บทที่ 2
แนวคิด ทฤษฎี เครื่องมือและวรรณกรรมที่เกี่ยวข้อง

โครงงานเรื่องการวิเคราะห์ข้อมูลเพื่อการพยากรณ์มูลค่าการส่งออกสินค้าประมงของไทยไปยังตลาดหลัก โดยใช้เทคนิคเหมืองข้อมูล ในบทนี้เป็นการนำเสนอเกี่ยวกับ แนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้อง ของการวิเคราะห์ข้อมูลเพื่อการพยากรณ์มูลค่าการส่งออกสินค้าประมงของไทยไปยังตลาดหลัก โดยใช้เทคนิคเหมืองข้อมูล ซึ่งได้รวบรวมการศึกษาเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางการศึกษา ประกอบด้วยรายละเอียดตามลำดับ ดังนี้

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics)

2.1.2 แนวคิดเกี่ยวกับการพยากรณ์ (Forecasting Concepts)

2.1.3 แนวคิดการส่งออกสินค้าประมงของไทย

2.2 ทฤษฎี

2.2.1 การทำเหมืองข้อมูล (Data Mining)

2.2.2 ทฤษฎีการประเมินความแม่นยำของแบบจำลอง (Model Evaluation Theory)

2.2.3 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์

2.2.4 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP–DM)

2.2.5 เทคนิค Random Forest

2.2.6 เทคนิค ARIMA (Auto Regressive Integrated Moving Average)

2.2.7 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression)

2.2.8 การปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)

2.2.9 ทฤษฎีภาษาที่ใช้สำหรับการพัฒนาเว็บไซต์

2.3 วรรณกรรมที่เกี่ยวข้อง

2.3.1 การส่งออกผลิตภัณฑ์ทูน่ากระป๋องของไทย

2.3.2 การพยากรณ์ปริมาณการส่งออกน้ำตาลรายเดือนของประเทศไทยด้วยตัวแบบอนุกรมเวลา

2.3.3 การพยากรณ์ปริมาณการส่งออกข้าวไทยไปจีน

2.3.4 การพยากรณ์ราคาข้าวเปลือกหอมมะลิ

2.3.5 การพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์ และส่วนประกอบของประเทศไทย

2.4 บทสรุป

2.1 แนวคิด

2.1.1 แนวคิดเกี่ยวกับการวิเคราะห์ข้อมูล (Data Analytics)

การดำเนินงานเรื่องการวิเคราะห์ข้อมูลเพื่อพยากรณ์มูลค่าการส่งออกสินค้าของวิสาหกิจขนาดกลางและขนาดย่อม โดยใช้เทคนิคการทำเหมืองข้อมูล ผู้จัดทำได้ศึกษาหลักการและทฤษฎีต่างๆ ซึ่งองค์ประกอบที่สำคัญ คือ การวิเคราะห์ข้อมูล (ดวงแก้ว เงินพูลทรัพย์, 2567) ซึ่งมีรายละเอียด ดังนี้

2.1.1.1 Data Analytics เป็นเครื่องมือสำหรับธุรกิจ (Business Intelligence) เพราะการที่กิจการบริหารงานโดยไม่ทราบข้อมูลที่จำเป็นก็เปรียบเหมือนคนตาบอดที่คอยคลำทางเพื่อไปยังจุดหมายซึ่งอาจจะถูกหรือผิดทาง บางครั้งอาจสรุปได้ว่าเป็นการเดินทางของกิจการที่ไร้จุดหมายปลายทาง ดังนั้น การทำ Data Analytics นี้ไม่จำเป็นต้องเป็นธุรกิจขนาดใหญ่เท่านั้น ธุรกิจขนาดกลางและเล็กก็สามารถทำได้เช่นกัน เรามาทำความเข้าใจรูปแบบของการวิเคราะห์ข้อมูล (Data Analytics)

2.1.1.2 การวิเคราะห์ข้อมูลแบบพื้นฐาน (Descriptive Analytics) เป็นการวิเคราะห์เพื่อแสดงผลของรายการทางธุรกิจ เหตุการณ์ หรือกิจกรรมต่างๆ ที่ได้เกิดขึ้น หรืออาจกำลังเกิดขึ้นในลักษณะที่ง่ายต่อการเข้าใจหรือต่อการตัดสินใจ ตัวอย่างเช่น รายงานการขาย รายงานผลการดำเนินงาน ปกติรายงานเหล่านี้เป็นรายงานที่เกี่ยวข้องกับการจัดทำข้อมูลทางการเงินหรืองบการเงินประจำปีและงบการเงินประจำปี เป็นหนึ่งในรายงานที่ใช้ในการวิเคราะห์ข้อมูลแบบพื้นฐานที่ผู้ใช้งบการเงินต้องการ เช่น กำไรขั้นต้นสูงขึ้นเนื่องจากต้นทุนการผลิตที่ต่ำลงจากผลของการขยายกำลังการผลิตที่เพิ่มขึ้นตามยอดขายที่เติบโตขึ้นโดยปัจจัยด้านราคาขายคงที่

2.1.1.3 การวิเคราะห์แบบเชิงวินิจฉัย (Diagnostic Analytics) เป็นการอธิบายถึงสาเหตุของสิ่งที่เกิดขึ้นปัจจัยต่างๆ และความสัมพันธ์ของปัจจัยหรือตัวแปรต่างๆ ที่มีความสัมพันธ์ต่อกันของสิ่งที่เกิดขึ้น ตัวอย่างเช่น ความสัมพันธ์ระหว่างยอดขายต่อกิจกรรมทางการขายของฝ่ายขายหรือแนวทางการทำตลาดของฝ่ายการตลาด ปกติเรื่องนี้มักจะเป็นการหาแนวทางใหม่ๆ ที่ช่วยเสริมให้การตัดสินใจของฝ่ายจัดการนั้นเป็นไปในทางที่ถูกต้องที่สุดในสภาพแวดล้อมปัจจุบัน เช่น

จำนวนยอดขายปี 25x2 เพิ่มขึ้น 5% จากปีก่อน เนื่องจากการเพิ่มกิจกรรมส่งเสริมการตลาดที่มีมากขึ้นและมีประสิทธิผลสูงจากการตอบรับของผู้บริโภคประกอบกับดัชนีความเชื่อมั่นของผู้บริโภคสูงขึ้น

2.1.1.4 การวิเคราะห์แบบพยากรณ์ (Predictive Analytics) เป็นการวิเคราะห์เพื่อพยากรณ์สิ่งที่กำลังจะเกิดขึ้นหรือน่าจะเกิดขึ้น โดยใช้ข้อมูลที่ได้เกิดขึ้นแล้วในอดีต (ส่วนใหญ่เป็นข้อมูลพื้นฐาน) กับแบบจำลองทางสถิติต่าง ๆ (สมมติฐาน) รวมถึงการใช้เทคโนโลยีปัญญาประดิษฐ์ต่างๆ (Artificial Intelligence) เป็นเครื่องมือช่วยในการจัดเรียงข้อมูลหรือเพื่อการวิเคราะห์ ตัวอย่างเช่น การพยากรณ์ยอดขาย การพยากรณ์ผลตอบแทนจากการลงทุนอาทิ จำนวนยอดขายปี 25x2 มีความเป็นไปได้ที่จะเพิ่มขึ้นอีก 5% จากปีก่อน ดัชนีความเชื่อมั่นของผู้บริโภคสูงขึ้น และแผนการขยายกำลังการผลิตจากเดิมที่ได้ลงทุนแล้วในปีก่อนหน้าสามารถเพิ่มกำลังการผลิตได้อีกเพราะยังมีกำลังการผลิตคงเหลือ ซึ่งสอดคล้องกับความต้องการสินค้าที่มีสูงขึ้นอย่างต่อเนื่องจากพฤติกรรมผู้บริโภคที่เปลี่ยนวิธีการสั่งซื้อสินค้ามาเป็นการซื้อผ่านช่องทางออนไลน์เพิ่มขึ้น

2.1.1.5 การวิเคราะห์แบบให้คำแนะนำ (Prescriptive Analytics) เป็นการวิเคราะห์ข้อมูลที่มีความซับซ้อนที่สุดและอาศัยความเชี่ยวชาญของผู้วิเคราะห์มากที่สุด เพราะเป็นทั้งการพยากรณ์สิ่งต่างๆ ที่จะเกิดขึ้น ข้อดี ข้อเสีย สาเหตุ และระยะเวลาของสิ่งที่จะเกิดขึ้น รวมถึงการให้คำแนะนำทางเลือกต่างๆ ที่มีอยู่ และผลของแต่ละทางเลือกแก่ฝ่ายจัดการ เพราะจะต้องใช้ข้อมูลพื้นฐานข้อมูลจากการวินิจฉัย ตลอดจนการคาดการณ์ต่างๆ อย่างละเอียดและระมัดระวังรอบคอบอย่างมาก เช่น กิจการควรชะลอแผนการลงทุนขยายกำลังการผลิตของโรงงานออกไปที่ไตรมาสที่ 3 ของปี เนื่องจากความไม่ชัดเจนของจำนวนความต้องการของสินค้าในช่วง 2 ปีข้างหน้า ด้วยคู่แข่งที่เริ่มเข้ามามากขึ้นและเทคโนโลยีการผลิตเปลี่ยนแปลงค่อนข้างเร็ว ตลอดจนอัตราดอกเบี้ยมีแนวโน้มลดลงส่งผลให้ต้นทุนทางการเงินในการดำเนินการที่น่าจะลดลงความชัดเจนน่าจะมีมากขึ้นในช่วงไตรมาสที่ 3 ของปีนี้

2.1.2 แนวคิดเกี่ยวกับการพยากรณ์ (Forecasting Concepts)

การพยากรณ์ (Forecasting) หมายถึง กระบวนการคาดการณ์หรือประมาณค่าของเหตุการณ์หรือข้อมูลที่จะเกิดขึ้นในอนาคต โดยอาศัยข้อมูลในอดีต ข้อมูลปัจจุบัน และปัจจัยแวดล้อมที่เกี่ยวข้อง เพื่อช่วยให้การตัดสินใจและการวางแผนมีความแม่นยำและมีประสิทธิภาพ

มากยิ่งขึ้น การพยากรณ์จึงเป็นเครื่องมือสำคัญที่ใช้ทั้งในภาคธุรกิจ เศรษฐกิจ การผลิต การตลาด และการบริหารจัดการทรัพยากร แนวคิดหลักของการพยากรณ์สามารถอธิบายได้ดังนี้

2.1.2.1 พื้นฐานของการพยากรณ์การพยากรณ์อาศัยข้อมูลจากอดีต (Historical Data) และสมมติฐานที่ว่า “รูปแบบ (Pattern) ที่เกิดขึ้นในอดีตจะยังคงเกิดขึ้นในอนาคต”

2.1.2.2 ประเภทของการพยากรณ์

1) การพยากรณ์ระยะสั้น (Short–term Forecasting) มุ่งเน้นการคาดการณ์ในช่วงเวลาไม่เกิน 1 ปี เช่น การพยากรณ์ยอดขายรายเดือน การจัดการสินค้า คงคลัง

2) การพยากรณ์ระยะกลาง (Medium–term Forecasting) ช่วงเวลา 1 – 3 ปี เช่น การวางแผนการผลิตหรือการวางแผนทรัพยากร

3) การพยากรณ์ระยะยาว (Long–term Forecasting) เกิน 3 ปีขึ้นไป เช่น การคาดการณ์ความต้องการพลังงานหรือแนวโน้มเศรษฐกิจ

2.1.2.3 วิธีการพยากรณ์

1) วิธีเชิงคุณภาพ (Qualitative Methods) อาศัยความคิดเห็น ประสบการณ์ และการตัดสินของผู้เชี่ยวชาญ เช่น Delphi Method หรือการสำรวจความคิดเห็น (Survey)

2) วิธีเชิงปริมาณ (Quantitative Methods) ใช้แบบจำลองทางสถิติและคณิตศาสตร์ เช่น การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เช่น Moving Average, Exponential Smoothing, ARIMA แบบจำลองเชิงสาเหตุ (Causal Models) เช่น Regression Analysis, Econometric Models (Amazon, 2568)

2.1.3 แนวคิดการส่งออกสินค้าประมงของไทย

การส่งออกสินค้าประมงของไทยมีความสำคัญอย่างยิ่งต่อเศรษฐกิจและการค้าระหว่างประเทศ เนื่องจากเป็นแหล่งรายได้หลักที่ช่วยสร้างดุลการค้าและสร้างมูลค่าเพิ่มให้กับสินค้าเกษตรและผลิตภัณฑ์แปรรูป นอกจากนี้ยังมีส่วนสำคัญในการสนับสนุนอุตสาหกรรมต่อเนื่องต่างๆ เช่น การขนส่ง บรรจุภัณฑ์ และการค้าปลีก อันเป็นการขยายโอกาสทางเศรษฐกิจของประเทศโดยรวมการส่งออกสินค้าประมงได้รับอิทธิพลจากปัจจัยหลายประการ ได้แก่ ปัจจัยด้านอุปสงค์ (Demand) ซึ่งมาจากความต้องการสินค้าประมงของตลาดโลก โดยเฉพาะประเทศคู่ค้าหลักอย่างญี่ปุ่น สหรัฐอเมริกา สหภาพยุโรป และจีน ปัจจัยด้านอุปทาน (Supply) ซึ่งขึ้นอยู่กับปริมาณการผลิตจากการเพาะเลี้ยงและการจับสัตว์น้ำ รวมถึงคุณภาพและมาตรฐานความปลอดภัยอาหาร ปัจจัยด้านการค้าและเศรษฐกิจ เช่น อัตรา

แลกเปลี่ยน สถานการณ์เศรษฐกิจโลก มาตรการทางการค้า ภาษี และกฎระเบียบด้านการนำเข้า ตลอดจนปัจจัยด้านสิ่งแวดล้อมและสังคม ไม่ว่าจะเป็นการเปลี่ยนแปลงสภาพภูมิอากาศ ความยั่งยืนของทรัพยากรประมง และข้อกำหนดด้านแรงงานที่เกี่ยวข้องในด้านแนวโน้มและทิศทางการส่งออก ประเทศไทยมีการมุ่งเน้นการผลิตสินค้าที่ได้มาตรฐานสากลและมีความปลอดภัยต่อผู้บริโภค ควบคู่ไปกับการพัฒนาผลิตภัณฑ์แปรรูปที่มีมูลค่าเพิ่ม เพื่อตอบสนองต่อความต้องการสินค้าพรีเมียมในตลาดโลก อีกทั้งยังมีการนำเทคโนโลยีและนวัตกรรมมาใช้เพื่อเพิ่มประสิทธิภาพในการผลิตและสนับสนุนระบบตรวจสอบย้อนกลับ (Traceability) รวมถึงการปรับตัวให้สอดคล้องกับมาตรฐานด้านสิ่งแวดล้อมและความยั่งยืน ซึ่งล้วนเป็นปัจจัยสำคัญที่ช่วยรักษาความสามารถในการแข่งขันของไทยในตลาดโลก (กรมประมง, 2567)

2.2 ทฤษฎี

2.2.1 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) หมายถึง กระบวนการสกัดและค้นหาความรู้จากข้อมูลจำนวนมาก เพื่อนำมาสร้างรูปแบบ ความสัมพันธ์ หรือแนวโน้มที่สามารถนำไปใช้ในการตัดสินใจได้ โดยอาศัยเทคนิคทางสถิติ คณิตศาสตร์ และปัญญาประดิษฐ์ เช่น การถดถอยเชิงเส้น (Linear Regression) และการจัดกลุ่ม (Clustering) การประยุกต์ใช้เหมืองข้อมูลในเชิงเศรษฐศาสตร์การเกษตรและการส่งออก สามารถช่วยในการพยากรณ์แนวโน้มความต้องการสินค้าและการเคลื่อนไหวของตลาดได้อย่างแม่นยำ

2.2.2 ทฤษฎีการประเมินความแม่นยำของแบบจำลอง (Model Evaluation Theory)

การประเมินความแม่นยำของแบบจำลอง (Model Evaluation Theory) เป็นขั้นตอนสำคัญในกระบวนการวิเคราะห์ข้อมูลที่ช่วยให้มั่นใจได้ว่าแบบจำลองที่สามารถอธิบายหรือคาดการณ์ข้อมูลได้อย่างมีประสิทธิภาพ ตัวชี้วัดมาตารฐานที่ใช้ในการวิเคราะห์โมเดล ดังนี้

Mean Absolute Percentage Error (MAPE) สามารถหาได้จากค่าเฉลี่ยของค่าร้อยละความคลาดเคลื่อนมาเฉลี่ยกัน ดังสมการ

$$\text{MAPE} = 1/n \sum_(i =)\^n \sum |e_i| / |A_i| = 1/n \sum_(i = 1)\^n \sum |F_i - A_i| / |A_i|$$

ค่า Mean Absolute Percentage Error (MAPE) ใช้การรายงานผลเป็นร้อยละทำให้สามารถทำความเข้าใจถึงระดับความแม่นยำได้ โดยไม่ต้องคำนึงถึงขนาดของสิ่งที่กำลังทำนายอยู่ (scale–independent) ทำให้รายงานให้ผู้มีส่วนได้ส่วนเสียเข้าใจได้ง่าย และสามารถตัดสินใจได้เร็ว

ข้อจำกัดด้านเทคนิคของ MAPE

1) ตัวหารที่ใกล้เคียง 0 ในกรณีที่ค่าจริง (Actual Demand) เป็น 0 จะไม่สามารถหาค่าได้ หรือในกรณีที่มีค่าใกล้เคียง 0 จะทำให้ค่าความคาดเคลื่อน มีค่าสูงผิดปกติ

2) สูตรไม่สมมาตร ในกรณีที่ค่าทำนายนั้น น้อยกว่าค่าจริง ค่า Absolute Percentage Error จะอยู่ที่ร้อยละ 0 – 100 แต่ในกรณีที่ค่าทำนายสูงกว่าค่าจริง ค่า Absolute Percentage Error มีโอกาสที่จะสูงกว่าร้อยละ 100 ดังนั้นการใช้ค่านี้ ในการเลือกโมเดล จะมีแนวโน้มที่จะได้โมเดลที่ทำนายค่าที่ต่ำกว่าปกติ

Mean Absolute Error (MAE) สามารถคำนวณได้จากการนำค่าสัมบูรณ์ของความคลาดเคลื่อน (Absolute Error) มาเฉลี่ยกัน ดังสมการ

$$\text{MAE} = \sum_(i =)\^n\sum|e_i|$$

วิธีการคำนวณ MAE จำเป็นต้องพิจารณาควบคู่กับขนาด (scale) ของสิ่งที่ทำนายด้วย การหารค่า MAE ด้วยค่าเฉลี่ยของค่าจริงหรือค่าร้อยละ MAE ของ MAE ดังสมการ

$$\text{ร้อยละ MAE} = (\sum_(i = 1)\^n\sum|e_i|)/(\sum_(i = 1)\^n\sum|A_i|)$$

Root Mean Square Error (RMSE) สามารถวัดได้โดยการนำค่ากำลังสองของความคลาดเคลื่อนมาเฉลี่ยกัน แล้วถอดรากที่ 2 ดังสมการ

$$\text{RMSE} = \sqrt{(1/n \sum_(i = 1)\^n\sum(A_i - F_i)\^2\ \ )}$$

ตัวชี้วัด Root Mean Square Error (RMSE) นั้นมีความซับซ้อนมากกว่าการคำนวณ MAE (Mean Square Error) เล็กน้อย และจะให้น้ำหนักความเสียหาย (Penalize) กับ กรณีที่ค่าความคลาดเคลื่อนมีความสุดโต่ง โดยตัวค่า RMSE นั้นสามารถหาเป็นร้อยละ ได้เช่นเดียวกับ MAE ดังสมการ

$$\text{ร้อยละ RMSE} = \sqrt{(1/n \sum_(i = 1)\^n\sum(A_i - F_i)\^2\ \ )}/(\text{RMSE})$$

ค่า RMSE นั้นจะได้รับความนิยมสูงในการใช้เพื่อฝึกฝน และเลือกแบบจำลอง เนื่องจากสูตรนั้นค่อนข้างง่ายต่อการหาค่าอนุพันธ์ (Derivative) ทำให้การปรับแบบจำลองนั้น สามารถทำได้อย่างเป็นธรรมชาติมากกว่า และเป็นค่าชี้วัดที่ใช้ในการแข่งขันสำหรับโจทย์ประเภท Regression จำนวนมากบน Platform Kaggle (Big data institute, 2566)

2.2.3 ทฤษฎีเกี่ยวกับการสร้างเว็บไซต์

2.2.3.1 การออกแบบเว็บไซต์ (Web Design) คือ กระบวนการสร้างและวางแผนการแสดงผลของเว็บไซต์ทั้งในด้านการใช้งาน (User Experience: UX) และลักษณะการมองเห็น (User Interface: UI) โดยทำให้เว็บไซต์นั้นมีความสวยงาม ใช้งานง่าย และสามารถตอบโจทย์ผู้ใช้ได้อย่างมีประสิทธิภาพ การออกแบบเว็บไซต์ประกอบด้วย

1) การออกแบบโดยมุ่งเน้นผู้ใช้ (User-Centered Design) สิ่งสำคัญของการออกแบบเว็บไซต์คือการให้ความสำคัญกับผู้ใช้เป็นหลัก โดยเริ่มจากการศึกษากลุ่มเป้าหมาย ความต้องการ และพฤติกรรมการใช้งาน เพื่อนำข้อมูลมาประยุกต์ใช้ในการออกแบบ ทั้งในด้านส่วนติดต่อผู้ใช้ (UI) และประสบการณ์ผู้ใช้ (UX) เพื่อให้สามารถใช้งานได้ง่าย ไม่ซับซ้อน และตอบสนองความต้องการได้อย่างแท้จริง

2) การออกแบบที่รองรับทุกอุปกรณ์ (Responsive Design) เว็บไซต์ควรถูกออกแบบให้สามารถปรับขนาดและรูปแบบการแสดงผลได้เหมาะสมกับทุกอุปกรณ์ โดยอาศัยเทคนิคการพัฒนา เช่น CSS Media Queries, Flexbox และ Grid เพื่อเพิ่มความสะดวกในการเข้าถึง

3) ความเรียบง่ายที่มีประสิทธิภาพ (Minimalism) การออกแบบเชิงมินิมอลมุ่งเน้นการลดทอนองค์ประกอบที่ไม่จำเป็น ใช้พื้นที่ว่าง (Negative Space) อย่างเหมาะสม และเน้นการจัดลำดับความสำคัญของเนื้อหาให้ชัดเจน เพื่อเสริมสร้างการรับรู้และประสบการณ์ใช้งานที่มีคุณภาพ

4) ประสิทธิภาพการโหลดหน้าเว็บไซต์ (Fast Page Loading) ความเร็วในการโหลดหน้าเว็บเป็นปัจจัยสำคัญต่อความพึงพอใจของผู้ใช้ เว็บไซต์ควรถูกออกแบบเพื่อลดเวลาในการโหลด เช่น การปรับขนาดไฟล์ภาพ การใช้ Lazy Loading และการจัดการ Cache ของเบราว์เซอร์

5) การใช้ภาพและวิดีโอเพื่อเสริมความเข้าใจ (Use of Images and Videos) สื่อมัลติมีเดียมีบทบาทสำคัญในการเพิ่มความน่าสนใจและความเข้าใจในเนื้อหา การเลือกใช้ภาพและวิดีโอควรคำนึงถึงคุณภาพ ความเหมาะสม และการจัดการไฟล์ให้มีขนาดที่ไม่กระทบต่อความเร็วของเว็บไซต์

6) การเข้าถึงสำหรับผู้ใช้ทุกกลุ่ม (Accessibility) การออกแบบเว็บไซต์ต้องคำนึงถึงการเข้าถึงของผู้ใช้ที่หลากหลาย รวมถึงผู้พิการ โดยใช้เทคนิคต่างๆ เช่น การใส่ Alt Text ให้ภาพ การใช้โครงสร้างหัวข้อที่ชัดเจน สีที่มีความคมชัด และการรองรับการใช้งานผ่านคีย์บอร์ด

7) แนวคิด Mobile First การออกแบบเว็บไซต์ควรเริ่มจากการรองรับการใช้งานบนอุปกรณ์เคลื่อนที่เป็นลำดับแรก เนื่องจากผู้ใช้ส่วนใหญ่เข้าถึงอินเทอร์เน็ตผ่านสมาร์ทโฟน การออกแบบจึงควรเน้นเฉพาะสิ่งที่จำเป็นและเหมาะสมกับหน้าจอขนาดเล็ก

8) การใช้แอนิเมชันและการโต้ตอบขนาดเล็ก (Animation & Microinteractions) การเคลื่อนไหวเล็ก ๆ บนเว็บไซต์ช่วยเพิ่มประสบการณ์เชิงโต้ตอบให้กับผู้ใช้ แต่ควรใช้ในปริมาณที่เหมาะสม เพื่อไม่ให้รบกวนการใช้งานหรือสร้างความซับซ้อนเกินไป

9) การประยุกต์ใช้ Chatbots และ AI การนำระบบแชทบอทที่ขับเคลื่อนด้วย AI มาประยุกต์ใช้บนเว็บไซต์ ช่วยให้สามารถโต้ตอบ ตอบคำถาม และให้บริการผู้ใช้ได้ตลอดเวลา โดยอาศัยเทคโนโลยีการประมวลผลภาษาธรรมชาติ (NLP) และ Machine Learning เพื่อเพิ่มประสิทธิภาพการให้บริการ

10) การรักษาความปลอดภัยของผู้ใช้ (Security) เว็บไซต์ควรให้ความสำคัญกับความปลอดภัยของข้อมูลผู้ใช้ เช่น การเข้ารหัสด้วย SSL/HTTPS การยืนยันตัวตนแบบหลายชั้น (Two-Factor Authentication) และมาตรการปกป้องข้อมูลส่วนบุคคลจากการรั่วไหล (Foxbith, 2567)

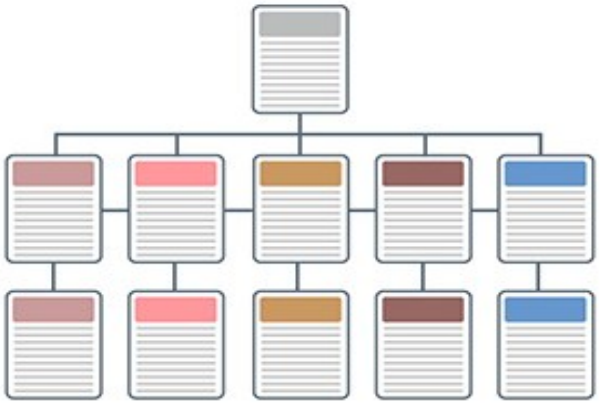
2.2.3.2 รูปแบบโครงสร้างของเว็บไซต์ หมายถึง การจัดระเบียบและกำหนดความสัมพันธ์ของหน้าเพจภายในเว็บไซต์ เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูได้อย่างมีประสิทธิภาพ โดย (En Fete, 2567) ได้แบ่งโครงสร้างเว็บไซต์ออกเป็น 4 รูปแบบ ดังนี้

1) Sequential Structure (โครงสร้างแบบเรียงลำดับ) โครงสร้างเว็บไซต์ลักษณะนี้เป็นการจัดเรียงหน้าเพจตามลำดับเชื่อมโยงกันเหมือนเส้นตรง โดยผู้ใช้งานสามารถเลือกการเข้าถึงหน้าเพจได้แบบไปข้างหน้า (Next) หรือถอยหลัง (Back) ตามลำดับที่กำหนด โครงสร้างดังกล่าวเหมาะสมกับเว็บไซต์ขนาดเล็กที่มีเนื้อหาเชิงลำดับขั้น เช่น เว็บไซต์คู่มือการใช้งาน หรือเว็บไซต์ที่มีลักษณะการนำเสนอเป็นขั้นตอนอย่างต่อเนื่อง

ภาพที่ 2.1 โครงสร้างแบบเรียงลำดับ

(ที่มา: https://enfete.co.th/th/website-structure/)

2) Hierarchical Structure (โครงสร้างแบบลำดับขั้น) โครงสร้างเว็บไซต์แบบลำดับขั้นเป็นการจัดระบบข้อมูลให้มีความสัมพันธ์กันในลักษณะต้นไม้ โดยมีหน้าแรก (Home Page) เป็นจุดศูนย์กลางในการกระจายไปยังหมวดหมู่หลัก และแตกแขนงต่อไปยังเนื้อหาย่อยตามลำดับชั้นที่กำหนด โครงสร้างนี้เหมาะสำหรับเว็บไซต์ที่มีปริมาณข้อมูลจำนวนมากและหลากหลายหมวดหมู่ เช่น เว็บไซต์องค์กร หรือเว็บไซต์ที่มีการแบ่งหมวดหมู่เนื้อหาอย่างชัดเจน



ภาพที่ 2.2 โครงสร้างแบบลำดับขั้น

(ที่มา: https://enfete.co.th/th/website-structure/)

3) Grid Structure (โครงสร้างแบบตาราง) โครงสร้างเว็บไซต์แบบตารางมีลักษณะการเชื่อมโยงเนื้อหาภายในเว็บไซต์เข้าหากันอย่างอิสระในหลายทิศทาง ผู้ใช้งานสามารถเลือกเข้าถึงข้อมูลได้จากหลายเส้นทางโดยไม่จำเป็นต้องเรียงตามลำดับขั้นตอนที่เคร่งครัด โครงสร้างนี้มักพบในเว็บไซต์ที่ต้องการความยืดหยุ่นในการนำเสนอ เช่น เว็บไซต์ข่าว หรือเว็บไซต์บล็อกที่มีการจัดหมวดหมู่ย่อยหลายกลุ่ม

ภาพที่ 2.3 โครงสร้างแบบตาราง

(ที่มา: https://enfete.co.th/th/website-structure/)

4) Web Structure (โครงสร้างแบบใยแมงมุม) โครงสร้างเว็บไซต์แบบใยแมงมุมเป็นการเชื่อมโยงหน้าเพจทั้งหมดภายในเว็บไซต์เข้าด้วยกันอย่างเสรี โดยผู้ใช้งานสามารถเข้าถึงเนื้อหาได้จากหลายเส้นทางโดยไม่จำกัดลำดับขั้นตอน อีกทั้งยังสามารถเชื่อมโยงไปยังเว็บไซต์ภายนอกได้ ลักษณะดังกล่าวเหมาะกับเว็บไซต์ที่ต้องการเน้นความอิสระและการเข้าถึงข้อมูลที่ครอบคลุม เช่น เว็บไซต์สารานุกรม หรือเว็บไซต์ที่เป็นฐานข้อมูลความรู้

ภาพที่ 2.4 โครงสร้างแบบใยแมงมุม

(ที่มา: https://enfete.co.th/th/website-structure/)

2.2.3.3 องค์ประกอบของเว็บไซต์ คือ องค์ประกอบสำคัญที่ช่วยให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้สะดวก และยังสร้างความน่าสนใจในการนำเสนอ ซึ่งองค์ประกอบหลักสามารถจำแนกได้ดังนี้ (เนตรนภิส คาดหมาย, 2564)

1) ส่วนหัว (Header) ส่วนหัวของเว็บไซต์ทำหน้าที่เป็นพื้นที่ในการแสดงสัญลักษณ์หรือเอกลักษณ์ของเว็บไซต์ เช่น โลโก้ ภาพแบนเนอร์ หรือสโลแกนที่สะท้อนถึงอัตลักษณ์และแนวทางขององค์กร รวมทั้งสามารถใช้เป็นพื้นที่ประชาสัมพันธ์หรือแนะนำภาพรวมของเว็บไซต์ เพื่อดึงดูดความสนใจและกระตุ้นให้ผู้ใช้งานติดตามเนื้อหาส่วนอื่นต่อไป

2) ส่วนท้าย (Footer) ส่วนท้ายของเว็บไซต์เป็นพื้นที่แสดงข้อมูลพื้นฐานและรายละเอียดที่สำคัญ อาทิ ข้อความเกี่ยวกับเว็บไซต์ ข้อความลิขสิทธิ์ (Copyright) รวมถึงข้อมูลการติดต่อ เช่น ที่อยู่ เบอร์โทรศัพท์ และอีเมล เพื่อสร้างความน่าเชื่อถือและอำนวยความสะดวกในการสื่อสารระหว่างผู้ใช้งานกับองค์กร

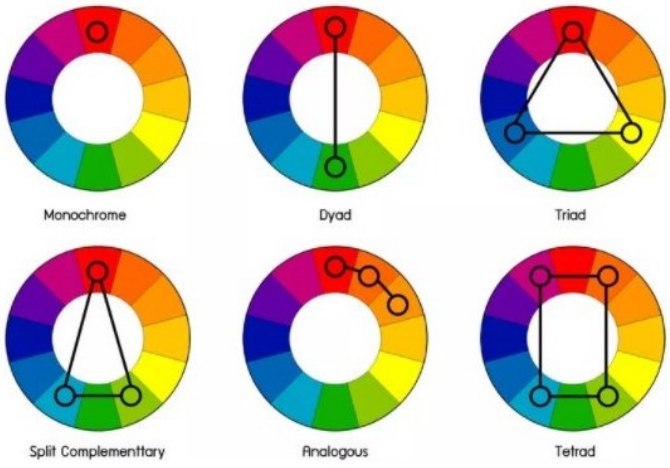
3) เนื้อหา (Contents) เนื้อหาเป็นหัวใจสำคัญของเว็บไซต์ ซึ่งครอบคลุมข้อมูล ข้อความ บทความ รูปภาพ ภาพเคลื่อนไหว เสียง และวิดีโอ โดยควรมีการจัดหมวดหมู่เนื้อหาอย่างชัดเจน เพื่อให้ผู้เข้าชมสามารถค้นหาและทำความเข้าใจข้อมูลได้ง่าย นอกจากนี้เนื้อหาควรมีคุณภาพและสอดคล้องกับวัตถุประสงค์ของเว็บไซต์

4) เมนูนำทาง (Navigator) เมนูนำทางเป็นองค์ประกอบที่ช่วยให้ผู้ใช้งานสามารถเข้าถึงข้อมูลหรือหมวดหมู่ต่าง ๆ ของเว็บไซต์ได้รวดเร็วและสะดวก โดยทั่วไปนิยมวางเมนูนำทางไว้บริเวณด้านบนของหน้าเว็บไซต์ หรือชิดด้านซ้าย/ขวา ทั้งนี้ การจัดวางควรมีความชัดเจนและเป็นระบบ เพื่อเพิ่มประสิทธิภาพในการค้นหาข้อมูล

5) ช่องว่าง (Space) ช่องว่างในเว็บไซต์มีบทบาทสำคัญในการจัดสมดุลของการนำเสนอข้อมูล ทำให้หน้าเว็บไซต์ดูเป็นระเบียบ ไม่หนาแน่นหรืออัดแน่นจนเกินไป การใช้พื้นที่ว่างอย่างเหมาะสมยังช่วยสร้างความสบายตาให้กับผู้เข้าชม และทำให้ประสบการณ์การใช้งานเว็บไซต์มีความราบรื่น

2.2.3.4 หลักในการเลือกใช้สีในหน้าเว็บไซต์ให้เหมาะสม โทนสีเป็นส่วนช่วยให้เว็บไซต์มีความสวยงาม น่าสนใจ ส่งผลต่อความรู้สึก และสามารถบ่งบอกถึงอารมณ์ไปยังผู้เข้าชมเว็บไซต์ โดยสีแต่ละสีจะมีความหมายและอารมณ์ที่ต่างกัน จึงควรเลือกใช้โทนสีให้เหมาะสมกับเว็บไซต์ เช่น สีส้ม เป็นสีแห่งความสร้างสรรค์ อบอุ่น สดใส มีสติปัญญา ความทะเยอทะยาน, สีแดง เป็นสีที่กระตุ้นระบบประสาทของเราได้รุนแรงที่สุด ให้ความรู้สึกเร้าใจ ตื่นเต้น ท้าทาย

ตื่นตัว, สีชมพู ให้ความรู้สึกอบอุ่น อ่อนโยน นุ่มนวล อ่อนหวาน ความรัก, สีเหลือง ให้ความรู้สึกแจ่มใส ความสดใส ความอบอุ่น ความร่าเริง ความสุขสว่างใส, สีเขียว เป็นสีที่เด่นที่สุดบนโลก ให้ความรู้สึกร่มเย็น สบายตา ผ่อนคลาย ปลอดภัย ทำให้เกิดความหวังและความสมดุล, สีน้ำเงิน เป็นสีที่สร้างความสุขุม เยือกเย็น หนักแน่นและละเอียดรอบคอบ น่าเชื่อถือ, สีฟ้า เป็นสีที่ให้ความรู้สึกสงบเยือกเย็น เป็นอิสระ ปลอดโปร่ง สบาย ปลอดภัย ใจเย็น สะอาด, สีม่วง เป็นสีแห่งผู้รู้ ให้ความรู้สึกมีเสน่ห์ น่าติดตาม เร้นลับ น่าค้นหา, สีขาว ให้ความรู้สึกบริสุทธิ์ สะอาด สดใส เบาบาง อ่อนโยน, สีทอง ให้ความรู้สึกความหรูหรา โอ่อ่า มีราคา สง่างาม เป็นต้น และควรคำนึงถึงความสอดคล้องของสีกับจุดประสงค์ของเว็บไซต์ สิ่งสำคัญที่สุดคือความเหมาะสมของการเลือกใช้สีได้สอดคล้องกับวัตถุประสงค์หรือเป้าหมายของเว็บไซต์ขึ้นอยู่กับความต้องการและควรเลือกใช้กลุ่มสีที่มีความกลมกลืนกัน (sogoodweb, 2563)



ภาพที่ 2.5 การเลือกใช้สีในหน้าเว็บไซต์

(ที่มา: https://www.wynnsoft-studio.com/Choosing-Color-in-Website-Design)

2.2.4 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

ภาพที่ 2.6 กระบวนการวิเคราะห์ข้อมูลด้วย (CRISP-DM)

(ที่มา: https://kamboonchob.medium.com/เข้าใจ-crisp-dm-ฉบับเร่งรัด-b0913050198f)

CRISP-DM ย่อมาจาก Cross-industry standard process for data miningซึ่งหมายถึง กระบวนการมาตรฐานที่ใช้สำหรับการทำเหมืองข้อมูล เพื่อทำการวิเคราะห์และนำไปใช้ประโยชน์ทางธุรกิจ มีอยู่ 6 ขั้นตอน คือ

2.2.4.1 การทำความเข้าใจธุรกิจ (Business Understanding) ขั้นตอนแรกมุ่งไปที่การทำความเข้าใจธุรกิจ ปัญหาและวัตถุประสงค์ของโครงการจากมุมมองทางธุรกิจ จากนั้นแปลงปัญหาให้อยู่ในรูปของโจทย์สำหรับการวิเคราะห์ข้อมูล และวางแผนการดำเนินงานเบื้องต้น

2.2.4.2 การทำความเข้าใจข้อมูล (Data Understanding) ขั้นตอนนี้เริ่มต้นด้วยการรวบรวมข้อมูล จากนั้นทำความเข้าใจ ตรวจสอบคุณภาพ และเลือกข้อมูลที่เก็บรวบรวมมาว่าจะใช้ข้อมูลใดบ้างในการวิเคราะห์

2.2.4.3 การเตรียมข้อมูล (Data Preparation) ขั้นตอนการเตรียมข้อมูล หมายถึงขั้นตอนทั้งหมดที่จะทำเพื่อให้ข้อมูลดิบที่เรารวบรวมมา กลายเป็นข้อมูลสมบูรณ์ที่พร้อมจะเข้าสู่โมเดลในขั้นตอนที่ 4 เช่น การสร้างตาราง การลบข้อมูลที่ไม่ต้องการออก การแปลงข้อมูลให้อยู่ในรูปแบบที่ต้องการ เป็นต้น

2.2.4.4 การสร้างโมเดล (Modeling) ในขั้นตอนนี้ เราจะเลือกและทดลองสร้างโมเดลหลายๆแบบที่น่าจะสามารถแก้ไขปัญหาที่ต้องการได้ จากนั้นค่อยๆปรับค่าพารามิเตอร์ในแต่ละโมเดล เพื่อให้ได้โมเดลที่เหมาะสมที่สุดมาใช้ในการแก้ไขปัญหา

2.2.4.5 การวัดประสิทธิภาพของโมเดล (Evaluation) ขั้นตอนการวัดประสิทธิภาพของโมเดลที่ได้จากขั้นตอนที่ 4 เพื่อวัดว่าโมเดลมีประสิทธิภาพเพียงพอต่อการนำไปใช้งานแล้วหรือไม่ ซึ่งโมเดลแต่ละประเภทก็จะมีตัววัดประสิทธิภาพที่แตกต่างกันออกไป

2.2.4.6 การนำโมเดลไปใช้งานจริง (Deployment) เป็นการนำโมเดลที่เหมาะสมที่สุดไปใช้งานจริง เพื่อวิเคราะห์และแก้ปัญหาที่ต้องการ

2.2.5 เทคนิค Random Forest

Random Forest เป็นอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning Algorithm) ที่ได้รับความนิยมอย่างแพร่หลาย โดยถูกพัฒนาขึ้นโดย Leo Breiman และ Adele Cutler หลักการสำคัญคือการรวมผลลัพธ์จากต้นไม้ตัดสินใจหลายต้น (Decision Trees) เพื่อให้ได้ผลลัพธ์สุดท้ายที่มีความแม่นยำมากขึ้น จุดเด่นของวิธีการนี้คือความสะดวกในการใช้งานและความยืดหยุ่น เนื่องจากสามารถประยุกต์ใช้ได้ทั้งในงานจำแนกประเภท (Classification) และการถดถอย (Regression) อัลกอริทึม Random Forest อยู่ภายใต้แนวคิดการเรียนรู้แบบกลุ่ม (Ensemble Learning) ซึ่งใช้ตัวจำแนกประเภทหลายตัว เช่น ต้นไม้ตัดสินใจ มารวมกันเพื่อสร้างผลลัพธ์ที่มีความน่าเชื่อถือมากกว่าแบบจำลองเพียงตัวเดียว วิธีการ Ensemble ที่เป็นที่รู้จักได้แก่ Bagging (Bootstrap Aggregation) และ Boosting โดยในปี ค.ศ. 1996 Leo Breiman ได้เสนอวิธีการ Bagging ซึ่งเป็นการสุ่มตัวอย่างข้อมูลจากชุดฝึกอบรมแบบมีการแทนที่ (With Replacement) เพื่อสร้างชุดข้อมูลย่อยจำนวนมาก แล้วนำไปฝึกแบบจำลองแยกกันอย่างอิสระ ผลลัพธ์จากแต่ละแบบจำลองจะถูกนำมาสรุปเป็นค่าเฉลี่ยในกรณีของการถดถอย หรือเลือกผลลัพธ์เสียงข้างมากในกรณีของการจำแนก วิธีนี้ช่วยลดความแปรปรวน (Variance) ของแบบจำลองโดยเฉพาะเมื่อข้อมูลมีสัญญาณรบกวนRandom Forest ถือเป็นการต่อยอดจากวิธีการ Bagging โดยเพิ่มการสุ่มคุณลักษณะ (Feature Randomness หรือ Random Subspace Method) เข้ามาในกระบวนการสร้างต้นไม้ตัดสินใจ ทำให้ต้นไม้แต่ละต้นไม่เหมือนกันและมีความสัมพันธ์ต่ำ (Low Correlation) ส่งผลให้ผลลัพธ์ของการรวมกันมีประสิทธิภาพสูงขึ้น ความแตกต่างสำคัญคือ ต้นไม้ตัดสินใจทั่วไปจะพิจารณาคุณลักษณะทั้งหมดที่มีอยู่ ในขณะที่ Random Forest จะสุ่มเลือกเฉพาะบางส่วนของคุณลักษณะเพื่อใช้ในการแบ่งโหนดแต่ละครั้ง ซึ่งช่วยลดความเสี่ยงจากการเกิด Overfitting

2.2.6 เทคนิค ARIMA (Auto Regressive Integrated Moving Average)

แนวคิดของแบบจำลอง ARIMAแบบจำลอง ARIMA (Autoregressive Integrated Moving Average) เป็นหนึ่งในเครื่องมือสำคัญที่ใช้ในการวิเคราะห์และพยากรณ์ข้อมูลอนุกรมเวลา (Time Series Data) ซึ่งเป็นข้อมูลที่ถูกเก็บรวบรวมตามลำดับเวลา เช่น รายเดือน รายไตรมาส หรือรายปี จุดเด่นของแบบจำลองนี้คือสามารถจัดการกับข้อมูลที่ไม่เป็นสถานี (Non–Stationary) ได้โดยอาศัยกระบวนการหาผลต่าง (Differencing) เพื่อปรับข้อมูลให้เป็นสถานี (Stationary) ก่อนที่จะนำไปสร้างแบบจำลอง

องค์ประกอบหลักของ ARIMA ประกอบด้วย 3 ส่วนหลัก ได้แก่

1) ส่วนอัตถิภาค (Autoregressive: AR) อธิบายความสัมพันธ์ของค่าปัจจุบันกับค่าของตัวเองในอดีต โดยใช้ตัวแปรที่มีค่าเลื่อนเวลา (Lagged Values) เข้ามาเป็นตัวทำนาย

$$Xt=c+\phi 1Xt-1+\phi 2Xt-2+\cdots+\phi pXt-p+\varepsilon t$$

x_t= ค่าปัจจุบัน

c = ค่าคงที่

$\emptyset_i$ = สัมประสิทธิ์ของตัวแปรอัตถิภาค (AR parameters)

ε_t= ค่าความคลาดเคลื่อน (White Noise)

2) ส่วนการอินทิเกรต (Integrated: I) เป็นการหาผลต่างของข้อมูล (Differencing) เพื่อลดผลกระทบจากแนวโน้ม (Trend) หรือฤดูกาล (Seasonality) ทำให้ข้อมูลมีค่าเฉลี่ยและความแปรปรวนคงที่

3) ส่วนค่าเฉลี่ยเคลื่อนที่ (Moving Average: MA) อธิบายความสัมพันธ์ของค่าปัจจุบันกับข้อผิดพลาด (Error Terms) ในอดีต ซึ่งช่วยปรับปรุงความแม่นยำของการพยากรณ์

$$X_t = c + \varepsilon_t + \theta_1\ \varepsilon_t\text{–}1+\theta_(2\varepsilon_t\ ) - 2 + \cdots + \theta_q\ \epsilon_t - q$$

θ_j = สัมประสิทธิ์ของตัวแปรค่าเฉลี่ยเคลื่อนที่ (MA parameters)

แบบจำลอง ARIMA จะรวมทั้ง AR และ MA เข้าด้วยกัน พร้อมการหาผลต่าง (Differencing) ระดับ d

$$\Delta\text{^}d\ x_t = c + \emptyset_1\ \Delta\text{^}d\ x_(t-1) + \cdots + \emptyset_p\ \Delta\text{^}d\ x_(t-p+\varepsilon_t\ )+\theta_1\ \ 〚\theta\varepsilon〛_(t-1) + \cdots + \theta_q\ \ 〚\theta\varepsilon〛_(t-q)$$

โดยที่

$\Delta \hat{} d\ x_t$= ข้อมูลที่ถูกหาผลต่าง ddd ครั้ง เพื่อทำให้เป็น Stationary หาก d = 0d = 0d = 0 จะกลายเป็น ARMA(p,q) หาก p=0 และ q=0 จะเหลือเพียงการหาผลต่างของข้อมูล

2.2.7 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression) คือวิธีการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์ระหว่างตัวแปรต้น (หนึ่งตัวแปรหรือมากกว่า) และตัวแปรตาม โดย Simple Linear Regression หรือการวิเคราะห์ข้อมูลถดถอยเชิงเส้นอย่างง่ายเป็นการวิเคราะห์ข้อมูลเมื่อมีตัวแปรต้นเพียง 1 ตัวแปร และ Multiple Linear Regression หรือการวิเคราะห์ข้อมูลถดถอยเชิงเส้นพหุคูณเป็นการวิเคราะห์ข้อมูลเมื่อมีตัวแปรต้นหลายตัว (datainvestigator, 2567)

2.2.7.1 การวิเคราะห์ข้อมูลถดถอยพหุคูณเชิงเส้น (Multiple Linear Regression) คือ แบบจำลองการพัฒนาต่อยอดจากแบบจำลอง Linear Regression เพื่อให้สามารถใช้ตัวแปรอิสระได้หลายตัว (X_1, X_2,...,X_n) ในการทำนายตัวแปรตามเพียงหนึ่งตัว แบบจำลองนี้ยังคงพิจารณาความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตาม แต่แทนที่จะเป็นเส้นตรงในสองมิติจะเป็นการค้นหาระนาบ (plane) ของคำตอบหรือไฮเปอร์เพลน (hyperplane) ที่เหมาะสมที่สุดในพื้นที่ที่มีหลายมิติตามจำนวนของตัวแปรสมการพื้นฐานสำหรับแบบจำลองการถดถอยเชิงเส้นพหุคูณแสดงได้ดังสมการต่อไปนี้

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

เมื่อ

Y คือ ตัวแปรตาม

$X_1, X_2, \ldots, X_n$ คือ ตัวแปรอิสระ n ตัว

β_0 คือ จุดตัดแกน Y

$\beta_1, \ 〖\beta〗_2, \ldots \ 〖,\beta〗_n$ คือ สัมประสิทธิ์ การถดถอยสำหรับตัวแปรอิสระแต่ละตัว

ε คือ ค่าความคลาดเคลื่อน (กรรณิการ์ จะกอ และปิยพันธ์ สุวรรณเวช, 2568)

2.2.8 การปรับเรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing) เป็นวิธีการพยากรณ์ข้อมูลอนุกรมเวลาแบบตัวแปรเดียว โดยสร้างค่าพยากรณ์จากค่าเฉลี่ยถ่วงน้ำหนักของข้อมูลในอดีต น้ำหนักของข้อมูลจะลดลงตามความเก่าของข้อมูลในลักษณะเอกซ์โปเนนเชียล ส่งผลให้ข้อมูลล่าสุดมีอิทธิพลต่อค่าพยากรณ์มากกว่าข้อมูลเก่า แนวคิดดังกล่าวยังสามารถขยายไปสู่แบบจำลองที่รองรับองค์ประกอบ แนวโน้ม (Trend) และฤดูกาล (Seasonality) ได้ แนวคิดการปรับ

เรียบแบบเอกซ์โปเนนเชียล (Exponential smoothing) ได้รับการพัฒนาตั้งแต่ช่วงทศวรรษ 1950 และต่อมาได้รับความนิยมอย่างแพร่หลายในงานวิเคราะห์ เนื่องจากสามารถคำนวณได้รวดเร็วและให้ผลพยากรณ์ที่มีความแม่นยำในหลายสาขา โดยเฉพาะงานด้านอุตสาหกรรม อีกทั้งยังถูกนำไปประยุกต์ใช้ในการประมวลผลสัญญาณ เพื่อปรับเรียบข้อมูลและลดสัญญาณรบกวนความถี่สูง โดยวิธีการปรับเรียบแบบเอกซ์โปเนนเชียลสามารถแบ่งเป็นหลายรูปแบบ ได้แก่

2.2.8.1 การปรับเรียบแบบเอกซ์โปเนนเชียลอย่างง่าย (Simple Exponential Smoothing)

2.2.8.2 การปรับเรียบแบบสองชั้น (Double Exponential Smoothing)

2.2.8.3 การปรับเรียบแบบสามชั้น หรือวิธีโฮลต์–วินเทอร์ส (Holt–Winters) ซึ่งแต่ละรูปแบบถูกออกแบบให้เหมาะสมกับลักษณะข้อมูลที่แตกต่างกัน โดยเฉพาะการมีหรือไม่มีแนวโน้มและฤดูกาล (Jim Frosts, 2564)

2.2.9 ทฤษฎีภาษาที่ใช้สำหรับการพัฒนาเว็บไซต์

ภาษาที่ใช้สำหรับพัฒนาเว็บไซต์ที่ได้รับความนิยมในปัจจุบัน ได้แก่ PHP, JSP, C# (ASP.NET), Python, Ruby และภาษาอื่น ๆ อีกหลายภาษา โดยภาษาเหล่านี้มักทำงานร่วมกับ HTML, CSS, JavaScript และเทคนิคต่าง ๆ เช่น AJAX เพื่อสร้างระบบเว็บไซต์ให้สามารถโต้ตอบและให้บริการได้อย่างมีประสิทธิภาพ เนื่องจากมีทางเลือกของภาษาสำหรับพัฒนาเว็บไซต์จำนวนมาก การตัดสินใจเลือกภาษาในบางครั้งอาจเกิดจากความชอบส่วนบุคคล ความนิยมของกระแสเทคโนโลยีใหม่ หรือความต้องการด้านความรวดเร็วในการพัฒนา อย่างไรก็ตาม การเลือกภาษาโดยไม่พิจารณาความเหมาะสมอย่างรอบด้าน อาจนำไปสู่ปัญหาในการพัฒนาระบบ เช่น ภาษาไม่สอดคล้องกับลักษณะงาน ภาษาไม่มีมาตรฐานที่ชัดเจน หรือระบบนิเวศ (Ecosystem) และเครื่องมือสนับสนุนมีจำกัด ดังนั้น จึงควรมีหลักเกณฑ์ในการคัดเลือกภาษาเพื่อให้เหมาะสมกับผู้พัฒนาและลักษณะงาน โดยสามารถพิจารณาได้ดังนี้

2.2.9.1 เลือกภาษาที่มีผู้ใช้งานและชุมชนผู้พัฒนาขนาดใหญ่ภาษาที่มีผู้ใช้งานจำนวนมากมักมีองค์ความรู้และแนวทางแก้ไขปัญหาที่เผยแพร่ไว้เป็นจำนวนมาก ทำให้ผู้พัฒนาสามารถค้นคว้า เรียนรู้ และแก้ไขปัญหาที่เกิดขึ้นได้สะดวกยิ่งขึ้น อีกทั้งยังมักมี Framework แนวคิด และเครื่องมือช่วยพัฒนาที่หลากหลาย ซึ่งช่วยเพิ่มประสิทธิภาพในการพัฒนา ตัวอย่างเช่น ภาษา PHP มี Framework เช่น Laravel และ Yii เป็นต้น

2.2.9.2 เลือกภาษาที่มีผู้สนับสนุนหรือองค์กรพัฒนาอย่างชัดเจนนอกจากความนิยมแล้ว ควรพิจารณาผู้พัฒนาและผู้สนับสนุนของภาษานั้น ๆ เนื่องจากภาษาที่มีองค์กรสนับสนุนมัก

ได้รับการพัฒนาและปรับปรุงอย่างต่อเนื่อง รวมถึงมีการแก้ไขข้อบกพร่องและอัปเดตด้านความปลอดภัยสม่ำเสมอ ตัวอย่างเช่น C# มี Microsoft เป็นผู้พัฒนาและสนับสนุน ส่วน Java มี Oracle ดูแลพัฒนา และ Python มีองค์กรและชุมชนขนาดใหญ่สนับสนุน ทำให้มั่นใจได้ว่าภาษามีแนวโน้มพัฒนาอย่างต่อเนื่องในระยะยาว

2.2.9.3 เลือกภาษาที่เหมาะสมกับลักษณะงานและความพร้อมของผู้พัฒนา แม้ภาษาจะเป็นที่นิยมและมีผู้สนับสนุนที่ดี แต่ไม่ได้หมายความว่าจะเหมาะสมกับทุกประเภทงานเสมอไป ภาษาบางภาษาอาจมีความซับซ้อนสูง หรือมีข้อจำกัดที่ส่งผลต่อการพัฒนาระบบ ดังนั้นควรพิจารณาความเหมาะสมของภาษาโดยเปรียบเทียบข้อดี –ข้อเสีย ความยากง่ายในการเรียนรู้ ความสามารถที่ตอบโจทย์งาน และความสอดคล้องกับทรัพยากรที่มีเพื่อให้การพัฒนาเป็นไปอย่างมีประสิทธิภาพและลดความเสี่ยงของปัญหาในอนาคต (บริษัท โปรซอฟท์ เว็บ จำกัด, 2559)

2.3 วรรณกรรมที่เกี่ยวข้อง

พีรธัช ถนอมจิตร์ (2564) ได้ทำการศึกษาเกี่ยวกับการส่งออกผลิตภัณฑ์ทูน่ากระป๋องของไทย โดยมีวัตถุประสงค์เพื่อ (1) วิเคราะห์สถานการณ์การส่งออกผลิตภัณฑ์ทูน่ากระป๋อง รวมถึงปัญหาและอุปสรรคของอุตสาหกรรมในประเทศไทย (2) ศึกษาความได้เปรียบโดยเปรียบเทียบในการส่งออกผลิตภัณฑ์ทูน่ากระป๋องของไทยไปยังตลาดผู้นำเข้าหลัก เมื่อเทียบกับประเทศคู่แข่ง (3) พยากรณ์การส่งออกผลิตภัณฑ์ทูน่ากระป๋องของไทยในช่วงปี พ.ศ. 2563 –2565 โดยใช้แบบจำลอง ARIMAข้อมูลที่ใช้ในการศึกษาคือข้อมูลทุติยภูมิจากสถิติการส่งออกสินค้าตามพิกัดศุลกากร 6 หลัก (160414) ย้อนหลัง 15 ปี (พ.ศ. 2548–2562) โดยใช้การวิเคราะห์เชิงพรรณนาร่วมกับการวิเคราะห์ดัชนีความได้เปรียบโดยเปรียบเทียบที่ปรากฏ (RCA) และดัชนีพลวัตความได้เปรียบโดยเปรียบเทียบ (DRCA) นอกจากนี้ ยังใช้แบบจำลอง ARIMA ในการพยากรณ์แนวโน้มการส่งออกผลการศึกษา พบว่า ประเทศไทยมีค่าความได้เปรียบโดยเปรียบเทียบสูงกว่าค่ามาตรฐาน (มากกว่า 1) อย่างต่อเนื่องตลอดช่วงเวลาที่ศึกษา และมีค่าความได้เปรียบมากกว่าคู่แข่งหลัก ได้แก่ จีน อินโดนีเซีย และฟิลิปปินส์ โดยเฉพาะในการส่งออกไปยังตลาดสำคัญ เช่น สหรัฐอเมริกา ออสเตรีย และญี่ปุ่น เมื่อพิจารณาศักยภาพการแข่งขัน พบว่าไทยอยู่ในสถานการณ์ “ดาวรุ่งพุ่งแรง” หลายปี แสดงถึงการมีศักยภาพสูงในตลาดโลก และจากผลการพยากรณ์พบว่าแนวโน้มการส่งออกผลิตภัณฑ์ทูน่ากระป๋องในช่วงปี พ.ศ. 2563 – 2565 จะมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง

ปิยะธิดา เฮืองฮุง (2567) ได้วิจัยเกี่ยวกับการพยากรณ์ปริมาณการส่งออกน้ำตาลรายเดือนของประเทศไทยด้วยตัวแบบอนุกรมเวลา โดยนำตัวแบบ ARIMA ใช้วิธีการเขียนสัญลักษณ์ตัวแบบอารีมาเชิงคูณในการสร้างตัวแบบ ข้อมูลอนุกรมเวลาต้องมีลักษณะกระบวนการคงที่ สารมารถทดสอบกระบวนการคงที่ด้วยการทดสอบ Unit Root Test วิธีการที่ได้รับความนิยม คือ วิธีทดสอบ Augmented Dickey–Fuller Test (ADF) ทำการแปลงข้อมูลข้อมูลด้วยการหาผลต่าง (Difference) และนำมาเปรียบเทียบประสิทธิภาพกับตัวพยากรณ์ Exponential Smoothing และ Holt–Winters โดยเปรียบเทียบด้วยเกณฑ์ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย RMSE ต่ำสุด ผลการวิจัยพบว่า ตัวแบบARIMA(1,1,0)(0,1,1) มีค่า RMSE ต่ำสุดเท่ากับ 44.5986

จิรฐา คำรูญ (2564) งานวิจัยฉบับนี้ใช้เทคนิค SARIMA ในการพยากรณ์ปริมาณและราคาส่งออกน้ำมันปาล์มดิบของไทย โดยใช้กระบวนการวิเคราะห์อนุกรมเวลา การตรวจสอบความนิ่งของข้อมูล และการประเมินโมเดลด้วยสถิติต่างๆ เช่น AIC, SIC, RMSE, MAPE ผลการศึกษาช่วยยืนยันถึงความเหมาะสมของ SARIMA สำหรับข้อมูลที่มีฤดูกาลและแนวคิดการเลือกโมเดลโดยใช้ตัวชี้วัดหลากหลาย อีกทั้งยังเน้นขั้นตอนการเตรียมข้อมูลและการทดสอบความนิ่งเพื่อให้ได้แบบจำลองที่แม่นยำที่สุด ข้อค้นพบนี้สามารถนำมาประยุกต์ใช้ในการวิเคราะห์และพยากรณ์ข้อมูลส่งออกสินค้าประมงของไทยที่มีลักษณะฤดูกาลเช่นเดียวกันในงานวิจัย

ธนาภรณ์ ผ่องศรี (2563) ได้วิจัยเกี่ยวกับการพยากรณ์ปริมาณการส่งออกข้าวไทยไปจีน กรณีศึกษาเชิงประจักษ์ด้วยวิธีการทางอนุกรมเวลา โดยศึกษาใช้อนุกรมเวลารายเดือน โดยใช้แบบจำลอง SARIMA(p,d,q) (P,D,Q)s ประกอบด้วย 4 ขั้นตอนขั้นตอนที่ 1) การทดสอบความหยุดนิ่งของข้อมูล (Stationary) ด้วยวิธี Unit root test เพื่อทดสอบว่าข้อมูลที่จะนำไปเข้ากระบวนการประมวลต้องมีความหยุดนิ่ง โดยทดสอบว่าข้อมูลมีความหยุดนิ่งหรือข้อมูลไม่มีความหยุดนิ่ง (Non–stationary) หรือไม่ ถ้าพบว่าข้อมูลยังไม่มีความหยุดนิ่งของข้อมูล จะต้องนำข้อมูลไปหาผลต่างอันดับที่ 1 (Yt–Yt–1) และถ้าข้อมูลยังไม่มีความหยุดนิ่ง ต้องนำข้อมูลไปหาผลต่างในอันดับที่สูงขึ้นไป จนกว่าข้อมูลจะมีความหยุดนิ่ง ซึ่งการทดสอบความหยุดนิ่งของข้อมูลในการศึกษาครั้งนี้มี 2 แบบ คือ แบบไม่มีฤดูกาล (Non–Seasonality) และแบบมีฤดูกาล (Seasonality) ขั้นตอนที่ 2) การประมาณค่าพารามิเตอร์ด้วยวิธี Maximum Likelihood เพื่อกำหนดแบบจำลอง SARIMA(p,d,q) (P,D,Q)s เบื้องต้นที่เหมาะสม และเลือกแบบจำลองที่ไม่มีปัญหาสหสัมพันธ์ในตัวเองของตัวคลาดเคลื่อน (Autocorrelation)ขั้นตอนที่ 3) ตรวจสอบความเหมาะสมของตัวแบบพยากรณ์/ประเมินผลกระทบ เพื่อให้ได้แบบจำลองที่มีคุณสมบัติเป็นตัวประมาณค่าที่ดี โดยถ้าหาได้

มากกว่า 1 แบบจำลอง จะเลือกแบบจำลองที่ดีที่สุด ซึ่งพิจารณาได้จากสถิติ AIC (Akaike Information Criterion), SIC (Schwarz Information Criterion), ค่าเฉลี่ยของรากที่สองของกำลังสองของความคลาดเคลื่อน (Root Mean Squared Error: RMSE) และค่าร้อยละ ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) ที่มีค่าต่ำสุด ขั้นตอนที่ 4) การพยากรณ์จากการคัดเลือกแบบจำลองที่มีความเหมาะสม แล้วนำแบบจำลองไปประมวลผลหาค่าพยากรณ์ปริมาณการส่งออกข้าวไทยไปจีนหรือประเมินผลกระทบต่อไปผลการวิจัยพบว่า ตัวแบบพยากรณ์ SARIMA (p,d,q) (P,D,Q)s ที่เหมาะสมต่อการพยากรณ์มากที่สุดคือ ตัวแบบพยากรณ์ SARIMA (2,0,4) (0,1,1)

เฉลิมวุฒิ คำเมือง และวชิรารักษ์ โอรสรัมย์ (2565) ได้วิจัยเกี่ยวกับการพยากรณ์ราคาข้าวเปลือกหอมมะลิ ด้วยเทคนิคการทำเหมืองข้อมูล โดยมีวัตถุประสงค์เพื่อศึกษาแบบจำลองทางคณิตศาสตร์เพื่อทำนายราคาข้าวเปลือกหอมมะลิ ข้อมูลที่ใช้ในการศึกษาราคาข้าวเปลือกหอมมะลิเฉลี่ยรายเดือน โดยใช้ตัวแบบพยากรณ์ วิธีบอกซเจนกินส์ (Box–Jenkins) วิธีแบ็กกิ้ง (Bagging) และ วิธีแรนดอมฟอเรสต์ (Random Forest) โดยวิธีแรนดอมฟอเรสต์ (Random Forest) ถูกควบคุมด้วย 3 ปัจจัยคือ 1.)ต้นไม้แต่ละต้นจะถูกสอน (Train) โดยการใช้เซตย่อยจากข้อมูลตัวอย่าง 2.) เมื่อต้นไม้โตขึ้นจะสามารถค้นหาโนด (Node) แต่ละโนดที่อยู่ในกิ่งที่ดีที่สุดของต้นไม้โดยใช้การสุ่มเลือกคุณสมบัติจาก N คุณสมบัติ 3.) ต้นไม้แต่ละต้นจะไม้มีการตัดออก แต่จะปล่อยให้ต้นไม้โตขึ้นไปเรื่อยๆ จนได้ผลลัพธ์ที่ดีที่สุดหลังจากการสร้างป่าแล้วทำการให้คะแนนโหวต โดยต้นไม้ภายในป่า หากต้นไม้ต้นใดได้คะแนนโหวตสูงสุดจะนำเอาต้นไม้นั้นออกมาสร้างเป็นโมเดล ผลการวิจัยพบว่าจากวิธีการพยากรณ์ที่เหมาะสมที่สุดคือ วิธีแรนดอมฟอเรสต์ (Random Forest)

อุรชา จัทรภา (2563) ได้ศึกษาการพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์ และส่วนประกอบของประเทศไทย โดยมีวัตถุประสงค์เพื่อค้นหาและนำเสนอตัวแบบที่เหมาะสมสำหรับการพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์และส่วนประกอบโดยรวมของประเทศไทย งานวิจัยนี้ได้ประยุกต์ใช้เทคนิค Exponential Smoothing ในการพยากรณ์มูลค่าการส่งออกของประเทศไทย โดยใช้ข้อมูลรายเดือน และพบว่าวิธีดังกล่าวสามารถสะท้อนแนวโน้มและทิศทางของข้อมูลได้อย่างเหมาะสม สำหรับพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์และส่วนประกอบโดยรวมของประเทศไทย และมูลค่าการส่งออกรถยนต์ อุปกรณ์และส่วนประกอบไปประเทศคู่ค้าที่มีมูลค่าการส่งออกจากประเทศไทยมากที่สุด 2 อันดับแรก ได้แก่ ประเทศออสเตรียและประเทศญี่ปุ่นด้วยวิธีอนุกรมเวลา วิธีการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ วิธีการเรียนรู้ของเครื่อง และวิธีพยากรณ์แบบผสม โดยใช้ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ในการ

เปรียบเทียบความแม่นยำของตัวแบบพยากรณ์ ข้อมูลที่ใช้ในการศึกษาวิธีการพยากรณ์ของงานวิจัยนี้เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม พ.ศ. 2551 ถึงเดือนธันวาคม พ.ศ. 2563 เป็นจำนวนทั้งหมด 156 เดือน วิธีที่ใช้ในการพยากรณ์ ได้แก่ Moving Average, Holt–Winters, SARIMA, Multiple Linear Regression, Artificial Neural Networks, Support Vector Regression, XGBoost, LSTM และวิธีพยากรณ์แบบผสม ผลการศึกษาพบว่า ตัวแบบพยากรณ์ที่มีความแม่นยำและเหมาะสมมากที่สุดสำหรับพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์และส่วนประกอบโดยรวมของประเทศไทย มูลค่าการส่งออกรถยนต์ อุปกรณ์และส่วนประกอบของประเทศไทยไปประเทศออสเตรียและประเทศญี่ปุ่น ได้แก่ ตัวแบบพยากรณ์ผสมวิธี LSTM–XGB วิธี SARIMA–XGB และวิธี LSTM–SARIMA มีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยเท่ากับ 6.63% 15.40% และ 6.27%

2.4 บทสรุป

จากการศึกษาทบทวนแนวคิด ทฤษฎี เครื่องมือ และวรรณกรรมที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลเพื่อการพยากรณ์มูลค่าการส่งออกสินค้าประมงของไทยไปยังตลาดหลัก โดยใช้เทคนิคเหมืองข้อมูล พบว่า การวิเคราะห์ข้อมูลในลักษณะนี้จำเป็นต้องอาศัยความเข้าใจในกระบวนการวิเคราะห์ข้อมูลอย่างเป็นระบบ โดยเฉพาะกระบวนการ CRISP–DM ซึ่งเป็นแนวทางมาตรฐานที่ได้รับการยอมรับในวงการเหมืองข้อมูล และสามารถนำเทคนิคทางสถิติและการเรียนรู้ของเครื่องมือ เช่น การประยุกต์ใช้แบบจำลองอนุกรมเวลาและเทคนิคเหมืองข้อมูลในบริบทของข้อมูลการส่งออก เนื่องจากงานวิจัยก่อนหน้าหลายฉบับได้แสดงให้เห็นว่าเทคนิคอย่าง ARIMA, Exponential Smoothing รวมถึง Random Forest สามารถอธิบายรูปแบบข้อมูลและให้ผลพยากรณ์ที่มีประสิทธิภาพในงานเศรษฐกิจและการค้าระหว่างประเทศได้ ดังนั้น งานวิจัยนี้จึงนำองค์ความรู้ดังกล่าวมาเป็นฐานในการกำหนดแนวทางการสร้างแบบจำลองและกรอบการประเมินผล และต่อยอดสู่การนำเสนอผลลัพธ์ในรูปแบบเว็บไซต์ ซึ่งเป็นสื่อกลางสำหรับเผยแพร่ข้อมูลและผลการพยากรณ์ให้ผู้ใช้งานเข้าถึงได้สะดวก และการสรุปตัวชี้วัดความแม่นยำของแบบจำลอง เพื่อเพิ่มประสิทธิภาพในการสื่อสารผลการศึกษาให้เกิดประโยชน์