การทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม

วิทยานิพนธ์

ของ

กิตติศักดิ์ ลาทอง

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล

ธันวาคม 2566

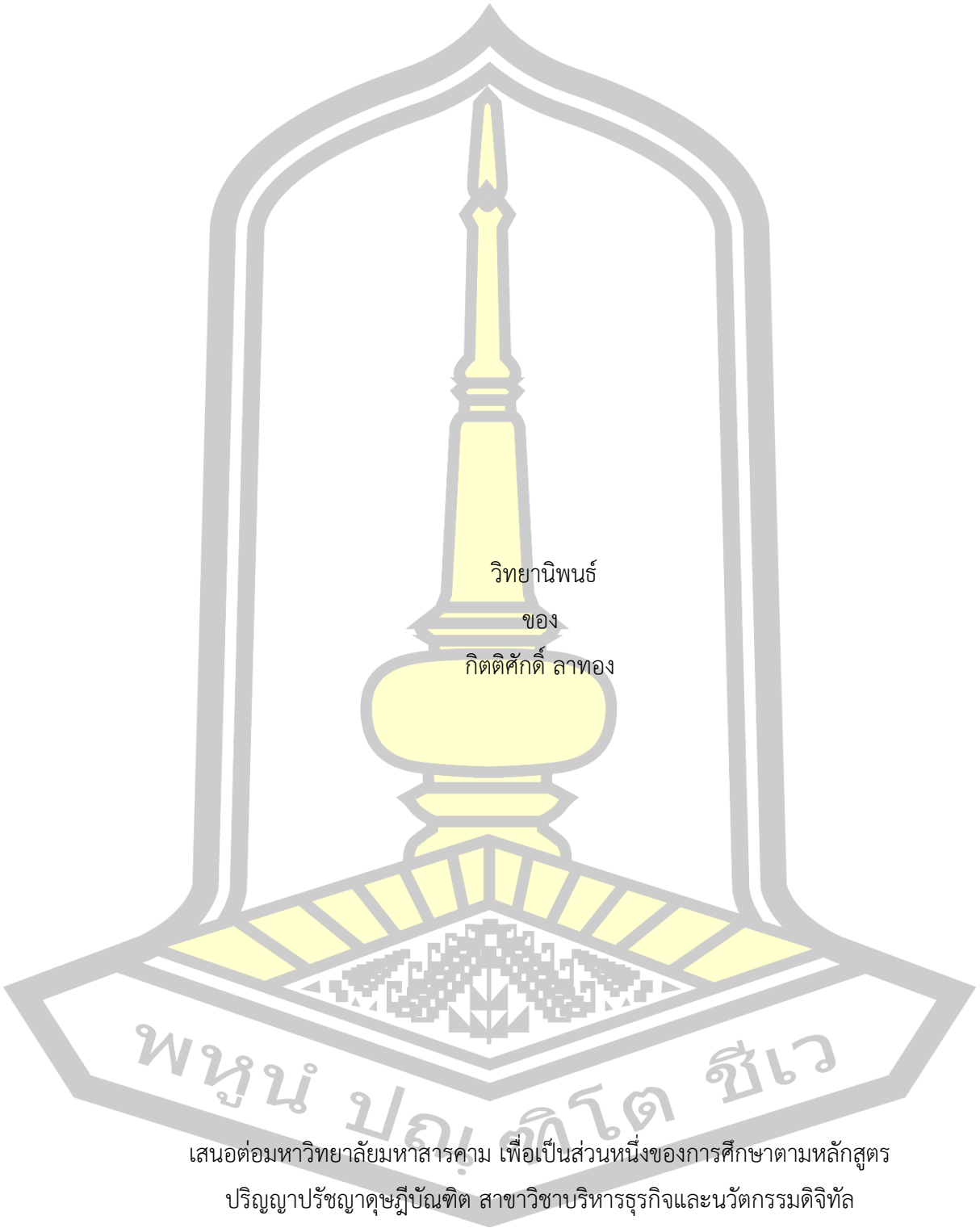
ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม

วิทยานิพนธ์

ของ

กิตติศักดิ์ ลาทอง



เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล

ธันวาคม 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine

Kittisak Lathong

A Thesis Submitted in Partial Fulfillment of Requirements

for Doctor of Philosophy (Business Administration and Digital Innovation)

December 2023

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายกิตติศักดิ์ ลาทอง แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รศ. ดร. ธีระวัฒน์ เจริญราษฎร์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รศ. ดร. กิตติพล วิแสง)

.....กรรมการ
(ดร. การันต์ กิจระการ)

.....กรรมการ
(ดร. กาญจนา หินเธาว์)

.....กรรมการ
(ดร. พงศธร ตันตระบัณฑิตย์)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา ปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล ของมหาวิทยาลัยมหาสารคาม

พหูน ปณฺฑิโต ชีเว

....

(ดร. ชลธิชา ธรรมวิญญู)
คณบดีคณะการบัญชีและการจัดการ

(รศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง การทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม
ผู้วิจัย กิตติศักดิ์ ลาทอง
อาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร. กิตติพล วิแสง
ปริญญา ปรัชญาดุษฎีบัณฑิต **สาขาวิชา** บริหารธุรกิจและนวัตกรรมดิจิทัล
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2566

บทคัดย่อ

การวิจัยนี้มีเป้าหมายในการใช้แบบจำลองการเรียนรู้เพื่อทำนายราคาค่าก่อสร้างบ้านพักแบบแนวราบ โดยใช้เทคนิคและแบบจำลองหลากหลาย เพื่อหาแบบจำลองที่มีประสิทธิภาพสูงสุดในการทำนายระหว่างแบบจำลองการเรียนรู้ทั้งหมด 10 แบบ โดยประเมินผลโดยใช้ตัวชี้วัดเช่น R2, MSE, RMSE, และ MAE ผลการทดลองเปิดเผยว่าโครงข่ายประสาทเทียม (ANN) เป็นแบบจำลองเชิงเดี่ยวที่มีประสิทธิภาพสูงสุดด้วยคะแนน R2 เท่ากับ 0.891 ตามมาด้วยการวิเคราะห์ถดถอยพหุคูณ (MLR) และต้นไม้ตัดสินใจ (DT) ที่มีความแม่นยำเท่ากันที่ 0.884 และ 0.864 ตามลำดับ ในแบบจำลองเครื่องจักรเรียนรู้เอ็กซ์ตรีม แบบจำลองการเรียนรู้แบบการโหวตสูงสุด (MVE) มีประสิทธิภาพสูงสุดด้วยความแม่นยำที่ 0.924 ที่สูงเป็นพิเศษ และแบบจำลองการเรียนรู้แบบ stacking และแบบเฉลี่ยมีความแม่นยำเท่ากับ 0.883 และ 0.871 ตามลำดับ ผลลัพธ์นี้เน้นถึงศักยภาพของวิธีการใหม่ที่แนะนำโดยกลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมในการเพิ่มความแม่นยำในการทำนายเหนือแบบจำลองเดี่ยวและนำเสนอความรู้มูลค่าสำหรับผู้เกี่ยวข้องในอุตสาหกรรมอสังหาริมทรัพย์ และแนะนำเส้นทางสำหรับการศึกษาเพิ่มเติมในการสำรวจรูปแบบแบบจำลองสถาปัตยกรรมและเทคนิคเพื่อความแม่นยำในการทำนายที่มากขึ้นสำหรับโครงสร้างสถาปัตยกรรมต่างๆ รวมถึงการรวมตัวแปรเพิ่มเติมในอนาคตเช่นกัน

คำสำคัญ : เครื่องจักรเรียนรู้เอ็กซ์ตรีม, แบบจำลองการทำนาย, บ้านพักอาศัยแนวราบ, การประมาณราคาค่าก่อสร้าง, การเรียนรู้ของเครื่องจักร

TITLE The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine

AUTHOR Kittisak Lathong

ADVISORS Associate Professor Kittipol Wisaeng , Ph.D.

DEGREE Doctor of Philosophy **MAJOR** Business Administration and Digital Innovation

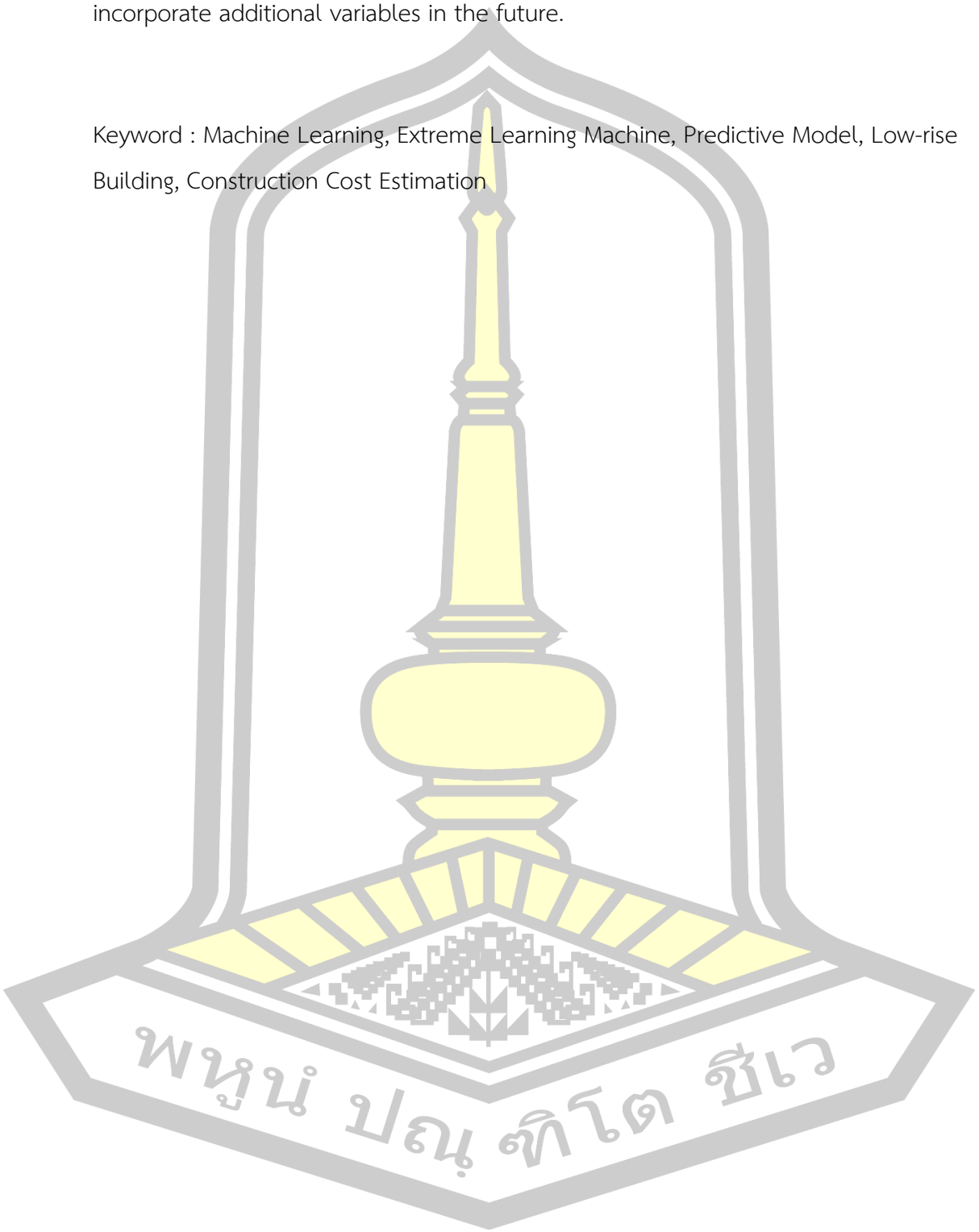
UNIVERSITY Mahasarakham University **YEAR** 2023

ABSTRACT

This research aims to employ learning model frameworks for predicting construction cost estimation of low-rise buildings. The study uses various techniques and models to identify the most effective predictor among all 10 learning models, evaluating their performance using indicators such as R2, MSE, RMSE, and MAE. Subsequently, model efficiency is assessed within two major groups: 1) The single-model or basic-model group includes models like Artificial Neural Network (ANN), Support Vector Machine (SVM), Multiple Linear Regression (MLR), Decision Tree (DT), and Random Forest (RF). 2) The ensemble learning model group consists of models like the Maximum Voted Ensemble (MVE), Average Ensemble, Stacking Ensemble, Bagging Ensemble, and Boosting Ensemble. The experimental results highlight the Artificial Neural Network (ANN) as the most proficient single-model predictor with an R2 Score of 0.891, followed by Multiple Linear Regression (MLR) and Decision Tree (DT) models, both with accuracies of 0.884 and 0.864, respectively. In the extreme learning machine model group, the Maximum Voted Ensemble (MVE) exhibits superior efficiency with an exceptional accuracy of 0.924. This discovery underscores the potential of the novel approach proposed by the extreme ensemble group, offering enhanced prediction accuracy compared to individual basic models and providing valuable insights for real estate stakeholders. Furthermore, the research suggests that further investigation into architectural model variations and additional techniques

could lead to even higher prediction accuracy for various architectural structures and incorporate additional variables in the future.

Keyword : Machine Learning, Extreme Learning Machine, Predictive Model, Low-rise Building, Construction Cost Estimation



กิตติกรรมประกาศ

วิทยานิพนธ์นี้ได้สำเร็จลุล่วงอย่างดีเพราะรับความกรุณาและเข้าใจจากอาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร. กิตติพล วิแสง อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้เวลามากมายในการให้คำปรึกษาและแนะนำที่เป็นประโยชน์ตลอดขั้นตอนการทำวิจัยนี้ พร้อมที่จะตรวจสอบและแก้ไขข้อผิดพลาดต่างๆ เพื่อให้วิทยานิพนธ์นี้เป็นสมบูรณ์อย่างเต็มที่ ด้วยความใส่ใจและความเมตตาที่อันมีอยู่ต่อผู้วิจัยเสมอ ผู้วิจัยรู้สึกถึงความสำคัญของความกรุณาและขอขอบพระคุณอย่างสูง

ขอบคุณท่านรองศาสตราจารย์ ดร. ธีระวัฒน์ เจริญราษฎร์ ที่ได้เป็นประธานกรรมการสอบและคณาจารย์ท่านอื่น ๆ ในคณะกรรมการสอบ ได้แก่ ดร. การันต์ กิจระการ ดร. พงศธร ตันตระบัณฑิตย์ ดร. ดวงรัตน์ ธารดำรง และ ดร. กาญจนา หินเธาว์ ที่ให้คำแนะนำอย่างดีตลอดตั้งแต่เริ่มต้นทำวิจัยจนสิ้นสุด

ขอขอบคุณคณาจารย์ทุกท่านที่ได้ให้ความรู้และการศึกษาที่เป็นประโยชน์กับผู้วิจัย ตลอดระยะเวลาการศึกษาในหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาบริหารธุรกิจและนวัตกรรมดิจิทัล มหาวิทยาลัยมหาสารคาม ซึ่งทำให้ผู้วิจัยได้พัฒนาทั้งในเรื่องความคิดและการปฏิบัติ ซึ่งจะเป็นประโยชน์ต่อผู้วิจัยในการทำงานในอนาคต

ขอขอบคุณเจ้าหน้าที่คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม รวมถึงเพื่อนร่วมรุ่นทั้งหมด โดยเฉพาะอย่างยิ่ง ดร. กุสุมา สร้อยทอง และ ดร. เสกสรรค์ วีระสุข ที่คอยให้ความช่วยเหลือและความสะดวกให้กับผู้วิจัยตลอด

ขอขอบคุณพ่อ แม่ และพี่ที่ได้ให้ความรักและความอบอุ่น การเลี้ยงดูที่ดี และการสนับสนุนที่ไม่รู้ลืม ที่ได้ช่วยเหลือให้ผู้วิจัยเข้ามาเป็นคนที่แข็งแกร่งพร้อมที่จะเผชิญกับอุปสรรคต่าง ๆ และบรรลุความสำเร็จได้ในวันนี้

ขอบคุณทุกท่านที่มีส่วนร่วมในความสำเร็จของวิทยานิพนธ์เล่มนี้ ผู้วิจัยจะไม่ลืมความเอาใจใส่ทุกคนในการช่วยเหลือ สำคัญที่สุดคือความกรุณาที่ท่านเสมอมาให้

ขอบคุณมหาวิทยาลัยมหาสารคาม สำหรับทุนสนับสนุนการวิจัย ประจำปีงบประมาณ 2566 เลขที่ 6603024 เป็นอย่างสูง ทำให้ผู้วิจัยได้มีโอกาสเผยแพร่งานวิจัยนี้ออกไปสู่สาธารณะอย่างกว้างขวาง

กิตติศักดิ์ ลาทอง

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ช

สารบัญ ..... ซ

สารบัญตาราง ..... ฎ

สารบัญภาพประกอบ ..... ฎ

บทที่ 1 บทนำ 1

ภูมิหลัง 1

ความมุ่งหมายของการวิจัย 9

ขอบเขตของการวิจัย 9

นิยามศัพท์เฉพาะ 10

ประโยชน์ที่คาดว่าจะได้รับ 10

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 11

แนวคิดเกี่ยวกับการประมาณราคาค่าก่อสร้างบ้านพักอาศัย 11

1. การประมาณราคาก่อสร้าง 11

2. แนวคิดและทฤษฎีที่เกี่ยวข้องกับการบริหารงานก่อสร้าง 14

2.1. ความหมายของการบริหารงานก่อสร้าง 14

2.2 ประเภทของงานก่อสร้าง 14

2.3 ขั้นตอนการก่อสร้างบ้าน 15

2.4 ขั้นตอนการบริหารงานก่อสร้าง 17

2.5 ทรัพยากรในงานก่อสร้าง 19

แนวคิดเกี่ยวกับเทคโนโลยีการเรียนรู้ของเครื่อง 19
1. ลักษณะการทำงานของการเรียนรู้ของเครื่อง 19
2. ประโยชน์ของเทคโนโลยีการเรียนรู้ของเครื่อง 22
3. ข้อจำกัดของเทคโนโลยีการเรียนรู้ของเครื่อง 24
เครื่องจักรเรียนรู้เอ็กซ์ตรีม 25
1. การเรียนรู้กลุ่มแบบโหวตสูงสุด 25
2. การเรียนรู้กลุ่มแบบเฉลี่ย 26
3. การเรียนรู้กลุ่มแบบสแต็กกิ้ง 28
4. การเรียนรู้กลุ่มแบบแบ๊กกิ้ง 29
5. การเรียนรู้กลุ่มแบบบูสต์ติง 32
โครงข่ายประสาทเทียม 33
1. ส่วนประกอบโครงข่ายประสาทเทียม 34
2. ลักษณะของโครงข่ายประสาทเทียม 34
ซัพพอร์ตเวกเตอร์แมชชีน 36
การวิเคราะห์ความถดถอย 40
ต้นไม้ตัดสินใจ 41
แรนดอมฟอเรสต์ 43
งานวิจัยที่เกี่ยวข้อง 45
บทที่ 3 วิธีดำเนินการวิจัย 55
ชุดข้อมูล 57
เครื่องมือที่ใช้ในงานวิจัย 58
การเก็บรวบรวมข้อมูล 59
การจัดกระทำกับข้อมูล 61
การสร้างแบบจำลอง 64

การทดสอบเปรียบเทียบประสิทธิภาพ 66
1. Mean Absolute Error (MAE) 66
2. Mean Square Error (MSE) 67
3. Root Mean Square Error (RMSE) 67
4. R square (R^2) 68
การปรับแก้ค่าทำนายการประมาณราคาค่าก่อสร้าง 69
บทที่ 4 ผลการวิเคราะห์ข้อมูล 71
ผลการวิเคราะห์ข้อมูลนำเข้าที่ใช้ในการวิจัย 71
ผลลัพธ์วิธีการดั้งเดิมหรือแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน 75
ผลลัพธ์วิธีการใหม่หรือแบบจำลองเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม 81
ผลการเปรียบเทียบประสิทธิภาพของวิธีการใหม่ที่นำเสนอในการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบกับวิธีการดั้งเดิม 83
ผลการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองในการการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบเพื่อหาวิธีการที่รวดเร็วและเหมาะสมที่สุด 87
บทที่ 5 สรุปผล อภิปราย และข้อเสนอแนะ 89
สรุปผลและอภิปราย 89
ข้อเสนอแนะ 92
บรรณานุกรม 93
ภาคผนวก 101
ภาคผนวก ก Source Code สำหรับสร้างแบบจำลองและวิเคราะห์แบบจำลองด้วยไพทอน 102
ประวัติผู้เขียน 119

สารบัญตาราง

ตาราง 1 การประมาณราคาแบบขึ้นกับผลิตภัณฑ์แบบต่างๆ ซึ่งมีการใช้อย่างแพร่หลาย 12

ตาราง 2 การประมาณราคาแบบขึ้นกับผลิตภัณฑ์ที่มีการใช้ไม่แพร่หลาย 13

ตาราง 3 ตัวอย่างวิธีการประมาณราคาแบบขึ้นกับกระบวนการ 14

ตาราง 4 คุณลักษณะของตัวแปรและการอ้างอิงที่มาจากการทบทวนวรรณกรรม..... 57

ตาราง 5 ดัชนีราคาวัสดุก่อสร้างและปัจจัยที่ใช้เป็นตัวคูณปรับฐาน..... 69

ตาราง 6 ผลลัพธ์แบบจำลองด้วยวิธีการวิเคราะห์สมการเชิงเส้นหลายตัวแปร 78

ตาราง 7 ประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน..... 79

ตาราง 8 ประสิทธิภาพของวิธีการใหม่หรือแบบจำลองเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม..... 82

ตาราง 9 สรุปผลประสิทธิภาพของอัลกอริทึมที่ใช้ในการวิจัย..... 84

พหุน ปณฺฑิโต ชีเว

สารบัญภาพประกอบ

ภาพประกอบ 1 การเปลี่ยนแปลงของการอนุญาตก่อสร้างอาคารโรงเรือน ระหว่างปี 2554-2564.... 2

ภาพประกอบ 2 จำนวนผู้ได้รับอนุญาตให้ก่อสร้างใหม่และต่อเติมหรือดัดแปลงสิ่งก่อสร้างและจำนวนสิ่งก่อสร้างที่ได้รับอนุญาต จำแนกตามประเภทสิ่งก่อสร้างเป็นรายภาค 2

ภาพประกอบ 3 พื้นที่สิ่งก่อสร้างที่เป็นอาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างใหม่และต่อเติมหรือดัดแปลง จำแนกตามประเภทของอาคารโรงเรือน 3

ภาพประกอบ 4 ความยาวสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างจำแนกตามชนิดของสิ่งก่อสร้าง 4

ภาพประกอบ 5 พื้นที่สิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างจำแนกตามชนิดของสิ่งก่อสร้าง..... 4

ภาพประกอบ 6 ความคลาดเคลื่อนในการประมาณราคาค่าก่อสร้างกับระยะเวลา 8

ภาพประกอบ 7 ความสัมพันธ์ของปัญญาประดิษฐ์ การเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก 20

ภาพประกอบ 8 ประเภทของการเรียนรู้ของเครื่อง..... 20

ภาพประกอบ 9 การเรียนรู้ของเครื่องแบบเสริมกำลัง..... 22

ภาพประกอบ 10 แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด 26

ภาพประกอบ 11 แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย..... 27

ภาพประกอบ 12 แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง 28

ภาพประกอบ 13 แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง 30

ภาพประกอบ 14 แบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง 31

ภาพประกอบ 15 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว..... 34

ภาพประกอบ 16 โครงสร้างโครงข่ายประสาทเทียมแบบหลายชั้น 35

ภาพประกอบ 17 แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน..... 37

ภาพประกอบ 18 แบบจำลองการวิเคราะห์ถดถอยพหุคูณ 40

ภาพประกอบ 19 แบบจำลองต้นไม้ตัดสินใจ 42

ภาพประกอบ 20 กรอบแนวคิดวิจัย 56
ภาพประกอบ 21 ตัวอย่างแบบแปลนบ้านจากโครงการแบบบ้านยิ้ม แบบบ้านฟรีเพื่อประชาชน 60
ภาพประกอบ 22 ตัวอย่างรูปด้านของบ้านจากโครงการแบบบ้านยิ้ม แบบบ้านฟรีเพื่อประชาชน 60
ภาพประกอบ 23 วิธีการวิเคราะห์ของแบบจำลอง 62
ภาพประกอบ 24 เทคนิคการทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ 65
ภาพประกอบ 25 ลักษณะข้อมูลตัวแปรอิสระ 72
ภาพประกอบ 26 ทัศนภาพข้อมูลตัวแปรอิสระ 73
ภาพประกอบ 27 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน 76
ภาพประกอบ 28 การวิเคราะห์ตัวประกอบ (Factor Analysis) 77
ภาพประกอบ 29 การเปรียบเทียบประสิทธิภาพของการทำนายราคาค่าก่อสร้าง 85
ภาพประกอบ 30 ผลการเปรียบเทียบวิธีการดั้งเดิมกับวิธีการใหม่ 86

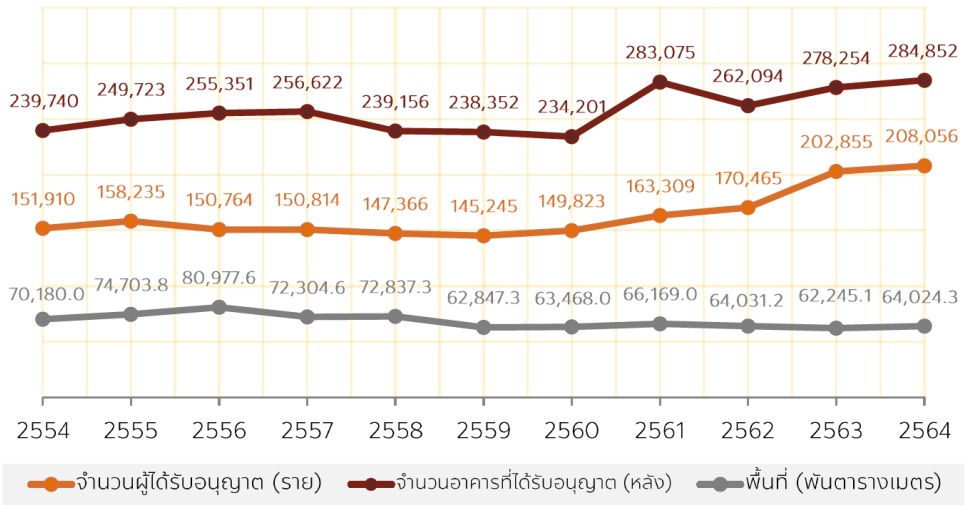
บทที่ 1

บทนำ

ภูมิหลัง

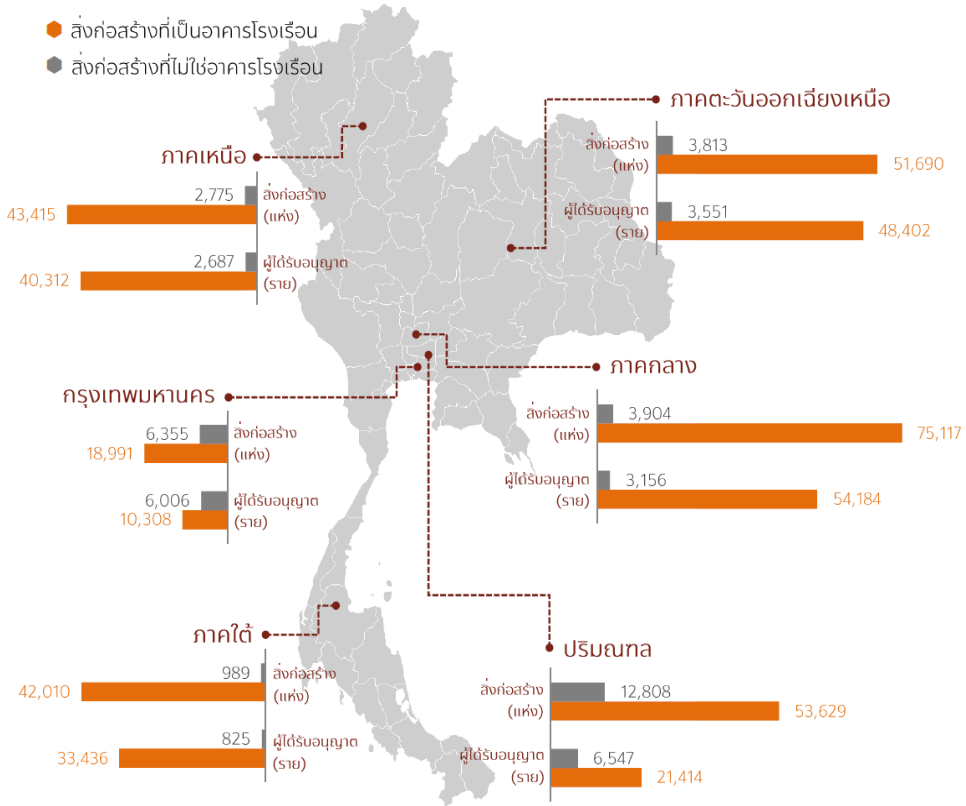
หนึ่งในสี่สิ่งที่จำเป็นสำหรับการดำรงอยู่ของมนุษย์คือที่อยู่อาศัย ซึ่งอัตราการเติบโตของประชากรในปัจจุบันที่สูงขึ้นความต้องการที่จะสร้างบ้านพักอาศัยหรือปรับปรุงที่อยู่อาศัยก็สูงขึ้นตามไปด้วย ซึ่งส่งผลโดยตรงกับอุตสาหกรรมอสังหาริมทรัพย์ที่จะมีการเติบโตขึ้น นำไปสู่การก่อสร้างบ้านพักอาศัยที่เพิ่มขึ้น ทั้งบ้านเดี่ยว ทาวน์เฮาส์ อาคารพาณิชย์ คอนโดมิเนียม และโครงสร้างอื่นๆ จากข้อมูลของกองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ (2564) พบว่าการขออนุญาตก่อสร้างอาคารโรงเรือนตั้งแต่ 2554 – 2564 ค่อนข้างมีความผันผวนสลับขึ้นลงในแต่ละปี อย่างไรก็ตาม ตั้งแต่ปี 2559 จำนวนผู้ได้รับอนุญาตให้ก่อสร้างอาคารโรงเรือนมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยในปี 2564 เพิ่มขึ้นร้อยละ 2.6 เมื่อเปรียบเทียบกับปี 2563 สำหรับจำนวนอาคารโรงเรือนที่ได้รับอนุญาตมีอัตราการเพิ่มขึ้นค่อนข้างคงที่ตั้งแต่ปี 2547 – 2557 หลังจากนั้นเริ่มลดลงอย่างต่อเนื่องจนถึงปี 2561 จำนวนอาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างมีการเปลี่ยนแปลงเพิ่มขึ้นถึงร้อยละ 20.9 จากปี 2560 ก่อนจะเพิ่มขึ้นอีกครั้งในปี 2563 โดยในปี 2564 จำนวนโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างเพิ่มขึ้นร้อยละ 2.4 เมื่อเปรียบเทียบกับปี 2563 ส่วนพื้นที่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างมีอัตราการเพิ่มขึ้นค่อนข้างคงที่ในปี 2554 – 2556 หลังจากนั้นมีการเปลี่ยนแปลงสลับขึ้นลงในแต่ละปี โดยในปี 2564 พื้นที่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างมีการเปลี่ยนแปลงเพิ่มขึ้นร้อย 2.9 เมื่อเปรียบเทียบกับปี 2563 (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังภาพประกอบ 1

สถิติตลอดปี 2564 มีผู้ได้รับอนุญาตให้ก่อสร้างใหม่และต่อเติมหรือดัดแปลงสิ่งก่อสร้างทั้งที่เป็นอาคารโรงเรือน และไม่ใช่อาคารโรงเรือนทั้งสิ้น 230,828 ราย โดยภาคกลางมีจำนวนผู้ได้รับอนุญาตฯ สูงสุด 57,340 ราย รองลงมาเป็นภาคตะวันออกเฉียงเหนือ 51,953 ราย ภาคเหนือ 42,999 ราย ภาคใต้ 34,261 ราย ปริมณฑล 27,961 ราย และกรุงเทพมหานคร 16,314 ราย แต่เมื่อพิจารณาจำนวนสิ่งก่อสร้างที่ได้รับอนุญาตทั้งสิ้น 315,496 แห่งพบว่าเป็นการก่อสร้างในภาคกลางสูงสุด 79,021 แห่ง รองลงมาเป็นปริมณฑท 66,437 แห่ง ภาคตะวันออกเฉียงเหนือ 55,503 แห่ง ภาคเหนือ 46,190 แห่ง ภาคใต้ 42,999 แห่ง และกรุงเทพมหานคร 25,346 แห่ง (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังภาพประกอบ 2



ที่มา: กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, 2564

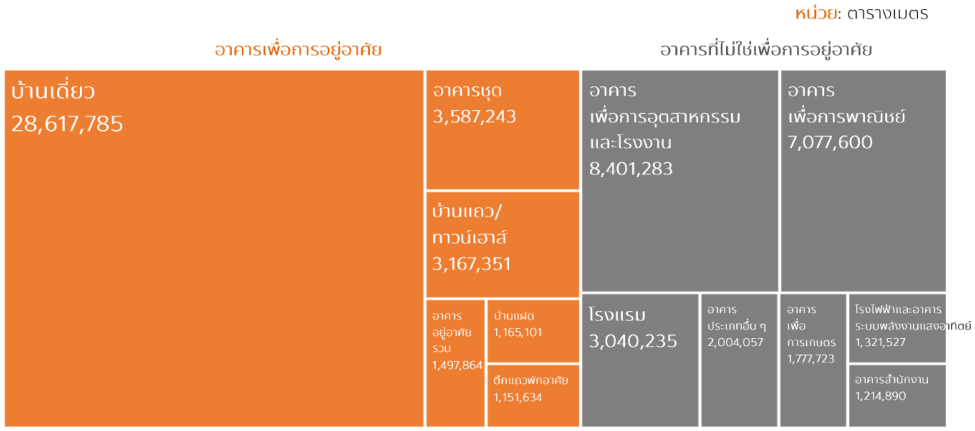
ภาพประกอบ 1 การเปลี่ยนแปลงของการอนุญาตก่อสร้างอาคารโรงเรือน ระหว่างปี 2554-2564



ที่มา: กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, 2564

ภาพประกอบ 2 จำนวนผู้ได้รับอนุญาตให้ก่อสร้างใหม่และต่อเติมหรือดัดแปลงสิ่งก่อสร้างและจำนวนสิ่งก่อสร้างที่ได้รับอนุญาต จำแนกตามประเภทสิ่งก่อสร้างเป็นรายภาค

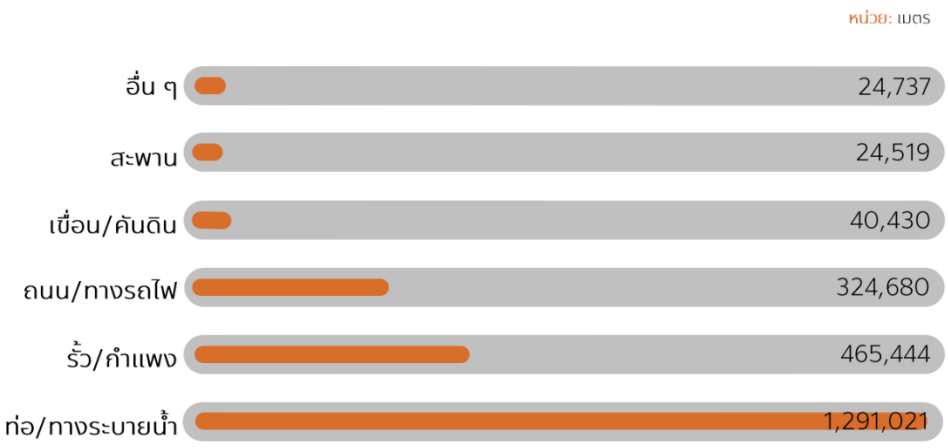
สิ่งก่อสร้างที่เป็นอาคารโรงเรือนตลอดทั้งปี 2564 มีสิ่งก่อสร้างที่เป็นอาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างเป็นพื้นที่รวมทั้งสิ้น 64,024,293 ตารางเมตร ในจำนวนนี้เป็นอาคารเพื่อการอยู่อาศัย 39,186,978 ตารางเมตร และอาคารที่ไม่ใช่เพื่อการอยู่อาศัย 24,837,315 ตารางเมตร เมื่อพิจารณาพื้นที่อาคารเพื่อการอยู่อาศัย พบว่า ส่วนใหญ่เป็นการได้รับอนุญาตให้ก่อสร้างบ้านเดี่ยวเป็นพื้นที่ 28,617,875 ตารางเมตร รองลงมาเป็นอาคารชุด (คอนโดมิเนียม) 3,587,243 ตารางเมตร อาคารอยู่อาศัยรวม (หอพัก,อพาร์ทเม้นต์) 1,497,864 ตารางเมตร บ้านแฝด 1,165,101 ตารางเมตร และตึกแถวพักอาศัย 1,151,634 ตารางเมตร (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังภาพประกอบ 3



ที่มา: กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, 2564

ภาพประกอบ 3 พื้นที่สิ่งก่อสร้างที่เป็นอาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างใหม่และต่อเติมหรือดัดแปลง จำแนกตามประเภทของอาคารโรงเรือน

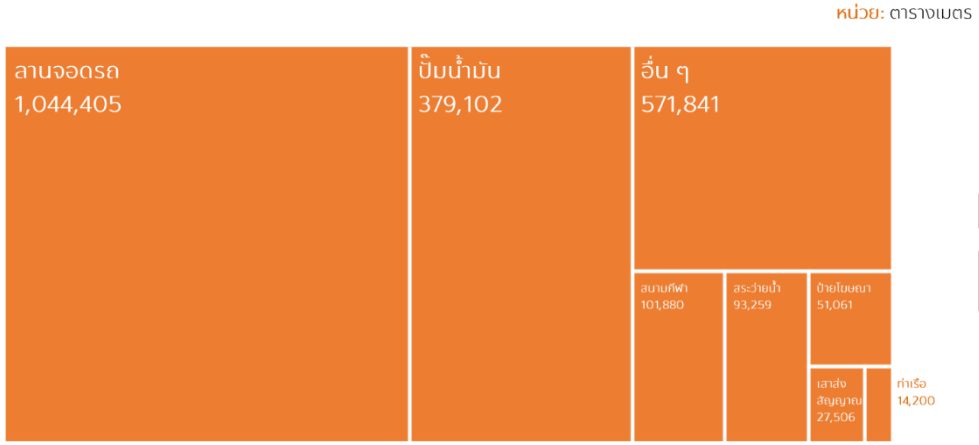
สำหรับอาคารที่ไม่ใช่เพื่อการอยู่อาศัย พบว่า เป็นการได้รับอนุญาตให้ก่อสร้างอาคารเพื่อการอุตสาหกรรมและโรงงาน 8,401,283 ตารางเมตร อาคารเพื่อการพาณิชย์ 7,077,600 ตารางเมตร โรงแรม 3,040,235 ตารางเมตร อาคารเพื่อการเกษตร 1,777,723 ตารางเมตร โรงไฟฟ้าและอาคารระบบพลังงานแสงอาทิตย์ 1,321,527 ตารางเมตร และอาคารสำนักงาน 1,214,890 ตารางเมตร ส่วนอาคารประเภทอื่น ๆ ได้รับอนุญาตให้ก่อสร้างเป็นพื้นที่รวม 2,004,057 ตาราง ดังภาพประกอบ 3 สิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่วยวัดเป็นความยาว ตลอดปี 2564 มีการอนุญาตให้ก่อสร้างสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่อยวัดเป็นความยาวรวมเป็นความยาวรวมทั้งสิ้น 2,170,831 เมตร โดยเป็นการก่อสร้างท่อหรือทางระบายน้ำ 1,291,021 เมตร รั้วหรือกำแพง 465,444 เมตรา ถนนหรือทางรถไฟ 324,680 เมตร เขื่อนหรือคันดิน 40,430 เมตร สะพาน 24,519 เมตร และสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่วยวัดเป็นความยาวชนิดอื่นๆ 24,737 เมตร (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังภาพประกอบ 4



ที่มา: กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, 2564

ภาพประกอบ 4 ความยาวสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างจำแนกตามชนิดของสิ่งก่อสร้าง

สิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่วยวัดเป็นพื้นที่ ตลอดปี 2564 มีการอนุญาตให้ก่อสร้างสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่วยวัดเป็นพื้นที่ เป็นพื้นที่รวมทั้งสิ้น 2,283,254 ตารางเมตร โดยเป็นการก่อสร้างลานจอดรถ 1,044,405 ตารางเมตร ปั๊มน้ำมัน 379,102 ตารางเมตร สนามกีฬา 101,880 ตารางเมตร สระว่ายน้ำ 93,259 ตารางเมตร ป้ายโฆษณา 51,061 ตารางเมตร เสาส่งสัญญาณ 27,506 ตารางเมตร ท่าเรือ 14,200 ตารางเมตร และสิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนและมีหน่วยวัดเป็นพื้นที่อื่นๆ 571,841 ตารางเมตร (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังภาพประกอบ 5



ที่มา: กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม, 2564

ภาพประกอบ 5 พื้นที่สิ่งก่อสร้างที่ไม่ใช่อาคารโรงเรือนที่ได้รับอนุญาตให้ก่อสร้างจำแนกตามชนิดของสิ่งก่อสร้าง

สิ่งก่อสร้างที่เป็นอาคารโรงเรือนเฉพาะเพื่อการพักอาศัย สามารถแบ่งออกได้เป็น 2 ประเภท คือ 1) สิ่งก่อสร้างประเภทของบ้านพักอาศัยแนวราบ และ 2) สิ่งก่อสร้างประเภทของบ้านอาศัยแนวดิ่งหรือแนวตั้ง

สิ่งก่อสร้างประเภทของบ้านพักอาศัยแนวราบหรืออสังหาริมทรัพย์แนวราบ มีลักษณะการการก่อสร้างอาคารที่ไม่สูงมากนัก และมักมีการขยายออกเป็นแนวกว้างหรือทางข้าง โดยส่วนใหญ่สามารถพบเจอได้ในในบริเวณรอบของตัวเมือง นอกจากนี้ที่อยู่อาศัยแนวราบยังเป็นที่นิยมกันอย่างแพร่หลาย สิ่งก่อสร้างประเภทของที่พักอาศัยแนวราบ ได้แก่ 1) บ้านเดี่ยว การออกแบบของบ้านเดี่ยวจะมีลักษณะที่ไม่มีส่วนผนังติดกับบ้านหลังอื่น ๆ เลย อีกทั้งยังมีพื้นที่ใช้สอยที่ไม่รวมตัวบ้าน และนิยมติดตั้งรั้วรอบบริเวณพื้นที่ใช้สอย นอกจากนี้หากบ้านเดี่ยวอยู่ในโครงการ ส่วนใหญ่จะมีพื้นที่ส่วนกลาง เช่น สระว่ายน้ำ สวนหย่อม และฟิตเนส เพื่ออำนวยความสะดวกสบายให้แก่ผู้อยู่อาศัย 2) ทาวน์เฮ้าส์ มีลักษณะการออกแบบเป็นตึกแถวซึ่งผนังของบ้าน จะติดกับบ้านหลังอื่น ๆ สมัยก่อนผู้คนมักชอบเรียกว่า “บ้านแถว” โดยปกติจะมีพื้นที่จอดรถและส่วนหย่อมหน้าบ้าน 3) ทาวน์โฮม มีลักษณะรูปแบบคล้ายกับทาวน์เฮ้าส์แต่จะความทันสมัยดูร่วมสมัยมากขึ้น และยังให้ความรู้สึกอบอุ่นเหมือนเป็นเจ้าของบ้านมากยิ่งขึ้น 4) บ้านแฝด มีลักษณะบ้านที่ถูกสร้างมาเป็นคู่ โดยมีผนังหรือส่วนใดส่วนหนึ่งของบ้านติดกันด้านหนึ่ง ซึ่งการออกแบบคล้ายกับบ้านเดี่ยวแต่มีเนื้อที่ของบ้านน้อยกว่า และ 5) อาคารพาณิชย์ ถูกออกแบบมาเพื่อให้สามารถอยู่อาศัย และประกอบธุรกิจได้ นิยมก่อสร้างในชุมชนเมือง สามารถเรียกอีกชื่อหนึ่งว่า “ตึกแถว” เป็นต้น

สิ่งก่อสร้างประเภทของบ้านอาศัยแนวดิ่งหรือแนวตั้ง มีลักษณะการก่อสร้างโดยการก่อชั้นสูงขึ้นไปต่อกันเป็นชั้น ๆ และมีการแบ่งจำนวนห้องภายในอาคารสำหรับการอยู่อาศัย เหมาะสำหรับผู้ที่ต้องการความสะดวกสบายในการใช้ชีวิตในตัวเมืองใหญ่ เนื่องจากมักก่อสร้างในทำเลที่สามารถเดินทางได้สะดวก สิ่งก่อสร้างประเภทของบ้านพักอาศัยแนวดิ่งหรือแนวตั้ง ได้แก่ 1) คอนโดมิเนียม มีลักษณะการก่อสร้างเป็นชั้น ๆ ต่อขึ้นไป และในแต่ละชั้นจะแบ่งเป็นห้องเพื่ออยู่อาศัย โดยอาจจะมีจำนวนห้องมากน้อยขึ้นอยู่ขนาดของห้อง และกลุ่มลูกค้าของทางโครงการ นอกจากนี้ยังมีสถานที่ส่วนกลางเพื่อเพิ่มความสะดวกสบายให้กับผู้อยู่อาศัย เช่น ห้องรับแขก ฟิตเนส สระว่ายน้ำ ซึ่งในบางโครงการอาจจะมีห้องดูหนัง ห้องสมุด หรืออื่น ๆ เพื่อตอบโจทย์ของผู้อยู่อาศัย 2) โรงแรม มีการก่อสร้างเป็นชั้น ๆ ต่อขึ้นไปคล้ายกับคอนโดมิเนียม แต่จะมีการแบ่งโซนประเภทต่าง ๆ ไว้ใช้ตามกิจกรรมของลูกค้าที่เข้ามาพัก เช่น ห้องอาหาร ห้องประชุม ห้องพักอาศัย ฟิตเนส และสระว่ายน้ำ นอกจากนี้ยังมีการแบ่งประเภทของราคาห้องพัก และระยะเวลาของผู้เข้าพักอีกด้วย 3) อพาร์ตเม้นต์ มีลักษณะคล้ายกับโรงแรม มักตั้งอยู่ในทำเลที่ดี สามารถเดินทางได้สะดวก ภายในอาคารจะแบ่งห้องสำหรับผู้เข้าพักตามขนาดของอาคาร

จากข้อมูลภาพประกอบ 2 และภาพประกอบ 3 พบว่าการขออนุญาตก่อสร้างทั้งหมด 315,496 แห่ง (230,828 ราย) และจังหวัดที่มีการก่อสร้างมากที่สุดคือจังหวัดกรุงเทพมหานครจำนวน 25,346 แห่ง (16,314 ราย) คิดเป็นร้อยละ 8.03 ของประเทศ เมื่อพิจารณาเฉพาะกลุ่มที่พักอาศัยทั้งหมด 39,186,978 ตารางเมตรซึ่งเป็นบ้านพักอาศัยแนวราบถึง 30,934,610 ตารางเมตร คิดเป็นร้อยละ 78.94 ดังนั้นผู้วิจัยจึงให้ความสำคัญเลือกกลุ่มตัวอย่างเป็นบ้านพักอาศัยแนวราบและเป็นจังหวัดกรุงเทพมหานครที่มีการก่อสร้างมากที่สุดและเป็นศูนย์กลางเศรษฐกิจของประเทศด้วย

การก่อสร้างบ้านพักอาศัยแนวราบได้ขึ้นอยู่กับหลายปัจจัยสำคัญ (วิศว์ ดวงแสงทอง 2559) เช่น การศึกษาความเป็นไปได้ของโครงการ การวางแผนดำเนินโครงการ การประมาณราคา การดำเนินการก่อสร้าง วางแผนทางการเงินและการเตรียมงบประมาณสำหรับการดำเนินการ ล้วนต้องสอดคล้องกัน ในระหว่างขั้นตอนการดำเนินการของโครงการ การประเมินต้นทุนจะถูกใช้เพื่อกำหนดงบประมาณโครงการ และเลือกวิธีการทำงานที่เหมาะสม ในการควบคุมโครงการ ให้ใช้งบประมาณที่ตั้งไว้เป็นเครื่องมือในการบริหารงบประมาณตามแผน นอกจากนี้ค่าใช้จ่ายสุดท้ายตอนจบโครงการสำหรับคิดงานเพิ่มลดจากการเปลี่ยนแปลงการก่อสร้างระหว่างดำเนินโครงการอีกด้วย ด้วยเหตุนี้การประมาณราคาจึงเป็นทักษะที่สำคัญอย่างยิ่งในการบริหารโครงการในอุตสาหกรรมการก่อสร้าง ที่หลีกเลี่ยงไม่ได้ว่าความแม่นยำของการประมาณราคาจะเป็นปัจจัยหนึ่งที่จะส่งผลต่อระดับความสำเร็จที่โครงการก่อสร้างทำได้ การเริ่มโครงการบัญชีแสดงการประมาณราคาค่าก่อสร้าง (Bill of Quantity, BOQ) มีความจำเป็นต่อการวิเคราะห์ความเป็นไปได้ของโครงการ ในการวางแผนการประมาณราคาใช้ในการวางแผนการเงินจัดเตรียมงบประมาณในการดำเนินการ ในการดำเนินโครงการใช้การประมาณราคาในการตั้งงบประมาณของโครงการตลอดจนเลือกวิธีการทำงานที่เหมาะสม ในการควบคุมโครงการงบประมาณที่ถูกจัดเตรียมไว้จะถูกนำมาใช้เป็นเครื่องมือในการควบคุมให้งบประมาณเป็นไปตามที่วางแผนไว้ และในการปิดโครงการราคาที่ประมาณการไว้จะถูกนามาเป็นฐาน ในการคิดงานลด-งานเพิ่ม

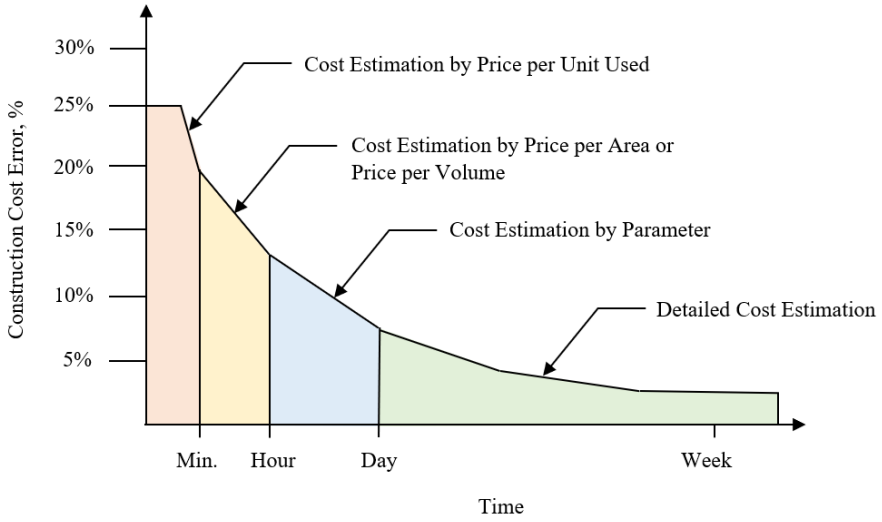
วิธีการประมาณราคาก่อสร้างแบ่งเป็นประเภทใหญ่ๆ ได้ 2 ประเภท ได้แก่ 1) การประมาณราคาอย่างหยาบหรือการประมาณราคาโดยสังเขป จะเป็นการประมาณราคาอย่างหยาบๆเพื่อศึกษาความเป็นไปได้ขั้นต้นก่อน ใช้เวลาน้อย แต่จะมีความคลาดเคลื่อนสูง และ 2) การประมาณราคาอย่างละเอียดเป็นการคำนวณปริมาณงานและราคางานอย่างละเอียด โดยทั่วไปจะทำการประมาณราคาแบบนี้เมื่อมีแบบก่อสร้างและข้อกำหนดการก่อสร้างแล้ว การประมาณราคาอย่างละเอียดนี้จะให้ผลค่อนข้างแม่นยำ แต่ใช้เวลาในการประมาณการนานกว่าการประมาณราคาขั้นต้น การประมาณราคาอย่างละเอียดโดยการแสดงบัญชีปริมาณวัสดุและราคาค่าก่อสร้าง จะมีความแม่นยำค่อนข้างสูงแต่ก็ใช้เวลาค่อนข้างมาก (สิทธิกร สิทธิการกูล 2564) แต่ในบางครั้งการเริ่มโครงการอาจต้องการข้อมูลเบื้องต้นในการตัดสินใจเริ่มโครงการเพื่อวิเคราะห์ความเป็นไปได้ของโครงการด้วย การประมาณราคา

แบบอย่างหยาบหรือโดยสังเขป เนื่องจากใช้เวลาประมาณราคาน้อยกว่า มีความรวดเร็วมากกว่า แต่มีความคลาดเคลื่อนสูง โดยคำนวณจากปริมาณเนื้องานทั้งหมดของวัสดุก่อสร้างที่ใช้แต่ละประเภท แล้วนำไปประมาณเพื่อหาค่าวัสดุ ค่าแรง ค่าดำเนินการต่าง ๆ ที่เกี่ยวข้องกับการก่อสร้าง แล้วรวมยอดเป็นค่าก่อสร้างอาคารและสิ่งปลูกสร้างทั้งหมด ซึ่งผลที่ได้จากการประมาณราคาก่อสร้างวิธีนี้ จะใกล้เคียงกับความเป็นจริงมากที่สุด ซึ่งปัญหาของผู้รับเหมาก่อสร้างส่วนใหญ่ที่ต้องเจอคือต้องใช้วิศวกรและสถาปนิกเร่งทำแบบก่อสร้าง ถอดแบบและประมาณราคา จัดทำบัญชีแสดงปริมาณวัสดุและราคาค่าก่อสร้าง (BOQ) เพื่อยื่นประมูลต่อเจ้าของบ้านหรือเจ้าของโครงการ ซึ่งต้องใช้ทรัพยากรมนุษย์และวัสดุอุปกรณ์เป็นจำนวนมาก รวมถึงใช้ระยะเวลาที่ค่อนข้างนานในแต่ละขั้นตอนหรือกระบวนการกว่าจะได้มาซึ่งแบบก่อสร้างและการถอดแบบและประมาณราคา ถ้าเป็นบริษัทไม่ได้ใหญ่มากโอกาสในการแข่งขันก็น้อยเนื่องจากแต่ละกระบวนการต้องใช้เวลาและกำลังคนในการทำแบบก่อสร้างและประมาณราคาพอสมควร กอปรกับบริษัทเล็กมีข้อจำกัดเรื่องของกำลังคน และ ทรัพยากรต่างๆ เช่น โปรแกรมคอมพิวเตอร์ที่ทันสมัยและสะดวกต่อการทำแบบ ต่อการถอดแบบและประมาณราคา เป็นต้น จนทำให้เลยกำหนดการยื่นประมูลงาน เสียโอกาสในการที่จะได้งาน เสียทรัพยากรต่างๆ ที่ทำไปแล้วแต่ไม่ได้สามารถยื่นได้ทันเวลา และเจ้าของโครงการก็ขาดโอกาสที่จะได้เห็นไอเดียและราคาเพื่อการแข่งขันที่หลากหลาย หรือบางครั้งพยายามเร่งเพื่อให้ทันเวลาโดยการทุ่มเททรัพยากรต่างๆ ที่มีแล้วแต่ไม่ได้งานก็ทำให้ทรัพยากรต่างๆ ที่ลงทุนไปแล้วเสียเปล่าไม่เกิดประโยชน์

ปัญญาประดิษฐ์ (Artificial Intelligence, AI) และการเรียนรู้ของเครื่องจักร (Machine Learning, ML) ได้รับการนำมาใช้ในหลายศาสตร์และแขนงอุตสาหกรรมและธุรกิจ (Cioffi et al., 2020; Kraus et al., 2020) การศึกษาที่มีผลกระทบได้แสดงให้เห็นถึงความทนทานของวิธีการปัญญาประดิษฐ์หรือการเรียนรู้ของเครื่องจักรในการทำนายและจำแนกปัจจัยต่าง ๆ เช่นอัตราดอกเบี้ย, อัตราดอกเบี้ยสินเชื่อบ้าน, ราคา เป็นต้น ในภาคอสังหาริมทรัพย์ (Kang et al., 2020; Mu et al., 2014) มีการทำนายราคาอสังหาริมทรัพย์ซึ่งเป็นปัญหาที่ซับซ้อนและไม่เชิงเส้น การทำนายได้รับผลกระทบจากหลายคุณสมบัติที่เป็นตัวแปรโดยตรงและอ้อม เช่น ปีก่อสร้าง, พื้นที่ห้องชุด, และอื่น ๆ (Ferlan et al., 2017).

จากวิธีดั้งเดิมของการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ ไม่ว่าจะเป็นการประมาณราคาอย่างหยาบหรือการประมาณราคาอย่างละเอียดล้วนมีความจำเป็นต้องใช้ทรัพยากรที่เป็นบุคคล เช่น วิศวกร สถาปนิก และเจ้าหน้าที่ประมาณราคา เป็นต้น รวมถึงทรัพยากรอื่นๆ เช่น คอมพิวเตอร์ เครื่องพิมพ์ กระดาษ และอื่นๆ อีกมากมาย เป็นต้น รวมถึงระยะเวลาที่ใช้เวลาค่อนข้างนานถึงจะได้ค่าที่ละเอียดและแม่นยำ ผู้วิจัยจึงสนใจใช้วิธีการใหม่โดยการทำให้คอมพิวเตอร์เรียนรู้ได้ด้วยตนเองโดยใช้ข้อมูลหรือการใช้งานเสมือนเป็นสมองของปัญญาประดิษฐ์ใช้การเรียนรู้ของเครื่องเอ็กซ์ตรีม (Extreme Machine Learning, XML) หรือการเรียนรู้ของเครื่องจักรขั้นสูงในการสร้าง

แบบจำลองเพื่อการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ มาใช้ในการประมาณราคาจากพื้นที่ของบ้านพักอาศัยเพื่อให้ได้ราคาค่าก่อสร้างโดยที่ไม่ต้องให้สถาปนิกและวิศวกรทำการออกแบบเลยแต่ให้ค่าใกล้เคียงกับการประมาณราคาดั้งเดิมที่มีความคลาดเคลื่อนประมาณร้อยละ 10 ของการประมาณราคาอย่างละเอียด ดังภาพประกอบ6



ภาพประกอบ 6 ความคลาดเคลื่อนในการประมาณราคาค่าก่อสร้างกับระยะเวลา
ดัดแปลงจาก (วิศว์ ดวงแสงทอง 2559; สิทธิกร สิทธิการกูล 2564)

การเรียนรู้ของเครื่องเป็นสาขาย่อยของปัญญาประดิษฐ์ที่ทำงานร่วมกับอัลกอริทึมและเทคโนโลยีในการดึงข้อมูลที่เป็นประโยชน์จากข้อมูล วิธีการเรียนรู้ของเครื่องมีความเหมาะสม ในการคำนวณโดยใช้ข้อมูลและพยายามประมวลผลข้อมูลด้วยตนเองจะเป็นไปไม่ได้โดยไม่ได้รับการสนับสนุนจากการเรียนรู้ของเครื่องในวิทยาการคอมพิวเตอร์พยายามแก้ปัญหาปัญหาตามอัลกอริทึมมากกว่าทางคณิตศาสตร์ล้วนๆ ดังนั้นจึงขึ้นอยู่กับการสร้างอัลกอริทึมที่อนุญาตให้การเรียนรู้ของเครื่องเรียนรู้เป็นวิธีการคาดการณ์แบบอัลกอริทึมที่สามารถประมวลผลข้อมูลเชิงปริมาณเพื่อคาดการณ์ได้ การเรียนรู้ของเครื่องทำให้เกิดการเปลี่ยนแปลงครั้งใหญ่ในอุตสาหกรรมต่างๆ และได้กลายเป็นเครื่องมือที่ทรงพลังในภาคการก่อสร้าง เนื่องจากสามารถทำให้กระบวนการทำงานเป็นอัตโนมัติได้ เทคโนโลยีการเรียนรู้ของเครื่องมีบทบาทสำคัญ โดยเฉพาะอย่างยิ่งเมื่อการประมวลผลข้อมูลจำนวนมากนำมูลค่าเพิ่มที่สำคัญมาสู่การประหยัดเวลาและเพิ่มทรัพยากรการประมวลผลสูงสุด ด้วยวิธีนี้การเรียนรู้ของเครื่องอาจเป็นวิธีที่เหมาะสมสำหรับอุตสาหกรรมการก่อสร้างในการคาดการณ์

จากที่กล่าวมาปัญหาที่เกิดขึ้นหากผู้ประกอบการ ผู้บริหารโครงการ หรือเจ้าของบ้านสามารถประมาณราคาก่อสร้างได้ใกล้เคียงกับการประมาณอย่างละเอียด ลดระยะเวลาที่ใช้ในการประมาณราคาและลดการใช้วิศวกร สถาปนิก หรือเจ้าหน้าที่เพื่อทำการประมาณได้ ก็จะเป็น

ประโยชน์ต่อการประมาณราคาและต่อองค์กร ผู้วิจัยมีความประสงค์ในการนำเสนอการพัฒนาแบบจำลองการทำนายการประมาณราคาบ้านพักอาศัยแนวราบโดยใช้เทคนิคการเรียนรู้ของเครื่องเอ็กซ์ตรีมเป็นเครื่องมือการทำนายการประมาณราคาบ้านพักอาศัยแนวราบจากข้อมูลต่างๆ ที่เป็นตัวแปรต่างๆ มาช่วยในการประมาณราคา ซึ่งสามารถสอนให้คอมพิวเตอร์เรียนรู้การทำงานได้จึงอาจเป็นทางเลือกที่ทำให้การประมาณราคามีความรวดเร็วมากขึ้น มีผลความคลาดเคลื่อนอยู่ในเกณฑ์ที่ยอมรับได้ลดการใช้ทรัพยากร ทำให้บริษัทหรือองค์ขนาดเล็กมีความสามารถในการแข่งขันมากขึ้นมีโอกาสเติบโตหรือได้ส่วนแบ่งการตลาดมากขึ้น และยังจะส่งผลให้เจ้าของโครงการมีตัวเลือกได้มากขึ้นได้ราคาที่มีความเหมาะสมมากขึ้น ดังนั้นงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อเสนอวิธีการใหม่ในการประมาณราคาบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีมแทนวิธีการดั้งเดิมที่ใช้มนุษย์ในการดำเนินการประมาณราคา รวมถึงเปรียบเทียบประสิทธิภาพวิธีการใหม่และวิธีการดั้งเดิมด้วย

ความมุ่งหมายของการวิจัย

1. เพื่อเสนอวิธีการใหม่ในการประมาณราคาบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม
2. เพื่อเปรียบเทียบประสิทธิภาพวิธีการใหม่กับวิธีการดั้งเดิมในทำนายการประมาณราคาบ้านพักอาศัยแนวราบ
3. เพื่อค้นหาวิธีการทำนายที่รวดเร็วและเหมาะสมที่สุดในการประมาณราคาบ้านพักอาศัยแนวราบ

ขอบเขตของการวิจัย

1. บ้านพักอาศัยแนวราบที่ออกแบบและรวบรวมโดยสำนักการโยธา กรุงเทพมหานคร ปี 2559 เป็นฐานข้อมูลเพื่อให้เครื่องได้ทำการเรียนรู้และฝึกฝน
2. ข้อมูลราคาเป็นข้อมูลเฉพาะราคาค่าวัสดุ ค่าแรง และค่าใช้จ่ายที่เกิดขึ้นในงานก่อสร้างไม่รวม ค่าดำเนินการ ภาษีมูลค่าเพิ่ม ที่ดินที่จะปลูกสร้างบ้านพักอาศัย อุปกรณ์เครื่องใช้ไฟฟ้า และของตกแต่งทั้งสิ้น
3. ราคาค่าก่อสร้างบ้านพักอาศัยแนวราบอ้างอิงจากกลุ่มงานประมาณราคา สำนักการโยธา กรุงเทพมหานคร ปี 2559

นิยามศัพท์เฉพาะ

การประมาณราคาค่าก่อสร้าง หมายถึง การคำนวณประมาณราคาค่าก่อสร้างทางตรง ได้แก่ ค่าวัสดุ ค่าแรงงาน ค่าเครื่องมือเครื่องจักร โดยยังคงความแข็งแรงตามมาตรฐาน และสามารถใช้สอยได้

บ้านพักอาศัยแนวราบ หมายถึง อสังหาริมทรัพย์ประเภทที่อยู่อาศัยที่มีการก่อสร้างตั้งอยู่บนที่ดินที่ได้รับการจัดสรร และมีเจ้าบ้านหรือผู้มีถือกรรมสิทธิ์เพียงคนเดียวในที่ดินนั้น ได้แก่ บ้านเดี่ยว บ้านแฝด ทาวน์เฮ้าส์ และทาวน์โฮม

เครื่องจักรเรียนรู้เอ็กซ์ตรีม หมายถึง อัลกอริทึมของคอมพิวเตอร์ที่มีการพัฒนาซึ่งเป็นส่วนหนึ่งของการเรียนรู้ของเครื่อง โดยอัลกอริทึมสร้างแบบจำลองทางคณิตศาสตร์จากข้อมูลตัวอย่างแบบบ้านพักอาศัยแนวราบเพื่อที่จะทำนายการประมาณราคาบ้านพักอาศัยแนวราบ การเรียนรู้ของเครื่องพัฒนามาจากการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยแบบจำลองที่สร้างมาจากชุดข้อมูลตัวอย่างแบบบ้านพักอาศัยแนวราบที่นำเข้าเพื่อการทำนายหรือตัดสินใจ

แบบจำลองการทำนาย หมายถึง กระบวนการทางคณิตศาสตร์หรือวิทยาการคอมพิวเตอร์ที่ใช้ข้อมูลที่มีอยู่เพื่อทำนายผลลัพธ์หรือเหตุการณ์ในอนาคต โดยใช้ข้อมูลที่มีอยู่ในการสร้างแบบจำลองทางคณิตศาสตร์ที่สามารถประมวลผลและวิเคราะห์ข้อมูลเหล่านั้นเพื่อทำนายผลลัพธ์ที่เป็นไปได้ในสถานการณ์ที่แตกต่างกัน

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้วิธีการใหม่ในการประมาณราคาบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม ซึ่งเป็นประโยชน์ต่อบริษัทที่รับออกแบบและก่อสร้างบ้านพักอาศัยแนวราบ เจ้าของบ้าน เจ้าของโครงการ วิศวกร สถาปนิก เจ้าหน้าที่ถอดแบบและประมาณราคา และนักพัฒนาอสังหาริมทรัพย์ที่จะได้คาดการณ์ราคาบ้านพักอาศัยแนวราบเพื่อกำหนดเป้าหมาย วางกลยุทธ์การตลาดหรือศึกษาความเป็นไปได้ที่จะก่อสร้างบ้านหรือโครงการได้

2. ได้ทราบประสิทธิภาพวิธีการใหม่เทียบกับวิธีการดั้งเดิมที่ใช้คนในการถอดแบบและประมาณราคาบ้านพักอาศัยแนวราบ และเป็นแนวทางในการตัดสินใจเลือกใช้วิธีการ

3. ได้ทราบวิธีการทำนายที่รวดเร็วและเหมาะสมที่สุดในการประมาณราคาบ้านพักอาศัยแนวราบ ทำให้สามารถลดต้นทุนค่าใช้จ่ายในการประมาณราคาค่าก่อสร้าง รวมทั้งลดค่าใช้จ่ายบุคลากร วิศวกร หรือเจ้าหน้าที่เพื่อการประมาณราคา และ ลดระยะเวลาในการประมาณราคา ทำให้สามารถประมาณราคาได้รวดเร็วและแม่นยำมากขึ้น

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

งานวิจัยเรื่องการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม (The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine) ผู้วิจัยได้ทำการค้นหาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง เพื่อเป็นแนวทางในการศึกษาดังต่อไปนี้

1. แนวคิดเกี่ยวกับการประมาณราคาค่าก่อสร้างบ้านพักอาศัย
2. แนวคิดเกี่ยวกับเทคโนโลยีการเรียนรู้ของเครื่อง (Machine Learning, ML)
3. การเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม (Extreme Learning Machine, XLM)
4. โครงข่ายประสาทเทียม (Artificial Neural Network, ANN)
5. ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM)
6. การวิเคราะห์ความถดถอย (Multiple Linear Regression, MLR)
7. ต้นไม้ตัดสินใจ (Decision Tree, DT)
8. แรนดอมฟอเรสต์ (Random Forest, RF)
9. งานวิจัยที่เกี่ยวข้อง

แนวคิดเกี่ยวกับการประมาณราคาค่าก่อสร้างบ้านพักอาศัย

1. การประมาณราคาก่อสร้าง

การประมาณราคา (Cost Estimate) หมายถึง การคำนวณหาปริมาณวัสดุ ค่าแรงและค่าดำเนินการที่ราคาใกล้เคียงกับค่าใช้จ่ายจริงมากที่สุด ในการแยกรายการวัสดุ ค่าแรง ค่าใช้จ่ายเครื่องมือเครื่องจักร และค่าใช้จ่ายอื่นที่เกี่ยวข้องกับงานโดยมีผลกับตัวแปรตามในด้านระยะเวลาของการทำงาน ดังนั้นการประมาณราคาจึงไม่ใช่ราคาที่แท้จริง แต่อาจใกล้เคียงกับราคาจริง

ประเภทของการประมาณราคาก่อสร้างแบ่งเป็นประเภทใหญ่ๆ ได้ 2 ประเภท (วิศว์ ดวงแสงทอง 2559) ดังนี้

1) การประมาณราคาอย่างหยาบ (Preliminary Estimating) หรือการประมาณราคาโดยสังเขป จะเป็นการประมาณราคาอย่างหยาบๆเพื่อศึกษาความเป็นไปได้ขั้นต้นก่อน เนื่องจากใช้เวลาน้อย แต่จะมีความคลาดเคลื่อนสูง

2) การประมาณราคาอย่างละเอียด (Detailed Estimating) เป็นการคำนวณปริมาณงานและราคางานอย่างละเอียด โดยทั่วไปจะทำการประมาณราคาแบบนี้เมื่อมีแบบและข้อกำหนดการก่อสร้างแล้ว การประมาณราคาอย่างละเอียดนี้จะให้ผลค่อนข้างแม่นยำ แต่ใช้เวลาในการประมาณการนานกว่าการประมาณราคาขั้นต้น

นอกจากการจำแนกประเภทวิธีการประมาณราคาด้วยความละเอียดในการประมาณการแล้วยังสามารถจำแนกประเภทวิธีการประมาณราคาด้วยการจำแนกตามผลิตภัณฑ์ (Product Based) และการจำแนกตามกระบวนการ (Process Based) (วิศว์ ดวงแสงทอง 2559)

รูปแบบการประมาณราคาแบบขึ้นกับผลิตภัณฑ์ (Product Based Cost Modeling) โดยที่ไม่ได้นำวิธีการในการก่อสร้างมาใช้ในการประมาณการ ซึ่งวิธีการในการประมาณการแบบขึ้นกับผลิตภัณฑ์นี้ สามารถแสดงดังตาราง 1 และตาราง 2

ตาราง 1 การประมาณราคาแบบขึ้นกับผลิตภัณฑ์แบบต่างๆ ซึ่งมีการใช้อย่างแพร่หลาย

วิธีการ	**รายละเอียด**
การประมาณราคาโดยไม่ต้องอาศัยการหาจำนวน (Judgment Without quantification)	เป็นการคิดราคาโดยอาศัยความเชี่ยวชาญ (Expertise) และปรีชาญาณ (Intuition) ของผู้ประมาณราคาในการประมาณราคา
หน่วยการใช้งาน (Functional Unit)	เป็นการตีราคาที่เหมาะสมตามการใช้งานของสิ่งปลูกสร้าง เช่น ราคาต่อพื้นที่จอดรถ
ราคาต่อพื้นที่ (Superficial)	ราคาต่อพื้นที่ เช่น บาทต่อตารางเมตร
การอินเทอร์โพเลท (Interpolation Method)	การใช้ข้อมูลจากโครงการที่ผ่านมามาทำการอินเทอร์โพเลทเพื่อหาราคาของโครงการปัจจุบัน
การวิเคราะห์ส่วนประกอบ (Elemental Analysis)	เกิดจากผลรวมของต้นทุนในแต่ละส่วนประกอบที่ได้ทำการเลือกในโครงการนั้นๆ
การประมาณการโดยคิดเฉพาะรายการที่สำคัญ (Significant Items)	ประมาณการโดยคิดเฉพาะรายการที่สำคัญต่องาน (Significant to Work)
การหาปริมาณโดยการประมาณ (Approximate Quantities)	เป็นการวัดและตีราคาโดยคิดแต่ละรายการก่อสร้างรวมเป็นกลุ่มย่อยๆ
การหาปริมาณโดยละเอียด (Detailed Quantities)	การคิดราคาแยกเป็นรายการย่อยๆ เช่น BOQ

ที่มา: วิศว์ ดวงแสงทอง (2559)

ตาราง 2 การประมาณราคาแบบขึ้นกับผลิตภัณฑ์ที่มีการใช้ไม่แพร่หลาย

วิธีการ	รายละเอียด
ราคาต่อหน่วยปริมาตร (Cube)	ราคาต่อหน่วยปริมาตร เช่น บาทต่อลูกบาศก์เมตร
Functional Area	คล้ายกับแบบราคาต่อพื้นที่ แต่มีการจำแนกพื้นที่ตามการใช้สอย ซึ่งการคิดราคาแบบนี้มีสมมติฐานว่าแต่ละพื้นที่การใช้สอยจะมีราคาไม่เท่ากัน
Superficial Perimeter	เป็นการใช้ราคาต่อพื้นที่ร่วมกับเส้นรอบวง เพื่อเพิ่มความแม่นยำในการคิดราคา
Story Enclosure	ใช้พื้นที่ล้อมรอบสิ่งปลูกสร้างในการคิดราคา เช่น พื้นที่กาแพง พื้นที่เพดาน ที่ล้อมรอบอาคารอยู่
Regression Analysis	เป็นการใช้ข้อมูลในอดีตมาทำการวิเคราะห์แบบถดถอย (Regression Analysis) เพื่อทำการหาความสัมพันธ์ทางคณิตศาสตร์ขององค์ประกอบหลักๆของโครงการ เพื่อทำการประมาณราคา
Probabilistic Treatments	ใช้ทฤษฎีความน่าจะเป็นร่วมในการคิดราคา ซึ่งจะมีการวิเคราะห์ความเสี่ยงที่อาจกระทบกับต้นทุนในการประมาณราคาด้วย
Expert Systems	ใช้ระบบผู้เชี่ยวชาญ (Expert System) โดยจะเป็นการดึงเอาความเชี่ยวชาญจากผู้เชี่ยวชาญออกมาใช้ ซึ่งความรู้ของผู้เชี่ยวชาญ จะถูกเก็บไว้ในคอมพิวเตอร์เพื่อที่จะสามารถนำมาใช้ในการประมาณราคาแบบผู้เชี่ยวชาญภายหลังได้

ความแม่นยำของรูปแบบในการประมาณราคาแบบขึ้นกับผลิตภัณฑ์นั้น จะอยู่ในช่วง ±10% ถึง ±40% (วิศว์ ดวงแสงทอง 2559)

รูปแบบการประมาณราคาแบบขึ้นกับกระบวนการ (Process Based Cost Modeling) เป็นการประมาณการราคาก่อสร้างโดยคำนึงถึงปัจจัยด้านกระบวนการทำงานร่วมด้วย ตัวอย่างวิธีการในการประมาณการราคาแบบขึ้นกับกระบวนการ สามารถแสดงดังตาราง 3 ดังนี้

ตาราง 3 ตัวอย่างวิธีการประมาณราคาแบบขึ้นกับกระบวนการ

วิธีการ	รายละเอียด
Operation Bill of Quantities (Skoyles, 1968)	เป็นการประมาณราคาที่ปรับปรุงจาก Bill of Quantities โดยการเพิ่มปัจจัยเกี่ยวกับวิธีการในการดำเนินงานของแต่ละรายการของ BOQ เข้าไปด้วย
Pre-established Critical Path Method Networks (Bowen, 1993)	เป็นรูปแบบในการประมาณราคาซึ่งใช้ Critical Path และ Injection Algorithm ซึ่งจะทำให้ทราบระยะเวลาในการก่อสร้าง รูปแบบแบบนี้จะใช้ในการทำการประมาณการต้นทุนในระยะเริ่มต้นโครงการ
Simulation (Bennett & Ormerod, 1984)	เป็นการจำลองการก่อสร้างที่จะเกิดขึ้นจริงในโปรแกรมคอมพิวเตอร์ เพื่อหาต้นทุนและระยะเวลาที่ใช้ในโครงการนั้นๆ

2. แนวคิดและทฤษฎีที่เกี่ยวข้องกับการบริหารงานก่อสร้าง

การบริหารงานก่อสร้าง (ConstructionManagement) เป็นเทคนิคการบริหารโครงการที่มีการพัฒนาจากศาสตร์สาขาหนึ่ง ในช่วงปี พ.ศ. 2503 ใช้ในการดำเนินการเพื่อให้บรรลุวัตถุประสงค์ของแผนงานที่กำหนดไว้ และดำเนินการให้สำเร็จภายในเวลาที่กำหนด

2.1. ความหมายของการบริหารงานก่อสร้าง

วิสูตร จิระดำเกิง (2563) ได้ให้ความหมายไว้ว่า การบริหารงานก่อสร้าง คือการจัดการการใช้ทรัพยากรต่าง ๆ ที่มีอยู่อย่างเหมาะสมและสมบูรณ์ที่สุด โดยทรัพยากรในที่นี้หมายถึงบุคลากร รวมถึงความเชี่ยวชาญและความสามารถที่มีอยู่ ความร่วมมือของทีมงาน เครื่องมือ เครื่องใช้ และสิ่งอำนวยความสะดวกต่าง ๆ ตลอดจนข้อมูล ระบบงาน เทคนิค เงินทุน และเวลา เพื่อให้การบริหารโครงการบรรลุวัตถุประสงค์ตามที่ได้มีการวางแผนไว้

2.2 ประเภทของงานก่อสร้าง

ประเภทของงานก่อสร้าง (Types of Construction Project) สามารถแบ่งออกได้เป็น 4 ประเภทดังนี้

1) งานก่อสร้างที่พักอาศัย (Residential Construction) เป็นงานก่อสร้างอาคารที่เกี่ยวกับการพักอาศัยแบบต่าง ๆ ได้แก่ บ้านเดี่ยว บ้านแถว (Town House) อาคารชุดพักอาศัยห้องเช่า

2) งานก่อสร้างเพื่อธุรกิจการค้า (Building Construction of Business) เน้นอาคารสำหรับประกอบธุรกิจการค้า ได้แก่ ศูนย์การค้า อาคารสำนักงาน ส่วนใหญ่จะมีภาคเอกชนเป็นผู้

ลงทุนในโครงการ ซึ่งออกแบบโดยสถาปนิกและวิศวกรผู้เชี่ยวชาญในแต่ละระบบงานตั้งแต่งานโครงสร้าง ระบบสุขาภิบาลและดับเพลิง ระบบไฟฟ้าและสื่อสาร ระบบปรับอากาศ และระบบอื่น ๆ ส่วนงานก่อสร้างมักจะมีผู้รับเหมาก่อสร้างหลัก (General Contractor หรือ Main Contractor) ทำงานในส่วนงานโครงสร้างและสถาปัตยกรรม โดยมีผู้รับเหมาช่วงหรือผู้รับเหมาชำนาญเฉพาะทางในงานระบบต่าง ๆ เข้าทำงานในแต่ละส่วน

3) งานก่อสร้างขนาดใหญ่หรืองานสาธารณูปโภค (Heavy Engineering Construction) ใช้เงินลงทุนค่อนข้างสูงและใช้เครื่องจักรหนัก เช่น งานก่อสร้างโครงข่ายถนนทางหลวงอุโมงค์ หรือเขื่อนไฟฟ้าพลังน้ำท่าเรือเป็นต้นโดยผู้ลงทุนมักเป็นหน่วยงานราชการหรือองค์กรภาครัฐสำหรับงานออกแบบมักทำโดยวิศวกรชำนาญการหรือผู้รับเหมาซึ่งต้องมีความชำนาญเฉพาะและมีขีดความสามารถในการทำงานอยู่ในระดับสูงทั้งด้านบุคลากร เทคโนโลยีเครื่องจักรและเงินทุนในการดำเนินงาน

4) งานก่อสร้างด้านอุตสาหกรรม (Industrial Construction) เป็นงานก่อสร้างเกี่ยวกับอาคารโรงงานออกแบบการติดตั้งเครื่องมือเครื่องจักรเช่นงานก่อสร้างโรงกลั่นน้ำมันโรงไฟฟ้าพลังงานความร้อนหรือพลังงานนิวเคลียร์โรงงานอุตสาหกรรมปิโตรเลียมเป็นต้นทำการออกแบบและก่อสร้างโดยวิศวกรผู้เชี่ยวชาญเฉพาะทางหลายสาขาที่เกี่ยวข้องกับงานก่อสร้างการจ้างสร้างอาคารลักษณะนี้มักจะเป็นลักษณะเหมารวมเบ็ดเสร็จทั้งออกแบบรวมก่อสร้าง (Design-BuildContractor) เพราะมีข้อจำกัดด้านความชำนาญชั้นสูงเฉพาะงาน

2.3 ขั้นตอนการก่อสร้างบ้าน

ขั้นตอนการก่อสร้างบ้านสามารถแบ่งออกได้เป็น 4 ขั้นตอนดังนี้ (คำนวณ คุณาพร, 2546)

1) การติดต่อสถาปนิก/บริษัทออกแบบเป็นการติดต่อกับบริษัทที่รับสร้างบ้านหรือรับออกแบบบ้านซึ่งสามารถจัดหาได้จากการตรวจสอบผลงานที่ผ่านมาจากสถานที่จริงหรือภาพผลงานเพื่อให้เห็นถึงคุณภาพและรูปแบบที่ตรงกับความชอบของผู้ที่ต้องการสร้างบ้านหรือติดต่อผ่านทางสมาคมสถาปนิกสยามเพื่อขอคำแนะนำต่าง ๆ สิ่งสำคัญ คือ เจ้าของบ้านพักอาศัยต้องบอกความต้องการโดยละเอียดแก่สถาปนิกเพื่อให้เกิดความเข้าใจที่ตรงกันโดยสถาปนิกจะคำนวณค่าออกแบบเป็นร้อยละตามงบประมาณในการก่อสร้างซึ่งกำหนด โดยสมาคมสถาปนิกสยามในพระบรมราชูปถัมภ์ในขั้นตอนของการออกแบบและตกแต่งภายในสามารถแบ่งออกเป็น 6 ขั้นตอนย่อย ดังนี้

1. การให้คำปรึกษาเป็นการพบปะพูดคุยกันระหว่างเจ้าของบ้านพักอาศัยและสถาปนิกเพื่อรวบรวมข้อมูลและความต้องการของเจ้าของบ้านพักอาศัย
2. การวางผังและแนวความคิด สถาปนิกจะนำเสนอภาพรวมของบ้านในด้านแนวความคิดรูปแบบและการแบ่งพื้นที่ใช้สอยเพื่อให้เจ้าของบ้านพักอาศัยมองเห็น

ภาพรวมว่าตรงตามความต้องการหรือไม่และมีการแก้ไขและพัฒนาไปสู่แบบร่างขั้นต่อไป

3. การพัฒนาแบบร่างขั้นต้น สถาปนิกจะนำแนวความคิดและภาพรวมที่ได้จากขั้นตอนก่อนหน้านี้มาพัฒนาโดยอาจจะนำเสนอในรูปแบบของหุ่นจำลอง (Model) ภาพร่าง (Sketch) หรือรูปทัศนียภาพ (Perspective) เพื่อให้เจ้าของบ้านพักอาศัยเข้าใจแบบได้ดียิ่งขึ้นและนำสิ่งที่ลูกค้าต้องการไปปรับปรุงและเพิ่มเติมในแบบต่อไป
4. การพัฒนาแบบร่างขั้นสุดท้ายสถาปนิกจะนำเสนองานในรูปแบบที่เหมือนจริงมากยิ่งขึ้นด้วยการทำหุ่นจำลอง (Model) หรือรูปทัศนียภาพ (Perspective) แสดงรายละเอียดของมุมมองต่าง ๆในบ้านเพื่อให้เจ้าของบ้านพักอาศัยได้แก้ไขในรายละเอียด
5. การกำหนด วัสดุตกแต่งภายในสถาปนิกจะนำข้อมูลของวัสดุมาให้เจ้าของบ้านพักอาศัยทำการเลือก โดยอ้างอิงจากแบบร่างครั้งก่อนและจัดทำเป็นแผ่นกำหนด วัสดุตกแต่ง (Material Board)เพื่อให้เจ้าของบ้านพักอาศัยได้เปรียบเทียบและเลือกดูได้ตามความต้องการ
6. การเขียนแบบรายละเอียดสถาปนิกจะทำการเขียนแบบรายละเอียดและส่งให้เจ้าของบ้านพักอาศัยตรวจดูเพื่อแก้ไขและพิจารณาก่อนจะดำเนินการเขียนแบบจริง

2) การขออนุญาตปลูกบ้านเป็นการนำแบบไปขออนุญาตจากองค์การบริหารส่วนท้องถิ่นโดยต้องกรอกเอกสารต่าง ๆ ได้แก่ กรอกคำขออนุญาตก่อสร้างอาคารดัดแปลงอาคารหรือรื้อถอนอาคาร (ข.1)แบบแปลนแผนผังหนังสือรับรองว่าเป็นผู้ออกแบบ/สถาปนิก/วิศวกรพร้อมสำเนาใบอนุญาตสำเนาโฉนดที่ดินที่จะก่อสร้างสำเนาบัตรประชาชนหรือทะเบียนบ้านเจ้าของบ้านพักอาศัย

3) การปลูกสร้างบ้าน เป็นขั้นตอนที่เจ้าของบ้านพักอาศัยจะได้เห็นงานจริงสถาปนิกและวิศวกรผู้ออกแบบรวมทั้งเจ้าของบ้านพักอาศัยจะต้องให้ความสำคัญและใส่ใจในรายละเอียดโดยสถาปนิกจะต้องคอยควบคุมดูแลให้การก่อสร้างดำเนินไปตามแบบที่สร้างเอาไว้และให้คำปรึกษาแก่ผู้รับเหมาเมื่อเกิดปัญหาในการก่อสร้างหน้างานซึ่งสามารถแบ่งงานออกเป็น 5 ด้าน ดังนี้

1. งานด้านโครงสร้างเริ่มตั้งแต่การทำโครงสร้างของฐานรากได้แก่การลงเสาเข็มและการหล่อตอม่อเพื่อรองรับโครงสร้างของเสาและคานงานโครงสร้างของพื้นและบันไดซึ่งจะต้องเชื่อมต่อกับเสาและคานที่ได้ทำไว้โดยการทำพื้นจะต้องเริ่มทาจากชั้นล่างขึ้นไปชั้นบนเพื่อความสะดวกในการทำงานและการลำเลียงวัสดุงานโครงสร้างของหลังคาซึ่งส่วนใหญ่มักจะทำเป็นโครงเหล็กเชื่อมต่อกับเสาและคานชั้นบนสุดและทำการมุงหลังคานอกจากนี้ยังมีงานโครงสร้างของรั้วด้วย

2. งานก่อสร้างตัวบ้านเป็นงานที่ทาต่อเนื่องจากงานโครงสร้างได้แก่การก่อผนังและการติดตั้งวงกบประตู - หน้าต่างซึ่งจะต้องทาควบคู่กันไปเนื่องจากวงกบประตู - หน้าต่างและผนังบ้านเป็นสิ่งที่ติดตั้งเชื่อมต่อเข้าด้วยกันหลังจากทำการติดตั้งวงกบเรียบร้อยแล้วจึงทำการฉาบแต่งผนังไปพร้อมกับการฉาบแต่งเสาและคานส่วนการติดตั้งบานประตู - หน้าต่างอาจทำการติดตั้งต่อเนื่องไปเลยหรือทยอยติดตั้งหรือทำในช่วงหลังก็ได้นอกจากนี้ ยังรวมถึงการเดินท่อน้าและสายไฟที่เป็นระบบฝังจุดที่ควรระวัง คือ ก่อนการฉาบแต่งผนังต้องแน่ใจว่าการวางท่อต่าง ๆ ทำไว้อย่างถูกต้องเรียบร้อยการวางตำแหน่งของสวิตซ์ไฟปลั๊กไฟก๊อกน้ำหรืออุปกรณ์ต่าง ๆ ที่เกี่ยวข้องมีความสัมพันธ์กับการเดินท่อภายในผนัง ซึ่งเป็นไปตามที่กำหนด และสอดคล้องกับแนวท่อที่เดินไว้
3. งานด้านสาธารณูปโภคได้แก่การวางท่อระบายน้ำและทำบ่อพักเพื่อระบายน้ำตามจุดต่าง ๆ ภายในอาคารการเดินสายไฟและท่อร้อยสายไฟ
4. งานด้านสุขาภิบาลได้แก่การวางท่อระบายน้ำและทาบ่อพักเพื่อระบายน้ำลงสู่ท่อระบายน้ำสาธารณะการวางท่อระบบบำบัดน้ำเสีย
5. งานตกแต่ง เป็นงานที่ให้ผลงานปรากฏแก่สายตาของผู้พบเห็นโดยมีวัตถุประสงค์เพื่อความสวยงามเรียบร้อยหรือโดดเด่นประทับใจร่วมอยู่ด้วยได้แก่การบุฝ้าเพดานการปูพื้นและผนังการทำสี การติดตั้งสุขภัณฑ์การติดตั้งดวงโคมไฟตลอดจนการติดตั้งเฟอร์นิเจอร์

2.4 ขั้นตอนการบริหารงานก่อสร้าง

ขั้นตอนการบริหารงานก่อสร้าง (วิสูตร จิระดำเกิง, 2563) ประกอบด้วย 4 ส่วนดังนี้

1) การวางแผน (Planning) คือ การคิดค้นและแสวงหาวิธีการทำงานก่อสร้างเพื่อให้เสร็จสมบูรณ์ตามเป้าหมายที่กำหนด ด้วยราคาที่ประหยัดและมีคุณภาพตามมาตรฐานภายในเวลาที่กำหนด ไว้จึงเป็นความพยายามที่จะใช้ทรัพยากรของงานก่อสร้างให้มีประสิทธิภาพและเหมาะสมที่สุดการวางแผนสามารถแบ่งออกเป็น

1. การจัดทำแผนงานและกำหนด เวลา (Schedule of Work) ในการก่อสร้างนิยมใช้ระบบ Bar Chart และ C.P.M.
2. การจัดทำแผนทางการเงินของโครงการ โดยยึดการรับและจ่ายเงินตลอดระยะเวลาของโครงการเป็นหลักเรียกว่า Cash Flow Forcast ซึ่งแสดงให้เห็นถึงปริมาณเงินจ่ายและรับ
3. การจัดทำแผนสำหรับการปฏิบัติงานรวมถึงแผนการจัดวางหรือติดตั้งเครื่องมือเครื่องจักรศูนย์บริหารงาน พื้นที่เตรียมและเก็บพัสดุ ทางเข้า–ออกและที่พักอาศัย

ของคนงานการจัดทำระบบการประสานงานและการติดต่อระหว่างผู้ที่เกี่ยวข้องในงานทั้งหมดตลอดจนแผนการจัดหาบุคลากรต่าง ๆ

2) การจัดแบ่งงาน (Organizing) เป็นขั้นตอนระหว่างการวางแผนและการจัดเตรียมทรัพยากรของงานก่อสร้างไว้ให้พร้อมที่จะใช้งานได้ตามเวลาและโอกาสที่ต้องการการจัดแบ่งงานต้องศึกษาถึงความพร้อมหรือข้อมูลดังต่อไปนี้

1. กำลังคนหรือบุคลากร
2. จำนวนเครื่องมือเครื่องจักร
3. แหล่งเงินสนับสนุนโครงการ
4. สาระสำคัญหรือเงื่อนไขสัญญา
5. แหล่งพลังงานและบริการ

3) การควบคุม (Controlling) เป็นวิธีการที่สามารถทำให้งานก้าวหน้าไปตามแผนงานได้อย่างมีประสิทธิภาพและราบรื่นการควบคุมจำเป็นต้องมีการเปรียบเทียบการปฏิบัติงานกับแผนงานที่วางไว้แล้ววิเคราะห์หาจุดบกพร่องเพื่อนำมาปรับปรุงการปฏิบัติงานต่อไปเป็นการดำเนินงานที่เกี่ยวข้องกับงานต่อไปนี้

1. การติดต่อประสานงาน (Communication) ระหว่างส่วนกลางกับส่วนสนามหรือระหว่างผู้ควบคุมงาน ผู้รับเหมาช่วง ผู้ขายวัสดุอุปกรณ์ ผู้ว่าจ้าง และระหว่างผู้ออกแบบกับผู้รับเหมาช่วง
2. การควบคุมการก่อสร้างให้เป็นไปตามแบบรูปรายการและหลักการประกอบด้วยการควบคุมการปฏิบัติงานของช่างเทคนิค ช่างฝีมือและแรงงานไร้ฝีมือทั้งของผู้รับเหมาทั่วไปและผู้รับเหมาช่วงรวมถึงการควบคุมตรวจสอบคุณภาพและการใช้วัสดุด้วย
3. การรายงานผลความก้าวหน้าของงานและการเบิกจ่ายวัสดุอุปกรณ์ รวมทั้งผลงานที่แรงงานทำได้เพื่อผลทางการควบคุมและสถิติสำหรับใช้หรือปรับปรุงการประมาณการหรือการวางแผนงานอื่นต่อไป

4) การประเมินผล (Evaluation) เป็นการติดตามและประเมินผลการปฏิบัติงานประกอบด้วย

1. การรวบรวมและศึกษาผลที่รายงานไว้ในด้านความก้าวหน้าของงานการใช้แรงงานการใช้เครื่องมือและวัสดุอุปกรณ์
2. หาทางแก้ไขปัญหาและอุปสรรคของการปฏิบัติงาน
3. ควบคุมให้การดำเนินงานเป็นไปตามแผนการใช้วัสดุและแรงงานตามงบประมาณและการรับจ่ายเงินตาม Cash Flow ปรับปรุงแผนงานเพื่อให้งานดำเนินไปอย่างรวดเร็วและสิ้นเปลืองทรัพยากรโดยเปล่าประโยชน์น้อยที่สุด

4. นำผลจากการประเมินมาใช้เป็นข้อมูลในการเตรียมงานและปรับปรุงงาน เพื่อการรับงานใหม่ต่อไป

2.5 ทรัพยากรในงานก่อสร้าง

ทรัพยากรในงานก่อสร้างประกอบด้วย (คมสัน เวนานนท์, 2545)

1) กำลังคน หมายถึง บุคลากรประกอบด้วยผู้บริหารสถาปนิกวิศวกรช่างเทคนิคช่างฝีมือและผู้ใช้แรงงานซึ่งร่วมกันดำเนินการก่อสร้างหากมีการบริหารงานและทำงานอย่างมีประสิทธิภาพมีวินัยและมีความรับผิดชอบในหน้าที่ที่ได้รับมอบหมายคุณภาพของงานก็จะเป็นที่ยอมรับและสำเร็จตามเป้าหมาย

2) วัสดุ คือ วัตถุดิบและอุปกรณ์สำหรับการก่อสร้างการเปลี่ยนแปลงราคาของวัสดุก่อสร้างจะมีผลกระทบอย่างยิ่งต่องานก่อสร้างทั้งนี้เพราะได้มีการกำหนด ราคาค่าก่อสร้างไว้แล้ว การจัดการในด้านวัสดุก่อสร้างมุ่งที่จะทำการควบคุมด้านปริมาณราคาคุณภาพการวางแผนและกำหนดเวลาที่จะใช้วัสดุการกำหนด มาตรการใช้วัสดุเพื่อให้มีการสูญเปล่าน้อยที่สุดรวมถึงการหาวิธีการเก็บรักษาและวิธีการควบคุมสต็อกวัสดุที่สถานที่ก่อสร้างด้วย

3) เครื่องมือคือเครื่องมือเครื่องใช้และเครื่องจักรที่นำมาใช้ในการก่อสร้าง

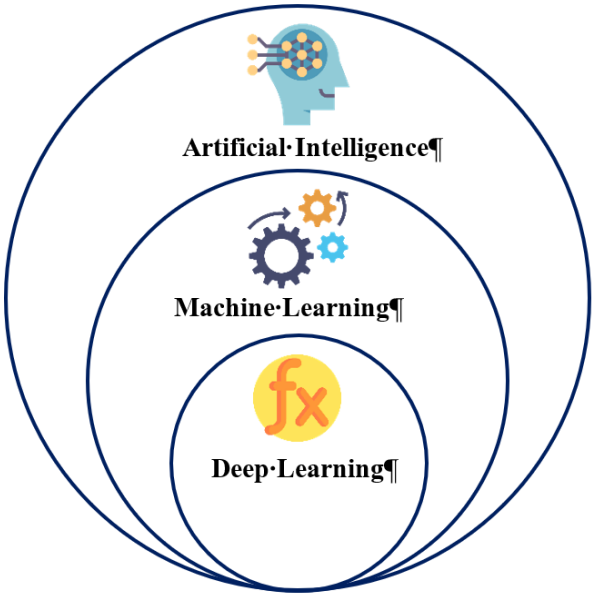
4) เงินคือตัวเงินตรารวมทั้งเครดิตต่าง ๆ ซึ่งเป็นทรัพยากรหลักของการก่อสร้างผู้ว่าจ้างมักจ่ายเงินตามผลงานเป็นงวดงานหรือตามระยะเวลาที่กำหนดไว้

แนวคิดเกี่ยวกับเทคโนโลยีการเรียนรู้ของเครื่อง

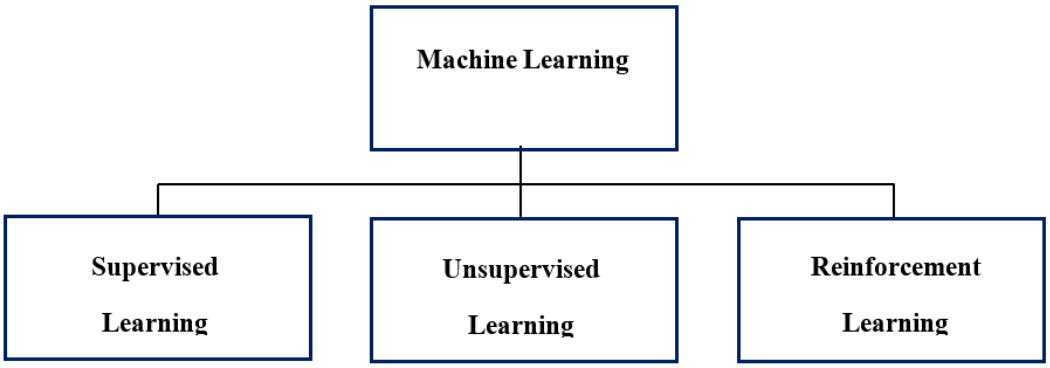
1. ลักษณะการทำงานของการเรียนรู้ของเครื่อง

ลักษณะการทำงานของการเรียนรู้ของเครื่องเป็นการทำให้เทคโนโลยีสามารถเรียนรู้ได้ด้วยตนเองจากข้อมูลที่กำหนดไว้ หรือการใช้งานเหมือนเป็นสมองของปัญญาประดิษฐ์ (Artificial Intelligence, AI) ซึ่งมีการเรียนรู้ของเครื่อง (Machine Learning, ML) เป็นส่วนหนึ่งของปัญญาประดิษฐ์ดังภาพประกอบ 7 เพื่อพัฒนาความสามารถในการคิดและตัดสินใจ โดยปัญญาประดิษฐ์สามารถเรียนรู้และพัฒนาได้อย่างอัตโนมัติจากประสบการณ์ โดยไม่ต้องมีการลงโปรแกรมใหม่โดยมนุษย์ ระบบงานนี้จะมุ่งเน้นด้านการพัฒนาโปรแกรมคอมพิวเตอร์ของปัญญาประดิษฐ์ในการเข้าถึงข้อมูลและนำข้อมูลมาใช้ด้วยตนเอง ซึ่งกระบวนการในการเรียนรู้และพัฒนาเริ่มต้นจากข้อมูลที่มนุษย์ป้อนให้ โดยข้อมูลนั้นเป็นข้อมูลนำเข้า (Input Data) และได้มาซึ่งผลลัพธ์ (Output Data) ยิ่งมนุษย์ให้ข้อมูลแก่การเรียนรู้ของเครื่องหรือปัญญาประดิษฐ์มากเท่าไหร่ปัญญาประดิษฐ์จะมีความรู้และประสบการณ์ในการหารูปแบบของข้อมูลเพื่อนำไปใช้ในการตัดสินใจในอนาคตได้อย่างถูกต้องมากขึ้นเท่านั้น โดยประเภทการเรียนรู้ของเครื่องสามารถแบ่งออกได้เป็น 3

ประเภท ได้แก่ การเรียนรู้แบบมีผู้สอน (Supervised Learning), การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ดังภาพประกอบ 8



ภาพประกอบ 7 ความสัมพันธ์ของปัญญาประดิษฐ์ การเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก



ภาพประกอบ 8 ประเภทของการเรียนรู้ของเครื่อง

1.1 Supervised Learning คือ วิธีการเรียนรู้ที่ง่ายที่สุด โดยปัญญาประดิษฐ์จะสามารถเรียนรู้จากข้อมูลที่มีการสอนหรือระบุมาให้ ว่าข้อมูลนี้คืออะไร ข้อมูลใดถูกหรือผิดภายใต้สถานการณ์ต่างๆ ข้อมูลที่มากขึ้นจะทำให้อัลกอริทึมของปัญญาประดิษฐ์พัฒนาและสามารถประมวลผลข้อมูลที่ถูกต้องในอนาคต เช่น การสอนความแตกต่างระหว่างสุนัขและแมว โดยมนุษย์นำภาพสุขนัขและแมวจำนวนมากมาระบุให้คอมพิวเตอร์ทราบว่าภาพใดคือสุนัขและภาพใดคือแมว คอมพิวเตอร์จะเรียนรู้

รูปร่างขนาด ลักษณะของสัตว์ทั้งสองประเภทจากรูปภาพจำนวนมาก จนสามารถแยกสุนัขและแมวได้ ในที่สุด นอกจากการเรียนรู้แบบ Supervised Learning ที่มีผู้สอนโดยปกติแล้ว ยังมีวิธีการเรียนรู้แบบ Deep Learning ที่มีรูปแบบการเรียนรู้จากข้อมูลที่มีการสอนเช่นกัน แต่ Deep Learning จะมีการกำหนดขั้นของข้อมูล (Layer) ที่แตกต่างกันเพื่อให้เครื่องสามารถเรียนรู้ข้อมูลได้มากขึ้น แม่นยำขึ้นโดยใช้เทคนิคการเพิ่มความลึกของข้อมูลที่แสดงออกมาตามจำนวนของชั้นของข้อมูลที่มากขึ้น เช่น ความสามารถ Face Recognition ที่สามารถวิเคราะห์ใบหน้าบุคคลได้จากหลายมุม หรือสามารถวิเคราะห์ใบหน้าในช่วงอายุที่แตกต่างกันได้

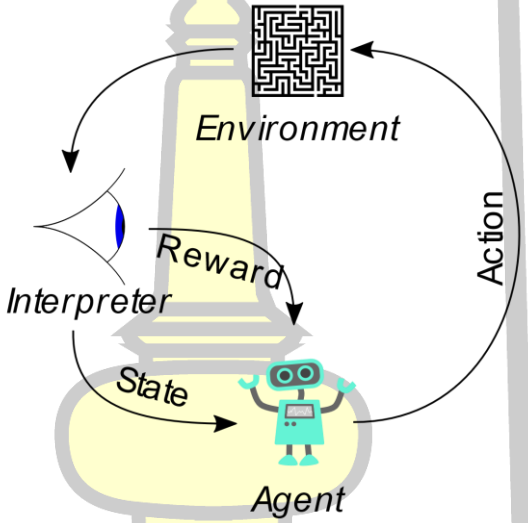
Supervised Learning คือวิธีการเรียนรู้ที่ข้อมูลอินพุตและเอาต์พุตถูกจัดเตรียมให้กับอัลกอริทึมซึ่งใช้ฟังก์ชันข้อผิดพลาดในการแมปอินพุตกับเอาต์พุต ฟังก์ชันนี้เป็นองค์ประกอบหลักในกระบวนการเรียนรู้ โดยมุ่งที่จะเข้าใกล้ฟังก์ชันพื้นฐานที่แท้จริงมากขึ้น คำว่าการเรียนรู้หมายถึงการปรับเปลี่ยนพารามิเตอร์ของอัลกอริทึมและการปรับน้ำหนักของตัวแปรอินพุต ตัวอย่างของวิธีการภายใต้การดูแลคืออัลกอริทึมการถดถอยและการจำแนกประเภท ในขณะที่การถดถอยใช้ตัวแปรต่อเนื่องเป็นอินพุตเพื่อรับเวกเตอร์เอาต์พุตที่เป็นตัวเลข รูปแบบการจำแนกประเภทจะกำหนดป้ายกำกับให้กับข้อมูลอินพุต

1.2 Unsupervised Learning คือการเรียนรู้จากข้อมูลที่ไม่มีการระบุโดยเฉพาะเจาะจงหรือไม่มีผู้สอน ซึ่งข้อมูลจะไม่ได้ถูกจัดประเภทหรือระบุป้ายกำกับข้อมูล การเรียนรู้แบบนี้ปัญญาประดิษฐ์จะอนุมานข้อมูลที่ได้รับและทำความเข้าใจถึงโครงสร้างที่ซ่อนอยู่ ดังนั้นปัญญาประดิษฐ์จะไม่สามารถระบุผลลัพธ์ที่ถูกต้องได้ทันที แต่จะใช้วิธีการสำรวจข้อมูลและอนุมานข้อมูล ซึ่งจะเพิ่มความถูกต้องแม่นยำตามระยะเวลาที่เพิ่มขึ้น หมวดหมู่ของอัลกอริทึมการเรียนรู้ของเครื่องที่ออกแบบมาเพื่อจัดการกับปัญหาที่เกี่ยวข้องกับข้อมูลที่ไม่สมบูรณ์ อัลกอริทึมเหล่านี้สร้างกฎตามความสัมพันธ์ระหว่างจุดข้อมูลโดยไม่ทราบผลลัพธ์ ตัวอย่างทั่วไปสำหรับการเรียนรู้แบบไม่มีผู้ดูแลคือการจัดกลุ่ม การจัดกลุ่มเป็นแนวทางที่ไม่มีผู้ดูแลโดยการจับคู่ข้อมูลที่ป้อนเข้าเป็นกลุ่มตามมาตรการความคล้ายคลึงกัน การวัดความคล้ายคลึงกันโดยทั่วไปรวมถึงการวัดระยะทางแบบยุคลิดและการวัดความสัมพันธ์ เช่น ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

1.3 Reinforcement Learning คือการเรียนรู้ตามสภาพแวดล้อมโดยคอมพิวเตอร์จะได้รับการสนับสนุนส่งเสริม หากผลลัพธ์จากการประมวลผลนั้นถูกต้อง ในทางกลับกันคอมพิวเตอร์จะได้รับการลงโทษ หากผลลัพธ์จากการประมวลผลนั้นผิดพลาด เช่น การเรียนรู้ของเครื่องดูดฝุ่นอัจฉริยะ โดยเครื่องดูดฝุ่นจะเรียนรู้เส้นทางการเดินตามสภาพแวดล้อม หากเครื่องดูดฝุ่นเดินไปตามเส้นทางที่ถูกต้องจะไม่พบสิ่งกีดขวาง ทำให้เครื่องดูดฝุ่นยังคงสามารถวิ่งต่อไปได้ แต่หากเครื่องดูดฝุ่นเดินไปตามเส้นทางที่ไม่ถูกต้องมีการชนกับผนังหรือสิ่งกีดขวาง จะทำให้เครื่องดูดฝุ่นไม่สามารถวิ่งต่อได้ปัญญาประดิษฐ์ของเครื่องดูดฝุ่นจะเรียนรู้ว่าเส้นทางใดคือเส้นทางที่ไปต่อได้หรือเส้นทางใดไปต่อ

ไม่ได้ และเมื่ออยู่ในสถานการณ์ที่เจอทางตันปัญญาประดิษฐ์จะเรียนรู้ว่าจะต้องเลี้ยวซ้าย เลี้ยวขวา หรือถอยหลังเพื่อที่จะวิ่งต่อไป

การเรียนรู้แบบกึ่งควบคุมดูแลเป็นการผสมผสานระหว่างการเรียนรู้แบบมีผู้สอนและแบบไม่มีผู้ดูแล ในขณะที่ข้อมูลอินพุตและเอาต์พุตถูกกำหนดให้กับอัลกอริทึม ข้อมูลเอาต์พุตบางส่วนอาจไม่ถูกต้องหรือไม่มีอยู่จริง การเรียนรู้การเสริมกำลังกำหนดประเภทการเรียนรู้โดยมีเป้าหมายเพื่อบรรลุวัตถุประสงค์บางอย่างในสภาพแวดล้อมแบบไดนามิกโดยการเรียนรู้ผ่านรางวัลและการลงโทษ การกระทำที่นำไปสู่การบรรลุเป้าหมายนั้นได้รับการตอบแทน ในขณะที่การกระทำที่ไม่สำเร็จนั้น ดังภาพประกอบ 3 หลังจากฝึกอัลกอริทึมอย่างเพียงพอแล้ว จะใช้การผสมผสานที่ดีที่สุดของการกระทำที่ให้รางวัลเพื่อให้ได้ประสิทธิภาพที่เหมาะสมที่สุด



ที่มา: https://en.wikipedia.org/wiki/Reinforcement_learning

ภาพประกอบ 9 การเรียนรู้ของเครื่องแบบเสริมกำลัง

2. ประโยชน์ของเทคโนโลยีการเรียนรู้ของเครื่อง

1) การประมวลผลข้อมูลจำนวนมากและทำงานซ้ำซ้อนเป็นประจำได้อย่างรวดเร็วและน่าเชื่อถือ จากการที่เทคโนโลยีสามารถประมวลผลข้อมูลจำนวนมากได้อย่างน่าเชื่อถือและสามารถทำงานอย่างต่อเนื่องโดยไม่เหน็ดเหนื่อย เช่น งานที่ต้องใช้การคำนวณ การประมวลผลข้อมูลและการปฏิบัติงานซ้ำๆ เป็นประจำ อีกทั้งการใช้เทคโนโลยียังช่วยลดความผิดพลาดที่เกิดจากมนุษย์ สามารถเพิ่มประสิทธิภาพของงาน และลดระยะเวลาในการทำงานได้

2) เพิ่มความฉลาดให้กับสินค้าและบริการ การใช้เทคโนโลยีปัญญาประดิษฐ์ในสินค้าหรือบริการเป็นการเพิ่มมูลค่าให้กับสินค้าหรือบริการนั้น ทำให้องค์กรสามารถขายสินค้าในราคาที่สูงขึ้นหรือให้บริการลูกค้าได้รวดเร็วขึ้น ส่วนในด้านของผู้บริโภคการเพิ่มเทคโนโลยีปัญญาประดิษฐ์ในสินค้า

หรือบริการทำให้ผู้บริโภคสามารถใช้สินค้าหรือบริการได้อย่างสะดวก รวดเร็วและมีประสิทธิภาพมากยิ่งขึ้น

3) เพิ่มระดับความแม่นยำ เทคโนโลยีปัญญาประดิษฐ์ยิ่งทำงานมากเท่าใด ความแม่นยำของระบบจะยิ่งมากขึ้น

4) การวิเคราะห์ข้อมูลเชิงลึก เทคโนโลยีปัญญาประดิษฐ์สามารถวิเคราะห์ข้อมูลในปริมาณมากและในเชิงลึกด้วยระยะเวลาอันสั้น

5) การพัฒนาระบบผ่านการเรียนรู้ เทคโนโลยีปัญญาประดิษฐ์สามารถปรับตัวโดยอัตโนมัติเพื่อรองรับความซับซ้อนและปริมาณของอัลกอริทึมที่มากขึ้นได้อย่างไร้ขีดจำกัด โดยเรียนรู้จากผลลัพธ์ที่ผ่านมาตามชุดคำสั่งของอัลกอริทึม

6) การได้รับประโยชน์สูงสุดจากข้อมูล เทคโนโลยีปัญญาประดิษฐ์สามารถนำข้อมูลมหาศาลที่ถูกรวบรวมและผ่านการประมวลไปใช้ประโยชน์ด้านต่างๆ เช่น การวิเคราะห์พฤติกรรมผู้บริโภคและความต้องการของลูกค้าอย่างต่อเนื่องและตลอดเวลา ช่วยให้ผู้บริหารองค์กรสามารถสร้างกลยุทธ์พัฒนาธุรกิจได้ตรงกับความต้องการของลูกค้า ยิ่งการประมวลผลข้อมูลใช้เทคโนโลยี ปัญญาประดิษฐ์มากเท่าใด ก็ยิ่งช่วยเพิ่มความได้เปรียบทางการแข่งขันมากขึ้น ซึ่งข้อมูลเหล่านั้นจะกลายเป็นสินทรัพย์ขององค์กรที่มูลค่ามหาศาล

จากบทความของกรรมการบริหาร Compliance, Governance and Oversight Counsel (CGOC) ได้กล่าวถึงประโยชน์ของเทคโนโลยีปัญญาประดิษฐ์ (Machine Learning) ในทางปฏิบัติสำหรับองค์กร CGOC ซึ่งเป็นองค์กรสภาที่จัดการประชุมทั่วสหรัฐอเมริกาและยุโรปเพื่อแก้ปัญหาความท้าทายทางกฎหมายด้านไอทีและธุรกิจในปัจจุบัน โดยมีผู้สนับสนุนการก่อตั้งเป็นบริษัท IBM ในปี 2019 ไว้ 5 ด้าน ได้แก่ 1) สามารถติดตามการเปลี่ยนแปลงของกฎระเบียบข้อบังคับในระดับโลกติดตามความคาดหวัง และความต้องการของธุรกิจในอุตสาหกรรมต่างๆ เพื่อสนับสนุนการดำเนินธุรกิจตามกฎระเบียบข้อกำหนดได้อย่างถูกต้องและรวดเร็ว 2) สามารถตรวจสอบและเฝ้าติดตามการดำเนินธุรกิจที่ขัดต่อนโยบายหรือกฎระเบียบข้อกำหนด โดยองค์กรระดับโลกสามารถระบุตัวพนักงานหรือคู่ค้าที่กระทำผิดได้อย่างรวดเร็ว เพื่อลดโอกาสในการถูกปรับจากการกระผิด โดยการรายงานด้วยตนเอง 3) สามารถระบุผลกระทบจากการเปลี่ยนแปลงกฎระเบียบข้อกำหนดต่างๆ เช่น การใช้ Machine Learning ในการระบุข้อตกลงกับลูกค้าที่ได้รับผลกระทบจาก กฎหมายที่เกี่ยวกับ Data Privacy ว่าด้วยเรื่องการคุ้มครองสิทธิของข้อมูลส่วนบุคคลของ EU Citizen (GDPR) 4) สามารถจัดประเภทข้อมูลโดยอัตโนมัติ ว่าข้อมูลใดควรเก็บและข้อมูลใดควรทิ้ง เพื่อเป็นการลดต้นทุนค่าใช้จ่ายและลดความเสี่ยงที่เกี่ยวข้องกับกฎหมาย Data Privacy 5) สามารถช่วยในการวิจัยทางกฎหมายได้รวดเร็วและแม่นยำมากยิ่งขึ้น เช่นการใช้ Machine Learning ในการพิจารณาความเสี่ยงที่อาจจะ

เกิดขึ้นและการตัดสินคดีทางกฎหมายในช่วงต้น เพื่อเป็นการประหยัดค่าใช้จ่ายให้กับองค์กรในการดำเนินคดีและลดความเสียหายที่จะเกิดขึ้นต่อชื่อเสียงขององค์กรให้น้อยที่สุด

ทั้งนี้สมาคมฟินเทคประเทศไทยได้กล่าวถึง ศักยภาพของปัญญาประดิษฐ์ในบทบาทของธนาคาร ในปี 2018 ว่าปัญญาประดิษฐ์มีบทบาทในการช่วยวิเคราะห์ข้อมูลโดยการวิเคราะห์แบบปัญญาประดิษฐ์ AI-driven สามารถตรวจสอบข้อมูลจำนวนมหาศาล นำไปสู่การจัดเรียงข้อมูลตามหมวดหมู่ รวมถึงสามารถทำการวิเคราะห์ข้อมูลแบบแบบเรียลไทม์ด้วยการเรียนรู้ของเครื่องที่สามารถนำข้อมูลหลังวิเคราะห์ไปใช้ในส่วนต่าง ๆ เช่น การสร้างรูปแบบความเสี่ยง การระบุไบโอเมตริกซ์ การตรวจสอบการทุจริต การสมัครบัตรเครดิต เป็นต้น และยังสามารถสร้าง Chatbot ผ่านเทคโนโลยีปัญญาประดิษฐ์เพื่อตอบโต้กับลูกค้า รวมถึงพนักงานภายในองค์กร หรือแม้กระทั่งความสามารถในการตรวจสอบความถูกต้องของข้อมูล โดยกระบวนการ Robotic process automation (RPA) ซึ่งใช้เทคนิคการทำซ้ำที่เป็นกิจวัตรของมนุษย์ เพื่อย้ำเตือนความถูกต้องทางด้านข้อมูล ส่งผลต่อการตรวจผลข้อมูลอย่างอัตโนมัติ

ดังนั้นจะเห็นได้ว่าเทคโนโลยีปัญญาประดิษฐ์นั้นมีประโยชน์กับธุรกิจหรืออุตสาหกรรมที่หลากหลาย ซึ่งหากจะเปรียบอุตสาหกรรมลิสซิ่งกับอุตสาหกรรมการเงินและธนาคาร จะพบว่ามีลักษณะการดำเนินธุรกิจบางส่วนที่คล้ายคลึงกัน ในรูปแบบการให้บริการทางการเงินกับลูกค้า ซึ่งหากนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ในอุตสาหกรรมลิสซิ่งก็จะมีประโยชน์และสามารถสร้างคุณค่าให้กับองค์กรได้เช่นกัน

3. ข้อจำกัดของเทคโนโลยีการเรียนรู้ของเครื่อง

1) เรียนรู้จากข้อมูลเท่านั้น เทคโนโลยีการเรียนรู้ของเครื่องเรียนรู้ ประมวลผล และพัฒนาจากข้อมูลที่มนุษย์ป้อนให้เท่านั้น หากข้อมูลที่ป้อนไม่ถูกต้อง ก็จะส่งผลต่อผลลัพธ์ที่ออกมาจากการประมวลของเทคโนโลยีการเรียนรู้ของเครื่อง

2) มีต้นทุนและค่าใช้จ่ายสูง เนื่องจากการพัฒนาเทคโนโลยีนั้นต้องใช้ความรู้เฉพาะด้านหลายแขนงและต้องใช้เวลาในการพัฒนาเทคโนโลยีการเรียนรู้ของเครื่องที่มีความซับ ซ้อน ให้เกิดประสิทธิภาพในการใช้งาน ซึ่งก่อให้เกิดต้นทุนด้านเวลา และต้นทุนค่าใช้จ่ายที่สูง

3) ขาดความคิดสร้างสรรค์ เทคโนโลยีการเรียนรู้ของเครื่องเหมาะสำหรับงานที่มีรูปแบบชัดเจนหรือเป็นการทำงานที่ซ้ำซ้อน เพื่อลดการทำงานของมนุษย์ที่อาจจะเกิดความผิดพลาด โดยปัจจุบันเทคโนโลยีการเรียนรู้ของเครื่องยังไม่มีการพัฒนากระบวนการคิด ความเข้าใจ ความรู้สึก หรือความสามารถทางปัญญาเทียบเท่ากับมนุษย์

4) การขาดความรู้ความสามารถของพนักงาน เกี่ยวกับเทคโนโลยี จากการเรียนรู้ของการเรียนรู้ของเครื่องบางประเภทที่ยังต้องอาศัยมนุษย์ในการป้อนข้อมูลหรือสอนการเรียนรู้ของเครื่องว่าข้อมูลนี้คืออะไรข้อมูลใดถูกหรือข้อมูลใดผิด ภายใต้สถานการณ์หรือเงื่อนไขต่างๆ ที่มนุษย์เป็นผู้

กำหนด ดังนั้นการขาดความรู้ของพนักงานในการใช้เครื่องมือการเรียนรู้ของเครื่องจึงอาจเป็นข้อจำกัดในการที่จะเริ่มสอนการเรียนรู้ของเครื่องให้สามารถประมวลผลข้อมูลที่ถูกต้องในอนาคต

เครื่องจักรเรียนรู้เอ็กซ์ตรีม

เครื่องจักรเรียนรู้เอ็กซ์ตรีม (Extreme Learning Machine, XLM) คือ การนำเสนอวิธีการที่ได้รับการพัฒนาและปรับปรุงการเรียนรู้ของเครื่องจักร (machine learning) ขั้นสูงเพื่อเพิ่มประสิทธิภาพของกระบวนการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ ซึ่งอย่างไรก็ตามการนำเสนอนี้มุ่งหวังที่จะทำให้ค่าความแม่นยำในการทำนายมีความสูงสุดเมื่อใช้อัลกอริทึมที่ได้รับการปรับแต่งและพัฒนาขึ้นเป็นแบบจำลองที่มีประสิทธิภาพที่สูงขึ้นในการทำนาย มีการเลือกใช้วิธีการเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมทั้งหมด 5 อัลกอริทึมที่ประกอบด้วย การเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) การเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning) การเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning), การเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) และ การเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) สำหรับการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบเพื่อให้แบบจำลองมีความสามารถในการทำนายที่เพิ่มขึ้นและความถูกต้องที่มากขึ้น

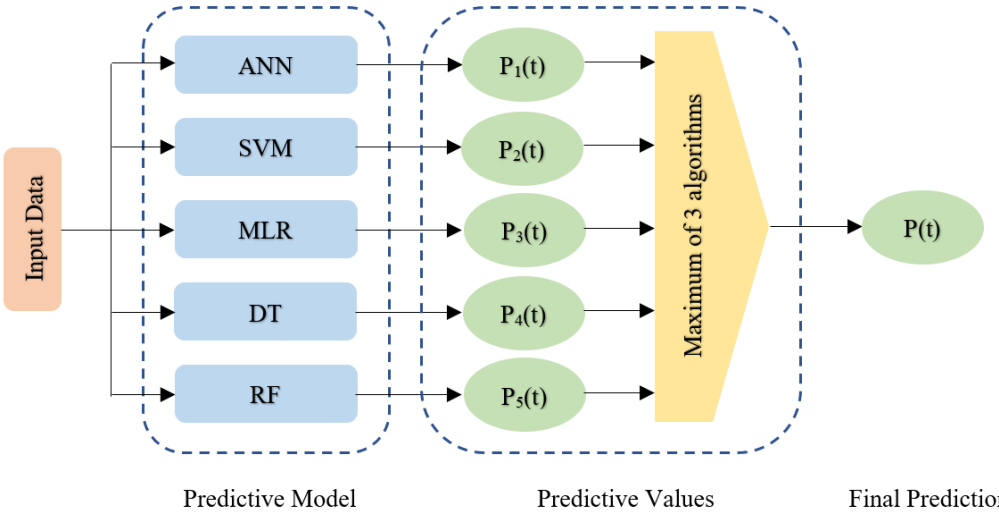
เครื่องจักรเรียนรู้เอ็กซ์ตรีมนี้เน้นถึงการปรับใช้หลักการของอัลกอริทึมทางสถิติและเทคนิคที่ซับซ้อนเพื่อเพิ่มประสิทธิภาพในกระบวนการทำนายราคาค่าก่อสร้างบ้านพักอาศัย ควบคุมความเสี่ยงและการลดความผิดพลาดในการทำนายมีความสำคัญเป็นพิเศษในส่วนนี้ โดยการเลือกใช้การเรียนรู้กลุ่ม 5 อัลกอริทึมดังที่กล่าวข้างต้นที่มีความหลากหลายและความเชื่อถือได้สูง ช่วยให้แบบจำลองสามารถจับความสัมพันธ์ของข้อมูลและคุณลักษณะที่มีผลต่อราคาค่าก่อสร้างบ้านพักอาศัยได้อย่างมีประสิทธิภาพ เป็นทั้งการรวบรวมความรู้จากหลากหลายแหล่งและการทดลองวิจัยในการปรับแต่งพารามิเตอร์เพื่อให้แบบจำลองมีความแม่นยำและความเชื่อถือในการทำนายที่มากที่สุดเมื่อเปรียบเทียบกับอัลกอริทึมทั่วไป

1. การเรียนรู้กลุ่มแบบโหวตสูงสุด

แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) ในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยนั้นมักจะเป็นการรวมคำทำนายจากแบบจำลองอัลกอริทึมทั่วไปๆ หรือ แบบจำลองการเรียนรู้แบบเชิงเดี่ยว(Individual Standard Model) (Surawatchayotin & Paireekreng, 2021) หลายรูปแบบเข้าด้วยกันเพื่อให้ได้คำทำนายสุดท้ายที่มีประสิทธิภาพสูงที่สุด ในวิธีการนี้นำแบบจำลองทั่วไปๆ เช่น โครงข่ายประสาทเทียม (Artificial Neural Networks) ซัพ

พอร์ตเวกเตอร์แมชชีน (Support Vector Machines) การสร้างสมการเชิงเส้นหลายตัวแปร (Multiple Linear Regression) ต้นไม้การตัดสินใจ (Decision Trees) หรือแรนดอมฟอเรสต์ (Random Forests) จะถูกฝึกการเรียนรู้เดี่ยวบนชุดข้อมูลเดียวกัน หลังจากแบบจำลองพื้นฐานเหล่านี้ฝึกการเรียนรู้แล้ว พวกเขาจะทำนายค่าตัวแปรเป้าหมาย ซึ่งในที่นี้ก็คือราคาของบ้านพักอาศัยแนวราบ วิธีการเรียนรู้กลุ่มโหวตสูงสุดจากนั้นจะเลือกค่าฐานนิยมจากคำทำนายที่ถูกนำเสนอโดยแบบจำลองพื้นฐานทั้งหมด ค่าฐานนิยมจะถูกเลือกเป็นคำทำนายสุดท้ายสำหรับราคาของบ้านพักอาศัยแนวราบ ซึ่งแสดงในภาพประกอบ 10

แนวคิดของวิธีการเรียนรู้กลุ่มแบบโหวตสูงสุดคือการนำความรู้ร่วมกันของแบบจำลองที่หลากหลาย แบบจำลองแต่ละรูปแบบอาจมีจุดแข็งและจุดอ่อนของตัวเอง และโดยการรวมกันของคำทำนายพวกเขาจะโหวตสูงสุดมุ่งหมายที่จะลดผลกระทบจากข้อผิดพลาดของแบบจำลองพื้นฐานแต่ละรูปแบบและสร้างคำทำนายที่แม่นยำและเสถียรมากขึ้น วิธีการนี้มีประสิทธิภาพในการใช้งานโดยเฉพาะเมื่อแบบจำลองพื้นฐานแต่ละรูปแบบมีความสัมพันธ์ที่น้อยในส่วนของความผิดพลาด ซึ่งหมายความว่าพวกเขาทำข้อผิดพลาดในกรณีที่แตกต่างกัน โดยการเลือกค่าฐานนิยม วิธีการเรียนรู้กลุ่มแบบโหวตสูงสุดสามารถใช้คำทำนายที่เห็นด้วยกันบ่อยที่สุดระหว่างแบบจำลอง เพื่อให้ได้การทำนายที่น่าเชื่อถือมากยิ่งขึ้นสำหรับราคาค่าก่อนสร้างบ้านพักอาศัยทั้งหมด



ภาพประกอบ 10 แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด

2. การเรียนรู้กลุ่มแบบเฉลี่ย

แบบจำลองกลุ่มแบบเฉลี่ย (averaging ensemble learning) เป็นวิธีที่ได้รับความนิยมอีกหนึ่งวิธีที่ใช้ในการรวมการทำนายจากแบบจำลองเชิงเดี่ยวหรือแบบจำลองทั่วไปหลายตัวเข้าด้วยกันเพื่อทำการทำนายสุดท้ายในงานทางการคาดคะเนแบบเชิงสถิต ตามที่แสดงในภาพประกอบ 11 เช่น

การคาดคะเนราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ วิธีนี้แบบจำลองทั่วไปเช่นเดียวกับการเรียนรู้กลุ่มแบบโหวตสูงสุด คือ โครงข่ายประสาทเทียม (Artificial Neural Networks) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) การสร้างสมการเชิงเส้นหลายตัวแปร (Multiple Linear Regression) ต้นไม้การตัดสินใจ (Decision Trees) และแรนดอมฟอเรสต์ (Random Forests)

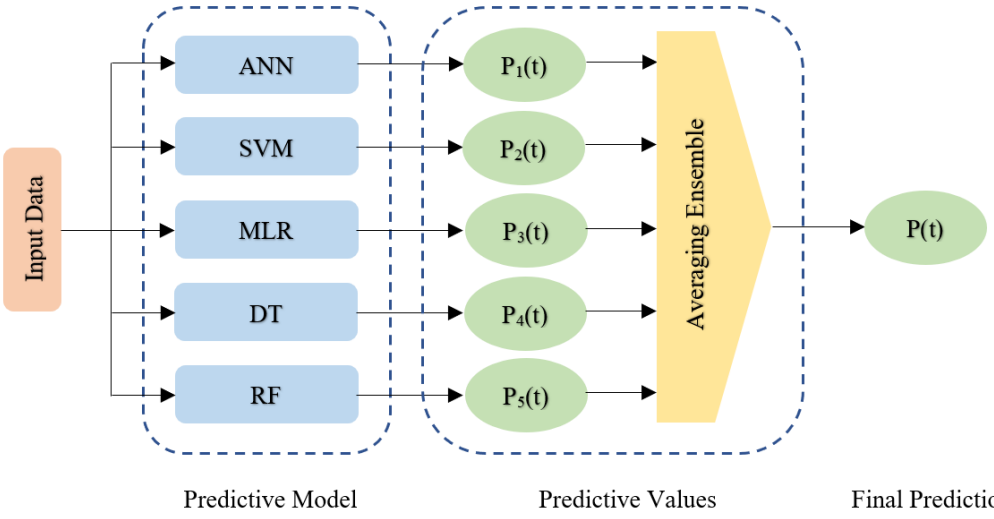
เมื่อแบบจำลองเชิงเดี่ยวหรือแบบจำลองพื้นฐานทั่วไปแต่ละตัวถูกฝึกอบรมแล้ว ก็จะทำการทำนายตัวแปรเป้าหมายหรือราคาค่าก่อสร้างบ้านพักอาศัยแนวราบในการแบบจำลองแบบกลุ่มโดยเฉลี่ยค่าทำนาย สุดท้ายจะได้รับโดยการหาค่าเฉลี่ยของค่าทำนายที่ได้จากแบบจำลองเดี่ยวทั้งหมด ทางคณิตศาสตร์สามารถแสดงเป็นดังสมการที่ 13

$$P(t) = P1(t) + P2(t) + \ldots + Pn(t) / N \qquad (13)$$

เมื่อ P(t) คือ แบบจำลองการทำนายสุดท้าย

P1(t), P2(t), ..., Pn(t) คือ แบบจำลองการทำนายของแต่ละแบบจำลองเดี่ยว

N คือ จำนวนของแบบจำลองการทำนายเดี่ยวที่ใช้ในแบบจำลองกลุ่ม



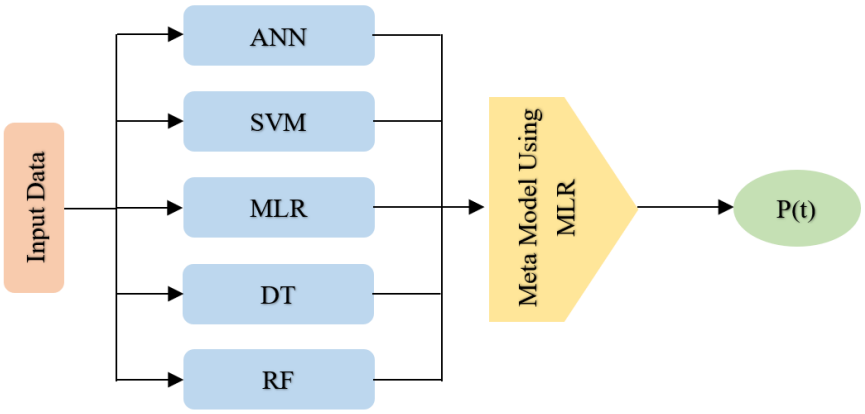
ภาพประกอบ 11 แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย

แนวคิดของการเรียนรู้กลุ่มแบบเฉลี่ยนั้นเกิดจากการนำประโยชน์ที่ได้จากการเรียนรู้สะสมของแบบจำลองต่าง ๆ มารวมกัน เพื่อลดความแปรปรวนในการทำนาย แต่ละแบบจำลองเดี่ยวๆ อาจมีข้อดีและข้อเสียที่แตกต่างกัน และผ่านการคาดคะเนโดยเฉลี่ย เป้าหมายของกลุ่มคือการสร้างการทำนายที่เสถียรและแม่นยำมากขึ้น การหาค่าเฉลี่ยช่วยให้ความคลาดเคลื่อนของแต่ละแบบจำลอง

ลดลงอย่างมีนัยสำคัญ ซึ่งทำให้ประสิทธิภาพการทำนายโดยรวมเพิ่มขึ้น เช่นเดียวกับการเรียนรู้กลุ่มแบบโหวตสูงสุด ค่าเฉลี่ยทั้งหมดจะมีประสิทธิภาพมากที่สุดเมื่อความสัมพันธ์ระหว่างแต่ละแบบจำลองต่ำ นั่นหมายความว่าแบบจำลองเหล่านั้นมีผลต่อการทำนายที่แตกต่างกันในสถานการณ์ต่าง ๆ โดยการหาค่าเฉลี่ยของการทำนายกลุ่มจะได้รับประโยชน์จากข้อมูลองค์กรที่เป็นเชิงลึกของแบบจำลองและได้การทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบที่น่าเชื่อถือและแม่นยำยิ่งขึ้น

3. การเรียนรู้กลุ่มแบบสแต็กกิ้ง

แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (Stacking ensemble learning) หรือที่มักเรียกกันว่า stacked generalization เป็นวิธีการเรียนรู้แบบจำลองการเรียนรู้กลุ่มที่ซับซ้อนซึ่งแสดงไว้ในภาพประกอบที่ 12 โดยผสมผสานการคาดการณ์จากแบบจำลองต่างๆ เข้าด้วยกันเพื่อเพิ่มประสิทธิภาพการทำนายที่ครอบคลุมในงานการเรียนรู้ของเครื่อง ซึ่งครอบคลุมการทำนายราคาบ้านพักอาศัยแนวราบสแต็กกิ้ง นอกเหนือไปจากวิธีการหาค่าเฉลี่ยหรือการลงคะแนนแบบง่ายๆ ด้วยการแนะนำแบบจำลองเมตาที่เรียนรู้ที่จะรวมเอาผลลัพธ์ของแบบจำลองพื้นฐานเข้าด้วยกัน



ภาพประกอบ 12 แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง

การเรียนรู้กลุ่มแบบสแต็กกิ้ง ประกอบด้วยขั้นตอนต่อไปนี้

1. แบบจำลองหลัก มีแบบจำลองหลักหลายแบบที่หลากหลายถูกฝึกอบรมแยกกันบนชุดข้อมูลเดียวกัน แบบจำลองหลักเหล่านี้อาจเป็นอัลกอริทึมที่แตกต่างกันหรือเป็นอัลกอริทึมเดียวกันแต่มีพารามิเตอร์แฮปเปอร์พารามิเตอร์ที่แตกต่างกัน

2. การตรวจสอบจากการคาดคะเน การสร้างชุดการตรวจสอบจากข้อมูลการฝึกอบรมแบบจำลองหลักทำการคาดคะเนบนชุดการตรวจสอบนี้ซึ่งจะถูกใช้เป็นป้อนเข้าสู่เมต้าแบบจำลอง

3. เมต้าแบบจำลอง หรือที่เรียกกันว่า "เบลนเดอร์" หรือ "ผู้รวม" ถูกฝึกอบรมโดยใช้คำทำนายจากแบบจำลองหลักบนชุดการตรวจสอบเป็นคุณสมบัติป้อนเข้า เมต้าแบบจำลองเรียนรู้วิธีการ

รวมคำทำนายเหล่านี้เพื่อสร้างการทำนายสุดท้ายสำหรับตัวแปรเป้าหมายหรือราคาค่าก่อสร้างบ้านพักอาศัยแนวราบนั่นเอง

4. การทำนายสุดท้าย เมื่อเมต้าแบบจำลองถูกฝึกอบรมแล้ว เขาจะทำการคาดคะเนบนชุดข้อมูลทดสอบ การคาดการณ์สุดท้ายสำหรับราคาค่าก่อสร้างจะได้รับจากผลลัพธ์ของเมต้าแบบจำลอง

ประโยชน์หลักของการเรียนรู้กลุ่มแบบสแต็กกิ้ง คือความสามารถในการจับความสัมพันธ์ระดับสูงระหว่างการทำนายของแบบจำลองหลัก ๆ โดยสามารถเรียนรู้ว่าเมื่อควรไว้วางใจแบบจำลองแต่ละรุ่นและกำหนดน้ำหนักที่แตกต่างกันให้กับการทำนายของพวกเขาโดยขึ้นอยู่กับประสิทธิภาพของพวกเขาบนชุดการตรวจสอบความถูกต้อง (validation set) นี้ช่วยให้การสะสมแบบสแต็กกิ้งสามารถดำเนินการดีกว่าแบบจำลองแต่ละรุ่นและการสะสมแบบเฉลี่ย/โหวต โดยการใช้ความสามารถของแบบจำลองแต่ละรุ่นให้เต็มที่มากขึ้น การสะสมแบบสแต็กกิ้ง เป็นเทคนิคที่ยืดหยุ่นและมีกำลังในการใช้งาน แต่ต้องการการดำเนินการและการปรับแต่งที่รอบคอบเพื่อป้องกันการเกิดปัญหาการเรียนรู้เกินไป (overfitting) การเรียนรู้กลุ่มแบบสแต็กกิ้งที่ถูกดำเนินการอย่างถูกต้องสามารถเพิ่มความแม่นยำในการทำนายและความทนทาน จึงเป็นทางเลือกที่นิยมในการแข่งขันทางด้านการเรียนรู้ของเครื่องจักรและการประยุกต์ใช้ในสถานการณ์จริง เช่น การทำนายราคาอสังหาริมทรัพย์ประเภทบ้านและหลายประเภทอื่น ๆ ด้วย

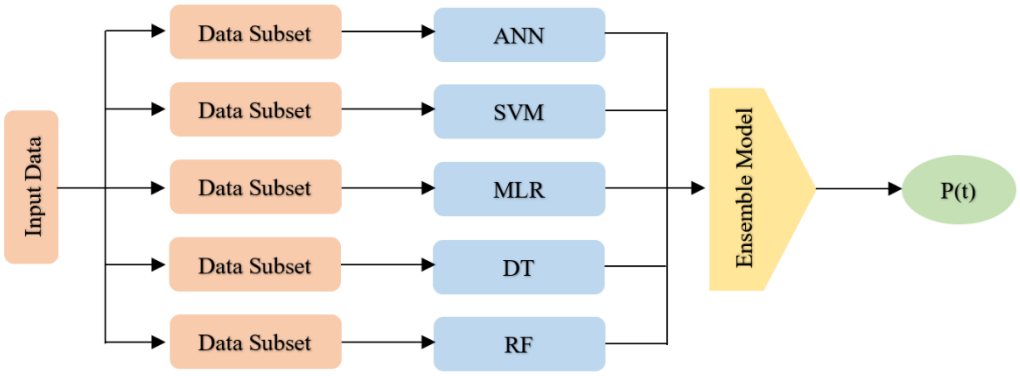
4. การเรียนรู้กลุ่มแบบแบ๊กกิ้ง

แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) หรือ Bootstrap Aggregating เป็นเทคนิคหนึ่งในการเพิ่มประสิทธิภาพในการทำนายของแบบจำลอง การเรียนรู้แบบกลุ่มประกอบด้วยการสร้างชุดข้อมูลย่อยจากชุดข้อมูลการฝึกอบรมเดิมผ่านกระบวนการ Bootstrap โดย Bootstrap คือกระบวนการสุ่มสร้างชุดข้อมูลย่อยแบบสุ่มทับซ้อนจากชุดข้อมูลต้นฉบับ แล้วใช้แบบจำลองเดิมหรือแบบจำลองหลายรูปแบบในการฝึกอบรมบนแต่ละชุดข้อมูลย่อยนี้ (Rahimi et al., 2023) ดังภาพประกอบ 13

ทุกชุดข้อมูลย่อยจะใช้สำหรับการฝึกอบรมแยกต่างหาก การทำนายของแต่ละแบบจำลองจะถูกรวมกันเพื่อสร้างการทำนายสุดท้าย ในกระบวนการนี้การเรียนรู้กลุ่มจะช่วยลดความผิดพลาดโดยการเฉลี่ยคำทำนายจากแบบจำลองหลายๆ ตัว เนื่องจากความสามารถในการทำนายของแต่ละแบบจำลองอาจแตกต่างกัน

ประสิทธิภาพของการเรียนรู้กลุ่มแบบแบ๊กกิ้ง มาจากความสามารถในการลดความเสี่ยงของ overfitting ที่อาจเกิดขึ้นในแบบจำลองทั่วไป ๆ เนื่องจากการใช้ชุดข้อมูลย่อยและการรวมคำทำนายจากแบบจำลองหลายรูปแบบ นอกจากนี้ยังช่วยปรับปรุงความแม่นยำและความเสถียรของการทำนาย เนื่องจากการเฉลี่ยคำทำนายแบบต่างๆ ช่วยลดความผิดพลาด"

เพื่อให้การเรียนรู้กลุ่มแบบแบ๊กกิ้ง ทำงานได้อย่างเต็มประสิทธิภาพ การใช้แบบจำลองฐานที่แตกต่างกันและการสร้างชุดข้อมูลย่อยจากการสุ่มทับซ้อนเป็นสิ่งสำคัญ และการจัดการกับปัญหาของข้อมูลไม่สมดุลในกรณีที่จำนวนข้อมูลคลาดเคลื่อนจากกัน



ภาพประกอบ 13 แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง

ขั้นตอนการทำงานของการเรียนรู้กลุ่มแบบแบ๊กกิ้ง มีดังต่อไปนี้

1. สร้างชุดข้อมูลย่อย (Bootstrap Samples) เริ่มต้นด้วยการสร้างชุดข้อมูลย่อยหลายชุดจากชุดข้อมูลการฝึกอบรมต้นฉบับ โดยกระบวนการที่เรียกว่า bootstrap ซึ่งเป็นการสุ่มเลือกข้อมูลมาจากชุดข้อมูลต้นฉบับโดยการสุ่มทับซ้อนกัน

2. การฝึกอบรมแบบจำลองหลายตัว จากนั้นจะใช้แต่ละชุดข้อมูลย่อยในการฝึกอบรมแบบจำลองต่าง ๆ ที่เราต้องการจะใช้ ซึ่งอาจเป็นแบบจำลองเดียวกันหรือแบบจำลองที่แตกต่างกัน การทำนายของแต่ละแบบจำลองจะถูกบันทึกเพื่อใช้ในขั้นตอนถัดไป

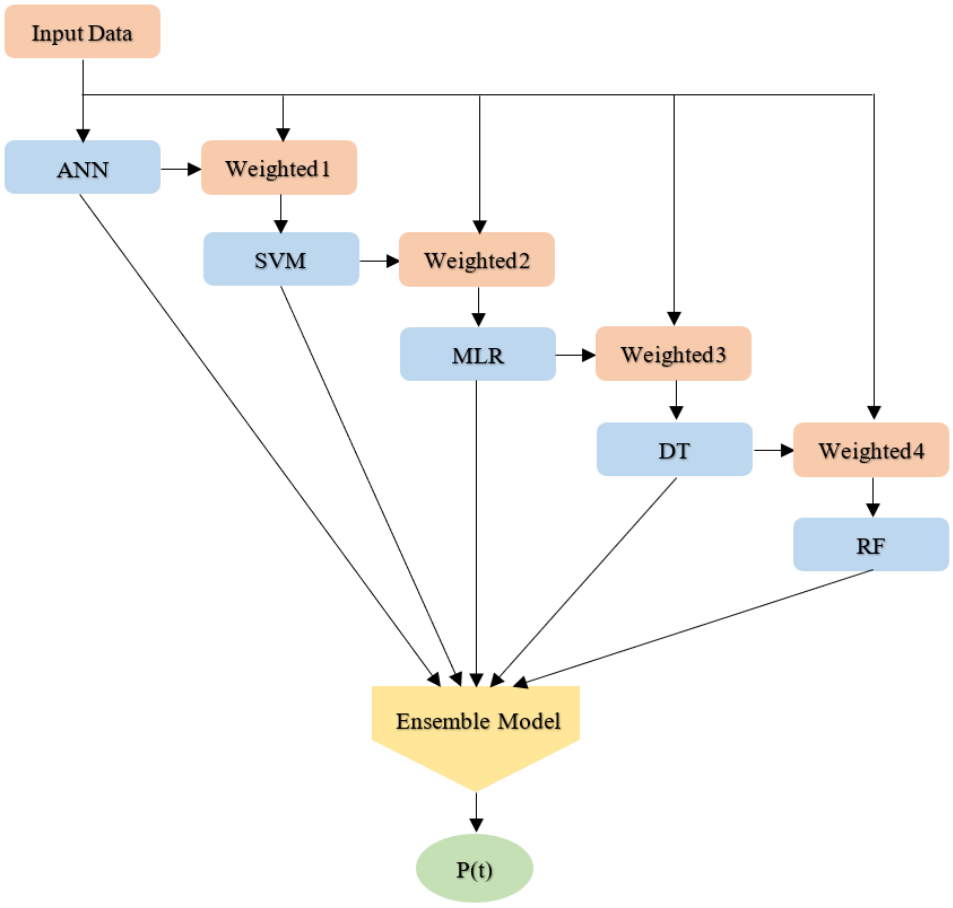
3. การรวมค่าการทำนาย (Aggregating Predictions) เมื่อแต่ละแบบจำลองฝึกอบรมและทำนายกับชุดข้อมูลย่อยเสร็จสิ้น การทำนายจากแต่ละแบบจะถูกรวมกันเพื่อสร้างค่าทำนายสุดท้าย ในกรณีของการการเรียนรู้กลุ่มแบบแบ๊กกิ้ง มักใช้การเฉลี่ยของค่าทำนายจากแต่ละแบบจำลอง

4. การประเมินผลกับชุดตรวจสอบ เพื่อประเมินประสิทธิภาพของการทำนายของการเรียนรู้จะใช้ชุดข้อมูลที่ไม่ได้ใช้ในการฝึกอบรมเรียกว่าชุดตรวจสอบ (Validation Set) และวัดค่าความแม่นยำของการทำนาย

5. การใช้งานบนชุดทดสอบ เมื่อได้รับผลการประเมินที่พอใจกับการเรียนรู้กลุ่มแบบแบ๊กกิ้ง การทำนายสุดท้ายสามารถทำได้โดยใช้การเรียนรู้กลุ่มที่ถูกฝึกอบรมจากข้อมูลทั้งหมดกับชุดข้อมูลทดสอบ (Test Set)

แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง ช่วยลดความผิดพลาดและความเสี่ยงของปัญหาการเรียนรู้เกินไป (overfitting) โดยการใช้ชุดข้อมูลย่อยและรวมคำทำนายจากแบบจำลองหลาย ๆ ตัวและมักนิยมในการปรับแบบจำลองที่ซับซ้อน เช่น แรนดอมฟอเรสต์ที่เป็นการเรียนรู้กลุ่มแบบ แบ๊กกิ้งของต้นแบบการสร้างต้นไม้การตัดสินใจ

ข้อได้เปรียบหากการใช้เทคนิคแบ๊กกิ้ง คือเป็นการช่วยลดความแปรปรวนและปัญหาการเรียนรู้เกินไปในการทำนาย โดยการใช้ประโยชน์จากความรู้ของแบบจำลองหลากหลายตัวที่ถูกฝึกด้วยชุดข้อมูลย่อยที่แตกต่างกัน การรวมการทำนายหลาย ๆ อันด้วยเทคนิคแบ๊กกิ้ง จะทำให้เกิดการทำนายสุดท้ายที่เสถียรและทนทานมากขึ้นซึ่งมักจะสามารถใช้กับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อนได้ดี อย่างแรนดอมฟอเรสต์ถือเป็นตัวอย่างที่นิยมใช้ในเทคนิคแบ๊กกิ้ง โดยที่จะมีการฝึกหลายต้นไม้ตัดสินใจบนชุดข้อมูลฝึกฝนย่อยที่แตกต่างกัน และผลการทำนายของพวกเขาจะถูกรวมกันเพื่อทำการทำนายสุดท้าย การใช้เทคนิคแบ๊กกิ้ง เป็นที่นิยมในการเรียนรู้ของเครื่อง เนื่องจากความง่ายและประสิทธิภาพในการเพิ่มประสิทธิภาพของแบบจำลองและลดปัญหาการเรียนรู้เกินไป



ภาพประกอบ 14 แบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง

5. การเรียนรู้กลุ่มแบบบูสต์ติง

การเรียนรู้แบบการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) เป็นเทคนิคหนึ่งที่ใช้ในการปรับปรุงประสิทธิภาพของแบบจำลองในการทำนาย โดยการเรียนรู้กลุ่มแบบนี้มุ่งเน้นการเรียนรู้แบบต่อเนื่องเพื่อปรับปรุงการทำนายของแบบจำลองในทุก ๆ รอบ การทำนายที่ผิดพลาดมากในรอบก่อนหน้าจะได้รับการชดเชยด้วยการทำนายที่ดีขึ้นในรอบถัดไป

ขั้นตอนหลักในการเรียนรู้แบบการเรียนรู้กลุ่มแบบบูสต์ติง ประกอบด้วยการฝึกอบรมแบบจำลองหลาย ๆ ตัวบนชุดข้อมูลการฝึกอบรมแต่ละรอบ แบบจำลองจะต้องแก้ไขความผิดพลาดของการทำนายจากรอบก่อนหน้า โดยให้ความสำคัญกับข้อมูลที่ถูกพลาดมากขึ้น ทำให้แบบจำลองสามารถโฟกัสในการจับความแม่นยำในข้อมูลที่ยากต่อการทำนาย (Rahimi et al., 2023) ดังภาพประกอบ 14

แบบจำลองที่ได้รับการฝึกอบรมจะมีน้ำหนักที่กำหนดโดยความผิดพลาดในการทำนายของแบบจำลองที่ผ่านมา แบบจำลองที่มีความผิดพลาดน้อยกว่าจะมีน้ำหนักสูงขึ้น เมื่อแบบจำลองถูกบวกเข้าไปในการเรียนรู้กลุ่ม การคำนวณความเสี่ยงโดยคำนวณน้ำหนักของแต่ละแบบจำลอง

การเรียนรู้กลุ่มแบบบูสต์ติง มีข้อได้เปรียบในการลดความผิดพลาดของแบบจำลองอย่างมีประสิทธิภาพ และมักนิยมในแบบจำลองที่ซับซ้อน เช่น ต้นไม้การตัดสินใจ (Decision Trees) ที่ถูกเรียกว่า Gradient Boosting Trees และแบบจำลองอื่น ๆ

ขั้นตอนของการเรียนรู้แบบการเรียนรู้กลุ่มแบบบูสต์ติง มีดังต่อไปนี้

1. ฝึกอบรมแบบจำลองพื้นฐาน (Base Models) เริ่มต้นด้วยการฝึกอบรมแบบจำลองพื้นฐาน (base models) หลาย ๆ ตัวบนชุดข้อมูลการฝึกอบรมเดิม แบบจำลองพื้นฐานเหล่านี้อาจเป็นต้นไม้การตัดสินใจ (Decision Trees) ที่มีความลึกต่างกันหรือแบบจำลองอื่น ๆ ที่เหมาะสม

2. การทำนายและคำนวณความผิดพลาด แบบจำลองพื้นฐานที่ถูกฝึกอบรมจะใช้สำหรับการทำนายข้อมูลการฝึกอบรม และคำนวณค่าความผิดพลาดของการทำนายโดยเปรียบเทียบกับค่าเป้าหมาย (target) ที่ถูกต้อง

3. การกำหนดน้ำหนักและปรับแบบจำลอง แบบจำลองที่มีการทำนายผิดพลาดมากขึ้นจะได้รับน้ำหนักสูงขึ้น โดยใช้น้ำหนักเหล่านี้ในการปรับแบบจำลอง เป้าหมายคือการแก้ไขความผิดพลาดในการทำนายของแบบจำลองที่ผ่านมา

4. การสร้างแบบจำลองใหม่และปรับปรุง แบบจำลองใหม่จะถูกสร้างขึ้นโดยใช้การปรับน้ำหนักในการทำนายจากแบบจำลองที่ผ่านมา การสร้างแบบจำลองใหม่นี้จะเน้นการแก้ไขความผิดพลาดของแบบจำลองก่อนหน้า ขั้นตอนนี้จะทำซ้ำจนกว่าการปรับปรุงของแบบจำลองจะไม่นำไปสู่การลดความผิดพลาดต่อไป

5. การรวมผลการทำนาย การทำนายจากแบบจำลองที่ผ่านการปรับปรุงจะถูกรวมกันเพื่อให้ได้การทำนายสุดท้าย วิธีการรวมผลอาจเป็นการเฉลี่ยค่าทำนายหรือใช้วิธีอื่น ๆ ที่เหมาะสม

6. การประเมินและการใช้งาน ในขั้นตอนสุดท้ายแบบจำลองที่ได้รับการปรับปรุงด้วยการเรียนรู้กลุ่มแบบบูสต์ติง จะถูกประเมินความสามารถในการทำนายบนชุดข้อมูลทดสอบ การใช้งานของแบบจำลองที่ปรับปรุงแล้วจะสามารถทำนายผลในข้อมูลใหม่ได้

ในกระบวนการการเรียนรู้กลุ่มแบบบูสต์ติง ความสำคัญอยู่ในการปรับปรุงและการทำนายใหม่ตลอดเวลาเพื่อปรับปรุงความแม่นยำของแบบจำลอง ข้อดีของการเรียนรู้กลุ่มแบบนี้คือสามารถปรับปรุงความสามารถในการทำนายของแบบจำลองได้ตามความยากลำบากของข้อมูลและสามารถใช้กับแบบจำลองที่ต่างกันได้

ข้อได้เปรียบที่สำคัญของเทคนิคการเรียนรู้กลุ่มแบบบูสต์ติง คือสามารถเพิ่มความแม่นยำในการทำนายอย่างมีนัยสำคัญเมื่อเปรียบเทียบกับแบบจำลองที่อ่อนแอเดี่ยว โดยการสร้างชุดแบบจำลองตามลำดับซึ่งสามารถแก้ไขข้อผิดพลาดของกันและกัน ทำให้เกิดแบบจำลองการเรียนรู้กลุ่มที่แข็งแกร่งซึ่งสามารถเข้าถึงความสัมพันธ์ที่ซับซ้อนในข้อมูลได้ ตัวอย่างของอัลกอริทึมการเรียนรู้แบบการเรียนรู้กลุ่มที่ได้รับความนิยมคือ Gradient Boosting Machines และ AdaBoost Gradient Boosting Machines ทำการลดค่าฟังก์ชันความสูญเสียแบบกรณีเดียวกันโดยการเพิ่มแบบจำลองอ่อนและ AdaBoost กำหนดน้ำหนักสูงกับข้อมูลที่ถูกจำแนกผิด เพื่อให้แบบจำลองอ่อนแอที่มาในขั้นต่อไปสามารถใช้ความสนใจในจุดดังกล่าวได้ การเรียนรู้แบบการเรียนรู้กลุ่มแบบบูสต์ติง เป็นเทคนิคที่ใช้กันอย่างแพร่หลายในการเรียนรู้เครื่องแบบจำลอง และมักมีประสิทธิภาพสูงกว่าแบบจำลองเดี่ยวและเทคนิคการเรียนรู้กลุ่มอื่น ๆ ทำให้เป็นเครื่องมือที่มีคุณค่าสำหรับการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบและงานที่ซับซ้อนในการทำนายแบบ regression อื่น ๆ ด้วย

โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมเป็นการจำลองการทำงานของสมองมนุษย์ ด้วยโปรแกรมคอมพิวเตอร์มาจากแนวคิดที่ต้องการให้คอมพิวเตอร์มีความสามารถในการเรียนรู้เหมือนมนุษย์เพื่อที่จะสามารถฝึกฝน และสอนให้สามารถนำไปแก้ปัญหาต่างๆได้โดยเป็นการจำลองการทำงานโครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาทที่เรียกว่านิวรอน (Neurons) ซึ่งมีส่วนของการประมวลผล เรียกว่านิวรอน ทุกๆ นิวรอนสามารถรับอินพุตได้จากหลายแหล่ง แต่สามารถส่งเอาต์พุตได้เพียงเส้นทางเดียว และเอาต์พุตแต่ละตัวจะแบ่งปันไปยังอินพุตของนิวรอนอื่นๆ ภายในโครงข่าย การติดต่อกันระหว่างนิวรอนทุกๆ อินพุตมีค่าน้ำหนักที่กำหนดระดับความสำคัญของการติดต่อภายใน และมีบทบาทสำคัญในการช่วยตัดสินใจในการประมวลผลข้อมูลต่างๆ ที่ผ่านมาในระบบ

ประสาทของมนุษย์ได้อย่างมีประสิทธิภาพ การติดต่อกันภายในระหว่างนิวรอนไม่ใช่ลักษณะการต่อแบบธรรมดาทุก ๆ อินพุตจะมีน้ำหนักเป็นตัวกำหนดกำลังของการติดต่อภายในและช่วยในการตัดสินใจ การทำงานของนิวรอนในบางโครงข่ายจะถูกกำหนดไว้ตายตัว แต่บางโครงข่ายสามารถที่จะปรับแต่งได้ซึ่งอาจจะเป็นการปรับแต่งจากภายนอกโครงข่ายหรือนิวรอนสามารถปรับได้ด้วยตัวเอง ในจุดนี้แสดงถึงความสามารถในการเรียนรู้และจดจำของโครงข่ายประสาทเทียม

1. ส่วนประกอบโครงข่ายประสาทเทียม

1) ข้อมูลป้อนเข้า (Input) เป็นข้อมูลเชิงปริมาณซึ่งเป็นตัวแปรอิสระสำหรับฟังก์ชั่นของโครงข่าย

2) ข้อมูลส่งออก (Output) เป็นผลลัพธ์ที่เกิดขึ้นจากกระบวนการเรียนรู้ของโครงข่าย

3) ค่าน้ำหนัก (Weights) เป็นค่าที่ได้จากการเรียนรู้ของโครงข่ายประสาทเทียม ค่านี้จะถูกเก็บเพื่อจดจำข้อมูลอื่นๆ อยู่ในรูปแบบเดียวกับที่ได้ทำการสอน (Train) โครงข่าย

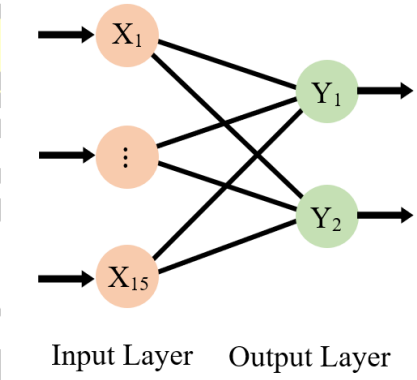
4) ฟังก์ชั่นผลรวม (Summation Function) เป็นฟังก์ชั่นผลรวมข้องข้อมูลป้อนเข้าและค่าน้ำหนัก

5) ฟังก์ชั่นการแปลง (Transfer Function) เป็นฟังก์ชั่นในการคำนวณการจำลองการทำงานของโครงข่ายประสาทเทียม

2. ลักษณะของโครงข่ายประสาทเทียม

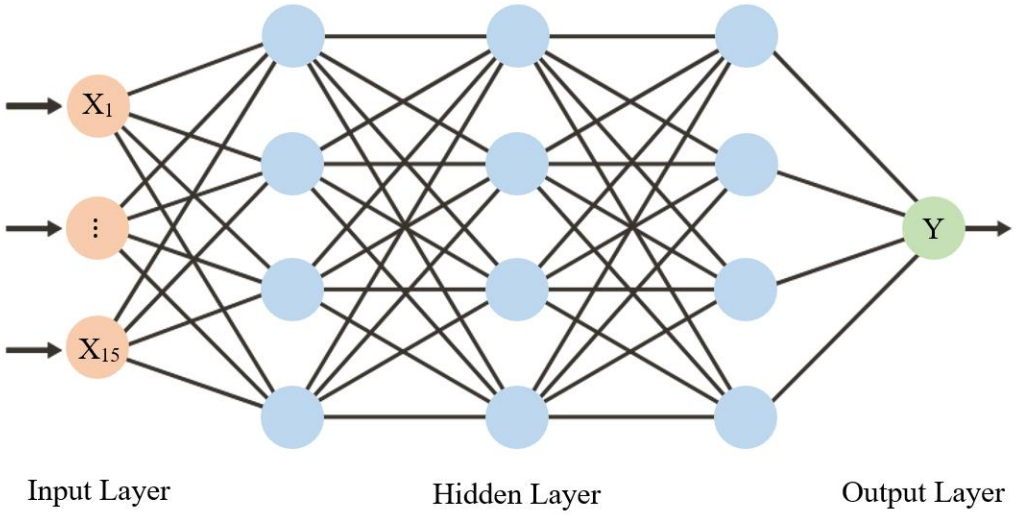
ลักษณะของโรงข่ายประสาทเทียมสามารถจำแนกได้ตามจำนวนชั้นแอบแฝง (Hidden Layer) ดังต่อไปนี้

1) โครงข่ายประสาทเทียมแบบชั้นเดียวเป็นโครงข่ายที่ไม่มีจำนวนชั้นแอบแฝงมีเพียงชั้นอินพุตและชั้นเอาต์พุตเท่านั้น ดังภาพประกอบ 15



ภาพประกอบ 15 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว

2) โครงข่ายประสาทเทียมแบบหลายชั้น เป็นโครงข่ายที่มีจำนวนชั้นแอบแฝงตั้งแต่หนึ่งชั้นขึ้นไป ดังภาพประกอบ 16



ภาพประกอบ 16 โครงสร้างโครงข่ายประสาทเทียมแบบหลายชั้น

จำนวนชั้นของโครงข่ายจะเป็นตัวกำหนดความสามารถในการแก้ปัญหาของโครงข่ายประสาทเทียม โดยทั่วไปโครงข่ายแบบชั้นเดียวจะสามารถแก้ปัญหาได้เพียงปัญหาเชิงเส้น ในขณะที่โครงข่ายแบบหลายชั้น สามารถที่จะแก้ปัญหาได้แทบทุกประเภท

โครงข่ายประสาทเทียม (Artificial Neural Network, ANN) เป็นแบบจำลองทางคณิตศาสตร์ที่จำลองโครงสร้างและการทำงานของระบบประสาทเทียมในสมองมนุษย์ เพื่อใช้ในการแก้ปัญหาต่าง ๆ โดยเฉพาะการเรียนรู้จากข้อมูล (Machine Learning) และการทำนาย (Prediction) ข้อมูล ขั้นตอนการทำงานของ ANN สามารถอธิบายได้ดังนี้

1. กำหนดโครงสร้างของโครงข่าย กำหนดจำนวนชั้นของโครงข่าย (Layers) และจำนวนเซลล์ประสาท (Neurons) ในแต่ละชั้น และกำหนดวิธีการเชื่อมต่อระหว่างเซลล์ประสาทในชั้นกันและกัน ซึ่งเรียกว่า Weight หรือค่าน้ำหนัก

2. Feedforward (Forward Propagation) ข้อมูลนำเข้า (Input) ถูกส่งผ่านชั้นแรก (Input Layer) และคำนวณค่าผ่านฟังก์ชันเชื่อมต่อระหว่างเซลล์ประสาทในแต่ละชั้น และข้อมูลถูกส่งผ่านชั้นต่อไปโดยนำค่าผ่านฟังก์ชัน Activation Function เพื่อให้ผลลัพธ์ออกมาเป็นค่าระหว่าง 0 กับ 1 หรือค่าอื่น ๆ ที่เหมาะสม

3. การคำนวณ Loss ประเมินความแตกต่างระหว่างผลลัพธ์ที่ได้จากแบบจำลองและค่าเป้าหมาย (Ground Truth) โดยใช้ฟังก์ชัน Loss Function ซึ่งบ่งบอกถึงประสิทธิภาพของแบบจำลองในการทำนาย

4. Backpropagation ใช้วิธีการ Backpropagation เพื่อปรับค่าน้ำหนัก (Weights) ของเซลล์ประสาทในแต่ละชั้นเพื่อลดค่า Loss โดยใช้ Gradient Descent และคำนวณ Gradient ของ Loss Function ตามค่าความคลาดเคลื่อนของการทำนายจากผลลัพธ์เป้าหมาย

5. การปรับค่าน้ำหนัก (Weights) และการเรียนรู้ ใช้ค่า Gradient ที่คำนวณได้จาก Backpropagation ในการปรับค่าน้ำหนักของเซลล์ประสาทในโครงข่าย เพื่อให้ค่า Loss ลดลง และการปรับค่าน้ำหนักทำให้แบบจำลองปรับตัวให้มีประสิทธิภาพในการทำนายข้อมูล

6. การทำนาย พร้อมที่แบบจำลองจะถูกนำไปใช้งาน ให้นำข้อมูลใหม่เข้าสู่โครงข่าย ผ่านชั้นแรกไปจนถึงชั้นสุดท้าย (Output Layer) เพื่อทำนายผลลัพธ์

การปรับแต่งและการตรวจสอบผลลัพธ์ ปรับแต่งพารามิเตอร์ต่างๆ เช่น จำนวนชั้นของโครงข่าย จำนวนเซลล์ประสาทในแต่ละชั้น และฟังก์ชัน Activation เพื่อปรับปรุงประสิทธิภาพของแบบจำลอง และประเมินผลลัพธ์ของแบบจำลอง ANN โดยใช้ข้อมูลทดสอบ (test data) เพื่อวัดประสิทธิภาพและความสามารถในการทำนายของแบบจำลอง

การปรับแต่งและการตรวจสอบผลลัพธ์เป็นขั้นตอนที่สำคัญเพื่อให้ได้แบบจำลองโครงข่ายประสาทเทียมที่มีประสิทธิภาพและสามารถใช้งานกับข้อมูลจริงได้อย่างมีประสิทธิภาพสูงสุด

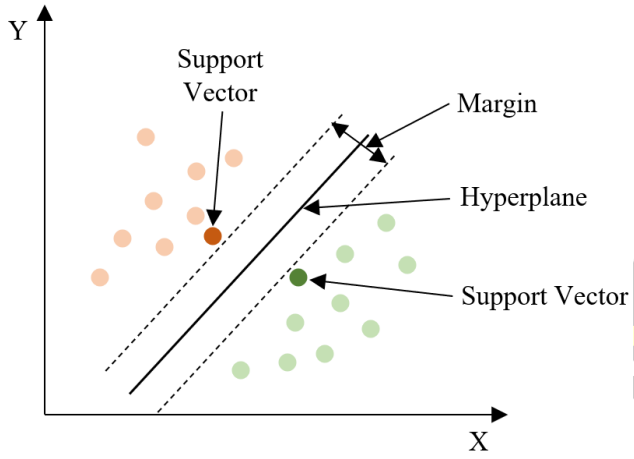
ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines, SVM) เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด อีกทั้งซัพพอร์ตเวกเตอร์แมชชีนยังมีความยืดหยุ่นและทำงานได้ดี โดยเฉพาะอย่างยิ่งเมื่อข้อมูลมีความซับซ้อน (หลาย Feature) แต่จำนวนตัวอย่างไม่มาก

จากภาพประกอบ 17 เป็นปัญหา Binary classification เราต้องการจำแนกข้อมูลออกเป็นสองพวก คือสีเขียวและสีแดง สิ่งที่ซัพพอร์ตเวกเตอร์แมชชีนทำ คือการหาเส้นแบ่งการตัดสินใจที่เป็นเส้นทึบ ซึ่งเส้นนี้จะเกิดขึ้นระหว่างกลางของเส้นประด้านซ้ายและขวา โดยมีเงื่อนไขว่าจะต้องหาคู่ของเส้นประที่กว้างที่สุดเท่าที่จะเป็นไปได้ ซัพพอร์ตเวกเตอร์แมชชีนจะพยายามหาเส้นแบ่ง (หรือพื้นผิวแบ่ง) ที่สามารถแบ่งกลุ่มข้อมูลออกเป็นสองกลุ่มได้อย่างเข้ากันได้มากที่สุด โดยที่ระยะห่างระหว่าง

ข้อมูลใกล้กับเส้นแบ่ง (หรือพื้นผิวแบ่ง) มากที่สุดเท่าที่เป็นไปได้ การเลือกเวกเตอร์สนับสนุน (Support Vectors) ตัวอย่างข้อมูลที่อยู่ใกล้เส้นแบ่งและส่งเสริมให้เส้นแบ่งแนวนอนหรือแนวตั้ง ถือเป็นเวกเตอร์สนับสนุน แบบจำลอง SVM จะใช้เวกเตอร์สนับสนุนเพื่อประมาณค่าพารามิเตอร์ในการสร้างเส้นแบ่งหรือพื้นผิวแบ่ง ต่อจากนั้นจะทำการแยกข้อมูล SVM จะใช้เวกเตอร์สนับสนุนเพื่อสร้างเส้นแบ่งหรือพื้นผิวแบ่งที่มีระยะห่างระหว่างข้อมูลใกล้สุดมากที่สุด นี่คือเส้นแบ่งหรือพื้นผิวแบ่งที่อยู่ระหว่างกลุ่มข้อมูลและมีระยะห่างระหว่างข้อมูลใกล้สุดมากที่สุด และสุดท้ายการทำนาย เมื่อมีข้อมูลใหม่เข้ามา SVM จะใช้เส้นแบ่งหรือพื้นผิวแบ่งที่ได้มาในขั้นตอนก่อนหน้าเพื่อทำนายกลุ่มข้อมูลที่ข้อมูลใหม่ตั้งอยู่

ประเภทของซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine, SVM) สามารถจำแนกออกได้เป็น 2 ประเภทและใช้สำหรับข้อมูลนำเข้าที่แตกต่างกัน ได้แก่ 1) SVM สำหรับการจำแนก (Classification SVM) ใช้สำหรับงานการจำแนกโดยที่เป้าหมายคือการแบ่งข้อมูลเป็นกลุ่ม เช่น การจำแนกรูปภาพว่าเป็นแมวหรือหมา และ 2) SVM สำหรับการทำนายตัวเลข (Regression SVM) ใช้สำหรับงานการทำนายตัวเลขโดยที่เป้าหมายคือการทำนายค่าตัวเลข เช่น การทำนายราคาของบ้านพักอาศัย ซึ่งในงานวิจัยครั้งนี้จะเลือกใช้ SVM สำหรับการทำนายที่เป็นตัวเลขหรือ Regression เพื่อทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ



ภาพประกอบ 17 แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีนเป็นอัลกอริทึมที่ใช้ในการแยกแยะข้อมูลที่ซับซ้อนโดยการสร้างเส้นแบ่งหรือพื้นผิวแบ่งในฟีเจอร์สเป็ซ (Feature Space) เพื่อแบ่งกลุ่มข้อมูลออกเป็นสองกลุ่ม วิธีการของซัพพอร์ตเวกเตอร์แมชชีนใช้เส้นที่เรียกว่าเส้นแบ่ง (Hyperplane) เพื่อแบ่งข้อมูลทั้งสองกลุ่มออก

จากกัน และเป้าหมายของซัพพอร์ตเวกเตอร์แมชชีนคือการหาเส้นแบ่งที่ดีที่สุดที่แยกข้อมูลสองกลุ่มได้อย่างมีประสิทธิภาพสูงสุด

เมื่อเรามีกลุ่มข้อมูลที่ต้องการแยกแยะ แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน จะค้นหาเส้นแบ่งหรือพื้นผิวแบ่งที่สามารถแบ่งกลุ่มข้อมูลออกเป็นสองกลุ่มโดยที่ระยะห่างระหว่างข้อมูลใกล้กับเส้นแบ่ง (หรือพื้นผิวแบ่ง) มากที่สุด ความห่างระหว่างข้อมูลใกล้กับเส้นแบ่งนี้เรียกว่า Margin และเส้นแบ่งที่มี Margin มากที่สุดเรียกว่าเส้นแบ่ง (Hyperplane) ที่ดีที่สุด

การค้นหาเส้นแบ่งที่ดีที่สุดในซัพพอร์ตเวกเตอร์แมชชีนนั้นทำได้โดยใช้เทคนิคเชิงคณิตศาสตร์ และเมื่อเราไม่สามารถแบ่งแยกข้อมูลได้โดยใช้เส้นตรงธรรมดา ซัพพอร์ตเวกเตอร์แมชชีนสามารถใช้เคิร์นิลเกอร์ (Kernel Function) เพื่อเปลี่ยนแปลงรูปแบบของข้อมูลให้เป็นรูปแบบที่เหมาะสมแก้ปัญหาได้ ส่วนข้อมูลที่อยู่บน Margin ของเส้นแบ่งถูกเรียกว่า Support Vectors ซึ่งเป็นข้อมูลที่มีบทบาทสำคัญในการกำหนดเส้นแบ่ง

แนวคิดหลักของซัพพอร์ตเวกเตอร์แมชชีนไม่ได้เน้นการจัดเรียงข้อมูลเชิงเส้น แต่เน้นการค้นหาเส้นแบ่งที่มี Margin มากที่สุด และกระบวนการการเรียนรู้ที่ผ่านการวนซ้ำเพื่อหาวิธีแก้ปัญหาที่ดีที่สุด การนำรูปแบบทางคณิตศาสตร์เข้ามาช่วยในการแยกแยะข้อมูลที่ซับซ้อนทำให้ซัพพอร์ตเวกเตอร์แมชชีนเป็นเครื่องมือที่สามารถจัดการกับการแยกแยะข้อมูลที่ซับซ้อนได้อย่างมีประสิทธิภาพและทันสมัย การเลือกคำนวณทางคณิตศาสตร์และการปรับแต่งพารามิเตอร์อาจจะทำให้การใช้งานยากขึ้นบ้าง แต่ซัพพอร์ตเวกเตอร์แมชชีนเป็นเครื่องมือที่สามารถนำมาใช้งานกับข้อมูลที่มีความซับซ้อนและมิติสูงได้อย่างมีประสิทธิภาพและถูกต้อง

ขั้นตอนของการสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนสำหรับงานแยกแยะ (classification) ข้อมูลสามารถสรุปได้ดังนี้

1. การเตรียมข้อมูล เริ่มต้นด้วยการเตรียมข้อมูลที่ต้องการใช้ในการสอน (training data) โดยมีข้อมูลแต่ละตัวมาพร้อมกับป้ายกำกับว่าเป็นกลุ่มใด (Class 0 หรือ Class 1 ย่อมาจากแต่ละกลุ่มข้อมูล) และ ปรับรูปแบบของข้อมูลให้เหมาะสมสำหรับซัพพอร์ตเวกเตอร์แมชชีน (บางครั้งจะต้องมีการปรับสเกลข้อมูล)

2. เลือก Kernel Function ที่เหมาะสมต่อปัญหาที่ต้องการแก้ ซึ่งสามารถเปลี่ยนแปลงรูปแบบของข้อมูลให้เหมาะสมกับการแยกแยะ

3. กำหนดพารามิเตอร์ C ที่กำหนดระดับของความผิดพลาดที่ยอมรับได้ เพื่อควบคุมการเรียนรู้เกินไปและความซับซ้อนของแบบจำลอง และ ในบางกรณี อาจต้องปรับพารามิเตอร์ของ Kernel Function อีกด้วย

4. การสร้างแบบจำลอง ใช้ข้อมูลที่เตรียมไว้ในขั้นตอนที่ 1 เพื่อสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนโดยคำนวณเส้นแบ่งหรือพื้นผิวแบ่งที่มี Margin มากที่สุด

5. การตั้งค่า Margin และการค้นหาเส้นแบ่ง หา Margin และเส้นแบ่งที่มี Margin มากที่สุดโดยคำนวณระยะห่างของข้อมูลถึง Margin ทั้งสองข้างของเส้นแบ่ง

6. การจัดเรียงข้อมูล เลือกข้อมูลที่อยู่บน Margin ทั้งสองข้างของเส้นแบ่ง โดยข้อมูลเหล่านี้จะเรียกว่า Support Vectors เนื่องจากมีบทบาทสำคัญในการกำหนดเส้นแบ่ง

7. การทำนาย เมื่อได้แบบจำลองซัพพอร์ตเวกเตอร์แมชชีนที่เรียนรู้แล้ว ใช้แบบจำลองนี้ในการทำนายกลุ่มข้อมูลใหม่ที่ไม่ได้มีป้ายกำกับ โดยให้แบบจำลองพิจารณาเส้นแบ่งหรือพื้นผิวแบ่งที่ได้

8. การปรับแต่งและการตรวจสอบผลลัพธ์ ปรับพารามิเตอร์ต่างๆ เช่น C และพารามิเตอร์ของ Kernel Function เพื่อปรับปรุงประสิทธิภาพของแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนและประเมินผลลัพธ์ของแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนโดยใช้ข้อมูลทดสอบ (test data) เพื่อวัดประสิทธิภาพและความสามารถในการทำนายของแบบจำลอง การปรับแต่งและการตรวจสอบผลลัพธ์เป็นขั้นตอนที่สำคัญเพื่อให้ได้แบบจำลองซัพพอร์ตเวกเตอร์แมชชีนที่มีประสิทธิภาพและสามารถใช้งานกับข้อมูลจริงได้อย่างมีประสิทธิภาพสูงสุด

ขั้นตอนของการสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนสำหรับงาน Regression จะมีความแตกต่างจากการใช้กับงานแยกแยะ (classification) บางส่วน ต่อไปนี้คือขั้นตอนหลักในการใช้ซัพพอร์ตเวกเตอร์แมชชีนสำหรับงาน Regression

1. การเตรียมข้อมูล เตรียมข้อมูลที่ต้องการใช้ในการสอน (training data) ซึ่งประกอบด้วยคู่ของข้อมูลเชิงตัวเลข (feature) และค่าเป้าหมายที่ต้องการทำนาย (target)

2. เลือก Kernel Function ที่เหมาะสมต่องาน Regression ตามลักษณะของข้อมูล ส่วนมากจะใช้ Linear Kernel, Polynomial Kernel, หรือ Radial Basis Function (RBF) Kernel

3. กำหนดพารามิเตอร์ C และพารามิเตอร์ของ Kernel Function ตามความเหมาะสมของงาน Regression

4. การสร้างแบบจำลอง ใช้ข้อมูลที่เตรียมไว้ในขั้นตอนที่ 1 เพื่อสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน Regression โดยหาเส้นแบ่งหรือพื้นผิวแบ่งที่เหมาะสมเพื่อให้ค่าเป้าหมายในข้อมูลที่ใกล้กับเส้นแบ่งมากที่สุด

5. การคำนวณค่าเป้าหมาย หลังจากสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน Regression เสร็จสิ้น ใช้แบบจำลองนี้ในการทำนายค่าเป้าหมายสำหรับข้อมูลที่ไม่ได้มีค่าเป้าหมาย โดยให้แบบจำลองคำนวณค่าเป้าหมายที่เหมาะสมตามพื้นผิวแบ่งหรือเส้นแบ่ง

6. การปรับแต่งและการตรวจสอบผลลัพธ์ ปรับพารามิเตอร์ต่างๆ โดยดูผลลัพธ์และประสิทธิภาพของแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน Regression ที่ได้ และประเมินผลลัพธ์ของแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน Regression โดยใช้ข้อมูลทดสอบ (test data) เพื่อวัดประสิทธิภาพและความสามารถในการทำนายค่าเป้าหมายของแบบจำลอง

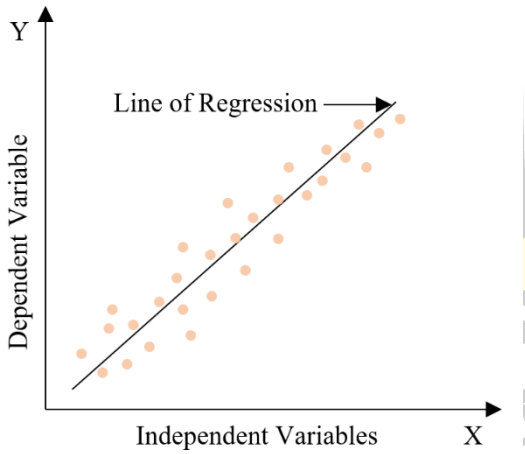
ความแตกต่างของแบบจำลองซัพพอร์ตเวกเตอร์แมชชีนระหว่าง SVM Regression และ SVM Classification อย่างหนึ่งคือใน SVM Regression เราจะไม่มีการคำนวณ Margin และเส้นแบ่งที่ดีที่สุด เนื่องจากเราไม่ได้มีกลุ่มข้อมูลที่ต้องการแยกแยะ แต่เราจะคำนวณค่าเป้าหมายในบริเวณใกล้กับเส้นแบ่งหรือพื้นผิวแบบเส้นตรงที่สร้างขึ้น ทำให้เราสามารถทำนายค่าตัวเลขที่เป็นเป้าหมาย ซึ่งเป็นตัวเลขแทนที่จะเป็นคลาสของกลุ่มได้

การวิเคราะห์ความถดถอย

การวิเคราะห์ถดถอยแบ่งตามลักษณะของข้อมูลและสมการถดถอย หรือรูปแบบการวิเคราะห์ถดถอย (regression analysis model) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ได้เป็น 4 ประเภท ดังนี้

1. รูปแบบการวิเคราะห์ถดถอยพหุคูณ (multiple regression analysis model) เป็นรูปแบบการวิเคราะห์ถดถอยใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรตามหนึ่งตัวและตัวแปรอิสระหลายตัวแปร โดยที่ Y เป็นตัวแปรตามและ X_1, X_2, ..., X_{15} เป็นตัวแปรอิสระ เมื่อตัวแปรมีความสัมพันธ์แบบเส้นตรงดังภาพประกอบ 18 และสามารถหาค่าตัวแปรตามได้ด้วยสมการ 9

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + e \qquad (9)$$



ภาพประกอบ 18 แบบจำลองการวิเคราะห์ถดถอยพหุคูณ

2. รูปแบบการวิเคราะห์ถดถอยโพลีโนเมียล (polynomial regression analysis model) เป็นรูปแบบการวิเคราะห์ถดถอยใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระตัวเดียวและตัวแปรตามตัวเดียว เมื่อมีความสัมพันธ์แบบเส้นโค้ง ดังสมการ 10

$$Y = b_0 + b_1X + b_2X^2 + \ldots + e \quad (10)$$

3. รูปแบบการวิเคราะห์ถดถอยพหุคูณมีตัวแปรดัมมี่ (multiple regression analysis model with dummy variables) เป็นรูปแบบการวิเคราะห์ถดถอย ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรตามหนึ่งตัวแปรกับตัวแปรอิสระที่เป็นตัวแปรนันเมตริกที่ได้รับการให้รหัสใหม่ (recode) หรือได้รับการเปลี่ยนรูป (transform) ให้เป็นตัวแปรดัมมี่ สามารถใช้ศึกษา ได้ทั้งกรณีที่ตัวแปรมีความสัมพันธ์แบบเส้นตรงและเส้นโค้ง ได้ผลการวิเคราะห์เช่นเดียวกันกับบรูปแบบการวิเคราะห์ความแปรปรวน (ANOVA model) ดังสมการ เมื่อมี D_1, D_2, ... D_{m-1} เป็นตัวแปรดัมมี่ที่สร้างขึ้นแทนตัวแปรต้น X ซึ่งเป็นตัวแปรนันเมตริกที่มี m ค่า ดังสมการ 11

$$Y = b_0 + b_1D_1 + b_2D_2 + \ldots + D_{m-1} + e \quad (11)$$

4. รูปแบบการวิเคราะห์ถดถอยพหุคูณมีเทอมปฏิสัมพันธ์ (multiple regression analysis model with interaction term) เป็นรูปแบบการวิเคราะห์ถดถอยใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรกรณีที่มีอิทธิพลทั้งอิทธิพลหลักจาก X, Z และอิทธิพลปฏิสัมพันธ์(X*Z) ต่อตัวแปรตาม ดังสมการ 12

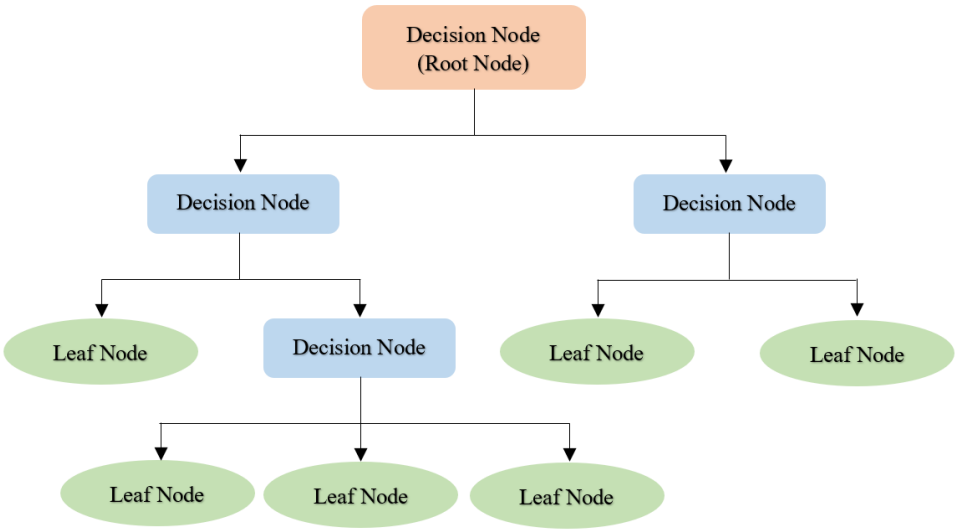
$$Y = b_0 + b_1X + b_2Z + b_2XZ + \ldots + e \quad (12)$$

นอกจากนี้ยังมีการพัฒนารูปแบบการวิเคราะห์ถดถอยเพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรในรูปเส้นโค้งตัว ‘S’ และรูปแบบแบบอื่นๆ เช่น การวิเคราะห์ถดถอยโลจิสติก (logistic regression analysis) การวิเคราะห์ล็อกลิเนียร์ (log-linear analysis) การวิเคราะห์โลจิท (logit analysis) เป็นต้น

ต้นไม้ตัดสินใจ

การสร้างแบบจำลองต้นไม้ตัดสินใจ (Decision Tree, DT) สำหรับการทำนายค่าตัวเลขในงานการทำนายแบบ regression เป็นอัลกอริทึมที่ทนทานและสามารถอธิบายได้ โดยอัลกอริทึมนี้จะสร้างแบบจำลองเป็นลักษณะของต้นไม้โดยแบ่งชุดข้อมูลออกเป็นส่วนย่อยๆ ตามค่าของคุณสมบัติต่างๆ ภายในชุดข้อมูล ที่และจุดแบ่งที่ช่วยลดความแปรปรวนหรือค่าเฉลี่ยของความผิดพลาดในตัว

แปรเป้าหมายภายในชุดย่อย จะถูกเลือกที่แต่ละโหนด กระบวนการนี้จะดำเนินการต่อไปจนกรณีที่มีเกณฑ์สำหรับการหยุด เช่นความลึกของต้นไม้สูงสุดหรือจำนวนตัวอย่างขั้นต่ำต่อใบ (leaf) การทำนายสำหรับแต่ละจุดข้อมูลจะถูกทำโดยการเดินทางผ่านต้นไม้การตัดสินใจจากโหนดรากไปยังโหนดใบที่ค่าตัวเลขที่เกี่ยวข้องกับโหนดใบจะถูกใช้เป็นการทำนายสุดท้ายสำหรับตัวแปรเป้าหมาย ดังตัวอย่างที่แสดงในภาพประกอบ 19 การใช้ต้นไม้การตัดสินใจในงาน regression จะสามารถอธิบายได้ง่าย ช่วยให้เข้าใจกระบวนการตัดสินใจอย่างชัดเจน การอธิบายปัจจัยที่มีผลต่อผลลัพธ์ทางตัวเลขที่ทำนายได้ อีกทั้งความสามารถในการปรับตัวให้เหมาะกับความสัมพันธ์ที่ไม่เป็นเส้นตรงทำให้เหมาะสำหรับการประยุกต์ใช้ในหลากหลายงาน เช่น การทำนายทางการเงิน หรือการวินิจฉัยทางการแพทย์ ที่ต้องการคำทำนายตัวเลขที่แม่นยำ อย่างไรก็ตาม เพื่อป้องกันปัญหาการเรียนรู้เกินไป เทคนิคการตัดสินใจเพาะกิจ (pruning) และการปรับพารามิเตอร์ (hyperparameter tuning) มักถูกนำมาใช้เพื่อควบคุมความซับซ้อนของแบบจำลองและให้ประสิทธิภาพในการแยกแยะที่แม่นยำและทนทาน



ภาพประกอบ 19 แบบจำลองต้นไม้ตัดสินใจ

ขั้นตอนของการสร้างแบบจำลองต้นไม้ตัดสินใจในการทำงานแบบ Regression มีดังนี้

1. เลือกคุณสมบัติ (Features) ที่ใช้ในการแยกข้อมูล คุณสมบัติเหล่านี้จะถูกใช้ในการแยกชุดข้อมูลออกเป็นส่วนย่อยๆ ตามค่าของคุณสมบัติที่เลือก ซึ่งความสัมพันธ์ของคุณสมบัติต่อตัวแปรเป้าหมายจะถูกนำมาใช้ในการแยกข้อมูลเป็นกลุ่มๆ

2. การแยกข้อมูล ในขั้นตอนนี้ต้นไม้ตัดสินใจจะแยกชุดข้อมูลออกเป็นส่วนย่อยๆ โดยใช้ค่าของคุณสมบัติที่เลือก และจะเลือกคุณสมบัติและจุดแบ่งที่ช่วยลดความแปรปรวนหรือค่าเฉลี่ยของความผิดพลาดในตัวแปรเป้าหมายภายในแต่ละชุดย่อย

3. ทำซ้ำขั้นตอนการแยกข้อมูล กระบวนการการแยกข้อมูลจะทำซ้ำเพื่อสร้างโหนดย่อยเพิ่มเติมบนแต่ละส่วนย่อยที่แยกออกมา ซึ่งกระบวนการนี้จะดำเนินต่อไปจนกรณีที่ใบหรือโหนดย่อยไม่สามารถแยกต่อไปได้อีกต่อไป

4. การหยุดเงื่อนไข การหยุดเงื่อนไขถูกกำหนดเพื่อป้องกันการสร้างต้นไม้ที่มีขนาดใหญ่เกินไปและการเรียนรู้เกินกว่าความจำเป็น ตัวอย่างเช่น ความลึกสูงสุดของต้นไม้หรือจำนวนตัวอย่างขั้นต่ำต่อใบ

5. การทำนาย เมื่อต้นไม้การตัดสินใจสร้างขึ้นแล้ว การทำนายสำหรับข้อมูลใหม่จะทำโดยการเดินทางผ่านต้นไม้ตัดสินใจจากโหนดรากไปยังโหนดใบที่เข้ากับข้อมูลใหม่ เมื่อเรามาถึงโหนดใบที่เกี่ยวข้อง เราจะใช้ค่าตัวเลขที่เกี่ยวข้องกับโหนดใบเป็นการทำนายสุดท้ายสำหรับตัวแปรเป้าหมาย

6. การป้องกันการเรียนรู้เกินไป เพื่อป้องกันการเรียนรู้เกินกว่าความจำเป็น (overfitting) บางครั้งอาจใช้เทคนิคการตัดสินใจเพาะกิจ (pruning) เพื่อลบส่วนของต้นไม้ที่ไม่สำคัญออก และการปรับแต่งพารามิเตอร์เพื่อควบคุมความซับซ้อนของต้นไม้เพื่อให้มีประสิทธิภาพในการแยกแยะที่ทั้งแม่นยำและทนทาน

ขั้นตอนเหล่านี้ช่วยให้เราสร้างแบบจำลองต้นไม้ตัดสินใจที่สามารถใช้ทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบได้อย่างมีประสิทธิภาพ

แรนดอมฟอเรสต์

แรนดอมฟอเรสต์ (Random Forest, RF) เป็นเทคนิคการเรียนรู้ร่วมที่มีการใช้สำหรับงานการทำนายแบบ regression ซึ่งเป็นการขยายต้นไม้การตัดสินใจเพื่อรวมกันเป็นต้นไม้หลายต้นหรือป่าเพื่อทำการทำนายที่แม่นยำและทนทานมากขึ้น เป็นการขยายส่วนต่อขยายแบบจำลองต้นไม้การตัดสินใจโดยการสร้างชุดของต้นไม้การตัดสินใจโดยใช้ชุดย่อยแบบสุ่มของข้อมูลการฝึกฝนและชุดย่อยแบบสุ่มของคุณสมบัติ แต่ละต้นไม้ในแรนดอมฟอเรสต์สำหรับ regression จะมีการทำนายอย่างอิสระกัน และการทำนายสุดท้ายจะได้มาจากการเฉลี่ยหรือการเลือกค่ากลางของการทำนายจากต้นไม้แต่ละต้น แนวคิดเชิงกลุ่มนี้ลดความเสี่ยงของการเรียนรู้เกินไปและเพิ่มความสามารถของแบบจำลองในการแยกแยะที่ดีสู่ข้อมูลใหม่ที่ไม่เคยเห็น

แรนดอมฟอเรสต์สำหรับ regression มีข้อดีหลายประการ เช่น การจัดการกับความสัมพันธ์ที่ไม่เป็นเส้นตรงระหว่างคุณสมบัติและตัวแปรเป้าหมาย การรับลูกเล่นที่ซับซ้อน และการลดผลกระทบของข้อมูลรบกวน การบริหารจัดการค่าแปรผิดปกติในการสร้างต้นไม้ยังทำให้แบบจำลองมีความทนทานมากขึ้นและเสถียรมากขึ้น นอกจากนี้แรนดอมฟอเรสต์ยังให้คะแนนความสำคัญของคุณสมบัติที่บ่งบอกถึงความสำคัญของแต่ละคุณสมบัติในกระบวนการทำนาย ข้อมูลเหล่านี้สามารถ

เป็นประโยชน์ในการเข้าใจว่าคุณสมบัติใดมีความสำคัญมากที่สุดในการกำหนดผลการทำนายแบบ regression

โดยรวมแล้ว การสร้างแบบจำลองแรนดอมฟอเรสต์เป็นอัลกอริทึมที่หลากหลายและมีประสิทธิภาพสำหรับงาน regression และเป็นทางเลือกที่ได้รับความนิยมในหลากหลายกลุ่ม เช่น การเงิน ด้านสุขภาพ และการค้าปลีก ที่ต้องการการทำนายตัวเลขที่แม่นยำเพื่อการตัดสินใจ

ขั้นตอนของการสร้างแบบจำลองแรนดอมฟอเรสต์สำหรับงาน Regression มีดังนี้

1. การเตรียมข้อมูล การเตรียมชุดข้อมูลที่ใช้ในการฝึกฝนแบบจำลอง รวมถึงตัวแปรเป้าหมาย (ค่าตัวเลขที่ต้องการทำนาย) และคุณสมบัติที่ใช้ในการทำนาย

2. สร้างชุดข้อมูลย่อย ๆ (Bootstrap) การสร้างชุดข้อมูลย่อยๆ โดยการสุ่มสลับตัวอย่างจากชุดข้อมูลเดิม โดยขนาดของชุดย่อยเท่ากับขนาดข้อมูลเดิม สร้างชุดย่อยเหล่านี้มากมาย

3. สร้างการตัดสินใจต้นไม้แบบสุ่ม ทำการสร้างแบบจำลองต้นไม้ตัดสินใจโดยใช้ชุดข้อมูลย่อยและชุดคุณสมบัติย่อยๆ โดยการเลือกคุณสมบัติแบบสุ่ม และจุดแบ่งที่ช่วยลดความผิดพลาดในการทำนายในแต่ละโหนด

4. ทำนายต้นไม้แต่ละต้น โดยให้แต่ละต้นไม้ในแรนดอมฟอเรสต์ทำนายตัวเลขตามข้อมูลในชุดทดสอบ ต้นไม้แต่ละต้นทำนายได้แบบอิสระ

5. รวมผลลัพธ์ รวมผลลัพธ์ที่ได้จากต้นไม้แต่ละต้น เช่น ให้ผลลัพธ์เป็นค่าเฉลี่ยหรือค่ากลางของการทำนายจากทุกต้นไม้

6. ทำการปรับค่าเพื่อป้องกันการเรียนรู้เกินไป สำหรับแต่ละต้นไม้สามารถทำการ pruning (การตัดส่วนที่ไม่สำคัญ) และปรับค่าพารามิเตอร์ เช่น จำนวนต้นไม้ที่ใช้หรือจำนวนคุณสมบัติที่เลือกเพื่อควบคุมความซับซ้อนของแบบจำลอง

7. ทดสอบและประเมิน ทำการทดสอบแบบจำลองแรนดอมฟอเรสต์ด้วยชุดข้อมูลทดสอบที่ไม่เคยเห็นมาก่อน เพื่อวัดประสิทธิภาพและความสามารถในการทำนายตัวเลข

8. วิเคราะห์ผลและความสำคัญของคุณสมบัติ ในขั้นตอนสุดท้าย คุณสามารถดูผลลัพธ์การทำนายและคะแนนความสำคัญของคุณสมบัติ เพื่อเข้าใจว่าแต่ละคุณสมบัติมีผลในการทำนายแบบไหนมากที่สุด

9. นำแบบจำลองในงานจริง เมื่อแบบจำลองแรนดอมฟอเรสต์ได้ผ่านการทดสอบและประเมิน จะสามารถนำแบบจำลองมาใช้ในงานจริงเพื่อทำนายค่าตัวเลขในการทำนายเชิงปัญหาที่คุณสนใจ อย่างเช่นการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ

ขั้นตอนนี้เป็นขั้นตอนการสร้างแบบจำลองแรนดอมฟอเรสต์เบื้องต้น ความซับซ้อนของแต่ละขั้นตอนอาจแตกต่างกันไปขึ้นอยู่กับแพคเกจ (package) หรือไลบรารี (library) ที่คุณใช้ในการ

สร้างแรนดอมฟอเรสต์แต่โดยรวมแล้ว ขั้นตอนเหล่านี้จะช่วยให้สร้างแบบจำลองแรนดอมฟอเรสต์สำหรับงาน Regression ได้อย่างเป็นประสิทธิภาพและแม่นยำ

งานวิจัยที่เกี่ยวข้อง

เทคนิคการเรียนรู้ของเครื่องถูกกำหนดให้เป็นระบบที่สามารถเรียนรู้จากข้อมูลได้ โดยทั่วไปจุดแข็งหลักของเทคนิคการเรียนรู้ของเครื่องสามารถระบุได้ดังนี้ ความสามารถในการจัดการกับความไม่แน่นอนในวิธีการ ความสามารถในการจัดการและดำเนินการกับข้อมูลที่ไม่สมบูรณ์ และความสามารถในการตัดสินใจและสรุปกรณีใหม่ตามประสบการณ์จากกรณีที่คล้ายคลึงกัน

พงศ์สยาม กันจินะ (2556) ได้ทำการศึกษาเกี่ยวกับการสร้างรูปแบบอาคารพักอาศัย โดยวิเคราะห์ตามสัดส่วนพื้นที่ใช้สอย โดยได้ทำการจำแนกประเภทพื้นที่ใช้สอยออกเป็น 8 ประเภท ได้แก่ 1) ห้องนอน 2) ห้องครัว 3) ห้องรับแขก 4) ทางเดินและโถง 5) ห้องน้ำ 6) พื้นที่จอดรถ 7)ระเบียง 8) พื้นที่ซักผ้า ซึ่งได้ทำการวิเคราะห์หาความสัมพันธ์ของตัวแปรประเภทพื้นที่แต่ละประเภทกับราคาที่ถอดปริมาณโดยละเอียด ซึ่งได้ทำการวิเคราะห์จากบ้านพักอาศัยคอนกรีตเสริมเหล็กสูงสองชั้นจำนวน 34 หลัง ซึ่งมีราคาตั้งแต่ 1.3 ล้านจนถึง 2.1 ล้านบาท และพื้นที่ใช้สอยตั้งแต่ 120 ตารางเมตรถึง 280 ตารางเมตรและรูปแบบบ้านเป็นแบบทั่วไปที่ใช้ในโครงการบ้านจัดสรร ซึ่งผลที่ได้ พบว่ารูปแบบสามารถประมาณราคาได้รวดเร็วกว่าการถอดราคาละเอียดมาก และมีความแม่นยำกว่าการประมาณราคาแบบคิดพื้นที่รวมอีกด้วย

ในงานวิจัยของ Ji et al. (2012) ได้ประยุกต์ใช้การปรับกรณีศึกษาเพื่อเพิ่มประสิทธิภาพ CBR พวกเขาแนะนำว่าการนำรูปแบบนี้ไปใช้ในโครงการประเภทอื่นเพื่อเป็นการวิจัยในอนาคต นอกจากนี้ พวกเขาเสนอว่าการใช้ปัจจัยเชิงคุณภาพมีผลกับรูปแบบ และแนะนำเป็นอย่างยิ่งให้พิจารณาความลำเอียงที่เกิดจากข้อมูลที่มาจากภูมิภาคต่างๆ นอกจากนี้ Tatari and Kucukvar (2011) ในการศึกษาของพวกเขา ได้ศึกษารูปแบบ BPANN และเปรียบเทียบกับการถดถอยในการประมาณต้นทุนของโครงการก่อสร้าง สถาปัตยกรรมที่ดีที่สุดของโครงข่ายประสาทเทียมถูกเลือกหลังจากกระบวนการทดลองและทดสอบข้อผิดพลาด ซึ่งในที่สุดโครงข่ายประสาทเทียมก็มีประสิทธิภาพที่ดีขึ้นเมื่อเทียบกับการวิเคราะห์การถดถอย ในงานวิจัยนี้ ขอแนะนำให้ใช้ชุดข้อมูลขนาดใหญ่ที่มีข้อมูลที่แม่นยำมากขึ้นในการวิจัยในอนาคต การประยุกต์ใช้รูปแบบไฮบริด (Modified PSO และ fuzzy neural network) ในการประมาณราคาโครงการก่อสร้างยังได้รับการตรวจสอบใน He et al. (2011) เป็นแนวทางใหม่ภายใน ซึ่งรูปแบบนี้สามารถนำไปใช้กับกรณีใหม่อื่นๆ นอกจากนี้ Arafa and Alqedra (2011) ได้ตรวจสอบรูปแบบ BPANN เพื่อคาดการณ์ต้นทุนของโครงการก่อสร้าง สถาปัตยกรรมที่มีแนวโน้มดีที่สุดจะถูกสร้างขึ้นหลังจากการทดลองหลายครั้ง พวกเขาแนะนำ

ให้มีชุดข้อมูลที่ใหญ่ขึ้นเป็นอินพุตสำหรับเครือข่าย เพื่อปรับปรุงประสิทธิภาพ Cheng et al. (2010) ได้บูรณาการโครงข่ายประสาทเทียมกับตรรกศาสตร์คลุมเครือเพื่อจัดการกับความไม่แน่นอนเป็นแนวทางใหม่ พวกเขาอ้างว่าโครงข่ายประสาทเทียมมีประสิทธิภาพมากกว่าโครงข่ายประสาทเทียมเพียงอย่างเดียวในการคาดการณ์ต้นทุนของโครงการก่อสร้างในระยะเริ่มต้นของโครงการ นอกจากนี้ยังมีหลักฐานที่เป็นรูปธรรมว่าโครงข่ายประสาทเทียมสามารถระบุการเชื่อมต่อทั้งแบบเชิงเส้นและแบบไม่เชิงเส้นในเลเยอร์ที่ซ่อนอยู่ได้ Cheng et al. (2009) ยังใช้ประโยชน์จาก อัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรม, Fuzzy logic และ โครงข่ายประสาทเทียม ไปพร้อม ๆ กันสำหรับการเพิ่มประสิทธิภาพระดับโลก การให้เหตุผลโดยประมาณ และการจับคู่อินพุต-เอาต์พุต ตามลำดับ วิธีการประมาณค่าใช้บังคับกับระยะเริ่มต้นของโครงการสำหรับกระบวนการตัดสินใจของผู้จัดการโครงการ

Pacifico and Ludermir (2013) ได้ศึกษา Evolutionary Extreme Learning Machine Based on Particle Swarm Optimization ถูก ได้นำเสนอวิธีการใช้ Particle Swarm (Optimization เข้ามาช่วย เช่นเดียวกันกับ Xu and Shu (2006) ในการหาค่าน้ำหนักและค่าเอนเอียงระหว่างชั้นข้อมูลนำเข้าและชั้นซ่อนที่เหมาะสมแทนการสุ่มค่า น้ำหนักและค่าแอนเอียงในรูปแบบเดิม ในงานวิจัยนี้ได้ทำการทดลองกับงานประเภททำนาย (Prediction) โดยใช้ข้อมูลผลผลิตข้าวโพดและปัจจัยที่มีต่อผลผลิตข้าวโพดในประเทศจีน

การประเมินต้นทุนของอาคารที่อยู่อาศัยโดยใช้การถดถอยเชิงเส้นแบบหลายปัจจัยได้รับการพิจารณาใน Alshemosi and Alsaad (2017) ซึ่งมีความแม่นยำถึง 92% ในตอนท้าย งานวิจัยได้แนะนำให้เปรียบเทียบผลลัพธ์กับงานวิจัยที่ใช้เทคนิคโครงข่ายประสาทเทียมเพื่อดูความแตกต่าง อันที่จริง การศึกษาได้สนับสนุนการใช้รูปแบบการประมาณราคาในโครงการก่อสร้างแทนวิธีการทั่วไป ใน Bala et al. (2014) สามารถมองเห็นการประยุกต์ใช้ Back-Propagation Artificial Neural Network (BPANN) เพื่อคาดการณ์ต้นทุนของโครงการก่อสร้างในไนจีเรียได้ อย่างไรก็ตาม รูปแบบนี้สามารถใช้ได้เฉพาะในอาคารประเภทสถาบันเท่านั้น ไม่มีอาคารประเภทอื่น หรือโครงการอื่นใดไม่สามารถประมาณได้โดยวิธีนี้ นอกจากนี้ เกณฑ์สำหรับประสิทธิภาพของรูปแบบพบข้อผิดพลาดในการคาดคะเนและไม่ได้คำนึงถึงวิธีการประเมินอื่นๆ นอกจากนี้ El-Sawalhi and Shehatto (2014) ได้ทำการสำรวจเพื่อตรวจสอบปัจจัยที่มีอิทธิพลมากที่สุดในกระบวนการประมาณราคา จากนั้นจึงพัฒนารูปแบบ โครงข่ายประสาทเทียม และดำเนินการวิเคราะห์ความอ่อนไหว พวกเขาได้ผลลัพธ์ที่น่าทึ่งด้วยโครงข่ายประสาท MLP ในขณะที่นำไปใช้ในช่วงเริ่มต้นของโครงการ นอกจากนี้ Roxas and Ongpeng (2014) ได้ใช้โครงข่ายประสาทเทียม เพื่อคาดการณ์ต้นทุนของโครงการก่อสร้างในฟิลิปปินส์ พวกเขาได้ข้อสรุปว่าบ่อยครั้ง โครงข่ายประสาทเทียม สามารถแสดงประสิทธิภาพที่ยอมรับได้แม้ว่าชุดข้อมูลที่มีอยู่จะไม่สมบูรณ์ อย่างไรก็ตาม ขอแนะนำให้ใช้ชุดข้อมูลที่สมบูรณ์ นอกจากนี้ Feng and Li (2013) ได้ใช้รูปแบบไฮบริดของโครงข่ายประสาทเทียม และ อัลกอริทึมขั้นตอนวิธีเชิง

พันธุกรรม เพื่อเอาชนะข้อเสียบางประการของโครงข่ายประสาทเทียม รวมถึงการบรรจบกันที่ช้าและติดอยู่กับค่าต่ำสุด นอกจากนี้ Kim and Shim (2014) ได้ใช้วิธีการไฮบริดของ CBR และ อัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรม ในระยะแรกๆ ของโครงการอาคารสูงเพื่อประเมินราคาด้วยวิธีที่ผิดพลาดน้อยกว่า การประยุกต์ใช้ อัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรม ได้ปรับปรุงกระบวนการของรูปแบบการประมาณได้สำเร็จโดยการกำหนดน้ำหนักของปัจจัยต้นทุน อย่างไรก็ตาม ขอแนะนำให้รวมหมวดหมู่ต้นทุนอื่นๆ สำหรับอาคารประเภทนี้ เช่น ค่าธรรมเนียมวิศวกรรมและเหตุการณ์ฉุกเฉิน โดยพิจารณาจากดัชนีสำหรับสถานที่ต่างๆ นำไปใช้ อัลกอริทึมแทน อัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรม เพื่อปรับปรุงน้ำหนัก การใช้รูปแบบกับชุดข้อมูลโครงการที่ใหญ่ขึ้น และการพิจารณาปัจจัยต้นทุนอื่นๆ ที่มีผลต่อกระบวนการประมาณราคา

Jeuschnigger (2018) ได้ศึกษา การคาดการณ์ราคาที่อยู่อาศัยในประเทศออสเตรเลียโดยเทคนิคการเรียนรู้ของเครื่อง วิทยานิพนธ์นี้ได้ทบทวนวรรณกรรมเกี่ยวกับที่อยู่อาศัยเพื่อศึกษาความสัมพันธ์ระหว่างราคาบ้านกับภาวะเศรษฐกิจ โดยเครื่องมือที่ใช้เป็นอัลกอริทึมการเรียนรู้ด้วยเครื่องเพื่อเพื่อคาดการณ์ราคาบ้านแต่ละแห่งในลอนดอนโดยใช้ข้อมูลที่เปิดเผยต่อสาธารณะเท่านั้น และพิจารณาว่าตัวแปรที่ระบุโดยวรรณกรรมสนับสนุนความถูกต้องในการคาดการณ์ของ อัลกอริทึม ดังนั้นจึงใช้การเรียนรู้ของเครื่องโดยใช้อัลกอริทึมการถดถอยที่มีการควบคุม 6 ชุดในชุดข้อมูลที่ถือราคาธุรกรรมที่อยู่อาศัยในลอนดอนตั้งแต่ปี 2546 ถึง พ.ศ. 2555 ผลการวิจัยพบว่าเทคนิคการเรียนรู้ของเครื่อง เช่น least square regression, elastic net, k-nearest neighbor, Bayesian regularization for feed-forward neural networks, random forests และ extreme gradient boosting. โดยรวมแล้ว พบว่า รูปแบบที่ไม่เป็นเชิงเส้นมีประสิทธิภาพดีกว่าตัวแบบเชิงเส้นตรง ในบรรดาอัลกอริทึมทั้งหมด เทคนิค Extreme gradient boosting (XGBoost หรือ XGB) จะให้ประสิทธิภาพที่ดีที่สุด ผลลัพธ์เพิ่มเติมพบว่าว่าตัวแปรที่เกี่ยวข้องกับที่ตั้งของที่อยู่อาศัยมีความสำคัญมากสำหรับรูปแบบที่สร้างขึ้น ในทางตรงกันข้าม ผลกระทบทางเศรษฐกิจมหภาคที่มีอยู่ในระดับดัชนีราคาที่อยู่อาศัยไม่สามารถตรวจสอบได้ในระดับราคาบ้านแต่ละแห่ง เนื่องจากข้อมูลและความสามารถในการคำนวณมีข้อจำกัดหลายประการ วิทยานิพนธ์นี้จึงไม่สามารถให้หลักฐานเพียงพอว่ารูปแบบการเรียนรู้ของเครื่องสามารถให้การคาดการณ์ที่เชื่อถือได้สำหรับราคาบ้านแต่ละหลังด้วยชุดข้อมูลที่กำหนด อย่างไรก็ตาม มันให้ภาพรวมโดยละเอียดของกระบวนการเรียนรู้ด้วยเครื่องและวิธีการที่อัลกอริทึมที่สำคัญสามารถนำมาใช้กับชุดข้อมูลที่อยู่อาศัยที่มีคุณลักษณะมากขึ้นหรือปัญหาทางเศรษฐกิจอื่นๆ ได้สำเร็จ

Jafarzadeh et al. (2014) ได้ศึกษาสถาปัตยกรรม โครงข่ายประสาทเทียม ที่เป็นไปได้ต่างๆ อย่างครอบคลุมด้วยอัตราการเรียนรู้ที่แตกต่างกัน และในที่สุดก็ได้เปรียบเทียบสถาปัตยกรรมเหล่านี้ และสรุปได้ว่าสิ่งที่ดีที่สุดคือโครงข่ายประสาท MLP ที่มีสองชั้นที่ซ่อนอยู่ ได้มาถึงข้อค้นพบที่

สำคัญในการวิจัยแล้วว่า จำนวนเซลล์ประสาทในเลเยอร์ที่ซ่อนอยู่ และพารามิเตอร์การเรียนรู้มีผลกระทบต่อความสามารถทั่วไปของเครือข่ายมากกว่าความสามารถด้านความแม่นยำ นอกจากนี้ จำนวนเซลล์ประสาทที่ซ่อนอยู่ยังมีประสิทธิภาพมากกว่าพารามิเตอร์การเรียนรู้ ในทางกลับกัน เครือข่ายมีความไวสูงต่อจำนวนอินพุต ดังนั้นยิ่งอินพุตมีมากขึ้น ยิ่งมีโอกาสเรียนรู้มากในเครือข่าย สุดท้ายการศึกษาแนะนำให้นำรูปแบบไปใช้ในอาคารประเภทอื่นๆ และเปรียบเทียบผลลัพธ์ปัจจุบันกับวิธีการประมาณต้นทุนอื่นๆ เช่น การถดถอยเชิงเส้นพหุคูณ นอกจากนี้ การพัฒนา โครงข่ายประสาทเทียม และ Support Vector Machine (SVM) สำหรับการคาดคะเนต้นทุนของโครงการก่อสร้างและกำหนดการถูกนำเสนอใน Wang et al. (2012) ซึ่ง SVM ได้แสดงให้เห็นประสิทธิภาพที่เหนือกว่า ดังนั้น โครงข่ายประสาทเทียม จึงเหมาะสมกว่าในข้อมูลตัวอย่างที่ไม่เป็นเชิงเส้น บทความนี้ยังแนะนำให้ใช้ชุด โครงข่ายประสาทเทียม และ SVM ในขณะที่ควรคำนึงถึงว่าการวางแผนล่วงหน้าถือเป็นปัจจัยสำคัญในความสำเร็จของโครงการ นอกจากนี้ El-Sawalhi and Shehatto (2014) ได้ดำเนินการสำรวจและดำเนินการวิเคราะห์ข้อมูลเพื่อแยกพารามิเตอร์อินพุตที่มีอิทธิพลหลักของรูปแบบคลุมเครือ พวกเขากล่าวว่าการใช้ฟังก์ชันสมาชิกภาพแบบสองด้านให้ผลลัพธ์ที่ดีกว่ารูปแบบอื่นๆ ที่ศึกษา พวกเขายังแนะนำว่าการเปรียบเทียบผลลัพธ์กับวิธีการเดี่ยวหรือวิธีผสมอื่น ๆ ก็มีประโยชน์เช่นกัน นอกจากนี้ Jin et al. (2012) ยังได้ใช้ประโยชน์จากความสามารถในการวิเคราะห์การถดถอยพหุคูณ (MRA) เพื่อแก้ไข CBR เพื่อเพิ่มความแม่นยำในการทำนาย พวกเขาแนะนำให้พิจารณาตัวแปรที่ระบุและตรวจสอบที่มาของอัตราความผิดพลาดที่เพิ่มขึ้น Son et al. (2012) ยังได้ใช้รูปแบบไฮบริดของการวิเคราะห์องค์ประกอบหลักและ Support Vector Regression (SVR) และเปรียบเทียบกับ SVR, โครงข่ายประสาทเทียม, Decision Tree และ Multiple Linear Regression (MLR) ซึ่งท้ายที่สุดแล้ว พวกเขาได้นำเสนอว่าอัลกอริทึม SVR มีประสิทธิภาพเหนือกว่า

Dimoski and Pettersen (2020) ได้ศึกษา การประมาณราคาก่อสร้างบ้านในประเทศนอร์เวย์ โดยได้ศึกษาการผลการใช้แมชชีนเลิร์นนิงในการประมาณราคาก่อสร้างบ้านพักอาศัยในประเทศนอร์เวย์ ผลการศึกษาพบว่า ในบรรดารูปแบบแมชชีนเลิร์นนิง ที่เลือกใช้ วิธีการ Elastic Net ให้ผลผลิตมากที่สุดในการคาดการณ์ราคารายไตรมาสที่แม่นยำ เมื่อเทียบกับ Norges Bank การคาดการณ์ของ Elastic Net นั้นมากกว่าแม่นยำถึง 29,6% ของไตรมาส แต่มีความแม่นยำน้อยกว่าในการประเมินโดยรวม การเบี่ยงเบนขนาดใหญ่ระหว่าง ปี 2561 และ 2562 ชี้ขาดผลประกอบการ หลังมีกฎเกณฑ์การจำนองฉบับใหม่แนะนำจาก Finanstilsynet Random Forest ทำนายการทำนายรายปีที่แม่นยำที่สุด แต่มีประสิทธิภาพดีกว่าธนาคาร Norges ถึงกระนั้น Random Forest ก็แซงหน้าทั้ง DNB และ SSB มาตลอดกระบวนการประเมินผล ผลที่ได้จาการศึกษาวิทยานิพนธ์นี้มีส่วนสนับสนุนงานวรรณกรรมที่มีอยู่หลายด้าน ประการแรก ทำได้ดีกว่าที่อยู่อาศัย ผู้เชี่ยวชาญ และแนวทางเศรษฐศาสตร์มหภาคแบบดั้งเดิมในการเลือกรูปแบบการคาดการณ์ ประการที่สอง ผลลัพธ์ของเรา

ระบุว่าตัวแบบเชิงเส้นเหมาะสมกว่าในช่วงเวลาที่สั้นกว่า ในขณะที่รูปแบบที่ไม่เป็นเชิงเส้นทำงานได้ดีกว่าในขอบเขตอันไกลโพ้น ประการที่สาม รูปแบบการเรียนรู้ของเครื่องมีระบุหนี้ครัวเรือนเป็นตัวแปรที่มีอิทธิพลมากที่สุดในการกำหนดราคาที่อยู่อาศัยในนอร์เวย์ โดยรวมแล้ว เราเชื่อว่าแมชชีนเลิร์นนิงอาจมีประโยชน์ในอนาคตการวิจัยทางเศรษฐศาสตร์มหภาคทางวิชาการและวิชาชีพ

Khalaf et al. (2020) ได้นำ Particle Swarm Optimization (PSO) มาใช้ในการประมาณราคาและระยะเวลาก่อสร้าง 60 โครงการในระยะแรก สิ่งที่อนุมานได้จากการศึกษานี้คือ PSO ทำงานได้ดีพร้อมผลลัพธ์ที่แม่นยำสูง ในขณะที่ต้องเจอกับพารามิเตอร์ที่มีความแปรปรวนหลากหลาย จุดแข็งของรูปแบบนี้คือขึ้นอยู่กับโครงการที่มีอยู่และมีความน่าเชื่อถือมากกว่าโครงการที่พิจารณาประมาณการและทดลอง อย่างไรก็ตามก็มีพยายามตรวจสอบรูปแบบด้วยพารามิเตอร์ที่กว้างขึ้นและนำไปใช้กับอาคารสีเขียวด้วย ในทางกลับกัน Jiang (2020) ได้ศึกษาการใช้ Artificial Neural Networks (ANN) ในการประมาณราคาโครงการก่อสร้าง และเปรียบเทียบผลลัพธ์กับวิธี Radial Basis Function Neural Network (RBFNN) และพบว่า โครงข่ายประสาทเทียม มีประสิทธิภาพดีกว่า จากนั้นจึงทำการศึกษาตามด้วยการเพิ่มประสิทธิภาพความถูกต้องของรูปแบบ และนำไปใช้กับโครงการประเภทอื่น และใช้วิธีอื่นในการพิจารณาปัจจัยด้านต้นทุน นอกจากนี้ Ji et al. (2019) ได้เสนอรูปแบบต้นทุนซึ่งเป็นรูปแบบตามปริมาณ ซึ่งผลลัพธ์จะถูกคูณด้วยราคาที่ต้องการในที่สุด แม้ว่ารูปแบบที่แนะนำจะมีประสิทธิภาพดีกว่าวิธี Case-Base Reasoning (CBR) เมื่อเปรียบเทียบกับรูปแบบนั้น แต่ก็มีความจำเป็นต้องทำการวิจัยเพิ่มเติมเพื่อเปรียบเทียบผลลัพธ์กับวิธีพารามิเตอร์เพิ่มเติมเพื่อตรวจสอบความน่าเชื่อถือของรูปแบบปัจจุบัน การศึกษานี้ยังใช้ประโยชน์จากกระบวนการสร้างพารามิเตอร์ ซึ่งบทบาทของกระบวนการนี้คือการสรุปปัจจัยต้นทุนที่มีประสิทธิภาพจำนวนมากลงในชุดของพารามิเตอร์ที่มีอิทธิพล ในอีกทางหนึ่ง Chandanshive and Kambekar (2019) ได้ตรวจสอบความสามารถของรูปแบบโครงข่ายประสาทเทียมแบบฟีดฟอร์เวิร์ดหลายชั้นด้วยอัลกอริทึมการเรียนรู้แบบแพร่กระจายย้อนกลับ (Backpropagation) เพื่อประเมินต้นทุนของโครงการก่อสร้าง 78 โครงการในอินเดีย ควบคู่ไปกับการทดสอบประสิทธิภาพของวิธีการหยุดก่อนกำหนดหรือแบบเบส์ ประสิทธิภาพของเครือข่ายพบปัญหาของรูปแบบที่มีความพอดีกับข้อมูลที่ใช้เรียนรู้ดีเกินไป นอกจากนี้ Wang (2017) ได้ใช้ตรรกศาสตร์คลุมเครือ (Fuzzy logic) เพื่อคาดการณ์ต้นทุนของโครงการก่อสร้าง เนื่องจากรูปแบบไม่ไดนามิกตามราคาตลาด จึงรู้สึกถึงความต้องการรูปแบบที่คล่องตัวมากขึ้น นอกจากนี้ Du and Li (2017) ได้ใช้การรวมเครือข่ายประสาทเทียม BP และอัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรมเพื่อประเมินต้นทุนของอาคารที่อยู่อาศัย หน้าที่ของอัลกอริทึมขั้นตอนวิธีเชิงพันธุกรรมคือการปรับปรุงประสิทธิภาพของ โครงข่ายประสาทเทียม โดยป้องกันไม่ให้ตกลงไปที่จุดสูงสุดในพื้นที่และเพิ่มความเร็วการบรรจบกัน นอกจากนี้ Amin (2017) ยังใช้ประโยชน์จากการวิเคราะห์การถดถอยพหุคูณเพื่อประเมินต้นทุนของอาคารที่อยู่อาศัย ในมุมมองของการวิจัย

พบว่า 92% ของต้นทุนอาคารที่พักอาศัยได้รับผลกระทบจากพื้นที่ที่ดินและพื้นที่อาคาร และอีก 8% ที่เหลือเกิดจากปัจจัยอื่นๆ

Kalliola et al. (2021) นำเสนอวิธีการและผลลัพธ์จากการปรับปรุงโมเดล MLP เพื่อการทำนายราคาอสังหาริมทรัพย์ในเฮลซิงกิ ฟินแลนด์ การปรับปรุงพารามิเตอร์ของโมเดลได้เสริมประสิทธิภาพได้เป็นอย่างดี (ค่า R2 ปรับปรุงขึ้น 0.05 และค่า RME ปรับปรุงขึ้น 2.5%) และสามารถพิจารณาได้ว่าเป็นขั้นตอนสำคัญในการพัฒนาแอปพลิเคชันทำนายอสังหาริมทรัพย์ อย่างไรก็ตาม วิธีการ ANN มีข้อด้อยและต้องถูกวิจารณะว่าเป็นที่วิจารณ์ ANN มีข้อบกพร่องในการอธิบาย เนื่องจากความสัมพันธ์ระหว่างอินพุตและเอาท์พุตไม่สามารถมองเห็นและอธิบายได้โดยตรง นอกจากนี้ มนุษย์ไม่สามารถแทรกแซงโดยตรงในความสัมพันธ์เหล่านั้นได้ อย่างไรก็ตาม การผลิตการประเมินราคาทรัพย์สินที่ยั่งยืนโดยไม่มีการแทรกแซงจากมนุษย์สามารถถือว่าเป็นประโยชน์ ANN สามารถสร้างผลลัพธ์ที่แม่นยำ ยืดหยุ่น และทั่วไปมากขึ้น หากมีข้อมูลเพียงพอเมื่อเทียบกับวิธีการอื่นๆ ดังนั้น สามารถพิจารณาได้ว่าเป็นทางออกที่ดีสำหรับปัญหาการทำนายราคา การวิเคราะห์ผลลัพธ์แสดงให้เห็นว่ากระบวนการปรับปรุงโมเดลได้เพิ่มประสิทธิภาพอย่างมีนัยสำคัญในแต่ละตัวชีวัด ผลลัพธ์การฝึกอบรมแสดงให้เห็นว่าไม่มีการทำลายหรือการฝึกเกินและการวิเคราะห์ความไวต่อการเปลี่ยนแปลงอธิบายถึงประสิทธิภาพที่ดีบนชุดทดสอบ การวิเคราะห์แสดงให้เห็นว่าผลลัพธ์สามารถปรับปรุงได้โดยการให้ความสำคัญกับการปรับปรุงโมเดลและการปรับพารามิเตอร์ การวิจัยได้แสดงให้เห็นว่าสามารถทำนายราคาอสังหาริมทรัพย์ในเฮลซิงกิ ฟินแลนด์ โดยใช้เทคโนโลยีเครือข่ายประสาทเชิงลึกและการเรียนรู้เชิงลึก และเทคโนโลยีนี้สามารถนำไปใช้ในงานการพยากรณ์ความสัมพันธ์ที่ไม่เชิงเส้นระหว่างอินพุตและเอาท์พุตได้

Elfahham (2019) พบว่าโครงการก่อสร้างต้องการลงทุนที่สูงและอาจใช้เวลาหลายปีในการดำเนินการเสร็จสิ้น ซึ่งทำให้มันอยู่ในกระบวนการเพิ่มราคาได้ ดัชนีต้นทุนการก่อสร้าง (CCI) เป็นดัชนีที่ให้ความเคลื่อนไหวของต้นทุนการก่อสร้าง สิ่งนี้มีประโยชน์ในการประเมิน การกำหนดราคาและการเสนอราคาโครงการก่อสร้างโดยเฉพาะอย่างยิ่งเมื่อมีอัตราเงินเฟ้อสูงในอียิปต์ ในงานวิจัยนี้ CCI ถูกคำนวณขึ้นบนพื้นฐานของดัชนีข่าวิศวกรรมในอเมริกา แต่มีการเปลี่ยนแปลงในรายการต้นทุนที่ใช้ในสูตรการคำนวณเพื่อให้เหมาะกับวัสดุก่อสร้างที่ใช้ในประเทศอียิปต์ โดยเฉพาะ; การศึกษาถูกจำกัดไว้ที่โครงสร้างคอนกรีตเท่านั้น ข้อมูลนั้นถูกสะสมมาจากหน่วยงานสถิติกลางสำหรับการเรียกรายและสถิติ ซึ่งเป็นหน่วยงานสถิติชั้นนำ การวิเคราะห์ข้อมูลที่สะสมได้ระบุว่ารายการต้นทุนหลักของโครงสร้างคอนกรีตคือ เหล็ก ปูนซีเมนต์ ทรายและกรวด และกระเบื้อง น้ำหนักของวัสดุเหล่านี้ในสูตรถูกพบว่าเป็น 54% 30% 10% และ 6% ตามลำดับ ดัชนี CCI ถูกคำนวณสำหรับช่วงเวลาการศึกษาตั้งแต่ปี 2002 ถึงปี 2018 โดยพิจารณาว่าปี 2010 เป็นปีฐาน วิธีการเครือข่ายประสาทเชิงลึก (Neural Network) การถดถอยเชิงเส้น (Linear Regression) และ

วิธีการลำดับเวลาถดถอย (Autoregression Time Series) ถูกใช้ในการทำนายดัชนีต้นทุนการก่อสร้างจนถึงปี 2025 ความสามารถในการทำนายของวิธีการทำนายทั้งสามวิธีจึงถูกประเมินโดยใช้การคำนวณค่าความผิดพลาดที่เฉลี่ยและการคำนวณค่าความผิดพลาดเฉลี่ยสัมประสิทธิ์ พบว่าการทำนายโดยใช้วิธีการลำดับเวลาถดถอยเป็นวิธีที่แม่นยำที่สุด เนื่องจากมีค่าความผิดพลาดเฉลี่ยของค่าสัมประสิทธิ์เฉลี่ยอยู่ที่ 3.5 ในขณะที่วิธีการถดถอยเชิงเส้นมีประสิทธิภาพน้อยลงมีค่าความผิดพลาดเฉลี่ยอยู่ที่ 17.5 นอกจากนี้ วิธีการถดถอยเชิงเส้นให้การทำนายเชิงเส้นซึ่งไม่เป็นปฏิบัติ

การประเมินต้นทุนของโครงการก่อสร้างอาคารในขั้นตอนแรกๆ ด้วยความแม่นยำที่สูงมีสำคัญในความสำเร็จของทุกๆ โครงการก่อสร้าง (Chandanshive & Kambekar, 2019) โดยขึ้นอยู่กับการสำรวจและคำแนะนำจากผู้ทำออกแบบและผู้รับเหมาก่อสร้าง ข้อมูลเซตของ 78 โครงการก่อสร้างอาคารได้รับจากเมืองหลวงขนาดใหญ่มุมไบ (อินเดีย) และภูมิภาคใกล้เคียงทางภูมิศาสตร์ พารามิเตอร์การออกแบบที่มีอิทธิพลมากที่สุดของต้นทุนโครงสร้างของอาคาร (หน่วยค่าเงินอินเดีย: INR) ได้ถูกระบุและกำหนดเป็นอินพุต และต้นทุนโครงสร้างโครงสร้างรวม (INR) แทนผลลัพธ์ของโมเดลเครือข่ายประสาทแบบผ่านสัญญาณหน้า (multilayer feed forward neural network) วัตถุประสงค์ของงานวิจัยนี้คือการพัฒนาโมเดลเครือข่ายประสาทแบบผ่านสัญญาณแบบหน้า (multilayer feed forward neural network) ที่ถูกฝึกด้วยอัลกอริทึม backpropagation สำหรับการทำนายต้นทุนการก่อสร้างอาคาร (INR) การใช้วิธี early stopping และ Bayesian regularization ได้ถูกนำมาใช้เพื่อเพิ่มประสิทธิภาพในการทำนายของเครือข่ายประสาท และเพื่อป้องกันการ overfitting ได้มีการสังเกตเห็นว่าในการทำนายต้นทุนการก่อสร้างการทำนายโดยวิธีการ regularization แบบ Bayesian มีระดับผลงานที่ดีกว่าการใช้วิธี early stopping ผลลัพธ์ที่ได้จากโมเดลเครือข่ายประสาทที่ถูกฝึกแสดงให้เห็นว่ามันสามารถทำนายต้นทุนของโครงการก่อสร้างอาคารในขั้นตอนแรกของการก่อสร้างได้ งานวิจัยนี้มีส่วนสำคัญในการจัดการการก่อสร้างและให้แนวคิดเกี่ยวกับงบประมาณการเงินทั้งหมดที่จะเป็นประโยชน์ต่อเจ้าของทรัพย์สินและนักลงทุนทางการเงินในการตัดสินใจและการจัดการการลงทุนของพวกเขาในวงการก่อสร้างที่เปลี่ยนแปลงได้

Pham et al. (2023) ได้ศึกษาการให้เฟรมเวิร์คการเรียนรู้ของเครื่องจักร (Machine Learning: ML) และการปรับปรุงที่รวดเร็วและแม่นยำ ที่ช่วยให้สามารถประเมินต้นทุนสำหรับการก่อสร้างได้เร็วขึ้น ซึ่งจะเสริมสร้างประสิทธิภาพในการดำเนินงานและความแข่งขันของบริษัทก่อสร้าง ชุดข้อมูลประกอบด้วยการกำหนดพารามิเตอร์ของอาคาร 10,000 รูปแบบที่ถูกสะสมมาจากกิจกรรมจริงตั้งแต่เริ่มจนจบในบริษัทคู่ค้าของเรา ถูกนำมาใช้ในการฝึกและตรวจสอบโมเดล ML เพื่อดำเนินการหลายงาน ในหมู่ 13 อัลกอริทึมสำหรับการ regression ของ ML ที่ถูกใช้โมเดล Artificial Neural Network (ANN), Gradient Boosting และ XGBoost ใช้ให้เหมาะที่สุดในการประมาณต้นทุนการก่อสร้างและทรัพยากรที่ต้องการ ด้วยความแม่นยำถึง 99% ในเวลาที่น้อยกว่าหนึ่งวินาที

ของเวลาการฝึก โมเดล ANN ยังถูกพัฒนาขึ้นเพื่อระบุตัวเลือกที่มีอยู่ของคุณสมบัติของอาคารภายใต้งบประมาณที่กำหนด ปัญหาการปรับปรุงภายใต้เงื่อนไขถูกแก้ไข ช่วยให้ลูกค้ากำหนดต้นทุนการก่อสร้างที่เหมาะสมตามความชอบของพวกเขา นอกจากนี้ ต้นทุนการก่อสร้างที่ได้รับการปรับปรุงจากงานวิจัยนี้ มีขนาดเล็กลง 7% จากข้อมูลจริง ซึ่งจะช่วยในการเพิ่มความแข่งขันของบริษัท งานวิจัยนี้เปิดโผการว่าโมเดล ML สามารถนำไปใช้อย่างมีประสิทธิภาพในวงการก่อสร้าง เพื่อปรับปรุงการทำงานเพื่อประหยัดต้นทุน และให้ข้อได้เปรียบที่เป็นประโยชน์ในการบริหารจัดการที่ขับเคลื่อนด้วยข้อมูล เช่นเดียวกันกับ El Mouna et al. (2023) พยายามแก้ปัญหานี้โดยการทำนายราคาบ้านโดยใช้การเรียนรู้เชิงเครื่องจักร โดยการเลือกใช้สามประเภทของการเรียนรู้เชิงเครื่องจักร: การถดถอยเชิงเส้น (Linear Regression: LN), การสร้างแบบสุ่ม (Random Forest: RF) และการบูสท์เกรด (Gradient Boosting: GB) เราทดสอบโมเดลของเราบนชุดข้อมูลอสังหาริมทรัพย์แห่งเมลเบิร์น ซึ่งประกอบด้วยการขายอสังหาริมทรัพย์ 34,857 รายการและ 21 ลักษณะ

Car-Pusic et al. (2020) นำเสนอการรวมโมเดลที่ใช้กระบวนการและข้อมูลเพื่อการทำนายต้นทุนในระยะแรกของโครงการก่อสร้าง โมเดล "time-cost" ของ Bromilow ถูกนำมาใช้เป็นโมเดลที่ใช้กระบวนการ และโมเดล neural network ที่ใช้การ regression ทั่วไป (GRNN) เป็นโมเดลที่ใช้ข้อมูล โมเดล GRNN ให้การทำนายที่แม่นยำที่สุดในหมู่ 3 โมเดลทำนายด้วย neural network ที่ถูกนำมาใช้ โดยมีค่าความคลาดเคลื่อนที่เฉลี่ย (MAPE) ประมาณ 0.73% และค่าสหสัมพันธ์ R2 ประมาณ 99.55% ค่าสหสัมพันธ์ระหว่างค่าที่ทำนายได้และค่าจริงคือ 0.998 โมเดลถูกออกแบบให้เป็นส่วนหนึ่งของระบบการทำนายต้นทุน (CPS) ซึ่งบทบาทคือการประมาณต้นทุนโครงการในระยะเริ่มต้น ผลลัพธ์ที่ได้รับถูกนำมาใช้เป็นอินพุตของ Cost Model (CM) ซึ่งเป็นส่วนหนึ่งของระบบสนับสนุนการตัดสินใจ (DSS) และส่วนหนึ่งของระบบข้อมูลการบริหารสิ่งสร้างอาคารที่กว้างขึ้น (BMIS) โมเดลสามารถเป็นประโยชน์สำหรับผู้ร่วมโครงการทุกคนในการทำนายต้นทุนในระยะเริ่มต้นของโครงการก่อสร้าง เป็นพิเศษในขั้นตอนของการเสนอราคาและทำสัญญาเมื่อมีปัจจัยหลายอย่างที่อาจจะกำหนดการดำเนินโครงการก่อสร้างที่ยังไม่ทราบอยู่

Srirutchataboon et al. (2021) ได้วิเคราะห์ข้อมูลราคาที่อยู่อาศัยที่ได้มาจากเว็บไซต์อสังหาริมทรัพย์ชั้นนำในประเทศไทยและ Open Street Maps (OSM) เพื่อระบุลักษณะที่มีผลต่อราคาที่อยู่อาศัยในประเทศไทยตั้งแต่ปี 2015 ถึง 2019 นอกจากนี้ เราขอเสนอโมเดลที่ใช้โครงสร้างการเรียนรู้จากผสาน (stacking ensemble learning) ซึ่งการทำนายได้โดยการรวมรายการเรียนรู้ฐานทั้ง 3 โมเดลปรับปรุงประสิทธิภาพประกอบด้วยโมเดลการเรียนรู้เชิงคอนโวลูชัน (convolutional neural network: CNN) โมเดลผสาน (เช่น random forests (RF), extreme gradient boosting (XGBoost) และ adaptive boosting (AdaBoost)) และเทคนิคการถดถอยเชิงเส้นเรียบง่ายโมเดล CNN ถูกใช้ในการสกัดลักษณะจากภาพบ้านซึ่งจากนั้นถูกผสานรวมกับลักษณะแบบดั้งเดิมเพื่อ

ประมาณราคาเริ่มต้น การทำนายจึงถูกปรับแต่งด้วยการถดถอยเชิงเส้น เมื่อเทียบกับโมเดลแต่ละตัว โมเดลที่เสนอได้รับ Mean Absolute Percentage Error (MAPE) ที่ 17.83% มีประสิทธิภาพสูงกว่ารายการอ้างอิงอื่น ๆ อย่างมีนัยสำคัญ เทียบด้วยโมเดลอื่น

Soltani et al. (2022) พบว่าวิธีการทำนายราคาที่อยู่อาศัยแบบดั้งเดิมนั้นหากันพิจารณาปัญหาความไม่สถิตในมิติเชิงพื้นที่และเวลาในข้อมูลปริมาณมาก ในงานวิจัยนี้มีการใช้สี่โมเดลเรียนรู้ของเครื่อง (Machine Learning: ML) เพื่อสำรวจผลกระทบของลักษณะต่าง ๆ เช่น คุณสมบัติของทรัพย์สินและคุณภาพของย่านบ้านต่อการเปลี่ยนแปลงของราคาที่อยู่อาศัยในระดับภูมิภาคที่แตกต่างกัน โดยใช้ชุดข้อมูลราคาที่อยู่อาศัยใน Metropolitan Adelaide, Australia ในระยะเวลา 32 ปี (1984-2016) โดยที่งานวิจัยนี้ขึ้นอยู่กับบันทึกธุรกรรมการขาย 428,000 รายการและตัวแปรอธิบาย 38 ตัวแปร แสดงให้เห็นว่าโมเดลที่ใช้การเรียนรู้เชิงต้นไม้ที่เชิงเลือก (Decision Tree) ที่ไม่เป็นเชิงเส้นมีประสิทธิภาพดีกว่าโมเดลที่เป็นเชิงเส้น นอกจากนี้ เทคนิคการเรียนรู้ของเครื่องแบบผสาน (ensemble machine learning) เช่น Gradient-Boosting และ Random Forest มีความสามารถที่ดีกว่าในการทำนายราคาที่อยู่อาศัยในอนาคต มีการเพิ่มตัวแปรของละความที่เชิงพื้นที่-เวลา (Spatiotemporal lag: ST-lag) เพื่อปรับปรุงความแม่นยำในการทำนายของโมเดล งานวิจัยนี้แสดงให้เห็นว่าตัวแปร ST-lag (หรือตัวบ่งชี้ที่เป็นประโยชน์ในมิติเชิงพื้นที่-เวลา) สามารถทำหน้าที่เป็นตัวระบมีความสำคัญในผลกระทบของมิติเชิงพื้นที่-เวลาในการประยุกต์ใช้ของ ML บทความนี้จะเป็นตัวกระตุ้นสำหรับงานวิจัยในอนาคตเกี่ยวกับแลมของตลาดอสังหาริมทรัพย์ในออสเตรเลีย โดยการใช้ประโยชน์จากเทคโนโลยีล่าสุดเพื่อพัฒนาโมเดลสำหรับธุรกิจและการประเมินทรัพย์สินที่ระดับภูมิภาคต่าง ๆ

Ahtesham et al. (2020) พบว่าราคาที่อยู่อาศัยเป็นภาพลักษณ์ที่สำคัญของเศรษฐกิจและช่วงค่าคุณค่าของมันเป็นปัญหาที่สำคัญสำหรับลูกค้าและพ่อค้าที่ดิน ราคาที่อยู่อาศัยเพิ่มขึ้นทุกปี ซึ่งในที่สุดนั้นเสริมทำให้จำเป็นต้องมีกลยุทธ์หรือเทคนิคที่สามารถทำนายราคาที่อยู่อาศัยในอนาคต มีปัจจัยบางประการที่มีอิทธิพลต่อราคาที่อยู่อาศัยรวมถึงเงื่อนไขทางกายภาพ สถานที่ จำนวนห้องนอนและอื่น ๆ การทำนายในทางเดิมทำโดยใช้ข้อมูลพื้นฐานเหล่านี้ อย่างไรก็ตามวิธีการทำนายทั้งหมดเหล่านี้ต้องการความรู้และประสบการณ์ที่เหมาะสมเกี่ยวกับเชิงดินแหลมนี้ เทคนิคของการเรียนรู้ของเครื่อง (Machine Learning) ได้เป็นแหล่งที่สำคัญของโอกาสที่ล้ำหน้าในการวิเคราะห์ ทำนาย และแสดงราคาที่อยู่อาศัย ในบทความนี้ โมเดลการทำนายแบบ Gradient Boosting XGBoost ถูกใช้เพื่อทำนายราคาที่อยู่อาศัย ข้อมูลที่มีอยู่แบบสาธารณะซึ่งมีบันทึก 38,961 รายการของเมืองการาซี ได้มาจากพอร์ทัลอสังหาริมทรัพย์เปิดให้บริการของปากีสถาน มีงานวิจัยมากมายที่ได้ทำการทำนายราคาที่อยู่อาศัยในหลายประเทศ อย่างไรก็ตาม มีจำนวนจำกัดของงานวิจัยที่ได้ทำการทำนายราคาที่อยู่อาศัยในปากีสถาน โมเดลการทำนายราคาที่อยู่อาศัยที่เราเสนอสามารถทำนายได้ด้วยความแม่นยำ 98%

การเกิดความผันผวนของราคาที่อยู่อาศัยในตลาดอสังหาริมทรัพย์เกิดขึ้นเนื่องจากผลของตัวแปรอื่นที่สัมพันธ์กับราคาที่อยู่อาศัย บางราคาไม่สามารถควบคุมได้หรือไม่รู้เมื่อจะเพิ่มขึ้นหรือลดลง ความผันผวนเหล่านี้สอดคล้องกับทฤษฎีในการตั้งราคาเชิงเฮโดนิก ซึ่งแสดงให้เห็นว่าการตั้งราคาของบ้านขึ้นอยู่กับลักษณะรวมรวม Sakri and Ali (2022) ได้ทำการวิจัยเชิงสำรวจแนวความคิดทางเศรษฐมวิทัศน์นี้โดยใช้อัลกอริทึมการเรียนรู้ของเครื่องในการพยากรณ์ราคาที่อยู่อาศัยอย่างแม่นยำ ในการสร้างโมเดลที่พยากรณ์ เราใช้ชุดข้อมูลที่เกี่ยวข้องกับที่อยู่อาศัยใน Ames ซึ่งรวมถึงคุณลักษณะอธิบาย 82 ลักษณะ และ 2930 รายการขายที่อยู่อาศัยใน Ames, Iowa, USA เราใช้การเลือกคุณสมบัติด้วยการตรวจหาค่าที่ดีที่สุดด้วยการเลือกคุณสมบัติแบบแผ่นได้ (particle swarm optimization) เพื่อแก้ไขปัญหามิติของชุดข้อมูลทั้งสอง สำหรับโมเดลที่ใช้การจัดลักษณะฐานสำหรับการแบ่งประเภทและโมเดลที่ใช้การถดถอยที่เน้นในการพยากรณ์ กับการลดมิติ เราพบว่าผลการทดลองแสดงให้เห็นว่าความแม่นยำในการทำนายราคาที่อยู่อาศัยเพิ่มขึ้นจาก 81.4% ไปเป็น 84.4% สำหรับโมเดลที่ใช้การแบ่งประเภท และความผิดพลาดในการทำนายราคาลดลงจาก 13.3% ไปเป็น 6.9% สำหรับโมเดลที่ใช้การถดถอยที่เน้นในการพยากรณ์ ผลการทดลองแสดงให้เห็นว่ามีความแตกต่างมากแค่ไหนระหว่างการใช้และไม่ใช้การเลือกคุณสมบัติในเชิงการทำนายราคาที่อยู่อาศัย ประโยชน์หลักของการนำมาใช้คุณสมบัติ คือ มันช่วยในการเลือกคุณลักษณะที่สัมพันธ์กับมากขึ้น เพื่อป้องกันการกระชับอย่างเกินไปของโมเดลด้วยคุณลักษณะที่ไม่เกี่ยวข้องมากมายเกินไป

พหูน ปณฺฑิโต ชีเว

บทที่ 3

วิธีดำเนินการวิจัย

งานวิจัยเรื่องการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม (The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine) การทำนายราคาอย่างแม่นยำเป็นสิ่งสำคัญในวงกว้างของอุตสาหกรรมอสังหาริมทรัพย์ งานวิจัยนี้นำเสนอวิธีการที่ครอบคลุมอย่างละเอียดสำหรับการทำนายราคาการก่อสร้างบ้านพักอาศัยแนวราบ ดังนั้นผู้วิจัยได้การดำเนินการศึกษาตามระเบียบวิธีวิจัย โดยการวิจัยแบ่งออกเป็นทั้งหมด 5 ระยะ ซึ่งประกอบด้วย ระยะที่ 1 เตรียมข้อมูลและจัดกระทำข้อมูล ระยะที่ 2 สร้างแบบจำลองการทำนายพื้นฐาน (base model) ระยะที่ 3 สร้างแบบจำลองเอ็นเซมเบิล (ensemble model) ระยะที่ 4 ประเมินประสิทธิภาพแบบจำลอง และระยะสุดท้ายระยะที่ 5 นำเสนอผลลัพธ์และสรุปผลงานวิจัย ดังภาพประกอบ 20

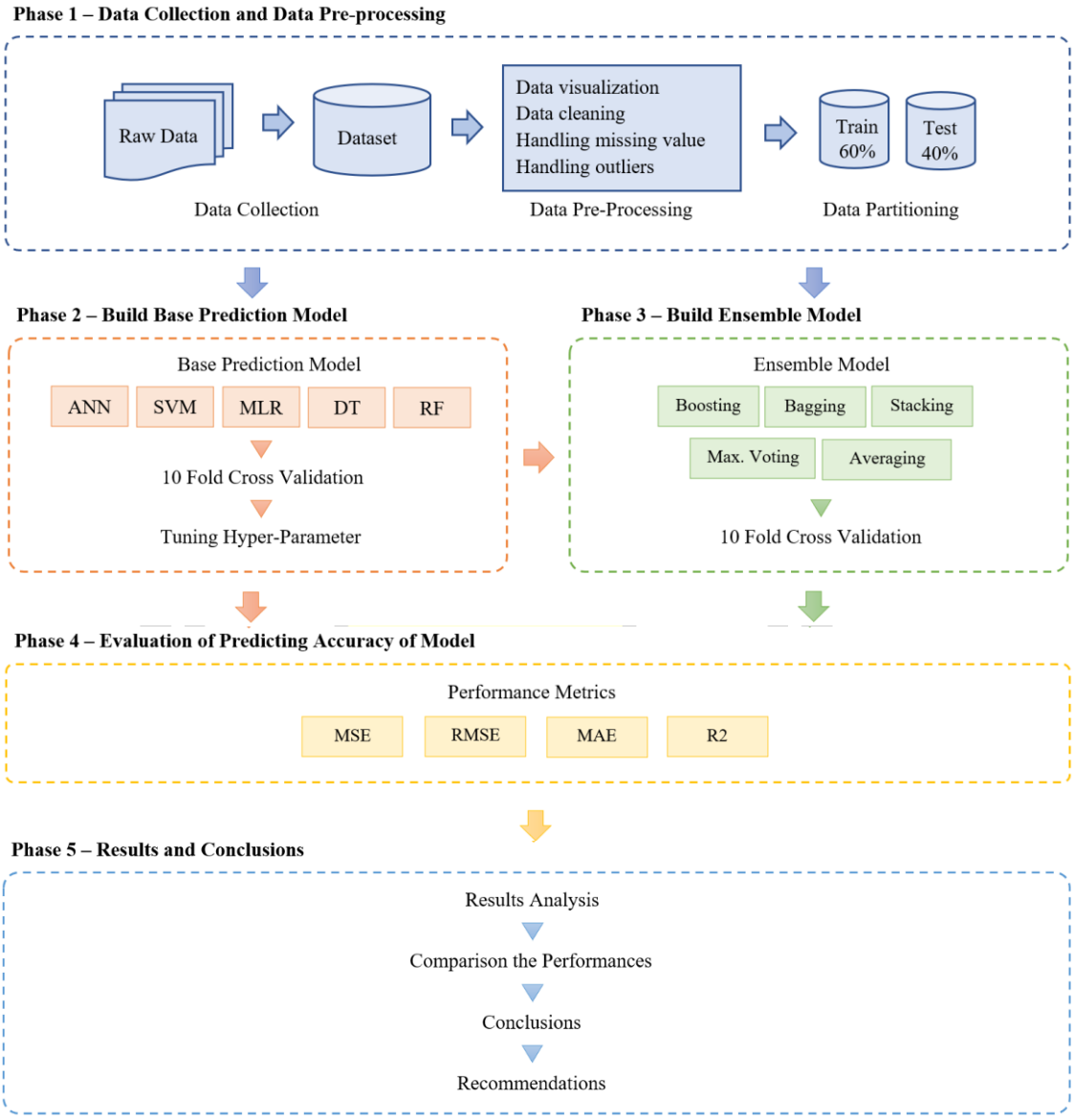
ระยะที่ 1 - การรวบรวมข้อมูลและการจัดเตรียมข้อมูลเป็นขั้นเริ่มต้น ในขั้นนี้มีรวบรวมข้อมูลที่ได้จากหน่วยงานประมาณราคา สำนักโยธา จังหวัดกรุงเทพมหานคร จากนั้นจัดเตรียมข้อมูลที่ครอบคลุม ซึ่งรวมถึงงานที่เกี่ยวกับการทำความสะอาดข้อมูล, การจัดการค่าที่ขาดหาย, การแก้ปัญหาข้อมูลผิดปกติ และการแบ่งข้อมูล

ระยะที่ 2 - การสร้างแบบจำลองการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ โดยเริ่มจากการสร้างแบบจำลองการทำนายพื้นฐาน (base model) โดยใช้แบบจำลอง Regression-Based Models ซึ่งประกอบด้วย Artificial Neural Networks (ANN), Support Vector Machines (SVM), Multiple Linear Regression (MLR), Decision Trees (DT), และ Random Forest (RF) ในขั้นนี้มีการใช้วิธีการ 10-fold cross-validation ร่วมกับการปรับพารามิเตอร์เพื่อเพิ่มประสิทธิภาพของแบบจำลองแต่ละแบบจำลองพื้นฐานด้วย

ระยะที่ 3 – การสร้างแบบจำลองเรียนรู้กลุ่มหรือเอ็นเซมเบิล (ensemble model) โดยใช้แบบจำลองพื้นฐาน (base model) เพื่อสร้างแบบจำลองการเรียนรู้กลุ่มผ่านเทคนิคต่าง ๆ เช่น Maximum Voting, Averaging, Stacking, Boosting, และ Bagging โดยยังคงใช้วิธีการ 10-fold cross-validation

ระยะที่ 4 - ประเมินความแม่นยำของแบบจำลองการทำนาย โดยทำการประเมินผลประสิทธิภาพของทั้งแบบจำลองพื้นฐานและแบบจำลองเรียนรู้กลุ่ม โดยใช้หลายตัวชี้วัดประสิทธิภาพดังนี้ Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), และ R-squared (R2)

ระยะที่ 5 - นำเสนอผลลัพธ์และสรุปผลงานวิจัย โดยรวมผลลัพธ์, การเปรียบเทียบแบบจำลอง, สรุปผล และการแนะนำเพื่อเลือกแบบจำลองที่เหมาะสมที่สุดสำหรับการทำนายต้นทุนการก่อสร้างของบ้านพักอาศัยแนวราบ การใช้ระเบียบวิธีการวิจัยนี้ช่วยในการวิเคราะห์อย่างเป็นระบบตลอดทุกขั้นตอนและทุกระยะของการวิจัย



ภาพประกอบ 20 กรอบแนวคิดวิจัย

ชุดข้อมูล

ชุดข้อมูลที่ใช้ในการดำเนินการสร้างโมเดลที่นำเสนอในงานวิจัยนี้ได้มาจากหน่วยงานประมาณราคา สำนักโยธา กรุงเทพมหานคร จำนวน 120 รายการ ซึ่งประกอบด้วยข้อมูลแบบบ้านที่เกี่ยวข้องกับบ้านพักอาศัยแนวราบและราคาที่เกี่ยวข้องกัน ชุดข้อมูลนี้ประกอบด้วยคุณลักษณะที่ถูกเลือกด้วยความรอบคอบ 15 ลักษณะ โดยเลือกจากการทบทวนอย่างละเอียดจากวรรณกรรมที่มีอยู่และความรู้ด้านความเชี่ยวชาญของผู้เชี่ยวชาญ ชุดข้อมูลและคุณลักษณะที่เกี่ยวข้องแสดงไว้ในตาราง 1 โดยคอลัมน์ด้านซ้ายระบุชื่อลักษณะ คอลัมน์กลางระบุคำอธิบายสรุปของแต่ละลักษณะ และคอลัมน์ด้านขวาระบุลักษณะที่ถูกเลือกจากการทบทวนวรรณกรรมที่อ้างอิง

การเลือกใช้ชุดข้อมูลจากหน่วยงานประมาณราคา สำนักโยธา กรุงเทพมหานครนี้ มีเหตุผลเนื่องมาจากข้อมูลสถิติที่ได้รับจากกองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม ปี 2564 พบว่ามีการขออนุญาตก่อสร้างทั้งหมด 315,496 แห่ง (รายละ 230,828) โดยจังหวัดที่มีการก่อสร้างมากที่สุดคือจังหวัดกรุงเทพมหานครจำนวน 25,346 แห่ง (รายละ 16,314) ซึ่งคิดเป็นร้อยละ 8.03 ของประเทศ หากพิจารณาเพียงกลุ่มที่เป็นบ้านพักอาศัยทั้งหมด ขนาดรวมของพื้นที่ทั้งหมดมีจำนวน 39,186,978 ตารางเมตร โดยบ้านพักอาศัยแนวราบเป็นส่วนที่ใหญ่สุดถึง 30,934,610 ตารางเมตร ซึ่งคิดเป็นร้อยละ 78.94 (กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ, 2564) ดังนั้นผู้วิจัยเลือกใช้กลุ่มตัวอย่างจากบ้านพักอาศัยแนวราบของจังหวัดกรุงเทพมหานครที่มีการก่อสร้างมากที่สุดและเป็นศูนย์กลางเศรษฐกิจของประเทศ และข้อมูลที่ใช้มาจากหน่วยงานกลุ่มงานประมาณราคา สำนักการโยธา กรุงเทพมหานคร เป็นข้อมูลที่มีความน่าเชื่อถือและเป็นมาตรฐาน

ตาราง 4 คุณลักษณะของตัวแปรและการอ้างอิงที่มาจากการทบทวนวรรณกรรม

คุณลักษณะ	รายละเอียด	อ้างอิง
Y	ราคาบ้านพักอาศัยแนวราบ	Sirithikon Sittikarnkul (2021), Ho et al. (2021), Park et al. (2022), Wit Dangsangthong (2016), Yang et al. (2021), Chandanshive and Kambekar (2019), Lim et al. (2016), Wang et al. (2021), (Soltani et al., 2022)
X1	จำนวนชั้น	Sirithikon Sittikarnkul (2021), Park et al. (2022), Ji et al. (2019), Roxas and Ongpeng (2014), Chandanshive and Kambekar (2019), Vineeth et al. (2018), Soltani et al. (2022)
X2	พื้นที่ใช้สอยทั้งหมด	Sirithikon Sittikarnkul (2021), Ho et al. (2021), Park et al. (2022), Ji et al. (2019), Roxas and Ongpeng (2014), Chandanshive and Kambekar (2019), Lim et al. (2016), Gao et al. (2022),

คุณลักษณะ	รายละเอียด	อ้างอิง
		Owusu-Manu et al. (2019), Montero et al. (2018), Wang et al. (2021)
X3	พื้นที่ห้องนอน	Wit Dangsangthong (2016), Sirithikon Sittikarnkul (2021), Lim et al. (2016), Jafari and Akhavian (2019), Choong (2018), Gao et al. (2022), Owusu-Manu et al. (2019), Vineeth et al. (2018), Wang et al. (2021), Soltani et al. (2022)
X4	พื้นที่ห้องน้ำ	Wit Dangsangthong (2016), Lim et al. (2016), Jafari and Akhavian (2019), Montero et al. (2018), Vineeth et al. (2018), Wang et al. (2021)
X5	พื้นที่ห้องนั่งเล่นและพักผ่อน	Wit Dangsangthong (2016), Vineeth et al. (2018), Wang et al. (2021)
X6	พื้นที่ห้องครัวและรับประทานอาหาร	Wit Dangsangthong (2016)
X7	พื้นที่ลานซักล้าง	Wit Dangsangthong (2016)
X8	พื้นที่เฉลียงหรือระเบียง	Wit Dangsangthong (2016)
X9	พื้นที่บันไดและโถงทางเดิน	Wit Dangsangthong (2016)
X10	พื้นที่หลังคาและกันสาดคอนกรีต	Ji et al. (2019), Sirithikon Sittikarnkul (2021), Soltani et al. (2022)
X11	พื้นที่หลังคากระเบื้อง	Sirithikon Sittikarnkul (2021), Soltani et al. (2022)
X12	พื้นที่โรงจอดรถ	Wit Dangsangthong (2016), Montero et al. (2018), Chandanshive and Kambekar (2019), Owusu-Manu et al. (2019), Park et al. (2022)
X13	ความสูงทั้งหมด	Ji et al. (2019), Yang et al. (2021), Wang et al. (2021), Sirithikon Sittikarnkul (2021), Park et al. (2022)
X14	ความสูงของหลังคา	ประสบการณ์วิชาชีพและผู้เชี่ยวชาญแนะนำ
X15	ความสูงเฉลี่ยแต่ละชั้น	Ji et al. (2019), Sirithikon Sittikarnkul (2021), Ho et al. (2021)

เครื่องมือที่ใช้ในงานวิจัย

ในการวิจัยเรื่องการทำนายการประมาณราคาการการก่อสร้างบ้างพักอาศัยแนวราบโดยใช้เครื่องจักเรียนรู้เอ็กซ์ตรีมได้มีการนำอุปกรณ์ฮาร์ดแวร์และซอฟแวร์มาใช้ในการวิจัย โดยเลือกใช้ภาษา

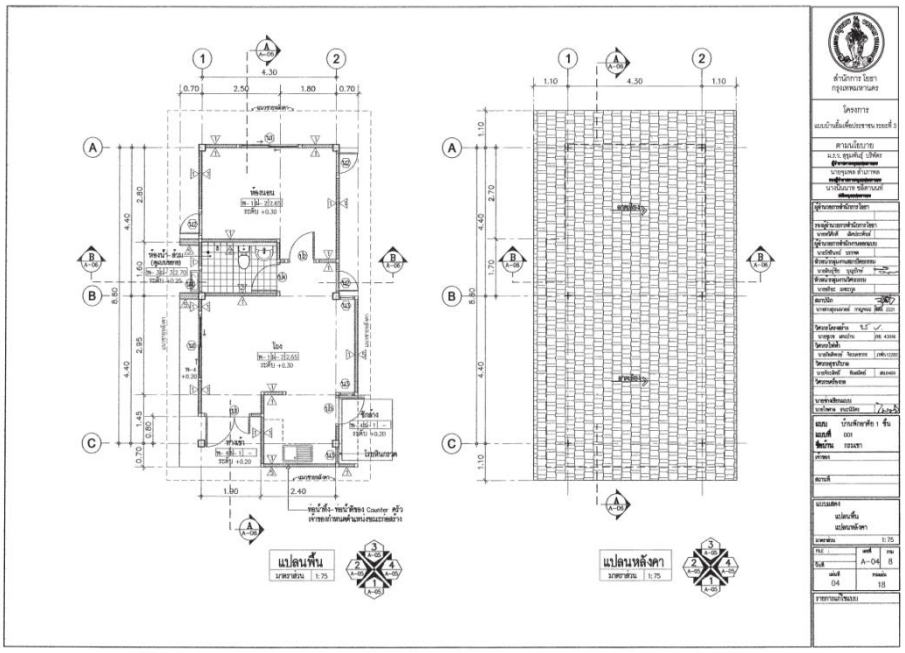
ไพทอน (python) ไพทอนเป็นภาษาโปรแกรมที่ใช้กันอย่างแพร่หลายในการพัฒนาซอฟต์แวร์ วิทยาศาสตร์ข้อมูล และการเรียนรู้เชิงลึก (Machine Learning) ซึ่งเหมาะแก่การทำงานทางด้านวิทยาศาสตร์ข้อมูลและการพัฒนาแบบจำลองทำนายราคา มีความเรียนรู้ง่ายและมีระบบหลากหลายให้เลือกใช้ และสามารถนำไปใช้งานบนแพลตฟอร์มต่างๆ ได้ โดยภาษา Python ยังสามารถใช้งานร่วมกับหลายๆ ระบบและมีความสามารถในการปรับแต่งส่วนต่างๆ ตามความต้องการของผู้วิจัย นอกจากนี้ยังเป็นภาษาที่รวดเร็วในการพัฒนาโปรแกรม

การเก็บรวบรวมข้อมูล

การดำเนินงานวิจัยในครั้งนี้ได้นำข้อมูลที่ใช้ในการศึกษาประกอบด้วยแบบบ้านพักอาศัยโครงสร้างคอนกรีตเสริมเหล็กที่ออกแบบและรวบรวมของกรุงเทพมหานครเพื่อใช้ประมาณราคา ซึ่งรูปแบบของบ้านที่ได้ทำการศึกษานั้นมีลักษณะดังนี้ 1) ฐานรากเป็นแบบฐานรากวางบนเสาเข็ม และใช้เสาเข็มแบบเสาเข็มคอนกรีตอัดแรงสี่เหลี่ยมตัน 2) โครงสร้างเป็นแบบโครงสร้างคอนกรีตเสริมเหล็กหล่อในที่ 3) ผนังเป็นผนังแบบก่ออิฐ-ฉาบปูน 4) ความสูงของบ้านไม่เกิน 3 ชั้น 5) ลักษณะหลังคาเป็นทรงปั้นหยา มีการใช้วัสดุมุงเป็นกระเบื้องซีแพคโมเนียหรือเป็นหลังคาคอนกรีต (ไม่รวมมุ้งลวด เหล็กดัด งานรั้ว งานจัดสวน ตกแต่งภายนอก ตกแต่งภายใน ระบบไฟฟ้า ประปาภายนอก) 6) บ้านที่มีห้องคนใช้จะนับเป็นอีก 1 ห้องนอน 7) บ้านที่มีห้องรับแขกอยู่แล้วและมีห้องนั่งเล่น ห้องพักผ่อน ห้องทำงาน ที่มีผนังไม่ถึง 4ด้าน จะนับห้องเหล่านี้เป็นห้องรับแขก 8) บ้านที่มีส่วนทานอาหารอยู่ในห้องครัวจะรวมนับเป็นห้องครัว 9) ห้องที่มีผนัง 4 ด้านและมีประตูเปิดเข้าออกได้ นับเป็นอีก1 ห้องนอน (เช่น ห้องทำงาน ห้องดูหนังฟังเพลง) 10) พื้นที่ส่วนจอดรถ จะคิดพื้นที่ถึงขอบเสาด้านใน 11) ความสูงของชั้นจะวัดจากระดับพื้นถึงระดับพื้นในชั้นถัดไป ยกเว้นความสูงหลังคาจะวัดจากระดับหลังอะเสจนถึงส่วนที่สูงที่สุดของหลังคา 12) กันสาดจะนับเป็นส่วนหนึ่งของหลังคาคอนกรีต

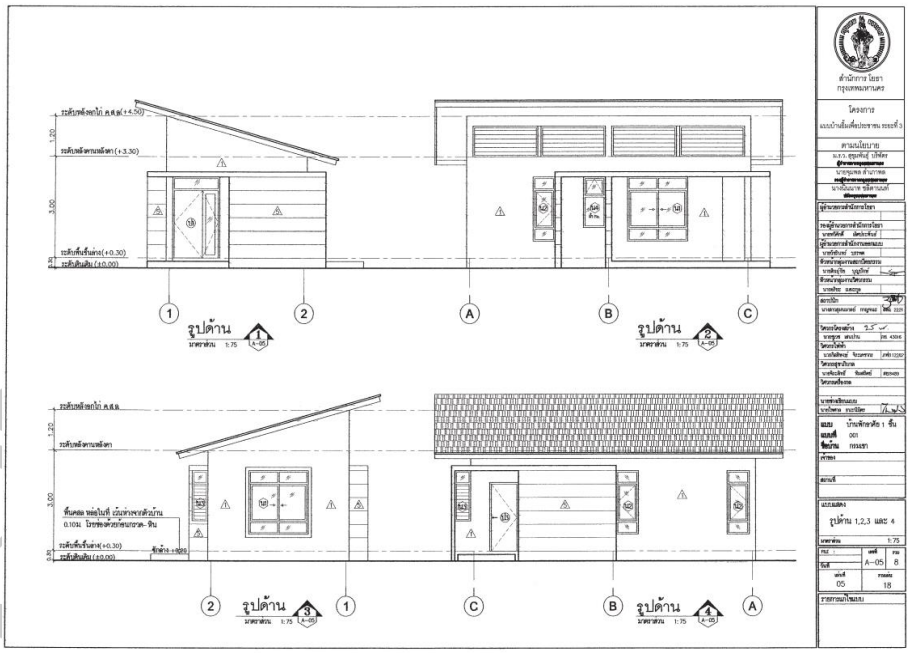
โดยข้อมูลที่ใช้ในการศึกษาได้ทำการรวบรวมจากข้อมูลจากโครงการแบบบ้านยิ้ม ซึ่งเป็นแบบบ้านพักอาศัยโครงสร้างคอนกรีตเสริมเหล็ก ที่ออกแบบของกรุงเทพมหานคร จำนวน 120 รายการ ตัวอย่างแบบบ้านดังภาพประกอบ 21 ถึงภาพประกอบ 22

พหุน ปณ ฑิโต ชีเว



ที่มา: กลุ่มงานประมาณราคา สำนักการโยธา กรุงเทพมหานคร (2564)

ภาพประกอบ 21 ตัวอย่างแบบแปลนบ้านจากโครงการแบบบ้านยิ้ม แบบบ้านฟรีเพื่อประชาชน



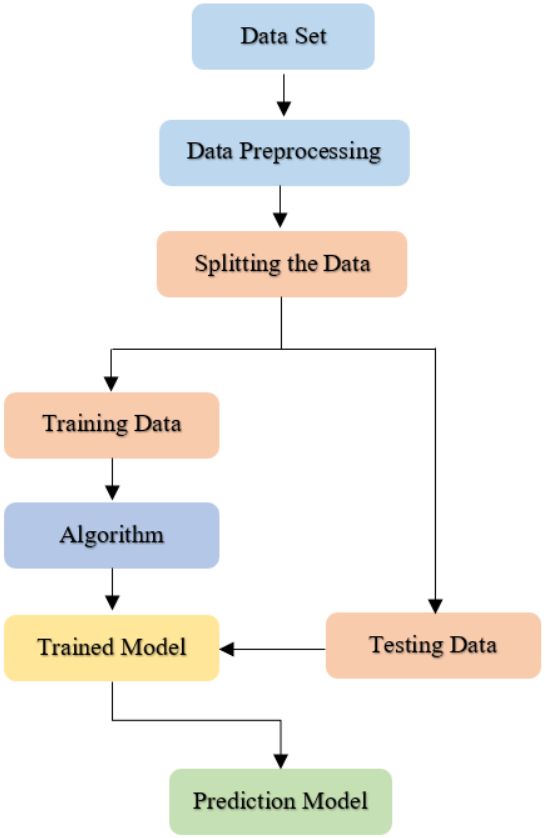
ที่มา: กลุ่มงานประมาณราคา สำนักการโยธา กรุงเทพมหานคร (2564)

ภาพประกอบ 22 ตัวอย่างรูปด้านของบ้านจากโครงการแบบบ้านยิ้ม แบบบ้านฟรีเพื่อประชาชน

การจัดกระทำกับข้อมูล

ขั้นตอนการประมวลผลข้อมูลที่สำคัญถูกดำเนินการอย่างละเอียดและรอบคอบ ซึ่งรวมถึงการลบข้อมูลที่ขาดหายไปและการแบ่งข้อมูลอย่างแม่นยำเพื่อเป็นชุดการฝึกและทดสอบ กระบวนการเหล่านี้รวมถึงการทัศนภาพของข้อมูลเพื่อดูและตรวจสอบลักษณะของข้อมูลและการทำความสะอาดข้อมูล เพื่อให้มั่นใจในการจัดการของค่าที่ขาดหายไปและค่าที่แปลกปลอมที่สุดโต่งก่อนการนำแบบจำลองไปใช้งานในขั้นตอนของวิธีการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ ความสนใจอยู่ที่การเลือกชุดข้อมูลคุณภาพสูง ที่คัดสรรจากแหล่งที่มีความเชื่อถือ ซึ่งงานวิจัยนี้เลือกจากหน่วยงานราชการของกรุงเทพมหานคร ประกอบด้วยคุณลักษณะ เช่น พื้นที่ห้อง ความสูงของอาคาร สิ่งอำนวยความสะดวก และอื่นๆ ซึ่งทั้งหมดมีผลต่อราคาของบ้านพักอาศัยแนวราบ เทคนิคการประมวลผลข้อมูลที่ละเอียดและรอบคอบทำให้ชุดข้อมูลพร้อมสำหรับการจัดการกับค่าที่แปลกปลอมที่สุดโต่ง แก้ไขค่าที่ขาดหายไป และค้นหาลักษณะที่มีความหมาย สุดท้าย การแบ่งข้อมูลเป็นชุดการฝึก การตรวจสอบ และการทดสอบ สนับสนุนการประเมินประสิทธิภาพของแบบจำลองอย่างแม่นยำ ซึ่งทำให้สามารถฝึกโมเดล ปรับค่าพารามิเตอร์และประเมินผลการดำเนินงาน

การสร้างแบบจำลองจะต้องการจัดกระทำกับข้อมูลเพื่อการวิเคราะห์ข้อมูล เพราะวิธีการสร้างแบบจำลองการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบเอ็กซ์ตรีม ประกอบด้วยทั้งหมด 5 ระยะดังที่กล่าวข้างต้น โดยระยะการจัดกระทำกับข้อมูลและการวิเคราะห์ข้อมูลถือมีความสำคัญอย่างมากในการสร้างแบบจำลอง ชุดข้อมูลได้รวบรวมและถูกแบ่งเป็นชุดฝึกอบรม การตรวจสอบความถูกต้องและการทดสอบ ข้อมูลเริ่มต้นจะถูกดำเนินการกระบวนการก่อนการวิเคราะห์ เพื่อลบข้อมูลชุดพิเศษและเพิ่มความคงทนให้กับชุดข้อมูลเพื่อปรับปรุงประสิทธิภาพของแบบจำลอง (Nusrat Jahan et al., 2023; Zelaya, 2019) วิธีการกระบวนการที่หลากหลายถูกนำมาใช้ เช่น การทำความสะอาดข้อมูล การจัดการข้อมูลที่ขาดหายไป การประมวลผลข้อมูลแบบหมวดหมู่ การเลือกคุณสมบัติและการปรับขนาดคุณสมบัติ ต่อมาแบบจำลองการเรียนรู้ของเครื่องจะถูกฝึกด้วยข้อมูลที่ผ่านกระบวนการจัดเตรียมข้อมูลแล้ว โดยมีวิธีการฝึกและทดสอบที่แตกต่างกันสำหรับแต่ละแบบจำลอง ในขณะที่ส่วนใหญ่ของแบบจำลองแสดงประสิทธิภาพที่ดี ก็จะมีบางแบบจำลองที่สามารถเข้ากันได้น้อยกว่า วิธีการกระบวนการที่ครอบคลุมครบถ้วนจะช่วยในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบได้อย่างแม่นยำ และช่วยในการเลือกแบบจำลองที่เหมาะสมที่สุดสำหรับงานนี้



ภาพประกอบ 23 วิธีการวิเคราะห์ของแบบจำลอง

การทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบอย่างแม่นยำเป็นสิ่งสำคัญในวงการอสังหาริมทรัพย์ งานวิจัยนี้นำเสนอกระบวนการวิธีจำนวน 6 ขั้นตอน ดังภาพประกอบ 23 สำหรับการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ ที่ประกอบด้วยขั้นตอนหลัก ๆ ดังนี้: การเลือกชุดข้อมูล, การเตรียมข้อมูลก่อนการวิเคราะห์, การแบ่งข้อมูล, การฝึกแบบจำลอง, การเลือกอัลกอริทึม, และการสร้างแบบจำลองทำนาย การสร้างขึ้นจากชุดข้อมูลที่ถูกจัดกระทำเรียบร้อยแล้วจากแหล่งข้อมูลที่เชื่อถือได้เป็นฐานพื้นฐานสำคัญของกระบวนการวิธีนี้ ผ่านกระบวนการเตรียมข้อมูลหลัก ๆ เน้นไปที่ขั้นตอนสำคัญ เช่น การจัดการกับข้อมูลชุดพิเศษ การปรับปรุงความสะอาดข้อมูล การแทนค่าข้อมูลที่ขาดหาย และการสร้างคุณสมบัติ การดำเนินการเหล่านี้รวมกันเสริมความคงทนของชุดข้อมูลในขณะที่เพิ่มประสิทธิภาพรวมของแบบจำลอง จากนั้น ชุดข้อมูลถูกแบ่งอย่างรอบคอบเป็นชุดฝึกอบรม การตรวจสอบความถูกต้อง และการทดสอบ ที่ช่วยให้เราสามารถประเมินประสิทธิภาพของแบบจำลองอย่างแม่นยำ จากนั้นจัดกระทำตัวอัลกอริทึมการเรียนรู้ของเครื่องในทางแตกต่างกัน แล้วทำการเลือกตัวเลือกที่เหมาะสมที่สุดสำหรับงานการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ

โดยนำอัลกอริทึมที่เลือกมาฝึกโดยใช้ข้อมูลฝึกอบรมที่เตรียมไว้ สร้างกระบวนการเรียนรู้เพื่อแยกความรู้จากชุดข้อมูล ในท้ายที่สุด จะสร้างแบบจำลองการทำนายเพื่อใช้สำหรับการทำนาย ราคาค่าก่อสร้างบ้านพักอาศัยแนวราบอย่างแม่นยำและเชื่อถือได้

1) การเลือกชุดข้อมูล ขั้นตอนแรกในกระบวนการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบเป็นการเลือกชุดข้อมูลที่มีคุณภาพและเกี่ยวข้องกับงานวิจัย เราจะเตรียมชุดข้อมูลที่ถูกเลือกมาอย่างรอบคอบจากแหล่งข้อมูลที่เชื่อถือได้ โดยให้ความสำคัญกับการรวบรวมคุณลักษณะที่ครอบคลุมที่สุดซึ่งมีผลต่อราคาของบ้านพักอาศัยแนวราบ ชุดข้อมูลประกอบไปด้วยคุณสมบัติต่าง ๆ เช่น ขนาดพื้นที่ของห้องต่าง ๆ จำนวนห้อง ความสูงของอาคาร และปัจจัยที่เกี่ยวข้องอื่น ๆ

2) เตรียมข้อมูล เพื่อเตรียมข้อมูลให้พร้อมสำหรับการสร้างแบบจำลอง เราจะนำเอาเทคนิคการเตรียมข้อมูลที่ดีมาใช้งาน การตรวจสอบและจัดการกับข้อมูลชุดพิเศษจะถูกนำมาดำเนินการเพื่อลดผลกระทบของข้อมูล ข้อมูลจะถูกปรับปรุงผ่านกระบวนการแก้ไขค่าผิดปกติ และการแก้ไขค่าข้อมูลที่ขาดหายไป โดยใช้การแทนค่าหรือมาตรการแก้ไขที่เหมาะสม เพื่อให้ชุดข้อมูลครบถ้วน Feature engineering จะถูกดำเนินการเพื่อสร้างคุณสมบัติที่มีความหมายจากคุณสมบัติที่มีอยู่เพื่อเสริมประสิทธิภาพในการทำนาย

3) แบ่งข้อมูล เพื่อประเมินประสิทธิภาพของแบบจำลองอย่างแม่นยำ เราจะแบ่งชุดข้อมูลเป็นส่วนย่อย ๆ ได้แก่ชุดฝึกอบรม การตรวจสอบความถูกต้อง และการทดสอบ ชุดข้อมูลที่ใช้ในการฝึกแบบจำลอง เช่น ชุดข้อมูลฝึกอบรม จะถูกใช้ในการฝึกแบบจำลอง ในขณะที่ชุดข้อมูลการตรวจสอบความถูกต้องจะถูกใช้ในการปรับปรุงพารามิเตอร์และหลีกเลี่ยงการเกิดการเรียนรู้ที่เกินไป (overfit) และชุดข้อมูลทดสอบจะเป็นชุดที่แยกออกมาเพื่อประเมินประสิทธิภาพของแบบจำลองในชุดข้อมูลสุดท้าย

4) การฝึกแบบจำลอง แบบจำลองการเรียนรู้ของเครื่องจะถูกฝึกโดยใช้ข้อมูลที่ผ่านกระบวนการเตรียมและแบ่งแล้ว มีการประเมินวิธีการเรียนรู้หลากหลาย เช่น โครงข่ายประสาทเทียม การสร้างแบบจำลองเชิงเส้นหลายตัวแปร ซัพพอร์ตเวกเตอร์แมชชีน ต้นไม้ตัดสินใจ หรือแรนดอมฟอเรสต์ เทคนิคการเลือกแบบจำลองที่เหมาะสมที่สุดจะต้องมีการฝึกแบบจำลองอย่างเข้มงวดบนข้อมูลฝึกอบรม ที่ช่วยให้แบบจำลองเรียนรู้รูปแบบและความสัมพันธ์ในข้อมูล

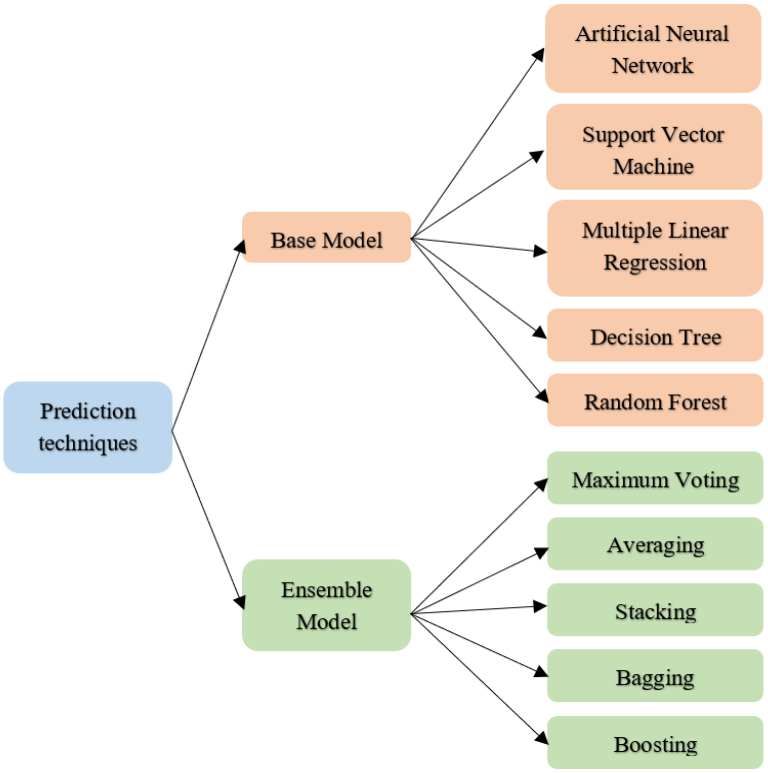
5) การเลือกอัลกอริทึม ความสำเร็จในการทำนายราคาของบ้านพักอาศัยแนวราบนั้นขึ้นกับการเลือกอัลกอริทึมที่เหมาะสมที่สุด การประเมินประสิทธิภาพของอัลกอริทึมต่าง ๆ จะถูกดำเนินการโดยพิจารณาอย่างรอบคอบ โดยใช้ตัวชี้วัดการประเมินที่เกี่ยวข้อง เช่น R-square (R2), Mean Squared Error (MSE), Root Mean Squared Error (MSE) หรือ Mean Absolute Error (MAE) บนชุดข้อมูลการตรวจสอบความถูกต้อง อัลกอริทึมที่แสดงความสามารถในการทำนายที่ดีกว่าและมีความสามารถในการสร้างกรอบการแยกแยะที่แข็งแกร่งถูกเลือกเป็นแบบจำลองการทำนาย

6) การสร้างแบบจำลองการทำนาย เมื่อเลือกอัลกอริทึมที่เหมาะสมเรียบร้อยแล้ว แบบจำลองการทำนายสุดท้ายจะถูกสร้างขึ้น แบบจำลองนี้จะพร้อมที่จะทำการทำนายราคาของบ้านพักอาศัยแนวราบอย่างแม่นยำตามคุณสมบัติที่ระบุเป็นข้อมูลเข้า กระบวนการวิธีที่ถูกกำหนดอย่างชัดเจนนี้จะเป็นเครื่องมือที่มีประสิทธิภาพสำหรับการทำนายราคาของบ้านพักอาศัยแนวราบอย่างแม่นยำ

ในสรุป การรวมขั้นตอนสำคัญตั้งแต่การเลือกและเตรียมข้อมูลไปจนถึงการฝึกแบบจำลองและการเลือกอัลกอริทึม กระบวนการทำนายราคาบ้านพักอาศัยแนวราบในแนวทางที่รอบคอบนี้สามารถทำนายราคาบ้านพักอาศัยแนวราบอย่างแม่นยำได้อย่างเป็นระบบ ด้วยการปฏิบัติตามวิธีการอย่างเป็นระบบนี้ ผู้เชี่ยวชาญด้านอสังหาริมทรัพย์ วิศวกร สถาปนิก เจ้าที่หน้าที่ถอดแบบและประมาณราคา นักลงทุน และวิเคราะห์สามารถทำการตัดสินใจด้วยความมั่นใจ จากการทำนายราคาบ้านพักอาศัยแนวราบที่แม่นยำและน่าเชื่อถือ

การสร้างแบบจำลอง

การสร้างแบบจำลองเพื่อทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบของบ้านพักอาศัยแนวราบ สามารถแบ่งออกเป็นสองแบบหลัก คือ 1) แบบจำลองเรียนรู้แบบเชิงเดี่ยว (Surawatchayotin & Paireekreng, 2021) หรือแบบจำลองพื้นฐาน และ 2) แบบจำลองเอ็กซ์ตรีมที่เป็นการเรียนรู้กลุ่ม (Rahimi et al., 2023) กลุ่มแรกคือแบบทางคณิตศาสตร์และสถิติศาสตร์ ที่ใช้ข้อมูลที่มีผลลัพธ์เป็นข้อมูลนำเข้าในการทำนาย แบบนี้เหมาะสำหรับการใช้พื้นที่ห้องและพารามิเตอร์อื่นๆ มีหลายเทคนิคที่เข้าร่วมอยู่ในแบบทางสถิติศาสตร์ ทั้งหมด 5 อัลกอริทึม คือ โครงข่ายประสาทเทียม (artificial neural network) ซัพพอร์ทเวกเตอ์แมชชีน (support vector machine) การวิเคราะห์ถดถอยพหุคูณ (multiple linear regression) ต้นไม้ตัดสิน (decision tree) และแรนดอมฟอเรสต์ (random forest) ส่วนกลุ่มที่สองเป็นใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีมแบบเรียนรู้กลุ่มซึ่งรวมเอาวิธีการทางสถิติมาผสมผสานกับคุณสมบัติที่เป็นเอกลักษณ์เพื่อรับมือกับข้อจำกัดของแบบบ้านแต่ละรูปแบบ การใช้วิธีการเรียนรู้กลุ่มเพิ่มประสิทธิภาพในการทำนายและคุณภาพของแบบจำลองอีก 5 อัลกอริทึม คือ แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) และ แบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) ดังภาพประกอบ 24



ภาพประกอบ 24 เทคนิคการทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ

การสร้างแบบจำลองทั้งหมด 10 อัลกอริทึมโดยใช้ภาษา Python ซึ่งมีฟังก์ชันสำหรับสร้างรูปแบบโครงข่ายประสาทเทียม (artificial neural network) ซัพพอร์ทเวกเตอร์แมชชีน (support vector machine) การวิเคราะห์ถดถอยพหุคูณ (multiple linear regression) ต้นไม้ตัดสิน (decision tree) และแรนดอมฟอเรสต์ (random forest) การเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) การเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning) การเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning) การเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) และการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning)

เนื่องจากระเบียบวิธีหรืออัลกอริทึม (Algorithm) ของแต่ละเทคนิคแตกต่างกันและสามารถแก้ปัญหาได้ในลักษณะที่แตกต่างกันด้วย และมีความเหมาะสมที่แตกต่างกันออกไป ขึ้นอยู่กับชุดข้อมูลที่ใช้งาน จึงจำเป็นต้องมีการเลือกรูปแบบโครงสร้าง ระเบียบวิธี และพารามิเตอร์ต่างๆ ที่เหมาะสม ในการศึกษานี้ได้ใช้พารามิเตอร์ต่างๆ จากนั้นจะกำหนดอัตราการเรียนรู้ของแบบจำลองเครื่องจักรเรียนรู้เอ็กซ์ตรีมระหว่างชุดฝึกสอนกับชุดทดสอบเพื่อหาอัตราการเรียนรู้ที่เหมาะสมที่สุด และในขั้นตอนนี้ ตัวแบบชุดฝึกสอนจะถูกฝึกอบรมและปรับแก้ตามข้อผิดพลาด และเปรียบเทียบกับตัวแบบชุดทดสอบเพื่อหาค่าความถูกต้องที่เหมาะสมที่สุดโดยใช้ค่านัยสำคัญจากการทดสอบ (Validation)

การทดสอบเปรียบเทียบประสิทธิภาพ

ทำการวิเคราะห์และเปรียบเทียบประสิทธิภาพของแบบจำลองสำหรับการทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ โดยมีขั้นตอนการดำเนินงาน และอธิบายได้ดังต่อไปนี้

1) ทำการวิเคราะห์ผลลัพธ์ที่ได้ของแต่ละอัลกอริทึม สำหรับพิจารณาการตั้งค่าพารามิเตอร์ที่ได้ประสิทธิภาพที่ดีที่สุด จากชุดข้อมูลที่ใช้ในการวิจัย โดยหลังจากได้ทำการวิเคราะห์และปรับปรุงพารามิเตอร์แล้ว

2) ทำการเปรียบเทียบผลลัพธ์ที่ได้ของแต่ละอัลกอริทึม เพื่อให้ได้ผลลัพธ์ที่ดีที่สุดของแต่ละอัลกอริทึมที่ใช้ในการวิจัย ที่ใช้การเรียนรู้กลุ่มแบบเครื่องจักรเรียนรู้เอ็กซ์ตรีม (Extreme Machine Learning)

3) ทำการเปรียบเทียบประสิทธิภาพและความแม่นยำในการการประมาณราคาที่ได้จากวิธีดั้งเดิมคือใช้คนในการประมาณราคากับวิธีการใหม่ที่ใช้การเรียนรู้ของเครื่องจักรเรียนรู้เอ็กซ์ตรีมในการทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ เพื่อให้ทราบว่าวิธีการใดให้ผลประสิทธิภาพและความแม่นยำมากที่สุด

งานวิจัยนี้ได้มีการระบุเครื่องมือการทำนายที่ดีที่สุดสำหรับการเรียนรู้ของเครื่องจักรเรียนรู้เอ็กซ์ตรีม ซึ่งใช้ตัวชี้วัดประสิทธิภาพดังนี้ Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE) และ R-Square (R2) ตัวชี้วัดประสิทธิภาพเหล่านี้อธิบายได้ดังต่อไปนี้

1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) เป็นตัววัดความคลาดเคลื่อนเฉลี่ยของค่าสัมบูรณ์ระหว่างค่าทำนายและค่าจริงของราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ โดยเอาค่าเฉลี่ยของความแตกต่างระหว่างทั้งสอง ค่าดังกล่าวนี้มีประโยชน์เมื่อมีความคลาดเคลื่อนในการทำนายที่เป็นรูปแบบเดียวกัน เนื่องจากมันจะจัดการกับความแตกต่างในข้อมูลทั้งหมดอย่างเท่าเทียมกัน

MAE มักถูกใช้งานในการแก้ปัญหาเชิงสถิติเพื่อประเมินความแม่นยำโดยรวมของการทำนาย โดยเฉพาะในงาน Regression ที่เป็นการทำนายค่าตัวเลข โดยที่ MAE จะสามารถแสดงถึงความถูกต้องของการทำนายทั้งหมดได้

ค่า MAE ที่น้อยกว่าจะแสดงถึงประสิทธิภาพในการทำนายที่ดีขึ้นในการทำนายราคาของบ้านพักอาศัยแนวราบและแสดงถึงความใกล้เคียงของค่าทำนายกับต้นทุนจริง ค่า MAE ที่น้อยกว่าแสดงถึงการสอดคล้องกันที่มากขึ้นระหว่างค่าทำนายและค่าจริง (Kumari & Toshniwal, 2021; Vignola et al., 2016) ดังสมการที่ 14 สมการนี้จะคำนวณความแตกต่างสัมบูรณ์ระหว่างค่าจริงและค่าทำนายสำหรับแต่ละรายการ จากนั้นคำนวณค่าเฉลี่ยของความแตกต่างสัมบูรณ์เหล่านี้สำหรับทุกรายการ เพื่อให้ได้วัดความแม่นยำโดยรวมของการทำนาย

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_{i(pred)} - y_{i(actual)}\right| \quad (14)$$

เมื่อ

MAE คือ ค่าความคลาดเคลื่อนเฉลี่ยเฉลี่ยสัมบูรณ์ หรือค่า MAE เป็นค่าเฉลี่ยของความแตกต่างสัมบูรณ์ระหว่างค่าจริงและค่าทำนาย

N คือ จำนวนรายการทั้งหมด

yi(actual) คือ ราคาจริงของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

yi(pred) คือ ราคาที่ทำนายของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

2. Mean Square Error (MSE)

Mean Square Error (MSE) เป็นตัวชี้วัดที่ใช้ประเมินความแม่นยำในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ หรือเรียกอีกชื่อหนึ่งว่า L2 Loss คำนวณโดยการหาค่าเฉลี่ยของความแตกต่างที่ถูกยกกำลังสองระหว่างค่าจริงและค่าทำนาย หรือกล่าวอีกอย่าง MSE คือวิธีการคำนวณความแตกต่างที่ถูกยกกำลังสองระหว่างราคาที่ทำนายไว้และราคาจริงของบ้านพักอาศัยแนวราบสำหรับแต่ละกรณีการทำนาย หลังจากการคำนวณนี้ กระบวนการคำนวณเฉลี่ยจะถูกนำมาใช้กับความแตกต่างที่ถูกยกกำลังสองเหล่านี้เพื่อให้ได้ผลลัพธ์ของ Mean Squared Error (MSE) (Chandanshive & Kambekar, 2019) ดังสมการที่ 15 ค่า MSE จะส่งผลให้ความคาดหวังที่ผิดพลาดในการทำนายที่มีขนาดใหญ่มากกว่าความคาดหวังที่ผิดพลาดที่มีขนาดเล็ก นั่นเป็นวิธีในการให้ค่าเชิงปริมาณในประสิทธิภาพของแบบจำลองการทำนาย ค่า MSE ที่ต่ำกว่าแสดงถึงประสิทธิภาพในการทำนายที่ดีขึ้น แสดงถึงความแตกต่างที่น้อยระหว่างราคาที่ทำนายและราคาจริงของบ้านพักอาศัยแนวราบ

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_{i(pred)} - y_{i(actual)})^2 \quad (15)$$

เมื่อ

MSE คือ ค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์

N คือ จำนวนรายการทั้งหมด

yi(actual) คือ ราคาจริงของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

yi(pred) คือ ราคาที่ทำนายของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

3. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) เป็นตัวชี้วัดที่ใช้บ่งบอกถึงความแม่นยำในการทำนายราคาของบ้านพักอาศัยแนวราบ คำนวณโดยการหาค่ารากที่สองของค่าเฉลี่ยของความแตกต่างที่ถูก

ยกกำลังสองระหว่างราคาที่ทำนายและราคาจริงของบ้านพักอาศัยแนวราบ (Chandanshive & Kambekar, 2019) ดังสมการที่ 16 ค่า RMSE มีความสำคัญในการจัดการกับข้อมูลที่มีข้อมูลแปลกปลอม (outliers) อย่างมีประสิทธิภาพ โดยสามารถระบุและกำจัดความแตกต่างที่แตกต่างไปจากคาดหวังอย่างสมบูรณ์ได้อย่างมีประสิทธิภาพ นอกจากนี้ RMSE ยังให้น้ำหนักมากกว่ากับความผิดพลาดที่มีขนาดใหญ่กว่า เป็นตัววัดความผิดพลาดหลักที่มีค่าสำหรับการประเมินประสิทธิภาพของแบบจำลองการทำนายราคาของบ้านพักอาศัยแนวราบ

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{i(pred)} - y_{i(actual)})^2} \quad (16)$$

เมื่อ

RMSE คือ ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยสัมบูรณ์

N คือ จำนวนรายการทั้งหมด

yi(actual) คือ ราคาจริงของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

yi(pred) คือ ราคาที่ทำนายของบ้านพักอาศัยแนวราบสำหรับรายการที่ i

4. R square (R^2)

R square (R^2) คือค่าสัมประสิทธิ์การกำหนด เป็นตัววัดที่ใช้ในการประเมินความเหมาะสมของแบบจำลองการทำนาย ค่านี้จะบ่งบอกถึงแบบแผนการแก้ไขของแบบจำลองว่ามีความเหมาะสมและแม่นยำแค่ไหนเมื่อเปรียบเทียบกับแบบจำลองที่ใช้ค่าเฉลี่ยของข้อมูลเป็นการทำนาย ดังสมการที่ 17

$$R^2 = 1 - \left(\frac{SSE}{SST}\right) \quad (17)$$

เมื่อ

SSE (Sum of Squared Errors) หรือผลรวมของกำลังสองของเศษที่เหลือ ซึ่งหมายถึงผลรวมของความแตกต่างระหว่างค่าที่ทำนายและค่าจริงของข้อมูล

SST (Total Sum of Squares) หรือผลรวมของกำลังสองแสดง ซึ่งหมายถึงผลรวมของความแตกต่างระหว่างค่าข้อมูลแต่ละค่ากับค่าเฉลี่ยของข้อมูลทั้งหมด

เมื่อค่า R2 มีค่าใกล้เคียง 1 แสดงถึงความเหมาะสมของแบบจำลองที่ดีกับข้อมูล ในขณะที่ค่า R2 ที่เข้าใกล้ 0 แสดงถึงความไม่เหมาะสมของแบบจำลองในการแสดงความคลาดเคลื่อนของข้อมูลที่มีความแตกต่างจากค่าเฉลี่ยของข้อมูลทั้งหมด

การปรับแก้ค่าทำนายการประมาณราคาค่าก่อสร้าง

ขั้นตอนสุดท้ายในกระบวนการ (Data Output) เกิดขึ้นหลังจากข้อมูลผ่านกระบวนการประมวลผลเรียบร้อยแล้ว ผลลัพธ์ที่ได้นั้นจะตรงกับวัตถุประสงค์ของผู้วิจัย ตัวแปรที่เกี่ยวข้องกับผลลัพธ์นี้จะเป็นเป้าหมายหรือตัวแปรที่ต้องการคำนวณในกระบวนการเรียนรู้ของเครื่อง ข้อมูลที่ถูกสร้างขึ้นจากแหล่งข้อมูลจะมีการจัดเก็บในช่วงเวลาที่แตกต่างจากปัจจุบัน เพื่อให้สอดคล้องกับบริบทของการวิจัย การปรับแก้ไขฐานข้อมูลเรื่องราคาค่าก่อสร้างมีความจำเป็นเพื่อให้ข้อมูลเข้ากับปีเดียวกันที่ใช้ในการศึกษา การปรับฐานนี้ได้ใช้ดัชนีราคาวัสดุก่อสร้างของประเทศ (Construction Materials Price Index, CMI) จากสำนักงานนโยบายและยุทธศาสตร์การค้า (2566) กระทรวงพาณิชย์เป็นปัจจัย (Factor) ที่ใช้ในกระบวนการปรับฐาน โดยการปรับฐานราคาจะถูกนำจากปีที่ใช้ในการคำนวณราคาเป้าหมายไปสู่ปีฐานที่เป็นพื้นฐานในการวัดดัชนีราคา เป็นการใช้ปีปัจจุบัน (พ.ศ. 2566) เป็นปีฐานในการคำนวณดัชนีราคาวัสดุ แหล่งข้อมูลที่แสดงและการประมาณค่าดัชนีราคาวัสดุสามารถใช้ได้ในตารางที่ 5 โดยการคำนวณตัวคูณสำหรับการปรับฐานจะเป็นไปตามสมการที่ 18

ตัวคูณปรับฐาน = ดัชนีราคาวัสดุเฉลี่ย ณ ปี2566/ดัชนีราคาเฉลี่ย ณ ปีที่ทำการคิดราคา *(18)*

ตาราง 5 ดัชนีราคาวัสดุก่อสร้างและปัจจัยที่ใช้เป็นตัวคูณปรับฐาน

รายการ	สัดส่วนน้ำหนัก ก.ค.66	เดือนกรกฎาคม 2566					เดือนมิถุนายน 2566			
		ดัชนี		อัตราการเปลี่ยนแปลง			ดัชนี	อัตราการเปลี่ยนแปลง		
		2566	2565	M/M	Y/Y	A/A		M/M	Y/Y	A/A
ดัชนี	100.0	112.5	113.1	-0.4	-0.5	0.4	113.0	-0.3	-0.9	0.4
หมวดไม้และผลิตภัณฑ์ไม้	5.01	130.7	122.6	0.2	6.6	6.7	130.4	0.0	6.4	6.7
หมวดซีเมนต์	11.47	100.0	98.8	0.0	1.2	2.9	100.0	-0.6	1.2	3.2
หมวดผลิตภัณฑ์คอนกรีต	16.55	110.4	109.5	0.0	0.8	1.6	110.4	0.0	0.8	1.8
หมวดผลิตภัณฑ์เหล็ก	29.68	135.6	141.4	-0.4	-4.1	-4.4	136.1	-0.3	-6.5	-4.4
หมวดกระเบื้อง	7.43	110.1	109.1	0.0	0.9	0.9	110.1	0.0	0.9	0.9
หมวดวัสดุฉาบผิว	3.23	108.1	103.5	-0.1	4.4	3.6	108.2	0.0	4.6	3.4
หมวดสุขภัณฑ์	2.15	100.6	99.0	0.0	1.6	1.3	100.6	0.2	1.6	1.2
หมวดไฟฟ้าและประปา	12.34	104.5	104.0	0.0	0.5	0.8	104.5	-0.4	0.0	0.8
หมวดวัสดุก่อสร้างอื่นๆ	12.13	99.8	103.1	-2.7	-3.2	1.0	102.6	-0.6	-0.2	1.7

ที่มา : กองดัชนีเศรษฐกิจการค้า สำนักงานนโยบายและยุทธศาสตร์การค้า กระทรวงพาณิชย์

หลังจากที่ได้ดำเนินการวิจัยตามที่กล่าวมาข้างต้นเรียบร้อยแล้ว ทำการสรุปผลลัพธ์ที่ได้จากการทดลองและการวิเคราะห์ โดยการสำรวจค่าความคลาดเคลื่อนที่เกิดขึ้นจากการใช้งานแบบจำลอง ปัจจัยที่มีผลต่อความคลาดเคลื่อนในการประมาณการ เป็นต้น

จากนั้นทำการสรุปผลลัพธ์ที่ได้จากการดำเนินการวิจัย ทำการวิจารณ์ค่าความคลาดเคลื่อนที่เกิดขึ้นจากการใช้งาน ปัจจัยที่ส่งผลต่อความคลาดเคลื่อนของการประมาณการจากรูปแบบ ตลอดจนข้อจำกัดของการนำรูปแบบที่ได้ในการวิจัยนี้ไปใช้งาน วิเคราะห์ผลเปรียบเทียบประสิทธิภาพของแบบจำลองของแต่ละอัลกอริทึมและเปรียบเทียบผลของวิธีการดั้งเดิมกับวิธีการใหม่ และผู้วิจัยสามารถสรุปผลการวิจัยและข้อเสนอแนะเพื่อเป็นแนวทางในการวิจัยต่อไป

พหุน ปณ ฑิโต ชีเว

บทที่ 4

ผลการวิเคราะห์ข้อมูล

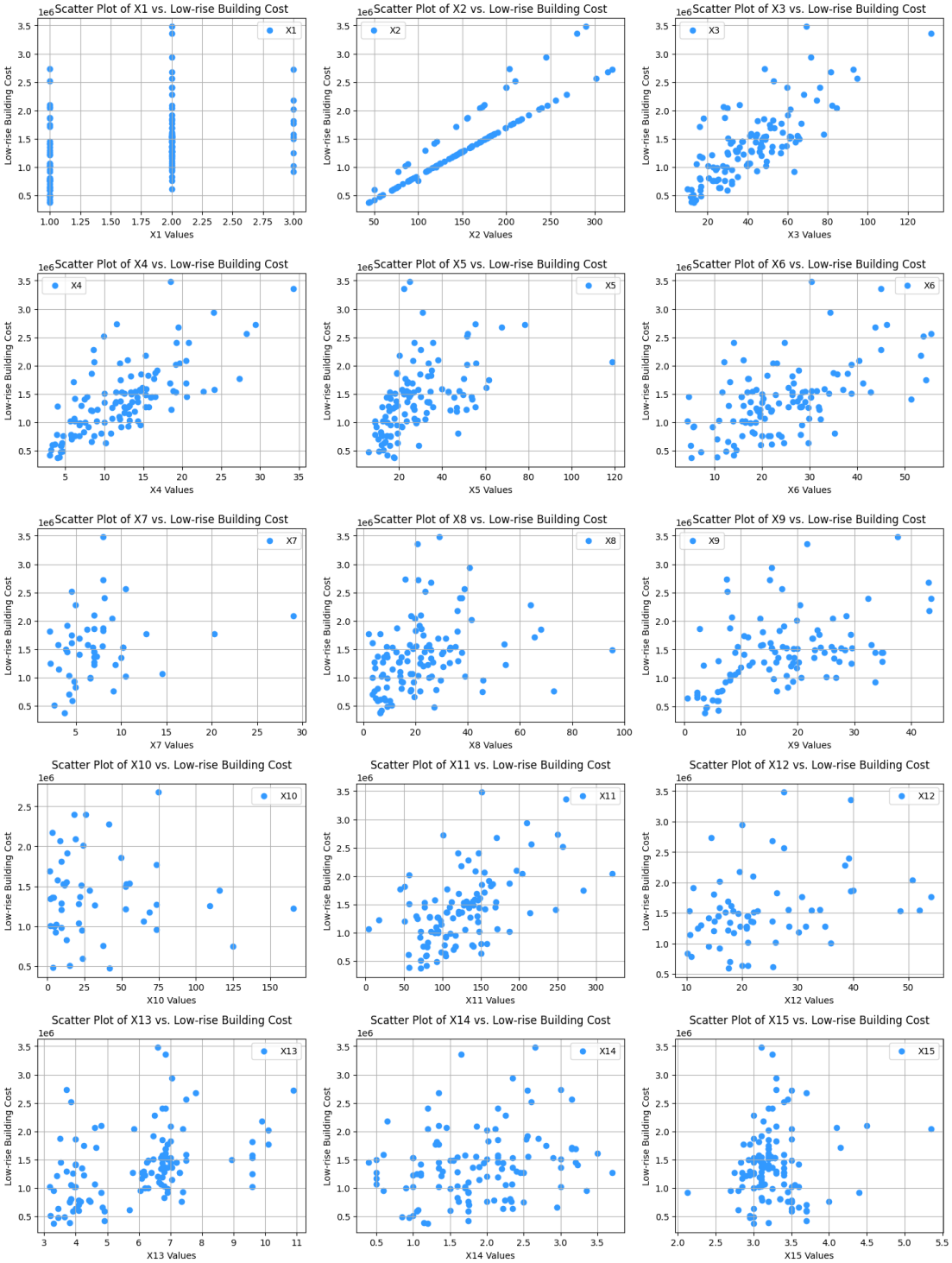
งานวิจัยเรื่องการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม (The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine) โดยใช้เทคนิคการเรียนรู้ของเครื่องต่าง ๆ ผู้วิจัยได้ดำเนินการวิจัยโดยศึกษาตามกรอบแนวคิดวิธีการวิจัยตามระยะต่างๆ ตลอดจนวัดผลประสิทธิภาพแบบจำลองเพื่อให้บรรลุวัตถุประสงค์ของการวิจัยที่ได้กำหนดไว้ได้ ซึ่งรายละเอียดผลการวิเคราะห์ข้อมูล ดังนี้

1. ผลการวิเคราะห์ข้อมูลนำเข้าที่ใช้ในการวิจัย
2. ผลลัพธ์วิธีการดั้งเดิมหรือแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน
3. ผลลัพธ์วิธีการใหม่หรือแบบจำลองเครื่องจักรเรียนรู้กลุ่มเอ็กซ์ตรีม
4. ผลการเปรียบเทียบประสิทธิภาพของวิธีการใหม่ที่นำเสนอในการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบกับวิธีการดั้งเดิม
5. ผลการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองในการการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบเพื่อหาวิธีการที่รวดเร็วและเหมาะสมที่สุด

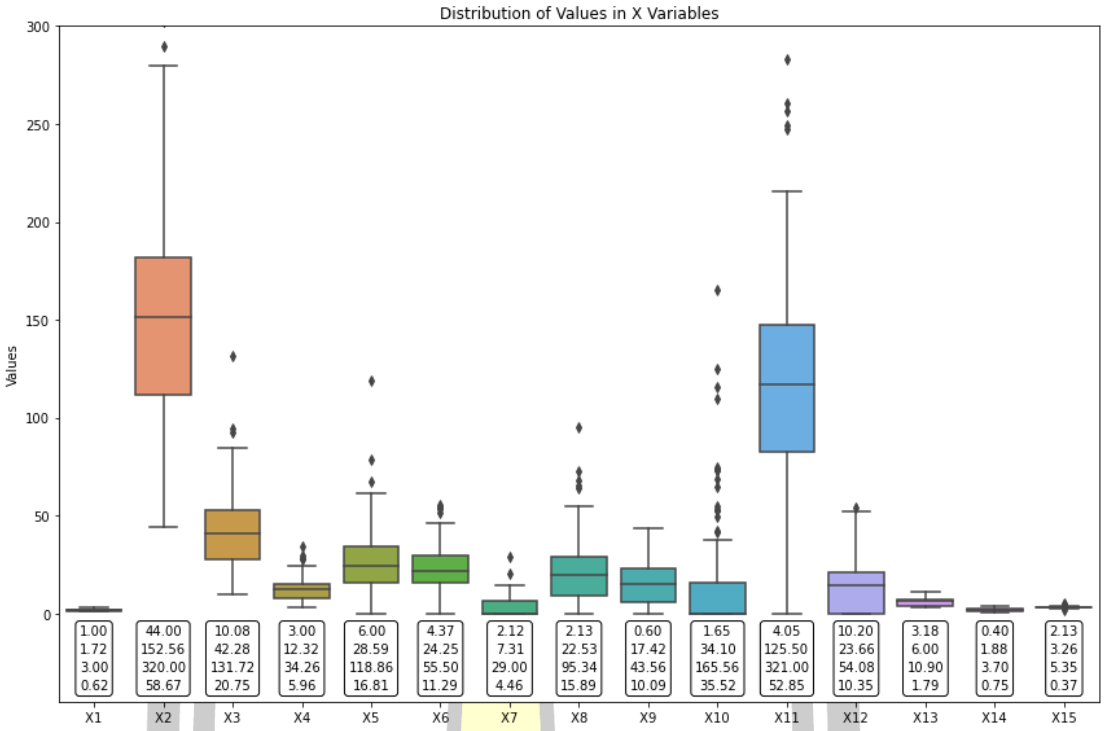
ผลการวิเคราะห์ข้อมูลนำเข้าที่ใช้ในการวิจัย

ในการนำแบบจำลองที่เสนอขึ้นมาประยุกต์ใช้ เราใช้ฐานข้อมูลเกี่ยวกับบ้านพักแนวราบและราคาที่เกี่ยวข้องที่ได้มาจากกลุ่มงานประมาณราคา สำนักงานโยธาของกรุงเทพฯ จำนวน 120 แบบ ซึ่งประกอบด้วยตัวแปรอิสระหรือคุณลักษณะที่ได้จากการทำการศึกษางานวิจัยที่ผ่านมาและจากประสบการณ์ทำงานรวมถึงคำแนะนำจากผู้เชี่ยวชาญทางด้านการถอดแบบและประมาณราคาทั้งสิ้น 15 ตัวแปร ดังที่กล่าวในบทที่ 3

ทำการสำรวจข้อมูลที่ได้จากการเก็บข้อมูลเพื่อดูแนวโน้มและลักษณะของข้อมูลตัวแปรอิสระดังภาพประกอบ 25 และทัศนภาพข้อมูลของแต่ละตัวแปรอิสระพบว่ามีค่าต่ำสุด (Min.) ค่าเฉลี่ย (Mean) ค่าสูงสุด (Max.) และค่าส่วนเบี่ยงเบนมาตรฐานดังภาพประกอบ 26



ภาพประกอบ 25 ลักษณะข้อมูลตัวแปรอิสระ



ภาพประกอบ 26 ทัศนภาพข้อมูลตัวแปรอิสระ

จากภาพประกอบ 26 พบว่าค่าตัวแปร X1 มีช่วงค่าตั้งแต่ 1.00 ถึง 3.00 โดยมีค่าเฉลี่ยที่ 1.72 มีค่าส่วนเบี่ยงเบนมาตรฐานที่ 0.62 ซึ่งแสดงถึงข้อมูลที่กระจายอยู่ใกล้กับค่าเฉลี่ย ชี้ให้เห็นถึงลักษณะที่เสถียรเพื่อวิเคราะห์และงานที่ต้องการความแม่นยำและการควบคุม

ตัวแปร X2 ครอบคลุมช่วงค่าตั้งแต่ 44.00 ถึง 320.00, และมีค่าเฉลี่ยที่ 152.56 ค่าส่วนเบี่ยงเบนมาตรฐานที่ 58.67 แสดงถึงความแปรผันและการกระจายของข้อมูลที่มีขนาดใหญ่ ข้อมูลกระจายอย่างกว้างนี้แสดงถึงข้อมูลที่ซับซ้อนและหลากหลายที่อาจจะต้องมีการสำรวจอย่างละเอียดเมื่อนำมาใช้ในแบบจำลองการวิเคราะห์

ตัวแปร X3 ครอบคลุมช่วงค่าตั้งแต่ 10.08 ถึง 131.72, และมีค่าเฉลี่ยที่ 42.28 ค่าส่วนเบี่ยงเบนมาตรฐานที่ 20.75 แสดงถึงการกระจายและความแปรผันที่มีขนาดใหญ่ ค่าเฉลี่ยที่สูงพร้อมกับค่าส่วนเบี่ยงเบนมาตรฐานที่ปานกลางชี้ให้เห็นถึงระดับความไม่สามารถคาดคะเนได้ซึ่งอาจจะต้องการการสำรวจเพิ่มเติม

ตัวแปร X4 ครอบคลุมช่วงค่าตั้งแต่ 3.00 ถึง 34.26, และมีค่าเฉลี่ยที่ 12.32 ค่าส่วนเบี่ยงเบนมาตรฐานที่ 5.96 แสดงถึงระดับความแปรผันปานกลาง ความสมดุลของค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงฟีเจอร์ X4 ที่มีลักษณะที่ค่อนข้างเสถียรพร้อมกับความแปรผันในระดับปานกลาง

ตัวแปร X5 มีช่วงค่าระหว่าง 6.00 ถึง 118.86 และค่าเฉลี่ยที่ 28.59 X5 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่อยู่ในระดับ 16.81 แสดงถึงการกระจายที่น่าสังเกต การกระจายที่กว้างนี้นั้นนำมาชี้ให้เห็นถึงการแปรผันและรูปแบบที่อาจจะต้องการการสำรวจเพิ่มเติมเพื่อเข้าใจแนวโน้มภายใน

ตัวแปร X6 มีช่วงค่าระหว่าง 4.37 ถึง 55.50, และค่าเฉลี่ยที่ 24.25 ค่าเบี่ยงเบนมาตรฐานที่ 11.29 แสดงถึงความแปรผันปานกลาง ความสมดุลของค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงฟีเจอร์ X6 ที่มีลักษณะที่เสถียรพร้อมกับระดับการแปรผันที่ปานกลาง

ตัวแปร X7 มีช่วงค่าระหว่าง 2.12 ถึง 29.00 และค่าเฉลี่ยที่ 7.31 X7 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 4.46 แสดงถึงการกระจายที่น่าสังเกต ค่าเฉลี่ยที่ต่ำและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงฟีเจอร์ X7 ที่เหมาะสำหรับการวิเคราะห์เนื่องจากมีการกระจายที่มีขนาดเล็กในช่วงค่า

ตัวแปร X8 มีช่วงค่าระหว่าง 2.13 ถึง 95.34 และค่าเฉลี่ยที่ 22.53 X8 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 15.89 แสดงถึงการกระจายที่มีขนาดใหญ่ การกระจายที่กว้างนี้นั้นนำมาชี้ให้เห็นถึงการแปรผันและความหลากหลายในการวิเคราะห์หรือแบบจำลองต่อไป

ตัวแปร X9 มีช่วงค่าระหว่าง 0.60 ถึง 43.56 และค่าเฉลี่ยที่ 17.42 X9 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 10.09 แสดงถึงการกระจายและความแปรผันที่มีขนาดใหญ่ ค่าเฉลี่ยที่ต่ำและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงรูปแบบที่อาจจะต้องการการสำรวจเพิ่มเติม

ตัวแปร X10 มีช่วงค่าระหว่าง 1.65 ถึง 165.56 และค่าเฉลี่ยที่ 34.10 X10 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 35.52 แสดงถึงการกระจายที่มีขนาดใหญ่ การกระจายที่แตกต่างนี้อาจมีผลกระทบต่อกระบวนการวิเคราะห์หรือแบบจำลองเนื่องจากช่วงค่าและความแปรผันที่กว้าง

ตัวแปร X11 มีช่วงค่าระหว่าง 4.05 ถึง 321.00 และค่าเฉลี่ยที่ 125.50 X11 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 52.85 แสดงถึงการกระจายที่มีขนาดใหญ่ ค่าเฉลี่ยที่ต่ำพร้อมกับค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงข้อมูลที่หลากหลายและมีการกระจายที่กว้างที่อาจจะมีผลกระทบต่อการวิเคราะห์ต่อไป

ตัวแปร X12 มีช่วงค่าระหว่าง 10.20 ถึง 54.08 และค่าเฉลี่ยที่ 23.66 X12 แสดงความแปรผัน ค่าเบี่ยงเบนมาตรฐานที่ 10.35 แสดงถึงความแปรผันปานกลาง ความสมดุลของค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงฟีเจอร์ X12 ที่มีลักษณะที่เสถียรพร้อมกับระดับการแปรผันที่ปานกลาง

ตัวแปร X13 มีช่วงค่าระหว่าง 3.18 ถึง 10.90 และค่าเฉลี่ยที่ 6.00 X13 แสดงความแปรผันที่น้อย ค่าเบี่ยงเบนมาตรฐานที่ 1.79 แสดงถึงการกระจายที่น้อย ฟีเจอร์ X13 แสดงถึงลักษณะที่มีความเสถียรและการกระจายที่คงที่

ตัวแปร X14 มีช่วงค่าระหว่าง 0.40 ถึง 3.70 และค่าเฉลี่ยที่ 1.88 X14 แสดงความแปรผันที่น้อย ค่าเบี่ยงเบนมาตรฐานที่ 0.75 แสดงถึงการกระจายที่น้อย ค่าเฉลี่ยที่ต่ำและค่าส่วนเบี่ยงเบนมาตรฐานนี้ชี้ให้เห็นถึงลักษณะที่เสถียรและการกระจายที่น้อย

ตัวแปร X15 มีช่วงค่าตั้งแต่ 2.13 ถึง 5.35, โดยมีค่าเฉลี่ยที่ 3.26 มีค่าส่วนเบี่ยงเบนมาตรฐานที่ 0.37 ซึ่งแสดงถึงข้อมูลที่มีการกระจายใกล้เคียงกันและมีความสมดุลในการวิเคราะห์และการใช้งาน

สรุปคือการวิเคราะห์และแสดงผลสำหรับตัวแปรอิสระแต่ละตัวแปร (X1 ถึง X15) ซึ่งช่วยให้เราเข้าใจลักษณะและคุณสมบัติของข้อมูลในแง่ทางค่าต่ำสุด ค่าเฉลี่ย ค่าสูงสุด ระดับความแปรผันและการกระจายที่เกิดขึ้นในข้อมูล ข้อมูลที่มีค่าเบี่ยงเบนมาตรฐานสูงอาจเป็นที่สนใจในการวิเคราะห์เพิ่มเติมเพื่อเข้าใจความแปรผันและความซับซ้อนของข้อมูล ในขณะที่ข้อมูลที่มีค่าเบี่ยงเบนมาตรฐานน้อยอาจเสถียรและมีความสมดุลมากขึ้น การเข้าใจลักษณะเหล่านี้เป็นสิ่งสำคัญสำหรับการตัดสินใจในการวิเคราะห์ การจัดการข้อมูลในงานวิจัยและการสร้างแบบจำลองการทำนายการประมาณราคาของการเรียนรู้ของเครื่องในแต่ละแบบจำลองและแต่ละออกอริทึม

ผลลัพธ์วิธีการดั้งเดิมหรือแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน

การวิเคราะห์สมการเชิงเส้นหลายตัวแปร (MLR) คือวิธีการวิเคราะห์และแบบจำลองความสัมพันธ์ระหว่างตัวแปรต้นหลายๆ ตัวแปรกับตัวแปรตามหนึ่งตัวแปรเพื่อทำนายหรืออธิบายผลตัวแปรตามด้วยตัวแปรต้น โดยพยายามหาค่าพารามิเตอร์ที่เหมาะสมเพื่อสร้างแบบจำลองทางสถิติที่สอดคล้องกับข้อมูลตัวอย่างที่ให้มา ในกรณีที่มีตัวแปรต้นหลายตัว (multiple independent variables) และตัวแปรตามหนึ่งตัว (dependent variable) เราสามารถใช้สมการเชิงเส้น (linear equation) เพื่อหาความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตามได้ดังสมการ 19

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_{15} X_{15} + \varepsilon \quad (19)$$

เมื่อ

Y คือ ตัวแปรตาม (dependent variable) ที่เราต้องการทำนายหรืออธิบาย

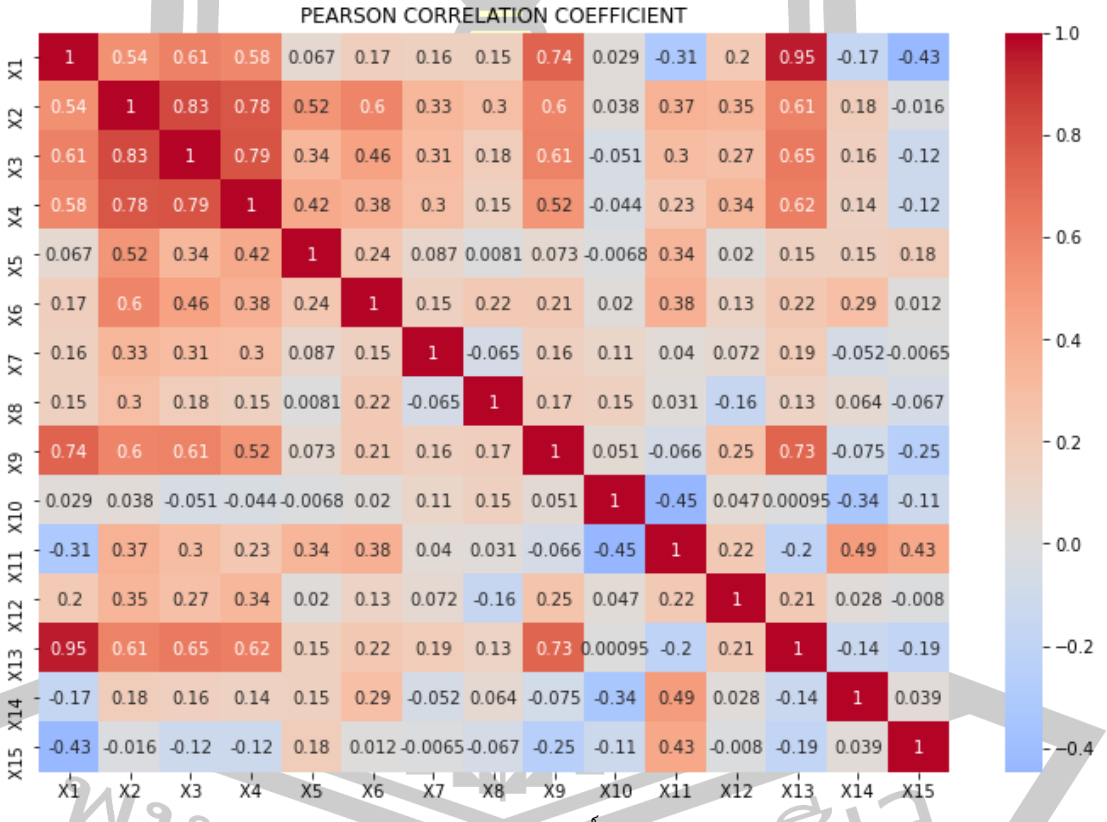
$X_1, X_2, X_3, \cdots, X_{15}$ คือ ตัวแปรอิสระ (independent variables) หลายตัวที่มีผลต่อตัวแปรตาม

$\beta_0, \beta_1, +\beta_2, +\beta_3, \cdots, \beta_{15}$ คือ พารามิเตอร์ (coefficients) ที่แสดงถึงค่าคงที่และความสัมพันธ์ของตัวแปรต้นแต่ละตัว

ε คือ ค่าคลาดเคลื่อนที่ไม่สามารถอธิบายได้โดยแบบจำลอง เป็นค่าความคลาดเคลื่อนทางสถิติ

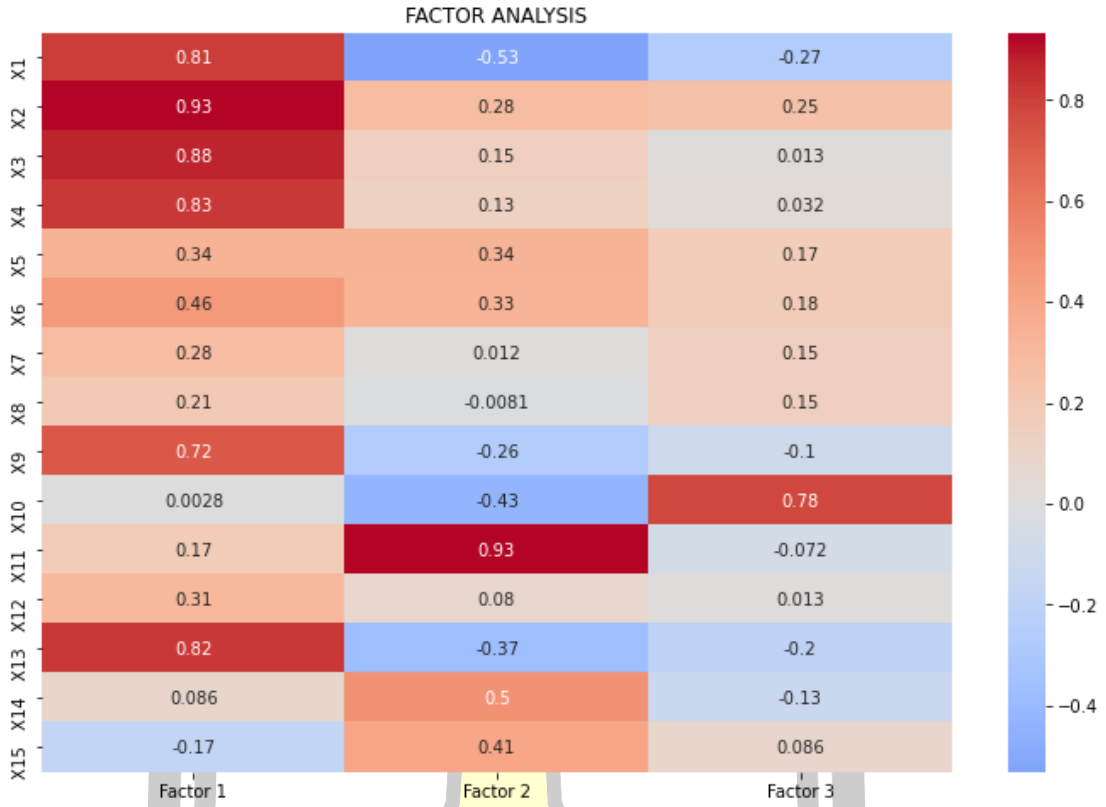
การวิเคราะห์และสร้างแบบจำลอง MLR เป็นกระบวนการที่ใช้ในการทำนายหรือวิเคราะห์ผลตัวแปรตามและตัวแปรต้นในกรณีที่มีหลายตัวแปรต้นที่มีผลต่อผลตัวแปรตาม และมีการคำนวณค่าพารามิเตอร์เพื่อกำหนดว่าแต่ละตัวแปรต้นมีผลอย่างไรต่อผลตัวแปรตาม และใช้แบบจำลองที่ได้สร้างขึ้นเพื่อทำนายหรือวิเคราะห์ผลตัวแปรตามในสถานการณ์ใหม่

การทดสอบค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation Coefficient) เป็นการวิเคราะห์หาความสัมพันธ์ของตัวแปรว่ามีความสัมพันธ์กันเองสูงหรือไม่โดยที่สังเกตุจากค่าที่มีค่าสูงกว่า 0.75 ขึ้นไป (สิทธิกร สิทธิการกูล et al., 2564) โดยให้ถือว่ามีความสัมพันธ์กันเองสูงจำเป็นจะต้องพิจารณาปรับโครงสร้างของตัวแปรก่อนนำไปใช้ในการสร้างแบบจำลองทำนายซึ่งพบว่าตัวแปรที่มีความสัมพันธ์กันสูง



ภาพประกอบ 27 ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

จากภาพประกอบ 27 ซึ่งพบว่าตัวแปรที่มีความสัมพันธ์กันสูงนั้นดังนี้ 1) X1:X13 = 0.95, 2) X2:X3 = 0.83, 3) X2:X4 = 0.78 และ X3:X4 = 0.79 ดังนั้นจะต้องทำการรวมตัวแปรด้วยวิธีการวิเคราะห์ตัวประกอบต่อไป



ภาพประกอบ 28 การวิเคราะห์ตัวประกอบ (Factor Analysis)

การวิเคราะห์ตัวประกอบ (Factor Analysis) เป็นเทคนิคทางสถิติที่ช่วยในการจัดกลุ่มหรือรวมกันของตัวแปรที่มีความสัมพันธ์กันเข้าไว้ในกลุ่มหรือปัจจัยเดียวกัน ตัวแปรที่ถูกจัดในปัจจัยเดียวกันจะมีความสัมพันธ์กันอย่างมีนัยสำคัญ การสัมพันธ์นี้อาจเป็นในทิศทางบวก (ทิศทางเดียวกัน) หรือทิศทางลบ (ทิศทางตรงกันข้าม) อย่างไรก็ตาม ตัวแปรที่ไม่มีความสัมพันธ์กันจะถูกแยกออกจากกันในการวิเคราะห์ตัวประกอบนั้น ๆ (นงลักษณ์ วิรัชชัย, 2542) (มณเฑียร รัตนศิริวงศ์วุฒิ et al., 2553)

จากภาพประกอบ 28 พบว่าสามารถจัดองค์ประกอบได้ 3 กลุ่มดังนี้

ปัจจัยที่ 1 (Factor 1) = 0.81*X1 + 0.93*X2 +0.88*X3 + 0.83*X4 + 0.46*X6 + 0.28*X7 + 0.21*X8 + 0.72*X9 + 0.31*X12 + 0.82*X13

ปัจจัยที่ 2 (Factor 2) = 0.34*X5 + 0.93*X11 + 0.50*X14 + 0.41*X15

ปัจจัยที่ 3 (Factor 3) = X10 (เนื่องจากมีแค่ตัวแปรเดียวจึงไม่จำเป็นต้องคูณตัวคูณ)

เมื่อสามารถจัดกลุ่มขององค์ประกอบได้แล้วนั้น นำตัวแปรในการสร้างแบบจำลองทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบต่อไป

ในการศึกษานี้เป็นการทดลองสร้างจำลองทั้งสิ้น 2 รูปแบบ คือ 1) นำเข้าตัวแปรทุกตัวแปรแบบไม่จัดองค์ประกอบ กับ 2) จัดองค์ประกอบตัวแปรอิสระเป็น 3 ปัจจัยก่อนการนำเข้า แล้วเลือกแบบจำลองการทำนายที่เหมาะสมที่สุดสำหรับผลลัพธ์แบบจำลองด้วยวิธีการวิเคราะห์สมการเชิงเส้นหลายตัวแปร (MLR) และผลของแบบจำลองทั้ง 2 รูปแบบนั้นสามารถสรุปค่าได้ดังแสดงในตาราง 6

ตาราง 6 ผลลัพธ์แบบจำลองด้วยวิธีการวิเคราะห์สมการเชิงเส้นหลายตัวแปร

แบบจำลอง	R2	MSE	RMSE	MAE
MLR-A	0.884	5.056e+10	224856.74	146446.13
MLR-FA	0.834	7.264e+10	269527.90	169969.03

ผลการทดสอบแบบจำลองแสดงให้เห็นถึงประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องจักรที่สร้างขึ้นในการอธิบายความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม ด้วยการใช้ตัวชี้วัดที่สำคัญได้แก่ R-squared (R2), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), และ Mean Absolute Error (MAE).ในการวัดประสิทธิภาพ

ในการทดสอบแบบจำลองหลายตัวแปร (Multiple Linear Regression - MLR) ที่มีการนำเอาตัวแปรต้นหลายตัว (X1 ถึง X15) มาวิเคราะห์ผลตัวแปรตาม ผลการทดสอบแสดงว่าแบบจำลองที่ใช้ตัวแปรต้นทุกตัวมีค่า R-squared (R2) สูงสุดที่ 0.884 เปรียบเทียบกับแบบจำลอง Factor-base (MLR-FA) ที่มีค่า R2 เท่ากับ 0.834 นอกจากนี้ แบบจำลองที่มีค่า R2 สูงสุดหรือ MLR-A ยังมีค่า Mean Squared Error (MSE) เท่ากับ 5.056e+10, Root Mean Squared Error (RMSE) เท่ากับ 224856.74, และ Mean Absolute Error (MAE) เท่ากับ 146446.13 อีกด้วย สำหรับแบบจำลอง Factor-base นอกจากค่า R2 เท่ากับ 0.834 ค่า MSE, RMSE และ MAE เท่ากับ 7.264e+10, 269527.90 และ 169969.03 ตามลำดับ

ค่าสมการสำหรับการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบ สามารถหาค่าราคาค่อก่อสร้างได้จากสมการที่ 20

*ราคาค่าก่อสร้าง = 1.379 + 0.254*X1 + 0.663*X2 – 0.049*X3 + 0.046*X4 – 0.083*X5 - 0.076*X6 – 0.030*X7 – 0.015*X8 – 0.034*X9 + 0.036*X10 + 0.106*X11 – 0.015*X12 -0.278*X13 + 0.017*X14 + 0.088*X15 (หน่วยล้านบาท)* *(20)*

ซึ่งสามารถนำสมการดังกล่าวไปใช้เป็นข้อมูลสำหรับการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบได้ โดยความถูกต้องร้อยละ 88.4

หลังจากสร้างแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐานที่ประกอบด้วยโครงข่ายประสาทเทียม (artificial neural network) ซัพพอร์ทเวกเตอร์แมชชีน (support vector machine) การวิเคราะห์ถดถอยพหุคูณ (multiple linear regression) ต้นไม้ตัดสิน (decision tree) และแรนดอมฟอเรสต์ (random forest) ผลลัพธ์ที่ได้ได้รับการวัดความประสิทธิภาพโดยใช้ตัวชี้ดังนี้ R-squared (R2), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) แสดงดังตาราง 7

ตาราง 7 ประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน

แบบจำลอง	R2	MSE	RMSE	MAE
โครงข่ายประสาทเทียม (ANN)	0.891	47679301000.00	218355.89	167152.98
ซัพพอร์ทเวกเตอร์แมชชี (SVM)	0.446	242250000000.00	492189.90	322532.74
การวิเคราะห์ถดถอยพหุคูณ (MLR)	0.884	50560550000.00	224856.74	146446.13
ต้นไม้ตัดสิน (DT)	0.864	59522890000.00	243973.13	126901.04
แรนดอมฟอเรสต์ (RF)	0.830	74476350000.00	272903.56	164069.51

ผลลัพธ์ในตารางที่ 8 จะช่วยให้เราเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐานในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ โดยที่เราสามารถวิเคราะห์ค่า R2 เพื่อดูความสอดคล้องของแบบจำลองกับข้อมูลจริง และสามารถดูค่า MSE, RMSE, และ MAE เพื่อประเมินความคลาดเคลื่อนของการทำนายราคาของแต่ละแบบจำลอง โดยค่า MSE, RMSE, และ MAE จะเป็นตัวชี้วัดที่แสดงความต่างหรือความคลาดเคลื่อนของค่าทำนายจากค่าจริงโดยเฉลี่ย

แบบจำลอง Artificial Neural Network (ANN) พบว่าค่า R-squared (R2) มีค่าประมาณ 0.891 หมายถึงว่าแบบจำลองสามารถอธิบายความคลาดเคลื่อนของข้อมูลได้ดีกว่า 89% โดยค่า R-squared มีค่าระหว่าง 0 ถึง 1 โดยค่ามากยิ่งดีและมีการสอดคล้องกับข้อมูลมากขึ้น ค่า Mean Squared Error (MSE) มีค่าประมาณ 47,679,301,000.00 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าทำนายจากข้อมูลจริง ค่า MSE น้อยกว่าหมายถึงความคลาดเคลื่อนน้อยของการทำนาย ค่า Root Mean Square Error (RMSE) มีค่าประมาณ 218,355.89 หมายถึงค่าสามารถใช้วัดความคลาดเคลื่อนเฉลี่ยของการทำนายราคาเมื่อนำมาเปรียบเทียบกับราคาจริง ค่า RMSE น้อย

กว่าหมายถึงความคลาดเคลื่อนน้อยของการทำนาย และค่า Mean Absolute Error (MAE) มีค่าประมาณ 167,152.98 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนที่มาจากการทำนายราคา โดยไม่สนใจทิศทาง ค่า MAE น้อยกว่าหมายถึงความคลาดเคลื่อนน้อยของการทำนาย

แบบจำลอง Support Vector Machine (SVM) พบว่าค่า R-squared มีค่าประมาณ 0.446 หมายถึงว่าแบบจำลองสามารถอธิบายความคลาดเคลื่อนของข้อมูลได้ประมาณ 44.6% เท่านั้น ค่า MSE มีค่าประมาณ 242,250,000,000.00 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าทำนายจากข้อมูลจริง ค่า RMSE มีค่าประมาณ 492,189.90 หมายถึงค่าสามารถใช้วัดความคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE มีค่าประมาณ 322,532.74 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนที่มาจากการทำนายราคา

แบบจำลอง Multiple Linear Regression (MLR) พบว่าค่า R-squared มีค่าประมาณ 0.884 หมายถึงว่าแบบจำลองสามารถอธิบายความคลาดเคลื่อนของข้อมูลได้ดีกว่า 88.4% ค่า MSE มีค่าประมาณ 50,560,550,000.00 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าทำนายจากข้อมูลจริง ค่า RMSE มีค่าประมาณ 224,856.74 หมายถึงค่าสามารถใช้วัดความคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE มีค่าประมาณ 146,446.13 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนที่มาจากการทำนายราคา

แบบจำลอง Decision Tree (DT) พบว่าค่า R-squared มีค่าประมาณ 0.864 หมายถึงว่าแบบจำลองสามารถอธิบายความคลาดเคลื่อนของข้อมูลได้ประมาณ 86.4% ค่า MSE มีค่าประมาณ 59,522,890,000.00 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าทำนายจากข้อมูลจริง ค่า RMSE มีค่าประมาณ 243,973.13 หมายถึงค่าสามารถใช้วัดความคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE มีค่าประมาณ 126,901.04 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนที่มาจากการทำนายราคา

แบบจำลอง Random Forest (RF) พบว่าค่า R-squared มีค่าประมาณ 0.830 หมายถึงว่าแบบจำลองสามารถอธิบายความคลาดเคลื่อนของข้อมูลได้ประมาณ 83.0% ค่า MSE มีค่าประมาณ 74,476,350,000.00 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าทำนายจากข้อมูลจริง ค่า RMSE มีค่าประมาณ 272,903.56 หมายถึงค่าสามารถใช้วัดความคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE มีค่าประมาณ 164,069.51 หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนที่มาจากการทำนายราคา

จากการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบพบว่าค่า R-squared สูงสุดคือโครงข่ายประสาทเทียมอยู่ที่ 0.891 แสดงถึงความสามารถในการอธิบายความเปลี่ยนแปลงของข้อมูลประมาณ 89.1% ซึ่งเป็นค่าที่สูง และค่า MSE, RMSE, และ MAE ที่ต่ำเป็นแสดงถึงความคลาดเคลื่อนของการทำนายที่น้อยลง ในขณะที่

แบบจำลองซัพพอร์ตเวกเตอร์แมชชีนมีค่า R-squared ที่ 0.446 ซึ่งเป็นค่าที่ต่ำ แสดงให้เห็นว่าแบบจำลองนี้มีความสามารถในการอธิบายความเปลี่ยนแปลงของข้อมูลเพียง 44.6% ซึ่งอาจแสดงถึงความพอใจในการศึกษาน้อยกว่าแบบจำลองอื่น ๆ ในรูปแบบนี้ แบบจำลอง ANN, MLR, DT, และ RF มีประสิทธิภาพที่ดีกว่า SVM ในการทำนายค่าก่อสร้างบ้านพักอาศัยแนวราบที่ได้รับการประเมินค่า R-squared และค่าคลาดเคลื่อนของการทำนายที่น้อยลง

ด้วยความสามารถในการทำนายและความแม่นยำของผลลัพธ์ แบบจำลองโครงข่ายประสาทเทียมให้ประสิทธิภาพที่ดีที่สุด โดยมีค่า R-squared สูงสุดและค่า RMSE และ MAE ที่ต่ำเมื่อเทียบกับแบบจำลองอื่น ๆ ส่วนแบบจำลอง SVM และ RF มีประสิทธิภาพที่ต่ำกว่า ANN และ MLR แต่ยังคงสามารถทำนายได้ในระดับที่ยอมรับได้ และแบบจำลอง DT มีความสามารถในการทำนายที่ดีในระดับเดียวกับ RF แม้จะไม่ได้ดีที่สุด แต่แบบจำลอง SVM ยังไม่เหมาะสมกับการใช้เพื่อทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบเนื่องจากประสิทธิภาพค่อนข้างต่ำ

ผลลัพธ์วิธีการใหม่หรือแบบจำลองเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม

หลังจากสร้างแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐานที่ประกอบด้วยโครงข่ายประสาทเทียม (artificial neural network) ซัพพอร์ทเวกเตอร์แมชชีน (support vector machine) การวิเคราะห์ถดถอยพหุคูณ (multiple linear regression) ต้นไม้ตัดสินใจ (decision tree) และแรนดอมฟอเรสต์ (random forest) แล้วจึงนำแบบจำลองพื้นฐานเหล่านี้มาสร้างเป็นแบบจำลองเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมหรือการเรียนรู้ของเครื่องจักรขั้นสูงซึ่งจะเป็นวิธีการใหม่ที่เรานำเสนอเพื่อการทำนายการประมาณราคาบ้านพักอาศัยแนวราบ ซึ่งประกอบด้วย แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) และแบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) โดยการวัดความประสิทธิภาพโดยใช้ตัวชี้วัดหลายประเภท ได้แก่ค่า R-squared (R2), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) เช่นเดียวกันกับแบบจำลองพื้นฐาน ซึ่งผลที่ได้ดังแสดงในตาราง 8

ผลลัพธ์ในตาราง 8 ทำให้เราเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองเรียนรู้กลุ่มเครื่องจักรเอ็กซ์ตรีมในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ โดยที่เราสามารถวิเคราะห์ค่า R^2 เพื่อดูความสอดคล้องของแบบจำลองกับข้อมูลจริง และสามารถดูค่า MSE, RMSE, และ MAE เพื่อประเมินความคลาดเคลื่อนของการทำนายราคาของแต่ละแบบจำลอง โดยค่า MSE, RMSE, และ MAE จะเป็นตัวชี้วัดที่แสดงความต่างหรือความคลาดเคลื่อนของค่าทำนายจากค่าจริงโดยเฉลี่ย

ตาราง 8 ประสิทธิภาพของวิธีการใหม่หรือแบบจำลองเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม

แบบจำลอง	R2	MSE	RMSE	MAE
แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning)	0.924	33325570000	182552.93	127356.65
แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning)	0.871	56266850000	237206.35	140085.86
แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning)	0.883	51074880000	225997.52	152669.31
แบบจำลองการเรียนรู้กลุ่มแบบแบ็กกิ้ง (bagging ensemble learning)	0.832	73493900000	271097.58	155363.36
แบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning)	0.846	67425130000	259663.50	147838.86

แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) มีค่า R2 ที่ 0.924 หมายถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ดีถึง 92.4% โดยค่า MSE เท่ากับ 33,325,570,000.00 หมายถึงความคลาดเคลื่อนของค่าทำนายจากข้อมูลจริงในรูปแบบของค่ากำลังสอง ค่า RMSE เป็น 182,552.93 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยของการทำนายราคา ส่วนค่า MAE เท่ากับ 127,356.65 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยที่เกิดจากการทำนายราคา

แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย(averaging ensemble learning) มีค่า R2 ที่ 0.871 หมายถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ถึง 87.1% และค่า MSE เป็น 56,266,850,000.00 หมายถึงความคลาดเคลื่อนของค่าทำนายจากข้อมูลจริงในรูปแบบของค่ากำลังสอง ค่า RMSE เท่ากับ 237,206.35 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE เท่ากับ 140,085.86 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยที่เกิดจากการทำนายราคา

แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning) มีค่า R2 ที่ 0.883 หมายถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ถึง 88.3% และค่า MSE เท่ากับ 51,074,880,000.00 หมายถึงความคลาดเคลื่อนของค่าทำนายจากข้อมูลจริงในรูปแบบของค่ากำลังสอง ค่า RMSE เท่ากับ 225,997.52 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE เท่ากับ 152,669.31 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยที่เกิดจากการทำนายราคา

แบบจำลองการเรียนรู้กลุ่มแบบแบ็กกิ้ง (bagging ensemble learning) มีค่า R2 ที่ 0.832 หมายถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ถึง 83.2% และค่า MSE เท่ากับ 73,493,900,000.00 หมายถึงความคลาดเคลื่อนของค่าทำนายจากข้อมูลจริงในรูปแบบของค่ากำลัง

สอง ค่า RMSE เท่ากับ 271,097.58 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE เท่ากับ 155,363.36 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยที่เกิดจากการทำนายราคา

แบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) มีค่า R2 ที่ 0.846 หมายถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ถึง 84.6% และค่า MSE เท่ากับ 67,425,130,000.00 หมายถึงความคลาดเคลื่อนของค่าทำนายจากข้อมูลจริงในรูปแบบของค่ากำลังสอง ค่า RMSE เท่ากับ 259,663.50 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยของการทำนายราคา และค่า MAE เท่ากับ 147,838.86 แสดงถึงค่าคลาดเคลื่อนเฉลี่ยที่เกิดจากการทำนายราคา

ในการประเมินแบบจำลองที่ใช้ในการทำนายข้อมูล พบว่าแบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุดมีค่า R2 ที่สูงถึง 0.924 ซึ่งสูงที่สุดจากทั้งหมด แสดงถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้อย่างดีถึง 92.4% พร้อมกับค่า MSE ที่เป็น 33,325,570,000.00 และค่า RMSE ที่ 182,552.93 และค่า MAE ที่ 127,356.65 อีกทั้ง แบบจำลอง Averaging,สแต็กกิ้ง, แบ๊กกิ้ง , และบูสต์ติง ก็มีความสามารถในการทำนายที่น่าสนใจ ซึ่งช่วยให้เลือกแบบจำลองที่เหมาะสมสำหรับงานที่กำลังดำเนินการได้อย่างเหมาะสมและมีประสิทธิภาพครอบคลุม

ผลลัพธ์จากการประเมินแต่ละแบบจำลองที่เสนอนี้ นำเสนอข้อมูลเชิงสถิติเพื่อให้เห็นความแม่นยำและประสิทธิภาพในการทำนายผลลัพธ์ข้อมูล ช่วยในการเลือกแบบจำลองที่เหมาะสมสำหรับงานประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบนี้

ผลการเปรียบเทียบประสิทธิภาพของวิธีการใหม่ที่นำเสนอในการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบกับวิธีการดั้งเดิม

แบบจำลองได้ถูกพัฒนาเพื่อการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบและแบบจำลองได้เรียนรู้โดยใช้ชุดข้อมูลที่รู้ผลล่วงหน้าอยู่แล้ว ในส่วนนี้ได้มีการวิเคราะห์เปรียบเทียบประสิทธิภาพของอัลกอริทึมทั้งหมด นอกจากนี้ยังมีการปรับแต่งพารามิเตอร์เพื่อให้เกิดการวิเคราะห์และประเมินประสิทธิภาพที่ดีและสูงที่สุด สำหรับแบบจำลองนี้มีการทบทวนอย่างละเอียดรอบครอบเกี่ยวกับตัวชี้วัดประสิทธิภาพสำหรับแบบจำลอง ซึ่งประกอบด้วยแบบจำลองพื้นฐานจำนวน 5 แบบจำลอง และแบบจำลองการเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมอีกจำนวน 5 แบบจำลอง แล้วทำการวิเคราะห์และการประเมินเปรียบเทียบดังแสดงในตารางที่ 10

ในการวิจัยนี้เป้าหมายคือการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบใช้แบบจำลองการเรียนรู้ของเครื่องจักรที่แตกต่างกันทั้งสิ้น 10 แบบจำลอง โดยใช้การประเมินประสิทธิภาพของแบบจำลองด้วยตัวชี้วัดทางสถิติที่แตกต่างกัน คือ R2, MSE, RMSE และ MAE ตามที่รายงานไว้ในบทที่ 3 ซึ่งได้ผลลัพธ์ของแต่ละแบบจำลองดังแสดงในตาราง 9 แสดงความแม่นยำในการทำนายที่ได้รับจากแต่ละแบบจำลอง

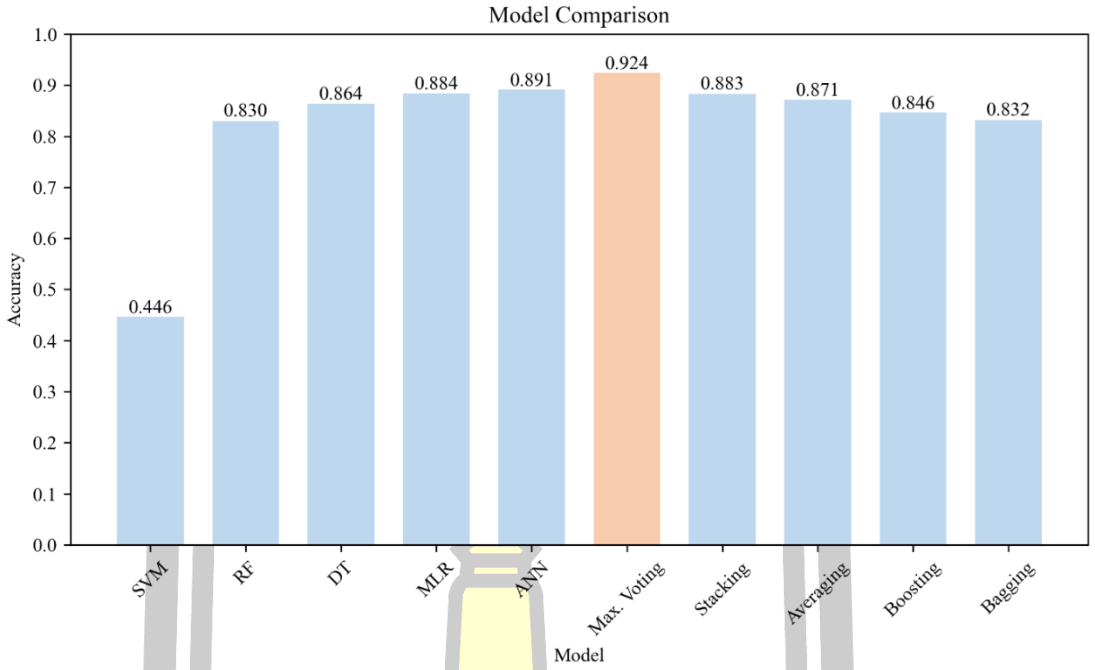
ตาราง 9 สรุปผลประสิทธิภาพของอัลกอริทึมที่ใช้ในการวิจัย

แบบจำลอง		R2	MSE	RMSE	MAE
วิธีการดั้งเดิม	ANN	0.891	47679301000.00	218355.89	167152.98
	SVM	0.446	242250000000.00	492189.90	322532.74
	MLR	0.884	50560550000.00	224856.74	146446.13
	DT	0.864	59522890000.00	243973.13	126901.04
	RF	0.830	74476350000.00	272903.56	164069.51
วิธีการใหม่	Max. Voting	0.924	33325570000.00	182552.93	127356.65
	Averaging	0.871	56266850000.00	237206.35	140085.86
	Stacking	0.883	51074880000.00	225997.52	152669.31
	Bagging	0.832	73493900000.00	271097.58	155363.36
	Boosting	0.846	67425130000.00	259663.50	147838.86

ในแบบจำลองประเภทแบบจำลองพื้นฐานอัลกอริทึมโครงข่ายประสาทเทียมเทียม (ANN) ขึ้นเป็นแบบจำลองพื้นฐานที่สำคัญด้วยความแม่นยำในการทำนายเท่ากับ 0.891 ซึ่งตามมาด้วยแบบจำลองการถดถอยเชิงเส้นหลายตัวแปร (MLR) ที่มีค่าความแม่นยำที่ดีรองลองมาคือ 0.884 และแบบจำลองต้นไม้การตัดสินใจ (DT) ที่ได้ความแม่นยำอยู่ที่ 0.864 แบบจำลองแรนดอมฟอเรสต์ (RF) มีความแม่นยำอยู่ที่ 0.830 ในขณะที่ซับพอร์ทเวกเตอร์แมซชีน (SVM) มีความแม่นยำที่ต่ำสุดที่ 0.446

สำหรับแบบจำลองการเรียนรู้กลุ่มโดยใช้เทคนิคต่าง ๆ เพื่อรวมการทำนายจากแบบจำลองพื้นฐานที่มีประสิทธิภาพไม่ดี พบว่าแบบจำลองแบบเรียนรู้ร่วมด้วยเทคนิคการโหวตสูงสุดมีค่าสูงสุด (Maximum voting ensemble learning) มีความแม่นยำที่สูงที่สุดเท่ากับ 0.924 มากกว่าทั้งแบบจำลองพื้นฐานและแบบจำลองการเรียนรู้ร่วมเทคนิคอื่น ๆ และรองลงมาคือการเรียนรู้กลุ่มแบบสแต็กกิ้ง ที่มีความแม่นยำเท่ากับ 0.883 โดยตามมาด้วยเทคนิคการเรียนรู้กลุ่มแบบเฉลี่ยที่ได้ค่าความแม่นยำเท่ากับ 0.871 แบบการเรียนรวมกลุ่มแบบบูสต์ติง ได้ความแม่นยำที่ 0.846 และการเรียนกลุ่มแบบ แบ๊กกิ้ง ได้ความแม่นยำที่ 0.832

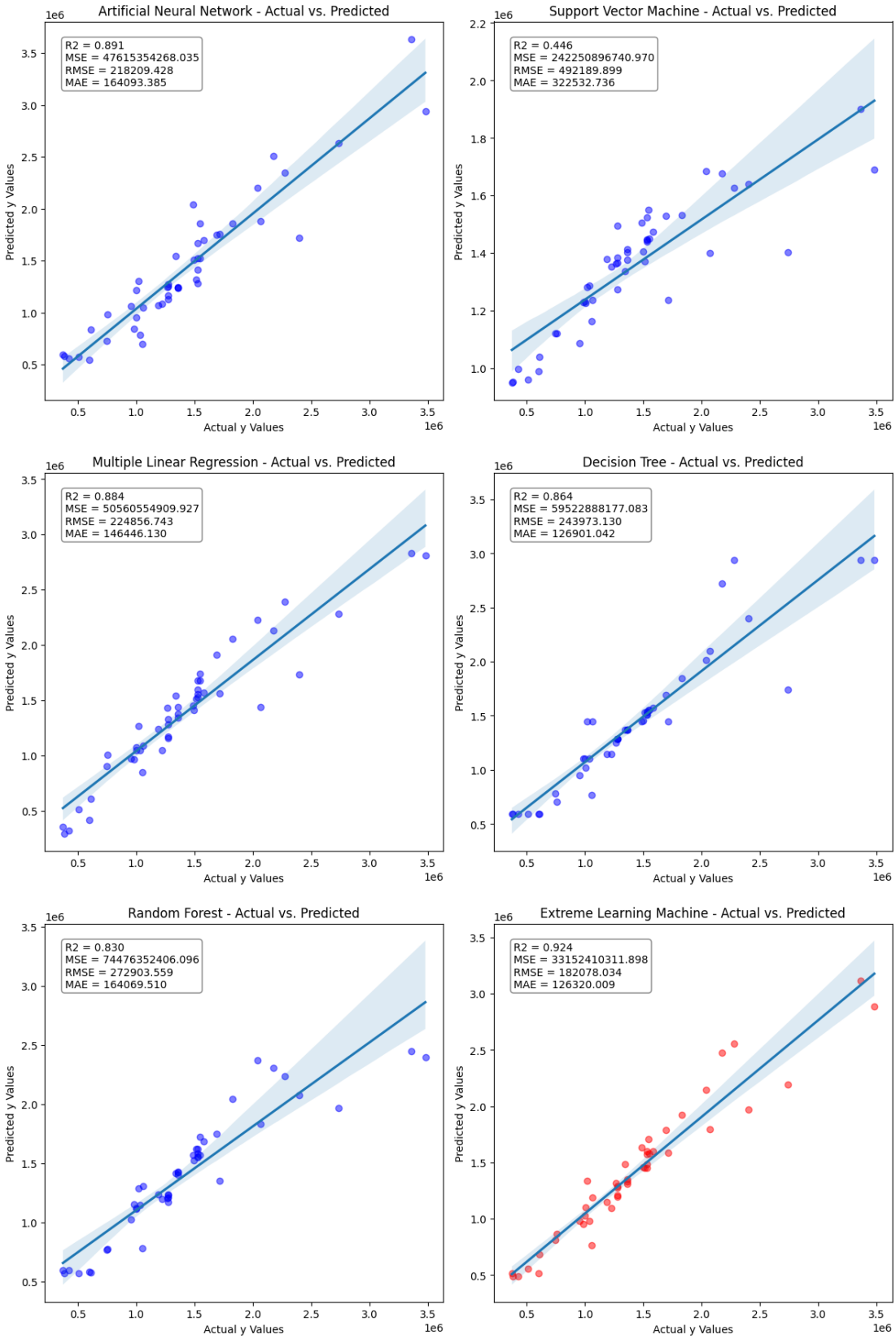
จากภาพประกอบ 29 แสดงผลกราฟเปรียบเทียบความแม่นยำของแบบจำลองทั้งหมดในรูปแบบที่สามารถมองเห็นได้อย่างชัดเจน ส่วนใหญ่ของอัลกอริทึมแสดงผลประสิทธิภาพที่ดีสำหรับทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ บ่งชี้ถึงศักยภาพในการทำนายราคาอย่างแม่นยำ และเมื่อเปรียบเทียบวิธีการใหม่และวิธีการดั้งเดิมพบว่าแนวโน้มการทำนายระหว่างค่าจริงกับค่าทำนายดังภาพประกอบ 30 ซึ่งวิธีการใหม่มีแนวโน้มความถูกต้องมากกว่าวิธีการดั้งเดิม



ภาพประกอบ 29 การเปรียบเทียบประสิทธิภาพของการทำนายราคาค่าก่อสร้าง

สรุปได้ว่างานวิจัยนี้ได้ใช้แบบจำลองการเรียนรู้ของเครื่องจักรที่เหมาะสม 10 รูปแบบของอัลกอริทึมและเทคนิคการเรียนรู้กลุ่มต่าง ๆ เพื่อทำนายราคาของบ้านพักอาศัยแนวราบ ผลการวิจัยเน้นให้เห็นถึงประสิทธิภาพของแบบการเรียนรู้กลุ่มแบบโหวตสูงสุด (Maximum voting ensemble learning) และแนะนำว่าเป็นตัวเลือกที่มีดีสำหรับการประยุกต์ใช้ในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ ช่วยให้วิศวกรและสถาปนิกประมาณการค่าก่อสร้างได้อย่างแม่นยำ นอกจากนี้แบบจำลองพื้นฐาน ANN, MLR, และ DT ยังมีความแม่นยำที่แข็งแกร่งซึ่งเป็นตัวเลือกที่คุ้มค่าสำหรับการพิจารณาด้วย

พหุน ปณฺ ฑิโต ชีเว



ภาพประกอบ 30 ผลการเปรียบเทียบวิธีการดั้งเดิมกับวิธีการใหม่

ผลการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลองในการการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบเพื่อหาวิธีการที่รวดเร็วและเหมาะสมที่สุด

จากวิธีการประมาณราคาค่าก่อสร้างอย่างหยาบที่มีค่าความคลาดเคลื่อนสูงเป็นเหตุให้เกิดความเสียหายจากการเสนอราคาที่ต่ำจนเกินไปทำให้ขาดทุนในที่สุด หรือเสนอราคาที่สูงจนเกินไปทำให้โอกาสในการได้งานมีโอกาสน้อย โดยบริษัทหรือห้างหุ้นส่วนจำกัดที่เข้าร่วมการประมูลงานหรือก่อสร้างงานบ้านพักอาศัยแนวรราบส่วนใหญ่เป็นบริษัทขนาดเล็กเพราะด้วยขนาดและตัวอาคารที่ไม่ใหญ่มากนักทำให้บริษัทขนาดกลางและขนาดใหญ่ไม่มีความคุ้มทุนในการมาร่วมประมูล และบริษัทขนาดเล็กก็มักจะไม่มีบุคลากรทางด้านวิศวกรรมเฉพาะด้านการถอดแบบและประมาณราคาหรือไม่มีบุคลากรทางด้านวิศวกรรมมากพอหรือไม่เวลามากพอในการประมาณราคาอย่างละเอียด บริษัทหรือห้างหุ้นส่วนจำกัดขนาดเล็กมักจะใช้วิธีการประมาณราคาค่าก่อสร้างอย่างหยาบในประมาณราคา และยื่นเสนอราคานั้นในการเข้าร่วมการประมูล ซึ่งส่งผลให้อาจเกิดการขาดทุนจากความคลาดเคลื่อนของวิธีการประมาณราคาโดยการประมาณราคาค่าก่อสร้างโดยใช้เครื่องจักรการเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมด้วยเทคนิคการเรียนรู้กลุ่มแบบโหวตสูงสุด (Maximum voting ensemble learning) เข้ามาเป็นตัวช่วย ทำให้สามารถควบคุมความคลาดเคลื่อนนั้นได้ โดยจากแบบจำลองการทำนายการประมาณราคาค่าก่อสร้างที่ได้เสนอนี้มีค่า R2= 0.924 ซึ่งหมายถึงความแม่นยำในการทำนายที่ร้อยละ92.4 ซึ่งเมื่อเทียบกับ กวี หวังนิเวศน์กุล (2556) ที่ได้นำเสนอไว้ว่าการประมาณราคาอย่างหยาบมีค่าความคลาดเคลื่อนอาจสูงถึงร้อยละ 50 และ สิทธิกร สิทธิการกูล (2564) ซึ่งได้นำเสนอวิธีการสร้างแบบจำลองด้วยโครงข่ายประสาทเทียมที่มีความแม่นยำในการทำนายที่ร้อยละ 91.4 และวิธีการสร้างแบบจำลองการถดถอยเชิงเส้นที่มีความแม่นยำในการทำนายร้อยละ 73.2 ซึ่งเมื่อนำเทคนิคการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบโดยใช้การเรียนรู้เครื่องจักรแบบเรียนกลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมจึงมีความแม่นยำสูงกว่าขึ้น อย่างไรก็ตาม งานวิจัยนี้ได้เลือกใช้เครื่องมือในการสร้างแบบจำลองด้วยภาษาไพทอนที่อาจจำเป็นต้องมีความรู้ด้านการเขียนซอร์สโค้ดแต่ไม่ได้มีความซับซ้อนมากนัก จึงสามารถที่จะนำไปใช้งานกับวิศวกรโยธา สถาปนิก หรือผู้ที่มีหน้าที่ในการประมาณราคาได้ ซึ่งจะเป็นประโยชน์ต่อการเผยแพร่เพื่อใช้งานต่อไป นอกจากนี้ ยังเป็นการพิสูจน์ถึงศักยภาพของการประยุกต์ใช้การเรียนรู้ของเครื่องจักรกับงานด้านการวิเคราะห์ข้อมูลในงานวิศวกรรมโยธาและบริหารงานก่อสร้างในประเทศไทย ซึ่งนอกจากจะทำให้เกิดประสิทธิภาพในการทำงานดังได้แสดงข้างต้นแล้วยังแสดงให้เห็นโอกาสในการทำงานด้านอื่นๆ อีกมากมาย

กระบวนการประมาณราคาที่ใช้กันทั่วไปหรือวิธีการดั้งเดิมนั้นเกี่ยวข้องกับการวิเคราะห์และประมาณราคาโดยผู้เชี่ยวชาญ เช่น วิศวกร สถาปนิก หรือบุคคลที่รับผิดชอบในการถอดแบบและประมาณราคา ได้แสดงให้เห็นว่าวิธีการดั้งเดิมนั้นมีทั้งข้อได้และข้อเสียในการประมาณราคา เช่น ถ้า

ประมาณราคาอย่างหยาบก็จะทำได้รวดเร็วแต่มีความคลาดเคลื่อนสูง แต่ถ้าประมาณราคาอย่างละเอียดจะมีความถูกต้องสูงแต่ใช้ระยะเวลาค่อนข้างนาน

ในการเปรียบเทียบกับวิธีการใหม่ที่ผู้วิจัยได้นำเสนอซึ่งคือการใช้เทคนิคการเรียนรู้กลุ่มแบบเครื่องจักรเรียนรู้เอ็กซ์ตรีม หรือ Extreme Learning Machine (XLM) ได้พิสูจน์ความแม่นยำในการประมาณราคาอย่างมีประสิทธิภาพและรวดเร็ว ซึ่งการวิเคราะห์ที่นำเสนอนี้เน้นใช้หลักการของเครื่องจักรเรียนรู้ที่เรียกว่า XLM เพื่อสร้างแบบจำลองที่เรียนรู้รูปแบบการประมาณราคาจากข้อมูลที่มีอยู่ ผลลัพธ์จากการทดสอบและการประเมินความแม่นยำของ XLM แสดงให้เห็นว่ามีความคลาดเคลื่อนเฉลี่ยอยู่ที่ประมาณร้อยละ 7.6 เมื่อเทียบกับผลการประมาณราคาจากวิธีการดั้งเดิม ดังภาพประกอบที่ 36 ถือเป็นระดับความคลาดเคลื่อนที่ยอมรับได้และในขอบเขตของการประมาณราคาแบบละเอียดอ้างอิงจากหลักการประมาณราคาของ กวี หวังนิเวศน์กุล (2556) ที่ควรจะน้อยกว่าร้อยละ 10

การใช้เทคนิค XLM ยังช่วยลดระยะเวลาในการดำเนินการถอดแบบและประมาณราคาอย่างมาก ทำให้ประหยัดเวลาและค่าใช้จ่ายในการเตรียมข้อมูลและวิเคราะห์ นอกจากนี้ เมื่อพิจารณาทางเลือกของแบบจำลองเพิ่มเติม คือ การใช้แบบจำลองเชิงเดี่ยวอย่างโครงข่ายประสาทเทียมก็พบว่าความคลาดเคลื่อนอยู่ในระดับประมาณร้อยละ 10.9 เมื่อเทียบกับการประมาณราคาแบบดั้งเดิม ดังภาพประกอบที่ 36 ซึ่งก็ยังคงเป็นทางเลือกที่น่าสนใจสำหรับการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบอีกด้วย

พหุน ปณ ฑิโต ชีเว

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

งานวิจัยเรื่องการทำนายการประมาณราคาการก่อสร้างบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม (The Prediction of Low-Rise Building Construction Cost Estimation Using Extreme Learning Machine) งานวิจัยนี้เป็นการศึกษาการสร้างแบบจำลองโดยใช้การเรียนรู้ของเครื่องจักรเพื่อช่วยในการประมาณราคาค่าก่อสร้างอย่างหยาบด้วย โดยทำการศึกษาการสร้างแบบจำลองการทำนายจากเทคนิค 2 กลุ่ม คือ แบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐาน และแบบจำลองการเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีม โดยมีวัตถุประสงค์ข้างต้น 3 วัตถุประสงค์ที่ได้แสดงไว้ในบทที่ 1 ซึ่งสามารถสรุปผลการวิจัยได้ตามวัตถุประสงค์ดังต่อไปนี้

สรุปผลและอภิปราย

ในการวิจัยนี้เพื่อเสนอวิธีการใหม่ในการประมาณราคาบ้านพักอาศัยแนวราบโดยใช้เครื่องจักรเรียนรู้เอ็กซ์ตรีม ได้นำเอาแบบจำลองการเรียนรู้ของเครื่องไปใช้งานทดแทนการคาดการณ์ราคาค่าก่อสร้างบ้านพักอาศัยแนวราบ จำนวน 10 อัลกอริทึมที่แตกต่างกันมาใช้งาน โดยแบ่งเป็นวิธีการดั้งเดิมหรืออัลกอริทึมพื้นฐาน 5 อัลกอริทึม และวิธีการใหม่ที่เป็นเครื่องจักรเรียนรู้เอ็กซ์ตรีมอีก 5 อัลกอริทึม โดยการค้นหาแบบจำลองที่มีประสิทธิภาพเหมาะสมที่สุดสำหรับการประมาณราคา ผลการวิเคราะห์ที่ละเอียดและรอบคอบถูกนำมาใช้ในการระบุแบบจำลองที่สำคัญที่สุดสำหรับแบบจำลองเครื่องจักรเรียนรู้พื้นฐานในวิธีการดั้งเดิมและแบบจำลองเครื่องจักรเรียนรู้เอ็กซ์ตรีมหรือวิธีการใหม่

สำหรับแบบจำลองเครื่องจักรเรียนรู้เอ็กซ์ตรีมหรือวิธีการใหม่ที่นำเสนอคือแบบจำลองการเรียนรู้กลุ่มที่มีค่าความแม่นยำสูงขึ้นและเกิดจากการรวมกลุ่มของแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐานหลายๆ ตัวรวมกัน สร้างเป็นแบบจำลองวิธีการใหม่ ประกอบด้วย แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (maximum voting ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย (averaging ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง (stacking ensemble learning) แบบจำลองการเรียนรู้กลุ่มแบบแบ๊กกิ้ง (bagging ensemble learning) และแบบจำลองการเรียนรู้กลุ่มแบบบูสต์ติง (boosting ensemble learning) พบว่าแบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด (MVE) ที่มีความแม่นยำสูงสุดในระดับ 0.924 มากกว่าแบบจำลองการเรียนรู้ของเครื่องจักรเทคนิคอื่นๆ หรือถึงของแบบจำลองการเรียนรู้กลุ่มด้วย แบบจำลองการเรียนรู้กลุ่มแบบสแต็กกิ้ง และการเรียนรู้กลุ่มแบบเฉลี่ยก็ได้แสดงความแม่นยำที่แข่งขันได้ด้วยคะแนนที่สูงใน

ระดับ 0.883 และ 0.871 ตามลำดับ ผลการทดลองเหล่านี้ชี้ให้เห็นว่าวิธีการการเรียนรู้กลุ่มเครื่องจักรเรียนรู้เอ็กซ์ตรีมสามารถเพิ่มประสิทธิภาพในการทำนายเกินความสามารถของแบบจำลองแต่ละตัวได้มาก

ผลการวิจัยนี้มีความสำคัญอย่างมากสำหรับอุตสาหกรรมการก่อสร้าง เนื่องจากการประเมินค่าใช้จ่ายอย่างแม่นยำเป็นสิ่งสำคัญในการรับรองความสำเร็จและความผลกำไรของโครงการ การใช้วิธีการเรียนรู้ของเครื่องจักรสามารถให้ความเข้าใจที่มีคุณค่าและช่วยลดความเสี่ยงของการเกิดค่าใช้จ่ายเกินกำหนดและความล่าช้า การใช้การเรียนรู้ของเครื่องจักรในการทำนายค่าใช้จ่ายการก่อสร้างมีความนิยมเพิ่มขึ้นในปีหลังๆ เนื่องจากความสามารถในการทำนายค่าได้อย่างแม่นยำและเชื่อถือได้ ในงานวิจัยนี้ เราได้สำรวจประสิทธิภาพของการเรียนรู้ของเครื่องจักรต่างๆในการทำนายค่าใช้จ่ายการก่อสร้างสำหรับบ้านพักอาศัยแนวราบในประเทศไทย

ผลการวิจัยของเราแสดงว่าแบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุดมีความเหมาะสมที่สุดในการประมาณราคาค่าการสร้างบ้านพักอาศัยแนวราบและทำนายได้อย่างรวดเร็วและประสิทธิภาพสูงกว่าอัลกอริทึมอื่นๆ ในเชิงความแม่นยำของการทำนาย เช่น R2, MAE, MSE และ RMSE นี่เป็นการแสดงว่าวิธีการเรียนรู้ของเครื่องจักรเรียนรู้ที่นำเสนอช่วยปรับปรุงความแม่นยำและความเชื่อถือได้ในการทำนาย โดยการรวมคุณสมบัติของแต่ละแบบจำลองเข้าด้วยกันและลดความเสี่ยงหรือ ความอ่อนไหวของแต่ละแบบจำลอง

เมื่อเปรียบเทียบประสิทธิภาพวิธีการใหม่หรือแบบจำลองเครื่องจักรเรียนรู้เอ็กซ์ตรีมในการทำนายการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบกับวิธีการดั้งเดิม พบว่าประสิทธิภาพของแต่ละแบบจำลองการเรียนรู้ของเครื่องจักรพื้นฐานในการทำนายราคาค่าก่อสร้างบ้านพักอาศัยแบบแนวราบพบว่าค่า R-squared สูงสุดคือโครงข่ายประสาทเทียมที่มีค่าอยู่ที่ 0.891 ซึ่งแสดงถึงความสามารถในการอธิบายความเปลี่ยนแปลงของข้อมูลประมาณ 89.1% ซึ่งเป็นค่าที่สูง และค่า MSE, RMSE, และ MAE ที่ต่ำเป็นแสดงถึงระดับความคลาดเคลื่อนของการทำนายที่น้อยลง แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน ค่า R-squared อยู่ที่ 0.446 ซึ่งต่ำกว่า แสดงให้เห็นว่าความสามารถในการอธิบายความเปลี่ยนแปลงของข้อมูลเพียง 44.6% เป็นไปได้ แสดงถึงแม่นยำที่น้อยกว่าแบบจำลองอื่น ๆ ในแบบจำลองนี้ ANN (Artificial Neural Network), MLR (Multiple Linear Regression), DT (Decision Tree), และ RF (Random Forest) มีประสิทธิภาพที่ดีกว่า SVM (Support Vector Machine) ในการทำนายค่าก่อสร้างบ้านพักอาศัยแบบแนวราบ ด้วยการประเมินค่า R-squared และค่าคลาดเคลื่อนของการทำนายที่น้อยลง ส่วนวิธีการใหม่หรือเครื่องจักรเรียนรู้เอ็กซ์ตรีมแบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุดหรือ MVE มีค่า R-squared สูงถึง 0.924 ซึ่งเป็นระดับสูงสุด แสดงถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้อย่างดีถึง 92.4% พร้อมกับค่า MSE ที่อยู่ที่ 33,325,570,000.00 และค่า RMSE ที่อยู่ที่ 182,552.93 และค่า MAE ที่อยู่ที่

127,356.65 นอกจากนี้ แบบจำลอง Averaging,สแต็กกิ้ง, แบ๊กกิ้ง, และบูสต์ติง ก็มีความสามารถในการทำนายที่น่าสนใจ ซึ่งช่วยในการเลือกแบบจำลองที่เหมาะสมสำหรับงานที่กำลังดำเนินการอยู่

ผลลัพธ์จากกระบวนการประเมินแต่ละแบบจำลองที่ได้นำเสนอนี้ นำเสนอข้อมูลทางสถิติเพื่อให้เห็นความแม่นยำและประสิทธิภาพในการทำนายผลลัพธ์ข้อมูล นี่ช่วยในการเลือกแบบจำลองที่เหมาะสมสำหรับงานประมาณราคาค่าก่อสร้างบ้านพักอาศัยแบบแนวราบดังกล่าว

การศึกษานี้ยังเน้นให้เห็นถึงความสำคัญของการเลือกคุณสมบัติและการปรับแต่งพารามิเตอร์เพื่อเพิ่มความแม่นยำในการทำนาย โดยพบว่าการเลือกชุดคุณสมบัติที่เหมาะสมและการปรับแต่งพารามิเตอร์ของอัลกอริทึมเป็นสิ่งสำคัญในการเพิ่มประสิทธิภาพของการทำนาย โดยรวมแล้ว การวิจัยนี้แสดงให้เห็นถึงความเป็นไปได้ของวิธีการเรียนรู้ของเครื่องจักรในการทำนายค่าใช้จ่ายการก่อสร้างบ้านพักอาศัยแนวราบในประเทศไทย การใช้วิธีการเหล่านี้สามารถให้การทำนายที่แม่นยำและเชื่อถือได้ ซึ่งเป็นประโยชน์ต่ออุตสาหกรรมการก่อสร้างในเชิงลดความเสี่ยงที่เกี่ยวข้องกับการบริหารโครงการ

จากการเปรียบเทียบกับวิธีการใหม่ที่ผู้วิจัยได้นำเสนอซึ่งคือการใช้เทคนิคการเรียนรู้กลุ่มแบบเครื่องจักรเรียนรู้เอ็กซ์ตรีม หรือ Extreme Learning Machine (XLM) ได้พิสูจน์ความแม่นยำในการประมาณราคาอย่างมีประสิทธิภาพและรวดเร็ว ซึ่งการวิเคราะห์ที่นำเสนอนี้เน้นใช้หลักการของเครื่องจักรเรียนรู้ที่เรียกว่า XLM เพื่อสร้างแบบจำลองที่เรียนรู้รูปแบบการประมาณราคาจากข้อมูลที่มีอยู่ ผลลัพธ์จากการทดสอบและการประเมินความแม่นยำของ XLM แสดงให้เห็นว่ามีความคลาดเคลื่อนเฉลี่ยอยู่ที่ประมาณร้อยละ 7.6 เมื่อเทียบกับผลการประมาณราคาจากวิธีการดั้งเดิมถือเป็นระดับความคลาดเคลื่อนที่ยอมรับได้และในขอบเขตของการประมาณราคาแบบละเอียดอ้างอิงจากหลักการประมาณราคาของควรจะน้อยกว่าร้อยละ 10

การใช้เทคนิค XLM ยังช่วยลดระยะเวลาในการดำเนินการถอดแบบและประมาณราคาอย่างมาก ทำให้ประหยัดเวลาและค่าใช้จ่ายในการเตรียมข้อมูลและวิเคราะห์ นอกจากนี้ เมื่อพิจารณาทางเลือกของแบบจำลองเพิ่มเติม คือ การใช้แบบจำลองการเรียนรู้ของเครื่องกจักรพื้นฐานอย่างโครงข่ายประสาทเทียมก็พบว่าความคลาดเคลื่อนอยู่ในระดับประมาณร้อยละ 10.9 เมื่อเทียบกับการประมาณราคาแบบอย่างหยาบ (preliminary cost estimation) ก็ยังคงเป็นทางเลือกที่น่าสนใจสำหรับการประมาณราคาค่าก่อสร้างบ้านพักอาศัยแนวราบอีกด้วย

ข้อเสนอแนะ

จากผลการวิจัยนี้ มีข้อเสนอแนะสำหรับการวิจัยในอนาคตได้ดังนี้:

1. การใช้ชุดข้อมูลที่หลากหลายและกว้างขึ้น ถึงแม้ชุดข้อมูลที่ใช้ในการวิจัยนี้จะได้นำเสนอข้อมูลที่มีค่าได้อย่างมีประสิทธิภาพ แต่การวิจัยในอนาคตควรพิจารณาการใช้ชุดข้อมูลที่มีขนาดใหญ่และหลากหลายมากขึ้นเพื่อประเมินประสิทธิภาพของอัลกอริทึมในบริบทที่แตกต่างกัน เช่น ภูมิภาคและประเภทของการก่อสร้าง

2. การตัวแปรต้นที่แตกต่างกัน ถึงแม้ว่าการวิจัยนี้จะพบว่ามีตัวแปรที่สำคัญที่เชื่อมโยงกับค่าใช้จ่ายในการก่อสร้างอย่างมีนัยสำคัญ แต่การวิจัยในอนาคตควรสำรวจคุณลักษณะอื่นๆ เช่น ตำแหน่งที่ตั้ง ประเภทวัสดุการก่อสร้าง และค่าแรงแต่ละพื้นที่ เพื่อเพิ่มความแม่นยำในการประเมินค่าใช้จ่าย

3. การนำไปใช้กับประเภทการก่อสร้างอื่นๆ ถึงแม้ว่าการวิจัยนี้จะเน้นไปที่การก่อสร้างอาคารบ้านพักอาศัยแนวราบในประเทศไทย แต่การวิจัยในอนาคตยังสามารถสำรวจผลของอัลกอริทึมเหล่านี้กับประเภทการก่อสร้างอื่น ๆ เช่น อาคารชั้นสูง อสังหาริมทรัพย์พาณิชย์ และโครงการสาธารณูปโภคได้

4. เปรียบเทียบอัลกอริทึมการเรียนรู้ของเครื่องที่แตกต่างกัน: การศึกษานี้เน้นไปที่อัลกอริทึมการเรียนรู้ของเครื่อง 4 แบบ แต่ยังมีอัลกอริทึมอื่นๆ ที่สามารถใช้ในการทำนายต้นทุนการก่อสร้างได้ เช่น การเรียนรู้ที่มีการเข้ารหัส (Encrypted learning), การเรียนรู้แบบแบ่งกลุ่ม (clustering) หรือการเรียนรู้แบบโครงข่ายเชิงลึก (deep learning) การศึกษาในอนาคตสามารถเปรียบเทียบผลการทำงานของอัลกอริทึมที่แตกต่างกันและระบุว่ามีอัลกอริทึมใดที่เหมาะสมที่สุดสำหรับการทำนายต้นทุนการก่อสร้างในประเทศไทยโดยพิจารณาความเหมาะสมในเชิงประสิทธิภาพ และความซับซ้อนของการปรับแต่งอัลกอริทึมในการใช้งานจริง

พหุน ปณฺฑิโต ชีเว

บรรณานุกรม

กวี หวังนิเวศน์กุล (2556). การประมาณราคางานวิศวกรรมก่อสร้าง. ซีเอ็ดยูเคชั่น.

กองสถิติเศรษฐกิจ สำนักงานสถิติแห่งชาติ. (2564). การประมวลข้อมูลพื้นที่การก่อสร้าง 2564. การประมวลข้อมูลพื้นที่การก่อสร้าง.

คมสัน เวนานนท์. (2545). ช่างควบคุมงานก่อสร้าง. กรมพัฒนาฝีมือแรงงาน.

นงลักษณ์ วิรัชชัย. (2542). ความสัมพันธ์โครงสร้างเชิงเส้น *LISREL:* สถิติวิเคราะห์สำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์. โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

พงศ์สยาม กันจินะ. (2556). แบบจำลองการประมาณราคาอาคารพักอาศัยคอนกรีตเสริมเหล็กโดยวิเคราะห์ตามสัดส่วนชนิดพื้นที่ใช้สอย = *Functional area based cost estimation model of residential buildings*. เชียงใหม่ : บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่.

มณเฑียร รัตนศิริวงศ์วุฒิ, ศจีมาจ ณ วิเชียร, & มนต์ชัย เทียนทอง. (2553). ปัจจัยที่มีอิทธิพลต่อสมรรถนะนักเทคโนโลยีสารสนเทศโดยการวิเคราะห์องค์ประกอบเชิงยืนยันลำดับที่สอง. วารสารเทคโนโลยีสารสนเทศ, *6*(12).

วิศว์ ดวงแสงทอง (2559). การประมาณราคาก่อสร้างบ้านพักอาศัยโดยเทคนิคโครงข่ายประสาทเทียม มหาวิทยาลัยเชียงใหม่].

วิสูตร จิระดำเกิง. (2563). การประมาณราคาก่อสร้าง (*Construction Cost Estimation)* (5th ed.).

สำนักงานนโยบายและยุทธศาสตร์การค้า (2566). ดัชนีราคาวัสดุก่อสร้างของประเทศ เดือนกรกฎาคม *2566*. กองดัชนีเศรษฐกิจการค้า สำนักงานนโยบายและยุทธศาสตร์การค้า กระทรวงพาณิชย์ ,.

สิทธิกร สิทธิการกูล (2564). การประมาณราคาค่าก่อสร้างสำหรับอาคารงานราชการโดยใช้เทคนิคการสร้างแบบจำลองพยากรณ์ *Construction cost estimation for government building using prediction modeling techniques*. มหาวิทยาลัยเชียงใหม่. บัณฑิตวิทยาลัย.

สิทธิกร สิทธิการกูล, ชินพัฒน์ บัวชาติ, & ดำรงศักดิ์ รินชุมภู. (2564). การประมาณราคาค่าก่อสร้างอาคารประเภทงานราชการโดยใช้เทคนิคการสร้างแบบจำลองการถดถอยเชิงเส้น. การประชุมวิชาการวิศวกรรมโยธาแห่งชาติครั้งที่ *26*.

Ahtesham, M., Bawany, N. Z., & Fatima, K. (2020). House price prediction using machine learning algorithm-the case of Karachi city, Pakistan. 2020 21st International Arab Conference on Information Technology (ACIT),

Alshemosi, A. M. B., & Alsaad, H. S. H. (2017). Cost estimation process for construction residential projects by using multifactor linear regression technique. *Criterion*, *6*(6), 7.

Amin, M. (2017). Development of cost estimation model for residential building. *Int J CivStructEng Res*, *5*(1), 1-4.

Arafa, M., & Alqedra, M. (2011). Early stage cost estimation of buildings construction projects using artificial neural networks. *Journal of Artificial Intelligence*, *4*(1), 63-75.

Bala, K., Ahmad Bustani, S., & Shehu Waziri, B. (2014). A computer-based cost prediction model for institutional building projects in Nigeria: an artificial neural network approach. *Journal of Engineering, Design and Technology*, *12*(4), 519-530.

Bennett, J., & Ormerod, R. N. (1984). Simulation applied to construction projects. *Construction Management and Economics*, *2*(3), 225-263.

Bowen, P. A. (1993). *A communication-based approach to price modelling and price forecasting in the design phase of the traditional building procurement process in South Africa* University of Port Elizabeth].

Car-Pusic, D., Petruseva, S., Zileska Pancovska, V., & Zafirovski, Z. (2020). Neural network-based model for predicting preliminary construction cost as part of cost predicting system. *Advances in Civil engineering*, *2020*, 1-13.

Chandanshive, V., & Kambekar, A. R. (2019). Estimation of building construction cost using artificial neural networks. *Journal of Soft Computing in Civil Engineering*, *3*(1), 91-107.

Cheng, M.-Y., Tsai, H.-C., & Hsieh, W.-S. (2009). Web-based conceptual cost estimates for construction projects using Evolutionary Fuzzy Neural Inference Model. *Automation in Construction*, *18*(2), 164-172.

Cheng, M.-Y., Tsai, H.-C., & Sudjono, E. (2010). Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry. *Expert Systems with Applications*, *37*(6), 4224-4231.

Choong, W. C. (2018). *Statistical Analysis Of Housing Prices In Petaling District Using Linear Functional Model* UTAR].

Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & De Felice, F. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability*, *12*(2), 492.

Dimoski, M., & Pettersen, M. (2020). *Predicting housing prices with machine learning: a macroeconomic analysis of the Norwegian housing market*

Du, Z., & Li, B. (2017). Construction project cost estimation based on improved BP Neural Network. 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA),

El-Sawalhi, N. I., & Shehatto, O. (2014). A neural network model for building construction projects cost estimating. *Journal of Construction Engineering and Project Management*, *4*(4), 9-16.

El Mouna, L., Silkan, H., Haynf, Y., Nann, M. F., & Tekouabou, S. C. (2023). A Comparative Study of Urban House Price Prediction using Machine Learning Algorithms. E3S Web of Conferences,

Elfahham, Y. (2019). Estimation and prediction of construction cost index using neural networks, time series, and regression. *Alexandria Engineering Journal*, *58*(2), 499-506.

Feng, G. L., & Li, L. (2013). Application of genetic algorithm and neural network in construction cost estimate. *Advanced materials research*, *756*, 3194-3198.

Ferlan, N., Bastic, M., & Psunder, I. (2017). Influential factors on the market value of residential properties. *Engineering Economics*, *28*(2), 135-144.

Gao, G., Bao, Z., Cao, J., Qin, A. K., & Sellis, T. (2022). Location-centered house price prediction: A multi-task learning approach. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *13*(2), 1-25.

He, X., Yao, Y., Xia, H. W., Su, Y. W., & Liu, G. (2011). Cost estimation of construction project using fuzzy neural network model embedded with modified particle optimizer. *Advanced materials research*, *243*, 6296-6301.

Ho, W. K., Tang, B.-S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research*, *38*(1), 48-70.

Jafari, A., & Akhavian, R. (2019). Driving forces for the US residential housing price: a predictive analysis. *Built Environment Project and Asset Management*, *9*(4), 515-529.

Jafarzadeh, R., Ingham, J. M., Wilkinson, S., González, V., & Aghakouchak, A. A. (2014). Application of artificial neural network methodology for predicting seismic

retrofit construction costs. *Journal of Construction Engineering and Management*, *140*(2), 04013044.

Jeuschnigger, D. (2018). *Forecasting housing prices employing machine learning techniques* University of Graz, Austria].

Ji, S.-H., Ahn, J., Lee, H.-S., & Han, K. (2019). Cost estimation model using modified parameters for construction projects. *Advances in Civil engineering*, *2019*.

Ji, S.-H., Park, M., & Lee, H.-S. (2012). Case adaptation method of case-based reasoning for construction cost estimation in Korea. *Journal of Construction Engineering and Management*, *138*(1), 43-52.

Jiang, Q. (2020). Estimation of construction project building cost by back-propagation neural network. *Journal of Engineering, Design and Technology*, *18*(3), 601-609.

Jin, R., Cho, K., Hyun, C., & Son, M. (2012). MRA-based revised CBR model for cost prediction in the early stage of construction projects. *Expert Systems with Applications*, *39*(5), 5214-5222.

Kalliola, J., Kapociute-Dzikiene, J., & Damasevicius, R. (2021). Neural network hyperparameter optimization for prediction of real estate prices in Helsinki. *PeerJ computer science*, *7*, e444.

Kang, J., Lee, H. J., Jeong, S. H., Lee, H. S., & Oh, K. J. (2020). Developing a forecasting model for real estate auction prices using artificial intelligence. *Sustainability*, *12*(7), 2899.

Khalaf, T. Z., Çağlar, H., Çağlar, A., & Hanoon, A. N. (2020). Particle swarm optimization based approach for estimation of costs and duration of construction projects. *Civil Engineering Journal*, *6*(2), 384-401.

Kim, S., & Shim, J. H. (2014). Combining case-based reasoning with genetic algorithm optimization for preliminary cost estimation in construction industry. *Canadian Journal of Civil Engineering*, *41*(1), 65-73.

Kraus, M., Feuerriegel, S., & Oztekin, A. (2020). Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research*, *281*(3), 628-641.

Kumari, P., & Toshniwal, D. (2021). Deep learning models for solar irradiance forecasting: A comprehensive review. *Journal of Cleaner Production*, *318*, 128566.

Lim, W. T., Wang, L., Wang, Y., & Chang, Q. (2016). Housing price prediction using neural networks. 2016 12th International conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD),

Montero, J.-M., Mínguez, R., & Fernández-Avilés, G. (2018). Housing price prediction: parametric versus semi-parametric spatial hedonic models. *Journal of Geographical Systems*, *20*, 27-55.

Mu, J., Wu, F., & Zhang, A. (2014). Housing value forecasting based on machine learning methods. Abstract and Applied Analysis,

Nusrat Jahan, P., Anik, T., Md, K., Md Shanon, M., & Khadiza Tul, K. (2023). Short-Term Rainfall Prediction Using Supervised Machine Learning. *Advances in Technology Innovation*, *8*(2), 111-120. https://doi.org/10.46604/aiti.2023.8364

Owusu-Manu, D.-G., Edwards, D. J., Donkor-Hyiaman, K. A., Asiedu, R. O., Hosseini, M. R., & Obiri-Yeboah, E. (2019). Housing attributes and relative house prices in Ghana. *International Journal of Building Pathology and Adaptation*, *37*(5), 733-746.

Pacifico, L. D., & Ludermir, T. B. (2013). Evolutionary extreme learning machine based on particle swarm optimization and clustering strategies. The 2013 International Joint Conference on Neural Networks (IJCNN),

Park, U., Kang, Y., Lee, H., & Yun, S. (2022). A stacking heterogeneous ensemble learning method for the prediction of building construction project costs. *Applied Sciences*, *12*(19), 9729.

Pham, T., Le-Hong, T., & Tran, X. (2023). Efficient estimation and optimization of building costs using machine learning. *International Journal of Construction Management*, *23*(5), 909-921.

Rahimi, N., Park, S., Choi, W., Oh, B., Kim, S., Cho, Y.-h., Ahn, S., Chong, C., Kim, D., & Jin, C. (2023). A Comprehensive Review on Ensemble Solar Power Forecasting Algorithms. *Journal of Electrical Engineering & Technology*, *18*(2), 719-733.

Roxas, C. L. C., & Ongpeng, J. M. C. (2014). An artificial neural network approach to structural cost estimation of building projects in the Philippines. DLSU Research Congress,

Sakri, S. B., & Ali, Z. (2022). Analysis of the dimensionality issues in house price forecasting modeling. 2022 Fifth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU),

Sirithikon Sittikarnkul. (2021). *Construction Cost Estimation for Government Building Using Prediction Modeling Techniques* Chiang Mai University].

Skoyles, E. R. (1968). *Introducing Bills of Quantities:(operational Format)*. Building Research Station.

Soltani, A., Heydari, M., Aghaei, F., & Pettit, C. J. (2022). Housing price prediction incorporating spatio-temporal dependency into machine learning algorithms. *Cities*, *131*, 103941.

Son, H., Kim, C., & Kim, C. (2012). Hybrid principal component analysis and support vector machine model for predicting the cost performance of commercial building projects using pre-project planning variables. *Automation in Construction*, *27*, 60-66.

Srirutchataboon, G., Prasertthum, S., Chuangsuwanich, E., Pratanwanich, P. N., & Ratanamahatana, C. (2021). Stacking ensemble learning for housing price prediction: a case study in Thailand. 2021 13th International Conference on Knowledge and Smart Technology (KST),

Surawatchayotin, A., & Paireekreng, W. (2021). The predictive model of higher education guidance using integrated techniques for imbalanced data of learner groups. *Journal of Information Science and Technology*, *11*(1), 65-78.

Tatari, O., & Kucukvar, M. (2011). Cost premium prediction of certified green buildings: A neural network approach. *Building and Environment*, *46*(5), 1081-1086.

Vignola, F., Derocher, Z., Peterson, J., Vuilleumier, L., Félix, C., Gröbner, J., & Kouremeti, N. (2016). Effects of changing spectral radiation distribution on the performance of photodiode pyranometers. *Solar energy*, *129*, 224-235.

Vineeth, N., Ayyappa, M., & Bharathi, B. (2018). House price prediction using machine learning algorithms. Soft Computing Systems: Second International

Conference, ICSCS 2018, Kollam, India, April 19–20, 2018, Revised Selected Papers 2,

Wang, S., Zhu, J., Yin, Y., Wang, D., Cheng, T. E., & Wang, Y. (2021). Interpretable multi-modal stacking-based ensemble learning method for real estate appraisal. *IEEE Transactions on Multimedia*.

Wang, X. (2017). Application of fuzzy math in cost estimation of construction project. *Journal of Discrete Mathematical Sciences and Cryptography*, *20*(4), 805-816.

Wang, Y.-R., Yu, C.-Y., & Chan, H.-H. (2012). Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models. *International Journal of Project Management*, *30*(4), 470-478.

Wit Dangsangthong. (2016). *Residential building cost estimation using artificial neural network approach* Chiang Mai University].

Xu, Y., & Shu, Y. (2006). Evolutionary extreme learning machine–based on particle swarm optimization. International Symposium on Neural Networks,

Yang, L., Liang, Y., Zhu, Q., & Chu, X. (2021). Machine learning for inference: Using gradient boosting decision tree to assess non-linear effects of bus rapid transit on house prices. *Annals of GIS*, *27*(3), 273-284.

Zelaya, C. V. G. (2019). Towards explaining the effects of data preprocessing on machine learning. 2019 IEEE 35th international conference on data engineering (ICDE),

ภาคผนวก

ภาคผนวก ก

Source Code สำหรับสร้างแบบจำลองและวิเคราะห์แบบจำลองด้วยไพทอน

การวิเคราะห์สหสัมพันธ์เพียร์สันและการวิเคราะห์ปัจจัยองค์ประกอบ

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from factor_analyzer import FactorAnalyzer

# Load data
pd.read_csv("C:/Users/20044837/Desktop/python/houseprice_noprice.csv")
df_n = pd.read_csv("D:/Python/houseprice.csv")

df_n = df_n.drop(['NO'], axis=1)
df_n = df_n.replace('-', 0)
df.fillna(0, inplace=True)

# Calculate correlation matrix
correlation_matrix = df.corr()

# Perform factor analysis
n_factors = 3
factor_analyzer = FactorAnalyzer(n_factors, rotation=None)
factor_analyzer.fit(df)
factor_loadings = factor_analyzer.loadings_

# Display correlation matrix
print("Correlation Matrix:")
print(correlation_matrix)

# Display factor loadings
print("Factor Loadings:")
factor_loadings_df = pd.DataFrame(factor_loadings, index=df.columns,
columns=[f'Factor {i+1}' for i in range(n_factors)])
print(factor_loadings_df)

# Plot Correllogram using Seaborn
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', center=0)
plt.title('Correllogram')
plt.show()

# Plot Correllogram for Factor Loadings using Seaborn
plt.figure(figsize=(12, 8))
sns.heatmap(factor_loadings_df, annot=True, cmap='coolwarm', center=0)
plt.title('Factor Loadings Correllogram')
plt.show()

# Plot Pairwise Plot (Scatterplot Matrix)
sns.pairplot(df)
plt.suptitle("Pairwise Plot (Scatterplot Matrix)", y=1.02, size=16)
plt.show()
```

การสร้างและวิเคราะห์แบบจำลอง ANN, MLR, SVM, DT และ RF

```
# 1. Base Model
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from factor_analyzer import FactorAnalyzer

# Load data
pd.read_csv("C:/Users/20044837/Desktop/python/houseprice_noprice.csv")
df_n = pd.read_csv("D:/Python/houseprice.csv")

df_n = df_n.drop(['NO'], axis=1)
df_n = df_n.replace('-', 0)
df_n.fillna(0, inplace=True)

# Split data into training and testing sets
X = df.iloc[:, :-1].values
y = df.iloc[:, -1].values

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.40,
random_state=20)

# Feature scaling
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Initialize and train the models
ann_model = MLPRegressor(hidden_layer_sizes=(500, 1000, 500),
max_iter=1000, activation='relu', solver='adam')
svm_model = SVR(kernel='linear', degree=10, C=1000, gamma='scale')
mlr_model = LinearRegression()
dt_model = DecisionTreeRegressor(random_state=20)
rf_model = RandomForestRegressor(n_estimators=100, random_state=20)

models = [ann_model, svm_model, mlr_model, dt_model, rf_model]

for model in models:
    model.fit(X_train, y_train)

# Make predictions on the test set and evaluate the performance
performance_metrics = []
for model in models:
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    model_names = model.__class__.__name__
    #model_names = ['ANN', 'SVM', 'MLR', 'DT', 'RF']
    performance_metrics.append({
        'MODEL': model_names,
```

```
        'R2': r2,
        'MSE': mse,
        'RMSE': rmse,
        'MAE': mae,
    })

# Display performance metrics in a table
table1_base = pd.DataFrame(performance_metrics)
print(table1_base)

# Save the performance metrics to an Excel file
#df_metrics.to_excel('performance_metrics.xlsx', index=False)
```

การสร้างและวิเคราะห์แบบจำลองการเรียนรู้กลุ่มแบบโหวตสูงสุด

```
# 2. Maximum Voting Ensemble
models = [ann_model, svm_model, mlr_model, dt_model, rf_model]
model_names = ['ANN', 'SVM', 'MLR', 'DT', 'RF']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Sort models based on R2 scores in descending order
sorted_models = [model for _, model in sorted(zip(r2_scores, models),
reverse=True)]

# Select the top 3 models based on R2 scores
top_3_models = sorted_models[:3]

# Perform ensemble voting using the top 3 models
ensemble = VotingRegressor(estimators=[(name, model) for name, model in
zip(model_names, top_3_models)])
ensemble.fit(X_train, y_train)
ensemble_predictions = ensemble.predict(X_test)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
```

```
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Display the top 3 models based on R2 score
print("Top 3 Models based on R2 score:")
for model in top_3_models:
    print(f"{model_names[models.index(model)]}")

#print("----------------------------")
#print("Ensemble Model Performance Metrics:")
#print(f"Ensemble R-squared (R2): {ensemble_r2:.3f}")
#print(f"Ensemble Mean Squared Error (MSE): {ensemble_mse:.3f}")
#print(f"Ensemble Root Mean Squared Error (RMSE): {ensemble_rmse:.3f}")
#print(f"Ensemble Mean Absolute Error (MAE): {ensemble_mae:.3f}")

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

table2_max_voting = pd.DataFrame(metrics_data)

# Add a row for the maximum voting ensemble metrics
table2_max_voting.loc['Ensemble'] = ['MAX VOTING', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]

print("----------------------------")
print("Performance Metrics Table:")
print(table2_max_voting)
```

การสร้างและวิเคราะห์แบบจำลองการเรียนรู้กลุ่มแบบเฉลี่ย

```
# 3. Averaging Ensemble
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)
```

```
    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Perform ensemble voting using Averaging ensemble algorithm
ensemble_predictions = np.mean([model.predict(X_test) for model in models],
axis=0)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Create a DataFrame to store performance metrics for each model
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

table3 = pd.DataFrame(metrics_data)

# Add a row for the ensemble model metrics
table3.loc['Ensemble'] = ['AVERAGING', ensemble_r2, ensemble_mse,
ensemble_rmse, ensemble_mae]

print("Performance Metrics Table:")
print(table3)
```

การสร้างและวิเคราะห์แบบจำลองการเรียนรู้กลุ่มแบบ boosting

```
# 4. Boosting Ensemble
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)
```

```
    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'Model': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Perform ensemble voting using the Boosting ensemble algorithm
boosting_models = [AdaBoostRegressor(base_model, n_estimators=10,
random_state=20) for base_model in models]
for model, name in zip(boosting_models, model_names):
    model.fit(X_train, y_train)

ensemble_predictions = np.mean([model.predict(X_test) for model in
boosting_models], axis=0)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['BOOSTING', ensemble_r2, ensemble_mse,
ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)
```

การสร้างและวิเคราะห์แบบจำลองการเรียนรู้กลุ่มแบบ bagging

```
# 5. Bagging Ensemble
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
```

```
    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'Model': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Perform ensemble voting using the Bagging ensemble algorithm
bagging_models = [BaggingRegressor(base_model, n_estimators=10,
random_state=20) for base_model in models]
for model, name in zip(bagging_models, model_names):
    model.fit(X_train, y_train)

ensemble_predictions = np.mean([model.predict(X_test) for model in
bagging_models], axis=0)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['BAGGING', ensemble_r2, ensemble_mse,
ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)
```

การสร้างและวิเคราะห์แบบจำลองการเรียนรู้กลุ่มแบบ stacking

```
# 6.1 Stacking Ensemble (Meta model = MLR)
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []
```

```
# Train and evaluate each base model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Create the meta-model (Linear Regression) for stacking
meta_model = LinearRegression()

# Initialize the data to hold the predictions of base models
num_models = len(models)
num_samples = X_train.shape[0]
num_features = num_models
meta_features_train = np.zeros((num_samples, num_features))

# Perform KFold cross-validation to get the meta-features for the meta-
model
kf = KFold(n_splits=10, random_state=20, shuffle=True)
for train_index, val_index in kf.split(X_train):
    for i, model in enumerate(models):
        model.fit(X_train[train_index], y_train[train_index])
        y_pred_val = model.predict(X_train[val_index])
        meta_features_train[val_index, i] = y_pred_val

# Train the meta-model using the meta-features
meta_model.fit(meta_features_train, y_train)

# Get the predictions of base models on the test set for stacking
meta_features_test = np.zeros((X_test.shape[0], num_features))
for i, model in enumerate(models):
    y_pred_test = model.predict(X_test)
    meta_features_test[:, i] = y_pred_test

# Make predictions using the meta-model
ensemble_predictions = meta_model.predict(meta_features_test)
```

```
# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['STACKING MLR', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)

# 6.2 Stacking Ensemble (Meta model = RF)
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each base model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Create the meta-model (Random Forest Regression) for stacking
meta_model = RandomForestRegressor(n_estimators=100, random_state=20)

# Initialize the data to hold the predictions of base models
num_models = len(models)
num_samples = X_train.shape[0]
num_features = num_models
```

```
meta_features_train = np.zeros((num_samples, num_features))

# Perform KFold cross-validation to get the meta-features for the meta-
model
kf = KFold(n_splits=10, random_state=20, shuffle=True)
for train_index, val_index in kf.split(X_train):
    for i, model in enumerate(models):
        model.fit(X_train[train_index], y_train[train_index])
        y_pred_val = model.predict(X_train[val_index])
        meta_features_train[val_index, i] = y_pred_val

# Train the meta-model using the meta-features
meta_model.fit(meta_features_train, y_train)

# Get the predictions of base models on the test set for stacking
meta_features_test = np.zeros((X_test.shape[0], num_features))
for i, model in enumerate(models):
    y_pred_test = model.predict(X_test)
    meta_features_test[:, i] = y_pred_test

# Make predictions using the meta-model
ensemble_predictions = meta_model.predict(meta_features_test)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['STACKING RF', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)

# 6.3 Stacking Ensemble (Meta model = SVM)
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each base model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)
```

```
    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Create the meta-model (SVR) for stacking
meta_model = SVR(kernel='linear')

# Initialize the data to hold the predictions of base models
num_models = len(models)
num_samples = X_train.shape[0]
num_features = num_models
meta_features_train = np.zeros((num_samples, num_features))

# Perform KFold cross-validation to get the meta-features for the meta-
model
kf = KFold(n_splits=10, random_state=20, shuffle=True)
for train_index, val_index in kf.split(X_train):
    for i, model in enumerate(models):
        model.fit(X_train[train_index], y_train[train_index])
        y_pred_val = model.predict(X_train[val_index])
        meta_features_train[val_index, i] = y_pred_val

# Train the meta-model using the meta-features
meta_model.fit(meta_features_train, y_train)

# Get the predictions of base models on the test set for stacking
meta_features_test = np.zeros((X_test.shape[0], num_features))
for i, model in enumerate(models):
    y_pred_test = model.predict(X_test)
    meta_features_test[:, i] = y_pred_test

# Make predictions using the meta-model
ensemble_predictions = meta_model.predict(meta_features_test)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['STACKING SVM', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]
```

```
# Display the DataFrame containing all performance metrics
print(df_metrics)

# 6.4 Stacking Ensemble (Meta model = DT)
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each base model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Create the meta-model (Decision Tree Regressor) for stacking
meta_model = DecisionTreeRegressor(random_state=20)

# Initialize the data to hold the predictions of base models
num_models = len(models)
num_samples = X_train.shape[0]
num_features = num_models
meta_features_train = np.zeros((num_samples, num_features))

# Perform KFold cross-validation to get the meta-features for the meta-
model
kf = KFold(n_splits=10, random_state=20, shuffle=True)
for train_index, val_index in kf.split(X_train):
    for i, model in enumerate(models):
        model.fit(X_train[train_index], y_train[train_index])
        y_pred_val = model.predict(X_train[val_index])
        meta_features_train[val_index, i] = y_pred_val
```

```
# Train the meta-model using the meta-features
meta_model.fit(meta_features_train, y_train)

# Get the predictions of base models on the test set for stacking
meta_features_test = np.zeros((X_test.shape[0], num_features))
for i, model in enumerate(models):
    y_pred_test = model.predict(X_test)
    meta_features_test[:, i] = y_pred_test

# Make predictions using the meta-model
ensemble_predictions = meta_model.predict(meta_features_test)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['STACKING DT', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)

# 6.5 Stacking Ensemble (Meta model = ANN)
models = [mlr_model, svm_model, dt_model, rf_model, ann_model]
model_names = ['MLR', 'SVM', 'DT', 'RF', 'ANN']

# Create lists to store performance metrics for each model
r2_scores = []
mse_scores = []
rmse_scores = []
mae_scores = []

# Train and evaluate each base model
for model, name in zip(models, model_names):
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)

    r2_scores.append(r2)
    mse_scores.append(mse)
    rmse_scores.append(rmse)
    mae_scores.append(mae)

# Create a DataFrame to display performance metrics in a table
metrics_data = {
    'MODEL': model_names,
    'R2': r2_scores,
    'MSE': mse_scores,
```

```
    'RMSE': rmse_scores,
    'MAE': mae_scores
}

df_metrics = pd.DataFrame(metrics_data)

# Create the meta-model (MLP Regressor) for stacking
meta_model = MLPRegressor(hidden_layer_sizes=(500, 1000, 500),
max_iter=1000, activation='relu', solver='adam')

# Initialize the data to hold the predictions of base models
num_models = len(models)
num_samples = X_train.shape[0]
num_features = num_models
meta_features_train = np.zeros((num_samples, num_features))

# Perform KFold cross-validation to get the meta-features for the meta-
model
kf = KFold(n_splits=10, random_state=20, shuffle=True)
for train_index, val_index in kf.split(X_train):
    for i, model in enumerate(models):
        model.fit(X_train[train_index], y_train[train_index])
        y_pred_val = model.predict(X_train[val_index])
        meta_features_train[val_index, i] = y_pred_val

# Train the meta-model using the meta-features
meta_model.fit(meta_features_train, y_train)

# Get the predictions of base models on the test set for stacking
meta_features_test = np.zeros((X_test.shape[0], num_features))
for i, model in enumerate(models):
    y_pred_test = model.predict(X_test)
    meta_features_test[:, i] = y_pred_test

# Make predictions using the meta-model
ensemble_predictions = meta_model.predict(meta_features_test)

# Calculate ensemble performance metrics
ensemble_r2 = r2_score(y_test, ensemble_predictions)
ensemble_mse = mean_squared_error(y_test, ensemble_predictions)
ensemble_rmse = np.sqrt(ensemble_mse)
ensemble_mae = mean_absolute_error(y_test, ensemble_predictions)

# Add the ensemble model's metrics to the DataFrame
df_metrics.loc[len(df_metrics)] = ['STACKING ANN', ensemble_r2,
ensemble_mse, ensemble_rmse, ensemble_mae]

# Display the DataFrame containing all performance metrics
print(df_metrics)
```

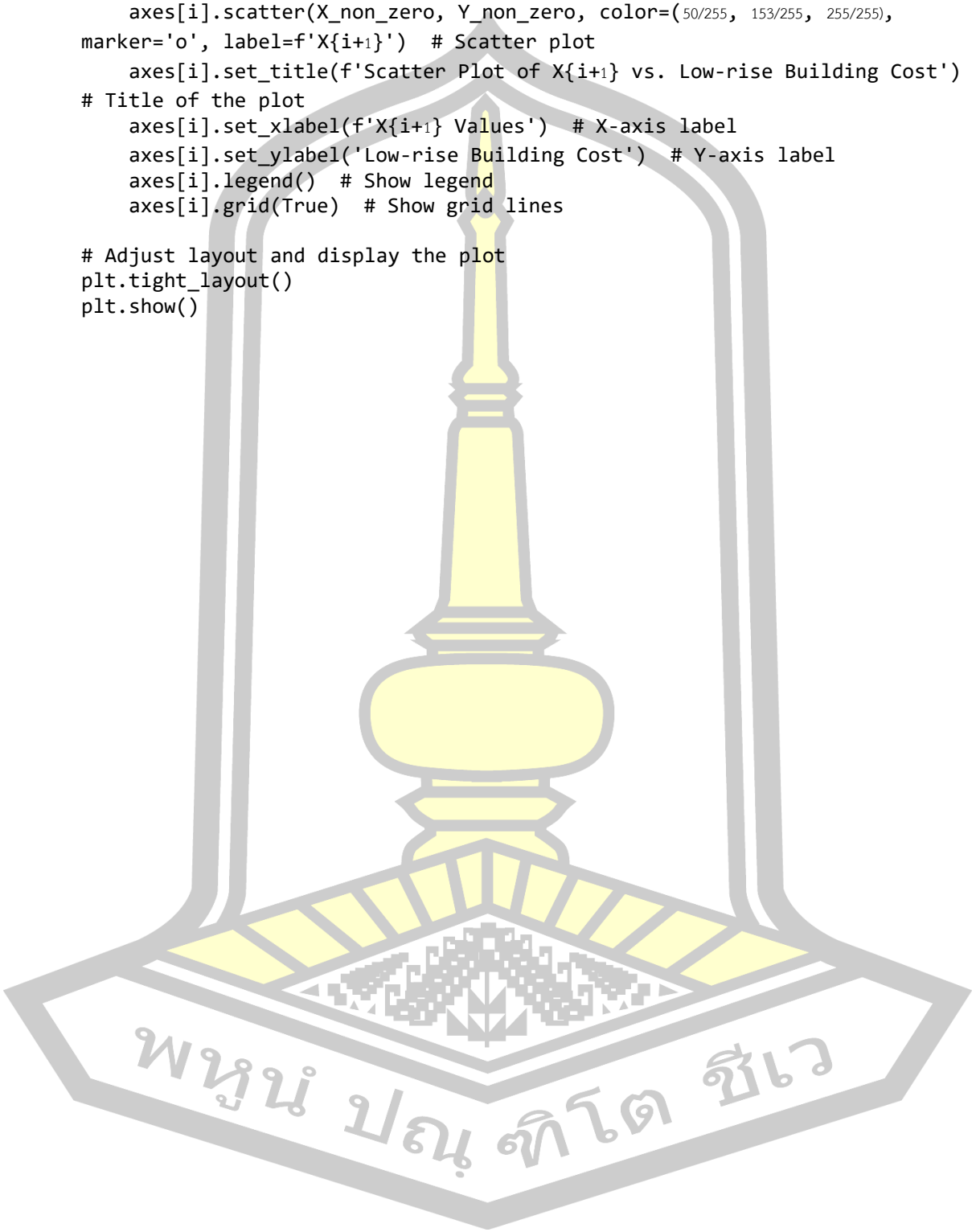
การทัศนภาพข้อมูล

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import seaborn as sns
```

```
# Load data
df = pd.read_csv("D:/ Python/houseprice.csv")
df = df.drop(['NO'], axis=1)

#df = df.dropna()
df = df.replace('-',0)
df.fillna(0, inplace=True)

# Extract X values (X1 to X15) and corresponding prices
x_columns = df.iloc[:, :-1] # Exclude the last column (Prices)
Y = df.iloc[:, -1] # Prices

num_x_values = x_columns.shape[1] # Number of X values (X1 to X15)

# Calculate statistics (max, min, mean, and standard deviation) for each X
variable
stats = []
for i in range(num_x_values):
    X = x_columns.iloc[:, i]  # Extract the ith X value
    non_zero_indices = X != 0  # Indices where X value is not 0

    X_non_zero = X[non_zero_indices]

    maximum = np.max(X_non_zero)
    minimum = np.min(X_non_zero)
    mean = np.mean(X_non_zero)
    std_dev = np.std(X_non_zero)

    stats.append((maximum, minimum, mean, std_dev))

# Create a DataFrame to display the statistics
stats_df = pd.DataFrame(stats, columns=['Min', 'Mean', 'Max', 'Std Dev'])

# Print the statistics DataFrame
print(stats_df)

# Export the tables to a CSV file
with open('Output55.csv', mode='w', newline='') as file:
    # Export table1
    stats_df.to_csv(file, index=False)

# Create box plots for each X variable
plt.figure(figsize=(12, 8))
sns.boxplot(data=x_columns)
#plt.xlabel('X Variables')
plt.ylabel('Values')
plt.title('Distribution of Values in X Variables')
plt.ylim(-45, 300) # Set y-axis limits between -20 and 300
plt.ticklabel_format(axis='y', style='plain', useOffset=False)  # Disable
scientific notation on y-axis

# Add mean, min, max, and std_dev values as text at the bottom of the plot
```

```
for i, (mean_value, min_value, max_value, std_dev_value) in
enumerate(zip(stats_df['Min'], stats_df['Mean'], stats_df['Max'],
stats_df['Std Dev'])):
    plt.text(i, -23,
f"{mean_value:.2f}\n{min_value:.2f}\n{max_value:.2f}\n{std_dev_value:.2f}",
             va='center', ha='center', color='black', fontsize=10,
linespacing=1.5, bbox=dict(facecolor='white', edgecolor='black',
boxstyle='round,pad=0.3'))

# Modify ylabel to include "mean:"
plt.ylabel('Values\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\nmean:',
rotation=0, labelpad=-20)

plt.tight_layout()
plt.show()

import matplotlib.pyplot as plt
import pandas as pd

# Assuming you have a DataFrame named 'df' with X1 to X15 columns and a
'Prices' column

# Extract X values (X1 to X15) and corresponding prices
x_columns = df.iloc[:, :-1].values  # Exclude the last column (Prices)
Y = df.iloc[:, -1].values  # Prices

num_x_values = x_columns.shape[1] # Number of X values (X1 to X15)

# Calculate the number of rows and columns for the subplots
num_rows = 5
num_cols = 3
num_plots = num_rows * num_cols

# Create a grid of subplots
fig, axes = plt.subplots(num_rows, num_cols, figsize=(15, 20))

# Flatten the axes array to iterate through it
axes = axes.flatten()

# Define a list of colors for the scatter plots
colors = ['blue', 'green', 'red', 'purple', 'orange', 'cyan', 'magenta',
'yellow', 'black', 'gray', 'brown', 'pink', 'lime', 'teal', 'indigo']

# Create scatter plots for each X value
for i in range(num_x_values):
    X = x_columns[:, i]  # Extract the ith X value
    non_zero_indices = X != 0  # Indices where X value is not 0

    X_non_zero = X[non_zero_indices]
    Y_non_zero = Y[non_zero_indices]

    row = i // num_cols
    col = i % num_cols
```

```
    # Create a scatter plot in the current subplot with a different color
    axes[i].scatter(X_non_zero, Y_non_zero, color=(50/255, 153/255, 255/255),
marker='o', label=f'X{i+1}')  # Scatter plot
    axes[i].set_title(f'Scatter Plot of X{i+1} vs. Low-rise Building Cost')
# Title of the plot
    axes[i].set_xlabel(f'X{i+1} Values')  # X-axis label
    axes[i].set_ylabel('Low-rise Building Cost')  # Y-axis label
    axes[i].legend()  # Show legend
    axes[i].grid(True)  # Show grid lines

# Adjust layout and display the plot
plt.tight_layout()
plt.show()
```



ประวัติผู้เขียน

ชื่อ	นายกิตติศักดิ์ ลาทอง
วันเกิด	16 เมษายน 2528
สถานที่เกิด	อุดรธานี
สถานที่อยู่ปัจจุบัน	207 ซอยเฉลิมพระเกียรติ ร.9 ซอย48 แยก19 แขวงดอกไม้ เขตประเวศ จังหวัดกรุงเทพมหานคร 10250
ตำแหน่งหน้าที่การงาน	Assist. Lead Civil & Building
สถานที่ทำงานปัจจุบัน	บริษัท ซีทีซีไอ (ประเทศไทย) จำกัด 825 อาคารไพโรจน์กิจจาทาวเวอร์ ชั้นที่ 19 ถนนเทพรัตน กม.4 แขวงบางนาเหนือ เขตบางนา จังหวัดกรุงเทพมหานคร 10260
ประวัติการศึกษา	พ.ศ. 2551 ปริญญาตรี วิศวกรรมศาสตรบัณฑิต (วศ.บ.) สาขาวิศวกรรมโยธา มหาวิทยาลัยศรีนครินทรวิโรฒ
	พ.ศ. 2555 ปริญญาโท รัฐประศาสนศาสตรมหาบัณฑิต (รป.ม.) สาขารัฐประศาสนศาสตร์ มหาวิทยาลัยรามคำแหง
	พ.ศ. 2557 ปริญญาตรี ปริญญาสาธารณสุขศาสตรบัณฑิต (ส.บ.) สาขาอาชีวอนามัยและความปลอดภัย มหาวิทยาลัยสุโขทัยธรรมาธิราช
	พ.ศ. 2558 ปริญญาโท วิศวกรรมศาสตรมหาบัณฑิต (วศ.ม.) สาขาวิศวกรรมโยธา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
	พ.ศ. 2566 ปริญญาปรัชญาดุษฎีบัณฑิต (ปร.ด.) สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล มหาวิทยาลัยมหาสารคาม
ทุนวิจัย	ทุนสนับสนุนการวิจัย มหาวิทยาลัยมหาสารคาม ประจำปีงบประมาณ 2566 ตามสัญญาเลขที่ 6603024 จำนวน 70,000 บาท
ผลงานวิจัย	-

พหูน ปณ ฑิโต ชีเว