TELKOM
DIGITAL TALENT
INCUBATOR 2020

Telkom University
Telkom Indonesia

DATA SCIENTIST

Muhammad Apriandito A.S.

Organized by:

CAE

TELKOM
DIGITAL TALENT
INCUBATOR 2020

Modul 3: Introduction to Machine Learning

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Module Overview

Topics

- What is Machine Learning?
- Application of Machine Learning
- Machine Learning Approach
- Type of Machine Learning
- Machine Learning Steps
- Deep Learning

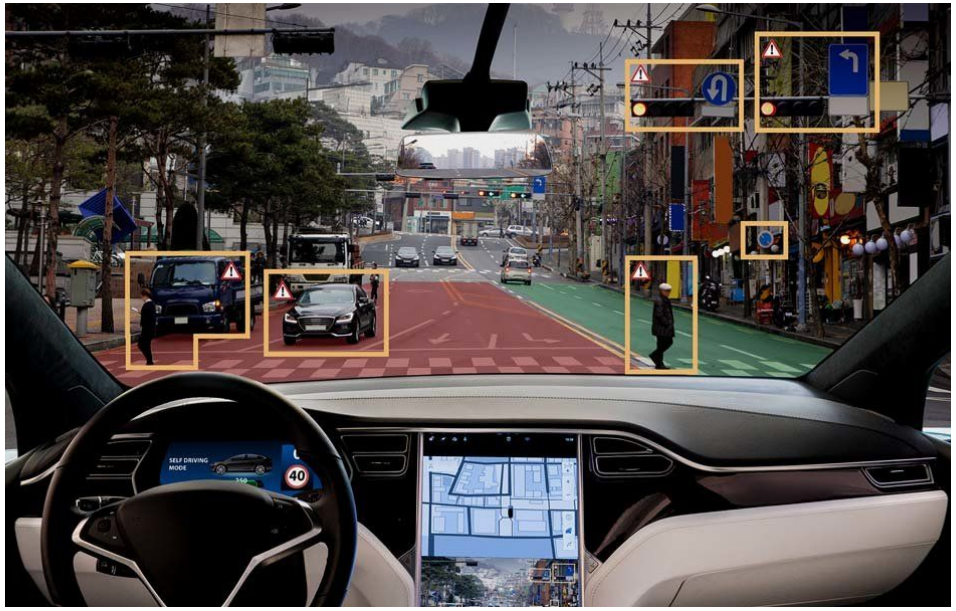
TELKOM DIGITAL TALENT INCUBATOR 2020

Module Objectives

- Understand what is Machine Learning, and Its Applications.
- Understand how Machine Learning algorithm work.
- Understand different type of Machine Learning

TELKOM DIGITAL TALENT INCUBATOR 2020

Our World Today



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

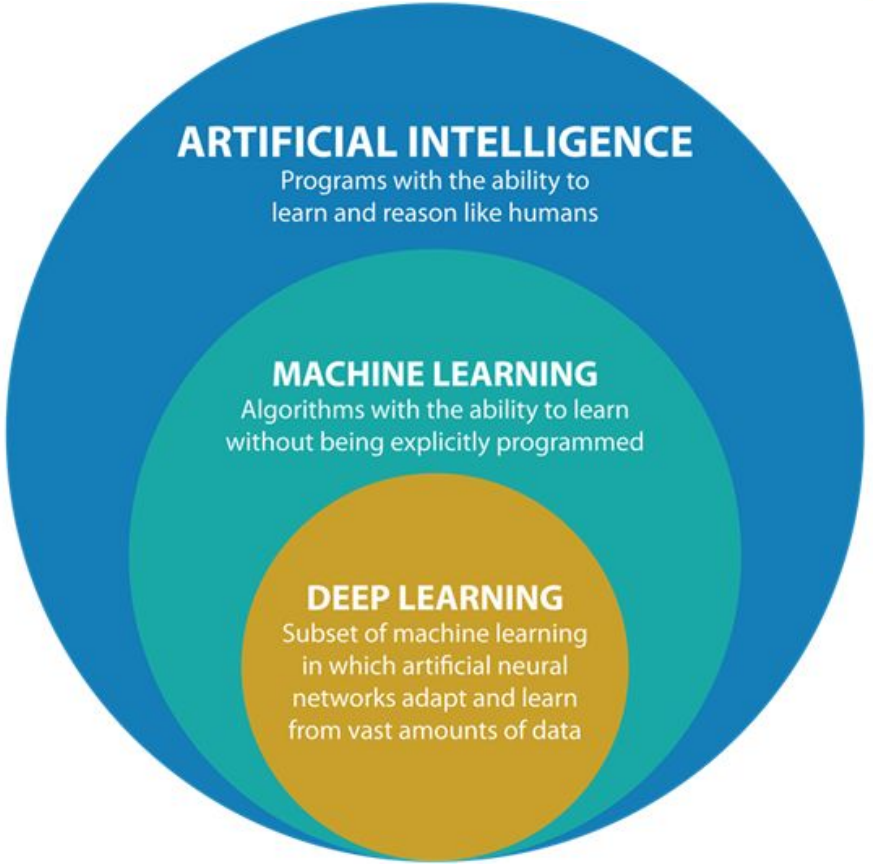
What is Machine Learning?

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

What is Machine Learning

- Machine Learning is a type of Artificial Intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.



TELKOM DIGITAL TALENT INCUBATOR 2020

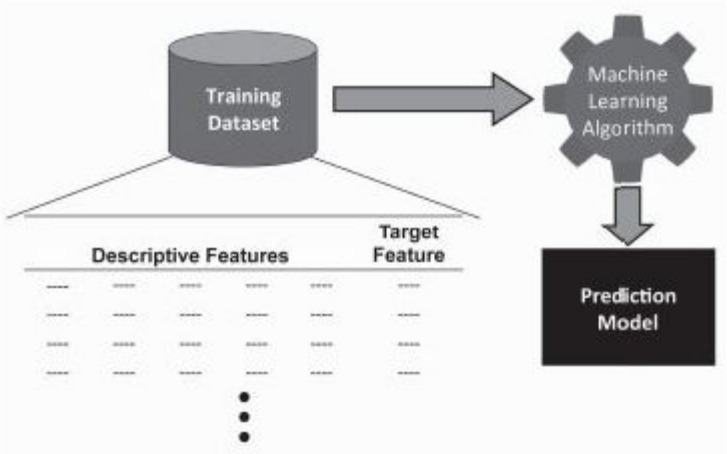
What is Machine Learning

Organized by:
CAE

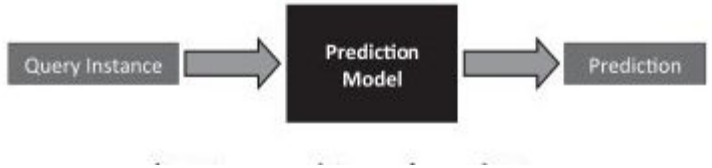
TELKOM DIGITAL TALENT INCUBATOR 2020

What is Machine Learning

- Machine learning algorithms automate the process of learning a model that captures the relationship between the descriptive features and the target feature in a dataset.



Learning from historical data



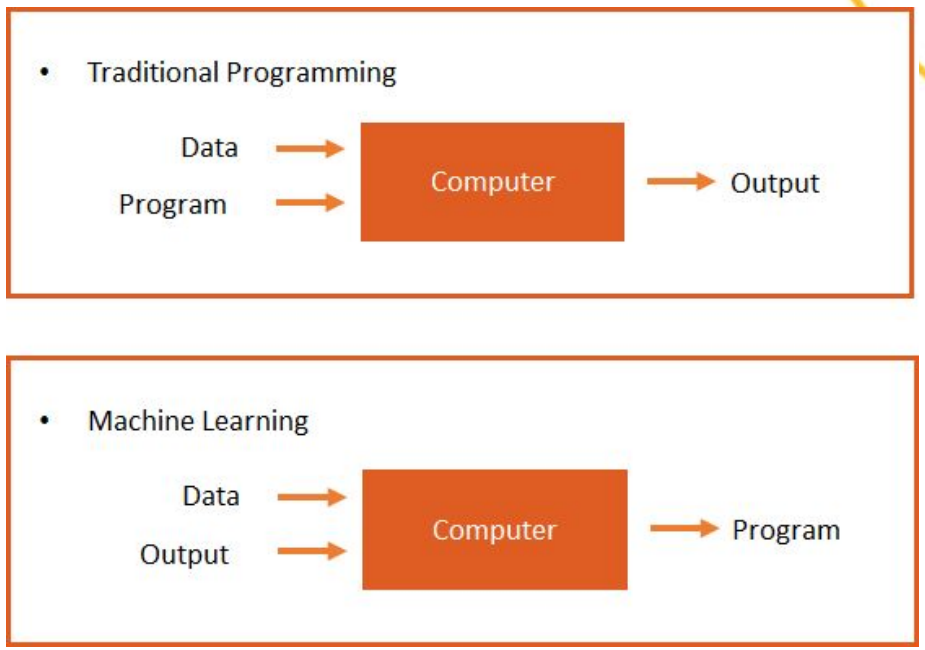
Using model to predict a new data

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Using Data to Answer Question

- Machine Learning is an idea to learn from examples and experience, without being explicitly programmed.
- Instead of writing code, we feed data to the generic algorithm, and it builds logic based on the data given.



TELKOM DIGITAL TALENT INCUBATOR 2020

Why we need Machine Learning?

- Pattern Identification
- Process Automation

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Pattern Identification

Lama bekerja (Tahun)	Gaji (Rupiah)
2	3.000.000
4	6.000.000
6	9.000.000
10	15.000.000
12	24.000.000
14	28.000.000

TELKOM DIGITAL TALENT INCUBATOR 2020

Pattern Identification

Lama bekerja (Tahun)	Gaji (Rupiah)
2	3.000.000
4	6.000.000
6	9.000.000
10	15.000.000
12	24.000.000
14	28.000.000

```
if (experience < = 10)
{ salary = experience * 1.5 * 100000}
else if(experience >10)
{ salary = experience * 2 * 100000}
```

TELKOM DIGITAL TALENT INCUBATOR 2020

Pattern Identification

Lama bekerja	Level Pekerjaan	Pendidikan Terakhir	Gaji (Rupiah)
2	3	Yes	4.500.000
4	3	No	6.000.000
6	4	No	7.500.000
10	5	Yes	18.000.000
12	5	No	15.000.000
14	6	No	18.000.000

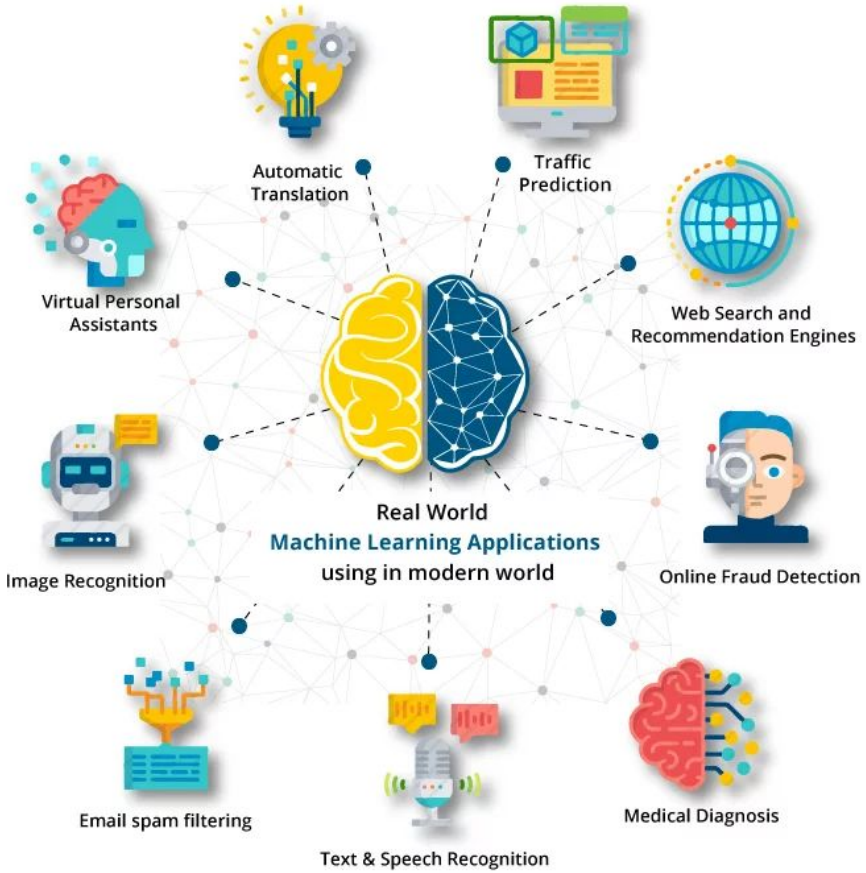
TELKOM DIGITAL TALENT INCUBATOR 2020

Pattern Identification

Lama bekerja	Level Pekerjaan	Pendidikan Terakhir	Gaji (Rupiah)
2	3	Yes	4.500.000
4	3	No	6.000.000
6	4	No	7.500.000
10	5	Yes	18.000.000
12	5	No	15.000.000
14	6	No	18.000.000

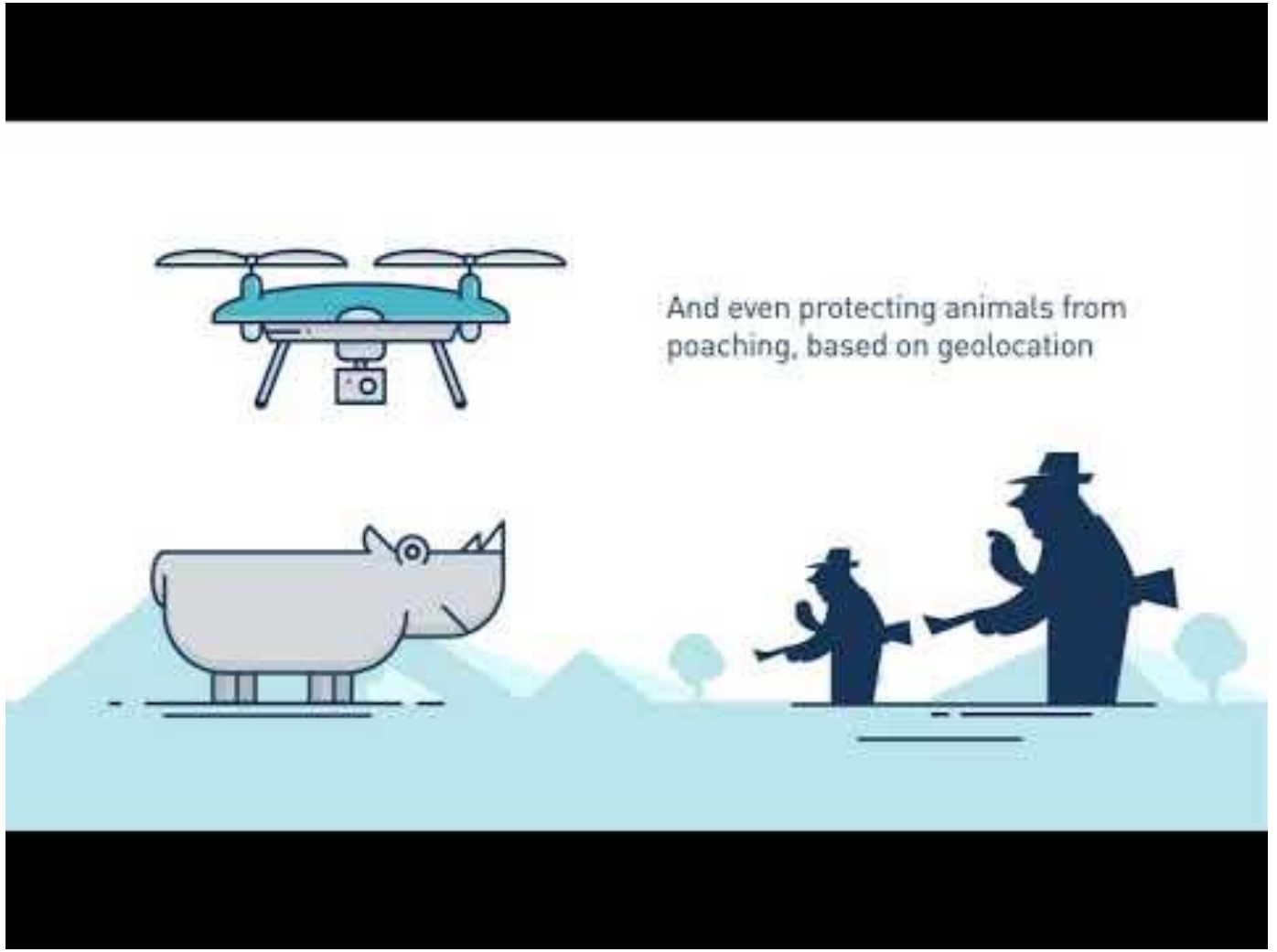
```
Salary = Experience * Magic_Number_1 + JobLevel * Magic_Number_2 +
Skill * Magic_Number_3 + Magic_Number_4
```

Application of Machine Learning



TELKOM DIGITAL TALENT INCUBATOR 2020

How Machine Learning Help Businesses

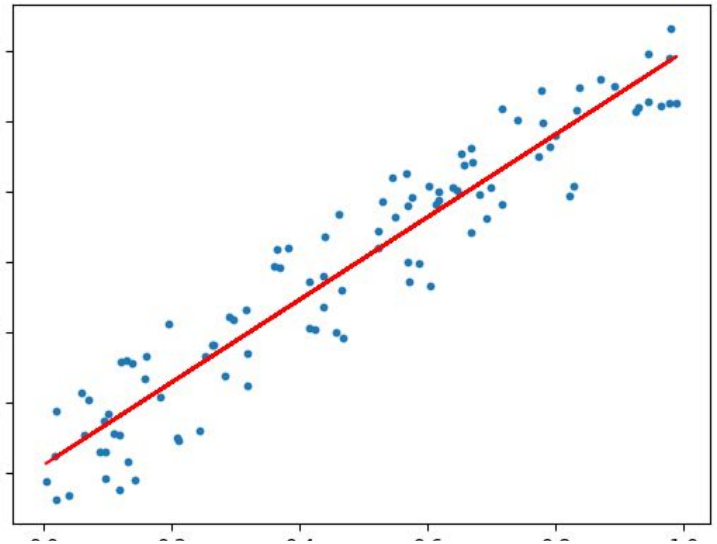


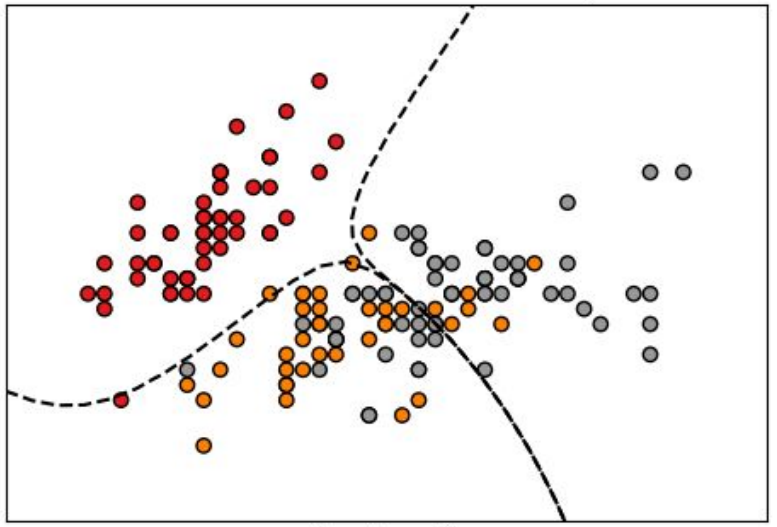
Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

How can a machine learn?

- Machine learning algorithms work by searching through a set of possible prediction models for the model that best captures the relationship between the descriptive features and target feature in a dataset.





Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Machine Learning Approach

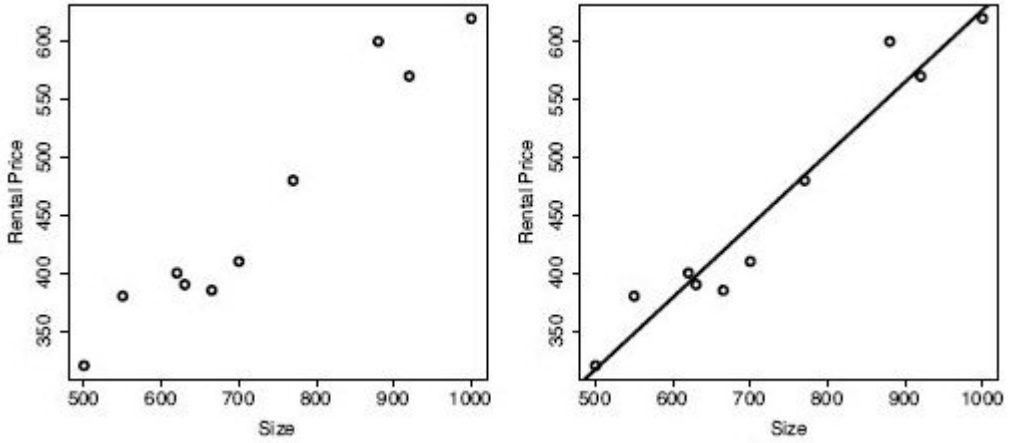
- Error Based Learning
- Information-based Learning
- Similarity-based Learning
- Probability-based Learning

TELKOM DIGITAL TALENT INCUBATOR 2020

Error Based Learning

- In error-based machine learning, Machine Learning perform a search for a set of parameters for a parameterized model that minimizes the total error across the predictions made by that model with respect to a set of training instances.

ID	Size	Floor	Broadband Rate	Energy Rating	Rental Price
1	500	4	8	C	320
2	550	7	50	A	380
3	620	9	7	A	400
4	630	5	24	B	390
5	665	8	100	C	385
6	700	4	8	B	410
7	770	10	7	B	480
8	880	12	50	A	600
9	920	14	8	C	570
10	1,000	9	24	B	620



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Information Based Learning

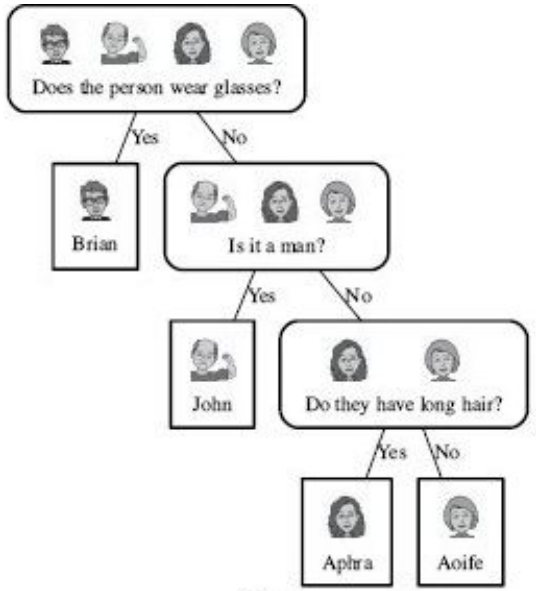
Brian

John

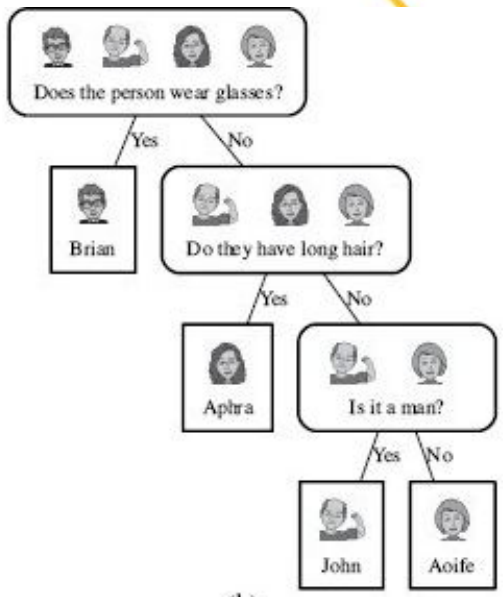
Aphra

Aoife

Man	Long Hair	Glasses	Name
Yes	No	Yes	Brian
Yes	No	No	John
No	Yes	No	Aphra
No	No	No	Aoife



(a)



(b)

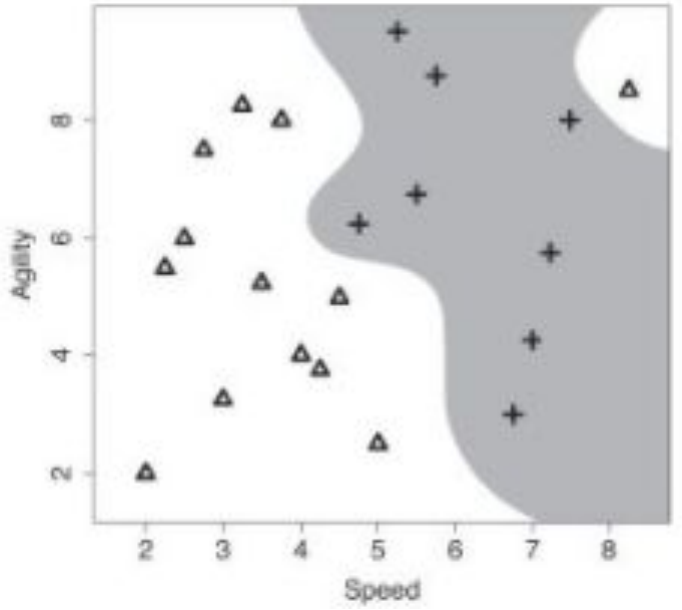
Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Similarity Based Learning

- Similarity-based approaches to machine learning come from the idea that the best way to make a predictions is to simply look at what has worked well in the past and predict the same thing again.

	Grrr!			Score
	✓	✗	✗	1
	✗	✓	✗	1
	✗	✓	✓	2



Organized by:
CAE

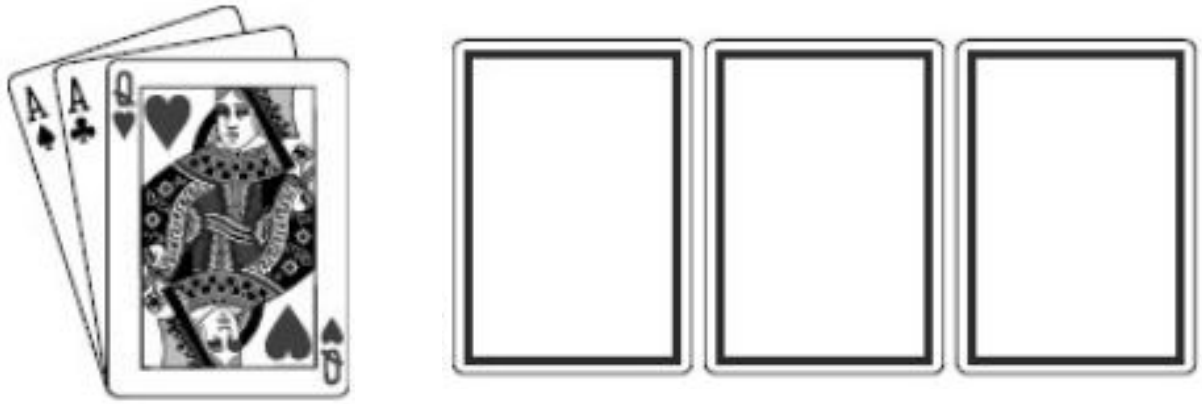
TELKOM DIGITAL TALENT INCUBATOR 2020

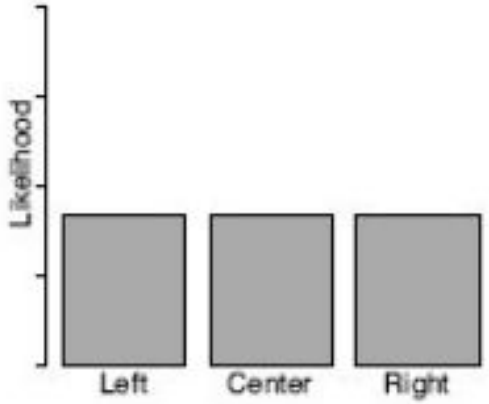
Probability Based Learning

Organized by:
CAE

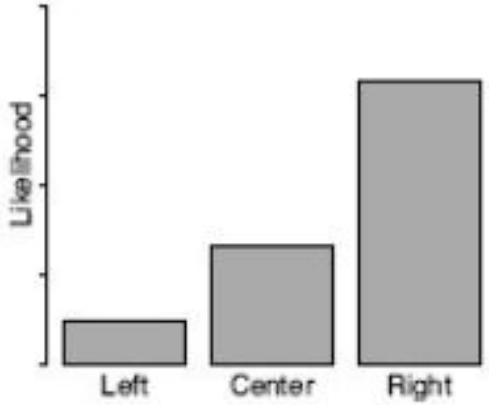
TELKOM DIGITAL TALENT INCUBATOR 2020

Probability Based Learning





Initial Likelihood



Revised Likelihood

Organized by:
CAE

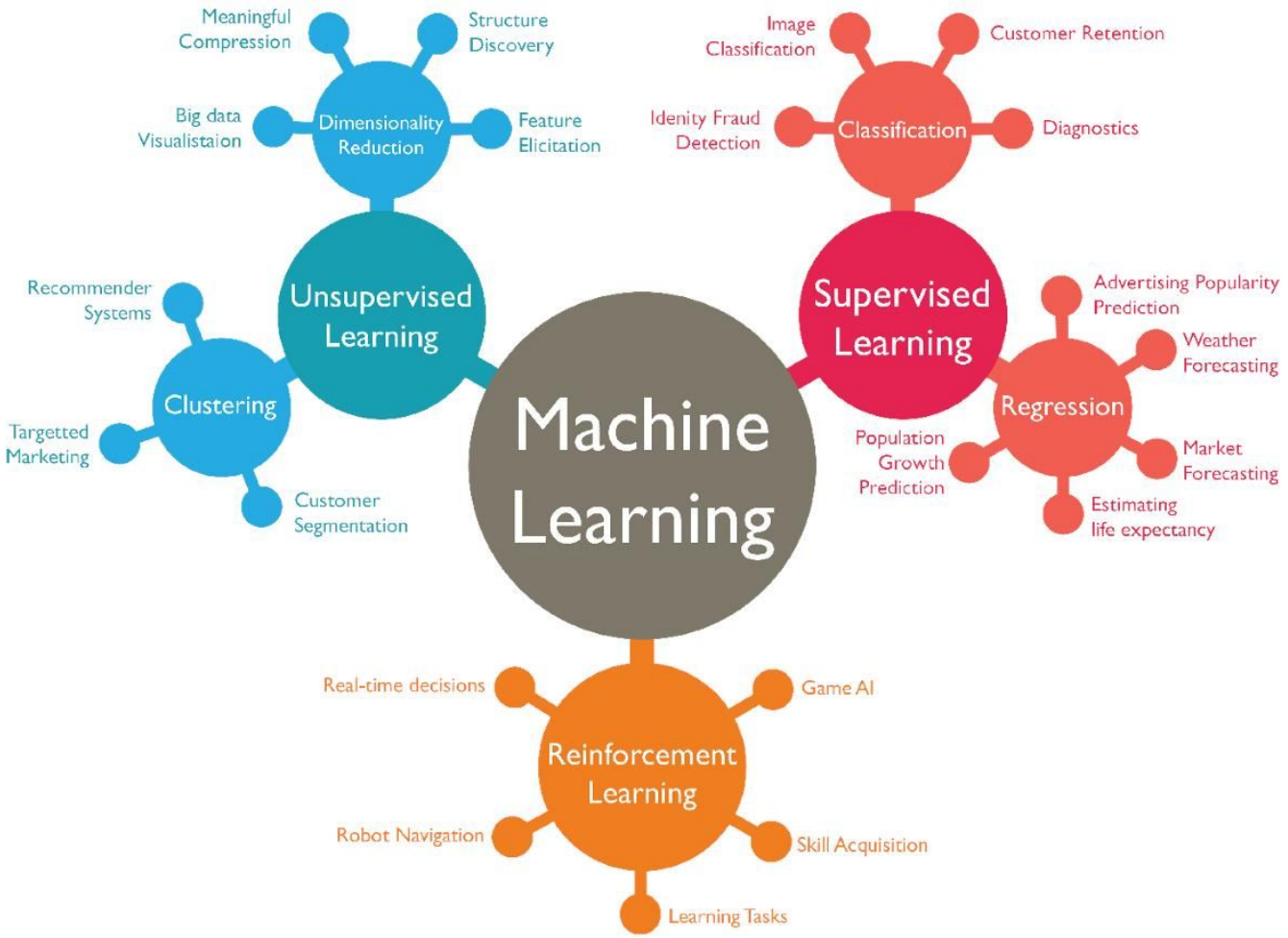
TELKOM DIGITAL TALENT INCUBATOR 2020

Type of Learning

- Supervised Learning.
- Unsupervised Learning.
- Reinforcement Learning.

TELKOM DIGITAL TALENT INCUBATOR 2020

Type of Learning



Organized by:

CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

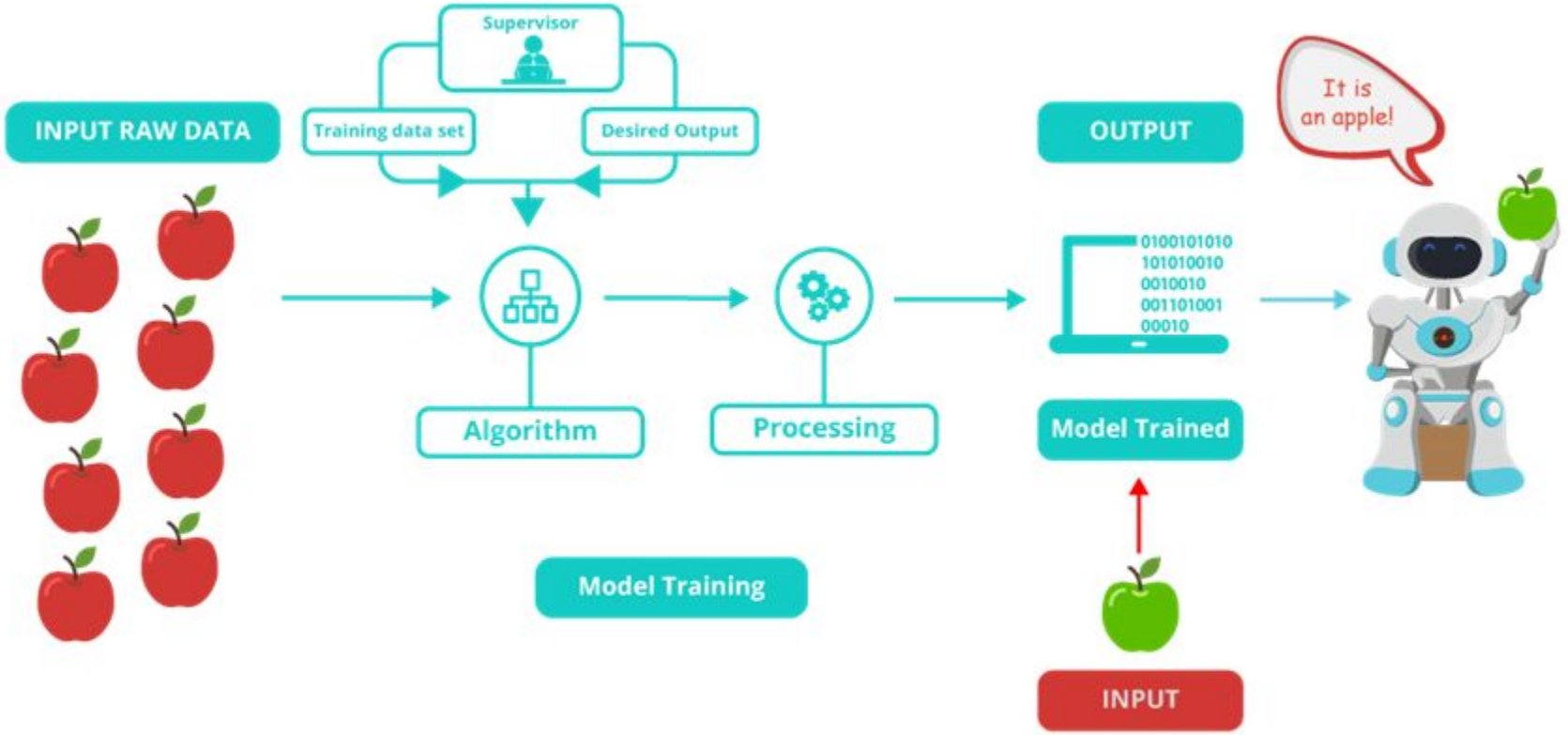
Supervised Learning

- Supervised Learning approach is indeed like human learning under the supervision of a teacher. The teacher provides good examples for the student to memorize, and the student then derives general rules from these specific examples.

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Supervised Learning



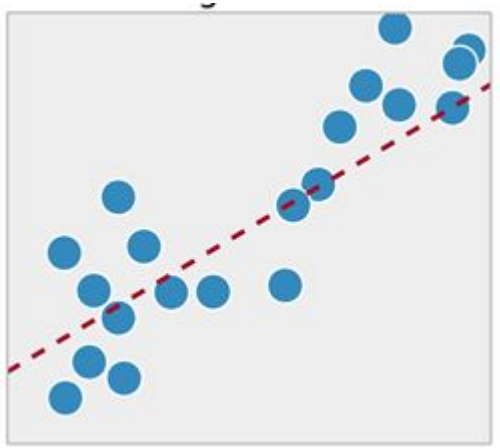
Organized by:

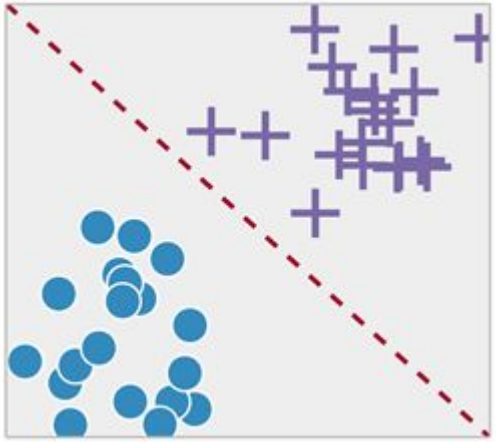
CAE

Supervised Learning Problem

- Regression : A regression problem is when the output variable is a real value, such as “weight” or “height.”

- Classification : A classification problem is when the output variable is a category or a group, such as “black” or “white” or “spam” and “no spam”.





TELKOM DIGITAL TALENT INCUBATOR 2020

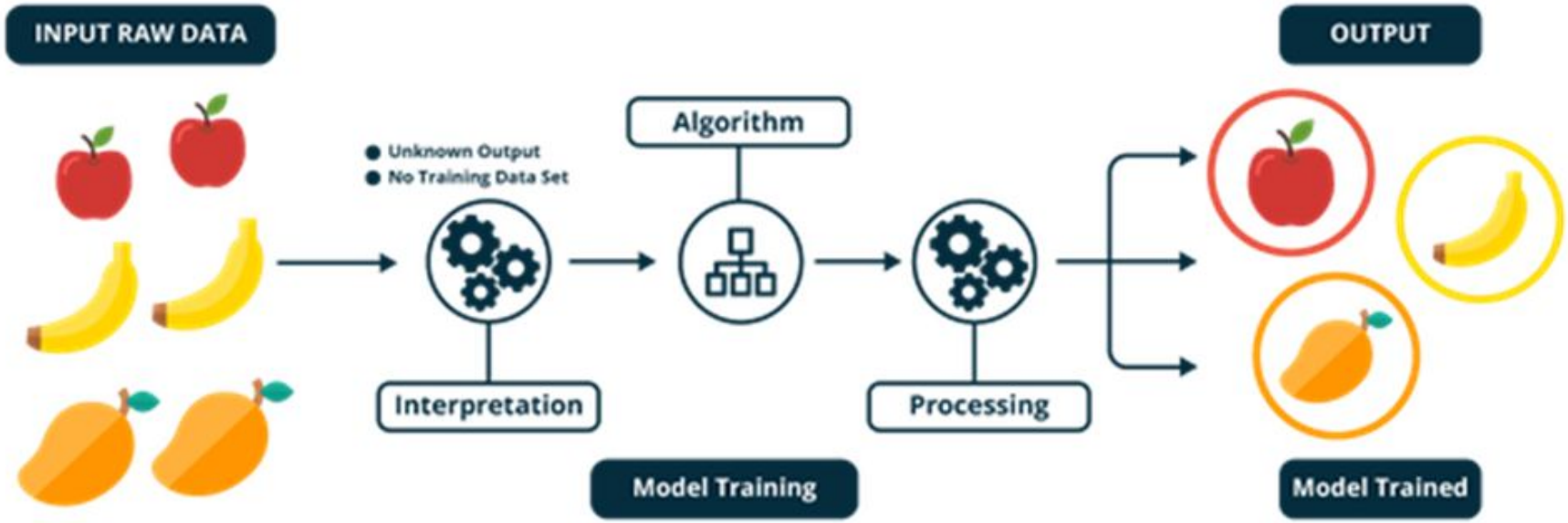
Unsupervised Learning

- Mathematically, Unsupervised learning is where you only have input data (X) and no corresponding output variables. The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data.

Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Unsupervised Learning



Organized by:

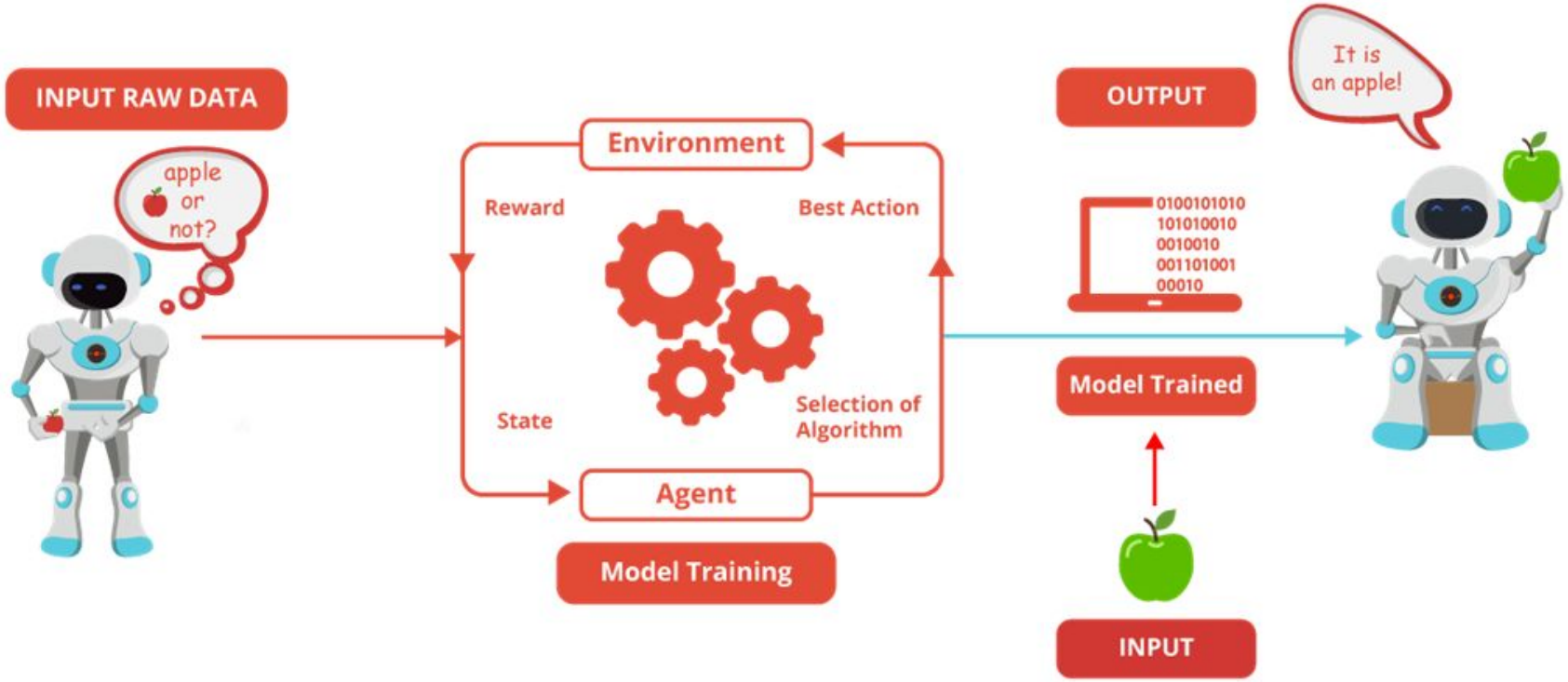
TELKOM DIGITAL TALENT INCUBATOR 2020

Unsupervised Learning Problem

- **Clustering**: Grouping a set of objects in such a manner that objects in the same group are more similar than to those object belonging to other groups.
- **Association**: Finding associations amongst items within large commercial databases.

Organized by:
CAE

Reinforcement Learning



Organized by:

TELKOM DIGITAL TALENT INCUBATOR 2020

Machine Learning Steps

- Gathering Data
- Preparing that Data
- Choosing a Model
- Training
- Evaluation
- Hyperparameter Tuning

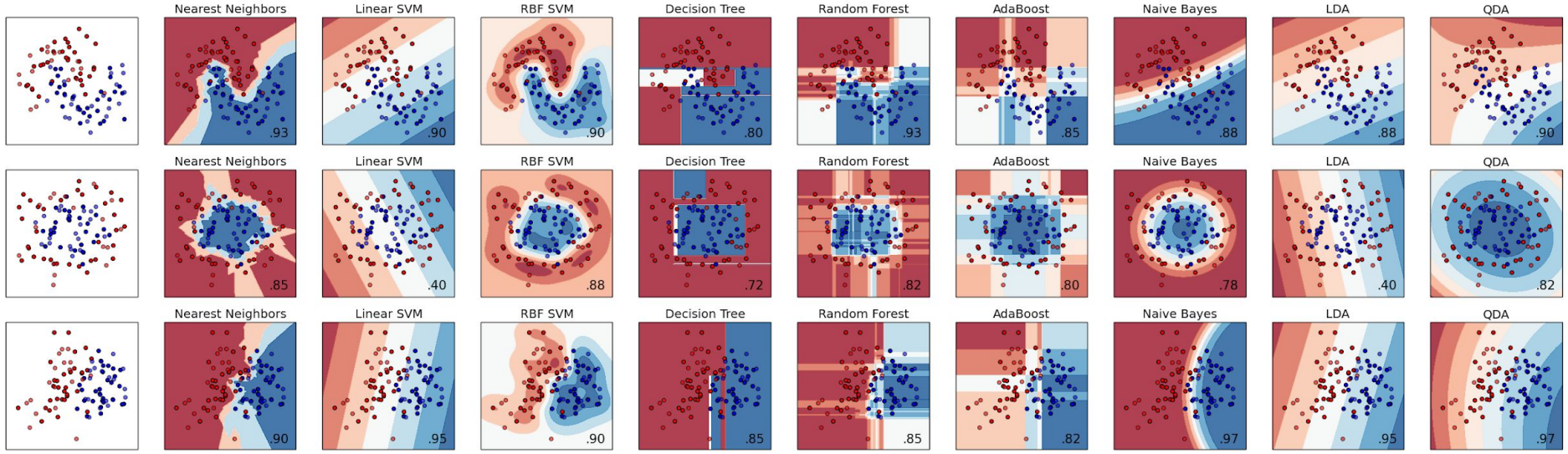
TELKOM DIGITAL TALENT INCUBATOR 2020

Machine Learning Steps

Organized by:
CAE

Overfitting and Underfitting

- Different machine learning algorithms encode different inductive biases. Because a machine learning algorithm encodes an inductive bias, it can induce models that generalize beyond the instances in a training dataset.



TELKOM DIGITAL TALENT INCUBATOR 2020

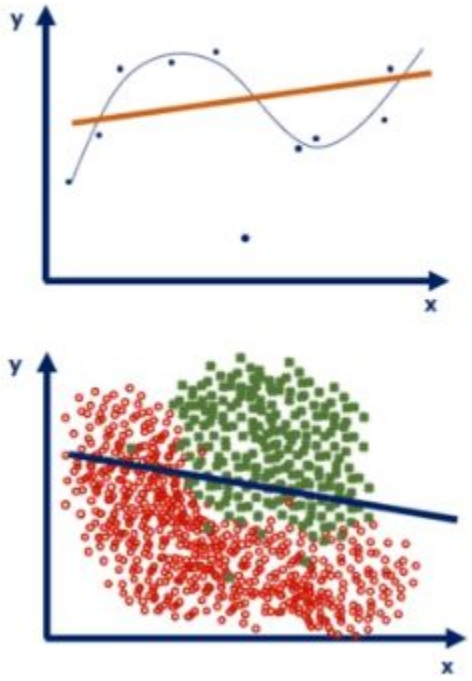
Overfitting and Underfitting

- There are two kinds of mistakes that an inappropriate inductive bias can lead to: **underfitting and overfitting .**
- **Underfitting** occurs when the prediction model selected by the algorithm is too simplistic to represent the underlying relationship in the dataset between the descriptive features and the target feature.
- **Overfitting** occurs when the prediction model selected by the algorithm is so complex that the model fits to the dataset too closely and becomes sensitive to noise in the data.

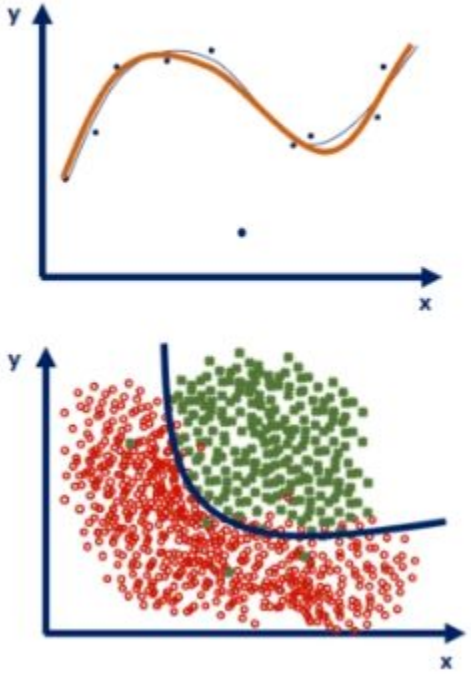
TELKOM DIGITAL TALENT INCUBATOR 2020

Overfitting and Underfitting

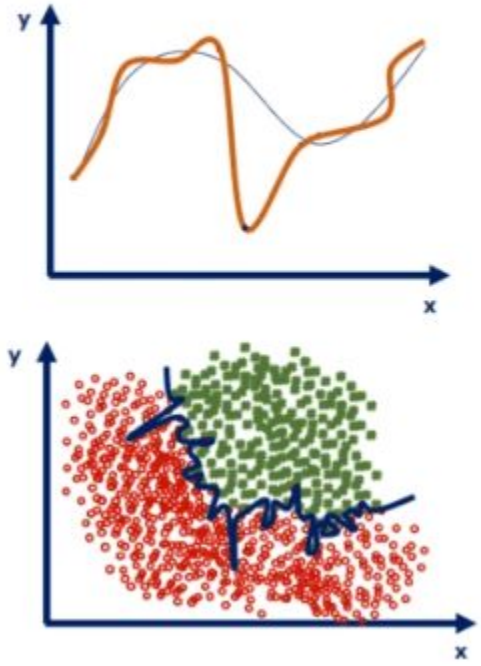
An **underfitted** model



A **good** model



An **overfitted** model



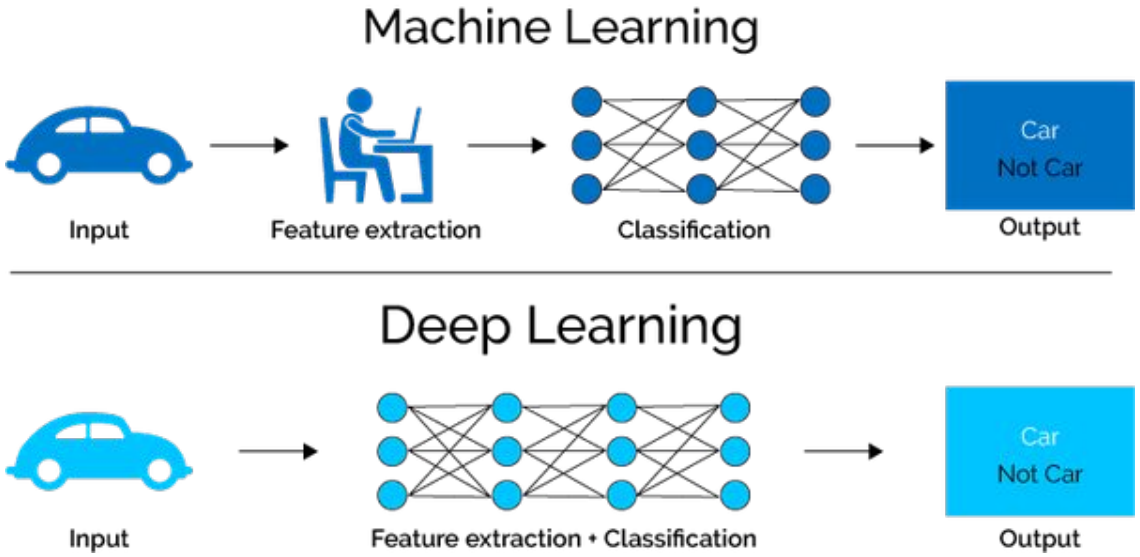
Organized by:

CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Deep Learning

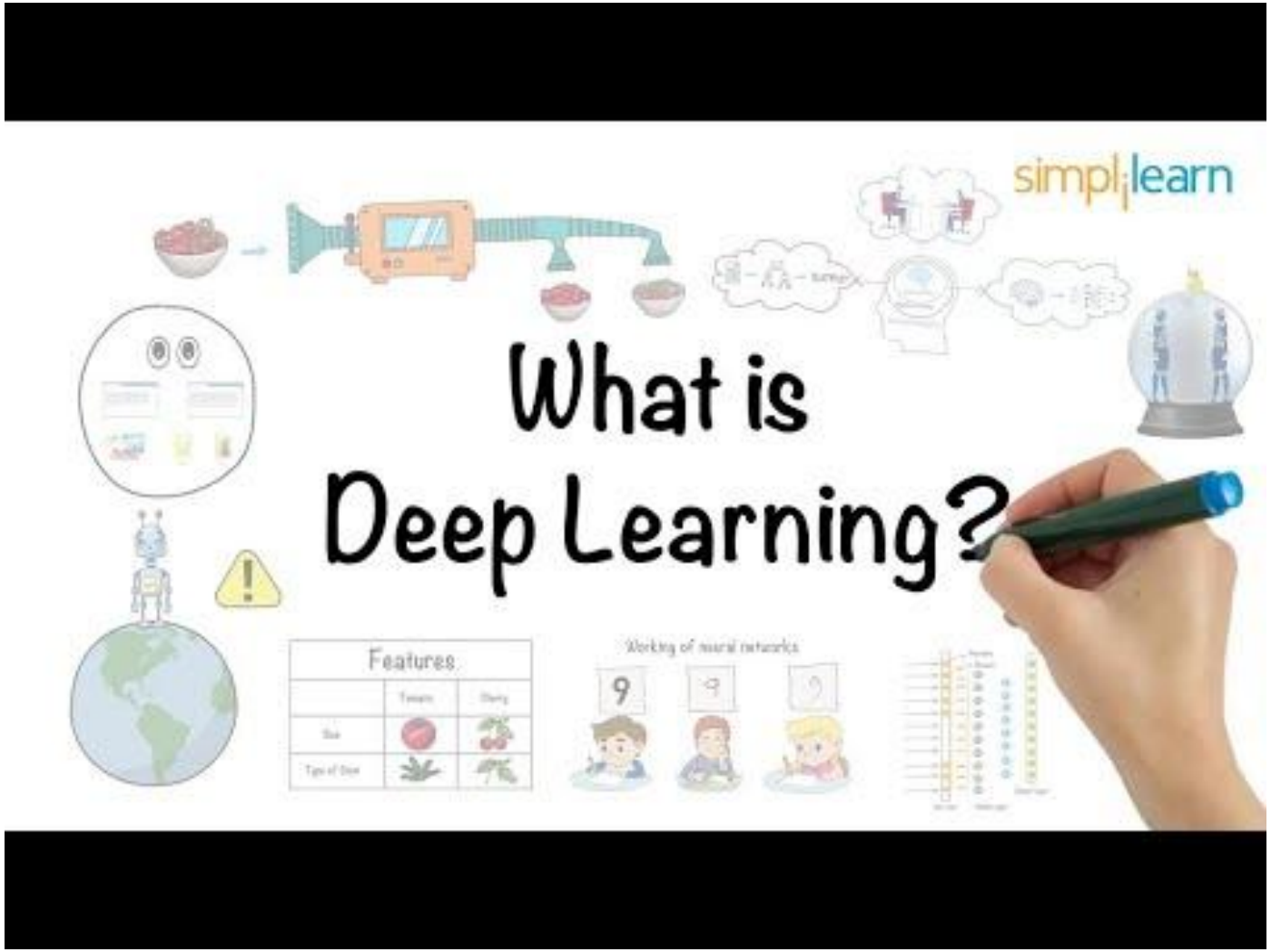
- Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

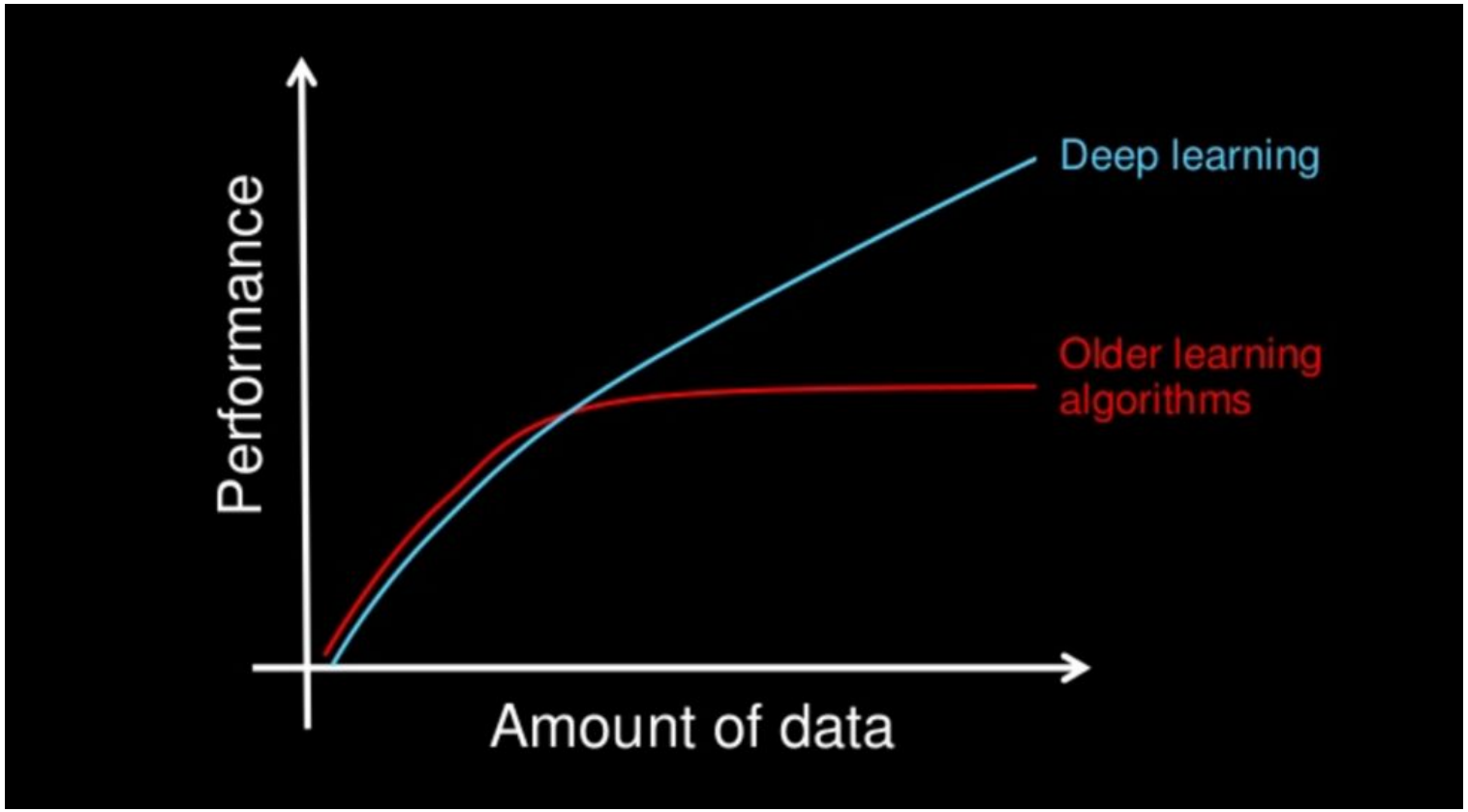
Deep Learning



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Why use Deep Learning



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Machine Learning vs Deep Learning

Classical Machine Learning:

- Interpretability and explainability are paramount
- Smaller amounts of relatively simple data
- Straightforward feature engineering
- Limited computational power
- Limited time, need for faster prototyping and operationalization
- Need for varied algorithm choices
- Accuracy of test dataset results is acceptable

Deep Learning:

- Very high accuracy is a priority (and primes over straightforward interpretability and explainability)
- Large amounts of precisely labeled data
- Complex feature engineering
- Powerful compute resources available (GPU acceleration)
- Augmentation and other transformations of the initial dataset will be necessary

TELKOM DIGITAL TALENT INCUBATOR 2020

Module Summary

- Machine Learning is a type of Artificial Intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.
- There are 3 types of machine learning, named Supervised Learning, Unsupervised Learning, Reinforcement Learning.
- Supervised learning is like a teacher teach a student.
- Unsupervised learning is like a student learning by himself.
- Underfitting occurs when the prediction model selected by the algorithm is too simplistic to represent the underlying relationship in the dataset between the descriptive features and the target feature.

TELKOM DIGITAL TALENT INCUBATOR 2020

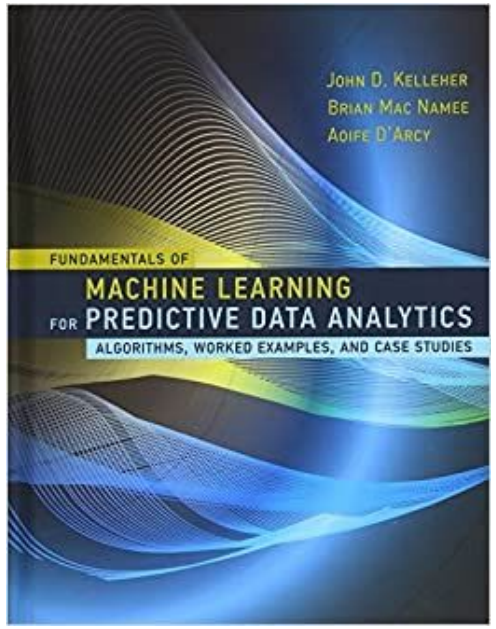
Module Summary

- Overfitting occurs when the prediction model selected by the algorithm is so complex that the model fits to the dataset too closely and becomes sensitive to noise in the data.

TELKOM DIGITAL TALENT INCUBATOR 2020

References/Additional Resources

- Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies - Kelleher John D. Mac Namee Brian D'Arcy Aoife
- Machine Learning Algorithms From Scratch - Matthew Tichenor



Organized by:
CAE

TELKOM DIGITAL TALENT INCUBATOR 2020

Assignment Week 3

- Find out the use of each type of Machine Learning in everyday life.
  - Write it as an article in Bahasa Indonesia
  - Post it on Medium, Blog, or other online writing media.
  - Not allowed to take other people's writings.
  - Make it as informative as possible
  - **Deadline: ???**

Organized by:
CAE