

การพัฒนาแบบจำลองเพื่อพยากรณ์ปริมาณ PM_{10} ในเขตกรุงเทพมหานคร โดยใช้โครงข่ายเรเดียลเบซิสฟังก์ชัน

เสาวลักษณ์ อร่ามพวงสานุวัต¹ และ พยุง มีสัจ²

¹ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ

²ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กรุงเทพฯ

Emails: saowalak.a@hotmail.com, pym@kmutnb.ac.th

บทคัดย่อ

บทความนี้เป็นการศึกษาแบบจำลองการพยากรณ์ปริมาณฝุ่นละอองขนาดเล็กไม่เกิน 10 ไมครอน (PM_{10}) ที่เป็นสารมลพิษทางอากาศที่เป็นปัญหาหลักในกรุงเทพมหานคร โดยใช้เทคนิคโครงข่ายประสาทเทียม (Neural Networks) ในรูปแบบของโครงข่ายเรเดียลเบซิสฟังก์ชัน (Radial Basis Function Network) เพื่อพยากรณ์ปริมาณ PM_{10} โดยอาศัยข้อมูลคุณภาพอากาศ ได้แก่ คาร์บอนมอนอกไซด์ (CO) โอโซน (O_3) ไนโตรเจนไดออกไซด์ (NO_2) ซัลเฟอร์ไดออกไซด์ (SO_2) ฝุ่นขนาดเล็กกว่า 10 ไมครอน (PM_{10}) และข้อมูลทางอุตุนิยมวิทยา ได้แก่ ปริมาณรังสีดวงอาทิตย์ (globe radiation) ปริมาณรังสีสุทธิ (net radiation) ความกดอากาศ (atmospheric pressure) ปริมาณน้ำฝน (rain) ความชื้นสัมพัทธ์ (relative humidity) อุณหภูมิ (temperature) ทิศทางลม (wind direction) และความเร็วลม (wind speed) โดยอาศัยข้อมูลจากสถานีดินแดง ซึ่งเป็นสถานีตรวจวัดคุณภาพอากาศของกรมควบคุมมลพิษ ในช่วงระยะเวลา 9 ปี ตั้งแต่ปี 2544-2552 มาเป็นข้อมูลในการสร้างรูปแบบการพยากรณ์

คำสำคัญ—การพยากรณ์; ฝุ่นขนาดเล็ก; โครงข่ายเรเดียลเบซิสฟังก์ชัน

1. บทนำ

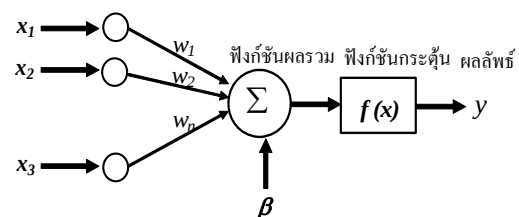
จากรายงานสถานการณ์คุณภาพอากาศของประเทศไทยโดยกรมควบคุมมลพิษ [1] ระบุว่าสารมลพิษทางอากาศที่เป็นปัญหาหลักที่สุดโดยเฉพาะในเขตเมืองใหญ่ คือฝุ่นละอองขนาดเล็กกว่า 10 ไมครอน (PM_{10}) เนื่องจากพบว่ามีระดับความเข้มข้นของฝุ่นสูง ในระดับที่อาจเกิดปัญหาสุขภาพกับประชาชนในเมืองใหญ่ โดยผลการตรวจวัดพบจำนวนครั้งที่เกินค่ามาตรฐาน 120 ไมโครกรัมต่อลูกบาศก์เมตร (มกก./ลบ.ม.) มากกว่าสารมลพิษทางอากาศชนิดอื่นๆ ซึ่งก่อให้เกิดปัญหาด้านสุขภาพแก่ประชากรทั้งผลระยะสั้นและเรื้อรัง เนื่องจากสามารถหายใจเข้าไปได้ (respirable particulate matter) ดังนั้นการติดตามตรวจสอบข้อมูลฝุ่นที่เป็นอันตรายต่อสุขภาพเช่นนี้ นับว่ามีความสำคัญมากในการวางแผน ป้องกัน แก้ไข และควบคุมมลภาวะทางอากาศให้ประสบผลสำเร็จในอนาคต [2] ซึ่งหน่วยงานภาครัฐ โดยกรมควบคุมมลพิษ ก็ได้ดำเนินการติดตามตรวจสอบคุณภาพอากาศของประเทศไทย โดยทำการตรวจวัดคุณภาพอากาศโดยใช้

สถานีตรวจวัดคุณภาพอากาศ แต่เนื่องจากสถานีตรวจวัดคุณภาพอากาศแบบถาวร มีข้อจำกัดในด้านการลงทุนและดำเนินการสูง จากปัญหาและข้อจำกัดดังกล่าว งานวิจัยนี้จึงจะทำการพยากรณ์สารมลพิษทางอากาศที่เป็นปัญหาหลัก คือฝุ่นละอองขนาดเล็กไม่เกิน 10 ไมครอน (PM_{10}) โดยพยากรณ์ปริมาณของ PM_{10} ในกรุงเทพมหานคร โดยใช้เทคนิคโครงข่ายเรเดียลเบซิสฟังก์ชัน (Radial Basis Function Network)

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 โครงข่ายประสาทเทียม (Artificial Neural Networks)

โครงข่ายประสาทเทียมมีการทำงานแบบขนาน เลียนแบบการประมวลผลของสมองมนุษย์ สามารถทำการเรียนรู้ชุดข้อมูลได้ทั้งแบบมีผู้สอน (Supervised Learning) และแบบไม่มีผู้สอน (Unsupervised Learning) [3] [4] โครงสร้างของโครงข่ายประสาทเทียมประกอบด้วยนิวรอนหลายนิวรอนเชื่อมต่อกันเป็นโครงข่าย การทำงานของแต่ละนิวรอนภายในโครงข่ายประสาทเทียมแสดงได้ดังรูปที่ 1



รูปที่ 1. การทำงานของนิวรอนภายในโครงข่ายประสาทเทียม

กำหนดให้ตัวแปร $x_1, x_2, \dots, x_n$ เป็นค่าข้อมูลเข้า ตัวแปร $w_1, w_2, \dots, w_n$ เป็นค่าน้ำหนัก ตัวแปร β เป็นค่าความโน้มเอียง ตัวแปร f เป็นฟังก์ชันกระตุ้น และตัวแปร y เป็นค่าผลลัพธ์จากนิวรอลดังสมการที่ 1

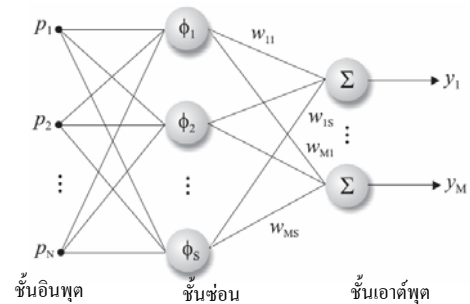
$$y = f[\sum(x_1w_1 + x_2w_2 + x_3w_3 + \dots + x_nw_n) + \beta] \quad (1)$$

ตัวอย่างการประยุกต์ใช้งานโครงข่ายประสาทเทียม [5] เช่น

- การจดจำรูปแบบ (pattern recognition) งานของการจดจำรูปแบบคือการกำหนดรูปแบบอินพุต (เช่น คลื่นเสียงพูด หรือลายมือเขียนตัวอักษร) ให้อยู่ในรูปของเวกเตอร์ลักษณะ (feature vector) ที่ซึ่งใช้แทนคลาสต่างๆ การประยุกต์ใช้งานที่เป็นที่รู้จักกันดี เช่น การจดจำลายมือตัวอักษร การจดจำเสียง
- การจับกลุ่ม (clustering/categorization) โดยปกติแล้ว ในงานการจับกลุ่มจะไม่มีข้อมูลล่วงหน้าให้สำหรับการฝึกสอน อัลกอริทึมการจับกลุ่มจะทำการค้นหาสถานะคล้าย (similarity) ระหว่างข้อมูลรูปแบบ และทำการจับกลุ่มรูปแบบที่คล้ายกันไว้ด้วยกัน
- การประมาณค่าฟังก์ชัน (function approximation) พิจารณาข้อมูลรูปแบบสำหรับการเรียนรู้ N ชุด ในรูปของ {เวกเตอร์อินพุต, เวกเตอร์เป้าหมาย} คือ $\{x_i, y_i\}, \{x_2, y_2\} \dots \{x_N, y_N\}$ ข้อมูลดังกล่าวได้มาจากฟังก์ชัน $y = f(x)$ ที่ไม่ทราบค่า การประมาณค่าฟังก์ชันคือการหาค่าประมาณ $\hat{y} = \hat{f}(x)$ ของฟังก์ชัน $f(x)$ ที่ซึ่งให้เอาต์พุตตามข้อมูลรูปแบบสำหรับการเรียนรู้
- การทำนาย (prediction/forecasting) กำหนดชุดค่าตัวอย่าง N ค่า คือ $\{y(t_1), y(t_2), \dots, y(t_N)\}$ เป็นค่า ณ เวลา $t_1, t_2, \dots, t_N$ ตามลำดับ การทำนายผลล่วงหน้า คือการทำนายค่าของ $y(t_{N+1})$ ซึ่งเป็นค่า ณ เวลา t_{N+1} ในอนาคต การทำนายผลมีการนำไปใช้กันมากในงานการตัดสินใจทางธุรกิจ วิทยาศาสตร์และวิศวกรรม ตัวอย่างการประยุกต์ใช้งาน เช่น การทำนายผลดัชนีตลาดหุ้น หรือการพยากรณ์อากาศล่วงหน้า
- การหาค่าเหมาะที่สุด (optimization) ปัญหาหลาย ๆ อย่างทางด้านคณิตศาสตร์ สถิติ วิศวกรรม วิทยาศาสตร์ การแพทย์และเศรษฐกิจ สามารถพิจารณาเป็นปัญหาการหาค่าเหมาะที่สุดได้ เป้าหมายของอัลกอริทึมการหาค่าเหมาะที่สุด คือการหาคำตอบที่เป็นไปตามเงื่อนไขข้อบังคับ และให้ค่าการประเมินจากฟังก์ชันวัตถุประสงค์ที่ดีที่สุด (สูงสุดหรือต่ำสุด)

2.2 โครงข่ายเรเดียลเบซิสฟังก์ชัน (Radial Basis Function Network)

โครงข่ายเรเดียลเบซิสฟังก์ชัน (Radial Basis Function Network หรือ RBF Network) เป็นโครงข่ายไปข้างหน้า (Feed Forward Network) ประเภทหนึ่ง ที่ได้รับการยอมรับว่ามีประสิทธิภาพสูงโครงข่ายหนึ่ง โครงข่าย RBF แตกต่างไปจากโครงข่ายเพอร์เซ็ปตรอนแบบหลายชั้น (Multi-layer Perceptron) ตรงที่โครงข่าย RBF นั้นมีชั้นซ่อนเร้นเพียงชั้นเดียว Broomhead และ Lowe [6] ถือเป็นผู้บุกเบิกนำเอา RBF มาประยุกต์ใช้ โครงข่ายเรเดียลเบซิสฟังก์ชันสามารถพิจารณาเป็นฟังก์ชันการส่ง (mapping function) ของความสัมพันธ์ระหว่างคู่รูปแบบอินพุตและเอาต์พุต ได้โดยการเรียนรู้ของโครงข่ายเป็นการปรับค่าน้ำหนักประสาทให้ได้ ฟังก์ชันการส่งที่เหมาะสมที่สุด



รูปที่ 2. เครือข่าย RBF

ดังนั้นสามารถกล่าวได้ว่าโครงข่าย RBF คือกระบวนการปรับเส้นโค้ง (curve fitting) ระหว่างข้อมูลอินพุตกับเอาต์พุตนั่นเอง โครงข่ายเรเดียลเบซิสฟังก์ชันจึงเหมาะในงานการประมาณค่าฟังก์ชัน และสามารถนำไปประยุกต์ใช้งานได้อย่างหลากหลาย

2.2.1 สถาปัตยกรรมของโครงข่าย RBF

โครงสร้างของโครงข่าย RBF โดยทั่วไปจะประกอบไปด้วยชั้นของนิวรอน 3 ชั้น [7] ได้แก่ ชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุต โดยที่ในชั้นอินพุตนั้น แต่ละอินพุตจะแทนคุณลักษณะของเวกเตอร์อินพุตเหมือนกับในโครงข่ายเพอร์เซ็ปตรอนแบบหลายชั้นทั่วๆ ไป ในที่นี้เวกเตอร์อินพุตมีขนาดเท่ากับ N และในชั้นซ่อน แต่ละนิวรอนในชั้นซ่อนจะมีฟังก์ชันถ่ายโอนซึ่งมีลักษณะพิเศษ ที่ซึ่งให้ผลตอบสนองของฟังก์ชันที่ขึ้นอยู่กับระยะห่างระหว่างอินพุตกับจุดศูนย์กลางของฟังก์ชัน กล่าวคือถ้าเวกเตอร์อินพุตอยู่ใกล้จุดศูนย์กลางมาก เอาต์พุตที่ได้จะมาก ถ้าเวกเตอร์อินพุตอยู่ห่างออกจากจุดศูนย์กลาง เอาต์พุตที่ได้จะลดลงตามลำดับ ในที่นี้จำนวนนิวรอนในชั้นซ่อนมีขนาดเท่ากับ S และในชั้นเอาต์พุต มีหน้าที่รวมเอาต์พุตที่ได้จากแต่ละนิวรอนในชั้นซ่อน โครงข่ายจะให้เอาต์พุตในรูปของเวกเตอร์ขนาดเท่ากับ M ดังรูปที่ 2

ดังนั้นจึงสามารถพิจารณาโครงข่าย RBF ว่าเป็นฟังก์ชันการส่งระหว่างปริภูมิของอินพุต $p \in R^{N \times 1}$ ไปยังปริภูมิของเอาต์พุต $y \in R^{M \times 1}$ ได้จากโครงข่าย RBF ในรูปข้างต้นจะได้ว่าเอาต์พุตของโครงข่ายมีค่าเท่ากับสมการ 2 และสมการ 3

$$y_i = \sum_{k=1}^S w_{ik} \phi_k(p, c_k) \quad (2)$$

$$y_i = \sum_{k=1}^S w_{ik} \phi_k(\|p, c_k\|) \quad (3)$$

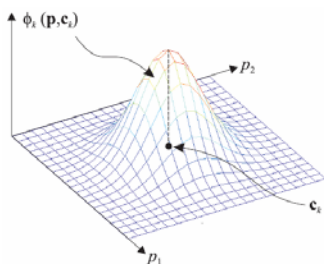
โดยที่ $\phi_k(\cdot)$ คือ ฟังก์ชันส่งค่าจาก R^+ ไปยัง R

$\|\cdot\|$ คือ ฟังก์ชันระยะทางแบบยูคลิด (Euclidean Norm)

w_{ik} คือ ค่าน้ำหนักประสาทในชั้นซ่อน

S คือ จำนวนนิวรอนในชั้นซ่อน

$c_k \in R^{N \times 1}$ คือ เวกเตอร์จุดศูนย์กลางของ RBF ในปริภูมิของเวกเตอร์อินพุต



รูปที่ 3. Radial Basis Function แบบเกาส์เซียน (แสดงในรูปแบบ 2 มิติ)

สำหรับแต่ละนิวรอนในชั้นซ่อน ค่าระยะทางยูคลิดระหว่างเวกเตอร์จุดศูนย์กลาง C_k กับเวกเตอร์อินพุต p จะถูกคำนวณ เอาต์พุตของนิวรอนในชั้นซ่อนนี้จะได้จากฟังก์ชัน $\phi_k(\cdot)$ ซึ่งเป็นฟังก์ชันแบบไม่เป็นเชิงเส้น สุดท้ายแล้วเอาต์พุตของโครงข่ายจะได้จากผลรวมของค่าน้ำหนักประสาทดกับเอาต์พุตของนิวรอนจากชั้นซ่อน ตัวอย่างของฟังก์ชัน $\phi_k(\cdot)$ ที่ใช้ในโครงข่าย RBF เช่น ฟังก์ชันเกาส์เซียน (Gaussian Function)

$$\text{Gaussian Function: } \phi(p) = e^{-p^2/\sigma^2} \quad (4)$$

โดยฟังก์ชันเกาส์เซียนเป็นฟังก์ชันที่นิยมใช้ใน RBF มากที่สุด ซึ่งรูปร่างของฟังก์ชันเกาส์เซียนแสดงได้ดังรูปที่ 3 โดยที่พารามิเตอร์ σ เป็นตัวควบคุมความกว้างของ RBF หรือเรียกว่าพารามิเตอร์การกระจาย (spread parameter)

เวกเตอร์ศูนย์กลาง C_k ของนิวรอนซ่อนตัวที่ k จะรับอินพุตจากเวกเตอร์ p ที่มีมิติเท่ากับ N พารามิเตอร์ σ_k ทำหน้าที่ควบคุมความกว้างของแต่ละ RBF โดยปกติแล้ว ถ้าเวกเตอร์อินพุต p มีระยะห่างจากจุดศูนย์กลาง C_k มากขึ้น กล่าวคือ $\|p - C_k\|$ มีค่ามากขึ้น ค่าที่ได้จากฟังก์ชัน ϕ_k จะลดลง ดังรูปที่ 3 โดยพื้นที่ของฟังก์ชัน ϕ เรียกว่าสนามรับ (receptive field) ของนิวรอนนั้นๆ [8] เอาต์พุต y_i ของโครงข่ายได้จากผลรวมของเอาต์พุตของฟังก์ชัน ϕ จากทุกนิวรอนในชั้นซ่อน โดยปกติแล้วเวกเตอร์จุดศูนย์กลาง C_k จะถูกเลือกจากปริภูมิของเวกเตอร์อินพุต ที่ซึ่งจะต้องมีจำนวนเวกเตอร์จุดศูนย์กลาง (หรือจำนวนนิวรอนในชั้นซ่อน) เพียงพอและครอบคลุมปริภูมิของอินพุตได้

2.2.2 การฝึกสอน โครงข่าย RBF

การฝึกสอนโครงข่าย RBF นั้นได้มีผู้นำเสนอไว้หลายรูปแบบ ในลักษณะเดียวกันกับการฝึกสอนโครงข่ายประสาทเทียมแบบอื่น ๆ การฝึกสอนโครงข่าย RBF ก็คือการค้นหาพารามิเตอร์ของโครงข่ายซึ่งประกอบไปด้วย

- ค่าน้ำหนักประสาทด w_{ik} , $i=1, \dots, M$, $k=1, \dots, S$
- เวกเตอร์จุดศูนย์กลางของนิวรอนในชั้นซ่อน C_k , $k=1, \dots, S$
- ค่าพารามิเตอร์การกระจาย $\sigma_{\phi k}$, $k=1, \dots, S$

ในงานวิจัยนี้ใช้วิธีการฝึกสอนแบบจุดศูนย์กลางคงที่ (fixed center) ซึ่งในการฝึกสอนแบบจุดศูนย์กลางคงที่ เวกเตอร์จุดศูนย์กลางจะถูกสุ่มมาจาก

เวกเตอร์อินพุต และจะไม่มีการเปลี่ยนแปลงตำแหน่งของจุดศูนย์กลางในระหว่างการฝึกสอน ในขณะที่การฝึกสอนแบบแบบเกรเดียนต์เฟ้นสุ่ม (stochastic gradient) จะมีการปรับพารามิเตอร์ของโครงข่ายทั้งหมด [9] และในการกำหนดจุดศูนย์กลาง RBF ให้คงที่ การฝึกสอนโครงข่ายจะให้ผลลัพธ์เป็นผลเฉลยรูปแบบปิด ดังนั้นการฝึกสอนจะสามารถทำได้อย่างรวดเร็ว จึงทำให้มีผู้สนใจนำเอาโครงข่าย RBF ไปใช้งานอย่างมากมาย

2.3 งานวิจัยที่เกี่ยวข้อง

มีงานวิจัยที่ได้พัฒนาแบบจำลองเพื่อทำการพยากรณ์ทางด้านคุณภาพอากาศ (Air Quality) เช่น งานวิจัยของ Chaloulakou และคณะ [10] ได้พัฒนาแบบจำลองสำหรับการทำนายปริมาณของค่าโอโซน (Ozone) ในเมือง Athens ประเทศกรีซ โดยใช้ความเร็วลม ทิศทางลม และความชื้นสัมพัทธ์เป็นตัวแปรอิสระในการวิเคราะห์และศึกษาเปรียบเทียบแบบจำลองที่สร้างด้วยเทคนิควิธี Neural Network และวิธี Multiple Regression ซึ่งผลที่ได้จากแบบจำลองด้วยวิธี Neural Network มีประสิทธิภาพดีกว่าวิธี Multiple Regression และงานวิจัยของ Jiang และคณะ [11] ได้พัฒนา Neural Network แบบ Multilayer Perceptron โดยมีวิธีการเรียนรู้แบบ Back-propagation สำหรับการนำมาใช้ในการทำนายดัชนีมลพิษอากาศ (Air Pollution Index, API) เป็นรายวัน ในเมืองเซี่ยงไฮ้ ประเทศจีน โดยที่ข้อมูลป้อนเข้าในแบบจำลองจะเป็นข้อมูลทางอุตุนิยมวิทยา คือ ความดันบรรยากาศ อุณหภูมิ ความชื้นสัมพัทธ์ ทิศทางลม และ ปริมาณน้ำฝน ซึ่งสามารถทำนายค่า PM_{10} , SO_2 และ NO_2 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ที่ 0.6056, 0.6993 และ 0.6300 ตามลำดับ นอกจากนี้งานวิจัยของ Kukkonen และคณะ [12] ยังได้ใช้โครงข่ายประสาทเทียมเพื่อทำนายการเกิด NO_2 และ PM_{10} ที่เมืองเฮลซิงกิ ประเทศฟินแลนด์ โดยนำข้อมูลปริมาณการจราจรในพื้นที่และข้อมูลทางด้านอุตุนิยมวิทยา ในช่วงปี 1996 ถึง 1999 พบว่าการใช้โครงข่ายประสาทเทียมสามารถให้ค่าอยู่ในช่วงที่ดี โครงข่ายประสาทเทียมจึงเหมาะสมที่จะนำมาพัฒนาแบบจำลองในการทำนายค่าระดับมลพิษอากาศในเขตเมือง

นอกจากนี้ยังมีงานวิจัยที่ได้มีการนำโครงข่าย Radial Basis Function มาประยุกต์ใช้ในงานต่างๆ เช่น Mellit และคณะ [13] ได้ทำการวิจัยโดยออกแบบนิวรอลเน็ตเวิร์ค เพื่อใช้ในการวิจัยการแผ่รังสีพลังงานแสงอาทิตย์ ด้วยวิธี Radial Basis Function (RBF) ที่มีนิวรอน 9 โหนดและ 1 ชั้นซ่อน ข้อมูลที่ใช้เทรนเครือข่ายเป็นข้อมูลค่าความเข้มการแผ่รังสีดวงอาทิตย์ในแต่ละวันกับค่าอุณหภูมิของอากาศในช่วงเวลานั้น ซึ่งวิธีการพิจารณาจะเป็นการพยากรณ์แบบ Time series จำนวนของข้อมูลที่ใช้เทรนระบบมีจำนวน 300 ค่า ซึ่งเป็นค่าของข้อมูลจริงที่จัดเก็บได้ บริเวณประเทศแอลจีเรียในช่วงปี ค.ศ. 1980-2000 โดยกำหนดค่าเฉลี่ยความผิดพลาดระบบหรือค่าผิดพลาดที่เป็นผลต่างระหว่างค่าจริงที่วัดได้ กับค่าที่ได้จากการพยากรณ์ไว้ไม่เกิน 1.5% และค่า Correlation coefficient สำหรับ Validation Data ที่ 98.9% ผลที่ได้ค่าความผิดพลาดเฉลี่ยอยู่ที่ 1.34% และงานวิจัยของ Declercq และคณะ [14] ได้ทำการวิจัยโดยทำการเปรียบเทียบชนิดของ

นิวรอลเน็ตเวิร์กที่นำมาใช้ในการพยากรณ์ 3 แบบ ได้แก่ Feed forward neural network แบบ Radial Based neural network และแบบ Elman neural network โดยการเทรนด้วยข้อมูลเป็นขั้นๆและข้อมูลแบบสุ่ม แบบ Feed forward neural network จะเลือก 1 ชั้นซ่อน และใช้ฟังก์ชันอินยัยแบบ Tanh เอคต์พุตแบบลิเนียร์ Feed forward neural network ใช้การเทรนแบบ Back propagation training rule จะได้ผลดีในเรื่องของเวลาที่ใช้เทรน การเริ่มต้นเทรนด้วยค่าที่ไม่เหมาะสมอาจเป็นผลให้ได้ค่าที่ผิดพลาดสูง โดยวิธีแก้ไขต้องกำหนดค่าเริ่มต้นของค่าน้ำหนักและค่าไบอัส สำหรับ Radial Basis neural network เลือกชั้นซ่อน 1 ชั้น และใช้ฟังก์ชันอินยัยแบบ Gaussian มีเอคต์พุตแบบลิเนียร์ เวลาที่ใช้ในการเทรนจะเร็วกว่า ส่วนเน็ตเวิร์กแบบ Elman neural network จะมีลักษณะเช่นเดียวกับ Feed forward neural network ใช้หลักการเทรนแบบ Back propagation เพียงแต่จะมีการป้อนกลับเพิ่มเข้ามาในชั้นของฮิดเด็น แบบนี้จะใช้เวลานานมากในการเทรนและค่าความผิดพลาดมีค่าสูงกว่าสองแบบแรก

3. วิธีการดำเนินการวิจัย

3.1 การรวบรวมและออกแบบการจัดรูปแบบข้อมูล

3.1.1 การรวบรวมและวิเคราะห์ชุดข้อมูล

ข้อมูลที่ใช้ในการวิจัย ได้แก่ คาร์บอนมอนอกไซด์ (CO) โอโซน (O₃) ไนโตรเจนไดออกไซด์ (NO₂) ซัลเฟอร์ไดออกไซด์ (SO₂) ฝุ่นขนาดเล็กกว่า 10 ไมครอน (PM₁₀) ปริมาณรังสีดวงอาทิตย์ (globe radiation) ปริมาณรังสีสุทธิ (net radiation) ความกดอากาศ (atmospheric pressure) ปริมาณน้ำฝน (rain) ความชื้นสัมพัทธ์ (relative humidity) อุณหภูมิ (temperature) ทิศทางลม (wind direction) และความเร็วม (wind speed) โดยข้อมูลที่นำมาใช้เป็นข้อมูลที่ได้จากสถานีดินแดง ซึ่งเป็นสถานีตรวจวัดคุณภาพอากาศของกรมควบคุมมลพิษ ในช่วงระยะเวลา 9 ปี ตั้งแต่วันที่ 1 มกราคม 2544 ถึง 31 ธันวาคม 2552 โดยเป็นข้อมูลอนุกรมเวลาที่มีลักษณะแบบไม่เชิงเส้น และทำการแบ่งชุดข้อมูลที่ใช้สำหรับการเรียนรู้ (Training Set) คิดเป็น 70% ของข้อมูลทั้งหมด เพื่อให้แบบจำลองได้เรียนรู้จากชุดข้อมูลจริง และข้อมูลชุดทดสอบ (Test Set) คิดเป็น 30% ของข้อมูลทั้งหมด

3.1.2 การแปลงค่าข้อมูล (Data transformation)

เป็นกระบวนการปรับขอบเขตของข้อมูล ให้อยู่ในช่วงที่เหมาะสมต่อการนำไปใช้งานในการสอนให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ สำหรับในโปรแกรมที่ใช้หลักการของโครงข่ายประสาทเทียมนั้น จะต้องแปลงค่าข้อมูลด้วยวิธีการนอร์มัลไลซ์ (Normalization) ซึ่งเป็นการลดค่าของข้อมูลให้อยู่ในขอบเขตที่น้อยลง เพื่อให้เหมาะสมกับฟังก์ชันที่ใช้งานของโครงข่ายประสาทเทียม

3.2 ขั้นตอนการสร้างแบบจำลองการพยากรณ์

สำหรับงานวิจัยนี้ มีลำดับขั้นตอนการทำงาน ดังนี้

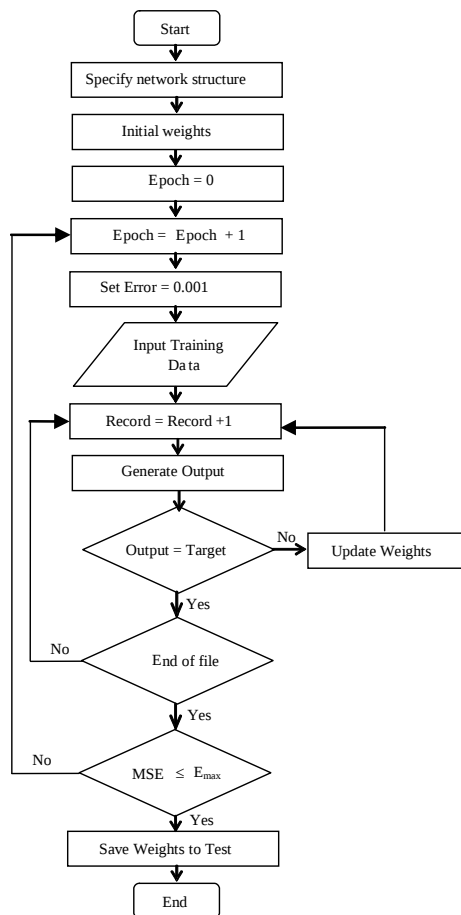
1. กำหนดให้ $X(t) = \{x_1, x_2, \dots, x_n\}$ คือ เซตของตัวแปรข้อมูลเข้าที่เวลา t เมื่อ n คือจำนวนตัวแปรเข้า $Y(t+1) = \{y_t\}$ คือเซตของข้อมูลตัวแปรผลลัพธ์ที่เวลา $t+1$ ทำการแบ่งข้อมูลเป็น 2 ชุด คือ ชุดสอน (Train set) และชุดทดสอบ (Test set)
2. ทำการเตรียมข้อมูล (preprocessing) โดยทำการแปลงข้อมูล (normalization) ให้อยู่ในช่วง $[1, 1]$
3. ทำการสอนโครงข่ายประสาทเทียม (neural network training) โดยใช้ตัวแปรข้อมูลเข้า $X(t)$ และตัวแปรผลลัพธ์ $Y(t+1)$ โดยใช้ข้อมูลชุดสอน
4. ทำการพยากรณ์โดยใช้ข้อมูลชุดทดสอบ และวัดค่าคลาดเคลื่อนเฉลี่ยกำลังสอง (Mean square error: MSE) เพื่อวัดความคลาดเคลื่อนของแบบจำลองในการพยากรณ์

3.3 การทดลองหาโครงสร้างของแบบจำลองโครงข่ายประสาทเทียมที่เหมาะสม ด้วยโปรแกรม MATLAB

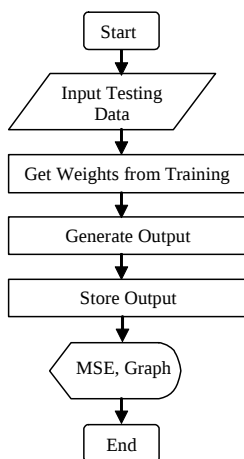
งานวิจัยนี้ได้ทำการสร้างแบบจำลองโครงข่ายประสาทเทียมเพื่อพยากรณ์ค่าของ PM₁₀ โดยที่ชนิดของฟังก์ชันที่ใช้ในแบบจำลอง คือ เรเดียลเบซิสฟังก์ชัน (Radial Basis Function) และได้ใช้วิธี Least Mean Squares ในการคัดเลือกข้อมูลนำเข้าแบบจำลอง ซึ่งโครงข่ายเรเดียลเบซิสฟังก์ชันที่ใช้ในงานวิจัยนี้ได้ใช้ชั้นซ่อน (hidden layer) จำนวน 1 ชั้น และใช้ฟังก์ชันอินยัยแบบ Gaussian มีเอคต์พุตแบบ linear โดยขั้นตอนที่ใช้ในการพยากรณ์ข้อมูลด้วยโครงข่ายเรเดียลเบซิสฟังก์ชัน มีดังนี้

1. กำหนดค่าเริ่มต้นของระบบ ได้แก่ จำนวนค่าข้อมูลในอดีตที่มีผลต่อข้อมูลเอคต์พุตของการพยากรณ์ข้อมูลด้วยรูปแบบเรเดียลเบซิสฟังก์ชัน
2. กำหนดข้อมูลอินพุตให้กับโครงข่าย $x = [x_1 \ x_2 \ \dots \ x_p]^T$ ซึ่งเป็นข้อมูลอินพุตของเรเดียลเบซิสฟังก์ชัน
3. แปลงข้อมูลแบบไม่เชิงเส้นจากอินพุต โดยในงานวิจัยนี้ เลือกฟังก์ชันที่ใช้ในการแปลงแบบเกาส์เซียน (Gaussian Function)
4. ฝึกหัดโครงข่ายเพื่อหาค่าพารามิเตอร์ที่เหมาะสมกับข้อมูล ในงานวิจัยนี้ใช้การฝึกหัดโครงข่ายด้วยอัลกอริทึมแบบ LMS (Least Mean Square) เพื่อคำนวณค่าเซ็นเตอร์และค่าน้ำหนักของโครงข่าย
5. คำนวณค่าเอคต์พุตของโครงข่าย
6. คำนวณค่าความคลาดเคลื่อนที่ได้จากการพยากรณ์

โดยขั้นตอนการทดสอบโครงข่าย (Testing Process) เริ่มจากการนำเข้าชุดข้อมูลทดสอบ และทำการคำนวณค่าผลลัพธ์จากค่าน้ำหนักที่ได้จากการสอน (Training) ซึ่งมีขั้นตอนการฝึกหัดโครงข่าย ดังรูปที่ 4 และการทดสอบโครงข่าย (Testing Process) มีขั้นตอนการทำงาน ดังรูปที่ 5



รูปที่ 4. การฝึกหัดโครงข่าย (Training Process)



รูปที่ 5. การทดสอบโครงข่าย (Testing Process)

4. ผลการดำเนินงาน

จากการทำงานในการพยากรณ์ PM_{10} ที่ใช้ข้อมูล ณ เวลา t เป็นข้อมูลเข้า และใช้ข้อมูลที่เวลาถัดไป $t+1$ เป็นข้อมูลผลลัพธ์ โดยทำการสอนโครงข่ายประสาทเทียมด้วยเรเดียลเบสฟังก์ชัน (Radial Basis Function) มีรายละเอียดดังต่อไปนี้

ตารางที่ 1. ผลการสร้างโมเดลพยากรณ์ PM_{10} ด้วยโครงข่าย RBF

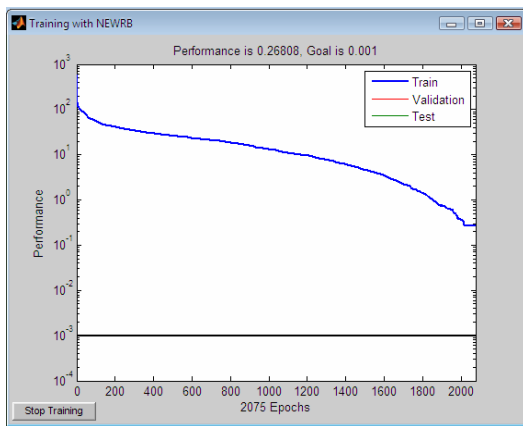
Radial Basis Function (radbas), Epoch = 5000 , Goals = 0.001		
Round Number	Spread	Performance (MSE)
1	0.1	$1.047916 \times 10^{-25} *$
2	0.5	2.779080×10^{-19}
3	1.0	0.0011868
4	1.5	5.420362×10^{-9}
5	2.0	0.001186
6	2.5	1.226605
7	3.0	9.1109157
8	3.5	27.080763

หมายเหตุ: * คือ Network Model ที่ให้ค่า MSE ต่ำที่สุด

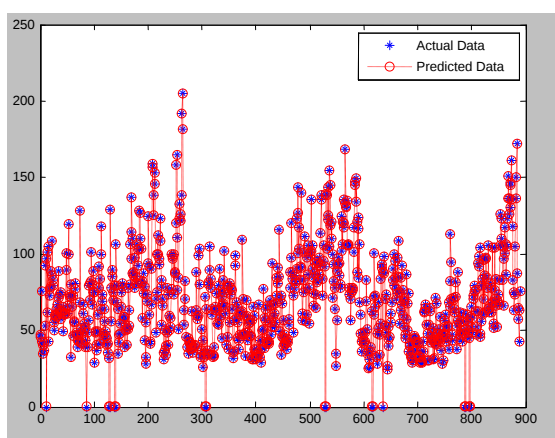
การสร้างแบบจำลองโครงข่ายประสาทเทียม ได้ทำการทดลองโดยกำหนดให้ชั้นข้อมูลเข้า (Input Layer) มีจำนวน 12 โหนด ส่วนในชั้นซ่อน (Hidden Layer) มีจำนวน 12 โหนด และชั้นผลลัพธ์ (Output Layer) จำนวน 1 โหนด โดยในการทดลองได้มีการกำหนดรูปแบบฟังก์ชันกระตุ้นชั้นซ่อน (Hidden Layer) แบบ Radial Basis Function และใช้ฟังก์ชันกระตุ้นชั้นแสดงผล (Output Layer) แบบ Linear โดยกำหนดจำนวนรอบในการฝึกสอน (Epochs) ไว้เท่ากับ 5,000 รอบ และกำหนดค่าความคลาดเคลื่อนเป้าหมาย (Goals) ไว้เท่ากับ 0.001

จากการทดลองสามารถแสดงผลการสร้างโมเดลพยากรณ์ ได้ดังตารางที่ 1 โดยพบว่าแบบจำลองโครงข่าย ซึ่งกำหนดให้มีชั้นข้อมูลเข้า (Input Layer) 12 โหนด ในชั้นซ่อน (Hidden Layer) ใช้ฟังก์ชันกระตุ้นแบบ Radial Basis Function จำนวน 12 โหนด ซึ่งกำหนดให้ค่าเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.1 และจำนวน 1 โหนดในชั้นแสดงผล (Output Layer) ใช้ฟังก์ชันกระตุ้นแบบ purelin (Linear) ให้ค่า MSE (Mean Square Error) เท่ากับ 1.047916×10^{-25} ซึ่งเป็นแบบจำลองที่ให้ค่าความผิดพลาดที่น้อยที่สุดเมื่อเทียบแบบจำลองอื่น ๆ ที่ได้ทำการทดลอง ดังตารางที่ 1

ซึ่งจากการทดลองที่ได้มีผลการเรียนรู้ข้อมูล แสดงได้ดังรูปที่ 6 และสามารถแสดงผลการเปรียบเทียบค่าของ PM_{10} ที่ได้จากการพยากรณ์โดยเทคนิคโครงข่ายประสาทเทียม (Predicted Data) กับค่าจริง (Actual Data) ได้ดังรูปที่ 7



รูปที่ 6. ผลการเรียนรู้ข้อมูล (Training data)



รูปที่ 7. การเปรียบเทียบค่าที่ได้จากการพยากรณ์กับค่าจริง

5. บทสรุป

งานวิจัยนี้ได้ทำการพัฒนาแบบจำลองเพื่อพยากรณ์ปริมาณ PM_{10} โดยใช้เทคนิคโครงข่ายประสาทเทียมในรูปแบบโครงข่ายเรเดียลเบสฟังก์ชัน (Radial Basis Function Networks) ซึ่งผลการดำเนินงานวิจัยพบว่าโครงข่ายประสาทเทียมสำหรับรูปแบบข้อมูล 12-12-1 ให้ผลการพยากรณ์คลาดเคลื่อนน้อยที่สุด โดยมีค่าคลาดเคลื่อนเฉลี่ยกำลังสอง (Mean Square Error: MSE) เท่ากับ 1.047916×10^{-25} ให้ผลเป็นที่น่าพอใจ ซึ่งแสดงว่าแบบจำลองโครงข่ายเรเดียลเบสฟังก์ชัน สามารถใช้ในการพยากรณ์ปริมาณ PM_{10} ได้เป็นอย่างดี ซึ่งข้อเสนอแนะสำหรับการวิจัยต่อไปในอนาคตนั้น อาจนำเทคนิคการสกัดตัวแปรข้อมูลเข้ามาใช้ เพื่อที่จะได้นำตัวแปรที่เหมาะสมและมีผลกระทบต่อกำหนดค่าของ PM_{10} โดยตรง มาใช้ในการสร้างแบบจำลองเพื่อพยากรณ์ปริมาณของ PM_{10} ต่อไป

6. กิตติกรรมประกาศ

ขอขอบคุณ สำนักจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม ที่กรุณาให้ความเคราะห์ในการใช้ข้อมูลสำหรับการวิจัย

7. เอกสารอ้างอิง

- [1] กรมควบคุมมลพิษ. สำนักจัดการคุณภาพอากาศและเสียง. **สถานการณ์และการจัดการปัญหามลพิษทางอากาศและเสียง ปี 2550**. กรุงเทพฯ : กรมควบคุมมลพิษ, 2551.
- [2] กรมควบคุมมลพิษ. สำนักจัดการคุณภาพอากาศและเสียง. **คู่มือการตรวจวัดฝุ่นละอองในบรรยากาศ**. กรุงเทพฯ: กรมควบคุมมลพิษ, 2551.
- [3] Richard J. Roiger, and Michael W. Geatz, "Data mining a tutorial based primer," Pearson Education, Inc., 2003.
- [4] Lippmann, R.P., "An introduction to computing with neural nets," IEEE ASSP Magazine, pp. 4–22, 1987.
- [5] อาทิตย์ ศรีแก้ว. **ปัญญาเชิงคำนวณ Computational Intelligence**. พิมพ์ครั้งที่ 1. สาขาวิชาวิศวกรรมไฟฟ้า สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี, 2552.
- [6] D. S. Broomhead and D. Lowe. "Multivariable functional interpolation and adaptive networks". Complex Systems, volume 2, pages 269–303, 1988.
- [7] M. M. Gupta, L. Jin and N. Homma, "Static and Dynamic Neural Networks: From Fundamentals to Advanced Theory", IEEE Press, New Jersey: John Wiley & Sons, Inc., pp. 81, 2003.
- [8] P.D. Wasserman. "Advanced Methods in Neural Computing". Van Nostrand Reinhold, New York, 1993.
- [9] F. M. Ham and I. Kostanic. "Principles of neurocomputing for science & engineering". McGraw-Hill, 2001.
- [10] Chaloulakou A, Saisana M, Spyrellis N., "Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens," The Science of the Total Environment, 313: pp. 1-13, 2003.
- [11] Jiang, D., et al., "Progress in developing an ANN model for air pollution index forecast," Atmospheric Environment 38: pp. 7055-7064, 2004.
- [12] Kukkonen, J., et al., "Extensive evaluation of neural network models for the prediction of NO_2 and PM_{10} concentrations, compared with a deterministic modeling system and measurements in central Helsinki," 2003.
- [13] Mellit, A., Benghanem, M. & Bendekhis, M., "Artificial neural network model for prediction solar radiation data: Application for sizing stand-alone photovoltaic power system", In Power Engineering Society, General Meeting, IEEE, June 12–16, 2005.
- [14] F. Declercq and R. De Keyser., "Comparative Study of Neural Predictors in Model Base Predictive control", Neural Networks for Identification, Control, Robotics, and Signal/Image Processing, 1996. Proceedings., pp.20-28, 1996.