

การเรียนรู้ของเครื่องเพื่อทำนายพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร Machine learning for Customer Behavior Prediction on Banking Product

ธีรินทร์ ฉายศิริโชติ¹ สุวรรณี อัสวกุลชัย^{2*}

¹ ภาควิชาวิศวกรรมอาหาร คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ถนนฉลองกรุง แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพมหานคร 10520

² หลักสูตรวิศวกรรมคอมพิวเตอร์และเทคโนโลยีการเงิน คณะวิศวกรรมศาสตร์ มหาวิทยาลัยหอการค้าไทย
126/1 ถนนวิภาวดีรังสิต แขวงรัชดาภิเษก เขตดินแดง กรุงเทพมหานคร 10400

Teerin Chaysirichote¹ Suwannee Adsavakulchai²

¹ Department of Food Engineering, Faculty of Engineering, King Mongkut's Institute of Technology
Ladkrabang Chalongkrung Road, Ladkrabang Sub-district, Ladkrabang District, Bangkok 10520

² Computer Engineering and Financial Technology Program, School of Engineering,
University of the Thai Chamber of Commerce

126/1 Vibhavadi Rangsit Road, Ratchadaphisek Sub-district, Din Daeng District, Bangkok 10400

*Corresponding author email: suwannee_ads@utcc.ac.th

(Received: May 14, 2022; Revised: May 16, 2023 ; Accepted: May 13, 2024)

บทคัดย่อ

บทความนี้มีวัตถุประสงค์ เพื่อศึกษาปัจจัยในการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร จากการนำเสนอผลิตภัณฑ์ผ่านโทรศัพท์ และเพื่อพัฒนาแบบจำลองพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคารด้วยการเรียนรู้ของเครื่อง กรณีศึกษาธนาคารแห่งหนึ่งในประเทศโปรตุเกส ในการศึกษาครั้งนี้ ใช้กระบวนการวิเคราะห์ข้อมูลมาตรฐาน เรียกว่า Cross-Industry Standard Process (CRISP) ประกอบด้วย 6 ขั้นตอน ได้แก่ เข้าใจปัญหาของธุรกิจ เข้าใจข้อมูล เตรียมข้อมูล พัฒนาแบบจำลอง การประเมิน การนำไปใช้จริง ศึกษาจากข้อมูลลูกค้า 41,188 คนและปัจจัย 20 รายการ เตรียมและคัดเลือกข้อมูลให้สมบูรณ์ด้วยกระบวนการ extract-transform-load พัฒนาแบบจำลองด้วยการเรียนรู้ของเครื่อง ใช้โปรแกรม Knime เป็นเครื่องมือ เลือก 3 อัลกอริทึม ได้แก่ Decision tree, Random forest และ Tree ensemble ทดสอบประสิทธิภาพอัลกอริทึม ด้วย accuracy, sensitivity และ specificity เพื่อหาอัลกอริทึมที่ดีที่สุด ผลการศึกษาพบว่า ปัจจัยที่มีผลต่อการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร 3 อันดับแรก ได้แก่ ระยะเวลาที่ใช้ในการคุยโทรศัพท์กับลูกค้า จำนวนวันที่ผ่านไปหลังจากที่ลูกค้าได้รับการติดต่อครั้งสุดท้ายจากแคมเปญก่อนหน้า และผลของแคมเปญการตลาดครั้งก่อน สำหรับแบบจำลองพฤติกรรมลูกค้าต่อตัดสินใจเลือกเงินฝากประจำด้วยการเรียนรู้ของเครื่อง พบว่า อัลกอริทึม Tree ensemble ให้ความถูกต้องในการพยากรณ์สูงสุด คิดเป็นร้อยละ 90.94 เมื่อเทียบกับอัลกอริทึม Decision tree และ Random forest ให้ความถูกต้องร้อยละ 87.54 และ 90.29 ตามลำดับ และ แสดงผลข้อมูลด้วยภาพ โดยใช้ Power BI โดยสรุป ผลจากการศึกษานี้เป็นประโยชน์ในการวางแผนการตลาด เพื่อเพิ่มจำนวนลูกค้าได้เป็นอย่างดี การศึกษาครั้งต่อไป การประยุกต์กับข้อมูลลูกค้าธนาคารในประเทศไทย

คำสำคัญ : การเรียนรู้ของเครื่อง ผลิตภัณฑ์ธนาคาร ต้นไม้การตัดสินใจ ป่าสุ่มแบบกลุ่ม กลุ่มต้นไม้การตัดสินใจ

ABSTRACT

This article aims to study the parameters for customer to make decision in bank product via telemarketing and to apply machine learning for customer behavior model of banking product. The technique in this study is the Cross-Industry Standard Process. There are 6 processes that are business understanding, data understanding, data preparation, model development, evaluation, and deployment. The data consisted of 41,188 customers and 20 parameters. To do the data preparation, feature selection i.e., data transfer, data reduction, data cleaning using Knime a tool. There are 3 algorithms for customer behavior model of banking product that are Decision tree, Random Forest and Tree ensemble. Algorithm performance testing by using accuracy, sensitivity, and specificity for the best algorithm. The results can demonstrate that the top three parameters impacting decision making are the length of time spent talking on the phone with customers, the number of days elapsed since the customer was last contacted from a previous campaign, and the results of previous marketing campaigns. The most accuracy algorithm for customer behavior prediction in bank product is Tree ensemble with accuracy 90.94% comparing with Decision tree and Random Forest tree with accuracy 87.54% and 90.29%, respectively. To do the data visualization using Power BI that is very beneficial for planning and leads to higher sales. On-going research, to apply this methodology with bank product marketing in Thailand.

Keywords: Machine learning, bank-marketing, decision tree, random Forest, tree ensemble.

1. บทนำ

ผลิตภัณฑ์ของธนาคาร ได้แก่ เงินฝาก การลงทุน ประกัน สินเชื่อและบัตรเครดิต สำหรับการตลาดทางโทรศัพท์เหมาะสำหรับสินค้าที่เป็นที่รู้จัก เช่น บัตรเครดิต ประกันชีวิต สินเชื่อ เป็นต้น เพื่อนำเสนอผลิตภัณฑ์ของธนาคารได้อย่างเหมาะสมตรงกับความต้องการของลูกค้า การคัดกรองลูกค้าเป็นกระบวนการที่ธนาคารได้ดำเนินการในการได้มาผู้ซึ่งจะมาเป็นลูกค้า มักจะใช้วิธีการโทรศัพท์สุ่มไปยังกลุ่มลูกค้าเป้าหมาย เช่น บัตรเครดิตธนาคาร ก็ต้องหาคนที่มียาได้เกิน 15,000 บาท ขึ้นไป ซึ่งในการเลือกกลุ่มเป้าหมายที่จะโทรศัพท์ เพื่อไปนำเสนอผลิตภัณฑ์ สามารถสร้างรายได้เพิ่มตามที่ให้แก่ธนาคาร [1]

ปัจจัยที่มีผลต่อพฤติกรรมตัดสินใจเลือกใช้บริการธนาคารพาณิชย์ ได้แก่ เพศ อายุ อาชีพ ชื่อเสียง ความน่าเชื่อถือ ความซื่อสัตย์ สภาพแวดล้อม เข้า

บริการสาขาใกล้บ้าน หรือที่ทำงานเป็นส่วนใหญ่ ช่องทางการจัดจำหน่าย กระบวนการให้บริการ และพนักงานของธนาคาร ความสุภาพอ่อนโยน ทั้งนี้ธนาคารต่างก็มุ่งหวังที่จะได้รับบริการที่ดีกับลูกค้า อย่างไรก็ตามการบริการต้องสอดคล้องกับความต้องการของลูกค้า ทำให้ลูกค้าเกิดความพึงพอใจสูงสุด [2,3]

ในยุคของข้อมูลข่าวสารผ่านสื่อออนไลน์ การวิเคราะห์พฤติกรรมและความต้องการของลูกค้า สามารถช่วยเจาะลงไปให้ลึกถึงระดับสนิยม เพื่อสร้างสามารถเข้าถึงความต้องการของลูกค้า การใช้ข้อมูลจากลูกค้า ทำให้เกิดประโยชน์ในการบริการ สร้างความประทับใจ ไม่เพียงให้ลูกค้าใช้บริการซ้ำ แต่ต้องบอกต่อ รวมถึงสื่อสารให้ตรง กับกลุ่มเป้าหมาย ปัจจุบันเป็นยุคของลูกค้าเฉพาะคน (Customize) ตอบโจทย์เฉพาะบุคคลมากขึ้น [4,5]

Cross-Industry Standard Process for Data Mining (CRISP-DM) เป็นกระบวนการวิเคราะห์ข้อมูล

มาตรฐานสำหรับเหมืองข้อมูล ประกอบด้วย 6 ขั้นตอน ได้แก่ เข้าใจปัญหาของธุรกิจ เข้าใจข้อมูล เตรียมข้อมูล พัฒนาแบบจำลอง การประเมิน การนำไปใช้จริง [6,7]

อัลกอริทึมต้นไม้การตัดสินใจ (Decision Tree) เป็นเทคนิคการเรียนรู้เชิงตัวแบบ (Supervised Learning) ใช้ในการวิเคราะห์ข้อมูลและสร้างโมเดลทำนายผลลัพธ์ ใช้จัดการข้อมูลเป็นกลุ่ม เช่น การจัดกลุ่มลูกค้าโดยคุณสมบัติต่างๆ หรือการแยกประเภทสิ่งของต่างๆ โดยมีข้อมูลคุณสมบัติของสิ่งของนั้นๆ เป็นตัวแปร input ในโมเดลต้นไม้การตัดสินใจ [8]

อัลกอริทึม random forest เป็นเทคนิคการเรียนรู้ใช้หลายๆ ต้นไม้การตัดสินใจในการทำนายผลลัพธ์ ใช้เทคนิคที่เรียกว่า "การสุ่มและการแบ่งข้อมูล" (random sampling and feature splitting) เพิ่มความแม่นยำของโมเดล โดยแต่ละต้นไม้ใน Random Forest จะมีความหลากหลายในการสร้างโมเดล และสร้างเสถียรภาพให้กับการทำนายผลลัพธ์ เมื่อผลลัพธ์จากแต่ละต้นไม้ถูกรวบรวมกัน จะทำให้ Random Forest มีความแม่นยำมากกว่า Decision Tree เดียว [9]

อัลกอริทึม Tree Ensemble เป็นเทคนิคการเรียนรู้เชิงตัวแบบ (Ensemble Learning) ที่ใช้หลายๆ โมเดลต้นไม้การตัดสินใจ หรือโมเดลประสาทเทียม เพื่อประมวลผลข้อมูล โดยการใช้หลายๆ โมเดลเพื่อทำนายผลลัพธ์จะช่วยลดความผิดพลาดในการทำนาย และเพิ่มความแม่นยำให้กับโมเดลที่สร้างขึ้น การรวมหลายๆ โมเดลเข้าด้วยกันยังช่วยลดโอกาสการเกิด overfitting ที่อาจเกิดขึ้นในโมเดลเดี่ยวๆ [10]

สำหรับการศึกษาในครั้งนี้ มีวัตถุประสงค์ เพื่อศึกษาปัจจัยในการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร และเพื่อพัฒนาแบบจำลองพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร ด้วยการเรียนรู้ของเครื่อง

2. วิธีการศึกษา

กระบวนการวิเคราะห์ข้อมูลมาตรฐาน CRISP [8] เพื่อวิเคราะห์ปัจจัยในการตัดสินใจเลือกผลิตภัณฑ์

ธนาคาร และพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคารด้วยการเรียนรู้ของเครื่อง มี 6 ขั้นตอน ได้แก่

2.1 การทำความเข้าใจปัญหาของธุรกิจ

การทำความเข้าใจถึงปัจจัยในการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร และพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร เพื่อนำมาวิเคราะห์ข้อมูลด้วยการเรียนรู้ของเครื่อง และสามารถระบุโอกาสในการจัดรายการส่งเสริมการขายให้อยู่ในรูปแบบที่เหมาะสม

2.2 การทำความเข้าใจข้อมูล

ศึกษารายละเอียดข้อมูล จาก <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing> ประกอบด้วยข้อมูลลูกค้า 41,188 คนและปัจจัย 20 รายการ ทำการตรวจสอบความถูกต้องของข้อมูล และรายละเอียดเพียงพอต่อการนำไปใช้ในการวิเคราะห์ด้วยการเรียนรู้ของเครื่อง

2.3 การเตรียมข้อมูล

หลังจากรวบรวมข้อมูล ในขั้นตอนนี้จะต้องมีการจัดเตรียมข้อมูลและคัดเลือกข้อมูลให้เป็นไปตามมาตรฐานเดียวกัน และคัดเลือกข้อมูลที่มีความถูกต้องสมบูรณ์มากที่สุด เพื่อให้ได้ผลการทดสอบที่มีประสิทธิภาพสูงสุด ขั้นตอนที่ใช้เวลานานที่สุด ถ้าข้อมูลที่ใช้นั้นไม่ถูกต้อง มีผิดพลาด ทำให้ผลลัพธ์คลาดเคลื่อนได้ การเตรียมข้อมูลแบ่งเป็น 3 ขั้นตอนย่อยดังนี้

1) การคัดเลือกข้อมูล (Data Selection) กำหนดเป้าหมายที่ทำการวิเคราะห์พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร เลือกใช้เฉพาะข้อมูลที่เกี่ยวข้องกับสิ่งที่จะวิเคราะห์ การกำหนดคุณลักษณะ และค่าที่เป็นไปได้ของคุณลักษณะ ศึกษาโครงสร้างของฐานข้อมูล และการจัดเตรียมข้อมูล คุณลักษณะที่คาดว่าจะนำมาใช้งาน เพื่อใช้ในการหาความรู้ และได้มีการกำหนดค่าที่เป็นไปได้ของคุณลักษณะ

2) การกลั่นกรองข้อมูล (Data Cleaning) ในการศึกษาครั้งนี้ พบข้อมูลที่ไม่ถูกต้อง เนื่องจากจากปัญหาในระหว่างการจัดเก็บข้อมูล เช่น กรอกข้อมูลไม่ครบ กรอกข้อมูลซ้ำซ้อน การกลั่นกรองข้อมูล การทำความสะอาดข้อมูลก่อน ลดขนาดข้อมูล

3) การแปลงรูปข้อมูล (Data Transformation) ขั้นตอนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้ในการวิเคราะห์ตามอัลกอริทึมของเหมืองข้อมูลที่เลือกใช้

2.4 การพัฒนาแบบจำลอง

แบบจำลองใช้ในการทำนาย ในการศึกษาครั้งนี้เลือกใช้ Decision tree, Random forest และ Tree ensemble ทั้งนี้เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด เมื่อทำขั้นตอนนี้แล้ว อาจมีการย้อนกลับไปทำขั้นตอน data preparation เพื่อแปลงข้อมูลบางส่วนให้เหมาะสมกับแต่ละเทคนิคด้วย นอกจากนี้ยังมีการประเมินแบบจำลองวิเคราะห์ข้อมูลที่ได้ ในรูปแบบความถูกต้องของแบบจำลอง เพื่อเป็นตัวบ่งชี้ความน่าเชื่อถือของแบบจำลองที่ได้

2.5 การประเมินประสิทธิภาพอัลกอริทึม

ขั้นตอนนี้เป็นการประเมินประสิทธิภาพของผลลัพธ์จากแบบจำลอง โดยนำผลลัพธ์ที่ได้มาวิเคราะห์ความถูกต้องจากข้อมูลที่นำมาทดสอบ ต้องอาศัยทักษะในการวิเคราะห์ข้อมูลเพื่อช่วยให้การวิเคราะห์ทำได้สะดวกและรวดเร็วขึ้น

2.6 การถ่ายทอดเพื่อนำไปปฏิบัติ

เป็นแนวทางไปใช้กับการวิเคราะห์พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร และจัดรายการส่งเสริมการขาย จึงมีการใช้เครื่องมือทำงานด้านกราฟิก เช่น การแสดงผลการวิเคราะห์ด้วยกราฟ รายงานรูปแบบต่างๆ เป็นต้น

3. ผลการศึกษา

3.1 ผลการทำความเข้าใจปัญหาของธุรกิจ

ข้อมูลนี้เกี่ยวข้องกับแคมเปญการตลาดผ่านโทรศัพท์ของธนาคารแห่งหนึ่งในประเทศโปรตุเกส แคมเปญการตลาดขึ้นอยู่กับประสิทธิภาพบ่อยครั้ง จำเป็นต้องมีการติดต่อกับลูกค้ารายเดียวกันมากกว่าหนึ่งครั้ง เพื่อเข้าถึงผลิตภัณฑ์เงินฝากประจำของธนาคาร

3.2 ผลการทำความเข้าใจข้อมูล

ใช้ข้อมูล open data จาก <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing> มีการจัดหมวดหมู่ลูกค้าสมัครเงินฝากประจำ มี 4 ชุดข้อมูล ได้แก่

1) ไฟล์ bank-additional-full.csv ประกอบด้วยตัวอย่างทั้งหมด 41,188 รายการ และมีปัจจัยจำนวน 20 รายการ เป็นข้อมูลตั้งแต่เดือนพฤษภาคม 2551 ถึงพฤศจิกายน 2553

2) ไฟล์ bank-additional.csv ประกอบด้วยข้อมูลที่สุ่มเลือกจากไฟล์ในข้อ 1) จำนวนร้อยละ 10 ของตัวอย่างทั้งหมดในข้อ 1) มีจำนวน 4,119 รายการ และมีปัจจัยจำนวน 20 รายการ

3) ไฟล์ bank-full.csv ประกอบด้วยตัวอย่างทั้งหมด และมีปัจจัยจำนวน 17 รายการ จัดเรียงตามวันที่ ชุดข้อมูลนี้เป็นข้อมูลชุดเก่าและมีปัจจัยน้อยกว่า(

4) ไฟล์ bank.csv ประกอบด้วยข้อมูลที่สุ่มเลือกจากไฟล์ในข้อ 3) จำนวนร้อยละ 10 ของตัวอย่างทั้งหมดในข้อ 3) มี และมีปัจจัยจำนวน 17 รายการ ชุดข้อมูลนี้เป็นข้อมูลชุดเก่าและมีปัจจัยน้อยกว่า(

ทั้งนี้ในการศึกษาครั้งนี้ เลือกใช้ข้อมูลชุดที่ 1 เท่านั้น

3.3 ผลการเตรียมข้อมูล

จากการสกัดข้อมูลจากลูกค้ากับการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร เพื่อเข้าสู่กระบวนการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมจะวิเคราะห์ต่อไป ศึกษาปัจจัยทั้งหมด 20 รายการ แสดงในตารางที่ 1

การเตรียมข้อมูลด้วยการลบข้อมูลซ้ำซ้อน แก้ไขข้อมูลผิดพลาด เช่น ข้อมูลที่มีผิดรูปแบบ ข้อมูลที่หายไป ข้อมูล Outlier ที่แปลกแยกจากคนอื่นผิดปกติค่อนข้างมาก เป็นต้น แสดงในตารางที่ 2

3.4 ผลการพัฒนาแบบจำลอง

นำข้อมูลจาก 3.3 มาทำการพัฒนาแบบจำลองวิเคราะห์หาความสัมพันธ์ และอัลกอริทึมพยากรณ์ พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร ได้แก่ Decision tree, Random forest และ Tree ensemble โดยใช้ Knime เป็นเครื่องมือ แสดงในรูปที่ 1-3 และเปรียบเทียบความถูกต้องของอัลกอริทึมทั้ง 3 แสดงในตารางที่ 3

3.5 ผลการประเมินประสิทธิภาพอัลกอริทึม

จากตารางที่ 3 อัลกอริทึม Tree ensemble ให้ค่าความถูกต้องในการพยากรณ์สูงสุด จึงนำมาทดสอบประสิทธิภาพอัลกอริทึม ด้วยการแบ่ง partition ข้อมูล

ปีที่ 19 ฉบับที่ 2 เดือน เมษายน – กันยายน พ.ศ. 2567

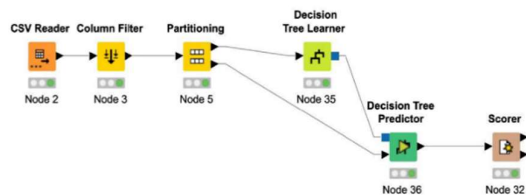
90% สำหรับ model creation และ 10% สำหรับ partitioning 70:30 แสดงในรูปที่ 4 และสรุปผลในตาราง model check เพื่อทำการทดสอบโมเดล เพิ่มจากทำ ที่ 4

ตารางที่ 1 ปัจจัยของลูกค้านับกับการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร

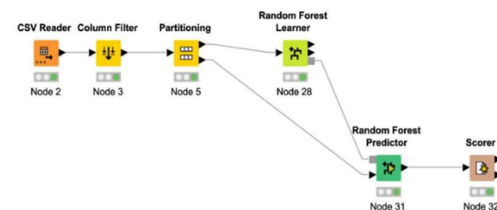
1. age	2. duration
3. job	4. campaign
5. marital	6. pdays
7. education	8. previous
9. default	10. poutcome
11. balance	12. employment variation rate
13. housing	14. consumer price index
15. loan	16. consumer confidence index
17. contact	18. euribor3m
19. day_of_week	20. number of employees

ตารางที่ 2 ผลการกรองข้อมูล (Data Cleaning)

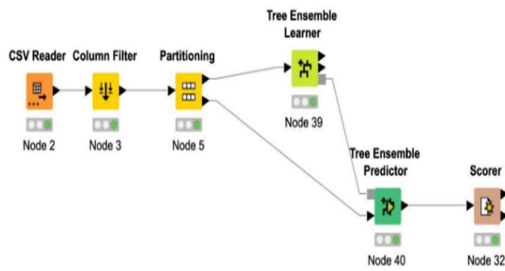
age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome	emp.var.rate	cons.price.idx	cons.conf.idx	euribor3m	nr.employed	y
41	technician	married	professional.course	no	yes	no	cellular	nov	mon	185	2	999	0	nonexistent	-1.1	94.767	-50.8	1.039	4963.6	no
31	housemaid	single	university.degree	no	no	no	telephone	nov	mon	152	5	999	0	nonexistent	-1.1	94.767	-50.8	1.039	4963.6	no
41	technician	married	professional.course	no	yes	no	cellular	nov	mon	545	2	999	0	nonexistent	-1.1	94.767	-50.8	1.039	4963.6	yes
31	housemaid	single	university.degree	no	no	no	cellular	nov	mon	159	4	999	0	nonexistent	-1.1	94.767	-50.8	1.039	4963.6	no
35	technician	divorced	basic.4y	no	no	no	cellular	nov	tue	363	1	999	0	nonexistent	-1.1	94.767	-50.8	1.035	4963.6	yes
35	technician	divorced	basic.4y	no	yes	no	cellular	nov	tue	514	1	9	4	success	-1.1	94.767	-50.8	1.035	4963.6	yes
33	admin.	married	university.degree	no	no	no	cellular	nov	tue	843	1	999	0	nonexistent	-1.1	94.767	-50.8	1.035	4963.6	yes
33	admin.	married	university.degree	no	yes	no	cellular	nov	tue	510	1	999	1	failure	-1.1	94.767	-50.8	1.035	4963.6	no
60	blue-collar	married	basic.4y	no	yes	no	cellular	nov	tue	347	2	4	1	success	-1.1	94.767	-50.8	1.035	4963.6	no
35	technician	divorced	basic.4y	no	yes	no	cellular	nov	tue	385	3	4	2	success	-1.1	94.767	-50.8	1.035	4963.6	yes
54	admin.	married	professional.course	no	no	no	cellular	nov	tue	1868	2	10	1	success	-1.1	94.767	-50.8	1.035	4963.6	yes
38	housemaid	divorced	university.degree	no	no	no	cellular	nov	wed	403	2	999	0	nonexistent	-1.1	94.767	-50.8	1.03	4963.6	yes
32	admin.	married	university.degree	no	no	no	telephone	nov	wed	651	1	999	1	failure	-1.1	94.767	-50.8	1.03	4963.6	yes
32	admin.	married	university.degree	no	yes	no	cellular	nov	wed	236	3	999	0	nonexistent	-1.1	94.767	-50.8	1.03	4963.6	no
38	entrepreneur	married	university.degree	no	no	no	cellular	nov	wed	144	2	999	0	nonexistent	-1.1	94.767	-50.8	1.03	4963.6	no
62	services	married	high.school	no	yes	no	cellular	nov	wed	154	5	999	0	nonexistent	-1.1	94.767	-50.8	1.03	4963.6	no
40	management	divorced	university.degree	no	yes	no	cellular	nov	wed	293	2	999	4	failure	-1.1	94.767	-50.8	1.03	4963.6	no
33	student	married	professional.course	no	yes	no	telephone	nov	thu	112	1	999	0	nonexistent	-1.1	94.767	-50.8	1.031	4963.6	yes
31	admin.	single	university.degree	no	yes	no	cellular	nov	thu	353	1	999	0	nonexistent	-1.1	94.767	-50.8	1.031	4963.6	yes
62	retired	married	university.degree	no	yes	no	cellular	nov	thu	329	1	999	2	failure	-1.1	94.767	-50.8	1.031	4963.6	yes
62	retired	married	university.degree	no	yes	no	cellular	nov	thu	208	1	1	6	success	-1.1	94.767	-50.8	1.031	4963.6	yes
34	student	single	unknown	no	yes	no	cellular	nov	thu	180	1	999	2	failure	-1.1	94.767	-50.8	1.031	4963.6	no
38	housemaid	divorced	high.school	no	yes	yes	cellular	nov	thu	360	1	999	0	nonexistent	-1.1	94.767	-50.8	1.031	4963.6	no
57	retired	married	professional.course	no	yes	no	cellular	nov	thu	124	6	999	0	nonexistent	-1.1	94.767	-50.8	1.031	4963.6	no
62	retired	married	university.degree	no	no	no	cellular	nov	thu	483	2	6	3	success	-1.1	94.767	-50.8	1.031	4963.6	yes



รูปที่ 1 ผลการใช้อัลกอริทึม Decision tree กับ
พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร



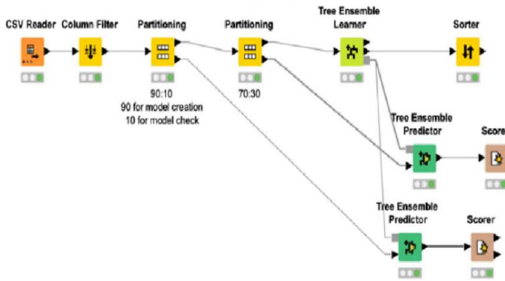
รูปที่ 2 ผลการใช้อัลกอริทึม Random forest กับ
พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร



รูปที่ 3 ผลการใช้อัลกอริทึม Tree ensemble กับ
พฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร

ตารางที่ 3 เปรียบเทียบความถูกต้องของอัลกอริทึมทั้ง 3

	Decision tree	Random forest	Tree ensemble
Accuracy (%)	87.54	90.29	90.94
Sensitivity (%)	92.71	91.72	92.34
Specificity (%)	46.14	65.23	68.44



รูปที่ 4 ผลการประเมินประสิทธิภาพอัลกอริทึม Tree
ensemble กับพฤติกรรมลูกค้าต่อผลิตภัณฑ์ธนาคาร

ตารางที่ 4 ผลการประเมินประสิทธิภาพอัลกอริทึม

	Tree ensemble	Tree ensemble + 10% data for testing
Accuracy (%)	90.94	91.23
Sensitivity (%)	92.34	92.71
Specificity (%)	68.44	71.42

จากรูปที่ 4 และตารางที่ 4 แสดงให้เห็นผลการ
ประเมินทดสอบประสิทธิภาพอัลกอริทึม Tree
ensemble ที่ให้ค่าความถูกต้องที่ร้อยละ 91.23

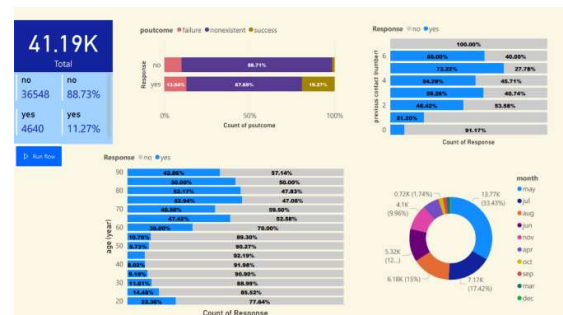
สำหรับปัจจัยที่มีผลต่อการตัดสินใจเลือกผลิตภัณฑ์
ธนาคาร เมื่อทำการจัดเรียงปัจจัย (sorter) พบว่า ปัจจัย
ที่มีผลต่อการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร 3 อันดับ
แรก ได้แก่ ระยะเวลาที่ใช้ในการคุยโทรศัพท์กับลูกค้า
(duration) จำนวนวันที่ผ่านไปหลังจากที่ลูกค้าได้รับการ
ติดต่อครั้งสุดท้ายจากแคมเปญก่อนหน้านี้ (pdays) และผล
ของแคมเปญการตลาดครั้งก่อน (poutcome) แสดงใน
ตารางที่ 5

ตารางที่ 5 ปัจจัยที่มีผลต่อการตัดสินใจเลือกผลิตภัณฑ์
ธนาคาร

Row ID	#splits ...	#splits ...	#splits ...	#candi...	#candi...	#candi...
duration	42	74	131	42	79	148
pdays	30	42	91	37	67	184
poutcome	28	50	65	43	88	176
previous	26	29	49	41	75	161
month	17	56	110	32	87	176
age	15	38	62	33	88	137
contact	13	25	46	40	76	164
job	11	31	54	38	81	153
default	6	8	34	40	73	170
education	4	16	34	51	86	157

3.6 ผลการถ่ายทอดเพื่อนำไปปฏิบัติ

จากอัลกอริทึมในการพยากรณ์พฤติกรรมลูกค้าต่อ
ผลิตภัณฑ์ธนาคาร ไปสู่การถ่ายทอดเพื่อนำไปปฏิบัติ
ด้วยการนำมาแสดงผลข้อมูลด้วยภาพ (data
visualization) โดยใช้ (Power BI แสดงในภาพที่ 5



ภาพที่ 5 ผลการทำ data visualization ด้วย Power BI

จากภาพที่ 5 การนำเสนอภาพพฤติกรรมลูกค้าต่อ
ผลิตภัณฑ์ธนาคาร ช่วยสนับสนุนการตัดสินใจได้ดีกว่า

ข้อมูลที่อยู่ในรูปของตาราง เนื่องจากภาพสามารถแสดงการรายละเอียด การเปรียบเทียบที่ชัดเจน โดยเฉพาะกราฟสีต่างๆ แยกให้เห็นชัดเจน

4. สรุปและวิจารณ์ผลการทดสอบ

ปัจจัยที่มีผลต่อการตัดสินใจเลือกผลิตภัณฑ์ธนาคาร 3 อันดับแรก ได้แก่ ระยะเวลาที่ใช้ในการคุยโทรศัพท์กับลูกค้า จำนวนวันที่ผ่านไปหลังจากที่ลูกค้าได้รับการติดต่อครั้งสุดท้ายจากแคมเปญก่อนหน้านี้ และผลของแคมเปญการตลาดครั้งก่อน ซึ่งเป็นประโยชน์ในการวางแผนการตลาด เพื่อเพิ่มจำนวนลูกค้าได้เป็นอย่างดี

แบบจำลองด้วยการเรียนรู้ของเครื่อง ในการที่ลูกค้าตัดสินใจเลือกเงินฝากประจำ จากการนำเสนอผลิตภัณฑ์ผ่านโทรศัพท์ กรณีศึกษา ธนาคารแห่งหนึ่งในประเทศโปรตุเกส พบว่า อัลกอริทึม Tree ensemble พยากรณ์ได้ถูกต้องสูงสุด คิดเป็นร้อยละ 90.94 เมื่อเทียบกับอัลกอริทึม Decision tree และ Random forest ที่มีถูกต้องร้อยละ 87.54 และ 90.29 ตามลำดับ สามารถนำมาแสดงผลข้อมูลด้วยภาพ โดยใช้ Power BI ช่วยสนับสนุนการตัดสินใจและนำไปใช้ในทางปฏิบัติได้เป็นอย่างดี

ข้อจำกัดในการศึกษาครั้งนี้ ได้แก่ การเข้าถึงข้อมูลลูกค้าของธนาคารในประเทศไทย เนื่องจากข้อมูลลูกค้าเป็นข้อมูลเฉพาะบุคคลและเป็นความลับ การศึกษานี้จึงไม่สามารถใช้ข้อมูลของประเทศไทย จำเป็นต้องใช้ข้อมูล open data เพื่อเป็นต้นแบบในการประยุกต์ต่อไป

ส่วนใหญ่การศึกษาพฤติกรรมของลูกค้า ยังคงใช้แบบสอบถามความพึงพอใจมากกว่าที่จะมีการวิเคราะห์เชิงลึก เพื่อให้ได้พฤติกรรมเชิงลึกจากข้อมูลการใช้บริการของลูกค้า ดังนั้น การเรียนรู้ของเครื่องด้วยการใช้ข้อมูลในอดีต เพื่อพยากรณ์พฤติกรรมเชิงลึกของลูกค้า (customer insight) สามารถนำไปใช้ประกอบการตัดสินใจในการนำเสนอผลิตภัณฑ์ธนาคารเพื่อตรงตามความต้องการของลูกค้าได้ดีขึ้น

สำหรับการศึกษาต่อไป ธนาคารต่างๆ ในประเทศไทยสามารถประยุกต์หลักการนี้ไปใช้กับข้อมูลลูกค้า

ธนาคารของตนเองได้ เพื่อให้สามารถเข้าถึงพฤติกรรมของลูกค้าได้มากยิ่งขึ้น เนื่องจากความต้องการของลูกค้ามีความเฉพาะบุคคลมากขึ้น จำเป็นที่ธนาคารต้องมีการปรับตัว เพื่อสามารถรองรับพฤติกรรมของลูกค้าที่เปลี่ยนแปลง และด้วยเทคโนโลยีที่สามารถเข้าถึงได้อย่างรวดเร็ว การเปรียบเทียบผลิตภัณฑ์ของแต่ละธนาคารทำได้สะดวก รวดเร็ว เพื่อตอบโจทย์ความต้องการมากที่สุด ถ้าธนาคารมีระบบที่สามารถให้ลูกค้าเลือกได้ตามความต้องการ ก็จะเพิ่มขีดความสามารถในการแข่งขันมากกว่าธนาคารอื่น รวมถึงมีช่องทางในการติดต่อทุกช่องทาง (omnichannel) ที่เข้าถึงลูกค้าได้ทุกที่ ทุกเวลาที่ เหมาะสมเฉพาะบุคคล เช่น เทคโนโลยีการเรียนรู้เชิงลึก (Deep learning) เป็นต้น

5. เอกสารอ้างอิง

- [1] Kasikorn Research Center. (2016, May 30). Credit Card Flat in 2015 but in 2016 Retain the Growth [Online]. Available :<https://www.kasikornresearch.com/th/K-Econ Analysis>
- [2] P. Anaboontarik, “Factors affecting the choice of using Commercial Banks in Mueang District, Chiang Mai Province,” Independent Study (IS), Master of Economics, Chiang Mai University, 2004.
- [3] K. Jitphuthai, “Factors Influencing Decision-Making Behavior in Choosing Commercial Bank Services in Bang Khae, Phasi Charoen, Nong Khaem, and Krathum Baen Districts” , Independent Study (IS) , Master of Business Administration, Siam University, 2009.
- [4] Matichon News. (2017, September 30). Krung Sri First Choice launched new card “Krung Sri Frist Choice Care” giving free accidental insurance with highest Protection 100,000 baht [Online]. Available :<https://www.matichon.co.th/news/663544>.

- [5] W Piyajinda, “Factors relating to the Decision-Making Behaviors of SMILE CASH Credit among Consumers of Bank of Ayudhaya Public Company Limited (Zone 5),” M.S. Thesis, Faculty of Business Administration. Khon Kean University, 2009.
- [6] P. Suthipongwirat, “Behaviors of Using Cash Card among Thai Consumers in Bangkok Metropolitan”. Independent Study of Master Business Administration, Faculty of Business Administration, Khon Kaen University, 2011.
- [7] C. Shearer, “The CRISP-DM model: The new blueprint for data mining,” *Journal of Data Warehousing*, vol. 5, no. 4, pp. 13–22, 2000.
- [8] N. Wongprachanukul, “A proper method for decision tree pruning in scientific data mining,” M.S. Thesis, Nakornratchasima, Suranaree University of Technology, 2005.
- [9] K.B. Jyothi, K.H. Bindu, and D. Suryanarayana, “A comparative study of random forest & K – nearest neighbors on HAR dataset using caret,” *Int.J. Innov. Res. Technol*, vol. 3, no. 9, pp. 6-9, 2017
- [10] Singh, A. (2019, 13 February). A Comprehensive Guide to Ensemble Learning (with Python codes) {online}. Available : <https://www.Analyticsvidhya.com/blog/2018/06/comprehendsive-guide-for-ensemble-models>.