

CREDIT RATING

Presented by: Poom Maneechai



OVERVIEW

- Purpose
- Data Description
- Data Preparation
- Model Implementation
- Conclusion





PURPOSE

- จำแนกระดับความน่าเชื่อถือของบริษัท (Rating)
- เปรียบเทียบความแม่นยำของโมเดลที่ใช้ในการจำแนกประเภทของระดับความน่าเชื่อถือ



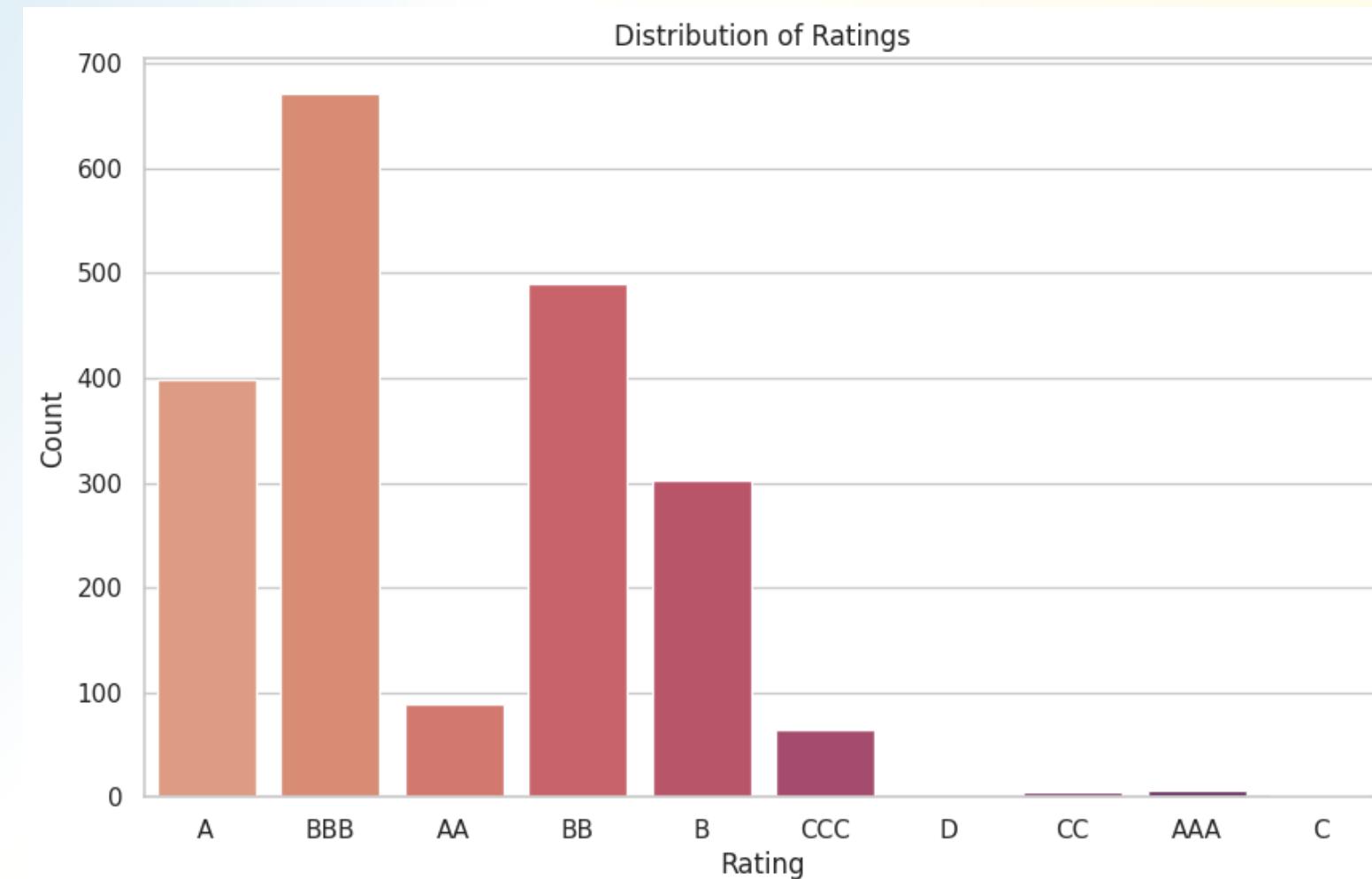
DATA DESCRIPTION

ขนาดข้อมูล

- 2029 Samples
- 31 parameters

Target Classes 10 ประเภท

- AAA
- AA
- A
- B
- BB
- BBB
- CCC
- CC
- C
- D



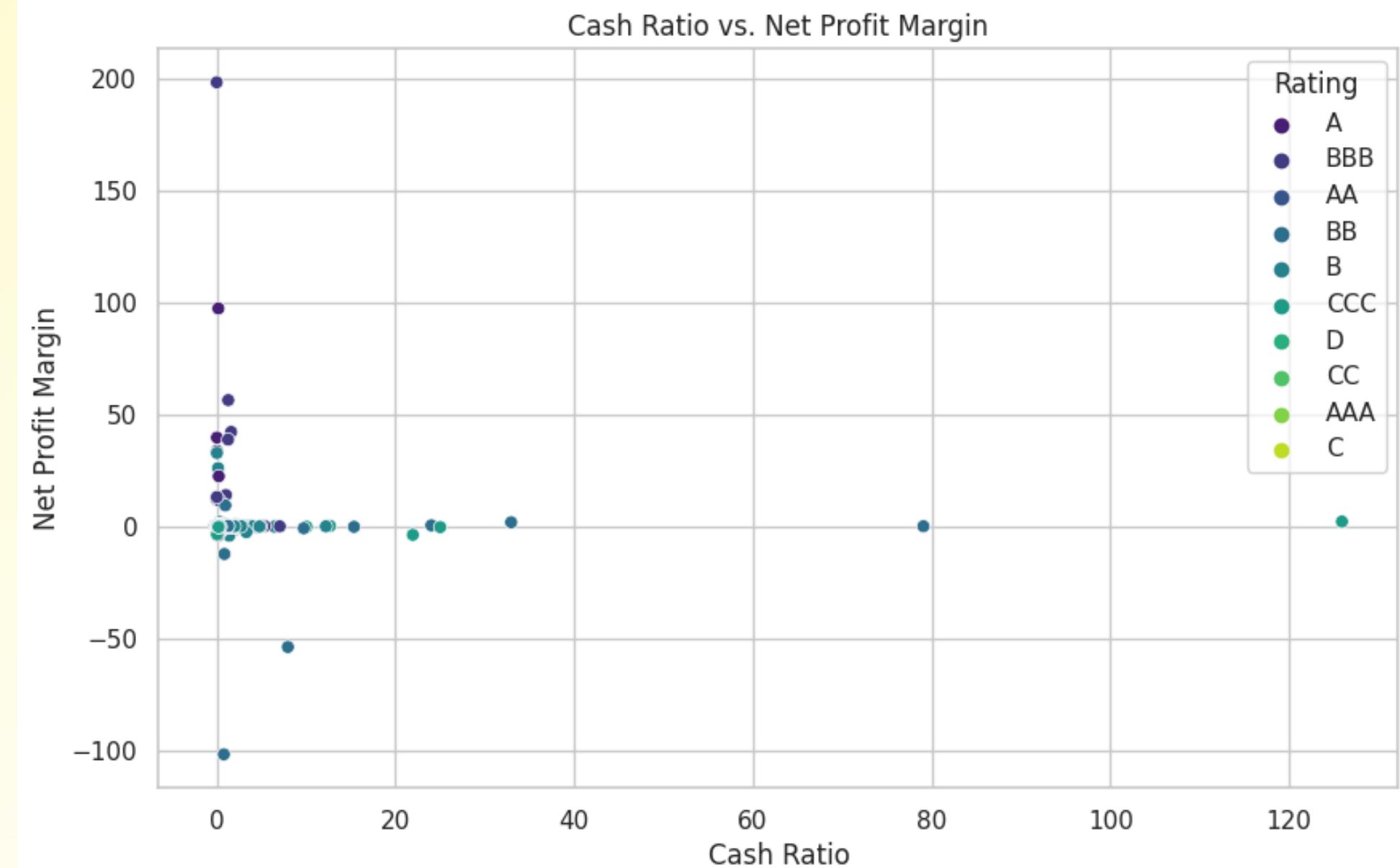
จำนวนข้อมูลแต่ละประเภท ของค่าเป้าหมาย หรือ Rating

EXAMPLE

Cash Ratio

And

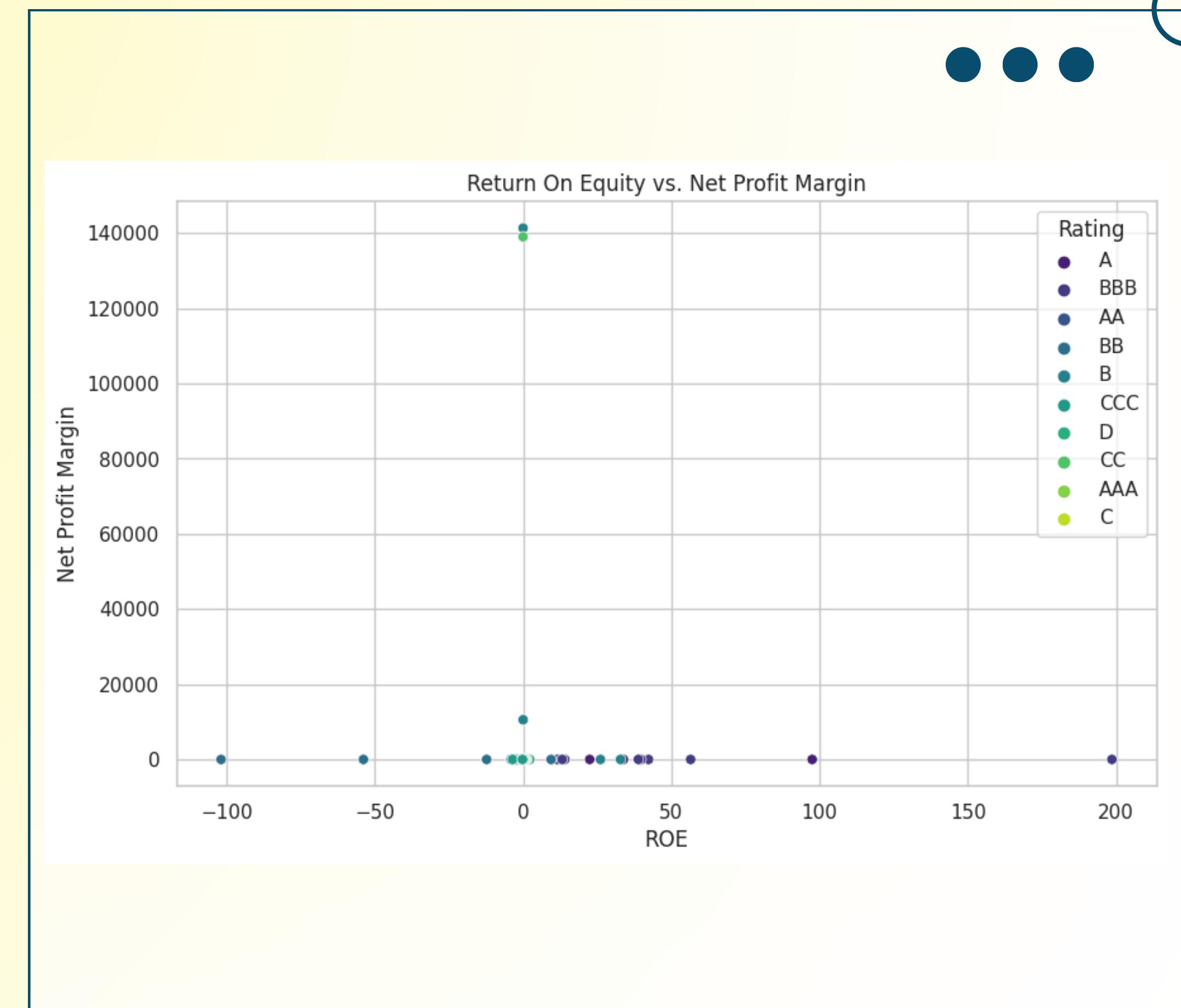
Net Profit Margin





EXAMPLE

Return On Equity
And
Net Profit Margin





DATA PREPARATION

Grouping

จัดกลุ่มข้อมูลใหม่ ให้เหมาะสม
กับปริมาณข้อมูล

Feature Interactions

สร้างฟีเจอร์ที่เหมาะสม โดย
การตั้งเงื่อนไขกับตัวฟีเจอร์
เพื่อหาความสัมพันธ์

Feature Aggregation

รวมกลุ่มของข้อมูลเข้ามาด้วย
กัน เพื่อนำมาหาค่าทาง
คณิตศาสตร์

Label Encoding

แปลงข้อมูลที่ไม่ได้เป็นตัวเลข
ให้อยู่ในรูปของตัวเลข

Feature Scaling

ปรับสเกลของข้อมูล เพื่อลด
การเหวี่ยงของตัวแปรที่มีค่า
มาก

Feature Selection

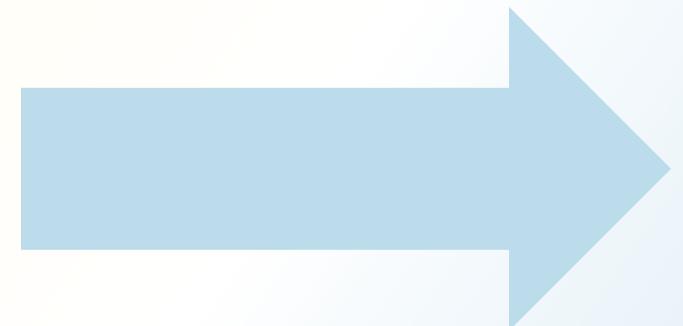
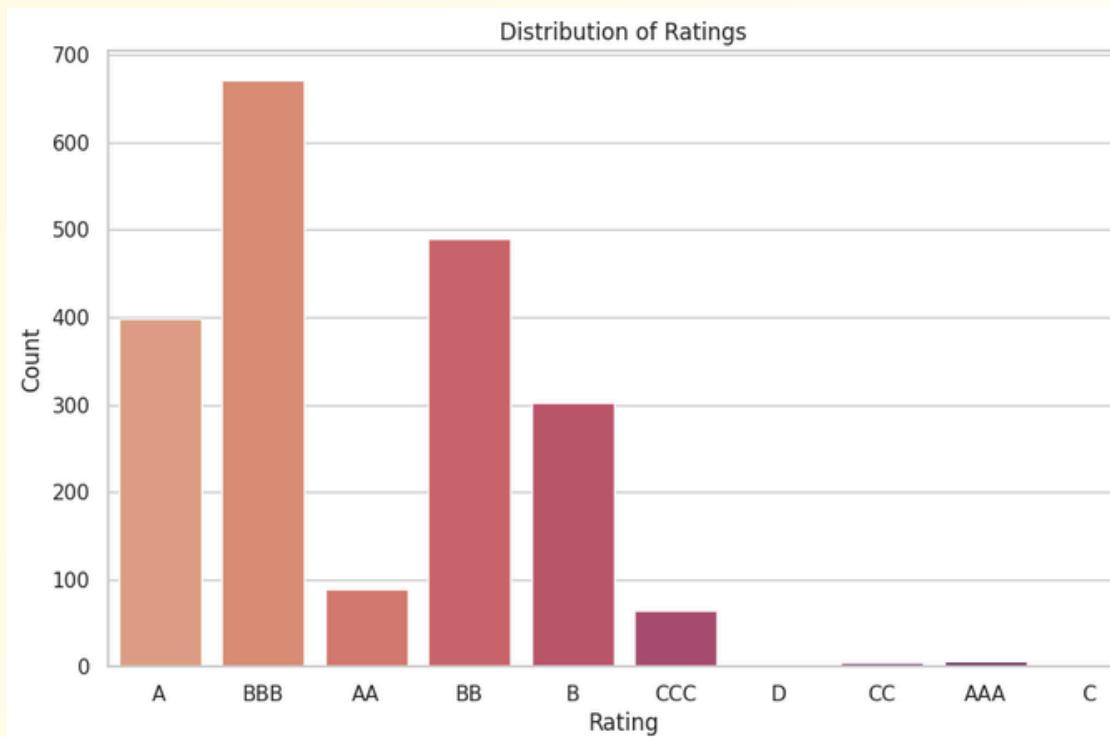
เลือกฟีเจอร์ที่เหมาะสม โดยใช้
Pearson Correlation



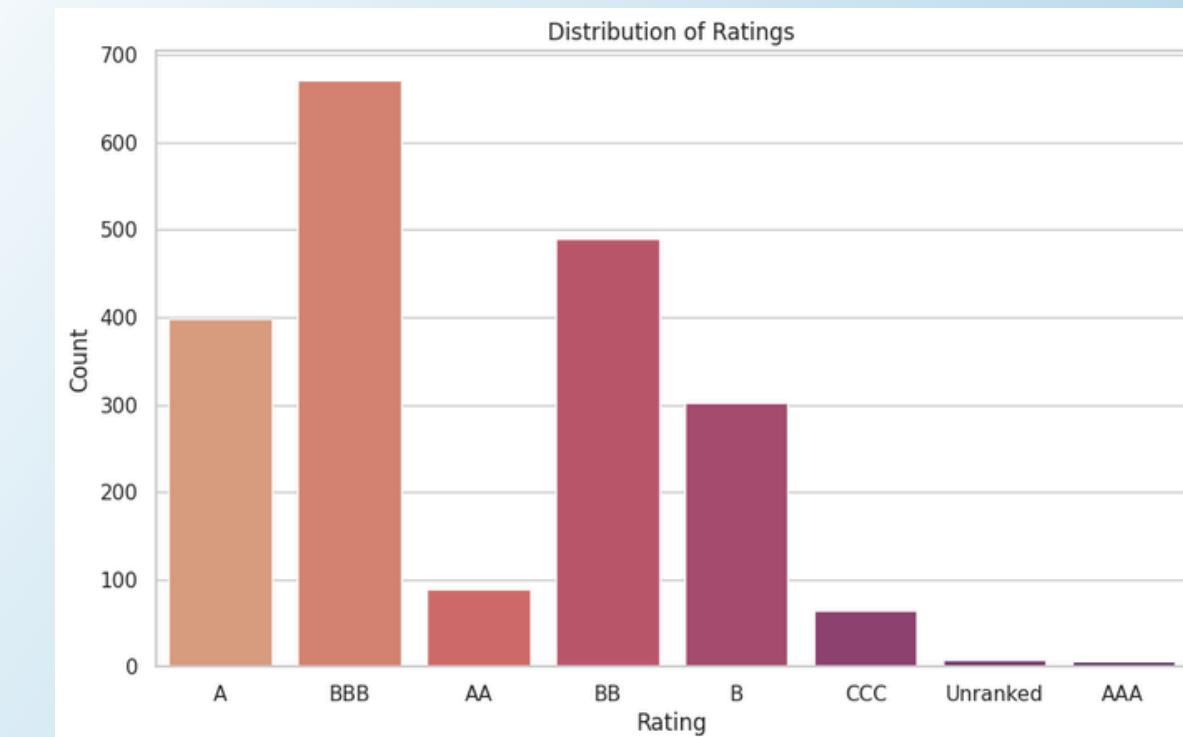
DATA PREPARATION

Grouping

Before



After



ก่อนจัดกลุ่ม

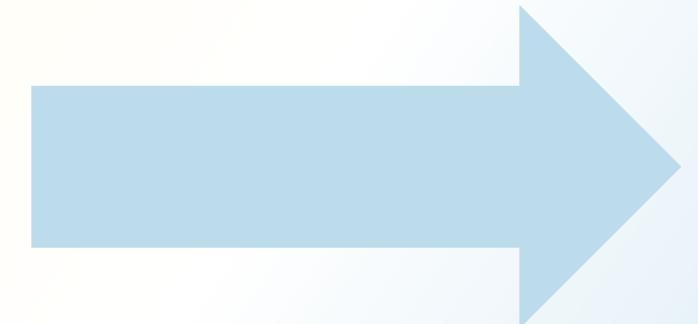
หลังจัดกลุ่ม โดยรวม CC,C,D เข้าเป็นกลุ่มเดียวกัน (Unrank)



DATA PREPARATION

Label Encoding

String



Integer

- ตัวอักษร
- ข้อความ

- 0,1,2...,n

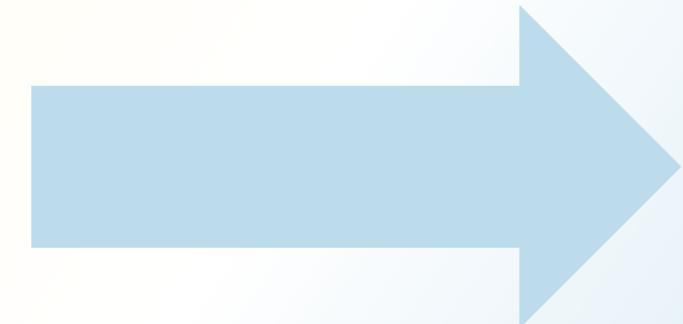


DATA PREPARATION

Feature Interaction

Create Feature

- สร้างฟีเจอร์จาก
ROE



New Features

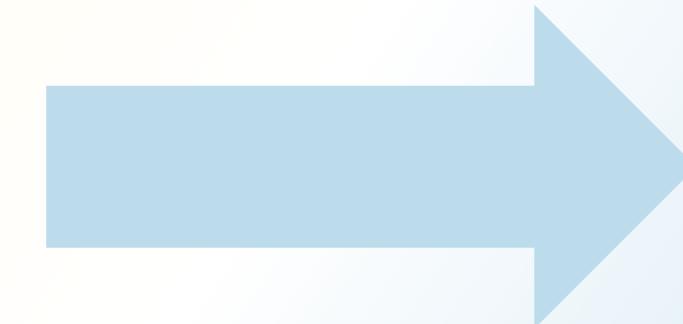
- isRoe15
- roe_quantile



DATA PREPARATION

Feature Aggregation

Aggregation



- Net Profit Margin

New Features

- Max
- Min
- Mean
- Standard deviation
- Different



DATA PREPARATION

Feature Scaling

Standard Scalar



Scaled Data

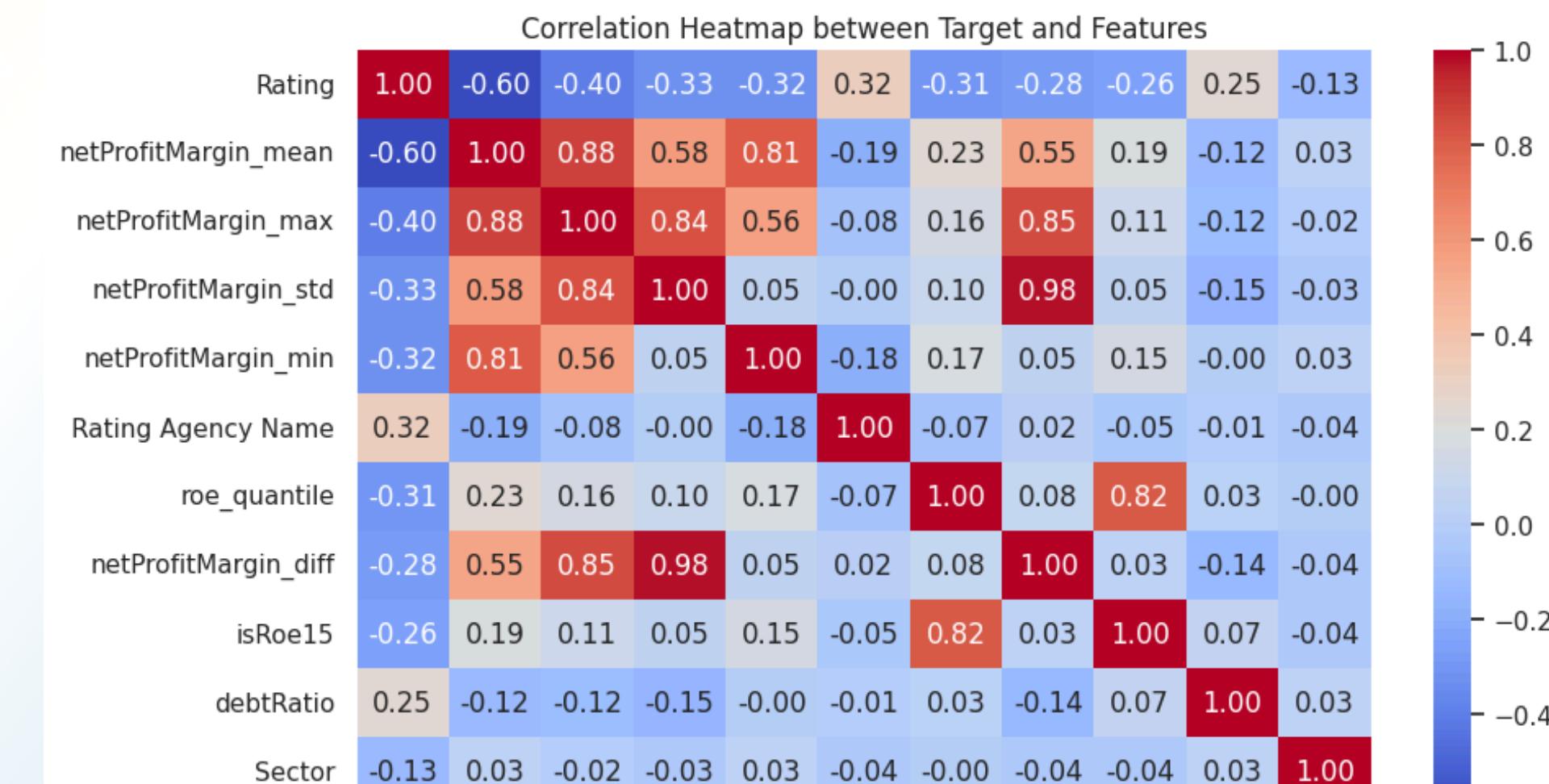
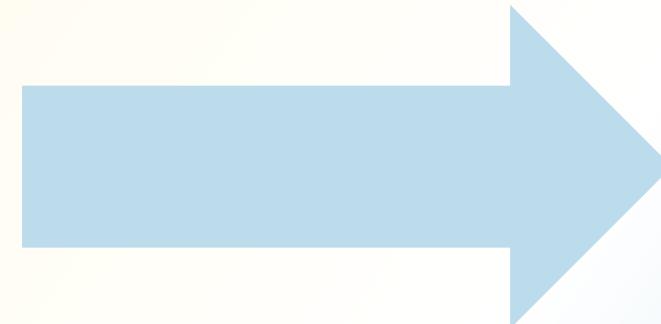


DATA PREPARATION

Feature Selection

Pearson Correlation

- เลือกฟีเจอร์ที่ค่ามากกว่า 0.1

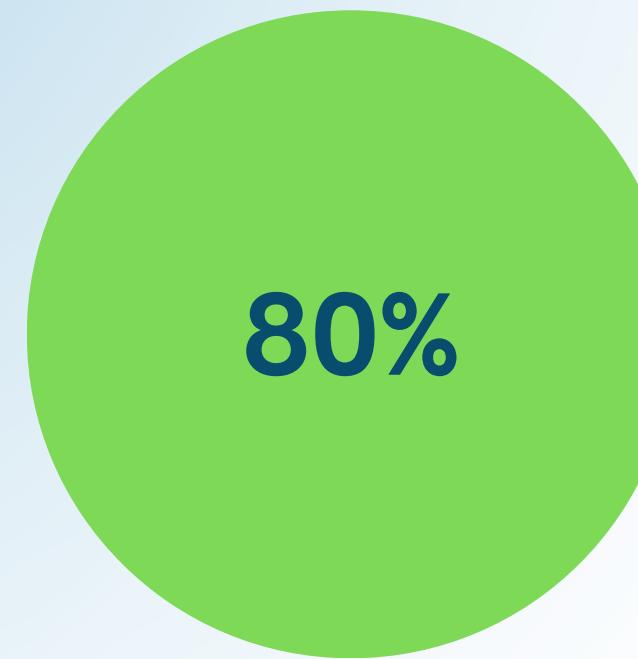


Rating netProfitMargin_mean netProfitMargin_max netProfitMargin_std netProfitMargin_min Rating Agency Name roe_quantile netProfitMargin_diff isRoe15 debtRatio Sector



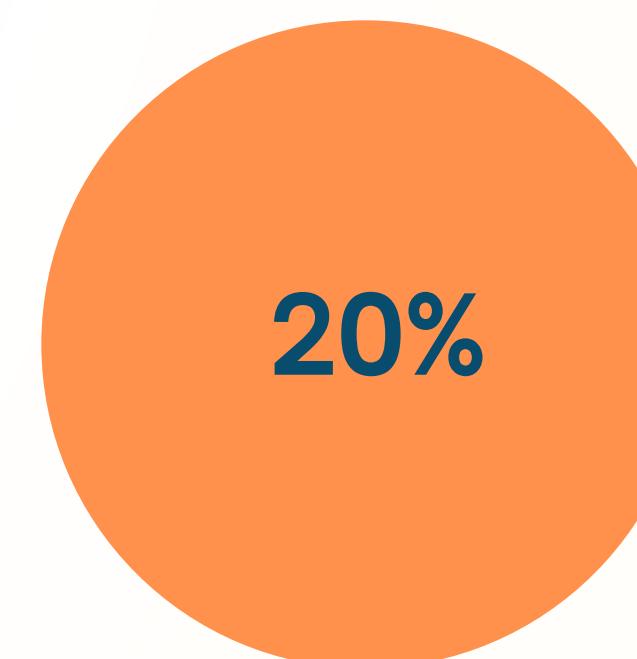
MODEL

...



Train Set

1623 sample



Test Set

406 samples

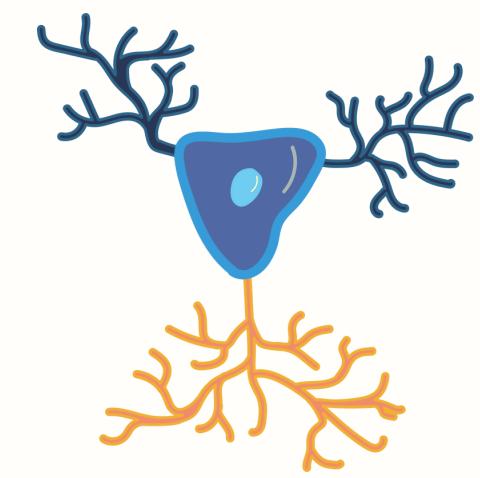
MODEL



Random Forest
SMOTE+ENN



Random Forest
Weight Classes



Deep Neuron
Network





MODEL

Random Forest SMOTE+ENN

SMOTE

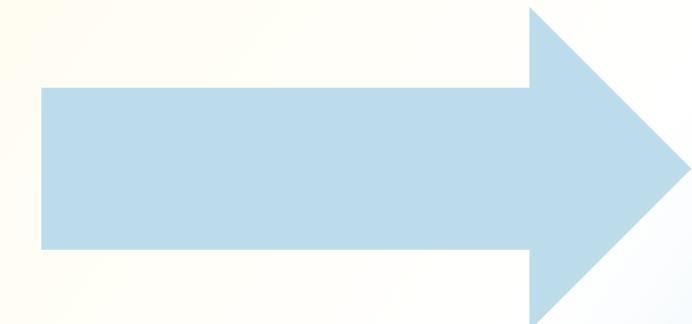
- Oversampling

ENN

- Undersampling

New Train Data

- 4195 samples from 1623 samples



Testing Results

	Classification Report:				
	precision	recall	f1-score	support	
0	1.00	1.00	1.00	1	
1	1.00	1.00	1.00	18	
2	1.00	1.00	1.00	80	
3	1.00	1.00	1.00	134	
4	1.00	1.00	1.00	98	
5	1.00	1.00	1.00	60	
6	1.00	1.00	1.00	13	
7	1.00	1.00	1.00	2	
accuracy				1.00	406
macro avg	1.00	1.00	1.00	1.00	406
weighted avg	1.00	1.00	1.00	1.00	406



MODEL

Random Forest Weight Classes

Weight Classes

```
weights = {  
    0:10,  
    1:1,  
    2:1,  
    3:1,  
    4:1,  
    5:1,  
    6:1,  
    7:10  
}
```



Testing Results

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	18
2	1.00	1.00	1.00	80
3	1.00	1.00	1.00	134
4	1.00	1.00	1.00	98
5	1.00	1.00	1.00	60
6	1.00	1.00	1.00	13
7	1.00	1.00	1.00	2
accuracy			1.00	406
macro avg	1.00	1.00	1.00	406
weighted avg	1.00	1.00	1.00	406



MODEL

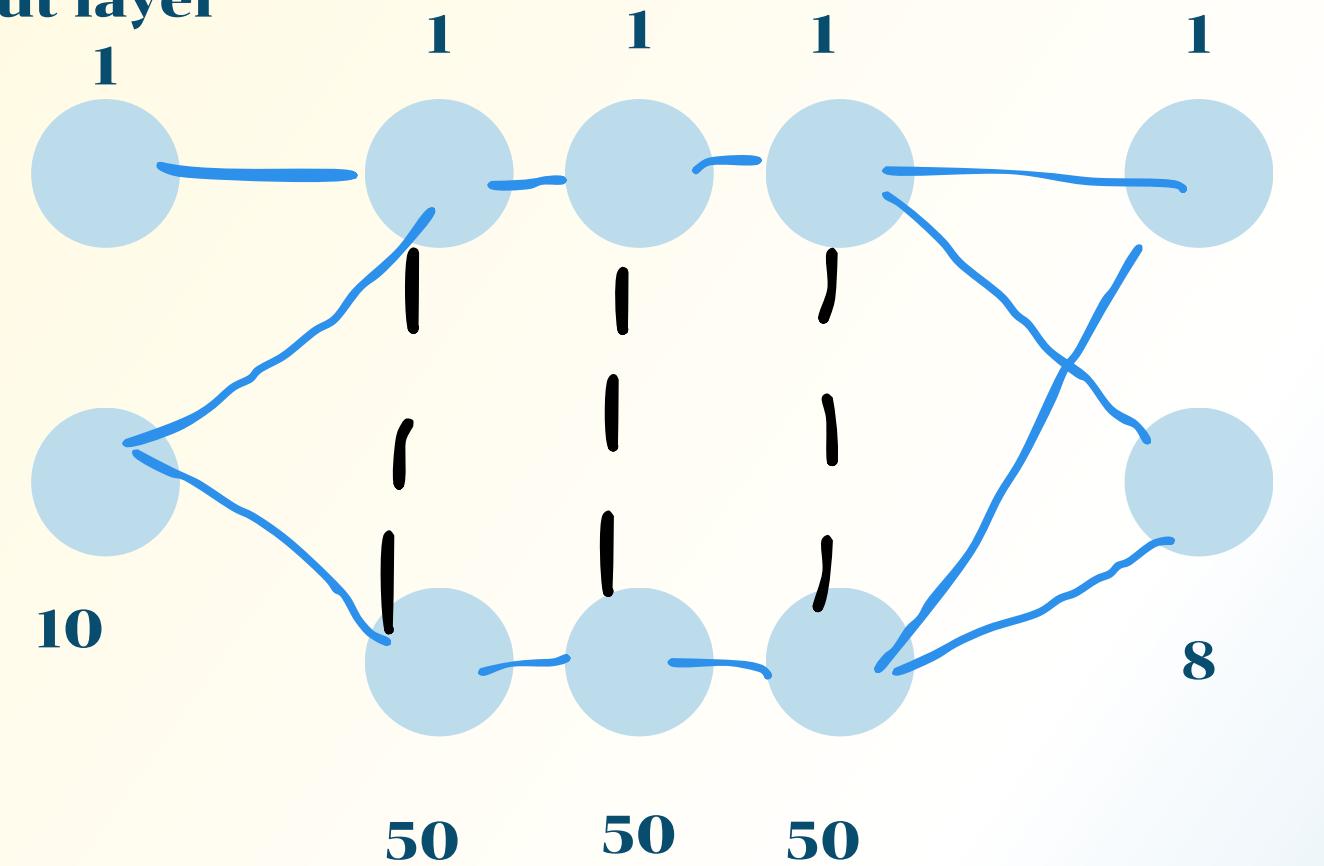
Deep Neuron Network



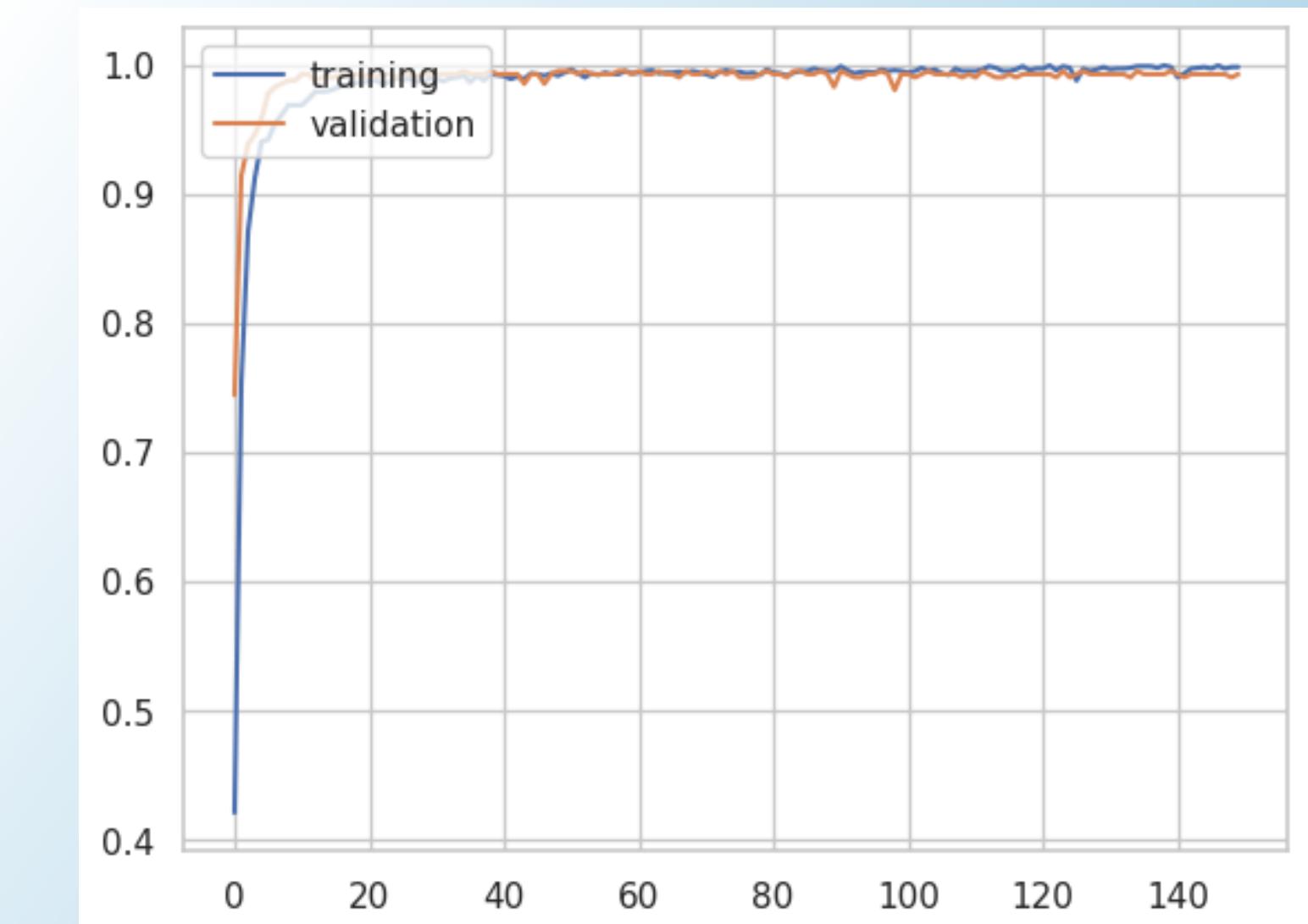
input layer

hidden layer

output layer



validation





MODEL

Deep Neuron Network



Test Result

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	18
2	1.00	1.00	1.00	80
3	1.00	1.00	1.00	134
4	1.00	1.00	1.00	98
5	1.00	1.00	1.00	60
6	0.86	0.92	0.89	13
7	0.00	0.00	0.00	2
accuracy			0.99	406
macro avg	0.86	0.87	0.86	406
weighted avg	0.99	0.99	0.99	406

RESULTS

	Test score	Validation score	Precision	Recall	F1-Score
Random Forest SMOTE+ENN	1	1	1	1	1
Random Forest Weight Classes	1	1	1	1	1
Deep Neural Network	0.99	0.99261	0.86	0.87	0.8649711

Thank you
for your
Attention

