

LINEAR DISCRIMINANT ANALYSIS

BY TAUTOLOGY

Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

Pros & Cons

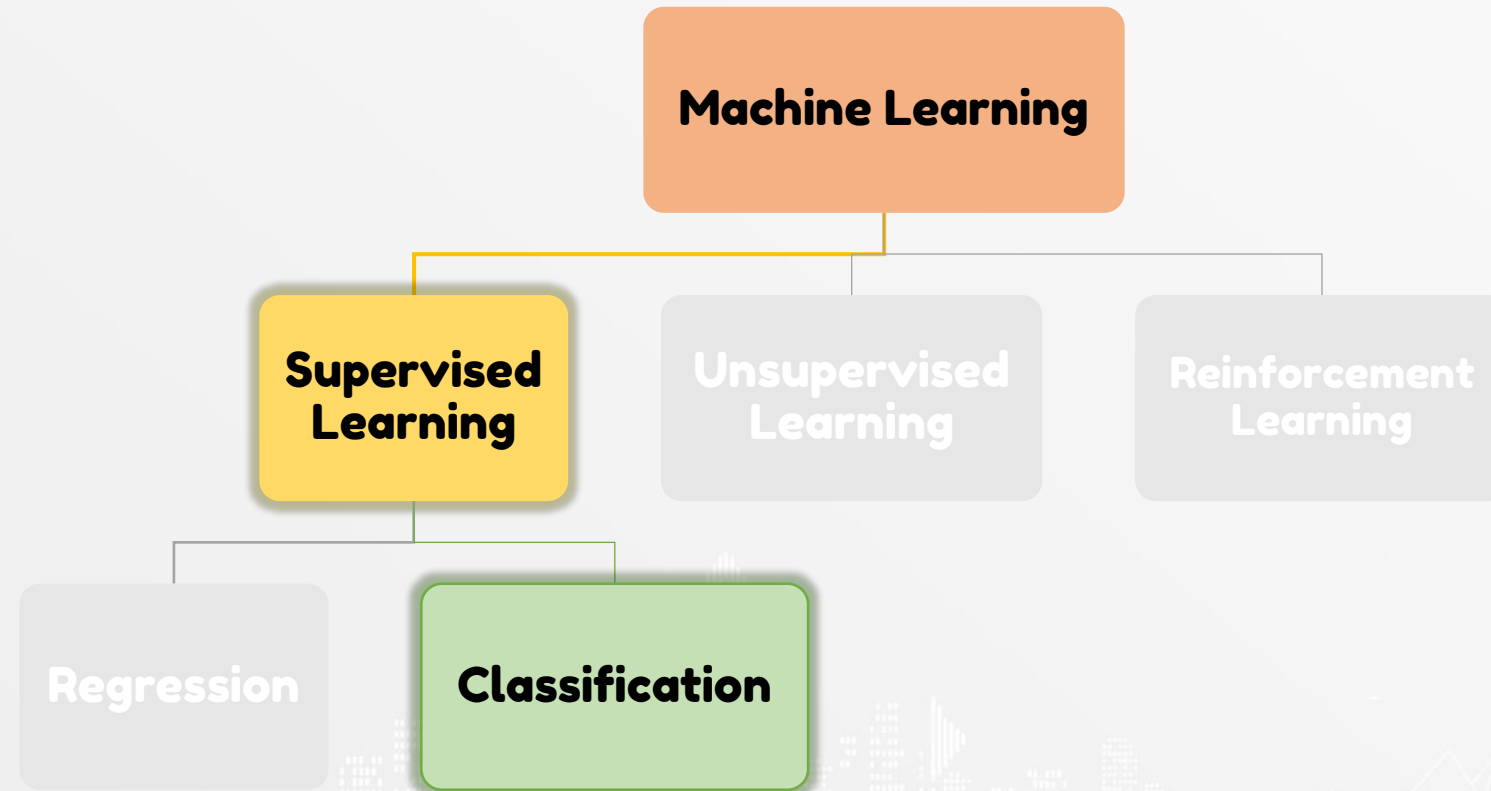
Workshop

What is Linear Discriminant Analysis?

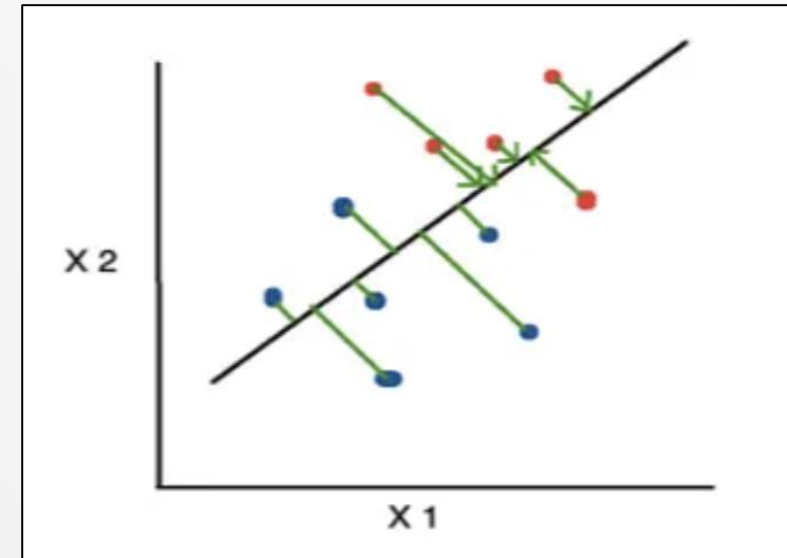
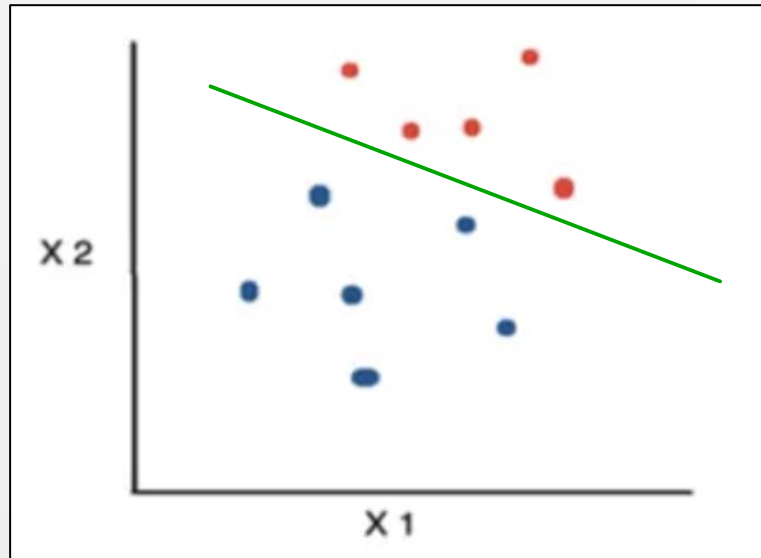
What is Linear Discriminant Analysis?

Linear Discriminant Analysis เป็นหนึ่งใน algorithm ประเภท supervised learning ที่ใช้สำหรับแก้ปัญหา classification โดยมีหลักการทำงาน คือ การ map ข้อมูลไปยังแกนใหม่เพื่อให้ข้อมูลประเภทเดียวกันกระจุกตัวกันมากที่สุด ในขณะที่จุดศูนย์กลางของข้อมูลคนละประเภทจะอยู่ไกลกันมากที่สุด

What is Linear Discriminant Analysis?

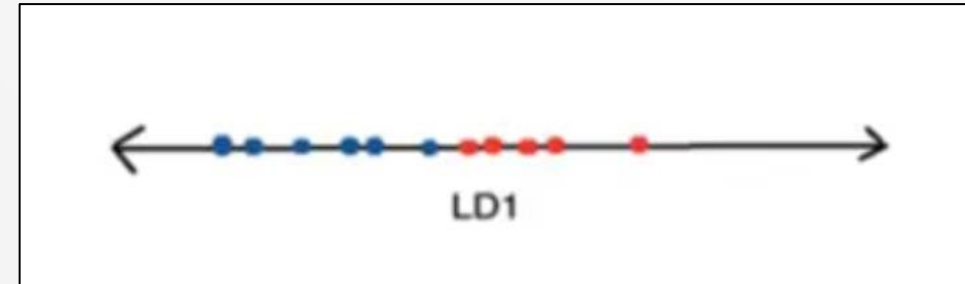
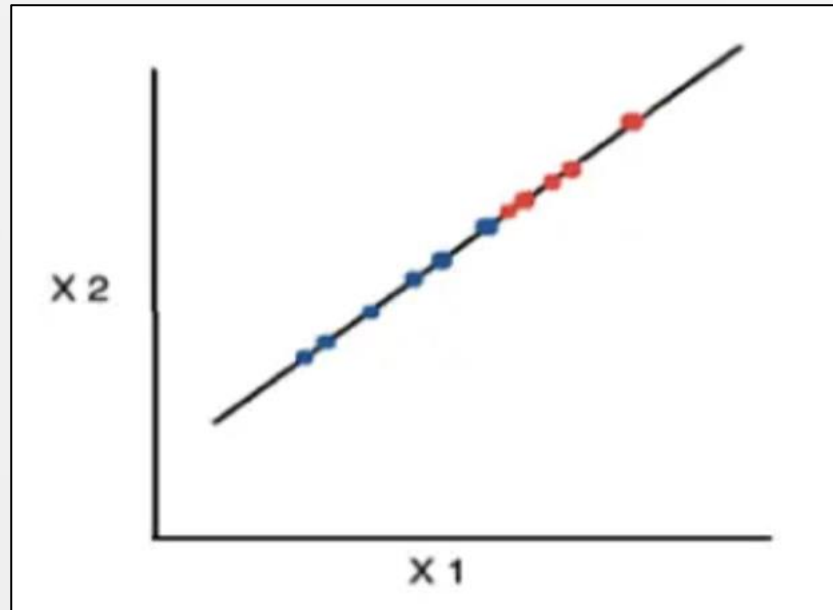


What is Linear Discriminant Analysis?



Ref : <https://www.knowledgehut.com/blog/data-science/linear-discriminant-analysis-for-machine-learning>

What is Linear Discriminant Analysis?



Ref : <https://www.knowledgehut.com/blog/data-science/linear-discriminant-analysis-for-machine-learning>

Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

Pros & Cons

Workshop

Assumption

Assumption

- Linear Separability
- ✗ • Gaussian Distribution ^{Normal}
- No Missing Features

→ ในกรณีที่ทำการแบ่ง class
เราสามารถใช้ logistic regression แทนได้

$$p \rightarrow q \neq \sim p \rightarrow \sim q$$

Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

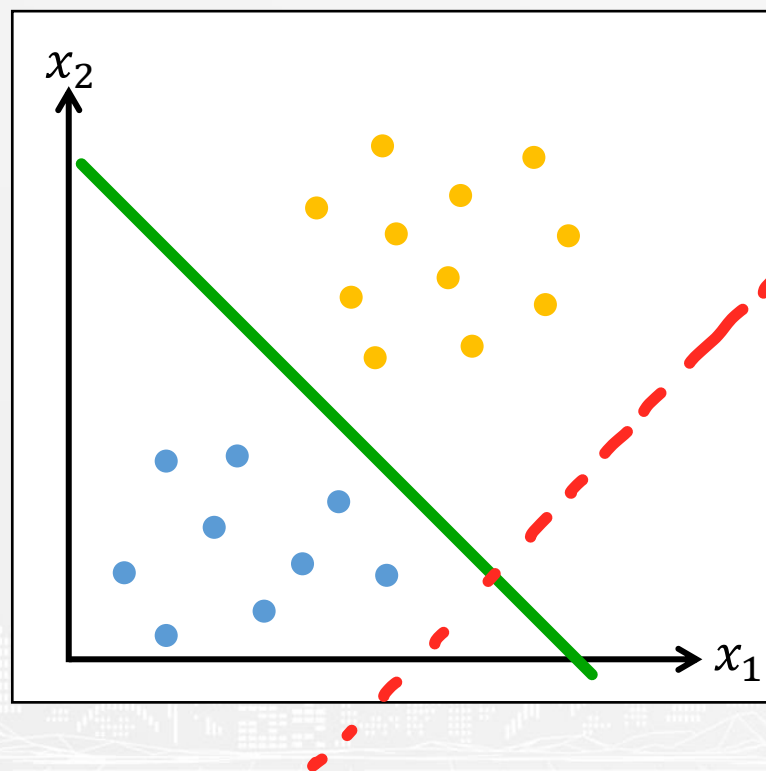
Pros & Cons

Workshop

Data for Linear Discriminant Analysis

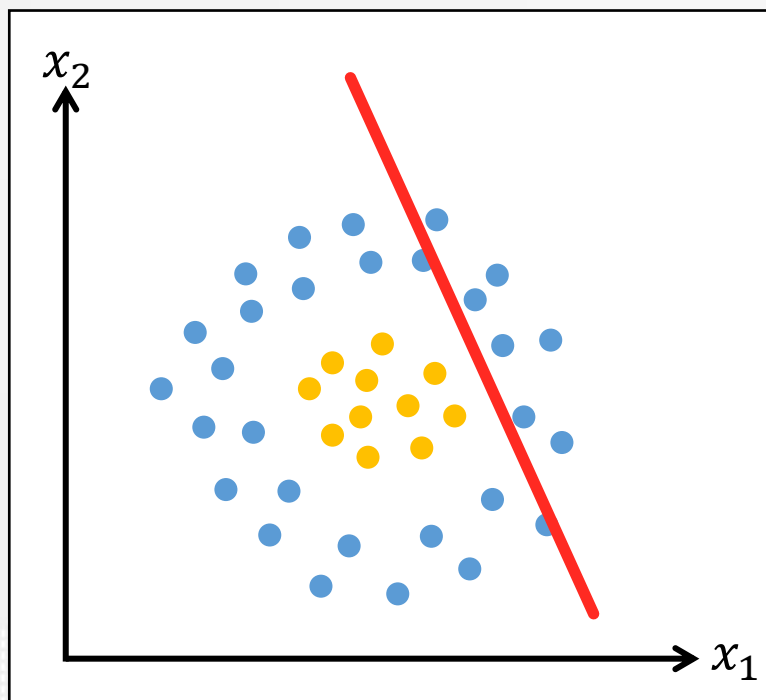
Data for Linear Discriminant Analysis

ตัวอย่างของข้อมูลที่เหมาะสมกับ LDA



Data for Linear Discriminant Analysis

ตัวอย่างของข้อมูลที่เหมาะสมกับ LDA



Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

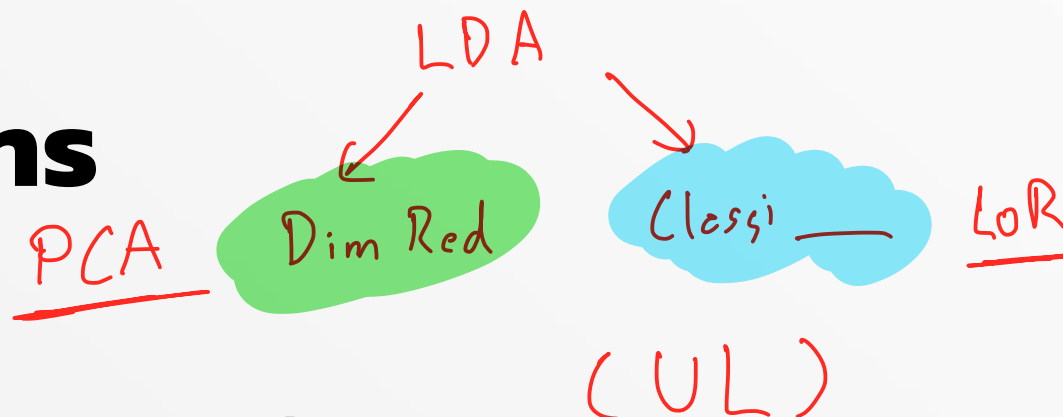
Pros & Cons

Workshop

Pros & Cons

ด้านบวก → จะมี algorithm อื่นๆ แทนที่
แต่ด้านลบก็มี
- $P(A) + LoR$ (ข้อ ๓)
- จะไม่ถูกแทนที่เมื่อเราขาดใช้สมมติฐาน LDA (specific)

Pros & Cons



ข้อดี

- เป็น algorithm เพียงตัวเดียวที่ทำทั้ง dimensionality reduction และ classification ในเวลาเดียวกัน
- สามารถจัดการได้กับทั้งข้อมูลที่เรียบง่าย และข้อมูลที่ซับซ้อน (ลอง)

ข้อเสีย

- เป็น algorithm ที่ซับซ้อน & ยากต่อการเข้าใจ

(Lagrange Multipliers)
↓
(Eigen Values & Eigen Vectors)

ข้อจำกัด

- โดยทั่วไปใช้งานได้กับข้อมูลที่สามารถแบ่งด้วยสมการเชิงเส้นได้

Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

Pros & Cons

Workshop

Workshop

Approval

- Abstract
- Why this project important?
- Who this project for?
- Credit Card Dataset
- What we learn from this project?

Abstract

สร้าง model เพื่อประเมินการอนุมัติเปิดบัตร credit โดย feature ที่นำมาใช้ เช่น

- เพศ
- อายุ
- รายได้ต่อปี
- ระดับการศึกษา
- อาชีพ



Why this project important?



- สามารถสร้างระบบอนุมัติการกู้ยืมเงิน
- สามารถสร้างระบบอนุมัติการรับพนักงาน
- สามารถสร้างระบบตรวจสอบคุณภาพสินค้า

Who is this project for?

- นักลงทุน
- ผู้ดูแลกองทุน
- นักวิจัยเชิงปริมาณ
- นักวิเคราะห์ข้อมูล

Credit Card Dataset

ข้อมูลของ Credit Card Approval



<https://www.kaggle.com/datasets/laotse/credit-card-approval>

Credit Card Dataset

Feature

- CODE_GENDER : เพศ
- FLAG_OWN_CAR : มีรถหรือไม่
- FLAG_OWN_REALTY : มีทรัพย์สินหรือไม่
- CNT_CHILDREN : จำนวนบุตร
- AMT_INCOME_TOTAL : รายได้ต่อปี
- NAME_EDUCATION_TYPE : ระดับการศึกษา
- NAME_FAMILY_STATUS : สถานะการแต่งงาน
- NAME_HOUSING_TYPE : สถานที่อยู่อาศัย
- DAYS_BIRTH : อายุ (วัน)
- DAYS_EMPLOYED : เวลาที่ทำงาน (วัน)

Credit Card Dataset

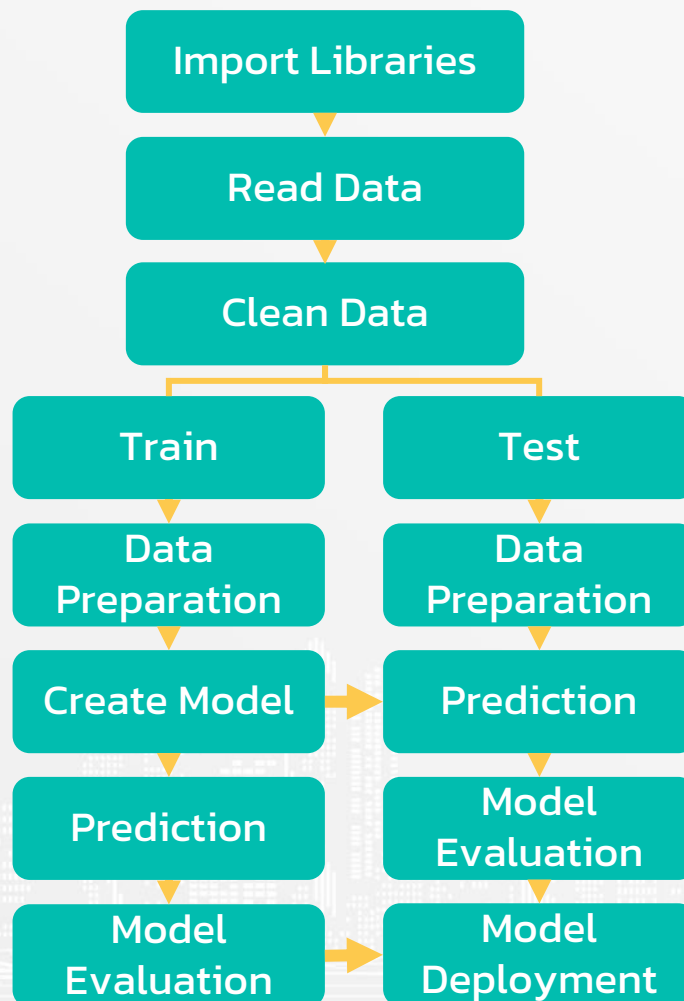
Feature

- FLAG_MOBIL : มีมือถือหรือไม่
- FLAG_WORK_PHONE : มีโทรศัพท์ที่ทำงานหรือไม่
- FLAG_PHONE : มีโทรศัพท์หรือไม่
- FLAG_EMAIL : มี e-mail หรือไม่
- JOB : อาชีพ
- BEGIN_MONTHS : อายุของข้อมูลที่เก็บ (เดือน)
- STATUS : สถานะการค้างชำระ

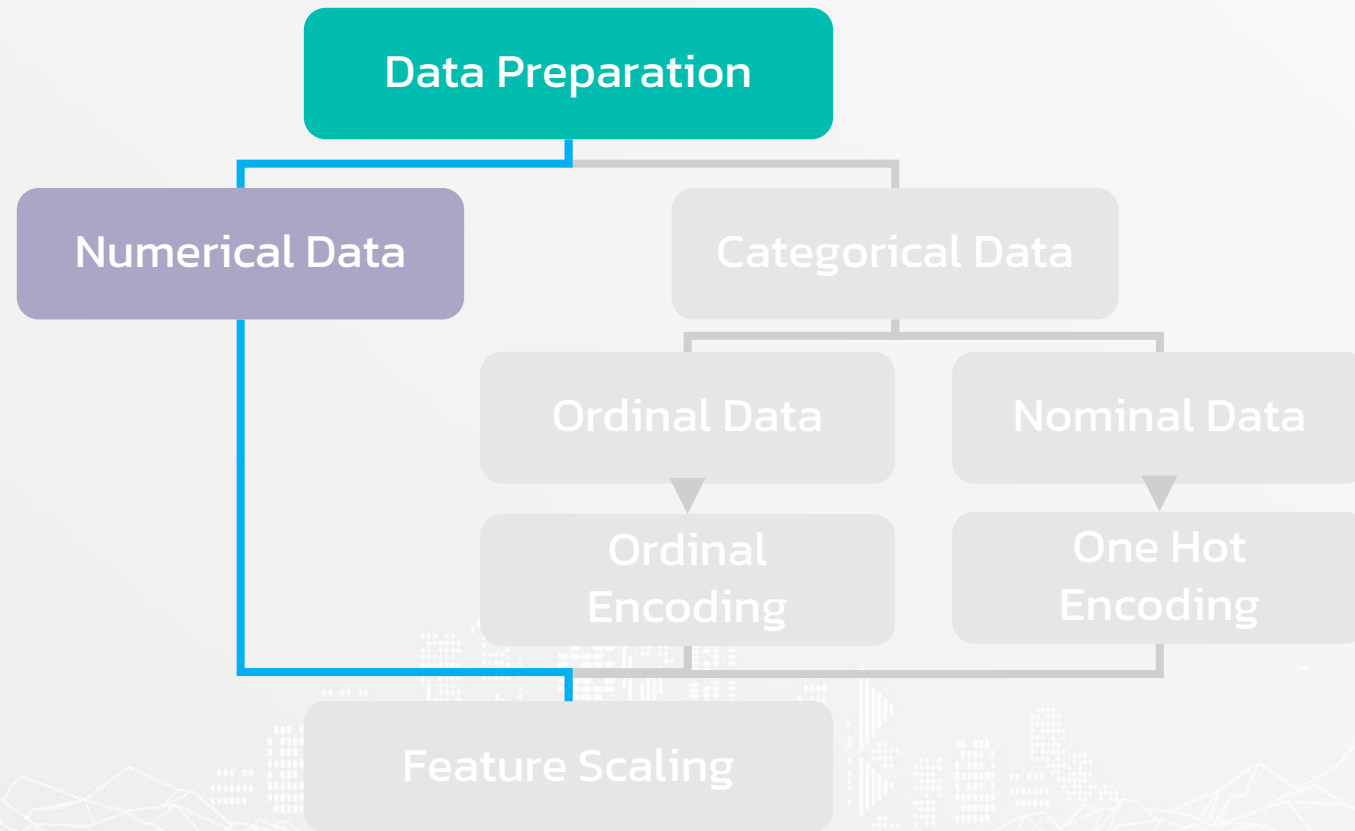
Target

- TARGET : 1 = เป็นบุคคลเสี่ยง, 0 = เป็นบุคคลไม่เสี่ยง

What we learn from this project?



Data Preparation





02. CREDIT CARD



credit_card_dataset.csv



credit_card_mc.ipynb



credit_card_md.ipynb



credit_card_model.pickle

Linear Discriminant Analysis

**What is Linear
Discriminant
Analysis?**

Assumption

**Data for Linear
Discriminant
Analysis**

Pros & Cons

Workshop