Intro

Machine Learning

Microsoft Student Partner
백진헌

Overview

1. Concept of Machine Learning
2. Supervised Learning
3. Unsupervised Learning

Concept of Machine Learning

✓Machine Learning?

“ Filed of study that gives computers the ability to learn without being explicitly programmed.” , Arthur Samuel (1959)

Concept of Machine Learning

✓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, improves with experience E.”, Tom Mitchell (1997)

Concept of Machine Learning

✓Machine Learning?

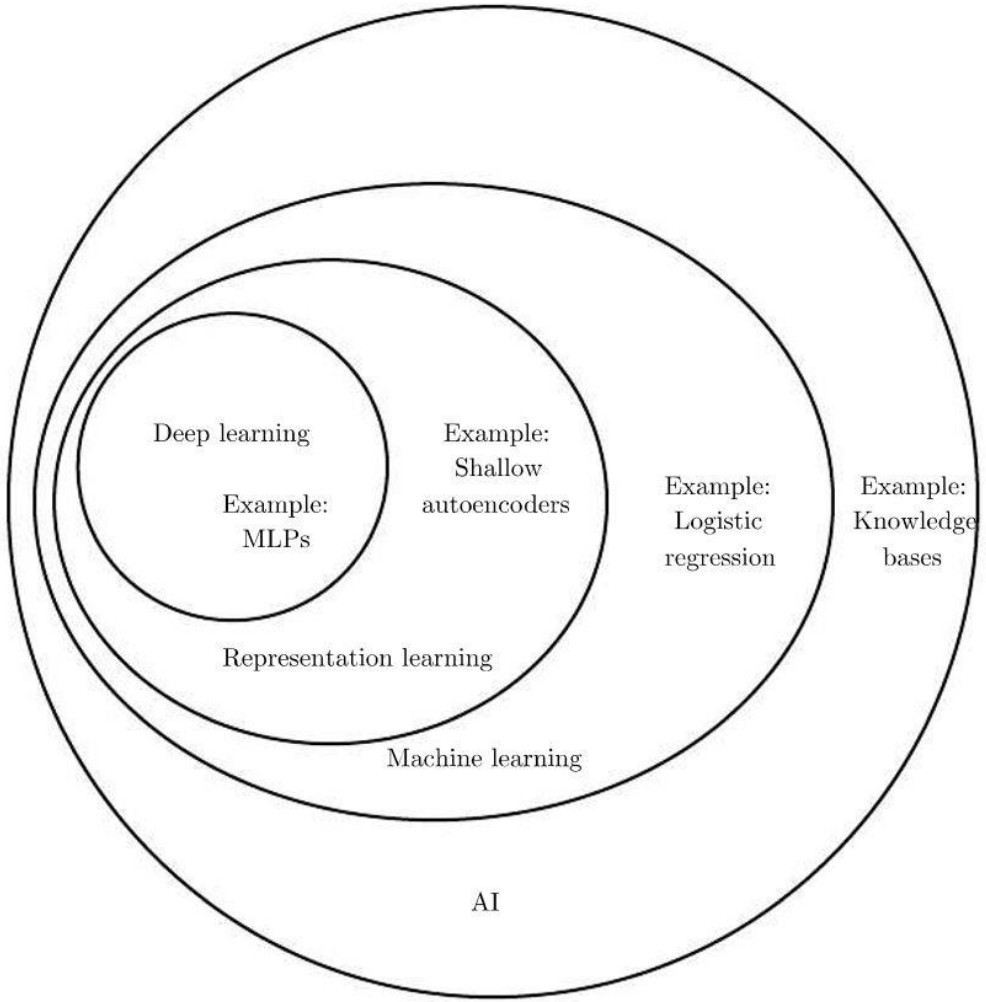
Ex) Spam mail filter,

E(Experience): Watching you label emails as spam or not spam.

T(Task): Classifying emails as spam or not spam.

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

Concept of Machine Learning



Concept of Machine Learning

✓AI

“ How to get this informal knowledge into a computer !”

Concept of Machine Learning

✓AI

“ How to get this informal knowledge into a computer !”

Knowledge base approach

Concept of Machine Learning

✓AI

None of these projects has led to a major success.

Concept of Machine Learning

✓Machine Learning

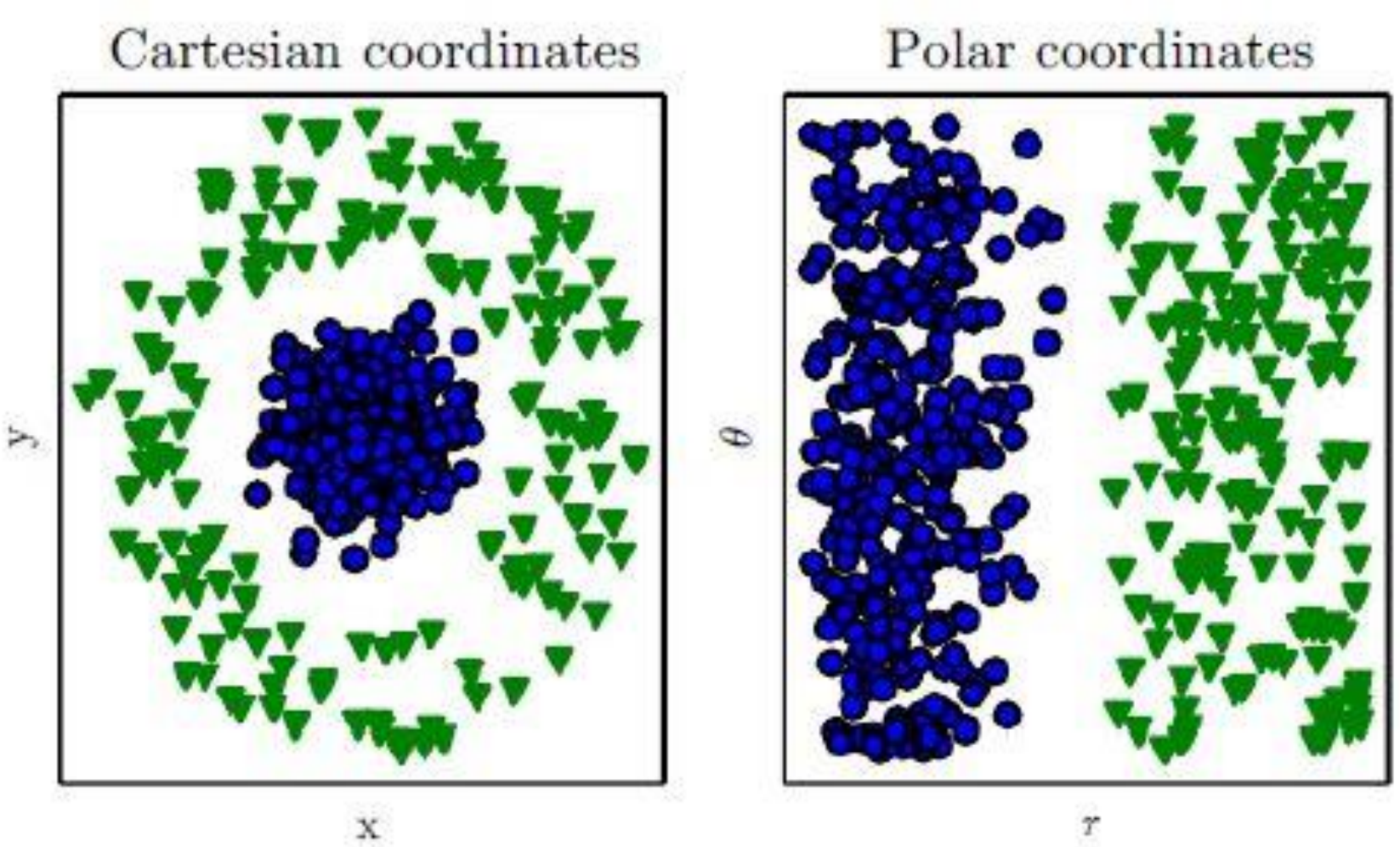
“Extract patterns from raw data.”

Concept of Machine Learning

✓Machine Learning

"Difficult to know what features should be extracted."

Concept of Machine Learning



Concept of Machine Learning

✓Representation Learning

"Discover not only the mapping from representation to output
but also the representation itself."

Concept of Machine Learning

✓**Representation Learning - Autoencoder**

Autoencoder: encoder + decoder

Encoder: input data into a different representation
Decode: new representation back into the original format

Concept of Machine Learning

✓Representation Learning

“What if difficult to obtain a representation”

Concept of Machine Learning

✓Deep Learning

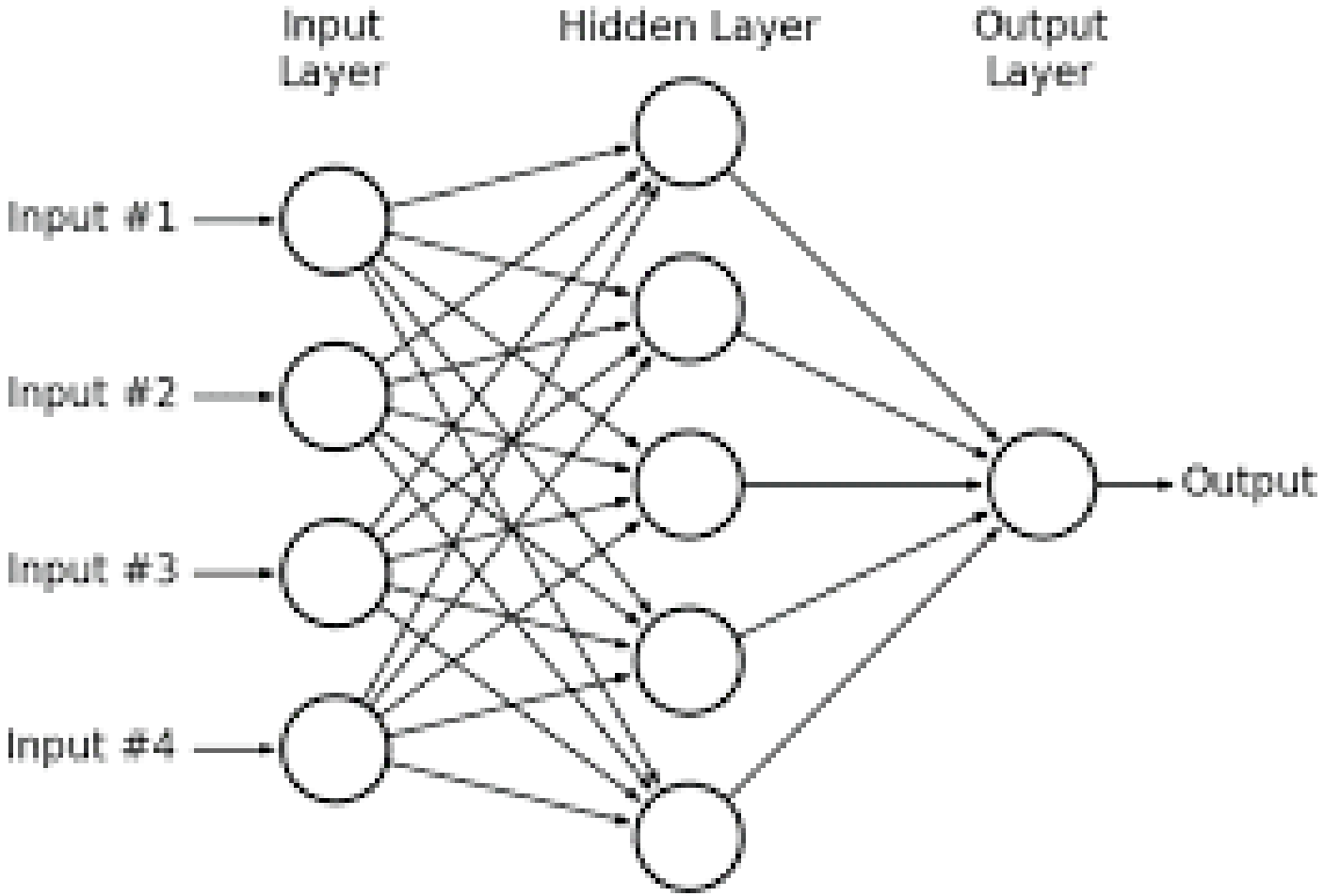
“Build complex concepts out of simpler concepts(representation)”

Concept of Machine Learning

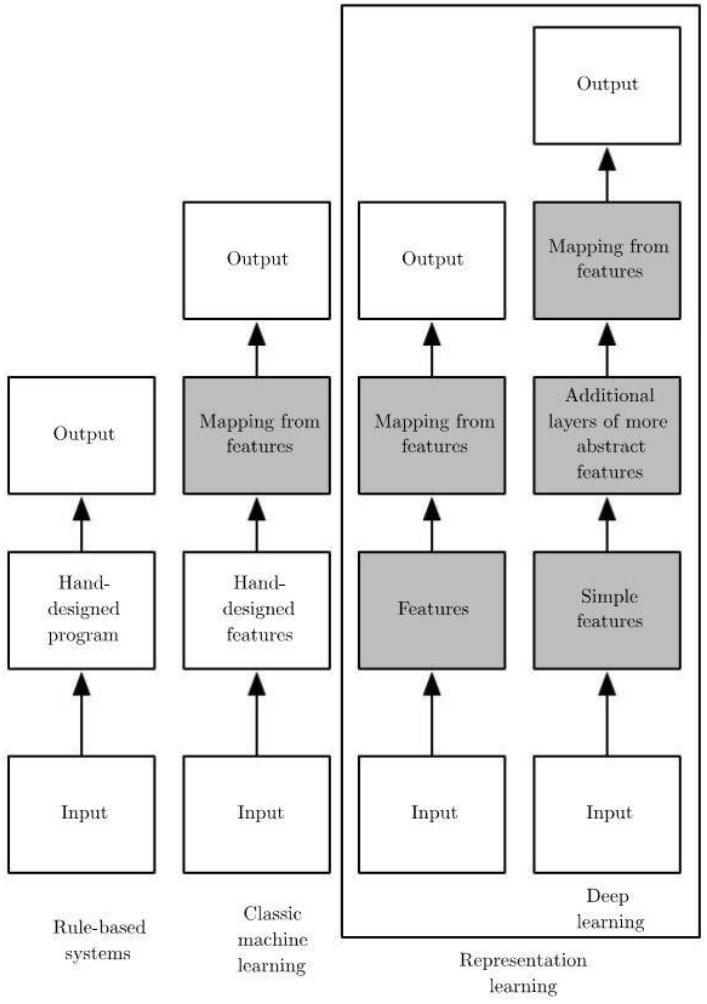
✓**Deep Learning - MLP (multilayer perceptron)**

“Each function provides a new representation”

Concept of Machine Learning



Concept of Machine Learning



Learning

✓**Supervised Learning**

: Experience a dataset containing features, but each example is also associated **with a label or target.**

✓**Unsupervised Learning**

: Experience a dataset containing many features, then learn useful properties of the **structure of this dataset.**

Supervised Learning

✓Supervised Learning

: Regression (Continuous Value)

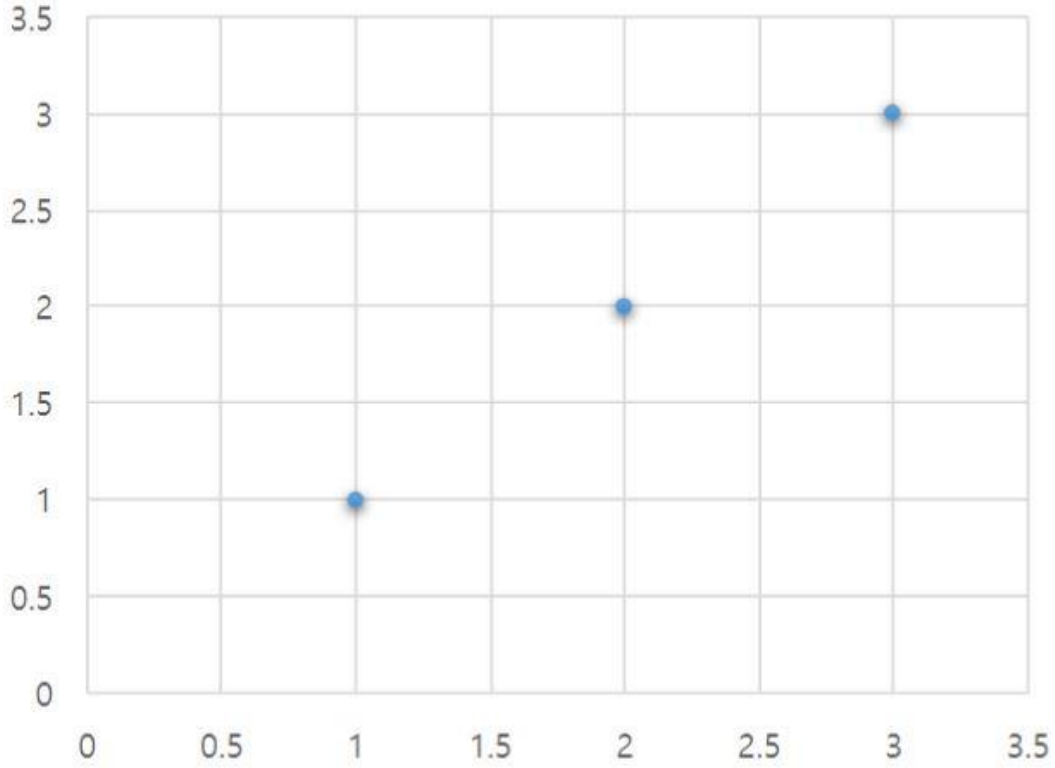
: Classification (Discrete Value)

In Deep Learning? CNN / RNN

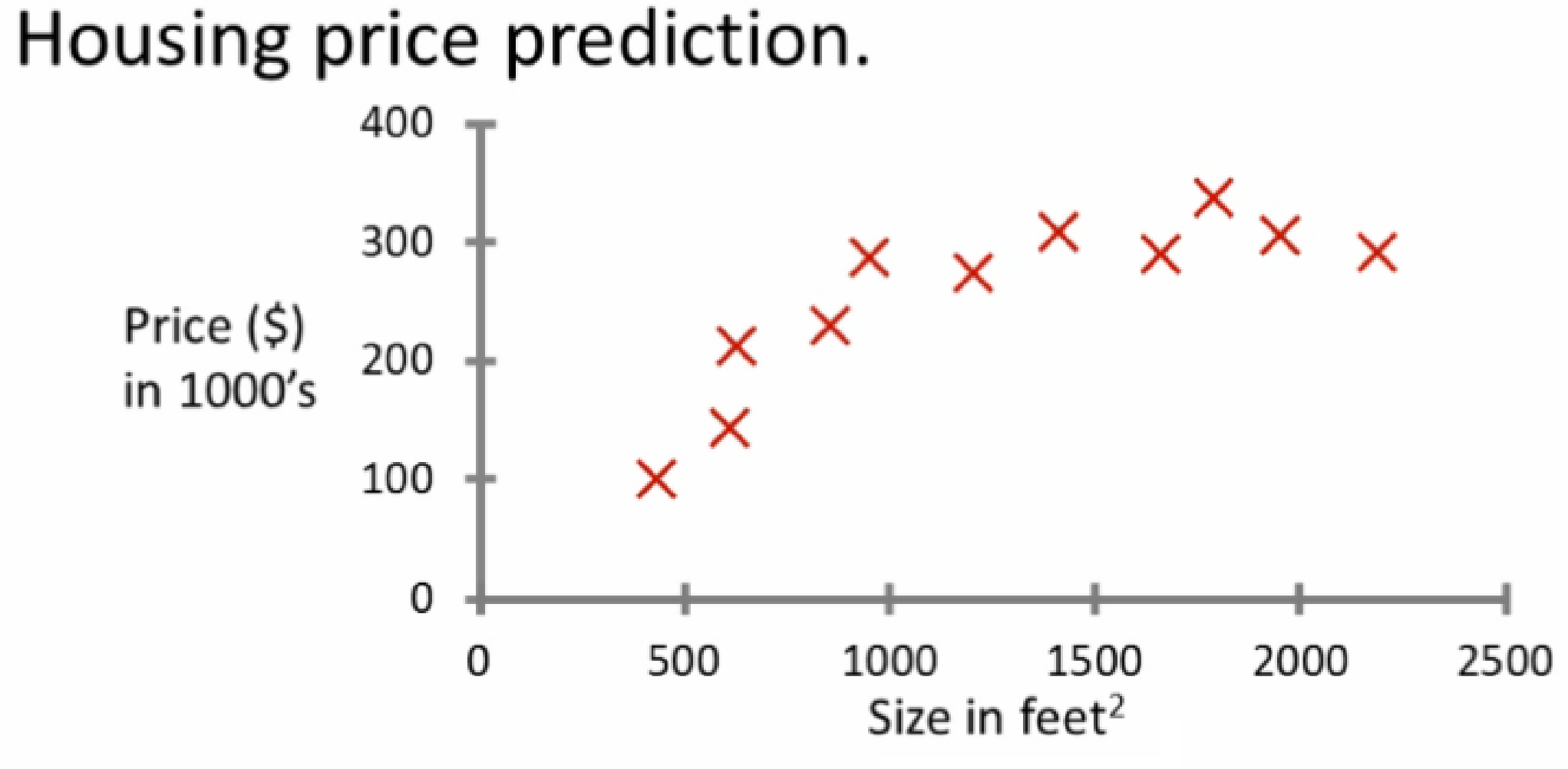
Supervised Learning

✓Supervised Learning - Regression

x	y
1	1
2	2
3	3

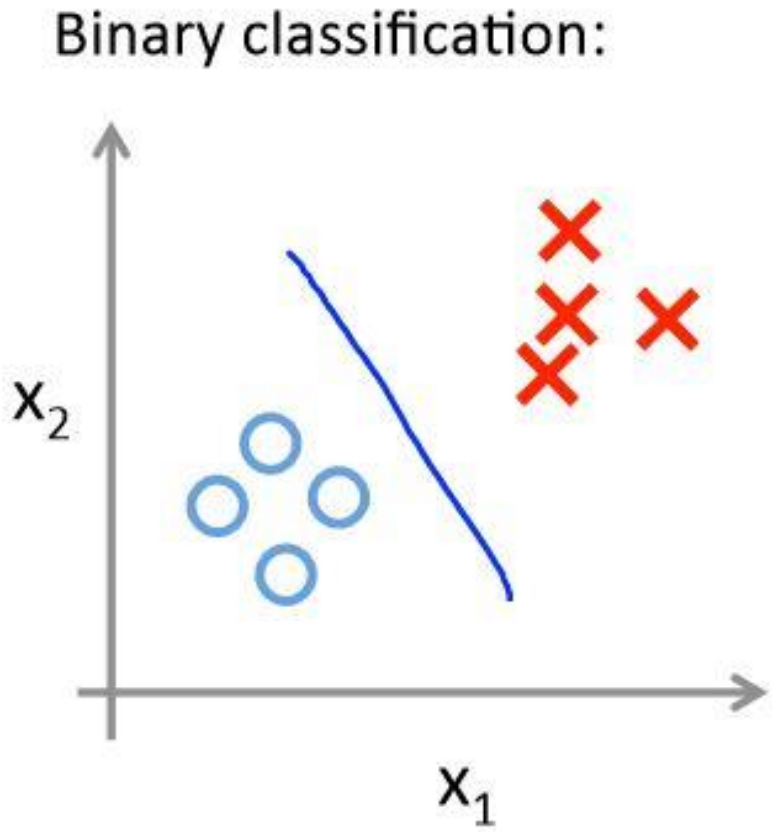


Supervised Learning



Supervised Learning

✓Supervised Learning - Classification



Unsupervised Learning

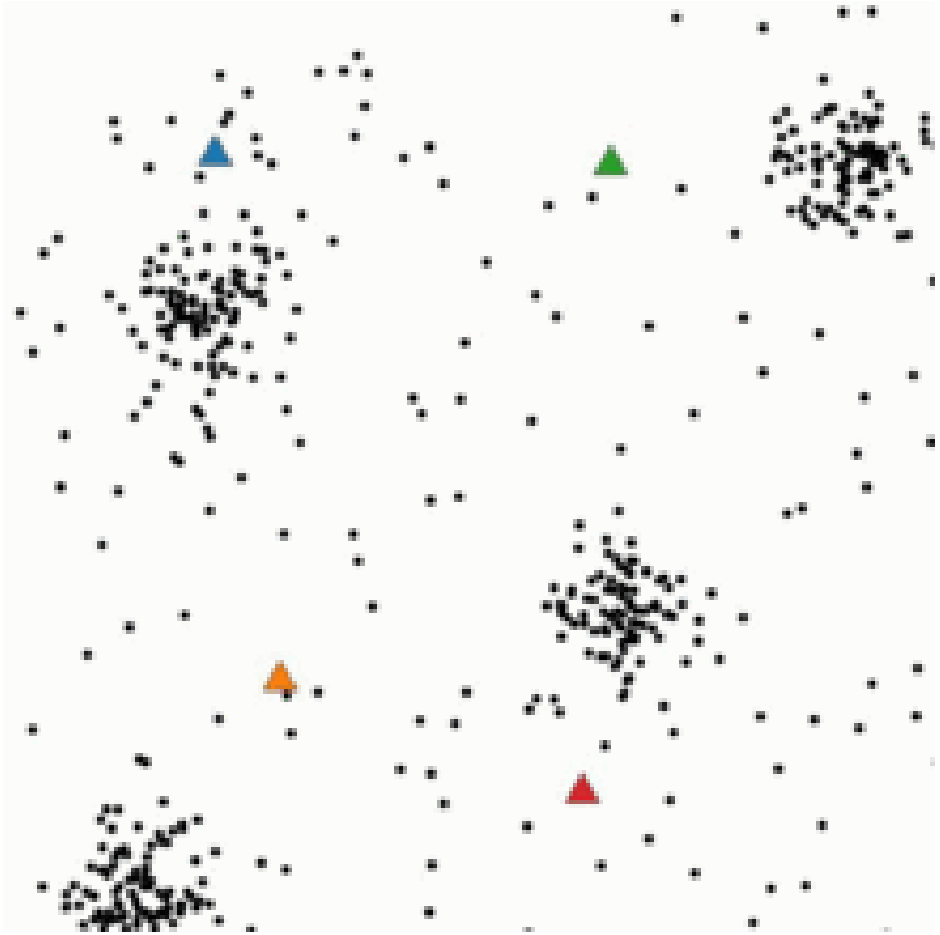
✓Unsupervised Learning

: Clustering

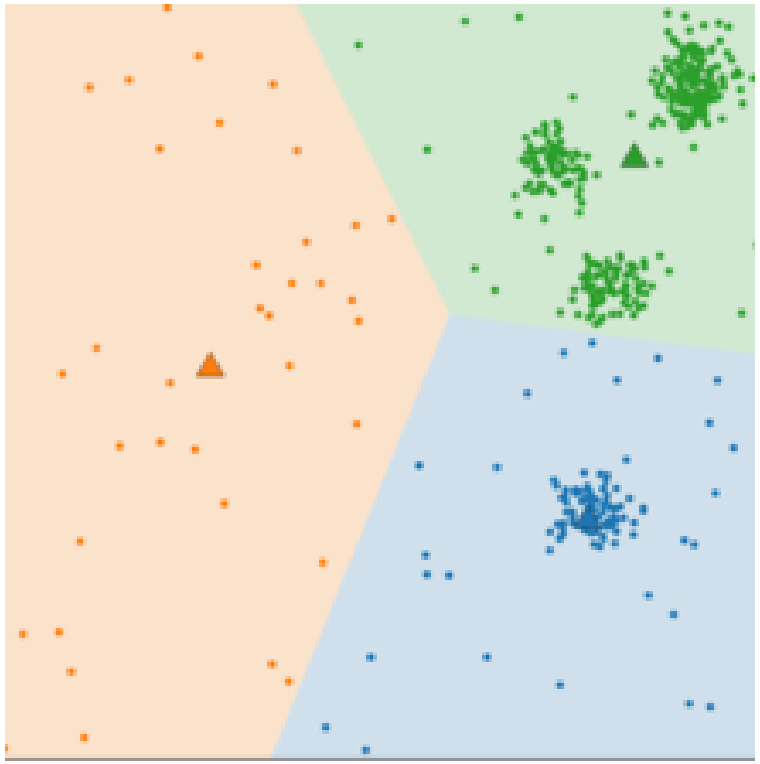
In Deep Learning? AutoEncoders

Learning

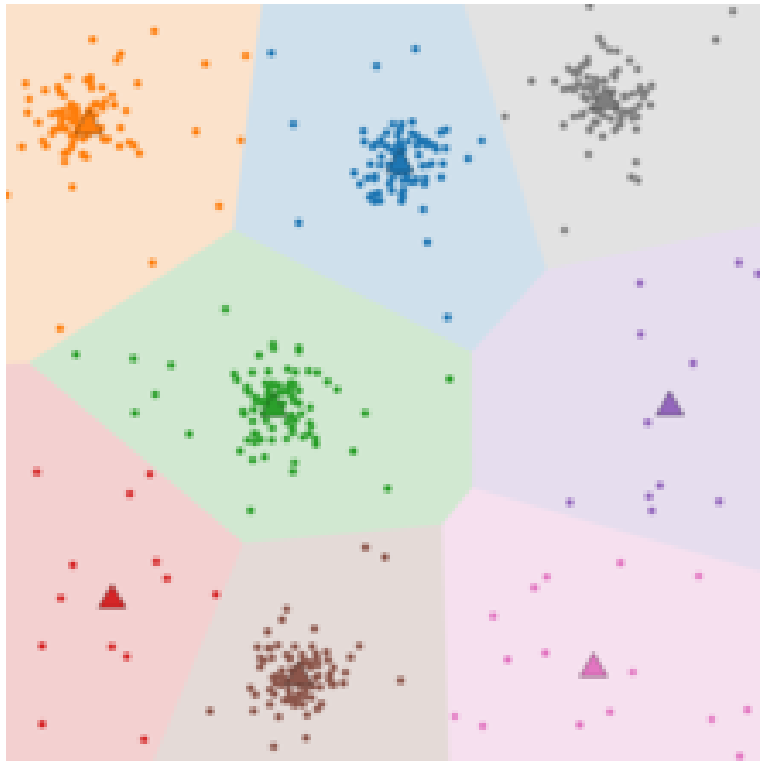
✓Unsupervised Learning - Clustering



Learning

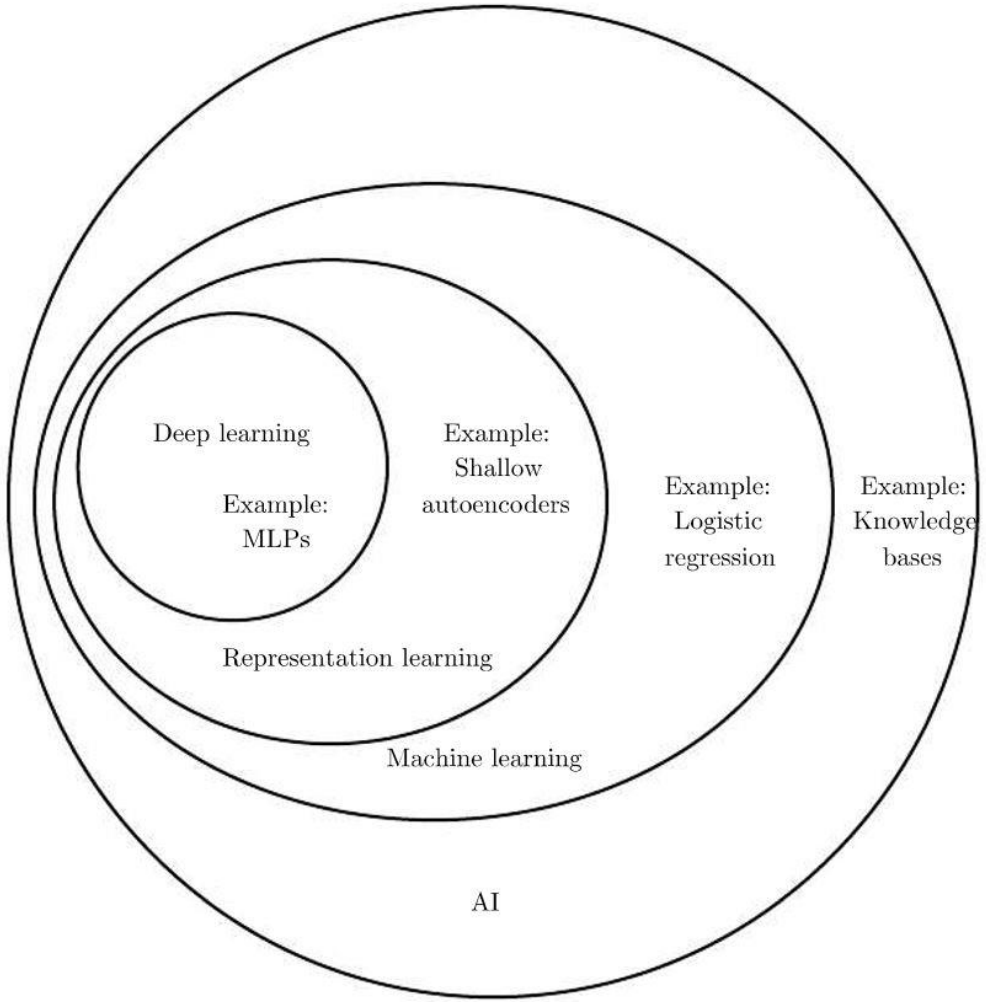


K = 3



K = 8

Warm Up


Deep learning
Example:
MLPs
Example:
Shallow
autoencoders
Example:
Logistic
regression
Example:
Knowledge
bases
Representation learning
Machine learning
AI

Warm Up

- ✓Supervised Learning
- ✓Unsupervised Learning

Supervised Learning & HOL

Machine Learning

Microsoft Student Partner
백진헌

Overview

1. **Linear Regression**
   (HOL with Azure Machine Learning)

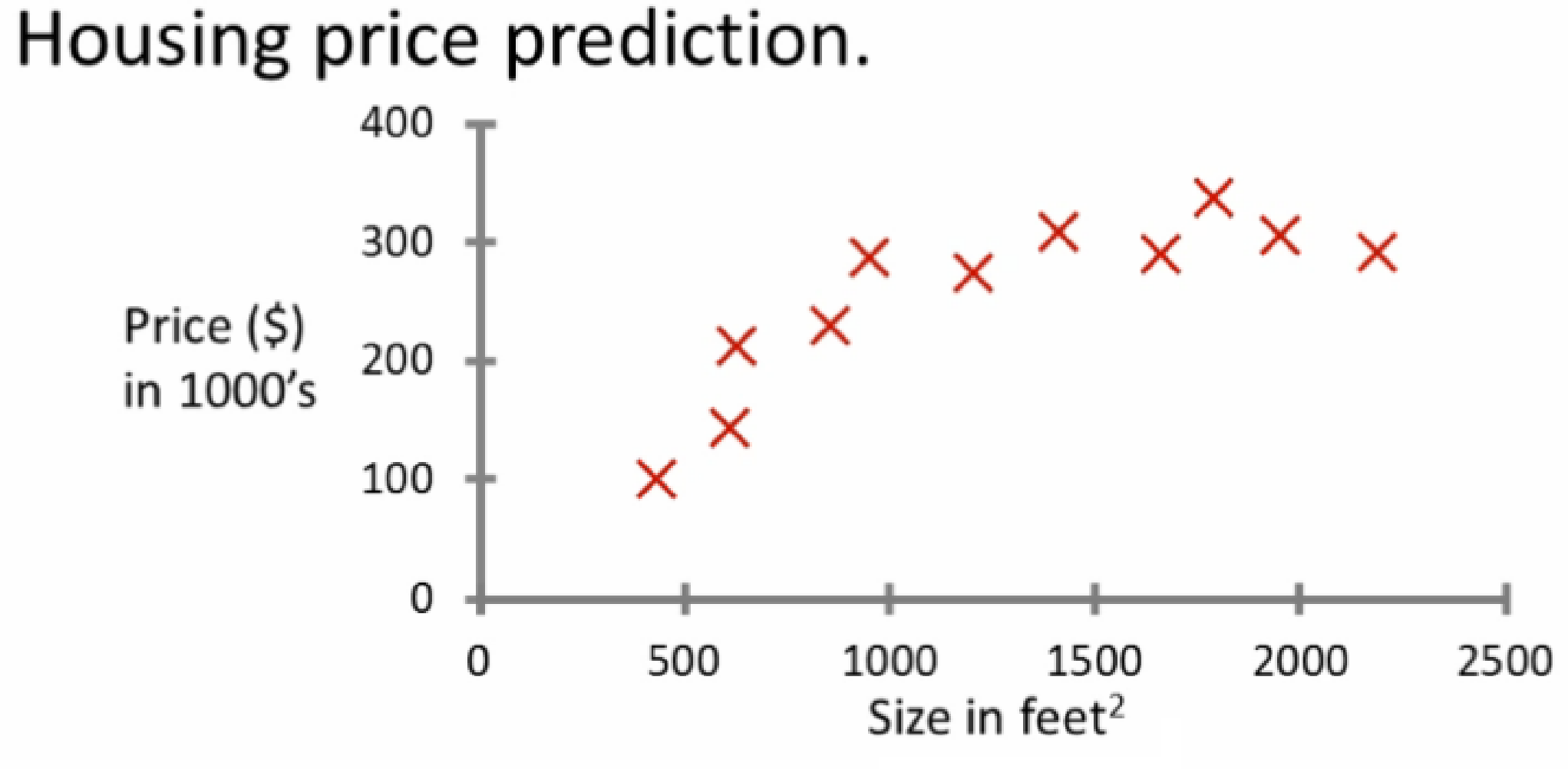
2. **Neural Network**
   (HOL with Azure Notebook)

Supervised Learning

[Machine Learning]

✓Regression
: Linear Regression

Supervised Learning

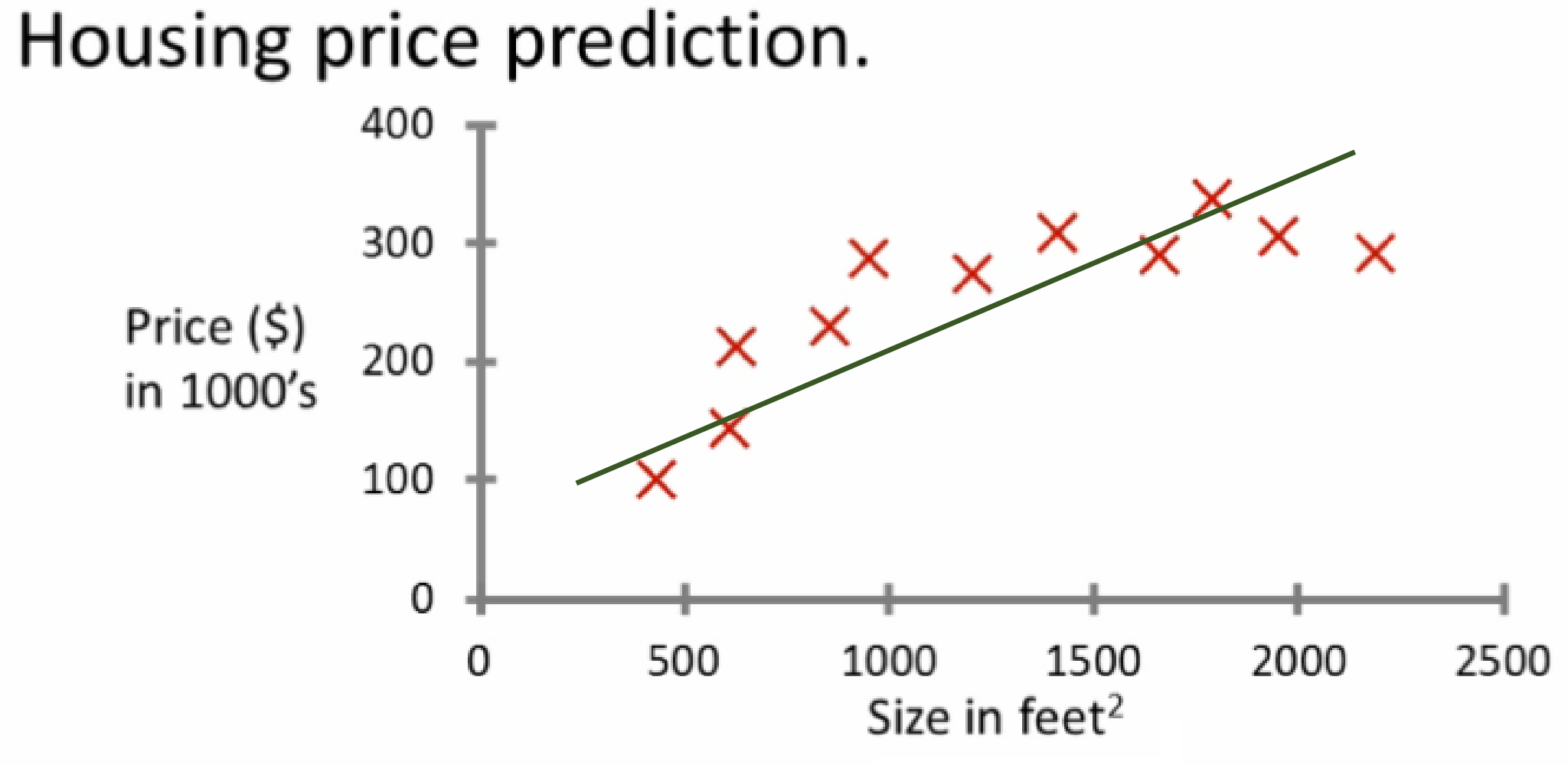


Supervised Learning

$$Y = W*X + b$$

X : Y를 예측할 때 필요한 유의미한 피처 (집의 크기)
W, b : Y를 예측할 때 사용되는 각 피처에 대한 무게 값
Y : 학습된 모델을 통해 예측한 값

Supervised Learning



Supervised Learning

$$Y = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$$

X : Y를 예측할 때 필요한 유의미한 피처 (집의 크기, 집의 층 수 등)
W, b : Y를 예측할 때 사용되는 각 피처에 대한 무게 값
Y : 학습된 모델을 통해 예측한 값

Supervised Learning

$$Y = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$$

그럼 우리가 학습을 통해 계산해야 하는 것은?

Supervised Learning

$$Y = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$$

계산해야 하는 W(weight) 는 어떻게 계산할까 ?

Supervised Learning

$$Y_{(predict)} = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$$

Y_predict : 학습된 모델을 통해 예측한 값

Y : 실제 값 (Label)

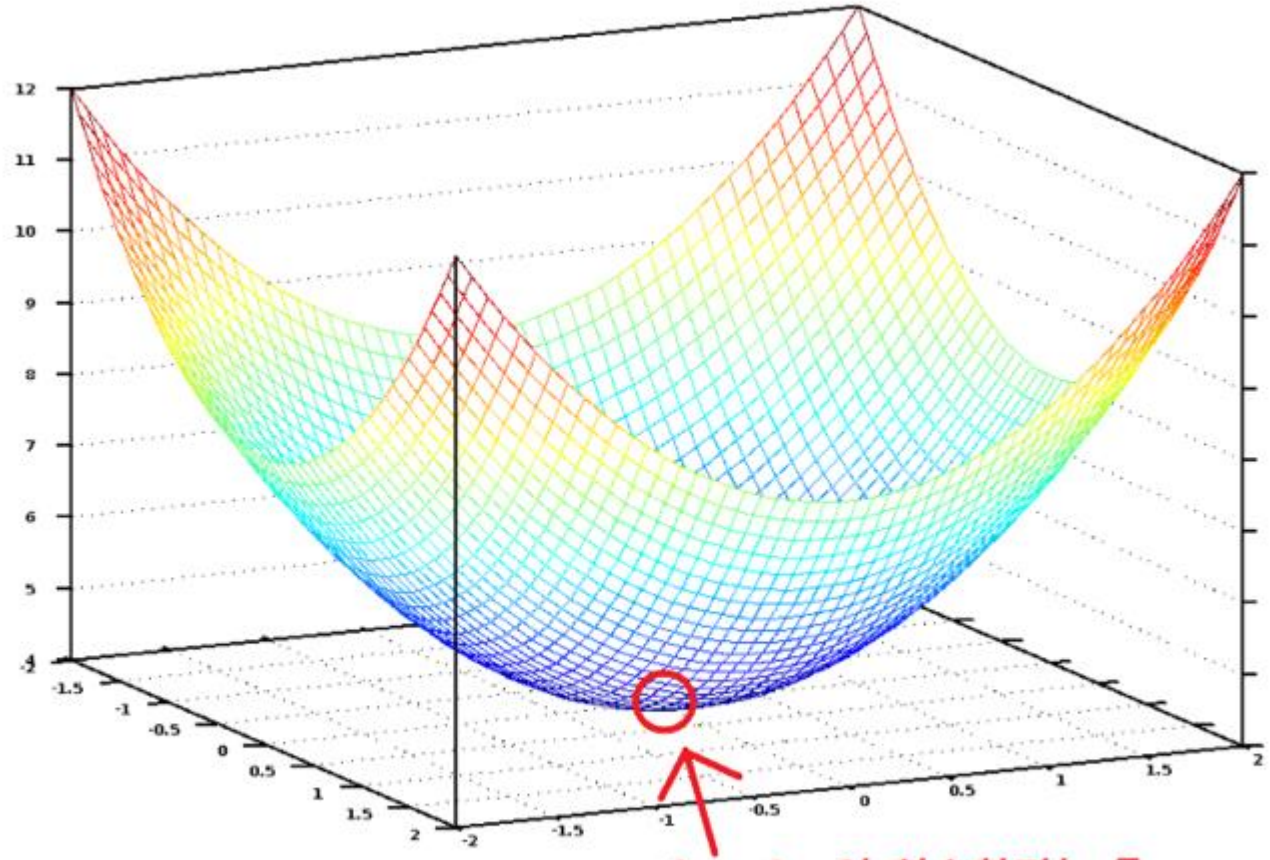
Supervised Learning

$$Y_{(predict)} = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$$

$$\text{Cost Function} : (Y - Y_{(predict)}) \text{ ^ } 2$$

Supervised Learning

✓Graph of Cost Function



Supervised Learning

- ✓Gradient, Gradient Descent

Supervised Learning

✓Derivative

$f'(x) > 0$
: positive direction -> increase

$f'(x) < 0$
: negative direction -> increase

✓Partial Derivative

Suppose $f: \mathbb{R}^n \to \mathbb{R}$

The *partial* derivative of $f(x_1, \dots, x_n)$ at x_i

$$\frac{\partial}{\partial x_i} f(x_1, \dots, x_n) = \lim_{h \to 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{h}$$

✓Partial Derivative

Suppose $f: \mathbb{R}^n \to \mathbb{R}$

The *partial* derivative of $f(x_1, \dots, x_n)$ at x_i

$$\frac{\partial}{\partial x_i} f(x_1, \dots, x_n) = \lim_{h \to 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{h}$$

$\frac{\partial f}{\partial x_i}$ 방향으로 가면, x_i 축에서 증가하는 방향으로 향한다.

✓Gradient

Suppose $f: \mathbb{R}^n \rightarrow \mathbb{R}$

Gradient $\nabla f(\boldsymbol{x})$

$$\nabla f = \left[\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_n}\right]$$

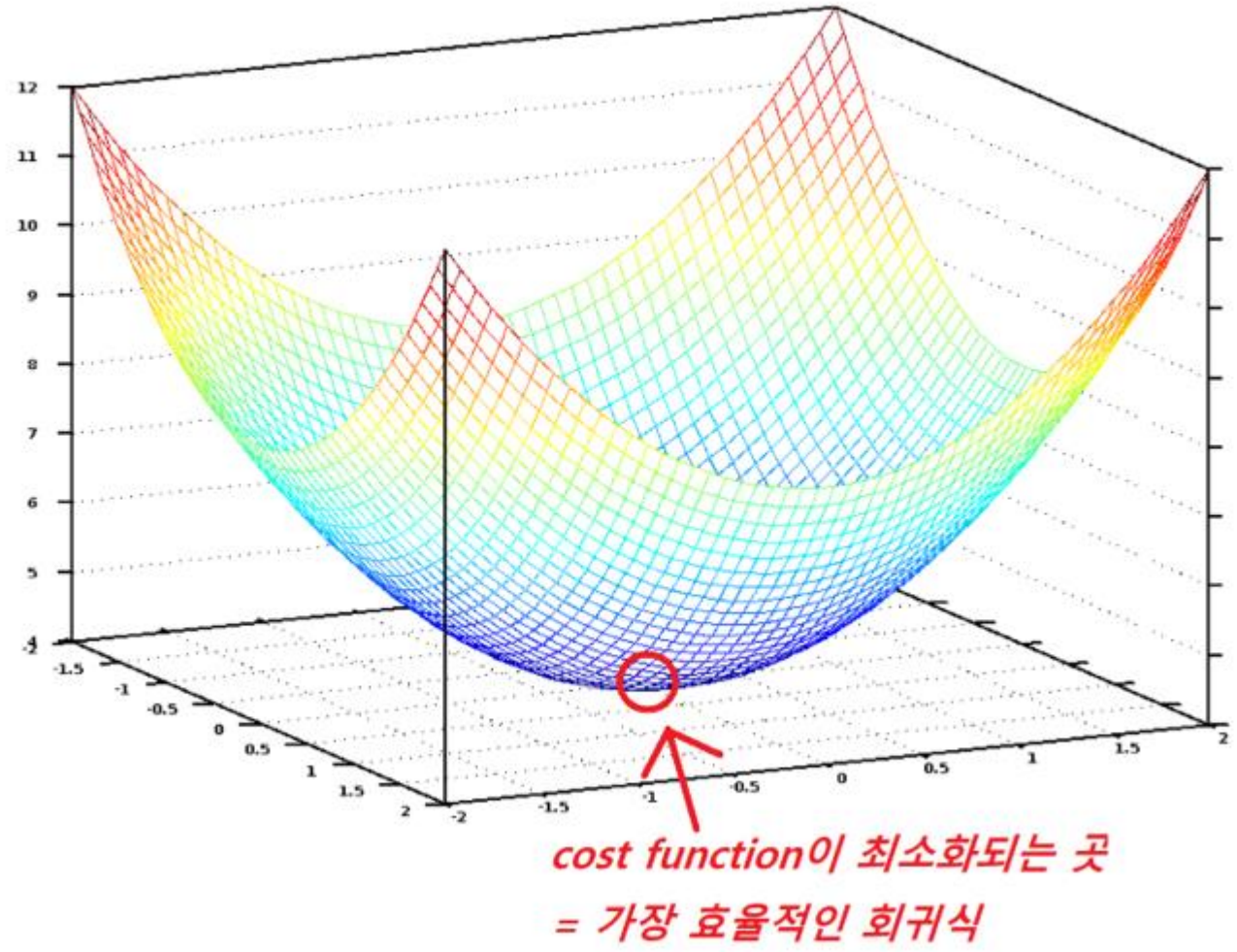
- A vector containing the partial derivatives of the function f at certain point $\boldsymbol{x}$
- Direction of **steepest ascent** (i.e., maximum increase of the function f) from $\boldsymbol{x}$
  - Thus, $-\nabla f$ is the direction of steepest descent (maximum decrease)

Supervised Learning

$Y_{(predict)} = (W1*X1) + (W2*X2) + \cdots + (Wn*Xn) + b$

Cost Function : $(Y - Y_{(predict)}) \wedge 2$

✓Gradient Descent



Supervised Learning

✓Stochastic Gradient Descent

"Gradient Descent is computationally expensive"

Supervised Learning

✓HOL

Automobile Price Data

Supervised Learning

✓HOL

“Azure Machine Learning Studio”

(https://studio.azureml.net/)

Supervised Learning

✓HOL

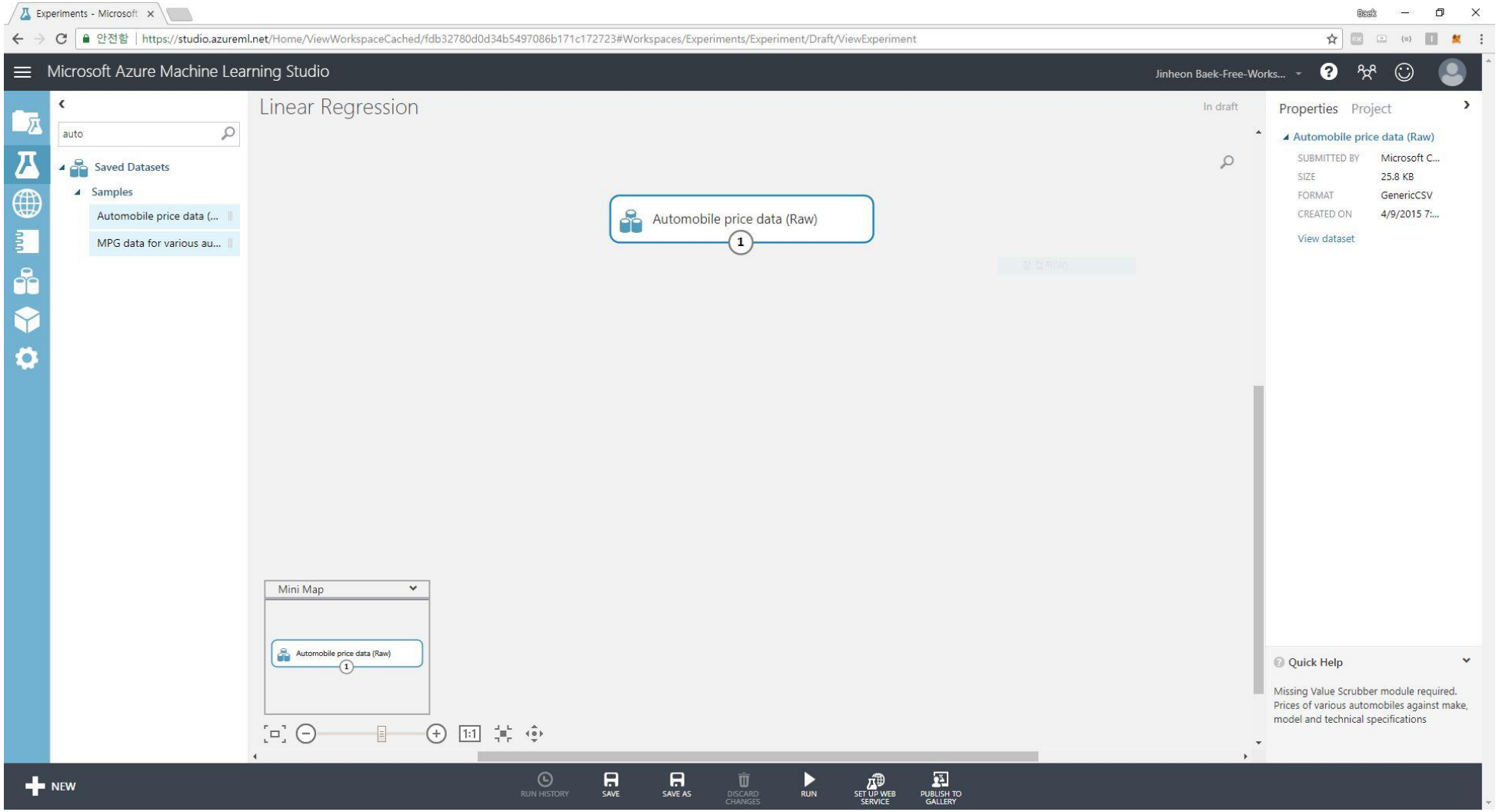
1) New

2) EXPERIMENT

3) Blank Experiment

Supervised Learning

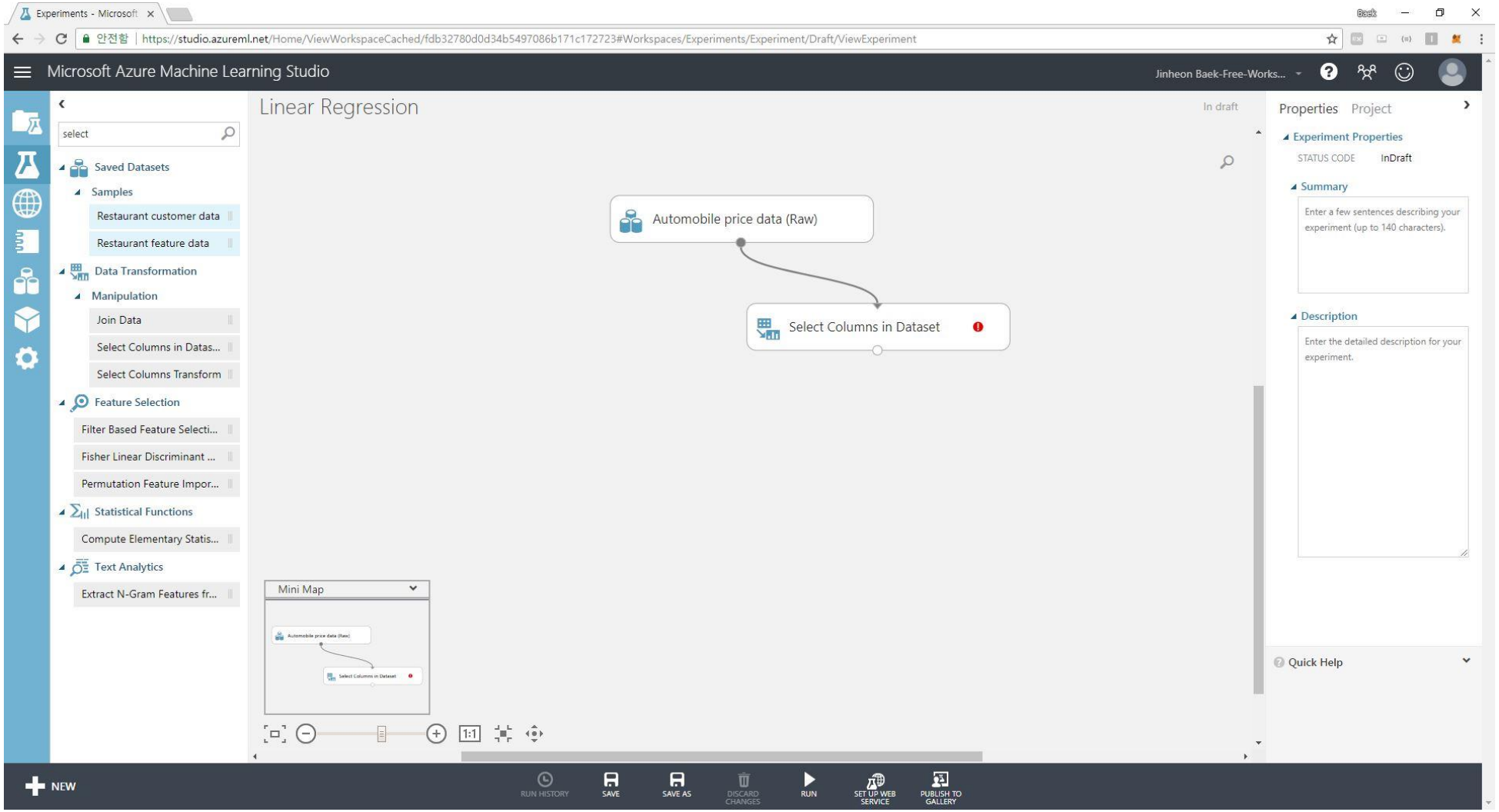
검색창: Automobile price data 선택



Supervised Learning

검색창: Select Columns in Dataset 선택

Automobile price data –〉 Select Columns ··· 이어주기

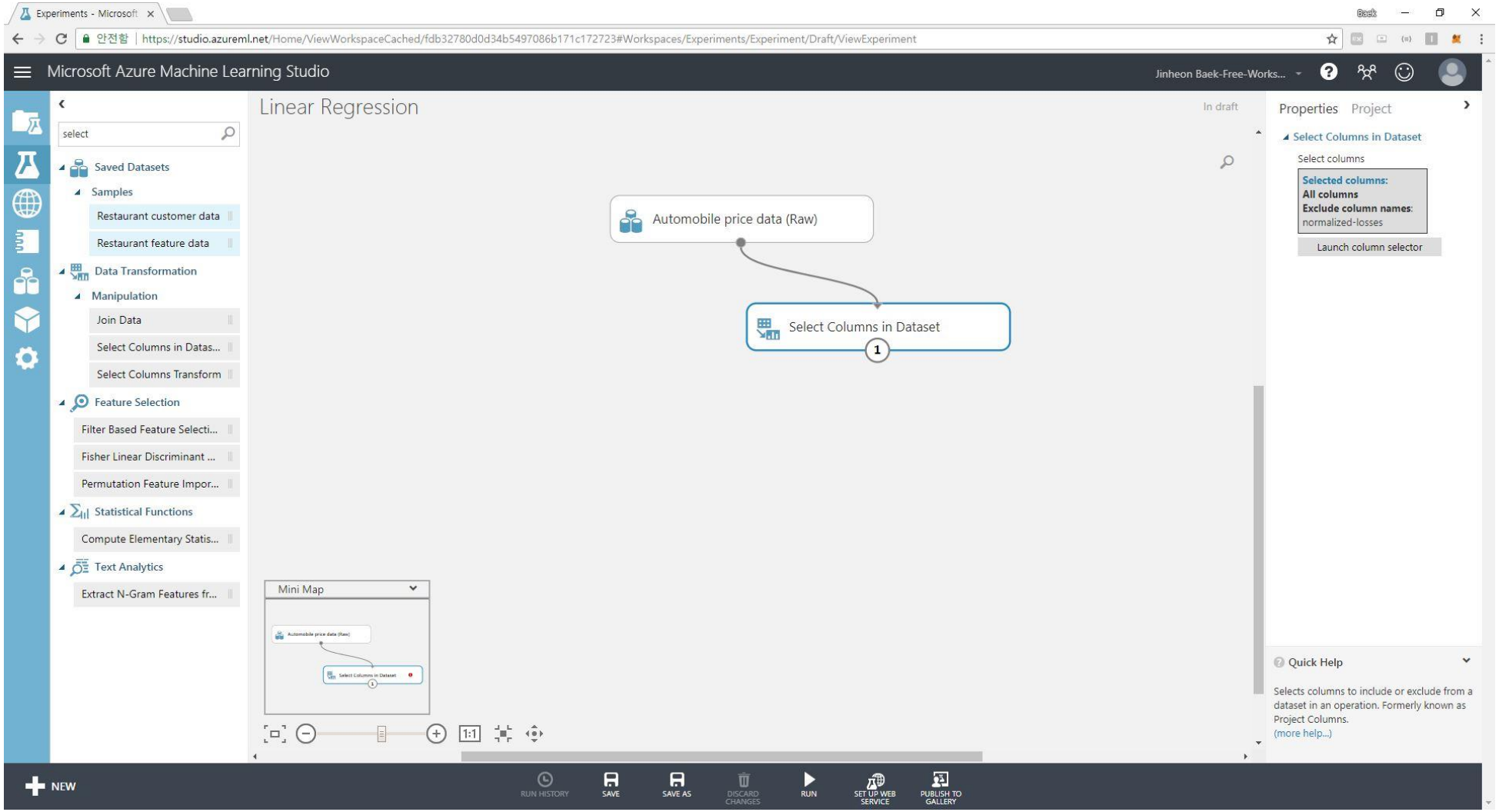


Supervised Learning

Select Columns in Dataset 클릭 (오른쪽 메뉴 확인)

Launch column selector 클릭

With rules –〉 All columns –〉 Exclude –〉 normalized–losses (column names)

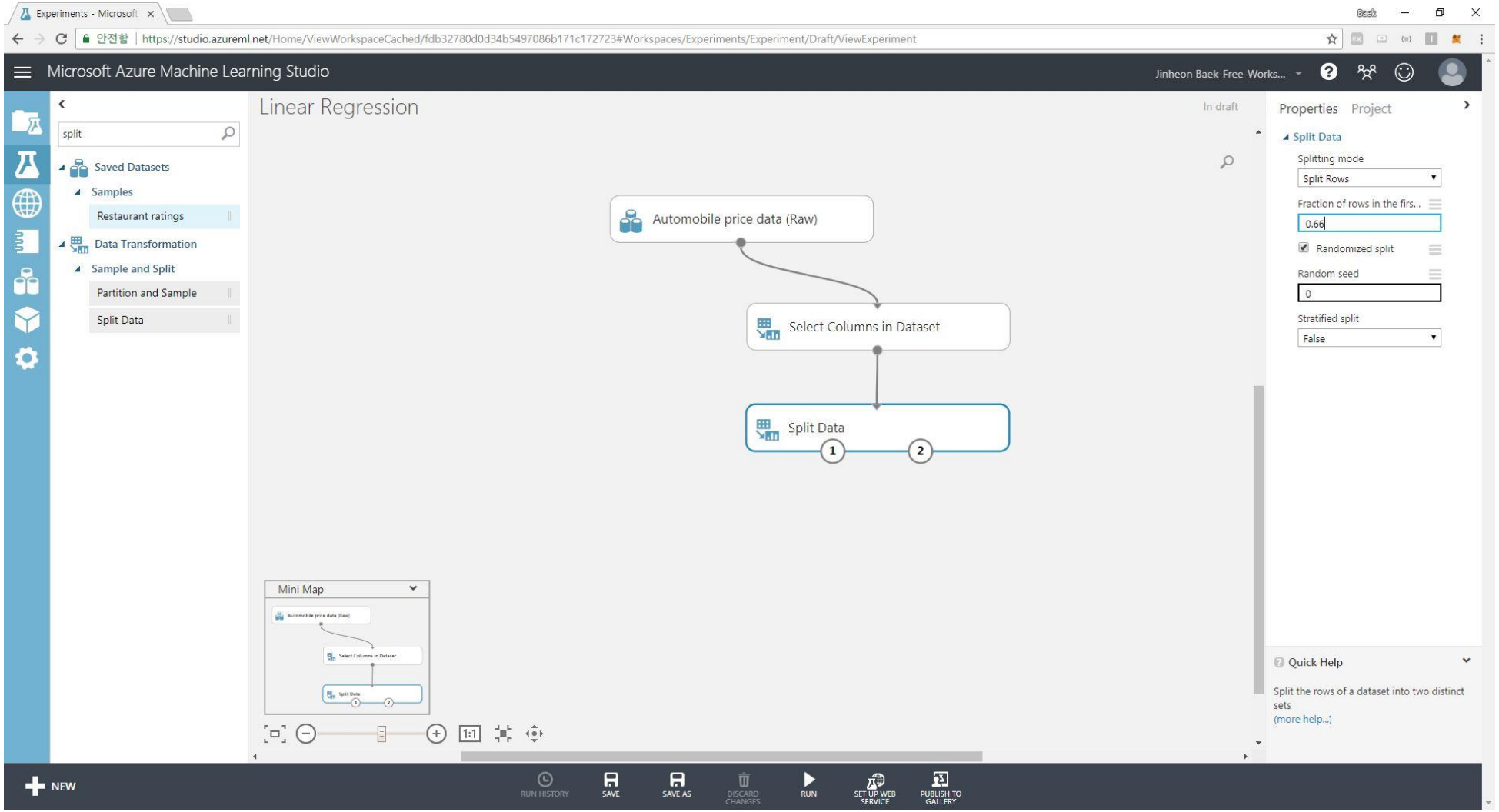


Supervised Learning

검색창: Split Data

Select Columns in Dataset –〉 Split Data (연결)

Split Data 오른쪽 메뉴 –〉 Fraction of rows … (0.66 설정)

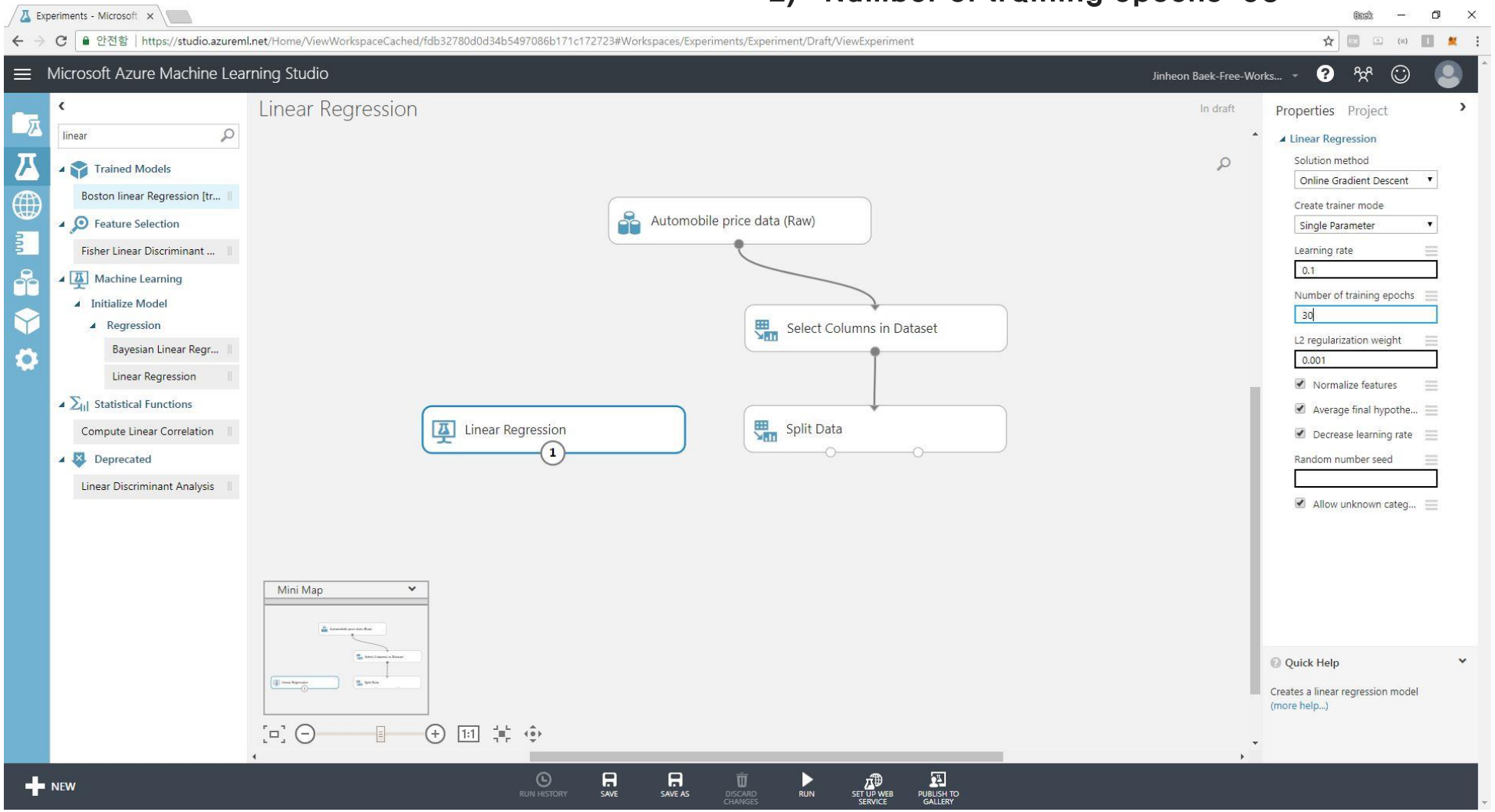


Supervised Learning

검색창: Linear Regression

Linear Regression 오른쪽 메뉴 –〉

1) Solution Method: Online Gradient Descent
2) Number of training epochs: 30

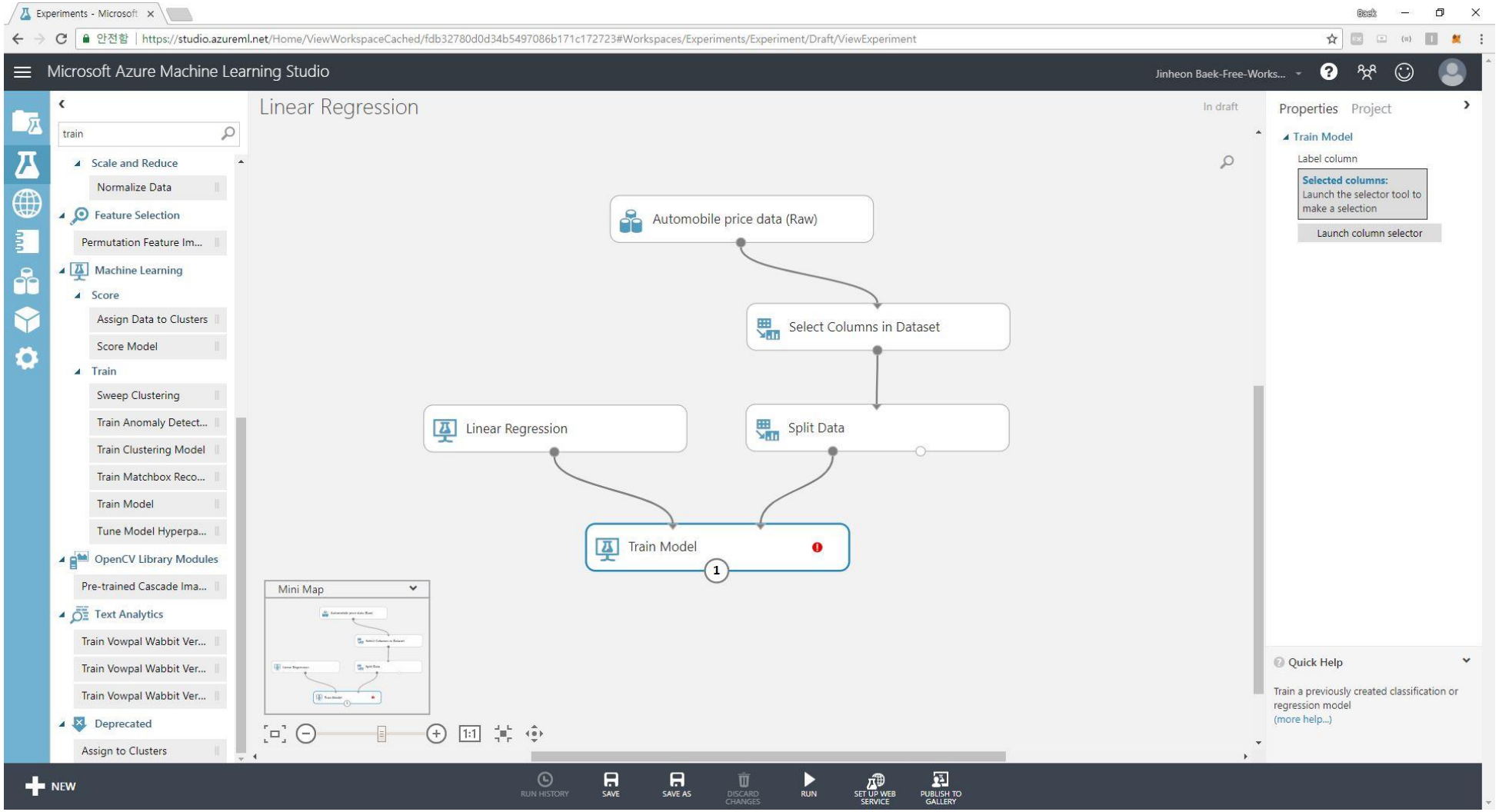


Supervised Learning

검색창: Train Model

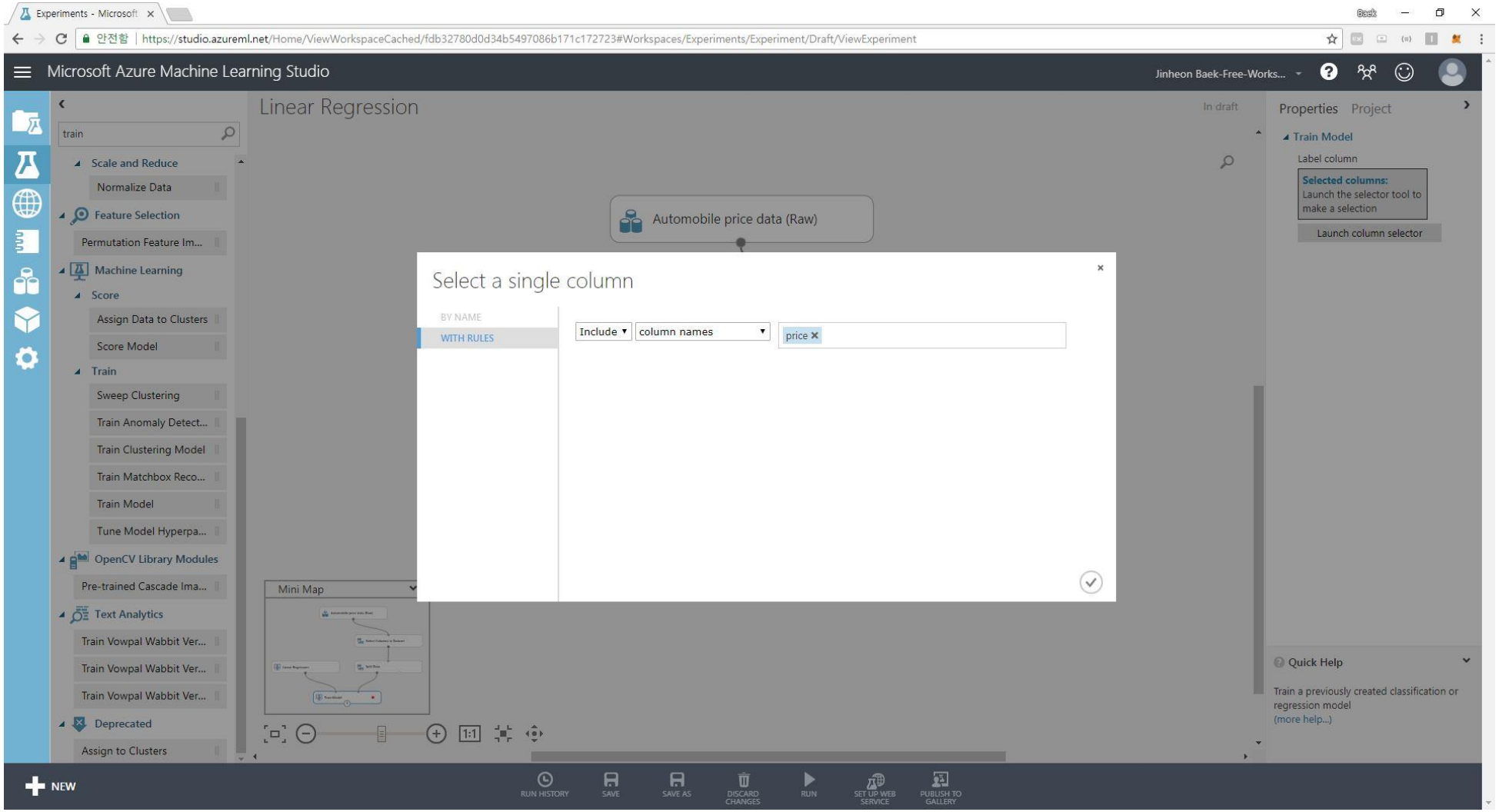
Linear Regression –〉 Train Model (왼쪽 연결)

Split Data –〉 Train Model (오른쪽 연결)



Supervised Learning

Train Model 오른쪽 메뉴 –〉 Launch column selector
With Rules –〉 Include –〉 column names –〉 price

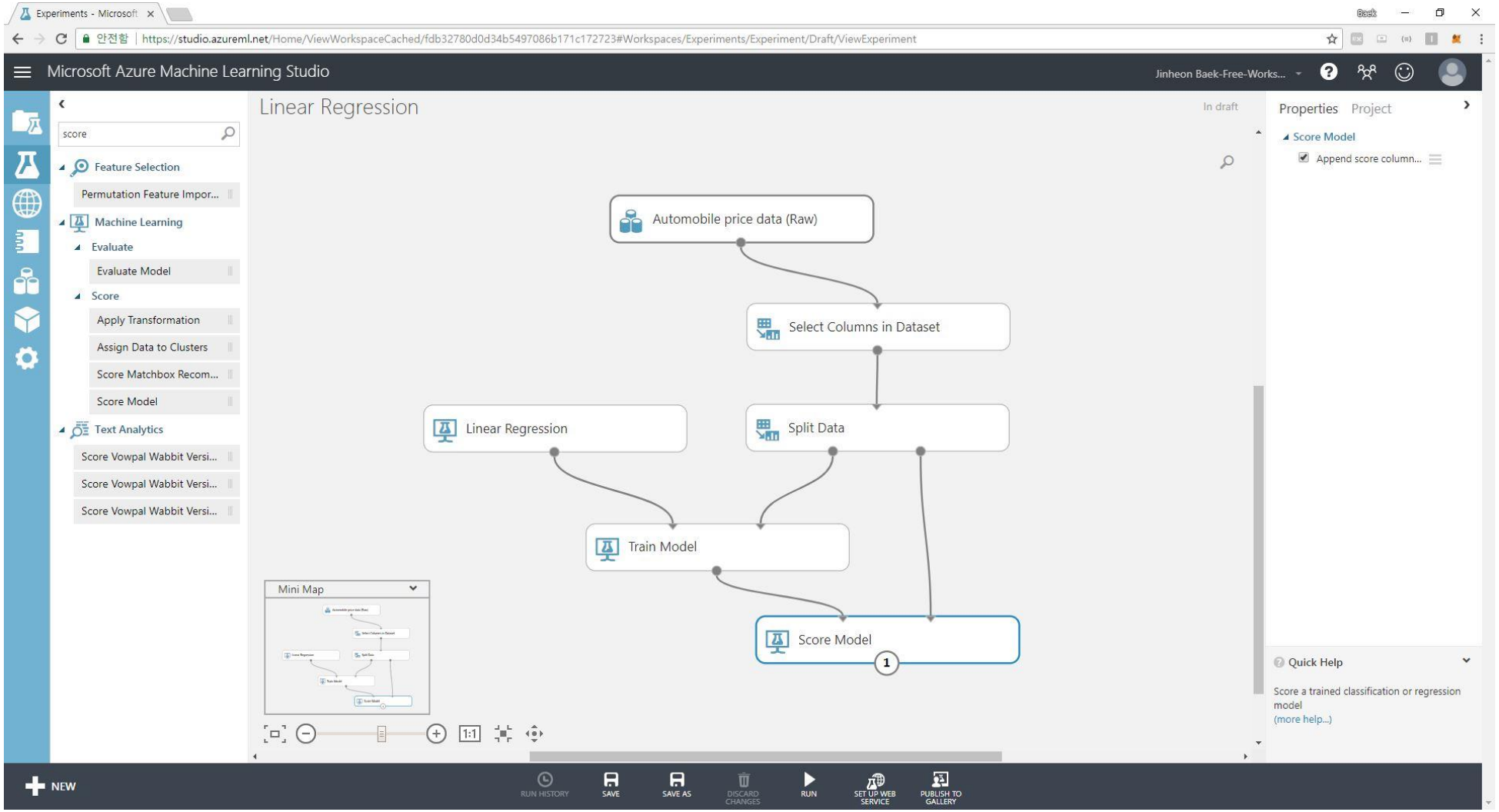


Supervised Learning

검색창: Score Model

Train Model –〉 Score Model (왼쪽 연결)

Split Data –〉 Score Model (오른쪽 연결)

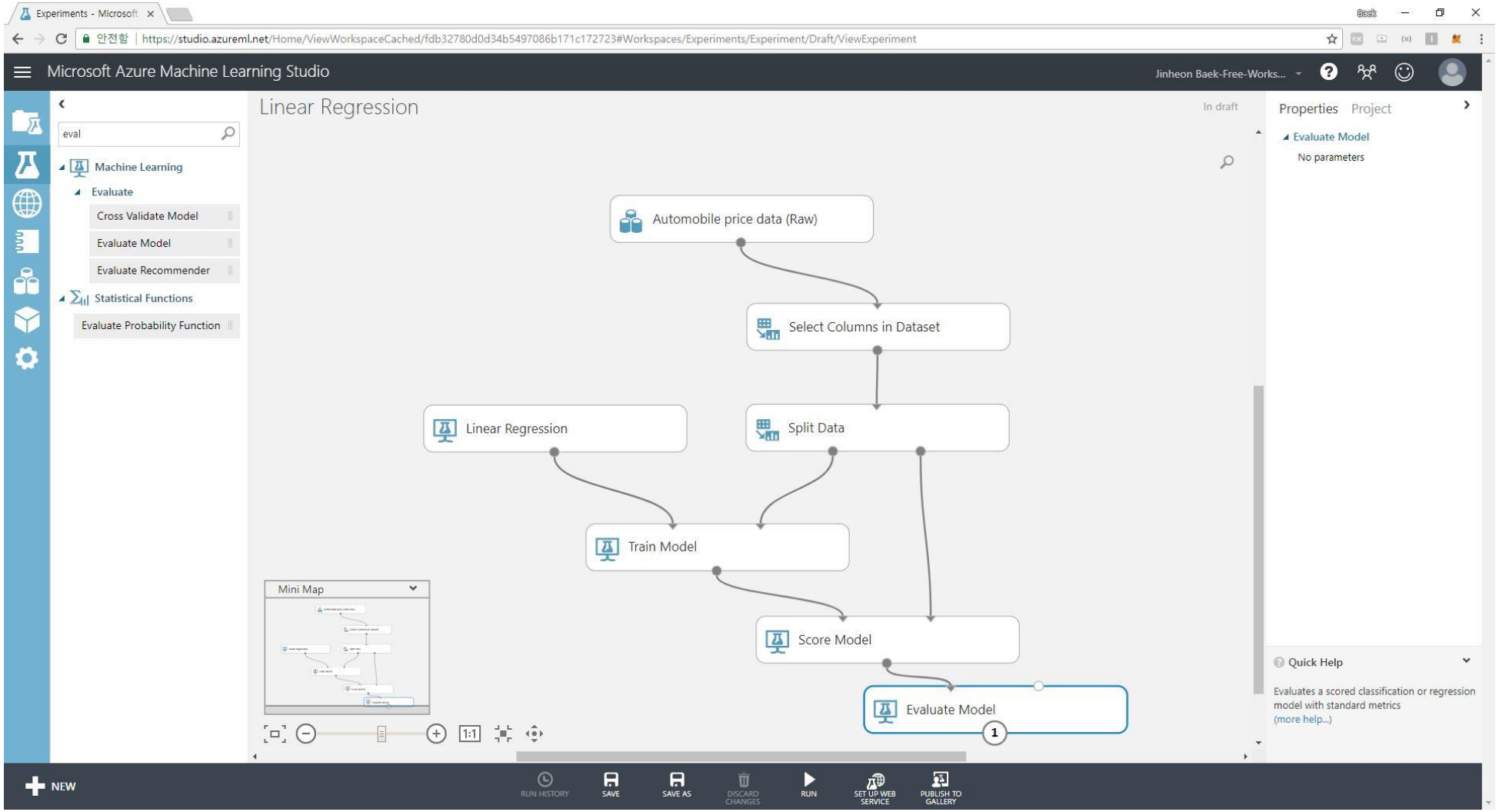


Supervised Learning

검색창: Evaluate Model

Score Model –〉 Evaluate Model (왼쪽 연결)

하단 Run 실행 후 결과 보기 !



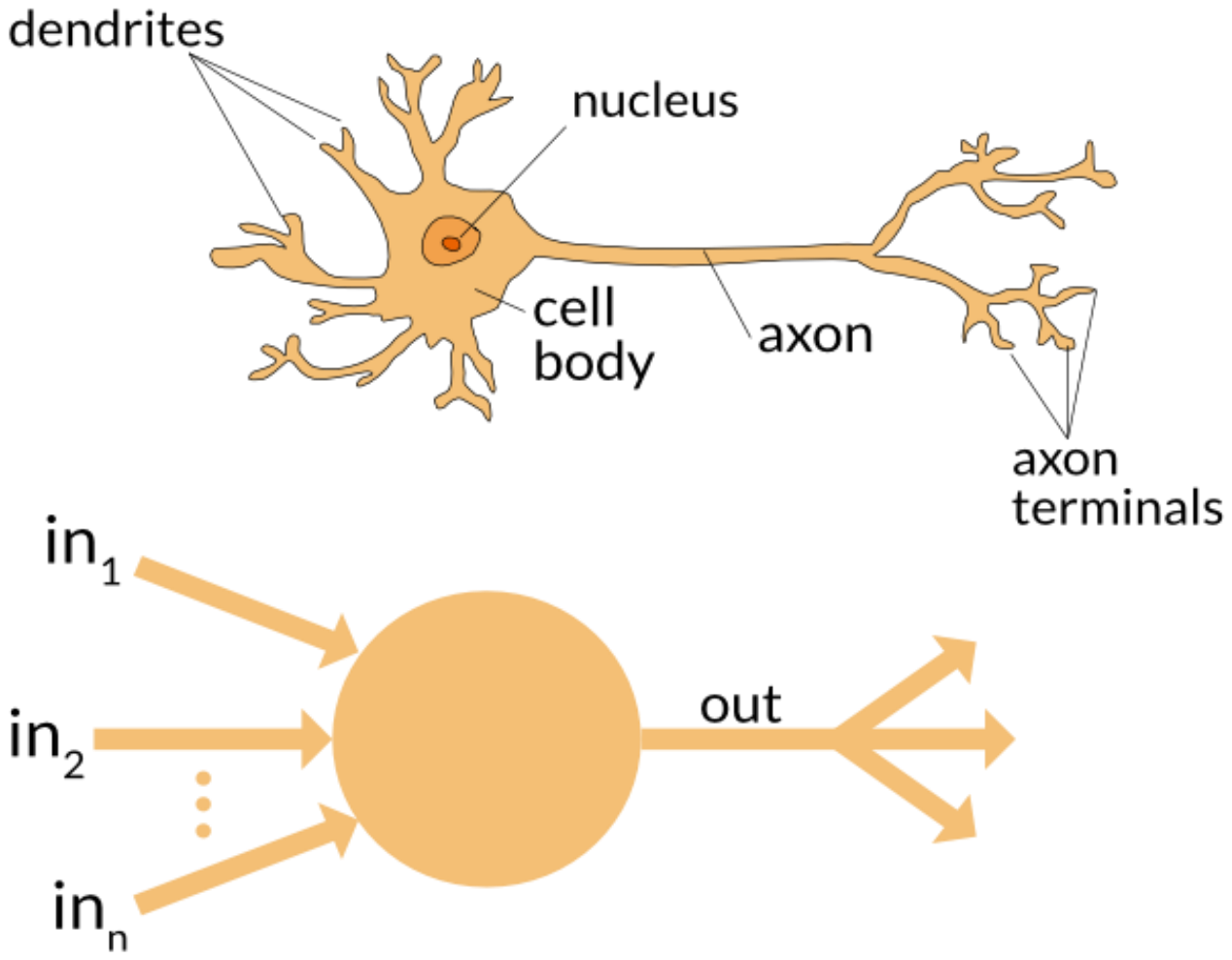
Supervised Learning

[Deep Learning]

✓Neural Network
: Simple MLP (Multi-Layer Perceptron)

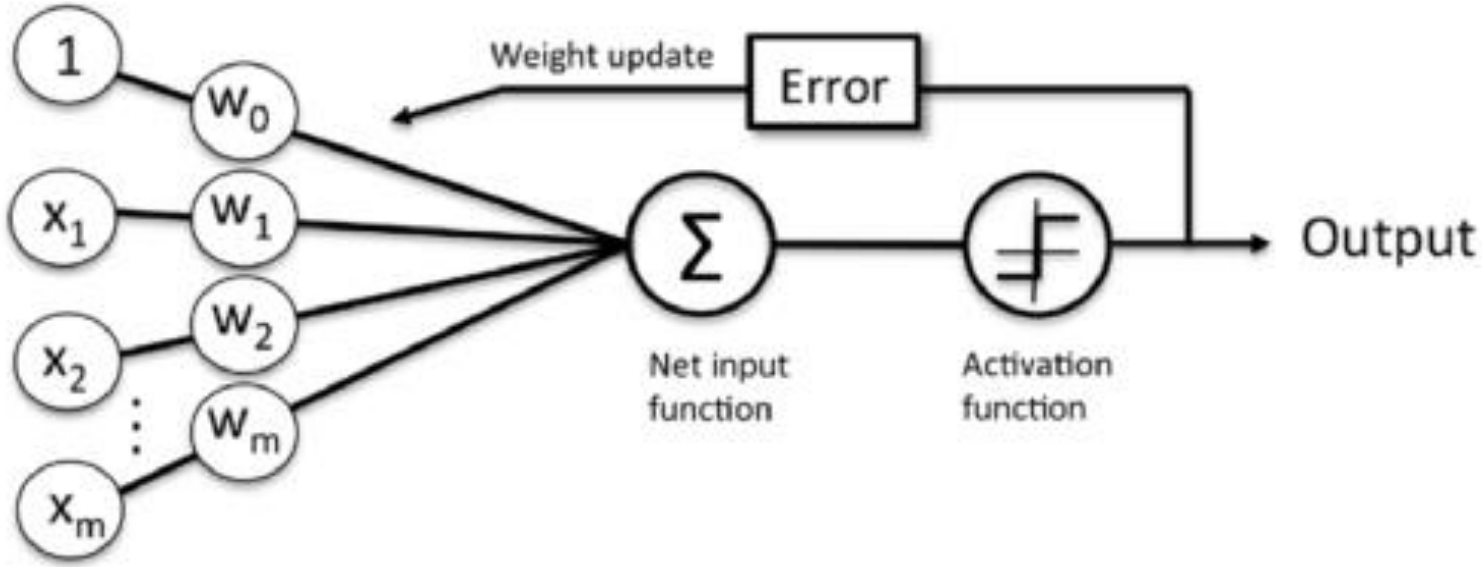
Note)
이건 mnist 로 하자!!!!
젠장, Neural Network 설명도 해야해…

Supervised Learning



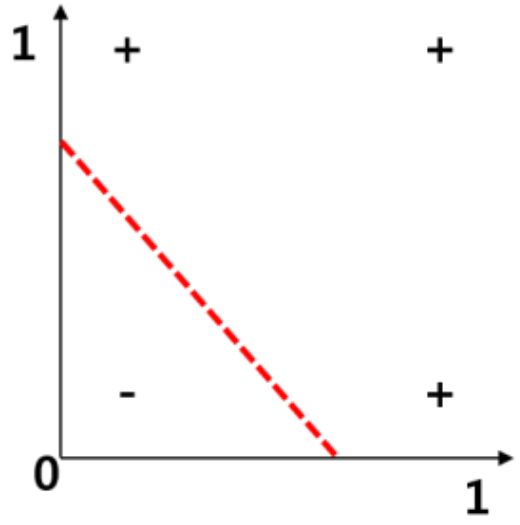
Supervised Learning

Single Layer Neural Network



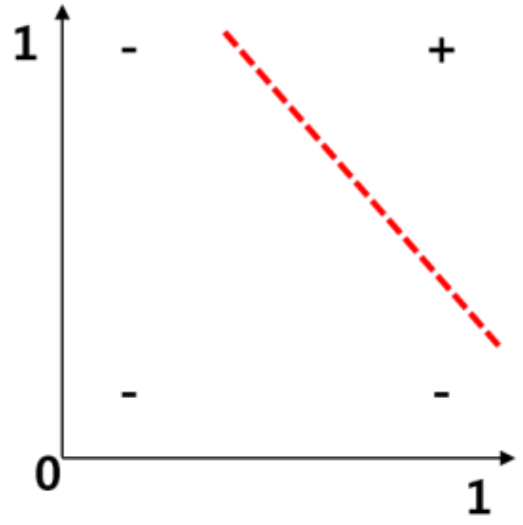
Supervised Learning

OR



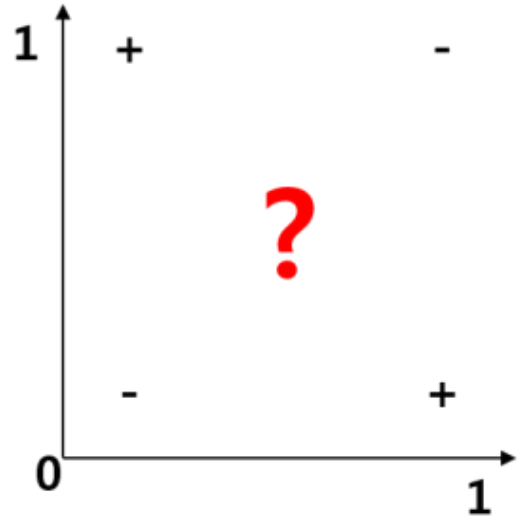
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

AND



x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

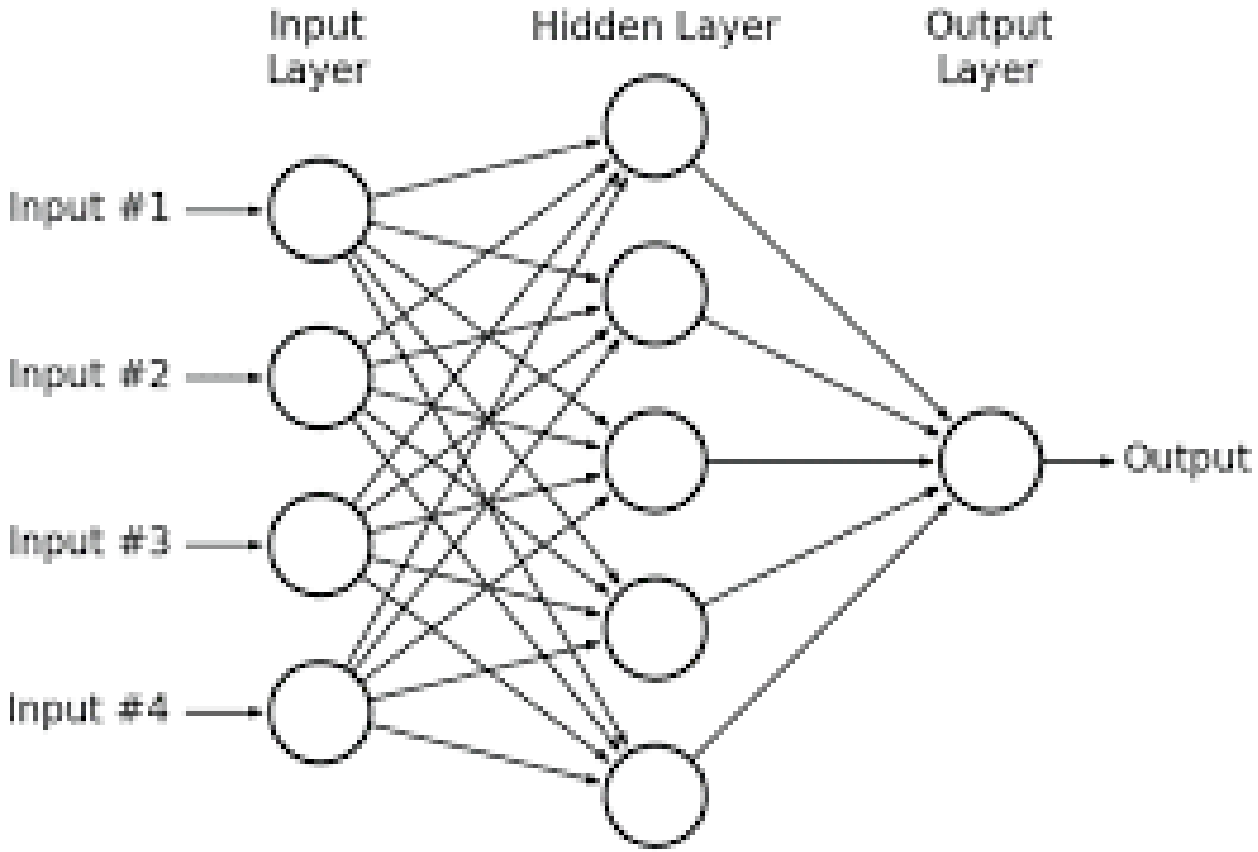
XOR



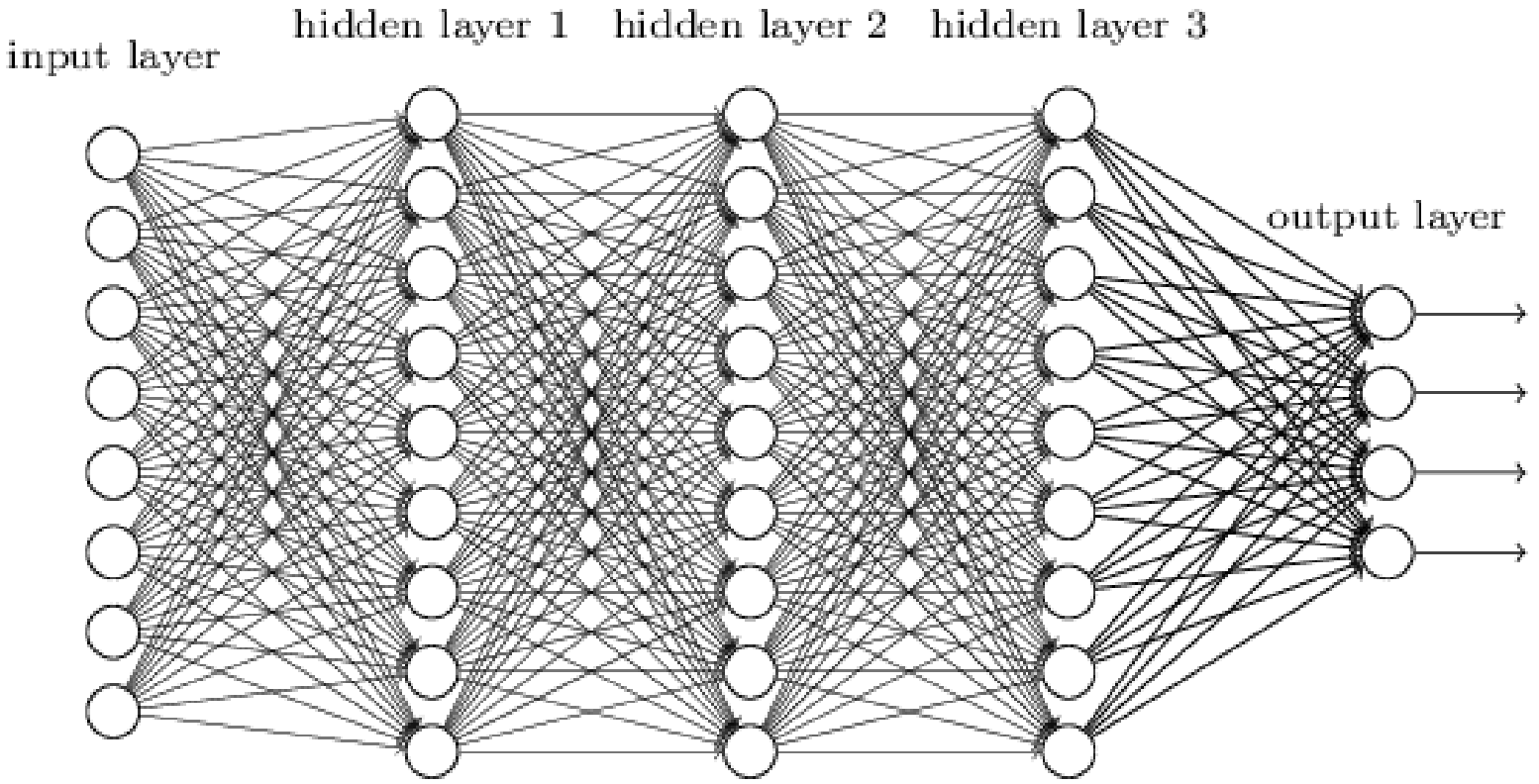
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

Supervised Learning

Multi–Layer Neural Network

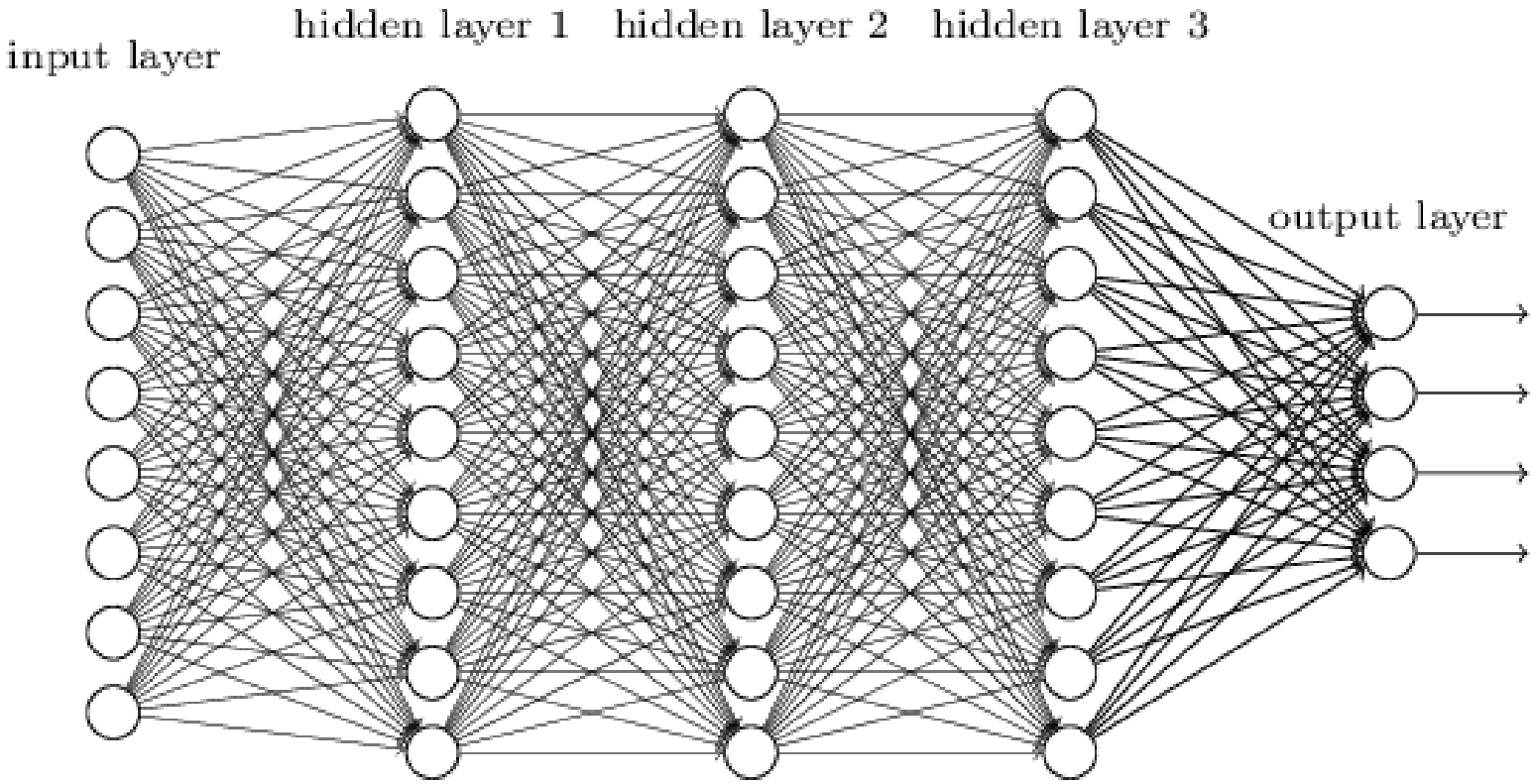


Neural-network



결과 예측

Neural-network



Layer 값 조정

Neural Network

✓HOL

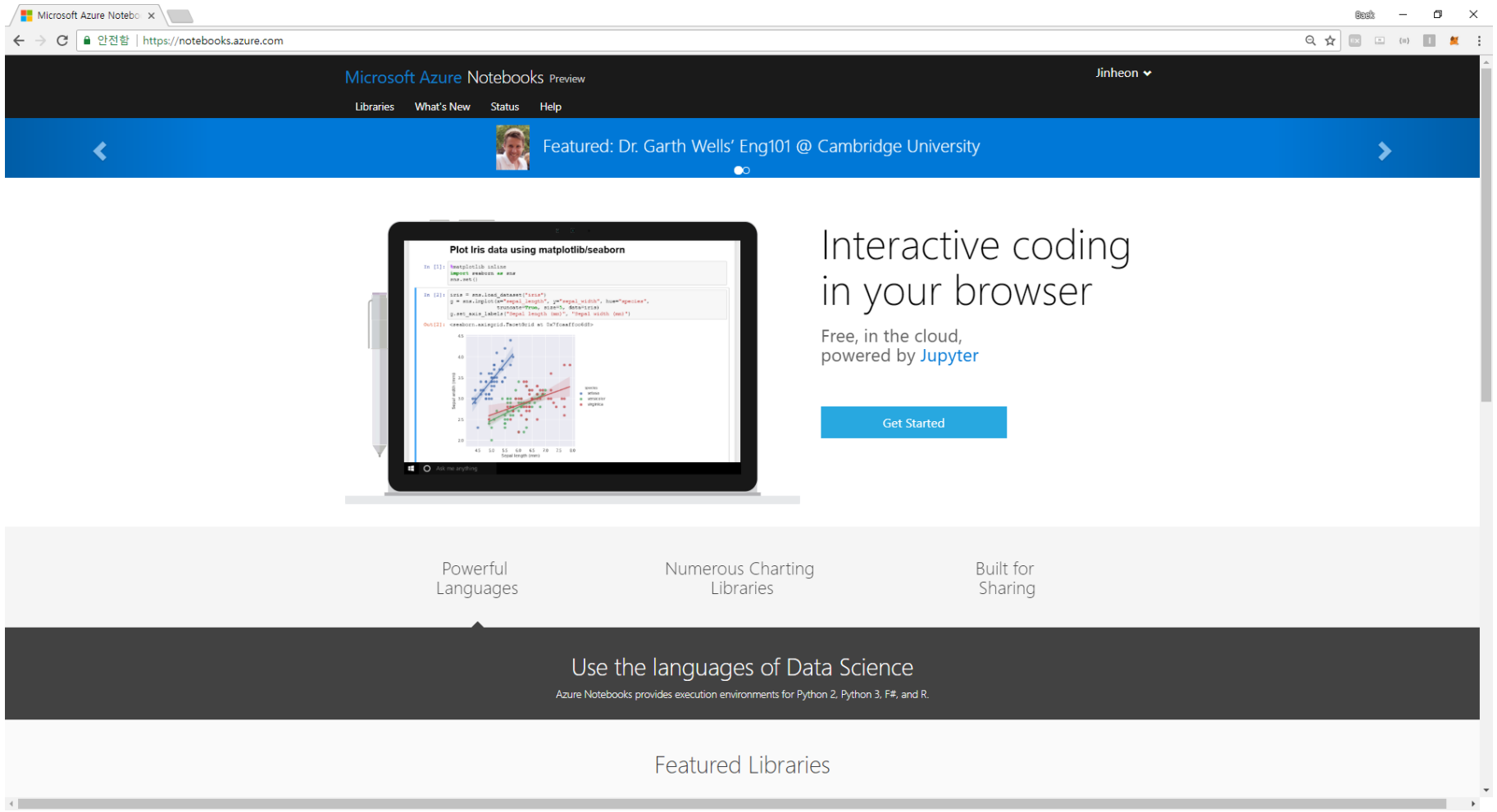
“MNIST Data Set ?”

Neural Network

28 x 28, 필기체 숫자 (0 ~ 9)

Neural Network

✓Azure Notebook (https://notebooks.azure.com)



Neural Network

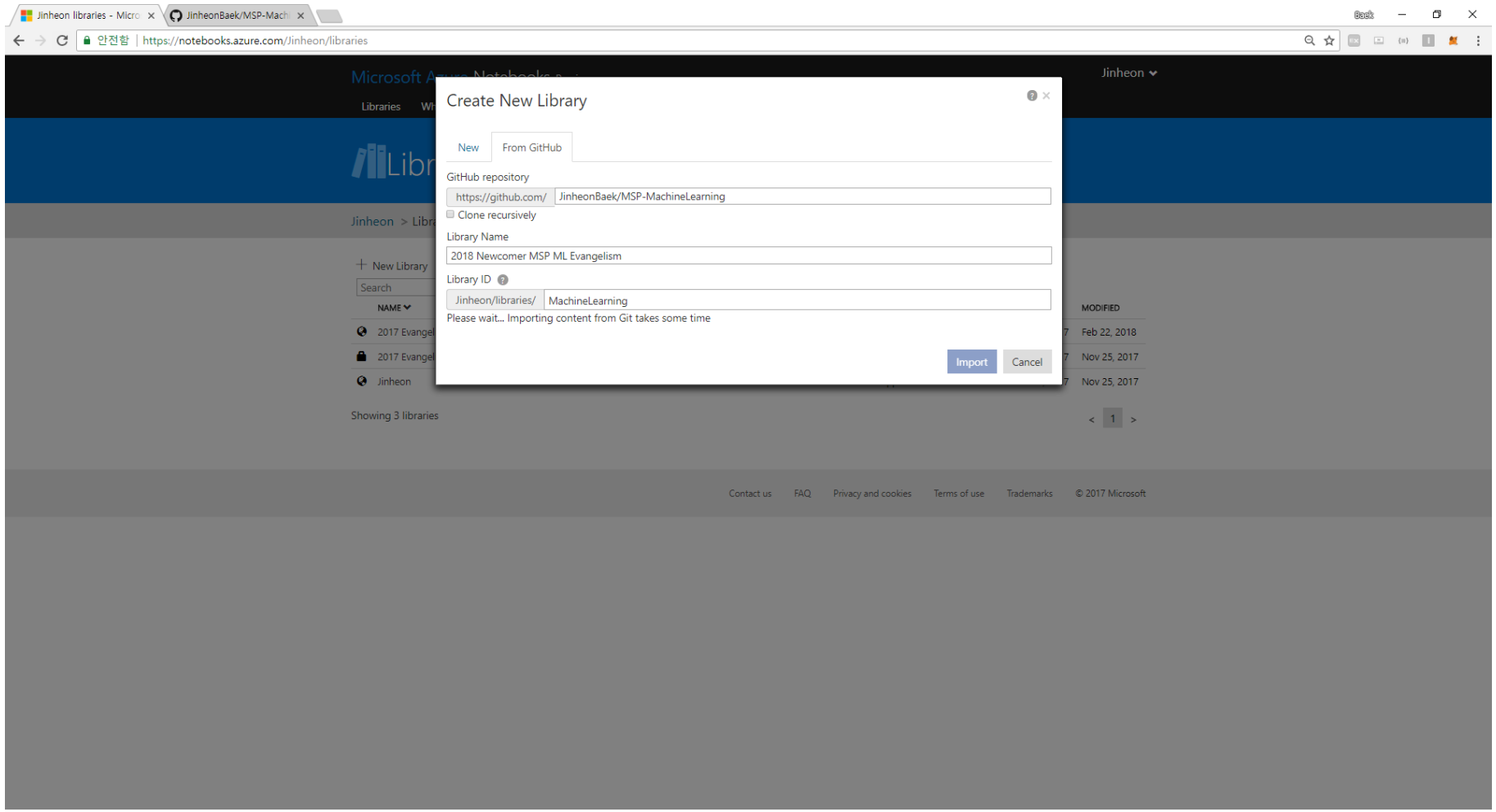
✓Azure Notebook (https://notebooks.azure.com)

1. Logic
2. Libraries 메뉴
3. New Library

Neural Network

GitHub repository: JinheonBaek/MSP-MachineLearning
Library Name (예시): 2018 Newcomer MSP ML Evangelism
Library ID (예시): MachineLearning

✓Azure Notebook (https://notebooks.azure.com)



Neural Network

✓HOL - Challenge

“Testing score is above 0.98”

Fitting

Machine Learning

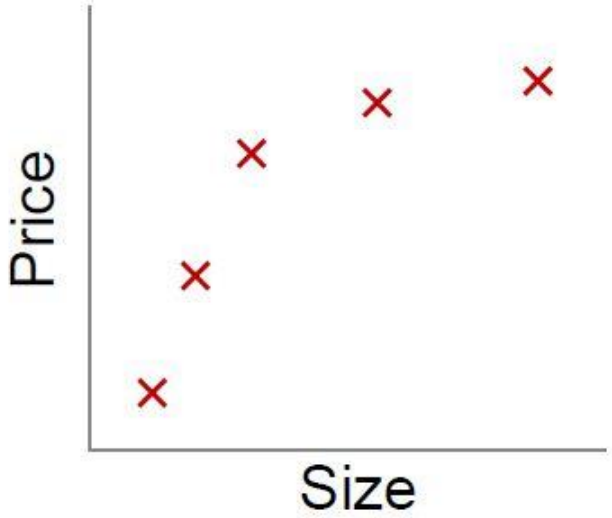
Microsoft Student Partner
백진헌

Fitting

✓Fitting

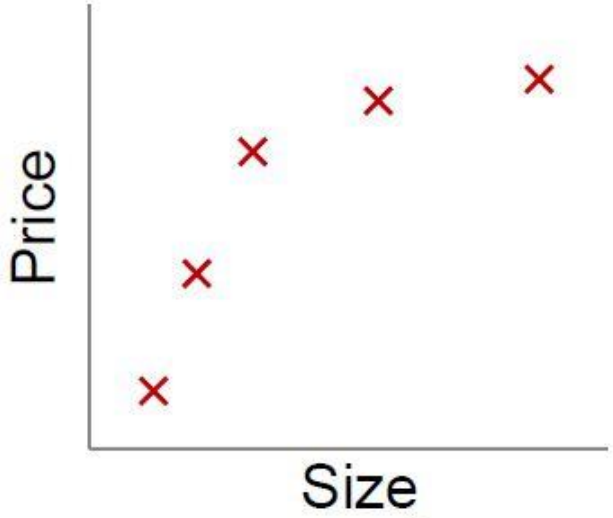
$\theta_0 + \theta_1 x$

High bias
(underfit)



$\theta_0 + \theta_1 x + \theta_2 x^2$

"Just right"

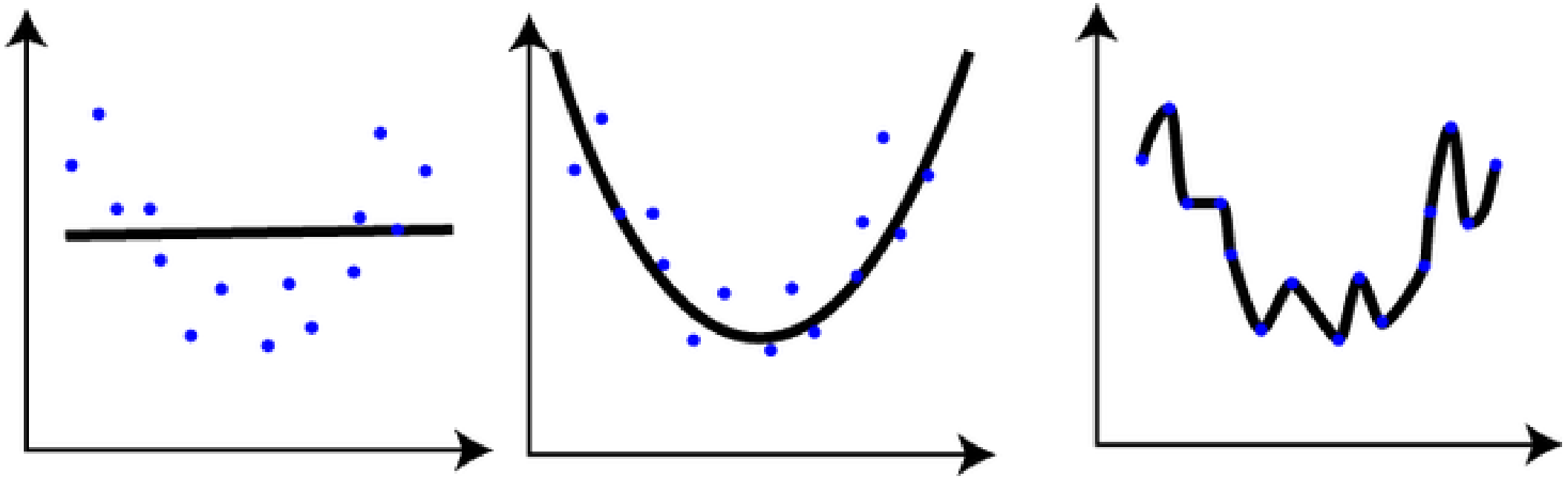


$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

High variance
(overfit)

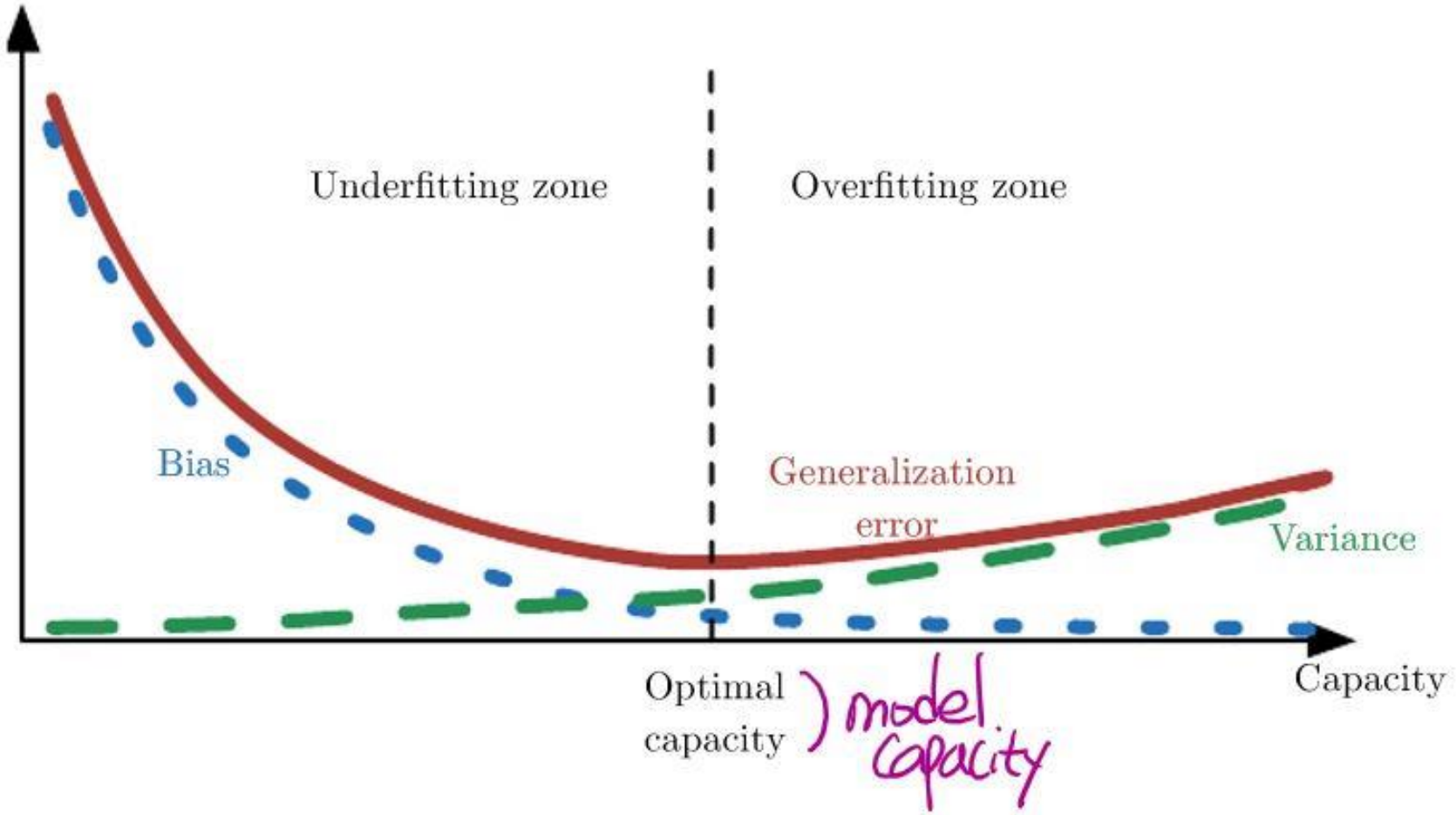
Fitting

✓Fitting



Fitting

✓**Fitting**



Fitting

✓Fitting HOL

“Connect Four”

Model

Machine Learning

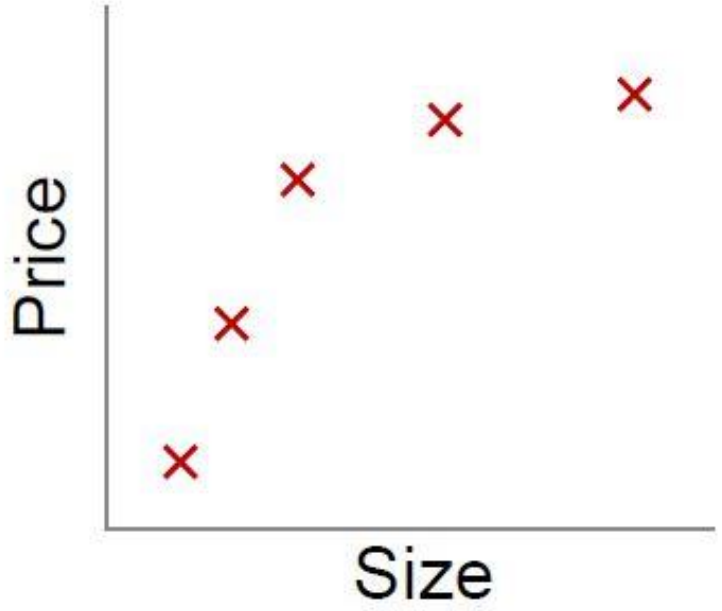
Microsoft Student Partner
백진헌

Models

- ✓Linear vs Nonlinear
- ✓Black box vs Descriptive
- ✓First-principle vs Data-driven

Models

✓Linear vs Nonlinear



$$\theta_0 + \theta_1 x$$

$$\theta_0 + \theta_1 x + \theta_2 x^2$$

Models

✓Blackbox vs Descriptive Model

Linear Regression

: Models are descriptive, because one can see which variables are weighed heaviest.

Neural Network

: Models are generally opaque.

Models

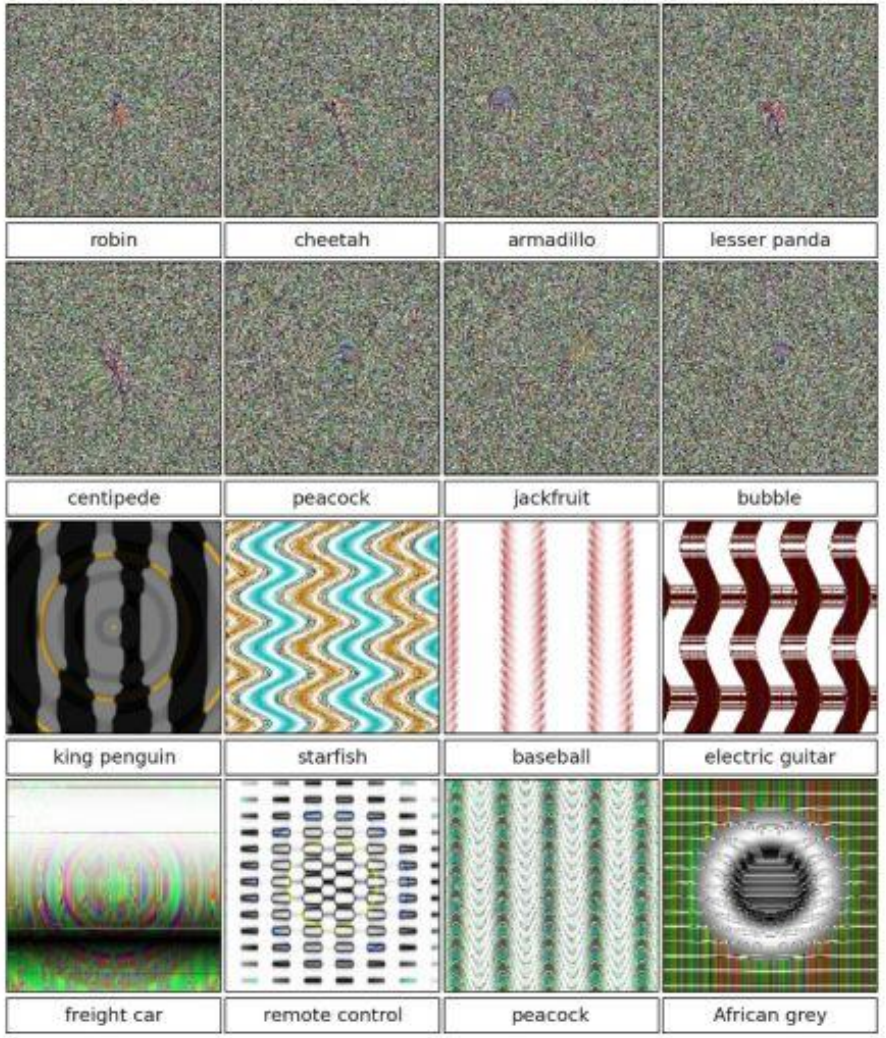
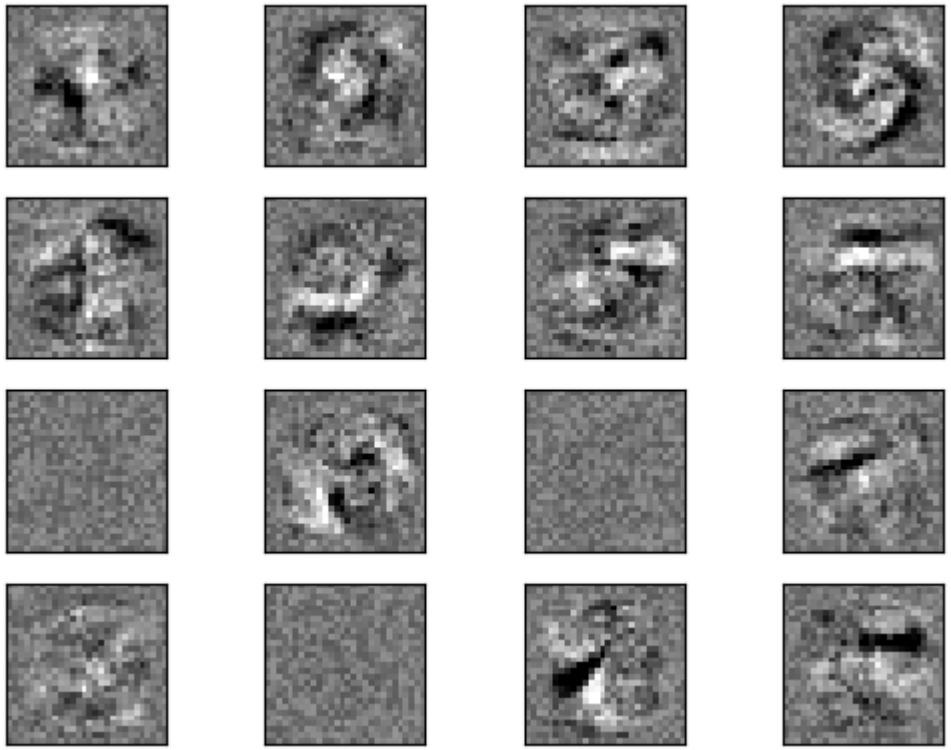


Figure 1. Evolved images that are unrecognizable to humans, but that state-of-the-art DNNs trained on ImageNet believe with $\geq$ 99.6% certainty to be a familiar object. This result highlights differences between how DNNs and humans recognize objects. Images are either directly (*top*) or indirectly (*bottom*) encoded.

Models



Models

✓First-principle vs Data-driven

First-principle
: Models based on a theoretical explanation.

Data-driven
: Models based on observed data correlations.

References

- ✓Deep Learning Book - MIT
- ✓Machine Learning - Coursera
- ✓Korea Univ. Cose432 (Data Science)

Q & A

Thanks