**การจำแนกพื้นที่จุดเสี่ยงไฟป่าโดยใช้สมการ Support vector machine**

**Random forest และ Gradient boosting machine**

**เจษฎา คำตะนิตย์ 1,ชลนที ชาญสมิง 2,มรกต วรชัยรุ่งเรือง3**

สาขาวิชาภูมิศาสตร์และภูมิสารสนเทศ คณะมนุษยศาสตร์และสังคมศาสตร์

มหาวิทยาลัยราชภัฏสวนสุนันทา

โทรศัพท์ 094-350-8769 e-mail : [s59122401014@ssru.ac.th](mailto:s59122401014@ssru.ac.th)

โทรศัพท์ 063-269-8983 e-mail : [s59122401028@ssru.ac.th](mailto:s59122401028@ssru.ac.th)

โทรศัพท์ 096-515-6223 e-mail : [morakot.wo@ssru.ac.th](mailto:morakot.wo@ssru.ac.th)

**บทคัดย่อ**

การศึกษานี้มีวัตถุประสงค์ เพื่อวิเคราะห์ปัจจัยที่ส่งผลต่อการเกิดไฟป่ามากที่สุด และวิเคราะห์อัลกอริทึมที่มีความสามารถจำแนกและทำนายจุดเกิดไฟป่าได้แม่นยำมากที่สุด ประกอบไปด้วยอัลกอริทึม 3 แบบ คือ Support Vector Machine (SVM), Random forest (RF) และ Gradient boosting machine (GBM) ด้วยวิธี การเรียนรู้ของเครื่องมือ (Machine learning) แบบมีผู้สอน (Supervised learning) ผ่านโปรแกรม RStudio โดยใช้ปัจจัย ความสูง, ความลาดชัน, ทิศด้านลาด, ประเภทของป่า, ดัชนีผลต่างพืชพรรณแบบนอมัลไลซ์ (Normalized difference vegetation index : NDVI), ดัชนีความแตกต่างของน้ำแบบนอมัลไลซ์ (Normalized Difference Water Index : NDWI), ดัชนีความชื้นในดิน (Soil Moisture Index : SMI), อุณหภูมิพื้นผิว (Land Surface Temperature LST ), อุณหภูมิ, ความชื้น, ปริมาณน้ำฝน ระยะห่างจากเส้นถนน และจุดความร้อน(Hotspot) มาวิเคราะห์ความสัมพันธ์ของปัจจัยด้วย การถดถอยเชิงเส้นสองตัวแปร (Regression binary) จากการวิเคราะห์การถดถอยเชิงเส้นสองตัวแปร (Regression binary) พบว่าปัจจัยที่เกี่ยวข้องมี ความสูง, อุณหภูมิ, ความชื้น, ทิศด้านลาดทิศตะวันออกเฉียงใต้, ทิศด้านลาดทิศตะวันตก, ป่าเบญจพรรณและป่าดิบเขา พบว่าอัลกอริทึม Support Vector Machine (SVM) นั้นมีความสามารถทำนายจุดเกิดไฟป่าได้แม่นยำที่ โดยที่ค่าความถูกอยู่ที่ 79.06977 เปอร์เซ็นต์ โดยที่อัลกอริทึม Random forest (RF) นั้นมีค่าความแม่นยำอยู่ที่ 76.74419 เปอร์เซ็นต์ และอัลกอริทึม Gradient boosting machine (GBM) มีความแม่นยำน้อยที่สุด อยู่ที่ 72.09 เปอร์เซ็นต์

**คำสำคัญ :** จุดเสี่ยงไฟป่า/ซัพพอร์ตเวกเตอร์แมชชีน/อัลกอริทึม

**Forest Fire Risk Identification and Classification by using equation Support vector machine, Random forest and Gradient boosting machine**

Jedsada khamtanit 1,Chonnatee Chansaming 2,Morakot Worachairungreung3

Branch of Geography and Geo-informatics, Faculty of Humanities and Social Sciences

Suan Sunandha Rajabhat University

tell 094-350-8769 e-mail : [s59122401014@ssru.ac.th](mailto:s59122401014@ssru.ac.th)

tell 063-269-8983 e-mail : [s59122401028@ssru.ac.th](mailto:s59122401028@ssru.ac.th)

tell 096-515-6223 e-mail : [morakot.wo@ssru.ac.th](mailto:morakot.wo@ssru.ac.th)

**Abstract**

The objective of this study was to analyze the most important casual factors for the incident of wildfire and the most accurate algorithms to identify and predict the wildfire incident. This study focused on exploring three main algorithms, namely: Support Vector Machine (SVM), Random Forest (RF) and Gradient Boosting Machine (GBM) using supervised learning as the machine learning task through RStudio program. In this study, height, slope, aspect, forest type, Normalized difference vegetation index (NDVI), Normalized Difference Water Index (NDWI), Soil Moisture Index (SMI), Land Surface Temperature (LST), temperature, humidity, rainfall, distance from the road line, and burned area were investigated to determine the relationship of factors using Binary Regression. The results of analyzing data with Binary Regression indicated that causal factors for the incident of wildfire were height, temperature, humidity, southeast aspect, west aspect, forest type: mixed deciduous forest and hill evergreen forest. The results showed that Support Vector Machine (SVM) was the algorithm that could provide the highest accurate results for predicting wildfire incident with an accuracy value of 79.06977%, followed by Random Forest (RF) algorithm with an accuracy value of 76.74419%. Gradient Boosting Machine (GBM) was the algorithm that could provide the lowest accurate results for predicting wildfire incident with an accuracy value of 72.09%.

**Keywords :** Fire Forest Risk/ Support Vector Machine/Algorithm

**บทนำ**

ภาคเหนือเป็นภูมิภาคที่มีลักษณะของภูมิประเทศที่เป็นภูเขาสลับซับซ้อนพื้นที่ส่วนใหญ่ของภาคเหนือจะเป็นพื้นที่ป่า ข้อมูลพื้นที่ป่าจากกรมป่าไม้ปี 2561 พบว่าประเทศไทยมีพื้นที่ป่าทั้งหมด 102,488,302.19 ไร่คิดเป็นร้อยละ 31.68 เปอร์เซ็นต์ของพื้นที่ประเทศไทย เมื่อดูเป็นรายภูมิภาคจะพบว่าภาคเหนือมีพื้นที่ป่ามากที่สุด จำนวน 38,533,429.40 ไร่ คิดเป็นร้อยละ 37.60 เปอร์เซ็นต์ รองลงมาคือภาคตะวันตกจำนวน 20,108,513.54 ไร่ คิดเป็นร้อยละ 19.62 เปอร์เซ็นต์ ภาคตะวันออกเฉียงเหนือจำนวน 15,750,098.53 ไร่ คิดเป็นร้อยละ 15.37 เปอร์เซ็นต์ ภาคกลางจำนวน 12,163,869.66 คิดเป็นร้อยละ 11.87 เปอร์เซ็นต์ ภาคใต้จำนวน 11,207,228.70 ไร่ คิดเป็นร้อยละ 10.93 เปอร์เซ็นต์ ภาคตะวันออกจำนวน 4,725,162.36 ไร่ คิดเป็นร้อยละ 4.61 เปอร์เซ็นต์ของพื้นที่ป่าในประเทศไทยตามลำดับ และจากข้อมูลกรมป่าไม้ จังหวัดเชียงใหม่เป็นจังหวัดที่มีพื้นที่ป่ามากที่สุดจำนวน 9,661,526.03 ไร่คิดเป็นร้อยละ 69.82 ของจังหวัด และจากข้อมูลยังพบว่า อำเภอแม่แจ่ม จังหวัดเชียงใหม่ เป็นอำเภอที่มีขนาดใหญ่มากที่สุดจำนวน 1,679,106.28 ไร่ และจากภาพดาวเทียม MODIS ในอำเภอแม่แจ่ม จังหวัดเชียงใหม่ยังพบจุดไฟป่ามากที่สุดจำนวน 71 จุด และจากข้อมูลพบจุดไฟป่ามากที่สุดในช่วงเดือนมีนาคม

ไฟป่าเป็นปัญหาที่เกิดขึ้นทุกปีในจังหวัดเชียงใหม่แม้จะมีการช่วยเหลือจากภาครัฐ แต่ปัญหาของการเกิดไฟป่านั้นไม่สามารถคาดการณ์หรือป้องกันไม่ให้เกิดได้ ทำได้เพียงแก้ไขปัญหาหลังจากเกิดไฟป่าไปแล้ว นอกจากนี้ไฟป่ายังเป็นปัญหาไปสู่การเกิดหมอกควันซึ่งเป็นปัญหาใหญ่ของจังหวัดเชียงใหม่ ซึ่งจังหวัดเชียงใหม่นั้นเป็นสถานที่ท่องเที่ยวที่สำคัญของประเทศและรายได้หลักของประเทศมากจากากรท่องเที่ยว แต่ปัญหาที่เกิดขึ้นมีส่วนทำให้เกิดผลกระทบต่อการท่องเที่ยวเป็นสาเหตุที่ทำให้นักท่องเที่ยวลดน้อยลงแม้จะมีการศึกษาโดยใช้หลักการซ้อนทับของพื้นที่(Overlay) แต่ผลลับที่ได้ไม่สามารถทำนายการเกิดพื้นที่ไฟป่าได้

ดังนั้นการศึกษาครั้งนี้ได้ทำการจำแนกพื้นที่เสี่ยงการเกิดไฟป่าโดยเลือกพื้นที่อำเภอแม่แจ่ม จังหวัดเชียงใหม่ เนื่องจากอำเภอแม่เป็นอำเภอที่มีขนาดใหญ่มากที่สุดและพบจุดเกิดไฟป่ามากที่สุดเป็นจำนวน 71 จุดในเดือนมีนาคม หลังจากการวิธีการวิเคราะห์ใช้ระบบสารสนเทศทางภูมิศาสตร์และการสำรวจระยะไกลมาวิเคราะห์โดยการนำอัลกอริทึม Support vector machine(SVM),Random forest(RF) และ Gradient boosting machine(GBM) โดยนำปัจจัยที่ส่งผลต่อการเกิดไฟป่าในพื้นที่ศึกษา อำเภอแม่แจ่ม จังหวัดเชียงใหม่ มาวิเคราะห์เพื่อหาอัลกอริทึมที่มีความแม่นยำและถูกต้องมากที่สุด เพื่อใช้เป็นแนวทางในการเลือกอัลกอริทึมมาใช้ในการคำนวณและวิเคราะห์พื้นที่ต่อไป

**วัตถุประสงค์**

1. วิเคราะห์ปัจจัยที่ส่งผลต่อการเกิดไฟป่าในพื้นที่ อำเภอแม่แจ่ม จังหวัดเชียงใหม่
2. วิเคราะห์อัลกอริทึมที่มีความสามารถจำแนกและทำนายจุดเกิดไฟป่าได้แม่นยำมากที่สุด

**ทบทวนวรรณกรรม**

ไฟป่าหมายถึงไฟที่ปราศจากการควบคุม และสามารถลุกลามได้อย่างอิสระ แล้วเผาผลาญเชื้อเพลิงธรรมชาติในป่าได้แก่อินทรียวัตถุที่สลายตัวแล้วและอินทรียวัตถุที่กำลังสลายตัว ใบไม้แห้ง หญ้า กิ่งก้านไม้ที่แห้ง ท่อนไม้ ตอไม้ วัชพืช ไม้พุ่ม ใบไม้สด และในระดับหนึ่งสามรถเผาผลาญต้นไม้ที่ยังมีชีวิตอยู่ สำหรับประเทศไทยนั้น เพื่อความเหมาะสมกับสภาพปัญหาและขอบเขตการจัดการไฟป่าจึงกำหนดคำนิยามของไฟป่าว่า “ไฟที่เกิดจากสาเหตุใดก็ตาม แล้วลุกไปได้โดยอิสระปราศจากการควบคุม ทั้งนี้ไม่ว่าไฟนั้นจะเกิดขึ้นในป่าธรรมชาติหรือสวนป่า”โดยไฟป่าที่เกิดจากการขึ้นจากการเผาไหม้ ซึ่งมีองค์ประกอบที่สำคัญ 3 ประการคือ ความร้อน เชื้อเพลิง และก๊าซออกซิเจน ดังนี้ 1. ความร้อน คือแหล่งความร้อนที่ทำให้เกิดไฟป่าแบ่งเป็น 2 ประเภทคือ แหล่งความร้อนตามธรรมชาติ และแหล่งความร้อนจากมนุษย์ซึ่งจุดไฟด้วยสาเหตุต่างกัน 2. เชื้อเพลิง อินทรียสารทุกชนิดที่ก่อให้เกิดไฟได้ เช่น ต้นไม้ กิ่งไม้ ใบไม้ นตอไม้ ลูกไม้ หญ้าและวัชพืช 3. ออกซิเจน ออกซิเจนเป็นองค์ประกอบทั่วไปในป่าจึงมีออกซิเจนการะจายอยู่อย่างสม่ำเสมอในป่า โดยองค์ประกอบทั้ง 3 ประการนี้เรียกว่า ทฤษฎีสามเหลี่ยมไฟ หากขาดองค์ประกอบใดองค์ประกอบหนึ่ง ไฟป่าจะไม่เกิดขึ้น (พัฒนะพงษ์, 2550)

จากทฤษฎีสามเหลี่ยมไฟส่งผลต่อพฤติกรรมของไฟป่า โดยปัจจัยหลักที่มีผลต่อพฤติกรรมของไฟป่า ที่มีอยู่ 3 ปัจจัย ได้แก่ ลักษณะเชื้อเพลิง ลักษณะอากาศ และลักษณะภูมิประเทศ ซึ่งสามารถอธิบายได้ดังนี้ 1. เชื้อเพลิง ลักษณะของเชื้อเพลิงที่แตกต่างกันจะมีผลให้พฤติกรรมของไฟป่าทั้งสามารถแยกลักษณะของเชื้อเพลิงที่มีผลต่อพฤติกรรมของไฟป่าได้ 4 ลักษณะ คือ 1.1 ความชื้นของเชื้อเพลิง หากเชื้อเพลิงมมีความชื้นสูงย่อมยากต่อการติดไฟ และการลุกลามก็เป็นไปได้ช้ากว่าความชื้นต่ำ 1.2 ขนาดของเชื้อเพลิง เชื้อเพลิงขนาดเล็ก เช่น กิ่งไม้เล็ก ใบไม้ สามารถลุกลามได้รวดเร็วกว่าเชื้อเพลิงขนาดใหญ่เช่น ท่อนไม้ขนาดใหญ่ ไม้ยืนต้นตาย 1.3 ปริมาณเชื้อเพลิง หากมีเชื้อเพลิงมาก ไฟย่อมมีความรุนแรง 1.4 ความต่อเนื่องของเชื้อเพลิง เป็นปัจจัยสำคัญที่ทำให้ไฟลุกลามได้เร็วหรือช้า 2. ลักษณะภูมิอากาศ ที่สำคัญและมีผล 2.1 อุณหภูมิ มีผลต่อความช้าเร็วในการแห้งของเชื้อเพลิง อุณหภูมิสูงเชื้อเพลิงย่อมแห้งได้เร็วกว่าอุณภูมิต่ำ 2.2 ความชื้น ถ้าอากาศมีความชื้นสูง ย่อมทำให้เชื้อเพลิงมีความชื้นสูงตามไปด้วยจึงยากต่อการติดไฟ 3. ภูมิประเทศ ที่มีอิทธิพล 3.1 ความลาดชัน เป็นตัวกำหนดอัตราความเร็ว และทิศทางของไฟ โดยปกติไฟจะลุกลามขึ้นไปตามความลาดชันเขาอยู่เสมอ 3.2 ทิศด้านลาด ด้านลาดชันขของภูเขาซึ่งรับแสงได้มากกว่า เชื้อเพลิงจะแห้งมากกว่าและลุกลามได้เร็วกว่า (พัฒนะพงษ์, 2550)

ซัพพอร์ตเวกเตอร์แมชชีน (Support vector machine : SVM) เป็นเทคนิคหนึ่งที่ใช้สำหรับการจำแนกประเภทหรือการจัดหมวดหมู่ข้อมูล โดยการวิเคราะห์จะเป็นแนวทางของการคำนวณหาเส้นแบ่งกลุ่มที่เป็นเชิงเส้นและ ไม่เป็นเชิงเส้นโดยพิจารณาจากจุดของข้อมูลที่อยู่ในกลุ่มที่แตกต่างกลุ่มเหล่านี้เรียกว่า ซัพพอร์ตเวกเตอร์ซึ่งจะนำมาหาสมาการเส้นแบ่งกลุ่มที่เหมาะสมต่อไป SVM จะมีทั้งการจำแนก ประเภทแบบไบนารี เพื่อสร้างสมาการเส้นตรงในการแบ่งข้อมูลออกเป็น 2 กลุ่ม และการจำแนกประเภทแบบหลายกลุ่ม เพื่อแก้ปัญหาการจำแนกประเภทที่มีมากกว่าสองกลุ่ม โครงสร้างของ SVM ประกอบด้วย 2 ส่วนหลักคือ การเพิ่มระยะการจำแนกมากสุด และการลดข้อผิดพลาดให้ต่ำสุด(ฌานิกา, 2561)

แรนดอมฟอเรส (Random Forest : RF) เป็นชุดการจำแนกแบบไม้ตัดแต่งกิ่ง ซึ่งถูกสร้างเพื่อนำข้อมูลฝึกสอนไปสุ่มเลือกตัวอย่างข้อมูลและคุณลักษณะของข้อมูลแล้วนำมาสร้างเป้นต้นไม้ตัดสินใจ ซึ่งมีตัวอย่างส่วนหนึ่งที่ไม่ถูกและจะถูกนำมาใช้ทดสอบในการตัดสินใจ

ซึ่งจากงานวิจัยได้มีการทำวิจัยโดยหาความสัมพันธ์ โดยค่าความสัมพันธ์(R2)จะต้องมีค่า 0.6 ขึ้นไปซึ่งจะเป็นค่าที่ยอมรับได้ในสายงานวิทยาศาสตร์ โดยค่ายิ่งไกล 1 มากเท่าไรจะยิ่งมีความสัมพันธ์กันมาก(Moksony,1990)

**วิธีดำเนินงานวิจัย**

1. **อุปกรณ์ที่ใช้ในการศึกษา**
   1. โปรแกรมสารสนเทศทางภูมิศาสตร์ และโปรแกรม RStudio version 3.6.1
   2. ข้อมูลจุดเกิดไฟป่า ปี 2562
   3. ภาพถ่ายดาวเทียม Landsat 8 เดือน มีนาคม ปี 2562
   4. ข้อมูลด้านพื้นที่ของอำเภอแม่แจ่มจังหวดเชียงใหม่ (ข้อมูลที่ตั้งและอาณาเขต ปริมาณน้ำฝน อุณหภูมิ เส้นถนน แบบจำลองภูมิประเทศเชิงเลข ข้อมูลการใช้ประโยชน์จากที่ดิน ลักษณะภูมิอากาศ )
2. **พื้นที่ศึกษา**

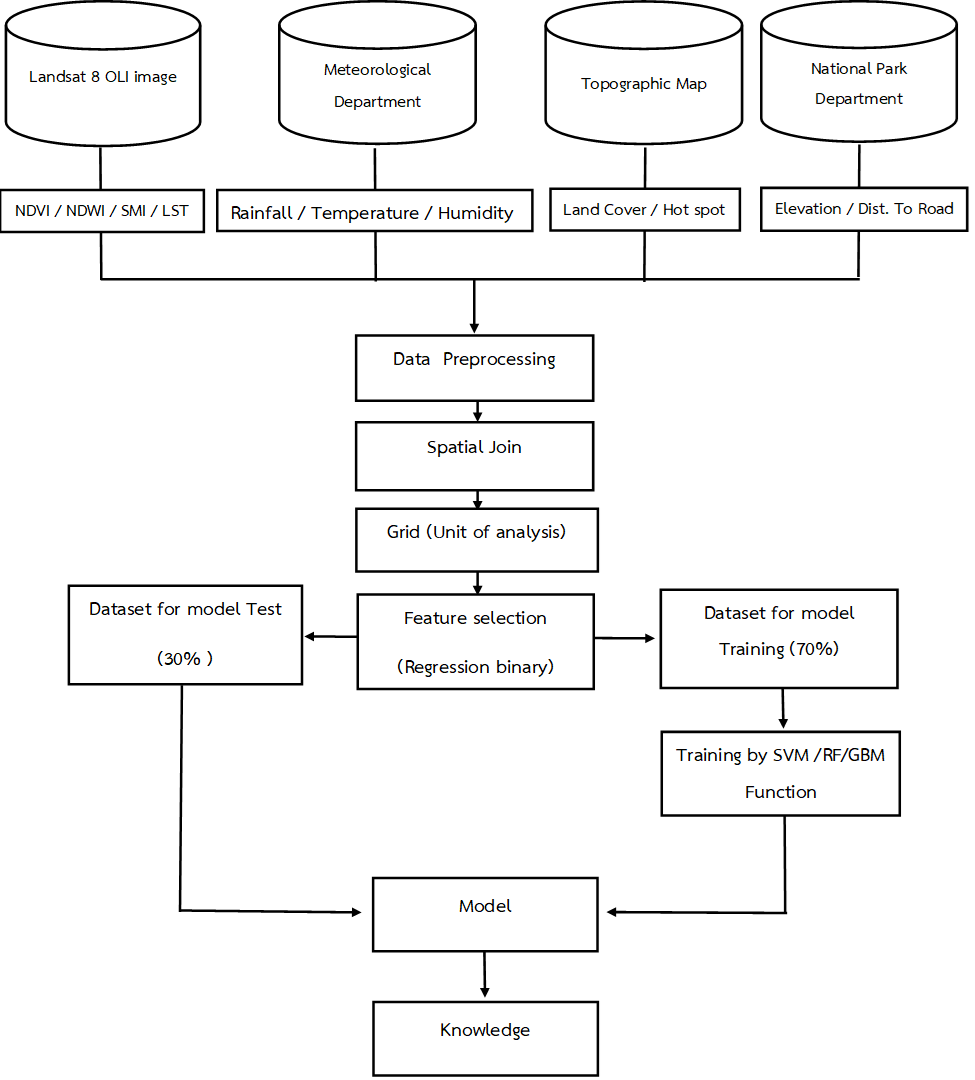
อำเภอแม่แจ่มมีอาณาเขตทิศเหนือติดกับอำเภอปาย จังหวัดแม่ฮ่องสอน ทิศตะวันออก ติดกับอำเภอสะเมิง อำเภอแม่วาง และอำเภอจอมทอง จังหวัดเชียงใหม่ ทิศใต้ติดกับอำเภอฮอด จังหวัดเชียงใหม่ ทิศตะวันตกติดกับอำเภอแม่ลาน้อย อำเภอขุนยวม และอำเภอเมืองแม่แม่ฮ่องสอน จังหวัด



**รูป 1** พื้นที่ศึกษา อำเภอแม่แจ่ม จังหวัดเชียงใหม่

**3. ขั้นตอนการเตรียมข้อมูล**

วัตถุประสงค์ของงานวิจัย วิเคราะห์อัลกอริทึมที่มีความสามารถจำแนกและทำนายจุดเกิดไฟป่าได้แม่นยำมากที่สุด โดยประกอบไปด้วยอัลกอริทึม 3 แบบ คือ Support Vector Machine (SVM), Random forest (RF) และ Gradient boosting machine (GBM) โดยจะมีวิธีดำเนินการดังนี้



**Objective 1**

**Objective 2**

**รูปที่** 2 แผนผังขั้นตอนการศึกษา

* 1. **ขั้นตอนการเตรียมข้อมูล** 
     1. ชุดข้อมูลภาพถ่าย ดาวเทียมสำรวจทรัพยากรธรรมชาติที่มีชื่อว่า Landsat-8 OLI/TIR path 131, row 48 ช่วงวันที่ 19 มีนาคม ปี พ.ศ. 2562 เนื่องจากเป็นช่วงที่อุณหภูมสูงที่สุดของปี สุ่มเสี่ยงเกิดไฟป่ามากที่สุด นำเข้ามาปรับแก้เชิงเลขในโปแกรมระบบสารสนเทศทางภูมิศาสตร์ (Raster calculator)
     2. ข้อมูลจุดเกิดไฟป่าช่วงเดือนเมษายน ปีพ.ศ. 2562 จากสำนักป้องกัน ปราบปราม และควบคุมไฟป่า คัดเลือกเฉพาะในพื้นที่ศึกษา
     3. ข้อมูลสิ่งปกคลุมดิน ประกอบด้วยข้อมูลประเภทป่าประเทศไทย จากกรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช คัดเลือกในพื้นที่ศึกษา
     4. ข้อมูลเส้นถนนประเทศไทย จาก WFPGeoNODE คัดเลือกเส้นถนนในพื้นที่ศึกษา แล้ว buffer ระยะห่างจากถนน 500, 1000, 1500, 2000 และ มากกว่า 2000 เมตร ตามลำดับ
     5. ข้อมูลภูมิประเทศ ประกอบด้วย แบบจำลองภูมิประเทศเชิงเลข Digital Elevation Model (DEM) จาก United States Geological Survey มาร้างความลาดชัน(Slope) และ ทิศด้านลาด(Aspect) ในโปรแกรมสารสนเทศทางภูมิศาสตร์
     6. ข้อมูลภูมิอากาศ ประกอบ ด้วยอุณหภูมิเฉลี่ยรายเดือนหน่วยเป็นองศาเซลเซียส ปริมาณน้ำฝนรายเดือนหน่วยมิลลิเมตร ปริมาณความชื้นหน่วยเป็นเปอร์เซ็น จากกรมอุตุนิยมวิทยา ทำการประมาณค่าในช่วงรูปแบบ Inverse Distance Weighted (IDW) จากจุดสถานีตรวจวัดอากาศ
  2. **ขั้นตอนการวิเคราะห์**
     1. นำข้อมูลปัจจัยทั้งหมดมาจัดเก็บอยู่ในรูปแบบกริดขนาด 30x30 เมตร ด้วยเครื่องมือ Spatial Join รวมข้อมูลปัจจัยทั้งหมดอยู่ในตารางกริดเดียวกัน จัดเก็บข้อมูลในรูปแบบตาราง
     2. วิเคราะห์ข้อมูลโดยใช้ Logistic Regression Analysis ในการวิเคราะห์ ด้วยโปรแกรม RStudio ใช้การถดถอยเชิงเส้นสองตัวแปร (Regression binary) หาความสัมพันธ์ ระหว่างปัจจัยที่ส่งผลต่อการเกิดไฟป่าจาก 14 ปัจจัย ได้แก่ ความสูง, ความลาดชัน, ทิศด้านลาด, ประเภทของป่า, ดัชนีผลต่างพืชพรรณแบบนอมัลไลซ์ (Normalized difference vegetation index : NDVI), ดัชนีความแตกต่างของน้ำแบบนอมัลไลซ์ (Normalized Difference Water Index : NDWI), ดัชนีความชื้นในดิน (Soil Moisture Index : SMI), อุณหภูมิพื้นผิว (Land Surface Temperature LST ), อุณหภูมิ, ความชื้น, ปริมาณน้ำฝน, ระยะห่างจากเส้นถนน, ปริมาณน้ำฝน, และจุดความร้อน(Hotspot) ด้วยการถดถอยเชิงเส้นสองตัวแปร (Regression binary) ดังสมการที่ 1

(1)

เมื่อ

= จุดความร้อน

𝑒 = ลอการิทึมธรรมชาติ มีค่าประมาณ 2.7182

𝛽 = ค่าพารามิเตอร์ตัวแบบ

𝑥 = ปัจจัยสิ่งแวดล้อม



***รูปที่ 3*** *แผนที่เฉพาะปัจจัยที่หาความสัมพันธ์แล้ว* *(a)สิ่งปกคลุมดิน (b)ทิศด้าน (c)อุณหภูมิ (d)ความสูง (e)ความชื้น*

3.2.3 นำปัจจัยทั้ง 5 ปัจจัยจาก การถดถอยเชิงเส้นสองตัวแปร (Regression binary) มาวิเคราะห์อัลกอริทึมที่มีความสามารถจำแนกและทำนายจุดเกิดไฟป่าที่ดีที่สุด ได้แก่ Support Vector Machine (SVM), Random forest (RF) และ Gradient boosting machine (GBM) โดยแบ่งข้อมูลพื้นที่ศึกษา เป็นสองส่วน ส่วนที่ 1 ให้เป็นชุดข้อมูลการสอน(Training data) และ ชุดที่ 2 เป็นข้อมูลตรวจสอบ(Test data) อัตราส่วนอยู่ที่ 70 ต่อ 30 เปอร์เซ็นต์ ตามลำดับ นำชุดข้อมูลการสอน(Training data) นำเข้ากระบวนการ การเรียนรู้ของเครื่องมือ (machine learning) โดยใช้วิธีการเรียนรู้แบบมีผู้สอน (Supervised learning) วิเคราะห์ด้วย อัลกอริทึม 3 แบบ คือ Support Vector Machine (SVM), Random forest (RF) และ Gradient boosting machine (GBM) ในการวิเคราะห์ การแบ่งประเภท (Classification) และทำนายผล (Regression)

* + 1. ตรวจสอบความแม่นยำของอัลกอริทึมทั้ง 3 แบบ โดยใช้ข้อมูล ข้อมูลตรวจสอบ (Test data) มาการตรวจสอบความถูกต้อง

**ผลการวิจัย**

จากการวิเคราะห์การถดถอยเชิงเส้นสองตัวแปร (Regression binary) พบว่าปัจจัยที่เกี่ยวข้องมีทั้งหมด 5 ปัจจัยประกอบด้วย ความสูง, อุณหภูมิ, ความชื้น, ทิศด้านลาด, ประเภทของป่า พบว่าค่า R2 เท่ากับ 0.8093 อยู่ในเกณฑ์ที่ยอมรับได้ โดยที่ค่า R2 เข้าใกล้ 1 แสดงว่าข้อมูลมีความสัมพันธ์กัน (Moksony F, 1990)

**ตารางที่ 1 ปัจจัยที่ส่งผลต่อการเกิดไฟป่าโดยแบบสมการ Logistic Regression**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | Estimate s | td.error stat | istic | p.value |
| (Intercept) | 16.1 | 7.41 | 2.17 | 3.17e- 2 |
| humidity | -0.674 | 0.051 | -13.2 | 9.99E-28 |
| Temperature | 0.586 | 0.259 | 2.26 | 2.49e- 2 |
| Northeast | -0.232 | 0.0639 | -3.63 | 3.72e- 4 |
| East | -0.357 | 0.0487 | -7.33 | 9.36E-12 |
| Souteast | -0.235 | 0.0533 | -4.4 | 1.89e- 5 |
| West | 0.128 | 0.0816 | 1.56 | 1.20e- 1 |
| Montane Forest | -0.412 | 0.0545 | -7.56 | 2.61E-12 |
| Pine Forest | -0.489 | 0.112 | -4.35 | 2.32e- 5 |
| Mixed Deciduous Forest | -0.164 | 0.0569 | -2.88 | 4.46e- 3 |
| Elevation | -0.172 | 0.11 | -1.56 | 1.21e- 1 |

ด้วยอัลกอริทึม 3 แบบ Support Vector Machine (SVM), Random forest (RF) และ Gradient boosting machine (GBM) พบว่าอัลกอริทึม Support Vector Machine (SVM) นั้นมีความสามารถทำนายจุดเกิดไฟป่าได้แม่นยำที่ โดยที่ค่าความถูกอยู่ที่ 79.06977 เปอร์เซ็นต์ โดยที่อัลกอริทึม Random forest (RF) นั้นมีค่าความแม่นยำอยู่ที่ 76.74419 เปอร์เซ็นต์ และอัลกอริทึม Gradient boosting machine (GBM) มีความแม่นยำน้อยที่สุด อยู่ที่ 72.09 เปอร์เซ็นต์



**รูปที่ 5** ตารางเปรียบเทียบ อัลกอริทึมทั้ง 3 แบบ

**รูปที่ 4** ตารางแบบจำลองการจำแนกไฟป่า

**อภิปายผล**

การศึกษาวิเคราห์ปัจจัยด้วยสมการ การถดถอยเชิงเส้นสองตัวแปร (Regression binary) ในพื้นที่ศึกษา อำเภอแม่แจ่ม จังหวัดเชียงใหม่ พบว่าปัจจัยในการทำแบบจำลองนั้นส่งผลต่อการเกิดไฟป่า มี 5 ปัจจัยประกอบด้วย ความสูง, อุณหภูมิ, ความชื้น, ทิศด้านลาด และประเภทของป่า ค่า R2 มีค่าเท่ากับ 0.8093 โดยที่ทิศด้านลาดที่สัมพันธ์กับการเกิดไฟป่าจะพบทิศด้านลาดดังนี้ ทิศด้านลาดทิศตะวันออกเฉียงเหนือ, ทิศด้านลาดทิศตะวันออก, ทิศด้านลาดตะวันออกเฉียงใต้ และทิศด้านลาดตะวันตก, ส่วนประเภทของป่าจะพบอยู่สามประเภทได้แก่ ป่าเบญจพรรณ, ป่าสนเขาและป่าดิบเขา อยู่ในเกณฑ์ที่ยอมรับได้ และพบว่า อัลกอริทึมที่ให้ค่าความแม่นยำมากที่สุด คือ Support Vector Machine (SVM) โดยค่าความถูกต้องอยู่ที่ 79.06977 เปอร์เซ็นต์

**ข้อเสนอแนะ**

ขอเสนอแนะในการนำไปใช้ประโยชน์

* 1. สามารถนำแบบจำลอง Support Vector Machine(SVM) ไปเป็นอัลกอริทึมสำหรับใช้ในการคำนวณ
  2. สามารถพัฒนาเพื่อใช้ในการป้องกันการเกิดไฟป่าได้

ขอเสนอแนะในการทำวิจัยครั้งต่อไป

* 1. สามารถใช้โปรแกรม python เพื่อดึงข้อมูลสภาพภูมิอากาศ เช่น ปริมาณน้ำฝน อุณหภูมิ ความชื้นเป็นรายวัน เพื่อมาใช้เป็นปัจจัยที่ใช้ในการวิเคราะห์เพื่อใช้ในการประเมินพื้นที่จุดเสี่ยงเป็นวันต่อวันได้ ส่วนข้อมูลเชิงพื้นที่อื่น มีการเปลี่ยนแปลงเล็กน้อยจึงไม่จำต้องเปลี่ยนบ่อย
  2. ควรใช้คอมพิวเตอร์ที่มีประสิทธิสูงในการทำงานเนื่องจากข้อมูลมีปริมาณของข้อมูลที่เยอะมาก หากใช้คอมพิวเตอร์ที่มีประสิทธิภาพต่ำจะทำให้การประมวลผลล่าช้าหรือผิดพลาดได้

**กิตติกรรมประกาศ**

ขอขอบคุณสำนักป้องกัน ปราบปราม และควบคุมไฟป่าจากกรมอุทยานแห่งชาติ สัตว์ป่าและพันธุ์พืชที่ให้ความอนุเคราะห์ที่ให้ข้อมูลจุดเกิดไฟป่าของจังหวัดเชียงใหม่ และปัจจัยที่มีผลต่อการเกิดไฟป่า

ขอขอบคุณกรมอุตินิยมวิทยาที่ให้ความอนุเคราะห์ข้อมูลอุณหภูมิ ความชื้น และปริมาณน้ำฝนย้อนหลังของภาคเหนือ

ขอขอบคุณ อาจารย์มรกต วรชัยรุ่งเรือง ที่ให้คำปรึกษาและชี้แนะแนวทางในการทำวิจัยครั้งนี้

**เอกสารอ้างอิง**

ฌานิกา สุขวัฒนวิจิตร. (2561). แบบจำลอง GA-SVM สำหรับการจำแนกการใช้ประโยชน์ที่ดินด้วยข้อมูลดาวเทียม

ระบบ SAR และ Optical. วารสารสมาคมสำรวจข้อมูลระยะไกลและสารสนเทศภูมิศาสตร์แห่งประเทศไทย, 18 (ฉบับพิเศษ), 1-10.

พัฒนะพงษ์ จันทร์คำ. (2550). การประยุกต์ใช้ระบบสารสนเทศภูมิศาสตร์ และการสำรวจข้อมูลระยะไกล

ประเมินพื้นที่เสี่ยงต่อการเกิดไฟป่า ในหน่วยจัดการแม่หวดของงป่าสาธิตแม่งาว อำเภองาว จังหวัดลำปาง. (วิทยานิพน์ปริญญามหาบัณฑิต, มหาวิทยาลัยเกษตรศาสตร์).

ภรัณยา ปาลวิสุทธิ.(2559).การเพิ่มประสิทธิภาพเทคนิคต้นไม้ตัดสินใจบนชุดข้อมูลที่ไม่สมดุล โดยวิธีการสุ่มเพิ่มตัวอย่างกลุ่มน้อยสำหรับข้อมูล การเป็นโรคติดอินเทอร์เน็ต. วรสารเทคโนโลยีสารสนเทศ,12(1), 57.

ศราวุธ ลักษวุธและโสภณวิชญ์ คำพิลัง. (2558). การจำแนกพื้นที่เสี่ยงต่อการเกิดไฟป่าโดยใช้กระบวนการ

วิเคราะห์ภาพเชิงวัตถุ กรณีศึกษาอุทยานแห่งชาติ น้ำตกตาดโตน ตำบลนาฝาย อำเภอเมือง จังหวัดชัยภูมิ. (วิทยานิพนธ์ปริญญามหาบัณฑิต, มหาวิทยาลัยมหาสารคาม).

Moksony, F. (1990). Small is beautiful. The use and interpretation of R2 in social research. Szociológiai Szemle, Special issue, 130-138.