

## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การวิจัยในหัวข้อเรื่อง “การใช้เทคนิคโครงข่ายประสาทเทียมหลายชั้น โดยพัฒนาโปรแกรม เพื่อพยากรณ์ปริมาณค่า PM2.5 ล่วงหน้า” ผู้วิจัยได้ทำการศึกษาทฤษฎีที่มีความจำเป็นต่องานวิจัย อีกทั้งได้ทำการศึกษางานวิจัยที่ผ่านมาที่มีความเกี่ยวข้องเพื่อนำมาเป็นพื้นฐานและ แนวคิดในการ พัฒนาสำหรับงานวิจัยนี้ ในการวิจัยพัฒนาการใช้เทคนิคโครงข่ายประสาทเทียมเพื่อทำนายปริมาณค่า PM2.5 โดยมีรายละเอียดของทฤษฎีและงานวิจัยที่เกี่ยวข้อง ดังต่อไปนี้

- 2.1 มลพิษทางอากาศ
- 2.2 การประมาณค่าระดับสารมลพิษทางอากาศ
- 2.3 โครงข่ายประสาทเทียม (Neural Networks)
- 2.4 แบบจำลองโครงข่ายประสาทเทียมแบบหลายชั้น
- 2.5 ไลบรารี Keras และ Tensorflow
- 2.6 ค่าความคลาดเคลื่อนสัมพัทธ์
- 2.7 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย
- 2.8 งานวิจัยที่เกี่ยวข้อง

#### 2.1 มลพิษทางอากาศ

มลพิษทางอากาศ (Air Pollution) เป็น ภาวะของอากาศที่มีสิ่งเจือปนอยู่ในปริมาณมากพอ และระยะเวลา นานพอ ที่จะทำให้เกิดผลเสียต่อสุขภาพของมนุษย์ สัตว์ พืช และวัสดุ โดยสิ่งเจือปนนี้ อาจเกิดขึ้นเองตามธรรมชาติ หรือจากการกระทำของคน ซึ่งอาจเป็นธาตุหรือสารประกอบทางเคมี และอาจอยู่ในรูปของก๊าซ หยดของเหลว และอนุภาคของแข็งระบบภาวะมลพิษทางอากาศ (Air Pollution System) จะประกอบด้วยส่วนประกอบ 3 ส่วนที่มีความสัมพันธ์กัน คือแหล่งกำเนิดมลพิษ ทางอากาศ บรรยากาศ และผลกระทบหรือผู้รับผลเสีย จากส่วนประกอบของระบบภาวะมลพิษทาง อากาศที่กล่าวมา จะเห็นได้ว่าในพื้นที่ใด ๆ ปริมาณและชนิดของสารมลพิษที่ถูกระบายออกจาก แหล่งกำเนิด สภาวะทางอุตุนิยมวิทยา และสภาพภูมิประเทศ จะเป็นตัวกำหนดถึงชนิด และปริมาณ หรือความเข้มข้นของสารมลพิษที่เจือปนอยู่ในบรรยากาศที่อยู่ห่างไกลออกไป และคุณภาพอากาศจะ

เป็นตัวกำหนดถึงลักษณะ และความรุนแรงของผลกระทบ หรือผลเสียหายที่เกิดขึ้นอีกทอดหนึ่ง (นพ ภาพร พานิช และ แสงสันต์ พานิช,2544)

2.1.1 อนุภาคมลสารหรือฝุ่นละออง อนุภาคมลสาร มลสารใด ๆ ในบรรยากาศหรือไอเสีย ซึ่งอยู่ในสภาพของแข็งหรือของเหลวที่อุณหภูมิ และความดันปกติ ยกเว้นไอน้ำซึ่งเกาะรวมกันเป็นกลุ่ม และมีเส้นผ่าศูนย์กลางตั้งแต่ 0.0002 ไมครอน แต่ไม่เกิน 500 ไมครอน อาจประกอบด้วยสาร นานาชนิด เช่น ซัลเฟต ไฮโดรคาร์บอน โลหะต่าง ๆ รวมทั้งละอองที่เกิดขึ้นตามธรรมชาติ ฝุ่นละอองที่เกิดขึ้นโดยปฏิกิริยาเคมีแสง (photochemical reaction) ฝุ่นละอองที่เกิดขึ้นเหล่านี้จะมีชื่อเรียก กันไปตามลักษณะการรวมตัวของฝุ่นละออง เช่น ควัน ฝุ่น หมอก เป็นต้น ขนาดและความหนาแน่น ของอนุภาคมลสาร เป็นองค์ประกอบสำคัญซึ่งควบคุมให้อนุภาคมลสารตกลงสู่พื้น ฝุ่นละอองที่มี ขนาดใหญ่อาจแขวนลอยในบรรยากาศได้เพียง 2-3 นาที แต่ฝุ่นละอองที่มีขนาดเล็กอาจแขวนลอยใน อากาศได้นานนับปี

ฝุ่นละอองแต่ละชนิด เมื่อแยกออกมาจากสารเดิมแล้ว จะทำให้มีพื้นที่รอบผิวของสารมากขึ้น เกิดมีช่องว่างมากขึ้น และขนาดของมลสารก็เปลี่ยนแปลงไปเช่นกัน ซึ่งการแตกแยกออกมานี้จะทำให้ ฝุ่นละอองที่เกิดจากสารนั้น ๆ มีคุณลักษณะด้านกายภาพ และด้านเคมีบางชนิดเปลี่ยนแปลงจาก เดิม เช่น มีขนาดเล็กกลงกว่าเดิม มีอัตราการฟุ้งกระจายเพิ่มขึ้น มีการละลายน้ำได้ดีขึ้นและทำปฏิกิริยากับ ออกซิเจนได้ดีขึ้นด้วย ซึ่งการที่มีฝุ่นละอองจะฟุ้งกระจายไปได้ไกลเท่าใด ขึ้นอยู่กับทิศทาง และ ความเร็วของกระแสลม ความชื้น และ อุณหภูมิ เช่น ถ้ามีความชื้นน้อย อุณหภูมิสูง และมีลมพัดแรง ก็จะทำให้ฝุ่นละอองฟุ้งกระจายไปได้ไกล

2.1.2 PM2.5 อนุภาคมลสารที่มีขนาดเล็กกว่า 2.5 ไมครอน หรืออาจเรียกว่าอนุภาคแขวนลอยสารแขวนลอยจะตกลงบนพื้นจากบรรยากาศโดยแรงโน้มถ่วงอย่างช้า ๆ และแขวนลอยอยู่ในอากาศเป็นเวลานาน

ประเทศไทยได้รับผลกระทบด้านชีวิตและสุขภาพ จากฝุ่นมลพิษแทบทุกภาคมาอย่างยาวนานกว่าทศวรรษและทวีความรุนแรงมากขึ้นในช่วงฤดูแล้ง โดยเฉพาะ 3-5 ปีล่าสุด โดยแหล่งกำเนิดมลพิษที่เป็นปัญหาและมีผลกระทบต่อชีวิตและสุขภาพของประชาชนมากที่สุดในช่วงฤดูแล้งคือ การเผาในพื้นที่เกษตรเพิ่มขึ้นจากการส่งเสริมอุตสาหกรรมเกษตร (โดยเฉพาะอุตสาหกรรม อ้อย ข้าวโพด และข้าว ในทุกภาคยกเว้นภาคใต้) และยังมีการทำเกษตรพืชเชิงเดี่ยวพื้นที่ภูเขา บ่อยครั้งที่ไฟลุกลามออกนอกพื้นที่ปลูกกลายเป็นสาเหตุสำคัญส่วนหนึ่งของไฟป่า(ภาคเหนือตอนบน) ตลอดจนมีการเพิ่มการเผาพื้นที่เกษตรจากการส่งเสริมอุตสาหกรรม เกษตรในประเทศเพื่อนบ้านด้วย

เช่นกัน ทำให้พื้นที่ทุกภาค ยกเว้นภาคใต้ได้รับผลกระทบจาก PM2.5 ในช่วงฤดูแล้งกันอย่างถ้วนทั่ว รวมทั้งกรุงเทพฯและปริมณฑลที่ได้รับ PM2.5 จากการเผาอุตสาหกรรมเกษตรเป็นหลักด้วยเช่นกัน เพราะกระแสลมที่พัดมวล PM2.5 จากแหล่งกำเนิดเข้าสู่กรุงเทพฯและปริมณฑลร่วมกับสภาพความกดอากาศ การกักขังอากาศไม่มีการถ่ายเทเอื้อให้มลพิษ PM2.5 ลอยแขวนในบรรยากาศอยู่นาน ทำให้สัดส่วนแหล่งกำเนิด PM2.5 มาจากการเผาอุตสาหกรรมเกษตรมีอิทธิพลมากกว่ามลพิษจากการจราจรหรือมลพิษจากปล่องโรงงานอุตสาหกรรม กล่าวโดยสรุปได้ว่ามลพิษ PM2.5 ในประเทศไทยในช่วงฤดูแล้งเกิดจากการเผาอุตสาหกรรมเกษตร อ้อย ข้าวโพด และข้าว เป็นหลัก และถูกซ้ำเติมด้วยการเผาอุตสาหกรรมเกษตรจากประเทศเพื่อนบ้าน ทั้งสาธารณรัฐกัมพูชา และสาธารณรัฐเมียนมาร์ (ศ.นพ.ชายชาญ โพธิรัตน์ แพทย์ผู้เชี่ยวชาญหน่วยโรคระบบการหายใจ เวชบำบัดวิกฤตและภูมิแพ้อายุรศาสตร์ คณะแพทยศาสตร์ มหาวิทยาลัยเชียงใหม่)

## 2.2 การประมาณค่าระดับสารมลพิษทางอากาศ

สถานการณ์คุณภาพอากาศ จะช่วยในการประเมินการเฝ้าระวัง วางแผนจัดการป้องกันและแก้ไขปัญหาคุณภาพอากาศได้อย่างเหมาะสม โดยเฉพาะบริเวณที่มีแนวโน้มระดับมลพิษทางอากาศสูงหรือเป็นบริเวณที่มีประชากรในกลุ่มที่อ่อนไหวต่อระดับมลพิษทางอากาศ เช่น บริเวณที่มีเด็กอาศัยหรือโรงเรียน เป็นต้น ซึ่งจำเป็นต้องทราบระดับมลพิษทางอากาศ ณ จุดที่เป็นตัวแทนพื้นที่ การตรวจวัดจึงเป็นวิธีการที่ใช้ดำเนินการศึกษาที่นิยมใช้ แต่การตรวจวัดจริงนั้นทำได้จำกัด เช่น สามารถตรวจวัดได้เพียงชั่วระยะเวลาหนึ่ง หรือบางครั้งต้องใช้เวลา ความไม่เพียงพอของจำนวนเครื่องมือและงบประมาณมีจำนวนจำกัด ดังนั้น การประมาณระดับสารมลพิษทางอากาศโดยการใช้แบบจำลองทางคณิตศาสตร์ จึงเป็นอีกแนวทางของการเฝ้าระวังคุณภาพอากาศ ลดข้อจำกัดของการตรวจวัดจริง (นพภาพร พานิช และ แสงสันต์ พานิช, 2544) ซึ่งแบบจำลองทางคณิตศาสตร์ด้านมลพิษทางอากาศแบ่งตามวิธีการได้ดังนี้

2.2.1 การใช้แบบจำลองทางด้านกายภาพ (Physical Model) การสร้างแบบจำลองของแหล่งกำเนิดและสภาพพื้นที่โดยรอบให้มีลักษณะเหมือนของจริงมากที่สุดโดยใช้วิธีย่อส่วนให้มีสัดส่วนอยู่ในระหว่าง 1 : 100 ถึง 1 : 5000 แล้วจึงศึกษาการแพร่กระจายของสารมลพิษจากแหล่งกำเนิดออกสู่บรรยากาศจริง ๆ โดยทำการศึกษาในอุโมงค์ลม (Wind Tunnel) แล้วเก็บตัวอย่างอากาศในบรรยากาศตามจุดต่าง ๆ ที่สนใจในอุโมงค์ลมที่ดำเนินการศึกษา วิธีนี้จะให้ผลถูกต้องใกล้เคียงกับความเป็นจริง แต่จะเสียค่าใช้จ่ายที่สูง

2.2.2 การใช้แบบจำลองทางคณิตศาสตร์ การใช้สมการทางคณิตศาสตร์ที่ได้มีการพัฒนามาจากพื้นฐานจากทฤษฎีการแพร่กระจายสารมลพิษทางอากาศ ซึ่งได้มีผู้คิดค้นปรับปรุงให้เหมาะสมกับการนำมาใช้งานเป็นวิธีที่นิยมใช้กันทั่วไป เนื่องจากเสียเวลาและค่าใช้จ่ายน้อยกว่าวิธีแรก แม้ว่าผลที่ได้จะถูกต้องแม่นยำน้อยกว่าการใช้แบบจำลองแบบย่อส่วนก็ตาม แต่ผลที่ได้มีความถูกต้องพอสมควรและเป็นที่ยอมรับ

2.2.3 การศึกษาสภาวะแวดล้อมทางอากาศโดยอาศัยหลักการของการใช้แบบจำลองทางคณิตศาสตร์ การใช้แบบจำลองทางคณิตศาสตร์เพื่อประเมินความเข้มข้นของมลพิษทางอากาศในบรรยากาศนั้น เป็นวิธีใช้กันอย่างแพร่หลายในการศึกษาผลกระทบที่เกิดจากแหล่งกำเนิดมลพิษทางอากาศซึ่งมีต่อผู้รับมลพิษซึ่งได้รับผลกระทบดังกล่าว การศึกษาผลกระทบที่ดีที่สุดนั้น ได้แก่ การตรวจวัดมลพิษทางอากาศด้วยเครื่องมือฟิสกส์ อาทิ การวัดฝุ่นแขวนลอยด้วยเครื่องวัดฝุ่นปริมาตรสูง (High Volume Air Sampler) หรือทางเคมี เช่น การวัดซัลเฟอร์ไดออกไซด์ โดยวิธีพาราโรซาริสัน แต่เนื่องจากการตรวจวัดทำได้จำกัด เช่น สามารถตรวจได้เพียงชั่วระยะเวลาหนึ่งหรือบางครั้งต้องใช้เวลานาน เช่น การวัดซัลเฟอร์ไดออกไซด์ต้องใช้เวลาราว 24 ชั่วโมง จึงจะได้ตัวอย่างหนึ่งสำหรับวิธีพาราโรซาริสัน ซึ่งค่าที่วัดได้ก็ถือเป็นค่าเฉลี่ย 24 ชั่วโมง แต่ในระหว่าง 24 ชั่วโมง นั้นหากต้องการทราบค่าในชั่วโมงใดชั่วโมงหนึ่งก็ไม่อาจจะทราบค่าได้ นอกจากจะใช้เครื่องมือที่มีราคาแพง ในลักษณะของการวัดแบบต่อเนื่อง (Continuous Sampler) ซึ่งก็มีข้อจำกัดอื่นอีกมาก

ส่วนการใช้แบบจำลองทางคณิตศาสตร์นั้น สามารถจะประเมินค่าเฉลี่ย 1 ชั่วโมง หรือ 24 ชั่วโมง หรือ 1 ปี ได้อย่างสมบูรณ์ แต่จำเป็นต้องรู้ข้อมูลของแหล่งกำเนิด และสภาพอุตุนิยมวิทยาอย่างถูกต้องที่สุด เพราะผลที่ได้ออกมาจะมีความถูกต้องไม่มากไปกว่าข้อมูลที่ป้อนเข้าไปเท่านั้น ในปัจจุบันไม่อาจกล่าวได้ว่า แบบจำลองทางคณิตศาสตร์มีความแม่นยำใการประเมินไม่เกิน 50 เปอร์เซ็นต์ ของความเข้มข้นที่เกิดขึ้นจริง เมื่อพิจารณาจากข้อดีข้อเสียของวิธีในการประเมินผลกระทบสิ่งแวดล้อมทั้งสองแบบจึงสมควรที่จะใช้ประกอบกันทั้งสองวิธี โดยสามารถจะนำข้อดีของแต่ละวิธีมาเสริมกันได้

### การตรวจวัดสารมลพิษในอากาศ

#### ข้อดี

1) ได้ค่าความเข้มข้นของมลพิษในอากาศที่ถูกต้อง ความผิดพลาดจะมีอยู่ในการวิเคราะห์เท่านั้น

2) หากตรวจวัดต่อเนื่องอย่างมีหลักเกณฑ์จะสามารถยืนยันผลกระทบต่อสิ่งแวดล้อมได้อย่างถูกต้องที่สุด (เฉพาะจุดที่ตรวจวัด)

### ข้อเสีย

- 1) การตรวจวัดต้องใช้เครื่องมือที่มีราคา และค่าใช้จ่ายแพง
- 2) ไม่อาจตั้งเครื่องมือไว้ต่อเนื่องเป็นเวลานาน ๆ และหลาย ๆ จุดให้เพียงพอ
- 3) ในบางกรณี เช่น ยังไม่ได้สร้างหรือมีแหล่งกำเนิดมลพิษ ก็ใช้วัดเฉพาะความเข้มข้นของมลพิษที่มีอยู่แล้ว เพื่อหวังผลในการเปรียบเทียบในอนาคตเท่านั้น แต่เอามาใช้ในการประเมินผลกระทบไม่ได้
- 4) หากไม่สามารถตรวจวัดได้เป็นเวลานาน และหลายจุดเพียงพอ จะทำให้การประเมินผลกระทบผิดพลาดได้ เนื่องจากความผันแปรของอุตุนิยมวิทยา ซึ่งอาจเกิดในช่วงเวลาที่ตรวจวัด และหากจุดตรวจวัดไม่เพียงพอก็อาจไม่เป็นตัวแทนในการพิจารณาผลกระทบสิ่งแวดล้อมได้ เช่น ในกรณีพื้นที่คาดว่าจะได้รับผลกระทบกว้างขวางมาก

### การใช้แบบจำลองทางคณิตศาสตร์

#### ข้อดี

- 1) สามารถใช้ประเมินผลกระทบได้ทุกกรณีที่ทราบ หรือคาดคะเนแหล่งกำเนิดมลพิษ (ที่ตั้ง อัตราการปล่อยสารมลพิษ ลักษณะการปล่อยสารมลพิษ) และสภาพอุตุนิยมวิทยาในช่วงเวลาที่ต้องการประเมิน (ความเร็วและทิศทางลม ความเสถียรของบรรยากาศ เป็นต้น) จึงสามารถใช้ประเมินผลกระทบจากแหล่งกำเนิดที่ยังมิได้สร้างหรือมีขึ้นได้ด้วย
- 2) มีค่าใช้จ่ายน้อยกว่าการตรวจวัดจริง แต่ให้ผลที่กว้างขวางและละเอียดกว่าเช่น ทำเส้นความเข้มข้น (Contour) ได้ทั้งบริเวณศึกษา
- 3) สามารถทราบผลได้อย่างรวดเร็ว จึงเป็นเครื่องมือในการประเมินผลกระทบเบื้องต้นที่ดี เช่น การเลือกจุดตรวจวัดที่เหมาะสม หรือประเมินผลกระทบสิ่งแวดล้อมที่เกิดขึ้นเนื่องจากการเปลี่ยนแปลงปรับปรุงแหล่งกำเนิด

## ข้อเสีย

1) หากแหล่งข้อมูลแหล่งกำเนิด หรือสภาพอุตุนิยมวิทยามีความคลาดเคลื่อน การประเมินก็จะผิดพลาดไปด้วย

- ความผิดพลาดของข้อมูลแหล่งกำเนิด เช่น อัตราการปล่อยสารมลพิษจะทำให้ผลที่คำนวณได้ผิดพลาดไปในสัดส่วนเดียวกัน

2) เนื่องจากผลที่ได้จากแบบจำลองทางคณิตศาสตร์มีความสวยงาม น่าเชื่อถือจึงอาจได้รับการยอมรับ และเชื่อถือมากเกินไปกว่าความถูกต้องของแบบจำลอง ซึ่งควรต้องมีการตรวจสอบความถูกต้อง และยืนยันการตรวจวัดจริงอยู่เสมอ

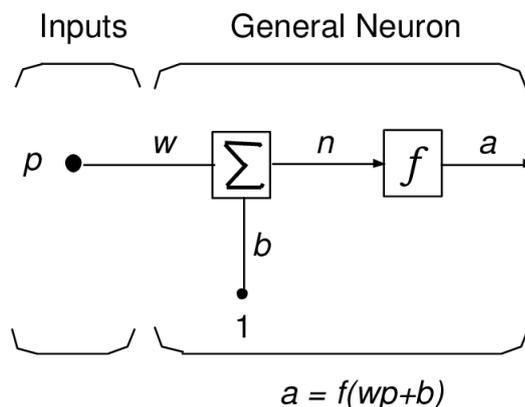
## วิธีการใช้ทั้งสองวิธีเสริมกัน

1) ใช้แบบจำลองทางคณิตศาสตร์ ซึ่งสามารถได้ผลโดยรวดเร็ว เพื่อแสดงจุดที่คาดว่าจะมีผลกระทบสูง แล้วจึงตั้งจุดตรวจวัดด้วยวิธีการเคมี หรือฟิสิกส์ที่จุดนั้น ๆ

2) นำผลการตรวจวัดที่ได้จริงนั้นมาเปรียบเทียบกับผลที่ได้จากแบบจำลองทางคณิตศาสตร์ เพื่อปรับแต่งข้อมูลที่ใช้กับแบบจำลองทางคณิตศาสตร์ ให้มีความถูกต้องยิ่งขึ้น

## 2.3 โครงข่ายประสาทเทียม (Neural Networks)

Neural Networks หรือโครงข่ายประสาทเทียมมีการทำงานแบบขนาน จำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ประกอบด้วยตัวแปรเข้า (Input) ค่าน้ำหนัก (Weight) ตัวแปรโน้มน้าเอียง (Bias) ฟังก์ชันกระตุ้น (Activation Function) และตัวแปรออก (Output) เรียนรู้แบบมีการสอน (Supervised Learning) และเรียนรู้แบบไม่มีการสอน (Unsupervised Learning) สามารถประยุกต์ใช้งาน ได้หลากหลาย ได้แก่ การจดจำรูปแบบ (Pattern Recognition) การจับกลุ่ม (Clustering) การประมาณค่า (Approximation) การทำนาย (Prediction) และการหาค่าที่เหมาะสมที่สุด (Optimization)



ภาพที่ 2.1 การทำงานของนิวรอนภายในโครงข่ายประสาทเทียม

(ที่มา : Manuel Carlos Gameiro da Silva, Carlos Alcobia, N.Martinho, João Esteves Ramos, " www.researchgate.net")

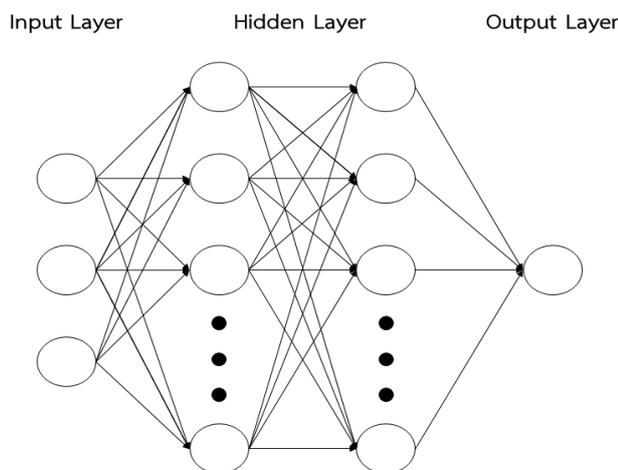
โครงข่ายประสาทเทียมแบบเพอร์เซพตรอนหลายชั้น (Multi-Layer Perceptron : MLP) ประกอบด้วยโครงข่าย 3 ชั้น ชั้นข้อมูลเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) โดยชั้นซ่อนสามารถมีได้มากกว่าหนึ่งชั้น

โครงข่ายประสาทเทียมที่ใช้การเรียนรู้แบบแพร่กระจายย้อนกลับ (Back Propagation Learning) มีหลักการเรียนรู้คือ การเปลี่ยนแปลงค่าถ่วงน้ำหนัก (Weight) เพื่อให้ได้ผลลัพธ์ที่ใกล้เคียงกับค่าจริงมากที่สุด โดยวิธีการแพร่ย้อนกลับ โดยพารามิเตอร์ที่มีผลต่อแบบจำลองคือ อัตราการเรียนรู้ และค่าโมเมนตัม นอกจากนี้แล้วชนิดของฟังก์ชันกระตุ้น (Activation Function) ก็มีผลต่อประสิทธิภาพของแบบจำลองเช่นกัน ซึ่งสามารถเลือกฟังก์ชันกระตุ้นได้หลายแบบ เช่น ฟังก์ชันล็อกซิก ฟังก์ชันแทนซิก และ ฟังก์ชันเพียวลิน

Radial Basis Function Network: RBF คือโครงข่ายแบบไปข้างหน้า (Feed Forward) ที่มีประสิทธิภาพสูง แตกต่างจากโครงข่ายเพอร์เซพตรอนแบบหลายชั้น (Multi-layer Perceptron: MLP) ที่มีชั้นซ่อนเพียงชั้นเดียว ซึ่งแต่ละนิวรอนในชั้นจะมีฟังก์ชันโอนถ่าย (Mapping Function) ลักษณะพิเศษ ระหว่างอินพุตและเอาต์พุตได้ โดยการเรียนรู้จะทำการปรับค่าน้ำหนักให้ได้ค่าฟังก์ชันการส่งที่เหมาะสมที่สุด ซึ่งผลตอบสนองของฟังก์ชันขึ้นอยู่กับระยะห่างระหว่างอินพุตกับจุดศูนย์กลางของฟังก์ชัน คือถ้าใกล้จุดศูนย์กลางเอาต์พุตจะมาก แต่ถ้าอยู่ห่างเอาต์พุตที่ได้จะลดลงตามลำดับ ดังนั้น RBF จึงเหมาะในงานการประมาณค่าฟังก์ชัน ฟังก์ชันที่นิยมใช้ใน BRF มากที่สุด คือฟังก์ชันเกาส์เซียน (Gaussian Function) โดยมีพารามิเตอร์การกระจาย (Spread Parameter) เป็นตัวควบคุมความกว้าง RBF

## 2.4 โครงข่ายประสาทเทียมแบบหลายชั้น

โครงข่ายประสาทเทียมแบบหลายชั้นแสดงได้ ดังภาพที่ 2.2 ใช้สำหรับงานที่มีความซับซ้อน มีการฝึกฝนเป็นแบบต้องมีผู้สอน (Supervised Learning) และ ใช้ขั้นตอนการเรียนรู้แบบแพร่ย้อนกลับ (Backpropagation) สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับประกอบด้วย 2 ส่วน คือ การส่งผ่านไปข้างหน้า (Forward Pass) การส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้าข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลเข้าและ จะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออกส่วนการส่งผ่านย้อนกลับค่าน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการแก้ข้อผิดพลาด (Error Correction) คือผลต่างของผลตอบที่แท้จริง (Actual Response) กับผลตอบเป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด (Error Signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อและ ค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย



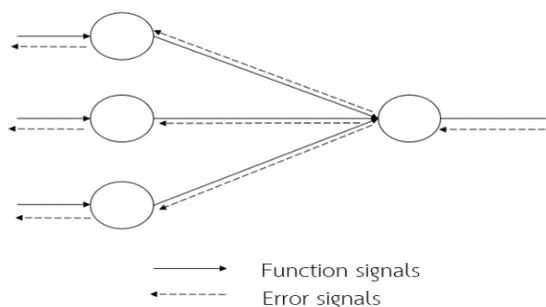
ภาพที่ 2.2 แบบจำลองโครงข่ายประสาทเทียมหลายชั้น

(ที่มา : Fausett, L., “Fundamentals of Neural Networks,”

Prentice Hall, Inc., New jersey. pp.461, 1994.)

สัญญาณที่มีในโครงข่ายประสาทเทียมแบบหลายชั้น มี 2 ประเภทคือ Function Signal และ Error Signal โดย Function Signal เป็นสัญญาณเข้าที่มาจากโหนดในชั้นก่อนหน้าและ จะส่งผ่านไปข้างหน้าจากโหนดหนึ่งไปสู่อีกโหนดหนึ่งและ Error Signal เป็นสัญญาณย้อนกลับที่เกิดขึ้นที่โหนดใน

ชั้นข้อมูลออกของโครงข่ายประสาทเทียมและ ถูกส่งผ่านย้อนกลับจากชั้นหนึ่งไปสู่อีกชั้นหนึ่ง ดังภาพที่ 2.3



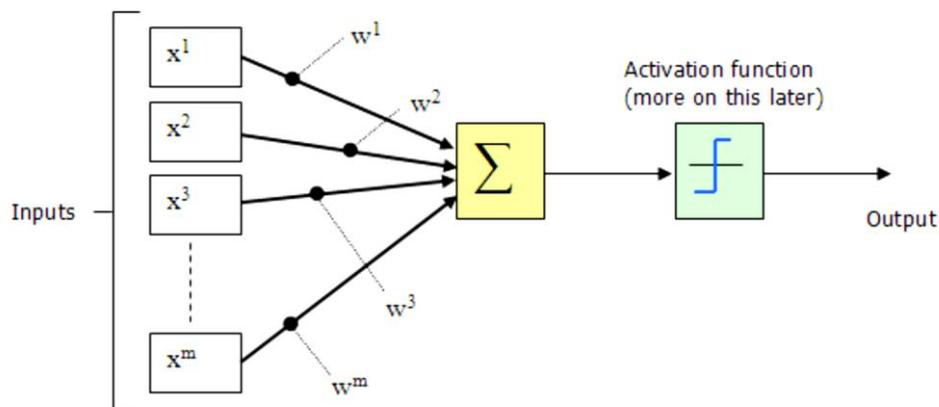
ภาพที่ 2.3 สัญญาณที่มีในโครงข่ายประสาทเทียมแบบหลายชั้น

(ที่มา : Haykin, S., “Neural Networks: A Comprehensive Foundation,”  
Prentice Hall. pp. 181, 1999.)

หลักการทำงานคือในแต่ละชั้นของชั้นซ่อนตัว (Hidden Layer) จะมีฟังก์ชันสำหรับคำนวณเมื่อได้รับสัญญาณจากโหนดในชั้นก่อนหน้านี้ เรียกว่า Activation Function โดยในแต่ละชั้นไม่จำเป็นต้องเป็นฟังก์ชันเดียวกันก็ได้ ชั้นซ่อนตัวนั้นมีหน้าที่สำคัญคือจะพยายามแปลงข้อมูลที่เข้ามาในชั้นนั้น ๆ ให้สามารถแยกแยะความแตกต่างโดยใช้เส้นตรงเส้นเดียวและ ก่อนที่ข้อมูลจะถูกส่งไปถึงชั้นข้อมูลออก (Output Layer) ในบางครั้งอาจจำเป็นต้องใช้ชั้นซ่อนตัวมากกว่า 1 ชั้นในการแปลงข้อมูลให้อยู่ในรูปเส้นตรงเส้นเดียว (Linearly Separable)

ในการคำนวณหาข้อมูลออกในปัญหาการจำแนกทำได้โดยการใส่ข้อมูลนำเข้าเข้าไปในโครงข่ายประสาทเทียมที่เราได้ทำการหาไว้แล้วจากนั้นให้ทำการเปรียบเทียบค่าของข้อมูลออกในชั้นข้อมูลออกและ ให้ทำการเลือกค่าของข้อมูลออกที่มีค่าสูงกว่าและ ทำการรับค่าของพยากรณ์ที่ตรงกับนิรอรลที่เลือกและ ให้นำค่ามาเปรียบเทียบกับค่าที่ยอมรับได้หากค่าอยู่ในช่วงที่รับได้ก็ให้ทำการรับข้อมูลชุดถัดไป แต่หากค่ามากกว่าค่าที่ยอมรับได้ ให้ทำการปรับค่าน้ำหนักและ ไบแอส ตามขั้นตอนที่ได้กล่าวไว้ข้างต้นเมื่อทำการปรับน้ำหนักเรียบร้อยแล้ว ให้ทำการรับข้อมูลชุดถัดไปและ ทำตามขั้นตอนซ้ำอีกรอบจนกระทั่งถึงข้อมูลชุดสุดท้ายและ เมื่อทำข้อมูลชุดสุดท้ายเสร็จจะนับเป็น 1 รอบของการคำนวณ จากนั้นจะทำการหาค่าผิดพลาดรวมเฉลี่ยจากค่าเฉลี่ยที่ได้เก็บค่าเอาไว้ เพื่อใช้ในการตรวจสอบว่าค่าโดยเฉลี่ยในการจำแนกนั้น มีค่าน้อยกว่าค่าผิดพลาดที่ยอมรับได้หรือไม่ถ้าใช่แสดงว่าโครงข่ายประสาทเทียมที่สร้างขึ้นนั้นสามารถให้ผลลัพธ์ที่ถูกต้องของทุก ๆ ข้อมูลแล้วจึงทำการจบการ

เรียนรู้ได้ แต่ถ้าไม่ใช่ให้กลับไปทำตามขั้นตอนแรก โดยเริ่มรับข้อมูลชุดที่ 1 ใหม่ ลักษณะการทำงานของโครงข่ายประสาทเทียม ดังภาพที่ 2.4



ภาพที่ 2.4 การทำงานของโครงข่ายประสาทเทียม

(ที่มา : Haykin, S., “Neural Networks: A Comprehensive Foundation,”  
Prentice Hall. pp. 33, 1999.)

ในแต่ละโหนดมีกระบวนการตัดสินใจดังสมการที่ 2.1

$$\begin{aligned} Y_i &= f(w_i^1 x_1 + w_i^2 x_2 + w_i^3 x_3 + \dots + w_i^m x_m) \\ &= f\left(\sum_j w_i^j x_j\right) \end{aligned}$$

โดยที่

$x_j$  = ข้อมูลนำเข้า

$w_j$  = น้ำหนัก (weight) ของแต่ละแขน

ในชั้นซ่อน (Hidden layer) มีฟังก์ชันในการคำนวณ (Activate function) ผู้วิจัยได้เลือกฟังก์ชันเส้นตรงที่ถูกปรับแก้ (Rectified Linear Unit , ReLU) ใช้ในการคำนวณซึ่งมีความไวในการเรียนรู้ได้ไวกว่าฟังก์ชันตัวอื่น ๆ ดังสมการที่ 2.2

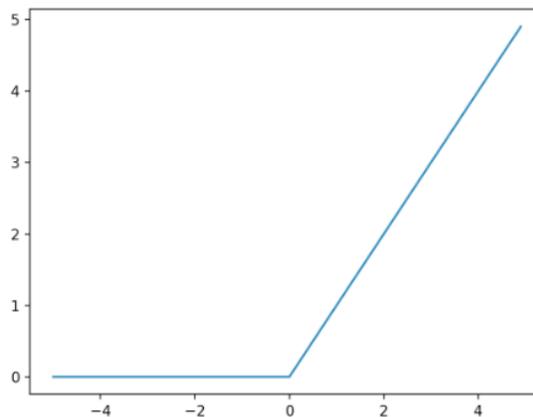
$$Y(x) = \max \{0, x\} = \{0 \text{ for } x \leq 0, x \text{ for } x > 0\} \quad (2.2)$$

โดยที่

$Y$  คือค่าที่ผ่านฟังก์ชันของชั้นซ่อนในแต่ละโหนด

$A$  คือผลรวมของแต่ละโหนด

ฟังก์ชันเส้นตรงที่ถูกปรับแก้ (Rectified Linear Unit , ReLU) เป็นฟังก์ชันในการเทรนโมเดลที่มีความเร็วในการเทรนโมเดลที่ไวกว่าฟังก์ชันอื่น ๆ ค่าจะเป็น 0 ก็ต่อเมื่อนำเข้ามานั้นมีค่าน้อยกว่าหรือเท่ากับ 0 แต่เมื่อนำเข้ามีค่ามากกว่า 0 ค่านำเข้าจะเท่ากับตัวมันเอง



ภาพที่ 2.5 กราฟฟังก์ชันเส้นตรงที่ถูกปรับแก้

## 2.5 ไบรารี Tensorflow และ Keras

2.5.1 ไบรารี TensorFlow คือโอเพ่นซอร์สไลบรารีสำหรับการสร้างโมเดลการเรียนรู้เชิงลึกที่พัฒนาโดย Google รองรับการทำงานแบบประมวลผลร่วมกันหลายๆ เครื่องและ สามารถใช้หน่วยประมวลผลกราฟิก (Graphics Processing Unit : GPU) ในการประมวลผลได้และมีอัลกอริทึมสำหรับการสร้างโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึกที่หลากหลายและ มีการนำไปประยุกต์ใช้กับงานวิจัยในหลายสาขาวิชาเช่น การรู้จำเสียง การประมวลผลภาพ สร้างหุ่นยนต์ ฯลฯ โดยการทำงานของ TensorFlow จะมีการสร้างกราฟสำหรับการประมวลผล ซึ่งกราฟจะประกอบด้วยกลุ่มของโหนด โดยกราฟจะแสดงการไหลของการประมวลผลข้อมูล

2.5.2 ไบรารี Keras คือ API สำหรับการพัฒนาโครงข่ายประสาทเทียมซึ่งเขียนด้วยภาษาไพธอนสามารถทำงานโดยอาศัยไลบรารีอย่าง TensorFlow โดย Keras สามารถรองรับการสร้างโมเดลการเรียนรู้ของเครื่องในรูปแบบโครงข่ายประสาทเทียมตลอดจนการสร้างโมเดลการเรียนรู้เชิงลึกโดยมีข้อดี คือง่ายต่อการเขียนและ ทำความเข้าใจ มีการทำงานแยกเป็นส่วนๆ สามารถแยกส่วนประกอบต่าง ๆ ในการสร้างโมเดลอย่างเช่น Neural Layers Cost Functions Optimizers Activation Functions และ ประกอบกันเป็นโมเดลใหม่ได้ สามารถพัฒนาฟังก์ชันหรือคลาสต่าง ๆ เพิ่มได้ง่าย โดยการพัฒนาทั้งหมดต้องใช้ภาษาไพธอนในการพัฒนา

## 2.6 ค่าความคลาดเคลื่อนสัมพัทธ์

ค่าความคลาดเคลื่อนสัมพัทธ์ (Relative Error) หรือ Percent Error คือเปอร์เซ็นต์ความผิดพลาดได้จากค่าความผิดพลาดสัมบูรณ์เปรียบเทียบกับค่าจริง ค่าจริง ดังสมการ 2.3

$$RE = \frac{| \text{Actual} - \text{Predict} |}{| \text{Actual} |} \times 100 \quad (2.3)$$

โดยที่

*Actual* คือค่าจริง

*Predict* คือค่าพยากรณ์

## 2.7 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย

ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) เป็นวิธีการหาค่าเฉลี่ยของความแตกต่างสัมบูรณ์ระหว่างค่าพยากรณ์และ ค่าจริง หากค่า MAE มีค่าน้อย แสดงว่าค่าพยากรณ์สามารถประมาณค่าได้ใกล้เคียงกับค่าจริง MAE สามารถคำนวณได้ดังสมการที่ 2.4

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad (2.4)$$

โดยที่

$N$  คือจำนวนข้อมูลทั้งหมด

$x_i$  คือค่าจริง

$\hat{x}_i$  คือค่าพยากรณ์

## 2.8 งานวิจัยที่เกี่ยวข้อง

เศรษฐ์ สัมภัตตะกุล และคณะ (2561) ศึกษาการวิเคราะห์ปริมาณฝุ่นละอองเชิงมวล PM2.5 และ PM10 ในบรรยากาศด้วยเครื่องตรวจวัดฝุ่นละอองไร้สายในพื้นที่ภาคเหนือของประเทศไทย โดยมีวัตถุประสงค์เพื่อสร้างและพัฒนาระบบฐานข้อมูลการเปลี่ยนแปลงสภาพภูมิอากาศ การตรวจวัดปริมาณฝุ่นละอองเชิงมวล PM2.5 และ PM10 ในอากาศใช้เครื่องตรวจจับฝุ่นละอองไร้สายและพัฒนาระบบฐานข้อมูลที่เรียกว่า Cloud Computing ติดตั้งสถานีตรวจวัดปริมาณละอองเชิงมวลเรียลไทม์เดือนกุมภาพันธ์ถึงพฤษภาคม (PM2.5, PM10) ในพื้นที่ภาคเหนือ จำนวนทั้งหมด 4 สถานี ได้แก่ (1) สถานีโรงเรียนยุพราช (2) สถานีอำเภอต๋อยสะเก็ด (3) สถานีตำบลแม่เหียะอำเภอเมืองมหาวิทยาลัยเชียงใหม่ (4) สถานีอำเภอนาน้อยจังหวัดน่าน จากการตรวจวัดปริมาณฝุ่นละอองขนาด

เล็ก PM2.5 และ PM10 ของเครื่องวัดฝุ่นละอองขนาดเล็ก DustDETEC ในอากาศต่อเนื่อง 24 ชั่วโมง ในช่วงเดือนมีนาคม พ.ศ.2560 จะเห็นได้ว่า ในช่วงของวันที่ 23-24 มีนาคม พ.ศ.2560 จะมีปริมาณของฝุ่นละอองขนาดเล็ก PM2.5 และ PM10 ที่ค่อนข้างสูงกว่าวันอื่น ๆ นำค่าฝุ่นละอองขนาดเล็ก PM2.5 และ PM10 เฉลี่ยรายวันที่วัดได้จากเครื่องวัดฝุ่นละอองขนาดเล็ก DustDETEC เปรียบเทียบกับข้อมูลจากสถานีตรวจวัดคุณภาพอากาศกรมควบคุมมลพิษ (สถานี โรงเรียนยุพราช) ระหว่างวันที่ 17 มีนาคม ถึงวันที่ 10 เมษายน พ.ศ.2560 พบว่าข้อมูลที่ได้มีแนวโน้มที่ใกล้เคียงกัน และเมื่อนำข้อมูลฝุ่นละอองขนาดเล็กจากเครื่องวัดฝุ่นละออง DustDETEC เปรียบเทียบกับข้อมูลจากกรมควบคุมมลพิษ และปริมาณจุดความร้อนที่เกิดขึ้นในช่วงเดือน มีนาคมและเมษายน พ.ศ.2560 พบว่ามีแนวโน้มเป็นไปในทิศทางเดียวกัน คือหากช่วงที่มีปริมาณจุดความร้อนสูง ก็จะส่งผลให้ค่าฝุ่นละอองขนาดเล็ก PM2.5 และ PM10 สูงขึ้นตามไปด้วย นอกจากนี้ได้ทำการพัฒนาระบบฐานข้อมูลการเปลี่ยนแปลงสภาพภูมิอากาศ รวมไปถึงการนำเสนอข้อมูลผ่านระบบรายงานผลออนไลน์จากสถานีตรวจวัดในแต่ละแห่ง เพื่อให้ผู้รับข้อมูลสามารถเข้าถึงข้อมูลได้อย่างรวดเร็วและทันสถานการณ์

เกษรารักษ์ ลีลาสิทธิกุล และคณะ (2561) ได้ทำการศึกษาเรื่อง ผลจากหมอกควันและมลพิษทางอากาศต่อระบบหัวใจและทางเดินหายใจในพื้นที่ภาคเหนือของประเทศไทย โดยมีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์ ระหว่างหมอกควัน (PM10, CO) กับสมรรถภาพทางปอดและความทนทานของระบบหัวใจและหลอดเลือดในประชากรที่อาศัยอยู่ในพื้นที่ภาคเหนือของประเทศไทย ได้แก่ จังหวัดเชียงใหม่ จังหวัดเชียงราย และจังหวัดน่าน โดยใช้วิธีศึกษาแบบ Cross-sectional study กลุ่มตัวอย่าง 450 คน ประกอบด้วย กลุ่มเด็ก (อายุเฉลี่ย  $12.27 \pm 1.45$ ) กลุ่มผู้ใหญ่ (อายุเฉลี่ย  $42.59 \pm 11.10$ ) และกลุ่มผู้สูงอายุ (อายุเฉลี่ย  $69.29 \pm 8.02$ ) ผลการศึกษาพบว่า สมรรถภาพปอด และความทนทานของระบบหัวใจและหลอดเลือดมีความสัมพันธ์เชิงลบกับระดับของก๊าซคาร์บอนมอนออกไซด์ (CO) ในบรรยากาศอย่างมีนัยสำคัญทางสถิติ นอกจากนี้ เมื่อควบคุมตัวแปร เช่น อายุ ระยะเวลาที่พักอาศัยในพื้นที่ และเพศพบว่าค่าความสัมพันธ์ดังกล่าวยังคงพบในก๊าซ CO ในบรรยากาศเช่นเดิม โดยผลกระทบต่อสุขภาพจากก๊าซ CO นั้น ทำให้เกิดอาการเวียนศีรษะ คลื่นไส้ อาเจียน อ่อนเพลีย และรู้สึกไม่มีแรง ยิ่งไปกว่านั้นก๊าซ CO ยังสามารถรวมตัวกับฮีโมโกลบินในเลือด กลายเป็นคาร์บอกซีฮีโมโกลบิน (HbCO) ทำให้ขัดขวางการขนส่งก๊าซออกซิเจนไปสู่เนื้อเยื่อต่าง ๆ ในร่างกายส่งผลต่อการลดลงของสมรรถภาพปอดและความสามารถในการทำกิจกรรม

พิมพ์พรรณ อัมพันธ์ทอง และปิยภัทร บุชบาบดินทร์ (2558) ได้ทำการศึกษาเรื่อง การพยากรณ์ PM10 ในบริเวณภาคเหนือตอนบนของประเทศไทยด้วยทฤษฎีเกรย์ โดยมีวัตถุประสงค์เพื่อ

พยากรณ์ปริมาณ PM10 ด้วยทฤษฎีระบบเกรย์ในบริเวณภาคเหนือตอนบนของประเทศไทย ประกอบด้วย 4 จังหวัด ได้แก่ จังหวัดเชียงใหม่ จังหวัดเชียงราย จังหวัดลำปาง และจังหวัดแม่ฮ่องสอน โดยใช้ข้อมูลปริมาณ PM10 รายเดือนตั้งแต่เดือนมกราคมถึงเดือนธันวาคมในแต่ละปี เริ่มตั้งแต่ปี พ.ศ.2551 ถึง 2556 ใช้ข้อมูลจากสถานีตรวจวัดคุณภาพอากาศ 4 สถานีในจังหวัดเชียงใหม่ จังหวัดเชียงราย จังหวัดลำปาง และจังหวัดแม่ฮ่องสอน ผลการศึกษาพบว่าคุณภาพอากาศในบริเวณที่ทำการศึกษามีปริมาณ PM10 สูงเกินกว่าเกณฑ์มาตรฐาน ซึ่งค่ามาตรฐานเฉลี่ยราย 24 ชั่วโมงไม่เกิน 120 ไมโครกรัมต่อลูกบาศก์เมตร ในส่วนของการพยากรณ์ปริมาณ PM10 ในบริเวณภาคเหนือตอนบนของประเทศไทยด้วยทฤษฎีเกรย์ GM (1,1) ถือเป็นทางเลือกที่ใช้ในการหาค่าพยากรณ์โดยพิจารณาข้อมูลในปี พ.ศ.2551-2556 จากสถานีตรวจวัดคุณภาพอากาศทั้ง 4 สถานี โดยค่าพยากรณ์ในอนาคต 4 ปีข้างหน้าคือ พ.ศ.2557 ถึง 2560 พบว่า ทฤษฎีเกรย์สามารถแสดงค่าพยากรณ์ได้ใกล้เคียงค่าความจริง (ในช่วง พ.ศ.2551-2556) และคุณภาพอากาศในบริเวณภาคเหนือตอนบนของประเทศไทยจะมีปริมาณ PM10 สูงในเดือนตุลาคมจนถึงเดือนเมษายนของทุกจังหวัด ยกเว้นในจังหวัดแม่ฮ่องสอน นอกจากนี้ การตรวจสอบความแม่นยำจากค่าทำนายเชื่อถือได้จากค่า MAE แสดงค่าใกล้เคียงกับค่าจริง ประกอบกับข้อมูลจริงของสองจังหวัด คือ จังหวัดแม่ฮ่องสอนมีชายแดนติดต่อกับประเทศเพื่อนบ้านมีการเกษตรแบบเดิม ๆ ทั้งในส่วนของประเทศไทยและประเทศเพื่อนบ้าน คือ การจุดไฟเผาหญ้าและการป้องกันไฟป่าที่ดี ทำให้เห็นว่ามลพิษทางอากาศปริมาณ PM10 มีแนวโน้มเพิ่มขึ้น ส่วนจังหวัดลำปาง พบว่าโรงงานไฟฟ้าแม่เมาะที่ใช้ถ่านหินเป็นพลังงานในกระบวนการผลิตกระแสไฟฟ้าซึ่งทุกครั้งมีการปล่อยกำมะถันประมาณร้อยละ 3-5 ก่อให้เกิด ปริมาณ PM10 สูงขึ้นทั้งนี้ จากค่าแนวโน้มของปริมาณ PM10 ที่เพิ่มขึ้น สามารถนำไปวางแผนในอนาคต เพื่อลดปัญหาคุณภาพอากาศในบริเวณภาคเหนือตอนบนของประเทศไทย

กาญจนา ทองบุญนาค และคณะ (2561) ได้ทำการศึกษาเรื่อง การพัฒนาแบบจำลองเพื่อพยากรณ์ปริมาณ PM10 ในพื้นที่จังหวัดเชียงใหม่ โดยใช้โครงข่ายเพอร์เซ็ปตรอนหลายชั้น โดยมีวัตถุประสงค์เพื่อ พัฒนาแบบจำลองเพื่อพยากรณ์ปริมาณ PM10 ในพื้นที่จังหวัดเชียงใหม่ โดยใช้โครงข่ายเพอร์เซ็ปตรอนหลายชั้น โดยใช้ข้อมูลปริมาณ PM10 จากสถานีตรวจวัดคุณภาพอากาศที่โรงเรียนยุพราชวิทยาลัย ซึ่งเป็นสถานีตรวจวัดคุณภาพอากาศของกรมควบคุมมลพิษ ในช่วงระยะเวลา 5 ปี ตั้งแต่ปี พ.ศ.2555 ถึง 2559 มาเป็นข้อมูลในการสร้างรูปแบบการพยากรณ์ให้มีความแม่นยำและมีประสิทธิภาพในการพยากรณ์โดยทำการปรับค่าพารามิเตอร์ของโครงข่ายประสาทเทียมให้มีความเหมาะสม จากการทดลองได้แสดงให้เห็นว่าจำนวน Hidden Node มีผลต่อความถูกต้อง

ของผลลัพธ์คือจำนวน Hidden Node มากจะทำให้ความถูกต้องมากขึ้น แต่ความเร็วในการ converge ก็จะใช้เวลามากเช่นกัน แต่เมื่อเปลี่ยนแปลงค่า Momentum Rate และ Learning Rate เพื่อทดสอบผลกระทบที่เกิดขึ้นโดย initialize weight แตกต่างกัน พบว่าความเร็วในการ converge จะใช้เวลาน้อยลงตามค่า Momentum Rate และ Learning Rate ที่เพิ่มขึ้นด้วย จากการทดสอบกับข้อมูลทั้งหมด สามารถแบ่งข้อมูลได้ 3 clusters และมีเปอร์เซ็นต์ความถูกต้องมากขึ้นเมื่อเทียบกับโครงข่ายประสาทเทียม (Neural Networks) ในรูปแบบของโครงข่ายเพอร์เซพตรอนแบบหลายชั้น (Multi-layer perceptron) โดยมีการเรียนรู้แบบ Back propagation จากการที่นำค่า Program Output ที่ได้จาก Neural Networks มาทำ Incremental Clustering ต่อนั้นทำให้ได้เปอร์เซ็นต์ความถูกต้องของผลลัพธ์เพิ่มมากขึ้น

ณัฐวุฒิ บางหลวง (2562) ได้ทำการศึกษาเรื่อง การพยากรณ์ราคามันสำปะหลังโดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น โดยมีจุดประสงค์เพื่อสร้างแบบจำลองการพยากรณ์สำหรับพยากรณ์ราคามันสำปะหลังในจังหวัดเชียงราย โดยใช้เทคนิคโครงข่ายประสาทเทียมหลายชั้น ได้ใช้ไลบรารี TensorFlow มาช่วยแก้ปัญหาในด้านโปรแกรมในการคำนวณของโครงข่ายประสาทเทียม และ ใช้ไลบรารี Keras ในการสร้างแบบจำลองโครงข่ายประสาทเทียมหลายชั้น เมื่อทำการทดลองพบว่าแบบจำลองโครงข่ายประสาทเทียมที่เหมาะสมกับการพยากรณ์ราคามันสำปะหลังมากที่สุด คือแบบจำลองที่มีจำนวนโหนดของข้อมูลนำเข้า (Input layer) 5 ข้อมูล ราคารับซื้อหัวมันสำปะหลัง ราคาเฉลี่ยที่รับซื้อหัวมันสำปะหลังที่มีปริมาณแบ่ง 25 และ 30 เปอร์เซ็นต์และ โรงงานที่รับซื้อมันสำปะหลังที่มีปริมาณแบ่ง 25 และ 30 เปอร์เซ็นต์ จำนวนชั้นซ่อน (Hidden layer) 3 ชั้น แต่ละชั้นซ่อนประกอบไปด้วยโหนดทั้งหมด (Hidden node) 8 โหนดและ มีชั้นข้อมูลออก (Output layer) 1 โหนด ให้ค่าความคลาดเคลื่อนสัมพัทธ์ที่น้อยที่สุดที่ 5.334 จำนวนรอบในการเรียนรู้ 150 รอบและเมื่อผู้วิจัยได้ทำการทดสอบจำนวนรอบที่เหมาะสมกับการเรียนรู้พบว่า จำนวนรอบที่ 2,000 นั้นให้ค่าความคลาดเคลื่อนเฉลี่ยสัมพัทธ์ที่น้อยที่สุดจากการทดลองจำนวนรอบตั้งแต่ 1,000 จนถึง 50,000 รอบให้ค่าความคลาดเคลื่อนสัมพัทธ์เฉลี่ยที่ 8.082 จากการใช้ข้อมูลในการเรียนรู้ 1 ปีแบ่งข้อมูลทดสอบไว้ 33 เปอร์เซ็นต์และ ทำการพยากรณ์ราคามันสำปะหลัง 2 เดือนสุดท้าย โดยใช้ข้อมูลที่ได้นี้มาจากสำนักเศรษฐกิจการเกษตรที่มีข้อมูลย้อนหลังและ ข้อมูลจากสมาคมโรงงานผลิตภัณฑ์มันสำปะหลังไทย ซึ่งข้อมูลย้อนหลังเพียง 1 ปี ในอนาคตถ้าได้เก็บข้อมูลไว้อย่างสม่ำเสมอก็จะช่วยให้มีข้อมูลในการเรียนรู้มากกว่าเดิม จะทำให้การพยากรณ์มีประสิทธิภาพมากขึ้นหรือถ้ามีปัจจัยใหม่ที่เกี่ยวข้องที่ทำให้การพยากรณ์ราคามันสำปะหลังเปลี่ยนไปก็ต้องทำการทดลองใหม่ทั้งหมดเพื่อหา

จำนวนชั้นซ้อน จำนวนโหนดในชั้นซ้อน และ จำนวนรอบในการเรียนรู้ที่เหมาะสม เพื่อสร้างโมเดล การพยากรณ์ราคามันสำปะหลังที่เหมาะสมกับการพยากรณ์

Pope & Dockery (2006) การศึกษาผลกระทบในระยะสั้นจากฝุ่นละอองขนาดเล็ก PM2.5 และ PM10 ต่ออัตราการตายรายวันของประเทศเนเธอร์แลนด์ โดยใช้ข้อมูลความเข้มข้นของฝุ่นละออง และข้อมูลอัตราการตายในปี 2008-2009 ใช้สถิติ Poisson regression analysis มีการควบคุมอิทธิพล ของข้อมูลด้าน long term ประกอบ ด้วยฤดูกาล อุบัติการณ์ของไข้หวัด ความแตกต่างด้านลักษณะ อุดมคติวันอาทิตย์ day of the week และช่วงสุดสัปดาห์ ผลการศึกษาพบว่าการเพิ่มขึ้นของความเข้มข้น ฝุ่นละอองขนาดเล็ก PM2.5 และ PM10 ทุก  $10 \mu\text{g}/\text{m}^3$  มีความสัมพันธ์กับทุกสาเหตุการตายและ การตายเฉพาะสาเหตุอย่างมีนัยสำคัญทางสถิติ (N.A.H.Janssen, 2013) สอดคล้องกับการศึกษาใน ระยะยาว ของ Pope & Dockery ได้ทำ systematic review เกี่ยวกับการตายที่มีความสัมพันธ์กับ การสัมผัสฝุ่นละอองใน ระยะยาว พบว่า การสัมผัสฝุ่นละอองขนาดเล็ก PM10 ที่เพิ่มขึ้นทุก  $20 \mu\text{g}/\text{m}^3$  เพิ่มความเสี่ยงในการ ตายจากทุกสาเหตุ เป็น 8 เท่า และ 2.1 เท่า เมื่อเทียบกับผู้ไม่ได้รับ สัมผัสฝุ่นละอองและมีความเสี่ยง ในการตายด้วยโรค Cardiopulmonary เพิ่มขึ้น 0.6 เท่า มีความเสี่ยงในการตายด้วยโรคมะเร็งปอด เป็น 81 เท่า เมื่อเทียบกับผู้ที่ไม่ได้รับสัมผัส

Byeong-Jae Lee (2014) การศึกษาของ Hyewon lee และคณะ ทำการศึกษาในเมือง 11 เมืองของสี่ประเทศคือ ญี่ปุ่น เกาหลี ไต้หวันและจีน แบบ time series analysis เพื่อหาความสัมพันธ์ ระหว่างฝุ่นละออง ขนาดเล็ก PM10 และ PM2.5 กับการตายด้วยโรกระบบทางเดินหายใจและโรค ระบบหัวใจและ หลอดเลือด ที่มีการบันทึกตามรหัสการจำแนกโรคระหว่างประเทศ ICD-9 และ ICD-10 โดยใช้ข้อมูล การตายที่ไม่ได้เกิดจากอุบัติเหตุ ICD-9 รหัส 001-799 ICD-10 รหัส A00-R99 โรค ระบบหัวใจและ หลอดเลือด ICD-9 รหัส 390-459 ICD-10 รหัส I00-I99 โรกระบบทางเดินหายใจ ICD-9 รหัส 460- 519 ICD-10 รหัส J00-J99 และมีการคาดการณ์ผลกระทบจากการทำ Me-ta analysis ผลการศึกษาพบว่าการเพิ่มขึ้นของความเข้มข้นของฝุ่นละอองขนาดเล็ก PM10 ทุก  $10 \mu\text{g}/\text{m}^3$  ใน 2 วันก่อนหน้า มีความสัมพันธ์กับการตายจากโรกระบบหัวใจและหลอดเลือดร้อยละ 0.69 อย่าง ไม่มีนัยสำคัญทางสถิติ และฝุ่นละอองขนาดเล็ก PM2.5 ที่เพิ่มขึ้นทุก  $10 \mu\text{g}/\text{m}^3$  ใน 2 วันก่อน หน้า มีความสัมพันธ์กับการเพิ่มขึ้นของอัตราการตายจากทุกสาเหตุร้อยละ 0.96 อย่างมีนัยสำคัญทาง สถิติ แต่พบความสัมพันธ์กับการเพิ่มขึ้นของการตายด้วยโรกระบบหัวใจและหลอดเลือดและโรกระบบ ทางเดินหายใจอย่างไม่มีนัยสำคัญทางสถิติ (Lee et al., 2015) ถึงอย่างไรก็ตามยังมีการศึกษาที่แสดง ถึงความสัมพันธ์อย่างมีนัยสำคัญทางสถิติ ระหว่างโรกระบบหัวใจและหลอดเลือดกับความเข้มข้นของ

ฝุ่นละอองขนาดเล็ก PM10 โดยพบว่าฝุ่นละอองขนาดเล็ก PM10 ที่เพิ่มขึ้นทุก 10  $\mu\text{g}/\text{m}^3$  ในวันก่อนหน้าที่จะเสียชีวิต มีความสัมพันธ์กับการเพิ่มขึ้นของการตายด้วยโรคหัวใจ (Cardiopulmonary) ร้อยละ 0.68 อย่างมีนัยสำคัญทางสถิติ