

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของปัญหา

แรงงานไทยเริ่มเดินทางไปยังต่างประเทศจำนวนมาก โดยใช้บริการจัดหางาน ตั้งแต่ปี 2518 เป็นต้นมา โดยเริ่มจากภูมิภาคตะวันออกกลางภายหลังการตั้งกลุ่มประเทศผู้ค้าน้ำมัน (OPEC) ทำให้กลุ่มประเทศดังกล่าวมั่งคั่งและเริ่มการพัฒนาประเทศด้วยการลงทุนในโครงสร้างพื้นฐานขนาดใหญ่ ประกอบกับนโยบายจัดส่งแรงงานไทยไปทำงานในต่างประเทศในช่วงแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติฉบับที่ 5 (2525-2529) ทำให้ชาวไทยจำนวนมากเดินทางไปทำงานในตะวันออกกลางอย่างต่อเนื่อง โดยส่วนใหญ่เป็นแรงงานไร้ฝีมือในภาคการผลิต ก่อสร้างและแรงงานทั่วไป แต่ได้เกิดปัญหาการหลอกลวงของนายหน้าและบริษัทจัดหางานจำนวนมาก ทำให้มีการตรา พ.ร.บ. จัดหางานและคุ้มครองคนหางาน พ.ศ. 2528 ขึ้น เพื่อแก้ปัญหาดังกล่าว แต่กลุ่มมิชชันนารีก็ได้มีการพัฒนากลไกหลายรูปแบบ ทำให้แรงงานจำนวนมากไม่จำเป็นต้องตกเป็นเหยื่อ ถูกหลอกลวง หรือต้องกู้หนี้ยืมสินมาเป็นค่านายหน้า และต่อมาก็ได้มีการใช้สื่อออนไลน์เพื่อหลอกลวงเหยื่อแรงงานไทยที่ไปทำงานในต่างประเทศเริ่มชะลอลงในทศวรรษที่ 2550 เป็นต้นมา โดยแรงงานไทยเริ่มลดลงจากจำนวน 161,917 คนในปี 2550 เหลือเพียง 134,101 คนในปี 2555 และ 114,437 คนในปี 2559 โดยเฉพาะในตลาดตะวันออกกลาง เนื่องมาจากปัจจัยหลายประการ เช่น ปัญหาขาดแคลนแรงงานในประเทศ ทำให้แรงงานบางส่วนไม่จำเป็นต้องไปทำงานในต่างประเทศ ปัญหาการแข่งขันจากแรงงานในกลุ่มประเทศเอเชียใต้ ที่ค่าแรงถูกกว่าไทย และปัญหาทางการเมืองในตะวันออกกลาง อย่างไรก็ตาม แม้ตลาดตะวันออกกลางจะชะลอตัวลง แต่ตลาดในภูมิภาคเอเชียตะวันออกได้เปิดกว้างมากขึ้น โดยเฉพาะไต้หวัน เกาหลีใต้ และญี่ปุ่น ทำให้ในปัจจุบันมีแรงงานไทยจำนวนราวปีละหนึ่งแสนคนเท่านั้น ที่เดินทางไปทำงานในต่างประเทศกรมการจ้างงานกระทรวงแรงงานได้รายงานว่ามีแรงงานไทยที่เดินทางไปทำงานต่างประเทศรุ่นใหม่จำนวน 88,154 คน ในปี 2565 โดยเป็นเพศชายถึง 68,460 คน เพศหญิง 19,694 คน โดยประเทศปลายทางสำคัญได้แก่ ไต้หวัน จำนวน 25,372 คน อิสราเอล 9,417 คน เกาหลีใต้ 8,571 คน ญี่ปุ่น 8,147 คน

มาเลเซีย 4,402 คน และสิงคโปร์ 3,064 คน แรงงานไทยที่ไปทำงานต่างประเทศเป็นเรื่องที่สำคัญมากในปัจจุบันเพราะช่วยลดปัญหาการว่างงานและเพิ่มรายได้ให้กับครอบครัวของแรงงานเองแต่ก็มีหลายปัจจัยที่ส่งผลต่อการที่แรงงานจะได้รับอนุญาตให้ไปทำงานต่างประเทศ เช่น เพศ จังหวัดต้นทาง วิธีการเดินทาง และประเทศปลายทาง รวมถึงเรื่องอัตราค่าแรงในไทยที่ต่ำ ทำให้แรงงานอยากหางานทำต่างประเทศมาก โดยการเคลื่อนย้ายแรงงานไปทำงานต่างประเทศเป็นปรากฏการณ์ที่มีบทบาทสำคัญต่อเศรษฐกิจและสังคมของประเทศไทย โดยเฉพาะอย่างยิ่งในยุคที่อัตราค่าแรงภายในประเทศอยู่ในระดับต่ำ ส่งผลให้แรงงานไทยมีแรงจูงใจในการไปทำงานในต่างประเทศเพิ่มขึ้นทั้งนี้การได้รับอนุญาตและผ่านกระบวนการเดินทางอย่างถูกต้องเป็นปัจจัยสำคัญที่กำหนดโอกาสและรูปแบบการเคลื่อนย้ายของแรงงานในขณะเดียวกัน ปัจจัยเชิงประชากร เช่น เพศ จังหวัดต้นทาง และรูปแบบการเดินทาง (วิธีการเดินทาง) รวมถึงประเทศปลายทางมีบทบาทสำคัญต่อการอนุมัติแรงงานออกไปทำงานต่างประเทศ การวิจัยนี้จึงสนใจศึกษาการพัฒนาแบบจำลองการพยากรณ์แรงงานไทยในต่างประเทศด้วยเทคนิคเหมืองข้อมูลเพื่อพัฒนาแบบจำลองการพยากรณ์แรงงานไทยในต่างประเทศด้วยเทคนิคเหมืองข้อมูลและเผยแพร่ข้อมูลผ่านเว็บไซต์ โดยการพยากรณ์แนวโน้มและรูปแบบการเคลื่อนย้ายแรงงานอย่างมีประสิทธิภาพซึ่งจะช่วยให้การวางนโยบายและการจัดการแรงงานในระดับชาติให้เหมาะสมและยั่งยืน

1.2 วัตถุประสงค์

1.2.1 พัฒนาแบบจำลองการพยากรณ์แรงงานไทยในต่างประเทศด้วยเทคนิคเหมืองข้อมูล

1.2.2 เพื่อเผยแพร่ข้อมูลผ่านเว็บไซต์

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1.3.1 ได้แบบจำลองที่สามารถใช้ในการพยากรณ์จำนวนแรงงานไทยในต่างประเทศ

1.3.2 ได้เผยแพร่ข้อมูลแรงงานไทยเดินทางออกไปทำงานในต่างประเทศบนเว็บไซต์

1.4 ขอบเขต

1.4.1 ขอบเขตการวิเคราะห์ข้อมูล ดำเนินการวิเคราะห์ข้อมูลตามกระบวนการมาตรฐานของการทำเหมืองข้อมูล (Cross – Industry Standard Process for Data Mining: CRISP – DM) ประกอบด้วย 6 ขั้นตอน ดังนี้

1.4.2 การทำความเข้าใจปัญหา (Problem Understanding) ช่วงปี 2564–2567 จำนวนแรงงานไทยที่เดินทางไปทำงานต่างประเทศมีการเปลี่ยนแปลงขึ้นลงตลอด ขึ้นอยู่กับ

ปัจจัยหลายอย่าง เช่น เพศ อายุ จังหวัดต้นทาง อาชีพ และประเทศที่ไปเพื่อตอบโจทย์ จึงสร้างแบบจำลองพยากรณ์จำนวนแรงงานรายเดือน เพื่อช่วยวางแผนและเตรียมการ

1.4.3 การทำความเข้าใจข้อมูล (Data Understanding) ข้อมูลที่นำมาใช้ในการพัฒนาแบบจำลองการพยากรณ์แรงงานไทยในต่างประเทศเป็นข้อมูลรายเดือนย้อนหลัง 48 เดือน ตั้งแต่ปี พ.ศ. 2564–2567 จำนวน 324,497 ระเบียบ

1.4.4 การเตรียมข้อมูล (Data Preparation)

1.4.4.1 นำเข้าข้อมูลสำหรับวิเคราะห์ข้อมูลเป็นข้อมูลรายเดือนย้อนหลัง 48 เดือน ตั้งแต่ปี พ.ศ. 2564–2567 จำนวน 324,497 ระเบียบ

1.4.4.2 ตรวจสอบค่าที่หายไป (Missing Values) แก้ไขค่าผิดปกติ เช่น อายุ 200 ปี หรือจำนวนแรงงานเยอะเกินจริง ทำการลบออกหรือปรับให้ถูกต้อง

1.4.4.3 กำหนดและตรวจสอบความถูกต้องของประเภทข้อมูล เช่น วันที่ให้อยู่ในรูปแบบ วัน-เดือน-ปี (Date) ตัวเลข (Numeric) และข้อมูลหมวดหมู่ (Categorical) เช่น เพศ ประเทศปลายทาง และวิธีการเดินทาง

1.4.4.4 สร้างตัวแปรใหม่ (Lag Features) สำหรับโมเดล Machine Learning เช่น Random Forest โดยใช้ข้อมูลจำนวนแรงงานย้อนหลัง 1 เป็นตัวแปรอิสระ

1.4.4.5 การแบ่งข้อมูลออกเป็น Training set (70%) และ Testing set (30%) รวมถึงการใช้ Walk-forward Validation เพื่อการประเมินผลในลักษณะอนุกรมเวลา

1.4.5 การสร้างแบบจำลอง (Modeling)

1.4.5.1 Multiple Linear Regression (MLR) ใช้วิธีนำข้อมูลแรงงานในอดีตมาสร้างเป็นสมการเส้นตรงทางคณิตศาสตร์ เพื่อหาแนวโน้มว่าในอนาคตจำนวนแรงงานจะเพิ่มขึ้นหรือลดลงเท่าไรตามระยะเวลาที่เปลี่ยนไป

1.4.5.2 Gradient Boosted Trees (GBT) ใช้วิธีสร้างต้นไม้ตัดสินใจ (Decision Tree) ขึ้นมาทีละต้น โดยต้นใหม่ที่สร้างขึ้นมาจะทำหน้าที่คอยแก้ไขจุดผิดพลาดของต้นก่อนหน้าไปเรื่อยๆ เพื่อให้การพยากรณ์จำนวนแรงงานมีความแม่นยำที่สุด

1.4.5.3 Random Forest Regression ใช้วิธีจำลองต้นไม้ตัดสินใจขึ้นมาจำนวนมาก (หลายร้อยต้น) ให้ช่วยกันวิเคราะห์ข้อมูลแรงงานย้อนหลัง แล้วนำคำตอบของทุกต้นมาหาค่าเฉลี่ยร่วมกัน เพื่อให้ได้ผลลัพธ์การพยากรณ์ที่น่าเชื่อถือและไม่เอนเอียง

1.4.5.4 k-Nearest Neighbors (k-NN) ใช้วิธีนำข้อมูลในปัจจุบันไปเปรียบเทียบกับข้อมูลในอดีต เพื่อค้นหาช่วงเวลาที่มีลักษณะเหตุการณ์คล้ายคลึงกันที่สุด (เช่น ยอดเดือนนี้

คล้ายกับเดือนเดียวกันของปีก่อน) แล้วนำตัวเลขของช่วงเวลานั้นมาใช้เป็นคำตอบในการพยากรณ์

1.4.6 การประเมินผล (Evaluation) เพื่อเปรียบเทียบความแม่นยำของแบบจำลองทั้งหมด จะใช้ตัวชี้วัดมาตรฐาน ดังนี้ RMSE (Root Mean Squared Error) วัดค่าความคลาดเคลื่อนโดยให้โทษหนักกับ error ที่สูงผิดปกติ MAE (Mean Absolute Error) หน่วยวัดค่าความผิดพลาดเฉลี่ยระหว่างค่าที่คาดการณ์และค่าจริง โดยคำนวณจากผลรวมของความแตกต่างสัมบูรณ์ระหว่างค่าที่คาดการณ์และค่าที่สังเกตได้ และผลการประเมินจะนำมาเปรียบเทียบเพื่อหาว่าโมเดลใดมีความแม่นยำสูงที่สุดและเหมาะสมที่สุดกับข้อมูลแรงงานไทย

1.4.7 การนำไปใช้งาน (Deployment) แสดงผลพยากรณ์แรงงานบนเว็บไซต์ ให้ผู้สนใจเข้าดูข้อมูลออนไลน์เพื่อใช้ประกอบการตัดสินใจได้ทันที

1.5 ขอบเขตการนำเสนอข้อมูล

1.5.1 ขอบเขตผู้ใช้งานทั่วไปบนเว็บไซต์

1.5.1.1 สามารถดูแนวโน้มอัตราการเดินทางของแรงงานไทยในแต่ละปีเพิ่มขึ้นจากการพยากรณ์ในลักษณะแผนภูมิสถิติการเดินทางไปทำงานต่างประเทศของแรงงานไทย

1.5.2.2 สามารถดูข้อมูลสารสนเทศในลักษณะตารางสรุปข้อมูลสถิติการเดินทางไปทำงานต่างประเทศของแรงงานไทย

1.5.2 ขอบเขตของผู้ดูแลระบบ

1.5.2.1 สามารถเผยแพร่ข้อมูลสารสนเทศได้

1.5.2.2 สามารถอัปเดตข้อมูลเกี่ยวกับอัตราการเดินทางไปทำงานต่างประเทศของแรงงานไทยเข้าไปในระบบได้

1.5.2.3 ผู้ดูแลสามารถเพิ่ม ลบ แก้ไข ข้อมูลการเดินทางไปทำงานต่างประเทศของแรงงานไทยได้

1.6 เครื่องมือที่ใช้

1.6.1 Hardware

1.6.1.1 CPU : AMD Ryzen 7 4800H

1.6.1.2 RAM : 16 GB

1.6.1.3 ระบบปฏิบัติการ : Windows 10 (64-bit)

1.6.2 Software

1.6.2.1 โปรแกรม AI Studio 2025

1.6.2.2 โปรแกรม Power BI

1.6.2.3 โปรแกรม Visual

1.6.2.4 ระบบปฏิบัติการ window 10

1.6.2.5 Microsoft Office 2016 Word Excel

1.7 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

1.7.1 สถานที่ใช้ในการดำเนินการศึกษา

มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา เชียงใหม่ 128 ถนน ห้วยแก้ว ตำบล
ช่างเผือก อำเภอ เมืองเชียงใหม่ เชียงใหม่ 50300

1.7.2 แหล่งรวบรวมข้อมูล

เว็บไซต์ กองบริหารแรงงานในไทยไปต่างประเทศและกระทรวงแรงงาน

1.8 ระยะเวลาในการดำเนินการ

ตารางที่ 1.1 แผนการดำเนินโครงการ

ปีการศึกษา แผนการดำเนินการ	2568							2569		
	มี.ย.	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.
1. ศึกษาและกำหนดความต้องการ	←		→							
2. ทำความเข้าใจข้อมูล		←	→							
3. การเตรียมข้อมูลและตรวจสอบ ความถูกต้อง		←	→							
4. การสร้างแบบจำลองข้อมูล			←	→						
5. การประเมินแบบจำลองข้อมูล					←	→				
6. การออกแบบเว็บไซต์				←	→					
7. การเขียนและทดสอบเว็บไซต์					←	→			←	→
8. ประเมินการใช้งานของเว็บไซต์									←	→
9. จัดทำเอกสารประกอบโครงการ	←									→

1.9 นิยามศัพท์เฉพาะ

1.9.1 แรงงานไทย คือ บุคคลสัญชาติไทยที่เดินทางไปทำงานในต่างประเทศตาม
ระยะเวลาที่กำหนด

1.9.2 แรงงานข้ามชาติ คือ บุคคลที่ย้ายถิ่นฐานจากประเทศต้นทางไปประกอบอาชีพ
ในอีกประเทศหนึ่ง

1.9.3 การพยากรณ์ คือ กระบวนการคาดคะเนค่าของข้อมูลในอนาคตจากข้อมูลในอดีตและปัจจุบัน โดยใช้วิธีทางสถิติหรือการทำเหมืองข้อมูล

1.9.4 การทำเหมืองข้อมูล (Data Mining) คือ กระบวนการสกัดหาความรู้และความสัมพันธ์จากข้อมูลปริมาณมาก เพื่อใช้ในการวิเคราะห์และตัดสินใจ

1.9.5 แบบจำลองการพยากรณ์ (Forecasting Model) คือ แบบจำลองทางคณิตศาสตร์หรือสถิติที่ใช้ความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม เพื่อคาดการณ์ค่าของตัวแปรตามในอนาคต

1.9.6 ค่ารากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error: RMSE) คือ ค่าที่ใช้วัดความแม่นยำของแบบจำลอง โดยยิ่งมีค่าน้อยแสดงว่าความคลาดเคลื่อนของการพยากรณ์โดยรวมต่ำการวิจัยนี้จะใช้ RMSE เพื่อเปรียบเทียบความแม่นยำของการพยากรณ์แรงงานไทยระหว่างโมเดลทั้ง 4 เทคนิค

1.9.7 ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) คือ ค่าที่ใช้วัดความคลาดเคลื่อนโดยเฉลี่ยระหว่างค่าที่พยากรณ์กับค่าจริง โดยไม่คำนึงถึงทิศทางของความคลาดเคลื่อนงานวิจัยนี้ใช้ MAE/MAPE เพื่อดูความคลาดเคลื่อนเฉลี่ยในรูปตัวเลขหรือเปอร์เซ็นต์ เพื่อให้เข้าใจง่ายว่าผลพยากรณ์คลาดเคลื่อนมากน้อยแค่ไหนเมื่อเทียบกับค่าจริง

1.9.8 CRISP-DM (Cross Industry Standard Process for Data Mining) คือ กระบวนการมาตรฐานในการทำเหมืองข้อมูล ซึ่งประกอบด้วย 6 ขั้นตอน ได้แก่ การทำความเข้าใจปัญหา การทำความเข้าใจข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง การประเมินผล และการนำไปใช้งาน โดยงานวิจัยนี้นำกระบวนการ CRISP-DM มาเป็นกรอบในการดำเนินงานทั้งหมด