**เรื่อง :** ระบบทำนายการพ้นสภาพของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ด้วยเทคนคิการทำเหมืองข้อมูล

https://ph02.tci-thaijo.org/index.php/scibru/article/view/242082

**วัตถุประสงค์ :** 1) เพื่อสร้างแบบจำลองการทำนายการพ้นสภาพของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ด้วยเทคนิคการทำเหมืองข้อมูล

2) เพื่อพัฒนาระบบทำนายการพ้นสภาพของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยราชภฏั บุรีรัมย์ ด้วยเทคนิคการทำเหมืองข้อมูล

**ขั้นตอนการดำเนินงาน :** ใช้คลังข้อมูล จากฐานข้อมูลการศึกษา ที่จัดเก็บข้อมูลของนักศึกษาตั้งแต่ปีการศึกษา 2556 - 2559 โดยใช้ข้อมูล ตัวอย่างจากนักศึกษาคณะวิทยาศาสตร์ ด้วยมีอัตราการพ้นสภาพมากที่สุดในมหาวิทยาลัยราชภัฏ บุรีรัมย์ จากนั้นนำข้อมูลจากหลายตารางมารวมกันผ่านกระบวนการเตรียมข้อมูลทำให้ได้ข้อมูลจำนวน 3,604 ระเบียน โดยใช้ อัลกอริทึมสำหรับการจำแนกประเภทเป็น J48 และใช้การวัดประสิทธิภาพการจำแนกประเภทโดยวิธีการตรวจสอบไขว้ (K-fold cross-validation) และนำ Model ที่ได้ไปปรับใช้กับ Web application โดยเลือกใช้ ภาษา PHP HTML JavaScript และระบบจัดการฐานข้อมูล MySQL

**การดำเนินงาน**

1. **วางกรอบการศึกษา**
   1. **การเตรียมข้อมูลก่อนประมวลผล (Data pre-processing)**

ต้องตรวจสอบความถูกต้องในเรื่องของรูปแบบข้อมูล ค่าของข้อมูลก่อนการประมวลผล

* 1. **อัลกอริทึมสำหรับการจำแนกประเภท (Classifier algorithm)**

ทำการจำแนกประเภท ด้วยอัลกอริทึม J48 เพื่อให้ได้ผลลัพธ์ที่มีความยืดหยุ่นสามารถนำไปใช้โดยแปลงเป็นกฎได้

* 1. **การวัดประสิทธิภาพการจำแนกประเภท (Evaluation classifier)**

โดยใช้วิธีการ ตรวจสอบไขว้ (K-fold cross-validation) เพื่อประเมินผลแบบจำลองจากค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-measure)

* 1. **การสร้างแบบจำลอง (Academic DSS model)**

เมื่อทำการสร้างแบบจำลองทำให้ได้กฎ จากอัลกอริทมึการจำแนกประเภทเพื่อนำไปสู่การพัฒนาระบบทำนายการพ้นสภาพของนักศึกษา

* 1. **การพฒันาและประเมินผลระบบ (Development and evaluation system)**

โดยเลือกใช้ ภาษา PHP HTML JavaScript และระบบจัดการฐานข้อมูล MySQL ในรูปแบบเว็บแอปพลิเคชัน (Web application)

1. **การรวบรวมข้อมูลเพื่อการประมวลผล**

ข้อมูล จากฐานข้อมูลการศึกษา ที่จัดเก็บข้อมูลของนักศึกษาตั้งแต่ปีการศึกษา 2556 – 2559

จำนวน 3,604 ระเบียน

1. **การคัดเลือกและแปลงข้อมูล (Data selection and transformation)**

กำหนดสมมติฐานของปัจจัยที่มีผลต่อการพ้นสภาพของนักศึกษาคือ ปัจจัยที่ เกี่ยวข้องกับผู้เรียน คือ สถานภาพครอบครัว อาชีพ/รายได้ของบิดาและมารดา ปัจจัยการเรียนในระดับ มัธยมศึกษา คือ ผลการเรียนเฉลี่ย แผนการเรียนที่จบ ขนาดโรงเรียน และปัจจัยระหว่างเรียนใน มหาวิทยาลัย คือ สาขาวิชา เกรดเฉลี่ยแต่ละภาคเรียน จำนวน 6 ภาคเรียน สถานการณ์กู้ยืม รวม 16 ปัจจัย

1. **การวิเคราะห์ข้อมูล**
   1. **การสร้างและทดสอบแบบจำลอง**

ใช้โปรแกรม WEKA 3.7.9เลือกใช้ J48 คือ อัลกอลิธึม C4.5 เพื่อใช้ในการเรียนรู้

จากนั้นสร้างแบบจำลองต้นไม้เพื่อการวิเคราะห์ปัจจัย โดย ตัวแบบที่ได้จะอยู่ในรูปของกฎการจำแนกประเภท (Classification rule) จากการเรียนรู้ด้วยชุดการ เรียนรู้ (Training set) แล้วนำไปทดสอบด้วยชุดทดสอบ (Test set) โดยใช้วิธีการตรวจสอบไขว้ กำหนดค่า k เป็น 10, 100 และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งเป็นร้อยละ

* 1. **การวัดค่าประสิทธิภาพของตัวแบบการพยากรณ์**

นำตัวแบบที่ได้จากชุดการ เรียนรู้มาทดสอบด้วยชุดข้อมูลทดสอบโดยการวัดค่าที่ได้ คือ ค่า Accuracy ค่า Precision ค่า Recall และค่า F-measure

1. **การพัฒนาและการประเมินผลระบบ**

**ผลลัพท์ :** 1) การใช้งานข้อมูลที่เก็บไว้ในฐานข้อมูลเพื่อการศึกษาของมหาวิทยาลัยราชภัฏบุรีรมัย์มาทำการวิเคราะห์ที่มีตารางจำนวนมาก และข้อมูลมีหลากหลายรูปแบบ รวมถึงการมีค่าผิดพลาด (Missing value) เก็บไว้จำนวนมาก ดังนั้นในขั้นตอนการเตรียมข้อมูลจึงต้องวางแผนเป็นอย่างดี ซึ่งต้องใช้เวลานาน และต้องกระทำอย่างระมัดระวัง 2) ความเหมาะสมของแอททริบิวส์ สำหรับการจัดประเภทข้อมูล (Data classification) ค่าข้อมูลจะต้องมีค่าที่วนซ้ำ และไม่หลากหลาย 3) อัตราการพ้นสภาพของนกัศึกษาชั้นปีที่ 1 มีอตัราที่สูงกว่าช้ัน ปีอื่น ดังนั้นมหาวิทยาลัยจะต้องให้ความสำคัญสำหรับนักศึกษาใหม่ทั้งในเรื่องของผลการเรียนและ พฤติกรรมของผู้เรียนด้วย

**ปัญหา :** Attribute GPA เนื่องด้วยตัว GPA มีข้อมูลที่หลากหลายมากซึ่งไม่เหมาะกับอัลกอริทึมการจำแนกอย่างมาก

**วิธีการแก้ไข :** ปรับให้อยู่ในรูปที่จำแนกได้ง่ายโดยแบ่งข้อมูลออกเป็นช่วงๆ โดย Week = GPA< 1.6 Medium = GPA 1.6-1.99 Good = GPA 2.0-2.5 Best = GPA>2.5

**เรื่อง :** วิเคราะห์และพัฒนาระบบแนะนําหนังสือคอมพิวเตอร์ แบบออนไลน์โดยใช้เทคนิคการกรองแบบอิงเนื้อหา

http://www.rpu.ac.th/Library\_web/doc/RC\_RR/2554\_ComBus\_Walainush.pdf

**วัตถุประสงค์ :** 1) เพื่อพัฒนาระบบแนะนําหนังสือคอมพิวเตอร์แบบออนไลน์โดยใช้เทคนิคการกรองแบบอิง เนื้อหามาใช้กับการแนะนําหนังสือให้ตรงกับความต้องการของผู้ใช้งานโดยอัตโนมัติได้

2) เพื่อประเมินความพึงพอใจของการพัฒนาระบบแนะนําหนังสือคอมพิวเตอร์แบบออนไลน์ที่พัฒนาขึ้น

**ขั้นตอนการดำเนินงาน และ เทคนิคหรือ theory ที่ใช้ :**

**ระบบให้การแนะนํา**

ระบบให้การแนะนํา.(Recommender System) เป็นเทคโนโลยี สมัยใหม่ที่ถูกนํามาใช้ในการแนะนําข้อมูลต่างๆ ที่คาดว่าผู้ใช้น่าจะสนใจ หรืออาจจะเป็นข้อมูลที่ ผู้ใช้ต้องการ และเหมาะสําหรับการดําเนินธุรกิจแบบพาณิชย์อิเล็กทรอนิกส์ (e-Commerce) โดยทั่วไประบบให้การแนะนําประกอบด้วย 3 ส่วน ดังต่อไปนี้

1. ส่วนข้อมูลพื้นฐานที่จําเป็นต้องใช้ในการประมวลผล เช่น โปรไฟล์ของผู้ใช้แต่ละคน
2. ส่วนการป้อนข้อมูล เป็นข้อมูลที่ได้จากการป้อนข้อมูลเข้ามาของผู้ใช้ เช่น การให้ คะแนนเรตติ้ง ซึ่งมีอยู่ 2 แบบ ดังนี้
   1. แบบชัดเจน.(Explicit) เรตติ้งแบบชัดเจนจะแสดงอยู่ในรูปของจํานวนตัวเลข ตามระดับความนิยมตั้งแต่ 1 ถึง 5, 1 ถึง 10 หรือระดับอื่นๆ ขึ้นอยู่กับการใช้งาน
   2. แบบไม่ชัดเจน (Implicit) ส่วนเรตติ้งแบบไม่ชัดเจนได้มาจากพฤติกรรมการใช้ งานของผู้ใช้ต่างๆ เช่น ประวัติการซื้อสินค้าหรือประวัติการเข้ามาใช้งานของผู้ใช้ในอดีตที่ผ่านมา
3. ส่วนอัลกอริธึมเป็นส่วนสําคัญที่สุดที่ใช้ประมวลข้อมูลเพื่อให้การแนะนําชิ้นข้อมูล ออกมา

Content-Based Filtering

เป็นวิธีที่นิยมใช้ในการค้นคืนข้อมูลที่ให้ความสนใจกับคุณภาพของเนื้อหาข้อมูลเป็นหลัก โดยจะสนใจว่าลักษณะข้อมูลนั้นตรงตามโปรไฟล์ของผู้ใช้หรือไม่ ซึ่งถ้าใช่ก็จะนําเสนอข้อมูลนั้น ทันที.แต่ถ้าไม่ใช่ก็จะไม่สนใจแม้ว่าข้อมูลนั้น จะมีลักษณะใกล้เคียงกับข้อมูลที่ผู้ใช้ ต้องการก็ตาม

โดยนําเนื้อหาของข้อมูล เช่น คําสําคัญ (Keywords), วลี (Phrases) หรือคุณลักษณะ (Feature) มา สร้างเป็นโปรไฟล์ของผู้ใช้แต่ละคน เพื่อค้นหาข้อมูลที่ผู้ใช้คนนั้นสนใจซึ่งวิธีการของ Content Based Filtering จะไม่ประสบกับปัญหาการให้เรตติ้งต่อชิ้นข้อมูลที่ไม่ทั่วถึง และปัญหาชิ้นข้อมูลที่ ยังไม่ได้ให้เรตติ้ง

**ขั้นตอนการทํางานของเทคนิคการกรองแบบอิงเนื้อหา**

3 ขั้นตอน คือ

1. การสร้างเมตริกซ์ของผู้ใช้ชิ้นข้อมูลจริง

นำข้อมูลจากฐานข้อมูลที่มีอยู่มาทำเป็น เมตริกซ์ผู้ใช้ชิ้นข้อมูล ซึ่งลักษณะเป็นตารางที่มีแถวเป็นผู้ใช้ และมีคอลัมน์เป็นชิ้นข้อมูลซึ่งก็คือ ชื่อเรื่องของหนังสือ และข้อมูลในแต่ละช่องเป็นเรตติ้ง ที่ผู้ใช้คนนั้นให้กับหนังสือเรื่องนั้นๆ โดยเรตติ้งที่ให้มีค่าเป็นจํานวนเต็มตั้งแต่ 1-5 และเป็น Null ถ้าผู้ใช้ไม่เคยให้เรตติ้งกับหนังสือเรื่องนั้น

1. การทํานายเรตติ้งด้วยเทคนิคการกรองแบบอิงเนื้อหา

นําข้อมูลพื้นฐานที่มี สองค่าด้วยกัน ได้แก่ คะแนนชื่อเรื่องและเนื้อหาหนังสือ , คะแนนชื่อผู้แต่งหนังสือ โดยทําการแบ่งค่าเรตติ้ง จาก 1-5 ออกเป็น 5 คลาส คือ ค่าเรตติ้ง 1-5 โดยเรียงลําดับคะแนนจากระดับน้อยซึ่งก็คือ 1 ไปจนถึง ระดับสูงซึ่งก็คือ 5 นําข้อมูลเนื้อหาของหนังสือในฐานข้อมูล และเมตริกซ์ผู้ใช้ชิ้นข้อมูลของแต่ละประเภทมาให้ Classifier เรียนรู้ประวัติการให้เรต ติ้งของผู้ใช้แล้วสร้างเป็นโปรไฟล์ของผู้ใช้แต่ละคน ซึ่งเก็บข้อมูลความน่าจะเป็นที่เนื้อหาต่างๆ ของชิ้นข้อมูลจะจัดอยู่ในคลาส ต่าง ๆ จากนั้นนําโปรไฟล์ของผู้ใช้แต่ละคนไปทํานายเรตติ้งให้กับชิ้นข้อมูลที่ผู้ใช้นั้นๆ ที่ยังไม่มีการให้เรตติ้งโดยใช้ Naïve Bayes Theorem

Naïve Bayes Theorem

การเรียนรู้แบบ Naïve Bayes เป็นอัลกอริธึมที่ง่ายและเร็วในการคํานวณ และเป็นการเรียนรู้ที่ อาศัยหลักทางสถิติและความน่าจะเป็น โดยผลลัพธ์ที่ให้ค่าความน่าจะเป็นสูงสุดเป็น คําตอบของ

Diagram

Description automatically generated with medium confidence

Text

Description automatically generated

1. การสร้างเมตริกซ์ผู้ใช้ชิ้นข้อมูลเทียม

จากเวคเตอร์ผู้ใช้ชิ้นข้อมูลเทียมของผู้ใช้ทีละคนที่ได้มาด้วยวิธีการข้างต้น จากนั้นนําเวคเตอร์เทียม

ที่ได้จากผู้ใช้ทั้งหมดนํามารวมกันเป็นเมตริกซ์ผู้ใช้ชิ้นข้อมูลเทียมที่ไม่มีความเบาบาง หรือก็ได้ทำ

การทำนายค่าเหล่านั้นผ่าน อัลกอริทึม Naïve Bayes

**Web Application**

**เครื่องมือที่ใช้ในการพัฒนา**

* HyperText Markup Language (HTML)
* Personal Home Page Tool (PHP)
* ระบบฐานข้อมูลด้วยมายเอสคิวแอล(MySql)

**การดำเนินงาน**

1. การศึกษาข้อมูลและปัญหาของระบบงานเดิม
   1. การศึกษาปัญหาของระบบงานเดิม
   2. การศึกษาเครื่องมือที่ทําการพัฒนาเว็บไซต์
2. การออกแบบระบบ
   1. แผนภาพกระแสข้อมูลระดับสูง(Context Diagram)
   2. แผนภาพการไหลของข้อมูล ระดับที่ 0 (Data Flow Diagram Level 0) ของระบบ
   3. แผนภาพแสดงความสัมพันธ์ (Entity Relationship Model)
   4. วิเคราะห์โครงสร้างฐานข้อมูลระบบ
   5. หน้าจอเปลี่ยนแปลงรหัสผ่านในการเข้าสู่ระบบ
   6. หน้าจอการลงคะแนนความชอบต่อหนังสือของผู้ใช้
   7. หน้าจอจัดการข้อมูลหนังสือ
3. การพัฒนาระบบ
4. การทดสอบระบบ
5. การประเมินผลระบบ

**ผลลัพท์ :** 1) การแสดงรายละเอียดหนังสือ การค้นหาข้อมูลหนังสือ การ แนะนําหนังสือโดยอัตโนมัติ รายงานการให้คะแนนหนังสือของผู้ใช้แต่เรื่อง ยอมรับความพึงพอใจ การใช้งานอยู่ในระดับดี 2) การแนะนําหนังสือให้ตรงกับความ ต้องการของผู้ใช้โดยอัตโนมัติ ยอมรับความพึงพอใจการใช้งานอยู่ในระดับดี 3) ความเหมาะสมของปริมาณข้อมูลที่นําเสนอในแต่ละหน้าจอการแสดงผลข้อมูลมีความเป็นรูปแบบและเป็ นมาตรฐานเดียวกันยอมรับความพึงพอใจการใช้งานอยู่ในระดับดี 4) ในการเข้าใช้งานการเตือนเมื่อพบข้อผิดพลาดในกรณีที่ ผู้ใช้ไม่ป้อนข้อมูลตามที่กําหนดความเหมาะสมของระบบในการรักษาความปลอดภัยยอมรับความ พึงพอใจการใช้งานอยู่ในระดับดี

**ปัญหา :** ระบบได้กําหนดให้ผู้ใช้งานทั่วไป สามารถทําการโหวตให้คะแนนหนังสืออย่างน้อย 5เล่มขึ้นไปเพราะถ้าลงคะแนนให้กับหนังสือต่ำกว่า 5 เล่ม ผลของการแนะนําหนังสือจะไม่ตรงกับ 58 ความต้องการของผู้ใช้ ซึ่งในทางกลับกันถ้าโหวตหนังสือย่างน้อย 5 เล่มขึ้นไป ระบบจะแนะนํา หนังสือได้ตรงกับความต้องการของผู้ใช้

**ปัญหา :** ในการพัฒนาระบบระบบแนะนําหนังสือคอมพิวเตอร์แบบออนไลน์โดยใช้เทคนิคการ กรองแบบอิงเนื้อหา ซึ่งต้องใช้เวลาในการจัดการข้อมูลหนังสือทั้งหมด 448 เรื่อง โดยทําการจัด หมวดหมู่ให้กับหนังสือแต่ละเรื่อง ตลอดจนถึงการจัดการข้อมูลในส่วนของผู้แต่ง โดยทําการ จําแนกผู้แต่งออกเป็น 2 กลุ่ม คือผู้แต่งที่ได้รับความนิยม ผู้แต่งที่ไม่ได้รับความนิยม ให้กับหนังสือ ทั้งหมด ซึ่งการดําเนินงานดังกล่าวมานี้ทําให้เสียเวลาเป็นอย่างมาก เพื่อที่จะได้มาซึ่งข้อมูลที่ สมบูรณ์ในการนํามาใช้กับเทคนิคการกรองแบบอิงเนื้อหา

**ปัญหา :** เทคนิคการกรองแบบอิงเนื้อหาจะเหมาะสมกับข้อมูลหนังสือประมาณ 600 เรื่อง แต่ถ้ามี ข้อมูลหนังสือมากกว่า 1000 เรื่อง ขึ้นไปจะทําให้ประสิทธิภาพการทํางานของระบบช้าลง ซึ่งส่งผล ให้ผู้ใช้งานต้องรอผลการแนะนําเป็นเวลานาน กว่าจะได้ผลลัพธ์ออกมา

**เรื่อง :** การทำนายผลสัมฤทธิ์ทางการเรียนของนักศึกษา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ที่ผ่านการคัดเลือกบุคคลเข้าศึกษาในระบบกลางการรับนักศึกษา

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjbtYSJlev2AhVVRmwGHS-pApUQFnoECA4QAQ&url=https%3A%2F%2Fdigital.lib.kmutt.ac.th%2Fjournal%2Floadfile.php%3FA\_ID%3D471&usg=AOvVaw2tGpFmXE5sWn7NHEL5Z1eZ

**วัตถุประสงค์ :** 1) เพื่อพัฒนาและตรวจสอบโมเดลผลสัมฤทธิ์ทางการ เรียนของนักศึกษามหาวิทยาลัยเทคโนโลยีพระจอมเกล้า ธนบุรีที่ผ่านการคัดเลือกเข้าศึกษาในระบบกลางการรับ นักศึกษา ภาคเรียนที่ 1 ปีการศึกษา 2553

2) เพื่อวิเคราะห์ผลการทำ นายผลสัมฤทธิ์ทางการ เรียนเมื่อมีการเปลี่ยนแปลงน้ำ หนักความสำ คัญของ องค์ประกอบที่ใช้ในการคัดเลือกบุคคลเข้าศึกษาใน มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีในระบบกลาง การรับนักศึกษาที่ต่างไปจากเดิม

**ขั้นตอนการดำเนินงาน :**

1. ทำ หนังสือขออนุญาตใช้ฐานข้อมูลทุติยภูมิคะแนน สอบคัดเลือกและฐานข้อมูลผลการเรียนของนักศึกษา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ชั้นปีที่ 1
2. นำ ข้อมูลมาเชื่อมโยงกับข้อมูลผู้ที่เข้าศึกษา อยู่จริงด้วยรหัสเลขประจำ ตัวและชื่อของนักศึกษาและ ทำ การบันทึกข้อมูลจากแหล่งข้อมูลทุติยภูมิประกอบด้วย ชื่อ-สกุล รหัสสาขาวิชา รหัสวิชาที่เป็นองค์ประกอบของ สาขาวิชานั้นพร้อมคะแนนแต่ละรายวิชา และผลสัมฤทธิ์ ทางการเรียนเฉลี่ยของนักศึกษาชั้นปีที่ 1 เมื่อสิ้นสุดภาค การเรียนที่ 1 ปีการศึกษา 2553 จากนั้นทำการกำ หนด รูปแบบน้ำ หนักความสำ คัญขององค์ประกอบที่ใช้ในการ คัดเลือกบุคคลเข้าศึกษาในคณะต่างๆ จำ นวน 10 รูปแบบ

**Table

Description automatically generated**

**Table

Description automatically generated**

**หมายเหตุ :** รูป แบบที่ 2-9 เป็นรูปแบบที่ผู้วิจัยได้กำ หนดขึ้นใหม่โดย เปลี่ยนแปลงน้ำ หนักความสำคัญขององค์ประกอบที่ใช้ใน การสอบคัดเลือกที่ต่างไปจากเดิม ส่วนรูปแบบที่ 10 เป็น รูปแบบที่กำ หนดขึ้นตามแนวทางการแก้ไขปัญหาเชิงรุก จากการประชุมเสวนาในหัวข้อ “ผลการสอบแอดมิชชัน ต่อการเรียนฟิสิกส์ของนิสิตและนักศึกษา”

และเนื่องจากแต่ละ รายวิชาของ O\_NET ในรูปแบบที่ 1 มีค่าเท่ากัน ผู้วิจัยจึงได้กำ หนดแต่ละรายวิชาของ O\_NET ในรูปแบบที่ 2-10 ให้มีน้ำ หนักเท่ากันเช่นเดียวกับรูปแบบที่ 1 ด้วย โดยให้ น้ำ หนักความสำคัญของวิชา PAT 73 (ความถนัดทาง วิศวกรรมศาสตร์) มากกว่า PAT 72 (ความถนัดทางด้าน วิทยาศาสตร์) เนื่องจากนักศึกษาคณะวิศวกรรมศาสตร์ ควรมีความถนัดวิศวกรรมศาสตร์มากกว่าวิชาอื่นๆ

1. วิเคราะห์ข้อมูลเบื้องต้นโดยใช้สถิติบรรยายคำ นวณ หาค่าสถิติเบื้องต้น ได้แก่ค่าเฉลี่ยค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์การกระจาย ค่าความเบ้ และค่าความ โด่ง เพื่อศึกษาลักษณะการกระจายของตัวแปรแต่ละตัว
2. วิเคราะห์ความสัมพันธ์ของตัวแปรด้วยการหา ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันโดยใช้โปรแกรม SPSS for Window version 16
3. วิเคราะห์ความเหมาะสมของโมเดลผลสัมฤทธิ์ ทางการเรียนในมหาวิทยาลัย โดยการวิเคราะห์ความ สอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์โดยโปรแกรม LISREL Version 8.72 for window ใช้สถิติทดสอบความ กลมกลืน ได้แก่การทดสอบค่าไค-สแควร์, ค่า p, ค่า GFI, ค่า RMR, ค่า RMSEA
4. วิเคราะห์อำ นาจของการทำ นายผลสัมฤทธิ์ทางการ เรียนในมหาวิทยาลัยโดยใช้ค่าสัมประสิทธิ์การทำ นาย ซึ่ง ได้จากผลการวิเคราะห์โมเดล LISREL

**Table

Description automatically generatedผลลัพท์การวิเคราห์ตาราง :**

1. พบว่า คณะวิศวกรรมศาสตร์ตัวแปร GPAX\_M6 และ GAT มีค่าเฉลี่ยค่อนข้างสูง สังเกต ได้จากลักษณะการแจกแจงของข้อมูลที่มีลักษณะเบ้ซ้าย (ค่าความเบ้เป็นลบ) ส่วนตัวแปร O\_NET, PAT, และ GPAX\_T1 มีลักษณะการแจกแจงเบ้ขวา (ค่าความเบ้เป็น บวก) นั่นคือ นักศึกษาส่วนใหญ่มีคะแนนค่อนข้างน้อย จึงทำ ให้ค่าเฉลี่ยของตัวแปรค่อนข้างต่ำ ส่วนการกระจาย ของข้อมูล พบว่าข้อมูลของตัวแปร O\_NET, PAT, และ GPAX\_T1 มีการกระจายมาก สังเกตได้จากค่าความโด่ง น้อยกว่าโค้งปกติ (ความโด่งมีค่าเป็นลบ) ส่วนข้อมูลของ ตัวแปร GPAX\_M6 และ GAT มีการกระจายของข้อมูล น้อย เนื่องจากมีความโด่งมากกว่าโค้งปกติ(ความโด่งมีค่า เป็นบวก)
2. ผลการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร สังเกตได้ ที่ใช้ในการวิเคราะห์โมเดลผลสัมฤทธิ์ทางการ เรียนของคณะวิศวกรรมศาสตร์และคณะวิทยาศาสตร์ จำ นวน 10 รูปแบบ โดยใช้สหสัมพันธ์แบบเพียร์สัน พบ ว่าตัวแปรสังเกตได้ส่วนใหญ่มีความสัมพันธ์กันอย่างมี นัยสำคัญที่ระดับ 0.01 โดยตัวแปรที่มีความสัมพันธ์กันมาก ที่สุดอย่างมีนัยสำคัญ คือ ผลการเรียนของนักศึกษาในชั้น ปีที่ 1 (GPAX\_T1) กับคะแนนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย (GPAX\_M6) แสดงว่าหาก นักศึกษามีคะแนนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษา ตอนปลาย (GPAX\_M6) สูง ก็จะมีคะแนนผลการเรียน ในชั้นปีที่ 1 (GPAX\_T1) สูงขึ้นด้วย นอกจากนั้นยังพบ ว่าเมทริกซ์สหสัมพันธ์ระหว่างตัวแปรสังเกตได้ของทุกรูป แบบ สอดคล้องกับผลการวิเคราะห์ค่าสถิติ Bartlett's test of sphericity และค่าดัชนีไกเซอร์-ไมเยอร์- ออลคิน ซึ่งแสดงให้เห็นว่าตัวแปรสังเกตได้ของกลุ่ม ตัวอย่างมีความสัมพันธ์ภายในต่อกันและเหมาะสมที่จะ นำ ไปใช้ในการพัฒนาโมเดลการทำ นายผลสัมฤทธิ์ทางการ เรียนของนักศึกษามหาวิทยาลัยเทคโนโลยีพระจอมเกล้า ธนบุรีที่ผ่านการคัดเลือกบุคคลเข้าศึกษาในระบบกลางการ รับนักศึกษาพร้อมกับตรวจสอบความสอดคล้องของโมเดล ที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ต่อไป
3. การนำ เสนอผลการวิเคราะห์โมเดลภายใต้รูปแบบ น้ำ หนักความสำคัญขององค์ประกอบทั้ง 10 รูปแบบ จะ ทำการนำ เสนอค่าอิทธิพลรวมของตัวแปรในโมเดล และ ค่าสถิติแสดงการตรวจสอบความสอดคล้องของโมเดลกับ ข้อมูลเชิงประจักษ์ ได้แก่ ค่าไค-สแคว์ (Chi-Square) องศาความเป็นอิสระ (df) ค่าความน่าจะเป็น (P-Value) ค่า RMSEA ค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) และค่าดัชนีวัดระดับความสอดคล้องกลมกลืน (GFI) ผล การวิเคราะห์โมเดลที่สอดคล้องกับข้อมูลเชิงประจักษ์ต้อง มีค่า P มากกว่า 0.05 ค่าไค-สแคว์/df น้อยกว่า 2 ค่า RMSEA และ RMR มีค่าใกล้0 ค่า GFI มีค่าใกล้1

**Table

Description automatically generated**

1. เมื่อพิจารณาค่าอิทธิพลของ ตัวแปรในโมเดลผลสัมฤทธิ์ทางการเรียนของนักศึกษา คณะวิศวกรรมศาสตร์พบว่า รูปแบบน้ำ หนักความสำคัญ ขององค์ประกอบทั้ง 10 รูปแบบ ตัวแปรความสามารถ ทางวิชาการในระดับมหาวิทยาลัยได้รับอิทธิพลทางตรง จากตัวแปรความสามารถทางวิชาการระดับโรงเรียนมีค่า อิทธิพลเท่ากับ 0.74 เมื่อพิจารณาค่าสถิติที่ใช้ในการตรวจ สอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์พบ ว่า ทุกรูปแบบมีความสอดคล้องของโมเดลกับข้อมูลเชิง ประจักษ์โดยมีค่าไค-สแคว์(Chi-Square) ตั้งแต่ 3.31- 3.69 องศาความเป็นอิสระมีค่าเท่ากับ 4 ค่าความน่าจะ เป็นมีค่าตั้งแต่ 0.449-0.507 ค่า RMSEA มีค่าเท่ากับ 0.000 ค่าดัชนีรากกำลังสองเฉลี่ยของเศษ (RMR) มีค่า ตั้งแต่ 0.015-0.016 และดัชนีความกลมกลืน (GFI) มีค่า เท่ากับ 0.99 และตัวแปรในโมเดลทั้งหมด สามารถทำ นาย ตัวแปรความสามารถทางวิชาการในระดับมหาวิทยาลัยได้ โดยมีค่าสัมประสิทธิ์การทำ นายเท่ากันมีค่าเท่ากับร้อยละ 55 ทุกรูปแบบ

**Table

Description automatically generated**

1. เมื่อพิจารณาค่าอิทธิพลของ ตัวแปรในโมเดลผลสัมฤทธิ์ทางการเรียนของนักศึกษา คณะวิศวกรรมศาสตร์พบว่า รูปแบบน้ำ หนักความสำคัญ ขององค์ประกอบทั้ง 10 รูปแบบ ตัวแปรความสามารถ ทางวิชาการในระดับมหาวิทยาลัยได้รับอิทธิพลทางตรง จากตัวแปรความสามารถทางวิชาการระดับโรงเรียนมีค่า อิทธิพลตั้งแต่ 0.85-0.90 เมื่อพิจารณาค่าสถิติที่ใช้ในการ ตรวจสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ พบว่า ทุกรูปแบบมีความสอดคล้องของโมเดลกับข้อมูล เชิงประจักษ์ โดยมีค่าไค-สแคว์ (Chi-Square) ตั้งแต่ 0.52-1.78 องศาความเป็นอิสระมีค่าเท่ากับ 3 ค่าความ น่าจะเป็นมีค่าตั้งแต่ 0.618-0.915 ค่า RMSEA มี ค่าเท่ากับ 0.000 ค่าดัชนีรากกำ ลังสองเฉลี่ยของเศษ (RMR) มีค่าตั้งแต่ 0.019-0.029 และดัชนีความกลมกลืน (GFI) มีค่าตั้งแต่ 0.99-1 และตัวแปรในโมเดลทั้งหมด สามารถทำ นายตัวแปรความสามารถทางวิชาการในระดับ มหาวิทยาลัยได้ โดยค่าสัมประสิทธิ์การทำ นายมีค่าตั้งแต่ ร้อยละ 72-81 ซึ่งรูปแบบที่สามารถทำ นายตัวแปรความ สามารถทางวิชาการในระดับมหาวิทยาลัยได้ดีที่สุด คือ รูปแบบที่ 7 ส่วนรูปแบบที่ 10 เป็นรูปแบบที่ทำ นายได้ น้อยที่สุด

**ผลลัพท์ :**

1. พบว่า องค์ประกอบที่ใช้ทำ นายผล สัมฤทธิ์ทางการเรียน คือ ความสามารถทางวิชาการในระดับโรงเรียน (ACA\_SCH) ประกอบด้วยตัวแปรสังเกต ได้ คือ คะแนนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษา ตอนปลาย (GPAX\_M6) คะแนนสอบทางการศึกษาแห่ง ชาติขั้นพื้นฐาน (O\_NET) การสอบวัดความถนัดทั่วไป (GAT) และการสอบวัดความถนัดวิชาชีพ (PAT)
2. การวิเคราะห์ความสอดคล้องของโมเดลผลสัมฤทธิ์ ทางการเรียนของนักศึกษามหาวิทยาลัยเทคโนโลยี พระจอมเกล้าธนบุรีกับข้อมูลเชิงประจักษ์ พบว่า โมเดลมี ความสอดคล้องกับข้อมูลเชิงประจักษ
3. จากการศึกษาการเปลี่ยนแปลงน้ำ หนักความสำคัญ ขององค์ประกอบที่ใช้ในการคัดเลือกบุคคลเข้าศึกษาใน มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรีที่ต่างไปจาก เดิม พบว่า ทุกคณะและทุกรูปแบบสามารถทำ นายผล สัมฤทธิ์ทางการเรียนของนักศึกษามหาวิทยาลัยเทคโนโลยี พระจอมเกล้าธนบุรีได้ดังรายละเอียดต่อไปนี้

3.1) คณะวิศวกรรมศาสตร์พบว่าทุกรูปแบบสามารถ อธิบายความแปรปรวนของตัวแปรความสามารถทาง วิชาการในระดับมหาวิทยาลัยได้เท่ากันคือร้อยละ 55 โดย มีตัวแปรสังเกตได้ 4 ตัวแปร ที่สามารถอธิบายความ แปรปรวนของตัวแปรความสามารถทางวิชาการในระดับ มหาวิทยาลัยและมีนัยสำคัญทางสถิติได้แก่ คะแนนเฉลี่ย สะสมตลอดหลักสูตรมัธยมศึกษาตอนปลาย (GPAX\_M6) คะแนนสอบทางการศึกษาขั้นพื้นฐาน (O\_NET) คะแนน การสอบวัดความถนัดทั่วไป (GAT) และคะแนนการสอบ วัดความถนัดวิชาชีพ (PAT) โดยคะแนนสอบทางการ ศึกษาขั้นพื้นฐาน (O\_NET) มีค่าน้ำ หนักมากที่สุด รองลง มาคือ คะแนนการสอบวัดความถนัดวิชาชีพ (PAT)

1. พบว่า ผลการทำ นายผลสัมฤทธิ์ทางการเรียนของ คณะวิศวกรรมศาสตร์ไม่มีการเปลี่ยนแปลง ทั้งนี้อาจเป็น เพราะว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรของทุก รูปแบบมีค่าใกล้เคียงกันมาก เมื่อนำ มาวิเคราะห์ด้วยโปร แกรมลิสเรล จึงทำ ให้ผลการทำ นายเท่ากันทุกรูปแบบ จึง กล่าวได้ว่าคณะวิศวกรรมศาสตร์สามารถใช้รูปแบบใดใน การกำ หนดน้ำ หนักความสำคัญก็ได้
2. เมื่อพิจารณาผลการทำ นายผลสัมฤทธิ์ทางการเรียน ของคณะวิทยาศาสตร์พบว่า รูปแบบที่สามารถทำ นายผล สัมฤทธิ์ทางการเรียนได้ดี เป็นรูปแบบที่มีการกำ หนดน้ำ หนักความสำคัญของคะแนนสอบทางการศึกษาขั้นพื้นฐาน (O\_NET) และคะแนนสอบวัดความถนัดวิชาชีพ (PAT) ให้มีค่ามาก อาจเป็นเพราะตัวแปรทั้งสอง มีความสัมพันธ์ กับผลสัมฤทธิ์ทางการเรียนในชั้นปีที่ 1 ในระดับปานกลาง ถึงค่อนข้างสูง และคะแนนที่ได้จากการสอบทั้งสอง เป็น ผลสัมฤทธิ์โดยตรงจากตัวผู้สอบเอง ทำ ให้สามารถวัดผล ได้ตรงกับความรู้ความสามารถที่แท้จริงของผู้สอบ เช่น เดียวกับการวัดผลสัมฤทธิ์ทางการเรียนในชั้นปีที่ 1 ซึ่ง เป็นการสอบเช่นกัน จึงกล่าวได้ว่า คะแนนจากการสอบ เป็นตัวทำ นายผลสัมฤทธิ์ทางการเรียนในชั้นปีที่ 1 ที่ดีที่สุด สอดคล้องกับงานวิจัยของมณีรัตน์กรุงแสนเมือง
3. รูปแบบการกำ หนดน้ำ หนักความสำ คัญขององค์ ประกอบทั้ง 10 รูปแบบในแต่ละคณะ ในทางปฏิบัติอาจ เป็นไปไม่ได้ทั้ง 10 รูปแบบ เนื่องจากบางรูปแบบที่ผู้วิจัย กำ หนดขึ้นได้กำ หนดให้ความสำ คัญขององค์ประกอบที่ เป็นคะแนนเฉลี่ยสะสมตลอดหลักสูตรมัธยมศึกษาตอน ปลายมากกว่าองค์ประกอบจากการสอบ เช่น รูปแบบที่ 6 ของคณะวิศวกรรมศาสตร์ และรูปแบบที่ 5 ของคณะ วิทยาศาสตร์ ซึ่งจากการสอบคัดเลือกที่ผ่านมาจะกำ หนด ให้องค์ประกอบจากคะแนนสอบมีน้ำ หนักความสำ คัญ มากกว่าองค์ประกอบอื่นๆ ดังนั้นในการกำ หนดน้ำ หนัก ความสำ คัญต้องมีการศึกษาวิจัยและติดตามผลอย่างต่อ เนื่อง เพราะการทำ นายผลสัมฤทธิ์ทางการเรียนดังกล่าว ไม่สามารถบอกได้ว่ารูปแบบใดดีกว่ากัน ต้องมีการศึกษา วิจัยต่อไป

**เรื่อง : การพัฒนาระบบแนะนำแหล่งท่องเที่ยวของจังหวัดมหาสารคาม**

**โดยใช้วิธีการกรองข้อมูลแบบผสมผสาน**

<http://research.rmu.ac.th/rdi-mis//upload/fullreport/1607006848.pdf>

**วัตถุประสงค์ :**

1. เพื่อพัฒนาระบบแนะนำแหล่งท่องเที่ยวของจังหวัดมหาสารคาม โดยใช้วิธีการกรองข้อมูล แบบผสมผสาน (Hybrid Filtering)

2. เพื่อทดสอบประสิทธิภาพของระบบที่พัฒนาขึ้น

3. เพื่อประเมินความพึงพอใจของระบบที่พัฒนาขึ้น

**ขั้นตอนการดำเนินงาน :**

1. การดำเนินกำรวิจัยและเก็บรวบรวมข้อมูล
   1. ขั้นตอนการวิเคราะห์ความต้องการของระบบ
   2. เก็บข้อมูลและรายละเอียดต่าง ๆ ของแหล่งท่องเที่ยวใน จังหวัดมหาสารคาม
   3. ศึกษาเครื่องมือที่ทำการพัฒนาเว็บไซต์
2. การออกแบบระบบ
   1. เขียนอธิบายรูปแบบของผังงาน (Flowchart)
   2. ออกแบบฐานขอมูล เป็นลักษณะ ER-Diagram
   3. ออกแบบ ส่วนนำเข้า (Input Design)
   4. ออกแบบข้อมูลออก (Output Design)
   5. ออกแบบจอภาพ (Screen Design)
   6. ออกแบบส่วนการแนะนำ
      1. การสร้างเมตริกซ์ของชิ้นข้อมูลจริงจากข้อมูลในฐานข้อมูล
      2. ให้ผู้ใช้ทำาการให้คะแนน (Vote) ต่อชิ้นข้อมูลต่อรายการที่ผู้ใช้พึงพอใจ ต่อสถานที่นั้น โดยสร้างเป็นเมทริกซ์ผู้ใช้กับชิ้นข้อมูล
      3. หาค่าความคล้ายคลึงของชิ้นข้อมูลที่มีความสัมพันธ์กัน โดยเลือกใช้วิธี Naïve-Bayes
      4. การทำนาย (Prediction) หลังจากที่ได้ความคล้ายระหว่างคู่รายการใดๆ แล้วทำการพยากรณ์ค่า ความชอบ ของผู้ใช้ต่อรายการใดรายการหนึ่ง โดยวิธี Weighted-sumนำค่าที่ได้จากการทำนาย มาทำการแนะนำให้แก่ผู้ใช้งานเป้าหมาย
3. การพัฒนาระบบ
4. การทดสอบระบบ
5. การประเมินผลระบบ
6. การวิเคราะห์ข้อมูล

**ผลการวิจัย:**

1. ผลการพัฒนาระบบแนะนำแหล่งท่องเที่ยวของจังหวัดมหาสารคาม โดยใช้วิธีการกรอง

ข้อมูลแบบผสมผสาน ผู้วิจัยได้ทำการประยุกต์โดยการนำเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

และเทคนิคการกรองข้อมูลแบบอิงเนื้อหา มาทำงานร่วมกัน ซึ่งจะให้คำแนะนำได้ดีกว่าการใช้เทคนิค

เดียว โดยการแนะนำข้อมูลให้แก่ผู้ใช้บนพื้นฐานของรายละเอียดที่ถูกประเมินของรายการข้อมูลที่เคย

ได้รับการให้คะแนนความพึงพอใจ (vote) ที่ผ่านมา ทำให้ได้ระบบแนะนำที่สามารถให้คำแนะนำที่

ใกล้เคียงกับความต้องการของผู้ใช้มากที่สุด

1. ผลการทดสอบประสิทธิภาพของระบบที่พัฒนาขึ้น

จากการพัฒนาระบบแนะนำแหล่งท่องเที่ยวของจังหวัดมหาสารคาม โดยใช้วิธีการกรอง

ข้อมูลแบบผสมผสาน พบว่าผู้เชี่ยวชาญมีความคิดเห็นต่อประสิทธิภาพของระบบโดยรวมอยู่ในระดับดี

( x = 4.15, s.d. = 0.54) เมื่อพิจารณาในเป็นรายด้าน พบว่า ประสิทธิภาพของระบบอยู่ในระดับดีทั้ง

4 ด้าน โดยมีค่าเฉลี่ยอยู่ระหว่าง 3.94 ถึง 4.24

1. ผลการประเมินความพึงพอใจของระบบที่พัฒนาขึ้น

ผลการประเมินความพึงพอใจของผู้ใช้งานทั่วไป แสดงให้เห็นว่า ผู้ใช้งานทั่วไปมีความพึง

พอใจในการใช้งานระบบแนะนำแหล่งท่องเที่ยวของจังหวัดมหาสารคาม โดยใช้วิธีการกรองข้อมูล

แบบผสมผสาน โดยภาพรวมและรายด้านอยู่ในระดับมากทุกด้าน

**ข้อเสนอแนะในการนำผลการวิจัยไปใช้**

1. การติดตั้งระบบบนเว็บไซต์ เพื่อให้การทำงานเป็นไปด้วยความสะดวก ลดความเสียหาย

ต่อข้อมูล จำเป็นอย่างยิ่งต้องศึกษาขอมูลเกี่ยวกับระบบเซิฟเวอร์ (Server System) ที่ให้บริการว่า

การบริการนั้นครอบคลุมกับการทำงานของระบบที่ได้ถูกออกแบบไว้หรือไม่ ทั้งนี้เพื่อลดปัญหาการ

เปลี่ยนแปลงซอฟต์แวร์ที่นำมาใช้ในการออกแบบระบบแนะนำ

1. ต้องมีการปรับปรุงฐานข้อมูลที่ใช้ในการสร้างตัวแบบอยู่เสมอ เพื่อความถูกต้องแม่นยำใน

**เรื่อง: การเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์**

**ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพ**

[**http://www.journal.msu.ac.th/upload/articles/article2154\_27225.pdf**](http://www.journal.msu.ac.th/upload/articles/article2154_27225.pdf)

**วัตถุประสงค์:**

1. เพื่อศึกษาปัจจัยที่มีผลต่อการสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์
2. พัฒนาแบบจำลองที่มีประสิทธิภาพในการพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ โดยใช้ข้อมูลนักเรียนที่เรียนครบตามระยะเวลาของหลักสูตรประกาศนียบัตรวิชาชีพ (ปวช.) ในปีการศึกษา 2557-2558
3. เพื่อช่วยในการพยากรณ์การสำเร็จการศึกษาในสถานศึกษาอาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ซึ่งสามารถช่วยส่งเสริมให้ครูแนะแนวและผู้บริหารมีข้อมูลสารสนเทศเพื่อวางแผนในการแนะแนวให้นักเรียน-นักศึกษาเข้าศึกษาต่อในสถานศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์

**วิธีการดำเนินการวิจัย:**

1. การเตรียมข้อมูล

การเตรียมข้อมูลของการศึกษาค้นคว้าอิสระครั้งนี้ ได้ข้อมูลมาจากการสำรวจข้อมูลจำนวนนักเรียนจากสถาน ศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ ประจำปีการ ศึกษา 2557-2558 จำนวน 12 แห่ง ดัง

1. การทำกระบวนการก่อนการสร้างแบบจำลอง

ในขั้นตอนนี้ผู้วิจัยได้ทำการแปลงข้อมูลและ ทำการวิเคราะห์ปัจจัย

1. การแปลงข้อมูล (Data transformation) จากข้อมูลที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนจากสถาน ศึกษาอาชีวศึกษาเอกชนในจังหวัดกาฬสินธุ์ ประจำปีการ ศึกษา 2557-2558 จำนวน 12 แห่ง มีทั้งหมด 20 ตัวแปร ดัง Table 2

**Table

Description automatically generatedTable

Description automatically generated**

2. การวิเคราะห์ปัจจัยเป็นการคัดเลือกแอตทริบิวต์ ที่สามารถเป็นตัวแทนของกลุ่มแอตทริบิวต์เพื่อลดจำนวน แอตทริบิวต์ในการพยากรณ์ซึ่งงานวิจัยนี้ได้นำเทคนิค Gain Ratio Attribute Evaluation มาใช้ในการวิเคราะห์ปัจจัยผลการ วิเคราะห์ทำการตัดแอตทริบิวต์ที่มีผลต่อการพยากรณ์น้อย ออกจำนวน 3 แอตทริบิวต์คือ Age, Number of siblings และ Province คงเหลือ 17 แอตทริบิวต์ ดัง Table 3

**Table

Description automatically generated**

1. การสร้างแบบจำลอง

การสร้างแบบจำลองพยากรณ์การสำเร็จการ ศึกษาของผู้ที่สนใจเข้าศึกษาต่อสถานศึกษาอาชีวศึกษา เอกชนในจังหวัดกาฬสินธุ์ เพื่อช่วยในการตัดสินว่าเมื่อเข้า ศึกษาต่อแล้วจะสามารถสำเร็จการศึกษาได้หรือไม่ด้วยตัวแปร ทั้ง 21 ตัวแปรที่ได้จากการสำรวจข้อมูลจำนวนนักเรียนในปี การศึกษา 2557-2558 นั้น ด้วยการทำเหมืองข้อมูลและเทคนิค ที่นำมาใช้ในการสร้างแบบจำลองมีจำนวน 6 เทคนิคคือ

* เทคนิค C4.5
* เทคนิค Random Forest
* เทคนิค Random Tree
* เทคนิค REP Tree
* เทคนิค k-NN
* เทคนิค SVM

1. การวัดประสิทธิภาพของแบบจำลอง

ในการวัดประสิทธิภาพของแบบจำลองนั้น ได้มี การใช้เทคนิคแบบ 10-fold cross validation โดยจะทำการ แบ่งข้อมูลออกเป็น 10 ชุดเท่าๆกัน จากนั้นจะทำการทดสอบ ทั้งหมด 10 รอบ โดยในแต่ละรอบจะใช้ข้อมูล 1 ชุดเป็นชุด ทดสอบและอีก 9 ชุดที่เหลือเป็นชุดฝึกสอน ในรอบต่อไปก็ใช้ ชุดข้อมูลถัดไปเป็นชุดทดสอบจนครบทั้ง 10 ชุดข้อมูล ซึ่ง ข้อมูลทั้งหมด จำนวน 1266 ข้อมูล แบ่งออกเป็น 10 ชุดข้อมูล ซึ่งคิดเป็นอัตราข้อมูลทดสอบต่อข้อมูลฝึก เป็นอัตราส่วน 10:90 การวิเคราะห์ประสิทธิภาพ คือ การวัดประสิทธิภาพ การทำงานในแต่ละขั้นตอนวิธี สามารถวัดได้จากผลของการ จำแนกกลุ่มข้อมูล โดยค่าของผลลัพธ์ที่ได้จากการจำแนกคือ

ค่า True Positive (TP) ค่า True Negative (TN) ค่า False Positive (FP) ค่า False Negative (FN) และสามารถหาค่า

Text, letter

Description automatically generated

**ผลการศึกษา:**

ในการศึกษาประสิทธิภาพของแบบจำลองนี้คณะ ผู้วิจัยได้นำเอาโปรแกรมWEKA เวอร์ชัน 3.9.1 มาเป็นเครื่อง มือมาใช้ในการสร้างแบบลองด้วยเทคนิค C4.5, Random Forest, Random Tree, REP Tree, k-NN, และ SVM วัด ประสิทธิภาพของแบบจำลองด้วย Accuracy, Precision และ Recall ผลการทดลองสามารถแสดงได้ดัง Figure 3

Table

Description automatically generated

Figure 3 แสดงให้เห็นค่า Accuracy ของแบบจำลอง ได้แก่เทคนิค C4.5 ให้ค่า Accuracy มากที่สุด 95.36% ต่อมา เทคนิค REP Tree ให้ค่า Accuracy 90.09% เทคนิค Random Forest ให้ค่า Accuracy 83.24% เทคนิค SVM ให้ค่า Accuracy 82.82% เทคนิค Random Tree ให้ค่า Accuracy 76.81% และเทคนิค k-NN ให้ค่า Accuracy 76.18% ตามลำดับ

ค่า Precision ของแบบจำลองได้แก่เทคนิค C4.5 ให้ ค่า Precision มากที่สุด 95.4% ต่อมาเทคนิค REP Tree ให้ ค่า Precision 89.9% เทคนิค Random Forest ให้ค่า Precision 82.7% เทคนิค SVM ให้ค่า Precision 81.6% เทคนิค Random Tree ให้ค่า Precision 75.9% และเทคนิค k-NN ให้ค่า Precision 74.4% ตามลำดับ

ค่า Recall ของแบบจำลองได้แก่เทคนิค C4.5 ให้ค่า Recall มากที่สุด 95.4% ต่อมา เทคนิค REP Tree ให้ค่า Recall 90.1%

เทคนิค Random Forest ให้ค่า Recall 83.2% เทคนิค SVM ให้ค่า Recall 82.8% เทคนิค Random Tree ให้ ค่า Recall 76.8% และเทคนิค k-NN ให้ค่า Recall 76.2% ตาม ลำดับ

**สรุปผลการทดลอง:**

จากผลการทดลอง สรุปได้ว่า C4.5 มีความเหมาะ สมในการพยากรณ์การสำเร็จการศึกษาในสถานศึกษา อาชีวศึกษาเอกชนจังหวัดกาฬสินธุ์ หากผู้ที่สนใจศึกษาหรือ สนใจที่จะพัฒนางานวิจัยนี้ ควรเพิ่มจำนวนข้อมูลชุดฝึกสอน ให้มากขึ้น และเพิ่มในลักษณะของคำแนะนำให้แก่นักเรียนว่า เหมาะสมที่จะเรียนประเภทวิชาอะไร เป็นต้น เพื่อที่จะได้เป็น แนวทางในเลือกตัดสินใจในการเลือกเรียนต่อไป