

บทความวิจัย (Research Article)



การวิเคราะห์เชิงปริมาณสารประกอบหลักในกาแฟด้วยเทคนิครามานสเปกโทรสโกปีร่วมกับการวิเคราะห์ข้อมูลโดยใช้การเรียนรู้ของเครื่อง

Quantitative Analysis of Major Compounds in Coffee Using Raman Spectroscopy Technique Combined with Machine Learning

กันตพงศ์ สุจริตพงษ์พันธ์¹, รุ่งโรจน์ จินตเมธาสวัสดิ์², มติ ห่อประทุม², และระพีพันธ์ แก้วอ่อน^{1*}

¹ ภาควิชาวิศวกรรมไฟฟ้า คณะ

วิศวกรรมศาสตร์และเทคโนโลยี

อุตสาหกรรม มหาวิทยาลัยศิลปากร

² ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และ

คอมพิวเตอร์แห่งชาติ, อุทยาน

วิทยาศาสตร์แห่งประเทศไทย

* ผู้รับผิดชอบบทความ

kaewon_r@su.ac.th

Received: 30 May 2024

Revised: 6 Jul 2024

Accepted: 16 Jul 2024

บทคัดย่อ

บทความนี้นำเสนอวิธีการวิเคราะห์สารคาเฟอีน (Caffeine) ซึ่งเป็นสารประกอบหลักในกาแฟ โดยจะทำการวิเคราะห์เชิงปริมาณของสารเคมีมาตรฐานด้วยเทคนิครามานสเปกโทรสโกปี (Raman Spectroscopy) ร่วมกับการวิเคราะห์ข้อมูลโดยใช้การเรียนรู้ของเครื่อง (Machine Learning) ในการทดลองนี้ได้ทำการเตรียมตัวอย่างคาเฟอีนมาตรฐานที่มีความเข้มข้น 500 ppm 1000 ppm 1500 ppm 2000 ppm 3000 ppm และ 4000 ppm ซึ่งปริมาณดังกล่าว เป็นปริมาณความเข้มข้นของคาเฟอีนที่สามารถพบได้ในตัวอย่างกาแฟจริง หลังจากนั้นจึงนำตัวอย่างมาวัดด้วยเทคนิครามานสเปกโทรสโกปี เพื่อได้ชุดข้อมูลสเปกตรัมรามาน อย่างไรก็ตาม ชุดข้อมูลสเปกตรัมรามานยังมีคุณภาพไม่เพียงพอ จึงจำเป็นต้องเพิ่มคุณภาพให้แก่ชุดข้อมูลรามานโดยการใช้เทคนิคการประมวลผลเบื้องต้น (Pre-Processing) ด้วยวิธีการปรับเส้นฐาน (Baseline Removal) การลบปรากฏการณ์รังสีคอสมิก (Cosmic ray Removal) การลดสัญญาณรบกวน (Smoothing Signal) การลบข้อมูลสเปกตรัมที่มีค่าผิดปกติ (Outlier Removal) และการปรับสัญญาณให้มีค่าอยู่ในช่วงที่เป็นมาตรฐาน (Amplitude Normalization) หลังจากนั้นจึงทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วน ด้วยอัตราส่วน 80:20 ชุดข้อมูลชุดแรกจำนวนร้อยละ 80 ของจำนวนข้อมูลทั้งหมด (Training Dataset) จะถูกนำมาใช้เพื่อฝึกสอนแบบจำลอง (Model Training) และชุดข้อมูลที่เหลือจำนวนร้อยละ 20 ของจำนวนข้อมูลทั้งหมด (Validation Dataset) จะถูกใช้ในการทดสอบและปรับพารามิเตอร์ให้แบบจำลองเพื่อเพิ่มประสิทธิภาพในการทำนายปริมาณของสารคาเฟอีน งานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้งหมด 18 แบบจำลอง ซึ่งใช้ ค่าความผิดพลาดเฉลี่ยสมบูรณ์ (Mean Absolute Error: MAE) ค่ารากที่สองเฉลี่ยสมบูรณ์ (Root Mean Square Error: RMSE) และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination: R-Square) เป็นตัวชี้วัดในการประเมินประสิทธิภาพของแบบจำลอง โดยแบบจำลองที่มีประสิทธิภาพสูงที่สุดผ่านตัวชี้วัดนี้คือ Extra Trees Regressor ซึ่งมีค่าความผิดพลาดเฉลี่ยสมบูรณ์ ค่ารากที่สองเฉลี่ยสมบูรณ์ และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ เป็น 710 ppm 914.05 ppm และ 0.34 ตามลำดับ เนื่องจากสัญญาณรามานมีความแปรปรวนของสัญญาณจากปัจจัยอื่นๆสูงโดยเฉพาะ การเกิดปรากฏการณ์ ฟลูออเรสเซนซ์ อย่างไรก็ตามวิธีการและผลการทดลองที่ได้จากงานนี้ แสดงแนวโน้มของการเปลี่ยนแปลงสัญญาณรามานเมื่อความเข้มข้นของสารคาเฟอีนมาตรฐานเปลี่ยนแปลงไป จึงแสดงให้เห็นถึงศักยภาพที่สามารถนำไปเป็นกระบวนการไปต่อยอดเพื่อเป็นแนวทางในการพัฒนาสารประกอบหลักใช้ในการหาปริมาณสารคาเฟอีนในกาแฟได้

คำสำคัญ: กาแฟ คาเฟอีน รามานสเปกโทรสโกปี การเรียนรู้ของเครื่อง

Abstract

This article presents a method for analyzing caffeine, which is a primary compound in coffee, by quantitatively analyzing standard chemical substances using Raman Spectroscopy technique together with data analysis using Machine Learning. In this experiment, standard caffeine samples were prepared at concentrations of 500 ppm, 1000 ppm, 1500 ppm, 2000 ppm, 3000 ppm, and 4000 ppm. Subsequently, these concentrations are comparable to those found in actual coffee samples. These samples were measured using Raman Spectroscopy technique to obtain a set of Raman spectra data. However, the quality of the Raman spectra data was found to be insufficient, necessitating enhancement of the data quality through preprocessing techniques such as Baseline Removal, Cosmic ray Removal, Smoothing Signal, Outlier Removal, and Amplitude Normalization. Following this, the dataset containing 80% of the total data (Training Dataset), was used to train model, while the remaining 20% (Validation Dataset) was used for testing and parameter adjustment to improve the model's prediction efficiency for caffeine quantitates. This research compared the performance of all 18 models using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Multiple Determination (R-Square) as performance indicators. The model with highest performance according to these indicators was the Extra Tree Regressor, with MAE, RMSE, and R-square values of 710 ppm, 914.05 ppm, and 0.34 respectively. The methods and experimental results from this study can be further developed for determining caffeine quantities in coffee.

Keywords: Wireless network, Internet of Things, Monitoring system

1. บทนำ

กาแฟเป็นหนึ่งในเครื่องดื่มที่นิยมบริโภคมากที่สุดในโลก เช่นเดียวกับชา และยังมีกาปลูกกาแฟในมากกว่า 50 ประเทศทั่วโลก จึงนับได้ว่าเป็นพืชเศรษฐกิจที่สำคัญในหลายประเทศ แม้ว่ากาแฟจะประกอบด้วยสารประกอบเคมีมากกว่า 1000 ชนิด แต่หนึ่งในสารประกอบหลักของกาแฟ คือ คาเฟอีน (Caffeine) ซึ่งมีผลต่อการกำหนดรสชาติความขมของกาแฟ ช่วยในการออกฤทธิ์กระตุ้นระบบประสาทส่วนกลาง ทำให้ผู้บริโภคมีการตื่นตัวตลอดเวลา และลดความเสี่ยงการเกิดโรคสมองเสื่อม [1] โดยปริมาณคาเฟอีนในกาแฟมีปริมาณ 0.6 ถึง 4% โดยมวลแห้งของกาแฟ และปริมาณคาเฟอีนจะขึ้นอยู่กับกระบวนการผลิต ซึ่งเป็นปัจจัยที่สำคัญในการตัดสินใจเลือกซื้อหรือการบริโภคกาแฟ ในปัจจุบันการตรวจวัดปริมาณสารสำคัญภายในกาแฟที่มีความแม่นยำ จำเป็นต้องใช้เทคนิคทางเคมีมาตรฐาน [2-3] เช่น Gas Chromatography Mass-Spectrometer (GC-MS) ซึ่งมีความแม่นยำในการตรวจวัดสูงแต่มีข้อจำกัดในเรื่องของระยะเวลาและราคาในการตรวจวัด ดังนั้นจึงมีความจำเป็นต้องพัฒนาเทคนิคการตรวจวัดปริมาณสารสำคัญในกาแฟ ที่มีความรวดเร็ว และราคาต่ำกว่าเทคนิคมาตรฐานที่มีอยู่ งานวิจัยนี้จึงได้มีการใช้เทคนิครามานสเปกโทรสโกปี ซึ่งที่ลดข้อจำกัดเหล่านั้น

รามานสเปกโทรสโกปี (Raman Spectroscopy) เป็นเทคนิคการวัดเชิงแสงแบบไม่ทำลายตัวอย่าง (Non-destructive) โดยอาศัยปรากฏการณ์การกระเจิงแสงรามาน (Raman scattering) โดยการยิงเลเซอร์ความเข้มสูงเข้าสู่สารตัวอย่าง ทำให้โมเลกุลของสารตัวอย่างถูกกระตุ้น โดยที่สารตัวอย่างแต่ละประเภทจะมีความไว (Sensitivity) ต่อแสงเลเซอร์ที่แตกต่างกัน ทำให้สัญญาณรามานที่กระเจิงออกมาจากสารตัวอย่าง หรือสเปกตรัม (Spectrum) มีการเปลี่ยนแปลงของความถี่และความเข้มในรูปแบบที่แตกต่างกัน จากหลักการดังกล่าวทำให้ได้ข้อมูลสเปกตรัมที่สามารถระบุประเภทของสารตัวอย่างได้อย่างรวดเร็ว และไม่ทำลายตัวอย่าง ส่งผลให้มีการใช้หลักการดังกล่าวในการวิเคราะห์เชิงคุณภาพของสารเคมีและวัสดุกันอย่างกว้างขวาง [4-5] อย่างไรก็ตาม การใช้เทคนิครามานสเปกโทรสโกปีในการวิเคราะห์เชิงปริมาณยังมีความท้าทายอยู่มากในปัจจุบัน เนื่องจากปริมาณของสารตัวอย่างที่เปลี่ยนไป จะไม่ส่งผลให้เกิด

การเปลี่ยนแปลงของสเปกตรัมอย่างชัดเจน ดังนั้นจึงจำเป็นต้องใช้เทคนิคการวิเคราะห์ข้อมูลโดยใช้การเรียนรู้ของเครื่อง (Machine Learning)

ปัจจุบัน การเรียนรู้ของเครื่อง (Machine Learning) ได้มีการนำมาใช้ในการวิเคราะห์ข้อมูลรามานสเปกโทรสโกปีทั้งการวิเคราะห์เชิงปริมาณ และเชิงคุณภาพกันอย่างแพร่หลาย [6-7] โดยการป้อนข้อมูลสเปกตรัมลงในอัลกอริทึมการเรียนรู้แบบจำลองการคำนวณสามารถสร้างขึ้นเพื่อส่งผลลัพธ์การตัดสินใจจากข้อมูลอินพุตเริ่มต้นโดยอัตโนมัติและรวดเร็ว การเรียนรู้ของเครื่องร่วมกับเทคนิครามานสเปกโทรสโกปี สามารถใช้ในการวิเคราะห์จัดกลุ่มสายพันธุ์ของกาแฟได้อย่างมีประสิทธิภาพ [8-9] และได้นำการเรียนรู้ของเครื่องประยุกต์ใช้ในการจำแนกประเภทของเครื่องดื่มกาแฟโดยใช้ข้อมูลสเปกตรัมของรามาน [10-11] อย่างไรก็ตามในปัจจุบันการวิเคราะห์เชิงคุณภาพโดยใช้ข้อมูลสเปกตรัมของรามานสามารถวิเคราะห์ได้อย่างมีประสิทธิภาพ แต่การวิเคราะห์เชิงปริมาณโดยใช้ข้อมูลจากสเปกตรัมของรามานยังขาดความแม่นยำ โดยในงานวิจัยนี้จะเป็นการพัฒนาวิธีการตรวจวัดปริมาณของสารคาเฟอีนมาตรฐานโดยใช้เทคนิครามานสเปกโทรสโกปี ซึ่งมีศักยภาพในการต่อยอดเพื่อตรวจวัดเชิงปริมาณของสารคาเฟอีนที่อยู่ในกาแฟ

2. วิธีการและทฤษฎีที่เกี่ยวข้อง

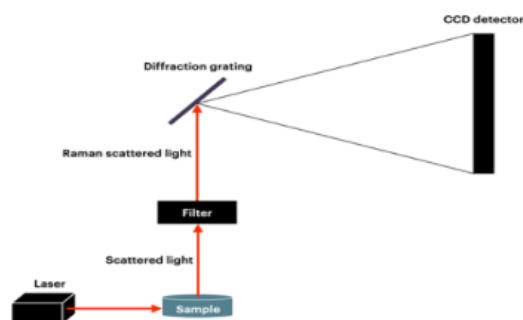
2.1 การเตรียมตัวอย่าง (Sample Preparation)

ในงานวิจัยนี้จะมีการเตรียมตัวอย่างสารคาเฟอีน ให้มีความเข้มข้นต่างๆ ตามต้องการ โดยสารคาเฟอีนที่ใช้จะเป็นสารแบบมาตรฐานที่มีความบริสุทธิ์ 99.96% (Loba Chemie Pvt Ltd, Mumbai, India) และจะถูกผสมกับตัวทำละลาย โดยตัวทำละลายจะประกอบด้วยเอทานอล (Ethanol) ผสมกับน้ำเกรด HPLC ในอัตราส่วน 1:1 ในการผสมสารคาเฟอีนมาตรฐานกับตัวทำละลาย จะผสมโดยใช้ปริมาณสาร 500 1000 1500 2000 3000 และ 4000 กรัม กับปริมาณตัวทำละลาย 1000 มิลลิลิตร ในหลอดทดลอง หลังจากนั้นจะทำการเขย่า เป็นเวลา 30 วินาที และทำการทำให้เกิดกระแสนวน (Vortex) เพื่อให้สารคาเฟอีนมาตรฐานละลายในตัวทำละลายได้ดีขึ้น ทำให้ได้สารละลาย

คาเฟอีนมาตรฐานที่มีความเข้มข้น 500 1000 1500 2000 3000 และ 4000 ppm ตามลำดับ

2.2 การวัดด้วยเทคนิครามานสเปกโทรสโกปี (Raman Spectroscopy)

เทคนิครามานสเปกโทรสโกปี (Raman Spectroscopy) ดังแสดงในรูปที่ 1 เป็นเทคนิคที่อาศัยปรากฏการณ์การกระเจิงแสงรามาน (Raman scatter) เกิดจากการยิงแสงเลเซอร์ความเข้มสูงเข้าสู่วัสดุ ทำให้โมเลกุลเกิดการกระตุ้นจนเกิดการสั่นของโมเลกุล และทำให้เกิดการกระเจิงแสง ในกรณีที่แสงที่กระเจิงออกมามีความยาวคลื่นเท่ากับความยาวคลื่นของแสงเลเซอร์ แสงส่วนนั้นจะถูกเรียกว่า Rayleigh Scattering แต่หากแสงที่กระเจิงออกมาเกิดการเปลี่ยนแปลงความยาวคลื่น แสงส่วนนั้นจะถูกเรียกว่า สัญญาณรามาน (Raman Scattering) และสามารถตรวจวัดได้โดยใช้ตัววัดแบบ CCD (Charge Coupled Device Detector) และเนื่องจากสัญญาณ Raman ที่กระเจิงออกมา นั้นสามารถมีความถี่ได้หลายค่า จึงจำเป็นต้องใช้อุปกรณ์ในการแยกความยาวคลื่น (Grating) ของสัญญาณรามาน เพื่อให้สามารถเก็บข้อมูลในรูปแบบสเปกตรัมได้



รูปที่ 1 หลักการวัดสารตัวอย่างด้วยเทคนิครามานสเปกโทรสโกปี (Raman Spectroscopy)

สำหรับตัวอย่างของสารที่เตรียมแล้ว จะถูกนำมาหยดลงบนอลูมินัมฟอยล์และปล่อยให้แห้งเป็นเวลา 30 นาที หลังจากนั้นจึงทำการเก็บข้อมูลสเปกตรัมรามานด้วยเครื่อง Raman Spectroscopy (Renishaw New Mills, Gloucestershire, UK) โดยใช้เลเซอร์ความยาวคลื่น 785 nm กำลังเลเซอร์ 250 mW ทำการกระตุ้นสารตัวอย่าง และทำการเก็บสเปกตรัมรามานที่ได้

ในช่วงเลขคลื่น 500 ถึง 1600 cm^{-1} ที่กำลังขยายภาพ 50 เท่า และความละเอียดของสเปกตรัมที่ 4 cm^{-1} สำหรับแต่ละตัวอย่าง จะทำการเก็บทั้งหมด 50 สเปกตรัม โดยแต่ละสเปกตรัมจะใช้ เวลาในการเก็บประมาณ 10 วินาที และจะทำการเก็บข้อมูล ทั้งหมด 3 ตัวอย่างต่อหนึ่งความเข้มข้น

2.3 การประมวลผลสเปกตรัมเบื้องต้น (Spectral Pre-Processing)

การประมวลผลสเปกตรัมเบื้องต้น เป็นวิธีการลดสิ่งรบกวน, ข้อมูลที่เกิดจากข้อผิดพลาดของอุปกรณ์ หรือข้อมูลที่ไม่เกี่ยวข้อง ออกจากข้อมูลสเปกตรัม โดยมีเป้าหมายเพื่อปรับปรุงคุณภาพ โดยรวมของชุดข้อมูล อย่างไรก็ตามเทคนิคเหล่านี้ไม่ได้ปรับปรุง ประสิทธิภาพของชุดข้อมูลเสมอไป แต่อาจจะเป็นการลบข้อมูลที่ สำคัญ ซึ่งทำให้ประสิทธิภาพของชุดข้อมูลลดลง ดังนั้นการเลือก วิธีการประมวลผลสเปกตรัมเบื้องต้น จึงจำเป็นต้องเลือกใช้ให้มีความเหมาะสมกับชุดข้อมูล สำหรับกระบวนการประมวลผล เบื้องต้นที่เหมาะสมกับข้อมูลสเปกตรัม มีขั้นตอนดังนี้

2.3.1 การปรับฐานของสเปกตรัม (Baseline Removal)

เนื่องจากสเปกตรัมรามานมักจะมีสัญญาณรบกวนพื้นหลัง (Background Noise) ซึ่งมักเกิดจากการกระเจิงของแสงรามาน ทำให้ค่าสัญญาณกระเจิงที่วัดได้เพิ่มขึ้นตามความถี่ของแสงรามาน [12] กระบวนการปรับฐานของสเปกตรัมจะช่วยในการลด สัญญาณกระเจิงที่เกิดขึ้น โดยในงานวิจัยนี้จะใช้อัลกอริทึมปรับ ฐานของสเปกตรัมแบบพหุนยกกำลังสอง (Second Order Polynomial Algorithm)

2.3.2 การลบรังสีคอสมิก (Cosmic Ray Removal)

เนื่องจากการวัดสัญญาณรามานจะใช้อุปกรณ์ Charge Couple Device Detector (CCD Detector) ในการตรวจวัด ซึ่งอุปกรณ์ ดังกล่าวมีความไวต่อรังสีคอสมิก ส่งผลให้เกิดการพุ่งขึ้นอย่าง รวดเร็วในสัญญาณ (Spikes) ซึ่งมีลักษณะคล้ายกับพีค (Peaks) ที่แสดงถึงคุณลักษณะที่สำคัญของสารตัวอย่าง ทำให้เกิดความ สับสนและอาจเกิดความผิดพลาดในการทำนายได้ ดังนั้น งานวิจัยนี้จึงใช้การหาค่าผลต่างระหว่างสัญญาณสองตำแหน่งที่ ติดกัน และหากมีค่าสัญญาณที่มีความต่างมากกว่าเกณฑ์ (Threshold) ที่กำหนด ก็แสดงว่าสัญญาณในบริเวณนั้นอาจเป็น Spikes และความเข้มสัญญาณ ณ ตำแหน่งดังกล่าวจะถูกแทนที่

ด้วยความเข้มสัญญาณเฉลี่ยภายในระยะหน้าต่าง (Window Size) โดยค่า Threshold และ Window Size ถูกกำหนดให้เป็น 1500 และ 5 ตามลำดับ

2.3.3 การกำจัดข้อมูลผิดปกติ (Outlier Removal)

เนื่องจากในระหว่างการวัดด้วยเทคนิครามานสเปกโทรสโกปี จะต้องทำการเลือกตำแหน่งการวัดบนสารตัวอย่าง ซึ่งความไม่ เป็นเนื้อเดียวกันของสารตัวอย่าง (Heterogeneity) ทำให้ในบาง บริเวณของการวัดมีค่าสัญญาณรามานที่ไม่ถูกต้อง ส่งผลให้ชุด ข้อมูลเกิดความแปรปรวน และคุณภาพของชุดข้อมูลลดลง ดังนั้น ในงานวิจัยนี้จึงต้องทำการกำจัดข้อมูลผิดปกติ (Outlier Removal) โดยจะใช้วิธีการหาอัตราส่วนสัญญาณต่อสัญญาณ รบกวน (Signal-To-Noise Ratio หรือ SNR) และทำการกำจัด สเปกตรัมที่มีค่า SNR สูงกว่าเกณฑ์ โดยตั้งค่าเกณฑ์เท่ากับ 0.5

2.3.4 การปรับสัญญาณให้เรียบ (Signal Smoothing)

หลังจากที่ทำการลบข้อมูลผิดปกติออกจากชุดข้อมูล จึงทำการ ปรับสัญญาณให้เรียบ (Signal Smoothing) เพื่อลดสัญญาณ รบกวนที่เกิดขึ้นจากความรบกวนสะสมในตัวขยายประจุ (Charge Amplifiers) ของ CCD Detector [13] โดยในงานวิจัยนี้จะทำ การปรับสัญญาณให้เรียบโดยใช้อัลกอริทึม Savitzky and Golang Smoothing เพื่อลดสัญญาณรบกวนโดยการตั้งค่าลำดับ พหุนาม (Polynomial Order) และขนาดหน้าต่างตัวกรอง (Filter Window Size) เป็น 2 และ 15 ตามลำดับ

2.3.5 การทำให้สเปกตรัมเป็นมาตรฐาน (Normalization)

เนื่องจากการที่จะนำสัญญาณรามานไปใช้ในการวิเคราะห์โดยใช้ การเรียนรู้ของเครื่องจะขึ้นอยู่กับความเข้มของสัญญาณรามาน โดยที่ปริมาณการกระเจิงของสาร และการเลือกตำแหน่งการวัด ส่งผลต่อความเข้มของสัญญาณและทำให้สเปกตรัมรามานของชุด ข้อมูลมีความไม่คงเส้นคงวา (Inconsistency) โดยในงานวิจัยนี้ ได้ทำการปรับค่าสัญญาณของสเปกตรัม ด้วยการคูณด้วยค่าคงที่ เพื่อให้ตำแหน่งที่มีความเข้มของสัญญาณสูงที่สุดมีค่าเท่ากันในทุกสเปกตรัม

2.4 การเรียนรู้ของเครื่อง (Machine Learning)

ในงานนี้ ได้ประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยมุ่งเน้นไปที่การเลือกใช้แบบจำลองการถดถอย

(Regression Model) แบบต่างๆที่ใช้กันอย่างแพร่หลาย มาทำการเปรียบเทียบประสิทธิภาพ โดยในงานวิจัยนี้ได้เปรียบเทียบประสิทธิภาพแบบจำลองที่มีการวิเคราะห์ข้อมูลที่แตกต่างกันทั้งหมด 5 กลุ่ม ประกอบด้วย 1) แบบจำลองในกลุ่ม Linear Models 2) แบบจำลองในกลุ่ม Support Vector Machine 3) แบบจำลองในกลุ่ม Naive Bayes 4) แบบจำลองในกลุ่ม Decision Trees และ 5) แบบจำลองในกลุ่ม Ensemble Methods ซึ่งทั้ง 5 กลุ่มนี้จะประกอบด้วยแบบจำลองทั้งหมด 18 แบบจำลอง โดยประเมินแบบจำลองผ่านตัวชี้วัดดังนี้

- ค่าความผิดพลาดเฉลี่ยสมบูรณ์ (Mean Absolute Error: MAE):

$$MAE = \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

- ค่ารากที่สองเฉลี่ยสมบูรณ์ (Root Mean Square Error: RMSE):

$$RMSE = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

- ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination: R²):

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (3)$$

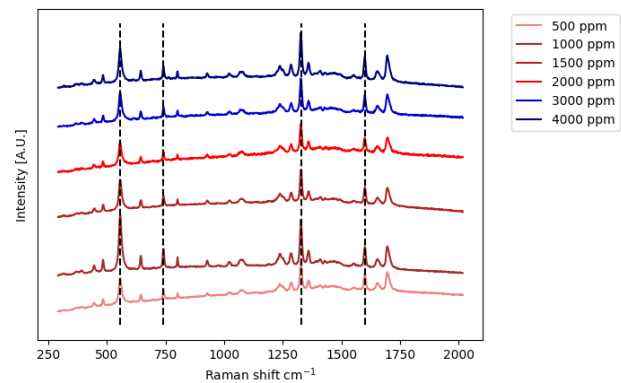
นอกจากนี้การวิเคราะห์ข้อมูลยังรวมถึงการตรวจสอบผลลัพธ์ของการทำนายของแบบจำลองผ่าน Prediction Plot ที่แสดงถึงความแม่นยำของแบบจำลองที่ได้ ทั้งนี้อาจจะมีแบบจำลองอื่นที่มีความสามารถในการเรียนรู้ข้อมูลที่มีความซับซ้อนกว่านี้ได้เช่น Multi-Layer Perceptron ที่เป็นอัลกอริทึมแบบโครงข่ายประสาทเทียม (Neural Network) มีการวิเคราะห์แบบไม่เป็นเชิงเส้น (Non-linear) หรือ แบบจำลอง LSTM แต่เนื่องจากจำนวนข้อมูลที่จำกัดทำให้มีโอกาสที่แบบจำลองจะเกิดการมีความเฉพาะเจาะจงกับชุดข้อมูลฝึกสอนมากเกินไป

(overfitting) และข้อมูลการวัดแต่ละจุดไม่ได้มีความเชื่อมโยงกัน (Independence) จึงทำให้แบบจำลองลักษณะไม่เหมาะสมกับชุดข้อมูลสเปกตรัมรามานของสารคาเฟอีนมาตรฐาน

3. ผลการทดลอง

3.1 การวัดสัญญาณรามาน (Raman Measurements)

รูปที่ 2 แสดงตัวอย่างสเปกตรัมรามานที่วัดได้ ของสารคาเฟอีนมาตรฐานที่ความเข้มข้นแตกต่างกัน โดยไม่ผ่านการประมวลผลเบื้องต้น จากรูปจะสังเกตได้ว่า สเปกตรัมของสารคาเฟอีนมาตรฐานนั้นจะมีลักษณะของสัญญาณรามานที่เหมือนกันในทุกความเข้มข้น และความเข้มของสัญญาณรามานไม่ได้สัมพันธ์กับความเข้มข้นของคาเฟอีนแบบเชิงเส้น



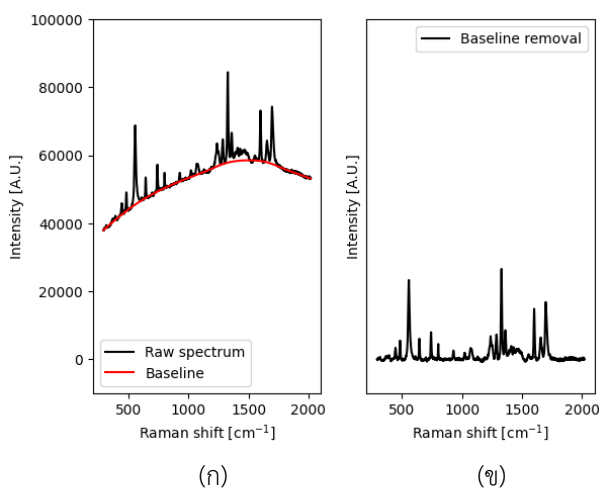
รูปที่ 2 สเปกตรัมรามานของสารคาเฟอีนมาตรฐานที่มีความเข้มข้นแตกต่างกัน

นอกจากนี้ สเปกตรัมรามานที่วัดได้มีระดับสัญญาณที่แตกต่างกัน และตำแหน่งการวัดที่แตกต่างกันก็มีผลต่อสเปกตรัมที่วัดได้ ทำให้ความเข้มของสเปกตรัมคาดเคลื่อน และทำให้การใช้ข้อมูลสเปกตรัมที่วัดได้โดยตรงมาทำการวิเคราะห์และประมวลผลไม่มีประสิทธิภาพ ดังนั้นจึงต้องทำการประมวลผลสัญญาณสเปกตรัมเบื้องต้น

3.2 การประมวลผลสัญญาณสเปกตรัมเบื้องต้น (Pre-processing)

3.2.1 การปรับฐานของสเปกตรัม (Baseline Removal)

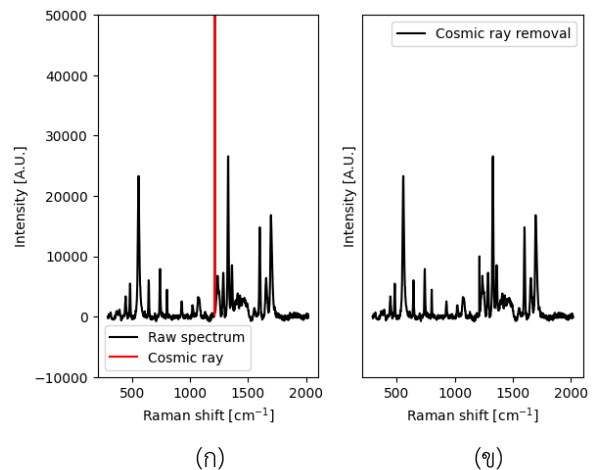
ตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐาน ก่อนการปรับฐานของสเปกตรัมแสดงในรูปที่ 3 (ก) จากรูปสังเกตได้ว่าแต่ละสเปกตรัมมีรูปแบบและระดับของฐานที่ต่างกัน ซึ่งระดับฐานที่ยกสูงขึ้นเกิดจากสัญญาณพื้นหลังจากการกระเจิงแสง และความถี่สูงขึ้นก็จะทำให้ระดับของฐานสูงขึ้นตามไปด้วย รูปที่ 3 (ข) แสดงตัวอย่างของสเปกตรัมของรามานของสารคาเฟอีนมาตรฐานที่ผ่านการปรับระดับของฐานแล้ว ทำให้ฐานของสเปกตรัมทั้งหมดอยู่ในระดับที่ใกล้เคียงกัน



รูปที่ 3 แสดงการเปรียบเทียบตัวอย่างสเปกตรัมของสารตัวอย่างที่วัดได้จากสารคาเฟอีนที่มีปริมาณ 500 และ 4000 ppm ตามลำดับ โดย (ก) และ (ข) แสดงสเปกตรัมก่อนและหลังการปรับฐานตามลำดับ

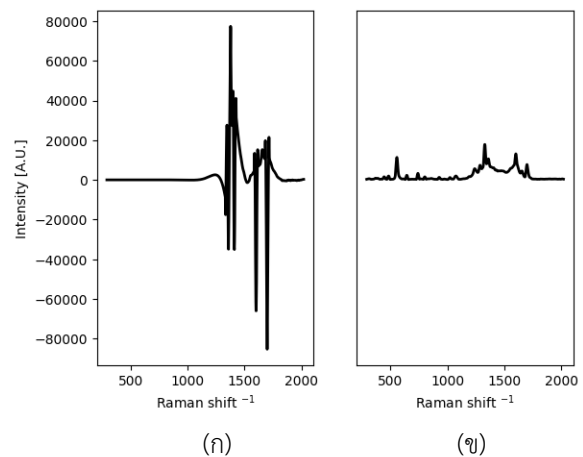
3.2.2 การลบรังสีคอสมิก (Cosmic Ray Removal)

ตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐานที่เกิดปรากฏการณ์รังสีคอสมิกแสดงในรูปที่ 4 (ก) เนื่องจาก CCD detector มีความไวต่อรังสีคอสมิก จึงทำให้เกิดการพุ่งขึ้นของสัญญาณอย่างรวดเร็ว ทำให้สเปกตรัมในบางตำแหน่งจะปรากฏให้เห็นค่าสัญญาณที่สูงผิดปกติ หรือที่เรียกว่า Spikes ซึ่งมีลักษณะคล้ายกับพีค (Peaks) ของสัญญาณรามาน แต่มีขนาดที่แคบกว่า รูปที่ 4 (ข) แสดงตัวอย่างของสัญญาณที่ทำการลบปรากฏการณ์รังสีคอสมิก จากรูปจะสังเกตเห็นว่า Spikes จะถูกแทนที่ด้วยค่าเฉลี่ยของสัญญาณในระยะหน้าต่าง (Window Size ที่กำหนดให้)



รูปที่ 4 ตัวอย่างสเปกตรัมรามานเปรียบเทียบระหว่าง

(ก) สเปกตรัมที่ปรากฏให้เห็น Spikes แสดงถึงปรากฏการณ์ของรังสีคอสมิก และ (ข) สเปกตรัมที่ได้หลังจากการลบ Spikes ออกไป



รูปที่ 5 ตัวอย่างสเปกตรัมรามานเปรียบเทียบระหว่าง (ก)

สเปกตรัมที่มีสัญญาณรบกวนสูงกว่าเกณฑ์ และ (ข) สเปกตรัมที่มีสัญญาณรบกวนน้อยกว่าเกณฑ์

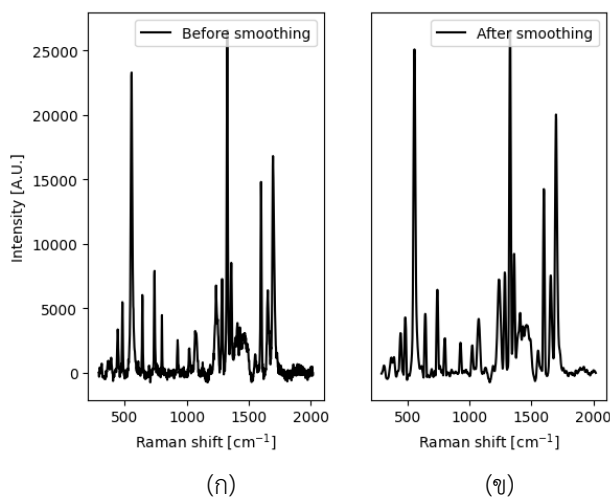
3.2.3 การกำจัดข้อมูลที่ผิดปกติ (Outlier Removal)

ตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐานที่มีค่าสัญญาณรบกวนสูงกว่าเกณฑ์แสดงในรูปที่ 5 (ก) ซึ่งเกิดจากตัวอย่างมีความไม่เป็นเนื้อเดียวกัน (Heterogeneous) ทำให้อาจเกิดการเลือกตำแหน่งการวัดที่ไม่เหมาะสม และทำให้สัญญาณรามานที่วัดได้ไม่แสดงถึงคุณลักษณะที่ถูกต้องของสารตัวอย่าง รูปที่ 5 (ข) แสดงตัวอย่างการลบสเปกตรัมรามานที่ผิดปกติออกจากชุด

ข้อมูลเพื่อเพิ่มคุณภาพให้แก่ชุดข้อมูล และทำให้ชุดข้อมูลปราศจากสเปกตรัมรบกวนที่มีสัญญาณรบกวนสูงกว่าเกณฑ์ที่กำหนด

3.2.4 การปรับสัญญาณให้เรียบ (Signal Smoothing)

ตัวอย่างสเปกตรัมของสารคาเฟอีนมาตรฐานก่อนปรับสัญญาณให้เรียบแสดงในรูปที่ 6 (ก) ซึ่งสังเกตได้ว่าสเปกตรัมยังมีสัญญาณรบกวนสูง รูปที่ 6 (ข) แสดงสเปกตรัมที่ผ่านการปรับสัญญาณให้เรียบ โดยใช้อัลกอริทึม Savitzky and Golang Smoothing ซึ่งจะเห็นได้ว่าสเปกตรัมมีสัญญาณรบกวนที่ลดลง และมีการแกว่งของสัญญาณ (Oscillation) ที่น้อยลงอย่างชัดเจน

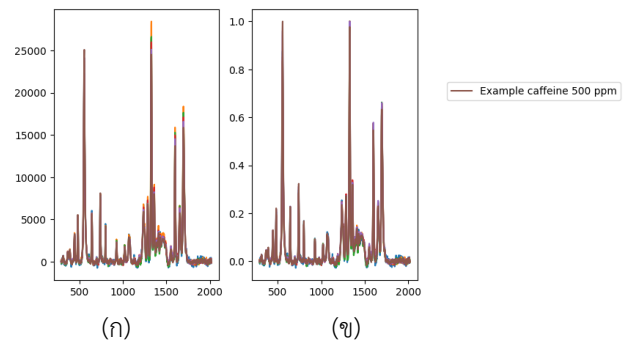


รูปที่ 6 ตัวอย่างสเปกตรัมรามานเปรียบเทียบระหว่าง (ก) สเปกตรัมก่อนการปรับสัญญาณให้เรียบ และ (ข) สเปกตรัมหลังการปรับสัญญาณให้เรียบ

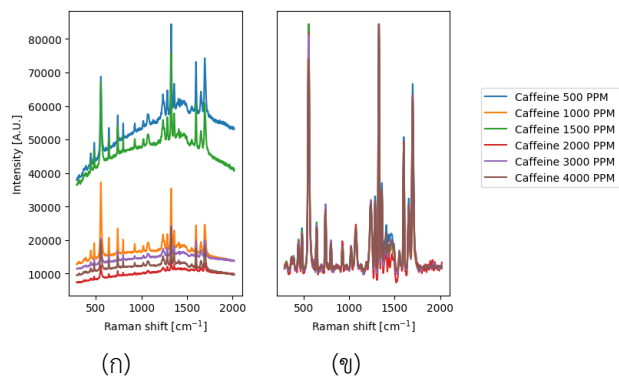
3.2.5 การทำให้สเปกตรัมเป็นมาตรฐาน (Normalization)

ตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐานก่อนการทำให้สเปกตรัมเป็นมาตรฐาน แสดงในรูปที่ 7 (ก) จากรูปสังเกตได้ว่า สำหรับตัวอย่างที่ต่างกันแต่มีความเข้มข้นของสารคาเฟอีนมาตรฐานเท่ากัน สัญญาณสเปกตรัมที่วัดได้อาจมีความเข้มไม่เท่ากัน ดังนั้นจึงไม่สามารถนำค่าความเข้มของสัญญาณที่วัดได้ไปใช้ในการวิเคราะห์ได้โดยตรง รูปที่ 7 (ข) แสดงตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐานที่ผ่านการทำให้สเปกตรัมเป็นมาตรฐาน ซึ่งสังเกตได้ว่าทุกสเปกตรัมจะมีค่าความเข้มสัญญาณ

อยู่ในช่วง 0 ถึง 1 A.U. และค่าสูงสุดของสเปกตรัมจะอยู่ในตำแหน่งเดียวกันเสมอ ทำให้สามารถนำสเปกตรัมเหล่านี้ไปวิเคราะห์ได้ต่อไป



รูปที่ 7 ตัวอย่างสเปกตรัมรามานเปรียบเทียบระหว่างสเปกตรัมที่ได้ (ก) ก่อนการทำให้สเปกตรัมเป็นมาตรฐาน และ (ข) หลังการทำให้สเปกตรัมเป็นมาตรฐาน



รูปที่ 8 แสดงการเปรียบเทียบตัวอย่างสเปกตรัมรามานที่ความเข้มข้นแตกต่างกัน ระหว่างสเปกตรัมที่ได้ (ก) ก่อนการประมวลผลเบื้องต้น และ (ข) หลังการประมวลผลเบื้องต้น

3.2.6 ผลลัพธ์สุดท้ายจากการประมวลผลสัญญาณสเปกตรัมเบื้องต้น

รูปที่ 8 (ก) แสดงตัวอย่างสเปกตรัมรามานของสารคาเฟอีนมาตรฐานที่ความเข้มข้น 500 ppm 1000 ppm 1500 ppm 2000 ppm 3000 ppm และ 4000 ppm ก่อนทำการประมวลผลสัญญาณเบื้องต้น จากรูปสังเกตได้ว่า สเปกตรัมที่ความเข้มข้นต่างกันมีรูปแบบและระดับฐานที่ต่างกัน ซึ่งเป็นผลมาจากปัจจัยรบกวนต่างๆ ตามที่ได้บรรยายไปแล้วในหัวข้อ

3.2.1 – 3.2.5 จึงทำให้ไม่สามารถนำชุดข้อมูลสเปกตรัมนี้มาทำการวิเคราะห์เชิงปริมาณได้ ดังนั้นจึงต้องใช้เทคนิคการประมวลผลเบื้องต้น เพื่อลดสิ่งรบกวนหรือข้อมูลที่ไม่เกี่ยวข้องออกจากข้อมูลสเปกตรัม ซึ่งหลังจากการประมวลผลข้อมูลสเปกตรัมแล้ว ทำให้สเปกตรัมรามาที่มีความคล้ายคลึงกันมากขึ้น ดังแสดงในรูปที่ 8 (ข)

3.3 การวิเคราะห์ปริมาณของสารคาเฟอีนมาตรฐาน
(Quantitative Analysis of Standard Caffeine Samples)

ตัวอย่างสเปกตรัมรามาทั้งหมด 1086 สเปกตรัม ที่ผ่านการประมวลผลล่วงหน้าแล้วโดยมีจำนวนสเปกตรัมที่มีความเข้มข้น 500 1000 1500 2000 3000 และ 4000 ppm จำนวน 150 193 185 189 184 และ 185 สเปกตรัม ตามลำดับ จะถูกนำมาแบ่งให้เป็นชุดข้อมูลฝึกสอน (Training set) และชุดข้อมูลการตรวจสอบ (Validation set) จำนวนร้อยละ 80 และร้อยละ 20 ของแต่ละความเข้มข้น ตามลำดับและใช้ในกระบวนการเรียนรู้ของเครื่องเพื่อสร้างแบบจำลองในการวิเคราะห์เชิงปริมาณสารคาเฟอีนมาตรฐาน โดยงานวิจัยนี้ได้ทดลองประสิทธิภาพแบบจำลอง ทั้งหมด 18 แบบ เพื่อเปรียบเทียบประสิทธิภาพของการวิเคราะห์เชิงปริมาณของสารคาเฟอีนมาตรฐาน ตารางที่ 1 และ 2 แสดงผลการเปรียบเทียบประสิทธิภาพของการวิเคราะห์ของทั้ง 18 แบบจำลอง ในกรณีที่มิใช่การประมวลผลสัญญาณเบื้องต้น และกรณีที่ใช่การประมวลผลสัญญาณเบื้องต้นตามลำดับ จากตารางทั้งสองจะพบว่า การประมวลผลสัญญาณเบื้องต้นช่วยให้ประสิทธิภาพในการวิเคราะห์สูงขึ้น อย่างไรก็ตาม การเปลี่ยนแปลงของสเปกตรัมรามาจากความเข้มข้นของสารที่เปลี่ยนไป มีรูปแบบที่ไม่ชัดเจน จึงยังทำให้ค่าความผิดพลาดในการทำนายสูง ทั้งนี้แบบจำลองที่มีประสิทธิภาพสูงสุดในการทำนายค่าความเข้มข้นของสารคาเฟอีนจากสเปกตรัมที่ผ่านการประมวลผลเบื้องต้นแล้ว คือ แบบจำลอง Extra Trees Regressor โดยมีค่าเฉลี่ยความผิดพลาดสมบูรณ์ (MAE) ค่ารากที่สองเฉลี่ยสมบูรณ์ (RMSE) และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R2) อยู่ที่ 710 ppm 914.05 ppm และ 0.34 ตามลำดับ

ตารางที่ 1 ผลการเปรียบเทียบประสิทธิภาพของการวิเคราะห์ของชุดข้อมูลที่ไม่ผ่านการประมวลผลสัญญาณเบื้องต้น

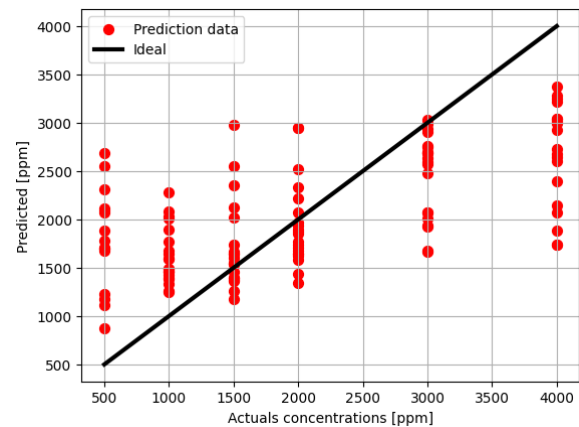
Model	MAE	RMSE	R2
Random Forest Regressor	822.76	1031.16	0.16
Extra Trees Regressor	818.86	1064.70	0.10
Gradient Boosting Regressor	861.81	1067.36	0.10
LightGBM	861.43	1092.33	0.07
AdaBoost Regressor	958.50	1098.67	0.05
K Neighbors Regressor	882.18	1106.50	0.03
Extreme Gradient Boosting	873.36	1126.40	0.00
Dummy Regressor	986.50	1168.31	-0.06
Decision Tree Regressor	986.50	1168.31	-0.45
Lasso Regression	943.47	1357.51	-4.09
Elastic Net	1008.40	1736.23	-4.09
Huber Regressor	1493.61	2525.65	-7.25
Bayesian Ridge	1042.96	2116.54	-9.71
Passive Aggressive Regressor	2048.47	3165.99	-12.47
Lasso Least Angle Regression	1103.52	2326.52	-14.37
Ridge Regression	1788.87	6651.25	-127.00
Linear Regression	2859.01	11522.9	-406.47
Orthogonal Matching Pursuit	2944.78	12400.6	-555.91

ตารางที่ 2 ผลการเปรียบเทียบประสิทธิภาพของการวิเคราะห์ของชุดข้อมูลที่ผ่านการประมวลผลสัญญาณเบื้องต้นแล้ว

Model	MAE	RMSE	R2
Extra Trees Regressor	710.40	914.05	0.3479
Gradient Boosting Regressor	712.10	919.37	0.3376
LightGBM	715.52	924.69	0.3288

Model	MAE	RMSE	R2
Random Forest Regressor	724.37	930.69	0.3238
AdaBoost Regressor	754.58	937.86	0.3103
K Neighbors Regressor	763.39	954.10	0.2861
Elastic Net	949.10	1120.61	0.027
Passive Aggressive Regressor	921.70	1146.74	-0.0182
Dummy Regressor	986.50	1168.32	-0.0617
Ridge Regression	876.68	1165.93	-0.2112
Decision Tree Regressor	930.78	1311.16	-0.3867
Lasso Regression	905.52	1239.66	-0.4791
Lasso Least Angle Regression	905.91	1240.63	-0.4814
Bayesian Ridge	901.84	1315.71	-0.8395
Huber Regressor	895.91	1423.75	-1.3927
Orthogonal Matching Pursuit	1793.18	4164.81	-24.93
Linear Regression	13977	37375.8	-1965
Lasso Least Angle Regression	168898	535077	-11340

รูปที่ 9 เป็นการวิเคราะห์ผลการทำนายความเข้มข้นของสารคาเฟอีนมาตรฐานเพื่อแสดงถึงภาพรวมความผิดพลาดของผลการทำนายปริมาณความเข้มข้นของสารคาเฟอีนจากสเปกตรัมรามานของสารคาเฟอีนมาตรฐานในชุดข้อมูลทดสอบ (Test Set) จำนวน 178 สเปกตรัม ณ ค่าความเข้มข้นต่างๆ โดยใช้แบบจำลอง Extra Tree Regressor เนื่องจากแบบจำลอง Extra Tree Regressor เป็นแบบจำลองที่มีประสิทธิภาพสูงสุด ในการทำนายปริมาณสารคาเฟอีน จึงเลือกแบบจำลองนี้วิเคราะห์ผลจากรูปแสดงให้เห็นว่า ในช่วง 1500 ppm ถึง 3000 ppm แบบจำลองมีแนวโน้มที่จะทำนายความเข้มข้นของสารคาเฟอีนได้ในบริเวณที่ใกล้เคียงกับค่าจริงอยู่บ้าง แต่ในบริเวณความเข้มข้นอื่น ค่าการทำนายยังห่างจากค่าจริงอยู่มาก



รูปที่ 9 Prediction Plot เปรียบเทียบระหว่างค่าความเข้มข้นของสารคาเฟอีนมาตรฐานที่ทำนายได้ (Y pred) เทียบกับค่าจริง (Y true) โดยใช้แบบจำลอง Extra Trees Regressor ที่สร้างได้จากข้อมูลที่ผ่านมาการประมวลผลสัญญาณเบื้องต้น

4. สรุปผลและอภิปรายผล

จากการวิเคราะห์เชิงปริมาณสารคาเฟอีนมาตรฐานโดยใช้เทคนิครามานสเปกโทรสโกปีรวมกับการวิเคราะห์ข้อมูลโดยใช้การเรียนรู้ของเครื่อง ทำให้เห็นถึงความเป็นไปได้ในการใช้วิธีที่เสนอในงานวิจัยนี้ ทำการวิเคราะห์เชิงปริมาณของสารคาเฟอีนที่อยู่ในกาแฟ อย่างไรก็ตาม รูปแบบการเปลี่ยนแปลงของสเปกตรัมรามานที่ไม่ชัดเจนเมื่อความเข้มข้นของสารเปลี่ยนแปลงไปนั้น ทำให้การวิเคราะห์เชิงปริมาณโดยใช้เทคนิครามานสเปกโทรสโกปี ยังมีความท้าทายอยู่มาก ดังจะเห็นได้จากผลการทดสอบในตารางที่ 2 ซึ่งแสดงถึงค่าความผิดพลาดของแบบจำลองที่ยังสูงอยู่ แม้สเปกตรัมจะผ่านการประมวลผลเบื้องต้นแล้วก็ตาม การพัฒนาแบบจำลองให้มีประสิทธิภาพสูงขึ้นจึงเป็นประเด็นวิจัยที่สำคัญในอนาคต ที่จะทำให้สามารถนำเทคนิครามานสเปกโทรสโกปีไปใช้ในอุตสาหกรรมกาแฟจริงได้

5. กิตติกรรมประกาศ

โครงการนี้ได้รับทุนอุดหนุนการวิจัยจากภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์และเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยศิลปากรและขอขอบคุณทีมวิจัยเทคโนโลยีเทระเฮิรตซ์ (เนคเทค-สวทช.) ที่กรุณาช่วยสนับสนุนอุปกรณ์ในงานวิจัย

6. ข้อเสนอแนะ

สามารถพัฒนาและทดสอบแบบจำลองการเรียนรู้ของเครื่องโดยการพัฒนาการประมวลผลสัญญาณเบื้องต้น เพื่อเพิ่มความแม่นยำในการวิเคราะห์เชิงปริมาณสารประกอบหลักในกาแฟ

7. เอกสารอ้างอิง

- [1] Geekabit. (25 August 2022). *5 Benefits of Wireless Network Monitoring for Your Business*. [Online] Available: <https://geekabit.co.uk/2019/06/11/5-benefits-wireless-network-monitoring-business/>
- [2] 7signal. (25 August 2022). *Why Monitor a Wireless Network from the End-User's Perspective?*. [Online] Available: <https://www.7signal.com/news/blog/why-monitor-a-wireless-network-from-the-end-users-perspective>
- [3] R. B. Johnson. *Evaluating the use of SNMP as a wireless network monitoring tool for IEEE 802.11 wireless networks*. Ph.D.thesis. Clemson University, 2009.
- [4] J. Yeo, M. Youssef, T. Henderson, and A. Agrawala, "An accurate technique for measuring the wireless side of wireless networks," *International Conference On Mobile Systems, Applications And Services: Papers presented at the 2005 workshop on Wireless traffic measurements and modeling*, Vol. 5, No. 5, pp. 13–18, 2005.
- [5] I. P. Mohottige, T. Sutjarittham, N. Raju, H. H. Gharakheili and V. Sivaraman, "Role of Campus WiFi Infrastructure for Occupancy Monitoring in a Large University," *2018 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*, Colombo, Sri Lanka, 2018, pp. 1-5, doi: 10.1109/ICIAfS.2018.8913341.
- [6] L. Chen, R. Wu and W. -K. Jia, "Exploring the Optimal Intercept Access Point Placement Problem in Software-Defined Networks," *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2021*, pp. 1-6, doi: 10.1109/CCNC49032.2021.9369594.
- [7] Y. Z. Jian, N. Ismail and M. S. Nabi, "Wireless Access Point Mapper (WAP-MAP): An Automated Wireless Access Point Plotting Web Application," *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2022*, pp. 1-7, doi: 10.1109/ASSIC55218.2022.10088395.
- [8] A. Thitinaruemit, S. Sitjongsatapom and S. Prongnuch, "Capacity Planning of Access Point Deployment for Indoor Positioning System," *2021 International Conference on Power, Energy and Innovations (ICPEI), Nakhon Ratchasima, Thailand, 2021*, pp. 167-170, doi: 10.1109/ICPEI52436.2021.9690679.
- [9] Y. Liu, F. R. Yu, X. Li, H. Ji, H. Zhang and V. C. M. Leung, "Self-optimizing interference management for non-orthogonal multiple access in ultra-dense networks," *2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 2018*, pp. 1-6, doi: 10.1109/WCNC.2018.8377038.
- [10] Hu, Q., Sellers, C., Kwon, J.S., & Wu, H. (2022). Integration of surface-enhanced Raman spectroscopy (SERS) and machine learning tools for coffee beverage classification. *Digital Chemical Engineering*, Vol. 3, 2020, doi: 10.1016/j.dche.2022.100020.

- [11] Chen, Long-Qi and Wenhua Huang. “153. APPLICATION OF RAMAN SPECTROSCOPY AND MACHINE LEARNING IN THE CLASSIFICATION OF ESOPHAGEAL SQUAMOUS CARCINOMA.” *Diseases of the Esophagus*: Vol 35, 2022.
- [12] Schulze, G., Jirasek, A., Yu, M. M. L., Lim, A., Turner, R. F. B., & Blades, M. W. “Investigation of Selected Baseline Removal Techniques as Candidates for Automated Implementation” *Applied spectroscopy*. 59, 545-74, Vol. 59, 2005.
- [13] Whitaker, D. A., & Hayes, K. “A Simple Algorithm for Despiking Raman Spectra” *Chemometrics and Intelligent Laboratory Systems*. Vol. 179, 2018.