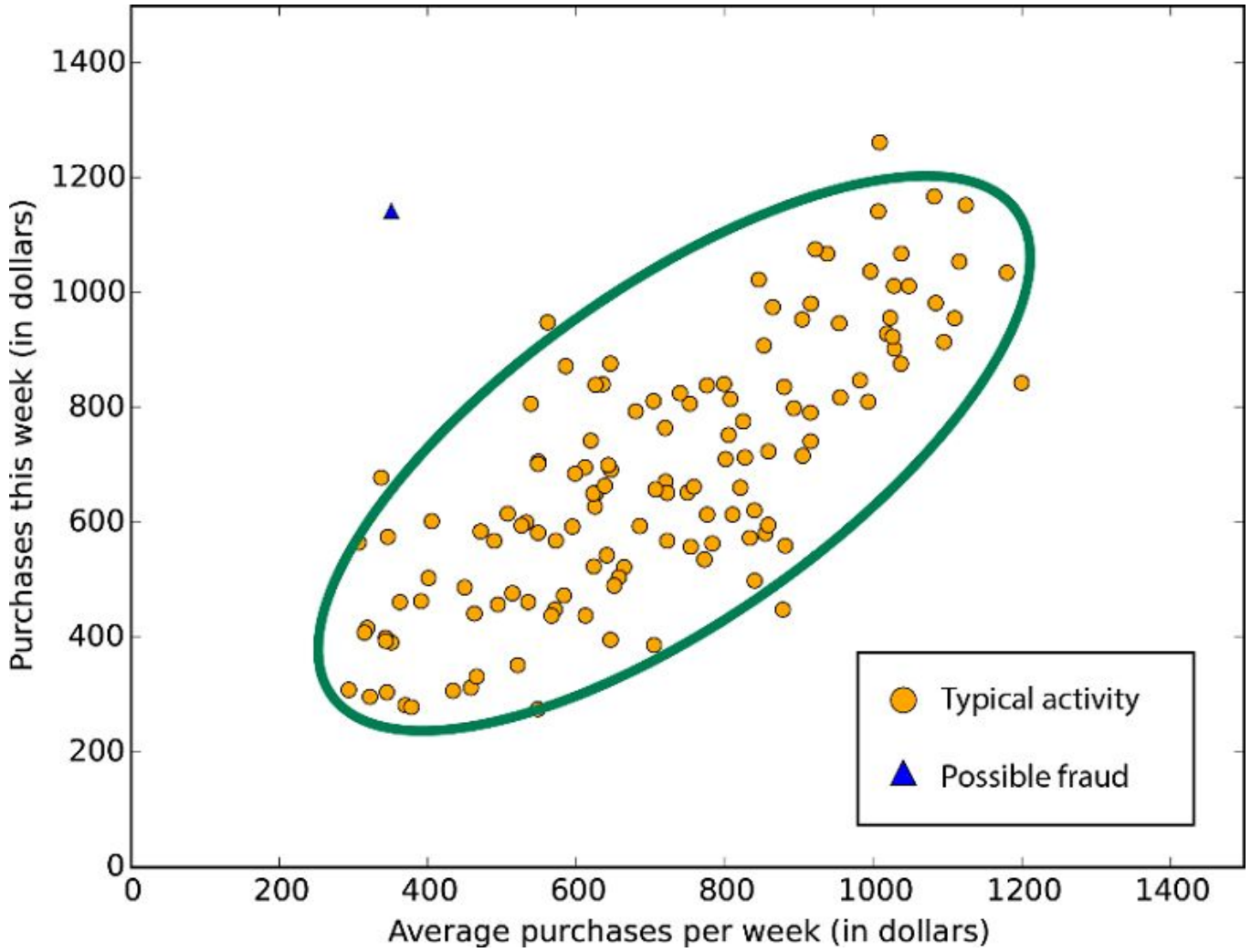
# Clustering



# Dimensionality reduction



# Anomaly detection

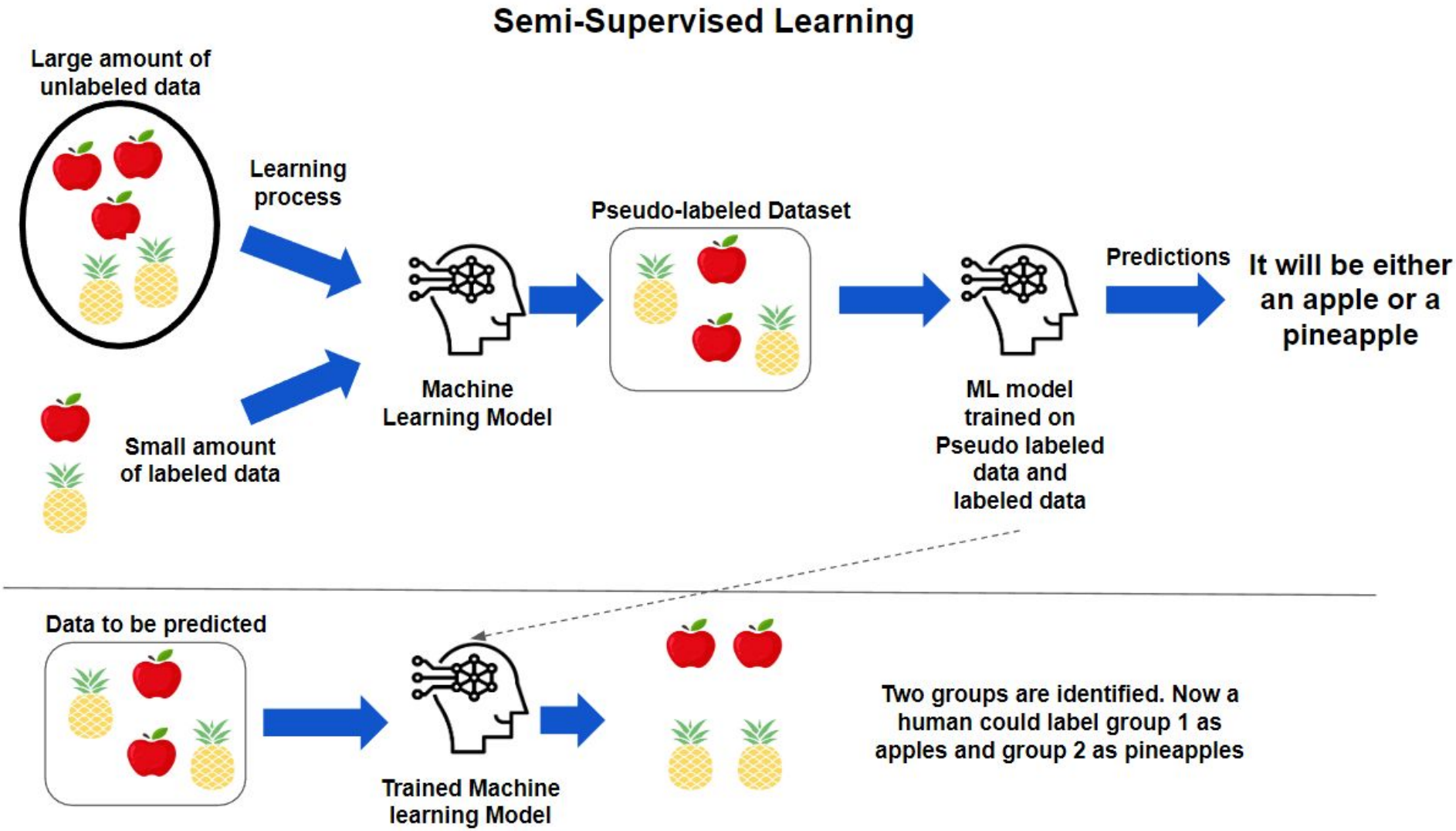


# Association rule learning

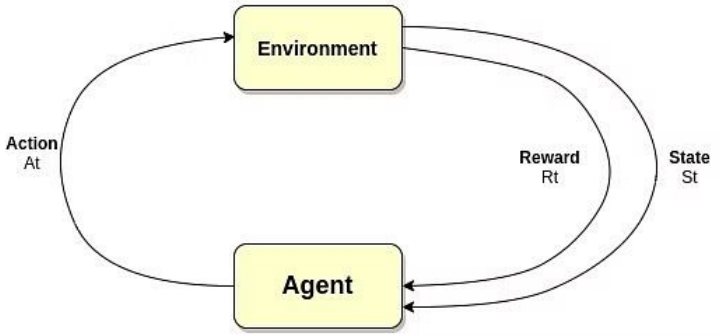
| ID | Items bought |
|---|---|
| 100 | A, B, E |
| 200 | B, D |
| 300 | B, C |
| 400 | A, B, D |
| 500 | A, C |
| 600 | A, B, C |

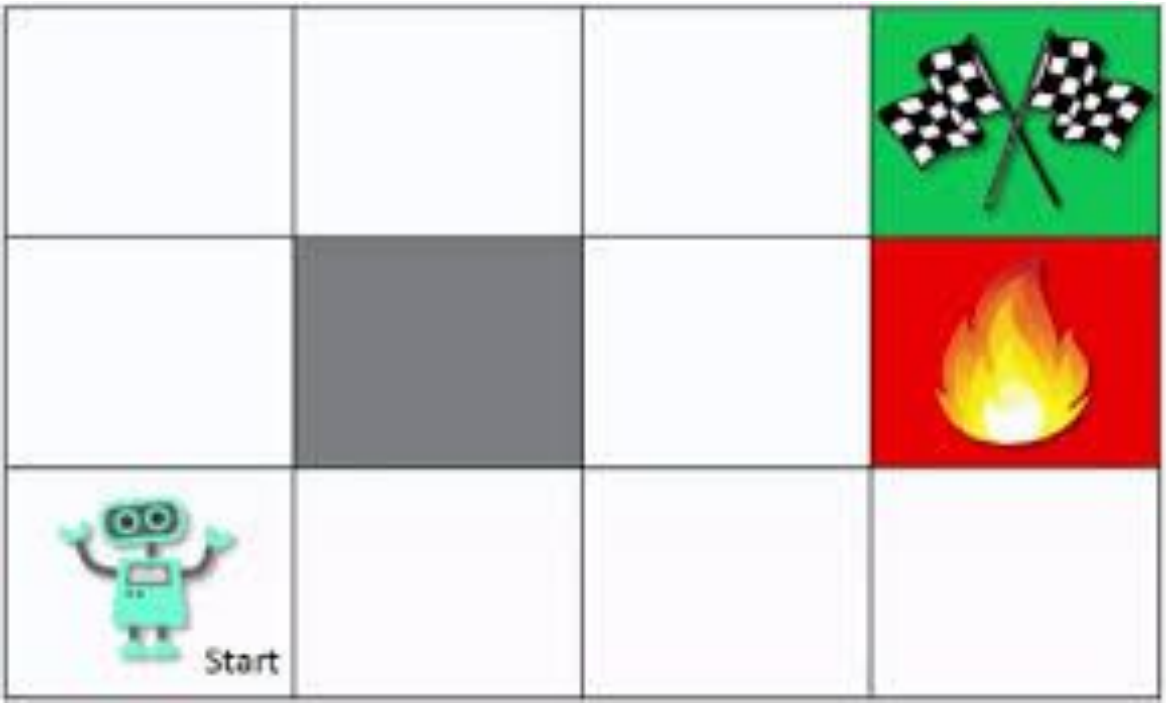
| Itemlist |
|---|
| A, B |
| A, C |
| B, C |
| B, D |

# Semi-supervised learning



# Reinforcement learning



# Main Challenges of Machine Learning



- Insufficient Quantity of Training Data
- Nonrepresentative Training Data
- Poor-Quality Data - cleaning data is important
- Irrelevant Features
- Overfitting the Training Data
- Underfitting the Training Data