การทำนายสมบัติสารด้วยการเรียนรู้ของเครื่องโดยใช้ลายนิ้วมือของโมเลกุล

FINGERPRINT-BASED MACHINE LEARNING PREDICTION OF CHEMICAL PROPERTIES

สุกฤษฎิ์ เกิดสวัสดิ์

เอกภาพ แพสุพัฒน์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมเคมี คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำนายสมบัติสารด้วยการเรียนรู้ของเครื่องโดยใช้ลายนิ้วมือของโมเลกุล

สุกฤษฎ์ เกิดสวัสดิ์

เอกภาพ แพสุพัฒน์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมเคมี คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

FINGERPRINT-BASED MACHINE LEARNING

PREDICTION OF CHEMICAL PROPERTIES

SUKIT KERDSAWT

EKKAPARB PAIRSUPAT

A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR

THE DEGREE OF BACHELOR OF ENGINEERING IN CHEMICAL ENGINEERING

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2023

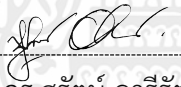
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์เรื่อง	การทำนายสมบัติสารด้วยการเรียนรู้ของเครื่องโดยใช้ลายนิ้วมือของโมเลกุล
โดย	นายสุกฤษฎิ์ เกิดสวัสดิ์ รหัสนักศึกษา 63010968
	นายเอกภาพ แพสุพัฒน์ รหัสนักศึกษา 63011091
อาจารย์ที่ปรึกษา	รศ.ดร.อมตะ อนันต์พินิจวัฒนา
ปริญญานิพนธ์	สาขาวิชาวิศวกรรมเคมี คณะวิศวกรรมศาสตร์
	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

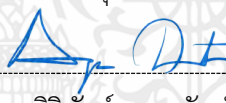
ปริญญานิพนธ์นี้ได้รับการพิจารณาอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมเคมี

คณะกรรมการตรวจสอบปริญญานิพนธ์

ประธานกรรมการ
(รศ.ดร.อมตะ อนันต์พินิจวัฒนา)



กรรมการ
(ผศ.ดร.สุรัตน์ อารีรัตน์)



กรรมการ
(ผศ.ศิริพันธ์ มุรธาธัญลักษณ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์เรื่อง	การทำนายสมบัติสารด้วยการเรียนรู้ของเครื่องโดยใช้ลายนิ้วมือของโมเลกุล		
โดย	นายสุกฤษฎิ์	เกิดสวัสดิ์	รหัสนักศึกษา 63010987
	นายเอกภาพ	แพสุพัฒน์	รหัสนักศึกษา 63011091
ปริญญา	วิศวกรรมศาสตรบัณฑิต		
สาขาวิชา	วิศวกรรมเคมี		
ปีการศึกษา	2566		
อาจารย์ที่ปรึกษา	รศ.ดร.อมตะ อนันต์พินิจวัฒนา		

บทคัดย่อ

การคัดเลือกสารที่มีสมบัติเหมาะสมสำหรับกระบวนการผลิตเป็นปัญหาที่มีมากอย่างยาวนานของอุตสาหกรรมปิโตรเคมี นอกจากการทดลองแล้ว การทำนายสมบัติสารจะใช้ความสัมพันธ์ระหว่างสมบัติกับโครงสร้างเชิงปริมาณ (QSPR) และการกระจายแบบกลุ่ม (Group Contribution Method) โดยในปัจจุบันมีการประยุกต์ใช้การเรียนรู้ของเครื่อง (Machine Learning) เพื่อทำนายสมบัติสารเนื่องมากจากมีความรวดเร็วและมีความแม่นยำที่มากกว่าวิธีการแบบดั้งเดิม เช่นงานวิจัยของ Nattasinee และ คณะ ที่ทำนายจุดเดือด (Normal Boiling Point) สำหรับสารอินทรีย์ที่มีองค์ประกอบของคาร์บอน และ ไฮโดรเจน พบว่ามีข้อจำกัดจากการมีคุณลักษณะที่ใช้แบ่งแยกโมเลกุลสารไม่สามารถใช้ได้กับคู่สารที่มีโครงสร้างคล้ายกัน งานวิจัยนี้จึงปรับปรุงการระบุคุณลักษณะของโมเลกุลสารด้วยลายนิ้วมือโมเลกุล โดยตัวเลขในลายนิ้วมือโมเลกุลสามารถแทนโครงสร้างย่อยที่มีอยู่ภายในโมเลกุลได้ นอกจากนี้ยังขยายขอบเขตกลุ่มสารที่ศึกษาในโครงงานนี้ซึ่งประกอบด้วย สารกลุ่มไฮโดรคาร์บอนที่มีคาร์บอนอะตอมตั้งแต่ 1-12 อะตอม และมีองค์ประกอบเป็น คาร์บอน ไฮโดรเจน ออกซิเจน และ ไนโตรเจน โดยมีจำนวนสารเพิ่มขึ้นจาก 560 ชนิด เป็น 11,177 ชนิด และมีค่าสัมประสิทธิ์การกำหนด (R^2) ของสำหรับแบบจำลองทำนายจุดเดือด (T_b) เป็น 0.655 และแบบจำลองทำนายความดันไอ ($\ln(P^{sat})$) เป็น 0.798

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Report Title	Fingerprint-Based Machine Learning Prediction of Chemical Properties		
By	Sukit	Kerdsawat	Student ID 63010987
	Ekkaparb	Pairsupat	Student ID 63011091
Degree	Bachelor of Engineering		
Program	Chemical Engineering		
Year	2566		
Advisor	Assoc. Prof. Dr. Amata Anantpinijwatna		

Abstract

Selection of suitable substances for production processes is a long-standing problem in the petrochemical industry. Besides the experiment, predicting substance properties will use quantitative structure-property relationships (QSPR) and Group Contribution Method. Currently, machine learning is applied to predict substance properties because it is faster and more accurate than traditional methods. For example, the research by Nattasinee et al. that predicted the normal boiling point for organic substances containing carbon and hydrogen found that they were limited by having features used to separate substance molecules that could not be used with pairs of substances with similar structures. Therefore, this research improves the featurization of substance molecules using molecular fingerprints. The numbers in a molecular fingerprint can represent substructures within a molecule. It also expands the scope of substances studied in this project, which includes hydrocarbon substances that have carbon atoms from 1-12 atoms and have elements of carbon, hydrogen, oxygen and nitrogen. The number of substances has increased from 560 species to 11,177 species and has a coefficient of determination (R^2) for the boiling point prediction model (T_b) is 0.655 and the vapor pressure prediction model ($\ln(P^{sat})$) is 0.798.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์เรื่อง การทำนายสมบัติสารด้วยการเรียนรู้ของเครื่องโดยใช้ลายนิ้วมือของโมเลกุล สำเร็จลุล่วงไปได้ด้วยดี เนื่องจากการได้รับความกรุณา และความช่วยเหลือจาก รศ.ดร.อมตะ อนันต์พินิจ อาจารย์ที่ปรึกษางานวิจัยที่คอยให้คำแนะนำและคำปรึกษาในทุกขั้นตอนการทำวิจัย ตลอดจนความรู้ทางวิชาการที่เป็นประโยชน์ต่องานวิจัยให้สำเร็จบรรลุตามวัตถุประสงค์

ขอขอบคุณคณาอาจารย์ บุคลากร และเพื่อนร่วมการศึกษา คณะวิศวกรรมศาสตร์ สาขาวิชาวิศวกรรมเคมี สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่อำนวยความสะดวกให้สถานที่ในการศึกษาและค้นคว้าทำการทดลอง ทั้งด้านอุปกรณ์การทดลอง สาธารณูปโภคและขอขอบคุณแหล่งค้นคว้าข้อมูลสำคัญที่เป็นประโยชน์แก่งานวิจัย ทั้งหนังสือ ตัวอย่างงานวิจัยต่างๆ เว็บไซต์ และบทความ วารสารทางวิชาการ ที่คณะผู้วิจัยได้นำมาใช้ประกอบการทำเล่มวิจัยฉบับนี้

สุดท้ายนี้ ทางคณะผู้จัดทำวิจัย ขอขอบคุณผู้มีพระคุณที่ช่วยเหลือและสนับสนุนทุนการวิจัยอันมีค่าแก่งานวิจัยนี้ และหวังว่างานวิจัยเล่มนี้จะเป็นประโยชน์และเป็นแนวทางให้กับผู้ที่สนใจศึกษา และหากผิดพลาดประการใดทางผู้วิจัยขออภัย มา ณ ที่นี้ด้วยพร้อมน้อมรับคำแก้ไขทุกประการ

คณะผู้วิจัย

นายสุกฤษฎิ์ เกิดสวัสดิ์

นายเอกภาพ แพสุพัฒน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อ II

Abstract III

กิตติกรรมประกาศ IV

สารบัญ V

สารบัญตาราง VIII

สารบัญรูป IX

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญ 1

1.2 วัตถุประสงค์ของโครงงาน 2

1.3 ขอบเขตการศึกษา 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 4

2.1 สารอินทรีย์และสมบัติสาร (Organic Compound and Properties) 4

2.1.1 สารอินทรีย์ (Organic Compound) 4

2.1.2 สมบัติของสารอินทรีย์: จุดเดือด (Boiling Point) 6

2.1.3 สมบัติของสารอินทรีย์: ความละลายได้ (Solubility) 6

2.1.4 สมบัติของสารอินทรีย์: ความจุความร้อนจำเพาะ (Specific Heat Capacity) 6

2.1.5 สมบัติของสารอินทรีย์: ความดันไอ (Vapor Pressure) 6

2.2 โครงสร้างของโมเลกุลสาร (Structure of Molecule) 7

2.2.1 Simplified Molecular-Input Line-Entry System (SMILES) 7

2.2.2 ลายนิ้วมือโมเลกุล (Molecular Fingerprint) 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.3 การเรียนรู้ของเครื่อง (Machine Learning) 9
2.3.1 อัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning Algorithm) 11
2.3.2 Cross Validation: K-fold 13
2.4 การประเมินผลแบบจำลอง (Model Evaluation) 14
2.5 งานวิจัยที่เกี่ยวข้อง (Related Work) 15
2.5.1 การใช้ลายนิ้วมือโมเลกุล และ Machine Learning เพื่อพัฒนาแบบจำลองทำนายสมบัติของ Ionic Liquid 15
2.5.2 วิธีการ Count-base Morgan Fingerprint ที่ใช้สำหรับบ่งบอกโครงสร้างโมเลกุลในการพัฒนา Machine Learning สำหรับการทำนายสมบัติของน้ำที่ปนเปื้อน 15
2.5.3 การใช้การเรียนรู้ของเครื่องเพื่อพัฒนาแบบจำลองทำนายสมบัติของสารอินทรีย์บริสุทธิ์ 17
2.5.4 การทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN (Path-Unifying Feed-Forward Interfaced Network) 19
บทที่ 3 วิธีการดำเนินงาน 20
3.1 การเก็บรวบรวมข้อมูล (Data Collecting) 20
3.2 การเตรียมข้อมูลและการแบ่งข้อมูล (Data Preparation & Data Splitting) 21
3.3 การฝึกฝนโมเดล (Model Training) 23
3.3.1 การฝึกฝนโมเดลเพื่อสร้างแบบจำลองการทำนายจุดเดือด 23
3.3.2 การฝึกฝนโมเดลเพื่อสร้างแบบจำลองการทำนายความดันไอและทำนายค่า Antoine Coefficients 23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 4 ผลการทดลองและวิเคราะห์ผลการทดลอง 25

4.1 แบบจำลองทำนายจุดเดือด 25

4.1.1 การเปรียบเทียบผลการทำนายระหว่างแบบจำลองจากงานวิจัย Nattasinee และคณะ กับแบบจำลองทำนายสมบัติที่ประยุกต์ลายนิ้วมือโมเลกุลมาใช้ 25

4.1.2 การประยุกต์แบบจำลองทำนายสมบัติที่ประยุกต์ลายนิ้วมือโมเลกุลมาใช้กับขอบเขตของสารอินทรีย์ที่ประกอบไปด้วยอะตอมของคาร์บอน ไฮโดรเจน ออกซิเจนและไนโตรเจน 28

4.2 แบบจำลองทำนายความดันไอ 32

4.2.1 วิเคราะห์การสร้างแบบจำลองทำนาย $\ln(P^{sat})$ 33

4.2.2 วิเคราะห์การสร้างแบบจำลองทำนาย Antoine Coefficients 39

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 43

5.1 สรุปผลการวิจัย 43

5.2 ปัญหา อุปสรรค และข้อเสนอแนะ 44

เอกสารอ้างอิง 45

ภาคผนวก 49

ภาคผนวก ก. โค้ดสำหรับการแปลง SMILES เป็น Morgan Fingerprint 50

และการสร้างแบบจำลอง 50

ภาคผนวก ข. จุดเดือดของสารอินทรีย์ที่ประกอบไปด้วย คาร์บอน ไฮโดรเจน ออกซิเจน และไนโตรเจนจากแหล่งอ้างอิงและการทำนาย 53

ภาคผนวก ค. ความดันไอของสารอินทรีย์ที่ประกอบไปด้วย คาร์บอน ไฮโดรเจน ออกซิเจน และไนโตรเจนจากแหล่งอ้างอิงและการทำนาย 87

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 2.1 การทดสอบประสิทธิภาพของแบบจำลองสำหรับ Ionic ของ Yi Ding และคณะ 15
ตารางที่ 2.2 ข้อมูลสมบัติที่เกี่ยวข้องกับน้ำที่ปนเปือนของ Shifa Zhong และ คณะ..... 16
ตารางที่ 2.3 การทดสอบประสิทธิภาพลายนิ้วมือของแต่ละชุดข้อมูลสำหรับน้ำปนเปื้อนของ Shifa Zhong และ คณะ..... 16
ตารางที่ 2.4 การประเมินประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point ของ งานวิจัยของ Nattasinee และคณะ..... 18
ตารางที่ 2.5 ผลลัพธ์การเปลี่ยน SMILES เป็นข้อมูลสำหรับการเรียนรู้และผลการทำนาย Normal Boiling Point ของ งานวิจัยของ Nattasinee และคณะ 18
ตารางที่ 3.1 แหล่งข้อมูลจาก Chemical Database of Chemical Engineering Design Library (ChEDL)..... 20
ตารางที่ 3.2 แสดงการตั้งค่าเพื่อหาลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการทำแบบจำลองทำนายจุดเดือด 22
ตารางที่ 3.3 แสดงการตั้งค่าเพื่อหาลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการทำแบบจำลองทำนายความดันไอ 22
ตารางที่ 4.1 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point ในขอบเขตอะตอมของคาร์บอนและไฮโดรเจน กับ Machine Learning Algorithmm ต่างๆ..... 26
ตารางที่ 4.2 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย Normal Boiling Point ที่ดีที่สุด. 27
ตารางที่ 4.3 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point งานวิจัยของ Nattasinee และคณะ กับงานในครั้งนี้ 27
ตารางที่ 4.4 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนายจุดเดือดของขอบเขต CHON..... 29
ตารางที่ *4.5* แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนายจุดเดือดที่ดีที่สุด..... 30
ตารางที่ 4.6 ตารางระหว่างหมู่ฟังก์ชันและค่าความคลาดเคลื่อน 31
ตารางที่ 4.7 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Vapor Pressure..... 33
ตารางที่ 4.8 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย $\ln(P^{sat})$ ที่ดีที่สุด 34
ตารางที่ 4.9 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Vapor Pressure ในหน่วย Pa 34
ตารางที่ 4.10 ตารางระหว่างหมู่ฟังก์ชันและค่าความคลาดเคลื่อน 37
ตารางที่ 4.11 เปรียบเทียบค่าความคลาดเคลื่อนของค่าที่ได้จากแบบจำลองทำนาย Antione eff. . 40
ตารางที่ 4.12 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย Antoine Coefficients ที่ดีที่สุด41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ **หน้า**

รูปที่ 2.1 กลุ่มสารที่มีองค์ประกอบคาร์บอนกับไฮโดรเจน 4
รูปที่ 2.2 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจนและออกซิเจน 5
รูปที่ 2.3 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจนและไนโตรเจน 5
รูปที่ 2.4 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจน ออกซิเจนและไนโตรเจน 5
รูปที่ 2.5 กลุ่มสารประกอบหมู่ฟังก์ชั่นที่มีอะตอมอื่นๆ 6
รูปที่ 2.6 Furfural, SMILES: c1cc(C=O)oc1 7
รูปที่ 2.7 การกำหนดตำแหน่งอะตอมภายในของ Butyramide 8
รูปที่ 2.8 การระบุโครงสร้างย่อยภายในของ Butyramide 9
รูปที่ 2.9 การกำจัดโครงสร้างย่อยที่ซ้ำกันของ Butyramide 9
รูปที่ 2.10 แสดงรูปแบบและชนิดการเรียนรู้ของ Machine Learning 10
รูปที่ 2.11 แผนผัง Machine Learning แบบง่าย 10
รูปที่ 2.12 Ridge Regression 11
รูปที่ 2.13 Random Forest 12
รูปที่ 2.14 Extreme Gradient Boosting 12
รูปที่ 2.15 K-Nearest Neighbors 13
รูปที่ 2.16 Deep Learning 13
รูปที่ 2.17 K-fold Cross Validation 14
รูปที่ 2.18 แผนผัง Machine Learning ฉบับงานวิจัยของ Nattasinee และคณะ 17
รูปที่ 2.19 กราฟแสดงประสิทธิภาพการทำนาย Normal Boiling Point ของงานวิจัยของ Nattasinee และคณะ 18
รูปที่ 2.20 แผนผังแสดงการสร้างแบบจำลองทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN ของงานวิจัย Santana และ คณะ 19
รูปที่ 2.21 กราฟแสดงประสิทธิภาพการทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN ของงานวิจัย Santana และ คณะ 19
รูปที่ 3.1 แผนผังการทำงานภายในการพัฒนา Propertes Prediction Model จาก Machine Learning 20
รูปที่ 3.2 Boxplot แสดงการกระจายตัวของช่วงข้อมูลปกติและนอกช่วงข้อมูลปกติ 21

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

รูปที่ 4.1 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองทำนาย Normal Boiling Point 25
รูปที่ 4.2 กราฟเปรียบเทียบผลการทำนาย Normal Boiling Point ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน จาก KNN, XGB, Ridge และ RF Algorithm 26
รูปที่ 4.6 กราฟเปรียบเทียบผลการทำนายจากงานวิจัย Nattasinee และคณะ กับงานวิจัยนี้ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน 27
รูปที่ 4.7 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองทำนาย Normal Boiling Point ที่ขอบเขตสารเป็น CHON 29
รูปที่ 4.8 กราฟเปรียบเทียบผลการทำนาย Normal Boiling Point ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน จาก KNN, RF, Ridge และ XGB Algorithm 30
รูปที่ 4.9 กราฟเปรียบเทียบผลการทำนายจุดเดือดจากแบบจำลองด้วย ขอบเขต CHON ที่ได้แยกตามหมู่ฟังก์ชั่นของ Test Set 31
รูปที่ 4.10 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองการทำนายความดันไอ 32
รูปที่ 4.11 กราฟเปรียบเทียบผลการทำนายความดันไอ จาก ก) DT ข. KNN ค) RF และ ง) XGB Algorithm 33
รูปที่ 4.12 กราฟเปรียบเทียบผลการทำนายความดันไอในหน่วย Pa จาก ก) DT ข) KNN ค) RF และ ง) XGB Algorithm 34
รูปที่ 4.13 กราฟเปรียบเทียบผลการทำนายความดันไอ จาก XGB Algorithm จาก Train (ซ้าย) และ Test (ขวา) Set 35
รูปที่ 4.14 กราฟเปรียบเทียบผลการทำนายความดันไอจากแบบจำลองที่ได้แยกตามหมู่ฟังก์ชั่นสำหรับ Training Set 36
รูปที่ 4.15 กราฟเปรียบเทียบผลการทำนายความดันไอจากแบบจำลองที่ได้แยกตามหมู่ฟังก์ชั่นสำหรับ Test Set 36
รูปที่ 4.16 กราฟเปรียบเทียบผลการทำนาย Antione Coefficient จาก XGB Algorithm จาก Train Set 37
รูปที่ 4.17 กราฟเปรียบเทียบผลการทำนาย Antione Coefficient จาก XGB Algorithm จาก Test Set 38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

รูปที่ 4.18 กราฟความสัมพันธ์ระหว่างอุณหภูมิกับความดันไอ จาก Antione Coefficient ทั้งจากฐานข้อมูลและค่าที่ได้จากการทำนายความดันไอของโมเลกุลสารของตัวอย่างผลลัพธ์ที่ดี 38
รูปที่ 4.19 กราฟความสัมพันธ์ระหว่างอุณหภูมิกับความดันไอ จาก Antione Coefficient ทั้งจากฐานข้อมูลและค่าที่ได้จากการทำนายความดันไอของโมเลกุลสารของตัวอย่างผลลัพธ์ที่ไม่ดี 38
รูปที่ 4.20 สารที่มีระยะห่างอุณหภูมิเยอะและการคำนวณนอกอุณหภูมิ 39
รูปที่ 4.21 สารที่มีระยะห่างอุณหภูมิน้อยและการคำนวณนอกอุณหภูมิ 39
รูปที่ 4.22 ค่า A ที่ได้จากแบบจำลอง 40
รูปที่ 4.23 ค่า B ที่ได้จากแบบจำลอง 40
รูปที่ 4.24 ค่า C ที่ได้จากแบบจำลอง 40
รูปที่ 4.25 กราฟเปรียบเทียบการคำนวณค่า $\ln(P^{sat})$ จากคำนวณด้วยค่าที่ได้จากการทำนายของแบบจำลอง 41
รูปที่ 4.26 กราฟเปรียบเทียบผลการทำนายความดันไอในหน่วย Pa จากคำนวณด้วยค่าที่ได้จากการทำนายของแบบจำลอง 41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

ในกระบวนการผลิตของโรงงานอุตสาหกรรมในปัจจุบันต้องมีการปรับตัวตามสถานการณ์ของโลก จึงจำเป็นต้องมีข้อมูลเกี่ยวกับสมบัติของสารที่ถูกใช้ในกระบวนการผลิต ตัวอย่างเช่น จุดเดือด มีความเกี่ยวข้องกับการจำลองในสมัยปัจจุบันซึ่งต้องการจุดเดือดและความหนาแน่นในสภาวะของเหลวของสาร เพื่อประมาณสมบัติอื่นๆ ที่จำเป็น [1] เช่น ความดันไอ สมดุลวัฏภาคของสารผสม จุดวาบไฟ [2] ซึ่งสามารถนำไปใช้หาความติดไฟได้เพื่อออกแบบทางด้านความปลอดภัยของกระบวนการผลิตได้ [3] สมบัติต่างๆของสารนั้นสามารถมีทั้งสมบัติที่ไม่ขึ้นและขึ้นกับสภาวะที่ดำเนินการของกระบวนการ (ความดัน อุณหภูมิ และอื่นๆ) และแผนผังผลิตภัณฑ์ของอุตสาหกรรมเคมีแสดงให้เห็นว่าสารอินทรีย์นั้นเกี่ยวข้องโดยตรงตั้งแต่ต้นน้ำจนถึงปลายน้ำ การที่จะได้มาซึ่งสมบัติใดๆของสารอินทรีย์หรือสารใดๆนั้นจำเป็นต้องมีการทดลองทดสอบสมบัติหรืออาจจะต้องทำการสังเคราะห์สารเคมีตัวใหม่ขึ้นมาเพื่อให้ได้สารใหม่ที่มีสมบัติตามที่ต้องการ ซึ่งการทดลองและการสังเคราะห์นั้นต้องใช้ทรัพยากรในการทดสอบที่ค่อนข้างมีราคาสูงและใช้เวลาเป็นอย่างมาก [4] ดังนั้นการพัฒนาแบบจำลองที่สามารถช่วยในการทำนายสมบัติของสารได้จึงเป็นที่ต้องการ

สมบัติสารสามารถทำนายได้ด้วยแบบจำลอง QSPR (Quantitative Structure Properties Relationship) ซึ่งเป็นการสร้างความสัมพันธ์ระหว่างคุณลักษณะทางโครงสร้างด้วยกันกับกลุ่มสมบัติของโมเลกุลสารเคมีที่สนใจด้วยความสัมพันธ์ทางคณิตศาสตร์และสถิติ [4,5] โดยวิธีการแบบดั้งเดิมที่ใช้นั้นเช่น Group Contribution ซึ่งมักถูกใช้กับสารผสมที่มีหมู่ฟังก์ชั่นภายในแตกต่างกันเล็กน้อย โดยใช้ผลรวมผลของอันตรกิริยาของแต่ละหมู่ฟังก์ชั่นโดยนิยมใช้ความสัมพันธ์เชิงเส้น [6] วิธีการนี้จะเกิดปัญหาขึ้นถ้าปฏิสัมพันธ์กันระหว่างหมู่ฟังก์ชั่นมีความซับซ้อนและสภาวะระบบมีการเปลี่ยนแปลงขึ้น

ความก้าวหน้าของการคำนวณทั้งในเรื่องของการเก็บข้อมูลและการประมวลผล ทำให้ Machine Learning ซึ่งเป็นศาสตร์ที่ทำให้คอมพิวเตอร์สามารถเรียนรู้ได้จากข้อมูลนั้นสามารถทำการคำนวณและจัดการข้อมูลที่มีปริมาณมากและซับซ้อนได้รวดเร็วและใช้ทรัพยากรได้อย่างประสิทธิภาพ [6] โดย Machine Learning ได้ถูกนำมาใช้ในการค้นคว้าและออกแบบโมเลกุลสารใหม่ๆ ช่วยในการวิจัย และยังสามารถนำมาใช้ในการทำแบบจำลองทำนายสมบัติของสารเคมีได้ดีกว่าวิธีการแบบ Group Contribution ในเรื่องความซับซ้อนของโมเลกุลสารและระบบ

จากงานวิจัย Nattasinee และคณะ [7] ซึ่งเกี่ยวกับกับการทำแบบจำลองที่ใช้ในการทำนาย Normal Boiling Point ด้วย Machine Learning โดยขอบเขตของแบบจำลองในงานวิจัยนี้ใช้ได้กับ Hydrocarbon ที่มีเพียงอะตอมของ C H เท่านั้น โดยเริ่มต้นทำการแปลงโครงสร้างสารเคมีซึ่งถูก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เขียนด้วย SMILES (Simplified Molecular-Input Line-Entry System) เป็นจำนวนอะตอมคาร์บอน พันธะเดี่ยว พันธะคู่ พันธะสาม กิ่ง และจำนวนวงอะโรมาติก ซึ่งนำไปใช้ไปในการทำแบบจำลอง Machine Learning ปัญหาที่เกิดขึ้นในงานวิจัยนั้นเกิดขึ้นอันเนื่องมาจากการแปลงข้อมูลโครงสร้างดังกล่าวไม่สามารถแยกแยะโครงสร้างที่คล้ายคลึงกันได้และวิธีการแปลงนี้ทำให้ขยายกลุ่มของสารเคมีที่ต้องการทำนายได้ยากมากยิ่งขึ้น

งานวิจัยในครั้งนี้ศึกษาการประยุกต์ใช้ลายนิ้วมือของโมเลกุล (Molecular Fingerprint) เพื่อแปลงโครงสร้างสารเคมีจาก SMILES ซึ่งทำการระบุทุกโครงสร้างย่อยภายในโมเลกุลดังกล่าว แทนที่การแปลงเดิมและการแปลงที่ไม่สามารถย้อนกลับได้ และได้ทำการปรับปรุง Machine Learning Algorithm ให้สอดคล้องกับการแปลงดังกล่าวเพื่อปรับปรุงประสิทธิภาพของแบบจำลองทำนายให้ดียิ่งขึ้น ทั้งนี้โครงงานนี้สามารถต่อยอดเพื่อนำไปใช้ในการสร้างแบบจำลองเพื่อทำนายสมบัติอื่นๆได้ด้วยเช่นกัน

1.2 วัตถุประสงค์ของโครงงาน

1. เพื่อศึกษาการแปลงโครงสร้างแบบ SMILE เป็นลายนิ้วมือโมเลกุล
2. เพื่อทำนายจุดเดือดและความดันไอของสารจากลายนิ้วมือโมเลกุลด้วยการเรียนรู้ของเครื่อง
3. เพื่อเพิ่มความแม่นยำของการทำนายจุดเดือดและความดันไอของสารด้วยเทคนิคด้านการเรียนรู้ของเครื่อง

1.3 ขอบเขตการศึกษา

1. ศึกษาสมบัติของสารอินทรีย์ที่มีองค์ประกอบเป็น คาร์บอน ไฮโดรเจน ออกซิเจน และไนโตรเจน เป็นองค์ประกอบและมีจำนวนคาร์บอนอะตอมตั้งแต่ 1-12 อะตอม
2. ศึกษาลายนิ้วมือโมเลกุลแบบ Morgan Fingerprint ซึ่งเป็นส่วนหนึ่งของเครื่องมือสำเร็จรูป RDKit ในโปรแกรมภาษาไพทอน (Python)
3. ศึกษาการทำนายจุดเดือดและความดันไอโดยใช้การแปลงโครงสร้างของสารด้วยลายนิ้วมือโมเลกุล
4. ศึกษาการทำนายจุดเดือดและความดันไอโดยใช้อัลกอริทึม Ridge Regression, K-Nearest Neighbors, Random Forest, Extreme Gradient Boosting และ Deep Learning ที่มีในไลบรารี่ (Library) สำหรับภาษาไพทอน (Python) ที่ชื่อว่า Scikit-learn, XGBoost และ PyTorch
5. ประเมินประสิทธิภาพของแบบจำลองด้วยสัมประสิทธิ์การกำหนด (Coefficient of Determination, R^2) ค่าความคลาดเคลื่อนแบบ MAE (Mean Absolute Error) %MAPE (Mean Absolute Percentage Error) และ RMSE (Root Mean Square Error)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. แบบจำลองทำนายสมบัติที่ตอบสนองกับความต้องการทั้งการศึกษาและอุตสาหกรรม
2. เรียนรู้เครื่องมือสำหรับการศึกษาสารเคมี (Cheminformatic) และวิธีการสร้างแบบจำลองการเรียนรู้ของเครื่อง
3. ประหยัดเวลาและค่าใช้จ่ายในการทดสอบหาสมบัติของสาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

ในการศึกษาเพื่อทำนายสมบัติของสารประกอบไปด้วยทฤษฎีและงานวิจัยที่เกี่ยวข้องต่างๆ ประกอบไปด้วย

2.1 สารอินทรีย์และสมบัติสาร (Organic Compound and Properties)

2.2 โครงสร้างของโมเลกุลสาร (Structure of Molecule)

2.3 การเรียนรู้ของเครื่อง (Machine Learning)

2.4 การประเมินผลแบบจำลอง (Model Evaluation)

2.5 งานวิจัยที่เกี่ยวข้อง (Related Work)

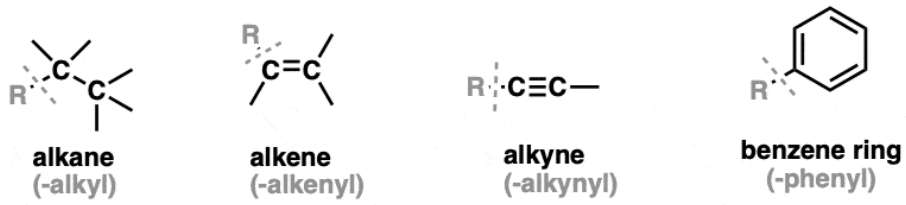
2.1 สารอินทรีย์และสมบัติสาร (Organic Compound and Properties)

2.1.1 สารอินทรีย์ (Organic Compound)

ไฮโดรคาร์บอน (Hydrocarbon) เป็นสารประกอบเคมีอินทรีย์ที่มีธาตุคาร์บอนและไฮโดรเจนเป็นองค์ประกอบหลัก ซึ่งเป็นองค์ประกอบหลักที่พบใน ปิโตรเลียม แก๊สธรรมชาติ ยางไม้ ถ่านหิน เป็นต้น นอกจากนี้ยังสามารถนำประยุกต์ใช้ในการผลผลิตอื่นๆได้ เช่น พลาสติก เส้นใย ปุ๋ย ตัวทำละลายต่างๆ

ไฮโดรคาร์บอนสามารถแบ่งออกได้เป็น 2 ประเภทหลักๆ ได้แก่ Aliphatic และ Aromatic. โดย Aliphatic เป็น Hydrocarbon ที่อะตอมของ C ต่อกันเป็นเส้นตรงหรือสายที่มีกิ่ง ในขณะที่ aromatic นั้นอะตอมของคาร์บอนจะต่อกันเป็นวงอะโรมาติก นอกจากนี้ไฮโดรคาร์บอนยังสามารถแบ่งออกจากกันด้วยหมู่ฟังก์ชั่นดังต่อไปนี้ [8–10]

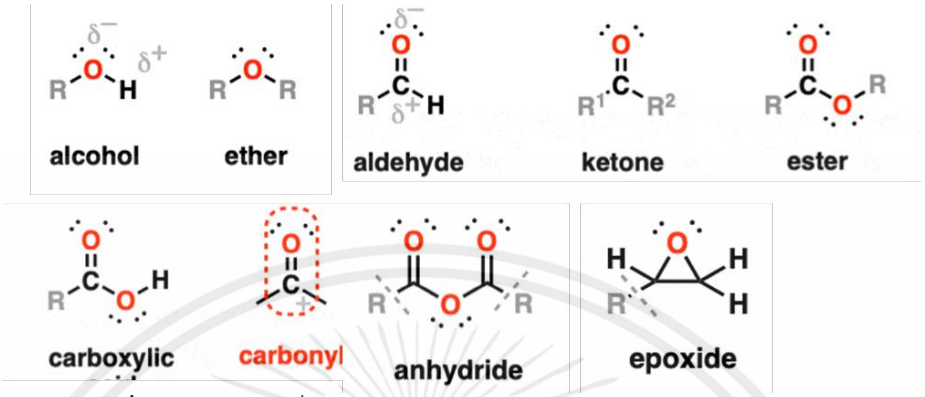
1. Hydrocarbon ที่มีอะตอมของคาร์บอน(C), ไฮโดรเจน(H) เท่านั้น ดังรูปที่ 2.1 อะตอมเชื่อมกันด้วยพันธะเดี่ยว คู่ หรือ สามเป็นต้น ซึ่งได้แก่ Alkane, Alkene, Alkyne และ Aromatic hydrocarbon



รูปที่ 2.1 กลุ่มสารที่มีองค์ประกอบคาร์บอนกับไฮโดรเจน

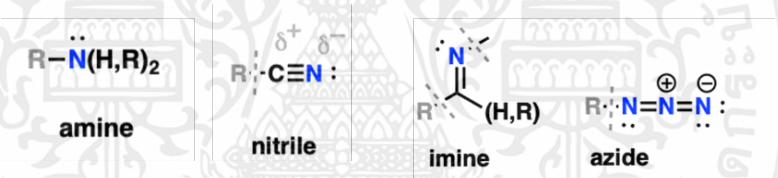
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Hydrocarbon ที่ประกอบด้วย C, H และออกซิเจน(O) ดังรูปที่ 2.2 อะตอมของออกซิเจน เชื่อมพันธะเดี่ยวหรือคู่กับอะตอมคาร์บอน ซึ่งได้แก่ Alcohol, Ether, Aldehyde, Ketone, Ester, Carboxylic Acid, Epoxide, Acid Anhydride และ Epoxide



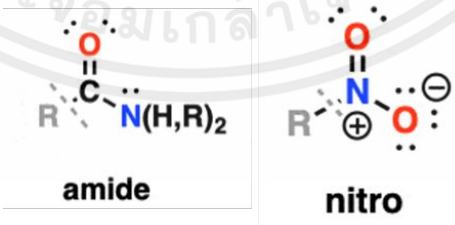
รูปที่ 2.2 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจนและออกซิเจน

3. Hydrocarbon ที่ประกอบด้วย C, H และไนโตรเจน(N) ดังรูปที่ 2.3 อะตอมของไนโตรเจนเชื่อมพันธะเดี่ยว คู่ หรือ สาม กับอะตอมคาร์บอน ซึ่งได้แก่ Amine, Nitrile, Imine, Azide



รูปที่ 2.3 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจนและไนโตรเจน

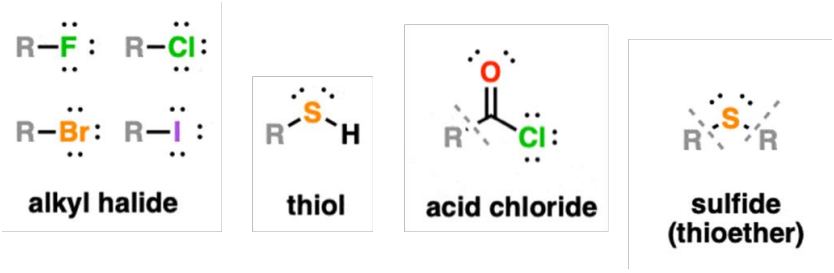
4. Hydrocarbon ที่ประกอบด้วย C, H, O และ N ดังรูปที่ 2.4 อะตอมออกซิเจนหรืออะตอมของไนโตรเจนอย่างน้อย 1 อะตอมจะเชื่อมพันธะกับอะตอมคาร์บอน ซึ่งได้แก่ Amide, Nitrile, Imine, Azo Compound, Azide



รูปที่ 2.4 กลุ่มสารที่มีองค์ประกอบคาร์บอน ไฮโดนเจน ออกซิเจนและไนโตรเจน

นอกจากนี้ยังมีประเภทของหมู่ฟังก์ชั่นที่มีอะตอมอื่นๆประกอบเพิ่มเติมซึ่งได้แก่ Sulfur, Halogen และอื่นๆ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5 กลุ่มสารประกอบหมู่ฟังก์ชั่นที่มีอะตอมอื่นๆ

2.2 สมบัติของสารอินทรีย์:

2.2.1 จุดเดือด (Boiling Point)

จุดเดือดของสาร คือ อุณหภูมิที่สารใดๆที่สถานะเปลี่ยนวัฏภาคจากของเหลวกลายเป็นไอ ซึ่งเกิดขึ้นเมื่อความดันไอของของเหลวสารดังกล่าวมีค่าอย่างน้อยหรือเท่ากับความดันที่อยู่รอบข้าง (ซึ่งปกติจะเป็นความดันบรรยากาศ) จึงมีอีกชื่อเรียกหนึ่งว่า Normal Boiling Point สำหรับการพิจารณาไฮโดรคาร์บอนเบื้องต้นนั้น ถ้ามีปัจจัยที่ทำให้พลังงานที่ต้องใช้สลายแรงดึงดูดระหว่างโมเลกุลออกจากกัน (Van Der Waals Force) นั้นสูงขึ้น จะทำให้มีจุดเดือดสูงขึ้นด้วย เช่น การเพิ่มความยาวสายไฮโดรคาร์บอน การมีหมู่ฟังก์ชันในสายไฮโดรคาร์บอน [11]

2.2.2 ความละลายได้ (Solubility)

ความละลายได้ คือ ความสามารถของสารที่จะถูกละลายในตัวทำละลายชนิดหนึ่ง ซึ่งวัดจากปริมาณความเข้มข้นอิ่มตัวของสารละลายระหว่างตัวถูกละลายและตัวทำละลายดังกล่าว โดยความละลายได้ดังกล่าวจะขึ้นกับชนิดของตัวทำละลายที่ใช้ อุณหภูมิ และความดัน [12]

2.2.3 ความจุความร้อนจำเพาะ (Specific Heat Capacity)

ความจุความร้อนจำเพาะ คือ ปริมาณพลังงานที่ใช้ในการเปลี่ยนแปลงอุณหภูมิหนึ่งหน่วย (°C / K) ต่อหนึ่งหน่วยของมวล ซึ่งบ่งบอกถึงปริมาณพลังงานที่สารนั้นสามารถกักเก็บพลังงานความร้อนไว้ได้ [13] โดยไม่ขึ้นกับปริมาณหรือขนาดของสาร ซึ่งปริมาณความจุความร้อนนี้จะมีค่าที่แตกต่างกันได้ตามวัฏภาคและอุณหภูมิเริ่มต้นของสารดังกล่าว และขึ้นกับแรงดึงดูดระหว่างโมเลกุลและความซับซ้อนของโมเลกุล

2.2.4 ความดันไอ (Vapor Pressure)

ความดันไอของสารบริสุทธิ์ คือ ลักษณะความดันที่อุณหภูมิใด ๆ ที่ไออยู่ในสภาวะสมดุลกับของเหลวหรือของแข็ง ความดันไอเป็นค่าที่บ่งบอกถึงความสามารถในการยึดติดกันเองของโมเลกุลสาร ถ้าโมเลกุลของสารที่เกาะติดกันเองได้ดีจะมีแรงดันไอต่ำ (มีแนวโน้มน้อยที่จะระเหยไปเป็นไอ) ในขณะที่สารที่เกาะติดกันไม่ดีจะมีความดันไอที่สูง

Antoine Equation เป็นสมการทางคณิตศาสตร์ที่แสดงความสัมพันธ์ระหว่างความดันไอและอุณหภูมิของสารบริสุทธิ์ที่เป็นของเหลวหรือของแข็ง ดังสมการ 2-1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\log P^{sat} = A - \frac{B}{T+C} \qquad (2\text{-}1)$$

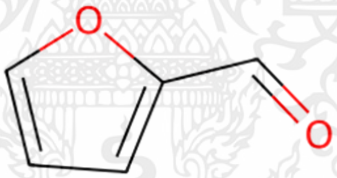
โดย P^{sat} แทนความดันไอสัมบูรณ์ของสารเคมี; T แทนอุณหภูมิของสารเคมี; A, B และ C เป็นค่า Antoine Coefficients ; และ log นิยมใช้เป็นลอการิทึมสามัญหรือลอการิทึมธรรมชาติ

ความดันไอของสารใดๆ จะเพิ่มขึ้นแบบไม่เป็นเส้นตรงตามอุณหภูมิตามสมการ Clausius-Clapeyron และจุดเดือดของของเหลวที่ความดันบรรยากาศ (จุดเดือดปกติ) คือ อุณหภูมิที่ความดันไอเท่ากับความดันบรรยากาศโดยรอบ เมื่ออุณหภูมิเพิ่มขึ้นเพียงเล็กน้อย แรงดันไอจะเพียงพอที่จะเอาชนะความดันบรรยากาศและทำให้ของเหลวเดือดกลายเป็นไอภายในเนื้อสาร การเกิดฟองอากาศในสภาวะที่สารเป็นของเหลวนั้น ต้องใช้ความดันไอที่สูงกว่าความดันอากาศ [14]

2.3 โครงสร้างของโมเลกุลสาร (Structure of Molecule)

2.3.1 Simplified Molecular-Input Line-Entry System (SMILES)

Simplified Molecular-Input Line-Entry System หรือ SMILES คือรูปแบบการเขียนโครงสร้างโมเลกุลสารเคมีให้อยู่ในรูปของข้อความของสัญลักษณ์ที่ทำให้คอมพิวเตอร์สามารถเข้าใจได้ ซึ่งเป็นรูปแบบการเขียนสามารถเรียนรู้ได้ง่ายและมีความยืดหยุ่นต่อรูปแบบสารเคมีที่หลากหลาย [15] เช่น SMILES ของ Furfural คือ c1cc(C=O)oc1 ตามรูปที่ 2.6



รูปที่ 2.6 Furfural, SMILES: c1cc(C=O)oc1

ในการเขียน SMILES ของสารเคมีใดๆนั้น สามารถใช้บ่งบอกลักษณะโครงสร้างภายในต่างๆได้หลักๆดังต่อไปนี้

1. อะตอม : SMILES จะแสดงอะตอมด้วยตัวอักษรตามดังในตารางธาตุ โดยอะตอมที่เกี่ยวข้องกับสารอินทรีย์จะไม่อยู่ใน [ ] ได้แก่ B, C, N, O, P, S, F, Cl, Br, และ I อะตอมที่ถูกเขียนมานั้นโดยปกติจะไม่มีการแสดงอะตอมของ H
2. พันธะ : SMILES จะแสดงถึงพันธะค่ (=) พันธะสาม (#) และโครงสร้างที่ไม่ได้เชื่อมกัน อย่างเช่น Ionic แต่ไม่นิยมแสดงพันธะเดี่ยว

เช่น CC (C-C, CH3-CH3), C=C (CH2=CH2), C#C (CH≡CH), CO (C-O, CH3-OH), C=O (H2C=O)

3. กิ่ง : SMILES ของโครงสร้างที่มีการต่อกันที่ไม่ใช่การต่อตรงเรื่อยๆ จะมีการเขียนสัญลักษณ์ของโครงสร้างที่ต่อแยกไว้ภาย ( )

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่น Isopropanol => CC(C)O, 2,2-Dimethylbutane => CC(C)(C)CC

4. วง : SMILES จะแสดงโครงสร้างที่เชื่อมกันเป็นวงด้วยตัวเลขหลังอะตอม การที่ในโครงสร้าง SMILES นั้นจะเชื่อมเป็นวงถึงกัน ต้องมีอะตอมสองตัวที่มีเลขต่อท้ายเหมือนกัน

เช่น Cyclopentane => C1CCCC1, Azetidine => C1CNC1

5. วงอะโรมาติก : SMILES จะแสดงโครงสร้างอะตอมที่อยู่ภายในวงอะโรมาติกด้วยตัวอักษรตัวพิมพ์เล็กทั้งหมด และเชื่อมกันเป็นวงด้วยตัวเลขหลังอะตอมที่เหมือนกัน

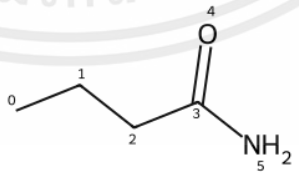
เช่น Benzene => c1ccccc1, Pyridine => n1ccccc1

2.3.2 ลายนิ้วมือโมเลกุล (Molecular Fingerprint)

ลายนิ้วมือโมเลกุล (Molecular Fingerprint) คือการแสดงลักษณะโครงสร้างของโมเลกุลโดยแสดงเป็นข้อความของบิต ซึ่งแต่ละบิตจะแสดงถึงการมีหรือไม่มีคุณลักษณะเฉพาะใดๆที่พบในโมเลกุล ลายนิ้วมือโมเลกุล (Molecular Fingerprint) มีรูปแบบการระบุโครงสร้างภายในของโมเลกุลที่แตกต่างกัน ซึ่งสามารถแบ่งออกเป็น 2 รูปแบบหลักๆ ต่อไปนี้ 1). รูปแบบที่กำหนดรูปแบบลักษณะโครงสร้างเฉพาะไว้ล่วงหน้า (Structural Keys) และ 2). รูปแบบที่ระบุลักษณะโครงสร้างที่แตกต่างกันภายในโมเลกุล (Generate Substructural Query) โดยวิธีการที่งานวิจัยนี้เลือกใช้จะเป็นรูปแบบที่ระบุลักษณะโครงสร้างที่แตกต่างกันภายในโมเลกุล ซึ่งมีรูปแบบวิธีการที่มีชื่อว่า ลายนิ้วมือมอร์แกน (Morgan Fingerprint)

ลายนิ้วมือมอร์แกน (Morgan Fingerprint) [16] เป็นหนึ่งในวิธีการในการระบุคุณลักษณะเฉพาะโครงสร้างย่อยภายในโมเลกุล โดยทำการวนซ้ำทั้งโครงสร้างเพื่อทำการระบุโครงสร้างที่พบที่ห่างออกจากอะตอมศูนย์กลางจากแต่ละตำแหน่งของโมเลกุลกับจำนวนอะตอมที่ห่างจากอะตอมศูนย์กลางดังกล่าวเป็นระยะรัศมีใดๆ ซึ่งสามารถแสดงได้ดังขั้นตอนการค้นหาโครงสร้างย่อยภายในได้ดังต่อไปนี้

1. กำหนดตำแหน่งของอะตอมทุกตัวภายในโมเลกุลที่ไม่ใช่ H

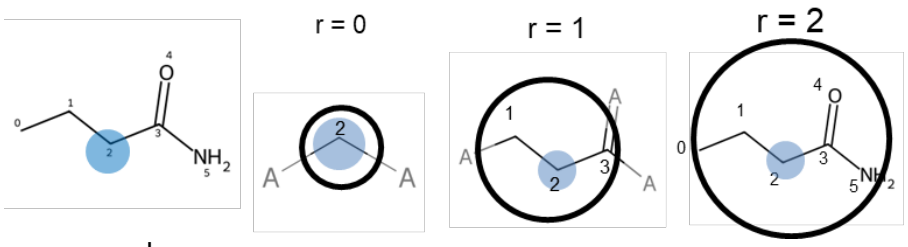


รูปที่ 2.7 การกำหนดตำแหน่งอะตอมภายในของ Butyramide

2. ระบุโครงสร้างที่อยู่ข้างเคียงอะตอมศูนย์กลางที่พิจารณาตามรัศมีที่กำหนด โดยเริ่มจากเลือกอะตอมตัวที่ระบุแล้วในขั้นตอนที่ 1 เป็นอะตอมศูนย์กลาง, พิจารณาอะตอมข้างเคียงโดยเริ่มจากรัศมีที่ 0 ถึงรัศมีที่กำหนด ซึ่งเป็นการเลือกอะตอมที่ห่างจากอะตอมศูนย์กลางตามรัศมี แล้วทำ

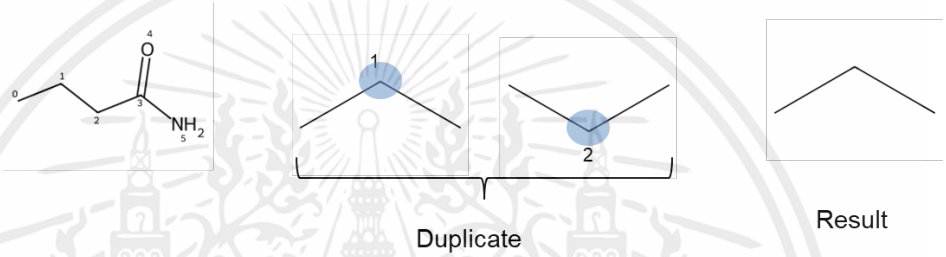
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กระบวนการดังกล่าวให้ทั่วทุกตำแหน่งอะตอมที่กำหนดไว้ภายในขั้นตอนที่ 1 รูปที่ 2.8 เป็นตัวอย่างในการระบุโครงสร้างย่อยรอบอะตอมศูนย์กลางตำแหน่งที่ 2 โดยมีรัศมี 2



รูปที่ 2.8 การระบุโครงสร้างย่อยภายในของ Butyramide

3. ทำการกำจัดโครงสร้างที่ซ้ำกันและจัดเก็บโครงสร้างย่อยที่พบทั้งหมดในลายนิ้วมือโมเลกุลของโมเลกุลสารดังกล่าว



รูปที่ 2.9 การกำจัดโครงสร้างย่อยที่ซ้ำกันของ Butyramide

Morgan Fingerprint สามารถแบ่งออกได้เป็น 2 ประเภทได้แก่ Binary Morgan Fingerprint (B-MF) และ Count-based Morgan Fingerprint (C-MF) [17] โดย B-MF เป็นลายนิ้วมือโมเลกุลที่แต่ละบิตนั้นแสดงถึงการมีคุณลักษณะเฉพาะดังกล่าว(โครงสร้างย่อย)ภายในโมเลกุล ในขณะที่ C-MF เป็นลายนิ้วมือโมเลกุลที่แต่ละบิตนั้นจะแสดงถึงการมีอยู่และจำนวนของคุณลักษณะเฉพาะ(โครงสร้างย่อย)ที่มีอยู่ในโมเลกุล

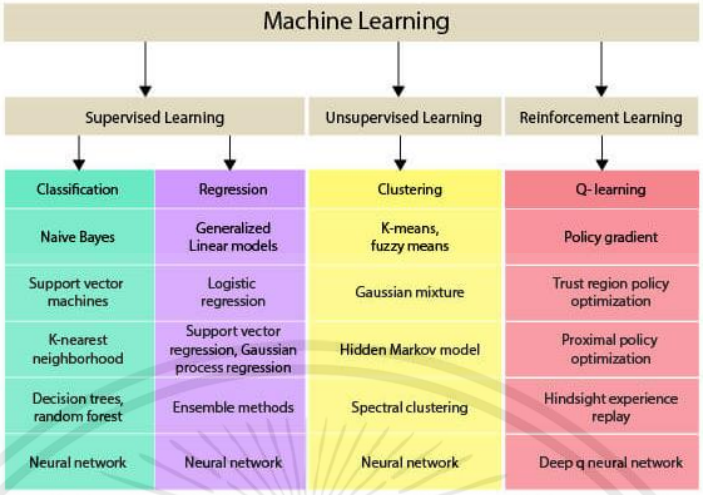
ในการศึกษา Morgan Fingerprint นั้นใช้โปรแกรมภาษาไพทอน (Python) โดยมีเครื่องมือ (Library) ที่มีชื่อว่า RDKit ซึ่งเป็นเครื่องมือที่ทำให้สามารถทำความเข้าใจโครงสร้างสารเคมีได้ (Cheminformatics) นำมาใช้ในการแบ่งโครงสร้างย่อยภายในโมเลกุลสารเป็น Molecular Fingerprint จากสารเคมีที่ป้อนเข้าข้อมูลโครงสร้างด้วย SMILES

2.4 การเรียนรู้ของเครื่อง (Machine Learning)

การเรียนรู้ของเครื่อง คือการที่ทำให้คอมพิวเตอร์เรียนรู้จากข้อมูลได้ด้วยตนเอง การเรียนรู้ของเครื่องเป็นส่วนหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence, AI) แบ่งออกเป็น 3 รูปแบบ คือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นการเรียนของเครื่องโดยรู้อาศัยข้อมูลที่มีการระบุผลเฉลยไว้แล้ว การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เป็นการเรียนรู้ของเครื่องโดยไม่มีการระบุคุณลักษณะของข้อมูล จะให้เรียนรู้ผ่านข้อมูลที่มีอยู่ และการเรียนรู้ตาม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

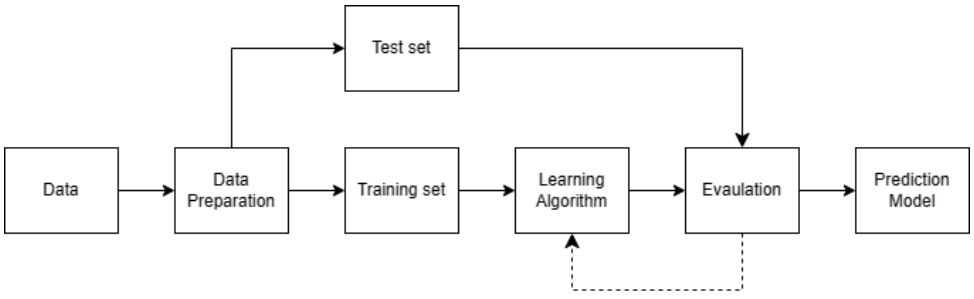
สภาพแวดล้อม (Reinforcement Learning) เป็นการเรียนรู้แบบลองผิดลองถูกเองภายใต้สถานการณ์ต่างๆเพื่อให้ระบบตัดสินใจ [18] สามารถแสดงได้ดังรูปที่ 2.10



รูปที่ 2.10 แสดงรูปแบบและชนิดการเรียนรู้ของ Machine Learning

ซึ่งในงานวิจัยในครั้งนี้ใช้วิธีการการเรียนรู้แบบมีผู้สอน (Supervised Learning) ซึ่งสามารถแบ่งออกได้เป็น 2 ประเภท ได้แก่ การวิเคราะห์แบบถดถอย (Regression) เป็นการทำให้คอมพิวเตอร์เรียนรู้เพื่อทำนายผลเฉลยออกมาเป็นตัวเลข และ การจำแนกหมวดหมู่ (Classification) เป็นการทำให้คอมพิวเตอร์เรียนรู้เพื่อทำนายผลเฉลยออกมาเป็นหมวดหมู่

การใช้เทคนิคแบบมีผู้สอนมีขั้นตอนดังนี้ โดยเริ่มจากการเก็บข้อมูล (Data Set) จากนั้นจะมีการจัดเตรียมข้อมูลที่เก็บมาได้ (Data Preparation) ซึ่งจะต้องมีการทำความสะอาดข้อมูลและระบุคุณลักษณะ (Features) ของข้อมูลให้เรียบร้อย ก่อนจะแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลชุดฝึก (Training Set) จะถูกใช้ในการสอนผ่านอัลกอริทึมในขั้นตอน Learning Algorithm ให้มีความเข้าใจในข้อมูลเพื่อสร้างแบบจำลองในการทำนาย และ ข้อมูลชุดทดสอบ (Test Set) ใช้ในทดสอบประสิทธิภาพของแบบจำลองที่ได้ (Evaluation) ซึ่งถ้าการประเมินประสิทธิภาพของแบบจำลองไม่เป็นที่พึงพอใจ สามารถปรับในส่วนของ Learning Algorithm ได้ถ้าประสิทธิภาพของแบบจำลองเป็นที่พึงพอใจแล้ว สามารถนำไปใช้ในเป็นแบบจำลองในการทำนายได้ (Prediction Model) สามารถแสดงได้ดัง**รูปที่ 2.11**



รูปที่ 2.11 แผนผัง Machine Learning แบบง่าย

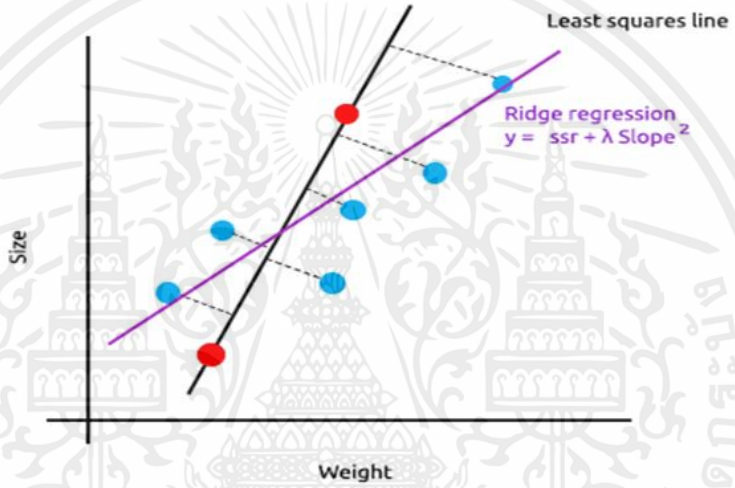
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.1 อัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning Algorithm)

คือ รูปแบบวิธีการในการเรียนรู้ข้อมูลของเครื่องเพื่อสร้างแบบจำลองต่างๆ โดยจะมีอัลกอริทึม (Algorithm) ที่หลากหลายดังต่อไปนี้

1. Ridge Regression (Ridge)

เป็นวิธีการที่คล้ายคลึงกับ Linear Regression ซึ่งเป็นการทำนายความสัมพันธ์ระหว่างข้อมูลเป็นเส้นตรง วิธีการนี้ปรับปรุงการหาความคลาดเคลื่อนของการทำนายให้ไม่ขึ้นกับข้อมูลที่ใช้ในการเรียนรู้มากเกินไป ทำให้สามารถรับมือกับข้อมูลใหม่ๆที่ไม่เคยเห็นได้ดีกว่า Linear Regression [19] โดย Linear regression จะลด SSR ให้ต่ำที่สุด แต่ Ridge Regression จะลดผลรวมระหว่าง SSR กับ พจน์ที่เกี่ยวข้องกับข้อมูลที่เรียนรู้ไปก่อนหน้า

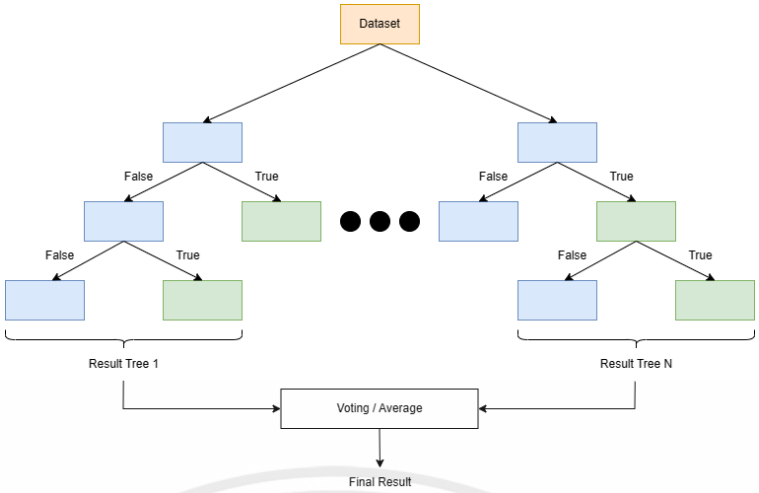


รูปที่ 2.12 Ridge Regression

2. Random Forest (RF)

เป็นวิธีการที่พัฒนาจาก Decision Tree โดย “ต้นไม้”ที่ได้จาก Decision Tree แต่ละครั้งนั้นจะแตกต่างกันไปตามข้อมูลที่ใช้ในการเรียนรู้ วิธีการนี้จึงได้ทำการสร้าง ต้นไม้ หลายๆต้น ให้เรียนรู้พร้อมๆกัน เพื่อลดความเคลื่อนในการทำนายข้อมูลใหม่ด้วยการสร้าง “ต้นไม้” ที่มีเงื่อนไขที่ใช้ในการแบ่งข้อมูลที่หลากหลายมากขึ้น สุดท้ายจะนำผลลัพธ์ที่ได้จากแต่ละ “ต้นไม้” มาเฉลี่ยหรือโหวตเพื่อหาค่าผลลัพธ์สุดท้ายที่ทำนายออกมาได้ [20]

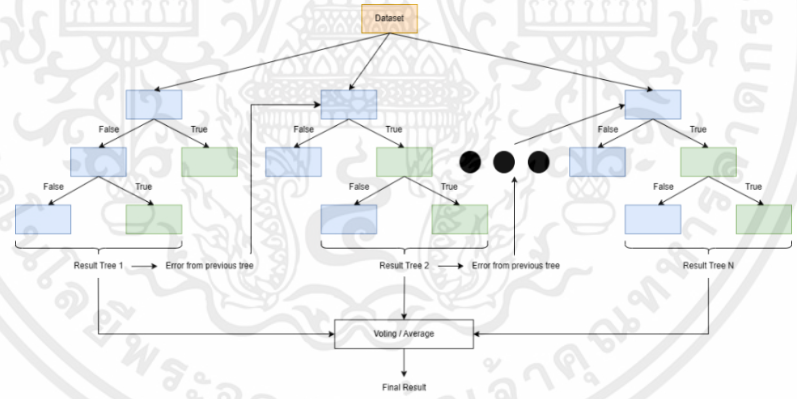
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 Random Forest

3. Extreme Gradient Boosting (XGBoost)

เป็นวิธีการที่พัฒนาจาก Gradient Boosting ซึ่งมีพื้นฐานจาก Decision Tree แต่ปรับปรุงให้ “ต้นไม้” จาก Decision Tree ให้เรียนรู้จากข้อผิดพลาดของ ต้นไม้ก่อน แล้วสร้างต้นไม้ใหม่ที่มีเงื่อนไขในการแบ่งข้อมูลที่ดีขึ้นจากต้นไม้ก่อนๆสุดท้ายจะนำผลทำนายของต้นไม้แต่ละต้นที่ได้ชั่งน้ำหนักเพื่อเฉลี่ยหรือโหวตว่าผลลัพธ์จากการทำนายเป็นอย่างไร โดยวิธีการนี้ทำเพื่อเพื่อลด Overfitting จากข้อมูลที่ได้เรียนรู้ ซึ่งเป็นข้อเสียของ Decision Tree [20,21]

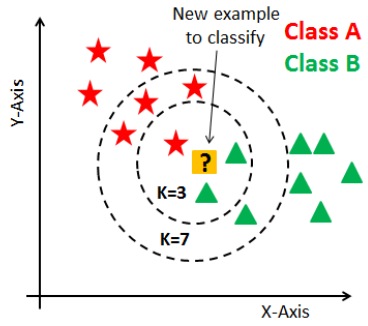


รูปที่ 2.14 Extreme Gradient Boosting

4. K-Nearest Neighbors

เป็นวิธีการวิเคราะห์ข้อมูลโดยอาศัยข้อมูลเพื่อนบ้านที่ใกล้ที่สุด K จุด ตามที่กำหนด โดยค่าของข้อมูลใหม่จะถูกทำนายโดยการหาค่าเฉลี่ยของระยะห่างกับเพื่อนบ้าน ที่เรียกว่า Euclidean Distance หรือกลุ่มของข้อมูลใหม่จะถูกทำนายจากการดูจำนวนข้อมูลบริเวณรอบ ๆ (Nearest Neighbor) ว่ามีกลุ่มไหนจำนวนเยอะกว่ากัน แล้วจึงเลือกกลุ่มที่มีจำนวนเยอะกว่า [22,23]

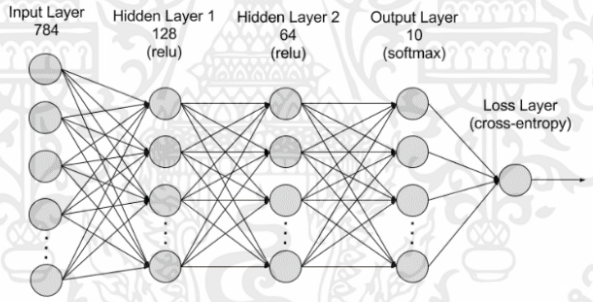
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 K-Nearest Neighbors

5. **Deep Learning (Neuron Network)**

เป็นวิธีการที่ทำให้คอมพิวเตอร์สามารถคิดได้แบบมนุษย์ โดยมีแรงบันดาลใจจากการทำงานของระบบประสาทในสมองมนุษย์ ซึ่งประกอบไปด้วยนิวรอน (Neurons) ที่เชื่อมต่อกันเป็นโครงข่ายจะมีทั้งหมด 3 ชั้นโดยการทำงานจะเริ่มจากการ ชั้นขาเข้า (Input Layer) จากนั้นส่งต่อไปยัง ชั้นซ่อน (Hidden Layer) ซึ่งจะประกอบไปด้วยนิวรอนจำนวนมาก นิวรอนจะประมวลผลและส่งต่อไปยังนิวรอนตัวอื่นๆในชั้นถัดไป กระบวนการนี้จะดำเนินต่อไปจนถึง ชั้นขาออก (Output Layer) จะแสดงผลลัพธ์ของการประมวลผล [24]

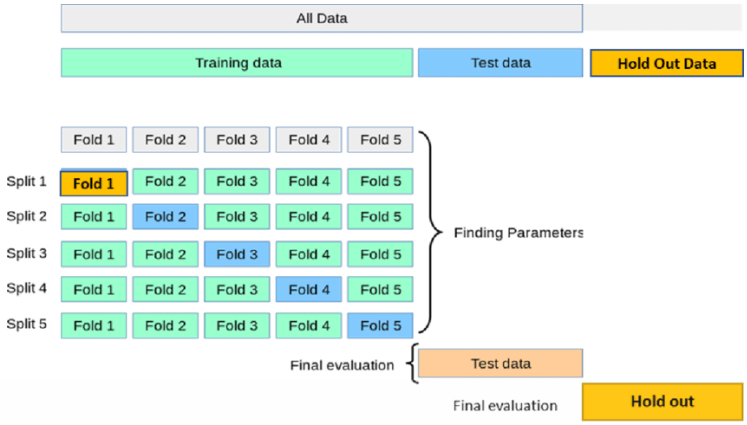


รูปที่ 2.16 Deep Learning

2.4.2 Cross Validation: K-fold

Cross Validation เป็นขั้นตอนในการเพิ่มความเชื่อมั่นว่าผลลัพธ์จากแบบจำลองที่ได้มาจาก Machine Learning Algorithm โดยในส่วนของ K-Fold นั้นจะทำการแบ่งข้อมูลเป็นจำนวน k ส่วน ซึ่งเอาไว้ใช้สำหรับข้อมูลสำหรับการเรียนรู้กับข้อมูลทดสอบ โดยเริ่มต้นจะใช้ข้อมูลชุดที่ 1 มาใช้ทดสอบ และข้อมูลส่วนที่เหลือจะถูกใช้ไปกับการเรียนรู้ทำแบบนี้ โดยเปลี่ยนให้ชุดข้อมูลทดสอบเป็น 2 ไปเรื่อยๆจนถึงชุดที่ K จะทำให้ได้ Model จำนวน K ซึ่งเป็นการทำให้ Machine Learning สามารถทำนายจากรูปแบบข้อมูลที่หลากหลายมากขึ้น สุดท้ายจะนำผลทำนายจากแต่ละ Model มาเฉลี่ยหาค่าที่ดีที่สุดออกมา [25] สามารถแสดงได้ดังรูปที่ 2.17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.17 K-fold Cross Validation

2.5 การประเมินผลแบบจำลอง (Model Evaluation)

หลังจากแบบจำลองได้เรียนรู้เรียบร้อย จะนำแบบจำลองที่ได้ทำนายผลออกมาด้วยชุดข้อมูล Test Set และเนื่องจากโจทย์ปัญหาในการทำแบบจำลองทำนายจุดเดือดนี้นั้น เป็นโจทย์ปัญหาแบบ Regression ซึ่งผลทำนายที่ได้ออกมาจากแบบจำลองจะเป็นเป็นตัวเลข ทำให้ต้องใช้ Mean Absolute Error (MAE) Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) และ Coefficient of Determination (R^2) ประเมินประสิทธิภาพของ Machine Learning Model โดยมีรายละเอียดดังต่อไปนี้

กำหนดให้ $\hat{y}_i$ คือผลการทำนายของค่า y

y_i คือค่าจริง

จะได้ว่า

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3-1}$$

มีขอบเขตของค่าตั้งแต่ 0-∞ ตัวอย่าง Error ที่ได้สามารถบอกได้ว่า สมบัติสารที่ทำนายได้ห่างจากค่าจริงไปกี่องศา โดยค่าที่ได้ยิ่งน้อยจะยิ่งแสดงว่าแบบจำที่ได้มีความแม่นยำมาก

$$\%\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100 \tag{3-2}$$

ปกตินั้นมีขอบเขตของค่าตั้งแต่ 0-100 แต่สามารถมีขอบเขตได้ตั้งแต่ 0-∞ ตัวอย่าง Error ที่ได้สามารถบอกได้ว่า สมบัติสารที่ทำนายได้ห่างจากค่าจริงไปกี่ % โดยค่าที่ได้ยิ่งน้อยจะยิ่งแสดงว่าแบบจำที่ได้มีความแม่นยำมาก

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(\hat{y}_i - y_i)^2}{n}} \tag{3-3}$$

มีขอบเขตของค่าตั้งแต่ 0-∞ ตัวอย่าง Error ที่ได้สามารถบอกได้ว่า สมบัติสารที่ทำนายได้ห่างจากค่าจริงไปกี่ % โดยค่าที่ได้ยิ่งน้อยจะยิ่งแสดงว่าแบบจำที่ได้มีความแม่นยำมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\sum_{i=1}^{n}(y_i - \bar{y})} \quad (3\text{-}4)$$

มีขอบเขตของค่าตั้งแต่ 0-1 ซึ่งค่าที่ใช้พิสูจน์ว่าแบบจำลองที่ได้นั้นเหมาะสมหรือไม่ ซึ่งยิ่งเข้าใกล้ 1 ยิ่งดี โดยทั่วไปควรมีค่ามากกว่า 0.6 จะถือว่าแบบจำลองที่ได้เป็นแบบจำลองที่ดี [26]

2.6 งานวิจัยที่เกี่ยวข้อง (Related Work)

2.6.1 การใช้ลายนิ้วมือโมเลกุล และ Machine Learning เพื่อพัฒนาแบบจำลองทำนายสมบัติของ Ionic Liquid

จากงานวิจัย Yi Ding และคณะ [27] สร้างแบบจำลองทำนาย Viscosity ของ Ionic Liquid โดยมีการประยุกต์ลายนิ้วมือโมเลกุลมอร์แกนกับ Machine Learning มีจำนวน 1,502 ตัวโดยมีจำนวน Ionic Liquid ทั้งหมด 89 ชนิด โดยที่มีขอบเขตของ Viscosity เป็น 8.28–142,000 cP ซึ่งได้ใช้ลายนิ้วมือมอร์แกนที่มีรัศมีถึง 7 และจำนวนบิตเป็น 4,632 บิต โดยใช้ Machine Learning Algorithm เป็น MLR, SVM, XGBoost โดยมีผลลัพธ์ดังตารางต่อไปนี้

ตารางที่ 2.1 การทดสอบประสิทธิภาพของแบบจำลองสำหรับ Ionic ของ Yi Ding และคณะ

Machine Learning Algorithm	MSE (log of Viscosity)	R^2 (log of Viscosity)
MLR	0.187	0.80
SVM	0.025	0.93
XGBoost	0.0091	0.97

งานวิจัยนี้สรุปว่าแบบจำลองระหว่าง Morgan Fingerprint-XGBoost เป็นวิธีที่มีประสิทธิภายในการพัฒนา QSAR model เพื่อใช้ทำนายความหนืดของ Ionic Liquid โดยวิธีการนี้สามารถสร้างข้อมูลที่สื่อถึงโครงสร้างและสามารถนำไปประยุกต์กับสภาวะของ Ionic Liquid ได้ง่าย

2.6.2 วิธีการ Count-base Morgan Fingerprint ที่ใช้สำหรับบ่งบอกโครงสร้างโมเลกุลในการพัฒนา Machine Learning สำหรับการทำนายสมบัติของน้ำที่ปนเปื้อน

จากงานวิจัยของ Shifa Zhong และ คณะ [17] ได้ศึกษาการประยุกต์และเปรียบเทียบลายนิ้วมือประเภท Count-Based Morgan Fingerprint (C-MF) กับ Binary Morgan Fingerprint (B-MF) โดยใช้ Machine Learning Algorithm ได้แก่ Ridge, SVM, KNN, RF, XGBoost, CatBoost ด้วยชุดข้อมูลที่หลากหลายเช่น Solubility จำนวน 1,395 ชนิด HO• จำนวน 1,374 ชนิด เป็นต้น ได้ผลลัพธ์ที่ดีที่สุดต่างแต่ละชุดข้อมูลดังตารางต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 ข้อมูลสมบัติที่เกี่ยวข้องกับน้ำที่ปนเปื้อนของ Shifa Zhong และ คณะ

Properties	Name of Properties
Solubility	Solubility Of Contaminants In Water
HO• / $k_{•HO}$	Second-Order Rate Constants of Contaminants Toward Oxidants of HO•
$SO_4^{•-}$ / $k_{SO_4^{•-}}$	Second-Order Rate Constants of Contaminants Toward Oxidants of $SO_4^{•-}$
K_{oc}	The Soil Organic Carbon-Normalized Sorption Coefficient of Chemicals
pK_d	Binding Affinity Data on The Dissociation Constant and Inhibition Constant
pIC_{50}	The Concentration for the 50% Maximum Inhibition
CCS_MH	Collision Cross Sections For[M+H]+
CCS_MNa	Collision Cross Sections For[M+Na]+
Lipo	Lipophilicity
FreeSolv	Free Solvation Database

ตารางที่ 2.3 การทดสอบประสิทธิภาพลายนิ้วมือของแต่ละชุดข้อมูลสำหรับน้ำปนเปื้อนของ Shifa Zhong และ คณะ

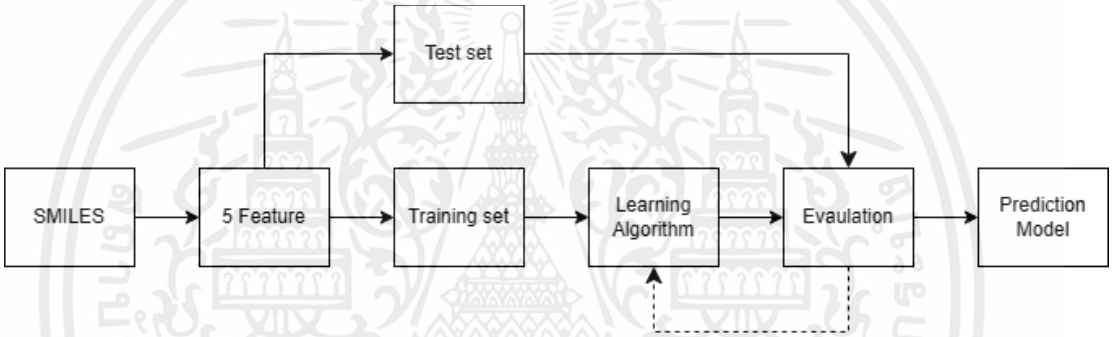
Dataset/Properties	Machine Learning Algorithm	Fingerprint Type	RMSE	R^2
Solubility	CatBoost	C-MF	0.692±0.033	0.909±0.009
HO• / $k_{•HO}$	CatBoost	C-MF	0.089±0.007	0.640±0.041
$SO_4^{•-}$ / $k_{SO_4^{•-}}$	CatBoost	B-MF	0.112±0.013	0.514±0.092
K_{oc}	Ridge	C-MF	0.507±0.019.	0.830±0.022
pK_d	CatBoost	C-MF	0.729±0.019	0.577±0.064
pIC_{50}	CatBoost	C-MF	0.636±0.025	0.675±0.023
CCS_MH	Ridge	C-MF	5.285±0.458	0.976±0.004
CCS_MNa	Ridge	C-MF	6.480±1.095	0.952±0.021
Lipo	Ridge	C-MF	0.683±0.020	0.676±0.018
FreeSolv	Ridge	C-MF	1.230±0.121	0.893±0.027

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากงานวิจัยแสดงให้เห็นว่า จากชุดข้อมูลทั้งหมด 10 ชุดนั้น C-MF มีประสิทธิภาพดีกว่าการใช้ B-MF อยู่ 9 ชุดข้อมูล โดยผลลัพธ์จากการตีความที่ได้จากแบบจำลองนั้นสามารถชี้ให้เห็นถึงผลจากจำนวนของกลุ่มอะตอม(โครงสร้างย่อย)ต่อผลลัพธ์ที่ทำนายได้ ตัวอย่างเช่น Feature_361 ในแบบจำลอง B-MF และ Feature_1673 ในแบบจำลอง C-MF ซึ่งบ่งบอกถึงหมู่ฟังก์ชั่น -OH โดยที่เมื่อหมู่ -OH เพิ่มขึ้นจะทำให้ค่าของสมบัติที่ต้องการทำนายมีค่าเพิ่มขึ้น โดยที่ C-MF สามารถระบุปรากฏการณ์ดังกล่าวได้ในขณะที่ B-MF ไม่สามารถอธิบายได้

2.6.3 การใช้การเรียนรู้ของเครื่องเพื่อพัฒนาแบบจำลองทำนายสมบัติของสารอินทรีย์บริสุทธิ์

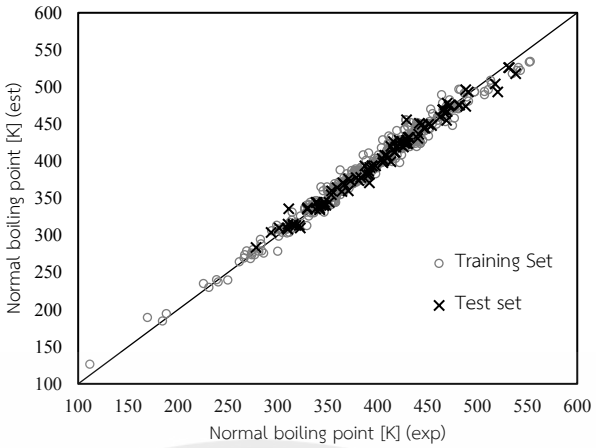
จากงานวิจัย Nattasinee และคณะ [7] แบบจำลองทำนายจุดเดือดที่ได้ศึกษามานั้น มีแผนผังในการสร้างแบบจำลองทำนายจุดเดือดของไฮโดรคาร์บอนได้ดังต่อไปนี้



รูปที่ 2.18 แผนผัง Machine Learning ฉบับงานวิจัยของ Nattasinee และคณะ

มีกำหนดคุณลักษณะ (Feature) จาก SMILES ซึ่งได้แก่ จำนวน C, Double Bond, Triple Bond, Branch และ Cyclic, มีการกำหนดตัวแปรตาม (Label) คือ T_b (Boiling Point) มีขอบเขตการศึกษาเป็น Hydrocarbon ที่มีอะตอม C, H เท่านั้น มีจำนวนข้อมูล 560 ชนิด และใช้ Machine Learning Algorithm เป็น Regression Algorithm ที่พัฒนาขึ้นมาใช้เอง มีผลการทำนายของแบบจำลองมีผลดังกราฟและตารางต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.19 กราฟแสดงประสิทธิภาพการทำนาย Normal Boiling Point ของงานวิจัยของ Nattasinee และคณะ

ตารางที่ 2.4 การประเมินประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point ของงานวิจัยของ Nattasinee และคณะ

Dataset	MAPE	RMSE	R^2
Train Set	0.016 (1.6%)	8.354	0.980
Test Set	0.015 (1.5%)	7.924	0.984
Total Set	0.016 (1.6%)	8.340	0.981

งานวิจัยนี้มุ่งเน้นที่การพัฒนาวิธีการใส่คุณลักษณะ (Feature) ให้กับแบบจำลองเช่นเดียวกันกับการปรับปรุงประสิทธิภาพแบบจำลองทำนายสมบัติ โดยมีปัญหาคือ การที่ไม่สามารถแยกความแตกต่างระหว่างโมเลกุลที่มีความคล้ายคลึงกัน ซึ่งเป็นปัญหาจากการใส่คุณลักษณะ (Feature) ดังกล่าวซึ่งได้แก่ จำนวนของ C, Double Bond, Triple Bond, Branch, Cyclic มีผลลัพธ์ของปัญหาดังตารางต่อไปนี้

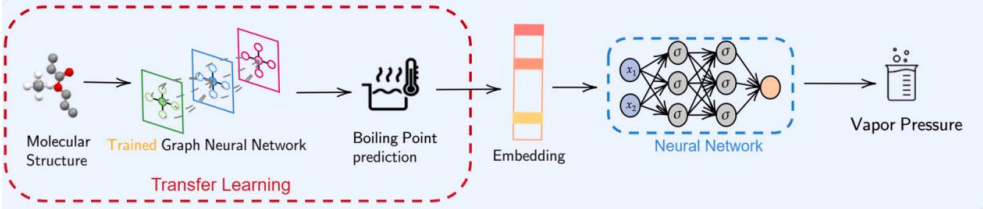
ตารางที่ 2.5 ผลลัพธ์การเปลี่ยน SMILES เป็นข้อมูลสำหรับการเรียนรู้และผลการทำนาย Normal Boiling Point ของ งานวิจัยของ Nattasinee และคณะ

SMILES	C	Double	Triple	Branch	Cyclic	T_b (exp)	T_b (test)
C1CCC=CCC1	7	1	0	0	1	388.15	375.85
CC1=CCCCC1	7	1	0	0	1	383.45	375.85
CC1CCC=CC1	7	1	0	0	1	375.85	375.85
CC1CCCC=C1	7	1	0	0	1	376.15	375.85
CCC1=CCCC1	7	1	0	0	1	379.45	375.85
CCC1CCC=C1	7	1	0	0	1	370.95	375.85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

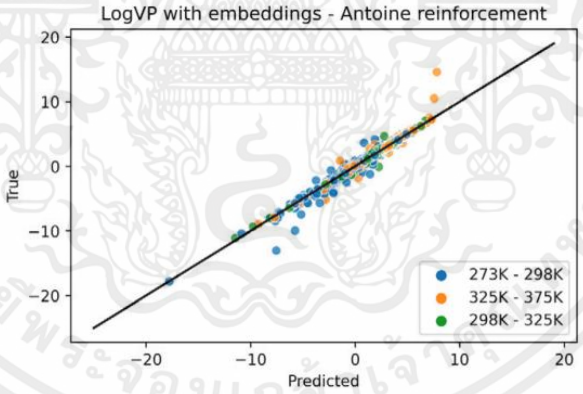
2.6.4 การทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN (Path-Unifying Feed-Forward Interfaced Network)

จากงานวิจัย Santana และ คณะ [28] แบบจำลองทำนายความดันไอที่ได้ศึกษามานั้น มีแผนผังในการสร้างแบบจำลองทำนายความดันไอของไฮโดรคาร์บอนที่มีผลลัพธ์ที่ดีที่สุดได้ดังต่อไปนี้



รูปที่ 2.20 แผนผังแสดงการสร้างแบบจำลองทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN ของงานวิจัย Santana และ คณะ

ขั้นตอนแรกมีการการคุณลักษณะจาก SMILES โดยใช้ข้อมูลกราฟของโมเลกุลสารดังกล่าวแปลงเป็นข้อมูลตัวเลขผ่าน Graph Neural Network ผนวกรวมเข้ากับอุณหภูมิของสารดังกล่าว มีการกำหนดตัวแปรตาม (Label) คือ P^{sat} ซึ่งอยู่ในรูปของ $\ln(P^{sat})$ ในหน่วย Pa มีขอบเขตการศึกษาเป็น Hydrocarbon ที่มีสถานะเป็นของเหลวตามสภาวะ มีจำนวนข้อมูล 1,851 molecules และใช้ Machine Learning Algorithm ที่ซับซ้อนมากกว่าปกติซึ่งได้แก่ Neural Network มีผลลัพธ์ของแบบจำลองดังต่อไปนี้



รูปที่ 2.21 กราฟแสดงประสิทธิภาพการทำนายความดันไอจาก Neural Network รูปแบบ PUFFIN ของงานวิจัย Santana และ คณะ

ผลการประเมินแบบจำลองในการทำนาย $\ln(P^{sat})$ ด้วยค่า MSE อยู่ที่ 0.1609 Pa งานวิจัยนี้มุ่งเน้นที่การพัฒนาแบบจำลองที่สามารถทำนายความดันไอได้ใกล้เคียงมากที่สุดจึงได้มีการใช้ Neural Network เข้ามาเพื่อแปลงข้อมูลโมเลกุลสารดังกล่าว ซึ่งผลจากการแปลงดังกล่าว ทำให้ผู้สร้างแบบจำลองเองก็ไม่สามารถเข้าคุณลักษณะที่ใส่เข้าไปใน Machine Learning Algorithm ได้ ถึงแม้ว่าจะมีความแม่นยำมากก็ตาม อันเนื่องมากจากคุณลักษณะที่ป้อนเข้าไปใน Neural Network มากจากการเปลี่ยนข้อมูลลักษณะโมเลกุลสารที่เป็นแบบกราฟ (Graph) ให้เป็นตัวเลขเพื่อนำไปหาความสัมพันธ์ต่อใน Neural Network

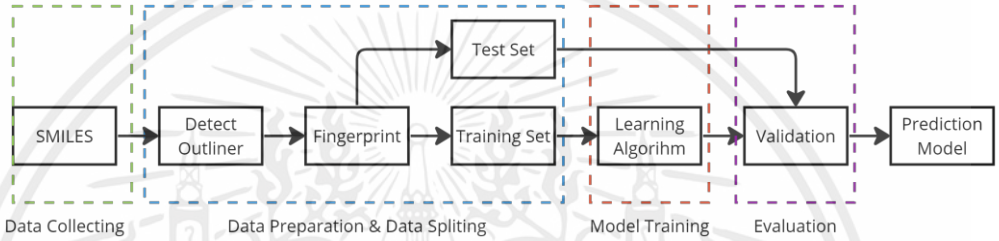
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงาน

ในงานวิจัยนี้มีกระบวนการขั้นตอนการดำเนินงานแบ่งออกเป็น 2 ส่วน ซึ่งเป็นขั้นตอนสำหรับกระบวนการสร้างแบบจำลองทำนายจุดเดือดกับความดันไอ โดยแต่ละส่วนมีรายละเอียดของ 3 ขั้นตอนย่อยแสดงตามแผนภาพในรูปที่ 3.1 ซึ่งแบ่งไปได้ต่อไปนี้

1. การเก็บรวบรวมข้อมูล (Data Collecting)
2. การเตรียมข้อมูลและการแบ่งข้อมูล (Data Preparation & Data Splitting)
3. การฝึกฝนโมเดล (Model Training)



รูปที่ 3.1 แผนผังการทำงานภายในการพัฒนา Propertes Prediction Model จาก Machine Learning

3.1 การเก็บรวบรวมข้อมูล (Data Collecting)

เริ่มต้นนำเข้าข้อมูลจากงานวิจัย Nattasinee และคณะ [7] ที่มีขอบเขตเป็นไฮโดรคาร์บอนที่มีอะตอมเพียงคาร์บอนและไฮโดรเจนเท่านั้นจำนวน 560 ชนิด และมีการนำเข้าข้อมูลเพิ่มเติมเพื่อขยายขอบเขตเป็นสารอินทรีย์ที่มีอะตอมคาร์บอน ไฮโดรเจน ออกซิเจนและไนโตรเจนจากแหล่งข้อมูล Chemical Database of Chemical Engineering Design Library (ChEDL) [29,30] ที่เป็นแหล่งข้อมูลที่รวบรวมข้อมูลจากแหล่งข้อมูลดังตารางที่ 3.1

ตารางที่ 3.1 แหล่งข้อมูลจาก Chemical Database of Chemical Engineering Design Library (ChEDL)

แหล่งที่มา	อ้างอิง	สมบัติสาร	จำนวนข้อมูล
CRC Handbook of Chemistry and Physics	[30]	จุดเดือด	5,542
CAS Common Chemistry	[30]		10,419
NIST Webbook	[30]		5,847
Wikidata	[30]		872
Yaws, “Thermophysical Properties of Chemicals and Hydrocarbons”	[30]		13,461
Joback, “Estimation of Pure-Component Properties from Group-Contributions”	[30]		23,068

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

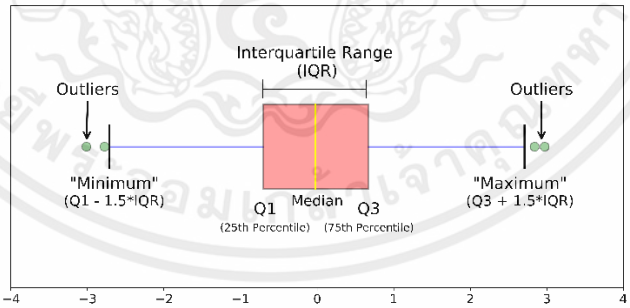
แหล่งที่มา	อ้างอิง	สมบัติสาร	จำนวนข้อมูล
Hall, K. R. Vapor Pressure and Antoine Constants for Hydrocarbons, and S, Se, Te, and Halogen Containing Organic Compounds. Springer, 1999.	[29]	ความดันไอ	6,346
Dykyj, J., and K. R. Hall. “Vapor Pressure and Antoine Constants for Oxygen Containing Organic Compounds”. 2000.	[29]		
Hall, K. R. Vapor Pressure and Antoine Constants for Nitrogen Containing Organic Compounds. Springer, 2001.	[29]		

3.2 การเตรียมข้อมูลและการแบ่งข้อมูล (Data Preparation & Data Splitting)

3.2.1 การเตรียมข้อมูลและการแบ่งข้อมูลของจุดเดือด

โดยจะทำการคัดกรองข้อมูลที่มีสมบัติที่ต้องการซึ่งได้แก่ จุดเดือด จากแต่ละแหล่งข้อมูล ทำการกำจัดข้อมูลที่ซ้ำกัน โดยข้อมูลที่ได้จะมีข้อมูลบางส่วนที่ไม่มี SMILES แต่จะมี CAS Number ซึ่งต้องใช้ Library ที่มีชื่อว่า PubChem ในการแปลงข้อมูล CAS Number เป็น SMILES

หลังจากนั้นคัดกรองข้อมูลที่ไม่เหมาะสมก่อนจะนำเข้าไปสู่การทำ Machine Learning ด้วยการใช้ Boxplot ซึ่งทำการกำจัดข้อมูลโดยใช้จุดเดือดเป็นเกณฑ์ ตัดข้อมูลที่ออกนอกช่วงข้อมูลที่ปกติ (นอก 1.5*IQR คิดเป็น 99.3%) แสดงได้ดังรูปที่ 3.2 ทำให้เหลือช่วงจุดเดือดอยู่ที่ 266.25 ถึง 682.52 K ซึ่งมีจำนวนข้อมูลทั้งสิ้น 11,177 ตัว โดยตารางข้อมูลสุดท้ายที่ได้ประกอบไปด้วย ชื่อสาร, SMILES และจุดเดือด



รูปที่ 3.2 Boxplot แสดงการกระจายตัวของช่วงข้อมูลปกติและนอกช่วงข้อมูลปกติ

หลังจากที่ได้ตารางข้อมูลแล้วจะทำการใช้เครื่องมือสำเร็จรูป (Library) ที่มีชื่อว่า RDKit เพื่อเก็บข้อมูลของโมเลกุลสารจากการป้อน SMILES เข้าไป หลังจากนั้นนำโมเลกุลสารที่ได้แปลงโครงสร้างโมเลกุลให้กลายเป็น Count-Based Morgan Fingerprint ด้วยชุดคำสั่ง “rdMolDescriptors.GetHashedMorganFingerprint()” โดยจะสามารถกำหนดรัศมีและจำนวนบิตซึ่งด้วยการทดลองปรับเปลี่ยนด้วยการตั้งค่าต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 แสดงการตั้งค่าเพื่อหาลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการทำแบบจำลองทำนายจุดเดือด

ตัวแปร	รายละเอียด
ตัวแปรควบคุม	ML Algorithm : XGBoost ที่มีการแบ่งข้อมูล : 90:10
ตัวแปรที่เปลี่ยนแปลง	MF_Radius = 2-5 MF_Bits = 512-4,096

หลังจากที่ได้ Fingerprint ที่เหมาะสมมาแล้วซึ่งเป็นตารางข้อมูลที่มีคุณลักษณะพร้อมแล้วจะทำการแบ่งชุดข้อมูลเป็นข้อมูลที่ใช้ในการเรียนรู้:ข้อมูลที่ใช้ในการทดสอบ เป็นอัตราส่วนจาก 80:20 หรือ 90:10 ขึ้นกับประสิทธิภาพแบบจำลองทำนายสมบัติที่ได้ออกมา

3.2.2 **การเตรียมข้อมูลและการแบ่งข้อมูลของความดันไอ**

โดยจะทำการคัดกรองข้อมูลที่มีสมบัติที่ต้องการซึ่งได้แก่ ความดันไอ จากแต่ละแหล่งข้อมูล ทำการกำจัดข้อมูลที่ซ้ำกัน โดยข้อมูลที่ได้จะไม่มี SMILES แต่จะมีเลข CAS ชื่อสาร ค่า A B C T_{min} และ T_{max} ตามลำดับซึ่งต้องใช้เครื่องมือสำเร็จรูปที่มีชื่อว่า PubChem ในการแปลงชื่อสารเป็น SMILES

หลังจากนั้นคัดกรองข้อมูลที่ไม่เหมาะสมก่อนจะนำเข้าไปสู่การทำ Machine Learning ด้วยการใช้ Boxplot โดยใช้ค่าความดันไอเป็นเกณฑ์ ตัดข้อมูลที่ออกนอกช่วงข้อมูลที่ปกติ (นอก 1.5*IQR คิดเป็น 99.3%) แสดงได้ดังรูปที่ 3.2 ทำให้เหลือช่วงของความดันไออยู่ที่ $6.925*10^{-7}$ ถึง 71.31 atm ซึ่งมีจำนวนข้อมูลทั้งสิ้น 1,787 ตัว โดยที่ช่วงของค่า A จะมีค่าอยู่ที่ 8.51 ถึง 86.16 ค่า B จะมีค่าอยู่ที่ 100 ถึง 122,723.96 และค่า C จะมีค่าอยู่ที่ -559.88 ถึง 1227.31 นอกจากนี้ยังทำการจัดการค่า T_{min} และ T_{max} ให้มีค่าอยู่ในช่วง 195 ถึง 493 K และ 299 ถึง 574 K ตามลำดับ สำหรับตารางข้อมูลสุดท้ายที่ได้ประกอบไปด้วย ชื่อสาร SMILES ค่า A B C T_{min} และ T_{max} ตามลำดับ

หลังจากที่ได้ตารางข้อมูลแล้วจะทำการใช้ Library ที่มีชื่อว่า RDKit เพื่อเก็บข้อมูลของโมเลกุลสารจากการป้อน SMILES เข้าไป หลังจากนั้นนำโมเลกุลสารที่ได้แปลงไปเป็น Count-Based Morgan Fingerprint ด้วยชุดคำสั่ง “rdMolDescriptors.GetHashedMorganFingerprint()” โดยจะสามารถกำหนดรัศมีและจำนวนบิตซึ่งด้วยการทดลองปรับเปลี่ยนด้วยการตั้งค่าต่อไปนี้

ตารางที่ 3.3 แสดงการตั้งค่าเพื่อหาลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการทำแบบจำลองทำนายความดันไอ

ตัวแปร	รายละเอียด
ตัวแปรควบคุม	ML Algorithm :XGBoost ที่มีการแบ่งข้อมูล : 80:20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแปร	รายละเอียด
ตัวแปรที่เปลี่ยนแปลง	MF_Radius =2-4 MF_Bits = 256-4,096

หลังจากที่ได้ Fingerprint ที่เหมาะสมมาแล้วซึ่งเป็นตารางข้อมูลที่มีคุณลักษณะพร้อมแล้วจะทำการแบ่งชุดข้อมูลเป็นข้อมูลที่ใช้ในการเรียนรู้:ข้อมูลที่ใช้ในการทดสอบ เป็นอัตราส่วนจาก 80:20 หรือ 90:10 ขึ้นกับประสิทธิภาพแบบจำลองทำนายสมบัติที่ได้ออกมา

3.3 การฝึกฝนโมเดล (Model Training)

Machine Learning Algorithm ที่ใช้ในการฝึกฝนโมเดลนั้น รวมไปถึงการใช้ K-fold ด้วยเพื่อทำให้การประเมินผลประสิทธิภาพของ Machine Learning Model ที่ได้ออกมานั้นดีขึ้น ต้องใช้เครื่องมือ (Library) ที่มีชื่อว่า sklearn (Scikit-learn) ซึ่งมีไว้ใช้สำหรับการทำ Machine Learning โดยใช้โปรแกรมภาษาไพทอน (Python)

3.3.1 การฝึกฝนโมเดลเพื่อสร้างแบบจำลองการทำนายจุดเดือด

ในการเพิ่มประสิทธิภาพของแบบจำลองทำนายสมบัตินี้จะฝึกฝนด้วยการใช้ชุดข้อมูลที่รวบรวมมาเพิ่มเติมทั้งหมด จำนวน 11,177 จุด มีขอบเขตเป็นสารอินทรีย์ที่มีอะตอมคาร์บอน ไฮโดรเจน ออกซิเจน และ ไนโตรเจนเป็นองค์ประกอบ วิธีการกำหนดคุณลักษณะสำหรับการฝึกฝนจะใช้ Count-based Morgan Fingerprint, หลังจากนั้นทำการแบ่งข้อมูลเป็นข้อมูลเรียนรู้:ข้อมูลทดสอบเป็น 90:10 โดยจะฝึกฝนด้วย Machine Learning Algorithm เป็น Ridge Regression, Random Forest, Extreme Gradient Boosting และ K-Nearest Neighbors

ในการรันแต่ละครั้งเพื่อให้ Machine Learning Algorithm ได้เรียนรู้นั้น จะต้องทำการเริ่มใช้ K-Fold ก่อนโดยงานวิจัยนี้ได้แบ่งออกเป็น 5 fold จากนั้นทำการให้เครื่องเรียนรู้ด้วย Algorithm ข้างต้น ทำให้ได้ Parameter ที่ดีที่สุดสำหรับชุดข้อมูลเรียนรู้ที่ได้ป้อนมา ทำให้ได้แบบจำลองที่เรียนผ่านการเรียนรู้มาแล้ว

3.3.2 การฝึกฝนโมเดลเพื่อสร้างแบบจำลองการทำนายความดันไอและทำนายค่า Antoine Coefficients

3.3.2.1 การทำนายค่าความดันไอ

ในการเพิ่มประสิทธิภาพของแบบจำลองทำนายสมบัตินี้จะฝึกฝนด้วยการใช้ชุดข้อมูลที่รวบรวมมาเพิ่มเติมทั้งหมด จำนวน 1,787 จุด มีขอบเขตเป็นสารอินทรีย์ในสถานะของเหลวที่มีอะตอมคาร์บอน ไฮโดรเจน ออกซิเจน และ ไนโตรเจนเป็นองค์ประกอบ วิธีการกำหนดคุณลักษณะสำหรับการฝึกฝนจะใช้ Count-Based Morgan Fingerprint, หลังจากนั้นทำการแบ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลเป็นข้อมูลเรียนรู้:ข้อมูลทดสอบเป็น 80:20 โดยจะฝึกฝนด้วย Machine Learning Algorithm เป็น Ridge Regression, Random Forest, Extreme Gradient Boosting, K-Nearest Neighbors

ในการรันแต่ละครั้งเพื่อให้ Machine Learning Algorithm ได้เรียนรู้นั้น จะต้องทำการเริ่มใช้ K-Fold ก่อนโดยงานวิจัยนี้ได้แบ่งออกเป็น 5 fold จากนั้นทำการให้เครื่องเรียนรู้ด้วย Algorithm ข้างต้น ทำให้ได้ parameter ที่ดีที่สุดสำหรับชุดข้อมูลเรียนรู้ที่ได้ป้อนมา ทำให้ได้แบบจำลองที่เรียนผ่านการเรียนรู้มาแล้ว

3.3.2.2 การทำนายค่า Antoine Coefficients

ในการเพิ่มประสิทธิภาพของแบบจำลองทำนายสมบัตินี้จะฝึกฝนด้วยการใช้ชุดข้อมูลที่รวบรวมมาเพิ่มเติมทั้งหมด จำนวน 1,787 จุด มีขอบเขตเป็นสารอินทรีย์ที่มีอะตอม คาร์บอน ไฮโดรเจน ออกซิเจน และ ไนโตรเจนเป็นองค์ประกอบ วิธีการกำหนดคุณลักษณะสำหรับการฝึกฝนจะใช้ Count-Based Morgan Fingerprint, หลังจากนั้นทำการแบ่งข้อมูลเป็นข้อมูลเรียนรู้: ข้อมูลทดสอบเป็น 80:20 โดยจะฝึกฝนด้วย Deep Learning

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

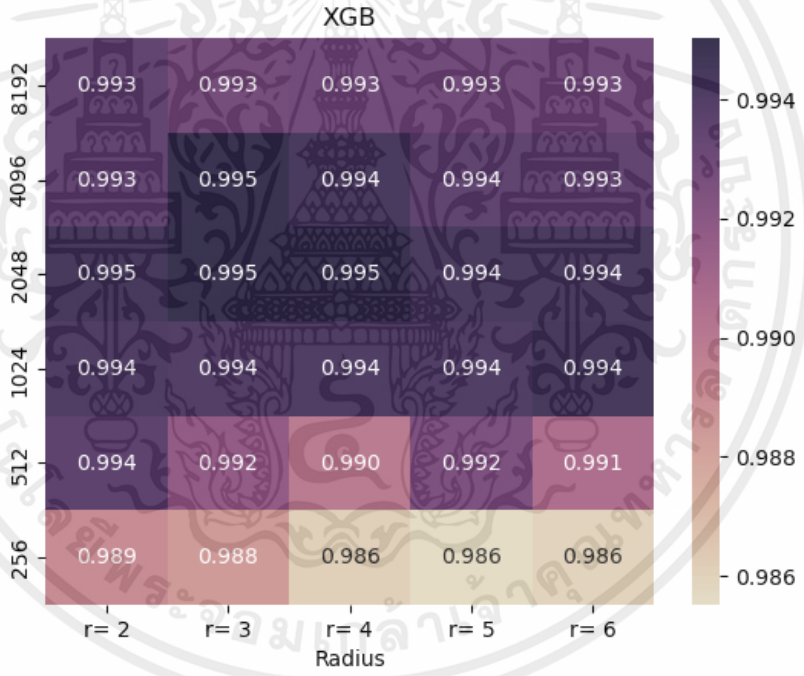
ผลการทดลองและวิเคราะห์ผลการทดลอง

ผลการดำเนินงานและวิเคราะห์ข้อมูลจากการพัฒนาแบบจำลองทำนายสมบัติของสารอินทรีย์ ได้ผลลัพธ์ดังต่อไปนี้

4.1 แบบจำลองทำนายจุดเดือด

4.1.1 การเปรียบเทียบผลการทำนายระหว่างแบบจำลองจากงานวิจัย Nattasinee และคณะ กับแบบจำลองทำนายสมบัติที่ประยุกต์ลายนิ้วมือโมเลกุลมาใช้

ผลการตรวจสอบลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการสร้างแบบจำลองจุดเดือด สามารถสังเกตได้จาก Heatmap ที่ทำการเปลี่ยนแปลงรัศมีและจำนวนบิต โดยยังคงใช้ Machine Algorithm ที่มีการตั้งค่าเหมือนเดิมโดยสามารถแสดงได้ดังภาพต่อไปนี้

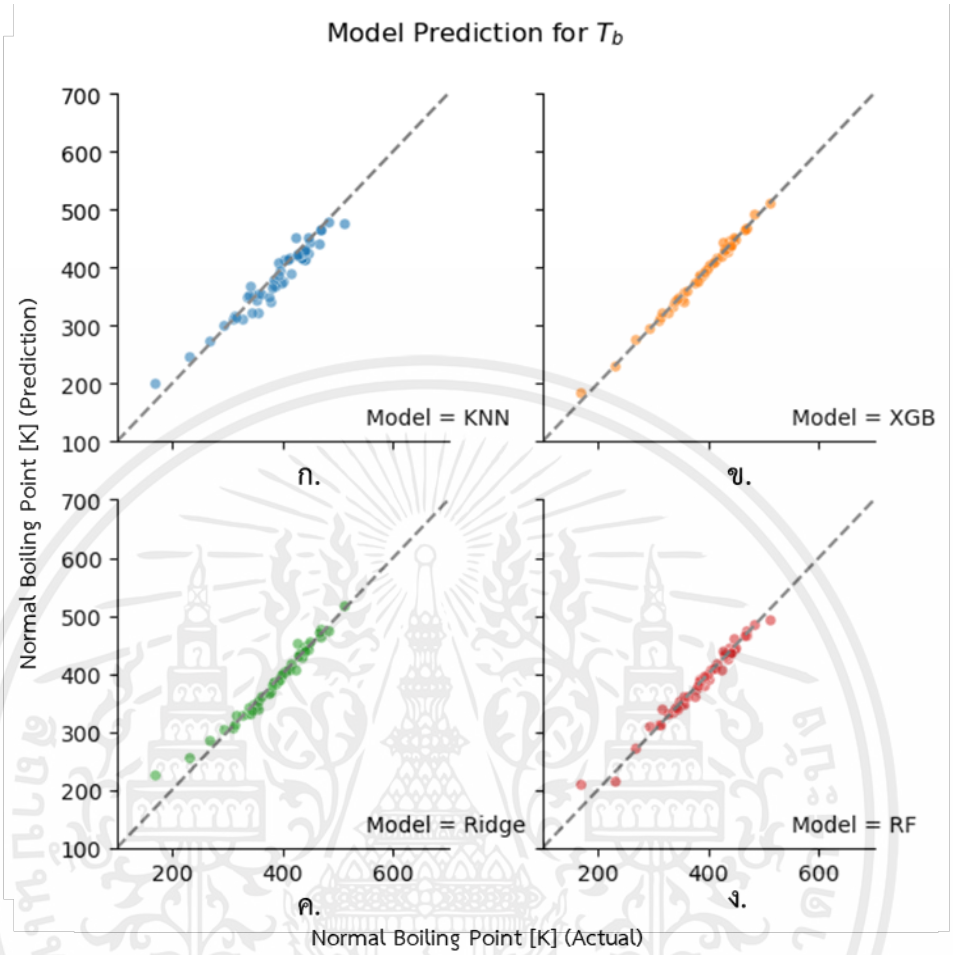


รูปที่ 4.1 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองทำนาย Normal Boiling Point

จากรูปที่ 4.1 สังเกตได้ว่าบริเวณที่สีอ่อนบ่งบอกถึงแบบจำลองที่มี R^2 ดีที่สุด แสดงว่าโมเดลดังกล่าวที่มีการตั้งค่า รัศมีและจำนวนบิต ในช่องดังกล่าวเป็นลายนิ้วมือโมเลกุลที่ดีที่สุดที่จะนำมาใช้ จากการคัดเลือกเบื้องต้นหาลายนิ้วมือโมเลกุลที่เหมาะสม จากรูปทำให้เลือกรัศมี 3 และจำนวนบิต 4,096 สำหรับการสร้างแบบจำลองทำนายจุดเดือด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากลายนิ้วมือโมเลกุลที่ได้มาพัฒนาแบบจำลองทำนายจุดเดือดด้วย Machine Learning Algorithm ต่างๆ ได้ผลลัพธ์และตารางดังต่อไปนี้



รูปที่ 4.2 กราฟเปรียบเทียบผลการทำนาย Normal Boiling Point ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน จาก KNN, XGB, Ridge และ RF Algorithm

ตารางที่ 4.1 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point ในขอบเขตอะตอมของคาร์บอนและไฮโดรเจน กับ Machine Learning Algorithmm ต่างๆ

Method	MAE	MAPE (%)	RMSE	R^2
Ridge	6.43	2.09	10.86	0.97
KNN	13.74	3.70	16.99	0.93
RF	5.91	1.83	9.227	0.98
XGB	3.20	0.84	4.700	0.99

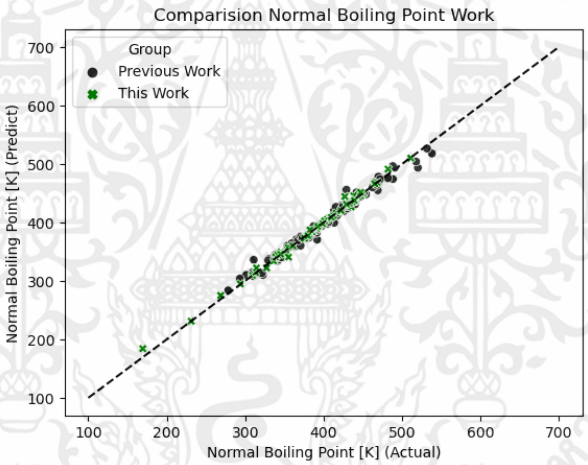
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับขอบเขต CH นั้น Algorithm ที่ดีที่สุดคือ XGB (รูปที่ 4.2ข) ที่มีการตั้งค่าดังตารางที่ 4.2 ซึ่งมีค่า Error จาก Test Set ผลดังต่อไปนี้ ค่า MAE = 3.20, ค่า MAPE = 0.84% และค่า R^2 = 0.99

ตารางที่ 4.2 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย Normal Boiling Point ที่ดีที่สุด

Algorithm	Hyperparameter	Training Control
XGB	Max Depth = 4 Learning Rate =0.2 N Estimators = 400	K-Fold=5

หลังจากนั้นนำผลลัพธ์ที่ได้จากงานในครั้งนี้เปรียบเทียบกับแบบจำลองทำนายของงานวิจัย Nattasinee และคณะ ซึ่งเป็นสารอินทรีย์ที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน กราฟและตารางแสดงประสิทธิภาพของแบบจำลองแต่ละงานวิจัยเป็นดังต่อไปนี้



รูปที่ 4.3 กราฟเปรียบเทียบผลการทำนายจากงานวิจัย Nattasinee และคณะ กับงานวิจัยนี้ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน

ตารางที่ 4.3 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Normal Boiling Point งานวิจัยของ Nattasinee และคณะ กับงานในครั้งนี้

งานวิจัย	MAE	MAPE (%)	RMSE	R^2
Nattasinee และคณะ	5.86	1.47	7.92	0.98
งานในครั้งนี้	3.20	0.84	4.70	0.99

จากกราฟรูปที่ 4.3 แสดงให้เห็นว่า แบบจำลองจากงานวิจัยนี้สามารถทำนายได้ใกล้เคียงกับเส้นสมบูรณ์แบบ (Perfect Line) ซึ่งบ่งบอกว่าค่าที่ทำนายเท่ากับค่าจริง การกระจายตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ของผลทำนายจากงานวิจัยนี้มีการกระจายตัวของผลทำนายน้อยกว่างานวิจัยของ Nattasinee และคณะ ซึ่งผลให้สามารถทำนายจุดเดือดได้แม่นยำดียิ่งขึ้นในขอบเขตของสารอินทรีย์ที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน ในขณะเดียวกันตารางแสดงผลการเปรียบเทียบก็สื่อสารออกมาในทิศทางเดียวกันว่า แบบจำลองจากงานวิจัยนี้ดีกว่างานวิจัยของ Nattasinee โดยเมื่อพิจารณาค่า MAE, MAPE, RMSE และ R^2 ของงานวิจัยนี้มีค่า MAE, MAPE, RMSE ที่ต่ำกว่าและมีค่า R^2 ที่สูงกว่า

จากตารางที่ 4.4 แสดงถึงปัญหาจากงานวิจัยของ Nattasinee และคณะ ซึ่งแสดงให้เห็นว่าโครงสร้างของสารเหล่านี้มีความคล้ายคลึงกันทำให้วิธีการของงานวิจัยนี้มีคุณลักษณะเหมือนกันส่งผลทำให้ทำนายจุดเดือดออกมาเท่ากัน จากตารางที่ 4.5 แสดงถึงการประยุกต์นำลายนิ้วมือโมเลกุลมาใช้ซึ่งทำให้สามารถแยกคุณลักษณะของโมเลกุลที่มีความคล้ายคลึงกันจากกันโดยทำให้สามารถสร้างคุณลักษณะที่แตกต่างกันได้ทำให้สามารถแยกทำนายของแต่ละโมเลกุลได้อย่างถูกต้อง

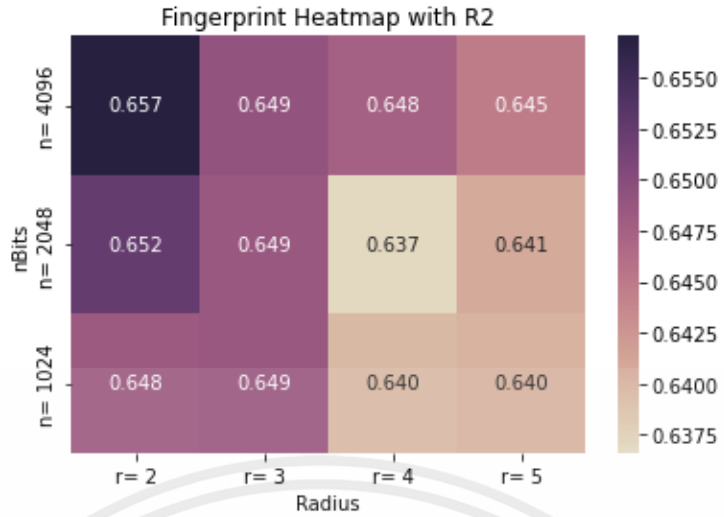
ตารางที่ 4.5 เปรียบเทียบการแก้ปัญหาคุณลักษณะขอบแบบจำลองทำนายจุดเดือดขอบเขตคาร์บอนและไฮโดรเจน

SMILES	Predict T_b, K Previous Work	Predict T_b, K This Work	Actual T_b , K
C1CCC=CCC1	375.85	387.34	388.15
CC1=CCCCC1	375.85	377.95	383.45
CC1CCC=CC1	375.85	370.31	375.85
CC1CCCC=C1	375.85	370.68	376.15

4.1.2 การประยุกต์แบบจำลองทำนายสมบัติที่ประยุกต์ลายนิ้วมือโมเลกุลมาใช้กับขอบเขตของสารอินทรีย์ที่ประกอบไปด้วยอะตอมของคาร์บอน ไฮโดรเจน ออกซิเจนและไนโตรเจน

ผลการตรวจสอบลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการสร้างแบบจำลองทำนาย Normal Boiling Point ที่มีขอบเขตอะตอมเพิ่มขึ้นเป็นสารอินทรีย์ที่ประกอบไปด้วยอะตอมของคาร์บอน ไฮโดรเจน ออกซิเจนและไนโตรเจน นั้นสามารถสังเกตได้จาก Heatmap ที่ทำการเปลี่ยนแปลงรัศมีและจำนวนบิต โดยยังคงใช้ Machine Algorithm ที่มีการตั้งค่าเหมือนเดิมโดยสามารถแสดงได้ดังภาพต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองทำนาย Normal Boiling Point ที่ขอบเขตสารเป็น CHON

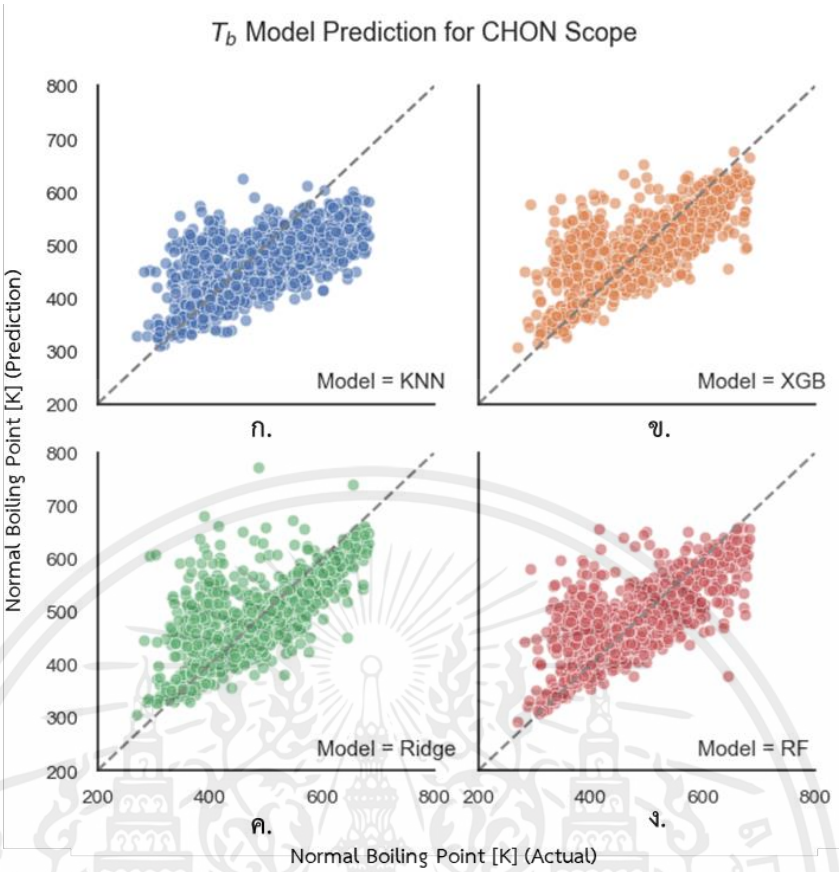
จากรูปที่ 4.4 บริเวณที่สีอ่อนบ่งบอกถึงแบบจำลองที่มี R^2 ดีที่สุด ทำให้เลือกรัศมี 2 และ จำนวนบิต 4,096 เป็นลายนิ้วมือโมเลกุลที่ดีที่จะนำมาใช้ในการพัฒนาแบบจำลองทำนายจุดเดือด (Normal Boiling Point) สำหรับขอบเขต CHON

จากลายนิ้วมือโมเลกุลที่ได้มานั้น นำมาใช้ในการหา Machine Learning Algorithm ที่ทำให้ได้ประสิทธิภาพการทำนายของแบบจำลองที่ได้ดีสุด ได้ผลลัพธ์เป็น กราฟและตารางที่แสดงประสิทธิภาพของแบบจำลองที่ขยายขอบเขตเพิ่มขึ้นดังต่อไปนี้

ตารางที่ 4.4 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนายจุดเดือดของขอบเขต CHON

Metrics / Algorithm	MAE		MAPE(%)		RMSE		R^2	
	Train	Test	Train	Test	Train	Test	Train	Test
Ridge	23.16	28.04	5.14	6.27	36.57	46.42	0.76	0.61
RF	16.15	26.96	3.56	5.94	26.47	44.49	0.87	0.64
XGB	20.42	26.72	4.51	5.94	30.38	43.40	0.83	0.66
KNN	0.31	39.27	0.07	8.16	3.02	55.06	0.99	0.45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 กราฟเปรียบเทียบผลการทำนาย Normal Boiling Point ในขอบเขตของสารที่ประกอบไปด้วยอะตอมของคาร์บอนและไฮโดรเจน จาก KNN, RF, Ridge และ XGB Algorithm

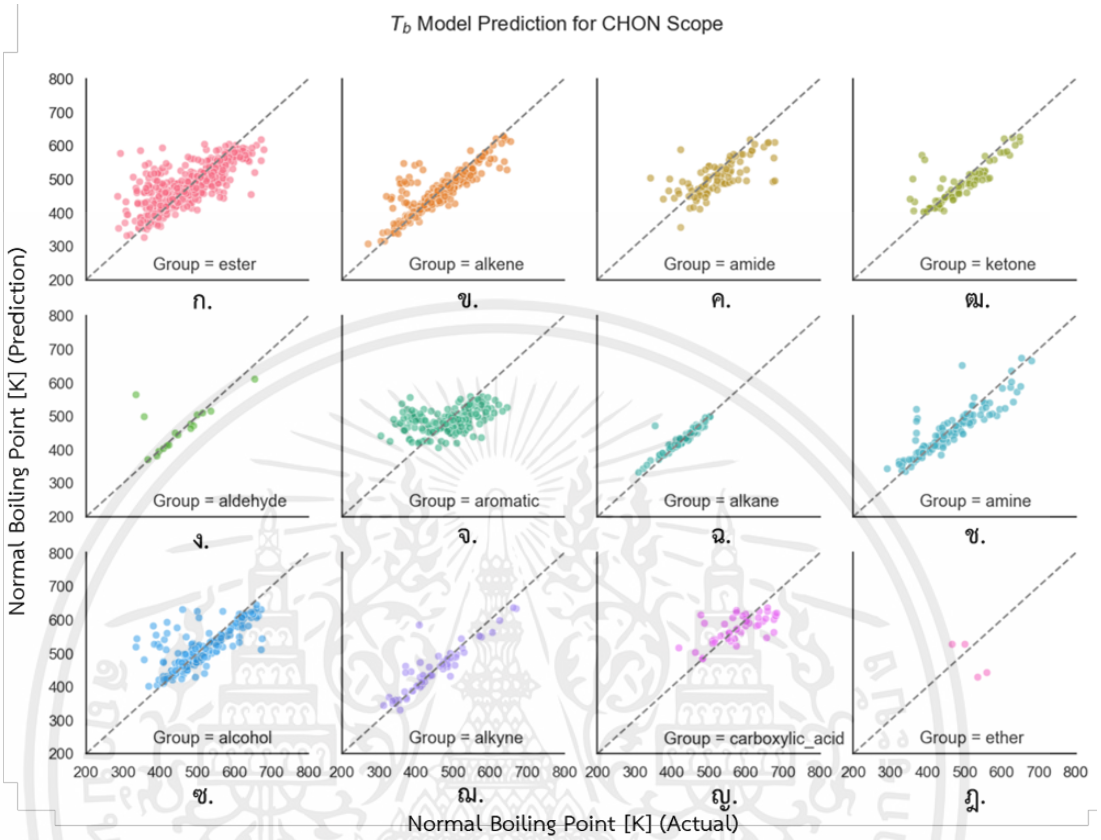
จากรูปที่ 4.5 กราฟแสดงให้เห็นว่า แบบจำลองทำนายจุดเดือดสำหรับขอบเขต CHON จากงานวิจัยนี้สามารถทำนายได้ มีความแปรปวนในการทำนายผลอยู่ปานกลาง ในขณะเดียวกัน ตารางที่ 4.5 แสดงผลการเปรียบเทียบยังคงแสดงให้เห็นว่า แบบจำลองทำนายจุดเดือดสำหรับขอบเขต CHON นั้น Algorithm ที่ดีที่สุดคือ XGB (รูปที่ 4.5ข) ที่มีการตั้งค่าดังตารางถัดไป โดย Error จาก Test Set มีผลดังต่อไปนี้ ค่า MAE MAPE RMSE และ R^2 มีค่าเท่ากับ 22.25 6.34% 45.43 และ 0.649 ตามลำดับ

ตารางที่ *4.5* แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนายจุดเดือดที่ดีที่สุด

Algorithm	Hyperparameter	Training Control
XGB	Max Depth = 7 Learning Rate = 0.05 N Estimators = 200	K-Fold=5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อวิเคราะห์ผลการทำนายจุดเดือดของแบบจำลองสำหรับขอบเขต CHON เพื่อดูประสิทธิภาพของแบบจำลองในการทำนายแยกตามหมู่ฟังก์ชั่นของสารอินทรีย์ของชุดข้อมูลทดสอบ ซึ่งแสดงจำนวน, MAPE, RMSE, R^2 ของสารแต่ละหมู่ฟังก์ชั่น สามารถแสดงผลได้ดังรูปต่อไปนี้



รูปที่ 4.6 กราฟเปรียบเทียบผลการทำนายจุดเดือดจากแบบจำลองด้วยขอบเขต CHON ที่ได้แยกตามหมู่ฟังก์ชั่นของ Test Set

จากรูปที่ 4.6 กราฟแสดงการเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองจุดเดือดจาก XGB Algorithm สามารถังเกตได้ว่า สำหรับเกณฑ์ที่ผู้วิจัยคาดการณ์ว่าแบบจำลองนี้สามารถทำนายได้ดี สามารถดูได้จาก MAPE ≤ 10% และ R^2 ≥ 0.5 [31,32] ทำให้กล่าวได้ว่าจากแบบจำลองทำนายจุดเดือดนี้สามารถทำนายสารประเภทเหล่านี้ได้ดีซึ่งได้แก่ Ester, Alkene, Ketone, Alkane, Amine, Alcohol และ Alkyne ในขณะที่หมู่ฟังก์ชั่นเหล่านี้แบบจำลองนี้จะทำนายได้ไม่ดีซึ่งได้แก่ Amide, Aldehyde, Aromatic, Carboxylic Acid และ Ether ซึ่งคาดว่าเกิดจากการที่มีข้อมูลของแต่ละหมู่ฟังก์ชั่นน้อยและไม่หลากหลายมากเพียงพอ ซึ่งอาจจะต้องมีหาข้อมูลและทำการปรับปรุงรูปแบบในการเรียนรู้เพื่อพัฒนาโมเดลให้ดีขึ้นในภายภาคหน้า

ตารางที่ 4.6 ตารางระหว่างหมู่ฟังก์ชันและค่าความคลาดเคลื่อน

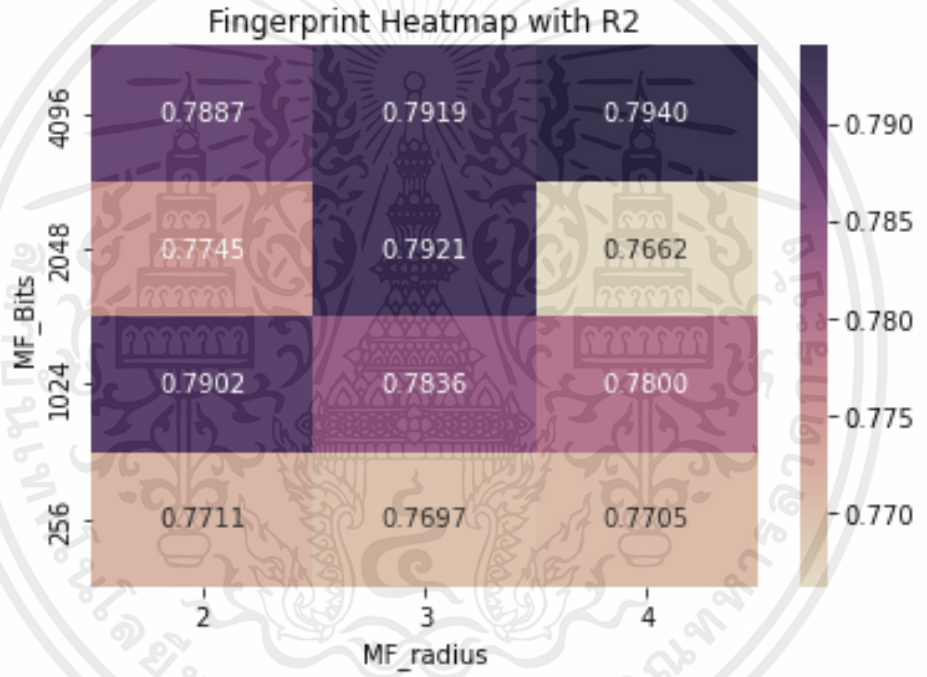
Group	MAPE (%)	R^2	Group	MAPE (%)	R^2
Ester	6.49	0.61	Alkane	1.53	0.884

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Group	MAPE (%)	R^2	Group	MAPE (%)	R^2
Alkene	4.19	0.78	Amine	5.23	0.80
Amide	7.79	0.42	Alcohol	6.97	0.53
Ketone	5.71	0.67	Alkyne	4.95	0.81
Aldehyde	8.27	0.20	Carboxylic Acid	6.42	0.42
Aromatic	9.78	0.18	Ether	14.71	-4.75

4.2 แบบจำลองทำนายความดันไอ

ผลการตรวจสอบลายนิ้วมือโมเลกุลที่เหมาะสมสำหรับการสร้างแบบจำลองทำนาย Vapor Pressure นั้นสามารถสังเกตได้จาก Heatmap ที่ทำการเปลี่ยนแปลงรัศมีและจำนวนบิต โดยยังคงใช้ Machine Algorithm ที่มีการตั้งค่าเหมือนเดิมโดยสามารถแสดงได้ดังภาพต่อไปนี้



รูปที่ 4.7 Heatmap เปรียบเทียบการเปลี่ยนแปลงของรัศมี และ บิต Count-Based Morgan Fingerprint สำหรับแบบจำลองการทำนายความดันไอ

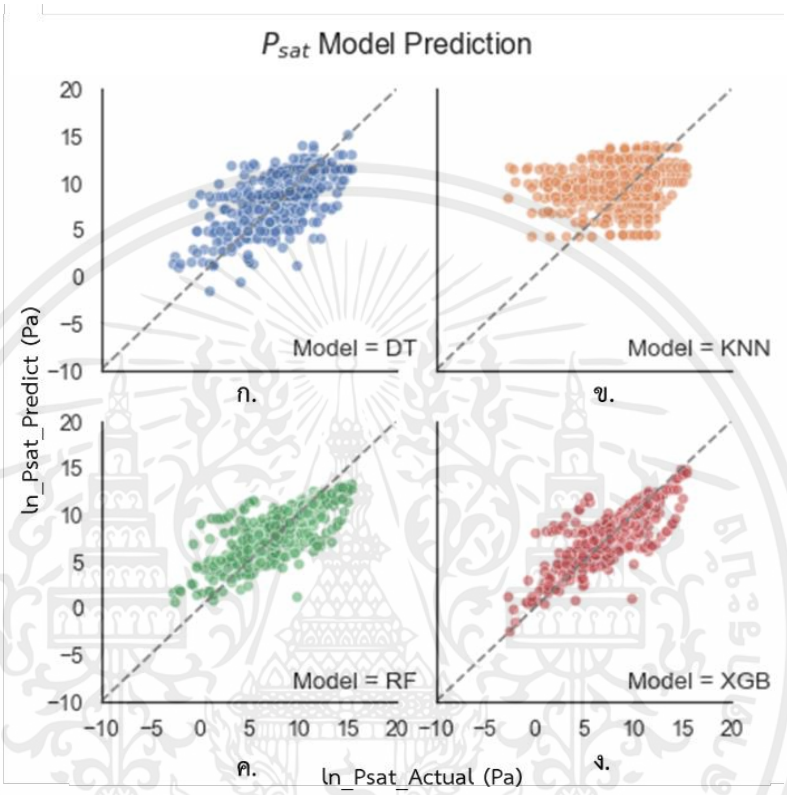
จากรูปที่ 4.7 โดยที่บริเวณที่สีอ่อนบ่งบอกถึงแบบจำลองที่มี R^2 ดีที่สุด ทำให้เลือกรัศมี 3 และ จำนวนบิต 2,048 เป็นลายนิ้วมือโมเลกุลที่ดีที่จะนำมาใช้ในการพัฒนาแบบจำลองทำนายความดันไอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 วิเคราะห์การสร้างแบบจำลองทำนาย ln(P^{sat})

4.2.1.1 วิเคราะห์ผลการทำนายความดันไอที่ได้จากแบบจำลอง

จากลายนิ้วมือโมเลกุลที่ได้มานั้น นำมาใช้ในการหา Machine Learning Algorithm ที่ทำให้ได้ประสิทธิภาพการทำนายของแบบจำลองที่ได้ดีสุด ได้ผลลัพธ์เป็น กราฟและตารางที่แสดงประสิทธิภาพของแบบจำลองที่ขยายขอบเขตเพิ่มขึ้นดังต่อไปนี้



รูปที่ 4.8 กราฟเปรียบเทียบผลการทำนายความดันไอ
จาก ก) DT ข. KNN ค) RF และ ง) XGB Algorithm

ตารางที่ 4.7 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Vapor Pressure

Method	MAE	MAPE (%)	RMSE	R^2
DT	0.80	15.00	1.38	0.68
KNN	1.76	22.30	2.42	-1.24
RF	0.63	15.90	1.10	0.73
XGB	0.59	13.60	1.06	0.80

จากรูปที่ 4.8 กราฟแสดงให้เห็นว่า แบบจำลองทำนายความดันไอจากงานวิจัยนี้สามารถทำนายได้ มีความแปรปวนในการทำนายผลอยู่ปานกลาง ในขณะเดียวกันตารางแสดงผลการเปรียบเทียบยังคงแสดงให้เห็นว่า แบบจำลองทำนายความดันไอนั้น Algorithm ที่ดีที่สุด

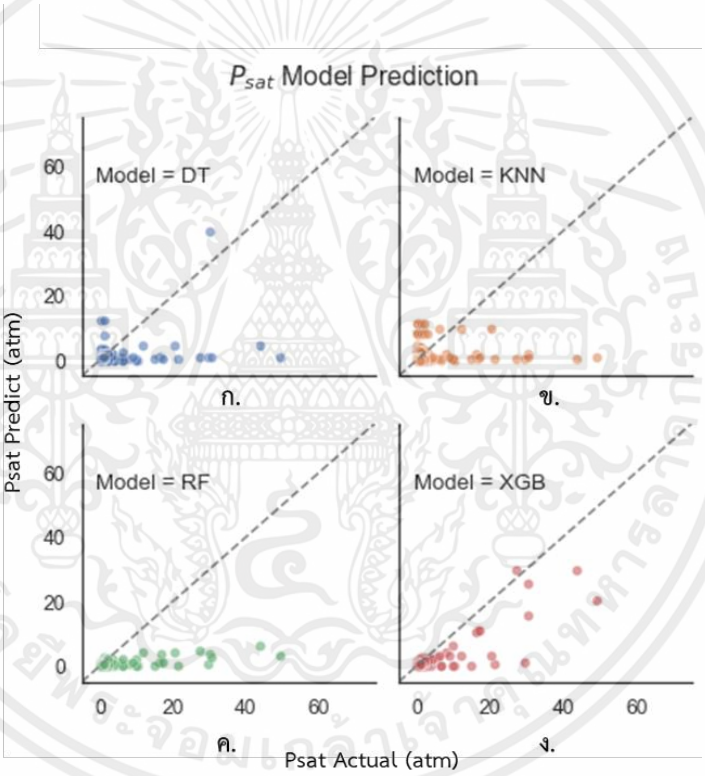
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คือ XGB (รูปที่ 4.8ข) ที่มีการตั้งค่าดังตารางถัดไป โดย Error จาก Test Set มีผลดังต่อไปนี้ ค่า MAE MAPE RMSE และ R^2 มีค่าเท่ากับ 0.590 13.6% 1.063 และ 0.798 ตามลำดับ

ตารางที่ 4.8 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย ln(P^{sat}) ที่ดีที่สุด

Algorithm	Hyperparameter	Training Control
XGB	Max Depth = 5 Learning Rate = None N Estimators = 400	K-Fold=5

ในขณะเดียวกัน สามารถเปลี่ยนหน่วยผลลัพธ์การทำนายให้อยู่ในรูปของความดันไอในหน่วย atm แสดงได้ดังกราฟและตารางดังต่อไปนี้



รูปที่ 4.9 กราฟเปรียบเทียบผลการทำนายความดันไอในหน่วย Pa จาก ก) DT ข) KNN ค) RF และ ง) XGB Algorithm

ตารางที่ 4.9 เปรียบเทียบประสิทธิภาพของแบบจำลองทำนาย Vapor Pressure ในหน่วย Pa

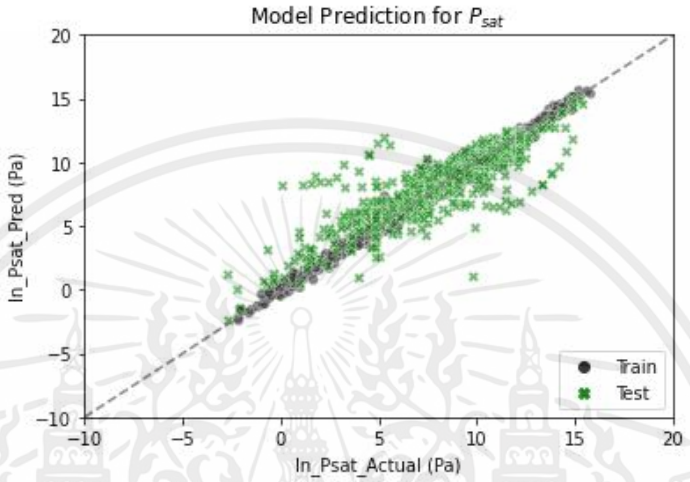
Method	MAE	MAPE (%)	RMSE	R^2
DT	0.48	725.90	2.79	-6.55
KNN	0.74	4887.00	3.29	-14.23
RF	0.44	161.40	2.88	-22.37

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Method	MAE	MAPE (%)	RMSE	R^2
XGB	0.38	373.60	2.40	-5.12

4.2.1.2 วิเคราะห์ผลการทำนายความดันไอแยกตามหมู่ฟังก์ชั่นของสาร

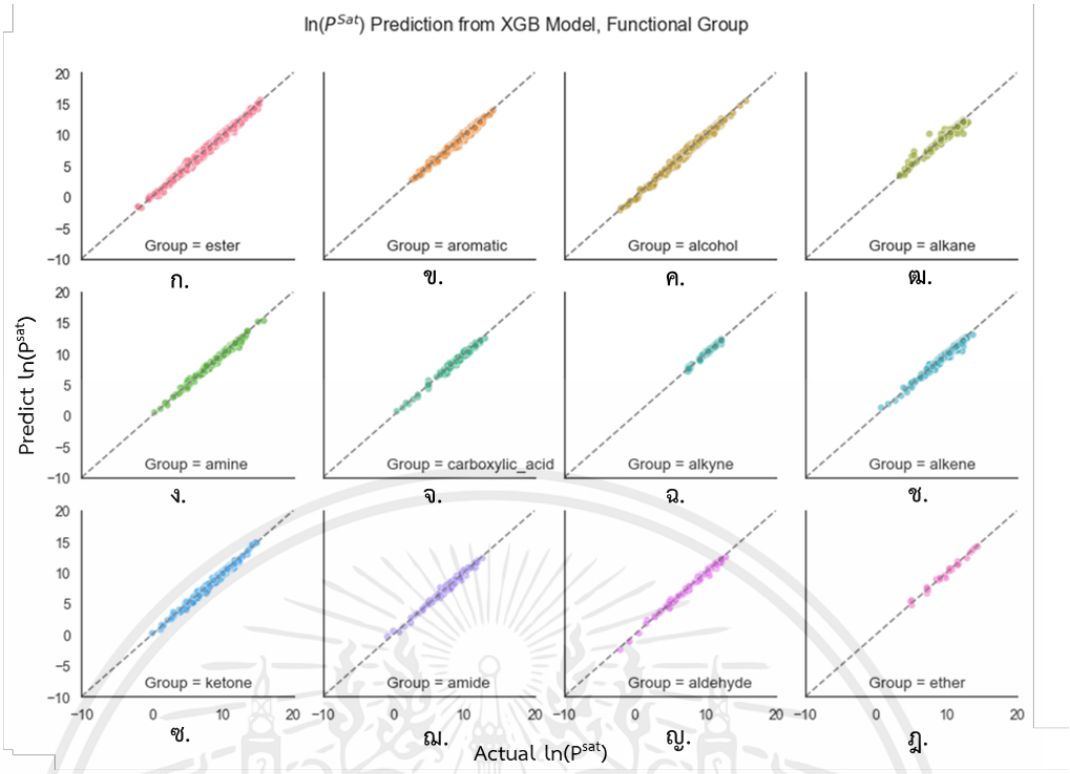
จากผลการทำนายความดันไอแบบข้างต้นจะได้ว่า XGB เป็น Algorithm ที่ดีที่สุดที่มีผลการทำนายของ Train และ Test Set ดังกราฟต่อไปนี้



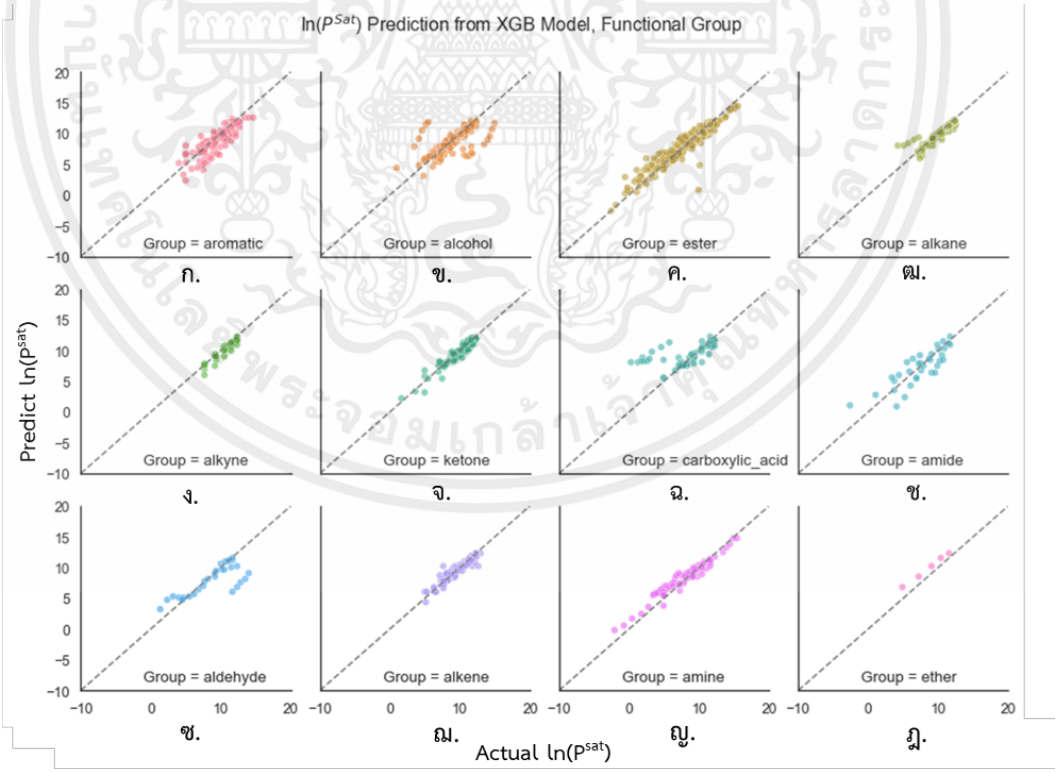
รูปที่ 4.10 กราฟเปรียบเทียบผลการทำนายความดันไอจาก XGB Algorithm จาก Train (ซ้าย) และ Test (ขวา) Set

หลังจากนั้นทำการแบ่งผลการทำนายทั้ง Train และ Test ออกเป็นตามหมู่ฟังก์ชั่นของโมเลกุลสารได้กราฟและประสิทธิภาพของแบบจำลองตามหมู่ฟังก์ชั่นสารได้รูปดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.11 กราฟเปรียบเทียบผลการทำนายความดันไอจากแบบจำลองที่ได้แยกตามหมู่ฟังก์ชั่น สำหรับ Training Set



รูปที่ 4.12 กราฟเปรียบเทียบผลการทำนายความดันไอจากแบบจำลองที่ได้แยกตามหมู่ฟังก์ชั่น สำหรับ Test Set

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

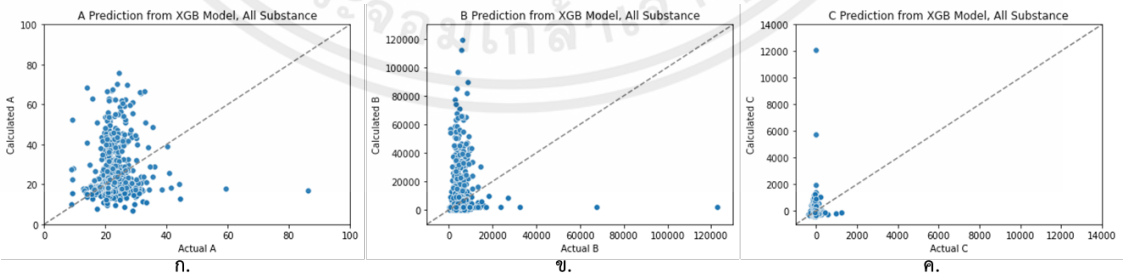
จากรูปที่ 4.11 และ รูปที่ 4.12 กราฟแสดงการเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองความดันไอจาก XGB Algorithm สามารถถังเกตได้ว่า สำหรับเกณฑ์ที่ผู้วิจัยคาดการณ์ว่าแบบจำลองนี้สามารถทำนายได้ดี สามารถดูได้จาก MAPE ≤ 10% และ $R^2 \geq 0.5$ [31,32] ทำให้กล่าวได้ว่าจาก แบบจำลองทำนายความดันไอนี้สามารถทำนายสารประเภทเหล่านี้ได้ดี ซึ่งได้แก่ Alcohol, Alkane, Alkyne, Ketone และ Alkene ในขณะที่หมู่ฟังก์ชั่นเหล่านี้ทำนายได้ไม่ดีซึ่งได้แก่ Aromatic, Ester, Carboxylic Acid, Amide, Aldehyde, Amine และ Ether เมื่อเปรียบเทียบกราฟที่ได้จากรูปที่ 4.12 กับ รูปที่ 4.11 จึงสามารถสรุปได้ว่าแบบจำลองนี้ยังไม่สามารถทำนายความดันไอของหมู่ Aromatic, Ester, Carboxylic Acid, Amide, Aldehyde, Amine และ Ether ยังไม่ดีมากพอ ซึ่งอาจจะต้องมีการปรับปรุงรูปแบบในการเรียนรู้เพื่อพัฒนาโมเดลให้ดีขึ้นในภายภาคหน้า

ตารางที่ 4.10 ตารางระหว่างหมู่ฟังก์ชันและค่าความคลาดเคลื่อน

Group	MAPE (%)	R^2	Group	MAPE (%)	R^2
Ester	18.43	0.90	Alkane	3.88	0.88
Alkene	3.24	0.93	Amine	21.40	0.90
Amide	30.08	0.71	Alcohol	8.87	0.61
Ketone	5.98	0.91	Alkyne	3.61	0.88
Aldehyde	23.32	0.55	Carboxylic Acid	186.93	0.19
Aromatic	10.30	0.68	Ether	20.08	0.60

4.2.1.3 การคำนวณค่า Antoine Coefficients จากความดันไอที่ได้จากแบบจำลอง

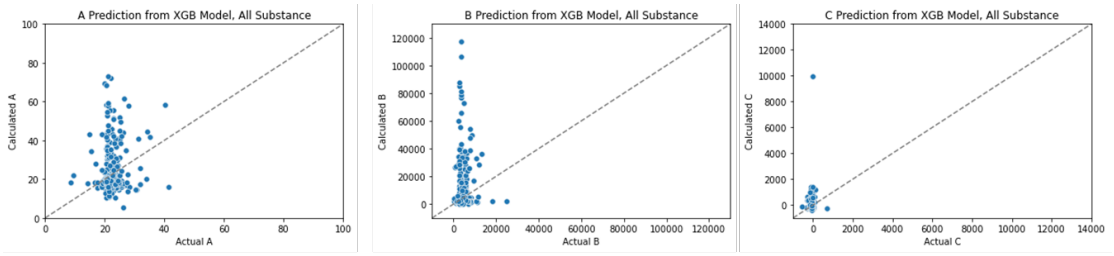
เมื่อได้แบบจำลองที่สามารถทำนายความดันไอได้นั้น เพื่อความสะดวกในการใช้งาน จึงได้มีการคำนวณเพื่อหา Antoine Coefficients ซึ่งได้มากจากการทำ Curve Fitting เพื่อให้ได้ค่า A B และ C ออกมา ซึ่งสามารถนำผลลัพธ์ความดันไอที่ทำนายได้จาก Train Set และ Test set ที่ได้จากแบบจำลองคำนวณได้ผลลัพธ์ดังต่อไปนี้



ก. ข. ค.

รูปที่ 4.13 กราฟเปรียบเทียบผลการทำนาย Antione Coefficient จาก XGB Algorithm จาก Train Set

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



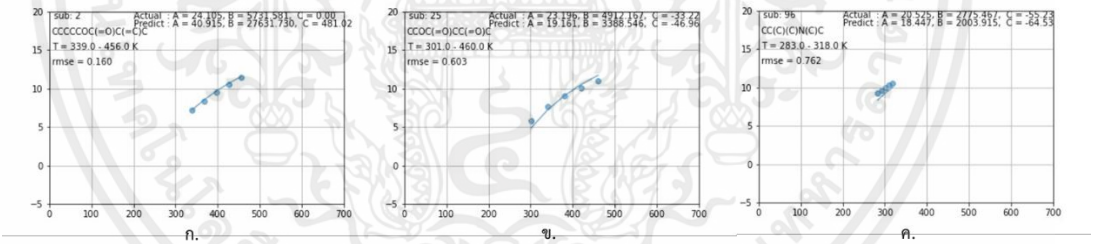
ก. ข. ค.

รูปที่ 4.14 กราฟเปรียบเทียบผลการทำนาย Antione Coefficient จาก XGB Algorithm จาก Test Set

จากรูปที่ 4.13 และ รูปที่ 4.14 สามารถสังเกตได้ว่า แม้ความดันไอที่ทำนายได้จะแม่นยำในส่วนของ Train Set (เช่น รูปที่ 4.13ก) แต่ Antione Coefficient ที่คำนวณได้นั้นมีการกระจายตัวออกไปจากค่าที่ได้จากฐานข้อมูล เช่นเดียวกันกับในส่วนของ Test Set (เช่น รูปที่ 4.14ก) ทำให้สามารถกล่าวได้ว่า Antione Coefficients ที่สามารถคำนวณได้จากความดันไอที่ทำนายจากแบบจำลองทำนายสมบัติความดันไอนี้นั้น เป็นชุดตัวเลข Antione Coefficients ใหม่

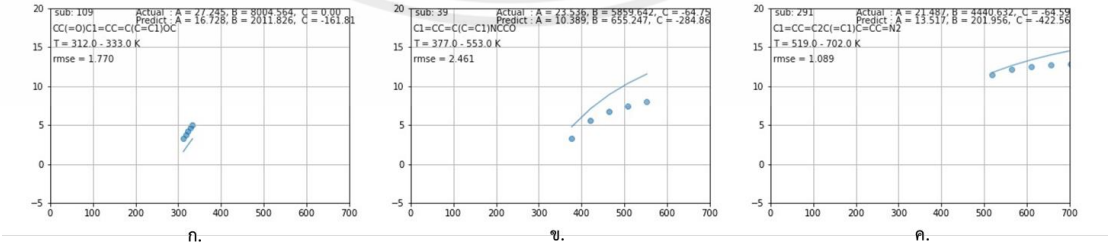
4.2.1.4 การตรวจสอบความดันไอที่ได้จาก Antoine Coefficients ชุดใหม่

เมื่อทำการคำนวณความดันไอจาก Antione Coefficients ชุดใหม่เปรียบเทียบกับจากฐานข้อมูล ตัวอย่างของกราฟแสดงความสัมพันธ์ระหว่างอุณหภูมิกับความดันไอจากการคำนวณผ่านสัมประสิทธิ์คนละชุดสามารถแสดงตัวอย่างสำหรับ Test Set ได้ผลดังต่อไปนี้



ก. ข. ค.

รูปที่ 4.15 กราฟความสัมพันธ์ระหว่างอุณหภูมิกับความดันไอ จาก Antione Coefficient ทั้งจากฐานข้อมูลและค่าที่ได้จากการทำนายความดันไอของโมเลกุลสารของตัวอย่างผลลัพธ์ที่ดี



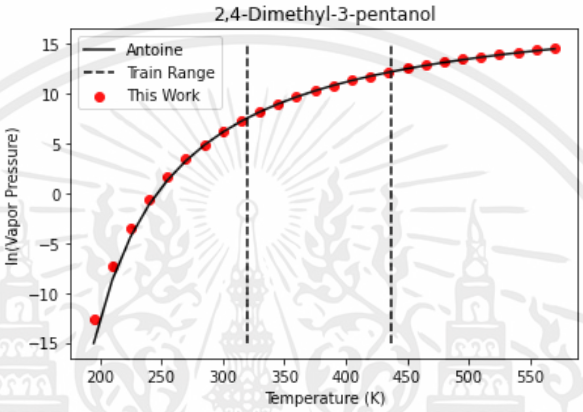
ก. ข. ค.

รูปที่ 4.16 กราฟความสัมพันธ์ระหว่างอุณหภูมิกับความดันไอ จาก Antione Coefficient ทั้งจากฐานข้อมูลและค่าที่ได้จากการทำนายความดันไอของโมเลกุลสารของตัวอย่างผลลัพธ์ที่ไม่ดี

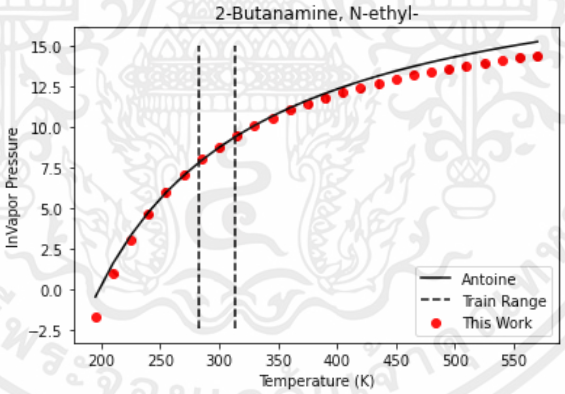
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.15 แสดงให้เห็นว่าค่า Antoine Coefficients ที่ได้มาเป็นค่าใหม่เนื่องจากเมื่อนำค่าที่ได้และจากแหล่งอ้างอิงมาคำนวณด้วยสมการ Antoine แล้วมีค่า $\ln(P^{sat})$ ใกล้เคียงกันแต่ว่าค่า Antoine Coefficients ที่ได้มาใหม่นั้นก็มีค่าที่ไม่ดีหลังจากคำนวณค่าจาก $\ln(P^{sat})$ ตามรูปที่ 4.16 เนื่องจากค่าที่ได้จากแบบจำลองนั้นทำนายไม่ดี จึงทำให้ได้ค่าที่ไม่สามารถนำมาใช้ได้

เมื่อลองนำอุณหภูมิที่อยู่นอกระยะมาคำนวณกับ Antoine Coefficient ที่ได้จากการคำนวณมาจากแบบจำลองและแหล่งอ้างอิงจะพบว่า ยิ่งค่าระยะห่างของอุณหภูมิมีค่ามากจะสามารถทำนายออกได้ตรงเทียบกับแหล่งอ้างอิงกว่าที่ค่าระยะห่างของอุณหภูมิน้อย เมื่อนำอุณหภูมิที่ต่ำที่สุดและสูงที่สุดมาจากแหล่งอ้างอิงของสารทั้งหมด แสดงได้ดังรูปที่ 4.17 และรูปที่ **4.18**



รูปที่ 4.17 สารที่มีระยะห่างอุณหภูมิเยอะและการคำนวณนอกอุณหภูมิ



รูปที่ 4.18 สารที่มีระยะห่างอุณหภูมิน้อยและการคำนวณนอกอุณหภูมิ

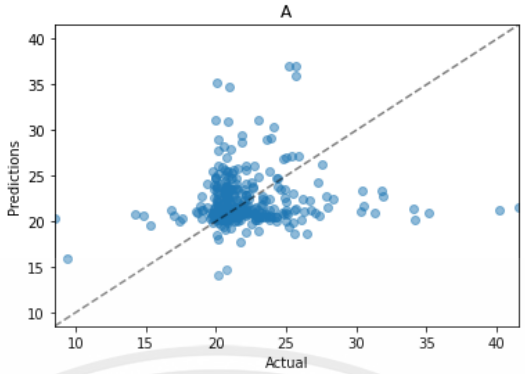
ซึ่งในงานวิจัยไม่ได้มีการยืนยันว่าเมื่อลองนำค่า Antoine Coefficients ใหม่มาคำนวณนอกระยะที่ได้จากแหล่งอ้างอิงว่าถูกต้อง ซึ่งสารที่จะนำมาค่ามาใช้จะสามารถใช้ได้แค่สารที่อยู่ในสภาวะของเหลวเท่านั้น

4.2.2 วิเคราะห์การสร้างแบบจำลองทำนาย Antoine Coefficients

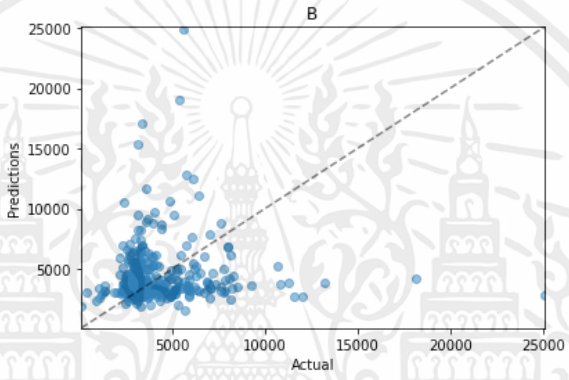
จากลายนิ้วมือโมเลกุลที่ได้มานั้น เมื่อนำไปเทรนด้วย Deep Learning แล้วนำแบบจำลองที่ได้มาใช้ในการทำนายค่า Antoine Coefficients หลังจากนั้นได้ทำการนำค่าที่ได้จากการทำนายจะนำค่าอุณหภูมิที่มี (T_{min} และ T_{max}) มาคำนวณหาค่า $\ln(P^{sat})$ หาว่าเมื่อนำค่า Antoine

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

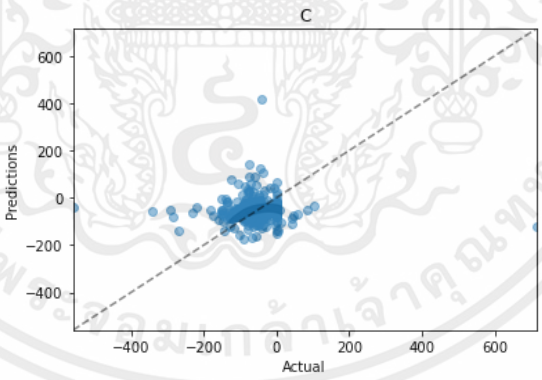
Coefficients กับค่าของจริงที่ได้นั้นมีค่าต่างกันเท่าไร โดยสามารถแสดงกราฟระหว่างค่าจริงกับค่าการทำนาย สามารถแสดงได้ดังรูปด้านล่าง



รูปที่ 4.19 ค่า A ที่ได้จากแบบจำลอง



รูปที่ 4.20 ค่า B ที่ได้จากแบบจำลอง



รูปที่ 4.21 ค่า C ที่ได้จากแบบจำลอง

ตารางที่ 4.11 เปรียบเทียบค่าความคลาดเคลื่อนของค่าที่ได้จากแบบจำลองทำนาย Antione eff.

Value	MAE	MAPE (%)	RMSE	R^2
A	2.63	11.81	3.98	-0.63
B	1939.86	53.61	3186.83	-1.00
C	51.32	2.37×10^{18}	85.94	-0.53

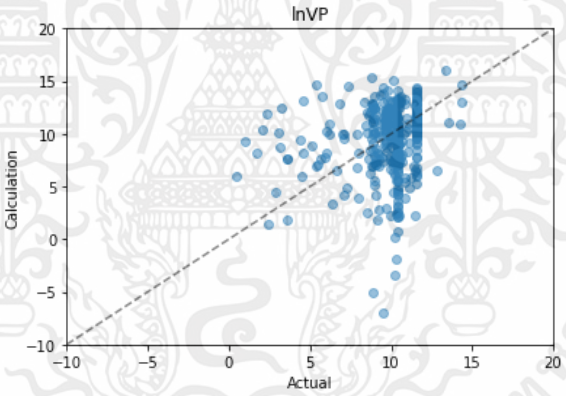
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.12 แสดงให้เห็นว่า แบบจำลองการทำนายค่า Antoine Coefficients มาแล้วนำไปคำนวณค่า ln(P^{sat}) นั้นมีค่า Error จาก Test Set ของค่า A B และ C ดังต่อไปนี้ MAE มีค่าเท่ากับ 2.63, 1939.86 และ 51.32 ค่า MAPE เท่ากับ 11.81%, 53.61% และ 2.37×10^{18}% ค่า RMSE เท่ากับ 3.98, 3186.83, และ 85.94 และค่า R^2 เท่ากับ -0.63, -1.00 และ -0.53 ตามลำดับ โดยมีการตั้งค่าของ Deep Learning ตามตารางที่ 4.12

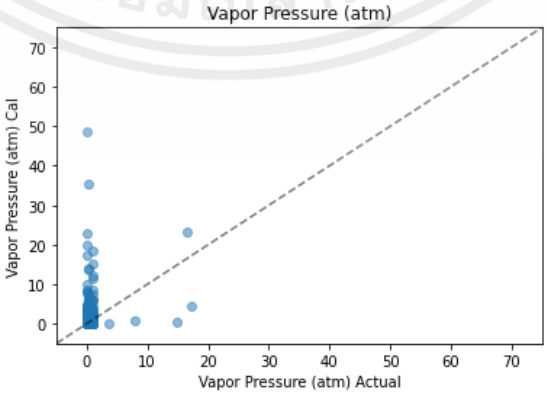
ตารางที่ 4.12 แสดงการตั้งค่าสำหรับการสร้างแบบจำลองทำนาย Antoine Coefficients ที่ดีที่สุด

Algorithm	Hyperparameter
Deep Learning	Hidden Layer = 2 Hidden Node = 1000 Learning Rate = 0.0001 Dropout = 0.2

ในขณะเดียวกันเมื่อนำค่า Antoine Coefficients มาคำนวณและเปลี่ยนหน่วยผลลัพธ์ให้อยู่ในรูปของความดันไอในหน่วย atm แสดงได้ดังกราฟและตารางดังต่อไปนี้



รูปที่ 4.22 กราฟเปรียบเทียบการคำนวณค่า ln(P^{sat}) จากคำนวณด้วยค่าที่ได้จากการทำนายของแบบจำลอง



รูปที่ 4.23 กราฟเปรียบเทียบผลการทำนายความดันไอในหน่วย Pa จากคำนวณด้วยค่าที่ได้จากการทำนายของแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลลัพธ์ที่จากโมเดลการทำนายค่า Antoine Coefficients เมื่อนำค่าที่ได้มาใช้ต่อในการคำนวณหาค่า ln(P^{sat}) จะพบว่าเราไม่สามารถควรนำค่าที่ได้จากโมเดลมาใช้ในการคำนวณ เนื่องจากว่าเมื่อดูค่าที่ได้มาแล้วมีค่าที่ MAE RMSE %MAPE สูงจึงไม่สามารถนำมาใช้ได้ การจะนำค่าที่ได้มาใช้ในการสร้างแบบจำลองควรจะเป็นค่าที่ได้จากการทดลองจริง ไม่ควรใช้ค่าที่นำผ่านการคำนวณมาใช้ในการสร้างแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

จากงานวิจัยนี้มุ่งเน้นในการนำ Molecular Fingerprint จากการแปลงข้อมูลโมเลกุลสารด้วย SMILES และพัฒนาแบบจำลองด้วย Machine Learning เพื่อทำนายสมบัติของสาร โดยในงานวิจัยนี้ได้พัฒนาแบบจำลองแบบจำลองทำนายจุดเดือด (Normal Boiling Point) ซึ่งเป็นสมบัติที่ไม่ขึ้นกับอุณหภูมิของสาร และ ทำนาย ความดันไอ (Vapor Pressure) ซึ่งเป็นสมบัติที่ขึ้นกับอุณหภูมิของสารนั้น โดยกำหนดขอบเขตเป็นสารอินทรีย์ที่มีอะตอมคาร์บอนตั้งแต่ 1-12 อะตอมและประกอบไปด้วยอะตอมคาร์บอน ไฮโดรเจน ออกซิเจนและไนโตรเจน

งานวิจัยนี้เริ่มต้นด้วยการเก็บรวบรวมข้อมูล ชื่อสาร และ สมบัติสารที่ต้องการทำนาย หลังจากนั้นทำการคัดแยกข้อมูลที่ไม่เหมาะสม สุดท้ายข้อมูลสารที่พร้อสำหรับงาน Normal Boiling Point และ Vapor Pressure มีทั้งหมด 11,177 และ 1,787 ตัว ตามลำดับ หลังจากนั้นทำการแปลงโครงสร้างสารให้อยู่ในรูป SMILES แล้วแปลงเป็น Morgan Fingerprint ที่เหมาะสม ขั้นตอนถัดไปทำการแบ่งข้อมูลไว้สำหรับชุดฝึก (Train Set) และชุดทดสอบ (Test Set) เป็น 80:20 หรือ 90:10

จากนั้นนำข้อมูลชุดฝึก (Train Set) ของข้อมูลจุดเดือดมาใช้ในการเทรนโมเดลเพื่อสร้างแบบจำลองทำนายจุดเดือด โดยมี Machine Learning Algorithm ที่หลากหลาย โดยได้ผลที่ดีที่สุดเป็น สำหรับงาน จุดเดือดสำหรับขอบเขต CH และ CHON คือ XGB สำหรับขอบเขต CH มีค่า MAE, %MAPE, R^2 เป็น 2.859, 0.834 และ 0.994 ตามลำดับ สำหรับขอบเขต CHON มีค่า MAE, %MAPE, R^2 เป็น 26.46, 5.832 และ 0.710 ตามลำดับ

หลังจากนั้นนำข้อมูลชุดฝึกของ (Train Set) ของขอมูลความดันไอมาทำการเทรนโมเดลเพื่อสร้างแบบจำลองทำนายความดันไอ โดยมี Machine Learning Algorithm ที่หลากหลาย โดยสามารถแบ่งการวิจัยออกเป็น การทำนายความดันไอโดยตรงกับการทำนาย Antoine Coefficients สำหรับการทำนายความดันไอในรูปของ $\ln(P^{sat})$ ได้ Machine Learning Algorithm ที่ดีที่สุดเป็น XGB โดยค่า MAE, %MAPE และ R^2 เป็น 0.590, 13.6 และ 0.798 ตามลำดับ สำหรับการทำนาย Antoine Coefficients ได้ Machine Learning Algorithm ที่ดีที่สุดเป็น Deep Learning โดยค่า MAE, %MAPE, R^2 ของ A B และC ดังต่อไปนี้ MAE มีค่าเท่ากับ 2.63, 1939.86 และ 51.32 ค่า %MAPE เท่ากับ 11.81, 53.61 และ 2.37×10^{18} และมีค่า R^2 เท่ากับ -0.63, -1.00 และ -0.53 ตามลำดับ

สำหรับการทำนายความดันไอ เมื่อวิเคราะห์ $\ln(P^{sat})$ พบว่าหมู่ฟังก์ชั่นของสารที่แบบจำลองทำนายความดันไอได้ดีนั้น ได้แก่หมู่ Alkane, Alkene, Alkyne, Amine, Ester และ Ketone สามารถในขณะที่หมู่ฟังก์ชั่นของสารที่แบบจำลองทำนายความดันไอได้ไม่ค่อยดีนั้น ได้แก่หมู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Alcohol, Aldehyde, Amide, Aromatic, Carboxylic Acid เมื่อทำการคำนวณค่า $\ln(P^{sat})$ หา Antoine Coefficients พบว่า Antoine Coefficients ที่ได้มานั้นเป็นตัวเลขชุดใหม่ สำหรับการสร้างแบบจำลองทำนายค่า Antoine Coefficients นั้น ผลการทำนายของแบบจำลองออกมามีความคลาดเคลื่อนมากจึงไม่สามารถวิเคราะห์ต่อได้

5.2 ปัญหา อุปสรรค และข้อเสนอแนะ

1. ตัวเลขแต่ละตัวภายในลายนิ้วมือโมเลกุลมีความเป็นไปได้ที่โครงสร้างย่อยที่ต่างกันอาจจะอยู่ในตำแหน่งเดียวกัน สามารถแก้ไขได้เบื้องต้นด้วยการเพิ่มจำนวนบิต แต่มีกรณีที่ไม่ว่าจะเพิ่มก็ไม่สามารถแยกออกจากกันได้
2. การสร้างแบบจำลองด้วย Deep Learning มีความซับซ้อนกว่า Machine Learning มากจนทำให้ผู้วิจัยไม่สามารถสร้างแบบจำลองทำนาย Antoine Coefficients ได้สำเร็จ
3. การสร้างแบบจำลองมีความรวดเร็ว แต่การหาแบบจำลองที่เหมาะสมจำเป็นต้องใช้เวลา จึงต้องวางแผนเพื่อเตรียมข้อมูลให้พร้อม
4. ถ้าต้องการให้แบบจำลองแม่นยำขึ้น อาจจะต้องแบ่งสร้างแบบจำลองที่แบ่งแยกขอบเขตให้เฉพาะเจาะจง ปัจจุบันขอบเขตเป็น CHON เปลี่ยนให้เป็นการเทรนเฉพาะกลุ่มหมู่ฟังก์ชั่นที่เลือกไว้ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] G. St. Cholakov, W.A. Wakeham, R.P. Stateva, Estimation of normal boiling points of hydrocarbons from descriptors of molecular structure, Fluid Phase Equilib 163 (1999) 21–42. https://doi.org/10.1016/S0378-3812(99)00207-1.

[2] D. Yaffe, Y. Cohen, Neural Network Based Temperature-Dependent Quantitative Structure Property Relations (QSPRs) for Predicting Vapor Pressure of Hydrocarbons, J Chem Inf Comput Sci 41 (2001) 463–477. https://doi.org/10.1021/ci000462w.

[3] S. Zeck, Thermodynamics in process development in the chemical industry - importance, benefits, current state and future development, Fluid Phase Equilib 70 (1991) 125–140. https://doi.org/10.1016/0378-3812(91)85029-T.

[4] S. Gupta, J.D. Olson, Industrial Needs in Physical Properties, Ind Eng Chem Res 42 (2003) 6359–6374. https://doi.org/10.1021/ie030170v.

[5] L.K. Tsou, S.H. Yeh, S.H. Ueng, C.P. Chang, J.S. Song, M.H. Wu, H.F. Chang, S.R. Chen, C. Shih, C.T. Chen, Y.Y. Ke, Comparative study between deep learning and QSAR classifications for TNBC inhibitors and novel GPCR agonist discovery, Scientific Reports 2020 10:1 10 (2020) 1–11. https://doi.org/10.1038/s41598-020-73681-1.

[6] S. Koutsoukos, F. Philippi, F. Malaret, T. Welton, A review on machine learning algorithms for the ionic liquid chemical space, Chem Sci 12 (2021) 6820–6843. https://doi.org/10.1039/D1SC01000J.

[7] N. Chorbngam, R. Chawuthai, A. Anantpinijwatna, Novel method for properties prediction of pure organic compounds using machine learning, in: 2021: pp. 431–437. https://doi.org/10.1016/B978-0-323-88506-5.50068-1.

[8] Hydrocarbon | Definition, Types, & Facts | Britannica, (n.d.). https://www.britannica.com/science/hydrocarbon (accessed October 29, 2023).

[9] Functional Groups in Organic Chemistry | ChemTalk, (n.d.). https://chemistrytalk.org/functional-groups-organic-chemistry/ (accessed October 31, 2023).

[10] Functional Groups In Organic Chemistry, (n.d.). https://www.masterorganicchemistry.com/2010/10/06/functional-groups-organic-chemistry/?fbclid=IwAR3xcs7ajGWcn4xa6mnbWYr38sWoacQuK2f6B25CHzgMicRMKFI0b70SoyU (accessed November 7, 2023).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[11] 3.5: Properties of Alkanes - Chemistry LibreTexts, (n.d.). https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Organic_Chemistry_(Morsch_et_al.)/03%3A_Organic_Compounds-_Alkanes_and_Their_Stereochemistry/3.05%3A_Properties_of_Alkanes (accessed October 31, 2023).

[12] Solubility – Introductory Chemistry, (n.d.). https://uen.pressbooks.pub/introductorychemistry/chapter/precipitation-reactions/ (accessed November 7, 2023).

[13] Specific heat capacity - Wikipedia, (n.d.). https://en.wikipedia.org/wiki/Specific_heat_capacity (accessed November 7, 2023).

[14] E. Voutsas, Estimation of the volatilization of organic chemicals from soil, Thermodynamics, Solubility and Environmental Issues (2007) 205–227. https://doi.org/10.1016/B978-044452707-3/50013-6.

[15] Usepa, Ocspp, Oppt, Rad, Sustainable Futures / P2 Framework Manual 2012 EPA-748-B12-001 Appendix F. SMILES Notation Tutorial, (n.d.). http://www.epa.gov/ncct/dsstox/MoreonSMILES.html#Tutorials. (accessed October 29, 2023).

[16] D. Rogers, M. Hahn, Extended-Connectivity Fingerprints, J Chem Inf Model 50 (2010) 742–754. https://doi.org/10.1021/ci100050t.

[17] S. Zhong, X. Guan, Count-Based Morgan Fingerprint: A More Efficient and Interpretable Molecular Representation in Developing Machine Learning-Based Predictive Regression Models for Water Contaminants' Activities and Properties, Environ Sci Technol (2023). https://doi.org/10.1021/acs.est.3c02198.

[18] What is Machine Learning? | IBM, (n.d.). https://www.ibm.com/topics/machine-learning (accessed October 29, 2023).

[19] 1.1. Linear Models — scikit-learn 1.3.2 documentation, (n.d.). https://scikit-learn.org/stable/modules/linear_model.html#ridge-regression-and-classification (accessed October 29, 2023).

[20] 1.11. Ensembles: Gradient boosting, random forests, bagging, voting, stacking — scikit-learn 1.3.2 documentation, (n.d.). https://scikit-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

learn.org/stable/modules/ensemble.html#random-forests-and-other-randomized-tree-ensembles (accessed October 29, 2023).

[21] XGBoost Documentation — xgboost 2.0.1 documentation, (n.d.). https://xgboost.readthedocs.io/en/stable/ (accessed October 31, 2023).

[22] 1.6. Nearest Neighbors — scikit-learn 1.3.2 documentation, (n.d.). https://scikit-learn.org/stable/modules/neighbors.html (accessed October 29, 2023).

[23] DIGI, (n.d.). https://digi.data.go.th/blog/what-is-k-nearest-neighbors/ (accessed March 23, 2024).

[24] นิวรัลเน็ตเวิร์กคืออะไร - คำอธิบายเกี่ยวกับนิวรัลเน็ตเวิร์กแบบเทียม - AWS, (n.d.). https://aws.amazon.com/th/what-is/neural-network/ (accessed March 8, 2024).

[25] K-Fold Cross Validation Technique and its Essentials - Analytics Vidhya, (n.d.). https://www.analyticsvidhya.com/blog/2022/02/k-fold-cross-validation-technique-and-its-essentials/ (accessed October 31, 2023).

[26] MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared — Which Metric is Better? | by Akshita Chugh | Analytics Vidhya | Medium, (n.d.). https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e (accessed October 31, 2023).

[27] Y. Ding, M. Chen, C. Guo, P. Zhang, J. Wang, Molecular fingerprint-based machine learning assisted QSAR model development for prediction of ionic liquid properties, J Mol Liq 326 (2021) 115212. https://doi.org/10.1016/j.molliq.2020.115212.

[28] V.V. Santana, C.M. Rebello, L.P. Queiroz, A.M. Ribeiro, N. Shardt, I.B.R. Nogueira, PUFFIN: A path-unifying feed-forward interfaced network for vapor pressure prediction, Chem Eng Sci 286 (2024) 119623. https://doi.org/10.1016/j.ces.2023.119623.

[29] Vapor Pressure (chemicals.vapor_pressure) — Chemicals 1.1.5 documentation, (n.d.). https://chemicals.readthedocs.io/chemicals.vapor_pressure.html#rc9de082557a5-6 (accessed March 21, 2024).

[30] Phase Change Properties (chemicals.phase_change) — Chemicals 1.1.5 documentation, (n.d.). https://chemicals.readthedocs.io/chemicals.phase_change.html#boiling-point (accessed March 21, 2024).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[31] P.K. Ozili, The acceptable R-square in empirical modelling for social science research, (2023).

[32] What Is MAPE? A Guide to Mean Absolute Percentage Error | Indeed.com, (n.d.). https://www.indeed.com/career-advice/career-development/what-is-mape (accessed March 25, 2024).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.
โค้ดสำหรับการแปลง SMILES เป็น Morgan Fingerprint และการสร้างแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โค้ดสำหรับการสร้างแบบจำลองและการแปลง SMILE เป็น Fingerprint

GitHub : https://github.com/Sawahiko/Fingerprint-Based-Machine-Learning-Prediction-of-Chemical-Properties

1) อ่านไฟล์และเลือกคอลัมน์ตัวแปรคุณลักษณะ

```
# Import Data

df = remove_outliers("../Data.xlsx", "Load_AllDataSetC12", 2)

X_data_excel= df[["SMILES"]]

Y_data= df["Tb"]
```

2) สร้างลายนิ้วมือโมเลกุลจาก SMILES ให้เหมาะสำหรับ Machine Learning

```
# Generate Fingerprint from SMILE

MF_radius = 3; MF_bit = 4096

X_data_use = X_data_excel.copy()

X_data_use["molecule"] = X_data_use["SMILES"].apply(lambda x: Chem.MolFromSmiles(x))

X_data_use["count_morgan_fp"] = X_data_use["molecule"].apply(lambda x:
rdMolDescriptors.GetHashedMorganFingerprint(x, radius=MF_radius, nBits=MF_bit, useFeatures=True,
useChirality=True))

X_data_use["arr_count_morgan_fp"] = 0

# Transfrom Fingerprint to Column in DataFrame

X_data_fp = []

for i in range(X_data_use.shape[0]):

    blank_arr = np.zeros((0,), dtype=np.int8)

    DataStructs.ConvertToNumpyArray(X_data_use["count_morgan_fp"][i],blank_arr)

    datafram_i = pd.DataFrame(blank_arr).T

    X_data_fp.append(datafram_i)

x_data_fp = pd.concat(X_data_fp, ignore_index=True)

y_data_fp = Y_data.copy()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) แบ่งข้อมูลเป็น Training Set และ Test Set

```
x_train_fp, x_test_fp, y_train_fp, y_test_fp = train_test_split(x_data_fp, y_data_fp, test_size=0.1,
random_state=42)
```

4) เตรียมพร้อมและเลือกใช้ Machine Learning Algorithm

```
# สำหรับ Ridge

from sklearn.linear_model import Ridge

def Ridge_M(x_train, y_train):

    ridge = Ridge(random_state=42)

    ridge.fit(x_train, y_train)

    return ridge
```

```
# สำหรับ XGB

from xgboost import XGBRegressor

def XGB (x_train, y_train):

    xgb = XGBRegressor(random_state=42)

    xgb.fit(x_train, y_train)

    return xgb
```

5) เทรนโมเดล

```
# สำหรับ Ridge

Model1 = Ridge_M (x_train_fp, y_train_fp)

y_pred_test1 = model1.predict(x_test_fp)

# สำหรับ XGB

Model2 = XGB(x_train_fp, y_train_fp)

y_pred_test2 = model2.predict(x_test_fp)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข.
จุดเดือดของสารอินทรีย์ที่ประกอบไปด้วย คาร์บอน ไฮโดรเจน ออกซิเจน และไนโตรเจนจากแหล่งอ้างอิงและการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.1 ค่าจุดเดือดของหมู่ฟังก์ชัน Alcohol ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C(=O)NCCO)O	Alcohol	601.96	538.88	482.41	586.09	586.70
2	CCCCCCC(C(C)CC)O	Alcohol	542.58	522.52	509.76	546.29	536.51
3	CCC(C)(C)CCCO	Alcohol	471.59	477.60	446.96	465.72	451.40
4	CC(C)CCCC(C)O	Alcohol	447.15	467.11	444.65	440.48	449.98
5	CC(C)N(CCO)CCO	Alcohol	556.12	530.87	524.28	539.44	531.96
6	CN(CC(=O)O)N=O	Alcohol	356.65	459.66	456.94	471.25	482.18
7	CCC(CCC(CC)O)O	Alcohol	566.12	529.08	477.43	566.61	554.88
8	CC1CCC(CC1C)O	Alcohol	462.15	471.98	437.79	465.28	463.20
9	CC(C)(CC(=O)CC(C)(C)O)O	Alcohol	637.29	602.37	480.90	632.84	614.61
10	C(O)O	Alcohol	406.84	404.44	411.67	462.44	475.92
11	CCCC(C)(CC(C)C)O	Alcohol	444.15	469.97	449.75	443.88	463.66
12	CCCCCCCC(CCO)O	Alcohol	612.32	588.15	577.04	605.98	617.50
13	C(CCCC(=O)O)CCCO	Alcohol	620.33	590.02	554.61	585.81	591.97
14	CCCC(C)CC(C)(C)O	Alcohol	408.15	484.87	458.44	480.51	473.07
15	CCC(C(C)C)C(C)O	Alcohol	437.33	471.15	430.13	434.84	448.32
16	C1CC2CC1C3C2C(CC3)C(=O)O	Alcohol	611.94	585.04	476.02	613.64	595.97
17	CC(C)(C)N(CCO)CCO	Alcohol	410.65	558.21	523.51	540.88	543.10
18	CCC(CO)(CO)C(=O)O	Alcohol	663.52	655.39	512.08	612.02	612.79
19	CN(CCO)N=O	Alcohol	436.26	427.80	416.55	446.98	439.14
20	CCCCCC(C)CCO	Alcohol	497.26	493.09	489.42	492.34	483.07
21	CCCC(CC)C(C)(C)O	Alcohol	450.65	486.31	460.69	486.01	473.07
22	CCC(=O)CO	Alcohol	433.15	426.57	396.11	399.02	421.87
23	CC(C)(C)CCCCCO	Alcohol	494.47	493.18	431.01	448.47	474.46
24	CC1CC1C(=O)O	Alcohol	464.15	447.04	463.87	461.45	468.32
25	CCCC(C)(CCC)O	Alcohol	434.15	445.96	455.76	438.18	443.22
26	C1C2CC3CC1CC(C2)(C3)C(=O)O	Alcohol	616.85	606.36	536.69	595.83	595.18
27	C1CCNC(C1)CCO	Alcohol	475.15	505.54	453.20	491.14	490.32
28	CC(CN(CCCCO)N=O)O	Alcohol	619.52	554.71	522.91	577.63	584.66
29	C1=CC=C2C(=C1)C=CC(=C2O)N	Alcohol	495.15	551.52	529.83	489.11	494.85
30	CC(CO)(CO)C1CCCCC1	Alcohol	629.08	570.17	505.70	578.54	618.73
31	C1CCCCCC(CCCC1)CO	Alcohol	607.24	557.84	534.65	566.43	579.01
32	C1CC1(C(=O)O)C(=O)O	Alcohol	612	610.05	580.46	604.84	604.38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.1 ค่าจุดเดือดของหมู่ฟังก์ชัน Alcohol ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	CC(C)CC(C)CCCO	Alcohol	496.82	489.38	452.83	474.95	474.46
34	CCN(CC)CCO	Alcohol	436.15	429.99	428.25	438.67	431.57
35	CCC(C)(CC)C(C)(C)O	Alcohol	491.24	486.34	448.55	478.34	472.80
36	CCC(CC)CCC(CC)O	Alcohol	519.7	500.12	501.33	513.71	500.81
37	CCCCC(CC)CCO	Alcohol	479.65	492.26	484.42	494.40	483.07
38	CC1(CC(CC(N1CCO)(C)C)O)C	Alcohol	654.5	622.30	495.59	632.20	624.69
39	C1=CC(=CC=C1C2=CC=C(C=C2)O)N	Alcohol	380.65	636.74	575.07	551.77	583.89
40	C(CO)NC(=O)N	Alcohol	536.99	535.59	456.73	520.48	538.07
41	CCCCCCCCC(C(C)C)O	Alcohol	542.58	523.04	503.97	542.32	535.74
42	C(CCCCO)CCCC(=O)O	Alcohol	643.21	610.05	568.46	619.09	607.01
43	C1CNCC1O	Alcohol	447.13	436.07	438.14	454.30	455.46
44	C(CNCCO)NCCO	Alcohol	621.58	610.68	521.88	544.40	577.69
45	CCC(CO)C(C)(C)C	Alcohol	471.15	462.89	430.56	456.78	449.74
46	C1CCC(CC1)(C#N)O	Alcohol	392.65	519.98	451.95	458.95	527.30
47	CCC(C)(CC(C)(C)C)O	Alcohol	491.24	455.22	459.40	473.20	465.67
48	CC(C)(C)N=NC1(CCCC1)O	Alcohol	659.19	610.55	449.72	562.66	606.66
49	CC1(C(C(C1O)(C)C)O)C	Alcohol	485.65	521.61	459.85	540.74	532.22
50	C(C(=O)O)N(C=O)O	Alcohol	567.03	563.89	460.52	482.39	519.46
51	CC(C)(CNC(C)(C)CO)N	Alcohol	504.15	509.77	419.77	516.80	518.09
52	CCCC(CCC)CCO	Alcohol	497.26	489.54	471.66	496.73	482.60
53	CCCC(C(=O)O)O	Alcohol	551.25	543.45	514.67	529.59	524.20
54	CC1CCC(C1)(C)O	Alcohol	462.79	448.26	420.90	409.53	438.69
55	C1CCC(CC1)C(CO)O	Alcohol	586.11	567.40	488.51	546.13	556.36
56	CC(CO)CO	Alcohol	468.15	476.39	474.40	476.85	484.93
57	CC(C(=O)N)O	Alcohol	486.38	476.41	444.78	451.90	468.92
58	CCCCN(CCCCO)N=O	Alcohol	333.15	523.79	497.96	519.61	518.94
59	CC(C)C1CCC(CC1)(C)O	Alcohol	481.65	499.79	492.96	513.46	498.62
60	CC(CN(CCO)CCO)O	Alcohol	418.15	602.51	543.69	619.25	592.96
61	CCCCC(C)C(C)(C)O	Alcohol	456.65	481.76	452.47	471.70	473.07
62	CCCC(CC)C(C)CO	Alcohol	466.15	485.29	459.71	475.60	474.83
63	C(C(=O)O)N=C(N)N	Alcohol	635.37	575.43	475.29	469.21	581.69
64	CCCCCCC(CCO)CO	Alcohol	612.32	580.34	528.33	586.72	550.15
65	CCC(C)(C)C(C)CCO	Alcohol	464.15	489.29	449.90	482.56	472.80
66	CCCCN(CCCC)CCCCO	Alcohol	578.78	562.84	488.05	562.56	552.04

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.1 ค่าจุดเดือดของหมู่ฟังก์ชัน Alcohol ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
67	CCC(C)(C)CC(C)(C)O	Alcohol	491.24	478.44	448.10	479.70	472.80
68	CC(C)(CC(C)(C)O)CO	Alcohol	483.15	521.88	483.60	533.04	528.37
69	C(C(CO)C=O)O	Alcohol	523.7	519.42	474.18	516.36	518.73
70	CCC(CC)(C(C)(C)C)O	Alcohol	447.15	463.19	456.73	478.59	471.97
71	CC(C)CC(C(C)C)O	Alcohol	430.65	455.25	432.87	436.07	448.32
72	CCCCCCCC(=O)O	Alcohol	512.15	506.47	559.88	527.72	524.01
73	C1CCC(CC1)CCO	Alcohol	481.15	488.58	448.51	483.54	475.71
74	CCCC(C(C)(C)C)O	Alcohol	429.15	449.00	438.37	445.55	448.47
75	CC1(C2CCC(C2)C1C(=O)O)C	Alcohol	582.56	559.88	506.37	579.58	567.32
76	CC(CC(C)C(=O)O)C(=O)O	Alcohol	649.9	616.65	584.74	630.06	618.09
77	CCC(CC)O	Alcohol	389.35	398.47	415.79	397.20	406.68
78	CCCC(C)(C)O	Alcohol	394.25	421.52	419.76	404.85	420.50
79	CC(C)(C(=O)O)C(C)(C)C(=O)O	Alcohol	498.15	657.98	554.56	649.56	627.37
80	CCCCCC(CCC)CO	Alcohol	490.65	498.71	491.65	504.96	512.10
81	C1C2CC3CC1CC(C2)(C3)CO	Alcohol	563.52	534.40	526.22	558.65	557.16
82	C1CCC(CC1)NC(=O)NCCO	Alcohol	671.46	642.44	536.73	612.38	632.76
83	CCC(C)(C)C(C)(CC)O	Alcohol	438.65	464.85	449.06	478.76	471.97
84	CCCC1CCCCC1O	Alcohol	474.15	498.67	491.21	490.34	491.12
85	C1CNCCC1CCO	Alcohol	500.65	504.22	465.54	491.58	496.94
86	CCC(C)(C)C(C(C)(C)C)O	Alcohol	460.65	495.00	468.74	497.93	497.99
87	CC1(CCC(CC1)C(C)(C)O)O	Alcohol	624.65	573.06	502.24	567.60	603.86
88	C(CO)N	Alcohol	443.95	413.17	404.22	453.42	449.47
89	C1CCCCC(CCCC1)O	Alcohol	557.21	525.25	566.41	523.26	527.40
90	C1CCCC(CCC1)CO	Alcohol	525.79	497.76	485.15	511.42	499.87
91	CC1(CC(C1)C(=O)O)C	Alcohol	477.15	503.52	458.11	483.12	494.03
92	CC(C)CC(C)(C)O	Alcohol	406.25	445.53	424.40	413.15	432.21
93	CCCC(C)(CCC)CO	Alcohol	494.47	482.47	442.63	451.03	474.18
94	CC(C)(CN(C)C)CO	Alcohol	461.15	450.73	398.04	447.03	443.81
95	CCC(C)CCCC(C)(C)O	Alcohol	516.91	500.20	488.09	506.40	498.35
96	C1CC2(CCC1CC2)O	Alcohol	497.08	488.37	455.11	471.94	474.95
97	C1CC1C(=O)O	Alcohol	456.15	456.83	496.51	445.29	462.28
98	CC1(C2CCC1(C(C2=O)O)C)C	Alcohol	597.29	550.31	489.71	554.32	583.67
99	CC1CCC(CC1)O	Alcohol	445.15	461.00	436.26	436.89	448.51
100	CCCCC(C)(C)O	Alcohol	416.15	441.55	433.46	411.17	432.72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.2 ค่าจุดเดือดของหมู่ฟังก์ชัน Aldehyde ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C)CCC=O	Aldehyde	394.15	390.51	385.22	394.10	395.35
2	CCCCCCCCCC=O	Aldehyde	481.65	474.06	475.30	470.51	478.32
3	CN(C)C1=C2C=C(C=C2C(=N1)N(C)C)C=O	Aldehyde	654.3	642.41	537.36	599.26	610.78
4	CCC(C=O)C(C)C	Aldehyde	407.54	409.57	409.77	409.36	404.50
5	CC(C)CC=O	Aldehyde	365.65	353.82	369.51	372.66	372.52
6	CCCCCCCCC(C)C=O	Aldehyde	499.5	489.46	474.94	502.69	505.20
7	CCC(C)(CC(C)C)C=O	Aldehyde	439.65	443.55	432.56	449.09	451.85
8	C1=CC=C(C(=C1)C=O)N	Aldehyde	355.65	483.17	509.48	503.56	501.27
9	CC(C)C1CC2CCC1C(C2)C=O	Aldehyde	535.06	508.86	501.66	536.38	515.53
10	CC(CC1CC2CCC1C2)C=O	Aldehyde	512.58	502.20	494.77	514.15	511.58
11	C1=CC(=C(C=C1O)C=O)O	Aldehyde	333.15	528.96	523.41	547.74	566.52
12	CC1CCCC(C1C=O)(C)C	Aldehyde	487.51	488.16	435.62	480.78	472.89
13	CC(CC=O)CC(C)(C)C	Aldehyde	446.15	424.37	433.13	448.11	446.80
14	CC(C)(CN(C)C)C=O	Aldehyde	417.63	422.93	383.73	413.71	413.13
15	C1CC(C1)C=O	Aldehyde	390.15	388.02	356.89	394.52	381.65
16	C1CCCC(CCC1)C=O	Aldehyde	482.27	468.51	460.40	447.12	464.04
17	CC1CCCC1C=O	Aldehyde	422.15	421.22	396.42	427.86	416.41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.3 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkane ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C)C(C)C(C)C(C)(C)C	Alkane	447.15	450.77	452.91	454.54	449.98
2	CCCCCCCCC	Alkane	423.95	415.62	455.83	421.39	419.21
3	CC(C)C1(CCCCC1)C	Alkane	447.75	449.18	434.49	437.62	435.67
4	CCCC(C)C(C)CC(C)C	Alkane	449.15	441.11	458.65	447.67	450.58
5	CCCC(CC)CC(C)C(C)C	Alkane	467.15	464.28	465.03	468.59	468.33
6	CCCCCCC(C)C(C)CC	Alkane	478.15	469.49	469.48	474.68	471.31
7	CC(C)(C)C(C)(C)C(C)(C)C	Alkane	467.15	448.51	463.25	460.61	452.09
8	CCC(C)(CC)CC(C)(C)C	Alkane	451.15	443.60	454.78	444.25	451.02
9	CCC1CCCC(C1)C	Alkane	420.40	423.34	416.16	421.26	419.95
10	CCCC(CC)C(C)C	Alkane	411.15	413.04	416.03	413.64	417.08
11	CCC(C)C(C(C)C)C(C)CC	Alkane	467.15	459.13	460.46	469.04	468.06
12	CCCC(C)(C)CCC(C)(C)C	Alkane	457.15	465.25	459.35	464.17	468.63
13	CCC(C)C(C)CC(C)C	Alkane	429.65	423.80	437.06	428.91	432.98
14	CCCC(C)(C)CC(C)CC	Alkane	443.15	446.92	446.56	445.98	451.00
15	CC(C)CCC1CCCC1	Alkane	442.15	443.58	447.48	446.78	443.15
16	C1C2CC3C2C1C3	Alkane	366.77	396.23	399.47	371.11	384.75
17	CCC(C)(C)CC(C)(C)C(C)C	Alkane	473.15	468.40	468.53	475.61	468.06
18	CCC(C)CC(C)C(C)(C)C	Alkane	447.15	443.00	453.28	449.36	450.73
19	CCC(C)(C)C(CC)(CC)CC	Alkane	483.15	468.96	464.97	477.64	470.53
20	CCC(CC)C(C(C)C)C(C)C	Alkane	463.15	462.76	459.36	466.29	468.71
21	CC(C)C(C)C(C)(C)C(C)C	Alkane	454.15	451.36	463.23	458.18	449.98
22	C1CC12CC23CC3	Alkane	376.59	395.99	385.95	381.10	390.02
23	CCCCC(C)CC(C)(C)CC	Alkane	463.15	467.16	463.77	466.95	468.47
24	CCC(C)(C)C(C)(C)CC	Alkane	443.15	432.65	448.14	435.06	432.98
25	CCCC(CC)C(C)CC	Alkane	438.15	431.08	438.15	434.69	436.68
26	CCCC(C)(CC)C(C)C(C)C	Alkane	471.15	474.61	465.16	470.60	468.33
27	CCC(C)(CC)C(C)CC(C)C	Alkane	470.15	463.19	464.92	466.32	467.72
28	CCCC(C)CCCC(C)(C)C	Alkane	463.15	465.56	464.15	464.98	467.67
29	CCC(C)(CC)CC	Alkane	391.35	394.49	421.95	384.57	391.43
30	CCCC(CCC)C(C)C(C)C	Alkane	467.15	465.84	460.31	469.05	468.98
31	CCCC(C)(CC)CC(C)(C)C	Alkane	467.15	460.71	462.50	466.45	468.63
32	CCCCCC(C)C(C)C(C)C	Alkane	473.15	466.91	464.94	470.95	468.33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.3 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkane ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	CC1C(C2CC1C3C2CCC3)C	Alkane	484.64	473.25	467.55	489.92	486.31
34	CC(C)CC(C)C(C)C(C)(C)C	Alkane	461.15	465.03	463.70	461.81	467.19
35	CCC(C)CCC(C)C(C)CC	Alkane	473.15	468.06	468.34	470.54	468.42
36	CCC(C)C(C)(C)CCC(C)C	Alkane	467.15	464.11	466.12	467.96	467.72
37	C1C2CC3CC1C3C2	Alkane	393.92	413.12	418.09	394.11	405.55
38	CC1CCCCC1(C)C	Alkane	418.35	423.81	417.58	414.71	415.90
39	CCCCC(CC)CC	Alkane	416.15	413.35	437.52	416.85	415.76
40	CCCC1CC1	Alkane	342.25	366.53	352.58	340.75	367.76
41	CCC(C)(C)C(C)(CC)C(C)C	Alkane	481.15	464.09	472.16	479.85	468.06
42	CCCCC(CCC)C(C)CC	Alkane	470.15	468.42	471.65	471.13	471.96
43	CCCC(CCC)CC(C)C	Alkane	447.15	442.33	449.73	448.95	451.90
44	CCCCC(C)(C)C(C)(C)CC	Alkane	475.15	469.78	468.63	474.43	469.82
45	C1CC2CCC3C2C1CC3	Alkane	448.22	462.93	467.86	449.13	450.19
46	CCC(C)C(C)(CC)C(C)C	Alkane	460.15	443.35	460.07	455.69	450.73
47	CC(C)C(C)CC(C)C(C)C	Alkane	448.15	445.02	453.40	449.05	449.98
48	CCCC(CC)(C(C)C)C(C)C	Alkane	473.15	463.37	467.01	474.13	468.33
49	CCCC(CC)(CC)C(C)C	Alkane	459.15	447.33	459.70	456.42	451.00
50	CCCC(C)C(C)C(C)(C)C	Alkane	449.15	450.64	457.68	449.54	450.64
51	CCC(C)C(C)(C)CC(C)(C)C	Alkane	469.15	459.49	464.36	465.63	467.26
52	CCCC(CC)(CC)C(C)(C)C	Alkane	473.15	469.20	468.19	475.41	469.82
53	CC(C)CC(C)(C(C)C)C(C)C	Alkane	467.15	458.64	462.24	465.20	467.19
54	CCC(C)(CC(C)C)CC(C)C	Alkane	458.15	456.19	457.28	464.78	466.53
55	CCC1CCCC1(C)C	Alkane	411.15	420.84	413.17	414.44	415.90
56	CCC(CC)(C(C)C)C(C)(C)C	Alkane	479.15	471.51	470.91	478.83	468.06
57	CCCCCCC(C)(C)CCC	Alkane	472.15	471.24	470.15	471.43	472.41
58	CCCC(C)CC	Alkane	364.15	374.12	384.00	364.44	371.70
59	C1CC2C3CCC4C1C2C43	Alkane	437.48	436.97	429.68	438.59	447.40
60	CC1CCCC2C1CCCC2	Alkane	353.65	471.74	493.13	474.03	473.38
61	CCC(CC)CC(C)(CC)CC	Alkane	470.15	466.56	468.59	470.10	469.11
62	CC(C)C1CCC(CC1)C(C)C	Alkane	483.90	484.18	474.44	476.66	473.55
63	CC(C)CC(C)C(C)(C)C	Alkane	421.55	420.97	432.51	429.83	432.22
64	CCCCC(C)(CC)C(C)CC	Alkane	474.15	468.62	469.12	473.23	468.47
65	CCCCC(CC)C(C)C(C)C	Alkane	470.15	468.56	465.32	469.74	468.98
66	CCCC(C)CC(CC)CC	Alkane	453.15	447.04	449.15	451.47	451.90

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.3 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkane ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
67	C1CCCC1	Alkane	322.45	354.58	402.96	315.41	345.43
68	CCCCC(CC)CCC(C)C	Alkane	469.15	467.28	471.15	471.49	471.31
69	CC(C)CCC(C)(C)C	Alkane	397.15	406.20	417.17	404.36	414.37
70	CCCC(CCC)C(C)(C)CC	Alkane	469.15	462.07	464.87	471.58	468.47
71	CCC(CC)C(C)C(C)(C)CC	Alkane	472.15	476.11	464.56	471.56	469.07
72	CCC(C)C(C)(CC)C(C)(C)C	Alkane	477.15	462.79	469.76	475.78	468.06
73	CCC(C)CC(C)CC(C)(C)C	Alkane	458.15	459.07	460.71	456.71	466.81
74	CCCC(C)(C)CC(CC)CC	Alkane	464.15	466.12	462.21	465.59	468.47
75	CCCC(C)(C)C(C)(CC)CC	Alkane	478.15	466.04	478.15	477.03	469.82
76	CCCC(CC(C)C)C(C)CC	Alkane	463.15	462.78	465.65	465.83	468.28
77	CC(C)CC(C)CC(C)C	Alkane	420.75	420.22	445.05	430.48	432.98
78	CCCC(C)CCCC(C)CC	Alkane	475.15	470.11	474.15	472.38	471.31
79	C1C2CC34CC1CC3(C2)C4	Alkane	447.30	452.82	509.65	451.41	452.28
80	C1CCC2C3CCC(C2C1)CC3	Alkane	502.52	497.68	505.90	502.05	500.30
81	CCCCC1CC1	Alkane	371.84	386.56	370.77	376.94	383.98
82	CCC(C)C(C)CC(C)C(C)C	Alkane	468.15	463.06	462.56	467.89	468.06
83	CCC(C)CCCC(CC)CC	Alkane	473.15	472.00	467.30	473.93	473.59
84	CCC(C(C)C)C(C)(C)C(C)C	Alkane	470.15	458.23	464.54	471.19	467.89
85	C1CCC2CC(C1)C2	Alkane	404.66	416.50	424.25	406.60	408.97
86	CC(C)CC(C)(C)C(C)C(C)C	Alkane	463.15	467.24	463.00	464.48	467.19
87	CCC(C)(CC)C(C)C	Alkane	417.85	415.93	429.42	410.17	414.37
88	CC1CC(CC(C1)C)C	Alkane	411.65	415.61	417.89	412.19	414.91
89	CC1CCC(C1)C(C)C	Alkane	415.69	423.03	437.28	415.80	417.36
90	CC1(CC1)C2(CC2)C	Alkane	396.60	404.86	386.96	395.47	401.23
91	CC(C)C(C)(C)C(C)(C)C	Alkane	439.25	438.63	445.09	430.15	432.22
92	CCC(C)(C)CC(C)C(C)(C)C	Alkane	460.15	462.28	468.25	469.75	468.06
93	CCC(C)(C(C)C)C(C)(C)C	Alkane	459.15	447.22	462.49	458.05	449.98
94	CCCC(C)C(CC)(CC)CC	Alkane	478.15	463.11	469.82	474.90	468.47
95	CCCCC(C)(C)C(C)C(C)C	Alkane	469.15	473.52	463.94	469.35	468.33
96	CCC(CC)C(CC)CC	Alkane	437.05	435.45	450.06	433.80	436.77
97	CCC1CCC(CC1)C	Alkane	425.75	425.54	434.37	421.31	419.95
98	CCCC(C)(C)CCC(C)C	Alkane	445.15	447.78	453.42	445.63	450.58
99	CCCCCCCC(C)CC	Alkane	461.25	454.24	462.46	460.52	454.60
100	C1CC2C1C3C2CC3	Alkane	393.92	404.48	416.44	395.36	405.55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkene ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	C1CCC/C=C\CC1	Alkene	416.15	421.40	450.25	394.14	409.68
2	CCCCC/C=C/C=C	Alkene	406.36	413.19	410.45	419.24	417.59
3	C=CC12CC3CC(C1)CC(C3)C2	Alkene	497.15	479.55	517.48	491.09	496.81
4	C1CC2CC1C3C2C=C3	Alkene	420.23	437.23	433.69	420.03	425.13
5	CC(C=C)N	Alkene	335.45	362.70	329.07	339.44	341.34
6	C1CC(=NC1)/C=C/C(=O)N	Alkene	568.11	569.30	433.61	506.21	551.92
7	CCC(=CC(C)(C)CC)C	Alkene	424.65	428.18	422.22	437.39	434.52
8	C=CCN(CCO)CC=C	Alkene	480.62	462.24	457.50	492.05	484.00
9	CC1=CCC(C1(C)C)C(=O)O	Alkene	514.65	529.56	477.95	532.80	528.39
10	CC1=CCCC1	Alkene	348.65	378.64	389.86	345.86	370.55
11	CC/C=C\CCCCCCCCO	Alkene	570.50	564.25	530.24	568.72	566.95
12	C(=C\C#N)\C#N	Alkene	499.44	481.47	417.98	496.48	482.08
13	C/C=C(\C)/C(C)O	Alkene	432.66	413.22	383.52	406.44	420.50
14	C=C1C2CC(C1=C)C=C2	Alkene	420.75	436.73	446.05	423.07	418.66
15	CCCCC/C=C\CCO	Alkene	501.86	501.27	461.49	482.46	491.12
16	CC1=CCC=C1	Alkene	346.65	368.81	353.48	344.27	367.93
17	CCC(=C)C#CC(C)(CC)O	Alkene	522.91	497.15	490.20	512.55	503.98
18	CCCC(CC=C)C(=O)O	Alkene	524.39	511.71	497.57	513.45	513.48
19	CC1=C(C(=O)CC1)C	Alkene	456.65	440.50	417.65	423.28	437.02
20	C/C=C(/C)\O	Alkene	387.34	391.82	365.52	375.18	394.07
21	C=CC1(CCC=CC1)OO	Alkene	512.87	484.03	407.80	521.90	482.89
22	CCCCCCCCCC(=C)C	Alkene	484.75	471.34	486.11	485.52	488.42
23	C1CC2CCC1CC=C2	Alkene	430.97	427.98	426.61	429.05	429.53
24	CCC#C/C=C/CCCCCCO	Alkene	579.50	561.58	499.78	557.70	567.72
25	CC(=C)C(=O)N	Alkene	488.00	409.43	411.49	435.01	423.89
26	C1=CC23C=CC=CC2(C=C1)C=C3	Alkene	508.14	504.04	475.24	514.64	508.96
27	CCC(=CC#C)CC	Alkene	376.80	383.62	377.46	382.80	391.17
28	CC(=C1C2C=CC1C(=C)C2=C)C	Alkene	495.91	499.73	466.97	486.45	486.45
29	CC(=C)CCCC(C)(C=C)O	Alkene	510.59	489.36	497.97	516.16	517.46
30	CC(C)(C=C)O	Alkene	370.15	396.46	372.61	393.79	391.39
31	C=C1CCC23C1(C(=O)CC2)C(=O)CC3	Alkene	642.87	594.13	553.86	611.50	618.04
32	CNC/C=C/CNC	Alkene	441.38	442.87	399.64	471.71	460.34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkene ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	C/C=C/1\CC2CC1C=C2	Alkene	420.93	417.88	416.97	422.52	421.68
34	CCCCC/C=C\C=C\C=O	Alkene	485.38	482.20	453.14	463.80	476.50
35	CC1=C(CN(N(C1)C)C)C	Alkene	432.32	434.32	418.68	443.53	459.27
36	C/C=C\C=C/C	Alkene	358.15	335.71	347.73	343.15	360.07
37	CCCC1C2CC(C1C=O)C=C2	Alkene	380.15	502.81	461.78	509.83	508.24
38	C1CC2CC2C=C1	Alkene	376.67	394.19	391.53	376.34	386.84
39	CC(CCO)C=C	Alkene	425.30	432.75	368.10	429.32	416.39
40	CC1(C2CC1C(=C)C(=O)C2)C	Alkene	508.70	469.96	486.16	501.13	486.45
41	CCC/C=C/C=C	Alkene	360.60	373.13	391.66	376.44	364.68
42	CC1=CC2=CC=CC=C2C1	Alkene	481.15	437.42	480.45	473.11	449.58
43	C1=CC(=O)C=CC1O	Alkene	514.75	519.24	446.56	486.05	489.05
44	CCCCCC1=C(CCC1=O)C	Alkene	416.65	515.93	497.19	514.51	529.33
45	C1CC12C3C=CC2C(=O)C3=O	Alkene	560.78	548.61	510.42	553.31	554.01
46	CC(CCC=C)CCO	Alkene	471.06	467.65	426.49	467.30	480.51
47	CC(C)C1C=CC=CC=C1	Alkene	449.26	443.66	427.23	448.05	442.90
48	CC/C=C(\C)/C1CC1	Alkene	402.45	395.02	384.99	395.05	400.11
49	CC1(C=C1)C#N	Alkene	422.22	417.04	335.77	401.49	420.50
50	CC(=CCCC=C)C	Alkene	392.30	385.46	391.00	388.65	390.89
51	C1C=CCC1CC(=O)O	Alkene	519.71	516.32	412.09	464.53	499.44
52	CC1=CCCC(C1)(C)C	Alkene	413.15	434.16	439.24	415.12	416.45
53	CCCC#CC(=C)C	Alkene	388.20	398.15	401.76	387.35	389.67
54	CCC=CCCC(=O)O	Alkene	509.43	486.92	460.48	469.58	494.59
55	CC/C=C\CO	Alkene	411.15	416.78	371.04	412.70	417.74
56	CC1=CCC2=CC=CC=C12	Alkene	471.15	423.80	476.35	485.99	481.18
57	C(#N)/C(=C(\C#N)/N)/N	Alkene	644.26	518.97	437.47	499.04	529.80
58	CC1=CC(=C(C=C1)C)NCC=C	Alkene	362.15	518.59	472.53	469.24	465.72
59	C1C2CC(=O)C1C=C2	Alkene	444.49	429.96	407.91	424.42	440.63
60	CCCCCCCC/C=C/C=O	Alkene	504.10	498.48	487.61	507.64	507.93
61	C=CC1=CC=CC=C1C=C	Alkene	473.15	411.88	457.56	467.57	436.79
62	C1CCCC/C=C/CCCC1	Alkene	496.01	482.11	512.53	488.91	484.07
63	CC(=CC=O)C	Alkene	407.15	363.05	342.69	369.56	373.09
64	C=CCNC1CCCC1	Alkene	430.65	429.08	460.66	452.53	435.74
65	CC(=CN1CCC1)C	Alkene	387.65	398.84	389.62	388.42	385.64
66	C1CC/C=C\C(=O)CC1	Alkene	482.38	449.33	455.03	476.19	461.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkene ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
67	C1C=CCN1N=O	Alkene	381.80	392.88	402.32	387.81	394.27
68	C(/C=C/CO)O	Alkene	404.15	473.86	421.00	489.84	491.54
69	C1CC=C/C(=C/2\CCCC=C2)/C1	Alkene	530.04	511.67	483.93	519.33	507.14
70	CC(C)CC1=CCCC1	Alkene	429.17	431.30	428.13	419.31	422.81
71	CC1=CCC(CC1)C(=C)C	Alkene	451.15	446.42	493.57	442.86	445.05
72	CC1CCC(CC1=O)C(=C)C	Alkene	360.65	488.32	472.03	497.48	487.44
73	CCCC/C=C\C	Alkene	371.55	372.35	390.65	368.94	376.36
74	C1=CC2=NC=CC3=NC=NC(=C1)N2 3	Alkene	636.00	637.57	482.69	477.21	554.78
75	C/C=C\N1CC1	Alkene	337.74	356.73	358.49	338.40	350.32
76	CCCCC/C=C\C	Alkene	398.75	392.38	409.27	393.64	400.02
77	C1C2C=CC1C3C2C(=O)N3	Alkene	513.72	516.45	456.61	508.27	509.61
78	CC/C(=C\C)/C=O	Alkene	389.58	390.45	388.69	398.94	395.63
79	CC(=CCCC(=C)C=C)C	Alkene	440.15	421.38	431.95	440.41	439.27
80	C/C=C/CCC1=CC=CC=C1	Alkene	474.15	472.46	474.13	481.60	488.30
81	C=CCCCCC#N	Alkene	458.52	453.58	436.97	455.23	452.45
82	CC1CC=C(C(=O)C1)C(C)C	Alkene	519.47	481.22	472.12	491.40	488.11
83	C1CCC(CC1)C2=CCCCC2	Alkene	522.07	505.51	498.49	505.81	506.97
84	C1C=CC(=C1)O	Alkene	429.43	429.59	337.72	413.80	419.98
85	CC(C)C(C)(C(C)C=C)O	Alkene	490.27	467.21	451.30	463.65	469.85
86	C=CCN1C=CN=C1	Alkene	413.70	419.54	466.11	453.87	428.73
87	CCC(=O)/C=C/C	Alkene	411.65	392.39	382.60	403.38	402.08
88	CC1(CC2CC1CC2=C)C	Alkene	425.15	436.75	444.89	440.13	436.23
89	CCCC1CCC=C1	Alkene	399.05	395.96	401.82	404.64	403.78
90	C12C3C4C1=C5C2C3=C45	Alkene	409.36	474.30	405.49	389.14	406.93
91	C1CC2C=CC1C3C2C(=O)C=CC3=O	Alkene	636.48	619.93	544.09	625.04	625.35
92	C=C1CC2CCC1C=C2	Alkene	425.86	423.55	428.95	425.04	420.82
93	C1CCC(=CC1)C(=O)O	Alkene	514.15	509.12	505.62	501.64	508.78
94	CC(=CCCC(=O)C)C	Alkene	446.65	431.18	428.82	444.40	445.77
95	CC1=CC=CC1	Alkene	345.93	381.45	368.10	345.30	367.93
96	CC/C=C\C=C	Alkene	346.25	363.22	346.07	342.69	341.40
97	CC(=C(CC(=O)O)C(=O)O)C	Alkene	654.70	584.16	600.36	645.06	614.95
98	CCCC=CCC1CCCC1=O	Alkene	538.54	514.73	469.58	509.70	516.27
99	CCCCCC/C=C\CCCCO	Alkene	570.50	565.99	570.50	569.27	560.32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.4 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkene ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
100	C1CCC(=CC1)CC(=O)O	Alkene	556.51	537.16	467.86	522.86	520.21

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.5 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkyne ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C)C(C#C)(C(C)C)O	Alkyne	436.15	436.80	450.92	465.20	460.14
2	CC(C)(C)C#CC=O	Alkyne	414.19	413.96	384.97	403.53	404.50
3	CC(C)CCC(C)(C#C)O	Alkyne	484.15	470.14	443.50	459.73	458.87
4	C#CC#CC#CC#CC#N	Alkyne	524.72	515.77	436.21	533.00	517.40
5	C#CCCC#CC#CCCCCO	Alkyne	574.46	570.84	456.14	574.37	551.38
6	CCCCCCC#CCCCC	Alkyne	483.16	483.62	483.15	483.94	485.54
7	C1CCCCC#CCCCC1	Alkyne	497.31	482.67	502.01	482.91	484.07
8	CCCCC(C)(C#CC)O	Alkyne	503.47	486.60	434.12	440.95	464.52
9	CC(=O)CC#CO	Alkyne	386.15	438.47	420.58	453.90	447.34
10	CCCCCCC#CCCO	Alkyne	529.58	512.14	461.02	526.57	503.91
11	CCC#CCC	Alkyne	354.15	361.27	391.56	343.45	360.83
12	CCCCCCCCCCC#C	Alkyne	488.15	466.91	501.63	488.61	483.32
13	CNCC#C	Alkyne	356.15	344.44	337.97	349.33	330.52
14	C#CC#CC#CC#CC#CC#C	Alkyne	490.40	501.33	461.55	501.64	486.33
15	CCCCC(CC)C(C#C)O	Alkyne	509.82	498.42	497.55	519.52	498.88
16	CC(C#CC1CCCC1)O	Alkyne	521.54	521.46	432.29	534.16	544.41
17	CCC#CC#C	Alkyne	336.00	352.52	350.44	341.58	350.84
18	CCCCC#CCCO	Alkyne	370.15	472.08	427.45	474.12	461.64
19	CC#CCCCCCCCC(=O)O	Alkyne	605.79	575.14	518.52	572.02	562.72
20	CC(C)CCCC(C)(C#C)O	Alkyne	460.65	490.17	474.14	514.02	491.99
21	CCCCC(C#C)O	Alkyne	435.15	447.62	424.13	435.33	432.57
22	C1CC#CCCC#C1	Alkyne	416.32	430.49	428.24	410.26	407.67
23	CC(C)(C#CC#CC(C)(C)O)O	Alkyne	624.30	575.06	468.10	571.44	600.11
24	C#CO	Alkyne	327.66	338.55	350.66	372.29	371.25
25	CCCCCCC#CCO	Alkyne	506.70	496.81	457.30	467.41	491.12
26	CCCCC#C	Alkyne	344.45	346.74	372.85	340.23	351.96
27	CC#CC(=O)C(C)(C)C	Alkyne	442.28	439.25	401.91	432.96	435.72
28	CCC(C)C1CCCCC1(C#C)O	Alkyne	571.14	553.98	488.02	579.12	552.50
29	CCC(C)(C)C#C	Alkyne	343.15	365.97	348.80	356.82	360.79
30	CCCCCCCCC#CCC	Alkyne	492.55	482.53	493.55	484.39	486.01
31	CC(C)(C#C)C#C	Alkyne	336.77	329.11	355.68	357.31	360.32
32	CCC(CC)C#CC=O	Alkyne	439.86	428.72	385.47	439.97	435.98

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.5 ค่าจุดเดือดของหมู่ฟังก์ชัน Alkyne ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	CCCC#CC1=CC=CC=C1	Alkyne	472.15	430.24	473.65	446.80	475.45
34	C1=CC=C(C=C1)C#CCO	Alkyne	410.65	450.28	462.82	443.70	474.31
35	CCC(C)(C#CC#CC(C)(CC)O)O	Alkyne	670.06	614.59	512.05	658.84	634.20
36	CCC(C)(C#CC(C)(CC)O)O	Alkyne	408.15	562.07	495.95	570.39	584.92
37	CC#CC(=O)C(C)C	Alkyne	422.19	410.80	392.83	425.36	415.65
38	CCCC#CCC(C)O	Alkyne	483.38	475.82	430.88	470.07	461.97
39	C#CCCCCO	Alkyne	419.18	430.29	413.67	444.22	424.52
40	CC1(CCCC(C#C1)(C)C)C	Alkyne	471.37	468.70	448.62	455.99	455.22
41	C#CCCC(=O)O	Alkyne	371.15	432.95	433.93	454.76	442.56
42	CCCC(C)(C#CC(C)(CCC)O)O	Alkyne	661.06	608.42	493.59	658.60	636.20
43	CCCCCC#CC	Alkyne	410.75	404.18	422.48	399.73	400.02
44	CCNCC#CCNCC	Alkyne	491.98	471.50	432.59	513.63	502.32
45	CCC(CC)(C#C)O	Alkyne	411.15	432.89	415.40	418.95	429.56
46	CCCCCC#C	Alkyne	372.85	366.77	400.49	374.12	368.68
47	CC#CC#CC#CC	Alkyne	409.64	429.77	403.86	398.51	403.42
48	CC(C#C)(C1CC1)O	Alkyne	418.65	446.71	418.85	433.10	435.01
49	CCCCCCCC#CCCC	Alkyne	483.16	483.62	483.15	483.94	485.54
50	C#CCCC#N	Alkyne	406.20	410.14	367.36	415.69	406.74
51	CC(C)(C)C#C	Alkyne	310.85	341.52	329.19	331.54	344.59
52	C#CCO	Alkyne	387.65	360.65	370.70	386.89	388.81
53	CCCCCC#CCO	Alkyne	483.82	476.78	423.28	467.80	467.30
54	C#CC1(CC1)C#C	Alkyne	373.00	359.71	356.44	365.57	363.19
55	CCCCCCC#CCCC	Alkyne	471.65	463.59	471.15	466.11	463.91
56	C#CCCCCCCO	Alkyne	464.94	470.35	448.77	480.51	462.09
57	CN(C)CC#CCN(C)C	Alkyne	416.52	399.52	405.18	431.27	415.18
58	C#CCNCC#C	Alkyne	367.29	384.01	330.31	398.64	386.74
59	CC(C)C#CC(C)C	Alkyne	390.76	402.93	378.93	386.03	390.31
60	CC(C#C)(C1=CC=CC=C1)O	Alkyne	490.65	479.06	463.86	461.12	434.08
61	CCC(=O)C#CC	Alkyne	399.75	397.40	383.44	401.22	402.08
62	CCC#CC(=O)O	Alkyne	468.51	479.12	462.47	461.76	468.32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.6 ค่าจุดเดือดของหมู่ฟังก์ชัน Amide ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CN1CCCC1=O	Amide	475.15	418.77	412.65	410.89	410.71
2	CNC(=O)C(=O)NC	Amide	499.20	519.94	502.32	553.78	540.63
3	CN(C)CCCN1CCCC1=O	Amide	513.90	513.19	487.85	505.65	512.02
4	CCCCCCC(=O)NO	Amide	555.98	524.72	476.28	526.88	521.89
5	CCC1CCCCN1C(=O)C=C(C)C	Amide	559.79	563.67	469.63	518.16	525.15
6	CN1CCC(=O)NCC1	Amide	489.91	468.06	419.02	473.67	476.98
7	C(C#N)NC=O	Amide	469.15	450.66	436.11	485.53	472.61
8	CC1=CC=CC=C1N(C)C(=O)C	Amide	533.15	505.23	479.33	483.72	486.98
9	CC(C)C(=O)N(C)C	Amide	448.65	393.11	424.10	424.74	415.46
10	C1CC1C(=O)NN	Amide	474.43	452.29	418.71	442.79	457.82
11	C=CC(=O)N1CN(CN(C1)C(=O)C=C)C(=O)C=C	Amide	674.54	660.66	585.20	582.33	563.38
12	C=CN1CCCCCC1=O	Amide	483.80	464.65	478.78	484.62	480.01
13	CCC(=O)NCCCN(C)C	Amide	499.12	519.87	454.77	497.73	507.27
14	C1CCCC(=O)NCCC1	Amide	536.04	533.60	508.19	511.23	517.20
15	CC(C)(C)C(=O)N	Amide	485.20	443.71	391.07	468.84	457.12
16	CCC(C)CC1(C(=O)NC(=O)N1)CC	Amide	676.22	624.34	550.24	623.64	612.06
17	CCN1C(=O)C=CC1=O	Amide	499.80	497.64	464.05	497.61	469.74
18	CC(C)C1(C(=O)NC1=O)C(C)C	Amide	600.08	515.11	456.77	573.96	528.32
19	CCCCCCCCC(=O)N1CC1	Amide	477.35	536.72	508.05	529.73	521.92
20	C1C(=O)NCC(=O)N1	Amide	548.08	480.24	489.96	548.25	540.63
21	CN(C)C(=O)N	Amide	407.08	427.57	491.15	468.58	443.48
22	CCC1(C(=O)NC(=O)N1)C	Amide	585.14	538.57	555.54	587.75	576.86
23	C1CCC(CC1)N2C(=O)CCC2=O	Amide	611.71	584.07	497.66	610.84	619.39
24	CN1CCCCCCCCCCC1=O	Amide	599.99	558.98	497.36	461.24	495.39
25	CC(=O)N(C)C(=O)C	Amide	468.15	435.45	446.20	446.50	445.44
26	CC1(C(=O)NC(=O)N1C)C	Amide	544.76	495.18	496.41	550.54	523.76
27	CN(C)CCCNC=O	Amide	448.15	468.67	402.32	459.56	476.18
28	CC1CN(CC(N1C(=O)C)C)N=O	Amide	531.13	521.61	499.58	519.01	511.19
29	CCN(CC)C(=O)C	Amide	458.65	420.24	407.56	413.35	416.41
30	CCC(=O)N(CC)CC	Amide	464.15	434.57	437.18	436.87	435.78
31	C1CNCCC1C(=O)N	Amide	531.38	514.52	487.20	532.02	537.25
32	C1C(CN(C1=O)CC(=O)N)O	Amide	646.73	630.70	485.51	584.35	596.30

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.6 ค่าจุดเดือดของหมู่ฟังก์ชัน Amide ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	CC=CC(=O)NCCN1CCNC1=O	Amide	658.21	648.12	530.18	613.63	609.03
34	CCCN1C(=O)C=CC1=O	Amide	522.68	514.95	470.17	513.57	504.09
35	C1C(=O)NC(=O)NC1=O	Amide	615.90	547.80	559.23	587.89	586.66
36	CC1C2=CC=CC=C2N(C1=O)C	Amide	546.15	458.51	489.36	487.49	487.87
37	CC1(C(=O)N(C(=O)N1)CCO)C	Amide	659.82	609.28	487.04	602.15	610.30
38	CN(C=C)C=O	Amide	422.15	376.66	368.05	365.70	357.36
39	C1CCN(CC1)C(=O)N	Amide	341.15	498.88	505.71	502.04	504.07
40	C1C(=O)NC(=O)N1	Amide	520.93	468.56	505.64	530.83	533.20
41	C1CC(=O)NC(=O)C1	Amide	522.41	513.61	510.87	530.86	515.63
42	CCN(CC)C(=O)CC1=CC=CC=C1	Amide	443.15	540.78	489.67	497.30	488.69
43	C1CC2C3CC(CNC3)CN2C(=O)C1	Amide	423.15	590.69	524.78	570.40	588.55
44	CCN(C(=O)NC)N=O	Amide	471.00	440.66	439.36	440.57	461.18
45	CCCCCC(=O)NO	Amide	533.10	504.69	474.95	499.46	494.12
46	CC1=NNC(=O)C1	Amide	485.28	482.94	425.02	485.00	480.35
47	CN1C(=O)C2CC=CCC2C1=O	Amide	610.90	537.27	480.89	575.88	575.73
48	C1CC(=O)N(C1=O)C2=CC=CC=C2	Amide	673.15	544.05	495.38	495.02	493.77
49	CCN1C(=O)CCC1=O	Amide	509.15	488.20	510.68	495.53	474.45
50	CCC(C)N(C(C)CC)C(=O)N	Amide	558.85	549.84	440.45	525.59	544.55
51	CCCCN1C(=O)CCC1=O	Amide	546.40	525.53	488.93	513.90	530.91
52	CC1(C(=O)NC(=O)N1)C	Amide	562.26	525.74	550.00	567.84	556.01
53	CC(C)NC(=O)CN1CCNCC1	Amide	590.06	530.41	506.72	562.37	576.06
54	CCN(CC)C(=O)N	Amide	452.84	465.22	412.40	472.14	465.33
55	CCCC1(CC(=O)NC1=O)C	Amide	582.35	532.94	509.78	574.31	554.28
56	CCN(CC)C(=O)C(C)(C)C	Amide	479.15	457.58	437.63	488.85	477.33
57	CCCN(C1=CC=CC=C1)C(=O)C	Amide	541.15	508.03	508.65	493.08	488.76
58	C1CCN(CC1)C(=O)CC#N	Amide	570.98	554.64	486.79	577.23	569.81
59	CNC(=O)N1CC1	Amide	414.74	441.70	419.71	463.10	467.24
60	CC(=CCN1CC2CCC(C1=O)N2C)C	Amide	584.38	566.72	510.54	520.61	521.31
61	CC(=O)N1CCCC=C1	Amide	445.18	443.42	418.76	459.94	451.10
62	CC1C(=O)NC(=O)N(C1=O)C	Amide	616.61	557.62	525.97	592.43	585.88
63	CN1CC2CCC(N2C)CC1=O	Amide	427.65	516.29	465.40	505.58	493.58
64	C(=O)N	Amide	493.15	394.14	399.06	440.42	430.55
65	CCCCCCCCCC(=O)N	Amide	554.80	554.91	495.03	521.52	557.88
66	CN1CCCCCCCC1=O	Amide	518.54	498.89	464.34	495.24	484.94

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.6 ค่าจุดเดือดของหมู่ฟังก์ชัน Amide ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
67	CC(C)N(C(=O)N)N=O	Amide	492.92	440.65	441.99	468.13	459.67
68	CCCCN(C1=CC=CC=C1)C(=O)C	Amide	554.15	528.06	494.24	494.22	490.35
69	C1CN(CC1C(=O)N)CC2=CC=CC=C2	Amide	394.65	608.77	486.66	504.77	503.13
70	CC(=O)N1CCCC1=O	Amide	391.15	482.42	375.08	436.36	466.91
71	CC(C)NC(=O)C	Amide	494.90	449.18	434.42	466.59	457.12
72	CCN(CC)C(=O)N(CC)N=O	Amide	563.15	488.75	463.82	496.08	496.85
73	C1CCC2C(C1)C(=O)N2	Amide	498.15	505.30	462.74	492.16	500.17
74	CC1CCCNC1=O	Amide	522.65	466.01	485.71	471.02	470.22
75	C1CC2CCC(=O)N(C1)C2	Amide	489.59	486.16	499.63	491.50	485.12
76	CC(=O)N1CCC(=O)CC1	Amide	491.15	497.98	460.69	511.28	502.84
77	CN1CC(=O)N(CC1=O)C	Amide	578.15	507.78	482.01	514.94	500.35
78	CC(C)C(=O)NC	Amide	417.60	441.59	505.65	474.52	457.60
79	CC1C=NNC1=O	Amide	475.63	435.00	405.95	447.00	453.48
80	CC(=O)NNC(=O)C	Amide	499.20	444.97	491.52	542.89	540.96
81	CC(CNC(=O)C)NC(=O)C	Amide	567.40	579.62	523.89	577.57	575.19
82	CC(C)N1C(=O)C=CC1=O	Amide	522.24	520.13	462.73	510.95	496.83
83	CN(C)C(=O)C#N	Amide	459.51	400.39	393.22	427.90	447.39
84	C(CC(=O)N)CN	Amide	490.05	489.62	449.10	528.59	530.26
85	CCCCCCNC(=O)CCCC	Amide	555.32	574.62	543.12	592.05	571.18
86	CCN(CC)C(=O)C=C	Amide	377.15	438.50	408.15	454.40	441.72
87	CC(C)(C)C1=NNC(=O)C1	Amide	550.69	521.81	428.29	515.68	506.26
88	CCCC(=O)N	Amide	489.15	434.73	430.69	452.73	446.44
89	C1CC2CC1C3C2NC3=O	Amide	514.56	522.58	455.38	527.06	536.34
90	C1CCC2(C1)C(=O)NC(=O)N2	Amide	623.70	596.57	524.58	613.88	604.66
91	CCCCCCCCCCCN(C(=O)N)N=O	Amide	676.40	610.16	525.10	566.16	496.68

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.7 ค่าจุดเดือดของหมู่ฟังก์ชัน Amine ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	C1CC(C1)N	Amine	355.15	381.05	356.19	336.86	370.29
2	CCNCCCN	Amine	429.15	431.64	423.71	424.07	424.59
3	CNCCN(C)CCN(C)CCNC	Amine	553.62	527.49	500.83	497.03	542.33
4	CC1(C2CCC1(C(C2)N)C)C	Amine	509.82	499.22	457.06	498.50	480.36
5	C1CNCCN1	Amine	419.15	427.04	508.91	443.99	424.00
6	CC(C)(C)CNC	Amine	383.82	384.00	345.07	376.78	383.63
7	C1CCCN2CCCN(C2)CC1	Amine	493.18	475.89	489.65	489.78	480.56
8	C1CC1N	Amine	323.65	361.02	337.44	326.06	366.49
9	CN1CCCCCCC1	Amine	423.57	424.25	462.93	426.09	421.89
10	CC1CCCC1N	Amine	396.15	413.89	389.77	399.61	405.61
11	CCCC(C)N	Amine	365.15	384.59	374.32	374.47	374.07
12	CCCCCCN(C)C	Amine	420.15	411.75	437.49	426.05	411.91
13	C1N2CN(CN1CN(C2)N=O)N=O	Amine	509.11	444.64	427.55	489.84	485.80
14	CC1(CC(=NO)CC(N1)(C)C)C	Amine	640.77	550.37	493.41	542.44	574.42
15	CCCCCNC	Amine	390.15	391.83	400.25	396.42	397.82
16	CCC(C)N(C)CC	Amine	376.15	382.25	371.06	376.94	377.31
17	CC1CCC(NC1)C	Amine	407.15	430.15	388.10	415.86	412.85
18	CC(CN1CCCCC1)N	Amine	466.65	461.46	433.15	442.20	443.39
19	CCCCNCCCCCCN	Amine	551.10	533.68	490.40	517.00	541.13
20	CN(C)C#N	Amine	436.65	355.86	339.55	398.20	385.32
21	CC1CCNCC1	Amine	403.15	406.12	435.09	395.69	406.67
22	C1CC2CC1CC2CN	Amine	468.25	458.83	430.66	460.98	444.29
23	CC1CCC(CC1N)N	Amine	515.03	489.82	436.98	465.60	475.84
24	CC1CNC(CN1)C	Amine	437.15	468.54	431.01	443.13	452.07
25	CCC1(CCNCC1)C	Amine	443.65	439.79	422.95	428.58	433.56
26	C1CCCNCCC1	Amine	447.65	437.77	429.42	422.95	432.55
27	CC(C)(C)N(C)C	Amine	363.05	362.40	354.94	354.99	362.60
28	C1CNC2=NCCCN2C1	Amine	514.22	513.98	413.94	441.00	490.69
29	CCNCCC(C)C	Amine	399.15	402.60	404.86	403.25	405.14
30	C1CC2=NCCCN2C1	Amine	369.15	472.10	431.96	439.65	452.19
31	CC(CC1CCCC1)NC	Amine	444.15	460.03	423.81	438.27	452.65
32	CC(C)(C)C1CCNCC1	Amine	470.39	455.64	458.77	460.04	447.77

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.7 ค่าจุดเดือดของหมู่ฟังก์ชัน Amine ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	C1CCC(C1)N	Amine	381.15	401.41	423.71	387.85	385.56
34	CC(C)N(C)C(C)C	Amine	383.65	385.98	373.37	375.43	377.13
35	CCC(CC)N	Amine	362.15	379.63	374.89	374.54	374.07
36	CC(C)NCCNC(C)C	Amine	443.15	451.88	437.58	515.81	476.13
37	C1CCC(CC1)N2CCCCC2	Amine	372.15	483.81	478.65	477.64	494.07
38	CCC(C)N	Amine	335.65	365.53	348.96	327.76	362.42
39	CC(C)C(C)N	Amine	358.65	371.78	336.79	367.66	372.39
40	C(CN)CNCCCNCCCN	Amine	650.92	632.81	552.88	584.44	591.48
41	CC(C)CN1CCCCC1	Amine	433.65	436.47	439.43	435.12	434.13
42	CCCNCCC#N	Amine	489.13	488.33	449.15	482.79	479.33
43	CC1CN(CCN1)CCNC(C)CN	Amine	626.93	609.20	568.87	630.18	636.61
44	CCC(C)NCC	Amine	370.15	383.69	383.75	381.30	392.14
45	C1CCC(C1)NCN	Amine	474.86	454.73	428.08	448.87	468.68
46	C1CCCCNCCC1	Amine	468.22	457.80	451.85	458.69	449.63
47	CCN(C)CCN(CC)CC	Amine	430.40	422.32	424.45	447.94	438.21
48	CCN1CCCC(C1)N	Amine	428.15	456.84	416.26	446.35	427.17
49	CNCCN	Amine	388.15	393.02	375.65	411.09	408.69
50	CC1CCNC(C1)C	Amine	423.19	423.12	402.59	414.11	412.85
51	CC(C)N(C)C	Amine	339.15	335.32	362.20	334.97	335.28
52	CCCCNCC#N	Amine	489.13	470.25	419.94	469.42	477.69
53	C(CCCN)CCCN	Amine	497.15	485.34	480.82	465.01	481.86
54	CC(CN)C#N	Amine	465.29	449.27	401.48	451.23	445.02
55	CCC(C)CC(CC)N	Amine	454.29	437.57	413.89	437.69	422.80
56	CC1CC(NC(C1)(C)C)C	Amine	429.15	483.87	415.58	427.18	443.15
57	CC1(CCCC(N1C)(C)C)C	Amine	420.15	491.37	444.73	444.55	441.52
58	CNCCCCNC	Amine	441.15	428.13	404.78	466.52	459.11
59	CC1CC(CCNC1)(C)C	Amine	473.46	462.06	427.87	458.17	447.77
60	CN1CC2CCC1CC2	Amine	412.83	406.96	433.06	420.51	414.34
61	CCNCCCN(CC)CCCNCC	Amine	586.94	533.95	530.57	499.37	501.98
62	CCCCN(CCCC)CCC#N	Amine	565.80	549.22	468.71	574.55	541.60
63	CCCNC(C)(C)C	Amine	386.15	386.76	371.70	394.31	399.59
64	CCCCCCNCCCCCC	Amine	509.15	518.85	496.07	512.12	500.01
65	CC(C)CNC1CCCCC1	Amine	466.15	485.86	484.44	490.59	490.49
66	CNC12CC3CC(C1)CC(C3)C2	Amine	521.51	533.13	499.20	485.04	486.49

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.7 ค่าจุดเดือดของหมู่ฟังก์ชัน Amine ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
67	C1CN2CCC1CC2	Amine	394.62	396.50	440.62	402.53	398.35
68	CC1CC(C(C1CN)N)(C)C	Amine	552.09	510.00	419.34	499.78	492.83
69	CCCCN1CCCC1	Amine	427.65	418.25	453.12	420.63	411.66
70	CCCCN(CCCC)CCN	Amine	513.37	481.27	456.56	473.03	482.27
71	CC(C)CCN	Amine	369.15	383.22	349.94	367.47	371.55
72	CNCCCCCCCCNC	Amine	522.15	508.25	472.10	507.93	535.88
73	CCCCNCCN	Amine	431.15	459.95	423.80	471.70	465.47
74	CC(C)N(CCN)C(C)C	Amine	394.15	449.78	416.88	389.50	409.41
75	C1CCN2CCCN(CC1)CC2	Amine	493.18	487.98	495.94	490.25	485.18
76	C1CCN(CC1)C#N	Amine	471.35	433.22	435.93	448.92	432.18
77	C1CCC2C(C1)CCCC2N	Amine	500.75	512.35	449.41	529.09	505.15
78	C1CC2CC1C3C2C(C(C3)CN)CN	Amine	629.70	582.40	517.64	516.98	544.72
79	C1CN2CCC1N3CCC2CC3	Amine	482.44	514.89	462.92	484.54	473.86
80	CC1(CCCCC1(C)N)C	Amine	493.41	459.86	415.39	444.09	450.51
81	C1CN2CCC1CC2C#N	Amine	514.91	479.32	453.32	513.83	506.30
82	CCNCCC(C)NCC	Amine	453.65	470.18	433.63	493.48	482.06
83	CC1(CN1C)C	Amine	329.15	355.97	329.51	334.39	337.88
84	CC1(C2CCC(C1C2)CN)C	Amine	369.65	481.04	448.80	479.31	487.49
85	CCCN(C)C1CC1	Amine	378.94	396.38	416.95	395.33	391.15
86	CC1CC(C(C1N)CN)(C)C	Amine	494.15	519.29	420.34	499.78	492.83
87	C1CCN2CCCN(C1)C2	Amine	438.88	435.83	453.46	416.41	419.87
88	CC1(CC2(CC(C1)(CNC2)N)C)C	Amine	594.70	537.34	483.02	503.71	522.00
89	CCCCNC(C)(C)C	Amine	389.15	408.85	389.37	415.16	412.80
90	CC1(CC(CC(C1)(C)CN)N)C	Amine	520.15	533.76	465.60	532.47	532.16
91	C(CN)CNCCN	Amine	494.15	501.85	482.22	497.53	484.94
92	C1CCCN2CCCN(CC1)CCC2	Amine	547.48	530.96	507.78	530.40	522.71
93	CC(C)N1CCN(CC1)C(C)C	Amine	474.65	463.23	467.41	461.51	460.88
94	CCCCCCC(C#N)N	Amine	556.81	518.13	460.82	541.58	528.43
95	CC(C)N(C#N)C(C)C	Amine	366.65	426.64	371.58	437.29	428.35
96	C1CNCCNCCCNCCNC1	Amine	680.98	647.59	581.42	637.32	666.42
97	CCN1CCCCC1	Amine	403.95	395.38	435.13	403.53	393.59
98	CN(C)C(N(C)C)(N(C)C)N(C)C	Amine	452.05	415.66	414.40	462.75	445.71
99	C1CCCN(CC1)CC#N	Amine	521.38	498.58	481.08	517.33	505.84
100	C1CNCCCNCCCNC1	Amine	601.01	585.25	537.72	571.70	584.38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.8 ค่าจุดเดือดของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CN(C)C1=CC=NC=C1	Aromatic	435.15	486.13	422.99	469.67	473.65
2	CC1=CC=NC2=CC=CC=C12	Aromatic	535.15	522.78	505.16	516.48	527.35
3	CC1=CC(=C(C2=CC=CC=C12)C)O	Aromatic	588.65	586.75	535.15	524.40	538.36
4	CC1=C2C(=CC=C1)C=CO2	Aromatic	463.65	473.95	518.18	487.79	488.60
5	CCC1=C(C(=CC=C1)CC)C	Aromatic	481.95	519.11	491.23	487.54	481.48
6	CCNC1=C(C=CC(=C1)C)C	Aromatic	495.65	500.80	473.55	489.49	496.62
7	CC(CO)C1=CC=CC=C1	Aromatic	497.65	463.51	481.75	485.26	480.15
8	CC1=CC=C(C=C1)CCC=O	Aromatic	496.15	498.59	468.53	488.37	489.13
9	CC1=CC(=CN1)C	Aromatic	441.15	435.61	449.04	442.30	464.56
10	C1=CC=C(C=C1)C(C(CO)O)O	Aromatic	458.15	578.74	485.77	501.13	507.31
11	C1CC2=C(C1)C=C(C=C2)N	Aromatic	521.15	521.46	515.10	498.61	507.25
12	CC(C(C)(C)C1=CC=CC=C1)O	Aromatic	470.15	508.82	481.34	483.42	478.05
13	C1CCC2=NC=CN=C2C1	Aromatic	382.65	417.50	448.08	450.02	431.36
14	CCC1=CC(=CC(=C1)O)CC	Aromatic	521.15	489.82	478.39	491.59	493.64
15	CCCC1=C(C=CC=C1C)C	Aromatic	480.75	496.69	485.87	489.45	486.31
16	CCC1=CC=CC=C1N	Aromatic	482.80	482.76	506.44	496.68	495.27
17	CC(=O)N1C=NN=N1	Aromatic	509.65	409.49	433.15	434.04	420.97
18	CCN(CC)C1=CC=C(C=C1)C	Aromatic	502.15	515.91	501.61	493.74	504.05
19	CC1=C(C=C(C=C1)C=O)C	Aromatic	499.65	466.29	491.12	483.39	480.58
20	CC1=CC=CNC1=O	Aromatic	562.15	492.84	476.67	522.99	547.22
21	C1=CC=C(C(=C1)C=NCCO)O	Aromatic	644.87	583.75	477.52	516.10	526.79
22	CC1=CC=NN1C	Aromatic	423.15	395.97	418.98	434.51	429.89
23	CC1=C(C(=CC=C1)CO)C	Aromatic	398.15	496.18	485.22	487.12	491.41
24	CC1=CC(=C2C=CC=NC2=C1)C	Aromatic	547.15	531.38	533.43	534.40	530.24
25	C1=CC(=C(C=C1N)N)N	Aromatic	613.15	546.25	534.80	545.65	550.44
26	CC1=CC2=CC=CC=C2C=C1	Aromatic	514.25	526.20	527.15	523.69	542.71
27	CCCCN1C=CN=C1	Aromatic	383.15	474.65	467.91	490.74	482.99
28	CNC1=NC(=NC(=N1)NC)NC	Aromatic	375.65	443.20	442.33	462.59	521.02
29	CCC(C1=CC=CC=C1)C(=O)NC(=O)N	Aromatic	389.95	621.39	488.75	488.16	489.60
30	CC1=NC=CN=C1C	Aromatic	429.15	441.84	416.94	422.06	426.45
31	C1(=O)NC(=O)NN1	Aromatic	546.60	561.85	490.84	560.96	559.95
32	CC1CC(NC2=CC=CC=C12)(C)C	Aromatic	532.15	446.32	471.93	485.79	472.54

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.8 ค่าจุดเดือดของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	CC1=CC(=C(C(=C1)C)CO)C	Aromatic	413.15	515.27	505.92	489.28	500.92
34	C1=CC=C2C=C(C=CC2=C1)O	Aromatic	558.15	524.19	538.87	538.03	533.65
35	CCC(C)(CC)C1=CC=CC=C1	Aromatic	478.15	478.02	483.39	498.01	483.28
36	CNNC1=CC=CC=C1	Aromatic	504.15	440.07	495.21	420.66	446.13
37	CC1=CC(=C(C=C1O)C)C(C)(C)C	Aromatic	537.15	517.00	513.40	498.95	506.51
38	CCC1=NC2=CC=CC=C2C(=C1)C	Aromatic	545.65	515.91	527.42	540.58	535.80
39	CCN1C(=O)NNC1=O	Aromatic	551.98	497.37	447.65	571.53	547.94
40	C1CC1NC2=NC(=NC(=N2)N)N	Aromatic	361.45	504.02	449.62	471.59	538.04
41	CC1=CC(=C(C(=C1O)C)C)O	Aromatic	568.15	551.51	518.67	548.90	550.11
42	CC(C)(C)C1=CC=C(C=C1)C#N	Aromatic	531.15	471.66	463.85	479.77	477.65
43	CCCN1C=CC=C1	Aromatic	418.65	424.86	426.60	460.40	426.87
44	CN1C=CC=C1C2=CN=CC=C2	Aromatic	554.15	514.52	507.59	504.94	504.91
45	CC1=CN=CN1	Aromatic	536.15	441.57	482.18	490.62	477.68
46	CCCCCCC1=CC=CC=C1O	Aromatic	408.65	524.01	507.66	498.38	498.08
47	CC1=C2C(=CC=C1)C=C(C=N2)C	Aromatic	542.15	511.51	528.30	538.80	527.54
48	CCCC1=CC=CC=C1N	Aromatic	499.15	493.09	501.33	496.51	500.56
49	CC1=CC=C(O1)C=O	Aromatic	460.15	404.88	409.15	390.98	405.88
50	CC1=CC(=CC=C1)CC(=O)O	Aromatic	394.65	508.83	503.82	528.50	513.47
51	CON1C=CC(=O)C=C1	Aromatic	457.83	484.55	437.80	477.67	463.85
52	CCC1=CN=C(C=C1)C	Aromatic	451.45	461.40	431.29	461.09	456.51
53	CC1=CN=CC2=CC=CC=C12	Aromatic	529.15	492.99	509.76	519.58	526.03
54	CC(=O)C(=O)C1=CC=CC=C1	Aromatic	495.15	478.51	495.43	488.59	491.78
55	CC(=O)CCCCC1=CC=CC=C1	Aromatic	541.65	520.10	493.51	491.64	492.94
56	C1=CC=C(C=C1)CNCCN	Aromatic	523.15	496.87	467.24	485.08	502.36
57	CC1=NC2=CC=CC=C2C=C1N	Aromatic	551.15	527.88	513.36	548.21	534.39
58	C1=CC2=C(C3=C(C=C2)N=CC=C3)N=C1	Aromatic	633.15	542.90	501.20	519.28	517.24
59	CCC1=C(NC(=C1C)C(=O)C)C	Aromatic	529.15	517.83	472.56	442.08	450.77
60	C1=CC=C2C(=C1)N=CN=N2	Aromatic	510.65	445.53	497.91	466.67	456.71
61	C1=CN=CC=C1C2=CC=NC=C2	Aromatic	578.15	546.13	517.26	523.78	524.06
62	CCC1=CC=C(C=C1)N	Aromatic	490.65	476.47	511.97	499.63	497.16
63	C1=CC=C(C=C1)C(=O)C2=CC=CN2	Aromatic	578.00	573.95	546.36	569.71	559.10
64	CN(C)/N=N/C1=C(NC=N1)C(=O)N	Aromatic	393.55	534.86	421.31	469.50	496.18
65	CNC(=O)NC1=NC2=CC=CC=C2N1	Aromatic	361.55	564.98	523.32	483.38	478.83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.8 ค่าจุดเดือดของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
66	CCNC1=CC=CC(=C1)C	Aromatic	494.15	496.28	479.53	491.03	495.81
67	CCC1=CC=CC2=C1NC=C2	Aromatic	503.15	529.17	526.08	522.58	525.45
68	CC(C)C1=CC(=CC=C1)O	Aromatic	501.15	489.38	483.85	490.10	491.70
69	CC(=O)NC1=CN=CC=C1	Aromatic	599.65	487.31	413.98	505.08	496.91
70	CC(C)(C)C1=CC=CC=N1	Aromatic	443.15	448.70	429.76	457.49	456.64
71	CCCCC(=O)C1=CC=NC=C1	Aromatic	376.65	486.96	467.15	478.10	468.34
72	CN1CCN(CC1)C2=CC=CC=C2	Aromatic	414.15	507.15	505.53	483.55	499.51
73	CC(C)C1=CN=CC=C1	Aromatic	450.65	436.44	465.46	470.18	456.05
74	CCC(=O)C1=CC=CN1	Aromatic	501.65	476.96	462.08	460.47	460.31
75	CCC1=NC=CN=C1C	Aromatic	330.15	436.88	425.30	438.43	427.46
76	C1=CNC(=C1)C=O	Aromatic	491.15	445.11	469.39	447.12	451.10
77	CCCCN1C=CC=CC1=O	Aromatic	504.05	516.13	462.40	484.39	502.79
78	CCC1CCC2=CC=CC=C12	Aromatic	485.15	477.36	470.69	490.54	488.71
79	C1CC2=C(C1)C(=CC=C2)N	Aromatic	508.15	507.98	513.61	498.49	507.25
80	CC1=CC=C(C=C1)N=C=O	Aromatic	344.15	463.92	492.99	484.43	479.13
81	CC1CC2=CC=CC=C2N1	Aromatic	389.15	427.58	490.84	476.26	490.85
82	CC1=C2C=CNC2=CC=C1	Aromatic	540.15	508.34	521.76	507.26	520.23
83	CCNC1=CC=CC2=CC=CC=C21	Aromatic	578.15	551.75	542.79	550.03	543.85
84	CN1C=CC=NC1=O	Aromatic	466.23	452.09	410.24	427.27	425.77
85	CCCCN1C=CC=N1	Aromatic	439.90	453.11	434.77	454.94	449.80
86	CC(=O)N1C=NC=N1	Aromatic	478.83	380.53	431.15	427.21	419.56
87	CCCCC(=O)C1=CC=CO1	Aromatic	381.65	430.97	443.29	441.12	431.91
88	C1=CC=C(C(=C1)CCO)O	Aromatic	441.15	498.69	497.66	507.38	516.93
89	CCC1=C2C=CC=CC2=NN1	Aromatic	563.15	463.31	515.63	522.55	519.89
90	CCC1=C(NC=C1C)C	Aromatic	472.15	479.73	460.43	450.85	467.84
91	CCCC(C)(C1=CC=CC=C1)O	Aromatic	489.15	489.16	473.88	480.60	484.16
92	CC(=NO)C1=CC=C(C=C1)O	Aromatic	621.87	526.23	512.75	538.91	536.19
93	C1=CC=C(C(=C1)N)N	Aromatic	530.15	494.96	510.33	542.87	522.28
94	CC1=CC=CC=C1C(=O)CC(C)C	Aromatic	521.15	490.05	488.22	498.78	495.86
95	CC1=CC(=C(C=C1C)O)C(C)(C)C	Aromatic	531.65	495.99	505.54	498.98	506.51
96	CC1=C(C(=CC=C1)C(C)(C)C)N	Aromatic	510.65	502.26	482.14	497.83	491.96
97	CC1=NN(C=C1N)C2=CC=CC=C2	Aromatic	587.65	556.02	511.01	525.88	505.69
98	CC(C)(C)C1=C(C=CC(=C1)O)O	Aromatic	546.15	505.00	515.33	555.54	542.54
99	C1CCC(C1)(CC2=CC=CC=C2)O	Aromatic	415.15	553.65	488.33	498.90	490.89

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.8 ค่าจุดเดือดของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
100	CCC1=CNC(=C1CC)C	Aromatic	475.65	490.16	458.17	444.60	467.72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.9 ค่าจุดเดือดของหมู่ฟังก์ชัน Carboxylic Acidของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C)COC(=O)CCC(=O)O	Carboxylic Acid	585.93	529.04	500.86	571.39	562.62
2	CC(C(=O)O)NC(=O)CN	Carboxylic Acid	635.64	634.22	561.19	639.25	626.15
3	C(CC(=O)O)C(C(=O)O)N	Carboxylic Acid	478.15	634.60	589.10	623.65	616.82
4	COC(=O)C12C3C4C1C5C2C3C45C(=O)O	Carboxylic Acid	645.29	639.54	515.40	611.06	597.19
5	CCCC(C)C(C(=O)O)N	Carboxylic Acid	576.92	576.67	506.38	572.57	574.16
6	CCCCCC(C(C)CC(=O)O)OC(=O)C	Carboxylic Acid	677.01	636.66	489.84	635.32	619.30
7	CC(C)CC(C(=O)O)NC=O	Carboxylic Acid	603.22	601.87	535.79	640.73	607.70
8	C(CO)C(C(=O)O)N	Carboxylic Acid	600.90	591.85	560.00	604.90	597.97
9	CCCCC(C(=O)O)N	Carboxylic Acid	554.48	558.52	537.08	562.33	563.33
10	C(C(=O)CC(=O)O)C(=O)O	Carboxylic Acid	573.15	622.96	612.04	622.61	630.43
11	CC(=O)NCC(=O)O	Carboxylic Acid	540.67	549.07	526.24	589.44	588.30
12	C1=CC(=CC=C1/C=C/C(=O)O)N	Carboxylic Acid	558.15	581.63	519.48	521.41	537.28
13	C1CC1(C(=O)O)N	Carboxylic Acid	516.14	501.25	517.11	532.43	541.84
14	CC(CC(=O)O)OC	Carboxylic Acid	481.49	487.11	439.84	465.44	486.67
15	CCC(C(=O)O)N	Carboxylic Acid	508.72	528.63	532.47	540.36	540.43
16	C1CCC(CC1)NCC(=O)O	Carboxylic Acid	597.87	578.12	481.00	580.71	582.22
17	CCCNCC(=O)O	Carboxylic Acid	509.68	505.27	488.28	526.22	530.95
18	CCOCCC(=O)O	Carboxylic Acid	481.93	447.99	448.15	469.63	484.46
19	COC1=CC(=C(C=C1C(=O)O)OC)OC	Carboxylic Acid	573.15	536.37	503.07	490.20	523.79
20	CCCCCCC(C(=O)O)N	Carboxylic Acid	600.24	598.58	529.07	598.88	594.31
21	CC1(C2CCC1(C(=O)C2C(=O)O)C)C	Carboxylic Acid	673.50	636.75	528.74	609.86	602.18
22	C1CCC(CC1)OC(=O)/C=C/C(=O)O	Carboxylic Acid	655.84	635.23	482.31	548.87	548.35
23	CCCOC(=O)/C=C\C(=O)O	Carboxylic Acid	567.65	538.50	498.61	550.43	543.74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.9 ค่าจุดเดือดของหมู่ฟังก์ชัน Carboxylic Acid ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
24	CN(C)NC(=O)CCC(=O)O	Carboxylic Acid	598.87	583.30	535.15	599.12	617.30
25	C(CO)N(CCO)CC(=O)O	Carboxylic Acid	679.19	619.66	519.69	608.51	613.82
26	CC1CCC(C(C1)OCC(=O)O)C(C)C	Carboxylic Acid	651.86	620.57	481.84	626.72	605.55
27	CC(C(=O)O)NC(=O)CCN	Carboxylic Acid	658.52	652.64	590.93	646.46	627.67
28	C(CCN)CC(C(=O)O)N	Carboxylic Acid	627.01	613.40	532.50	574.23	583.52
29	CC(C)CCCCCOC(=O)/C=C\C(=O)O	Carboxylic Acid	681.61	629.63	518.88	657.37	622.42
30	CC(COC(=O)/C=C\C(=O)O)OC(=O)/C=C\C(=O)O	Carboxylic Acid	486.37	771.72	547.81	516.58	591.14
31	CC(C)(C)CC(C(=O)O)N	Carboxylic Acid	574.13	567.78	503.08	573.42	572.48
32	CC(CC(=O)O)C(=O)C	Carboxylic Acid	514.65	528.09	489.55	530.53	533.07
33	C(C(=O)O)NCC(=O)O	Carboxylic Acid	632.31	604.01	558.78	605.71	606.40
34	CC(C)(C)OC(=O)NCCCC(=O)O	Carboxylic Acid	656.19	610.74	552.31	613.89	637.03
35	C(C(=O)O)N	Carboxylic Acid	463.40	462.78	427.99	495.20	503.26
36	C1CC2CC1C(C2N)C(=O)O	Carboxylic Acid	609.09	577.91	495.32	594.59	590.96
37	CC(C(=O)O)NC(=O)C	Carboxylic Acid	563.11	574.91	534.36	596.23	606.15
38	C1CC(=O)NC1C(=O)O	Carboxylic Acid	591.16	603.59	534.45	619.55	612.04
39	C1=CC=C(C=C1)C(=O)NOCC(=O)O	Carboxylic Acid	417.65	626.88	522.21	504.64	518.74
40	C(C(CO)N)C(=O)O	Carboxylic Acid	600.90	571.81	525.06	593.23	597.30
41	CC(C)CC(C(=O)O)N	Carboxylic Acid	554.04	570.00	547.97	560.96	560.47
42	C1CNCCC1C(=O)O	Carboxylic Acid	550.49	519.45	483.83	551.13	545.12
43	CNC(=O)C12C3C1C4C2(C4C3)C(=O)O	Carboxylic Acid	674.70	645.36	485.75	618.21	563.29
44	CN(CCN(C)CC(=O)O)CC(=O)O	Carboxylic Acid	573.15	618.70	521.02	652.08	623.70
45	C(CNCCNC(=O)O)N	Carboxylic Acid	632.38	586.04	487.08	562.94	586.66
46	CCCC(C(=O)O)N	Carboxylic Acid	531.60	538.49	539.74	552.83	551.97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.10 ค่าจุดเดือดของหมู่ฟังก์ชัน Ester ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CCOC1=CC=C(C=C1)C(=O)OCC	ester	548.15	507.53	488.73	501.43	502.28
2	CCCCC(CC)COC(=O)CCC	ester	488.15	518.13	505.67	523.64	521.01
3	CCCC(=O)OCCC(C)C	Ester	452.15	459.99	452.14	461.31	456.51
4	C1C(O1)COCCCCO	Ester	508.05	506.93	443.36	501.85	490.13
5	C=C1CC(=O)O1	Ester	399.25	364.47	380.87	353.51	360.49
6	COC1CCCCC1	Ester	406.15	404.02	416.93	408.90	407.27
7	CCCCC(C)CCC(=O)OC	Ester	486.18	485.91	471.00	463.25	477.22
8	CC(=O)CCC1=CC(=C(C=C1)O)OC	Ester	460.65	538.88	515.59	493.82	484.87
9	C=CCOCC1=CC=CC=C1	Ester	477.65	421.41	468.43	458.16	436.53
10	CC(C)CCC1CCOC1=O	Ester	470.04	462.11	450.32	464.01	459.11
11	COC(=O)N1C=CC2CCC1C2	Ester	493.09	491.20	478.23	488.61	481.95
12	COC(=O)/N=N/C(=O)OC	Ester	556.76	476.85	466.15	450.40	464.25
13	CCCOCC(COOC(C)(C)C)O	Ester	584.17	570.49	543.89	564.36	580.80
14	CN(C)C1=CC=CC=C1OC	Ester	483.15	454.14	478.23	486.27	482.67
15	CCOC(CN(C)C)OCC	Ester	443.65	436.04	424.76	437.82	447.35
16	CCCCOC(=O)CCC(=O)OCCCC	Ester	547.65	568.06	515.42	530.32	552.88
17	CCC(=O)OC	Ester	352.95	361.91	394.59	372.07	381.06
18	C=CCOCCO	Ester	431.65	417.28	416.51	440.01	435.90
19	CCCCOCC(COC)O	Ester	481.15	524.96	488.81	537.27	561.31
20	CCOC(=O)CC(C)CC(=O)OCC	Ester	544.40	532.93	534.91	519.76	533.47
21	C=CC1OCC2(CO1)COC(OC2)C=C	Ester	382.15	548.48	472.65	539.19	532.96
22	COC(=O)C1=CCCN(C1)N=O	Ester	517.91	485.86	492.01	497.31	479.66
23	CCC1CC(=C(O1)C)C(=O)OCC	Ester	537.97	508.68	471.24	483.65	500.91
24	CCCC(=O)OCC(COC(=O)CCC)O	Ester	659.46	641.23	553.79	572.30	602.05
25	C1COCCN1	Ester	401.15	419.93	376.69	379.79	393.81
26	C1CCOC(C1)CN	Ester	455.91	442.12	412.14	432.24	434.68
27	CC1=CC=CC=C1OC	Ester	444.15	449.55	471.84	487.70	484.04
28	CCC(C)(C)ON=O	Ester	396.59	364.43	375.91	369.16	364.85
29	CCC(C)CC(=O)OC	Ester	410.80	432.33	422.14	418.38	425.21
30	CCOC1OC(C(O1)(C)C)(C)C	Ester	488.26	479.00	425.38	465.15	474.94
31	CCC(=O)OC(C)C1=CC=CC=C1	Ester	364.65	481.59	499.70	490.12	484.59
32	CNCC1CCCO1	Ester	428.65	421.16	406.67	430.30	424.43

ตารางที่ ข.10 ค่าจุดเดือดของหมู่ฟังก์ชัน Ester ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน		Algorithm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

			แหล่งอ้างอิง	Ridge	KNN	RF	XGB
33	CC1CC2C(O2)CC1C(=O)OCC=C	Ester	541.54	499.53	494.23	527.62	525.00
34	C1CCOC(C1)O	Ester	452.68	457.20	401.50	446.87	442.87
35	C(CCOCCC#N)COCCC#N	Ester	677.40	657.15	530.62	603.30	588.04
36	CCOC(=O)C1=CC(=CC=C1)N=C=O	Ester	486.15	503.67	528.87	515.60	505.30
37	COC(OC)(OC)OCC=C	Ester	442.89	428.87	397.85	479.09	451.78
38	CC(C1=CC=CC=C1OC)O	Ester	398.15	474.61	484.24	487.15	480.47
39	CCCCCCCCC(=O)OCC=C	Ester	360.15	509.54	513.51	528.42	517.96
40	C1C2(COC(O1)OC2)CO	Ester	532.17	512.06	455.17	518.16	506.86
41	CCCCCC(=O)CC(=O)OC	Ester	517.61	512.78	497.80	513.92	521.53
42	CC/C=C/CC(=O)OCC	Ester	439.65	432.00	440.22	443.21	444.99
43	CCCCOC(=O)/C=C/CC	Ester	467.90	462.48	464.46	462.31	459.84
44	COC(=O)CCC(C(=O)OC)C(=O)OC	Ester	579.74	554.77	519.95	570.87	572.11
45	CC1CCC(C(=C1)OC(=O)C)C(C)C	Ester	550.96	515.61	488.49	535.68	520.84
46	CNC(=O)OC1CC2CCC(C1)N2C	Ester	562.31	536.03	477.55	550.52	554.98
47	C#CCOC(=O)CCCCC(=O)OCC#C	Ester	570.84	562.00	538.65	578.61	554.14
48	CC(=O)NC1C=CC(=O)O1	Ester	505.04	531.16	442.66	496.36	516.27
49	C/C=C/C1=CC(=C2C(=C1OC)OCO2)OC	Ester	576.65	538.02	526.15	524.59	528.75
50	CCCCOC(=O)CCCC(=O)C	Ester	540.49	522.13	486.87	551.15	542.86
51	CC(C)OC1=CCCC=C1	Ester	455.02	459.29	431.06	442.49	445.05
52	CC(=O)CC(=O)NC1=CC(=CC=C1)OC	Ester	411.15	568.71	508.87	512.21	522.91
53	CCCN(CCC)C(=O)CCC(=O)OCC	Ester	598.69	584.98	535.27	579.29	563.40
54	CC(CN1CCOCC1)O	Ester	510.84	485.18	463.80	448.75	462.27
55	CC1=CC(=CC=C1)OC(=O)N(C)N=O	Ester	408.35	515.02	490.99	466.59	450.56
56	CCCCCOCC(COOC(C)(C)C)O	Ester	629.93	604.97	549.57	589.87	591.88
57	CCCCCCCC1CCCC(=O)O1	Ester	399.15	535.87	545.53	529.32	521.80
58	COCCNC(=O)NCCOC	Ester	558.81	591.80	458.74	583.15	589.54
59	CC(C)OC(C)OC(C)C	Ester	399.65	417.96	405.82	437.09	431.74
60	CCCCOC(=O)C1=CC=CC(=C1)C	Ester	535.45	532.96	513.79	500.84	514.52
61	CCOC(=O)CC1CCCC1=O	Ester	546.84	527.72	502.95	486.70	520.65
62	COCCOCCOC1OCCO1	Ester	369.15	501.84	498.73	446.01	504.15
63	COC1C=CC(O1)(C(=O)OC)OC	Ester	392.65	535.74	471.18	486.37	492.88
64	COCCC(=O)OC	Ester	415.95	417.70	415.29	413.69	415.94
65	CCCC(OCC)OCC	Ester	416.15	419.65	438.09	420.78	432.64

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

66	CC(=O)OC	Ester	329.85	343.44	357.31	366.30	377.25
67	CCOC(=O)NCCC(C)C	Ester	491.15	491.96	458.56	490.98	489.53
68	C1COCN1CCO	Ester	461.25	460.04	446.14	434.00	447.14
69	CC1=NN(C(=C1)OC(=O)N(C)C)C(C)C	Ester	390.90	536.48	432.63	438.46	434.74
70	CCOCCOCCOCCOCC	Ester	518.08	497.12	564.05	502.92	512.47
71	CCCCOC(=O)C=O	Ester	443.76	460.11	443.18	443.96	454.05
72	CCC(C)(C)OC(=O)C(C)(C)C	Ester	480.16	466.16	478.00	486.17	475.30
73	CCCCC(=O)OCC1=CC=CO1	Ester	501.15	477.11	460.45	467.00	437.73
74	CCOC(=O)CCCCCC#N	Ester	565.82	545.99	485.81	534.48	542.51
75	CCCCCCCC(=O)OC	Ester	466.05	470.04	466.16	463.39	460.66
76	COC(=O)N1C2CCC1CC2	Ester	466.78	478.85	471.06	467.82	462.03
77	C=C1C2C(=C)C(=O)C(C1=C)O2	Ester	515.52	486.02	427.83	490.95	492.99
78	CCCCCC(C(=O)C)C(=O)OCC	Ester	562.93	534.68	504.89	583.20	560.70
79	CC(C(=O)CC(=O)OC)(OC)OC	Ester	536.34	538.60	494.24	508.42	511.54
80	CN(C)CCOC(=O)CCC(=O)OCCN(C)C	Ester	615.48	601.43	518.72	586.53	571.95
81	C1CCC2C=CC(C1)O2	Ester	435.04	436.49	423.44	425.53	433.71
82	CC1C(=O)C(=C(O1)C)OC(=O)C	Ester	516.15	518.73	480.60	479.09	489.11
83	COCN=C=O	Ester	360.65	380.77	370.59	413.13	413.46
84	CCC(=O)C(C)CC(C)C1CO1	Ester	515.08	493.02	478.15	515.94	490.92
85	CC(=O)C1=CC2=C(C=CC=C2OC)OC1	Ester	401.15	540.70	502.28	492.47	519.21
86	CCOC(=O)C(C)C	Ester	383.25	386.46	384.66	401.15	410.83
87	CC(C)(/N=N/C(C)(C)OC)OC	Ester	570.22	476.18	407.89	479.66	461.13
88	CCO/C=C/C1=CC=CC=C1	Ester	497.65	457.91	496.26	485.97	473.21
89	COC1=NC=C(C(=N1)N)CO	Ester	502.15	503.65	423.22	464.62	457.19
90	C/C=C\1/C=CC2C13C(O3)CCN2	Ester	534.43	495.35	463.12	564.83	565.74
91	CC(=O)CCOC	Ester	390.29	404.92	377.41	413.32	409.39
92	CCC(=O)OC=C	Ester	364.35	366.74	369.82	371.12	375.04
93	CCCO/C=C/C	Ester	363.46	384.78	357.30	371.09	375.51
94	C1CCC(CC1)C2CCOC(=O)C2	Ester	540.06	534.34	476.47	516.37	517.76
95	CCOC(=NOC)C	Ester	435.40	411.29	388.38	410.15	418.86
96	CCC(C)(OO)OOC(C)(CC)OO	Ester	292.15	604.68	453.14	581.49	577.91
97	CCC(=O)OC(C)C(C)CC=C(C)C	Ester	535.54	498.47	497.60	519.54	513.30
98	CC(C)(COC(=O)C=C)COC(=O)C=C	Ester	369.15	516.41	531.46	539.18	547.54
99	CCOC(=O)CC(C)N	Ester	467.19	449.73	418.15	467.91	453.14
100	C1COCCN1CC(CO)O	Ester	603.02	548.13	479.96	511.97	566.74

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.11 ค่าจุดเดือดของหมู่ฟังก์ชัน Ketone ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	CC(C)CCN(C(C)C(=O)C)N=O	Ketone	534.35	488.06	490.41	512.57	509.64
2	CC1(C(=O)C(=O)C(=O)C1(C)C)C	Ketone	620.07	569.63	536.18	576.81	573.35
3	CC(=O)CN(CC=C)CC=C	Ketone	465.19	432.88	430.86	473.30	450.49
4	CC1CCCC1C(=O)C	Ketone	443.65	450.36	439.75	432.28	441.70
5	CCC(C)C1CCCCC1=O	Ketone	515.33	496.20	465.04	497.51	480.07
6	CC1=CC(=C(C=C1)C(=O)C)O	Ketone	518.15	511.04	522.14	496.49	513.74
7	CCCCCC(=O)CC(=O)C	Ketone	513.26	505.07	455.59	485.24	514.98
8	CC1CC1C(=O)C	Ketone	392.82	403.31	386.24	406.78	404.29
9	C1CC2C3C4CCC(C3C1C2=O)C4=O	Ketone	627.42	590.95	556.59	618.49	625.67
10	C1C2CC3CN(C2)CC1C3=O	Ketone	506.00	490.35	495.96	484.24	488.02
11	C1C2CC(=O)C1CC2=O	Ketone	513.15	500.89	473.44	499.30	495.97
12	CCCC(C(=O)C)C(=O)C	Ketone	489.94	480.31	475.15	481.06	478.83
13	C1CC(=O)CC1=O	Ketone	469.59	447.80	433.37	433.59	445.55
14	C1CN2CCC1C(=O)C2	Ketone	462.44	442.53	444.08	430.18	450.17
15	CCCC(C)C(=O)C	Ketone	416.65	416.67	421.52	413.41	415.95
16	C1CCC2C(C1)CCCC2=O	Ketone	526.78	511.49	478.07	428.97	500.46
17	CCCC(=O)CC(C)C	Ketone	428.15	438.62	438.19	434.95	437.54
18	C1CC2CC1C(=O)C2=O	Ketone	513.15	511.44	478.12	492.94	488.43
19	CC1CCCC(C1=O)C	Ketone	448.15	449.34	445.26	452.71	447.62
20	CC(=O)C1=CC=CC=C1N	Ketone	360.65	476.74	510.96	502.27	501.49
21	C1CCC(=O)C(=O)C1	Ketone	467.15	437.98	434.90	444.50	443.05
22	CN1CCC(=O)CC1	Ketone	437.09	419.38	408.90	426.39	429.57
23	C1CCC(=O)CCCCC(=O)C1	Ketone	605.34	542.88	513.03	563.58	580.46
24	CCCN1CCC(=O)CC1	Ketone	482.85	460.96	459.78	465.06	463.30
25	C1CCC(=O)CCC(=O)C1	Ketone	551.04	502.82	551.04	534.36	524.44
26	CC12CC1(C(=O)CC2=O)C	Ketone	532.24	520.72	488.44	523.01	499.64
27	CC(=O)CN(C)N=O	Ketone	420.83	389.11	428.01	419.12	417.29
28	CC(=O)C#N	Ketone	365.45	390.07	379.08	409.14	432.82
29	CC1(CC(C(=O)C(=O)C1=O)(C)C)C	Ketone	647.22	598.71	554.16	603.99	612.81
30	CC1(CCC(=O)CC1)C	Ketone	470.25	454.27	447.13	454.18	449.83
31	CC(C)(C)CC(=O)C(C)(C)C	Ketone	434.03	484.19	466.72	486.84	475.20
32	CC(=O)C1=CC(=CC=C1)N	Ketone	562.65	494.18	511.74	503.69	502.91

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.11 ค่าจุดเดือดของหมู่ฟังก์ชัน Ketone ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
33	C12C3C4C1C5C6C4C(=O)C3C6C2C5=O	Ketone	605.94	599.22	474.68	612.98	621.53
34	CCNCCCC(=O)C	Ketone	463.80	450.08	425.94	446.31	450.49
35	CC(C)CC(=O)C(C)C	Ketone	420.65	430.15	425.20	432.73	437.05
36	C1C2C1C(=O)C(=O)C2=O	Ketone	553.82	549.40	490.46	507.29	527.51
37	CC(=O)C(C)(C)N=O	Ketone	428.04	429.52	391.17	443.62	423.43
38	CCCCCCC(=O)C	Ketone	445.65	435.90	455.87	450.20	446.63
39	CCCCCCCCC(=O)C(=O)C	Ketone	383.15	528.43	500.24	597.41	571.65
40	CC1CC(C(=O)C1)C	Ketone	425.65	434.83	417.33	420.19	419.70
41	C1CCC(=O)CC1	Ketone	428.55	418.68	444.29	425.24	425.88
42	CCC(C)(C)C(=O)C	Ketone	404.65	416.26	395.77	406.29	415.46
43	CCCCCC(=O)CCCC	Ketone	484.15	479.73	479.65	472.93	478.11
44	CC1CC(=O)C1(C)C	Ketone	434.16	428.25	410.37	407.02	415.65
45	CC1(CC(C(=O)CC1=O)(C)C)C	Ketone	579.40	554.79	522.40	574.05	559.82
46	CC(=O)C(=O)C	Ketone	361.15	405.94	388.94	398.14	406.84
47	CC1CCCC(=O)C1C	Ketone	351.15	448.13	441.76	434.55	441.70
48	CN(C)N1C2CCCC1CC(=O)C2	Ketone	543.12	491.95	471.51	521.41	527.62
49	CCCC(=O)CCC(=O)C	Ketone	490.38	479.98	445.09	492.57	491.76
50	C1C2CC3CC1C(C2)C3=O	Ketone	488.89	482.64	461.97	476.26	482.03
51	CC1(C2CC(C1=O)C(C2=O)(C)C)C	Ketone	595.81	522.75	534.59	575.50	580.25
52	CC1(CC(C(=O)C1=O)(C)C)C	Ketone	552.25	522.54	513.25	517.42	514.14
53	CCC(=O)C1=CC(=CC=C1)N	Ketone	441.65	500.12	501.13	500.72	502.91
54	CCC(=O)CCCC(C)C	Ketone	455.65	447.50	444.59	453.11	457.31
55	C/C(=C/C(=O)C)/N	Ketone	443.10	455.57	398.42	391.31	405.48
56	CCC1C2CCCCC2CC1=O	Ketone	540.72	524.54	486.46	507.07	518.59
57	CCC(C)CC(=O)C	Ketone	417.65	416.82	423.41	413.10	415.59
58	CC(=O)C1CC1(C)C	Ketone	404.15	429.73	397.39	409.78	415.65
59	C1CCC2(C1)C(=O)C3(C2=O)CCCC3	Ketone	647.98	596.83	545.22	633.97	626.18
60	CCC(C)C(=O)CC	Ketone	407.65	413.69	416.05	412.96	414.19
61	CC1CC(C(=O)C1=O)C	Ketone	506.01	496.10	454.21	478.13	460.61
62	CC1=CC(=C(C=C1)O)C(=O)C(C)C	Ketone	524.15	502.73	481.93	482.05	485.18
63	CCCN(CC(=O)C)N=O	Ketone	466.59	436.17	453.97	447.70	447.65
64	CC(C)CC(=O)C	Ketone	389.65	394.80	403.12	401.76	402.25
65	C1CCN(CC1)CC2CCCCC2=O	Ketone	392.15	562.13	506.03	571.22	558.97

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.11 ค่าจุดเดือดของหมู่ฟังก์ชัน Ketone ของแหล่งอ้างอิงและจากการทำนาย (ต่อ)

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
66	CC1(CC(=CC(=O)C1)N)C	Ketone	546.92	538.15	475.54	487.81	498.17
67	C1CCCC(=O)CCC1	Ketone	469.15	458.74	457.78	476.42	440.27
68	CC(=O)C1=CC=C(C=C1)N	Ketone	567.15	491.63	507.10	506.98	504.80
69	C1CC(=O)C2CCC(=O)C1C2	Ketone	567.45	536.17	494.91	554.24	557.71
70	CC1CC(=O)CC(C1)(C)C	Ketone	462.15	482.74	456.95	464.69	470.30
71	CCC1CCC(C1=O)CC	Ketone	483.95	471.08	446.70	470.73	466.51
72	CC(C)CC(=O)CC(=O)C(C)(C)C	Ketone	555.35	545.26	517.21	565.16	570.08
73	CCCC(=O)CC	Ketone	396.65	390.85	408.93	407.92	402.43
74	CCC1CCCC1=O	Ketone	430.65	439.66	370.15	408.26	426.13
75	CC1(CCC(C1=O)(C)C)C	Ketone	484.43	479.19	464.30	451.42	459.34
76	CC(C)N1CCC(=O)CC1	Ketone	482.41	452.06	421.70	451.11	458.77
77	CCCCCCCCC(=O)C	Ketone	483.15	475.96	487.64	475.97	480.75
78	C1CN1C2=CC(=O)C(=CC2=O)N3CC3	Ketone	635.70	593.13	468.64	594.35	603.62
79	CC1(CC(=O)CC(C1)(C)C)C	Ketone	469.65	513.91	501.39	471.62	494.30
80	CCCCCCCCCC(=O)CC	Ketone	528.03	511.02	495.17	513.00	520.15
81	CC1(CC(=CC(=O)C1)NC(C)(C)C)C	Ketone	612.85	576.88	518.83	590.71	583.58
82	CC1CCC(C(=O)C1)C	Ketone	447.00	457.07	446.12	450.75	447.62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ข.12 ค่าจุดเดือดของหมู่ฟังก์ชัน Ether ของแหล่งอ้างอิงและจากการทำนาย

ลำดับ	SMIELS	หมู่ฟังก์ชัน	แหล่งอ้างอิง	Algorithm			
				Ridge	KNN	RF	XGB
1	C1CC(=O)OC1=O	Ether	534.15	420.47	447.00	452.91	430.39
2	CCCCC(=O)OC(=O)CCCC	Ether	500.15	521.57	473.47	525.13	528.01
3	CC(C)(C)C(=O)OC(=O)C(C)(C)C	Ether	466.15	508.46	474.73	527.52	526.99
4	C1CC(=O)OC(=O)C1	Ether	560.15	440.50	464.15	453.73	443.69

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค.
ความดันไอของสารอินทรีย์ที่ประกอบไปด้วย คาร์บอน ไฮโดรเจน ออกซิเจน และไนโตรเจนจากแหล่งอ้างอิงและการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.1 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Alcohol ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	C(CCO)CO	Alcohol	402	7.58	7.66	7.32	7.77	9.48
2	C(CCO)CO	Alcohol	432.25	8.97	8.06	8.60	9.37	9.48
3	C(CCO)CO	Alcohol	462.5	10.19	9.00	9.40	9.91	9.49
4	C(CCO)CO	Alcohol	492.75	11.25	10.56	10.83	10.73	9.49
5	C(CCO)CO	Alcohol	523	12.20	11.26	11.57	11.81	9.49
6	CCCC(C(C)(C)C)O	Alcohol	323	7.59	7.62	7.56	7.38	10.13
7	CCCC(C(C)(C)C)O	Alcohol	350.25	8.83	7.10	8.19	8.69	10.13
8	CCCC(C(C)(C)C)O	Alcohol	377.5	9.89	9.99	9.60	9.93	10.13
9	CCCC(C(C)(C)C)O	Alcohol	404.75	10.81	10.60	10.58	10.80	10.13
10	CCCC(C(C)(C)C)O	Alcohol	432	11.61	11.62	11.42	11.69	10.13
11	CC(CC(C)(C)O)O	Alcohol	293	0.96	-1.48	3.44	4.53	10.18
12	CC(CC(C)(C)O)O	Alcohol	340.5	5.42	1.33	3.82	5.97	10.18
13	CC(CC(C)(C)O)O	Alcohol	388	8.31	7.85	7.59	7.99	10.18
14	CC(CC(C)(C)O)O	Alcohol	435.5	10.34	9.73	9.52	10.12	10.18
15	CC(CC(C)(C)O)O	Alcohol	483	11.85	10.38	11.06	11.68	10.18
16	CC1=C(C(=CC=C1)C(C)(C)C)O	Alcohol	375	7.16	7.19	6.85	6.81	10.12
17	CC1=C(C(=CC=C1)C(C)(C)C)O	Alcohol	407.25	8.56	8.24	8.13	8.39	10.12
18	CC1=C(C(=CC=C1)C(C)(C)C)O	Alcohol	439.5	9.72	9.45	9.56	9.81	10.12
19	CC1=C(C(=CC=C1)C(C)(C)C)O	Alcohol	471.75	10.69	10.59	10.49	10.74	10.12
20	CC1=C(C(=CC=C1)C(C)(C)C)O	Alcohol	504	11.51	11.53	11.44	11.41	10.12
21	C1=CC=C(C=C1)NCCO	Alcohol	377	4.77	7.15	6.75	3.29	8.57
22	C1=CC=C(C=C1)NCCO	Alcohol	421	7.09	8.69	8.91	5.50	8.58
23	C1=CC=C(C=C1)NCCO	Alcohol	465	8.90	11.53	10.48	6.79	8.58
24	C1=CC=C(C=C1)NCCO	Alcohol	509	10.35	10.22	11.09	7.60	8.58
25	C1=CC=C(C=C1)NCCO	Alcohol	553	11.53	10.22	11.33	7.84	8.58
26	CC(C)(C=C)O	Alcohol	289	7.66	7.53	7.53	6.52	10.00
27	CC(C)(C=C)O	Alcohol	309.75	8.91	7.60	8.13	8.80	10.00
28	CC(C)(C=C)O	Alcohol	330.5	9.97	9.90	9.50	9.73	10.01
29	CC(C)(C=C)O	Alcohol	351.25	10.88	10.21	9.99	10.81	10.01
30	CC(C)(C=C)O	Alcohol	372	11.67	11.42	10.83	11.24	10.02
31	CC(C)CCCCCO	Alcohol	364	7.61	8.92	8.15	7.84	9.51
32	CC(C)CCCCCO	Alcohol	389.75	8.96	8.96	9.42	9.33	9.51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.1 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Alcohol ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CC(C)CCCCCO	Alcohol	415.5	10.05	10.77	10.48	10.33	9.51
34	CC(C)CCCCCO	Alcohol	441.25	10.94	11.17	11.28	11.01	10.18
35	CC(C)CCCCCO	Alcohol	467	11.69	11.17	11.55	11.50	10.18
36	CC(C)C(C(C)C)O	Alcohol	320	7.56	10.27	7.79	7.50	10.13
37	CC(C)C(C(C)C)O	Alcohol	349.25	9.18	7.10	8.93	9.34	10.13
38	CC(C)C(C(C)C)O	Alcohol	378.5	10.42	10.22	10.13	10.30	10.13
39	CC(C)C(C(C)C)O	Alcohol	407.75	11.40	10.95	11.07	11.28	10.13
40	CC(C)C(C(C)C)O	Alcohol	437	12.20	11.56	11.52	12.31	10.13
41	CCCC1(CCCC1)O	Alcohol	344	7.04	9.14	7.24	7.38	9.74
42	CCCC1(CCCC1)O	Alcohol	369.75	8.62	7.39	8.58	8.86	9.74
43	CCCC1(CCCC1)O	Alcohol	395.5	9.82	10.05	9.60	9.45	9.74
44	CCCC1(CCCC1)O	Alcohol	421.25	10.76	10.95	10.67	10.60	9.74
45	CCCC1(CCCC1)O	Alcohol	447	11.52	11.50	11.24	11.20	9.74
46	CC[C@@H](CC1=CC=CC=C1)CO	Alcohol	353	4.09	4.82	5.59	4.84	9.89
47	CC[C@@H](CC1=CC=CC=C1)CO	Alcohol	363	4.79	4.82	5.86	5.71	9.89
48	CC[C@@H](CC1=CC=CC=C1)CO	Alcohol	373	5.44	4.82	5.88	6.19	9.89
49	CC[C@@H](CC1=CC=CC=C1)CO	Alcohol	383	6.05	7.19	6.81	6.57	9.89
50	CC[C@@H](CC1=CC=CC=C1)CO	Alcohol	393	6.62	7.19	7.04	7.67	9.89
51	CCCC(CC)C(C)CO	Alcohol	329	5.99	7.18	6.37	4.62	11.76
52	CCCC(CC)C(C)CO	Alcohol	363.5	7.79	8.02	7.90	7.00	12.32
53	CCCC(CC)C(C)CO	Alcohol	398	9.27	7.53	9.60	8.99	12.95
54	CCCC(CC)C(C)CO	Alcohol	432.5	10.52	10.59	10.95	10.55	13.63
55	CCCC(CC)C(C)CO	Alcohol	467	11.58	11.50	11.57	11.32	13.63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.2 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Aldehyde ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	C/C=C/C=O	Aldehyde	288	8.12	9.11	8.64	8.27	8.59
2	C/C=C/C=O	Aldehyde	310	9.13	4.81	8.74	9.12	8.59
3	C/C=C/C=O	Aldehyde	332	10.03	11.47	9.69	9.66	8.59
4	C/C=C/C=O	Aldehyde	354	10.83	10.38	11.13	11.02	8.59
5	C/C=C/C=O	Aldehyde	376	11.55	12.29	11.80	11.26	8.59
6	C1=COC(=C1)C=O	Aldehyde	366	9.22	7.36	8.37	9.58	10.67
7	C1=COC(=C1)C=O	Aldehyde	383.25	9.89	12.05	10.24	10.78	10.67
8	C1=COC(=C1)C=O	Aldehyde	400.5	10.50	12.05	10.96	11.05	10.67
9	C1=COC(=C1)C=O	Aldehyde	417.75	11.04	11.19	11.12	11.20	10.67
10	C1=COC(=C1)C=O	Aldehyde	435	11.53	9.60	11.10	11.67	10.67
11	CC(C)C=O	Aldehyde	286	11.49	9.11	7.67	6.22	7.59
12	CC(C)C=O	Aldehyde	301.25	12.19	6.63	7.26	6.86	6.88
13	CC(C)C=O	Aldehyde	316.5	12.81	7.38	8.09	7.70	7.56
14	CC(C)C=O	Aldehyde	331.75	13.36	7.38	8.45	8.23	8.23
15	CC(C)C=O	Aldehyde	347	13.86	8.91	9.29	9.13	8.86
16	CCCCCCCCCCC=O	Aldehyde	288	1.25	1.35	0.83	3.31	7.07
17	CCCCCCCCCCC=O	Aldehyde	299.25	2.20	2.75	1.88	4.73	8.00
18	CCCCCCCCCCC=O	Aldehyde	310.5	3.07	2.75	2.83	5.39	8.03
19	CCCCCCCCCCC=O	Aldehyde	321.75	3.89	3.78	3.58	5.27	8.07
20	CCCCCCCCCCC=O	Aldehyde	333	4.65	4.86	4.42	5.21	8.11
21	CC=O	Aldehyde	219	7.59	4.74	5.90	7.92	8.56
22	CC=O	Aldehyde	242.5	9.14	7.61	7.17	8.53	8.57
23	CC=O	Aldehyde	266	10.37	9.19	8.56	9.80	8.59
24	CC=O	Aldehyde	289.5	11.37	9.19	9.51	10.21	8.60
25	CC=O	Aldehyde	313	12.20	10.39	10.19	10.26	8.61
26	C/C(=C\C1=CC=CC=C1)/C=O	Aldehyde	348	4.50	6.03	5.24	4.86	8.59
27	C/C(=C\C1=CC=CC=C1)/C=O	Aldehyde	359.25	5.40	7.16	6.27	5.44	8.59
28	C/C(=C\C1=CC=CC=C1)/C=O	Aldehyde	370.5	6.15	7.15	6.69	5.76	8.59
29	C/C(=C\C1=CC=CC=C1)/C=O	Aldehyde	381.75	6.79	7.17	7.32	6.55	8.59
30	C/C(=C\C1=CC=CC=C1)/C=O	Aldehyde	393	7.34	7.17	7.51	7.18	8.59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.3 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Alkane ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CCC(C)CCC(C)CC	Alkane	326	7.57	7.59	7.58	7.59	10.39
2	CCC(C)CCC(C)CC	Alkane	360	9.14	9.17	9.14	9.25	10.97
3	CCC(C)CCC(C)CC	Alkane	394	10.37	10.38	10.39	10.35	11.44
4	CCC(C)CCC(C)CC	Alkane	428	11.37	11.39	11.36	11.17	11.44
5	CCC(C)CCC(C)CC	Alkane	462	12.19	12.20	12.20	12.03	11.44
6	CCCCCCCC(C)C	Alkane	332	7.57	8.47	7.85	7.02	8.40
7	CCCCCCCC(C)C	Alkane	366.25	9.14	8.54	8.86	9.01	8.84
8	CCCCCCCC(C)C	Alkane	400.5	10.38	10.39	10.30	10.27	10.09
9	CCCCCCCC(C)C	Alkane	434.75	11.39	11.38	11.24	11.12	11.09
10	CCCCCCCC(C)C	Alkane	469	12.21	11.69	11.79	11.86	11.13
11	CC(C)C	Alkane	188	7.40	6.69	6.69	7.25	12.28
12	CC(C)C	Alkane	210.5	9.01	7.36	7.90	8.30	12.59
13	CC(C)C	Alkane	233	10.28	9.14	8.77	9.40	12.60
14	CC(C)C	Alkane	255.5	11.29	10.52	10.11	10.59	12.60
15	CC(C)C	Alkane	278	12.13	11.31	11.20	10.86	12.60
16	C1CCCCC1	Alkane	291	7.58	9.17	8.43	8.27	10.33
17	C1CCCCC1	Alkane	320.25	9.04	8.58	9.63	9.33	10.34
18	C1CCCCC1	Alkane	349.5	10.20	10.45	10.99	10.55	10.35
19	C1CCCCC1	Alkane	378.75	11.15	10.40	11.28	11.36	10.38
20	C1CCCCC1	Alkane	408	11.94	10.84	11.07	11.88	10.41
21	CCCCCCC(C)CCC	Alkane	340	7.20	7.16	6.96	6.68	8.53
22	CCCCCCC(C)CCC	Alkane	370	8.54	8.44	8.39	8.12	8.57
23	CCCCCCC(C)CCC	Alkane	400	9.68	9.65	9.61	10.07	9.81
24	CCCCCCC(C)CCC	Alkane	430	10.67	10.13	10.51	10.75	10.84
25	CCCCCCC(C)CCC	Alkane	460	11.52	11.53	11.56	11.54	11.10
26	C1CCCC1	Alkane	236	7.43	6.92	7.79	7.39	10.29
27	C1CCCC1	Alkane	264	9.10	9.91	9.86	9.02	10.29
28	C1CCCC1	Alkane	292	10.40	10.64	10.66	10.16	10.37
29	C1CCCC1	Alkane	320	11.45	12.28	11.78	10.77	10.54
30	C1CCCC1	Alkane	348	12.30	12.52	11.72	11.69	10.73
31	CCCCCC(C)C(C)C	Alkane	329	7.61	7.61	7.59	7.59	9.26
32	CCCCCC(C)C(C)C	Alkane	363.25	9.16	9.16	9.14	9.24	9.37

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.3 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน alkane ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CCCCCC(C)C(C)C	Alkane	397.5	10.39	10.39	10.37	10.36	9.51
34	CCCCCC(C)C(C)C	Alkane	431.75	11.38	11.38	11.30	11.47	9.65
35	CCCCCC(C)C(C)C	Alkane	466	12.20	12.20	12.19	12.21	9.75
36	C1CC12CC2	Alkane	276	10.09	9.91	10.34	9.73	9.73
37	C1CC12CC2	Alkane	293	10.82	10.64	10.77	10.21	9.73
38	C1CC12CC2	Alkane	310	11.45	12.28	11.71	10.90	10.30
39	C1CC12CC2	Alkane	327	12.01	11.72	11.77	11.24	11.59
40	C1CC12CC2	Alkane	344	12.51	12.03	11.74	11.51	11.59
41	C1CCCC2(CC1)CCCCC2	Alkane	389	4.15	8.86	8.38	8.04	8.71
42	C1CCCC2(CC1)CCCCC2	Alkane	399.25	4.94	8.86	8.54	8.03	8.71
43	C1CCCC2(CC1)CCCCC2	Alkane	409.5	5.70	8.86	8.66	8.44	8.71
44	C1CCCC2(CC1)CCCCC2	Alkane	419.75	6.42	9.09	9.12	9.05	8.71
45	C1CCCC2(CC1)CCCCC2	Alkane	430	7.10	9.09	9.42	9.50	8.71
46	CC(C)CC(C)CC(C)C	Alkane	317	7.61	7.60	7.56	6.43	9.51
47	CC(C)CC(C)CC(C)C	Alkane	349.75	9.16	9.17	9.04	8.05	9.51
48	CC(C)CC(C)CC(C)C	Alkane	382.5	10.39	10.41	10.14	8.98	9.51
49	CC(C)CC(C)CC(C)C	Alkane	415.25	11.38	11.38	11.07	10.31	9.51
50	CC(C)CC(C)CC(C)C	Alkane	448	12.20	11.53	12.06	11.40	9.51
51	CC(C)CCC(C)C(C)C	Alkane	322	7.58	7.60	7.59	7.54	9.91
52	CC(C)CCC(C)C(C)C	Alkane	355.75	9.15	9.16	9.15	9.08	9.91
53	CC(C)CCC(C)C(C)C	Alkane	389.5	10.38	10.39	10.38	10.48	9.91
54	CC(C)CCC(C)C(C)C	Alkane	423.25	11.37	11.39	11.36	11.29	10.61
55	CC(C)CCC(C)C(C)C	Alkane	457	12.19	12.20	12.17	11.95	10.61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.4 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน alkene ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CC(C)C=C	Alkene	213	7.40	7.26	7.83	7.93	11.01
2	CC(C)C=C	Alkene	236.75	8.97	7.26	8.09	8.48	11.01
3	CC(C)C=C	Alkene	260.5	10.20	10.40	9.97	9.59	11.01
4	CC(C)C=C	Alkene	284.25	11.20	10.40	10.52	10.62	11.01
5	CC(C)C=C	Alkene	308	12.02	11.39	11.51	11.58	11.01
6	C/C=C/C#N	Alkene	254	4.92	9.26	7.31	6.10	8.71
7	C/C=C/C#N	Alkene	289.25	7.25	10.53	9.02	7.49	9.68
8	C/C=C/C#N	Alkene	324.5	9.01	10.53	10.49	9.32	9.93
9	C/C=C/C#N	Alkene	359.75	10.39	12.43	11.83	10.11	10.82
10	C/C=C/C#N	Alkene	395	11.50	12.07	11.28	11.00	10.82
11	CC/C=C/CC	Alkene	253	7.57	7.36	7.55	7.74	10.85
12	CC/C=C/CC	Alkene	280.5	9.14	9.33	9.15	9.01	10.85
13	CC/C=C/CC	Alkene	308	10.37	10.31	10.18	10.25	10.85
14	CC/C=C/CC	Alkene	335.5	11.37	11.52	11.40	11.21	10.85
15	CC/C=C/CC	Alkene	363	12.20	12.20	11.94	11.82	10.85
16	CC(C)CCC(=C)C	Alkene	288	7.62	7.62	7.66	7.66	10.14
17	CC(C)CCC(=C)C	Alkene	318.5	9.17	9.17	9.17	9.41	10.15
18	CC(C)CCC(=C)C	Alkene	349	10.39	10.39	10.39	10.51	10.17
19	CC(C)CCC(=C)C	Alkene	379.5	11.37	11.38	11.36	11.52	10.18
20	CC(C)CCC(=C)C	Alkene	410	12.19	12.20	12.22	12.12	10.19
21	CC(C)/C=C\C(C)C	Alkene	285	7.96	7.61	7.59	7.82	10.14
22	CC(C)/C=C\C(C)C	Alkene	312.75	9.36	9.14	9.15	9.15	10.14
23	CC(C)/C=C\C(C)C	Alkene	340.5	10.49	10.37	10.23	10.30	10.14
24	CC(C)/C=C\C(C)C	Alkene	368.25	11.42	11.36	10.84	11.36	10.14
25	CC(C)/C=C\C(C)C	Alkene	396	12.20	11.98	11.79	12.02	10.14
26	CC/C(=C\C)/C	Alkene	253	7.58	7.36	7.55	7.59	10.15
27	CC/C(=C\C)/C	Alkene	280.75	9.15	9.33	9.08	8.97	10.15
28	CC/C(=C\C)/C	Alkene	308.5	10.38	10.31	10.16	10.15	10.15
29	CC/C(=C\C)/C	Alkene	336.25	11.38	11.39	11.17	11.11	10.15
30	CC/C(=C\C)/C	Alkene	364	12.20	12.20	11.81	11.84	10.15
31	CCC(C)C(=C)C	Alkene	267	7.61	7.25	7.63	7.84	10.13
32	CCC(C)C(=C)C	Alkene	295.5	9.16	9.16	9.16	9.25	10.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.4 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน alkene ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CCC(C)C(=C)C	Alkene	324	10.38	10.40	10.44	10.53	10.13
34	CCC(C)C(=C)C	Alkene	352.5	11.37	11.39	11.39	11.36	10.13
35	CCC(C)C(=C)C	Alkene	381	12.19	12.20	12.11	12.17	10.13
36	CC(C)(C)CC=C	Alkene	255	7.59	5.79	6.19	6.79	4.41
37	CC(C)(C)CC=C	Alkene	283.75	9.16	6.79	7.10	8.53	4.41
38	CC(C)(C)CC=C	Alkene	312.5	10.40	9.61	8.24	9.88	4.41
39	CC(C)(C)CC=C	Alkene	341.25	11.39	8.24	8.77	11.11	4.41
40	CC(C)(C)CC=C	Alkene	370	12.20	11.94	10.49	12.00	4.41
41	CC1=CCCCC1	Alkene	310	8.86	9.14	9.24	9.28	10.99
42	CC1=CCCCC1	Alkene	328.25	9.66	10.39	10.30	10.24	10.99
43	CC1=CCCCC1	Alkene	346.5	10.35	10.83	10.94	10.55	10.99
44	CC1=CCCCC1	Alkene	364.75	10.97	11.52	11.38	11.10	10.99
45	CC1=CCCCC1	Alkene	383	11.51	12.07	11.86	11.78	10.99
46	C=CC1CCC=CC1	Alkene	329	9.04	8.67	8.85	9.24	10.68
47	C=CC1CCC=CC1	Alkene	345.75	9.72	9.59	10.02	9.54	10.68
48	C=CC1CCC=CC1	Alkene	362.5	10.32	10.71	10.39	10.26	10.68
49	C=CC1CCC=CC1	Alkene	379.25	10.87	11.50	11.22	10.96	10.68
50	C=CC1CCC=CC1	Alkene	396	11.36	11.53	11.60	11.15	10.68
51	CCC(C)/C(=C\C)/C	Alkene	290	7.57	7.61	7.59	7.46	10.15
52	CCC(C)/C(=C\C)/C	Alkene	321.25	9.14	9.15	9.15	8.71	10.15
53	CCC(C)/C(=C\C)/C	Alkene	352.5	10.37	10.38	10.39	9.72	10.15
54	CCC(C)/C(=C\C)/C	Alkene	383.75	11.37	11.38	11.38	11.09	10.15
55	CCC(C)/C(=C\C)/C	Alkene	415	12.20	12.20	12.20	11.76	10.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.5 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Alkyne ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CCC#CC	Alkyne	245	7.59	8.98	8.82	7.74	10.71
2	CCC#CC	Alkyne	271.75	9.17	10.15	10.19	9.51	10.71
3	CCC#CC	Alkyne	298.5	10.40	10.63	11.42	10.72	10.71
4	CCC#CC	Alkyne	325.25	11.39	12.11	12.01	11.20	10.71
5	CCC#CC	Alkyne	352	12.21	12.43	11.57	11.95	10.71
6	CCCCCCC#C	Alkyne	301	7.61	6.46	7.43	8.03	10.86
7	CCCCCCC#C	Alkyne	332	9.15	9.17	8.62	9.48	10.86
8	CCCCCCC#C	Alkyne	363	10.37	10.00	10.12	10.35	10.87
9	CCCCCCC#C	Alkyne	394	11.37	11.38	11.12	11.20	10.87
10	CCCCCCC#C	Alkyne	425	12.20	11.38	11.71	11.83	10.87
11	CC#CC(C)(C)C	Alkyne	265	7.62	7.62	7.50	7.07	10.71
12	CC#CC(C)(C)C	Alkyne	294	9.18	9.08	8.99	8.63	10.71
13	CC#CC(C)(C)C	Alkyne	323	10.41	10.40	10.37	10.21	10.71
14	CC#CC(C)(C)C	Alkyne	352	11.40	11.39	11.16	10.84	10.71
15	CC#CC(C)(C)C	Alkyne	381	12.21	11.28	11.83	11.95	10.71
16	CCCCCCCCCCC#C	Alkyne	439	10.17	10.10	9.99	10.22	10.88
17	CCCCCCCCCCC#C	Alkyne	451.5	10.54	9.97	10.21	10.51	10.88
18	CCCCCCCCCCC#C	Alkyne	464	10.90	10.90	10.62	11.00	10.88
19	CCCCCCCCCCC#C	Alkyne	476.5	11.23	11.22	11.08	11.09	10.88
20	CCCCCCCCCCC#C	Alkyne	489	11.53	10.70	11.26	11.61	10.88
21	CC(C)(C)C#C	Alkyne	230	7.56	5.06	6.25	6.11	11.94
22	CC(C)(C)C#C	Alkyne	255.5	9.14	9.19	7.68	7.53	11.94
23	CC(C)(C)C#C	Alkyne	281	10.38	9.33	9.05	9.11	11.94
24	CC(C)(C)C#C	Alkyne	306.5	11.37	11.40	10.26	10.14	11.94
25	CC(C)(C)C#C	Alkyne	332	12.19	10.57	11.26	11.09	11.94
26	CCC(C)CC#C	Alkyne	272	7.58	7.58	7.75	7.78	10.14
27	CCC(C)CC#C	Alkyne	301	9.14	9.15	9.16	9.40	10.14
28	CCC(C)CC#C	Alkyne	330	10.37	10.38	10.38	10.76	10.14
29	CCC(C)CC#C	Alkyne	359	11.37	11.38	11.51	11.41	10.14
30	CCC(C)CC#C	Alkyne	388	12.19	12.21	12.08	12.31	10.14
31	CCC(CC)C#C	Alkyne	267	7.59	7.58	7.74	7.57	10.77
32	CCC(CC)C#C	Alkyne	295.5	9.15	9.16	9.15	9.08	10.77

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.5 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Alkyne ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CCC(CC)C#C	Alkyne	324	10.38	10.29	10.31	10.49	10.77
34	CCC(CC)C#C	Alkyne	352.5	11.38	11.39	11.46	11.32	10.77
35	CCC(CC)C#C	Alkyne	381	12.20	12.20	12.14	12.27	10.77
36	CC(C)C#C	Alkyne	225	7.62	7.50	7.79	7.86	10.14
37	CC(C)C#C	Alkyne	249.5	9.18	8.98	9.09	9.13	10.14
38	CC(C)C#C	Alkyne	274	10.41	10.40	10.32	10.12	10.15
39	CC(C)C#C	Alkyne	298.5	11.39	10.63	11.38	11.38	10.15
40	CC(C)C#C	Alkyne	323	12.21	12.11	11.84	11.84	10.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.6 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Amide ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	C1CCC(=O)NCC1	Amide	346	3.63	3.57	3.60	3.55	9.50
2	C1CCC(=O)NCC1	Amide	397	6.74	6.87	6.55	5.75	9.50
3	C1CCC(=O)NCC1	Amide	448	8.86	10.33	10.03	8.57	9.50
4	C1CCC(=O)NCC1	Amide	499	10.41	10.33	10.84	9.39	9.50
5	C1CCC(=O)NCC1	Amide	550	11.58	11.62	11.81	11.10	9.50
6	CCN(CC)C(=O)C1=CC=CC(=C1)C	Amide	298	4.01	-0.55	2.06	0.91	5.81
7	CCN(CC)C(=O)C1=CC=CC(=C1)C	Amide	324.5	5.07	1.52	3.51	2.57	5.81
8	CCN(CC)C(=O)C1=CC=CC(=C1)C	Amide	351	5.97	5.10	5.05	4.34	5.81
9	CCN(CC)C(=O)C1=CC=CC(=C1)C	Amide	377.5	6.75	4.35	5.70	5.45	5.81
10	CCN(CC)C(=O)C1=CC=CC(=C1)C	Amide	404	7.42	7.01	7.01	6.80	5.81
11	CCN(CC)C(=O)N(CC)CC	Amide	245	-2.67	1.35	1.73	1.18	8.48
12	CCN(CC)C(=O)N(CC)CC	Amide	273.75	0.90	1.35	1.75	2.84	8.48
13	CCN(CC)C(=O)N(CC)CC	Amide	302.5	3.62	5.17	3.44	5.13	8.48
14	CCN(CC)C(=O)N(CC)CC	Amide	331.25	5.77	4.86	4.92	6.59	8.48
15	CCN(CC)C(=O)N(CC)CC	Amide	360	7.50	5.54	6.57	7.97	8.48
16	CNC=O	Amide	369	7.82	8.02	8.44	7.67	9.52
17	CNC=O	Amide	394.75	8.96	8.11	8.94	9.47	9.53
18	CNC=O	Amide	420.5	9.93	9.24	10.20	10.20	9.53
19	CNC=O	Amide	446.25	10.78	10.33	10.68	10.48	9.53
20	CNC=O	Amide	472	11.51	10.33	11.14	11.17	9.53
21	CC(=O)N	Amide	381	7.18	6.79	7.32	7.99	9.61
22	CC(=O)N	Amide	408.75	8.55	8.60	8.46	8.81	9.62
23	CC(=O)N	Amide	436.5	9.71	8.60	10.12	11.09	9.62
24	CC(=O)N	Amide	464.25	10.69	11.24	11.07	11.55	9.62
25	CC(=O)N	Amide	492	11.53	11.24	11.57	12.28	9.62
26	CC(=C)C(=O)NC(C)(C)C	Amide	391	8.96	5.62	6.57	6.52	7.97
27	CC(=C)C(=O)NC(C)(C)C	Amide	401.5	9.39	6.87	7.00	6.50	7.97
28	CC(=C)C(=O)NC(C)(C)C	Amide	412	9.80	6.87	7.40	7.52	7.97
29	CC(=C)C(=O)NC(C)(C)C	Amide	422.5	10.18	6.87	7.87	8.41	7.97
30	CC(=C)C(=O)NC(C)(C)C	Amide	433	10.53	7.87	8.50	8.50	7.97
31	CCCCCCCCNC(=O)C(C)O	Amide	404	3.22	6.87	6.96	5.78	9.50
32	CCCCCCCCNC(=O)C(C)O	Amide	420.5	4.34	6.87	7.43	6.07	9.50
33	CCCCCCCCNC(=O)C(C)O	Amide	437	5.38	10.30	9.15	8.20	9.50
34	CCCCCCCCNC(=O)C(C)O	Amide	453.5	6.34	10.30	9.63	8.73	9.50
35	CCCCCCCCNC(=O)C(C)O	Amide	470	7.24	11.15	10.08	9.33	9.50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.7 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Amine ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CCC(C)NCC	Amine	283	7.84	9.33	8.91	7.78	11.32
2	CCC(C)NCC	Amine	290.5	8.26	9.17	9.25	8.38	11.32
3	CCC(C)NCC	Amine	298	8.66	9.33	9.47	8.95	11.32
4	CCC(C)NCC	Amine	305.5	9.03	9.99	9.68	9.07	11.32
5	CCC(C)NCC	Amine	313	9.38	9.99	9.85	9.26	11.32
6	CC(C)CNCC(C)C	Amine	268	4.90	5.73	6.02	3.92	9.66
7	CC(C)CNCC(C)C	Amine	304	7.21	7.62	7.38	6.36	9.66
8	CC(C)CNCC(C)C	Amine	340	8.98	8.79	9.64	8.58	9.66
9	CC(C)CNCC(C)C	Amine	376	10.38	11.38	11.18	9.76	9.66
10	CC(C)CNCC(C)C	Amine	412	11.52	12.20	11.73	10.76	9.66
11	CNC1=CC=NC=C1	Amine	313	3.40	4.87	5.98	5.61	10.13
12	CNC1=CC=NC=C1	Amine	320.5	3.90	5.18	6.22	5.89	10.13
13	CNC1=CC=NC=C1	Amine	328	4.38	5.18	6.53	6.30	10.13
14	CNC1=CC=NC=C1	Amine	335.5	4.84	6.66	7.26	6.52	10.13
15	CNC1=CC=NC=C1	Amine	343	5.28	7.30	7.77	6.56	10.13
16	C1CCCNCC1	Amine	489	13.24	11.69	11.73	12.60	10.97
17	C1CCCNCC1	Amine	521.75	13.80	10.39	12.35	13.37	10.97
18	C1CCCNCC1	Amine	554.5	14.32	11.51	12.89	13.91	10.97
19	C1CCCNCC1	Amine	587.25	14.82	11.54	13.06	14.91	10.97
20	C1CCCNCC1	Amine	620	15.29	13.05	13.38	14.91	10.98
21	CC(C)(C)N(C)C	Amine	283	8.34	9.33	8.45	9.23	12.60
22	CC(C)(C)N(C)C	Amine	291.75	8.79	9.17	8.70	9.57	12.60
23	CC(C)(C)N(C)C	Amine	300.5	9.21	8.40	9.14	10.14	12.60
24	CC(C)(C)N(C)C	Amine	309.25	9.60	8.40	9.55	10.27	12.60
25	CC(C)(C)N(C)C	Amine	318	9.96	8.91	9.79	10.45	12.60
26	CC1=CC(=CC=C1)N	Amine	394	8.90	5.96	8.06	9.08	4.40
27	CC1=CC(=CC=C1)N	Amine	414.75	9.69	5.81	8.67	8.64	4.40
28	CC1=CC(=CC=C1)N	Amine	435.5	10.39	5.81	9.42	8.91	4.40
29	CC1=CC(=CC=C1)N	Amine	456.25	11.00	11.86	10.05	9.85	4.40
30	CC1=CC(=CC=C1)N	Amine	477	11.54	11.54	10.45	10.25	4.40
31	CNC	Amine	201	6.43	10.76	9.04	8.41	10.11
32	CNC	Amine	220.75	8.14	10.76	8.72	9.27	10.11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.7 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Amine ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CNC	Amine	240.5	9.50	10.76	9.60	9.96	11.71
34	CNC	Amine	260.25	10.61	11.38	10.60	11.13	12.28
35	CNC	Amine	280	11.53	12.26	11.01	11.38	12.60
36	C(CCCCCCN)CCCCCN	Amine	313	-2.19	0.96	2.00	-0.01	11.45
37	C(CCCCCCN)CCCCCN	Amine	323.25	-0.85	1.51	2.55	0.64	11.45
38	C(CCCCCCN)CCCCCN	Amine	333.5	0.41	2.84	4.15	1.88	11.45
39	C(CCCCCCN)CCCCCN	Amine	343.75	1.60	3.65	4.67	2.56	11.45
40	C(CCCCCCN)CCCCCN	Amine	354	2.71	3.65	5.36	3.60	11.45
41	CC(C)CN	Amine	255	7.39	10.52	9.97	8.71	10.10
42	CC(C)CN	Amine	263.5	7.96	9.98	10.33	9.01	10.12
43	CC(C)CN	Amine	272	8.48	11.31	10.55	9.37	10.15
44	CC(C)CN	Amine	280.5	8.97	11.31	10.63	9.24	10.17
45	CC(C)CN	Amine	289	9.41	10.93	10.90	9.77	10.19
46	CNC1=CN=CC=C1	Amine	313	3.48	4.87	6.24	5.77	10.58
47	CNC1=CN=CC=C1	Amine	320.5	4.00	5.18	6.49	6.05	10.58
48	CNC1=CN=CC=C1	Amine	328	4.49	5.18	6.74	6.46	10.58
49	CNC1=CN=CC=C1	Amine	335.5	4.96	6.66	7.45	7.01	10.58
50	CNC1=CN=CC=C1	Amine	343	5.41	7.30	7.87	7.05	10.58
51	CC1=CC(=C(C(=C1)C)N)C	Amine	341	4.86	7.46	6.62	6.13	12.46
52	CC1=CC(=C(C(=C1)C)N)C	Amine	375	6.82	7.22	8.13	7.30	12.43
53	CC1=CC(=C(C(=C1)C)N)C	Amine	409	8.43	9.45	9.47	8.96	12.43
54	CC1=CC(=C(C(=C1)C)N)C	Amine	443	9.76	10.39	10.62	9.91	12.43
55	CC1=CC(=C(C(=C1)C)N)C	Amine	477	10.89	11.52	11.44	10.84	12.20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.8 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CC1=CC(=NC2=CC=CC=C12)C	Aromatic	458	9.34	9.39	9.46	9.27	11.39
2	CC1=CC(=NC2=CC=CC=C12)C	Aromatic	479.5	10.00	9.98	10.17	9.81	11.39
3	CC1=CC(=NC2=CC=CC=C12)C	Aromatic	501	10.59	10.46	10.77	10.19	11.39
4	CC1=CC(=NC2=CC=CC=C12)C	Aromatic	522.5	11.12	11.05	11.32	10.90	11.39
5	CC1=CC(=NC2=CC=CC=C12)C	Aromatic	544	11.60	11.56	11.62	11.25	11.39
6	CCC1=C(C=C(C(=C1)C)C)C	Aromatic	360	7.14	6.47	7.20	7.07	10.16
7	CCC1=C(C=C(C(=C1)C)C)C	Aromatic	389.75	8.54	8.85	8.38	8.61	10.16
8	CCC1=C(C=C(C(=C1)C)C)C	Aromatic	419.5	9.67	9.16	9.46	9.54	10.16
9	CCC1=C(C=C(C(=C1)C)C)C	Aromatic	449.25	10.60	11.17	10.57	10.81	10.16
10	CCC1=C(C=C(C(=C1)C)C)C	Aromatic	479	11.38	11.36	11.20	11.39	10.16
11	C1=CC=C2C(=C1)C3=CC=CC=C23	Aromatic	338	3.86	8.77	6.52	5.23	9.69
12	C1=CC=C2C(=C1)C3=CC=CC=C23	Aromatic	355.5	4.93	4.07	5.81	6.60	9.69
13	C1=CC=C2C(=C1)C3=CC=CC=C23	Aromatic	373	5.88	4.07	6.26	6.91	9.69
14	C1=CC=C2C(=C1)C3=CC=CC=C23	Aromatic	390.5	6.74	7.59	6.64	7.88	9.69
15	C1=CC=C2C(=C1)C3=CC=CC=C23	Aromatic	408	7.51	11.44	7.62	8.73	9.69
16	CC1=CC=NC=C1	Aromatic	348	9.16	8.63	8.61	9.02	12.60
17	CC1=CC=NC=C1	Aromatic	376	10.23	9.77	9.05	9.74	12.61
18	CC1=CC=NC=C1	Aromatic	404	11.12	10.82	10.93	11.37	12.61
19	CC1=CC=NC=C1	Aromatic	432	11.88	11.74	11.52	11.49	12.61
20	CC1=CC=NC=C1	Aromatic	460	12.53	12.02	11.80	12.08	12.42
21	CCC1=CC(=C(C=C1)C)C	Aromatic	342	7.60	6.47	7.64	6.99	10.14
22	CCC1=CC(=C(C=C1)C)C	Aromatic	380.25	9.18	9.17	9.34	8.74	10.14
23	CCC1=CC(=C(C=C1)C)C	Aromatic	418.5	10.41	10.41	10.42	10.10	10.14
24	CCC1=CC(=C(C=C1)C)C	Aromatic	456.75	11.39	12.14	11.49	11.08	10.14
25	CCC1=CC(=C(C=C1)C)C	Aromatic	495	12.20	12.21	12.08	12.03	10.15
26	CC1=CC(=C(C=C1C)C)C	Aromatic	355	7.61	6.47	7.32	7.63	8.99
27	CC1=CC(=C(C=C1C)C)C	Aromatic	391.25	9.15	9.17	8.97	9.04	8.99
28	CC1=CC(=C(C=C1C)C)C	Aromatic	427.5	10.37	10.41	10.21	10.09	8.99
29	CC1=CC(=C(C=C1C)C)C	Aromatic	463.75	11.37	10.64	10.99	11.00	8.99
30	CC1=CC(=C(C=C1C)C)C	Aromatic	500	12.20	10.65	11.71	11.61	8.99
31	C1=CC=C(C=C1)N=C=O	Aromatic	283	4.88	1.92	2.50	2.50	10.12
32	C1=CC=C(C=C1)N=C=O	Aromatic	326.25	7.43	4.89	5.18	4.24	10.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.8 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Aromatic ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	C1=CC=C(C=C1)N=C=O	Aromatic	369.5	9.33	7.57	7.74	6.73	10.12
34	C1=CC=C(C=C1)N=C=O	Aromatic	412.75	10.79	10.71	9.41	8.88	10.12
35	C1=CC=C(C=C1)N=C=O	Aromatic	456	11.96	9.28	10.12	10.04	10.12
36	C1=CC=C2C=CC=C2C=C1	Aromatic	373	6.76	4.07	6.19	7.01	9.51
37	C1=CC=C2C=CC=C2C=C1	Aromatic	385.5	7.33	6.68	6.42	7.58	9.51
38	C1=CC=C2C=CC=C2C=C1	Aromatic	398	7.86	7.59	6.95	8.00	9.51
39	C1=CC=C2C=CC=C2C=C1	Aromatic	410.5	8.33	11.44	7.45	8.87	9.51
40	C1=CC=C2C=CC=C2C=C1	Aromatic	423	8.77	11.74	7.97	8.68	9.51
41	CC1=CN=C(C=C1)C	Aromatic	358	9.16	9.27	8.96	9.47	9.71
42	CC1=CN=C(C=C1)C	Aromatic	376.25	9.86	10.29	9.60	9.76	9.71
43	CC1=CN=C(C=C1)C	Aromatic	394.5	10.48	10.63	10.41	10.40	9.71
44	CC1=CN=C(C=C1)C	Aromatic	412.75	11.04	11.44	11.10	11.32	9.71
45	CC1=CN=C(C=C1)C	Aromatic	431	11.55	11.74	11.49	11.36	9.71
46	C1=CC=NC=C1	Aromatic	327	7.83	8.77	7.94	8.95	10.59
47	C1=CC=NC=C1	Aromatic	352.5	9.00	4.07	7.31	9.94	10.62
48	C1=CC=NC=C1	Aromatic	378	9.99	10.63	7.86	10.66	10.64
49	C1=CC=NC=C1	Aromatic	403.5	10.83	11.44	9.57	11.94	10.65
50	C1=CC=NC=C1	Aromatic	429	11.55	11.74	10.08	11.60	10.65
51	CC1=CC=CC=C1	Aromatic	286	7.58	4.94	5.39	5.94	12.60
52	CC1=CC=CC=C1	Aromatic	317	9.15	7.61	6.80	7.65	12.60
53	CC1=CC=CC=C1	Aromatic	348	10.39	9.17	7.95	8.86	12.60
54	CC1=CC=CC=C1	Aromatic	379	11.39	10.63	9.23	10.04	12.61
55	CC1=CC=CC=C1	Aromatic	410	12.22	11.44	10.63	11.38	12.61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.9 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Carboxylic Acid ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CCCCCCCC(=O)O	Carboxylic Acid	417	0.11	8.68	9.04	8.15	9.95
2	CCCCCCCC(=O)O	Carboxylic Acid	441.25	1.80	9.99	9.64	8.88	10.05
3	CCCCCCCC(=O)O	Carboxylic Acid	465.5	3.24	11.83	10.25	9.79	10.16
4	CCCCCCCC(=O)O	Carboxylic Acid	489.75	4.49	11.83	10.79	10.62	10.27
5	CCCCCCCC(=O)O	Carboxylic Acid	514	5.58	11.86	11.44	11.32	10.35
6	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	Carboxylic Acid	450	1.16	8.03	9.58	8.17	10.70
7	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	Carboxylic Acid	455	1.60	8.60	9.56	8.29	10.70
8	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	Carboxylic Acid	460	2.04	8.60	9.60	8.43	10.70
9	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	Carboxylic Acid	465	2.46	8.60	9.64	8.43	10.70
10	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	Carboxylic Acid	470	2.88	8.60	9.74	8.79	10.70
11	CC/C=C\C(=O)O	Carboxylic Acid	359	7.60	7.57	7.70	7.84	9.76
12	CC/C=C\C(=O)O	Carboxylic Acid	384.5	8.86	8.85	9.01	8.73	9.76
13	CC/C=C\C(=O)O	Carboxylic Acid	410	9.92	9.91	9.90	10.05	9.76
14	CC/C=C\C(=O)O	Carboxylic Acid	435.5	10.83	10.85	10.68	10.86	9.76
15	CC/C=C\C(=O)O	Carboxylic Acid	461	11.61	11.54	11.39	11.60	9.76
16	C(CC(=O)O)C(=O)O	Carboxylic Acid	458	4.85	5.93	8.26	8.49	10.12
17	C(CC(=O)O)C(=O)O	Carboxylic Acid	471.5	6.97	5.93	8.58	8.85	10.12
18	C(CC(=O)O)C(=O)O	Carboxylic Acid	485	8.80	5.93	9.24	8.99	10.12
19	C(CC(=O)O)C(=O)O	Carboxylic Acid	498.5	10.40	5.93	9.23	9.88	10.12
20	C(CC(=O)O)C(=O)O	Carboxylic Acid	512	11.81	9.44	9.53	10.71	10.12
21	CCCCCC(CC)C(=O)O	Carboxylic Acid	386	6.49	7.60	7.96	6.76	9.83
22	CCCCCC(CC)C(=O)O	Carboxylic Acid	408.25	7.57	8.68	8.56	8.10	9.83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.9 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Carboxylic Acid ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
23	CCCCCC(CC)C(=O)O	Carboxylic Acid	430.5	8.54	8.68	9.28	8.86	9.83
24	CCCCCC(CC)C(=O)O	Carboxylic Acid	452.75	9.41	9.99	10.00	9.68	9.83
25	CCCCCC(CC)C(=O)O	Carboxylic Acid	475	10.20	11.14	10.53	10.33	9.83
26	CC(C)CC(=O)O	Carboxylic Acid	363	8.09	7.57	7.55	7.98	12.60
27	CC(C)CC(=O)O	Carboxylic Acid	388.25	9.30	8.68	8.87	9.19	12.60
28	CC(C)CC(=O)O	Carboxylic Acid	413.5	10.32	9.91	9.80	10.49	12.60
29	CC(C)CC(=O)O	Carboxylic Acid	438.75	11.19	10.81	10.88	11.23	12.60
30	CC(C)CC(=O)O	Carboxylic Acid	464	11.94	11.54	11.28	11.76	12.60
31	CCCCC(=O)O	Carboxylic Acid	372	8.10	7.69	7.85	7.58	10.11
32	CCCCC(=O)O	Carboxylic Acid	395.25	9.21	7.60	8.76	8.80	10.11
33	CCCCC(=O)O	Carboxylic Acid	418.5	10.16	9.91	9.89	10.05	10.11
34	CCCCC(=O)O	Carboxylic Acid	441.75	10.98	11.78	10.60	10.80	10.11
35	CCCCC(=O)O	Carboxylic Acid	465	11.69	11.54	11.11	11.55	10.12
36	CCCCCCCCC(=O)O	Carboxylic Acid	419	7.58	8.68	8.82	7.73	8.74
37	CCCCCCCCC(=O)O	Carboxylic Acid	453.5	9.12	9.99	9.18	8.93	9.49
38	CCCCCCCCC(=O)O	Carboxylic Acid	488	10.35	11.15	10.15	10.25	10.67
39	CCCCCCCCC(=O)O	Carboxylic Acid	522.5	11.36	11.53	11.25	10.97	10.88
40	CCCCCCCCC(=O)O	Carboxylic Acid	557	12.21	11.63	11.54	11.84	11.49
41	C(=O)O	Carboxylic Acid	278	7.58	12.26	9.41	7.56	12.39
42	C(=O)O	Carboxylic Acid	308	9.11	7.02	7.95	8.03	12.60
43	C(=O)O	Carboxylic Acid	338	10.34	8.96	9.05	8.15	12.60
44	C(=O)O	Carboxylic Acid	368	11.35	10.41	10.18	9.59	11.22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.9 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Carboxylic Acid ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
45	C(=O)O	Carboxylic Acid	398	12.20	10.63	9.76	10.87	12.61
46	C1=CC=C(C=C1)CCC(=O)O	Carboxylic Acid	375	4.87	5.55	7.70	5.44	11.42
47	C1=CC=C(C=C1)CCC(=O)O	Carboxylic Acid	419.5	7.09	5.86	8.82	7.97	11.42
48	C1=CC=C(C=C1)CCC(=O)O	Carboxylic Acid	464	8.87	8.60	9.70	9.33	11.42
49	C1=CC=C(C=C1)CCC(=O)O	Carboxylic Acid	508.5	10.32	14.04	10.62	10.65	11.42
50	C1=CC=C(C=C1)CCC(=O)O	Carboxylic Acid	553	11.53	14.04	11.41	11.57	11.42
51	C(CCC(=O)O)CCC(=O)O	Carboxylic Acid	436	4.86	8.95	6.96	5.81	8.50
52	C(CCC(=O)O)CCC(=O)O	Carboxylic Acid	480.75	7.01	5.93	7.75	7.58	8.51
53	C(CCC(=O)O)CCC(=O)O	Carboxylic Acid	525.5	8.78	9.44	9.40	9.46	8.51
54	C(CCC(=O)O)CCC(=O)O	Carboxylic Acid	570.25	10.26	9.44	10.20	11.52	8.52
55	C(CCC(=O)O)CCC(=O)O	Carboxylic Acid	615	11.52	11.51	11.33	12.30	8.52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.10 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Ester ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CCCCCOC(=O)C(=C)C	Ester	339.00	7.20	7.17	7.29	7.22	8.78
2	CCCCCOC(=O)C(=C)C	Ester	368.25	8.54	8.79	8.28	8.46	9.51
3	CCCCCOC(=O)C(=C)C	Ester	397.50	9.69	9.52	9.41	9.27	9.51
4	CCCCCOC(=O)C(=C)C	Ester	426.75	10.67	10.34	10.33	10.64	9.51
5	CCCCCOC(=O)C(=C)C	Ester	456.00	11.54	11.41	11.14	11.38	9.51
6	CC(C)OC	Ester	260.00	9.64	9.58	9.37	9.66	10.59
7	CC(C)OC	Ester	276.25	10.42	9.79	9.76	9.95	10.62
8	CC(C)OC	Ester	292.50	11.10	9.79	10.09	10.80	10.62
9	CC(C)OC	Ester	308.75	11.70	11.56	11.20	11.55	10.62
10	CC(C)OC	Ester	325.00	12.22	11.39	11.51	11.98	10.62
11	CCCCCCOC(=O)CCC(=O)C	Ester	363.00	4.80	5.92	5.84	4.73	8.95
12	CCCCCCOC(=O)CCC(=O)C	Ester	407.25	7.17	7.58	7.65	7.42	8.95
13	CCCCCCOC(=O)CCC(=O)C	Ester	451.50	8.97	9.03	9.45	9.05	8.96
14	CCCCCCOC(=O)CCC(=O)C	Ester	495.75	10.39	11.54	11.05	10.55	8.96
15	CCCCCCOC(=O)CCC(=O)C	Ester	540.00	11.53	11.54	11.78	11.55	8.97
16	CCOC(=O)C(=O)OCC	Ester	300.00	4.14	5.30	5.09	6.07	8.60
17	CCOC(=O)C(=O)OCC	Ester	339.00	6.81	9.41	6.20	7.96	8.60
18	CCOC(=O)C(=O)OCC	Ester	378.00	8.80	7.55	8.22	9.12	8.60
19	CCOC(=O)C(=O)OCC	Ester	417.00	10.33	8.95	9.34	9.78	8.60
20	CCOC(=O)C(=O)OCC	Ester	456.00	11.54	10.38	10.55	11.25	8.60
21	CC(=O)OCCC(=O)OC	Ester	343.00	6.44	7.14	6.46	7.47	9.59
22	CC(=O)OCCC(=O)OC	Ester	345.75	6.63	7.14	7.09	7.47	9.59
23	CC(=O)OCCC(=O)OC	Ester	348.50	6.82	7.14	7.14	7.40	9.59
24	CC(=O)OCCC(=O)OC	Ester	351.25	7.00	7.14	7.17	7.40	9.59
25	CC(=O)OCCC(=O)OC	Ester	354.00	7.18	7.14	7.26	7.86	9.59
26	CCCCCCC1COCCC1O	Ester	293.00	-2.64	2.09	0.74	-2.42	11.68
27	CCCCCCC1COCCC1O	Ester	313.00	-0.69	2.95	2.55	0.39	11.68
28	CCCCCCC1COCCC1O	Ester	333.00	1.01	6.50	5.13	2.60	11.68
29	CCCCCCC1COCCC1O	Ester	353.00	2.52	7.19	5.80	4.79	11.68
30	CCCCCCC1COCCC1O	Ester	373.00	3.86	6.44	6.96	5.57	11.68
31	CCOC(=O)C(=C)C	Ester	285.00	7.16	7.49	7.12	7.23	10.16
32	CCOC(=O)C(=C)C	Ester	311.25	8.52	8.57	8.23	8.38	10.16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.10 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Ester ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CCOC(=O)C(=C)C	Ester	337.50	9.68	10.29	9.18	10.21	10.16
34	CCOC(=O)C(=C)C	Ester	363.75	10.66	10.81	10.22	10.60	10.16
35	CCOC(=O)C(=C)C	Ester	390.00	11.52	10.59	11.31	11.52	10.16
36	COC(=O)C1=CC=CC=C1C(=O)OC	Ester	495.00	10.07	9.48	9.71	9.55	9.15
37	COC(=O)C1=CC=CC=C1C(=O)OC	Ester	502.00	10.30	9.48	9.96	10.27	9.15
38	COC(=O)C1=CC=CC=C1C(=O)OC	Ester	509.00	10.51	10.50	10.20	10.27	9.15
39	COC(=O)C1=CC=CC=C1C(=O)OC	Ester	516.00	10.71	10.50	10.57	10.32	9.15
40	COC(=O)C1=CC=CC=C1C(=O)OC	Ester	523.00	10.89	10.50	10.69	10.68	9.15
41	CCOC(=O)CC(=O)C	Ester	301.00	4.85	4.41	5.29	5.79	10.17
42	CCOC(=O)CC(=O)C	Ester	340.75	7.22	4.73	5.92	7.71	10.17
43	CCOC(=O)CC(=O)C	Ester	380.50	9.05	9.41	9.01	8.98	10.17
44	CCOC(=O)CC(=O)C	Ester	420.25	10.50	11.13	9.83	9.99	10.18
45	CCOC(=O)CC(=O)C	Ester	460.00	11.69	10.38	10.31	11.02	10.18
46	C/1CC2C(O2)CC/C=C1	Ester	355.00	7.85	10.15	10.10	8.80	8.63
47	C/1CC2C(O2)CC/C=C1	Ester	364.50	8.31	10.56	10.29	9.41	8.63
48	C/1CC2C(O2)CC/C=C1	Ester	374.00	8.73	10.19	10.68	9.93	8.63
49	C/1CC2C(O2)CC/C=C1	Ester	383.50	9.12	11.17	10.84	10.21	8.63
50	C/1CC2C(O2)CC/C=C1	Ester	393.00	9.48	11.39	10.94	10.36	8.63
51	CCOC(=O)C	Ester	271.00	7.95	7.49	7.63	8.73	10.40
52	CCOC(=O)C	Ester	296.50	9.36	9.38	8.60	9.23	10.41
53	CCOC(=O)C	Ester	322.00	10.50	8.55	9.55	10.12	10.42
54	CCOC(=O)C	Ester	347.50	11.43	11.46	10.72	11.46	10.42
55	CCOC(=O)C	Ester	373.00	12.22	12.29	11.53	11.81	10.43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.11 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Ketone ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	C[C@@]12CC[C@@H](C1(C)C)CC2=O	Ketone	451	10.79	10.67	10.62	10.95	10.09
2	C[C@@]12CC[C@@H](C1(C)C)CC2=O	Ketone	464.5	11.14	11.52	11.05	11.23	10.09
3	C[C@@]12CC[C@@H](C1(C)C)CC2=O	Ketone	478	11.47	11.52	11.17	11.59	10.09
4	C[C@@]12CC[C@@H](C1(C)C)CC2=O	Ketone	491.5	11.77	11.52	11.19	11.87	10.09
5	C[C@@]12CC[C@@H](C1(C)C)CC2=O	Ketone	505	12.06	11.53	11.25	12.22	10.09
6	CC(=O)C1CC1	Ketone	361	10.72	9.93	10.16	10.05	8.52
7	CC(=O)C1CC1	Ketone	365	10.86	9.70	10.25	10.10	8.52
8	CC(=O)C1CC1	Ketone	369	10.99	11.22	10.58	10.17	8.52
9	CC(=O)C1CC1	Ketone	373	11.12	11.46	10.84	10.23	8.52
10	CC(=O)C1CC1	Ketone	377	11.25	12.05	11.03	10.74	8.53
11	CCC(C)CC(=O)C	Ketone	336	8.68	8.94	8.87	9.88	10.15
12	CCC(C)CC(=O)C	Ketone	356.5	9.57	9.93	9.63	9.83	10.15
13	CCC(C)CC(=O)C	Ketone	377	10.36	11.35	10.59	10.96	10.15
14	CCC(C)CC(=O)C	Ketone	397.5	11.06	11.57	11.35	11.38	10.15
15	CCC(C)CC(=O)C	Ketone	418	11.70	11.52	11.75	11.68	10.16
16	CC(CO)C(=O)C	Ketone	382	8.62	10.32	10.46	8.86	8.11
17	CC(CO)C(=O)C	Ketone	402.25	9.49	10.92	11.01	9.45	8.11
18	CC(CO)C(=O)C	Ketone	422.5	10.26	12.19	11.62	10.32	8.11
19	CC(CO)C(=O)C	Ketone	442.75	10.95	10.62	11.34	10.77	8.11
20	CC(CO)C(=O)C	Ketone	463	11.57	11.54	11.44	11.12	8.11
21	CCCCC(=O)C	Ketone	303	7.62	7.59	7.42	7.54	10.11
22	CCCCC(=O)C	Ketone	333.75	9.15	9.54	8.94	8.72	10.12
23	CCCCC(=O)C	Ketone	364.5	10.37	9.70	10.10	10.20	10.13
24	CCCCC(=O)C	Ketone	395.25	11.37	11.29	11.35	11.26	10.14
25	CCCCC(=O)C	Ketone	426	12.19	12.20	11.96	12.02	10.14
26	CC(=O)C1=CC2=CC=CC=C2C=C1	Ketone	393	4.88	4.89	7.55	5.92	10.13
27	CC(=O)C1=CC2=CC=CC=C2C=C1	Ketone	438.5	7.07	10.34	9.27	8.46	10.13
28	CC(=O)C1=CC2=CC=CC=C2C=C1	Ketone	484	8.85	10.34	10.54	9.89	10.13
29	CC(=O)C1=CC2=CC=CC=C2C=C1	Ketone	529.5	10.32	11.52	11.48	11.33	10.13
30	CC(=O)C1=CC2=CC=CC=C2C=C1	Ketone	575	11.55	11.52	11.95	11.85	10.13
31	CC(=CC(=O)C)C	Ketone	292	6.86	7.58	7.67	7.52	10.00
32	CC(=CC(=O)C)C	Ketone	321.25	8.48	9.03	9.15	8.40	10.01

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.11 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Ketone ของแหล่งอ้างอิงและการทำนาย (ต่อ)

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
33	CC(=CC(=O)C)C	Ketone	350.5	9.77	10.38	10.21	10.12	10.01
34	CC(=CC(=O)C)C	Ketone	379.75	10.82	11.35	11.26	11.28	10.01
35	CC(=CC(=O)C)C	Ketone	409	11.70	12.20	11.89	11.53	10.01
36	CCC(=O)C1=CC=CC=C1	Ketone	405	8.91	8.93	7.96	8.95	12.61
37	CCC(=O)C1=CC=CC=C1	Ketone	426.75	9.71	8.54	8.85	9.35	12.61
38	CCC(=O)C1=CC=CC=C1	Ketone	448.5	10.40	10.34	10.45	11.01	12.61
39	CCC(=O)C1=CC=CC=C1	Ketone	470.25	11.01	10.34	11.03	11.80	12.42
40	CCC(=O)C1=CC=CC=C1	Ketone	492	11.54	11.52	11.25	12.07	12.42
41	CCCCCCC(=O)C	Ketone	342	7.57	8.97	8.62	7.92	9.95
42	CCCCCCC(=O)C	Ketone	374.5	9.11	9.00	9.48	9.20	10.09
43	CCCCCCC(=O)C	Ketone	407	10.34	11.39	10.78	10.32	10.24
44	CCCCCCC(=O)C	Ketone	439.5	11.35	11.97	11.34	11.26	10.36
45	CCCCCCC(=O)C	Ketone	472	12.20	11.83	11.56	11.97	10.43
46	CCCCCCCC(=O)C	Ketone	357	7.57	7.17	7.50	7.53	9.58
47	CCCCCCCC(=O)C	Ketone	391.75	9.12	9.60	9.08	9.48	9.71
48	CCCCCCCC(=O)C	Ketone	426.5	10.35	10.37	10.28	10.33	9.89
49	CCCCCCCC(=O)C	Ketone	461.25	11.36	11.15	11.49	11.20	10.08
50	CCCCCCCC(=O)C	Ketone	496	12.20	11.15	11.61	11.99	10.23
51	CCCCCC(=O)CC	Ketone	239	1.57	4.81	5.22	2.25	10.13
52	CCCCCC(=O)CC	Ketone	266.25	3.62	4.81	5.55	3.54	10.14
53	CCCCCC(=O)CC	Ketone	293.5	5.38	5.24	6.18	5.54	10.14
54	CCCCCC(=O)CC	Ketone	320.75	6.90	7.58	7.45	6.69	10.14
55	CCCCCC(=O)CC	Ketone	348	8.23	8.97	8.89	8.28	10.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ค.12 ค่าความดันไอที่อุณหภูมิต่างๆของหมู่ฟังก์ชัน Ether ของแหล่งอ้างอิงและการทำนาย

ลำดับ	SMILES	หมู่ฟังก์ชัน	อุณหภูมิ	แหล่งอ้างอิง	Algorithm			
					DT	RF	XGB	KNN
1	CC1CCC(=O)OC1=O	Ether	366	4.85	7.23	7.02	6.93	10.00
2	CC1CCC(=O)OC1=O	Ether	413.25	7.14	8.99	8.92	8.68	10.82
3	CC1CCC(=O)OC1=O	Ether	460.5	8.92	10.40	10.34	10.33	10.82
4	CC1CCC(=O)OC1=O	Ether	507.75	10.34	11.53	11.54	11.63	10.82
5	CC1CCC(=O)OC1=O	Ether	555	11.51	11.53	12.61	12.38	10.82

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้