

การวิเคราะห์ปัจจัยที่มีผลต่อการผันแปรของหุ้นกลุ่มอุปโภคบริโภคโดยใช้เทคนิค

การทำเหมืองข้อมูล

Analysis of factors affecting the variation of consumption stocks using by data mining.

ณัฐวุฒิ ชญานิชกุล¹ สมชาย เล็กเจริญ²

บทคัดย่อ

วัตถุประสงค์ของการวิจัยนี้ เพื่อศึกษาปัจจัยที่มีผลกระทบต่อการแปรผันของราคาหุ้นกลุ่มอุปโภคบริโภค และเพื่อเปรียบเทียบประสิทธิภาพ แบบจำลองโมเดลต้นไม้การตัดสินใจและหารูปแบบที่เหมาะสม เพื่อใช้ในการตัดสินใจ โดยการนำข้อมูลหุ้นของกลุ่มอุปโภคบริโภค 11 องค์ประกอบ ได้แก่แนวโน้มราคาหุ้น (Trend), มูลค่าหลักทรัพย์ตามราคาตลาด (Market Capitalization MC), อัตราส่วนราคาตลาดต่อกำไรสุทธิ (P/E), อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี (P/BV), มูลค่าตามบัญชีต่อหุ้น (Book Value per Share BVPS), อัตราผลตอบแทนจากเงินปันผล (Dividend Yield DivY), อัตราการหมุนเวียนการซื้อขาย (Turnover Ratio), อัตราผลตอบแทนผู้ถือหุ้น (Return On Equity ROE), ผลตอบแทนต่อสินทรัพย์ (Return on Assets ROA), สินทรัพย์สุทธิ (Asset) และราคา ณ เวลาปิดตลาด (Close) ตั้งแต่ปี พ.ศ. 2555 ถึง พ.ศ. 2559 โดยผู้วิจัยเลือกใช้โปรแกรมสำเร็จรูปในการวิเคราะห์ทางสถิติมาวิเคราะห์โดยใช้เทคนิคการวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) มาทดสอบสมมติฐานที่กำหนดไว้ในการศึกษา ในการหากระดับนัยสำคัญที่ 0.05 และการพยากรณ์เปรียบเทียบประสิทธิภาพการวิเคราะห์ผลโดยการนำตัวแปรมาจำลองโมเดลต้นไม้การตัดสินใจ (Decision Tree) ชนิด J48, REPTree, RandomTree, RandomForest, LMT, และ Hoeffding Tree กับโปรแกรม Weka

ผลการทดลองพบว่าแบบจำลองโมเดลต้นไม้การตัดสินใจชนิด RandomTree มีประสิทธิภาพดีที่สุด มีความแม่นยำอยู่ที่ 100% และค่าความคลาดเคลื่อนเฉลี่ยอยู่ที่ 0.00 ปัจจัยที่มีผลต่อการเปลี่ยนราคาหุ้นคือ อัตราผลตอบแทนจากเงินปันผล (Dividend Yield) รองลงมาคือ อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี (P/BV) และสุดท้ายคือ มูลค่าตามบัญชีต่อหุ้น Book Value per Share (BVPS) โดยมีค่า R Square = 0.73 ซึ่งการวิจัยครั้งนี้มีประโยชน์ต่อการตัดสินใจซื้อหุ้นกลุ่มอุปโภคบริโภค

คำสำคัญ: การวิเคราะห์การถดถอยพหุคูณ, การทำเหมืองข้อมูล, ต้นไม้ตัดสินใจ, หุ้นกลุ่มอุปโภคบริโภค, การวิเคราะห์ทางสถิติ

¹นักศึกษาระดับปริญญาโท หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ วิทยาลัยเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยรังสิต อีเมล nattawut.c56@rsu.ac.th

²อาจารย์ประจำ หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ วิทยาลัยเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยรังสิต อีเมล somchai.l@rsu.ac.th

Abstract

The objectives of this research were to study the factors that affect stock prices of consumer stocks in order to compare and to find suitable decision models. The factors in this study were following Trend, Market Capitalization (MC), Price to Earnings Ratio (P/E), Price to Book Ratio (P/BV), Book Value per Share (BVPS), Dividend Yield (DivY.), Turnover Ratio, Return On Equity (ROE), Return on Equity Asset Return on Assets (ROA), Asset Value, and Close Price. According to the data for consumer stocks from 2012 to 2016, The researcher had selected the statistical analysis program by using Multiple regression analysis to test hypotheses of the study. The significance level was 0.05. For the forecasting of the information of property, the researcher used decision tree such as algorithms J48, REPTree, RandomTree, RandomForest, LMT, and Hoffmann tree models with Weka program.

This study shows the Random Tree model had the best performance. Its accuracy was 100% and the average deviation is 0.00. The factor that has the most significant effect is Dividend Yield. The Price to Book Ratio (P/ BV) came in the second. Then the Book Value per Share (BVPS) came in the last with R Square value of 0.73.

Keyword: Multiple Regression Analysis, Data mining, Decision Tree, Consumer stock, Statistical Analysis

บทนำ

การลงทุนในประเทศไทยในอดีต ผู้ลงทุนส่วนใหญ่มักเน้นไปที่การออม การฝากเงินไว้กับธนาคารพาณิชย์หรือสถาบันการเงินต่างๆ ยิ่งผู้ลงทุนฝากมากดอกเบี้ยที่ได้รับก็ยิ่งมีมาก และนั่นเป็นวิธีที่ปลอดภัยและความเสี่ยงน้อย แต่อาจจะใช้เวลาเวลานานกว่าจะได้ดอกเบี้ย การลงทุนกับตราสารทุน หรืออนุพันธ์ยังไม่เป็นที่นิยม เนื่องจากมีความเสี่ยงที่สูงมาก ทั้งตัวเลือกการลงทุนที่น้อย และการผันผวนของค่าเงิน ถึงแม้ว่าได้เงินตอบแทนที่สูงและรวดเร็วกว่าการออม แต่ในปัจจุบัน การลงทุนนั้นได้แพร่หลายมากขึ้น ทั้งการลงทุนในอสังหาริมทรัพย์ ตราสารทุนหรือหุ้น อนุพันธ์ ตราสารหนี้หรือ หุ้นกู้ การออมและการเกรงกำไรทอง ทำให้มีนักลงทุนสนใจที่จะลงทุนเพิ่มขึ้นจำนวนมาก ถึงแม้ความเสี่ยงจะยังมีเหมือนเดิมก็ตาม และในปัจจุบันการลงทุนในหุ้นเป็นการลงทุนที่เป็นที่นิยมและให้ผลตอบแทนที่ดีกว่าการออม ผู้ลงทุนสามารถเริ่มลงทุนด้วยเงินทุนที่ไม่สูงมากและสามารถทำกำไรได้รวดเร็วกว่าการออม ถึงกระนั้นการลงทุนในหุ้นนั้นก็ยังมีความเสี่ยงสูงเนื่องจากการผันผวนตามปัจจัยต่างๆ ผู้เริ่มต้นลงทุนมักจะมีพฤติกรรมการลงทุน (กนกวรรณ ศรีนวล และ สุทธินันท์ พรหมสุวรรณ, 2559) ที่แตกต่างกันขึ้นอยู่กับสภาพแวดล้อม สังคม และประสบการณ์ โดยผู้เริ่มต้นลงทุนมักจะประสบปัญหา ในการวิเคราะห์และการตัดสินใจเลือกซื้อหุ้น สิ่งที่นักลงทุนส่วนใหญ่ควรปฏิบัติคือการวิเคราะห์ความเสี่ยงในการลงทุน เพื่อเพิ่มประสิทธิภาพในการตัดสินใจเลือกลงทุน

นักลงทุนส่วนใหญ่มักหาข่าวสารและข้อมูลที่มีผลต่อราคาหุ้นมาใช้เป็นตัวแปรในการวิเคราะห์เพื่อหาปัจจัยและโอกาส ที่จะได้ผลตอบแทนจากการลงทุนมากที่สุด โดยการวิเคราะห์ที่มี 2 รูปแบบที่ คือ การวิเคราะห์ปัจจัยพื้นฐานและการวิเคราะห์ปัจจัยทางเทคนิค

การวิเคราะห์ปัจจัยพื้นฐาน (Fundamental analysis) เป็นการวิเคราะห์ภาวะเศรษฐกิจ ภาวะอุตสาหกรรม อัตราสภาพคล่อง ความสามารถในการดำเนินงาน ความสามารถในการทำกำไร อัตราการชำระหนี้ ภาวะบริษัท เพื่อนำมาใช้ในการกำหนดมูลค่าที่แท้จริงของหลักทรัพย์ (การวิเคราะห์ปัจจัยพื้นฐาน, 2558)

การวิเคราะห์ปัจจัยทางเทคนิค (Technical Analysis) เป็นการศึกษาพฤติกรรมของราคาหุ้น หรือพฤติกรรมของตลาดในอดีตโดยใช้หลักสถิติ เพื่อนำมาใช้คาดการณ์พฤติกรรมราคาหุ้นในอนาคต และช่วยให้ผู้ลงทุนหาจังหวะการลงทุนที่เหมาะสม โดยข้อมูลหลักที่ใช้ ในการวิเคราะห์ทางเทคนิคได้แก่ ระดับราคา และปริมาณการซื้อขายหุ้น (การวิเคราะห์ทางเทคนิค, 2558) เป็นสาเหตุให้ผู้ลงทุนในการซื้อหุ้นกลุ่มอุปโภคบริโภค ยกต่อการตัดสินใจในการซื้อหุ้น

ผู้วิจัยได้นำเอาข้อมูลหรือตัวแปรที่มีเกี่ยวข้องกับ ความผันแปรของหุ้นมาวิเคราะห์ ผ่านอัลกอริทึมต่างๆ ด้วยเทคนิคการทำเหมืองข้อมูล (Data mining) เพื่อเปรียบเทียบและหาความแม่นยำของอัลกอริทึมและนำผลลัพธ์มาสนับสนุนการตัดสินใจลงทุนหุ้นในกลุ่มอุปโภคบริโภค ในปัจจุบันการทำเหมืองข้อมูลได้นำเอาไปประยุกต์ใช้กับงานในหลากหลายด้านการวิเคราะห์ในครั้งนี้เป็นงานวิเคราะห์หุ้นในกลุ่มอุปโภคบริโภค (Consumer) ซึ่งจะแบ่งแยกย่อยได้ 3 หมวด ได้แก่ หมวดแฟชั่น (Fashion) หมวดของใช้ในครัวเรือนสำนักงาน (Home & Office) และหมวดของใช้ส่วนตัวเวชภัณฑ์ (Personal) โดยนำข้อมูลจาก SETSMART (SET Market Analysis and Reporting Tool) เป็นบริการระบบข้อมูลตลาดหลักทรัพย์ฉบับออนไลน์ ผลิตโดยตลาดหลักทรัพย์แห่งประเทศไทย มาวิเคราะห์การผันแปรของราคาหุ้นเพื่อช่วยสนับสนุนตัดสินใจในการลงทุน

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาปัจจัยที่มีความสัมพันธ์ต่อการเปลี่ยนแปลงของราคาหุ้นมากที่สุด
2. เพื่อเปรียบเทียบประสิทธิภาพ แบบจำลองโมเดลต้นไม้การตัดสินใจและหารูปแบบที่เหมาะสม เพื่อใช้ในการตัดสินใจ

เนื้อหาที่เกี่ยวข้อง

การทำเหมืองข้อมูล (Data Mining) คือการกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบความสัมพันธ์ ที่ซ่อนอยู่ในชุดข้อมูลนั้น (ภควัต คุปต์ธนโรจน์, 2554)

เทคนิคการทำเหมืองข้อมูลแบบการจัดแบ่งประเภทของข้อมูล (Classification) เป็นการจัดแบ่งประเภทของข้อมูล โดยหาชุดต้นแบบหรือชุดของการทำงานที่อธิบายและแบ่งประเภทข้อมูล วัตถุประสงค์เพื่อให้สามารถใช้เป็นต้นแบบทำนายประเภทของวัตถุหรือข้อมูลที่ ไม่มีการระบุประเภทหรือชนิดของข้อมูล ซึ่งต้นแบบสร้างจากการวิเคราะห์ชุดของข้อมูลฝึกสอน (Training Data) โดยอาจจะเป็นกลุ่มข้อมูลที่มีการระบุประเภทหรือกลุ่มเรียบร้อยแล้ว รูปแบบของต้นแบบแสดงได้หลายแบบเช่น Classification Rules, Decision Trees หรือ Neural Networks เป็นต้น (อดุลย์ ยิ้มงาม, 2561)

แบบจำลองโมเดลต้นไม้การตัดสินใจ (Decision Tree) ที่มีลักษณะเหมือนแผนภูมิองค์กรหรือรากของต้นไม้ โดยที่แต่ละโหนดแสดง Attribute แต่ละกิ่งแสดงผลในการทดสอบ และสีฟองหรือโหนดหรือใบจะแสดงคลาสที่กำหนดไว้ วิธีนี้ใช้เวลาในการเรียนรู้สั้นกว่าวิธีอื่น ผลลัพธ์ที่ได้เขียนให้เป็นกฎที่เข้าใจได้ง่าย และมีความถูกต้องสูงเมื่อเทียบกับวิธีอื่น (สุพจน์ ผ่องสวัสดิ์กุล, 2554)

การวิเคราะห์การถดถอยเชิงพหุคูณ (Multiple Linear Regression Analysis) ซึ่งเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตาม 1 ตัว เพื่อศึกษาว่ามีตัวแปรอิสระตัวใดบ้างที่ร่วมกันทำนายหรือพยากรณ์หรืออธิบายการผันแปรของตัวแปรตามได้ เขียนเป็นสมการดังนี้

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$$

X_i คือ ค่าของตัวแปรอิสระแต่ละตัว

Y คือ ค่าของตัวแปรตาม

β_0 คือ ค่าคงที่ (Constant) ของสมการถดถอย

β_i คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficient) ของตัวแปรอิสระ X_i แต่ละตัว (จะใช้สัญลักษณ์ b_i สำหรับค่าที่ได้จากตัวอย่าง และ สำหรับค่าประมาณหรือตัวทำนาย) β_i หรือ b_i จะแสดงอัตราการเปลี่ยนแปลงของค่า x_i ต่อค่า y ดังนี้ คือ ถ้าค่า x_i เปลี่ยนไป 1 หน่วย จะทำให้ค่า y เปลี่ยนไป b_i หน่วย

ϵ คือ ค่าความคลาดเคลื่อน (Error or Residual) (สุทิน ชนะบุญ, 2560)

วิธีดำเนินการวิจัย

1. การเตรียมข้อมูล

รวบรวมข้อมูลหุ้นกลุ่มอุปโภคบริโภค (Consumer) หมวดแฟชั่น (Fashion) หมวดของใช้ในครัวเรือนสำนักงาน (Home & Office) และหมวดของใช้ส่วนตัวเวชภัณฑ์ (Personal) จาก www.setsmart.com และเก็บบันทึกค่าตัวแปรทางการเงินที่สำคัญดังนี้ แนวโน้มราคาหุ้น, มูลค่าหลักทรัพย์ตามราคาตลาด, อัตราส่วนราคาตลาดต่อกำไรสุทธิ, อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี, มูลค่าตามบัญชีต่อหุ้น, อัตราผลตอบแทนจากเงินปันผล, อัตราการหมุนเวียนการซื้อขาย, อัตราผลตอบแทนผู้ถือหุ้น, ผลตอบแทนต่อสินทรัพย์, สินทรัพย์สุทธิและราคา ณ เวลาปิดตลาด ตั้งแต่ปี พ.ศ. 2555 ถึง พ.ศ. 2559 เป็นระยะเวลา 5 ปี ทั้งหมด 40 บริษัทลงในโปรแกรม Notepad และบันทึกนามสกุลไฟล์เป็น .CSV

2. การหาค่าปัจจัยที่มีผลต่อการเปลี่ยนแปลงของราคาหุ้นกลุ่มอุปโภคบริโภค

นำข้อมูลหุ้นกลุ่มอุปโภคบริโภคที่เก็บบันทึกไว้แบบไฟล์ .CSV นำเข้ามาใส่ในโปรแกรมวิเคราะห์สถิติข้อมูลและใช้การวิเคราะห์ข้อมูลด้วยวิธีแบบสมการถดถอยพหุคูณ (Multiple Regression Analysis) โดยนักวิจัยได้ใช้ ค่าราคา ณ เวลาตลาดปิด ตั้งเป็นตัวแปรตาม (Dependent Variable) และนำค่า มูลค่าหลักทรัพย์ตามราคาตลาด, อัตราส่วนราคาตลาดต่อกำไรสุทธิ, อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี, มูลค่าตามบัญชีต่อหุ้น, อัตราผลตอบแทนจากเงินปันผล, อัตราการหมุนเวียนการซื้อขาย, อัตราผลตอบแทนผู้ถือหุ้น, ผลตอบแทนต่อสินทรัพย์ และสินทรัพย์สุทธิ มาใส่เป็นตัวแปรอิสระ (Independent Variable) และเก็บผลลัพธ์จากการคำนวณเพื่อมาสรุปผลของตัวแปรที่มีผลต่อการเปลี่ยนแปลงของหุ้นกลุ่มอุปโภคบริโภค

3. การทดสอบ

นำข้อมูลหุ้นกลุ่มอุปโภคบริโภคที่เตรียมไว้ มาประมวลผลแบบการวิเคราะห์ข้อมูลด้วยเทคนิคการทำเหมืองข้อมูลกับโปรแกรมวิเคราะห์ข้อมูลด้วยเทคนิคการทำเหมืองข้อมูล Weka ใช้เทคนิคการทำเหมืองข้อมูลแบบการจัดแบ่งประเภทของข้อมูล รูปแบบต้นไม้การตัดสินใจ (Decision Tree) โดยอัลกอริทึมที่ใช้ทดสอบมีดังนี้ J48, REPTree, RandomTree, RandomForest, LMT, และ Hoeffding Tree และโปรแกรม Weka จะทำการหาค่าความแม่นยำ

(Accuracy) และหาค่าความคลาดเคลื่อน (Mean Absolute Error : MAE) ของอัลกอริทึมนี้เพื่อทำการเปรียบเทียบ และหาแบบจำลองที่ดีที่สุด จากวิธีการทดสอบหาค่าพยากรณ์ความถูกต้องของข้อมูลว่ามีความถูกต้องมากน้อยเพียงใด โดยคิดเป็นค่าร้อยละ

ผลการวิจัย

ผลจากการทดสอบแบ่งออกเป็น 2 ส่วนตามวัตถุประสงค์ดังนี้

1. ผลจากการศึกษาหาปัจจัย ผู้วิจัยได้นำผลจากการวิเคราะห์ด้วยเทคนิคการวิเคราะห์การถดถอยเชิงเส้นตรงพหุคูณ (Multiple Regression Analysis) ที่มีผลต่อการเปลี่ยนแปลงของราคาหุ้นกลุ่มอุบโภคบริโภค ได้ผลตามตารางที่ 1 ตารางที่ 2 และตารางที่ 3

ตารางที่ 1 โมเดลสรุปผล

R	R Square	Adjusted R Square	Std. Error of the Estimate
.85	0.73	0.72	28.63

การอ่านค่าในตารางที่ 1

จากตารางที่ 1 พบว่า R Square = 0.73 หรือ 73% แสดงว่า โมเดลนี้ สามารถพยากรณ์ได้ถูกต้อง 73%

ตารางที่ 2 การวิเคราะห์ความแปรผันของหุ้นกลุ่มอุบโภคบริโภค

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	384270.87	8	48033.86	58.59	.000 ^b
Residual	143460.68	175	819.78		
Total	527731.55	183			

* p < .05

จากตารางที่ 2 จะพบว่าพยากรณ์นี้มีแนวโน้มเชื่อถืออยู่ ระดับนัยสำคัญทางสถิติ 0.05

ตารางที่ 3 การวิเคราะห์ค่าสัมประสิทธิ์

Model	Unstandardized Coefficients		Standardized Coefficients		T	Sig.
	B	Std. Error	Beta			
(Constant)	-6.71	4.14			-1.62	0.11
P/E	0.01	0.03	0.01		0.31	0.76
P/BV	1.21	0.68	0.08		1.79	0.08
BVPS	0.84	0.04	0.82		19.49	0.00
DivY	2.64	0.79	0.14		3.33	0.00
Turnover Ratio	0.00	0.02	0.00		-0.05	0.96
ROE	-0.02	0.02	-0.04		-0.88	0.38
ROA	0.43	0.25	0.08		1.70	0.09
NAVASSET	0.00	0.00	-0.09		-2.23	0.03

* p < .05

จากตารางที่ 3 การทดสอบวิเคราะห์หาค่าสัมประสิทธิ์ โดยการนำข้อมูล ราคา ณ เวลาปิดตลาด (Close), อัตราส่วนราคาตลาดต่อกำไรสุทธิ (P/E), อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี (P/BV), มูลค่าตามบัญชีต่อหุ้น Book Value per Share (BVPS), อัตราผลตอบแทนจากเงินปันผล DIVIDEND YIELD (DivY), อัตราการหมุนเวียนการซื้อขาย Turnover Ratio, อัตราผลตอบแทนผู้ถือหุ้น Return On Equity (ROE), ผลตอบแทนต่อสินทรัพย์ Return on Assets (ROA) และสินทรัพย์สุทธิ (Asset) จากตารางที่ 3 วิเคราะห์โดยใช้การวิเคราะห์การถดถอยเชิงพหุคูณ (Multiple Linear Regression Analysis) เขียนเป็นสมการดังนี้

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

และเมื่อนำมาใส่ค่าตัวแปรจากตารางที่ 3 จะได้สมการดังนี้ $Y = -6.71 + (0.01*P/E) + (1.21*P/BV) + (0.84*BVPS) + (2.64*DivY) + (0.00*Turnover Radio) + (-0.02*ROE) + (0.43*ROA)$

จากผลการวิเคราะห์ค่าความแปรปรวน ตัวแปรของหุ้นกลุ่มอุปโภคบริโภคที่มีผลต่อการเปลี่ยนแปลงของราคามากที่สุดคือ อัตราผลตอบแทนจากเงินปันผล (Dividend Yield) รองลงมาคือ อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี (P/BV) และสุดท้ายคือ มูลค่าตามบัญชีต่อหุ้น Book Value per Share (BVPS)

ตารางที่ 4 การเปรียบเทียบประสิทธิภาพของแบบจำลองโมเดลแต่ละชนิด

Algorithm	Accuracy	Precision	Recall	Mean Absolute Error
Random Tree	100.00	1.00	1.00	0.00
Random Forest	100.00	1.00	1.00	0.17
REPTree	81.15	0.79	0.87	0.30
LMT	62.30	0.61	0.78	0.47
Hoeffding Tree	60.20	0.63	0.6	0.42
J48	61.78	0.82	0.36	0.45

2. จากตารางที่ 4 ผลการเปรียบเทียบประสิทธิภาพของแบบจำลองของโมเดล ทั้งหมด 6 แบบ แบบจำลองของโมเดลต้นไม้ตัดสินใจ (Decision Tree) อัลกอริทึม J48, REPTree, RandomTree, RandomForest, LMT และ Hoeffding Tree ผ่านโปรแกรมทำเหมืองข้อมูล Weka ได้ผลการทดสอบดังตารางที่ 4 โดยแบบจำลองโมเดลที่ทำการวิเคราะห์ทั้งหมด ได้ผลลัพธ์ที่ดีที่สุด คือ อัลกอริทึมชนิด RandomTree ซึ่งมีค่าความแม่นยำอยู่ที่ 100% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.00 รองลงมาเป็น อัลกอริทึมชนิด RandomForest มีค่าความแม่นยำอยู่ที่ 100% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.17 อัลกอริทึมชนิด REPTree มีค่าความแม่นยำอยู่ที่ 81.15% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.30 อัลกอริทึมชนิด LMT มีค่าความแม่นยำอยู่ที่ 62.30% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.47 อัลกอริทึมชนิด Hoeffding Tree มีค่าความแม่นยำอยู่ที่ 60.20% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.42 และอัลกอริทึมชนิด J48 มีค่าความแม่นยำอยู่ที่ 61.78% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.45 สรุปว่าการพยากรณ์ข้อมูลนี้ ใช้โมเดล Random Tree สามารถพยากรณ์ได้แม่นยำ 100% และมีค่าผิดพลาดที่ 0.00

อภิปรายผล

ผลการวิจัย เพื่อหาปัจจัยที่มีผลกระทบต่อการแปรผันของราคาหุ้นกลุ่มอุปโภคบริโภคและการเปรียบเทียบของแบบจำลองโมเดลด้วยวิธีการทำเหมือนข้อมูลของหุ้นกลุ่มอุปโภคบริโภค ในตลาดหลักทรัพย์แห่งประเทศไทย ทั้งหมด 40 บริษัท โดยใช้ข้อมูลย้อนหลัง 5 ปี นับจากปี 2555 – 2559 มีประเด็นอภิปรายดังนี้

1. จากการศึกษาและการทดสอบเกี่ยวกับปัจจัยที่ส่งผลต่อการเปลี่ยนแปลงของราคาหุ้นกลุ่มอุปโภคบริโภคของตลาดหลักทรัพย์แห่งประเทศไทย โดยการวิเคราะห์แบบการถดถอยพหุคูณ (Multiple Regression Analysis) ปรากฏว่า ปัจจัยที่มีผลกระทบมากที่สุด คือ อัตราผลตอบแทนจากเงินปันผล Dividend Yield รองลงมาคือ อัตราส่วนราคาตลาดต่อมูลค่าตามบัญชี (P/BV) และสุดท้ายคือ มูลค่าตามบัญชีต่อหุ้น Book Value per Share (BVPS)

แบบจำลองโมเดลต้นไม้ตัดสินใจ (Decision Tree) ชนิด RandomTree มีค่าความแม่นยำอยู่ที่ 100% และมีค่าความคลาดเคลื่อน อยู่ที่ 0.00 (จากค่าในตารางที่ 4)

ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลวิจัยไปใช้

1. นักวิจัยและผู้สนใจด้านงานวิจัยสามารถนำผลการวิจัยนี้ไปพัฒนาและประยุกต์ให้งานวิจัยของนักวิจัยมีประสิทธิภาพ

2. ในการใช้งานจริงงานวิจัยชิ้นนี้เป็นเครื่องมือช่วยสนับสนุนการตัดสินใจในการลงทุนหุ้นของกลุ่มอุปโภคบริโภคที่อ้างอิงจากการวิเคราะห์ข้อมูลจากปี พ.ศ. 2555 – 2559 และผลลัพธ์จะมีการเปลี่ยนแปลงเสมอเมื่อผู้ใช้งานเพิ่มตัวแปรและข้อมูลใหม่เข้าไปในการวิเคราะห์

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ในการศึกษาครั้งถัดไปควรเพิ่มจำนวนตัวแปรและความถี่ของข้อมูลที่ใช้ในการวิเคราะห์เพื่อให้ได้ผลลัพธ์ที่ละเอียดและแม่นยำมากกว่าเดิม

2. การวิเคราะห์หุ้นกลุ่มอุปโภคบริโภคโดยการใช้เทคนิคเหมืองข้อมูลครั้งนี้ได้ใช้แบบจำลองโมเดลต้นไม้ตัดสินใจเพียงชนิดเดียว เพื่อความถูกต้องและแม่นยำที่สุดควรทดสอบโดยใช้แบบจำลองที่หลากหลาย แล้วนำมาเปรียบเทียบกัน

กิตติกรรมประกาศ

งานวิจัยฉบับนี้เสร็จสมบูรณ์ได้ตามเป้าหมายเพราะได้รับการช่วยเหลือและคำแนะนำจากหลายส่วนขอขอบคุณ ผศ.ดร.สมชาย เล็กเจริญ สำหรับคำปรึกษาและเทคนิคการดำเนินงาน ตลอดจนการศึกษาค้นคว้างานวิจัยนี้สำเร็จไปได้ด้วยดีและขอขอบคุณระบบข้อมูลจาก SETSMART ที่เก็บรวบรวมข้อมูลทั้งหมดของตลาดหลักทรัพย์แห่งประเทศไทยทั้งหมดบนอินเทอร์เน็ต

เอกสารอ้างอิง

- การวิเคราะห์ปัจจัยพื้นฐาน. (2558). ตลาดหลักทรัพย์แห่งประเทศไทย สืบค้นเมื่อ 22 ธันวาคม 2560, จาก www.set.or.th หรือ https://www.set.or.th/education/th/begin/stock_content04.pdf
- การวิเคราะห์ทางเทคนิค. (2558). ตลาดหลักทรัพย์แห่งประเทศไทย สืบค้นเมื่อ 22 ธันวาคม 2560, จาก www.set.or.th หรือ https://www.set.or.th/education/th/begin/stock_content06.pdf
- กนกวรรณ ศรีนวล และ สุทธินันท์ พรหมสุวรรณ. (2559). การศึกษาปัจจัยส่วนบุคคล พฤติกรรมการลงทุน และ แรงจูงใจในการลงทุนของนักลงทุนไทยที่มีอิทธิพลต่อการตัดสินใจลงทุนในตลาดหลักทรัพย์แห่งประเทศไทย. มหาวิทยาลัยกรุงเทพ
- ภควัต คุปต์ธนโรจน์. (2554). การทำเหมืองข้อมูลการขายต่อเนื่องของบริการบริหารเงินสดกรณีศึกษาธนาคารพาณิชย์แห่งหนึ่ง. มหาวิทยาลัยธรรมศาสตร์
- สุทิน ชนะบุญ. (2560). สถิติและการวิเคราะห์ข้อมูลในงานวิจัยด้านสุขภาพเบื้องต้น. สำนักงานสาธารณสุข จังหวัดขอนแก่น
- สุพจน์ ผ่องสวัสดิ์กุล. (2554). การสร้างแบบจำลองการจัดการความรู้ของศูนย์บริการด้วยเทคนิคการทำเหมืองข้อมูลกรณีศึกษา หน่วยงานเทคโนโลยีสารสนเทศแห่งหนึ่ง. มหาวิทยาลัยธรรมศาสตร์
- อดุลย์ ยิ้มงาม. (2561). การทำเหมืองข้อมูล (Data Mining). หัวหน้าแผนกเครือข่ายสารสนเทศรังสิตศูนย์คอมพิวเตอร์ สืบค้นเมื่อ 25 มกราคม 2561, จาก <http://compcenter.bu.ac.th/news-information/data-mining>