



Best Paper Award: Technology - Interdisciplinary Track

โครงการสนับสนุนทุนวิจัยด้านตลาดทุน ประจำปี 2567/2568

หัวข้อ การวิเคราะห์โดยใช้การประมวลผลภาษาธรรมชาติ (NLP) จากข่าวที่เผยแพร่สาธารณะ
ในการประเมินความเสี่ยงในการผิดนัดชำระหนี้ของหุ้นกู้ในประเทศไทย

โดย

ธนกร โฆษณิขการ ระดับปริญญาโท หลักสูตรวิศวกรรมการเงิน คณะพัฒนาการเศรษฐกิจ สถาบันบัณฑิตพัฒนบริหารศาสตร์
อาจารย์ที่ปรึกษา รศ.ดร.ปริญดา สุขเจริญสิน

11 มิถุนายน 2568



ปัญหาและแรงจูงใจ



ความท้าทายที่สำคัญที่ตลาดทุนต้องเผชิญ

01

หุ้นกู้เป็นพื้นฐานของการขยายตัวทางเศรษฐกิจ

- แนวทางในการระดมทุนที่สำคัญต่อตลาดทุนไทย
- บางบริษัทอาจไม่สามารถชำระหนี้ได้ตามสัญญา

02

จำนวนบริษัทจดทะเบียนมาก

- เฝ้ารวังมากกว่า 800 บริษัทจดทะเบียน ต้องใช้พื้นที่และการจัดการมาก
- ตลาดขยายตัวและ ซับซ้อนทำให้อาจจะพลาดปัญหาที่แท้จริง

03

ประเมินผลพร้อม ๆ กันทุก ๆ บริษัท

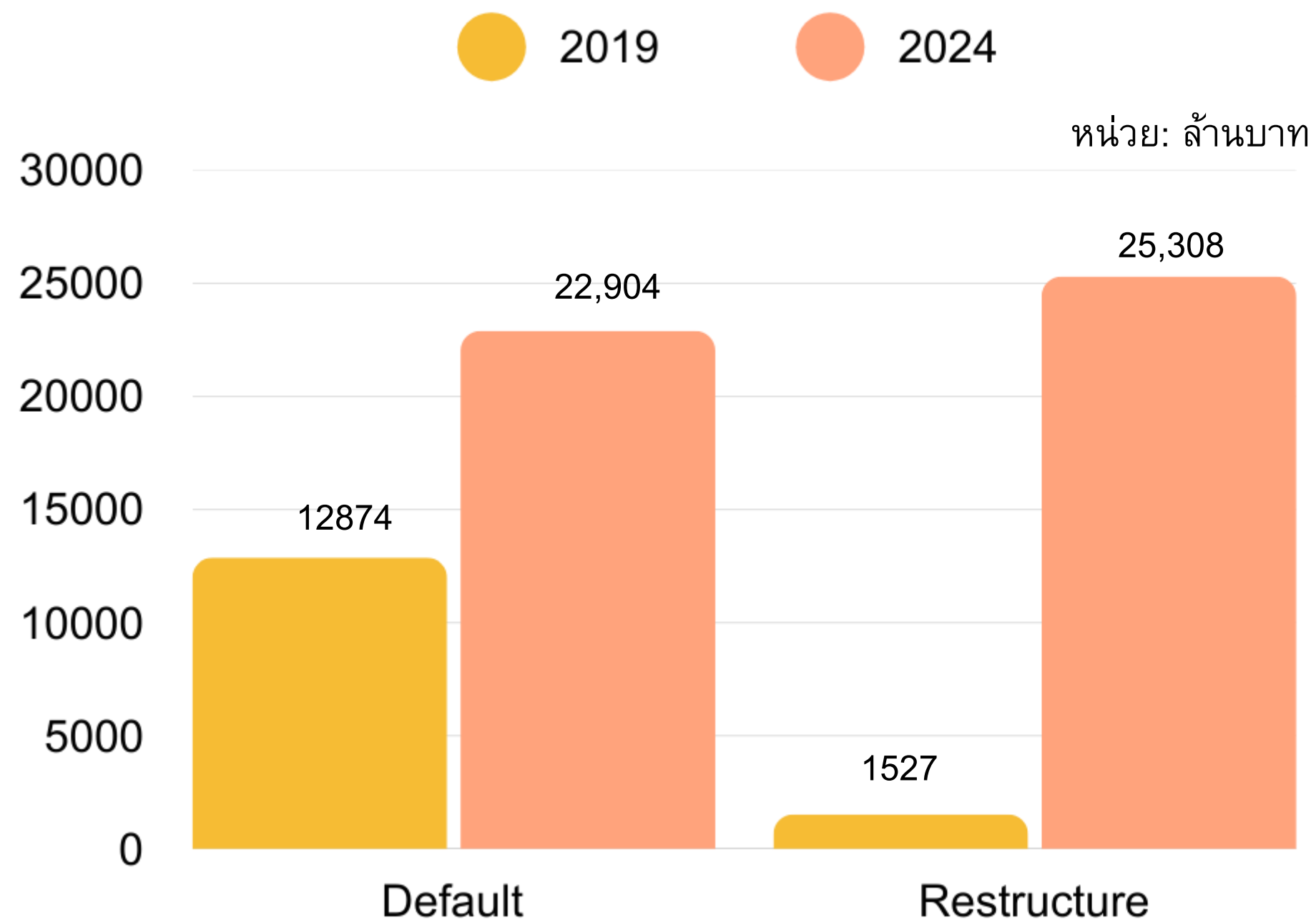
- หน่วยงานกำกับดูแลต้องประเมินบริษัทอย่างสม่ำเสมอเพื่อให้สามารถตรวจจับสัญญาณความเสี่ยงทางการเงินได้ตั้งแต่ระยะแรก
- การประเมินความเสี่ยงอย่างทันที่่วงที่และแม่นยำมีความสำคัญอย่างยิ่งต่อความมั่นคงของตลาด



ปัญหาและแรงจูงใจ



มูลค่าหุ้นกู้มีปัญหาเพิ่มขึ้นอย่างมีนัยสำคัญหลังจากวิกฤตโควิด





01

พัฒนารูปแบบการ
ประเมินความเสี่ยงที่
ขับเคลื่อนด้วย AI

การผสมรวม LLMs และ ML
เพื่อวิเคราะห์ข้อมูลทางการเงิน
ที่ไม่เป็นระเบียบ
หรือ soft information

02

เสริมสร้างการทำนาย
ความเสี่ยงการผิดนัด
ชำระหนี้ของหุ้นกู้

โดยใช้การวิเคราะห์อารมณ์
ความสัมพันธ์ของหน่วยงาน
และเรื่องราวทางการเงิน

03

หาตัวบ่งชี้ที่ส่งผลต่อ
การผิดนัดชำระหนี้ทั้ง
เชิงบวกและลบ

สำหรับหน่วยงานกำกับดูแล เพื่อ
ส่งเสริมการบริหารความเสี่ยงเชิงรุก
จากตัวบ่งชี้ที่ได้จากงานวิจัย

04

สนับสนุนกรอบการ
ติดตามความเสี่ยงที่มี
ความยืดหยุ่น

สำหรับตลาดทุนไทย เพื่อ
ปรับปรุงการตรวจจับความเสี่ยง
ในระยะเริ่มต้นและเสริมสร้าง
ความมั่นใจของนักลงทุน



ผลลัพธ์ที่คาดหวัง



01

โมเดลการทำนายความเสี่ยงที่ขับเคลื่อนด้วย AI

- โมเดลที่สามารถวิเคราะห์ข้อมูลเชิงคุณภาพได้
- โมเดลที่สามารถใช้งานได้อย่างมีประสิทธิภาพมากกว่า วิธีการดั้งเดิม

02

การตรวจจับสัญญาณความเสี่ยงทางการเงินในระยะแรกที่ดีขึ้น

- การผสมรวมข่าวสาร ความรู้ลึกของตลาด และข้อมูลที่เปิดเผยของบริษัท

03

ได้ตัวบ่งชี้ของสัญญาณในการเกิดปัญหาของหุ้นกู้

- สัญญาณทั้งเชิงบวกและลบทำให้หน่วยงานรับผิดชอบตรวจสอบจับความเสี่ยงได้อย่างทันท่วงที



โมเดลสถิติแบบดั้งเดิม

Machine Learning

Natural Language Processing (NLP)

Large Language Models (LLMs)

Merton model
Merton (1974)

Regression models
Kim et al. (2020)

TF-IDF
George & Murugesan, (2024)

Multimodal financial risk modeling
Cao et al. (2024)

Distance-to-Default
Black & Scholes (1980)

XGBoost
Zhang & Chen (2021)

BERTopic
Tang et al. (2024)

LLMs for financial statement analysis
Kim et al. (2024)

Corporate bond
default history
Giesecke et al. (2010)

SMOTE
Zhang & Chen (2021)

Text in financial reports
Tsai & Wang (2016)

FinBERT: Finance-specific LLM
Liu & Jia (2024)

งานวิจัยนี้

Machine Learning

Natural Language Processing (NLP)

Large Language Models (LLMs)



การออกแบบวิจัย



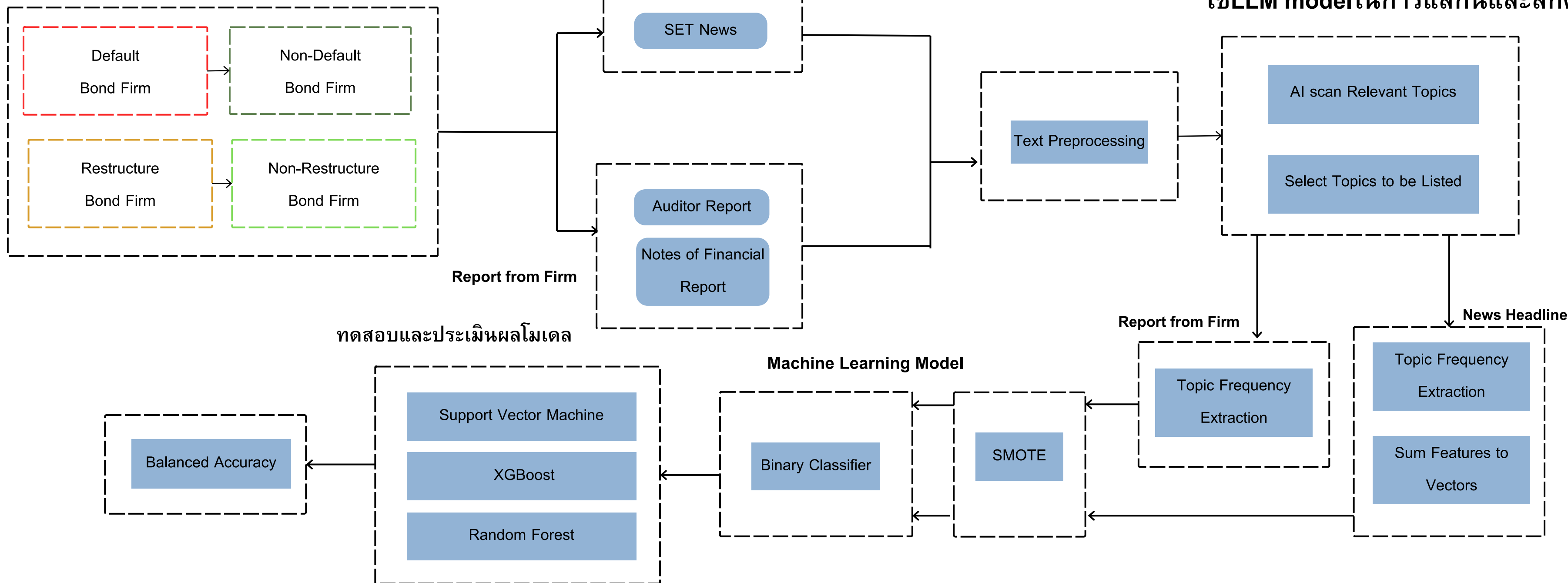
Research Framework

รวบรวมข้อมูลบริษัทที่มีปัญหา
และคัดเลือกบริษัทเปรียบเทียบ

กระบวนการเก็บรวบรวมและดึงข้อมูลอัตโนมัติ

Preprocessing ข้อมูล

ใช้LLM modelในการสแกนและสกัดvector





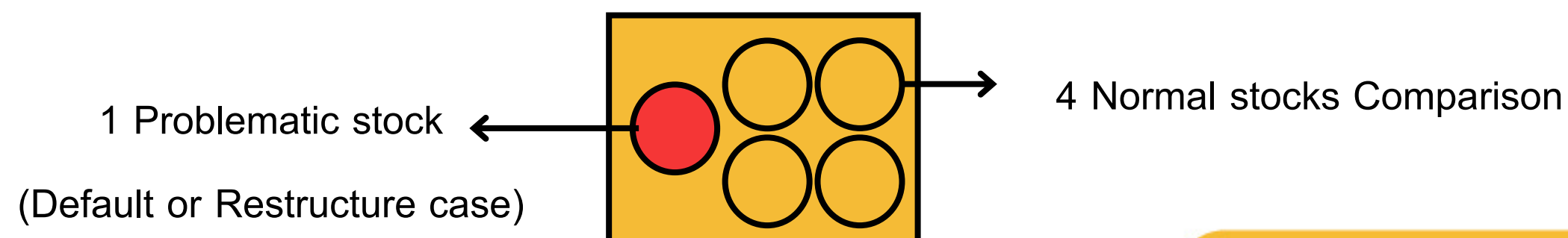
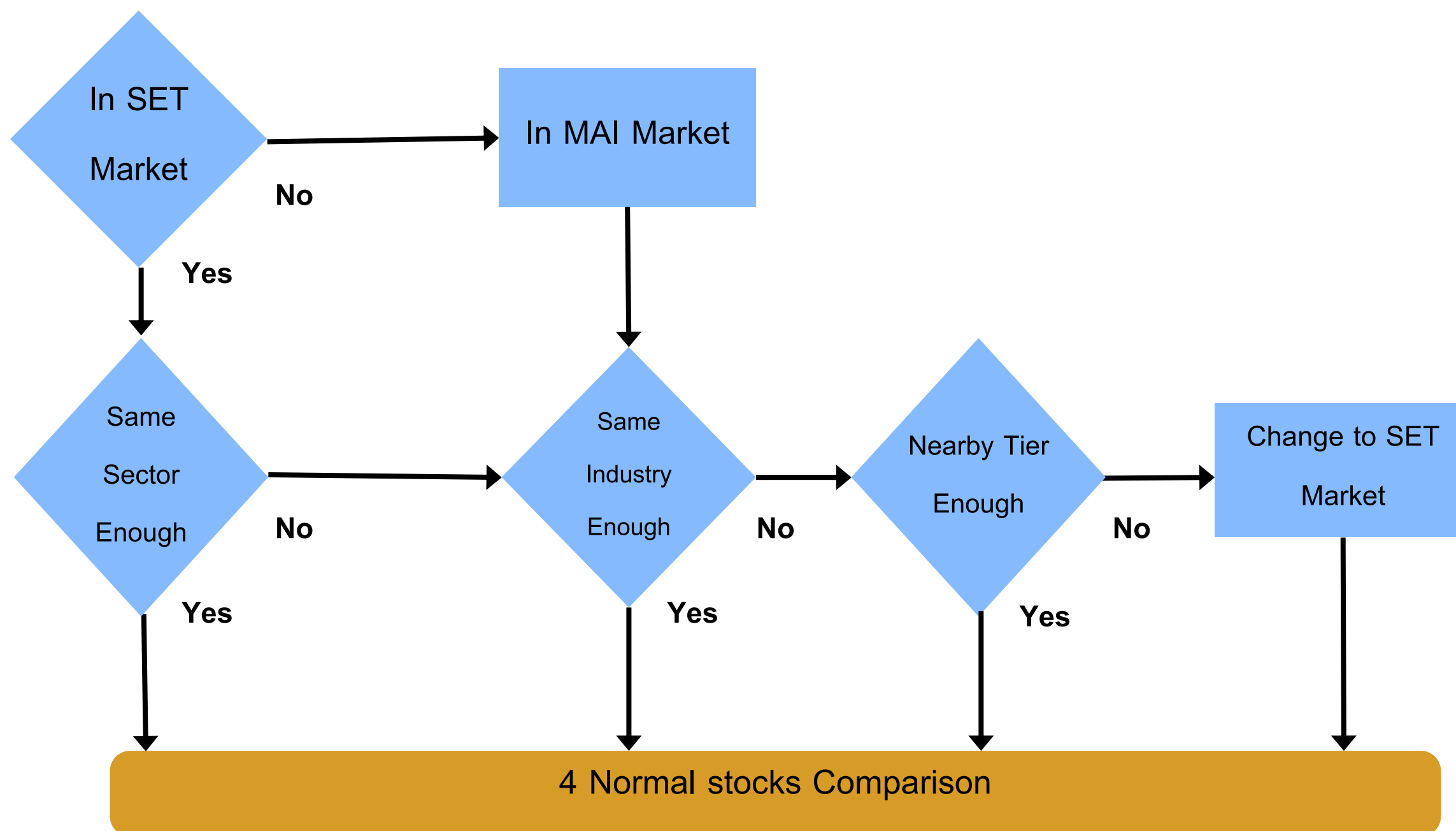
การคัดเลือกข้อมูล



วิธีเลือกหุ้นและเกณฑ์

SET ลำดับชั้นตาม market cap.
มากกว่า 100,000 ล้านบาท,
30,000–100,000 ล้านบาท,
10,000–30,000 ล้านบาท,
3,000–10,000 ล้านบาท,
และต่ำกว่า 3,000 ล้านบาท

MAI ลำดับชั้นตาม market cap.
แบ่งเป็น 2 ชั้น
มากกว่า 1,500 ล้านบาท และ
ต่ำกว่า 1,500 ล้านบาท





การคัดเลือกข้อมูล



แหล่งข้อมูลและกระบวนการดึงข้อมูลอัตโนมัติ

แหล่งข้อมูล

01



02



กระบวนการดึงข้อมูลอัตโนมัติ

01



Instant Data Scraper
Chrome Extension



การออกแบบวิจัย



การแยกความถี่หัวข้อข่าว

หัวข้อ

1

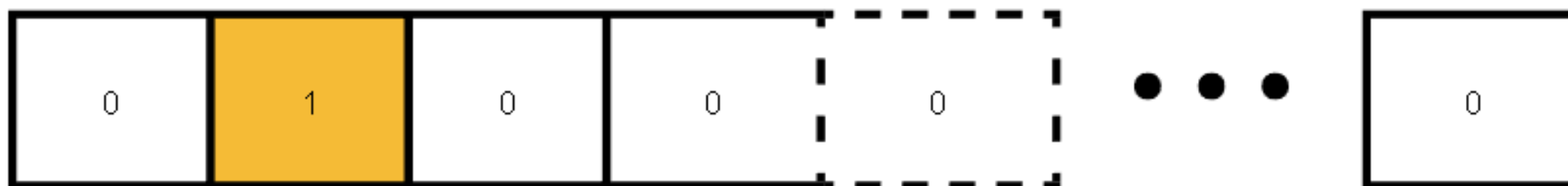
2

3

4

19

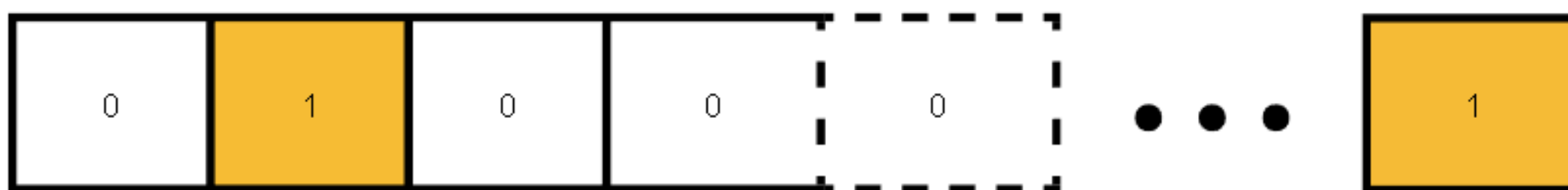
ข่าวที่ 1



--หัวข้อ A

แจ้งมติที่ประชุมคณะกรรมการของบริษัท ครั้งที่ 3/2567 เรื่อง การงดจ่ายปันผล (Topic 2)

ข่าวที่ 2



--หัวข้อ B

แจ้งมติที่ประชุมคณะกรรมการของบริษัท ครั้งที่ 4/2567 เรื่อง การงดจ่ายปันผล (Topic 2) คดีข้อพิพาท (Topic 19)

ขั้นตอนหาว่าแต่ละข่าวอยู่ในหัวข้อไหนบ้างถูกวิเคราะห์โดยใช้



1.5 flash

ได้ออกมาเป็น vector



การออกแบบวิจัย



การแยกความถี่หัวข้อข่าว
หัวข้อ 1 2 3 4 19

ข่าวที่ 1	0	1	0	0	0	• • •	0
ข่าวที่ 2	0	0	0	1	0	• • •	0
ข่าวที่ 3	0	1	0	0	0	• • •	0
ข่าวที่ 4	1	0	0	0	0	• • •	0
					• • •		
1 หุ้่น รวมจำนวนข่าว	1	2	0	1	0	• • •	0



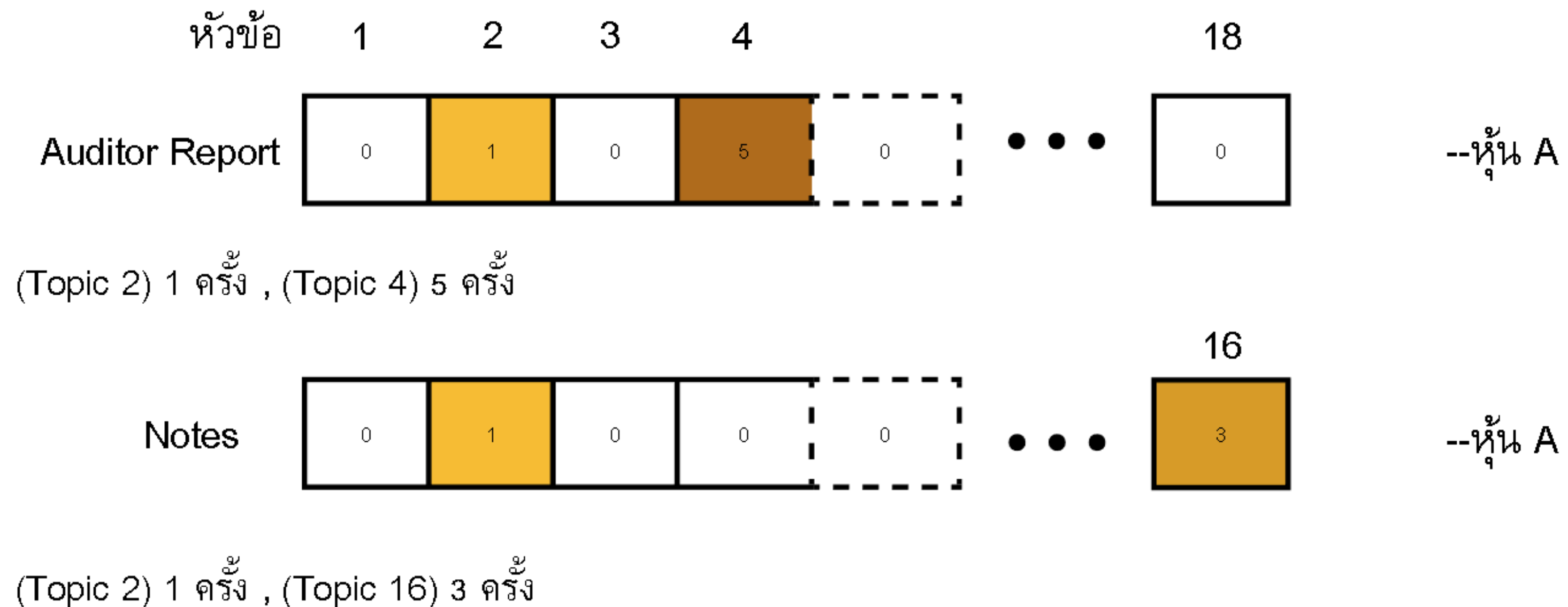
1-90 วัน หรือ
91-365 วันก่อน



การออกแบบวิจัย



การแยกความถี่หัวข้อ ในรายงานผู้สอบบัญชี(Auditor Report) และหมายเหตุรายงานงานทางการเงิน(



ขั้นตอนหาว่าแต่ละข่าวอยู่ในรายงานผู้สอบบัญชี (Auditor Report) และหมายเหตุรายงานงานทางการเงิน (Notes) มีเนื้อหาในหัวข้อไหนบ้างถูกวิเคราะห์โดยใช้

ได้ออกมาเป็น vector

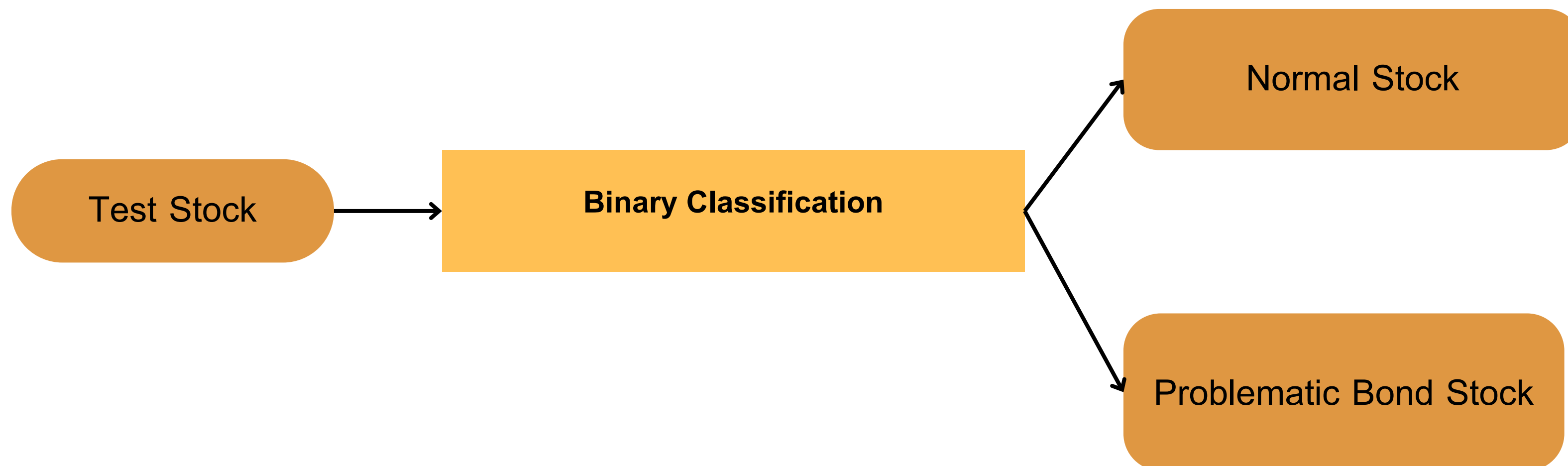
Gemini 1.5 flash



การออกแบบวิจัย



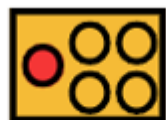
Machine Learning Binary Classification



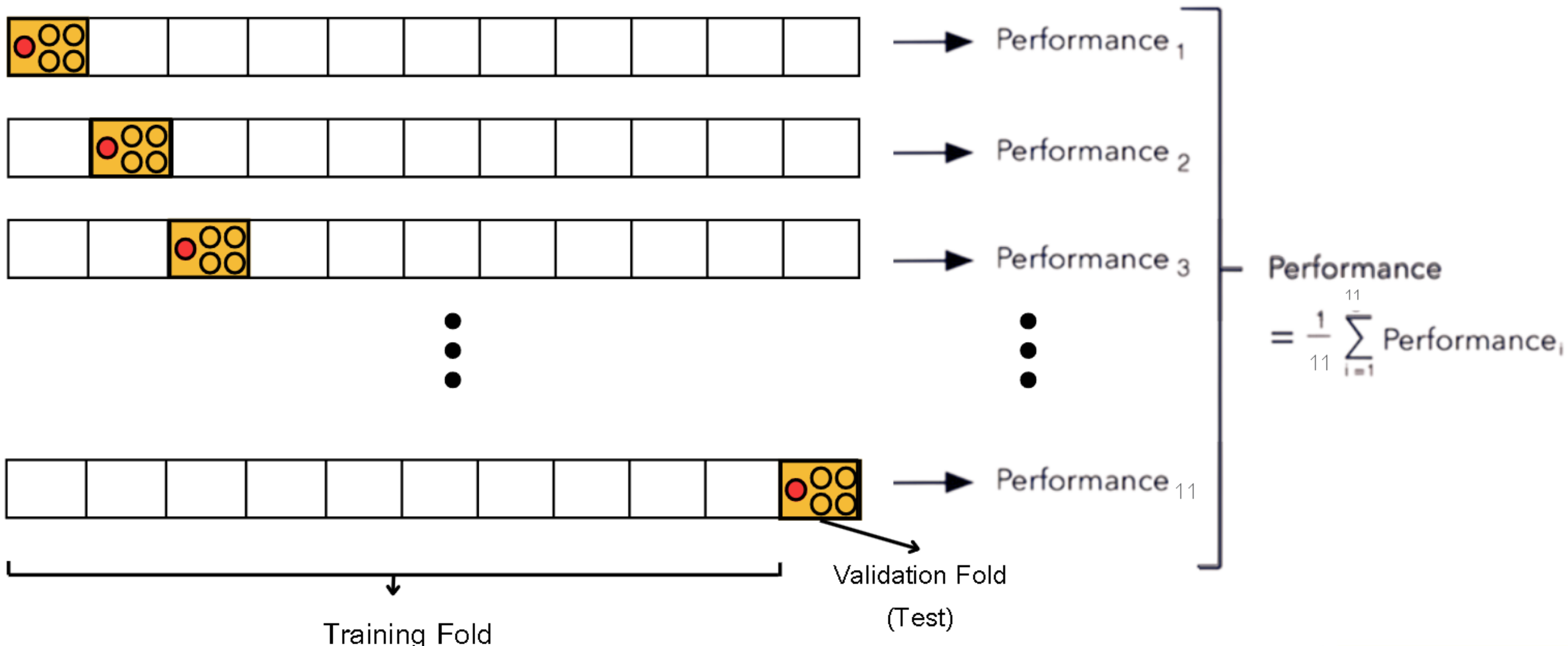


การออกแบบวิจัย

11 fold cross validation



1 บริษัทที่มีปัญหา , 4 บริษัทที่ไม่มีปัญหา



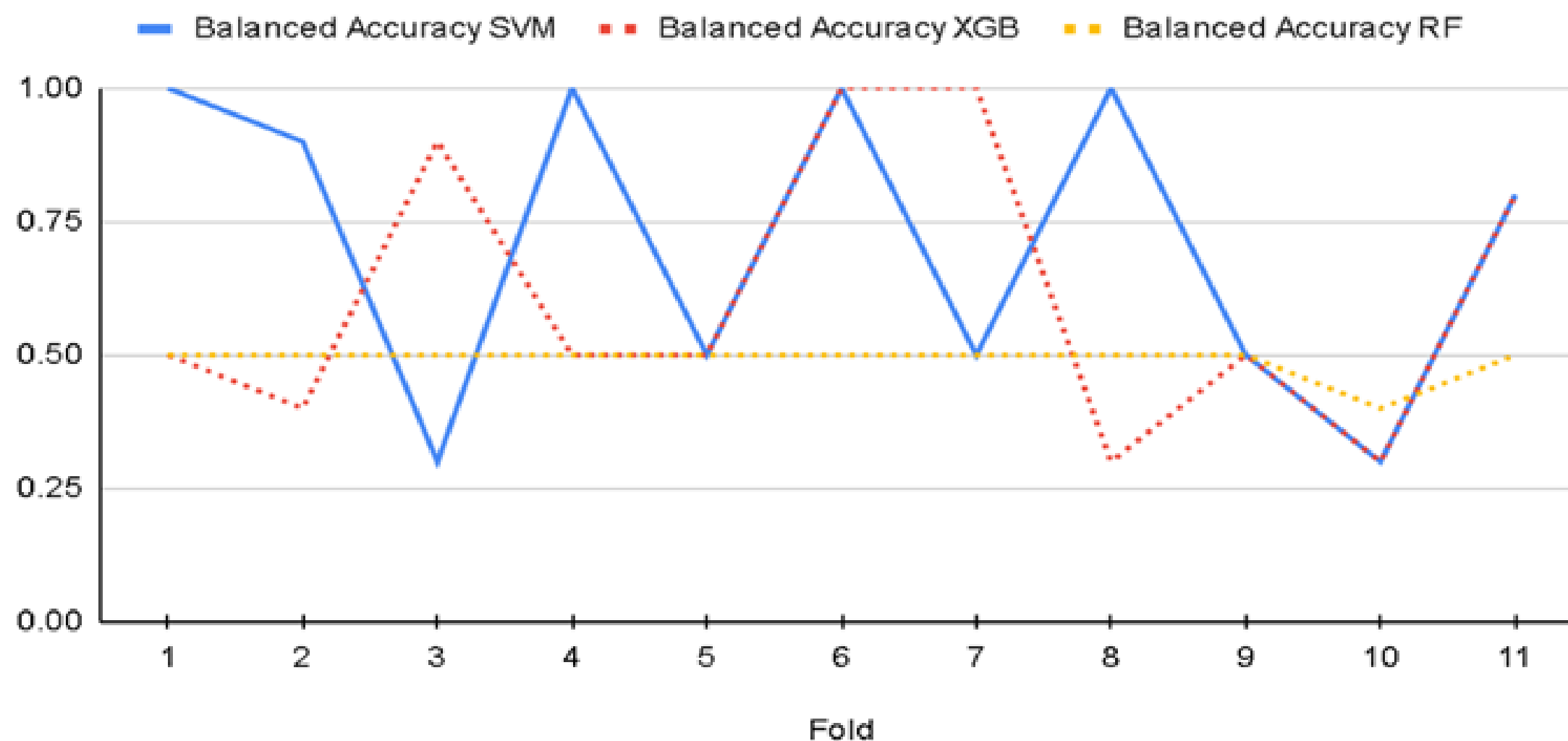


ผลการศึกษาเชิงประจักษ์



ประเมินโมเดล machine learning โดย SVM ได้ความแม่นยำสูงสุด

Balanced Accuracy SVM, XGB, RF

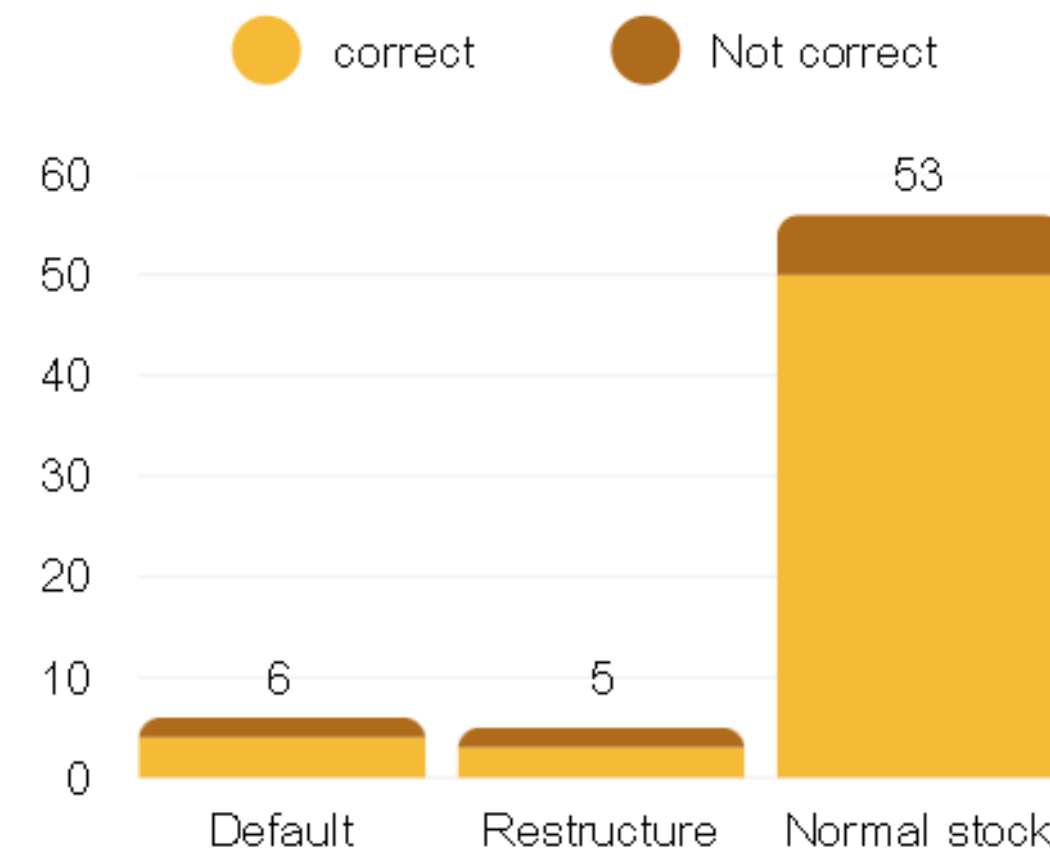


Average Balanced Accuracy

XGBoost = 0.6091

Random Forest = 0.4909

SVM = 0.7091





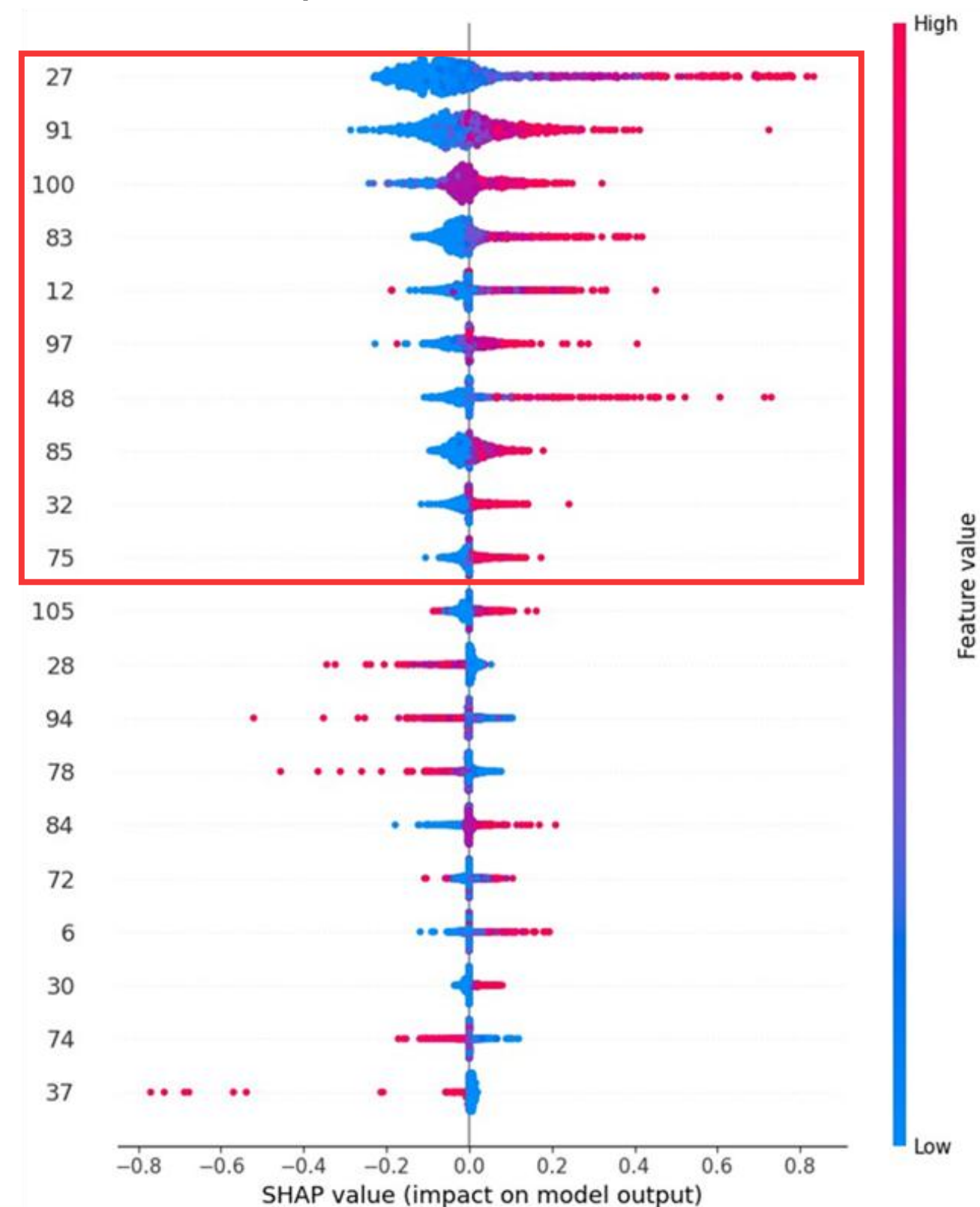
ผลการศึกษาเชิงประจักษ์



คุณสมบัติ (feature) ของข่าวจาก SET ที่มีผลกระทบเชิงบวกต่อเหตุการณ์ที่มีปัญหา

FeatureID	Key	Data
27	การเพิ่มทุน	ข่าวจากSETช่วงเวลา: 91–365 วันก่อนเกิดปัญหา
32	การเปลี่ยน CFO	ข่าวจากSETช่วงเวลา: 91–365 วันก่อนเกิดปัญหา
12	การแต่งตั้ง CEO คนใหม่	ข่าวจากSETช่วงเวลา: 1–90 วันก่อนเกิดปัญหา

จากคุณสมบัติ (feature) ทั้งหมด 19 features x 2 ช่วงเวลา = 0 - 37





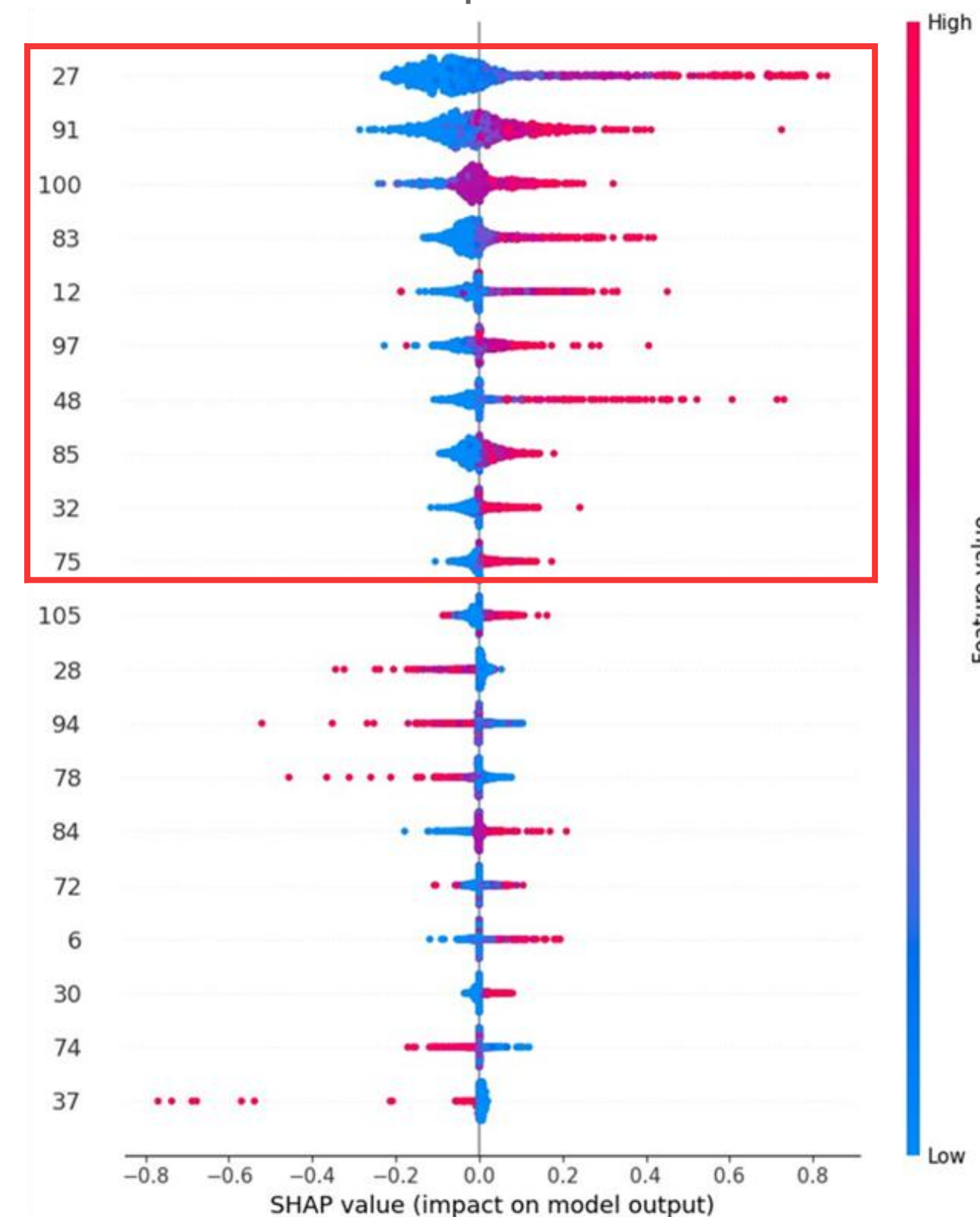
ผลการศึกษาเชิงประจักษ์



คุณสมบัติ (feature) ของรายงานผู้สอบบัญชี ที่มีผลกระทบเชิงบวกต่อเหตุการณ์ที่มีปัญหา

FeatureID	Key	Data
48	ความถี่ของคดีความทางกฎหมาย	รายงานผู้สอบบัญชี ช่วง 1 ปีก่อนเกิดปัญหา

จากคุณสมบัติ (feature) ทั้งหมด 18 features x 2 ช่วงเวลา = 38 - 73





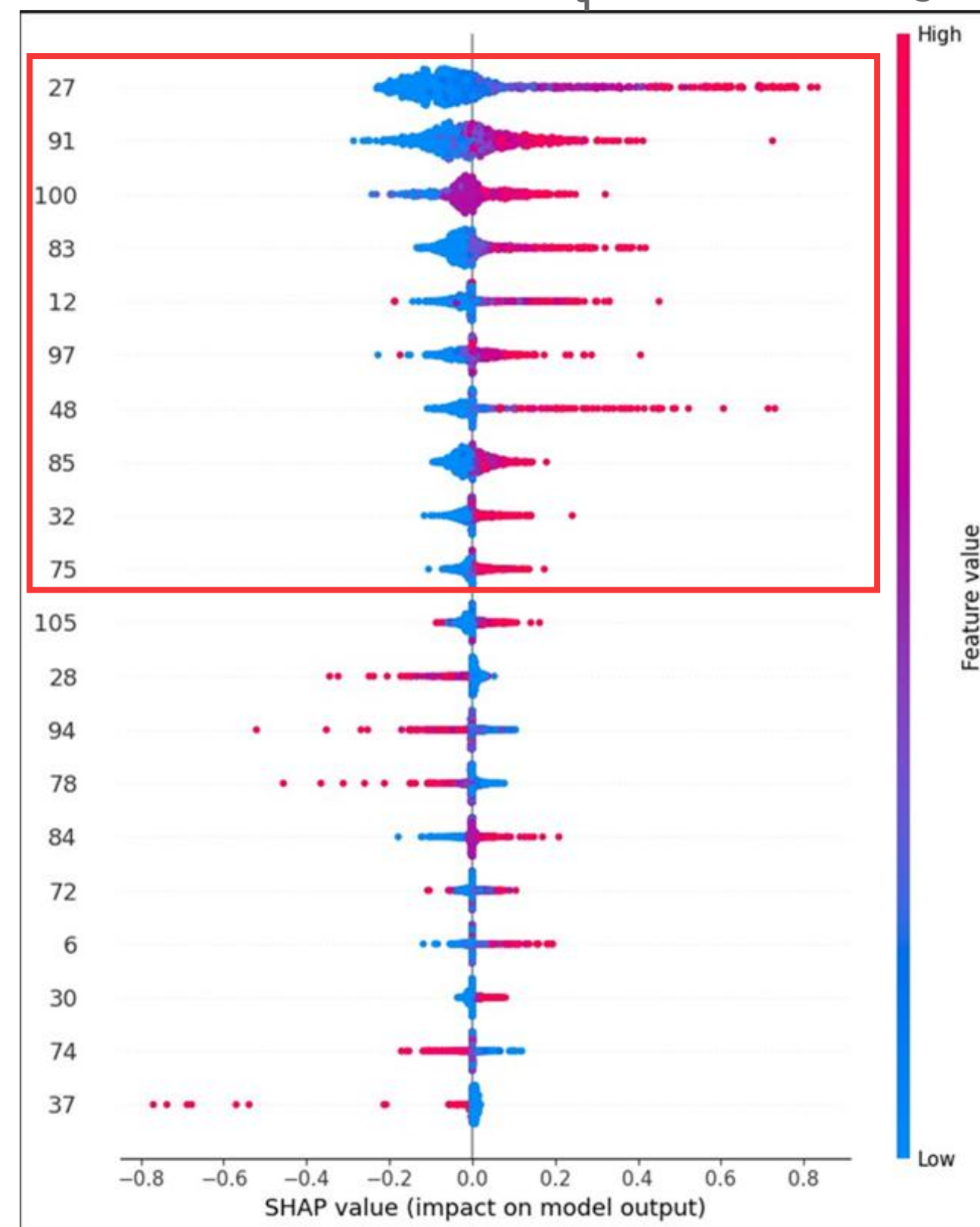
ผลการศึกษาเชิงประจักษ์



คุณสมบัติ (feature) ของหมายเหตุประกอบงบเงิน ที่มีผลกระทบเชิงบวกต่อเหตุการณ์ที่มีปัญหา

FeatureID	Key	Data
91	ความถี่ในการขายทรัพย์สิน	หมายเหตุรายงานทางการเงิน(Notes)ช่วง 2-3 ปีก่อน
75	ความถี่ในการขายทรัพย์สิน	หมายเหตุรายงานทางการเงิน(Notes)1 ปีก่อน
100	ความถี่ในการเปลี่ยนนโยบายบัญชี	หมายเหตุรายงานทางการเงิน(Notes)ช่วง 2-3 ปีก่อน
97	ความถี่ในการให้ความช่วยเหลือบริษัทย่อย	หมายเหตุรายงานทางการเงิน(Notes)ช่วง 2-3 ปีก่อน
83	ความถี่ในการจัดตั้งบริษัทย่อยใหม่	หมายเหตุรายงานทางการเงิน(Notes)1 ปีก่อน

จากคุณสมบัติ (feature) ทั้งหมด 16 features x 2 ช่วงเวลา = 74 - 105



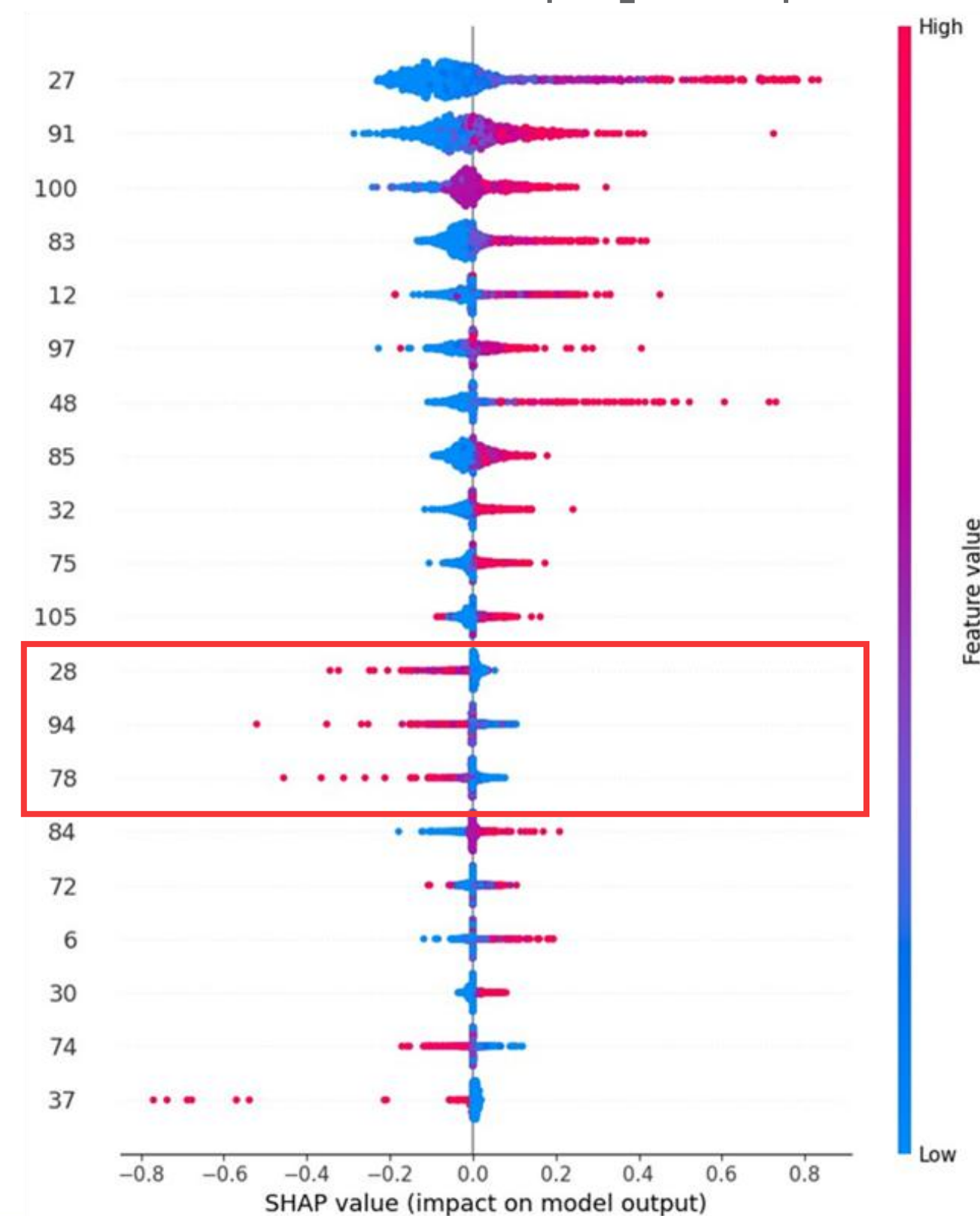


ผลการศึกษาเชิงประจักษ์



คุณสมบัติ (feature) ที่มีผลกระทบเชิงลบต่อเหตุการณ์ที่มีปัญหาด้านหุ่นกึ่งของหุ่น

FeatureID	Key	Data
94	ความถี่ในการเข้าซื้อกิจการ	หมายเหตุประกอบงบการเงิน: 2-3 ปีก่อนเกิดปัญหา
78	ความถี่ในการเข้าซื้อกิจการ	หมายเหตุประกอบงบการเงิน1 ปีก่อนเกิดปัญหา
28	การลงทุน	ข่าวจากSETช่วงเวลา: 91-365 วันก่อนเกิดปัญหา





สรุปผลการวิจัยและข้อเสนอแนะ



การประยุกต์ใช้ของนักลงทุน

สามารถสังเกตตัวบ่งชี้ที่ส่งผลต่อการตัดสินใจชำระหุ้นกู้ เพื่อประเมินความเสี่ยงด้วยตนเอง

ตัวบ่งชี้จากผลการวิจัยที่ส่งผลต่อการตัดสินใจชำระหุ้นกู้

ข่าว	หมายเหตุประกอบงบการเงิน	รายงานผู้สอบบัญชี
- การเพิ่มทุน	- การขายทรัพย์สิน	- คดีความทางกฎหมาย
- การเปลี่ยน CFO	- การเปลี่ยนนโยบายบัญชี	
- การแต่งตั้ง CEO คนใหม่	- การจัดตั้งบริษัทย่อยใหม่	



สรุปผลการวิจัยและข้อเสนอแนะ



ประโยชน์ในการต่อยอดในเชิงกำกับดูแล



ก.ล.ต. สามารถนำโมเดลไปใช้ในการ **ตรวจจับความเสี่ยงล่วงหน้า** โดยโมเดลสามารถพัฒนาให้เป็น **ระบบอัตโนมัติ**



สามารถเตือนเป็นระดับจากมากไปน้อยเป็นขั้น ๆ ได้โดยการพัฒนาจาก Binary เป็น Regression Model



สามารถนำข้อมูลข่าวรายวันของ SET เข้าไปในโมเดล หรือรวมไปถึงการพัฒนาไปครอบคลุมถึงข่าวอื่น ๆ



Thank you



www.set.or.th