บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน (Particulate Matter 2.5 – PM 2.5)

ฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน[3] (Particulate Matter 2.5) หรือ PM 2.5 คือ ฝุ่นละอองขนาดเล็กเทียบได้ว่ามีขนาดประมาณ 1 ใน 25 ส่วนของเส้นผ่านศูนย์กลางเส้นผมมนุษย์ เล็กจนขนจมูกของมนุษย์ที่ทำหน้าที่กรองฝุ่นนั้นไม่สามารถกรองได้ จึงแพร่กระจายเข้าสู่ทางเดินหายใจ กระแสเลือด และเข้าสู่อวัยอื่นๆ ในร่างกายได้ ตัวฝุ่นเป็นพาหะนำสารอื่นเข้ามาด้วย เช่น แคดเมียม ปรอท โลหะหนัก และสารก่อมะเร็งอื่นๆ

สาเหตุที่ทำให้เกิด PM2.5 (Particulate Matter 2.5)

ฝุ่นละอองขนาดไม่เกิน 2.5 ไมครอน (PM2.5) มาจากสองแหล่งกำเนิดใหญ่ๆ คือ แหล่งกำเนิดโดยตรง ได้แก่ การเผาในที่โล่ง การคมนาคมขนส่ง การผลิตไฟฟ้า อุตสาหกรรมการผลิต การรวมตัวของก๊าซอื่นๆ ในบรรยากาศ โดยเฉพาะซัลเฟอร์ไดออกไซด์ (SO2) และออกไซด์ของไนโตรเจน (NOx) รวมทั้งสารพิษอื่นๆ ที่ล้วนเป็นอันตรายต่อร่างกายมนุษย์ เช่น สารปรอท (Hg), แคดเมียม (Cd), อาร์เซนิก (As) หรือโพลีไซคลิกอะโรมาติกไฮโดรคาร์บอน (PAHs)

เกณฑ์ของดัชนีคุณภาพอากาศ

เกณฑ์ตรวจเช็คดัชนีคุณภาพอากาศ (Air Quality Index : AQI) โดยประเทศไทยแบ่งดัชนีคุณภาพอากาศเป็น 5 ระดับ ตั้งแต่ 0 ถึง 201 ขึ้นไป โดยใช้สีเป็นตัวเปรียบเทียบระดับของผลกระทบต่อสุขภาพ

ตารางที่ 2.1 เกณฑ์ของดัชนีคุณภาพอากาศ

AQI	คุณภาพอากาศ	สีที่ใช้	ข้อความแจ้งเตือน
0 - 25	ดีมาก	ฟ้า	เหมาะสำหรับกิจกรรมกลางแจ้งและการท่องเที่ยว
26 - 50	ดี	เขียว	สามารถทำกิจกรรมกลางแจ้งและการท่องเที่ยวได้ตามปกติ

Computer Science Department
Faculty of Informatics, Mahasarakham University

ตารางที่ 2.1 เกณฑ์ของดัชนีคุณภาพอากาศ(ต่อ)

AQI	คุณภาพอากาศ	สีที่ใช้	ข้อความแจ้งเตือน
51 - 100	ปานกลาง	เหลือง	สามารถทำกิจกรรมกลางแจ้งได้ตามปกติ แต่ถ้าเป็นผู้ที่ต้องดูแลสุขภาพเป็นพิเศษ หากมีอาการเกี่ยวกับระบบทางเดินหายใจและระคายเคืองตา ไม่ควรทำกิจกรรมกลางแจ้งนาน
101 - 200	เริ่มมีผลกระทบต่อสุขภาพ	ส้ม	ควรเฝ้าระวังสุขภาพ ถ้ามีอาการเกี่ยวกับระบบทางเดินหายใจ ระคายเคืองตา ไม่ควรทำกิจกรรมกลางแจ้งนาน หรือใช้อุปกรณ์ป้องกัน ส่วนผู้ที่ต้องดูแลสุขภาพเป็นพิเศษ แล้วมีอาการทางสุขภาพ เช่น ไอ หายใจลำบาก ตาอักเสบ แน่นหน้าอก ปวดศีรษะ หัวใจเต้นไม่เป็นปกติ คลื่นไส้อ่อนเพลีย ควรปรึกษาแพทย์
201 ขึ้นไป	มีผลกระทบต่อสุขภาพ	แดง	ทุกคนควรหลีกเลี่ยงกิจกรรมกลางแจ้งทุกอย่างหลีกเลี่ยงพื้นที่ที่มีมลพิษทางอากาศสูง หรือใช้อุปกรณ์ป้องกันตนเองหากมีความจำเป็น หากมีอาการทางสุขภาพควรปรึกษาแพทย์

2.1.2 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม[11] คือ ระบบการคำนวณที่สร้างเลียนแบบการทำงานของระบบสมองมนุษย์ซึ่งมีปรับเปลี่ยนตัวเองต่อการตอบสนองของอินพุตตามกฎของการเรียนรู้ (learning rule) เพื่อใช้ประโยชน์ในการคาดคะเนเหตุการณ์จากข้อมูลที่มีอยู่

การประมวลผลต่าง ๆ เกิดขึ้นในหน่วยประมวลผลย่อย เรียกว่า โหนด (node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ ระหว่างโหนดที่เชื่อมต่อกัน จำลองมาจากการเชื่อมต่อของใยประสาท และแกนประสาทในระบบประสาทของสมองมนุษย์ ภายในโหนดจะมีฟังก์ชันกำหนดสัญญาณส่งออกที่เรียกว่า ฟังก์ชันการแปลง (transfer function) ซึ่งทำหน้าที่เปรียบเสมือนกระบวนการทำงานในเซลล์ ซึ่งโครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบ ดังนี้

Computer Science Department Faculty of Informatics, Mahasarakham University

1. ข้อมูลอินพุต (input) เป็นข้อมูลที่เป็นตัวเลข หากเป็นข้อมูลเชิงคุณภาพ ต้องแปลงให้อยู่ในรูปเชิงปริมาณที่โครงข่ายประสาทเทียมยอมรับได้

2. ข้อมูลเอาต์พุต (output) คือ ผลลัพธ์ที่เกิดขึ้นจริง จากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม

3. ค่าน้ำหนัก (weights) คือ สิ่งที่ได้จากการเรียนรู้ของโครงข่ายประสาทเทียม หรือเรียกอีกอย่างหนึ่งว่า ค่าความรู้ (knowledge) ค่านี้จะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูลอื่น ๆ ที่อยู่ในรูปแบบเดียวกัน

4. ฟังก์ชันผลรวม (Summation function: S) เป็นผลรวมของข้อมูลป้อนเข้า (p_i) และค่าน้ำหนัก (w_i)

$$S = \sum_{i=1}^{n} p_i w_i \qquad (1)$$

5. ฟังก์ชันการแปลง (transfer function) ฟังก์ชันการแปลง เป็นส่วนที่ทำหน้าที่รวมค่าเชิงตัวเลขจากเอาต์พุตของนิวรอล แล้วทำการตัดสินใจว่าจะส่งสัญญาณเอาต์พุตออกไปในรูปใด ฟังก์ชันการแปลงสามารถเป็นได้ทั้งแบบเชิงเส้นหรือไม่เป็นเชิงเส้น การเลือกใช้ฟังก์ชันการแปลงจะขึ้นอยู่กับลักษณะของระบบ ที่นำเอาโครงข่ายประสาทเทียมไปประยุกต์ใช้

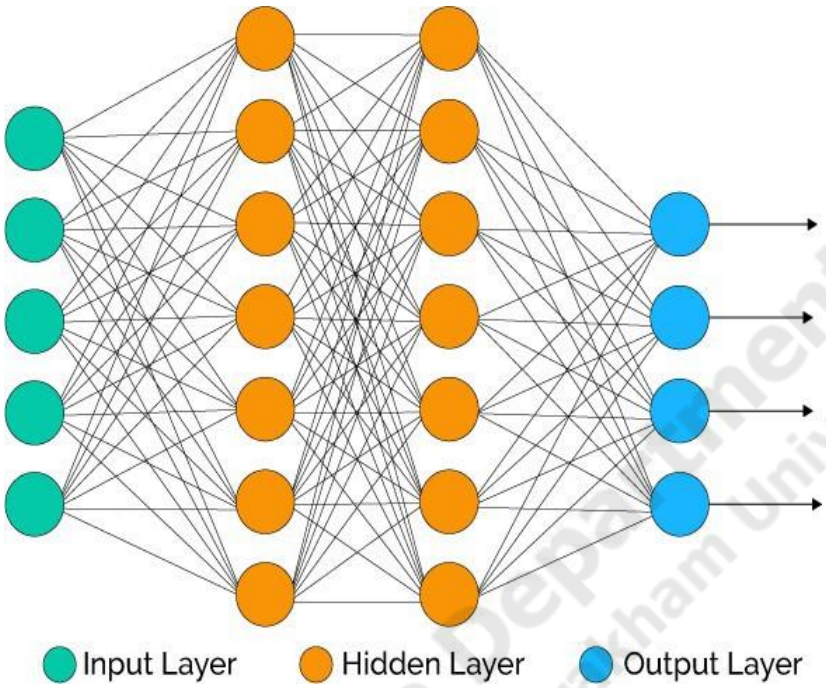
ส่วนประกอบของโครงข่ายประสาทเทียม มี 3 ส่วน ได้แก่

1.Input Layer ที่ชั้นนี้จะเป็นข้อมูล input ของเรานั่นเอง จำนวนของโหนดขึ้นอยู่กับจำนวนของ input ว่ามีข้อมูลอะไรบ้างที่สามารถนำเข้ามาประมวลผลในโมเดล เช่น ถ้าข้อมูลของลูกค้าเป็น input ซึ่งประกอบด้วย อายุ เพศ จังหวัดที่อาศัย รวมทั้งสิ้น 4 อย่าง input layer ก็จะมี 4 โหนด (ปกติแล้วใน Machine Learning เราจะเรียกปัจจัยที่นำมาวิเคราะห์เหล่านี้ว่า feature)

2.Hidden Layer เป็นชั้นที่อยู่ระหว่างกลาง ซึ่งจะมีผลอย่างมากต่อประสิทธิภาพในการเรียนรู้ของโมเดล ซึ่ง hidden layer นั้นจะมีกี่ชั้นก็ได้ แล้วแต่ผู้สร้าง โดยแต่ละชั้นจะมีจำนวนของ Neuron เท่าไหร่ก็ได้เช่นกัน ซึ่งการเพิ่มชั้นและจำนวน neuron ก็จะส่งผลต่อการทำงานของโมเดล ในส่วนของ hidden layer มีการทำงานเปรียบเสมือนส่วนที่เรียนรู้ข้อมูลเชิงลึก หรือ deep learning

3.Output Layer ชั้นที่เราจะนำเอาข้อมูลจากการคำนวณไปใช้ จำนวนของโหนดในชั้นนี้ ขึ้นอยู่กับรูปแบบของ output ที่จะเอาไปใช้ ซึ่งสามารถยกตัวอย่างง่ายๆ เช่น ถ้างานที่ทำเป็น Regression ก็กำหนดให้ output layer เป็นแบบ 1 โหนด เพราะต้องการคำตอบค่าเดียว ถ้าเป็นหลายค่าก็เพิ่มไปตามที่ต้องการ เช่น ในบางงานอาจจะ predict หาตำแหน่งของภาพในแกน x และ y พร้อมๆ กัน ในกรณีนี้สามารถใส่ output layer เป็น 2 โหนด เป็นต้น

Computer Science Department Faculty of Informatics Mahasarakham University



ภาพประกอบที่ 2.1 ส่วนประกอบของโครงข่ายประสาทเทียม

สถาปัตยกรรมของโครงข่ายประสาทเทียม

ในทางคณิตศาสตร์โครงข่ายประสาทเทียมสามารถแสดงเป็นกราฟกำกับน้ำหนัก สถาปัตยกรรมโครงข่ายประสาทเทียม ที่พบมากที่สุดคือ

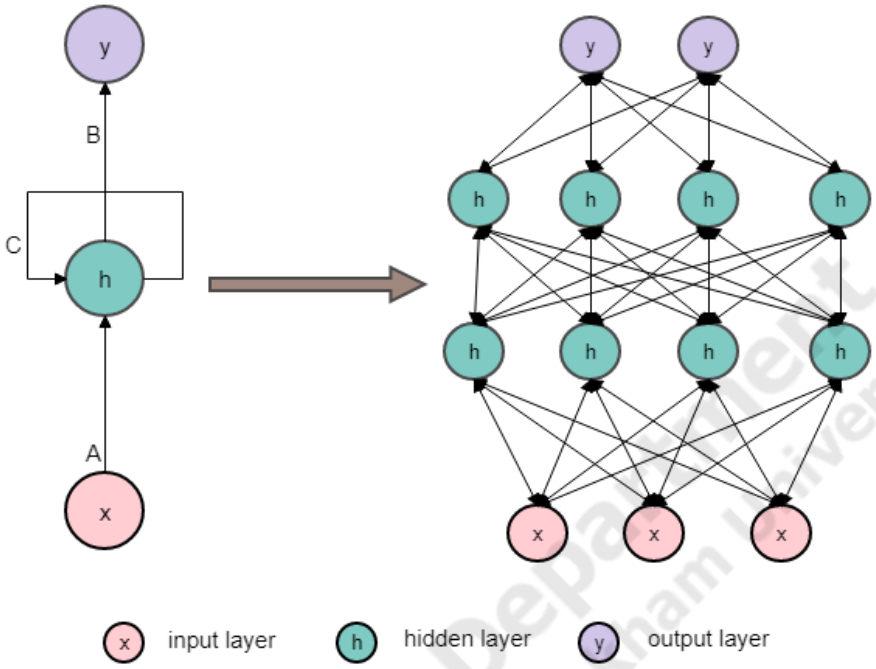
Single-layer neural network - เครือข่ายแบบชั้นเดียวเครือข่ายที่มีเซลล์ประสาทการคำนวณชั้นเดียวที่ประมวลผลสัญญาณอินพุตในทิศทางไปข้างหน้า

Multilayer neural network – เครือข่ายแบบหลายชั้นสำหรับป้อนข้อมูลเครือข่ายที่มีการเชื่อมต่อสองชั้นหรือมากกว่าพร้อมน้ำหนักที่ประมวลผลอินพุตในทิศทางไปข้างหน้า

2.1.3 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) หรือ โครงค่ายประสาทเทียมแบบวกกลับ คือโมเดลที่ถูกสร้างมาเพื่อแก้ปัญหาที่มีการเกิดขึ้นซ้ำๆ เป็นโครงค่ายประสาทเทียมแบบหลายชั้น ซึ่ง RNN จะมีการนำผลลัพธ์ก่อนหน้าวกกลับมาใช้เป็น input ในการคำนวณผลลัพธ์ใหม่ในเวลาถัดไปเสมอ

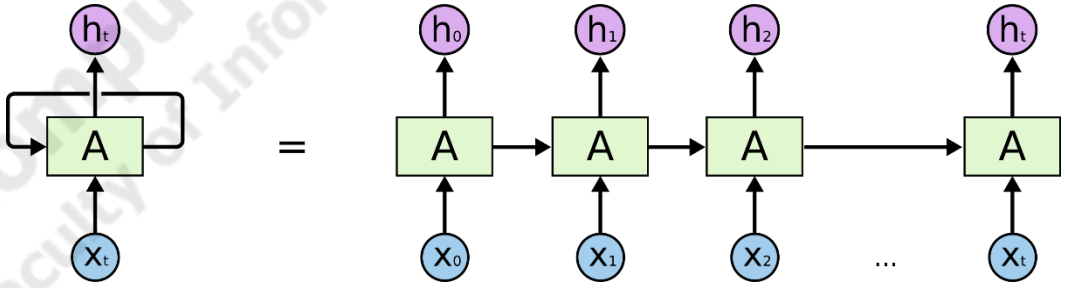
Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 2.2 Recurrent Neural Network

2.1.4 Long Short-Term Memory (LSTM)

Long Short-Term Memory[10] หรือ LSTM เป็นโครงข่ายประสาทเทียมแบบหนึ่งที่ถูกออกแบบมาสำหรับการประมวลผลลำดับ โดย LSTM นั้นจัดว่าเป็นโครงข่ายประเภท RNN นั่นคือการนำใช้ Neural Network (NN) ที่มีการนำ output ของตัวมันเองก่อนหน้านี้กลับมาใช้อีกครั้ง และการใช้ LSTM จะเป็นการแก้ปัญหาของ RNN ที่มีต่อการประมวลผลลำดับที่ยาวของข้อมูล



ภาพประกอบที่ 2.3 Long Short-Term Memory

2.1.5 ฟังก์ชันกระตุ้น

ฟังก์ชันกระตุ้น เป็นตัวกำหนดความสัมพันธ์ระหว่างอินพุตและเอาต์พุตของโหนดและเครือข่าย โดยทั่วไปฟังก์ชันกระตุ้นจะแนะนำระดับความไม่เชิงเส้นซึ่งมีค่าสำหรับแอปพลิเคชัน โครงข่ายประสาทเทียม ส่วนใหญ่ ระบุเงื่อนไขทั่วไปสำหรับฟังก์ชันต่อเนื่องที่จะมีคุณสมบัติเป็นฟังก์ชันกระตุ้น

Computer Science Department
Faculty of Informatics, Mahasarakham University

ฟังก์ชัน differentiable ใด ๆ สามารถมีคุณสมบัติเป็นฟังก์ชันการเปิดใช้งานในทางทฤษฎี ในทางปฏิบัติมีการเปิดใช้งานฟังก์ชันการเปิดใช้งาน“ แบบมีพฤติกรรมดี” เพียงเล็กน้อย (จำกัด ขอบเขตเพิ่มขึ้นแบบจำเจและแตกต่างกัน) เหล่านี้รวมถึง

1.The sigmoid (logistic) function :

$$f(x) = (1 + \exp(-x))^{-1} ; \quad (2)$$

2.The hyperbolic tangent (tanh) function:

$$f(x) = (\exp(x) - \exp(-x))/(\exp(x) + \exp(-x)) ; \quad (3)$$

3.The sine or cosine function:

$$f(x) = \boldsymbol{sin\, x} \text{ or } f(x) = \boldsymbol{cos\, x} \quad (4)$$

4.The linear function:

$$f(x) = x \quad (5)$$

2.1.6 รูปแบบของกระบวนการเรียนรู้

ทฤษฎีของโครงข่ายประสาทเทียมกระบวนการเรียนรู้ (ดร.วราวุธ ม.ป.ป.)สำหรับการปรับน้ำหนักช่วยให้ผลลัพธ์ใกล้เคียงกับเป้าหมายมากขึ้น กระบวนการเรียนรู้ประกอบด้วย 2 ลักษณะ ได้แก่

Supervised

สันนิษฐานว่าจะมีข้อเสนอแนะอยู่ในระหว่างกระบวนการเรียนรู้และตัวอย่างแต่ละรูปแบบที่ใช้ในการฝึกอบรมเครือข่ายรวมถึงรูปแบบการป้อนข้อมูลพร้อมกับเป้าหมายหรือรูปแบบผลลัพธ์ที่ต้องการคำตอบที่ถูกต้อง

Unsupervised

เครือข่ายไม่มีข้อเสนอแนะเกี่ยวกับผลลัพธ์ที่ต้องการหรือถูกต้อง ไม่มีคำแนะนำที่จะนำเสนอรูปแบบเป้าหมาย ดังนั้นระบบจะต้องเรียนรู้โดยการค้นพบและปรับให้เข้ากับคุณสมบัติโครงสร้างในรูปแบบการป้อนข้อมูลนั่นคือโดยการปรับให้เข้ากับระเบียบทางสถิติหรือการจัดกลุ่มของรูปแบบจากตัวอย่างการฝึกการป้อนข้อมูล

2.1.7 การแบ่งชุดข้อมูลเป็นชุดสอนและชุดเรียนรู้

โดยทั่วไปแล้วการตัวอย่างการชุดสอนและชุดเรียนรู้ จำเป็นสำหรับการสร้างเครื่องพยากรณ์โครงข่ายประสาทเทียม ตัวอย่าง ชุดสอน ใช้สำหรับการพัฒนาแบบจำลองโครงข่ายประสาทเทียม และนำตัวอย่าง ชุดเรียนรู้ มาใช้เพื่อประเมินความสามารถในการพยากรณ์ของแบบจำลอง บางครั้งหนึ่งในสามที่เรียกว่าตัวอย่างการตรวจสอบก็ถูกใช้เพื่อหลีกเลี่ยงปัญหาการ overfitting หรือ

Computer Science Department
Faculty of Informatics, Mahasarakham University

เพื่อกำหนดจุดหยุดของกระบวนชุดสอนเป็นเรื่องปกติที่จะใช้ ชุดเรียนรู้หนึ่งชุดสำหรับการตรวจสอบความถูกต้องและการทดสอบโดยเฉพาะกับชุดข้อมูลขนาดเล็ก ในมุมมองของเราการเลือกตัวอย่างการชุดสอนและชุดเรียนรู้ อาจส่งผลต่อประสิทธิภาพของโครงข่ายประสาทเทียม

2.1.8 การทำข้อมูลให้เป็นมาตรฐาน (Data Normalization)

ฟังก์ชันการเปิดใช้งานแบบไม่เชิงเส้น เช่น ฟังก์ชันลอจิสติกมักจะมีบทบาทการแบนในการจำกัด หรือบีบเอาท์พุทที่เป็นไปได้จากโหนดไปยังโดยทั่วไปแล้ว การทำข้อมูลให้เป็นมาตรฐานนั้นมักจะทำก่อนที่กระบวนการฝึกอบรมจะเริ่มขึ้น ดังที่ได้กล่าวไว้ก่อนหน้านี้เมื่อใช้ฟังก์ชันถ่ายโอนแบบไม่เชิงเส้นที่โหนดเอาต์พุต ค่าเอาต์พุตที่ต้องการต้องถูกเปลี่ยนเป็นช่วงของเอาต์พุตจริงของเครือข่าย แม้ว่าจะมีการใช้ฟังก์ชั่นถ่ายโอนข้อมูลเชิงเส้น แต่ก็อาจเป็นประโยชน์ในการสร้างมาตรฐานเอาต์พุตเช่นเดียวกับอินพุตเพื่อหลีกเลี่ยงปัญหาการคำนวณเพื่อตอบสนองความต้องการอัลกอริทึมและเพื่ออำนวยความสะดวกในการเรียนรู้เครือข่าย

1.ตามการปรับสภาพช่องสัญญาณ: แชนเนลถูกกำหนดให้เป็นชุดขององค์ประกอบในตำแหน่งเดียวกันกับเวกเตอร์อินพุตทั้งหมดในชุดการฝึกอบรมหรือการทดสอบ นั่นคือแต่ละช่องสามารถคิดว่าเป็นตัวแปรอินพุต“ อิสระ” การปรับสภาพช่องสัญญาณตามแนวนั้นจะถูกดำเนินการทีละคอลัมน์ถ้าใส่เวกเตอร์อินพุตลงในเมทริกซ์ มันจะทำให้ตัวแปรอินพุตแต่ละตัวเป็นปกติ

2.ข้ามการปรับสภาพของช่องสัญญาณ: การทำให้เป็นมาตรฐานชนิดนี้จะดำเนินการสำหรับเวกเตอร์อินพุตแต่ละรายการอย่างอิสระนั่นคือการทำให้เป็นมาตรฐานอยู่ในองค์ประกอบทั้งหมดในรูปแบบข้อมูล

3.การทำให้เป็นมาตรฐานของแชนเนลผสม: ตามชื่อที่แนะนำวิธีนี้จะใช้การผสมระหว่างและตามมาตรฐาน

4.การทำให้เป็นมาตรฐานภายนอก: ข้อมูลการฝึกอบรมทั้งหมดจะถูกทำให้เป็นมาตรฐานในช่วงที่ระบุ

ตัวเลือกของวิธีการข้างต้นมักจะขึ้นอยู่กับองค์ประกอบของเวกเตอร์อินพุต สำหรับปัญหาการพยากรณ์อนุกรมเวลาการทำให้เป็นมาตรฐานภายนอกมักเป็นขั้นตอนการทำให้ปกติที่เหมาะสมเท่านั้น การสังเกตที่ล้าหลังจากแหล่งเดียวกันนั้นถูกใช้เป็นตัวแปรอินพุตและสามารถรักษาโครงสร้างระหว่างแชนเนลเหมือนในซีรีย์ดั้งเดิม อย่างไรก็ตามสำหรับปัญหาการคาดการณ์เชิงสาเหตุควรใช้วิธีการทำให้เป็นมาตรฐานตามช่องสัญญาณเนื่องจากตัวแปรอินพุตเป็นตัวแปรอิสระที่ใช้ในการทำนายตัวแปรตาม ใช้วิธีการปรับสภาพข้ามช่องสัญญาณสำหรับข้อมูลอนุกรมเวลาซึ่งอาจสร้างปัญหาร้ายแรงซึ่งข้อมูลเดียวกันในรูปแบบการฝึกอบรมที่แตกต่างกันจะถูกทำให้เป็นมาตรฐานแตกต่างกัน ดังนั้นข้อมูลที่มีค่าในโครงสร้างพื้นฐานของอนุกรมเวลาดั้งเดิมอาจสูญหาย

Computer Science Department
Faculty of Informatics Mahasarakham University

สำหรับวิธีการปรับสภาพมาตรฐานแต่ละประเภทที่กล่าวถึงข้างต้นมักใช้สูตรต่อไปนี้:

- linear transformation to [0,1]: $x_n=(x_0-x_{min})/(x_{max}-x_{min})$ (6)
- linear transformation to [a,b]: $x_n=(b-a)(x_0-x_{min})/(x_{max}-x_{min})+a$ (7)
- statistical normalization: $x_n=(x_0-\overline{x})/s$ (8)
- simple normalization: $x_n=x_0/x_{max}$ (9)

โดยที่ x_n และ x_0 แสดงถึงข้อมูลที่เป็นมาตรฐานและเป็นต้นฉบับ x_{min} , x_{max} , $\overline{x}$ และ s คือค่าต่ำสุด, สูงสุด, ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานตามคอลัมน์หรือแถวตามลำดับ

2.1.9 การนำเสนอข้อมูลในรูปแบบแผนภาพ (Data Visualization)

การนำเสนอข้อมูลในรูปแบบแผนภาพ หรือ Data Visualization เป็นสิ่งหนึ่งถูกนำมาใช้แสดงแทนคำพูด เป็นการใช้ภาพเพื่อแสดงข้อมูลในเชิงปริมาณที่วัดได้ ไม่ว่าจะเป็นตัวเลข แผนภูมิ กราฟ และอื่นๆอีกมากมาย คำว่า Data คือ ข้อมูล ส่วน Visualization คือ การมองเห็น เมื่อนำมารวมกันแล้วหมายถึง ข้อมูลที่มองเห็นได้ด้วยตานั่นเอง ตัวอย่างเช่นประโยชน์ของ Data Visualization ก็คือทำให้ข้อมูลในเชิงปริมาณดูน่าสนใจ เข้าใจง่าย เห็นภาพรวมได้ชัดเจน ง่ายต่อการจดจำ และนิยมนำมาใช้ประกอบในการรายงาน การวิเคราะห์ สรุปผล อย่างแพร่หลาย

2.1.10 อนุกรมเวลา (Time Series)

อนุกรมเวลา[12] คือ เซตของข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง ตัวอย่างเช่น ดัชนีตลาดหลักทรัพย์ในแต่ละวันเมื่อปิดทำการซื้อขายในแต่ละวัน รายไตรมาส รายรับในแต่ละปีของบริษัทแห่งหนึ่ง เป็นต้น

ข้อมูลอนุกรมเวลา (Time Series data) คือ ชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลาเป็นช่วง ๆ อย่างต่อเนื่องกัน เช่น ข้อมูลยอดขายสินค้าที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ เดือน ข้อมูลรายได้ที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ ปี เป็นต้น ข้อมูลอนุกรมเวลาอาจอยู่ในลักษณะที่เป็นข้อมูลรายปี หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์

ดังนั้นการวิเคราะห์อนุกรมเวลาจึงเข้ามามีบทบาทช่วยในการตัดสินใจ เทคนิคที่ใช้ช่วยในการควบคุมการดำเนินการในปัจจุบันและในการวางแผนความต้องการในอนาคต คือ การพยากรณ์ (forecasting) ซึ่งการพยากรณ์นั้นทำได้หลายวิธี แต่ละวิธีต่างมีเป้าหมายเดียวกัน คือ ทำนายเหตุการณ์ในอนาคต

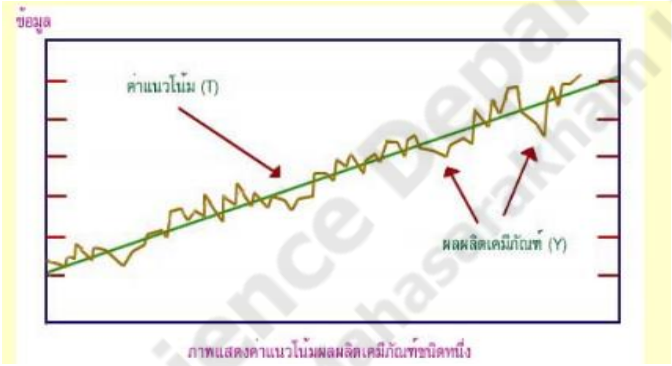
Computer Science Department Faculty of Informatics, Mahasarakham University

องค์ประกอบของอนุกรมเวลา

ในการวิเคราะห์อนุกรมเวลา ผู้วิเคราะห์จะแยกองค์ประกอบต่าง ๆ ที่ประกอบกันขึ้นเป็นอนุกรมเวลา โดยจะมีการเปลี่ยนแปลงไปตามอิทธิพลต่าง ๆ เช่น การเปลี่ยนแปลงการผลิต เทคโนโลยี สภาพอากาศ เป็นต้น การอธิบายถึงองค์ประกอบของการแปรผันของอนุกรมเวลาแบ่งเป็น 4 ส่วน ดังนี้

1. ค่าแนวโน้ม (Secular trend) แทนด้วย T_t

เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะราบ แนวโน้ม อาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่าแนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ควรเป็นข้อมูลรายปี และควรมีข้อมูลอย่างน้อย 15 ปี ซึ่งจะแสดงทิศทางของอนุกรมเวลา



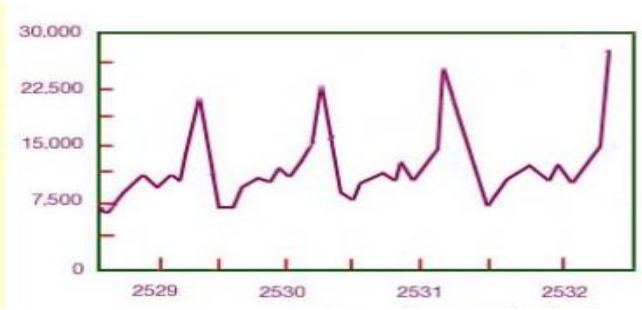
ภาพประกอบที่ 2.4 กราฟแสดงค่าแนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง

จากกราฟ Y แทนข้อมูลอนุกรมเวลาของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง ในช่วงเวลา 15 ปี ค่าแนวโน้มแทนด้วยกราฟเส้นตรง ซึ่งซ้อนอยู่บนเส้นกราฟของ Y เส้นกราฟแสดงแนวโน้มนอกจากจะมีลักษณะเป็นเส้นตรงดังรูปแล้วอาจมีลักษณะเป็นเส้นโค้ง เช่น เส้นโค้งเอ็กโพเนนเชียลหรือพาราโบลาก็ได้

2. การเปลี่ยนแปลง (Seasonal Variation) แทนด้วย S_t

เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล หน่วยของระยะเวลาสำหรับข้อมูลอาจเป็นรายชั่วโมง รายวัน รายสัปดาห์ รายเดือน สำหรับข้อมูลรายปีไม่มีการแปรผันตามฤดูกาล การเปลี่ยนแปลงตามฤดูกาลนั้นกำหนดระยะเวลาการเกิดซ้ำในรอบหนึ่ง ๆ ได้ค่อนข้างแน่นอน ตัวอย่างเช่น ยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง

Computer Science Department
Faculty of Informatics Mahasarakham University



ภาพประกอบที่ 2.5 กราฟแสดงยอดขายรายเดือนของห้างสรรพสินค้าแห้งหนึ่ง

จากกราฟ จะเห็นว่ายอดขายของห้างสูงประมาณเดือนธันวาคมของทุกปี ซึ่งเป็นเทศกาลคริสต์มาส และปีใหม่ ประชาชนจึงมีการจับจ่ายใช่สอยมาก ส่วนในราวเดือนพฤษภาคมของทุกปียอดขายจะต่ำกว่าในเดือนอื่น ๆ ที่เป็นเช่นนี้เพราะเป็นช่วงเปิดภาคเรียน ประชาชนต้องเตรียมเงินไว้สำหรับค่าใช้จ่ายในการศึกษาของบุตรหลาน

3. การเปลี่ยนแปลงตามวัฏจักร (Cyclical Variation) แทนด้วย C_t

การเปลี่ยนแปลงตามวัฏจักร มีการเปลี่ยนแปลงเคลื่อนไหวในลักษณะซ้ำ ๆ กันและจะมีลักษณะคล้ายคลึงกับการเปลี่ยนแปลงตามฤดูกาล จะต่างกันก็ตรงที่การเปลี่ยนแปลงตามวัฏจักรแต่ละรอบจะใช้ระยะเวลาที่นานกว่า คือ ตั้งแต่ 5 ปีขึ้นไป

ข้อมูลที่มีการเปลี่ยนแปลงตามวัฏจักรในทางธุรกิจ เรียกว่า "วัฏจักรธุรกิจ" โดยทั่วไปประกอบด้วย ระยะเจริญรุ่งเรือง ระยะฝืดเคือง ระยะตกต่ำ และระยะขยายตัว (recovery)

ภาพประกอบที่ 2.6 วัฏจักรธุรกิจ

จากกราฟ การเกิดระยะต่าง ๆ เหล่านี้ เกิดอย่างต่อเนื่องเป็นวัฏจักร และแต่ละรอบของวัฏจักรมีระยะเวลาไม่แน่นอน

4. การเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (Irregular Variation) แทนด้วย I_t

เป็นการเปลี่ยนแปลงของข้อมูลอนุกรมเวลาที่เกิดจากเหตุการณ์ที่เราไม่สามารถ

Computer Science Department
Faculty of Informatics, Mahasarakham University

คาดการณ์ได้ล่วงหน้า เช่น การเกิดไฟไหม้ในโรงงาน การเกิดอุทกภัย การนัดหยุดงานของคนงาน แผ่นดินไหว เป็นต้น ซึ่งเหตุการณ์เหล่านี้เป็นสิ่งที่เกิดขึ้นโดยบังเอิญไม่คาดคิดมาก่อน เป็นการเปลี่ยนแปลงที่เป็นเชิงสุ่ม (random variation) เพราะไม่ได้อยู่ภายใต้เงื่อนไขที่เรากำหนด

2.1.11 การพยากรณ์

การพยากรณ์ คือ การประมาณ หรือ การคาดคะเนว่าอะไรจะเกิดขึ้นในอนาคต เช่น การพยากรณ์ยอดขายของ 3 ปีข้างหน้า การพยากรณ์มีบทบาทสำคัญกับทุกด้าน ทั้งหน่วยงานของรัฐบาล และเอกชน รัฐบาลต้องประมาณ หรือ พยากรณ์รายได้ รายจ่ายในปีหน้า เพื่อนำมาวางแผน เอกชนต้องพยากรณ์ยอดขาย เพื่อนำมาวางแผนการผลิต สินค้าคงคลัง แรงงาน เป็นต้น

การพยากรณ์จำแนกออกได้เป็น 2 ประเภทใหญ่ๆ คือ

1. การพยากรณ์เชิงปริมาณ

เป็นการพยากรณ์ที่ต้องอาศัยสถิติข้อมูลเชิงปริมาณในอดีตมาใช้เป็นฐานการพยากรณ์ ดังนั้นควรนำมาใช้ในกรณีที่

- ข้อมูลในอดีตสามารถหาได้
- ข้อมูลมีจำนวนเพียงพอ
- ประวัติศาสตร์จะซ้ำรอย

วิธีการพยากรณ์เชิงปริมาณ ได้แก่

1.1 วิธีอนุกรมเวลา (TIME-SERIES METHODS)

- วิธีค่าเฉลี่ยเคลื่อนที่ (MOVING AVERAGE)
- วิธีเอกโปเนนเชียล (EXPONENTIAL SMOOTHING)

1.2 วิธีวิเคราะห์ความสัมพันธ์ (CASUAL METHOD)

- วิธีวิเคราะห์สมการถดถอย (REGRESSION ANALYSIS)

2. การพยากรณ์เชิงคุณภาพ

เป็นการพยากรณ์ที่อาศัยข้อมูล เชิงพรรณนา เช่น ความรู้สึกวิจารณญาณ ทัศนคติ ความคิดเห็นส่วนตัว ประสบการณ์เป็นฐานในการพยากรณ์ วิธีนี้เหมาะสำหรับกรณีที่สถิติข้อมูลเชิงปริมาณในอดีตมีไม่เพียงพอหรือไม่สามารถรวบรวมได้

วิธีการพยากรณ์เชิงคุณภาพ มีหลายวิธี เช่น

2.1 วิธีเดลฟี (DELPHI METHOD)

2.2 วิธีสอบถามผู้บริหารระดับสูง

Computer Science Department Faculty of Informatics, Mahasarakham University

2.3 วิธีสอบถามพนักงานระดับปฏิบัติ

2.4 วิธีสอบถามผู้เชี่ยวชาญ

2.5 วิธีสำรวจตลาด

2.1.12 การวิเคราะห์ทางสถิติ

การวิเคราะห์ทางสถิติเป็นวิธีที่ใช้สำหรับการตรวจสอบข้อมูลการวิเคราะห์ปัจจัยและการตรวจสอบประสิทธิภาพของ ANN ในการตรวจสอบประสิทธิภาพข้อมูลการทำนายจากแบบจำลองจะถูกนำไปเปรียบเทียบกับข้อมูลการตรวจสอบจริง การวิเคราะห์การเบี่ยงเบนของการทำนายที่ใช้กันอย่างแพร่หลายดังต่อไปนี้

1. ค่าคาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\frac{(T_i - F_i)^2}{100} \quad (10)$$

โดยที่ N = จำนวนข้อมูล

T = ข้อมูลเป้าหมาย

F = ข้อมูลพยากรณ์

2. ค่าเฉลี่ยของรากที่สองของความคาดเคลื่อน (Root Mean Square Error : RMSE)

$$RMSE = \sqrt{(\frac{1}{N}\sum_{i=1}^{N}(T_i - F_i)^2)} \quad (11)$$

โดยที่ N = จำนวนข้อมูล

T = ข้อมูลเป้าหมาย

F = ข้อมูลพยากรณ์

3. ค่าเฉลี่ยของเปอร์เซ็นต์ของความคาดเคลื่อน (Mean Absolute Percentage Error : MAPE)

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{(T_i - F_i)^2}{T_i}\right| \times 100 \quad (12)$$

โดยที่ N = จำนวนข้อมูล

T = ข้อมูลเป้าหมาย

F = ข้อมูลพยากรณ์.

Computer Science Department
Faculty of Information, Mahasarakham University

2.1.13 ฟลัทเตอร์ (Flutter Framework)

Flutter Framework คือ SDK (Software Development Kit) สำหรับพัฒนาแอปพลิเคชัน บน Mobile ซึ่งพัฒนาโดย Google ในปี 2016 Flutter นั้นสามารถสร้างทั้ง iOS และ Android ได้ด้วยการเขียนเพียงครั้งเดียว ภาษาที่ใช้เขียน Flutter คือ ภาษา Dart เป็นหลัก โดยการเขียนใน Flutter นั้นจะมี Component (ใน Flutter จะเรียกว่า Widget) ต่างๆที่สามารถเรียกใช้ได้สะดวก และเป็น Framework ที่จัดอยู่ในหมวด hybrid-native

ข้อดีของ Flutter

- สามารมั่นใจได้ว่าหน้าตาของแอปพลิเคชันที่สร้างขึ้นจะเหมือนกันไม่ว่าจะเป็น Android หรือ iOS
- แอปพลิเคชันที่สร้างขึ้นไม่จำเป็นจะต้องเหมือนกันทั้งหมดซะทีเดียวเพราะ Flutter สามารถ design แยกแต่ละ Platform ได้
- ประสิทธิภาพดีกว่าเนื่องจากตัวภาษา Dart นั่น compile เป็น machine code โดยไม่ต้องรันบน JS Engine

ข้อเสียของ Flutter

- ใช้ภาษา Dart ซึ่งคนส่วนใหญ่ยังไม่คุ้นเคยกับ syntax มากนัก
- เนื่องด้วยการที่ไม่ได้ใช้ Component พื้นฐานของ OS ทำให้เวลามีการเปลี่ยนแปลง UI ตัวแอปพลิเคชัน จะไม่ได้รับการ update ในทันที

2.1.14 ไพธอน (Python)

Python คือ ชื่อภาษาที่ใช้ในการเขียนโปรแกรมภาษาหนึ่ง ซึ่งถูกพัฒนาขึ้นโดยไม่ยึดติดกับแพลตฟอร์มใดๆ Python เป็น Open Source ทำให้ผู้พัฒนาโปรแกรมหลายๆคนสามารถที่จะนำ Python มาพัฒนาต่อยอดโดยไม่เสียค่าใช้จ่าย เนื่องจาก python มีความสามารถที่สูงขึ้นจึงและครอบคลุมลักษณะงานที่หลากหลาย

ข้อดีของ Python

- ไวยากรณ์อ่านง่าย
- การเขียนโค้ดด้วย Python สามารถนำไปรันบนระบบปฏิบัติการได้หลากหลาย
- มีฟังชันก์สนับสนุนฐานข้อมูลอย่าง MySQL , Oracle และ Sybase เป็นต้น
- เป็น Open Source

Computer Science Department Faculty of Informatics Mahasarakham University

ข้อเสียของ Python

- การทำงานของ Python จะช้ากว่าการทำงานจากโปรแกรมที่ผ่านการแปลโปรแกรมเป็นภาษาเครื่อง

Libery ที่ใช้กับ Python มีดังนี้

1. Pandas

Pandas นั้นคือ Library หนึ่งในภาษา Python ที่ทำให้เราสามารถจัดการข้อมูลต่างๆ ได้ง่ายขึ้น เช่น การโหลดข้อมูลไฟล์ CSV เข้ามาแล้วแสดงข้อมูลให้ออกมาในรูปแบบคล้ายกับ Table โดยมีการแบ่งข้อมูลเป็น Row กับ Column

2. Sklearn

Sklearn เป็นไลบรารีฟรีในภาษา Python สำหรับการพัฒนาโปรแกรมโดยใช้การเรียนรู้ของเครื่องจุดเด่นคือฟังก์ชันในการแบ่งประเภทข้อมูลการแบ่งกลุ่มข้อมูลการวิเคราะห์การถดถอยหลายอย่างไม่ว่าจะเป็นซัพพอร์ตเวกเตอร์แมชชีนการเรียนรู้ต้นไม้ตัดสินใจและการแบ่งกลุ่มข้อมูลแบบเคมีน

3. Keras

Keras คือ deep learning framework สำหรับ python ที่ทำให้วิธีของเทรน โมเดลด้วย deep learning ง่ายขึ้น โดยมีฟังก์ชั่น build-in ในการใช้งาน convolution network สำหรับงาน computer visionและงาน recurrentnetwork สำหรับงาน sequenceprocessing หมายความว่าKerasเหมาะสำหรับการสร้าง deep learning model ในทุกๆ แบบจาก generative adversarialnetworkไปจนถึงneuralTuringmachine Keras อยู่ภายใต้ MIT license ที่หมายความว่าเราสามารถใช้งานมันได้อย่างอิสระเพื่อธุรกิจที่หวังกำไรได้ โดยไม่ต้องจ่ายเงินให้ใคร และยังใช้ได้กับทุกเวอร์ชั่นของ python จาก 2.7 ถึง 3.6 อีกด้วย

4. Flask

Flask คือ web framework ที่เขียนขึ้นมาสำหรับ Python เพื่อใช้ร่วมกัน webserver เช่น Apache และได้รับการยอมรับจาก community we pages ชั้นนำเช่น Pinterest, LinkedIn เป็นต้น โดย Flask ถูกเรียกว่า micro framework เพราะว่า มันไม่ต้องการเครื่องมือ หรือ library อะไรมาก อีกทั้ง ไม่จำเป็นต้องมี database ด้วย แต่อย่างไรก็ตาม Flask ก็ยังรองรับการเพิ่ม extensions พิเศษได้ ถ้ามันรองรับ Flask

Computer Science Department Faculty of Informatics, Mahasarakham University

2.2 งานวิจัยและระบบงานที่เกี่ยวข้อง

2.2.1 Forecasting the behavior of multivariate time series using neural networks

บทความนี้[2] Kanad Chakraborty , Kishan Mehrotra , Chilukuri K. Mohan และ Sanjay Ranka ต้องการนำเสนอวิธีโครงข่ายประสาทเทียมในการวิเคราะห์อนุกรมเวลาหลายตัวแปร การสังเกตราคาแป้งในสามเมืองในโลกจริงถูกใช้เป็นเกณฑ์มาตรฐานในการทดลองของเรา เครือข่ายการเชื่อมต่อ Feedforward ได้รับการออกแบบเพื่อจำลองราคาแป้งในช่วงเวลาระหว่างเดือนสิงหาคม 2515 ถึงพฤศจิกายน 2523 สำหรับเมืองต่างๆในเมืองบัฟฟาโล มินนิอาโปลิส และแคนซัสซิตี้ ความสำเร็จที่โดดเด่นได้รับความสำเร็จในการฝึกอบรมเครือข่ายเพื่อเรียนรู้เส้นโค้งราคาสำหรับแต่ละเมืองเหล่านี้และในการทำนายราคาอย่างแม่นยำ ผลลัพธ์ของเราแสดงให้เห็นว่าวิธีโครงข่ายประสาทเทียมเป็นคู่แข่งชั้นนำที่มีวิธีการสร้างแบบจำลองทางสถิติ

2.2.2 Designing an artificial neural network for forecasting tourism time series

บทความนี้[1] Alfonso Palmer , Juan José Montaño และ Albert Sesé มีจุดประสงค์เพื่อให้ข้อมูลเบื้องต้นเกี่ยวกับหลักการทางทฤษฎีของโครงข่ายประสาทเทียม และ step-by-step methodology สำหรับการออกแบบเครือข่ายประสาทสำหรับการคาดการณ์อนุกรมเวลาการท่องเที่ยว อนุกรมเวลาที่สอดคล้องกับค่าใช้จ่ายด้านการท่องเที่ยวในหมู่เกาะแบลีแอริก (สเปน) ซึ่งเป็นหนึ่งในจุดหมายปลายทางการท่องเที่ยวที่สำคัญของโลกถูกใช้เป็นข้อมูลเพื่อแสดงให้เห็นถึงกระบวนการนี้ กฎการปฏิบัติและการอภิปรายจำนวนมากระหว่างผู้เขียนซึ่งครอบคลุมทั้งนักวิจัยเชิงวิชาการและผู้ปฏิบัติได้ถูกนำมารวมไว้ในบทความเพื่อให้ ANN สามารถนำไปใช้ได้อย่างประสบความสำเร็จ สุดท้ายผลที่ได้จากการศึกษาครั้งนี้ให้ข้อมูลสำหรับนักวิจัยที่สนใจใช้โครงข่ายประสาทเทียม กับการพยากรณ์ข้อมูลการท่องเที่ยว

2.2.3 Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile

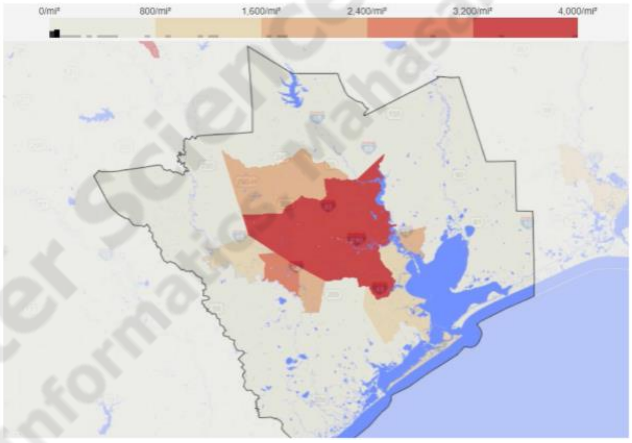
บทความนี้[6] Patricio Pérez , Alex Trier , Jorge Reyes ได้ทำการเปรียบเทียบการทำนายที่สร้างขึ้นโดยวิธีการที่ต่างกันสามวิธี ได้แก่ โครงข่ายประสาทเทียมหลายชั้น , Linear Regression และ Persistence โดยมีการวัดค่าความเข้มข้นของ PM 2.5 เฉลี่ยรายชั่วโมงที่จุดคงที่ในย่านใจกลางเมืองซานติอาโก ประเทศชิลี และมุ่งเน้นความสนใจไปที่ข้อมูลสำหรับเดือนที่มีค่าความเข้มข้นที่สูงขึ้น ตั้งแต่พฤษภาคมถึงกันยายนในปี 1994 และ 1995 ซึ่งแสดงให้เห็นว่าเป็นไปได้ที่จะทำนายความเข้มข้นในทุก ๆ ชั่วโมงของวันโดยปรับฟังก์ชันเป็น 24 ชั่วโมง เพื่อวัดค่าความเข้มข้นเฉลี่ยที่

Computer Science Department
Faculty of Informatics, Mahasarakham University

วัดได้ในวันก่อนหน้า โดยรวมแล้วโครงข่ายประสาทเทียมจะให้ผลลัพธ์ที่ดีที่สุด ข้อผิดพลาดในการทำนายเพิ่มขึ้นจาก 30% สำหรับชั่วโมงก่อนหน้าเป็น 60% สำหรับช่วงเวลาหลัง เพื่อปรับปรุงการทำนายผลกระทบของการลดเสียงรบกวนการจัดเรียงข้อมูลใหม่และการพิจารณาตัวแปรทางอุตุนิยมวิทยาอย่างชัดเจน

2.2.4 Temporal characteristics and forecasting of PM2.5 concentration based on historical data in Houston , USA

บทความนี้[5] Jianbang Du , Fengxiang Qiao , Lei Yu ได้นำเสนอถึงลักษณะชั่วคราวของความเข้มข้นใน PM2.5 ของแต่ละพื้นที่ทางภูมิศาสตร์ในเมืองฮูสตัน รัฐเท็กซัส ประเทศสหรัฐอเมริกาจากพื้นดินในชนบทไปยังพื้นที่ชายฝั่งและเพื่อสร้างแบบจำลองการพยากรณ์ที่สามารถทำ PM2.5 รายวันและรายชั่วโมง การคาดการณ์การวิจัยครั้งนี้วิเคราะห์ข้อมูลความเข้มข้นของ PM2.5 ในอดีตตั้งแต่ปี 2008 ถึง 2017 ที่เก็บรวบรวมจากห้าภูมิภาคในฮูสตันโดยคณะกรรมการคุณภาพสิ่งแวดล้อมเท็กซัส



ภาพประกอบที่ 2.7 ความหนาแน่นของ PM2.5 ในเมืองฮูสตัน

ความแม่นยำของการพยากรณ์ได้รับการปรับปรุงตามการรวมกันของปัจจัยต่างๆเมื่อเปรียบเทียบกับการพยากรณ์ตาม ปัจจัยความเข้มข้นรายชั่วโมง และปัจจัยความเข้มข้นรายเดือน ข้อดีของการพยากรณ์รวมดังกล่าวรวมถึงการลดความซับซ้อนในการคำนวณ และการพยากรณ์ความเข้มข้นรายชั่วโมง ตารางการค้นหาสำหรับแต่ละลักษณะของคุณลักษณะทางโลกนั้นถูกจัดเตรียมไว้ซึ่งสามารถนำไปใช้ในการพยากรณ์อื่น ๆ ได้อย่างง่ายดาย

Computer Science Department
Faculty of Informatics Mahasarakham University

2.2.5 Application of fuzzy time series models for forecasting pollution concentration

บทความนี้ D. Doman´ skaa , M. Wojtylak ได้นำเสนอแบบจำลองเพื่อทำนายความเข้มข้นของอนุภาคฝุ่น PM10, PM2.5, SO_2 , NO, CO และ O_3 สำหรับจำนวนชั่วโมงที่เลือกล่วงหน้าจะถูกนำเสนอ โดยวิธีนี้ต้องการข้อมูลจำนวนมากโดยเฉพาะข้อมูลการพยากรณ์อากาศข้อมูลสภาพอากาศจริงและข้อมูลมลพิษ โดยการจับคู่ข้อมูลการพยากรณ์กับข้อมูลการคาดการณ์ที่คล้ายกันในชุดข้อมูลในอดีตเป็นไปได้ที่จะรับข้อมูลสภาพอากาศจริงและผ่านข้อมูลมลพิษนี้ เพื่อรวมคะแนนเวลาด้วยการคาดการณ์ที่คล้ายกันข้อมูลที่กำหนดโดยฟังก์ชั่นระยะทางตัวเลขที่คลุมเครือจะถูกสร้างขึ้นจากข้อมูลการพยากรณ์ครอบคลุมข้อมูลการคาดการณ์และข้อมูลจริง เมื่อใช้ฟังก์ชันระยะทางอีกครั้งข้อมูลจริงจะถูกเปรียบเทียบกับฟัซซี่เพื่อกำหนดระดับการเป็นสมาชิก แบบจำลองถูกจัดทำในลักษณะที่ทุกคนข้อมูลที่มักจะไม่ชัดเจนวุ่นวายไม่แน่ใจสามารถนำมาใช้ รูปแบบที่ใช้ในโปแลนด์โดยสถาบันอุตุนิยมวิทยาและการจัดการน้ำและโดย Voivodship Inspector เพื่อสิ่งแวดล้อม การคุ้มครองจิต มันคาดการณ์ความเข้มข้นมลพิษที่เลือกสำหรับทุกพื้นที่ของโปแลนด์

2.2.6 Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation

บทความนี้ Xiao Feng , Qi Li , Yajie Zhu , Junxiong Hou , Lingyan Jin , Jingjie Wang ต้องการนำเสนอการผสมผสานระหว่างการวิเคราะห์วิถีมวลอากาศและการแปลงเวฟเล็ตเพื่อปรับปรุงโครงข่ายประสาทเทียม ที่แม่นยำในการคาดการณ์ความเข้มข้นเฉลี่ยรายวันของ PM2.5 ล่วงหน้าสองวัน แบบจำลองดังกล่าวได้รับการพัฒนาจากสถานีตรวจวัดมลพิษทางอากาศ 13 แห่งในปักกิ่ง โดยมีเส้นทางมวลอากาศ ระยะทางระหว่างพื้นที่ใกล้เคียง ความเร็วลมและทิศทางถูกพิจารณาเป็นพารามิเตอร์ในการคำนวณค่าตัวบ่งชี้มลพิษทางอากาศ อนุกรมเวลาเดิมของความเข้มข้น PM2.5 จะถูกเปลี่ยนโดยการแปลงเวฟเล็ตเป็นซีรีย์ย่อยไม่กี่ชุดที่มีความแปรปรวนต่ำ แบบจำลองทางภูมิศาสตร์ที่ใช้วิถีและการแปลงเวฟเล็ตสามารถเป็นเครื่องมือที่มีประสิทธิภาพในการปรับปรุงความแม่นยำในการพยากรณ์ PM2.5 ค่าเฉลี่ยข้อผิดพลาดกำลังสอง (RMSE) ของโมเดลไฮบริดสามารถลดลงได้โดยเฉลี่ยมากถึง 40 เปอร์เซ็นต์ โดยเฉพาะอย่างยิ่ง PM2.5 วันที่สูงคาดว่าจะใช้การสลายตัวของเวฟเล็ตและอัตราการตรวจจับ (DR) สำหรับเกณฑ์การแจ้งเตือนที่กำหนดของรุ่นไฮบริดสามารถเข้าถึง 90% โดยเฉลี่ย วิธีนี้แสดงให้เห็นถึงศักยภาพ

Computer Science Department Faculty of Informatics Mahasarakham University

2.2.7 แอปพลิเคชัน AirCare

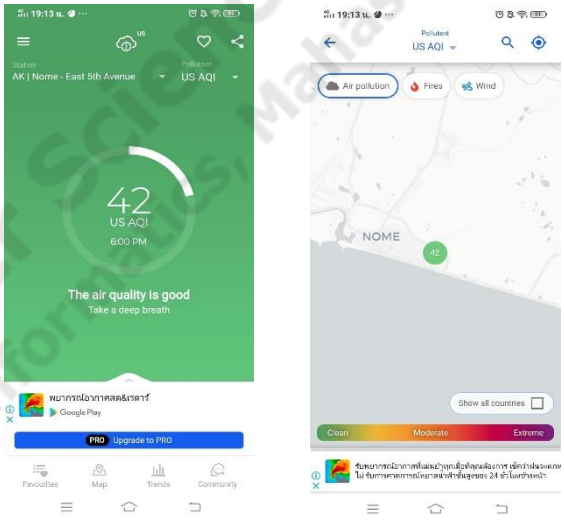
แอปพลิเคชัน AirCare Gorjan Jovanovski เป็นผู้นำเสนอแอปพลิเคชันคอยเก็บข้อมูลมลพิษทางอากาศในอเมริกาเหนือ ยุโรป และออสเตรเลีย มลพิษที่แอปพลิเคชันได้ทำการเก็บข้อมูลได้แก่ PM10 , PM2.5 , CO และ NO2 เป็นต้น นอกจากจะเก็บข้อมูลเหล่านี้แล้ว แอปพลิเคชันยังได้ทำข้อมูลให้ออกมาดูได้ง่ายเหมาะกับผู้ใช้งานที่พึ่งเริ่มใช้

ข้อดี

- สามารถเลือกได้ว่าจะดูค่ามลพิษทางอากาศแบบใดได้บาง
- มีการทำให้ Data Visualization ทำให้ข้อมูลดูได้ง่าย
- เลือกดูแบบ Map เพื่อดูได้ว่าอากาศบริเวณไหนมีมลพิษมากน้อย

ข้อเสีย

- พื้นที่บริการยังไม่ครอบคลุม ใช้งานได้เฉพาะบางจุดเท่านั้น
- ไม่สามารถดูข้อมูลย้อนหลังได้
- สามารถดูได้ทีละข้อมูล



ภาพประกอบที่ 2.8 ตัวอย่างของแอปพลิเคชั่น AirCare

2.2.8 แอปพลิเคชัน Breezometer

แอปพลิเคชัน Breezometer Breezometer ได้นำเสนอแอปพลิเคชันคอยเก็บข้อมูลและพยากรณ์มลพิษทางอากาศซึ่งแอปพลิเคชันนี้มีบริการถึง 94 ประเทศ โดยแอปพลิเคชันจะเก็บข้อมูลมลพิษทางอากาศโดยรวมเหมือนกับข้อก่อนหน้า แต่แอปพลิเคชันนี้จะดูที่มีได้พร้อมกัน มีการเก็บข้อมูล

Computer Science Department
Faculty of Informatics, Mahasarakham University

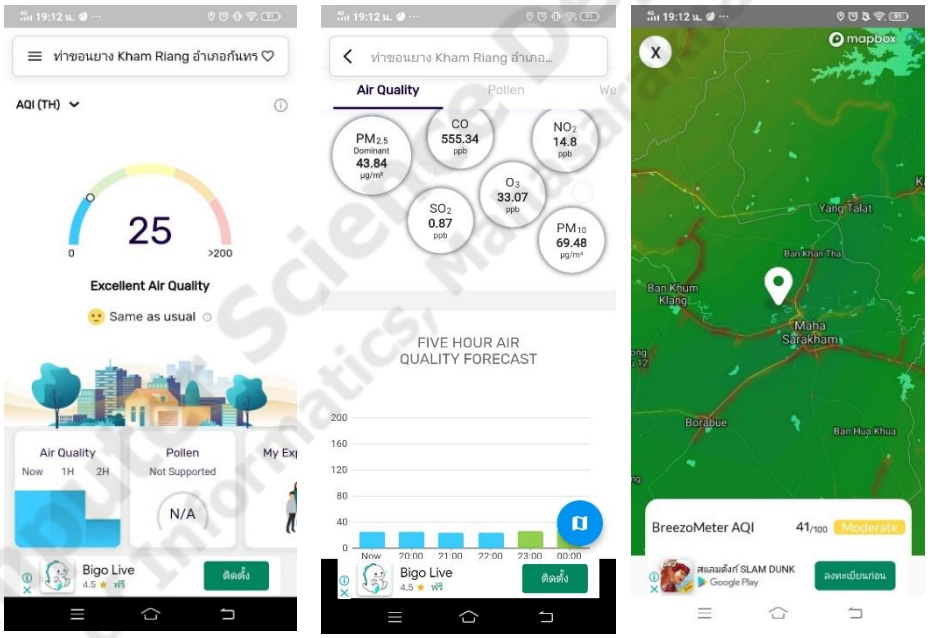
ย้อนหลัง และการพยากรณ์สามารถได้ถึง 5 ช่วงเวลา ตัวแอปพลิเคชันยังได้ทำข้อมูลให้ออกมาดูได้ง่าย เหมาะกับผู้ใช้งานที่พึ่งเริ่มใช้

ข้อดี

- สามารถเลือกได้ว่าจะดูค่ามลพิษทางอากาศแบบใดได้บาง หรือดูพร้อมกันทั้งก็ได้
- มีการทำให้ Data Visualization ทำให้ข้อมูลดูได้ง่าย
- เลือกดูแบบ Map เพื่อดูได้ว่าอากาศบริเวณไหนมีมลพิษมากน้อย
- พื้นที่บริการครอบคลุมมากกว่าแอปพลิเคชันแรก

ข้อเสีย

- การใช้งานดูซับซ้อนขึ้น



ภาพประกอบที่ 2.9 ตัวอย่างของแอปพลิเคชั่น AirCare

2.2.9 แอปพลิเคชัน AirVisual

แอปพลิเคชัน AirVisual ได้นำเสนอแอปพลิเคชันคอยเก็บข้อมูลและพยากรณ์มลพิษทางอากาศซึ่งครอบคลุมมากที่สุดจาก 2 แอปพลิเคชันที่ผ่านมา แต่จะเก็บมลพิษทางอากาศน้อยกว่า 2 แอปพลิเคชันแรก แอปพลิเคชันนี้มีการเก็บข้อมูลย้อนหลังและมีการพยากรณ์ล่วงถึง 7 วัน ตัวแอปพลิเคชันยังได้ทำข้อมูลให้ออกมาดูได้ง่ายเหมาะกับผู้ใช้งานที่พึ่งเริ่มใช้

ข้อดี

- สามารถเลือกได้ว่าจะดูค่ามลพิษทางอากาศแบบใดได้บาง หรือดูพร้อมกันทั้งก็ได้
- มีการทำให้ Data Visualization ทำให้ข้อมูลดูได้ง่าย
- เลือกดูแบบ Map เพื่อดูได้ว่าอากาศบริเวณไหนมีมลพิษมากน้อย
- พื้นที่บริการครอบคลุมทุกพื้นที่

ข้อเสีย

- ไม่มี

ภาพประกอบที่ 2.10 ตัวอย่างแอปพลิเคชัน AirVisual

Computer Science Department
Faculty of Informatics, Mahasarakham University