



การศึกษาเปรียบเทียบการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจด้วย
การเรียนรู้เชิงลึกสำหรับข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนโดยใช้ข้อมูลจำลอง

A Simulation Study to Compare ECG Classification Using
Deep Learning for ECG Artifacts

ณัฐชัย บวรมงคลศักดิ์* และ อัครินทร์ ไพบูลย์พานิช

Nattachai Bovornmongkolsak* and Akarin Phaibulpanich

ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย เขตปทุมวัน กรุงเทพฯ ประเทศไทย

Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Pathumwan, Bangkok, Thailand

*Corresponding Author E-mail: nntch.b@gmail.com

บทคัดย่อ

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบตัวแบบในการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจโดยใช้ตัวแบบการเรียนรู้เชิงลึก 3 ตัวแบบได้แก่ (1) โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น (2) โครงข่ายคอนโวลูชันเต็มรูป และ (3) โครงข่ายแบบเรซินิคัลหรือเรสเน็ต ชุดข้อมูลคลื่นไฟฟ้าหัวใจถูกจำลองขึ้นโดยใช้โปรแกรม Neurokit2 และทำการเพิ่มสิ่งแปลกปนลงในข้อมูล 4 แบบได้แก่ Wandering baseline, Muscle tremor, AC interference และ Motion artifacts และเปรียบเทียบประสิทธิภาพของแต่ละตัวแบบในการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจปกติกับคลื่นไฟฟ้าหัวใจรบกวนแล้วเปรียบเทียบผลโดยใช้วิธีครอสวาไลเดชัน แบ่งข้อมูลเป็น 10 ส่วนแล้วพิจารณา ค่าความถูกต้อง ค่าความแม่นยำ และค่าความครบถ้วน จากผลการทดสอบพบว่า การจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจโดยใช้โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น จะมีประสิทธิภาพลดลงเมื่อข้อมูลมีสิ่งแปลกปน โดยสิ่งแปลกปนแบบ Motion artifacts จะลดประสิทธิภาพของตัวแบบมากที่สุด โดยค่าเฉลี่ยความถูกต้อง (Accuracy) เท่ากับ 0.609 และ 0.601 ลดลงจาก 0.690 มีค่าเฉลี่ยความแม่นยำ (Precision) เท่ากับ 0.563 ลดลงจาก 0.620 และ ค่าเฉลี่ยความครบถ้วน (Recall) เท่ากับ 0.979 และ 0.918 ลดลงจาก 0.990 สำหรับตัวแบบโครงข่ายคอนโวลูชันเต็มรูป เมื่อข้อมูลมีสิ่งแปลกปนแบบ Motion artifacts จะมีค่าเฉลี่ยความถูกต้องเท่ากับ 0.941 และ 0.972 ลดลงจาก 0.999 มีค่าเฉลี่ยความแม่นยำเท่ากับ 0.884 และ 0.978 ลดลงจาก 0.999 และมีค่าเฉลี่ยความครบถ้วนเท่ากับ 0.898 และ 0.972 ลดลงจาก 0.999 สำหรับตัวแบบโครงข่ายแบบเรซินิคัล เมื่อข้อมูลมีสิ่งแปลกปนแบบ Motion artifacts จะมีค่าเฉลี่ยความถูกต้องเท่ากับ 1.00 และ 0.999 ค่าเฉลี่ยความแม่นยำเท่ากับ 1.00 และ 0.999 และค่าเฉลี่ย



ความครบถ้วนเท่ากับ 1.00 และ 1.00 โดยพบว่าค่าประสิทธิภาพไม่ต่างจากข้อมูลคลื่นไฟฟ้าหัวใจที่ไม่มีสิ่งแปลกปนอย่างมีนัยสำคัญ แต่มีค่าเบี่ยงเบนมาตรฐานสูงในบางกรณี

คำสำคัญ: การจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจ ข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปน การเรียนรู้เชิงลึก ข้อมูลอนุกรมเวลา

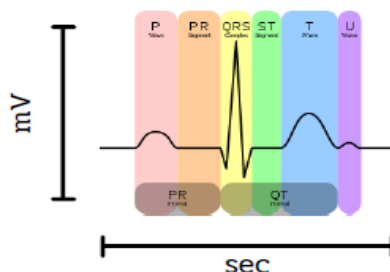
Abstract

The objective of this research is to study and compare ECG classification models using three deep learning models: (1) multilayer perceptrons, (2) fully convolutional networks, and (3) residual networks. The ECG datasets were simulated using Neurokit2 program, with 4 types of artifacts including wandering baseline, muscle tremor, AC interference, and motion artifact. The comparison of the performance of each model was conducted using 10-fold cross validation. The obtained accuracy, precision, and recall were investigated. The results showed that the multilayer perceptrons were less effective when the datasets contained artifacts. The motion artifact was the most effective the mean accuracy was 0.609 and 0.601 decreased from 0.690, the mean precision was 0.563 decreased from 0.620 and mean recall was 0.979 and 0.918 decreased from 0.990. The fully convolutional networks when the data contained motion artifacts, the mean accuracy was 0.941 and 0.972 decreased from 0.999, the mean precision was 0.884 and 0.978 decreased from 0.999 and the mean of recall was 0.898 and 0.972 decreased from 0.999. The residual networks when the data contained motion artifacts, the mean accuracy was 1.00 and 0.99 the mean precision was 1.00 and 0.99 and the mean of recall was 1.00 and 1.00. The performance was not significantly different with cleaned ECG dataset, but a high standard deviation was detected in some cases.

Keywords: ECG classification, ECG artifacts, Deep Learning, Time Series

1. บทนำ

ข้อมูลคลื่นไฟฟ้าหัวใจ (ECG หรือ EKG หรือ Electrocardiogram) เป็นข้อมูลอนุกรมเวลา (Time Series) ที่แสดงให้เห็นถึงการนำไฟฟ้าบริเวณต่าง ๆ ของกล้ามเนื้อหัวใจ ซึ่งจะสอดคล้องกับการบีบและคลายตัวของกล้ามเนื้อหัวใจ ดังนั้นข้อมูลคลื่นไฟฟ้าหัวใจจึงเป็นข้อมูลที่มีความสำคัญมากต่อการวินิจฉัยโรคหัวใจ โดยแพทย์จะวินิจฉัยจากรูปร่างของคลื่นไฟฟ้าหัวใจแต่ละส่วนใน 1 บีท หรือ รอบหัวใจเต้น (Beat หรือ Cardiac cycle) เนื่องจากรูปร่างของคลื่นไฟฟ้าหัวใจ 1 บีท ประกอบด้วยจังหวะการบีบและคลายตัวของหัวใจ ตำแหน่งต่าง ๆ เช่น P wave แสดงการบีบตัวของกล้ามเนื้อหัวใจห้องบน QRS จะแสดงการบีบตัวของกล้ามเนื้อหัวใจห้องล่าง เป็นต้น โดยองค์ประกอบเหล่านี้มีความสำคัญในการวินิจฉัยโรคต่าง ๆ ที่เกี่ยวข้อง เช่น ถ้าหากผู้ป่วยเป็นโรคกล้ามเนื้อหัวใจตาย (Myocardial infarction) ลักษณะของข้อมูลคลื่นไฟฟ้าหัวใจในส่วน ST Segment จะมีการยกตัวขึ้นสูง (ST-Elevation) เป็นต้น



รูปที่ 1 ภาพตัวอย่างแสดงองค์ประกอบแต่ละส่วนของข้อมูลคลื่นไฟฟ้าหัวใจ 1 บีท
แกน x แสดงข้อมูลเวลาหน่วยเป็นวินาที (sec) แกน y แสดงค่าความต่างศักย์ไฟฟ้า มีหน่วยเป็น มิลลิโวลต์ (mV)
(ที่มา: “Electrocardiography” 2021)

การจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจ คือ การใช้ข้อมูลคลื่นไฟฟ้าหัวใจเพื่อจำแนกผู้ป่วยที่มีปัญหาการทำงานของหัวใจ เช่น การจำแนกผู้ป่วยโรคกล้ามเนื้อหัวใจขาดเลือด (Myocardial Ischemia หรือ MI) การจำแนกผู้ป่วยที่มีภาวะหัวใจเต้นเร็วผิดปกติ (Supraventricular tachycardia หรือ SVT) เป็นต้น มีงานวิจัยจำนวนมากนำเสนอวิธีการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจ แต่ยังไม่มีงานวิจัยใดที่สามารถนำผลลัพธ์ไปใช้ได้จริงทางการแพทย์ เนื่องจากข้อมูลคลื่นไฟฟ้าหัวใจมีสิ่งแปลกปน (ECG artifacts) อัลกอริทึมจึงจำแนกประเภทข้อมูลผิดพลาด โดยข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปน คือ คลื่นไฟฟ้าหัวใจที่ถูกรบกวน ทำให้ภาพของคลื่นไฟฟ้าหัวใจที่ปรากฏมีองค์ประกอบที่ไม่ชัดเจน โดยสิ่งแปลกปนที่พบบ่อยครั้งแบ่งเป็น 4 ประเภท (เหมวรรณ และ โชติรัตน์, 2557) ได้แก่ 1. Wandering baseline สาเหตุโดยทั่วไปมาจากการเคลื่อนที่ของร่างกาย หรือ ขั้วไฟฟ้า ลักษณะของเส้นมาตรฐานแกว่งขึ้นลงช้า ๆ 2. Muscle tremor สาเหตุโดยทั่วไปมาจากการเกร็งของกล้ามเนื้อ โดยลักษณะจะมีการเปลี่ยนแปลงอย่างรวดเร็วในเวลาสั้นๆ และเส้นมาตรฐานไม่คงที่ 3. AC interference สาเหตุโดยทั่วไปมาจากการติดขั้วไฟฟ้าไม่ดี การรั่วของกระแสไฟฟ้า โดยลักษณะจะมีการเปลี่ยนแปลงอย่างรวดเร็วเป็นเวลานาน 4. Motion artifact สาเหตุโดยทั่วไปมาจากร่างกายสั่น การเคลื่อนที่ของร่างกาย หรือการหายใจ โดยลักษณะจะมีการเปลี่ยนแปลงเส้นมาตรฐานและค่าสูงสุดของแอมพลิจูด

Apandi et al.(2020) ได้ศึกษาผลกระทบของสิ่งแปลกปนต่อประสิทธิภาพในการค้นพบตำแหน่งองค์ประกอบของข้อมูลคลื่นไฟฟ้าหัวใจแต่ละวิธี โดยใช้ฐานข้อมูล MIT-BIH Arrhythmia Database ในชุดข้อมูลประกอบด้วย บันทึกข้อมูลคลื่นไฟฟ้าหัวใจของ 48 คน แต่ละคนบันทึกข้อมูล 30 นาที มีความถี่สัญญาณที่ 360 เฮิรตซ์ มีข้อมูล 109,505 รอบการเต้นหัวใจ (beat) และทำการเพิ่มสิ่งแปลกปนให้กับข้อมูลคลื่นไฟฟ้าหัวใจ โดยสิ่งแปลกปนที่เพิ่มได้แก่ baseline wander, muscle artifacts และ electrode motion และมีการทดสอบระดับของสิ่งแปลกปนหลายระดับ ในงานวิจัยได้ทำการทดสอบ 3 อัลกอริทึมในการระบุตำแหน่ง QRS ของคลื่นไฟฟ้าหัวใจ ได้แก่ Pan Tompkins, WQRS และ Hamilton และวัดประสิทธิภาพด้วยค่าความไว (Sensitivity) ค่าพยากรณ์ผลบวก (Positive predictivity)



จากการทดสอบพบว่าสิ่งแปลกปนมีผลกระทบต่อประสิทธิภาพในการระบุตำแหน่ง QRS ในทุกอัลกอริทึม โดยเฉพาะสิ่งแปลกปนชนิด electrode motion ที่ทำให้ประสิทธิภาพของอัลกอริทึมลดลงมากที่สุด

อย่างไรก็ตาม ยังไม่มีงานวิจัยที่ศึกษาเปรียบเทียบประสิทธิภาพของตัวแบบที่ใช้เทคนิคการเรียนรู้เชิงลึกในการจำแนกข้อมูลคลื่นไฟฟ้าหัวใจ โดยการเรียนรู้เชิงลึกจะประกอบด้วยการประมวลผลหลายชั้นเพื่อเรียนรู้ตัวแทนข้อมูลหรือคุณลักษณะพิเศษจากข้อมูล ต่างจากวิธีการจำแนกข้อมูลที่ใช้อัลกอริทึมต่าง ๆ เพื่อหาคุณลักษณะพิเศษ ซึ่งยังมีปัญหาเนื่องจากยังไม่มีอัลกอริทึมใด สามารถกำจัดสิ่งแปลกปนทุกประเภทออกหมดได้ ผู้วิจัยจึงสนใจศึกษาตัวแบบการเรียนรู้เชิงลึกสำหรับการจำแนกประเภทข้อมูลอนุกรมเวลา เนื่องจากข้อมูลคลื่นไฟฟ้าหัวใจมีลักษณะเป็นข้อมูลอนุกรมเวลา สำหรับเทคนิคการเรียนรู้เชิงลึกเพื่อจำแนกข้อมูลอนุกรมเวลามีหลายแนวทาง (Fawaz et. al., 2019) โดยมี 2 แนวทางหลักได้แก่ (1) Generative Models ซึ่งเป้าหมายหลักของแนวทางนี้จะเป็นการสร้างตัวแปรเพื่อนำไปใช้ต่อ ผู้วิจัยจึงไม่นำแนวทางนี้มาศึกษา และ (2) Discriminative Models ซึ่งประกอบด้วยวิธี Feature Engineering และ วิธี End-to-End โดยงานวิจัยนี้จะศึกษาตัวแบบประเภท End-to-End เนื่องจากต้องการศึกษาเฉพาะผลกระทบของสิ่งแปลกปนที่มีต่อตัวแบบเท่านั้น ไม่ได้มุ่งเน้นไปที่การปรับแต่งโครงข่ายโดยใช้ความรู้เฉพาะทางเพื่อให้ตัวแบบมีประสิทธิภาพสูง โดยในปี 2017 Wang, Yan และ Oates (Wang, Yan, & Oates, 2017) ได้นำเสนอตัวแบบ ที่เรียบง่ายและไม่ต้องการเตรียมข้อมูลหรือการหาคุณลักษณะพิเศษ 3 ตัวแบบ ได้แก่ (1) โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (Multilayer Perceptrons: MLPs) (2) โครงข่ายคอนโวลูชันเต็มรูป (Fully Convolutional Networks: FCNs) และ (3) โครงข่ายแบบเรสซิดิวลหรือเรสเนต (Residual Network: ResNet) โดยงานวิจัยได้ทดสอบประสิทธิภาพของตัวแบบทั้งสามโดยใช้ชุดข้อมูลทั้งหมด 44 ชุดจากคลังข้อมูลด้านอนุกรมเวลา UCR Time Series Dataset (Chen et al., 2015) และพบว่าการใช้เทคนิคการเรียนรู้เชิงลึกมีประสิทธิภาพใกล้เคียงตัวแบบ COTE ซึ่งใช้ทรัพยากรในการคำนวณจำนวนมากและมีความซับซ้อนจนไม่เหมาะกับการนำไปใช้งานจริง การใช้เทคนิคการเรียนรู้เชิงลึกจึงเป็นทางเลือกที่น่าสนใจในการจำแนกข้อมูลอนุกรมเวลา

จากที่กล่าวมาแสดงให้เห็นว่าการจำแนกข้อมูลคลื่นไฟฟ้าหัวใจที่มีการแปลกปนมีความสำคัญ และส่งผลกระทบต่อการใช้งานจริงทางการแพทย์ และยังคงเป็นปัญหาที่ไม่มีงานวิจัยใด สามารถตรวจจับสิ่งผิดปกติในข้อมูลคลื่นไฟฟ้าหัวใจที่สิ่งแปลกปนได้อย่างถูกต้อง แต่ยังไม่มียงานวิจัยที่ศึกษาเปรียบเทียบประสิทธิภาพของตัวแบบในการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนโดยใช้เทคนิคการเรียนรู้เชิงลึก งานวิจัยนี้ผู้วิจัยจึงสนใจที่จะทำการศึกษา อย่างไรก็ตามการจะนำตัวแบบที่ใช้เทคนิคการเรียนรู้เชิงลึกมาใช้ จำเป็นต้องมีข้อมูลจำนวนมาก ซึ่งการจะรวบรวมข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนแต่ละชนิด สามารถทำได้ยาก งานวิจัยนี้จึงจำลองข้อมูลเพื่อศึกษาโดยใช้โปรแกรม Neurokit2 (Makowski et al., 2020) เพื่อจำลองข้อมูลมาใช้ในการเปรียบเทียบประสิทธิภาพของตัวแบบในงานวิจัยนี้



2. วัตถุประสงค์

เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบในการจำแนกประเภทข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนโดยใช้เทคนิคการเรียนรู้เชิงลึก

3. วิธีดำเนินงานวิจัย

3.1 กลุ่มตัวอย่าง

กลุ่มตัวอย่างทั้งหมดมีจำนวน 10,000 ตัวอย่าง แบ่งเป็นข้อมูลคลื่นไฟฟ้าหัวใจปกติ 5,000 ตัวอย่าง และคลื่นไฟฟ้าหัวใจกรณีกล้ามเนื้อหัวใจขาดเลือด 5,000 ตัวอย่าง มีจำนวนจุดเวลา (time length) เท่ากับ 1080 จุด (3 วินาที โดยมีความถี่สัญญาณ 360 ตัวอย่างต่อวินาที) และอัตราการเต้นของหัวใจ จะสุ่มจากการแจกแจงต่อเนื่องแบบยูนิฟอร์ม $U(60, 100)$

3.2 การจำลองข้อมูลคลื่นไฟฟ้าหัวใจ

3.2.1 ข้อมูลคลื่นไฟฟ้าหัวใจปกติ

จำลองข้อมูลโดยใช้โปรแกรม Neurokit2 ตามข้อกำหนดในข้อ 3.1

3.2.2 ข้อมูลคลื่นไฟฟ้าหัวใจกรณีกล้ามเนื้อหัวใจขาดเลือด

1) จำลองข้อมูลโดยใช้โปรแกรม Neurokit ตามข้อกำหนดใน 3.1

2) ปรับแต่งข้อมูล (data augmentation) ในส่วน ST-Segments เพิ่มเติม โดยจะเพิ่มค่า 0.5 เท่าของค่าเดิมและลดลงตามลำดับจนไม่มีการเพิ่มค่าที่ตำแหน่ง T เพื่อให้สอดคล้องกับลักษณะการเกิด ST-Elevation

3.3 การเพิ่มสิ่งแปลกปนในข้อมูลคลื่นไฟฟ้าหัวใจ

3.3.1 สิ่งแปลกปนแบบ Wandering baseline

ในการเพิ่มสิ่งแปลกปนชนิดนี้ จะสร้างตัวแปรโดยใช้สมการคลื่นไซน์ (sine wave) และสุ่มความถี่เชิงมุม (angular velocity) จากการแจกแจงแบบยูนิฟอร์ม $U\{0.25, 0.50, 0.75\}$ รอบต่อวินาที และกำหนดขนาด Amplitude โดยคูณด้วย 0.50 หรือ 1.00

3.3.2 สิ่งแปลกปนแบบ Muscle tremor

ในการเพิ่มสิ่งแปลกปนชนิดนี้ จะสุ่มตำแหน่งในการเพิ่มสิ่งแปลกปนจากการแจกแจงแบบยูนิฟอร์ม $U\{0, 1, \dots, 720\}$ โดยจะสร้างตัวแปรโดยใช้ตัวแปรสุ่มจากฟังก์ชันการแจกแจงปกติมาตรฐาน ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐาน จะสุ่มจากการแจกแจงแบบยูนิฟอร์ม $U\{0.10, 0.15\}$

3.3.3 สิ่งแปลกปนแบบ AC interference

ในการเพิ่มสิ่งแปลกปนชนิดนี้ จะสร้างตัวแปรโดยใช้ตัวแปรสุ่มจากฟังก์ชันการแจกแจงปกติมาตรฐาน ซึ่งมีค่าเฉลี่ยเท่ากับ 0 และค่าเบี่ยงเบนมาตรฐาน จะสุ่มจากการแจกแจงแบบยูนิฟอร์ม $U\{0.10, 0.15\}$



3.3.4 สิ่งแปลกปนแบบ Motion artifacts

ในการเพิ่มสิ่งแปลกปนชนิดนี้ จะสุ่มตำแหน่งในการเพิ่มสิ่งแปลกปนจากการแจกแจงไม่ต่อเนื่องแบบยูนิฟอร์ม $U\{0,1,...,720\}$ และสร้างตัวแปรโดยใช้สมการคลื่นไซน์ (sine wave) และสุ่มความถี่เชิงมุม จากการแจกแจงแบบยูนิฟอร์ม $U\{0.25, 0.50, 0.75\}$ รอบต่อวินาที และกำหนดขนาด Amplitude โดยคูณด้วย 0.50 หรือ 1.00

3.4 การจำแนกประเภทข้อมูลอนุกรมเวลา

สำหรับโครงสร้างของโครงข่ายประสาทเทียมแต่ละตัวแบบที่ใช้ในการทดสอบ จะใช้โครงสร้างจากงานวิจัยของ Wang et al. ที่กล่าวข้างต้น โดยแต่ละตัวแบบมีโครงสร้างดังนี้

3.4.1 โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น

ประกอบด้วยโครงข่ายประสาทเทียมทั้งหมด 4 ชั้น ในชั้นที่ 1 ถึง ชั้นที่ 3 จะมีจำนวนเพอร์เซ็ปตรอน 500 หน่วย และชั้นที่ 4 คือชั้นข้อมูลขาออก แต่ละชั้นเชื่อมโยกันโดยใช้ฟังก์ชันเรกติไฟต์เชิงเส้น (ReLU) เป็นฟังก์ชันกระตุ้น มีการใช้เทคนิค การหยุดเรียนรู้กลางคัน (Drop-out) กำหนด dropout rate เท่ากับ 0.1, 0.2, 0.2 และ 0.3 ตามลำดับ

3.4.2 โครงข่ายคอนโวลูชันเต็มรูป

ประกอบด้วย ชั้นคอนโวลูชัน 3 ชั้น ชั้นการรวม 1 ชั้น และ ชั้นข้อมูลขาออก ออก โดยจำนวนตัวกรองแต่ละชั้นเท่ากับ 128, 128 และ 256 ตามลำดับ แต่ละชั้นเชื่อมโยกันโดยใช้ฟังก์ชันเรกติไฟต์เชิงเส้น (ReLU) เป็นฟังก์ชันกระตุ้น มีการใช้เทคนิค Batch Normalization

3.4.3 โครงข่ายแบบเรซิดิวลหรือเรสเนท

ประกอบด้วย เรซิดิวล บล็อก 3 บล็อกโดยแต่ละบล็อกมีรายละเอียดดังนี้ แต่ละบล็อกประกอบด้วย คอนโวลูชันชนิด 1D ทั้งหมด 3 เลเยอร์ โดยจำนวนตัวกรองของทุกเลเยอร์ในแต่ละบล็อกเท่ากับ 64, 128 และ 128 ตามลำดับ มีการเชื่อมโยกันโดยใช้ฟังก์ชันเรกติไฟต์เชิงเส้น (ReLU) เป็นฟังก์ชันกระตุ้น และมีการใช้เทคนิค Batch Normalization

3.5 การศึกษาเปรียบเทียบตัวแบบ

ในการเปรียบเทียบประสิทธิภาพของตัวแบบ เพื่อเป็นการตรวจสอบการเกิด Overfitting จะใช้วิธีการสวาลิเดชันโดยแบ่งชุดข้อมูลเป็น 10 ส่วน และพิจารณาค่าดังนี้

3.5.1 ค่าความถูกต้อง (Accuracy)

สำหรับการเปรียบเทียบค่าความถูกต้อง ถ้ามีค่าเข้าใกล้ 1.00 จะมีประสิทธิภาพดีที่สุด โดยสามารถคำนวณได้ดังนี้

$$\text{ค่าความถูกต้อง} = \frac{\text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TP)} + \text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TN)}}{\text{จำนวนข้อมูลทั้งหมด}}$$



3.5.2 ค่าความแม่นยำ (Precision)

สำหรับการเปรียบเทียบค่าความแม่นยำ ถ้ามีค่าเข้าใกล้ 1.00 จะมีประสิทธิภาพดีที่สุด โดยสามารถคำนวณได้ดังนี้

$$\text{ค่าความแม่นยำ} = \frac{\text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TP)}}{\text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TP) + จำนวนที่ทำนายคลาสนั้นแต่ผลลัพธ์คือคลาสนั้น (FP)}}$$

3.5.3 ค่าความครบถ้วน (Recall)

สำหรับการเปรียบเทียบค่าความครบถ้วน ถ้ามีค่าเข้าใกล้ 1.00 จะมีประสิทธิภาพดีที่สุด โดยสามารถคำนวณได้ดังนี้

$$\text{ค่าความครบถ้วน} = \frac{\text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TP)}}{\text{จำนวนที่ทำนายคลาสนั้นและผลลัพธ์คือคลาสนั้น (TP) + จำนวนที่ทำนายคลาสนั้นแต่ผลลัพธ์คือคลาสนั้น (FN)}}$$

3.6 สรุปผลวิจัยในแต่ละกรณี

4. ผลการวิจัยและวิเคราะห์ผล

4.1 การเปรียบเทียบค่าความถูกต้อง (Accuracy)

ตารางที่ 1 ตารางการเปรียบเทียบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความถูกต้อง

Artifacts	Amplitude*	Accuracy		
		MLPs	FCNs	ResNet
		Avg. (std.)	Avg. (std.)	Avg. (std.)
None	-	0.690 (0.031)	0.999 (0.002)	0.999 (0.001)
Wandering Baseline	0.50	0.735 (0.052)	0.944 (0.149)	0.946 (0.155)
	1.00	0.761 (0.057)	0.974 (0.057)	1.000 (0.000)
Muscle tremor	0.15	0.661 (0.039)	0.992 (0.022)	1.000 (0.001)
	0.20	0.665 (0.026)	0.997 (0.004)	0.998 (0.002)
AC interference	0.15	0.657 (0.035)	0.997 (0.002)	0.998 (0.002)
	0.20	0.676 (0.032)	0.992 (0.007)	0.996 (0.003)
Motion artifacts	0.50	0.609 (0.024)	0.941 (0.153)	1.000 (0.000)
	1.00	0.601 (0.021)	0.972 (0.056)	0.999 (0.001)

* Amplitude ในที่นี้หมายถึง ขนาดการแกว่งตัวของสิ่งแปลกปน



จากตารางที่ 1 ผลวิเคราะห์เป็นดังนี้

การจำแนกประเภทข้อมูลด้วยวิธี ResNet และ FCNs มีประสิทธิภาพใกล้เคียงกัน แต่ ResNet มีค่าความถูกต้องสูงกว่าเล็กน้อย และวิธี MLPs มีประสิทธิภาพต่ำสุด

กรณีข้อมูลคลื่นไฟฟ้าหัวใจมีสิ่งแปลกปน จะส่งผลต่อค่าความถูกต้องในตัวแบบ MLPs โดยสิ่งแปลกปนแบบ Motion artifacts ส่งผลให้ค่าความถูกต้องลดลงมากที่สุด โดยค่าเฉลี่ยความถูกต้องเท่ากับ 0.609 และ 0.601 ลดลงจาก 0.690 สำหรับตัวแบบ FCNs และ ResNet พบว่าค่าความถูกต้องไม่ต่างจากข้อมูลคลื่นไฟฟ้าหัวใจที่ไม่มีสิ่งแปลกปนอย่างมีนัยสำคัญ แต่ค่าความถูกต้องของ FCNs ลดลงเล็กน้อยเมื่อมีสิ่งแปลกปนแบบ Wandering baseline ค่าเฉลี่ยความถูกต้องเท่ากับ 0.944 และ 0.974 ลดลงจาก 0.999 และ Motion artifacts ค่าเฉลี่ยความถูกต้องเท่ากับ 0.941 และ 0.972 ลดลงจาก 0.999

4.2 ค่าความแม่นยำ (Precision)

ตารางที่ 2 ตารางการเปรียบเทียบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความแม่นยำ

Artifacts	Amplitude*	Precision		
		MLPs	FCNs	ResNet
		Avg. (std.)	Avg. (std.)	Avg. (std.)
None	-	0.620 (0.027)	0.999 (0.003)	0.999 (0.003)
Wandering Baseline	0.50	0.665 (0.054)	0.952 (0.150)	0.899 (0.316)
	1.00	0.782 (0.098)	0.973 (0.081)	1.000 (0.001)
Muscle tremor	0.15	0.599 (0.032)	0.998 (0.002)	0.999 (0.003)
	0.20	0.601 (0.021)	0.995 (0.007)	0.996 (0.003)
AC interference	0.15	0.596 (0.025)	0.995 (0.004)	0.997 (0.004)
	0.20	0.61 (0.028)	0.988 (0.014)	0.993 (0.006)
Motion artifacts	0.50	0.563 (0.021)	0.884 (0.312)	1.000 (0.000)
	1.00	0.563 (0.021)	0.978 (0.064)	0.999 (0.002)

* Amplitude ในที่นี้หมายถึง ขนาดการแกว่งตัวของสิ่งแปลกปน

การจำแนกประเภทข้อมูลด้วยวิธี ResNet และ FCNs มีประสิทธิภาพใกล้เคียงกัน แต่ ResNet มีค่าความแม่นยำสูงกว่าเล็กน้อย และวิธี MLPs มีประสิทธิภาพต่ำสุด

กรณีข้อมูลคลื่นไฟฟ้าหัวใจมีสิ่งแปลกปน จะส่งผลต่อค่าความแม่นยำในตัวแบบ MLPs โดยสิ่งแปลกปนแบบ Motion artifacts ส่งผลให้ค่าความแม่นยำลดลงมากที่สุด โดยค่าเฉลี่ยความแม่นยำเท่ากับ 0.563 ลดลงจาก 0.620 สำหรับตัวแบบ FCNs และ ResNet พบว่าค่าความแม่นยำมีประสิทธิภาพลดลงและมีค่าเบี่ยงเบนมาตรฐานสูงในบางกรณี FCNs มีค่าเฉลี่ยความแม่นยำเท่ากับ 0.884 และค่าเบี่ยงเบน 0.312 เมื่อมีสิ่งแปลกปนแบบ Motion artifacts และ ResNet มีค่าเฉลี่ยความแม่นยำเท่ากับ 0.899 และค่าเบี่ยงเบน 0.316 เมื่อมีสิ่งแปลกปนแบบ Wandering Baseline



4.3 ค่าความครบถ้วน (Recall)

ตารางที่ 3 ตารางการเปรียบเทียบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าความครบถ้วน

Artifacts	Amplitude*	Recall		
		MLPs	FCNs	ResNet
		Avg. (std.)	Avg. (std.)	Avg. (std.)
None	-	0.990 (0.003)	0.999 (0.002)	1.000 (0.001)
Wandering Baseline	0.50	0.964 (0.016)	0.981 (0.059)	0.891 (0.314)
	1.00	0.777 (0.173)	0.983 (0.054)	1.000 (0.001)
Muscle tremor	0.15	0.989 (0.004)	0.984 (0.044)	1.000 (0.001)
	0.20	0.988 (0.003)	0.999 (0.001)	1.000 (0.001)
AC interference	0.15	0.989 (0.003)	0.999 (0.002)	1.000 (0.001)
	0.20	0.984 (0.005)	0.996 (0.004)	0.999 (0.001)
Motion artifacts	0.50	0.979 (0.011)	0.898 (0.316)	1.000 (0.000)
	1.00	0.918 (0.032)	0.972 (0.089)	1.000 (0.000)

* Amplitude ในที่นี้หมายถึง ขนาดการแกว่งตัวของสิ่งแปลกปน

การจำแนกประเภทข้อมูลด้วยวิธี ResNet และ FCNs มีประสิทธิภาพใกล้เคียงกัน แต่ ResNet มีค่าความครบถ้วนสูงกว่าเล็กน้อย และวิธี MLPs มีประสิทธิภาพต่ำสุด

กรณีข้อมูลคลื่นไฟฟ้าหัวใจมีสิ่งแปลกปน จะส่งผลต่อค่าความครบถ้วนในตัวแบบ MLPs โดยสิ่งแปลกปนแบบ Wandering baseline จะส่งผลให้ค่าความครบถ้วนลดลงมากที่สุดโดยมีค่าเฉลี่ยความครบถ้วนเท่ากับ 0.964 และ 0.777 ลดลงจาก 0.990 สำหรับตัวแบบ FCNs และ ResNet พบว่าค่าความครบถ้วนมีประสิทธิภาพลดลงและมีค่าเบี่ยงเบนมาตรฐานสูงในบางกรณี โดยตัวแบบ FCNs มีค่าเฉลี่ยความครบถ้วนเท่ากับ 0.898 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.316 เมื่อข้อมูลมีสิ่งแปลกปนแบบ Motion artifacts และตัวแบบ ResNet มีค่าเฉลี่ยความครบถ้วนเท่ากับ 0.891 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.314 เมื่อข้อมูลมีสิ่งแปลกปนแบบ Wandering Baseline

5. สรุปผลการศึกษา

โดยภาพรวมแล้ว ประสิทธิภาพในการจำแนกประเภทข้อมูลกรณีข้อมูลคลื่นไฟฟ้าหัวใจ พบว่าตัวแบบ ResNet มีประสิทธิภาพสูงที่สุด รองลงมาเป็นวิธี FCNs โดยมีประสิทธิภาพต่างกันเล็กน้อย และตัวแบบ MLPs มีประสิทธิภาพต่ำสุด

ข้อมูลคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนจะส่งผลต่อประสิทธิภาพในการจำแนกประเภทข้อมูล โดยพิจารณาจากค่าความถูกต้อง ค่าความแม่นยำ และค่าความครบถ้วน โดยสิ่งแปลกปนแบบ Motion artifacts จะส่งผลต่อประสิทธิภาพในการจำแนกประเภทข้อมูลมากที่สุด มีสาเหตุมาจากค่าความแม่นยำที่ลดลง สิ่งแปลกปนแบบ Muscle



tremor และ AC interference จะส่งผลต่อประสิทธิภาพใกล้เคียงกัน สิ่งแปลกปนแบบ Wandering baseline จะส่งผลต่อค่าความครบถ้วนที่ลดลง และมีค่าเบี่ยงเบนมาตรฐานที่สูง แต่ในภาพรวมจะไม่ส่งผลต่อค่าความถูกต้อง

ในอนาคตจะทำการศึกษาเปรียบเทียบประสิทธิภาพของตัวแบบโดยใช้ข้อมูลคลื่นไฟฟ้าหัวใจจากฐานข้อมูล MIT-BIH Arrhythmia ซึ่งเป็นข้อมูลจริง และนอกจากการศึกษาประสิทธิภาพแล้ว จะศึกษาการสร้างคุณลักษณะพิเศษของข้อมูลในแต่ละตัวแบบ เพื่อความเข้าใจในการทำงานของแต่ละตัวแบบ และสามารถพัฒนาโครงสร้างที่มีประสิทธิภาพ สามารถทำความเข้าใจได้ และสามารถนำไปใช้ได้จริง

6. กิตติกรรมประกาศ

ทางผู้วิจัยขอขอบคุณผู้พัฒนาเครื่องมือในการสร้างโครงข่ายประสาทเทียม และการทดสอบตัวแบบ โดยใช้ภาษา python รวมถึงโปรแกรม Neurokit ที่พัฒนาเครื่องมือการจำลองข้อมูลคลื่นไฟฟ้าหัวใจ ทำให้วิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี

7. เอกสารอ้างอิง

- เหมวรรณ ศิริวิทย์, & โชติรัตน์ รัตนานัทธนะ. (2557). การค้นพบกายสัณฐานวิทยาในคลื่นไฟฟ้าหัวใจที่มีสิ่งแปลกปนโดยใช้การทำเหมืองข้อมูลอนุกรมเวลา. สืบค้นจาก <http://cuir.car.chula.ac.th/handle/123456789/45441>
- Apandi, Z. F. M., Ikeura, R., Hayakawa, S., & Tsutsumi, S. (2020). An analysis of the effects of noisy electrocardiogram signal on heartbeat detection performance. *Bioengineering*, 7(2), 53.
- Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., & Batista, G. (2015). The ucr time series classification archive. Retrieved from https://www.cs.ucr.edu/~eamonn/time_series_data/UCR%20Time%20Series%20Classification%20Archive.pdf
- Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4), 917-963.
- Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., ... & Chen, S. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 1-8.
- Wang, Z., Yan, W., & Oates, T. (2017, May). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1578-1585). IEEE.
- Wikipedia. (Online). Electrocardiography. Retrieved February 28, 2021 from <https://en.wikipedia.org/wiki/Electrocardiography>