บทที่ 1

บทนำ

1.1 หลักการและเหตุผล

การฟื้นตัวในระยะข้างหน้าจะมีความแตกต่างกันในแต่ละกลุ่มธุรกิจและครัวเรือนที่สำคัญ หากการระบาดของโควิด-19 กินระยะเวลายาวนานขึ้น ความแตกต่างของการฟื้นตัว จะยิ่งทวีความรุนแรงมากขึ้น ซึ่งในท้ายที่สุดจะไปซ้ำเติมความเปราะบางของโครงสร้างเศรษฐกิจไทยที่เป็นต้นทุนอยู่เดิม (Rachot Leingchan, 2564: ออนไลน์) ซึ่งมหาวิทยาลัยราชภัฏสกลนครมีนโยบายและพันธกิจในการพัฒนาท้องถิ่นอย่างยั่งยืนบนพื้นฐานเศรษฐกิจพอเพียงเพื่อพัฒนาท้องถิ่น โดยมีโครงการน้อมนำศาสตร์พระราชาสู่การพัฒนาท้องถิ่น ยกระดับรายได้ซึ่งได้ดำเนินการมาตั้งแต่ปี 2561 จนถึงปัจจุบัน (มรสน, 2560: ออนไลน์) ในการพัฒนาเศรษฐกิจของคนในชุมชนซึ่งมหาวิทยาลัยราชภัฏสกลนคร ได้รับผิดชอบ 19 ตำบลในจังหวัดสกลนคร โดยมีการเก็บข้อมูลเศรษฐกิจไว้ในเป็นฐานข้อมูล ส่วนกลางของมหาวิทยาลัยราชภัฏกลนคร ซึ่งปัจจุบันยังไม่มีการวิเคราะห์ปัจจัยที่ส่งผลต่อเศรษฐกิจครัวเรือนที่จะสามารถสนับสนุนการตัดสินใจหรือการวางแผนการพัฒนาชุมชน

เทคโนโลยีในปัจจุบันมีหลากหลายเทคโนโลยีและมีหลากหลายศาสตร์ที่จะสามารถนำมาวิเคราะห์ข้อมูลขนาดใหญ่ได้โดยเฉพาะอย่างยิ่งหลักการทำเหมืองข้อมูลซึ่งเป็นหลักการในการวิเคราะห์เป็นหลักการของสถิติขั้นสูงที่จะสามารถวิเคราะห์ข้อมูลขนาดใหญ่ได้ โดยมีนักวิจัยหลากหลายศาสตร์ที่ได้นำเทคนิคการทำเหมืองข้อมูลนำมาประยุกต์ใช้งาน เช่น งานวิจัยของ นิภาพร ชนะมาร และพรรณี สิทธิเดช (2557) ได้ศึกษาการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนิสิตที่ศึกษาและสำเร็จการศึกษาแล้วในหลักสูตรเดียวกัน จำนวน 180 ระเบียน โดยใช้เทคนิคการคัดเลือกคุณสมบัติที่สำคัญ แล้วสร้างตัวแบบการพยากรณ์ด้วยเทคนิค BPNN และเทคนิค SVMs เมื่อทดลองใช้เทคนิคการรวมกลุ่ม ด้วยวิธี Bagging ร่วมกับ BPNN และ SVMs พบว่าผลการพยากรณ์ของ Baggingร่วมกับ BPNN (Bagging BPNN) มีค่าความผิดพลาดอยู่ใน ระดับต่ำสุด (RMSE=0.1051) งานวิจัยของวรายุทธ พลาศรี (2556) ได้ศึกษาการศึกษาปัจจัยที่มีผลต่อความยากจนของครัวเรือนในชนบท กรณีศึกษาจังหวัดมหาสารคาม การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาเศรษฐกิจของครัวเรือนในชนบท สถานการณ์ความยากจน ลักษณะของครัวเรือนที่ยากจน และปัจจัยที่มีผลต่อความยากจนของครัวเรือนในชนบทจังหวัดมหาสารคาม โดยกลุ่มประชากรที่ใช้ในการศึกษาคือ ครัวเรือนที่อยู่ ในเขตพื้นที่ชนบทจังหวัดมหาสารคาม จำนวน 180,328 ครัวเรือน ขนาดกลุ่มตัวอย่างเท่ากับ 400 ครัวเรือน และงานวิจัยของ อัจจิมา มณฑาพันธุ์ (2562) งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเกี่ยวกับการเปรียบเทียบวิธีการคัดเลือกคุณลักษณะที่สำคัญเพื่อนำมาใช้ในการ ปรับปรุงการพยากรณ์การเป็นมะเร็งเต้านม โดยใช้วิธีการคัดเลือก คุณลักษณะจากเทคนิคต่าง ๆ จำนวน 7 เทคนิค ได้แก่ เทคนิค Correlation Based Feature Selection เทคนิค Information Gain เทคนิค Gain Ratio เทคนิค Chi-Square เทคนิค Forward Selection เทคนิค Backward Elimination และเทคนิค Evolutionary Selection หลังจากคัดเลือกคุณลักษณะ ที่สำคัญจึงนำผลที่ได้จากแต่ละเทคนิคมาคำนวณหาค่าประสิทธิภาพในการพยากรณ์การเป็นมะเร็งเต้านมโดยการใช้ทคนิคซัพพอร์ต เวกเตอร์แมชชีน ผลการทดลองพบว่าร้อยละของความถูกต้องในการพยากรณ์การเป็นมะเร็งเต้านม จากจำนวนคุณลักษณะของข้อมูลทั้งหมด 30 คุณลักษณะเท่ากับ 91.39%

จากที่กล่าวมาข้างต้นผู้วิจัยจึงมีความสนใจที่จะศึกษาวิเคราะห์ปัจจัยสำคัญของเศรษฐกิจครัวเรือนด้วยเทคนิค Gain Ratio Feature Selection โดยใช้โปรแกรม RapidMiner เพื่อให้ได้ปัจจัยที่สำคัญสำหรับการวางแผน และการพัฒนาด้านเศรษฐกิจของชุมชนท้องถิ่นเพื่อให้สามารถมีคุณภาพชีวิตที่ดีขึ้น

1.2 วัตถุประสงค์ของโครงงาน

เพื่อวิเคราะห์ปัจจัยสำคัญของข้อมูลเศรษฐกิจครัวเรือนด้วยการคัดเลือกคุณสมบัติแบบ Gain Ratio Feature Selection

1.3 ขอบเขตของโครงงาน

1.3.1 ด้านข้อมูล

ข้อมูลที่ใช้ในการศึกษาครั้งนี้ คือข้อมูลประชากรจากภาคครัวเรือนเฉพาะครัวเรือน ของจังหวัดสกลนคร ซึ่งมี 12 หมู่บ้าน โดยช่วงเวลาที่ทำการเก็บรวบรวมข้อมูล คือ ปี พ.ศ. 2563 – 2564 และจากฐานข้อมูลเศรษฐกิจครัวเรือน (สำนักวิทยบริการและเทคโนโลยีสารสนเทศ, 2563: ออนไลน์)

1.3.1.1 ประชากรที่ใช้ในการศึกษา ได้แก่ ครัวเรือนใน 12 หมู่บ้าน จำนวน 2,909 ครัวเรือน มีทั้งหมด 17,933 ระเบียน

1.3.1.2 กลุ่มตัวอย่างที่ใช้ในการศึกษาและวิเคราะห์ข้อมูล ได้แก่ กลุ่มครัวเรือน ที่ได้มาจากการเลือกแบบเจาะจง (Purposive Sampling) จำนวน 1,751 ระเบียน (ครัวเรือน)

ข้อมูลเศรษฐกิจครัวเรือนถูกจัดเก็บออกเป็น 10 ส่วน ดังนี้

ส่วนที่ 1 ข้อมูลทั่วไปครัวเรือน

ส่วนที่ 2 ทรัพย์สินของครัวเรือน

ส่วนที่ 3 อาชีพและรายได้ของครัวเรือน

ส่วนที่ 4 รายจ่ายของครัวเรือน

ส่วนที่ 5 หนี้สินของครัวเรือน

ส่วนที่ 6 ผลกระทบจากสถานการณ์การระบาดของโรคติดเชื้อไวรัสโคโรน่า 2019 (COVID - 19)

ส่วนที่ 7 การใช้เทคโนโลยีสารสนเทศ

ส่วนที่ 8 การเข้าร่วมการละเล่น การฟ้อน การรำ พิธีกรรมตามวิถีวัฒนธรรมชุมชน

ส่วนที่ 9 การเข้าร่วมโครงการที่ผ่านมาย้อนหลัง 3 ปี

ส่วนที่ 10 ข้อคิดเห็นและข้อเสนอแนะเพิ่มเติม

1.3.2 ด้านเทคนิค

การศึกษาครั้งนี้ได้ประยุกต์ใช้เทคนิคทางด้านการทำเหมืองข้อมูล สำหรับพัฒนาหาตัวแบบที่เหมาะสมในการวิเคราะห์ปัจจัยสำคัญต่อข้อมูลเศรษฐกิจครัวเรือน ดังนี้

1.3.2.1 เทคนิค Gain Ratio Feature Selection เป็นเทคนิคที่ใช้ในการวิเคราะห์ปัจจัยสำคัญ

1.3.2.2 เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เพื่อมาหาประสิทธิภาพเปรียบเทียบระหว่างปัจจัยเริ่มต้น และการวิเคราะห์ปัจจัยสำคัญจากเทคนิค Gain Ratio Feature Selection มาเปรียบเทียบประสิทธิภาพจากนั้นจะได้มาซึ่งปัจจัยที่เหมาะสมที่สุด

1.3.2.3 กระบวนการ CRIPS-DM (Cross Reference Industry Standard for Data Mining) เป็นกระบวนการที่ใช้ในการดำเนินงานวิจัย

1.3.3 ด้านเครื่องมือในการพัฒนา

**1.3.3.1 ซอฟต์แวร์**

การศึกษาครั้งนี้ได้ทำการทดลองดำเนินการผ่านโปรแกรม RapidMiner Studio เวอร์ชัน 9.10 เป็นโปรแกรมที่ออกแบบมาสำหรับการวิเคราะห์ข้อมูล (mypccrack, 2565 : ออนไลน์)

**1.3.3.2 ฮาร์ดแวร์**

1) เครื่องคอมพิวเตอร์ Notebook Asus

1.1 หน่วยประมวลผล ADM Ryzen 5 2500U with Radeon Vega Mobile Gfx 2.00 GHz

1.2 หน่วยความจำหลัก (RAM): 8.00 GB

1.3 ระบบปฏิบัติการ (OS): Windows 10 64-bit

1.4 แผนการดำเนินงาน

**ตารางที่ 1.1** ระยะเวลาการดำเนินงาน

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **กิจกรรม** |  |  |  | | **ระยะเวลาในการดำเนินงาน (พ.ศ. 2565)** | | | | | | |  |
|  | **ม.ค** | | | **ก.พ** | | **มี.ค** | **เม.ย** | **พ.ค** | **มิ.ย** | **ก.ค** | **ส.ค** | **ก.ย** |
| 1) กำหนดหัวข้อและนำเสนอหัวข้อ |  | | |  | |  |  |  |  |  |  |  |
| 2) ค้นหาปัญหา โอกาสและเป้าหมาย |  | | |  | |  |  |  |  |  |  |  |
| 3) ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง |  | | |  | |  |  |  |  |  |  |  |
| 4) เสนอเค้าโครงโครงงาน |  | | |  | |  |  |  |  |  |  |  |
| 5) ศึกษาและวิเคราะห์ข้อมูล |  | | |  | |  |  |  |  |  |  |  |
| 6) ทำความเข้าใจข้อมูลและเตรียมข้อมูล |  | | |  | |  |  |  |  |  |  |  |
| 7) ดำเนินการพัฒนาโมเดล |  | | |  | |  |  |  |  |  |  |  |
| 8) ประเมินประสิทธิภาพการพัฒนาโมเดล |  | | |  | |  |  |  |  |  |  |  |
| 9) จัดทำเอกสารประกอบโครงงาน |  | | |  | |  |  |  |  |  |  |  |
| 10) นำเสนอโครงงานจบ |  | | |  | |  |  |  |  |  |  |  |
| 11) รายงานด้วยเล่มสมบูรณ์ |  | | |  | |  |  |  |  |  |  |  |

1.5 ประโยชน์ที่คาดว่าจะได้รับ

ได้ปัจจัยที่สำคัญที่ส่งผลต่อเศรษฐกิจครัวเรือนเพื่อสนับสนุนหรือเป็นข้อมูลประกอบการตัดสินใจในการพัฒนาชุมชนท้องถิ่นสำหรับนักวิจัย