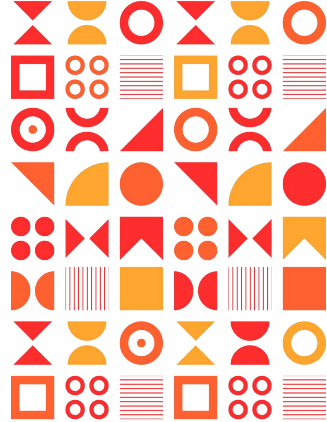


Introduction to machine learning?

semicolon





What is Machine Learning?

semicolon



What is Machine Learning?

Machine Learning is the science (and art) of programming computers so they can learn from data.

What is Machine Learning?

Machine Learning is the science (and art) of programming computers so they can learn from data.

Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

What is Machine Learning?

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**, improved with experience **E**.

Tom Mitchell, 1998

What is Machine Learning?

Elements:

- Task T
- Experience E
- Performance measure P

Detect the dialog act of an utterance

T: Classify the dialog act of an utterance

E: A set of utterances labeled with their dialog acts

P: The fraction of utterances correctly classified

What is Machine Learning?

Elements:

- Task T
- Experience E
- Performance measure P

Filter spam emails

T: *Classifying emails as spam or not spam*

E: *A set of labelled emails as spam or not spam*

P: *The number (or fraction) of emails correctly classified as spam/not spam*

What is Machine Learning?

Elements:

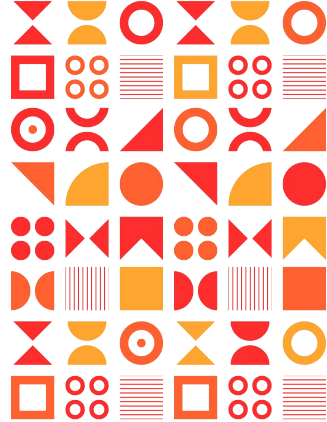
- Task T
- Experience E
- Performance measure P

Self-driving cars

T: Drive on public highways using vision sensors

E: Sequence of images and steering commands from human drivers

P: Average distance traveled before an error



**What part of the elements seem
to matter the most?**

semicolon





**What part of the elements seem to
matter the most?**

Answer: The Experience E

semicolon



Experience E

1. Supervised learning
2. Unsupervised learning

Others:

Reinforcement learning, etc

Experience E

1. Supervised learning
2. Unsupervised learning

Others:
Reinforcement learning,
etc

*Train a model using
labelled data*

*Example: image
classification, dialog
act classification.*

Experience E

1. Supervised learning
2. Unsupervised learning

Others:
Reinforcement learning,
etc

*Train a model using
unlabelled data*

*Example: market
segmentation, Social
network analysis.*

Experience E

1. Supervised learning
2. Unsupervised learning

Others:

Reinforcement learning,
etc

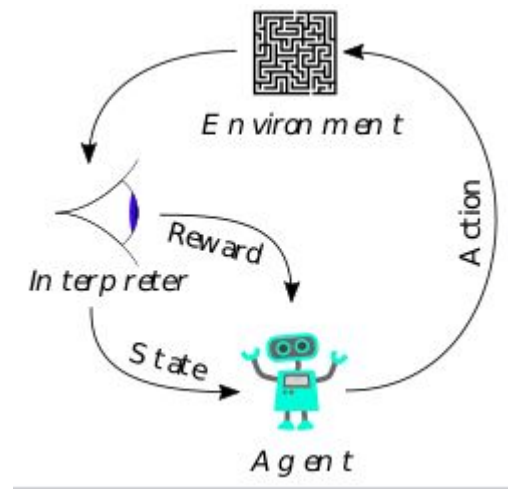


Image source: Wikipedia

*Example: game of
ping pong, chess*

Supervised learning

In supervised learning, the “right answers” are given.

That is, every data in the dataset has a class label it belongs to.

semicolon _____

Supervised learning

However, in supervised learning, a problem could take either of the two forms:

- Regression
- Classification

Supervised learning

In supervised learning, there's also need to define features. Features are the characteristics of the instances that the model uses for predictions

Features for whether an image is a dog or a cat

- *Size*
- *Color*

Features for whether a breast cancer image is benign or malignant

- *Age*
- *Tumor size*

Supervised learning

In supervised learning, there's also need to define features. Features are the characteristics of the instances that the model uses for predictions

Features for house price prediction:

- *Overall condition of the house*
- *Neighborhood*
- *Condition of the basement*
- *Number of bedrooms*
- *Construction date*
- *Number of schools in within 2 km*
- *Condition of the kitchen*

Regression problem - Supervised learning

In regression problems, predicted valued output are continuous.

An example is a problem of predicting housing prices based on land size, location, etc

***ML algorithms include:
Linear regression, etc***

Regression problem - Supervised learning

In regression problems, predicted values and output are continuous.

An example is a problem of predicting housing prices based on land size, location, etc

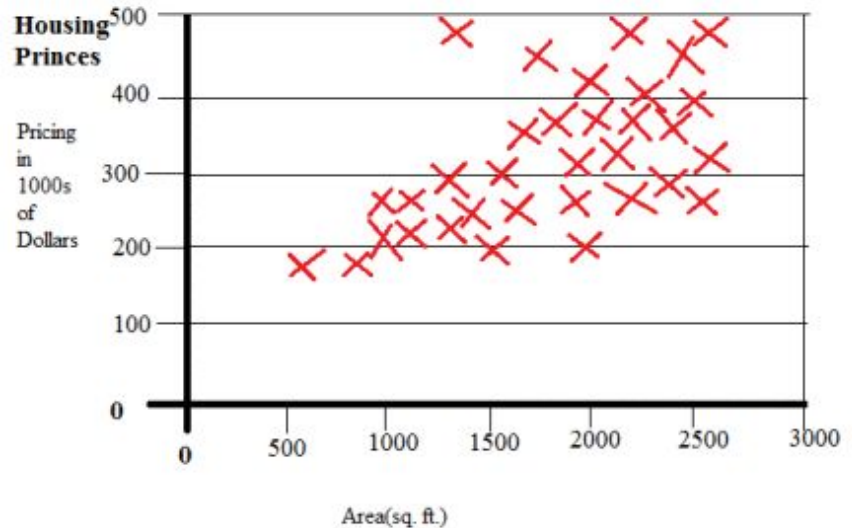


Image source: Wingshore

Regression problem - Supervised learning

In regression problems, predicted values and output are continuous.

Let's predict the price of a new input value of 1250 ft

We plot a linear straight line for this purpose

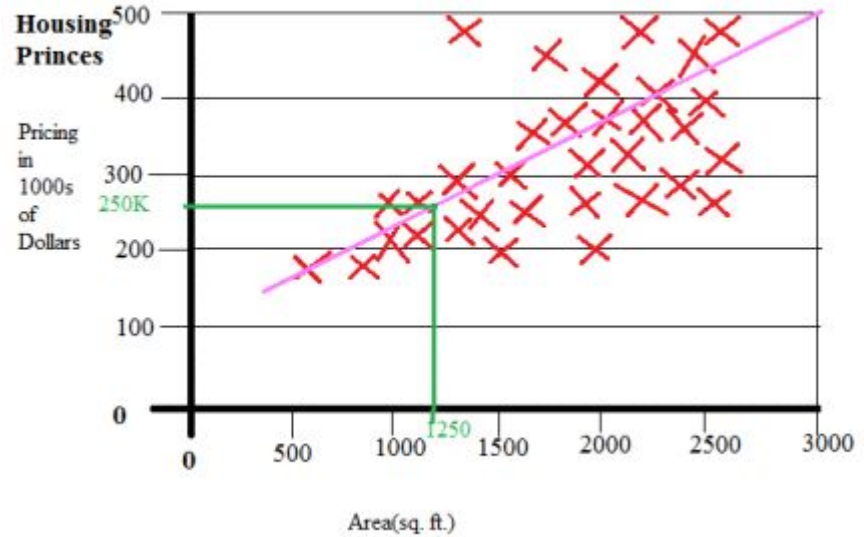


Image source: Wingshore

Classification problem - Supervised learning

In classification problems, predicted values are discrete.

An example is a CT scan image for breast cancer as benign or malignant, or classifying emails as spam or not spam.

ML algorithms include:

SVM, Logistic regression, KNN, Neural networks, etc

Classification problem - Supervised learning

In classification problems, predicted values or output are discrete.

An example is a CT scan image for breast cancer as benign or malignant, or classifying emails as spam or not spam.

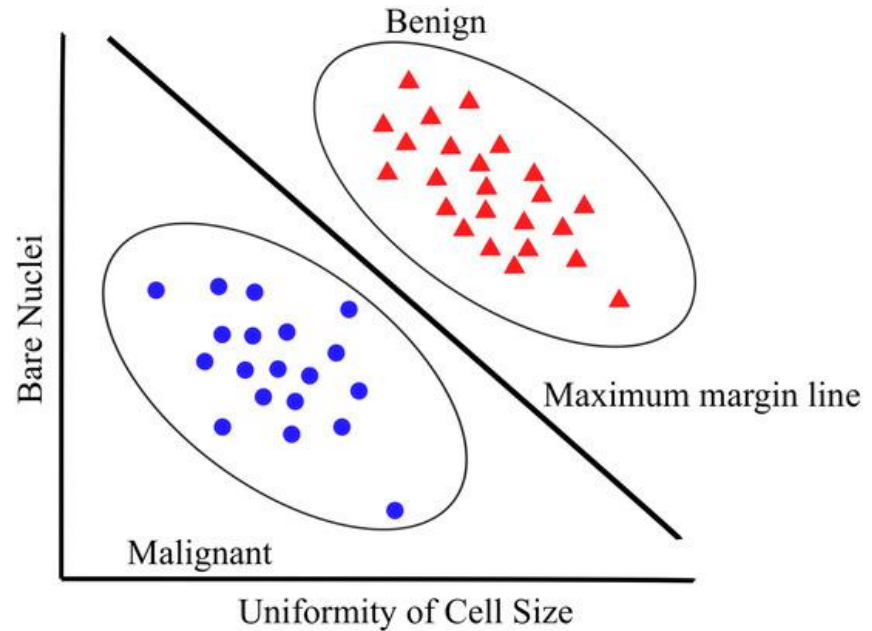


Image source: mdpi.com

semicolon

Unsupervised learning (Clustering)

In unsupervised learning, no answers (class labels) are given.

Here, we tell the program:

“Here is the dataset. Can you find some structure in it?”

Examples include:

Market segmentation, Social network analysis, Astronomical data analysis, etc.

semicolon



Unsupervised learning (Clustering)

In unsupervised learning, no answers (class labels) are given.

Examples include:

*Market segmentation,
Social network analysis,
Astronomical data
analysis, etc.*

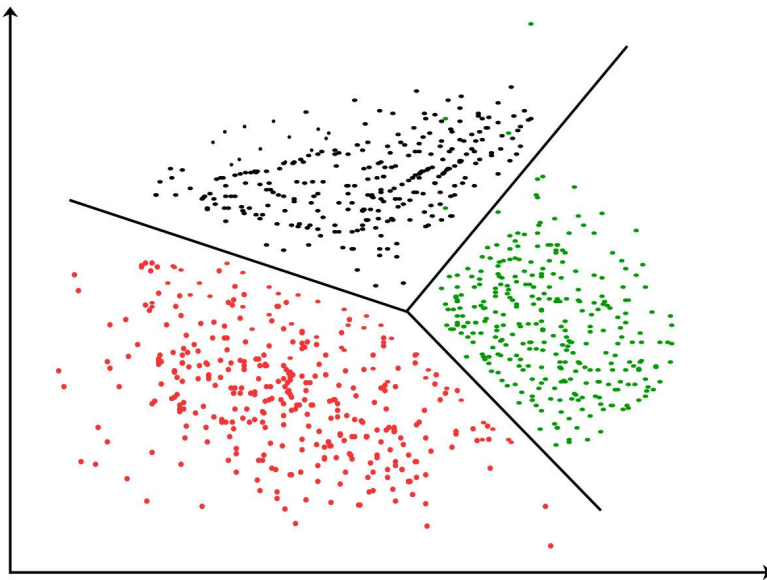


Image source: GeeksforGeeks

Any questions?

semicolon

