

The Comparison of Data Classification Efficiency to Predict the Decision-Making in Future Elections Among Thai Businesspersons Using Data Mining Techniques

Kittisak Sangthong

Faculty of Management Technology,
Rajamangala University of Technology Srivijaya,
Nakhon Si Thammarat Campus, Thailand

Putwimon Chotchara

Department of Management, Prince of
Songkla University, Songkla, Thailand

Ladawan Jantawong

Faculty of Science and Technology,
Rajamangala University of Technology Srivijaya,
Nakhon Si Thammarat Campus, Thailand

Arun Aiadrit

Faculty of Science and Technology,
Rajamangala University of Technology
Srivijaya, Nakhon Si Thammarat Campus,
Thailand

Abstract

This research aims to compare the efficiency of models used for predicting decision-making in future elections among Thai businesspersons. Data were collected regarding the election decision-making from businesspersons across 7 regions in Thailand. The dataset comprises 6 attributes and 2,917 records. Rapid Miner Studio program version 10.1 was used to identify the number of relational attributes and models, and 10-Fold Cross Validation was employed to evaluate the following models: Decision Tree, Naïve Bayes, Logistic Regression, Deep Learning, and Random Forest. The results indicate that the accuracies of Random Forest, Decision Tree, Naïve Bayes, and Logistic Regression are 88.85%, 88.51%, 88.46%, and 88.11% respectively. This research could be useful for predicting and analyzing election campaigns of political parties and developing an information system to support policymaking in developing countries.

Keywords

Decision tree, naïve bayes, logistic regression, random forest

CORRESPONDING AUTHOR

Kittisak Sangthong, Faculty of Management Technology Rajamangala University of Technology Srivijaya, Nakhon Si Thammarat Campus, 80110 Thailand. Email: Singhanakhon_1@hotmail.com
© College of Local Administration, Khon Kaen University. All rights reserved.

การเปรียบเทียบประสิทธิภาพการจำแนก ข้อมูลเพื่อทำนายการตัดสินใจเลือกตั้งใน อนาคตของนักธุรกิจไทย โดยใช้เทคนิค วิธีการทำเหมืองข้อมูล

กิตติศักดิ์ แสงทอง

คณะเทคโนโลยีการจัดการ
มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย
วิทยาเขตนครศรีธรรมราช (สไใหญ่)

พชรวิมล คชรัตน์

หลักสูตรการศึกษาดุษฎีบัณฑิต
สาขาการจัดการ มหาวิทยาลัยสงขลานครินทร์

ลดาวัลย์ จันทวงศ์

คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย
วิทยาเขตนครศรีธรรมราช (สไใหญ่)

อรุณ เอียดฤทธิ์

คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย
วิทยาเขตนครศรีธรรมราช (สไใหญ่)

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพโมเดลการจำแนกข้อมูลเพื่อทำนายการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจไทย โดยใช้ข้อมูลจากการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจทั้ง 7 ภูมิภาคทั่วประเทศไทย มีข้อมูลจำนวน 6 แอตทริบิวต์ และจัดเก็บจำนวน 2,917 ชุด การวิเคราะห์ข้อมูลใช้โปรแกรม Rapid Miner 10.1 ในการวิเคราะห์ปัจจัยที่เกี่ยวข้องและทดสอบประสิทธิภาพด้วยวิธี 10-Fold Cross Validation กับตัวแบบวิธีต้นไม้ตัดสินใจ (Decision Tree) วิธี Bayes วิธีการถดถอยโลจิสติกส์ (Logistic Regression) และวิธีแรนดอมฟอเรส (Random Forest) ผลการวิจัย พบว่า การจำแนกประเภทที่มีค่าความถูกต้องสูงสุดโดยวิธีแรนดอมฟอเรส มีค่าร้อยละ 88.85 รองลงมาคือ วิธีต้นไม้ตัดสินใจ มีค่าความถูกต้องร้อยละ 88.51 วิธี Bayes มีค่าความถูกต้องร้อยละ 88.46 และวิธีที่ถูกต้องน้อยที่สุด คือ วิธีการถดถอยโลจิสติกส์ มีค่าความถูกต้องร้อยละ 88.11ตามลำดับ งานวิจัยนี้สามารถนำไปใช้ทำนายเพื่อวิเคราะห์แนวโน้มในการหาเสียงเลือกตั้งของพรรคการเมือง และพัฒนาระบบสารสนเทศเพื่อใช้ประกอบการพิจารณาในการกำหนดนโยบายเพื่อการพัฒนาประเทศต่อไป

คำสำคัญ

ตัวแบบวิธีต้นไม้ตัดสินใจ, วิธีแบบเบย์, วิธีการถดถอยโลจิสติกส์, วิธีแรนดอมฟอเรส

บทนำ

บรรดาประเทศที่มีการปกครองในระบอบประชาธิปไตย (Democracy) จำเป็นอย่างยิ่งที่จะต้องมีการเลือกตั้ง เพราะว่าการปกครองที่เป็นประชาธิปไตยนั้นจะต้องยึดถืออำนาจอธิปไตยเป็นของปวงชนประเทศที่เป็นประชาธิปไตยจำเป็นต้องมีการเลือกตั้งเพื่อเปิดโอกาสให้บุคคลสามารถเสนอตัวเข้ารับใช้ส่วนรวมโดยการรับสมัครเลือกตั้ง และเปิดโอกาสให้ ประชาชนใช้สิทธิในการที่จะเลือกบุคคลที่ตนต้องการให้เป็นผู้ปกครอง หรือเป็นผู้ใช้สิทธิเป็นปากเป็นเสียงแทนตนในสภา (Riera & Cantú, 2022) การเลือกตั้งจึงถือเป็นเครื่องมือสำคัญของประชาชนที่จะลิขิตอนาคตของประเทศและยืนยันหลักการ “อำนาจอธิปไตยเป็นของประชาชน” ในขณะเดียวกัน ด้วยเหตุผลทางโครงสร้างและความสัมพันธ์ทางอำนาจที่เกิดขึ้นระหว่างอำนาจทางการเมืองกับอำนาจของประชาชนแล้ว ยังมีอำนาจทางเศรษฐกิจก็มีส่วนสำคัญในระบอบการปกครองแบบประชาธิปไตย เป็นปัจจัยหรือช่องทางที่สำคัญที่เปิดโอกาสให้กลุ่มธุรกิจหรือนักธุรกิจเข้ามาเป็นมีบทบาทในทางการเมือง สามารถผลักดันนโยบายสาธารณะต่าง ๆ โดยเฉพาะนโยบายด้านเศรษฐกิจ ส่งผลทำให้เศรษฐกิจของประเทศเติบโตอย่างรวดเร็ว ซึ่งทั้งสองกลุ่มมีการพึ่งพาอาศัยซึ่งกันและกัน เพื่อนำไปสู่ข้อสรุปที่เป็น ภาพผ่านการเลือกตั้งทางการเมืองภายในประเทศ (Good & Schwabger, 2022) อย่างไรก็ตามการเลือกตั้งในประเทศไทยก็ได้มี “รัฐธรรมนูญแห่งราชอาณาจักรไทย แก้ไขเพิ่มเติม (ฉบับที่ 1) พุทธศักราช 2564” ที่ถือเป็นเครื่องมือสำคัญในระบอบประชาธิปไตยอีกครั้งหนึ่ง เพราะมีสิ่งที่น่าสนใจเกี่ยวกับการตัดสินใจเลือกตั้งในอนาคต ก็คือรัฐธรรมนูญฉบับนี้ได้มีการกำหนดให้เลือกผู้สมัครสมาชิกสภาผู้แทนราษฎร และการเลือกตั้งพรรคการเมือง แยกออกจากกันโดยใช้บัตรเลือกตั้งเป็น 2 ใบ ซึ่งหนีไม่พ้นที่นักการเมืองและพรรคการเมืองจะต้องมีการค้นหาวิธีการด้านการตลาด การขาย การโฆษณา การสร้างสัมพันธ์ที่ดีกับลูกค้าซึ่งในทางการเมืองคือ “ผู้มีสิทธิออกเสียงเลือกตั้ง” เพื่อมุ่งสร้างความประทับใจผ่านการดูแลเอาใจใส่ผู้มีสิทธิออกเสียงทั้งในช่วงก่อนระหว่าง และภายหลังเลือกตั้ง คล้ายกับการให้บริการของบริษัทเอกชน (De Clercq, 2019) อาจจะส่งผลให้ “เงิน” กลายเป็นปัจจัยสำคัญในการทำงานทางการเมืองของนักการเมืองและพรรคการเมือง และนำไปสู่การพึ่งพาอาศัยการสนับสนุนทางการเงินจากบริษัทเอกชน ในช่วงที่ผ่านมาประเทศไทยมีนักธุรกิจที่มีฐานะทางเศรษฐกิจสูง หรืออาจเรียกว่าเป็น “ชนชั้นนำทางธุรกิจ” ที่มี อิทธิพลหรือบทบาทในกระบวนการตัดสินใจที่มีผลต่อสังคมนั้น อาจแบ่งอย่างกว้าง ๆ ออกเป็น 2 ระดับ ได้แก่ นักธุรกิจชั้นนำระดับชาติ และนักธุรกิจชั้นนำระดับท้องถิ่น (Plangsree, 2022) ซึ่งจากการทบทวนแนวคิดและการศึกษาวิจัยในอดีต ที่เกี่ยวข้องกับบทบาททางการเมืองของกลุ่มนักธุรกิจและ ความสัมพันธ์ระหว่างทุนกับการเมืองดังที่กล่าวมา ทำให้ทราบว่าบทบาททางการเมืองของกลุ่มธุรกิจในการเมืองไทยนั้น มีหลายลักษณะความสัมพันธ์ระหว่างกลุ่มนักธุรกิจและนักการเมือง หรือพรรคการเมืองก็มีความซับซ้อน มีความแตกต่างกันในแต่ละยุคสมัย และยังขึ้นอยู่กับประเภทของธุรกิจของกลุ่มนั้น ๆ ด้วย อย่างไรก็ตามการศึกษาอย่างเจาะลึกเพื่อทราบถึงการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจของไทย ตามรัฐธรรมนูญฉบับแก้ไข พ.ศ. 2560 ยังไม่มี จึงไม่สามารถที่จะทำให้เห็นภาพความเชื่อมโยงกันระหว่างนักธุรกิจกับการเลือกตั้งระดับชาติที่กำลังจะเกิดขึ้น ซึ่งคณะผู้วิจัยมองว่านับวันเครือข่ายทางธุรกิจของประเทศไทยจะมีความหลากหลายขึ้น ทั้งในแง่จำนวน และประเภทของกลุ่มธุรกิจที่เข้าไปเกี่ยวข้อง ตลอดจนการสร้างความสัมพันธ์เชื่อมโยงระหว่างกลุ่มธุรกิจด้วยกันเอง และระหว่างกลุ่มธุรกิจกับผู้ที่มีอำนาจตัดสินใจในภาครัฐก็มีความสลับซับซ้อนมากกว่าในอดีตเป็นอย่างมาก ซึ่งแนวโน้มการเลือกตั้งในอนาคต ด้วยระบอบ

ประชาธิปไตยและระบบทุนนิยมจะส่งผลให้กลุ่มนักธุรกิจยังคงมีบทบาททางการเมืองต่อไป เนื่องจากมีความพร้อมทางการเงินทุนในการสนับสนุนกิจกรรมทางการเมืองต่าง ๆ และยังคงมีการพัฒนาความสัมพันธ์ระหว่างกลุ่มนักธุรกิจ และนักการเมืองที่ซับซ้อนมากขึ้น

ด้วยเหตุนี้ การศึกษามุมมองของกลุ่มนักธุรกิจในการตัดสินใจการเลือกตั้งในอนาคตครั้งนี้จึงเป็นเรื่องสำคัญมาก เพราะนักธุรกิจเป็นผู้ถือครองปัจจัยความมั่งคั่ง ซึ่งอาจปรับเปลี่ยนอำนาจทางการเมืองได้ง่าย ทั้งนี้ นักธุรกิจเป็นผู้ถือครองและควบคุมเครือข่ายอุปถัมภ์ในชุมชนท้องถิ่น ซึ่งประกอบไปด้วยระบบการซื้อขายต่าง ๆ ซึ่งทั้งหมดนี้เกิดจากการประกอบการของนักธุรกิจและการคลุกคลีอยู่กับประชาชนในพื้นที่ ความสัมพันธ์ระหว่างนักธุรกิจกับประชาชนในชุมชน ดำเนินไปในลักษณะพึ่งพาอาศัยกันและกัน ทั้งนี้ หากมุมมองของนักธุรกิจเกี่ยวกับการตัดสินใจเลือกตั้งในอนาคตไปในทิศทางใด ก็เชื่อได้ว่ามีบทบาทสำคัญในการผลักดันให้เกิดการเปลี่ยนแปลงอำนาจทางการเมืองในทิศทางที่เปิดกว้างมากขึ้น ผลการวิจัยครั้งนี้เชื่อว่าจะก่อให้เกิดผลดีต่อการบริหารบ้านเมืองที่มุ่งต่อประโยชน์ของประชาชนในอนาคต ดังนั้นในการวิจัยครั้งนี้ คณะผู้วิจัยจึงได้นำข้อมูลการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจทั่วทั้งประเทศไทย โดยแบ่งพื้นที่ออกเป็น 7 ภูมิภาค คือ ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ ภาคตะวันออกเฉียงเหนือ ภาคตะวันออก ภาคตะวันตก กรุงเทพมหานคร และภาคใต้ มีกลุ่มตัวอย่างจำนวน 2,917 ตัวอย่าง โดยเป็นการนำข้อมูลมาศึกษาเทคนิควิธีการทำนายความสอดคล้องของการเลือกตั้งกับข้อมูลการตัดสินใจเลือก โดยใช้เทคนิคการทำเหมืองข้อมูล ซึ่งเป็นกระบวนการค้นหาความรู้ ซึ่งเป็นรูปแบบและความสัมพันธ์ที่น่าสนใจในข้อมูล จากฐานข้อมูลขนาดใหญ่ที่ซับซ้อน สามารถตอบปัญหาและค้นพบความรู้ใหม่จากฐานข้อมูลที่มีอยู่ โดยมีเทคนิคการกำหนดกลุ่มข้อมูล ได้แก่ การจัดกลุ่ม (Clustering) เป็นการรวมกลุ่มของข้อมูลที่มีลักษณะคล้ายกันหรือใกล้เคียงกันไว้ในกลุ่มเดียวกัน การหาความสัมพันธ์ (Association) เป็นการแสดงความสัมพันธ์ของข้อมูลทั้งสองชุดหรือมากกว่าไว้ (Jaroenpuntaruk, 2015)

บทความนี้เป็นการนำเสนอการเปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลเพื่อการทำนายระหว่างผลการตัดสินใจเลือกและข้อมูลของนักธุรกิจ โดยใช้เทคนิควิธีทั้งหมด 4 แบบ ได้แก่ วิธีต้นไม้ตัดสินใจ (Decision Tree) วิธี Bayes วิธีถดถอยโลจิสติกส์ (Logistic Regression) และวิธีแรนดอมฟอเรสต์ (Random Forest) เพื่อหาเทคนิควิธีที่มีความถูกต้องและมีความแม่นยำมากที่สุดและนำไปใช้ในการพัฒนารูปแบบหรือวิธีการจัดการข้อมูลเพื่อใช้ประกอบการพิจารณาให้มีความเหมาะสมและเป็นประโยชน์ต่อการพัฒนาประเทศไทย

วัตถุประสงค์การวิจัย

เพื่อเปรียบเทียบประสิทธิภาพโมเดลการจำแนกข้อมูลเพื่อทำนายการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจไทย

แนวคิดและทฤษฎี

ในประเทศประชาธิปไตย พรรคการเมืองเป็นช่องทางที่จะแสดงออกความต้องการของประชาชนหรือรวบรวมข้อเสนอต่าง ๆ สู่รัฐบาล เป็นเครื่องมือที่ทำให้การเลือกตั้งเป็นไปด้วยดีและมีประสิทธิภาพเป็นเครื่องมือในการควบคุมการบริหารราชการแผ่นดิน ฯลฯ ซึ่งส่วนมากเป็นกิจกรรมและบทบาทในทาง

ส่งเสริมการปกครองให้บรรลุปรัชญาของระบอบประชาธิปไตยยิ่งขึ้น ในส่วนของประเทศไทย ถึงขณะนี้ การต่อสู้ทางการเมืองยังคงจบลงโดยการรัฐประหาร หลายสิบปีที่ผ่านมา พรรคการเมืองไทยมีสถานภาพ ที่ไม่แน่นอน ในหลายช่วงเวลาพรรคการเมืองดำรงอยู่ได้ก็เนื่องด้วยกองทัพและผู้นำทหารยินยอม บ่อยครั้งมีการใช้กำลังความรุนแรงเข้ายึดอำนาจทางการเมืองและยุติบทบาทของพรรคการเมืองที่ถูกต้อง ตามกฎหมาย ปรากฏการณ์ดังกล่าวส่งผลให้เกิดการแสดงพลังทางการเมืองโดยมีผู้ถือธงชาติรวมกลุ่ม เดินขบวนเรียกร้องให้ยุติการทำร้ายประเทศไทยมาอย่างต่อเนื่อง ซึ่งกลุ่มดังกล่าวประกอบด้วยคน หลากหลายอาชีพ และส่วนหนึ่งเป็นกลุ่มที่เรียกว่า The silent majority คือ กลุ่มนักธุรกิจจำนวนหนึ่งที่ ร่วมอยู่ในขบวนการดังกล่าวด้วย ซึ่งโดยทั่ว ๆ ไปนักธุรกิจมักจะระมัดระวังในการแสดงออก แต่เมื่อถึงจุด วิฤติสำคัญ นักธุรกิจ และสถาบันทางธุรกิจก็จะออกมาเรียกร้องหรือร่วมขบวนเพื่อต้องการที่จะให้ “การเมืองนิ่ง” และ “มีเสถียรภาพ” ส่งผลให้การประกอบธุรกิจโดยเฉพาะอย่างยิ่งการลงทุนจาก ต่างประเทศและการท่องเที่ยวสามารถเจริญรุ่งเรืองต่อไปได้ นอกจากนี้ยังมีประเด็นสำคัญอีกประการ หนึ่ง คือ การเรียกร้องให้มีการเลือกตั้ง เพราะถือว่าเป็นวิธีการที่ดีที่สุดในการระบอบประชาธิปไตย การ เลือกตั้งนั้นเป็นกิจกรรมที่แสดงออกถึงเจตนาของการเข้าไปมีส่วนร่วมทางการเมืองของประชาชนทุกกลุ่ม อาชีพ เพื่อต้องการที่จะเลือกผู้แทนที่มีความคิด ความเห็น นโยบาย และอุดมการณ์คล้ายคลึงกับตนมาก ที่สุด เพื่อเข้าไปทำหน้าที่แทนตนเพื่อประโยชน์แก่ตนให้เกิดความกินดีอยู่ดีในฐานะที่เป็นเจ้าของประเทศ (limmane, 2015)

การทำเหมืองข้อมูล (Data Mining) เป็นการวิเคราะห์ข้อมูลเพื่อแยกประเภท จำแนกรูปแบบ และความสัมพันธ์ของ ข้อมูลจากฐานข้อมูลที่มีขนาดใหญ่ หรือคลังข้อมูล โดยมีวิธีต่าง ๆ หลายวิธีซึ่ง รูปแบบการทำเหมืองข้อมูลนั้นได้รวบรวม ความรู้จากหลายแขนงเข้าไว้ด้วยกันประกอบด้วย การเรียนรู้ ของเครื่องมือ (Machine Learning) ร่วมกับวิทยาศาสตร์ สารสนเทศ (Information Science) สถิติ (Statistics) เทคนิคการสร้างภาพ (Visualization Techniques) ทำให้ค้นพบ รูปแบบความรู้ใหม่ ๆ และ สามารถนำเสนอความรู้ในลักษณะที่เข้าใจได้ง่าย ในอีกความหมายหนึ่งการทำเหมืองข้อมูล คือกระบวนการ ที่กระทำกับข้อมูลจำนวนมากเพื่อสกัดสารสนเทศ รูปแบบ และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูล ขนาดใหญ่ (Han et al., 2022) ซึ่งมีหลายเทคนิค แต่ในการวิจัยครั้งนี้นำมาใช้ 4 เทคนิค ดังนี้

ต้นไม้ตัดสินใจ (Decision Tree) เป็นเทคนิคหนึ่งของ Classification ที่ได้รับความนิยมนำมาใช้ ในการแบ่งประเภทหรือแยกหมวดหมู่ของข้อมูลโดย มีลักษณะคล้ายโครงสร้างต้นไม้โดย จะแสดงการเก็บ คุณสมบัติของคุณลักษณะ แสดงเงื่อนไขการทดสอบ โดยแสดงกลุ่มข้อมูล ที่กำหนดไว้ซึ่งจะมีการเรียนรู้ จากข้อมูลการสอนในระบบ (Training Data) และใช้การคำนวณค่า Information Gain สำหรับ แต่ละ คุณลักษณะโดยแบ่งข้อมูลตามลำดับ แต่ข้อจำกัด คือ ไม่สามารถจำแนกข้อมูลได้อย่างต่อเนื่องจะการ จำแนกข้อมูลที่มีอยู่ในกฎเท่านั้น

วิธีแรนดอมฟอเรส (Random Forest) เป็นโครงสร้างข้อมูลแบบต้นไม้ตัดสินใจ (Decision Tree) โดยจะใช้ วิธีการสุ่มตัวแปรแล้วนำผลที่ได้ข้อมูลที่ซ้ำกันออก ซึ่งวิธีการนี้ช่วยจำแนกข้อมูลได้อย่าง ต่อเนื่อง และผลการพยากรณ์มีความแม่นยำและลดการความซ้ำซ้อนของข้อมูลทำให้สามารถสร้างตัว ทำนายที่เหมาะสมและแม่นยำ

วิธีการถดถอยโลจิสติกส์ (Logistic Regression) เป็นเทคนิคการวิเคราะห์ตัวแปรเชิงพหุ เป็นการ ทำนายค่าความน่าจะเป็นของการเกิดเหตุการณ์ หรือไม่เกิดเหตุการณ์ ความสัมพันธ์จะอยู่ในรูปแบบของ สมการเส้นถดถอย (Regression Equation) มีตัวแปรตามที่มีค่าเพียง 2 ค่า คือ เป็น 0 และ 1 ตัวแปรอิสระ

อาจจะมีตัวเดียวหรือหลายตัวก็ได้ ค่าพยากรณ์หรือค่าตัวแปรจะมีค่าอยู่ที่ 0 กับ 1 โดยที่ถ้าค่าความน่าจะเป็น >0.5 ถือว่าค่าพยากรณ์คือ 1 แต่ถ้าค่าความน่าจะเป็น <0.5 ถือว่าค่าพยากรณ์ คือ 0

วิธี Bayes อาศัยหลักการความน่าจะเป็น (Probability) ตามทฤษฎีเบย์ (Bayes Theorem) ซึ่งมีอัลกอริทึมที่ไม่ซับซ้อน เป็นขั้นตอนวิธีในการจำแนกข้อมูล โดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ เป็นการจำแนกข้อมูลโดยใช้ความน่าจะเป็นและคำนวณการแจกแจงความน่าจะเป็นตามสมมติฐานที่ตั้งให้กับข้อมูลจากการคำนวณตัวอย่างใหม่ที่ได้จะถูกนำมาปรับเปลี่ยนการแจกแจง ซึ่งมีผลต่อการเพิ่มหรือลดความน่าจะเป็นของข้อมูล ข้อมูลใหม่ที่เกิดขึ้นและตัวแบบที่ตั้งไว้ให้กับข้อมูลจะถูกปรับเปลี่ยนไปตามข้อมูลใหม่โดยผนวกกับข้อมูลเดิมที่มี

ระเบียบวิธีวิจัย

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย คือ นักธุรกิจไทย ทั้งนี้เพราะนักธุรกิจมีภาระหน้าที่ที่สำคัญโดยต้องมีส่วนร่วมในการพัฒนาเศรษฐกิจอย่างยั่งยืน การยกระดับคุณภาพชีวิตให้แก่แรงงาน ชุมชน และสังคมตามหลักการเพื่อผลประโยชน์ของสถานประกอบการและการพัฒนาของสังคม และสอดคล้องกับนโยบายของรัฐบาล อย่างไรก็ตามในสถานการณ์วิกฤตทางการเมืองในประเทศ อาจส่งผลให้เกิดภาวะวิกฤตทางเศรษฐกิจของประเทศตามมา และลุกลามไปยังต่างประเทศ กล่าวคือ นักธุรกิจมีหน้าที่หลักคือพัฒนาการผลิตประกอบธุรกิจ และมีบทบาทสำคัญทางการเมืองผ่านการเลือกตั้ง เพื่อให้ประเทศมีรัฐบาลที่เข้าใจนโยบายทางเศรษฐกิจ สามารถพัฒนาประเทศผ่านการส่งออกสินค้าอุปโภคและบริโภคให้แก่ประเทศ ดังนั้น นักธุรกิจหลาย ๆ รายจึงได้ปรับปรุงรายการการลงทุนประกอบธุรกิจตามนโยบายของพรรคการเมืองที่ร่วมกันจัดตั้งรัฐบาล (Teeravakin, 2009) ทั้งนี้คัดเลือกกลุ่มตัวอย่างในการวิจัยครั้งนี้ กำหนดจากนักธุรกิจที่จดทะเบียนพาณิชย์ ณ สำนักงานพาณิชย์จังหวัดทุกจังหวัดทั่วประเทศ โดยใช้สูตรทาร์ยามาเน่ ในแต่ละภูมิภาคของประเทศ โดยแบ่งเป็น 1) ภาคเหนือ มี 16 จังหวัด กลุ่มตัวอย่างจำนวน 400 ชุด 2) ภาคตะวันออกเฉียงเหนือ มี 20 จังหวัด กลุ่มตัวอย่าง จำนวน 400 ชุด 3) ภาคกลาง มี 15 จังหวัด กลุ่มตัวอย่าง จำนวน 405 ชุด 4) ภาคตะวันออกมี 7 จังหวัด กลุ่มตัวอย่าง จำนวน 406 ชุด 5) ภาคตะวันตก มี 4 จังหวัด กลุ่มตัวอย่าง จำนวน 400 ชุด 6) ภาคใต้ มี 14 จังหวัด กลุ่มตัวอย่าง จำนวน 406 ชุด และ 7) พื้นที่กรุงเทพมหานคร กลุ่มตัวอย่าง จำนวน 500 ชุด รวมจำนวนทั้งหมด 2,917 ชุด

วิธีดำเนินงานวิจัย

ขั้นตอนดำเนินการวิจัยได้ใช้กระบวนการทำเหมืองข้อมูลของ CRISP-DM (Sinsomboonthong, 2015) ดังต่อไปนี้

1) การทำความเข้าใจการวิจัย โดยการศึกษาข้อมูลจากองค์ประกอบของการพิจารณาการตัดสินใจเลือกตั้ง และข้อมูลการตัดสินใจเลือกของนักธุรกิจไทย โดยกำหนดรายชื่อคุณลักษณะที่นำมาสร้างตัวต้นแบบ ซึ่งมีจำนวน 13 ปัจจัย เพื่อใช้ในการทดลอง

2) การเตรียมข้อมูลสำหรับการทำเทคนิคดาต้าไมนิง (Data Mining) หลังจากที่ได้ข้อมูลมาแล้วมีขั้นตอนดังนี้

การคัดเลือกข้อมูล (Data Preprocessing) เป็นการนำข้อมูลมาผ่านกระบวนการ ตรวจสอบเพิ่ม และปรับปรุงรายการข้อมูลที่ไม่มีความสัมพันธ์กันออกจากชุดข้อมูล ซึ่งการวิเคราะห์และหาความสัมพันธ์ของตัวแปรจากปัจจัยที่รวบรวมมาจำนวน 13 ปัจจัยรวมการตัดสินใจการเลือกตั้งในอนาคต (จำนวน 2,917 รายการ) เมื่อทำการวิเคราะห์หาความสัมพันธ์ของปัจจัยทั้งหมดด้วยวิธีการวิเคราะห์สหสัมพันธ์เชิงอันดับของ สเปียร์แมน (Spearman Rank Correlation) ซึ่งเป็นเทคนิคทางสถิติที่ใช้หาขนาดและทิศทางของความสัมพันธ์ระหว่างการจัดลำดับของหลายสิ่งด้วยสองทางเลือก โดยพบว่า ปัจจัยที่มีความสัมพันธ์และส่งผลต่อการตัดสินใจการเลือกตั้งในอนาคตของนักธุรกิจไทย มีจำนวน 6 ปัจจัย ได้แก่ เพศ อายุ ภูมิภาค ความรู้ความสามารถของผู้สมัคร ส.ส. พรรคที่มีความจงรักภักดี และความซื่อสัตย์ และพรรคที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน โดยค่า R จะมีความสัมพันธ์กัน อย่างมีนัยสำคัญที่ระดับ .05 และ .001 เพราะเป็นระดับนัยสำคัญทางสถิติ คือ ค่าที่แสดงความมั่นใจต่อการสรุปผลได้อย่างถูกต้อง และมีโอกาสที่จะเกิดความคลาดเคลื่อนในการสรุปผลมีเพียง .05 และ .001 ซึ่งโดยทั่วไปแล้วในการวิจัยจะมีการกำหนดค่าของระดับความเชื่อมั่นไว้เท่ากับ 95% 99% หรือ 99.99% ดังแสดงในตารางที่ 1 (Table 1)

Table 1. Characteristics Prototyping Details and Data Analysis.

Factor	Variable	Data Encryption	Correlation	
			R	Sig.
Sex	Sex	1 = Male 2 = Female	.073	.000***
Age	Age	1 = 18-25 years old 2 = 25-35 years old 3 = 36-45 years old 4 = 46-55 years old 5 = 56-65 years old 6 = 65 years old and over	.193	.000***
Domicile (Region)	Domicile	1 = Bangkok 2 = Central 3 = Western 4 = Eastern 5 = Southern 6 = Northern 7 = Northeast (Isan)	.077	.000***
Education Level	Study	1 = Middle school or less 2 = High school/vocational certificate 3 = Associate degree/ high vocational certificate 4 = Bachelor's degree 5 = Post-graduate degree	-.012	.516
Reputation of MP candidate	Reputable	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.031	.092
Knowledge and abilities of MP candidates	Knowledge	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.043	.022*
Past performance of MP candidates	Past Works	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.012	.503

Factor	Variable	Data Encryption	Correlation	
			R	Sig.
Loyalty to the institution and honesty of political parties	Royal	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.122	.000***
The ability of political parties to raise money for national development	Earn Money	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	.029	.118
Reliable and trustworthy political parties led by reliable leaders and executive committee	Board Director	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	.020	.269
Parties with practical policies and transparent budgets.	Executable	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.083	.000***
Parties with populist policies	Populist	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	.030	.106
Parties with national reform policy	Reform	Little = Has a minimal effect on decision-making. Moderate = Has a reasonable influence on decision-making. Lot = Has a significant effect on decision-making.	-.018	.323
Decision	Decision	1 = (Vote) 2 = (No Vote)		

Source: Statistical package

2.2) การทำข้อมูลให้สมบูรณ์ (Data Cleaning) จากการตรวจสอบทุกข้อมูลไม่มีความผิดพลาด

2.3) การเปลี่ยนรูปแบบ (Data Transformation) เพื่อให้มีความเหมาะสมและพร้อมที่จะนำไปใช้งาน

3) การสร้างตัวต้นแบบเลือกใช้เทคนิคการสร้างต้นแบบโดยใช้แบบจำแนกข้อมูล จัดทำตัวแบบให้เป็นมาตรฐานเพื่อให้ได้ผลลัพธ์ที่ดีที่สุดเลือกใช้เทคนิคที่แตกต่างกันสำหรับปัญหาการทำเหมืองข้อมูลที่เหมือนกันได้ โดยใช้เทคนิควิธีโครงสร้างต้นไม้การตัดสินใจ (Decision Tree) วิธี Bayes วิธีถดถอยโลจิสติกส์ (Logistic Regression) และวิธีเร้นดอมฟอเรส (Random Forest) ด้วยวิธีการ 10-Fold Cross Validation โดยแบ่งข้อมูลออกเป็น 10 ชุดเท่ากัน ใช้ 1 ชุดเป็นกลุ่มทดสอบข้อมูลอีก 9 ชุดเป็นกลุ่มที่ใช้ในการเรียนรู้และทำการวนซ้ำจำนวน 10 รอบ โดยเปลี่ยนกลุ่มทดสอบไปเรื่อย ๆ จนครบดังภาพที่ 1-4 (Figures 1-4)

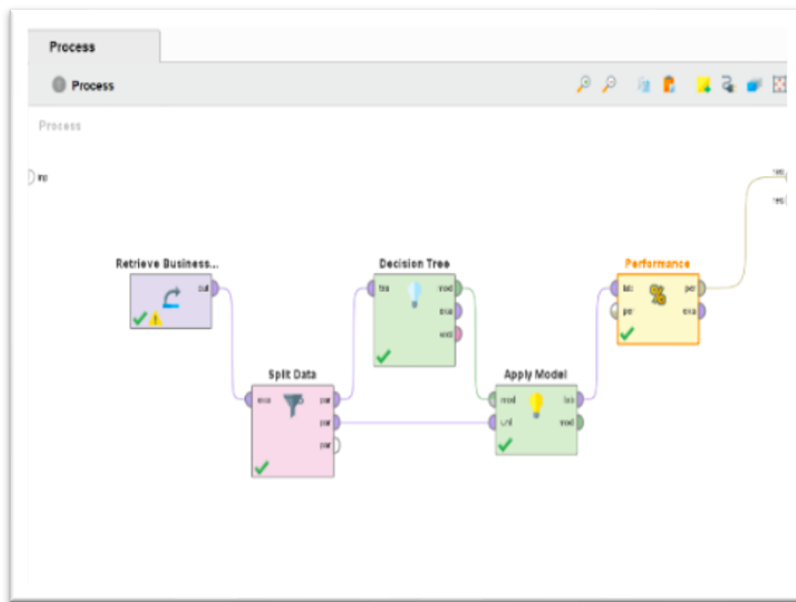


Figure 1. Decision Tree Model

Source: Rapid Miner Studio 10.1

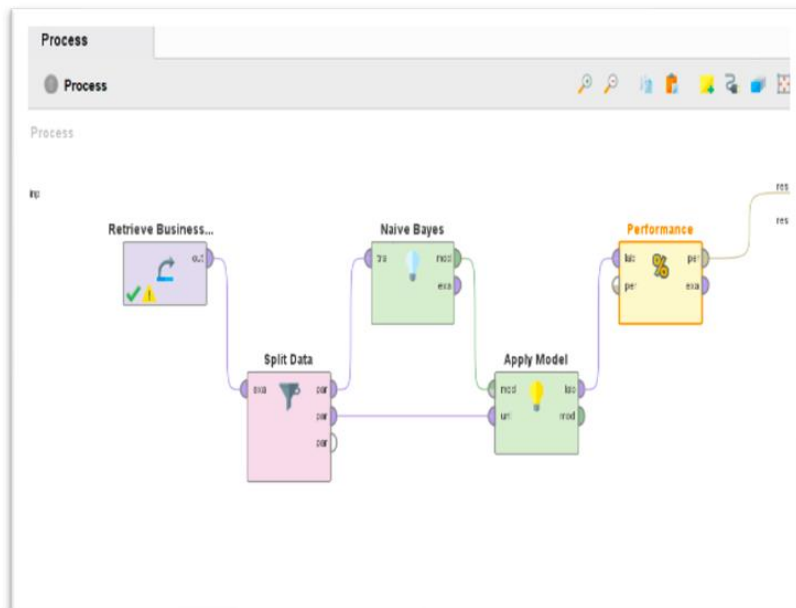


Figure 2. Decision Tree Model

Source: Rapid Miner Studio 10.1

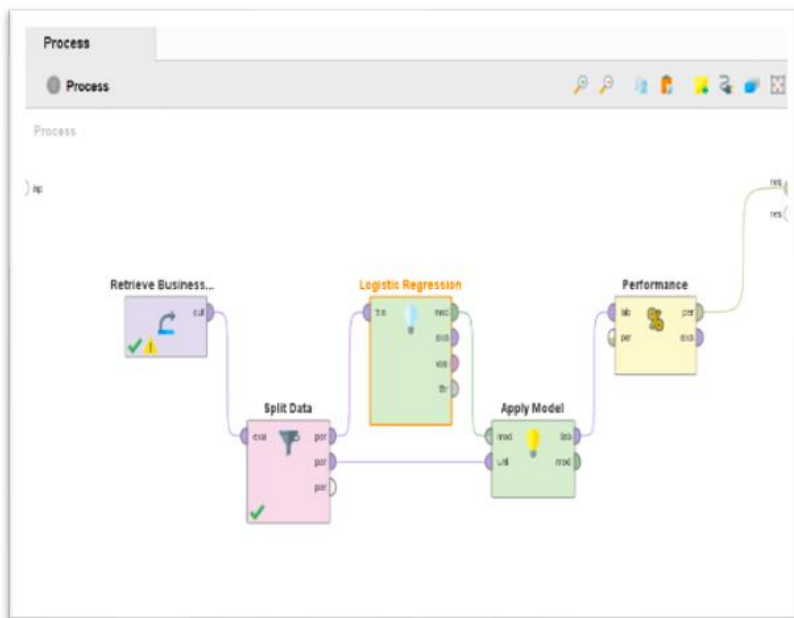


Figure 3. Logistic Regression Model
Source: Rapid Miner Studio 10.1

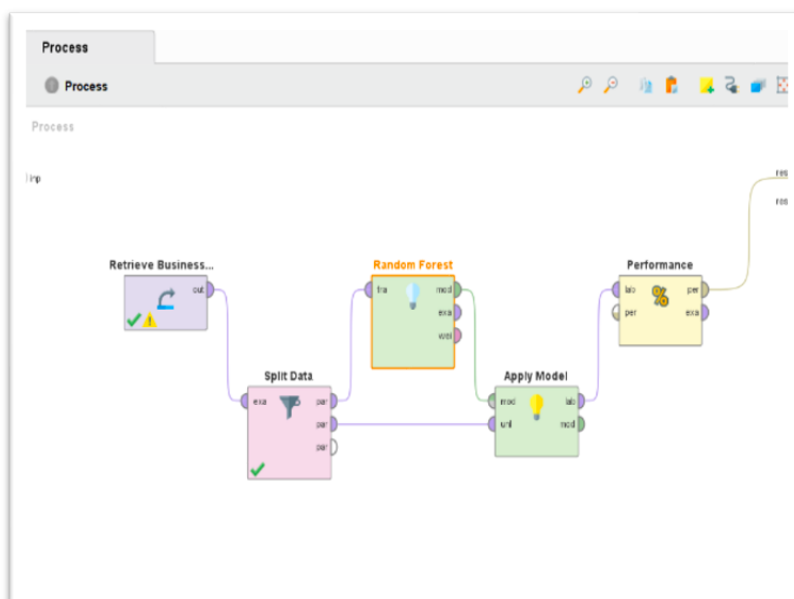


Figure 4. Random Forest Model
Source: Rapid Miner Studio 10.1

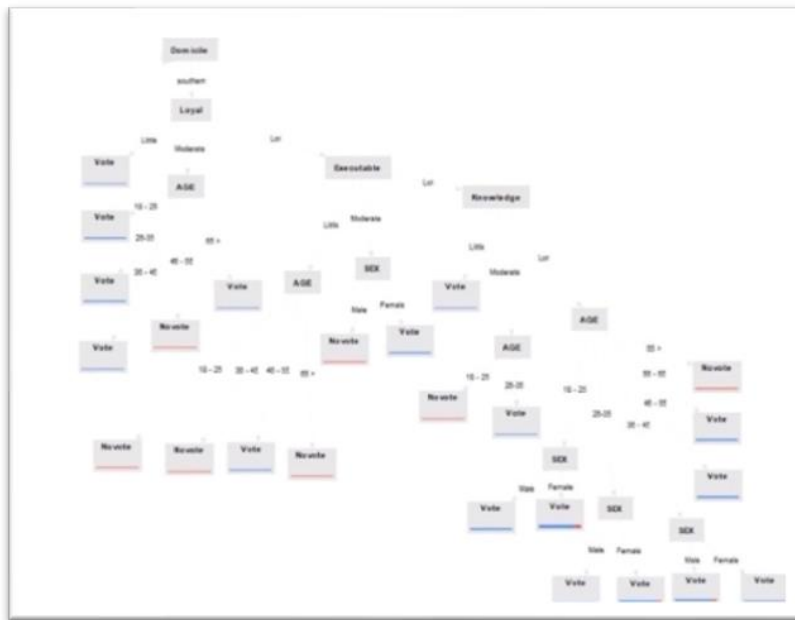


Figure 7. Random Forest Model for the South

Source: Rapid Miner Studio 10.1



Figure 8. Random Forest Model for the Northeast

Source: Rapid Miner Studio 10.1

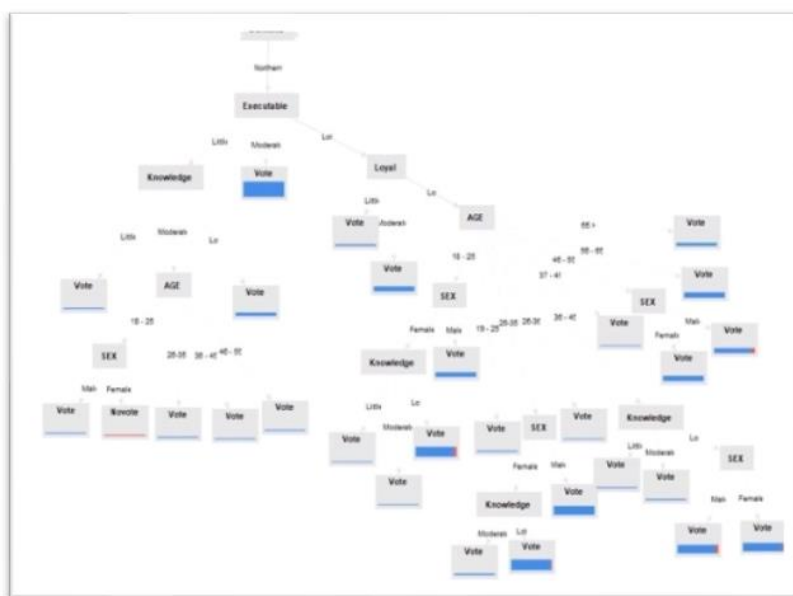


Figure 9. Random Forest Model for the North

Source: RapidMiner Studio 10.1

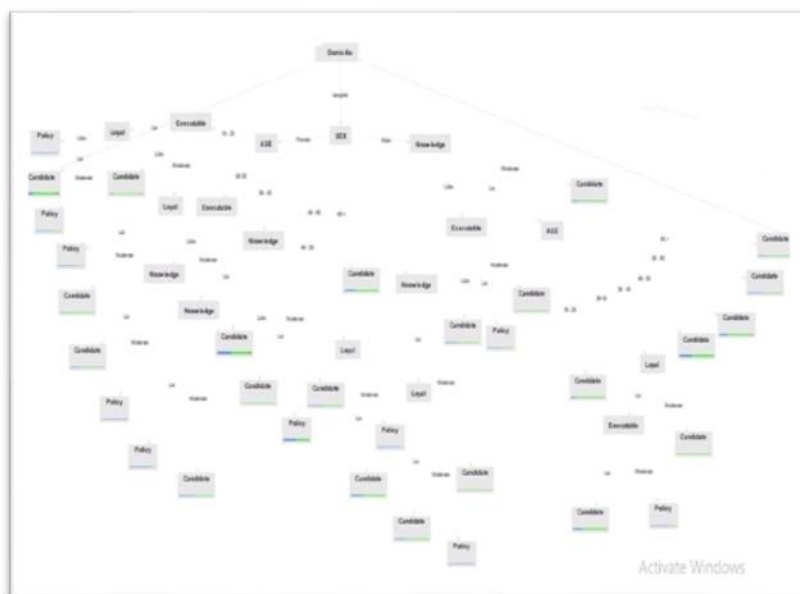


Figure 10. Random Forest Model for the Central

Source: RapidMiner Studio 10.1



Figure 11. Random Forest Model for Bangkok

Source: Rapid Miner Studio 10.1

4) การประเมินตัวแบบหาตัวแบบที่ได้บรรลุวัตถุประสงค์สร้างกฎเกณฑ์ที่สำคัญของปัญหาการวิจัยพิจารณาเพื่อตัดสินใจจากประโยชน์ที่ได้โดยวิเคราะห์ด้วยวิธี 10-Fold Cross Validation เพื่อให้ได้ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ที่ได้ค่าจากประมวลผลตามตัวแบบ (Model)

5) ขั้นตอนการใช้งานสร้างตัวแบบที่ยังไม่สามารถทำให้สมบูรณ์แบบได้หรือการสร้างรายงานที่เหมาะสมไปใช้งานจริงเพื่อวิเคราะห์และแก้ปัญหาที่เกิดขึ้นตามความต้องการ

เครื่องมือที่ใช้ในการเก็บข้อมูล

1) ในการเก็บข้อมูลจากแบบสำรวจการตัดสินใจเลือกตั้งในอนาคต การบันทึกข้อมูลผ่านโปรแกรมสำเร็จรูปทางสถิติ โดยบันทึกข้อมูลเก็บไว้ในฐานข้อมูล Businessman Decision

2) ใช้โปรแกรม Rapid Miner Studio 10.1 ในการวิเคราะห์ข้อมูลการทำนายการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจไทย

การวิเคราะห์ข้อมูล

นำข้อมูลมาทดสอบผลการทำนายตามตัวแบบ (Model) และเปรียบเทียบประสิทธิภาพด้วยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree) วิธี Bayes วิธีถดถอยโลจิสติกส์ (Logistic Regression) และวิธีแรนดอมฟอร์เรส (Random Forest)

สถิติที่ใช้ในการวิจัย

ในการพิจารณาผลการทำนายการตัดสินใจเลือกตั้งในอนาคตจะดูผลจากค่าร้อยละสูงที่สุดประกอบด้วย

1) ค่าความถูกต้อง (Accuracy) คือจำนวนที่ทำนายถูกของทุกคลาส ใช้ $Accuracy = \frac{TP+TN}{N}$

2) ค่าความแม่นยำ (Precision) คือผลการคำนวณที่ทำนายถูกร้อยละเท่าใด ใช้ $Precision = \frac{TP}{TP+FP}$

3) ค่าความระลึก (Recall) คือจำนวนที่ทำนายถูกที่ตัวใช้ $Recall = \frac{TP}{TP+FN}$

ค่า TP (True Positive) คือ จำนวนผู้ตัดสินใจเลือกองค์ประกอบของการเลือกตั้งและได้ตอบเลือก
ค่า FP (False Positive) คือ จำนวนผู้ตัดสินใจเลือกองค์ประกอบของการเลือกตั้งแต่ได้ตอบไม่เลือก
ค่า FN (False Negative) คือ จำนวนผู้ที่ไม่ได้ตัดสินใจเลือกองค์ประกอบของการเลือกตั้งแต่ได้ตอบเลือก
และค่า TN (True Negative) คือ จำนวนผู้ที่ไม่ได้ตัดสินใจเลือกองค์ประกอบของการเลือกตั้งและได้ตอบไม่เลือก

การนำโมเดลต้นแบบไปใช้งานจริงได้นั้น จำเป็นต้องทราบประสิทธิภาพของโมเดลต้นแบบก่อน โดยทั่วไปแล้วจะมีตัววัดที่นิยมใช้กันในงานวิจัยและการทำงานต่าง ๆ อยู่ 4 ค่า ดังนี้ ค่าความแม่นยำ (Precision) คือ ค่าที่ดูสิ่งที่ทำนายออกมาแล้วทายถูกได้กี่เปอร์เซ็นต์ ค่าความระลึก (Recall) คือ จำนวนที่ทำนายถูกที่ตัว เป็นการวัดความถูกต้องของโมเดล ค่าความถ่วงดุล (F-measure) คือ ค่าเฉลี่ยของค่าความแม่นยำและค่าความระลึก ค่าความถูกต้อง (Accuracy) คือ จำนวนข้อมูลที่ทำนายถูกทุกคลาส เป็นการวัดความถูกต้องของโมเดล โดยพิจารณารวมทุกคลาส โดยต้องมีค่าสูงที่สุด

จากการทดลอง สามารถอธิบายผลการทดลองได้ดังนี้

1) ผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูล ซึ่งเมื่อทำการเลือกปัจจัยที่เกี่ยวข้องกับชุดข้อมูลแล้ว ได้นำชุดข้อมูลทั้งหมด 2,917 ชุด มาทดสอบประสิทธิภาพของโมเดลการจำแนกข้อมูล เพื่อการทำนายด้วยเทคนิควิธีต้นไม้ตัดสินใจ (Decision Tree) วิธี Bayes วิธีถดถอยโลจิสติก (Logistic Regression) และวิธีแรนดอมฟอเรส (Random Forest) เพื่อทำการเปรียบเทียบประสิทธิภาพในการทำจำแนกข้อมูล ซึ่งได้ค่าประสิทธิภาพ ดังแสดงในตารางที่ 2 (Table 2.)

Table 2. Comparison of Data Classification Efficiency to Predict of 4 Techniques

Algorithms	Accuracy (%)	Precision (%)	Recall (%)
Random forest	88.85	99.23	89.43
Decision tree	88.51	99.81	88.64
Naïve Bayes	88.46	100.00	88.46
Logistic Regression	88.11	99.61	88.62

Source: Rapid Miner Studio 10.1

จาก Table 2. พบว่า ตัวแบบที่สร้างจากเทคนิค วิธีแรนดอมฟอเรส (Random Forest) มีค่าความถูกต้องมากที่สุดคือร้อยละ 88.85 ค่าความแม่นยำร้อยละ 99.23 และค่าความระลึกร้อยละ 89.43 รองลงมา คือ วิธีต้นไม้ตัดสินใจ (Decision Tree) มีค่าความถูกต้องสูงคือร้อยละ 88.51 ค่าความแม่นยำร้อยละ 99.81 และค่าความระลึกร้อยละ 88.64 วิธี Bayes มีค่าความถูกต้องสูงร้อยละ 88.46 ค่าความ

แม่นยำร้อยละ 100.00 และค่าความระสีร้อยละ 88.46 และวิธีถดถอยโลจิสติกส์ (Logistic Regression) มีค่าความถูกต้องน้อยที่สุดคือร้อยละ 88.11 ค่าความแม่นยำร้อยละ 99.61 และค่าความระสีร้อยละ 88.62 ตามลำดับ

2) การสร้างโมเดล

ในการศึกษาค้นคว้าครั้งนี้คณะผู้วิจัยได้ทำการสร้างโมเดลด้วยโดยใช้เทคนิควิธีโครงสร้างต้นไม้การตัดสินใจ (Decision Tree) วิธีแบบเบย์ (Naïve Bayes) วิธีถดถอยโลจิสติกส์ (Logistic Regression) และวิธีแรนดอมฟอเรส (Random Forest) โดยใช้ปัจจัยได้จากการหาความสัมพันธ์ทั้งหมด 6 ปัจจัย ประกอบด้วย เพศ อายุ ภูมิลำเนา ความรู้ความสามารถของผู้สมัคร ส.ส. พรรคที่มีความจงรักภักดี และความซื่อสัตย์ และพรรคที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง ผลลัพธ์ของต้นไม้ตัดสินใจ ดังภาพที่ 5-12 (Figures 5-12)

จากภาพที่ 5-12 โมเดลที่สร้างจากเทคนิคแรนดอมฟอเรส (Random Forest) โดยใช้โปรแกรม Rapid Miner Studio 10.1 พบว่า ปัจจัยที่เป็นโหนดหลักที่มีอิทธิพลต่อการตัดสินใจเลือกตั้งในอนาคตของนักธุรกิจไทย คือ ภูมิลำเนา (ภาค) เป็นตัวกำหนดการทำงานของโมเดล ซึ่งนักธุรกิจที่มีภูมิลำเนาอยู่ในกรุงเทพฯ ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ ภาคตะวันออก ภาคตะวันตก และภาคใต้ ต่างก็มีมุมมองว่าพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน จะส่งผลต่อการตัดสินใจเลือกตั้งมาก ในขณะที่นักธุรกิจบางส่วนมีมุมมองว่า การที่พรรคการเมืองมีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจนก็จะส่งผลต่อการตัดสินใจเลือกตั้งพอสมควร แต่กระนั้นก็ตามจะพิจารณาถึงความรู้ความสามารถของผู้สมัคร ส.ส. ด้วย หากมีความรู้ความสามารถสูงก็จะส่งผลต่อการตัดสินใจเลือกตั้งมาก ในทางตรงข้ามหากมีความรู้ความสามารถน้อยก็จะตัดสินใจไม่เลือก นอกจากนี้ผู้สมัคร ส.ส. ที่มีระดับความรู้ความสามารถอยู่ในระดับปานกลาง นักธุรกิจก็มีมุมมองที่จะพิจารณาถึงพรรคการเมืองที่มีความจงรักภักดี และมีความซื่อสัตย์ หากมีระดับพอสมควรหรือมีระดับมาก ก็จะส่งผลต่อการตัดสินใจเลือกตั้ง แต่หากมีบ้างก็จะตัดสินใจไม่เลือกตั้ง ทั้งนี้เมื่อพิจารณาถึงกฎที่ถูกสร้างขึ้นพบว่า มีเฉพาะค่าคุณลักษณะประจำทั้ง 6 ปัจจัย คือ เพศ (Sex) อายุ (Age) ภูมิลำเนา (ภาค) (Domicile) พรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน (Executable) ความรู้ความสามารถของผู้สมัคร ส.ส. (Knowledge) และพรรคการเมืองที่มีความจงรักภักดี และมีความซื่อสัตย์ (Royal) ซึ่งจะนำเสนอกฎที่ได้พอเป็นตัวอย่าง มีดังนี้

กฎข้อที่ 1 ถ้านักธุรกิจที่มีภูมิลำเนาอยู่ในพื้นที่กรุงเทพฯ และเห็นด้วยกับพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน เท่ากับมีผลต่อการตัดสินใจเลือกพรรคการเมืองนั้นมาก ด้วยค่าความเชื่อมั่น ที่ 0.99 ค่าสนับสนุนที่ 0.47 และค่าสหสัมพันธ์ที่ 2.32

กฎข้อที่ 2 ถ้านักธุรกิจที่มีภูมิลำเนาอยู่ในพื้นที่กรุงเทพฯ และเห็นด้วยกับพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน มีผลต่อการตัดสินใจเลือกพรรคการเมืองนั้นบ้าง ด้วยค่าความเชื่อมั่น ที่ 0.99 ค่าสนับสนุนที่ 0.45 และค่าสหสัมพันธ์ที่ 2.48

กฎข้อที่ 3 ถ้านักธุรกิจที่มีภูมิลำเนาอยู่ในพื้นที่กรุงเทพฯ และเห็นด้วยกับพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน มีผลต่อการตัดสินใจเลือกพรรคการเมืองนั้นพอสมควร และหากผู้สมัคร ส.ส. ที่มีความรู้ความสามารถอยู่ในระดับมาก มีผลต่อการตัดสินใจเลือก ด้วยค่าความเชื่อมั่น ที่ 0.99 ค่าสนับสนุนที่ 0.42 และค่าสหสัมพันธ์ที่ 2.68

Table 3. Association Rule of Decision Making in the Future Election for Thai Businesspersons

Association Rules	Confidence Level	Key Performance Indicators	Pearson Correlation
If domicile = Bangkok and executable = Lot then “Vote”	0.99	0.47	2.32
If domicile = Bangkok and executable = Little then Vote	0.99	0.45	2.48
If domicile = Bangkok and executable = Moderate and knowledge = Lot then “Vote”	0.99	0.42	2.68
If domicile = Bangkok and executable = Moderate and knowledge = Little, then “NoVote”	0.98	0.40	2.63
If domicile = Bangkok and executable = Moderate and knowledge = Moderate and loyal = Little then “NoVote”	0.97	0.38	2.62

Source: RapidMiner Studio 10.1

กฎข้อที่ 4 ถ้านักธุรกิจที่มีภูมิลำเนาอยู่ในพื้นที่กรุงเทพฯ และเห็นด้วยกับพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน มีผลต่อการตัดสินใจเลือกพรรคการเมืองนั้นพอสมควร และหากผู้สมัคร ส.ส. ที่มีความรู้ความสามารถน้อย มีผลต่อการตัดสินใจไม่เลือก ด้วยค่าความเชื่อมั่น ที่ 0.98 ค่าสนับสนุนที่ 0.40 และค่าสหสัมพันธ์ที่ 2.63

กฎข้อที่ 5 ถ้านักธุรกิจที่มีภูมิลำเนาอยู่ในพื้นที่กรุงเทพฯ และเห็นด้วยกับพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน มีผลต่อการตัดสินใจเลือกพรรคการเมืองนั้นพอสมควร และหากผู้สมัคร ส.ส. ที่มีความรู้ความสามารถพอสมควรและมีความจงรักภักดีและซื่อสัตย์น้อย มีผลต่อการตัดสินใจไม่เลือก ด้วยค่าความเชื่อมั่น ที่ 0.97 ค่าสนับสนุนที่ 0.38 และค่าสหสัมพันธ์ที่ 2.62

ผลการวิจัย

งานวิจัยนี้เป็นการนำเหมืองข้อมูลมาประยุกต์ใช้ในการตัดสินใจ (Decision Data Mining) โดยมีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพแบบจำลองการจำแนกข้อมูลเพื่อทำนายการตัดสินใจ การเลือกตั้งในอนาคตของนักธุรกิจไทย โปรแกรมที่ใช้ในการวิเคราะห์ข้อมูลการ คือ Rapid Miner Studio 10.1 เทคนิคดาต้าไมนนิ่ง 4 วิธี ได้แก่ การจำแนกข้อมูลด้วยวิธีต้นไม้ตัดสินใจ วิธีแบบเบย์ วิธีแรนดอมฟอรัล และวิธีถดถอยโลจิสติกส์ สถิติที่ใช้ในการวิจัย ได้แก่ สถิติการทำนายโดยพิจารณา ค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก โดยต้องมีค่าสูงที่สุด ผลการทดสอบประสิทธิภาพด้วยวิธีการ 10-Fold cross-validation เพื่อทำการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูล 4 เทคนิค จากผลการทดลอง พบว่า วิธีแรนดอมฟอรัลให้ค่าความถูกต้องสูงที่สุด คิดเป็นร้อยละ 88.85 รองลงมาคือ วิธีต้นไม้ตัดสินใจ คิดเป็นร้อยละ 88.51 และวิธีแบบเบย์ คิดเป็นร้อยละ 88.46 และวิธีถดถอยโลจิสติกส์ คิดเป็นร้อยละ 88.11 ตามลำดับ ซึ่งจากการศึกษางานวิจัยข้างต้น พบว่า เทคนิคการทำเหมืองข้อมูลเป็นเทคนิค

ที่มีความสำคัญในการค้นหาสารสนเทศหรือความรู้ที่ซ่อนอยู่ในสารสนเทศเพื่อนำมาใช้ประโยชน์ในการตัดสินใจทางด้านศาสตร์ต่าง ๆ เช่น ด้านการศึกษา ด้านการเงินการธนาคาร หรือด้านการตลาด แต่ยังไม่มีการวิจัยใดที่นำเทคนิคการทำเหมืองข้อมูลมาเป็นเทคนิคในการค้นหาการตัดสินใจของผู้มีสิทธิ์เลือกตั้งในระดับท้องถิ่นหรือระดับชาติ ดังนั้นจากผลการวิจัยนี้สามารถชี้ชัดได้ว่า ตัวแบบซึ่งได้จากการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลในงานวิจัยนี้ เป็นตัวแบบที่เหมาะสมที่จะนำไปพัฒนาเพื่อให้สามารถนำไปใช้ทำนายเพื่อวิเคราะห์แนวทางในการเสนอแนะ และนำไปใช้ประโยชน์ประกอบการพิจารณาการเลือกตั้งในอนาคต ซึ่งข้อมูลที่สำคัญนี้สามารถสะท้อนไปยังพรรคการเมือง และนักการเมืองที่อยู่ในระดับต่าง ๆ ต่อไปได้ ทั้งนี้โดยทั่วไปนักธุรกิจที่มักจะระมัดระวังในการแสดงออก ส่วนใหญ่ต้องการทำมาหากินโดยไม่เกี่ยวข้องกับความเสี่ยงทางการเมือง จนบางครั้งเกิดความรู้สึกว่านักธุรกิจเป็นกลุ่มบุคคลที่ไม่มีส่วนรับรู้ต่อการขึ้นลงของสังคม แต่กระนั้นก็ตามในบรรดานักธุรกิจต่างก็มีความต้องการที่ให้ “การเมืองนิ่ง” และ “มีเสถียรภาพ” เพื่อให้การประกอบธุรกิจโดยเฉพาะอย่างยิ่งการลงทุนจากต่างประเทศและการท่องเที่ยวสามารถเจริญรุ่งเรืองต่อไปได้”

ข้อเสนอแนะ

งานวิจัยนี้สามารถนำภูมิความสัมพันธ์ และกฎการตัดสินใจของนักธุรกิจที่มีการตัดสินใจเลือกพรรคการเมืองที่มีนโยบายสามารถนำไปสู่การปฏิบัติได้จริง และมีแหล่งที่มาของงบประมาณชัดเจน และพรรคการเมืองที่มีความจงรักภักดี และมีความซื่อสัตย์ รวมถึงความรู้ความสามารถของผู้สมัคร ส.ส. ไปประยุกต์ใช้ในการกำหนดแนวนโยบายของพรรคการเมือง และการคัดเลือกกว่าที่ผู้สมัคร ส.ส. ในนามพรรค เพื่อให้สามารถเป็นที่ยอมรับของประชาชนทุกภาคส่วนรวมทั้งนักธุรกิจ ทั้งนี้เพราะมีส่วนสำคัญอย่างยิ่งในการสร้างสังคมไทยไปสู่การลงทุน การสร้างงาน การสร้างสังคมสมัยใหม่ที่ทันกับโลก สำหรับการทํารววจัยในครั้งต่อไปนั้น คณะผู้วิจัยขอเสนอว่า ควรพิจารณาเทคนิคการสร้างตัวแบบที่ใช้ในการจำแนกประเภทข้อมูลที่หลากหลาย และเพิ่มชุดข้อมูลที่ทำการทดสอบให้มากขึ้น เช่น การเปรียบเทียบมุมมองของนักธุรกิจไทยกับนักธุรกิจต่างชาติที่มาลงทุนในประเทศไทยเกี่ยวกับการเลือกตั้งในอนาคต เพื่อให้ข้อมูลที่กว้างขึ้น นอกจากนี้ในส่วนของการทำความสะอาดข้อมูล (Data Cleaning) นั้น จัดเป็นขั้นตอนที่สำคัญและใช้เวลาค่อนข้างมาก และสิ่งที่ต้องระวังให้มากคือความสอดคล้องแบบนัยของข้อมูลในแต่ละคุณลักษณะซึ่งอาจจะส่งผลถึงการทำนายที่มีความถูกต้องและแม่นยำ

เอกสารอ้างอิง

- De Clercq, D., Haq, I. U., Azeem, M. U., & Ahmad, H. N. (2019). The relationship between workplace incivility and helping behavior: roles of job dissatisfaction and political skill. *The Journal of psychology*, 153(5), 507-527.
- Good, M.C. & Schwepker, C. H. (2022). Business-to-business salespeople and political skill: Relationship building, deviance, and performance. *Journal of Business Research*, 139(2022), 32-43.
- Han, J., Pei, J. & Tong, (2022). *Data mining concepts and techniques* (4th ed.). Morgan Kaufman Publishers. Retrieved from https://gsmis.snru.ac.th/e-thesis/file_att1/2023031063426423118_fulltext.pdf (in Thai)

- Jaroenpuntaruk, W. (2015). *Date Warehouse, Data Mining and Business Intelligence*. Nonthaburi: Sukhothai Thammathirat Open University. (in Thai).
- limmane, A. (2015). *State society and change: A consideration on power, policy, and relationship network* (2nd ed). Bangkok: Siamparitas. (in Thai)
- Plangsree, S. (2022). *Factors affecting people's decision to vote for the council member and administrators of Nakhon Phanom local government organizations* (master's thesis). Sakon Nakhon Rajabhat University, Sakon Nakhon, Thailand.
- Riera, P., & Cantú, F. (2022). Electoral systems and ideological voting. *European Political Science Review*, 14(4), 463-481.
- Sinsomboonthong, S. (2015). *Data Mining* (1st ed). Bangkok: Chamchuri Product. (in Thai)
- Teeravakin, L. (2009). *Role of business persons of political*. Retrieved from <https://mgronline.com/daily/detail/9520000053345>
- Witten, I. H., Frank, E. & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (3rd ed.). Burlington, Massachusetts: Morgan Kaufmann Publishers.

