

NAÏVE BAYES

BY TAUTOLOGY

Naïve Bayes



Introduction

Introduction

What is
Naïve Bayes?

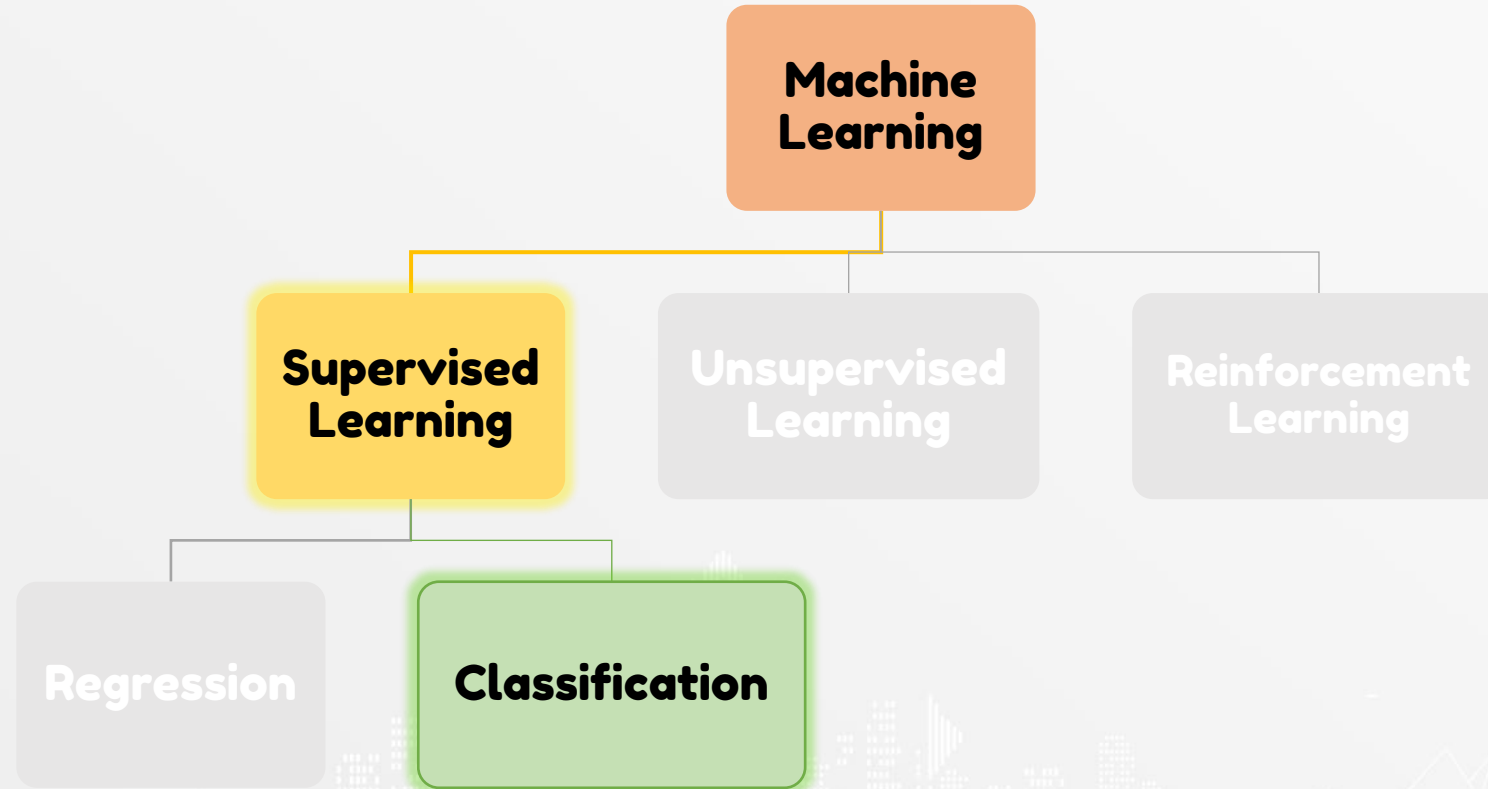
Pros & Cons

Real World
Application

What is Naïve Bayes?

Naïve Bayes เป็นหนึ่งใน algorithm ประเภท supervised learning
ที่ใช้สำหรับแก้ปัญหา classification โดยมีหลักการทำงานคือการ
พิจารณาความน่าจะเป็นซึ่งมีเงื่อนไขว่าทุกตัวแปรต้นเป็นอิสระต่อกัน

What is Naïve Bayes?



What is Naïve Bayes?

work	education	status
active	study	rich
active	not study	rich
lazy	study	rich
lazy	not study	poor
active	not study	poor
lazy	not study	poor

active

study

?

$\text{argmax}(P(\text{rich}|\text{active, study}), P(\text{poor}|\text{active, study}))$

What is Naïve Bayes?

$$C = \operatorname{argmax}_m \left(\frac{P(C_m) \prod_{d=1}^p P(x_d|C_m)}{P(x_1, x_2, \dots, x_p)} \right)$$

$$C = \operatorname{argmax}_m (P(C_m) \prod_{d=1}^p P(x_d|C_m))$$

where

- ◆ C_m be the m^{th} class
- ◆ p be the number of feature

Introduction

**What is
Naïve Bayes?**



Pros & Cons



**Real World
Application**



Pros & Cons

ข้อดี

- เป็น algorithm เพียงตัวเดียว ที่ใช้หลักการความน่าจะเป็น

ข้อเสีย

- ไม่สามารถตีความผลลัพธ์ได้ของ model ได้

ข้อจำกัด

- สมมติฐานความเป็นอิสระต่อกันของตัวแปรต้น

Introduction

**What is
Naïve Bayes?**



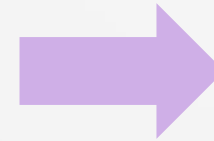
Pros & Cons



**Real World
Application**



Real World Application



วินิจฉัยโรคหัวใจบนเว็บไซต์หนึ่ง

โดยจะพิจารณาจากข้อมูล 15 ประเภท
เช่น เพศ อายุ น้ำหนัก ส่วนสูง ความดัน
โลหิต ประเภทการเจ็บหน้าอก อัตราการ
เต้นของหัวใจสูงสุด เป็นต้น

อ้างอิง : [2018, Dulhare] Prediction System For Heart Disease Using Naive Bayes



โดยพิจารณาจาก ความยาวของ
description แต่ละ post จำนวน
follower จำนวน following จำนวน
tweet ที่กด liked เป็นต้น

facebook/tautologyai
www.tautology.live

Introduction

**What is
Naïve Bayes?**



Pros & Cons



**Real World
Application**

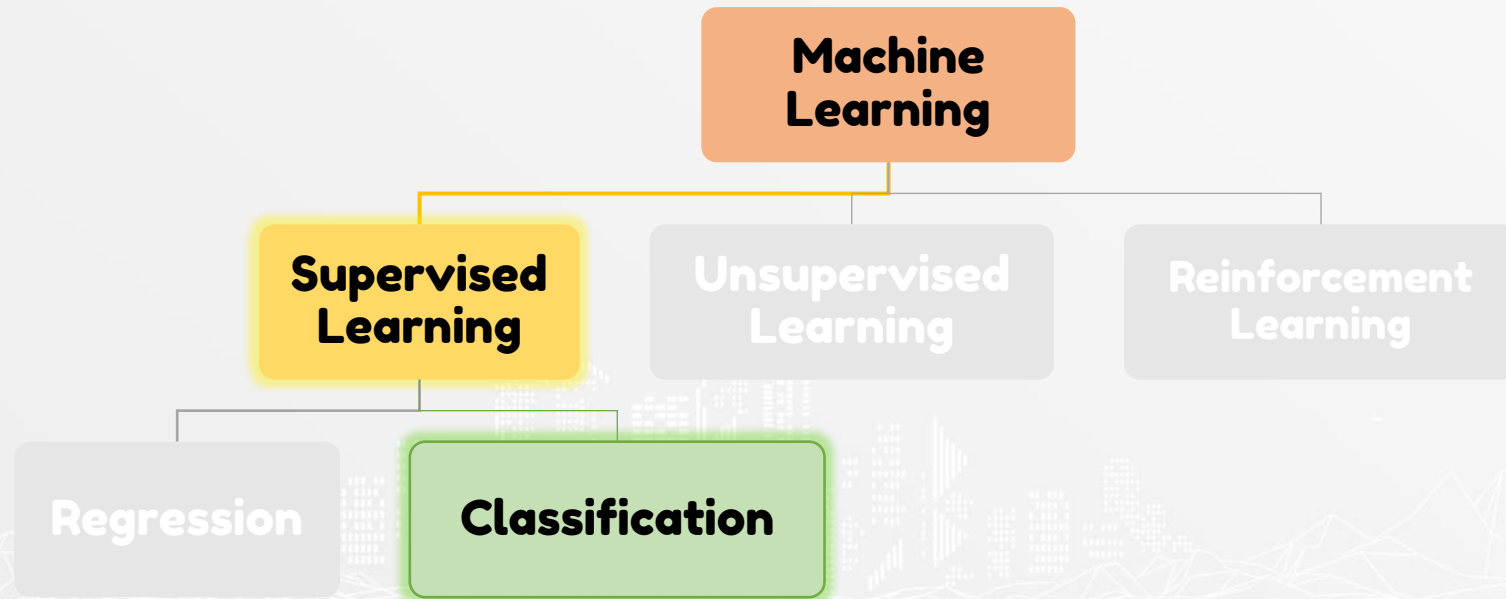


Naïve Bayes



Naïve Bayes

Naïve Bayes เป็นหนึ่งใน algorithm ประเภท supervised learning



Concept of Supervised Learning

Data \Rightarrow **Model** \Rightarrow **Prediction**



Model Creation

Model Creation

Assumption

Real Face of the
Model

How to Create Model
(Math)

How to Create Model
(Code)

Further Reading

Assumption

- All Feature are Independent
- No Missing Features

Model Creation

Assumption



Real Face of the
Model



How to Create Model
(Math)



How to Create Model
(Code)



Further Reading



Real Face of the Model

Naïve Bayes คือ การจำแนกประเภทของข้อมูล โดยใช้ทฤษฎีความน่าจะเป็นของเบส์ ที่เพิ่มสมมติฐานว่า คุณลักษณะต่าง ๆ (Feature) ไม่ขึ้นต่อกัน

work	education	status
active	study	rich
active	not study	rich
lazy	study	rich
lazy	not study	poor
active	not study	poor
lazy	not study	poor

active

study

?

$\text{argmax}(P(\text{rich}|\text{active, study}), P(\text{poor}|\text{active, study}))$

Real Face of the Model

$$C = \operatorname{argmax}_m \left(\frac{P(C_m) \prod_{d=1}^p P(x_d | C_m)}{P(x_1, x_2, \dots, x_p)} \right)$$

$$C = \operatorname{argmax}_m (P(C_m) \prod_{d=1}^p P(x_d | C_m))$$

where

- ◆ C_m be the m^{th} class
- ◆ p be the number of feature

Model Creation

Assumption



**Real Face of the
Model**



**How to Create Model
(Math)**



**How to Create Model
(Code)**



Further Reading



How to Create Model (Math)

- ☐ Step 1 : หาความน่าจะเป็นจาก dataset
- ☐ Step 2 : เลือกตอบ class ที่ให้ค่าความน่าจะเป็นสูงที่สุด

How to Create Model (Math)

ตัวอย่างการคำนวณ Naïve Bayes

sex	education	acceptation
female	master	yes
male	bachelor	no
female	bachelor	yes
male	doctor	yes
female	master	no

ตารางแสดงการรับเข้าทำงานของบริษัท

How to Create Model (Math)

☑ **Step 1** : หาความน่าจะเป็นจาก dataset

sex = female
education = master

$$\operatorname{argmax}(P(\text{yes}|\text{female, master}), P(\text{no}|\text{female, master}))$$

How to Create Model (Math)

☑ **Step 1** : หาความน่าจะเป็นจาก dataset

$$\begin{aligned} &P(\text{yes}|\text{female}, \text{master}) \\ &= P(\text{yes})P(\text{female}|\text{yes})P(\text{master}|\text{yes}) \\ &= \frac{3}{5} \times \frac{2}{3} \times \frac{1}{3} = \frac{2}{15} \end{aligned}$$

$$\begin{aligned} &P(\text{no}|\text{female}, \text{master}) \\ &= P(\text{no})P(\text{female}|\text{no})P(\text{master}|\text{no}) \\ &= \frac{2}{5} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{10} \end{aligned}$$

How to Create Model (Math)

☑ **Step 2** : เลือกตอบ class ที่ให้ค่าความน่าจะเป็นสูงที่สุด

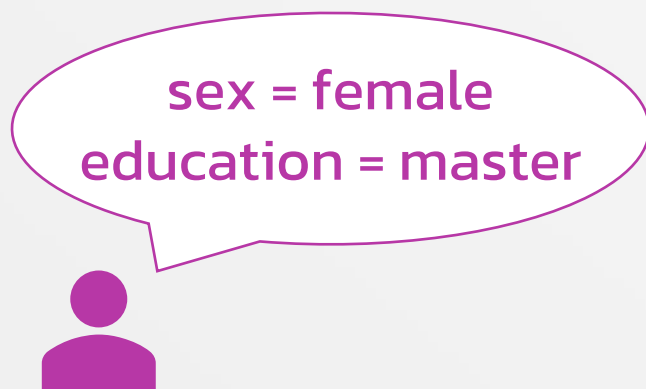
$$\operatorname{argmax}(P(\text{yes}|\text{female, master}), P(\text{no}|\text{female, master}))$$

$$= \operatorname{argmax}\left(\frac{2}{15}, \frac{1}{10}\right)$$

$$= \text{yes}$$

How to Create Model (Math)

☑ **Step 2** : เลือกตอบ class ที่ให้ค่าความน่าจะเป็นสูงที่สุด



Model Creation

Assumption



**Real Face of the
Model**



**How to Create Model
(Math)**



**How to Create Model
(Code)**



Further Reading



How to Create Model (Code)

ตัวอย่าง Code สำหรับ Naïve Bayes

sex	education	acceptation
0	1	yes
1	0	no
0	0	yes
1	2	yes
0	1	no

ตารางแสดงการรับเข้าทำงานของบริษัท

How to Create Model (Code)

- Code สำหรับสร้าง model จากข้อมูลของเราโดยที่

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} yes \\ no \\ yes \\ yes \\ no \end{bmatrix}$$

```
1 clf = GaussianNB()  
2 clf.fit(X, y)
```

GaussianNB()

How to Create Model (Code)



Code for this section



Open File
Model Creation.ipynb

Model Creation

Assumption



**Real Face of the
Model**



**How to Create Model
(Math)**



**How to Create Model
(Code)**



Further Reading



Further Reading

- Bayes' Theorem

Model Creation

Assumption



**Real Face of the
Model**



**How to Create Model
(Math)**



**How to Create Model
(Code)**



Further Reading



Naïve Bayes



Prediction

Prediction

Naïve Bayes คือ การจำแนกประเภทของข้อมูล โดยใช้ทฤษฎีความน่าจะเป็นของเบส์ ที่เพิ่มสมมติฐานว่า คุณลักษณะต่าง ๆ (Feature) ไม่ขึ้นต่อกัน

work	education	status
active	study	rich
active	not study	rich
lazy	study	rich
lazy	not study	poor
active	not study	poor
lazy	not study	poor

active

study

?

$\text{argmax}(P(\text{rich}|\text{active}, \text{study}), P(\text{poor}|\text{active}, \text{study}))$

Prediction

1-Sample

Multi-Sample

Code


1-Sample

ตัวอย่างการคำนวณ \hat{y}

sex	education		\hat{y}
0	1	→	?

1-Sample

sex	education
0	1



\hat{y}
yes

$$\begin{aligned} & \operatorname{argmax}(P(\text{yes}|\text{female, master}), P(\text{no}|\text{female, master})) \\ &= \operatorname{argmax}\left(\frac{2}{15}, \frac{1}{10}\right) \\ &= \text{yes} \end{aligned}$$

Prediction

1-Sample



Multi-Sample



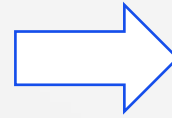
Code



Multi-Sample

ตัวอย่างการคำนวณ \hat{y}

sex	education
0	1
0	2
1	0
1	1



\hat{y}
?
?
?
?

Multi-Sample

sex = female
education = master

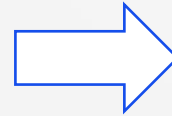
$$\operatorname{argmax}(P(\text{yes}|\text{female, master}), P(\text{no}|\text{female, master}))$$

$$= \operatorname{argmax}\left(\frac{2}{15}, \frac{1}{10}\right)$$

$$= \text{yes}$$

Multi-Sample

sex	education
0	1
0	2
1	0
1	1



\hat{y}
yes
?
?
?

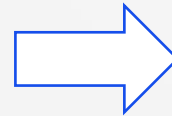
Multi-Sample

sex = female
education = doctor

$$\begin{aligned} & \operatorname{argmax}(P(\text{yes}|\text{female, doctor}), P(\text{no}|\text{female, doctor})) \\ &= \operatorname{argmax}\left(\frac{2}{15}, 0\right) \\ &= \text{yes} \end{aligned}$$

Multi-Sample

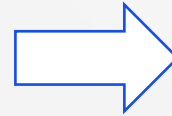
sex	education
0	1
0	2
1	0
1	1



\hat{y}
yes
yes
?
?

Multi-Sample

sex	education
0	1
0	2
1	0
1	1



\hat{y}
yes
yes
no
no

Prediction

1-Sample



Multi-Sample



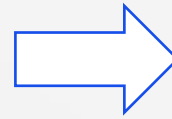
Code



Code

ตัวอย่าง code สำหรับการคำนวณ \hat{y}

sex	education
0	1
0	2
1	0
1	1



\hat{y}
?
?
?
?

Code

- Code สำหรับสร้าง model จากข้อมูลของเราโดยที่

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 2 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

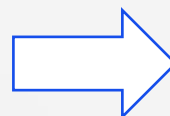
```
1 clf.predict(X)
```

```
array(['yes', 'yes', 'no', 'no'], dtype='<U3')
```

Code

ดังนั้น เราจะได้ \hat{y} สำหรับข้อมูลชุดนี้คือ

sex	education
0	1
0	2
1	0
1	1



\hat{y}
yes
yes
no
no

Code



Code for this section



Open File
Model Creation.ipynb

Prediction

1-Sample



Multi-Sample



Code



Naïve Bayes



Workshop

AI in Agriculture

- Abstract
- Why this project important?
- Who this project for?
- Plant Disease Dataset
- What we learn from this project?

Abstract

สร้าง model เพื่อจำแนกว่าพืชเป็นโรคหรือไม่ โดยพิจารณาจากใบพืช



Why this project important?



- ◆ สามารถสร้างระบบตรวจโรคพืชที่ทำงานได้อย่างทันท่วงที
- ◆ สามารถนำความรู้ไปต่อยอดเพื่อสร้าง smart farm
- ◆ สามารถนำไปต่อยอดตรวจโรคในพืชชนิดอื่น ๆ

Who this project is for?

- ◆ เกษตรกรที่สนใจ AI กับการเกษตร
- ◆ ผู้ควบคุม/วางแผนการผลิต
- ◆ นักวิเคราะห์ข้อมูล



Plant Disease Dataset



<http://archive.ics.uci.edu/ml/datasets/Leaf?ref=datanews.io#>

Plant Disease Dataset

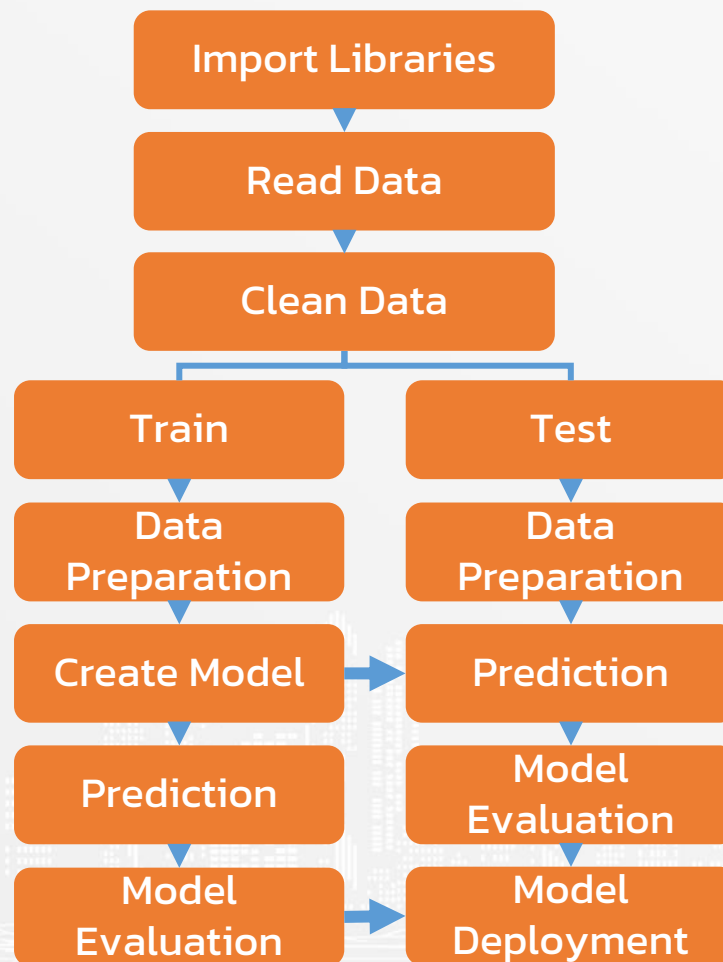
Feature



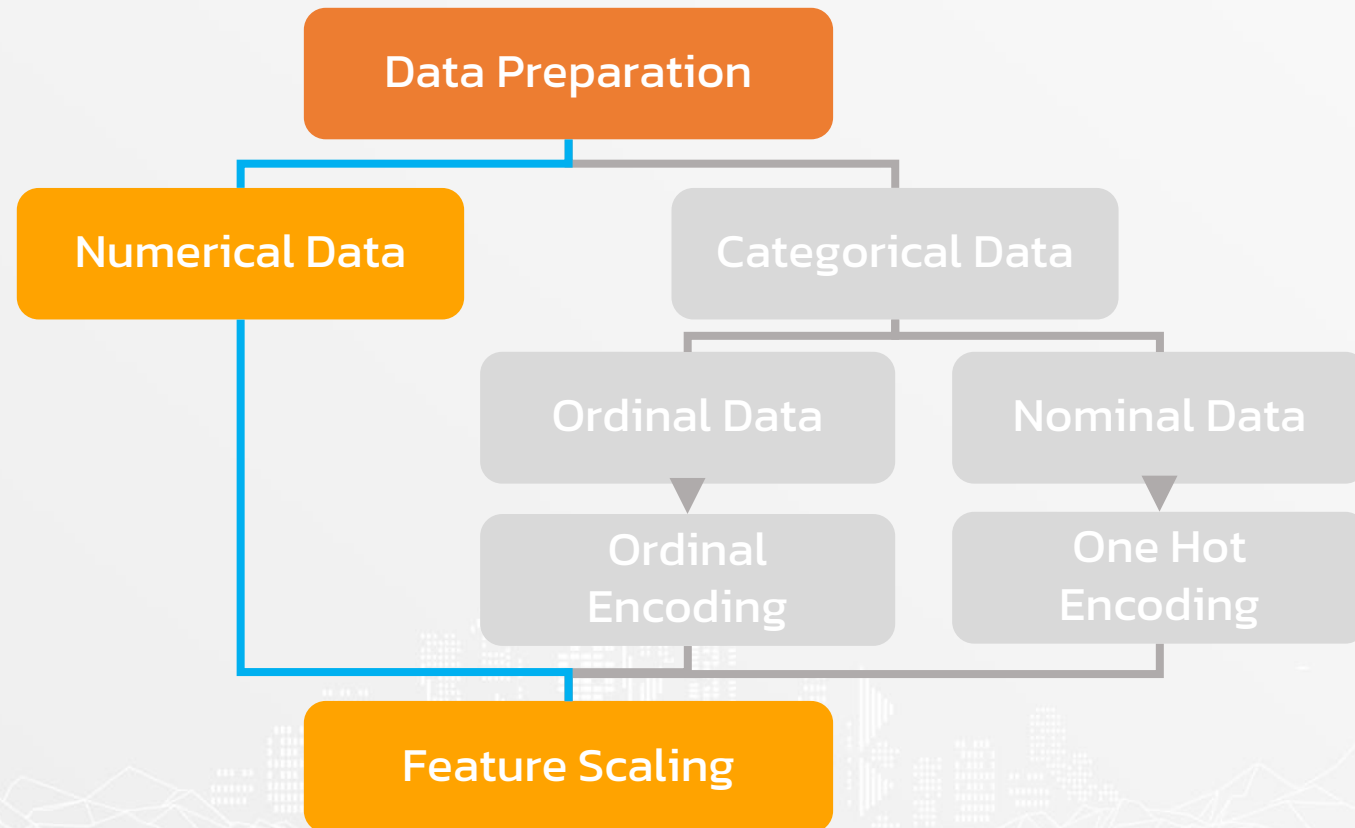
Target

- target : การเป็นโรคของพืช (healthy, unhealthy)

What we learn from this project?



Data Preparation



02. PLANT DISEASE

 **dataset**

 **plant_disease_mc.ipynb**

 **plant_disease_md.ipynb**

 **plant_disease_model.pickle**

Naïve Bayes

