

บทที่ 2

แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงการเรื่อง การพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในประเทศไทย มีความสำคัญต่อการบริหารจัดการด้านเกษตรและเศรษฐกิจ เนื่องจากราคาข้าวโพดเลี้ยงสัตว์มีความผันผวนสูงและส่งผลกระทบต่อเกษตรกร ผู้ประกอบการอุตสาหกรรมอาหารสัตว์ และผู้กำหนดนโยบาย ดังนั้นการศึกษานี้จึงจำเป็นต้องมีการทบทวนแนวคิด ทฤษฎี เครื่องมือ และงานวิจัยที่เกี่ยวข้อง เพื่อเป็นพื้นฐานสำหรับการสร้างโมเดลการพยากรณ์ที่เหมาะสมและมีประสิทธิภาพโดยได้รวบรวมการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

- 2.1.1 แนวคิดเกี่ยวกับการพยากรณ์ (Forecasting Concept)
- 2.1.2 แนวคิดเกี่ยวกับราคาสินค้าเกษตร
- 2.1.3 แนวคิดการวิเคราะห์อนุกรมเวลา (Time Series Analysis)
- 2.1.4 แนวคิดเกี่ยวกับการสร้าง Visualization

2.2 ทฤษฎีที่เกี่ยวข้อง

- 2.2.1 ทฤษฎีการวิเคราะห์อนุกรมเวลา (Time Series Analysis Theory)
- 2.2.2 ทฤษฎีการวิเคราะห์การถดถอย (Regression Analysis Theory)
- 2.2.3 ทฤษฎีการเรียนรู้ของเครื่อง (Machine Learning Theory)
- 2.2.4 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์
- 2.2.5 ทฤษฎีเกี่ยวกับเครื่องมือและเทคโนโลยีที่ใช้ในการพัฒนา

2.3 วรรณกรรมที่เกี่ยวข้อง

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

- 2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)
- 2.4.2 แบบจำลองอนุกรมเวลา (Time Series)

2.5 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการพยากรณ์ (Forecasting Concept)

1. การพยากรณ์ (Forecasting) หมายถึงการคาดการณ์ถึงสิ่งใดสิ่งหนึ่งที่จะเกิดขึ้นในช่วงเวลาอนาคต โดยนำค่าพยากรณ์ที่ได้มาซึ่งประโยชน์ในการตัดสินใจ การพยากรณ์มีบทบาทสำคัญทั้งในหน่วยงานภาครัฐและเอกชน สำหรับภาคเกษตรกรรม การพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ถือเป็นเครื่องมือสำคัญที่จะช่วยให้เกษตรกร ผู้ประกอบการ และผู้บริหารภาครัฐ สามารถวางแผนการผลิต การตลาด และการบริหารทรัพยากรได้อย่างมีประสิทธิภาพ ช่วยลดความเสี่ยงจากความผันผวนของราคาข้าวโพด และสามารถเตรียมรับมือกับความไม่แน่นอนที่อาจเกิดขึ้นในอนาคต ความสำคัญของการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ การพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์มีบทบาทสำคัญในหลายด้าน ได้แก่

1. ช่วยวางแผนการผลิตเกษตรกรสามารถปรับปริมาณการปลูกข้าวโพดให้สอดคล้องกับราคาที่คาดการณ์ เพื่อให้ต้นทุนและผลตอบแทนมีความเหมาะสม
2. ช่วยในการบริหารทรัพยากรทำให้ทราบว่าควรจัดสรรทรัพยากร เช่น เมล็ดพันธุ์ ปุ๋ย แรงงาน และพื้นที่เพาะปลูกเท่าใด เพื่อให้สอดคล้องกับราคาที่คาดการณ์
3. ช่วยในการวางแผนการตลาดและการจำหน่ายผู้ประกอบการสามารถกำหนดเวลาจำหน่ายหรือขนส่งข้าวโพดให้ตรงกับช่วงราคาสูงสุด ลดการขาดทุนจากราคาต่ำ
4. ช่วยลดความเสี่ยงทางการเงินการพยากรณ์ราคาช่วยให้เกษตรกรและผู้ประกอบการสามารถวางแผนการขายหรือทำสัญญาซื้อขายล่วงหน้า เพื่อป้องกันความเสี่ยงจากความผันผวนของราคา
5. ช่วยในการวางแผนการลงทุนหากราคาคาดว่าจะสูงขึ้น ผู้ประกอบการสามารถลงทุนเพิ่ม เช่น ขยายพื้นที่เพาะปลูก หรือซื้อเครื่องจักรเพิ่ม เพื่อรองรับความต้องการในอนาคต

2. ประเภทของการพยากรณ์ การพยากรณ์สามารถแบ่งออกเป็น 2 ประเภทหลัก คือ

2.1 การพยากรณ์เชิงคุณภาพ (Qualitative Forecasting) เป็นการใช้ความคิดเห็นและประสบการณ์ของผู้เชี่ยวชาญ เช่น เกษตรกรรุ่นเก่า ผู้ประกอบการ หรือผู้เชี่ยวชาญด้านการตลาด มาช่วยคาดการณ์ราคา เหมาะสำหรับกรณีที่มีข้อมูลราคาหรือผลผลิตจำกัด เช่น การพยากรณ์ราคาสำหรับพันธุ์ข้าวโพดใหม่ หรือช่วงที่มีการเปลี่ยนแปลงนโยบายของรัฐเทคนิคที่นิยมใช้ เช่น Delphi Method การรวบรวมความคิดเห็นผู้เชี่ยวชาญ

หลายๆคนและปรับแนวคิดจนเกิดความเห็นร่วมScenario Analysis การสร้างสถานการณ์ต่างๆของตลาดเพื่อคาดการณ์ผลกระทบต่อราคา

2.2 การพยากรณ์เชิงปริมาณ (Quantitative Forecasting) ใช้ข้อมูลตัวเลขและสถิติจากอดีต เช่น ราคาข้าวโพดรายเดือน ปริมาณผลผลิต การส่งออก และข้อมูลเศรษฐกิจมาวิเคราะห์เพื่อคาดการณ์ราคาในอนาคตแบ่งออกเป็น 2 วิธีหลักการ

2.2.1 พยากรณ์ด้วยอนุกรมเวลา (Time Series Forecasting) สมมติว่าราคาข้าวโพดในอนาคตขึ้นอยู่กับราคาที่เกิดขึ้นในอดีต สามารถใช้เทคนิค เช่น ค่าเฉลี่ยเคลื่อนที่ (Moving Average), การถดถอยเชิงเส้น (Linear Regression), หรือ ARIMA

2.2.2 การพยากรณ์เชิงสาเหตุและตัวแปรพหุคูณ (Causal and Multivariate Forecasting) สมมติว่าการเปลี่ยนแปลงของราคาไม่ได้ขึ้นอยู่กับเวลาเพียงอย่างเดียว แต่ขึ้นอยู่กับปัจจัยภายนอก (Explanatory Variables) อื่น ๆ ที่มีความสัมพันธ์กัน เช่น ราคาน้ำมันดิบ, ปริมาณผลผลิต, สภาพอากาศ หรือราคาตลาดโลก วิธีการนี้จะนำปัจจัยเหล่านี้มาวิเคราะห์ร่วมกันเพื่อคาดการณ์ราคา โดยแบ่งเป็น 2 กลุ่มหลัก

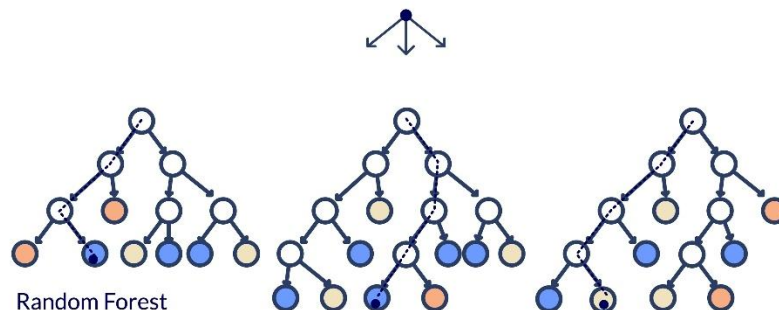
1. กลุ่มสถิติดั้งเดิม เช่น การวิเคราะห์การถดถอยพหุคูณ (Multiple Linear Regression) ซึ่งเหมาะกับข้อมูลที่มีความสัมพันธ์แบบเส้นตรง

2. กลุ่มวิธีการเรียนรู้ของเครื่อง (Machine Learning) เช่น Random Forest และ Support Vector Regression (SVR) ซึ่งเป็นการประยุกต์แนวคิดของ Support Vector Machine (SVM) สำหรับงานพยากรณ์ค่าเชิงปริมาณ โดยมีจุดเด่นในการจัดการกับข้อมูลที่มีความซับซ้อน และความสัมพันธ์แบบไม่เป็นเชิงเส้น (Non-linear Relationship) ได้ดีกว่าวิธีทางสถิติแบบดั้งเดิม

2.2.3 ทฤษฎีการเรียนรู้ของเครื่อง (Machine Learning Theory) การเรียนรู้ของเครื่อง (Machine Learning) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) ที่มุ่งเน้นการพัฒนาอัลกอริทึมให้สามารถเรียนรู้รูปแบบจากข้อมูล (Pattern Recognition) และทำการพยากรณ์หรือตัดสินใจได้โดยไม่ต้องเขียนโปรแกรมสั่งการแบบระบุเจาะจง สำหรับโครงการนี้ ผู้วิจัยเลือกใช้เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Learning) ประเภทการถดถอย (Regression) เพื่อพยากรณ์ราคาข้าวโพดซึ่งเป็นข้อมูลเชิงปริมาณที่มีความต่อเนื่อง โดยมีรายละเอียดทฤษฎีของแบบจำลองที่เลือกใช้ ดังนี้

1. Random Forest เป็นอัลกอริทึมแบบ Ensemble Learning ที่นำเสนอโดย Breiman (2001) ซึ่งทำงานโดยการสร้างต้นไม้ตัดสินใจ (Decision Trees) จำนวน

มากแล้วนำผลลัพธ์จากต้นไม้แต่ละต้นมาหาค่าเฉลี่ย (Averaging) เพื่อให้ได้ผลการพยากรณ์ที่แม่นยำ

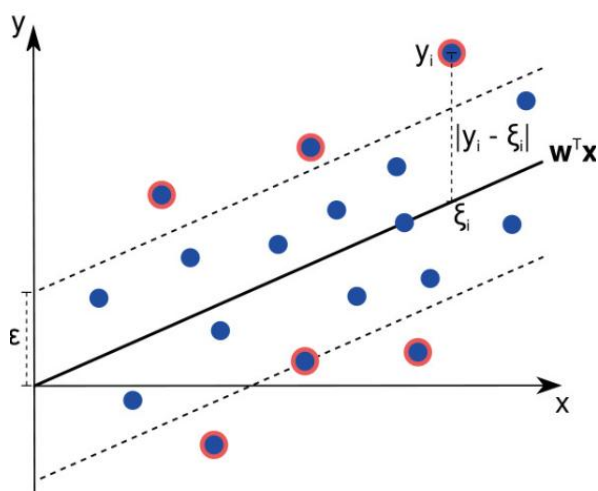


ภาพที่ 2.1 โครงสร้างการทำงานของแรนดอมฟอเรสต์ (Random Forest Structure)

หลักการสำคัญ อัลกอริทึมนี้ใช้วิธีการที่เรียกว่า Bagging (Bootstrap Aggregating) ในการสุ่มตัวอย่างข้อมูลเพื่อสร้างต้นไม้แต่ละต้น และมีการสุ่มเลือกคุณลักษณะ (Random Feature Selection) ในการแตกกิ่ง ทำให้ลดปัญหาการจดจำข้อมูลที่มากเกินไป (Overfitting) และช่วยลดความคลาดเคลื่อนในการพยากรณ์ได้อย่างมีประสิทธิภาพ

2. Support Vector Regression (SVR) เป็นอัลกอริทึมที่ประยุกต์มาจากอัลกอริทึม Support Vector Machine (SVM) โดยปรับเปลี่ยนเป้าหมายจากการจำแนกประเภทข้อมูล (Classification) มาเป็นการวิเคราะห์การถดถอย (Regression) เพื่อใช้ในการพยากรณ์ค่าข้อมูลที่มีความต่อเนื่อง (Continuous Values) เช่น การพยากรณ์ราคาสินค้า โดยมีพื้นฐานมาจากทฤษฎีการเรียนรู้ทางสถิติ (Vapnik, 1995) เป้าหมายหลักของ SVR คือการค้นหา "เส้นหรือระนาบแนวโน้มที่เหมาะสมที่สุด" (Optimal Hyperplane) ที่ครอบคลุมจุดข้อมูลส่วนใหญ่ให้อยู่ภายในขอบเขตความคลาดเคลื่อนที่ยอมรับได้ (เรียกว่า Epsilon-tube หรือ ϵ -tube) ให้ได้มากที่สุด โดยพยายามลดความคลาดเคลื่อนให้เหลือน้อยที่สุด ซึ่งจุดข้อมูลที่อยู่บนเส้นขอบหรืออยู่นอกขอบเขต จะถูกนำมาใช้เป็นตัวกำหนดระนาบ เรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vectors)

ฟังก์ชันเคอร์เนล (Kernel Function) เนื่องจากข้อมูลราคาสินค้าเกษตรมักมีความซับซ้อน ผันผวน และไม่มีความสัมพันธ์แบบเส้นตรงเสมอไป (Non-linear) โมเดล SVR จึงใช้เทคนิคที่เรียกว่า Kernel Trick เพื่อแปลงมิติของข้อมูลไปยังมิติที่สูงขึ้น (Higher-dimensional space) ทำให้โมเดลสามารถค้นหาความสัมพันธ์ที่ซับซ้อน และสร้างแบบจำลองที่มีความแม่นยำสูงในการทำนายราคาล่วงหน้าได้



ภาพที่ 2.2 แบบจำลองการวิเคราะห์การถดถอยซัพพอร์ตเวกเตอร์

(Support Vector Regression)

2.1.2 แนวคิดเกี่ยวกับราคาสินค้าเกษตร

ราคาสินค้าเกษตร คือ มูลค่าหรือจำนวนเงินที่ผู้ซื้อยินดีจ่ายเพื่อให้ได้สินค้าเกษตรหนึ่งหน่วยในช่วงเวลาหนึ่ง ราคาสินค้าเกษตรมีความผันผวนสูงและขึ้นอยู่กับปัจจัยหลายประการ ทั้งปัจจัยทางด้านกลไกอุปสงค์และอุปทาน (Demand and Supply) ปัจจัยด้านสภาพภูมิอากาศ นโยบายการแทรกแซงจากภาครัฐ และต้นทุนการผลิต

1. ปัจจัยที่มีผลต่อราคาสินค้าเกษตร

1.1 ปริมาณผลผลิต (Supply) ราคาสินค้าเกษตรมักขึ้นอยู่กับปริมาณผลผลิตที่เกษตรกรผลิตได้ หากผลผลิตมาก ราคาจะลดลง แต่หากผลผลิตน้อย ราคาจะสูงขึ้น

1.2 ความต้องการของตลาด (Demand) ความต้องการสินค้าเกษตรจากผู้บริโภค ผู้ประกอบการแปรรูป หรือภาคอุตสาหกรรมมีผลต่อราคาสินค้า หากความต้องการสูง ราคาจะสูงตาม

1.3 ต้นทุนการผลิต (Cost of Production) ราคาสินค้าเกษตรยังสะท้อนต้นทุนการผลิต เช่น ราคาน้ำมัน ค่าแรงงาน ค่าขนส่ง และค่าเมล็ดพันธุ์ หากต้นทุนสูง ราคาสินค้าเกษตรมักปรับตัวสูงตาม

1.4 สภาพอากาศและฤดูกาล (Climate and Seasonality) การเปลี่ยนแปลงของสภาพอากาศ เช่น ฝนตกชุกแล้ง หรืออุณหภูมิผิดปกติ สามารถส่งผลต่อผลผลิต และทำให้ราคาผันผวน

2. แนวคิดเกี่ยวกับการวิเคราะห์ราคา การวิเคราะห์ราคาสินค้าเกษตรมีความสำคัญต่อการวางแผนการผลิต การตลาด และการลงทุน เกษตรกรและผู้ประกอบการสามารถใช้ข้อมูลราคาย้อนหลัง มาวิเคราะห์แนวโน้มและคาดการณ์ราคาล่วงหน้า เพื่อปรับกลยุทธ์ให้เหมาะสม ตัวอย่างเช่น การใช้ อนุกรมเวลา (Time Series Analysis) หรือ การวิเคราะห์เชิงสาเหตุ (Causal Analysis) เพื่อคาดการณ์ราคาข้าวโพดเลี้ยงสัตว์ในอนาคต การเข้าใจแนวคิดเกี่ยวกับราคาสินค้าเกษตรช่วยให้เกษตรกรและผู้ประกอบการสามารถตัดสินใจได้อย่างมีประสิทธิภาพ เช่น เลือกลงเวลาขายให้ได้ราคาสูงสุด วางแผนปริมาณการผลิตให้สอดคล้องกับความต้องการของตลาด ประเมินความเสี่ยงจากความผันผวนของราคา

2.1.3 แนวคิดการวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นศาสตร์ที่มุ่งศึกษาข้อมูลซึ่งถูกจัดเก็บตามลำดับเวลา เช่น ข้อมูลรายวัน รายเดือน หรือรายปี โดยข้อมูลแต่ละจุดจะมีความสัมพันธ์เชื่อมโยงกับข้อมูลในอดีต ทำให้การวิเคราะห์อนุกรมเวลา มีความแตกต่างจากข้อมูลทั่วไปที่ไม่ขึ้นกับเวลา จุดมุ่งหมายของการวิเคราะห์อนุกรมเวลา คือ การค้นหารูปแบบ (Pattern) ที่ปรากฏในข้อมูล เพื่อนำไปสร้างแบบจำลองทางสถิติที่สามารถใช้ในการทำนายค่าในอนาคต (Forecasting) ได้อย่างมีประสิทธิภาพ (Box & Jenkins, 1976)

1. องค์ประกอบหลักของอนุกรมเวลา โดยทั่วไป การวิเคราะห์อนุกรมเวลาจะอธิบายข้อมูลผ่านองค์ประกอบหลัก 4 ส่วน (Makridakis et al., 1998)

1.1 แนวโน้ม (Trend, T) คือทิศทางของข้อมูลในระยะยาว เช่น ราคาข้าวโพดอาจมีแนวโน้มสูงขึ้นเมื่อเวลาผ่านไปเนื่องจากต้นทุนการผลิตเพิ่มขึ้น หรือความต้องการข้าวโพดสูงขึ้น

1.2 ฤดูกาล (Seasonality, S) คือรูปแบบความผันผวนที่เกิดซ้ำตามฤดูกาล เช่น ราคาข้าวโพดมักสูงขึ้นในช่วงปลายฤดูแล้งหรือต้นฤดูฝน เพราะผลผลิตน้อยลง

1.3 วัฏจักร (Cyclic, C) คือรูปแบบความผันผวนที่เกิดขึ้นเป็นรอบ ๆ แต่ไม่จำเป็นต้องเกิดซ้ำตามฤดูกาล เช่น ราคาข้าวโพดอาจเพิ่มขึ้นเป็นรอบ 3-5 ปี ราคาข้าวโพดอาจเพิ่มขึ้นเป็นรอบหลายปี ตามความต้องการในอุตสาหกรรม

1.4 ความผันผวนแบบสุ่ม (Irregular/Random, I) คือความเปลี่ยนแปลงที่เกิดขึ้นโดยไม่สามารถคาดการณ์ได้ เช่น ผลกระทบจากภัยธรรมชาติ หรือเหตุการณ์เศรษฐกิจโลกที่ส่งผลกระทบต่อราคา

2. เทคนิคการวิเคราะห์อนุกรมเวลา

2.1 ค่าเฉลี่ยเคลื่อนที่ (Moving Average, MA) วิธีนี้ใช้ค่าเฉลี่ยของข้อมูลย้อนหลังจำนวนหนึ่งมาเป็นค่าพยากรณ์ในอนาคต เพื่อให้ข้อมูลเรียบขึ้น เหมาะสำหรับการหาทิศทางโดยรวมของราคา แต่ไม่เหมาะหากข้อมูลมีฤดูกาลหรือความผันผวนสูง

2.2 การถดถอยเชิงเส้น (Linear Regression) วิธีนี้ใช้สมการเชิงเส้นระหว่างเวลา (t) และข้อมูลราคา เพื่อหาทิศทาง การเปลี่ยนแปลงระยะยาว เหมาะสำหรับการวิเคราะห์แนวโน้มโดยรวม แม้ไม่สามารถจับฤดูกาลได้ชัดเจน แต่ช่วยให้เห็นทิศทางการเปลี่ยนแปลงในอนาคต

2.3 แบบจำลอง ARIMA และ SARIMA ARIMA (Autoregressive Integrated Moving Average) เป็นแบบจำลองที่นิยมมากที่สุดสำหรับการวิเคราะห์อนุกรมเวลา โดยอาศัยสมการเชิงสถิติที่รวมเอาความสัมพันธ์เชิงเวลาและการทำให้ข้อมูลนิ่งเข้าด้วยกัน หากข้อมูลมีฤดูกาลที่ชัดเจน สามารถใช้ SARIMA (Seasonal ARIMA) ซึ่งเป็นการต่อยอดเพื่อตอบโจทย์ข้อมูลที่มี seasonality เช่น ราคาข้าวโพดเลี้ยงสัตว์ที่มีฤดูเก็บเกี่ยวและฤดูขาดแคลน

2.4 การทำให้เรียบแบบเอ็กซ์โพเนนเชียล (Exponential Smoothing / Holt-Winters) เป็นวิธีที่ให้น้ำหนักกับข้อมูลล่าสุดมากกว่าอดีต สามารถขยายไปสู่แบบ Holt หรือ Holt-Winters ซึ่งรองรับข้อมูลที่มีแนวโน้มและฤดูกาลได้ จึงมักถูกนำมาใช้ในการพยากรณ์ราคาสินค้าเกษตรรายเดือนที่มีความผันผวนตามฤดูกาล

2.1.4 แนวคิดเกี่ยวกับการสร้าง Visualization

การสร้าง Visualization หรือการนำเสนอข้อมูลด้วยภาพ (Data Visualization) เป็นเครื่องมือสำคัญที่ช่วยในการวิเคราะห์และสื่อสารข้อมูลเชิงสถิติและเชิงเวลา การนำเสนอข้อมูลในเชิงภาพช่วยให้ผู้วิเคราะห์และผู้ใช้ข้อมูลเข้าใจแนวโน้ม ความผันผวน และรูปแบบที่เกิดขึ้นจากข้อมูลได้ง่ายขึ้นกว่าการดูเพียงตัวเลขในตาราง (Few, 2012; Tufte, 2001) โดยเฉพาะอย่างยิ่งเมื่อทำการวิเคราะห์ข้อมูลอนุกรมเวลา (Time Series Analysis) ซึ่งเป็นข้อมูลที่มีการเปลี่ยนแปลงตามลำดับเวลา การใช้ Visualization ที่เหมาะสมจึงมีความจำเป็นต่อการตีความและการพยากรณ์ในอนาคต

1. ความสำคัญของ Visualization

1.1 ทำความเข้าใจข้อมูลที่ซับซ้อน ข้อมูลราคาข้าวโพดเลี้ยงสัตว์มีจำนวนมาก และมีความต่อเนื่องหลายปี การใช้กราฟ เช่น กราฟเส้น (Line Chart) ช่วยให้เห็นแนวโน้มและการเปลี่ยนแปลงของข้อมูลได้ดีกว่าตารางตัวเลข (Cleveland, 1993)

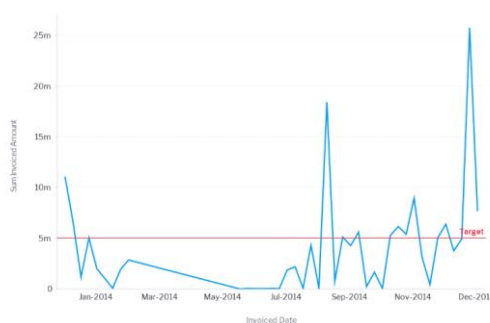
1.2 การค้นหาความสัมพันธ์และแนวโน้ม Visualization สามารถแสดงการเคลื่อนไหวของราคา เช่น การเพิ่มขึ้นหรือลดลง รวมถึงการใช้ค่าเฉลี่ยเคลื่อนที่ (Moving Average) หรือกราฟ Seasonal Plot เพื่อตรวจสอบรูปแบบฤดูกาล (Montgomery et al., 2015)

1.3 การตรวจจับความผิดปกติ กราฟกระจาย (Scatter Plot) หรือ Boxplot สามารถใช้ตรวจสอบค่าผิดปกติ เช่น ราคาที่สูงหรือต่ำเกินไป ซึ่งอาจสะท้อนถึงเหตุการณ์เศรษฐกิจหรือนโยบายภาครัฐ (Cole, 2016)

1.4 การสื่อสารและสนับสนุนการตัดสินใจ Visualization ช่วยแปลงผลลัพธ์เชิงสถิติที่ซับซ้อนให้เข้าใจง่าย ทำให้ผู้บริหาร เกษตรกร หรือผู้กำหนดนโยบายสามารถใช้ข้อมูลในการวางแผนได้อย่างมีประสิทธิภาพ (วรเดช, 2563)

2. รูปแบบ Visualization ที่เหมาะสมกับ Time Series

2.1 แผนภูมิแบบเส้น (Line Chart) แผนภูมิแบบเส้นเป็นกราฟที่ประกอบด้วยแกน x แสดงเวลา และแกน y แสดงค่าของตัวแปร เช่น ราคาข้าวโพดเลี้ยงสัตว์ในแต่ละเดือน จุดข้อมูลเรียงต่อเนื่องตามเวลาและเชื่อมด้วยเส้นตรง แผนภูมิแบบเส้นเหมาะสำหรับการแสดงแนวโน้มของข้อมูลต่อเนื่องตามเวลา เช่น ราคาข้าวโพดรายเดือนหรือรายวัน เพื่อให้เห็นการเปลี่ยนแปลงของราคาในช่วงเวลาต่าง ๆ ช่วยในการสังเกต จุดสูงสุด จุดต่ำสุด แนวโน้มระยะสั้นและระยะยาว



ภาพที่ 2.3 แผนภูมิแบบเส้น (Line Chart)

ที่มา : www.yellowfinbi.com/blog/how-to-choose-best-data-visualization-for-reporting

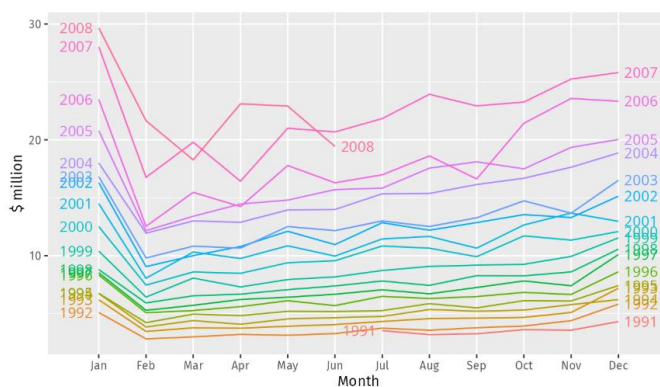
2.2 แผนภูมิพื้นที่ (Area Chart) แผนภูมิพื้นที่เป็นกราฟที่คล้ายกับแผนภูมิแบบเส้น แต่ด้านใต้เส้นถูกเติมสีหรือแรเงา เพื่อเน้น ปริมาณรวมหรือการสะสมของตัวแปร แผนภูมิประเภทนี้เหมาะสำหรับการเปรียบเทียบ การเปลี่ยนแปลงสะสม ของราคาหรือผลผลิต เช่น การเปลี่ยนแปลงราคาสะสมของข้าวโพดรายเดือน



ภาพที่ 2.4 แผนภูมิพื้นที่ (Area Chart)

ที่มา : www.yellowfinbi.com/blog/how-to-choose-best-data-visualization-for-reporting

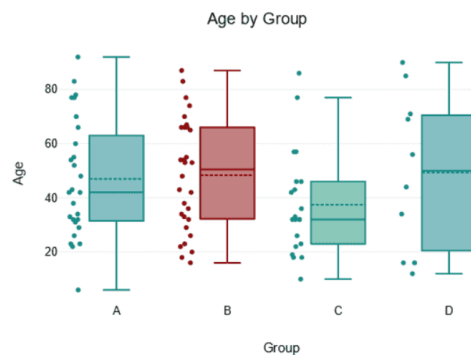
2.3 แผนภูมิฤดูกาล (Seasonal Plot) แผนภูมิฤดูกาลเป็นการแสดงข้อมูลแยกตาม เดือนหรือช่วงเวลาในปี เพื่อวิเคราะห์ รูปแบบฤดูกาลของข้อมูล โดยแต่ละเส้นในกราฟแทนข้อมูลของปีหนึ่ง แผนภูมิประเภทนี้ช่วยให้สังเกตเห็น ความแตกต่างของราคาข้ามเดือนและข้ามปี ได้ชัดเจน Seasonal Line Plot แสดงค่าเฉลี่ยของแต่ละเดือนเปรียบเทียบข้ามปี Seasonal Subseries Plot: แสดงแยกเดือนแต่ละปี เพื่อให้เห็นการเปลี่ยนแปลงภายในเดือน



ภาพที่ 2.5 แผนภูมิฤดูกาล (Seasonal Plot)

ที่มา : <https://otexts.com/fpp2/seasonal-plots.html>

2.4 แผนภูมิกล่อง (Boxplot) แผนภูมิกล่องเป็นกราฟที่ประกอบด้วย ค่ามัธยฐาน (Median), ควอร์ไทล์ (Quartiles), whiskers และ outliers เหมาะสำหรับตรวจสอบ การกระจายตัวและความแปรปรวนของข้อมูลตามช่วงเวลา เช่น เดือนหรือไตรมาส Boxplot แบบรายเดือน แสดงการกระจายราคาข้าวโพดในแต่ละเดือน Boxplot แบบรายไตรมาส แสดงความแปรปรวนตามฤดูกาล



ภาพที่ 2.6 แผนภูมิกล่อง (Boxplot)

ที่มา : <https://numiqo.com/tutorial/box-plot>

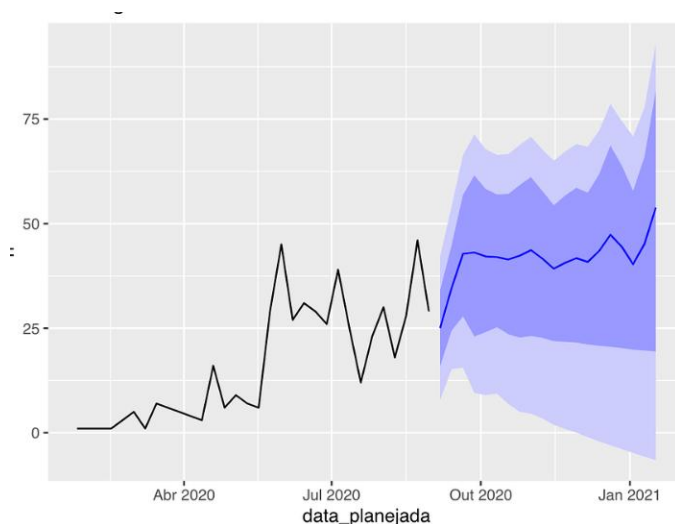
2.5 แผนภูมิความร้อน (Heatmap) แผนภูมิความร้อนเป็นการแสดง ค่าตัวแปร ในรูปของสี โดยแกน x แสดงเดือน แกน y แสดงปี และค่าของเซลล์แทนราคา เหมาะสำหรับการสังเกตรูปแบบ ฤดูกาลและค่าผันผวนข้ามปี Heatmap แสดงค่าเฉลี่ยรายเดือนของแต่ละปี ใช้สีแตกต่างกันเพื่อนำราคาสูงและต่ำ ดูว่าช่วงไหนของปีราคาข้าวโพดเลี้ยงสัตว์สูงหรือต่ำเป็นประจำและตรวจสอบความผิดปกติของราคาข้ามปี



ภาพที่ 2.7 แผนภูมิความร้อน (Heatmap)

ที่มา : www.researchgate.net/figure/This-heatmap-is-a-visualization-of-the-comparison-between-the-different-forecasting_fig4_369924476

2.6 แผนภูมิพยากรณ์ (Forecast Plot) แผนภูมิพยากรณ์เป็นการแสดงค่าพยากรณ์ควบคู่กับค่าจริง พร้อมช่วงความเชื่อมั่น (Confidence Interval) เพื่อให้เห็น ความแม่นยำของการพยากรณ์และแนวโน้มในอนาคต โครงสร้างได้แก่ เส้นราคาจริง (Observed) เส้นค่าพยากรณ์ (Forecast Mean) ช่วงความเชื่อมั่น (Shaded Area) ตัวอย่างการใช้งาน เช่น เปรียบเทียบราคาจริงกับค่าที่พยากรณ์ วิเคราะห์ช่วงเวลาที่มีความผันผวนสูง



ภาพที่ 2.8 แผนภูมิพยากรณ์ (Forecast Plot)

ที่มา : <https://forum.posit.co/t/best-way-to-present-forecast-plot/87798>

2.7 แดชบอร์ด (Dashboards) การรวบรวมข้อมูลจากแหล่งต่าง ๆ เพื่อนำมาจัดเรียงและสรุปเป็นภาพ โดยใช้แผนภูมิ กราฟ หรือรูปแบบ Visualization ต่าง ๆ เพื่อช่วยในการนำเสนอข้อมูล ปัจจุบัน แดชบอร์ด (Dashboard) ถือเป็นเครื่องมือ Data Visualization ที่นิยมใช้สำหรับการนำเสนอข้อมูลแบบ Real-time ผ่านซอฟต์แวร์หรือเครื่องมือจัดการและวิเคราะห์ข้อมูลต่าง ๆ เช่น เครื่องมือด้านการตลาด เครื่องมือบริหารจัดการข้อมูล เครื่องมือติดตามและดูแลเว็บไซต์ หรือเครื่องมือสำหรับวิเคราะห์ข้อมูลเชิงธุรกิจ เป็นต้น (เพ็ญศรีศราพรทิติจิตต์ โสภุชัน, 2563)



ภาพที่ 2.9 แดชบอร์ด (Dashboards)

ที่มา : <https://grafana.com/blog/2021/09/27/how-to-visualize-real-time-data-from-an-iot-smart-home-weather-station-with-grafana-dashboards/>

2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 ทฤษฎีการวิเคราะห์อนุกรมเวลา (Time Series Analysis Theory)

ทฤษฎีนี้เน้นการวิเคราะห์ชุดข้อมูลที่บันทึกตามลำดับเวลา (Time Series) เพื่อเข้าใจ รูปแบบการเปลี่ยนแปลง และใช้สำหรับการพยากรณ์อนาคต ข้อมูลอนุกรมเวลาประกอบด้วยองค์ประกอบสำคัญ 4 ส่วน (Montgomery, Jennings, & Kulahci, 2015) ได้แก่ 1. แนวโน้ม (Trend Component) แสดงทิศทางระยะยาวของข้อมูล เช่น ราคาข้าวโพดที่เพิ่มขึ้นหรือลดลงตามปี 2. ฤดูกาล (Seasonal Component) รูปแบบซ้ำในช่วงเวลาที่กำหนด เช่น ราคาสูงในเดือนเก็บเกี่ยวและต่ำในเดือนปลูก 3. รอบระยะยาว (Cyclical Component) การเปลี่ยนแปลงเป็นรอบหลายปี เช่น วัฏจักรเศรษฐกิจ 4. ความแปรปรวนแบบสุ่ม (Random/Irregular Component) การเปลี่ยนแปลงที่เกิดขึ้นโดยไม่คาดคิด เช่น ภัยธรรมชาติหรือปัจจัยเศรษฐกิจเฉพาะช่วง

แบบจำลอง SARIMA (Seasonal Autoregressive Integrated Moving Average) เป็นแบบจำลองทางสถิติที่ใช้สำหรับพยากรณ์ข้อมูลอนุกรมเวลาที่มีอิทธิพลของ "ฤดูกาล" เข้ามาเกี่ยวข้อง(เช่นราคา ผลผลิตทางการเกษตรที่จะขึ้นลงตามฤดูเก็บเกี่ยว) ซึ่งประกอบด้วยส่วน AR (Autoregressive), MA (Moving Average), การหาผลต่าง (Integration) และส่วนของฤดูกาล (Seasonal)

$$\Phi P(Bs)\phi p(B)(1-B)d(1-Bs)Dyt=\theta Q(Bs)\theta q(B)\epsilon t$$

$\Phi_P(B^s)$ = ส่วนพารามิเตอร์ Autoregressive แบบฤดูกาล (Seasonal AR)

$\varphi_p(B)$ = ส่วนพารามิเตอร์ Autoregressive แบบปกติ (Non-seasonal AR)

$(1-B)^d$ = การหาผลต่างในระดับปกติ (Non-seasonal Differencing)

$(1-B^s)^D$ = การหาผลต่างในระดับฤดูกาล (Seasonal Differencing)

y_t = ข้อมูลอนุกรมเวลา (ราคาข้าวโพด) ณ เวลา t

$\Theta_Q(B^s)$ = ส่วนพารามิเตอร์ Moving Average แบบฤดูกาล (Seasonal MA)

$\theta_q(B)$ = ส่วนพารามิเตอร์ Moving Average แบบปกติ (Non-seasonal MA)

ε_t = ค่าความคลาดเคลื่อนเชิงสุ่ม (White Noise Error) ณ เวลา t

2.2.2 ทฤษฎีการวิเคราะห์การถดถอย (Regression Analysis Theory)

การวิเคราะห์การถดถอยเป็นเทคนิคทางสถิติที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variables) และตัวแปรตาม (Dependent Variable) ในกรณีราคาข้าวโพดสามารถใช้ Regression เพื่อประเมินผลของปัจจัยต่าง ๆ เช่น ปริมาณผลผลิต อัตราแลกเปลี่ยน ราคาน้ำมัน หรือปัจจัยภายนอกอื่น ๆ ต่อราคา

1. ประเภทโมเดลที่นิยม

1.1 Linear Regression (การถดถอยเชิงเส้น) เป็นโมเดลพื้นฐานที่สุดใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ 1 ตัว (X) กับตัวแปรตาม 1 ตัว (Y) โดยสมมติว่าความสัมพันธ์เป็นเส้นตรง

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

พยากรณ์ราคาข้าวโพดตามปริมาณผลผลิตเพียงตัวเดียว เช่น ราคาข้าวโพดขึ้นกับผลผลิตรวมของประเทศ

1.2 Multiple Regression (การถดถอยเชิงพหุ) ขยายจาก Linear Regression โดยใช้ ตัวแปรอิสระหลายตัว เพื่ออธิบายตัวแปรตาม

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

ราคาข้าวโพด (Y) ขึ้นกับปัจจัยหลายอย่าง เช่น ปริมาณผลผลิต (X_1), ราคาน้ำมัน (X_2) และพื้นที่เพาะปลูก (X_3) สามารถวิเคราะห์ปัจจัยหลายตัวพร้อมกัน และควบคุมตัวแปรรอบวงนได้

Y คือ ตัวแปรตาม (Dependent Variable) เช่น ราคาข้าวโพด

X คือ ตัวแปรอิสระ (Independent Variable) เช่น ปริมาณผลผลิต

β_0 คือ ค่าจุดตัดแกน Y (Intercept)

β_1, β_2 คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficients)

ϵ คือ ค่าความคลาดเคลื่อน (Error Term)

2.2.3 ทฤษฎีการเรียนรู้ของเครื่อง (Machine Learning Theory) นอกจากสมการเชิงเส้น โครงการนี้ยังประยุกต์ใช้อัลกอริทึมการเรียนรู้ของเครื่องเพื่อจับความสัมพันธ์ที่ซับซ้อนได้แก่

1. Support Vector Regression (SVR) เป็นการประยุกต์อัลกอริทึม SVM มาใช้กับงานพยากรณ์ตัวเลขต่อเนื่อง เป้าหมายคือการสร้างระนาบ (Hyperplane) ที่ครอบคลุมจุดข้อมูลให้อยู่ในขอบเขตความคลาดเคลื่อนที่ยอมรับได้ และใช้ Kernel Function ในการแปลงมิติข้อมูลเพื่อจัดการกับข้อมูลที่ไม่เป็นเส้นตรง

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b$$

โดยที่ α_i, α_i^* คือตัวคูณลากรองจ์ และ $K(x_i, x)$ คือฟังก์ชันเคอร์เนล

2. Random Forest Regression เป็นเทคนิคการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) ที่สร้างต้นไม้ตัดสินใจ (Decision Trees) จำนวนหลายๆ ต้น แล้วนำผลลัพธ์จากการทำนายของทุกต้นมาหาค่าเฉลี่ย เพื่อลดปัญหาการเรียนรู้เกิน (Overfitting)

$$\hat{y} = \frac{1}{B} \sum_{b=1}^B T_b(x)$$

โดยที่ B คือจำนวนต้นไม้ทั้งหมด และ $T_b(x)$ คือผลการพยากรณ์จากต้นไม้ต้นที่ b

2.2.4 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์ สามารถออกแบบโครงสร้างได้หลายรูปแบบ โดยขึ้นอยู่กับเนื้อหาและกลุ่มเป้าหมายของผู้ใช้งาน การเลือกโครงสร้างที่เหมาะสมจะช่วยให้ผู้ใช้เข้าถึงข้อมูลสำคัญ เช่น แผนภูมิราคารายวัน ค่าพยากรณ์ และรายงานสรุป ได้อย่างสะดวก

โครงสร้างเรียงลำดับ (Sequential Structure) เป็นโครงสร้างแบบธรรมดาที่ใช้กันมาก เพราะเป็นโครงสร้างที่เข้าใจง่ายไม่ซับซ้อน มีการจัดลำดับการลิงค์ข้อมูลแต่ละหน้าเป็นแบบลำดับ ซึ่งทำให้ผู้ใช้จะต้องเข้าสู่เนื้อหาอื่นๆ ในลักษณะเส้นตรง โดยมีปุ่ม “เดินหน้า”

หรือ “ถอยหลัง” เป็นเครื่องมือหลักในการกำหนดทิศทางการเข้าสู่แต่ละหน้า เว็บไซต์ส่วนใหญ่ที่ใช้โครงสร้างแบบนี้ก็จะเป็นเว็บไซต์ประเภทองค์กรบริษัทขนาดเล็ก



ภาพที่ 2.10 โครงสร้างเรียงลำดับ (Sequential Structure)

ที่มา : <https://enfete.co.th/th/website-structure/>

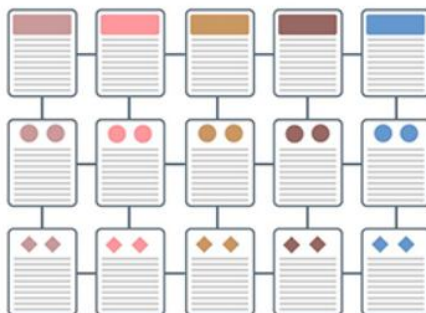
โครงสร้างแบบลำดับชั้น (Hierarchical Structure) เป็นโครงสร้างที่ดีที่สุด และเป็นที่นิยมใช้กันอย่างมาก เพราะเป็นโครงสร้างที่สามารถจัดการกับข้อมูลที่มีความซับซ้อนได้เป็นอย่างดี โดยลักษณะของโครงสร้างแบบลำดับชั้นนี้เป็นลักษณะแนวคิดเดียวกับการจัดแผนภูมิองค์กร มีการแบ่งเนื้อหาเป็นส่วนๆ แต่ละส่วนก็จะมีการนำเสนอรายละเอียดย่อย ๆ ที่ลดหลั่นกันมาจากบนลงล่าง โดยจะมีจุดรวมจุดเดียวกัน



ภาพที่ 2.11 โครงสร้างแบบลำดับชั้น (Hierarchical Structure)

ที่มา : <https://enfete.co.th/th/website-structure/>

โครงสร้างแบบตาราง (Matrix Structure) โครงสร้างแบบตารางเป็นโครงสร้างที่มีความซับซ้อนกว่าโครงสร้างแบบด้านบนออกแบบมาเพื่อเพิ่มความยืดหยุ่นให้ผู้ใช้เข้าสู่เนื้อหาได้ง่ายขึ้นโดยเพิ่มการเชื่อมโยงระหว่างเนื้อหาแต่ละส่วนเข้าหากันในปัจจุบันเว็บไซต์ส่วนใหญ่จะเลือกใช้โครงสร้างแบบตารางเพื่อความสะดวกและทันสมัยเพื่อผู้ใช้งานเว็บไซต์สามารถใช้งานเว็บไซต์สะดวกและเข้าสู่เนื้อหาต่างๆ ภายในเว็บด้วยตัวเองได้อย่างง่ายขึ้น



ภาพที่ 2.12 โครงสร้างแบบตาราง (Matrix Structure)

ที่มา : <https://enfete.co.th/th/website-structure/>

โครงสร้างแบบใยแมงมุม (Web / Network Structure) โครงสร้างแบบใยแมงมุมเป็นโครงสร้างที่มีความยืดหยุ่นมากที่สุด เพราะทุกหน้าบนเว็บไซต์สามารถเชื่อมโยงหากันได้ทั้งหมดทั้งภายในเว็บไซต์หรือแม้กระทั่งเชื่อมโยงไปสู่เว็บไซต์ภายนอกได้ ทำให้ผู้ใช้สามารถเลือกที่จะเข้าสู่เนื้อหาได้อย่างอิสระด้วยตนเอง.



ภาพที่ 2.13 โครงสร้างแบบใยแมงมุม (Web / Network Structure)

ที่มา : <https://enfete.co.th/th/website-structure/>

2.2.4.1 ส่วนประกอบหลักในหน้าเว็บไซต์ โครงงานพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ควรมีส่วนประกอบหลักที่ชัดเจน เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูลสำคัญได้อย่างสะดวกและเข้าใจง่าย โดยสามารถแบ่งออกเป็นองค์ประกอบดังนี้

Header (ส่วนหัวของเว็บไซต์) ประกอบด้วยโลโก้โครงการ เพื่อสร้างการจดจำและระบุเอกลักษณ์ของโครงการชื่อเว็บไซต์ ชัดเจน เช่น “ระบบพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในประเทศไทย” เมนูหลัก (Main Menu) เช่น Dashboard, ข้อมูลย้อนหลัง,

พยากรณ์ราคา เพื่อให้ผู้ใช้งานเข้าถึงหมวดหมู่สำคัญได้ทันทีที่ควรจัดวางให้ มองเห็นชัดเจนและไม่รกลายตา

Navigation(ระบบนำทาง)แถบเมนูที่ช่วยให้ผู้ใช้งานเลือกเข้าถึงหน้าต่างๆ ของเว็บไซต์ได้ง่ายตัวอย่างเมนู ได้แก่ Dashboard แสดงภาพรวมราคาข้าวโพดล่าสุดและค่าพยากรณ์ ข้อมูลย้อนหลังตารางและกราฟราคาข้าวโพดรายเดือนหรือรายปี พยากรณ์ราคาค่าพยากรณ์ราคาล่วงหน้า พร้อมช่วงความเชื่อมั่น การออกแบบเมนูควร กระชับ เข้าใจง่าย และรองรับการใช้งานเดสก์ท็อป

Content (เนื้อหาหลัก) พื้นที่แสดง ข้อมูลสำคัญของโครงการ ได้แก่ แผนภูมิ Line, Bar, Boxplot, Heatmap แสดงแนวโน้มและความผันผวนของราคาข้าวโพด ตารางข้อมูลย้อนหลัง แสดงราคาตามเดือน ปี หรือจังหวัด แผนที่ แสดงพื้นที่เพาะปลูกข้าวโพดในแต่ละจังหวัด เพื่อวิเคราะห์เชิงพื้นที่รายงานสรุปและข้อเสนอแนะ สำหรับเกษตรกรหรือผู้สนใจตลาดเนื้อหาควรจัดเรียงให้ เข้าใจง่าย สามารถอ่านและวิเคราะห์ข้อมูลได้ทันที

Search Function (ฟังก์ชันค้นหา) ช่วยให้ผู้ใช้งาน ค้นหาข้อมูลตามช่วงเวลา, จังหวัดหรือหมวดหมู่ข้อมูลที่น่าสนใจตัวอย่างเช่น ค้นหาราคาข้าวโพดเลี้ยงสัตว์ในจังหวัดเชียงใหม่ปี 2566 ค้นหาค่าพยากรณ์ราคาล่วงหน้า 3 เดือน ฟังก์ชันค้นหาควรตอบสนองเร็วและให้ผลลัพธ์แม่นยำ

Visual Content (เนื้อหาภาพและกราฟิก) ใช้กราฟและแผนภูมิเป็นเครื่องมือหลักในการสื่อสารข้อมูล เช่น Line Chart แสดงแนวโน้มราคาข้าวโพดตามช่วงเวลา Bar Chart เปรียบเทียบราคาหรือปริมาณผลผลิตระหว่างจังหวัด Boxplotตรวจสอบค่าผิดปกติหรือความผันผวนรายเดือน Heatmap วิเคราะห์ฤดูกาลหรือช่วงราคาที่สูง-ต่ำ การเลือกใช้กราฟต้อง เหมาะสมกับประเภทข้อมูลและวัตถุประสงค์ของผู้ใช้

2.2.4.2 ทฤษฎีสีในการออกแบบติดต่อประสานงานผู้ใช้ การเลือกใช้สีในการแสดงผลข้อมูลมีความสำคัญต่อการตีความและการรับรู้ข้อมูลโดยสามารถนำหลักการทฤษฎีสีมาประยุกต์ใช้กับกราฟและแดชบอร์ดของโครงการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ได้ดังนี้

Warm Colors (โทนสีร้อน) ประกอบด้วย สีแดง ส้ม เหลืองใช้สำหรับแสดงราคาสูงหรือแจ้งเตือนสถานการณ์สำคัญ เช่น ราคาข้าวโพดที่สูงเกินปกติ สีร้อนช่วยดึงดูดความสนใจและเน้นความสำคัญของข้อมูล (Few, 2013; Ware, 2013)

Cool Colors (โทนสีเย็น) ประกอบด้วย สีฟ้า เขียว ม่วงอ่อนใช้สำหรับแสดงราคาปกติหรือต่ำ สีเย็นให้ความรู้สึกสงบและเป็นกลาง เหมาะสำหรับการแสดงข้อมูลปกติที่ไม่ต้องการเน้นแจ้งเตือน (Few, 2013)



ภาพที่ 2.14 โทณสี (Colors)

ที่มา : www.jallearnmedia.com/blog/color-theory/

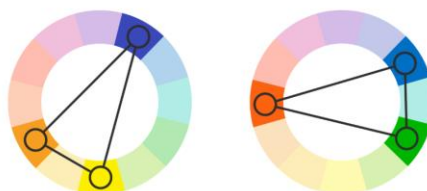
Complementary Colors (สีตรงข้าม) ใช้คู่สีที่อยู่ตรงข้ามบนวงล้อสี เช่น แดง-เขียว, น้ำเงิน-ส้ม ใช้สำหรับ เน้นความแตกต่างของราคาหรือแนวโน้ม เช่น การเปรียบเทียบราคาสูงสุดกับราคาต่ำสุดช่วยให้ผู้ใช้งานมองเห็นความแตกต่างได้ชัดเจน (Munzner, 2014)



ภาพที่ 2.15 Complementary Colors (สีตรงข้าม)

ที่มา : www.jallearnmedia.com/blog/color-theory/

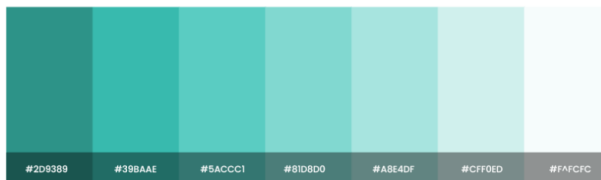
Analogous Colors (สีใกล้เคียงกัน) ประกอบด้วยสีที่อยู่ติดกันบนวงล้อสี เช่น ฟ้า-น้ำเงิน-เขียวใช้สำหรับ กราฟที่แสดงหลายตัวแปรพร้อมกัน เช่น การแสดงราคาข้าวโพดในหลายจังหวัดในกราฟเดียวสีใกล้เคียงช่วยให้ข้อมูลหลายชุดกลมกลืนและอ่านง่าย (Ware, 2013)



ภาพที่ 2.16 Analogous Colors (สีใกล้เคียงกัน)

ที่มา : www.jallearnmedia.com/blog/color-theory/

Monochromatic Colors (สีโทนเดียว)ใช้เฉดสีเดียวแต่มีความเข้ม-อ่อนต่างกันเหมาะสำหรับ แสดงค่าผลรวมหรือแนวโน้มรวม เช่น ราคากลางหรือราคาสะสมรายเดือนช่วยให้ผู้ใช้งานโฟกัสที่แนวโน้มโดยไม่ถูกดึงดูดด้วยสีหลายสี (Few, 2013; Munzner, 2014)



ภาพที่ 2.17 Monochromatic Colors (สีโทนเดียว)

ที่มา : <https://thewhitelabelagency.com/monochromatic-colors-guide/>

2.2.5 ทฤษฎีเกี่ยวกับเครื่องมือและเทคโนโลยีที่ใช้ในการพัฒนา ในการพัฒนาระบบพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในครั้งนี้ ผู้จัดทำได้เลือกใช้เครื่องมือและเทคโนโลยีที่ทันสมัยและเหมาะสมกับลักษณะงาน เพื่อให้ระบบทำงานได้อย่างมีประสิทธิภาพ โดยมีรายละเอียดของเครื่องมือแต่ละประเภท ดังนี้

1. โปรแกรม RapidMiner Studio คือแพลตฟอร์มซอฟต์แวร์สำหรับการทำวิทยาศาสตร์ข้อมูล (Data Science) ที่รองรับกระบวนการทำเหมืองข้อมูล (Data Mining) อย่างครบวงจร ตั้งแต่การเตรียมข้อมูล (Data Preparation), การเรียนรู้ของเครื่อง (Machine Learning), ไปจนถึงการวัดประสิทธิภาพโมเดล (Model Evaluation) ความสำคัญในโครงการใช้เป็นเครื่องมือหลักในการสร้างและทดสอบแบบจำลองพยากรณ์ราคาทั้ง 3 เทคนิค (Linear Regression, Random Forest, SVM) โดยอาศัยการทำงานแบบ Visual Workflow ที่เชื่อมต่อกับ Operator ต่างๆ เข้าด้วยกัน ทำให้สามารถวิเคราะห์ผลลัพธ์ได้อย่างแม่นยำ

2. โปรแกรม IBM SPSS Statistics เป็นซอฟต์แวร์สำเร็จรูปทางสถิติที่มีประสิทธิภาพสูงและได้รับความนิยมอย่างแพร่หลายในงานวิจัยและการวิเคราะห์ข้อมูล รองรับการจัดการข้อมูลขนาดใหญ่และการคำนวณทางสถิติที่ซับซ้อน ตั้งแต่สถิติพื้นฐานไปจนถึงสถิติขั้นสูง การวิเคราะห์สถิติเชิงพรรณนา (Descriptive Statistics) ใช้สำหรับวิเคราะห์ข้อมูลเบื้องต้นของราคาข้าวโพดและปัจจัยแวดล้อม เช่น การหาค่าเฉลี่ย (Mean), ส่วนเบี่ยงเบนมาตรฐาน (S.D.), ค่าสูงสุด-ต่ำสุด เพื่อทำความเข้าใจลักษณะการกระจายตัวของข้อมูล การวิเคราะห์ความสัมพันธ์ (Correlation Analysis) ใช้ตรวจสอบความสัมพันธ์ระหว่างตัวแปรต้น (เช่น ราคา น้ำมัน, ราคาปุ๋ย, ราคาต้นทุนสำหรับ) กับตัวแปรตาม (ราคาข้าวโพด) เพื่อคัดเลือกตัวแปรที่มี

นัยสำคัญทางสถิติ (Feature Selection) ก่อนนำไปเข้าสู่กระบวนการสร้างแบบจำลองพยากรณ์ในลำดับต่อไป

3.โปรแกรม XAMPP XAMPP คือชุดซอฟต์แวร์จำลองเซิร์ฟเวอร์ (Web Server Solution Stack) แบบ Open Source ที่รวบรวมเครื่องมือสำคัญสำหรับการพัฒนาเว็บไซต์ไว้ในที่เดียว ประกอบด้วย Apache ทำหน้าที่เป็นเว็บเซิร์ฟเวอร์ (Web Server) สำหรับประมวลผลและส่งข้อมูลหน้าเว็บไซต์ MySQL ทำหน้าที่เป็นระบบจัดการฐานข้อมูล (Database Management System) PHP ภาษาโปรแกรมสำหรับประมวลผลฝั่งเซิร์ฟเวอร์ ใช้สำหรับจำลองเครื่องคอมพิวเตอร์ส่วนบุคคลให้ทำหน้าที่เป็นเซิร์ฟเวอร์ เพื่อทดสอบการทำงานของเว็บไซต์ Dashboard และระบบจัดการข้อมูลก่อนนำไปใช้งานจริง

4.ระบบจัดการฐานข้อมูล MySQL และ phpMyAdmin คือระบบจัดการฐานข้อมูลเชิงสัมพันธ์ (Relational Database Management System: RDBMS) ที่ได้รับความนิยมสูง มีความเสถียรและรวดเร็ว โดยใช้งานผ่านเครื่องมือ phpMyAdmin ซึ่งเป็นโปรแกรมบนเว็บที่ช่วยให้ผู้ดูแลระบบสามารถจัดการฐานข้อมูล สร้างตาราง (Tables) และบริหารจัดการข้อมูล (CRUD Operations) ได้ผ่านกราฟิกอินเทอร์เฟซ ใช้สำหรับจัดเก็บข้อมูลสำคัญทั้งหมดของระบบ ได้แก่ ข้อมูลผู้ดูแลระบบ, ข้อมูลราคาข้าวโพดย้อนหลัง, ผลลัพธ์การพยากรณ์ และบทความข่าวสารต่างๆ

5.เฟรมเวิร์ก Bootstrap (Bootstrap Framework) คือ Front-end Framework ที่พัฒนาด้วย HTML, CSS และ JavaScript ซึ่งเป็นที่นิยมในการพัฒนาเว็บไซต์แบบ Responsive Design กล่าวคือ สามารถแสดงผลได้อย่างสวยงามและเหมาะสมกับทุกขนาดหน้าจอ ไม่ว่าจะเป็นคอมพิวเตอร์ แท็บเล็ต ใช้สำหรับการออกแบบส่วนติดต่อผู้ใช้งาน (User Interface: UI) ของเว็บไซต์ Dashboard การจัดวาง Layout ตารางข้อมูล และกราฟแสดงผล เพื่อให้เกษตรกรสามารถใช้งานระบบได้ง่ายและสะดวกที่สุด

6.ภาษา PHP คือภาษาคอมพิวเตอร์ในลักษณะ Server-Side Scripting ที่ถูกออกแบบมาสำหรับใช้ในการสร้างเว็บเพจแบบพลวัต (Dynamic Web Page) สามารถทำงานร่วมกับ HTML และฐานข้อมูล MySQL ได้เป็นอย่างดี ใช้เขียนคำสั่งประมวลผลหลักของระบบ เช่น การคำนวณข้อมูล, การดึงข้อมูลพยากรณ์จากฐานข้อมูลมาแสดงผล, ระบบตรวจสอบสิทธิ์การเข้าใช้งาน (Login) และการเชื่อมต่อระหว่างหน้าเว็บกับฐานข้อมูล

7. โปรแกรม Visual Studio Code (VS Code) คือโปรแกรมแก้ไขซอร์สโค้ด (Source Code Editor) ที่พัฒนาโดย Microsoft รองรับการทำงานเขียนโปรแกรมหลากหลายภาษา มีส่วนขยาย (Extensions) ที่ช่วยอำนวยความสะดวกในการเขียนโค้ด ใช้เป็นเครื่องมือหลักในการเขียนและแก้ไขโค้ดภาษา HTML, PHP, CSS และ SQL ตลอดระยะเวลาการพัฒนา

2.3 วรรณกรรมที่เกี่ยวข้อง

ธนกร สุทธิสนธิ (2566) งานวิจัยนี้นำเสนอการพัฒนาแบบพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์รายเดือนในประเทศไทยด้วยการใช้โมเดลผสม (Hybrid Model) ซึ่งรวมเทคนิคหลายแบบเข้าด้วยกัน เพื่อเพิ่มความแม่นยำและลดข้อผิดพลาดในการทำนายราคาข้าวโพดโดยข้อมูลที่ใช้เป็นชุดข้อมูลราคาข้าวโพดรายเดือนย้อนหลังหลายปี งานวิจัยได้เลือกใช้การรวมกันระหว่าง SARIMA (Seasonal ARIMA) เพื่อจับลักษณะข้อมูลที่มีฤดูกาล และ Support Vector Regression (SVR) ที่ช่วยจัดการกับความซับซ้อนและความไม่เชิงเส้นของข้อมูล ผลการทดลองพบว่าโมเดลผสมนี้สามารถพยากรณ์ราคาข้าวโพดได้อย่างแม่นยำสูงกว่าโมเดล SARIMA หรือ SVR เดี่ยวๆ โดยค่าความผิดพลาดของโมเดล (เช่น MAPE, RMSE) ลดลงอย่างมีนัยสำคัญ งานวิจัยยังแนะนำให้ใช้โมเดลนี้ในการสนับสนุนการตัดสินใจของเกษตรกรและผู้ประกอบการในตลาดข้าวโพด เพื่อบริหารความเสี่ยงด้านราคาและวางแผนการผลิตได้อย่างมีประสิทธิภาพ

ภวิดา จันทันแก้ว และ นววรรณ สุนทรภิชช์ (2566) ศึกษาการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในประเทศไทยโดยใช้แบบจำลองถดถอยเชิงเส้นหลายตัวแปร พร้อมเปรียบเทียบกับวิธีถดถอยแบบบริดจ์และลาซโซ่ ข้อมูลที่ใช้เป็นราคาข้าวโพดย้อนหลังตั้งแต่ปี 2545-2562 ผู้วิจัยสร้างพีเจอร์ใหม่ ได้แก่ อัตราการเปลี่ยนแปลงราคามันสำปะหลังและข้าวโพดย้อนหลังเพื่อเพิ่มประสิทธิภาพโมเดล ผลการศึกษาพบว่าแบบจำลองถดถอยเชิงเส้นที่มีการคัดเลือกพีเจอร์ มีความแม่นยำสูงสุด ($R\text{-squared} = 0.94$) และค่าคลาดเคลื่อนต่ำกว่าแบบจำลองอื่น ๆ งานวิจัยนี้ชี้ให้เห็นถึงความสำคัญของการเลือกพีเจอร์ในการพัฒนาการพยากรณ์ราคา และสามารถนำไปใช้ประโยชน์ในการวางแผนด้านการเกษตรได้อย่างมีประสิทธิภาพ

สิรินทรา สามชัยรัตน์ และคณะ (2565) งานวิจัยนี้นำเสนอการพัฒนาแบบจำลองการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในประเทศไทยโดยใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ซึ่งมุ่งเน้นการจับลักษณะความสัมพันธ์เชิงเวลาที่ซับซ้อนและความผันผวนของราคาข้าวโพดในตลาด ผลผลิตและราคาพืชเศรษฐกิจชนิดนี้มีความผันผวนสูงและได้รับผลกระทบจากหลายปัจจัย เช่น สภาพอากาศ ภาวะเศรษฐกิจโลก ราคาน้ำมัน และนโยบายการค้า

งานวิจัยจึงนำเสนอการใช้โมเดล LSTM (Long Short-Term Memory) และ GRU (Gated Recurrent Unit) ซึ่งเป็นเครือข่ายโครงข่ายประสาทเทียมที่ออกแบบมาเพื่อจัดการกับข้อมูลอนุกรมเวลาได้อย่างมีประสิทธิภาพ โดยเฉพาะอย่างยิ่งสำหรับการจับความสัมพันธ์ในข้อมูลระยะยาวและจัดการกับข้อมูลที่มีความไม่แน่นอนสูง ข้อมูลที่ใช้ในการฝึกสอนโมเดลประกอบด้วยข้อมูลราคาข้าวโพดรายเดือนย้อนหลังหลายปี รวมถึงข้อมูลเสริมอื่นๆ เช่น ราคา น้ำมัน สภาพอากาศ อัตราแลกเปลี่ยน และข้อมูลเศรษฐกิจที่อาจส่งผลกระทบต่อราคาข้าวโพดในตลาด งานวิจัยเปรียบเทียบประสิทธิภาพของโมเดล deep learning กับแบบจำลองสถิติแบบดั้งเดิม เช่น ARIMA และเทคนิคการเรียนรู้เครื่องแบบ SVR พบว่าโมเดล LSTM ให้ผลการพยากรณ์ที่แม่นยำกว่า โดยเฉพาะในช่วงเวลาที่ราคามีความผันผวนสูง โมเดลจึงเหมาะสำหรับนำไปใช้ในการสนับสนุนการตัดสินใจของเกษตรกร ผู้ประกอบการ และผู้กำหนดนโยบายในตลาดข้าวโพด

Azis et al. (2023) งานวิจัยนี้ศึกษาประสิทธิภาพของวิธี Support Vector Regression (SVR) ในการพยากรณ์ราคาสินค้าเกษตรหลักระดับชาติ เช่น ข้าว ข้าวโพด และพืชเศรษฐกิจอื่น ๆ โดยใช้ชุดข้อมูลรายเดือนย้อนหลังของสินค้าต่าง ๆ ในระดับประเทศ นักวิจัยได้รวบรวมข้อมูลราคาสินค้าเกษตรรายเดือนจากฐานข้อมูลรัฐบาลและหน่วยงานที่เกี่ยวข้องในช่วงเวลาหลายปี เพื่อทดสอบความแม่นยำของโมเดล SVR ในการทำนายราคาล่วงหน้า ผลการวิเคราะห์พบว่า SVR สามารถทำนายราคาสินค้าเกษตรได้อย่างแม่นยำ โดยมีข้อดีในเรื่องความสามารถจัดการกับข้อมูลที่มีความซับซ้อนและความแปรปรวนสูง นอกจากนี้ SVR ยังมีความยืดหยุ่นและสามารถปรับแต่งพารามิเตอร์เพื่อเพิ่มประสิทธิภาพของโมเดลได้อย่างเหมาะสม งานวิจัยนี้จึงสนับสนุนการใช้ SVR เป็นเครื่องมือหนึ่งในการวางแผนและตัดสินใจเชิงนโยบายสำหรับภาคเกษตร โดยเฉพาะอย่างยิ่งในเรื่องการจัดการความเสี่ยงด้านราคา และช่วยให้เกษตรกรและผู้ประกอบการสามารถเตรียมความพร้อมและวางแผนได้ดียิ่งขึ้น

James et al. (2021) An Introduction to Statistical Learning (2nd ed.). Springer. หนังสือเล่มนี้เป็นตำราพื้นฐานด้าน Machine Learning และสถิติสำหรับงานพยากรณ์และการสร้างแบบจำลอง อธิบายหลักการแบ่งข้อมูลออกเป็นชุดฝึก (training set) และชุดทดสอบ (test set) เพื่อใช้ประเมินประสิทธิภาพของโมเดลกับข้อมูลที่ไม่เคยเห็นมาก่อน โดยแนวปฏิบัติที่ใช้กันทั่วไปคือแบ่งข้อมูลประมาณ 70–90% สำหรับการฝึก และส่วนที่เหลือสำหรับการทดสอบ หนังสือเล่มนี้ได้รับการยอมรับอย่างกว้างขวางในงานวิชาการและเหมาะใช้อ้างอิงหลักการเลือกสัดส่วนการแบ่งข้อมูลในงานวิจัยและโครงการวิเคราะห์ข้อมูล

Lindauer, M., & Hutter, F. (2022) ได้นำเสนอแนวทางการเพิ่มประสิทธิภาพให้กับ อัลกอริทึมการเรียนรู้ของเครื่องผ่านเครื่องมือ SMAC3 ซึ่งเน้นการหาค่าไฮเปอร์พารามิเตอร์ที่เหมาะสมที่สุด (Hyperparameter Optimization) โดยใช้วิธีการเพิ่มประสิทธิภาพแบบเบย์เซียน (Bayesian Optimization) งานวิจัยนี้ชี้ให้เห็นว่าพารามิเตอร์ของอัลกอริทึมส่งผลกระทบต่อประสิทธิภาพของแบบจำลองในการประมวลผลกับชุดข้อมูลและแอปพลิเคชันที่แตกต่างกัน นอกจากนี้ยังมีการนำแบบจำลองต้นไม้ตัดสินใจแบบสุ่ม (Random Forest) มาประยุกต์ใช้ในฐานะแบบจำลองตัวแทนที่มีความยืดหยุ่นสูง ผลการศึกษาเน้นย้ำว่าการปรับแต่งค่าพารามิเตอร์อย่างเป็นระบบช่วยให้แบบจำลองสามารถทำงานได้ดีขึ้นอย่างก้าวกระโดดภายในระยะเวลาการประมวลผลที่จำกัด ซึ่งเป็นฐานรากสำคัญในการพัฒนาการพยากรณ์ที่มีความแม่นยำสูงในปัจจุบัน

Mao, S., & Soonthornphisaj, N. (2024) งานวิจัยนี้นำเสนอการพยากรณ์ราคาข้าวโพดในประเทศไทยโดยใช้เทคนิค Ensemble Learning ซึ่งเป็นการรวมโมเดลหลายๆ แบบเพื่อเพิ่มความแม่นยำของการทำนาย โดยใช้ข้อมูลราคาข้าวโพดรายเดือนย้อนหลังตั้งแต่ปี 2545 ถึง 2566 รวมทั้งหมด 260 ตัวอย่าง พร้อมคุณลักษณะ 53 ตัวที่เกี่ยวข้อง ในขั้นตอนการวิเคราะห์งานวิจัยได้ใช้วิธี Recursive Feature Elimination (RFE) ร่วมกับการวัดความสัมพันธ์แบบ Pearson Correlation เพื่อลดจำนวนคุณลักษณะลงให้เหลือเฉพาะข้อมูลที่มีผลกระทบต่อราคาข้าวโพดจริงๆ จากนั้นใช้เทคนิค 10-fold cross-validation เพื่อฝึกสอนและทดสอบโมเดลต่างๆ ได้แก่ Support Vector Regression (SVR), Bagging-SVR, Random Forest และ Bagging Regression Tree ผลการทดลองพบว่า Bagging-SVR มีประสิทธิภาพสูงสุด โดยให้ค่า $R^2 = 0.961$, MAE = 0.234 และ RMSE = 0.315 ซึ่งดีกว่าโมเดล SVR เดี่ยวและ Random Forest แบบดั้งเดิมอย่างมีนัยสำคัญ งานวิจัยชี้ให้เห็นว่าเทคนิค Ensemble Learning โดยเฉพาะ Bagging-SVR สามารถช่วยเพิ่มความแม่นยำในการพยากรณ์ราคาข้าวโพดในประเทศไทยได้อย่างมีประสิทธิภาพ และเหมาะสำหรับใช้สนับสนุนการตัดสินใจของเกษตรกร ผู้ประกอบการ และผู้กำหนดนโยบายในตลาดข้าวโพด

2.4 เครื่องมือในการออกแบบและวิเคราะห์ข้อมูล

2.4.1 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM) เป็นมาตรฐานการทำเหมืองข้อมูล (Data Mining) ที่นิยมใช้ทั่วโลก เนื่องจากเป็น Workflow ที่ชัดเจนและยืดหยุ่น สามารถปรับใช้กับโครงการวิเคราะห์ข้อมูลทุกประเภท รวมถึงการพยากรณ์ราคาสินค้าเกษตร เช่น ข้าวโพดเลี้ยงสัตว์ (Shearer, 2000) CRISP-DM ประกอบด้วย 6 ขั้นตอนหลัก ดังนี้

1. Business Understanding (ความเข้าใจเชิงธุรกิจ) เป็นขั้นตอนเริ่มต้นที่ต้องทำความเข้าใจ วัตถุประสงค์ของโครงการ และ ความต้องการของผู้ใช้งาน

2. Data Understanding (ความเข้าใจข้อมูล) ทำการสำรวจและตรวจสอบข้อมูล เพื่อเข้าใจรูปแบบ ขอบเขต และคุณภาพข้อมูลขั้นตอนนี้รวมถึงการตรวจสอบ ค่าที่หายไป (Missing Values), ค่าผิดปกติ (Outliers), การแจกแจงข้อมูล และความสัมพันธ์ระหว่างตัวแปร

3. Data Preparation (การเตรียมข้อมูล) เป็นขั้นตอน ทำความสะอาดและปรับแต่งข้อมูล ให้พร้อมใช้กับโมเดลรวมถึงการจัดรูปแบบข้อมูล, การรวมหลายแหล่งข้อมูล, การแปลงข้อมูล (Feature Engineering) และการสร้างตัวแปรใหม่

4. Modeling (การสร้างแบบจำลอง) นำข้อมูลที่เตรียมแล้วไป สร้างโมเดลทางสถิติหรือแมชชีนเลิร์นนิง เพื่อพยากรณ์หรือวิเคราะห์ข้อมูลเลือกเทคนิคที่เหมาะสมตามประเภทปัญหา (เช่น การพยากรณ์เป็น Time Series หรือ Regression)

5. Evaluation (การประเมินผล) ตรวจสอบ ความแม่นยำของโมเดล และวัดประสิทธิภาพการพยากรณ์ ใช้ตัวชี้วัด เช่น RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) หรือ MAPE (Mean Absolute Percentage Error)

รากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) เป็นเกณฑ์วัดความคลาดเคลื่อนที่นิยมใช้ในการประเมินความแม่นยำของตัวแบบพยากรณ์ โดยคำนวณจากรากที่สองของค่าเฉลี่ยความแตกต่างกำลังสองระหว่างค่าจริงและค่าพยากรณ์ มีสูตรคำนวณดังนี้

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

ค่าเฉลี่ยร้อยละของความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error: MAPE) เป็นเกณฑ์ที่ใช้วัดความคลาดเคลื่อนในรูปแบบเปอร์เซ็นต์ ซึ่งช่วยให้เปรียบเทียบความแม่นยำได้ง่าย มีสูตรคำนวณดังนี้

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error: MAE) เป็นเกณฑ์วัดประสิทธิภาพที่คำนวณจากค่าเฉลี่ยของผลรวมความแตกต่างระหว่างค่าจริงกับค่าพยากรณ์ในรูปของค่าสัมบูรณ์ (Absolute Value) โดยไม่นำเครื่องหมายบวกหรือลบมาคิด ทำให้ทราบว่าโดยเฉลี่ยแล้วโมเดลพยากรณ์คลาดเคลื่อนไปจากความเป็นจริงกี่หน่วย (เช่น กีบาท)

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

6. Deployment (การนำไปใช้จริง) นำโมเดลที่ผ่านการประเมินแล้วไปใช้งานจริงในระบบ หรือสร้าง แดชบอร์ดสำหรับการติดตามผลแบบเรียลไทม์

2.4.2 แบบจำลองอนุกรมเวลา (Time Series)

2.4.2.1 SARIMA (Seasonal ARIMA) แบบจำลองนี้ขยาย ARIMA ให้รองรับฤดูกาล (Seasonality) เหมาะสำหรับข้อมูลราคาข้าวโพดที่มีแนวโน้มตามฤดูกาล เช่น ราคาสูงช่วงปลายปีและลดช่วงต้นปีสามารถจับ แนวโน้ม + ฤดูกาล ได้ชัดเจน การประเมินประสิทธิภาพ RMSE, MAE, MAPE (Hyndman, R. J., & Athanasopoulos, G. 2018).

2.4.2.2 Linear Regression แบบจำลองเชิงเส้น ใช้ ตัวแปรอิสระ (เช่น ราคาย้อนหลัง, ปริมาณผลผลิต, ราคาปุ๋ย) ทำนายราคาข้าวโพด เหมาะสำหรับศึกษาความสัมพันธ์ระหว่างราคากับปัจจัยอื่น ๆ จุดเด่น คือเข้าใจง่าย, วิเคราะห์ผลกระทบของแต่ละตัวแปรได้ การประเมินประสิทธิภาพ RMSE, MAE (Montgomery, D. C., Peck, E. A., & Vining, G. G. 2012).

2.4.2.3 Random Forest แบบจำลอง Machine Learning ที่รวมหลาย Decision Trees เพื่อลดความคลาดเคลื่อน ใช้ข้อมูลหลายตัวแปร (เช่น ราคาย้อนหลัง, ปริมาณผลผลิต, ปัจจัยอากาศ) ทำนายราคาข้าวโพด จุดเด่น คือจับความสัมพันธ์ที่ซับซ้อนได้ดี, ลด overfitting ด้วยการรวมหลาย Trees การประเมินประสิทธิภาพ RMSE, MAE (Breiman, L. 2001).

2.4.2.4 Support Vector Regression (SVR) แบบจำลอง Machine Learning ที่ประยุกต์ใช้แนวคิดพื้นฐานมาจาก Support Vector Machine (SVM) แต่นำมาปรับใช้กับข้อมูลตัวเลขที่มีค่าต่อเนื่อง (Continuous Data) เพื่อพยากรณ์ราคาข้าวโพดจากตัวแปรหลายมิติ แบบจำลองนี้สามารถจัดการกับความสัมพันธ์ที่ไม่เป็นเส้นตรง (Non-linear relationships) ผ่านการใช้เคอร์เนล (Kernel Trick) จุดเด่นคือ มีความแม่นยำสูงเมื่อเผชิญกับข้อมูลที่มีความซับซ้อน และสามารถจัดการกับค่าผิดปกติ (Outliers) ได้อย่างมีประสิทธิภาพ การประเมินประสิทธิภาพใช้ค่า RMSE, MAE, MAPE

2.5 บทสรุป

จากการทบทวนแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลเพื่อการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ พบว่าการใช้กระบวนการวิเคราะห์ข้อมูลตามขั้นตอนของ Cross-Industry Standard Process for Data Mining (CRISP-DM) เป็นแนวทางที่เหมาะสม เนื่องจากเป็นมาตรฐานสากลที่สามารถปรับใช้ได้กับทุกประเภทของข้อมูล โดยโครงการนี้ได้เลือกใช้เทคนิคการพยากรณ์อนุกรมเวลา ได้แก่ SARIMA, Linear Regression, Random Forest และ Support Vector Regression (SVR) เพื่อตรวจสอบและเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ นอกจากนี้ ผลการวิเคราะห์และพยากรณ์ยังถูกนำเสนอผ่าน เว็บไซต์และแดชบอร์ดเชิงโต้ตอบ (Interactive Dashboard) โดยใช้เทคโนโลยี HTML, CSS และ JavaScript ทำให้ผู้ใช้งานสามารถเข้าถึงข้อมูลย้อนหลังและข้อมูลการพยากรณ์ได้อย่างสะดวก รวดเร็ว และมีประสิทธิภาพ ซึ่งช่วยสนับสนุนการตัดสินใจเชิงธุรกิจของเกษตรกร ผู้ประกอบการ และหน่วยงานที่เกี่ยวข้องกับตลาดสินค้าเกษตรได้อย่างเป็นรูปธรรม