DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

การพยากรณ์ปริมาณการเข้าชมเว็บไซต์สำหรับธุรกิจอีคอมเมิร์ซ โดยใช้เทคนิคการเรียนรู้ของเครื่องและสร้างระบบแจ้งเตือนโดยอัตโนมัติ

สิทธิ โง้ววัฒนา

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาตรมหาบัณฑิต
สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่
วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์
มหาวิทยาลัยธุรกิจบัณฑิตย์
ปีการศึกษา 2565

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

FORECASTING WEB TRAFFIC FOR E-COMMERCE WEBSITE USING MACHINE LEARNING AND AUTOMATED PUSH NOTIFICATIONS

SITTHI NGOWWATTANA

A Thematic Paper Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Engineering
Department of Big Data Engineering,
College of Innovative Technology and Engineering
Dhurakij Pundit University
Academic Year 2022

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ใบรับรองสารนิพนธ์

วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์

วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่

หัวข้อสารนิพนธ์ การพยากรณ์ปริมาณการเข้าชมเว็บไซต์สำหรับธุรกิจอีคอมเมิร์ซโดยใช้เทคนิคการเรียนรู้ของเครื่องและสร้างระบบแจ้งเตือนโดยอัตโนมัติ

เสนอโดย สิทธิ โง้ววัฒนา

สาขาวิชา วิศวกรรมข้อมูลขนาดใหญ่

อาจารย์ที่ปรึกษาสารนิพนธ์ ดร.เอกสิทธิ์ พัชรวงศ์ศักดา

ได้พิจารณาเห็นชอบโดยคณะกรรมการสอบสารนิพนธ์แล้ว

วฤษายย์ ประธานกรรมการ

(รองศาสตราจารย์ ดร.วฤษาย์ ร่มสายหยุด)

เอกสิทธิ์ พัชรวงศ์ศักดา กรรมการที่ปรึกษาสารนิพนธ์

(ดร.เอกสิทธิ์ พัชรวงศ์ศักดา)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ดวงใจ จิตคงชื่น)

วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ รับรองแล้ว

คณบดีวิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์

(ดร.ชัยพร เขมะภาตะพันธ์)

วันที่ 31 เดือน กรกฎาคม พ.ศ. 2566

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

หัวข้อสารนิพนธ์	การพยากรณ์ปริมาณการเข้าชมเว็บไซต์สำหรับธุรกิจอีคอมเมิร์ซ โดยใช้เทคนิคการเรียนรู้ของเครื่องและสร้างระบบแจ้งเตือนโดยอัตโนมัติ
ชื่อผู้เขียน	สิทธิ โง้ววัฒนา
อาจารย์ที่ปรึกษา	ดร. เอกสิทธิ์ พัชรวงศ์ศักดา
หลักสูตร	วิศวกรรมข้อมูลขนาดใหญ่
ปีการศึกษา	2565

บทคัดย่อ

งานวิจัยนี้นำเสนอการพยากรณ์ปริมาณการเข้าชมเว็บไซต์สำหรับธุรกิจ e-commerce โดยใช้เทคนิคการเรียนรู้ของเครื่องในการพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) ซึ่งจะช่วยคาดการณ์แนวโน้มปริมาณการเข้าชมเว็บไซต์ในอนาคต เมื่อปริมาณการเข้าชมเว็บไซต์เกินกว่าเกณฑ์ที่กำหนดไว้ ระบบจะแจ้งเตือนผ่านแอปพลิเคชั่นให้ผู้ดูแลระบบรับทราบเพื่อทำการปรับปรุงเซิร์ฟเวอร์ทันที นี้จะช่วยให้สามารถจัดการเตรียมทรัพยากรและเซิร์ฟเวอร์ได้อย่างมีประสิทธิภาพ นอกจากนี้การพยากรณ์ปริมาณการเข้าชมเว็บไซต์อย่างแม่นยำ ยังสามารถใช้ในการวางแผนการตลาดและโปรแกรมส่งเสริมการขาย ช่วยให้ธุรกิจสามารถจัดสรรทรัพยากรได้อย่างเหมาะสมและเพิ่มประสบการณ์ใช้งานที่ดีสำหรับผู้ใช้งาน

คำสำคัญ: ปริมาณการเข้าชมเว็บไซต์, อนุกรมเวลา, อีคอมเมิร์ซ, การเรียนรู้ของเครื่อง

อาจารย์ที่ปรึกษา

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

Thematic Paper Title	FORECASTING WEB TRAFFIC FOR E-COMMERCE WEBSITE USING MACHINE LEARNING AND AUTOMATED PUSH NOTIFICATIONS
Author	SITTHI NGOWWATTANA
Thematic Paper Advisor	Dr. Eakasit Pacharawongsakda
Program	Big Data Engineering
Academic Year	2022

ABSTRACT

This research presents a framework for time series forecasting of web traffic on an e-commerce website, with a focus on optimizing server resource allocation and enhancing performance. Leveraging time series forecasting techniques, the study predicts future web traffic patterns, enabling businesses to efficiently manage server infrastructure and allocate resources effectively. The framework incorporates automated push notifications, alerting the system administrator when the forecasted web traffic exceeds a predefined threshold. This proactive approach allows for timely server upscaling to ensure optimal website performance. The accurate forecasting of web traffic also facilitates strategic planning for marketing campaigns, enabling businesses to allocate resources efficiently. By leveraging time series forecasting, businesses can improve operational efficiency, maintain competitiveness, and deliver a seamless user experience on their e-commerce websites.

Keywords: Web Traffic, Time Series, E-Commerce, Machine Learning

Advisor

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

กิตติกรรมประกาศ

สารนิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีโดยการให้ความช่วยเหลือของ ดร.เอกสิทธิ์ พัชรวงศ์ศักดา ซึ่งเป็นอาจารย์ที่ปรึกษาสารนิพนธ์ ที่ได้กรุณาให้คำแนะนำ ตรวจสอบ และแก้ไขข้อบกพร่องต่าง ๆ มาโดยตลอด เพื่อให้สารนิพนธ์ฉบับนี้สมบูรณ์ ผู้เขียนจึงขอกราบขอบพระคุณไว้ ณ โอกาสนี้

ขอขอบคุณ รศ.ดร.วฤษาย์ ร่มสายหยุด ที่กรุณาให้เกียรติเป็นประธานโดยมี ผศ.ดร.ดวงใจ จิตคงชื่น เป็นกรรมการในการสอบสารนิพนธ์ ซึ่งได้กรุณาตรวจ แก้ไขสารนิพนธ์ฉบับ นี้ให้ถูกต้องสมบูรณ์ยิ่งขึ้น รวมทั้งได้ให้คำแนะนำแนวทางที่เป็นประโยชน์ต่องานวิจัย และขอขอบคุณ นางสาวกุลธิดา รอดบุญ ที่ให้ความสะดวกด้านอำนวยการ และประสานงาน ในการทำสารนิพนธ์ให้กับผู้เขียนมาโดยตลอด

หากมีสิ่งใดที่ผู้วิจัยได้ทำผิดพลาดหรือบกพร่องประการใด ผู้วิจัยต้องกราบขออภัยเป็นอย่างสูงมา ณ โอกาสนี้ หวังเป็นอย่างยิ่งว่าสารนิพนธ์ฉบับนี้ จะเป็นพื้นฐานในการต่อยอดองค์ความรู้ของผู้ที่สนใจ ศึกษางานในด้านนี้ต่อไป

สุดท้ายนี้ขอขอบคุณ บิดา มารดา ครอบครัวและเพื่อนๆ ที่คอยช่วยส่งเสริม สนับสนุนและให้กำลังใจ ทำให้การศึกษาวิจัยในครั้งนี้สำเร็จลุล่วงไปด้วยดี

สิทธิ โง้ววัฒนา

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญภาพ..... ญ

บทที่

1. บทนำ..... 1
   - 1.1 ที่มาและความสำคัญของปัญหา..... 1
   - 1.2 วัตถุประสงค์ของการศึกษาหรือวิจัย..... 1
   - 1.3 ขอบเขตของงานวิจัย..... 2
   - 1.4 ประโยชน์ที่คาดว่าจะได้รับ..... 2
2. แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง..... 3
   - 2.1 ความหมายและประเภทของ Web Traffic..... 3
   - 2.2 ความหมายและประเภทของการพยากรณ์..... 4
   - 2.3 พยากรณ์แบบอนุกรมเวลา (Time Series Forecasting)..... 5
   - 2.4 Autoregressive Integrated Moving Average (ARIMA)..... 7
   - 2.5 Prophet..... 8
   - 2.6 งานวิจัยที่เกี่ยวข้อง..... 10
3. ระเบียบวิธีวิจัย..... 14
   - 3.1 การรวบรวมข้อมูล..... 14
   - 3.2 การจัดเตรียมข้อมูล..... 15
   - 3.3 การสร้างตัวแปร (Feature Engineering)..... 16
   - 3.4 การสร้างโมเดล..... 17
   - 3.5 เครื่องมือที่ใช้ในงานวิจัย..... 24

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

สารบัญ (ต่อ)

บทที่ **หน้า**

4. ผลการวิจัย..... 26
4.1 ผลการวัดประสิทธิภาพความถูกต้องของโมเดล..... 26
4.2 ผลการดำเนินการในส่วนของการแจ้งเตือนผ่านแอปพลิเคชั่น..... 28
5. สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ..... 30
5.1 สรุปผลการศึกษา..... 30
5.2 ข้อเสนอแนะ..... 30
บรรณานุกรม..... 32
ภาคผนวก..... 34
ก การดึงข้อมูลจาก Google Analytics เพื่อนำมาใช้งาน..... 35
ประวัติผู้เขียน..... 38

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

สารบัญตาราง

ตารางที่ **หน้า**

2.1 รูปแบบแต่ละช่วงเวลาของฤดูกาล..... 6
3.1 ตารางเทศกาลของ e-commerce holidays 17
4.1 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล ARIMA (Baseline model)..... 26
4.2 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล Prophet แบบ Univariate (Baseline Model) 27
4.3 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล Prophet แบบ Multivariate 28

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

สารบัญภาพ

ภาพที่ **หน้า**

2.1 องค์ประกอบของอนุกรมเวลา 5
2.2 ตัวอย่างสมการของโมเดล ARIMA 8
2.3 ขั้นตอนการทำงานของงานวิจัย (1) 10
2.4 ผลการพยากรณ์ของ ARIMA 11
2.5 ผลการพยากรณ์ของ LSTM RNN 11
2.6 ผลการพยากรณ์ของการรวม ARIMA และ LSTM RNN เข้ากันด้วย DWT 12
2.7 ขั้นตอนการทำงานของงานวิจัย (2) 12
2.8 ผลการทดลองของงานวิจัย (2) 13
2.9 ผลการทดลองของงานวิจัย (3) 13
3.1 Dashboard ของ Google Analytics 14
3.2 เว็บไซต์ Google Merchandise Store 15
3.3 ตัวอย่างชุดข้อมูลที่รวบรวมมาเพื่อทำการวิจัย 15
3.4 แสดงการพลอตอนุกรมเวลาของชุดข้อมูลทั้งหมด 16
3.5 แสดงผลการพยากรณ์ของ Prophet แบบ Univariate 18
3.6 แสดงผลการพยากรณ์ของ Prophet แบบ Univariate (2) 19
3.7 แสดงผลการพยากรณ์ของ ARIMA 20
3.8 แสดงผลการพยากรณ์ของ Prophet แบบ Multivariate 21
3.9 แสดงผลการพยากรณ์ของ Prophet แบบ Multivariate (2) 22
3.10 ตัวอย่างการผลลัพธ์การแจ้งเตือนผ่าน API ของ Line Notify 23
3.11 ตัวอย่างการใช้งาน Python ผ่าน Command Prompt 24
3.12 หน้าตาการใช้งาน Google Colaboratory 25
3.13 หน้าตาเว็บไซต์ของไลน์ โนติฟาย (Line Notify) 25
4.1 ตัวอย่างการแจ้งเตือนผ่านทางแอปพลิเคชั่นไลน์ (Line) 29

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ณ ปัจจุบันในยุคที่เต็มไปด้วยเทคโนโลยีและความสะดวกสบายในการเชื่อมต่ออินเตอร์เน็ต การทำธุรกิจออนไลน์หรือ e-commerce เป็นส่วนสำคัญและเติบโตอย่างรวดเร็วในเศรษฐกิจโลก การพยากรณ์ปริมาณการเข้าชมเว็บไซต์เป็นสิ่งสำคัญที่ไม่สามารถละเว้นได้ในการที่จะประสบความสำเร็จในการดำเนินธุรกิจ e-commerce

การพยากรณ์ปริมาณการเข้าชมเว็บไซต์ช่วยให้ธุรกิจสามารถวางแผนและดำเนินการในอนาคตอย่างมีประสิทธิภาพ การพยากรณ์ที่แม่นยำช่วยให้ธุรกิจสามารถจัดสรรทรัพยากรและการตลาดให้เหมาะสมและช่วยให้การตัดสินใจทางธุรกิจเป็นไปตามความต้องการและการเติบโตของตลาด e-commerce อย่างมีประสิทธิภาพ

การพยากรณ์ปริมาณการเข้าชมเว็บไซต์ในอนาคตเป็นที่สนใจสำหรับธุรกิจ e-commerce เนื่องจากสามารถนำข้อมูลไปใช้ในการพัฒนาและปรับปรุงผลิตภัณฑ์และบริการ การพยากรณ์ที่แม่นยำช่วยให้ธุรกิจสามารถตอบสนองความต้องการของลูกค้าได้อย่างเหมาะสมและสร้างความพึงพอใจ นอกจากนี้ยังช่วยให้ธุรกิจสามารถตระหนักถึงโอกาสในการเติบโตของธุรกิจ e-commerce อย่างมีประสิทธิภาพ

ในด้านของการทำธุรกิจออนไลน์ การพยากรณ์ปริมาณการเข้าชมเว็บไซต์มาวิเคราะห์ข้อมูล จะสามารถทำให้เข้าใจพฤติกรรมของผู้ใช้งานเว็บไซต์มากยิ่งขึ้น

งานวิจัยนี้จึงนำเสนอการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ในอนาคตด้วยวิธีอนุกรมเวลาและส่งข้อความแจ้งเตือนผ่านทางแอปพลิเคชั่น

1.2 วัตถุประสงค์ของการศึกษาหรือวิจัย

1.2.1 ศึกษาความสามารถและประสิทธิภาพของ Prophet ในการทำนายการเข้าชมเว็บไซต์ในภาค e-commerce การศึกษานี้จะวิเคราะห์และทดสอบความแม่นยำและประสิทธิภาพของ Prophet ในการทำนายปริมาณการเข้าชมเว็บไซต์ e-commerce โดยเปรียบเทียบผลลัพธ์ที่ได้กับโมเดลทางสถิติอื่นๆ

1.2.2 เพื่อให้แนวทางและคำแนะนำที่เกี่ยวข้องสำหรับธุรกิจ e-commerce ผลลัพธ์จากการวิจัยนี้จะมีความสำคัญในการสร้างความเข้าใจและแนะนำแนวทางการปฏิบัติที่มีประสิทธิภาพสำหรับธุรกิจที่ต้องการปรับปรุงกระบวนการทำนายปริมาณการเข้าชมเว็บไซต์ ซึ่งสามารถใช้เพื่อเพิ่มประสิทธิภาพในการวางแผนกลยุทธ์การตลาด เป็นการจัดสรรทรัพยากรให้พอดีต่อการใช้งาน และช่วยให้ธุรกิจสามารถตอบสนองความต้องการของลูกค้าเพื่อที่จะเติบโตในตลาด e-commerce ได้อย่างมีประสิทธิภาพ

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

1.3 ขอบเขตของงานวิจัย

1.3.1 ข้อมูลการเข้าใช้งานเว็บไซต์ Google Merchandise Store ตั้งแต่วันที่ 1 มกราคม 2559 ถึงวันที่ 5 พฤษภาคม 2566

1.3.2 เตรียมข้อมูลให้พร้อมสำหรับการนำมาวิเคราะห์และสร้างโมเดลสำหรับการพยากรณ์ปริมาณการเข้าชมเว็บไซต์

1.3.3 สร้างโมเดลการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ 14 วันล่วงหน้า

1.3.4 สร้างระบบแจ้งเตือนล่วงหน้า 14 วัน เพื่อให้ผู้เกี่ยวข้องได้มีเวลามากพอที่จะดำเนินการ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 เพื่อให้ได้โมเดลตามหลักการอนุกรมเวลา ซึ่งนำไปพยากรณ์ปริมาณการเข้าชมเว็บไซต์ในอนาคต เพื่อบริหารและจัดการในการจัดสรรทรัพยากร รวมทั้งวางแผนด้านการตลาดได้อย่างมีประสิทธิภาพ

1.4.2 นำเทคนิคการพยากรณ์จากงานวิจัยไปประยุกต์ใช้กับชุดข้อมูลในธุรกิจต่างๆต่อได้

1.4.3 ได้เครื่องมือสำหรับแจ้งเตือนผลการพยากรณ์ผ่านทางแอปพลิเคชั่น

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บทที่ 2
แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

สารนิพนธ์นี้มีวัตถุประสงค์เพื่อนำการพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) มาประยุกต์ใช้ในการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ล่วงหน้าสำหรับธุรกิจ e-commerce โดยมีผลลัพธ์เป็นการสร้างระบบแจ้งเตือนผ่านทางแอปพลิเคชั่น เพื่อที่แจ้งเตือนผู้ดูแลระบบให้เตรียมการให้พร้อมรับมือกับปริมาณการเข้าชมเว็บไซต์ที่มากขึ้น จึงจำเป็นต้องศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

2.1 ความหมายและประเภทของ Web Traffic

2.2 ความหมายและประเภทของการพยากรณ์

2.3 พยากรณ์แบบอนุกรมเวลา (Time Series Forecasting)

2.4 Autoregressive Integrated Moving Average (ARIMA)

2.5 Prophet

2.6 งานวิจัยที่เกี่ยวข้อง

2.1 ความหมายและประเภทของ Web Traffic

2.1.1 ความหมายของ web traffic

Web traffic หมายถึงจำนวนข้อมูลที่ถูกรับและส่งผ่านเครือข่ายอินเตอร์เน็ตของเว็บไซต์ในระยะเวลาที่กำหนด โดยมักใช้เพื่อวัดปริมาณการเข้าชมของเว็บไซต์ในแต่ละวัน จำนวน web traffic สามารถสื่อถึงความนิยมและความสำเร็จของเว็บไซต์ได้ในบางครั้ง ยิ่งมีจำนวน web traffic มากขึ้น มักแปลว่ามีผู้เข้าชมหรือผู้ใช้งานมากขึ้นเช่นกัน

การวัด web traffic ช่วยให้ผู้ดูแลเว็บไซต์ได้รับข้อมูลเกี่ยวกับการเข้าชมและการใช้งานของผู้ใช้ ซึ่งสามารถใช้ข้อมูลเหล่านี้ในการวิเคราะห์และปรับแก้กลยุทธ์เพื่อเพิ่มประสิทธิภาพและสร้างประสบการณ์ที่ดีต่อผู้ใช้งาน

2.1.2 ประเภทของ web traffic

(1) Organic Traffic

เป็นการเข้าชมเว็บไซต์ที่เกิดจากผลการค้นหาธรรมชาติในเครื่องมือค้นหา เช่น ผลการค้นหาใน Google ผู้เข้าชมประเภทนี้มักเป็นบุคคลที่มีความสนใจในเนื้อหาที่เว็บไซต์หรือแอปพลิเคชันเว็บมีอยู่แล้ว

(2) Referral Traffic

เป็นการเข้าชมที่เกิดจากการคลิกลิงก์ที่อยู่ในเว็บไซต์อื่นๆ ที่ชี้ไปยังเว็บไซต์ปลายทาง เช่น ลิงก์ที่อยู่ในโพสต์บนสื่อสังคมออนไลน์ หรือลิงก์ที่อยู่ในบทความบนเว็บไซต์อื่น

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

(3) Direct Traffic

เป็นการเข้าชมที่เกิดจากผู้ใช้งานที่พิมพ์ URL ของเว็บไซต์โดยตรงในแถบที่อยู่ของเบราว์เซอร์ หรือใช้บุ๊คมาร์กหรือประวัติการเรียกดูเว็บ ไม่ได้เกิดจากการค้นหาบน Search Engine

(4) Social Media Traffic

เป็นการเข้าชมที่เกิดจากการคลิกลิงก์ที่อยู่ในโพสต์หรืออัพเดตทางสื่อสังคมออนไลน์ เช่น Facebook, Twitter, Instagram เป็นต้น

(5) Paid Traffic

เป็นการเข้าชมที่เกิดจากการโฆษณาและการตลาดทางออนไลน์ที่เสียค่าใช้จ่าย เช่น การโฆษณาบนเครื่องมือค้นหาหรือโฆษณาบนโซเชียลมีเดีย

(6) Campaign Traffic

เป็นการเข้าชมที่เกิดจากแคมเปญการตลาดที่เฉพาะเจาะจง เช่น การส่งอีเมล์โปรโมชั่นหรือการโฆษณาทางอีเมล์

2.2 ความหมายและประเภทของการพยากรณ์

2.2.1 ความหมายของการพยากรณ์

การพยากรณ์เป็นกระบวนการที่ใช้ข้อมูลที่มีอยู่ในอดีตหรือปัจจุบันเพื่อทำนายสิ่งที่จะเกิดขึ้นในอนาคต ซึ่งสามารถนำไปใช้ในการวิเคราะห์เพื่อให้เห็นแนวโน้มของการเปลี่ยนแปลงในสิ่งที่เราต้องการพยากรณ์ โดยใช้การวิเคราะห์และการสร้างโมเดลทางสถิติและคณิตศาสตร์ เพื่อประมาณค่าหรือเหตุการณ์ที่อาจเกิดขึ้นในอนาคต โดยสามารถนำไปใช้ในการวางแผนสำหรับการดำเนินงาน อีกทั้งยังช่วยในการประกอบการตัดสินใจเพื่อให้บรรลุวัตถุประสงค์ของการดำเนินการ

2.2.2 ประเภทของการพยากรณ์ตามระยะเวลา

การพยากรณ์ตามระยะเวลาสามารถแบ่งได้ตามระยะเวลาในการพยากรณ์ออกเป็น 3 ประเภท ดังนี้

(1) การพยากรณ์ระยะสั้น (Short-term Forecasting) เป็นการพยากรณ์ในระยะเวลาไม่เกิน 1 ปี เช่น การพยากรณ์จำนวนผู้ใช้งานในช่วง 1 เดือนล่วงหน้า

(2) การพยากรณ์ระยะปานกลาง (Intermediate Forecasting) เป็นการพยากรณ์ในระยะเวลาระหว่าง 3 เดือนถึง 3 ปี เช่น การพยากรณ์เศรษฐกิจของปีหน้า

(3) การพยากรณ์ระยะยาว (Long-term Forecasting) เป็นการพยากรณ์ในระยะเวลามากกว่า 3 ปีขึ้นไป เช่น การพยากรณ์การเติบโตของอุตสาหกรรมในอีก 5 ปีข้างหน้า

2.2.3 ประเภทของการพยากรณ์ตามคุณลักษณะ

การพยากรณ์ตามคุณลักษณะสามารถแบ่งออกเป็น 2 ประเภท ดังนี้

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

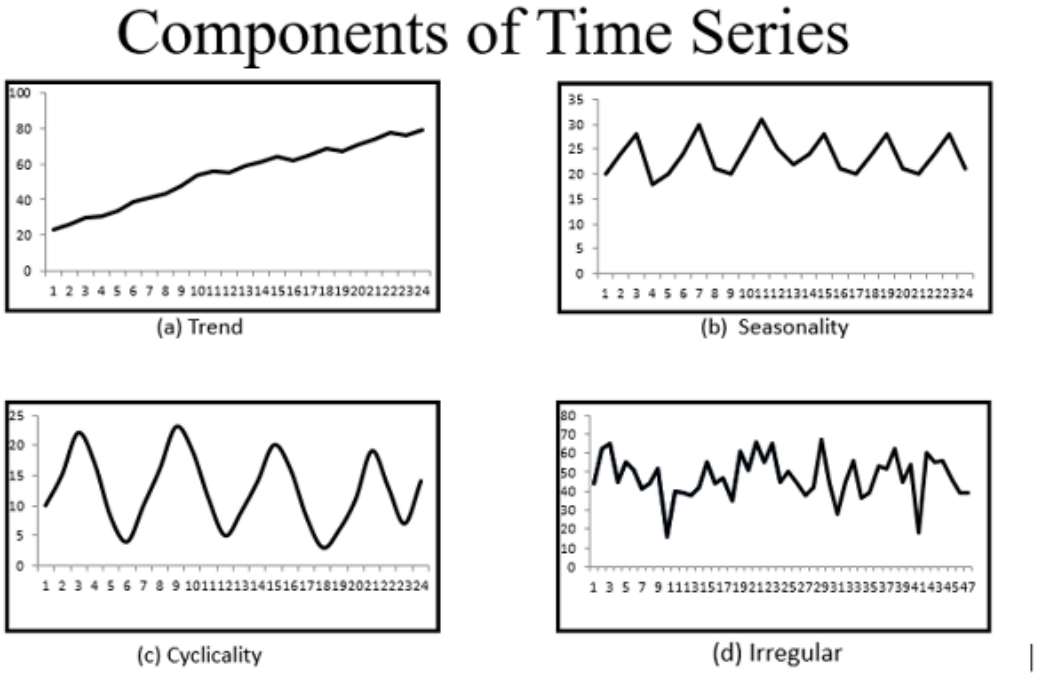
(1) วิธีการเชิงคุณภาพ (Qualitative Methods) เป็นการพยากรณ์โดยอาศัยผู้เชี่ยวชาญในการคาดการณ์ระยะสั้น โดยไม่ใช้โมเดลทางคณิตศาสตร์ จึงตรวจสอบความแม่นยำได้ยากกว่าการพยากรณ์เชิงปริมาณ

(2) วิธีการเชิงปริมาณ (Quantitative Methods) จะใช้สถิติที่มีความหมายและข้อมูลเก่าเพื่อคาดการณ์แนวโน้มในอนาคตในระยะยาว

2.3 พยากรณ์แบบอนุกรมเวลา (Time Series Forecasting)

ในการศึกษานี้จะขอทำการยกตัวอย่างจากการเก็บข้อมูลปริมาณการเข้าชมเว็บไซต์และกิจกรรมอื่นๆ ภายในกรอบระยะเวลาหนึ่งสม่ำเสมอโดยมีช่วงห่างในการเก็บข้อมูลที่เท่ากัน

องค์ประกอบของอนุกรมเวลา (Time Series Components) การพยากรณ์แบบอนุกรมเวลา มีสมมติฐานว่าปริมาณการเข้าชมเว็บไซต์ มีการเปลี่ยนแปลงตามระยะเวลาที่ผ่านไป ดังนั้นการ คำนวณด้วยสมการทางคณิตศาสตร์ ตัวแปรที่มีผลต่อปริมาณการเข้าชมเว็บไซต์คือเวลา (Time) โดยทั่วไป ปริมาณการเข้าชมเว็บไซต์ เกิดจากปัจจัย 4 อย่าง ดังนี้ แนวโน้ม (Trend), ฤดูกาล (Seasonal), วัฏจักร (Cycle) และความผิดปกติ (Irregular)



ภาพที่ 2.1 องค์ประกอบของอนุกรมเวลา

ที่มา: www.analyticsvidhya.com

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

แนวโน้ม (Trend) หมายถึง ปริมาณการเข้าชมเว็บไซต์มีลักษณะเพิ่มขึ้นหรือลดลงอย่างค่อยเป็นค่อยไป โดยส่วนมากอิทธิพลของแนวโน้มมีสาเหตุมาจากเหตุการณ์ต่างๆ การเปลี่ยนแปลงกลุ่มเป้าหมายของเว็บไซต์ การโฆษณา เทคโนโลยีที่มีผลต่อการเข้าชม รวมทั้งกิจกรรมทางการตลาดของเว็บไซต์

ฤดูกาล (Seasonal) หมายถึง ปริมาณการเข้าชมเว็บไซต์ที่เปลี่ยนแปลงมีลักษณะเป็นรูปแบบซ้ำๆ กันจากผลกระทบของฤดูกาลโดยสามารถสังเกตเห็นรูปแบบ Pattern ที่ชัดเจนและเกิดตามรูปแบบนั้นๆ ซ้ำในช่วงเวลาถัดมา โดยหน่วยของระยะเวลาเป็นได้ทั้งรายปี รายเดือน รายวัน หรือรายชั่วโมง ส่วนมากลักษณะปริมาณการเข้าชมเว็บไซต์แบบฤดูกาลมีอิทธิพลมาจากเทศกาลหรือวันหยุดต่างๆ

ตารางที่ 2.1 รูปแบบแต่ละช่วงเวลาของฤดูกาล

รูปแบบช่วงเวลา	ช่วงของฤดูกาล	ระยะในแต่ละรูปแบบ
ปี	ไตรมาส	4 ไตรมาส
ปี	เดือน	12 เดือน
ปี	สัปดาห์	52 สัปดาห์
เดือน	สัปดาห์	4 สัปดาห์
เดือน	วัน	28 – 31 วัน
สัปดาห์	วัน	7 วัน
วัน	ชั่วโมง	24 ชั่วโมง

วัฏจักร (Cycle) หมายถึง ปริมาณการเข้าชมเว็บไซต์ที่มีลักษณะลดลงหรือเพิ่มขึ้น โดยไม่เกี่ยวข้องกับฤดูกาล สามารถมีระยะเวลายาวเท่าหรือยาวกว่าฤดูกาล เช่น การลดลงหรือเพิ่มขึ้นของผู้เข้าชมเว็บไซต์ในระยะยาวเท่ากับหรือยาวกว่าหนึ่งปี

ความผิดปกติ (Irregular) หมายถึง ความเป็นธรรมชาติของข้อมูลที่มีความผิดปกติหรือสิ่งที่ไม่สามารถคาดเดาได้ล่วงหน้า อาจเกิดจากปัจจัยที่ไม่สามารถควบคุมได้หรือเหตุการณ์ที่ไม่คาดคิด เช่น สภาวะฉุกเฉิน สภาวะภัยพิบัติ หรือสภาวะเหตุการณ์ไม่สงบ เป็นต้น ซึ่งสามารถทำให้เกิดการเปลี่ยนแปลงของปริมาณการเข้าชมเว็บไซต์ หากพบเหตุการณ์เช่นนี้ในการพยากรณ์จะตัดข้อมูลที่ผิดปกติทิ้งไปเลย ก่อนที่จะทำการพยากรณ์ต่อไป

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

2.4 Autoregressive Integrated Moving Average (ARIMA)

ARIMA (Autoregressive Integrated Moving Average) เป็นโมเดลเชิงสถิติที่ใช้เพื่อการพยากรณ์และวิเคราะห์อนุกรมเวลา (Time Series Analysis) โดยเฉพาะ โดยมีความสามารถในการจัดการกับอิทธิพลของแนวโน้ม (Trend) และสภาวะซ้ำซ้อน (Seasonality) ซึ่งเป็นความสามารถที่สำคัญในการพยากรณ์และการวิเคราะห์แนวโน้มของข้อมูลตามเวลาในอนาคต

ARIMA ประกอบด้วยสามส่วนหลัก คือ AR (Autoregressive), I (Integrated), และ MA (Moving Average) ดังต่อไปนี้

AR (Autoregressive) เป็นส่วนที่ใช้ในการจัดการกับแนวโน้ม (Trend) โดยพิจารณาปริมาณการเข้าชมเว็บไซต์ก่อนหน้าเป็นตัวแปรอิสระในการพยากรณ์ค่าความต้องการสินค้าปัจจุบัน โดยใช้ปริมาณการเข้าชมเว็บไซต์ในอดีตและพารามิเตอร์ที่เกี่ยวข้อง เช่น ปริมาณการเข้าชมเว็บไซต์ในเดือนก่อนหน้าหรือในไตรมาสก่อนหน้า

I (Integrated) เป็นส่วนที่ใช้ในการจัดการกับองค์ประกอบของการแนวโน้ม (Trend) โดยการเปลี่ยนแปลงแนวโน้มให้อยู่ในรูปของอนุกรมเวลาที่เป็นแนวเดียวกัน ด้วยการทำความแตกต่าง (Difference) ระหว่างปริมาณการเข้าชมเว็บไซต์ปัจจุบันและในเวลาก่อนหน้า

MA (Moving Average) เป็นส่วนที่ใช้ในการจัดการกับสภาวะซ้ำซ้อน (Seasonality) โดยใช้ค่าเฉลี่ยเคลื่อนที่ของปริมาณการเข้าชมเว็บไซต์ในช่วงเวลาที่กำหนด เพื่อปรับปรุงและลดสภาวะซ้ำซ้อนออกจากข้อมูล

ARIMA ช่วยให้เราสามารถพยากรณ์ปริมาณการเข้าชมเว็บไซต์ในอนาคตได้ โดยพิจารณาแนวโน้มที่เกิดขึ้นในอดีต และปรับปรุงด้วยการตรวจสอบและตั้งค่าพารามิเตอร์ที่เหมาะสม เช่น จำนวนของระยะเวลาที่นำไปใช้ในการพยากรณ์ (lag order) และการเลือกโมเดล ARIMA ที่เหมาะสมสำหรับข้อมูลที่มีแนวโน้ม ฤดูกาล และวัฏจักรต่าง ๆ ภายในนั้น

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

รูปแบบ	สมการของ (Y)
ARIMA(1,1,0)/AR(1)	$\Delta Y_t = \alpha + \phi_1 \Delta Y_{t-1} + \varepsilon_t$
ARIMA(2,1,0)/AR(2)	$\Delta Y_t = \alpha + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \varepsilon_t$
ARIMA(0,1,1)/MA(1)	$\Delta Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$
ARIMA(0,1,2)/MA(2)	$\Delta Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$
ARMA(1,1,1)	$\Delta Y_t = \alpha + \phi_1 \Delta Y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$

ภาพที่ 2.2 ตัวอย่างสมการของโมเดล ARIMA

ที่มา: https://medium.com/qunt-i-love-u/การพยากรณ์ข้อมูลอนุกรมเวลาด้วยเทคนิค-arima-ด้วย-python-44809eb8e990

2.5 Prophet

Prophet เป็นโอเพนซอร์สำหรับการพยากรณ์ข้อมูลแบบอนุกรมเวลา สามารถใช้งานได้ทั้งบน R และ Python ซึ่งถูกพัฒนาโดยกลุ่มนักวิทยาศาสตร์ข้อมูลภายในองค์กร Facebook (ที่รู้จักในปัจจุบันว่า Meta) ได้นำมาใช้ในงานพยากรณ์ภายใน Facebook ในหลายโครงการ และได้ทำการปล่อยให้ใช้งานกันแบบสาธารณะในปี 2017 อัลกอริทึมของ Prophet จะทำงานได้ดีกับข้อมูลที่มีหลายฤดูกาล มีปริมาณข้อมูลมาก และอธิบายคุณลักษณะของฤดูกาลของข้อมูลได้

Sean J. Taylor และ Ben Letham คือกลุ่มผู้พัฒนา Prophet ซึ่งมีความต้องการให้ Prophet เป็นโมเดลที่ใช้งานได้ง่าย โดยที่ไม่จำเป็นต้องมีความรู้เกี่ยวกับรายละเอียดของโมเดลพื้นฐานมาก อธิบายผลได้ง่าย โดยการแบ่งการทำงานออกเป็นส่วน ซึ่งมีคุณลักษณะเด่นๆของ Prophet มีดังนี้

มีความยืดหยุ่น คือสามารถจัดการกับงานพยากรณ์อนุกรมเวลาได้หลายรูปแบบ เช่น ข้อมูลที่ไม่มีความสม่ำเสมอหรือข้อมูลที่ขาดหาย จัดการกับ outlier ได้ดี เป็นต้น

การจำแนกแนวโน้ม คือสามารถตรวจจับและจัดการแนวโน้มในขข้อมูลได้หลายรูปแบบ ไม่ว่าเป็น Linear หรือ Non-Linear รวมถึงการเปลี่ยนแปลงของแนวโน้ม

การจำแนกฤดูกาล คือการที่ Prophet สามารถจัดการรูปแบบฤดูกาลได้หลายช่วงเวลา เช่น รายเดือน หรือรายปี โดยที่จะนำรูปแบบดังกล่าวไปใช้ในการพยากรณ์

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

การวิเคราะห์ผลกระทบจากวันหยุด คือช่วยให้ผู้ใช้ระบุเหตุการณ์ วันหยุด เทศกาล ที่กำหนดเองและคาดว่าจะส่งผลกระทบต่อข้อมูล เพื่อให้โมเดลรับรู้ถึงความแตกต่างระหว่างฤดูกาลหรือวัฏจักรที่เกิดจากวันหยุดหรือเหตุการณ์พิเศษ

การแสดงผลและการวิเคราะห์ข้อมูล คือความสามารถในการรายงานผลการพยากรณ์ ออกมาในรูปแบบกราฟต่างๆ เช่น แนวโน้มของผลพยากรณ์รายวัน รายสัปดาห์ รายเดือน และรายปี, แสดงความสำคัญของวันหยุดต่างๆที่ส่งผลต่อการพยาการณ์ เป็นต้น

Prophet มีพื้นฐานมาการวิเคราะห์ จาก 3 ส่วนประกอบหลัก ซึ่งอ้างอิงมาจาก Additive model คือ แนวโน้มของพฤติกรรมของข้อมูล (Trend), พฤติกรรมของข้อมูลที่ไม่ได้เป็นเชิงเส้นจะถูกอธิบายด้วยฤดูกาล (Seasonality) ในช่วงระยะเวลารายปี รายเดือน รายวัน, และส่วนสุดท้ายคือ ผลกระทบที่เกิดจากวันหยุด (Effect of holidays) ซึ่งจะช่วยให้โมเดลเรียนรู้ ช่วยทำให้มีความแม่นยำสูงขึ้น ดังสมการต่อไปนี้

$$y(t) = g(t) + s(t) + h(t) + et \qquad (1)$$

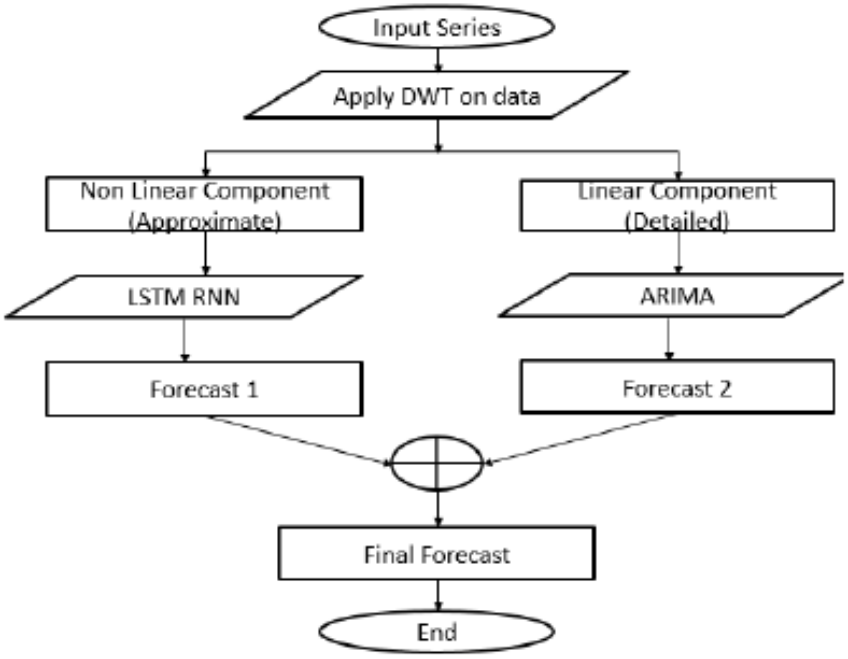
โดยที่ g(t) คือ แนวโน้ม (Trend)
s(t) คือ ฤดูกาล (Seasonality)
h(t) คือ ผลกระทบจากวันหยุด (Effect of holidays)
et คือ ความผิดปกติหรือความไม่แน่นอน (Error)

Prophet นำเสนอโมเดลสำหรับการพยากรณ์ข้อมูลแบบอนุกรมเวลาที่ใช้ saturating growth model และ piecewise linear model สำหรับการวิเคราะห์แนวโน้ม (Trend), ใช้อนุกรมฟูเรียร์เป็นฐานในการวิเคราะห์ฤดูกาล (Seasonality) เพื่อใช้ในการสร้างโมเดลที่ยืดหยุ่นในการตรวจจับและวิเคราะห์ฤดูกาลของชุดข้อมูล และในส่วนของผลกระทบจากวันหยุด (Effect of holidays) ที่มีผลต่อข้อมูลนอกเหนือจากวันหยุดที่เกิดขึ้นซ้ำในช่วงเวลาเดียวกันของปี ยังมีเหตุการณ์จำนวนมากและวันหยุดที่มีวันที่ไม่แน่นอนในแต่ละปี ที่ส่งผลกระทบต่อพฤติกรรมของข้อมูล นอกจากนี้วันหยุดยังแตกต่างกันตามแต่ละประเทศ ซึ่งอ้างอิงจากปฏิทินจันทรคติ Prophet ให้ความสำคัญกับวันหยุดสากลและวันหยุดของแต่ละประเทศ และสมมติให้เหตุการณ์ วันหยุด เป็นอิสระต่อกัน และนำมาใช้งานคล้ายกับการวิเคราะห์ฤดูกาล

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

2.6 งานวิจัยที่เกี่ยวข้อง

Tejas Shelatkar, Stephen Tondale, Swaraj Yadav, and Sheetal Ahir, (2020) ได้นำเสนอการพยากรณ์ข้อมูลอนุกรมเวลาด้วยการ ใช้ Discrete Wavelet Transform (DWT) แบ่งข้อมูลที่ได้มาจาก Wikipedia Pageview API ออกเป็น 2 ส่วน ในขั้นตอนการจัดเตรียมข้อมูล ทำให้ได้ส่วนที่เป็น Detail และ Approximate เพื่อที่จะแบ่งให้ ARIMA และ LSTM RNN ช่วยกันพยากรณ์ และจะนำผลการพยากรณ์กลับมารวมกันด้วยวิธี Invert Discrete Wavelet Transform (iDWT) ทำให้ได้ผลการพยากรณ์ที่แม่นยำมากขึ้น



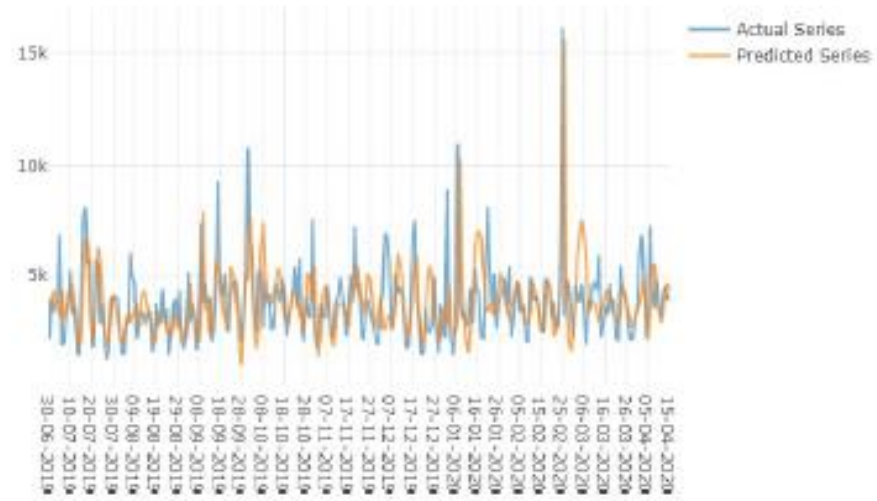
ภาพที่ 2.3 ขั้นตอนการทำงานของงานวิจัย (1)

ที่มา: Tejas Shelatkar, Stephen Tondale, Swaraj Yadav, and Sheetal Ahir, (2020)

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ภาพที่ 2.4 ผลการพยากรณ์ของ ARIMA

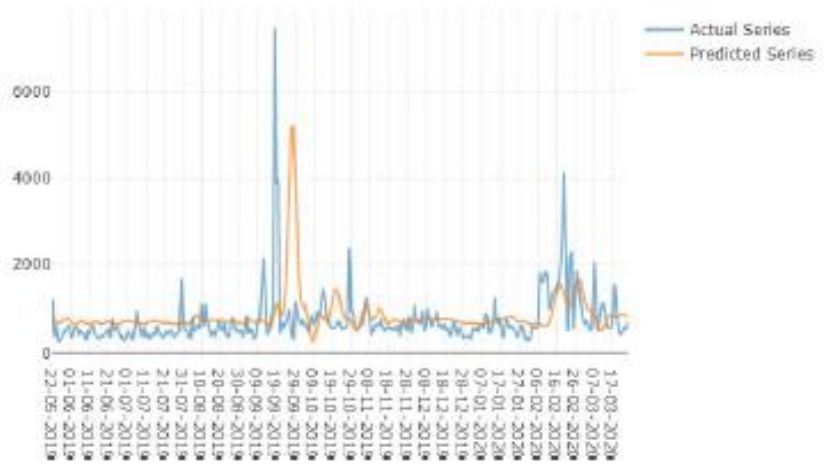
ที่มา: Tejas Shelatkar, Stephen Tondale, Swaraj Yadav, and Sheetal Ahir, (2020)



ภาพที่ 2.5 ผลการพยากรณ์ของ LSTM RNN

ที่มา: Tejas Shelatkar, Stephen Tondale, Swaraj Yadav, and Sheetal Ahir, (2020)

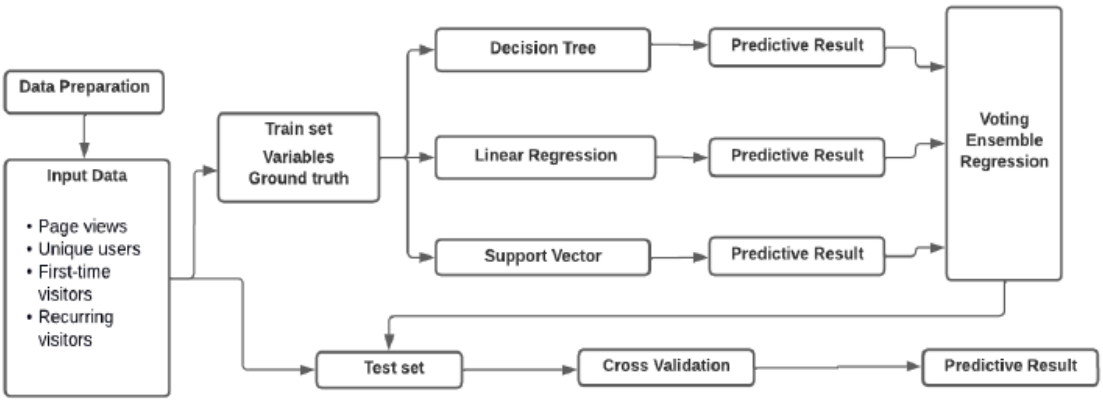
DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



ภาพที่ 2.6 ผลการพยากรณ์ของการรวม ARIMA และ LSTM RNN เข้ากันด้วย DWT

ที่มา: Shelatkar, T., Tondale, S., Yadav, S., & Ahir, S. (2020).

D. Sikka and C. N. S. Vinoth Kumar (2023) ได้นำเสนอการพยากรณ์ข้อมูลอนุกรมเวลาด้วยการใช้ ensemble model ที่ประกอบไปด้วยโมเดลหลายแบบได้แก่ Decision Tree, Multiple Linear Regression และ Support Vector Machine ในการพยากรณ์ข้อมูล pageview และ visitors ในการหาจำนวน unique visitors จาก statforecasting.com จากโมเดลดังกล่าวได้ค่าความแม่นยำสูงถึง 99.96% และค่าความคลาดเคลื่อนเพียง 0.24%



ภาพที่ 2.7 ขั้นตอนการทำงานของงานวิจัย (2)

ที่มา: D. Sikka and C. N. S. Vinoth Kumar (2023)

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

MODEL USED	R^2	MAPE
3-Algorithm Voting Regressor Ensemble Model	0.9996	0.0024
Linear Regression	0.9712	0.0041
Support Vector	0.9856	0.0037
Decision Trees	0.9843	0.0098

ภาพที่ 2.8 ผลการทดลองของงานวิจัย (2)

ที่มา: D. Sikka and C. N. S. Vinoth Kumar (2023)

N. Petluri and E. Al-Masri (2018) ได้นำเสนอการพยากรณ์อนุกรมเวลาด้วยวิธีการสร้างตัวแปรใหม่ (Feature Engineering) จากโมเดลเดิมที่มีคือ RNN seq2seq โดยที่ผู้วิจัยได้ตั้งสมมติฐานขึ้นมาว่าการสร้างตัวแปรใหม่ขึ้นมา จะช่วยส่งเสริมการเรียนรู้ให้กับโมเดล ซึ่งจะทำให้ประสิทธิภาพของโมเดลที่มีอยู่นั้นสูงขึ้นได้

Model and its features	Training set(days)	SMAPE
Existing model	804	0.351
	740	0.354
Model with only 7 day median and page popularity	804	0.350
Model with 7, 30, 90, 180 days median as independent features	804	0.349
Model with median of medians with Fibonacci sequence of variable window length	804	0.349
Model with all the newly added features	740	0.351

ภาพที่ 2.9 ผลการทดลองของงานวิจัย (3)

ที่มา: N. Petluri and E. Al-Masri (2018)

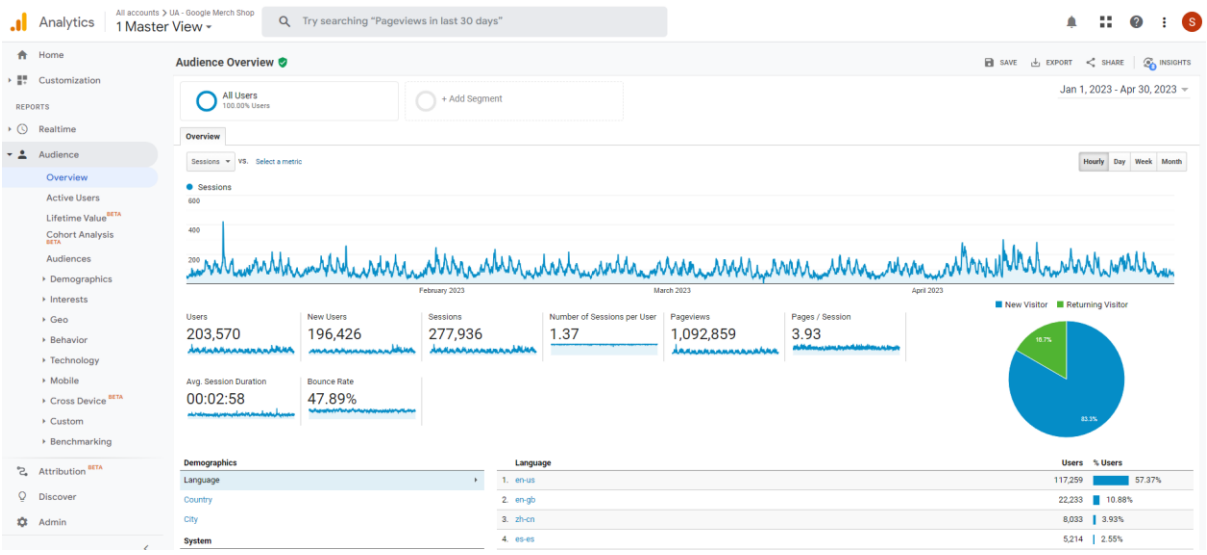
DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บทที่ 3
ระเบียบวิธีวิจัย

สารนิพนธ์นี้มีวัตถุประสงค์เพื่อนำการพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) มาประยุกต์ใช้ในการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ล่วงหน้าสำหรับธุรกิจ e-commerce โดยมีผลลัพธ์เป็นการสร้างระบบแจ้งเตือนผ่านทางแอปพลิเคชั่น เพื่อที่แจ้งเตือนผู้ดูแลระบบให้เตรียมการให้พร้อมรับมือกับปริมาณการเข้าชมเว็บไซต์ที่มากขึ้น โดยมีขั้นตอนการดำเนินการดังนี้

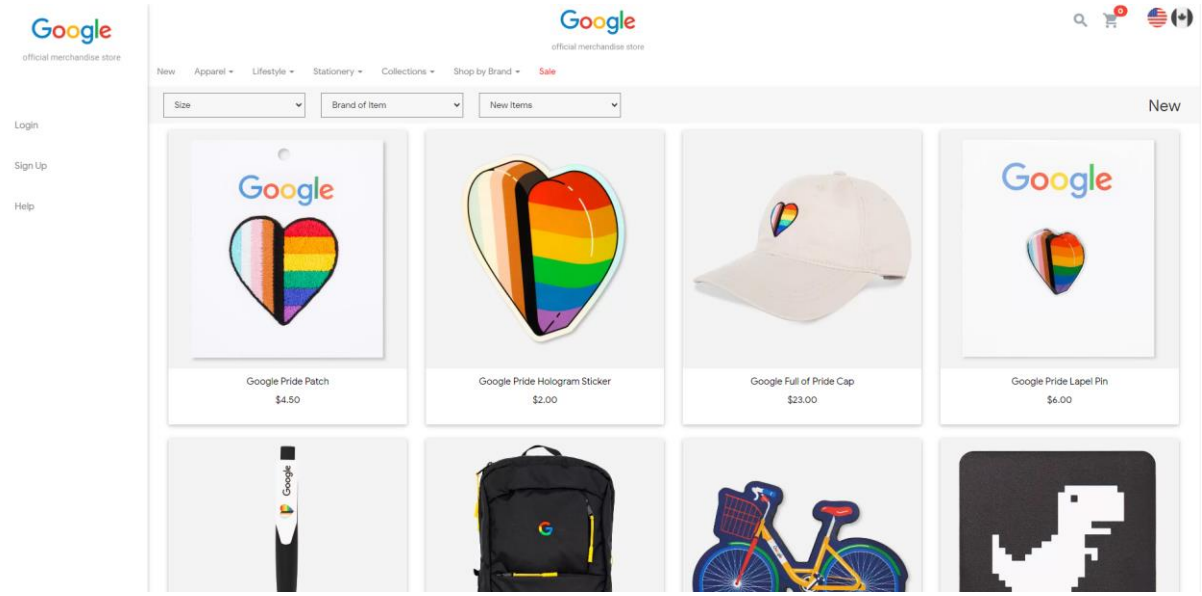
3.1 การรวบรวมข้อมูล

งานวิจัยนี้ได้ใช้ข้อมูล Web Traffic จาก Google Merchandise Store ตั้งแต่วันที่ 1 มกราคม 2559 ถึง 5 พฤษภาคม 2566 ซึ่งถูกรวบรวมโดย Google Analytics ซึ่งเป็นเครื่องมือวิเคราะห์และรายงานข้อมูลที่ให้บริการโดย Google เพื่อวัดและวิเคราะห์การใช้งานเว็บไซต์ โดย Google Analytics จะเก็บข้อมูลเกี่ยวกับการเข้าชมเว็บไซต์ เช่น จำนวนผู้เข้าชม เวลาการใช้งาน การเรียกดูหน้าเว็บ และพฤติกรรมของผู้ใช้ ซึ่งข้อมูลเหล่านี้จะถูกนำมาวิเคราะห์ให้ทราบถึงสถิติการใช้งานและประสิทธิภาพของเว็บไซต์



ภาพที่ 3.1 Dashboard ของ Google Analytics

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



ภาพที่ 3.2 เว็บไซต์ Google Merchandise Store (https://shop.googlemerchandisestore.com/)

	A	B
1	date	sessions
2	01/01/16	1929
3	01/02/16	1530
4	01/03/16	1567
5	01/04/16	1938
6	01/05/16	1840
7	01/06/16	1852
8	01/07/16	1920
9	01/08/16	1734
10	01/09/16	1466
11	01/10/16	1592

ภาพที่ 3.3 ตัวอย่างชุดข้อมูลที่รวบรวมมาเพื่อทำการวิจัย

3.2 การจัดเตรียมข้อมูล

3.2.1 การเปลี่ยนรูปแบบข้อมูล

เนื่องจากบางครั้งข้อมูลอาจจะไม่ได้อยู่ในรูปแบบที่เหมาะสมสำหรับการใช้งาน เช่น ข้อมูลวันที่และข้อมูลอื่นๆ ถูกเก็บเป็นข้อความ (String) แทนที่จะอยู่ในรูปแบบที่สามารถนับหรือคำนวณได้ จึงต้องทำการเปลี่ยนแปลงรูปแบบของข้อมูลให้อยู่ในรูปแบบที่ถูกต้อง

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

3.2.2 การตั้งชื่อคอลัมน์

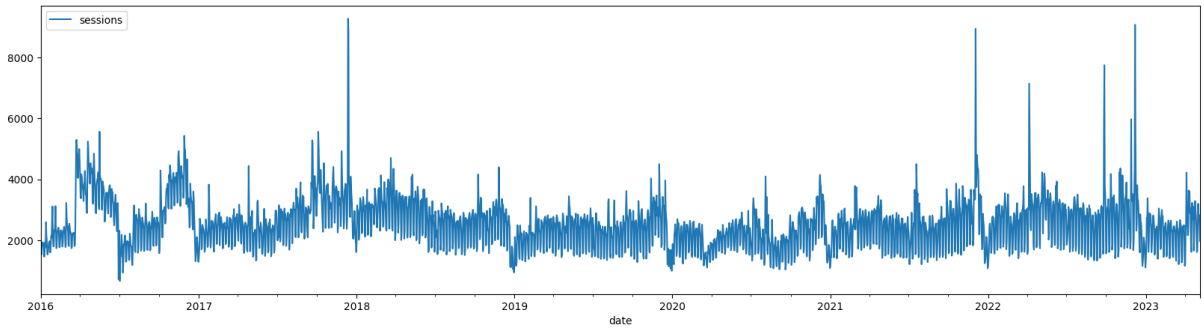
เนื่องจากข้อกำหนดการใช้งาน Prophet นั้นต้องเปลี่ยนชื่อคอลัมน์ที่เก็บข้อมูลเวลา วันที่ เป็น ‘ds’ ซึ่งมาจากคำว่า datestamp หรือ datetime stamp ส่วนคอลัมน์ที่เก็บข้อมูลที่ต้องการพยากรณ์เป็น ‘y’ เพื่อแสดงถึงตัวแปรที่คาดว่าจะมีความเปลี่ยนแปลงตามเวลา (dependent variable) ที่ต้องการทำนาย

3.2.3 การลบข้อมูลที่ไม่เกี่ยวข้อง

หากมีคอลัมน์ที่ไม่จำเป็นต่อการพยากรณ์ เราสามารถลบข้อมูลเหล่านั้นออกจาก DataFrame ได้ เพื่อให้ข้อมูลเหล่านั้นไม่ส่งผลกระทบต่อการทำนาย ซึ่งการลบคอลัมน์ที่ไม่เกี่ยวข้องจะช่วยลดขนาดของข้อมูลที่ต้องประมวลผล และช่วยให้โมเดลทำงานได้เร็วขึ้น

3.2.4 การจัดเรียงข้อมูล

ถ้าข้อมูลไม่ได้เรียงตามลำดับเวลา ควรจัดเรียงลำดับข้อมูลเพื่อให้เข้ากับการพยากรณ์ การเรียงลำดับข้อมูลตามเวลาจะช่วยให้โมเดลสามารถทำนายตามลำดับเวลาได้ถูกต้อง และผลลัพธ์ที่ได้จะมีความสอดคล้องกับลำดับเวลาที่ต้องการพยากรณ์ สามารถช่วยให้การพยากรณ์มีความแม่นยำและมีประสิทธิภาพสูงขึ้น



ภาพที่ 3.4 แสดงการพลอตอนุกรมเวลาของชุดข้อมูลทั้งหมด

3.3 การสร้างตัวแปร (Feature Engineering)

เป็นกระบวนการที่ผู้วิเคราะห์ข้อมูลทำเพื่อสร้างคุณลักษณะ (features) หรือตัวแปรที่เหมาะสมและมีประโยชน์ในการสร้างโมเดล (models) หรือวิเคราะห์ข้อมูล ซึ่งคุณลักษณะเหล่านี้จะถูกนำเข้าไปใช้ในการประมวลผลและการตัดสินใจในการแยกแยะข้อมูล การทำ feature engineering มีเป้าหมายเพื่อเพิ่มประสิทธิภาพของโมเดล ลดความซับซ้อน และเพิ่มความเข้าใจในข้อมูล

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

3.3.1 is_weekend

สร้างตัวแปรขึ้นมาใหม่โดยอ้างอิงจากวันที่ เพื่อแสดงว่าวันที่เหล่านั้นเป็นวันเสาร์-อาทิตย์หรือไม่ โดยที่จะแสดงผลเป็น 1 ถ้าไม่ใช่วันเสาร์-อาทิตย์ จะแสดงผลเป็น 0

3.3.2 week_no

สร้างตัวแปรขึ้นมาใหม่โดยอ้างอิงจากวันที่ เพื่อแสดงว่าเป็นสัปดาห์ที่เท่าไรในแต่ละปี ซึ่งจะมีค่าอยู่ที่ระหว่าง 1-52

3.3.2 avg_session_duration

รวบรวมข้อมูลระยะเวลาเฉลี่ยของการเข้าชมเว็บแต่ละครั้งเพิ่มเติม จาก Google Analytics ซึ่งจะมีหน่วยเป็นวินาที

3.3.4 สร้างตัวแปร e-commerce holidays

สร้างตัวแปรขึ้นมาใหม่จากการค้นหาข้อมูลของวันที่ของวันเทศกาลสำหรับธุรกิจ e-commerce ในประเทศสหรัฐอเมริกา เพื่อช่วยเพิ่มรายละเอียดและการเรียนรู้ให้โมเดลในการพยากรณ์ได้ ที่เพิ่มเติมจากวันหยุดที่มีอยู่ในฐานข้อมูลเดิมของ Prophet โดยที่มีวันเทศกาลต่างๆตามตารางดังนี้

ตารางที่ 3.1 ตารางเทศกาลของ e-commerce holidays

เทศกาลและวันหยุด	วันที่
St. Patrick's Day	17 มีนาคม ของทุกปี
Easter	ในเดือนเมษายน วันที่ไม่แน่นอน
Halloween	31 ตุลาคม ของทุกปี
Singles’ Day	11 พฤศจิกายน ของทุกปี
Black Friday	วันศุกร์ที่สี่ของเดือนพฤศจิกายน
Cyber Monday	วันจันทร์แรกที่ถัดจาก Black Friday
Super Saturday	วันเสาร์ก่อนที่จะถึงเทศกาลคริสมาสต์
Christmas Eve	24 ธันวาคม ของทุกปี
New Year's Eve	31 ธันวาคม ของทุกปี

3.4 การสร้างโมเดล

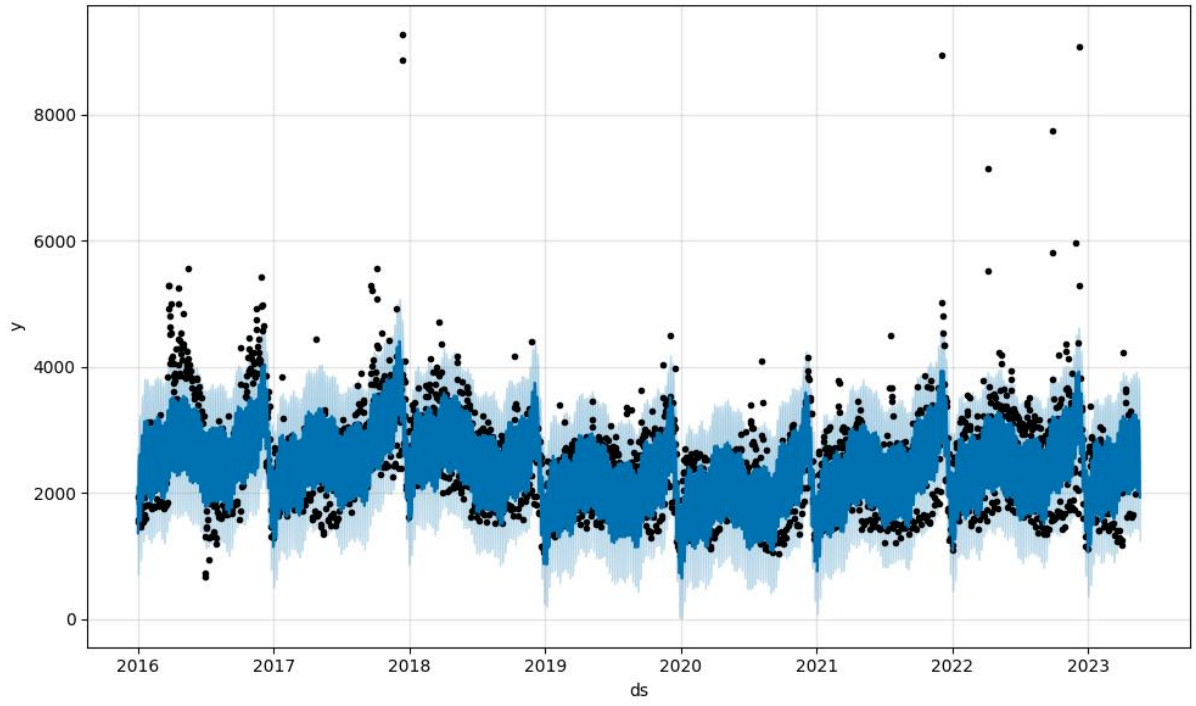
หลังจากได้ข้อมูลปริมาณการเข้าชมเว็บไซต์แล้ว ก็นำไปสร้างโมเดลโดยงานวิจัยนี้ได้ทดลองกับโมเดล Prophet และ ARIMA โดยจะเปรียบเทียบกันด้วยวิธีการพยากรณ์แบบตัวแปรเดี่ยว (Univariate) เพราะว่า ARIMA สามารถพยากรณ์แบบ Univariate ได้เท่านั้น เพื่อที่จะเลือกว่าโมเดลไหนเหมาะสมกับการ

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

พยากรณ์ในครั้งนี้ โดยจะคัดเลือกจาก Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) และ Root Mean Square Error (RMSE) และความสามารถในการพัฒนาต่อไปได้ของโมเดล

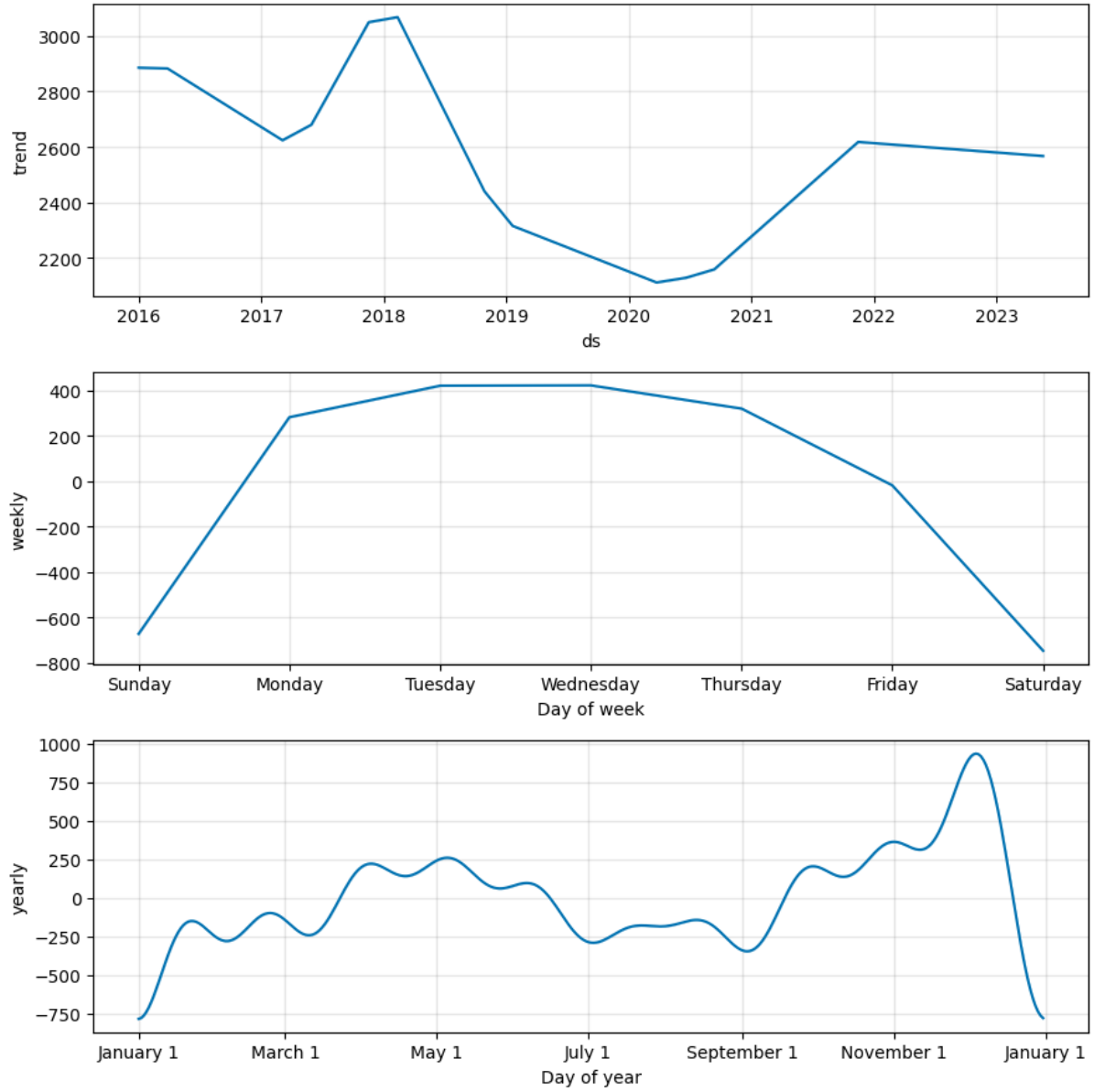
3.4.1 โมเดล Prophet แบบ Univariate (Baseline Model)

โดยวิธีการเริ่มนำชุดข้อมูลสร้างโมเดลเพื่อทำการพยากรณ์ล่วงหน้า 14 วัน โดยค่าพารามิเตอร์ที่ใช้สำหรับการสร้างโมเดลได้แก่ seasonality_mode='multiplicative' และ make_future_dataframe (period=14, freq='D')



ภาพที่ 3.5 แสดงผลการพยากรณ์ของ Prophet แบบ Univariate

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

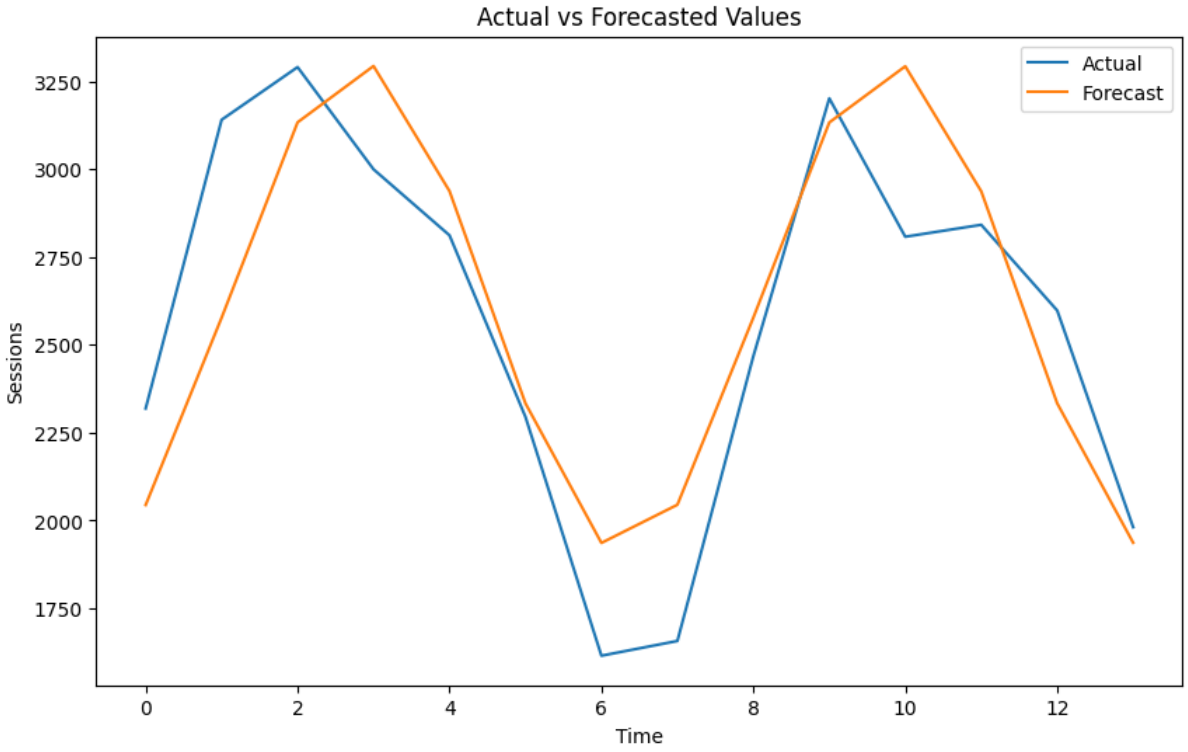


ภาพที่ 3.6 แสดงผลการพยากรณ์ของ Prophet แบบ Univariate (2)

3.4.2 โมเดล ARIMA (Baseline Model)

โดยวิธีการเริ่มนำชุดข้อมูลสร้างโมเดลเพื่อทำการพยากรณ์ล่วงหน้า 14 วัน ค่าพารามิเตอร์ที่ใช้ในการสร้างโมเดล ARIMA โดยใช้ค่า (p, d, q) เท่ากับ (2, 1, 3) ตามลำดับ

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



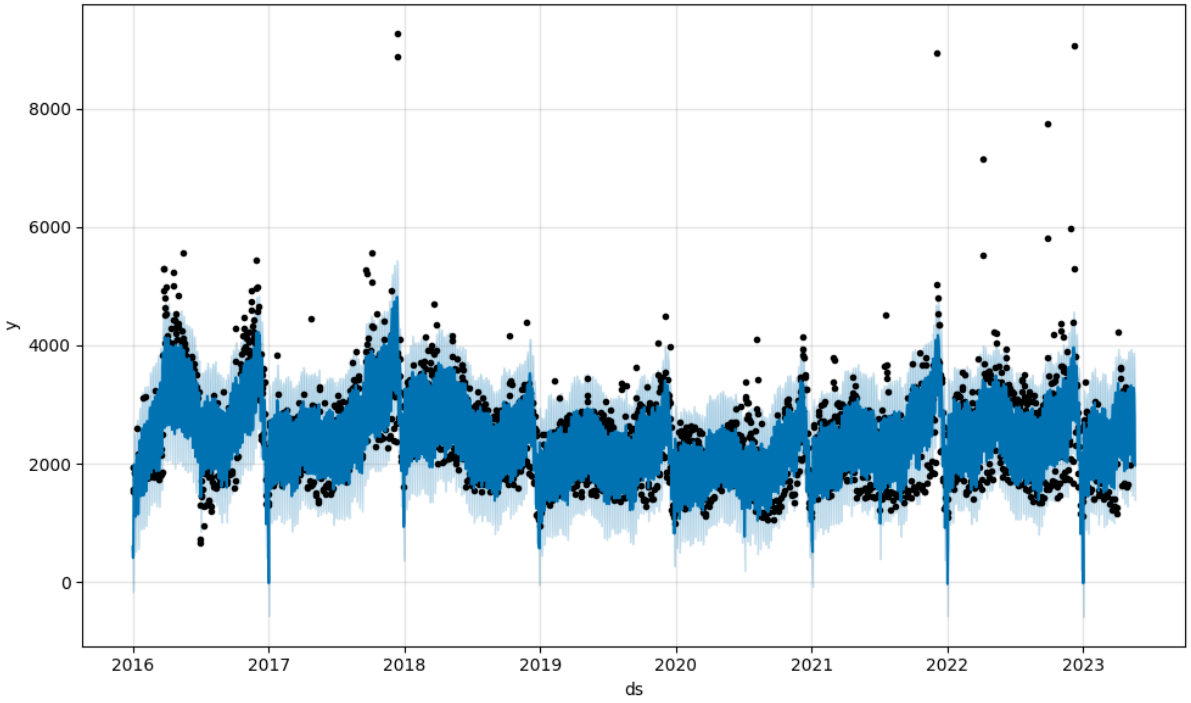
ภาพที่ 3.7 แสดงผลการพยากรณ์ของ ARIMA

3.4.3 โมเดล Prophet (Multivariate)

โดยวิธีการเริ่มนำชุดข้อมูลสร้างโมเดลเพื่อทำการพยากรณ์ล่วงหน้า 14 วัน โดยค่าพารามิเตอร์ที่ใช้สำหรับการสร้างโมเดลได้แก่ seasonality_mode='multiplicative' และ make_future_dataframe (period=14, freq='D') เหมือนกับในโมเดลแบบ Prophet (Univariate) แต่เพิ่มเติมในส่วนของการสร้าง holidays และ e-commerce holidays เข้ามาตามพารามิเตอร์ได้แก่ make_holidays_df(country = 'US') รวมทั้งทำการสร้างฐานข้อมูลของ e-commerce holidays ขึ้นมา

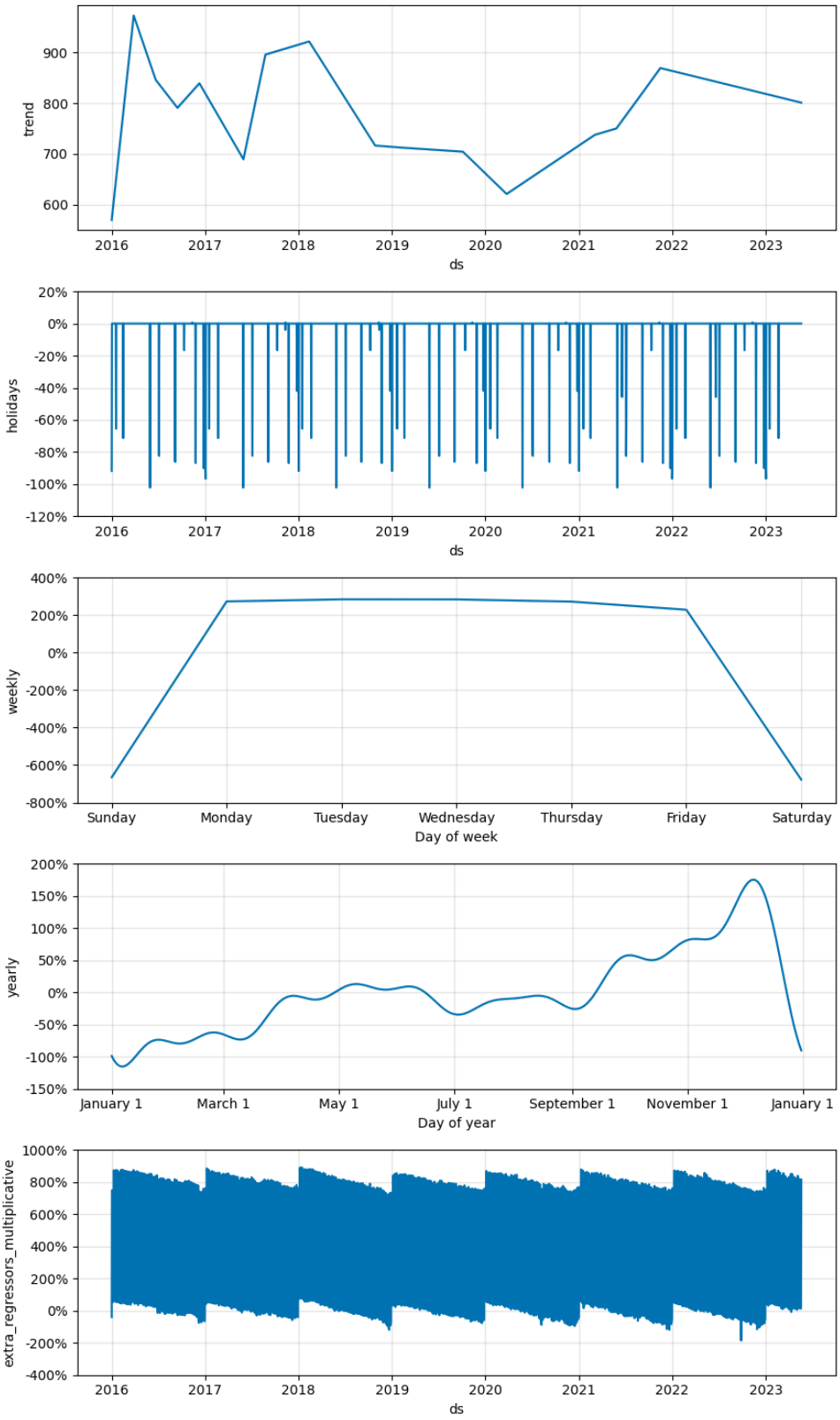
เมื่อนำ e-commerce holidays ทั้งหมดที่สร้างขึ้นมามารวมเข้าด้วยกันเป็น dataframe เพื่อที่จำเก็บไว้เป็นตัวแปรที่จะใช้งาน และเพิ่มตัวแปร is_weekend และ week_no มาช่วยในการพยากรณ์ รวมทั้งใช้คำสั่ง add_regressor ให้กับตัวแปรที่สร้างขึ้นมาใหม่ เพื่อที่โมเดลจะได้นำค่าตัวแปรเหล่านี้ไปช่วยในการพยากรณ์แบบ multivariate

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



ภาพที่ 3.8 แสดงผลการพยากรณ์ของ Prophet แบบ Multivariate

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



ภาพที่ 3.9 แสดงผลการพยากรณ์ของ Prophet แบบ Multivariate (2)

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

3.4.4 การประเมินผล

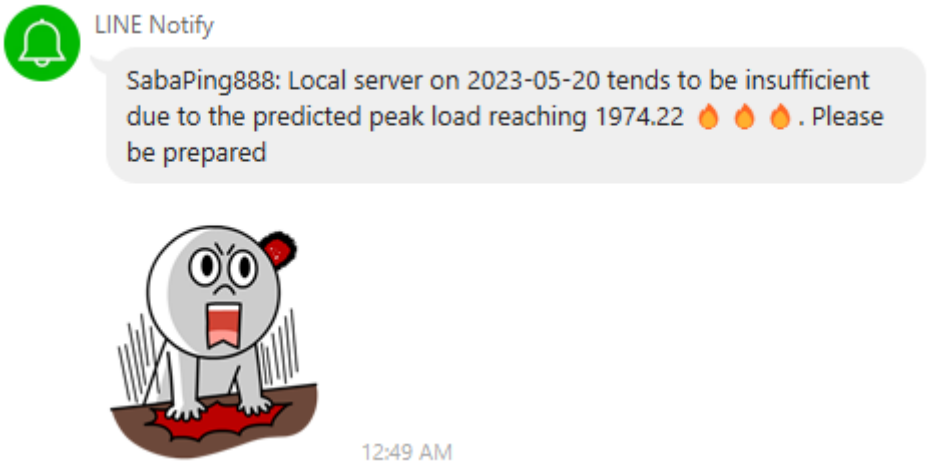
ขั้นตอนนี้เป็นการวัดประสิทธิภาพของผลลัพธ์ที่ได้ว่ามีความน่าเชื่อถือมากน้อยเพียงใด โดยทำการวัดผลจากค่าความคลาดเคลื่อน (Error) ต่างๆ เช่น Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) และ Root Mean Square Error (RMSE)

3.4.5 การกำหนดเกณฑ์ที่จะนำผลลัพธ์ไปใช้งาน

ขั้นตอนนี้เป็นการกำหนดเกณฑ์ที่จะใช้ในการนำไปใช้ในการแจ้งเตือน โดยในงานวิจัยนี้เลือกใช้วิธีการหา Outlier จาก Interquartile range (IQR) เพื่อที่จะนำมาใช้เป็นเกณฑ์ ซึ่งจะเลือกใช้ค่า upper bound คือการนำค่าใน ควอไทล์ที่ 3 มาบวกกับ 1.5 เท่าของ IQR

3.4.6 การนำไปใช้งาน

เป็นการนำโมเดลไปใช้งานในการแจ้งเตือนผู้เกี่ยวข้อง เมื่อค่าที่พยากรณ์ได้มีค่าสูงกว่าเกณฑ์ที่ตั้งไว้ โดยเป็นการส่งข้อความผ่านแอปพลิเคชั่นไลน์ (Line) ด้วยการใช้ API ของ Line Notify เพื่อให้ผู้เกี่ยวข้องได้เตรียมพร้อมในการดำเนินการขั้นต่อไป



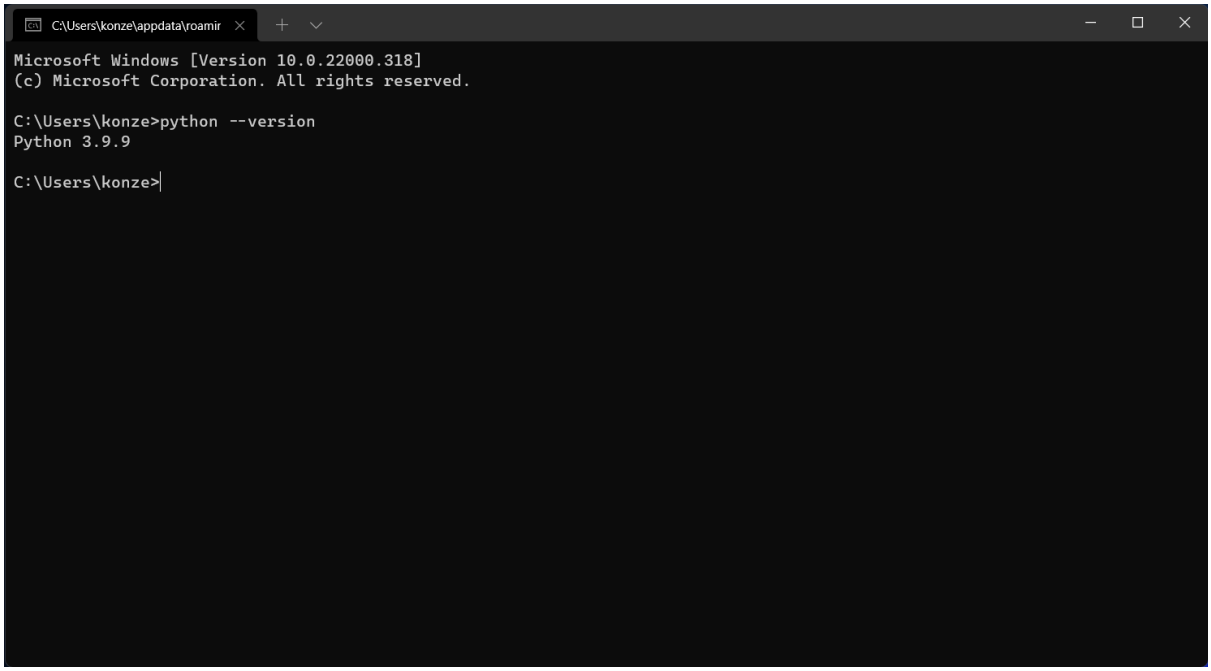
ภาพที่ 3.10 ตัวอย่างผลลัพธ์การแจ้งเตือนผ่าน API ของ Line Notify

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

3.5 เครื่องมือที่ใช้ในงานวิจัย

3.5.1 ภาษาไพทอน (Python programming language)

ภาษาไพทอน (Python) เป็นภาษาโปรแกรมที่ได้รับความนิยมสูงและใช้กันอย่างแพร่หลายในวงการโปรแกรมมิ่ง โดยเฉพาะในการพัฒนาซอฟต์แวร์และการวิเคราะห์ข้อมูล ภาษาไพทอนมีวัตถุประสงค์หลักในการสื่อสารกับคอมพิวเตอร์อย่างง่ายโดยมีไวยากรณ์ที่เข้าใจง่ายและอ่านได้เป็นอย่างดี ด้วยเหตุนี้จึงส่งผลให้เกิดชุมชนที่ใหญ่ขึ้น ทำให้มี ไลบรารี ที่เป็นประโยชน์ในงานสาขาต่างๆมากมาย ให้เลือกใช้



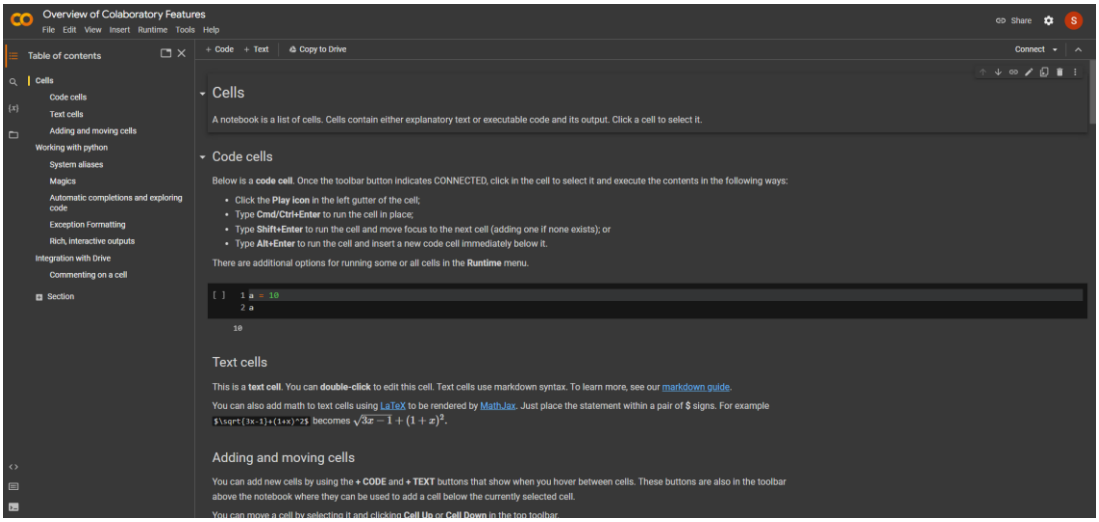
ภาพที่ 3.11 ตัวอย่างการใช้งาน Python ผ่าน Command Prompt

3.5.2 Google Colaboratory

Google Colaboratory เป็นแพลตฟอร์มในรูปแบบของสมุดบันทึกออนไลน์ที่ใช้ในการเขียนและเรียกใช้โค้ด Python ผ่านเว็บเบราว์เซอร์ แพลตฟอร์มนี้ให้บริการโดย Google และเป็นแหล่งที่มาของเครื่องมือและทรัพยากรที่สามารถใช้ในการเรียนรู้และพัฒนาโค้ด Python ได้อย่างง่ายดาย นอกจาก Python แล้วยังรองรับภาษาโปรแกรมอื่น ๆ เช่น R, Julia และอื่น ๆ อีกด้วย

Google Colaboratory มีข้อดีที่สำคัญคือไม่จำเป็นต้องติดตั้ง Python หรือเครื่องมือพิเศษอื่น ๆ ลงในเครื่อง ทุกอย่างถูกเก็บรักษาและทำงานบนเซิร์ฟเวอร์ของ Google ทำให้สามารถเขียนและเรียกใช้โค้ด Python ได้ทุกที่ที่มีการเชื่อมต่ออินเตอร์เน็ตเท่านั้น นอกจากนี้ยังสามารถแชร์สมุดบันทึกให้ผู้อื่นดูและแก้ไขร่วมกันได้อีกด้วย

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

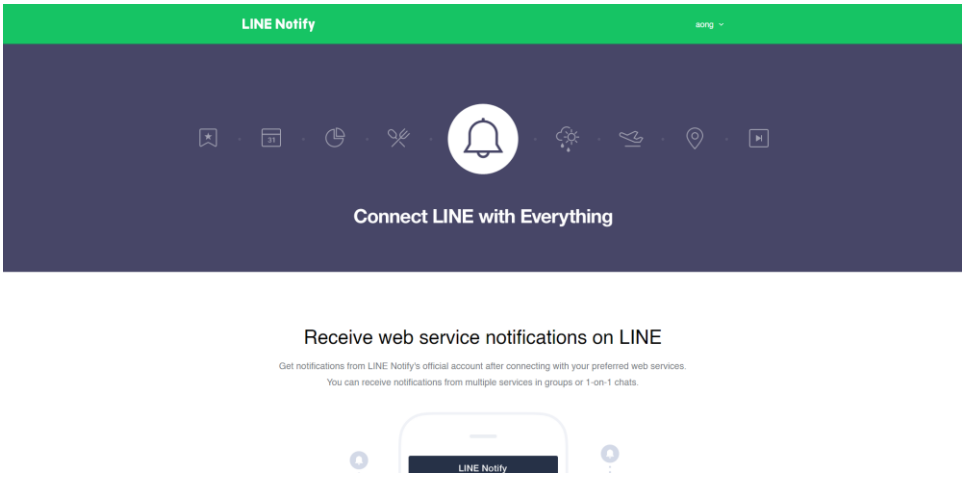


ภาพที่ 3.12 หน้าตาการใช้งาน Google Colaboratory

3.5.3 ไลน์โนติฟาย (Line Notify)

Line Notify เป็นบริการที่ให้คุณส่งการแจ้งเตือนหรือข้อความไปยังแชทบอทของ Line ผ่านการใช้งาน API ของ Line Notify ซึ่งเป็นบริการเสริมของแอปพลิเคชัน Line โดยสามารถนำมาใช้เพื่อแจ้งเตือนผู้ใช้เกี่ยวกับข้อมูลหรือเหตุการณ์ต่าง ๆ ที่สำคัญได้ เช่น การแจ้งเตือนเมื่อมีการทำรายการทางการเงิน, การแจ้งเตือนข่าวสารหรือข้อมูลใหม่ หรือแม้กระทั่งการแจ้งเตือนสถานะของระบบ เป็นต้น

การใช้งาน Line Notify ต้องทำการลงทะเบียนและสร้าง Token ในเว็บไซต์ของ Line Notify ก่อน เมื่อได้รับ Token จากการลงทะเบียน ถึงจะสามารถใช้งาน API ของ Line Notify เพื่อส่งข้อความไปยังแชทบอทของ Line ผ่านการเรียกใช้ HTTP request



ภาพที่ 3.13 หน้าตาเว็บไซต์ของไลน์ โนติฟาย (Line Notify)

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บทที่ 4
ผลการวิจัย

สารนิพนธ์นี้มีวัตถุประสงค์เพื่อนำการพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) มาประยุกต์ใช้ในการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ล่วงหน้าสำหรับธุรกิจ e-commerce โดยมีผลลัพธ์เป็นการสร้างระบบแจ้งเตือนผ่านทางแอปพลิเคชั่น เพื่อที่แจ้งเตือนผู้ดูแลระบบให้เตรียมการให้พร้อมรับมือกับปริมาณการเข้าชมเว็บไซต์ที่มากขึ้น โดยมีรายละเอียดของผลการศึกษาดังต่อไปนี้

4.1 ผลการวัดประสิทธิภาพความถูกต้องของโมเดล

งานวิจัยนี้ใช้ข้อมูลรายวันของ Google Merchandise Store ตั้งแต่วันที่ 1 มกราคม 2559 จนถึงวันที่ 6 พฤษภาคม 2566 โดยนำเข้าตัวแปรต่างๆ เช่น วันที่ ปริมาณการเข้าชมเว็บไซต์ และระยะเวลาเฉลี่ยในการเข้าชมเว็บไซต์แต่ละครั้ง รวมทั้งสร้างตัวแปรเพิ่มขึ้นมา เช่น is_weekend, week_no และ e-commerce holidays และทำการสร้างโมเดล Prophet (Multivariate) ซึ่งทำการแบ่งข้อมูลวัดผลด้วยวิธีการ Cross Validation มาทำการวัดผลโดยใช้ Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) และ Root Mean Square Error (RMSE) มาทำการวัดผลเทียบกันกับโมเดลต่างๆ โดยพิจารณาจากข้อมูลที่มีก่อนหน้า ได้ผลดังตารางที่ 4.1 ถึงตารางที่ 4.3

ตารางที่ 4.1 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล ARIMA (Baseline model)

Horizon	RMSE	MAE	MAPE
1 day	63.168	63.168	0.032
2 days	575.544	575.149	0.257
3 days	809.479	657.036	0.290
4 days	816.197	719.306	0.309
5 days	758.835	666.633	0.259
6 days	432.060	419.788	0.158
7 days	262.949	208.036	0.090
8 days	524.485	433.950	0.204
9 days	790.917	701.154	0.335
10 days	967.215	838.617	0.397

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ตารางที่ 4.1 (ต่อ)

Horizon	RMSE	MAE	MAPE
11 days	973.001	842.239	0.388
12 days	826.681	737.679	0.317
13 days	560.770	506.388	0.202
14 days	281.397	230.790	0.095

ตารางที่ 4.2 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล Prophet แบบ Univariate (Baseline Model)

Horizon	RMSE	MAE	MAPE
2 days	322.849	267.576	0.151
3 days	500.680	412.833	0.238
4 days	471.871	385.048	0.190
5 days	384.408	332.434	0.152
6 days	411.181	356.904	0.180
7 days	444.612	403.361	0.231
8 days	378.911	328.384	0.185
9 days	541.692	395.526	0.166
10 days	564.653	445.163	0.190
11 days	472.972	387.102	0.175
12 days	337.286	280.647	0.126
13 days	410.776	334.005	0.135
14 days	411.785	324.345	0.130

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ตารางที่ 4.3 ตารางแสดงผลการทดสอบประสิทธิภาพของโมเดล Prophet แบบ Multivariate

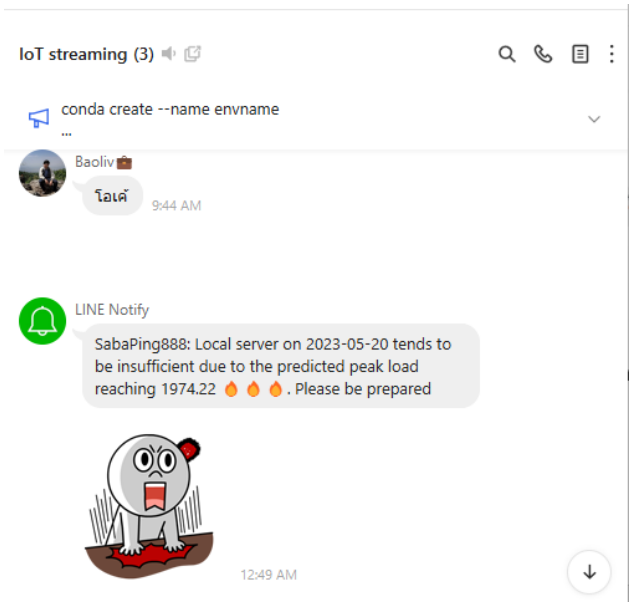
Horizon	RMSE	MAE	MAPE
2 days	421.397	331.615	0.180
3 days	371.928	295.820	0.172
4 days	290.346	235.363	0.128
5 days	326.201	291.792	0.147
6 days	448.867	365.931	0.191
7 days	360.739	300.145	0.164
8 days	277.244	245.521	0.134
9 days	396.295	314.753	0.142
10 days	355.742	279.775	0.130
11 days	303.217	233.536	0.101
12 days	270.910	219.458	0.092
13 days	304.040	245.407	0.100
14 days	279.129	224.805	0.090

จากตารางผลการทดสอบของโมเดล Prophet (Multivariate) มีประสิทธิภาพในการพยากรณ์ปริมาณการเข้าใช้เว็บไซต์ใน 14 วันล่วงหน้า ได้ดีที่สุด โดยที่ค่า RMSE เท่ากับ 279.129 ค่า MAE เท่ากับ 224.805 และ MAPE เท่ากับ 9% และภาพรวมของค่า Error ต่างๆมีความนิ่งกว่าโมเดลอื่นๆ

4.2 ผลการดำเนินการในส่วนของการแจ้งเตือนผ่านแอปพลิเคชั่น

ผลลัพธ์ของการพยากรณ์จะเป็นค่าคาดการณ์ปริมาณผู้เข้าชมเว็บ จะถูกนำมาวัดกับเกณฑ์ที่ตั้งไว้ ถ้าผลลัพธ์มีค่ามากกว่าเกณฑ์ ระบบจะทำการแจ้งเตือนผ่านแอปพลิเคชั่นไลน์ (Line) ด้วยการใช้ API ของ Line Notify เพื่อที่แจ้งเตือนผู้ปฏิบัติงานให้เตรียมการให้พร้อมรับมือกับปริมาณการเข้าชมเว็บไซต์ที่มากขึ้น และไม่ส่งการแจ้งเตือนในกรณีที่ผลลัพธ์ไม่เกินเกณฑ์ที่ตั้งไว้ เพื่อจะได้ไม่ก่อความรำคาญให้กับผู้ปฏิบัติงาน

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์



ภาพที่ 4.1 ตัวอย่างผลลัพธ์การแจ้งเตือนผ่านทางแอปพลิเคชั่นไลน์ (Line)

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บทที่ 5
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สารนิพนธ์นี้มีวัตถุประสงค์เพื่อนำการพยากรณ์แบบอนุกรมเวลา (Time Series Forecasting) มาประยุกต์ใช้ในการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ล่วงหน้าสำหรับธุรกิจ e-commerce โดยมีผลลัพธ์เป็นการสร้างระบบแจ้งเตือนผ่านทางแอปพลิเคชั่น เพื่อที่แจ้งเตือนผู้ดูแลระบบให้เตรียมการให้พร้อมรับมือกับปริมาณการเข้าชมเว็บไซต์ที่มากขึ้น โดยสามารถสรุปผลการวิจัยได้ดังนี้

5.1 สรุปผลการศึกษา

การใช้วิธีการวิเคราะห์อนุกรมเวลาเพื่อพยากรณ์ปริมาณการเข้าชมเว็บไซต์เป็นหนึ่งในวิธีที่สามารถประยุกต์ใช้ในทางปฏิบัติเพื่อจัดเตรียมทรัพยากรให้เพียงพอสำหรับการรองรับปริมาณการเข้าชมเว็บไซต์ที่มีการเพิ่มขึ้น โดยการทำนายปริมาณการเข้าชมเว็บไซต์จะช่วยให้ผู้ดูแลระบบสามารถปรับแผนการขยายทรัพยากรและออกแบบระบบให้เหมาะสมกับความต้องการของผู้ใช้งาน ซึ่งมีประโยชน์ในการป้องกันปัญหาเว็บไซต์ล่ม และเพิ่มประสบการณ์การใช้บริการที่ดีขึ้นสำหรับผู้ใช้งาน รวมถึงลดความเสี่ยงในการสูญเสียผู้ใช้งานที่อาจเสียใจและไปใช้บริการกับคู่แข่งได้

จากการทดลองพบว่าในการประเมินประสิทธิภาพของโมเดลที่ใช้ในการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ พบว่า Prophet แบบ Multivariate มีประสิทธิภาพที่ดีกว่า ARIMA และ Prophet แบบ Univariate การใช้วิธี ARIMA ในการพยากรณ์อนุกรมเวลาเป็นวิธีที่ได้รับความนิยมในอดีต แต่ในการทดลองนี้พบว่า Prophet แบบ Multivariate มีประสิทธิภาพที่ดีกว่า ARIMA โดยสามารถให้ผลการพยากรณ์ที่แม่นยำขึ้นและเหมาะสมกับข้อมูลการเข้าชมเว็บไซต์มากขึ้น และในการทดลองนี้ ยังพบว่า Prophet แบบ Multivariate ให้ผลการพยากรณ์ที่ดีกว่า Prophet แบบ Univariate ซึ่งหมายความว่าการนำเข้าข้อมูลเพิ่มเติมเช่นตัวแปรอื่นๆ ที่สอดคล้องกับปริมาณการเข้าชมเว็บไซต์ สามารถเสริมประสิทธิภาพในการทำนายได้

นำผลลัพธ์ที่ได้จากโมเดล ไปพัฒนาเป็นระบบการแจ้งเตือนอัตโนมัติ ในกรณีที่ผลการพยากรณ์ล่วงหน้า 14 วัน สูงกว่าเกณฑ์ที่กำหนด เพื่อเป็นตัวช่วยในการตัดสินใจให้แก่ผู้ดูแลระบบ

5.2 ข้อเสนอแนะ

5.2.1 ชุดข้อมูลที่นำมาใช้งานเป็นข้อมูลของเว็บไซต์ Google Merchandise Store ซึ่งทาง Google เปิดให้บริการซื้อขายของที่ระลึก รวมทั้งยังใช้งานเป็นเว็บไซต์ตัวอย่างในการศึกษาวิธีใช้งาน Google Analytics ข้อมูลจึงอาจจะมีความคลาดเคลื่อนบ้างในบางวัน

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

5.2.2 พิจารณาหาตัวแปรหรือปัจจัยอื่นๆที่ส่งผลต่อการพยากรณ์ นอกเหนือจากวันหยุดและ e-commerce holidays ได้แก่ กิจกรรมส่งเสริมการขายและการตลาดของแพลตฟอร์ม เทศกาลอื่นๆที่อาจจะส่งผลต่อการเข้าชมเว็บไซต์ เช่น Pride Month เป็นต้น

5.2.3 สามารถพัฒนาประสิทธิภาพของโมเดลได้ด้วยการทดลองปรับค่าพารามิเตอร์ การสร้างตัวแปรใหม่ๆ หรือทดลองกับโมเดลประเภทอื่น เพื่อให้ได้ผลการพยากรณ์ปริมาณการเข้าชมเว็บไซต์ที่ความแม่นยำสูงขึ้น

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บรรณานุกรม

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

บรรณานุกรม

[1] Nuthdanai Wangpratham, “*การพยากรข้อมูลอนุกรมเวลาด้วยเทคนิค ARIMA ด้วย Python,*” **[**ออนไลน์**]**. https://medium.com/qunt-i-love-u/การพยากรณ์ข้อมูลอนุกรมเวลาด้วยเทคนิค-arima-ด้วย-python-44809eb8e990. (เข้าถึงเมื่อ: 25 มิถุนายน 2566)

[2] G. Rafferty, *Forecasting Time Series Data with Prophet - Build, improve, and optimize time series forecasting model using Meta’s advanced forecasting tool.* Packt Publishing, 2023.

[3] N. Petluri and E. Al-Masri, "Web Traffic Prediction of Wikipedia Pages," *2018 IEEE International Conference on Big Data (Big Data)*, Seattle, WA, USA, 2018, pp. 5427-5429, doi: 10.1109/BigData.2018.8622207.

[4] S. Shukla, “Various Techniques to Detect and Isolate Time Series Components Using Python,” [Online]. https://www.analyticsvidhya.com/blog/2023/02/various-techniques-to-detect-and-isolate-time-series-components-using-python/ (Accessed: Jun. 25, 2023).

[5] T. Shelatkar, S. Tondale, S. Yadav and S. Ahir, “Web traffic time series forecasting using ARIMA and LSTM RNN,” ITM Web of Conferences (Vol. 32, p. 03017). EDP Sciences.

[6] D. Sikka and C. N. S. Vinoth Kumar, "Website Traffic Time Series Forecasting Using Regression Machine Learning," *2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)*, Bhopal, India, 2023, pp. 246-250, doi: 10.1109/CSNT57126.2023.10134631.

[7] S. J. Taylor and B. Letham, “Forecasting at scale,” The American Statistician, 72(1), pp. 37-45, Sep. 2017, doi: https://doi.org/10.7287/peerj.preprints.3190v2.

[8] E. Zunic, K. Korjenic, S. Delalic, and Z. Subara, “Comparison Analysis of Facebook's Prophet, Amazon's DeepAR+ and CNN-QR Algorithms for Successful Real-World Sales Forecasting,” *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 13, no. 2, Apr. 2021, doi: 10.5121/ijcsit.2021.13205.

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ภาคผนวก

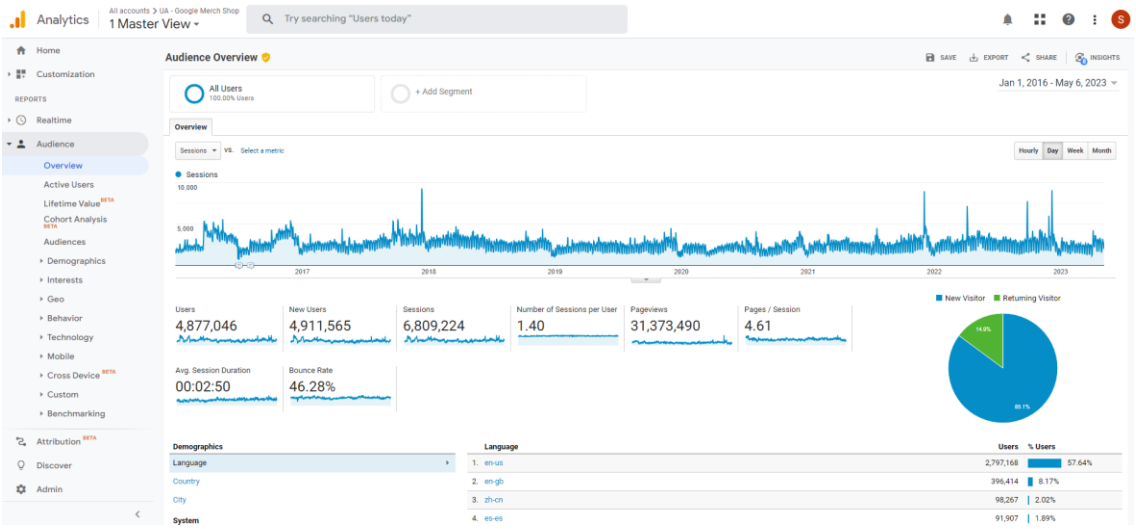
DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ภาคผนวก ก

การดึงข้อมูลจาก Google Analytics เพื่อนำมาใช้งาน

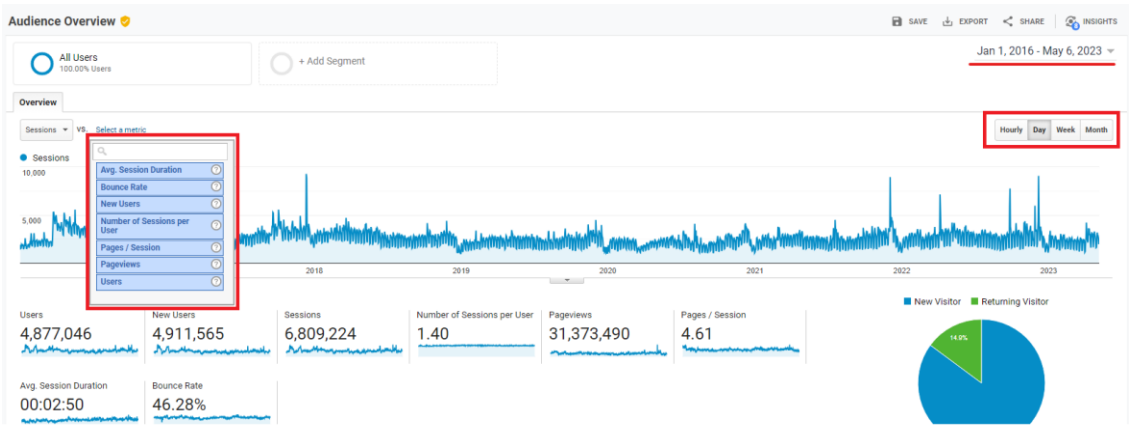
DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

เข้ามาที่หน้าเว็บไซต์ของ Google Analytics จะพบกับแดชบอร์ด (Dashboard) ที่แสดงข้อมูลของเว็บไซต์ จากภาพที่ 1 จะเป็นข้อมูลของ Google Merchandise Store จากประเทศสหรัฐอเมริกา ซึ่งในหน้าแดชบอร์ดนี้จะแสดงค่าวัดผลต่างๆที่เกี่ยวข้องกับ Web Traffic ที่จำเป็นในการวิเคราะห์ข้อมูล



ภาพที่ 1 Dashboard จากเว็บไซต์ Google Analytics

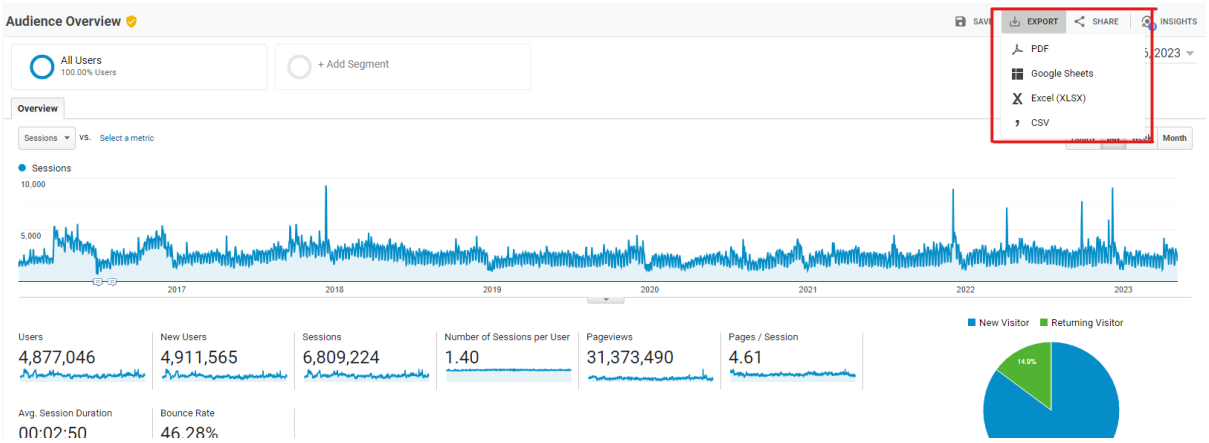
จากนั้นสามารถเลือกช่วงเวลาเริ่มต้นและเวลาสิ้นสุด ค่าวัดผลที่เราสนใจ และความละเอียดของเวลาในหน่วยต่างๆได้แก่รายชั่วโมง รายวัน รายสัปดาห์ และรายเดือน ดังภาพที่ 2



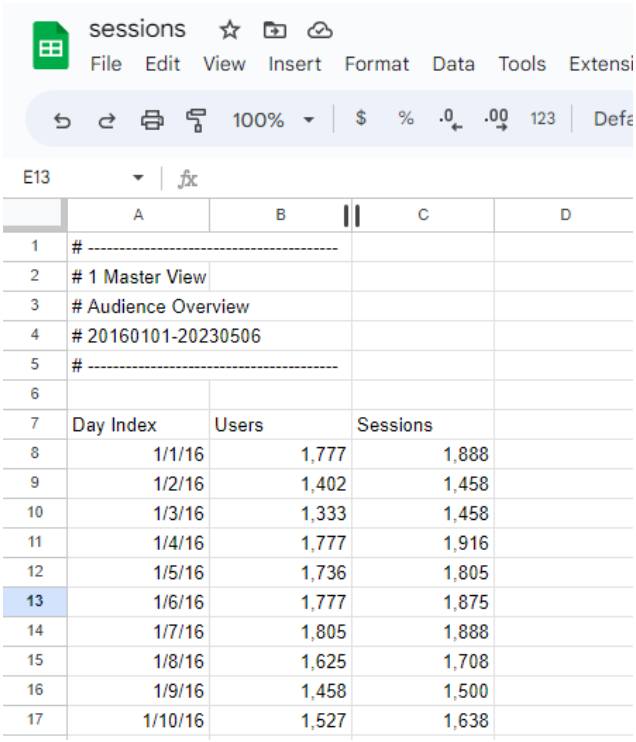
ภาพที่ 2 ตัวอย่างการเลือกค่าวัดผล ช่วงเวลาและความละเอียดของเวลา

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

เมื่อเลือกช่วงเวลา ความละเอียดของช่วงเวลาและค่าวัดผลแล้ว สามารถนำข้อมูลออกมาด้วยการ Export และเลือก Format เพื่อให้เหมาะสมกับงานที่ใช้ ซึ่งสามาเลือกได้ 4 ประเภทดังนี้ PDF, Google Sheet, Excel (xlsx) และ CSV ดังภาพที่ 3 และ 4



ภาพที่ 3 การ Export ข้อมูลออกมาใน Format ต่างๆเพื่อนำข้อมูลไปใช้



ภาพที่ 4 ตัวอย่างการ Export ข้อมูลออกมาเป็น Google Sheet

DPU DHURAKIJ PUNDIT UNIVERSITY มหาวิทยาลัยธุรกิจบัณฑิตย์

ประวัติผู้เขียน

ชื่อ – นามสกุล สิทธิ โง้ววัฒนา

ประวัติการศึกษา

พ.ศ. 2554 - วิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมเครื่องกล มหาวิทยาลัยเกษตรศาสตร์

ประสบการณ์ทำงาน (เรียงจากใหม่ไปหาเก่า)

พ.ศ. 2566 - Business Intelligence บริษัท ซ้อปปี้ (ประเทศไทย) จำกัด

พ.ศ. 2564 - Piping Engineer บริษัท ทีทีซีแอล จำกัด (มหาชน)