**ประเภท 2 SW\_Dev**

**CE63-18**

**ห้องสอบ ECC-808**

**หุ่นยนต์อัตโนมัติสำหรับการตรวจจับความผิดปกติในการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์**

**Automatic Disorder Trading Detection for Stock Market**

**ภัทรวุฒิ วงศ์ประกรณ์กุล**

**สรวิศ อรรถรุ่งโรจน์**

**สาริศ พึ่งอาศัย**

**รายงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตร์บัณฑิต**

**สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**ปีการศึกษา 2563**

นายภัทรวุฒิ วงศ์ปกรณ์กุล 61015060

นายสรวิศ อรรถรุ่งโรจน์ 61015082

นายสาริศ พึ่งอาศัย 61015085

ผศ.บัณฑิต พัสยา อาจารย์ที่ปรึกษา

ปีการศึกษา 2563

**บทคัดย่อ**

ในปัจจุบัน การลงทุนในตลาดหลักทรัพย์ได้รับความสนใจอย่างมาก มีผู้ที่ริเริ่มเข้ามาลงทุนในตลาดหลักทรัพย์นี้มากมายหลากหลายอาชีพเพราะเป็นหนทางหารายได้เสริมอีกทางหนึ่งที่หากลงทุนได้ถูกต้องจะมีกำไรที่ดี เส้นทางตลาดหลักทรัพย์นี้นั้นมีความเสี่ยงและค่อนข้างผันผวน  คณะผู้วิจัยจึงคิดค้นโครงงาน “หุ่นยนต์อัตโนมัติสำหรับการตรวจจับความผิดปกติในการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์” โดยมีจุดประสงค์หลักเพื่อเป็นหนึ่งในเครื่องมือที่ช่วยในการตัดสินใจการซื้อขายหลักทรัพย์ให้แก่ผู้ที่สนใจได้ โมเดลการเรียนรู้นี้จะคำนวณหุ้น SET50 โดยจะใช้ข้อมูลในการฝึกฝนเป็นตัวชี้วัดทางเทคนิคที่คำนวณมาจากการเปลี่ยนแปลงของราคาหุ้นในแต่ละช่วงเวลาของวัน  นอกจากนี้ยังพัฒนาเว็บแอพพลิเคชันสำหรับแสดงการเปลี่ยนแปลงในด้านต่างๆของ SET50 พร้อมกับทำนายแนวโน้มของหุ้นในอนาคตว่าจะเป็นเช่นไร

I

**กิตติกรรมประกาศ**

ที่ปริญญานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีนั้น ต้องขอขอบพระคุณเป็นอย่างสูง ด้วยความช่วยเหลือของ ผศ.บัณฑิต พัสยา อาจารย์ที่ปรึกษาโครงงาน ผู้ชี้แนะแนวทางและออกความคิดเห็นต่างๆเกี่ยวกับแนวทางการดำเนินโครงงาน ซึ่งเป็นประโยชน์อย่างยิ่งในการทำงานวิจัยในหลายๆด้าน อีกทั้งยังช่วยแก้ไขปัญหาและอุปสรรคต่างๆที่ต้องพบเจอระหว่างการทำวิจัยด้วย ขอขอบคุณอาจารย์ท่านอื่นๆทุกท่าน ที่ได้ให้ความรู้และความเห็นรวมถึงเทคนิคต่างๆในแนวทางอื่นๆ มีผลให้สามารถพัฒนาการวิจัยได้ดี ครอบคลุม และแม่นยำมากขึ้น รวมถึงความช่วยเหลือ การระดมความคิด และความคิดเห็นของเพื่อนๆในสาขาในการทำโครงงานวิจัยนี้

สุดท้ายนี้ขอขอบพระคุณ ครอบครัวและทุกท่าน ที่ได้มอบโอกาสในการศึกษา และคอยช่วยเหลือข้าพเจ้าคณะผู้วิจัยตลอดจนสามารถสำเร็จการศึกษา

ภัทรวุฒิ วงศ์ปกรณ์กุล

สรวิศ อรรถรุ่งโรจน์

สาริศ พึ่งอาศัย

II

**สารบัญ**

**หน้า**

บทคัดย่อ I

กิตติกรรมประกาศ II

สารบัญ III

สารบัญตาราง V

สารบัญรูป VI

บทที่ 1 บทนำ 1

1.1 ความเป็นมาของปัญหา 1

1.2 วัตถุประสงค์ของโครงงาน 1

1.3 ขอบเขตของโครงงาน 2

1.4 วิธีการดำเนินงาน 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 เอกสารและทฤษฎีที่เกี่ยวข้อง 4

2.1 การเรียนรู้ของเครื่อง**(Machine Learning) 4**

**2.2** กลุ่มหลักทรัพย์ในตลาดหลักทรัพย์ของประเทศไทย 10

2.3 ตัวชี้วัดทางเทคนิค(Technical Indicator) 11

2.4 ไลบรารีสำหรับกระบวนการการเรียนรู้ของเครื่อง 13

2.5 งานวิจัยที่เกี่ยวข้อง 15

บทที่ 3 การออกแบบและพัฒนา 19

3.1 การออกแบบและพัฒนาโมเดลการจำแนกกลุ่มแนวโน้มของหลักทรัพย์ 19

III

**สารบัญ(ต่อ)**

**หน้า**

บทที่ 4 การทดลองและผลการทดลอง 24

4.1 ผลการเปรียบเทียบโมเดล Decision tree , Random forest และ XGBoost 24

4.2 ผลการทดลอง 24

บทที่ 5 บทสรุปและแนวทางการพัฒนา 28

5.1 บทสรุป 28

5.2 ปัญหาและอุปสรรค 28

5.3 แนวทางในการพัฒนาต่อ 28

IV

**สารบัญตาราง**

**ตาราง หน้า**

1.1 ระยะเวลาการดำเนินงาน 3

4.1 แสดงการเปรียบเทียบค่า Accuracy ของโมเดล Decision tree , Random forest และ XGBoost 24

4.2 แสดงการเปรียบเทียบค่า Precision ของโมเดล Decision tree , Random forest และ XGBoost 24

4.3 แสดงการเปรียบเทียบค่า Recall ของโมเดล Decision tree , Random forest และ XGBoost 25

4.4 แสดงการเปรียบเทียบค่า F1-Score ของโมเดล Decision tree , Random forest และ XGBoost 25

V

**สารบัญรูป**

**รูป หน้า**

2.1 แผนภาพการทำงานของ การเรียนรู้แบบมีผู้สอน 6

2.2 แผนภาพแสดงโครงข่ายประสาทเทียม 6

2.3 ตัวอย่างแผนผัง Decision Tree 9

2.4 ตัวอย่างแผนผัง Random forest 9

2.5 ตัวอย่างแผนผัง XGBoost 10

2.6 ตลาดหลักทรัพย์แห่งประเทศไทย 10

2.7 Scikit-learn 14

2.8 Pandas 14

2.9 ผลจากการทำนายด้วยวิธีโรฟัสซีระยะเวลา 1 เดือนโดยใช้เวลาการเรียนรู้ 100 วัน 16

2.10 ผลจากการทำนายด้วยวิธีโรฟัสซีระยะเวลา 1 เดือนโดยใช้เวลาการเรียนรู้ 500 วัน 16

2.11 ผลลัพธ์จากการทำนายโดยใช้ Z-Score Model 18

3.1 ตารางแสดงข้อมูลหลังจากจัดชุดข้อมูล 19

3.2 ค่าเริ่มต้นของโมเดล Decision tree 21

3.3 ค่าเริ่มต้นของโมเดล Random forest 22

3.4 ค่าเริ่มต้นของโมเดล XGBoost 22

VI

**สารบัญรูป(ต่อ)**

**รูป หน้า**

3.5 ผลการทดสอบของโมเดล Decision tree 23

3.6 ผลการทดสอบของโมเดล Random forest 23

3.7 ผลการทดสอบของโมเดล XGBoost 23

4.1 ตัวอย่างผลลัพธ์จากการทำนายในช่วง มีนาคม 2562 26

4.2 ตัวอย่างผลลัพธ์ของการทำนาย ในช่วงที่หลักทรัพย์มีความเปลี่ยนแปลงสูง 26

VII

**บทที่ 1**

**บทนำ**

**1.1 ความเป็นมาของปัญหา**

ตลาดหลักทรัพย์แห่งประเทศไทย (Stock Exchange of Thailand)อธิบายการลงทุนในหลักทรัพย์ว่า ผู้ถือหลักทรัพย์จะเป็นส่วนหนึ่งของกิจการได้รับผลตอบแทนเป็นเงินปันผลตามจำนวนหลักทรัพย์ที่ถืออยู่เมื่อกิจการมีกำไร ซึ่งได้รับเงินปันผลที่มากกว่าการลงทุนจากการฝากเงินในธนาคาร และถ้ากิจการมีผลประกอบการที่ดี ราคาหลักทรัพย์เติบโต จะสามารถทำกำไรได้จากส่วนต่างของราคาหลักทรัพย์ได้ด้วย แต่ราคาหลักทรัพย์ของแต่ละกิจการนั้นมีการเปลี่ยนแปลงที่ไม่แน่นอนตามสภาวะของเศรษฐกิจหรือตามปัจจัยอื่น ณ เวลานั้นๆ ถ้ากิจการมีปัญหาจนถึงขั้นล้มละลาย  ผู้ถือหลักทรัพย์จะ ได้รับเงินลงทุนคืนหลังจากกิจการจ่ายภาระผูกพันธ์แก่เจ้าหนี้และผู้ถือหุ้นบุริมสิทธิเรียบร้อยแล้ว ซึ่งถือเป็นความเสี่ยงของการลงทุนในหลักทรัพย์นอกจากนี้ หลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยนั้นยังมีจำนวนที่มากจึงเป็นเรื่องยากที่เราจะบริหารบัญชีซื้อขายหลักทรัพย์ คณะผู้วิจัยยังคิดหาวิธีการที่จะมาเป็นเครื่องมือช่วยสำหรับผู้ลงทุนในหลักทรัพย์ ในปัจจุบันการศึกษาและการนำไปใช้เกี่ยวกับการเรียนรู้ของเครื่อง (machine learning) กำลังเป็นที่นิยม ทำให้คณะผู้วิจัยสนใจที่จะศึกษาและหาแนวทางเพื่อช่วยเหลือสำหรับการลงทุนในตลาดหลักทรัพย์โดยใช้ Machine learning โดยในงานวิจัยนี้จะเป็นการศึกษาและคาดการณ์แนวโน้มและความผิดปกติของหุ้นตัวหนึ่งซึ่งก็คือ SET50 ให้สามารถลงทุนได้อย่างปลอดภัยที่สุด

**1.2 วัตถุประสงค์ของโครงงาน**

1. ศึกษาและเปรียบเทียบโมเดลการเรียนรู้ของเครื่องในการวิเคราะห์แนวโน้มของตลาดหลักทรัพย์ในรูปแบบต่างๆ
2. ศึกษาและเปรียบเทียบการเลือกใช้ฟีทเจอร์ (features) ที่แตกต่างกันสำหรับการใช้โมเดลใน Machine learning เพื่อหาโมเดลที่เสถียรที่สุด
3. สร้างและพัฒนาเว็บแอปพลิเคชันแนะนำในการตัดสินใจซื้อขายหลักทรัพย์แก่ผู้ลงทุนในระยะสั้นให้ได้กำไรและปลอดภัยมากที่สุด
4. ศึกษาและเปรียบเทียบผลจากการจำลองการซื้อขายหลักทรัพย์ เพื่อวิเคราะห์ในกรณีที่ตลาดหลักทรัพย์มีความผิดปกติ

**1.3 ขอบเขตของโครงงาน**

1. ใช้ชุดข้อมูลหลักทรัพย์ในตลาดหลักทรัพย์ของประเทศไทยตั้งแต่เดือน  มกราคม พ.ศ. 2562 จนถึงเดือนพฤศจิกายน พ.ศ.2563 ในการฝึกฝนโมเดล
2. ใช้ชุดข้อมูลหลักทรัพย์ในตลาดหลักทรัพย์ในช่วงเวลาปัจจุบันเพื่อใช้ในการตรวจสอบความผิดปกติในการซื้อขายหลักทรัพย์ของหลักทรัพย์ใน SET50
3. ระบบจะมีการตรวจจับความผิดปกติเพื่อแนะนำทิศทางของหลักทรัพย์ เฉพาะ SET50 ของตลาดหลักทรัพย์เท่านั้น

**1.4 วิธีการดำเนินงาน**

1. ศึกษาข้อมูลเรื่อง ดัชนี SET50
2. ศึกษาข้อมูลเรื่องปริมาณการซื้อขาย(Volume)
3. ศึกษาข้อมูลเรื่องกราฟแท่งเทียน(Candlestick Chart)
4. ศึกษาข้อมูลเรื่อง Moving average(MA)
5. ศึกษาข้อมูลเรื่อง Moving Average Convergence Divergence(MACD)
6. ศึกษาข้อมูลเรื่อง Relative Strength Index (RSI)
7. ศึกษาข้อมูลเรื่อง Machine learning
8. ศึกษาอัลกอรึทึมของการจำแนกกลุ่ม
9. ทดลองใส่ฟีทเจอร์ให้ Machine learning
10. ออกแบบการทำงานตรวจจับความผิดปกติในการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์
11. ทดลองและสรุปผลการทดลอง
12. จัดทำเอกสารปริญญานิพนธ์

**ตาราง 1.1 ระยะเวลาการดำเนินงาน**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| หัวข้อกิจกรรม | เดือน | | | |
| ส.ค. | ก.ย. | ต.ค. | พ.ย. |
| 1) ศึกษาข้อมูลเรื่อง ดัชนี SET50 |  |  |  |  |
| 2) ศึกษาข้อมูลเรื่องปริมาณการซื้อขาย(Volume) |  |  |  |
| 3) ศึกษาข้อมูลเรื่องกราฟแท่งเทียน(Candlestick Chart) |  |  |  |
| 4) ศึกษาข้อมูลเรื่อง Moving average(MA) |  |  |  |
| 5) ศึกษาข้อมูลเรื่อง Moving Average Convergence Divergence(MACD) |  |  |  |
| 6) ศึกษาข้อมูลเรื่อง Relative Strength Index (RSI) |  |  |  |
| 7) ศึกษาข้อมูลเรื่อง Machine learning |  |  |  |
| 8) ศึกษาอัลกอริทึมของการจำแนกกลุ่ม |  |  |
| 9) ทดลองใส่ฟีดเจอร์ให้ Machine learning |  |  |  |
| 10) ออกแบบการทำงานระบบตรวจจับความผิดปกติในการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์ |  |  |  |
| 15) ทดลองและสรุปผลการทดลอง |  |  |  |  |
| 16) จัดทำเอกสารปริญญานิพนธ์ |  |  |  |  |

**1.5 ประโยชน์ที่คาดว่าจะได้รับ**

1. มีความรู้ ความเข้าใจเกี่ยวกับกระบวนการทำงานในการเรียนรู้ของเครื่อง
2. มีความรู้ ความเข้าใจเกี่ยวกับพื้นฐานของการซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์
3. มีความรู้ ความเข้าใจในการออกแบบและพัฒนาการทำงานของแบบแอพพลิเคชันที่สามารถคาดการณ์แนวโน้มของตลาดหลักทรัพย์ได้

**บทที่ 2**

**เอกสารและทฤษฎีที่เกี่ยวข้อง**

**2.1 การเรียนรู้ของเครื่อง(Machine Learning)**

**2.1.1 การเรียนรู้ของเครื่อง**

การเรียนรู้ของเครื่อง หรือ Machine Learning เป็นส่วนการเรียนรู้ที่ถูกใช้งานเสมือนเป็นสมองของ ปัญญาประดิษฐ์(Artificial Intelligence) เราอาจพูดได้ว่า ปัญญาประดิษฐ์ใช้ การเรียนรู้ของเครื่องในในการยกระดับความรู้ของตนเอง โดยมักจะใช้เรียกโมเดลที่ปัญญาประดิษฐ์นั้นเป็นผู้เรียนรู้เอง ไม่ได้เกิดจากการเขียนโดยใช้มนุษย์ มนุษย์มีหน้าที่เพียงเขียนโปรแกรมให้ ปัญญาประดิษฐ์ (เครื่อง) เรียนรู้จากข้อมูลเท่านั้น โดยในส่วนที่เหลือนั้นเครื่องจะทำหน้าที่ประมวลผลจัดการดูแลเอง

การเรียนรู้ของเครื่อง จะเรียนรู้จากสิ่งที่ส่งเข้าไปกระตุ้นแล้วจดจำเอาไว้เป็นมันสมอง ส่งผลลัพธ์ออกมาเป็นตัวเลข หรือ โค๊ด(Code) ที่ส่งต่อไปแสดงผลหรือให้ตัว ปัญญาประดิษฐ์ นำไปแสดงการกระทำเอง และในส่วนของ การเรียนรู้ของเครื่อง ตัวเครื่องนั้นก็สามารถเอาไปใช้งานได้หลายรูปแบบ ซึ่งต้องอาศัยกลไกที่เป็นโปรแกรมหรือเรียกว่า อัลกอริทึม ที่มีหลากหลายรูปแบบ โดยมี นักวิทยาการข้อมูล(Data Scientist) เป็นผู้ออกแบบ อย่างไรก็ตามในการทำงานจริง นักวิทยาการข้อมูลจำเป็นต้องออกแบบตัวแปรต่างๆ และยังต้องหาอัลกอริทึมรูปแบบอื่นๆ มาเป็นคู่เปรียบเทียบ เพื่อมองหา อัลกอริทึม ที่เหมาะสมที่สุดในการใช้งานจริง

**2.1.2 กระบวนจัดการเพื่อนการเรียนรู้ของเครื่อง**

ในการที่จะนำข้อมูลของมาให้เครื่องเรียนรู้นั้น จำเป็นต้องจัดหาข้อมูลที่ถูกต้องแม่นยำ จึงต้องหาแหล่งข้อมูลที่เชื่อถือได้ จากนั้นจึงต้องมีการจัดการจัดเตรียมข้อมูลเหล่านั้นให้สามารถส่งเข้าไปเรียนรู้ในเครื่องได้

**2.1.2.1 ค้นหาและรวบรวมข้อมูล**

ค้นหาข้อมูลจากแหล่งข้อมูลที่น่าเชื่อถือ และนำข้อมูลเหล่านั้นมาเก็บไว้ให้ตรงหมวดหมู่ของตนเอง จัดเก็บไว้เป็นชุดข้อมูลตามประเภทของข้อมูลนั้นๆ

**2.1.2.2 จัดเตรียมชุดข้อมูล**

จัดเตรียมชุดข้อมูลให้เป็นไปในรูปแบบที่เราต้องการ ซึ่งการจัดเตรียมนี้อาจจะเป็นการ ลบ เพิ่ม สลับที่ ย้ายที่ ทำความสะอาด ชุดข้อมูลต่างๆให้อยู่ในรูปแบบใดๆ ซึ่งรูปแบบต่างๆนั้นจะขึ้นอยู่กับว่าผู้ศึกษาต้องการผลลัพธ์เช่นไรหลังการนำไปให้เครื่องเรียนรู้

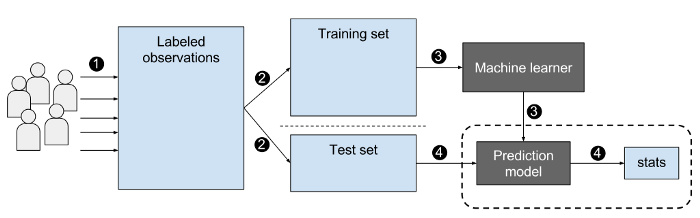
**2.1.2.3 แบ่งชุดข้อมูล**

แบ่งชุดข้อมูลที่เก็บรวบรวมมาได้เป็น 3 ส่วน คือ ชุดข้อมูลสำหรับฝึกฝน (training set) ชุดข้อมูลสำหรับทดสอบ (testing set) และชุดข้อมูลสำหรับตรวจสอบ (validation set)

**2.1.2.4 ศึกษาและใช้งานประเภทของการเรียนรู้ที่เหมาะสม**

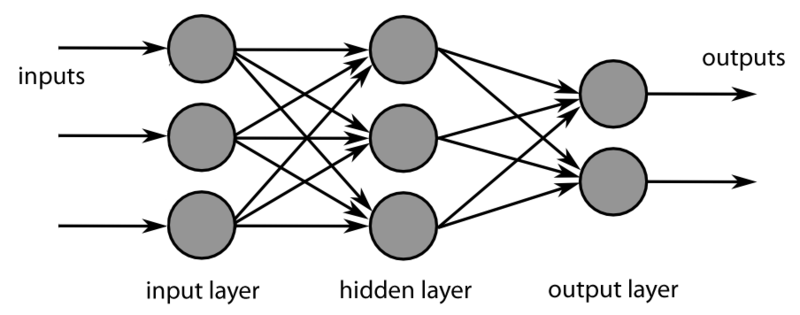
ประเภทของการเรียนรู้ที่ใช้แบ่งเป็น 2 ประเภท ได้แก่

* + - 1. **การเรียนรู้แบบมีผู้สอน (Supervised learning)** เป็นการเรียนรู้จากชุดข้อมูลที่มีคำถามคู่กับคำตอบ หรือก็คือมีคำตอบที่ชัดเจนมาแล้วให้เครื่องได้เรียนรู้ให้เกิดผลลัพธ์เช่นนั้น นอกจากนี้ยังมีการเรียนรู้ความคล้าย (similarity and metric learning) โดยเครื่องจะมีตัวอย่างของคู่ที่ถูกมองว่าคล้ายมากและคู่ที่ถูกมองว่าคล้ายน้อย เครื่องจะต้องหาฟังก์ชันความคล้ายออกมาที่สามารถทำนายได้ว่าวัตถุใหม่นั้นมีความคล้ายมากน้อยเพียงใด



**รูป 2.1 แผนภาพการทำงานของ การเรียนรู้แบบมีผู้สอน**

* + - 1. **โครงข่ายประสาทเทียม (Artificial neural networks)** เป็นอัลกอริทึมที่ได้แรงบันดาลใจมาจากโครงสร้างและการทำงานของเซลล์ประสาทในสมอง การคำนวณของโครงข่ายประสาทเทียมถูกสร้างเป็นโครงสร้างของการเชื่อมต่อของประสาทเทียมแต่ละตัว ประมวลผลข้อมูลโดยหลักการการเชื่อมต่อ



**รูป 2.2 แผนภาพแสดงโครงข่ายประสาทเทียม**

**2.1.2.5 ฝึกฝนโมเดล**

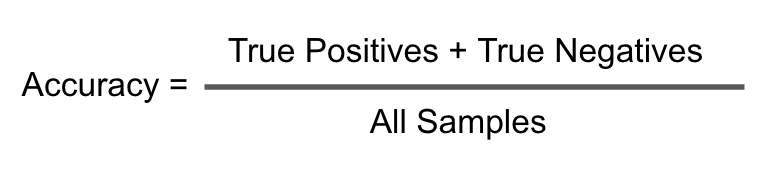
นำชุดข้อมูลสำหรับฝึกฝนที่เตรียมไว้ซึ่งจัดเรียงเรียบร้อยแล้วใส่เข้าไปในโมเดลที่เลือกเพื่อให้เกิดการเรียนรู้

**2.1.2.6 ทดสอบโมเดล**

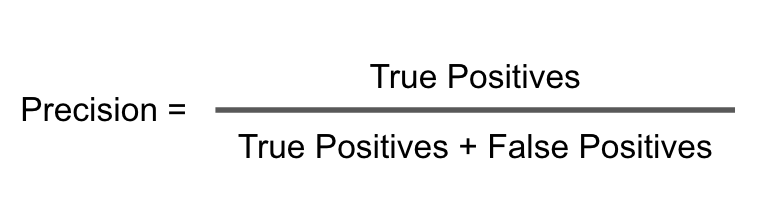
ใช้ชุดข้อมูลสำหรับตรวจสอบในการตรวจสอบหาข้อผิดพลาดในโมเดลของเรา ผลลัพธ์ถูกต้องตามที่ควรหรือไม่ มีความแม่นยำมากน้อยเพียงใด หากใส่ชุดข้อมูลใหม่ที่ไม่เคยพบจะทำนายได้หรือไม่ ซึ่งวิธีทดสอบที่นิยมใช้ในการตรวจสอบข้อมูลที่โมเดลยังไม่เคยพบเจอมาก่อนคือวิธี hold-out เป็นการแบ่งชุดข้อมูลสำหรับฝึกฝน 80 เปอร์เซ็นต์ และชุดข้อมูลสำหรับทดสอบ 20 เปอร์เซ็นต์ ซึ่งขึ้นอยู่กับว่าผู้จัดทำจะแบ่งกี่เปอร์เซ็นต์

* + - 1. **การประเมินผลของโมเดล**

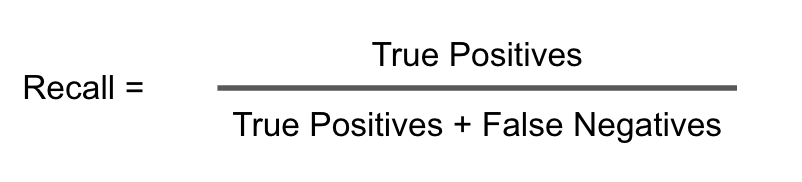
1. Accuracy คือสัดส่วนเปอร์เซ็นต์ความถูกต้อง คือจำนวนที่ทำนายถูก/จำนวนทั้งหมด ดังสมการต่อไปนี้

 (2.1)

1. Precision คือความแม่นยำผลทำนาย (สนใจผลทำนาย หรือ Prediction) คำนวณโดยสัดส่วนที่ทำนายว่าเป็น Positive กับ Negative กี่เปอร์เซ็นต์ ตามสมการต่อไปนี้

 (2.2)

1. Recall คือตัวประเมินผลของโมเดลที่มีการวัดความแม่นยำโดยเทียบที่ผลลัพธ์และความเป็นจริง โดยสัดส่วนต่างๆนั้นแบ่งได้ตามสมการต่อไปนี้

 (2.3)

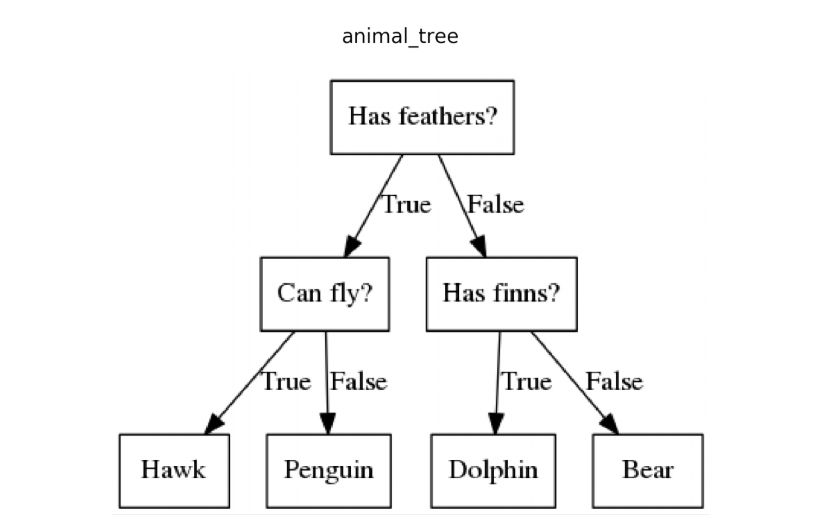
1. F1 Score เป็นตัวประเมินโมเดลที่เป็นค่าแสดงประสิทธิภาพ โดยการนำเอา Precision กับ Recall มาคำนวณหาค่าเฉลี่ย ซึ่งหากว่าได้ค่าสูงจะถือว่ามีประสิทธิภาพดี โดยมีสมการดังนี้

(2.4)

**2.1.3 อัลกอริทึมการจำแนกกลุ่ม**

**2.1.3.1 Decision Tree**

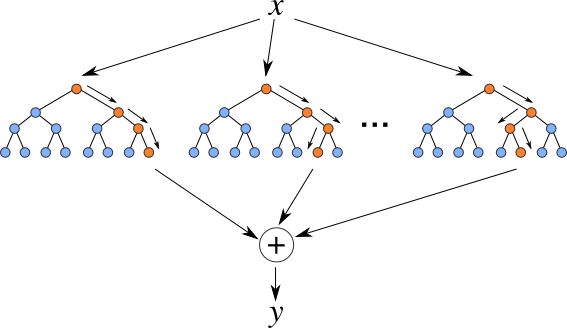
Decision tree หรือต้นไม้แห่งการตัดสินใจ เป็นอัลกอริทึมพื้นฐานตัวหนึ่งของ Machine Learning จัดอยู่ในกลุ่ม การเรียนรู้แบบมีผู้สอน เป็นการเรียนรู้จากคุณลักษณะของข้อมูล (Attributes) แล้วสร้างการตัดสินใจออกมารูปร่างคล้ายกับลำต้นและกิ่งไม้ โดยปกติใช้กับการจำแนกหรือแยกแยะ (Classification) ว่าข้อมูลจัดอยู่ในกลุ่มใด โดยผลลัพธ์จะมีสองกลุ่มหรือมากกว่าก็ได้



**รูป 2.3 ตัวอย่างแผนผัง Decision Tree**

**2.1.3.2 Random Forest**

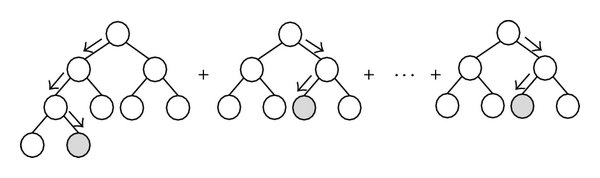
Random forest เป็นอัลกอริทึมชนิดหนึ่งที่ประกอบด้วย Decision tree หลายต้น ซึ่งเป็นตัวที่ถูกพัฒนาขึ้นมาจาก Decision tree โดยจะนำมาช่วยในการตัดสินใจที่ละเอียดยิ่งขึ้น เนื่องจากเกิดการจำลองหลายครั้ง และจะตัดสินใจเลือกผลลัพธ์ที่ดีที่สุดออกมาใช้ในการตัดสินใจ



**รูป 2.4 ตัวอย่างแผนผัง Random forest**

**2.1.3.3 Extreme Gradient Boosting (XGBoost)**

เป็นอัลกอริทึมที่นำเอา Decision Tree มาฝึกฝนต่อเนื่องกันหลายๆต้น โดยที่แต่ละต้น จะเรียนรู้จากความผิดพลาดของต้นก่อนหน้า ทำให้ความแม่นยำของในการทำนายจะ แม่นยำสูงขึ้น ซึ่งจะหยุดการเรียนรู้เมื่อไม่มีความผิดพลาดส่งต่อมาจากต้นก่อนหน้าแล้ว

****

**รูป 2.5 ตัวอย่างแผนผัง XGBoost**

**2.2 กลุ่มหลักทรัพย์ในตลาดหลักทรัพย์ของประเทศไทย**

**2.2.1 ตลาดหลักทรัพย์แห่งประเทศไทย (The Stock Exchange of Thailand: SET)**

****

**รูป 2.6 ตลาดหลักทรัพย์แห่งประเทศไทย**

ตลาดหลักทรัพย์แห่งประเทศไทย ([อังกฤษ](https://th.wikipedia.org/wiki/%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2%E0%B8%AD%E0%B8%B1%E0%B8%87%E0%B8%81%E0%B8%A4%E0%B8%A9): Stock Exchange of Thailand, SET) เป็น[ตลาดหลักทรัพย์](https://th.wikipedia.org/wiki/%E0%B8%95%E0%B8%A5%E0%B8%B2%E0%B8%94%E0%B8%AB%E0%B8%A5%E0%B8%B1%E0%B8%81%E0%B8%97%E0%B8%A3%E0%B8%B1%E0%B8%9E%E0%B8%A2%E0%B9%8C)ของ[ประเทศไทย](https://th.wikipedia.org/wiki/%E0%B8%9B%E0%B8%A3%E0%B8%B0%E0%B9%80%E0%B8%97%E0%B8%A8%E0%B9%84%E0%B8%97%E0%B8%A2) จัดตั้งขึ้นโดย[พระราชบัญญัติ](https://th.wikipedia.org/wiki/%E0%B8%9E%E0%B8%A3%E0%B8%B0%E0%B8%A3%E0%B8%B2%E0%B8%8A%E0%B8%9A%E0%B8%B1%E0%B8%8D%E0%B8%8D%E0%B8%B1%E0%B8%95%E0%B8%B4)ตลาดหลักทรัพย์แห่งประเทศไทย [พ.ศ. 2517](https://th.wikipedia.org/wiki/%E0%B8%9E.%E0%B8%A8._2517) ทำหน้าที่เป็นตลาดรองเพื่อแลกเปลี่ยนซื้อขาย[ตราสารทุน](https://th.wikipedia.org/wiki/%E0%B8%95%E0%B8%A3%E0%B8%B2%E0%B8%AA%E0%B8%B2%E0%B8%A3%E0%B8%97%E0%B8%B8%E0%B8%99) ของบริษัทต่าง ๆ ที่ขึ้นทะเบียนไว้ และ เพื่อให้สามารถระดมเงินทุนเพิ่มเติมจากสาธารณะได้โดยสะดวก และยังได้มีการผูกรวมหุ้นราคาของหุ้นสามัญ 50 ตัวสำคัญเอาไว้เรียกว่า ดัชนีราคาหลักทรัพย์ SET50 เพื่อให้ง่ายต่อการซื้อขาย และมีความคล่องตัว ปัจจุบันการดำเนินงานของตลาดหลักทรัพย์แห่งประเทศไทย อยู่ภายใต้พระราชบัญญัติหลักทรัพย์และตลาดหลักทรัพย์ [พ.ศ. 2535](https://th.wikipedia.org/wiki/%E0%B8%9E.%E0%B8%A8._2535)

**2.3 ตัวชี้วัดทางเทคนิค(Technical Indicator)**

**2.3.1 ปริมาณการซื้อขาย(Volume)**

ปริมาณการซื้อขายของหลักทรัพย์นั้นช่วยแสดงให้เราเห็นถึงความต้องการซื้อขายของผู้คนในหลักทรัพย์ตัวนั้นๆ หากราคาหลักทรัพย์ปรับสูงขึ้นโดยที่ปริมาณการซื้อขายเพิ่มขึ้นตามมาด้วยจะเป็นการผลักดันให้ราคาหลักทรัพย์สูงขึ้น หรือหากว่าราคาหลักทรัพย์นั้นมีราคาที่เพิ่มสูงขึ้น แต่ทว่าราคาของหลักทรัพย์กลับปรับตัวลงในเวลาต่อมา อีกทั้งปริมาณการซื้อขายของหลักทรัพย์ก็ปรับลดตัวลงด้วย สิ่งนี้แสดงให้เห็นว่าการปรับตัวลงของหลักทรัพย์นั้นจะเกิดขึ้นเป็นการชั่วคราว หลังจากนั้นราคาหลักทรัพย์จะมีการปรับตัวสูงขึ้นอีกครั้ง

การปรับตัวขึ้นลงของราคาหลักทรัพย์และปริมาณการซื้อหลักทรัพย์ที่ไม่สัมพันธ์กัน อาจจะกำลังสื่อให้เห็นถึงความผิดปกติของหลักทรัพย์ได้ ซึ่งถ้าหากเราตัดสินใจซื้อ-ขายในหลักทรัพย์นั้นอาจจะทำให้เราขาดทุนได้

**2.3.2 มูลค่าตามราคาตลาด (Market Capitalization)**

มูลค่าของราคาตลาดหลักทรัพย์สามารถบอกเราถึงความน่าลงทุน มูลค่าของตลาด หรือขนาดของตลาดหลักทรัพย์ได้ โดยคำนวณได้จาก

มูลค่าของราคาตลาด = ราคาปิดของหลักทรัพย์ x จำนวนหลักทรัพย์จดทะเบียน (2.5)

**2.3.3 เส้นค่าเฉลี่ยเคลื่อนที่ (Moving Average: MA)**

Moving Average หรือเส้นค่าเฉลี่ยเคลื่อนที่ เป็นเครื่องมือเทคนิคที่ใช้วิเคราะห์แนวโน้ม หรือทิศทางของตลาดซึ่งจะช่วยประกอบการตัดสินใจในการซื้อขายหลักทรัพย์นั้นๆ ซึ่งจะคำนวณจากระยะเวลาที่เราสามารถกำหนดได้ ซึ่งแบ่งออกได้เป็นหลายประเภทเพื่อใช้ประกอบการตัดสินใจได้ดียิ่งขึ้นในแต่ละรูปแบบของตลาดหลักทรัพย์

**2.3.3.1 Simple Moving Average (SMA)**

เป็นการหาค่าเฉลี่ยตามจำนวนวันที่เรากำหนด เช่น หากเราต้องการหา SMA5 เราก็นำราคาปิดของ 5 วันนั้นที่เราต้องการมารวมกันแล้วหารด้วยจำนวนวัน ซึ่งในที่นี้ก็คือ 5

(2.6)

**2.3.3.2 Exponential Moving Average (EMA)**

เนื่องจาก EMA เป็นเส้นค่าเฉลี่ยที่คำนวณโดยให้ความสำคัญกับราคาช่วงแรกๆน้อยและให้ความสำคัญกับราคาช่วงหลังๆมาก จึงเป็นสูตรที่แกว่งตัวตามราคาปัจจุบัน ทำให้เหมาะในการดูสัญญาณซื้อขาย ซึ่งมีสูตรการคำนวณดังนี้ โดยกำหนดให้ n คือระยะเวลาที่กำหนด(วัน)

(2.7)

**2.3.4 Moving Average Convergence Divergence (MACD)**

Moving Average Convergence Divergence เป็นการหาค่าเฉลี่ยโดยคำนวณจากระยะห่างของเส้นค่าเฉลี่ยสองเส้นเพื่อหาทิศทางแนวโน้มของราคาหลักทรัพย์ โดยเส้นที่นิยมใช้คือ เส้นค่าเฉลี่ยของ EMA12 และ เส้นค่าเฉลี่ยของ EMA 25 หรือ 26 ซึ่งสูตรคำนวณคือ

(2.8)

**2.3.5 Relative Strength Index (RSI)**

RSI เป็นตัววัดความเร็วในการเคลื่อนไหวและทิศทางของราคาหลักทรัพย์ ซึ่งจะแสดงค่าเฉลี่ยของช่วงที่กำหนด เพื่อนำไปใช้ในการบอกถึงแรงส่งของหลักทรัพย์ได้ ซึ่งมีสูตรการคำนวณดังนี้

โดยกำหนดให้ : average gain คือ ราคาปิดบวกในแต่ละวันที่กำหนดรวมกัน / จำนวนวันที่กำหนด

Average loss คือ ราคาปิดลบในแต่ละวันที่กำหนดรวมกัน / จำนวนวันที่กำหนด

(2.9)

**2.4 ไลบรารีสำหรับกระบวนการการเรียนรู้ของเครื่อง**

**2.4.1 Scikit-learn**

Scikit-learn คือ Open Source Library สำหรับทำเหมืองข้อมูล และการวิเคราะห์ข้อมูลด้วยภาษาไพธอน มีอัลกอริทึมในการสร้าง โมเดลการเรียนรู้ของเครื่องเช่น โมเดลจำแนกข้อมูล (Classification), โมเดล แบบถดถอย (Regression) และ โมเดลจัดกลุ่มข้อมูล (Clustering) ตลอดจนการเตรียมข้อมูลต่าง ๆ เช่น การทำ Normalization หรือ Standardization ตลอดจนการจัดการกับข้อมูลที่ผิดปกติเช่นข้อมูลที่ไม่ได้กำหนดค่า (Missing Value) เป็นต้น



**รูป 2.7 Scikit-learn**

**2.4.2 Pandas**

Pandas นั้นคือไลบรารีหนึ่งในภาษาไพทอนที่ทำให้เราสามารถจัดการข้อมูลต่างๆได้ง่ายและเป็นระเบียบยิ่งขึ้น หนึ่งในการทำงานที่สำคัญคือ Pandas สามารถนำข้อมูลไฟล์ CSV เข้ามาแล้วแสดงข้อมูลให้ออกมาในรูปแบบ ตาราง เราเรียกสิ่งนี้ว่า Data Frame

****

**รูป 2.8 Pandas**

**2.5 งานวิจัยที่เกี่ยวข้อง**

**2.5.1 การพยากรณ์ราคาหลักทรัพย์ด้วยวิธีการนิวโรฟัซซี**

งานวิจัยนี้นําข้อมูลมาพยากรณ์โดยใช้แบบจําลอง โครงข่ายประสาทเทียมแบบฟัซซีด้วยหลักการ Adaptive Neuro–Fuzzy Inference System (ANFIS) โดยทฤษฎีพื้นฐานของ ฟัซซีลอจิก (Fuzzy Logic) ซึ่งสามารถจัดการกับปัญหาที่มีลักษณะของข้อมูลคลุมเครือได้เป็นอย่างดี

โดยจะป้อนข้อมูลด้วยหลักการจำลองข้อมูลซึ่งเป็นรูปแบบที่ใช้กันแพร่หลายสําหรับงานที่พบว่ามีรูปแบบของข้อมูลที่อยู่ในลักษณะของ อนุกรมเวลา (Time Series) โดยมีการจัดเก็บข้อมูลในช่วงเวลาที่กําหนดไว้ในอดีต ซึ่งจะใช้พยากรณ์ถึงสิ่งที่ต้องการทราบในช่วงเวลาถัดไป โดยจะจัดข้อมูลนําเข้าแบบจําลองตาม สมการดังต่อไปนี้

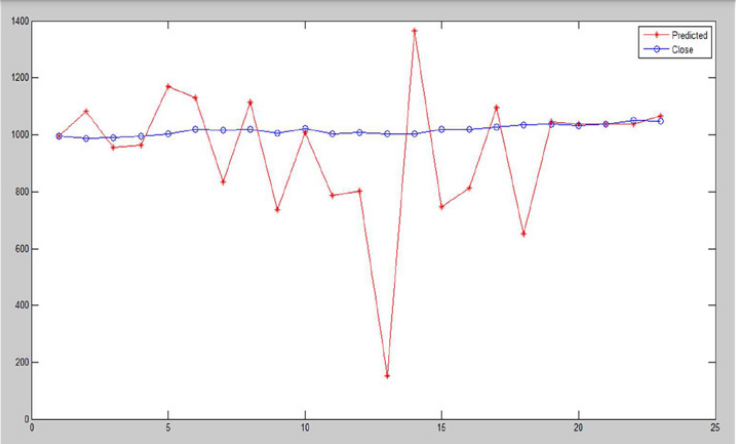
(2.10)

โดย D คือ จํานวนของชุดข้อมูลป้อนเข้า

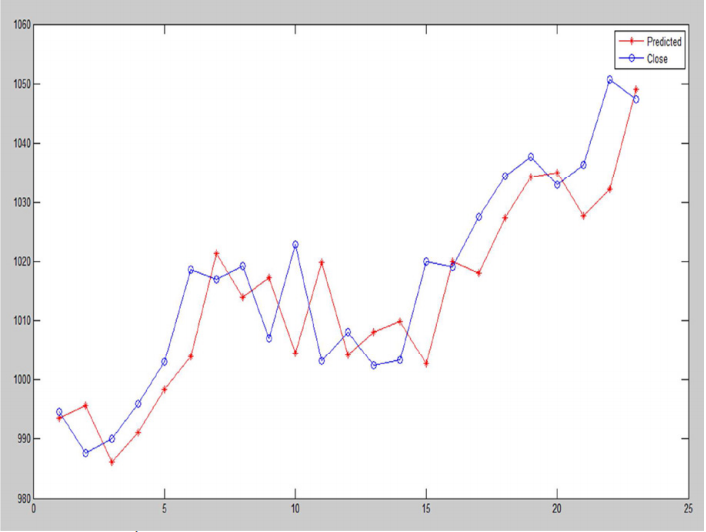
∆ คือ เวลาย้อนหลังในอดีตของแต่ละช่วง มีค่าเท่ากับ P

P คือ เวลาที่จะทําการพยากรณ์

t คือ แนวโน้มเวลา

****

**รูป 2.9 ผลจากการทำนายด้วยวิธีโรฟัสซีระยะเวลา 1 เดือนโดยใช้เวลาการเรียนรู้ 100 วัน**

****

**รูป 2.10 ผลจากการทำนายด้วยวิธีโรฟัสซีระยะเวลา 1 เดือนโดยใช้เวลาการเรียนรู้ 500 วัน**

**2.5.2 การวิเคราะห์ความเสี่ยงด้วยแบบจำลอง Altman ในกลุ่มบริษัทที่อยู่ในกลุ่มอุตสาหกรรมบริการ หมวดธุรกิจการท่องเที่ยวในตลาดหลักทรัพย์แห่งประเทศไทย**

Altman Z-Score ถูกพัฒนาขึ้นมาโดยมีวัตถุประสงค์เพื่อวัสถานการณ์ทางการเงินของบริษัทเพื่อคาดการณ์โอกาสล้มเหลวของกิจการล่วงหน้า โดยมีความแม่นยำกว่าร้อยละ 82 ของกิจการที่ล้มละลายใน 1 ปี ซึ่งมีสูตรคำนวณดังนี้

*(2.11)*

โดย Z คือ ค่าดัชนีวัดภาวะล้มละลายของธุรกิจ

คือ อัตราส่วนความคล่องตัว = เงินทุนหมุนเวียน/สินทรัพย์รวม

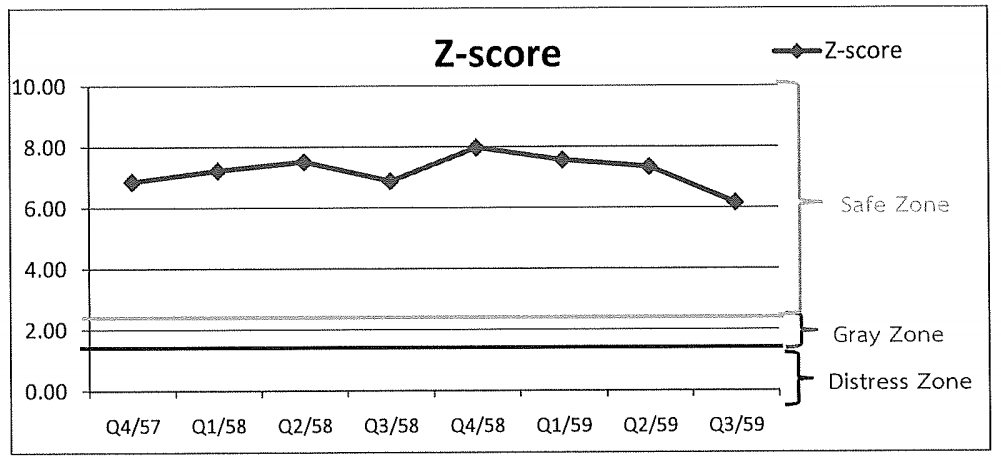
คือ เงินทุนสะสมจากแหล่งภายใน = กำไรสะสม/สินทรัพย์รวม

คือ ความสามารถหากำไร = กำไรก่อนดอกเบี้ยและภาษี/สินทรัพย์รวม

คือ สัดส่วนแห่งภาระผูกพันทางการเงิน = มูลค่าตลาดของหุ้นสามัญ/มูลค่าตาม บัญชีของหนี้สินรวม

คือ ความสามารถในการใชประโยขน์จากสินทรัพย์ของธุรกิจ = ขาย/สินทรัพย์รวม

การวิจัยนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Analysis) ซึ่งเครื่องมือที่ใช้ก็คือ Z-Score Model โดยสร้างกระดาษทำการในโปรแกรมไมโครซอฟท์เอ็กเซล(Microsoft Excel) มาคำนวณอัตราส่วนทางการเงินทั้ง 5 ส่วน และนำมาแทนค่าในสมการแบบจำลแง Z-Score



**รูป 2.11 ผลลัพธ์จากการทำนายโดยใช้ Z-Score Model**

**บทที่ 3**

**การออกแบบและพัฒนา**

**3.1 การออกแบบและพัฒนาโมเดลการจำแนกกลุ่มแนวโน้มของหลักทรัพย์**

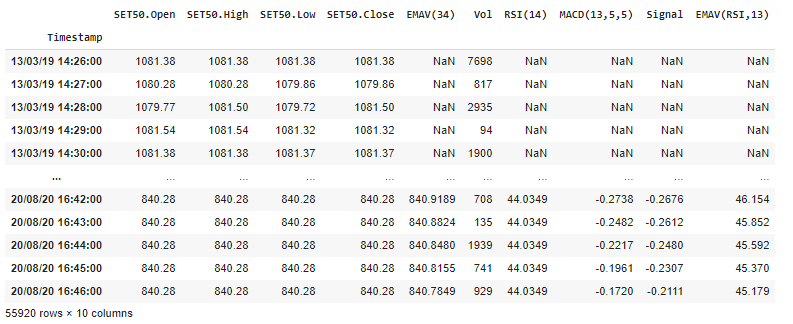
**3.1.1 เก็บรวบรวมข้อมูล**

ใช้ชุดข้อมูลหลักทรัพย์จากตลาดหลักทรัพย์แห่งประเทศไทยตั้งแต่เดือน มกราคม พ.ศ.2562 จนถึงเดือนพฤศจิกายน พ.ศ.2563 โดยเป็นข้อมูลที่แสดงเกี่ยวกับตัวชี้วัดในด้านต่างๆซึ่งนำมาใช้ในการหาทิศทางของหลักทรัพย์

**3.1.2 จัดเตรียมชุดข้อมูล**

**3.1.2.1 เลือกค่าตัวชี้วัดที่ต้องการจากชุดข้อมูลมาจัดเป็นชุดข้อมูลใหม่**

ในชุดข้อมูลที่คณะผู้วิจัยมีนั้นจะมีตัวชี้วัดอยู่หลายตัว ซึ่งคณะผู้วิจัยไม่ได้ใช้ตัวชี้วัดทั้งหมดนั้น จึงต้องแบ่งออกมาเฉพาะข้อมูลที่ต้องการแล้วจัดเป็นชุดข้อมูลใหม่ ซึ่งตัวชี้วัดที่ใช้นั้นได้แก่ ราคาปิด , EMA เฉลี่ยย้อนหลัง 34 นาที , MACD เฉลี่ยย้อนหลัง 13 นาที , RSI , ค่าเฉลี่ยย้อนหลังของ RSI 13 นาที และ Signal Line(MACD เฉลี่ยย้อนหลัง 9 นาที)



**รูป 3.1 ตารางแสดงข้อมูลหลังจากจัดชุดข้อมูล**

**3.1.2.2 ตรวจสอบชุดข้อมูลที่จัดเตรียมมาใหม่**

ชุดข้อมูลที่ได้จัดเรียงมาใหม่นั้นมีความบกพร่องในส่วนข้อมูลของตัวชี้วัด เราจะต้องตรวจสอบค่าตัวชี้วัดเหล่านั้นว่ามีค่าเป็น NaN , Null หรือไม่ หากพบเจอจะทำการลบข้อมูลในแถวนั้นทั้งหมด

**3.1.2.3 เพิ่มค่ากลุ่มแนวโน้มของหลักทรัพย์ให้กับชุดข้อมูล**

โดยจะทำนายการเปลี่ยนแปลงของหลักทรัพย์โดยพิจารณาเงื่อนไขต่อไปนี้

1. ถ้าตัวชี้วัด MACD สูงกว่า Signal จะมีโอกาสที่หลักทรัพย์จะมีราคาพุ่งสูงขึ้นหรือไม่เกิดการเปลี่ยนแปลง
2. หากตรงตามเงื่อนไขข้อที่ 1 จากนั้นหากว่าตัวชี้วัด RSI มากกว่า ค่าเฉลี่ยย้อนหลังของตัวชี้วัด RSI 13 นาที จะมีโอกาสขึ้นหรือไม่เกิดการเปลี่ยนแปลง
3. หากตรงตามเงื่อนไขข้อที่ 2 จากนั้นหากว่าตัวชี้วัด RSI ต่ำกว่า 70 จะแนะนำให้ซื้อหลักทรัพย์ในขณะนั้น
4. ถ้าตัวชี้วัด MACD ต่ำกว่า Signal จะมีโอกาสที่หลักทรัพย์จะมีราคาลดต่ำลงหรือไม่เกิดการเปลี่ยนแปลง
5. หากตรงตามเงื่อนไขข้อที่ 4 จากนั้นหากว่าตัวชี้วัด RSI ต่ำกว่ากว่า ค่าเฉลี่ยย้อนหลังของตัวชี้วัด RSI 13 นาที จะมีโอกาสที่หลักทรัพย์จะมีราคาลดต่ำลงหรือไม่เกิดการเปลี่ยนแปลง
6. หากตรงตามเงื่อนไขข้อที่ 5 จากนั้นหากว่าตัวชี้วัด RSI สูงกว่า 30 และ ราคาปิดน้อยกว่าราคาปิดเฉลี่ย จะแนะนำให้ขายหลักทรัพย์ในขณะนั้น
7. ถ้าไม่ตรงตามเงื่อนไขใดๆก่อนหน้าให้ถือหลักทรัพย์ไว้

**3.1.3 แบ่งชุดข้อมูล**

ใช้ชุดข้อมูลหลักทรัพย์ของตลาดหลักทรัพย์แห่งประเทศไทยตั้งแต่เดือนมีนาคม พ.ศ. 2562 จนถึงเดือนสิงหาคม พ.ศ. 2563 โดยแบ่งเป็นสองชุด ชุดแรกสำหรับฝึกฝนแบ่งเป็น 80% ชุดที่สองสำหรับทดสอบแบ่งเป็น 20%

**3.1.4 ศึกษาโมเดลการจำแนกกลุ่มหลักทรัพย์ที่เหมาะสม**

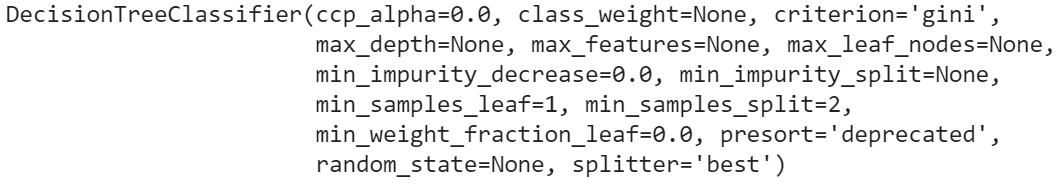
ศึกษาโมเดลในการจำแนกกลุ่มหลักทรัพย์ เพื่อหาโมเดลที่เหมาะสมตามแต่ละเงื่อนไขที่ใช้งาน ซึ่งโมเดลที่ทำการศึกษามีดังนี้

1. Dicision tree
2. Random forest
3. XGBoost

**3.1.5 ทดสอบโมเดลการจำแนกกลุ่มหลักทรัพย์**

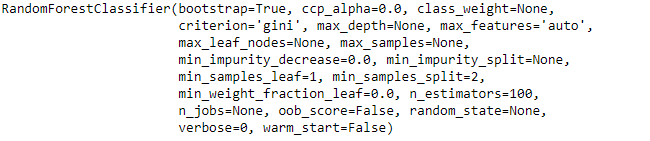
ฝึกฝนและทดสอบโมเดลจำแนกกลุ่มที่ได้เลือกศึกษามาฝึกฝนชุดข้อมูลสำหรับฝึกฝน ซึ่งค่าเริ่มต้นจากไลบรารีใน scikit-learn และ ผลทดสอบของแต่ละโมเดลมีดังนี้

1. **Dicision tree**



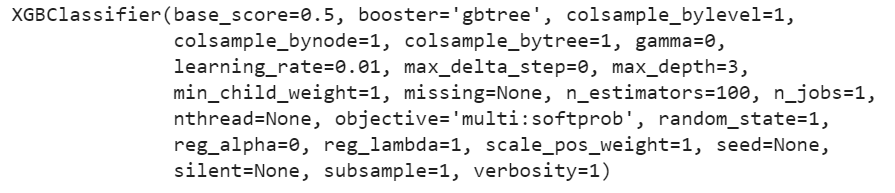
**รูป 3.2 ค่าเริ่มต้นของโมเดล Decision tree**

1. **Random forest**



**รูป 3.3 ค่าเริ่มต้นของโมเดล Random forest**

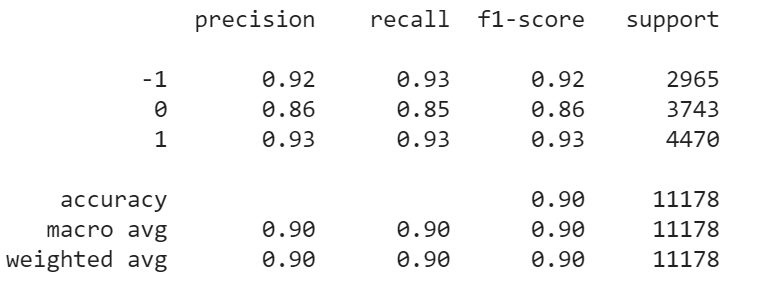
1. **XGBoost**



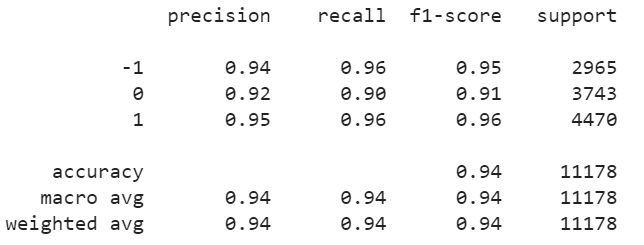
**รูป 3.4 ค่าเริ่มต้นของโมเดล XGBoost**

**3.1.6 ทดสอบโมเดลการจำแนกกลุ่มแนวโน้มของหลักทรัพย์**

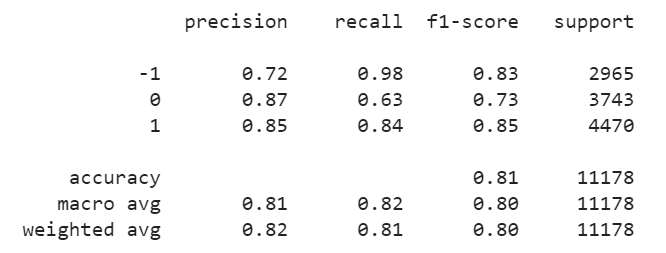
ฝึกฝนโมเดลการจำแนกกลุ่มแนวโน้มของหลักทรัพย์ด้วยชุดข้อมูลสำหรับการฝึกฝน แล้วจากนั้นก็นำไปทดสอบด้วยชุดข้อมูลสำหรับทดสอบ ซึ่งผลที่ได้ของแต่ละโมเดลเป็นดังต่อไปนี้



**รูป 3.5 ผลการทดสอบของโมเดล Decision tree**



**รูป 3.6 ผลการทดสอบของโมเดล Random forest**



**รูป 3.7 ผลการทดสอบของโมเดล XGBoost**

**บทที่ 4**

**การทดลองและผลการทดลอง**

**4.1 ผลการเปรียบเทียบโมเดล Decision tree , Random forest และ XGBoost**

ผลการเปรียบเทียบระหว่างโมเดลต่างๆโดยจะเปรียบเทียบกันผ่านตัวประเมินผลของโมเดล 4 ตัวที่ได้กล่าวไว้ข้างต้น คือ Accuracy , Precision , Recall และ F1-Score ซึ่งแต่ละโมเดลจะมีผลลัพธ์ที่นำมาเปรียบเทียบกัน 3 อย่างคือ การซื้อ การขาย และถือ หลักทรัพย์ไว้

**ตาราง 4.1 แสดงการเปรียบเทียบค่า Accuracy ของโมเดล Decision tree , Random forest และ XGBoost**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Decision tree | Random forest | XGBoost |
| Accuracy | 0.90 | 0.94 | 0.81 |

**ตาราง 4.2 แสดงการเปรียบเทียบค่า Precision ของโมเดล Decision tree , Random forest และ XGBoost**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Decision tree | Random forest | XGBoost |
| Precision | | |
| ซื้อ | 0.93 | 0.95 | 0.85 |
| ถือ | 0.86 | 0.92 | 0.87 |
| ขาย | 0.92 | 0.95 | 0.72 |

**ตาราง 4.3 แสดงการเปรียบเทียบค่า Recall ของโมเดล Decision tree , Random forest และ XGBoost**

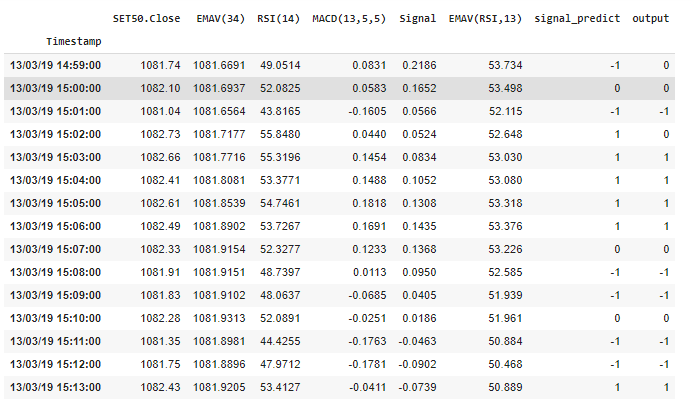
|  |  |  |  |
| --- | --- | --- | --- |
|  | Decision tree | Random forest | XGBoost |
| Recall | | |
| ซื้อ | 0.93 | 0.96 | 0.84 |
| ถือ | 0.85 | 0.90 | 0.63 |
| ขาย | 0.93 | 0.96 | 0.98 |

**ตาราง 4.4 แสดงการเปรียบเทียบค่า F1-Score ของโมเดล Decision tree , Random forest และ XGBoost**

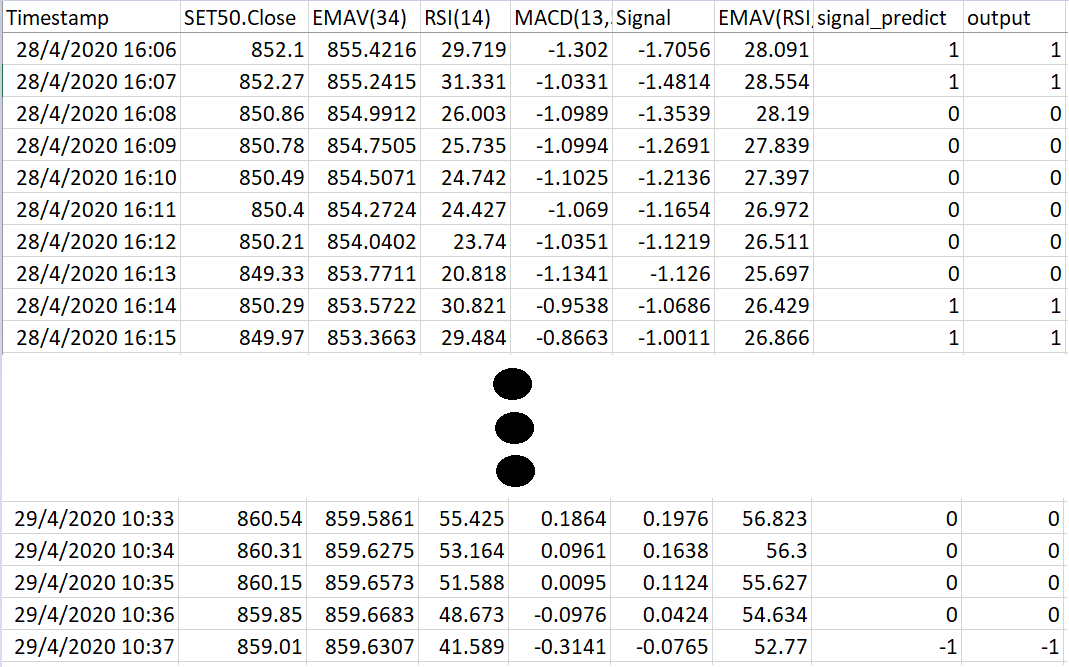
|  |  |  |  |
| --- | --- | --- | --- |
|  | Decision tree | Random forest | XGBoost |
| F1-Score | | |
| ซื้อ | 0.93 | 0.96 | 0.85 |
| ถือ | 0.86 | 0.91 | 0.73 |
| ขาย | 0.92 | 0.95 | 0.83 |

จากผลการเปรียบเทียบในแต่ละโมเดลดังที่ได้แสดงไปในตารางข้างต้นนั้น แสดงให้เห็นว่าโมเดล Random Forest มีค่าความถูกต้องแม่นยำสูงที่สุดในเกือบทุกๆด้าน เราจึงตัดสินใจเลือกใช้โมเดล Random forest ในการทำนายทิศทางการเปลี่ยนแปลงของหลักทรัพย์

**4.2 ผลการทดลอง**



รูป 4.1 ตัวอย่างผลลัพธ์จากการทำนายในช่วง มีนาคม 2562



รูป 4.2 ตัวอย่างผลลัพธ์ของการทำนาย ในช่วงที่หลักทรัพย์มีความเปลี่ยนแปลงสูง

จากรูปที่ 4.1 และ 4.2 แสดงผลลัพธ์จากการทำนายของสองช่วงเวลาที่ต่างกันคือ ช่วงเวลาที่หลักทรัพย์มีความเปลี่ยนแปลงไม่มาก หรือไม่ชัดเจนจนเราสังเกตเห็นได้โดยง่าย และอีกช่วงเวลาคือช่วงที่มีการเปลี่ยนแปลงของมูลค่าหลักทรัพย์อย่างชัดเจนจากการเว้นช่วงเวลาการซื้อขายยืดออกไปยาวกว่านั้นจะสามารถคาดการณ์ถึงทิศทางได้แม่นยำมากขึ้น

**บทที่ 5**

**บทสรุปและแนวทางการพัฒนา**

**5.1 บทสรุป**

การฝึกฝนโมเดลทำออกมาได้ตรงตามที่คาดหมาย ไม่พบปัญหาใด ในส่วนของโมเดลนั้นก็มีความแม่นยำค่อนข้างสูง ใกล้เคียงกับความเป็นจริงเหมาะสมที่จะนำมาใช้ได้ แต่ในด้านของเงื่อนไขสำหรับการจำแนกความผิดปกติในตลาดหลักทรัพย์ซึ่งใช้ในการคาดการณ์ทิศทางการเปลี่ยนแปลงของหลักทรัพย์ในตลาดหลักทรัพย์ยังไม่เป็นที่น่าพอใจ เนื่องจากยังไม่ครอบคลุมในจุดที่ตลาดหลักทรัพย์มีความเคลื่อนไหวเพียงเล็กน้อย

**5.2 ปัญหาและอุปสรรค**

* 1. เนื่องจากเงื่อนไขในการจำแนกความผิดปกติในตลาดหลักทรัพย์ยังไม่ครอบคลุมในส่วนที่ตลาดหลักทรัพย์มีความเปลี่ยนแปลงไม่ชัดเจนมากนัก ทำให้ผลการตรวจสอบในช่วงนั้นคลาดเคลื่อนออกไปจากที่คาดการณ์ไว้เล็กน้อย

**5.3 แนวทางในการพัฒนาต่อ**

* 1. โดยจะพัฒนาเงื่อนไขในการจำแนกความผิดปกติในตลาดหลักทรัพย์ให้สามารถครอบคลุมในจุดที่หลักทรัพย์มีการเคลื่อนไหวแค่เพียงเล็กน้อยได้ใกล้เคียงความจริงมากที่สุด
  2. จะพัฒนาแอพพลิเคชันหรือเว็บแอพ เพื่อแสดงข้อมูลความผิดปกติของหลักทรัพย์ดัชนีเซ็ท50 ในตลาดหลักทรัพย์แห่งประเทศไทย