วัลลภ คุ้มประดิษฐ์. "การพยากรณ์อัตราแลกเปลี่ยนเงินบาทต่อดอลลาร์สหรัฐด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน". รายงานสืบเนื่องจากการประชุมวิชาการระดับชาติของนักเศรษฐศาสตร์ ครั้งที่ 12. เมื่อวันที่ 13 กรกฎาคม 2561 จัดโดยคณะเศรษฐศาสตร์ มหาวิทยาลัยรามคำแหงร่วมกับ 8 สถาบัน. หน้า 214-231

การพยากรณ์อัตราแลกเปลี่ยนเงินบาทต่อดอลลาร์สหรัฐฯ ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน

Exchange Rate of Thai Baht against the US Dollar Forecasting using Support Vector Machines Technique

วัลลภ คุ้มประดิษฐ์[1]
Wonlope Khumpradith[1]

บทคัดย่อ

ปัจจุบันเทคนิคทางด้าน "การเรียนรู้ของเครื่อง" (Machine Learning) ถูกนำมาประยุกต์ใช้อย่างแพร่หลายทั้งสาขาทางด้านวิทยาศาสตร์ และสังคมศาสตร์ เพื่อทำการดึงสาระสำคัญจากข้อมูลและสารสนเทศ ที่มีอยู่อย่างมากมายนำไปวิเคราะห์และสังเคราะห์ อีกทั้งนำไปใช้ในการอธิบายและจำแนกรูปแบบ (Classification) ต่างๆ ผ่านวิธี "การรู้จำแบบ" (Pattern Recognition) ตามวัตถุประสงค์ของงานวิจัย รวมถึงยังสามารถนำมาใช้ในการทำนาย/การคาดการณ์ (Prediction/Forecasting) จากการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ ทั้งนี้การประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องมือมีอิทธิพลอย่างมากต่อการศึกษาทางด้านเศรษฐศาสตร์และการวิเคราะห์ทางธุรกิจ โดยเฉพาะการนำเทคนิคดังกล่าวไปใช้ในการทำนายและการคาดการณ์ตัวแปรทางเศรษฐกิจต่างๆ อีกทั้งงานศึกษาในเรื่องการพยากรณ์อัตราแลกเปลี่ยน ถือได้ว่าเป็นเรื่องที่ท้าทายต่อแนวคิดหลัก ความมีประสิทธิภาพของตลาด (Market Efficiency) ที่สะท้อนจากสมมติฐานเกี่ยวกับการเคลื่อนไหวของอัตราแลกเปลี่ยนและราคาสินทรัพย์ทางการเงินในตลาดหลักทรัพย์ (Foreign Exchange Rate and Financial Asset Prices) มีการเคลื่อนไหวตามลักษณะ Random Walk การศึกษานี้มุ่งเน้นศึกษาเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เพื่อ (1) พยากรณ์อัตราแลกเปลี่ยนสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐฯ (BHT/USD) รายวันด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน และ (2) เพื่อเปรียบเทียบประสิทธิภาพการพยากรณ์อัตราแลกเปลี่ยนด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีนกับแบบจำลอง Random Walk และแบบจำลอง Autoregressive Integrated Moving Average (ARIMA)

ผลการศึกษาพบว่า การพยากรณ์อัตราแลกเปลี่ยนสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ เมื่อพิจารณาจากความถูกต้องของทิศทางในการคาดการณ์ (Mean Directional Accuracy) แบบจำลอง SVM เกือบทุกกรณีมีประสิทธิภาพสูงกว่าแบบจำลอง ARIMA และแบบจำลอง Random walk แต่หาก

[1] คณะเศรษฐศาสตร์ มหาวิทยาลัยรามคำแหง 2086 ถนน รามคำแหง แขวง หัวหมาก เขต บางกะปิ กรุงเทพมหานคร 10240
* Faculty of Economics, Ramkhamhaeng University, E-mail address: wonlope.kh@gmail.com

พิจารณาด้วยค่า RMSE พบว่า แบบจำลอง ARIMA มีประสิทธิภาพสูงสุด รองลงมาคือ แบบจำลอง SVM และแบบจำลอง Random Walk

คำสำคัญ: อัตราแลกเปลี่ยน, ซัพพอร์ตเวกเตอร์แมชชีน, การพยากรณ์

Abstract

Currently, there is a new statistical technique called "Machine Learning" which has been applied widely in the field of science and social science. This technique is use for analyzing and synthesizing data to spotlight its significant information. Moreover, it is used to describe and classify patterns of data through the "Pattern Recognition" according to the purpose of researches. This technique can also create algorithms which used to learn information directly from data and then predict future outputs. The application of these learning techniques has a great influence on economic studies and business analytics especially in forecasting economic variables such as exchange rate. The study of forecasting exchange rate is a challenge to the traditional thought. The main idea of the believe is when the market is efficiency, foreign exchange rate and financial asset prices follow a random walk. As a result, this study is focus on using machine learning technique to (1) predict daily exchange rates of BHT/USD by Support Vector Machines technique (SVM), and (2) compare the efficiency of the forecasting models among Support Vector Machines technique, Random Walk model and Autoregressive Moving Average model (ARMA).

The result of the study shows that the forecast of the BHT/USD exchange rate based on directional accuracy illustrate that SVM techniques are more efficient than ARMA models and random walk models. On the other hand, for RMSE value, ARMA models are most efficient comparing to SVM and Random Walk models.

Keyword: Exchange Rate, Support Vector Machine (SVM), Forecasting

บทนำ

ปัจจุบันเทคนิคทางด้าน "การเรียนรู้ของเครื่อง" (Machine Learning) ถูกนำมาประยุกต์ใช้อย่างแพร่หลายทั้งสาขาทางด้านวิทยาศาสตร์ และสังคมศาสตร์ เพื่อทำการดึงสาระสำคัญจากข้อมูลและสารสนเทศ ที่มีอยู่อย่างมากมายนำไปวิเคราะห์และสังเคราะห์ อีกทั้งนำไปใช้ในการอธิบายและจำแนกรูปแบบ (Classification) ต่างๆ ผ่านวิธี "การรู้จำแบบ" (Pattern Recognition) ตามวัตถุประสงค์ของงานวิจัย รวมถึงยังสามารถนำมาใช้ในการทำนาย/การคาดการณ์ (Prediction/Forecasting) จากการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้

ทั้งนี้ การประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องมือมีอิทธิพลอย่างมากต่อการศึกษาทางด้านเศรษฐศาสตร์และการวิเคราะห์ทางธุรกิจ โดยเฉพาะการนำเทคนิคดังกล่าวไปใช้ในการทำนายและการคาดการณ์ตัวแปรทางเศรษฐกิจต่างๆ เช่น งานศึกษาของ Sheta, (2015) ทำการศึกษาและเปรียบเทียบประสิทธิภาพของเทคนิคต่างๆ จากการเรียนรู้ของเครื่องในการคาดการณ์ดัชนีตลาดหลักทรัพย์ Plakandaras (2015) ใช้เทคนิค Support Vector Machines (SVMs) และ Artificial Neural Networks (ANNs) ในการพยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศรายวันและรายเดือน และเช่นเดียวกับงานศึกษาของ Theofilatos (2012), Soulas, (2013), Dumitru (2013), Xhaja (2014), และ Nur Izzati Abdul Rashid (2016) ที่ศึกษาเทคนิคการเรียนรู้ของเครื่องต่างๆ เพื่อพยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศ รวมถึง Guo (2012) ศึกษาเทคนิคดังกล่าวเพื่อการคาดการณ์ราคาน้ำมันดิบในตลาดโลก

การศึกษานี้มุ่งเน้นศึกษาเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เพื่อทำนาย/พยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศ โดยเน้นไปที่เทคนิคซัพพอร์ตเวกเตอร์แมชชีนที่งานศึกษาข้างต้นได้นำไปใช้ในการพยากรณ์ตัวแปรทางเศรษฐกิจและตัวแปรทางการเงินต่างๆ ได้อย่างมีประสิทธิภาพ การศึกษานี้เน้นไปที่การพยากรณ์อัตราแลกเปลี่ยนเงินต่างประเทศรายวัน เนื่องด้วยอัตราแลกเปลี่ยนเงินตราต่างประเทศ (Foreign Exchange Rates) มีบทบาทสำคัญอย่างมากต่อการค้าและการลงทุนระหว่างประเทศ ทั้งในมุมของภาคธุรกิจและในมุมของการบริหารเสถียรภาพทางเศรษฐกิจของผู้กำกับดูแลหรือภาครัฐ การเปลี่ยนแปลงและความผันผวนของอัตราแลกเปลี่ยนส่งผลโดยตรงต่อกำไรและขาดทุนของภาคธุรกิจที่ทำการค้าและการลงทุนระหว่างประเทศ ขณะเดียวกันในระดับภาพรวมของประเทศ การเปลี่ยนแปลงและความผันผวนดังกล่าวย่อมส่งผลกระทบต่อเสถียรภาพทางเศรษฐกิจ เช่น ระดับราคาสินค้า หนี้สาธารณะ ความเชื่อมั่นของนักลงทุน และตัวแปรเศรษฐกิจต่างๆ ที่สามารถต่อการเจริญเติบโตทางเศรษฐกิจและอัตราเงินเฟ้อของประเทศได้ ทั้งนี้ เงินตราต่างประเทศมีบทบาทในการเป็นสื่อกลางการแลกเปลี่ยนสินค้าและบริการระหว่างประเทศแล้ว ยังเป็นสินค้าชนิดหนึ่งในตลาดที่มีราคาเป็นอัตราแลกเปลี่ยนเงินตราตามสกุลของแต่ละประเทศเปรียบเทียบกับประเทศอื่น โดยการเปลี่ยนแปลงเป็นไปตามทฤษฎีอุปสงค์-อุปทาน และการเก็งกำไรของเงินตราต่างประเทศ ดังนั้น การพยากรณ์แนวโน้มการเปลี่ยนแปลงของอัตราแลกเปลี่ยนย่อมสามารถที่จะป้องกันความเสี่ยง หรือสามารถแสวงหากำไร/ขาดทุนจากความผันผวนดังกล่าวได้

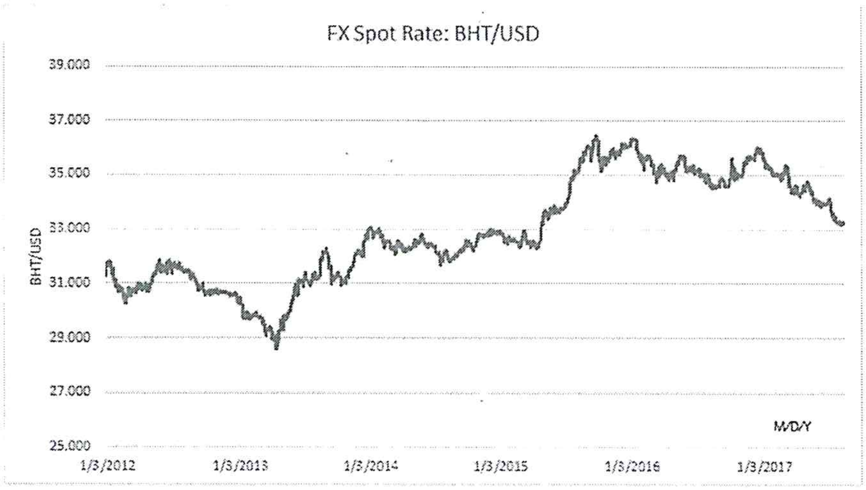
นอกจากนี้ งานศึกษาในเรื่องการพยากรณ์อัตราแลกเปลี่ยน ถือได้ว่าเป็นเรื่องที่ท้าทายต่อแนวคิดหลักของทฤษฎีความมีประสิทธิภาพของตลาด (Market Efficiency) ที่สะท้อนจากสมมติฐานเกี่ยวกับการเคลื่อนไหวของอัตราแลกเปลี่ยนและราคาสินทรัพย์ทางการเงินในตลาดหลักทรัพย์ (Foreign Exchange Rate and Financial Asset Prices) มีการเคลื่อนไหวตามลักษณะ Random Walk (Fama, 1965) กล่าวอีกนัยหนึ่งคือ การเคลื่อนไหวของราคาหรืออัตราแลกเปลี่ยนไม่สามารถที่จะคาดการณ์ได้อย่างถูกต้อง (Unpredictable) หรือกล่าวได้ว่า ไม่มีแบบจำลองใดที่มีประสิทธิภาพหรือสามารถพยากรณ์ได้ดีกว่าแบบจำลอง Random Walk (RW Model) อย่างไรก็ตามมีงานศึกษาหลายงานศึกษา (Lo & MacKinaly, 1999) พยายามสร้างแบบจำลองและค้นคว้าหาเทคนิคทางสถิติอย่างต่อเนื่องเพื่อทำพยากรณ์อัตราแลกเปลี่ยน และพิสูจน์ผลการพยากรณ์ดังกล่าวมีประสิทธิภาพสูงกว่า RW Model ภายใต้เงื่อนไขของแบบจำลองที่สร้างขึ้น เช่น แบบจำลอง Markov Switching แบบจำลอง Vector Autoregressive (VAR) แบบจำลอง Vector Error Correction (VECM) แบบจำลอง Time-Varying parameters ต่างๆ แบบจำลอง Generalized Autoregressive Conditional Heteroskedasticity (GARCH) แบบจำลอง Bayesian VAR รวมถึงใช้เทคนิค Machine Learning เช่น Artificial Neural Networks (ANN) แบบจำลอง Support Vector Machine (SVM) มาใช้ในการทำนาย/พยากรณ์อัตราแลกเปลี่ยนในปัจจุบัน

ดังนั้นวัตถุประสงค์ของงานศึกษาคือ 1) เพื่อพยากรณ์อัตราแลกเปลี่ยนสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ (BHT/USD) รายวัน ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน และ 2) เพื่อเปรียบเทียบประสิทธิภาพการพยากรณ์อัตราแลกเปลี่ยนด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน กับแบบจำลอง Random Walk และแบบจำลอง Autoregressive Moving Average (ARMA)

วิธีการศึกษา

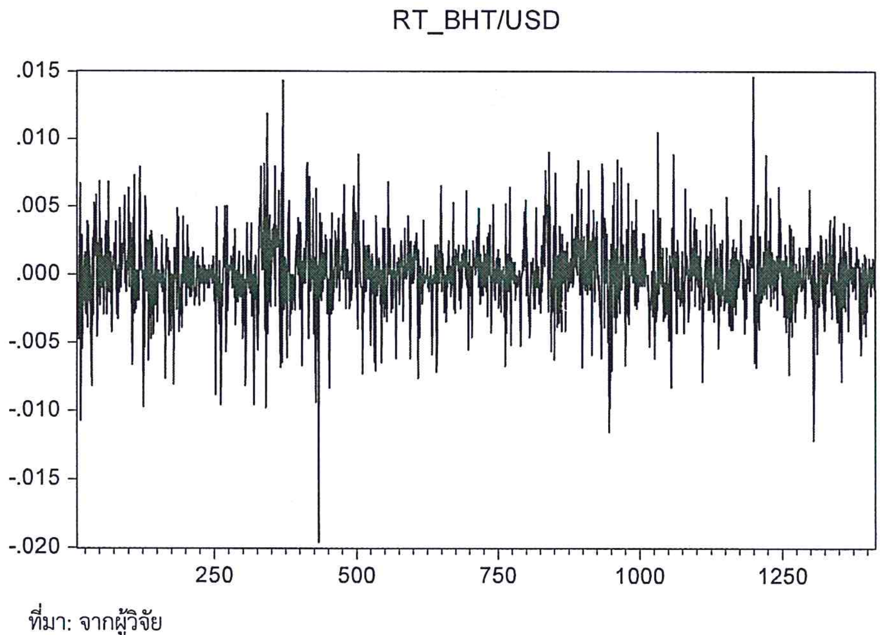
การศึกษานี้ใช้ข้อมูลตั้งแต่ปี มกราคม 2012 - สิงหาคม 2017 โดยนำข้อมูลมาจากธนาคารแห่งประเทศไทย โดยมีจำนวนข้อมูลทั้งสิ้น 1,413 วัน ดังแสดงในแผนภาพด้านล่าง

แผนภาพที่ 1 อัตราแลกเปลี่ยนรายวันของสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ



ที่มา: ธนาคารแห่งประเทศไทย

แผนภาพที่ 2 อัตราผลตอบแทนของอัตราแลกเปลี่ยนรายวันของสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ $\left(R_t = \frac{P_t - P_{t-1}}{P_{t-1}}\right)$



ที่มา: จากผู้วิจัย

ตัวแปรที่ใช้ในการศึกษาแบ่งออกเป็น 2 ลักษณะ คือ 1) ตัวแปรภายใน (Endogenous variable และ 2) ตัวแปรภายนอก (Exogenous variable) และเนื่องด้วยการทดสอบ Unit Root ข้างต้นพบว่า อัตราแลกเปลี่ยนสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ เป็นข้อมูลที่ไม่นิ่ง (Non-stationary) ซึ่งจำเป็นต้องทำการ first different $(P_t - P_{t-1})$ และทดสอบ Unit Root ซึ่งพบว่า ข้อมูลดังกล่าวมีคุณสมบัติความนิ่ง อย่างไรก็ตามแทนที่จะใช้ first different $(P_t - P_{t-1})$ ในการศึกษานี้จึงใช้ $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ ซึ่งก็คือ อัตราผลตอบแทนของอัตราแลกเปลี่ยน โดยตัวแปรภายใน คือ อัตราผลตอบแทนของอัตราแลกเปลี่ยน ช่วงเวลาที่ t ขณะที่ตัวแปรภายนอก คือ อัตราผลตอบแทนของอัตราแลกเปลี่ยนในช่วงเวลาที่ผ่านมา โดยกำหนดช่วงเวลาออกเป็น 5 ช่วงเวลา (Windows) คือ 1) ช่วงเวลาที่ผ่านมา 1 วัน (w_1) 2) ช่วงเวลาที่ผ่านมา 1-3 วัน (w_3) 3) ช่วงเวลาที่ผ่านมา 1-5 วัน (w_5) 4) ช่วงเวลาที่ผ่านมา 1-7 วัน (w_7) และ 5) ช่วงเวลาที่ผ่านมา 1-15 วัน (w_{15})

การศึกษานี้แบ่งข้อมูลการศึกษาออกเป็น 2 ลักษณะ กล่าวคือ 1) กำหนด Training Data ร้อยละ 90 ของข้อมูลทั้งหมด (1,272 วัน) และ Testing Data เท่ากับร้อยละ 10 ของข้อมูลทั้งหมด (141 วัน) และ 2) กำหนด Training Data ร้อยละ 80 ของข้อมูลทั้งหมด (1,130 วัน) และ Testing Data เท่ากับร้อยละ 10% ของข้อมูลทั้งหมด (283 วัน) โดยอ้างอิงจาก Alamili, Mohamand. (2011) และทำการเปรียบเทียบผลการศึกษาโดยใช้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (Root Mean Square Error: RMSE) และพิจารณาความถูกต้องของทิศทางการพยากรณ์กับข้อมูลจริง (Mean Directional Accuracy) ตามงานศึกษาของ Blaskowitz, O., & Herwartz, H. (2011)

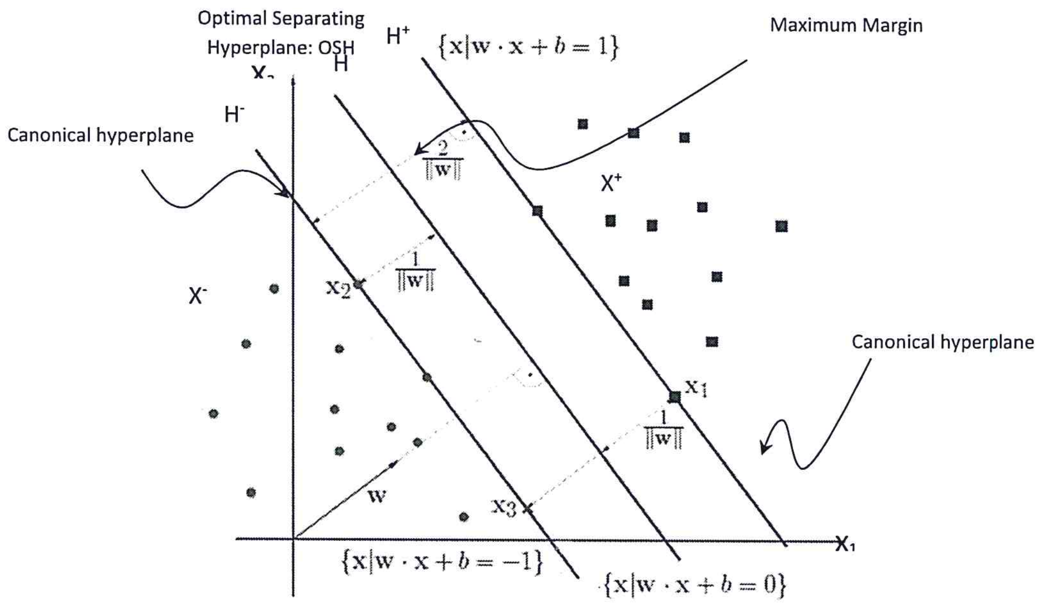
ทฤษฎีซัพพอร์ตเวกเตอร์แมชชีนสำหรับการจำแนก (Theory of Support Vector Machines for Classification)

ซัพพอร์ตเวกเตอร์แมชชีน (SVMs) เป็นเทคนิคหนึ่งที่ใช้ในการจำแนกประเภทข้อมูล (Classifier) ที่คิดค้นและพัฒนาจาก Vapnik ในปี 1992 การจำแนกประเภทต่างๆ เป็นศิลปะของการรู้จำรูปแบบที่มีลักษณะที่แน่นอนของข้อมูลชุดหนึ่งๆ (Recognizing certain pattern) โดยสามารถจำแนกได้ในหลายคลาส (Multiple classes) ซึ่งการจำแนกของ SVMs จะต้องมีการจำแนกระหว่างข้อมูลที่นำเข้าเพื่อทำการเรียนรู้ (Training data set) และข้อมูลที่ทำการทดสอบแบบจำลอง (Validation set) ทั้งนี้ SVMs ใช้ Hyperplane[2] ในการแบ่งข้อมูลเรียนรู้ (Training set) ที่มีอยู่ออกเป็น 2 ประเภท (Binary Classification) ออกจากกันได้อย่างสมบูรณ์ในลักษณะเชิงเส้น กล่าวอีกนัยหนึ่งคือ ข้อมูลทั้งหมดที่มีตำแหน่งอยู่เหนือ Hyperplane จะเป็นประเภท Positive Class และข้อมูลทั้งหมดที่มีตำแหน่งอยู่ใต้

[2] Hyperplane หมายถึง ลักษณะทั่วไปของระนาบหนึ่งๆ (An hyperplane is a generalization of a plane) ทั้งนี้หาก ใน $\mathbb{R}^1$ จะเรียกกว่า จุด (Point) ซึ่งอยู่บนเส้นจำนวนจริง ถ้าหากใน $\mathbb{R}^2$ จะเรียกว่า เส้น (Line) ที่อยู่บนแกน 2 มิติ ถ้าหากใน $\mathbb{R}^3$ จะเรียกว่า ระนาบ (plane) ที่อยู่บนแกน 3 มิติ และถ้าหาก $\mathbb{R}^n$ โดยที่ $n > 3$ จะเรียกว่า Hyperplane

Hyperplane จะเป็น Negative Class ซึ่งการที่จะหา Hyperplane มาแบ่งข้อมูลทั้งสองประเภทได้อย่างสมบูรณ์หรือไม่นั้น ขึ้นอยู่กับการกระจายของข้อมูล โดยสามารถยกตัวอย่างได้ดังแผนภาพด้านล่าง

แผนภาพที่ 3 Hyperplane H ที่แบ่งข้อมูลประเภท X^+ และ X^- ได้อย่างสมบูรณ์



จากแผนภาพข้างต้น เมื่อทำการพลอตความสัมพันธ์ของ (x_1, x_2) ใดๆ คู่ลำดับ (x_1, x_2) นั้นสามารถจำแนกได้ออกเป็นสองลักษณะ นั่นคือ (1) คู่ลำดับ (x_1, x_2) ที่อยู่ในคลาส X^+ ซึ่งมีลักษณะเป็นสี่เหลี่ยม และ (2) คู่ลำดับ (x_1, x_2) ที่อยู่ในคลาส X^- ซึ่งมีลักษณะเป็นวงกลม โดยการจำแนกข้อมูลได้ 2 คลาสตามแผนภาพข้างต้นทำได้โดยการหา Optimal Separating Hyperplane: OSH ซึ่งได้มาจากการพิจารณาขอบกั้นของเส้นแบ่งข้อมูล (Margin) ระหว่าง Hyperplane H^+ และ H^-

การจำแนกข้อมูลออกเป็น 2 ประเภท (Binary classification) โดยทั่วไปประกอบไปด้วย 2 กรณี คือ 1) จำแนกได้อย่างเป็นเชิงเส้น (Linearly separable) และ 2) ไม่สามารถจำแนกได้อย่างเป็นเชิงเส้น (Linearly inseparable) ซึ่งแผนภาพข้างต้นแสดงให้เห็นถึงการจำแนกในลักษณะกรณีแรกคือ จำแนกได้อย่างเป็นเชิงเส้น (Linearly separable) สำหรับกรณีที่สองจะต้องมีการใช้ Kernel function เพื่อแก้ไขปัญหาดังกล่าว ทำได้โดยการเปลี่ยนมิติของข้อมูลให้สูงขึ้นเพื่อช่วยในการเรียงตัวของข้อมูลใหม่ที่เรียกว่า Higher Dimensional Space หรือ Feature space ดังจะกล่าวในรายละเอียดถัดไป

จากแผนภาพข้างต้น กำหนดให้ข้อมูลสามารถจำแนกได้ออกเป็น 2 คลาสได้แก่ positive class และ negative class ซึ่ง $\mathbf{x}_i$ หรือ (x_1, x_2) หรือ ซึ่งสอดคล้องกับคลาสของ y_i นั่นคือ

$$\mathbf{x}_i \in \mathbb{R}^n, \qquad y_i \in \{-1, 1\}$$

โดยที่ $\mathbf{x}_i$ ซึ่งสอดคล้อง (corresponding to) กับ y_i ซึ่งเท่ากับ 1 หากเป็น positive class ขณะที่ y_i กล่าวอีกนัยหนึ่งคือ กำหนดให้ข้อมูลเรียนรู้ประกอบด้วยข้อมูล n จำนวน นั่นคือ

$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)$ โดยที่ $\mathbf{x}_i$ เป็นเวกเตอร์ของข้อมูลเรียนรู้ใดๆ และ y_i เป็นค่าแทนประเภทของข้อมูล $\mathbf{x}_i$

โดยที่ $y_i = 1$ เมื่อ $\mathbf{x}_i$ เป็นประเภทบวก (Positive class)

$y_i = -1$ เมื่อ $\mathbf{x}_i$ เป็นประเภทลบ (Negative class)

ทั้งนี้สามารถยกตัวอย่างจากสมการเส้นตรงเพื่อพิจารณาสมการของ Hyperplane ดังนี้ กำหนดให้สมการเส้นตรง (Linear Equation) คือ

$$x_2 = 2x_1 + b \qquad \text{หรือ} \qquad x_2 - 2x_1 - b = 0$$

สามารถเขียนให้อยู่สมการไฮเปอร์เพลน (Equation of an hyperplane) ได้คือ

$$\mathbf{w}^T\mathbf{x} + (-b) = 0$$

โดยที่

$$\mathbf{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \qquad \text{และ} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\mathbf{w}^T\mathbf{x} + (-b) = (-2)\times x_1 + (1)\times x_2 - b = -2x_1 + x_2 - b = 0$$

กล่าวอีกนัยหนึ่ง $\mathbf{w}$ คือ เวกเตอร์หน้าสัมประสิทธิ์ของตัวแปร และ $\mathbf{x}$ คือเวกเตอร์ของตัวแปร และ b คือค่าคงที่ โดยสมการ Hyperplane $\mathbf{w}^T\mathbf{x} + (-b) = 0$ นี้ นำไปพิจารณาภายใต้เงื่อนไขที่สามารถจำแนกข้อมูลได้เป็น 2 คลาสที่เป็นประเภทบวก (Positive Class) นั่นคือ $\mathbf{w}^T\mathbf{x}_i + b = 1$ หรือเส้น Hyperplane H^+ และที่เป็นประเภทลบ (Negative Class) นั่นคือ $\mathbf{w}^T\mathbf{x}_i + b = -1$ หรือเส้น Hyperplane H^-

กรณีทั่วไปที่สามารถจำแนกเป็นเชิงเส้นได้ Hyperplane $f(\mathbf{x}_i)$ ใดๆ จะต้องอยู่ภายใต้เงื่อนไขดังนี้

$$\mathbf{w}^T\mathbf{x}_i + b \geq 1, \qquad y_i = 1$$

$$\mathbf{w}^T\mathbf{x}_i + b \leq -1, \qquad y_i = -1$$

นอกจากนี้ Optimal Separating Hyperplane สามารถแบ่งข้อมูลทั้งสองประเภทได้อย่างสมบูรณ์ หมายถึง ระยะห่างจาก Hyperplane ไปยังข้อมูลประเภท + จะต้องมีระยะที่ใกล้ที่สุด พร้อมๆ กับระยะห่างจาก Hyperplane ไปยังข้อมูลประเภท – ที่ใกล้ที่สุดเช่นกัน กล่าวคือ Margin ควรจะมีค่าที่มากที่สุดเท่าที่จะทำได้ (Maximal Margin)

ทั้งนี้หากนำ y_i หรือคลาสที่เป็นลบ (Negative class) นำไปคูณเข้าทั้งสองข้างของสมการ $\mathbf{w}^T\mathbf{x}_i + b \leq -1$ โดยที่ค่า y_i คือ (-1) จะได้

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq y_i(-1) \quad \text{หรือ} \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \quad \text{สำหรับ } \mathbf{x}_i \text{ ที่อยู่ใน Class -1}$$

เช่นเดียวกันหากนำ y_i หรือคลาสที่เป็นบวก (Positive class) นำไปคูณเข้าทั้งสองข้างของสมการ $\mathbf{w}^T\mathbf{x}_i + b \geq 1$ (โดยที่ค่า y_i คือ (1) จะได้

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq y_i(1) \quad \text{หรือ} \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \quad \text{สำหรับ } \mathbf{x}_i \text{ ที่อยู่ใน Class 1}$$

ทั้งนี้เงื่อนไขที่กำหนดข้างต้นหรือขอบเขต (Boundary) สามารถเขียนใหม่ได้ด้วยรูปแบบดังต่อไปนี้

$$\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b\right) y_i \geq 1 \qquad \text{สำหรับ } \mathrm{i} = 1, 2, \ldots, \mathrm{n}$$

ขณะที่ Magin ของ H^+ และ H^- สามารถแสดงได้คือ[3]

$$m = \frac{2}{\|\mathbf{w}\|}$$

ดังนั้น ภายใต้ฟังก์ชันการตัดสินใจ คือ การหา Optimal separating hyperplane โดยการหาค่าสูงสุดของ Margin ภายใต้เงื่อนไขขอบเขต (Boundary) สามารถแสดงได้จากปัญหาการหาค่าสูงสุดดังนี้

$$\text{Maximize: } \frac{2}{\|\mathbf{w}\|} \qquad \text{Subject to: } \left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b\right) y_i \geq 1$$

จากสมการข้างต้นเป็นการ Maximize Margin ภายใต้เงื่อนไขข้างต้น ทั้งนี้สามารถพิจารณาได้อีกรูปแบบหนึ่งคือการ Minimize ระยะทางของ w ภายใต้เงื่อนไขข้างต้น[4] ดังนี้

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\|^2 \qquad \text{Subject to: } \left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b\right) y_i \geq 1$$

ภายใต้ปัญหาการหาค่าที่เหมาะสมภายใต้เงื่อนไข (Constrained optimization problem) สามารถดำเนินการภายใต้การกำหนดสมการวัตถุประสงค์ (Objective Function) นั่นก็คือ การ Minimize ระยะทางของ $\|\mathbf{w}\|$ และนำLagrange Multiplier (α_i) คูณเข้ากับสมการเงื่อนไขซึ่งก็คือ $\left(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b\right) y_i \geq 1$ ตามวิธีการของ Lagrange Function ในรูปแบบ Primal Formulationได้ดังนี้[5]

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}.\mathbf{w} - \sum_{i=1}^{N} \alpha_i \left[y_i \left(\mathbf{w}.\mathbf{x}_i + b\right) - 1 \right] \tag{1}$$

โดยที่ α_i คือ Lagrange multipliers, $\alpha_i \geq 0$ ทำการหาค่าอนุพันธ์ของ Lagrange Function เทียบกับ (Take partial derivative with respect to) b, α_i และ $\mathbf{w}$ และกำหนดให้เท่ากับ 0 ภายใต้ First order condition จะได้

$$\frac{\partial L(\mathbf{w}, b, \alpha)}{\partial b} = -\sum_{i=1}^{N} \alpha_i y_i (\mathbf{w}, b) = 0 \tag{2}$$

$$\frac{\partial L(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i = 0 \tag{3}$$

หรือ สมการที่ 2 ก็คือ

$$\sum_{i=1}^{N} \alpha_i y_i (\mathbf{w}, b) = 0 \tag{4}$$

และสมการที่ 3 ก็คือ

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i \tag{5}$$

---

[3] $\|\mathbf{w}\|$ คือ Norm ซึ่งเป็นความยาวหรือขนาดของเวกเตอร์ w หรือ รากที่สองของ $w^T w$

[4] $\|\mathbf{w}\|^2 = \left(\sqrt{w.w}\right)^2 = w$

[5] อ้างอิงจาก Campbell, Colin. & Yiming Ying. (2011).

ทำการสมการที่ 4 และ 5 แทนค่าลงใน $L(\mathbf{w},b,\alpha)$ ซึ่งจะได้ dual formulation ซึ่งรู้จักกันคือ *Wolfe dual* ดังนี้

$$\begin{aligned} W(\alpha) &= \frac{1}{2}\sum_{i=1}^{N}\alpha_i y_i \mathbf{x}_i^T \sum_{j=1}^{N}\alpha_j y_j \mathbf{x}_j + \sum_{i=1}^{N}\alpha_i\left(1 - y_i\sum_{j=1}^{N}\alpha_j y_j \mathbf{x}_j^T \mathbf{x}_i + b\right) \\ &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_i \mathbf{x}_i^T \mathbf{x}_j + \sum_{i=1}^{N}\alpha_i - \sum_{i=1}^{N}\alpha_i y_i \sum_{j=1}^{N}\alpha_j y_j \mathbf{x}_j^T \mathbf{x}_i - b\sum_{i=1}^{N}\alpha_i y_i \\ &= -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_i \mathbf{x}_i^T \mathbf{x}_j + \sum_{i=1}^{N}\alpha_i \end{aligned}$$

$$W(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j (\mathbf{x}_i.\mathbf{x}_j) \tag{6}$$

และทำการหาค่าอนุพันธ์ $W(\alpha)$ เทียบ α ภายใต้เงื่อนไข

$$\alpha_i \geq 0 \quad \sum_{i=1}^{m}\alpha_i y_i = 0 \tag{7}$$

ทั้งนี้ Primal formulation เกี่ยวข้องกับปัญหาการหาค่าต่ำสุด (Minimization) ขณะที่ dual formulation เกี่ยวข้องกับการหาค่าสูงสุด (Maximization) นอกจากนี้ สมการที่ 6 คือ Quadratic สำหรับหาค่าพารามิเตอร์ α_i^* ซึ่งสามารถอธิบายได้จาก Quadratic programming (QP) ขณะที่สมการ 7 คือ Constrained quadratic programming ซึ่งสามารถหาค่าที่เหมาะสมของ α_i^* ได้จากการแก้ปัญหา Quadratic programming[6] เมื่อได้ α_i^* นำไปแทนในสมการที่ 4 และสมการที่ 5 และทำการหาค่าที่เหมาะสมของ $\mathbf{w}^*$ และ b ทั้งนี้กรณีที่ $\alpha_i > 0$ และสามารถหาผลลัพธ์ได้ภายใต้เงื่อนไขดังกล่าว จุดที่อยู่บนเส้น Hyperplane จะเรียกว่า Support Vector ซึ่งแสดงในภาพคือ x_1 x_2 และ x_3 และพิจารณาจากสมการการตัดสินใจ (Optimized decision function) $f(\mathbf{x}_i) = \text{sgn}(\mathbf{w}^T\mathbf{x}_i + b)$ จะสามารถเขียนได้อยู่ในรูปดังนี้

$$f(\mathbf{x}) = \text{sgn}(\mathbf{w}.\mathbf{x} + b) = \text{sgn}\left[\left(\sum_{i=1}^{N}\alpha_i y_i \mathbf{x}_i\right).\mathbf{x} + b\right]$$

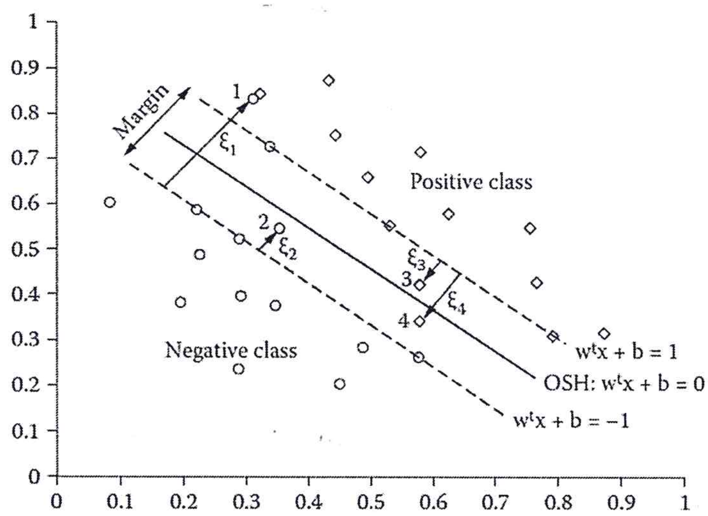
กรณีนี้จะเรียกอีกอย่างว่า Hard-Margin Support Vector Machines ซึ่งเป็นกรณีที่ Hyperplane สามารถจำแนกข้อมูลได้ออกเป็น 2 กลุ่มอย่างสมบูรณ์ โดยที่ไม่มีข้อมูลจุดใดที่ถูกจัดเป็นประเภท Positive Class ไปอยู่ใน Negative Class และเช่นเดียวกันในทางตรงกันข้าม

การคำนวณซัพพอร์ตเวกเตอร์แมชชีน (SVMs) เพื่อหา Hyperplane สำหรับกรณีที่ไม่สามารถจำแนกได้อย่างเป็นสมบูรณ์ (SVMs for linearly inseparable case)

[6] การแก้ไขปัญหา Quadratic programming (QP) สามารถใช้วิธีการจาก Sequential Minimal Optimization Algorithm: SMO Algorithm) ซึ่งถูกค้นพบจาก John Platt ในปี 1998

ในกรณีปกติทั่วไปจะพบว่าข้อมูลไม่สามารถจำแนกได้อย่างสมบูรณ์กล่าวคือ อาจมีข้อมูลที่จัดเป็นประเภท Positive Class ไปอยู่ใน Negative Class หรือข้อมูลที่จัดเป็นประเภท Negative Class ไปอยู่ใน Positive Class ดังที่แสดงในกรณีของ Hard margin support vector machines ซึ่งสามารถแสดงดังแผนภาพด้านล่าง

แผนภาพที่ 4 Hyperplane กรณีที่ไม่สามารถจำแนกได้อย่างเป็นสมบูรณ์



จากแผนภาพจะเห็นได้ชัดว่า กรณีดังกล่าวมีความเป็นไปได้ยากที่จะสามารถสร้าง Hyperplane ที่สามารถจำแนกได้ออกเป็น 2 คลาสโดยปราศจากความคาดเคลื่อนเลย นำมาซึ่งการหา Hyperplane ที่กำหนดให้มีค่าความคลาดเคลื่อนร่วมด้วยซึ่งเรียกว่า Soft margin support vector machines

1) Slack variable-based "Soft margin" technique

วิธีการนี้เป็นการเพิ่มตัวแปร เพื่อที่จะหา Optimal separating hyperplane โดยพิจารณาความคลาดเคลื่อนร่วมด้วย ดังนั้นเงื่อนไขหรือขอบเขตใหม่สามารถแสดงได้ดังนี้

$$(\mathbf{w}.\mathbf{x}_i + b)y_i \geq 1 - \xi_i, \quad \xi_i \geq 0, \; i = 1, 2, ..., N$$

โดยที่ ξ_i คือ slack variable ที่เป็นตัววัดความระยะห่างของข้อมูลที่ถูกจัดผิดประเภทกับขอบเขตของ Hyperplane ดังนั้นปัญหาค่าที่เหมาะสมสามารถแสดงได้ดังนี้

$$\text{Minimize: } \frac{1}{2}\|\mathbf{w}\| + C\sum_{i=1}^{N}\xi_i$$

$$\text{Subject to: } (\mathbf{w}.\mathbf{x}_i + b)y_i \geq 1 - \xi_i, \quad \xi_i \geq 0$$

โดยที่ C คือ penalizing factor ที่กำหนด trade-off ระหว่างความคาดเคลื่อนของ training set (training error) แล Margin นั่นเอง

2) Kernel Function-based nonlinear mapping

ดังที่กล่าวไว้ข้างต้นว่า การแปลงข้อมูลให้มีมิติที่สูงขึ้นหรือที่เรียกว่า feature space เป็นวิธีหนึ่งที่จะสามารถหา optimal separating hyperplane ได้ ซึ่งวิธีการดังกล่าวอาศัย Kernel Function จากตัวอย่างข้างต้น $\mathbf{x}_i$ เป็นจุดข้อมูล (data points) ซึ่งปรากฏใน inner product[7] หากข้อมูลอยู่ในปริภูมิที่มีมิติสูงกว่า ที่เรียกว่า Feature Space เราสามารถ mapping ได้ดังนี้

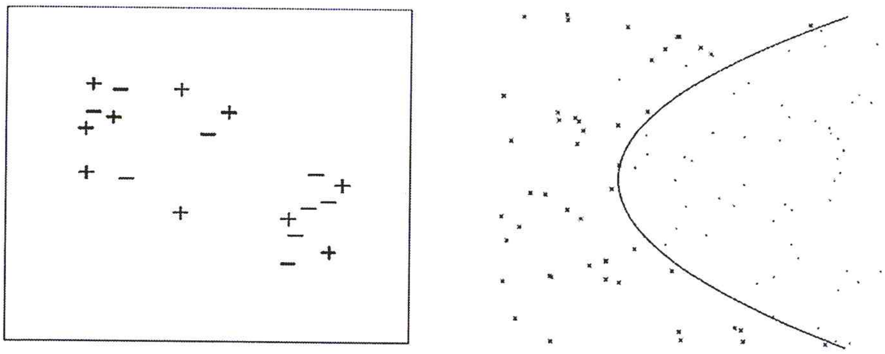
$$\mathbf{x}_i.\mathbf{x}_j \rightarrow \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)$$

โดยที่ $\Phi(.)$ คือ mapping function โดย $\Phi(.)$ จะขึ้นอยู่กับการกำหนดฟังก์ชันของ Kernel (Kernel function)

เคอร์เนล (Kenel)

เมื่อข้อมูลเรียนรู้มีการวางตัวเรียงกันในลักษณะตำแหน่งที่ไม่สามารถแบ่งกลุ่มได้โดยเส้นตรงตามวิธีของ SVM ลักษณะเชิงเส้น การแก้ไขปัญหาทำได้โดยการเปลี่ยนมิติของข้อมูลให้สูงขึ้นเพื่อช่วยในการเรียงตัวของข้อมูลใหม่ที่เรียกว่า Higher Dimensional Space

แผนภาพที่ 5 การเปลี่ยนมิติของข้อมูลให้สูงขึ้น (Higher Dimensional Space)



จากแผนภาพข้างต้น หากใช้ Linear SVM ในการจำแนกประเภทของข้อมูลจะไม่สาสามารถทำได้ ทั้งนี้การเปลี่ยนมิติให้สูงขึ้นตาม $\Phi(.)$ จะสามารถทำการจำแนกประเภทข้อมูลได้ดังภาพด้านขวามือ ซึ่งการหาฟังก์ชัน $\Phi(.)$ ได้จากการเลือกของ Kernel:

$$K(\mathbf{x}_i.\mathbf{x}_j) = \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j) \qquad i, j = 1,2,3,\ldots,\mathrm{N}$$

ดังนั้นปัญหาค่าที่เหมาะสมสามารถแสดงได้ดังนี้

Maximize: $W(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \left(\Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)\right)$

Subject to: $0 \le \alpha_i \le C \qquad i, j = 1,2,3,\ldots,\mathrm{N}$

---

[7] Inner product หมายถึงการคูณภายในของเวกเตอร์ ใน $\mathbb{R}^n$ โดยที่ เวกเตอร์ $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ และ $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ โดยที่ผลคูณภายใน (Euclidean inner product) ของ $\mathbf{u}$ และ $\mathbf{v}$ เขียนแทนด้วย $\mathbf{u}.\mathbf{v} = \sum_{i=1}^{n} u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$

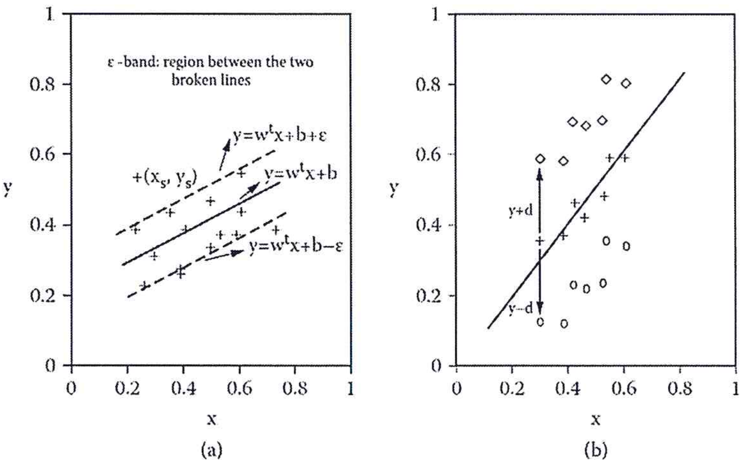
โดยที่ C คือ penalizing factor และ $\Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)$ คือ Kernel Function

เนื่องด้วย Kernel เป็นการคูณกันภายใน (Inner product) ระหว่างเวกเตอร์ $\mathbf{x}_i$ กับ $\mathbf{x}_j$ ในรูปแบบของฟังก์ชันต่างๆ ซึ่งโดยทั่วไปที่นิยมใช้กันมีอยู่ 3 ชนิดคือ 1) โพลิโนเมียล (Polynomial) 2) ซิกมอยด์ (Sigmoid) และ 3) เรเดียลเบสิสฟังก์ชัน (Radial Basis Function: RBF) หรือเรียกว่า Gaussian Kernel ซึ่งการศึกษานี้ใช้แบบเรเดียลเบสิสฟังก์ชัน

$$K(\mathbf{x}_i.\mathbf{x}_j) = \exp\left(-\frac{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2}{2\sigma^2}\right)$$

วิธีการของ SMVs สามารถนำไปใช้ได้กับการแก้ปัญหาในรีเกรสชั่น โดยกำหนดให้เซตของข้อมูลคือ $D = \{(x_i, y_i)\}_{i=1}^{N}$ โดยที่ x_i คือ sample vector ซึ่งสอดคล้องกับ y_i และ N คือจำนวนข้อมูล ซึ่งกรอบแนวคิดพื้นฐานที่นำมาประยุกต์ใช้คือ ε-band และ ε-insensitive loss function ดังแสดงรายละเอียดดังแผนภาพด้านล่างดังนี้

แผนภาพที่ 6 การประยุกต์ SVMs สำหรับรีเกรสชั่น



กำหนดให้ฟังก์ชันหนึ่งตัวแปรในแผนภาพ (a) ε-band หมายถึงขนาดหรือระยะทางระหว่างเส้นประซึ่งสามารถหาได้จากการเคลื่อนของ ε-band ขึ้นและลง ขณะที่แผนภาพ (b) เป็นการเปลี่ยนจากปัญหา Regression มาสู่ปัญหาการจำแนกรูปแบบ กล่าวคือแต่ละ $\mathbf{x}_i$ ซึ่งสอดคล้องกับ y_i ที่ถูกบวกด้วย “d” และเกิด $(\mathbf{x}_i', y_i^1)$ ซึ่งจำแนกได้เป็น Class ที่ 1 ขณะที่แต่ละ $\mathbf{x}_i$ ซึ่งสอดคล้องกับ y_i ที่ถูกลบด้วย “d” และเกิด new sample $(\mathbf{x}_i'', y_i^{-1})$ ซึ่งจำแนกได้เป็น Class ที่ 2

ทั้งนี้ ε-insensitive loss function สามารถแสดงได้ดังนี้

$$L(y_i, y_i^* | \varepsilon) = \begin{cases} |y_i^* - y_i| - \varepsilon, & if\, |y_i^* - y_i| \geq \varepsilon \\ 0, & otherwise \end{cases}$$

โดยที่ y_i^* คือค่าประมาณ จากปัญหารีเกรสชั่น สามารถนำไปสู่ ε-insensitive loss function ได้ดังนี้

$$\text{Minimize:}\quad \frac{1}{2}\|\mathbf{w}\|^2+\frac{C}{N}\sum_{i=1}^{N}L\left(y_i,y_i^*\middle|\varepsilon\right)$$

โดยที่ C คือ penalizing factor ที่กำหนด จะสามารถเขียนอีกในรูปแบบหนึ่งในรูปของ slack variable ξ_i และ ξ_i^* โดยที่ ξ_i คือ ค่าความคลาดเคลื่อนที่มากกว่า ε และ ξ_i^* คือ ค่าความคลาดเคลื่อนที่น้อยกว่า ε ได้ดังนี้

$$\text{Minimize:}\quad L\left(\mathbf{w},b,\xi_i,\xi^*\right)(\frac{1}{2}\|\mathbf{w}\|^2+\frac{C}{N}\sum_{i=1}^{N}\left(\xi_i+\xi_i^*\right)$$

$$\text{Subject to:}\quad \left(\mathbf{w}.\mathbf{x}_i+b\right)-y_i\le\varepsilon+\xi_i^*,\qquad i,j=1,2,3,\ldots,\mathrm{N}$$

$$y_i-\left(\mathbf{w}.\mathbf{x}_i+b\right)\le\varepsilon+\xi_i,\qquad i,j=1,2,3,\ldots,\mathrm{N}$$

$$\xi_i^*\ge 0,\ i,j=1,2,3,\ldots,\mathrm{N}$$

ทั้งนี้การหาผลเฉลยสามารถหาได้จาก Wolfe dual และแก้ไขปัญหา Quadratic programming (QP) จาก Sequential Minimal Optimization Algorithm: SMO Algorithm) John Platt ในปี 1998

ผลการศึกษา

การทดสอบ Stationary ของข้อมูลอัตราแลกเปลี่ยนสกุลเงินบาทต่อสกุลเงินดอลลาร์สหรัฐ พบว่า ที่ระดับ Level เกิดความไม่นิ่งของข้อมูล (Non-stationary) ขณะที่ทดสอบข้อมูลอัตราผลตอบแทนดังกล่าวที่ระดับ Level เกิดความนิ่งของข้อมูล (Stationary) ดังนั้นการศึกษานี้จึงใช้ข้อมูลผลตอบแทนดังกล่าวแทนการใช้ข้อมูลอัตราแลกเปลี่ยนเพื่อไม่ให้เกิดปัญหาสมการถดถอยที่ไม่แท้จริง (Spurious Regression)

ตารางที่ 1 สรุปเปรียบเทียบผลการพยากรณ์ของ BHT/USD ในแบบจำลองต่างๆ

BHT/USD		Testing Data (10%,141 วัน)		Testing Data (20%, 283 วัน)	
		RMSE	Mean Directional Accuracy	RMSE	Mean Direction Accuracy
(1) SVM Model	R_t = f (R_{t-1})	0.006391619	51.063%	0.005309446	49.469%
	R_t = f (R_{t-1}, R_{t-2}, R_{t-3})	0.004957886	55.319%	0.005697685	49.469%
	R_t = f (R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, R_{t-5})	0.004826151	46.808%	0.005338232	46.996%
	R_t = f (R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4} ,R_{t-5}, R_{t-6}, R_{t-7})	0.006038214	49.645%	0.005811392	47.349%
	R_t = f (R_{t-1}, R_{t-2}, R_{t-3}, R_{t-4}, ,R_{t-5}, R_{t-6}, $R_{t-7,\ldots}$, R_{t-15})	0.004764158	51.773%	0.005150985	51.943%

BHT/USD	Testing Data (10%,141 วัน)		Testing Data (20%, 283 วัน)	
	RMSE	Mean Directional Accuracy	RMSE	Mean Direction Accuracy
(2) Random Walk Model	0.006282999	46.099%	0.007146612	45.936%
(3.1) ARIMA(2,0,2)	0.004489412	49.645%		
(3.2) ARIMA(2,0,1)			0.004912313	49.823%

หากพิจารณาในกรณีแรกที่กำหนด Testing Data/Out-of-sample forecasting เป็นร้อยละ 10 ของข้อมูลทั้งหมด (141 วันหรือประมาณ 4.5 เดือน) พบว่า เมื่อพิจารณาจาก RMSE จะพบว่า แบบจำลอง SVM มีค่า RMSE ต่ำกว่าแบบจำลอง Random Walk แต่แบบจำลอง ARIMA จะให้ค่า RMSE ต่ำที่สุด อย่างไรก็ตามหากพิจารณาจากความถูกต้องของทิศทางในการคาดการณ์ (Directional accuracy) แบบจำลอง SVM ทุกแบบจำลองจะสามารถพยากรณ์ได้แม่นยำกว่าแบบจำลอง ARIMA และแบบจำลอง Random Walk ซึ่งแบบจำลอง SVM ที่มีความถูกต้องของการพยากรณ์ทิศทางมากที่สุด คือ แบบจำลอง SVM ณ ช่วงเวลาที่ผ่านมา 1-3 วัน (w_3)

ในกรณีแรกที่กำหนด Testing Data/Out-of-sample forecasting เป็นร้อยละ 20 ของข้อมูลทั้งหมด (283 วันหรือประมาณ 9.5 เดือน) พบว่า เมื่อพิจารณาจาก RMSE จะพบว่า แบบจำลอง ARIMA จะมีค่า RMSE ต่ำที่สุด รองลงมาคือแบบจำลอง SVM และแบบจำลอง Random Walk อย่างไรก็ตามหากพิจารณาจากความถูกต้องของทิศทางในการคาดการณ์ (Mean Directional Accuracy) แบบจำลอง SVM ณ ช่วงเวลาที่ผ่านมา 1-15 วัน (w_{15}) จะให้ค่าความถูกต้องของการพยากรณ์ทิศทางมากที่สุดและมากกว่าแบบจำลอง Random walk ในทุกกรณี

ข้อเสนอแนะ

การศึกษาวิธีการพยากรณ์ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน มีการกำหนดค่าพารามิเตอร์ในฟังก์ชันต่างๆ เพื่อหาค่าที่เหมาะสมที่สุด ซึ่งการกำหนดค่าดังกล่าวจำเป็นต้องอ้างอิงจากงานวิจัยที่เกี่ยวข้อง โดยมีการกำหนดที่หลากหลายช่วง ซึ่งผลการศึกษาจึงขึ้นอยู่กับค่าพารามิเตอร์นั้นด้วยเช่นกัน ทั้งนี้ในการศึกษาครั้งต่อไป สามารถนำเทคนิคซัพพอร์ตเวกเตอร์แมชชีนไปใช้กับแบบจำลองอื่นๆ เช่น SVM-GRARCH ,LS-SVM เพื่อให้การพยากรณ์ตัวแปรต่างๆ มีความแม่นยำมากยิ่งขึ้น

เอกสารอ้างอิง

Adhikari, Ratnadip. & R. K. Agrawal. (2013). **An Introductory Study on Time Series Modeling and Forecasting.**

Alamili, Mohamand. (2011). **Exchange Rate Prediction Using Support Vector Machines: A Comparison with Artificial Neural Network.** Master's Thesis. Delft University of Technology.

Blaskowitz, O., & Herwartz, H. (2011). On economic evaluation of directional forecasts. **International Journal of Forecasting,** 27(4), 1058-1065.

Campbell, Colin. & Yiming Ying. (2011). **Learning with Support Vector Machines.** Morgan & Claypool Publishers series in Synthesis Lectures on Artificial Intelligence and Machine Learning.

Cao, Lijuan. (2003). Support Vector Machines Experts for Time Series Forecasting. **Neurocomputing** 51 (2003) 321 – 339.

Dean, Jared. (2014). **Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners.** John Wiley & Sons, Inc., Canada.

Dumitru, Ciobanu. & Bar Mary Violeta. (2013). On the Prediction of Exchange Rate Dollar/Euro with an SVM Model. **Revista Economica** 65:2 (2013).

Fletcher, Tristan. (2012). **Machine Learning for Financial Market Prediction.** Ph.D Thesis, University College London.

Guo, Xiaopeng., DaCheng Li., & Anhui Zhang (2012). Improved Support Vector Machine Oil Price Forecast Model Based on Genetic Algorithm Optimization Parameters. **AASRI Procedia.** Volume 1, 2012, Pages 525-530.

Huang, Shian-Chang. et al. (2010). Chaos-based Support Vector Regressions for Exchange Rate Forecasting. **Expert Systems with Applications** 37 (2010) 8590–8598.

James, Gareth. et al. (2014). **An Introduction to Statistical Learning with Application in R.** Springer, New York.

José Antonio Sanabria Montañez. (2011). **A Contribution to Exchange Rate Forecasting Based on Machine Learning Techniques.** Doctoral Thesis, Universitat Ramon Llull.

Kim, Kyoung-jae. (2003). Financial Time Series Forecasting Using Support Vector Machines. **Neurocomputing.** 55, 307-309.

Liang, Yizeng., et al. (2011). **Support Vector Machines and Their Apllication in Chmistry and Biotechnology.** CRC Press, Taylor and Francis Group, USA.

Lo, Andrew W. & MacKinlay, A. Craig. (1999). **A Non-Random Walk Down Wall Street.** Princeton University Press, Princeton New Jersey,

Nur Izzati Abdul Rashid. et al. (2016). Exchange Rate Forecasting Using Modified Empirical Mode Decomposition and Least Squares Support Vector Machine. **International Journal of Advances in Soft Computing and Its Application.** Vol. 8, No. 3, December 2016.

Pérez-Cruz, Fernando. et al. (2003). Estimating GARCH Models Using Support Vector Machines. Research Paper. **Quantitative Finance.** Volume 3 (2003) 1–10.

Plakandaras, Bill. (2015). Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques. **Journal of Forecasting**. April 2015.

Platt, John C. (1988). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. **Microsoft Research,** Technical Report MSR-TR-98-14.

Shafiee, Mahboobeh. et al. (2013). Forecasting Stock Returns using Support Vector Machine and Decision Tree: a Case Study in Iran Stock Exchange. **International Journal of Economy, Management and Social Sciences,** 2(9) September 2013, 746-751.

Sheta, Alan F. et al. (2015). A Comparison between Regression, Artificial Neural Networks and Support Vector Machine for Predicting Stock Market Index. **International Journal of Advanced Research in Artificial Intelligence (IJARAI).** Vol. 4, No.7, 2015.

Soulas, Eleftherios. & Dennis Shasha. (2013). Online Machine Learning Algorithms for Currency Exchange Prediction.

Theofilatos, Konstantinos. et al. (2012). Modeling and Trading the EUR/USD Exchange Rate Using Machine Learning Techniques. **ETASR - Engineering, Technology & Applied Science Research.** Vol. 2, No. 5, 2012, 269-272.

Ullrich, Christian. (2009). **Forecasting and Hedging in the Foreign Exchange Markets.** Springer-Verlag Berlin Heidelberg.

Witten, Ian H., & Eibe Frank. (2005) **Data Mining: Practical Machine Learning Tools and Techniques.** Elsevier Inc, USA.

Xhaja, Denada., Ana Ktona., & Gazmira Brahusi. (2014). Currency Exchange Rate Forecasting Using Machine Learning Techniques. **International Journal of**

Computer and Information Technology, Volume 03 – Issue 05, September 2014.

Yusof, Yuhanis. & Zuriani Mustaffa. (2016). A Review on Optimization of Least Squares Support Vector Machine for Time Series Forecasting. **International Journal of Artificial Intelligence & Applications** (IJAIA), Vol. 7, No. 2, March 2016.