

การวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดอาชญากรรมด้วยวิธีการทำเหมืองข้อมูล

An Analysis of Factors Associated of Crime Occurred by Data Mining Method

พรพิรุณ โองอินทร์ (Pornpiroon Ong-in)* ดร. วรรณิา สิริแสงตระกูล (Dr.Wanna Sirisangtragul)**

ดร. พุชยดี สิริแสงตระกูล (Dr.Pusadee Seresangtakul)***

บทคัดย่อ

การศึกษานี้นำเสนอการวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดอาชญากรรมในพื้นที่เขตเมืองและนอกเขตเมืองด้วยการทำเหมืองข้อมูล โดยนำขั้นตอนวิธีริปปอร์ (Ripper Algorithm) มาทำการศึกษา ข้อมูลที่นำมาศึกษาเป็นข้อมูลอาชญากรรมจังหวัดขอนแก่น ซึ่งได้จากระบบข้อมูลสารสนเทศ Polis ของสำนักงานตำรวจแห่งชาติ ประกอบด้วย 26 สถานี 5 ประเภทคดี ได้แก่ คดียาเสพติด คดีการพนัน คดีลักทรัพย์ คดีข่มขืนกระทำชำเรา และคดีข่มขืน รงขณะเมาสุรา ผลการศึกษาพบว่า กฎที่ได้จากความสัมพันธ์เกี่ยวกับคดีแต่ละประเภะนั้นมีเวลาและสถานที่เกิดเหตุแตกต่างกันตามเขตพื้นที่ ในคดีลักทรัพย์ สถานที่เกิดเหตุซึ่งพบในเขตเมืองเป็นแหล่งค้าขายหรือแหล่งสาธารณะซึ่งมีผู้คนอาศัยอยู่มาก แต่นอกเขตเมืองจะเป็นเคหะสถาน คดีข่มขืนกระทำชำเราเกิดเหตุในเขตเมืองและนอกเขตเมืองซึ่งเป็นอำเภอขนาดใหญ่เป็นช่วงเวลา 18.00-02.59 น. และนอกเขตเมืองซึ่งเป็นอำเภอขนาดเล็กจะพบช่วงเวลา 09.00-11.59 น. ส่วนคดีข่มขืนกระทำชำเราและคดีการพนันพบการเรียนรู้กฎเฉพาะนอกเขตเมือง ในเขตเมืองไม่พบการเรียนรู้กฎ ประโยชน์ที่ได้จากการศึกษานี้จะสามารถนำไปใช้เป็นแนวทางเพื่อวางแผนป้องกันและปราบปรามอาชญากรรมในพื้นที่ซึ่งมีแนวโน้มเกิดขึ้นในอนาคตได้อย่างมีประสิทธิภาพ

ABSTRACT

This study presented an analysis of the factors associated with the incidence of crime in urban and suburban areas using data mining. The crime data were from 26 police stations in Khon Kaen province. The data were collected from Polis, which is information system of the Royal Thai Police Office. There are 5 crime types including: drug, gambling, theft, rape and drunken driving cases. In this study, the crime data were analyzed and learned using the Ripper algorithm (JRip). The results showed that crime types have associated with time and location where the crime types occurred in different areas. In urban, the theft crime type often occurred in trading places or public areas whereas in suburban it occurred in residence areas. The drunken driving type often occurred in the evening during 06.00 p.m-02.59 a.m. in large districts, but in the small districts, it occurred in the morning during 09.00-11.59 a.m. The rape and gambling cases only had suburban rule learning. In urban, it could not find rule learning. The benefits of this study will be used as guidance for planning, prevention and suppression of crime in areas which are likely to occur in the future more effectively.

คำสำคัญ: อาชญากรรม การทำเหมืองข้อมูล การเรียนรู้กฎ

Key Words: Crime, Data mining, Rule learning

* นักศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

** อาจารย์ สาขาวิชานิติวิทยาศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

*** ผู้ช่วยศาสตราจารย์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

บทนำ

สภาพสังคมไทยในปัจจุบันนั้น ถึงแม้เศรษฐกิจจะเจริญเติบโต แต่ในทางกลับกัน สังคมไทยนั้นประสบกับปัญหาความยากจน การว่างงาน ค่าครองชีพสูง ความเสื่อมโทรมทางศีลธรรม และปัญหาสุขภาพจิต (ธนินทร์, 2551) มาสู่ปัญหาอาชญากรรม ซึ่งส่งผลกระทบต่อความปลอดภัยในชีวิตและทรัพย์สินของประชาชนในสังคม ดังนั้น ผู้พิทักษ์สันติราษฎร์จะต้องป้องกันและปราบปรามอาชญากรรม เช่น การระงับเหตุการณ์หรือโอกาสการกระทำผิดกฎหมาย ในส่วนของการป้องกันและปราบปรามอาชญากรรมนั้น เจ้าหน้าที่ตำรวจจะเน้นไปที่งานสายตรวจ เพื่อเป็นการสอดส่องดูแลความสงบสุขเรียบร้อยของประชาชนจึงต้องมีการวางแผนนโยบายในงานด้านนี้เกี่ยวกับสถานที่ในการเข้าไปตรวจตราการกระทำผิดกฎหมาย แต่เนื่องจากกำลังพลของเจ้าหน้าที่ตำรวจในบางหน่วยงานมีจำนวนจำกัด จึงจำเป็นต้องมีการวางแผนงานของสายตรวจให้มีประสิทธิภาพ ซึ่งในการวางแผนงานสายตรวจในบางท้องที่ยังไม่สามารถพิจารณาพื้นที่จุดเสี่ยงและเวลาการเกิดเหตุอาชญากรรมได้เท่าที่ควรจากการศึกษาที่ผ่านมาพบว่า มีการนำเทคนิคการทำเหมืองข้อมูลมาใช้ในการวิเคราะห์หาความสัมพันธ์ของปัจจัยที่ส่งผลต่อการเกิดปัญหาอาชญากรรมทางเพศ (Hewitt, Beauregard, 2014) อีกทั้งยังมีการวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดอาชญากรรมกับผู้หญิง ซึ่งใช้ขั้นตอนวิธีอปริโอริ (Apriori algorithm) ในการค้นหารูปแบบความสัมพันธ์ระหว่างผู้เสียหายและผู้กระทำความผิด (Bansal, Bhambhu, 2013) และในการศึกษาความสัมพันธ์ของคุณลักษณะของอุบัติการณ์ความเสี่ยงในโรงพยาบาลโดยใช้วิธีต้นไม้ตัดสินใจ (Decision Tree) และหากฎความสัมพันธ์ (Association Rules) แต่พบว่ากฎความสัมพันธ์ที่ได้ไม่มีความแตกต่างจากสถิติพื้นฐาน จึงได้มีการใช้วิธีการเรียนรู้กฎ (Rule Learning) ซึ่งให้กฎความสัมพันธ์ที่น่าสนใจและสามารถนำไปใช้ประโยชน์ได้มากกว่า (กิริติ, กรชวัล, 2557)

ดังนั้น การศึกษานี้จึงได้เลือกนำเสนอการวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดคดีอาญาในพื้นที่เขตเมืองและนอกเขตเมืองด้วยการทำเหมืองข้อมูล โดยนำวิธีการเรียนรู้กฎมาทำการวิเคราะห์ข้อมูลเพื่อศึกษาความสัมพันธ์ของประเภทคดีอาญา 5 ประเภทที่เกิดขึ้นในเขตพื้นที่จังหวัดขอนแก่น ได้แก่ คดียาเสพติด คดีข่มขืนกระทำชำเรา คดีลักทรัพย์ คดีการพนัน คดีขบขี้รถขณะเมาสุรา โดยกฎความสัมพันธ์ที่ได้นั้นสามารถนำมาใช้วิเคราะห์เพื่อหาแนวโน้มการก่อเหตุอาชญากรรมและนำมาวางแผนปราบปรามอาชญากรรมในอนาคตได้อย่างมีประสิทธิภาพ

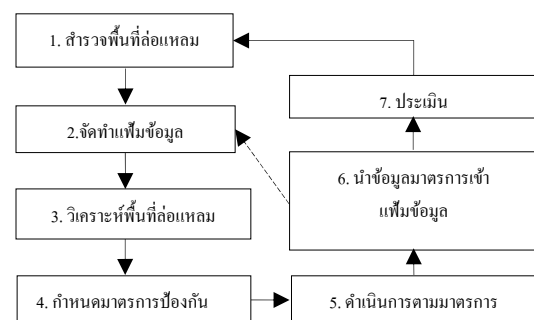
วัตถุประสงค์การวิจัย

ศึกษาการทำเหมืองข้อมูลเพื่อวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดคดีอาญาในพื้นที่เขตเมืองและนอกเขตเมือง

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การป้องกันและปราบปรามอาชญากรรม

อาชญากรรมเป็นการกระทำ หรือคว้นการกระทำที่ฝ่าฝืนกฎหมาย ผู้กระทำความผิดต้องถูกรัฐดำเนินกลงโทษตามบทบัญญัติของกฎหมาย ในการป้องกันปราบปรามอาชญากรรมนั้น มีหลักการที่นำมาใช้ปฏิบัติได้ตามภาพที่ 1 ดังนี้



ภาพที่ 1 การป้องกันอาชญากรรมเชิงรุกระดับสถานีตำรวจ (สำนักงานตำรวจแห่งชาติ, 2556)

1. การสำรวจพื้นที่ล่อแหลม โดยการแบ่งพื้นที่รับผิดชอบให้มีขนาดเหมาะสมกับกำลังพล
2. จัดทำแฟ้มข้อมูลที่ได้ข้อมูลจากขั้นตอนที่ 1

3. ทำการวิเคราะห์พื้นที่ล่อแหลม ระบบรักษาความปลอดภัย และตรวจตราพื้นที่ช่วงเวลาที่สามารถเกิดเหตุได้

4. วางมาตรการป้องกัน ซึ่งบางมาตรการต้องอาศัยการบูรณาการของงานหลายฝ่าย เช่น เจ้าหน้าที่ตำรวจ เจ้าของพื้นที่ ผู้ใช้พื้นที่ เทคโนโลยี และองค์กรต่างๆ

5. มอบหมายให้ผู้รับผิดชอบพื้นที่ไปดำเนินการ โดยมีหัวหน้างานปราบปรามเป็นผู้ควบคุมให้คำแนะนำ

6. รวบรวมข้อมูลมาตรการที่ทำนำเข้าแฟ้มข้อมูลพื้นที่ล่อแหลม ตามขั้นตอนที่ 2

7. ประเมินผลโดยวัดระดับความหวาดกลัวอาชญากรรมของประชาชนในพื้นที่ ประกอบกับสถิติคดีที่เกิดขึ้น

8. สืบสวนพื้นที่ล่อแหลมอื่นๆ ในพื้นที่รับผิดชอบต่อไป

ภารกิจสำคัญในงานป้องกันและปราบปรามอาชญากรรม คือ งานสายตรวจ โดยทั่วไปขั้นตอนการดำเนินงานสายตรวจแบ่งออกได้เป็น 3 ส่วนคือ ขั้นตอนก่อนออกตรวจ ได้แก่ การแบ่งเขตตรวจ การจัดรูปแบบของสายตรวจ การหาข้อมูลและการวิเคราะห์เพื่อวางแผนการตรวจ และการวางแผนการตรวจ ส่วนของขั้นตอนขณะออกตรวจ ได้แก่ การออกปฏิบัติงานสายตรวจ การควบคุมตรวจสอบสายตรวจ และสุดท้ายเป็นขั้นตอนหลังการตรวจ ได้แก่ การปฏิบัติภายหลังการตรวจ และการประเมินผล ซึ่งระบบของงานสายตรวจนั้นสามารถดำเนินการไปพร้อมกับความเปลี่ยนแปลงที่เกิดขึ้นในยุคโลกาภิวัตน์ได้อย่างมีประสิทธิภาพ

การทำเหมืองข้อมูล (Data Mining)

เป็นการวิเคราะห์ข้อมูลเพื่อให้ได้ความรู้ที่ซ่อนอยู่จากฐานข้อมูลขนาดใหญ่ (สุชาติ, 2545) โดยมีกระบวนการดังนี้

1. การเตรียมข้อมูล (Pre-Processing) เป็นการทำให้ข้อมูลมีความถูกต้องและเหมาะสม เปลี่ยนแปลงข้อมูลให้อยู่ในรูปแบบที่สามารถนำมาวิเคราะห์ได้ในกระบวนการนี้มีการคัดเลือกข้อมูล (Data Selection) โดยเลือกเฉพาะข้อมูลที่น่าสนใจจะทำวิเคราะห์ การ

กลั่นกรองข้อมูล (Data Cleaning) เป็นการนำข้อมูลที่ไม่ถูกต้องหรือมีความซ้ำซ้อนออกไป การแปลงข้อมูล (Data Transformation) เป็นการ ทำข้อมูล ให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้ถูกต้องตามวิธีการทำเหมืองข้อมูลที่ใช้

2. การทำเหมืองข้อมูล (Data Mining) ต้องเลือกวิธีการทำเหมืองข้อมูลให้เหมาะสมกับงานที่ต้องการวิเคราะห์ ซึ่งมีงานหลายประเภท เช่น การค้นหากฎความสัมพันธ์ ซึ่งกฎที่ได้สามารถนำไปวิเคราะห์หรือทำนายผลต่อไป การจำแนกประเภทข้อมูล (Classification) เป็นการสร้างโมเดลเพื่อจัดแบ่งกลุ่มข้อมูลที่มีความแตกต่างกัน การพยากรณ์ (Prediction) เป็นการทำนายผลของข้อมูลที่น่าจะเกิดขึ้นในอนาคต การจัดกลุ่มข้อมูล (Clustering) เป็นการแบ่งกลุ่มข้อมูลที่มีลักษณะเหมือนกันอยู่ในกลุ่มเดียวกัน

3. ขั้นตอนหลังกระบวนการประมวลผล (Post-Processing) เป็นการนำความรู้ที่ได้จากการทำเหมืองข้อมูลมาทดสอบเพื่อวิเคราะห์ว่าความรู้ที่ได้มีความถูกต้องตรงตามความต้องการหรือไม่ ซึ่งอาจต้องกลับไปขั้นตอนการทำเหมืองข้อมูลใหม่ เพื่อให้ได้ความรู้ที่ต้องการ

ขั้นตอนวิธีริปเปอร์ (Ripper Algorithm)

เป็นการเรียนรู้กฎ ซึ่งมีกระบวนการตัดแต่งกิ่งที่เพิ่มขึ้นซ้ำๆ เพื่อลดข้อผิดพลาดที่เกิดจากการสร้างกฎขึ้น (Cohen, 1995) มี 4 ขั้นตอน

ขั้นตอนที่ 1 เป็นการเจริญเติบโต (Growth) สร้างลำดับของกฎเฉพาะ โดยการเพิ่มกฎจนกระทั่งได้กฎที่น่าพอใจแล้วจึงจะหยุดเพิ่ม

ขั้นตอนที่ 2 เป็นการตัดแต่งกิ่ง (Pruning) โดยตัดกฎที่ลดประสิทธิภาพการทำงานของการเรียนรู้กฎออก

ขั้นตอนที่ 3 การเพิ่มประสิทธิภาพ (Optimization) มีการเพิ่มคุณลักษณะเข้าไปในแต่ละกฎเดิมหรือกฎที่ถูกสร้างขึ้นใหม่ในขั้นตอนที่ 1 และ 2

ขั้นตอนที่ 4 การเลือกกฎ (Selection) เป็นการเลือกกฎที่ดีที่สุดเก็บไว้ ส่วนกฎอื่นๆ จะถูกตัดออกไป โดยค่าความยาวของคำอธิบายที่น้อยที่สุด (Minimum

Description Length) จะเป็นเกณฑ์ในการหยุดกระบวนการค้นหากฎ (Thangaraj, Vijayalakshmi, 2013) โดยกระบวนการทำงานแสดงตามภาพที่ 2

```

Ruleset Ripper(D) { //D เป็นชุดข้อมูลทั้งหมด
  RS = {} //RS เป็นชุดกฎว่าง
  Ci ordered in increasing prior probability
  //ชุดกฎสำหรับแต่ละคลาส Ci เรียงลำดับคลาสตามค่าความน่าจะเป็น
  for p = 1 to K - 1 {
    Pos = Cp, Neg = Cp+1, ..., Ck
    //Pos เป็นชุดข้อมูลประเภทบวก, Neg เป็นชุดข้อมูลประเภทลบ
    RSp = {} // RSp เป็นชุดกฎว่างที่จัดเก็บกฎจากกระบวนการเติบโตและตัดแต่งกฎ
    while D contains positive samples {
      Divide D into Grow set G and Prune set P
      //G เป็นกฎที่เติบโต, P เป็นกฎที่ถูกตัดออก
      r = GrowRule(G)
      PruneRule(r,P)
      if CalculateError(r) > 0.5 {
        break
      } else {
        RSp = RSp + r
        Remove examples covered by r from D
      }
    }
    for i = 1 to 2 {
      OptimizeRuleset(RSp, D)
      SimplifyRuleset(RSp, D)
    }
    RS = RS + RSp
  } //RS เป็นชุดกฎที่ได้จากการเรียนรู้กฎทั้งหมด
  return RS
}
    
```

ภาพที่ 2 กระบวนการเรียนรู้กฎด้วยวิธีริปเปอร์ (Ata, Yildiz, 2012)

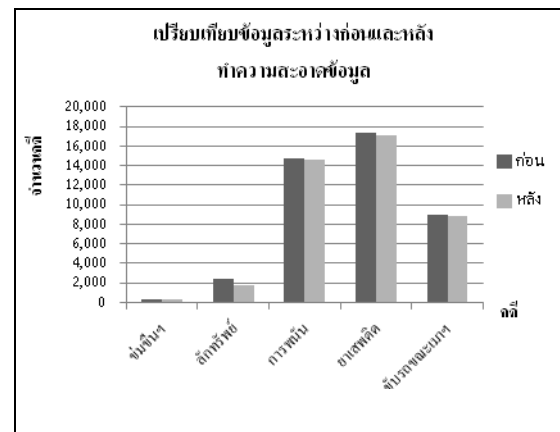
วิธีการวิจัย

ขั้นตอนการเก็บรวบรวมข้อมูล

จากข้อมูลคืออายุ 5 ประเภทซึ่งได้จากรายงานข้อมูลระบบ Polis (สำนักงานตำรวจแห่งชาติ, 2552) ของสำนักงานตำรวจแห่งชาติ โดยเลือกตัวอย่างการวิจัยครั้งนี้ในเขตพื้นที่จังหวัดขอนแก่น ตั้งแต่ปี พ.ศ.2553-2556 จำนวน 43,731 เรกคอร์ด

ขั้นตอนการเตรียมข้อมูล

ก่อนจะนำข้อมูลไปวิเคราะห์ จำเป็นต้องมีการเตรียมข้อมูลให้มีความถูกต้องก่อน โดยเลือกใช้ข้อมูลที่มีคุณภาพตัดทิ้งข้อมูลที่มีค่าคุณลักษณะขาดหายไป ข้อมูลที่มีค่าผิดปกติจะนำมาทำความสะอาดข้อมูลก่อน เช่น ข้อมูลอายุผู้กระทำความผิดที่มีค่าน้อยกว่า 0 ปีหรือมากกว่า 100 ปี ต้องนำข้อมูลวันเกิดมาทำการตรวจสอบก่อน เป็นต้น



ภาพที่ 3 เปรียบเทียบข้อมูลก่อนและหลังทำความสะอาดข้อมูล

จากภาพที่ 3 เป็นการเปรียบเทียบข้อมูลก่อนและหลังทำความสะอาดข้อมูล ซึ่งข้อมูลที่สามารถนำมาทดสอบได้จริงเป็นจำนวน 42,691 เรกคอร์ด คิดเป็นร้อยละ 97.62 ของข้อมูลทั้งหมด โดยได้นำไปวิจัยซึ่งครอบคลุมถึงลักษณะสภาพทางกายภาพที่ก่อให้เกิดอาชญากรรมรวมถึงลักษณะผู้กระทำความผิด มาทำการแทนค่าด้วยรหัสตามตารางที่ 1 จากนั้นนำข้อมูลมาวิเคราะห์หาความสัมพันธ์ของปัจจัยทั้งหมดที่ส่งผลต่อการเกิดคดีอาญาในแต่ละพื้นที่ด้วยวิธีการวิเคราะห์สหสัมพันธ์

(Correlation Analysis) ผลการศึกษาพบว่า มีข้อมูล 7 ปัจจัยที่มีความสัมพันธ์ต่อการเกิดคดีอาญา ได้แก่ ประเภทคดี สถานีดำรวจในแต่ละพื้นที่ที่รับผิดชอบ ประเภทวัน เวลาที่เกิดเหตุ สถานที่เกิดเหตุ และปัจจัยเกี่ยวกับผู้กระทำผิดเช่น เพศ อายุ อาชีพ

ตารางที่ 1 ตัวอย่างการแทนค่ารหัสของตัวแปร

คำอธิบาย ตัวแปร	ตัวแปร นำเข้า	รหัส
สถานีดำรวจ	STATION	1=เมืองขอนแก่น 2=น้ำพอง 3=กระนวน 4=ชำสูง 5= พระยืน 6=อุบลรัตน์ 7=บ้านฝาง 8=เขาสวนกวาง 9=พล 10=ชนบท 11=หนองสองห้อง 12=เวียงน้อย 13=เวียงใหญ่ 14=บ้านไผ่ 15=บ้านแฮด 16=โนนศิลา 17=มัญจาคีรี 18=โลกโพธิ์ชัย 19=เปือยน้อย 20=ชุมแพ 21=ภูเวียง 22=หนองเรือ 23=สีชมพู 24=ภูผาม่าน 25=หนองนาคำ 26=เวียงเก่า
จำนวน ประชากร	POPULATION	1=น้อยกว่า 100,000 คน 2=100,001 - 200,001 คน. 3=มากกว่า 200,000 คนขึ้นไป
ความหนาแน่น ของประชากร	DENSITY	1=0-86 คน/ตร.กม. 2=87-171 คน/ตร.กม. 3=172-257 คน/ตร.กม.
ประเภทคดี	CRIME	1=ข่มขืนกระทำชำเรา 2=ลักทรัพย์ 3=การพนัน 4=ยาเสพติด 5=ขั้บรถขณะเมาสุรา
วันในสัปดาห์	DAY	1=จันทร์ 2=อังคาร 3=พุธ 4=พฤหัสบดี 5=ศุกร์ 6=เสาร์ 7=อาทิตย์
ประเภทวัน	HOLIDAY	0=วันปกติ 1=วันหยุด
เวลาที่เกิดเหตุ	OCCUR _TIME	1=00.00-02.59 น. 2=03.00-05.59 น. 3=06.00-08.59 น. 4=09.00-11.59 น.
สถานที่เกิดเหตุ	OCCUR _PLACE	1=ร้านอาหารและสถานบันเทิง 2=ริมเส้นทางคมนาคม
เพศผู้กระทำผิด	OFFENDER _SEX	1=ชาย 2=หญิง
อายุผู้กระทำผิด	OFFENDER _AGE	1=5-12 ปี 2=13-15 ปี 3=16-18 ปี 4=19-24 ปี 5=25-30 ปี 6=31-35 ปี
อาชีพ ผู้กระทำผิด	OFFENDER _OCCUPY	0=ไม่ทราบ 1=รับราชการ 2=ทหาร 3=ตำรวจ

ขั้นตอนการวิเคราะห์ข้อมูล

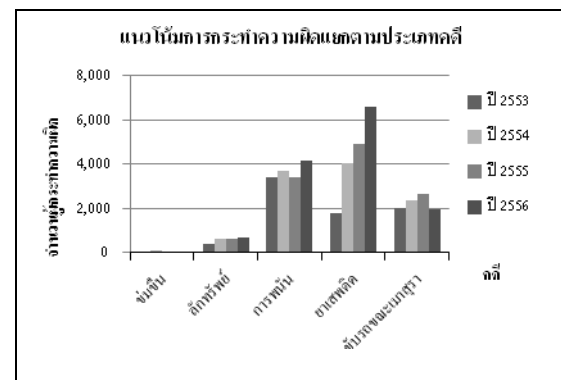
ในขั้นตอนนี้ ได้มีการวิเคราะห์ข้อมูลจากการหาสถิติทั่วไป และใช้การทำเหมืองข้อมูลเพื่อวิเคราะห์หาความสัมพันธ์ระหว่างประเภทคดี เวลา สถานที่เกิดเหตุ ลักษณะผู้กระทำผิด โดยนำหลักการขั้นตอนวิธี Ripper มาใช้วิเคราะห์ข้อมูล ซึ่งในโปรแกรม Weka นั้น เป็นการเลือกใช้วิธี JRip (Xu, 2014) จากนั้นนำกฎที่ได้จากการศึกษามาสอบถามผู้เชี่ยวชาญ เพื่อค้นหากฎที่น่าสนใจและสามารถนำไปใช้วางแผนแก้ปัญหาอาชญากรรมได้จริง

ผลการวิจัย

จากที่ได้กล่าวมาข้างต้น ในส่วนนี้แบ่งผลการวิเคราะห์ข้อมูลออกเป็น 2 ขั้นตอน ดังต่อไปนี้

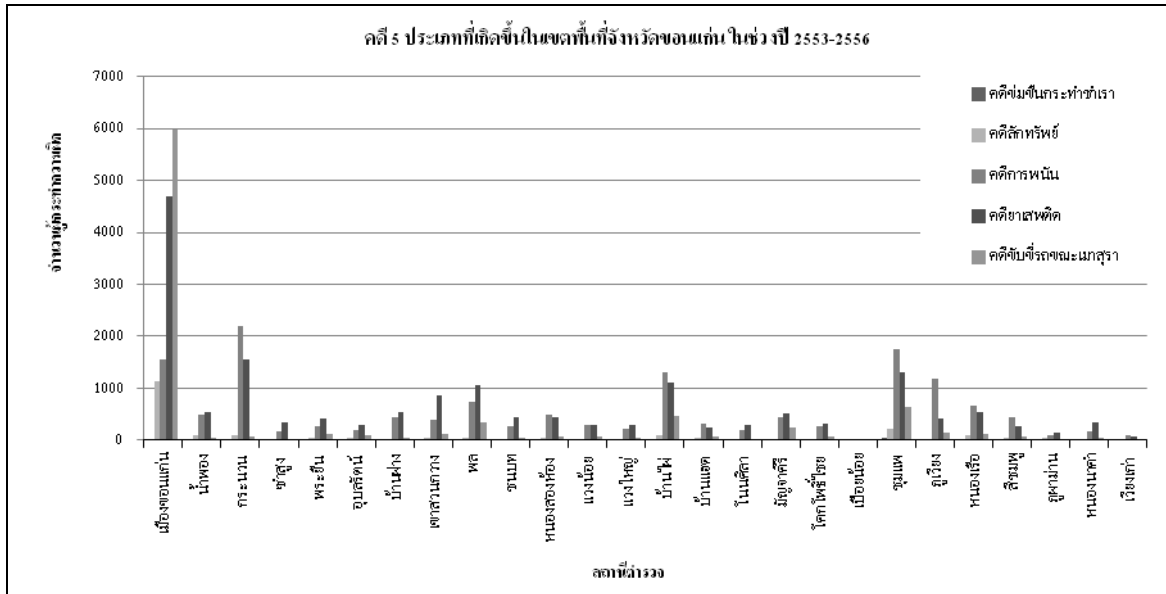
ผลวิเคราะห์จากสถิติ

จากข้อมูลคดีที่นำมาทำการวิจัย 5 ประเภทคดี ได้แก่ คดียาเสพติด คดีข่มขืนกระทำชำเรา คดีลักทรัพย์ คดีการพนัน และคดีขั้บรถขณะเมาสุรา ซึ่งเกิดขึ้นในเขตพื้นที่จังหวัดขอนแก่นช่วงปี พ.ศ.2553-2556 นั้น



ภาพที่ 4 แนวโน้มการกระทำผิดแยกตามประเภทคดีในปี 2553-2556

จากภาพที่ 4 พบว่า มีแนวโน้มการเกิดคดียาเสพติด และคดีการพนันเพิ่มขึ้นเรื่อยๆ และการกระทำผิดในคดียาเสพติดเกิดขึ้นมากที่สุด ส่วนคดีขั้บรถขณะเมาสุราในปี 2553-2555 มีแนวโน้มเพิ่มขึ้น แต่ในปี 2556 มีจำนวนลดลง



ภาพที่ 5 จำนวนผู้กระทำความผิดแต่ละประเภทคดีแยกตามสถานีตำรวจ

ตารางที่ 2 จำนวนประชากร โดยเฉลี่ย พ.ศ.2553-2556

อำเภอ	เฉลี่ยจำนวนประชากร	อำเภอ	เฉลี่ยจำนวนประชากร
เมืองขอนแก่น	393,228	อุบลรัตน์	44,199
ชุมแพ	122,868	แก้งน้อย	42,177
น้ำพอง	110,835	เขาสวนกวาง	37,944
บ้านไผ่	101,169	พระยืน	34,429
หนองเรือ	93,162	บ้านแฮด	32,570
พล	87,270	แก้งใหญ่	29,460
กระนวน	78,575	โนนศิลา	26,408
สีชมพู	78,239	โคกโพธิ์ไชย	25,494
หนองสองห้อง	78,182	ซำสูง	23,697
ภูเวียง	71,899	หนองนาคำ	23,659
มีนบุรี	71,788	ภูผาม่าน	22,674
บ้านฝาง	54,085	เปือยน้อย	19,904
ชนบท	48,760	เวียงเก่า	19,870

จากภาพที่ 5 แสดงให้เห็นว่า ในเขตพื้นที่ สก.เมืองขอนแก่น มีอาชญากรรมเกิดขึ้นมากที่สุด รองลงมา เป็นเขตพื้นที่ สก.ชุมแพ สก.กระนวน, สก.บ้านไผ่ และ สก.พล มีอาชญากรรมลดน้อยลงตามลำดับ ส่วนเขตพื้นที่ สก.เปือยน้อย มีอาชญากรรมเกิดขึ้นน้อยที่สุด และเมื่อเปรียบเทียบกับข้อมูลจำนวนประชากรโดยเฉลี่ยตามตารางที่ 2 ซึ่งอำเภอเมืองขอนแก่นมีจำนวน

ประชากรโดยเฉลี่ยมากที่สุด รองลงมาเป็นอำเภอชุมแพ อำเภอเมือง อำเภอบ้านไผ่ มีจำนวนประชากรโดยเฉลี่ยน้อยลงตามลำดับ ซึ่งเขตพื้นที่อำเภอเหล่านี้มีพื้นที่ขนาดใหญ่ ข้อมูลนี้สามารถบ่งชี้ได้ว่า เขตพื้นที่ซึ่งเป็นเขตเมืองหรือเขตพื้นที่ซึ่งมีขนาดใหญ่ มีประชากรอยู่อาศัยมาก จะมีปัญหาอาชญากรรมมากกว่าเขตพื้นที่นอกเมืองซึ่งมีประชากรอยู่อาศัยน้อยกว่า

เมื่อวิเคราะห์จำนวนผู้กระทำความผิดในแต่ละคดีพบว่า คดีขี้น้ำร้อน/พิษ ยาเสพติด และลักทรัพย์ จะพบมากที่สุด ในเขตพื้นที่ สก.เมืองขอนแก่น คดีการพนันจะพบมากที่สุด ในเขตพื้นที่ สก.กระนวน แสดงให้เห็นว่า การกระทำผิดอาชญากรรมบางอย่างนั้น อาจมีปัจจัยอื่นซึ่งส่งผลให้เกิดการกระทำความผิด นอกเหนือจากจำนวนประชากรในแต่ละพื้นที่ได้ ซึ่งข้อมูลเหล่านี้ไม่สามารถทราบได้จากการวิเคราะห์สถิติทั่วไปว่าปัจจัยหรือสาเหตุที่ทำให้คดีอาญาบางคดีเกิดขึ้นในพื้นที่เขตนอกเมืองมีจำนวนมากกว่าเขตเมือง

ผลวิเคราะห์จากการทำเหมืองข้อมูล

จากการทำเหมืองข้อมูลด้วยโปรแกรม Weka (JRip) ซึ่งเป็นขั้นตอนการเรียนรู้กฎ ทำให้ได้ตัวอย่างกฎความสัมพันธ์ เช่น

1. (OFFENDER_SEX = 1) and (OCCUR_PLACE = 9) and (STATION = 2 0) and (OFFENDER_AGE = 4) and (OCCUR_TIME = 4) => CRIME=10 (4.0/1.0)
2. (OFFENDER_SEX = 1) and (OFFENDER_AGE = 2) and (STATION = 1) and (OCCUR_TIME = 7) => CRIME=11 (16.0/5.0)
3. (STATION = 1) and (OCCUR_TIME = 3) and (OCCUR_PLACE = 36) => CRIME=11 (13.0/6.0)
4. (STATION = 1) and (OCCUR_PLACE = 34) and (OCCUR_TIME = 6) and (OFFENDER_AGE = 5) => CRIME=11 (11.0/3.0)
5. (OFFENDER_SEX = 2) and (OCCUR_TIME = 6) and (OFFENDER_AGE = 8) => CRIME=19 (515.0/34.0)
6. (OFFENDER_SEX = 2) and (OCCUR_TIME = 5) and (STATION = 22) => CRIME=19 (10.0/2.0)
7. (OFFENDER_SEX = 2) and (STATION = 3) and (OCCUR_PLACE = 9) => CRIME=19 (312.0/28.0)
8. (OFFENDER_SEX = 1) and (OCCUR_PLACE = 2) and (STATION = 20) and (OCCUR_TIME = 1) and (OFFENDER_OCCUPY = 6) => CRIME=24 (55.0/15.0)
9. (OFFENDER_SEX = 1) and (OCCUR_PLACE = 2) and (OFFENDER_AGE = 8) and (OCCUR_TIME = 4) and (OFFENDER_OCCUPY = 6) => CRIME=24 (30.0/8.0)
10. (OCCUR_PLACE = 2) and (OFFENDER_SEX = 1) and (STATION = 20) and (OFFENDER_OCCUPY = 6) and (OCCUR_TIME = 7) => CRIME=24 (162.0/42.0)

จากกฎความสัมพันธ์ที่น่าสนใจ แสดงให้เห็นว่าคดีข่มขืนกระทำชำเรา จากการเรียนรู้กฎของ JRip นั้นเป็นความสัมพันธ์ที่เกิดขึ้นนอกเขตเมือง ผู้กระทำความผิดไม่เลือกช่วงเวลาในการก่อเหตุ สถานที่ก่อเหตุเป็นบริเวณ บ้านพัก เคหะสถาน และแหล่งชุมชน ผู้กระทำความผิดเป็นเพศชายอยู่ในช่วงอายุ 13-24 ปี

คดีลักทรัพย์นั้น มีกฎความสัมพันธ์ที่น่าสนใจเกิดขึ้นในเขตและนอกเขตเมือง โดยในเขตเมืองนั้น ผู้กระทำความผิดไม่เลือกช่วงเวลาในการก่อเหตุ สถานที่เกิดเหตุเป็นแหล่งเศรษฐกิจและสาธารณะเช่น ตลาด ร้านค้า โรงพยาบาล ส่วนนอกเขตเมืองเป็นเคหะสถาน ผู้กระทำความผิดเป็นเพศชาย โดยอายุผู้กระทำความผิดในเขตเมืองอยู่ในช่วงอายุ 13-30 ปี ส่วนนอกเขตเมืองอยู่ในช่วงอายุ 13-15 ปี จะพบว่า ในเขตเมืองนั้นมีอายุสูงกว่านอกเขตเมือง

คดีการพนันนั้น มีกฎความสัมพันธ์ที่น่าสนใจเกิดขึ้นนอกเขตเมือง ผู้กระทำความผิดเลือกช่วงเวลาในการก่อเหตุ ช่วงเวลา 12.00 – 17.59 น. สถานที่เกิดเหตุเป็นแหล่งที่พักอาศัยของประชากร เช่น แหล่งชุมชน เคหะสถาน ผู้กระทำความผิดเป็นเพศหญิง อายุ 25-50 ปี ประกอบอาชีพเกษตรกรกรรม

คดีขักรถขณะเมาสุรา มีกฎความสัมพันธ์ที่น่าสนใจเกิดขึ้นในเขตและนอกเขตเมือง โดยจะพบผู้กระทำความผิดในเขตเมืองช่วงเวลา 18.00-23.59 น. ส่วนนอกเขตเมืองซึ่งเป็นอำเภอที่มีจำนวนประชากรมาก จะเกิดเหตุช่วงเวลา 18.00-02.59 น. และนอกเขตเมืองซึ่งเป็นอำเภอที่มีจำนวนประชากรน้อย จะเกิดเหตุช่วงเวลา 09.00-11.59 น. สถานที่เกิดเหตุเป็นแหล่งชุมชนและเส้นทางคมนาคม อายุและอาชีพผู้กระทำความผิดนอกเขตเมืองซึ่งเป็นอำเภอที่มีจำนวนประชากรมากอยู่ในช่วงอายุ 41-45 ปี ประกอบอาชีพรับจ้าง แต่พื้นที่นอกเขตเมืองซึ่งเป็นอำเภอที่มีจำนวนประชากรน้อย ผู้กระทำความผิดประกอบอาชีพเกษตรกรกรรม

ดังนั้น จะเห็นได้ว่า การกระทำความผิดอาชญากรรมแต่ละประเภทจะมีลักษณะการเกิดเหตุเช่น

เวลา สถานที่ที่เกิดเหตุจะแตกต่างกันตามลักษณะเขตพื้นที่ โดยนอกเขตเมืองจะมีลักษณะการกระทำความคิดที่คล้ายกัน แต่ในเขตเมืองการกระทำความคิดเกิดขึ้นลักษณะหลากหลายแตกต่างกัน จึงให้กฎที่น่าสนใจเกี่ยวกับคดีลักทรัพย์ และขับริชณะเมาสุราเท่านั้น ส่วนนอกเขตเมืองให้กฎที่น่าสนใจในแต่ละประเภทคดีมากกว่า เช่น คดีข่มขืนกระทำชำเรา ลักทรัพย์ การพนัน และขับริชณะเมาสุรา ซึ่งผลจากการวิเคราะห์ข้อมูลที่ได้จากการเรียนรู้กฎ นี้เป็นความสัมพันธ์ที่ได้ นอกเหนือจากข้อมูลสถิติพื้นฐาน สามารถนำไปใช้ให้เกิดประโยชน์ในด้านการวางแผน สนับสนุนนโยบายต่างๆ เพื่อสังเกต ตรวจตรา และดูแลความสงบสุขเรียบร้อยของชุมชนได้ และยังเป็นการป้องกันโอกาสที่จะก่ออาชญากรรมได้อีกทางหนึ่ง

อภิปรายและสรุปผลการวิจัย

งานวิจัยนี้ได้ศึกษาการวิเคราะห์ปัจจัยที่มีความสัมพันธ์ต่อการเกิดคดีอาญาในพื้นที่เขตเมืองและนอกเขตเมือง จังหวัดขอนแก่น โดยได้นำข้อมูลคดีอาญา 5 ประเภทมาทำการวิเคราะห์ เพื่อให้เจ้าหน้าที่ผู้ดูแลความสงบเรียบร้อยของสังคมสามารถวิเคราะห์ถึงสถานที่และเวลาที่เกิดเหตุในแต่ละประเภทคดี และกลุ่มเสี่ยงที่น่าจะมีโอกาสก่อเหตุอาชญากรรมได้สะดวกและรวดเร็วขึ้น

การศึกษาปัจจัยที่มีความสัมพันธ์ต่อการเกิดคดีอาญาในพื้นที่เขตเมืองและนอกเขตเมือง ด้วยวิธีการเรียนรู้กฎ จะให้ข้อมูลที่เป็นประโยชน์ นอกเหนือจากการวิเคราะห์สถิติพื้นฐาน เนื่องจากข้อมูลที่นำมาทำการวิเคราะห์เป็นข้อมูลพื้นฐานเกี่ยวกับการก่อเหตุอาชญากรรม จึงทำให้วิเคราะห์ได้เพียงบางส่วน จึงควรมีการเก็บบันทึกข้อมูลให้มีความสมบูรณ์และเพิ่มตัวแปรเกี่ยวกับคุณลักษณะของผู้กระทำความผิดและผู้เสียหายหรือเหยื่อของการเกิดอาชญากรรมในแต่ละครั้ง เพื่อให้การวิเคราะห์ข้อมูลอาชญากรรมทำได้มีประสิทธิภาพมากยิ่งขึ้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับการสนับสนุนข้อมูลจากศูนย์เทคโนโลยีสารสนเทศกลาง สำนักงานตำรวจแห่งชาติ

เอกสารอ้างอิง

กิริติ กระจำง, กรชวัล ชายผา. การวิเคราะห์ทางสถิติและการทำเหมืองข้อมูลสำหรับระบบบริหารจัดการความเสี่ยงในโรงพยาบาล. การประชุมวิชาการวิทยาการคอมพิวเตอร์และวิศวกรรมคอมพิวเตอร์นานาชาติ ครั้งที่ 18; 30 กรกฎาคม- สิงหาคม 2557; ขอนแก่น.
 ขอนแก่น: ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น; 2557. หน้า 97-102.

ชินนทร์ ภูวานคำ. ปัญหาอาชญากรรม [ออนไลน์] 2551 [อ้างเมื่อ 24 ตุลาคม 2557]. จาก <https://www.13nr.org/posts/335795>

สุชาดา กิระนันท์. คำถามคำตอบเกี่ยวกับการทำเหมืองข้อมูลเบื้องต้น [ออนไลน์] 2545 [อ้างเมื่อ 24 กันยายน 2557]. จาก http://web.kku.ac.th/wichuda/Knowledge/6DataMining/Datamining_Suchada.pdf

สำนักงานตำรวจแห่งชาติ. เชื่อมต่อเข้าสู่ระบบ VPN [ออนไลน์] 2552 [อ้างเมื่อ 4 สิงหาคม 2557]. จาก <https://sites.google.com/site/kmthaipolice/home>

สำนักงานตำรวจแห่งชาติ, บทที่ 2 การป้องกันอาชญากรรมเชิงรุก ในคู่มือการบริหารงานป้องกันและปราบปรามอาชญากรรม [ออนไลน์] 2556 [อ้างเมื่อ 19 กันยายน 2557]. จาก <http://www.p1.police.go.th/Download/56/คู่มือการบริหารงานป้องกันปราบปรามอาชญากรรม.pdf>

- Ata, S., Yildiz, OT. Searching for the Optimal Ordering of Classes in Rule Induction. 21st International Conference on Pattern Recognition (ICPR 2012); 2012 November 11-15; Tsukuba. United States: IEEE Computer Society; 2012. p. 1277-1280.
- Bansal, D., Bhambhu, L. Usage of Apriori Algorithm of Data Mining as an Application to Grievous Crimes against Women. International Journal of Computer Trends and Technology 2013; 4(9): 3194-3199.
- Cohen, WW. Fast Effective Rule Induction. In Proceedings of the Twelfth International Conference on Machine Learning; 1995 July 9-12; Tahoe. California: Morgan Kaufmann; 1995. p. 115-123.
- Explorable. com. Statistical Correlation [online] 2009 [cited 2014 Oct 24]. Available from: <https://explorable.com/statistical-correlation>
- Hewitt, A., Beauregard, E. Sexual crime and place: The impact of the environmental context on sexual assault outcomes. Journal of Criminal Justice 2014; 42(5): 375-383.
- Thangaraj, M., Vijayalakshmi, CR. Performances Study on Rule-based Classification Techniques across Multiple Database Relations. International Journal of Applied Information Systems (IJ AIS) 2013; 5(4): 1-7.
- Xu, X., Class JRip [online] 2014 [cited 2014 Oct 13]. Available from: <http://weka.sourceforge.net/doc/stable/weka/classifiers/rules/JRip.html>