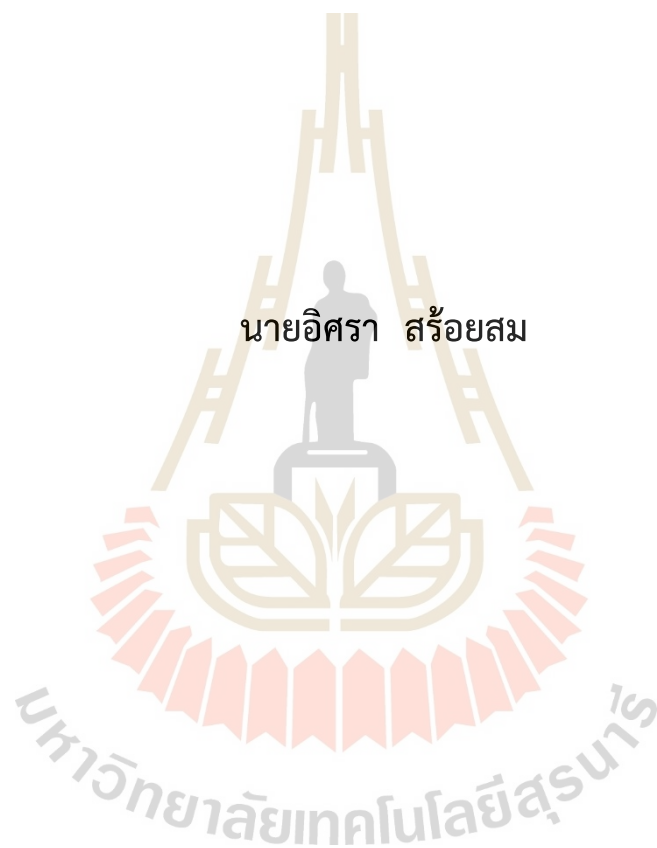


การวิเคราะห์และค้นหาพื้นที่ในประเทศไทยที่ให้พลังงานแสงอาทิตย์สูงตลอดปี
โดยการจัดกลุ่มแบบเคมีนส์



นายอิสรา สร้อยสม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมโทรคมนาคมและคอมพิวเตอร์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2565

THE ANALYSIS AND SEARCH FOR AREA IN THAILAND WITH HIGH
SOLAR ENERGY THROUGHOUT THE YEAR USING
K-MEANS CLUSTERING



A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Engineering in Telecommunication and Computer Engineering
Suranaree University of Technology
Academic Year 2022

การวิเคราะห์และค้นหาพื้นที่ในประเทศไทยที่ให้พลังงานแสงอาทิตย์สูงตลอดปี
โดยการจัดกลุ่มแบบเคมินส์

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นักวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์



(รศ. ดร.กิตติศักดิ์ เกิดประสพ)
ประธานกรรมการ



(รศ. ดร.นิตยา เกิดประสพ)
กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)



(ดร.กีระชาติ สุขสุทธิ)
กรรมการ



(ดร.รติพร จันทร์กลั่น)
กรรมการ



(รศ. ดร.ฉัตรชัย โชติษฐียงกูร)
รองอธิการบดีฝ่ายวิชาการและประกันคุณภาพ



(รศ. ดร.พรศิริ จงกล)
คณบดีสำนักวิชาวิศวกรรมศาสตร์

อิศรา สร้อยสม : การวิเคราะห์และค้นหาพื้นที่ในประเทศไทยที่ให้พลังงานแสงอาทิตย์สูงตลอดปีโดยการจัดกลุ่มแบบเคมีนส์ (THE ANALYSIS AND SEARCH FOR AREA IN THAILAND WITH HIGH SOLAR ENERGY THROUGHOUT THE YEAR USING K-MEANS CLUSTERING) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร. นิตยา เกิดประสพ, 50 หน้า

คำสำคัญ : การจัดกลุ่มแบบเคมีนส์/พลังงานแสงอาทิตย์/การวิเคราะห์ภาพเงา

ประเทศไทยเป็นประเทศเขตร้อนที่มีการให้ค่าพลังงานแสงอาทิตย์ที่เฉลี่ยสูงทั้งประเทศ ทำให้ประเทศไทยมีพื้นที่เหมาะกับการนำพลังงานแสงอาทิตย์มาใช้งานเป็นอย่างมาก อย่างไรก็ตามสำหรับทุกพื้นที่ในประเทศไทยไม่ได้มีการให้พลังงานแสงอาทิตย์สูงทั้งปี งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อศึกษาพื้นที่ที่ให้พลังงานแสงอาทิตย์สูงในหนึ่งปี โดยมีการใช้งานการวิเคราะห์ภาพเงา (Silhouette Analysis) ในการกำหนดตัวแปร k เริ่มต้น เพื่อใช้ในการทำการแบ่งกลุ่มข้อมูลแบบเคมีนส์ จากนั้นแบ่งข้อมูลออกเป็นชุดของข้อมูลตามจำนวนเดือน และทำการอินเตอร์เซคข้อมูลที่ได้จากการแบ่งกลุ่มของชุดของข้อมูลเพื่อหาข้อสรุปของพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปีของประเทศไทย



มหาวิทยาลัยเทคโนโลยีสุรนารี

สาขาวิชา วิศวกรรมคอมพิวเตอร์
ปีการศึกษา 2565

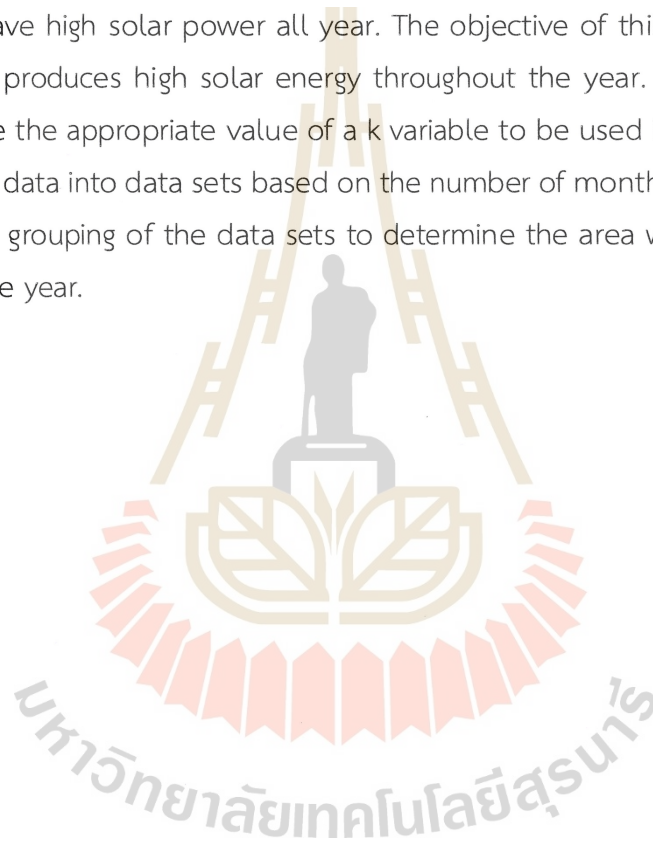
ลายมือชื่อนักศึกษา.....อิศรา
ลายมือชื่ออาจารย์ที่ปรึกษา.....ดร. นิตยา

ISARA SOISOM : THE ANALYSIS AND SEARCH FOR AREA IN THAILAND WITH HIGH SOLAR ENERGY THROUGHOUT THE YEAR USING K-MEANS CLUSTERING.

THESIS ADVISOR : ASSOC. PROF. NITAYA KERDPRASOP, Ph.D., 50 PP.

Keywords : K-means clustering/Solar Energy/Silhouette Analysis

Thailand is a tropical country with high average solar power across the country. This makes Thailand a very suitable area for using solar energy. However, not all areas in Thailand have high solar power all year. The objective of this research is to study the area that produces high solar energy throughout the year. Silhouette analysis is used to define the appropriate value of a k variable to be used in k-means clustering. We divide the data into data sets based on the number of months. Then, intersect the data from the grouping of the data sets to determine the area with high solar energy throughout the year.



School of Computer Engineering
Academic Year 2022

Student's Signature Isara
Advisor's Signature Nitaya Kerdprasop

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จลุล่วงด้วยดี เนื่องจากได้รับการช่วยเหลือจากหลายฝ่ายอย่างดียิ่ง ทั้งทางด้านวิชาการและทางด้านการดำเนินงานวิจัย ผู้เขียนขอขอบพระคุณผู้มีพระคุณทุกท่าน ดังต่อไปนี้

อาจารย์ รศ. ดร.นิตยา เกิดประสพ และอาจารย์ รศ. ดร.กิตติศักดิ์ เกิดประสพ ที่คอยแนะนำ และให้คำปรึกษาในเรื่องต่าง ๆ ทั้งเรื่องการใช้ชีวิตในมหาวิทยาลัย, งานผู้ช่วยสอน ให้คำปรึกษาเรื่อง เอกสารสำคัญ และวิทยานิพนธ์สามารถสำเร็จได้ เพราะการสนับสนุนของท่านอาจารย์ทั้งสองท่าน

คุณปราณี กฐินใหม่ เลขาธิการสาขาวิชาวิศวกรรมคอมพิวเตอร์ที่ให้ความช่วยเหลือในการ ประสานงานระหว่างศึกษา ทั้งเอกสารแต่งตั้งผู้สอน กำหนดการส่งเอกสารสำคัญต่าง ๆ พร้อม คำแนะนำหลายอย่างที่ทำให้กระบวนการหลาย ๆ อย่างผ่านไปได้อย่างถูกต้อง และรวดเร็ว

ขอขอบคุณนักศึกษา ร่วมชั้นเรียนทั้งปริญญาโท และปริญญาเอกทุกคนที่ให้คำแนะนำในทุก ด้าน ทั้งเกี่ยวกับการศึกษาก็ดี หรือเรื่องในชีวิตประจำวันก็ดี

ขอขอบคุณ มหาวิทยาลัยเทคโนโลยีสุรนารี ที่เปิดโอกาสทางการศึกษา และการเข้าถึง บุคลากรที่มีคุณภาพ

และสุดท้ายนี้ขอขอบคุณ นายอุทิศ สร้อยสม และนางผาสุข สร้อยสม บิดามารดา ที่คอยให้ กำลังใจ และการสนับสนุนต่าง ๆ ทั้งเงินทุน และค่าเล่าเรียน

อิสรา สร้อยสม

มหาวิทยาลัยเทคโนโลยีสุรนารี

สารบัญ

หน้า

บทคัดย่อ (ภาษาไทย).....	ก
บทคัดย่อ (ภาษาอังกฤษ).....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่	
1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของงานวิจัย.....	1
1.2 วัตถุประสงค์.....	3
1.3 ขอบเขตและข้อจำกัดในงานวิจัย.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	3
2 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1 การแบ่งข้อมูลแบบเคมีนส์ (k-means clustering).....	4
2.1.1 ความเป็นมาของการแบ่งข้อมูล (clustering).....	4
2.1.2 การหาระยะห่างของการแบ่งข้อมูล.....	5
2.1.3 ทฤษฎีของการแบ่งข้อมูลแบบเคมีนส์.....	7
2.1.4 ข้อเสียของการแบ่งข้อมูลแบบเคมีนส์.....	8
2.2 การวิเคราะห์ภาพเงา (Silhouette Analysis).....	8
2.3 ทฤษฎีการอินเตอร์เซก (Intersection) และความเกี่ยวข้องกับงานวิจัย.....	10
2.4 ความเป็นมาของแผนภาพกล่อง (Box Plot).....	11
2.5 ความสำคัญของพลังงานแสงอาทิตย์.....	14
2.5.1 พลังงานแสงอาทิตย์.....	14
2.5.2 เทคโนโลยีพลังงานไฟฟ้าแสงอาทิตย์โซลาร์เซลล์.....	15
2.5.3 เทคโนโลยีพลังงานแสงอาทิตย์เพื่อผลิตน้ำร้อน.....	16

สารบัญ (ต่อ)

	หน้า
2.5.4 เทคโนโลยีพลังงานแสงอาทิตย์เพื่อการอบแห้ง.....	17
2.6 งานวิจัยที่เกี่ยวข้อง.....	18
3 ขั้นตอนการวิจัย.....	22
3.1 แผนภาพขั้นตอนการดำเนินงานวิจัย.....	22
3.2 แหล่งที่มาของข้อมูล และลักษณะของข้อมูล.....	23
3.3 การกำหนดค่าอินพุต.....	25
3.4 การหาตัวแปร k สำหรับการแบ่งกลุ่มแบบเคมีนส์ ด้วยการวิเคราะห์ภาพเงา.....	26
3.5 ขั้นตอนการแบ่งกลุ่มแบบเคมีนส์.....	27
3.6 การอินเตอร์เซคข้อมูลที่ได้จากการแบ่งกลุ่มแบบเคมีนส์.....	28
3.7 ขั้นตอนการวิเคราะห์ผล.....	29
3.7.1 การวิเคราะห์การแบ่งกลุ่มแบบเคมีนส์ในแต่ละชุดกลุ่มข้อมูล.....	29
3.7.2 การวิเคราะห์ผลจากการนำข้อมูลมาระบายสีบนรูปแผนที่จำลอง.....	29
3.7.3 การเปรียบเทียบผลลัพธ์ระหว่างงานวิจัย และการคลัสเตอร์แบบปกติ.....	30
4 ผลการวิจัยและอภิปรายผล.....	32
4.1 ผลลัพธ์การหาค่าตัวแปร k.....	32
4.2 ผลลัพธ์การแบ่งกลุ่มแบบเคมีนส์ในแต่ละชุดกลุ่มข้อมูล.....	33
4.2.1 ชุดกลุ่มข้อมูล มกราคม.....	33
4.2.2 ชุดกลุ่มข้อมูล กุมภาพันธ์.....	33
4.2.3 ชุดกลุ่มข้อมูล มีนาคม.....	34
4.2.4 ชุดกลุ่มข้อมูล เมษายน.....	35
4.2.5 ชุดกลุ่มข้อมูล พฤษภาคม.....	36
4.2.6 ชุดกลุ่มข้อมูล มิถุนายน.....	36
4.2.7 ชุดกลุ่มข้อมูล กรกฎาคม.....	37
4.2.8 ชุดกลุ่มข้อมูล สิงหาคม.....	38
4.2.9 ชุดกลุ่มข้อมูล กันยายน.....	38
4.2.10 ชุดกลุ่มข้อมูล ตุลาคม.....	39
4.2.11 ชุดกลุ่มข้อมูล พฤศจิกายน.....	39

สารบัญ (ต่อ)

	หน้า
4.2.12 ชุดกลุ่มข้อมูล ธันวาคม.....	40
4.3 วิเคราะห์ผลจากการนำข้อมูลมาระบายสีบนรูปแผนที่จำลอง.....	41
4.4 วิเคราะห์ผลจากการนำข้อมูลมาเปรียบเทียบ.....	44
4.5 สรุปการอภิปรายผล.....	46
5 สรุปและข้อเสนอแนะ.....	47
5.1 สรุปผลการวิจัย.....	47
5.2 ข้อเสนอแนะ.....	48
เอกสารอ้างอิง.....	49
ประวัติผู้เขียน.....	50



สารบัญตาราง

ตารางที่	หน้า
2.1	ตัวอย่างตำแหน่งแต่ละแกนของแต่ละจุด สำหรับการคำนวณหาระยะห่าง.....6
2.2	สรุปความสัมพันธ์ของงานวิจัยที่เกี่ยวข้องในรูปแบบตารางแบบกะทัดรัด.....21
3.1	ตัวอย่างลักษณะของข้อมูลความเข้มรังสีแสงอาทิตย์ ส่วนพื้นที่สำรวจ.....23
3.2	ตัวอย่างลักษณะของข้อมูลความเข้มรังสีแสงอาทิตย์ ส่วนค่าพลังงานเฉลี่ย.....24
3.3	ชื่อ, คำอธิบาย และหน่วยของแต่ละฟิลล์ ในฐานะข้อมูลความเข้มรังสีแสงอาทิตย์.....24
3.4	ตารางแสดงการกำหนดค่าอินพุตข้อมูล ในฐานะข้อมูลความเข้มรังสีแสงอาทิตย์.....25
3.5	ตัวอย่างชุดของข้อมูลหลังจากแบ่งกลุ่ม ตัวอย่างกลุ่ม Average.....26
3.6	ตัวอย่างชุดของข้อมูลหลังจากแบ่งกลุ่ม ตัวอย่างกลุ่ม January.....26
3.7	ตัวอย่างลักษณะของข้อมูลหลังจากการแบ่งกลุ่มแบบเคมีนส์ ของกลุ่ม January.....27
3.8	ตัวอย่างของการอินเตอร์เซคข้อมูลค่าพลังงานแสงอาทิตย์ ส่วนพื้นที่สำรวจ.....28
3.9	ตัวอย่างของการอินเตอร์เซคข้อมูลค่าพลังงานแสงอาทิตย์ ส่วนค่าพลังงานเฉลี่ย.....28
4.1	ค่าเฉลี่ยของ silhouette score ในแต่ละตัวแปร.....32
4.2	แสดงผลลัพธ์ของงานวิจัย โดยมีผลลัพธ์พื้นที่มากกว่า 10 เปอร์เซ็นต์.....42
4.3	แสดงผลลัพธ์ของงานวิจัย โดยมีผลลัพธ์พื้นที่น้อยกว่า 10 เปอร์เซ็นต์.....44
4.4	แสดงผลลัพธ์ของงานวิจัยในการหาค่าพื้นที่ที่มีพลังงานแสงอาทิตย์สูงสุด.....45

สารบัญรูป

รูปที่		หน้า
2.1	ตัวอย่างการแบ่งข้อมูล.....	4
2.2	ตัวอย่าง Silhouette Scores ของค่า k ที่ 2 – 3.....	9
2.3	ตัวอย่าง Silhouette Scores ของค่า k ที่ 4 – 5.....	10
2.4	การอินเตอร์เซกระหว่างเซต A และ B กรณีมีสมาชิกร่วมกัน.....	11
2.5	การอินเตอร์เซกระหว่างเซต A และ B กรณีไม่มีสมาชิกร่วมกัน.....	11
2.6	ลักษณะของแผนภาพกล่อง.....	12
2.7	การกระจายตัวแบบปกติ (Normal Distribution).....	13
2.8	การกระจายตัวแบบการเอียงเอนไปทางฝั่งข้างมาก (Positive Distribution).....	13
2.9	การกระจายตัวแบบเอียงเอนไปทางฝั่งข้างน้อย (Negative Distribution).....	13
2.10	ตำแหน่งของค่าผิดปกติ.....	14
2.11	ขนาดของแผนภาพกล่องในแต่ละช่วงที่ควรจะเป็น.....	14
3.1	แผนภาพขั้นตอนการดำเนินงานวิจัย.....	22
3.2	ตัวอย่างผลลัพธ์ที่ได้หลังจากการแบ่งกลุ่มแบบเคมีนส์ ของชุดข้อมูลกลุ่มใด ๆ.....	29
3.3	แผนที่จำลองที่ใช้สำหรับงานวิจัยนี้.....	30
4.1	แผนภาพกล่องของค่าเฉลี่ยพลังงานแสงอาทิตย์ในค่าตัวแปร k ที่เหมาะสม.....	32
4.2	แผนภาพกล่องของชุดกลุ่มข้อมูล มกราคม.....	33
4.3	แผนภาพกล่องของชุดกลุ่มข้อมูล กุมภาพันธ์.....	34
4.4	แผนภาพกล่องของชุดกลุ่มข้อมูล มีนาคม.....	34
4.5	แผนภาพกล่องของชุดกลุ่มข้อมูล เมษายน.....	35
4.6	แผนภาพกล่องของชุดกลุ่มข้อมูล พฤษภาคม.....	35
4.7	แผนภาพกล่องของชุดกลุ่มข้อมูล มิถุนายน.....	36
4.8	แผนภาพกล่องของชุดกลุ่มข้อมูล กรกฎาคม.....	37
4.9	แผนภาพกล่องของชุดกลุ่มข้อมูล สิงหาคม.....	37
4.10	แผนภาพกล่องของชุดกลุ่มข้อมูล กันยายน.....	38
4.11	แผนภาพกล่องของชุดกลุ่มข้อมูล ตุลาคม.....	39

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.12	แผนภาพกล่องของชุดกลุ่มข้อมูล พศจิกายน.....40
4.13	แผนภาพกล่องของชุดกลุ่มข้อมูล ธันวาคม.....40
4.14	แผนที่จำลองประเทศไทยที่แสดงผลพีธที่ได้จากงานวิจัย.....41
4.15	แผนที่จำลองประเทศไทยที่แสดงผลพีธที่ได้จากงานวิจัย เมื่อลบพื้นที่จังหวัดที่มีพื้นที่ จากผลลัพธ์น้อย (เทา).....43



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของงานวิจัย

ประเทศไทยเป็นประเทศที่มีค่าเฉลี่ยปริมาณพลังงานแสงอาทิตย์สูงตลอดปี การสร้างพลังงานทดแทนอย่างพลังงานแสงอาทิตย์ย่อมเป็นสิ่งที่ประเทศไทยสนใจเป็นอย่างมาก แต่สำหรับในประเทศไทยที่ทุกพื้นที่ไม่สามารถให้ค่าพลังงานแสงอาทิตย์ที่ใกล้เคียงกันได้ตลอดปี โดยในแต่ละเดือนในหนึ่งปีอาจให้ค่าพลังงานแสงอาทิตย์ที่ต่างกันมาก ส่งผลทำให้กำลังในการผลิตพลังงานจากแสงอาทิตย์ไม่ได้ตามเป้าหมาย ทำให้ไม่เหมาะสมแก่การลงทุนในพื้นที่ส่วนนั้น ๆ และเนื่องจากแต่ละช่วงเวลาในหนึ่งปีไม่เป็นค่าพลังงานแสงอาทิตย์ที่สูงไปตลอดปี ซึ่งมีผลทำให้ไม่สามารถแบ่งได้ว่าพื้นที่ใดบ้างที่จะเป็นกลุ่มพื้นที่ที่สามารถให้ค่าพลังงานแสงอาทิตย์สูง เพื่อเป็นประโยชน์สำหรับความคุ้มค่าต่อการลงทุนเกี่ยวกับการนำพลังงานแสงอาทิตย์มาใช้ในพื้นที่ที่ต้องการโดยให้เกิดประสิทธิภาพที่ดีที่สุด แม้ว่าเป็นการลงทุนสำหรับครัวเรือนก็ตาม โดยส่งผลให้ประเทศไทยเป็นประเทศที่เหมาะสมแก่การนำพลังงานแสงอาทิตย์มาใช้ให้เกิดประโยชน์ได้สูงที่สุดต่อไป

ในงานวิจัยนี้ต้องการที่จะทราบว่าพื้นที่ที่สามารถให้พลังงานแสงอาทิตย์ที่สูงตลอดหนึ่งปีมีพื้นที่ใดบ้าง โดยใช้ข้อมูลข้อมูลความเข้มรังสีอาทิตย์ระดับตำบลสำหรับประเทศไทย จากปี พ.ศ. 2560 (นัยนา บุญนาค, 2563) มาทำการวิเคราะห์กับการแบ่งกลุ่มข้อมูลแบบเคมีนส์ (k-means Clustering) ซึ่งเป็นวิธีของการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งเป็นส่วนหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) โดยมีการแบ่งกลุ่มของข้อมูลออกเป็น 12 ช่วงตามจำนวนของเดือนในหนึ่งปี โดยก่อนที่ทำการใช้การแบ่งกลุ่มข้อมูลแบบเคมีนส์จำเป็นต้องมีการกำหนดตัวแปร k ซึ่งเป็นค่าที่ใช้ในการระบุจำนวนกลุ่มสำหรับการทำการแบ่งกลุ่มข้อมูลแบบเคมีนส์ แต่สำหรับข้อมูลใด ๆ ตัวแปร k ที่เหมาะสมกับการแบ่งกลุ่มของแต่ละข้อมูลย่อมมีความแตกต่างกัน ดังนั้นในงานวิจัยนี้จึงนำเสนอการวิเคราะห์ภาพเงา (Silhouette Analysis) ซึ่งเป็นกระบวนการในการหาตัวแปร k ที่เหมาะสมที่สุดสำหรับการแบ่งกลุ่มข้อมูลเพื่อนำไปสู่ผลลัพธ์ที่เหมาะสมที่สุด โดยนำค่าเฉลี่ยค่าพลังงานแสงอาทิตย์ในปี พ.ศ.2560 ที่เป็นตัวแทนของทั้ง 12 ช่วงข้อมูลมาเข้ากระบวนการวิเคราะห์ภาพเงาเพื่อใช้ตัวแปร k นี้ เป็นตัวแปรสำหรับทุกช่วงของข้อมูล เพื่อให้เป็นมาตรฐานเดียวกันในทุกช่วงข้อมูล จากนั้นจึงนำ 12 ช่วงข้อมูลมาแบ่งกลุ่มข้อมูล ทำให้แต่ละช่วงจะได้จำนวนกลุ่มเท่ากับตัวแปร k ที่นำมาใช้งาน และเพื่อที่จะทราบว่าในแต่ละหนึ่งช่วงข้อมูลของการแบ่งกลุ่มมีค่าพลังงานเฉลี่ยกลุ่มใดในช่วงนั้นมีค่าที่ดีที่สุด แล้วจึงจะนำข้อมูลจากกลุ่มในช่วงนั้นมาทำการอินเตอร์เซค (Intersection) ข้อมูลที่เหมือนกันของทั้ง 12 ช่วงจากกลุ่มของแต่ละช่วงข้อมูลที่มีค่าเฉลี่ยพลังงานแสงอาทิตย์สูงสุด ทำให้ได้ผลลัพธ์ออกมาเป็นค่าข้อมูลพื้นที่ที่มีค่าพลังงานแสงอาทิตย์

สูงในหนึ่งปี โดยโปรแกรม Orange Data mining (Demšar, et al., 2013) ซึ่งเป็นซอฟต์แวร์สำหรับการวิจัยการเรียนรู้ของเครื่อง และการทำเหมืองข้อมูล (Data mining)

งานวิจัยนี้จะแสดงให้เห็นถึงการนำการแบ่งกลุ่มข้อมูลแบบเคมีนส์ที่อยู่ในหมวดหมู่ของการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) โดยที่การเรียนรู้แบบไม่มีผู้สอนไม่จำเป็นต้องมีการกำหนดเป้าหมาย (Target) ของข้อมูล หมายความว่าจำเป็นต้องใช้การวิเคราะห์ และพิจารณาผลลัพธ์มากกว่าการเรียนรู้แบบมีผู้สอน (Supervised Learning) ที่มีการกำหนดเป้าหมายของข้อมูล ดังนั้นเพื่อให้การอธิบายผลลัพธ์มีความเหมาะสม จึงมีการใช้แผนภาพกล่อง (Box plot) ที่สามารถแสดงผลจากการแบ่งกลุ่มข้อมูลแบบเคมีนส์ สำหรับการวิเคราะห์ และช่วยในการหาค่ากลุ่มที่ต้องการเพื่อนำมาใช้สำหรับการอินเทอร์เน็ต จากนั้นนำผลที่ได้ที่เป็นรูปแบบตารางที่ผ่านอินเทอร์เน็ตมาลงสีในแผนที่จำลองประเทศไทยเพื่อวิเคราะห์ผลจากรูปภาพ สุดท้ายนี้การวิเคราะห์ผลทั้งหมดนี้สามารถให้บทสรุปของการหาค่าพลังงานแสงอาทิตย์ที่สูงที่สุดในหนึ่งปี โดยสามารถแสดงให้เห็นได้ชัดเจนและเหมาะสมที่สุดที่งานวิจัยนี้

การแบ่งกลุ่มข้อมูลแบบเคมีนส์นับเป็นขั้นตอนที่สำคัญสำหรับงานวิจัยนี้ จากงานวิจัยที่ใช้รูปแบบข้อมูลที่คล้ายคลึงกัน ดังเทคนิคของงานวิจัยของ Shuai Li, Hongjie Ma, and Weiyi Li (2017) ที่ได้ดำเนินการแบ่งกลุ่มข้อมูลแบบเคมีนส์ มาแก้ปัญหาเกี่ยวกับเกณฑ์คัดเลือกข้อมูล และสามารถทำงานขยายมาตราส่วนจากเดือนเป็นปี แม้จำนวนข้อมูลจะน้อยกว่าวิธีดั้งเดิมก็ตาม ต่อมาในงานวิจัยของ Hengxu Zhang, Yongji Cao, et al., (2018), Khalil Benmouiz, Ali Cheknane (2013), Nuño Basurto, Ángel Arroyo, Rafael Vega, et al., (2019), และ Pedro Chévez, Dante Barbero, et al., (2017) มีการนำการแบ่งกลุ่มข้อมูลแบบเคมีนส์มาใช้เป็นองค์ประกอบของงานวิจัยเพื่อปรับปรุงคุณภาพ เนื่องจากมีความซับซ้อนที่น้อยกว่าการแบ่งกลุ่มแบบอื่น อย่างเช่นงานวิจัยของ Peijie Lin, Zhouning Peng, et al. (2018) ที่ได้ผลลัพธ์ที่ดีขึ้นจากการนำการแบ่งกลุ่มข้อมูลแบบเคมีนส์ไปเป็นหนึ่งในขั้นตอน เมื่อเทียบกับกระบวนการก่อนปรับปรุง หรือกระบวนการเปรียบเทียบในแต่ละงานวิจัยดังกล่าว สุดท้ายในงานวิจัยของ Masoud Masoumi (2021) เป็นงานวิจัยที่มีการใช้กระบวนการแบ่งกลุ่มข้อมูลแบบเคมีนส์ และการวิเคราะห์ภาพเงา เป็นงานวิจัยที่ใหม่ล่าสุดที่งานวิจัยนี้ศึกษา นับเป็นต้นกำเนิดให้แก่งานวิจัยเล่มนี้ด้วย และจากการศึกษางานวิจัยข้างต้นเหล่านี้ทำให้พบว่า การแบ่งกลุ่มข้อมูลแบบเคมีนส์ยังเป็นกระบวนการที่เหมาะสมในปัจจุบัน เห็นได้จากมีงานวิจัยต่าง ๆ ในช่วงหลายปีนี้ยังมีการใช้งานการแบ่งกลุ่มข้อมูลแบบเคมีนส์ ไม่ว่าจะด้านการวิเคราะห์ หรือการปรับแต่งขั้นตอนกระบวนการ ซึ่งก่อให้เกิดผลลัพธ์ไปในเชิงบวกทั้งสิ้น จึงเป็นสาเหตุที่ทำให้งานวิจัยนี้เลือกใช้การแบ่งกลุ่มข้อมูลแบบเคมีนส์เป็นอัลกอริทึมหลัก

1.2 วัตถุประสงค์

1.2.1. เพื่อศึกษาการหาพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปีด้วยการแบ่งกลุ่มข้อมูลแบบเคมีนส์

1.2.2. เพื่อสร้างรูปแบบของการหาข้อมูล หรือกระบวนการที่เหมาะสม สำหรับการหาพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปี

1.2.3. เพื่อแสดงพื้นที่ผลลัพธ์ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปีที่งานวิจัยนี้ค้นพบ

1.3 ขอบเขตและข้อจำกัดในงานวิจัย

1.3.1. งานวิจัยนี้ใช้ข้อมูลพลังงานแสงอาทิตย์ปี พ.ศ.2560 (นัยนา บุญนาค, 2563) เพียงหนึ่งชุดข้อมูลเท่านั้น

1.3.2. ใช้อัลกอริทึมการแบ่งกลุ่มข้อมูลแบบเคมีนส์ (k-means clustering)

1.3.3. ใช้โปรแกรม Orange Data mining (Demšar, et al., 2013) เป็นเครื่องมือในการจัดกลุ่มข้อมูล

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1. สามารถใช้ประเมินความเหมาะสมในการหาพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปี

1.4.2. สามารถใช้เปรียบเทียบความแตกต่างของการวิเคราะห์ของพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปีระหว่างมนุษย์ และคอมพิวเตอร์

1.4.3. สามารถนำกระบวนการวิจัยนี้ไปใช้เพื่อวิเคราะห์ของพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงในหนึ่งปี สำหรับปีอื่น ๆ

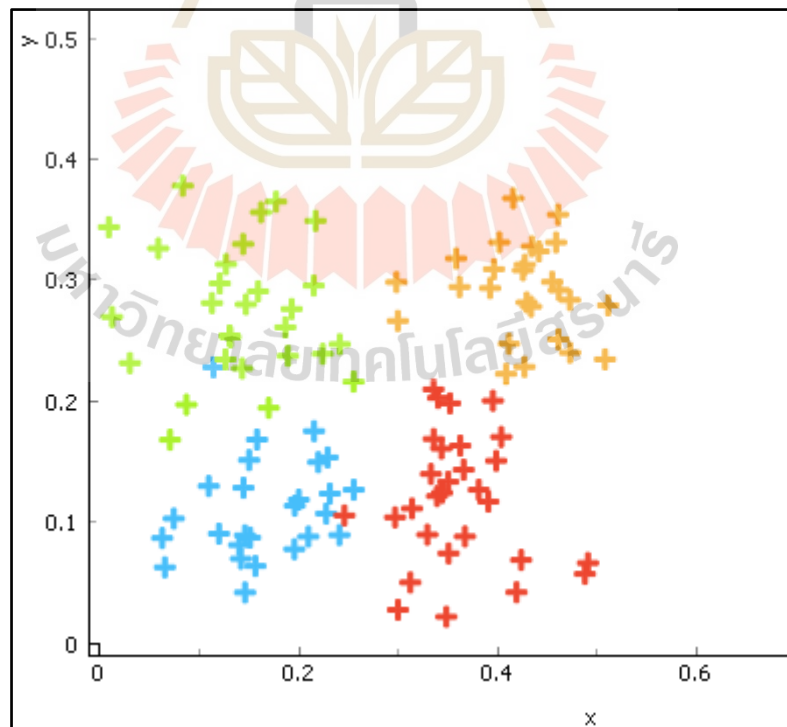
1.4.4. สามารถนำกระบวนการวิจัยนี้ไปประยุกต์เพื่อวิเคราะห์ข้อมูลในรูปแบบอื่น ๆ

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง

2.1 การแบ่งข้อมูลแบบเคมีนส์ (k-means clustering)

2.1.1 ความเป็นมาของการแบ่งข้อมูล (Clustering)

การแบ่งข้อมูล เป็นกระบวนการการจัดสรรข้อมูลให้อยู่เป็นกลุ่ม หรือหมวดหมู่จากข้อมูลที่กระจัดกระจายกัน การแบ่งข้อมูล คือการหาจุดรวมของข้อมูลในกลุ่มนั้น และกำหนดจุดแตกต่างกันระหว่างกลุ่มข้อมูล ได้ผลลัพธ์ดังแสดงดังรูปที่ 2.1 และเนื่องจากความต้องการในการจัดระเบียบข้อมูลอย่างรวดเร็ว การแบ่งข้อมูล เป็นหนึ่งในกระบวนการหาคำตอบที่ตอบสนองความต้องการใช้งานในสาขาต่าง ๆ เช่น ปัญญาประดิษฐ์ (Artificial Intelligence), ชีววิทยา, การค้า, การจัดการความสัมพันธ์, การบีบอัดข้อมูล, การทำเหมืองข้อมูล (Data mining), การดึงข้อมูล, การประมวลผลภาพ, การเรียนรู้ของเครื่อง (Machine learning), การตลาด, การแพทย์, การจดจำรูปแบบ, จิตวิทยา และด้านสถิติอื่น ๆ (Yadav, and Sharma, 2013)



รูปที่ 2.1 ตัวอย่างการแบ่งข้อมูล

2.1.2 การหาระยะห่างของการแบ่งข้อมูล

Yadav, J., and Sharma, M. (2002) ได้กล่าวถึงระยะห่างของการแบ่งข้อมูลเป็นระยะที่ข้อมูลนั้น ๆ จะถูกเลือกกว่าควรอยู่กลุ่มข้อมูลใด โดยมีสมการดังต่อไปนี้มาช่วยในการหาระยะห่างของการแบ่งข้อมูล ดังนี้

1. สมการ Euclidean distance

มีรูปสมการดังนี้

$$d(i, j) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (2-1)$$

เมื่อ $i = x_i$ และ $j = y_i$ เป็นจุด 2 จุดบน n -มิติ

2. สมการ Manhattan distance

มีรูปสมการดังนี้

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{in} - x_{jn}| \quad (2-2)$$

3. สมการ Minkowski distance

เป็นการรวมกันของสมการ Euclidean distance กับสมการ Manhattan distance จึงได้รูปสมการดังนี้

$$d(i, j) = (|x_{i1} - x_{j1}|^p + |x_{i2} - x_{j2}|^p + \dots + |x_{in} - x_{jn}|^p)^{1/p} \quad (2-3)$$

เมื่อ p เป็นจำนวนเต็มบวก ระยะทางดังกล่าวเรียกว่า L_p norm แทนที่ Manhattan distance เมื่อ $p = 1$ และแทนที่ Euclidean distance เมื่อ $p = 2$

โดยจากสมการ Minkowski distance เป็นตั้งสมการที่สามารถปรับรูปแบบได้ โดยการเปลี่ยนค่า p มีความยืดหยุ่นในการใช้งาน นอกจากนี้สามารถยกตัวอย่างสมการ Euclidean distance และสมการ Manhattan distance แบบ 1 มิติ, 2 มิติ, และ 3 มิติ โดยที่มิติที่ 1 แทนด้วย x , มิติที่ 2 แทนด้วย y , และมิติที่ 3 แทนด้วย z ได้ดังนี้

1 มิติ [x]

- Euclidean distance

$$d(i, j) = \sqrt{(x_i - x_j)^2} \quad (2-4)$$

- Manhattan distance

$$d(i, j) = |x_i - x_j| \quad (2-5)$$

2 มิติ $[x, y]$

- Euclidean distance

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2-6)$$

- Manhattan distance

$$d(i, j) = |x_i - x_j| + |y_i - y_j| \quad (2-7)$$

3 มิติ $[x, y, z]$

- Euclidean distance

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \quad (2-8)$$

- Manhattan distance

$$d(i, j) = |x_i - x_j| + |y_i - y_j| + |z_i - z_j| \quad (2-9)$$

ตารางที่ 2.1 ตัวอย่างตำแหน่งแต่ละแกนของแต่ละจุด สำหรับการคำนวณหาระยะห่าง

จุด	ตำแหน่งบนแกน		
	x	y	z
i	7	4	6
j	5	1	8

เพื่อแสดงให้เห็นการคำนวณของ Euclidean distance และ Manhattan distance อย่างชัดเจน โดยกำหนดค่าตัวอย่างดังตารางที่ 2.1 แล้วนำค่าในตารางนี้ไปแทนค่าในสมการ โดยยกตัวอย่างเป็นแบบ 3 มิติได้ดังนี้

- Euclidean distance

แทนค่าในสมการ 2-8 ได้

$$d(i, j) = \sqrt{(7 - 5)^2 + (4 - 1)^2 + (6 - 8)^2}$$

$$d(i, j) = \sqrt{4 + 9 + 4}$$

$$d(i, j) = \sqrt{17}$$

$$d(i, j) = 4.123$$

- Manhattan distance

แทนค่าในสมการ 2-9 ได้

$$d(i, j) = |7 - 5| + |4 - 1| + |6 - 8|$$

$$d(i, j) = 2 + 3 + 2$$

$$d(i, j) = 7$$

2.1.3 ทฤษฎีของการแบ่งข้อมูลแบบเคมีนส์

การแบ่งข้อมูลแบบเคมีนส์ เป็นหนึ่งในกระบวนการหาคำตอบของการเรียนรู้ของเครื่องในรูปแบบของการเรียนรู้แบบไม่มีผู้สอน ใช้สำหรับการแบ่งกลุ่มข้อมูลออกมา k กลุ่ม โดยจะแทนค่าแต่ละกลุ่มด้วยค่าเฉลี่ยในแต่ละกลุ่มซึ่งมาจากระยะห่างจากจุดศูนย์กลางกลางของแต่ละกลุ่ม (Centroid) โดยมีกระบวนการดังนี้ (Kanungo, et al., 2002, quoted in Masoumi, 2021)

1. กำหนดตัวแปร k และจุดศูนย์กลางกลางของแต่ละกลุ่ม
2. นำข้อมูลทุกจุดมาคำนวณหาระยะทางระหว่างจุดข้อมูลนั้นกับจุดศูนย์กลางกลางของแต่ละกลุ่ม โดยที่ค่าระยะทางที่น้อยที่สุดในเส้นทางของจุดศูนย์กลางกลางของแต่ละกลุ่มดังกล่าวให้ถือว่าอยู่ในกลุ่มเดียวกับจุดศูนย์กลางกลางของแต่ละกลุ่มนั้น
3. หาค่าเฉลี่ยทุกกลุ่ม แล้วกำหนดจุดศูนย์กลางกลางของแต่ละกลุ่มขึ้นมาใหม่
4. ทำวนซ้ำที่ 2. ใหม่ จนกว่าค่าจุดศูนย์กลางกลางของแต่ละกลุ่มนั้น ๆ ไม่มีการเปลี่ยนแปลงค่า

2.1.4 ข้อเสียของการแบ่งข้อมูลแบบเคมีนส์

Yadav, J., and Sharma, M. (2013) ได้กล่าวถึงข้อเสียของการแบ่งข้อมูลแบบเคมีนส์ไว้ดังนี้

1. จุดศูนย์กลางกลางของแต่ละกลุ่มที่กำหนดในตอนแรกมีความอ่อนไหวมาก มีผลให้การวิเคราะห์ที่เกิดคลาดเคลื่อนได้ง่าย
2. ไม่มีกฎในการกำหนดค่า k ที่ตายตัว ส่งผลให้ผลลัพธ์ที่ออกมาในค่า k ที่แตกต่างกัน มีความแตกต่างกันโดยสิ้นเชิง
3. ง่ายต่อการได้รับผลกระทบจากจุดที่ไม่ปกติ เมื่อข้อมูลมีการกระจายกันมาก
4. มีโอกาสเกิดปัญหาด้าน Dead Unit

2.2 การวิเคราะห์ภาพเงา (Silhouette Analysis)

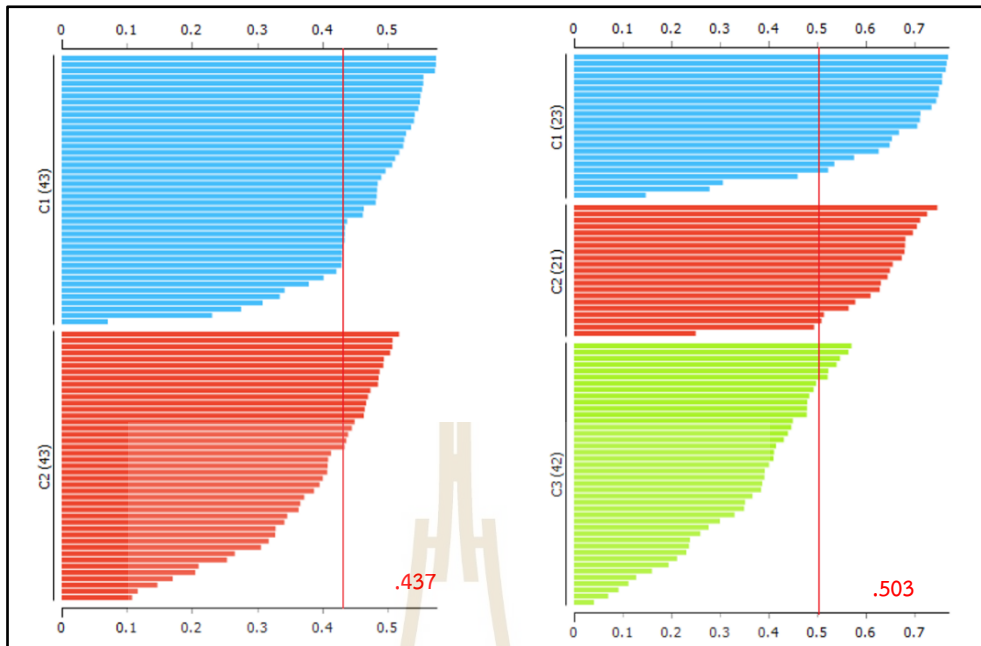
การกำหนดค่าของตัวแปร k ในการแบ่งข้อมูลแบบเคมีนส์มีความท้าทายว่าตัวแปร k ที่ต้องใช้ว่าควรใช้ค่าใดจึงเหมาะสมสำหรับชุดข้อมูล โดยการวิเคราะห์ภาพเงา เป็นหนึ่งหนทางที่สามารถหาค่าตัวแปร k ที่เหมาะสมกับการแบ่งข้อมูลแบบเคมีนส์ (Kodinariya, and Makwana, 2013 quoted in Masoumi, 2021) โดยเริ่มต้นกำหนดช่วงของค่า k ออกมาก่อน จากนั้นหาค่า Silhouette Scores ซึ่งค่านี้นบอกถึงความเหมาะสมข้อมูลนั้นกับกลุ่มที่เป็นสมาชิกอยู่และค่า Silhouette Score สามารถมีค่าได้ระหว่าง -1 กับ 1 โดยค่าที่ใกล้ 1 มากให้ถือว่าข้อมูลนั้นมีความเหมาะสมที่จะเป็นสมาชิกกลุ่มนั้นมาก โดยแสดงดังสมการที่ 2-10, 2-11 และ 2-12 (Rousseeuw, 1987 quoted in Masoumi, 2021)

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2-10)$$

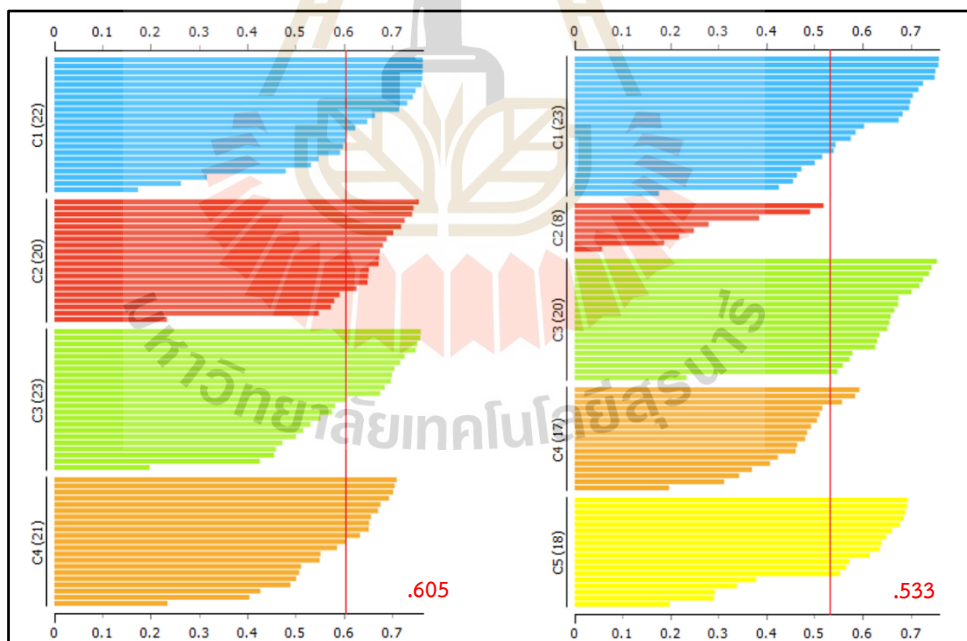
$$s(i) = \begin{cases} 1 - \frac{a(i)}{b(i)}, & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, & \text{if } a(i) > b(i) \end{cases} \quad (2-11)$$

$$-1 \leq s(i) \leq 1 \quad (2-12)$$

เมื่อ $a(i)$ คือ ค่าเฉลี่ยของระยะทางของข้อมูลจุดนั้นกับสมาชิกทุกตัวในกลุ่มเดียวกัน และ $b(i)$ คือ ค่าเฉลี่ยของระยะทางของข้อมูลนั้นกับสมาชิกทุกตัวในกลุ่มที่ใกล้ที่สุด โดยค่า $s(i)$ จากสมการเป็นค่าที่ได้เฉพาะจุดข้อมูลนั้น ๆ จากนั้นนำค่า Silhouette Scores ทุกค่าของแต่ละค่าตัวแปร k มาหาค่าเฉลี่ยแล้วเปรียบเทียบว่าค่าตัวแปร k ใดที่เหมาะสมกับชุดข้อมูลดังกล่าว โดยมีผลเชิงคุณภาพแสดงให้เห็นเป็นค่าตัวเลขใน Silhouette Scores โดยแสดงเป็นตัวอย่างดังรูปที่ 2.2 – 2.3



รูปที่ 2.2 ตัวอย่าง Silhouette Scores ของค่า k ที่ 2 – 3



รูปที่ 2.3 ตัวอย่าง Silhouette Scores ของค่า k ที่ 4 – 5

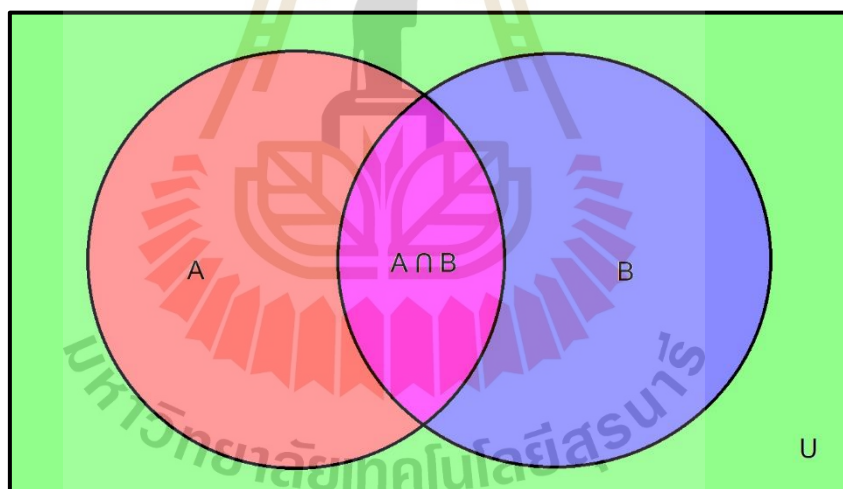
เมื่อสังเกตจากรูปที่ 2.2 – 2.3 พบว่าค่าเฉลี่ย Silhouette Scores ของ k ที่ 4 มีค่าที่มากที่สุด เมื่อเทียบกับค่าเฉลี่ย Silhouette Scores (เทียบจากเส้นสีแดงบนรูปที่ 2.2 – 2.3) ของค่า k อื่น ๆ

โดยหาค่าของ Silhouette Scores ของข้อมูลนั้นมีมาก หมายความว่าข้อมูลนั้นอยู่ใกล้กับจุดศูนย์กลางกลางของแต่ละกลุ่มในกลุ่มนั้นมาก ทำให้ผู้ที่ทำการแบ่งข้อมูลแบบเคมีนส์ได้ค่า k ที่เหมาะสมสำหรับการใช้งานในการแบ่งข้อมูล

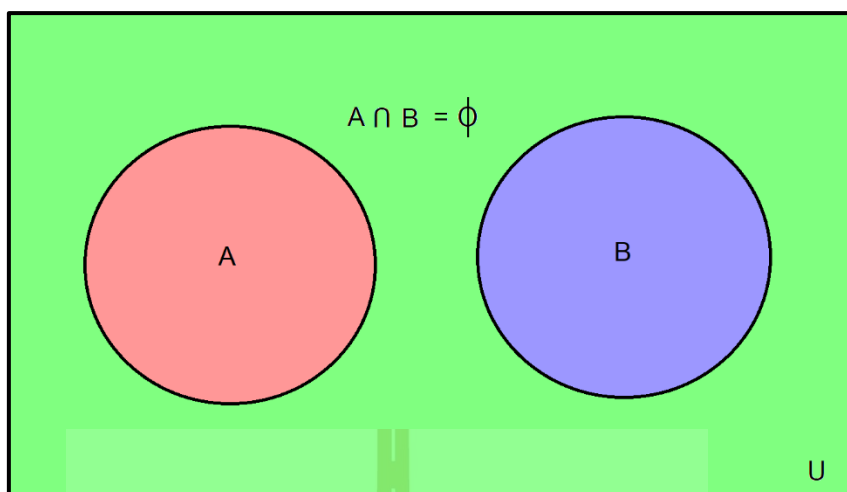
2.3 ทฤษฎีการอินเตอร์เซค (Intersection) และความเกี่ยวข้องกับงานวิจัย

นิยามของการอินเตอร์เซคคือ หากกำหนดให้เซต A อินเตอร์เซค กับเซต B โดยผลลัพธ์ที่ได้เป็นสมาชิกร่วมกันระหว่างเซต A และเซต B หากไม่มีสมาชิกร่วมกันระหว่างเซต A และเซต B ผลลัพธ์ที่ได้เป็นเซตว่าง (Šen, and Verescagin, 2002) โดยสามารถแสดงประกอบคำอธิบายได้ดังรูปที่ 2.4 – 2.5 และมีสมบัติการอินเตอร์เซคไว้ดังนี้

1. $A \cap B = B \cap A$
2. $A \cup \emptyset = \emptyset$
3. $A \cap A' = \emptyset$;เมื่อ A' คืออินเวอร์ส (Inverse) ของ A
4. $A \cap A = A$



รูปที่ 2.4 การอินเตอร์เซคระหว่างเซต A และ B กรณีมีสมาชิกร่วมกัน



รูปที่ 2.5 การอินเตอร์เซกระหว่างเซต A และ B กรณีไม่มีสมาชิกร่วมกัน

ในงานวิจัยนี้ การอินเตอร์เซกมีความสำคัญต่อการรวมข้อมูลเป็นอย่างมาก และถือเป็นกระบวนการสำคัญของงานวิจัยนี้ เนื่องจากงานวิจัยนี้มีข้อมูลหลายชุดที่ต้องการการอินเตอร์เซก เพื่อใช้สำหรับวิเคราะห์ และสรุปผลการวิจัย

2.4 ความเป็นมาของแผนภาพกล่อง (Box Plot)

Box Plot หรืออีกชื่อ Whisker Plot เป็น Data Visualization ประเภทหนึ่งที่ใช้สำหรับการอธิบายลักษณะของกลุ่มข้อมูล โดยจะมาในรูปแบบกราฟแท่งเบื้องต้นที่สามารถช่วยให้การวิเคราะห์ข้อมูล (Data Analysis) เบื้องต้นนั้นสามารถอ่าน และทำความเข้าใจได้ง่ายขึ้น โดยมีการแสดงข้อมูลดังรูปที่ 2.6 และคำอธิบายดังนี้ (Mcleod, 2019; keawkuay, 2020)

Minimum Score คือค่าที่น้อยที่สุดของชุดข้อมูลนั้น

Maximum Score คือค่าที่มากที่สุดของชุดข้อมูลนั้น

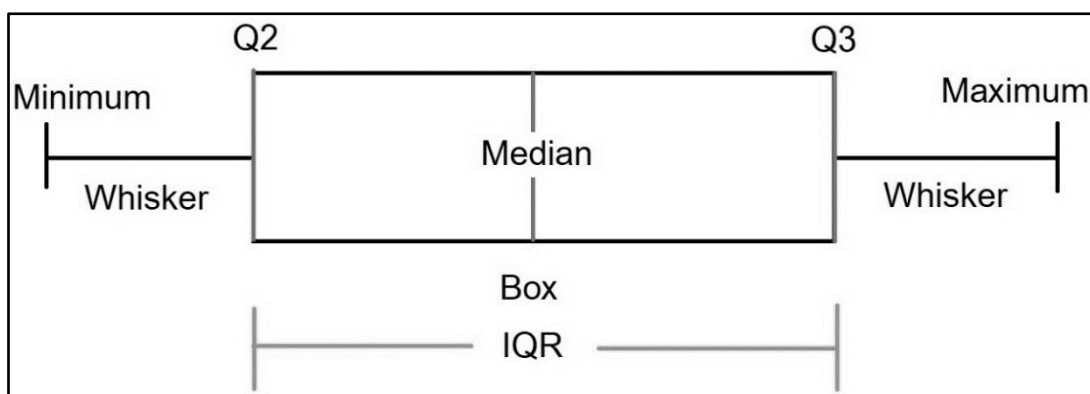
Lower Quartile (Q2) คือค่าของข้อมูลที่มีขนาดประมาณ 25% หรือหนึ่งในสี่ของชุดข้อมูลนั้น

Upper Quartile (Q3) คือค่าของข้อมูลที่มีขนาดประมาณ 75% หรือสามในสี่ของชุดข้อมูลนั้น

Median คือค่าของข้อมูลที่มีขนาดประมาณ 50% หรือหนึ่งในสองของชุดข้อมูลนั้น

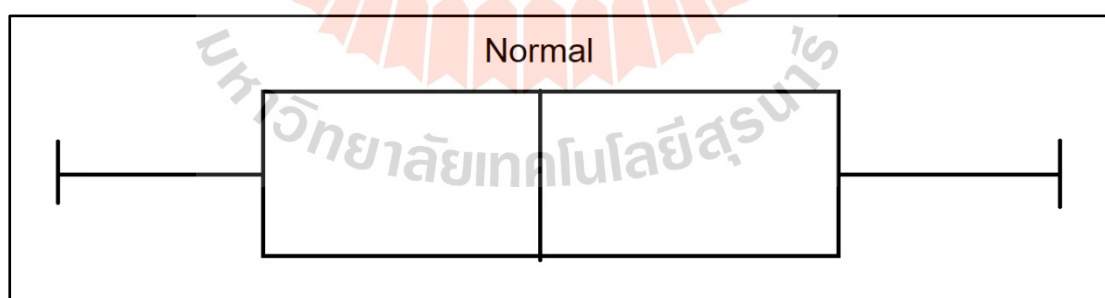
Whiskers คือค่านอกตัวแผนภาพกล่อง เป็นค่าระหว่าง Minimum Score ถึง Lower Quartile และเป็นค่าระหว่าง Upper Quartile ถึง Maximum Score

The Interquartile Range (IQR) คือค่าตัวแผนภาพกล่อง โดยในแผนภาพกล่องจะมีค่าอยู่ระหว่างประมาณ 25% - 75% หรือคือค่าระหว่าง Lower Quartile กับ Upper Quartile

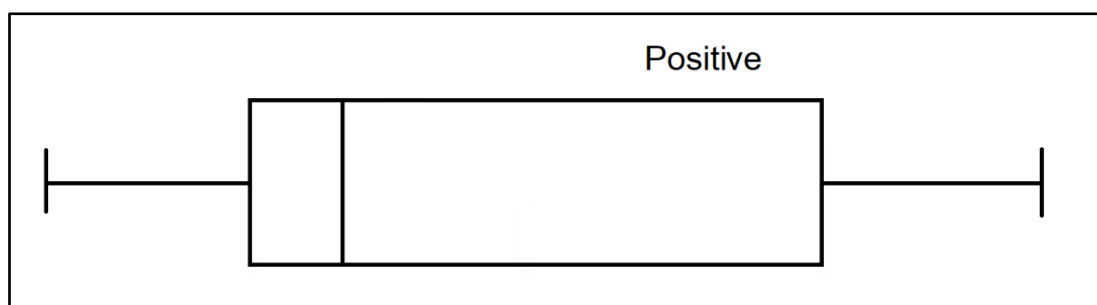


รูปที่ 2.6 ลักษณะของแผนภาพกล่อง

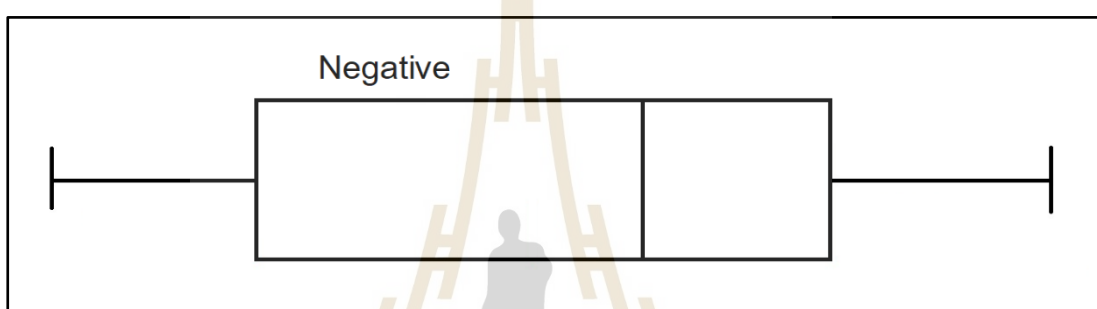
การสังเกตว่าแผนภาพกล่องแบบใดเป็นกลุ่มของข้อมูลที่เหมาะสม สามารถสังเกตได้จากการกระจายตัว (Distribution) โดยสังเกตได้จากเส้น Median ว่าอยู่ใกล้เคียงกับตรงกลางของ Box หรือ IQR หรือไม่ หากใช้หมายความว่าข้อมูลชุดดังกล่าวมีการกระจายตัวของข้อมูลแบบปกติ (Normal Distribution) แสดงดังรูปที่ 2.7 โดยที่ค่าช่วงของ Q2 และ Q3 มีขนาดที่ใกล้เคียงกัน แต่ถ้าไม่แสดงว่าข้อมูลมีการเอียงเอนไปทางฝั่งใดฝั่งหนึ่ง หากข้อมูลมีการเอียงเอนไปทางฝั่งข้างมาก (Positive Distribution) นั้นหมายความว่าค่าของ Median อยู่ใกล้กับ Lower Quartile ทำให้ช่วงของ Q3 มีช่วงที่กว้างมากกว่ามาก ๆ แสดงดังรูปที่ 2.8 หรือหากข้อมูลมีการเอียงเอนไปทางฝั่งข้างน้อย (Negative Distribution) นั้นหมายความว่าค่าของ Median อยู่ใกล้กับ Upper Quartile ทำให้ช่วงของ Q2 มีช่วงที่กว้างมากกว่ามาก ๆ แสดงดังรูปที่ 2.9 (McLeod, 2019)



รูปที่ 2.7 การกระจายตัวแบบปกติ (Normal Distribution)



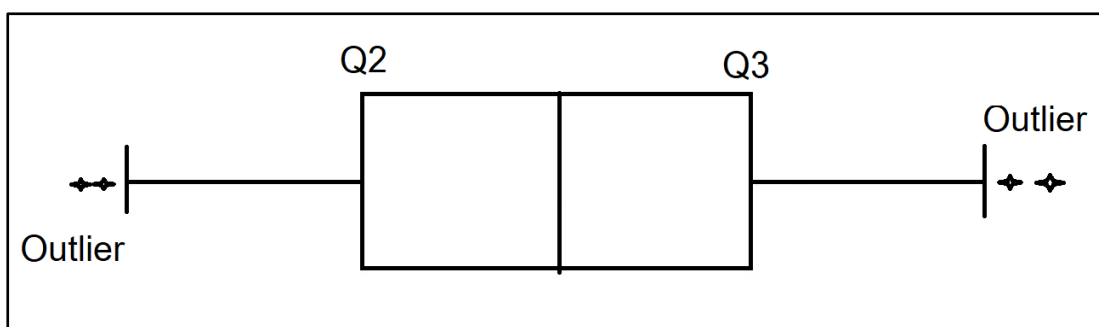
รูปที่ 2.8 การกระจายตัวแบบการเอียงเอนไปทางฝั่งข้างมาก (Positive Distribution)



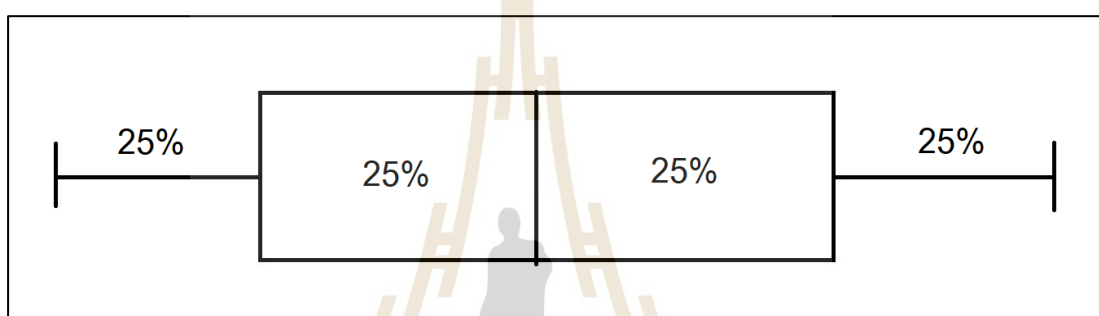
รูปที่ 2.9 การกระจายตัวแบบเอียงเอนไปทางฝั่งข้างน้อย (Negative Distribution)

แผนภาพกล่องสามารถแสดงค่าผิดปกติ (Outliers) ที่อยู่ในชุดข้อมูล โดยที่ค่าผิดปกติเหล่านั้นมักมีจำนวนไม่มาก เนื่องจากค่าเหล่านี้เป็นค่าที่อยู่นอก Minimum Score หรือ Maximum Score และค่าเหล่านั้นยังไม่อยู่ Whiskers หรือ Box อีกด้วย กล่าวอีกความหมายคือเป็นข้อมูลที่อยู่ นอกเหนือแผนภาพกล่องนั้น ๆ โดยข้อมูลที่เข้าข่ายค่าผิดปกติของชุดข้อมูล สามารถหาค่าข้อมูล เหล่านั้นได้จากการนำค่าเหล่านั้นมาเทียบกับเงื่อนไขการหาค่าผิดปกติ เมื่อข้อมูลนั้นมีค่าน้อยกว่า $Q2 - (1.5 \times IQR)$ หรือข้อมูลนั้นมีค่ามากกว่า $Q3 + (1.5 \times IQR)$ หากเป็นจริงแสดงว่าค่านั้นเป็นค่า ผิดปกติโดยค่าเหล่านั้นแสดงได้ดังรูปที่ 2.10 (Mcleod, 2019)

ประโยชน์ของแผนภาพกล่องในด้านการศึกษาและการใช้งานโดยแผนภาพกล่องนั้นมีการ แสดงให้เห็นถึงช่วงข้อมูลทั้งสี่ช่วงตั้งแต่จุดที่ต่ำที่สุดถึงจุดที่สูงที่สุด โดยแต่ละช่วงประมาณ 25% ของ ข้อมูลทั้งหมด (Mcleod, 2019) แสดงดังรูปที่ 2.11 และสามารถทราบถึงลักษณะของกลุ่มข้อมูลนั้น ได้แม้ไม่ทราบหรือเคยเห็นข้อมูลในกลุ่ม เหมาะสมสำหรับใช้แสดงลักษณะของกลุ่มข้อมูลที่มีปริมาณ ของข้อมูลมาก ช่วยให้สามารถคาดคะเนค่าเฉลี่ยจากการวิเคราะห์ลักษณะของแผนภาพกล่อง และ ด้วยลักษณะการแสดงผลดังกล่าวส่งผลให้สามารถวิเคราะห์ได้โดยไม่ต้องผ่านการอ่านข้อมูลจากด้าน ใน



รูปที่ 2.10 ตำแหน่งของค่าผิดปกติ



รูปที่ 2.11 ขนาดของแผนภาพกล่องในแต่ละช่วงที่ควรจะเป็น

2.5 ความสำคัญของพลังงานแสงอาทิตย์

2.5.1 พลังงานแสงอาทิตย์

พลังงานแสงอาทิตย์เป็นพลังงานที่เกิดขึ้นมาตามธรรมชาติ เป็นพลังงานที่สามารถนำมาหมุนเวียนใช้งานได้อย่างไม่สิ้นสุด เนื่องจากแสงอาทิตย์ได้ปล่อยพลังงานความร้อนลงมาถึงโลกเป็นพลังงานสะอาด ไม่มีมลพิษ และยังสามารถนำมาใช้ประโยชน์ได้อย่างหลากหลาย เช่น พลังงานไฟฟ้าแสงอาทิตย์, พลังงานความร้อนสำหรับผลิตน้ำร้อน, และพลังงานความร้อนสำหรับผลิตอาหารตากแห้ง (พลังงานแสงอาทิตย์กับการนำมาใช้ประโยชน์, (2020); สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย [TRECA], ม.ป.ป.)

พลังงานแสงอาทิตย์ถือเป็นต้นกำเนิดของพลังงานทางธรรมชาติหลายอย่าง เช่น พลังงานลม, และพลังงานน้ำ โดยที่ในแต่ละปีมีพลังงานแสงอาทิตย์มายังโลกประมาณ 173,000 เทราวัตต์ (พันล้านกิโลวัตต์) คิดเป็นประมาณ 35% ของพลังงานแสงอาทิตย์ทั้งหมดที่เข้ามาในชั้นบรรยากาศโลก และลงมาถึงพื้นผิวโลก ด้วยค่าปริมาณพลังงานแสงอาทิตย์นี้มีค่าเทียบเท่ากับพลังงานไฟฟ้าที่ใช้ทั่วโลกรวมกัน 10,000 เท่า (Chandler, 2011, quoted in Banpu NEXT Co., Ltd., 2020) และถือเป็นพลังงานที่มีการใช้งานอย่างแพร่หลายมากกว่า 60 ปี โดยใช้แผงโซลาร์เซลล์ ในการผลิตกระแสไฟฟ้าที่ได้มาจากแสงอาทิตย์ โดยใช้ตั้งแต่จานดาวเทียมบนอวกาศมาจนถึงระดับครัวเรือน

เช่น ระบบผลิตไฟฟ้าพลังงานแสงอาทิตย์ที่ติดตั้งบนหลังคา (Solar Rooftop) และระบบผลิตไฟฟ้าพลังงานแสงอาทิตย์แบบติดตั้งบนพื้นดิน (Solar Farm) พลังงานแสงอาทิตย์เป็นพลังงานที่ให้มลพิษน้อย และประหยัดทรัพยากรมากกว่าการผลิตไฟฟ้ารูปแบบอื่น ๆ (Banpu NEXT Co., Ltd., 2020)

ปัจจุบันการติดตั้งระบบแผงโซลาร์เซลล์ได้รับความสนใจเป็นอย่างมากจากการเห็นความสำคัญเกี่ยวกับพลังงานสะอาด และเทคโนโลยีโซลาร์เซลล์ที่พัฒนาไปมาก ทำให้ได้แผงโซลาร์เซลล์ที่มีประสิทธิภาพ และต้นทุนต่ำ โดยในประเทศสหรัฐอเมริกามีการผลิตไฟฟ้าจากพลังงานแสงอาทิตย์สูงถึง 97.2 กิกะวัตต์ต่อปี ซึ่งเพียงพอกับ 18.9 ล้านครัวเรือนในประเทศสหรัฐอเมริกา (Solar Energy Industries Association [SEIA], 2021) และที่ประเทศญี่ปุ่นได้สร้างโรงไฟฟ้าจากพลังงานแสงอาทิตย์แบบลอยน้ำแห่งแรกของโลกขึ้นที่จังหวัดไอจิ ทางตอนกลางของเกาะฮอนชู ในทะเลสาบ และอ่างเก็บน้ำหลายแห่ง โดยในปี ค.ศ. 2019 โรงไฟฟ้าพลังงานแสงอาทิตย์แบบลอยน้ำดังกล่าวมีขนาดใหญ่เป็นลำดับที่ 73 จาก 100 ทั่วโลก และมีกำลังการผลิตถึง 50% ของโรงไฟฟ้าพลังงานแสงอาทิตย์แบบลอยน้ำทั่วโลก โดยมีกำลังการผลิตประมาณ 246 เมกะวัตต์ (Broom, 2019, quoted in Banpu NEXT Co., Ltd., 2020)

2.5.2 เทคโนโลยีพลังงานไฟฟ้าแสงอาทิตย์โซลาร์เซลล์

เทคโนโลยีพลังงานไฟฟ้าแสงอาทิตย์โซลาร์เซลล์ มีการแบ่งออกเป็นสามระบบดังนี้ (สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

1. เซลล์แสงอาทิตย์แบบอิสระ (Photovoltaic (PV) Stand-alone system) เป็นระบบที่ออกแบบมาให้ใช้สำหรับชนบทที่ไม่มีระบบสายส่งไฟฟ้า ประกอบไปด้วยแผงโซลาร์เซลล์, และแบตเตอรี่ โดยมีหลักการทำงานเบื้องต้นแบ่งเป็นสองช่วงคือ ช่วงเวลากลางวันเก็บพลังงานแสงอาทิตย์และชาร์จแบตเตอรี่ พร้อมกับการใช้งานพลังงานไฟฟ้าไปด้วย ส่วนช่วงเวลากลางคืนไม่สามารถเก็บพลังงานแสงอาทิตย์ได้ จึงเปลี่ยนการใช้งานพลังงานไฟฟ้าไปด้วยแบตเตอรี่แทน (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014i)

2. เซลล์แสงอาทิตย์แบบต่อกับระบบจำหน่าย (Photovoltaic Grid connected system) เป็นระบบที่ออกแบบมาให้ใช้สำหรับใช้งานในเขตเมืองที่มีไฟฟ้าเข้าถึง มีแตกต่างจากเซลล์แสงอาทิตย์แบบอิสระ (Photovoltaic Stand-alone system) กลางวันมีการใช้งานใช้พลังงานไฟฟ้าจากแผงโซลาร์เซลล์โดยตรง ส่วนกลางคือจะใช้พลังงานไฟฟ้าจากระบบครัวเรือนปกติ (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014g)

3. เซลล์แสงอาทิตย์แบบผสมผสาน (Photovoltaic Hybrid system) เป็นระบบที่ออกแบบมาให้ใช้สำหรับใช้งานร่วมกับอุปกรณ์ผลิตไฟฟ้าประเภทอื่น ๆ เช่น เครื่องยนต์ดีเซล, ระบบผลิตไฟฟ้าพลังงานลม, และระบบผลิตไฟฟ้าพลังงานน้ำ เป็นต้น หลักการคือผลิตพลังงานไฟฟ้าพลังงานแสงอาทิตย์ ร่วมกับร่วมกับอุปกรณ์ผลิตไฟฟ้าประเภทอื่น ๆ แล้วชาร์จไว้ที่แบตเตอรี่ และมีการใช้งานพลังงานไฟฟ้าร่วมไปพร้อมกัน หากพลังงานไฟฟ้าไม่เพียงพอต่อการใช้งาน ระบบจะดึงพลังงานจากแบตเตอรี่มาใช้แทน (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014h)

2.5.3 เทคโนโลยีพลังงานแสงอาทิตย์เพื่อผลิตน้ำร้อน

เทคโนโลยีพลังงานแสงอาทิตย์เพื่อผลิตน้ำร้อนแบ่งออกเป็นสามชนิดดังนี้ (สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

1. การผลิตน้ำร้อนชนิดไหลเวียนตามธรรมชาติ (Thermosiphon System) เป็นการผลิตน้ำร้อนชนิดที่มีถังเก็บอยู่สูงกว่าแผงรับแสงอาทิตย์ โดยใช้หลักการของ Thermosiphon เมื่อน้ำมีอุณหภูมิสูงขึ้นจากแสงอาทิตย์ทำให้น้ำที่ร้อนมีความหนาแน่นน้อยลง ส่งผลให้น้ำเย็นไหลเข้ามาแทนที่ เหมาะสำหรับที่พักอาศัยขนาดเล็ก (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014c; Phanniti Co., Ltd., 2019)

2. การผลิตน้ำร้อนชนิดใช้ปั๊มน้ำหมุนเวียน (Force Circulation System) ใช้หลักการเดียวกับการผลิตน้ำร้อนชนิดไหลเวียนตามธรรมชาติ แต่จะใช้เครื่องปั๊มน้ำส่งน้ำไปที่แผงรับแสงอาทิตย์แทนการใช้หลักการของ Thermosiphon เหมาะสำหรับสถานที่ที่มีการใช้งานอย่างต่อเนื่อง เช่น โรงแรม, โรงพยาบาล, และอุตสาหกรรมบางแห่ง (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014a; สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

3. การผลิตน้ำร้อนชนิดผสมผสาน (Hybrid Solar Hot Water System) เป็นการใช้น้ำร้อนที่ได้จากการผลิตด้วยพลังงานแสงอาทิตย์ ร่วมกับน้ำร้อนจากแหล่งผลิตอื่น นิยมเป็นน้ำร้อนที่ทิ้งจากเครื่องปรับอากาศ โดยอุปกรณ์แลกเปลี่ยนความร้อน (Heat Exchanger) เพื่อลดพื้นที่แผงรับแสงอาทิตย์ และนำทรัพยากรน้ำร้อนจากกระบวนการอื่นมาวนใช้งานเพื่อให้เกิดความคุ้มค่า (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014b; พลังงานแสงอาทิตย์กับการนำมาใช้ประโยชน์, 2020)

2.5.4 เทคโนโลยีพลังงานแสงอาทิตย์เพื่อการอบแห้ง

ในปัจจุบันเทคโนโลยีพลังงานแสงอาทิตย์เพื่อการอบแห้ง มีการยอมรับอยู่สามระบบ ดังนี้ (พลังงานแสงอาทิตย์กับการนำมาใช้ประโยชน์, 2020; สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

2.5.4.1. การอบแห้งระบบ Passive คือการอบที่ใช้พลังงานแสงอาทิตย์ และกระแสลมในการอบแห้ง มีสามประเภทดังนี้ (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014f; พลังงานแสงอาทิตย์กับการนำมาใช้ประโยชน์, 2020; สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.) ได้แก่

- 1.) การตากแห้งแบบธรรมชาติ
- 2.) เครื่องอบแห้งแบบได้รับแสงอาทิตย์โดยตรง เป็นเครื่องที่ทำให้การรับพลังงานแสงอาทิตย์มีประสิทธิภาพด้วยวัสดุที่โปร่งใส ดูดกลืนความร้อนไว้ในตู้อบ โดยอากาศร้อนภายในเครื่องอบแห้งทำให้เกิดการหมุนเวียนของอากาศเพื่อช่วยถ่ายเทอากาศขึ้นได้ดี

3.) เครื่องอบแห้งพลังงานแสงอาทิตย์แบบผสม เป็นเครื่องอบแห้งที่ใช้พลังงานแสงอาทิตย์ทั้งทางตรงตามธรรมชาติ และทางอ้อมจากแผงรับรังสีแสงอาทิตย์ ทำให้อากาศภายในเครื่องร้อนอย่างรวดเร็ว

2.5.4.2. การอบแห้งระบบ Active คือระบบอบแห้งที่มีเครื่องควบคุมการไหลเวียนของอากาศ ทำให้ความชื้นในอากาศถูกพัดออกไปอย่างรวดเร็ว ส่งผลให้การอบใช้ระยะเวลาที่สั้นกว่าการอบแห้งระบบ Passive นิยมใช้สำหรับอบแห้งอาหาร (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014d; พลังงานแสงอาทิตย์ กับการนำมาใช้ประโยชน์, 2020; สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

2.5.4.3. การอบแห้งระบบ Hybrid คือระบบอบแห้งที่ใช้พลังงานความร้อนจากพลังงานแสงอาทิตย์คู่กับการให้ความร้อนจากแหล่งพลังงานอื่นในกรณีที่พลังงานแสงอาทิตย์ไม่เพียงพอ และมีเครื่องควบคุมการไหลเวียนของอากาศ ส่งผลให้การอบใช้ระยะเวลาที่สั้นกว่าการอบแห้งระบบ Active (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน, 2014e; พลังงานแสงอาทิตย์ กับการนำมาใช้ประโยชน์, 2020; สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย, ม.ป.ป.)

2.6 งานวิจัยที่เกี่ยวข้อง

การศึกษาพื้นที่ที่ให้พลังงานสูงสุดตลอดปีในประเทศไทย ด้วยการแบ่งกลุ่มแบบเคมีนส์ มีงานวิจัยที่เป็นแนวทางให้งานวิจัยนี้ที่ศึกษาแนวทาง โดยประกอบไปด้วยงานวิจัยดังนี้

งานวิจัยของ Shuai Li, Hongjie Ma, and Weiyi Li (2017) นำเสนอเกี่ยวกับค่าการแผ่รังสีแสงอาทิตย์ประจำวัน (Daily solar radiation: DSR) ซึ่งค่าดังกล่าวส่งผลต่อการออกแบบ และควบคุมของระบบกักเก็บพลังงานแสงอาทิตย์ โดยปรกตินิยมใช้กระบวนการ Typical Solar Radiation Year (TSRY) ที่มีพื้นฐานมาจากการแบ่งกลุ่มแบบเคมีนส์ และ discrete-time Markov chain (DTMC) เพื่อวิเคราะห์ประวัติอัตราส่วนของความปลอดโปร่งของท้องฟ้า (clear-sky ratio: CSR) ในพื้นที่ตัวแทนสี่แห่ง และกำหนดเวกเตอร์ที่ถูกอธิบายด้วยคุณลักษณะหกมิติของ ค่าการแผ่รังสีแสงอาทิตย์ประจำวัน และอัตราส่วนของความปลอดโปร่งของท้องฟ้า จากนั้นใช้การแบ่งกลุ่มแบบเคมีนส์ เพื่อแบ่งกลุ่มออกเป็นสี่กลุ่ม ต่อจากนั้นใช้ discrete-time Markov chain สร้างโมเดลสำหรับการประมาณค่าการเปลี่ยนแปลงของค่าการแผ่รังสีแสงอาทิตย์ประจำวัน จากนั้นใช้กระบวนการ Typical Solar Radiation Year เพื่อเปรียบเทียบ

งานวิจัยของ Hengxu Zhang, Yongji Cao, Yi Zhang, and Vladimir Terzija (2018) นำเสนอเกี่ยวกับการวิเคราะห์ และพิจารณาเกี่ยวกับความผันผวนของพลังงานทดแทน โดยให้ความสำคัญของการประเมินทรัพยากรที่เพียงพอ เหมาะสมสำหรับแผนการลดความผันผวนทางพลังงาน เพื่อประเมินและเห็นลักษณะเชิงพื้นที่ของพลังงานลม และพลังงานแสงอาทิตย์ในภูมิภาคจากมุมมองของการทำเหมืองข้อมูล โดยการวิเคราะห์ข้อมูลจากอุทุนิยมวิทยาาระยะยาวของความเร็วลม, การแผ่รังสีแสงอาทิตย์, และอุณหภูมิแวดล้อมในระยะยาว มาประมวลผลผ่านระบบการสร้าง

เสมือนจริง (via virtual generation systems) จากนั้นใช้การแบ่งกลุ่มแบบเคมีนส์ เพื่อลดมิติ และจัดรูปแบบการทำงานร่วมกันเชิงพื้นที่

งานวิจัยของ Khalil Benmouiza, Ali Chekane (2013) นำเสนอเกี่ยวกับการพยากรณ์การแผ่รังสีดวงอาทิตย์ในแนวนอนทั่วโลกรายชั่วโมงโดยอิงจากการแบ่งกลุ่มแบบเคมีนส์ และโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) โดยใช้การแบ่งกลุ่มแบบเคมีนส์ใช้ในการดึงข้อมูลเพื่อสร้างแบบจำลอง Time series และหารูปแบบของพื้นที่อินพุตโดยการจัดกลุ่มข้อมูล นอกจากนี้มีการใช้ โครงข่ายประสาทเทียมแบบไม่เชิงเส้น (nonlinear autoregressive: NAR) เป็นหนึ่งในกระบวนการของโครงข่ายประสาทเทียม ในการสร้างแบบจำลอง Time series โดยกระบวนการดังกล่าวข้างต้นเพื่อให้ได้ผลลัพธ์ที่ดีขึ้น

งานวิจัยของ Nuño Basurto, Ángel Arroyo, Rafael Vega, Héctor Quintián, José Luis Calvo-Rolle, and Álvaro Herrero (2019) นำเสนอเกี่ยวกับการนำพลังงานแสงอาทิตย์ที่เป็นพลังงานหมุนเวียนที่มีการกักเก็บได้อย่างจำกัด เพิ่มประสิทธิภาพช่วงของการรวมกักเก็บพลังงานแสงอาทิตย์ด้วยวิธีที่เรียกว่า Hybrid Intelligent System (HIS) โดยการใช้การเรียนรู้ของเครื่องอย่างโครงข่ายประสาทเทียม และใช้การแบ่งกลุ่ม แล้วนำแบบจำลองทั้งหมดมาเปรียบเทียบกับข้อมูลจริง เพื่อค้นหาแบบจำลองที่ดีที่สุด

งานวิจัยของ Masoud Masoumi (2021) นำเสนอเกี่ยวกับการนำพลังงานไฮบริดอย่าง Wave-wind หรือพลังงานคลื่นลมมานำเสนอ เนื่องด้วยจากความต้องการลดต้นทุนระยะยาวสำหรับการผลิตพลังงานนอกชายฝั่ง โดยที่ประสิทธิภาพของระบบบางส่วนขึ้นอยู่กับคุณสมบัติของคลื่นและลมของสถานที่ติดตั้ง โดยการใช้การเรียนรู้ของเครื่องแบบการเรียนรู้แบบไม่มีผู้สอน และการวิเคราะห์ภาพเงา ในการจำแนกภูมิภาคต่าง ๆ โดยมีเป้าหมายคือการสร้างเครื่องมือการประเมินเบื้องต้นเพื่อปรับปรุงการตัดสินใจในกระบวนการออกแบบสำหรับระบบไฮบริด Wave-wind ที่มีการปรับตัวสูงภายในแต่ละภูมิภาคตามข้อมูลที่บันทึกไว้ วัตถุประสงค์เพื่อการแบ่งกลุ่มสำหรับพิจารณาคุณสมบัติของคลื่น และลม เพื่อเปรียบเทียบผลลัพธ์ นอกจากนี้ยังสามารถให้ข้อมูลสถิติโดยละเอียดของแต่ละกลุ่มที่หาได้จากการแบ่งกลุ่ม

งานวิจัยของ Pedro Chévez, Dante Barbero, Irene Martini, and Carlos Discoli (2017) นำเสนอเกี่ยวกับการตรวจหาพื้นที่ของการใช้ไฟฟ้าในที่อยู่อาศัยที่เป็นพื้นที่เดียวกัน โดยงานวิจัยนี้ระบุปัจจัยทางสังคมและประชากรหลักที่ส่งผลกระทบต่อความต้องการไฟฟ้า และที่ตั้งทางภูมิศาสตร์ของแต่ละพื้นที่สำหรับการใช้ไฟฟ้าในพื้นที่เดียวกัน ด้วยการแบ่งกลุ่มแบบเคมีนส์ เพื่อการคาดการณ์การใช้ไฟฟ้าในที่อยู่อาศัยที่เป็นพื้นที่เดียวกันซึ่งสามารถอธิบายได้ตามลักษณะทางประชากรและสังคม เช่น จำนวนห้องในบ้าน, จำนวนคนในบ้าน, เพอร์เซ็นต์ของความไม่พอใจในความต้องการพื้นฐานของแต่ละบ้าน, ความครอบคลุมของเครือข่ายก๊าซ, ประเภทที่อยู่อาศัย, และคุณภาพของที่อยู่อาศัย เป็นต้น ด้วยข้อมูลนี้สามารถช่วยในการเสนอกยุทธ์ประสิทธิภาพพลังงานและรวมพลังงานทางเลือกทดแทนโดยใช้เกณฑ์ที่เหมาะสมสำหรับแต่ละพื้นที่ในนโยบายที่แตกต่างกัน

งานวิจัยของ Peijie Lin, Zhouning Peng, Yunfeng Lai, Shuying Cheng, Zhicong Chen, and Lijun Wu (2018) นำเสนอเกี่ยวกับการใช้พลังงานแสงอาทิตย์ที่เป็นพลังงานสะอาด ไม่ก่อให้เกิดมลพิษสำหรับการผลิตไฟฟ้าจากเซลล์แสงอาทิตย์สามารถใช้พลังงานหมุนเวียนได้อย่างมีประสิทธิภาพ เนื่องจากความไม่แน่นอนของสภาพอากาศทำให้กำลังไฟฟ้าที่ผลิตไฟฟ้าจากเซลล์แสงอาทิตย์มีการผันผวน และไม่สม่ำเสมออย่างมาก ดังนั้นวิธีการทำนายพลังงานผลิตไฟฟ้าจากเซลล์แสงอาทิตย์ที่เชื่อถือได้สามารถลดข้อเสียของการผลิตไฟฟ้าจากเซลล์แสงอาทิตย์ ซึ่งมีความสำคัญอย่างมากต่อการบำรุงรักษา และซ่อมแซมโรงไฟฟ้า ด้วยทำนายแบบไฮบริดแบบใหม่ที่รวมการแบ่งกลุ่มแบบเคมีนส์, grey relational analysis (GRA), และโครงข่ายประสาทเทียมของ Elman สำหรับการทำนายการผลิตไฟฟ้าจากเซลล์แสงอาทิตย์ในระยะสั้น แบบจำลองที่สร้างขึ้นใช้ปัจจัยอุตุนิยมวิทยาหลายตัวแปร และชุดข้อมูลพลังงานในอดีตเป็นเวลาสองปี การแบ่งกลุ่มแบบเคมีนส์ที่ปรับปรุงแล้วนำไปใช้กับการจัดกลุ่มชุดข้อมูลพลังงานในอดีต และ grey relational analysis ใช้ในการวิเคราะห์วันที่มีความคล้ายคลึงกันที่เหมาะสมที่สุดของวันที่คาดการณ์ ส่วนและโครงข่ายประสาทเทียมของ Elman นำมาใช้สำหรับพัฒนาความสัมพันธ์แบบไม่เชิงเส้นระหว่างปัจจัยอุตุนิยมวิทยาหลายตัวแปร และข้อมูลกำลังไฟฟ้า

สามารถแสดงความสัมพันธ์ของงานวิจัยที่เกี่ยวข้องได้ดังตารางที่ 2.2. ดังนี้

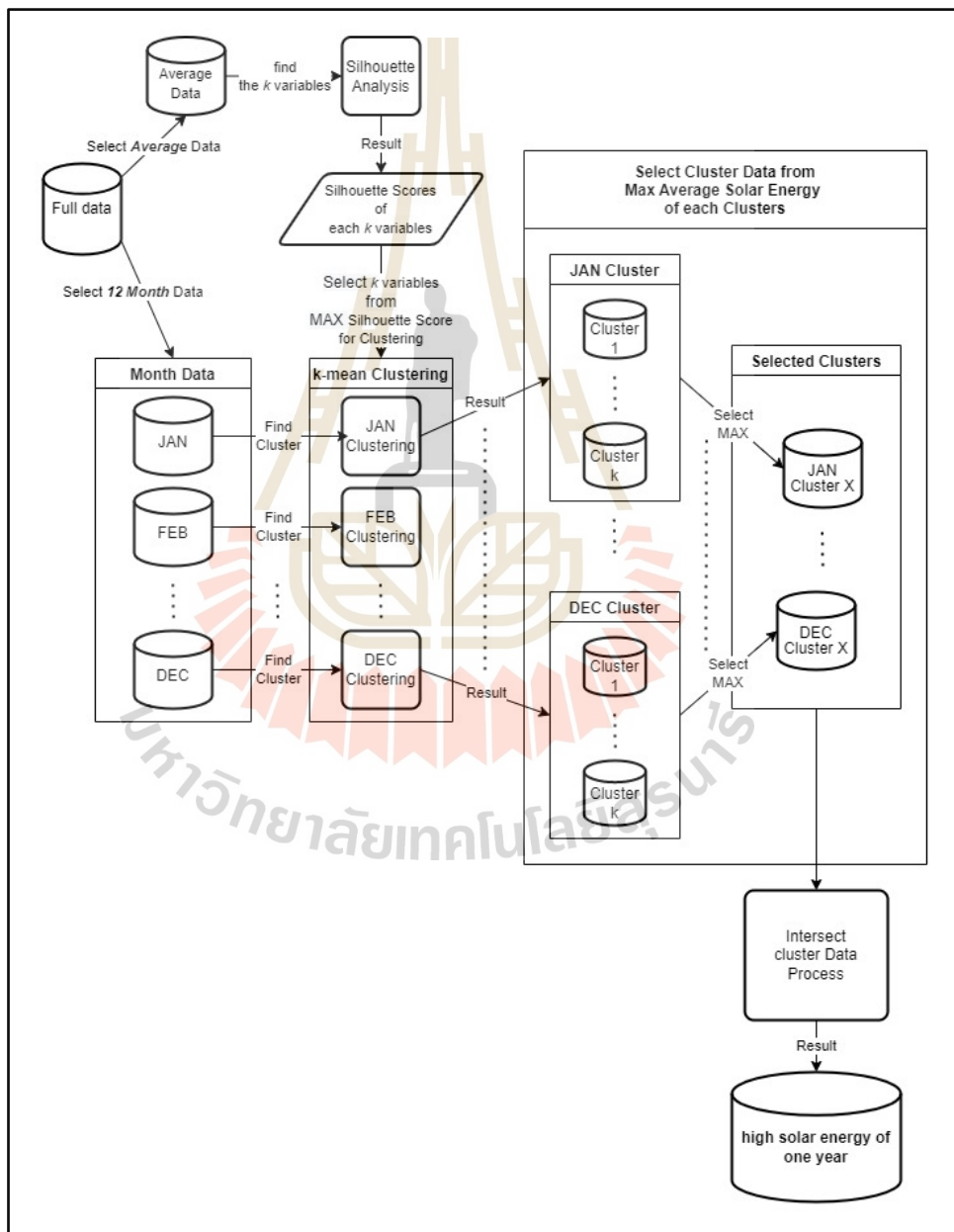
- เมื่อ
- a เป็นงานวิจัยของ Shuai Li, Hongjie Ma, and Weiyi Li (2017)
 - b เป็นงานวิจัยของ Hengxu Zhang, Yongji Cao, et al., (2018)
 - c เป็นงานวิจัยของ Khalil Benmouiz, Ali Cheknane (2013)
 - d เป็นงานวิจัยของ Nuño Basurto, Ángel Arroyo, Rafael Vega, et al. (2019)
 - e เป็นงานวิจัยของ Masoud Masoumi (2021)
 - f เป็นงานวิจัยของ Pedro Chévez, Dante Barbero, et al. (2017)
 - g เป็นงานวิจัยของ Peijie Lin, Zhouning Peng, et al. (2018)
 - h เป็นงานวิจัยของวิทยานิพนธ์นี้

ตารางที่ 2.2. สรุปความสัมพันธ์ของงานวิจัยที่เกี่ยวข้องในรูปแบบกะทัดรัด

กระบวนการทำงาน	งานวิจัยที่เกี่ยวข้อง							
	a	b	c	d	e	f	g	h
1. ลักษณะข้อมูล								
ข้อมูลเกี่ยวกับพลังงานแสงอาทิตย์	✓	✓	✓	✓			✓	✓
ข้อมูลเกี่ยวกับพลังงานลม	✓	✓			✓			
ข้อมูลเกี่ยวกับพลังงานคลื่น					✓			
ข้อมูลเกี่ยวข้องอื่น ๆ	✓					✓		
2. กระบวนการทางการเรียนรู้ของเครื่องที่ใช้								
k-means clustering	✓	✓	✓	✓	✓	✓	✓	✓
Silhouette technique					✓			✓
discrete-time Markov chain	✓							
via virtual generation systems		✓						
Artificial Neural Networks			✓	✓				
Hybrid Intelligent System				✓				
Regression				✓				
grey relational analysis							✓	
3. ลักษณะการวิเคราะห์ และแสดงผลข้อมูล								
วิเคราะห์ด้วยแผนภาพกล่อง		✓		✓	✓	✓		✓
วิเคราะห์ด้วยรูปภาพ		✓		✓	✓	✓		✓
วิเคราะห์ด้วยกราฟ	✓	✓	✓	✓	✓	✓	✓	
วิเคราะห์ด้วยตาราง	✓	✓		✓	✓	✓	✓	✓

บทที่ 3 ขั้นตอนการวิจัย

3.1 แผนภาพขั้นตอนการดำเนินงานวิจัย



รูปที่ 3.1 แผนภาพขั้นตอนการดำเนินงานวิจัย

จากรูปที่ 3.1 ได้แสดงขั้นตอนของงานวิจัยนี้ทั้งหมดโดยสังเกตจากรูปภาพในส่วนของ Full data เป็นส่วนการทำงานเริ่มต้นจากข้อมูลต้นฉบับ แบ่งออกเป็นสองส่วนคือ ส่วนค่าพลังงานเฉลี่ย (Average Data) และส่วนข้อมูลเดือนทั้ง 12 (Month Data) จากนั้นส่วนพลังงานเฉลี่ยมาใช้ในการหาค่าตัวแปร k ด้วยการวิเคราะห์ภาพเงา จากนั้นในส่วนข้อมูลเดือนทั้ง 12 นำมาแยกส่วนในแต่ละเดือนในการทำคลัสเตอร์ โดยมีการนำค่าตัวแปร k ในขั้นตอนก่อนหน้ามาใช้งาน เพื่อให้ในแต่ละเดือนที่ถูกคลัสเตอร์ได้จำนวนกลุ่มออกมาเท่ากัน ต่อจากนี้เลือกกลุ่มที่มีค่าเฉลี่ยสูงสุดของแต่ละเดือน (Selected Cluster) ที่ได้มาอินเตอร์เซกตรวมข้อมูลค่าเฉลี่ยสูงสุดของแต่ละเดือนทั้ง 12 เดือน จากนั้นนำข้อมูลที่ได้ไปวิเคราะห์ โดยรายละเอียด และผลลัพธ์ที่เกี่ยวข้องกับแต่ละขั้นตอนได้ถูกแสดงภายในหัวข้อหลังจากนี้ตามลำดับ

3.2 แหล่งที่มาของข้อมูล และลักษณะของข้อมูล

ข้อมูล ฐานข้อมูลความเข้มรังสีแสงอาทิตย์ ระดับตำบล ปี พ.ศ. 2560 เป็นการเก็บข้อมูลพลังงานแสงอาทิตย์โดย กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน กระทรวงพลังงาน อับโหลด โดย นัยนา บุญนาค ในปี พ.ศ. 2563 บนเว็บไซต์ data.go.th โดยสำนักงานพัฒนารัฐบาลดิจิทัล (องค์การมหาชน) (สพร.) โดยแสดงตัวอย่างข้อมูลบางส่วนโดยแยกส่วนข้อมูลเพื่อให้เห็นอย่างชัดเจนเป็นสองส่วน โดยที่ทั้งสองส่วนนั้นเป็นข้อมูลที่เรียงต่อกันดังตารางที่ตารางที่ 3.1 และตารางที่ 3.2 ตามลำดับ และแสดงคำอธิบายลักษณะของข้อมูลในแต่ละฟิลล์ในตารางที่ 3.3 โดยฟิลล์จังหวัด, อำเภอ และตำบล คือฟิลล์ตำแหน่งของพื้นที่ ส่วน latitude และ Longitude คือค่าที่อยู่ของพื้นที่ตำบลดังกล่าวในโลก โดยในแต่ละพื้นที่มีค่าพลังงานแสงอาทิตย์ในแต่ละเดือน และค่าเฉลี่ยพลังงานแสงอาทิตย์ในหนึ่งปี โดยที่หน่วยของค่าพลังงานแสงอาทิตย์ของข้อมูลนี้เป็น MJ/m²-day (เมกะจูลต่อตารางเมตรต่อวัน)

ตารางที่ 3.1 ตัวอย่างลักษณะของข้อมูลความเข้มรังสีแสงอาทิตย์ ส่วนพื้นที่สำรวจ

จังหวัด	อำเภอ	ตำบล	latitude	Longitude
นราธิวาส	เมืองนราธิวาส	โคกเคียน	6.49	101.76
อุดรดิตถ์	พิชัย	พญาแมน	17.19	100.02
มุกดาหาร	ดอนตาล	โพธิ์ไทร	16.39	104.85
สกลนคร	เมืองสกลนคร	จู้ด่อน	17.11	104.16

ตารางที่ 3.2 ตัวอย่างลักษณะของข้อมูลความเข้มรังสีแสงอาทิตย์ ส่วนค่าพลังงานเฉลี่ย

January	February	March	April	May	June	July	August	September	October	November	December	Average
15.7	19.2	20.0	20.8	19.7	19.0	18.8	19.3	19.0	17.5	14.5	13.6	18.1
16.5	18.9	20.1	22.0	20.2	18.4	17.0	16.5	16.9	17.5	16.7	15.7	18.0
16.9	18.7	19.9	20.8	19.8	18.5	17.1	16.1	16.9	18.0	16.8	16.1	18.0
16.6	18.4	19.6	20.8	19.4	17.9	16.8	16.6	17.5	18.4	16.9	16.1	17.9

ตารางที่ 3.3 ชื่อ, คำอธิบาย และหน่วยของแต่ละฟิลล์ ในฐานข้อมูลความเข้มรังสีแสงอาทิตย์

ฟิลล์	คำอธิบาย	หน่วย
	ชื่อจังหวัดในประเทศไทย	-
อำเภอ	ชื่ออำเภอของแต่ละจังหวัดในประเทศไทย	-
ตำบล	ชื่อตำบลของแต่ละอำเภอในประเทศไทย	-
latitude	ค่าละติจูดของตำบล	-
Longitude	ค่าลองจิจูดของตำบล	-
January	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนมกราคม	MJ/m ² -day
February	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนกุมภาพันธ์	MJ/m ² -day
March	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนมีนาคม	MJ/m ² -day
April	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนเมษายน	MJ/m ² -day
May	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนพฤษภาคม	MJ/m ² -day
June	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนมิถุนายน	MJ/m ² -day
July	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนกรกฎาคม	MJ/m ² -day
August	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนสิงหาคม	MJ/m ² -day
September	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนกันยายน	MJ/m ² -day
October	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนตุลาคม	MJ/m ² -day
November	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนพฤศจิกายน	MJ/m ² -day
December	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในเดือนธันวาคม	MJ/m ² -day
Average	ค่าเฉลี่ยพลังงานแสงอาทิตย์ในหนึ่งปี	MJ/m ² -day

3.3 การกำหนดค่าอินพุต

การกำหนดค่าอินพุตนั้นมีจุดประสงค์เพื่อให้ผลลัพธ์ที่ออกมา นั้น เพื่อให้สามารถวิเคราะห์ได้ง่ายขึ้น โดยสำหรับค่าฟิลล์ต่าง ๆ ที่ปรากฏคำอธิบายดังตารางที่ 3.3 สามารถกำหนดค่าต่าง ๆ ไว้ดังคำอธิบายที่แสดงในตารางที่ 3.4 ดังนี้

ตารางที่ 3.4 ตารางแสดงการกำหนดค่าอินพุตข้อมูล ในฐานข้อมูลความเข้มรังสีแสงอาทิตย์

ฟิลล์	การกำหนดค่า			คำอธิบายเพิ่มเติม
	อินพุต	คงค่า	นำออก	
จังหวัด		✓		สำหรับการวิเคราะห์ผล
อำเภอ		✓		สำหรับการวิเคราะห์ผล
ตำบล		✓		สำหรับการวิเคราะห์ผล
latitude			✓	ไม่มีความจำเป็นในการวิเคราะห์ผล
Longitude			✓	ไม่มีความจำเป็นในการวิเคราะห์ผล
January	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
February	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
March	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
April	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
May	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
June	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
July	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
August	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
September	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
October	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
November	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
December	✓			สำหรับการแบ่งกลุ่มแบบเคมีนส์
Average	✓			สำหรับการวิเคราะห์ภาพเงา

ตารางที่ 3.5 ตัวอย่างชุดของข้อมูลหลังจากแบ่งกลุ่ม ตัวอย่างกลุ่ม Average

จังหวัด	อำเภอ	ตำบล	Average
นราธิวาส	เมืองนราธิวาส	โคกเคียน	18.1
อุตรดิตถ์	พิชัย	พญาแมน	18.0
มุกดาหาร	ดอนตาล	โพธิ์ไทร	18.0
สกลนคร	เมืองสกลนคร	จิวด่อน	17.9
หนองบัวลำภู	โนนสัง	โคกใหญ่	17.9

ตารางที่ 3.6. ตัวอย่างชุดของข้อมูลหลังจากแบ่งกลุ่ม ตัวอย่างกลุ่ม January

จังหวัด	อำเภอ	ตำบล	January
นราธิวาส	เมืองนราธิวาส	โคกเคียน	15.7
อุตรดิตถ์	พิชัย	พญาแมน	16.5
มุกดาหาร	ดอนตาล	โพธิ์ไทร	16.9
สกลนคร	เมืองสกลนคร	จั่วต่อน	16.6
หนองบัวลำภู	โนนสัง	โคกใหญ่	16.5

การกำหนดค่าในรูปแบบอินพุตถูกใช้สำหรับฟิลล์ที่มีค่าค่าเฉลี่ยพลังงานแสงอาทิตย์ เพื่อนำฟิลล์ดังกล่าวมาใช้สำหรับการการแบ่งกลุ่มแบบเคมินส์ โดยมีฟิลล์ที่ถูกการกำหนดค่าแบบคงค่าติดตามไปด้วย โดยการกำหนดค่าแบบคงค่ามีไว้สำหรับแสดงผลลัพธ์ของการแบ่งกลุ่มแบบเคมินส์ที่มาจากฟิลล์ค่าเฉลี่ยพลังงานแสงอาทิตย์ในแต่ละเดือน นอกจากนี้ค่าฟิลล์ Average นั้นมีการใช้งานสำหรับเพียงขั้นตอนการหาตัวแปร k สำหรับการแบ่งกลุ่มแบบเคมินส์ ด้วยการวิเคราะห์ภาพเงาเท่านั้น ดังตารางที่ 3.5 และ 3.6 สามารถแสดงค่าตัวอย่างของชุดข้อมูลหลังจากแบ่งกลุ่มเพื่อสามารถสังเกตความแตกต่างได้ โดยรูปแบบของกลุ่มข้อมูลอื่น ๆ นั้นมีลักษณะเช่นเดียวกัน โดยมีความแตกต่างเพียงแค่ค่าของอินพุตที่แต่ละกลุ่มเท่านั้นดังนี้

3.4 การหาตัวแปร k สำหรับการแบ่งกลุ่มแบบเคมินส์ ด้วยการวิเคราะห์ภาพเงา

การหาค่าตัวแปร k ที่เหมาะสมสำหรับการแบ่งกลุ่มแบบเคมินส์ งานวิจัยนี้ใช้การวิเคราะห์ภาพเงา โดยหาจากชุดข้อมูล Average โดยค่าตัวแปร k ที่หาได้จากกระบวนการนี้มาจากการหาค่า silhouette score ของข้อมูลในแต่ละกลุ่มจำนวน k กลุ่มมาเฉลี่ยกัน จากนั้นเลือกค่าตัวแปร k ที่มีค่าเฉลี่ยของ silhouette score ที่มากที่สุด เพื่อใช้สำหรับกำหนดค่าคงที่ของ k ในการแบ่งกลุ่มของชุดกลุ่มข้อมูลอื่น ๆ

เหตุผลที่ใช้ชุดข้อมูล Average มาเป็นตัวแทนสำหรับกำหนดค่า k นั้นมาจากสมมติฐานที่ว่าค่าของชุดข้อมูล Average มาจากค่าเฉลี่ยของทั้ง 12 ชุดกลุ่มข้อมูลในเดือนต่าง ๆ ซึ่งถือได้ว่าเป็นค่าตัวแปร k กลางที่เหมาะสมสำหรับใช้ในชุดข้อมูลอื่นได้ด้วย เนื่องจากหากไม่ใช้ค่า k กลางนี้ แล้วหาค่า k ในทุกชุดข้อมูลแทน สำหรับงานวิจัยนี้คาดว่าไม่เหมาะสม เพราะว่าหากค่า k ที่แตกต่างกันย่อมส่งผลถึงจำนวนข้อมูลระหว่างการแบ่งกลุ่มของแต่ละชุดกลุ่มข้อมูลที่อาจมีความแตกต่างกันมากเกินไป ส่งผลให้ค่าผลลัพธ์สุดท้ายมีความเสี่ยงที่ไม่สามารถหาค่าตอบได้สูงกว่าการใช้ค่าตัวแปร k กลางที่หาได้จากขั้นตอนนี้ และสามารถลดเวลาในการหาค่าตัวแปร k ที่จะใช้ในแต่ละชุดข้อมูล

3.5 ขั้นตอนการแบ่งกลุ่มแบบเคมีนส์

จากหัวข้อที่ 3.4 นำค่า k ที่ได้จากระบวนการการวิเคราะห์ภาพเงา มาใช้สำหรับกำหนดค่า k ซึ่งเป็นจำนวนกลุ่มที่ใช้แบ่งค่าของแต่ละชุดข้อมูลออกจากกันตามจำนวน k กลุ่มที่กำหนด โดยสามารถแสดงลักษณะผลที่ได้ออกมาได้ดังตารางที่ 3.7 ดังนี้

ตารางที่ 3.7 ตัวอย่างลักษณะของข้อมูลหลังจากการแบ่งกลุ่มแบบเคมีนส์ ของกลุ่ม January

จังหวัด	อำเภอ	ตำบล	January	cluster
นราธิวาส	เมืองนราธิวาส	โคกเคียน	15.7	1
อุตรดิตถ์	พิชัย	พญาแมน	16.5	1
มุกดาหาร	ดอนตาล	โพธิ์ไทร	16.9	3
สกลนคร	เมืองสกลนคร	จิวัดอน	16.6	2
หนองบัวลำภู	โนนสัง	โคกใหญ่	16.5	1

จากตารางที่ 3.7 เมื่อทำการแบ่งกลุ่มแบบเคมีนส์ ในตารางเกิดฟิลล์ cluster ขึ้นมาเพื่อแสดงว่าข้อมูลดังกล่าวอยู่ในกลุ่มใด ตามตัวเลขของ cluster ที่ได้จากการแบ่งกลุ่มแบบเคมีนส์ จากนั้นหาค่าเฉลี่ยของค่าพลังงานแสงอาทิตย์ในแต่ละกลุ่ม cluster เพื่อหากลุ่มที่มีค่าเฉลี่ยสูงที่สุด โดยนับว่ากลุ่มข้อมูลที่ได้กลายเป็นกลุ่มพื้นที่ที่มีพลังงานแสงอาทิตย์สูงในชุดข้อมูลเดือนนั้น ๆ ส่วนกลุ่มข้อมูลที่เหลือในชุดข้อมูลให้เป็นกลุ่มพื้นที่ที่ไม่มีพลังงานแสงอาทิตย์สูง

3.6 การอินเตอร์เซคข้อมูลที่ได้จากการแบ่งกลุ่มแบบเคมีนส์

นำผลข้อมูลที่อยู่ในกลุ่มของค่าพลังงานแสงอาทิตย์สูงในชุดของกลุ่มข้อมูลที่แบ่งออกตามหัวข้อที่ 3.3 นั้นเข้ามาอินเตอร์เซคข้อมูลกันโดยใช้ค่าของฟิลล์จังหวัด, อำเภอ, และตำบลในการอินเตอร์เซค เฉพาะข้อมูลที่เหมือนกันในทุกชุดกลุ่มข้อมูลเท่านั้นจึงสามารถใช้เป็นผลการวิจัยได้ เพื่อใช้สำหรับในการวิเคราะห์ผลต่อไป โดยตัดฟิลล์ cluster ของทุกชุดกลุ่มของข้อมูลออก เหลือไว้เพียงค่าพลังงานในแต่ละเดือน และชื่อจังหวัด, อำเภอ, และตำบล ของแต่ละข้อมูลนั้น ๆ โดยแบ่งแสดงได้สองส่วนดังตารางที่ 3.8 และ ตารางที่ 3.9 ตามลำดับดังนี้

ตารางที่ 3.8 ตัวอย่างของการอินเตอร์เซคข้อมูลค่าพลังงานแสงอาทิตย์ ส่วนพื้นที่สำรวจ

จังหวัด	อำเภอ	ตำบล
นราธิวาส	เมืองนราธิวาส	โคกเคียน
อุตรดิตถ์	พิชัย	พญาแมน
มุกดาหาร	ดอนตาล	โพธิ์ไทร

ตารางที่ 3.9 ตัวอย่างของการอินเตอร์เซกข้อมูลค่าพลังงานแสงอาทิตย์ ส่วนค่าพลังงานเฉลี่ย

January	February	March	April	May	June	July	August	September	October	November	December
15.7	19.2	20.0	20.8	19.7	19.0	18.8	19.3	19.0	17.5	14.5	13.6
16.9	18.7	19.9	20.8	19.8	18.5	17.1	16.1	16.9	18.0	16.8	16.1
16.6	18.4	19.6	20.8	19.4	17.9	16.8	16.6	17.5	18.4	16.9	16.1

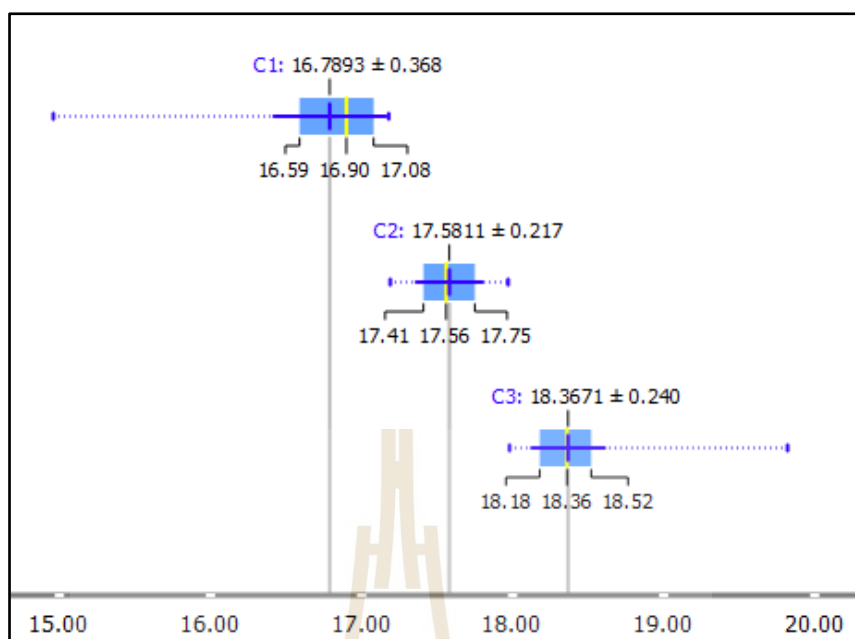
ข้อควรระวังสำหรับขั้นตอนนี้คือ หากผลลัพธ์ทุกฟิลล์ของทุกแถวมีค่าว่างเกิดขึ้นแม้เพียงหนึ่งจุด แสดงว่าการอินเตอร์เซกข้อมูลที่ได้จากการแบ่งกลุ่มแบบเคมินส์นี้เกิดความผิดพลาด หรือผลลัพธ์ที่ได้ออกมามีค่าน้อยมาก ๆ โดยสังเกตจากการทำกระบวนการนี้ซ้ำ ๆ ความผิดพลาดดังกล่าวนี้เกิดขึ้นได้โดย การกำหนดค่าในโปรแกรมผิดพลาด หรือเลือกกลุ่มพื้นที่ที่มีพลังงานแสงอาทิตย์สูงผิดพลาด เมื่อเกิดปัญหาให้แก้ไขการกำหนดค่าโปรแกรม และเลือกกลุ่มพื้นที่ที่มีพลังงานแสงอาทิตย์สูงที่ถูกต้อง

3.7 ขั้นตอนการวิเคราะห์ผล

การวิเคราะห์ผลในงานวิจัยนี้แบ่งออกเป็นหัวข้อการวิเคราะห์ต่าง ๆ ได้ดังนี้

3.7.1 การวิเคราะห์การแบ่งกลุ่มแบบเคมินส์ในแต่ละชุดกลุ่มข้อมูล

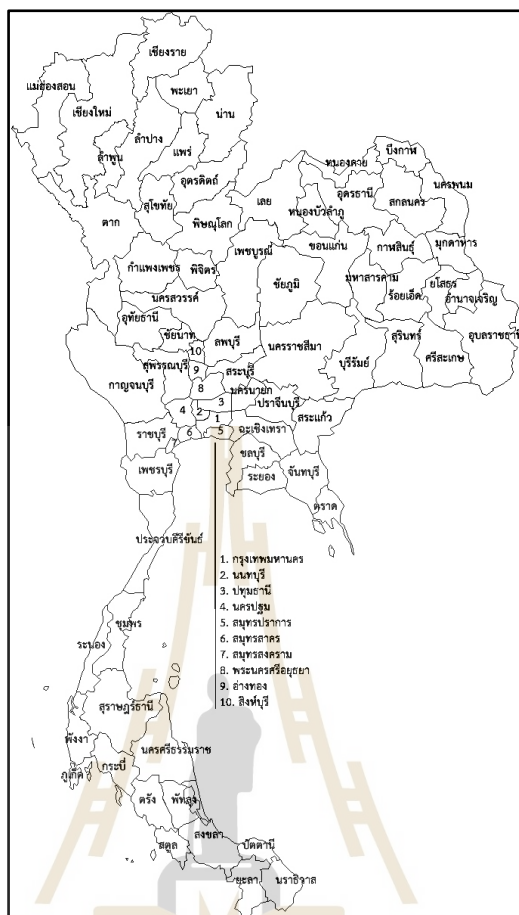
การวิเคราะห์ในส่วนนี้มีการใช้แผนภาพกล่องในการช่วยวิเคราะห์ค่าของการแบ่งกลุ่มของแต่ละชุดข้อมูล เพื่อแสดงให้เห็นถึงลักษณะของการกระจายของข้อมูลในแต่ละกลุ่มของแต่ละชุดข้อมูล เนื่องจากแผนภาพกล่องสามารถแสดงค่าต่าง ๆ ให้เห็นจากกราฟแท่งที่ปรากฏ โดยสามารถช่วยให้วิเคราะห์ผลลัพธ์ว่าในชุดกลุ่มข้อมูลในแต่ละเดือนหลังจากผ่านกระบวนการแบ่งกลุ่มแบบเคมินส์ข้อมูลอยู่ในลักษณะใด ดังตัวอย่างรูปที่ 3.2



รูปที่ 3.2 ตัวอย่างผลลัพธ์ที่ได้หลังจากการแบ่งกลุ่มแบบเคมินส์ ของชุดข้อมูลกลุ่มใด ๆ

3.7.2 การวิเคราะห์ผลจากการนำข้อมูลมาระบายสีบนรูปแผนที่จำลอง

การวิเคราะห์ส่วนนี้ใช้หลังจากการรวมข้อมูลที่ได้จากการแบ่งกลุ่มแบบเคมินส์ เหตุผลที่ระบายสีบนรูปแผนที่จำลองเนื่องจาก การสังเกตผลจากตารางโดยตรงนั้นอาจสร้างความ สับสนให้กับการวิเคราะห์อื่นเนื่องมาจากข้อมูลอาจมีมาก และป้องกันความไม่ชัดเจนในตำแหน่งพื้นที่ ที่มีพลังงานแสงอาทิตย์สูง สามารถเพิ่มประสิทธิภาพในการวิเคราะห์ และแสดงผล โดยการระบายสี ลงไปในแผนที่จำลองให้ตรงกับชื่อจังหวัดที่ได้ผลจากการรวมข้อมูลในทุกชุดกลุ่มข้อมูล โดยมีการไล่ ระดับข้อสีเพื่อแสดงให้เห็นว่าพื้นที่จังหวัดดังกล่าวได้จำนวนพื้นที่พลังงานแสงอาทิตย์สูงเป็นจำนวน มากแค่ไหน เมื่อสีแดงคือจำนวนพื้นที่ที่เป็นผลลัพธ์มีมาก และสีเหลืองคือจำนวนพื้นที่ที่เป็นผลลัพธ์มี น้อย โดยมีจุดประสงค์เพื่อวิเคราะห์ในมุมมองกว้าง จากการระบายสีในแผนที่จำลอง สามารถแสดงให้เห็น ถึงพื้นที่ที่ผลลัพธ์แสดงออกมาว่าหาเทียบลงไปแผนที่แล้ว ผลที่ได้มีลักษณะเป็นแบบใด โดยแสดง แผนที่จำลองที่ใช้ในงานวิจัยนี้ดังรูปที่ 3.3



รูปที่ 3.3 แผนที่จำลองที่ใช้สำหรับงานวิจัยนี้

3.7.3 การเปรียบเทียบผลลัพธ์ระหว่างงานวิจัย และการคลัสเตอร์แบบปกติ

ในงานวิจัยนี้ใช้การเปรียบเทียบผลลัพธ์ระหว่างงานวิจัย และการคลัสเตอร์แบบปกติ โดยการคลัสเตอร์แบบปกติหมายถึง การนำข้อมูลไปใช้สำหรับการแบ่งกลุ่มแบบเคมีนส์ โดยไม่ผ่านกระบวนการต่าง ๆ ที่งานวิจัยนี้ใช้ และเป็นการแบ่งกลุ่มแบบเคมีนส์จากข้อมูลโดยตรง นอกจากนี้ให้มีการเปรียบเทียบกับผลลัพธ์ก่อนทำการวิจัยด้วย โดยการเปรียบเทียบนี้สามารถแสดงให้เห็นว่าผลลัพธ์ของงานวิจัยนี้เหมาะสมหรือไม่ หากได้ผลลัพธ์เป็นไปตามเกณฑ์อย่างน้อยสองในสามข้อ นับว่างานวิจัยนี้ประสบความสำเร็จ โดยมีเกณฑ์ดังนี้

1. จำนวนผลลัพธ์ หากจำนวนผลลัพธ์น้อยลงไปเมื่อเปรียบเทียบกับผลลัพธ์ก่อนทำการวิจัย และได้จำนวนผลลัพธ์น้อยกว่าการคลัสเตอร์แบบปกติ หมายความว่างานวิจัยนี้สามารถจำกัดพื้นที่ที่มีพลังงานแสงอาทิตย์สูงได้ดีกว่าการคลัสเตอร์แบบปกติ

2. ค่าเฉลี่ยของพลังงานแสงอาทิตย์ หากได้ค่าที่สูงขึ้นเมื่อเปรียบเทียบกับผลลัพธ์ก่อนทำการวิจัย และได้จำนวนผลลัพธ์มากกว่าการคลัสเตอร์แบบปกติ หมายความว่างานวิจัยนี้สามารถจำกัดพื้นที่ที่มีพลังงานแสงอาทิตย์สูงของงานวิจัยนี้มีประสิทธิภาพที่เหมาะสม

3. ค่าเบี่ยงเบนมาตรฐานของพลังงานแสงอาทิตย์ หากได้ค่าที่ลดลงเมื่อเปรียบเทียบกับผลลัพธ์ก่อนทำการวิจัย และได้จำนวนผลลัพธ์มากกว่าการคลัสเตอร์แบบปกติ หมายความว่าค่าพลังงานแสงอาทิตย์สูงของงานวิจัยนี้ ในช่วงเวลาทั้ง 12 เดือนมีค่าพลังงานแสงอาทิตย์ที่สม่ำเสมออยู่ แม้ถูกจำกัดพื้นที่ให้น้อยลง



บทที่ 4 ผลการวิจัยและอภิปรายผล

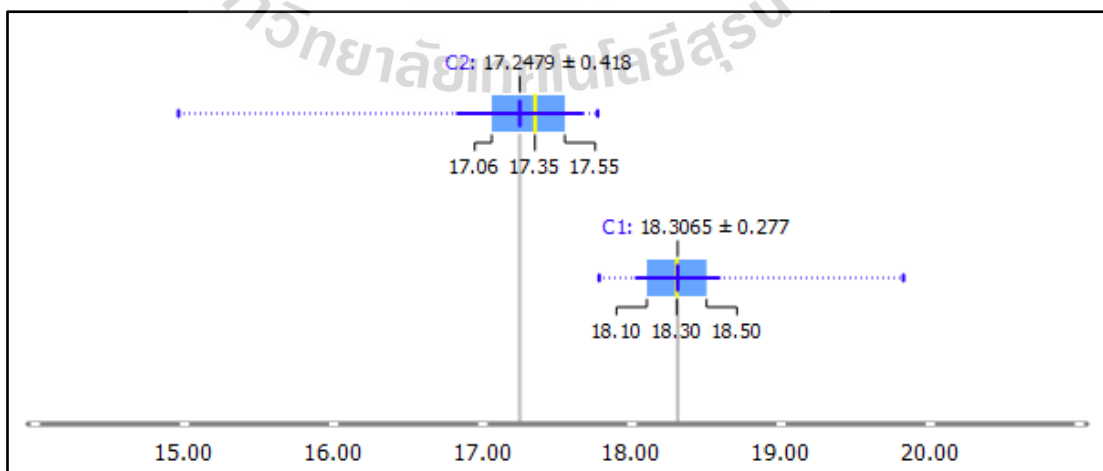
4.1 ผลลัพธ์การหาค่าตัวแปร k

การหาค่าตัวแปร k เป็นกระบวนการเริ่มต้นของงานวิจัยนี้ ดังแสดงในตารางที่ 4.1. ได้แสดงให้เห็นถึงค่าตัวแปร k ที่เหมาะสมกับการนำมาใช้ในงานวิจัยนี้ และรูปที่ 4.1. แสดงแผนภาพกล่องที่ใช้ค่า k ที่เหมาะสมที่สุดจากตารางที่ 4.1 เมื่อนำไปแบ่งกลุ่มแบบเคมีนส์

จากรูปที่ 4.1 สามารถบอกได้ถึงลักษณะของข้อมูลเมื่อนำมาแสดงในรูปแบบแผนภาพกล่องที่สามารถพบได้ในอนาคตของทุกชุดข้อมูลนี้ เนื่องจากทุกกลุ่มข้อมูลที่ถูกแบ่งออกมานั้นมีค่าเฉลี่ยพลังงานนี้เป็นจุดรวมของข้อมูล จึงคาดการณ์ได้ว่าเมื่อทุกชุดกลุ่มข้อมูลจะมีลักษณะข้อมูลที่เป็นผลลัพธ์ออกมามีความคล้ายคลึงกับในตารางที่ 4.1

ตารางที่ 4.1 ค่าเฉลี่ยของ silhouette score ในแต่ละตัวแปร

k	Average silhouette score
2	0.633
3	0.583
4	0.548
5	0.545



รูปที่ 4.1 แผนภาพกล่องของค่าเฉลี่ยพลังงานแสงอาทิตย์ในค่าตัวแปร k ที่เหมาะสม

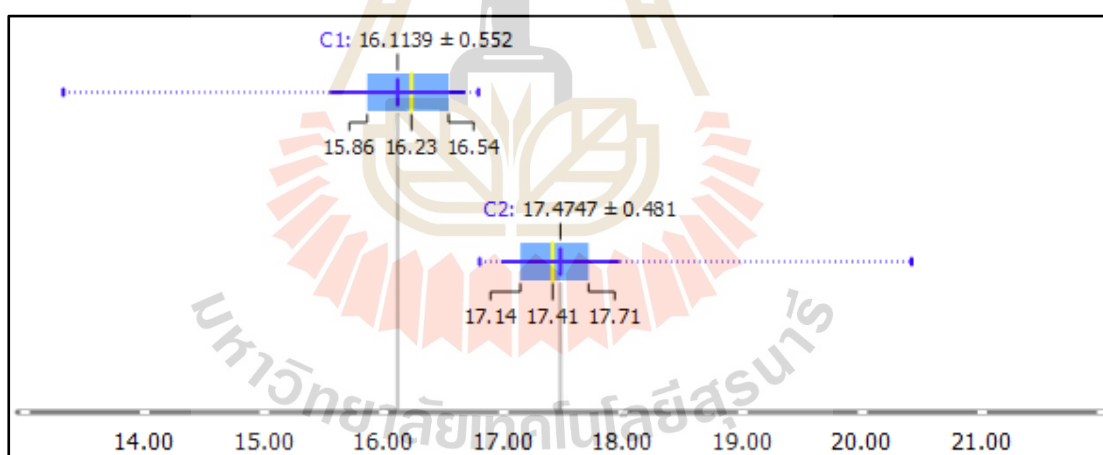
4.2 ผลลัพธ์การแบ่งกลุ่มแบบเคมินส์ในแต่ละชุดกลุ่มข้อมูล

4.2.1 ชุดกลุ่มข้อมูล มกราคม

1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Positive Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 16.8/20.5
4. ค่า Lower Quartile/Median/Upper Quartile: 17.14/17.41/17.71
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.4747/0.481

4.2.2 ชุดกลุ่มข้อมูล กุมภาพันธ์

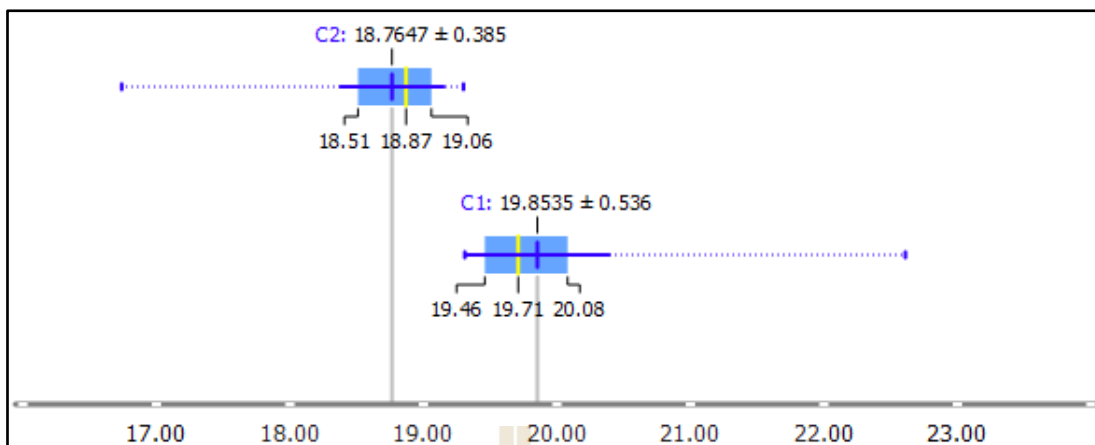
1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Positive Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 19.3/22.6
4. ค่า Lower Quartile/Median/Upper Quartile: 19.46/19.71/20.08
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 19.85/0.54



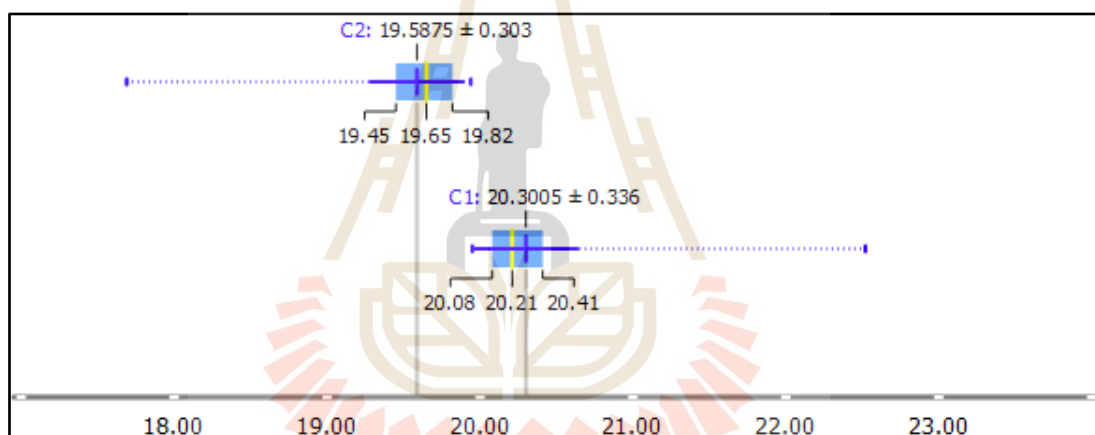
รูปที่ 4.2 แผนภาพกล่องของชุดกลุ่มข้อมูล มกราคม

4.2.3 ชุดกลุ่มข้อมูล มีนาคม

1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Positive Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 19.7/22.5
4. ค่า Lower Quartile/Median/Upper Quartile: 20.08/20.21/20.41
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 20.30/0.34



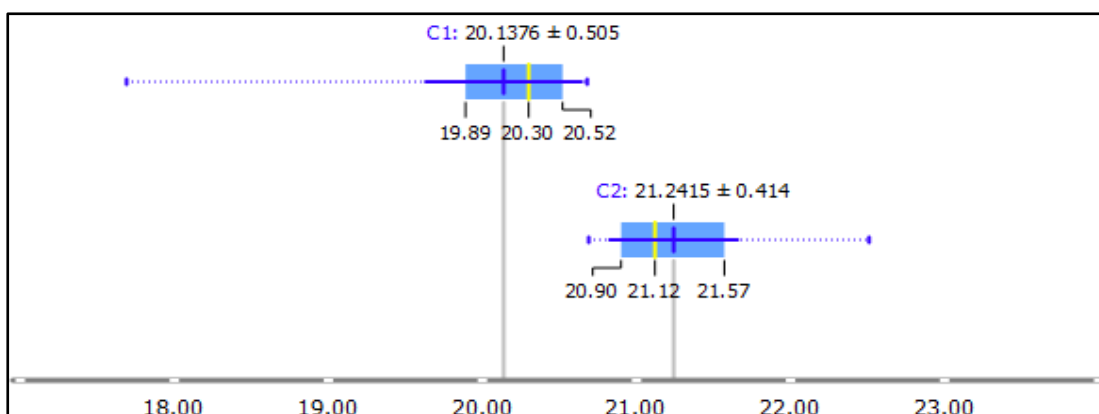
รูปที่ 4.3 แผนภาพกล่องของชุดกลุ่มข้อมูล กุมภาพันธุ์



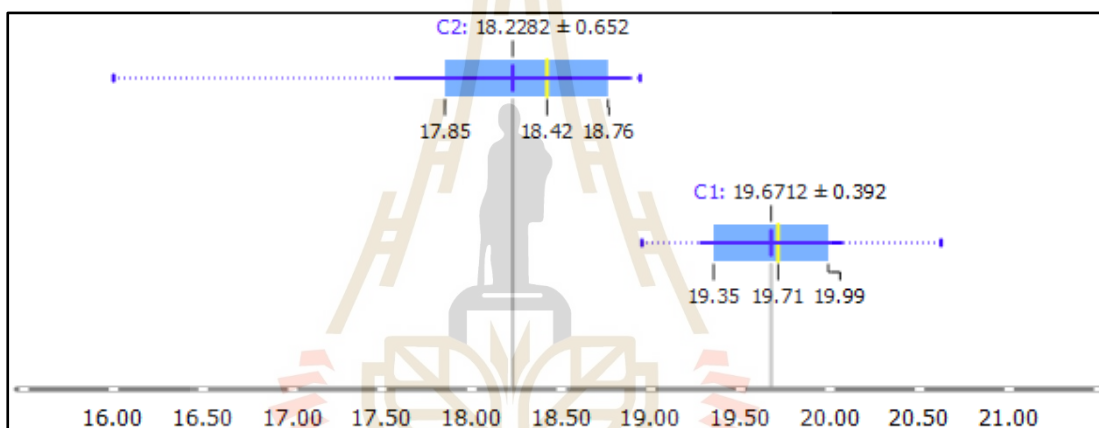
รูปที่ 4.4 แผนภาพกล่องของชุดกลุ่มข้อมูล มีนาคม

4.2.4 ชุดกลุ่มข้อมูล เมษายน

1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Positive Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 20.7/22.5
4. ค่า Lower Quartile/Median/Upper Quartile: 20.90/21.12/21.57
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 21.24/0.41



รูปที่ 4.5 แผนภาพกล่องของชุดกลุ่มข้อมูล เมษายน



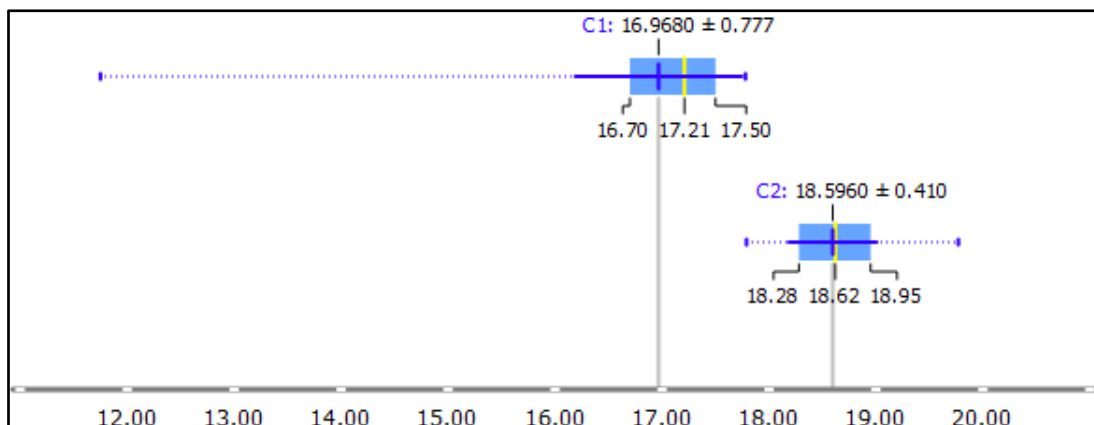
รูปที่ 4.6 แผนภาพกล่องของชุดกลุ่มข้อมูล พฤษภาคม

4.2.5 ชุดกลุ่มข้อมูล พฤษภาคม

1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 18.9/20.6
4. ค่า Lower Quartile/Median/Upper Quartile: 19.35/19.71/19.99
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 19.67/0.39

4.2.6 ชุดกลุ่มข้อมูล มิถุนายน

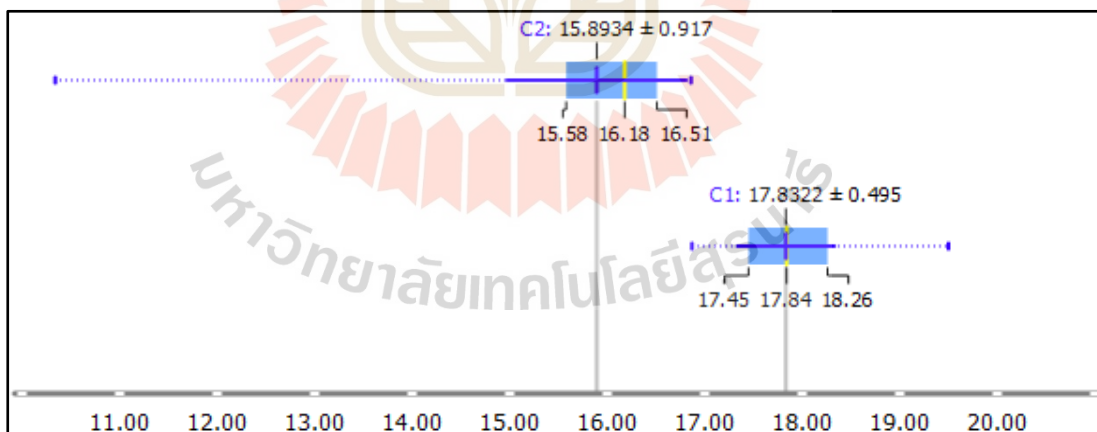
1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 17.8/19.9
4. ค่า Lower Quartile/Median/Upper Quartile: 18.28/18.62/18.95
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 18.60/0.41



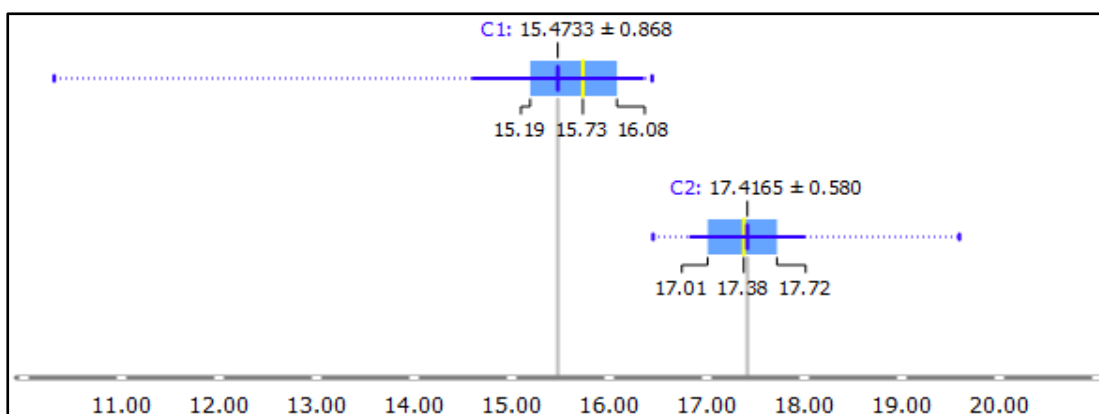
รูปที่ 4.7 แผนภาพกล่องของชุดกลุ่มข้อมูล มิถุนายน

4.2.7 ชุดกลุ่มข้อมูล กรกฎาคม

1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 16.9/19.7
4. ค่า Lower Quartile/Median/Upper Quartile: 17.45/17.84/18.26
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.83/0.50



รูปที่ 4.8 แผนภาพกล่องของชุดกลุ่มข้อมูล กรกฎาคม



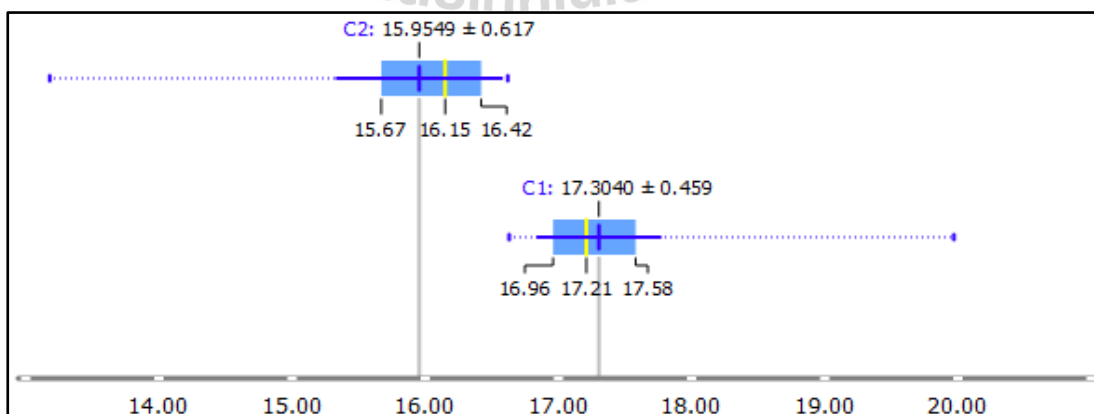
รูปที่ 4.9 แผนภาพกล่องของชุดกลุ่มข้อมูล สิงหาคม

4.2.8 ชุดกลุ่มข้อมูล สิงหาคม

1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 16.6/19.7
4. ค่า Lower Quartile/Median/Upper Quartile: 17.01/17.38/17.72
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.42/0.58

4.2.9 ชุดกลุ่มข้อมูล กันยายน

1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Positive Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 17.7/20.0
4. ค่า Lower Quartile/Median/Upper Quartile: 16.96/17.21/17.58
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.30/0.46



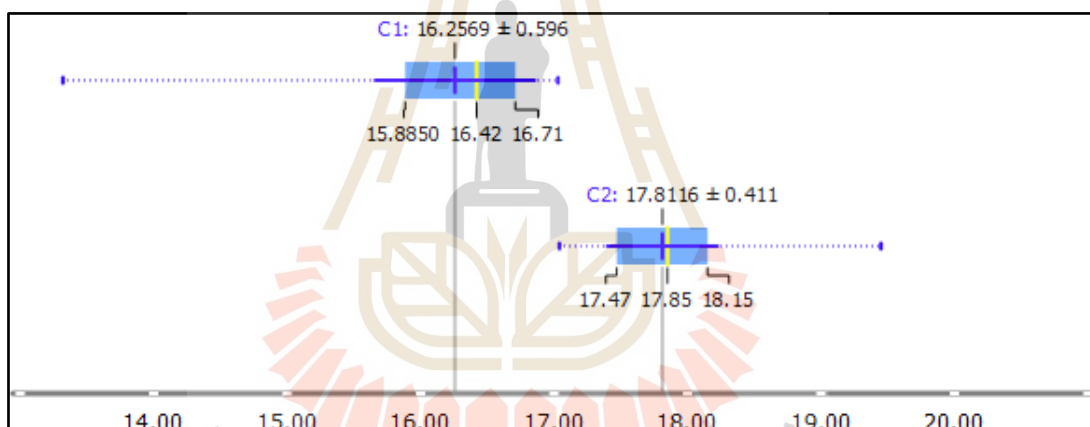
รูปที่ 4.10 แผนภาพกล่องของชุดกลุ่มข้อมูล กันยายน

4.2.10 ชุดกลุ่มข้อมูล ตุลาคม

1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 17.0/19.5
4. ค่า Lower Quartile/Median/Upper Quartile: 17.47/17.85/18.15
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.81/0.41

4.2.11 ชุดกลุ่มข้อมูล พฤศจิกายน

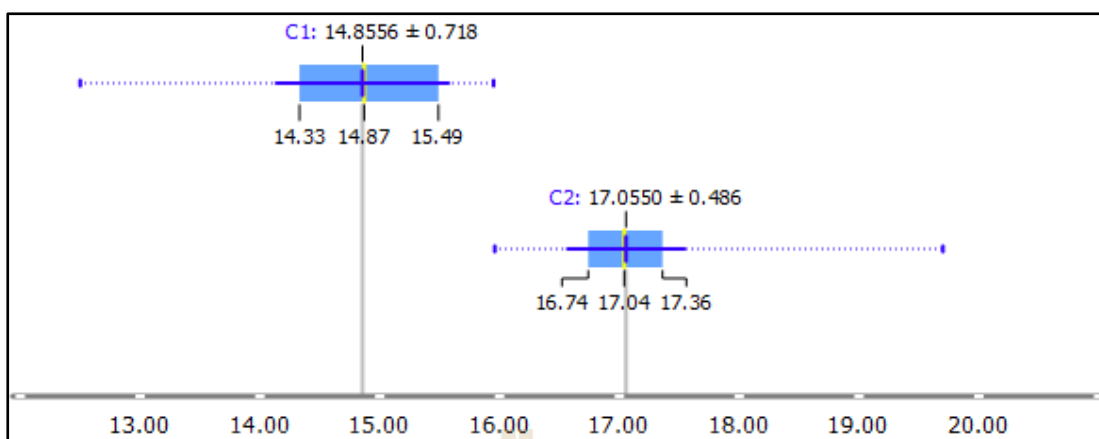
1. กลุ่มข้อมูลที่เลือก: C2
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 17.0/19.5
4. ค่า Lower Quartile/Median/Upper Quartile: 16.74/17.04/17.36
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 17.06/0.49



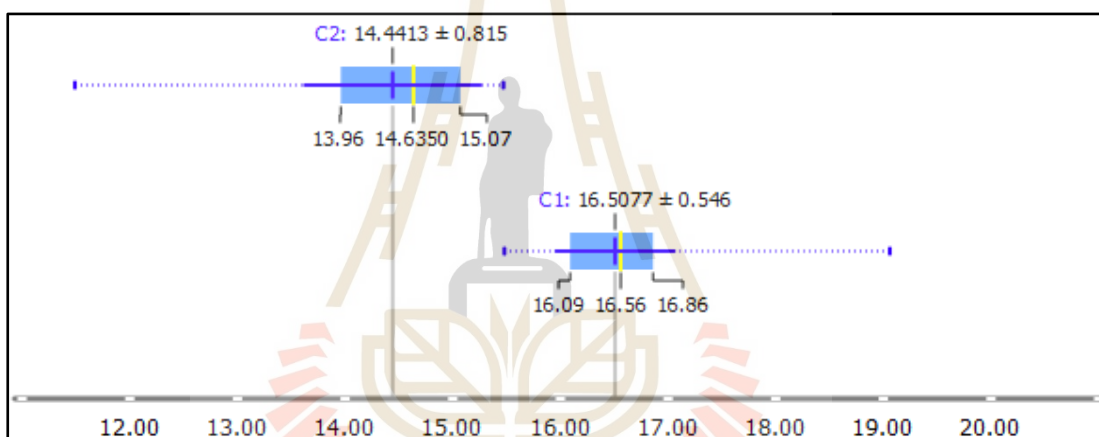
รูปที่ 4.11. แผนภาพกล่องของชุดกลุ่มข้อมูล ตุลาคม

4.2.12 ชุดกลุ่มข้อมูล ธันวาคม

1. กลุ่มข้อมูลที่เลือก: C1
2. ลักษณะการกระจายตัวของแผนภาพกล่อง: Normal Distribution
3. ค่าน้อยที่สุด/ค่ามากที่สุด ของกลุ่มข้อมูลโดยประมาณ: 15.6/19.2
4. ค่า Lower Quartile/Median/Upper Quartile: 16.09/16.56/16.86
5. ค่าเฉลี่ย/ค่าเบี่ยงเบนมาตรฐาน: 16.51/0.55



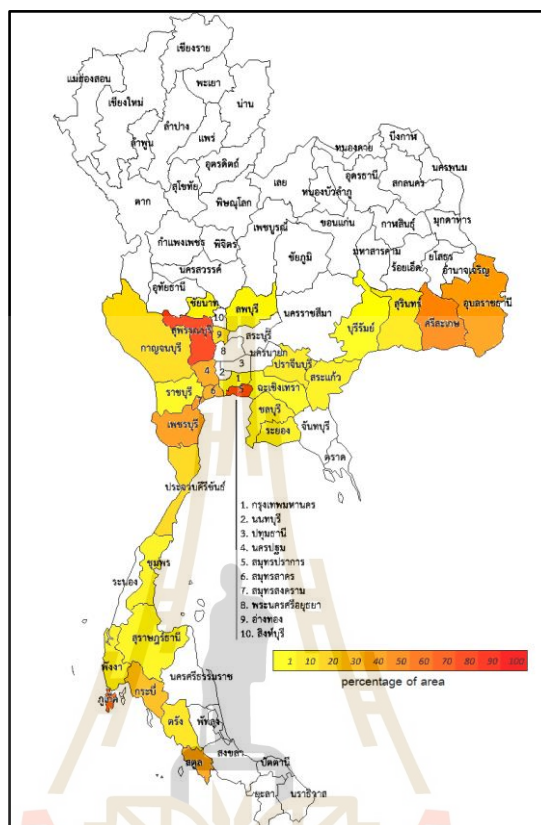
รูปที่ 4.12 แผนภาพกล่องของชุดกลุ่มข้อมูล พฤศจิกายน



รูปที่ 4.13 แผนภาพกล่องของชุดกลุ่มข้อมูล ธันวาคม

จากลักษณะของแผนภาพกล่องในตั้งแต่รูปที่ 4.2 ถึงรูปที่ 4.13 มีลักษณะของชุดกลุ่มข้อมูลที่ใกล้เคียงกัน โดยในทุกชุดข้อมูลในแต่ละกลุ่มข้อมูล กลุ่มข้อมูลที่มีค่าเฉลี่ยต่ำที่สุดของแต่ละชุดข้อมูลมี Whiskers ในส่วน Minimum Score มีลักษณะที่ยาวมากกว่า Whiskers ในด้านตรงกันข้าม และในกลุ่มข้อมูลที่มีค่าเฉลี่ยสูงที่สุดของแต่ละชุดข้อมูลมี Whiskers ในส่วน Maximum Score ในลักษณะที่ยาวมากกว่า Whiskers ในด้านตรงกันข้าม และเป็นเช่นนี้ในหลายกลุ่มชุดข้อมูล และการกระจายตัวของแผนภาพกล่องในกลุ่มที่เลือกของชุดกลุ่มข้อมูลมกราคม ถึงชุดกลุ่มข้อมูลเมษายน เป็นแบบ Positive Distribution ส่วนชุดกลุ่มข้อมูลอื่น ๆ มีการกระจายตัวของแผนภาพกล่องเป็นแบบ Normal Distribution เป็นส่วนมาก ซึ่งลักษณะดังกล่าวนี้มีความคล้ายกับแผนภาพกล่องของค่าเฉลี่ยพลังงานแสงอาทิตย์ ในรูปที่ 4.1

4.3 วิเคราะห์ผลจากการนำข้อมูลมาระบายสีบนรูปแผนที่จำลอง



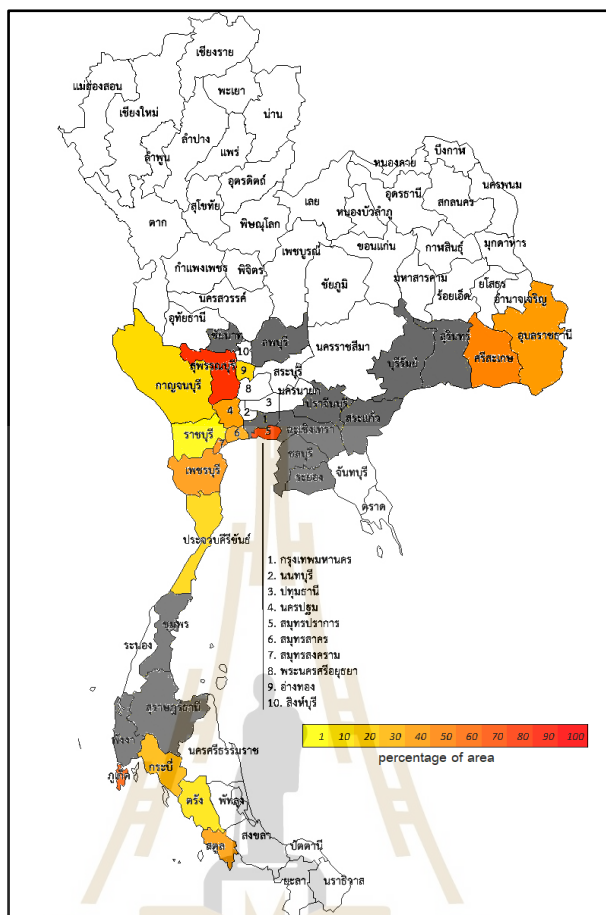
รูปที่ 4.14 แผนที่จำลองประเทศไทยที่แสดงผลลัพธ์ที่ได้จากงานวิจัย

ตารางที่ 4.2 แสดงผลลัพธ์ของของงานวิจัย โดยมีผลลัพธ์พื้นที่มากกว่า 10 เปอร์เซ็นต์

จังหวัด	จำนวนตำบลทั้งหมด	จำนวนตำบลผลลัพธ์	เปอร์เซ็นต์ของผลลัพธ์
สุพรรณบุรี	110	89	80.91
สมุทรปราการ	50	37	74.00
ภูเก็ต	17	11	64.71
ศรีสะเกษ	206	100	48.54
สมุทรสงคราม	36	17	47.22
เพชรบุรี	93	39	41.94
อุบลราชธานี	219	86	39.27
สตูล	36	14	38.89
นครปฐม	106	39	36.79
สมุทรสาคร	40	14	35.00
กระบี่	53	15	28.30
ประจวบคีรีขันธ์	48	8	16.67
อ่างทอง	73	11	15.07
กาญจนบุรี	98	14	14.29
ตรัง	87	9	10.34

จากรูปที่ 4.14 มีการลงสีในแผนที่จำลองของประเทศไทยโดยมีการไล่ระดับจำนวนพื้นที่ที่เป็นผลลัพธ์ของจังหวัดนั้น ๆ จากพื้นที่มากไปน้อย โดยมีการไล่ระดับสีจากจากสีแดงไปสู่สีเหลืองตามลำดับ โดยแสดงให้เห็นจำนวนผลลัพธ์ที่ได้ออกมาจากการนำกลุ่มข้อมูลที่เลือกจากทุกกลุ่มชุดข้อมูลมาเข้าสู่กระบวนการอินเทอร์เน็ตเซคเรียบร็อยแล้ว โดยจากรูปภาพนี้แสดงให้เห็นว่า ภาคตะวันออกเฉียงเหนือตอนใต้, ภาคตะวันออก, ภาคกลางตอนล่าง, ภาคตะวันตก, และภาคใต้ เป็นพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงที่สุด แต่จากผลลัพธ์ที่แสดงดังตารางที่ 4.2 และตารางที่ 4.3 โดยผลลัพธ์ดังกล่าวนี้สามารถนำไปวิเคราะห์ใหม่ได้ ซึ่งเป็นผลทำให้ผลลัพธ์ที่สามารถสังเกตได้จากการระบายสีบนแผนที่จำลองนั้นต่างไปจากเดิม

ดังตารางที่ 4.2 และตารางที่ 4.3 เป็นผลลัพธ์ของงานวิจัยทั้งสองตาราง แต่ผลลัพธ์นั้นมีทั้งแบบได้พื้นที่เปอร์เซ็นต์มากกว่า 10 เปอร์เซ็นต์ และพื้นที่เปอร์เซ็นต์ที่น้อยกว่า 10 เปอร์เซ็นต์ โดยจำนวนตำบลผลลัพธ์เป็นค่าจำนวนที่มาจากอินเทอร์เน็ตเซคพื้นที่ของจังหวัดนั้น ๆ โดยที่เปอร์เซ็นต์ของจำนวนผลลัพธ์เป็นผลลัพธ์ที่เปรียบเทียบจากจำนวนตำบลทั้งหมดในจังหวัด และจำนวนตำบลผลลัพธ์ ซึ่งในงานวิจัยเล่มนี้ได้พิจารณาว่าจังหวัดที่ได้พื้นที่ผลลัพธ์น้อยกว่า 10 เปอร์เซ็นต์นั้นเป็นผลลัพธ์แบบเดียวกับการไม่เป็นผลลัพธ์ เนื่องจากเป็นผลลัพธ์ที่งานวิจัยนี้พิจารณาว่าเป็นจำนวนผลลัพธ์ที่น้อยเกินไป ทำให้ผู้ที่นำข้อมูลนี้ไปใช้งานอาจไม่ได้พื้นที่ของจังหวัดที่มีค่าพลังงานแสงอาทิตย์สูง กล่าวอีกนัยนั้นแสดงว่ามีพื้นที่ที่จังหวัดนั้นมีพื้นที่พลังงานแสงอาทิตย์สูงน้อยเกินไปจึงนำออกจากการระบายสี แต่งานวิจัยนี้ยังคงผลลัพธ์ที่หาได้ไว้ดังเดิม แต่เสมือนว่าพื้นที่ผลลัพธ์จังหวัดนั้น ๆ ไม่เคยมีมาก่อน ดังแสดงในรูปที่ 4.15 ที่แสดงพื้นที่ตำบลของจังหวัดที่ถูกนำออกไปด้วยสีเทา ทำให้เห็นพื้นที่จังหวัดที่มีค่าพลังงานแสงอาทิตย์สูงที่สุดได้ชัดเจนมากขึ้น โดยจากรูปภาพนี้แสดงให้เห็นว่า ภาคใต้ตอนล่างฝั่งอันดามัน, ภาคตะวันตก, จังหวัดอุบลราชธานี, และจังหวัดศรีสะเกษ ยังคงเป็นผลลัพธ์ที่ได้ของพื้นที่ที่มีพลังงานแสงอาทิตย์สูงที่สุด หากนำรูปที่ 4.14 มาเทียบกับรูปที่ 4.15 พบว่ามีพื้นที่จังหวัดที่ได้จำนวนผลลัพธ์น้อย และถูกนำออกไปจากรูปภาพค่อนข้างมาก



รูปที่ 4.15 แผนที่จำลองประเทศไทยที่แสดงผลลัพธ์ที่ได้จากงานวิจัย เมื่อลบพื้นที่จังหวัดที่มีพื้นที่จากผลลัพธ์น้อย (เทา)

ตารางที่ 4.3 แสดงผลลัพธ์ของของงานวิจัย โดยมีผลลัพธ์พื้นที่น้อยกว่า 10 เปอร์เซ็นต์

จังหวัด	จำนวนตำบลทั้งหมด	จำนวนตำบลผลลัพธ์	เปอร์เซ็นต์ของผลลัพธ์
สุรินทร์	159	14	8.81
ปราจีนบุรี	65	5	7.69
พังงา	48	3	6.25
ชลบุรี	92	5	5.43
ระยอง	58	3	5.17
สระแก้ว	59	3	5.08
ชุมพร	70	3	4.29
ชัยนาท	53	2	3.77
ลพบุรี	124	4	3.23
กรุงเทพมหานคร	160	5	3.13
สุราษฎร์ธานี	131	4	3.05
ราชบุรี	104	2	1.92
ฉะเชิงเทรา	93	1	1.08
บุรีรัมย์	189	1	0.53

4.4 วิเคราะห์ผลจากการนำข้อมูลมาเปรียบเทียบ

การเปรียบเทียบผลลัพธ์สามารถสังเกตได้จากการนำค่าเฉลี่ยผลลัพธ์ต่าง ๆ ที่ได้จากงานวิจัยนี้มาเปรียบเทียบกับข้อมูลค่าเฉลี่ยต่าง ๆ และจำนวนผลลัพธ์ก่อนทำการวิจัย และหลังทำการวิจัย โดยที่ในส่วนหลังการทำการวิจัยมีการแสดงผลของการนำข้อมูลที่คลัสเตอร์แบบปกติ โดยไม่มีการอินเตอร์เซค และทำการคลัสเตอร์ข้อมูลทั้งหมดครั้งเดียวมาเปรียบเทียบกัน และมีค่าชุดข้อมูลสองแบบ โดยแบบแรกเป็นชุดข้อมูลค่าเฉลี่ยพลังงานแสงอาทิตย์ แบบที่สองเป็นชุดข้อมูลค่าพลังงานแสงอาทิตย์ 12 เดือน โดยที่ทุกระบวนการเป็นการหาผลลัพธ์ค่าพลังงานแสงอาทิตย์ที่สูงที่สุด ซึ่งแสดงดังตารางที่ 4.4 ดังนี้

ส่วนแรกจากตารางที่ 4.4 แสดงให้เห็นที่ประเภทชุดข้อมูลค่าเฉลี่ยพลังงานแสงอาทิตย์ ซึ่งเป็นชุดข้อมูลที่ใช้ในการหาค่าตัวแปร k ที่ใช้ในงานวิจัยนี้ หากนำมาหาค่าพลังงานแสงอาทิตย์สูงสุดพบว่า ค่าเฉลี่ยพลังงาน และค่าต่ำที่สุดมีค่าที่มากขึ้น ค่าเบี่ยงเบนมาตรฐานลดลง ส่วนค่าสูงสุดคงเดิม แสดงให้เห็นว่าผลลัพธ์ของชุดข้อมูลค่าเฉลี่ยพลังงานแสงอาทิตย์ได้ผลลัพธ์เป็นไปตามที่การทดลองนี้ต้องการ

ตารางที่ 4.4 แสดงผลลัพธ์ของงานวิจัยในการหาค่าพื้นที่ที่มีพลังงานแสงอาทิตย์สูงสุด

ลำดับ	ประเภทชุดข้อมูล	กระบวนการ	พื้นที่ผลลัพธ์	ค่าผลลัพธ์			
				AVG	STD	MAX	MIN
ก่อนทำการวิจัย							
1	ค่าเฉลี่ยพลังงานแสงอาทิตย์	-	7 416	17.896	0.617	19.820	14.960
2	ค่าพลังงานแสงอาทิตย์ 12 เดือน	-	7 416	17.896	0.878	20.476	13.694
หลังทำการวิจัย							
3	ค่าเฉลี่ยพลังงานแสงอาทิตย์	คลัสเตอร์แบบปกติ	4 543	18.306	0.277	19.820	17.780
4	ค่าพลังงานแสงอาทิตย์ 12 เดือน	คลัสเตอร์แบบปกติ	4 307	18.326	0.500	20.476	16.499
5	ค่าพลังงานแสงอาทิตย์ 12 เดือน	งานวิจัยนี้	558	18.661	0.464	20.397	17.739

ส่วนที่สองในส่วนชุดข้อมูลค่าพลังงานแสงอาทิตย์ 12 เดือนในส่วนก่อนทำวิจัย และหลังการทำวิจัยของกระบวนการคลัสเตอร์แบบปกติเพื่อหาพื้นที่ที่ค่าพลังงานแสงอาทิตย์สูงสุดพบว่า ค่าเฉลี่ย และค่าต่ำที่สุดมีค่าที่มากขึ้น ค่าเบี่ยงเบนมาตรฐานลดลง ส่วนค่าสูงสุดคงเดิม แสดงว่าผลลัพธ์ของชุดข้อมูลค่าเฉลี่ยพลังงานแสงอาทิตย์ได้ผลลัพธ์เป็นไปตามที่การทดลองนี้ต้องการเช่นกัน กับผลลัพธ์ของชุดข้อมูลค่าเฉลี่ยพลังงานแสงอาทิตย์

ส่วนที่สามในส่วนชุดข้อมูลค่าพลังงานแสงอาทิตย์ 12 เดือนผลลัพธ์หลังทำการวิจัยของกระบวนการคลัสเตอร์แบบปกติ และวิธีการของงานวิจัยนี้พบว่า ค่าเฉลี่ยพลังงานแสงอาทิตย์ และค่าต่ำที่สุดมีค่าที่มากขึ้น ค่าเบี่ยงเบนมาตรฐานที่ลดลง แต่ค่าสูงสุดที่ควรเท่าเดิมกลับลดลงเป็นผลมาจากการอินเตอร์เซคข้อมูลที่อาจทำให้ค่าสูงที่สุดที่ควรได้ออกจากผลลัพธ์เนื่องมาจาก ข้อมูลพื้นที่ดังกล่าวไม่ตรงตามกฎการอินเตอร์เซคข้อมูล โดยข้อมูลที่สามารถเป็นผลลัพธ์ได้นั้นจำเป็นต้องมีอยู่ในทุกกลุ่ม

ชุดข้อมูลทั้ง 12 เดือน ทำให้ค่าสูงสุดต่ำกว่าค่าก่อนทำการวิจัย และต่ำกว่ากว่ากระบวนการคลัสเตอร์แบบปกติ

ส่วนสุดท้ายเมื่อนำจำนวนพื้นที่ผลลัพธ์ที่ได้จากกระบวนการของงานวิจัยนี้มาเปรียบเทียบกับพื้นที่ผลลัพธ์ก่อนทำการวิจัย และพื้นที่ผลลัพธ์หลังทำการวิจัยด้วยกระบวนการคลัสเตอร์แบบปกติ พบว่า จำนวนพื้นที่ผลลัพธ์ที่ได้จากกระบวนการของงานวิจัยนี้ได้น้อยกว่ากระบวนการคลัสเตอร์แบบปกติมาก แต่ค่าผลลัพธ์ต่าง ๆ กลับดีกว่า หรือใกล้เคียง แสดงให้เห็นถึงความสามารถของกระบวนการในงานวิจัยนี้ในการหาพื้นที่ที่มีค่าพลังงานแสงอาทิตย์สูงสุดที่สามารถคัดเลือกพื้นที่ที่เหมาะสมที่สุดออกมาได้ หากนำไปใช้งานจริงจำนวนพื้นที่ผลลัพธ์ที่มากกว่าสามารถส่งผลเสียต่อการตัดสินใจในการเลือกพื้นที่ที่มีพลังงานแสงอาทิตย์สูงอย่างแน่นอน

4.5 สรุปการอภิปรายผล

ส่วนแรกในการหาค่า k ที่เหมาะสมสามารถวิเคราะห์ได้จากตารางที่ 4.1 พบว่า ค่าเฉลี่ยของ silhouette score ในแต่ละตัวแปร k เป็นค่าที่ผกผันกับตัวแปร k ที่เพิ่มขึ้น แสดงให้เห็นว่าจำนวนกลุ่มของงานวิจัยนี้อาจมีประสิทธิภาพ และความน่าเชื่อถือลดลงหากใช้ค่าตัวแปร k ที่มาก และไม่เหมาะสม ทั้งนี้ผลดังกล่าวอาจไม่เป็นไปตามที่กล่าวข้างต้นหากเปลี่ยนชุดข้อมูล

ส่วนที่สองผลลัพธ์จากการคลัสเตอร์ในแต่ละชุดข้อมูลพลังงานแสงอาทิตย์ทั้ง 12 เดือนดังแสดงดังรูปที่ 4.2 ถึง 4.13 พบว่าผลลัพธ์ที่ออกมาจากงานวิจัยนี้ทั้ง 12 ชุดมีความคล้ายคลึงกันมาก โดยเฉพาะในกลุ่มที่มีค่าเฉลี่ยพลังงานแสงอาทิตย์ที่สูง และสามารถสังเกตค่าเฉลี่ยของผลลัพธ์ต่าง ๆ ของชุดข้อมูลพลังงานแสงอาทิตย์ทั้ง 12 เดือนได้ที่ตาราง 4.4 ในลำดับที่ 5

ส่วนที่สามผลลัพธ์จากการอินเตอร์เซกกลุ่มข้อมูลที่มีค่าเฉลี่ยพลังงานแสงอาทิตย์สูงสุดจากการคลัสเตอร์ในแต่ละชุดข้อมูลพลังงานแสงอาทิตย์ทั้ง 12 เดือนพบว่า ได้ผลลัพธ์เป็นพื้นที่จังหวัดทั้งหมด 29 จังหวัดดังแสดงที่รูปที่ 4.14 และที่ตารางที่ 4.2 ถึง 4.3 แต่มีการระบายสีใหม่โดยระบายเฉพาะส่วนที่มีผลลัพธ์พื้นที่ต่ำบลมากกว่า 10 เปอร์เซ็นต์ของพื้นที่ทั้งหมดในจังหวัดนั้น ๆ ดังแสดงในรูปที่ 4.15 โดยมีพื้นที่จังหวัดที่เหลือในรูปภาพอยู่ 15 จังหวัด ซึ่งสอดคล้องกันกับตารางที่ 4.2 สาเหตุที่วิเคราะห์ผลรูปแบบพื้นที่จังหวัดนั้นเพื่อให้สามารถสังเกตพื้นที่ที่มีพลังงานแสงอาทิตย์สูงได้ง่ายกว่าการระบายสีที่ละพื้นที่ต่ำบล

ส่วนสุดท้ายการเปรียบเทียบผลลัพธ์ของงานวิจัยนี้สามารถสังเกตเห็นอย่างชัดเจนว่า จำนวนพื้นที่ผลลัพธ์จากการทดลองนี้สามารถทำการหาพื้นที่ได้น้อยกว่าจำนวนพื้นที่ผลลัพธ์ที่มาจากวิธีการคลัสเตอร์แบบปกติ โดยจำนวนพื้นที่ผลลัพธ์จากการทดลองนี้แตกต่างกันประมาณ 7.12 เท่าของพื้นที่ผลลัพธ์ที่มาจากวิธีการคลัสเตอร์แบบปกติ และเป็นพื้นที่ผลลัพธ์ 7.52 เปอร์เซ็นต์ของจำนวนพื้นที่ทั้งหมดโดยประมาณ ขณะที่พื้นที่ผลลัพธ์ที่มาจากวิธีการคลัสเตอร์แบบปกติได้จำนวนพื้นที่ผลลัพธ์เป็น 58.05 เปอร์เซ็นต์ของจำนวนพื้นที่ทั้งหมดโดยประมาณ

บทที่ 5

สรุปและข้อเสนอแนะ

5.1. สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เป็นการค้นหาพื้นที่ในประเทศไทยที่สามารถให้พลังงานแสงอาทิตย์ที่สูงได้ตลอดทั้งปี จึงมีการนำหลักการของการแบ่งข้อมูลแบบเคมีนส์ และการวิเคราะห์ภาพเงา มาใช้เพื่อค้นหาผลลัพธ์ที่ต้องการ โดยสามารถแสดงรายละเอียดได้ดังนี้

1) การแบ่งข้อมูล และกำหนดค่าอินพุตสำหรับการหาพื้นที่ในประเทศไทยที่สามารถให้พลังงานแสงอาทิตย์ที่สูงได้ตลอดทั้งปี เนื่องจากในข้อมูลที่ถูกแบ่งออกในแต่ละส่วนจะถูกใช้งานในขั้นตอนที่แตกต่างกัน และการกำหนดค่าอินพุตแต่ละส่วนนั้นจะใช้สำหรับการวิเคราะห์ที่แตกต่างกัน

2) การวิเคราะห์ภาพเงาสามารถทำได้โดยการใช้ค่าเฉลี่ยของค่าพลังงานแสงอาทิตย์ทุกพื้นที่ทั้ง 12 เดือน มาเป็นอินพุต โดยสาเหตุที่ไม่ใช้การวิเคราะห์ภาพเงาในทุก ๆ การแบ่งข้อมูลนั้นสามารถส่งผลกระทบต่อจำนวนกลุ่มผลลัพธ์ที่ไม่เหมือนกันในแต่ละกลุ่มข้อมูล และสามารถทำให้ค่าของจำนวนข้อมูลในแต่ละกลุ่มข้อมูลมีความแตกต่างกันมากเกินไป นั้นส่งผลให้ผลลัพธ์มีโอกาสที่จะไม่แสดงผล แต่การใช้ค่าเฉลี่ยของค่าพลังงานแสงอาทิตย์ทุกพื้นที่ทั้ง 12 เดือน มาร่วมกับการวิเคราะห์ภาพเงา สามารถลดโอกาสจะไม่แสดงผลได้ และเนื่องจากทุกกลุ่มข้อมูลมีค่าตัวแปร k ที่เหมือนกัน สามารถทำให้ค่าของจำนวนข้อมูลในแต่ละกลุ่มข้อมูลมีความแตกต่างกันน้อยลงได้ด้วย โดยค่าผลลัพธ์นี้แสดงดังบทที่ 4 หัวข้อที่ 4.1 โดยผลลัพธ์คือค่า k เท่ากับ 2

3) การนำข้อมูลที่ถูกแบ่งออกมาเข้ากระบวนการของการแบ่งกลุ่มแบบเคมีนส์ โดยข้อมูลที่ถูกแบ่งออกมามีจำนวนทั้งหมด 12 กลุ่ม ตามจำนวนเดือนในหนึ่งปี โดยใช้ค่า k ที่ได้จากการวิเคราะห์ภาพเงามาแบ่งกลุ่มข้อมูลในแต่ละกลุ่มของแต่ละเดือน แล้วทำการแยกกลุ่มข้อมูลที่ได้จากการแบ่งกลุ่มแบบเคมีนส์ในแต่ละกลุ่มของแต่ละเดือน โดยเลือกกลุ่มที่มีค่าเฉลี่ยของพลังงานแสงอาทิตย์ที่สูงที่สุดในแต่ละเดือน จากนั้นนำกลุ่มข้อมูลนั้นมาอินเตอร์เซกกันเพื่อเป็นผลลัพธ์

การวิเคราะห์ผลของงานวิจัยนี้มีการแสดงผลของแต่ละกลุ่มข้อมูลของแต่ละเดือนโดยใช้แผนภาพกล่อง เพื่อให้เห็นว่าในแต่ละกลุ่มที่ถูกแบ่งออกมาของแต่ละเดือนมีลักษณะของข้อมูลเป็นอย่างไรโดยแสดงดังรูปที่ 4.2 – 4.13 ตั้งแต่เดือนมกราคม ถึงเดือนธันวาคมตามลำดับ ผลลัพธ์ที่ออกมาทุกกลุ่มข้อมูลของแต่ละเดือนมีลักษณะที่คล้ายกันกับแผนภาพกล่องของค่าเฉลี่ยพลังงานแสงอาทิตย์

พื้นที่ที่พลังงานแสงอาทิตย์สูงที่สุดที่เป็นผลลัพธ์การวิจัยนี้ ได้ผลลัพธ์เป็นพื้นที่จังหวัดทั้งหมด 29 จังหวัด โดยเป็นจำนวนประมาณ 7.52 เปอร์เซ็นต์ของพื้นที่ตำบลทั้งหมดในประเทศไทยตามที่แสดงดังรูปที่ 4.14 และแสดงดังตารางที่ 4.2 – 4.3 แล้วผลลัพธ์ของจังหวัดที่มีพื้นที่ที่มีค่าพลังงาน

แสงอาทิตย์ที่สูงสุดโดยการลบจังหวัดที่ได้พื้นที่ผลลัพท์น้อยกว่า 10 เปอร์เซ็นต์ออกไป ได้ผลลัพท์ 15 จังหวัด และมีจำนวน 503 ตำบล โดยเป็นจำนวนประมาณ 6.78 เปอร์เซ็นต์ของพื้นที่ตำบลทั้งหมดในประเทศไทย โดยแสดงดังรูปที่ 4.15 และแสดงดังตารางที่ 4.2

จากกระบวนการของงานวิจัยนี้ เมื่อเปรียบเทียบกับการคลัสเตอร์แบบปกติพบว่า ผลลัพท์การหาพื้นที่ที่มีพลังงานแสงอาทิตย์สูงของงานวิจัยนี้สามารถจำกัดพื้นที่ที่มีพลังงานแสงอาทิตย์สูงได้ โดยได้ผลลัพท์ลดลงมาจากการคลัสเตอร์แบบปกติประมาณ 7.12 เท่าสำหรับข้อมูลในงานวิจัยนี้ โดยที่มีเฉลี่ยค่าพลังงานแสงอาทิตย์, ค่าสูงที่สุด และต่ำที่สุดที่เพิ่มขึ้น และค่าเบี่ยงเบนมาตรฐานที่ลดลง ซึ่งเป็นค่าที่มีความเหมาะสมมากขึ้นสำหรับงานวิจัยนี้ ดังแสดงในตารางที่ 4.4 ในลำดับที่ 4 และ 5

5.2. ข้อเสนอแนะ

ในส่วนแรกงานวิจัยนี้มีข้อมูลที่มีการใช้ข้อมูลพลังงานแสงอาทิตย์แค่ในระยะเวลาหนึ่งปีเท่านั้น นอกจากนี้ยังไม่มีการพัฒนาเครื่องมือสำหรับการแสดงผล, ในส่วนการแบ่งชุดข้อมูล, และการอินเตอร์เซกตรวมชุดข้อมูล โดยเป็นส่วนที่ยังต้องไข่มุขเป็นผู้กระทำกระบวนการเหล่านี้ และส่วนสุดท้ายการนำมาใช้งานที่ยังไม่มีผลการนำผลลัพท์ที่ได้จากการวิจัยนี้ไปใช้งานจริง และยังไม่มีการผู้เชี่ยวชาญที่สามารถระบุได้ว่าผลลัพท์ที่ออกมามีความถูกต้องหรือเหมาะสมมากน้อยอย่างไร งานวิจัยนี้คาดการณ์ว่าหากส่วนที่ขาดไปมีครบถ้วนสามารถทำให้งานวิจัยนี้มีผลลัพท์ที่เหมาะสมมากกว่านี้



เอกสารอ้างอิง

- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014a). การผลิตน้ำร้อนชนิดใช้ปั้มน้ำหมุนเวียน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=554&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014b). การผลิตน้ำร้อนชนิดผสมผสาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=553&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014c). การผลิตน้ำร้อนชนิดไหลเวียนตามธรรมชาติ. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=555&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014d). การอบแห้งระบบ Active. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=551&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014e). การอบแห้งระบบ Hybrid. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=550&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014f). การอบแห้งระบบ Passive. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=552&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014g). เซลล์แสงอาทิตย์แบบต่อกับระบบจำหน่าย (PV Grid connected system). กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=557&filename=index.
- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014h). เซลล์แสงอาทิตย์แบบผสมผสาน (PV Hybrid system). กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewt_news.php?nid=556&filename=index.

- กระทรวงพลังงาน. กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน. (2014i). เซลล์แสงอาทิตย์แบบอิสระ (PV stand alone system). กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน (พพ.) กระทรวงพลังงาน. Retrieved September 30, 2021, from https://www.dede.go.th/ewtadmin/ewt/dede_web/ewt_news.php?nid=558&filename=index.
- นัยนา บุญนาค. (2020, July 14). ข้อมูลข้อมูลความเข้มรังสีอาทิตย์ระดับตำบลสำหรับประเทศไทย - Open Government Data of Thailand. https://data.go.th/dataset/https-www-dede-go-th-more_news-php-cid-547-filename-index.
- พลังงานแสงอาทิตย์ กับการนำมาใช้ประโยชน์. (2020). Energy Next. Retrieved September 30, 2021, from <https://energynext.co.th/2020/06/พลังงานแสงอาทิตย์-กับกา/>.
- สมาคมพลังงานทดแทนสู่ชุมชนแห่งประเทศไทย [TRECA]. (ม.ป.ป.). พลังงานแสงอาทิตย์. TRECA. Retrieved September 30, 2021, from <http://reca.or.th/solar/>.
- สำนักงานพัฒนารัฐบาลดิจิทัล. (2015). open government data of Thailand. Retrieved September 30, 2021, from <https://data.go.th/>.
- Banpu NEXT Co., Ltd. (2020, July 20). "พลังงานแสงอาทิตย์" ทรัพยากรมากประโยชน์ที่ไม่มีวันหมด สู่ประโยชน์มหาศาลที่ช่วยโลกอย่างยั่งยืนด้วยการนำมาผลิตไฟฟ้า. Banpu NEXT. Retrieved September 30, 2021, from <https://www.banpunext.co.th/articles/พลังงานแสงอาทิตย์-ทรัพยากรมากประโยชน์ที่ไม่มีวันหมด/>.
- Basurto, N., Arroyo, Á., Vega, R., Quintián, H., Calvo-Rolle, J. L., & Herrero, Á. (2019). A hybrid intelligent system to forecast solar energy production. *Computers & Electrical Engineering*, 78, 373-387.
- Benmouiza, K., & Cheknane, A. (2013). Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. *Energy Conversion and Management*, 75, 561-569.
- Broom, D. (2019, March 22). How Japan became the world leader in Floating solar power. World Economic Forum. Retrieved September 30, 2021, from <https://www.weforum.org/agenda/2019/03/japan-is-the-world-leader-in-floating-solar-power/>.
- Chandler, D. L. (2011, October 26). Shining brightly. MIT News | Massachusetts Institute of Technology. Retrieved September 30, 2021, from <https://news.mit.edu/2011/energy-scale-part3-1026>.
- Chávez, P., Barbero, D., Martini, I., & Discoli, C. (2017). Application of the k-means clustering method for the detection and analysis of areas of homogeneous residential electricity consumption at the Great La Plata region, Buenos Aires, Argentina. *Sustainable cities and society*, 32, 115-129.

- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevár, T., Milutinovič, M., & Zupan, B. (2013). Orange: data mining toolbox in Python. *the Journal of machine Learning research*, 14(1), 2349-2353.
- Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y., (2002). A local search approximation algorithm for k-means clustering. In: *Proceedings of the Eighteenth Annual Symposium on Computational Geometry*, pp. 10-18.
- keawkuay, B. pairin. (2020, August 20). *Box Plot การแสดงผลแบบกล่องอันทรงพลัง*. getcodecamp.com. Retrieved September 29, 2021, from <https://getcodecamp.com/2020/08/20/box-plot-การแสดงผลแบบกล่องอันท/>.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, 1(6), 90-95.
- Li, S., Ma, H., & Li, W. (2017). Typical solar radiation year construction using k-means clustering and discrete-time Markov chain. *Applied Energy*, 205, 720-731.
- Lin, P., Peng, Z., Lai, Y., Cheng, S., Chen, Z., & Wu, L. (2018). Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. *Energy Conversion and Management*, 177, 704-717.
- Masoumi, M. (2021). Ocean data classification using unsupervised machine learning: Planning for hybrid wave-wind offshore energy devices. *Ocean Engineering*, 219, 108387.
- McLeod, S. (2019, July 19). *Box plots (also known as box and whister plots)*. Box Plot | Simply Psychology. Retrieved September 29, 2021, from <https://www.simplypsychology.org/boxplots.html>.
- Phanniti Co., Ltd. (2019). Solar hot water heater. phanniti.com. Retrieved September 30, 2021, from https://www.phanniti.com/Solar_hot_water_heater/58e46c720a431a00015ff8d9?gclid=Cj0KCQjwwNWKBhDAARIsAJ8HkhduHli_dS2lO8PTsEwxam7vPE2RTTHIS09PNrruMdE6sqa7Tfno7WYyAqEJELw_wcB.
- Solar Energy Industries Association [SEIA]. (2021). U.S. Solar Market Insight. SEIA. Retrieved September 30, 2021, from <https://www.seia.org/us-solar-market-insight>.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53-65.
- Šen Aleksandr, & Veresčajin Nikolaj K. (2002). *Basic set theory*. American Math. Soc.

- Yadav, J., & Sharma, M. (2013). A Review of K-mean Algorithm. *Int. J. Eng. Trends Technol*, 4(7), 2972-2976.
- Zhang, H., Cao, Y., Zhang, Y., & Terzija, V. (2018). Quantitative synergy assessment of regional wind-solar energy resources based on MERRA reanalysis data. *Applied energy*, 216, 172-182.



ประวัติผู้เขียน

นายอิสรา สร้อยสม เกิดเมื่อวันที่ 16 สิงหาคม พ.ศ. 2540 ที่จังหวัดเพชรบุรี แล้วย้ายที่อยู่มาเป็นจังหวัดระยองในปี พ.ศ. 2541 เริ่มเข้าเรียนในระดับประถมศึกษาในปี พ.ศ. 2548 ที่โรงเรียนบ้านตีนเนิน อำเภอบ้านค่าย จังหวัดระยอง แล้วเข้ารับการศึกษาในระดับมัธยมต้นในปี พ.ศ. 2553 และมัธยมปลายในปี พ.ศ. 2556 ที่โรงเรียนบ้านค่าย อำเภอบ้านค่าย จังหวัดระยอง แล้วในปี พ.ศ. 2559 ได้เริ่มเข้าศึกษาระดับปริญญาตรี ในสาขาวิชาวิศวกรรมคอมพิวเตอร์ สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี อำเภอเมือง จังหวัดนครราชสีมา และได้สำเร็จการศึกษาเมื่อปี พ.ศ. 2563

ภายหลังจากสำเร็จการศึกษาระดับปริญญาตรี ได้เข้าเรียนต่อระดับปริญญาโทในปี พ.ศ. 2563 ในสาขาวิชาวิศวกรรมโทรคมนาคมและคอมพิวเตอร์ สำนักวิชาวิศวกรรมศาสตร์มหาวิทยาลัยเทคโนโลยีสุรนารี อำเภอเมือง จังหวัดนครราชสีมา ภายใต้อาจารย์ รศ. ดร.นิตยา เกิดประสพ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และอาจารย์ รศ. ดร.กิตติศักดิ์ เกิดประสพ