บทที่ 4

ผลการดำเนินงาน

จากการวิเคราะห์ข้อมูล การวิเคราะห์การเปรียบเทียบประสิทธิภาพแนวโน้มการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564 –2566 เพื่อใช้สำหรับเผยแพร่ข้อมูลบนเว็บไซต์ มีวัตถุประสงค์เพื่อการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564–2566 สำหรับเผยแพร่ข้อมูล บนเว็บไซต์เพื่อให้ได้เว็บไซต์ที่รวมแหล่งความรู้ทั่วไปเกี่ยวกับข้อมูล และแสดงกระบวนการ วิเคราะห์ข้อมูลที่สามารถนำไปใช้งานได้อย่างมีประสิทธิภาพและประสิทธิผล โดยผู้จัดทำ โครงการได้ทำการวิเคราะห์ข้อมูล และนำเทคโนโลยีเข้าช่วยในการเสนอข้อมูล ทำให้มีความ สะดวกรวดเร็ว และสามารถใช้งานได้จริงจนสำเร็จลุล่วงตามเป้าหมาย

4.1 ผลการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลโครงงานเรื่องการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564–2566 ผู้วิเคราะห์ข้อมูลได้ดำเนินการศึกษาและวิเคราะห์ข้อมูลจาก www.doe.go.th ด้วยขั้นตอนกระบวนการวิเคราะห์ข้อมูล CRISP–DM อย่างละเอียดเพื่อให้ ข้อมูลมีความสมบูรณ์ และมีความถูกต้องมากที่สุด จากนั้นผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ ข้อมูลให้อยู่ในรูปแบบของ Data Mining โดยการจำแนกข้อมูลออกเป็นประเภทต่าง ๆ ด้วย เทคนิคการทำเหมืองข้อมูล ในรูปแบบของเทคนิคอนุกรมเวลา หรือที่ เรียกว่า Time series Analysis โดยจำแนกข้อมูลออกเป็น 3 หมวดหมู่คือ ผู้มีงานทำ ผู้ว่างงานและผู้ที่รอฤดูกาล

4.2 ผลการวิเคราะห์ออกแบบระบบ

ในการวิเคราะห์ข้อมูลโครงงานเรื่องการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564–2566 ผู้วิเคราะห์ข้อมูลได้ดำเนินการศึกษาและวิเคราะห์ข้อมูลจาก www.doe.go.th เพื่อให้ได้ข้อมูลสารสนเทศจากการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564–2566 ผู้วิเคราะห์ข้อมูลจัดทำการนำเสนอข้อมูลสารสนเทศแบบ Visualization ได้เผยแพร่ ข้อมูลสารสนเทศนี้บน Web Browser ให้กับผู้ใช้งานทั่วไป ซึ่งทำให้ผู้ใช้งานทั่วไปเข้าใจได้อย่าง สะดวกรวดเร็วยิ่งขึ้น รวมถึงมีความถูกต้อง ลดความซ้ำซ้อนจากข้อมูลที่มีจำนวนมากมหาศาล เพิ่มประสิทธิภาพให้กับการศึกษา

ค้นคว้าและทันต่อเวลา ผู้วิเคราะห์ข้อมูลจึงได้ทำการวิเคราะห์ข้อมูล และจัดทำเว็บไซต์ตามวัตถุประสงค์ที่ได้กำหนดไว้ และมีผลการดำเนินงานดังนี้

ภาพที่ 4.1 แสดงหน้าแรกของเว็บไซต์

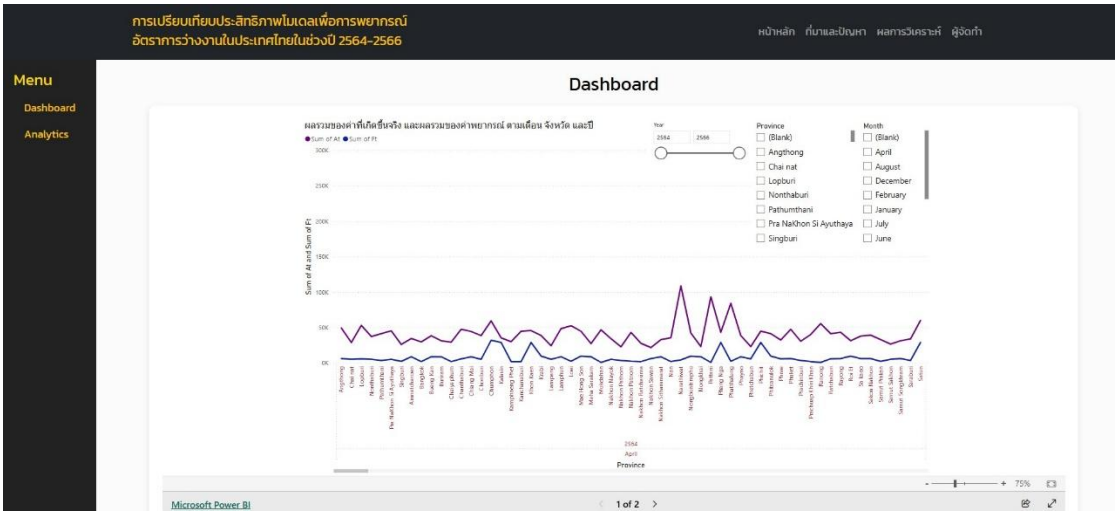
จากภาพที่ 4.1 แถบเมนูสามารถใช้งานได้ตามความต้องการของการใช้งาน ซึ่งเป็นหน้าสำหรับผู้ใช้งานทั่วไปสามารถเลือกดูและศึกษาค้นคว้าเข้าไปใช้งานในส่วนของเว็บไซต์ได้

ภาพที่ 4.2 แสดงหน้าข่าวสารตลาดแรงงาน

จากภาพที่ 4.2 เมื่อผู้ใช้คลิกที่ปุ่ม "READ MORE" บนหน้าแรก เว็บไซต์จะทำการเชื่อมต่อไปยังหน้าใหม่ที่มีข้อมูลเพิ่มเติมเกี่ยวกับข่าวสารตลาดแรงงาน

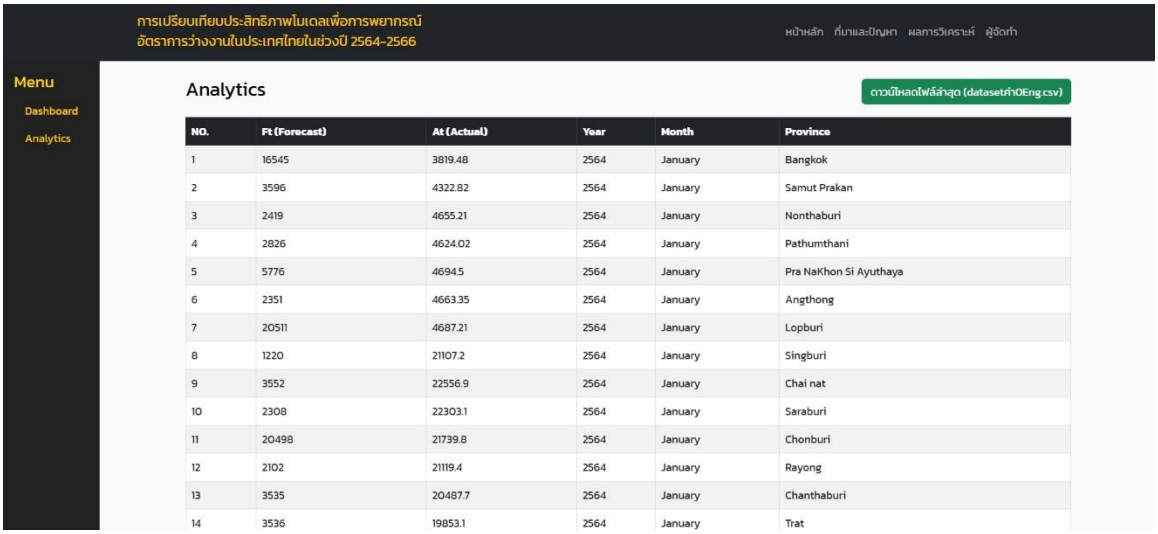
ภาพที่ 4.3 แสดงหน้าเว็บไซต์ที่มาของการวิเคราะห์ปัญหา

จากภาพที่ 4.3 หน้าแสดงหน้าเว็บไซต์ที่มาของการวิเคราะห์ปัญหา อธิบายรายละเอียดที่มาของการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564-2566



ภาพที่ 4.4 แสดงหน้าเว็บไซต์ผลการวิเคราะห์ข้อมูล Dashboard

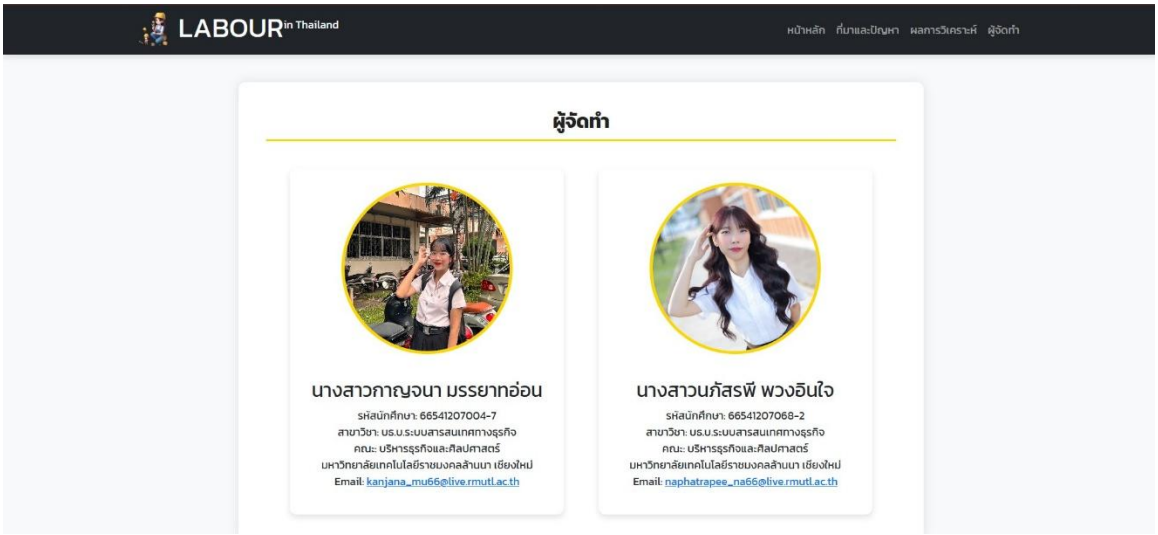
จากภาพที่ 4.4 หน้าแสดงหน้าเว็บไซต์ผลการวิเคราะห์ข้อมูล แสดงผลแบบ Dashborad ที่ผ่านการนำเข้าข้อมูลจากโปรแกรม Power Bi



NO.	Ft (Forecast)	At (Actual)	Year	Month	Province
1	16545	3819.48	2564	January	Bangkok
2	3596	4322.82	2564	January	Samut Prakan
3	2419	4655.21	2564	January	Nonthaburi
4	2826	4624.02	2564	January	Pathumthani
5	5776	4694.5	2564	January	Pra NaKhon Si Ayuthaya
6	2351	4663.35	2564	January	Angthong
7	20511	4687.21	2564	January	Lopburi
8	1220	21107.2	2564	January	Singburi
9	3552	22556.9	2564	January	Chai nat
10	2308	22303.1	2564	January	Saraburi
11	20498	21739.8	2564	January	Chonburi
12	2102	21119.4	2564	January	Rayong
13	3535	20487.7	2564	January	Chanthaburi
14	3536	19853.1	2564	January	Trat

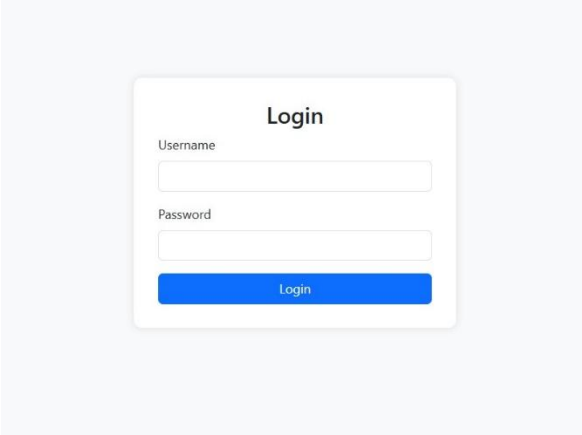
ภาพที่ 4.5 แสดงหน้าเว็บไซต์ผลการวิเคราะห์ข้อมูล Analytics

จากภาพที่ 4.4 หน้าแสดงหน้าเว็บไซต์ผลการวิเคราะห์ข้อมูล แสดงผล ตารางข้อมูล Analytics ตารางนี้นำเสนอข้อมูลเกี่ยวกับอัตราการจ้างงานในจังหวัดต่าง ๆ ของไทย โดยเปรียบเทียบค่าพยากรณ์กับค่าจริง ให้ผู้ใช้สามารถดาวน์โหลดข้อมูลที่แสดงอยู่ในตารางได้



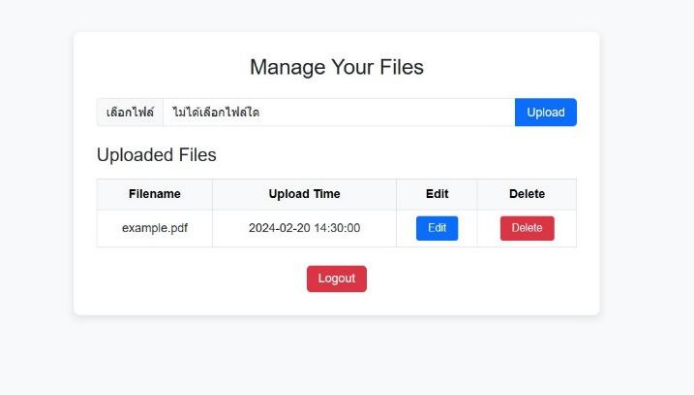
ภาพที่ 4.6 แสดงหน้าเว็บไซต์ผู้จัดทำ

จากภาพที่ 4.6 หน้าแสดงประวัติผู้จัดทำเว็บไซต์ และข้อมูลที่ต้องการติดต่อรวมไปถึงหน่วยงานที่เกี่ยวข้องกับโครงการนี้



ภาพที่ 4.7 แสดงหน้าหน้าเว็บไซต์การเข้าสู่ระบบ

จากภาพที่ 4.7 หน้าแสดงการเข้าสู่ระบบเพื่อไปจัดการกับข้อมูลภายในเว็บไซต์ โดย จำกัดการเข้าถึงข้อมูลเฉพาะแอดมินเท่านั้น



ภาพที่ 4.8 แสดงหน้าเว็บไซต์หน้าอัพโหลดไฟล์ข้อมูล

จากภาพที่ 4.8 แสดงหน้าเว็บไซต์หน้าอัพโหลดไฟล์ข้อมูลเข้าสู่เว็บไซต์ โดย จำกัดการเข้าถึงข้อมูลเฉพาะแอดมินเท่านั้น และข้อมูลที่ทำการอัพโหลดต้องเป็นข้อมูลที่ผ่านกระบวนการ CRISP-DM

4.3 อภิปรายผล

จากวัตถุประสงค์เพื่อการเปรียบเทียบประสิทธิภาพโมเดลเพื่อการพยากรณ์อัตราการว่างงานในประเทศไทยในช่วงปี 2564-2566 สำหรับเผยแพร่ข้อมูลความเป็นไปได้บนเว็บไซต์ ผู้วิเคราะห์ได้ ศึกษาปัญหา และเก็บรวบรวมข้อมูลซึ่งผู้วิเคราะห์ข้อมูลได้ทำการวิเคราะห์ในรูปแบบของ Data Mining โดยใช้กระบวนการวิเคราะห์ข้อมูลด้วย Cross Industry Standard Process for Data Mining หรือ CRISP-DM เพื่อใช้ในการวิเคราะห์ตามขอบเขตของโครงงานดังนี้

4.2.1 อภิปรายผลการดำเนินโครงงาน

4.2.1.1 อภิปรายผลจากการวิเคราะห์ข้อมูล

1) กระบวนการศึกษาทำความเข้าใจธุรกิจ (Business Understanding) ผลการศึกษาพบว่าปัญหาของข้อมูลคือ ข้อมูลสถิติความต้องการแรงงานรายจังหวัดปี 2564-2566 จำนวน 288 ไฟล์ ทำให้ไม่สามารถทำความเข้าใจกับข้อมูลที่มีความซับซ้อนได้อย่างรวดเร็วและนำเสนอหรือเผยแพร่ ข้อมูลให้กับบุคคลภายนอก หรือกลุ่มผู้ใช้ข้อมูลได้รับรู้ยังไม่มีประสิทธิภาพที่ดีนัก

2) การทำความเข้าใจข้อมูล (Data Understanding) ผู้วิเคราะห์ข้อมูลทำการรวบรวมข้อมูลเพื่อตรวจสอบรายละเอียดและปริมาณ จากเว็บไซต์ www.doe.go.th ซึ่งเป็นเว็บไซต์ของกองบริหารข้อมูลตลาดแรงงาน สามารถเชื่อถือได้ เป็นช่องทางให้ ผู้ใช้บริการทั้งภาคประชาชน ภาคธุรกิจเอกชน รวมถึงหน่วยงานของรัฐ สามารถค้นหา และ เข้าถึงข้อมูลที่มีคุณภาพของภาครัฐได้ง่าย

3) การเตรียมข้อมูล (Data Preparation) ผู้วิเคราะห์ข้อมูลนำ Dataset มาทำการคัดเลือก ข้อมูล และทำการ Data Cleaning แก้ไขข้อมูลที่ผิดพลาด ทำการจัดหมวดหมู่เพื่อความถูกต้อง ผู้วิเคราะห์ข้อมูลดำเนินการจัดกลุ่มแบ่งข้อมูล 3 แอตทริบิวต์ คือแอทริบิวต์ผู้มีงานทำ ผู้ว่างงาน และผู้ที่รอฤดูกาลเพื่อแบ่งข้อมูลอย่างชัดเจนเพื่อให้ได้ข้อมูลที่สมบูรณ์

4) การสร้างโมเดล (Modeling) ผู้วิเคราะห์ข้อมูลได้วิเคราะห์ข้อมูลด้วย การทำเหมืองข้อมูล ใช้เทคนิคอนุกรมเวลาหรือที่เรียกว่า Time series Analysis โดยใช้โมเดล ARIMA ,Holt's winters และ K-NN โดยใช้โปรแกรม Rapid miner studio ด้วยชุดข้อมูล 8 แอตทริบิวต์ ข้อมูลหลักที่ใช้ในการพยากรณ์คือ ผู้ว่างงาน ดังตารางที่ 4.1

ตารางที่ 4.1 แสดงตารางข้อมูลที่นำมาใช้ รายละเอียดของแอตทริบิวต์

ชื่อแอตทริบิวต์	รายละเอียด
ปี พ.ศ.	ปีพ.ศ.ที่นำข้อมูลมาวิเคราะห์ ตั้งแต่ พ.ศ. 2564-2566
เดือน	เดือนที่นำมาวิเคราะห์ 12 เดือน
จังหวัด	จังหวัดที่นำข้อมูลมาวิเคราะห์ 77 จังหวัด
ผู้มีงานทำ	บุคคลที่อยู่ในกำลังแรงงานและมีสถานะทำงานหรือประกอบอาชีพในช่วงเวลาที่กำหนด
ผู้ว่างงาน	บุคคลที่อยู่ในช่วงอายุและสถานะที่สามารถทำงานได้ แต่ในขณะสำรวจข้อมูลยังไม่มีงานทำ
ผู้ที่รอฤดูกาล	บุคคลที่อยู่ในช่วงอายุที่สามารถทำงานได้ และเคยมีงานทำมาก่อน แต่ในช่วงเวลาที่สำรวจข้อมูล ไม่มีงานทำเนื่องจากลักษณะงานของพวกเขาขึ้นอยู่กับฤดูกาลหรือช่วงเวลาที่กำหนด
รวม	ผลรวมของจำนวนผู้มีงานทำ ผู้ว่างงาน และผู้ที่รอฤดูกาล

ได้ผลลัพธ์ตามโมเดลดังต่อไปนี้

4.1 โมเดล ARIMA (AutoRegressive Integrated Moving Average) เป็นเทคนิคการพยากรณ์อนุกรมเวลาที่ได้รับความนิยมอย่างแพร่หลาย โดยผสมผสานระหว่างการถดถอยอัตโนมัติ (AR) การทำให้ข้อมูลอยู่ในระดับคงที่ (I) และค่าเฉลี่ยเคลื่อนที่ (MA) เพื่อสร้างโมเดลที่สามารถพยากรณ์ข้อมูลในอนาคตได้อย่างมีประสิทธิภาพ (โรม, 2561) ผู้วิเคราะห์ได้นำเอาโมเดล ARIMA มาใช้ในการวิเคราะห์ข้อมูลอัตราการว่างงานในประเทศไทยปี2564-2566 โดยใช้โปรแกรม Rapid miner studio ในการวิเคราะห์ข้อมูล แสดงผลดังภาพที่ 4.6

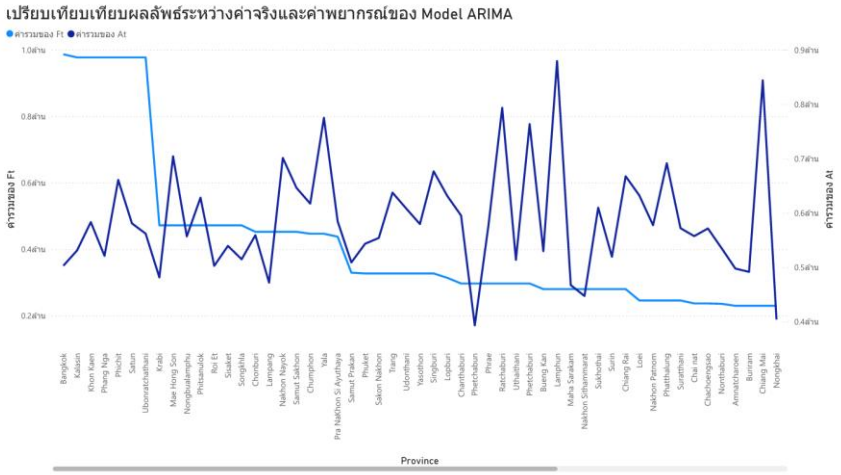
ArimaModel

```
Forecast Model trained on the following time series:
Name of time series: Unemployed Number of values: 100

Resulting Forecast Model:
Arima Model (p: 2,d: 0, q: 1)
AR Coefficients: [0.045428433651755834, 0.06130871707854149]
MA Coefficients: [0.2913706691306534]
constant: 5084.820205059702, length of residuals: 1
```

ภาพที่ 4.9 แสดงผลลัพธ์ ของการพยากรณ์โดยใช้โมเดล ARIMA

จากภาพที่ 4.9 ผลลัพธ์ของโมเดล ARIMA โดยใช้แอตทริบิวต์ ผู้ว่างงาน(Unemployed) ในการพยากรณ์ ค่า AR Coefficients คือการถดถอยอัตโนมัติเท่ากับ 0.0454 และ0.0613 หมายความว่าความสัมพันธ์ของข้อมูลในอดีตมีอิทธิพลต่อค่าพยากรณ์อย่างมีนัยสำคัญ ค่า MA Coefficient คือค่าเฉลี่ยเคลื่อนที่เท่ากับ 0.2913 หมายความว่าข้อผิดพลาดในอดีตส่งผลกระทบอย่างมากต่อค่าพยากรณ์ในปัจจุบัน ค่า constant คือค่าคงที่เท่ากับ 5084.82 หมายความว่า ข้อมูลส่วนใหญ่อยู่ใกล้เคียงกับค่านี้ ซึ่งบ่งบอกถึงความเหมาะสมของโมเดลในการพยากรณ์ โมเดล ARIMA สามารถอธิบายแนวโน้มของจำนวนผู้ว่างงานได้อย่างมีนัยสำคัญ ผ่านองค์ประกอบของความสัมพันธ์เชิงเวลาและข้อผิดพลาดในอดีต

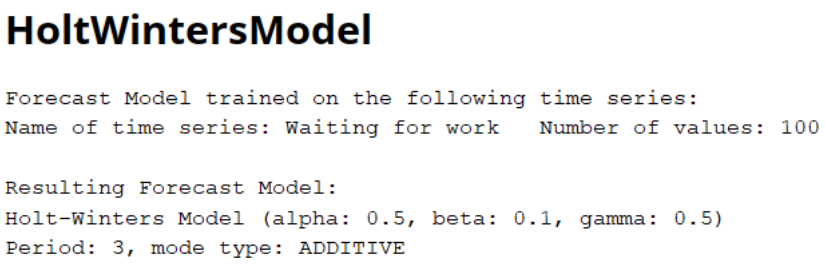


ภาพที่ 4.10 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงแบบรายจังหวัด

จากภาพที่ 4.10 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงของจำนวนผู้ว่างงานในแต่ละจังหวัดของประเทศไทยทั้ง 77 จังหวัด โดยใช้โปรแกรม Power BI สำหรับการแสดงผล และใช้ RapidMiner ในการพยากรณ์ข้อมูล กราฟเส้นนี้ประกอบด้วย แกน X ซึ่งแสดงรายชื่อจังหวัดใน

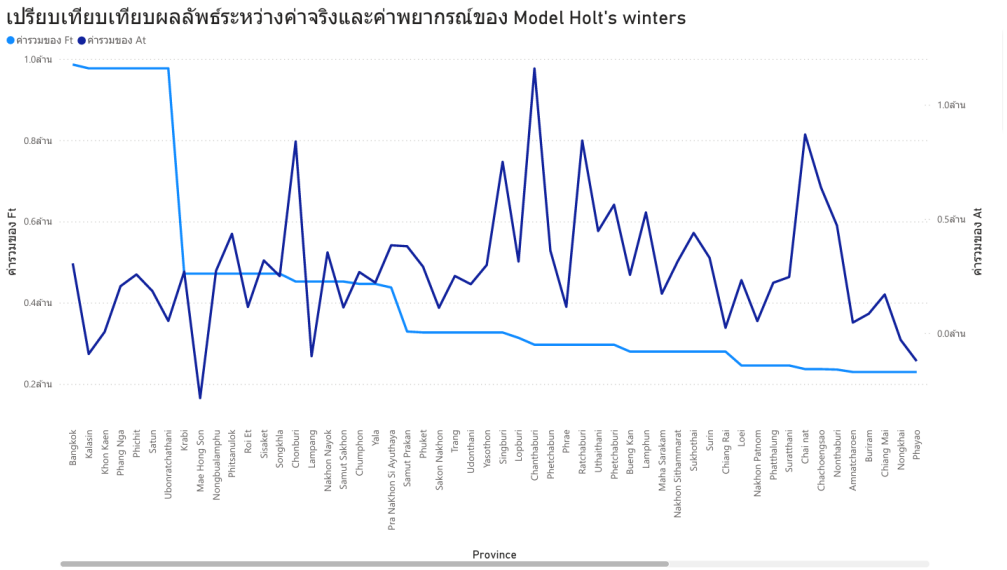
ประเทศไทย แกน Y หลัก แสดงค่าพยากรณ์จำนวนผู้ว่างงานที่ได้จาก RapidMiner และ แกน Y รอง แสดงค่าจริงของจำนวนผู้ว่างงาน จากกราฟพบว่า เส้นค่าพยากรณ์ (สีน้ำเงินอ่อน) มีแนวโน้มแตกต่างจาก เส้นค่าจริง (สีน้ำเงินเข้ม) ในบางจังหวัด ซึ่งสะท้อนให้เห็นถึงข้อผิดพลาดของโมเดลการพยากรณ์ อย่างไรก็ตาม ในบางพื้นที่ โมเดลสามารถคาดการณ์ได้ใกล้เคียงกับค่าจริง ขณะที่บางจังหวัดมีข้อผิดพลาดค่อนข้างสูง ข้อมูลนี้บ่งชี้ว่าโมเดลอาจต้องได้รับการปรับปรุงเพิ่มเติม เช่น การปรับพารามิเตอร์ของโมเดล หรือการเพิ่มตัวแปรอื่น ๆ เพื่อเพิ่มความแม่นยำในการพยากรณ์

4.2 Holt–Winters เป็นเทคนิคการพยากรณ์ที่ใช้สำหรับข้อมูลอนุกรมเวลาที่มีแนวโน้ม (trend) และฤดูกาล (seasonality) โดยอาศัยหลักการของ Exponential Smoothing ซึ่งช่วยให้ค่าพยากรณ์มีความแม่นยำขึ้นโดยคำนึงถึงข้อมูลในอดีต (รังสิมา, 2566) ผู้วิเคราะห์ได้นำเอาโมเดล Holt–Winters มาใช้ในการวิเคราะห์ข้อมูลอัตราการว่างงานในประเทศไทยปี2564–2566 โดยใช้โปรแกรม Rapid miner studio ในการวิเคราะห์ข้อมูล แสดงผลดังภาพที่ 4.8



ภาพที่ 4.11 แสดงผลลัพธ์ ของการพยากรณ์โดยใช้โมเดล Holt's winter

จากภาพที่ 4.11 ผลลัพธ์ของโมเดล Holt's winter โดยใช้แอตทริบิวต์ ผู้รอฤดูกาล(Waiting for work) ในการพยากรณ์ ค่าAlpha คือค่าที่ควบคุมน้ำหนักของข้อมูลในอดีตเท่ากับ0.5 หมายความว่าโมเดลให้ความสำคัญกับข้อมูลปัจจุบันมากกว่าอดีต ค่าbeta คือค่าที่ควบคุมแนวโน้มเท่ากับ 0.1 หมายความว่าโมเดลปรับแนวโน้มช้าและ ค่าgamma คือค่าที่ควบคุมฤดูกาลเท่ากับ 0.5 หมายความว่าโมเดลปรับฤดูกาลเร็ว



ภาพที่ 4.12 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงแบบรายจังหวัด

จากภาพที่ 4.12 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงของจำนวนผู้ว่างงานในแต่ละจังหวัดของประเทศไทยทั้ง 77 จังหวัด โดยใช้โปรแกรม Power BI สำหรับการแสดงผล และใช้ Rapid Miner ในการพยากรณ์ข้อมูล กราฟเส้นนี้ประกอบด้วยแกน X ซึ่งแสดงรายชื่อจังหวัดในประเทศไทย แกน Y หลัก แสดงค่าพยากรณ์จำนวนผู้ว่างงานที่ได้จาก Rapid Miner และแกน Y รอง แสดงค่าจริงของจำนวนผู้ว่างงาน จากกราฟพบว่า ค่าพยากรณ์ (เส้นสีน้ำเงินอ่อน) มีแนวโน้มแตกต่างจากค่าจริง (เส้นสีน้ำเงินเข้ม) ในบางจังหวัด สะท้อนให้เห็นถึงความคลาดเคลื่อนของโมเดล แม้ว่าจะมีบางพื้นที่ที่โมเดลสามารถพยากรณ์ได้ใกล้เคียงกับค่าจริง แต่ในบางจังหวัดค่าพยากรณ์มีความราบเรียบกว่ามาก ขณะที่ค่าจริงมีความผันผวนสูง ซึ่งอาจบ่งชี้ว่าโมเดล Holt's Winters ไม่สามารถจับลักษณะของข้อมูลในบางพื้นที่ได้อย่างแม่นยำ

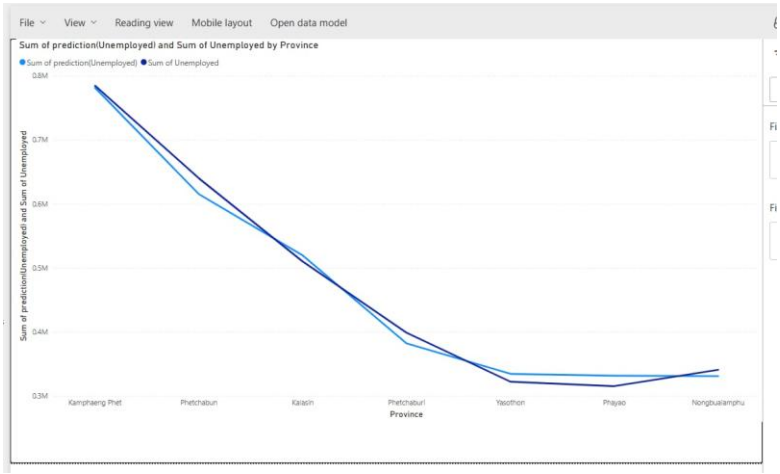
4.3 K-Nearest Neighbors (K-NN) เป็นหนึ่งในอัลกอริธึมของ Machine Learning ที่ใช้หลักการของ ความคล้ายคลึงกัน (Similarity) ในการจำแนกประเภทหรือทำนายค่า จุดเด่นของโมเดลนี้คือ ไม่ต้องมีการฝึกโมเดลล่วงหน้า (Lazy Learning) แต่จะใช้ข้อมูลในอดีตมาคำนวณหาค่าที่ใกล้เคียงที่สุดเพื่อตัดสินผลลัพธ์ (พรเทพ, 2562) ผู้วิเคราะห์ได้นำเอาโมเดล Holt-Winters มาใช้ในการวิเคราะห์ข้อมูลอัตราการว่างงานในประเทศไทยปี2564-2566 โดยใช้โปรแกรม Rapid miner studio ในการวิเคราะห์ข้อมูล แสดงผลดังภาพที่ 4.13

KNNRegression

```
Weighted 7-Nearest Neighbour model for regression.
The model contains 2772 examples with 7 dimensions.
```

ภาพที่ 4.13 แสดงผลลัพธ์ ของการพยากรณ์โดยใช้โมเดล K-NN

จากภาพที่ 4.13 ผลลัพธ์ของโมเดล K-NNโดยใช้แอตทริบิวต์ ผู้ว่างงาน (Unemployee) ในการพยากรณ์ เมื่อทำการ Cross validation แล้วได้ลัพธ์เท่ากับ 7 dimensions ซึ่งเป็นค่าใกล้เคียง 7 จังหวัดที่มีค่าพยากรณ์และค่าจริงที่มีความใกล้เคียงกัน ได้แก่ จังหวัด กำแพงเพชร เพชรบูรณ์ กาฬสิน เพชรบุรี ยโสธร พะเยา และหนองบัวลำภู



ภาพที่ 4.14 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงแบบรายจังหวัด

จากภาพที่ 4.14 แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงของจำนวนผู้ว่างงานในแต่ละจังหวัดของประเทศไทยทั้ง 77 จังหวัด โดยใช้โปรแกรม Power BI สำหรับการแสดงผล และใช้ RapidMiner ในการพยากรณ์ข้อมูล กราฟเส้นนี้ประกอบด้วย แกน X ซึ่งแสดงรายชื่อจังหวัดในประเทศไทย แกน Y หลัก แสดงค่าพยากรณ์จำนวนผู้ว่างงานที่ได้จาก Rapid Miner และ แกน Y

รอง แสดงค่าจริงของจำนวนผู้ว่างงาน จากกราฟพบว่า กราฟนี้แสดงการเปรียบเทียบค่าพยากรณ์และค่าจริงของจำนวน ผู้ว่างงาน (Unemployed) ใน 77 จังหวัดของประเทศไทย โดยใช้ โมเดล K-NN ในการพยากรณ์ พบว่าโมเดลให้ผลลัพธ์ที่ดีที่สุดที่ 7 dimensions ซึ่งสอดคล้องกับ 7 จังหวัด ที่มีค่าพยากรณ์ใกล้เคียงค่าจริงมากที่สุด ได้แก่ กำแพงเพชร เพชรบูรณ์ กาฬสินธุ์ เพชรบุรี ยโสธร พะเยา และหนองบัวลำภู โดยเส้นสองเส้นในกราฟ (สีน้ำเงินอ่อนและสีน้ำเงินเข้ม) แสดงให้เห็นว่าค่าพยากรณ์และค่าจริงในจังหวัดเหล่านี้มีแนวโน้มใกล้เคียงกันมาก

5) การวัดประสิทธิภาพของโมเดล (Evaluation) การทดสอบ ประสิทธิภาพของแบบจำลองได้ดำเนินการวัดค่า RMSE ค่านี้จะให้ความสำคัญกับค่าผิดพลาดที่มาก โดยค่า RMSE ที่ต่ำแสดงว่าโมเดลมีความแม่นยำสูง ,MAE ค่านี้จะให้ความสำคัญกับขนาดของความคลาดเคลื่อนโดยรวม ค่า MAE ที่ต่ำแสดงว่าโมเดลมีความแม่นยำสูง MAPE วัดความแม่นยำของโมเดลพยากรณ์, Relative Error วัดค่าความคลาดเคลื่อนมีขนาดใหญ่เมื่อเทียบกับค่าจริงมากน้อยเพียงใด ค่าที่ต่ำแสดงว่าโมเดลมีความแม่นยำสูง และSquared Error ค่าความคลาดเคลื่อนกำลังสองระหว่างค่าที่ทำนายได้กับค่าจริง ดังตารางที่ 4.2

ตารางที่ 4.2 ตารางเปรียบเทียบประสิทธิภาพระหว่างโมเดล

Model	RMSE	MAE	MAPE	Relative Error	Squared Error
ARIMA	4231.482	3250.121	17%	462.16%	554147953.621
Holt's winters	5101.080	4073.976	6%	575.72%	86922612.346
K-NN	71327.397	10770.948	9%	109.39%	11599905173.965

จากตารางที่ 4.2 แสดงการเปรียบเทียบประสิทธิภาพของโมเดลพยากรณ์ 3 รูปแบบ ได้แก่ ARIMA, Holt's Winters และ K-NN โดยใช้ตัวชี้วัดด้านความแม่นยำและข้อผิดพลาด เช่น RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), Relative Error และ Squared Error เพื่อประเมินผลการพยากรณ์ของแต่ละโมเดล โมเดลที่มีค่าความคลาดเคลื่อนต่ำกว่าย่อมแสดงถึงความแม่นยำที่ดีกว่า

5.1 ARIMA RMSE (4231.482) และ MAE (3250.121) มีค่าความคลาดเคลื่อนเฉลี่ยต่ำที่สุดในทุกตัวชี้วัด ซึ่งแสดงว่าโมเดลนี้เหมาะสมสำหรับการคาดการณ์ข้อมูลชุดนี้ MAPE (17%) มีความแม่นยำในระดับดีพอสมควรโดยความคลาดเคลื่อนเฉลี่ยคิดเป็น 17% ของค่าจริง

Relative Error (462.16%) ค่าความคลาดเคลื่อนสัมพัทธ์อยู่ในระดับสูง Squared Error (554147953.621) มีค่าความคลาดเคลื่อนยกกำลังสองต่ำที่สุด แสดงว่าโมเดลนี้สามารถลดค่าคลาดเคลื่อนขนาดใหญ่ได้ดีกว่าโมเดลอื่น

5.2 Holt's winters RMSE (5101.080) และ MAE (4073.976) มีค่าคลาดเคลื่อนมากกว่า ARIMA แต่ยังน้อยกว่า K-NN MAPE (6%) มีค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยเป็นเปอร์เซ็นต์ที่ต่ำที่สุดในตาราง ซึ่งบ่งชี้ว่าโมเดลนี้ทำงานได้ดีมากในแง่ของเปอร์เซ็นต์ความแม่นยำ Relative Error (575.72%) ค่าคลาดเคลื่อนสัมพัทธ์สูงกว่าของ ARIMA Squared Error (86922612.346) มีความคลาดเคลื่อนขนาดใหญ่มากกว่า ARIMA

5.3 K-NN RMSE (71327.397) และMAE (10770.948) มีค่าความคลาดเคลื่อนสูงมากเมื่อเทียบกับ ARIMA และ Holt's Winters MAPE (9%)อยู่ในระดับกลาง ค่าความคลาดเคลื่อนสัมบูรณ์คิดเป็น 9% ของค่าจริง Relative Error (109.39%) มีค่าความคลาดเคลื่อนสัมพัทธ์ต่ำที่สุดในตาราง ซึ่งแสดงว่าโมเดลนี้ทำงานได้ดีเมื่อเทียบในเชิงสัดส่วน Squared Error (11599905173.965) มีค่าความคลาดเคลื่อนยกกำลังสองสูงที่สุดในตาราง

6) การนำโมเดลไปใช้งานจริง (Deployment) ผู้วิเคราะห์ข้อมูลได้ทำการศึกษาความรู้พื้นฐาน ทางด้านภาษาที่ใช้ในการพัฒนาเว็บไซต์คือ ภาษาเอชทีเอ็มแอล (HTML) ภาษาพีเอชพี (PHP) และภาษาซีเอสเอส (CSS) เมื่อศึกษาความรู้พื้นฐานในภาษาต่าง ๆ และการวิเคราะห์ข้อมูลด้วยเทคนิคที่กล่าวมาข้างต้น แล้วจึงทำการวิเคราะห์ข้อมูลและ ออกแบบเว็บไซต์ โดยเว็บไซต์มีส่วนประกอบต่าง ๆ ดังนี้ ส่วนของหน้าแรกที่รวบรวมเมนู ต่าง ๆไว้ ส่วนของที่มาและปัญหาของการวิเคราะห์ข้อมูล ส่วนของการแสดงผล Dashboard ผ่าน Power Bi ส่วนของการวิเคราะห์ข้อมูลด้วยโมเดลARIMA และ ส่วนของผู้จัดทำ ผู้ใช้งานทั่วไปบนเว็บบราวเซอร์

1) สามารถดูข้อมูลสารสนเทศในลักษณะตารางสรุปข้อมูลสถิติอัตราการว่างงานในประเทศไทยในช่วงปี 2564 -2566

2) สามารถแสดงผล แผนภาพสรุปข้อมูล Data visualization

3) สามารถกรองข้อมูลสารสนเทศผ่านระบบการค้นหาแสดงผลแผนภาพสรุปข้อมูล เช่น ปี 2564-2566 และภูมิภาค

4) สามารถเลือกการแสดงอัตราการว่างงาน แต่ละจังหวัดได้

4.4 บทสรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของโมเดลพยากรณ์อัตราการว่างงานในประเทศไทยระหว่างปี 2564–2566 โดยใช้กระบวนการวิเคราะห์ข้อมูล CRISP–DM และเทคนิค Data Mining เพื่อให้สามารถทำนายแนวโน้มของอัตราการว่าง งานได้อย่างแม่นยำ นอกจากนี้ ผลการวิเคราะห์ยังถูกนำเสนอผ่าน เว็บไซต์ ในรูป Dashboard เพื่อให้ผู้ใช้สามารถเข้าถึงข้อมูลได้สะดวก

การศึกษานี้ได้เปรียบเทียบ ประสิทธิภาพของโมเดลหลัก 3 โมเดล ได้แก่ ARIMA, Holt–Winters และ K–NN ผลการวิเคราะห์พบว่า โมเดล ARIMA มีค่าคลาดเคลื่อนต่ำที่สุดทั้ง ในแง่ RMSE และ MAE ทำให้สามารถพยากรณ์แนวโน้มของข้อมูลได้แม่นยำที่สุด แม้ว่า จะมีค่า Relative Error สูง แต่ค่าคลาดเคลื่อนกำลังสองที่ต่ำบ่งชี่ว่าโมเดลนี้สามารถ จัดการกับข้อมูลที่มีแนวโน้มระยะยาวได้ดี ในขณะที่ Holt–Winters เป็นโมเดลที่มีความ แม่นยำสูงโดยเฉพาะในข้อมูลที่มีฤดูกาล โดยมีค่า MAPE ต่ำสุดที่ 6% ซึ่งแสดงให้เห็นว่า โมเดลนี้สามารถจับแนวโน้มและฤดูกาลได้อย่างมีประสิทธิภาพ แม้ว่าค่าคลาดเคลื่อน RMSE และ MAE จะสูงกว่า ARIMA แต่ก็ยังต่ำกว่า K–NN ทำให้เป็นทางเลือกที่ดีสำหรับ ข้อมูลที่มีลักษณะเป็นวัฏจักรตามฤดูกาล ในทางกลับกัน K–NN มีค่าคลาดเคลื่อนสัมพัทธ์ สูงสุดในทุกตัวชี้วัด ทำให้ไม่เหมาะสมกับการพยากรณ์ข้อมูลอนุกรมเวลา

เพื่อให้ข้อมูลที่ได้จากการวิเคราะห์สามารถนำไปใช้งานได้ง่าย งานวิจัยนี้ได้พัฒนาเว็บไซต์ เพื่อแสดงผลข้อมูล โดยใช้ HTML, CSS, PHP และ Power BI เว็บไซต์ช่วยให้ผู้ใช้ สามารถดูสถิติอัตราการว่างงานในช่วงปี 2564–2566 พร้อมทั้งสามารถวิเคราะห์แนว โน้ม เปรียบเทียบข้อมูลในแต่ละจังหวัด และเข้าถึงข้อมูลผ่าน Dashboard ที่ออกแบบมา ให้ใช้งานง่าย จากผลการศึกษาสามารถสรุปได้ว่า ARIMA เป็นโมเดลที่ดีที่สุดสำหรับการพยากรณ์แนว โน้มอัตราการว่างงาน ในขณะที่ Holt–Winters เหมาะสำหรับข้อมูลที่มีลักษณะเป็นฤดูกาล ส่วน K–NN อาจไม่เหมาะสมสำหรับงานพยากรณ์อนุกรมเวลา เนื่องจากมีค่าคลาดเคลื่อน สูง โครงการนี้จึงประสบความสำเร็จในการ นำเทคโนโลยีมาช่วยวิเคราะห์และเผยแพร่ ข้อมูลอัตราการว่างงานได้อย่างแม่นยำ ซึ่งสามารถนำไปใช้ประโยชน์ในการวางแผน นโยบายแรงงานและการบริหารทรัพยากรบุคคลได้อย่างมีประสิทธิภาพ