การพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ

Condominium Price Forecasts in Bangkok for Loan Assessment

ภควัต อมรวิลาศ

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2567

KMITL-2024-SC-M-017-020

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Condominium Price Forecasts in Bangkok for Loan Assessment

PAKAWAT AMORNVILAS

AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE AND ANALYTICS
KMITL DIGITAL ANALYTICS AND INTELLIGENCE CENTER SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2024

KMITL-2024-SC-M-017-020

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2024

SCHOOL OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อการศึกษาค้นคว้าอิสระ การพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ
ชื่อนักศึกษา นาย ภควัต อมรวิลาศ
รหัสประจำตัว 65056068
ปริญญา วิทยาศาสตรมหาบัณฑิต (วิทยาการข้อมูลและการวิเคราะห์)
ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง
พ.ศ. 2567
อาจารย์ที่ปรึกษาการศึกษาค้นคว้าอิสระ รองศาสตราจารย์ ดร.ฉัฐไชย์ ลีนาวงศ์

บทคัดย่อ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานคร เพื่อช่วยในการตัดสินใจในการให้สินเชื่อ โดยมุ่งเน้นสร้างโมเดลในการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ เพื่อลดขั้นตอนการส่งเจ้าหน้าที่ไปประเมินราคาคอนโดมิเนียม ซึ่งทำการเก็บรวบรวมข้อมูลข้อมูลจากเว็บไซต์ zmyhome.com ด้วยวิธี Web Scraping โดยเป็นข้อมูลโครงการคอนโดมิเนียมจำนวน 9,041 ห้อง ซึ่งมีแอตทริบิวต์ทั้งหมด 16 แอตทริบิวต์ และทำการเปรียบเทียบประสิทธิภาพของแบบจำลองระหว่าง Decision Tree, Random Forest, XGBoost และ LightGBM ผลการศึกษาพบว่า แบบจำลอง XGBoost มีประสิทธิภาพดีที่สุด โดยให้ค่า Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) และ R-Squared (R^2) เท่ากับ 8.4135%, 9,444.8167 และ 0.9527 ตามลำดับ ซึ่งผลการศึกษาแสดงถึงความเป็นไปได้ในการนำเทคนิค Machine Learning มาใช้ในการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Independent Study Title	Condominium Price Forecasts in Bangkok for Loan Assessment
Student	PAKAWAT AMORNVILAS
Student ID	65056068
Degree	Master of Science (Data Science and Analytics) KMITL Digital Analytics and Intelligence Center
Year	2024
Independent Study Advisor	Assoc. Prof. Dr. Chartchai Leenawong

Abstract

The objective of this study is to forecast condominium prices in Bangkok to assist in loan decision-making. The focus is on developing a model to predict condominium prices in Bangkok to appraise loan values, thereby reducing the need to send personnel for on-site price evaluations. Data was collected from the website zmyhome.com using Web Scraping, encompassing 9,041 condominium projects with a total of 16 attributes. The study compares the performance of models between Decision Tree, Random Forest, XGBoost, and LightGBM. The results indicate that the XGBoost model performs the best, with Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R-Squared (R2) values of 8.4135%, 9,444.8167, and 0.9527, respectively. The findings suggest the feasibility of using Machine Learning techniques to forecast condominium prices in Bangkok for loan appraisal purposes.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การศึกษาค้นคว้าอิสระครั้งนี้สำเร็จลุล่วงไปได้ด้วยดีด้วยความกรุณาจากหลายท่าน ผู้ศึกษาขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ฉัฐไชย์ ลีนาวงศ์ อาจารย์ที่ปรึกษาเป็นอย่างสูงที่ได้ให้ทั้งความรู้ คำชี้แนะ คำปรึกษาตลอดจนตรวจสอบชี้แนะแนวทาง และแก้ไขปรับปรุงข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่ตลอดระยะเวลาที่ทำการศึกษาจนการศึกษาครั้งนี้เสร็จสมบูรณ์

ขอกราบขอบพระคุณคณาจารย์คณะวิทยาศาสตร์ สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์ ศูนย์วิเคราะห์ข้อมูลดิจิทัลอัจฉริยะพระจอมเกล้าลาดกระบัง ที่ได้ถ่ายทอดวิชาความรู้อัน เป็นประโยชน์ยิ่งให้แก่ผู้ศึกษา รวมถึงเจ้าหน้าที่สาขาวิชาทุกท่านที่ให้ความช่วยเหลือในการติดต่อ ประสานงานในเรื่องต่าง ๆ

สุดท้ายนี้ ขอกราบขอบพระคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง สถานที่ที่ถ่ายทอดความรู้แก่นักศึกษา

ภควัต อมรวิลาศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อ ..... ก

Abstract ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฉ

สารบัญภาพ ..... ช

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของการศึกษา 2

1.3 ขอบเขตของการศึกษา 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 2

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 3

2.1 สินเชื่อ 3

2.2 การประเมินค่าทรัพย์สิน 5

2.3 แบบจำลอง 7

2.4 การประเมินผล 10

2.5 งานวิจัยที่เกี่ยวข้อง 13

บทที่ 3 วิธีดำเนินงานการศึกษา 15

3.1 แผนผังขั้นตอนการดำเนินงาน 15

3.2 ข้อมูล 16

3.3 การสำรวจข้อมูล 19

3.4 การเตรียมข้อมูล 20

3.5 แบบจำลอง 25

3.6 ประเมินผลแบบจำลอง 27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

บทที่ 4 ผลการศึกษาและการอภิปรายผล.....29

4.1 ผลลัพธ์การสร้างต้นแบบ 29

4.2 เปรียบเทียบประสิทธิภาพ 31

บทที่ 5 สรุปผลการศึกษาและข้อเสนอแนะ 32

5.1 สรุปผลการการศึกษา 32

5.2 ข้อเสนอแนะ..... 33

5.3 ข้อจำกัดในการศึกษา..... 33

เอกสารอ้างอิง 34

ประวัติผู้เขียน..... 35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

3.1 ตัวอย่างข้อมูลคอนโดมิเนียมในกรุงเทพมหานคร 18
3.2 จำนวนข้อมูลที่ไม่สมบูรณ์ของแต่ละแอตทริบิวต์ 24
3.3 พารามิเตอร์ของการจำแนกประเภทแบบต้นไม้ตัดสินใจ 26
3.4 พารามิเตอร์ของการจำแนกประเภทแบบป่าสุ่ม 26
3.5 พารามิเตอร์ของการจำแนกประเภทแบบ XGBoost 26
3.6 พารามิเตอร์ของการจำแนกประเภทแบบ LightGBM 27
4.1 ประสิทธิภาพของต้นแบบที่สร้างจากแบบจำลอง Decision Tree 29
4.2 ประสิทธิภาพของต้นแบบที่สร้างจากแบบจำลอง Random Forest 30
4.3 ประสิทธิภาพของต้นแบบที่สร้างจากแบบจำลอง XGBoost 30
4.4 ประสิทธิภาพของต้นแบบที่สร้างจากแบบจำลอง LightGBM 31
4.5 เปรียบเทียบผลลัพธ์ของแบบจำลองทั้ง 4 แบบ 31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

ภาพที่ **หน้า**

2.1 ต้นไม้ตัดสินใจ 8
2.2 การรวมกลุ่มกันของโครงสร้างต้นไม้ตัดสินใจ 9
2.3 อัลกอริทึม Level-wise 10
2.4 อัลกอริทึม Leaf-wise 10
2.5 ขั้นตอนการวิเคราะห์ความแม่นตรงของตัวแบบทำนาย K-fold cross validation 11
2.6 ตัวอย่างกราฟของต้นแบบการถดถอยเชิงเส้น 12
3.1 แผนผังขั้นตอนการดำเนินการ 15
3.2 ข้อมูลจากเว็บไซต์ zmyhome.com 16
3.3 สัมประสิทธิ์สหสัมพันธ์ของแต่ละแอตทริบิวต์ 19
3.4 ปีที่สร้างเสร็จต่อราคาคอนโดมิเนียมต่อตารางเมตร 20
3.5 พื้นที่ใช้สอยต่อราคาคอนโดมิเนียมต่อตารางเมตร 21
3.6 พื้นที่ใช้สอยต่อราคาคอนโดมิเนียมต่อตารางเมตรตั้งแต่ 0 ถึง 20 ตารางเมตร 21
3.7 ความถี่ของราคาคอนโดมิเนียมต่อตารางเมตร 22
3.8 ความถี่ขอราคาคอนโดมิเนียมต่อตารางเมตรตั้งแต่ 0 ถึง 50,000 บาทต่อตารางเมตร 23
3.9 จำนวนห้องนอนต่อจำนวนห้องน้ำ 25
3.10 ช่วงความถี่ของราคาคอนโดมิเนียมต่อตารางเมตรหลังลบค่าสุดโต่ง 24
3.11 ค่าสหสัมพันธ์ของราคาเฉลี่ยของคอนโดมิเนียม 25
3.12 ตัวอย่างการแบ่งข้อมูลด้วยวิธี K-Fold cross validation 28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

เนื่องจากในปัจจุบันก่อนธนาคารให้สินเชื่อต้องผ่านขั้นตอนการประเมินราคาทรัพย์สินที่ผู้กู้ต้องการซื้อโดนจำเป็นต้องส่งเจ้าหน้าที่ไปประเมินราคาทรัพย์สินที่ผู้กู้ต้องการซื้อ เพื่อให้ธนาคารสามารถนำข้อมูลนี้ไปพิจารณาปล่อยวงเงินสินเชื่อได้ การศึกษาค้นคว้าอิสระนี้จะนำเสนอการพยากรณ์ราคาคอนโดมิเนียมเพื่อประเมินมูลค่าสินเชื่อ ในบทนี้จะกล่าถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ของการศึกษา ขอบเขตของการศึกษา และประโยชน์ที่คาดว่าจะได้รับ โดยมีรายละเอียดดังนี้

1.1 ความเป็นมาและความสำคัญของปัญหา

จากในช่วงไม่กี่ปีที่ผ่านมากรุงเทพมหานครเป็นหนึ่งในเมืองที่มีแหล่งการดำเนินธุรกิจจำนวนมาก ทำให้เกิดการย้ายถิ่นฐานของประชากรทั้งในปริมณฑลและต่างจังหวัด เข้ามาทำงานในกรุงเทพมหานคร ส่งผลให้เกิดความต้องการที่อยู่อาศัยประเภทคอนโดมิเนียมที่เพิ่มสูงขึ้น เมื่อเทียบกับที่อยู่อาศัยในลักษณะพื้นที่ราบซึ่งมีอยู่จำนวนจำกัด และมีราคาสูงโดยเฉพาะอย่างยิ่งในกลุ่มคนเจเนอเรชัน Y ยังมีอัตราความต้องการในการซื้อคอนโดมิเนียมสูงที่สุด เมื่อเทียบกับกลุ่มคนเจเนอเรชันอื่น ๆ เพราะอยู่ในช่วงวัยของการเริ่มต้นทำงานและต้องการที่อยู่อาศัยใกล้ระบบขนส่งมวลชนที่สะดวก การเป็นเจ้าของคอนโดมิเนียมมีหลายวิธี เช่น การซื้อเงินสดหรือการผ่อนชำระ ซึ่งคนเจเนอเรชัน Y นั้นมีแนวโน้มที่จะเลือกการผ่อนชำระมากกว่าการซื้อเงินสด และถ้าหากต้องการผ่อนชำระจะต้องมีการขอสินเชื่อจากธนาคารซึ่งจะต้องผ่านการประเมินมูลค่าคอนโดมิเนียมก่อนเพื่อให้ได้รับวงเงินสินเชื่อ

อย่างไรก็ตาม ในปัจจุบันการกู้สินเชื่อคอนโดมิเนียมต้องผ่านขั้นตอนการประเมินราคาคอนโดมิเนียม หนึ่งในกระบวนการที่สำคัญคือการส่งเจ้าหน้าที่ไปประเมินราคาทรัพย์สินที่ผู้กู้ต้องการซื้อ เพื่อให้ธนาคารสามารถนำข้อมูลนี้มาพิจารณาปล่อยวงเงินสินเชื่อได้ โดยทั่วไปถ้าราคาประเมินสูงเท่าไหร่วงเงินกู้ก็จะสูงขึ้นตามไปด้วย ในทางกลับกัน ถ้าธนาคารประเมินราคาต่ำ วงเงินสินเชื่อที่ได้ก็จะน้อยลง โดยสาเหตุที่ธนาคารต้องทำการประเมินราคาคอนโดมิเนียมก่อน ที่จะพิจารณาวงเงินกู้ ก็เพราะว่าธนาคารจำเป็นที่จะต้องทราบมูลค่าที่แท้จริงของคอโดมิเนียม ว่ามีมูลค่าตามที่ขอวงเงินมาหรือไม่ โดยการประเมินมูลค่าคอนโดมิเนียมสามารถพิจารณาได้จากข้อมูลหลาย ๆ อย่าง เช่น ราคาประเมินของราชการ รูปแบบคอนโดมิเนียม การแต่งห้อง สภาพอาคาร และราคาตลาด เป็นต้น กระบวนการนี้ใช้ทั้งทรัพยากรบุคคล เวลา และค่าใช้จ่ายในการส่งเจ้าหน้าที่ไปประเมินราคา แต่ถ้าหากทางธนาคารเปลี่ยนมาใช้ข้อมูลคุณลักษณะต่าง ๆ ของคอนโดมิเนียมเพื่อใช้ในการหามูลค่าคอนโดมิเนียมก็จะช่วยลดระยะเวลาในการประเมิน รวมถึงช่วยลดค่าใช้จ่ายในการส่งเจ้าหน้าที่ไปประเมินราคาอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้นการศึกษาค้นคว้าอิสระครั้งนี้ จึงได้จัดทำการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine learning) การศึกษาค้นคว้าอิสระนี้ประกอบด้วยหลายขั้นตอน เริ่มตั้งแต่การรวบรวมข้อมูลด้วยเทคนิคการขูดเว็บ (Web Scraping) การสำรวจข้อมูล การเตรียมข้อมูล และการสร้างแบบจำลอง ซึ่งการศึกษาครั้งนี้จะเป็นประโยชน์ทั้งทางธนาคาร และผู้บริโภค ทางธนาคารก็สามารถลดค่าใช้จ่ายในการในการส่งเจ้าหน้าที่ไปประเมินราคา และลดเวลาในการดำเนินงาน ส่วนทางผู้บริโภคก็ได้รับความสะดวกสบาย และความรวดเร็วในการได้รับบริการจากทางธนาคารอีกด้วย

1.2 วัตถุประสงค์ของการศึกษา

วัตถุประสงค์ของการศึกษาค้นคว้าอิสระครั้งนี้ มีวัตถุประสงค์ ดังนี้

1. ศึกษาและระบุปัจจัยที่เกี่ยวข้องกับราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ
2. สร้างแบบจำลองเพื่อพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานคร เพื่อช่วยในการตัดสินใจในการให้สินเชื่อ
3. เปรียบเทียบประสิทธิภาพของแบบจำลองระหว่าง Decision Tree, Random Forest, XGBoost และ LightGBM

1.3 ขอบเขตของการศึกษา

ข้อมูลของการศึกษาค้นคว้าอิสระครั้งนี้ เป็นข้อมูลราคาคอนโดมิเนียมในกรุงเทพมหานครที่รวบรวมข้อมูลจากเว็บไซต์ zmyhome.com ณ วันที่ 1 มีนาคม พ.ศ. 2567 โดยข้อมูลที่ศึกษาเป็นข้อมูลคอนโดมิเนียมในกรุงเทพมหานครเป็นรายห้อง เพื่อศึกษาปัจจัยที่เกี่ยวข้องกับราคาคอนโดมิเนียมและประเมินมูลค่าสินเชื่อ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อปรับปรุงกระบวนการประเมินราคาคอนโดมิเนียมในกรุงเทพมหานครสำหรับการประเมินมูลค่าสินเชื่อให้มีประสิทธิภาพ ความรวดเร็ว และความแม่นยำมาก โดยใช้การพยากรณ์ราคาคอนโดมิเนียมแทนการส่งเจ้าหน้าที่ไปประเมินราคาในสถานที่จริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ ได้ทำการศึกษาเอกสารจากหนังสือ ตำรา วารสาร บทความ และกฎระเบียบต่าง ๆ ตลอดจนผลงานวิจัยทุกประเภทที่เกี่ยวข้อง เพื่อรวบรวมทฤษฎีและแนวคิดต่าง ๆ มาสรุปเป็นกรอบความคิดในโดยเนื้อหาในบทนี้ประกอบด้วยหัวข้อหลัก ๆ คือ สินเชื่อ การประเมินค่าทรัพย์สินม แบบจำลอง การประเมินผล และงานวิจัยที่เกี่ยวข้อง

2.1 สินเชื่อ

การให้สินเชื่อเป็นธุรกรรมหลักที่ส่งผลสำคัญอย่างยิ่งต่อผลกำไรและความเสี่ยงของธนาคารการมีมาตรฐานการให้สินเชื่อที่ไม่เหมาะสม การบริหารพอร์ตสินเชื่อที่ไม่ถูกต้องหรือการไม่ใส่ใจติดตามการเปลี่ยนแปลงของสถานการณ์อย่างเพียงพออาจส่งผลให้คุณภาพสินเชื่อด้อยลง และเป็นเหตุให้ความเสี่ยงของธนาคารเพิ่มสูงขึ้น ดังนั้น ธนาคารจึงต้องมีวิธีการในการให้สินเชื่อที่รัดกุมและเหมาะสม

สุนีย์ ศักรนันทน์ (2534) ได้อธิบายว่าแนวคิดเบื้องต้นที่ใช้ในการวิเคราะห์สินเชื่อประกอบไปด้วยสองส่วน ส่วนแรก คือ วัตถุประสงค์การให้สินเชื่อและนโยบายของผู้ให้สินเชื่อ โดยผู้ให้ สินเชื่อจะระบุวัตถุประสงค์ในการกู้สินเชื่อแต่ละประเภทไว้ อีกทั้งกำหนดนโยบายสำหรับใช้เป็นแนวทางประกอบการให้สินเชื่อเพื่อให้กิจการบรรลุตามวัตถุประสงค์ตามเป้าหมายที่ตั้งไว้ และ ดำเนินการภายในระยะเวลาที่กำหนด หากผู้กู้มีวัตถุประสงค์ในการกู้ที่ไม่สอดคล้องกับวัตถุประสงค์ และนโยบายของผู้ให้สินเชื่อก็มีความเป็นไปได้ที่จะถูกปฏิเสธการให้สินเชื่อดังกล่าว ส่วนที่สอง คือ กฎหมายที่เกี่ยวกับการวิเคราะห์สินเชื่อ เช่น พระราชบัญญัติห้ามเรียกดอกเบี้ยเกินอัตราพระราชบัญญัติธนาคารแห่งประเทศไทยและพระราชบัญญัติดอกเบี้ยให้กูยืมของสถาบันการเงิน เป็นต้น นอกจากนี้ยังมีบทบัญญัติในประมวลกฎหมายแพ่งและพาณิชย์ในการกำหนดความสัมพันธ์ ของเจ้าหนี้ ลูกหนี้ และบุคคลที่เกี่ยวข้องไว้ เช่น บทบัญญัติในเรื่องหนี้ การฝากทรัพย์ การจำนอง และการจำนำ เป็นต้น ซึ่งการตัดสินใจให้สินเชื่อจะต้องเป็นไปตามกฎหมายและบทบัญญัติที่เกี่ยวข้อง

หลักเกณฑ์สำคัญที่สถาบันการเงินนิยมใช้เป็นหลักในการวิเคราะห์สินเชื่อประกอบ ไปด้วยหลักเกณฑ์ C’s policy และหลักเกณฑ์ P’s policy ซึ่งถือเป็นการวิเคราะห์เชิงคุณภาพ (Qualitative) เพื่อดูคุณภาพของผู้ขอสินเชื่อตามหลักวิเคราะห์ (ชนินทร์ พิทยาวิวิธ 2534) โดยมีรายละเอียด ดังนี้

หลักเกณฑ์ C’s policy ประกอบไปด้วย

1) คุณสมบัติของผู้กู้ (Character) เป็นการวิเคราะห์เกี่ยวกับคุณลักษณะภายในของผู้กู้ ซึ่งประกอบไปด้วย พฤติกรรม บุคลิก ความรับผิดชอบและอุปนิสัยส่วนตัว รวมไปถึงสถานภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมรส ความมั่นคงในอาชีพและความซื่อสัตย์สุจริต โดยคุณสมบัติเหล่านี้จะช่วยสะท้อนถึงแนวโน้ม การชำระหนี้ในอนาคต

2) ความสามารถในการชำระหนี้ (Capacity) เป็นปัจจัยที่ใช้ประกอบการพิจารณาความสามารถในการชำระหนี้ของผู้กู้ เช่น รายได้ ซึ่งประกอบไปด้วย เงินเดือน ค่าจ้าง หรือเงินได้อื่น ๆ รวมไปถึงแผนการใช้จ่ายเงิน จำนวนสมาชิกในครอบครัว การศึกษา ประสบการณ์การทำงาน และภาระหนี้สินที่มีอยู่ในปัจจุบัน เป็นต้น

3) เงินทุนหรือส่วนร่วมทุน (Capital) แม้การพิจารณาความสามารถในการชำระหนี้จะช่วยให้ทราบได้ว่าผู้กู้มีความสามารถในการชำระหนี้ได้หรือไม่ แต่เงินทุนหรือส่วนร่วมทุนจะช่วยให้ทราบ ว่าผู้กู้จะสามารถชำระหนี้ได้เป็นจำนวนเงินเท่าใด ยิ่งผู้กู้มีเงินทุนหรือส่วนร่วมทุนมาก ก็จะยิ่งแสดงถึงความมั่นคงทางการเงินของผู้กู้ตามไปด้วย โดยส่วนใหญ่สถาบันการเงินจะพิจารณาจากเงินออมและการถือครองกรรมสิทธิ์ในทรัพย์สินเป็นหลัก

4) หลักประกัน (Collateral) เป็นปัจจัยสำคัญที่ช่วยลดความเสี่ยงในการเกิดหนี้สูญ สินเชื่อที่มีความเสี่ยงในการกู้เงินสูงสุดคือสินเชื่อที่ไม่มีหลักประกัน รองลงมาคือสินเชื่อที่ใช้บุคคลค้ำประกัน โดยสินเชื่อที่ใช้หลักทรัพย์ค้ำประกันถือเป็นสินเชื่อที่มีความเสี่ยงต่ำสุด เช่น ที่ดิน อาคาร และสิ่งปลูกสร้าง การมีหลักประกันในสินเชื่อจะช่วยป้องกันความผันผวนของความสามารถในการชำระหนี้ของลูกหนี้ได้

5) สภาพเศรษฐกิจ (Condition) การเปลี่ยนแปลงในภาวะเศรษฐกิจทั้งระยะสั้นและระยะยาวจะทำให้ผู้กู้มีความสามารถในการชาระหนี้ที่แตกต่างกัน และอาจส่งผลให้เกิดหนี้ด้อยคุณภาพขึ้นได้ เช่น การเปลี่ยนแปลงทางการเมือง การเปลี่ยนแปลงอัตราดอกเบี้ย ภาวะเงินเฟ้อ การว่างงาน การเปลี่ยนแปลงทางสภาพภูมิอากาศหรือการเปลี่ยนแปลงทางเทคโนโลยีที่มีผลกระทบต่อผลผลิต เป็นต้น

หลักเกณฑ์ P’s policy ซึ่งประกอบไปด้วย

1) ปัจจัยตัวบุคคล (People) โดยจะพิจารณาในส่วนความรับผิดชอบของผู้กู้ ความสำเร็จในการทำงาน และความสามารถในการชาระหนี้ มาใช้ประกอบการพิจารณาสินเชื่อ เช่น ข้อมูลเกี่ยวกับประวัติการชำระหนี้ของสถาบันการเงินเดิม ข้อมูลจากสานักงานข้อมูลเครดิตแห่งชาติ เป็นต้น

2) ความมุ่งหมายในการกู้ยืม (Purpose) หรือ การพิจารณาวัตถุประสงค์ในการกู้ยืมเพื่อให้ทราบว่าผู้กู้จะนำเงินไปใช้เพื่อวัตถุประสงค์ใด โดยส่วนใหญ่วัตถุประสงค์ในการกู้ยืมจะประกอบไปด้วย 3 วัตถุประสงค์หลัก คือ การกู้ยืมเงินเพื่อใช้ลงทุนหรือใช้เป็นทุนหมุนเวียนในกิจการ การกู้ยืมเพื่อนำไปชำระหนี้สิน และการกู้ยืมเพื่อใช้ในการอุปโภคบริโภคหรือซื้อสินค้าและบริการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) การชำระหนี้ (Payment) เป็นการพิจารณาความสามารถในการชำระของผู้กู้ว่าควรจะกู้ยืมเป็นจำนวนเงินเท่าใด จำนวนเงินงวดที่ต้องชำระหนี้ และระยะเวลาในการชำระหนี้ที่เหมาะสม เพื่อให้เป็นไปตามเงื่อนไขในการกู้ยืม โดยการพิจารณาจะใช้ข้อมูลลูกหนี้ในอดีต เพื่อนำมาคำนวณความสามารถในการชำระหนี้ในอนาคต

4) หลักการป้องกัน (Protection) เป็นการพิจารณาถึงปัจจัยในการป้องกันความเสี่ยงที่อาจเกิดหนี้ด้อยคุณภาพขึ้นในอนาคต เช่น การพิจารณาในส่วนการถือครองทรัพย์สินของผู้กู้ จำนวนเงินออม คุณสมบัติในการชำระหนี้ของผู้กู้ร่วม และความน่าเชื่อถือของผู้ค้ำประกัน เป็นต้น

5) องค์ประกอบรวม (Perspective) เป็นการพิจารณาองค์ประกอบโดยร่วมของผู้กู้ เช่น ความเสี่ยงที่จะเกิดหนี้สูญ ความยุ่งยากในการเรียกเก็บเงิน ต้นทุน และค่าใช้จ่ายในการฟ้องดำเนินคดี เป็นต้น

2.2 การประเมินค่าทรัพย์สิน

แนวคิดเกี่ยวกับการประเมินราคาทรัพย์สิน (สำนักงานประเมินราคาทรัพย์สิน 2535) ในการประเมินราคาทรัพย์สินนั้นสำนักประเมินราคาทรัพย์สินได้อาศัยอำนาจหน้าที่ตามประมวลกฎหมายที่ดินและระเบียบที่เกี่ยวข้องกับการประเมินราคา เช่น ประมวลกฎหมายที่ดิน ฉบับแก้ไขเพิ่มเติม 19 กันยายน พ.ศ. 2534 มาตรา 104 และมาตรา 105 และระเบียบคณะกรรมการกำหนดราคาประเมินทุนทรัพย์ว่าด้วยหลักเกณฑ์และวิธีการในการกำหนดราคาประเมินทุนทรัพย์ ของอสังหาริมทรัพย์เพื่อเรียกเก็บค่าธรรมเนียมจดทะเบียนสิทธิและนิติกรรม พ.ศ. 2535 นอกจากนี้แล้วการประเมินราคาทุนทรัพย์จะต้องคำนึงถึงราคาตลาดเป็นสำคัญ นั่นคือราคาประเมินจะต้องใกล้เคียงกับราคาตลาดให้มากที่สุด กระบวนการประเมินราคา (ไพโรจน์ซึงศิลป์ 2538) การประเมินราคาที่ดิน มีกระบวนการ 5 ขั้นตอน ดังนี้

ขั้นที่ 1 กำหนดปัญหา

ผู้ประเมินราคาจะต้องเข้าใจถึงปัญหาอย่างชัดเจน โดยสิ่งที่ต้องคำนึงถึง ได้แก่ การระบุตำแหน่งที่ตั้งของทรัพย์สิน ส่วนได้เสียในทรัพย์สิน ลักษณะของการเป็นเจ้าของ ภาระผูกพันกับสถาบันการเงินต่าง ๆ ภาระจำยอมเรื่องสาธารณูปโภคต่าง ๆ ทางเดิน ทางรถยนต์ เป็นต้น รวมทั้งวันที่ที่ทำการประเมิน และวัตถุประสงค์ในการประเมินราคา หรือจุดมุ่งหมายในการประเมินราคาเพื่ออะไร เป็นต้น

ขั้นที่ 2 วางแผนการประเมินราคา

เมื่อกำหนดปัญหาแล้วผู้ประเมินราคาต้องทำการสำรวจเบื้องต้น เพื่อดูลักษณะ ขอบเขต และปริมาณงานที่จะต้องทำซึ่งต้องมีการวางแผนในเรื่องข้อมูลและแหล่งของข้อมูลที่ต้องใช้ในการประเมิน ซึ่งแหละข้อมูลนั้น สามารถให้ได้จากกรมที่ดิน หน่วยงานราชการที่ตั้งอยู่ในท้องถิ่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผู้ใหญ่บ้านหรือกำนัน บริษัทนายหน้า วารสารเกี่ยวกับอสังหาริมทรัพย์ เป็นต้น นอกจากนี้ต้องมีการกำหนดบุคลากรที่ต้องการให้เหมาะสม รวมทั้งเวลาและค่าธรรมเนียมต่าง ๆ

ขั้นที่ 3 การรวบรวมข้อมูลและวิเคราะห์ข้อมูล

ข้อมูลที่ใช้ในการประเมินประกอบด้วย ข้อมูลทั่วไป ได้แก่ ข้อมูลระดับภาค และระดับเมือง แนวโน้มการพัฒนาภูมิภาค และปัจจัยทางเศรษฐกิจ ข้อมูลเกี่ยวกับชุมชน ได้แก่ สภาพการใช้ประโยชน์ที่ดินในชุมชนนั้น การแบ่งเขตตามกฎหมาย ผังเมือง สาธารณูปโภคต่าง ๆ ระบบการคมนาคม โรงเรียน ศูนย์การค้า ฯลฯ ข้อมูลเกี่ยวกับตัวทรัพย์สิน เป็นข้อมูลเกี่ยวกับที่ดิน เช่น ที่ตั้งของทรัพย์สิน เจ้าของกรรมสิทธิ์เนื้อที่ดิน ข้อมูลเกี่ยวกับอาคาร เช่น อายุอาคาร สภาพอาคารและข้อมูลเกี่ยวกับที่ตั้ง เช่น ขนาด รูปร่าง ทำเลที่ตั้งถนน ทางระบายน้ำ ประปา ไฟฟ้า และโทรศัพท์ ซึ่งเป็นปัจจัยภายในของตัวทรัพย์สิน และข้อมูลเฉพาะเกี่ยวกับต้นทุน รายได้และราคาขาย เช่น ต้นทุนการก่อสร้างอาคาร ราคาขาย และข้อมูลที่จำเป็นต่อการประเมินด้วยวิธีคำนวณรายได้

ขั้นที่ 4 วิธีการประเมินราคา

วิธีการประเมินราคา ซึ่งถือเป็นหลักของกระบวนการประเมินราคา โดยทั่วไปมักมีวิธีการประเมิน 3 วิธี คือ

1. การประเมินราคาโดยใช้วิธีเปรียบเทียบราคาตลาด (Market Approach)
2. การประเมินราคาโดยใช้วิธีต้นทุน (Cost Approach)
3. การประเมินราคาโดยวิธีคำนวณจากรายได้ (Income Approach)

ในการประเมินราคา สามารถใช้วิธีการประเมินราคาทั้ง 3 วิธีซึ่งการประเมินนั้น จะต้องประเมินมูลค่า ณ ภาวะปกติ แต่บางครั้งในการประเมินราคาอาจไม่จำเป็นต้องใช้วิธีการประเมินราคาทั้ง 3 วิธี ขึ้นอยู่กับสภาพของทรัพย์สินเป็นหลัก

ขั้นที่ 5 การหาความสัมพันธ์ของวิธีการประเมินราคาทั้ง 3 วิธี

การประเมินราคาขั้นสุดท้ายขั้น ในกระบวนการประเมินราคา คือ การพิจารณาดูความสัมพันธ์ของมูลค่าทรัพย์สินที่ได้ใช้การประเมินราคาทั้ง 3 วิธีดังกล่าว โดยคำนึงถึงวัตถุประสงค์ ในการประเมินราคาประเภทของทรัพย์สิน และข้อจำกัดของข้อมูล ผู้ประเมินต้องใช้ดุลยพินิจ ให้ความสำคัญกับวิธีที่เหมาะสม และน่าเชื่อถือมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 แบบจำลอง

2.3.1 ต้นไม้ตัดสินใจ (Decision Tree) เป็นการนำข้อมูลมาสร้างแบบจำลองมีลักษณะเป็นผังงาน (Flowchart) เหมือนโครงสร้างต้นไม้ เป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ สร้างแบบจาลองขึ้นมาจากข้อมูลที่นำมาใช้เรียนรู้โดยต้นไม้ตัดสินใจสามารถนา มาใช้ในการทำนายค่าต่าง ๆ ได้โดยผลการทำนายจะขึ้นอยู่กับตัวแปรต้น รูปแบบของต้นไม้ประกอบด้วยโหนดแรกสุด ที่เรียกว่าโหนดราก (Root Node) จากโหนดรากก็จะแตกออกเป็นโหนดลูกซึ่งมีกิ่งในการเชื่อม ระหว่างโหนด และโหนดลูกก็จะมีลูกของตัวเองซึ่งที่โหนดระดับสุดท้ายจะเรียกว่าโหนดใบ (Leaf Node) แต่ละโหนดของโหนดรากและโหนดลูกจะแสดงค่าคุณลักษณะ (Attribute) ที่ใช้ทดสอบ ข้อมูล ส่วนโหนดใบจะแสดงกลุ่ม (Class) ที่กำหนดไว้เมื่อมีข้อมูลที่ต้องการทำนายการทำงานจะเริ่มต้นจากโหนดราก ซึ่งจะนำค่าคุณลักษณะต่าง ๆ ของข้อมูลนั้นไปเปรียบเทียบกับคุณลักษณะ ของโหนด และทำการตัดสินใจว่า จะเดินทางไปตามกิ่งใด หลังจากนั้น จะเดินทางผ่านโหนดลูกและ ทำการเปรียบเทียบคุณลักษณะไปเรื่อย ๆ จนกระทั่ง สุดท้ายไปถึงโหนดใบ ก็จะได้กลุ่มที่ถูกกำหนด วิธีการสร้างต้นไม้ตัด สินใจอันดับแรกจะหาคุณลักษณะที่สำคัญที่สุดมาแบ่งข้อมูลโดย คุณลักษณะนี้จะถูกตั้งให้เป็นโหนดรากจากโหนดรากจะสร้างเส้นทางเชื่อมหรือกิ่งไปยังโหนดลูก โดยจำนวนเส้นทางเชื่อมจะเท่ากับจำนวนค่าที่เป็นไปได้ของคุณลักษณะของโหนดรากถ้าโหนดลูก เป็นกลุ่มของข้อมูลที่อยู่ในกลุ่มเดียวกันทั้งหมดให้หยุดการสร้างต้นไม้ แต่ถ้าโหนดลูกมีข้อมูลของหลายกลุ่มปะปนกัน ต้องสร้างโหนดลูกเพื่อจำแนกข้อมูลต่อไป โดยวนกลับไปทำขั้นตอนแรกซ้ำ เพื่อเลือกคุณลักษณะที่สำคัญที่สุดเป็นตัวแบ่งข้อมูลต่อไป สำหรับความสำคัญของคุณลักษณะ สามารถหาได้จากการคำนวณค่าการเพิ่มสารสนเทศ (Information Gain) ต้นไม้ตัดสินใจสร้างกิ่งสาขาจะพิจารณาจากค่าความจริงของคุณลักษณะโดยค่าที่ใช้จะมาจากการคำนวณจากค่าการเพิ่มสารสนเทศการสร้างต้นไม้ตัดสินใจ C4.5 ใช้ค่ามาตรฐานอัตราส่วนเกิน (Gain Ratio) เพื่อเลือกคุณลักษณะที่จะใช้เป็นรากหรือโหนด ถ้าให้ชุดข้อมูล M ประกอบด้วย ค่าที่เป็นไปได้คือ $\{m_1, m_2, ..., m_n\}$ และให้ความน่าจะเป็นที่จะเกิดค่า m_1 มีค่าเท่ากับ $P(m_i)$ จะได้ ว่าค่าเกนสารสนเทศของ M เขียนแทนด้วย I(M) คำนวณได้ดังสมการ

$$I(M) = \sum_{i=1}^{n} -P(m_i)\, log_2 P(m_i) \quad (1)$$

ถ้าให้ข้อมูลสอน คือ T และคุณลักษณะที่เป็นโหนด คือ x และมีค่าทั้งหมดที่เป็นไปได้ n ค่าโหนดปัจจุบันจะแบ่งตัวอย่าง T ออกตามกิ่งเป็น $\{t_1, t_2, ..., t_n\}$ ตามค่าที่เป็นไปได้ของ x ดังนั้น จึงสามารถคำนวณค่าเกนสารสนเทศหลังจากแบ่งตามคุณลักษณะและค่ามาตรฐานเกน (Gain) ของคุณลักษณะ x ได้ดังสมการ

$$I_x(T) = \sum_{i=1}^{n} \frac{|t_i|}{|T|} I(t_i) \quad (2)$$

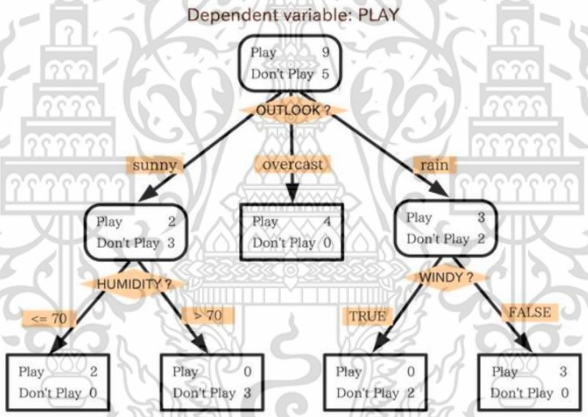
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$Gain(x) = I(T) = I_x(T) \quad (3)$$

จากนั้นคำนวณค่าสารสนเทศของการแบ่งแยก (Split nformation) ของคุณลักษณะแต่ละตัว ถ้าให้ T คือ ชุดของตัวอย่าง เมื่อแบ่งตัวอย่างนี้ตามคุณลักษณะ x จะได้ชุดของตัวอย่างยอยู่ในแต่ละกิ่ง คือ $\{t_1, t_2, ..., t_n\}$ จำนวน n ชุดตามค่าที่เป็นไปได้ในคุณสมบัติ x เมื่อคำนวณค่าสารสนเทศของการแบ่งแยกได้ดังสมการ

$$Split\ Information = \sum_{i=1}^{n} \frac{|t_i|}{|T|} log_2 \frac{|t_i|}{|T|} \quad (4)$$

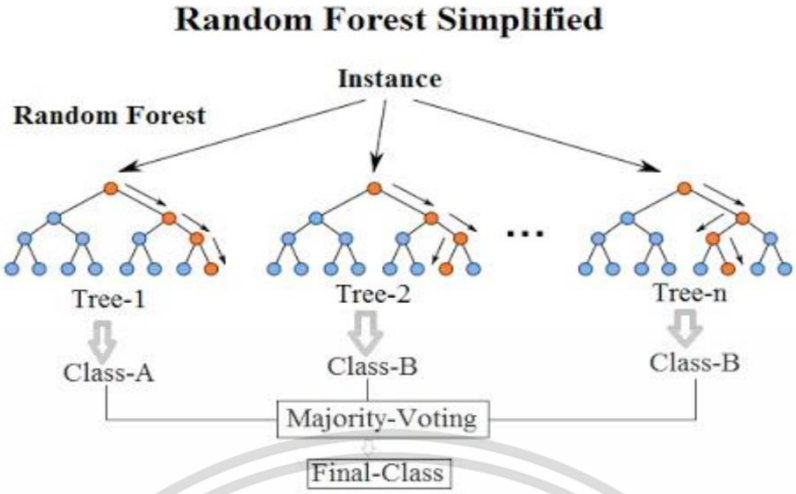
คำนวณค่ามาตรฐานอัตราส่วนเกน (Gain ratio) ได้จาก Gain Ratio = Gain – Split Information ท้ายสุดจึงเลือกค่า Gain ratio สูงสุดเป็นคุณลักษณะเริ่มต้นและเลือกคุณสมบัติถัดไปตามค่า Gain ratio น้อยลงตามลำดับ แสดงในภาพที่ 2.1



ภาพที่ 2.1 ต้นไม้ตัดสินใจ

2.3.2 ป่าสุ่ม (Random Forest) เกิดจากการรวมกลุ่มกันของโครงสร้างต้นไม้ ซึ่งค่าความคลาดเคลื่อนโดยรวมของป่าไม้จะถูกเปลี่ยนให้เป็นค่าลิมิต ทำให้จำนวนของต้นไม้ในป่าเพิ่มขึ้น ค่าความคลาดเคลื่อนโดยรวมจะขึ้นกับความมั่นคง (Strength) ของต้นไม้แต่ละต้น รวมถึงความสัมพันธ์กันระหว่างต้นไม้เหล่านั้น โดยจะใช้วิธีการสุ่มเลือกคุณสมบัติเพื่อการแบ่งแยกโหนด ทำให้ค่าความผิดพลาดลดลงขั้นตอนวิธีนี้จะมีประสิทธิภาพมากเมื่อนำไปใช้วิเคราะห์เกี่ยวกับการประมาณการขนาดใหญ่ เราสามารถสร้างแบบจำลองที่ใช้ต้นไม้หลาย ๆ ต้นในการตัดสินใจเพื่อนำมาประมวลผล ซึ่งมีความแม่นยำสูง สามารถจัดการข้อมูลได้มากและเหมาะสำหรับข้อมูลที่มีความสำคัญ แสดงในภาพที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

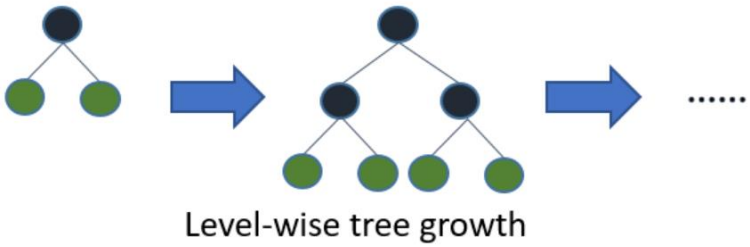


ภาพที่ 2.2 การรวมกลุ่มกันของโครงสร้างต้นไม้ตัดสินใจ

ในขั้นตอนการทำงานของแรนดอมฟอเรส จะทำการจำแนกต้นไม้หลาย ๆ ต้น ซึ่งในต้นไม้แต่ละต้นจะมีการแบ่งเป็นคลาส โดยที่ต้นไม้แต่ละต้นจะถูกสร้างขึ้นจากกลุ่มตัวอย่างที่แตกต่างกันจากกระบวนการของต้นไม้ตัดสินใจ (Tree Classification Algorithm) ถูกสร้างขึ้นจนกลายเป็นป่า (forest) จนกระทั่งวิเคราะห์การตัดสินใจจากต้นไม้แต่ละต้นที่อยู่ในป่า ดังนั้นสามารถสรุป ได้ว่า อัลกอริทึมแรนดอมฟอเรส เป็นอัลกอริทึมประเภทหนึ่งของอัลกอริทึมต้นไม้ตัดสินใจที่มีลักษณะแบบไม่ตัดแต่งกิ่ง (Unpruned) หรือต้นไม้ถดถอย (Regression Trees) ซึ่งถูกสร้างจากการนำข้อมูลฝึกสอนไปสุ่มเลือกตัวอย่างข้อมูลและคุณลักษณะข้อมูลแล้วนำมาสร้างเป็นต้นไม้ตัดสินใจ ซึ่งมีตัวอย่างส่วนหนึ่งที่ไม่ถูกเลือกจะถูกนำมาใช้ในการทดสอบต้นไม้ตัดสินใจ เรียกข้อมูลส่วนนี้ว่า Out-of-Bag (OOB) ซึ่งวิธีการนี้เรียกว่า Bagging ผลลัพธ์ที่ได้อย่างอิสระจากต้นไม้ตัดสินใจแต่ละต้นถูกนำมาคิดเป็นผลการโหวตที่มากที่สุด อัลกอริทึมแรนดอมฟอเรสไม่จำเป็นต้องมีข้อมูลทดสอบเพื่อประมาณความผิดพลาดเพราะข้อมูล Out-of-Bag นั้นถูกนำมาใช้ทดสอบต้นไม้ตัดสินใจ

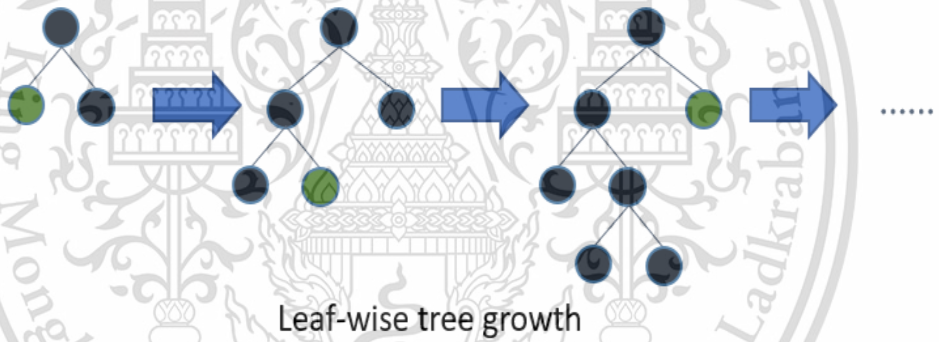
2.3.3 Extreme Gradient Boosting (XGBoost) เป็นการเรียนรู้หลายแบบร่วมกันที่นำเอาต้นไม้การตัดสินใจหลาย ๆ ต้นที่มีพื้นฐานจากอัลกอริทึมกาเดียนบูทติ้งที่ออกแบบมาให้สามารถปรับขนาดได้สูงวัตถุประสงค์เพื่อลดฟังก์ชันการสูญเสียให้น้อยที่สุด โดยจะเรียนรู้จากความผิดพลาดก่อนหน้าทำให้มีความถูกต้อง ความแม่นยำมากขึ้น โดย XGBoost จะใช้อัลกอริทึม Level-wise ที่เป็นการค้นหาตามแนวลึก (Depth-first) แสดงในภาพที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.3 อัลกอริทึม Level-wise

2.3.4 Light Gradient Boosting Machine (LightGBM) เป็นเฟรมเวิร์กที่มีประสิทธิภาพสูงของการเรียนรู้หลายแบบร่วมกันโดยใช้ต้นไม้ตัดสินใจ LightGBM จะใช้อัลกอริทึม Leaf-wise คือเป็นการค้นหาตามค่าดีสุดดีที่สุดก่อน (Best-first) แสดงในภาพที่ 2.4 เพื่อลดฟังก์ชันการสูญเสียได้มากกว่าแบบ Level-wise ที่เป็นการค้นหาตามแนวลึก (Depth-first) ซึ่งเป็นวิธีที่อยู่ในต้นไม้ตัดสินใจทั่วไปทำให้อัลกอริทึม Leaf-wise มีความถูกต้อง ความแม่นยำ และความเร็วที่มากกว่าอย่างไรก็ตามอัลกอริทึม Leaf-wise ก็มีโอกาสที่จะเกิด Overfitting ได้เมื่อใช้กับข้อมูลที่มีขนาดเล็ก



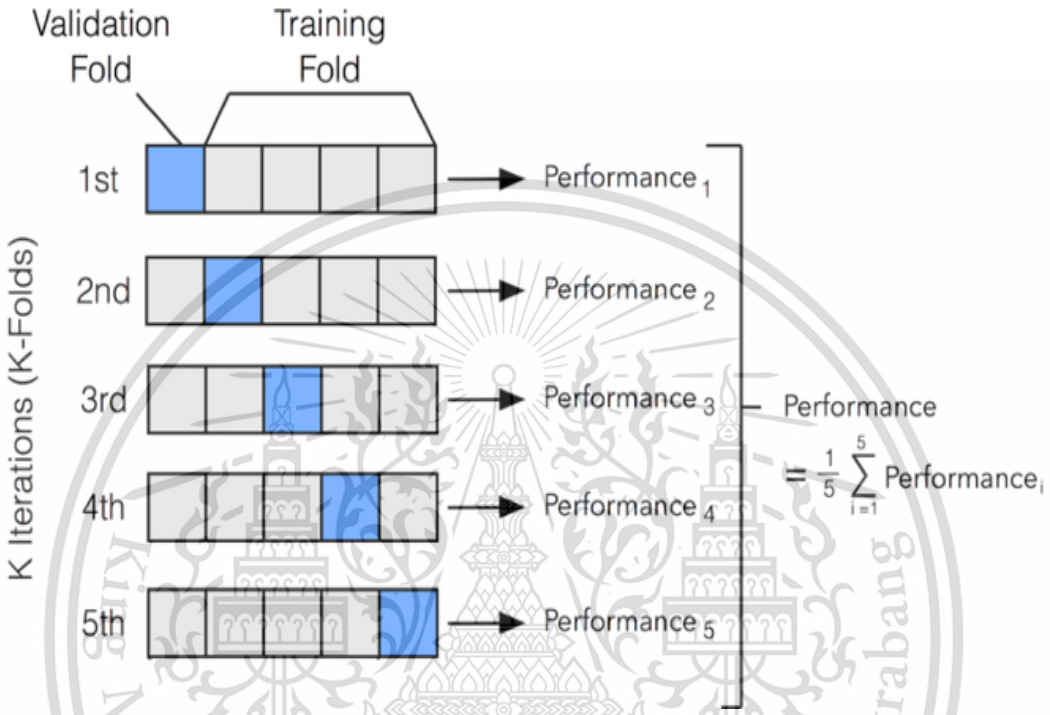
ภาพที่ 2.4 อัลกอริทึม Leaf-wise

2.4 การประเมินผล

2.4.1 การวิเคราะห์ความแม่นตรงของตัวแบบทำนาย K-fold Cross Validation เป็นการวิเคราะห์ความแม่นตรงของตัวแบบทำนาย K-fold cross validation วิธีการตรวจสอบค่าความผิดพลาดในการคาดการณ์ของตัวแบบทำนาย โดยการแบ่งข้อมูลออกเป็นกลุ่มจำนวน K กลุ่ม (K-fold) ในตอนแรกเลือกข้อมูลกลุ่มที่ 1 เป็นข้อมูลชุดทดสอบ และข้อมูลชุดที่เหลือจะเป็นข้อมูลชุดสอน นำข้อมูลไปจัดหมวดหมู่ จากนั้นจะสลับข้อมูลกลุ่มที่ 2 มาเป็นชุดทดสอบและข้อมูลกลุ่มอื่น ๆ ที่เหลือเป็นชุดทดสอบ สลับอย่างนี้ไปเรื่อย ๆ จนครบ K กลุ่ม ในขั้นตอนสุดท้ายจะหาค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐานของค่าความถูกต้องในแต่ละกลุ่ม โดยวิธีการนี้ข้อมูลทุกตัวจะได้เป็นทั้งชุดทดสอบ และชุดสอน ซึ่งจะดีกว่าการแบ่งข้อมูลไว้เป็น 2 ส่วนคือ Training Data กับ Testing Data ที่ทำให้ไม่ทราบได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ว่าข้อมูลใดคือ ข้อมูลที่ควรจะนำมาเป็น Training Data แต่ Cross Validation จะสามารถหา Training Data ที่ดีที่สุดโดยการแบ่งข้อมูลและนำทุกส่วนมาเป็น Training Data และยังสามารถใช้เปรียบเทียบได้อีกว่าควรใช้วิธีไหนที่เหมาะสมที่สุดในการสร้างโมเดล (รัชพล กลัดชื่น และ จรัญ แสนราช 2561) แสดงในภาพที่ 2.5



ภาพที่ 2.5 ขั้นตอนการวิเคราะห์ความแม่นตรงของตัวแบบทำนาย K-fold cross validation

2.4.2 การวัดประสิทธิภาพของแบบจำลอง

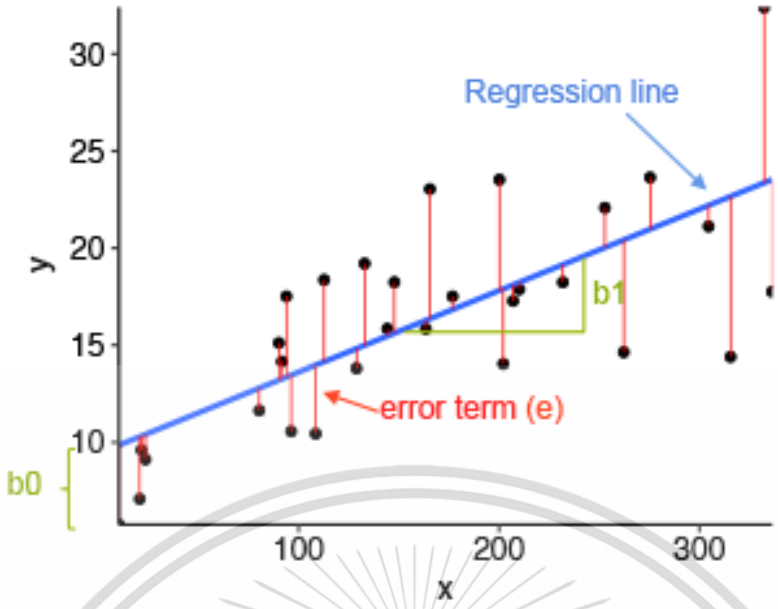
การสร้างต้นแบบ การพยากรณ์ราคาคอนโดมิเนียมจำเป็นต้องมีการคัดเลือกต้นแบบที่ดีที่สุดไม่ว่าจะเป็นในการสร้างต้นแบบเดียวกันหรือระหว่างต้นแบบอื่น ๆ ซึ่งในปัญหาการถดถอยจะทำการสร้างต้นแบบเพื่อคาดการณ์ค่าที่ต้องการจากข้อมูล โดยสามารถเขียนสมการเพื่อคำนวณค่าความผิดพลาดของต้นแบบได้ดังนี้

$$Error = actual - prediction \quad (5)$$

โดยที่ $Error$ คือ ค่าความผิดพลาดของต้นแบบ
$prediction$ คือ ค่าที่คาดการณ์ได้จากต้นแบบ
$actual$ คือ ค่าที่ใส่ให้กับผู้เรียนรู้

ในการเรียนรู้ของเครื่อง เรียกสมการนี้ว่า Loss function โดยเป้าหมายของการสร้างต้นแบบ คือ การทำให้ตัวแปรนี้มีค่าน้อยที่สุด ยกตัวอย่างกราฟของต้นแบบการถดถอยเชิงเส้น ดังแสดงดังภาพที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 2.6 ตัวอย่างกราฟของต้นแบบการถดถอยเชิงเส้น

จากภาพที่ 2.6 แสดงถึงกราฟของต้นแบบการถดถอยเชิงเส้นที่คาดการณ์ค่า y ด้วยตัวแปร x จุดสีดำ คือ ค่าที่ใส่ให้กับผู้เรียนรู้ เส้นสีน้ำเงิน คือ ค่าคาดการณ์ที่ได้จากต้นแบบการถดถอยเชิงเส้น และเส้นสีแดง คือ ค่าผิดพลาดที่ต้นแบบต้องการทำให้มีค่าน้อยที่สุด

ในการคำนวณค่าความผิดพลาดจากต้นแบบ จากกราฟจะเห็นได้ว่าค่าผิดพลาดที่ได้จะมีทั้งค่าบวก และค่าลบ ซึ่งถ้าคำนวณโดยปกติอาจจะทำให้เกิดการหักล้างกันเอง จึงต้องมีการใช้เมตริกต่าง ๆ ในการช่วยคำนวณค่าผิดพลาดของต้นแบบโดยการศึกษาในครั้งนี้จะใช้เมตริก ดังนี้

1) Mean Absolute Percentage Error (MAPE)

เมตริก MAPE เป็นเมตริกที่จะบ่งบอกถึงความแตกต่างระหว่างค่าเฉลี่ย และค่าที่คาดการณ์ได้ในรูปแบบร้อยละโดยสามารถเขียนสมการคำนวณเมตริก MAPE ได้ดังนี้

$$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\frac{|Error|}{|actual|}\right) * 100 \quad (6)$$

2) Root Mean Square Error (RMSE)

เมตริก RMSE เป็นเมตริกที่เป็นรากที่สองของเมตริก MSE เพื่อให้ได้ค่าที่มีหน่วยเดียวกับข้อมูลที่ใส่ให้ผู้เรียนรู้ ทำให้สามารถตีความประสิทธิภาพจากเมตริกได้ง่ายขึ้น โดยสามารถเขียนสมการการคำนวณเมตริก RMSE ได้ดังนี้

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Error)^2} \quad (7)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) R-Squared (R^2)

เมตริก R^2 เป็นเมตริกที่บ่งบอกถึงค่าความแปรปรวนของต้นแบบที่สร้างขึ้นว่าสามารถอธิบายได้จากความแปรปรวนที่เกิดขึ้นทั้งหมดเป็นสัดส่วนเท่าใด โดยสามารถเขียนสมการการคำนวณเมตริก R^2 ได้ดังนี้

$$R^2 = \left(\frac{\sum_{i=1}^{n}(Error)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right) \quad (8)$$

โดยที่ $y_1, y_2, \ldots, y_n$ คือ ค่าที่ใส่ให้กับผู้เรียนรู้

$\bar{y}$ คือ ค่าที่ใส่ให้กับผู้เรียนรู้

2.5 งานวิจัยที่เกี่ยวข้อง

Shuzlina Abdul Rahman (2563) ศึกษาเรื่อง Advanced Machine Learning Algorithms for House Price Prediction: Case Study in Kuala Lumpur ทำนายราคาบ้านโดยอ้างอิงจากข้อมูลที่อยู่อาศัยของกรุงกัวลาลัมเปอร์ โดยใช้แบบจำลองทั้งหมด 2 วิธี ได้แก่ Light Gradient Boosting Machine (LightGBM) และ Extreme Gradient Boosting (XGBoost) มาเปรียบเทียบกับ ผลลัพธ์แสดงให้เห็นว่าโมเดล XGBoost จากนั้นจากประเมินประสิทธิภาพของแบบจำลองโดยใช้ Root Mean Squared Error (RMSE) เท่ากับ 0.197 และ 0.210 ตามลำดับ นอกจากนี้การวิจัยในอนาคตอาจพิจารณาถึงสถานที่อื่น ๆ ในการศึกษา เนื่องจากการศึกษานี้มุ่งเน้นเฉพาะสถานที่ในภูมิภาคกรุงกัวลาลัมเปอร์ การวิจัยในอนาคตอาจขยายพื้นที่การวิจัยนี้ออกไปทั้งประเทศมาเลเซียและการศึกษานี้ยังสามารถช่วยในการทำนายราคาอสังหาริมทรัพย์อื่น ๆ ได้ในอนาคต

Zhou Yichen (2560) ศึกษาเรื่อง Housing Sale Price Prediction Using Machine Learning Algorithms วัตถุประสงค์ของงานวิจัยนี้คือการสร้างโมเดลเพื่อทำนายราคาขายบ้านและค้นหาปัจจัยสำคัญที่ส่งผลต่อราคาขายบ้าน โดยใช้แบบจำลองทั้งหมด 4 วิธี ได้แก่ linear regression, lasso regression, Random Forest และ XGboost สำหรับแบบจำลอง linear regression ไม่เป็นไปตามสมมติฐานของความเท่าเทียมกันของความแปรปรวน ดังนั้นเราจึงไม่สามารถใช้โมเดลเชิงเส้นได้ เพื่อแก้ปัญหานี้จึงใช้แบบจำลอง lasso regression แต่ค่า RMSE และ R^2 ได้ผลลัพธ์ไม่ค่อยดี ส่วนแบบจำลอง Random Forest ไดค่า RMSE ของแบบจำลองนี้ค่อนข้างต่ำทั้งในข้อมูลสำหรับสร้างต้นแบบ (train set) และข้อมูลสำหรับทดสอบประสิทธิภาพของต้นแบบ (test set) และค่า R^2 ใน train set ได้ผลลัพธ์ดี แต่ใน test set ค่อนข้างต่ำ ซึ่งอาจแสดงให้เห็นว่าโมเดล Random Forest มีปัญหา overfitting สุดท้ายแบบจำลอง XGboost ให้ผลลัพธ์ดีที่สุดได้ค่า RMSE และ R^2 เท่ากับ 0.0523 และ 0.8673 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Quang Truong, Minh Nguyen, Hy Dang, and Bo Mei (2562) ศึกษาเรื่อง Housing price prediction via Improved machine learning techniques โดยใช้ข้อมูลที่อยู่อาศัยในกรุงปักกิ่งที่มีข้อมูลมากกว่า 300,000 รายการ และมีตัวแปรทั้งหมด 26 ตัวแปร ซึ่งข้อมูลที่นำมาใช้มีข้อมูลสูญหายมากกว่า 50% และใช้แบบจำลอง Random Forest, XGBoost, LightGBM, Hybrid Regression และ Stacked Generalization Regression และวัดประสิทธิภาพด้วยวิธี Root Mean Squared Logarithmic Error (RMSLE) ผลการศึกษา พบว่าแบบจำลองที่ให้ค่า RMSLE ที่ดีที่สุดของข้อมูล train set ได้แก่ Random Forest และ Hybrid Regression จึงสรุปว่าแบบจำลองที่ดีที่สุดใน test set คือ Stacked Generalization Regression แต่ ในอนาคตผู้วิจัยจะหาปัจจัยที่ทำให้แบบจำลองมีประสิทธิภาพและความเร็วมากขึ้น

พสธร วิทยาปรีชาพล (2563) ศึกษาเรื่องการประเมินราคาเสนอขายห้องชุดด้วย Deep Neural Network และ K-Means Clustering Algorithm โดยข้อมูลที่นำมาศึกษาได้นำมาจาก ZmyHome ซึ่งประกอบไปด้วยข้อมูลจานวน 11,062 ชุด ปัจจัยที่นำมาศึกษาแบ่งเป็นปัจจัยด้านตัวโครงการ (Internal) ปัจจัยด้านทำเล (Location) และปัจจัยด้านสภาพแวดล้อม (Neighborhood) สำหรับเครื่องมือในการวิเคราะห์ผู้ได้ใช้ Gap Statistics สาหรับการหาค่า K ที่เหมาะสมของ K-Means และใช้ Correlation Matrix กับ Loss Function ในการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ของแต่ละตัวแปรกับความผิดพลาดของโมเดลตามลำดับ สำหรับความแม่นยาในการทานายราคาได้มีการเปรียบเทียบระหว่างการคัดเลือก คอนโดมิเนียมด้วย K-Means และระยะทาง พบว่า ทั้งสองวิธีไม่ได้มีความแม่นยำในการทำนายราคา เสนอขายแตกต่างกันอย่างมีนัยสำคัญ แต่การทำนายราคาโดยใช้ K-Means ในการคัดเลือกคอนโดมิเนียมมีอัตราการทำนายราคาสำเร็จที่มากกว่า

Ali Alshamsi (2565) ศึกษาเรื่อง Prediction of Dubai Apartment Prices Using Machine Learning วัตถุประสงค์เพื่อพัฒนาแบบจำลอง Machine Learning ที่สามารถทำนายราคาในอนาคตของอพาร์ทเม้นท์ในดูไบได้ โดยใช้แบบจำลองทั้งหมด 6 วิธี ได้แก่ Linear Regression, Decision Tree, Random Forest, K Nearest Neighbors, Support Vector Regression และ Boosted Regression Model มาเปรียบเทียบกัน และวัดประสิทธิภาพด้วย R^2 และได้ค่าดังนี้ 0.46, 0.55, 0.73, 0.21, 0.30 และ 0.68 ตามลพดับ โดยแบบจำลองที่ให้ผลลัพธ์ดีที่สุดคือ Random Forest ซึ่งให้ค่า R^2 สูงที่สุด

Jumpol Polvichai (2566) ศึกษาเรื่อง Predict Condominium Prices in Bangkok Based on Ensemble Learning Algorithm with various factors โดยดึงข้อมูลเกี่ยวกับคอนโดมิเนียมทั้งหมดที่แสดงอยู่ในเว็บไซต์อสังหาริมทรัพย์ยอดนิยมของไทย ข้อมูลที่รวบรวมครอบคลุมปีการประเมิน ตั้งแต่ปี 2008 ถึง 2019 โดยใช้แบบจำลองทั้งหมด 4 วิธี ได้แก่ Linear Regression, Random Forest, Gradient Boosting Regressor และ XGBoost จากนั้นจากประเมินประสิทธิภาพของแบบจำลองโดยใช้ Mean Absolute Percentage Error (MAPE) โดยโมเดลที่ให้ค่า MAPE ต่ำที่สุดคือ XGBoost รองลงมาคือ Gradient Boosting Regressor, Random Forest และ Linear Regression ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

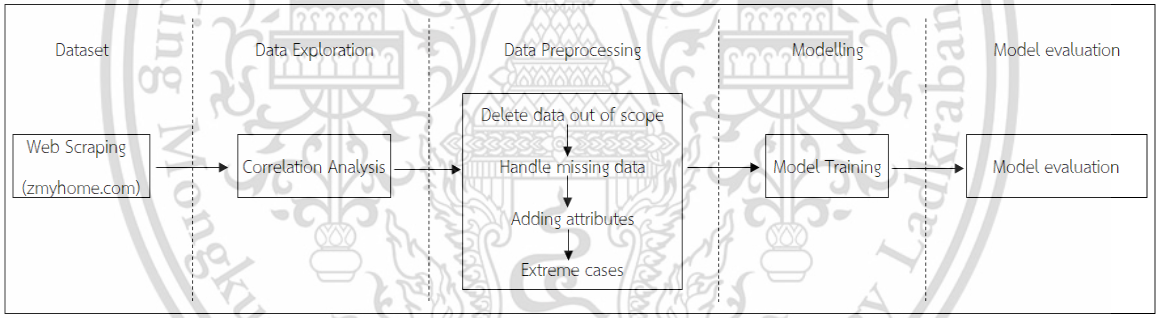
บทที่ 3
วิธีดำเนินงานการศึกษา

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ โดยเนื้อหาในบทนี้ประกอบด้วยหัวข้อหลัก ๆ คือ แผนผังขั้นตอนการดำเนินงาน ข้อมูล การเตรียมข้อมูล แบบจำลอง และผลลัพธ์และการอภิปราย

3.1 แผนผังขั้นตอนการดำเนินงาน

สามารถเขียนสรุปเป็นแผนขั้นตอนการดำเนินงานได้ ดังภาพที่ 3.1 ซึ่งประกอบด้วย

1. ข้อมูล
2. การสำรวจข้อมูล
3. การเตรียมข้อมูล
4. แบบจำลอง
5. ประเมินผลแบบจำลอง



ภาพที่ 3.1 แผนผังขั้นตอนการดำเนินการ

3.2 ข้อมูล

ในการศึกษาครั้งนี้ทำการเก็บรวบรวมข้อมูลจากเว็บไซต์ zmyhome.com ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมข้อมูลราคาอสังหาริมทรัพย์ทั้งบ้านและคอนโดมิเนียม ซึ่งมีปัจจัยที่ใช้ในการศึกษาครบถ้วนมากที่สุดเมื่อเปรียบเทียบกับเว็บไซต์ที่เกี่ยวกับอสังหาริมทรัพย์อื่น ๆ ในประเทศไทย ในการเก็บรวบรวมใช้เทคนิค Web Scraping เพื่อรวบรวมข้อมูลที่เกี่ยวกับ ทำเลที่ตั้ง ลักษณะโครงการ และราคาของคอนโดมิเนียมในกรุงเทพหานคร แสดงในภาพที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 3.2 ข้อมูลจากเว็บไซต์ zmyhome.com

3.4.1 ข้อมูลคอนโดมิเนียมในกรุงเทพมหานคร

ในการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ ข้อมูลนี้เก็บรวบรวมจากเว็บไซต์ zmyhome.com โดยเก็บข้อมูลแต่ละห้องแยกจากกัน เนื่องจากราคาของแต่ละห้องสามารถแตกต่างกันได้แม้ว่าจะอยู่ในคอนโดมินียมเดียวกัน โดยทำการเก็บจำนวนทั้งสิ้น 9,041 ห้อง แสดงตัวอย่างในตารางที่ 3.1 ซึ่งประกอบด้วยข้อมูล ดังนี้

- ชื่อของโครงการคอนโดมิเนียม (Condominium name) ข้อมูลชื่อของโครงการคอนโดมิเนียมนั้น ๆ
- พื้นที่ตั้งของโครงการ (Condo area) ข้อมูลเขตที่โครงการคอนโดมิเนียมตั้งอยู่
- จังหวัด (Province) ข้อมูลจังหวัดที่โครงการคอนโดมิเนียมตั้งอยู่
- ถนน (Road) ข้อมูลถนนแถวที่โครงการคอนโดมิเนียมตั้งอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ปีที่สร้างเสร็จ (Year completed) ข้อมูลปีที่โครงการคอนโดมิเนียมสร้างเสร็จ
- พื้นที่โครงการ (Project area) ข้อมูลพื้นที่โครงการทั้งหมด หน่วยเป็นตารางเมตร
- จำนวนชั้นทั้งหมด (Total of floor) ข้อมูลชั้นสูงสุดของโครงการคอนโดมิเนียม
- ชั้นที่อยู่อาศัย (Floor) ข้อมูลชั้นของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- ราคาคอนโดมิเนียมต่อตารางเมตร (Price per square meter) ข้อมูลราคาต่อตารางเมตรของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- จำนวนห้องนอน (bedrooms) ข้อมูลจำนวนห้องนอนทั้งหมดของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- จำนวนห้องน้ำ (bathrooms) ข้อมูลจำนวนห้องน้ำทั้งหมดของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- พื้นที่ใช้สอย (Floor space) ข้อมูลพื้นที่ใช้สอยทั้งหมดของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- ยูนิตทั้งหมด (Total of units) ข้อมูลจำนวนห้องที่มีทั้งหมดในโครงการคอนโดมิเนียม
- ค่าส่วนกลาง (Common area fee) ข้อมูลค่าส่วนกลางของห้องนั้น ๆ ในโครงการคอนโดมิเนียม
- พื้นที่จอดรถ (Parking space) ข้อมูลพื้นที่ ที่สามารถจอดรถได้ทั้งหมดของโครงการคอนโดมิเนียม
- ระยะห่างจากรถไฟฟ้า (Distance from the BTS) ข้อมูลระยะทางจากโครงการคอนโดมิเนียมถึงรถไฟฟ้าที่ใกล้ที่สุด หน่วยเป็นกิโลเมตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 ตัวอย่างข้อมูลคอนโดมิเนียมในกรุงเทพมหานคร

Condominium name	Condo area	Province	Road	Year completed	Project area	Floor	Total of floor
ศุภาลัย พรีมา ริวา	ยานนาวา	กรุงเทพมหานคร	ถนนพระรามที่ 3	2558	800	34	47
ศุภาลัย เอลีท พญาไท	ราชเทวี	กรุงเทพมหานคร	ถนนศรีอยุธยา	2561	4013	23	31
เอลลิโอ เดล เนสท์	บางนา	กรุงเทพมหานคร	ถนนสุขุมวิท 103	2563	2100.7	9	35
ศุภาลัยพรีเมียร์เจริญนคร	คลองสาน	กรุงเทพมหานคร	ถนนสมเด็จเจ้าพระยา	2564	1911	14	26
ยู เกษตรนวมินทร์	บางเขน	กรุงเทพมหานคร	ถนนลาดปลาเค้า	2560	800	4	8
ลิปป์ ลาดพร้าว 20	จตุจักร	กรุงเทพมหานคร	ถนนลาดพร้าว	2557	3926	31	60

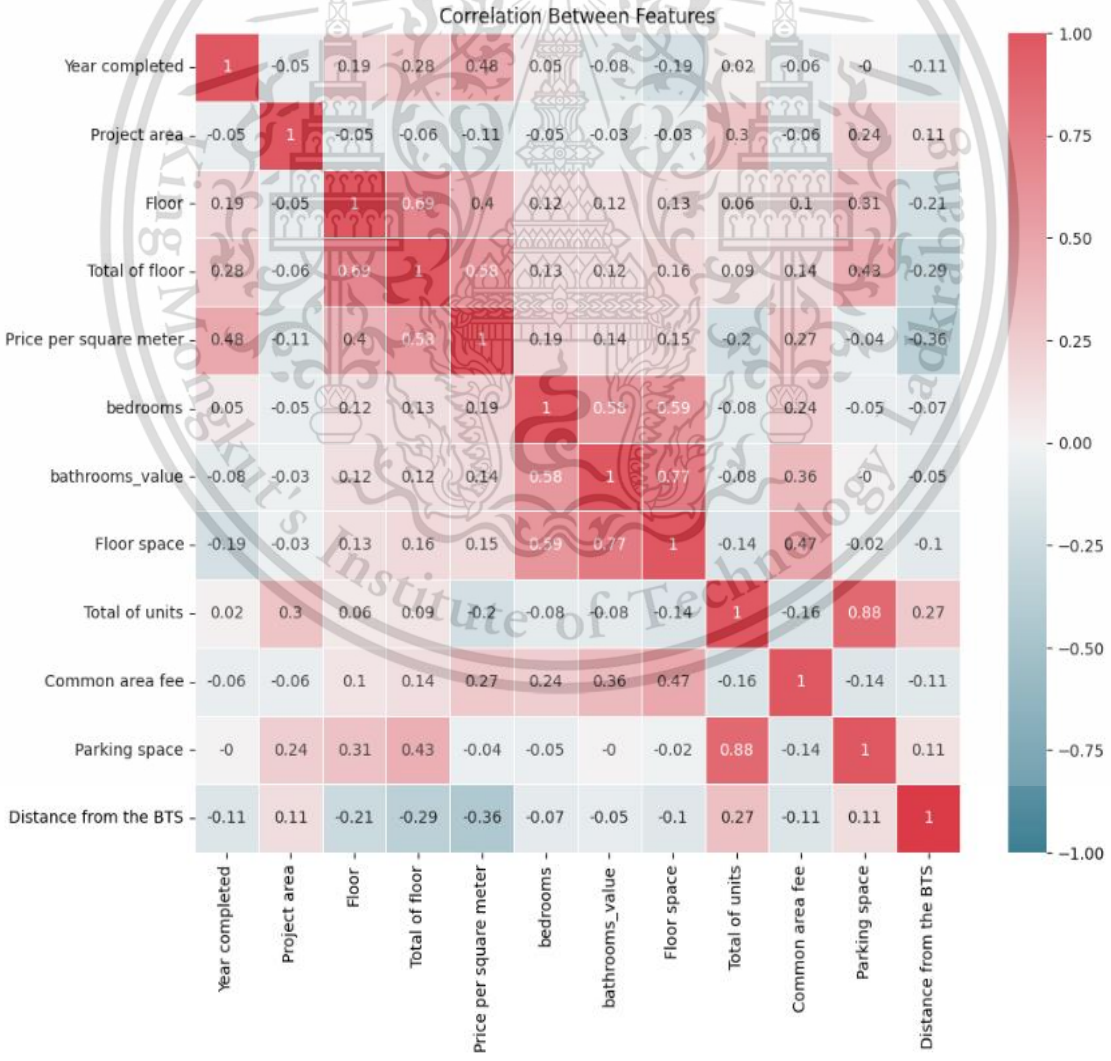
ตารางที่ 3.1 (ต่อ)

Price per square meter	bedrooms	bathrooms	Floor space	Total of units	Common area fee	Parking space	Distance from the BTS
84921	1	1	50.81	69	1270	40	0.9
14787	2	1	44	285	1100	102	0.8
95769	0	1	36	446	1296	219	0.7
137779	1	1	61	285	1525	102	0.8
61132	0	1	21.5	1959	538	550	1.1
10000	0	1	30.3	322	1182	121	0.5

3.3 การสำรวจข้อมูล

ในการศึกษาครั้งนี้ ได้ทำการศึกษาประสิทธิภาพในการสนับสนุนการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครของแต่ละแอตทริบิวต์ผ่านขั้นตอนการสำรวจข้อมูล พบว่าแต่ละแอตทริบิวต์มีระดับนัยสำคัญในการพยากรณ์ผลลัพธ์ที่แตกต่างกัน ยกตัวอย่างจากการสำรวจข้อมูล เช่น คอนโดมิเนียมที่มีจำนวนชั้นมากกว่า (High Rise) มักจะมีราคาที่สูงกว่าคอนโดมิเนียมที่มีจานวนชั้นน้อย (Low Rise)

นอกเหนือจากนี้แล้วการวิเคราะห์สหสัมพันธ์ (Correlation analysis) ของแอตทริบิวต์ทั้งหมด จากภาพที่ 3.3 พบว่าจำนวนชั้นทั้งหมด ปีที่สร้างเสร็จ ชั้นที่อยู่อาศัย และระยะห่างจากรถไฟฟ้ามีนัยสำคัญต่อการพยากรณ์ข้อมูลสูง แต่แอตทริบิวต์ที่มีค่าสหสัมพันธ์น้อยนั้นก็คือ พื้นที่จอดรถทั้งหมดของคอนโดมิเนียม ซึ่งค่าเท่ากับ -0.04 แสดงว่าพื้นที่จอดรถทั้งหมดของคอนโดมิเนียมส่งผลน้อยต่อราคาคอนโดมิเนียมต่อตารางเมตร ซึ่งพิจารณานำข้อมูลพื้นที่จอดรถทั้งหมดของคอนโดมิเนียมออกจากตัวแบบพยากรณ์



ภาพที่ 3.3 สัมประสิทธิ์สหสัมพันธ์ของแต่ละแอตทริบิวต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

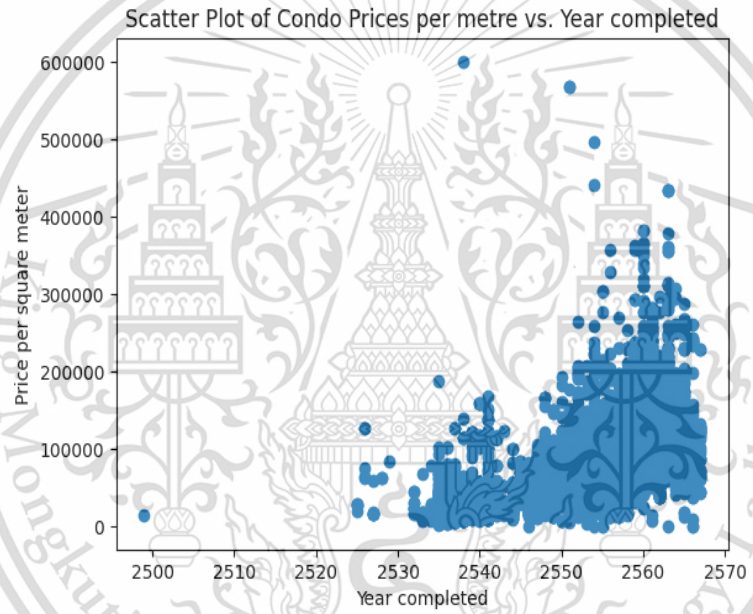
3.4 การเตรียมข้อมูล

จากการเก็บรวบรวมข้อมูลราคาคอนโดมิเนียมในกรุงเทพมหานครจากเว็บไซต์ zmyhome.com ด้วยเทคนิค Web Scraping พบว่ามีข้อมูลที่ผิดปกติหลายอย่าง จึงทำการประมวลผล และปรับปรุงข้อมูล แบ่งออกเป็น 4 ส่วน ดังนี้

3.4.1 ลบข้อมูลที่อยู่นอกขอบเขตของการศึกษา

จากการตรวจสอบข้อมูลที่รวบรวม พบข้อมูลที่อยู่นอกขอบเขตของการศึกษา ดังนี้

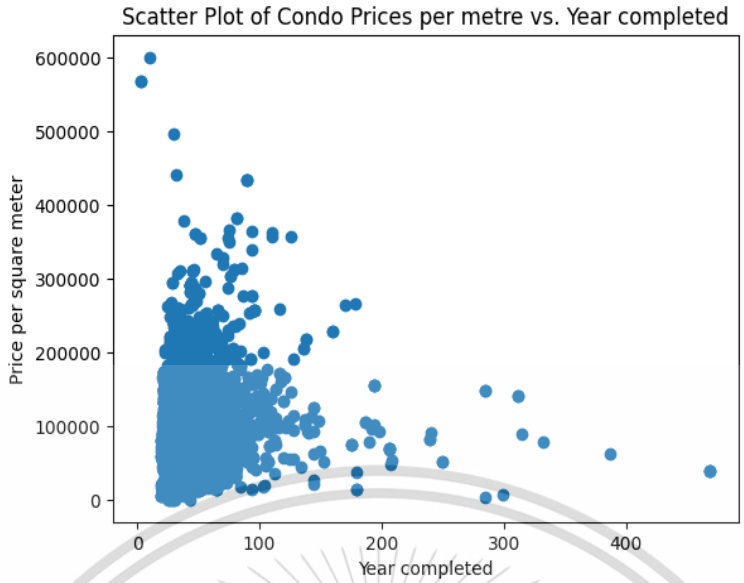
1) ปีที่คอนโดนิเนียมสร้างเสร็จ พบว่ามีข้อมูลปีที่สร้างเสร็จบางส่วนที่เก่าเกินไป ซึ่งอยู่นอกขอบเขตของการศึกษา จึงตัดสินใจลบข้อมูลปีที่สร้างเสร็จที่น้อยกว่าปี พ.ศ. 2530 ออกจากข้อมูลทั้งหมด 21 ห้อง แสดงในภาพที่ 3.4



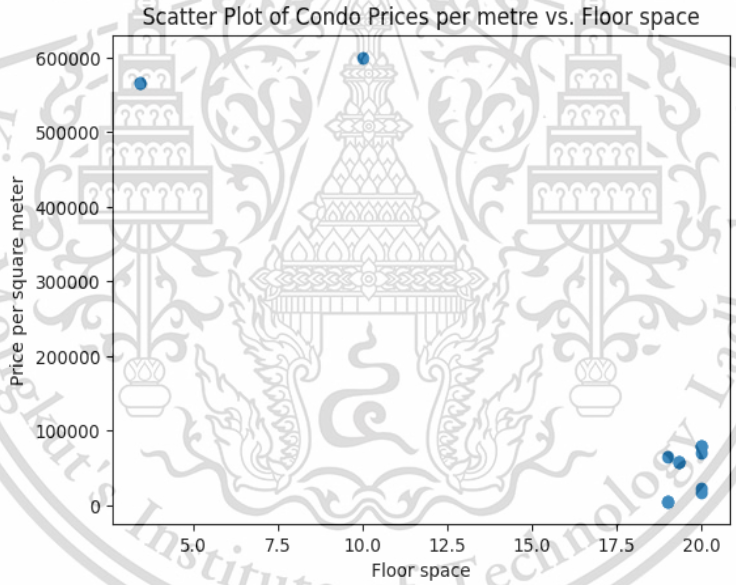
ภาพที่ 3.4 ปีที่สร้างเสร็จต่อราคาคอนโดมิเนียมต่อตารางเมตร

2) พื้นที่ใช้สอย พบว่ามีพื้นที่ใช้สอยบางส่วนที่น้อยเกินไป ซึ่งอยู่นอกขอบเขตของการศึกษา จึงตัดสินใจลบข้อมูลพื้นที่ใช้สอยที่มีค่าน้อยกว่า 18 ตารางเมตร ออกจากข้อมูลทั้งหมด 3 ห้อง แสดงในภาพที่ 3.5 และภาพที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



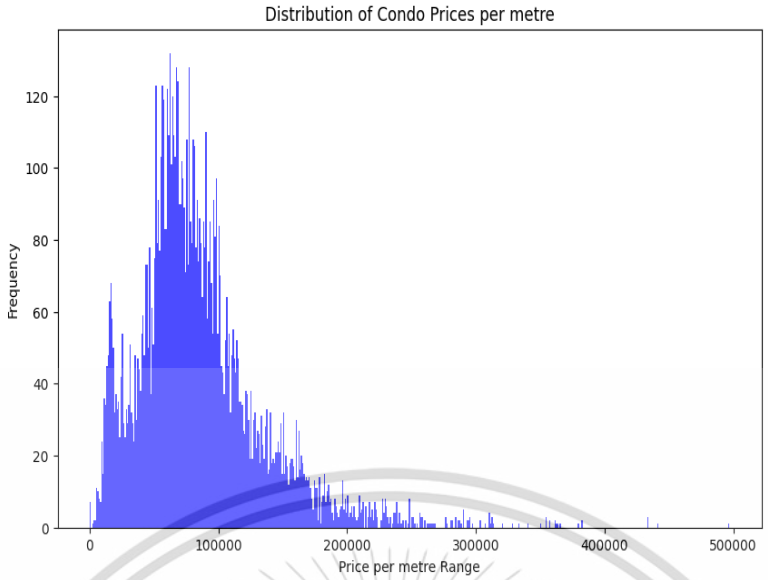
ภาพที่ 3.5 พื้นที่ใช้สอยต่อราคาคอนโดมิเนียมต่อตารางเมตร



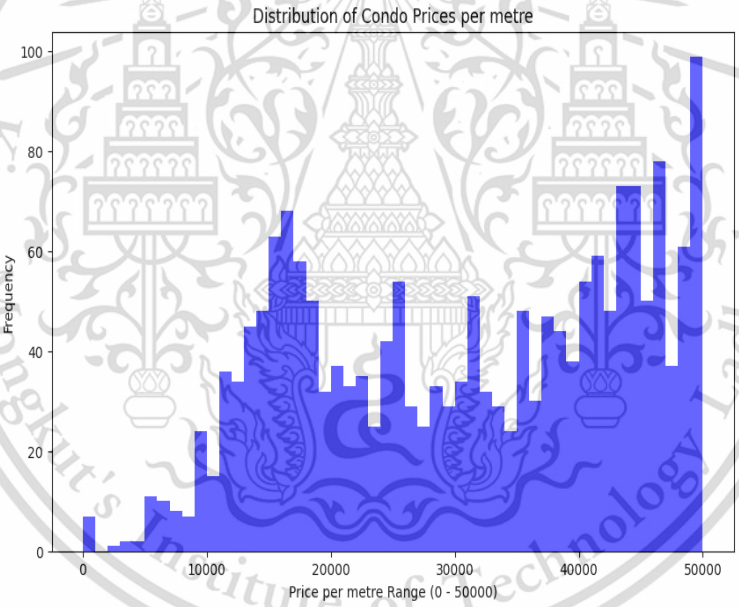
ภาพที่ 3.6 พื้นที่ใช้สอยต่อราคาคอนโดมิเนียมต่อตารางเมตรตั้งแต่ 0 ถึง 20 ตารางเมตร

3) ราคาคอนโดมิเนียมต่อตารางเมตร พบว่ามีราคาคอนโดมิเนียมต่อตารางเมตรบางส่วนที่มีราคาน้อยเกินไป ซึ่งอยู่นอกขอบเขตของการศึกษา คือราคาหลักร้อยลงมา แสดงในภาพที่ 3.7 และภาพที่ 3.8 จึงตัดสินใจลบข้อมูลราคาคอนโดมิเนียมต่อตารางเมตรที่มีราคาน้อยกว่า 7,000 บาทต่อตารางเมตร ออกจากข้อมูลทั้งหมด 33 ห้อง เนื่องจากอ้างอิงข้อมูลจากมติชนออนไลน์ พบว่าราคาคอนโดมิเนียมต่อตารางเมตรที่ถูกที่สุดในกรุงเทพมหานคร คือ คอนโดนิรันดร์เรสซิเดนซ์ ซึ่งข้อมูลที่เก็บรวบรวมมาคอนโดนิรันดร์เรสซิเดนซ์ในการศึกษานี้มีราคาเริ่มที่ 7,000 บาทต่อตารางเมตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



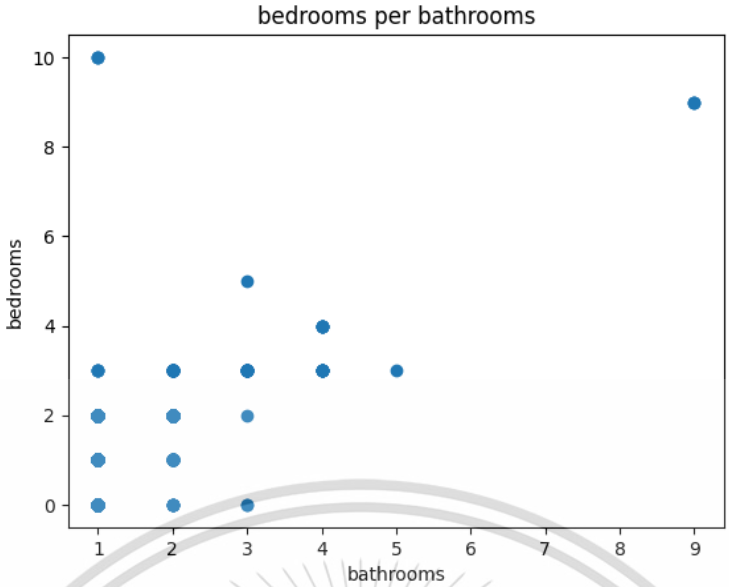
ภาพที่ 3.7 ความถี่ของราคาคอนโดมิเนียมต่อตารางเมตร



ภาพที่ 3.8 ความถี่ขอราคาคอนโดมิเนียมต่อตารางเมตรตั้งแต่ 0 ถึง 50,000 บาทต่อตารางเมตร

4) จำนวนห้องนอนต่อจำนวนห้องน้ำ พบว่ามีจำนวนห้องนอนต่อจำนวนห้องน้ำบางส่วนที่มากเกินไป ซึ่งอยู่นอกขอบเขตของการศึกษา เช่นมี 10 ห้องนอน 1 ห้องน้ำ หรือมี 9 ห้องนอน 9 ห้องน้ำ จึงตัดสินใจลบข้อมูลจำนวนห้องนอนต่อจำนวนห้องน้ำที่มากกว่า 5 ห้อง ออกจากข้อมูลทั้งหมด 6 ห้อง แสดงในภาพที่ 3.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.9 จำนวนห้องนอนต่อจำนวนห้องน้ำ

3.4.2 จัดการกับข้อมูลที่ขาดหาย

จากตารางที่ 3.2 แสดงจำนวนข้อมูลที่ไม่สมบูรณ์อยู่จำนวนหนึ่ง (Missing value) ได้แก่ พื้นที่โครงการ จำนวนชั้นทั้งหมด ชั้นที่อยู่อาศัย ยูนิตทั้งหมด ค่าส่วนกลาง และระยะห่างจากรถไฟฟ้า ดังนั้นจึงได้ทำการปรับปรุงข้อมูลที่ไม่สมบูรณ์ เพื่อให้สามารถแทนข้อมูลที่ไม่สมบูรณ์ได้มากที่สุด ดังนี้ พื้นที่โครงการ ค่าส่วนกลาง และระยะห่างจากรถไฟฟ้า เป็นข้อมูลที่สามารถแทนค่าได้ด้วยวิธีหาค่าเฉลี่ย ซึ่งการแทนด้วยค่าเฉลี่ยนั้นจะทำการแทนค่าเฉลี่ยทั้งหมดสามรอบ เพื่อเติมข้อมูลให้สมบูรณ์มากที่สุด โดยแต่ละรอบมีรายละเอียด ดังนี้

รอบที่หนึ่ง ทำการจับกลุ่มพื้นที่ตั้งของโครงการ และถนนที่เหมือนกัน แล้วหาค่าเฉลี่ย

รอบที่สอง ทำการจับกลุ่มพื้นที่ตั้งของโครงการเหมือนกันแล้วหาค่าเฉลี่ย

รอบที่สาม ทำการจับกลุ่มถนนที่เหมือนกันแล้วหาค่าเฉลี่ย

ส่วนข้อมูล จำนวนชั้นทั้งหมด ชั้นที่อยู่อาศัย และยูนิตทั้งหมด ไม่สามารถแทนค่าทางสถิติได้ เนื่องจากเป็นข้อมูลรายละเอียดเฉพาะเจาะจงของแต่ละห้องในคอนโดมิเนียมนั้น ๆ ดั้งนั้นจากขั้นตอนจัดการข้อมูลที่ขาดหายหลังจากปรับปรุงข้อมูลแล้วยังเหลือข้อมูลที่ไม่สามารถทำให้สมบูรณ์ได้ จึงตัดสินใจลบออกจากข้อมูลออกทั้งหมด 222 ห้อง

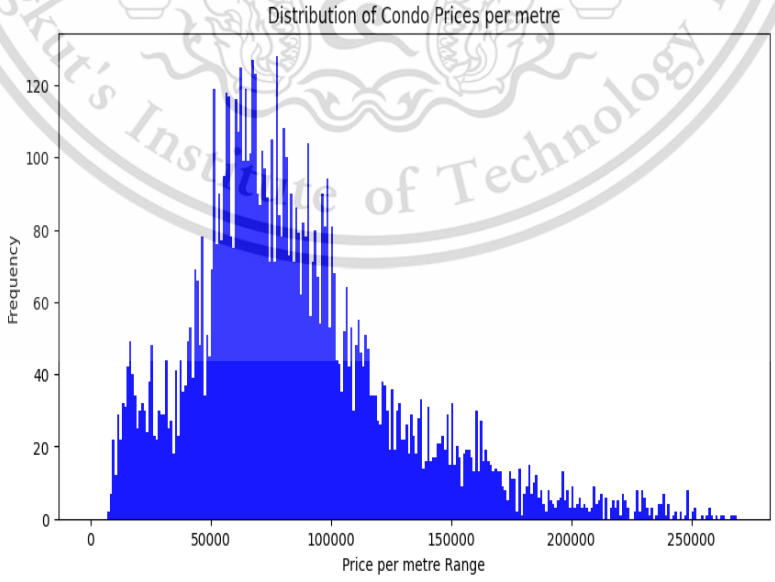
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 จำนวนข้อมูลที่ไม่สมบูรณ์ของแต่ละแอตทริบิวต์

แอตทริบิวต์	จำนวนข้อมูลที่ไม่สมบูรณ์
ปีที่สร้างเสร็จ (Year completed)	0
พื้นที่โครงการ (Project area)	3237
จำนวนชั้นทั้งหมด (Total of floor)	10
ชั้นที่อยู่อาศัย (Floor)	78
ราคาคอนโดมิเนียมต่อตารางเมตร (Price per square meter)	0
จำนวนห้องนอน (bedrooms)	0
จำนวนห้องน้ำ (bathrooms)	0
พื้นที่ใช้สอย (Floor space)	0
ยูนิตทั้งหมด (Total of units)	57
ค่าส่วนกลาง (Common area fee)	2964
ระยะห่างจากรถไฟฟ้า (Distance from the BTS)	351

3.4.3 ลบค่าสุดโต่ง

จากภาพที่ 3.7 จะสังเกตได้ว่ามีค่าสุดโต่งในช่วงราคาคอนโดมิเนียมต่อตารางเมตร จึงกำหนดจำนวนเปอร์เซ็นต์ที่ยอมรับของค่าที่ถูกต้อง (Threshold) เท่ากับ 4 ซึ่งมีค่าเท่ากับ 282,202 บาทต่อตารางเมตร เพื่อลบค่าสุดโต่งออกจากข้อมูลทั้งหมด 315 ห้อง แสดงในภาพที่ 3.10

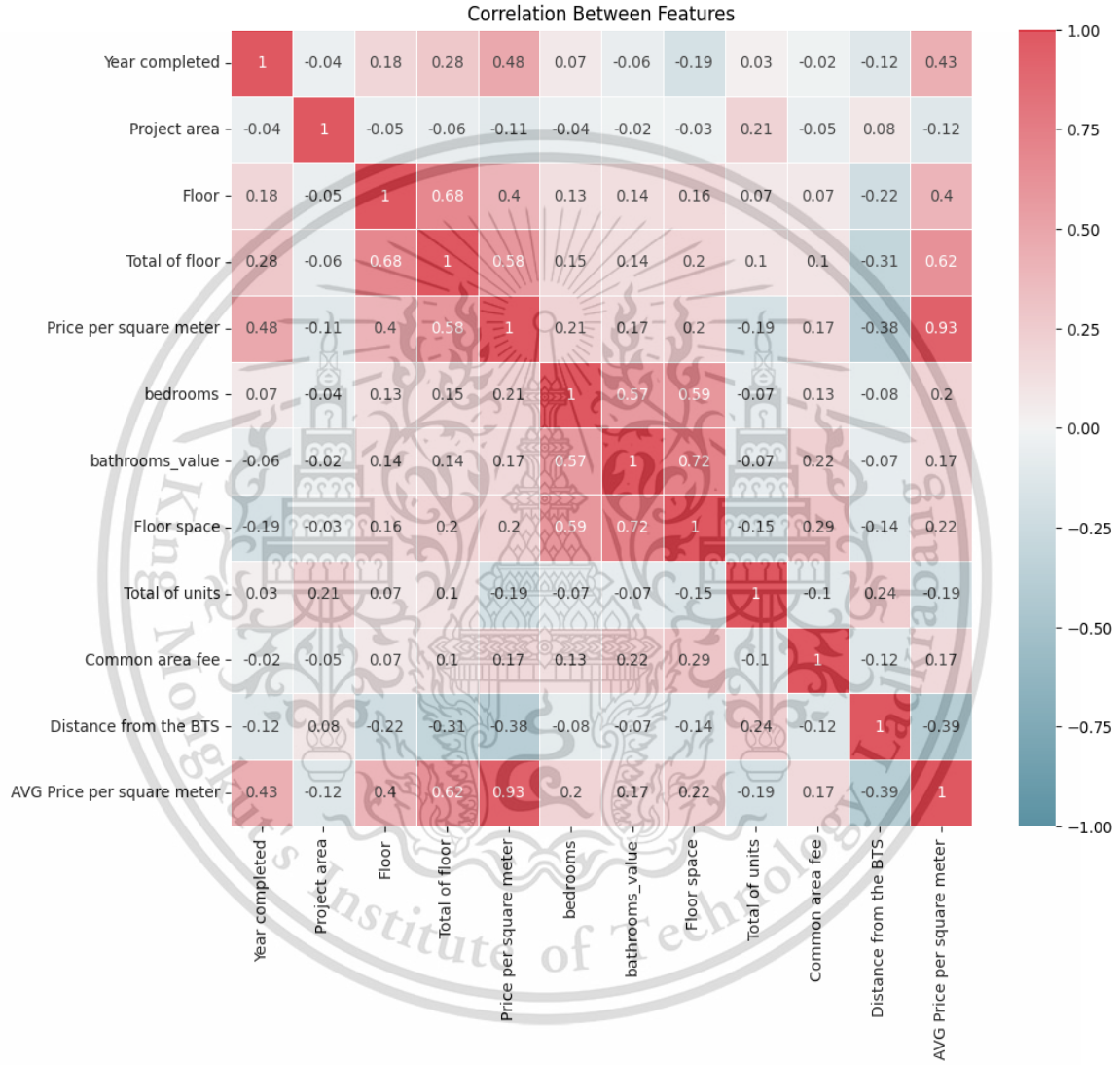


ภาพที่ 3.10 ช่วงความถี่ของราคาคอนโดมิเนียมต่อตารางเมตรหลังลบค่าสุดโต่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.3 เพิ่มแอตทริบิวต์

จากวิธีการประเมินราคา ผู้ประเมินมักจะใช้วิธีการเปรียบเทียบกับราคาตลาด จึงเห็นว่าการเพิ่มแอตทริบิวต์ราคาเฉลี่ยของคอนโดมิเนียมต่อตารางเมตร ซึ่งเกิดจากการจับกลุ่มของพื้นที่ตั้งของโครงการ ถนน และจำนวนชั้นทั้งหมด ที่เหมือนกันแล้วหาค่าเฉลี่ยจะเป็นแอตทริบิวต์ที่ส่งผลต่อค่าสหสัมพันธ์ของราคาคอนโดมิเนียมต่อตารางเมตร แสดงในภาพที่ 3.11



ภาพที่ 3.11 ค่าสหสัมพันธ์ราคาเฉลี่ยของคอนโดมิเนียมต่อตารางเมตร

3.5 แบบจำลอง

แบบจำลอง (Model) ในการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ ได้เปรียบเทียบกัน 4 แบบจำลอง คือ ต้นไม้ตัดสินใจ (Decision Tree) ป่าสุ่ม (Random Forest) Extreme Gradient Boosting (XGBoost) และLight Gradient Boosting (LightGBM)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.1 การจำแนกประเภทแบบต้นไม้ตัดสินใจ (Decision Tree)

การกำหนดพารามิเตอร์ของการจำแนกประเภทแบบต้นไม้ตัดสินใจ มีรายละเอียดแสดงในตารางที่ 3.3

ตารางที่ 3.3 พารามิเตอร์ของการจำแนกประเภทแบบต้นไม้ตัดสินใจ

ชื่อพารามิเตอร์	ความหมาย	ค่า
Max Depth	ความสูงของต้นไม้ตัดสินใจ	None
Min Samples leaf	จำนวนตัวอย่างขั้นต่ำที่ต้องใช้ในการแยกใบ	0.00759
Min Samples Split	จำนวนตัวอย่างขั้นต่ำที่ต้องใช้ในการแยกโหนด	0.00018

3.5.2 การจำแนกประเภทแบบป่าสุ่ม (Random Forest)

การกำหนดพารามิเตอร์ของการจำแนกประเภทแบบป่าสุ่ม มีรายละเอียดแสดงในตารางที่ 3.4

ตารางที่ 3.4 พารามิเตอร์ของการจำแนกประเภทแบบการต้นไม้ป่าสุ่ม

ชื่อพารามิเตอร์	ความหมาย	ค่า
Max Depth	ความสูงของต้นไม้ตัดสินใจ	None
Number of Trees	จำนวนแบบจำลองต้นไม้ตัดสินใจ	100
Min Samples leaf	จำนวนตัวอย่างขั้นต่ำที่ต้องใช้ในการแยกใบ	0.00346
Min Samples Split	จำนวนตัวอย่างขั้นต่ำที่ต้องใช้ในการแยกโหนด	0.00063

3.5.3 การจำแนกประเภทแบบ Extreme Gradient Boosting (XGBoost)

การกำหนดพารามิเตอร์ของการจำแนกประเภทแบบ Extreme Gradient Boosting มีรายละเอียดแสดงในตารางที่ 3.5

ตารางที่ 3.5 พารามิเตอร์ของการจำแนกประเภทแบบ XGBoost

ชื่อพารามิเตอร์	ความหมาย	ค่า
Max Depth	ความสูงของต้นไม้ตัดสินใจ	6
Min Child Weight	จำนวนผลรวมข้อมูลที่ต่ำที่สุดที่อยู่ในโหนดใบ	1
Number of Trees	จำนวนแบบจำลองต้นไม้ตัดสินใจ	100
Gamma	จำนวนต่ำสุดของ loss reduction ที่จะใช้แบ่งโหนดใบของต้นไม้ตัดสินใจ	0
Learning Rate	อัตตราการเรียนรู้ของแบบจำลองโดยมีหน้าที่ป้องกันไม่ให้แบบจำลอง overfit มากเกินไป	0.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5.4 การจำแนกประเภทแบบ Light Gradient Boosting (LightGBM)

การกำหนดพารามิเตอร์ของการจำแนกประเภทแบบ Light Gradient Boosting มีรายละเอียดแสดงในตารางที่ 3.6

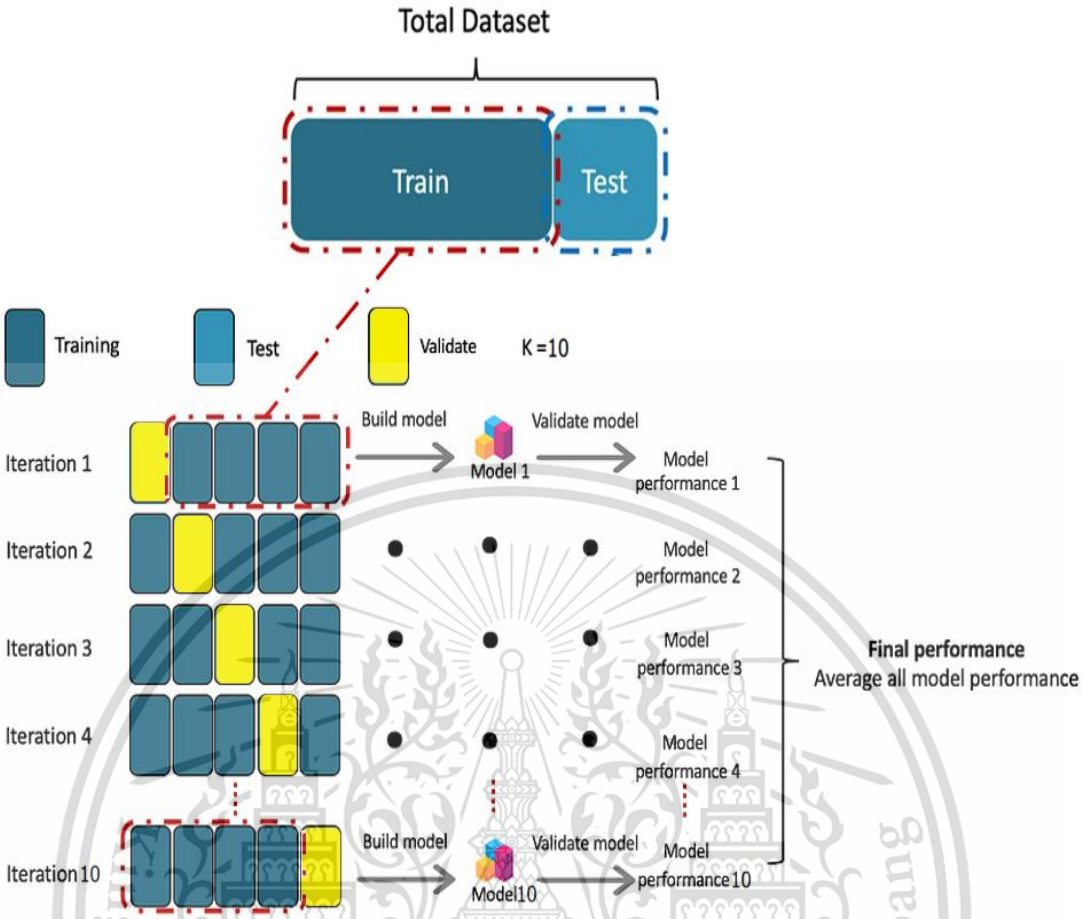
ตารางที่ 3.6 พารามิเตอร์ของการจำแนกประเภทแบบ LightGBM

ชื่อพารามิเตอร์	ความหมาย	ค่า
Max Depth	ความสูงของต้นไม้ตัดสินใจ	-1
Min Child Weight	จำนวนผลรวมข้อมูลที่ต่ำที่สุดที่อยู่ในโหนดใบ	0.001
Number of Trees	จำนวนแบบจำลองต้นไม้ตัดสินใจ	100
Number of Leaves	จำนวนใบในต้นไม้ เพื่อควบคุมความซับซ้อนของโมเดล	31
Learning Rate	อัตตราการเรียนรู้ของแบบจำลองโดยมีหน้าที่ป้องกันไม่ให้แบบจำลอง overfit มากเกินไป	0.1

3.6 ประเมินผลแบบจำลอง

หลังจากลบข้อมูลที่อยู่นอกขอบเขตของการศึกษา จัดการกับข้อมูลที่ขาดหาย ลบค่าสุดโต่ง เพิ่มแอตทริบิวต์ และกำหนดพารามิเตอร์ของแบบจำลองแล้ว เหลือข้อมูลที่จะนำไปเข้าแบบจำลองทั้งหมด 8,441 ห้อง โดยการศึกษานี้จะใช้แบบจำลอง 4 แบบ ในการสร้างต้นแบบ ได้แก่ Decision Tree, Random forest, XGBoost และ LightGBM โดยทำการแบ่งข้อมูลเป็นข้อมูลสำหรับสร้างต้นแบบ (train set) 80 ส่วน และข้อมูลสำหรับทดสอบประสิทธิภาพของต้นแบบ (test set) 20 ส่วน หลังจากนั้นนำข้อมูลต้นแบบ 80 ส่วน มาแบ่งข้อมูลแบบ K-Fold cross validation เป็นข้อมูลสำหรับสร้างต้นแบบ 90 ส่วน และข้อมูลสำหรับตรวจสอบ (validation set) 10 ส่วน แสดงในภาพที่ 3.12 โดยกำหนดค่า K เท่ากับ 10 คือการทำการสร้างและทดสอบต้นแบบทั้งหมด 10 ครั้ง โดยทำการสลับข้อมูลสำหรับสร้างต้นแบบและข้อมูลสำหรับตรวจสอบไปแบบไม่ซ้ำกัน จากนั้นวัดประสิทธิภาพต้นแบบจากการนำผลลัพธ์การสร้างต้นแบบทั้ง 10 ครั้ง มาหาค่าเฉลี่ยแล้วนำมาเปรียบเทียบประสิทธิภาพกับข้อมูลสำหรับทดสอบที่แบ่งเอาไว้ตั้งแต่แรก จากนั้นจะทำวัดผลด้วยวิธีการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 4 แบบด้วย Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) และR-Squared (R^2) แล้วคัดเลือกแบบจำลองการทำนายผลการเรียนรู้ที่ให้ประสิทธิ์ภาพของแบบจำลองดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 3.12 ตัวอย่างการแบ่งข้อมูลด้วยวิธี K-Fold cross validation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการศึกษาและการอภิปรายผล

บทนี้จะกล่าวถึงผลการศึกษาและการอภิปรายผล ซึ่งมีวัตถุประสงค์เพื่อเปรียบเทียบแบบจำลองที่ให้ผลลัพธ์การพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อที่ดีที่สุด โดยเนื้อหาในบทนี้ประกอบด้วยหัวข้อหลัก ๆ คือ ผลลัพธ์การสร้างต้นแบบ และเปรียบเทียบประสิทธิภาพ

4.1 ผลลัพธ์การสร้างต้นแบบ

การสร้างต้นแบบจากแบบจำลอง Decision Tree, Random Forest, XGBoost และ LightGBM ด้วยวิธีแบ่งข้อมูลแบบ K-Fold cross validation โดยใช้ค่า K เท่ากับ 10 คือการทำการสร้าง และทดสอบต้นแบบ 10 ครั้ง โดยมีตัวอย่างผลลัพธ์การสร้างต้นแบบของแบบจำลองทั้ง 4 แบบจำลอง ดังต่อไปนี้

4.1.1 ผลลัพธ์การสร้างต้นแบบของแบบจำลอง Decision Tree

การทดสอบประสิทธิภาพของแบบจำลอง Decision Tree ด้วยวิธีแบ่งข้อมูลแบบ K-Fold cross validation โดยใช้ค่า K เท่ากับ 10 แสดงในตารางที่ 4.1

ตารางที่ 4.1 ตัวอย่างผลลัพธ์ของต้นแบบที่สร้างจากแบบจำลอง Decision Tree

Model	Test			
	Condominium name	actual	prediction	Error
Decision Tree	คันทรี่ คอมเพล็กซ์ บางนา	34,357	38,935	4,578
	ลิฟวิ่งเนสท์ รามคำแหง	70,811	75,065	4,254
	เอ็ม จตุจักร	151,691	162,753	11,062
	ลุมพินี วิลล์ อ่อนนุช 46	55,518	53,333	-2,185
	แบงค์คอก เฟลิซ สาทรตากสิน	98,660	119,667	21,007
	บ้านนวธารา ริเวอร์ไลฟ์	57,057	62,188	5,131
	ไอคอนโด สุขุมวิท 103	57,002	57,813	811
	สุภาพงษ์ เพลส	68,101	37,643	-30,458
	เย็นสบายดี ลาดพร้าว 15	88,009	96,498	8,489
	ลุมพินี สวีท เพชรบุรีมักกะสัน	141,292	139,759	-1,533

4.1.2 ผลลัพธ์การสร้างต้นแบบของแบบจำลอง Random Forest

การทดสอบประสิทธิภาพของแบบจำลอง Random Forest ด้วยวิธีแบ่งข้อมูลแบบ K-Fold cross validation โดยใช้ค่า K เท่ากับ 10 แสดงในตารางที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ตัวอย่างผลลัพธ์ของต้นแบบที่สร้างจากแบบจำลอง Random Forest

Model	Test			
	Condominium name	actual	prediction	Error
Random Forest	คันทรี่ คอมเพล็กซ์ บางนา	35,039	38,935	3,896
	ลิฟวิ่งเนสท์ รามคำแหง	72,877	75,065	2,188
	เอ็ม จตุจักร	149,795	162,753	12,958
	ลุมพินี วิลล์ อ่อนนุช 46	56,036	53,333	-2,703
	แบงค์คอก เฟลิซ สาทรตากสิน	97,441	119,667	22,226
	บ้านนวธารา ริเวอร์ไลฟ์	56,229	62,188	5,959
	ไอคอนโด สุขุมวิท 103	56,772	57,813	1,041
	สุภาพงษ์ เพลส	65,116	37,643	-27,473
	เย็นสบายดี ลาดพร้าว 15	90,659	96,498	5,839
	ลุมพินี สวีท เพชรบุรีมักกะสัน	139,825	139,759	-66

4.1.3 ผลลัพธ์การสร้างต้นแบบของแบบจำลอง XGBoost

การทดสอบประสิทธิภาพของแบบจำลอง XGBoost ด้วยวิธีแบ่งข้อมูลแบบ K-Fold cross validation โดยใช้ค่า K เท่ากับ 10 แสดงในตารางที่ 4.3

ตารางที่ 4.3 ตัวอย่างผลลัพธ์ของต้นแบบที่สร้างจากแบบจำลอง XGBoost

Model	Test			
	Condominium name	actual	prediction	Error
XGBoost	คันทรี่ คอมเพล็กซ์ บางนา	32,566	38,935	6,369
	ลิฟวิ่งเนสท์ รามคำแหง	68,372	75,065	6,693
	เอ็ม จตุจักร	162,211	162,753	542
	ลุมพินี วิลล์ อ่อนนุช 46	57,992	53,333	-4,659
	แบงค์คอก เฟลิซ สาทรตากสิน	106,516	119,667	13,151
	บ้านนวธารา ริเวอร์ไลฟ์	58,223	62,188	3,965
	ไอคอนโด สุขุมวิท 103	57,266	57,813	547
	สุภาพงษ์ เพลส	42,828	37,643	-5,185
	เย็นสบายดี ลาดพร้าว 15	93,803	96,498	2,695
	ลุมพินี สวีท เพชรบุรีมักกะสัน	139,266	139,759	493

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.4 ผลลัพธ์การสร้างต้นแบบของแบบจำลอง LightGBM

การทดสอบประสิทธิภาพของแบบจำลอง LightGBM ด้วยวิธีแบ่งข้อมูลแบบ K-Fold cross validation โดยใช้ค่า K เท่ากับ 10 แสดงในตารางที่ 4.4

ตารางที่ 4.4 ตัวอย่างผลลัพธ์ของต้นแบบที่สร้างจากแบบจำลอง LightGBM

Model	Test			
	Condominium name	actual	prediction	Error
LightGBM	คันทรี่ คอมเพล็กซ์ บางนา	34,725	38,935	4,210
	ลิฟวิ่งเนสท์ รามคำแหง	66,634	75,065	8,431
	เอ็ม จตุจักร	160,746	162,753	2,007
	ลุมพินี วิลล์ อ่อนนุช 46	55,071	53,333	-1,738
	แบงค์คอก เฟลิซ สาทรตากสิน	103,568	119,667	16,099
	บ้านนวธารา ริเวอร์ไลฟ์	56,284	62,188	5,904
	ไอคอนโด สุขุมวิท 103	56,947	57,813	866
	สุภาพงษ์ เพลส	55,442	37,643	-17,799
	เย็นสบายดี ลาดพร้าว 15	89,633	96,498	6,865
	ลุมพินี สวีท เพชรบุรีมักกะสัน	139,938	139,759	-179

4.2 เปรียบเทียบประสิทธิภาพ

หลังจากได้ผลลัพธ์การสร้างต้นแบบของแบบจำลอง Decision Tree, Random Forest, XGBoost และ LightGBM แล้วได้นำผลลัพธ์ทั้งหมด 10 ครั้งมาหาค่าเฉลี่ย และนำมาเปรียบเทียบประสิทธิภาพกับข้อมูลสำหรับทดสอบ โดยสามารถเรียงลำดับแบบจำลองที่มีประสิทธิภาพมากไปน้อยได้ดังนี้ XGBoost, LightGBM, Random Forest และ Decision Tree แสดงในตารางที่ 4.5

ตารางที่ 4.5 เปรียบเทียบผลลัพธ์ของแบบจำลองทั้ง 4 แบบ

Model	Test		
	MAPE	RMSE	R^2
Decision Tree	13.1593%	13,895.5953	0.8978
Random Forest	12.0594%	13,032.3396	0.9101
XGBoost	8.4135%	9,444.8167	0.9527
Light GBM	10.6188%	11,021.8749	0.9357

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการศึกษาและข้อเสนอแนะ

บทนี้จะกล่าวถึงสรุปผลการศึกษาและข้อเสนอแนะ โดยเนื้อหาในบทนี้ประกอบไปด้วยหัวข้อหลัก ๆ คือ สรุปผลการศึกษา ข้อเสนอแนะ และข้อจำกัดในการศึกษา

5.1 สรุปผลการการศึกษา

ในการศึกษานี้ได้นำเสนอการสร้างต้นแบบในการพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ โดยจะช่วยธนาคารลดขั้นตอนการส่งเจ้าหน้าที่ไปประเมินราคาคอนโดมิเนียมในกรุงเทพมหานคร ที่ผู้กู้มีความประสงค์จะซื้อได้โดยใช้เทคนิค Machine Learning เข้ามาช่วยในการพยากรณ์ราคาแทน ซึ่งหารวบรวมข้อมูลได้ในปริมาณมากและข้อมูลมีคุณภาพจะทำให้การการพยากรณ์ราคามีประสิทธิภาพมากขึ้นด้วย

การสร้างต้นแบบจะใช้ข้อมูลต่าง ๆ ที่เกี่ยวกับข้อมูลคอนโดมิเนียมโดมีตัวแปรทั้งหมด 17 ตัวแปร ได้แก่ ชื่อของโครงการคอนโดมิเนียม พื้นที่ตั้งของโครงการ จังหวัด ถนน ปีที่สร้างเสร็จ พื้นที่โครงการ จำนวนชั้นทั้งหมด ชั้นที่อยู่อาศัย ราคาคอนโดมิเนียมต่อตารางเมตร จำนวนห้องนอน จำนวนห้องน้ำ พื้นที่ใช้สอย ยูนิตทั้งหมด ค่าส่วนกลาง พื้นที่จอดรถ และระยะห่างจากรถไฟฟ้า โดยการศึกษาครั้งนี้ได้มีการเพิ่มตัวแปร ราคาเฉลี่ยของคอนโดมิเนียมต่อตารางเมตร ซึ่งส่งผลต่อราคาคอนโดมิเนียมต่อตารางเมตร

การตรวจสอบข้อมูลเพื่อที่จะสามารถสร้างต้นแบบได้อย่างมีประสิทธิภาพ พบว่า มีทั้งข้อมูลที่ไม่สมบูรณ์ ได้แก่ พื้นที่โครงการ ค่าส่วนกลาง และระยะห่างจากรถไฟฟ้าจัดการด้วยวิธีแทนด้วยค่าเฉลี่ย และข้อมูลที่มีค่าผิดปกติได้แก่ ปีที่คอนโดนิเนียมสร้างเสร็จ พื้นที่ใช้สอย ราคาคอนโดมิเนียมต่อตารางเมตร จำนวนห้องนอน และจำนวนห้องน้ำจัดการด้วยวิธีตัดข้อมูลออกแล้วจะได้ข้อมูลสุดท้ายก่อนที่จะนำไปสร้างต้นแบบ

เมื่อนำข้อมูลไปสร้างต้นแบบขั้นแรกจะเข้าสู่ขั้นตอนการแบ่งข้อมูลด้วย K-fold cross validation ในการศึกษาครั้งนี้ได้กำหนดให้ K เท่ากับ 10 แล้วทำการสร้างต้นแบบ 10 ครั้งตามที่แบ่งข้อมูลไว้แล้วนำผลลัพธ์ที่ได้มาเฉลี่ยกันแล้วมาเปรียบเทียบประสิทธิภาพ กับข้อมูลสำหรับทดสอบโดยสามารถเรียงลำดับแบบจำลองที่มีประสิทธิภาพมากไปน้อยได้ดังนี้ XGBoost, LightGBM, Random Forest และ Decision Tree โดยมีค่า MAPE อยู่ที่ 8.4135%, 10.6188%, 12.0594% และ 13.1593% ตามลำดับ ค่า RMSE อยู่ที่ 9,444.8167, 11,021.8749, 13,032.3396 และ 13,895.5953 ตามลำดับ และค่า R^2 อยู่ที่ 0.9527, 0.9357, 0.9101 และ 0.8978 ตามลำดับ สรุปได้ว่าแบบจำลอง XGBoost มีประสิทธิภาพดีที่สุดเมื่อเทียบกับแบบจำลองที่อื่น ๆ และปัจจัยที่ส่งผลต่อราคาคอนโดมิเนียมในกรุงเทพมหานครเพื่อประเมินมูลค่าสินเชื่อ ได้แก่ ราคาเฉลี่ยของคอนโดมิเนียมต่อตารางเมตร จำนวนชั้นทั้งหมด และปีที่สร้างเสร็จ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อเสนอแนะ

สำหรับการศึกษาต่อไป ควรมีการวิเคราะห์เพิ่มเติมเกี่ยวกับประสิทธิภาพของตัวแบบในการทำนายราคาคอนโดมิเนียมในกรุงเทพมหานคร โดยเฉพาะการวิเคราะห์ว่าแบบจำลองมีความแม่นยำในการพยากรณ์ราคากับพื้นที่ตั้งของโครงการหรือถนนแถวไหนมากที่สุด เพื่อที่จะสามารถเริ่มใช้แบบจำลองในการพยากรณ์ราคากับคอนโดมิเนียมในพื้นที่นั้น ๆ ก่อน ส่วนพื้นที่ตั้งของโครงการที่มีประสิทธิภาพต่ำควรพิจารณาเพิ่มเติมว่าปัญหาเกิดขึ้นจากสาเหตุอะไร เช่น ข้อมูลน้อยเกินไป หรือมีข้อมูลที่ผิดปกติมากเกินไป นอกจากนี้อาจพิจารณาเพิ่มประสิทธิภาพของแบบจำลองด้วยการใช้แบบจำลองอื่น ๆ รวมถึงการเก็บรวบรวมข้อมูลจากเว็บไซต์อื่น ๆ มาเพิ่มเติมเพื่อให้ได้ข้อมูลที่ครอบคลุมและหลากหลายมากขึ้น

5.3 ข้อจำกัดในการศึกษา

1. การพยากรณ์ราคาคอนโดมิเนียมในงานศึกษาครั้งนี้ ยังสามารถพยากรณ์ได้แค่คอนโดมิเนียมในจังหวัดกรุงเทพหานครเท่านั้น เนื่องจากข้อมูลคอนโดมิเนียมในจังหวัดอื่น ๆ มีค่อนข้างน้อย

2. ราคาคอนโดมิเนียมต่อตารางเมตรที่เกิดจากการพยากรณ์ในการศึกษานี้ อาจจะไม่ใช่มูลค่าสินเชื่อที่แท้จริงจะต้องมีการปรับลดลงจากราคาที่พยากรณ์ได้ เนื่องจากราคาคอนโดมิเนียมต่อตารางเมตรเกิดจากการเก็บรวบรวมมาจากเว็บไซต์ ไม่ใช่ราคาที่มาจากทางผู้ประเมินจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

Alshamsi, A. (2022). Prediction of Dubai Apartment Prices Using Machine Learning. Rochester Institute of Technology. Retrieved from https://repository.rit.edu/theses/11179

IHRE, A., & ENGSTRÖM, I. (2019). Predicting house prices with machine learning methods. School of Electrical Engineering and Computer Science, Sweden. Retrieved from https://www.diva-portal.org/smash/get/diva2:1354741/FULLTEXT01.pdf

Quang Truong, Minh Nguyen, Hy Dang, & Bo Mei. (2020). Housing Price Prediction via Improved Machine Learning Techniques. Elsevier. Retrieved from https://doi.org/10.1016/j.procs.2020.06.111

Xu, X., & Zhang, Y. (2021). House price forecasting with neural networks. Intelligent Systems with Applications. Retrieved from https://doi.org/10.1016/j.iswa.2021.200052

จิรพล สังข์โพธิ์, และ ศราวุธ แรมจันทร์. (2563). การพยากรณ์ราคาคอนโดมิเนียมในกรุงเทพมหานครโดยเทคนิค Web Mining. วารสารศรีนครินทรวิโรฒวิจัยและพัฒนา, 12(24), 15 - 27. เข้าถึงได้จาก https://www.researchgate.net/publication/353035849_PREDICTING_CONDOMINIUM_PRICE_IN_BANGKOK_USING_WEB_MINING_TECHNIQUES_In_Thai

พสธร วิทยาปรีชาพล. (2563). การประเมินราคาเสนอขายห้องชุดด้วย deep neural network และ K-means clustering algorithm. มหาวิทยาลัยธรรมศาสตร์ https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:184130

พัชริดา เสนีย์วงศ์. (2565). การพยากรณ์อัตราการขายของโครงการหมู่บ้านจัดสรรด้วยเทคนิคการเรียนรู้ของเครื่อง. มหาวิทยาลัยธรรมศาสตร์ https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:306376

พีรภัทร วัสแสง. (2564). การพยากรณ์ราคาขายทรัพย์สินรอการขายประเภทที่อยู่อาศัยในพื้นที่กรุงเทพมหานครโดยใช้เทคนิคการเรียนรู้ของเครื่อง. มหาวิทยาลัยธรรมศาสตร์. https://digital.library.tu.ac.th/tu_dc/frontend/Info/item/dc:275006

มนตรี ล้อเลิศสกุล , และ อานนท์ ศักดิ์วรวิชญ์. (2560). แบบจำลองการวิเคราะห์ถดถอยเชิงพื้นที่เพื่อพยากรณ์ราคาขายเฉลี่ยต่อตารางเมตรของอาคารชุดในกรุงเทพมหานคร. วารสารวิชาการสถาปัตยกรรมศาสตร์, 121 - 138. เข้าถึงได้จาก https://so01.tci-thaijo.org/index.php/AJA/article/view/115431/89174

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นาย ภควัต อมรวิลาศ
วัน เดือน ปีเกิด	29 พฤษภาคม 2539
ที่อยู่ปัจจุบัน	28/100 ถนน กาญจนาภิเษก ตำบล บางม่วง อำเภอ บางใหญ่ นนทบุรี 11140
ประวัติการศึกษา	(2561) วิทยาศาสตรบัณฑิต สาขาสถิติ เกรดเฉลี่ย 2.70 (มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ) (2566) วิทยาศาสตร์มหาบัณฑิต สาขาวิทยาการข้อมูลและการวิเคราะห์ (สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้