



การทำแผนที่การใช้ที่ดินด้วยภาษา R และการเรียนรู้ของเครื่องจักร:
การวิเคราะห์ข้อมูลจากการสำรวจระยะไกลแบบพาสซีฟและแอคทีฟในเมืองขอนแก่น

LAND USE MAPPING WITH R AND MACHINE LEARNING:
ANALYZING PASSIVE AND ACTIVE REMOTE SENSING DATA IN KHON KAEN CITY

วันวิสาข์ อินทปาสาน¹, ชาติชาย ไวยสุระสิงห์^{2*} และชุตินา ไวยสุระสิงห์³

¹นักศึกษา หลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

²รองศาสตราจารย์ ภาควิชาวิศวกรรมโยธา ศูนย์วิจัยและพัฒนาโครงสร้างพื้นฐานอย่างยั่งยืน

คณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น

³อาจารย์ คณะศิลปศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

*Corresponding author; E-mail: fceews@kku.ac.th

บทคัดย่อ

การศึกษานี้มุ่งเน้นที่การจำแนกภาพจากดาวเทียม Sentinel-2, Sentinel-1 และภาพจากดาวเทียมทั้งสองที่ผสมกัน โดยใช้ภาษา R และเทคนิคการเรียนรู้ของเครื่องจักรหลากหลายแบบ ได้แก่ วิธีการที่ใกล้เคียงที่สุด (KNN), โครงข่ายประสาทเทียม (ANN), เครื่องจักรที่สนับสนุนเวกเตอร์ (SVM), ป่าแบบสุ่ม (RF) และต้นไม้การตัดสินใจ (DT) เทคนิคเหล่านี้ถูกนำไปใช้กับข้อมูลภาพจากดาวเทียมทั้งแบบ passive และ active จากเมืองขอนแก่นในปี ค.ศ. 2020 เพื่อประเมินความแม่นยำในการจำแนก ผลการทดสอบแสดงว่าภาพจากดาวเทียม Sentinel-2 สามารถจำแนกได้ด้วยความแม่นยำสูงสุดที่ 92%, ซึ่งเป็นผลการทดสอบที่มีความแม่นยำสูงสุด นอกจากนี้เรายังพบว่าการเพิ่มปริมาณข้อมูลการฝึกอบรมสามารถเพิ่มความแม่นยำในการจำแนกของภาพจากดาวเทียม Sentinel-2 และ Sentinel-1 ผลการศึกษาของเรายืนยันว่าภาษา R มีศักยภาพที่จะนำไปสู่การประยุกต์ใช้ในการวิเคราะห์การเจริญเติบโตของเมืองหากมีข้อมูลที่เหมาะสม ซึ่งสะท้อนถึงบทบาทสำคัญของการเรียนรู้ของเครื่องจักรสามารถเล่นในการส่งเสริม

คำสำคัญ: ภาษา R; เทคนิคจักรกลเรียนรู้; ข้อมูลรับรู้ระยะไกลเชิงผสมระหว่างแบบพาสซีฟและแบบแอคทีฟ; การใช้ที่ดิน

Wanwisa Inthapasan¹, Chattichai Waisurasingha^{2*} and Chutima Waisurasingha³

¹Master Student in Civil Engineering Program, Fac. of Engineering, KhonKhon University, Thailand.

²Associate Professor, Sustainable Infrastructure Research and Development Center, Fac. of Engineering, KhonKhon University, Thailand.

³Lecturer, Faculty of Liberal Arts, King Mongkut's Institute of Technology Ladkrabang, Thailand.

ABSTRACT

This study focuses on the classification of Sentinel-2, Sentinel-1, and their hybrid satellite imagery using the R programming language and a range of machine learning techniques, including K-nearest neighbors (KNN), Artificial Neural network (ANN), Support Vector Machine (SVM), Random Forest (RF), and Decision Tree (DT). These techniques were applied to both passive and active satellite image data from Khon Kaen city in 2020 to evaluate their classification accuracy. Results showed that Sentinel-2 satellite images could be classified with a peak accuracy of 92%, the highest among the tested data sets. We also found that increasing the volume of training data could potentially enhance the classification accuracy for both Sentinel-2 and Sentinel-1 imagery. The findings underscore the potential of using R for further applications in urban growth analysis, given the right data. This highlights the significant role that machine learning can play in advancing our understanding of urban landscapes through remote sensing data analysis.

KEYWORD: R-Language; Machine Learning Technique; Hybrid of Passive and Active Remotely Sensed Data; Land use/cover map

1. บทนำ

งานศึกษาในอดีต[1][2][3][4]ได้ชี้ชัดตรงกันว่า การวิเคราะห์การเปลี่ยนแปลงการใช้ที่ดินจากข้อมูลการรับรู้ระยะไกลโดยใช้ข้อมูลภาพถ่ายดาวเทียมและแผนที่ ซึ่งเป็นภาพถ่ายดาวเทียมสำรวจทรัพยากรโลกแบบพาสดิฟที่เก็บข้อมูลการสะท้อนของแสงจากดวงอาทิตย์และการแผ่รังสีของวัตถุต่าง ๆ บนผิวโลก โดยการวิเคราะห์ข้อมูลดาวเทียมและแผนที่ในช่วงที่ผ่านมาใช้ข้อมูลในช่วงคลื่นที่ตามองเห็นและช่วงคลื่นอินฟราเรดใกล้ (visible and infrared band) มาเป็นข้อมูลหลักในการจำแนก การศึกษาถึงการเปลี่ยนแปลงการใช้ที่ดินในมิติการผสมผสานข้อมูลภาพถ่ายดาวเทียมแบบพาสดิฟและแบบแอคทีฟจึงเป็นประเด็นที่น่าสนใจ โดยเฉพาะอย่างยิ่งการวิเคราะห์ในบริบทของพื้นที่เมืองขอนแก่นที่เป็นเมืองแห่งแรกในเขตภูมิภาคของประเทศไทยที่ประกาศตัวในการขับเคลื่อนเป็นนครอัจฉริยะ (Smart City) ซึ่งการศึกษาของการวิเคราะห์การเปลี่ยนแปลงการใช้ที่ดินของเมืองขอนแก่นได้ถูกวิเคราะห์โดย[5]ได้มีการวิเคราะห์ข้อมูลภาพถ่ายดาวเทียมแบบพาสดิฟด้วยเทคนิคการจำแนกด้วยจักรกลเรียนรู้ (Machine Learning) โดยใช้ตัวจำแนกแบบจักรกลเวกเตอร์ค้ำยัน (Support Vector Machine: SVM) ซึ่งให้ความถูกต้องที่สูงกว่า 90% ซึ่ง สูงกว่าการใช้ตัวจำแนกแบบความน่าจะเป็นได้สูงสุด (Maximum Likelihood Classifier: MLC) โดยผลการศึกษาของ[5]ได้แสดงให้เห็นถึงการประยุกต์ใช้ ตัวจำแนกแบบจักรแห่งการเรียนรู้ (Machine Learning) ด้วยตัวจำแนกแบบจักรกลเวกเตอร์ค้ำยัน (Support Vector Machine) มาพิจารณาข้อมูลที่ผสมผสานระหว่างแบบพาสดิฟและแบบแอคทีฟ ที่เป็นการรวมกันของข้อมูลภาพถ่ายดาวเทียมและแผนที่เข้ากับข้อมูลภาพถ่ายดาวเทียมเรดาร์ช่องเปิดสังเคราะห์ในช่วงคลื่น C band นั้น สามารถช่วยเพิ่มความถูกต้องโดยรวมของการจำแนกได้เกินกว่า 90% ซึ่งสูงกว่าการใช้เพียงข้อมูลแผนที่เพียงอย่างเดียว โดยเหตุนี้เองจึงเป็นที่น่าสนใจอย่างยิ่ง ที่จะศึกษาถึงขีดความสามารถในการประยุกต์ใช้ตัวจำแนกแบบจักรแห่งการเรียนรู้ประเภทอื่นๆ อย่าง Single Decision Tree, Random Forest เข้ามาศึกษาเปรียบเทียบกับตัวจำแนกแบบจักรกลเวกเตอร์ค้ำยัน

ในการศึกษาวิจัยครั้งนี้ จึงสนใจอย่างยิ่งที่จะศึกษาการจำแนกข้อมูลภาพถ่ายดาวเทียมด้วยภาษา R ในการวิเคราะห์การเปลี่ยนแปลงการใช้ที่ดิน โดยเน้นไปที่เมืองขอนแก่น ทั้งนี้เนื่องจากมีคาดการณ์กันว่า ในอนาคต 2 ใน 3 ของจำนวนประชากรโลก จะอาศัยอยู่ในชุมชนเมือง ส่งผลให้ในปัจจุบันที่ทิศทางการบริหารงานของรัฐบาลไทย มีแนวนโยบายในการขับเคลื่อนเมืองใหญ่ๆ ในประเทศเป็นนครอัจฉริยะ หรือ หรือที่เรียกว่า “สมาร์ท ซิตี้” (Smart City) และเมืองขอนแก่นเองก็ประกาศตัวเอง ในฐานะเมืองแห่งแรกในภาคตะวันออกเฉียงเหนือของประเทศไทยที่จะพัฒนาไปสู่ความเป็นนครอัจฉริยะ

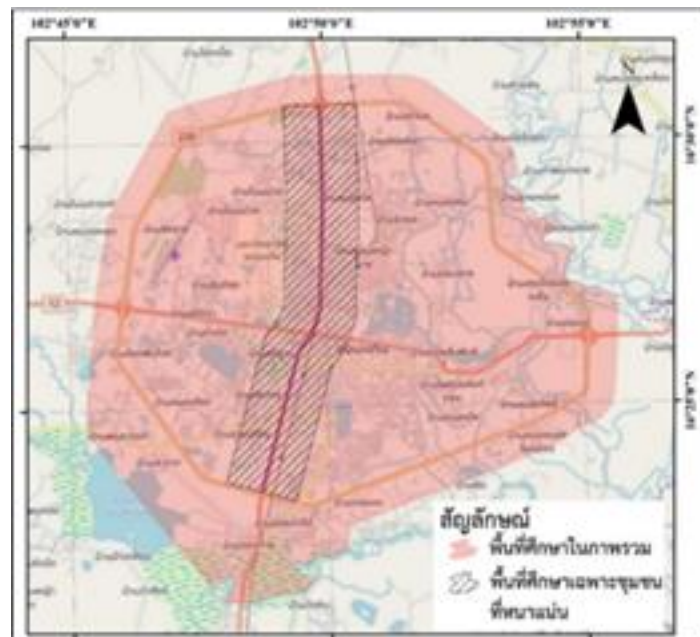
2. วัตถุประสงค์การวิจัย

เพื่อวิเคราะห์ความถูกต้องของการจำแนกข้อมูลภาพถ่ายดาวเทียมเชิงผสมผสานทั้งแบบพาสซีฟและแบบแอคทีฟในกรณีต่างๆ ด้วยการประยุกต์ใช้ตัวจำแนกแบบจักรแห่งการเรียนรู้ (Machine Learning) แบบ K-nearest neighbors, แบบ Artificial Neural Network, จักรกลเวกเตอร์ค้ำยัน (SVM), แบบ Random Forest และ แบบ Decision Tree โดยเลือกจำแนกประเภทการใช้ที่ดินของเมืองขอนแก่น ปี ค.ศ. 2020 เป็นกรณีศึกษา

3. ข้อมูลวิจัย

3.1 พื้นที่ทำการวิจัย

การศึกษานี้จะได้มีการใช้ข้อมูลภาพถ่ายดาวเทียม Sentinel-1 และ Sentinel-2 โดยมีการวิเคราะห์ความถูกต้องของการจำแนกจะเป็นการประยุกต์ใช้ข้อมูลจาก Google MapTM และข้อมูลการลงสนามจริง และศึกษาใน โครงการนี้ อยู่ภายใต้เขตผังเมืองรวม อำเภอเมือง จังหวัดขอนแก่น ดังรูปที่ 1 (พื้นที่ระบายสีชมพู)



รูปที่ 1 พื้นที่ศึกษาภายในเขตผังเมืองรวมจังหวัดขอนแก่น อำเภอเมือง จังหวัดขอนแก่น

3.2 ข้อมูลเกี่ยวกับดาวเทียมที่ใช้ในการศึกษา

ดาวเทียม Sentinel-1

ดาวเทียม Sentinel-1 เป็นภารกิจอวกาศที่ได้รับทุนจากสหภาพยุโรป และดำเนินการโดย องค์การอวกาศยุโรป(ESA) ภายใต้ชื่อโครงการ Copernicus ประกอบด้วยดาวเทียม 2 โดยดาวเทียม Sentinel-1A ได้ถูกส่งขึ้นสู่วงโคจร ณ วันที่ 3 เมษายน 2014, 21:02:31 UTC เป็นดาวเทียม เป็นเรดาร์ช่องเปิดสังเคราะห์(SAR) ในระบบ C band จะตรวจสอบสภาพแวดล้อมหลายด้านตั้งแต่การตรวจจับและติดตาม การรั่วไหลของน้ำมันและการทำแผนที่น้ำแข็งในทะเล เพื่อตรวจสอบความเคลื่อนไหวในพื้นที่ผิวดิน และการทำแผนที่เพื่อเปลี่ยนแปลงการใช้ที่ดิน สามารถโคจรครบรอบวงโคจรได้ในเวลา 12 วัน ในภายหลัง ดาวเทียม Sentinel-1B ถูกส่งขึ้นสู่วงโคจรภายหลัง ณ วันที่ 25 เมษายน 2016, 21:02 UTC เพื่อใช้งานร่วมกับ Sentinel-1A เพื่อส่งข้อมูลในการสำรวจทรัพยากรต่าง ๆ ได้ดีขึ้น โดยวิหังโคจรของทั้งสองนั้น โคจร 180 องศาต่อ Sentinel-1A เพื่อเพิ่มการครอบคลุมและการจัดส่งข้อมูล สำหรับการสำรวจการเปลี่ยนแปลงทรัพยากรของเรา ลักษณะการทำงานของดาวเทียม Sentinel-1 มีจุดเด่นที่สามารถทำแผนที่พื้นที่ส่วนกลางของโลกได้ ในโหมด Interferometric Wide swath ทุก ๆ 12 วัน ในการผ่านเพียงครั้งเดียว (ขึ้นหรือลง) กลุ่มดาวเทียมสองดวงมีวงจรการทำซ้ำที่แน่นอน 6 วันที่เส้นศูนย์สูตร เนื่องจากระยะเวลายาวของวงโคจรมีความแตกต่างกันไปตามละติจูด อัตราการทบทวนซ้ำจะสูงกว่าละติจูดที่มากกว่าเส้นศูนย์สูตร ทั้งนี้ดาวเทียม SENTINEL-1 ใช้เครื่องเรดาร์รับส่งสังเคราะห์ C-band แบบเดี่ยวที่ทำงานที่ความถี่กลาง 5.405 GHz ประกอบด้วยเสาอากาศแบบแอททิฟที่ใช้งานได้เต็มรูปแบบทำให้สามารถสแกนได้อย่างรวดเร็วในระดับความสูงและมุมสูงความจุในการจัดเก็บข้อมูล 1 410 Gb และ 520 Mbit / s X-band อุปกรณ์ C-SAR รองรับการทำงานในโพลาไรซ์แบบคู่ (HH + HV, VV + VH) ผ่านหัวขั้วส่งสัญญาณ (สลับไปยัง H หรือ V) และสองสายรับขนานสำหรับโพลาไรซ์ H และ V ข้อมูลโพลาไรซ์แบบคู่นี้มีประโยชน์สำหรับการจำแนกประเภทที่ดินและการใช้งานในทะเลน้ำแข็ง โหมดที่ปราศจากความขัดแย้งหลัก คือ IW เนื้อพื้นดินและ WV เนื้อมหาสมุทรเปิด SENTINEL-1 ในรูปแบบสี่โหมด

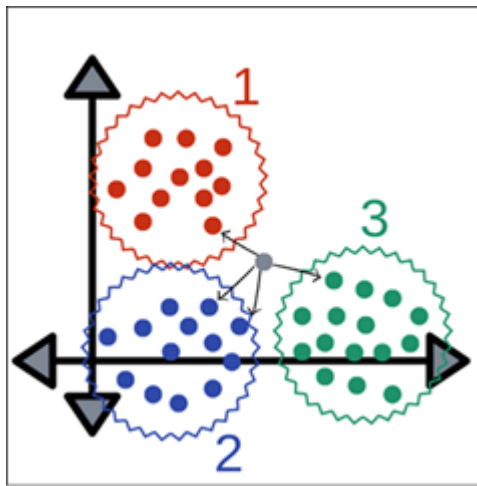
ดาวเทียม Sentinel-2

ดาวเทียม Sentinel-2 เป็นภารกิจสังเกตการณ์จากโลก ที่พัฒนาโดยองค์การอวกาศยุโรป(ESA) ซึ่งเป็นส่วนหนึ่งของโครงการ Copernicus เพื่อทำการสังเกตการณ์บนบกเพื่อสนับสนุนการบริการ เช่นการตรวจสอบป่าการตรวจหาการเปลี่ยนแปลงพื้นที่ปกคลุมและการจัดการภัยพิบัติทางธรรมชาติ ประกอบด้วยสองดาวเทียมที่เหมือนกัน Sentinel-2A และ Sentinel-2B ซึ่งแต่ละดวงจะมีเครื่องวัดหลายสเปกตรัมเดี่ยว (MSI) ที่มี 13 ช่องสเปกตรัมในรังสีอินฟราเรดที่มองเห็น / ไกล (VNIR) และช่วงคลื่นอินฟราเรดสั้น (SWIR) ในระยะสั้น โดย MSI ใช้แนวคิด ระบบการกวาดภาพที่สแกนภาพตามแนวทิศทางที่ดาวเทียมโคจร (Pushbroom) และการออกแบบของมันเป็นถูกขับเคลื่อนด้วยความกว้าง 290 กม. พร้อมด้วยสมรรถนะทางเรขาคณิตและสเปกตรัมสูงที่ต้องใช้ในการวัดซึ่งมีรูรับแสงขนาด 150 มม. (6 นิ้ว) และการออกแบบเลนส์(Anastigmat) สามกระจกที่มีความยาวโฟกัสประมาณ 600 มม. (24 นิ้ว). มุมมองในทันทีคือประมาณ 21 องศาโดย 3.5 องศากระจกเป็นรูปสี่เหลี่ยมผืนผ้าและทำจากซิลิคอนคาร์ไบด์ซึ่งเป็นเทคโนโลยีที่คล้ายคลึงกับผู้ที่อยู่ในภารกิจ Gaia ดาวเทียม Sentinel-2A ได้ถูกส่งขึ้นสู่วงโคจร ณ วันที่ 23 มิถุนายน 2015, 01:51:58 UTC เป็นดาวเทียมที่มีกล้องถ่ายภาพหลายมุมที่มีความละเอียดสูงและมีแถบสเปกตรัม 13 สาย จะทำการสังเกตการณ์ภาคพื้นดินเพื่อสนับสนุนการให้บริการเช่นการตรวจสอบป่าการตรวจสอบการเปลี่ยนแปลงพื้นที่ดินและการจัดการภัยพิบัติทางธรรมชาติ โดย

ต่อมาได้มีการส่ง Sentinel-2B ขึ้นสู่วงโคจร ณ วันที่ 7 มีนาคม 2560, 01:49:24 UTC จะทำการโคจร 180 องศา โดยลักษณะของดาวเทียม Sentinel-2 มีจุดเด่นที่มีกล้องถ่ายภาพหลายมุมที่มีความละเอียดสูงและมีแถบสเปกตรัม 13 สายเช่นกัน จะให้ข้อมูลสำหรับการเกษตรและป่าไม้และอื่น ๆ เพื่อให้สามารถทำนายผลผลิตได้

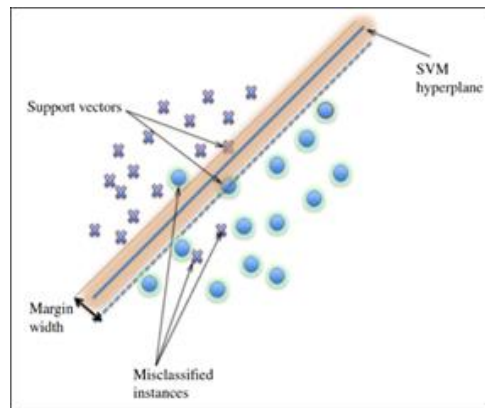
การจำแนกประเภทข้อมูลแบบควบคุม (supervised classification)

ตัวจำแนกแบบค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว (K-nearest neighbors: KNN) นับเป็นเทคนิคที่มีวิธีการไม่ซับซ้อนและเข้าใจได้ง่ายที่สุดที่ใช้ในการจำแนกประเภทข้อมูล โดยหลักการทำงาน คือ จะใช้หลักการเปรียบเทียบความคล้ายคลึงกันของข้อมูลที่สนใจกับข้อมูลอื่นว่ามีความคล้ายคลึงหรืออยู่ใกล้กับข้อมูลใดมากที่สุด k ตัว จากนั้นจะทำการตัดสินใจว่า คำตอบของข้อมูลที่สนใจนั้นควรเป็นคำตอบเดียวกับข้อมูลที่อยู่ใกล้ที่สุด k ตัวนั้น ทั้งนี้ k คือความถี่ของข้อมูลที่อยู่ใกล้กับข้อมูลที่สนใจดังรูปที่ 2



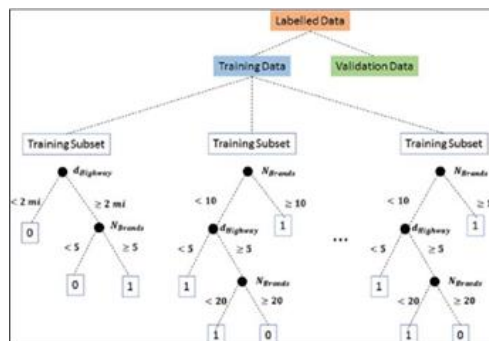
รูปที่ 2 หลักการจำแนกประเภทแบบค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว (K-nearest neighbors: KNN)

ตัวจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines Classifier: SVMC) เป็นการจำแนกข้อมูลแบบไม่ได้อาศัยพารามิเตอร์ทางสถิติ (non-parametric statistical) ดังนั้น การจำแนกด้วยวิธีนี้จึงไม่ได้อาศัยรากฐานของการกระจายตัวของข้อมูล (data distribution) มาเป็นหลักในการคิด โดยเหตุนี้ การประยุกต์ใช้เทคนิคการจำแนกด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนที่สามารถจำแนกข้อมูลได้แม้จำนวนพื้นที่ตัวอย่างมีอยู่ค่อนข้างจำกัด โดยการจำแนกด้วยวิธีนี้จะได้อาศัยข้อมูลที่มีอยู่อย่างค่อนข้างจำกัดนั้นในการสร้าง hyperplane เพื่อช่วยในการจำแนกข้อมูล โดยวิธี SVMC นี้ต้องอาศัยฟังก์ชันแกน (kernel function) ทั้งนี้ kernel function ที่ได้รับความนิยมสูงที่สุดซึ่งนำมาใช้ในการศึกษานี้คือ linear kernel ที่อาศัยการแบ่งกลุ่มด้วยเส้นตรงดังรูปที่ 3



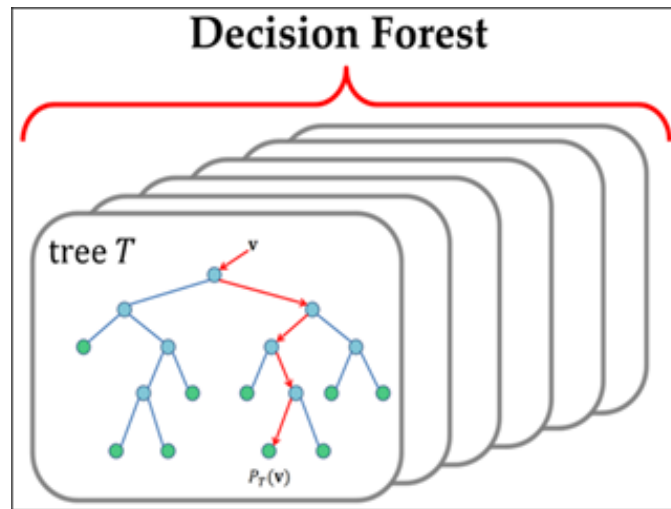
รูปที่ 3 หลักการจำแนกประเภทแบบซัพพอร์ตเวกเตอร์แมชชีน [6]

ตัวจำแนกแบบ Decision Tree เป็น model แบบ rule-based คือ สร้างกฎ if-else จากค่าของแต่ละ feature โดยไม่มีสมการมากำกับความสัมพันธ์ระหว่าง feature & target สิ่งที่สำคัญในการสร้าง Decision Tree คือ การเลือก split ค่า feature แต่ละครั้ง จะต้อง minimise ค่าของ cost function ให้น้อยที่สุด (regression-mse, classification-impurity, entropy) ซึ่งมีหลักการดังในรูปที่ 4



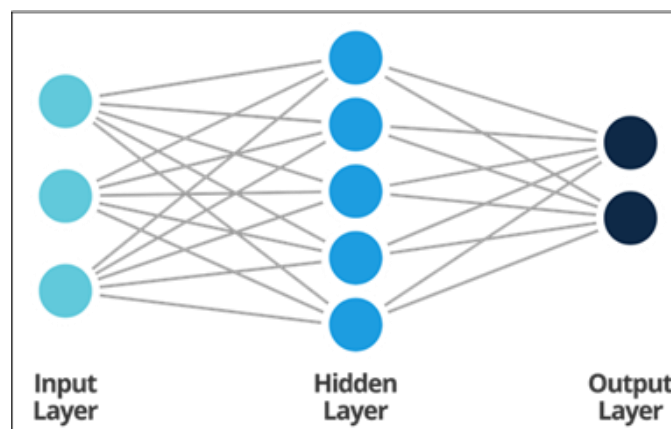
รูปที่ 4 หลักการจำแนกประเภทแบบ Decision Tree

ตัวจำแนกแบบ Random Forest เป็น model ที่นำ Decision Tree หลายๆ tree มา Train ร่วมกัน (ตั้งแต่ 10 ต้น ถึง มากกว่า 1000 ต้น) โดยที่แต่ละ tree จะได้รับ feature และ data เป็น subset ของ feature และ data ทั้งหมด แบบ random ตอนทำ prediction ก็ให้แต่ละ Decision Tree ทำ prediction ของใครของมัน และเลือกผล final prediction จากค่า prediction ที่ได้รับการโหวตมากที่สุด technique ดังกล่าวเรียกว่า bagging หรือ bootstrapping ดังรูปที่ 5



รูปที่ 5 หลักการจำแนกประเภทแบบ Random Forest

ตัวจำแนกแบบโครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) เป็นวิธีการที่มีพื้นฐานมาจากการเลียนแบบการทำงานของสมองมนุษย์ซึ่งประกอบด้วยนิวรอน (Neurons) ที่ใช้ในการเชื่อมโยงเพื่อแก้ปัญหาต่างๆ โดยมีโครงสร้างการทำงาน ตามรูปที่ 6 ซึ่งมีข้อมูลเข้า (Input) 1 ชั้น และข้อมูลออก (Output) 1 ค่าจะเรียกว่าเพอร์เซพตรอน (Perceptron)



รูปที่ 6 หลักการจำแนกประเภทแบบโครงข่ายประสาทเทียม

3.4 การประเมินความถูกต้องของการจำแนกข้อมูลภาพถ่ายดาวเทียม (accuracy assessment of satellite imagery classification)

การประเมินความถูกต้องของการจำแนกประเภทจำเป็นต้องใช้ข้อมูลที่ใช้สำหรับอ้างอิงในการตรวจสอบความถูกต้อง ข้อมูลที่ได้ในการศึกษาวิจัยครั้งนี้จะได้อาจจากการสำรวจภาคสนามหรือจากภาพถ่ายทางอากาศที่มีเวลาใกล้เคียงกันกับภาพถ่ายดาวเทียม โดยองค์ประกอบพร้อมด้วยรายละเอียดสามารถอธิบายได้ดังนี้

ข้อมูลอ้างอิงเปรียบเทียบ (reference data)

ในการประเมินความถูกต้องของการจำแนกข้อมูลภาพดาวเทียม จำเป็นอย่างยิ่งที่ต้องมีข้อมูลมาใช้อ้างอิงในการตรวจสอบความถูกต้อง ข้อมูลที่ใช้อ้างอิงอาจจะได้มาจาก 2 แหล่งที่มา คือ

ภาพถ่ายดาวเทียมความละเอียดสูงจากซอฟต์แวร์ Google Earth ซึ่งส่วนสำคัญคือถ่ายภาพของรายละเอียดบนพื้นผิวภูมิประเทศ ความเข้มของสีต่างๆ ในรูปถ่ายทางอากาศ จะบอกถึงความแตกต่างกันทั้งทางด้านภูมิศาสตร์กายภาพ พืชพรรณธรรมชาติ และก็จะเปลี่ยนไปตามฤดูกาล เราจะใช้ภาพถ่ายที่ถ่ายไว้ในเวลาใกล้เคียงกับภาพดาวเทียมที่ใช้จำแนกประเภท เพื่อตรวจสอบว่าการจำแนกข้อมูลภาพดาวเทียมนั้นมีความใกล้เคียงกับภาพถ่ายดาวเทียมความละเอียดสูงมากน้อยเพียงใด เพื่อเพิ่มความน่าเชื่อถือให้กับข้อมูลที่เราจำแนก

การเก็บข้อมูลภาคสนาม (ground truth data) การเก็บข้อมูลจะสุ่มเก็บในแต่ละจุดในพื้นที่ศึกษา เพื่อใช้อ้างอิงในการตรวจสอบการประเมินความถูกต้องของภาพดาวเทียม การสุ่มเลือกตัวอย่างควรทำอย่างน้อย 50 จุด ในแต่ละประเภทชุดข้อมูล [7] แต่อย่างไรก็ตาม บางประเภทข้อมูลพื้นที่น้อยมาก จึงทำให้จำนวนจุดตรวจสอบแม้จะกระจายทั่วทั้งพื้นที่ศึกษาก็ไม่สามารถมีจำนวนจุดตรวจสอบเกินกว่า 50 จุดได้ และจากการศึกษางานวิจัย [8] ได้ตรวจสอบแผนที่เมืองไนโรบี ประเทศเคนย่า โดยมีขนาดพื้นที่ศึกษาทั้งหมด 713 ตารางกิโลเมตร โดยสุ่มเก็บตัวอย่างทั้งหมด 482 จุด ซึ่งคิดเป็นจำนวนประมาณ 7 จุด ต่อ 10 ตารางกิโลเมตร ดังนั้น สำหรับบางประเภทการใช้ที่ดินที่มีขนาดพื้นที่ไม่มากนักที่ไม่สามารถเก็บตัวอย่างได้ถึง 50 จุด จึงจะใช้เกณฑ์การพิจารณาต่อขนาดของพื้นที่ในลักษณะเดียวกันกับของ [8] ดังนั้นการเก็บตัวอย่างแต่ละประเภทการใช้ที่ดินจึงมีไม่เท่ากันขึ้นอยู่กับปริมาณและความเหมาะสมของประเภทการใช้ที่ดินบริเวณพื้นที่ศึกษาด้วย

เมทริกซ์ความคลาดเคลื่อน (error matrix)

การสร้างตารางเมทริกซ์ความคลาดเคลื่อนดังรูปที่ 7 เพื่อหาความสัมพันธ์ระหว่างข้อมูลอ้างอิงที่มีกับผลที่ได้จากการจำแนกประเภทให้ตรงกัน โดยข้อมูลที่ได้จากการจำแนกประเภทจะจัดให้อยู่ในหลัก (columns) ของตารางและข้อมูลที่ใช้อ้างอิงจะจัดให้อยู่ในแถว (rows) ของเมทริกซ์ความคลาดเคลื่อน [9] ในการสร้างตารางเมทริกซ์ความคลาดเคลื่อนสามารถหาค่าความถูกต้องโดยรวม (overall accuracy) ความถูกต้องของผู้ใช้ (user's accuracy) ความถูกต้องของผู้ผลิตแผนที่ (producer's accuracy) และค่า kappa index ความถูกต้องของผู้ใช้ (user's accuracy) สามารถคำนวณได้จากตัวเลขที่ได้จากจุดภาพการจำแนกประเภทในแต่ละกลุ่มหารด้วยค่าทั้งหมดของจุดภาพของการจำแนกประเภทในแถวนั้น (ผลรวมในแถว) และความถูกต้องของผู้ผลิตแผนที่ (producer's accuracy) สามารถคำนวณได้จาก ตัวเลขที่ได้จากจุดภาพการจำแนกประเภทในแต่ละกลุ่มหารด้วยค่าทั้งหมดของจุดภาพของการจำแนกประเภทในคอลัมน์นั้น (ผลรวมในคอลัมน์)

		j = Columns (Reference)			Row Total
		1	2	r	x_{i+}
i = Rows (Classification)	1	x_{11}	x_{12}	x_{1k}	x_{1+}
	2	x_{21}	x_{22}	x_{2k}	x_{2+}
	r	x_{k1}	x_{k2}	x_{kk}	x_{k+}
Column Total x_{+j}		x_{+1}	x_{+2}	x_{+3}	x_{+4}

รูปที่ 7 ตัวอย่างการสร้างตารางเมทริกซ์ความคลาดเคลื่อน

ค่า overall accuracy และ kappa index

การวิเคราะห์ค่า overall accuracy และ kappa จะถูกใช้ในการประเมินความถูกต้องของการจำแนกประเภท ซึ่งจะทำการวิเคราะห์ในตารางเมทริกซ์ความคลาดเคลื่อน โดยใช้เทคนิคทางสถิติช่วยในการคำนวณ overall accuracy คำนวณได้โดยนำค่าผลรวมในแนวทแยงหลักหารด้วยจำนวนจุดภาพทั้งหมดที่ใช้ในการอ้างอิงในตารางเมทริกซ์ความคลาดเคลื่อน () การวิเคราะห์ค่า kappa index จะใช้ในการประเมินความถูกต้อง [7] และค่า kappa index ในเชิงสถิติจะเขียนอยู่ในรูปซึ่งถือว่าเป็นค่าความถูกต้อง [8-9] และค่า kappa index นี้สามารถคำนวณได้จาก สมการที่ 1

$$\hat{K} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})} \quad (1)$$

กำหนดให้

r	คือ	จำนวนแถวของตาราง error matrix
x _{ii}	คือ	ค่าในแถวที่ i และคอลัมน์ที่ i
x _{i+}	คือ	ผลรวมของค่าอ้างอิงในแต่ละแถว (แสดงค่าทางขวามือของตาราง error matrix)
x _{+i}	คือ	ผลรวมของค่าอ้างอิงในแต่ละคอลัมน์ (แสดงค่าทางส่วนล่างของตาราง error matrix)
N	คือ	ผลรวมจากทุกค่าอ้างอิง

3.5 ภาษา R

R เป็นแพลตฟอร์มซอฟต์แวร์รวมแหล่งข้อมูลฟรีสำหรับการเขียนโปรแกรม การจัดการข้อมูล การคำนวณ และการแสดงผลกราฟฟิก (R Development Core Team 2005) จากข้อมูลของ Main Donald and Braun (2003) R เป็นภาษาที่พัฒนามาจากภาษา S ซึ่ง

ได้รับการพัฒนาโดย Ross Ihaka และ Robert Gentleman จาก University of Auckland ประเทศนิวซีแลนด์ ภาษ R เปิดตัวภายใต้ใบอนุญาต GNU GPL ในปี 1995 และมีการพัฒนาดังแต่นั้นมา ขณะนี้ R ได้รับการดูแลและพร้อมใช้งานบนแพลตฟอร์มคอมพิวเตอร์หลายแพลตฟอร์ม เช่น Windows, GNU/Linux และ Mac OS X

4. วิธีการศึกษา

4.1 พื้นที่ทำการศึกษา

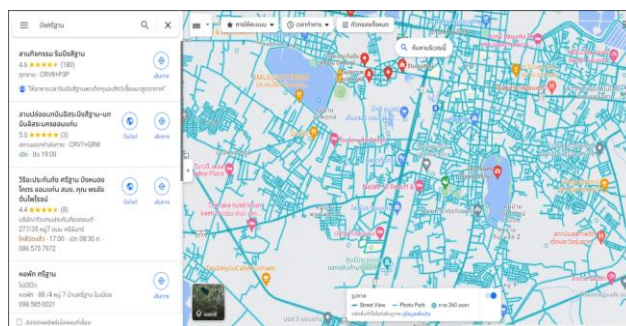
ในงานวิจัยนี้ได้ทำการศึกษาพื้นที่เมืองขอนแก่น และได้มีการรวบรวมและการนำเข้าข้อมูลภาพถ่ายดาวเทียม ภาพถ่ายดาวเทียมที่นำมาใช้ในงานวิจัย ได้แก่ ภาพถ่ายดาวเทียม Sentinel-1 จำนวน 3 ภาพ ภาพถ่ายดาวเทียม Sentinel-2 จำนวน 1 ภาพ ดังแสดงในตารางที่ 1

ตารางที่ 1 รายละเอียดข้อมูลภาพถ่ายดาวเทียม Sentinel-1 และ Sentinel-2

ลำดับที่	ดาวเทียม	ระบบการถ่ายภาพ	วันที่บันทึกข้อมูลภาพ
1	Sentinel-1	C-band	7 เมษายน 2020
2	Sentinel-1	C-band	25 สิงหาคม 2020
3	Sentinel-1	C-band	27 ธันวาคม 2020
4	Sentinel-2	MSI	27 ธันวาคม 2020

4.2 ซอฟต์แวร์ที่ใช้ในงานวิจัย

ซอฟต์แวร์ที่ใช้เตรียมการสร้างข้อมูลชุดสอนการเรียนรู้ (Training Data) และสร้างข้อมูลชุดทดสอบการเรียนรู้ (Test Data) มี 2 ซอฟต์แวร์ คือ Google Earth และ ArcGIS ขณะที่ซอฟต์แวร์ Google Earth: เป็นซอฟต์แวร์ที่พัฒนาโดยบริษัทกูเกิล ที่ให้บริการดูภาพถ่ายทางอากาศพร้อมทั้งแผนที่ เส้นทาง และผังเมืองซ้อนทับลงในแผนที่ รวมทั้งระบบ จีไอเอส ในรูปแบบ 3 มิติ สามารถทำงานผ่านเครื่องคอมพิวเตอร์ส่วนบุคคล ทั้ง PC และ Notebook รวมทั้ง Smart Phone และ Tablet และซอฟต์แวร์ ArcGIS: เป็นซอฟต์แวร์ด้าน GIS สำหรับการสร้าง แก้ไข วิเคราะห์ จัดเก็บ และแบ่งปันข้อมูลร่วมกัน ช่วยใช้ในการตัดสินใจ แสดงผลได้ทั้งแบบ 2 มิติ และ 3 มิติ นอกจากนี้ซอฟต์แวร์ที่ใช้ในการเตรียมข้อมูลภาพถ่ายดาวเทียมได้ใช้ซอฟต์แวร์ SNAP: เป็นซอฟต์แวร์ที่ผลิตโดยองค์การอวกาศยุโรป (European Space Agency: ESA) ใช้สำหรับวิเคราะห์และประมวลผลข้อมูลภาพถ่ายดาวเทียม และสำคัญที่สุดในการศึกษานี้คือ การซอฟต์แวร์ที่ใช้ในการประมวลผลภาพถ่ายดาวเทียมที่เป็นซอฟต์แวร์ RStudio: เป็นซอฟต์แวร์ที่ทำงานร่วมกับโปรแกรม R แต่เอื้อและเหมาะสมกับการใช้งานและการพัฒนาให้สะดวกมากขึ้น

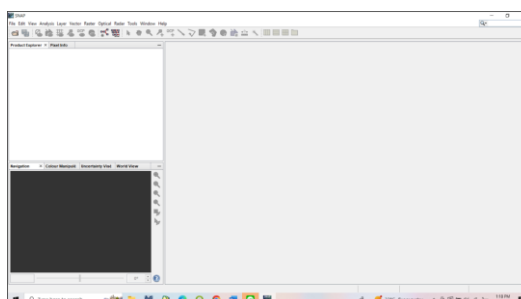


(ก)

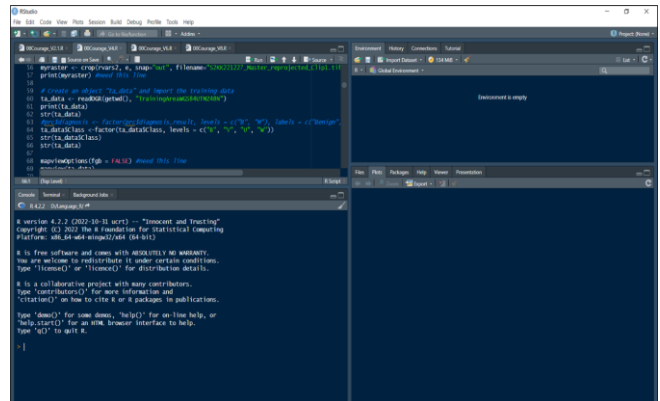


(ข)

รูปที่ 8 (ก) ซอฟต์แวร์ Google Earth, (ข) ซอฟต์แวร์ ArcGIS



รูปที่ 9 ซอฟต์แวร์ SNAP



รูปที่ 10 ซอฟต์แวร์ RStudio

4.3 ขั้นตอนการดำเนินงานวิจัย

การสร้างข้อมูลชุดสอนการเรียนรู้ (Training Data) และสร้างข้อมูลชุดทดสอบการเรียนรู้ (Test Data)

กำหนดพื้นที่ตัวอย่างเพื่อใช้เป็นตัวแทนของข้อมูลแต่ละประเภท เพื่อใช้สำหรับการวิเคราะห์ค่าทางสถิติกับกลุ่มพื้นที่ตัวอย่าง เพื่อหาค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน และค่าสหสัมพันธ์เพื่อใช้ค่าสถิตินั้นในการกำหนดประเภทข้อมูล ในการศึกษาได้มีการใช้ชุดข้อมูล Training Data ได้จากการรวบรวมข้อมูลด้วยการลงพื้นที่ศึกษาจริง และ Google Earth

การจำแนกประเภทการใช้ที่ดิน (land use classification)

การศึกษานี้ ใช้การเปรียบเทียบความถูกต้องของแผนที่การใช้ที่ดินที่ได้จากการใช้ผสมผสานภาพถ่ายดาวเทียม Sentinel-1 และ Sentinel-2 มาผสมผสานในแบบต่างๆ และจำแนกด้วยตัวจำแนกแบบจักรกลเรียนรู้ ได้แก่ แบบ K-nearest neighbors, แบบ Artificial Neural Network, จักรกลเวกเตอร์ค้ำยัน (SVM), แบบ Random Forest และ แบบ Decision Tree

การประเมินความถูกต้อง (accuracy assessment)

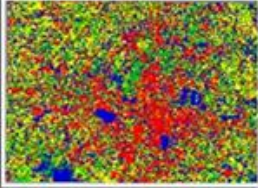
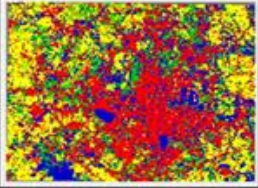
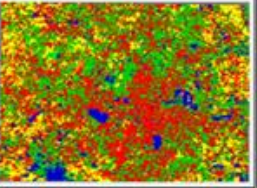
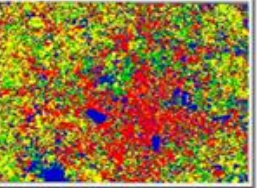
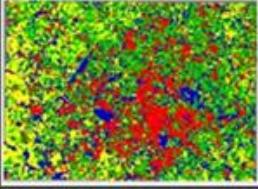
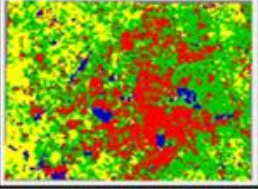
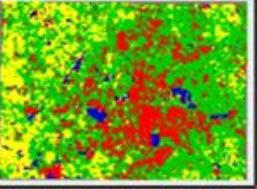
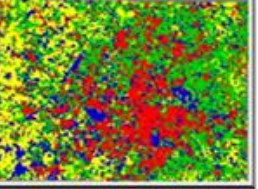
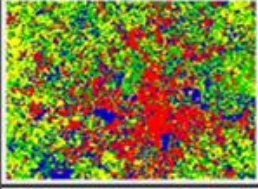
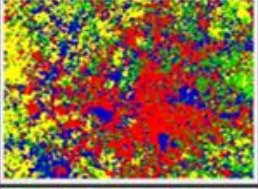
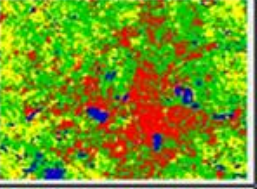
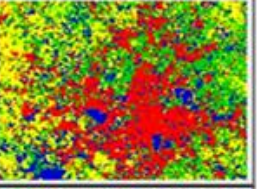
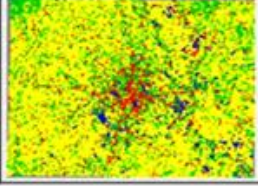
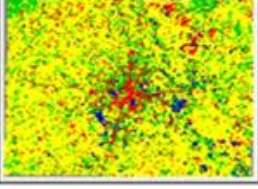
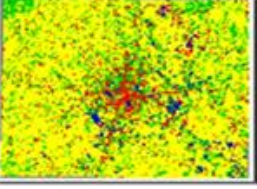
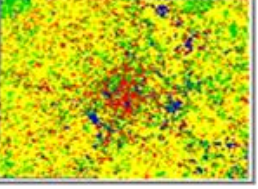
การประเมินความถูกต้องของการจำแนกประเภทการใช้ที่ดินในแต่ละปี (ค.ศ. 2008 ค.ศ. 2015 และ ค.ศ. 2018) ได้ใช้ข้อมูลอ้างอิงผ่านการแปลภาพถ่ายดาวเทียมความละเอียดสูงของแต่ละปีใน Google Earth™ ผ่านฟังก์ชันย้อนเวลาโดยอาศัยวิธีการแบบสุ่มเลือก (random) จำนวนอย่างน้อย 75 จุดสำหรับแต่ละประเภทการใช้ที่ดิน แล้ววิเคราะห์ผ่านเมตริกซ์ความคลาดเคลื่อน (error matrix) เพื่อประเมินค่าความถูกต้องโดยรวม (overall accuracy) ค่าความถูกต้องของผู้ใช้ (user's accuracy) ค่าความถูกต้องของผู้ผลิต (producer's accuracy) และค่าดัชนีแคปปา (kappa index) โดยค่าต่างๆ เหล่านี้ถือเป็นตัวชี้วัดที่สำคัญในการประเมินคุณภาพของการจำแนกประเภทการใช้ที่ดิน

5. ผลการทดลอง

ผลการวิเคราะห์การจำแนกการใช้ที่ดิน

ผลการวิเคราะห์การจำแนกการใช้ที่ดินแต่ละประเภท (Bare Land: B, Vegetation Area: V, Urban Area: U และ Water: W) ตามแต่ละกรณี ซึ่งมีดังนี้ กรณี 1: ใช้ภาพถ่ายดาวเทียม Sentinel-2, กรณี 2: ใช้ภาพถ่ายดาวเทียม Sentinel-1, กรณี 3: ใช้ภาพถ่าย

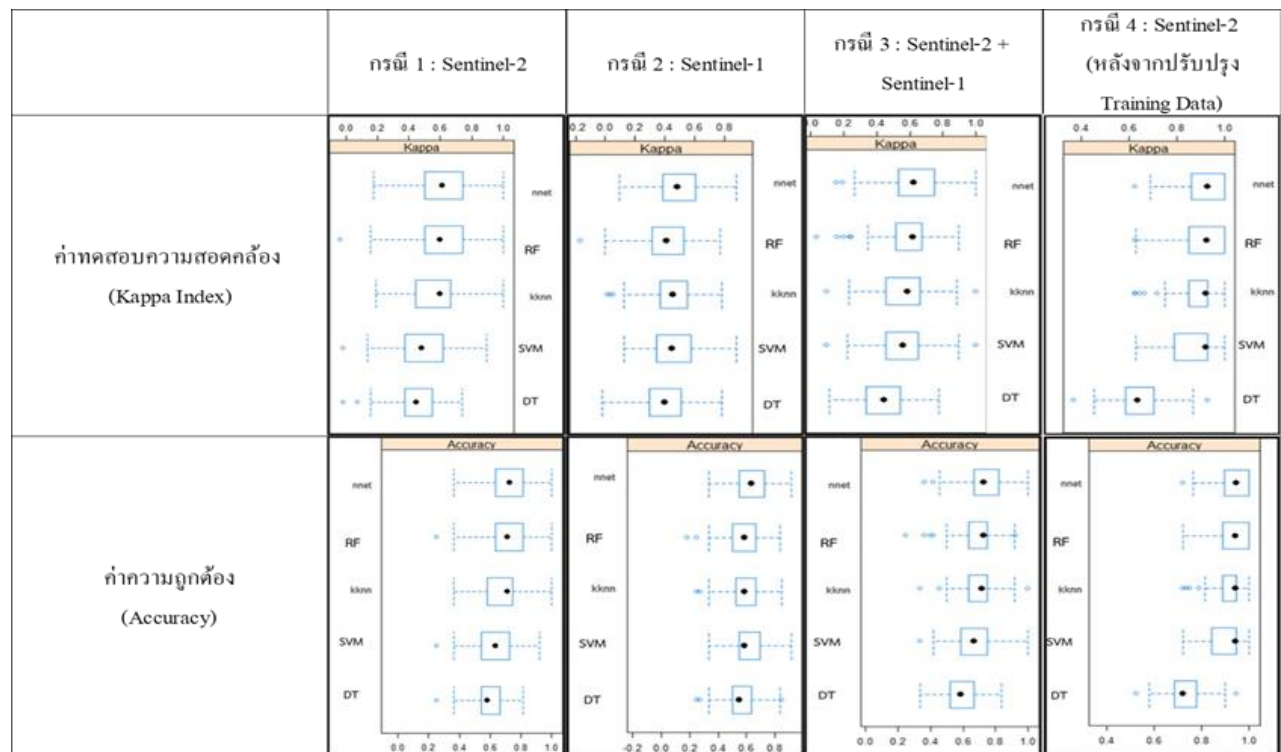
ดาวเทียม Sentinel-2 รวมกับ Sentinel-1 และกรณี4: ใช้ภาพถ่ายดาวเทียม Sentinel-2 (หลังจากปรับปรุง Training Data) ดังแสดงในรูปที่ 11

	K-Nearest Neighbors (KNN)	Artificial Neural Network (ANN)	จักรกลเวกเตอร์ค้ำยัน (SVM)	Random Forest (RF)
กรณี 1 : Sentinel-2				
กรณี 2 : Sentinel-1				
กรณี 3 : Sentinel-2 + Sentinel-1				
กรณี 4 : Sentinel-2 (หลังจากปรับปรุง Training Data)				

รูปที่ 11 การเปรียบเทียบผลลัพธ์จากการจำแนกประเภทในแต่ละกรณี

ผลการวิเคราะห์ค่าทดสอบความสอดคล้อง (Kappa Index) และค่าความถูกต้อง (Accuracy)

ผลการวิเคราะห์ค่าทดสอบความสอดคล้อง (Kappa Index) และค่าความถูกต้อง (Accuracy) การใช้ที่ดินแต่ละประเภท (Bare Land: B, Vegetation Area: V, Urban Area: U และ Water: W) ตามแต่ละกรณี ซึ่งมีดังนี้ กรณี1: ใช้ภาพถ่ายดาวเทียม Sentinel-2, กรณี2: ใช้ภาพถ่ายดาวเทียม Sentinel-1, กรณี3: ใช้ภาพถ่ายดาวเทียม Sentinel-2 รวมกับ Sentinel-1 และกรณี4: ใช้ภาพถ่ายดาวเทียม Sentinel-2 (หลังจากปรับปรุง Training Data) ดังแสดงในรูปที่ 12



รูปที่ 12 การเปรียบเทียบผลค่าทดสอบความสอดคล้อง (Kappa Index) และค่าความถูกต้อง (Accuracy) ตามแต่ละกรณี

6. สรุปผลและอภิปรายผล

สรุปผล

จากรูปที่ 8 ตารางแสดงผลการเปรียบเทียบการจำแนกประเภทการใช้ที่ดินในแต่ละกรณีด้วยภาษา R แสดงให้เห็นว่าเมื่อมีการปรับปรุงสร้างข้อมูลชุดสอนการเรียนรู้ (Training Data) และสร้างข้อมูลชุดทดสอบการเรียนรู้ (Test Data) สามารถจำแนกการใช้ที่ดินได้ออกมาชัดเจนมากกว่ากรณีอื่นๆ ที่ชุดข้อมูลไม่ได้มีจำนวนที่มาก จากรูปที่ 9 ตารางแสดงผลการเปรียบเทียบผลค่าทดสอบความสอดคล้อง (Kappa Index) และค่าความถูกต้อง (Accuracy) แสดงให้เห็นว่า ในการจำแนกการใช้ที่ดิน 4 ประเภท ได้แก่ พื้นที่ว่างเปล่า พื้นที่ปกคลุมด้วยพืชพรรณ พื้นที่แหล่งชุมชน และแหล่งน้ำ ใช้ตัวจำแนกแบบ K-nearest neighbors, แบบ Artificial Neural Network, จักรกลเวกเตอร์ค้ำยัน (SVM) และแบบ Random Forest โดยผลการประเมินความถูกต้องของผลการจำแนกที่ใช้ภาพถ่ายดาวเทียม Sentinel-2 พบว่ามีค่าความถูกต้องโดยรวมเท่ากับ 71% และค่าทดสอบความสอดคล้องโดยรวมเท่ากับ 0.6 ส่วนผลการประเมินความถูกต้องของผลการจำแนกที่ใช้ภาพถ่ายดาวเทียม Sentinel-1 พบว่ามีค่าความถูกต้องโดยรวมเท่ากับ 55% และค่าทดสอบความสอดคล้องโดยรวมเท่ากับ 0.45 ส่วนผลการประเมินความถูกต้องของผลการจำแนกที่ใช้ภาพถ่ายดาวเทียม Sentinel-2 รวมกับ Sentinel-1 พบว่ามีค่าความถูกต้องโดยรวมเท่ากับ 65% และค่าทดสอบความสอดคล้องโดยรวมเท่ากับ 0.55 ในขณะที่ผลการประเมินความถูกต้องของผลการจำแนกที่ใช้ภาพถ่ายดาวเทียม Sentinel-2 หลังจากการปรับชุดข้อมูล Training Data พบว่ามีค่าความถูกต้องโดยรวมเท่ากับ 92% และค่าทดสอบความสอดคล้องโดยรวมเท่ากับ 0.9 ซึ่งค่าความถูกต้องนั้นจะบ่งบอกการจำแนกประเภทการใช้ที่ดิน จากงานวิจัยนี้สามารถนำไปสู่การวิเคราะห์การใช้ประโยชน์ที่ดิน หรือการขยายตัวของเมืองในอนาคตได้ว่า

เมืองจะมีการขยายไปในทิศทางไหน ยิ่งแผนที่การใช้ที่ดินที่มีการจำแนกออกมาได้อย่างแม่นยำมากที่สุด การวางแผนในอนาคตจะยิ่งทำได้ง่ายมากยิ่งขึ้น

7. อภิปรายผล

การใช้ชุดคำสั่งในภาษา R เพื่อประโยชน์ในการจำแนกการใช้ที่ดินเป็นการวิเคราะห์ที่ให้ผลลัพธ์ที่ค่อนข้างมากในการจำแนกประเภทการใช้ที่ดิน และมีความคล่องตัวสูงในการปรับเปลี่ยน ตัวแปร และมีจุดเด่นในมิติของการนำมาใช้เพื่อสนับสนุนการจำแนกข้อมูลด้วยวิธีรหัสเปิด (Opensource) ซึ่งทำให้ลดข้อจำกัดในเรื่องของการจัดหาทรัพยากรโปรแกรมที่ในปัจจุบันมีราคาก่อนข้างสูง ยิ่งไปกว่านั้นการวิเคราะห์ภาษา R ยังทำให้เกิดการสื่อสารได้โดยตรงกับนักวิจัยชั้นนำของโลกได้อีกด้วย โดยมีกลุ่มของผู้เชี่ยวชาญที่ ณ ปัจจุบันนี้กำลังให้ความสนใจที่จะนำภาษา R ไปวิเคราะห์ข้อมูลภาพถ่ายดาวเทียมต่างๆ ให้เกิดเป็นรูปแบบใหม่เป็นจำนวนมาก และผลลัพธ์ที่ได้จากการศึกษาวิจัยนี้เอง ได้สะท้อนให้เห็นถึงขีดความสามารถของภาษา R ในการเป็นทางเลือกให้กับวิศวกร ในการนำมาใช้เพื่อการสำรวจระยะไกลได้เป็นอย่างดี ในปัจจุบันที่ประเทศไทยกำลังก้าวเข้าสู่ยุคสังคมแห่งนวัตกรรม เพื่อให้ประเทศไทยสามารถ ที่จะยืนบนขาของตนเองได้ อย่างน้อยการพัฒนาในรูปแบบของการใช้ Machine Learning ในการจำแนก โดยที่ลดการนำเข้าซอฟต์แวร์อันเป็นการสูญเสียรายได้ให้กับประเทศชาติจึงเป็นสิ่งจำเป็น การศึกษาวิจัยนี้ได้สะท้อนให้เห็นถึงขีดความสามารถของการนำเอาภาษา R ที่มีลักษณะของการเข้าชุดคำสั่งโดยการเลือกชุดความรู้ที่อยู่ในรูปแบบของ Library ให้เหมาะสม พร้อมกันนี้ในการศึกษาครั้งนี้ได้มีการนำเอาชุดคำสั่งที่เกี่ยวข้องกับการพัฒนาในการจำแนกมารวบรวมเอาไว้ โดยผลลัพธ์ได้สะท้อนให้เห็นว่า ภาษา R สามารถนำมาใช้ในการจำแนกข้อมูลภาพถ่ายดาวเทียมแบบควบคุมได้เป็นอย่างดี ข้อค้นพบที่ได้จากการศึกษานี้คือ ชุดคำสั่งที่สามารถนำไปประยุกต์ใช้กับการศึกษาวิจัยอื่นๆ ได้ในอนาคต พร้อมด้วยกระบวนการวิเคราะห์ข้อมูลที่น่าสนใจ

ผลประโยชน์ทับซ้อน

ผู้เขียนขอประกาศว่าบทความนี้ไม่มีผลประโยชน์ทับซ้อน

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณทุนอุดหนุนงานวิทยานิพนธ์จากคณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น และขอกราบขอบพระคุณผู้บริหาร คณาจารย์ เจ้าหน้าที่ ของคณะวิศวกรรมศาสตร์ มหาวิทยาลัยขอนแก่น ที่ได้ให้ความอนุเคราะห์เครื่องมือ สถานที่ และอุปกรณ์ที่ใช้ในการศึกษาวิจัยครั้งนี้

เอกสารอ้างอิง

- [1] M. G. C. Gómez, Joint use of Sentinel-1 and Sentinel-2 for land cover classification: A machine learning approach, 2017, 18, pp. 72.
- [2] E. Kositsakulchai *et al.*, Machine-learning classification of debris-covered glaciers using a combination of Sentinel-1/-2 (SAR/optical), Landsat 8 (thermal) and digital elevation data, *J. Sci. Technol. Kasetsart Univ.*, 2018, 7 (3) pp. 107365. doi: 10.1016/j.geomorph.2020.107365.

- [3] E. Kositsakulchai, S. Yodjaroen, and Y. Phankamolail, Assessment of the impact of land use change on runoff in Lam Phachi Basin using satellite data and SWAT model [in Thai, with English Abstract], *J. Sci. Technol. Kasetsart Univ.*, 2018, 7 (3), pp. 11–26.
- [4] H. Alifu, J. F. Vuillaume, B. A. Johnson, and Y. Hirabayashi, *Machine-learning classification of debris-covered glaciers using a combination of Sentinel-1/-2 (SAR/optical), Landsat 8 (thermal) and digital elevation data*, 2020, 369, Elsevier B.V.
- [5] แทนรัฐ สุจาวี, ชาติชาย ไวยสุระสิงห์, พนกฤษณ คลังบุญครอง, อนันต์ ศรีเดช, ชูติมา ไวยสุระสิงห์, and ศิโรรัตน์ พัฒนไพโรจน์, การประยุกต์ใช้ตัวจำแนกแบบจักรแห่งการเรียนรู้ในการวิเคราะห์ข้อมูลภาพถ่ายดาวเทียมเชิงผสมผสานระหว่างแบบพาสซีฟและแบบแอคทีฟเพื่อประเมิน การเปลี่ยนแปลงการใช้ที่ดิน: กรณีศึกษาเมืองขอนแก่น, *วารสารวิจัย มข*, 2564, 21 (1) pp. 14–25.
- [6] M. N. Murty and R. Raghava, “Linear support vector machines, *SpringerBriefs Comput. Sci.*, 2016, 9783319410623, pp. 41–56. doi: 10.1007/978-3-319-41063-0_4.
- [7] Hay AM. Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*. 1979, 45(4), pp. 529-33.
- [8] Mundia CN, Aniya M. Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. *International Journal of Remote Sensing*. 2005, 26 (13), pp.2831-49.
- [9] Congatton RG, Green K. Assessing the accuracy of remotely sensed data: principles and practices. 2nd ed. Boca Raton: CRC Press/Taylor & Francis; 2009, p 183.