

Research Article

การเปรียบเทียบตัวแบบจำแนกเสียงเปิดวาล์วของอุปกรณ์ดักไอน้ำ

Comparison of classification model for steam trap valve opening sound

ปณณนุช ดำนงค์^{1*} พิมพภา ตานินพงศ์¹ อนุชา พรหมวังขวา² และ จักรเมธ บุตรกระจำ³

Punyanut Damnong^{1*} Phimphaka Taninpong¹ Anucha Promwungkwa and Jakramate Bootkrajang³

¹ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ 239 ถ.ห้วยแก้ว ต.สุเทพ อ.เมือง เชียงใหม่ 50200 ประเทศไทย

²ภาควิชาวิศวกรรมเครื่องกล คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่ 239 ถ.ห้วยแก้ว ต.สุเทพ อ.เมือง เชียงใหม่ 50200 ประเทศไทย

³ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ 239 ถ.ห้วยแก้ว ต.สุเทพ อ.เมือง เชียงใหม่ 50200 ประเทศไทย

¹Department of Statistics, Faculty of Science, Chiang Mai University, 239 Huay Kaew Road, Suthep, Muang district, Chiang Mai 50200 Thailand

²Department of Mechanical Engineering, Faculty of Engineering, Chiang Mai University, 239 Huay Kaew Road, Suthep, Muang district, Chiang Mai 50200 Thailand

³Department of Computer Science, Faculty of Science, Chiang Mai University, 239 Huay Kaew Road, Suthep, Muang district, Chiang Mai 50200 Thailand

*E-mail: punyanutdamnong@gmail.com

Received: 29/06/2021; Revised: 30/10/2021; Accepted: 16/11/2021

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบการจำแนกเสียงเปิดวาล์วของอุปกรณ์ดักไอน้ำและเปรียบเทียบประสิทธิภาพของตัวแบบที่ใช้วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) และวิธีหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory: LSTM) ทำการบันทึกเสียงการทำงานของอุปกรณ์ดักไอน้ำ และแยกคุณลักษณะของเสียง โดยใช้วิธีอัตราค่าตัดศูนย์ (zero crossing rate) วิธี spectral centroid วิธี spectral rolloff วิธีสัมประสิทธิ์เซปสตรัมบนความถี่แบบเมล (mel-frequency cepstral coefficient) และการแปลงฟูเรียร์ระยะสั้น (short-time fourier transform) นอกจากนี้ยังทำการเปรียบเทียบประสิทธิภาพของตัวแบบโดยใช้ชุดข้อมูล 2 ชุด ได้แก่ ชุดข้อมูลที่ข้อมูลไม่สมดุล กับชุดข้อมูลสมดุลที่ทำการแก้ไขปัญหาคือข้อมูลไม่สมดุลด้วยวิธีผสมผสาน (hybrid sampling) ผลการวิจัยพบว่าตัวแบบ SVM เมื่อใช้ Kernel function เป็น polynomial 3rd degree และตัวแบบ LSTM ที่

เรียนรู้จากชุดข้อมูลที่สมดุลให้ค่า F1 สูงกว่าการเรียนรู้จากชุดข้อมูลที่ไม่สมดุล โดยมีค่า F1 เท่ากันเท่ากับ 66.67% แต่ตัวแบบ SVM มีค่าความเที่ยงสูงกว่าตัวแบบ LSTM ซึ่งมีค่าเท่ากับ 63.64% และ 52.94% ในขณะที่ตัวแบบ LSTM มีค่าการเรียกคืนสูงกว่าตัวแบบ SVM ซึ่งมีค่าเท่ากับ 90.00% และ 70.00% ตามลำดับ

คำสำคัญ: การจำแนกเสียง, อุปกรณ์ดักไอน้ำ, ซัพพอร์ตเวกเตอร์แมทรีน, หน่วยความจำระยะสั้นแบบยาว

Abstract

This research aims to create steam trap valve opening sound classification models using two classification methods including support vector machine (SVM) and long short-term memory (LSTM), and to compare the performance of the models. This study employs five feature extraction methods including zero-crossing rate, spectral centroid, Mel-frequency cepstral coefficient, spectral rolloff, and short-term Fourier transform. In addition, this study compares the performance of the models using two datasets including imbalanced dataset and balanced dataset which is resolved by using hybrid sampling method. The results show that SVM with Polynomial kernel function and LSTM provide higher F1 score when learning from balanced dataset than using imbalanced dataset. Moreover, SVM and LSTM provide the same F1 score of 66.67%. However, SVM provides higher precision than LSTM with value of 63.64% and 52.94%. In addition, LSTM gives higher recall than SVM with value of 90.00% and 70.00%, respectively.

Keywords: sound classification, steam trap, support vector machine, long short-term memory

บทนำ

อุปกรณ์ดักไอน้ำ (steam trap) เป็นวาล์วอัตโนมัติที่ช่วยในการระบายไอน้ำควบแน่น (condensate) ติดตั้งในหม้อไอน้ำซึ่งใช้ในกระบวนการของการผลิตในโรงงานอุตสาหกรรมบางประเภท หม้อไอน้ำเป็นหนึ่งในอุปกรณ์ที่สำคัญของระบบไอน้ำ ระบบไอน้ำเป็นระบบที่ผลิตไอน้ำประกอบด้วยอุปกรณ์และระบบย่อยต่าง ๆ ได้แก่ หม้อไอน้ำ ระบบส่งจ่ายไอน้ำ และระบบนำกลับไอน้ำควบแน่น (Ministry of Energy, 2010) ซึ่งระบบนำกลับไอน้ำควบแน่นนั้นจำเป็นต้องระบายไอน้ำที่เหลือจากการผลิตและไอน้ำควบแน่นให้ทัน หากระบายไอน้ำควบแน่นออกไม่ทันจะส่งผลให้มีไอน้ำควบแน่นค้างในระบบไอน้ำ ซึ่งอาจส่งผลต่อคุณภาพของผลิตภัณฑ์ ดังนั้นอุปกรณ์ดักไอน้ำ จึงช่วยในการระบายไอน้ำควบแน่นโดยสามารถปรับปริมาณการระบายไอน้ำควบแน่นได้ตามปริมาณไอน้ำควบแน่นที่เกิดขึ้นจริง (Department of Industrial Works, 2010) นอกจากนี้ยังทำหน้าที่ป้องกันการรั่วซึมของน้ำ และระบายก๊าซและอากาศออกจากระบบโดยไม่สูญเสียไอน้ำ หากอุปกรณ์ดักไอน้ำทำงานไม่มีประสิทธิภาพจะส่งผลต่อระบบไอน้ำ

น้ำอย่างมาก (Ministry of Energy, 2010) ดังนั้นจึงจำเป็นที่จะต้องทำการตรวจสอบคุณภาพในการทำงานของอุปกรณ์ดักไอน้ำอย่างสม่ำเสมอ วิธีการตรวจสอบคุณภาพการทำงานของอุปกรณ์ดักไอน้ำมีหลายวิธี เช่น การทดสอบด้วยการใช้น้ำรดอุปกรณ์ดักไอน้ำ การใช้เครื่องวัดอุณหภูมิทดสอบอุปกรณ์ดักไอน้ำ การเผาดูการระบายของอุปกรณ์ดักไอน้ำ การใช้เครื่องทดสอบอัลตราโซนิกทดสอบอุปกรณ์ดักไอน้ำ และการใช้เครื่องตรวจวัดอุปกรณ์ดักไอน้ำ (TM5) ข้อดีของการใช้เครื่องตรวจวัดอุปกรณ์ดักไอน้ำ คือ สามารถตรวจวัดและประมวลผลได้อย่างถูกต้องแม่นยำ เชื่อถือได้ และสามารถทำซ้ำได้ (Ministry of Energy, 2011) แต่ข้อเสียคือ มีค่าใช้จ่ายสูง ดังนั้นวิธีที่มีค่าใช้จ่ายต่ำกว่าคือ ตรวจสอบการทำงานของอุปกรณ์ดักไอน้ำโดยการฟังเสียงการเปิดวาล์วเพื่อระบายไอน้ำความดันของอุปกรณ์ดักไอน้ำ ซึ่งการตรวจจับเสียงการเปิดวาล์วของอุปกรณ์ดักไอน้ำโดยใช้เทคนิคการรู้จำเสียงอาจเป็นทางเลือกหนึ่งในการตรวจสอบคุณภาพของอุปกรณ์ดักไอน้ำแทนการซื้อเครื่องตรวจวัดอุปกรณ์ดักไอน้ำ

จากงานวิจัยที่เกี่ยวข้องกับการจำแนกเสียงหรือการตรวจจับความผิดปกติของสัญญาณเสียง พบว่าตัวแบบที่นิยมใช้มีหลายวิธี ได้แก่ วิธี Support Vector Machine (SVM) (Kaewtai, 2009; Giannakopoulos, 2015) วิธี k-Nearest Neighbor (k-NN) (Kaewtai, 2009; Giannakopoulos, 2015) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) (Kaewtai, 2009; Galván-Tejada, et al., 2016) วิธี Hidden Markov Model (HMM) (Giannakopoulos, 2015) วิธี Random Forest (RF) (Galván-Tejada, et al., 2016) วิธีโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Network: RNN) (Lim et al., 2019) และ วิธี Extreme Gradient Boosting (XGBoost) (Li et al., 2018) โดยพบว่าวิธี SVM จำแนกเพลงไทยเดิมได้ดีที่สุดให้ค่าความถูกต้องเท่ากับ 84.95% และสามารถจำแนกเสียงเพลงกับเสียงพูดได้ค่าความถูกต้องถึง 94.6% (Kaewtai, 2009; Giannakopoulos, 2015) สำหรับวิธี RNN สามารถจำแนกเสียงกรนได้ค่าความถูกต้องสูงเช่นเดียวกัน โดยมีค่าความถูกต้องถึง 98.9% (Lim et al., 2019) นอกจากนี้ในงานวิจัยที่มีการเปรียบเทียบประสิทธิภาพของวิธี SVM และ Naïve Bayes สำหรับการจำแนกกลุ่มข้อความ (Hassan et al., 2012) พบว่าตัวแบบ SVM ให้ค่า micro-average and macro-average F-measure สูงกว่าตัวแบบ Naïve Bayes ดังนั้นงานวิจัยนี้จึงสนใจเปรียบเทียบประสิทธิภาพของวิธี SVM และ LSTM ซึ่ง LSTM เป็นตัวแบบ RNN ชนิดพิเศษที่สามารถเรียนรู้อนุกรมเวลาระยะยาวในการจำแนกเสียงเปิดวาล์วของอุปกรณ์ดักไอน้ำ และในการสร้างตัวแบบการจำแนกเสียงจำเป็นต้องทำการแยกคุณลักษณะของเสียง (feature extraction) จากเสียงที่บันทึกก่อน ซึ่งวิธีแยกคุณลักษณะของเสียงที่นิยมใช้ได้แก่ spectral centroid (Kaewtai, 2009; Giannakopoulos, 2015) อัตราค่าตัดศูนย์ (zero crossing rate: ZCR) (Kaewtai, 2009; Giannakopoulos, 2015; Lim et al., 2019) สัมประสิทธิ์เซปสตรัมบนความถี่แบบเมล (mel-frequency cepstral coefficient: MFCC) (Kaewtai, 2009; Giannakopoulos, 2015; Galván-Tejada, et al., 2016; Lim et al., 2019) และการแปลงฟูเรียร์ระยะสั้น (short-time fourier transform: STFT) (Lim et al., 2019) นอกจากนี้ยังมีอีกหลายวิธี ได้แก่ ค่าสัมประสิทธิ์จากการประมาณค่าเชิงเส้น (linear predictive coefficient: LPC) (Kaewtai, 2009) power spectrum (Kaewtai, 2009) energy (Giannakopoulos, 2015) entropy of energy (Giannakopoulos, 2015) spectral spread (Giannakopoulos, 2015) spectral entropy (Giannakopoulos, 2015) spectral flux (Giannakopoulos,

2015) spectral rolloff (Giannakopoulos, 2015) chroma vector (Giannakopoulos, 2015) chroma deviation (Giannakopoulos, 2015) statistical features (Galván-Tejada, et al., 2016) และ wavelet packet energy (Li et al., 2018)

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบจำแนกเสียงเปิดวาล์วของอุปกรณ์คักไอน้ำจากเครื่องกำเนิดไอน้ำ และเปรียบเทียบประสิทธิภาพของตัวแบบของวิธี SVM กับวิธีหน่วยความจำระยะสั้นแบบยาว (long short-term memory: LSTM) โดยใช้เกณฑ์ค่า F1 (F1-score) ที่สูงที่สุด นอกจากนี้งานวิจัยนี้ทำการเปรียบเทียบประสิทธิภาพของตัวแบบทั้งสองโดยใช้ชุดข้อมูล 2 ชุด ได้แก่ ชุดข้อมูลที่ข้อมูลเสียงที่ไม่สมดุลซึ่งเป็นข้อมูลที่มีเสียงเปิดวาล์วน้อยกว่าเสียงที่ไม่ใช่เสียงเปิดวาล์ว กับชุดข้อมูลสมดุลที่ทำการแก้ไขปัญหาค่าข้อมูลเสียงที่ไม่สมดุลด้วยวิธีผสมผสาน (hybrid sampling) เพื่อหาตัวแบบที่เหมาะสมที่สุดสำหรับจำแนกเสียงเปิดวาล์วของอุปกรณ์คักไอน้ำ

ระเบียบวิธีวิจัย

1. ข้อมูลและการเก็บรวบรวมข้อมูล

ข้อมูลเสียงที่ใช้ในการศึกษาได้จากการบันทึกเสียงการทำงานของอุปกรณ์คักไอน้ำ โดยสร้างท่อใส่อุปกรณ์บันทึกเสียงและนำไปติดกับท่อที่ปล่อยคอนเดนเสท อากาศ และก๊าซ ออกจากระบบไอน้ำซึ่งเป็นบริเวณที่อยู่ใกล้กับอุปกรณ์คักไอน้ำ แล้วทำการบันทึกเป็นระยะเวลา 1 วัน (24 ชั่วโมง)

2. การเตรียมข้อมูล

การเตรียมข้อมูลซึ่งเป็นสัญญาณเสียงความยาว 24 ชั่วโมงเพื่อใช้ในการสร้างตัวแบบการจำแนกเสียง มีกระบวนการดังนี้ 1) ทำการปรับลดค่า sampling rate ของเสียง ซึ่งก็คือ การสุ่มตัวอย่างสัญญาณอนาล็อกซึ่งเป็นสัญญาณไฟฟ้ามาแปลงเป็นสัญญาณดิจิทัลที่คอมพิวเตอร์เข้าใจ ซึ่งสัญญาณเสียงที่ได้จากการบันทึกมีค่า sampling rate เท่ากับ 44,100 Hz จะทำการปรับลดเหลือ 16,000 Hz 2) เปลี่ยนสัญญาณเสียงจากระบบสเตอริโอเป็นระบบโมโน และกรองค่าความถี่ให้อยู่ในช่วง 3,000 Hz ถึง 7,000 Hz 3) ทำการตัดแบ่งไฟล์เสียงออกเป็นไฟล์เสียงที่มีความยาวไฟล์เสียงละ 2.5 วินาที บันทึกไฟล์เสียงให้อยู่ในรูปแบบไฟล์สัญญาณเสียงนามสกุล wav (.wav) ได้จำนวนไฟล์ทั้งหมด 24,510 ไฟล์ 4) ทำการแยกคุณลักษณะของไฟล์เสียงที่ตัดแบ่งโดยใช้วิธีที่อธิบายในหัวข้อ 3. จะได้ชุดข้อมูลที่มีข้อมูลทั้งหมด 24,510 แถว ในจำนวนนี้เป็นเสียงเปิดวาล์วจำนวน 31 แถวและไม่ใช่เสียงเปิดวาล์วจำนวน 24,479 แถว 4) สุ่มตัวอย่างข้อมูลโดยใช้วิธีสุ่มตัวอย่างแบบชั้นภูมิ โดยสุ่มตัวอย่างจากข้อมูลเสียงเปิดวาล์วและข้อมูลที่ไม่ใช่เสียงเปิดวาล์ว คิดเป็นร้อยละ 70 ของจำนวนข้อมูลในแต่ละกลุ่ม โดยการสุ่มอย่างง่าย รวมกันได้ชุดข้อมูลสำหรับการเรียนรู้ (training set) ส่วนที่เหลืออีกร้อยละ 30 ของข้อมูลแต่ละกลุ่มจะเป็นชุดข้อมูลสำหรับการทดสอบ (testing set)

นำชุดข้อมูลสำหรับการเรียนรู้มาแบ่งออกเป็น 2 ส่วน ส่วนที่หนึ่งสุ่มตัวอย่างร้อยละ 80 ของข้อมูลชุดนี้เป็นชุดข้อมูลการเรียนรู้ (training set) และที่เหลือร้อยละ 20 ของข้อมูลชุดนี้เป็นชุดข้อมูลตรวจสอบ (validation set) จำนวนข้อมูลของแต่ละชุดข้อมูลแสดงดังตารางที่ 1

ตารางที่ 1 จำนวนข้อมูลตัวอย่างของแต่ละชุดข้อมูล

กลุ่มข้อมูล	Training Set	Validation Set	Testing Set	รวม
เสียงเปิดวาล์ว	17	4	10	31
ไม่ใช่เสียงเปิดวาล์ว	13,709	3,428	7,342	24,479
รวม	13,726	3,432	7,352	24,510

3. การแยกคุณลักษณะของเสียง

งานวิจัยนี้ใช้วิธีการแยกคุณลักษณะของเสียงทั้งหมด 5 วิธี ได้แก่ ZCR spectral centroid spectral rolloff MFCC และ STFT โดยมีรายละเอียดของแต่ละวิธีดังนี้

3.1 อัตราค่าตัดศูนย์ (zero crossing rate: ZCR)

อัตราค่าตัดศูนย์ (ZCR) (Tzanetakis & Cook, 2002) คืออัตราการข้ามผ่านศูนย์ภายในเฟรม t กล่าวคือ อัตราค่าตัดศูนย์ของสัญญาณเสียง คือ อัตราของสัญญาณข้ามเส้นศูนย์แอมพลิจูดโดยเปลี่ยนจากค่าบวกเป็นค่าลบ หรือเปลี่ยนจากค่าลบเป็นค่าบวก อัตราค่าตัดศูนย์ของแต่ละเฟรมคำนวณโดยใช้สมการ (1)

$$ZCR_t = \sum_{i=1}^N \frac{|\text{sign}(x_i) - \text{sign}(x_{i-1})|}{2N} \quad (1)$$

โดยที่ $\text{sign}(x_i)$ คือ ฟังก์ชันที่ให้ค่าสัญญาณโดเมนเวลา x_i ซึ่งมีค่าเป็น 1 เมื่อ x_i อยู่เหนือเส้นสัญญาณศูนย์ หรือมีค่าเป็น -1 ถ้า x_i อยู่ต่ำกว่าเส้นสัญญาณศูนย์ นอกจากนี้มีค่าเป็น 0 ถ้า x_i อยู่บนเส้นสัญญาณศูนย์

N คือ ความยาวของสัญญาณภายในเฟรม t

i คือ หมายเลขช่วงความถี่ ซึ่งมีค่าอยู่ในช่วง 1 ถึง N

3.2 Spectral Centroid

Spectral Centroid (Tzanetakis & Cook, 2002) เป็นการวัดตำแหน่งและรูปร่างของสเปกตรัมอย่างง่าย ที่แสดงถึงค่าศูนย์กลางของสเปกตรัมจากการแปลงฟูเรียร์ระยะสั้น (Short-Time Fourier Transform) Spectral Centroid (C_t) ของเฟรม t คำนวณโดยใช้สมการ (2)

$$C_t = \frac{\sum_{i=1}^N M_t[i] \times i}{\sum_{i=1}^N M_t[i]} \quad (2)$$

โดยที่ $M_t[i]$ คือ การแปลงฟูเรียร์ที่เฟรม t และช่วงความถี่ที่ i

N คือ ความยาวของสัญญาณในเฟรม t

3.3 Spectral rolloff

Spectral rolloff คือ การวัดความเบ้ของรูปร่างสเปกตรัมซึ่งมีลักษณะเบ้ขวา สำหรับสเปกตรัมที่เบ้ขวาจะเป็นส่วนที่มีค่า spectral rolloff สูง กำหนดให้เป็นองค์ประกอบทางความถี่รอง ภายใต้อัตรา 85 เปอร์เซ็นต์ของการกระจายขนาดของสเปกตรัม (Tzanetakis & Cook, 2002; Burred & Lerch, 2004) คำนวณโดยใช้สมการ (3)

$$\sum_{i=1}^K |X_t[i]| \leq 0.85 \sum_{i=1}^{N/2} |X_t[i]| \quad (3)$$

โดยที่ $X_t[i]$ คือ แมกนิจูดของสเปกตรัมที่สอดคล้องกัน

i คือ หมายเลขช่วงความถี่

K คือ ความถี่ spectral rolloff ในเฟรม t

N คือ ความยาวของสัญญาณในเฟรม t

3.4 สัมประสิทธิ์เซปสตรัมบนความถี่แบบเมล (mel-frequency cepstral coefficient: MFCC)

เซปสตรัม (cepstral) เป็นการแปลงโคไซน์แบบไม่ต่อเนื่อง (discrete cosine transform) ของลอการิทึมจากสเปกตรัมสัญญาณในช่วงสั้น ๆ สัมประสิทธิ์เซปสตรัมบนความถี่แบบเมลเป็นเทคนิคที่ปรับปรุงจากเซปสตรัม ด้วยการปรับสเกลของสเปกตรัมให้อยู่บนสเกลที่เหมาะสมสำหรับการฟังของมนุษย์โดยสังเกตจากลักษณะของสัญญาณเสียง สัญญาณเสียงในช่วงความถี่ต่ำจะมีความสำคัญมากกว่าช่วงความถี่สูง จึงออกแบบสเกลของสเปกตรัมให้สามารถเก็บรายละเอียดของสัญญาณเสียงช่วงความถี่ต่ำได้มากกว่า เรียกการออกแบบนี้ว่าสเกลเมล (mel scale) โดยมีขั้นตอนในการคำนวณค่าสัมประสิทธิ์เซปสตรัมบนสเกลเมล (Junto, 2015; Kopanyapipat, 2015) ดังนี้

ในการคำนวณค่า MFCC จะเริ่มจาก 1) แปลงสัญญาณเสียงให้อยู่ในโดเมนของความถี่ โดยวิธีการแปลงฟูเรียร์แบบเร็ว (fast fourier transform: FFT) 2) คำนวณหาพลังงานสเปกตรัมที่ผ่านตัวกรอง (mel scale filtering) ขั้นตอนนี้นำความถี่ที่ได้จากการคำนวณสเปกตรัม (ค่า FFT) มาหาขนาดกำลังสอง ได้ $|x(k)|^2$ ส่งผ่านชุดตัวกรองแบบสามเหลี่ยมในสเกลเมล เพื่อเน้นความสำคัญของความถี่ที่อยู่ในช่วงกลางของชุดตัวกรองแต่ละตัวกรองตามสมการ (4) โดยที่ความถี่กลางของตัวกรองแต่ละชุดนั้นเกิดจากการแปลงค่าความถี่ปกติ (f) ให้อยู่บนสเกลเมล ($Mel(f)$) ตามสมการ (5) 3) การคำนวณสัมประสิทธิ์เซปสตรัมบนสเกลเมล (MFCC) จะนำลอการิทึม ($\log(.)$) ของพลังงานมาผ่านการแปลงโคไซน์แบบไม่ต่อเนื่อง (discrete cosine transform) ทำให้ได้ค่าสัมประสิทธิ์เซปสตรัมบนสเกลเมล c ลำดับที่ m ตามสมการ (6)

$$E_j = \sum_{k=0}^{n/2-1} \phi_j(k) |x(k)|^2; 0 \leq j \leq J \quad (4)$$

เมื่อ ϕ_j คือ ค่าประจำตัวกรองที่ j

$x(k)$ คือ สเปกตรัม

และ E_j คือ ค่าพลังงานสเปกตรัมที่ผ่านตัวกรอง

$$\text{Mel}(f) = 1125 \times \ln\left(1 + \frac{f}{700}\right) \quad (5)$$

โดยที่ $\text{Mel}(f)$ คือ ความถี่แบบเมล (Mel-Frequency)

และ f คือ ความถี่ปกติ หน่วยเป็น เฮิรตซ์ (Hz)

$$c_m = w_t(m) \sum_{j=1}^J \log_{10}(E_j) \cos\left(\frac{\pi}{J}(j - 0.5)m\right) ; m = 0, 1, \dots, J - 1 \quad (6)$$

$$\text{เมื่อ } w_t(m) = \begin{cases} \frac{1}{\sqrt{J}}, m = 0 \\ \frac{\sqrt{2}}{J}, 1 \leq m < J \end{cases}$$

โดยที่ $w_t(m)$ คือ ฟังก์ชันเงื่อนไข

j คือ อันดับของตัวกรอง

J คือ จำนวนตัวกรอง

3.5 การแปลงฟูเรียร์ระยะสั้น (short-time fourier transform: STFT)

การแปลงฟูเรียร์ระยะสั้น เป็นอีกหนึ่งวิธีที่มีประสิทธิภาพสำหรับการประมวลผลสัญญาณเสียง ซึ่งจะแปลงสัญญาณเสียงจากโดเมนเวลาไปเป็นโดเมนเวลาและความถี่ โดยอาศัยฟังก์ชันหน้าต่าง สามารถระบุแอมพลิจูดที่ซับซ้อนเทียบกับเวลา และความถี่ของสัญญาณเสียง (Colen, 1995) ในการคำนวณ STFT จะเริ่มต้นจากการแบ่งสัญญาณเสียงออกเป็นช่วงสั้น ๆ แล้วคูณแต่ละส่วนตามฟังก์ชันหน้าต่าง เนื่องจากการแปลงฟูเรียร์แบบไม่ต่อเนื่องทางเวลา (discrete-time fourier transform: DTFT) จะได้สูตรคำนวณ STFT (Ahmadizadeh, 2014) ดังสมการ (7)

$$X(\omega, n_0) = \sum_{n=-N}^{N-1} w[n]x[n + n_0]\exp(-j\omega n) \quad (7)$$

โดยที่ $(x[n_0 - N], \dots, x[n_0 + N - 1])^T$ คือ ส่วนของสัญญาณเสียง

$w[n]$ คือ ฟังก์ชันหน้าต่าง

และ $X(\omega, n_0)$ คือ STFT เป็นฟังก์ชันของทั้งความถี่ ω และเวลา n_0

4. การแก้ไขปัญหาข้อมูลไม่สมดุล

เนื่องจากข้อมูลเสียงในชุดข้อมูลเรียนรู้เป็นข้อมูลที่ไม่สมดุลในแต่ละกลุ่มจึงทำการแก้ปัญหาด้วยวิธีการปรับข้อมูลตัวอย่างแบบผสมผสาน (hybrid sampling) เป็นการนำเทคนิควิธีสุ่มเกิน (Oversampling) ซึ่งเป็นการสุ่มเพิ่มจำนวนข้อมูลในกลุ่มส่วนน้อยให้มีจำนวนใกล้เคียงหรือเท่ากับจำนวนของกลุ่มส่วนมาก และวิธีสุ่มลด (Undersampling) ซึ่งเป็นวิธีการสุ่มลดจำนวนข้อมูลในกลุ่มส่วนมากให้มีจำนวนใกล้เคียงหรือเท่ากับจำนวนในกลุ่มส่วนน้อย มาทำงานร่วมกัน (Kesornsit, et al., 2018) โดยมีขั้นตอนในการปรับข้อมูลตัวอย่างดังต่อไปนี้

- 1) หาค่ากลางของจำนวนข้อมูลเพื่อเป็นจำนวนในการสุ่มตัวอย่างโดยใช้จำนวนข้อมูลในกลุ่มส่วนมาก และจำนวนข้อมูลในกลุ่มส่วนน้อย ซึ่งสามารถคำนวณได้จากสมการที่ (8)

$$\text{center} = \frac{(\text{class}_0 + \text{class}_1)}{2} \quad (8)$$

เมื่อ class_0 คือ จำนวนข้อมูลของกลุ่มส่วนมาก ซึ่งเป็นกลุ่มข้อมูลเสียงที่ไม่ใช่เสียงเปิดวาล์ว

และ class_1 คือ จำนวนข้อมูลของกลุ่มส่วนน้อย ซึ่งเป็นกลุ่มข้อมูลเสียงเปิดวาล์ว

- 2) ใช้เทคนิควิธีการสุ่มเกินเพื่อเพิ่มจำนวนข้อมูลในกลุ่มเสียงเปิดวาล์วโดยใช้วิธีการสุ่มตัวอย่างแบบง่ายในการสุ่มข้อมูลเสียงเปิดวาล์วให้มีจำนวนเพิ่มขึ้นจนเท่ากับค่ากลางที่คำนวณได้จากข้อ 1
- 3) ใช้เทคนิควิธีการสุ่มลดเพื่อลดข้อมูลในกลุ่มเสียงที่ไม่ใช่เสียงเปิดวาล์วโดยใช้วิธีการสุ่มตัวอย่างแบบง่ายในการสุ่มข้อมูลที่ไม่ใช่เสียงเปิดวาล์วให้มีจำนวนเท่ากับค่ากลางที่คำนวณได้จาก 1
- 4) รวมข้อมูลที่ได้ในข้อ 2) และข้อ 3) ข้อมูลที่ได้จะเป็นข้อมูลที่มีจำนวนข้อมูลเสียงเปิดวาล์วและข้อมูลที่ไม่ใช่เสียงเปิดวาล์วเท่ากัน เพื่อใช้ในการสร้างตัวแบบและทดสอบตัวแบบต่อไป

5. เทคนิคการจำแนกเสียง

งานวิจัยนี้สร้างตัวแบบการจำแนกเสียงและทำการเปรียบเทียบตัวแบบจากวิธีซัพพอร์ตเวกเตอร์แมชชีน และวิธีหน่วยความจำระยะสั้นแบบยาว โดยมีรายละเอียดแต่ละวิธีดังนี้

5.1 วิธีซัพพอร์ตเวกเตอร์แมชชีน (support vector machine: SVM)

ซัพพอร์ตเวกเตอร์แมชชีน เป็นการเรียนรู้แบบมีผู้สอน ซึ่งนำมาช่วยแก้ปัญหาการจำแนกกลุ่มของข้อมูล โดยอาศัยหลักการของการหาค่าเหมาะที่สุด (optimization) เพื่อประมาณค่าน้ำหนักของตัวแบบเชิงเส้น เพื่อให้ได้เส้นตรงที่ใช้แบ่งกลุ่มข้อมูล (hyperplane) ที่เหมาะสมที่สุด แนวคิดของวิธี SVM คือการแปลงค่าของข้อมูลใน input space ไปยังพื้นที่ที่คุณลักษณะ (feature space) โดยอาศัย kernel function แล้วหาเส้นตรงที่ใช้แบ่งข้อมูลทั้งสองออกจากกัน โดยเส้นตรงที่แบ่งข้อมูลสองกลุ่มที่ดีที่สุดเป็นเส้นตรงที่มีระยะขอบมากที่สุด (Pinmuang & Thonhkam, 2017) แสดงดังรูปที่ 1 และสมการสำหรับการตัดสินใจ (Thamsiri & Meesad, 2011) แสดงดังสมการที่ (9)

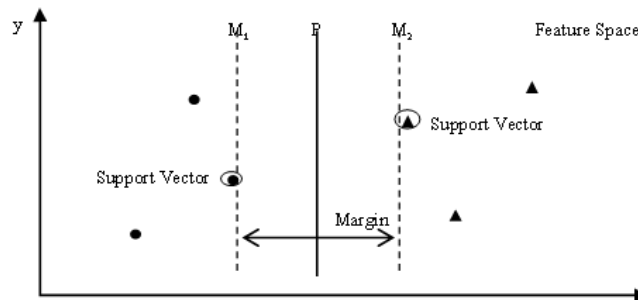
$$y = \text{sign}(\sum_{i=1}^{n_s} w_i \phi_i(x) \phi_i(x_i) + b) \quad (9)$$

$$\phi(x) = [\phi_1(x_1), \phi_2(x_2), \dots, \phi_n(x_n)]^T \quad (10)$$

โดย y คือ ค่ากลุ่มของข้อมูลมีค่าเป็น 1 เมื่อข้อมูลอยู่ในกลุ่มที่สนใจ (เป็นเสียงเปิดวาล์ว) และมีค่าเป็น -1 เมื่อข้อมูลอยู่ในกลุ่มที่ไม่สนใจ (ไม่ใช่เสียงเปิดวาล์ว)

w_i คือ ค่าน้ำหนัก (weight)

- x คือ ค่าคุณลักษณะหรือข้อมูลนำเข้า
 x_i คือ ซัพพอร์ตเวกเตอร์
 n_s คือ จำนวนซัพพอร์ตเวกเตอร์
 และ b คือ ค่าเอนเอียง (bias)



รูปที่ 1 เส้นตรงในการจำแนกข้อมูลและระยะขอบ (Margin)

แต่ข้อมูลจริงที่พบส่วนใหญ่ นั้น ข้อมูลทั้ง 2 กลุ่มไม่สามารถแบ่งได้ด้วยสมการเส้นตรง ดังนั้นจึงต้องมีเครื่องมือมาช่วยให้ข้อมูลเหล่านั้นเรียงตัวใหม่ในพื้นที่หลักมิติ ซึ่งเครื่องมือเหล่านั้นคือ kernel function ดังสมการ (11)

$$K(x_i, x_j) = \phi(x_i)\phi(x_j) \quad (11)$$

ทั้งนี้ kernel function ที่นิยมใช้มีอยู่ 3 ชนิดด้วยกัน (Kaewtai, 2009) ได้แก่

- โพลีโนเมียล (polynomial)

$$K(x_i, x_j) = (\langle x_i^T x_j \rangle + 1)^d \quad (12)$$

เมื่อ d คือ ค่าเลขยกกำลัง

- เรเดียลเบสฟังกชัน (radial basis function: RBF)

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (13)$$

เมื่อ σ คือ ค่าพารามิเตอร์

- ซิกมอยด์ (sigmoid)

$$K(x_i, x_j) = \tanh(\alpha x_i^T x_j + c) \quad (14)$$

เมื่อ α และ c คือ ค่าพารามิเตอร์

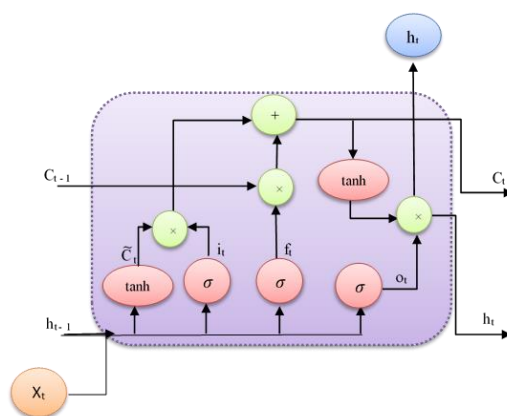
ดังนั้นสมการสำหรับการจำแนกข้อมูลสามารถแสดงดังสมการที่ (15)

$$f(x) = \text{sign}\left(\sum_{i=1}^{n_s} w_i K(x_i, x_j) + b\right) \quad (15)$$

เมื่อ $f(x)$ คือ ฟังก์ชันที่ใช้ในการจำแนกกลุ่มของ SVM โดยหาก $f(x)$ มีค่าเป็นบวก จะจำแนกเสียงเป็นเสียงเปิดวาล์ว และหาก มีค่าเป็นลบ จะจำแนกเสียงเป็นไม่ใช่เสียงเปิดวาล์ว

5.2 วิธีหน่วยความจำระยะสั้นแบบยาว (long short-term memory: LSTM)

หน่วยความจำระยะสั้นแบบยาว (LSTM) เป็นโครงข่ายประสาทเทียมแบบวนกลับชนิดพิเศษที่สามารถเรียนรู้อนุกรมเวลาระยะยาว หน่วยความจำระยะสั้นแบบยาวได้รับการออกแบบมาเพื่อหลีกเลี่ยงปัญหาการจดจำในระยะยาว โดยออกแบบให้มีเกต (gate) ทั้งหมด 3 เกต ได้แก่ อินพุตเกต (input gate) ฟอว์เกตเกต (forget gate) และ เอาต์พุตเกต (output gate) (Apaydin et al., 2020) นอกจากนี้ยังมีสถานะอินพุตของเซลล์ (input cell state: $\tilde{C}_t$) หน่วยความจำระยะสั้นแบบยาวมีโครงสร้างเป็นลูกโซ่ที่คล้ายกับโครงข่ายประสาทเทียมแบบวนกลับอย่างง่าย แต่มีความแตกต่างกันในโครงสร้างของตัวแบบที่ซับซ้อนกว่า เช่น เกต 3 เกต สำหรับทำหน้าที่ควบคุมการไหลของข้อมูลเข้าหรือข้อมูลออก และมีเซลล์สำหรับจดจำช่วงเวลาต่าง ๆ แสดงดังรูปที่ 2 (Cui et al., 2018)



รูปที่ 2 โครงสร้างของหน่วยความจำระยะสั้นแบบยาว (LSTM)

สำหรับการคำนวณค่าอินพุตเกต (i_t) ฟอว์เกตเกต (f_t) เอาต์พุตเกต (o_t) และ สถานะอินพุตของเซลล์ (input cell state: $\tilde{C}_t$) ของเวลา t จะคำนวณจากสมการ (16) ถึง (19)

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (16)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (17)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (18)$$

$$\tilde{C}_t = \tanh(W_C x_t + U_C h_{t-1} + b_C) \quad (19)$$

โดยที่ W_p , W_h , W_o และ W_C เป็นเมทริกซ์น้ำหนักเลเยอร์ซ่อน (hidden layer) ข้อมูลเข้าของแต่ละเกทและสเตท
 U_p , U_h , U_o และ U_C เป็นเมทริกซ์น้ำหนักที่เชื่อมต่อสถานะเอาต์พุตของเซลล์ในอดีตกับแต่ละเกทและสเตท
 b_p , b_h , b_o และ b_C เป็นเวกเตอร์ค่าความเอนเอียงของแต่ละเกทและสเตท

และ σ_g คือ Activation Function ของเกท ซึ่งจะป็นฟังก์ชัน Sigmoid
 จะใช้สมการ (20) ถึง (21) เพื่อคำนวณสถานะเอาต์พุตของเซลล์ (C_t) และเอาต์พุตเลเยอร์ (h_t)

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (20)$$

$$h_t = o_t \times \tanh(C_t) \quad (21)$$

โดยที่ $\tanh$ เป็น Activation Function ของเอาต์พุตเลเยอร์

สำหรับ Activation Function ของเอาต์พุตเลเยอร์ที่ใช้ในงานวิจัยนี้มี 3 ชนิด ได้แก่ rectified linear unit (relu), hyperbolic tangent (tanh) และ sigmoid ซึ่งเป็นฟังก์ชันทั้งปวงที่ใช้งานสำหรับ LSTM สามารถคำนวณได้ดังสมการที่ (22) ถึง สมการที่ (24)

- Sigmoid

$$g(z) = \frac{1}{1 + e^{-z}} \quad (22)$$

- Tanh

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (23)$$

- ReLU

$$g(z) = \max(0, z) \quad (24)$$

6. การกำหนดพารามิเตอร์ของตัวแบบ

ในการสร้างตัวแบบการจำแนกเสียงเปิดวาล์วของอุปกรณ์คักไอน้ำ โดยใช้วิธีการทั้งสองวิธีจำเป็นต้องกำหนดและปรับค่าพารามิเตอร์เพื่อให้ได้ตัวแบบที่ให้ประสิทธิภาพในการจำแนกเสียงดีที่สุด การปรับค่าพารามิเตอร์ของตัวแบบพิจารณาจากการศึกษางานวิจัยที่เกี่ยวข้องกับการศึกษาในครั้งนี้ (Albon, 2017; Navlani, 2019)

โดยวิธี SVM จะทำการปรับค่าพารามิเตอร์ของตัวแบบทั้งหมด 3 ค่า ได้แก่ kernel function, C และ gamma แสดงค่าพารามิเตอร์ที่ปรับดังต่อไปนี้

-

- kernel function คือ การแปลงข้อมูลอินพุตของชุดข้อมูลที่กำหนดให้อยู่ในรูปแบบที่ต้องการ ในงานวิจัยนี้ใช้ kernel function 4 ฟังก์ชันได้แก่ linear, polynomial of 3rd degree, sigmoid และ radial basis function
- C คือ ค่าที่กำหนดให้มีจุดข้อมูลละเมิดเส้นขอบ โดยส่งผลต่อความกว้างของเส้นขอบ ในงานวิจัยนี้กำหนด 4 ค่า ได้แก่ 10, 100, 500 และ 1,000
- gamma คือ การกำหนดการแพร่กระจายของ kernel จึงเป็นขอบเขตการตัดสินใจ หากมีค่าต่ำ เส้นโค้งของขอบเขตการตัดสินใจจะต่ำ ในทางตรงข้ามหากมีค่าสูงขอบเขตการตัดสินใจจะกว้าง ดังนั้นเส้นโค้งของขอบเขตการตัดสินใจจะสูง ซึ่งจะสร้างเป็นขอบเขตการตัดสินใจรอบ ๆ จุดข้อมูล ในงานวิจัยนี้กำหนด 7 ค่า ได้แก่ 0.000001 0.00001 0.0001 0.001 0.01 0.1 และ 1.0 ยกเว้น gamma ของ linear function จะมีค่าเท่ากับ 1

วิธี LSTM จะทำการสร้างตัวแบบจำลองตามลำดับ (sequential model) เป็นการเรียงซ้อนของชั้นเชิงเส้น โดยมีชั้นของข้อมูลเข้า (input layer) และชั้นของข้อมูลผลการวิเคราะห์ (output layer) นอกจากนี้ยังมีฟังก์ชันสำหรับการเรียนรู้ และการทำนาย ซึ่งจะมีการปรับค่าพารามิเตอร์ (Zhu & Chollet, 2020; Brownlee, 2017; Lim et al., 2019; TensorFlow, 2021) ดังนี้

ชั้นของข้อมูลเข้า (input layer) ซึ่งเป็นการกำหนดลักษณะต่าง ๆ ของข้อมูลเข้า จะกำหนดคุณสมบัติของชั้นข้อมูลเข้า 4 ค่า ได้แก่ units, input shape, dropout และ return sequences มีการกำหนดค่าดังต่อไปนี้

- units กำหนดค่าเท่ากับ 64
- input shape กำหนดค่าเป็น (1, 536) โดยค่าแรกคือ จำนวนแถวของข้อมูลเข้าในแต่ละรอบ ซึ่งในแต่ละรอบจะนำเข้า 1 แถว และค่าที่สอง คือจำนวนของตัวแปรหรือจำนวนคุณลักษณะ (feature) ของข้อมูล ซึ่งมี 536 ค่า
- dropout กำหนดให้มีค่าตั้งแต่ 0.1 จนถึง 0.5 เพิ่มทีละ 0.1
- return sequences กำหนดให้มีค่าเป็นจริง (true)

ชั้นของข้อมูลผลการวิเคราะห์ (output layer) จะต้องกำหนดคุณสมบัติของชั้นผลการวิเคราะห์ 2 ค่า ได้แก่ units และ activation function โดย units กำหนดค่าเท่ากับ 1 เพราะเป็นการสร้างตัวแบบเพื่อจำแนกกลุ่ม 2 กลุ่มจะแสดงค่าความน่าจะเป็นที่จะเป็นกลุ่ม 1 หรือกลุ่มที่สนใจ ส่วน activation function ของ output layer ใช้ 3 ฟังก์ชัน ได้แก่ tanh, relu และ sigmoid ดังแสดงในสมการที่ (22) ถึง สมการที่ (24)

ฟังก์ชันสำหรับการเรียนรู้จะใช้ฟังก์ชัน compile เป็นการกำหนดค่ากระบวนการเรียนรู้ของตัวแบบ มีการกำหนดค่า 3 ค่า ได้แก่ loss, optimizer และ metrics โดยกำหนดค่าดังต่อไปนี้

- loss เป็นการกำหนดฟังก์ชันการสูญเสีย (loss function) ให้เป็น binary_crossentropy เนื่องจากเป็นปัญหาการจำแนกกลุ่ม 2 กลุ่ม ผลลัพธ์การทำนายของตัวแบบจะบอกค่าความน่าจะเป็นว่ามีโอกาสที่จะเป็นกลุ่ม 1 โดยปกติจะกำหนดให้ 1 แทนกลุ่มที่สนใจและ 0 แทนกลุ่มที่ไม่สนใจ
- optimizer ใช้อัลกอริทึม ADAM ซึ่งอัลกอริทึมนี้สามารถปรับค่า learning rate สำหรับพารามิเตอร์ในแต่ละครั้งได้ ดังนั้นจึงปรับค่า learning rate ทั้งหมด 6 ค่า ได้แก่ 0.1 0.01 0.001 0.0001 และ 0.00001

- metrics เป็นการกำหนดเกณฑ์การประเมินผลตัวแบบ งานวิจัยนี้ใช้เกณฑ์ accuracy คือให้แสดงในรูปแบบของค่าความถูกต้อง

ฟังก์ชันสำหรับการทำนายจะใช้ฟังก์ชัน fit เป็นฟังก์ชันในการเรียนรู้โดยใช้ข้อมูลเรียนรู้ (training set) ซึ่งทำการกำหนดค่าอาร์กิวเมนต์ 2 ค่า ได้แก่ epochs และ batch size โดยกำหนดค่าดังต่อไปนี้

- epochs เท่ากับ 100 เป็นการวนซ้ำข้อมูล x และ y ในการเรียนรู้ตัวแบบโดยแต่ละ Epoch จะทำให้ Loss ลดลง ในขณะที่ Accuracy เพิ่มขึ้น
- batch size เท่ากับ 64 เป็นการกำหนดจำนวนตัวอย่างต่อการอัปเดตค่า gradients

ในการแยกคุณลักษณะของเสียงจากโดยใช้วิธีการ 5 วิธีและการสร้างตัวแบบทั้งสองวิธี เขียนโปรแกรมโดยใช้ภาษา Python บน Google Colab (<https://colab.research.google.com/>)

7. การวัดประสิทธิภาพของตัวแบบ

งานวิจัยนี้ใช้เกณฑ์การวัดประสิทธิภาพของตัวแบบ ได้แก่ ค่าความเที่ยง (precision) ค่าการเรียกคืน (recall) และค่า F1 (F1-score) โดยมีทางเลือกในการสรุปผลการทำนายดังตารางที่ 2

ตารางที่ 2 Confusion Matrix

ค่าทำนาย	ค่าจริง	
	เสียงเปิดวาล์ว	ไม่ใช่เสียงเปิดวาล์ว
เสียงเปิดวาล์ว	TP	FP
ไม่ใช่เสียงเปิดวาล์ว	FN	TN

โดยที่ true positive (TP) คือ ข้อมูลที่ค่าจริงเป็นเสียงเปิดวาล์ว ทำนายถูกว่าเป็นเสียงเปิดวาล์ว
 true negative (TN) คือ ข้อมูลที่ค่าจริงไม่ใช่เสียงเปิดวาล์ว ทำนายถูกว่าไม่ใช่เสียงเปิดวาล์ว
 false positive (FP) คือ ข้อมูลที่ค่าจริงไม่ใช่เสียงเปิดวาล์ว ทำนายผิดว่าเป็นเสียงเปิดวาล์ว
 false negative (FN) คือ ข้อมูลที่ค่าจริงเป็นเสียงเปิดวาล์ว ทำนายผิดว่าไม่ใช่เสียงเปิดวาล์ว

7.1 ค่าความเที่ยง (precision)

คำนวณหาการทำนายที่ถูกต้องเพื่อวัดความแม่นยำของตัวแบบ โดยคิดจากสัดส่วนของจำนวนการทำนายเสียงเปิดวาล์วถูกต้อง (true positive) เทียบกับผลรวมของจำนวนการทำนายว่าเป็นเสียงเปิดวาล์วทั้งหมด

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (25)$$

7.2 ค่าการเรียกคืน (recall)

คำนวณหาความถูกต้องในการทำนายเชิงเปิดวาล์วของตัวแบบโดยคิดจากสัดส่วนของจำนวนการทำนายว่าเป็นเชิงเปิดวาล์วถูกต้อง (true positive) เทียบกับจำนวนเชิงเปิดวาล์วในข้อมูลชุดนั้น

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (26)$$

7.3 ค่า F1 (F1-score)

ค่าเฉลี่ยฮาร์โมนิก (harmonic mean) ระหว่างค่าความเที่ยง (precision) และค่าการเรียกคืน (recall) และเป็นค่าที่ใช้เพื่อหาตัวแบบที่มีความสมดุลที่ดีระหว่างค่าความเที่ยง และค่าการเรียกคืน คำนวณได้จากสมการที่ (24)

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (27)$$

8. ผลและวิจารณ์ผลการทดลอง

จากการทดลองสร้างตัวแบบด้วยวิธี SVM และวิธี LSTM โดยทำการปรับค่าพารามิเตอร์ตามที่กล่าวไว้ในหัวข้อ 6 ดำเนินการสร้างตัวแบบโดยใช้ชุดข้อมูลการเรียนรู้ที่ข้อมูลไม่สมดุลกับชุดข้อมูลการเรียนรู้ที่ข้อมูลสมดุลที่ปรับแก้ด้วยการใช้วิธีผสมผสาน ผลการศึกษาเมื่อทดสอบด้วยข้อมูลชุดทดสอบแสดงดังตารางที่ 3 และตารางที่ 4

ตารางที่ 3 ตารางผลการวิเคราะห์ของข้อมูลชุดทดสอบที่สร้างตัวแบบด้วยวิธี SVM

พารามิเตอร์			ชุดการเรียนรู้ข้อมูลที่ไม่สมดุล			ชุดการเรียนรู้ข้อมูลสมดุล		
kernel	gamma	C	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)
linear	1	100	62.50	50.00	55.56	50.00	60.00	54.55
linear	1	500	62.50	50.00	55.56	50.00	60.00	54.55
linear	1	1000	62.50	50.00	55.56	50.00	60.00	54.55
poly	0.000001	10	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.000001	100	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.000001	500	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.000001	1000	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.00001	10	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.00001	100	0.00	0.00	0.00	0.00	0.00	0.00
poly	0.00001	500	0.00	0.00	0.00	50.00	10.00	16.67
poly	0.00001	1000	0.00	0.00	0.00	28.57	20.00	23.53

poly	0.0001	10	0.00	0.00	0.00	27.78	50.00	35.71
poly	0.0001	100	0.00	0.00	0.00	50.00	70.00	58.33
poly	0.0001	500	100.00	10.00	18.18	53.85	70.00	60.87
poly	0.0001	1000	50.00	10.00	16.67	58.33	70.00	63.64
poly	0.001	10	100.00	30.00	46.15	58.33	70.00	63.64
poly	0.001	100	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.001	500	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.001	1000	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.01	10	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.01	100	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.01	500	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.01	1000	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.1	10	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.1	100	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.1	500	80.00	40.00	53.33	63.64	70.00	66.67
poly	0.1	1000	80.00	40.00	53.33	63.64	70.00	66.67
poly	1	10	80.00	40.00	53.33	63.64	70.00	66.67
poly	1	100	80.00	40.00	53.33	63.64	70.00	66.67
poly	1	500	80.00	40.00	53.33	63.64	70.00	66.67
poly	1	1000	80.00	40.00	53.33	63.64	70.00	66.67
sigmoid	0.000001	10	0.00	0.00	0.00	13.16	100.00	23.26
sigmoid	0.000001	100	0.00	0.00	0.00	40.00	80.00	53.33
sigmoid	0.000001	500	0.00	0.00	0.00	44.44	80.00	57.14
sigmoid	0.000001	1000	33.33	10.00	15.38	44.44	80.00	57.14
sigmoid	0.00001	10	0.00	0.00	0.00	40.00	80.00	53.33
sigmoid	0.00001	100	33.33	10.00	15.38	44.44	80.00	57.14
sigmoid	0.00001	500	50.00	30.00	37.50	50.00	70.00	58.33
sigmoid	0.00001	1000	57.14	40.00	47.06	54.55	60.00	57.14
sigmoid	0.0001	10	33.33	10.00	15.38	47.06	80.00	59.26

sigmoid	0.0001	100	55.56	50.00	52.63	53.85	70.00	60.87
sigmoid	0.0001	500	41.18	70.00	51.85	42.11	80.00	55.17
sigmoid	0.0001	1000	36.84	70.00	48.28	33.33	80.00	47.06
sigmoid	0.001	10	27.27	30.00	28.57	9.28	90.00	16.82
sigmoid	0.001	100	25.00	30.00	27.27	9.18	90.00	16.67
sigmoid	0.001	500	23.08	30.00	26.09	9.18	90.00	16.67
sigmoid	0.001	1000	23.08	30.00	26.09	9.18	90.00	16.67
sigmoid	0.01	10	0.00	0.00	0.00	0.47	70.00	0.94
sigmoid	0.01	100	0.00	0.00	0.00	0.47	70.00	0.94
sigmoid	0.01	500	0.00	0.00	0.00	0.47	70.00	0.94
sigmoid	0.01	1000	0.00	0.00	0.00	0.47	70.00	0.94
sigmoid	0.1	10	50.00	10.00	16.67	0.27	60.00	0.55
sigmoid	0.1	100	50.00	10.00	16.67	0.27	60.00	0.55
sigmoid	0.1	500	50.00	10.00	16.67	0.27	60.00	0.55
sigmoid	0.1	1000	50.00	10.00	16.67	0.27	60.00	0.55
sigmoid	1	10	33.33	10.00	15.38	0.40	80.00	0.79
sigmoid	1	100	25.00	10.00	14.29	0.40	80.00	0.79
sigmoid	1	500	25.00	10.00	14.29	0.40	80.00	0.79
sigmoid	1	1000	25.00	10.00	14.29	0.40	80.00	0.79
rbf	0.000001	10	0.00	0.00	0.00	22.50	90.00	36.00
rbf	0.000001	100	0.00	0.00	0.00	42.11	80.00	55.17
rbf	0.000001	500	33.33	10.00	15.38	44.44	80.00	57.14
rbf	0.000001	1000	33.33	10.00	15.38	50.00	80.00	61.54
rbf	0.00001	10	0.00	0.00	0.00	42.11	80.00	55.17
rbf	0.00001	100	33.33	10.00	15.38	47.06	80.00	59.26
rbf	0.00001	500	57.14	40.00	47.06	54.55	60.00	57.14
rbf	0.00001	1000	57.14	40.00	47.06	50.00	60.00	54.55
rbf	0.0001	10	33.33	10.00	15.38	43.75	70.00	53.85
rbf	0.0001	100	57.14	40.00	47.06	53.85	70.00	60.87

rbf	0.0001	500	62.50	50.00	55.56	53.85	70.00	60.87
rbf	0.0001	1000	62.50	50.00	55.56	53.85	70.00	60.87
rbf	0.001	10	50.00	40.00	44.44	50.00	70.00	58.33
rbf	0.001	100	50.00	40.00	44.44	50.00	70.00	58.33
rbf	0.001	500	50.00	40.00	44.44	50.00	70.00	58.33
rbf	0.001	1000	50.00	40.00	44.44	50.00	70.00	58.33
rbf	0.01	10	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.01	100	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.01	500	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.01	1000	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.1	10	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.1	100	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.1	500	0.00	0.00	0.00	0.00	0.00	0.00
rbf	0.1	1000	0.00	0.00	0.00	0.00	0.00	0.00
rbf	1	10	0.00	0.00	0.00	0.00	0.00	0.00
rbf	1	100	0.00	0.00	0.00	0.00	0.00	0.00
rbf	1	500	0.00	0.00	0.00	0.00	0.00	0.00
rbf	1	1000	0.00	0.00	0.00	0.00	0.00	0.00

ตารางที่ 4 ตารางผลการวิเคราะห์ของข้อมูลชุดทดสอบที่สร้างตัวแบบโดยใช้วิธี LSTM

พารามิเตอร์			ชุดการเรียนรู้ข้อมูลที่ไม่สมดุล			ชุดการเรียนรู้ข้อมูลสมดุล		
Learning rate	Activation func.	Dropout	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)
0.1	tanh	0.1	0.00	0.00	0.00	20.59	70.00	31.82
0.1	tanh	0.2	0.00	0.00	0.00	2.01	80	3.91
0.1	tanh	0.3	0.00	0.00	0.00	23.33	70.00	35.00
0.1	tanh	0.4	0.00	0.00	0.00	21.05	80.00	33.33
0.1	tanh	0.5	0.00	0.00	0.00	17.50	70.00	28.00
0.1	relu	0.1	0.00	0.00	0.00	0.39	80	0.78

0.1	relu	0.2	0.00	0.00	0.00	0.29	60	0.59
0.1	relu	0.3	0.00	0.00	0.00	0.47	100	0.93
0.1	relu	0.4	0.00	0.00	0.00	2.12	50	4.07
0.1	relu	0.5	0.00	0.00	0.00	2.34	90	4.57
0.1	sigmoid	0.1	33.33	20.00	25.00	25.93	70	37.84
0.1	sigmoid	0.2	50.00	30.00	37.50	31.58	60	41.38
0.1	sigmoid	0.3	23.08	30.00	26.09	29.17	70	41.18
0.1	sigmoid	0.4	33.33	20.00	25.00	29.17	70	41.18
0.1	sigmoid	0.5	33.33	10.00	15.38	25	60	35.29
0.01	tanh	0.1	0.00	0.00	0.00	33.33	60.00	42.86
0.01	tanh	0.2	0.00	0.00	0.00	33.33	60.00	42.86
0.01	tanh	0.3	10.00	10.00	10.00	25.64	100.00	40.82
0.01	tanh	0.4	50.00	20.00	28.57	38.89	70.00	50.00
0.01	tanh	0.5	0.00	0.00	0.00	42.11	80.00	55.17
0.01	relu	0.1	0.00	0.00	0.00	33.33	70.00	45.16
0.01	relu	0.2	0.00	0.00	0.00	23.08	90.00	36.73
0.01	relu	0.3	0.00	0.00	0.00	40.91	90.00	56.25
0.01	relu	0.4	0.00	0.00	0.00	38.89	70.00	50.00
0.01	relu	0.5	33.33	20.00	25.00	40.00	80.00	53.33
0.01	sigmoid	0.1	57.14	40.00	47.06	33.33	60.00	42.86
0.01	sigmoid	0.2	50.00	40.00	44.44	44.44	80.00	57.14
0.01	sigmoid	0.3	50.00	40.00	44.44	36.84	70.00	48.28
0.01	sigmoid	0.4	50.00	50.00	50.00	42.11	80.00	55.17
0.01	sigmoid	0.5	50.00	50.00	50.00	38.89	70.00	50.00
0.001	tanh	0.1	40.00	20.00	26.67	35.29	60.00	44.44
0.001	tanh	0.2	40.00	20.00	26.67	38.89	70.00	50.00
0.001	tanh	0.3	50.00	40.00	44.44	52.94	90.00	66.67
0.001	tanh	0.4	50.00	50.00	50.00	41.18	70.00	51.85
0.001	tanh	0.5	50.00	50.00	50.00	42.11	80.00	55.17

0.001	relu	0.1	33.33	10.00	15.38	38.89	70.00	50.00
0.001	relu	0.2	0.00	0.00	0.00	38.89	70.00	50.00
0.001	relu	0.3	60.00	60.00	60.00	37.50	60.00	46.15
0.001	relu	0.4	60.00	30.00	40.00	42.11	80.00	55.17
0.001	relu	0.5	62.50	50.00	55.56	38.89	70.00	50.00
0.001	sigmoid	0.1	50.00	60.00	54.55	35.29	60.00	44.44
0.001	sigmoid	0.2	50.00	60.00	54.55	38.89	70.00	50.00
0.001	sigmoid	0.3	60.00	60.00	60.00	40.00	60.00	48.00
0.001	sigmoid	0.4	50.00	60.00	54.55	41.18	70.00	51.85
0.001	sigmoid	0.5	50.00	60.00	54.55	41.18	70.00	51.85
0.0001	tanh	0.1	50.00	40.00	44.44	37.50	60.00	46.15
0.0001	tanh	0.2	45.45	50.00	47.62	37.50	60.00	46.15
0.0001	tanh	0.3	50.00	60.00	54.55	38.89	70.00	50.00
0.0001	tanh	0.4	45.45	50.00	47.62	35.29	60.00	44.44
0.0001	tanh	0.5	38.46	50.00	43.48	38.89	70.00	50.00
0.0001	relu	0.1	55.56	50.00	52.63	37.50	60.00	46.15
0.0001	relu	0.2	46.15	60.00	52.17	41.18	70.00	51.85
0.0001	relu	0.3	54.55	60.00	57.14	41.18	70.00	51.85
0.0001	relu	0.4	33.33	10.00	15.38	38.89	70.00	50.00
0.0001	relu	0.5	54.55	60.00	57.14	38.89	70.00	50.00
0.0001	sigmoid	0.1	54.55	60.00	57.14	37.50	60.00	46.15
0.0001	sigmoid	0.2	60.00	60.00	60.00	35.29	60.00	44.44
0.0001	sigmoid	0.3	54.55	60.00	57.14	38.89	70.00	50.00
0.0001	sigmoid	0.4	60.00	60.00	60.00	41.18	70.00	51.85
0.0001	sigmoid	0.5	55.56	50.00	52.63	42.11	80.00	55.17
0.00001	tanh	0.1	50.00	50.00	50.00	38.89	70.00	50.00
0.00001	tanh	0.2	44.44	40.00	42.11	40.91	90.00	56.25
0.00001	tanh	0.3	40.00	20.00	26.67	40.00	80.00	53.33
0.00001	tanh	0.4	0.00	0.00	0.00	33.33	90.00	48.65

0.00001	tanh	0.5	0.00	0.00	0.00	11.11	100.00	20.00
0.00001	relu	0.1	100.00	10.00	18.18	8.79	80.00	15.84
0.00001	relu	0.2	57.14	40.00	47.06	7.30	100.00	13.61
0.00001	relu	0.3	45.45	50.00	47.62	1.44	100	2.84
0.00001	relu	0.4	25.00	10.00	14.29	1.22	100	2.41
0.00001	relu	0.5	0.00	0.00	0.00	1.15	100	2.27
0.00001	sigmoid	0.1	40.00	20.00	26.67	42.11	80.00	55.17
0.00001	sigmoid	0.2	60.00	30.00	40.00	42.11	80.00	55.17
0.00001	sigmoid	0.3	50.00	10.00	16.67	40.00	80.00	53.33
0.00001	sigmoid	0.4	50.00	10.00	16.67	36.00	90.00	51.43
0.00001	sigmoid	0.5	0.00	0.00	0.00	24.32	90.00	38.30

จากตารางที่ 3 และตารางที่ 4 สามารถสรุปผลที่ได้ค่า F1 สูงที่สุดของแต่ละตัวแบบที่สร้างจากข้อมูลการเรียนรู้ทั้ง 2 ชุดข้อมูลได้ดังตารางที่ 5

ตารางที่ 5 เปรียบเทียบประสิทธิภาพของตัวแบบที่สร้างจากชุดข้อมูลการเรียนรู้ทั้ง 2 ชุด โดยแสดงค่า F1 สูงที่สุด

ตัวแบบ	ข้อมูลที่ไม่สมดุล			ข้อมูลที่สมดุล		
	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)
SVM	62.50	50.00	55.56	63.64	70.00	66.67
LSTM	60.00	60.00	60.00	52.94	90.00	66.67

จากตารางที่ 5 จะพบว่าตัวแบบที่เรียนรู้จากชุดข้อมูลที่สมดุลให้ค่า F1 สูงกว่าตัวแบบที่เรียนรู้จากชุดข้อมูลที่ไม่สมดุล สำหรับการสร้างตัวแบบจากชุดข้อมูลการเรียนรู้ที่ไม่สมดุล ตัวแบบ LSTM จะมีค่า F1 และการเรียกคืนมากกว่าตัวแบบ SVM โดยมีค่าเท่ากับ 60.00% แต่ค่าความเที่ยงต่ำกว่าตัวแบบ SVM เล็กน้อยโดยมีค่าเท่ากับ 60.00% สำหรับชุดข้อมูลเรียนรู้ที่สมดุล ผลการศึกษาพบว่า ตัวแบบ SVM ให้ค่า F1 เท่ากับตัวแบบ LSTM โดยมีค่าเท่ากับ 66.67% และมีความเที่ยงสูงกว่าตัวแบบ LSTM ซึ่งมีค่าเท่ากับ 63.64% แต่มีการเรียกคืนต่ำกว่าตัวแบบ LSTM มีค่าเท่ากับ 70.00% โดยฟังก์ชันเคอร์เนลที่ให้ผลลัพธ์ดีที่สุดคือ polynomial 3rd degree เมื่อกำหนดค่า gamma เป็น 0.001, 0.01, 0.1 และ 1 กำหนดค่า C เป็น 10, 100, 500 และ 1,000 ส่วนตัวแบบ LSTM ที่สร้างจากชุดข้อมูลเรียนรู้ที่สมดุลมีประสิทธิภาพดีที่สุดเมื่อกำหนดค่า learning rate ของ optimizer ADAM เท่ากับ 0.001 ใช้ activation function ของ output layer เป็นฟังก์ชัน tanh และกำหนดค่า dropout เท่ากับ 0.3 โดยมีค่าการเรียกคืนสูงที่สุดเท่ากับ 90.00%

ค่าความเที่ยงต่ำที่สุดเท่ากับ 52.94% และค่า F1 เท่ากับ 66.67% แม้ว่าค่า F1 ของทั้งสองตัวแบบจะมีค่าเท่ากันแต่พบว่าตัวแบบ LSTM สามารถทำนายเสียงเปิดวาล์วได้ถูกต้องมากกว่าตัวแบบ SVM แต่ให้ค่าความเที่ยงน้อยกว่าตัวแบบ SVM เมื่อพิจารณาค่าความเที่ยงและค่าการเรียกคืนพร้อมกัน จึงได้ค่า F1 เท่ากัน

อย่างไรก็ตามจากตารางที่ 3 และตารางที่ 4 จะพบว่าตัวแบบที่มีประสิทธิภาพดีที่สุด สามารถกำหนดค่าพารามิเตอร์ที่แตกต่างกันได้ โดยจะให้ค่า F1 ค่าความเที่ยง และค่าการเรียกคืนเท่ากัน

สรุปผลการทดลอง

การวิจัยนี้สร้างตัวแบบการจำแนกเสียงเปิดวาล์วของอุปกรณ์คักไอน้ำ โดยใช้วิธีการจำแนกกลุ่ม 2 วิธี ได้แก่ วิธี SVM และวิธี LSTM และกำหนด activation function ของ output layer สามฟังก์ชัน ได้แก่ tanh relu และ sigmoid และใช้การแยกคุณลักษณะของเสียง 5 วิธี ได้แก่ อัตราค่าตัดศูนย์ spectral centroid spectral rolloff สัมประสิทธิ์เชปสตรัมบนความถี่แบบเมล และการแปลงฟูเรียร์ระยะสั้น นอกจากนี้ได้ทำการทดลองสร้างตัวแบบจำแนกกลุ่มโดยใช้ชุดข้อมูลเรียนรู้ 2 ชุด ได้แก่ ชุดข้อมูลเรียนรู้ที่ไม่สมดุลกับชุดข้อมูลเรียนรู้ที่สมดุลที่ได้ปรับแก้ด้วยวิธีผสมผสาน ผลการวิเคราะห์จำแนกกลุ่มโดยใช้ข้อมูลชุดทดสอบพบว่าตัวแบบ SVM และ LSTM ที่สร้างจากชุดข้อมูลเรียนรู้ที่สมดุลมีค่า F1 สูงกว่าการฝึกสร้างจากชุดข้อมูลเรียนรู้ที่ไม่สมดุล โดยตัวแบบทั้งสองให้ค่า F1 เท่ากันเท่ากับ 66.67% อย่างไรก็ตามพบว่าตัวแบบ SVM มีค่าความเที่ยงมากกว่า ในขณะที่ตัวแบบ LSTM มีค่าการเรียกคืนมากกว่า ดังนั้นตัวแบบ LSTM ทำนายผลเสียงเปิดวาล์วถูกต้องมากกว่าตัวแบบ SVM แต่ทำนายเสียงที่ไม่ใช่เสียงเปิดวาล์วผิดเป็นเสียงเปิดวาล์วมากกว่าตัวแบบ SVM เนื่องจากมีค่าความเที่ยงต่ำกว่า เมื่อนำค่าความเที่ยงและค่าการเรียกคืนมาคำนวณค่า F1 แล้วจึงได้ค่าเท่ากัน ทั้งนี้ตัวแบบ SVM ที่เหมาะสมคือตัวแบบที่ใช้ kernel function คือ polynimail 3rd degree และสามารถกำหนดค่า C ไม่น้อยกว่า 100 และ gamma ไม่น้อยกว่า 0.001 สำหรับตัวแบบ LSTM ที่เหมาะสมมีจำนวนโหนดใน Input Layer เท่ากับ 64 unit กำหนดค่า dropout เท่ากับ 0.3 และ output layer มีโหนด 1 unit กำหนด activation function ของ output layer เป็น tanh และกำหนดค่า learning rate เท่ากับ 0.001 อย่างไรก็ตามพบปัญหาในการวิจัยคือขณะบันทึกเสียงอุปกรณ์คักไอน้ำจะมีเสียงรบกวนจากเครื่องจักรในโรงงาน ควรศึกษาวิธีการบันทึกเสียงที่ลดเสียงรบกวนได้ หรือสร้างพื้นที่ที่ลดเสียงรบกวนจากเครื่องจักรรอบ ๆ อุปกรณ์คักไอน้ำ สำหรับการวิจัยในอนาคตอาจพิจารณาวิธีการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) สำหรับการจำแนกข้อมูลอาจทำให้ความแม่นยำของตัวแบบจำแนกเสียงสูงขึ้นและอาจใช้วิธีการปรับค่าพารามิเตอร์โดยวิธีการ Grid Search

กิตติกรรมประกาศ

ขอขอบพระคุณบัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่สำหรับทุนผู้ช่วยสอน/ผู้ช่วยวิจัย และขอขอบพระคุณคณะวิทยาศาสตร์และคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่ที่ให้การสนับสนุนการทำวิจัย

เอกสารอ้างอิง

- Ahmadizadeh, M. (2014). *An Introduction to Short-Time Fourier Transform (STFT)*.
<https://sharif.edu/~ahmadizadeh/courses/advstdyn/Short-Time%20Fourier%20Transform.pdf>
- Albon, C. (2017). *SVC Parameters When Using RBF Kernel*. https://chrisalbon.com/code/machine_learning/support_vector_machines/svc_parameters_using_rbf_kernel/
- Apaydin, H., Feizi, H., Sattari, M. T., Colak, M. S., Shamshirband, S., & Chau, K.W. (2020). Comparative Analysis of Recurrent Neural Network Architectures for Reservoir Inflow Forecasting. *Water*, 12(5), 18.
<https://doi.org/10.3390/w12051500>
- Brownlee, J. (2016). *Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras*.
<https://machinelearningmastery.com/sequence-classification-lstm-recurrent-neural-networks-python-keras/>
- Brownlee, J. (2017). *Dropout with LSTM Networks for Time Series Forecasting*.
<https://machinelearningmastery.com/use-dropout-lstm-networks-time-series-forecasting/>
- Brownlee, J. (2017). *How to Reshape Input Data for Long Short-Term Memory Networks in Keras*.
<https://machinelearningmastery.com/reshape-input-data-long-short-term-memory-networks-keras/>
- Burred, J. J., & Lerch, A. (2004). Hierarchical Automatic Audio Signal Classification. *Journal of The Audio Engineering Society*, 52(7/8), 724-739.
- Cohen, L. (1995). *Time-frequency analysis*. Englewood Cliffs, New Jersey: Prentice hall.
- Cui, Z., Ke, R., & Wang, Y. (2018). Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction. *arXiv preprint arXiv: 1801.02143*.
- Department of Industrial Works. (2010). *Safety Technology Bureau, Manual and maintenance of the boiler*.
<https://php.diw.go.th/safety/wp-content/uploads/2015/01/boiler2.pdf> (in Thai)
- Galván-Tejada, C. E., Galván-Tejada, J. I., Celaya-Padilla, J. M., Delgado-Contreras, J. R., Magallanes-Quintanar, R., Martinez-Fierro, M. L., Garza-Veloz, M. F., López-Hernández, Y., & Gamboa-Rosales, H. (2016). An analysis of audio features to develop a human activity recognition model using genetic algorithms, random forests, and neural networks. *Mobile Information Systems*, 2016.
- Giannakopoulos, T. (2015). pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLoS One*, 10(12), e0144610. <https://doi.org/10.1371/journal.pone.0144610>

- Hassan, S., Rafi, M., & Shaikh, M. S. (2011). Comparing SVM and naïve Bayes classifiers for text categorization with Wikitology as knowledge enrichment. In *2011 IEEE 14th International Multitopic Conference* (pp. 31-34). Karachi, Pakistan. <https://doi.org/10.1109/INMIC.2011.6151495>
- Junto, A. (2015). *Sound-based road traffic detection using artificial intelligent approach* [Unpublished master's thesis]. Suranaree University of Technology. (in Thai)
- Kaewtai, O. (2009). *Thai Classical Music Classification* [Unpublished Master's Thesis]. Silpakorn University. (in Thai)
- Kesornsit, W., Lorchirachoonkul, V., & Jitthavech, J. (2018). Imbalanced Data Problem Solving in Classification of Diabetes Patients. *KKU Research journal*, 18(3), 11. (in Thai)
- Kittinaradorn, C. (2020). *Neural Network Programming*. <https://guopai.github.io/ml-blog15.html> (in Thai)
- Kopanyapipat, R. (2015). *Noise robust speech recognition for thai language using N-based limabeam algorithm*. [Unpublished Master's Thesis]. Thammasat University. (in Thai)
- Li, T., Chen, X. R., Tang, H., & Xu, X. K. (2018). Identification of the Normal/Abnormal Heart Sounds Based on Energy Features and Xgboost. In *Biometric Recognition: 13th Chinese Conference, CCBR 2018, Urumqi, China, 11-12 August, 2018, Proceeding 13* (pp. 536-544). https://doi.org/10.1007/978-3-319-97909-0_57
- Lim, S., Jang S., Lim, J., & Ko, J. (2019). Classification of Snoring Sound Based on a Recurrent Neural Network. *Expert Systems with Applications*, 123, 237-245. <https://doi.org/10.1016/j.eswa.2019.01.020>
- Ministry of Energy. (2010). *Industrial Steam System*. https://www2.dede.go.th/bhrd/old/Download/file_handbook/Pre_Build/Build_12.pdf (in Thai)
- Ministry of Energy. (2011). *Steam Trap Management*. https://www2.dede.go.th/bhrd/old/Download/file_handbook/Pre_T_H/Heat_4.pdf (in Thai)
- Navlani, A. (2019). *Support Vector Machines with Scikit-learn*. <https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python#tuning>
- Pinmuang, N., & Thongkam, J. (2017). Classifying Thai opinions on online media using text mining. *J Sci Technol MSU*, 37(3), 372-379. (in Thai)
- TensorFlow. (2021). *tf.keras.metrics.binary_crossentropy*. https://www.tensorflow.org/api_docs/python/tf/keras/metrics/binary_crossentropy
- Thamsiri, D., & Meesad, P. (2011). Ensemble Data Classification Based on Decision Tree, Artificial Neuron Network and Support Vector Machine Optimized by Genetic Algorithm. *The Journal of KMUTNB*, 21(2), 293-303. (in Thai)

Tzanetakis, G., & Cook, P. (2002). Musical Genre Classification of Audio Signals. *IEEE Transactions on Speech and Audio Processing*, 10(5), 293-302. <https://doi.org/10.1109/TSA.2002.800560>

Zhu, S., & Chollet, F. (2020). *Working with RNNs*. https://keras.io/guides/working_with_rnn/