

การคาดการณ์ราคาดัชนีหุ้น

62090500421-1, วุฒิภัทร, 62090500431-2, นฤดา, 62090500434-3, ปวีณนุช,
62090500441-4, สุตาภัทร, 62090500442-5, อภิษฎา, 62090500448-6, โยชิตา

บทคัดย่อ

รายงานนี้ได้นำเสนอวิธีการทำนายแนวโน้มและคาดการณ์ว่าดัชนีหุ้นจะขึ้นหรือลง ตั้งแต่เดือนมกราคม 2014 ถึง ธันวาคม 2021 เป็นเวลา 6 ปี ในรูปแบบประเภท Ensemble Classification และ Regression ถัดไปคณะผู้จัดทำก็ได้ใช้วิธีการ Feature Selection เพื่อหา Technical Indicators ที่เหมาะสมมาเบื้องต้นมา 5 Indicator จากทั้งหมด 10 Indicator โดยโมเดลของ Ensemble Classification ที่ใช้ จะประกอบไปด้วยหลาย ๆ อัลกอริทึมรวมกัน คือ Decision Tree, Random Forest, Artificial Neural Networks และ Logistic regression ในส่วนของ Regression จะใช้ Linear Regression และ AdaBoost Regressor เพื่อมาทำนายแนวโน้มและคาดการณ์ราคาปิดของตลาดในอนาคต โดยคณะผู้จัดทำจะลองนำวิธีการต่าง ๆ มาปรับและนำมาใช้กับหุ้นต่าง ๆ โดยหุ้นที่ผู้จัดทำเลือกมา คือหุ้น 4 ตัวที่ได้รับการคัดเลือกโดยอาจารย์ที่ปรึกษารายวิชา ได้แก่ Dow Jones Industrial Average (DJIA), Nikkei 225 Stock Average (N225), Heng Seng Index (HSI) และ SSE Composite Index (SSE) หลังจากได้ผ่านการทำ Ensemble Classification หรือ Regression แล้วทางคณะผู้จัดทำก็จะนำผลลัพธ์มาวิเคราะห์และพิจารณาว่าสุดท้ายแล้วควรตัดสินใจว่าควรซื้อหรือขายหุ้นตัวนั้นหรือไม่ โดยจากการที่คณะผู้จัดทำได้ทำการทดลองและทำนายผลลัพธ์ของดัชนีหุ้นของวันที่ 8 ธันวาคม 2021 ที่ได้จากการทดลองและทำนายทั้งหมดสรุปได้ว่า หุ้น DJIA, NIK, HSI และ SSE ทั้งหมดมีแนวโน้มว่าดัชนีกำลังจะขึ้น เหมาะสมที่นักลงทุนจะขายในช่วงเวลานี้

1. บทนำ

ในปัจจุบันผู้คนเริ่มหันมาให้ความสนใจกับการลงทุนเกี่ยวกับหุ้นมากขึ้น ทำให้มีการศึกษาเรื่องการทำนายแนวโน้มของหุ้นด้วยหลากหลายเครื่องมือ ซึ่งการทำนายในที่นี้นั้นใช้งานเพื่อเป็นการเพิ่มความมั่นใจให้แก่ผู้ลงทุนก่อนตัดสินใจซื้อหรือขายหุ้นชนิดนั้น อย่างไรก็ตามการทำนายแนวโน้มของหุ้นค่อนข้างมีความซับซ้อนเนื่องจากในตลาดหุ้นมีความผันผวนสูง และในตอนนี้ก็ได้มีการทำการศึกษาและมีการวิจัยออกมาอย่างมากมาย แต่ในรายงานเล่มนี้ทางคณะผู้จัดทำให้ความสำคัญกับเทคนิคที่เป็นการนำข้อมูลย้อนหลังของหุ้นแต่ละตัวผ่านอัลกอริทึม เพื่อใช้ทำนายแนวโน้มของหุ้นแล้วสามารถสร้างโมเดลออกมา 2 ประเภท ได้แก่ Classification และ Regression ในส่วนของความต่างของ Classification กับ Regression คือ ลักษณะของ Target ที่ต่างกัน Target ของ Classification คือ ข้อมูลที่เป็นลักษณะกลุ่ม หรือข้อมูลที่ไม่มีความต่อเนื่องกัน โดย Target ของทางคณะผู้จัดทำ จะใช้ชื่อคอลัมน์ว่า UpDown ซึ่งเป็นคอลัมน์ที่ได้มาจากการเปรียบเทียบค่า close ของวันนี้กับวันก่อนหน้า หากค่า close ของวันนี้มีค่ามากกว่าวันก่อนหน้าจะเก็บเป็น 1 และหากค่า close ของวันนี้มีค่าน้อยกว่าวันก่อนหน้าจะเก็บเป็น 0 และในส่วน Target ของ Regression คือ ข้อมูลที่มีความต่อเนื่อง ในที่นี้คือคอลัมน์ชื่อ Future Close

ในรายงานเล่มนี้ทางคณะผู้จัดทำได้เลือกใช้ทั้ง 2 โมเดลด้วยกันคือ Classification และ Regression โดยในส่วนของ Classification ทางคณะผู้จัดทำได้คัดเลือกมา 4 อัลกอริทึม ได้แก่ Decision tree, Random Forest, MLP Artificial Neural Network (ANN) และ Logistic Regression ถัดไปในส่วนของ Regression ทางคณะผู้จัดทำก็ได้คัดเลือกและนำมาใช้ทั้งหมด 2 อัลกอริทึม นั่นคือ AdaBoostRegressor และ Linear regression โดยทั้ง 2 โมเดล ที่ทางคณะผู้จัดทำได้ใช้นั้นเพื่อสำหรับทำนายแนวโน้มของดัชนีหุ้นทั้ง 4 ตัว ได้แก่ DJIA, N225, HSI และ SSE โดยมีระยะเวลาของหุ้นที่ใช้ย้อนหลังคือ ช่วงระยะเวลาดังแต่เดือนมกราคม 2014 จนถึง ธันวาคม 2021 นับเป็นเวลา 6 ปี โดยทางคณะผู้จัดทำมีเป้าหมายในการจัดทำรายงานนี้ขึ้นเพื่อเป็นแนวทางในการตัดสินใจสำหรับนักลงทุนหรือผู้ที่สนใจลงทุนต่อไป

2. งานที่เกี่ยวข้อง

2.1. การวิเคราะห์และการใช้ Technical Indicator

Technical Indicators เป็นเครื่องมือในการวัดความแกว่งตัว เพื่อช่วยในการชี้ทิศของตลาดระยะสั้นถึงปานกลาง สามารถบอกแนวโน้มของราคาว่ายังคงเข้มแข็ง หรือกำลังอ่อนแรงลงได้ ด้วยเทคโนโลยีในปัจจุบัน Technical Indicators นั้นมีอยู่มากมาย มากกว่า 100 ตัว แต่ทางผู้จัดทำก็ได้เลือก Technical Indicators ที่คาดว่าจะเหมาะสมกับ Model และ ตลาดหุ้นที่เราต้องการจะ Predict มาทั้งหมด 10 ตัว ดังนี้

1. Moving Average (MA) หรือ เส้นค่าเฉลี่ยเคลื่อนที่ เป็นเครื่องมือที่ใช้วิเคราะห์แนวโน้มหรือทิศทางของตลาด โดยใช้ข้อมูลของดัชนีหุ้นย้อนหลังตามระยะเวลาที่กำหนดมาคำนวณค่าเฉลี่ย โดยทางคณะผู้จัดทำได้กำหนดไว้ 6 ปีคือ เดือนมกราคม 2014 จนถึงเดือนธันวาคม 2021
2. Exponential Moving Average (EMA) เป็นเส้นค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียล โดยจะมองความสัมพันธ์ของราคาหุ้นย้อนหลังแบบถ่วงน้ำหนักในรูปแบบของเลขชี้กำลังโดยให้ความสำคัญกับราคาสุดท้ายมากที่สุด ดังนั้นเส้นค่าเฉลี่ยแบบ EMA จึงเคลื่อนไหวได้เร็วกว่า และตอบสนองต่อการเปลี่ยนแปลงของราคาไว ทั้งยังมีการตั้งค่าวันในกรณีที่ต้องการมองภาพแนวโน้มระยะยาวมากขึ้น โดยทางคณะผู้จัดทำได้ตั้งค่า EMA ไว้ที่ 50 วัน เพื่อใช้สะท้อนการเคลื่อนไหวของราคาหุ้นระยะกลาง และ 200 วัน เพื่อใช้สะท้อนการเคลื่อนไหวของราคาหุ้นระยะยาว

$$EMA_{Today} = Value_{Today} \times \frac{(Smoothing)}{1 + Days} + EMA_{Yesterday} \times \left[1 - \left(\frac{Smoothing}{1 + Days} \right) \right]$$

3. Simple Moving Average (SMA) เป็นเส้นค่าเฉลี่ยเคลื่อนที่รูปแบบหนึ่ง โดยการคำนวณเส้นแบบ SMA จะให้ความสำคัญหรือน้ำหนักกับราคาย้อนหลังเท่า ๆ กันทุกวัน กล่าวคือเอาราคาในอดีตมาหาค่าเฉลี่ยโดยการหารด้วยจำนวนวัน ทั้งยังมีการตั้งค่า SMA อีกด้วยเพื่อหาแนวโน้มของดัชนีหุ้น โดยทางคณะผู้จัดทำได้ตั้งค่า SMA ไว้ที่ 50 วัน เพื่อคำนวณค่าเฉลี่ยเคลื่อนที่อย่างง่ายในระยะกลาง และ 200 วัน เพื่อคำนวณค่าเฉลี่ยเคลื่อนที่อย่างง่ายในระยะยาว หลังจากใช้ SMA ในการหาแนวโน้มของราคาหุ้น พบว่า SMA สามารถตอบสนองต่อการเปลี่ยนแปลงของราคาได้อย่างดีในระยะกลาง แต่อาจตอบสนองในระยะยาวได้ช้ากว่า

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n}$$

โดยที่ A_n คือ ราคาปิดของวันที่ n

n คือ จำนวนวัน หรือ แท่งเทียนที่ใช้คำนวณ

4. Average Directional Index (ADX) เป็น Indicator ที่ใช้วัดความแข็งแกร่งของสภาวะแนวโน้มของตลาด กล่าวคือมันจะวิเคราะห์แนวโน้มว่าดัชนีหุ้นมีโอกาสจบลงหรือยังและวิเคราะห์ว่าสภาวะตลาดปัจจุบันมีความเป็นไซด์เวย์หรือไม่ ซึ่งจะแตกต่างจาก Indicator ตัวอื่นๆ ที่มีวัตถุประสงค์ในการให้สัญญาณเทรดตรงๆ ทำให้ ADX เหมาะสำหรับการเทรดระยะยาว
5. On Balance Volume (OBV) เป็น Indicator ที่ใช้สำหรับวัดความแข็งแกร่งของแนวโน้ม ซึ่งเป็นตัวบอกจำนวนในช่วงเวลาเฉพาะ ที่บอกถึงจำนวนหุ้น สัญญาณ หรือขนาดในการซื้อขายข้อมูลที่ถูกรวบรวมและให้จากตลาดการแลกเปลี่ยน กล่าวคือ OBV จะช่วยบอกว่ามีนักลงทุนเข้ามาซื้อขายหุ้น หรือออกจากตลาดมากน้อยเพียงใด ในบางครั้ง OBV จะแสดงการเปลี่ยนแปลงเกิดขึ้นก่อนราคา เช่น ราคายังคงขยับตัวสูงขึ้นแต่ OBV กลับลดลงแสดงว่าแรงซื้อเริ่ม

เสียลงแล้ว นอกจากนี้ OBV จะช่วยยืนยันแนวโน้มของราคาได้ ทั้งในระยะสั้น และระยะปานกลาง และเป็นสัญญาณเตือนว่าอาจมีการเปลี่ยนแปลงทิศทางของแนวโน้มเกิดขึ้นได้เมื่อ OBV เคลื่อนไหวไปคนละทิศทางกับราคา

$$OBV = OBV_{prev} + \begin{cases} volume & \text{if } close > close_{prev} \\ 0 & \text{if } close = close_{prev} \\ -volume & \text{if } close < close_{prev} \end{cases}$$

If $close > close_{[-1]}$ then
 $OBV = OBV_{[-1]} + volume$
 else If $close < close_{[-1]}$ then
 $OBV = OBV_{[-1]} - volume$
 else
 $OBV = OBV_{[-1]}$

6. Average True Range (ATR) เป็นการบอกระดับความผันผวนของตลาด ส่วนมาก ATR ที่จะใช้บอกแนวโน้มของราคา หรือระดับราคาการซื้อขายสุดโต่ง Overbought หรือ Oversold และมักจะนำไปใช้อ้างอิงร่วมกับ Indicator ที่บอกแนวโน้มของราคา เพื่อยืนยันแนวโน้มให้ถูกต้องชัดเจนมากขึ้น

$$ATR = \left(\frac{1}{n} \right) \sum_{(i=1)}^{(n)} TR_i$$

where:

TR_i = A particular true range

n = The time period employed

7. Relative Strength Index (RSI) เป็น Indicator ที่ใช้ในการวิเคราะห์ข้อมูลของกราฟทางเทคนิค จะมีค่าตัวเลขให้ใช้งานระหว่าง 0-100 ที่ใช้บอกถึงสถานะของตลาดและราคาของสินทรัพย์ตัวนั้นๆ ว่า ซื้อมากเกินไป(Overbought) หรือ ขายมากเกินไป(Oversold) โดยถ้าสถานะ Overbought คือ $RSI > 70$ ราคาหุ้นจะแพงขึ้น ทำให้นักลงทุนสามารถใช้เป็นจังหวะในการขายได้ และถ้าหากมีสถานะ Oversold คือ $RSI < 30$ ราคาหุ้นจะถูกลง ทำให้นักลงทุนสามารถพิจารณาหาจังหวะในการเข้าซื้อได้ ทั้งนี้ RSI ยังสามารถช่วยบอกถึงสัญญาณเตือนการกลับตัวของราคาอีกด้วย เพื่อให้ให้นักลงทุนได้ตัดสินใจในการเข้าซื้อและทำกำไร

$$RSI = 100 - \left(\frac{100}{1 + \frac{Average\ gain}{Average\ loss}} \right)$$

โดยที่ Average gain คือ ค่าเฉลี่ยของอัตราผลตอบแทนที่เป็นบวกย้อนหลัง 14 แท่งเทียน

Average loss คือ ค่าเฉลี่ยของอัตราผลตอบแทนที่เป็นลบย้อนหลัง 14 แท่งเทียน

หมายเหตุ: ส่วนใหญ่การคำนวณจะนิยมใช้ช่วง 14 แท่งเทียน แต่นักลงทุนสามารถกำหนดจำนวนแท่งเทียนที่ใช้ในการคำนวณได้ โดยทางคณะผู้จัดทำได้กำหนดเป็น RSI14

8. William Percent Range (WILLR) หรือ W%R ใช้สำหรับวัดโมเมนตัมการแกว่งตัวของราคา สามารถใช้จับจังหวะในการเทรดได้เป็นอย่างดีเครื่องมือ Williams %R เป็นการวัดรอบการแกว่งตัวของราคา โดยใช้ราคาปิดปัจจุบัน เทียบกับช่วง High-Low ในอดีต (ย้อนหลัง 14 วัน) เพื่อดูว่า ราคาปิดปัจจุบันนั้นอยู่ในช่วงต่ำ หรือสูง เมื่อเทียบกับรอบการแกว่งตัวที่ผ่านมา

$$Williams\ PercentR = \frac{Highest\ High - Close}{Highest\ High - Lowest\ Low}$$

โดยที่	Highest High	คือ ราคาสูงสุดของจำนวนแท่งที่เราใช้ในการคำนวณ ถ้าเป็นค่าเริ่มต้นจะเท่ากับ 14 วัน
	Close	คือ ราคาปิดของแท่งปัจจุบัน
	Lowest Low	คือ ราคาต่ำสุดของจำนวนแท่งที่ใช้ในการคำนวณ ถ้าเป็นค่าเริ่มต้นจะเท่ากับ 14 วัน

9. Moving Average Convergence Divergence (MACD) เป็นการประยุกต์เอาลักษณะของกราฟเส้น Moving Average ที่วาดขึ้นมา 2 เส้นพร้อม ๆ กันมาคำนวณเป็น Indicator ตัวใหม่คือ MACD โดยการคำนวณค่า MACD จึงเป็นการวัดระยะห่างระหว่างเส้นค่าเฉลี่ยเคลื่อนที่ 2 เส้น เพื่อดูว่าเส้นค่าเฉลี่ยเคลื่อนที่ 2 เส้นนี้กำลังเคลื่อนที่ลู่ออกเข้าหากันหรือแยกออกจากกัน โดย Signal Line และ MACD Histogram ซึ่งเป็น Indicator ที่ใช้วิเคราะห์และให้ข้อมูลเกี่ยวกับเส้น MACD อีกต่อหนึ่ง ณ จุดนี้ โดยแท้จริงแล้ว Signal Line และ MACD Histogram ไม่ใช่ Indicator ที่ใช้สำหรับวิเคราะห์กราฟของราคาหุ้น แต่เป็น Indicator ที่ใช้วิเคราะห์ลักษณะที่เกิดขึ้นของเส้น MACD ที่วาดขึ้นมาได้ (Indicator ของ MACD)

$$MACD = EMA(12) - EMA(26)$$

โดยที่	MACD	คือ EMA (12) ที่หักออกด้วย EMA (26)
	EMA (12) และ EMA (26)	คือ เส้นค่าเฉลี่ยเคลื่อนที่ที่ลักษณะเส้นเป็นแบบ Exponential โดยใช้ราคาปิดย้อนหลัง 12 วัน และ 26 วันตามลำดับ

10. Rate-of-Change (ROC) หรือ ROC indicator เป็นเครื่องมือที่วัดโมเมนตัมโดยเนื้อแท้ โดยดูจากเปอร์เซ็นต์การเปลี่ยนแปลงการขึ้นลงของราคาตามระยะเวลาที่กำหนด

วิธีการคำนวณ Price ROC

$$ROC = \frac{(Closing\ Price_T - Closing\ Price_{T-n})}{Closing\ Price_{T-n}}$$

โดยที่	Closing Price _T	คือ ราคาปิดล่าสุด
	Closing Price _{T-n}	คือ ราคาปิดเมื่อ n วันก่อนหน้า
	n	คือ จำนวนระยะเวลา ขึ้นอยู่กับนักลงทุนว่าจะใช้ระยะเวลาเท่าใด สามารถเลือกใช้ตั้งแต่ 1-200 วันหรือมากกว่านั้น

โดยการดูแนวโน้มนั้นเพียงสังเกตว่า

- $ROC > 0$ = แนวโน้มขาขึ้น
- $ROC < 0$ = แนวโน้มขาลง

2.2. โมเดลในการทำนายข้อมูล

2.2.1. การทำงานของ Decision Tree ในระบบจะเริ่มต้นด้วยการกำหนดความลึกของต้นไม้ ซึ่งจะต้องปรับไปเรื่อย ๆ ตามความเหมาะสมของ Dataset ต่อมาจะแบ่งข้อมูลเป็น 2 ชุด ชุดแรกคือ Training dataset และ ชุดที่สองคือ Test dataset โดยระบบจะเริ่มเรียนรู้จาก Training dataset ก่อน ซึ่งจะมี method ที่ระบบจะได้ใช้ คือ Top-Down Method กล่าวคือ ระบบจะทำ Decision tree เริ่มจากด้านบนที่เป็น Root Node ไหลลงไปยังสุดท้ายคือ Leaf Node หลังจากที่เราได้เริ่มเรียนรู้กับข้อมูล Training dataset ไปซักระยะ ระบบจะค่อย ๆ ปรับและแบ่ง Node โดยใช้ Technical Indicator เป็นตัวกำหนดว่าจะแบ่งไป Yes หรือ No และควรแบ่งด้วยค่า Technical Indicator เท่าไหร่จึงจะเหมาะสม โดยส่วนที่จะตัดสินใจว่าตัวไหนควรอยู่ด้านบน หรืออยู่ด้านล่างจะใช้ฟังก์ชันที่เรียกว่า Discrete Splitting Function ฟังก์ชันนี้จะใช้ข้อมูลต่าง ๆ ไม่ว่าจะเป็น Information Gain หรือ Gini index เพื่อตัดสินใจว่า Technical Indicators ไหนควรอยู่ด้านบนสุด (Root Node) หลังจากนั้นระบบจะนำไปใช้กับข้อมูล Test Dataset เพื่อทดลองทำนายว่าข้อมูลออกมาเป็นอย่างไร ยกตัวอย่างให้เห็นภาพ เริ่มต้น Root Node อาจจะมีเงื่อนไขเป็น Technical Indicators RSI > 50 ถ้าตรงตามเงื่อนไข จะไปทางด้านซ้าย เงื่อนไขที่สองอาจจะเป็น SMA > 80 ถ้าตรงตามเงื่อนไขอีก ก็จะไปทางด้านซ้ายเหมือนเดิม ทำอย่างนี้ตามลำดับไปเรื่อย ๆ จนถึง Leaf Node เท่าที่ความลึกของ Decision Tree กำหนดไว้ตอนแรก จนสุดท้ายจะได้ผลลัพธ์ว่า ขึ้น หรือ ลง

2.2.2. ถัดไปการทำงานของ Random Forest คือวิธีการนำโมเดล Decision Tree มาหลาย ๆ โมเดล เมื่อจัดประเภทของข้อมูลในแต่ละโมเดลเสร็จ จะนำผลของทุกโมเดลมารวมกันแล้วดูค่า mode ที่มีค่ามากที่สุดออกมาเป็นผลลัพธ์วิธีนี้ใช้เพื่อแก้ปัญหา overfitting ของโมเดล Decision Tree และสามารถเพิ่มความแม่นยำในการพยากรณ์ให้กับโมเดลมากขึ้น

2.2.3. Artificial Neural Networks หรือ ANN นั้นมีแนวคิดริเริ่มจากการนำค่า target มาคำนวณกับค่า weight ในแต่ละ node เมื่อได้ผลรวมออกมาจะนำไปคำนวณผ่าน activation function ที่สามารถปรับใช้ให้เหมาะสมกับ feature target ที่ใช้ โดย Artificial Neural Networks สามารถปรับ hidden layer ได้ตามความเหมาะสมกับประเภท data set เมื่อโมเดลให้ output ออกมาหลังจากนั้นโมเดลจะนำ output มาหาค่า Loss function ออกมาจะได้ Loss score แล้วนำค่านี้มาทำการ optimizing แล้วนำไป update ค่า weight การทำงานของ Artificial Neural Networks จะเป็นแบบ feed forward ไปเรื่อย ๆ ไม่มีการย้อนกลับจนได้ผลลัพธ์ที่ optimal ออกมา

$$Output = F\left(\sum_{i=1}^n x_i \bullet w_i\right)$$

2.2.4. การทำงานของ Logistic Regression จะทำนายกับคอลัมน์ y ที่เป็น Binary outcome เท่านั้น ก็คือ 0 กับ 1 เช่น คอลัมน์ UpDown ของคณะผู้จัดทำที่มีค่าแค่ 0 (หูลง) กับ 1 (หูลขึ้น) แต่ถ้าค่าในคอลัมน์ y ไม่ได้เป็น 0 กับ 1 ก็จะมีการใช้ความน่าจะเป็นและ sigmoid function เข้ามาช่วยในการคำนวณ

2.2.5. การทำงานของ Linear Regression เป็นการทำนายข้อมูลด้วยเส้นตรง โดยมีหลักการคือ ทางคณะผู้จัดทำจะนำข้อมูลที่ต้องการเทรนมาใส่ในโมเดล แล้วจะมีการเปรียบเทียบค่า X กับ y โดยเมื่อทำการประมวลผลให้ได้ฟังก์ชันสมการเส้นตรงออกมา แล้วหลังจากนั้นให้ทางคณะผู้จัดทำจะนำโมเดลที่ได้นี้มาใส่ข้อมูลที่เราต้องการเพื่อทำนายค่าออกมา แล้วนำค่าที่ได้จากการทำนายมาคำนวณหาค่า Accuracy

$$Y \approx \beta_0 + \beta_1 X$$

2.2.6. AdaBoost เป็นการใช้เทคนิค Ensembled method ซึ่งอัลกอริทึมที่ใช้จะเป็น Decision Tree ที่มีการแบ่งข้อมูลครั้งเดียว อัลกอริทึมนี้จะสร้างแบบจำลองและให้ค่า weight เท่ากันกับทุกข้อมูลที่น่าเชื่อถือแล้วทำการ classified ออกมา 1 โมเดล หลังจากนั้นจะกำหนดค่า weight ของค่าที่ทำการ classified ผิดให้มีค่ามากขึ้นแล้วทำการ classify

ใหม่ให้เป็นอีกโมเดลที่ได้จากโมเดลเดิมในขั้นตอนนี้ค่า weight ที่ classified มาผิดจะถูกโฟกัสเพราะค่า weight มีค่ามากกว่าตัวอื่นทำให้โมเดลจะทำการแบ่งข้อมูลออกมาโดยได้แบ่งข้อมูลตัวที่มีค่า weight มากออกมาอีกครั้งเพื่อทำให้ error ลดลงขั้นตอนนี้แต่ละโมเดลยังเป็น weak learner อยู่จึงต้องมีการ นำหลาย ๆ โมเดลที่แบ่งข้อมูลออกมานี้มาทำการหาค่าเฉลี่ย weight อีกครั้งเพื่อให้ได้โมเดลที่ classify ได้แม่นยำมากขึ้นขั้นตอนนี้เรียกว่า strong learner AdaBoost จะมีความยืดหยุ่นมากซึ่งปรับใช้ได้กับทุก Learning algorithm ไม่ว่าจะเป็น Classification หรือ Regression โดยใช้เทคนิคของการคำนวณซ้ำ แล้วปรับค่าน้ำหนัก ทำให้ลด Bias Error ลงได้ และไม่ทำให้เกิดการ over fitting เกินไป

3. แผนการทำงาน

3.1. วัตถุประสงค์

1. เพื่อคาดการณ์ดัชนีขึ้นหรือลงของหุ้นทั้ง 4 ตัว ได้แก่ DJIA, N225, HSI และ SSE
2. เพื่อทำนายแนวโน้มและคาดการณ์ค่าราคาปิดของตลาดในอนาคต โดยใช้ข้อมูลราคาปิดในตลาดตั้งแต่เดือนมกราคม 2014 ถึง ธันวาคม 2021

3.2. การรวบรวมและเตรียมข้อมูล

ทางคณะผู้จัดทำจะทำการทดลองโดยใช้ข้อมูลของดัชนีหุ้น 4 ตัว คือ DJIA, N225, HIS และ SSE โดยทางคณะผู้จัดทำจะดึงข้อมูลดัชนีหุ้นมาจากเว็บไซต์ Yahoo! Finance ผ่านการใช้คำสั่ง `get_data_yahoo` ซึ่งเป็น API ข้อมูลดัชนีหุ้นที่ดึงมาทั้งหมดจะเป็นข้อมูลตั้งแต่วันที่ 1 เดือนมกราคม 2014 จนถึง วันที่ 31 เดือนธันวาคม 2021 นับเป็นเวลา 6 ปี หลังจากนั้นทางคณะผู้จัดทำจะแบ่งข้อมูลของดัชนีหุ้นที่ได้รับมานี้เป็น 2 ชุด ซึ่งก็คือ (1) Training Dataset (2) Test Dataset ซึ่งการแบ่งข้อมูลนี้เป็นการแบ่งเพื่อสำหรับการเตรียมความพร้อมก่อนจะนำข้อมูลไป Train โดยข้อมูลในส่วนของ Training Dataset มีไว้สำหรับเรียนรู้ และในส่วนของ Test Dataset มีไว้สำหรับทดลองและเปรียบเทียบเพื่อหาค่าความแม่นยำ

3.3. Feature Selection

3.3.1 Classification

1. Target ที่ต้องการจะนำไปใช้กับ classification เรากำหนดจากการเทียบราคาปิดของวันนี้และราคาปิดวันก่อนหน้า
2. เมื่อเราได้ Target ที่ต้องการจะนำไปวิเคราะห์ Feature selection โดยใช้วิธี Recursive Feature Elimination (RFE)
3. วิธีนี้จะเป็นการเลือก Feature ที่เหมาะสมมาใช้เองโดยไม่ได้กำหนดจำนวน feature ที่จะนำมาใช้ แต่จะแสดง ranking มาให้ว่าตัวไหนเหมาะสมที่สุด

3.3.2 Regression

1. Target ที่จะนำไปใช้สำหรับ Regression คือคอลัมน์ที่ชื่อ Future Close หรือค่า Close ของวันถัดไป
2. เมื่อได้ Target ที่ต้องการแล้วจะนำไปวิเคราะห์ Feature selection โดยใช้วิธี Recursive Feature Elimination (RFE)

โดยในส่วนของ Regression ทางคณะผู้จัดทำได้ทำการ Drop Technical indicator บางส่วนออก ได้แก่ Moving Average (MA) และ Exponential Moving Average (EMA) เพราะ เมื่อนำไปคำนวณแล้วค่าที่ออกมามีความผิดปกติ นั่นคือการพบว่ามี ค่าความแม่นยำที่สูงจนเกินความเป็นจริง

3.4. Classification & Regression

Target ของ Classification คือ ข้อมูลที่เป็นลักษณะกลุ่ม หรือข้อมูลที่ไม่มีความต่อเนื่องกัน ดังนั้นทางคณะผู้จัดทำจะใช้ข้อมูล ดัชนีหุ้นของแต่ละวันมาเปรียบเทียบกันแล้วหลังจากนั้นจะนำมาทำนาย โดยกระบวนการคือคณะผู้จัดทำจะใช้ราคาปิดของวัน หนึ่งเปรียบเทียบกับราคาปิดของวันก่อนหน้าว่าดัชนีหุ้นตัวนี้ราคาขึ้นหรือลง โดยจะกำหนดให้ 1 เป็นราคาขึ้น และ 0 เป็นราคา ลง หลังจากนั้นจะนำมาเก็บใส่คอลัมน์ UpDown และให้อัลกอริทึมต่าง ๆ ใน Classification หาผลลัพธ์ในการทำนาย และใน ส่วนของ Regression นั้น Target โดยใช้ค่าราคาปิดของวันถัดไปได้เลย

หลังจากนั้นจะนำ Training Dataset ไปเรียนรู้กับอัลกอริทึมต่าง ๆ จากทั้ง 2 โมเดลเมื่อเรียนรู้เสร็จระบบจะทำการทดลองและ ทำนายราคาดัชนีหุ้นจาก Test Dataset สุดท้ายจะนำข้อมูลที่ทำนายเปรียบเทียบกับข้อมูลจริงและสรุปผลลัพธ์ออกมาเป็นค่าที่ สามารถแสดงความแม่นยำต่าง ๆ

3.5. Evaluation

โดยในส่วนนี้จะเป็นการอธิบายความหมายและสูตรคำนวณต่าง ๆ ที่ใช้ในการคำนวณหาค่าความแม่นยำต่าง ๆ

3.5.1 Classification

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

โดย True Positive (TP) คือ จำนวนที่โปรแกรมทำนายว่าดัชนีหุ้นขึ้นแล้วผลลัพธ์ออกมาเป็นค่าถูกต้อง ตรงกัน

โดย False Positive (FP) คือ จำนวนที่โปรแกรมทำนายว่าดัชนีหุ้นขึ้นแล้วผลลัพธ์ออกมาเป็นค่าไม่ถูกต้อง ต่างกัน

โดย True Negative (TN) คือ จำนวนที่โปรแกรมทำนายว่าดัชนีหุ้นลงแล้วผลลัพธ์ออกมาเป็นค่าถูกต้อง ตรงกัน

โดย False Negative (FN) คือ จำนวนที่โปรแกรมทำนายว่าดัชนีหุ้นลงแล้วผลลัพธ์ออกมาเป็นค่าไม่ถูกต้อง ต่างกัน

3.5.1.1 Accuracy

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\%$$

3.5.1.2 Precision

$$Precision = \frac{TP}{TP + FP}$$

3.5.1.3 Recall

$$Recall = \frac{TP}{TP + FN}$$

3.5.1.4 F1-Score

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

3.5.2 Regression

3.5.2.1 Root Mean Squared Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}$$

3.5.2.2 R-squared

$$R^2 = 1 - \frac{SSE}{SST}$$

$$Sum\ of\ Squares\ Error \rightarrow SSE = \sum (y - y')^2$$

$$Sum\ of\ Squares\ Total \rightarrow SST = \sum (y - \bar{y})^2$$

3.6. การปรับค่าพารามิเตอร์

การปรับค่าพารามิเตอร์ทางคณะผู้จัดทำจะทำการทดลองและปรับค่าพารามิเตอร์ตามที่เหมาะสม จะคัดเลือกพยายามปรับเพื่อให้แต่ละอัลกอริทึมมีค่า Accuracy >= 0.5 โดยแต่ละโมเดลและอัลกอริทึมจะใช้พารามิเตอร์ที่ต่างกันโดยทางคณะผู้จัดทำได้คัดเลือกและทำการทดลองโดยมีค่าพารามิเตอร์ของแต่ละอัลกอริทึมดังนี้

Decision tree

max_depth[none, 3]

criterion['gini','entropy']

min_samples_leaf[2,5]

max_depth[none,2]

min_samples_leaf[0.35,2]

splitter['random','best']

Random Forest

n_estimators[100, 200, 300, 400, 500, 600]

random_state[none,2, 42]

max_depth[none,2,15]

Artificial Neural Networks

max_iter[200,2000]

random_state[none,0]

activation['relu','logistic']

Logistic Regression

random_state[none,0]

max_iter[100,1000]

3.7. การแสดงค่าผลลัพธ์

ผลลัพธ์ของ Classification ทางคณะผู้จัดทำจะนำข้อมูล Accuracy ของแต่ละอัลกอริทึมมาเพื่อเรียงลำดับ และเลือกผลลัพธ์ของการทำนายว่าดัชนีหุ้นจะขึ้นหรือลงเฉพาะข้อมูลที่มี Accuracy สูงที่สุด 3 ตัว หลังจากนั้นจะนำผลลัพธ์ของทั้ง 3 อัลกอริทึมมาทำการ Majority Vote และนำผลลัพธ์ที่ได้จากการสรุป Majority Vote มาเพื่อทำนายผลลัพธ์ของดัชนีหุ้นในวันที่ 6 ธันวาคม 2021

4. การทดลอง

4.1. วิธีการประเมินผลและข้อมูลที่ใช้

1. Data Acquisition and Preparation: ดาวน์โหลดข้อมูลของหุ้นที่สนใจ ซึ่งประกอบด้วย open, low, high และ close ในช่วงเวลาที่สนใจ เพื่อนำมาคำนวณกับ 11 Indicator ที่คณะผู้จัดทำได้เลือกไว้ จาก TA-Lib ในส่วนของ dataset จะใช้ข้อมูลตั้งแต่วันที่ 1 เดือนมกราคม 2014 ถึง 31 ธันวาคม 2021 หรือประมาณ 6 ปี โดยจะทำการทดสอบโดยแบ่งเป็น 2 ช่วงเวลา คือระยะเวลา 4 ปี นั่นก็คือ วันที่ 1 เดือนมกราคม 2016 ถึง 31 ธันวาคม 2021 และ ระยะเวลา 6 ปี นั่นก็คือ วันที่ 1 เดือนมกราคม 2014 ถึง 31 ธันวาคม 2021 เมื่อนำมาคำนวณกับฟังก์ชัน generateUpDown แล้วจะให้ค่าอยู่ในช่วง [0,1] โดยเก็บค่าในคอลัมน์ที่ชื่อว่า UpDown ก่อนจะนำไปเข้ากระบวนการถัดไป
2. Feature Selection: เป็นขั้นตอนการคัดเลือก feature ที่เหมาะสมจาก 10 Indicator โดยคณะผู้จัดทำจะนำโมเดล RFE มาคัดกรอง feature โดยจะเลือกอัลกอริทึมที่มีความเกี่ยวข้องกับโมเดลของคณะผู้จัดทำที่สนใจมาอย่างละ 1 อัลกอริทึม โดยในการทดลองนี้คณะผู้จัดทำเลือกใช้ Linear Regression มาเป็นตัวแทนในการคัดเลือก feature
3. Training model: หลังจากขั้นตอน feature selection คณะผู้จัดทำจะนำ dataset มาทำการ split ให้เป็น 2 ส่วนคือ ส่วนของ training data และ testing data ในอัตราส่วน 75:25 ในส่วนของ training ไป

เข้าอัลกอริทึมต่าง ๆ โดย classification ที่คณะผู้จัดทำใช้จะมี Decision tree, random forest, Artificial neuron network และ Logistic Regression และ ในส่วนของ Regression คณะผู้จัดทำจะใช้ Linear Regression และ AdaBoostRegressor

4. Testing model: หลังจากคณะผู้จัดทำเอา dataset ไป train เสร็จคณะผู้จัดทำจะนำโมเดลไป test กับ ข้อมูล testing data ดูผลการทดลองพร้อมทั้งคำนวณค่า accuracy ในแต่ละโมเดล
5. Evaluation: ดูผลลัพธ์ทั้ง 6 โมเดลเพื่อมาทำการวิเคราะห์ราคาขึ้น-ลงของหุ้นชนิดนั้น ๆ โดยผ่านฟังก์ชัน FinalPredict เพื่อนำมาตัดสินใจในการซื้อ-ขาย

แสดงผลการเปรียบเทียบ

ตารางแสดงผลการการเปรียบเทียบค่า Accuracy แบบ 4 ปี และ 8 ปี

4 ปี	HSI			N225			SSE			DJI		
	Acc	F1	ROC	Acc	F1	ROC	Acc	F1	ROC	Acc	F1	ROC
DCT	0.472	0.5	0.468	0.46	0.54	0.454	0.51	0.67	0.500	0.58	0.73	0.486
RF	0.51	0.58	0.513	0.47	0.54	0.411	0.5	0.63	0.475	0.58	0.7	0.557
ANN	0.51	0.57	0.472	0.44	0.56	0.433	0.51	0.66	0.501	0.58	0.73	0.531
LGT	0.5	0.6	0.465	0.48	0.54	0.488	0.51	0.51	0.505	0.59	0.73	0.534

8 ปี	HSI			N225			SSE			DJI		
	Acc	F1	ROC	Acc	F1	ROC	Acc	F1	ROC	Acc	F1	ROC
DCT	0.5	0.37	0.512	0.5	0.67	0.500	0.51	0.67	0.480	0.59	0.59	0.500
RF	0.5	0.53	0.48	0.52	0.61	0.564	0.5	0.63	0.475	0.52	0.63	0.503
ANN	0.56	0.6	0.54	0.53	0.64	0.538	0.51	0.66	0.501	0.59	0.74	0.495
LGT	0.53	0.61	0.48	0.56	0.54	0.578	0.51	0.63	0.505	0.59	0.74	0.509

4.2. ผลการทดลอง

ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น HSI ช่วงเวลา 4 ปี

Classification

DCT

ตารางที่ 1 ตาราง Confusion Matrix ของ DCT เป็นเวลา 4 ปี

40	47
55	51

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.47150259067357514

Prediction recall Score: 0.4704510952071134

Prediction F1_score: 0.5

Prediction Precision: 0.5204081632653061

RF

ตารางที่ 2 ตาราง Confusion Matrix ของ RF

35	52
42	64

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5129533678756477

Prediction recall Score: 0.5030362177401865

Prediction F1_score: 0.5765765765765765

Prediction Precision: 0.5517241379310345

ANN

ตารางที่ 3 ตาราง Confusion Matrix ของ ANN

36	51
43	63

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5129533678756477

Prediction recall Score: 0.5040663630448926

Prediction F1_score: 0.5727272727272728

Prediction Precision: 0.5526315789473685

LGT

ตารางที่ 4 ตาราง Confusion Matrix ของ LGT

28	59
38	68

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.49740932642487046

Prediction recall Score: 0.4816742572110171

Prediction F1_score: 0.5836909871244635

Prediction Precision: 0.5354330708661418

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

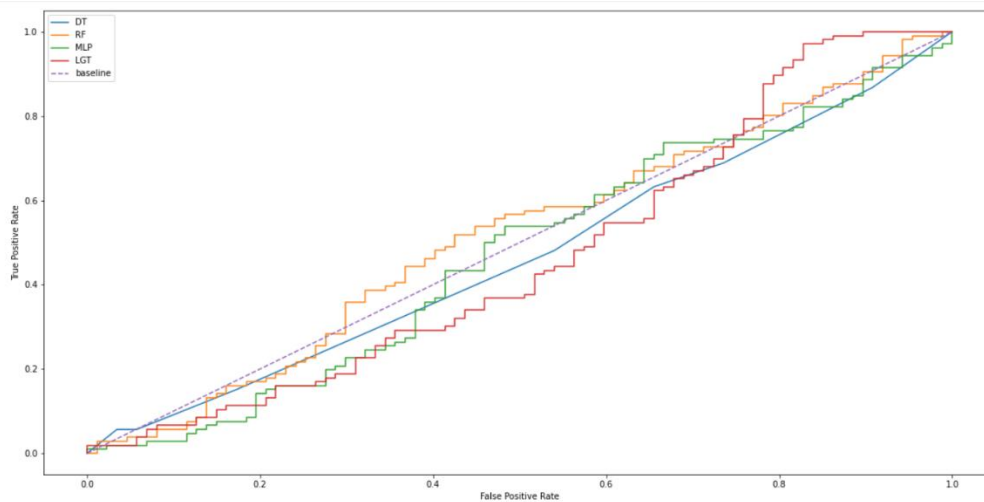
DTC: ROC AUC=0.468

RF: ROC AUC=0.513

MLP: ROC AUC=0.472

LGT: ROC AUC=0.465

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 5 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 4 ปี ของ HSI

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.472	0.47	0.52	0.5	0.468
RF	0.51	0.5	0.55	0.58	0.513
ANN	0.51	0.5	0.55	0.57	0.472
LGT	0.5	0.48	0.54	0.6	0.465

Regression

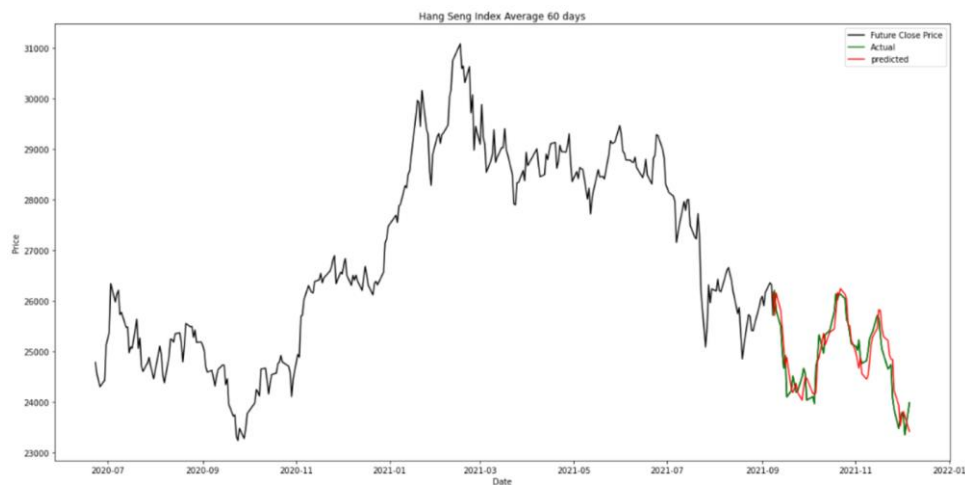
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 388.16883402096437

R2 Score: 0.9506438878262707

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



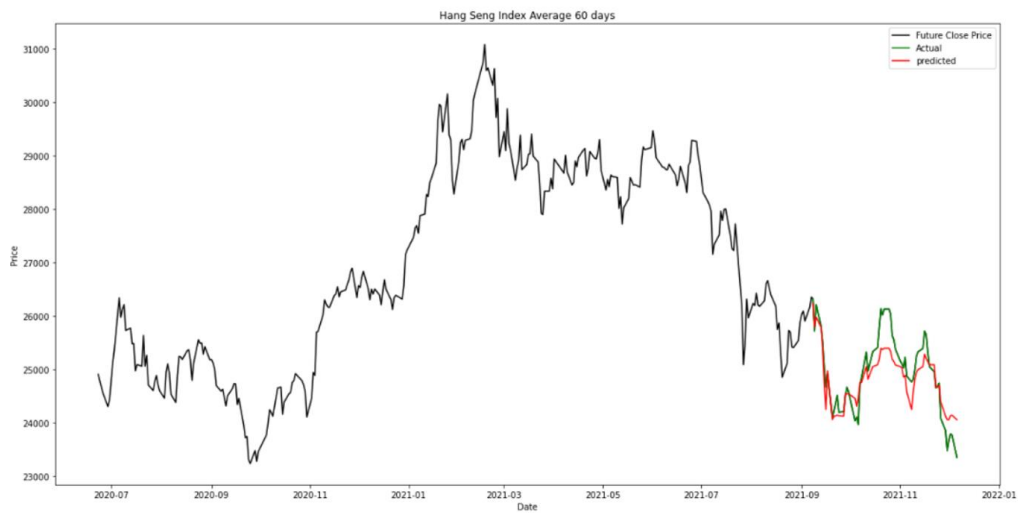
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 523.0869679770188

R2 Score: 0.9091182623755502

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น HSI ช่วงเวลา 8 ปี

Classification

DCT

ตารางที่ 6 ตาราง Confusion Matrix ของ DCT

154	51
170	64

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.49658314350797267

Prediction recall Score: 0.5123618928496978

Prediction F1_score: 0.3667621776504298

Prediction Precision: 0.5565217391304348

RF

ตารางที่ 7 ตาราง Confusion Matrix ของ RF

94	111
110	124

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.49658314350797267

Prediction recall Score: 0.4942255576401918

Prediction F1_score: 0.5287846481876332

Prediction Precision: 0.5276595744680851

ANN

ตารางที่ 8 ตาราง Confusion Matrix ของ ANN

94	111
110	124

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5603644646924829

Prediction recall Score: 0.5552637064832187

Prediction F1_score: 0.6053169734151329

Prediction Precision: 0.5803921568627451

LGT

ตารางที่ 9 ตาราง Confusion Matrix ของ LGT

65	140
68	166

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5261958997722096

Prediction recall Score: 0.5132374400667084

Prediction F1_score: 0.6148148148148148

Prediction Precision: 0.5424836601307189

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

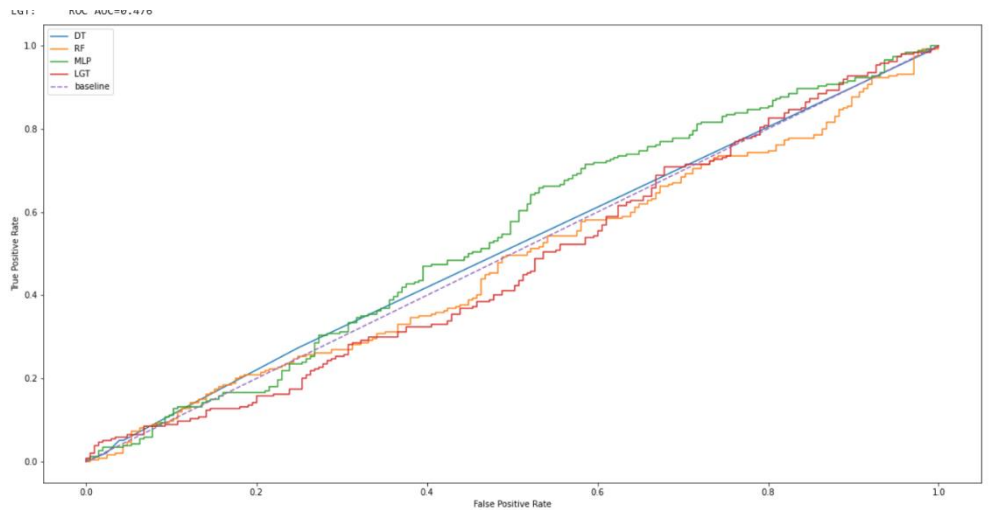
DTC: ROC AUC=0.512

RF: ROC AUC=0.479

MLP: ROC AUC=0.541

LGT: ROC AUC=0.476

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 10 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 8 ปี ของ HSI

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.5	0.51	0.57	0.37	0.512
RF	0.5	0.5	0.53	0.53	0.48
ANN	0.56	0.56	0.58	0.6	0.54
LGT	0.53	0.51	0.54	0.61	0.48

Regression

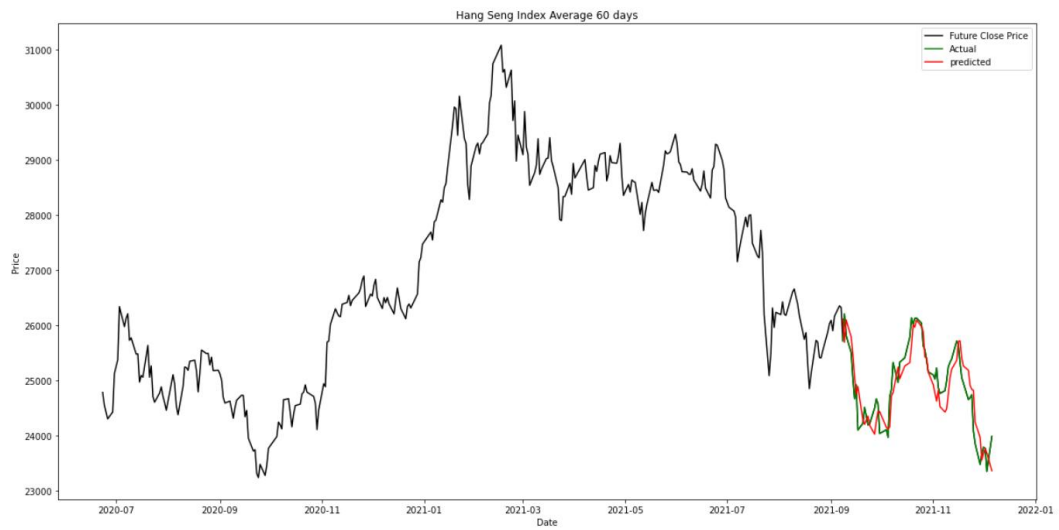
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 366.2097719954414

R2 Score: 0.9831058898499216

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



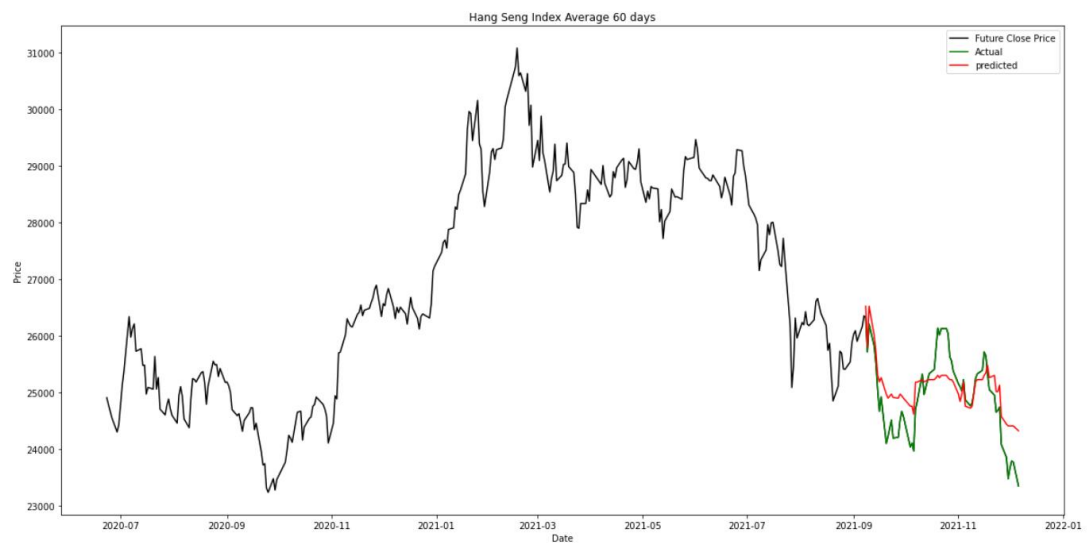
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 690.5791668961314

R2 Score: 0.9399451979522063

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น N225 ช่วงเวลา 4 ปี

Classification

DCT

ตารางที่ 11 ตาราง Confusion Matrix ของ DCT

56	39
58	27

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.4611111111111114

Prediction recall Score: 0.4535603715170278

Prediction F1_score: 0.35761589403973504

Prediction Precision: 0.4090909090909091

RF

ตารางที่ 12 ตาราง Confusion Matrix ของ RF

27	68
28	57

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.4666666666666667

Prediction recall Score: 0.47739938080495353

Prediction F1_score: 0.5428571428571428

Prediction Precision: 0.456

ANN

ตารางที่ 13 ตาราง Confusion Matrix ของ ANN

16	79
22	63

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.4388888888888889

Prediction recall Score: 0.4547987616099071

Prediction F1_score: 0.5550660792951542

Prediction Precision: 0.44366197183098594

LGT

ตารางที่ 14 ตาราง Confusion Matrix ของ LGT

30	65
29	56

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.4777777777777778

Prediction recall Score: 0.4873065015479876

Prediction F1_score: 0.5436893203883494

Prediction Precision: 0.4628099173553719

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

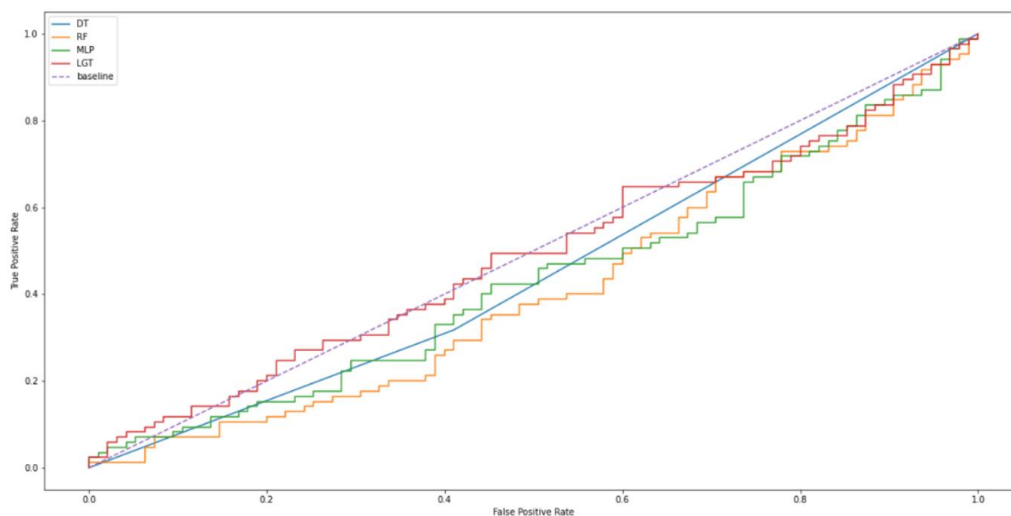
DTC: ROC AUC=0.454

RF: ROC AUC=0.411

MLP: ROC AUC=0.433

LGT: ROC AUC=0.488

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 15 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 4 ปี ของ N225

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.46	0.48	0.41	0.54	0.454
RF	0.47	0.48	0.46	0.54	0.411
ANN	0.44	0.45	0.44	0.56	0.433
LGT	0.48	0.49	0.46	0.54	0.488

Regression

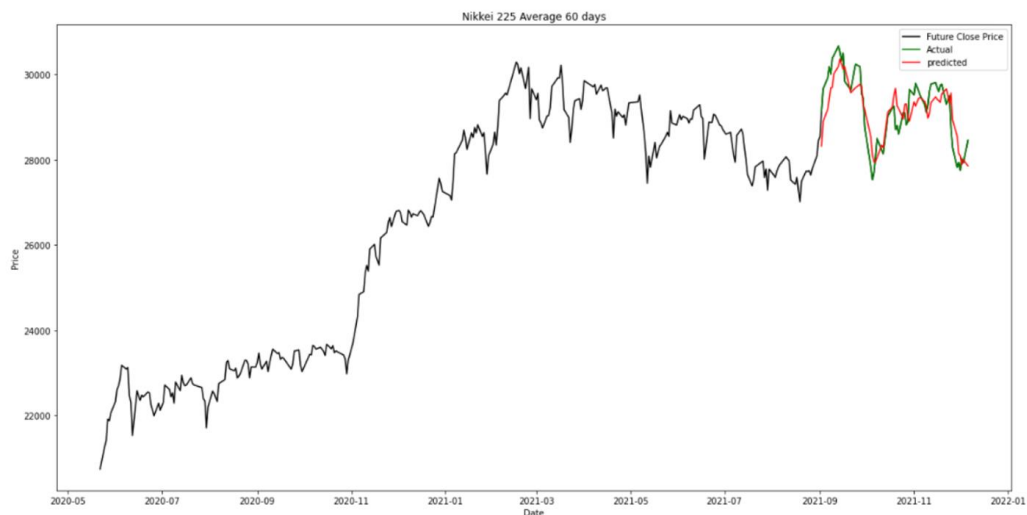
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 373.4221359871856

R2 Score: 0.9885726967339857

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



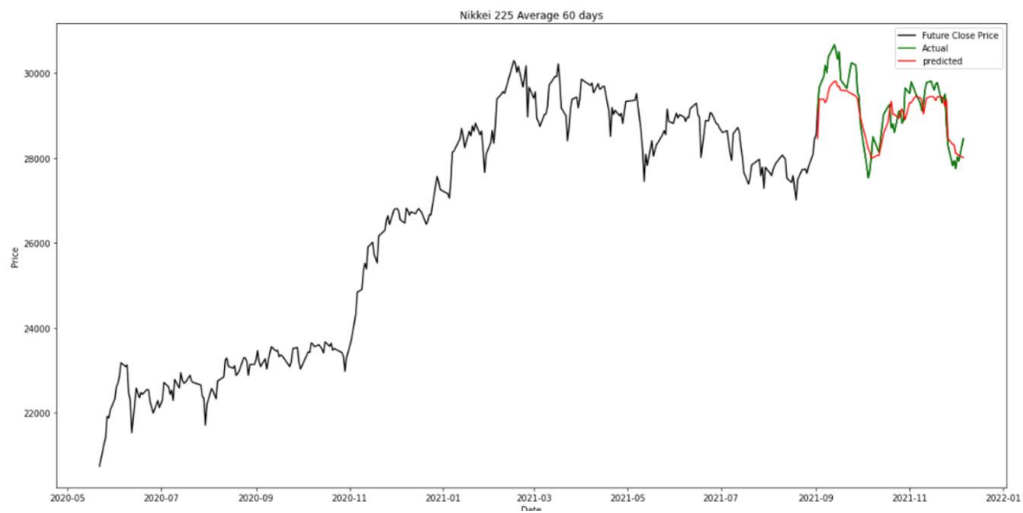
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 468.08636577696717

R2 Score: 0.9820445779324927

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น N225 ช่วงเวลา 8 ปี

Classification

DCT

ตารางที่ 16 ตาราง Confusion Matrix ของ DCT

0	67
0	67

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5

Prediction recall Score: 0.5

Prediction F1_score: 0.6666666666666666

Prediction Precision: 0.5

RF

ตารางที่ 17 ตาราง Confusion Matrix ของ RF

19	48
16	51

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5223880597014925

Prediction recall Score: 0.5223880597014925

Prediction F1_score: 0.6144578313253012

Prediction Precision: 0.5151515151515151

ANN

ตารางที่ 18 ตาราง Confusion Matrix ของ ANN

18	49
13	54

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5373134328358209

Prediction recall Score: 0.5373134328358209

Prediction F1_score: 0.6352941176470589

Prediction Precision: 0.5242718446601942

LGT

ตารางที่ 19 ตาราง Confusion Matrix ของ LGT

24	43
16	51

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5597014925373134

Prediction recall Score: 0.5597014925373134

Prediction F1_score: 0.6335403726708074

Prediction Precision: 0.5425531914893617

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

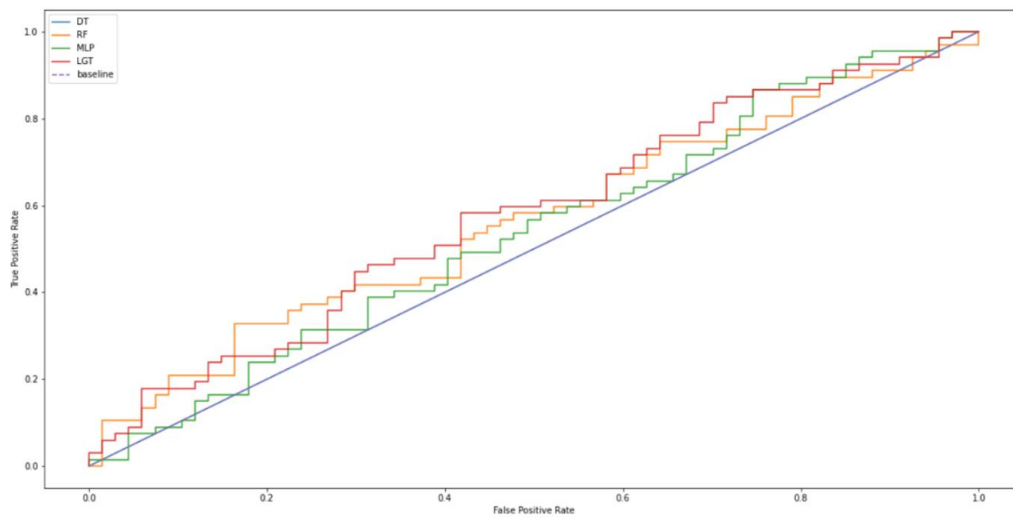
DTC: ROC AUC=0.500

RF: ROC AUC=0.564

MLP: ROC AUC=0.538

LGT: ROC AUC=0.578

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 20 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 8 ปี ของ N225

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.5	0.5	0.5	0.67	0.500
RF	0.52	0.52	0.52	0.61	0.564
ANN	0.53	0.54	0.52	0.64	0.538
LGT	0.56	0.56	0.63	0.54	0.578

Regression

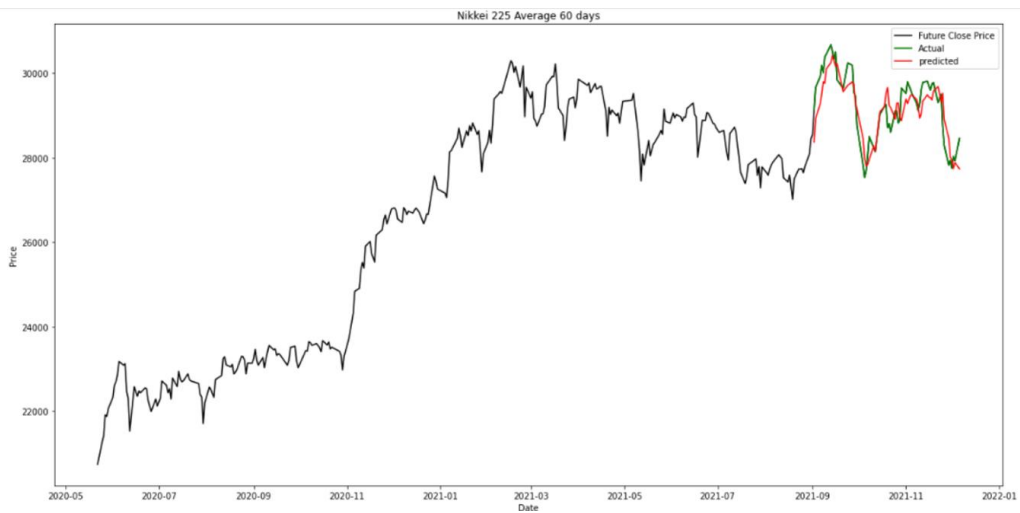
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 406.05278514110705

R2 Score: 0.9861167025979451

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



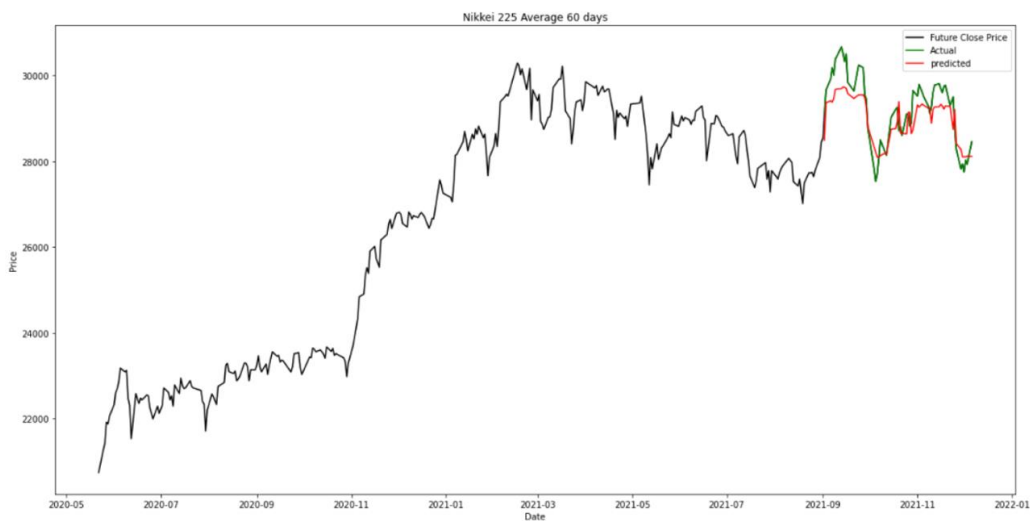
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 484.8258985161621

R2 Score: 0.9802075634371717

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น SSE ช่วงเวลา 4 ปี

Classification

DCT

ตารางที่ 21 ตาราง Confusion Matrix ของ DCT

0	67
0	69

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5073529411764706

Prediction recall Score: 0.5

Prediction F1_score: 0.673170731707317

Prediction Precision: 0.5073529411764706

RF

ตารางที่ 22 ตาราง Confusion Matrix ของ RF

12	55
13	56

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5

Prediction recall Score: 0.4953493402552455

Prediction F1_score: 0.6222222222222222

Prediction Precision: 0.5045045045045045

ANN

ตารางที่ 23 ตาราง Confusion Matrix ของ ANN

7	60
6	63

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5147058823529411

Prediction recall Score: 0.508760545100584

Prediction F1_score: 0.6562499999999999

Prediction Precision: 0.5121951219512195

LGT

ตารางที่ 24 ตาราง Confusion Matrix ของ LGT

13	54
13	56

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5073529411764706

Prediction recall Score: 0.5028120268224097

Prediction F1_score: 0.6256983240223463

Prediction Precision: 0.509090909090909

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

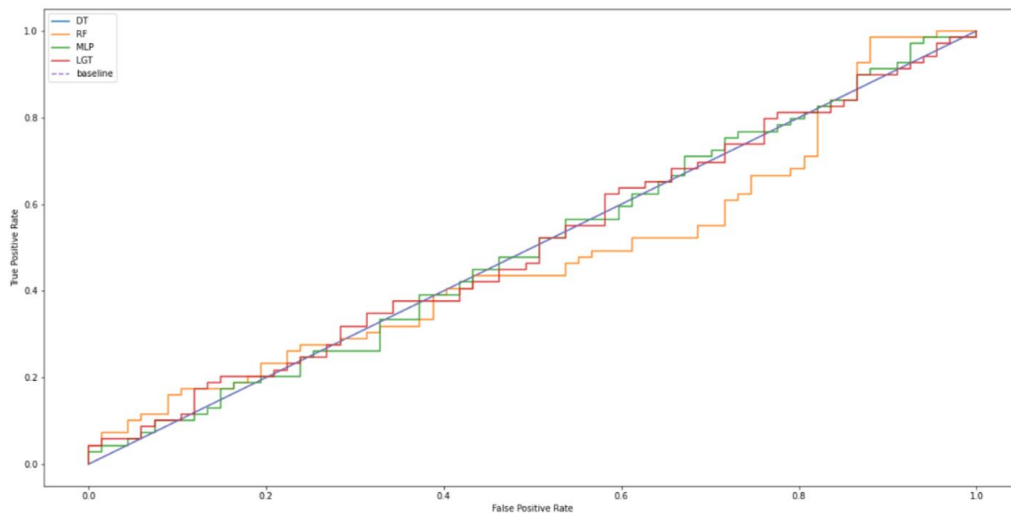
DTC: ROC AUC=0.500

RF: ROC AUC=0.475

MLP: ROC AUC=0.501

LGT: ROC AUC=0.505

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 25 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 4 ปี ของ SSE

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.51	0.5	0.51	0.67	0.500
RF	0.5	0.5	0.5	0.63	0.475
ANN	0.51	0.51	0.51	0.66	0.501
LGT	0.51	0.5	0.63	0.51	0.505

Regression

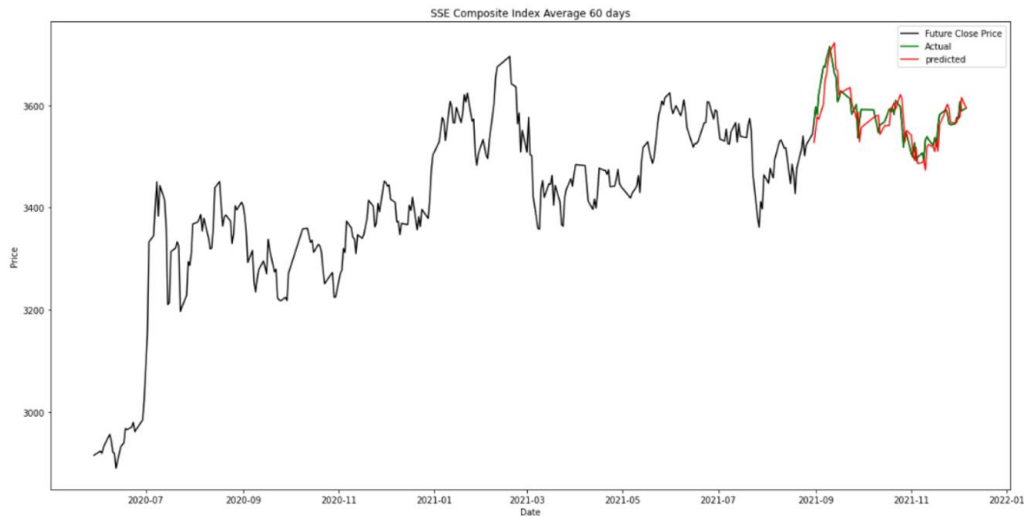
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 38.86793684459121

R2 Score: 0.980373183047781

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



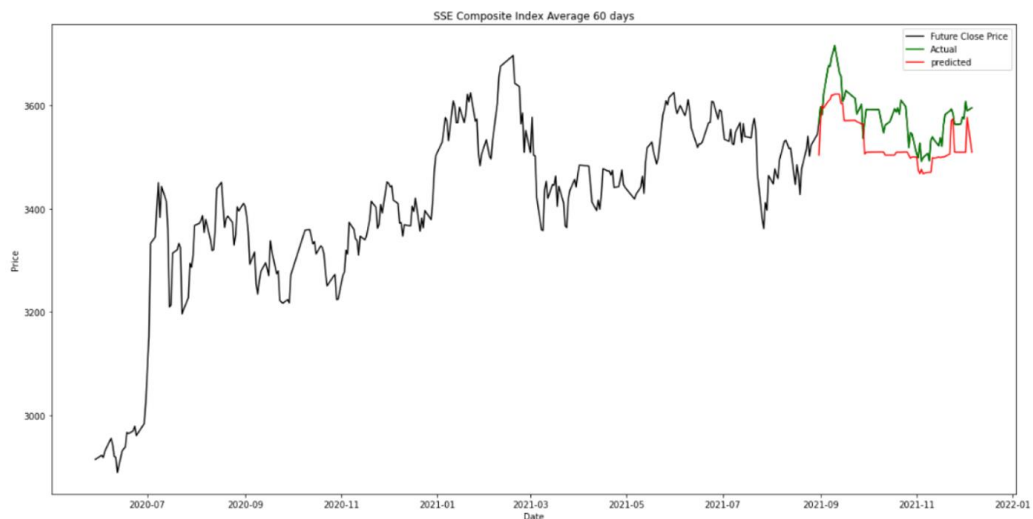
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 55.09967628764117

R2 Score: 0.9605574495190661

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น SSE ช่วงเวลา 8 ปี

DCT

ตารางที่ 26 ตาราง Confusion Matrix ของ DCT

0	67
0	69

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5073529411764706

Prediction recall Score: 0.5

Prediction F1_score: 0.673170731707317

Prediction Precision: 0.5073529411764706

RF

ตารางที่ 27 ตาราง Confusion Matrix ของ RF

12	55
13	56

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5

Prediction recall Score: 0.4953493402552455

Prediction F1_score: 0.6222222222222222

Prediction Precision: 0.5045045045045045

ANN

ตารางที่ 28 ตาราง Confusion Matrix ของ ANN

7	60
6	63

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5147058823529411

Prediction recall Score: 0.508760545100584

Prediction F1_score: 0.6562499999999999

Prediction Precision: 0.5121951219512195

LGT

ตารางที่ 29 ตาราง Confusion Matrix ของ LGT

13	54
13	56

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5073529411764706

Prediction recall Score: 0.5028120268224097

Prediction F1_score: 0.6256983240223463

Prediction Precision: 0.509090909090909

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

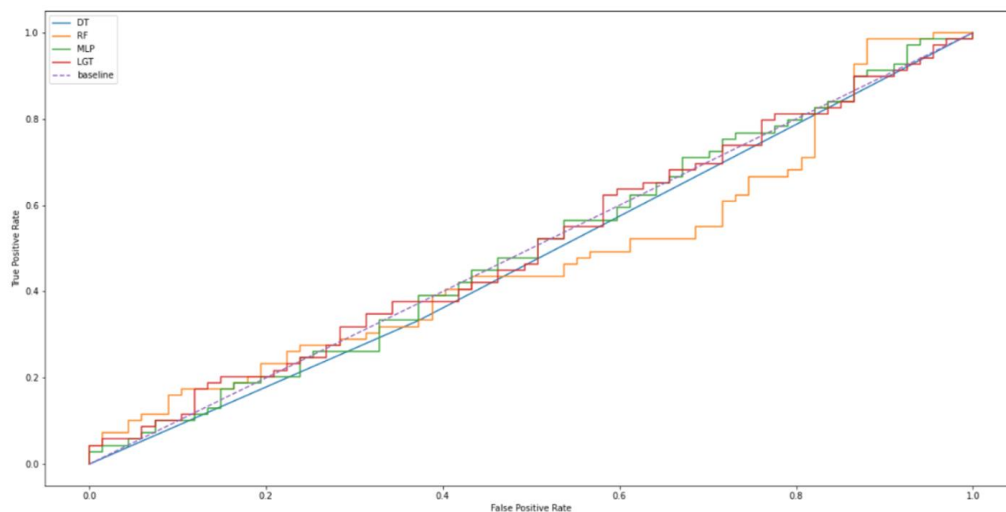
DTC: ROC AUC=0.480

RF: ROC AUC=0.475

MLP: ROC AUC=0.501

LGT: ROC AUC=0.505

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 30 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 8 ปี ของ SSE

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.51	0.5	0.51	0.67	0.480
RF	0.5	0.5	0.5	0.63	0.475
ANN	0.51	0.51	0.51	0.66	0.501
LGT	0.51	0.5	0.51	0.63	0.505

Regression

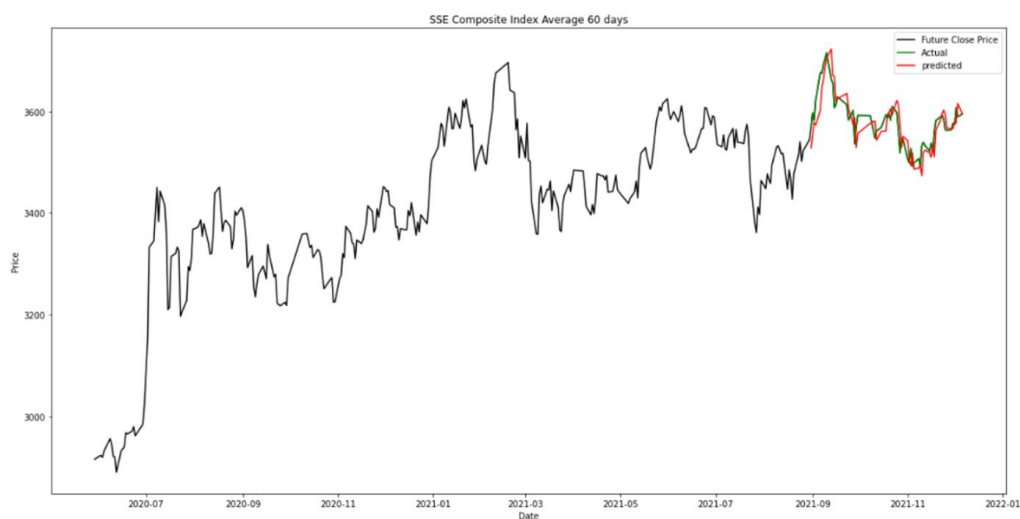
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 38.867936844591114

R2 Score: 0.980373183047781

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



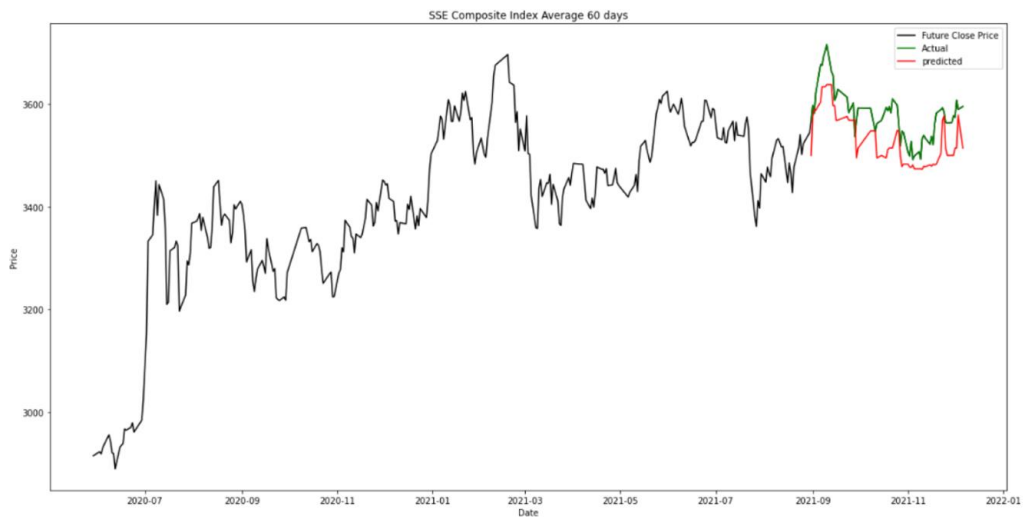
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 54.34987740836997

R2 Score: 0.9616236176019866

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น DJI ช่วงเวลา 4 ปี

DCT

ตารางที่ 31 ตาราง Confusion Matrix ของ DCT

0	137
0	187

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5771604938271605

Prediction recall Score: 0.5

Prediction F1_score: 0.7318982387475538

Prediction Precision: 0.5771604938271605

RF

ตารางที่ 32 ตาราง Confusion Matrix ของ RF

33	104
31	156

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5833333333333334

Prediction recall Score: 0.5375502556696202

Prediction F1_score: 0.697986577181208

Prediction Precision: 0.6

ANN

ตารางที่ 33 ตาราง Confusion Matrix ของ ANN

2	135
0	187

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5833333333333334

Prediction recall Score: 0.5072992700729927

Prediction F1_score: 0.7347740667976425

Prediction Precision: 0.5807453416149069

LGT

ตารางที่ 34 ตาราง Confusion Matrix ของ LGT

10	127
5	182

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5925925925925926

Prediction recall Score: 0.5231273664077443

Prediction F1_score: 0.7338709677419355

Prediction Precision: 0.5889967637540453

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

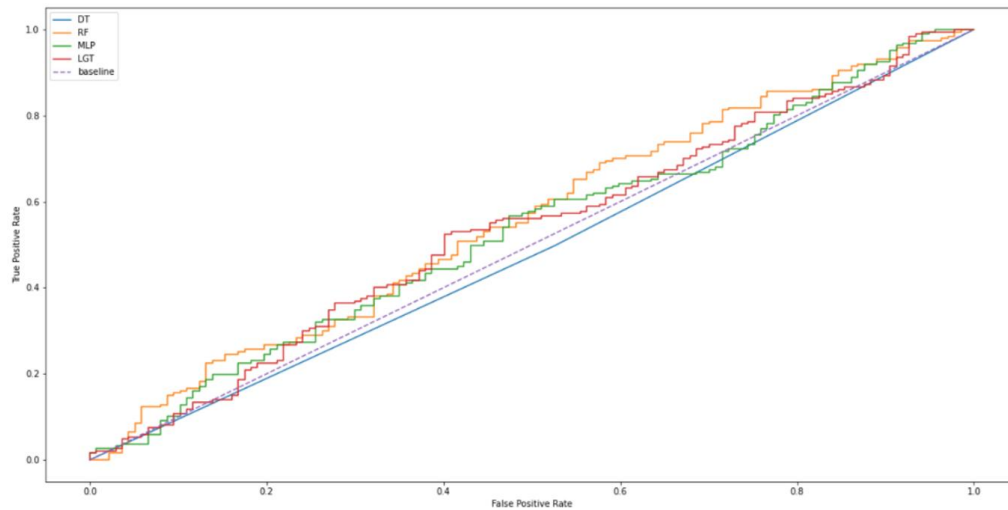
DTC: ROC AUC=0.486

RF: ROC AUC=0.557

MLP: ROC AUC=0.531

LGT: ROC AUC=0.534

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 35 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 4 ปี ของ DJI

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.58	0.5	0.58	0.73	0.486
RF	0.58	0.54	0.6	0.7	0.557
ANN	0.58	0.51	0.58	0.73	0.531
LGT	0.59	0.52	0.59	0.73	0.534

Regression

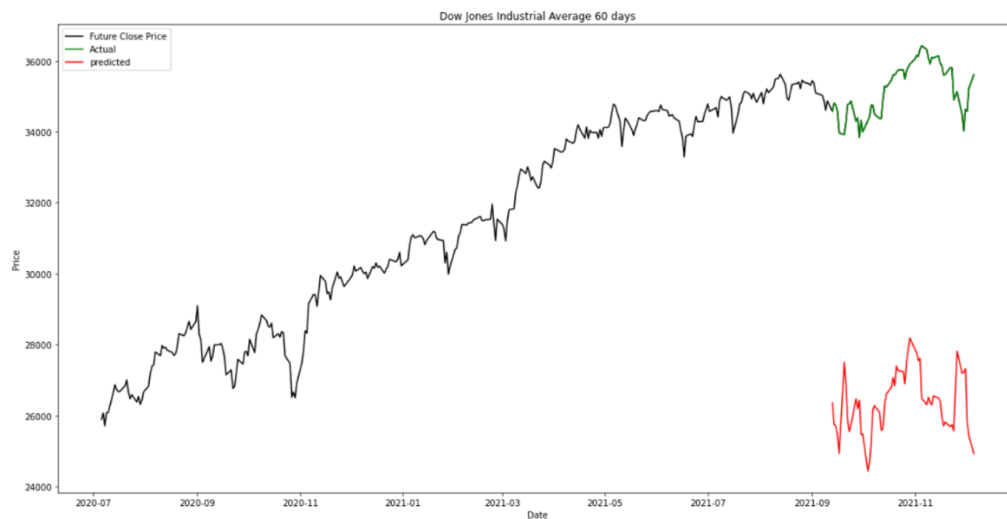
Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 4178.703881362111

R2 Score: 0.1320630127751622

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



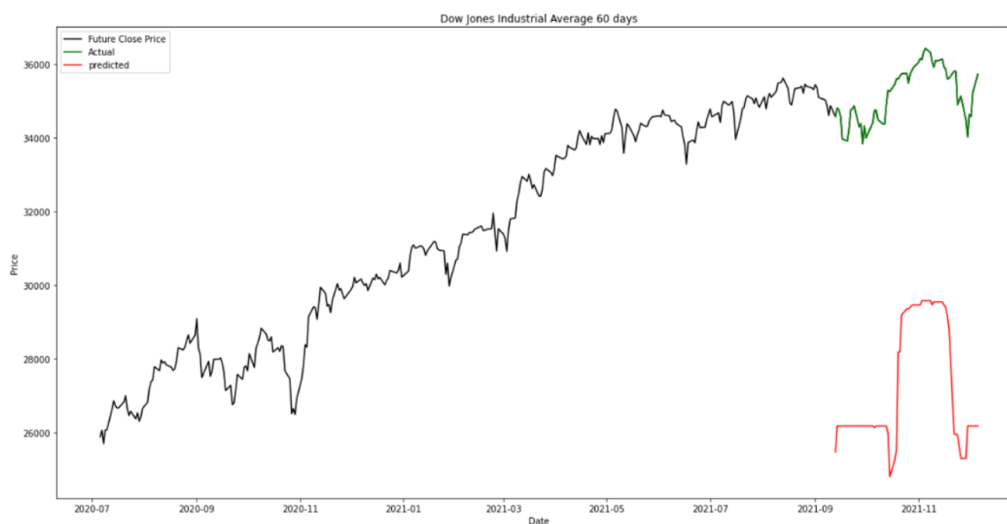
AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 3807.6187943672817

R2 Score: 0.2793706820130435

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการใช้ข้อมูลดัชนีหุ้น DJI ช่วงเวลา 8 ปี

DCT

ตารางที่ 36 ตาราง Confusion Matrix ของ DCT

0	185
0	265

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5888888888888889

Prediction recall Score: 0.5
Prediction F1_score: 0.7412587412587412
Prediction Precision: 0.588888888888889

RF

ตารางที่ 37 ตาราง Confusion Matrix ของ RF

51	134
80	185

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5244444444444445
Prediction recall Score: 0.4868944416114227
Prediction F1_score: 0.6335616438356165
Prediction Precision: 0.5799373040752351

ANN

ตารางที่ 38 ตาราง Confusion Matrix ของ ANN

0	185
0	265

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.588888888888889
Prediction recall Score: 0.5
Prediction F1_score: 0.7412587412587412
Prediction Precision: 0.588888888888889

LGT

ตารางที่ 39 ตาราง Confusion Matrix ของ LGT

0	185
0	265

โดยคำนวณค่าความแม่นยำได้เป็น

Prediction Accuracy Score: 0.5888888888888889

Prediction recall Score: 0.5

Prediction F1_score: 0.7412587412587412

Prediction Precision: 0.5888888888888889

ROC AUC

และสุดท้ายสรุปออกมาเป็นกราฟ ROC AUC จะได้มาเป็นค่า

NoSkill: ROC AUC=0.500

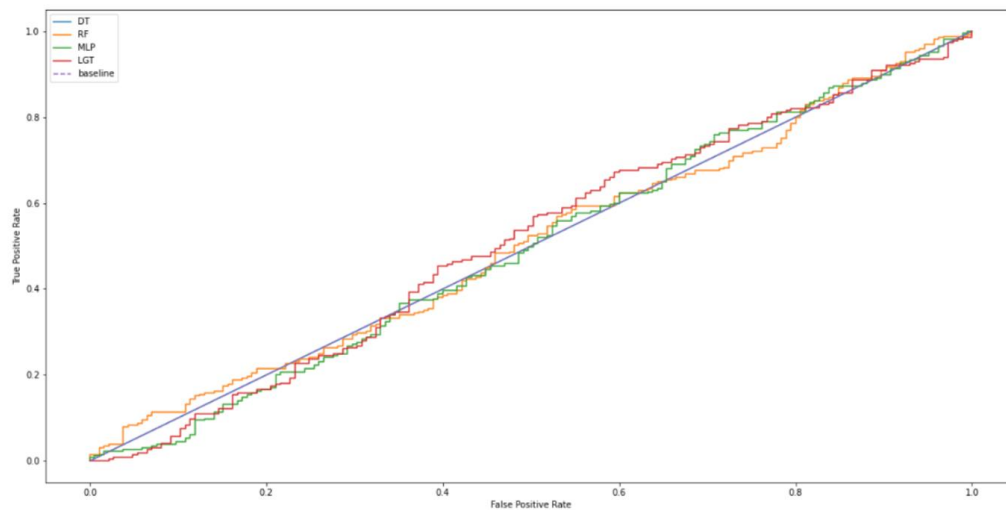
DTC: ROC AUC=0.500

RF: ROC AUC=0.503

MLP: ROC AUC=0.495

LGT: ROC AUC=0.509

และเมื่อทำการ Plot ออกมาเป็นกราฟจะได้ผลลัพธ์ดังนี้



และสุดท้ายคือตารางสรุปผลลัพธ์ความแม่นยำของ Classification

ตารางที่ 40 ผลลัพธ์ความแม่นยำของ Classification ในข้อมูล 8 ปี ของ DJI

	Accuracy	Recall	Precision	F-score	ROC CURVE
DCT	0.59	0.5	0.74	0.59	0.500
RF	0.52	0.49	0.58	0.63	0.503
ANN	0.59	0.5	0.59	0.74	0.495
LGT	0.59	0.5	0.59	0.74	0.509

Regression

Linear Regression

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 338.92974256836624

R2 Score: 0.9961215498682767

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



AdaBoost

โดยคำนวณค่าความแม่นยำได้เป็น

RMSE: 882.7512956563919

R2 Score: 0.9736903194112346

เมื่อทำการ Plot ตามช่วงเวลาจะได้เป็น



ผลลัพธ์จากการทดลองทำนายดัชนีหุ้นทั้ง 4 ตัว ของวันที่ 8 ธันวาคม 2021

HSI

DT ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

RF ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

ANN ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

LGT ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

โดยเมื่อวัดจากความแม่นยำจะนำค่าของ DT ออกและเหลือเพียง RF, ANN และ LGT

เมื่อทำการ Majority Vote จะได้ผลลัพธ์เป็น ลง

N225

DT ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

RF ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

ANN ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

LGT ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

โดยเมื่อวัดจากความแม่นยำจะนำค่าของ DT ออกและเหลือเพียง RF, ANN และ LGT

เมื่อทำการ Majority Vote จะได้ผลลัพธ์เป็น ลง

SSE

DT ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

RF ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

ANN ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

LGT ทำนายว่าดัชนีราคาหุ้นกำลังจะลง

โดยเมื่อวัดจากความแม่นยำจะนำค่าของ DT ออกและเหลือเพียง RF, ANN และ LGT

เมื่อทำการ Majority Vote จะได้ผลลัพธ์เป็น ลง

DJI

DT ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

RF ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

ANN ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

LGT ทำนายว่าดัชนีราคาหุ้นกำลังจะขึ้น

โดยเมื่อดูผลลัพธ์โดยรวมแล้วดัชนีราคาหุ้นทุกตัวมีผลลัพธ์เดียวกัน

จึงไม่ต้องทำการ Majority Vote และมีผลลัพธ์เป็น ขึ้น

4.3. การวิเคราะห์ผล

การศึกษาครั้งนี้มีวัตถุประสงค์หลัก 2 อย่างด้วยกัน คือ (1) เพื่อที่คาดการณ์ดัชนีขึ้นหรือลงของหุ้นทั้ง 4 ตัว ได้แก่ DJIA, N225, HSI และ SSE (2) เพื่อทำนายแนวโน้มและคาดการณ์ค่าราคาปิดของตลาดในอนาคต โดยใช้ข้อมูลราคาปิดในตลาดตั้งแต่วันที่ 1 มกราคม 2014 ถึง ธันวาคม 2021 และข้อมูลที่ใช้จะเป็นข้อมูลและดัชนีหุ้นจากเว็บไซต์ Yahoo! Finance โดยโมเดลที่ใช้ในการคาดการณ์คือ Classification และ Regression ในส่วนของ Classification จะใช้ข้อมูลราคาปิดของวันนั้นเป็น Target เทียบกับราคาปิดของวันก่อนหน้า หลังจากนั้นจะทำการคำนวณแบบ Majority Vote จาก Algorithm ต่าง ๆ ประกอบด้วย Decision tree, Random

Forest, MLP Artificial Neural Network (ANN) และ Logistic Regression โดยคัดเลือกเฉพาะ Algorithm ที่มีค่า Accuracy สูงที่สุด 3 ตัว ถัดไปในส่วนของ Regression จะใช้ราคาปิดของตลาดในแต่ละวันเป็น Target และใช้ AdaBoostRegressor Algorithm และ Linear regression Algorithm ในการทำนายราคาปิดของวันถัดไป โดยจากการศึกษาทางผู้จัดทำได้ทำการให้ระบบทำการเรียนจาก Training Dataset ทั้งหมด 75% และให้ระบบทำนายจาก Unseen Data ด้วย Test Dataset ทั้งหมด 25% โดยข้อมูลที่ใช้จะมาจากช่วงเวลาตั้งแต่เดือนมกราคม 2014 ถึง ธันวาคม 2021 เมื่อทำการทำนายหรือคาดการณ์ ผลสรุปเมื่อเทียบกับระหว่าง Classification และ Regression แล้วพบว่าการทำนายด้วยโมเดล Classification มีค่า Accuracy ที่แม่นยำน่าเชื่อถือมากกว่า โดยผลสรุปสุดท้ายทางคณะผู้จัดทำได้ทำการทดลองและทำนายผลลัพธ์ของดัชนีหุ้นของวันที่ 8 ธันวาคม 2021 ที่ได้จากการทดลองและทำนายทั้งหมดสรุปได้ว่า หุ้น N225, HSI และ SSE มีแนวโน้มที่จะลง และหุ้น DJIA มีแนวโน้มที่จะขึ้น โดยเมื่อเทียบดัชนีราคาปิดของหุ้นวันที่ 7 ธันวาคม 2021 เทียบกับ ราคาปิดของหุ้นวันที่ 8 ธันวาคม 2021 นั้น หุ้นทั้ง 4 ตัวมีราคาปิดที่สูงขึ้น ดังนั้นสรุปได้ว่าการทำนายของทางคณะผู้จัดทำถูกต้องเพียง 1 ใน 4 ของหุ้นทั้งหมด

5. สรุปผลการวิจัย

จากการได้ทดลองในครั้งนี้ผู้จัดทำได้เลือกใช้วิธีการในการทำนายดัชนีของหุ้นด้วยการใช้ข้อมูลดัชนีหุ้นย้อนหลังจาก เดือนมกราคม 2014 ถึง เดือนธันวาคม 2021 โดยมีหุ้น 4 ตัวที่ทางอาจารย์ที่ปรึกษาวิทยานิพนธ์แนะนำ ได้แก่ DJIA, N225, HIS, SSE และในส่วนถัดมาผู้จัดทำได้เลือกใช้ 2 โมเดลในการทำนาย คือ Classification และ Regression โดยโมเดลแรกทางผู้จัดทำได้ใช้ 4 อัลกอริทึม กับ Classification คือ (1) Decision tree (2) Random Forest (3) MLP Artificial Neural Network (ANN) และ (4) Logistic Regression โดยหลังจากที่ได้ผลลัพธ์การทำนายจะนำผลลัพธ์จากทั้ง 4 อัลกอริทึม มาคัดเลือกเพื่อหาผลลัพธ์ที่มีความแม่นยำมากที่สุดเพื่อทำการ Majority Vote หาว่าสุดท้ายแล้วดัชนีหุ้นวันถัดไปจะขึ้นหรือลง ถัดมาในส่วนของโมเดลที่ 2 ทางผู้จัดทำได้เลือกใช้ 2 อัลกอริทึม กับ Regression คือ (1) AdaBoostRegressor และ (2) Linear regression เพื่อทำนายหาว่าราคาปิดวันถัดไปจากตลาดหุ้น โดยทางผู้จัดทำได้ทำการทดลองและทำนายดัชนีของหุ้นในวันที่ 8 ธันวาคม 2021 ได้ผลลัพธ์เป็นทำนายถูกต้องเพียงหุ้นเดียวคือ DJIA และได้ผลสรุปว่าการทำนายดัชนีราคาหุ้นด้วยเพียงแค่ข้อมูลของราคาที่มีการขึ้นลงตลอดเวลานั้นมีความไม่แม่นยำไม่เพียงพอสำหรับการลงทุน เพราะการที่ดัชนีราคาหุ้นจะขึ้นหรือลงมีปัจจัยมาจากหลายอย่าง ไม่ว่าจะเป็นแนวโน้มของบริษัท การออกสินค้าในตลาด หรือข่าวลือต่าง ๆ ไม่ว่าจะเป็นข่าวลือที่ดี หรือข่าวลือที่ไม่ดี ดังนั้น ทางคณะผู้จัดทำขอสรุปไว้ว่าการทำนายการขึ้นลงของราคาดัชนีหุ้นนั้นส่งผลไม่ถูกต้องมากเท่าที่ควร แต่สามารถใช้เพื่อสำหรับการศึกษา หรือเป็นอีกหนึ่งในแหล่งอ้างอิงประกอบกับปัจจัยอื่น ๆ ควบคู่ เพื่อตัดสินใจในการลงทุนได้

ข้อเสนอแนะในการศึกษาครั้งถัดไป ทางคณะผู้จัดทำแนะนำให้ศึกษา AdaBoost เป็นพิเศษ เนื่องจาก AdaBoost เป็น Algorithm ที่น่าสนใจเหมาะสมกับข้อมูลที่มีความซับซ้อนเป็นพิเศษ นอกจากนี้ยังเป็น Meta-Algorithm ที่สามารถปรับใช้ได้กับทั้ง Classification และ Regression ทางคณะผู้จัดทำจึงมีความสนใจในการศึกษา AdaBoost เพิ่มเติม เพื่อให้ได้ทราบว่า การใช้ AdaBoost Classifier หรือ AdaBoost Regressor แบบไหนถึงจะมีความถูกต้องและเหมาะสมกับข้อมูลมากกว่า

6. เอกสารอ้างอิง

- [1] Sushant Ratnaparkhi and Milind Paradkar. 2017. **Use Decision Trees in Machine Learning to Predict Stock Movements.** [Online]. Available: <https://blog.quantinsti.com/use-decision-trees-machine-learning-predict-stock-movements/> [Accessed: 5 December 2021]
- [2] Saran Khotsathian. 2018. **Machine Learning: ANN คืออะไร.** [Online]. Available: <https://medium.com/@sarankhotsathian/machine-learning-ann-คืออะไร-3527a9aa0c8c> [Accessed: 5 December 2021]
- [3] Witchapong Daroontham. 2018. **เจาะลึก Random Forest !!!— Part 2 of “รู้จัก Decision Tree, Random Forest, และ XGBoost!!!”.** [Online]. Available: <https://medium.com/@witchapongdaroontham/เจาะลึก-random-forest-part-2-of-รู้จัก-decision-tree-random-forest-และ-xgboost-79b9f41a1c1c> [Accessed: 5 December 2021]

- [4] Yuth Kaiyawan. 2012. **หลักการและการใช้การวิเคราะห์การถดถอยโลจิสติกสำหรับการวิจัย Principle and Using Logistic Regression Analysis for Research**. [Online]. Available: <http://rdi.rmuts.ac.th/rmutsvrj/download/year4-issue1-2555/p1.pdf> [Accessed: 5 December 2021]
- [5] cway investment. 2018. **Adaptive Boosting Algorithm**. [Online]. Available: <https://medium.com/cw-quantlab/adaptive-boosting-algorithm-a761f0a0b264> [Accessed: 5 December 2021]
- [6] Jason Brownlee. 2016. **Linear Regression for Machine Learning**. [Online]. Available: <https://machinelearningmastery.com/linear-regression-for-machine-learning/> [Accessed: 5 December 2021]
- [7] Shipra Saxena. 2021. **Introduction to AdaBoost Algorithm with Python Implementation**. [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/03/introduction-to-adaboost-algorithm-with-python-implementation/> [Accessed: 7 December 2021]
- [8] Daddy trader. 2018. **MACD ถ้าอยากใช้ ต้องรู้อะไรบ้าง?**. [Online]. Available: <https://www.finnomena.com/daddy-trader/macd/> [Accessed: 7 December 2021]
- [9] Jason Fernando. 2021. **Relative Strength Index (RSI)**. [Online]. Available: <https://www.investopedia.com/terms/r/rsi.asp> [Accessed: 7 December 2021]
- [10] Amazon Polly. **Rate of Change (ROC)**. [Online]. Available: https://school.stockcharts.com/doku.php?id=technical_indicators:rate_of_change_roc_and_momentum [Accessed: 7 December 2021]
- [11] Anshul Saini. 2021. **AdaBoost Algorithm – A Complete Guide for Beginners**. [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/09/adaboost-algorithm-a-complete-guide-for-beginners/> [Accessed: 7 December 2021]
- [12] Daddy trader. 2018. **Moving Average ถ้าอยากใช้ ต้องรู้อะไรบ้าง?**. [Online]. Available: <https://www.finnomena.com/daddy-trader/moving-average/> [Accessed: 7 December 2021]
- [13] Online Trading. 2020. **รู้จัก EMA คือ อะไร ? และสอนใช้จับจังหวะซื้อขายด้วย EMA**. [Online]. Available: <https://knowledge.bualuang.co.th/knowledge-base/whatisema/> [Accessed: 7 December 2021]
- [14] Admin-forex. 2017. **Simple Moving Average**. [Online]. Available: <https://www.forex.co.th/simple-moving-average-sma-คืออะไร/> [Accessed: 7 December 2021]
- [15] Digital Advisor. 2021. **RSI คืออะไร? วิเคราะห์แนวโน้มหุ้นขาขึ้นหรือขาลงได้อย่างไร**. [Online]. Available: <https://knowledge.bualuang.co.th/knowledge-base/rsi/> [Accessed: 7 December 2021]
- [16] PPTVOnline. 2021. **รู้จัก RSI Indicator เครื่องมือวิเคราะห์สัญญาณการลงทุนสำหรับมืออาชีพ**. [Online]. Available: <https://www.pptvhd36.com/news/ประชาสัมพันธ์/145893> [Accessed: 7 December 2021]
- [17] ADMIRAL MARKETS. 2021. **ADX Indicator คืออะไร : การใช้ ADX ในการเทรด**. [Online]. Available: <https://admiralmarkets.sc/th/education/articles/forex-indicators/adx-forex-indicator> [Accessed: 7 December 2021]
- [18] ForextThai. **ATR คือ อะไร?**. [Online]. Available: <https://forexthai.in.th/atr-คืออะไร/> [Accessed: 7 December 2021]
- [19] Thai Broker Forex. **การใช้ Indicator ต่างๆ : William % R**. [Online]. Available: <https://thaibrokerforex.com/การใช้-indicator-ต่างๆ-william-r/> [Accessed: 7 December 2021]