

## *Lecture 1*

# Machine Learning and Data Mining

**Dr. Le Huu Ton**

# *Outline*



**Course Introduction**



**What is ML-DM?**



**Different Machine Learning Algorithms**



**Questions and Answers**

# *Outline*



**Course Introduction**



**What is ML-DM?**



**Different Machine Learning Algorithms**



**Questions and Answers**

# What is Artificial Intelligent?

## Course Evaluation:

Final Examination: 50%

Attendance: 10%

Exercises or Group Presentation: 40%

## Contact

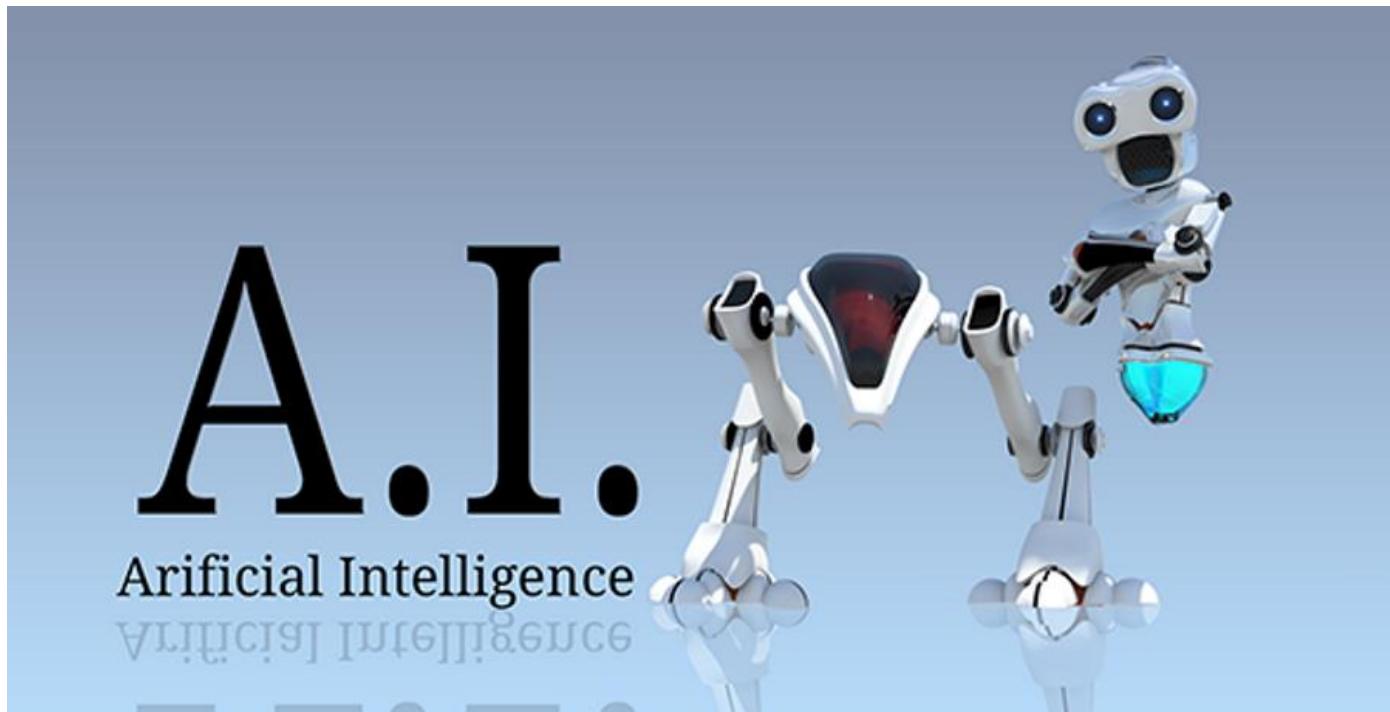
E-mail: lehuuton@gmail.com

le-huu.ton@usth.edu.vn

Mobile: 0385794025

# What is Artificial Intelligent?

## What is Artificial Intelligent?



# What is Artificial Intelligent?



- Object detection
- Face recognition
- Image enhancement
- Motion analysis
- ...

# What is Artificial Intelligent?

# What is Artificial Intelligent?

## What is Artificial Intelligent?

Artificial intelligence is a sub-field of computer science. Its goal is to enable the development of computers that are able to do things normally done by people -- in particular, things associated with people acting intelligently.

# What is Artificial Intelligent?

## Strong AI vs Weak AI

**Weak AI:** aiming at building a system that behaves like human



IBM's Deep Blue beat the World Chess Champion

**Strong AI:** aiming at building a system that not only behaves like human but also think like human



Google Deep Learning- Alpha Go defeat the World Go Champion

# Artificial Intelligent - Machine Learning

How to achieve Artificial Intelligent???



**Machine learning: The machine build its intelligence by learning and experience.**

# What is Machine Learning

*“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, Carnegie Mellon University

## Example

A computer learn to play chess by watching thousands of chess matches

Task T: Playing chess

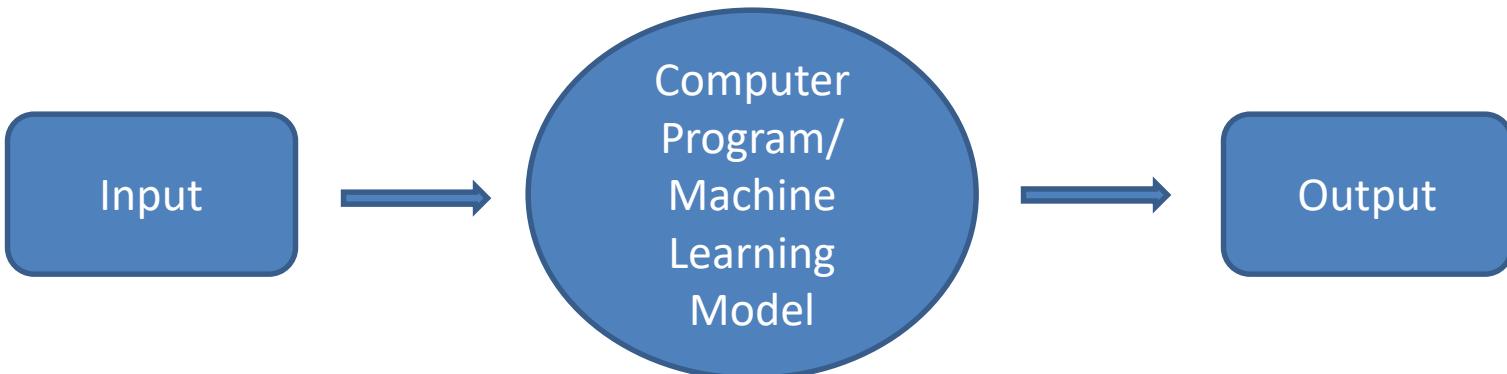
Experience E: Watching chess matches

Measure P: How many games it wins vs human or other computer

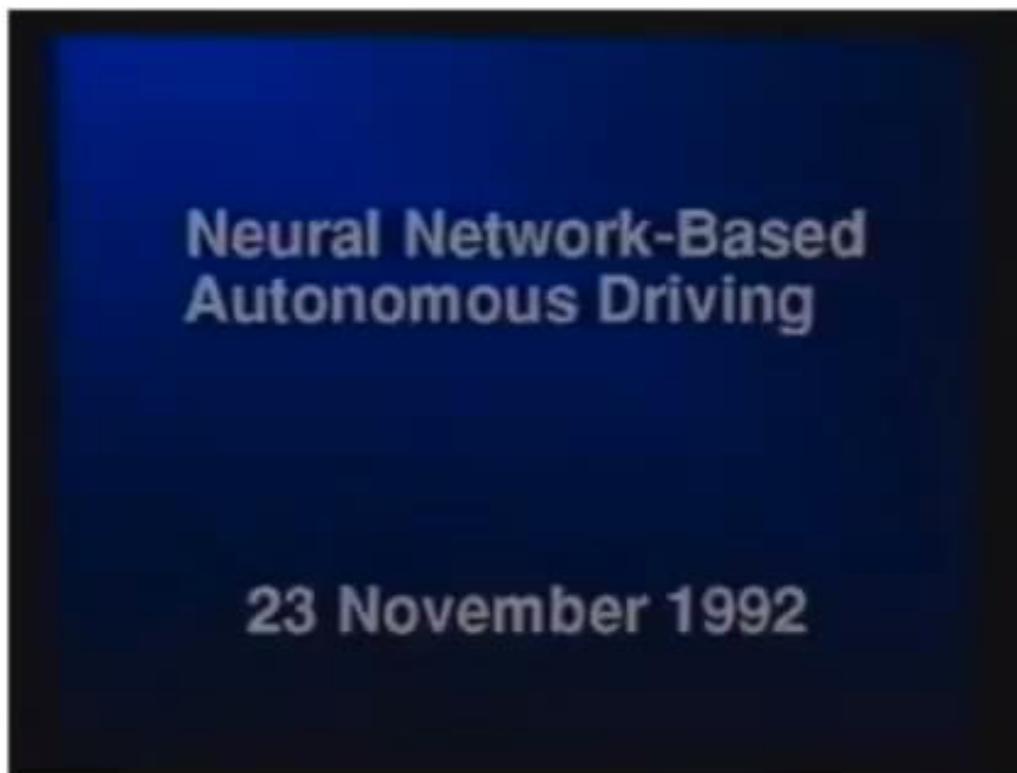
# Introduction to Machine Learning (3/5)

Find the biggest number among three inputs:  $a$ ,  $b$ ,  $c$

Traditional Approach	Machine Learning Approach	
<pre>If (a&gt;b)&amp;&amp;(a&gt;c):     Output=a; Else if (b&gt;a)&amp;&amp;(b&gt;c)     Output=b; Else     Output =c;</pre>	Input (3,5,1) (1,3,5) (73,12,77) (-12,43,67) ...	Output 5 5 77 67
Computer Program	Machine Learning Model	



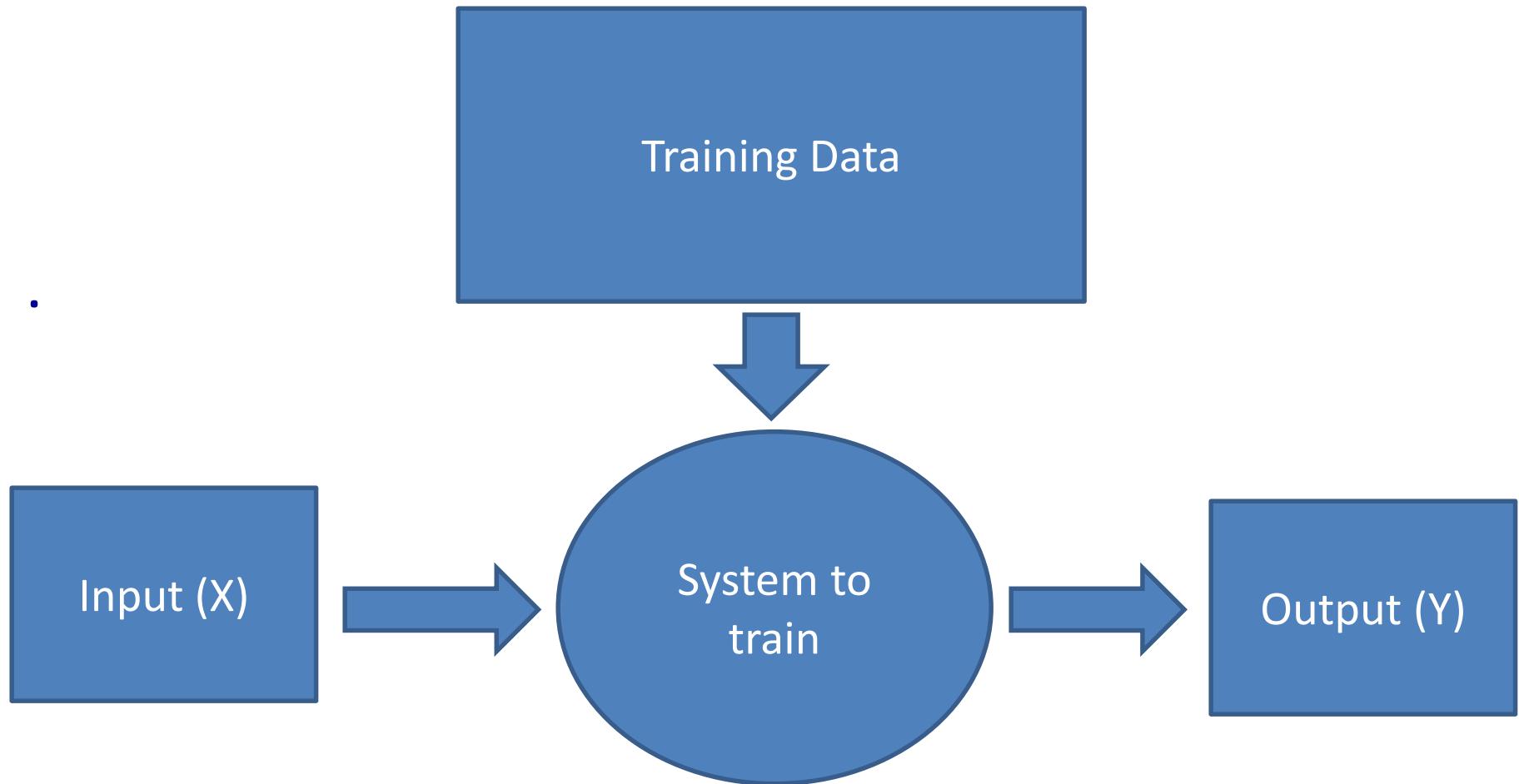
# Machine Learning



[Courtesy of Dean Pomerleau]

Autonomous Driving System [2]

# Machine Learning



# Machine Learning

## Unsupervised Learning, Supervised Learning and Semi-Supervised Learning

**Supervised Learning :** algorithms are trained using labeled examples, such as an input where the desired output is known. Predict output for given input.

+ Regression: The output is continuous, it can be any value

+ Classification: The output is discrete, it only can be some specific values

# Supervised Learning

**Example 1: Given a bundle of emails, which is already labeled as spam or normal email. Predict if an email is spam or not**

**Example 2: Given the housing price for different houses in Hanoi. Predict the housing price of a given house**

Square (m <sup>2</sup> )	Location	Price (billion VND)
30	Thanh Xuan	2.5
43	Thanh Xuan	3.4
25	Ba Dinh	6.1
51	Ba Dinh	13
33	Ba Dinh	?????

# Machine Learning

**Unsupervised Learning:** is used against data that has no historical labels, the training data does not provide the correct answer or the real output of the training input

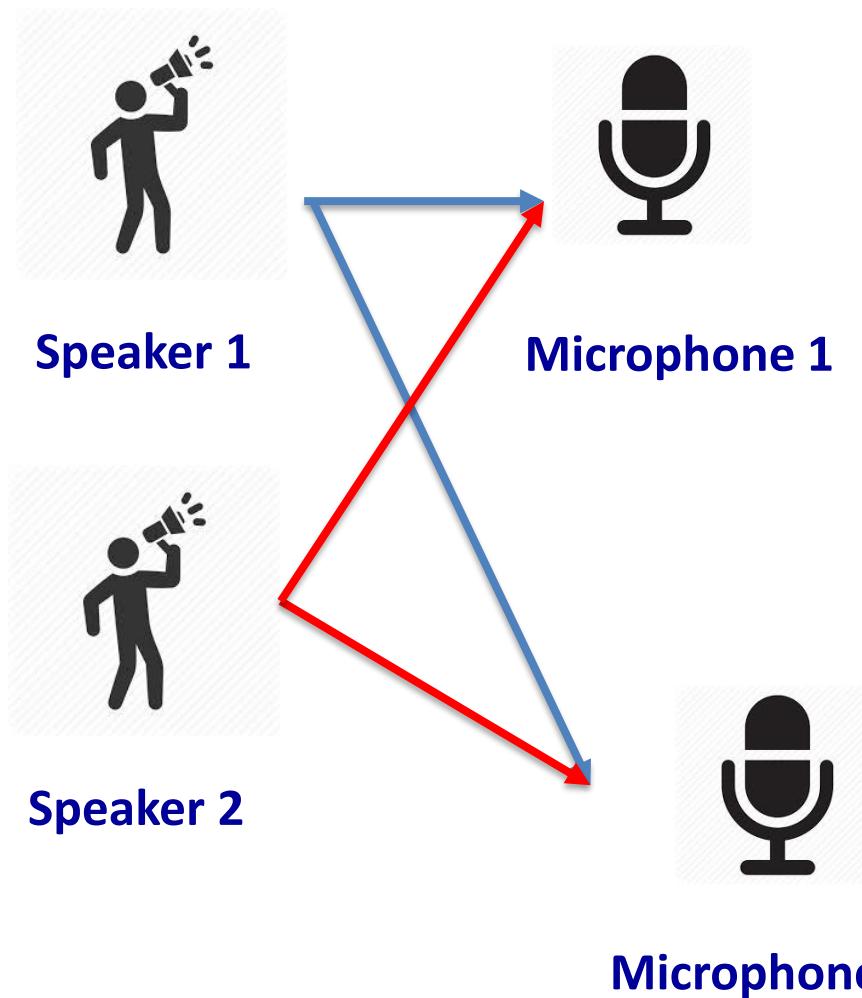
- Group similar inputs into the same group, cluster

**Semi- Supervised Learning:** is used for the same applications as supervised learning. But it uses both labeled and unlabeled data for training – typically a small amount of labeled data with a large amount of unlabeled data

**Reinforcement learning:** The system is not told which is the correct output, as in most forms of machine learning, but instead must discover which output yield the most reward by trying them.

# Machine Learning

## Example of Unsupervised Learning: Cocktail Party Problem



# Machine Learning

Speakers [3]



Recorders



Microphone 1



Microphone 2



Microphone 3

Outputs



Output 1



Output 2



Output 3

# Machine Learning

**Exercise:**

**Determine the Task T, Performance P, Experience E for the cocktail party problem**

.

# Machine Learning

Method	Training Data	Targets- Goals
Supervised Learning	Labeled Data Input-Output	Regression, Classification
Un-Supervised Learning	Unlabeled Data Input	Clustering
Semi-Supervised Learning	Both Labeled and Unlabeled Data	Regression, Classification
Reinforcement Learning	Input- Indirect Output (rewards)	Regression, Classification

# Machine Learning

## More Exercises:

- 1. Given following example, determine if it is supervised- unsupervised learning, regression or classification. Identify experience E, Task T, and performance P, what is the input, the output.**
  1. Collect data of mid-term score, attendance and final score in Machine Learning Class from previous year student. Predict if a student from this year pass the final exam of Machine learning or not, based on his mid-term score and attendance
  2. Same situation, but predict the final score of the student.
  3. Same situation, but in order to divided the class into 3 groups projects, such as students with similar level are in the same group
  4. You have bunch of photos of 6 people but without information who is on which one and want to divide this dataset into 6 folders, each with photos of one individual
- 2. If you want to build a system to predict the price of Iphone 14s in the next 6 months. What do you do? (which kind of learning? What is the input, what is the output, which data can be used to train the system,...)**

# Q & A

It's time for QUESTION

.

## References

- [1] <http://www.livescience.com/49007-history-of-artificial-intelligence.html>
- [2] <https://www.youtube.com/watch?v=iIP4aPDTBPE>
- [3]