บทที่ 3

วิธีการดำเนินการวิจัย

การดำเนินการวิจัยนี้ผู้วิจัยได้ศึกษาทฤษฎีและ งานวิจัยที่เกี่ยวข้องที่มีความจำเป็นต่อการวิจัย ดังได้นำไว้ในบทที่ผ่านมา เพื่อนำมาใช้สำหรับการสร้างแบบจำลองโครงข่ายประสาทเทียม โดยผู้วิจัยเริ่มต้นจากการศึกษาข้อมูล PM2.5 ศึกษาการสร้างแบบจำลองโดยใช้โครงข่ายประสาทเทียม รวมไปถึงการศึกษาการเขียนภาษาที่ต้องนำมาใช้ในการเขียนโปรแกรม คือ python โดยผู้วิจัยได้แบ่งขั้นตอนการดำเนินงานวิจัยได้ดังนี้

3.1 เตรียมข้อมูลเพื่อสร้างแบบจำลอง

3.2 การสร้างแบบจำลองโครงข่ายประสาทเทียม

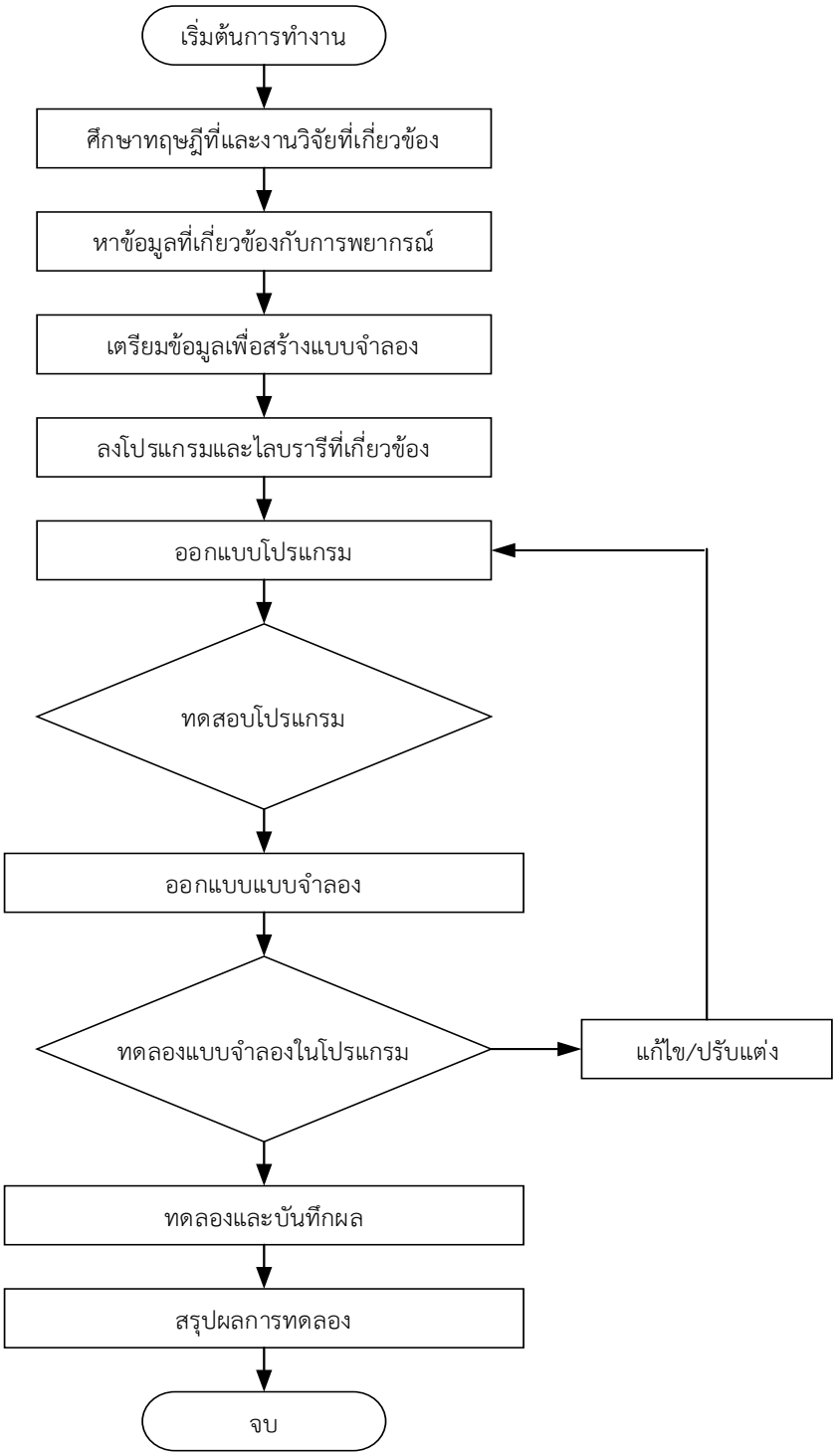
3.3 หน่วยวัดผลที่ใช้ในการทดลอง

3.4 การเตรียมการในการพัฒนาโปรแกรม

3.5 ตัวอย่างผลการทดลอง

3.6 คำสั่งที่ใช้ในโปรแกรมด้วยไลบรารี Keras ในการสร้างแบบจำลองโครงข่ายประสาทเทียม

ในการดำเนินการวิจัยผู้วิจัยได้ทำการเขียนโปรแกรมการทำงานต่าง ๆ เพื่อให้ได้ประสิทธิภาพสูงสุดในการดำเนินงานวิจัยนี้ ซึ่งผู้วิจัยได้แบ่งขั้นตอนต่าง ๆง ในการดำเนินงานเพื่อศึกษาหาวิธีการที่เหมาะสมกับการพยากรณ์ปริมาณ PM2.5 ในจังหวัดเชียงราย โดยขั้นตอนต่าง ๆ ผู้วิจัยได้จัดทำเอาไว้แสดงในแผนผังการดำเนินงานวิจัยดังภาพที่ 3.1



ภาพที่ 3.1 แผนผังการดำเนินงานวิจัยการพยากรณ์ปริมาณ PM2.5 ในจังหวัดเชียงรายโดยใช้โครงข่ายประสาทเทียม

จากภาพที่ 3.1 แสดงแผนผังการดำเนินงานวิจัยการพยากรณ์ปริมาณ PM2.5 ในจังหวัดเชียงรายโดยใช้โครงข่ายประสาทเทียม จากแผนผังการดำเนินงานวิจัยผู้วิจัยได้เริ่มจากการศึกษาเรื่อง

แบบจำลองโครงข่ายประสาทเทียมประกอบกับการศึกษางานวิจัยที่เกี่ยวข้องอื่น ๆ ทำให้ผู้วิจัยทราบว่าแบบจำลองโครงข่ายประสาทเทียมสามารถคาดการณ์ปริมาณ PM2.5ได้ ผู้วิจัยได้แบ่งข้อมูลออกเป็น 2 ส่วนคือชุดข้อมูลฝึกฝน (Training set) คิดเป็นร้อยละ 70 จากข้อมูลทั้งหมด เพื่อใช้สำหรับแบบจำลองโครงข่ายประสาทเทียมแบบหลายชั้น ส่วนชุดข้อมูลทดสอบ (Test set) คิดเป็นร้อยละ 30 จากข้อมูลทั้งหมด เพื่อใช้ในการพยากรณ์ปริมาณ PM2.5 ซึ่งข้อมูลนำเข้า 6 ข้อมูล โดยใช้โปรแกรมที่ผู้วิจัยพัฒนาขึ้นมา เป็นโปรแกรมภาษา Python โดยใช้ไลบรารี Keras เป็นไลบรารีในการพยากรณ์ปริมาณ PM2.5 แบ่งชุดข้อมูลฝึกฝนเป็นข้อมูลในการทดสอบคือ 0.30 ซึ่งการแบ่งข้อมูลไปทดสอบนี้ผู้วิจัยได้ทำการทดลองหาค่าคลาดเคลื่อนสัมพันธ์เฉลี่ยเพื่อหาประสิทธิภาพของแบบจำลอง จากนั้นกำหนดค่า Training cycle ตั้งแต่ 200 ถึง 1400 รอบ เพื่อทำการทดลองหาแบบจำลองโครงข่ายประสาทเทียมที่ให้ค่าคลาดคลาดเคลื่อนสัมพันธ์เฉลี่ยน้อยที่สุด ทดลองโดยการกำหนดจำนวนชั้นซ่อนตั้งแต่ 1 ถึง 4 ชั้น และจำนวนโหนดในชั้นซ่อนเท่ากัน ตั้งแต่ 6 ถึง 30 โหนด ในการทดลองผู้วิจัยได้ทำการเปรียบเทียบค่าในแต่ละวันโดยใช้ค่าความคลาดเคลื่อนสัมพัทธ์ จากนั้นนำมาเฉลี่ยเพื่อให้ได้ค่าความคลาดเคลื่อนสัมพัทธ์น้อยที่สุดพบว่าการทดลองโดยการกำหนดจำนวนชั้นซ่อน 4 ชั้น และจำนวนโหนดในชั้นซ่อน โหนดเท่ากันทุกชั้น ให้ค่าความคลาดเคลื่อนสัมพันธ์เฉลี่ยน้อยที่สุด ผู้วิจัยจึงได้นำแบบจำลองที่มีจำนวนชั้นซ่อน 4 ชั้น และจำนวนโหนดในชั้นซ่อน 30 โหนดเป็นแบบจำลองตั้งต้นเพื่อพยากรณ์ปริมาณ PM2.5

3.1 การเตรียมข้อมูลเพื่อสร้างแบบจำลอง

ทำการรวบรวมข้อมูลปริมาณมลพิษทางอากาศในพื้นที่จังหวัดเชียงราย ข้อมูลที่ใช้ในการทดลองอ้างอิงมาจากเว็บไซต์ wunderground ซึ่งได้แสดงข้อมูลทางอุตุนิยมวิทยาที่เกี่ยวข้อง ได้แก่ อุณหภูมิ จุดน้ำค้าง ความชื้น ความเร็วลม ความกดอากาศ และใช้ข้อมูลปริมาณ PM2.5 จากเว็บไซต์ berkeleyearth ในช่วงระยะเวลา 4 ปี 6 เดือน ตั้งแต่วันที่ 3 มีนาคม 2559 ถึง 24 กันยายน 2563 มาเป็นข้อมูลในการสร้างรูปแบบการพยากรณ์

ตารางที่ 3.1 รายละเอียดข้อมูลที่นำมาใช้ในการพยากรณ์

ชื่อข้อมูล	ความหมาย
Temperature	อุณหภูมิเฉลี่ยในแต่ละวัน
Dew Point	จุดน้ำค้างเฉลี่ยในแต่ละวัน
Humidity	ความชื้นเฉลี่ยในแต่ละวัน
Wind Speed	ความเร็วเฉลี่ยในแต่ละวัน
day	วันที่บันทึกข้อมูล
month	เดือนที่บันทึกข้อมูล
PM2.5	ปริมาณฝุ่นขนาดเล็กกว่า 2.5 ไมครอน

กระบวณการเตรียมข้อมูลมี 3 ขั้นตอน คือ

1. การคัดเลือกข้อมูล โดยข้อมูลที่นำมาใช้ในการวิจัย ได้แก่ อุณหภูมิ จุดน้ำค้าง ความชื้น ความเร็วลม ความกดอากาศ เดือน และ ปริมาณฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน

2. การกรองข้อมูล โดยการกรองข้อมูลที่ไม่ถูกต้องหรือไม่สมบูรณ์ ซึ่งพบว่าข้อมูลบางเรคคอร์ดมีค่าว่าง จึงคัดค่าว่างออก

3. การแปลงข้อมูล โดยข้อมูลที่ได้จากเว็บไซต์เป็นข้อมูลที่อยู่ในรูปแบบ ตารางผู้วิจัยจึงบันทึกค่าที่ได้จากเว็บไซต์บันทึกลงในไฟล์เอ็กเซลล์ จากนั้นแปลงข้อมูลให้อยู่ในรูปแบบ CSV เพื่อนำไปใช้ในการประมวลผลในโปรแกรมต่อไป

ตารางที่ 3.2 ตัวอย่างข้อมูลนำเข้าที่ 15 วัน

No.	Temperature	Dew Point	Humidity	Wind Speed	day	month	PM2.5
1	23.1	16.9	71	5.7	3	3	46.50667
2	24.4	17.1	67.8	14.6	4	3	37.76667
3	22.8	15.5	63.6	6	5	3	40.00417
4	24.8	17.4	66.7	4	6	3	50.18333

ตารางที่ 3.2 ต่อ

No.	Temperature	Dew Point	Humidity	Wind Speed	day	month	PM2.5
5	25.1	17.1	65.8	6.2	7	3	51.9375
6	26.2	18.2	64.6	5.1	8	3	51.82917
7	26.6	17.4	61.2	6.2	9	3	51.30417
8	25.9	17.3	61.4	4.5	10	3	64.7913
9	24.9	17.3	61.1	5.3	11	3	59.21667
10	22.8	15.6	43.5	5.4	12	3	61.93333
11	27.4	19.3	63.9	14.5	13	3	54.65652
12	26.2	18.6	65.4	5.4	14	3	68.0087
13	27.3	19.8	66.4	5	15	3	68.10417
14	28.1	18.7	60.5	5.2	16	3	71.74091
15	28.7	17.9	51.4	7.4	17	3	70.50417

จากตารางที่ 3.2 ผู้วิจัยได้ใช้ข้อมูลจาก อุณหภูมิ จุดน้ำค้าง ความชื้น ความเร็วลม ความกดอากาศ เดือน และปริมาณ PM2.5 เพื่อทำการเปรียบค่าจริงกับการพยากรณ์ในการเรียนรู้แต่ละรอบ ซึ่งข้อมูลดังกล่าวเป็นข้อมูลเกี่ยวข้องกับปริมาณ PM2.5 ที่ผู้วิจัยจะการพยากรณ์

3.2 การสร้างแบบจำลองโครงข่ายประสาทเทียม

ดัชนีคุณภาพอากาศที่ใช้อยู่ในประเทศไทย คำนวณโดยเทียบจากมาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไปสารมลพิษทางอากาศมี 6 ประเภท ได้แก่ ก๊าซโอโซน (O_3) เฉลี่ย 1 ชั่วโมง ก๊าซไนโตรเจนไดออกไซด์ (NO_2) เฉลี่ย 1 ชั่วโมง ก๊าซคาร์บอนมอนอกไซด์ (CO) เฉลี่ย 8 ชั่วโมง ก๊าซซัลเฟอร์ไดออกไซด์ (SO_2) เฉลี่ย 24 ชั่วโมง ฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM10) เฉลี่ย 24 ชั่วโมง และ PM2.5 เฉลี่ย 24 ชั่วโมง ทั้งนี้ ดัชนีคุณภาพอากาศที่คำนวณได้ของสารมลพิษทางอากาศประเภทใดมีค่าสูงสุด จะใช้เป็นดัชนีคุณภาพอากาศของวันนั้น

ดัชนีคุณภาพอากาศของประเทศไทยแบ่งเป็น 5 ระดับ คือ ตั้งแต่ 0 ถึง มากกว่า 300 ซึ่งแต่ละระดับจะใช้สีเป็นสัญลักษณ์เปรียบเทียบระดับของผลกระทบต่อสุขภาพอนามัย ตามตารางที่ 3.3 โดยดัชนีคุณภาพอากาศ 100 จะมีค่าเทียบเท่ามาตรฐานคุณภาพอากาศในบรรยากาศโดยทั่วไป หาก

ดัชนีคุณภาพอากาศมีค่าสูงเกินกว่า 100 แสดงว่าค่าความเข้มข้นของสารมลพิษทางอากาศมีค่าเกินมาตรฐานและคุณภาพอากาศในวันนั้นจะเริ่มมีผลกระทบต่อสุขภาพอนามัยของประชาชน

ตารางที่ 3.3 เกณฑ์ของดัชนีคุณภาพอากาศสำหรับประเทศไทย

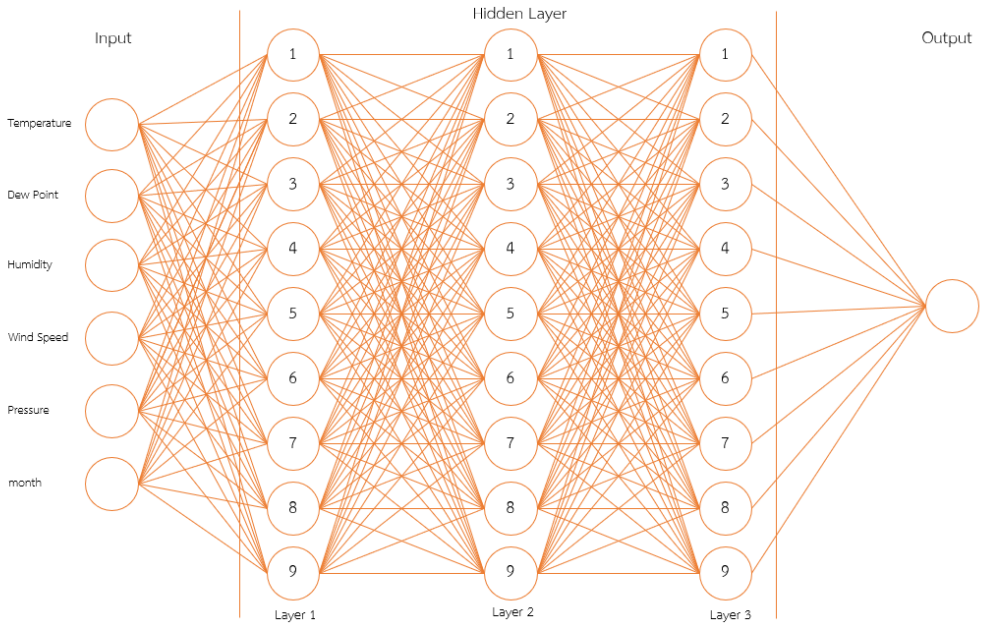
AQI	ความหมาย	สีที่ใช้	แนวทางการป้องกันผลกระทบ
0-50	คุณภาพดี	ฟ้า	ไม่มีผลกระทบต่อสุขภาพ
51-100	คุณภาพปานกลาง	เขียว	ไม่มีผลกระทบต่อสุขภาพ
101-200	มีผลกระทบต่อสุขภาพ	เหลือง	ผู้ป่วยโรคระบบทางเดินหายใจ ควรหลีกเลี่ยงการออกกำลังกายนอกอาคารบุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุ ไม่ควรทำกิจกรรมภายนอกอาคารเป็นเวลานาน
201-300	มีผลกระทบต่อสุขภาพมาก	ส้ม	ผู้ป่วยโรคระบบทางเดินหายใจ ควรหลีกเลี่ยงกิจกรรมภายนอกอาคารบุคคลทั่วไป โดยเฉพาะเด็กและผู้สูงอายุ ควรจำกัดการออกกำลังภายนอกอาคาร
มากกว่า 300	อันตราย	แดง	บุคคลทั่วไป ควรหลีกเลี่ยงการออกกำลังกายรอกอาคาร สำหรับผู้ป่วยโรคระบบทางเดินหายใจ ควรอยู่ภายในอาคาร

3.2.1 นำข้อมูลที่ใช้ในการทดลองทางอุตุนิยมวิทยาที่เกี่ยวข้อง ได้แก่ อุณหภูมิ จุดน้ำค้าง ความชื้น ความเร็วลม ความกดอากาศ เดือน และปริมาณ PM2.5 มาเป็นข้อมูลในการสร้างรูปแบบการพยากรณ์

3.2.2 นำข้อมูลแบ่งออกเป็น 2 ส่วนคือชุดฝึกฝน (Training set) คิดเป็นร้อยละ 70 จากข้อมูลทั้งหมด เพื่อใช้สำหรับแบบจำลองโครงข่ายประสาทเทียม ส่วนชุดข้อมูลทดสอบ (Test set) คิดเป็นร้อยละ 30 จากข้อมูลทั้งหมดโดยใช้โปรแกรมภาษา Python โดยใช้ไลบรารี Keras เป็นไลบรารีในการพยากรณ์ปริมาณ PM2.5

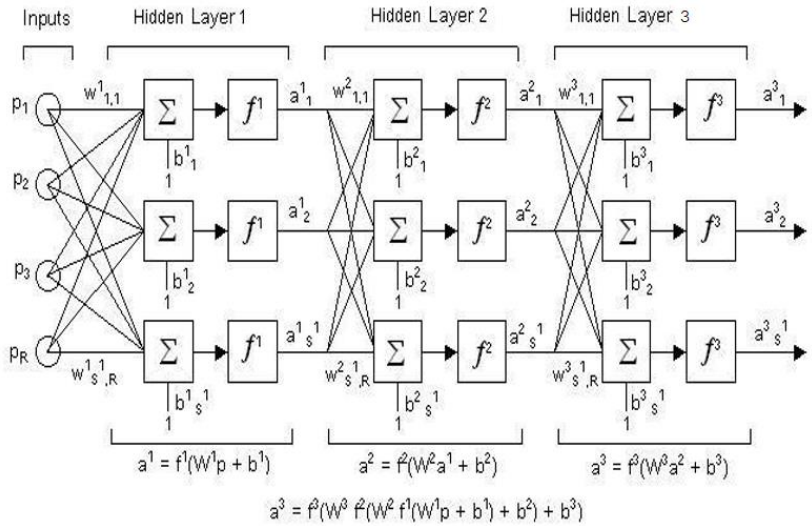
3.2.3 ทำการแบ่งชุดข้อมูลฝึกฝนเป็นข้อมูลในการทดสอบ และทดลองกำหนดจำนวนชั้น Hidden layer ตั้งแต่ 1 ชั้น ถึง 4 ชั้น และจำนวนโหนดใน Hidden layer ตั้งแต่ 6 ถึง 30 โหนด ในการทดลองได้ทำการเอาค่าที่พยากรณ์ที่ได้มาเปรียบเทียบเกณฑ์ของดัชนี PM2.5 แต่ละวันโดยใช้

อ้างอิงจากตารางที่ 3.4 และค่าความคลาดเคลื่อนสัมพัทธ์ จากนั้นนำค่าความคลาดเคลื่อนมาหาค่าเฉลี่ยเพื่อจะหาค่าความคลาดเคลื่อนเฉลี่ยสัมพันธ์ที่น้อยที่สุด



ภาพที่ 3.2 ตัวอย่างแบบจำลองโครงข่ายประสาทเทียมที่มีชั้นซ่อน 3 ชั้น มีโหนดในแต่ละชั้น 9 โหนด

การประมวลผลของโครงข่ายประสาทเทียมเมื่อมีข้อมูลนำเข้ามาจะประมวลผลโดยคูณกับค่าถ่วงน้ำหนัก ซึ่งแทนความสำคัญของข้อมูลนำเข้าซึ่งผลรวมของข้อมูลจะถูกส่งต่อไปยังฟังก์ชันการถ่ายโอน โดยการวิจัยนี้ใช้ฟังก์ชันเส้นตรงถูกปรับแก้ (Rectified Linear Unit , ReLU) เป็นฟังก์ชันในการถ่ายโอน ดังภาพที่ 3.3



ภาพที่ 3.3 การคำนวณประมวลผลข้อมูล

(ที่มา : พยุง มีสัจ. “โครงข่ายประสาทเทียม”. เทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีสารสนเทศพระจอมเกล้าพระนครเหนือ.)

โดยที่ p_i คือข้อมูลนำเข้า

$w^{1,2,3}$ คือค่าถ่วงน้ำหนัก (Weight) ที่เชื่อมระหว่างชั้นข้อมูลด้านอินพุตและ
ชั้นซ่อนที่ 1 , 2 และ 3

$b^{1,2,3}$ คือค่าไบแอส (Bias) ในชั้นซ่อนที่ 1 , 2 และ 3

$\mathrm{f}^{1,2,3}$ คือฟังก์ชันก์การถ่ายโอนในชั้นซ่อนที่ 1 , 2 และ 3

$a^{1,2,3}$ คือค่าผ่านฟังก์ชันการถ่ายโอน

การทดลองในวิจัยนี้จะทำการปรับเปลี่ยนจำนวนโหนดในแต่ละชั้นจากการลองผิดลองถูก (Trial and error) เพื่อหาแบบจำลองโครงข่ายประสาทเทียมที่มีจำนวนชั้นซ่อนและ จำนวนโหนดที่ทำให้ค่าพยากรณ์ใกล้เคียงกับค่าจริงมากที่สุด

3.3 หน่วยวัดผลที่ใช้ในการทดลอง

ค่าความผิดพลาดสัมพัทธ์ (Relative Error) (Steven, C & Raymond, P, 2010) ใช้ในการสรุปผลในแต่ละวันเพื่อเฉลี่ยหาค่าที่น้อยที่สุดเพื่อเลือกแบบจำลองที่มีความเหมาะสมที่สุดในการพยากรณ์ราคามันสำปะหลัง ดังสมการที่ 3.1

$$RE = \frac{|\ Actual\ -\ \Pr edict\ |}{|\ Actual\ |} \times 100 \qquad (3.1)$$

โดยที่ $Actual$ คือค่าจริง

$Predict$ คือค่าพยากรณ์

ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) (นวภัทรา หนูนาค, 2555) ใช้ในการสรุปประสิทธิภาพของแบบจำลองโครงข่ายประสาทเทียมดังสมการที่ 3.2

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|\ x_t - x_t\ | \qquad (3.2)$$

โดยที่ N คือจำนวนข้อมูลทั้งหมด

x_t คือค่าจริง

x_t คือค่าพยากรณ์

เกณฑ์ของดัชนี PM2.5 (กรมควบคุมมลพิษ) ใช้ในการแยกระดับข้อมูลที่พยากรณ์มาได้จากแบบจำลองโครงข่ายประสาทเทียมดังตารางที่ 3.4

ตารางที่ 3.4 เกณฑ์ของดัชนี PM2.5 อ้างอิงจากกรมควบคุมมลพิษ

PM2.5	ความหมาย
0-25	ดีมาก
26-50	ดี
51-100	ปานกลาง
101-200	เริ่มมีผลกระทบต่อสุขภาพ
201 ขึ้นไป	มีผลกระทบต่อสุขภาพ

จากตารางที่ 3.4 เป็นเกณฑ์ของดัชนี PM2.5 ของกรมควบคุมมลพิษ โดยมีทั้งหมด 5 ระดับด้วยกัน ระดับที่ 1 คือดีมากปริมาณ PM 2.5 อยู่ที่ 0-25 ระดับที่ 2 คือ ดีปริมาณ PM2.5 อยู่ที่ 26-50 ระดับ 3 คือ ปานกลางปริมาณ PM2.5 อยู่ที่ 51-100 ระดับที่ 4 คือ เริ่มมีผลกระทบต่อสุขภาพปริมาณ PM2.5 อยู่ที่ 101-200 และระดับที่ 5 คือ มีผลกระทบต่อสุขภาพปริมาณ PM2.5 อยู่ที่ 201 ขึ้น

3.4 การเตรียมการในการพัฒนาโปรแกรม

3.4.1 สภาพทั่วไปของคอมพิวเตอร์ที่ผู้วิจัยใช้งาน

คอมพิวเตอร์ที่ใช้งานเป็นคอมพิวเตอร์โน๊ตบุ๊ค CPU Core i-5 gen 6 การ์ดจอ NVDIA GeForce 920M 2GB RAM 8 GB

ระบบปฏิบัติการใช้ Windows 10 64 bit (ไลบรารีของ Keras รองรับระบบปฏิบัติการ 64 bit เท่านั้น)

3.4.2 ขั้นตอนการเตรียมการพัฒนา

1. ทำการติดตั้งโปรแกรม Anaconda (64 bit) และติดตั้ง Package ต่าง ๆ สำหรับ Anaconda

2. ทำการติดตั้งภาษา Python 3.7 ต้องเป็น Python เวอร์ชัน 3 ขึ้นไป (ไลบรารี Keras สนับสนุนตั้งแต่ Python 2.7 ขึ้นไป)

3. ทำการติดตั้งไลบรารี Tensorflow และ Keras

4. ทำการติดตั้งโปรแกรม Visual Studio Code ซึ่งผู้วิจัยได้ใช้โปรแกรม Visual Studio Code ในการเรียนและคอมไพล์โปรแกรม

3.5 ตัวอย่างผลการทดลอง

ผู้วิจัยได้จัดข้อมูลนำเข้าให้อยู่ในรูปแบบ Comma separate values (CSV) การทดลองตัวอย่างนี้ผู้วิจัยได้ทดลอง จำนวนรอบในการประมวลผล 1200 รอบ โดยการแบ่งข้อมูลเป็น 2 ชุดคือ ข้อมูลที่ใช้สำหรับการเรียนรู้ และข้อมูลที่ใช้สำหรับการพยากรณ์ ได้ทำการทดลองที่มีข้อมูลนำเข้า 6 ข้อมูล จำนวนชั้นซ่อนมีทั้งหมด 4 ชั้น ทุกชั้นมี 30 โหนดเท่ากัน จำนวนข้อมูลออก 1 ชั้น ในการทดลองในแต่ละครั้งค่าพยากรณ์ข้อมูลจะถูกจัดอยู่ในรูปแบบไฟล์ CSV จากนั้นนำปริมาณ PM2.5 ที่พยากรณ์ได้ในแต่ละวันหาค่าความคลาดเคลื่อนสัมพันธ์แล้วเฉลี่ยค่าความคลาดเคลื่อนสัมพันธ์ เพื่อให้ได้ค่าความผิดพลาดในการเรียนรู้ของโครงข่ายประสาทเทียมและความผิดพลดในการทดสอบ

ตารางที่ 3.5 ผลการทดลอง 30 วัน โดยแยกสีตามเกณฑ์ของดัชนีตารางที่ 3.4

d/m/y	PM2.5	Prediction
15/5/2019	25.84167	17.20784
16/5/2019	29.075	20.17537
17/5/2019	29.92917	15.94886

ตารางที่ 3.5 ต่อ

d/m/y	PM2.5	Prediction
18/5/2019	35.74583	27.24649
19/5/2019	48.25833	42.66168
20/5/2019	58.43333	40.46808
21/5/2019	39.02941	26.57197
22/5/2019	22.34167	32.22175
23/5/2019	22.1875	21.61069
24/5/2019	24.94167	11.65393
25/5/2019	23.06667	18.44446
26/5/2019	20.19583	15.66673
27/5/2019	23.30833	17.18107
28/5/2019	20.92083	14.84738
29/5/2019	21.2	17.61892
30/5/2019	11.93333	20.50463
31/5/2019	8.491667	15.43902
1/6/2019	12.525	12.65684
2/6/2019	13.60833	18.63173
3/6/2019	13.90417	19.20008
4/6/2019	14.57083	13.59154
5/6/2019	11.2875	12.26919
6/6/2019	11.37917	19.81785
7/6/2019	14.3	19.56124
8/6/2019	15.9375	15.33717
9/6/2019	16.59167	17.7409
10/6/2019	13.1375	12.96011
11/6/2019	8.219048	12.17561
12/6/2019	12.11304	10.89421

ตารางที่ 3.5 ต่อ

d/m/y	PM2.5	Prediction
18/5/2019	13/5/2019	9.766667

จากตารางที่ 3.5 เป็นผลการพยากรณ์ปริมาณ PM2.5 เป็นระยะเวลา 30 วัน ด้วยโครงข่ายประสาทเทียมที่อยู่ในรูปแบบ CSV โดย d/m/y คือวันเดือนปี PM2.5 คือ ค่าจริงของปริมาณ PM2.5 และ Prediction คือ PM2.5 ที่พยากรณ์ได้ จากตารางจะเห็นได้ว่า ค่าจริง และค่าที่พยากรณ์มาได้อยู่ในระดับเกณฑ์ดัชนีเดียวกัน 25 วัน ไม่อยู่ระดับเกณฑ์ดัชนีเดียวกัน 5 วัน

3.6 คำสั่งที่ใช้ในการโปรแกรมด้วยไลบรารี Keras ในการสร้างแบบจำลองโครงข่ายประสาทเทียม

จากการออกแบบ แบบจำลองโครงข่ายประสาทเทียมที่ได้กล่าวไว้ในตัวอย่างผลการทดลองได้ใช้จำนวนข้อมูลนำเข้า 6 ข้อมูล จำนวนชั้นซ่อน 4 ชั้น จำนวนโหนดในชั้นซ่อน 11 โหนดข้อมูล โดยมี Activation Function เป็น Relu และมีจำนวนข้อมูลออก 1 ข้อมูล ใช้ Loss function เป็นรากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Squared Error) Optimizer เป็น Adam และกำหนด Metrics เป็น Metrics.mae เพื่อวัดประสิทธิภาพของแบบจำลองโดยใช้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) ดังภาพที่ 3.4

```
def create_model():
    # create model
    model = Sequential()
    model.add(Dense(11, input_dim=X_train.shape[1], activation='relu'))
    model.add(Dense(11, activation='relu'))
    model.add(Dense(11, activation='relu'))
    model.add(Dense(11, activation='relu'))
    model.add(Dense(1))
    # Compile model
    model.compile(optimizer ='adam', loss = 'mean_squared_error',
                  metrics =[metrics.mae])
    return model
```

ภาพที่ 3.4 ตัวอย่างคำสั่งสร้างแบบจำลองโครงข่ายประสาทเทียม

ใช้คำสั่ง fit ในการเรียนรู้ข้อมูลโดยกำหนด X_train คือข้อมูลสำหรับการเรียนรู้แบบจำลอง และ y_train เพื่อให้แบบจำลองเปรียบเทียบผลต่างจากการคำนวณโดย Loss function จากนั้นจึงทำการปรับค่าถ่วงน้ำหนักกลับมาผ่าน optimizer ตั้งค่า epochs หรือจำนวนรอบที่จะทำการเรียนรู้ในที่นี้กำหนดไว้ที่ 500 รอบ กำหนด batch_size เพื่อกำหนดให้ X_train เรียนรู้ครั้งละ 140 ข้อมูล ระหว่างการเรียนรู้จะแบ่งข้อมูลสำหรับการตรวจสอบออกมาต่างหาก จะใช้ X_test และ y_test แบ่งออกมาเป็นข้อมูลทดสอบ ดังภาพที่ 3.5

```
history = model.fit(X_train, y_train, validation_data=(X_test,y_test), epochs=500, batch_size=140)
```

ภาพที่ 3.5 ตัวอย่างคำสั่งการเรียนรู้ของแบบจำลองโครงข่ายประสาทเทียม