บทความวิจัย

แบบจำลองการพยากรณ์แนวโน้มของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐแบบรายวันด้วยข้อมูลขั้วความรู้สึกของข่าวสารออนไลน์

พรพิมล ชัยวุฒิศักดิ์*
ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 09 8914 6526 อีเมล: pornpimol.ch@kmitl.ac.th
DOI: 10.14416/j.kmutnb.2023.10.005
รับเมื่อ 23 สิงหาคม 2564 แก้ไขเมื่อ 2 พฤศจิกายน 2564 ตอบรับเมื่อ 11 พฤศจิกายน 2564 เผยแพร่ออนไลน์ 25 ตุลาคม 2566
© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพยากรณ์แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐในตลาดซื้อขายแลกเปลี่ยนสกุลเงินระหว่างประเทศ และการทดสอบสมมติฐานเชิงเป็นเหตุเป็นผลระหว่างด้วยวิธีแกรนเจอร์ โดยข้อมูลที่นำมาใช้ในการศึกษาประกอบด้วยข้อมูลแบบมีโครงสร้าง ได้แก่ เส้นค่าเฉลี่ยเคลื่อนที่ ราคาทองคำและราคาน้ำมันดิบ และข้อมูลแบบไม่มีโครงสร้างได้แก่ สัดส่วนของจำนวนข่าวขั้วความรู้สึกเชิงบวก เชิงลบ และเป็นกลางของข่าวสารที่เกี่ยวข้องกับอัตราการแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ ซึ่งได้เก็บรวบรวมในระหว่างวันที่ 2 เดือนตุลาคม พ.ศ. 2563 ถึงวันที่ 10 เดือนกุมภาพันธ์ พ.ศ. 2564 จำนวน 92 รายการ โดยข้อมูลข่าวสารที่ได้จะถูกนำมาประมวลผลโดยวิธีเทคนิคการวิเคราะห์ขั้วความรู้สึกและวิธีการพยากรณ์แนวโน้มที่ใช้ในการศึกษานี้ได้แก่ 1) โครงข่ายประสาทเทียม 2) ซัพพอร์ตเวกเตอร์แมทชีน และ 3) การจำแนกประเภทแบบการสุ่มป่าไม้ นอกจากนี้ได้ทำการเปรียบเทียบการวัดประสิทธิภาพของตัวแบบพยากรณ์ คือ วิธีการแบ่งข้อมูลสำหรับสอนและวัดผลและวิธีการตรวจสอบแบบไขว้ ผลการศึกษาข่าวที่มีขั้วความรู้สึกเชิงบวกและข่าวที่มีขั้วความรู้สึกเป็นกลางมีความสัมพันธ์กับแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ และตัวแบบพยากรณ์แนวโน้มที่สร้างขึ้นด้วยวิธีการจำแนกประเภทแบบการสุ่มป่าไม้จากการแบ่งข้อมูลวิธีวิธีการแบ่งข้อมูลสำหรับสอนและวัดผลให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความถูกต้อง คิดเป็นร้อยละ 71.43 ทั้งนี้ตัวแบบการจำแนกประเภทแบบการสุ่มป่าไม้ สามารถนำมาใช้พยากรณ์อัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐเพื่อเป็นแนวทางการลงทุนการซื้อขายได้อย่างดี

คำสำคัญ: แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ การวิเคราะห์ขั้วความรู้สึก การสุ่มป่าไม้ โครงข่ายประสาทเทียม ซัพพอร์ตเวกเตอร์แมทชีน การทดสอบสมมติฐานเชิงเป็นเหตุเป็นผล

การอ้างอิงบทความ: พรพิมล ชัยวุฒิศักดิ์, "แบบจำลองการพยากรณ์แนวโน้มของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐแบบรายวันด้วยข้อมูลขั้วความรู้สึกของข่าวสารออนไลน์," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 33, ฉบับที่ 4, หน้า 1–14, เลขที่บทความ 234-175353, ต.ค.–ธ.ค. 2566.

Research Article

A Trend Forecasting Model of Daily Euro/USD Exchange Rate Using Sentiment Analysis on Online News

Pornpimol Chaiwuttisak*
Department of Statistics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

* Corresponding Author, Tel. 09 8914 6526, E–mail: pornpimol.ch@kmitl.ac.th
DOI: 10.14416/j.kmutnb.2023.10.005
Received 23 August 2021; Revised 2 November 2021; Accepted 11 November 2021; Published online: 25 October 2023
© 2023 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

The objectives of this research were to forecast the movement of the dollar-euro exchange rates in the Foreign Exchange (FOREX) Market and examine the relationship of the polarity of the news articles and the EUR to USD exchange rate using Granger Causality test. The variables used in the study consisted of structured variables such as a moving averages indicator, gold prices and crude oil prices, while the unstructured variables were the proportions of the positive, negative and neutral polarities of online news related to the EUR/USD exchange rates. Daily data was collected from October 2, 2020 to February 10, 2021, a total of 92 items in the study. The news articles were processed using the sentiment analysis and the following methods for predictive modeling: 1) Artificial Neural Network 2) Support Vector Machine and 3) Random Forest Classification. Two techniques for model evaluation are 1) Train-Test Split and 2) Cross Validation. The results showed that the positive, negative and neutral sentiments were related to the EUR to USD exchange rate trends. Moreover, The Random Forest based on Train-Test Split yielded the best results with accuracy accounting for 71.43%. Concisely, the Random Forest model can be used as an effective trading guide for forecasting the movement of the USD-EUR exchange rates.

Keywords: Trend of USD/Euro Exchange Rate, Random Forest, Sentiment Analysis, Artificial Neural Network, Support Vector Machine, Granger Causality Test

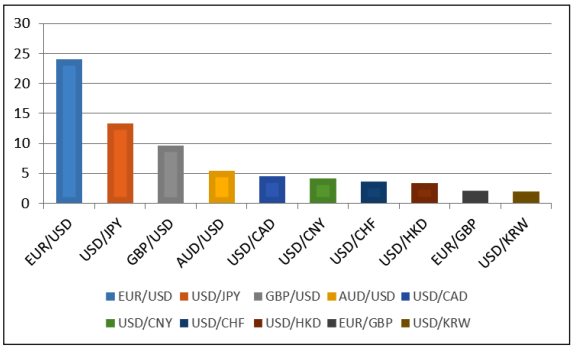
Please cite this article as: P. Chaiwuttisak, "A trend forecasting model of daily Euro/USD exchange rate using sentiment analysis on online news," *The Journal of KMUTNB*, vol. 33, no. 4, pp. 1–14, ID. 234-175353, Oct.–Dec. 2023 (in Thai).

1. บทนำ

ในปัจจุบันการแลกเปลี่ยนสกุลเงินระหว่างประเทศในตลาดซื้อขายอัตราแลกเปลี่ยน (Foreign Exchange Market; Forex) กำลังเป็นที่สนใจของนักลงทุนอย่างมาก เนื่องจากตลาดซื้อขายอัตราแลกเปลี่ยนนั้นมีสภาพคล่องทางการเงินสูง โดยมีเงินหมุนเวียนเฉลี่ยประมาณ 4 ล้านล้านเหรียญดอลลาร์สหรัฐต่อวัน มีการขึ้นลงของราคาอย่างต่อเนื่องและรวดเร็ว เมื่อเทียบกับตลาดหุ้นขนาดใหญ่ที่เป็นอันดับต้นๆ ของโลก ตัวอย่างเช่น ตลาดหุ้นนิวยอร์กมีเงินหมุนเวียนเพียงแค่ 22.4 พันล้านดอลลาร์สหรัฐต่อวันเท่านั้น [1] ส่งผลให้มีการลงทุนในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศเป็นจำนวนมาก

การลงทุนในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศเป็นการเก็งกำไรเกี่ยวกับความแข็งแกร่งของคู่สกุลหนึ่งเทียบกับอีกคู่สกุลเงินหนึ่ง ซึ่งนักลงทุนส่วนใหญ่มักจะเลือกลงทุนในคู่สกุลเงินที่ใหญ่และมีสภาพคล่องสูง เช่น EUR/USD, USD/JPY, GPY/USD, AUD/USD, USD/CHF และ USD/CAD และเมื่อสำรวจการจัดอันดับคู่เงิน พบว่า EUR/USD เป็นคู่เงินที่ได้รับความนิยมมากที่สุด โดย [2] ได้ทำการจัดอันดับ 10 คู่เงินที่มีการลงทุนมากที่สุด ดังรูปที่ 1 ได้แก่ 1) การลงทุนในคู่เงิน EUR/USD มากที่สุดเป็นอันดับ 1 คิดเป็นร้อยละ 34 2) USD/JPY คิดเป็นร้อยละ 18 3) GBP/USD คิดเป็นร้อยละ 13 4) AUD/USD คิดเป็นร้อยละ 7 5) USD/CAD คิดเป็นร้อยละ 6 6) USD/CNY คิดเป็นร้อยละ 6 7) USD/CHF คิดเป็นร้อยละ 5 8) USD/HKD คิดเป็นร้อยละ 5 9) EUR/GBP คิดเป็นร้อยละ 3 และ (10) USD/KRW คิดเป็นร้อยละ3 ตามลำดับ

การขึ้นลงของคู่สกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) ในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศนั้นมีผลมาจากอิทธิพลหลายอย่าง มีการศึกษามากมายที่กล่าวถึงอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) โดยที่ [3] ได้ศึกษาและพบว่าราคาน้ำมันดิบในตลาดโลกนั้นมีผลต่ออัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐเป็นอย่างมาก [4] ได้กล่าวว่าราคาทองคำและค่าเงินดอลลาร์มีความสัมพันธ์แบบ



รูปที่ 1 อันดับคู่เงินที่ได้รับการลงทุนมากที่สุด 10 อันดับแรก [2]

ผกผันกล่าว คือ การเพิ่มขึ้นของราคาทองคำจะส่งผลให้ค่าเงินดอลลาร์ลดลง จากงานวิจัยทั้งสองสามารถสรุปได้ว่าทั้งราคาน้ำมัน และราคาทองคำมีผลต่อคู่สกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ

อิทธิพลของข่าวสารออนไลน์เป็นอีกหนึ่งปัจจัยที่มีผลต่อการขึ้นลงของอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศนอกเหนือจากปัจจัยที่กล่าวมาในข้างต้น ทั้งนี้ข่าวบางอย่างอาจส่งผลกระทบต่ออัตราแลกเปลี่ยนในทางบวกหรือทางลบได้ ตัวอย่างเช่นข่าวการแพร่ระบาดของเชื้อไวรัสโควิด 19 ส่งผลกระทบต่อการขึ้นลงของคู่สกุลเงินอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) ในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศอย่างเห็นได้ชัดโดยราคามีแนวโน้มที่ลดลง โดย [5] รายงานว่า นักลงทุนยังคงเทขายสกุลเงินยูโรอย่างต่อเนื่องจนทำให้เงินยูโรร่วงลงไปสู่ระดับราคาต่ำที่สุดนับตั้งแต่ พ.ศ. 2563 โดย EUR/USD ร่วงลงไปสู่ระดับ 1.1630 [6] หรือข่าวการลดอัตราดอกเบี้ยของธนาคารสหรัฐอเมริกา ดังนั้นความน่าเชื่อถือของข่าวจากสื่อออนไลน์จึงมีความสำคัญอย่างมากในการตัดสินใจลงทุนคู่สกุลเงินในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศ

ดังนั้นงานวิจัยนี้ศึกษาการขึ้นลงของคู่สกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ ในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศพิจารณาอิทธิจากข่าวของสื่อออนไลน์ร่วมกับปัจจัยอื่นๆ ได้แก่ อัตราแลกเปลี่ยนในตลาด และเครื่องมือ

ทางเทคนิคที่ใช้ในการวิเคราะห์แนวโน้ม เพื่อนำไปสร้างตัวแบบพยากรณ์ใช้ทำนายการขึ้นลงของคู่สกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ ในตลาดซื้อขายอัตราแลกเปลี่ยนสกุลเงินระหว่างประเทศโดยอาศัยวิธีการเรียนรู้ของเครื่องมือ (Machine Learning)

1.1. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1.1.1 เทคนิคการจำแนกกลุ่มข้อมูล

งานวิจัยของ Taboada [7] ได้อธิบายการจำแนกข้อมูล (Classification) เป็นหนึ่งในเทคนิคการทำเหมืองข้อมูล ซึ่งจะเป็นการสร้างโมเดลหรือตัวจำแนกข้อมูล (Classifier) เพื่อทำนายหมวดหมู่ของข้อมูล (Categories/Class) เทคนิคการจำแนกข้อมูล (Classification) นั้นเป็นกระบวนการทางสถิติเพื่อจัดประเภทและวิเคราะห์หารูปแบบของกลุ่มข้อมูลใหม่ โดยทำการแบ่งข้อมูลออกเป็น 2 กลุ่มหลักๆ คือ กลุ่มข้อมูลสอน (Training Dataset) เป็นชุดข้อมูลที่มีบทบาทในการสร้างโมเดลจำแนกประเภทข้อมูลขึ้นมา และกลุ่มข้อมูลทดสอบ (Test Dataset) เป็นชุดข้อมูลประเมินความถูกต้องของโมเดลโดยกระบวนการสร้างตัวโมเดลจำแนกประเภทข้อมูล การสร้างโมเดลการจำแนกประเภท การประเมินผลโมเดลเป็นขั้นตอนตรวจสอบความถูกต้องโดยใช้ข้อมูลทดสอบ และปรับปรุงโมเดลจนกว่าจะได้ค่าความถูกต้อง

1.1.1.1 โครงข่ายประสาทเทียม

งานวิจัยของพรเทพ [8] ได้อธิบายโครงข่ายประสาทเทียม (Artificial Neuron Network; ANN) เป็นศาสตร์แขนงหนึ่งของทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) มีรูปแบบโครงสร้างและการประมวลผลเหมือนกับสมองของสิ่งมีชีวิตซึ่งมีปรับเปลี่ยนตัวเองต่อการตอบสนองของอินพุตตามกฎของการเรียนรู้ (Learning Rule) หลังจากที่โครงข่ายได้เรียนรู้สิ่งที่ต้องการแล้วโครงข่ายนั้นจะสามารถทำงานที่กำหนดไว้ได้เช่น มีความสามารถในการเรียนรู้การจดจำรูปแบบ(Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) เป็นต้น

โครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบ [9] ได้แก่

1) ข้อมูลป้อนเข้า (Input) เป็นข้อมูลเชิงปริมาณ หากเป็นข้อมูลเชิงคุณภาพต้องทำการแปลงให้อยู่ในรูปของข้อมูลเชิงปริมาณ ข้อมูลที่นำเข้ามาประมวลผลตามรูปแบบของปัญหาที่ต้องการแก้ไขหรือพิจารณา ซึ่งข้อมูลนำเข้าจะมีความสอดคล้องและสัมพันธ์กับปัญหา

2) ข้อมูลส่งออก (Output) คือผลลัพธ์ที่เกิดขึ้นจริงจากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม

3) ค่าน้ำหนัก (Weights) เป็นค่าที่ได้จากการเรียนรู้จากโครงข่ายประสาทเทียมหรือเรียกว่า "ค่าความรู้" ซึ่งถูกกำหนดให้กับข้อมูลนำเข้าแต่ละตัว ค่าน้ำหนักที่แตกต่างกันแสดงถึงระดับความสำคัญของข้อมูลนำเข้าซึ่งถือเป็นองค์ประกอบที่สำคัญมากในโครงข่ายประสาท

4) ฟังก์ชันผลรวม (Summation Function; S) เป็นฟังก์ชันที่ทำหน้าที่รวมค่าน้ำหนักที่ได้จากโครงข่ายของข้อมูลนำเข้า แทนด้วยสัญลักษณ์ a_i และค่าน้ำหนักในแต่ละชั้น แทนด้วยสัญลักษณ์ a_i เพื่อใช้สำหรับสรุปผลความสัมพันธ์ระหว่างข้อมูลนำเข้าทั้งหมด ซึ่งฟังก์ชันผลรวมพิจารณาจากสมการที่ 1

$$S = \left(\sum_{i=1}^{n} a_i w_i \right) + b \tag{1}$$

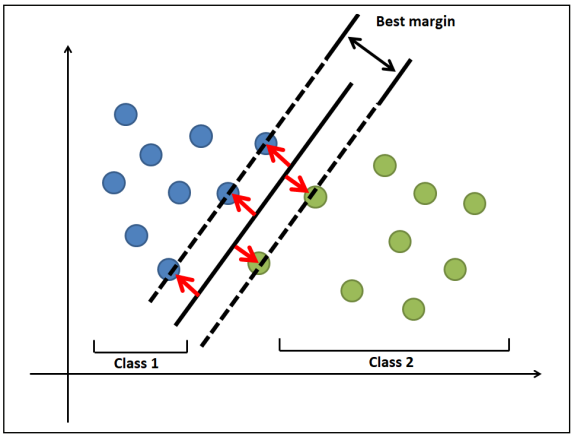
โดยที่ n คือ จำนวนข้อมูลนำเข้าในชั้นรับข้อมูล (Input Layer)

b คือ ความเอนเอียงที่เพิ่มเข้าไปในฟังก์ชันผลรวม

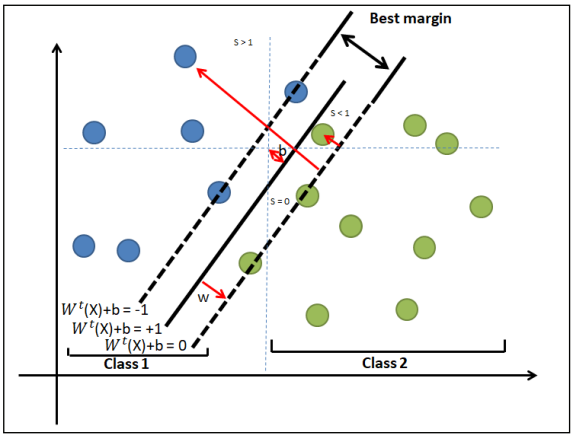
5) ฟังก์ชันการแปลง (Transfer Function) เป็นการคำนวณการจำลองการทำงานของโครงข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน (Sigmoid Function), ฟังก์ชันไฮเปอร์โบลิกแทนเจนต์ (Hyperbolic Tangent Function)

1.1.1.2 ซัพพอร์ตเวกเตอร์แมทชีน

ซัพพอร์ตเวกเตอร์แมทชีน (Support Vector Machines) เป็นหนึ่งในวิธีการเรียนรู้แบบมีผู้สอน ใช้เพื่อการแบ่งประเภทข้อมูลและการวิเคราะห์การถอดถอย เมื่อมีข้อมูลฝึกมาให้และแต่ละข้อมูลถูกจัดอยู่ในประเภทใดประเภทหนึ่งจากสองประเภท ซัพพอร์ตเวกเตอร์แมทชีนจะสร้างแบบจำลองที่สามารถพยากรณ์ได้ว่าตัวอย่างใหม่นี้จะตกอยู่ในกลุ่มใด ซึ่งเป็นอัลกอรึทึมที่ได้ถูกพัฒนามาจากเพอร์เซปตรอน



รูปที่ 2 การแบ่งเส้นกึ่งกลางระหว่างกลุ่มทั้งสอง



รูปที่ 3 การแบ่งโดยใช้ตัวแปรอนุโลม (Slack Variable)

และมีการเพิ่มลักษณะการยืดหยุ่นด้วยการพยายามปรับเส้นแบ่งเพื่อที่จะให้เกิดระยะขอบ (Margin) มากๆ อัลกอริทึมซัพพอร์ตเวกเตอร์แมทชีน ถือเป็นอัลกอริทึมที่มีประสิทธิภาพมากในการจำแนกข้อมูล แนวคิด คือ การนำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) ต่อจากนั้นนำจุดขอบ (Convex Hull) มาลากเส้นขอบเชื่อมต่อกัน หลังจากลากเส้นเชื่อมขอบแล้ว ก็จะทำการสร้างเส้นตรงที่ขนานกันระหว่างจุดขอบทั้งสองกลุ่ม ต่อไปสร้างเส้นแบ่งตรงกึ่งกลางระหว่างสองกลุ่มโดยพยายามให้ระยะขอบมากที่สุด เพื่อที่จะสามารถแบ่งแยกข้อมูลทั้งสองได้ ดังรูปที่ 2 และในบางกรณีเพื่อให้เกิดระยะขอบมากๆ จะใช้ตัวแปรอนุโลม (Slack Variable) แต่อาจเกิดการทำนายผิดพลาดไปบ้างเพื่อให้ระยะการแบ่งที่มากที่สุด วิธีนี้ถือว่าเป็นวิธีที่ยืดหยุ่น [10], [11] ดังรูปที่ 3

วิธีการของซัพพอร์ตเวกเตอร์แมทชีน จะถูกอธิบายด้วยวิธีการคำนวณ โดยกำหนดให้ $(x_i, y_i),...,(x_n, y_n)$ เป็นข้อมูลตัวอย่างที่ใช้สำหรับการสอน กำหนดให้ n คือ จำนวนของข้อมูลตัวอย่าง m คือ จำนวนมิติของข้อมูลเข้าและ y คือ ผลลัพธ์ที่แทนประเภทหรือแทนกลุ่มข้อมูลที่มีค่า +1 หรือ –1 โดยที่ +1 แทนข้อมูลกลุ่มที่ 1) และ –1 แทนข้อมูลกลุ่มที่ 2) อธิบายดังสมการที่ (2)

$$(x_i, y_i),...,(x_n, y_n) \text{ เมื่อ } x \in R^m, y \in \{+1, -1\} \quad (2)$$

สำหรับปัญหาเชิงเส้น 2 มิติที่เป็นพื้นฐานของการจำแนกข้อมูล สามารถคำนวณได้ดังสมการที่ (3)

$$(W * x) + b \quad (3)$$

เมื่อ W คือ น้ำหนัก และคือค่าไบแอส สมการใช้สำหรับจำแนกประเภทของข้อมูล ด้วยวิธีการซัพพอร์ตเวกเตอร์แมทชีน ดังสมการที่ (4)

$$(W * x) + b > 0 \text{ ถ้า } y_i = +1 \text{ และ}$$
$$(W * x) + b < 0 \text{ ถ้า } y_i = -1 \quad (4)$$

นอกจากฟังก์ชันเชิงเส้นแล้วในข้อมูลที่มีมิติสูง และข้อมูลไม่สามารถแบ่งแยกได้โดยง่าย ซัพพอร์ตเวกเตอร์แมทชีน ยังมีฟังก์ชัน (Kernel Function) ที่สามารถนำมาประยุกต์ใช้ในการแก้ไขปัญหาได้หลายวิธี เช่นฟังก์ชิกมอยด์ (Sigmoid Function), ฟังก์ชันเรเดียลเบสิส (Radial Basis Function) และฟังก์ชันพหุนาม (Polynomial Function) เป็นต้น

1.1.1.3 การจำแนกประเภทแบบการสุ่มป่าไม้

เทคนิคการจำแนกประเภทแบบการสุ่มป่าไม้ (Random Forest) ถูกนำเสนอครั้งแรกใน พ.ศ. 2538 โดย Tin Kam ซึ่งต่อมาเทคนิคการจำแนกประเภทแบบการสุ่มป่าไม้ ถูกต่อยอด

โดย Leo Breiman เป็นเทคนิคการสุ่มเลือกใช้ข้อมูลและคุณลักษณะ ต้นไม้ตัดสินใจ (Decision Tree) ซึ่งถูกสร้างจากการนำข้อมูลไปสุ่มเลือกตัวอย่างแบบใส่คืน (Sampling With Replacement) แล้วนำมาสร้างเป็นต้นไม้ (Tree) ซึ่งจะมีตัวอย่างส่วนหนึ่งไม่ถูกเลือก ข้อมูลส่วนนี้เรียกว่า Out-of-Bag (OOB) แล้วจะถูกนำมาใช้ในการทดสอบ ต้นไม้ตัดสินใจ (Decision Tree) วิธีการดังกล่าวนี้เรียกว่า Bagging ผลลัพธ์ที่ได้อย่างอิสระจากต้นไม้ตัดสินใจ (Decision Tree) แต่ละต้นถูกนำมาคิดเป็นผลการโหวตที่มากที่สุด เทคนิคการจำแนกประเภทแบบการสุ่มป่าไม้ ไม่จำเป็นต้องมีข้อมูลทดสอบเพื่อประมาณความผิดพลาดเพราะ OOB นั้นได้ถูกนำมาใช้ทดสอบต้นไม้ตัดสินใจแล้ว กระบวนของการจำแนกประเภทแบบการสุ่มป่าไม้ [12]

1.1.2 การทดสอบสมมติฐานเชิงเป็นเหตุเป็นผล (Granger Causality Test)

เป็นแนวคิดและวิธีทดสอบ โดยมีตัวแปร 2 ตัวแปร คือ X และ Y ในลักษณะที่เป็นข้อมูลอนุกรมเวลา นั้นคือถ้าการเปลี่ยนแปลงของ X เป็นต้นเหตุของการเปลี่ยนแปลงของ Y แล้ว X ก็ควรที่จะเกิดขึ้นก่อน Y

จากงานวิจัยของ Pindyck [13] ได้สรุปว่า ถ้า X เป็นเหตุผลการเปลี่ยนแปลงใน Y เงื่อนไขสองประการจะต้องเกิดขึ้นคือ เงื่อนไขประการแรกคือ X ควรจะช่วยในการทำนาย Y นั่นก็คือในการถดถอยของ Y กับค่าที่ผ่านมาของ Y นั้น ค่าที่ผ่านมาของ X ซึ่งทำหน้าที่เป็นตัวแปรอิสระควรที่จะมีส่วนช่วยในการเพิ่มอำนาจในการอธิบายของสมการถดถอยอย่างมีนัยสำคัญ เงื่อนไขประการที่สอง Y ไม่ควรช่วยในการทำนาย X เหตุผลก็คือถ้า X ช่วยทำนาย Y และ Y ช่วยทำนาย X ก็น่าจะมีตัวแปรอื่นอีกหนึ่งตัวหรือมากกว่าที่เป็นสาเหตุให้เกิดการเปลี่ยนแปลงทั้งใน X และ Y

การทดสอบสมมติฐานเชิงเป็นเหตุเป็นผลมีสมติฐานหลัก (H_0) คือ X ไม่ได้เป็นตัวแปรเป็นเหตุเป็นผลของ Y (X does not Granger Cause Y) สมการที่ (5) เป็นสมการถดถอยที่ไม่ได้ระบุข้อจำกัด (Unrestricted Regression)

$$Y_t = \sum_{i=1}^{p} \theta_i Y_{t-1} + \sum_{i=1}^{p} \gamma_i X_{t-1} + \varepsilon_t \quad (5)$$

สมมติฐานหลัก (H_0) และสมมติฐานรอง (H_a) ในเชิงสถิติสามารถเขียนได้ดังนี้

$H_0 : \gamma_1 = \gamma_2 = ... = \gamma_p = 0$

$H_a : H_0$ ไม่เป็นจริง

โดยที่ค่าสถิติของการทดสอบจะเป็นการใช้ค่าสถิติ F-statistic โดยสมการในการคำนวณดังสมการที่ (6)

$$F_{q,(n-k)} = \frac{(RSS_r - RSS_{ur})/q}{RSS_{ur}/(n-k)} \quad (6)$$

ถ้าค่า F-statistic ที่คำนวณได้สูงกว่าค่าวิกฤติ [Prob.< a] แสดงว่าสมติฐานหลัก (H_0) ถูกปฏิเสธหมายความว่า X เป็นเหตุเป็นผลการเปลี่ยนแปลงของ Y และในทำนองเดียวกันถ้าต้องการทดสอบสมติฐานหลัก (H_0) ว่า Y ไม่ได้เป็นเหตุเป็นผลของ X (Y does not Granger Cause X) ก็สามารถทำได้ด้วยกระบวนการทดสอบอย่างเดียวกับข้างต้นเพียงแต่สลับเปลี่ยนแปลงจำลองข้างต้นจาก X มาเป็น Y และจาก Y มาเป็น X ดังสมการที่ (7) และ (8)

$$X_t = \sum_{i=1}^{p} \theta_i X_{t-1} + \sum_{i=1}^{p} \gamma_i Y_{t-1} + u_t \quad (7)$$

$$X_t = \sum_{i=1}^{p} \theta_i X_{t-1} + u_t \quad (8)$$

จากสมการข้างต้นที่กล่าวมาจะเกตุได้ว่า ρ = ค่าคาบเวลาล้าหลัง (Lag Value) เป็นตัวแปรอิสระ ดังนั้นเพื่อที่จะได้แน่ใจว่าผลลัพธ์ที่ได้มานั้นไม่อ่อนไหวไปกับค่าของ p โดยทั่วไปแล้วจะทำการทดสอบ ณ ค่าของ ρ ที่แตกต่างกัน 2–3 ค่า

1.1.3 การวิเคราะห์ขั้วความรู้สึก

การวิเคราะห์ขั้วความรู้สึก (Sentiment Analysis) เป็นการวิเคราะห์ขั้วความรู้สึกผ่านตัวอักษร และเป็นอีกกระบวนการหนึ่งที่เกี่ยวกับการทำการประมวลผลภาษาธรรมชาติ (Natural Language Processing) ซึ่งจะทำการวิเคราะห์ข้อความเชิงคำนวณเพื่อระบุความรู้สึก โดยมีวัตถุประสงค์เพื่อวิเคราะห์ความคิดเห็นของผู้บริโภค หรือวิเคราะห์หัวข้อทัศนคติต่างๆ ที่สื่อออกมาผ่านทางข้อความเพื่อนำไปวิเคราะห์แก้ไข หรือใช้เป็นแนวทางในการดำเนินการ

ต่อไปในอนาคต

ในงานวิจัยของ Angiani [14] กล่าวว่าขั้นพื้นฐานของการวิเคราะห์ขั้วความรู้สึก โดยการทำนายขั้วความรู้สึกของประโยคที่กำหนดเพื่อดูว่าประโยคนั้นแสดงความรู้สึกออกมาในเชิงบวก เชิงลบหรือเป็นกลาง โดยจะมีขั้นตอนคือ ขั้นแรก รวบรวมข้อมูล ขั้นที่สองจัดเตรียมข้อมูล ก่อนที่จะนำข้อมูลนั้นไปใช้ ได้แก่ การลบเครื่องหมายวรรคตอน (Clean Text) การตัดคำ (Tokenize) การลดคำ (Stemming) และคำหยุด (Stop Word)

เมื่อทำการจัดเตรียมข้อมูลเป็นที่เรียบร้อยแล้วขั้นต่อไปจะเข้าสู่กระบวนการวิเคราะห์ขั้วความรู้สึก โดยการวิเคราะห์ขั้วความรู้สึกมีหลายวิธีที่ใช้ในการประเมินคุณลักษณะความสำคัญของข้อความโดยการให้น้ำหนักค่าของคำที่พบเป็นส่วนมาก และวิธีที่ได้รับความนิยมมากที่สุด ได้แก่ คุณลักษณะความถี่ (Feature Frequency; FF) การนับความถี่ของคำ (Term Frequency Inverse Document Frequency; TF-IDF) และการปรากฏของคำคุณลักษณะ (Feature Presence; FP) ด้วยคุณลักษณะความถี่ (FF) คือจำนวนครั้งที่เกิดขึ้นในการนับความถี่ของคำ (TF-IDF) [15] โดยมีสูตรในการคำนวณดังสมการที่ (9)

$$TF-IDF=(FF)\left(Log\left(\frac{N}{DF}\right)\right) \tag{9}$$

โดยที่ N คือ จำนวนเอกสาร

DF คือ จำนวนเอกสารที่มีคำคุณลักษณะ

FF คือ ความถี่ของคำในเอกสาร

Valence Aware Dictionary For Sentiment Reasoning; VADER เป็นแพ็คเกจที่อยู่ใน Natural Language Toolkit; NLTK เป็นชุดโปรแกรมสำหรับ การประมวลผลภาษาธรรมชาติ (NLP) สำหรับภาษาอังกฤษโดยเขียนในภาษาไพทอน (Python) ซึ่ง VADER จะถูกใช้ร่วมกับ Sentiment Lexicon ซึ่งเป็นพจนานุกรมที่มีรายการคุณลักษณะของคำศัพท์ประมาณ 7,500 รายการ ส่วนคำใดที่ไม่ได้อยู่ในพจนานุกรมจะให้ค่าเป็นศูนย์ ซึ่งในกฎของไวยากรณ์ทางภาษานอกจากคำศัพท์ที่เป็นขั้วความรู้สึกแล้ว ยังมีโครงสร้างหรือศัพท์บางคำที่มีความเป็นกลางแต่แท้จริงแล้วคำพวกนั้นสามารถเปลี่ยนขั้วความรู้สึกของคำได้ เช่น “ไม่”, “แต่” หรือปรับเปลี่ยนประโยคให้มีความเข้มขึ้นเช่น “มาก”, “อย่างยิ่ง” ใน VADER นักพัฒนาซอฟต์แวร์ได้รวมกฎ Heuristic ไว้เพื่อจัดการกับกรณีเครื่องหมาย วรรคตอน การใช้อักษรพิมพ์ใหญ่ กริยาวิเศษณ์ และคำสันธานและ VADER สามารถบอกได้ว่าเป็นขั้วความรู้สึกเชิงบวกหรือขั้วความรู้สึกเชิงลบ โดยคำนวณค่า Compound ที่จะมีช่วงอยู่ระหว่าง [0,1] มีสูตรในการคำนวณค่า Compound ดังสมการที่ (10)

$$x=\frac{x}{\sqrt{x^2+\alpha}} \tag{10}$$

โดยที่ x คือ ผลรวมของคะแนนของค่าที่อยู่ภายในข้อความ

a คือ ค่าแอลฟ่าเท่ากับ 15 (ค่าคงที่)

โดย Compound เป็นค่าในการกำหนดเกณฑ์ว่าเนื้อหาข่าวนั้นจะมีขั้วความรู้สึกทางใด โดยที่ค่า Compound สามารถแบ่งแยกได้ดังนี้

- ค่า Compound มากกว่า 0.5 เป็นขั้วความรู้สึกเชิงบวก (1)
- ค่า Compound น้อยว่า –0.5 เป็นขั้วความรู้สึกเชิงลบ (–1)
- ค่า Compound อยู่ระหว่าง –0.5 ถึง 0.5 เป็นขั้วความรู้สึกเป็นกลาง (0)

ตัวอย่างข้อความข่าว

ข่าวที่ 1 “EUR/USD Price Analysis: Daily chart shows failed breakout”

แปลว่า การวิเคราะห์ราคา EUR/USD: กราฟรายวันแสดงให้เห็นว่าเส้นราคาไม่ทะลุแนวกำแพง (หมายความว่าราคาอาจล่วงลงมาจากจุดปัจจุบัน) แสดงว่าข้อความนี้มีผลขั้วความรู้สึกเชิงลบต่อหุ้นของอัตราแลกเปลี่ยนสกุลเงิน EUR/USD เมื่อเทียบกับค่า Compound = –0.51 ในเครื่องมือ VADER จะได้ว่ามีความสอดคล้องกัน

ข่าวที่ 2 “EUR/USD: Sidelined near 1.11760 as US Vice Presidential debate kicks off”

แปลว่า แนวต้านใกล้ 1.11760 ขณะที่การหาลือของ

ประธานาธิบดีสหรัฐอเมริกาเริ่มขึ้น (แม้ราคาจะเข้าใกล้จุดที่ทำการซื้อขายแต่การหารือของประธานาธิบดีกำลังเริ่มขึ้น จึงไม่แนะนำให้ทำการซื้อขาย) แสดงว่าข้อความนี้มีผลขั้วความรู้สึกเป็นกลางต่อหุ้นของอัตราแลกเปลี่ยนสกุลเงิน EUR/USD เมื่อเทียบกับค่า Compound = 0.00 ในเครื่องมือ VADER จะได้ว่ามีความสอดคล้องกัน

ข่าวที่ 3 “The EUR/USD move higher comes amidst the improved tone in the risk complex, with EU stock indices hitting new record highs and US yields receding further from daily tops”

แปลว่า การขยับขึ้นของค่าเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) โดยดัชนีหุ้นของสหภาพยุโรปแตะระดับสูงสุดเป็นประวัติศาสตร์และผลตอบแทนสหรัฐลดลงจากระดับสูงสุดรายวัน (ดัชนีหุ้นสหภาพยุโรปสูงขึ้นซึ่งเป็นแนวโน้มที่ดีต่อคู่เงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) แสดงว่าข้อความนี้มีผลขั้วความรู้สึกเชิงบวกต่อหุ้นของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) เมื่อเทียบกับค่า Compound = 0.65 ในเครื่องมือ VADER จะได้ว่ามีความสอดคล้องกัน

1.1.4 งานวิจัยที่เกี่ยวข้อง

จากงานวิจัยของ Nassirtoussi [16] ได้ศึกษาทิศทางของอัตราแลกเปลี่ยนระหว่างประเทศในช่วงระยะเวลาสั้นของคู่สกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) และคาดการณ์ความสัมพันธ์ของข่าวสารกับอัตราแลกเปลี่ยนระหว่างประเทศคู่สกุลเงิน (EUR/USD) มีความสัมพันธ์กัน โดยได้ศึกษาหัวข้อข่าวที่ถูกนำเสนอออกมาราย 2 ชั่วโมง ส่วนข้อมูลของอัตราแลกเปลี่ยนคู่สกุลเงิน (EUR/USD) เก็บข้อมูลทุก 2 ชั่วโมง ทั้งหมด 8 รายการ คือ Open, High, Low, Close ของทั้ง Bid Price (ราคาขาย) และ Ask Price (ราคาซื้อ) โดยมีกระบวนการ 3 ขั้นตอนๆ ดังนี้ 1) Pre-Processing 2) Machine Leaning และ 3) Evaluation ซึ่งในขั้นที่ 1 Pre-Processing จะทำการขุดข้อความและ Multi-Layer ทั้ง 3 Layer ได้แก่ Semantic Abstraction Layer, Sentiment Integration Layer และ Synchronous Targeted Feature; STFR และขั้นที่ 2 Machine Leaning ได้แก่ Support Vector Machine; SVM, K-Nearest Neighbors; K-NN และ Naïve Bayes ผลการศึกษา พบว่าความสัมพันธ์ของข่าวสารเชิงบวกและข่าวสารเชิงลบกับอัตราแลกเปลี่ยนระหว่างประเทศคู่สกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) และความถูกต้องของเทคนิค SVM ให้ผลที่ดีที่สุดที่ร้อยละ 83.33

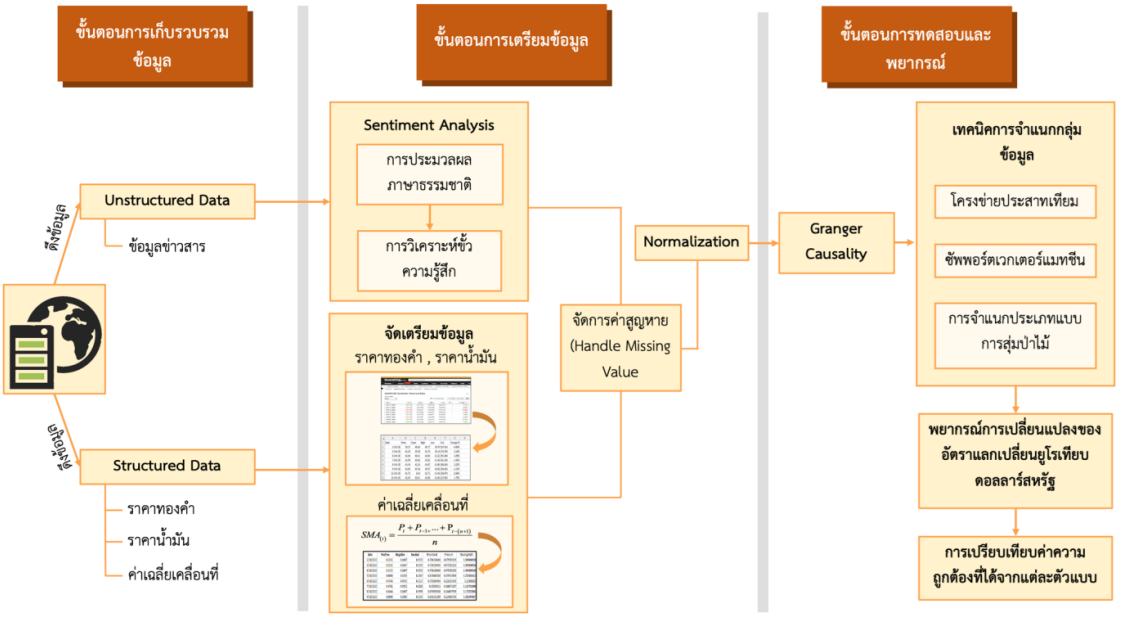
จากงานวิจัยของ Islam [17] ได้ศึกษาแนวทางการทำนายสกุลเงิน FOREX โดยนักวิจัยส่วนใหญ่ให้ความสนใจในโมเดลโครงข่ายประสาทเทียมเป็นแนวทางเทคนิคการเพิ่มประสิทธิภาพ อย่างไรก็ตามเป็นที่น่าแปลกใจที่การประมวลผลภาษาธรรมชาติไม่ได้ถูกนำไปใช้อย่างมากมายเมื่อเทียบกับแนวทางอื่นๆ

จากงานวิจัยของวิกานดา [18] ได้ทำการพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทยโดยใช้วิธีการประมวลผลภาษาธรรมชาติทำการศึกษาที่จะพยากรณ์ทิศทางของราคาหุ้นรายวันจากข้อความข่าวภาษาไทย โดยใช้การประมวลผลภาษาธรรมชาติ (Natural Language Processing; NLP) ซึ่งนำข้อความข่าวจาก 6 แหล่ง คือ 1) สำนักข่าวกรุงเทพธุรกิจ 2) สำนักข่าวหุ้นอินไซด์ 3) สำนักข่าวข่าวหุ้น 4) สำนักข่าวโพสต์ทูเดย์ 5) สำนักข่าว Innnews และ 6) สำนักข่าวหุ้นสมาร์ท ทำการตัดข้อความข่าวแยกเป็นคำใช้ Library Pythainlp ซึ่งในวิจัยใช้ 9 วิธี คือ 'cfcut', 'deepcut', 'etcc', 'longest', 'multi_cut', 'newmm', 'ssg', 'tcc' และ 'trie' และหุ้นที่ทำการพยากรณ์ 5 หุ้น ได้แก่ 1) INTUCH 2) TOP 3) BH 4) CPALL และ 5) KBANK สรุปผลการรพยากรณ์ หุ้น INTUCH บริษัท อินทัช โฮลดิ้งส์ จำกัด มีร้อยละความถูกต้องมากสุดที่ 100 และหุ้นอีก 4 ตัว มีร้อยละความถูกต้องเท่ากับ 66.67

2. วัสดุ อุปกรณ์และวิธีการวิจัย

2.1 ระเบียบวิธีการวิจัย

การวิเคราะห์การเคลื่อนไหวของแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) โดยใช้การเรียนรู้ของเครื่องมือผ่านการวิเคราะห์ขั้วความรู้สึกบนข่าวสารออนไลน์ มีขั้นตอนซึ่งสามารถแบ่งออกเป็นขั้นตอน



รูปที่ 4 ขั้นตอนการดำเนินงาน

ดังรูปที่ 4 การวิเคราะห์แนวโน้มของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) โดยใช้การเรียนรู้ของเครื่องมือผ่านการวิเคราะห์ขั้วความรู้สึกบนข่าวสารออนไลน์ โดยอาศัยข้อมูลทุติยภูมิ (Secondary Data) ที่เก็บรวบรวมจากแหล่งข้อมูลต่างๆ ระหว่าง วันที่ 2 เดือนตุลาคม พ.ศ. 2563 ถึง วันที่ 10 เดือนกุมภาพันธ์ พ.ศ. 2564 ซึ่งแหล่งข้อมูลที่ใช้ได้แก่ เว็บไซต์ www.investing.com เป็นแพลตฟอร์มที่ให้บริการข้อมูลเกี่ยวกับตลาดการเงินแบบเรียลไทม์ มีราคาซื้อขายล่าสุด กราฟ เครื่องมือทางการเงิน ข่าวสาร และบทวิเคราะห์สำหรับตลาดแลกเปลี่ยนจำนวน 250 แห่งทั่วโลก

หลังจากการรวบรวมเนื้อหาข้อมูลข่าวสาร ต่อมาทำการนำเนื้อหาข่าวเข้าสู่กระบวนการตัดคำ การลดคำและกำจัดคำหยุด แล้วเข้าสู่กระบวนการวิเคราะห์ขั้วความรู้สึก ใช้วิธีการนับความถี่ของคำ (Term Frequency Inverse Document Frequency; TF-IDF) และใช้เครื่องมือ VADER เพื่อคำนวณหาค่า Compound สำหรับใช้พิจารณาขั้วความรู้สึกของแต่ละข่าวและคำนวณหาสัดส่วนของข่าวขั้วความรู้สึกเชิงบวก ข่าวขั้วความรู้สึกเชิงลบ และข่าวขั้วความรู้สึกเป็นกลางในแต่ละวัน

ส่วนข้อมูลที่มีโครงสร้าง ได้แก่ ข้อมูลราคาทอง ข้อมูลราคาน้ำมัน Brent ข้อมูลเส้นค่าเฉลี่ยเคลื่อนที่ และข้อมูลราคาย้อนหลังของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) รายวัน สามารถเก็บรวบรวมจากเว็บไซต์ www.investing.com ซึ่งเป็นเว็บไซต์ที่สามารถทำการดาวน์โหลดไฟล์ราคาทองย้อนหลังได้ ข้อมูลที่ถูกดาวน์โหลดจะอยู่ในรูปแบบไฟล์โปรแกรม Microsoft Excel นามสกุล xlsx จากนั้นจัดเตรียมข้อมูลและปรับข้อมูลโดยวิธี Normalization ด้วย Min-Max Scaling เนื่องจากตัวเลขของข้อมูลแต่ละตัวแปรอิสระมีค่าที่แตกต่างกันมาก

เมื่อจัดการข้อมูลจะได้ชุดข้อมูลใหม่หลังการทำ Normalization จะได้ตัวแปรที่ใช้ในการศึกษาเพื่อทำการทดสอบตัวแปรต้นเหตุ (Granger Causality Tests) ว่าตัวแปรข่าวที่มีขั้วความรู้สึกเชิงบวก ตัวแปรข่าวที่มีขั้วความรู้สึกเชิงลบ และตัวแปรข่าวที่มีขั้วความรู้สึกเป็นกลางเป็นเหตุต่อการเปลี่ยนแปลงแนวโน้มของอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) หรือไม่ จากนั้นทำการวิเคราะห์สถิติเชิงพรรณนา และ สร้างตัวแบบพยากรณ์โดยอาศัยเทคนิคการจำแนกกลุ่มได้แก่ โครงข่ายประสาทเทียม

ซัพพอร์ตเวกเตอร์แมทชีน และการจำแนกประเภทแบบการสุ่มป่าไม้

2.2 ตัวแปรที่ใช้ในการศึกษา

ตัวแปรอิสระ (X) และตัวแปรตาม (Y) แสดงรายละเอียดดังตารางที่ 1 โดยจำนวนระเบียนเท่ากับ 92 ระเบียน

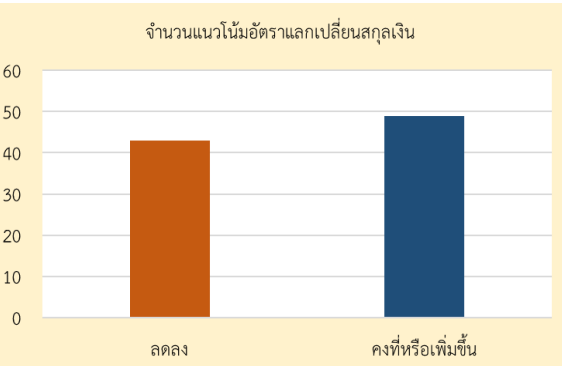
ตารางที่ 1 รายละเอียดตัวแปรอิสระและตัวแปรตาม

ชื่อตัวแปร	คำอธิบายตัวแปร	มาตรวัด	ระดับค่า
ตัวแปรอิสระ (X)			
x_1	สัดส่วนข่าวที่มีขั้วความรู้สึกเชิงบวก (Positive)	มาตรอัตราส่วน	
x_2	สัดส่วนข่าวที่มีขั้วความรู้สึกเชิงลบ (Negative)	มาตรอัตราส่วน	
x_3	สัดส่วนข่าวที่มีขั้วความรู้สึกเป็นกลาง (Neutrally)	มาตรอัตราส่วน	
x_4	ราคาทอง หน่วยราคาทองเป็นดอลลาร์สหรัฐต่อทรอยออนซ์ (USD/Oz.)	มาตรอัตราส่วน	
x_5	ราคาน้ำมันดิบ Brent หน่วยราคาดอลลาร์สหรัฐต่อบาร์เรล	มาตรอัตราส่วน	
x_6	เส้นค่าเฉลี่ยเคลื่อนที่ Moving Average	มาตรอัตราส่วน	
ตัวแปรตาม (Y)			
Y	แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD)	นามบัญญัติ	0 = ลดลง 1 = คงที่หรือเพิ่มขึ้น

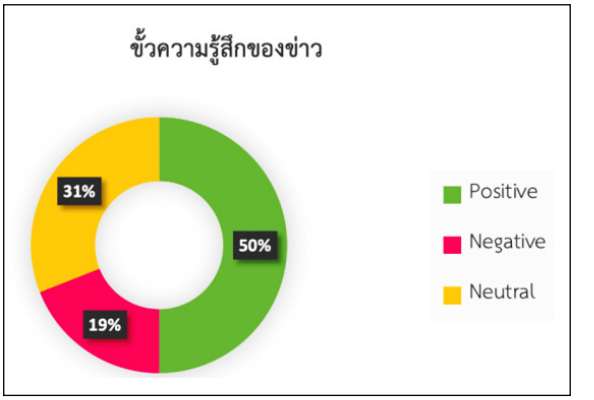
3. ผลการทดลอง

ข้อมูลที่นำมาใช้ในการสร้างตัวแบบมีจำนวน 92 ระเบียน โดยตัวแปรสัดส่วนของข่าวที่มีขั้วความรู้สึกเชิงบวก เชิงลบและเป็นกลางได้มาจากการวิเคราะห์ขั้วความรู้สึกของข่าวออนไลน์จำนวน 1,432 ข่าว

จำนวนแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) มีทั้งหมด 92 ค่า โดยแบ่งเป็นแนวโน้มลดลง 43 ค่า และแนวโน้มคงที่หรือเพิ่มขึ้น 49 ค่า



รูปที่ 5 จำนวนแนวโน้มของอัตราแลกเปลี่ยน (EUR/USD)



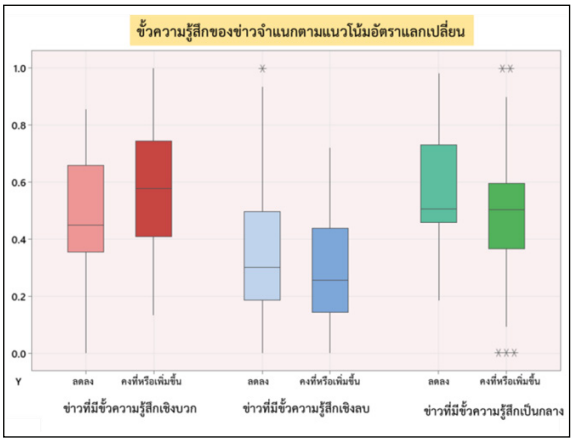
รูปที่ 6 การจำแนกประเภทข้อความทั้ง 3 ขั้วความรู้สึก

โดยชุดข้อมูลไม่มีปัญหา Imbalance Data นั่นคือสัดส่วนแนวโน้มคงที่/เพิ่มขึ้นมีสัดส่วนใกล้เคียงกับสัดส่วนแนวโน้มลดลง แสดงดังรูปที่ 5

ผลการวิเคราะห์การประมวลผลภาษาทางธรรมชาติข้อความข่าวจากสื่อสังคมออนไลน์ที่เกี่ยวข้องหรือส่งผลต่อแนวโน้ม และพยากรณ์แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) จำนวน 1,432 ข่าว โดยข่าวที่มีขั้วความรู้สึกเชิงบวก ข่าวที่มีขั้วความรู้สึกเชิงลบ และข่าวที่มีขั้วความรู้สึกเป็นกลางมีสัดส่วนดังรูปที่ 6 และจำนวนข่าวจำแนกตามขั้วความรู้สึกของข่าวและแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) นั้นคือ แนวโน้มลดลงและแนวโน้มคงที่หรือเพิ่มขึ้น ดังตารางที่ 2

ตารางที่ 2 จำนวนข่าวจำแนกตามขั้วความรู้สึกของข่าวจำแนกและแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD)

ขั้วความรู้สึกของข่าว	แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD)		รวม
	ลดลง	คงที่หรือเพิ่มขึ้น	
ข่าวที่มีขั้วความรู้สึกเชิงบวก	302	407	709
ข่าวที่มีขั้วความรู้สึกเชิงลบ	144	129	273
ข่าวที่มีข้าวความรู้สึกเป็นกลาง	234	216	450
รวม	680	752	1,432



รูปที่ 7 คะแนนขั้วความรู้สึกของข่าวจำแนกตามขั้วความรู้สึกของข่าวและแนวโน้มอัตราแลกเปลี่ยน

กราฟ Box-Plot แสดงขั้วความรู้สึกของข่าวจำแนกตามแนวโน้มอัตราแลกเปลี่ยน ซึ่งแกนแนวตั้ง (Y) แทน คะแนนขั้วความรู้สึก ดังรูปที่ 7 โดยข่าวที่มีขั้วความรู้สึกเชิงบวกที่อัตราแลกเปลี่ยนมีแนวโน้มลดลงมีค่ามัธยฐานต่ำกว่าข่าวที่มีขั้วความรู้สึกเชิงบวกที่อัตราแลกเปลี่ยนมีแนวโน้มคงที่หรือเพิ่มขึ้นหมายความว่าข่าวที่มีขั้วความรู้สึกเชิงบวกส่วนใหญ่ให้ผลแนวโน้มอัตราแลกเปลี่ยนเป็นแนวโน้มคงที่หรือเพิ่มขึ้น ส่วนข่าวที่มีขั้วความรู้เชิงลบที่อัตราแลกเปลี่ยนมีแนวโน้มลดลงมีค่ามัธยฐานสูงกว่าที่อัตราแลกเปลี่ยนมีแนวโน้มคงที่หรือเพิ่มขึ้นนั้นคือข่าวที่มีขั้วความรู้สึกเชิงลบส่วนใหญ่ให้ผลแนวโน้มของอัตราแลกเปลี่ยนแนวโน้มลดลงและข่าวที่มีขั้วความรู้สึกเป็นกลางจะเห็นว่าข่าวที่มีขั้วความรู้สึกเป็นกลางที่อัตราแลกเปลี่ยนมีแนวโน้มลดลงและแนวโน้มคงที่หรือเพิ่มขึ้นมีค่า มัธยฐานใกล้เคียงกันหมายถึง ข่าวที่มีขั้วความรู้สึกเป็นกลางจะให้ผลแนวโน้มอัตราแลกเปลี่ยนแนวโน้มลดลงและแนวโน้มคงที่หรือเพิ่มขึ้นเท่าๆ กัน

ตัวอย่างกลุ่มคำจากข่าวที่มีขั้วความรู้สึกเชิงบวก ได้แก่ High, Gain, Support, Fresh, Vaccine, Positive, Beyond และ Top เป็นต้น ส่วนตัวอย่างกลุ่มคำจากข่าวที่มีขั้วความรู้สึกเชิงลบ ได้แก่ low, Drop, Covid, Back, Risk, Lockdown, Bear, Resist และ Lose เป็นต้น และตัวอย่างกลุ่มคำจากข่าวที่มีขั้วความรู้สึกเป็นกลาง ได้แก่ Elect, Remain, Ahead, Near, Around, Pair, Market และ New เป็นต้น

ตารางที่ 3 ผลการทดสอบสมมติฐานเชิงเป็นเหตุเป็นผล (Granger Causality Tests)

ตัวแปรขั้วความรู้สึกกับแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD)	F-test	*p*-value
ข่าวที่มีขั้วความรู้สึกเชิงบวกกับแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD)	2.77	0.04
ข่าวที่มีขั้วความรู้สึกเชิงลบกับแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD)	5.40	0.01
ข่าวที่มีขั้วความรู้สึกเป็นกลางกับแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD)	1.78	0.04

จากตารางที่ 3 จะเห็นว่าค่า p-value = 0.04 น้อยกว่า a = 0.05 จึงปฏิเสธสมมติฐานหลัก H_0 นั่นคือตัวแปรข่าวที่มีขั้วความรู้สึกเชิงบวก ตัวแปรข่าวที่มีขั้วความรู้สึกเชิงลบ และตัวแปรข่าวที่มีขั้วความรู้สึกเป็นกลางเป็นสาเหตุการเปลี่ยนแปลงของแนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD)

ตารางที่ 4 ผลการเปรียบเทียบวิธีการพยากรณ์ทั้ง 3 ตัวแบบ

การทดสอบประสิทธิภาพแบบแบ่งข้อมูลแบบวิธี Split Test				
ตัวแบบ	**Accuracy**	**Precision**	**Recall**	**F1-Score**
โครงข่ายประสาทเทียมแบบหลายชั้น (2 Layer)	60.14 %	56.25 %	69.23 %	62.07 %
ซัพพอร์ตเวกเตอร์แมทชีน	60.71 %	55.56 %	76.92 %	64.52 %
การจำแนกประเภทแบบการสุ่มป่าไม้	71.43 %	63.16 %	92.31 %	74.99 %
การทดสอบประสิทธิภาพแบบ Cross Validation				
โครงข่ายประสาทเทียมแบบหลายชั้น (2 Layer)	59.78 %	62.50 %	61.22 %	61.86 %
ซัพพอร์ตเวกเตอร์แมทชีน	59.78 %	66.66 %	48.98 %	56.47 %
การจำแนกประเภทแบบการสุ่มป่าไม้	55.43 %	60.00 %	48.98 %	53.93 %

สรุปผลการทดสอบจากตารางที่ 4 พบว่า การทดสอบประสิทธิภาพแบบแบ่งข้อมูลแบบวิธี Split Test และการทดสอบประสิทธิภาพแบบ Cross Validation จะได้ว่าตัวแบบที่สร้างขึ้นด้วยการจำแนกประเภทแบบการสุ่มป่าไม้แบบวิธี Split Test โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลเรียนรู้ (Training Set) และข้อมูลชุดทดสอบ (Test Set) ด้วยสัดส่วน 70 : 30 ให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความถูกต้องร้อยละ 71.43 ค่าความแม่นยำร้อยละ 63.16 ค่าความระลึกร้อยละ 92.31 และค่าความถ่วงดุลร้อยละ 74.99

4. อภิปรายผลและสรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์ขั้วความรู้สึก กับแนวโน้มของอัตราแลกเปลี่ยน และสร้างตัวแบบพยากรณ์แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) ในตลาดซื้อขายแลกเปลี่ยนสกุลเงินระหว่างประเทศซึ่งอาศัยข้อมูลแบบไม่มีโครงสร้างได้แก่ สัดส่วนของข่าวขั้วความรู้สึกเชิงบวก เชิงลบและเป็นกลางที่ได้จากการวิเคราะห์ขั้วความรู้สึก และข้อมูลแบบมีโครสร้างได้แก่ราคาทองคำ น้ำมัน และค่าเฉลี่ยเคลื่อนที่ โดยวิธีโครงข่ายประสาทเทียม (Artificial Neural Network) ซัพพอร์ตเวกเตอร์แมทชีน (Support Vector Machine) และการจำแนกประเภทแบบการสุ่มป่าไม้ (Random Forest) และประเมินประสิทธิภาพจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F1-Score) และทำการสรุปผลของข้อมูลโดยพิจารณาจากค่าความถูกต้อง (Accuracy) เนื่องจากสามารถบอกได้ว่าตัวแบบที่ใช้ในการพยากรณ์แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบดอลลาร์สหรัฐ (EUR/USD) นั้นมีความถูกต้องมากน้อยเพียงใด

ผลการศึกษาพบว่า สัดส่วนของจำนวนข่าวที่มีขั้วความรู้สึกเชิงบวกสูงสุด ตามด้วยข่าวขั้วที่มีความรู้สึกเป็นกลางและข่าวที่มีขั้วความรู้สึกเชิงลบ นอกจากนี้ผลการทดสอบสมมติฐานเชิงเป็นเหตุเป็นผลของข้อมูลข่าวสารกับแนวโน้มอัตราแลกเปลี่ยน สกุลเงิน (EUR/USD) สรุปผลได้ว่าข่าวที่มีขั้วความรู้สึกเชิงบวก ข่าวที่มีขั้วความรู้สึกเชิงลบ และข่าวที่มีขั้วความรู้สึกเป็นกลาง เป็นสาเหตุการเปลี่ยนแปลงของแนวโน้มอัตราแลกเปลี่ยนสกุลเงิน (EUR/USD) ซึ่งสอดคล้องกับงานวิจัย [16] ที่คาดการณ์ความสัมพันธ์ของหัวข้อข่าวสารทั้งข่าวเชิงบวกและข่าวเชิงลบกับอัตราแลกเปลี่ยน ระหว่างประเทศคู่สกุลเงิน (EUR/USD) ผลคือ ความสัมพันธ์ดังกล่าวมีอยู่จริงและข่าวสารมีผลหรือมีความสัมพันธ์กับอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) ในราย 2 ชั่วโมง

ตัวแบบการจำแนกประเภทแบบการสุ่มป่าไม้ ด้วยวิธี Split Test ซึ่งเป็นการแบ่งข้อมูล ออกเป็น 2 ชุด คือ ชุดข้อมูลเรียนรู้ (Training Set) และข้อมูลชุดทดสอบ (Test Set) ให้ค่าความถูกต้อง (Accuracy) มากที่สุดอยู่ที่ร้อยละ 71.43 รองลงมาเป็นซัพพอร์ตเวกเตอร์แมทชีนและโครงข่ายประสาทเทียมแบบหลายชั้น โดยมีค่าความถูกต้องอยู่ที่ร้อยละ 60.71 และ 60.14 ตามลำดับ จึงนำตัวแบบที่เหมาะสมนี้

ไปพยากรณ์แนวโน้มอัตราแลกเปลี่ยนสกุลเงินยูโรเทียบกับดอลลาร์สหรัฐ (EUR/USD) เพื่อเป็นแนวทางช่วยในการตัดสินใจของนักลงทุนหรือผู้ที่สนใจที่จะลงทุนในตลาดซื้อขายแลกเปลี่ยนสกุลเงินระหว่างประเทศ

ผลการศึกษาเป็นไปตามที่ [19] กล่าวไว้ว่า การจำแนกประเภทแบบการสุ่มป่าไม้มีประสิทธิภาพที่ดีกว่าในแง่ของความแม่นยำและการจดจำโดยเฉพาะอย่างยิ่งกับชุดทดสอบที่มีขนาดเล็ก เนื่องจากขนาดตัวอย่างในงานวิจัยครั้งนี้มีเพียง 92 ตัวอย่าง อย่างไรก็ตาม แนวทางการปรับปรุงในอนาคตควรเพิ่มจำนวนข้อมูลที่ใช้ในการศึกษา โดยทั่วไปแล้วควรจะไม่ต่ำกว่า 1,000 ตัวอย่าง

เอกสารอ้างอิง

[1] Moneyhub. (2016, August). *Forex*. [Online]. (in Thai). Available: https://moneyhub.in.th/article/forex-fraud/.

[2] Bank for international settlements. (2013). *Triennial Central Bank Survey*. [Online]. Available: https://www.bis.org/publ/rpfx13fx.pdf.

[3] M. Vochozka, Z. Rowland, P. Suler, and J. Marousek, "The influence of the international price of oil on the value of the EUR/USD exchange rate," *Journal of Competitiveness*, vol. 12, no. 2, pp. 167–190, 2020.

[4] S. Sindhu, "A study on impact of select factors on the price of Gold," *IOSR Journal of Business and Management*, vol. 8, no. 4, pp. 84–93, 2013.

[5] Messi. (2019, September). *EUR/USD before Non Farm announcement*. [Online]. (in Thai). Avaiable: https://www.thaifrx.com/eur-usd.

[6] TradingView ND. *EURUSD Forex Charts*. [Online]. (in Thai). Available: https://th.tradingview.com/symbols/EURUSD.

[7] M. Taboada and J. Grieve, "Analyzing Appraisal Automatically," in *AAAI Spring Symposium on Exploring Attitude and Affect in Text*, 2004, pp. 158–161.

[8] P. Puangprakhon and S. Narupiti, "Application of artificial neural networks and neighboring sections relationship in the short-term travel time prediction on urban roadways," presented at 25th National Convention on Civil Engineering, Chonburi, Thailand, July 15–17, 2020 (in Thai).

[9] G. W. Rwin, K. Warwick, and K. J. Hunt, *Neural Network Applications in Control*. United Kingdom: Stevenage, 1995.

[10] A. Hirunyawanakul, "A technique to improve image recognition performance of handwritten signature," M.S. thesis, Department of Computer Engineering, Faculty of Engineering, Suranaree University of Technology, 2014 (in Thai).

[11] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18–28, 1998.

[12] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[13] R. S. Pindyck and D. L. Rubinfeld, *Econometric Models and Economic Forecasts*, New York: McGraw-Hill, 1998.

[14] G. Angiani, L. Ferrari, T. Fontanini, P. Fornacciari, E. Iotti, F. Magliani, and S. A. Manicardi, "A comparison between preprocessing techniques for sentiment analysis in twitter," *KDWeb*, 2016, pp. 1–11.

[15] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Computer Science*, vol. 17, pp. 26–32, 2013.

[16] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah,

and D. C. L. Ngo, "Text mining of news-headlines for FOREX market prediction: A multi-layer dimension reduction algorithm with semantics and sentiment," *Expert Systems with Application*, vol. 42, no. 1, pp. 306–324, 2015.

[17] S. M. Islam, E. Hossain, A. Rahman, M. S. Hossain, and K. Andersson, "A review on recent advancements in FOREX currency prediction," Algorithms, vol. 13, no. 8, pp. 186, 2020.

[18] W. Phaphan and A. Pimpisal, "The predictions of a daily stock price direction from the Thai news content by using natural language processing," *The Journal of Applied Science*, vol. 19, no. 1, pp. 59–79, 2020.

[19] D. Jiang, T. Han, Q. Zhao, and L. Wang, "Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 8, pp. 2681–2693, 2017.