AI – Machine Learning

I. Level of AI

1. AI

2. AGI

3. ASI

II. Data in AI / ML

A good dataset? Some problems when choosing dataset?

* Wrong data
* Missing data
* Outliers data
* Unbalanced data
* Lack of variability in data

III. Keyword

1. Machine Learning

- Features engineering

Input -> Feature extraction -> classification -> output

2. Deep Learning

- Learn features by themselves

Input -> Feature extraction + classification -> output

3. Pre-processing

* Computer vision: resize
* Graph: Scale (-1 -> 1)
* Text: Tokenization

IV. Supervised – Unsupervised – Reinforcement Learning

1. Supervised Learning

- Learning on labeled data. The goal is to learn a general rule that maps inputs to outputs.

Ex: regression, classification, etc.

- Learning to creaate f(x) -> y

- Regression: mapping input value to continous output variable

Ex: housing price prediction, bitcoin value prediction, scoring, etc.

- Classification: mapping input value to discrete or categorical values.

Ex: object detection, image segmentation, etc.

1.1 Prolems: Overfitting and Underfitting

For Overfitting:

* In Regression: Finding much more complex function than the given data.
* In Classification: Wrong prediction.

For Underfitting:

* In Regression: Finding much simplier function than the given data.
* In Classification: Predict right properties but wrong object.

1.2 Reasons

For Overfitting:

* Training data is no cleaned and contains noise.
* Complex model but small dataset.

For Underfitting:

* Lack of features
* The model is too simple
* Small size of dataset

1.3 Solutions

For Overfitting:

* Spend more time for pre-processing data
* Regularization (most used)
* Get/Generate more data

For Underfitting:

* Spend more time for feature engineering (ML)
* Increase model complexity
* Increase the training duration (# epochs)

2. Unsupervised Learning

- Learning on unlabeled data. The goal is to cluster the similar examples to classes.

Ex: segmentation, recommendation system, etc.

3. Reinforcement Learning

* Neural network
* Policy network
* Mechanism
* Reward shaping

Ex: self-driving car, gaming, etc.

V. Deep Learning

Unit = neuron ?

1. Deep neural network

Input layer -> hidden layer (1 -> n) -> output layer.