

การพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตโดยใช้วิธี
วิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล

Forecasting the amount of Internet Banking Using Time Series Data

Mining Techniques

ณรงค์ฤทธิ์ สุคนธสิงห์

Narongrit Sukonthasing

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

38 ถนนเพชรเกษม แขวงบางหว้า เขตภาษีเจริญ กรุงเทพมหานคร 10160

Email: narongrit_rit@hotmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านบริการธนาคารบนอินเทอร์เน็ตของธนาคารพาณิชย์ในประเทศไทย โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล 3 เทคนิค ได้แก่ 1) การถดถอยเชิงเส้น 2) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น และ 3) ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย ข้อมูลอนุกรมเวลาที่นำมาใช้ในการศึกษาเป็นข้อมูลปริมาณการชำระเงินผ่านบริการธนาคารบนอินเทอร์เน็ต ตั้งแต่ปี 2553 ถึง 2558

จากผลการเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านบริการธนาคารบนอินเทอร์เน็ตกับชุดข้อมูลในปี 2558 พบว่า แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น มีประสิทธิภาพโดยรวมสูงที่สุด โดยมีค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ เท่ากับ 8.70% ลำดับถัดมาได้แก่ แบบจำลองการพยากรณ์ด้วยเทคนิควิธีการถดถอยเชิงเส้น โดยมีค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ เท่ากับ 16.75%

คำสำคัญ: การพยากรณ์ ธนาคารบนอินเทอร์เน็ต การวิเคราะห์อนุกรมเวลา เทคนิควิธีเหมืองข้อมูล

Abstract

The purposes of this research were to develop the model of forecasting the amount of internet banking using the time series data mining techniques by 3 methods, as follows; 1) Linear Regression, 2) Multi-Layer Perceptron and 3) Support Vector Machine for Regression. The time series data for study was the amount of internet banking from year 2010 to 2015

The comparing the effectiveness of forecasting's model of amount of internet banking with time series testing data in 2015 found that the multilayer perceptron was the highest performance model for forecasting at 8.70% of MMRE, next, the linear regression forecasting was at 16.75% of MMRE.

Keywords: Forecasting, Internet Banking, Time Series Analysis, Data Mining technique

บทนำ

ปัจจุบันเทคโนโลยีอินเทอร์เน็ตเข้ามามีบทบาทต่อชีวิตประจำวันของประชากรจำนวนมาก และได้ขยายวงกว้างออกไปมากขึ้น เทคโนโลยีอินเทอร์เน็ตได้เข้าไปมีบทบาทในทุกสาขาอาชีพ โดยมีการนำเทคโนโลยีอินเทอร์เน็ตมาใช้เพื่อประโยชน์ต่อองค์กรในรูปแบบต่างๆ เช่น การศึกษา การทำธุรกิจ การประชาสัมพันธ์ การโฆษณาสินค้า การค้าขาย การติดต่อสื่อสาร เป็นต้น เทคโนโลยีอินเทอร์เน็ตส่งผลให้เกิดการเปลี่ยนแปลงหลายอย่างต่อบุคคลและสังคม เช่น วิธีการติดต่อสื่อสาร วิธีการหาความรู้ วิธีการทำธุรกิจ วิธีการชำระเงิน เป็นต้น สถาบันการเงินเป็นหน่วยงานหนึ่งในสังคมที่มีบริการระบบการชำระเงินที่มีการพัฒนามาโดยลำดับ การพัฒนาดังกล่าวมีความก้าวหน้ามากขึ้น โดยเฉพาะอย่างยิ่งในสังคมเมือง ซึ่งมีโครงสร้างพื้นฐาน เช่น ระบบสื่อสาร โทรคมนาคมที่ทันสมัย ประชากรมีฐานะทางเศรษฐกิจที่ดีและมีจำนวนผู้ใช้บริการที่มากพอที่จะคุ้มค่าการลงทุนในระบบและอุปกรณ์ต่างๆ ระบบการชำระเงิน เช่น บัตรเครดิต บัตรเดบิต บัตรสมาร์ทการ์ด เป็นต้น ธุรกิจการชำระเงินผ่านบริการ Mobile banking และ Internet banking พบว่า จำนวนบัญชีลูกค้าที่ใช้บริการ, ปริมาณรายการที่ทำธุรกรรม, และมูลค่ารายการ มีจำนวนเพิ่มขึ้นอย่างต่อเนื่อง (สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์, 2559) โดยเฉพาะในกลุ่มประชากรที่มีรายได้แน่นอนในเมืองใหญ่ ธนาคารอิเล็กทรอนิกส์ เกี่ยวกับการให้บริการทำธุรกรรมทางการเงินต่างๆ ผ่านอุปกรณ์หรือระบบอิเล็กทรอนิกส์ เช่น โทรศัพท์มือถือ อินเทอร์เน็ต เป็นต้น มีการให้บริการ เช่น การฝากเงิน ถอนเงิน โอนเงิน สอบถามยอดเงิน เป็นต้น มีผู้ใช้บริการธนาคารอิเล็กทรอนิกส์ ที่เพิ่มมากขึ้นอย่างต่อเนื่อง เนื่องจากธนาคารอิเล็กทรอนิกส์ทำให้เกิดความรวดเร็วและสะดวกสบายในการทำธุรกรรมมากขึ้นและประหยัดทรัพยากร ลักษณะการให้บริการลูกค้า สามารถชำระค่าบริการต่างๆ ได้ทุกวัน ตลอด 24 ชม. โดยไม่ต้องเดินทาง เช่น ค่าโทรศัพท์มือถือ ค่าบัตรเครดิต ค่าใช้บริการอินเทอร์เน็ต ผ่อนชำระค่าบ้าน ค่าเช่าซื้อรถยนต์ เป็นต้น

งานวิจัยชิ้นนี้ผู้วิจัยเลือกชุดข้อมูล ธุรกิจการชำระเงินผ่านบริการ Internet banking ในช่วงปี พ.ศ. 2553 ถึง 2558 ซึ่งเป็นชุดข้อมูลที่ทางรัฐบาลได้นำออกมาเผยแพร่ ผ่านทางเว็บไซต์ www.data.go.th คือเว็บไซต์ที่บริหารจัดการโดยกระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร โดยสำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) เป็นเว็บไซต์ศูนย์กลางชุดข้อมูลเปิดที่หน่วยงานภาครัฐได้ทำการรวบรวมจากหน่วยงานต่างๆ นำออกมาเผยแพร่ซึ่งมีหมวดหมู่ชุดข้อมูลทั้งหมด 17 หมวดหมู่และชุดข้อมูลทุกชุดจะต้องมีคำอธิบายและอยู่ในรูปแบบที่กำหนด โดยมีวัตถุประสงค์หลักเพื่อให้ผู้ใช้บริการไม่ว่าจะเป็นประชาชนทั่วไปตลอดจนหน่วยงานภาครัฐและเอกชน สามารถสืบค้นและเข้าถึงข้อมูลที่มีคุณภาพ สะดวก รวดเร็ว และสามารถนำไปใช้ประโยชน์เพื่อสร้างนวัตกรรมใหม่และมีส่วนร่วมในการปรับปรุงชุดข้อมูลให้มีคุณภาพมากยิ่งขึ้น โดยการแสดงความคิดเห็นและการร้องขอข้อมูลที่ต้องการ www.data.go.th ขึ้นตามมติคณะรัฐมนตรีปี 2556 เพื่อยกระดับการให้บริการประชาชน และหน่วยงานภาครัฐและเอกชน ลดความเหลื่อมล้ำทางสังคมและรองรับการเปลี่ยนเข้าสู่ประชาคมอาเซียน (สำนักงานรัฐบาลอิเล็กทรอนิกส์, 2559)

งานวิจัยนี้เลือกใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) 3 เทคนิค ได้แก่ การถดถอยเชิงเส้น (Linear Regression) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) และซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) โดยผลลัพธ์ของการประมวลผลข้อมูลอนุกรมเวลาจะอยู่ในรูปของแบบจำลองการพยากรณ์ธุรกรรมการเงินผ่านบริการ Internet banking และสามารถนำแบบจำลองที่ได้ไปทำนายธุรกรรมการเงินผ่านบริการ Internet banking เพื่อให้สถาบันการเงินนำไปใช้ในการวางแผนบริหารและจัดการเกี่ยวกับฮาร์ดแวร์และซอฟต์แวร์ที่เกี่ยวข้องกับบริการผ่าน Internet banking ได้อย่างมีประสิทธิภาพต่อไป

วัตถุประสงค์

1. เพื่อเปรียบเทียบประสิทธิภาพ และเลือกเทคนิควิธีที่เหมาะสม ในการวิเคราะห์อนุกรมเวลาของมูลค่าธุรกรรมการเงินผ่านอินเทอร์เน็ต
2. เพื่อสร้างรูปแบบการพยากรณ์มูลค่าธุรกรรมทางการเงินผ่านอินเทอร์เน็ต โดยวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล

เอกสารและงานวิจัยที่เกี่ยวข้อง

ธนาคารอิเล็กทรอนิกส์ (Internet banking)

ธนาคารอิเล็กทรอนิกส์ หมายถึง การประกอบธุรกิจการพาณิชย์อิเล็กทรอนิกส์ เกี่ยวกับการให้บริการทำธุรกรรมทางการเงินต่างๆ ผ่านอุปกรณ์หรือระบบอิเล็กทรอนิกส์ เช่น โทรศัพท์มือถือ หรืออินเทอร์เน็ต มีการให้บริการเช่น การฝากเงิน ถอนเงิน โอนเงิน หรือ สอบถามยอดเงิน เป็นต้น โดยในอนาคตการให้บริการของธนาคารอิเล็กทรอนิกส์ ยังสามารถพัฒนาได้อีกเรื่อยๆ เพื่อรองรับความต้องการในการใช้บริการของผู้ใช้บริการธนาคารอิเล็กทรอนิกส์ ที่เพิ่มมากขึ้นอย่างต่อเนื่อง เนื่องจากธนาคารอิเล็กทรอนิกส์ทำให้เกิดความรวดเร็วและสะดวกสบายในการทำธุรกรรมมากขึ้นและประหยัดทรัพยากร (Wikipedia,2559)

ข้อมูลอนุกรมเวลา (Time Series Data)

ข้อมูลอนุกรมเวลา หมายถึง ชุดข้อมูลที่รวบรวมและจัดเก็บตามลำดับ ต่อเนื่องกันภายใต้การเพิ่มขึ้นของเวลา ซึ่งข้อมูลจะต้องถูกเก็บรวบรวมอย่างละเอียดและต่อเนื่องเป็นระยะเวลาติดต่อกันตามช่วงเวลาที่ต้องการ เช่น ข้อมูลปริมาณการใช้บัตรเครดิตของสถาบันการเงิน ข้อมูลปริมาณธุรกรรมการชำระเงินผ่านบริการ Mobile banking เป็นต้น ซึ่งจะกำหนดเวลาในการจัดเก็บเริ่มต้น และกำหนดเวลาในการจัดเก็บสิ้นสุดเท่า ๆ กันทุกปี การจัดเก็บข้อมูลอนุกรมเวลามีวัตถุประสงค์สำคัญเพื่อการสร้างแบบจำลองในการทำนายปริมาณที่จะเกิดขึ้นในอนาคต (Time Series Model) โดยการวิเคราะห์อนุกรมเวลา (Time Series Analysis) (Hoshmand, 2009)

การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์อนุกรมเวลา หมายถึง การทำนายปริมาณข้อมูลที่จะเกิดขึ้นในอนาคตซึ่งอาศัยข้อมูลอนุกรมเวลาในอดีตเป็นเทคนิควิธีที่ได้รับความนิยม เป็นการศึกษาถึงความสัมพันธ์ระหว่างข้อมูลกับช่วงเวลาที่เกี่ยวข้อง ช่วงเวลาอาจหมายถึง วัน เดือน ไตรมาส หรือปี อย่างไรก็ตาม การวิเคราะห์อนุกรมเวลาเพื่อการพยากรณ์ปริมาณข้อมูลที่จะเกิดขึ้นในอนาคต จำเป็นต้องวิเคราะห์ถึงความสัมพันธ์ต่าง ๆ ที่ขึ้นต่อกันของข้อมูล เช่น ส่วนของแนวโน้ม (Trend Component) ส่วนของฤดูกาล (Seasonal Component) ส่วนของวัฏจักร (Cyclical Component) และ ส่วนของเหตุการณ์ผิดปกติ (Irregular Component) จุดมุ่งหมายที่สำคัญของการวิเคราะห์อนุกรมเวลา คือ วิเคราะห์หาตัวแปรที่เหมาะสมกับการสร้างแบบจำลองในการทำนาย (Time Series Model) ซึ่งค่าพยากรณ์ดังกล่าวมีความเกี่ยวข้องกับข้อมูลในอดีต (Hoshmand, 2009)

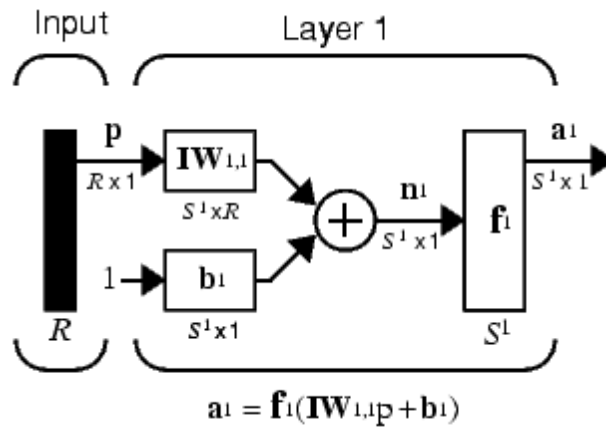
การวิเคราะห์การถดถอย (Linear Regression)

การวิเคราะห์การถดถอยเป็นวิธีการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ซึ่งเทคนิคนี้จะอาศัยความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรเพื่อใช้ในการทำนาย โดยลักษณะของสมการ Linear regression จะแสดงให้เห็นถึงค่าความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระในลักษณะของข้อมูลเชิงปริมาณ โดยสามารถเขียนได้ดังนี้ (Cai, Hall, & others, 2006)

$$Y = a + bx_1 + cx_2 + \dots \quad (1)$$

โครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multi-Layer Perceptron: MLP)

โครงข่ายประสาทเทียมแบบหลายชั้น คือ ลักษณะการเชื่อมต่อแบบโยไปข้างหน้าแบบทั่วถึง (Fully connected feed-forward nets) สามารถมีจำนวนชั้นตั้งแต่หนึ่งชั้นขึ้นไป (Frias-Martinez, Sanchez, & Velez, 2006) ถูกพัฒนาขึ้นมาเพื่อแก้จุดอ่อนของโครงข่ายประสาทเทียมแบบชั้นเดียว (Single layer perceptron) ให้ความสามารถในการคำนวณที่สูงขึ้น โดยโครงข่ายประสาทเทียมแบบหลายชั้นจะประกอบด้วย ชั้นข้อมูลเข้า (Input layer) ตามจำนวนแอตริบิวต์ของชุดข้อมูล ชั้นซ่อน (Hidden Layer) ความเหมาะสมของการกำหนดขึ้นอยู่กับทดสอบประสิทธิภาพ และชั้นผลลัพธ์ (Output Layer)



รูปที่ 1 โครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multi-Layer Machine for Regression)

ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression)

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine for Regression: SVM) เป็นอัลกอริทึมที่ใช้ในการจำแนกกลุ่มข้อมูล ด้วยวิธีการหาระนาบการตัดสินใจ (Decision Hyperplane) หรือไฮเปอร์เพลนที่เหมาะสม สำหรับการแบ่งข้อมูล 2 ส่วนจากกัน SVM ยังถูกนำมาประยุกต์ใช้ในการสร้างสมการในประมาณการค่าฟังก์ชันเชิงเส้น $f(x)$ ที่ใช้แทนระนาบตัดสินใจ ซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย หรือ SVR จะเป็นการนำข้อมูลปัจจุบันและข้อมูลในอดีตจำนวนหนึ่งมาเรียนรู้ (Training) เพื่อให้ทราบถึงรูปแบบสำหรับคาดการณ์ผลซึ่งจะเกิดขึ้นในอนาคต (Shevade, Keerthi, Bhattachryya, & Murthy, 2000; Smola & Scholkopf, 2004)

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มุ่งเน้นการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยดำเนินการวิจัยตามขั้นตอน 3 ขั้นตอนดังนี้

ศึกษาปัญหาและวิเคราะห์ข้อมูล

การพยากรณ์แนวโน้มปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตเป็นวิธีศึกษาปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตในอนาคต ทั้งนี้เพื่อที่จะนำผลการพยากรณ์มาใช้ในการวางแผนส่งเสริมการใช้บริการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตของสถาบันการเงิน หากทราบปริมาณการชำระเงินล่วงหน้า นอกจากนี้ยังนำมาใช้วางแผนยุทธศาสตร์ด้านเทคโนโลยีสารสนเทศ เพื่อรองรับปริมาณการใช้บริการที่เกิดขึ้นในอนาคต ผู้วิจัยจึงพัฒนางานวิจัยนี้ขึ้นเพื่อวัตถุประสงค์หลักในการสร้างแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต เพื่อสร้างเป็นรูปแบบการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตด้วยวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล และเพื่อเปรียบเทียบประสิทธิภาพ และเลือกเทคนิควิธีที่เหมาะสมในการวิเคราะห์ข้อมูลอนุกรมเวลาของปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ชุดข้อมูลที่ผู้วิจัยนำมาใช้เป็นชุดข้อมูล

ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยแบ่งเป็นรายเดือนในช่วงปี พ.ศ. 2553 ถึง ปี พ.ศ. 2558 ที่รัฐบาลเผยแพร่ข้อมูลผ่านเว็บไซต์บริการศูนย์กลางข้อมูลภาครัฐ (Open Government Data) (สำนักงานรัฐบาลอิเล็กทรอนิกส์ องค์การมหาชน, 2559)

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพ และเลือกเทคนิควิธีที่เหมาะสมในการวิเคราะห์ข้อมูลอนุกรมเวลาของปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต อีกประการหนึ่งเพื่อสร้างรูปแบบการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ด้วยวิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูลตั้งแต่ปี พ.ศ. 2553 ถึง ปี พ.ศ. 2558

การเตรียมข้อมูล (Data Preparation)

ผู้วิจัยสร้างแบบจำลองเพื่อใช้ในการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) 3 เทคนิค ได้แก่ ได้แก่ การถดถอยเชิงเส้น (Linear Regression) แบบจำลองโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) และซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) โดยผลลัพธ์ของการประมวลผลข้อมูลอนุกรมเวลาจะอยู่ในรูปแบบจำลองของการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ผู้วิจัยแบ่งข้อมูลออกเป็น 2 ส่วน คือ 1) ชุดข้อมูลเรียนรู้ (Training Data Set) วิเคราะห์โดยใช้การวัดรากของความคลื่อนที่กำลังสอง (Root Mean Square Error: RMSE) และค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error: MAE) เพื่อแสดงการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อใช้ประมาณค่าปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตสำหรับ 2) ชุดข้อมูลทดสอบ (Testing Data Set) โดยใช้วิธีการประมาณการความแม่นยำในการพยากรณ์ด้วยค่าความคลาดเคลื่อนสัมพัทธ์ (Magnitude of Relative Error: MRE) สำหรับทดสอบประสิทธิภาพของแบบจำลองโดยแยกแต่ละเดือนในช่วงของชุดข้อมูลทดสอบ และใช้ค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ (Mean Magnitude of Relative Error: MMRE) เพื่อใช้ในการทดสอบประสิทธิภาพโดยรวมของแบบจำลองการพยากรณ์สำหรับชุดข้อมูลทดสอบ

ชุดข้อมูลจะประกอบไปด้วยปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ตั้งแต่ปี พ.ศ.2553 ถึง ปีพ.ศ.2558 จำนวน 6 ปี โดยในแต่ละชุดข้อมูลจะแบ่งข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลที่จะนำมาสอนระบบเพื่อสร้างแบบจำลองการทำนาย (Training Data Set) โดยใช้ข้อมูลในระหว่างปี พ.ศ.2553 ถึง ปีพ.ศ.2557 และแบ่งข้อมูลที่เหลือในปี 2558 สำหรับใช้เป็นข้อมูลในการทดสอบ (Testing Data Set) แบบจำลองการพยากรณ์

การวิเคราะห์ข้อมูล (Data Analysis)

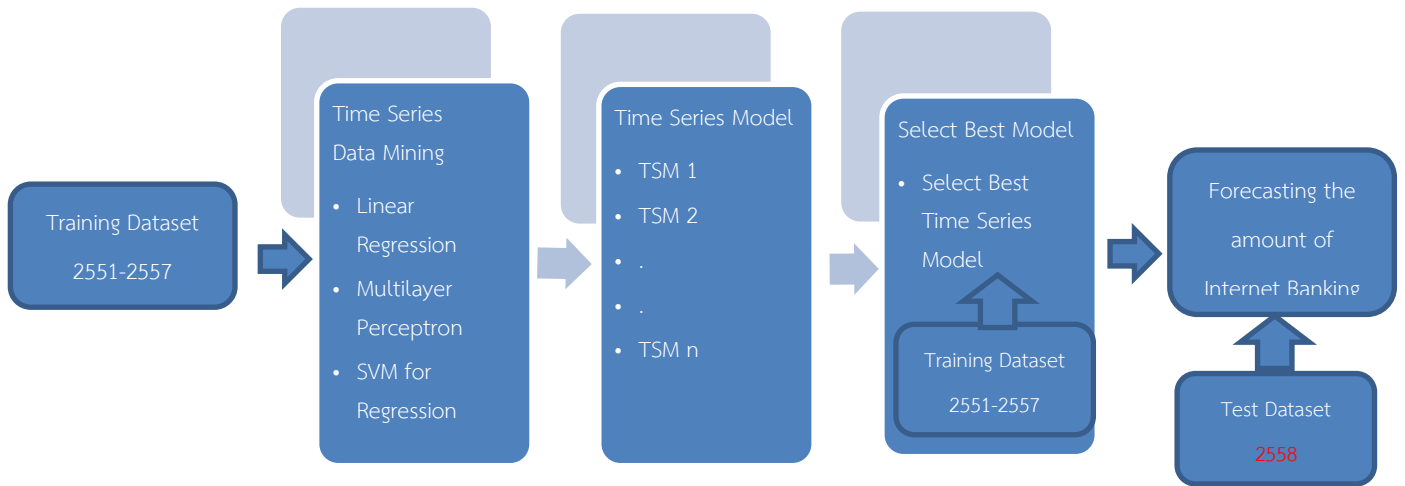
ผู้วิจัยเลือกการใช้โปรแกรม WEKA Version 3.9.0 เพื่อวิเคราะห์ข้อมูลและสร้างแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) วิธีการทั้งหมดจะให้ผลลัพธ์ในรูปแบบของแบบจำลอง (Model) ซึ่งถือเป็นลักษณะของการแทนความรู้ (Knowledge Representation) แบบหนึ่ง

สำหรับรูปแบบการพยากรณ์จะมีลักษณะเป็นแบบ Sliding Window ซึ่งเป็นการจัดเรียงข้อมูลแบบอนุกรมเวลาโดยสร้างชุดข้อมูลสำหรับการสร้างแบบจำลองการพยากรณ์ผู้วิจัยจะนำข้อมูลในปี พ.ศ.2553 ถึง ปี พ.ศ.2557 เป็นชุดข้อมูลทดสอบแบ่งเป็นชุดข้อมูลย้อนหลัง (Lagged) 1) 1 ปี 2) 2 ปี 3) 3 ปี 4) 4 ปี โดยผลลัพธ์ในการทดสอบประสิทธิภาพของชุดข้อมูลในปี พ.ศ.2558 เป็นเป้าหมาย (Target) ในการพยากรณ์ประสิทธิภาพของการพยากรณ์ในแต่ละอัลกอริทึมจะเป็นดัชนีชี้วัดว่าอัลกอริทึมใดเหมาะสมกับการพยากรณ์ในเดือนใด รายละเอียดชุดข้อมูลตัวอย่างแสดงได้ดังตารางที่ 1

ตารางที่ 1 รายละเอียดการจัดเรียงข้อมูลแบบอนุกรมเวลาโดยข้อมูลมีลักษณะเป็นเดือน 12 เดือน

เดือน	ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต (พันล้านบาท)					
	2553	2554	2555	2556	2557	2558
	Training Data Set/Variation Data Set					Testing Data Set
มกราคม	551	716	959	1241	1826	1768
กุมภาพันธ์	549	632	1058	1289	1645	1646
มีนาคม	674	781	1184	1338	1870	1914
เมษายน	672	682	1046	1663	1712	1785
พฤษภาคม	666	739	1260	1677	1703	1843
มิถุนายน	669	756	1239	1737	1676	2060
กรกฎาคม	668	742	1253	1784	1844	2064
สิงหาคม	636	846	1278	1745	1679	2162
กันยายน	683	772	1177	1974	1837	2328
ตุลาคม	666	708	1233	1722	1590	2132
พฤศจิกายน	708	644	1190	1647	1387	2000
ธันวาคม	749	762	1233	1732	1731	2180
รวม	7891	8780	14110	19549	20500	23882

จากข้อมูลในตาราง ผู้วิจัยนำข้อมูลอนุกรมเวลาทั้งหมดไปทดสอบสร้างแบบจำลอง โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมือนข้อมูล ทั้ง 3 เทคนิค (Linear Regression, Multilayer Perceptron, Support Vector Machine for Regression) เพื่อทดสอบประสิทธิภาพของการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ทั้งนี้ผู้วิจัยพิจารณาค่า RMSE และ MAE มาเป็นดัชนีชี้วัดแบบจำลองการทำนายที่มีความเหมาะสมกับการพยากรณ์แต่ละเดือนสำหรับการทดสอบประสิทธิภาพของแบบจำลอง โดยแยกแต่ละเดือนในชุดข้อมูลทดสอบ ผู้วิจัยพิจารณาค่า MRE และ MMRE เพื่อเลือกแบบจำลองที่เหมาะสม (Weiss & Indurkha, 1998) ในการนำไปใช้งาน



รูปที่ 2 กระบวนการวิเคราะห์แบบจำลองสำหรับการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต

ผลการวิจัย

ผลการวิจัยพบว่า การวัดประสิทธิภาพแบบจำลองเพื่อใช้ในการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) ตามวิธีการดำเนินงานวิจัยด้วยเทคนิค 3 วิธี กับการทดสอบชุดข้อมูลเรียนรู้ (Training Data Set) และชุดข้อมูลทดสอบ (Testing Data Set) ตามขั้นตอนวิธีวิจัย สามารถแบ่งออกเป็น 2 ส่วน ดังนี้

การเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายกับชุดข้อมูล

จากชุดข้อมูลที่ผู้วิจัยเลือกปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ซึ่งเป็นชุดข้อมูลสำหรับการสร้างแบบจำลองการพยากรณ์ ผู้วิจัยนำข้อมูลในปี พ.ศ.2553 – 2557 เป็นชุดข้อมูลทดสอบ แบ่งเป็น ชุดข้อมูลย้อนหลัง (Lagged) 1) 1 ปี 2) 2 ปี 3) 3 ปี 4) 4 ปี นำมาสร้างแบบจำลองในการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต เพื่อวิเคราะห์ประสิทธิภาพที่ได้จากเทคนิควิธีเหมืองข้อมูลทั้ง 3 วิธี โดยการเปรียบเทียบความคลาดเคลื่อนด้วยวิธีการ MAE และ RMSE จากผลการทดลองสามารถแสดงได้ดังตาราง

ตารางที่ 2 การเปรียบเทียบประสิทธิภาพแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตโดยใช้ชุดข้อมูลย้อนหลัง (Lagged)

Month Lagged	Time Series Data Mining techniques					
	Linear Regression		Multilayer Perceptron		SMOreg	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Lag 12	<u>101.84</u>	<u>127.37</u>	164.01	230.15	98.45	127.26
Lag 24	107.57	128.33	38.91	48.81	91.50	126.10
Lag 36	121.83	149.44	<u>29.60</u>	<u>37.49</u>	54.24	104.26
Lag 48	160.55	183.49	53.18	62.07	<u>3.0</u>	<u>5.76</u>

จากตารางที่ 2 ผลการทำนายของแต่ละเทคนิควิธีเหมืองข้อมูลกับการใช้ชุดข้อมูล โดยมีจำนวนเดือนย้อนหลัง (Lagged) ที่แตกต่างกันเมื่อใช้จ่าย MAE และ RMSE เป็นดัชนีชี้วัดประสิทธิภาพ จากการทดลองแสดงให้เห็นว่า เมื่อสร้างแบบจำลองด้วยเทคนิคการถดถอยเชิงเส้น ด้วยข้อมูลย้อนหลัง 12 เดือน (Lag 12) มีประสิทธิภาพสูงที่สุด สำหรับแบบจำลองโครงข่ายประสาทเทียม ด้วยข้อมูลย้อนหลัง 36 เดือน (Lag 36) มีประสิทธิภาพสูงที่สุด และแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย ด้วยข้อมูลย้อนหลัง 48 เดือน (Lag 48) มีประสิทธิภาพสูงที่สุด

จากผลการทดลอง ผู้วิจัยจึงเลือกชุดข้อมูลย้อนหลังที่ทำให้ได้ค่าประสิทธิภาพดีที่สุด มาใช้สร้างแบบจำลองเพื่อการทำนายปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต บนชุดข้อมูลทดสอบ (Testing Data Set) ซึ่งเป็นปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ปี พ.ศ.2558

การเปรียบเทียบประสิทธิภาพแบบจำลองการทำนายในแต่ละเดือน

จากการนำแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต มาทดสอบกับชุดข้อมูลปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ในปี พ.ศ.2558 โดยแยกปริมาณออกเป็นรายเดือนจำนวน 12 เดือน ด้วยวิธีวิเคราะห์หอนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining Techniques) โดยใช้อัลกอริทึม 3 แบบ และคำนวณหาค่าความคลาดเคลื่อนสัมพัทธ์ (Magnitude of Relative Error) สามารถแสดงผลการทดลองได้ดังตารางต่อไปนี้

ตารางที่ 3 การเปรียบเทียบประสิทธิภาพแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตในแต่ละเดือน

ชุดข้อมูล		Time Series Data Mining techniques					
ปี 2558		LinearRegression (Lag 12)		MultilayerPerceptron (Lag 36)		SMOreg (Lag 48)	
Month	Actual	Predict	MRE	Predict	MRE	Predict	MRE
มกราคม	1768	1647.74	0.07	1501.22	0.15	1644.81	0.07
กุมภาพันธ์	1646	1662.38	0.01	1648.28	0.00	1847.83	0.12
มีนาคม	1914	1654.72	0.14	1741.70	0.09	1910.13	0.00
เมษายน	1785	1652.27	0.07	1746.64	0.02	1649.22	0.08
พฤษภาคม	1843	1648.73	0.11	1831.51	0.01	1945.67	0.06
มิถุนายน	2060	1645.54	0.20	1902.10	0.08	1611.71	0.22
กรกฎาคม	2064	1642.36	0.20	1803.16	0.13	1160.24	0.44
สิงหาคม	2162	1639.28	0.24	1866.00	0.14	249.87	0.88
กันยายน	2328	1636.26	0.30	1872.84	0.20	102.81	0.96
ตุลาคม	2132	1633.32	0.23	1885.34	0.12	3.88	1.00
พฤศจิกายน	2000	1630.45	0.18	1968.38	0.02	167.70	0.92
ธันวาคม	2180	1627.65	0.25	1947.48	0.11	-11.64	1.01
MMRE		16.75%		8.70%		47.85%	

เมื่อพิจารณาผลการทดลองจากตารางที่ 3 พบว่าการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล (Time Series Data Mining techniques) ในภาพรวมของปี 2558 แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) โดยใช้ชุดข้อมูลย้อนหลัง 36 เดือน (Lag 36) มีค่าประสิทธิภาพโดยรวมสูงที่สุด มีค่า MMRE = 8.70% ถ้าดับถัดมาได้แก่ แบบจำลองการพยากรณ์ด้วยเทคนิคการถดถอยเชิงเส้น (Linear Regression) ใช้ชุดข้อมูลย้อนหลัง 12 เดือน (Lag 12) โดยมีค่า MMRE = 16.75%

เมื่อพิจารณาในรายละเอียดของผลลัพธ์การพยากรณ์แต่ละเดือนในชุดข้อมูลทดสอบปี 2558 พบว่าแบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น (Multilayer Perceptron) มีประสิทธิภาพสูงที่สุดจำนวน 10 เดือนจากชุดข้อมูลทั้งหมด 12 เดือน โดยมีค่า

MRE อยู่ระหว่าง 0.00 – 0.20 และ แบบจำลองการพยากรณ์ด้วยเทคนิคการถดถอยเชิงเส้น (Linear Regression) มีค่าประสิทธิภาพสูงที่สุด จำนวน 1 เดือน โดยมีค่า MRE เท่ากับ 0.07 และเทคนิควิธีชัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอย (Support Vector Machine for Regression) มีค่าประสิทธิภาพสูงที่สุด จำนวน 2 เดือน โดยมีค่า MRE อยู่ระหว่าง 0.00-0.07

สรุปผลการวิจัย และข้อเสนอแนะที่ได้รับจากการวิจัย

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อสร้างแบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต โดยใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยเทคนิคเหมืองข้อมูล โดยมีขั้นตอนในการดำเนินการวิจัยแบ่งออกเป็น 3 ขั้นตอนดังนี้ 1) ศึกษาปัญหาและวิเคราะห์ข้อมูล 2) การเตรียมข้อมูลและ 3) การวิเคราะห์ข้อมูล

ผลการวิจัยสามารถสรุปได้ว่า แบบจำลองการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตกับชุดข้อมูลในปี 2558 โดยนำเสนอข้อมูลเป็นรายเดือนพบว่า แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น มีค่า MMRE = 8.70% มีค่าประสิทธิภาพโดยรวมสูงที่สุด ถัดมาได้แก่ แบบจำลองการพยากรณ์ด้วยเทคนิคการถดถอยเชิงเส้น มีค่า MMRE = 16.75%

เมื่อพิจารณาจากผลการวิจัยโดยละเอียดจะพบว่า ประเด็นสำคัญในการพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตกับชุดข้อมูลในปี 2558 พบว่าเดือนมกราคม แบบจำลองการพยากรณ์ด้วยวิธีเทคนิคการถดถอยเชิงเส้นและวิธีชัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอยเชิงเส้น มีค่า MRE = 0.07 มีค่าประสิทธิภาพสูงที่สุด เดือนกุมภาพันธ์ แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น มีค่า MRE = 0.00 มีค่าประสิทธิภาพสูงที่สุด เดือนมีนาคม แบบจำลองการพยากรณ์ด้วยวิธีชัพพอร์ตเวกเตอร์แมชชีนสำหรับการถดถอยเชิงเส้น มีค่า MRE = 0.00 มีค่าประสิทธิภาพสูงที่สุด เดือนเมษายน ถึงธันวาคม แบบจำลองการพยากรณ์ด้วยเทคนิควิธีโครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น มีค่า MRE ระหว่าง 0.01-0.20 มีค่าประสิทธิภาพสูงที่สุด จากผลการวิจัยดังกล่าวจะเห็นว่าแต่ละเดือนค่า MRE ที่มีประสิทธิภาพสูงที่สุด จะเปลี่ยนแปลงตามแบบจำลองการพยากรณ์และข้อมูลย้อนหลัง (Lagged) ดังนั้นสถาบันการเงินสามารถนำผลลัพธ์ของแบบจำลองพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ต ไปใช้ประโยชน์ โดยเลือกแบบจำลองการพยากรณ์ด้วยวิธีเทคนิคที่มีค่าประสิทธิภาพสูงที่สุดแต่ละเดือนตามผลการวิจัย จะส่งผลให้การพยากรณ์ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตมีความแม่นยำสูงขึ้น

เมื่อพิจารณาผลลัพธ์จากการวิจัยและการนำไปประยุกต์ใช้ในงานด้านแผนการประชาสัมพันธ์เชิงรุก เพื่อกระตุ้นลูกค้าของสถาบันการเงินให้มาชำระเงินผ่านธนาคารบนอินเทอร์เน็ตมากขึ้น ในเวลาที่ปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตลดลง นอกจากนี้ ยังพบว่าปริมาณการชำระเงินผ่านธนาคารบนอินเทอร์เน็ตมีแนวโน้มเพิ่มขึ้นเกือบทุกเดือน สถาบันการเงินสามารถนำผลลัพธ์จากการวิจัย

ไปวางแผนเทคโนโลยีสารสนเทศในด้านซอฟต์แวร์ ฮาร์ดแวร์ ระบบรักษาความปลอดภัยในการชำระเงิน
เพื่อรองรับลูกค้าที่เพิ่มขึ้นในอนาคต

เอกสารอ้างอิง

ภาษาอังกฤษ

Berry, Michael J. A., and Gordon S. Linnoff.(2004) Data Mining Techniques For Marketing, Sale and
Customer Relationship Management. New York: Wiley Publishing

Cai, T.T., Hall, P., & others. (2006). Prediction in functional linear regression. The Annals of Statistics,
34(5), 2159-2179.

Frias-Martinez, E., Sanchez, A., & Velez, J. (2006). Support vector machines versus multi-layer
perceptrons for efficient off-line signature recognition. Engineering Applications of Artificial
Intelligence, 19(6), 693-704.

Hoshmand, A. R. (2009). Business Forecasting, Second Edition: A Practical Approach. Routledge.

Shevade, S. K., Keetthi, S. S., Bhattacharyya, C., & Murthy, K. R. K. (2000). Improvements to the
SMO algorithm for SVM regression. IEEE Transactions on Neural Networks, 11(5), 1188-
1193.

Weiss, S. M., & Indurkha, N. (1998). Predictive Data Mining: A Practical Guide. San Francisco, CA,
USA: Morgan Kaufmann Publishers Inc.

เว็บไซต์

ธนาคารแห่งประเทศไทย. สถิติ. (ออนไลน์) เว็บไซต์ <https://www.bot.or.th> (เข้าถึงเมื่อวันที่ 1 กันยายน
2559).

Wikipedia.ธนาคารอิเล็กทรอนิกส์.(ออนไลน์) [https://th.wikipedia.org/wiki/ธนาคาร
อิเล็กทรอนิกส์](https://th.wikipedia.org/wiki/ธนาคารอิเล็กทรอนิกส์).(เข้าถึงเมื่อวันที่ 1 กันยายน 2559).

สำนักงานพัฒนาธุรกรรมทางอิเล็กทรอนิกส์ (องค์การมหาชน).E-Banking(ออนไลน์)
<https://standard.etda.or.th/wp/?p=219>. (เข้าถึงเมื่อวันที่ 25 กันยายน 2559).