การวิเคราะห์น้ำหนักรถไฟด้วย AI
Train Weight Analysis with AI

กรองกาญจน์ ชื่นชม ชลสิทธิ์ อัครธิติโรจน์ ธิวร์ อิษฎาวงศ์ และ รศ. ดร. บุญชัย แสงเพชรงาม
[1,2,3] ภาควิชาวิศวกรรมโยธา คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อ
การวิเคราะห์น้ำหนักของรถไฟเป็นองค์ประกอบสำคัญที่ส่งผลต่อความปลอดภัยและประสิทธิภาพของระบบขนส่งทางราง โครงงานนี้มีวัตถุประสงค์เพื่อพัฒนาระบบประเมินน้ำหนักของรถไฟจากสัญญาณเซนเซอร์ โดยใช้ทั้งวิธีการแบบ manual ผ่านโปรแกรม MATLAB และการประเมินโดยใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) ซึ่งประกอบด้วยการฝึกสอนโมเดลด้วยเทคนิค LSTM และ Isolation Forest และการจัดกลุ่มข้อมูลด้วย clustering ได้แก่ k-means และ LSTM โดยใช้ข้อมูลที่ได้จากการประเมินแบบ manual เป็นฐานสำหรับการเรียนรู้ โมเดลที่ได้จะถูกนำไปทดสอบกับไฟล์ใหม่เพื่อตรวจสอบความแม่นยำของระบบ AI เปรียบเทียบกับค่าจริงจากวิธี manual ผลการศึกษาจะช่วยประเมินความเป็นไปได้ในการนำ AI มาใช้แทนแรงงานมนุษย์ในการประเมินน้ำหนักรถไฟในอนาคต
คำสำคัญ: น้ำหนักรถไฟ, การวิเคราะห์, วิธีแมนนวล, ปัญญาประดิษฐ์, การเปรียบเทียบ

Abstract (Subject without numbering)

Train weight analysis plays a critical role in ensuring the safety and efficiency of railway transportation systems. This project aims to develop a system for estimating train weight using sensor signals, employing both a manual approach via MATLAB and an automated approach using artificial intelligence (AI). The AI model is trained using techniques such as Long Short-Term Memory (LSTM) and K-means clustering, with training data derived from manual evaluations. The trained model is then tested on new, unseen data to evaluate its accuracy against the ground truth obtained from manual calculations. The results of this study help assess the feasibility of applying AI to replace human-based train weight evaluations in real-world railway systems.

Keywords: Train Weight, Analysis, Manual Method, Artificial Intelligence (AI), Comparison

1.บทนำ

ในการขนส่งทางราง น้ำหนักของรถไฟแต่ละขบวนเป็นข้อมูลสำคัญที่ใช้ในการบริหารจัดการขบวนรถ การวางแผนโลจิสติกส์ และการควบคุมความปลอดภัย การประเมินน้ำหนักอย่างแม่นยำจึงเป็นสิ่งจำเป็น โดยเฉพาะในระบบขนส่งที่มีปริมาณการเดินรถสูง โครงงานนี้มีวัตถุประสงค์เพื่อพัฒนาระบบที่สามารถประเมินน้ำหนักของรถไฟจากข้อมูลสัญญาณที่ได้จากเซนเซอร์ภาคสนาม โดยแบ่งออกเป็นสองส่วนหลัก ได้แก่ วิธีการประเมินแบบแมนนวล (manual) และการประเมินโดยใช้ปัญญาประดิษฐ์ (AI) ในขั้นตอนแรกจะใช้โปรแกรม MATLAB วิเคราะห์ไฟล์ข้อมูลที่บันทึกจากเซนเซอร์ เพื่อหาน้ำหนักของรถไฟจริงในแต่ละขบวน ซึ่งจะถือเป็นค่ามาตรฐาน (ground truth) สำหรับฝึกสอนโมเดล AI (supervised LSTM) จากนั้นจะนำโมเดลที่ฝึกแล้วไปประเมินข้อมูลไฟล์ใหม่ที่ไม่เคยใช้ในการฝึกสอน เพื่อเปรียบเทียบความแม่นยำกับวิธี manual

1.1 วัตถุประสงค์ของโครงงานนี้ประกอบด้วย

1.1.1. เพื่อประยุกต์ใช้เทคโนโลยี AI ในงานวิศวกรรมการขนส่งทางราง

1.1.2. เพื่อเพิ่มประสิทธิภาพและความปลอดภัยของระบบขนส่งทารางๆ

1.1.3. เพื่อหา loading factor ทำให้ทราบอัตราการใช้งานบรรทุกน้ำหนัก

1.1.4. เพื่อหาชนิดของรถไฟที่ผ่าน

1.2. ขอบเขตของการศึกษานี้จะมุ่งเน้นไปที่การเปรียบเทียบความแม่นยำของ AI กับการประเมินแบบ manual ในสองด้านหลัก ได้แก่

1.2.1. ความถูกต้องในการประเมินน้ำหนัก

1.2.2. ความถูกต้องในการนับจำนวนคันรถ

1.3 ผลลัพธ์ที่คาดว่าจะได้รับจากการศึกษานี้

คือความสามารถในการออกแบบและพัฒนาระบบ AI ที่สามารถประเมินน้ำหนักของรถไฟได้อย่างแม่นยำ โดยไม่จำเป็นต้องอาศัยแรงงานมนุษย์ และสามารถนำไปต่อยอดใช้ในการวิเคราะห์ข้อมูลในระบบขนส่งทางรางได้ในอนาคต

2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2.1 ข้อมูลทั่วไปของรถไฟ

2.1.1 ส่วนประกอบของรถไฟ

รถไฟเป็นระบบขนส่งที่ประกอบด้วยหัวรถจักร ตู้โดยสาร และตู้สินค้า โดยหัวรถจักรมักมีล้อ 12 ล้อ ตู้โดยสารทั่วไป และตู้สินค้ามี 8 ล้อ มี 8 ล้อ ทั้งนี้เพื่อรองรับน้ำหนักบรรทุกและความเสถียรในการเดินรถ

2.2 การบันทึกสัญญาณน้ำหนักล้อรถไฟ

สถานีรถไฟที่ใช้ในการเก็บข้อมูลในโครงงานนี้ คือ สถานีพานทอง ซึ่งเป็นสถานีรถไฟระดับชั้นที่ 3 ตั้งอยู่ในตำบลพานทอง อำเภอพานทอง จังหวัดชลบุรี ลักษณะของขบวนรถไฟที่เดินทางผ่านสถานีนี้ส่วนใหญ่เป็นขบวนรถไฟบรรทุกสินค้าทั่วไป รวมถึงรถไฟบรรทุกน้ำมัน และก๊าซ การติดตั้งเซนเซอร์เพื่อวัดแรงกระทำจากล้อรถไฟ (Strain Gauge) ดำเนินการติดตั้งที่ช่องสัญญาณ Ch1, Ch2, Ch3 และ Ch4 โดย Ch1 และ Ch2 ใช้สำหรับขาออกจากกรุงเทพมหานครมุ่งหน้าไปยังแหลมฉบัง ในขณะที่ Ch3 และ Ch4 ใช้สำหรับขาเข้าจากแหลมฉบังกลับเข้าสู่กรุงเทพมหานคร

2.2.1 วิธีการบันทึกสัญญาณใช้ Strain gauge

ติดตั้งใต้รางเพื่อตรวจวัดแรงเฉือนเมื่อรถไฟแล่นผ่าน ซึ่งแปลงเป็นค่า Strain และบันทึกด้วย Data Logger เพื่อนำมาคำนวณเป็นค่าน้ำหนักล้อ

2.2.2 รูปแบบของข้อมูล

ข้อมูลถูกบันทึกในไฟล์ .ks2 ซึ่งอาจอยู่ในรูปแบบ V/V หรือ Tonf หากเป็น V/V ต้องทำการ calibrate ก่อนใช้งาน โดยข้อมูลประกอบด้วยค่าเวลาและค่าสัญญาณจากเซนเซอร์แต่ละช่อง (Ch1 ถึง Ch4)

2.3 การแปลงข้อมูลให้อยู่ในรูปแบบกราฟ

ใช้โปรแกรม MATLAB เพื่ออ่านไฟล์ .ks2 และแสดงข้อมูลในรูปแบบกราฟ เพื่อให้ง่ายต่อการวิเคราะห์จุดพีค (peak) และรูปแบบของน้ำหนัก

2.4 การประยุกต์ใช้งาน AI กับการตรวจจับ peak และแยกคันรถไฟ

2.4.1 การตรวจจับ peak ของล้อรถไฟด้วย AI

ใช้เทคนิค Deep Learning เช่น LSTM ซึ่งเป็นโครงข่ายประสาทเทียมที่เหมาะสำหรับข้อมูลตามลำดับเวลา เพื่อหาจุดพีคของล้อรถไฟได้อย่างแม่นยำ นอกจากนี้ยังมี Isolation Forest สำหรับตรวจจับ anomaly ซึ่งเหมาะกับการหา peak ที่มีลักษณะโดดเด่นจากข้อมูลทั่วไป

2.4.2 การแบ่งจำนวนคันจาก 1 ขบวนรถไฟด้วย AI

ใช้การจัดกลุ่มแบบ Unsupervised Learning เช่น K-means และ DBSCAN เพื่อแยกล้อออกเป็นคันรถ โดยอิงจากตำแหน่งล้อและความหนาแน่นของข้อมูล เพื่อให้สามารถระบุจำนวนคันได้อย่างแม่นยำโดยไม่ต้องระบุล่วงหน้า

2.5 การประมวลผลสัญญาณ (Signal Processing)

ใช้การประมวลผลสัญญาณเพื่อแปลงค่าที่วัดได้จาก strain gauge ให้เป็นน้ำหนัก โดยต้องกรองสัญญาณรบกวนและใช้เทคนิคต่าง ๆ เพื่อคัดแยกข้อมูลที่เกี่ยวข้องกับแรงกระทำจริงจาก noiseเพื่อให้ง่ายต่อการนำไปวิเคราะห์หรือใช้งานต่อในระบบอื่น ๆ

2.6 งานวิจัยที่เกี่ยวข้อง

2.6.1 Estimating the load weight of freight trains using machine learning (2023)

เป็นงานวิจัยจาก KTH Royal Institute of Technology ประเทศสวีเดน ได้พัฒนาระบบ RMD (Railcar Mass Determination)ซึ่งเป็นอัลกอริธึมสำหรับการประเมินน้ำหนักของตู้รถไฟขนส่งสินค้า ด้วย Machine Learning การศึกษาได้ทดลองใช้วิธีการเรียนรู้ของเครื่องหลายรูปแบบ ได้แก่ Polynomial Regression, K-Nearest Neighbors, Regression Tree, Random Forest และ Support Vector Regression

2.6.2 Application of AI in railway freight loading(1999)Geng, Zhang, Zhu และ Zhong (1999)

ได้นำเสนอระบบผู้เชี่ยวชาญ (Expert System) ที่พัฒนาขึ้นเพื่อช่วยในการตัดสินใจเกี่ยวกับการบรรทุกสินค้าบนรถไฟ งานวิจัยนี้มุ่งเน้นการวางแผนและจัดเรียงโหลดของรถไฟโดยคำนึงถึงข้อจำกัดด้านเทคนิค เช่น น้ำหนักสูงสุดของตู้สินค้า จุดศูนย์ถ่วง ความมั่นคงของขบวนรถ และการใช้พื้นที่อย่างมีประสิทธิภาพ

2.6.3 A Systematic Review on K-Means Clustering Techniques

บทความนี้ได้ทบทวนเทคนิคต่าง ๆ ในการพัฒนาและปรับปรุงอัลกอริธึม K-means ซึ่งเป็นวิธีการจัดกลุ่มข้อมูลแบบ Unsupervised Learning จุดเด่นของ K-means คือการแบ่งข้อมูลออกเป็นกลุ่มที่มีความคล้ายคลึงกันสูงภายในกลุ่ม และมีความแตกต่างกันระหว่างกลุ่ม โดยอาศัยหลักการลดระยะห่างของข้อมูลจาก centroid บทความยังกล่าวถึงข้อจำกัด เช่น ความไวต่อ outliers, การเกิด empty clusters และปัญหาในการเลือกตำแหน่งเริ่มต้นของ centroid ซึ่งเป็นความท้าทายในการใช้งานจริง

2.6.4 Railway freight car axle load spectra and its implications for track maintenance (2002)

งานวิจัยโดย Zarembski และ Resor (2002) ได้ศึกษาข้อมูลโหลดของเพลารถไฟเพื่อวิเคราะห์ลักษณะการกระจายของน้ำหนักที่กระทำต่อรางรถไฟ โดยใช้ข้อมูลจากระบบตรวจวัดอัตโนมัติ (Weigh-in-Motion Systems) ผลการวิเคราะห์พบว่าความแปรผันของโหลดส่งผลต่อการออกแบบ การบำรุงรักษา และการตรวจจับความเสียหายของโครงสร้างพื้นฐานของรางรถไฟ งานวิจัยนี้เน้นถึงความจำเป็นในการมีระบบประเมินน้ำหนักที่แม่นยำเพื่อวางแผนการซ่อมบำรุงอย่างมีประสิทธิภาพ

3 ระเบียบวิธีวิจัย

3.1 การวิเคราะห์ข้อมูลด้วยเอลกอลิทึมของMatlab

ในการประมวลผลข้อมูลที่ได้จากการทดลอง ได้ใช้โปรแกรม MATLAB ในการวิเคราะห์และสร้างกราฟ โดยเริ่มจากการอ่านข้อมูลที่ได้จากไฟล์ .ks2 ซึ่งเป็นรูปแบบเฉพาะของข้อมูลสัญญาณที่บันทึกได้จากเครื่องมือวัด จากนั้นทำการแปลงหน่วยของข้อมูลให้อยู่ในรูปของน้ำหนักที่เหมาะสม โดยมีหน่วยเป็นตันแรง (Tonf)

หลังจากนั้น ได้ทำการวิเคราะห์หาค่าต่ำสุดของสัญญาณ (valleys) ในแต่ละช่องสัญญาณ โดยใช้งานฟังก์ชัน findpeaks ของ MATLAB ซึ่งสามารถปรับให้ตรวจจับจุดต่ำสุดได้ด้วยการกลับค่าข้อมูล จากนั้นจึงกำหนดช่วงค่าที่ต้องการวิเคราะห์ และแสดงผลในรูปแบบกราฟที่เข้าใจง่าย

กราฟผลลัพธ์ประกอบด้วย 2 กราฟ ได้แก่ กราฟของช่องสัญญาณที่ 1 ซึ่งใช้สีฟ้า (น้ำเงินอ่อน) โดยกำหนดช่วงแกน Y อยู่ระหว่าง 3.5 ถึง 8 Tonf และกราฟของช่องสัญญาณที่ 2 ซึ่งใช้สีชมพู โดยกำหนดช่วงแกน Y อยู่ระหว่าง -9 ถึง -2 Tonf ตามรูปที่ 2.1ทั้งสองกราฟแสดงความสัมพันธ์ระหว่างน้ำหนักกับเวลา โดยจุดต่ำสุดที่ตรวจพบจะแสดงด้วยวงกลมสีแดงเพื่อให้มองเห็นได้ชัดเจน

จากการวิเคราะห์ดังกล่าว ค่าที่จุดต่ำสุดทั้งหมดในทั้งสองช่องสัญญาณสามารถนำมารวมกัน เพื่อประมาณค่าน้ำหนักรวมของล้อรถไฟที่ผ่านตำแหน่งจุดตรวจวัดในช่วงเวลานั้นได้อย่างแม่นยำ

3.2 การกำหนดโมเดล AI

3.2.1 การกำหนดโมเดล AI จำนวนพีค

ในการคัดแยกพีคที่เกี่ยวข้องกับการกดของล้อรถไฟบนเซนเซอร์น้ำหนัก ได้มีการเลือกใช้แบบจำลองทางปัญญาประดิษฐ์ ได้แก่ **LSTM (Long Short-Term Memory)** ซึ่งเป็น deep learning และ **Isolation Forest** ซึ่งเป็น machine learning ซึ่งมีคุณสมบัติเด่นในการวิเคราะห์ข้อมูลที่อยู่ในรูปแบบลำดับเวลา (time series) แบบจำลอง LSTM เป็นโครงข่ายประสาทเทียมชนิดหนึ่งที่เหมาะสมสำหรับการเรียนรู้จากข้อมูลลำดับต่อเนื่อง โดยสามารถจดจำบริบทก่อนหน้าและแยกแยะรูปแบบของพีคที่เกิดจากล้อรถไฟกับพีคที่เกิดจากสัญญาณรบกวนได้อย่างมีประสิทธิภาพ ส่วน Isolation Forest ถูกนำมาใช้ควบคู่เพื่อคัดกรองและตรวจจับค่าที่มีลักษณะผิดปกติจากชุดข้อมูลทั้งหมด ซึ่งช่วยเสริมให้ระบบสามารถจำแนกพีคที่สำคัญได้แม่นยำยิ่งขึ้น

3.2.2 การกำหนดโมเดล AI จำนวนคัน

ในการวิเคราะห์จำนวนคัน เราได้ใช้โมเดล K-Means, DBSCAN ซึ่งเป็น Algorithm ประเภท Unsupervised ใช้สำหรับกรณีที่สามารถประมาณจำนวนกลุ่ม (จำนวนคันรถ) ได้ล่วงหน้า (ต้องรู้จำนวนคันรถก่อนแบ่งคัน) ซึ่งจะทำการจัดกลุ่มพีคโดยอ้างอิงจากค่าศูนย์กลางของแต่ละกลุ่ม ส่วนโมเดล **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)** เหมาะสำหรับกรณีที่ข้อมูลมีความหนาแน่นไม่เท่ากัน และจำนวนกลุ่มไม่แน่นอนล่วงหน้า โดยอาศัยระยะห่างระหว่างจุดและความหนาแน่นของข้อมูลเป็นเกณฑ์หลักในการจัดกลุ่ม

3.3 การเตรียมข้อมูลสำหรับ AI

3.3.1 การเตรียมข้อมูลสำหรับการคัดแยกพีค

ในการตรวจจับจุดพีคล้อรถไฟ ระบบได้มีการเตรียมข้อมูลแบบมีป้ายกำกับ (Label)สำหรับใช้ฝึกโมเดลLSTMแบบ Supervisedเพื่อให้สามารถเรียนรู้ลักษณะของพีคที่เกิดขึ้นจริงได้อย่างแม่นยำเหตุผลที่เลือกใช้ข้อมูลเดียวกันในทั้งสองโมเดลเป็นเพราะต้องการ ประเมินความแม่นยำของ Isolation Forest โดยการเปรียบเทียบกับคำตอบจริง (label) เช่นเดียวกับ LSTM ซึ่งช่วยให้สามารถวัดผลของโมเดลแบบไม่มีผู้สอนได้อย่างเป็นรูปธรรม โดยใช้เกณฑ์เดียวกัน ในงานวิจัยนี้ได้กำหนดป้ายกำกับข้อมูลโดยใช้เลข 0 แทนตำแหน่งข้อมูลน้ำหนักที่ ไม่ใช่ล้อรถไฟ และเลข 1 แทนตำแหน่งที่ ตรงกับล้อรถไฟ ซึ่งมีความเกี่ยวข้องโดยตรงกับสัญญาณน้ำหนักที่เกิดขึ้นจากแรงกดของล้อบนราง กระบวนการกำหนด label ดังกล่าวอ้างอิงจากผลการวิเคราะห์เบื้องต้นที่ได้จาก MATLAB โดยเฉพาะการตรวจจับจุดต่ำสุดของสัญญาณ (valleys) ด้วยฟังก์ชัน findpeaks และการพิจารณารูปแบบของสัญญาณในบริเวณที่ล้อรถไฟผ่าน จุดที่มีความสอดคล้องกับตำแหน่งล้อจะถูกกำหนดเป็น label 1 ส่วนจุดอื่นที่ไม่เตรงกับเงื่อนไขดังกล่าวจะถูกกำหนดเป็น label 0 การเตรียมข้อมูลแบบนี้ช่วยให้โมเดลสามารถเรียนรู้ความแตกต่างระหว่างตำแหน่งที่เป็นล้อจริงกับสัญญาณรบกวนหรือค่าที่ไม่เกี่ยวข้องได้อย่างแม่นยำ และส่งผลให้ประสิทธิภาพของโมเดลดีขึ้นเมื่อถูกนำไปใช้งานจริง

time	ch1	ch2	ch3	ch4	label1	label2	label3	label4
0	1.421963	6.429238	6.251481	9.826184	0	0	0	0
0.0005	1.424589	6.432991	6.254605	9.829838	0	0	0	0
0.001	1.418795	6.426917	6.248609	9.823395	0	0	0	0
0.0015	1.422062	6.430109	6.250924	9.826295	0	0	0	0
0.002	1.424353	6.434136	6.253542	9.829572	0	0	0	0
0.0025	1.42076	6.432018	6.252076	9.826188	0	0	0	0
0.003	1.417972	6.431748	6.250979	9.825601	0	0	0	0

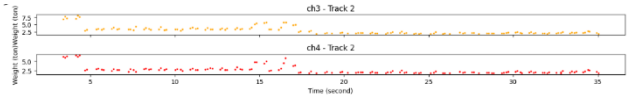
รูปที่ 3.1 ตารางตัวอย่างข้อมูลที่ระบุ label

3.3.2 การเตรียมข้อมูลสำหรับการแบ่งจำนวนคัน

เนื่องจากแบบจำลอง AI ที่ใช้ในการวิเคราะห์จำนวนคันรถไฟ ได้แก่ LSTM K-Means และ DBSCAN ล้วนเป็นเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) จึงไม่จำเป็นต้องเตรียมข้อมูลแบบมีป้ายกำกับ (Label) สำหรับใช้ฝึกโมเดลเหมือนกับเทคนิคแบบ Supervised

3.3.2.1 การดึงและกรองตำแหน่งพีค

สำหรับการแบ่งจำนวนคันรถไฟ ระบบจะอ้างอิงตำแหน่งของจุดพีค ที่ได้จากการ คัดพีคแบบ Manual (Signal Processing) โดยผู้เชี่ยวชาญ ไม่ได้ใช้พีคที่ตรวจจับจาก AI มาใช้ในขั้นตอนนี้ โดยพีคดังกล่าวได้จากข้อมูลน้ำหนักของล้อรถไฟในรูปแบบ Time Series ซึ่งถูกประมวลผลใน MATLAB และแปลงค่ากลับด้าน (Invert) เพื่อให้ลักษณะของพีคหันขึ้นด้านบน เพื่อให้ง่ายต่อการกรองและหาตำแหน่งที่มีความถี่ของพีคสูง ตำแหน่งพีคเหล่านี้จะถูกนำมาใช้เป็น Input สำหรับการจัดกลุ่มล้อในแต่ละคัน ด้วยวิธี DBSCAN หรือ K-Means โดยไม่มีการใช้ข้อมูล Label และไม่มีการนำผลจาก AI ที่ใช้คัดพีคมาใช้ต่อในขั้นตอนการจัดกลุ่ม



รูปที่ 3.2 ตัวอย่างการจัดเตรียมข้อมูลสำหรับ Unsupervise

3.4 การแบ่งข้อมูล Data Set

ในการฝึกและประเมินประสิทธิภาพของแบบจำลอง AI (สำหรับ supervised LSTM และ Isolation Forest) ได้มีการแบ่งชุดข้อมูลทั้งหมดออกเป็น 3 ส่วน ได้แก่ ชุดฝึก (Training Set) ร้อยละ 70 ชุดตรวจสอบ (Validation Set) ร้อยละ 15 และชุดทดสอบ (Test Set) ร้อยละ 15 โดยการแบ่งข้อมูลดำเนินการในลักษณะต่อเนื่องตามลำดับเวลา (Time-based Split) เพื่อรักษาลักษณะความสมจริงของข้อมูลที่เป็นสัญญาณตามเวลา (Time Series)

ชุดฝึก (Training Set) คือชุดข้อมูลหลักที่ใช้สำหรับการเรียนรู้ของโมเดล โดยโมเดลจะศึกษารูปแบบของสัญญาณน้ำหนักที่เกี่ยวข้องกับการที่ล้อรถไฟกดลงบนเซนเซอร์ (เช่น ช่วงเวลาที่มีแรงกดสูงหรือต่ำ) เพื่อให้สามารถแยกแยะลักษณะของล้อและไม่ใช่ล้อได้ในภายหลัง

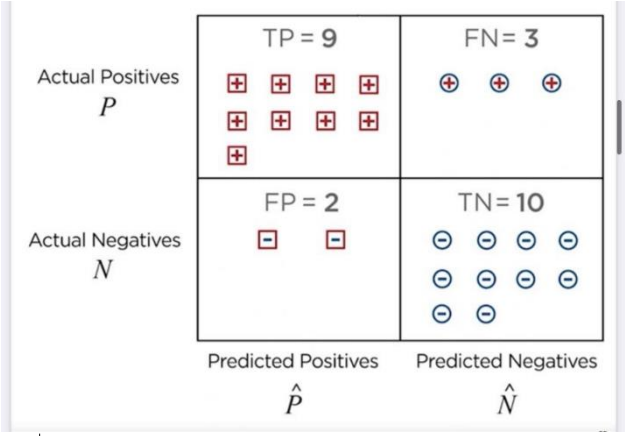
3.5 การประเมินผล AI

เมื่อสิ้นสุดกระบวนการฝึก โมเดลจะถูกนำไปทดสอบกับชุดข้อมูลที่ไม่เคยพบมาก่อน (Test Set) เพื่อประเมินความสามารถในการตรวจจับพีคของล้อรถไฟอย่างแม่นยำ โดยใช้ตัวชี้วัดหลักสองตัว ได้แก่ Accuracy และ F1-score ในการวัดประสิทธิภาพของโมเดล

Accuracy ใช้สำหรับวัดสัดส่วนของจำนวนการทำนายที่ถูกต้องเมื่อเปรียบเทียบกับข้อมูลทั้งหมด

F1-score ใช้สำหรับวัดความสมดุลระหว่างค่าความแม่นยำ (Precision) และค่าความสามารถในการเรียกคืนข้อมูลที่เกี่ยวข้อง (Recall) ซึ่งเหมาะสมอย่างยิ่งในกรณีที่ข้อมูลมีความไม่สมดุล

การประเมินในลักษณะนี้ช่วยสะท้อนความสามารถของโมเดลในการนำไปใช้งานจริงในการตรวจจับพีคที่เกิดจากล้อรถไฟได้อย่างมีประสิทธิภาพ



รูปที่ 3.3 ตัวอย่างการประเมินผลลัพของโมเดล AI

3.6 การเก็บข้อมูลการทดลอง Signal processing

ในการศึกษาครั้งนี้ได้ออกแบบการทดลองเพื่อเปรียบเทียบประสิทธิภาพระหว่างการประมวลผลโดยมนุษย์ (Human) กับระบบปัญญาประดิษฐ์ (AI) ในการวัดน้ำหนักและนับจำนวนคันรถไฟโดยมี-3.6.1 เงื่อนไขของการทดลอง

3.6.1 มนุษย์ (Human): ให้ผู้เชี่ยวชาญทำการวัดน้ำหนัก

และนับจำนวนคันตามกระบวนการที่ใช้ในปัจจุบัน ซึ่งอาศัยการอ่านค่าจากกราฟหรือเครื่องมือโดยตรง การประเมินของมนุษย์จะถูกใช้เป็น ค่ามาตรฐาน (Reference) และถือว่ามี Accuracy สูงสุด เนื่องจากเป็นการอ่านค่าจากประสบการณ์และวิธีที่ผ่านการยอมรับในภาคสนาม

3.6.2 AI (Artificial Intelligence): ใช้ระบบ AI หรือระบบอัตโนมัติที่พัฒนาในงานวิจัยนี้ในการวัดน้ำหนักและนับจำนวนคัน โดยใช้โมเดล เช่น LSTM Autoencoder, Isolation Forest, K-Means หรือ DBSCAN ตามลำดับขั้นตอนที่ออกแบบไว้ในบทก่อนหน้า

files name	จำนวนคัน Ch2 (manual)	จำนวคัน Ch3 (manual)	%error AI Ch2	%error AI Ch3
1				
2				
3				

ตารางที่3.1 ตัวอย่างการเปรียบเทียบ Signal processing

4 สรุปผลการทดลอง และประโยชน์

4.1 สรุปผลการทดลองในการใช้ AI วิเคราะห์น้ำหนักรถไฟ

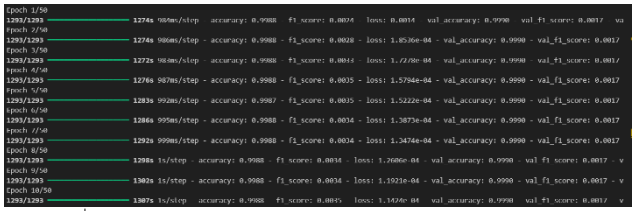
4.1.1 ผลการทดลองการตรวจจับ peak ของล้อรถไฟด้วย AI (การนับจำนวนล้อรถไฟ)

วิธีที่ 1: Unsupervised LSTM Autoencoder จากการทดลองใช้ LSTM Autoencoder

File_name	Humen1	Humen2	AI1	AI2	%	%
TEST#0010_A0047.KS2	0	129	0	134	0.00%	3.88%
TEST#0010_A0048.KS2	0	115	0	139	0.00%	20.87%
TEST#0010_A0051.KS2	91	0	89	0	2.20%	0.00%
TEST#0010_A0052.KS2	96	0	96	0	0.00%	0.00%
TEST#0010_A0053.KS2	50	0	50	0	0.00%	0.00%
TEST#0010_A0054.KS2	0	131	0	142	0.00%	8.40%
TEST#0010_A0056.KS2	0	80	0	81	0.00%	1.25%
TEST#0010_A0058.KS2	0	137	0	146	0.00%	6.57%
TEST#0010_A0060.KS2	0	144	0	142	0.00%	1.39%
TEST#0010_A0061.KS2	112	0	112	0	0.00%	0.00%
TEST#0010_A0063.KS2	114	0	114	0	0.00%	0.00%
TEST#0010_A0072.KS2	80	0	80	0	0.00%	0.00%
TEST#0010_A0073.KS2	0	109	0	110	0.00%	0.92%
TEST#0010_A0075.KS2	128	0	128	0	0.00%	0.00%
TEST#0010_A0082.KS2	145	0	146	0	0.69%	0.00%
TEST#0010_A0084.KS2	0	134	0	134	0.00%	0.00%
TEST#0010_A0085.KS2	0	126	0	126	0.00%	0.00%
TEST#0010_A0088.KS2	106	0	106	0	0.00%	0.00%
TEST#0010_A0090.KS2	0	128	0	128	0.00%	0.00%
TEST#0010_A0094.KS2	0	114	0	114	0.00%	0.00%
TEST#0010_A0097.KS2	0	130	0	130	0.00%	0.00%
TEST#0010_A0100.KS2	0	114	0	121	0.00%	6.14%
TEST#0010_A0101.KS2	111	0	117	0	5.41%	0.00%
TEST#0010_A0104.KS2	0	131	0	134	0.00%	2.29%
TEST#0010_A0106.KS2	0	102	0	68	0.00%	33.33%
TEST#0010_A0110.KS2	0	77	0	86	0.00%	11.69%
TEST#0010_A0112.KS2	78	0	78	0	0.00%	0.00%
TEST#0010_A0113.KS2	145	0	146	0	0.69%	0.00%
TEST#0010_A0114.KS2	78	0	78	0	0.00%	0.00%

รูปที่ 4.1 ตัวอย่างผลลัพธ์ในการใช้ Unsupervised LSTM Autoencoder ในการหา peak เปรียบเทียบกันหลายไฟล์

วิธีที่ 2: Supervised LSTM
การใช้โมเดล LSTM แบบมีการกำกับเพื่อทำนายจุด peak ให้ผลลัพธ์ที่ต่ำกว่าคาดแม้พยายามปรับ loss และ threshold



รูปที่ 4.2ตัวอย่างการแสดงผลการทำงานในการประเมินของ supervised LSTM

วิธีที่ 3: Isolation Forest

Isolation Forest สามารถจับพีคได้ครบ (Recall = 1.0000) แต่ให้ Precision ต่ำมากจนไม่สามารถใช้งานได้จริง

```
📊 Classification Report:
              precision    recall  f1-score   support

           0     0.0000    0.0000    0.0000   2774116
           1     0.0005    1.0000    0.0009      1290

    accuracy                         0.0005   2775406
   macro avg     0.0002    0.5000    0.0005   2775406
weighted avg     0.0000    0.0005    0.0000   2775406


🧩 Confusion Matrix:
[[      0 2774116]
 [      0    1290]]
```

รูปที่ 4.3 ตัวอย่างการแสดงผลการทำงานในการประเมินของ Isolation Forest

สรุปภาพรวม:

1. LSTM แบบ Unsupervised ให้ผลลัพธ์ดีที่สุดในกลุ่ม AI
2. การวิเคราะห์ด้วยมนุษย์ยังแม่นยำกว่า โดยเฉพาะในสัญญาณที่ซับซ้อน

4.1.2 ผลการทดลองการแบ่งจำนวนคันจาก 1 ขบวนรถไฟด้วย AI

วิธีที่ 1 : K-means Clustering

K-means สามารถแบ่งจำนวนคันได้ใกล้เคียงกับความจริง (Error < 10%) แต่ยังแบ่งผิดตำแหน่งจากโครงสร้างล้อจริง

File_name	number_of_cars_Humen1	number_of_cars_Humen2	kmean1	kmean2	%1	%2
TEST#0009_A0000.KS2	36	0	34	0	5.56%	0.00%
TEST#0009_A0001.KS2	36	36	39	35	8.33%	2.78%
TEST#0009_A0003.KS2	0	36	0	36	0.00%	0.00%
TEST#0009_A0004.KS2	35	0	35	0	0.00%	0.00%
TEST#0009_A0005.KS2	21	0	19	0	9.52%	0.00%
TEST#0010_A0000.KS2	0	28	0	28	0.00%	0.00%
TEST#0010_A0005.KS2	32	0	33	0	3.13%	0.00%
TEST#0010_A0006.KS2	0	20	0	21	0.00%	5.00%
TEST#0010_A0020.KS2	0	36	0	36	0.00%	0.00%
TEST#0010_A0021.KS2	36	0	36	0	0.00%	0.00%
TEST#0010_A0022.KS2	0	36	0	37	0.00%	2.78%
TEST#0010_A0023.KS2	21	0	20	0	4.76%	0.00%
TEST#0010_A0024.KS2	0	20	0	17	0.00%	15.00%
TEST#0010_A0026.KS2	0	15	0	15	0.00%	0.00%
TEST#0010_A0027.KS2	36	0	37	0	2.78%	0.00%
TEST#0010_A0031.KS2	0	20	0	22	0.00%	10.00%
TEST#0010_A0032.KS2	36	0	31	0	13.89%	0.00%
TEST#0010_A0034.KS2	0	35	0	34	0.00%	2.86%
TEST#0010_A0034.KS2	0	35	0	34	0.00%	2.86%
TEST#0009_A0000.KS2	36	0	34	0	5.56%	0.00%
TEST#0009_A0001.KS2	36	36	39	35	8.33%	2.78%
TEST#0009_A0003.KS2	0	36	0	36	0.00%	0.00%
TEST#0009_A0004.KS2	35	0	35	0	0.00%	0.00%
TEST#0009_A0005.KS2	21	0	19	0	9.52%	0.00%
TEST#0010_A0000.KS2	0	28	0	28	0.00%	0.00%
TEST#0010_A0005.KS2	32	0	33	0	3.13%	0.00%
TEST#0010_A0006.KS2	0	20	0	21	0.00%	5.00%
TEST#0010_A0020.KS2	0	36	0	36	0.00%	0.00%
TEST#0010_A0021.KS2	36	0	36	0	0.00%	0.00%
TEST#0010_A0022.KS2	0	36	0	37	0.00%	2.78%
TEST#0010_A0023.KS2	21	0	20	0	4.76%	0.00%
TEST#0010_A0024.KS2	0	20	0	17	0.00%	15.00%
TEST#0010_A0026.KS2	0	15	0	15	0.00%	0.00%
TEST#0010_A0027.KS2	36	0	37	0	2.78%	0.00%
TEST#0010_A0031.KS2	0	20	0	22	0.00%	10.00%
TEST#0010_A0032.KS2	36	0	31	0	13.89%	0.00%
TEST#0010_A0034.KS2	0	35	0	34	0.00%	2.86%

รูปที่ 4.4 ตัวอย่างผลลัพธ์ในการใช้ K-means ในการหาจำนวน

วิธีที่ 2 : DBSCAN

DBSCAN ให้ผลลัพธ์ความแม่นยำต่ำกว่า K-means อย่างชัดเจน โดยเฉพาะกับช่องสัญญาณที่มีล้อห่างกัน

File_name	number_of_cars_Humen1	number_of_cars_Humen2	dbscan1	dbscan2	%1	%2
TEST#0009_A0000.KS2	36	0	30	0	13.89%	0.00%
TEST#0009_A0001.KS2	36	36	24	19	27.78%	50.00%
TEST#0009_A0003.KS2	0	36	0	19	0.00%	50.00%
TEST#0009_A0004.KS2	35	0	28	0	17.14%	0.00%
TEST#0009_A0005.KS2	21	0	9	0	52.38%	0.00%
TEST#0010_A0000.KS2	0	28	0	22	0.00%	21.43%
TEST#0010_A0005.KS2	32	0	27	0	18.75%	0.00%
TEST#0010_A0006.KS2	0	20	0	16	0.00%	15.00%
TEST#0010_A0020.KS2	0	36	0	31	0.00%	11.11%
TEST#0010_A0021.KS2	36	0	29	0	16.67%	0.00%
TEST#0010_A0022.KS2	0	36	0	25	0.00%	33.33%
TEST#0010_A0023.KS2	21	0	13	0	38.10%	0.00%
TEST#0010_A0024.KS2	0	20	0	18	0.00%	10.00%
TEST#0010_A0026.KS2	0	15	0	8	0.00%	46.67%
TEST#0010_A0027.KS2	36	0	27	0	33.33%	0.00%
TEST#0010_A0031.KS2	0	20	0	10	0.00%	45.00%
TEST#0010_A0032.KS2	36	0	26	0	27.78%	0.00%
TEST#0010_A0034.KS2	0	35	0	29	0.00%	20.00%

รูปที่ 4.5 ตัวอย่างผลลัพธ์ในการใช้ DBSCAN ในการหาจำนวน cluster

สรุป: มนุษย์ยังสามารถแยกคันรถได้แม่นยำกว่า AI clustering เนื่องจากเข้าใจโครงสร้าง เช่น ล้อคู่, หัวรถ 6 ล้อ ฯลฯ ซึ่ง AI ไม่สามารถเรียนรู้ได้จากระยะห่างเพียงอย่างเดียว

4.2 ผลวิเคราะห์การทดลอง

จากการทดลองใช้โมเดล AI เพื่อวิเคราะห์จุดพีคล้อรถไฟและแบ่งจำนวนคันในขบวน พบปัญหาของแต่ละโมเดล ดังนี้

Unsupervised LSTM Autoencoder

1. Reconstruction error ยังไม่สามารถแยกแยะพีคได้ชัดเจนในบางบริบท
2. ลักษณะของพีคมีความหลากหลายสูง (ขนาด รูปร่าง ตำแหน่ง)
3. ข้อมูลมีความไม่สมดุลสูง (peak มีสัดส่วนน้อยมาก) ทำให้โมเดลมี bias ไปทาง non-peak

Supervised LSTM

1. จำนวน label = 1 มีน้อยมากเมื่อเทียบกับ label = 0 ทำให้โมเดลเรียนรู้เป็น 0 ตลอดเวลา
2. พีคเกิดขึ้นแบบเฉียบพลันและสั้นเกินไป ทำให้ LSTM มองข้ามได้ง่าย
3. แม้ปรับ loss function และ threshold แล้ว ค่า F1-score ยังคงต่ำมาก (≈ 0.002–0.004)

Isolation Forest

1. ค่า Recall เท่ากับ 1.0000 แต่ Precision ต่ำมาก (0.0005) เพราะโมเดลทำนายทุกตำแหน่งเป็นพีค
2. โมเดลไม่สามารถแยกจุดที่เป็นพีคจริงออกจากความผิดปกติอื่น ๆ ได้
3. ไม่สามารถจับลักษณะ pattern แบบต่อเนื่องในเวลาได้ เนื่องจากเป็นโมเดล anomaly detection แบบจุดเดี่ยว (point anomaly)

K-means Clustering

1. โมเดลไม่ได้พิจารณาโครงสร้างล้อของรถไฟ เช่น 4 ล้อต่อคัน หรือ 6 ล้อในหัวรถจักร
2. แบ่งกลุ่มผิดพลาด เช่น ล้อจากคันเดียวกันถูกแบ่งแยก (under-cluster) หรือรวมคันต่างกันเข้าด้วยกัน (over-cluster)

DBSCAN

1. ค่าความผิดพลาดสูงในหลายกรณี โดยเฉพาะช่องสัญญาณ ch3
2. หากช่วงเวลาระหว่างล้อในคันเดียวกันไม่สม่ำเสมอ หรือเกินค่า eps ที่ตั้งไว้ จะถูกแบ่งออกจากกันทันที แม้จะอยู่ในคันเดียวกัน

4.3 ประโยชน์ของโครงงาน

แม้ผลการทดลองของโมเดล AIจะยังไม่สามารถนำไปใช้แทนผู้เชี่ยวชาญในทางปฏิบัติได้โดยตรง แต่โครงงานนี้ได้สร้างประโยชน์ในหลายด้าน ได้แก่

1. เป็นกรณีศึกษาเชิงวิศวกรรมที่แสดงให้เห็นขีดจำกัดของการประยุกต์ใช้ AI กับปัญหาที่มีโครงสร้างเฉพาะ
2. สร้างชุดข้อมูล ตัวอย่างสคริปต์ และกระบวนการวิเคราะห์ที่สามารถนำไปใช้ต่อยอดในงานวิจัยหรือโครงการในอนาคต
3. พัฒนา pipeline สำหรับการประมวลผลสัญญาณจริงจากเซนเซอร์ ซึ่งสามารถนำไปประยุกต์ใช้ในงานตรวจสอบความปลอดภัยของรางรถไฟ
4. สร้างแนวคิดการใช้ AI อย่างมีเหตุผลในระบบวิศวกรรม ไม่ใช่เพียงการตามเทคโนโลยี แต่เน้นผลลัพธ์ที่สอดคล้องกับความเป็นจริงและความปลอดภัย

โครงงานนี้จึงนับว่าเป็นรากฐานที่สำคัญในการประเมินความเป็นไปได้ และขีดจำกัดของ AI ในงานระบบราง ซึ่งสามารถใช้ต่อยอดได้ในอนาคตทั้งในเชิงวิชาการและเชิงอุตสาหกรรม นอกจากนี้ ผลลัพธ์จากการวิเคราะห์แบบ manual ยังสามารถนำมาใช้เป็น indirect product เพื่อการวิเคราะห์เชิงสถิติและวางแผนการเดินรถ เช่น การหาความน่าจะเป็นของจำนวนคันรถในแต่ละขบวน การวิเคราะห์ปริมาณการขนส่งสินค้าขาเข้า(แหลมฉบังไปกทม.)เปรียบเทียบกับขาออก(กทม.ไปแหลมฉบัง) หรือการตรวจจับขบวนที่มีพฤติกรรมผิดปกติ ซึ่งช่วยเพิ่มมูลค่าของข้อมูลที่ได้จากระบบวิเคราะห์แมนนวล แม้จะไม่ได้ใช้ AI โดยตรง

4.3 Indirect products

ในการศึกษานี้ ได้มีการนำข้อมูลที่ได้จากการตรวจจับรถไฟในแต่ละวันมาวิเคราะห์ในเชิงสถิติ เพื่อหาแนวโน้มของจำนวนขบวนรถไฟที่วิ่งผ่านในแต่ละเดือน โดยทำการเปรียบเทียบจำนวนรอบของรถไฟในแต่ละเดือน เพื่อวิเคราะห์ถึงความแตกต่างและหาปัจจัยที่อาจส่งผลต่อความถี่ของการเดินรถ เช่น ปัจจัยด้านฤดูกาล ความต้องการขนส่ง หรือเหตุขัดข้องทางเทคนิคต่าง ๆ\n\nนอกจากนี้ ได้มีการวิเคราะห์ความน่าจะเป็นของจำนวนคันรถในแต่ละขบวน โดยพิจารณาว่าขบวนรถไฟหนึ่งขบวนมีแนวโน้มที่จะประกอบด้วยคันรถจำนวนเท่าใด และเหตุใดจึงเกิดความแตกต่างดังกล่าว ทั้งนี้ยังมีการตรวจสอบว่ามีกรณีใดบ้างที่ขบวนรถไฟมีการบรรทุกน้ำหนักเกินกว่าค่ากำหนด โดยวิเคราะห์จากข้อมูลในช่วงเวลา 1 วันที่ผ่านมา เพื่อประเมินความถี่ของเหตุการณ์ดังกล่าวและวางแนวทางการตรวจสอบหรือแจ้งเตือนในอนาคต

4.4 ข้อสรุปเชิงวิศวกรรม

AIยังไม่สามารถทดแทนตรรกะเชิงวิศวกรรมของมนุษย์ได้โดยสมบูรณ์ โดยเฉพาะงานที่ต้องการความแม่นยำในโครงสร้าง (เช่น แบ่งคันรถ)แต่สามารถนำมาใช้สนับสนุนการตัดสินใจได้อย่างมีประสิทธิภาพในงาน routine หรืองานที่มีข้อมูลจำนวนมาก

5.สรุปผลการวิจัย

งานวิจัยนี้มีเป้าหมายเพื่อศึกษาความเป็นไปได้ของการนำโมเดลปัญญาประดิษฐ์ (AI) มาใช้ในการตรวจจับพีคล้อรถไฟและวิเคราะห์จำนวนคันรถไฟในหนึ่งขบวน โดยมีการทดลองใช้โมเดลหลายรูปแบบ ได้แก่ LSTM ทั้งแบบ Supervised และ Unsupervised, Isolation Forest รวมถึง Clustering Algorithms อย่าง K-Means และ DBSCAN พร้อมทั้งเปรียบเทียบกับวิธีการแบบดั้งเดิมที่ใช้ Signal Processing โดยผู้เชี่ยวชาญ

5.1 การตรวจจับพีคล้อรถไฟ

1. LSTM (Supervised) ให้ค่า accuracy สูง แต่ค่า F1-score ต่ำมาก (< 0.01) เนื่องจากข้อมูลไม่สมดุล ทำให้โมเดลไม่สามารถตรวจจับพีคได้แม่นยำ และเกิด bias ไปทาง non-peak
2. LSTM Autoencoder (Unsupervised) สามารถจับพีคได้บางส่วนโดยไม่ต้องใช้ label โดยเฉพาะพีคขนาดใหญ่ มี F1-score สูงกว่าแบบ Supervised และ False Positive ต่ำ อย่างไรก็ตาม ยังมีข้อจำกัดในการตรวจจับพีคที่มีลักษณะเฉพาะหรือขนาดเล็ก
3. Isolation Forest ซึ่งเป็นโมเดลประเภท anomaly detection ไม่สามารถประยุกต์ใช้กับลักษณะพีคของล้อรถไฟได้อย่างแม่นยำ เนื่องจากพีคไม่ได้เป็น anomaly ตามนิยามของโมเดล ทำให้ผลลัพธ์ไม่สอดคล้องกับความเป็นจริง โดยเฉพาะเมื่อใช้กับข้อมูลขนาดใหญ่และหลากหลาย

สรุป: วิธีที่ให้ผลแม่นยำที่สุดในการตรวจจับพีคยังคงเป็น Signal Processing ที่ออกแบบโดยผู้เชี่ยวชาญ เนื่องจากสามารถตีความบริบทของล้อรถไฟได้ดีกว่า AI และสามารถจัดการกับความหลากหลายของสัญญาณได้ดีกว่า

5.2 การวิเคราะห์จำนวนคันรถไฟ

K-Means และ DBSCAN ไม่สามารถเข้าใจโครงสร้างของรถไฟจริง เช่น ล้อ 4 ล้อต่อคัน หรือการจัดเรียงล้อแบบคู่ ทำให้ไม่สามารถแบ่งคันรถได้อย่างแม่นยำ ผลลัพธ์ที่ได้แม้จะใกล้เคียงในเชิงตัวเลข แต่ขาดความถูกต้องเชิงโครงสร้าง

สรุป: การแบ่งคันรถด้วย AI ยังไม่สามารถแทนที่การนับด้วยมนุษย์ได้ เนื่องจากโมเดลไม่เข้าใจตรรกะของโครงสร้างรถไฟ และไม่ได้ถูกออกแบบให้สะท้อนธรรมชาติของปัญหา

5.3 แนวทางการปรับปรุง

จากข้อผิดพลาดที่เกิดขึ้น
สามารถเสนอแนวทางการปรับปรุงได้ดังนี้
ด้านข้อมูล (Data)

1. ทำ preprocessing เช่น ขยายพีค (amplify), smooth noise
2. เพิ่ม label หรือเปลี่ยนเป็นการมาร์กช่วงเวลา (Group label) รอบพีคแทนรายจุด
3. เพิ่มฟีเจอร์เชิงเวลาจากสัญญาณ เช่น rolling std, amplitude, slope

ด้านโมเดล (Model)

1. เปลี่ยนไปใช้โมเดลที่เหมาะกับลักษณะของข้อมูล เช่น Hybrid CNN-LSTM
2. ใช้ semi-supervised หรือ weakly-supervised แทนการเรียนรู้แบบ fully-supervised
3. ใช้ post-processing เพื่อลด false positive เช่น ตรวจสอบความต่อเนื่องของพีค หรือระยะเวลาระหว่างพีค

ด้านการออกแบบการทดลอง

1. วาง baseline เปรียบเทียบกับการวิเคราะห์แบบ Manual อย่างชัดเจน
2. สร้างชุด test ที่หลากหลาย ครอบคลุม noise และลักษณะพีคต่าง ๆ
3. พิจารณาใช้ AI ในลักษณะเป็นเครื่องมือสนับสนุน (supportive tool) ควบคู่กับตรรกะจากผู้เชี่ยวชาญ

อ้างอิง

1. Geng, G., Zhang, H., Zhu, J., & Zhong, C. (1999). Application of AI in railway freight loading. Expert Systems with Applications, 16(1), 63-72.
2. Patel, D. S., & Panchal, M. H. (2018). A systematic review on K-means clustering techniques.
3. Kongpachith, E. (2023). Estimating the load weight of freight trains using machine learning. Degree project in technology, second cycle, KTH Royal Institute of Technology, Stockholm, Sweden.