

The Development of a Forecast Model for Predicting Product Orders Using Data Mining Techniques

Phatarapon Voraphacha

Department Division of Information and Digital Technology Management, Faculty of
Industrial Technology, Phranakhon Rajabhat University

Corresponding author: phatarapon.v@pnru.ac.th

Received
16/12/2023

Reviewed
23/04/2024

Revised
25/04/2024

Accepted
26/04/2024

Abstract

The objective of this research is to test a forecasting model for predicting product orders and determine the efficiency of the system. The researcher created using the Cross Validation 10 Folds method through data mining techniques and find satisfaction survey results of entrepreneurs who have tried using the forecasting model system to predict product orders using data mining techniques. From the 5-level score criteria using basic statistics, namely finding the average and standard deviation. It was found that the Decision Tree technique (J48) had the highest accuracy and was higher than other techniques. It is calculated as 84.84 percent, including the lowest error value, calculated as 15.16 percent, followed by the technique for searching for friends near the house (K-Nearest Neighbors), which has an accuracy value calculated as 81.60 percent, with an error value calculated in hundreds. 18.40 each and the Naïve Bayes technique has the least accuracy value, accounting for 74.11 percent, which has the highest error value, accounting for 25.89 percent.

Therefore, the Cross Validation 10 Folds method through Decision Tree technique (J48) is used to classify learning data of product orders. There were satisfaction scores regarding the development of the order forecasting model system using data mining techniques on the efficiency of the system. From a trial of 50 stores in the central region, the average was 4.12 and the standard deviation was 0.10. The interpretation score was good.

Keyword: Data mining techniques; Forecasting models; Product orders

การพัฒนาตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล

ภัทรพล วรประชา

สาขาวิชาการจัดการเทคโนโลยีดิจิทัลและสารสนเทศ คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏ
พระนครศรี

ผู้นิพนธ์หลัก: phatarapon.v@pnru.ac.th

บทคัดย่อ

งานวิจัยในครั้งนี้ มีวัตถุประสงค์ เพื่อทดสอบตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อสินค้า และหาประสิทธิภาพของระบบ ที่ผู้วิจัยสร้างขึ้นด้วยวิธีการ Cross Validation 10 Folds ผ่านเทคนิคเหมืองข้อมูล และหาผลสำรวจความพึงพอใจของผู้ประกอบการที่ได้ทดลองใช้ระบบตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล จากเกณฑ์คะแนน 5 ระดับ โดยใช้สถิติพื้นฐาน คือ การหาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน พบว่า เทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) มีค่าความถูกต้องมากที่สุดและสูงกว่าเทคนิคอื่น คิดเป็นร้อยละ 84.84 รวมถึงมีค่าผิดพลาดน้อยที่สุดคิดเป็นร้อยละ 15.16 รองลงมาเป็นเทคนิคค้นหาเพื่อนใกล้บ้าน (K-Nearest Neighbors) มีค่าความถูกต้องคิดเป็นร้อยละ 81.60 มีค่าผิดพลาดคิดเป็นร้อยละ 18.40 และเทคนิคนาอิวเบย์ (Naïve Bayes) มีค่าความถูกต้องน้อยที่สุดคิดเป็นร้อยละ 74.11 ซึ่งมีค่าผิดพลาดมากที่สุดคิดเป็นร้อยละ 25.89

ดังนั้นจึงนำวิธีการ Cross Validation 10 Folds ผ่านเทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) ไปใช้ในการจำแนกประเภทข้อมูลการเรียนรู้ของคำสั่งซื้อสินค้า โดยมีผลคะแนนความพึงพอใจต่อการพัฒนาระบบตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อโดยใช้เทคนิคเหมืองข้อมูลต่อประสิทธิภาพของระบบ จากการทดลองใช้ของผู้ประกอบการในเขตภาคกลาง จำนวน 50 ร้านค้า มีค่าเฉลี่ยอยู่ที่ 4.12 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.10 ระดับคะแนนแปลผลอยู่ในเกณฑ์ดี

คำสำคัญ: เทคนิคเหมืองข้อมูล; ตัวแบบการพยากรณ์คาดการณ์; คำสั่งซื้อสินค้า

บทนำ

Office of the National Economic and Social Development Council Office of the Prime Minister (2022) ได้กล่าวถึงระบบเศรษฐกิจของประเทศไทยในปัจจุบันเป็นการขับเคลื่อนเศรษฐกิจโดยประสิทธิภาพที่อาศัยประสิทธิภาพของภาคการผลิตและคุณภาพสินค้าในการพัฒนาขีดความสามารถในการแข่งขันทางเศรษฐกิจ โดยให้ความสำคัญกับการลงทุนพัฒนาปัจจัยสนับสนุน อาทิ โครงสร้างพื้นฐาน การศึกษา การฝึกอบรมแรงงาน ประสิทธิภาพของตลาดแรงงาน ขนาดของตลาด การพัฒนาตลาดการเงิน ความพร้อมของเทคโนโลยี ซึ่งแม้ประเทศไทยจะมีการพัฒนาปัจจัยสนับสนุนต่าง ๆ ดังกล่าวมาอย่างต่อเนื่อง แต่ก็ยังคงประสบปัญหา ด้านประสิทธิภาพการใช้ทรัพยากร รวมทั้งยังมีอุปสรรคในการยกระดับประสิทธิภาพของตลาดสินค้า ตลาดแรงงาน และประสิทธิภาพของภาครัฐ ที่มีความล่าช้าเมื่อเปรียบเทียบกับหลายประเทศที่เริ่มพัฒนาในช่วงเวลาเดียวกันและสามารถก้าวสู่การเป็นประเทศรายได้สูงได้สำเร็จไปแล้วในช่วงก่อนหน้านี้ ส่งผลให้ประเทศไทยติดกับดักประเทศรายได้ปานกลางมาเป็นเวลานาน จากการจัดสรรทรัพยากรระหว่างภาคเศรษฐกิจที่ผ่านมาที่ทำให้รูปแบบการขยายตัวทางเศรษฐกิจของไทยในช่วงที่ผ่านมาไม่สามารถขับเคลื่อนสู่การเป็นประเทศรายได้สูง อีกทั้ง ยังไม่สามารถตอบสนองต่อโอกาสและทิศทางแนวโน้มการเปลี่ยนแปลงในระดับโลกได้อย่างเต็มศักยภาพ แม้ว่าประเทศไทยจะมีจุดแข็งในการรักษาเสถียรภาพทางเศรษฐกิจ โดยการใช้กลไกทางการคลัง และการบริหารจัดการนโยบายภาครัฐ แต่ยังคงมีปัญหามาจากปัจจัยเชิงโครงสร้าง อาทิ การพึ่งพาต่างประเทศในสัดส่วนสูง ทั้งเงินลงทุน เทคโนโลยี ปัจจัยการผลิต ตลาดสำหรับ การส่งออก แต่บทบาทและอำนาจต่อรองในห่วงโซ่มูลค่าโลกอยู่ในระดับที่ค่อนข้างต่ำรวมทั้งขีดจำกัดเชิงผลิตภาพของเศรษฐกิจโดยรวมที่เป็นอุปสรรคต่อการยกระดับรายได้ และส่งผลให้เศรษฐกิจไทยมีความอ่อนไหวต่อแนวโน้มการเปลี่ยนแปลงหลายประการ

ความก้าวหน้าของเทคโนโลยีสารสนเทศได้มีบทบาทที่สำคัญต่อวิถีชีวิตและสังคมของมนุษย์ เทคโนโลยีสารสนเทศได้สร้างการเปลี่ยนแปลง และโอกาสให้แก่องค์กรและชุมชน เช่น เปลี่ยนโครงสร้างความสัมพันธ์ และการแข่งขันในอุตสาหกรรม ปรับโครงสร้างการดำเนินงานขององค์กรเพิ่มประสิทธิภาพในการผลิต และบริการ เป็นต้น เนื่องจากเทคโนโลยีสารสนเทศก่อให้เกิดรูปแบบใหม่ในการติดต่อสื่อสาร และมีปฏิสัมพันธ์ระหว่างบุคคลทำให้มีการพัฒนา และกระจายตัวของภูมิปัญญา ซึ่งต้องอาศัยบุคคลที่มีความรู้ และความเข้าใจในการใช้งานเทคโนโลยี โดยที่ผู้บริหารจะต้องเตรียมความพร้อมสำหรับองค์กรและปัจจุบัน การใช้เหมืองข้อมูลเป็นเครื่องมือหนึ่งที่ถูกนำมาใช้ในยุคของไทยแลนด์ 4.0 ซึ่งการนำเหมืองข้อมูลเข้าไปช่วยสนับสนุนการตัดสินใจให้กับผู้บริหารองค์กร ผู้ประกอบการ และชุมชนชุมชน Phakdeewattanukul, K. (2010) ได้ศึกษาระบบสนับสนุนการตัดสินใจนี้จะช่วยพัฒนาประสิทธิภาพการทำงานส่วนบุคคล โดยเฉพาะงานที่เกี่ยวข้องกับการตัดสินใจ ซึ่งเป็นงานหลักของผู้บริหาร ระบบจะช่วยจัดเตรียมสารสนเทศที่มีประสิทธิภาพ เพื่อใช้ประกอบการตัดสินใจพัฒนาประสิทธิภาพของการแก้ไขปัญหาโดยช่วยให้ผู้ตัดสินใจสามารถแก้ไขปัญหาได้รวดเร็วและถูกต้องมากยิ่งขึ้น และยังสามารถช่วยตัดสินใจทั้งโครงสร้างและปัญหาที่ไม่มีโครงสร้างได้อย่างมีประสิทธิภาพ Assawamekin, N. (2018) ค้นพบว่าระบบอาจจะมีการจัดเตรียมสารสนเทศเกี่ยวกับการตัดสินใจของ

ปัญหาในลักษณะเดียวกับในอดีต และผลที่ได้รับจากการตัดสินใจนั้น ๆ เพื่อพิจารณาประกอบการตัดสินใจของผู้ใช้ระบบ ซึ่งช่วยให้การตัดสินใจมีความถูกต้อง รวดเร็ว และน่าเชื่อถือมากยิ่งขึ้น ช่วยอำนวยความสะดวกในการติดต่อสื่อสาร

จากความสำคัญดังกล่าว ผู้วิจัยจึงเล็งเห็นถึงการนำความสามารถทางเทคโนโลยีสารสนเทศมาช่วยในการพัฒนาระบบพยากรณ์คาดการณ์ ในการช่วยเหลือผู้ประกอบการในการตัดสินใจผลิตสินค้าตามจำนวนที่สอดคล้องกับความต้องการของผู้บริโภคโดยใช้เทคโนโลยีที่เหมาะสมที่สุด ผู้วิจัยได้ดำเนินการสำรวจความต้องการและความคิดเห็นของผู้ประกอบการร้านค้าในเขตภาคกลาง พบว่ามีอัตราการคงเหลือของการขายจากสต็อกสินค้าแตกต่างกันไป บางปีเหลือเป็นจำนวนมากซึ่งทำให้มีต้นทุนที่ค้างอยู่ อีกทั้งสินค้าที่เหลืออยู่ยังมีคุณภาพที่ลดลงเมื่อนำไปขายในปีถัดไป และบางปีเกิดปัญหาสินค้าไม่เพียงพอในการจำหน่าย จึงทำให้สูญเสียโอกาสทางรายได้เป็นจำนวนมากทำให้ร้านค้าไม่สามารถคาดการณ์ จำนวนที่จะสั่งซื้อในแต่ละประเภท ด้วยปัญหาที่พบจากการสำรวจความต้องการและความคิดเห็นของผู้ประกอบการร้านค้าดังกล่าว ผู้วิจัยจึงทำการพัฒนาตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อ โดยนำเทคนิคเหมืองข้อมูลเข้ามาช่วยในการบริหารจัดการ เพื่อเป็นระบบช่วยเหลือการสั่งซื้อเข้ามาเป็นสต็อกสินค้าของร้านค้า และช่วยสนับสนุนการตัดสินใจในการสั่งซื้อของร้านค้าให้เพียงพอกับความต้องการของผู้บริโภค และไม่ให้สินค้าคงค้างเหลือในสต็อกสินค้าเป็นจำนวนมาก

วิธีการวิจัย

1) ศึกษากระบวนการทำเหมืองข้อมูล โดยผู้วิจัยได้ศึกษากระบวนการที่เป็นมาตรฐานในการดำเนินการที่เรียกว่า คริสป์-ดีเอ็ม (CRISP DM – Cross Industry Standard Process for Data Mining) เพื่อช่วยให้ทำงานมีประสิทธิภาพขึ้น ประกอบไปด้วย 6 ขั้นตอน คือ ความเข้าใจธุรกิจ (Business Understanding), ความเข้าใจข้อมูล (Data Understanding), การเตรียมข้อมูล (Data Preparation), การจัดทำแบบจำลอง (Modeling), การประเมินผล (Evaluation) และการนำเสนอแบบจำลองไปใช้งาน (Deployment) Wilailak, S., Charoenphantarak, W. and Wichadakul, D. (2015)

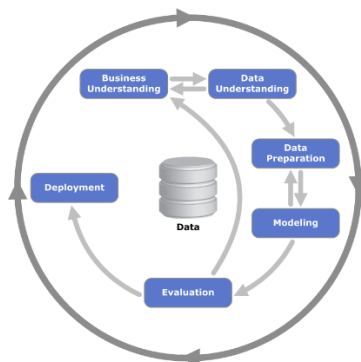


Figure 1 Data mining process, steps of CRISP-DM (CRISP DM).

2) รวบรวมข้อมูลจากกลุ่มตัวอย่าง กรณีศึกษาร้านเสริมภัณฑ์จำหน่ายชุดนักเรียน โดยมีข้อมูลแบบของชุดนักเรียน ประเภทของชุดนักเรียน และขนาดของชุดนักเรียน ที่มีการจัดเก็บย้อนหลังในระยะเวลา 3 ปี โดยแบ่งชุดข้อมูลเพื่อสร้างแบบจำลอง (Training data set) หลังจากที่ได้ออกแบบจำลอง และทำการทดสอบแบบจำลองด้วยข้อมูลอีกกลุ่ม (Testing data set) จำนวน 587 รายการ จากนั้นนำข้อมูลที่ได้ และแปลงข้อมูลโดยแทนชนิดของรายการสินค้าทั้งหมด 45 รายการด้วย X01, X02, X03, ,X45 ตามลำดับ และแปลงข้อมูลของขนาดของสินค้า และข้อมูลทั้งหมดให้อยู่ในรูปแบบของไฟล์ CSV สำหรับเตรียมนำไปใช้ และทดสอบตัวแบบด้วยโปรแกรม WEKA แสดงตัวอย่างใน Figure 2

NO	Product Type	Size	OUT_Buy_2561	IN_STOCK_2561	Stock_Re_maining	OUT_Buy_2562	IN_STOCK_2562	Stock_Re_maining	OUT_Buy_2563	IN_STOCK_2563	Stock_Re_maining
1	X01	30	22	36	14	47	36	3	48	48	3
2	X01	32	57	60	3	82	84	5	81	84	8
3	X01	34	75	84	9	35	36	10	45	36	1
4	X01	36	74	84	10	85	84	9	56	48	1
5	X01	38	90	96	6	89	84	1	54	60	7
6	X01	40	101	108	7	115	108	0	121	144	23
7	X01	42	46	48	2	77	84	9	105	96	0
8	X01	44	55	60	5	62	60	3	105	108	6
9	X01	46	50	60	10	32	24	2	49	48	1
10	X01	48	24	24	0	14	24	10	30	24	4
11	X01	50	11	12	1	4	12	9	9	0	0
12	X01	52	2	12	10	2	0	8	7	0	1
13	X01	54	3	12	9	2	0	7	8	12	11
14	X02	32	2	12	10	6	0	4	8	12	8
15	X02	34	9	12	3	6	12	9	11	12	10
16	X02	36	15	24	9	16	12	5	22	24	7
17	X02	38	23	24	1	57	60	4	39	36	1
18	X02	40	35	36	1	73	72	0	76	84	8
19	X02	42	61	72	11	53	48	6	49	48	5
20	X02	44	22	24	2	35	36	3	37	36	2
21	X02	46	7	12	5	16	12	1	15	24	10
22	X02	48	7	12	5	9	12	8	7	0	1
23	X02	50	4	12	8	3	0	5	9	24	20
24	X03	32	15	24	9	12	12	9	14	12	7
25	X03	34	16	24	8	15	12	5	15	12	2
:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:
586	X45	2529	10	12	2	1	0	1	2	12	11
587	X45	2530	1	12	11	2	0	9	2	0	7

Figure 2 Representation of sales/order data. and stored in .CSV file format.

3) สร้างระบบคาดการณ์คำสั่งซื้อสินค้าจากตัวแบบที่ได้ Chankaew, K. and Piriyaophonsiri, J. (2017) โดยผู้วิจัยเปรียบเทียบประสิทธิภาพของตัวแบบโดยใช้เทคนิคการจำแนกข้อมูลทั้งหมด 3 เทคนิค โดยอ้างอิง ดังต่อไปนี้

3.1) เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นหนึ่งในเทคนิคการทำเหมืองข้อมูลในรูปแบบวิธีการจัดหมวดหมู่ที่รู้จักกันดีที่สุดโดยมักใช้ ตรวจสอบข้อมูล และสร้างต้นไม้เพื่อการพยากรณ์ สำหรับโครงสร้างของต้นไม้ตัดสินใจ จะมีลักษณะคล้ายโครงสร้างต้นไม้ทั่วไป โดยการแตกแขนงไปตามเงื่อนไขหรือเส้นทางของกิ่งไม้ และข้อมูลที่คาดคะเนไว้ว่าจะเกิดขึ้น ซึ่งจะใช้กฎในรูปแบบ “ถ้า (เงื่อนไข) แล้ว (ผลลัพธ์)” (If-then Rule) มาประกอบกรสร้างโครงสร้างต้นไม้ตัดสินใจ สำหรับโครงสร้างต้นไม้ตัดสินใจจะประกอบด้วย

- โหนดภายใน (Internal Node) คือโหนดที่แสดงถึง คุณลักษณะ (Feature) ที่นำมาใช้ในการแบ่งกลุ่มของข้อมูลซึ่ง มีโหนดราก (Root Node) อยู่บนสุดของโครงสร้าง ซึ่งเป็นโหนดที่มีอิทธิพลต่อการจำแนกกลุ่มมากที่สุด

- กิ่ง (Branch) เป็นตัวเชื่อมระหว่างโหนดที่ใช้เป็นเงื่อนไขหรือ ทางเลือกของการกระทำซึ่งมาจากผลลัพธ์ แต่ละตัวของทุกตัว ทำนาย (Predictor) หรือคุณสมบัติ (Feature)

- โหนดใบ (Leaf Node) เป็นโหนดที่แสดงผลลัพธ์ของเงื่อนไข หรือการกระทำตามเงื่อนไขที่เกิดขึ้น ต้นไม้ตัดสินใจ เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ ซึ่งเมื่อมีข้อมูลที่ต้องการจัดกลุ่มก็นำคุณลักษณะต่าง ๆ ของข้อมูลนั้นไปเทียบกับเส้นทางใน ต้นไม้จนกระทั่งคลาสปลายทางซึ่งก็คือกลุ่มของข้อมูลที่เหมือนกันภายในต้นไม้จะประกอบไปด้วย

- โหนด (Node) ซึ่งแต่ละโหนดจะมีคุณลักษณะเป็นตัวทดสอบ

- กิ่งของต้นไม้ (Branch) แสดงถึงค่าที่เป็นไปได้ของคุณลักษณะที่ถูกเลือกทดสอบ

- ใบ (Leaf) ซึ่งเป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจแสดงถึงกลุ่มของข้อมูล (Class)

ก็คือผลลัพธ์ที่ได้จากการทำนายโหนดที่อยู่บนสุดของต้นไม้เรียกว่าโหนดราก (Root Node) โครงสร้างของต้นไม้ตัดสินใจแสดงดัง Figure 3

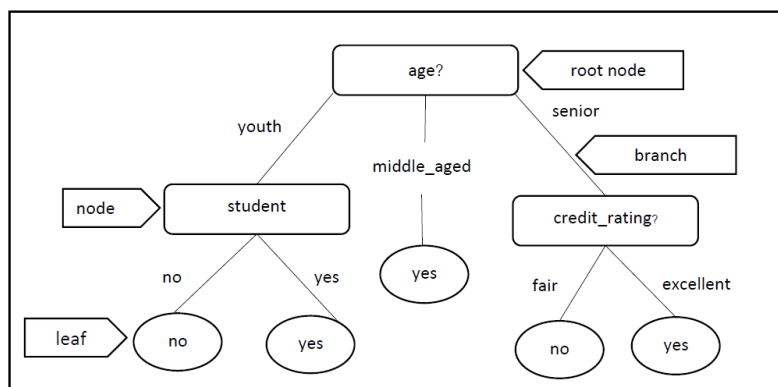


Figure 3 Decision tree example.

3.2) เทคนิคนาอีฟเบย์ (Naïve Bayes) การจำแนกประเภทข้อมูลที่ใช้หลักความน่าจะเป็นอยู่บนพื้นฐานของ Bayes' Theorem และสมมติฐานที่ให้การเกิดของเหตุการณ์ต่าง ๆ เป็นอิสระต่อกัน ถ้ากำหนดให้ $P(h)$ ความน่าจะเป็นที่จะเกิดเหตุการณ์ h และ $P(h|D)$ คือ ความน่าจะเป็นที่จะเกิดเหตุการณ์ h เมื่อเกิดเหตุการณ์ D จากตัวแปรที่กำหนด เราสามารถตรวจจับเหตุการณ์ที่พิจารณาจากเหตุการณ์ต่าง ๆ ได้ดังสมการต่อไปนี้

$$P(h | D) = P(D | h) * P(h) / P(D)$$

จากสมการข้างต้น สามารถสร้างคำนวณความน่าจะเป็นของการจำแนกประเภทแบบเบย์ดังสมการ

$$P(d | h) = P(a_1, \dots, a_n | h) = P(a_1 | h)$$

3.3) เทคนิคค้นหาเพื่อนใกล้บ้าน (K-Nearest Neighbors) เป็นวิธีการแบ่งคลาสสำหรับใช้จัดหมวดหมู่ข้อมูล (Classification) ใช้หลักการเปรียบเทียบข้อมูลที่สนใจกับข้อมูลอื่นว่ามีความคล้ายคลึงมากน้อยเพียงใด หากข้อมูลที่กำลังสนใจนั้นอยู่ใกล้ข้อมูลใดมากที่สุด ระบบจะให้คำตอบเป็นเหมือนคำตอบของข้อมูลที่อยู่ใกล้ที่สุดนั้นลักษณะการทำงานแบบไม่ได้ใช้ข้อมูลชุดเรียนรู้ (Training Data) ในการสร้างแบบจำลองแต่จะใช้ข้อมูลนี้มาเป็นตัวแบบจำลอง เทคนิคนี้จะตัดสินใจว่าข้อมูลมีความคล้ายคลึงหรือใกล้เคียงกับคลาสใด โดยการตรวจสอบข้อมูลบางจำนวน (K) เหมาะสำหรับข้อมูลแบบตัวเลขเพื่อหาวิธีการวัดระยะห่างของแต่ละ Attribute ในข้อมูลให้ได้

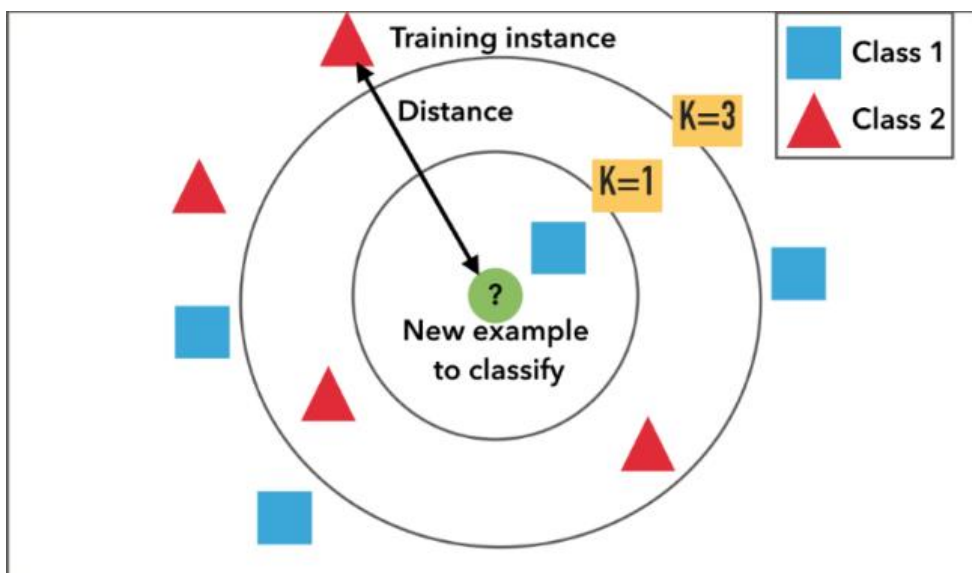


Figure 4 Behavior of K-Nearest Neighbors (K-NN).

4) ทดสอบการคาดการณ์ของระบบโดยทำการเลือกตัวแบบที่ได้ทำการทดสอบระบบที่มีความแม่นยำสูงสุดจากเทคนิคที่ใช้งาน

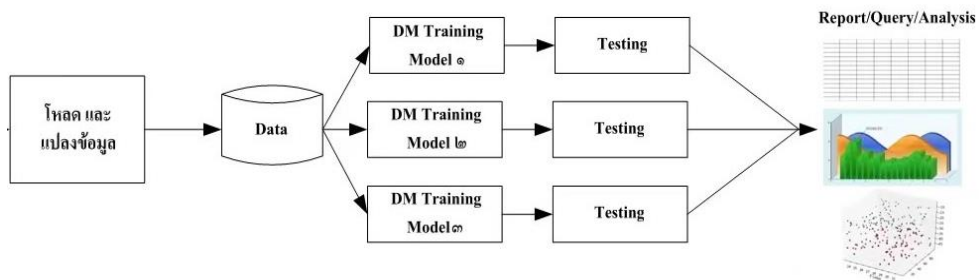


Figure 5 Testing the data import model.

5) นำตัวแบบที่เลือกมาใช้ในการคาดการณ์การสั่งซื้อชุดนักเรียนของร้านเสริมภัณฑ์จากปัจจุบัน โดยทำทดสอบการคาดการณ์โดยใช้ชุดข้อมูลจากการเรียนรู้

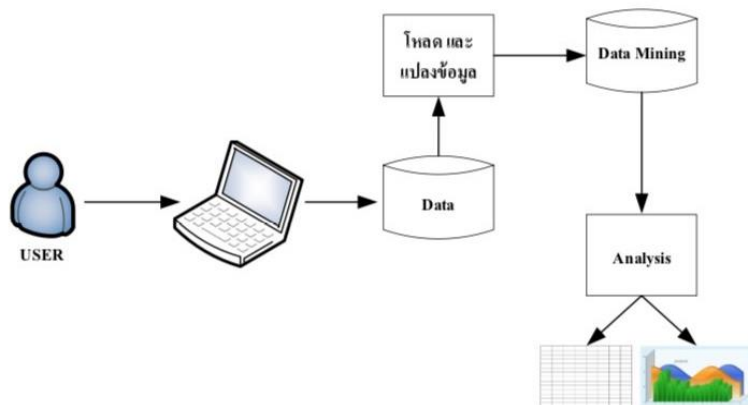


Figure 6 Implementing the model.

6) หาประสิทธิภาพของของระบบ และความพึงพอใจที่มีต่อระบบพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล

ผลการวิจัยและอภิปรายผลการวิจัย

จากการดำเนินงานวิจัย เรื่อง การพัฒนาตัวแบบการพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล ผู้วิจัยสามารถสรุปผลการวิจัยออกเป็น 4 ส่วน ประกอบด้วย

1) ผลการทดสอบตัวแบบการคาดการณ์โดยใช้ชุดข้อมูลจากการเรียนรู้และข้อมูลการทดสอบ การสร้างและทดสอบตัวแบบโดยใช้โปรแกรม WEKA ด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)), เทคนิคนาอิวเบย์ (Naive Bayes) และเทคนิคค้นหาเพื่อนใกล้บ้าน (K-Nearest Neighbors) โดยตัวแบบที่ได้จะอยู่ในรูปของกฎการจำแนกข้อมูลเพื่อจำแนกรายการสั่งซื้อสินค้าเข้าสู่คลังสินค้าตามประเภทของสินค้า และตามขนาดของรายการสินค้า แสดงตาม Table 1 ดังนี้

Table 1 Performance values from testing the model using the Cross Validation 10 Folds method.

Model Testing	Correct Value	Accuracy Value	Remembrance Value	Counterbalance Value
Decision Tree	84.84	87.10	84.50	86.40
Naive Bayes	74.11	78.70	78.70	75.66
K-Nearest Neighbors	81.60	81.60	80.44	81.88

จาก Table 1 แสดงผลการจำแนกประเภทข้อมูลคำสั่งซื้อสินค้า ที่พัฒนาตัวแบบด้วยวิธีการ Cross Validation 10 Folds โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)), เทคนิคนาอิวเบย์ (Naive Bayes) และเทคนิคค้นหาเพื่อนใกล้บ้าน (K-Nearest Neighbors) พบว่าเทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) มีค่าคะแนนร้อยละสูงที่สุดในทุกด้าน โดยมีค่าความถูกต้องเท่ากับร้อยละ 84.84 ค่าความแม่นยำเท่ากับร้อยละ 87.10 ค่าความระลึกเท่ากับร้อยละ 84.50 และค่าความถ่วงดุลเท่ากับร้อยละ 86.40 รองลงมาเป็น เทคนิคค้นหาเพื่อนใกล้บ้าน (K-Nearest Neighbors) มีค่าความถูกต้องเท่ากับร้อยละ 82.65 ค่าความแม่นยำเท่ากับร้อยละ 81.60 ค่าความระลึกเท่ากับร้อยละ 80.44 และค่าความถ่วงดุลเท่ากับร้อยละ 81.88 และเทคนิคนาอิวเบย์ (Naive Bayes) มีค่าคะแนนร้อยละน้อยที่สุด โดยมีค่าความถูกต้องเท่ากับร้อยละ 79.77 ค่าความแม่นยำเท่ากับร้อยละ 78.70 ค่าความระลึกเท่ากับร้อยละ 78.70 และค่าความถ่วงดุลเท่ากับร้อยละ 75.66

2) ผลการเปรียบเทียบตัวแบบการคาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล แสดงตาม Table 2 ดังนี้

Table 2 Results of comparing the accuracy and error values from data classification with Decision Tree technique (J48), Naive Baye technique, and K-Nearest Neighbor technique.

Classification Techniques	Number	Correctness	Percentage	Erroneousness	Percentage
Decision Tree	587	498	84.84	89	15.16

Classification Techniques	Number	Correctness	Percentage	Erroneousness	Percentage
Decision Tree	587	498	84.84	89	15.16
Naive Bayes	587	435	74.11	152	25.89
K-Nearest Neighbors	587	479	81.60	108	18.40

จาก Table 2 แสดงผลการเปรียบเทียบ พบว่าการจำแนกประเภทข้อมูลคำสั่งซื้อสินค้า โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) ผลการจำแนกประเภทข้อมูลมีความถูกต้องคิดเป็นร้อยละ 84.84 และมีค่าผิดพลาดคิดเป็นร้อยละ 15.16 เทคนิคนาอิวเบย์ (Naive Bayes) ผลการจำแนกประเภทข้อมูลมีความถูกต้องคิดเป็นร้อยละ 74.11 และมีค่าผิดพลาดคิดเป็นร้อยละ 25.89 และเทคนิคความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) ผลการจำแนกประเภทข้อมูลมีความถูกต้องคิดเป็นร้อยละ 81.60 และมีค่าผิดพลาดคิดเป็นร้อยละ 18.40 ตามลำดับ แสดงให้เห็นว่าเทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) มีค่าความถูกต้องสูงที่สุด และค่าผิดพลาดน้อยที่สุด เมื่อเปรียบเทียบกับเทคนิคนาอิวเบย์ (Naive Bayes) และเทคนิคความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor)

3) ผลการหาประสิทธิภาพของระบบพยากรณ์คาดการณ์คำสั่งซื้อสินค้า โดยใช้เทคนิคเหมืองข้อมูล

จากผลการทดสอบตัวแบบ ผู้วิจัยจึงเลือกใช้ การจำแนกประเภทข้อมูลคำสั่งซื้อสินค้า โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)) ทำการวัดค่าประสิทธิภาพของตัวแบบการจำแนกประเภทข้อมูลของเทคนิคมาใช้ในการคาดการณ์การสั่งซื้อปี พ.ศ. 2565 และจำนวนที่ขายได้จริงในปี พ.ศ. 2565 เพื่อทำการทดสอบประสิทธิภาพ โดยใช้ข้อมูลของปี พ.ศ. 2565 ซึ่งประกอบไปด้วยข้อมูลรายการสินค้าจำนวน 587 รายการ โดยมีจำนวนที่สั่งซื้อ แสดงดัง Figure 7

ลำดับ	ประเภทสินค้า	ขนาด	พ.ศ. 2565
			จำนวนชุดนักเรียนที่สั่งซื้อ
1	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 30	60
2	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 32	84
3	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 34	48
4	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 36	60
5	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 38	60
6	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 40	120
7	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 42	108
8	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 44	108
9	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 46	60
10	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 48	36
11	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 50	12
12	เสื้อเชิ้ตนักเรียนชายเสริ	เบอร์ 52	12

Figure 7 Example of forecasting orders for school uniforms in 2022, 587 items, Seriphan Shop.

จาก Figure 7 แสดงรายการการคาดการณ์คำสั่งซื้อสินค้า จำนวน 587 รายการ ร้านเสรีรักษ์ เมื่อได้ข้อมูลจำนวนการคาดการณ์ที่จะสั่งซื้อ ผู้วิจัยจึงต้องนำจำนวนที่จะตัดสินใจในการสั่งซื้อแต่ละรายการมาทำการหักกับจำนวนสต็อกสินค้าคงเหลือในปี พ.ศ. 2564 โดยหักจากจำนวนสินค้าที่มีจะสั่งเพิ่มเป็นจำนวนโหล (12 ชิ้น) ต่อรายการสินค้าเนื่องจากการคาดการณ์คำสั่งซื้อจะออกมาในรูปแบบจำนวนชิ้น

Table 3 Example of the performance of the forecasting system for predicting orders and actual sales in 2022, Seriphan Shop.

No	Product	Size	2021		2022		Accuracy (Percent age)
			Remainin g	Quanti ty Sold	Order Quantit y	Rem ainin g	
1	Male student shirt	Size 30	0	49	60	11	100
2	Male student shirt	Size 32	6	85	84	5	100
3	Male student shirt	Size 34	2	47	48	3	100
4	Male student shirt	Size 36	4	55	60	9	100
...							
...							
586	School skirt, 6 pleats, durian, blue look	Navy blue 2529	10	2	0	8	100
587	School skirt, 6 pleats, durian, blue look	Navy blue 2530	7	2	0	5	100

จาก Table 3 แสดงจำนวนที่สั่งซื้อกับจำนวนที่ขายได้จริงใน พ.ศ. 2565 โดยมีการแสดงประเภทสินค้าจำนวน 587 รายการ ร้านเสรีรักษ์ แบ่งเป็น ขนาดของประเภทสินค้า, จำนวนสินค้า

คงเหลือในปี พ.ศ. 2564, จำนวนชุดนักเรียนที่ขายได้ในปี พ.ศ. 2565, จำนวนชุดนักเรียนที่สั่งซื้อในปี พ.ศ. 2565 (โดยมีการสั่งซื้อมาเป็นจำนวนโหลและแปลงข้อมูลเข้าเป็นจำนวนชิ้น), จำนวนสินค้าคงเหลือในปี พ.ศ. 2565 และร้อยละของความถูกต้อง ซึ่งจากข้อมูลในตารางดังกล่าวพบว่ามีความถูกต้อง และสินค้าคงเหลือพอกับ สต็อกสินค้า

4) ผลความพึงพอใจที่มีต่อระบบพยากรณ์และคาดการณ์คำสั่งซื้อสินค้า โดยใช้เทคนิคเหมืองข้อมูล

ผู้วิจัยได้นำระบบตัวแบบพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูลไปให้ผู้ประกอบการร้านค้าในพื้นที่ภาคกลาง จำนวน 50 ร้าน ทดลองใช้งานการพัฒนาระบบตัวแบบพยากรณ์คาดการณ์คำสั่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล และได้ทำการประเมินความพึงพอใจที่มีต่อประสิทธิภาพของระบบ โดยสรุปผลความพึงพอใจได้ตาม Table 4 ดังต่อไปนี้

Table 4 Satisfaction evaluation results from the trial of the development of a product order forecasting model system using data mining techniques on system performance.

N = 15

Satisfaction Measurement Issues	Average ($\bar{x}$)	Standard Deviation (S.D.)
1. Accessing the system		
1.1 The program's capabilities meet the needs of the user.	4.12	0.10
1.2 The program can be run according to the steps.	4.13	0.11
1.3 The nature of the program is easy to user.	4.11	0.10
1.4 The program is as effective as the user desires.	4.12	0.12
1.5 The program provides data security.	4.11	0.11
2. Forecasting product orders		
2.1 The forecast results predict orders from the program according to your needs.	4.11	0.11
2.2 The nature of the program is easy to forecast orders.	4.13	0.11
2.3 The program's effective order forecasting can be used appropriately.	4.12	0.11
Satisfaction with overall system performance	4.12	0.10

จาก Table 4 แสดงผลสรุปความพึงพอใจในการใช้การพัฒนาระบบตัวแบบพยากรณ์คาดการณ์ค่าส่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล จากจำนวนร้านค้าที่ได้ทดลองใช้ตัวแบบการพยากรณ์คาดการณ์ที่ผู้วิจัยสร้างขึ้น จำนวน 50 ราย ได้ผลการประเมินความพึงพอใจต่อประสิทธิภาพของระบบ มีค่าเฉลี่ยเท่ากับ 4.12 ส่วนเบี่ยงเบนมาตรฐาน 0.10 แสดงว่าระบบที่พัฒนาขึ้นนั้นอยู่ในระดับความพึงพอใจอยู่ในระดับดี โดยการเข้าใช้งานระบบมีประสิทธิภาพ มีความสะดวก และตรงกับความต้องการของผู้ใช้งาน มีการรักษาความปลอดภัยของข้อมูล และผลของการพยากรณ์คาดการณ์มีความน่าเชื่อถือ จากผลการประเมินความพึงพอใจเห็นไปทิศทางเดียวกันว่าระบบนั้นมีความเหมาะสมกับการใช้งานจริง

สรุปผลการวิจัย

งานวิจัยในครั้งนี้นำเสนอการพัฒนาตัวแบบการพยากรณ์คาดการณ์ค่าส่งซื้อสินค้าโดยใช้เทคนิคเหมืองข้อมูล จากปัญหาของงานวิจัยในเรื่องอัตราการคงเหลือของการขายจากสต็อกสินค้าเป็นข้อมูลที่ได้จากความต้องการของผู้ประกอบการร้านค้าในเขตภาคกลาง โดยเลือกใช้ 3 โมเดลประกอบไปด้วย Decision Tree (J48), Naive Bayes และ K-Nearest Neighbors ใช้โปรแกรม Weka สำหรับสร้างโมเดล และวัดประสิทธิภาพ โดยการวัดประสิทธิภาพใช้ 10 Folds Cross Validation และมีผลคะแนนความพึงพอใจต่อการพัฒนาระบบตัวแบบพยากรณ์คาดการณ์ค่าส่งซื้อ โดยใช้เทคนิคเหมืองข้อมูลต่อประสิทธิภาพของระบบมีค่าเฉลี่ยอยู่ที่ 4.12 อยู่ในเกณฑ์ดี โดยผลการทดลองพบว่า Decision Tree (J48) ให้ค่าความถูกต้องมากที่สุดและสูงกว่าเทคนิคอื่น คิดเป็นร้อยละ 84.84 โดยงานวิจัยนี้เป็นงานวิจัยที่เป็นประโยชน์ เนื่องจากสามารถนำโมเดลที่ได้มาช่วยทำนายผลลัพธ์ที่อาจเกิดขึ้นในอนาคตได้ ซึ่งสอดคล้องกับงานวิจัยเรื่องการเปรียบเทียบอังกอร์ทีมเหมืองข้อมูลเพื่อวิเคราะห์ปัจจัยที่ส่งผลกระทบต่อระดับการเรียนของนักศึกษา Pharasamrit, Y., Phubunop, J. and Phongsiri, W. (2013) ประกอบด้วยเทคนิค C4.5, Naive Bayes และ K-Nearest Neighbor เทคนิคที่ให้ค่าประสิทธิภาพสูงที่สุดจะถูกนำมาใช้ในการค้นหาปัจจัยที่ส่งผลกระทบต่อระดับผลการเรียนของนักศึกษา ผลการเปรียบเทียบพบว่าแบบ C4.5 ให้ค่าประสิทธิภาพสูงที่สุด ร้อยละ 73.55 ซึ่งมากกว่าแบบ k-Nearest Neighbor และแบบ Naive Bayes ซึ่งมีค่าร้อยละ 66.63 และร้อยละ 49 ตามลำดับ ผลการค้นหาปัจจัยที่ส่งผลกระทบต่อระดับผลการเรียนของนักศึกษา พบว่า ตัวแปรที่มีความสำคัญเรียงจากมากไปหาน้อย คือ ชั้นปีจำนวนพี่น้องที่กำลังศึกษา อายุ จำนวนพี่น้องทั้งหมด และสาขาวิชา นอกจากการค้นหาปัจจัยแล้วสามารถนำกฎการจำแนกข้อมูลที่ได้มาใช้ในการพัฒนาระบบพยากรณ์ระดับผลการเรียนของนักศึกษาได้อีกด้วย

ปัจจุบันมีนักวิจัยทางการศึกษาหลายท่านได้ให้ความสนใจในการนำเทคนิคเหมืองข้อมูลมาใช้วิเคราะห์ข้อมูลของงานวิจัยของ Ahmad, N. B. Hj. Shamsuddin and S. M. (2010) ได้ทำการเปรียบเทียบประสิทธิภาพการจำแนกรูปแบบการเรียนรู้ของผู้เรียนโดยใช้เทคนิคต้นไม้ตัดสินใจเบย์และกฎความสัมพันธ์ ผลการวิจัยพบว่า เทคนิคต้นไม้ตัดสินใจมีประสิทธิภาพค่าความถูกต้องในการจำแนกเท่ากับร้อยละ 93.58 สูงกว่าเทคนิคอื่น ๆ สามารถนำกฎการจำแนกที่ได้ไปใช้ในการจำแนก

ข้อมูลรูปแบบการเรียนรู้ของผู้เรียนได้อย่างมีประสิทธิภาพ และยังสามารถเชื่อมโยงกับงานวิจัยของ Pantoo, O. (2016) ที่นำเทคนิคต้นไม้ตัดสินใจไปใช้ในการวิเคราะห์รูปแบบการเรียนรู้ VARK สำหรับ นักศึกษาระดับปริญญาตรี มหาวิทยาลัยราชภัฏเชียงใหม่ ผลการวิจัยพบว่า ประสิทธิภาพการจำแนก ข้อมูลมีประสิทธิภาพค่าความถูกต้องคิดเป็นร้อยละ 79.74 ซึ่งผลที่ได้สามารถนำไปพัฒนากิจกรรมการเรียนรู้ให้เหมาะสมกับรูปแบบการเรียนรู้ของผู้เรียนแต่ละบุคคลได้เป็นอย่างดี จากรายละเอียดดังกล่าว ข้างต้น สามารถสรุปได้ว่าการจัดการเรียนการสอนควรมุ่งเน้นการพัฒนาผู้เรียนตามศักยภาพและความถนัดรวมถึงแตกต่างระหว่างบุคคลโดยเน้นผู้เรียนเป็นสำคัญ สำหรับการวิเคราะห์รูปแบบการเรียนรู้ปัจจุบันได้มีการนำทฤษฎีเหมืองข้อมูลเข้ามาใช้ในการจำแนกข้อมูลกันอย่างแพร่หลาย เช่น การใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree (J48)), เทคนิคนาอิวเบย์ (Naïve Bayes) และ เทคนิคความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) เป็นวิธีหนึ่งที่สำคัญในการจำแนกประเภทข้อมูล ซึ่งเทคนิคเหล่านี้ง่ายต่อการปรับเปลี่ยนเป็นกฎการจำแนกประเภทข้อมูลและให้ผลการพยากรณ์ที่แม่นยำ จึงเป็นที่นิยมในการนำมาใช้ในการจำแนกรูปแบบการเรียนรู้ด้านต่าง ๆ กันอย่างแพร่หลาย

กิตติกรรมประกาศ

สาขาวิชาการจัดการเทคโนโลยีดิจิทัลและสารสนเทศ คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏพระนคร (contract no.02.003/65).

References

- Ahmad, N. B. Hj. Shamsuddin and S. M. (2010). A comparative analysis of mining techniques for Automatic detection of student's learning style. *Intelligent Systems Design and Applications, Cairo, Egypt*, pp. 877-882.
- Assawamekin, N. (2018). *Basic principles of software engineering*, Publisher: SE-ED Education, Bangkok.
- Chankaew, K. and Piriyaiphonsiri, J. (2017). Decision support system for purchasing life insurance. Hat Yai National and International Academic Conference, Hat Yai University, No.8, pp. 370-380.
- Office of the National Economic and Social Development Council Office of the Prime Minister. (2022). *National Economic and Social Development Plan No. 13 (2023-2027)*. Royal Gazette.
- Pakdeewattanakul, K. (2010). *Decision support system scripture and expert system*, Publisher: KTP.
- Pantoo, O. (2016). *Synthesis of an adaptive e-learning model with a mentor system electronics Using the VARK learning model analyzed using data mining methods. Doctor of Philosophy in Information Technology*, Graduate School King Mongkut's University of Technology Klao Phra Nakhon Nuea.
- Pharasamrit, Y., Phubunop, J. and Phongsi, W. (2013). Comparison of data mining algorithms to analyze factors affecting students' learning levels. *Journal of Science and Technology Mahasarakham University*, Vol. 9, pp. 281-289.
- Wilailak, S., Charoenphantarak, W. and Wichadakul, D. (2015). Using data mining techniques to predict student academic performance Demonstration School of Kasetsart University Kamphaeng Saen Campus Educational Research and Development Center. *Veridian E-Journal Science and Technology Silpakorn University*, Vol. 2, No. 2, pp. 1-17.