การคาดการณ์การบำรุงรักษาสะพานโดยใช้วิธีป่าแบบสุ่ม
Forecasting Bridge Maintenance Using a Random Forest Model

อนวัช ทิพย์ประเสริฐ[1]
รองศาสตราจารย์ ดร.สุพรชัย อุทัยนฤมล[2]
ดร.สุกิจ ยินดีสุข[3]
[1]นักศึกษาระดับบัณฑิตศึกษา ภาควิชาวิศวกรรมโยธา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[2]ภาควิชาวิศวกรรมโยธา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
[3]สำนักสำรวจและออกแบบ กรมทางหลวง

บทคัดย่อ

สะพานถือเป็นโครงสร้างพื้นฐานทางด้านการคมนาคมที่มีความสำคัญต่อการเติบโตทางเศรษฐกิจของประเทศไทย ความสามารถในการคาดการณ์ความเสียหายที่เกิดขึ้นกับโครงสร้างสะพานสามารถช่วยวางแผนการซ่อมแซม บำรุงรักษาสะพานในอนาคตได้ แนวทางในการตรวจสอบสภาพความเสียหายของสะพานโดยทั่วไปจะใช้วิธีตรวจสอบโดยการสังเกตด้วยตา (Visual Inspection) มีเครื่องมือเฉพาะทางในการวัดค่าที่ต้องการ ซึ่งวิธีการนี้จะมีข้อเสียคือใช้กำลังคน เงินทุน และเวลามาก ดังนั้น งานวิจัยนี้จะมุ่งเน้นศึกษาการคาดการณ์ความเสียหายของสะพานโดยใช้ วิธีป่าแบบสุ่ม (Random Forest) โดยการสร้างแบบจำลองจากฐานข้อมูลสะพาน กรมทางหลวง การสร้างแบบจำลองได้แบ่งข้อมูลออกเป็น 2 ส่วน คือ ข้อมูลสำหรับฝึกสอนแบบจำลอง และข้อมูลสำหรับทดสอบแบบจำลอง โดยการใช้วิธี K-Folds Cross Validation ค่าพารามิเตอร์หลักในการเป็นปัจจัยนำเข้าแบบจำลอง ได้แก่ อายุของสะพาน จราจรรายวันเฉลี่ยรายปี ร้อยละรถบรรทุกหนัก ความยาวของสะพาน จำนวนช่วงของสะพาน ระดับทางหลวง และความกว้างผิวจราจร ผลลัพธ์ค่าความแม่นยำที่ได้จากแบบจำลองป่าแบบสุ่มอยู่ที่ร้อยละ 81.00

คำสำคัญ: บำรุงรักษาสะพาน, วิธีป่าแบบสุ่ม, การคาดการณ์

Abstract

A bridge is one of the important transportation infrastructures for promoting the economic growth of Thailand. The ability to predict the damage condition of a bridge structure could aid in future bridge repair and maintenance planning. Guidelines for examining the condition of bridge damage by visual inspection are commonly used for bridge inspection. The disadvantage of this method is that it requires a lot of manpower, money, and time. Therefore, this research will focus on studying bridge damage condition prediction using the random forest method. By creating a model from the bridge inventory database from the Department of Highways. The main parameters for model input were bridge age, annual average daily traffic (AADT), percentage of heavy trucks, length of the

bridge, number of bridge spans, highway hierarchy, and traffic surface width. The accuracy of the prediction result obtained from the random forest model was 81.00%.

Keywords: Bridge Maintenance, Random Forest, Forecasting

บทนำ

สะพานถือเป็นโครงสร้างพื้นฐานทางด้านการคมนาคมที่มีความสำคัญต่อการเติบโตทางเศรษฐกิจของประเทศไทย ไม่กี่ปีที่ผ่านมา จำนวนสะพานที่ทรุดโทรมได้เพิ่มขึ้นอย่างมาก ทำให้เกิดความกังวลเกี่ยวกับปัญหาที่อาจจะเกิดขึ้นกับสะพานเหล่านั้น ไม่ว่าจะเป็นความปลอดภัยหรือความสามารถในการรองรับการเชื่อมโยงโครงข่ายทางด้านการคมนาคม ความเสียหายที่เกิดกับสะพานสามารถหลีกเลี่ยงได้ถ้ามีการบำรุงรักษาเชิงป้องกันอย่างเป็นระบบ การปฏิบัติงานซ่อมบำรุงสะพานในแต่ละครั้ง จะสามารถทำให้การใช้งานสะพานเป็นไปอย่างมีประสิทธิภาพ การตรวจสอบตามช่วงระยะเวลา ของส่วนประกอบต่าง ๆ ของโครงสร้างควรจะได้รับการปฏิบัติอย่างระมัดระวังและเป็นระบบในการที่จะระบุตำแหน่งที่ต้องได้รับการดูแลก่อนที่จะกลายเป็นปัญหาใหม่ในการซ่อมแซม การบำรุงรักษาโครงสร้างสะพาน คือ การทำให้สะพานอยู่ในสภาพที่พร้อมใช้งานและปลอดภัย โดยทั่วไปการตรวจสอบประสิทธิภาพของสะพานจะใช้วิธี การตรวจสอบโดยใช้สายตาเพื่อประเมินลักษณะสภาพการใช้งานเบื้องต้น การบำรุงรักษาสะพาน เป้าหมายหลักคือการแก้ไขปัญหาที่เกิดขึ้นอยู่ในปัจจุบัน ป้องกันความเสียหายที่อาจเกิดขึ้นในอนาคต หรือทำโครงสร้างเดิมที่มีอยู่ในดีขึ้นเพื่อลดค่าใช้จ่ายในการดูแลรักษาในระยะยาว

ในปัจจุบันกรมทางหลวงมีการรวบรวมข้อมูลบัญชีสะพานเพื่อให้ง่ายต่อการบำรุงรักษาสะพาน โดยการสำรวจ ตรวจสอบพฤติกรรมของสะพาน ประเมินความเสียหายที่เกิดขึ้น ผ่านทางระบบบริหารงานบำรุงสะพาน (Bridge Maintenance and Management System: BMMS) ซึ่งการรวบรวมข้อมูลจะใช้วิธีให้แขวงทางหลวง สำรวจ ตรวจสอบสภาพของสะพานที่อยู่ในความรับผิดชอบและส่งข้อมูลกลับมาที่ส่วนกลาง จากนั้นนำข้อมูลที่รวบรวมได้มาประเมิน หาแนวทางการแก้ไขที่เหมาะสมต่อไป จะเห็นได้ว่า แนวทางการป้องกันหรือแก้ไขปัญหาที่เกิดขึ้นกับสะพานมีหลายขั้นตอน อีกทั้งการดำเนินการในแต่ละขั้นตอนมีระยะเวลาในการดำเนินการค่อนข้างมาก ซึ่งในบางครั้งสะพานที่อยู่ในสะถานะเสียหายวิกฤต อาจไม่มีงบซ่อมแซมที่ทันการณ์ ทำให้ส่งผลเสียต่อผู้ใช้เส้นทางนั้น ๆ ในการคมนาคม ดังนั้นการลดระยะเวลาในการประเมินสภาพปัจจุบันหรือความเสียหายที่เกิดขึ้นกับสะพานจึงเป็นสิ่งจำเป็น

งานวิจัยนี้จึงมุ่งเน้นศึกษาเพื่อพัฒนาแบบจำลองในการคาดการณ์ความเสียหายที่เกิดขึ้นกับสะพานขนาดเล็กในความรับผิดชอบของกรมทางหลวง เพื่อให้สามารถวางแผนในการจัดการบำรุง รักษา และป้องกันปัญหาที่อาจเกิดขึ้นในอนาคตได้

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาปัจจัยสำคัญที่มีผลต่อความเสียหายที่เกิดขึ้นกับโครงสร้างสะพานโดยใช้วิธีป่าแบบสุ่ม
2. เพื่อประยุกต์ใช้วิธีป่าแบบสุ่มในการคาดการณ์เสียหายที่เกิดขึ้นกับโครงสร้างสะพาน

ขอบเขตการวิจัย

1. ศึกษาปัจจัยที่มีผลต่อความเสียหายที่เกิดขึ้นกับสะพานขนาดเล็ก คือ สะพานประเภท Plank Girder และ Slab type ซึ่งในความรับผิดชอบของกรมทางหลวง ในช่วงปี พ.ศ. 2557-2564

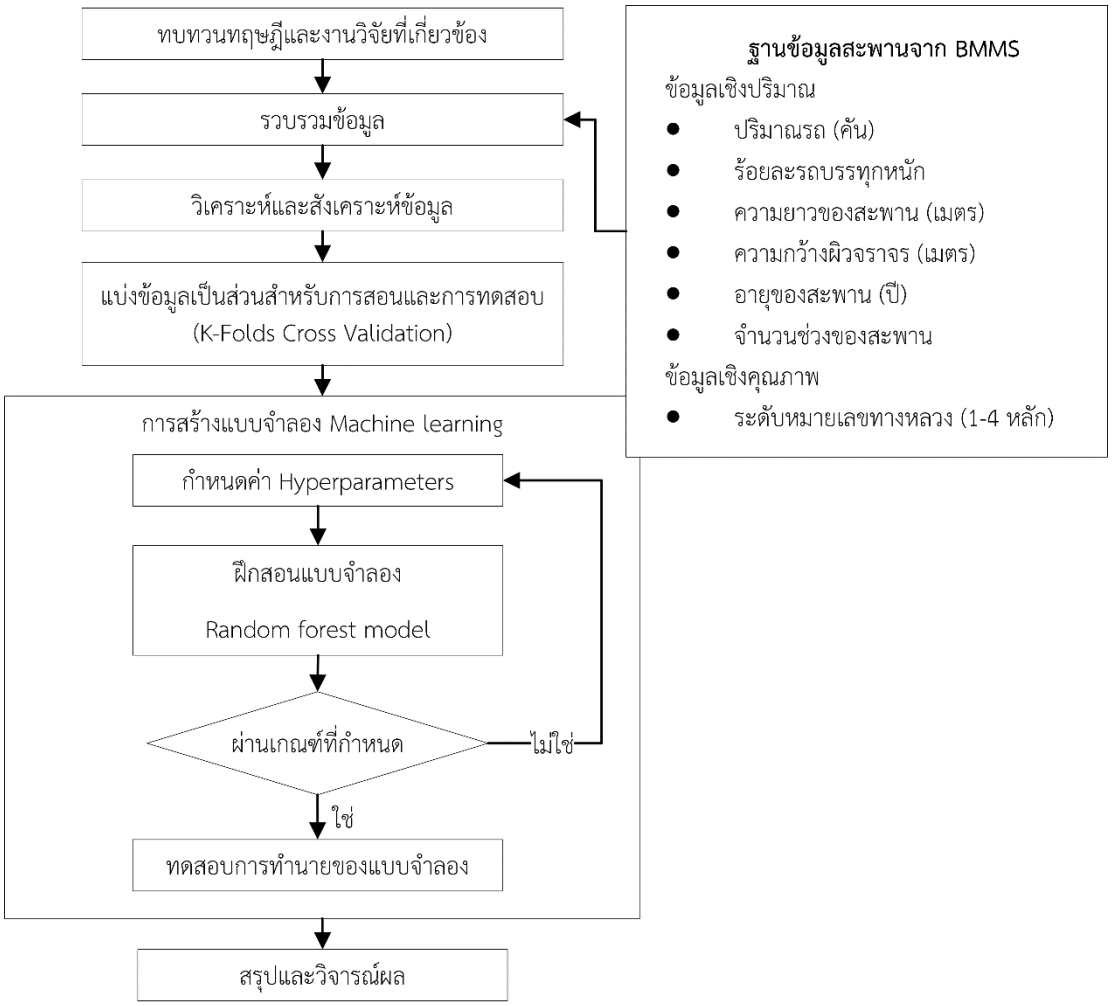
2. สร้างแบบจำลองป่าแบบสุ่ม ที่ใช้คาดการณ์ความเสียหายที่เกิดกับโครงสร้างสะพาน และคาดการณ์แนวโน้มของสะพานที่ต้องการการบำรุงรักษาในอนาคต

วิธีดำเนินการวิจัย

1. ผังการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัยมีรายละเอียดดังนี้ดังแสดงใน

ภาพที่ **1**



ภาพที่ 1 แผนผังการสร้างแบบจำลองป่าแบบสุ่ม

2. ตัวแปรที่ศึกษาและการรวบรวมข้อมูล

ในการวิจัยนี้ เป็นการทำนายระดับความเสียหายที่เกิดกับสะพานโดยใช้วิธีป่าแบบสุ่ม โดยใช้ข้อมูลปัจจัยที่เกี่ยวข้องในการสอนแบบจำลอง ซึ่งแบ่งเป็น 2 ประเภทหลัก ๆ คือ

1) ข้อมูลที่เกี่ยวข้องกับการจราจร เนื่องจากการเสื่อมสภาพของสะพานเกี่ยวข้องโดยตรงกับการใช้งานสะพาน ได้แก่ ข้อมูลปริมาณรถรายวัน, ข้อมูลร้อยละรถบรรทุกหนัก และระดับหมายเลขทางหลวง

2) ข้อมูลทางกายภาพของสะพาน เนื่องจากประเภทของสะพานที่แตกต่างกันการใช้งานและความทนทานของสะพานย่อมแตกต่างกัน ได้แก่ ความยาวของสะพาน, ความกว้างผิวจราจร, จำนวนช่วงของสะพาน และอายุของสะพาน

ในการรวบรวมข้อมูลจะใช้ข้อมูลสะพานจากฐานข้อมูลการบำรุงสะพานของกรมทางหลวง โดยจะจำแนกตามระดับความเสียหายของสะพาน (สำนักก่อสร้างสะพาน กรมทางหลวง, n.d.) โดยแบ่งเป็น 4 ระดับความเสียหายได้แก่ ดีมาก (5), ดีพอใช้ (4), พอใช้ (3), ควรซ่อมแซม (2) ซึ่งไล่ระดับตั้งแต่ 5 ถึง 2 ระดับการใช้งานสะพานสามารถแบ่งออกได้เป็น 4 ประเภท คือ ระดับการใช้งานโดยรวม (Overall Condition Rating: OCR), ระดับการใช้งานโครงสร้างสะพานส่วนบน (Superstructure Condition Rating: SUP_CR) เช่น พื้นสะพาน เป็นต้น ระดับการใช้งานโครงสร้างสะพานส่วนล่าง (Substructure Condition Rating: SUB_CR) เช่น คานรัดหัวเสา, ค้ำยันตอม่อ, เสาตอม่อ เป็นต้น ระดับการใช้งานโครงสร้างส่วนประกอบรอง (Secondary Component Condition Rating: SEC_CR) ได้แก่ หูช้าง, เชิงลาดสะพาน, ผิวทาง เป็นต้น

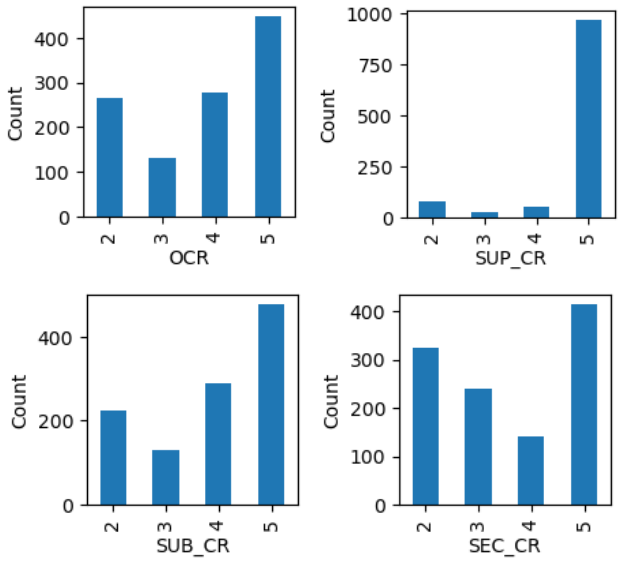
3. ข้อมูลนำเข้า

ชุดข้อมูลนำเข้าจะมีจำนวนข้อมูลตามที่แสดงในภาพที่ 2 โดยแบ่งเป็นจำนวนข้อมูลตามระดับความเสียหายรวม (ซ้ายบน), จำนวนข้อมูลตามระดับความเสียหายโครงสร้างส่วนบน (ขวาบน), จำนวนข้อมูลตามระดับความเสียหายโครงสร้างส่วนล่าง (ซ้ายล่าง), จำนวนข้อมูลตามระดับความเสียหายโครงสร้างรอง (ขวาล่าง) การแบ่งข้อมูลสำหรับทดสอบจะใช้วิธี K-Fold Cross-Validation (Brownlee, 2018)

4. แบบจำลองป่าแบบสุ่ม (Random Forest Model)

ในการทำวิจัยนี้ผู้ทำวิจัยได้ใช้ภาษาโปรแกรม Python ในการสร้างแบบจำลองป่าแบบสุ่ม เขียนบน Jupyter Notebook โดยมีไลบารี่ (Library) ที่เรียกใช้งานเกี่ยวกับ Machine Learning หลัก ๆ ได้แก่ Keras และ Scikit-learn (Géron, 2019)

การกำหนดค่า Hyperparameters ของวิธีป่าแบบสุ่มจะใช้วิธีการค้นหาแบบกริดดังที่แสดงในตารางที่ **1** โดย Hyperparameters ประกอบไปด้วย 3 พารามิเตอร์หลัก คือ Number of Features, Number of Trees และ Tree Depth



ภาพที่ 2 จำนวนข้อมูลทั้งหมดแยกตามระดับความเสียหายของสะพาน

ตารางที่ 1 Hyperparameters ที่เหมาะสมในการสร้างแบบจำลองป่าแบบสุ่ม

Hyperparameters	Grid Values	Optimize Value
Number of Features	1 ถึง 8	3
Number of Trees	1 ถึง 1000	100
Tree Depth	1 ถึง None	17

ผลการวิจัย

ในการประเมินประสิทธิภาพของแบบจำลองจะใช้วิธีคำนวณและแสดงผลออกมาในรูป Confusion Matrix ซึ่งเป็นเทคนิคในการสรุปผลประสิทธิภาพของอัลกอริทึมประเภทการจำแนกหมวดหมู่ ในปัญหาด้านการจำแนกหมวดหมู่ การใช้เพียงแค่ค่าความแม่นยำ (Accuracy) อย่างเดียวในชุดข้อมูลเป้าหมายที่ไม่เท่ากันอาจนำไปสู่การเข้าใจผิดได้ (Brownlee, 2016) จึงต้องนำค่าที่ได้มาคำนวณค่า Accuracy, Precision, Recall และ F1-Score (Korstanje, 2021) โดยมีวิธีการคำนวณคือ $\text{Accuracy}=\frac{\text{number of correct predictions}}{\text{number of all predictions}}$, $\text{Precision}=\frac{\text{number of correctly predicted positive instances}}{\text{number of total positive predictions you made}}$, $\text{F1-Score}=2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$ $\text{Recall}=\frac{\text{number of correctly predicted positive instances}}{\text{number of total positive instances in the dataset}}$ และ Support คือ จำนวนข้อมูลทั้งหมด ดังแสดงในตารางที่ 2 ค่าความแม่นยำในการทำนายของโครงสร้างโดยรวม, ตารางที่ 3 ค่าความแม่นยำในการทำนายของโครงสร้างส่วนบน, ตารางที่ 4 ค่าความแม่นยำในการทำนายของโครงสร้างส่วนล่าง และตารางที่ 5 ค่าความแม่นยำในการทำนายของโครงสร้างรอง

ตารางที่ 2 ค่าความแม่นยำในการทำนายของโครงสร้างโดยรวม

	PRECISION	RECALL	F1-SCORE	SUPPORT
OCR 2	0.75	0.79	0.77	266
OCR 3	0.61	0.36	0.45	131
OCR 4	0.72	0.83	0.77	276
OCR 5	0.95	0.94	0.94	448
ACCURACY			0.81	1121
MACRO AVG	0.76	0.73	0.73	1121
WEIGHTEDAVG	0.80	0.81	0.80	1121

ตารางที่ 3 ค่าความแม่นยำในการทำนายของโครงสร้างส่วนบน

	PRECISION	RECALL	F1-SCORE	SUPPORT
SUP 2	0.33	0.11	0.17	79
SUP 3	0.00	0.00	0.00	23
SUP 4	0.00	0.00	0.00	51
SUP 5	0.87	0.98	0.92	968
ACCURACY			0.85	1121
MACRO AVG	0.30	0.27	0.27	1121
WEIGHTED AVG	0.78	0.85	0.81	1121

ตารางที่ 4 ค่าความแม่นยำในการทำนายของโครงสร้างส่วนล่าง

	PRECISION	RECALL	F1-SCORE	SUPPORT
SUB 2	0.68	0.77	0.72	224
SUB 3	0.54	0.35	0.42	129
SUB 4	0.73	0.82	0.77	290
SUB 5	0.95	0.91	0.93	479
ACCURACY			0.79	1121
MACRO AVG	0.73	0.71	0.71	1121
WEIGHTED AVG	0.79	0.79	0.79	1121

ตารางที่ 5 ค่าความแม่นยำในการทำนายของโครงสร้างรอง

	PRECISION	RECALL	F1-SCORE	SUPPORT
SEC 2	0.46	0.47	0.46	324
SEC 3	0.46	0.34	0.39	241
SEC 4	0.42	0.12	0.19	141
SEC 5	0.60	0.82	0.69	415
ACCURACY			0.53	1121
MACRO AVG	0.49	0.44	0.43	1121
WEIGHTED AVG	0.51	0.53	0.50	1121

จากตารางที่ 2 เมื่อพิจารณาแยกตามหมวดหมู่ระดับความเสียหายแล้วจะเห็นได้ว่า ในระดับความเสียหายที่มีปริมาณข้อมูลน้อยจะมีค่า F1-Score ต่ำมาก เมื่อเทียบกับระดับความเสียหายที่มีปริมาณข้อมูลมาก เช่น OCR 5 แต่เมื่อดูค่าเปรียบเทียบแบบ Weighted Avg. แล้วจะพบว่า เมื่อมีการกระจายค่าน้ำหนักของปริมาณข้อมูลเข้าไปร่วมด้วย ผลลัพธ์ที่ได้ไม่แตกต่างจากความแม่นยำรวมมากนัก ตามตารางที่ 3 เมื่อชุดข้อมูลนำเข้าที่มีระดับความเสียหายแตกต่างกันมาก ดังจะดูได้จาก Support เกือบทั้งหมดไปตกอยู่ที่ SUP 5 ทำให้เมื่อพิจารณาค่า F1-Score แยกเป็นรายระดับความเสียหายแล้ว ตัวแบบจำลองไม่สามารถทำนายข้อมูลในระดับความเสียหายที่มีปริมาณข้อมูลต่ำกว่าได้เลย

จากตารางที่ 4 เมื่อพิจารณา Support พบว่าการกระจายตัวของปริมาณข้อมูลคล้ายกับระดับความเสียหายโครงสร้างโดยรวม และความแม่นยำที่ได้มีค่าใกล้เคียงกันมาก และจากตารางที่ 5 พบว่าในโครงสร้างประกอบรอง การกระจายตัวของข้อมูลไม่แตกต่างกับโครงสร้างโดยรวมมากนัก แต่เมื่อพิจารณาความแม่นยำและ Weighted Avg จะเห็นว่าค่าที่ได้จากการคาดการณ์ด้วยแบบจำลองมีค่าต่ำมาก แสดงว่าการคาดการณ์การบำรุงสะพานของโครงสร้างส่วนนี้ ไม่สัมพันธ์กับชุดข้อมูลนำเข้า

อภิปรายผลการวิจัย

จากการใช้แบบจำลองป่าแบบสุ่มในการคาดการณ์ประสิทธิภาพของสะพาน พบว่าผลลัพธ์ที่ได้ออกมาจากแบบจำลองสามารถคาดการณ์ได้โดยมีความแม่นยำอยู่ที่ 0.81 สำหรับโครงสร้างโดยรวมของสะพาน ทั้งนี้เมื่อพิจารณาแยกออกมาในแต่ละส่วนประกอบย่อยของสะพาน จะพบว่า ในโครงสร้างส่วนบน เช่น แผ่นพื้น แบบจำลองป่าแบบสุ่มจะคาดการณ์ระดับความเสียหายอยู่ที่ระดับ 5 เป็นส่วนใหญ่ ซึ่งส่งผลอย่างมากเนื่องจากจำนวนชุดข้อมูลนำเข้าที่ไม่สมดุลกันจะยากต่อการฝึกสอนแบบจำลอง (Brownlee, 2020) และอาจหมายความว่า แบบจำลองมีความคลาดเคลื่อนสูงเนื่องจากไม่สามารถทำนายระดับความเสียหายที่เหลือได้เลย หรือโครงสร้างในส่วนนี้มีความเสียหายต่ำ สำหรับโครงสร้างส่วนประกอบรอง เช่น หูช้าง, กำแพงกันดิน ผลลัพธ์จากการคาดการณ์ความเสียหายมีความแม่นยำต่ำมาก สาเหตุที่เป็นไปได้อาจเกิดจากชุดปัจจัยข้อมูลนำเข้าอาจไม่สัมพันธ์กับระดับความเสียหาย นั่นคือ ชุดข้อมูลนำเข้าไม่มีความเกี่ยวข้องกับประสิทธิภาพของสะพานในส่วนนี้ ทำให้ความแม่นยำในคาดการณ์ในส่วนนี้มีค่าน้อยและท้ายสุดในการคาดการณ์ความเสียหาย

ที่เกิดขึ้นกับโครงสร้างส่วนล่าง ความแม่นยำที่ได้จากการคาดการณ์ในส่วนนี้ไม่แตกต่างจากการคาดการณ์ประสิทธิภาพของโครงสร้างโดยรวมของสะพาน

ข้อเสนอแนะ

1. ในการสร้างแบบจำลองป่าแบบสุ่มมีข้อควรระวังเรื่องข้อมูลนำเข้าเป็นอย่างมาก เนื่องจากแบบจำลองค่อนข้างอ่อนไหวต่อชุดข้อมูลนำเข้า หากใช้ชุดข้อมูลที่มีความน่าเชื่อถือต่ำ จะส่งผลอย่างมากต่อประสิทธิภาพของแบบจำลอง จึงควรคัดกรองข้อมูลก่อนนำมาใช้ในแบบจำลอง

2. หากต้องการเพิ่มประสิทธิภาพของแบบจำลองในอนาคต สามารถทำได้โดยเพิ่มจำนวนชุดข้อมูลนำเข้าที่ผ่านการคัดกรองเรียบร้อยแล้ว และปรับพารามิเตอร์ให้เข้ากับชุดข้อมูล สิ่งเหล่านี้จะช่วยให้สามารถประยุกต์ใช้แบบจำลองในการคาดการณ์ประสิทธิภาพ การบำรุงรักษาสะพานในอนาคตได้

เอกสารอ้างอิง

สำนักก่อสร้างสะพาน กรมทางหลวง. (n.d.). *คู่มือการตรวจสอบสะพานกรมทางหลวง (BMMS)*. กรุงเทพฯ: สำนักก่อสร้างสะพาน กรมทางหลวง.

Brownlee, J. (2016, November 17). What is a Confusion Matrix in Machine Learning. *MachineLearningMastery.Com*. from https://machinelearningmastery.com/confusion-matrix-machine-learning/.

Brownlee, J. (2018, May 22). A Gentle Introduction to k-fold Cross-Validation. *MachineLearningMastery.Com*. from https://machinelearningmastery.com/k-fold-cross-validation/.

Brownlee, J. (2020, February 16). Why Is Imbalanced Classification Difficult? *MachineLearningMastery.Com*. from https://machinelearningmastery.com/imbalanced-classification-is-hard/.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Second edition). O'Reilly.

Korstanje, J. (2021, August 31). *The F1 score*. Medium. from https://towardsdatascience.com/the-f1-score-bec2bbc38aa6.

Visualizing cross-validation behavior in scikit-learn. (n.d.). Scikit-Learn. Retrieved November 26, 2022, from https://scikit-learn/stable/auto_examples/model_selection/plot_cv_indices.html.