

**โมเดลเพื่อทำนายความสามารถในการชำระหนี้เงินกู้ กรณีศึกษาสำนักงานส่งเสริมสวัสดิการ  
และสวัสดิภาพครูและบุคลากรทางการศึกษา จังหวัดอุดรธานี**

**The Model of Predicting Settlement Ability: A Case Study in Office of the Welfare Promotion  
Commission for Teachers and Educational Personnel, Udonthani Province, Thailand**

ญาณินี สิทธิสาร (Yaninee Sitthisan)\* ดร.บุษดี ศิริแสงตระกูล (Dr.Pusadee Seresangtakul)\*\*

**บทคัดย่อ**

งานวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อนำเสนอโมเดลทำนายความสามารถในการชำระหนี้เงินกู้ของสมาชิกสำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา จังหวัดอุดรธานี ผู้วิจัยได้ศึกษาและรวบรวมปัจจัยที่มีคาดว่าจะมีผลต่อการความสามารถในการชำระหนี้เงินกู้ จำนวน 8 ปัจจัย ผู้วิจัยได้ทำการวิเคราะห์ความสัมพันธ์ของปัจจัยโดยใช้ การวิเคราะห์การถดถอยพหุคูณ ผลการศึกษาพบว่าปัจจัยที่มีความสัมพันธ์กับความสามารถในการชำระหนี้เงินกู้ได้แก่ จำนวนบุตร เพศ อายุ รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ การทำประกันเงินกู้ จำนวนเงินงวด ข้อมูลได้จากการรวบรวมประวัติการกู้เงินระหว่างปี พ.ศ. 2547 - 2557 จำนวนทั้งสิ้น 11,732 รายการ การสร้างโมเดลเพื่อทำนายความสามารถในการชำระหนี้เงินกู้ โดยใช้ต้นไม้การตัดสินใจ เครือข่ายความเชื่อเบย์ และโครงข่ายประสาทเทียม ผลการศึกษาพบว่าประสิทธิภาพการทำนายของโมเดลโครงข่ายประสาทเทียมพบค่าความแม่นยำสูงสุดที่ 98.77 โมเดลเครือข่ายความเชื่อเบย์ ที่ค่าความแม่นยำ 98.76 และโมเดลต้นไม้การตัดสินใจ ที่ค่าความแม่นยำ 98.74 ตามลำดับ เมื่อเปรียบเทียบประสิทธิภาพการทำนายแล้ว พบว่าโมเดลโครงข่ายประสาทเทียมมีประสิทธิภาพการทำนายสูงที่สุด

**ABSTRACT**

The purpose of this research is to present a model to predict the ability to repay a loan and to Members of the Office for the Promotion of the Benefits and Welfare of Teachers and Educational Personnel Udon Thani Province. In the study, the researchers have studied and summarized the factors that are expected to affect the ability to pay loan amount 8 factors. We analyzed the relationship of factors by using regression. The results showed that the factors that effect to loan paying ability were number of children, gender, age, salary, loan, debt, Loan insurance and period amounts. In order to construct the loan paying prediction model, we collected history of the loan data from 2004 to 2014, totally 11,732 items. We constructed Model for predict repayment ability by using decision tree, Bayesian belief network and artificial neural network. The results showed that predictive efficiency; the accuracy values were 98.77, 98.76 and 98.74 with artificial neural network, Bayesian belief network and decision tree, respectively. By comparing predictive efficiency, neural network model had the highest predictive efficiency.

**คำสำคัญ:** โครงข่ายประสาทเทียม ต้นไม้การตัดสินใจ เครือข่ายความเชื่อเบย์

**Keywords:** Artificial neural network, Decision tree, Bayesian belief network

\*นักศึกษาคณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

\*\*ผู้ช่วยศาสตราจารย์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

## บทนำ

การมาปนกิจสงเคราะห์ช่วยเหลือเพื่อนครูและบุคลากรทางการศึกษาเรียกโดยย่อว่า ข.พ.ค. (สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา, 2559) มีเจตนารมณ์เพื่อเป็นสวัสดิการช่วยเหลือสงเคราะห์สมาชิกในการจัดการศพให้เป็นไปอย่างสมเกียรติของผู้ที่เคยเป็นครูรวมทั้งสงเคราะห์ครอบครัวของสมาชิก สมาชิกได้รับประโยชน์ในด้านสวัสดิการเป็นอย่างดี ในปี พ.ศ.2547 สภาวะเศรษฐกิจมีความผันผวนส่งผลกระทบต่อความเป็นอยู่ของสมาชิก สำนักงานส่งเสริมสวัสดิภาพและสวัสดิการครูและบุคลากรทางการศึกษา ซึ่งเป็นหน่วยงานช่วยเหลือดูแลได้ร่วมมือกับสถาบันการเงินในการลงนามบันทึกข้อตกลงโครงการสวัสดิการเงินกู้ ข.พ.ค. ขึ้นเพื่อสร้างสวัสดิการให้กับสมาชิก สามารถนำเงินจำนวนหนึ่งมาใช้จ่ายเพื่อพัฒนาคุณภาพชีวิตของตนเองและครอบครัว ปัจจุบัน ข.พ.ค. ดำเนินงานของโครงการสวัสดิการเงินกู้ ข.พ.ค. มาแล้วทั้งสิ้น 7 โครงการ (สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา, 2554)

ปัจจุบันการอนุมัติเงินกู้โดยไม่มีภาระประเมินความเสี่ยงด้านเครดิตของสมาชิก ข.พ.ค. มีผลทำให้เกิดหนี้ค้างชำระเงินกู้เป็นจำนวนมาก การประเมินผ่านการตัดสินใจและหลังการตัดสินใจซึ่งทำโดยเจ้าหน้าที่สินเชื่อ ใช้ประสบการณ์ของตนเองและการวิเคราะห์ข้อมูล ดังนั้นการตัดสินใจให้สินเชื่อเงินกู้ซึ่งขึ้นกับบุคคลและใช้เวลานานเนื่องจากสมาชิกที่ได้รับเครดิตสินเชื่อผิดนัดการชำระเงินกู้ตามสัญญาข้อตกลงที่ทำไว้ ซึ่งจากการตรวจสอบข้อมูลและการติดตามทวงหนี้กับสมาชิกที่ค้างชำระทำให้ทราบว่าสมาชิกที่กู้เงินมีหนี้สินกับสถาบันการเงินอื่น เป็นผลให้ไม่สามารถชำระหนี้ให้แก่หน่วยงานได้ ทำให้หน่วยงานสำนักงาน สกสค. ที่ทำเอกสารบันทึกข้อตกลงกับสถาบันทางการเงินที่อนุมัติสินเชื่อให้กับสมาชิก ข.พ.ค. โดยที่หากมีสมาชิก ข.พ.ค. ที่กู้เงินค้างชำระ 3 เดือน หรือคิดเป็น 1 ไตรมาส สำนักงาน สกสค. จะต้องทวงเงินจ่ายแทนสมาชิก ข.พ.ค. ที่ค้างชำระ ขณะนี้หน่วยงานได้ทวงจ่ายเงินแทนมาเป็นเวลากว่า 7 ปี (สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา, 2554) คิดเป็นจำนวนเงินหลายร้อยล้านบาท ส่งผลกระทบต่อความมั่นคงทางการเงินของหน่วยงานจึงจำเป็นต้องมีการวิเคราะห์เครดิตสมาชิกเพื่อช่วยตัดสินใจในการที่จะให้เครดิตกับสมาชิกหรือไม่

ผู้วิจัยได้ศึกษาปัจจัยที่มีผลต่อการชำระหนี้เงินกู้ ข.พ.ค. โดยอาศัยข้อมูลในอดีตของสมาชิก ข.พ.ค. ที่กู้เงินตั้งแต่ปี พ.ศ. 2547 – 2557 ด้วยการนำเทคนิคเหมืองข้อมูล (Data Mining Technique) มาประยุกต์ใช้กับงานวิจัยประกอบด้วย เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เครือข่ายความเชื่อเบย์ (Bayesian Belief Network) และโครงข่ายประสาทเทียม (Artificial neural network) ซึ่งเทคนิคดังกล่าว มีหลักเกณฑ์และกระบวนการคำนวณที่แตกต่างกัน ทำให้ผลลัพธ์ที่ได้จะมีความแตกต่างกันออกไปด้วย ผลลัพธ์จากงานวิจัยจะทำให้ทราบว่า การประเมินความเสี่ยงของผู้ที่คาดว่าจะมีผลต่อการเกิดหนี้ค้างชำระ จะต้องพิจารณาจากปัจจัยใดบ้าง จึงจะสามารถลดจำนวนการเกิดหนี้ค้างชำระเงินกู้ ข.พ.ค. รวมถึงสามารถลดความเสี่ยงในการปล่อยกู้ของลูกค้ายที่มีความประสงค์ที่จะกู้เงินกับสถาบันทางการเงินต่าง ๆ ในอนาคต

## วัตถุประสงค์ของงานวิจัย

1. เพื่อศึกษาปัจจัยที่มีผลต่อการชำระหนี้เงินกู้ของสมาชิก ข.พ.ค.
2. เพื่อสร้างโมเดลใช้ทำนาย ความสามารถในการชำระหนี้เงินกู้ ข.พ.ค.

## วิธีการวิจัย

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

#### 1. การวิเคราะห์การถดถอยพหุคูณ (Multiple regression analysis)

การวิเคราะห์การถดถอยพหุคูณ (Rodgers, 1988) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรต้นที่ทำกับตัวแปรตาม โดยวิธีการคัดเลือกแบบขั้นตอน (Stepwise Selection) การคัดเลือกแบบนี้เป็นการผสมผสานระหว่างวิธีการคัดเลือกตัวแปรการทำนายทั้งแบบก้าวหน้าและแบบถอยหลังเข้าด้วยกัน ซึ่งในขั้นแรกจะมีการเลือกตัวแปรที่มีค่าสัมประสิทธิ์สหสัมพันธ์กับตัวแปรตามสูงที่สุดเข้าสมการก่อน หลังจากนั้นก็จะทดสอบตัวแปรที่ไม่ได้อยู่ในสมการว่ามีตัวแปรใดบ้างที่สามารถเข้ามาอยู่ในการสมการด้วยวิธีการคัดเลือกแบบก้าวหน้า (Forward Selection) พร้อมทั้งจะมีการทดสอบตัวแปรที่อยู่ในสมการด้วยว่า ตัวแปรการทำนายที่อยู่ในสมการตัวใดมีโอกาสที่จะถูกขจัดออกจากสมการด้วยวิธีการคัดเลือกแบบถอยหลัง (Backward Selection) โดยจะกระทำการคัดเลือกผสมทั้งสองวิธีนี้ในทุกขั้นตอนจนกระทั่งไม่มีตัวแปรใดที่ถูกคัดออกจากสมการและไม่มีตัวแปรใดที่จะถูกนำเข้ามาเข้าสมการกระบวนการก็จะยุติและได้สมการถดถอยที่มีสัมประสิทธิ์การพยากรณ์สูงสุด สามารถคำนวณได้จากสมการที่ (1)

$$r = \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{[N \sum y^2 - (\sum x)^2][N \sum x^2 - (\sum y)^2]}} \quad (1)$$

โดย  $x$  คือ ตัวแปรอิสระ  
 $y$  คือ ตัวแปรตาม  
 $N$  คือ ขนาดของกลุ่มตัวอย่าง

#### 2. ต้นไม้ตัดสินใจ (Decision tree)

ต้นไม้ตัดสินใจ (Surjeet, 2012) เป็นการเรียนรู้จากการคาดคะเนเหตุการณ์ต่างๆ ที่อาจเกิดขึ้น โดยอาศัยเงื่อนไขเป็นตัวช่วยในการตัดสินใจ ค่า Information gain จะถูกนำมาใช้ในการเลือกคุณลักษณะ (attribute) ในแต่ละโหนดของต้นไม้ตัดสินใจ คุณลักษณะตัวใดที่มีค่า Information gain สูงสุด จะเป็นตัวที่ถูกเลือก จำนวนจากสมการที่ (2)

$$Gain(S, A) = E(S) - \sum_{v=value(A)} \frac{|S_v|}{|S|} E(S_v) \quad (2)$$

โดย  $S$  คือ ตัวอย่างที่ประกอบด้วยชุดข้อมูลตัวแปรต้นและตัวแปรตาม  
 $A$  คือ ตัวแปรต้นที่พิจารณา  
 $E$  คือ เอนโทรปีของตัวอย่าง  
Value (A) คือ เซตของค่าของ  $A$  ที่เป็นไปได้  
 $S_v$  คือ ตัวอย่างที่  $A$  มีค่า  $v$  ทั้งหมด

### 3. โมเดลเครือข่ายความเชื่อเบย์ (Bayesian Belief Network)

เครือข่ายความเชื่อเบย์ เป็นแบบจำลองกราฟของความน่าจะเป็น (Gang M, 2010) อธิบายได้ว่าโหนดใดบ้างที่ขึ้นกับโหนดอื่น โหนดใดบ้างที่ไม่ขึ้นกับโหนดอื่น พร้อมทั้งอธิบายความสัมพันธ์ระหว่างตัวแปร แล้วแสดงผลออกมาในภาพแบบโมเดล ซึ่งสามารถอธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไข (Condition Independent) ระหว่างตัวแปร คำนวณได้จากสมการที่ (3)

$$P(Y_i | X) = \frac{P(Y_i \cap X)}{\sum_{i=1}^n P(X | Y_i) P(Y_i)} \quad (3)$$

|     |                 |                                                                             |
|-----|-----------------|-----------------------------------------------------------------------------|
| โดย | $P(Y_i   X)$    | คือ ความน่าจะเป็นที่เหตุการณ์ $Y_i$ จะเกิดขึ้นถ้าเหตุการณ์ $X$ เกิดขึ้นแล้ว |
|     | $P(Y_i \cap X)$ | คือ ความน่าจะเป็นของ $Y_i$ เมื่อเกิดเหตุการณ์ $X$                           |
|     | $P(X   Y_i)$    | คือ ความน่าจะเป็นที่เหตุการณ์ $X$ จะเกิด ถ้าเหตุการณ์ $Y_i$ เกิดขึ้นแล้ว    |
|     | $P(Y_i)$        | คือ ความน่าจะเป็นที่เหตุการณ์ $Y_i$ จะเกิด                                  |
|     | $n$             | คือ จำนวนเหตุการณ์ทั้งหมด                                                   |

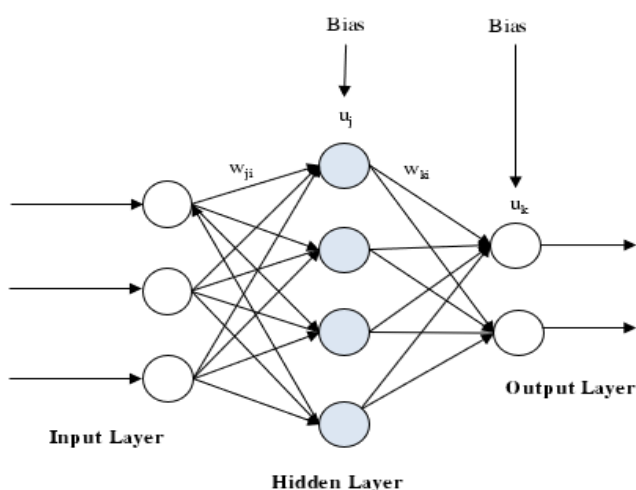
### 4. โครงข่ายประสาทเทียม (Artificial neural network)

โครงข่ายประสาทเทียม (Fayyad, 1996) เป็นการจำลองการทำงานของโครงข่ายประสาทของมนุษย์ โดยการทำงานของเทคนิคนี้ จะประกอบด้วยชั้นของข้อมูลเบื้องต้น 3 ชั้น ได้แก่ ชั้นนำเข้า (input layer) ชั้นซ่อน (hidden layer) และชั้นส่งออก (output layer)

การปรับสอนโครงข่าย นิยมใช้แบบโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back Propagation Neural Network) ซึ่งการปรับสอนประกอบด้วย 3 ขั้นตอน คือ

- 1) ขั้นตอนการคำนวณไป ข้างหน้า (feed forward) จากชั้นอินพุตไปยังชั้นซ่อน และไปสู่ชั้นเอาต์พุต
- 2) ขั้นตอนการคำนวณและการ แทนค่ากลับของผลรวมของผลคูณของผลลัพท์ กับค่าเป้าหมาย
- 3) ขั้นตอนการปรับค่าน้ำหนัก (weight) ไปเรื่อย ๆ เป็นขั้นตอนการคำนวณไปข้างหน้าเพียงขั้นตอนเดียว

โครงสร้างของโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ แสดงได้ดังภาพที่ 1



ภาพที่ 1 โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ 1 ชั้นซ่อน

โดยสามารถคำนวณได้จากสมการที่ (4)

$$u_j^{(l)} = \sum_{i=1}^p w_{ji}^{(l-1)} y_i^{(l-1)} + Bias \quad (4)$$

|     |                |                                                                                                                      |
|-----|----------------|----------------------------------------------------------------------------------------------------------------------|
| โดย | $y_i^{(l-1)}$  | แทนค่าเอาต์พุตโหนดที่ $i$ ในชั้นก่อนหน้า ( $l-1$ )                                                                   |
|     | $w_{ji}^{(l)}$ | แทนค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างโหนดที่ $j$ ชั้นที่ 1 กับโหนดที่ $i$ ชั้นที่ ( $l-1$ )                           |
|     | $n$            | แทนจำนวนนิวรอนของข้อมูลนำเข้า                                                                                        |
|     | $p$            | แทนค่าจำนวนโหนดชั้นที่ ( $l-1$ )                                                                                     |
|     | $u_j^{(l)}$    | แทนค่าผลลัพธ์ที่ได้จากการคำนวณ และจะต้องนำค่าดังกล่าวไปผ่านฟังก์ชันกระตุ้นเพื่อ<br>ได้เอาต์พุตที่ต้องการ $y_j^{(l)}$ |
|     | $Bias$         | แทนค่าความโน้มเอียง                                                                                                  |

ถ้าผลลัพธ์เป็นค่าตัวเลขที่ต่อเนื่อง ให้คำนวณด้วย Continuous Function ซึ่งนิยมใช้ Sigmoid function ดังสมการที่ (5)

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (5)$$

|        |        |                            |
|--------|--------|----------------------------|
| โดยที่ | $f(x)$ | เป็นฟังก์ชันกระตุ้น        |
|        | $e$    | แทนฟังก์ชันเอกซ์โพเนนเชียล |
|        | $x$    | แทนค่าข้อมูลนำเข้า         |

#### งานวิจัยที่เกี่ยวข้อง

การใช้เหมืองข้อมูลทำนายการจำแนกประเภทผู้สมัครบัตรเครดิต Yap Bee Wah และ Irma Rohaiza Ibrahim (Bee Wah, Y, 2010) โดยใช้แบบจำลองการทำนาย สามารถช่วยในกระบวนการประเมินการคัดเลือกลูกค้าที่จะสมัครบัตรเครดิต โดยใช้เทคนิคเหมืองข้อมูล (Data Mining) 3 วิธี ได้แก่ Logistic regression (LR) model, Classification and regression tree (CART) model และ Neural Network (NN) model ผู้สมัครบัตรเครดิตซึ่งข้อมูลที่น่ามาเป็นข้อมูลในอดีตและปัจจุบัน ตัวอย่างเช่น เพศ อายุ สถานภาพการสมรส เงินเดือน รายได้เสริม ประเภทที่อยู่อาศัย และระดับการศึกษา ผลที่ได้จากการเปรียบเทียบ Classification and regression tree (CART) model สามารถทำนายถูกต้องแม่นยำสูงกว่าวิธีอื่น

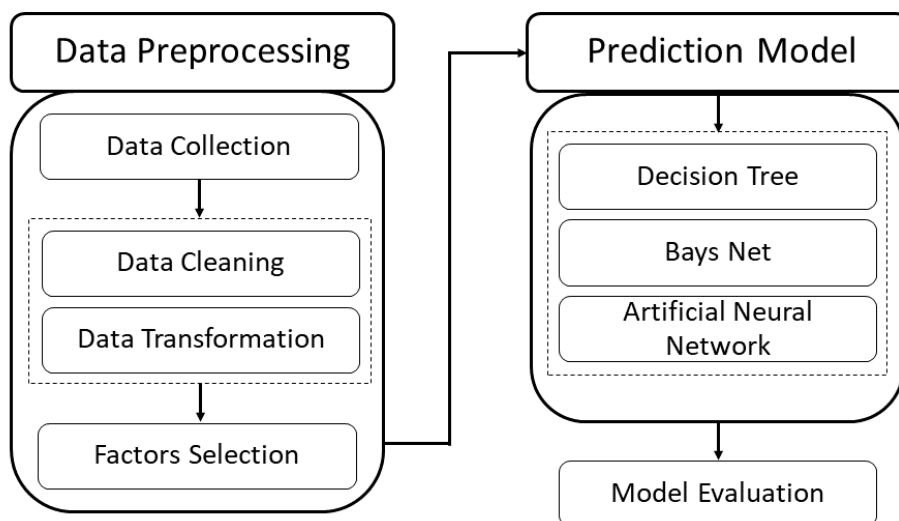
Wei Li และ Jibiao Liao (Wei Li, 2011) ได้มีการศึกษาข้อมูลที่จะทำให้เกิดผลกระทบปัญหาหนี้ให้น้อยที่สุด จึงได้สร้างโดยใช้รูปแบบเทคโนโลยีเหมืองข้อมูล จากเทคนิค C5.0 decision tree, neural network, chi-squared automatic interaction detector, stepwise logistic model และ classification and regression tree ซึ่งวิธีที่สามารถจำแนกประเภทผู้สมัครบัตรเครดิตที่ไม่ผิดนัดชำระ ที่ได้ผลลัพธ์ได้ดีที่สุดคือเทคนิค C5.0 decision tree

Wei Li และคณะ (Wei Li, 2010) ได้ศึกษาการจำแนกประเภทของลูกค้าที่มีบัตรเครดิตเป็น 4 รูปแบบ เพื่อเป็นข้อมูลพื้นฐานในการสร้างระบบทำนายความสามารถในการชำระหนี้ของลูกค้าบัตรเครดิตให้แก่ธนาคาร โดยใช้รูปแบบ C5.0, Neural network, Chi-squared automatic in traction detector และ Classification and regression tree ข้อมูลประกอบด้วย รายได้ส่วนตัวต่อเดือนและรายได้ครอบครัวต่อเดือน ค่าเฉลี่ยของจำนวนครั้งที่รู้ดบัตรต่อเดือนและสถานะเครดิตส่วนตัวมาใช้เป็นตัวแปร ผลที่ได้จากการจำแนกประเภทลูกค้าบัตรเครดิตโดยวิธี tree decision tree model C5.0 ความถูกต้องสูงกว่าวิธีอื่น

จากข้อมูลของธนาคารชูดาน Eiman kambal และคณะ (Eiman kambal, 2013) ได้พบว่าหนึ่งในปัจจัยความสำเร็จที่สำคัญขององค์กรสินเชื่อและธนาคารในการให้กู้ยืมโดยเฉพาะการประเมินความคุ้มค่าของผู้กู้ รูปแบบการจัดลำดับเครดิตได้ถูกนำมาใช้โดยนักวิจัยจำนวนมากที่จะปรับปรุงกระบวนการของการประเมินเครดิต โดยให้ผู้กู้ยืมเงินในอนาคตอยู่บนพื้นฐานของที่น่าจะเป็นของการชำระหนี้ การทดลองสร้างแบบจำลองการจัดลำดับเครดิตที่เหมาะสมสำหรับธนาคารชูดาน คือ Decision Tree (DT) และเครือข่ายประสาทเทียม (ANN)

### วิธีการดำเนินงาน

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อการชำระหนี้เงินกู้ของสมาชิก ช.พ.ค.และเพื่อสร้างโมเดลทำนายความสามารถในการชำระหนี้เงินกู้ โดยเทคนิคต้นไม้ตัดสินใจ เครือข่ายความเชื่อเบย์ และโครงข่ายประสาทเทียม แล้วนำผลการดำเนินงานมาเปรียบเทียบเพื่อหาโมเดลที่ให้ผลลัพธ์ที่มีค่าความถูกต้องแม่นยำสูงสุด ซึ่งมีขั้นตอนในการทำงานดังภาพที่ 2



ภาพที่ 2 วิธีการดำเนินการวิจัย

### การรวบรวมข้อมูล

การดำเนินงานวิจัยผู้วิจัยได้ทำการรวบรวมข้อมูลจากข้อมูลสมาชิกที่กู้เงินจากโครงการสวัสดิการการเงินกู้ทั้งหมด 7 โครงการ ตั้งแต่ปี พ.ศ. 2547 - 2557 ในสังกัดสำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา จังหวัดอุดรธานี รวมทั้งสิ้น 11,732 ราย มีผู้ที่สามารถชำระหนี้รายเดือนเป็นปกติจำนวน 7,699 ราย และผู้ที่ไม่สามารถชำระหนี้รายเดือนได้จำนวน 4,033 ราย

เพื่อการสร้างโมเดลในการทำนายความสามารถในการชำระหนี้เงินกู้ ในส่วนปัจจัยที่เกี่ยวข้องกับการกู้เงินมีคุณลักษณะทั้งหมด 14 คุณลักษณะ และผลการชำระ 1 คุณลักษณะ รวมเป็น 15 คุณลักษณะ ซึ่งปัจจัยที่คัดเลือกมานี้ได้จากการรวบรวมข้อมูลรายละเอียดการกู้เงินของสมาชิก ช.พ.ค. ตั้งแต่ปี พ.ศ. 2547 - 2557 ซึ่งสามารถแสดงรายละเอียดได้ดังตารางที่ 1

**ตารางที่ 1** ปัจจัยที่คาดว่าจะมีผลต่อความสามารถในการชำระหนี้เงินกู้และการชำระหนี้ของตัวแปร

| ลำดับ | ปัจจัย                              | ตัวอย่างข้อมูล                                 | การชำระหนี้   |
|-------|-------------------------------------|------------------------------------------------|---------------|
| 1     | สถานภาพการสมรส (Marital Status)     | โสด (1) สมรส (2) หย่า (3)                      | 1, 2, 3       |
| 2     | จำนวนบุตร (Children)                | 1 คน (1) 2 คน (2) 3 คน (3) 4 คน (4)            | 1, 2, 3, 4    |
| 3     | เพศ (Gender)                        | ชาย (1) หญิง (2)                               | 1, 2          |
| 4     | อายุ (Age)                          | 25, 26, 27,                                    | ข้อมูลจริง    |
| 5     | อาชีพ (Occupation)                  | ลูกจ้าง (1) พนักงานราชการ (2) รับราชการ (3)    | 1, 2, 3       |
| 6     | รายได้ต่อเดือน (Monthly salary)     | 15,000.00, 25,240.00, 53,080.00                | ปรับค่าข้อมูล |
| 7     | จำนวนเงินกู้ (Loan)                 | 600,000.00, 1,200,000.00, 3,000,000.00         | ปรับค่าข้อมูล |
| 8     | วิธีการชำระหนี้ (Payment method)    | ชำระด้วยตนเอง (1) หักเงินเดือน ณ ที่จ่าย (2)   | 1, 2          |
| 9     | ภาระหนี้ (Debt)                     | ไม่เคยกู้เงิน (1) มีภาระหนี้เงินกู้ (2)        | 1, 2,         |
| 10    | การทำประกัน (Insurance description) | ทำประกัน (1) ไม่ทำประกัน (2)                   | 1, 2          |
| 11    | จำนวนเงินงวด (Period)               | 3,800.00, 7,500.00, 23,200.00                  | ปรับค่าข้อมูล |
| 12    | ระดับการศึกษา (Degree)              | ป.ตรี (1) ป.โท (2) ป.เอก (3)                   | 1, 2, 3       |
| 13    | อาชีพคู่สมรส (Spouses)              | ไม่มีคู่สมรส(1) รับราชการ(2) ธุรกิจส่วนตัว (3) | 1, 2, 3       |
| 14    | วัตถุประสงค์การกู้ (Objective)      | เพื่อการศึกษา (1) ชำระหนี้ (2) สร้างบ้าน (3)   | 1, 2, 3...    |
| 15    | ค้างชำระเงินกู้รายเดือน (Result)    | ไม่ค้างชำระเงินกู้ (1) ค้างชำระเงินกู้ (2)     | 1, 2          |

### การทำข้อมูลให้สมบูรณ์

ข้อมูลทั้งหมด 16,026 มีค่าข้อมูลที่ยังไม่สมบูรณ์ ซึ่งหมายถึงค่าข้อมูลที่ขาดหายไป เนื่องจากการบันทึกข้อมูลในปัจจุบัน มีความคลาดเคลื่อนจากความเป็นจริง ซึ่งจะต้องทำให้ข้อมูลทั้งหมดมีความสมบูรณ์ก่อนนำไปสร้างตัวแบบ ซึ่งมีเงื่อนไขดังนี้

1.2.1 ตัดข้อมูลสมาชิก ช.พ.ค. ที่กู้เงินนั้น ๆ ทั้ง ในกรณีที่เอกสารคำขอกู้ไม่ครบสมบูรณ์

ตัดข้อมูลสมาชิก ช.พ.ค. ที่กู้เงินนั้น ๆ ทั้ง ในกรณีที่ผู้กู้ย้ายออกจากจังหวัดอุดรธานี 1.2.2

1.2.3 การตัดข้อมูลบางคอลัมน์ที่ไม่น่าสนใจออก เช่น ข้อมูลผู้รับผลประโยชน์ของผู้กู้ ข้อมูลเลขสมาชิก ช.พ.ค. ข้อมูลจำนวนเงินสมทบกองทุน เป็นต้น

1.2.4 เพิ่มข้อมูลบางรายการที่ผู้ยื่นกู้กรอกข้อมูลไม่สมบูรณ์ เช่น เพศ อายุ เป็นต้น

1.2.5 การตรวจสอบการพิมพ์ผิดให้ถูกต้องอย่างละเอียด และรอบคอบ ซึ่งข้อมูลที่พิมพ์ผิดนี้เกิดจากการบันทึกข้อมูลที่ผิดพลาดของเจ้าหน้าที่ผู้ตรวจสอบเอกสารการยื่นกู้

### การปรับบรรทัดฐานข้อมูล

เนื่องด้วยข้อมูลมีความแตกต่างกันในเรื่องของช่วงของข้อมูล ซึ่งประกอบด้วยค่าที่มี 1 หลัก (หลักหน่วย) และค่าที่มี 7 หลัก (หลักล้าน) ดังตัวอย่างในตารางที่ 2 เพื่อไม่ให้เกิดความแตกต่างก่อนการนำข้อมูลจำเป็นต้องมีการปรับช่วงของข้อมูลให้อยู่ในบรรทัดฐานเดียวกัน โดยใช้หลักการ Max-Min Normalization (เอกสิทธิ์ พัทธวงค์ศักดิ์ดา, 2557) จะทำให้ได้ข้อมูลอยู่ในช่วงระหว่าง 0 ถึง 1 ซึ่งสามารถปรับช่วงค่าข้อมูลได้ดังสมการที่ (6)

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A \quad (6)$$

ข้อมูลคอลลัมน์รายได้ (income) มีค่าต่ำสุด 15,000 บาท (min) และมีค่ามากที่สุด 62,820 บาท (max) ซึ่งข้อมูลที่ต้องการแปลงคือ 49,420 บาท ต้องการแปลงให้ข้อมูลนี้อยู่ในช่วงใหม่ คือ [0,1] เพราะฉะนั้น 49,420 บาท จะมีค่าใหม่ ซึ่งสามารถปรับช่วงค่าข้อมูลรายได้ต่อเดือนตัวอย่างแสดงรายละเอียดได้ดังตารางที่ 2

ตารางที่ 2 แสดงตัวอย่างปัจจัยข้อมูล

| รายได้ต่อเดือน (Monthly salary) | จำนวนเงินกู้ (Loan) | จำนวนเงินงวด (Period) |
|---------------------------------|---------------------|-----------------------|
| 49,420.00                       | 2,400,000.00        | 18,400.00             |
| 62,820.00                       | 3,000,000.00        | 24,400.00             |
| 15,000.00                       | 600,000.00          | 3,800.00              |

จากตารางที่ 2 สามารถปรับค่าบรรทัดฐานของรายได้ต่อเดือนในแถวที่ 1 ได้ดังนี้

$$\text{รายได้ต่อเดือน}' = \frac{49,420 - 15,000}{62,820 - 15,000} (1 - 0) + 0 = 0.7197$$

ดังนั้นจะได้ค่า 'รายได้ต่อเดือน' ใหม่ที่ถูกปรับช่วงค่าเท่ากับ 0.7197 และเมื่อมีการปรับบรรทัดฐานทั้งหมด จะทำให้ได้ผลลัพธ์ดังตารางที่ 3

ตารางที่ 3 แสดงตัวอย่างปัจจัยข้อมูลปรับบรรทัดฐานช่วงค่า

| รายได้ต่อเดือน (Monthly salary) | จำนวนเงินกู้ (Loan) | จำนวนเงินงวด (Period) |
|---------------------------------|---------------------|-----------------------|
| 0.7197                          | 0.7500              | 0.07087               |
| 1.0000                          | 1.0000              | 1.0000                |
| 0.0000                          | 0.0000              | 0.0000                |



### ารคัดเลือกปัจจัย

จากปัจจัยทั้ง 14 ปัจจัยที่ผู้วิจัยได้คัดเลือกมาเพื่อสร้างโมเดล ผู้วิจัยทำการวิเคราะห์เพื่อหาความสัมพันธ์ของข้อมูลปัจจัยกับความสามารถในการชำระหนี้เงินกู้โดยใช้การวิเคราะห์การถดถอยพหุคูณของตัวแปร ซึ่งผลการวิเคราะห์ปรากฏในตารางที่ 4

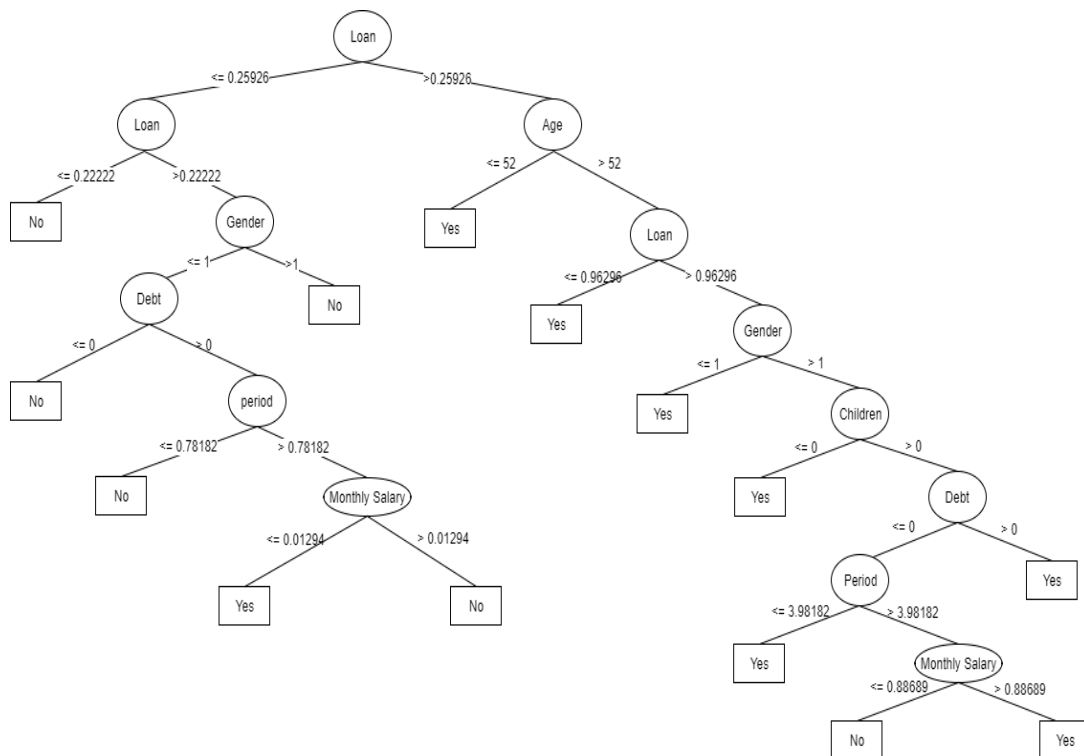
**ตารางที่ 4** แสดงผลการวิเคราะห์การถดถอยพหุคูณ

| ลำดับ | ปัจจัย                              | ระดับนัยสำคัญ |
|-------|-------------------------------------|---------------|
| 1     | จำนวนบุตร (Children)                | 0.000         |
| 2     | จำนวนเงินงวด (Period)               | 0.000         |
| 3     | อายุ (Age)                          | 0.007         |
| 4     | จำนวนเงินกู้ (Loan)                 | 0.000         |
| 5     | รายได้ต่อเดือน (Monthly salary)     | 0.000         |
| 6     | ภาระหนี้ (Debt)                     | 0.000         |
| 7     | การทำประกัน (Insurance description) | 0.000         |
| 8     | เพศ (Gender)                        | 0.008         |

จากตารางที่ 4 แสดงผลการวิเคราะห์ โดยการศึกษาในครั้งนี้กำหนดระดับนัยสำคัญที่ 0.01 และ 0.05 หมายถึงค่าความเชื่อมั่นน้อยกว่าหรือเท่ากับ 0.01 มีค่าความเชื่อมั่นที่ 99 % และ 0.05 มีค่าความเชื่อมั่นที่ 95 % ซึ่งพบว่ามี 8 ปัจจัยที่มีค่าอยู่ในระดับที่กำหนดซึ่งได้แก่ จำนวนบุตร เพศ อายุ รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ การทำประกันเงินกู้ จำนวนเงินงวด ซึ่งปัจจัยเหล่านี้คาดว่าจะมีผลต่อความสามารถในการชำระหนี้เงินกู้

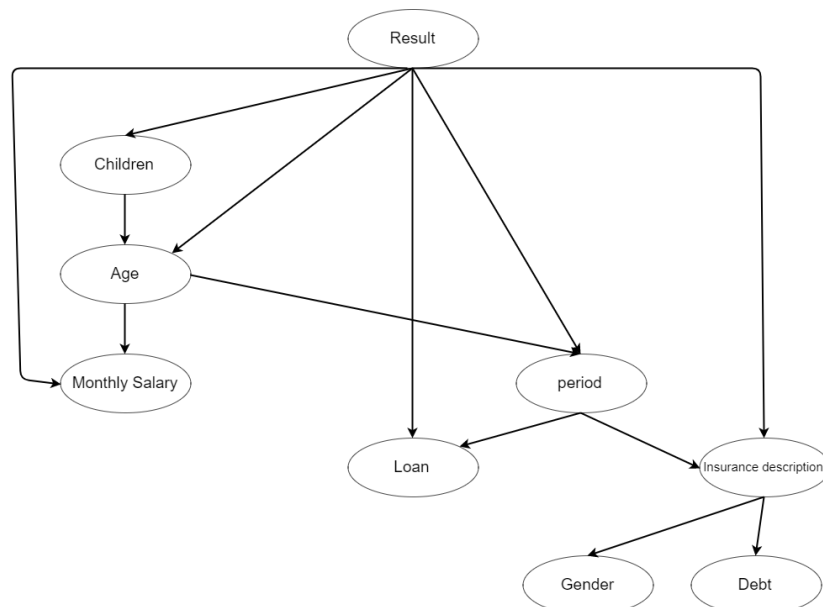
### การสร้างโมเดล

ผู้วิจัยได้ทำการสร้างโมเดลด้วยวิธี ต้นไม้ตัดสินใจ โมเดลเครือข่ายความเชื่อเบย์ และวิธีโครงข่ายประสาทเทียม โดยใช้ K-Fold Cross Validation ส่วนการสร้างโมเดลด้วยวิธีโครงข่ายประสาทเทียมนั้น มีการกำหนดค่าเรียนรู้ตั้งแต่ 0.1 - 0.3 ขึ้นแสดงผลเป็น 1 ค่าโมเมนตัมในชั้นซ่อนและชั้นแสดงผลตั้งแต่ 0.1 - 0.2 ฟังก์ชันในการกระตุ้นชั้นซ่อนและชั้นแสดงผล เป็นแบบ multilayer perceptron จำนวนรอบของการเรียนรู้ 500 รอบ จากปัจจัยทั้งหมด 8 ปัจจัยประกอบด้วย จำนวนบุตร เพศ อายุ รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ การทำประกันเงินกู้ จำนวนเงินงวด ทำการปรับค่าไปเรื่อย ๆ จนถึงค่าสูงสุดที่กำหนดไว้ ผลลัพธ์โมเดลต้นไม้ตัดสินใจ แสดงได้ดังภาพที่ 3 ผลลัพธ์โมเดลเครือข่ายความเชื่อเบย์ แสดงได้ดังภาพที่ 4 และค่าพารามิเตอร์ของวิธีโครงข่ายประสาทเทียม แสดงได้ดังตารางที่ 5



ภาพที่3 โมเดลต้นไม้ตัดสินใจ

จากภาพปัจจัย รายได้ต่อเดือน จำนวนเงินงวด จำนวนบุตร เพศ และอายุ มีผลต่อสัมพันธ์เนื่องกับจำนวนเงินกู้ในการชำระหนี้



ภาพที่4 โมเดลเครือข่ายความเชื่อเบย์

จากรูปปัจจัย เพศ อายุ จำนวนบุตร รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ จำนวนเงินงวด การทำประกันเงินกู้ มีความน่าจะเป็นที่จะเกิดการค้างชำระหนี้เงินกู้

**ตารางที่ 5** ค่าพารามิเตอร์ในการทดลองของวิธีโครงข่ายประสาทเทียม

| ลำดับ | ค่าพารามิเตอร์                          | ANN |
|-------|-----------------------------------------|-----|
| 1     | จำนวน Input                             | 8   |
| 2     | จำนวน Output                            | 2   |
| 3     | จำนวนชุดในการเรียนรู้และทดสอบ           | 10  |
| 4     | อัตราการเรียนรู้ในชั้นซ่อนและชั้นแสดงผล | 0.3 |
| 5     | ค่าโมเมนตัมในชั้นซ่อนและชั้นแสดงผล      | 0.2 |
| 6     | จำนวนรอบ                                | 500 |

#### ผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อการชำระหนี้เงินกู้ของสมาชิก ช.พ.ค. เพื่อสร้างโมเดลในการทำนายการชำระหนี้เงินกู้ของสมาชิก ช.พ.ค. จากการศึกษาครั้งนี้พบว่าปัจจัยที่มีผลต่อการชำระหนี้เงินกู้ของสมาชิก ช.พ.ค. มีทั้งหมด 8 ปัจจัยประกอบไปด้วย จำนวนบุตร เพศ อายุ รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ การทำประกันเงินกู้ จำนวนเงินงวด ข้อมูลดังกล่าวได้ถูกใช้ในการสร้างโมเดลเพื่อทำนาย โดยวิธีโครงข่ายประสาทเทียม ดันไม่ตัดสินใจ และโมเดลเครือข่ายความเชื่อเบย์ เมื่อเปรียบเทียบประสิทธิภาพผลการทำนายพบว่า การทำนายด้วยวิธีโครงข่ายประสาทเทียม มีค่าความแม่นยำ 98.77 % ที่สูงกว่าวิธีอื่น ตามด้วยวิธีเครือข่ายความเชื่อเบย์ มีค่าความแม่นยำ 98.76 % และวิธีดันไม่ตัดสินใจ ที่มีค่าความแม่นยำ 98.74 % ตามลำดับ ผลการเปรียบเทียบค่าความแม่นยำของทั้งสามวิธีได้แสดงดังตารางที่ 6

**ตารางที่ 6** เปรียบเทียบค่าความแม่นยำ (%) ของโมเดล

| Model                     | K-Fold | TP (%) | TN (%) | Accuracy (%) |
|---------------------------|--------|--------|--------|--------------|
| Artificial neural network | 10 k   | 98.7   | 1.3    | 98.77        |
| Bayesian Belief Network   | 10 k   | 98.7   | 1.3    | 98.76        |
| Decision tree             | 10 k   | 98.7   | 1.3    | 98.74        |

#### อภิปรายและสรุปผลการวิจัย

การสร้างโมเดลในการทำนายความสามารถในการชำระหนี้เงินกู้ มีการคัดเลือกปัจจัยทั้งหมด 8 ปัจจัยประกอบด้วย จำนวนบุตร เพศ อายุ รายได้ต่อเดือน จำนวนเงินกู้ ภาระหนี้ การทำประกันเงินกู้ และจำนวนเงินงวด ซึ่งทั้ง 8 ปัจจัยนี้มีผลต่อการชำระหนี้เงินกู้ ผลการทำนายจากโมเดลด้วยวิธี ดันไม่ตัดสินใจ เครือข่ายความเชื่อเบย์ และวิธีโครงข่ายประสาทเทียม พบว่าจำนวนเงินกู้ที่มาก หากผู้กู้มีอายุน้อย และมีรายได้ต่อเดือนน้อย จะทำให้ผู้กู้ไม่สามารถชำระหนี้เงินกู้ได้ ทำให้เกิดหนี้ค้างชำระ จากการเรียนรู้และการทดสอบด้วยโมเดลวิธีโครงข่ายประสาทเทียม มีค่าความผิดพลาดน้อยที่สุดและมีประสิทธิภาพในการทำนายแม่นยำสูงสุด

## เอกสารอ้างอิง

- กิตติ ภักดีวัฒนะกุล.คัมภีร์ระบบสนับสนุนการตัดสินใจ และระบบผู้เชี่ยวชาญ. กรุงเทพฯ: เค ทีพีคอมพ์ แอนด์ คอนซัลท์.2551
- กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และธนาวิทย์ รักธรรมานนท์. การใช้เทคนิคดาต้าไมน์นิ่งเพื่อพัฒนาคุณภาพ การศึกษาคณะวิศวกรรมศาสตร์คณะวิศวกรรมศาสตร์. ว. NECTEC Technical 2557 ;10, 31-33.
- พยุชน พานิชย์กุล.การพัฒนาแบบดาต้าไมน์นิ่งโดยใช้ Decision Tree. โครงการพัฒนาระบบงาน ปริญญา (วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ แขนงวิชาวิทยาการ สารสนเทศ) สถาบัน เทคโนโลยีพระจอมเกล้าคุณทหารลาดกระบัง, 2548.
- มหาวิทยาลัยเชียงใหม่. (2559). เหมืองข้อมูล. [อ้างเมื่อ 2 มิถุนายน 2559] จาก [http:// archive.lib.cmu.ac.th/ful/T/2556/comp40856ps\\_ch2.pdf](http://archive.lib.cmu.ac.th/ful/T/2556/comp40856ps_ch2.pdf).
- ยุทธ ไถยวรรณ.หลักการและการใช้การวิเคราะห์การถดถอยโลจิสติกสำหรับการวิจัย.วารสารวิจัยมหาวิทยาลัย เทคโนโลยีราชมงคลศรีวิชัย 2555;4(1), 1-12;
- วรรัตน์ สงฆ์เป็น. เอกสารประกอบการสอนวิชา 322756 Data Mining. ขอนแก่น: ภาควิชาวิทยาการ คอมพิวเตอร์ มหาวิทยาลัยขอนแก่น.2558
- สิรภัทร เชี่ยวชาญวัฒนา.เอกสารประกอบการสอนวิชา 322752 ข่ายงานประสาทเทียม (Artificial Neural Networks). ขอนแก่น: ภาควิชาวิทยาการคอมพิวเตอร์มหาวิทยาลัยขอนแก่น.2552.
- สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา. ประวัติสำนักงานคณะกรรมการ สกสศ.2559 [อ้างเมื่อ 7 มีนาคม 2559], จาก <http://www.otep.go.th/?p=otep-page&id=1>.
- สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและบุคลากรทางการศึกษา.คู่มือการปฏิบัติงานการกู้เงินซื้อ โครงการสวัสดิการเงินกู้ ข.พ.ค. ปี 2554. อุดรธานี: สำนักงานส่งเสริมสวัสดิการและสวัสดิภาพครูและ บุคลากรทางการศึกษา จังหวัดอุดรธานี.2554.
- เอกสิทธิ์ พัทธวงศ์ศักดิ์. การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้าไมน์เบื้องต้น. พิมพ์ครั้งที่ 2 . กรุงเทพฯ: เอเชีย ดิจิตอลการพิมพ์.2557
- Bee Wah, Y., Ibrahim, I.R. Using data mining predictive models to classify credit card applicants. 2010 6th International Conference on Advanced Information Management and Service (IMS); 2010, 30 November - 02 December; Seoul, Korea; 2010. pp.394-398.
- Eiman kambal, Izzeldin Oaman, Methag Taha, Noon Mohammed, Sara Mohammed. Credit Scoring Using Data Mining Techniques with Particular Reference to Sudanese Banks. 2013 International Conference on Computing, Electrical and Electronic Engineering (ICCEEE); 2013, 26-28 August; Khartoum, sudan; 2013. pp.378-383;
- GangMa, BinLi FangfangYang .Application of Bayesian Network in Improving Customer Credit Precision. 2010 International Conference on Management and Service Science (MASS); 2010 August 24-26; Wuhan: IEEE; 2010. pp. 1-3;2010
- H. Abdou and J. Pointon.Credit Scoring , Statistical Techniques and Evaluation Criteria. A Review of the Literature, Intelligent Systems in Accounting, Finance & Management 2011; 18(2-3): pp. 59-88.

- Mehran, S. Learning Limited Dependence Bayesian Classifiers. KDD-96 Proceedings, 2–4 August (pp. 81–106).  
Portland: Oregon.n
- Michell, Tom M. Machine Learning. New Jersey : McGraw Hill. 1997.
- Quinlan, J. R. C4.5: program for machine learning, California : Morgan Kaufmann. 1993.
- Rodgers, J. L.; Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. The American Statistician 1988 ;42 (1): 59–66.
- Surjeet, K., & Yadav, S.P. (2012) Data Mining A Prediction for Performance Improvement of Engineering Students using Classification. World of Computer Science and Information Technology Journal, 5-9 January (pp. 25-28). Malaysia: IIUM.
- U. Fayyad, G. Piatetsky-Shapiro & P. Smyth. (1996). Data Mining to Knowledge Discovery in Database. American Associate for Artificial Intelligence. American Association for Artificial Intelligence. 1996. (pp. 37-54). USA: California
- Wei Li, Xuemei Wu, Yayun Sun, Qianju Zhang. (2010). Credit Card Customer Segmentation and Target Marketing Based on Data Mining. 2010 International Conference on Computational Intelligence and Security (CIS); 2010, 11-14 December; Nanning, pp. 378-383.
- Wei Li, Jibiao Liao. (2011). An Empirical Study on Credit Scoring Model for Credit Card by Using Data Mining Technology. 2011 Seventh International Conference on Computational Intelligence and Security (CIS); 2011, 3-4 December; Hainan, pp. 1279-1282.