การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล

โดย
นายสุเมธัส เนียมแก้ว

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์สิ่งแวดล้อม แผน ก แบบ ก 1 ระดับปริญญามหาบัณฑิต
ภาควิชาวิทยาศาสตร์สิ่งแวดล้อม
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2562
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล

โดย
นายสุเมธัส เนียมแก้ว

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์สิ่งแวดล้อม แผน ก แบบ ก 1 ระดับปริญญามหาบัณฑิต
ภาควิชาวิทยาศาสตร์สิ่งแวดล้อม
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2562
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ANALYSIS OF SPATIAL RELATIONSHIP PATTERN FOR PREDICTION OF DRUG SMUGGLING BASED KNOWLEDGE DISCOVERY

By

MR. Sumethat NIAMKAEO

A Thesis Submitted in Partial Fulfillment of the Requirements

for Master of Science (ENVIRONMENTAL SCIENCE)

Department of ENVIRONMENTAL SCIENCE

Graduate School, Silpakorn University

Academic Year 2019

Copyright of Graduate School, Silpakorn University

หัวข้อ	การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล
โดย	สุเมธัส เนียมแก้ว
สาขาวิชา	วิทยาศาสตร์สิ่งแวดล้อม แผน ก แบบ ก 1 ระดับปริญญามหาบัณฑิต
อาจารย์ที่ปรึกษาหลัก	ผู้ช่วยศาสตราจารย์ ดร. ดิเรกฤทธิ์ บัวเวช

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ได้รับพิจารณาอนุมัติให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต

.....คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร.จุไรรัตน์ นันทานิช)

พิจารณาเห็นชอบโดย

.....ประธานกรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.รัฐพล อ้นแฉ่ง)

.....อาจารย์ที่ปรึกษาหลัก
(ผู้ช่วยศาสตราจารย์ ดร.ดิเรกฤทธิ์ บัวเวช)

.....อาจารย์ที่ปรึกษาร่วม
(ผู้ช่วยศาสตราจารย์ ดร.อรวรรณ เชาวลิต)

.....ผู้ทรงคุณวุฒิภายนอก
(รองศาสตราจารย์ พันตำรวจเอก ดร.พงษ์พิษณุ ภักดีณรงค์)

60311201 : วิทยาศาสตร์สิ่งแวดล้อม แผน ก แบบ ก 1 ระดับปริญญามหาบัณฑิต
คำสำคัญ : การลักลอบลำเลียงยาเสพติด, การค้นพบความรู้ในฐานข้อมูล, ความสัมพันธ์เชิงพื้นที่, ต้นไม้การตัดสินใจ, เบย์อย่างง่าย, การคาดการณ์

นาย สุเมธัส เนียมแก้ว: การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล อาจารย์ที่ปรึกษาวิทยานิพนธ์ : ผู้ช่วยศาสตราจารย์ ดร. ดิเรกฤทธิ์ บัวเวช

ในปัจจุบันปริมาณผู้ใช้ยาเสพติดเพิ่มสูงขึ้น สาเหตุของผู้ใช้ยาเสพติดที่เพิ่มสูงขึ้นมาจากสภาวะสังคมที่บีบคั้น ในประเทศไทยเกิดการลักลอบลำเลียงยาเสพติดตามแนวชายแดน จังหวัดเชียงใหม่ และจังหวัดเชียงรายพบสถิติการลักลอบลำเลียงยาเสพติดสูงที่สุด ใน ในการวิจัยครั้งนี้ได้เลือกพื้นที่ศึกษา 8 อำเภอที่อยู่ติดกับประเทศเพื่อนบ้าน ได้แก่ อำเภอแม่จัน อำเภอแม่ฟ้าหลวง อำเภอแม่สาย จังหวัดเชียงราย อำเภอฝาง อำเภอเชียงดาว อำเภอแม่อาย อำเภอไชยปราการ และอำเภอเวียงแหง จังหวัดเชียงใหม่ การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และเปรียบเทียบวิธีการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ (Decision tree) และวิธี Naïve Bayes ผลการศึกษาพบว่า ปัจจัยตำแหน่งการลักลอบลำเลียงยาเสพติดมีค่า Information gain มากที่สุดเท่ากับ 0.518 ทำให้ปัจจัยตำแหน่งการลักลอบลำเลียงยาเสพติดถูกใช้เป็นโหนดเริ่มต้น พื้นที่อำเภอแม่จัน อำเภอแม่สาย อำเภอแม่อาย และอำเภอฝาง มีความสัมพันธ์กับปัจจัยฤดูกาล การลักลอบลำเลียงยาเสพติดในอำเภอไชยปราการมีความสัมพันธ์กับระยะทางจากจุดตรวจรัศมี 500 เมตร อำเภอแม่ฟ้าหลวง การลักลอบลำเลียงยาเสพติดมีความสัมพันธ์กับปัจจัยการใช้ประโยชน์ที่ดิน และประเภทของยาเสพติดที่จับกุมได้ ผลการศึกษาจากการเปรียบเทียบวิธีที่ใช้ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดพบว่าวิธีต้นไม้ตัดสินใจเป็นวิธีที่ดีที่สุด ซึ่งเห็นได้จากค่าทางสถิติที่ได้จากขั้นตอนการฝึกหัดข้อมูล การทดสอบข้อมูล และการคาดการณ์มีค่าสูงกว่า โดยมีค่าเท่ากับร้อยละ 63.5 57.1 และ 47.7 ตามลำดับ พื้นที่ที่ถูกคาดการณ์ความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดเรียงลำดับจากมากที่สุดไปยังน้อยที่สุดได้แก่ อำเภอไชยปราการ อำเภอแม่จัน อำเภอแม่สาย อำเภอแม่ฟ้าหลวง อำเภอแม่อาย อำเภอฝาง อำเภอเวียงแหง และอำเภอเชียงดาว มีความเสี่ยงจากมากไปน้อยตามลำดับ

60311201 : Major (ENVIRONMENTAL SCIENCE)
Keyword : DRUG SMUGGLING, KNOWLEDGE DISCOVERY FROM DATABASE, SPATIAL RELETIONSHIP, DECISION TREE, NAÏVE BAYES, PREDICTION

MR. SUMETHAT NIAMKAEO : ANALYSIS OF SPATIAL RELATIONSHIP PATTERN FOR PREDICTION OF DRUG SMUGGLING BASED KNOWLEDGE DISCOVERY THESIS ADVISOR : ASSISTANT PROFESSOR DIREKRIT BUAVETCH

The number of drug users has been growing upward which was caused by oppressed social condition. In Thailand, drug smuggling occurs through the borderline of Northern Thailand. Chiang Mai and Chiang Rai provinces showed the highest statistics of drug trafficking. In this research, the study areas of 8 districts, adjacent to neighboring country, were chosen. These include Mae Chan, Mae Fa Luang, and Mae Sai located in Chiang Rai Province, Fang, Chiang Dao, Mae Ai, Chai Prakan, and Wiang Haeng situated in Chiang Mai Province. This study was aimed to discover spatial relationship of factors related to narcotic smuggling using data mining-based decision tree technique and to compare prediction methods between decision tree and Naïve Bayes. The results showed that the factor of geographic locations of drug smuggling arrest illustrated the highest information gain of 0.518, hence it was designated as an initial node. Drug smuggling in Mae Chan, Mae Sai, Mae Ai and Fang districts was related to season factor. The distance from checkpoint showed a spatial relationship with drug smuggling arrests in the Chai Prakan district. In Mae Fah Luang district, drug trafficking was relevant to landuse characteristics and drug exhibited. The results of comparing prediction method demonstrated that the decision tree method was the best prediction method in this case. It was confirmed by better statistic values derived from data training data testing and prediction approach as 63.5, 57.1, and 47.7, respectively. The predicted drug trafficking risk areas were classified from highest to lowest risk level as follows; Chai Prakan, Mae Chan, Mae Sai, Mae Fah Luang, Mae Ai, Fang, Wiang Haeng and Chiang Dao.

กิตติกรรมประกาศ

ขอขอบพระคุณผู้อนุเคราะห์ข้อมูลที่ใช้ในการศึกษาครั้งนี้ ได้แก่ กรมพัฒนาที่ดินที่ได้อนุเคราะห์ข้อมูลขอบเขตการปกครอง และการใช้ประโยชน์ที่ดิน สำนักปลัดกระทรวงคมนาคมสำหรับการอนุเคราะห์ข้อมูลเส้นทางคมนาคม เจ้าหน้าที่ปฏิบัติการในพื้นที่ สำหรับคำแนะนำเกี่ยวกับการลักลอบลำเลียงยาเสพติด และนางสาวเบญจมาศ เดชศิริ สำหรับข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดในปี พ.ศ. 2554 ถึงปี พ.ศ. 2559 และสุดท้ายวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ดีเพราะได้ความกรุณาจาก ผศ.ดร.อรประภา ภุมมะกาญจนะ โรแบร์ และผศ.ดร.อรวรรณ เชาวลิต ซึ่งเป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ที่ให้ความช่วยเหลือ และให้คำแนะนำที่เป็นประโยชน์อย่างยิ่งต่อผู้วิจัย ส่งผลให้วิทยานิพนธ์เล่มนี้ถูกต้องและสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณในความกรุณาของทุกท่านเป็นอย่างสูง

สุเมธัส เนียมแก้ว

มหาวิทยาลัยศิลปากร

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ญ

สารบัญภาพ ..... ฎ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญ 1

1.2 วัตถุประสงค์ของการศึกษา 4

1.3 ประโยชน์ที่ได้รับจากการศึกษา 4

1.4 ขอบเขตการศึกษา 4

บทที่ 2 ทบทวนวรรณกรรม 6

2.1 สถานการณ์ยาเสพติด 6

2.2 การค้นพบความรู้จากฐานข้อมูล (Knowledge Discovery from Database) 7

2.2.1 การสร้างคลังข้อมูล 8

2.2.2 การทำเหมืองข้อมูล 9

2.2.2.1 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) 11

2.2.2.2 การเรียนรู้แบบมีผู้สอน (Supervised learning) 12

2.2.3 การคาดการณ์ และการแสดงผลการทำเหมืองข้อมูล 17

บทที่ 3 วิธีการดำเนินงานวิจัย 18

3.1 ข้อมูล 20

3.2 อุปกรณ์ที่ใช้ในการวิจัย..... 21
3.3 วิธีการศึกษา..... 21
3.3.1 การรวบรวมข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติด 22
3.3.2 ขั้นตอนการปรับข้อมูลให้อยู่ในรูปฐานข้อมูลระบบสารสนเทศทางภูมิศาสตร์ 24
3.3.3 การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด..... 25
3.3.3.1 การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วยการทำเหมืองข้อมูลวิธีต้นไม้การตัดสินใจ..... 26
3.3.3.2 การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes 28
3.3.4 ขั้นตอนการแสดงผลการศึกษา..... 31
บทที่ 4 ผลการศึกษา..... 32
4.1 ผลการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ 32
4.1.1 ผลการคำนวณหาความสำคัญของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด เพื่อใช้เป็นตัวกำหนดโหนดเริ่มต้น 32
4.1.2 ผลของความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด 33
4.2 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ (Decision tree) และวิธีเบย์อย่างง่าย (Naïve Bayes) 36
4.2.1 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ 36
4.2.1.1 ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด..... 36
4.2.1.2 ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด..... 37

4.2.1.3 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ 38
4.2.2 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes 40
4.2.2.1 ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด 41
4.2.2.2 ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด 41
4.2.2.3 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes 42
4.2.3 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes 44
4.2.3.1 การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes จากร้อยละความถูกต้อง 45
4.2.3.2 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes รูปแบบเชิงพื้นที่ 46
บทที่ 5 สรุปและข้อเสนอแนะ 49
5.1 สรุป 49
5.2 ข้อเสนอแนะ 50
รายการอ้างอิง 51
ประวัติผู้เขียน 56

สารบัญตาราง

หน้า

ตารางที่ 1 สถิติการจับกุมคดียาเสพติด 2

ตารางที่ 2 ประเภทการทำเหมืองข้อมูล 11

ตารางที่ 3 ข้อมูลที่ใช้ในการศึกษา 20

ตารางที่ 4 ตัวอย่างข้อมูลที่รวบรวมจากข่าวการจับกุมยาเสพติด 23

ตารางที่ 5 ค่า Information gain ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด 33

ตารางที่ 6 ค่า Gini index ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด 33

ตารางที่ 7 ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการฝึกหัดข้อมูล 37

ตารางที่ 8 ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการทดสอบข้อมูล 38

ตารางที่ 9 ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการคาดการณ์ 39

ตารางที่ 10 รายละเอียดผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยต้นไม้การตัดสินใจ 40

ตารางที่ 11 ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการฝึกหัดข้อมูล 41

ตารางที่ 12 ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการทดสอบข้อมูล 42

ตารางที่ 13 ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการคาดการณ์ 43

ตารางที่ 14 รายละเอียดผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes 44

ตารางที่ 15 เกณฑ์การกำหนดระดับความเสี่ยงจากค่า F-measure 46

สารบัญภาพ

หน้า

ภาพที่ 1 เส้นทางการลักลอบลำเลียงยาเสพติดเข้ามาในประเทศ 2

ภาพที่ 2 เส้นทางการลักลอบลำเลียงยาเสพติดภายในประเทศเพื่อย้ายจุดพักยา 3

ภาพที่ 3 แผนที่ขอบเขตพื้นที่ศึกษา 5

ภาพที่ 4 ความสัมพันธ์ของเหตุการณ์ B และ A ที่เกิดขึ้นไม่พร้อมกัน 16

ภาพที่ 5 ขั้นตอนการดำเนินงานวิจัย 18

ภาพที่ 6 แผนที่ปัจจัยเชิงพื้นที่ที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด 25

ภาพที่ 7 การจัดรูปแบบในการคำนวณค่า Information gain ด้วย RapidMiner 27

ภาพที่ 8 การจัดรูปแบบในการคำนวณค่า Gini index ด้วย RapidMiner 27

ภาพที่ 9 การจัดรูปแบบในการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วย RapidMiner 27

ภาพที่ 10 การเลือกวิธีการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด 29

ภาพที่ 11 การฝึกหัดข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด29

ภาพที่ 12 การทดสอบข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด 30

ภาพที่ 13 การคาดการณ์จากข้อมูลที่ไม่ปรากฏในคลังข้อมูลที่ผ่านการฝึกหัดข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด 30

ภาพที่ 14 ความสัมพันธ์เชิงพื้นที่ของปัจจัยการลักลอบลำเลียงยาเสพติด 34

ภาพที่ 15 กราฟเปรียบเทียบค่าความถูกต้องของวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes 45

ภาพที่ 16 แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากวิธีต้นไม้การตัดสินใจ 47

ภาพที่ 17 แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากวิธี Naïve Bayes..... 47

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

ในปัจจุบันปัญหาหนึ่งที่สำคัญของประเทศไทย คือ ปัญหาทางด้านยาเสพติด ซึ่งเป็นปัญหาที่ส่งผลกระทบต่อทางด้านภาพลักษณ์ของประเทศ และส่งผลให้เกิดปัญหาอาชญากรรมด้านอื่น ๆ ตามมา การควบคุมยาเสพติดในประเทศไทยได้มีการจัดตั้งพระราชบัญญัติยาเสพติดให้โทษ และบทลงโทษสำหรับผู้ค้า และผู้เสพ (กองควบคุมวัตถุเสพติด, 2559) อย่างไรก็ตามมีกฎหมายควบคุมปัญหาทางด้านยาเสพติดไม่ได้ทำให้แนวโน้มของปัญหาที่เกิดจากยาเสพติดลดลง เนื่องจากยังมีการผลิตอย่างต่อเนื่องและมีการลักลอบลำเลียงยาเสพติดจากประเทศเพื่อนบ้านเข้ามาในประเทศ ทำให้ปัญหาดังกล่าวไม่หมดไปจากประเทศ โดยส่วนใหญ่ปัญหาการลักลอบลำเลียงยาเสพติดเกิดขึ้นตามแนวชายแดนของประเทศไทย เนื่องจากแหล่งผลิตอยู่ในประเทศเพื่อนบ้าน และมีแหล่งพักยาตั้งอยู่ตามแนวชายแดนของทั้ง 2 ฝั่งของประเทศ ทำให้สามารถลักลอบลำเลียงยาเสพติดจากประเทศเพื่อนบ้านมายังประเทศไทยได้สะดวกยิ่งขึ้น จึงได้มีการศึกษาความสัมพันธ์ของปัจจัยเชิงพื้นที่ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่ตามแนวชายแดน โดยพื้นที่จังหวัดเชียงรายมีพื้นที่ติดต่อกับสาธารณรัฐแห่งสหภาพพม่า และสาธารณรัฐประชาธิปไตยประชาชนลาว กลุ่มงานยุทธศาสตร์และข้อมูลเพื่อการพัฒนา (สำนักงานจังหวัดเชียงราย, 2561)

จากรายงานของสำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติดในปี 2560 (สำนักงานคณะกรรมการป้องกันและปราบปรามยาเสพติด, 2560) พบว่าสถานการณ์ยาเสพติดในพื้นที่ภาคเหนือตอนบนของประเทศไทยยังมีการลักลอบนำเข้ายาเสพติดอย่างต่อเนื่อง พื้นที่หลักในการลักลอบลำเลียงยาเสพติดเข้าสู่พื้นที่ภาคเหนือตอนบนของประเทศไทย คือ อำเภอแม่ฟ้าหลวง อำเภอแม่สาย จังหวัดเชียงราย และ อำเภอฝาง อำเภอแม่อาย จังหวัดเชียงใหม่ เนื่องจากอยู่ใกล้แหล่งผลิตและแหล่งพักยาในเขตพื้นที่ประเทศเพื่อนบ้าน วิธีการลักลอบลำเลียงอยู่ในรูปแบบต่าง ๆ เช่น การเดินเท้า รถจักรยานยนต์ รถยนต์กระบะปิคอัพ รถยนต์นั่งส่วนบุคคล และรถบรรทุกพืชผลทางการเกษตร

จากตารางที่ 1 แสดงข้อมูลสถิติการจับกุมคดียาเสพติด ในพื้นที่ภาคเหนือตอนบนพบว่าจังหวัดเชียงใหม่ และจังหวัดเชียงรายมีสถิติการจับกุมอยู่ใน 2 อันดับแรก สาเหตุเป็นเพราะทั้ง 2 จังหวัดมีอาณาเขตติดกับประเทศเพื่อนบ้าน ส่งผลให้พื้นที่หลักของการลักลอบลำเลียงยาเสพติดอยู่ใน 2 จังหวัดนี้ พบสถิติการจับกุมมีค่าสูงกว่าจังหวัดอื่น ๆ การศึกษาครั้งนี้จึงเลือกพื้นที่อำเภอแม่ฟ้าหลวง อำเภอแม่สาย อำเภอแม่จัน อำเภอฝาง อำเภอแม่อาย รวมถึงพื้นที่ในอำเภอเชียงดาว อำเภอไชยปราการ อำเภอเวียงแหง เพิ่มเติมเนื่องจากทั้งสามอำเภอนี้ มีเส้นทางการเดินทางติดต่อกับอำเภอที่กล่าวมาข้างต้น ส่งผลให้ 3 อำเภอนี้เป็นทางผ่านในการลักลอบลำเลียงยาเสพติด

ตารางที่ 1 *สถิติการจับกุมคดียาเสพติด*

เดือน	เชียงใหม่	เชียงราย	แพร่	แม่ฮ่องสอน	น่าน	พะเยา	ลำปาง	ลำพูน	รวม
ต.ค.59	601	302	260	47	168	101	197	127	1,803
พ.ย.59	n/a	394	213	55	180	130	194	138	1,304
ธ.ค.59	646	396	148	49	170	144	174	185	1,912
ม.ค.60	502	413	210	43	164	159	n/a	160	1,651
รวม	1,749	1,505	831	194	682	534	565	610	6,670

ที่มา: สำนักงาน ป.ป.ส. (2560)

หมายเหตุ สถิติการจับกุมคดียาเสพติดเป็นการรวบรวมจากทุกคดีที่เกิดขึ้นจากยาเสพติด

ภาพที่ 1 *เส้นทางการลักลอบลำเลียงยาเสพติดเข้ามาในประเทศ*

ที่มา: สำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติด, 2560

ภาพที่ 2 *เส้นทางการลักลอบลำเลียงยาเสพติดภายในประเทศเพื่อย้ายจุดพักยา*
ที่มา: สำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติด, 2560

จากภาพที่ 1 และภาพที่ 2 แสดงเส้นทางการลักลอบลำเลียงยาเสพติด ทั้งจากการลักลอบลำเลียงยาเสพติดจากประเทศเพื่อนบ้านเข้าสู่ประเทศไทย และการลักลอบลำเลียงยาเสพติดภายในประเทศ จากภาพที่แสดงทั้ง 2 จะเป็นส่วนช่วยในการพิจารณาเลือกพื้นที่ศึกษาร่วมกับข้อมูลทางด้านสถิติในตารางที่ 1

การศึกษาครั้งนี้รวบรวมข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติดในช่วงปี พ.ศ. 2554 ถึงปี พ.ศ. 2560 จากข่าวสารทางสื่อต่าง ๆ เช่น อินเตอร์เน็ต โทรทัศน์ และหนังสือพิมพ์ เป็นต้น ข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดที่รวบรวมจะถูกทำการปรับแต่งข้อมูลด้วยการจัดโครงสร้างให้อยู่ในรูปของตารางฐานข้อมูล (.dbf) ซึ่งเป็นการจัดทำฐานข้อมูลที่เกี่ยวข้องกับการลักลอบลำเลียงยาเสพติด จากนั้นดำเนินการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ และการคาดการณ์การลักลอบลำเลียงยาเสพติดจากข้อมูลดังกล่าวข้างต้นด้วยวิธี Knowledge Discovery from Database (KDD) กระบวนการหลักของ Knowledge Discovery นี้ประกอบด้วย การทำคลังข้อมูล (Data warehouse) การทำเหมืองข้อมูล (Data mining) และการแสดงผล (Visualization) การประยุกต์ใช้ Knowledge Discovery ทำให้ได้มาซึ่งผลลัพธ์ความสัมพันธ์เชิง

พื้นที่ของปัจจัยที่มีความสัมพันธ์ต่อการลักลอบลำเลียงยาเสพติด และสำหรับการคาดการณ์การลักลอบลำเลียงยาเสพติดนั้น ผู้วิจัยดำเนินการศึกษาอัลกอริทึมการทำเหมืองข้อมูลที่เหมาะสมระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ที่สามารถนำมาใช้ในการคาดการณ์การลักลอบลำเลียงยาเสพติดได้ จากการทบทวนวรรณกรรมผู้วิจัยพบว่า ที่ผ่านมายังไม่มีการประยุกต์ใช้ Knowledge Discovery ในการศึกษาการลักลอบลำเลียงยาเสพติด

1.2 วัตถุประสงค์ของการศึกษา

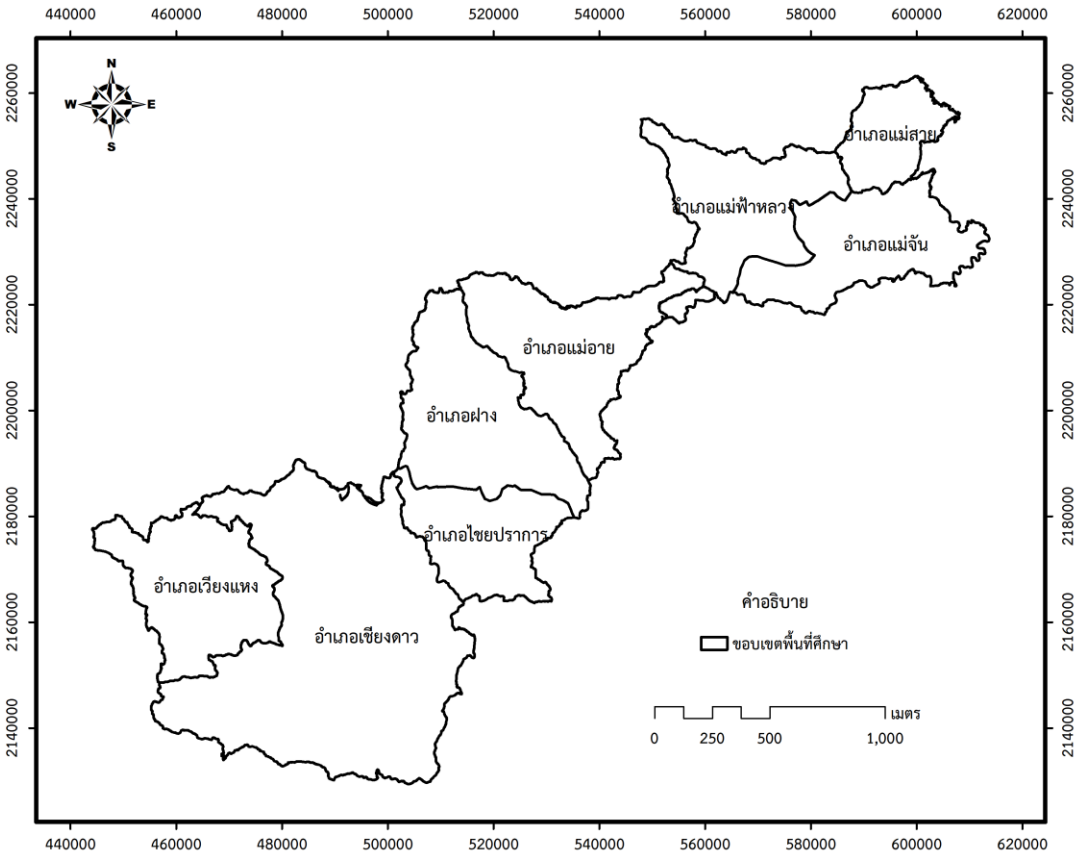
1. เพื่อศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ
2. เพื่อเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ (Decision tree) และวิธีเบย์อย่างง่าย (Naïve Bayes)

1.3 ประโยชน์ที่ได้รับจากการศึกษา

ประโยชน์ที่ได้รับจากการศึกษาเรื่อง การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติด พบว่ามีประโยชน์ต่อหน่วย และเจ้าหน้าที่เกี่ยวข้องในการตัดสินใจด้านการกำหนดมาตราการเพื่อพร้อมการจับกุมการลักลอบลำเลียงยาเสพติดในพื้นที่ศึกษา อีกทั้งการศึกษาครั้งนี้จะเป็นแนวทางในการศึกษาปัญการลักลอบลำเลียงยาเสพติดต่อไป

1.4 ขอบเขตการศึกษา

ในการศึกษาการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่และการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยหลักการ Knowledge Discovery from Database นี้ จะใช้การรวบรวมข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติด (ไม่รวมข่าวที่มีการจับกุมผู้ที่กระทำความผิดจากการเสพ) จากสื่อต่าง ๆ เช่น หนังสือพิมพ์ และอินเตอร์เน็ต เป็นต้น ทั้งที่เป็นสื่อในระดับท้องถิ่น และระดับชาติ ข่าวการจับกุมการลักลอบลำเลียงยาเสพติดที่รวบรวมอยู่ในช่วงปี พ.ศ.2554 ถึงปี พ.ศ.2560 พื้นที่ที่ศึกษากำหนดจากการรายงานจากสำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติด (สำนักงานคณะกรรมการป้องกันและปราบปรามยาเสพติด, 2560) ดังที่กล่าวไว้ในหัวข้อที่ 1.1 ความเป็นมาและความสำคัญของปัญหา ภาพที่ 3 แสดงแผนที่ขอบเขตการศึกษา



ภาพที่ 3 *แผนที่ขอบเขตพื้นที่ศึกษา*

มหาวิทยาลัยศิลปากร

บทที่ 2
ทบทวนวรรณกรรม

2.1 สถานการณ์ยาเสพติด

สถานการณ์ยาเสพติดในประเทศจากรายงานของ United Nations Office on Drug and Crime (2018) ที่ได้สำรวจจากประชากรโลกช่วงอายุ 15 ปี ถึง 64 ปี พบว่าในปี 2016 มีผู้ใช้ยาเสพติดจำนวน 275 ล้านคนทั่วโลก เพิ่มขึ้นจากปี 2015 สูงถึง 20 ล้านคน และคาดว่าแนวโน้มของผู้ใช้ยาเสพติดจะมีแนวโน้มเพิ่มขึ้นเป็น 346 ล้านคน (United Nations Office on Drugs and Crime, 2018) จากรายงานดังกล่าวจะสังเกตได้ว่าอุปสงค์ยาเสพติดมีแนวโน้มที่เพิ่มขึ้น ดังนั้นอุปทานยาเสพติดจะต้องเพิ่มขึ้นด้วย แนวโน้มความต้องการยาเสพติดที่เพิ่มขึ้นในระดับโลกเป็นผลมาจากความต้องการในระดับภูมิภาคที่เพิ่มขึ้น เช่นเดียวกับประเทศไทยที่มีความต้องการยาเสพติดเพิ่มขึ้น เนื่องจากสังคมในปัจจุบันมีแรงผลักดันให้เกิดการใช้ยาเสพติดเพื่อแสวงหาความเปลี่ยนแปลงทางอารมณ์ในสภาวะที่สังคมบีบคั้น และเลือกการใช้ยาเสพติดเป็นการแก้ไขปัญหาเฉพาะหน้า (สำนักงานคณะกรรมการป้องกันและปราบปรามยาเสพติด, 2560) ประเทศไทยไม่มีแหล่งผลิตยาเสพติดจึงนำไปสู่การลักลอบลำเลียงยาเสพติดตามแนวชายแดน และจากรายงานของสำนักงานคณะกรรมการป้องกัน และปราบปรามยาเสพติด (2560) พบว่าแนวชายแดนในพื้นที่ภาคเหนือตอนบนของประเทศไทยเป็นแนวชายแดนที่มีสถิติการลักลอบลำเลียงยาเสพติดมากที่สุด โดยมีอำเภอที่มีอาณาเขตติดกับแนวชายแดนดังต่อไปนี้ อำเภอแม่จัน อำเภอแม่ฟ้าหลวง อำเภอแม่สาย จังหวัดเชียงราย อำเภอฝาง อำเภอเชียงดาว อำเภอแม่อาย อำเภอไชยปราการ และอำเภอเวียงแหง จังหวัดเชียงใหม่ ในการศึกษาครั้งนี้จึงได้กำหนดขอบเขตการศึกษาที่ครอบคลุมพื้นที่ของอำเภอดังกล่าว การศึกษาเกี่ยวกับการลักลอบลำเลียงยาเสพติดพบว่าได้มีการใช้กระบวนการลำดับชั้นเชิงวิเคราะห์ (Analytic Hierarchy Process; AHP) ในการกำหนดความสำคัญของปัจจัยการลักลอบลำเลียงยาเสพติดที่ประกอบด้วยปัจจัยทางกายภาพ ปัจจัยด้านยาเสพติด และปัจจัยด้านสังคมประชากร (Robert O., Kumsap C., & Janpengpen A., 2018) ซึ่งกระบวนการลำดับชั้นเชิงวิเคราะห์เป็นกระบวนการที่วิเคราะห์ปัจจัยทีละคู่ โดยต้องมีโครงสร้างของลำดับชั้นความสำคัญที่แน่นอน การลักลอบลำเลียงกัญชา และฝิ่นในประเทศเม็กซิโก (Medel M., Lu Y., & Chow E, 2015) ได้ทำการใช้หลักการ Cost surface ที่เป็นหลักการการแสดงสมรรถภาพของการลักลอบลำเลียงกัญชา และฝิ่นในรูปแบบของต้นทุนในการลักลอบลำเลียงบนเส้นทางคมนาคม เพื่อทำการคาดการณ์เส้นทางการลักลอบลำเลียงกัญชา และฝิ่นจากความสัมพันธ์ของปัจจัยทางกายภาพ ปัจจัยสังคม และประชากร และปัจจัยทางด้านยาเสพติด

2.2 การค้นพบความรู้จากฐานข้อมูล (Knowledge Discovery from Database)

งานวิจัยฉบับนี้เป็นการศึกษาเกี่ยวกับ Knowledge Discovery from Database ของสถานการณ์การลักลอบลำเลียงยาเสพติดในพื้นที่ในบางอำเภอของจังหวัดเชียงใหม่ และจังหวัดเชียงราย Knowledge Discovery from Database คือ การตอบสนองต่อข้อมูลปริมาณมหาศาลที่จัดเก็บไว้ในฐานข้อมูลปฏิบัติการ และฐานข้อมูลทางวิทยาศาสตร์ โดยทั่วไปจะรู้จักกันอย่างแพร่หลายในชื่อของ “การทำเหมืองข้อมูล” แต่ที่จริงแล้วการทำเหมืองข้อมูลเป็นเพียงองค์ประกอบหนึ่งของกระบวนการ Knowledge Discovery from Database ที่มีขนาดใหญ่เท่านั้น Knowledge Discovery from Database เป็นกระบวนการระดับสูงในการรวบรวมข้อมูลจากการทำเหมืองข้อมูลและการกลั่นกรองข้อมูล ซึ่งจะทำการตีความจากข้อมูลที่รวบรวมมา และผนวกเข้ากับข้อมูลเดิมที่มีอยู่ ทำให้ได้มาซึ่งชุดความรู้ใหม่ขึ้น (Harvey JM. & Jiawei H., 2009)

Knowledge Discovery from Databases มีหลักความเชื่อในการค้นหาคุณสมบัติ และความสัมพันธ์ของข้อมูลจากฐานข้อมูลขนาดใหญ่ได้แก่ หลักความเชื่อ valid, novel, useful และ understandable ซึ่งแต่ละหลักความเชื่อให้คำจำกัดความไว้ดังนี้ valid คือ ข้อมูลเดิมที่มีความถูกต้องแน่นอนเพื่อใช้ในการผนวกกับข้อมูลใหม่ หลักความเชื่อ novel คือ ข้อมูลใหม่ที่ได้จากการรวบรวม และจะสามารถนำไปใช้หาคุณสมบัติ และความสัมพันธ์ได้ หลักความเชื่อ useful จะใช้ในการหาคุณสมบัติ และความสัมพันธ์ แต่สุดท้ายทุกอย่างของการค้นหาคุณสมบัติ และความสัมพันธ์ของข้อมูลจะต้องมีความเป็น understandable ซึ่งหมายถึงทุกขั้นตอนจะต้องเข้าใจได้ง่าย (Fayyad U., Piatetsky-Shapiro G., & Smyth P., 1996) หลักความเชื่อที่กล่าวมาข้างต้นทั้งหมดนี้ ถือว่าเป็นพื้นฐานของการทำเหมืองข้อมูล ตัวอย่างเช่น ข้อมูลปัจจัยฤดูกาลที่มีอยู่แล้ว เชื่อว่าจะมีการลักลอบในฤดูหนาวมากที่สุด (Valid) เมื่อทำการรวบรวมข้อมูลใหม่ (Novel) นำมารวมกับข้อมูลเดิมจะได้คุณสมบัติ (Useful) และความสัมพันธ์ที่เข้าใจได้ง่าย (Understandable) ว่าฤดูหนาวจะมีการลักลอบลำเลียงมากที่สุด

นอกจากนี้กระบวนการ Knowledge Discovery from Database ได้มีการใช้หลักการทางสถิติมาใช้ในการวิเคราะห์ความสัมพันธ์เชิงสถิติ ตัวอย่างเช่น การวิเคราะห์แบบปัวส์ซองที่ใช้ในการวิเคราะห์แบบการแจกแจงไม่ปกติ เพื่อหาความสัมพันธ์เชิงสถิติเหตุการณ์ต่อเหตุการณ์ นอกจากนั้นยังมีหลักการของ Machine Learning มาใช้ในกระบวนการ Knowledge Discovery from Database ซึ่ง Machine Learning คือ การทำให้ระบบคอมพิวเตอร์เรียนรู้ด้วยตัวเองได้จากการนำข้อมูลเข้า การจดจำรูปแบบของข้อมูล การค้นหาตัวเลข และการตั้งสมมุติฐานทางวิทยาศาสตร์ หลักการที่กล่าวมาข้างต้นจัดอยู่ในขั้นตอนของการทำเหมืองข้อมูล เพื่อใช้ในการวิเคราะห์ข้อมูลทุติย

ภูมิในฐานข้อมูลที่มีขนาดใหญ่ การทำเหมืองข้อมูลใช้ในการหาความสัมพันธ์ของข้อมูลที่สนใจในฐานข้อมูล และนำการวิเคราะห์ไปใช้ประโยชน์ต่อไป (Hand DJ., 1998) ตัวอย่างเช่น การใช้งานอัลกอริทึม เครือข่ายแบบเบย์เซียนกับต้นไม้การตัดสินใจ สามารถนำข้อมูลที่รวบรวมไปสอนให้คอมพิวเตอร์รับรู้ ตัวอย่างเช่น ข้อมูลปัจจัยเชิงพื้นที่ของการลักลอบลำเลียงยาเสพติดนำเข้าสู่กระบวนการจดจำรูปแบบของคอมพิวเตอร์ โดยใช้หลักการตั้งสมมุติฐานทางวิทยาศาสตร์ให้คอมพิวเตอร์ได้เรียนรู้ข้อมูลปัจจัยเชิงพื้นที่ของการลักลอบลำเลียงยาเสพติด เมื่อมีข้อมูลปัจจัยเชิงพื้นที่ของการลักลอบลำเลียงยาเสพติดใหม่คอมพิวเตอร์จะสามารถจำแนกข้อมูลได้ ตัวอย่างการนำไปใช้ประโยชน์ของหลักการ Knowledge Discovery from Database ซึ่ง Knowledge Discovery from Database ได้ถูกใช้ในการวิเคราะห์ความสัมพันธ์ระหว่างการฆ่าตัวตายกับลักษณะสภาพอากาศที่เปลี่ยนแปลงในประเทศเม็กซิโก จากการสร้างคลังข้อมูลประชากรชาวเม็กซิโกที่ฆ่าตัวตายจากการแขวนคอตั้งแต่ปี ค.ศ. 2005 ถึง ปี ค.ศ. 2012 ข้อมูลดังกล่าวได้นำเข้าสู่กระบวนการทำเหมืองข้อมูลด้วยอัลกอริทึมการหากฎความสัมพันธ์ (Association rule) ผลการศึกษาพบว่าในวันที่ไม่มีฝนตกอุณหภูมิในช่วง 30 องศาเซลเซียส ถึง 40 องศาเซลเซียส จะพบผู้ชายชาวเม็กซิโกฆ่าตัวตายมากที่สุด โดยแสดงค่าสถิติความเชื่อมั่นที่มีค่าเท่ากับ 0.86 (Fernández-Arteaga V. et al., 2016) นอกจากนั้นหลักการ Knowledge Discovery from Database ได้ถูกนำมาใช้ในการจัดการผู้โดยสารสายการบินที่ประเทศไต้หวัน โดยใช้การทำเหมืองข้อมูลด้วยอัลกอริทึมต้นไม้การตัดสินใจในการหาความสัมพันธ์จากข้อมูลผู้โดยสารที่ใช้สายการบินไต้หวัน เพื่อใช้ในการตัดสินใจในการจัดอันดับให้บริการกับผู้โดยสาร (Wong JY & Chung PH, 2007)

Knowledge Discovery from Databases มีขั้นตอนการศึกษา 3 ขั้นตอนดังนี้ (2.2.1) การสร้างคลังข้อมูล (2.2.2) การทำเหมืองข้อมูล และ (2.2.3) การคาดการณ์ และการแสดงผลการทำเหมืองข้อมูล (Harvey JM. & Jiawei H., 2009) สามารถอธิบายขั้นตอนการศึกษาทั้ง 3 ขั้นตอนได้ดังนี้

2.2.1 การสร้างคลังข้อมูล

คลังข้อมูล คือ การเก็บรวบรวมข้อมูลที่มีจุดมุ่งหมายเพื่อสร้างความรู้ในการตัดสินใจและรวดเร็วยิ่งขึ้นในด้านการบริหาร การจัดการ และการวิเคราะห์ข้อมูลในคลังข้อมูล การที่จะสนับสนุนการตัดสินใจให้ถูกต้องนั้น จะต้องมีข้อมูลที่ถูกต้องทั้งเชิงพื้นที่ และเชิงเวลา รวมถึงการมีต้นทุนที่เหมาะสม (Jarke M., Lenzerini M., Vassiliou Y., & Vassiliadis P., 2003) นอกจากนี้คลังข้อมูล หมายถึง ฐานข้อมูลขนาดใหญ่ขององค์กร หรือหน่วยงาน ซึ่งเก็บรวบรวมข้อมูลจาก

ฐานข้อมูลระบบงานประจำวัน หรือเรียกอีกอย่างว่า operational database และฐานข้อมูลอื่นภายนอกองค์กร หรือเรียกว่า external database โดยข้อมูลที่ถูกจัดเก็บในคลังข้อมูลนั้น มีวัตถุประสงค์ในการนำมาใช้งาน นอกจากนี้ข้อมูลที่รวมอยู่ในคลังข้อมูลสามารถเรียกใช้งานได้จากอินเตอร์เฟสแบบกราฟิกได้โดยตรง (Graphical User Interface: GUI) เพื่อให้พร้อมสำหรับการจัดการข้อมูล และนำข้อมูลที่ได้ไปวิเคราะห์ต่อไป (เศรษฐพงค์ มะลิสุวรรณ, 2553) การทำคลังข้อมูลเป็นการเก็บรวบรวมข้อมูลปริมาณมากตั้งแต่อดีตจนถึงปัจจุบัน ส่วนมากมักพบในงานด้านธุรกิจ ตัวอย่างเช่นการทำคลังข้อมูลสินค้าที่ขายในร้านสะดวกซื้อ วัตถุประสงค์ของการทำคลังข้อมูล เพื่อให้ผู้บริหารตัดสินใจได้ง่าย ว่าร้านสะดวกควรนำสินค้าประเภทไหนวางขายมากที่สุด เพื่อให้ได้กำไรตามเป้าหมายที่ตั้งไว้ ตัวอย่างการทำคลังข้อมูลทางด้านอาชญากรรมเช่น การทำแผนเพื่อเก็บไว้เป็นหลักฐาน จะทำให้ทราบตำแหน่งที่เกิดเหตุ และการทำบันทึกประจำวัน จะเป็นการรวบรวมคดีที่เกิดขึ้นในแต่ละวัน

2.2.2 การทำเหมืองข้อมูล

ได้มีผู้ให้ความหมายเกี่ยวกับการทำเหมือง Hand, Mannila and Smyth (2001) อธิบายความหมายของการทำเหมืองข้อมูลว่า เป็นการวิเคราะห์ข้อมูลจำนวนมากเพื่อหาความสัมพันธ์และการสรุปข้อมูลให้สามารถเข้าใจได้ และเป็นประโยชน์ต่อผู้ทำการเก็บรวบรวมข้อมูล (Hand DJ., Mannila H., & Smyth P., 2001) Cabena, Hadjinian, Stadler, Verhees and Zanasi (1998) ได้ให้ความหมายไว้ว่า การทำเหมืองข้อมูลเป็นการนำวิธีการจาก Machine learning การจดจำรูปแบบ วิธีการทาง สถิติ และฐานข้อมูล เพื่อทำการสกัดข้อมูลจากฐานข้อมูลขนาดใหญ่ได้ (Cabena P., Hadjinian P., Stadler R., Verhees J., & Zanasi A., 1998) ทั้งนี้ การทำเหมืองข้อมูลระบบสารสนเทศทางภูมิศาสตร์เพื่อการวิเคราะห์การลักลอบลำเลียงยาเสพติดนั้นหมายถึง การสกัดข้อมูลที่สำคัญเชิงพื้นที่ที่เก็บรวบรวม และนำไปวิเคราะห์ทางภูมิศาสตร์ เพื่อระบุตำแหน่งของเหตุการณ์การลักกลอบลำเลียงยาเสพติด ตัวอย่างเช่น การเก็บรวบรวมพิกัดที่มีการจับกุมการลักลอบลำเลียงยาเสพติดทั้งหมด นำมารวมกับปัจจัยการใช้ประโยชน์ที่ดิน แล้วนำหลักการของ Machine learning มาใช้ในการดำเนินการให้คอมพิวเตอร์เรียนรู้ความสัมพันธ์ของพิกัดการจับกุมการลักลอบลำเลียงยาเสพติดกับการใช้ประโยชน์ที่ดิน เพื่อวิเคราะห์หาการใช้ประโยชน์ที่ดินที่มีความเสี่ยงต่อการลักลอบลำเลียงมากที่สุด การทำเหมืองข้อมูล เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูลจากเดิมที่มีการจัดเก็บข้อมูลอย่างง่ายๆ มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูล

สารสนเทศมาใช้ได้ วิวัฒนาการของการทำเหมืองข้อมูลได้เริ่มต้นในปีค.ศ. 1960 (ช่วงเวลาที่มีการพัฒนาระบบสารสนเทศทางภูมิศาสตร์) มีการนำข้อมูลมาจัดเก็บอย่างเหมาะสมในอุปกรณ์ที่น่าเชื่อถือและป้องกันการสูญหายได้เป็นอย่างดี ต่อมาในปีค.ศ. 1980 ได้มีการนำข้อมูลที่จัดเก็บมาสร้างความสัมพันธ์ต่อกันในข้อมูลเพื่อประโยชน์ในการนำไปวิเคราะห์ และการตัดสินใจอย่างมีคุณภาพ จากนั้นค.ศ. 1990 มีการรวบรวมข้อมูลขนาดใหญ่โดยครอบคลุมทุกด้านขององค์กรลงในฐานข้อมูลเพื่อช่วยสนับสนุนการตัดสินใจ และได้มีการพัฒนาอย่างต่อเนื่องจนกระทั่งปีค.ศ. 2000 ได้มีการนำข้อมูลจากฐานข้อมูลมาวิเคราะห์และประมวลผล โดยการสร้างแบบจำลองและความสัมพันธ์ทางสถิติ (สิทธิชัย คำคง, 2558)

การทำเหมืองข้อมูลสามารถแบ่งออกได้เป็น 2 ประเภทได้แก่ (1) การสร้างตัวแบบในการทำนาย (Predictive data mining) หรือเรียกอีกอย่างหนึ่งว่าเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised learning) และ (2) การสร้างตัวแบบในการบรรยายหรือการพรรณนา (Descriptive data mining) หรือ เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) การทำเหมืองข้อมูลอธิบายเพิ่มเติมได้ดังนี้ (1) การสร้างตัวแบบในการทำนาย หรือเทคนิคแบบมีผู้สอน คือการนำข้อมูลในอดีตมาสร้างตัวแบบเพื่อทำการทำนายอนาคต จากการใช้ข้อมูลฝึกหัด (Training data) ซึ่งข้อมูลทั้งหมดจะมีคุณสมบัติที่ใช้ในการทำนายอัลกอริทึมประเภทนี้ การสร้างตัวแบบในการทำนายจะมุ่งเน้นในการจำแนกข้อมูลออกเป็นกลุ่มตามคุณสมบัติของข้อมูล ซึ่งถ้าค่าคุณสมบัติของข้อมูลมีค่าไม่ต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่า การจำแนกประเภท หรืออธิบายจากตัวอย่างได้ว่าการแยกปัจจัยที่ไม่ส่งผลกับการลักลอบลำเลียงยาเสพติดออกจากปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด แต่ถ้าค่าคุณสมบัติของข้อมูลมีค่าต่อเนื่อง จะเรียกกระบวนการที่ใช้จำแนกข้อมูลว่า การถดถอย หรือการพยากรณ์ ตัวอย่างเช่น การหาความสัมพันธ์ของข้อมูลที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการจัดกลุ่มของข้อมูลตำแหน่งของการจับกุมเพื่อศึกษาการกระจายตัวของข้อมูลเป็นต้น (2) การสร้างตัวแบบในการบรรยาย หรือการพรรณนา หรือเทคนิคแบบไม่มีผู้สอน คือการนำข้อมูลมาวิเคราะห์กฎความสัมพันธ์ด้วยระบบกฎความสัมพันธ์ (Rules based system) หรือการจัดกลุ่ม (Clustering) จุดมุ่งหมายของเทคนิคแบบไม่มีผู้สอน ตัวอย่างเช่นการทำนายการจัดกลุ่มของปัจจัยที่มีความสัมพันธ์ต่อการลักลอบลำเลียงยาเสพติดกลุ่มเดียวกัน (สายชล สินสมบูรณ์ทอง, 2558) ซึ่งทั้ง 2 ประเภท แบ่งย่อยได้ดัง ตารางที่ 2 ในการศึกษาครั้งนี้จะดำเนินการศึกษาวิเคราะห์หาเทคนิคการทำเหมืองข้อมูลที่เหมาะสมต่อการหาความสัมพันธ์เชิงพื้นที่ และการคาดการณ์การลักลอบลำเลียงยาเสพติดรายละเอียดของการทำเหมืองข้อมูลแต่ละเทคนิคสามารถอธิบายได้ดังนี้

ตารางที่ 2 *ประเภทการทำเหมืองข้อมูล*

เทคนิคการทำเหมืองข้อมูล	วิธีในการวิเคราะห์	ตัวอย่างอัลกอริทึมที่ใช้
เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)	- การหากฎความสัมพันธ์ (Association rule)	Apriori
	- การแบ่งกลุ่มข้อมูล (Clustering)	K-means
เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised learning)	- การประมาณค่าข้อมูล (Regression)	Regression
	- การจำแนกประเภทข้อมูล (Classification)	Decision tree Bayesian network K-Nearest Neighbors Neural Network

2.2.2.1 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) ซึ่งวิธีที่ใช้ในการวิเคราะห์ของเทคนิคการเรียนรู้แบบไม่มีผู้สอน แบ่งเป็น 2 วิธี คือ (1) การหากฎความสัมพันธ์ (Association rule) ซึ่งเป็นกฎที่ใช้ในการทำนายความสัมพันธ์ของสิ่งที่สนใจ และหาค่าความสัมพันธ์ของคุณลักษณะที่มีอยู่ต่อคุณลักษณะอื่น ๆ อัลกอริทึมที่ใช้ส่วนใหญ่จะเป็น Apriori ซึ่งเป็นอัลกอริทึมพื้นฐานที่ใช้ในการหาความสัมพันธ์ของข้อมูลโดยใช้หลักการค้นหาแบบนับทรานแซคชัน Apriori จะทำการสร้าง และตรวจสอบกลุ่มของเหตุการณ์ที่เกิดขึ้นซ้ำทีละชั้น (Ester M., Kriegel H.-P., & Sander J., 1997) ตัวอย่างเช่น การศึกษาความสัมพันธ์ของพารามิเตอร์ในปุ๋ยหมัก เพื่อหาพารามิเตอร์ที่เหมาะสมต่อการย่อยสลาย (Calçada DB., Rezende SO., & Teodoro MS., 2019) และการศึกษาความสัมพันธ์ของปัจจัยที่ส่งผลให้มีการจราจรติดขัด เพื่อตรวจสอบความสัมพันธ์ปัจจัยที่ส่งผลต่อปัญหาจราจร (Xu C., Bao J., Wang C., & Liu P., 2018) และ (2) การแบ่งกลุ่มข้อมูล (Clustering) เป็นแนวคิดที่สำคัญในการรวมกลุ่มสิ่งที่สนใจเข้าด้วยกันด้วยการสังเกตจากสิ่งที่

คล้ายกันนำมาจัดกลุ่มเดียวกัน โดยอัลกอริทึมที่ใช้ส่วนใหญ่ คือ K-means ซึ่งเป็นวิธีในการศึกษาข้อมูลเป็นการหาพาร์ทิชัน หรือช่วงการแบ่งกลุ่มของข้อมูลที่ดีที่สุดของข้อมูลแล้วนำข้อมูลที่แบ่งมารวมกันซึ่งจะเป็นการลดผลรวมของข้อผิดพลาด ทำให้ K-means เป็นที่นิยมในการใช้ในการศึกษาการแบ่งกลุ่ม เพื่อสังเกตการกระจายตัวของข้อมูล (Hamfelt A., Karlsson M., Thierfelder T., & Valkovsky V., 2011) ตัวอย่างเช่น การศึกษาการจัดกลุ่มเพื่อติดตามตรวจสอบกระบวนการทางเคมีของโรงงานอุตสาหกรรม (Thomas MC., Zhu W., & Romagnoli JA., 2018) และการศึกษาการจัดกลุ่มทำการจัดกลุ่มโครงสร้างของดินเพื่อใช้หาปฏิสัมพันธ์ของดิน (Ge Q., Xiong F., Xie L., Chen J., & Yu M., 2019)

2.2.2.2 การเรียนรู้แบบมีผู้สอน (Supervised learning) ซึ่งวิธีที่ใช้ในการวิเคราะห์ของเทคนิคการเรียนรู้แบบมีผู้สอน แบ่งเป็น 2 วิธี (1) การประมาณค่าการถดถอย (Regression) เป็นการทำนายจากข้อมูลตัวเลข เพื่อนำมาประมาณค่าการถดถอยเชิงเส้น ค่าที่ได้จะเป็นค่าเชิงสถิติสามารถนำไปประยุกต์ใช้ในการศึกษาทางด้านสิ่งแวดล้อมได้ เช่น การศึกษาการประมาณค่าการถดถอยวิเคราะห์ปริมาณการกักเก็บน้ำในดิน เพื่อคำนวณปริมาณ และการกักเก็บน้ำตามประเภทของดิน (Nguyen PM. et al., 2017) และการศึกษาการคาดการณ์ปริมาณของหอยสองฝาที่อาศัยอยู่ในทะเล ด้วยเทคนิคการเรียนรู้แบบมีผู้สอนแบบการประมาณค่าถดถอย (Oliveira MM. et al., 2017) และ (2) การจำแนกข้อมูล (Classification) คือ การใช้การจำแนกข้อมูลแต่ละชุดลงในหมวดหมู่ที่ต้องการวิเคราะห์เพื่อเป็นตัวต้นแบบ หรือหมายถึงข้อมูลที่ใช้ในการฝึกสอนคอมพิวเตอร์ ในการสร้างตัวจำแนกที่ใช้ประกอบการคาดการณ์ข้อมูล (Kesavaraj G. & Sukumaran S., 2013) การจำแนกข้อมูลมีอัลกอริทึมด้วยกัน 4 อัลกอริทึม ได้แก่ อัลกอริทึมเครือข่ายแบบเบย์ (Bayesian Network) อัลกอริทึมต้นไม้การตัดสินใจ (Decision tree) อัลกอริทึม K-Nearest Neighbors (K-NN) และอัลกอริทึม Neural Network ตัวอย่างของการจำแนกข้อมูลในรูปแบบการเรียนรู้แบบมีผู้สอนทางด้านสิ่งแวดล้อมและสุขภาพ ได้แก่ การศึกษาการคาดการณ์พื้นที่เสี่ยงต่อการเกิดโรคมาลาเลียด้วยการจำแนกข้อมูลโดยการใช้เครือข่ายเบย์ ได้ผลมาซึ่งพื้นที่เสี่ยงต่อโรคมาลาเรียในพื้นที่ภาคเหนือของประเทศไทย (Peter H. et al., 2018) การศึกษาที่เป็นการใช้การจำแนกข้อมูลด้วยอัลกอริทึมต้นไม้การตัดสินใจในการศึกษาความสัมพันธ์ของการเปลี่ยนแปลงระดับน้ำใต้ดิน ปริมาณน้ำฝน และระดับอ่างเก็บน้ำ เพื่อนำความสัมพันธ์ของปัจจัยทั้งสามสร้างแบบจำลองในการคาดการณ์ดินถล่มในประเทศจีน (Zhao Y., Li Y., Zhang L., & Wang Q., 2016) ในการศึกษาครั้งนี้มุ่งเน้นที่เทคนิคการเรียนรูแบบมีผู้สอนวิธีการจำแนกข้อมูลด้วยอัลกอริทึมต้นไม้การตัดสินใจ และวิธี Naïve Bayes ซึ่ง

Naïve Bayes เป็นวิธีวิเคราะห์ Bayesian Network แบบง่ายในการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล สามารถอธิบายรายละเอียดของอัลกอริทึมต้นไม้การตัดสินใจ และอัลกอริทึม Naïve Bayes ได้ดังนี้

ต้นไม้การตัดสินใจ เป็นกระบวนการของการตัดสินใจที่มีรูปแบบในลักษณะของกราฟต้นไม้ที่ประกอบด้วยโหนดเริ่มต้น และเส้นความสัมพันธ์ โดยโหนดเริ่มต้นภายในแทนลักษณะประจำที่นำมาใช้แยกกลุ่มข้อมูล เส้นเชื่อมโยงแทนค่า หรือผลของการทดสอบ โหนดถัดมาจะแทนค่าของคลาสหรือการกระจายของคลาส (Quinlan JR., 1986) ตัวอย่างเช่น ฤดูหนาวมีการลักลอบลำเลียงยาเสพติดช่วงเวลาเช้า 100 ครั้ง จุดยอดจะเป็นฤดู กับช่วงเวลา ก้านที่สร้างจะเชื่อมโยงจากฤดูถึงช่วงเวลา การสร้างต้นไม้การตัดสินใจมีขั้นตอนในการสร้างประกอบด้วย 2 ขั้นตอน ได้แก่ ขั้นตอนที่ 1 เป็นการสร้างต้นไม้เริ่มจากการสร้างโหนดเริ่มต้นที่มีการฝึกหัดข้อมูลทั้งหมด โหนดจะสามารถกำหนดโหนดเริ่มต้นได้ด้วยค่า Information gain หรือค่า Gini index ซึ่งค่า Information gain และค่า Gini index เป็นค่าที่แสดงถึงความสำคัญของโหนดแต่ละโหนด โดยต้องทราบถึงค่าเอนโทรปีที่เป็นค่าการกระจายตัวของข้อมูลหาได้จากสมการที่ 1 ซึ่งค่า Information gain และค่า Gini index มีค่าตั้งแต่ 0.000 ถึง 1.000 ยิ่งค่าเข้าใกล้ 1.000 มาก แสดงว่าโหนดนั้นมีความสำคัญมากที่สุด ค่า Information gain และค่า Gini index คำนวณได้จากสมการที่ 2 และสมการที่ 3 ตามลำดับ

$$\text{entropy}(C_n) = -p(C_n) - \log(C_n) \qquad \text{สมการที่ 1}$$

โดยที่

$\text{entropy}(c_n)$ คือ ค่าเอนโทรปีของปัจจัยที่ n เช่น ค่าเอนโทรปีของปัจจัยมูลค่ายาเสพติดมีค่าค่อนข้างสูง เนื่องจากปัจจัยมูลค่าในแต่ละพื้นที่มีมูลค่าที่แตกต่างกันค่อนข้างมาก

$p(c_n)$ คือ ค่าความน่าจะเป็นของปัจจัยที่ n เช่น ความน่าจะเป็นของปัจจัยมูลค่ายาเสพติดต่อพื้นที่แต่ละพื้นที่

$$IG = \text{entropy}(\text{parent}) - \sum p(C_n) \times \text{entropy}(C_n) \qquad \text{สมการที่ 2}$$

โดยที่

IG คือ ค่า Information gain

entropy(parent) คือ ค่าเอนโทรปีของโหนดเริ่มต้น เช่น จากการศึกษาโหนดเริ่มต้นคือ ตำแหน่งพิกัดการจับกุม คำนวณได้จากสมการที่ 1

entropy(c_1) คือ ค่าเอนโทรปีของปัจจัยที่ 1 เช่น ปัจจัยฤดูกาล ปัจจัยรัศมีจากจุดตรวจ และปัจจัยการใช้ประโยชน์ที่ดิน เป็นปัจจัยที่เชื่อมโยงกับโหนดเริ่มต้นโดยตรง คำนวณได้จากสมการที่ 1

entropy(c_n) คือ ค่าเอนโทรปีของปัจจัยที่ n เช่น ปัจจัยประเภทของยาเสพติดเชื่อมโยงกับปัจจัยการใช้ประโยชน์ที่ดิน โดยมีโหนดเริ่มต้น คือ ตำแหน่งพิกัดการจับกุม คำนวณได้จากสมการที่ 1

$p(c_1)$ คือ ค่าความน่าจะเป็นของปัจจัยที่ 1 เช่น ค่าความน่าจะเป็นของปัจจัยฤดูกาลที่ส่งผลต่อมูลค่าของยาเสพติดในแต่ละพื้นที่

$p(c_n)$ คือ ค่าความน่าจะเป็นของปัจจัยที่ n เช่น ค่าความน่าจะเป็นของปัจจัยประเภทของยาเสพติดที่ส่งผลต่อมูลค่าของยาเสพติดในแต่ละพื้นที่

$$\text{Gini index} = 1 - \sum(p_i)^2 \qquad \text{สมการที่ 3}$$

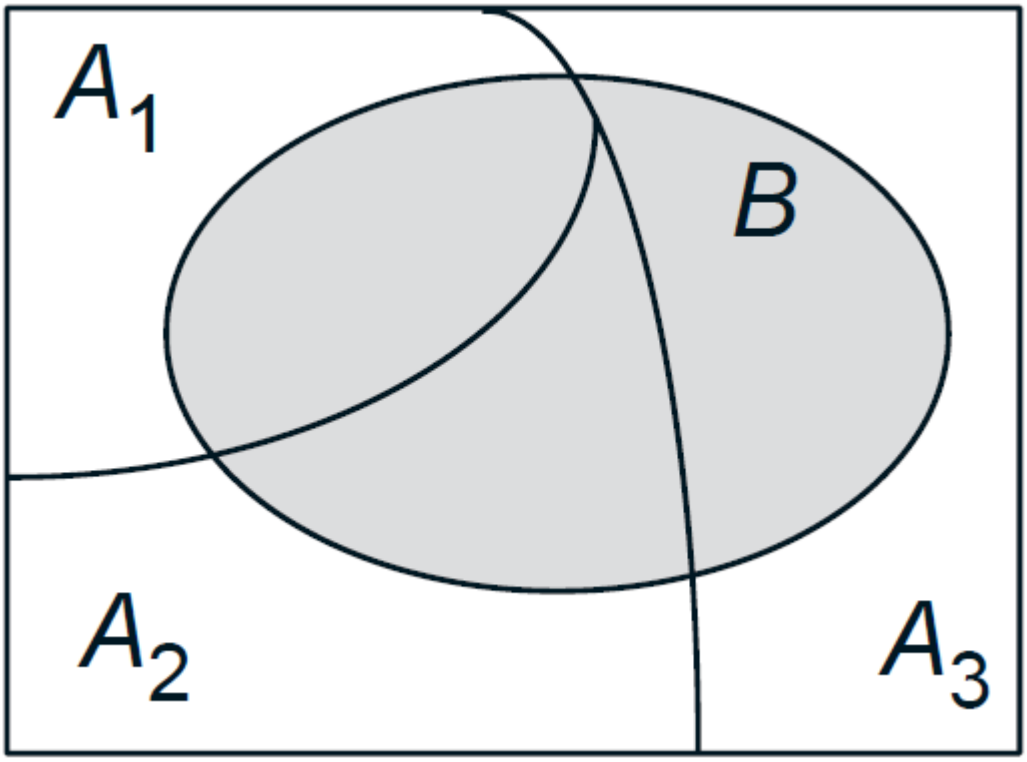
โดยที่

p_i คือ ค่าความน่าจะเป็นของปัจจัยที่ i เช่น การหาค่า Gini index ของปัจจัยรัศมีจากจุดตรวจต้องใช้ค่าความน่าจะเป็นของปัจจัยรัศมีจากจุดตรวจที่ส่งผลต่อมูลค่าของยาเสพติดในแต่ละพื้นที่

อัลกอริทึมที่ใช้ในการทำงานของต้นไม้ตัดสินใจที่เป็นพื้นฐานคือ Iterative Dichotomiser 3 (ID3) ซึ่งเป็นวิธีที่ใช้ในการแบ่งกลุ่มเพื่อให้สามารถสร้างแผนภูมิต้นไม้ที่ดีที่สุด (Quinlan JR., 1986) นอกจาก ID3 ยังมีอัลกอริทึมอื่น เช่น Classification And Regression Tree (CART) และ CHi-squared Automatic Interaction Detector (CHAID) เป็นต้น แนวคิดของID3 คือ การหาตัวแบ่งการจำแนกประเภทข้อมูล (Classifier) ที่ดีที่สุด ด้วยหลักการทฤษฎีของข้อมูล และการพยายามลดจำนวนครั้งใน

การตัดสินใจให้น้อยที่สุด ดังนั้นในการเลือกคุณลักษณะ เพื่อใช้ในการตัดสินใจของแต่ละโหนดจะพยายามเลือกคุณลักษณะที่จะให้สารสนเทศเพิ่มเติมสูงสุด (Highest information gain) จากการใช้เอนโทรปี (Entropy) หรือลักษณะที่บ่งชี้ระดับการสุ่ม ให้เป็นตัววัดปริมาณข้อมูลก่อน และหลังการจำแนก ถ้าเอนโทรปีมีค่าต่ำลง แสดงว่าข้อมูลชุดนั้นมีการกระจัดกระจายน้อยลง หรือข้อมูลถูกจำแนกจนเป็นระเบียบมากขึ้น ทำให้ได้สารสนเทศจากข้อมูลชุดนั้นมากขึ้นด้วย (Quinlan JR., 1986) ต่อมาได้มีการพัฒนาอัลกอริทึม ID3 เป็น อัลกอริทึม C4.5 ซึ่งให้ผลลัพธ์เป็นต้นไม้การตัดสินใจที่เป็นวิธีการสร้างต้นไม้เช่นเดียวกับ ID3 แต่เพิ่มกระบวนการปรับปรุงต้นไม้เพื่อให้มีลักษณะที่ดียิ่งขึ้น (Quinlan JR., 1994) ในปัจจุบันอัลกอริทึม C5.0 เป็นอัลกอริทึมล่าสุดของวิธีต้นไม้การตัดสินใจ ซึ่งเป็นอัลกอริทึมที่ใช้หน่วยความจำน้อยกว่า และแสดงผลการวิเคราะห์ที่รวดเร็วกว่า C4.5 แต่การศึกษาครั้งนี้ได้นำอัลกอริทึม C4.5 มาใช้ในการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด เนื่องจากการศึกษาครั้งนี้ได้ใช้ซอฟต์แวร์สำเร็จรูป ซึ่งมีเพียงอัลกอริทึม C4.5 เป็นอัลกอริทึมล่าสุดของซอฟต์แวร์ หากต้องการใช้ C5.0 จำเป็นต้องเขียนโปรแกรมภาษา R เพื่อเรียกใช้งานในซอฟต์แวร์ R package

Naïve Bayes เป็นการจำแนกข้อมูลด้วยหลักการจากทฤษฎีของเบย์ (Bayes theorem) ซึ่งมีนักคณิตศาสตร์ชื่อโทมัส เบย์ เป็นคนคิดค้น Naïve Bayes เป็นการใช้แบบจำลองทางคณิตศาสตร์ที่แสดงชุดของตัวแปรที่มีความเป็นอิสระต่อกัน สามารถอธิบายรายละเอียดของหลักการทฤษฎีของเบย์ได้ว่า ทฤษฎีของเบย์เป็นการใช้หลักความน่าจะเป็นแบบมีเงื่อนไขในการพัฒนา และศึกษาสมมติฐานที่ถูกต้องที่สุด โดยใช้ความรู้ในอดีตที่ได้รวบรวมไว้มาตั้งสมมติฐานสมมติฐาน และนำมาวิเคราะห์หาสมมติฐานที่ถูกต้องที่สุด (นพมาศ ปักเข็ม, 2558) จากภาพที่ 4 อธิบายได้ว่ามีเหตุการณ์ B และ A เกิดขึ้น โดยเหตุการณ์ A เป็นเหตุการณ์ที่เกิดขึ้นไม่พร้อมกัน ซึ่งเหตุการณ์ A สามารถเกิดขึ้นได้มากกว่า 3 เหตุกาณ์ และ B เป็นเหตุการณ์ที่เกิดขึ้นเป็นปกติ การคำนวณจะต้องมีทั้งเหตุการณ์ B และA_i (i=1,2,3,...,K) โดยจะสามารถคำนวณหาความน่าจะเป็นแบบมีเงื่อนไขของเหตุการณ์หนึ่งใน A_i เมื่อเหตุการณ์ B เกิดขึ้นแล้วได้ดังสมการที่ 4 (Bertsekas DP. & Tsitsiklis JN., 2008)



ภาพที่ 4 *ความสัมพันธ์ของเหตุการณ์ B และ A ที่เกิดขึ้นไม่พร้อมกัน*

ที่มา: Bertsekas DP. and Tsitsiklis JN.,2008

$$P(Ai|B) = [P(B|Ai) \times P(Ai)]/P(B)$$ สมการที่ 4

โดยที่

P คือ ความน่าจะเป็นที่คำนวณได้

A_i คือ เหตุการณ์ที่เกิดขึ้น i จำนวนเหตุการณ์

B คือ เหตุการณ์ที่เกิดขึ้นแล้ว

ตัวอย่างเช่น ให้ B แทนเหตุการณ์ลักลอบลำเลียงยาเสพติด และ A แทนเหตุการณ์ที่จะเกิดขึ้น หรือเหตุการณ์ที่เกิดขึ้นพร้อมกับ B เช่น เกิดการปะทะระหว่างเจ้าหน้าที่กับผู้ลักลอบ และเกิดความเป็นอันตรายต่อผู้ที่อยู่อาศัยในเส้นทางการลักลอบลำเลียงยาเสพติด เป็นต้น

ในการนำหลักการทฤษฎีของเบย์ไปประยุกต์ใช้เพื่อการคาดการณ์มีตัวอย่างดังนี้ การศึกษาจุดความร้อน (Hotspot) ของอาชญากรรมโดยใช้วิธีการสร้างแบบจำลองเบย์เชิงพื้นที่ กรณีการศึกษาเกี่ยวกับอาชญากรรมรุนแรงในมหานคร เขตโตรอนโต (Law J., Quick M., & Chan PW., 2015) นอกจากนี้ยังมีการศึกษาที่เกี่ยวกับ จุดความร้อนอีกงานวิจัยหนึ่ง ซึ่งเป็นการศึกษาความ

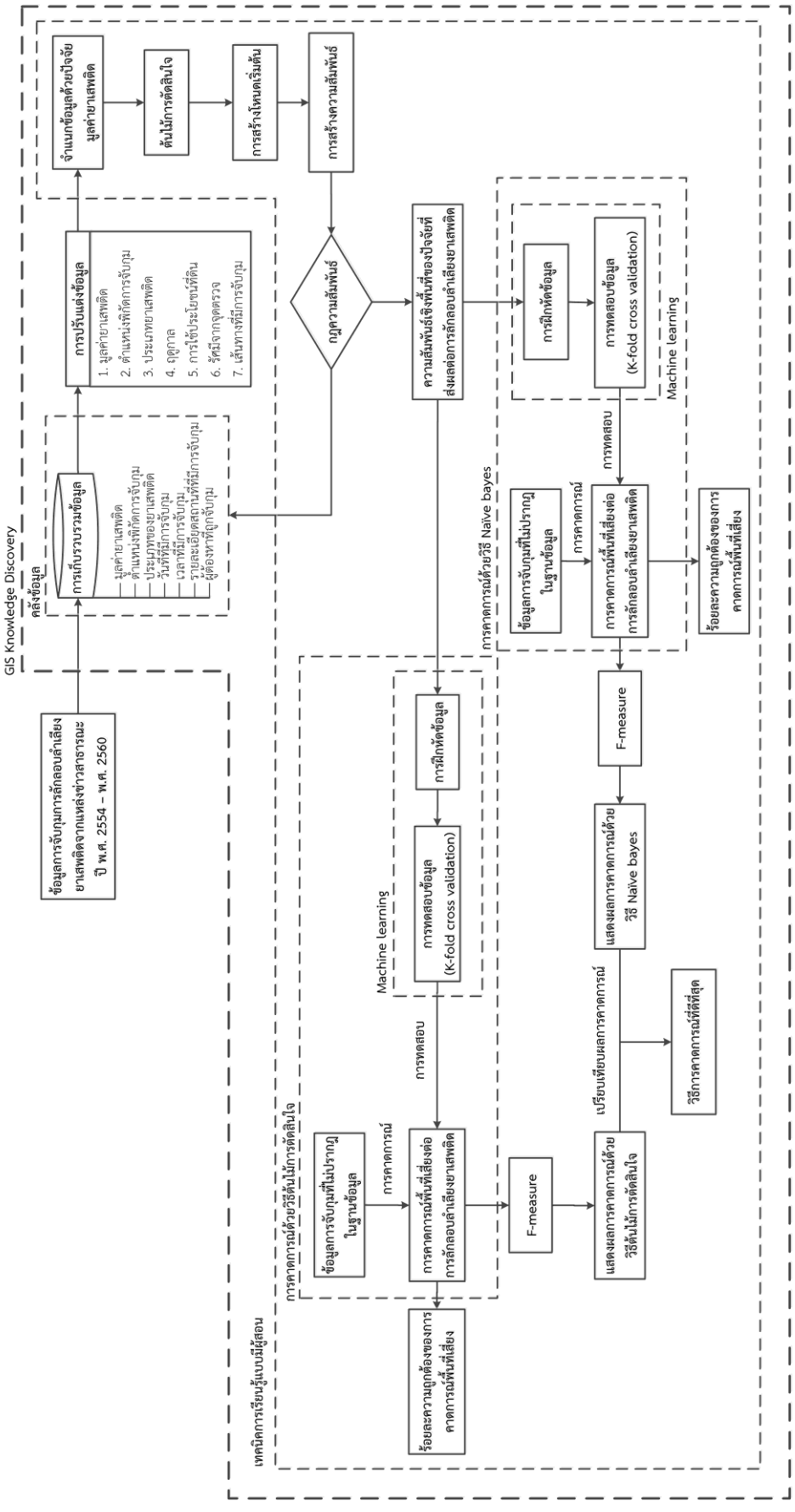
แปรปรวนของพื้นที่และเวลาที่ส่งผลต่อความเสี่ยงของอาชญากรรมประเภทการลักทรัพย์ด้วยหลักการของเบย์ (Li G., Haining R., Richardson S., & Best N., 2014) และนอกจากนำหลักการของเบย์มาทำการศึกษาจุดความร้อนแล้วยังมีการนำไปประยุกต์ใช้ร่วมกับวิธีระบาดวิทยาเชิงพื้นที่ในการศึกษาอาชญากรรมรุนแรงในซานอัลโตนิโอ (Spark CS., 2011)

2.2.3 การคาดการณ์ และการแสดงผลการทำเหมืองข้อมูล

การแสดงผลการทำเหมืองข้อมูล คือ กระบวนการที่ใช้ในการแสดงผลจากการทำเหมืองข้อมูล ซึ่งรูปแบบการแสดงผลจะขึ้นอยู่กับวัตถุประสงค์ของการศึกษา ในการศึกษาครั้งนี้ได้นำซอฟต์แวร์ RapidMiner และซอฟต์แวร์ Weka มาใช้ในการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล และแสดงผลของการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล โดยที่ซอฟต์แวร์ RapidMiner ถูกใช้ในวัตถุประสงค์ที่ 1 เพื่อศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ ผลที่ได้ออกมาในรูปแบบกราฟความสัมพันธ์ในรูปต้นไม้ ส่วนวัตถุประสงค์ที่ 2 ได้ใช้ซอฟต์แวร์ Weka ในการแสดงผลของการศึกษาวิธีที่ใช้ในการทำเหมืองข้อมูลที่เหมาะสมระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ต่อการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด ซึ่งผลจาก Weka แสดงในรูปแบบค่าทางสถิติเชิงตัวเลข (ดูผลการศึกษาในตารางที่ 7 ถึงตารางที่ 15)

บทที่ 3
วิธีการดำเนินงานวิจัย

การศึกษาครั้งนี้ได้ทำการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วย Knowledge Discovery from Database ซึ่งการศึกษาครั้งนี้ได้แสดงผลการศึกษาในรูปแบบความสัมพันธ์เชิงพื้นที่ของปัจจัยที่มีผลต่อการลักลอบลำเลียงยาเสพติดและการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดในพื้นที่ อำเภอแม่จัน อำเภอแม่สาย และอำเภอแม่ฟ้าหลวง จังหวัดเชียงราย อำเภอฝาง อำเภอแม่อาย อำเภอเชียงดาว อำเภอไชยปราการ และอำเภอเวียงแหง จังหวัดเชียงใหม่ โดยข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติดรวบรวมได้จากหนังสือพิมพ์ท้องถิ่น หนังสือพิมพ์ระดับชาติ สื่อออนไลน์ อินเตอร์เน็ต และรายงานการประชุมสถานการณ์ยาเสพติด ข่าวการจับกุมการลักลอบลำเลียงยาเสพติดดังกล่าวทำให้ทราบถึงข้อมูลคุณลักษณะต่าง ๆ เช่น เพศ ปริมาณยาเสพติด ตำแหน่งที่จับได้ และวันเวลา ในการสร้างคลังข้อมูลจึงได้รวบรวมข้อมูลที่กล่าวมาข้างต้นในรูปแบบรายปี ตั้งแต่ปี พ.ศ. 2554 จนถึงปี พ.ศ. 2560 ส่วนการคาดการณ์ได้ใช้ข้อมูลในปี พ.ศ. 2561 ซึ่งข้อมูลที่เก็บรวบรวมในรูปคลังข้อมูลได้ถูกนำมาปรับแต่งให้อยู่ในรูปแบบตารางคุณลักษณะ (.dbf) เมื่อปรับแต่งข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดเรียบร้อยแล้ว ทำการจำแนกข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดด้วยมูลค่าของยาเสพติดที่จับกุมได้ และนำข้อมูลที่ผ่านการจำแนกเข้าสู่กระบวนการการทำเหมืองข้อมูลโดยใช้วิธีต้นไม้การตัดสินใจในการหาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่มีผลต่อการลักลอบลำเลียงยาเสพติด เมื่อได้ความสัมพันธ์เชิงพื้นที่ของปัจจัยที่มีผลต่อการลักลอบลำเลียงยาเสพติดแล้ว นำความสัมพันธ์เชิงพื้นที่ดังกล่าวมาคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด และเปรียบเทียบวิธีที่ใช้ในการการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ทั้งนี้การศึกษารูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วย Knowledge Discovery from Database ประกอบไปด้วย 4 ขั้นตอน คือ (1) ขั้นตอนการรวบรวมข้อมูล (2) ขั้นตอนการปรับข้อมูลให้อยู่ในรูปฐานข้อมูลระบบสารสนเทศทางภูมิศาสตร์ (3) ขั้นตอนการวิเคราะห์ และ (4) ขั้นตอนการแสดงผล โดยมีวิธีการศึกษาตามกรอบแนวความคิดการวิจัยดังภาพที่ 3.1 และมีรายละเอียดดังต่อไปนี้



ภาพที่ 5 ขั้นตอนการดำเนินงานวิจัย

3.1 ข้อมูล

ข้อมูลที่ใช้ในการดำเนินการวิจัยครั้งนี้ประกอบด้วยข้อมูลการเกิดเหตุการณ์จับกุมการลักลอบลำเลียงยาเสพติด ขอบเขตการปกครอง เส้นทางคมนาคม และการใช้ประโยชน์ที่ดิน ข้อมูลเหล่านี้บางส่วนได้รับความอนุเคราะห์จากการวิจัยเรื่อง การคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยสถิติภูมิศาสตร์ (Dechsiri B., 2017) โดยมีรายละเอียดดังตารางที่ 3

ตารางที่ 3 *ข้อมูลที่ใช้ในการศึกษา*

ข้อมูล	รายละเอียดข้อมูล	แหล่งข้อมูล / ปี
ข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติด	ข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติดได้จากการจำแนกข้อมูลข่าว รูปแบบ (.xlsx)	แหล่งข่าวสาธารณะ / ปี พ.ศ. 2554 ถึงปี พ.ศ. 2560
ขอบเขตการปกครอง	ขอบเขตการปกครองในพื้นที่อำเภอแม่จัน อำเภอแม่สาย และอำเภอแม่ฟ้าหลวง จังหวัดเชียงราย อำเภอฝาง อำเภอแม่อาย อำเภอเชียงดาว อำเภอไชยปราการ และอำเภอเวียงแหง จังหวัดเชียงใหม่ มาตราส่วน 1:25,000 รูปแบบ (.shp)	กรมพัฒนาที่ดิน / ปี พ.ศ. 2558 (จังหวัดเชียงใหม่) และปี พ.ศ. 2555 (จังหวัดเชียงราย)
เส้นทางคมนาคม	ข้อมูลทางหลวงแผ่นดิน ทางหลวงชนบท และถนนท้องถิ่น มาตราส่วน 1:25,000 รูปแบบ (.shp)	สำนักปลัดกระทรวงคมนาคม / ปี พ.ศ. 2558
การใช้ประโยชน์ที่ดิน	การแบ่งการใช้ประโยชน์ที่ดิน มาตราส่วน 1:25,000 รูปแบบ (.shp)	กรมพัฒนาที่ดิน / ปี พ.ศ. 2558 (จังหวัดเชียงใหม่) และปี พ.ศ. 2555 (จังหวัดเชียงราย)

3.2 อุปกรณ์ที่ใช้ในการวิจัย

อุปกรณ์ที่ใช้ดำเนินการวิจัยประกอบด้วย (1) ซอฟต์แวร์การวิเคราะห์ข้อมูล RapidMiner Studio (2) ซอฟต์แวร์การวิเคราะห์ข้อมูล Weka 3.8.4 (3) ซอฟต์แวร์ระบบสารสนเทศทางภูมิศาสตร์ ArcMap 10.3 (4) ซอฟต์แวร์ Microsoft Excel 2016 (5) ซอฟต์แวร์ระบบพิกัดตำแหน่งทางภูมิศาสตร์ Google Earth Pro 7.3 และ (6) คอมพิวเตอร์ระบบปฏิบัติการ Window10 Pro Intel® Core™ i5-7200U RAM 4 GB

3.3 วิธีการศึกษา

วิธีการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ประกอบด้วย 3 ส่วน ดังนี้ (3.3.1) การรวบรวมข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติด (3.3.2) ขั้นตอนการปรับข้อมูลให้อยู่ในรูปฐานข้อมูลระบบสารสนเทศทางภูมิศาสตร์ และ (3.3.3) การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด แสดงขั้นตอนการศึกษาดังภาพที่ 5 โดยมีรายละเอียดดังนี้

มหาวิทยาลัยศิลปากร

3.3.1 การรวบรวมข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติด

ข้อมูลที่ใช้ในการศึกษาครั้งนี้ประกอบด้วย (1) ข้อมูลทุติยภูมิ ได้แก่ ข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติดจากหนังสือพิมพ์ท้องถิ่น หนังสือพิมพ์ระดับชาติ สื่อออนไลน์อินเตอร์เน็ต และรายงานการประชุมสถานการณ์ยาเสพติด ข่าวการจับกุมการลักลอบลำเลียงยาเสพติดทำให้ได้ทราบถึงข้อมูลตำแหน่งพิกัดการจับกุมในระบบพิกัด UTM WGS 1984 ซึ่งข้อมูลพิกัดการจับกุมได้จากการหาตำแหน่งผ่านซอฟต์แวร์ Google Earth Pro วันที่ที่มีการจับกุม เวลาที่มีการจับกุม รายละเอียดเกี่ยวกับสถานที่ที่มีการจับกุม ผู้ต้องหาที่ถูกจับกุม ประเภทของยาเสพติด และมูลค่ายาเสพติด โดยการรวบรวมข้อมูลที่กล่าวมาข้างต้น เป็นขั้นตอนของการจัดทำคลังข้อมูลในรูปแบบรายปี ตั้งแต่ปี พ.ศ. 2554 จนถึงปี พ.ศ. 2560 ตัวอย่างข้อมูลแสดงดังตารางที่ 4 (2) ข้อมูลปฐมภูมิ ที่รวบรวมได้แก่ ข้อมูลที่ได้จาการสัมภาษณ์เจ้าหน้าที่ที่ทำงานในพื้นที่ศึกษา การสัมภาษณ์เป็นการสัมภาษณ์เกี่ยวกับคุณลักษณะที่ผู้กระทำความผิดทำการลักลอบลำเลียงยาเสพติด คุณลักษณะที่ได้สอบถามไปได้แก่ ฤดูกาล การใช้ประโยชน์ที่ดิน และเส้นทางที่ส่วนใหญ่ใช้ในการลักลอบลำเลียง ข้อมูลเชิงพื้นที่ในรูปแบบเวกเตอร์ที่ได้รวบรวมมาจะเป็นข้อมูลการใช้ประโยชน์ที่ดิน ข้อมูลเส้นทางถนน และข้อมูลขอบเขตการปกครอง

มหาวิทยาลัยศิลปากร

ตารางที่ 4 *ตัวอย่างข้อมูลที่รวบรวมจากข่าวการจับกุมยาเสพติด*

Value	X, Y	Exhibit	Date	Time	Locale	Offender
16,500,000	510238, 2186339	D_1	2 พฤษภาคม 2554	12.00	สวนลิ้นจี่ บ.สันมะกอกหวาน หมู่ 9 ต.แม่งอน อ.ฝาง	ขบวนการค้ายาเสพติด ชาวจีนฮ่อ จำนวน 3 คน
1,716,264	514246, 2172072	D_6	25 สิงหาคม 2554	13.00	ด่านตรวจผาหงส์ ต.ศรีดงเย็น อ.ไชยปราการ	พลทหารนเรศ คำสิงห์ อายุ 23 ปี
1,250,000	592903, 2253582	D_1	20 กุมภาพันธ์ 2554	13.00	ถนนเลี่ยงเมืองทางเชื่อมต่อระหว่าง บ.น้ำจำ บ.สันผักฮี้ หมู่ 5 ต.โป่งผา อ.แม่สาย	นายโฮ เอง ชอว์ (Ho eng chaw)
530,000	589147, 2227353	D_1	17 ธันวาคม 2554	00.14	ห้องพักธารทิพย์ วิลล่า ต.แม่จัน อ.แม่จัน เชียงราย	นายไก่ (ไม่ทราบนามสกุล) อายุ 18 ปี

ที่มา: Desiri (2554 - 2559)

หมายเหตุ Value คือ มูลค่าของยาเสพติดที่จับกุมได้จากการลักลอบลำเลียงยาเสพติด

X,Y คือ ตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติด

Exhibit คือ ประเภทของสารเสพติดโดยที่ (D_1) ยาบ้า (D_2) ไอซ์ (D_3) ฝิ่น (D_4) กระท่อม (D_5) มีของกลางมากกว่า 1 ชนิด (D_6) เฮโรอีน และ (D_7) สารเคมีที่ใช้ในการผลิตยาเสพติด

Date คือ วัน เดือน ปี ที่มีการจับกุมการลักลอบลำเลียงยาเสพติด

Time คือ เวลาที่มีการจับกุมการลักลอบลำเลียงยาเสพติด

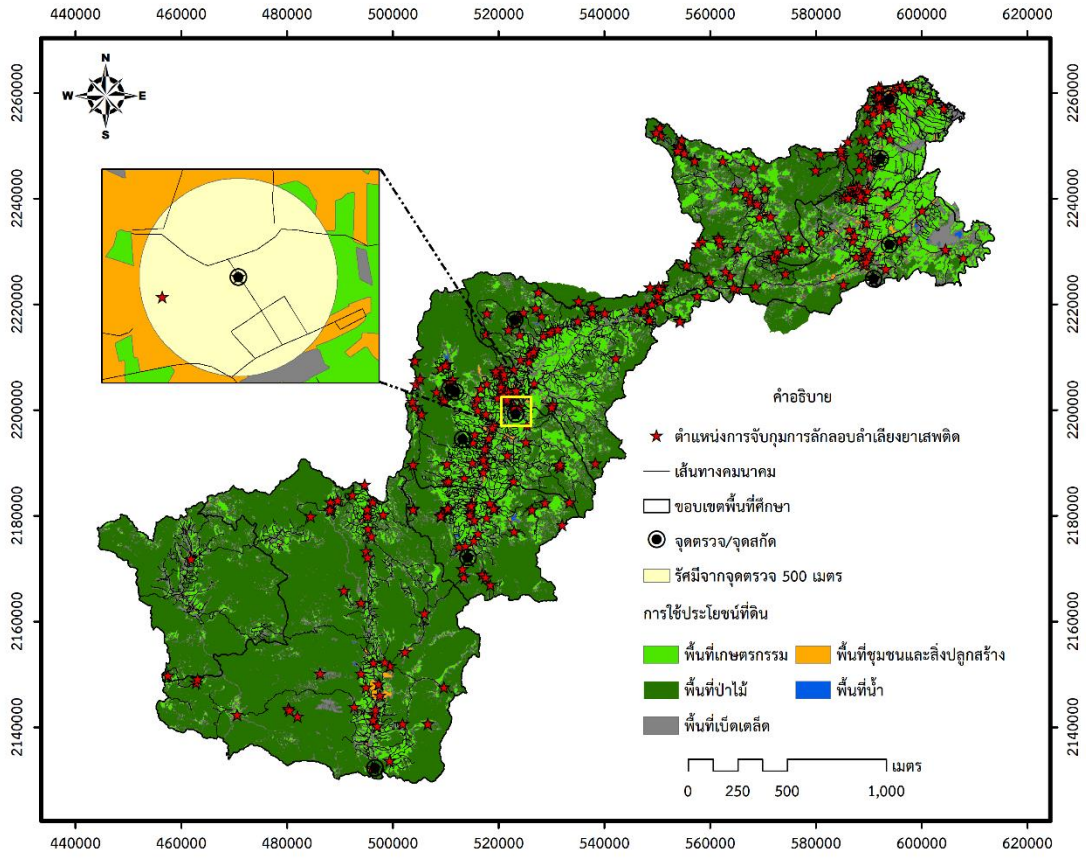
Locale คือ รายละเอียดสถานที่ที่สามารถจับกุมได้

Offender คือ ชื่อผู้กระทำความผิด

3.3.2 ขั้นตอนการปรับข้อมูลให้อยู่ในรูปฐานข้อมูลระบบสารสนเทศทางภูมิศาสตร์

ข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดที่เก็บรวบรวมได้ถูกทำการปรับแต่งข้อมูลให้อยู่ในรูปแบบตาราง (.dbf) โดยที่จำแนกปัจจัยเชิงพื้นที่ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดตามคุณลักษณะได้ดังนี้ (1) Value คือ มูลค่ายาเสพติดที่จับกุมได้ ซึ่งจากการรวบรวมข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติด พบว่าการกระจายของข้อมูลที่เป็นมูลค่าค่อนข้างมาก จึงได้มีการดัดแปลงด้วยการจัดกลุ่มช่วงของมูลค่ายาเสพติด กลุ่มของมูลค่ายาเสพติดที่จัดได้แบ่งออกเป็น 5 ช่วง ดังนี้ ช่วงที่ 1 มีมูลค่าตั้งแต่ 1 บาท ถึง 1 ล้านบาท ช่วงที่ 2 มีมูลค่าตั้งแต่ 1 ล้านบาท ถึง 10 ล้านบาท ช่วงที่ 3 มีมูลค่าตั้งแต่ 10 ล้านบาท ถึง 100 ล้านบาท ช่วงที่ 4 มีมูลค่าตั้งแต่ 100 ล้านบาท ถึง 1,000 ล้านบาท และช่วงที่ 5 มีมูลค่าตั้งแต่ 1,000 ล้านบาทขึ้นไป โดยที่มูลค่ายาเสพติดจะเป็นปัจจัยที่ใช้ในการจำแนกช่วงของข้อมูลการลักลอบลำเลียงยาเสพติดที่ใช้ในการทำเหมืองข้อมูลการลักลอบลำเลียงยาเสพติด (2) Geographic location คือ ตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติด โดยทำการจัดกลุ่มพิกัดดังกล่าวตามการกระจายของข้อมูลข่าวการลักลอบลำเลียงยาเสพติดออกเป็น 5 กลุ่ม แบ่งตามอำเภอได้ดังนี้ อำเภอแม่จัน และอำเภอแม่สาย (L_1) อำเภอเชียงดาว และอำเภอเวียงแหง (L_2) อำเภอไชยปราการ (L_3) อำเภอแม่อาย และอำเภอฝาง (L_4) และกลุ่มสุดท้ายอำเภอแม่ฟ้าหลวง (L_5) (3) Exhibit คือ ประเภทของยาเสพติด จากการรวบรวมข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดทั้งหมด ข้อมูลเหล่านี้ได้ทำการจำแนกยาเสพติดเป็น 7 ประเภท ดังนี้ ยาบ้า (D_1) ไอซ์ (D_2) ฝิ่น (D_3) กระท่อม (D_4) เฮโรอีน (D_5) สารตั้งต้นในการผลิต (D_6) และยาเสพติดที่มากกว่า 1 ชนิด (D_7) (4) Season คือ ฤดูกาล ที่ถูกทำการปรับแต่งข้อมูลเป็น 3 ช่วงฤดูกาลได้แก่ ฤดูร้อน (Summer) ที่อยู่ในช่วงเดือนกุมภาพันธ์ถึงเดือนพฤษภาคม ฤดูฝน (Rainy) ที่อยู่ในช่วงเดือนมิถุนายนถึงเดือนตุลาคม และฤดูหนาว (Winter) ที่อยู่ในช่วงเดือนพฤศจิกายนถึงเดือนมกราคม (5) Land use คือ การใช้ประโยชน์ที่ดิน ปัจจัยการใช้ประโยชน์ที่ดินเป็นปัจจัยที่เกิดขึ้นจากการซ้อนทับข้อมูลระหว่างตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติดกับการใช้ประโยชน์ที่ดินที่ได้รับจากกรมพัฒนาที่ดิน โดยใช้ข้อมูลการใช้ประโยชน์ที่ดินของจังหวัดเชียงใหม่ปีพ.ศ. 2558 และจังหวัดเชียงรายปีพ.ศ. 2555 การใช้ประโยชน์ที่ดินถูกแบ่งเป็นระดับที่ 1 ซึ่งคือ พื้นที่เกษตรกรรม (Agriculture) พื้นที่ชุมชนและสิ่งปลูกสร้าง (Urban) พื้นที่ป่าไม้ (Forest) พื้นที่น้ำ (Water) และพื้นที่เบ็ดเตล็ด (Miscellaneous) (6) Radius คือ รัศมีจากจุดตรวจ 500 เมตร เนื่องจากจุดตรวจส่วนใหญ่ที่ได้รวบรวมจากข่าวการจับกุมเป็นพื้นที่ชุมชน จึงกำหนดให้รัศมีเฉลี่ยอยู่ที่ 500 เมตร จากจุดตรวจ

เพื่อให้ครอบคลุมพื้นที่ชุมชนในบริเวณนั้น โดยที่การจับกุมการลักลอบลำเลียงยาเสพติดที่อยู่ในรัศมี 500 เมตร ให้ตัวแปรเป็น T จากนั้นนำปัจจัยเชิงพื้นที่ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดเข้าสู่ซอฟต์แวร์ ArcMap เพื่อทำการรวมชั้นข้อมูลของปัจจัยเชิงพื้นที่ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดแสดงดังภาพที่ 6

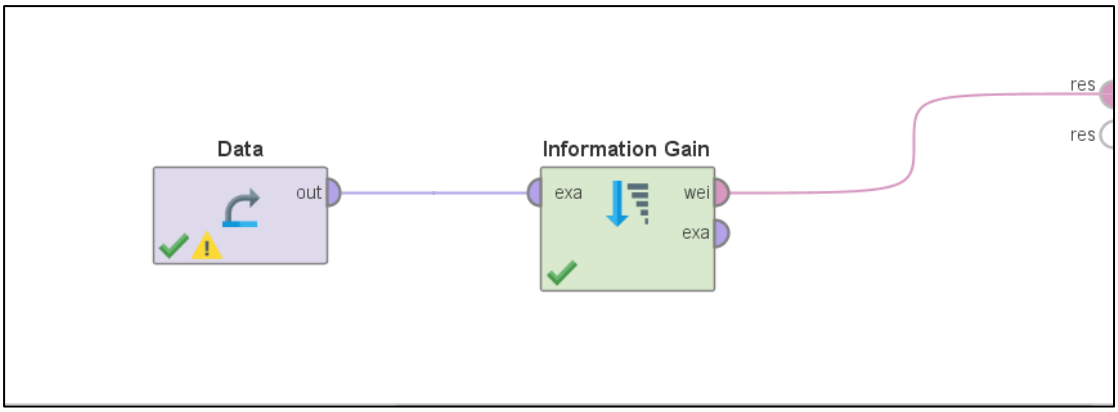


ภาพที่ 6 *แผนที่ปัจจัยเชิงพื้นที่ที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด*

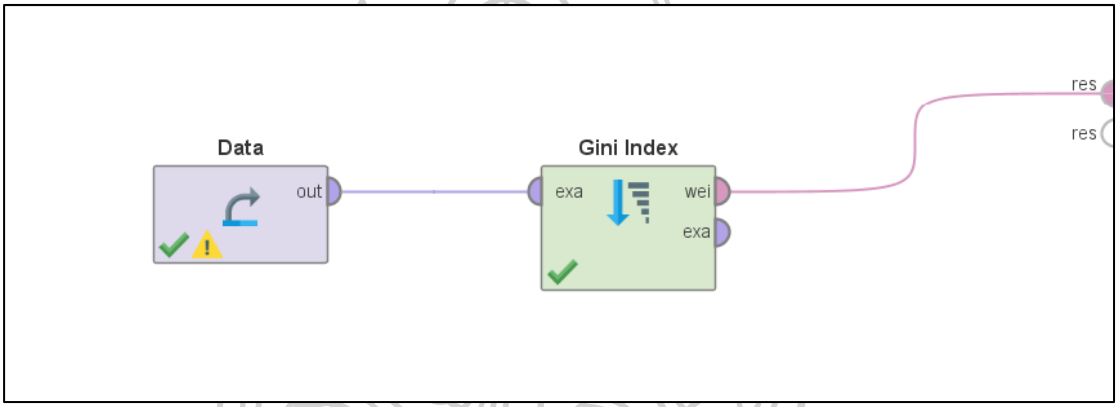
3.3.3 การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดและการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด

หลังจากปรับแต่งข้อมูลปัจจัยเชิงพื้นที่ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดแล้ว ทำการนำข้อมูลเข้าสู่กระบวนการวิเคราะห์ข้อมูล แบ่งออกเป็น (3.3.3.1) การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วยการทำเหมืองข้อมูลวิธีต้นไม้การตัดสินใจ และ (3.3.3.2) การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes โดยมีรายละเอียดดังนี้

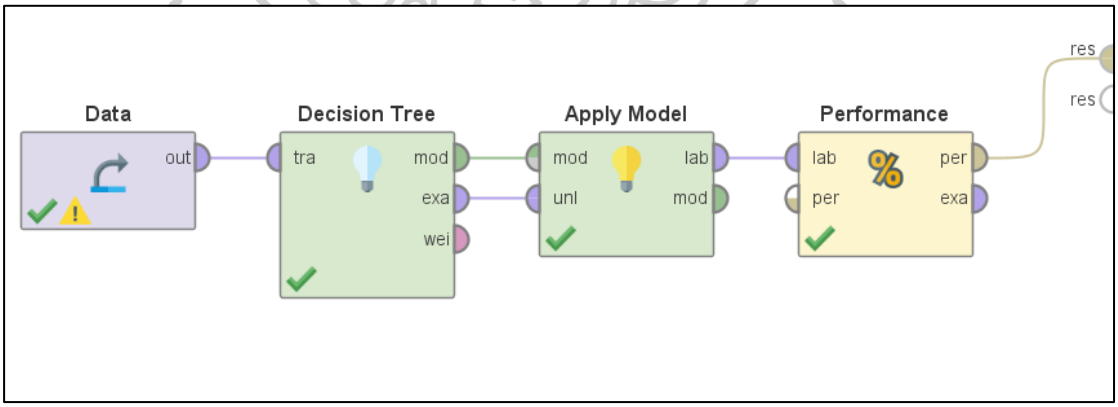
3.3.3.1 การศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วยการทำเหมืองข้อมูลวิธีต้นไม้การตัดสินใจ ซึ่งวิธีการทำเหมืองข้อมูลด้วยวิธีต้นไม้การตัดสินใจเป็นการทำเหมืองประเภทการเรียนรู้แบบมีผู้สอน การทำเหมืองข้อมูลโดยใช้การจำแนกแบบต้นไม้ตัดสินใจประกอบด้วยสองขั้นตอน คือ การสร้างโหนด และการสร้างความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด ในกระบวนการสร้างโหนด คือ การเปลี่ยนปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดให้เป็นโหนดก่อนที่จะสร้างความสัมพันธ์เชิงพื้นที่ของแต่ละโหนด การฝึกหัดข้อมูลได้ถูกนำไปใช้เพื่อหาโหนดเริ่มต้นที่ได้จากความสำคัญของปัจจัยยิ่งมีความสำคัญมากจะถูกนำไปใช้เป็นโหนดเริ่มต้น ซึ่งการศึกษาครั้งนี้คำนวณค่าความสำคัญของปัจจัยออกมาในรูปของค่า Information gain และค่า Gini index ในขั้นตอนที่สองเป็นการสร้างความสัมพันธ์ของแต่ละโหนด โดยมีโหนดเริ่มต้นเป็นโหนดหลักของการสร้างความสัมพันธ์เชิงพื้นที่ การศึกษาครั้งนี้ได้ประยุกต์ใช้ซอฟต์แวร์ RapidMiner ในการคำนวณค่า Information gain และค่า Gini index แสดงการวางรูปแบบการคำนวณค่าทั้งสองดังภาพที่ 7 และภาพที่ 8 นอกจากนี้ RapidMiner ถูกนำมาใช้ศึกษาความสัมพันธ์เชิงพื้นที่ด้วยการสร้างต้นไม้การตัดสินใจ โดยใช้อัลกอริทึมในการวิเคราะห์เป็น C4.5 ขั้นตอนการวิเคราะห์ (1) นำข้อมูลปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดเข้าสู่หน้าต่าง Process (2) ทำการเลือกตัววิเคราะห์ Decision tree, Apply Model และ Performance ทั้งหมดที่เลือกมาเป็นองค์ประกอบของการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด โดยที่ Decision tree คือ ส่วนของการวิเคราะห์เพื่อศึกษาความสัมพันธ์เชิงพื้นที่ของการลักลอบลำเลียงยาเสพติด Apply Model คือ ส่วนที่ช่วยในการหมุนเวียนชุดของข้อมูลในการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ของการลักลอบลำเลียงยาเสพติดด้วย Decision tree และ Performance ช่วยในการคำนวณร้อยละความถูกต้องของการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ของการลักลอบลำเลียงยาเสพติดด้วย Decision tree และขั้นสุดท้าย (3) นำส่วนประกอบทั้งหมดเชื่อมเข้าด้วยกัน เพื่อทำการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ แสดงการวางรูปแบบการสร้างต้นไม้การตัดสินใจดังภาพที่ 9



ภาพที่ 7 *การจัดรูปแบบในการคำนวณค่า Information gain ด้วย RapidMiner*

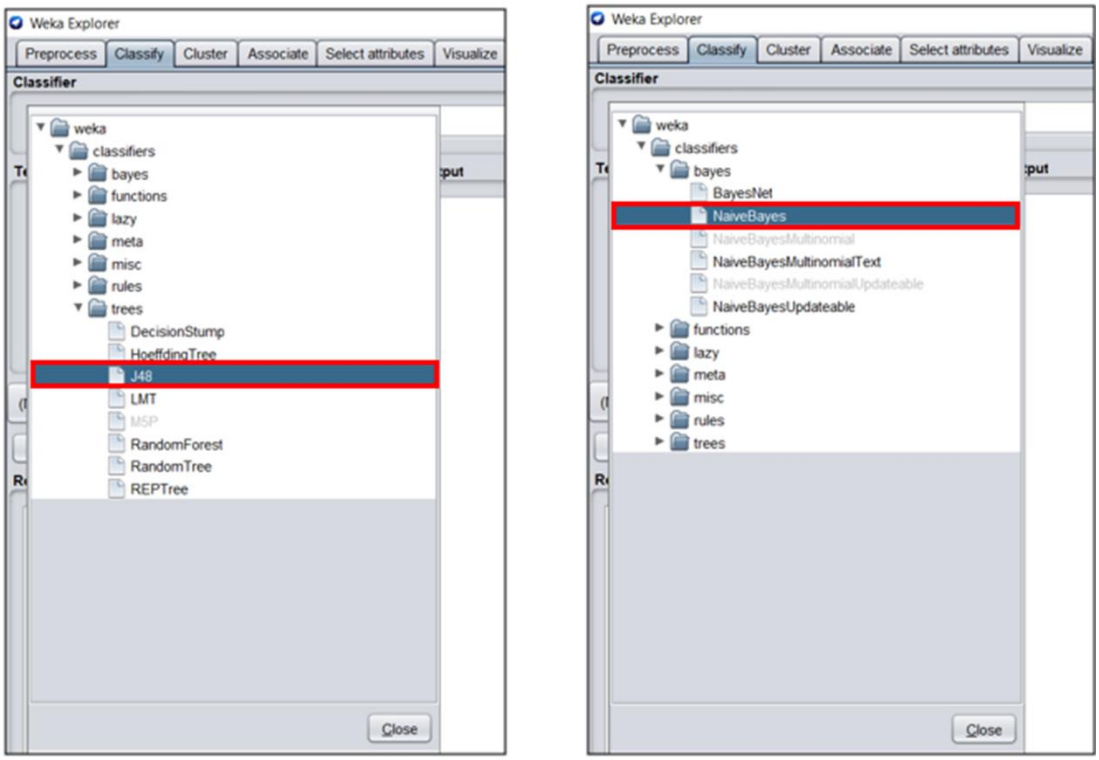


ภาพที่ 8 *การจัดรูปแบบในการคำนวณค่า Gini index ด้วย RapidMiner*

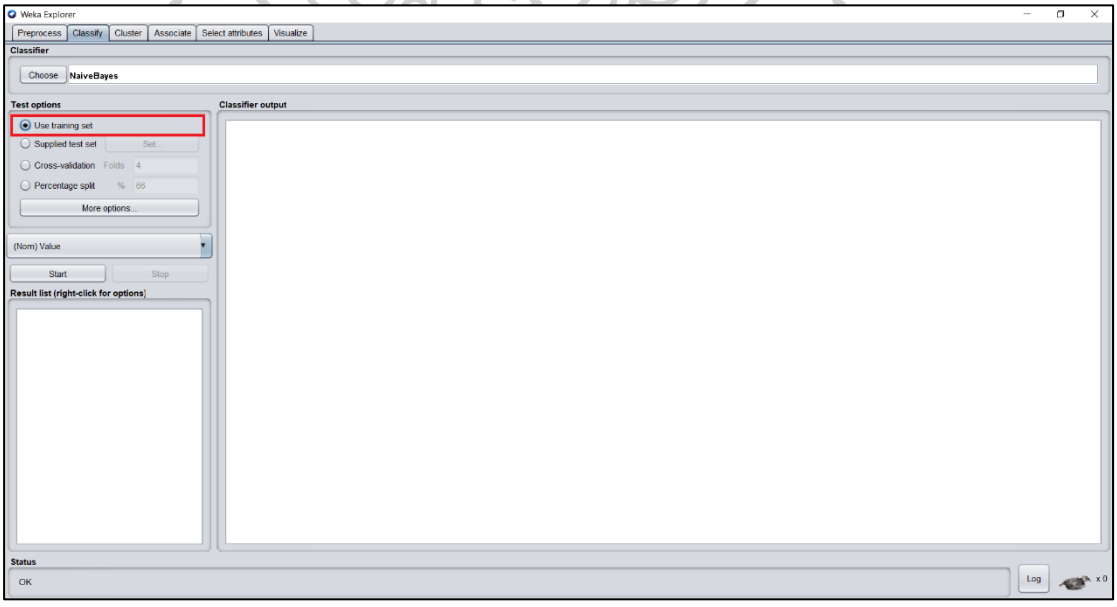


ภาพที่ 9 *การจัดรูปแบบในการวิเคราะห์ความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วย RapidMiner*

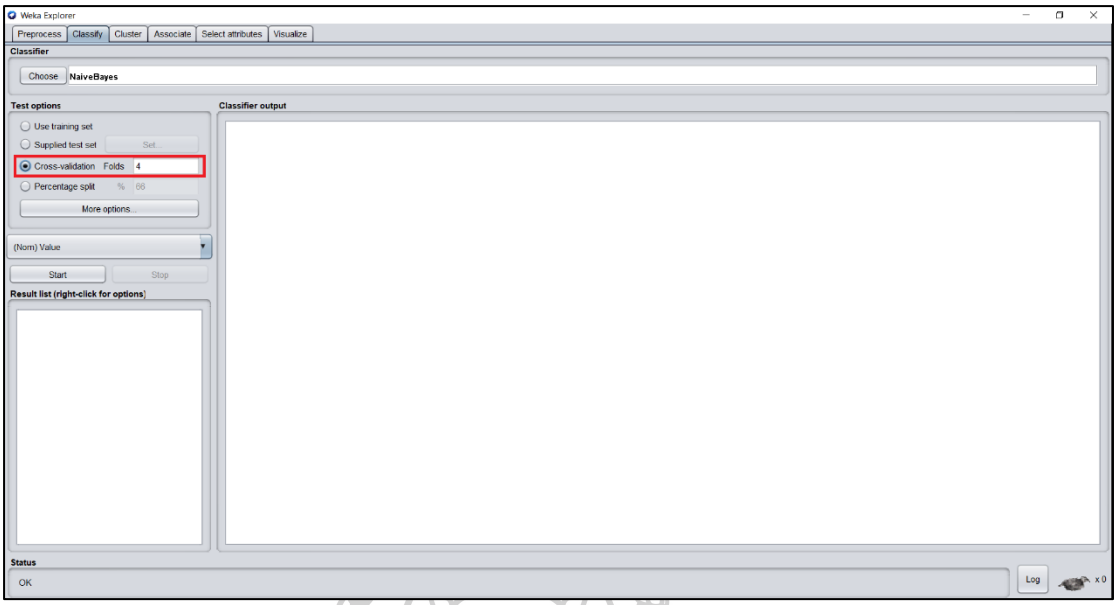
3.3.3.2 การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ได้นำซอฟต์แวร์ Weka มาใช้กับการศึกษาครั้งนี้ การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes แบ่งออกเป็น 2 ส่วน ส่วนที่ 1 คือ การคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ อัลกอริทึม C4.5 (J48) แสดงหน้าต่างการเลือกวิธีต้นไม้การตัดสินใจดังภาพที่ 10 (ก) และส่วนที่ 2 คือ การคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes แสดงหน้าต่างการเลือกวิธี Naïve Bayes ดังภาพที่ 10 (ข) ซึ่งขั้นตอนในการวิเคราะห์ทั้ง 2 ส่วน ได้ใช้ข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดที่ได้จากขั้นตอนในหัวข้อที่ 3.3.3.1 วิธีการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดทั้ง 2 วิธี แตกต่างกันที่วิธีที่ใช้ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด แต่ในส่วนของขั้นตอนการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดเหมือนกัน โดยมีรายละเอียดขั้นตอนของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดดังนี้ ขั้นตอนที่ 1 เป็นการฝึกหัดข้อมูล เพื่อให้ได้ค่าความถูกต้องของข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด แสดงการเลือกการฝึกหัดข้อมูลดังภาพที่ 11 ขั้นตอนที่ 2 เป็นการทดสอบข้อมูล โดยจะทำการแบ่งข้อมูลออกเป็น 4 ส่วน ในการทำ K-fold cross validation เพื่อให้ได้มาซึ่งค่าความถูกต้องของข้อมูลที่ผ่านการฝึกหัดข้อมูล แสดงการเลือกการทดสอบข้อมูลดังภาพที่ 12 และขั้นตอนสุดท้ายเป็นการนำเข้าข้อมูลการจับกุมการลักลอบลำเลียงยาเสพติดที่ไม่ปรากฏในคลังข้อมูลที่นำไปฝึกหัดข้อมูล เพื่อใช้ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด แสดงการเลือกข้อมูลที่ใช้คาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดดังภาพที่ 13 ซึ่งผลที่ได้จากการคาดการณ์อยู่ในช่วงระหว่าง 0.0 ถึง 1.0 ค่าความถูกต้องจะแปรผันตามค่าที่ได้



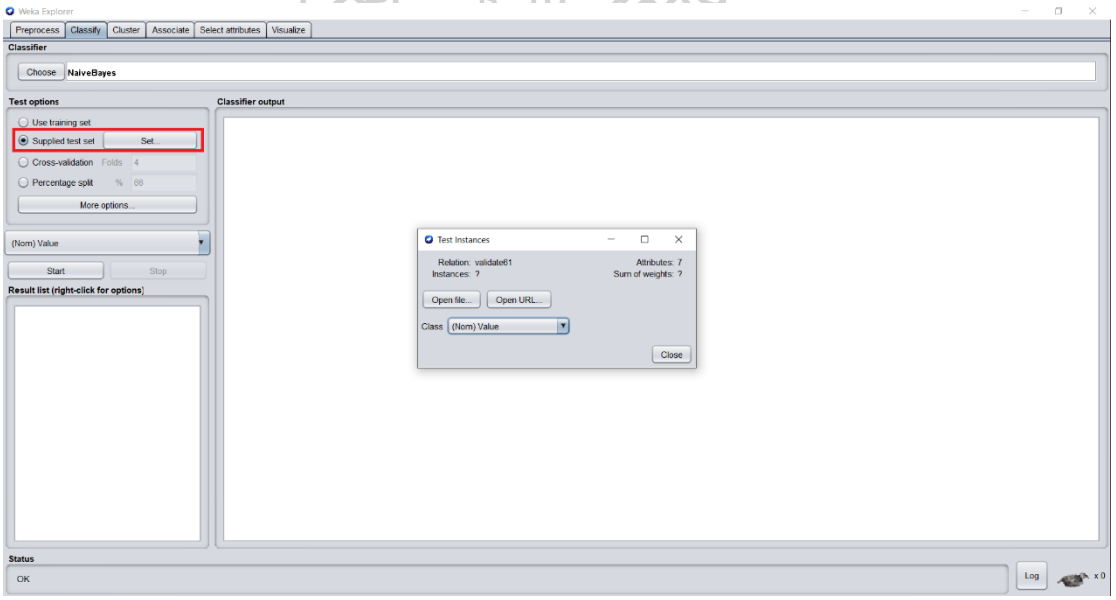
ภาพที่ 10 *การเลือกวิธีการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด*
(ก) วิธีต้นไม้การตัดสินใจ (ข) วิธี Naïve Bayes



ภาพที่ 11 *การฝึกหัดข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด*



ภาพที่ 12 *การทดสอบข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด*



ภาพที่ 13 *การคาดการณ์จากข้อมูลที่ไม่ปรากฏในคลังข้อมูลที่ผ่านการฝึกหัดข้อมูลความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด*

3.3.4 ขั้นตอนการแสดงผลการศึกษา

ขั้นตอนแสดงผลการศึกษาแบ่งออกเป็น 2 ส่วน คือ (1) ผลจากการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดแสดงในรูปแบบแผนภาพต้นไม้ที่มีการเชื่อมความสัมพันธ์เชิงพื้นที่ของแต่ละปัจจัยเข้าด้วยกัน (2) ผลจากการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayesแสดงในรูปแบบตารางค่าสถิติร้อยละความน่าจะเป็น และร้อยละความถูกต้อง เพื่อทำการเปรียบเทียบวิธีที่เหมาะสมในการคาดการณ์การลักลอบลำเลียงยาเสพติด นอกจากนั้นนำผลการคาดการณ์ดังกล่าวมาแสดงในรูปแบบแผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติด

บทที่ 4
ผลการศึกษา

การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วย Knowledge Discovery from Database ได้ผลการศึกษาตามวัตถุประสงค์ของการศึกษา ได้แก่ (4.1) ผลการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ และ (4.2) ผลการศึกษาวิธีที่ใช้ในการทำเหมืองข้อมูลที่เหมาะสมต่อการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด โดยมีรายละเอียดดังนี้

4.1 ผลการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ

4.1.1 ผลการคำนวณหาความสำคัญของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดเพื่อใช้เป็นตัวกำหนดโหนดเริ่มต้น

ขั้นตอนการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดของการศึกษาครั้งนี้ได้ใช้ค่า Information gain และค่า Gini index ในการกำหนดโหนดเริ่มต้นของความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด ซึ่งค่าทั้ง 2 ได้จากการคำนวณจากการจำแนกข้อมูลปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดด้วยปัจจัยมูลค่ายาเสพติด จากการคำนวณทั้ง 2 ค่าโดยใช้ปัจจัยมูลค่ายาเสพติดเป็นปัจจัยในการจำแนกข้อมูลพบว่าปัจจัยตำแหน่งพิกัดการจับกุม ปัจจัยประเภทของยาเสพติด ปัจจัยการใช้ประโยชน์ที่ดิน ปัจจัยฤดูกาล ปัจจัยรัศมีจากจุดตรวจ และปัจจัยเส้นทางที่มีการจับกุม มีความสำคัญจากมากไปน้อยตามลำดับดังตารางที่ 5 และตารางที่ 6 ดังนั้นปัจจัยที่ถูกกำหนดเป็นโหนดเริ่มต้นของความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด คือ ปัจจัยตำแหน่งพิกัดการจับกุม แสดงให้เห็นว่าพื้นที่ที่มาการลักลอบลำเลียงในแต่ละอำเภอมีความสัมพันธ์ต่อมูลค่าของยาเสพติดที่จับกุมได้ เช่น พื้นที่อำเภอเชียงดาวที่มีลักษณะภูมิประเทศเป็นเขาสูงทำให้การปฏิบัติงานของเจ้าหน้าที่ที่เกี่ยวข้องเป็นไปอย่างยากลำบาก ผู้ต้องหาจึงมีการลักลอบลำเลียงยาเสพติดในมูลค่าสูง

ตารางที่ 5 *ค่า Information gain ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด*

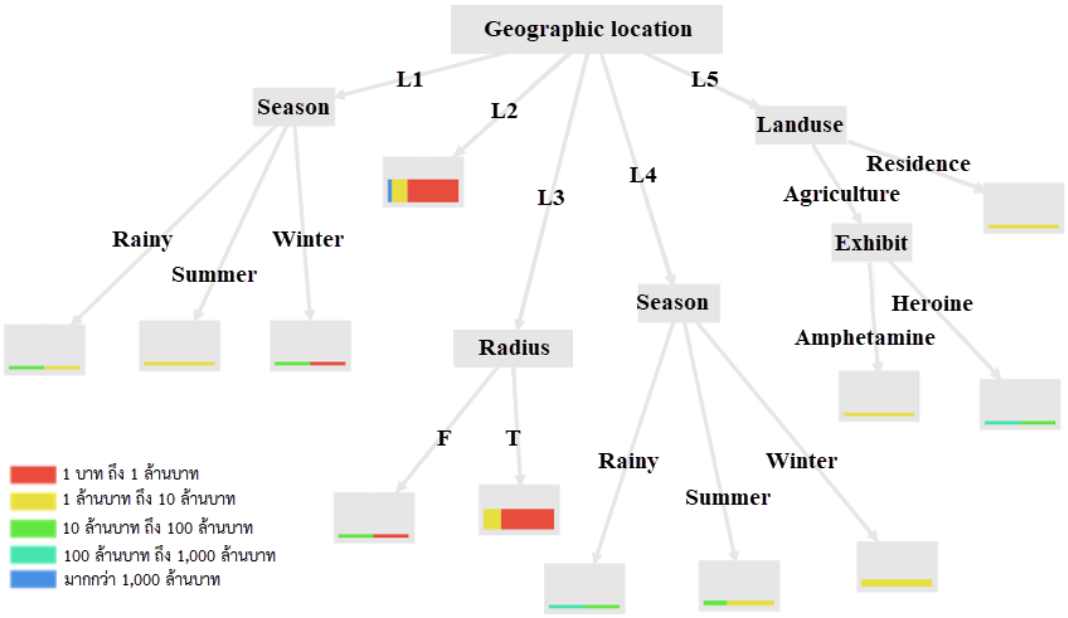
ปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด	ค่า Information gain
ตำแหน่งพิกัดการจับกุม	0.518
ประเภทของยาเสพติด	0.348
การใช้ประโยชน์ที่ดิน	0.194
ฤดูกาล	0.179
รัศมีจากจุดตรวจ	0.135
เส้นทางที่มีการจับกุม	0.060

ตารางที่ 6 *ค่า Gini index ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด*

ปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด	ค่า Gini index
ตำแหน่งพิกัดการจับกุม	0.169
ประเภทของยาเสพติด	0.102
การใช้ประโยชน์ที่ดิน	0.057
ฤดูกาล	0.038
รัศมีจากจุดตรวจ	0.038
เส้นทางที่มีการจับกุม	0.019

4.1.2 ผลของความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด

หลังจากได้ผลของการกำหนดโหนดเริ่มต้นแล้ว เข้าสู่ขั้นตอนของการสร้างความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด แสดงความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดดังภาพที่ 14



หมายเหตุ L1 อำเภอแม่จัน และอำเภอแม่สาย L2 อำเภอเชียงดาว และอำเภอเวียงแหง L3 อำเภอไชยปราการ L4 อำเภอแม่อาย และอำเภอฝาง L5 อำเภอแม่ฟ้าหลวง F รัศมีนอกจุดตรวจ 500 เมตร T รัศมีระหว่างจุดตรวจ 500 เมตร

ภาพที่ 14 *ความสัมพันธ์เชิงพื้นที่ของปัจจัยการลักลอบลำเลียงยาเสพติด*

จากความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดพบว่ามีความถูกต้องร้อยละ 74.1 และมีรายละเอียดดังนี้ ปัจจัยตำแหน่งพิกัดการจับกุมในอำเภอแม่จัน อำเภอแม่สาย (L1) อำเภอแม่อาย และอำเภอฝาง (L4) มีความสัมพันธ์กับปัจจัยฤดูกาล เมื่อสังเกตจากพื้นที่อำเภอดังกล่าวพบว่า ฤดูฝนเป็นฤดูที่มีช่วงมูลค่าของยาเสพติดสูงสุดอยู่ในช่วง 10 ล้านบาท ถึง 100 ล้านบาท เนื่องจากในฤดูฝนจะมีปริมาณยาเสพติดที่ถูกลักลอบลำเลียงต่อครั้งมากกว่าฤดูร้อนและฤดูหนาว เพื่อเป็นการลดความเสี่ยงจากการขาดทุนที่มาจากความเสียหายของยาเสพติดขณะลำเลียง ปัจจัยตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติดในพื้นที่อำเภอเชียงดาว และอำเภอเวียงแหง (L2) พบว่าทั้ง 2 อำเภอนี้มีช่วงของมูลค่ายาเสพติดที่จับกุมได้สูงสุดอยู่ในช่วงมากกว่า 1 พันล้านบาท เนื่องจากการกระจายตัวของปัจจัยมูลค่ายาเสพติดค่อนข้างสูงทำให้ค่าเอนโทรปีสูงตามไปด้วย ส่งผลให้ในพื้นที่ดังกล่าวไม่สามารถจำแนกข้อมูลความสัมพันธ์เชิงพื้นที่ร่วมกับปัจจัยอื่นได้ ปัจจัยตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติดในพื้นที่อำเภอไชยปราการ (L3) มีความสัมพันธ์กับปัจจัยรัศมีจากจุดตรวจ 500 เมตร และจากความสัมพันธ์ดังกล่าวพบว่าช่วงมูลค่ายาเสพติดที่จับกุมได้สูงสุดอยู่นอกรัศมีจากจุดตรวจ 500 เมตร มีค่าอยู่ในช่วง 10 ล้านบาท ถึง 1 พันล้านบาท สาเหตุที่พบช่วงมูลค่ายาเสพติดนอกรัศมีจุดตรวจมากกว่าในรัศมีจุดตรวจ เนื่องมาจากปริมาณ

ยาเสพติดที่ลักลอบลำเลียงผ่านจุดตรวจมีปริมาณน้อยกว่าการลักลอบลำเลียงนอกรัศมีจุดตรวจ เพื่อลดความเสี่ยงต่อการเป็นที่ต้องสงสัยของเจ้าหน้าที่ และปัจจัยตำแหน่งพิกัดการจับกุมการลักลอบลำเลียงยาเสพติดในพื้นที่อำเภอแม่ฟ้าหลวง (L5) มีความสัมพันธ์กับปัจจัยการใช้ประโยชน์ที่ดิน และปัจจัยประเภทของยาเสพติด จากการสังเกตพบว่าในพื้นที่อำเภอแม่ฟ้าหลวงมีช่วงมูลค่ายาเสพติดในพื้นที่การใช้ประโยชน์ที่ดินประเภทชุมชน และสิ่งปลูกสร้างอยู่ในช่วง 1 บาท ถึง 10 ล้านบาท ในขณะที่การใช้ประโยชน์ที่ดินประเภทเกษตรกรรมช่วงมูลค่ายาเสพติดสูงสุดอยู่ในช่วง 100 ล้านบาท ถึง 1 พันล้านบาท และประเภทยาเสพติดที่จับได้ส่วนใหญ่เป็นยาเสพติดประเภทเฮโรอีน ซึ่งสอดคล้องกับข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดในพื้นที่อำเภอแม่ฟ้าหลวงที่รวบรวมได้ พบว่าในพื้นที่อำเภอแม่ฟ้าหลวงมีการจับกุมยาเสพติดในพื้นที่การใช้ประโยชน์ที่ดินประเภทเกษตรกรรมตามประเภทของยาเสพติดที่จับได้เป็นประเภทยาบ้าร้อยละ 55 และประเภทเฮโรอีนร้อยละ 31 ถึงแม้ยาเสพติดประเภทเฮโรอีนมีร้อยละของข้อมูลการจับกุมที่น้อยกว่ายาเสพติดประเภทยาบ้า แต่ถ้าคำนวณจากต้นทุนการผลิตยาเสพติดประเภทเฮโรอีนจะมีมูลค่าที่สูงกว่ายาเสพติดประเภทยาบ้า นอกจากนี้พบว่ามีปัจจัยหนึ่งที่ไม่ถูกแสดงในผลการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด คือ ปัจจัยเส้นทางที่มีการจับกุม ซึ่งสามารถสังเกตได้จากค่า Information gain และค่า Gini index ที่แสดงให้เห็นว่าปัจจัยเส้นทางที่มีการจับกุมมีความสำคัญในการสร้างความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดน้อยที่สุด และเมื่อสังเกตจากข้อมูลการจับกุมที่รวบรวมได้พบว่าเส้นทางที่ใช้ร้อยละ 85.0 เป็นเส้นทางรอง และอีกร้อยละ 15.0 เป็นเส้นทางภูมิประเทศ จากข้อมูลดังกล่าวสามารถตีความได้ว่าเมื่อมีการลักลอบลำเลียงยาเสพติดเกิดขึ้นในสภาวะที่ปัจจัยตำแหน่งพิกัดการจับกุม ปัจจัยประเภทของยาเสพติด ปัจจัยการใช้ประโยชน์ที่ดิน ปัจจัยฤดูกาล และรัศมีจากจุดตรวจมีการเปลี่ยนแปลง แต่ปัจจัยที่ไม่มีการเปลี่ยนแปลง คือ ปัจจัยเส้นทางที่มีการจับกุมการลักลอบลำเลียงยาเสพติดที่ส่วนใหญ่เกิดขึ้นในเส้นทางรอง ถึงแม้ปัจจัยเส้นทางการลักลอบลำเลียงยาเสพติดไม่มีความสัมพันธ์กับปัจจัยอื่น แต่ปัจจัยเส้นทางการลักลอบลำเลียงยาเสพติดเป็นปัจจัยที่เอื้อต่อการลักลอบลำเลียงยาเสพติดตามการศึกษาของ Medel, Lu, and Chow (2015) ที่แสดงให้ว่าปัจจัยเส้นทางคมนาคมมีส่วนช่วยในการลดต้นทุนการลักลอบลำเลียงยาเสพติด

4.2 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ (Decision tree) และวิธีเบย์อย่างง่าย (Naïve Bayes)

จากความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดที่ได้ในหัวข้อที่ 4.1.2 ถูกนำมาใช้ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด ซึ่งผลของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแบ่งเป็น (4.2.1) ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ และ (4.2.2) ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes เมื่อได้ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดทั้ง 2 วิธีแล้ว ผลที่ได้ดังกล่าวจะถูกนำมาแสดงเชิงพื้นที่ในรูปแบบแผนที่ระดับความเสี่ยงต่อการลักลอบลำเลียงยาเสพติด และนำมาเปรียบเทียบวิธีการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ในหัวข้อที่ (4.2.3) โดยผลการศึกษาทั้งหมดมีรายละเอียดดังต่อไปนี้

4.2.1 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ

ในส่วนของผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจแบ่งออกเป็น (4.2.1.1) ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด (4.2.1.2) ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด และ (4.2.1.3) ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ ซึ่งผลการคาดดการณพื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีรายละเอียดดังนี้

4.2.1.1 ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงดังตารางที่ 7 ซึ่งสามารถอธิบายตามตารางที่ 7 ได้ว่าผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances คือ ค่าร้อยละความถูกต้องของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับร้อยละ 63.475 (2) Incorrectly Classified Instances คือ ค่าร้อยละความผิดพลาดของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับร้อยละ 36.525 (3) Kappa statistic คือ ค่าสัมประสิทธิ์ตัวชี้วัดทางสถิติความสอดคล้องของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับ 0.274 ซึ่งค่า Kappa statistic มีค่าอยู่ในช่วง 0.000 - 1.000 (4) Mean absolute error คือ ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ของการคาดการณ์พื้นที่

เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับ 0.201 (5) Root mean squared คือ ค่าเฉลี่ยความคาดเคลื่อนกำลังสองของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับ 0.317 (6) Relative absolute error คือ ค่าความคาดเคลื่อนสัมพัทธ์ของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีค่าเท่ากับร้อยละ 84.7 และ (7) Root relative absolute error คือ ค่าความคาดเคลื่อนสัมพัทธ์ของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่มีการถอดรากที่ 2 มีค่าเท่ากับร้อยละ 92.1 ซึ่งค่า Relative absolute error และค่า Root relative absolute error ยิ่งมีค่ามากแสดงว่าผลของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีความคาดเคลื่อนมากเช่นกัน

ตารางที่ 7 *ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการฝึกหัดข้อมูล*

การคาดการณ์ด้วยต้นไม้การตัดสินใจ	**ผลการคาดการณ์**
Correctly Classified Instances	63.5 %
Incorrectly Classified Instances	36.5 %
Kappa statistic	0.274
Mean absolute error	0.201
Root mean squared	0.317
Relative absolute error	84.7 %
Root relative absolute error	92.2 %

4.2.1.2 ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงดังตารางที่ 8 ซึ่งสามารถอธิบายตามตารางที่ 8 ได้ว่าผลของขั้นตอนการทดสอบข้อมูลด้วย K-fold cross validation จากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances มีค่าเท่ากับร้อยละ 57.1 (2) Incorrectly Classified Instances มีค่าเท่ากับร้อยละ 42.9 (3) Kappa statistic มีค่าเท่ากับ 0.144 (4) Mean absolute error มีค่าเท่ากับ 0.223 (5) Root mean squared มีค่าเท่ากับ 0.346 (6) Relative

absolute error มีค่าเท่ากับร้อยละ 93.9 และ (7) Root relative absolute error มีค่าเท่ากับร้อยละ 100.7 ซึ่งแต่ละค่าได้อธิบายความหมายไว้ในหัวข้อที่ 4.2.1.1

ตารางที่ 8 *ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการทดสอบข้อมูล*

การคาดการณ์ด้วยต้นไม้การตัดสินใจ	**ผลการคาดการณ์**
Correctly Classified Instances	57.1 %
Incorrectly Classified Instances	42.9 %
Kappa statistic	0.144
Mean absolute error	0.223
Root mean squared	0.346
Relative absolute error	93.9 %
Root relative absolute error	100.7 %

4.2.1.3 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจแสดงดังตารางที่ 9 ซึ่งสามารถอธิบายตามตารางที่ 9 ได้ว่าผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances มีค่าเท่ากับร้อยละ 47.7 (2) Incorrectly Classified Instances มีค่าเท่ากับร้อยละ 52.3 (3) Kappa statistic มีค่าเท่ากับ 0.194 (4) Mean absolute error มีค่าเท่ากับ 0.297 (5) Root mean squared มีค่าเท่ากับ 0.397 (6) Relative absolute error มีค่าเท่ากับร้อยละ 96.5 และ (7) Root relative absolute error มีค่าเท่ากับร้อยละ 102.3 ซึ่งแต่ละค่าได้อธิบายความหมายไว้ในหัวข้อที่ 4.2.1.1

ตารางที่ 9 *ผลการคาดการณ์ด้วยต้นไม้การตัดสินใจขั้นตอนการคาดการณ์*

การคาดการณ์ด้วยต้นไม้การตัดสินใจ	**ผลการคาดการณ์**
Correctly Classified Instances	47.7 %
Incorrectly Classified Instances	52.3 %
Kappa statistic	0.194
Mean absolute error	0.297
Root mean squared	0.397
Relative absolute error	96.5 %
Root relative absolute error	102.3 %

นอกจากนี้ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจยังแสดงผลของประสิทธิภาพในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด โดยสามารถอธิบายค่าที่แสดงในตารางที่ 10 ได้ดังนี้ Class ที่แสดงในตารางที่ 10 คือ ตำแหน่งอำเภอที่มีการจับกุมการลักลอบลำเลียงยาเสพติด (L_1-L_5) สามารถอธิบายผลโดยแบ่งตามอำเภอได้ดังนี้ อำเภอเชียงดาว และอำเภอเวียงแหง (L_2) มีค่าอัตราส่วนของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่เป็นจริงต่อเหตุการณ์ที่เกิดขึ้นจริงทั้งหมด (TP Rate) เท่ากับ 0.000 ซึ่งค่า TP Rate มีค่าเท่ากับค่าความแม่นยำของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด (Recall) เสมอ และมีค่าอัตราส่วนของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่ไม่เป็นจริงต่อเหตุการณ์ที่เกิดขึ้นจริงทั้งหมด (FP Rate) เท่ากับ 0.034 มีค่าที่แสดงถึงค่าความถูกต้องของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่เป็นจริง (Precision) มีค่าเท่ากับ. 0.000 แสดงว่าการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่เป็นจริงไม่เกิดขึ้นในพื้นที่ L_2 ค่าที่แสดงถึงประสิทธิภาพของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด (F-measure) มีค่าเท่ากับ 0.000 นอกจากค่าสถิติที่กล่าวมาข้างต้นยังมีค่าสถิติที่แสดงในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดอีก 3 ค่า ได้แก่ ค่าวัดคุณภาพ 2 ระดับของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด (MCC) ซึ่งค่า MCC แสดงทั้งค่าบวกและลบ ยิ่งค่ามีลบมากแสดงว่าค่าความคาดเคลื่อนมากตามไปด้วย ค่าที่ได้จากกราฟความสัมพันธ์ระหว่าง TP Rate กับ FP

Rate (ROC Area) และสุดท้ายค่าพื้นที่ใต้กราฟของกราฟความสัมพันธ์ระหว่างค่า Precision กับค่า Recall (PRC Area) ในการศึกษาครั้งนี้มีค่าความแม่นยำรวมของการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดเท่ากับ 0.652 (ร้อยละ 65.2) และได้มีการใช้ค่า F-measure ที่เป็นค่าแสดงประสิทธิภาพของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่แสดงในรูปแบบของค่าเฉลี่ยของค่า Precision และค่า Recall จากค่า F-measure ของอำเภอที่เหลือมีดังนี้ อำเภอแม่อาย และอำเภอฝางมีค่าเท่ากับ 0.000 อำเภอแม่ฟ้าหลวงมีค่าเท่ากับ 0.222 อำเภอแม่จัน และอำเภอแม่สายมีค่าเท่ากับ 0.412 และสุดท้ายอำเภอไชยปราการมีค่าเท่ากับ 0.676 ค่า F-measure ที่ได้จะถูกนำไปสร้างแผนที่ความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดเพื่อนำเสนอต่อไปในหัวข้อที่ 4.2.3

ตารางที่ 10 *รายละเอียดผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยต้นไม้การตัดสินใจ*

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
0.000	0.034	0.000	0.000	0.000	-0.062	0.459	0.118	L_2
0.000	0.068	0.000	0.000	0.000	-0.082	0.540	0.121	L_4
0.500	0.095	0.143	0.500	0.222	0.225	0.770	0.107	L_5
0.368	0.174	0.467	0.368	0.412	0.210	0.361	0.387	L_1
0.742	0.412	0.622	0.742	0.676	0.333	0.607	0.561	L_3
0.477	0.260	0.437	0.477	0.450	0.213	0.518	0.408	Weighted Avg.

4.2.2 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes

ในส่วนของผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes แบ่งออกเป็น (4.2.2.1) ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด (4.2.2.2) ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติด และ (4.2.2.3) ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes ซึ่งผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีรายละเอียดดังนี้

4.2.2.1 ผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงดังตารางที่ 11 ซึ่งสามารถอธิบายตามตารางที่ 11 ได้ว่าผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances มีค่าเท่ากับร้อยละ 58.2 (2) Incorrectly Classified Instances มีค่าเท่ากับร้อยละ 41.8 (3) Kappa statistic มีค่าเท่ากับ 0.216 (4) Mean absolute error มีค่าเท่ากับ 0.206 (5) Root mean squared มีค่าเท่ากับ 0.333 (6) Relative absolute error มีค่าเท่ากับร้อยละ 86.8 และ (7) Root relative absolute error มีค่าเท่ากับร้อยละ 96.7 ซึ่งแต่ละค่าได้อธิบายความหมายไว้ในหัวข้อที่ 4.2.1.1

ตารางที่ 11 *ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการฝึกหัดข้อมูล*

การคาดการณ์ด้วย Naïve Bayes	**ผลการคาดการณ์**
Correctly Classified Instances	58.2 %
Incorrectly Classified Instances	41.8 %
Kappa statistic	0.216
Mean absolute error	0.206
Root mean squared	0.333
Relative absolute error	86.8 %
Root relative absolute error	96.7 %

4.2.2.2 ผลของขั้นตอนการทดสอบข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงดังตารางที่ 12 ซึ่งสามารถอธิบายตามตารางที่ 12 ได้ว่าผลของขั้นตอนการทดสอบข้อมูลด้วย K-fold cross validation จากการใช้วิธีต้นไม้การตัดสินใจในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances มีค่าเท่ากับร้อยละ 55.1 (2) Incorrectly Classified Instances มีค่าเท่ากับร้อยละ 44.9 (3) Kappa statistic มีค่าเท่ากับ 0.158 (4) Mean absolute error มีค่าเท่ากับ 0.211 (5) Root mean squared มีค่าเท่ากับ 0.341 (6) Relative

absolute error มีค่าเท่ากับร้อยละ 89.0 และ (7) Root relative absolute error มีค่าเท่ากับร้อยละ 99.3 ซึ่งแต่ละค่าได้อธิบายความหมายไว้ในหัวข้อที่ 4.2.1.1

ตารางที่ 12 *ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการทดสอบข้อมูล*

การคาดการณ์ด้วย Naïve Bayes	**ผลการคาดการณ์**
Correctly Classified Instances	55.1 %
Incorrectly Classified Instances	44.9 %
Kappa statistic	0.158
Mean absolute error	0.211
Root mean squared	0.341
Relative absolute error	89.0 %
Root relative absolute error	99.3 %

4.2.2.3 ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes แสดงดังตารางที่ 13 ซึ่งสามารถอธิบายตามตารางที่ 13 ได้ว่าผลของขั้นตอนการฝึกหัดข้อมูลจากการใช้วิธี Naïve Bayes ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดแสดงค่าเชิงสถิติทั้งหมด 7 ค่า ประกอบด้วย (1) Correctly Classified Instances มีค่าเท่ากับร้อยละ 36.3 (2) Incorrectly Classified Instances มีค่าเท่ากับร้อยละ 63.7 (3) Kappa statistic มีค่าเท่ากับ 0.178 (4) Mean absolute error มีค่าเท่ากับ 0.297 (5) Root mean squared มีค่าเท่ากับ 0.389 (6) Relative absolute error มีค่าเท่ากับร้อยละ 94.8 และ (7) Root relative absolute error มีค่าเท่ากับร้อยละ 98.4 ซึ่งแต่ละค่าได้อธิบายความหมายไว้ในหัวข้อที่ 4.2.1.1

ตารางที่ 13 *ผลการคาดการณ์ด้วย Naïve Bayes ขั้นตอนการคาดการณ์*

การคาดการณ์ด้วยต้นไม้การตัดสินใจ	**ผลการคาดการณ์**
Correctly Classified Instances	36.3 %
Incorrectly Classified Instances	63.7 %
Kappa statistic	0.178
Mean absolute error	0.297
Root mean squared	0.389
Relative absolute error	94.8 %
Root relative absolute error	98.4 %

นอกจากนี้ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes ยังแสดงผลของประสิทธิภาพในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดดังตารางที่ 14 ซึ่งได้อธิบายความหมายของค่าสถิติในตารางไว้ในหัวข้อที่ 4.2.1.3 ดังนั้นในหัวข้อนี้จะแสดงผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดเฉพาะค่า F-measure ของแต่ละอำเภอมีดังนี้ อำเภอเชียงดาว และอำเภอเวียงแหงมีค่าเท่ากับ 0.349 อำเภอแม่อาย และอำเภอฝางมีค่าเท่ากับ 0.079 อำเภอแม่ฟ้าหลวงมีค่าเท่ากับ 0.335 อำเภอแม่จัน และอำเภอแม่สายมีค่าเท่ากับ 0.348 และสุดท้ายอำเภอไชยปราการมีค่าเท่ากับ 0.455 ค่า F-measure ที่ได้จะถูกนำไปสร้างแผนที่ความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดเพื่อนำเสนอต่อไปในหัวข้อที่ 4.2.3

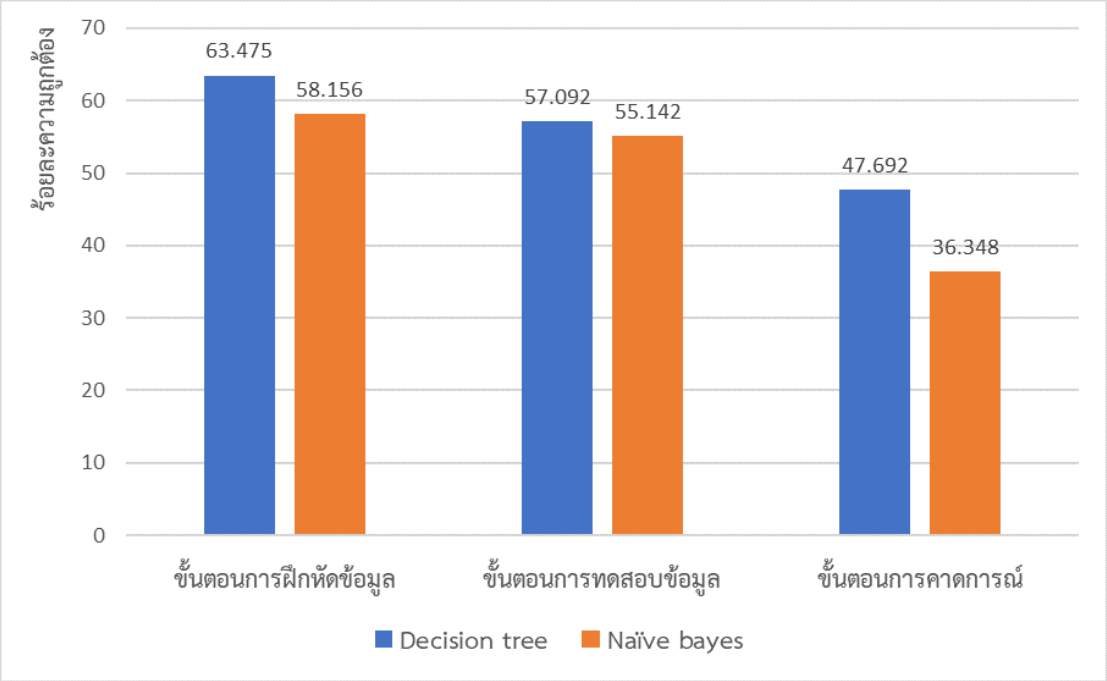
ตารางที่ 14 *รายละเอียดผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes*

TP Rate	FP Rate	Precision	Recall	F-measure	MCC	ROC Area	PRC Area	Class
0.309	0.129	0.400	0.309	0.349	0.198	0.639	0.336	L_2
0.045	0.012	0.333	0.045	0.079	0.084	0.617	0.209	L_4
0.430	0.186	0.274	0.430	0.335	0.205	0.686	0.272	L_5
0.299	0.149	0.415	0.299	0.348	0.169	0.597	0.348	L_1
0.581	0.346	0.374	0.581	0.455	0.210	0.635	0.345	L_3
0.363	0.185	0.372	0.363	0.342	0.181	0.631	0.317	Weighted Avg.

4.2.3 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes

จากผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดในหัวข้อที่ 4.2.2 นำมาสู่การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes ซึ่งในหัวข้อนี้แบ่งออกเป็น (4.2.3.1) การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจกับวิธี Naïve Bayes จากร้อยละความถูกต้อง และ (4.2.3.2) การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes รูปแบบเชิงพื้นที่

4.2.3.1 การเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes จากร้อยละความถูกต้องพบว่า ร้อยละความถูกต้องของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจมีค่ามากกว่าร้อยละความถูกต้องของการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes ในทุกขั้นตอนของกระบวนการวิเคราะห์ แสดงดังภาพที่ 17 ดังนั้นวิธีที่เหมาะสมในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดของการศึกษาครั้งนี้ คือ วิธีต้นไม้การตัดสินใจ ทั้งนี้ผลของการคาดการณ์ด้วยวิธีดังกล่าวยังจัดว่าอยู่ในระดับปานกลาง ซึ่งเกิดได้จากหลายสาเหตุเช่น ปริมาณข้อมูลไม่เพียงพอ การกระจายตัวของข้อมูลบางข้อมูลไม่สามารถจำแนกข้อมูลได้ และการกำหนดจำนวนปัจจัยที่น้อยเกินไป เป็นต้น



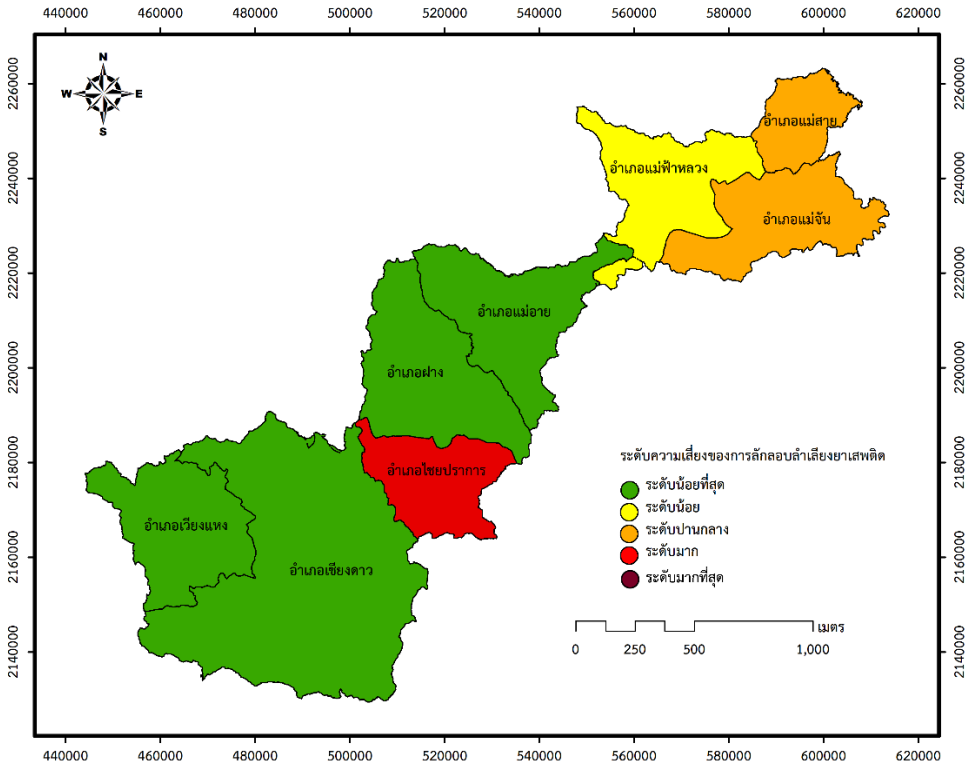
ภาพที่ 15 *กราฟเปรียบเทียบค่าความถูกต้องของวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes*

4.2.3.2 ผลการเปรียบเทียบผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดระหว่างวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes รูปแบบเชิงพื้นที่ ได้จากค่า F-measure ที่แสดงถึงประสิทธิภาพของการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดที่นำมาจัดช่วงของค่า F-measure ตั้งแต่ 0.000 ถึง 1.000 ได้ช่วงของระดับความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดดังตารางที่ 15

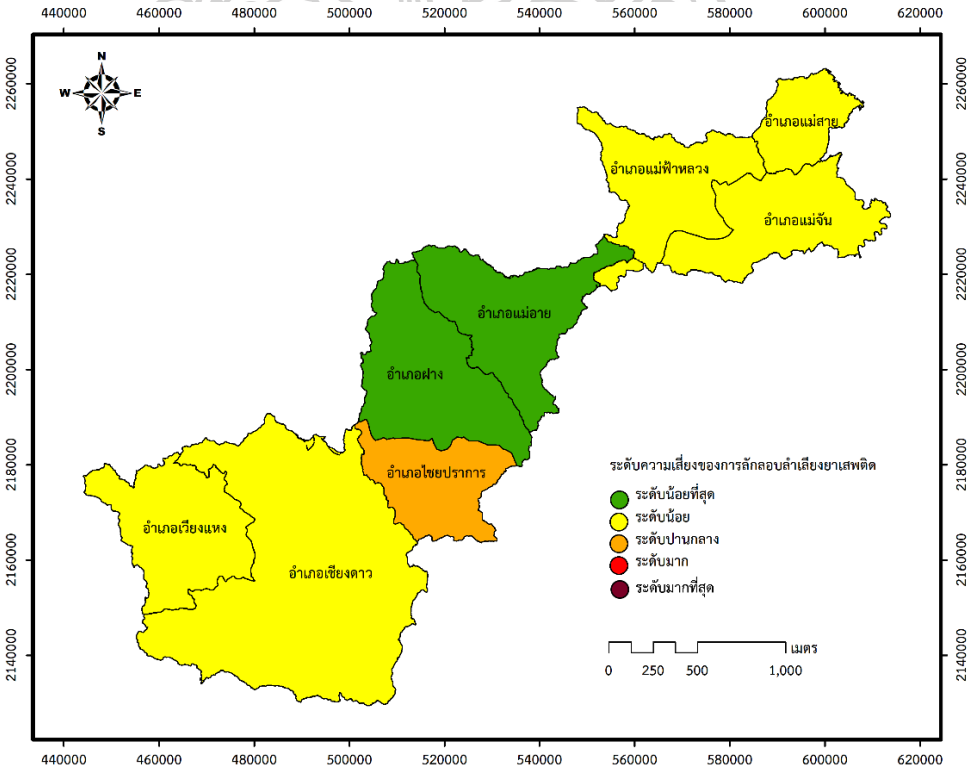
ตารางที่ 15 *เกณฑ์การกำหนดระดับความเสี่ยงจากค่า F-measure*

ค่า F-measure	**ระดับความเสี่ยง**
0.000 – 0.200	น้อยที่สุด
0.201 – 0.400	น้อย
0.401 – 0.600	ปานกลาง
0.601 – 0.800	มาก
0.801 – 1.000	มากที่สุด

จากนั้นนำค่า F-measure ที่ได้จากการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดจากตารางที่ 10 และตารางที่ 14 มาจัดลำดับตามเกณฑ์ระดับความเสี่ยงแล้วพบว่าการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจมีค่า F-measure มากที่สุดที่อำเภอไชยปราการมีค่าเท่ากับ 0.676 ซึ่งจัดอยู่ในระดับความเสี่ยงระดับมาก (สีแดง) และค่าน้อยที่สุดอยู่ที่อำเภอเชียงดาว อำเภอเวียงแหง อำเภอแม่อาย และอำเภอฝาง ทั้ง 4 อำเภอมีค่าเท่ากับ 0.000 ซึ่งจัดอยู่ในระดับความเสี่ยงระดับน้อยที่สุด (สีเขียว) แสดงแผนที่ระดับความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจดังภาพที่ 15 ส่วนการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes มีค่า F-measure มากที่สุดอยู่ที่อำเภอไชยปราการมีค่าเท่ากับ 0.455 ซึ่งจัดอยู่ในระดับความเสี่ยงระดับปานกลาง (สีส้ม) และค่าน้อยที่สุดอยู่ที่อำเภอแม่อาย และอำเภอฝางมีค่าเท่ากับ 0.079 ซึ่งจัดอยู่ในระดับความเสี่ยงระดับน้อยที่สุด (สีเขียว) แสดงแผนที่ระดับความเสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes ดังภาพที่ 16



ภาพที่ 16 *แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากวิธีต้นไม้การตัดสินใจ*



ภาพที่ 17 *แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากวิธี Naïve Bayes*

จากภาพที่ 16 แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ และภาพที่ 17 แผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes พบว่าสามารจัดลำดับระดับความเสี่ยงจากมาไปน้อยได้ดังนี้ อำเภอไชยปราการ อำเภอแม่จัน อำเภอแม่สาย อำเภอเชียงดาว อำเภอเวียงแหง อำเภอแม่อาย และอำเภอฝาง และเมื่อเปรียบเทียบระหว่างแผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจกับแผนที่ระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดจากการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธี Naïve Bayes พบว่าอำเภอแม่อาย อำเภอฝาง และอำเภอแม่ฟ้าหลวงมีระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดที่เหมือนกัน ส่วนอำเภอไชปราการ อำเภอเวียงแหง อำเภอเชียงดาว อำเภอแม่จัน และอำเภอแม่สายมีระดับความเสี่ยงของการลักลอบลำเลียงยาเสพติดที่แตกต่าง เนื่องจากใช้วิธีในการคาดการณ์ที่แตกต่างกัน ทำให้ค่า F-measure ที่นำมาทำแผนที่ระดับความเสี่ยงมีค่าอยู่ในช่วงที่แตกต่างกัน แต่จากการเปรียบเทียบค่าความถูกต้องพบว่าวิธีต้นไม้การตัดสินใจมีค่าความถูกต้องที่มากกว่าวิธี Naïve Bayes

มหาวิทยาลัยศิลปากร

บทที่ 5
สรุปและข้อเสนอแนะ

ในการศึกษาเรื่อง การวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูล (Knowledge discovery from database) มีบทสรุปดังนี้

5.1 สรุป

จากการวิเคราะห์รูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูลทั้งการศึกษาความสัมพันธ์เชิงพื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดในพื้นที่จังหวัดเชียงใหม่ และจังหวัดเชียงรายด้วยหลักการ Knowledge Discovery โดยใช้วิธีต้นไม้การตัดสินใจ และการศึกษาวิธีที่ใช้ในการทำเหมืองข้อมูลที่เหมาะสมต่อการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดพบว่า ข้อมูลข่าวการจับกุมการลักลอบลำเลียงยาเสพติดทั้งหมด 564 ข่าว ที่ได้จากการเก็บรวบรวมตั้งแต่ปี พ.ศ. 2554 ถึงปี พ.ศ. 2560 ได้ถูกนำมาศึกษาความสัมพันธ์เชิงพื้นที่ในรูปแบบของการกำหนดปัจจัยจากข้อมูลข่าวการจับกุมทำให้ทราบว่าตำแหน่งพื้นที่ที่มีการจับกุมมีความสำคัญสูงที่สุด เนื่องจากการลักลอบลำเลียงยาเสพติดจำเป็นต้องใช้ลักษณะภูมิประเทศเพื่อให้เอื้อต่อการลักลอบลำเลียงยาเสพติด ซึ่งในแต่ละอำเภอมีความสัมพันธ์กับปัจจัยที่ส่งผลต่อการลักลอบลำเลียงอื่น ๆ ที่แตกต่างกันขึ้นอยู่กับข้อมูลที่รวบรวมได้ เช่น ในอำเภอแม่จัน และอำเภอแม่สายมีความสัมพันธ์กับปัจจัยฤดูกาล เนื่องจากข้อมูลฤดูกาลที่มีการจับกุมมีนัยสำคัญที่สุดในอำเภอดังกล่าว จากนั้นได้นำความสัมพันธ์เชิงพื้นที่พื้นที่ของปัจจัย ที่ส่งผลต่อการลักลอบลำเลียงยาเสพติดมาใช้ในการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดด้วยวิธีต้นไม้การตัดสินใจ และวิธี Naïve Bayes เพื่อศึกษาความเหมาะสมของวิธีที่ใช้ในคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดปรากฏว่าวิธีต้นไม้การตัดสินใจเป็นวิธีที่เหมาะสมกับการศึกษาครั้งนี้มากกว่าวิธี Naïve Bayes เพราะเมื่อเปรียบเทียบค่าร้อยละความถูกต้องในขั้นตอนการฝึกหัดข้อมูลปัจจัยการลักลอบลำเลียงยาเสพติด การทดสอบข้อมูลปัจจัยกการลักลอบลำเลียงยาเสพติด และการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดพบว่าวิธีต้นไม้การตัดสินใจมีค่าร้อยละความถูกต้องมากกว่าวิธี Naïve Bayes ทั้ง 3 ขั้นตอน ทั้งนี้ผลการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดทั้ง 2 วิธีให้ผลอยู่ในระดับปานกลาง ซึ่งเกิดจากปริมาณข้อมูลที่ไม่เพียงพอ การกระจายตัวของข้อมูลที่รวบรวมได้มีการกระจายตัวค่อนข้างสูง เช่น ปัจจัยมูลค่ายาเสพติดที่รวบรวมได้มีมูลค่าที่อยู่

ในช่วงที่กระจายตัวกันมาก และจำนวนของปัจจัยที่ใช้ในการศึกษาครั้งนี้มีน้อยเกินไปทำให้ความละเอียดของปัจจัยน้อยน้อยตามไปด้วย

การศึกษาครั้งนี้ผู้วิจัยคาดว่าผลการศึกษารูปแบบความสัมพันธ์เชิงพื้นที่ เพื่อการคาดการณ์การลักลอบลำเลียงยาเสพติดด้วยการค้นพบความรู้ในฐานข้อมูลจะเป็นประโยชน์ต่อหน่วยงาน และเจ้าหน้าที่ที่เกี่ยวข้องเพื่อนำไปใช้ร่วมกับการตัดสินใจในการตั้งจุดตรวจ และจุดสกัด เพื่อเป็นการลดปัญหายาเสพติดในประเทศ เช่น ในอำเภอไชยปราการที่มีระดับความเสี่ยงมากที่สุด เนื่องจากเป็นอำเภอทางผ่านในการลักลอบลำเลียงยาเสพติด ดังนั้นจึงควรมีการตั้งจุดตรวจสกัดตามเส้นทางคมนาคมเพิ่มเติม

5.2 ข้อเสนอแนะ

จากที่สรุปผลในหัวข้อที่ 5.1 พบว่าปัญหาของการศึกษาครั้งนี้อยู่ที่ปริมาณข้อมูล ดังนั้นควรมีการนำข้อมูลจากแหล่งอื่น ๆ เช่น การเก็บข้อมูลภาคสนามของตำแหน่งการลักลอบลำเลียงยาเสพติด หรือการได้รับทราบถึงข้อมูลส่วนที่เป็นความลับที่มีรายละเอียดมากกว่าข้อมูลจากข่าวการลักลอบลำเลียงยาเสพติด เช่น ข้อมูลภาพจากกล้องวงจรปิดจากตำแหน่งที่เกิดเหตุ ซึ่งจากการเพิ่มข้อมูลภาคสนาม และข้อมูลที่เป็นความลับจะทำให้ปริมาณข้อมูลเพิ่มขึ้น อีกทั้งยังเพิ่มการกำหนดปัจจัยในพื้นที่ให้มีรายละเอียดมากขึ้นอีกด้วย การเพิ่มรายละเอียด และปริมาณข้อมูลจะทำให้การศึกษาความสัมพันธ์เชิงพื้นที่พื้นที่ของปัจจัยที่ส่งผลต่อการลักลอบลำเลียงยาเสพติด และการคาดการณ์พื้นที่เสี่ยงต่อการลักลอบลำเลียงยาเสพติดมีผลร้อยละความถูกต้อง และความแม่นยำมากขึ้น ในส่วนของการระบุตำแหน่งการจับกุมการลักลอบลำเลียงยาเสพติดในการศึกษานี้ใช้ซอฟต์แวร์ Google Earth Pro จึงทำให้ตำแหน่งการจับกุมการลักลอบลำเลียงยาเสพติดที่ได้มีความแม่นยำน้อยกว่าการเก็บข้อมูลภาคสนามของตำแหน่งการจับกุมการลักลอบลำเลียงยาเสพติดจากพื้นที่การจับกุมการลักลอบลำเลียงยาเสพติด และหากต้องการเพิ่มความถูกต้องให้มากขึ้นข้อมูลที่รวบรวมควรมีความมันต่อเหตุการณ์ เช่น ข้อมูลตำแหน่งการจับกุมที่ได้จากการลงพื้นที่ภาคสนาม และข้อมูลภาพถ่ายทางอากาศจากการบินถ่ายภาพด้วยอากาศยานไร้คนขับ เป็นต้น

รายการอ้างอิง

Bertsekas DP., & Tsitsiklis JN. (2008). *Introduction to Probability*. Cambridge, MA: MIT Press.

Cabena P., Hadjinian P., Stadler R., Verhees J., & Zanasi A. (1998). *Discovering Data Mining: From Concepts to Implementation*: Prentice Hall.

Calçada DB., Rezende SO., & Teodoro MS. (2019). Analysis of green manure decomposition parameters in northeast Brazil using association rule networks. *Computers and Electronics in Agriculture, 159*, 34-41.

Dechsiri B. (2017). *DRUG TRAFFICKING ESTIMATION BASED GEO-STATISTICS.* (Master's thesis). Silpakorn University,

Ester M., Kriegel H.-P., & Sander J. (1997). Spatial data mining: A database approach. In *Advances in Spatial Databases, Computer Science 1262* (pp. 47-66). Berlin: Springer.

Fayyad U., Piatetsky-Shapiro G., & Smyth P. (1996). From data mining to knowledge discovery: An overview. In P.-S. G. Fayyad U., Smyth P., and Ulthurusamy R. (Ed.), *Advances in Knowledge Discovery and data Mining* (pp. 1-34). Cambridge, MA: MIT Press.

Fernández-Arteaga V., Tovilla-Zárate CA., F. A., González-Castro TB., Juárez-Rojop IE., López-Narváez L., & Yazmín Hernández-Díaz. (2016). Association between completed suicide and environmental temperature in a Mexican population, using the Knowledge Discovery in Database approach. *COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE, 135*, 219-224.

Ge Q., Xiong F., Xie L., Chen J., & Yu M. (2019). Dynamic interaction of soil – Structure cluster. *Soil Dynamics and Earthquake Engineering, 123*, 16-30.

Hamfelt A., Karlsson M., Thierfelder T., & Valkovsky V. (2011). Beyond K-means: Clusters Identification for GIS. In C. C. Popovich V., Devogele T., Schrenk M., and Korolenko K. (Ed.), *Information Fusion and Geographic Information Systems* (Vol. 5, pp. 93-105). Berlin, Heidelberg: Springer.

Hand DJ. (1998). Data mining Statistic and more. *American Statistician, 52*, 112-118.

Hand DJ., Mannila H., & Smyth P. (2001). *Principles of Data Mining*. Cambridge, MA: MIT Press.

Harvey JM., & Jiawei H. (2009). *Geographic Data Mining and Knowledge Discovery*. Boca Raton, FL: Taylor & Francis Group.

Jarke M., Lenzerini M., Vassiliou Y., & Vassiliadis P. (2003). Data Warehouse Practice: An Overview. . In L. M. Jarke M., Vassiliou Y., and Vassiliadis P. (Ed.), *Fundamentals of Data Warehouses*. Berlin, Heidelberg: Springer.

Kesavaraj G., & Sukumaran S. (2013). *A study on classification techniques in data mining.* Paper presented at the Communications and Networking Technologies, Tiruchengode, India.

Law J., Quick M., & Chan PW. (2015). Analyzing Hotspots of Crime Using a Bayesian Spatiotemporal Modeling Approach: A Case Study of Violent Crime in the Greater Toronto Area. *Geographical Analysis, 47*, 1-19.

Li G., Haining R., Richardson S., & Best N. (2014). Space–time variability in burglary risk: A Bayesian spatio-temporal modelling approach. *Journal of Spatial Statistics, 9*, 180-191.

Medel M., Lu Y., & Chow E. (2015). Mexico's drug networks: Modeling the smuggling routes towards the United States. *Applied Geography, 60*, 240-247.

Nguyen PM., Haghverdi A., Pue J., Botula YD., Le KV., Waegeman W., & Cornelis WM. (2017). Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. *Biosystems Engineering, 153*, 12-27.

Oliveira MM., Camanho AS., Walden JB., Miguéis VL., Ferreira NB., & Gaspar MB. (2017). Forecasting bivalve landings with multiple regression and data mining techniques: The case of the Portuguese Artisanal Dredge Fleet. *Marine Policy, 84*, 110-118.

Peter H., Kasantikul R., Hasan I., Rattanabumrung C., Rungrun P., Suksopee N., . . . Niruntasuk N. (2018). Spatiotemporal Bayesian Networks for Malaria Prediction: Case Study of Northern Thailand. *Journal of Artificial Intelligence in Medicine, 84*, 127-138.

Quinlan JR. (1986). *Induction of Decision Trees*. Boston: Kluwer Academic.

Quinlan JR. (1994). *C4.5: Programs for Machine Learning*. Boston: , Kluwer Academic.

Robert O., Kumsap C., & Janpengpen A. (2018). Simulation of counter drugs operations based on geospatial technology for use in a military training simulator. *International Journal of Simulation and Process Modelling, 13*(4).

Spark CS. (2011). Violent crime in San Antoio, Texas: An application of spatial epidemiological method. *Journal of Spatial and Spatio-temporal Epidemiology, 2*, 301-309.

Thomas MC., Zhu W., & Romagnoli JA. (2018). Data mining and clustering in chemical process databases for monitoring and knowledge discovery. *Journal of Process Control, 67*, 160-175.

United Nations Office on Drugs and Crime. (2018). World Drug Report 2018. Retrieved from https://www.unodc.org/wdr2018

Wong JY, & Chung PH. (2007). Managing valuable Taiwanese airline passengers using knowledge discovery in database techniques. *Journal of Air Transport Management, 13*, 362-370.

Xu C., Bao J., Wang C., & Liu P. (2018). Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. *Journal of Safety Research, 67*, 65-75.

Zhao Y., Li Y., Zhang L., & Wang Q. (2016). Groundwater level prediction of landslide based on classification and regression tree. *GEODESY AND GEODYNAMICS, 7*(5), 348-355.

เศรษฐพงค์ มะลิสุวรรณ. (2553). คลังข้อมูล (Data warehouse). Retrieved from http://www.scimath.org

กองควบคุมวัตถุเสพติด. (2559). Introduction and Historical Backgound. Retrieved from http://narcotic.fda.moph.go.th/english/

นพมาศ ปักเข็ม. (2558). เอกสารประกอบการสอนวิชาการทำเหมืองข้อมูล เวอร์ชั่น 1.0. In. มหาวิทยาลัยทักษิณ, พัทลุง.

สายชล สินสมบูรณ์ทอง. (2558). การทำเหมืองข้อมูล. กรุงเทพมหานคร: บริษัท จามจุรีโปรดักส์ จำกัด.

สำนักงานคณะกรรมการป้องกันและปราบปรามยาเสพติด. (2560). สถานการณ์ยาเสพติดพื้นที่ภาคเหนือห้วงเดือนมกราคม 2560. Retrieved from https://www.oncb.go.th

สำนักงานจังหวัดเชียงราย. (2561). ข้อมูลทั่วไปจังหวัด เชียงราย. Retrieved from http://www.chiangrai.net

สิทธิชัย คำคง. (2558). การทำเหมืองข้อมูล. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, กรุงเทพมหานคร: สาขาวิชาครุศาสตร์เทคโนโลยีคณะครุศาสตร์อุตสาหกรรม.

มหาวิทยาลัยศิลปากร

ประวัติผู้เขียน

ชื่อ-สกุล	สุเมธัส เนียมแก้ว
วัน เดือน ปี เกิด	17 พฤษภาคม 2538
สถานที่เกิด	ราชบุรี
วุฒิการศึกษา	วิทยาศาสตร์สิ่งแวดล้อม คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร (2559)
ที่อยู่ปัจจุบัน	538/11 หมู่ 1 ตำบลไร่ใหม่ อำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์
ผลงานตีพิมพ์	Niamkaeo, S., & Robert, O. (2020). Spatial Relationship of Drug Smuggling in Northern Thailand Using GIS-based Knowledge Discovery. Environment and Natural Resources Journal, xx-xx. Retrieved from https://ph02.tci-thaijo.org/index.php/ennrj/article/view/240574

มหาวิทยาลัยศิลปากร