

ระบบทำนายข้อมูลล่วงหน้าโดยใช้วิธี

Artificial Neural Network

แบบจำลองโครงข่ายประสาทเทียม

1. ทฤษฎีแบบจำลองโครงข่ายประสาทเทียม

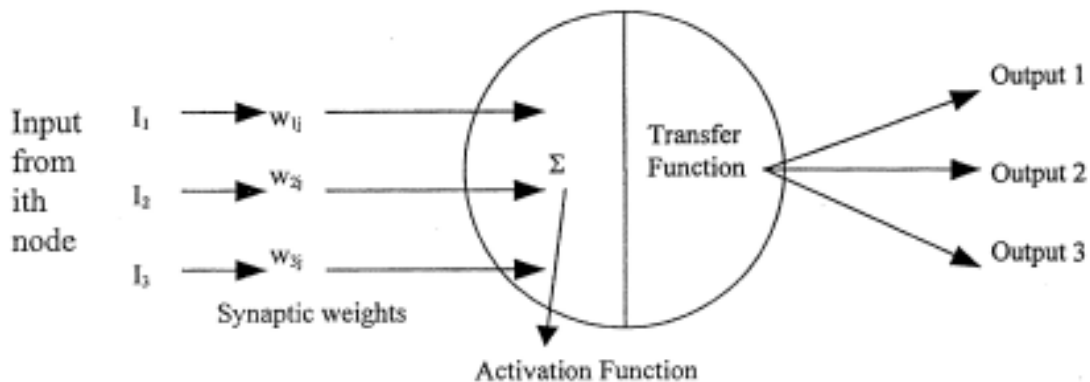
ระบบโครงข่ายประสาทเทียม (Artificial Neural Networks, ANNs) เป็นส่วนประกอบอย่างหนึ่ง ในระบบปัญญาประดิษฐ์ (Artificial Intelligence) ซึ่งมีโครงสร้างทางสถาปัตยกรรม และการทำงาน คล้ายคลึงกับเซลล์สมองและระบบประสาทของมนุษย์ โดยได้นำข้อดีของเซลล์สมองมาใช้คือ ความสามารถในการจำแนกลักษณะสิ่งของที่มีความใกล้เคียงกัน ความสามารถในการเรียนรู้จากประสบการณ์ ความสามารถในการแปลความหมายของสัญลักษณ์และภาพ ซึ่งผลลัพธ์ที่ออกมาขึ้นอยู่กับข้อมูลที่นำเข้า เชื่อถือได้ ถึงแม้ว่าข้อมูลที่ป้อนเข้าไปนั้นจะมีความผิดพลาดอยู่บ้าง หรือมีความไม่สมบูรณ์ของข้อมูลที่ป้อนเข้าไป เมื่อนำระบบโครงข่ายประสาทเทียมมาทำงานร่วมกับคอมพิวเตอร์ที่มีความสามารถในการคำนวณค่าทางคณิตศาสตร์ได้อย่างรวดเร็วและถูกต้อง สามารถทำให้ระบบโครงข่ายประสาทเทียมมีศักยภาพในการทำงานสูงขึ้น โดยระบบโครงข่ายประสาทเทียมใหม่ที่ได้มีคุณสมบัติพิเศษเฉพาะได้แก่ ความสามารถในการจำลองปัญหาโดยไม่จำเป็นต้องทราบรูปแบบการกระจายตัวของข้อมูล โดยผลลัพธ์ที่ออกมามีความผิดพลาดอยู่ในเกณฑ์ที่จำกัดและสามารถยอมรับได้ จากคุณสมบัติเหล่านี้ทำให้ระบบโครงข่ายประสาทเทียมมีความสามารถในการนำมาใช้ในการแก้ปัญหาได้ใกล้เคียงกับการคำนวณของเซลล์สมองและระบบประสาทของมนุษย์ แต่ในระบบโครงข่ายประสาทเทียมมีความสามารถในการคำนวณสูงกว่า โดยมีรูปแบบการทำงานผ่านฟังก์ชันทางคณิตศาสตร์ที่ไม่ซับซ้อน โดยการทำงานของระบบไม่ได้ทำงานซ้ำๆ ตามชุดคำสั่งเหมือนดังโปรแกรมคอมพิวเตอร์ทั่วไป และผลลัพธ์ที่ออกมามีค่าอยู่ในเกณฑ์ที่น่าพอใจ ถึงแม้ว่าจะมีความผิดพลาดอยู่บ้างก็ตาม แต่ระบบโครงข่ายประสาทเทียมยังมีความสามารถที่จะทำงานได้แม้ว่าข้อมูลที่ป้อนจะมีความบกพร่อง ผิดพลาด หรือขาดความสมบูรณ์ในตัวข้อมูลนั้นๆ ซึ่งเมื่อนำไปเปรียบเทียบกับแบบจำลองทางคอมพิวเตอร์ที่มีข้อจำกัดอยู่หลายอย่าง เช่นรูปแบบการทำงาน หรือการแก้ปัญหาที่แน่นอน ถ้าข้อมูลที่ป้อนเข้าในแบบจำลองมีความผิดพลาด ผลลัพธ์ที่ได้จากแบบจำลองทางคอมพิวเตอร์อื่นๆก็จะมี ความผิดพลาด หรือไม่ใกล้เคียงกับความเป็นจริง ซึ่งจะแตกต่างกับแบบจำลองระบบโครงข่ายประสาทเทียมดังที่ได้กล่าวมาแล้วในข้างต้น

ระบบโครงข่ายประสาทเทียมที่ใช้ในครั้งนี้นำประกอบด้วย Back Propagation, Time Delay Neural Networks (TDNN) และ Continuous Adaptive Time Neural Networks (CATNN) วิธีการที่เป็นพื้นฐานและนิยมใช้กันมากเนื่องจากง่ายที่จะเข้าใจและสะดวกต่อการใช้งาน คือ Back Propagation จึงขอกล่าวรายละเอียดในที่นี้ สำหรับวิธีการอื่นๆจะอธิบายโดยสังเขปต่อไป

2. โครงข่ายประสาทเทียมชนิด Back Propagation

โครงข่ายประสาทเทียมชนิด Back Propagation โดยทั่วไปแล้วมีโครงสร้างแสดงดังภาพที่ 1 โดยจะมีการกำหนดข้อมูลด้านเข้าในชั้นรับข้อมูลในชั้นแรก และจะถูกคูณด้วยค่าถ่วงน้ำหนัก เมื่อเข้าสู่ระบบ Back Propagation ค่าต่างๆที่ได้มาจะถูกรวมเข้าด้วยกัน หลังจากนั้นโครงข่ายประสาทเทียมจะมีการคำนวณโดยผ่าน Activation Function ในอันดับสุดท้ายจะมีการส่งผ่านผลที่ได้ออกทางชั้นแสดงผล

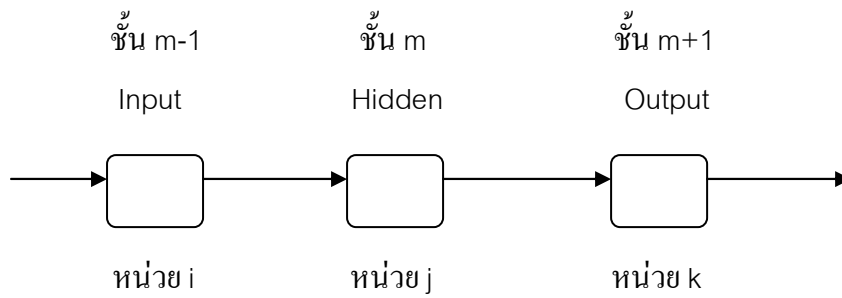
วิธีการ Back Propagation (BP) ใช้ค่าข้อมูลด้านเข้าและข้อมูลด้านออก โดยชุดข้อมูลด้านเข้าจะถูกนำเข้าสู่ระบบโครงข่ายประสาทเทียมเพื่อให้ได้มาซึ่งข้อมูลด้านออก หลักจากนั้นจะนำผลลัพธ์ที่ได้จากการคำนวณโดยระบบโครงข่ายประสาทเทียมไปเปรียบเทียบกับผลลัพธ์จริง ซึ่งถ้าผลลัพธ์ที่ออกมาไม่มีความแตกต่างระหว่างผลลัพธ์ที่ได้จากการคำนวณกับผลลัพธ์ที่ได้จากการวัดจริงระบบโครงข่ายประสาทเทียมจะไม่มีการเรียนรู้ แต่ในทางกลับกันถ้ามีความแตกต่างระหว่างผลลัพธ์ค่าถ่วงน้ำหนักจะถูกเปลี่ยนแปลงย้อนหลังเพื่อลดค่าความแตกต่าง และเกิดการเรียนรู้ใหม่



ภาพที่ 1 ลักษณะโครงสร้างของ Back Propagation

ที่มา: รัชชัย, 2543

รูปแบบของหน่วยและชั้นที่นำมาใช้หาความสัมพันธ์ของตัวแปรนำเข้าและตัวแปรด้านออกของทฤษฎีการเรียนรู้ชนิด Back Propagation เมื่อพิจารณาหน่วยเพียงหน่วยเดียวของชั้นใดๆ ซึ่งมีการเชื่อมโยงของหน่วยระหว่างชั้น ดังภาพที่ 2 จะสามารถที่จะแสดงเป็นตัวแปร และสมการต่างๆตามลำดับชั้นตอนได้ดังนี้



ภาพที่ 2 แสดงการเชื่อมโยงของหน่วยใดๆระหว่างชั้น
ที่มา: Tsoukalas L.H. และ R.E. Uhrig, 1997

กำหนดให้

$w_{ji,m}(n)$ = ค่าน้ำหนักการเชื่อมต่อจากหน่วย i th ในชั้น $m-1$ ไปยังหน่วย j th ในชั้น m ในรอบการคำนวณที่ n th

$o_{j,m}$ = ค่าของข้อมูลด้านออกที่หน่วย j th ในชั้น m ($m=1,2,\dots,l$)

$I_i = o_{i,0}$ = ค่าของข้อมูลนำเข้าที่หน่วย i th ในชั้นแรก

T_j = ค่าของข้อมูลด้านออกที่ต้องการ (Desired or Target Output)

n_m = จำนวนหน่วยทั้งหมดของชั้น m th ($m=1,2,\dots,l$)

$o_{i,m-1} = I_{j,m}$ = ค่าของข้อมูลนำเข้าที่หน่วย j th ในชั้น m ($m=1,2,\dots,l$)

$\theta_{j,m}$ = ค่าเบี่ยงเบนที่หน่วย j th ในชั้น m

$N_{j,m}$ = ผลรวมของผลคูณระหว่างค่าน้ำหนักกับข้อมูลนำเข้ารวมกับค่าเบี่ยงเบน

Rumelhart และคณะ (1986) ได้เสนอให้ใช้กลุ่มของฟังก์ชันไม่เป็นเชิงเส้นที่เรียกว่า “กึ่งเชิงเส้น” (Semi Linear) ซึ่งทำให้ผลลัพธ์มีค่าไม่ลดลง และสามารถหาอนุพันธ์ไปยังผลรวมได้ดังสมการที่ (3)

$$o_{j,m} = f(N_{j,m}) \quad (3)$$

จากข้อกำหนด $N_{j,m}$ = ผลรวมของผลคูณระหว่างค่าน้ำหนักกับข้อมูลนำเข้ารวมกับค่าเบี่ยงเบน จะสามารถแสดงได้ดังสมการที่ (4)

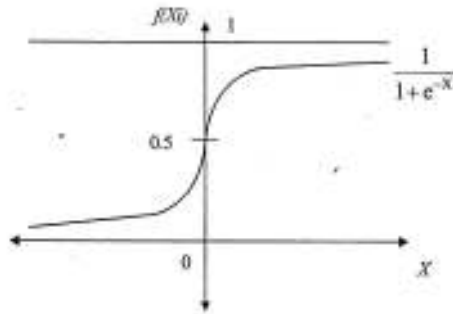
$$N_{j,m} = \sum_{i=1}^{n_{m-1}} w_{ji,m} \cdot o_{i,m-1} + \theta_{j,m} \quad (4)$$

ฟังก์ชันแปลง Logistic Sigmoidal Transform Function แสดงดังภาพที่ 3 ที่ใช้มีรูปแบบของสมการดังสมการที่ (5)

$$o_{j,m} = f(N_{j,m}) = \frac{1}{1 + e^{-N_{j,m}}} \quad (5)$$

จากคุณสมบัติของ Logistic Sigmoidal Transform Function เป็น Semi-Linear Function เป็นฟังก์ชันที่มีขอบเขตที่แน่นอน ที่มีช่วงอยู่ระหว่าง 0-1 ซึ่งสามารถหาค่าอนุพันธ์ได้ และค่าอนุพันธ์ของฟังก์ชันที่ได้จะอยู่ในรูปของผลลัพธ์เพียงอย่างเดียวดังจะแสดงในสมการที่ (6)

$$\begin{aligned} f(N_{j,m}) &= \frac{1}{(1 + e^{-N_{j,m}})} \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= \frac{\partial(1 + e^{-N_{j,m}})^{-1}}{\partial(N_{j,m})} \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= -1(1 + e^{-N_{j,m}})^{-2} \cdot \frac{\partial(1 + e^{-N_{j,m}})}{\partial(N_{j,m})} \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= -1(1 + e^{-N_{j,m}})^{-2} \cdot e^{-N_{j,m}} \cdot \frac{\partial(-N_{j,m})}{\partial(N_{j,m})} \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= \left[\frac{1}{(1 + e^{-N_{j,m}})} \right] \cdot \left[\frac{e^{-N_{j,m}}}{(1 + e^{-N_{j,m}})} \right] \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= f(N_{j,m}) \cdot \left[\frac{e^{-N_{j,m}} + 1 - 1}{(1 + e^{-N_{j,m}})} \right] \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= f(N_{j,m}) \cdot \left[1 - \frac{1}{(1 + e^{-N_{j,m}})} \right] \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= f(N_{j,m}) \cdot [1 - f(N_{j,m})] \\ \frac{\partial f(N_{j,m})}{\partial(N_{j,m})} &= o_{j,m} \cdot (1 - o_{j,m}) \quad (6) \end{aligned}$$



ภาพที่ 3 Logistic Sigmoidal Function

ที่มา: Tsoukalas L.H. และ R.E. Uhrig, 1997

ชุดของข้อมูลที่คำนวณสุดท้ายแล้ว เมื่อนำไปเปรียบเทียบกับข้อมูลที่ได้จากการวัดจริง จะได้ผลลัพธ์ข้อค่าความผิดพลาดที่เกิดขึ้นจากการคำนวณซึ่งอยู่ในรูปของผลรวมของผลต่างกำลังสอง (The Mean Squared Difference Between the Output and the Target E_p) ซึ่งจะแสดงดังรูปสมการที่ (7)

$$E_p = \frac{1}{2} \sum_{j=1}^{n_l} (T_j - o_{j,l})^2 \quad (7)$$

โดยที่ผลรวมของความผิดพลาดทั้งหมดจะแสดงดังสมการที่ (8)

$$E = \sum_p E_p \quad (8)$$

โดยทฤษฎีการเรียนรู้แบบ Back Propagation ที่นำมาหาค่าความผิดพลาดต่ำสุด (Minimize Error) เพื่อใช้ในการปรับแก้ค่าน้ำหนัก และค่าเบี่ยงเบนระหว่างการเชื่อมโยงในรูปแบบสมการอนุพันธ์อันดับที่หนึ่ง ดังสมการที่ (9)

$$\frac{\partial E_p}{\partial w_{j,m}} = \frac{\partial E_p}{\partial N_{j,m}} \cdot \frac{\partial N_{j,m}}{\partial w_{j,m}} \quad (9)$$

กำหนดให้

$$\delta_{j,m} = \frac{\partial E_p}{\partial N_{j,m}} \quad (10)$$

คำนวณค่าสมการอนุพันธ์จาก สมการที่ (4) เทียบกับ $w_{ij,m}$ จะได้

$$\frac{\partial N_{j,m}}{\partial w_{ij,m}} = o_{i,m-1} \quad (11)$$

จากการรวมสมการที่ (9) (10) และ (11) จะได้

$$\frac{\partial E_p}{\partial w_{ji,m}} = \delta_{j,m} \cdot o_{i,m-1} \quad (12)$$

$$\frac{\partial E_p}{\partial \theta_{j,m}} = \frac{\partial E_p}{\partial N_{j,m}} \cdot \frac{\partial N_{j,m}}{\partial \theta_{j,m}} \quad (13)$$

จากสมการที่ (10) จะได้ $\delta_{j,m} = \frac{\partial E_p}{\partial N_{j,m}}$

จากนั้นทำการคำนวณค่าอนุพันธ์สมการที่ (4) เทียบกับ $\theta_{j,m}$ ซึ่งแสดงในสมการที่ (14)

$$\frac{\partial N_{j,m}}{\partial \theta_{j,m}} = 1 \quad (14)$$

จากการรวมสมการที่ (11) (13) และ (14) จะได้

$$\frac{\partial E_p}{\partial \theta_{j,m}} = \delta_{j,m} \quad (15)$$

จากสมการที่ (10) จะได้ $\delta_{j,m} = \frac{\partial E_p}{\partial N_{j,m}}$

จากกฎของ Chain Rule จะได้

$$\delta_{j,m} = \frac{-\partial o_{j,m}}{\partial N_{j,m}} \cdot \frac{\partial E_p}{\partial o_{j,m}} \quad (16)$$

ค่าอนุพันธ์ของฟังก์ชันผลรวมของผลคูณของตัวแปรนำเข้ากับค่าน้ำหนักของหน่วยน้ำหนัก แต่ละหน่วยในชั้นแสดงผล หาได้จากสมการที่ (6)

$$\frac{\partial o_{j,m}}{\partial (N_{j,m})} = \frac{\partial f(N_{j,m})}{\partial (N_{j,m})} = f'(N_{j,m}) = o_{j,m}(1-o_{j,m}) \quad (17)$$

จากนั้นคำนวณค่าสมการอนุพันธ์ของสมการที่ (7) เทียบกับ $o_{j,m}$ จะได้ดังนี้

$$\frac{\partial E_p}{\partial o_{j,m}} = -(T_j - o_{j,m}) \quad (18)$$

จากการรวมสมการที่ (16) (17) และ (18) จะได้

$$\delta_{j,m} = -o_{j,m}(1-o_{j,m}) \cdot (T_j - o_{j,m}) \quad (19)$$

ค่าอนุพันธ์ของฟังก์ชันผลรวมของผลคูณของตัวแปรนำเข้ากับค่าน้ำหนักของหน่วยน้ำหนักแต่ละหน่วยในชั้นแฝง หาได้จากสมการที่ (6)

$$\frac{\partial o_{j,m}}{\partial (N_{j,m})} = \frac{\partial f(N_{j,m})}{\partial (N_{j,m})} = f'(N_{j,m}) = o_{j,m}(1-o_{j,m}) \quad (20)$$

จากกฎของ Chain Rule จะได้

$$\frac{\partial E_p}{\partial o_{j,m}} = \frac{\partial E_p}{\partial (N_{k,m+1})} \cdot \frac{\partial (N_{k,m+1})}{\partial o_{j,m}} \quad (21)$$

จากสมการที่ (4) เมื่อเปลี่ยนหน่วยและชั้นเป็น k และ m+1 ตามลำดับจะมีการเปลี่ยนแปลงเป็นดังสมการที่ (22)

$$N_{k,m+1} = \sum_{k=1}^{n_{m+1}} w_{kj,m+1} \cdot o_{j,m} + \theta_{k,m+1} \quad (22)$$

หาค่าสมการอนุพันธ์สมการที่ (22) เทียบกับ $o_{j,m}$ จะผลลัพธ์เป็น

$$\frac{\partial (N_{k,m+1})}{\partial o_{j,m}} = \sum_{k=1}^{n_{m+1}} w_{kj,m+1} \quad (23)$$

จากสมการที่ (10) เมื่อเปลี่ยนหน่วยและชั้นเป็น k และ m+1 ตามลำดับ จะผลลัพธ์เป็น

$$\delta_{k,m+1} = \frac{\partial E_p}{\partial (N_{k,m+1})} \quad (24)$$

จากการรวมสมการที่ (21) (23) และ (24) ได้ผลดังนี้

$$\frac{\partial E_p}{\partial o_{j,m}} = \sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \quad (25)$$

จากการรวมสมการที่ (16) (20) และ (25) ได้ผลดังนี้

$$\delta_{j,m} = o_{j,m} (1 - o_{j,m}) \cdot \left[\sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \right] \quad (26)$$

กำหนดให้ค่า η เป็นค่าอัตราการเรียนรู้ (Learning Rate) คำนวณน้ำหนักที่ต้องการปรับแก้ สำหรับนำไปใช้ในการคำนวณในรอบถัดไป (รอบที่ n+1) หาได้ดังนี้

$$\Delta w_{ji,m}(n+1) = -\eta \cdot \frac{\partial E_p}{\partial w_{ji}} \quad (27)$$

เมื่อแทนค่า $\frac{\partial E_p}{\partial w_{ji,m}}$ จากสมการที่ (12) จะได้ผลลัพธ์เป็น

$$\Delta w_{ji,m}(n+1) = -\eta \cdot \delta_{j,m} \cdot o_{i,m-1} \quad (28)$$

เมื่อแทนค่า $\delta_{j,m}$ จากสมการที่ (19) จะได้ค่านำหนักที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ n+1) ในชั้นแสดงผลดังนี้

$$\Delta w_{ji,m}(n+1) = \eta \cdot [o_{j,m} \cdot (1 - o_{j,m}) \cdot (T_j - o_{j,m})] \cdot o_{i,m-1} \quad (29)$$

เมื่อแทนค่า $\delta_{j,m}$ จากสมการที่ (26) จะได้ค่านำหนักที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ n+1) ในชั้นแฝงดังนี้

$$\Delta w_{ji,m}(n+1) = -\eta \left[o_{j,m} (1 - o_{j,m}) \left[\sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \right] \right] \cdot o_{i,m-1} \quad (30)$$

ค่าเบี่ยงเบนที่ต้องปรับแก้ สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ n+1) หาได้ดังนี้

$$\Delta \theta_{j,m}(n+1) = -\eta \frac{\partial E_p}{\partial \theta_{j,m}} \quad (31)$$

เมื่อแทนค่า $\frac{\partial E_p}{\partial \theta_{j,m}}$ จากสมการที่ (15) จะได้ผลลัพธ์เป็น

$$\Delta \theta_{j,m}(n+1) = -\eta \delta_{j,m} \quad (32)$$

เมื่อแทนค่า $\delta_{j,m}$ จากสมการที่ (19) จะได้ค่าเบี่ยงเบนที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ n+1) ในชั้นแสดงผลดังนี้

$$\Delta\theta_{j,m}(n+1) = \eta[o_{j,m} \cdot (1 - o_{j,m}) \cdot (T_j - o_{j,m})] \quad (33)$$

เมื่อแทนค่า $\delta_{j,m}$ จากสมการที่ (26) จะได้ค่าเบี่ยงเบนที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ $n+1$) ในชั้นแฝงดังนี้

$$\Delta\theta_{j,m}(n+1) = -\eta \left[o_{j,m} (1 - o_{j,m}) \cdot \left[\sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \right] \right] \quad (34)$$

โดยที่ค่า η คือค่าอัตราการเรียนรู้ (Learning Rate) ซึ่งเป็นค่าคงที่ในช่วง 0 ถึง 1 ค่าอัตราการเรียนรู้มีผลต่อความเร็วในการเรียนรู้ เพื่อหาค่าชุดของค่าน้ำหนัก และค่าเบี่ยงเบนที่เหมาะสม จากหลักการเรียนรู้โดยจะมีการปรับแก้ค่าน้ำหนักระหว่างการเชื่อมโยงของหน่วยต่างๆที่เกิดขึ้นในระบบโครงข่ายประสาทเทียม โดยการปรับแก้ใช้ค่าสมการอนุพันธ์ลำดับที่หนึ่ง ของค่าความผิดพลาดทั้งหมดเทียบกับค่าน้ำหนัก โดยที่ยังมีความบกพร่องในการหาค่าชุดค่าน้ำหนักและค่าเบี่ยงเบนที่เหมาะสมในการใช้งาน คือระยะเวลาที่ใช้ในการคำนวณเพื่อเข้าสู่ค่าของค่าน้ำหนัก และค่าเบี่ยงเบนที่เหมาะสม ต้องใช้เวลาในการคำนวณที่ช้ายาวนาน และอาจจะได้ค่าชุดค่าน้ำหนักและค่าเบี่ยงเบนที่ไม่ใช่ชุดที่ทำให้ค่าอนุพันธ์ระหว่างค่าความผิดพลาดเทียบกับค่าน้ำหนักแต่ละค่า และค่าถ่วงน้ำหนักแต่ละค่ามีค่าต่ำสุดที่ไม่แท้จริงเรียกว่า Local Minima of the Error Surface แต่เราต้องการค่าต่ำสุดที่แท้จริง Global Minima ดังนั้นเพื่อหลีกเลี่ยงปัญหาดังกล่าว และเป็นการปรับปรุงคุณภาพของระบบโครงข่ายประสาทเทียมให้มีประสิทธิภาพดียิ่งขึ้น จึงได้นำเทอมของ Momentum Term เพื่อนำมาใช้ในการปรับปรุงค่าน้ำหนักและค่าเบี่ยงเบนที่ใช้ในการคำนวณใหม่ในรอบต่อไป โดยการบวกเพิ่มตัวโมเมนตัมเทอม ค่าโมเมนตัมเทอมที่นำมาใช้จะช่วยให้กระบวนการเรียนรู้ของโครงข่ายระบบประสาทเทียมลดการสั่น (Oscillation) และยังเป็นลดเวลาที่ใช้ในการคำนวณเพื่อเข้าสู่ค่าชุดของค่าน้ำหนักและค่าเบี่ยงเบนที่เหมาะสม นอกจากนั้นยังทำให้การเรียนรู้ไม่ตกอยู่ในสถานะภาพของจุดต่ำสุดของค่าความผิดพลาดที่ไม่แท้จริง หรือไม่ทำให้การเรียนรู้ข้ามจุดต่ำสุดที่แท้จริงของค่าความผิดพลาด ไปสู่จุดต่ำสุดของค่าความผิดพลาดที่ไม่แท้จริง

ค่า α เป็นค่าโมเมนตัมแฟกเตอร์ (Momentum Factor) ซึ่งเป็นค่าคงที่ในช่วง 0 ถึง 0.9 โดยนำมาใช้คูณกับผลจากการเปลี่ยนแปลงค่าน้ำหนักในรอบก่อนหน้า หรือการเปลี่ยนแปลงของค่าน้ำหนักในรอบก่อนหน้า เพื่อนำมาใช้ในการพิจารณาทิศทางการปรับแก้ค่าน้ำหนัก และค่าเบี่ยงเบนในรอบปัจจุบัน

ค่าน้ำหนักที่ต้องปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ $n+1$) เมื่อนำเอาผลของค่าโมเมนต์มาพิจารณาพร้อมด้วยจะมีรูปแบบของสมการดังนี้

$$\Delta w_{ji,m}(n+1) = -\eta \cdot \frac{\partial E_p}{\partial w_{ji}} + \alpha \Delta w_{ji,m}(n) \quad (35)$$

จากสมการที่ (12) เมื่อแทนค่า $\frac{\partial E_p}{\partial w_{ji,m}}$ จะได้ผลลัพธ์เป็น

$$\Delta w_{ji,m}(n+1) = -\eta \cdot \delta_{j,m} \cdot o_{i,m-1} + \alpha \Delta w_{ji,m}(n) \quad (36)$$

จากสมการที่ (19) เมื่อแทนค่า $\delta_{j,m}$ จะได้ค่าน้ำหนักที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ $n+1$) ในชั้นแสดงผลดังนี้

$$\Delta w_{ji,m}(n+1) = \eta \cdot [o_{j,m} \cdot (1 - o_{j,m}) \cdot (T_j - o_{j,m})] \cdot o_{i,m-1} + \alpha \Delta w_{ji,m}(n) \quad (37)$$

จากสมการที่ (26) เมื่อแทนค่า $\delta_{j,m}$ จะได้ค่าน้ำหนักที่ต้องการปรับแก้สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ $n+1$) ในชั้นแฝงดังนี้

$$\Delta w_{ji,m}(n+1) = -\eta \left[o_{j,m} (1 - o_{j,m}) \left[\sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \right] \right] \cdot o_{i,m-1} + \alpha \Delta w_{ji,m}(n) \quad (38)$$

ค่าเบี่ยงเบนที่ต้องปรับแก้ สำหรับนำไปใช้คำนวณในรอบถัดไป (รอบที่ $n+1$) เมื่อนำผลของโมเมนต์มาพิจารณาด้วยจะได้ผลลัพธ์ดังนี้

$$\Delta \theta_{j,m}(n+1) = -\eta \frac{\partial E_p}{\partial \theta_{j,m}} + \alpha \Delta \theta_{j,m}(n) \quad (39)$$

จากสมการที่ (15) เมื่อแทนค่า $\frac{\partial E_p}{\partial \theta_{j,m}}$ จะได้ผลลัพธ์เป็น

$$\Delta\theta_{j,m}(n+1) = -\eta\delta_{j,m} + \alpha\Delta\theta_{j,m}(n) \quad (40)$$

จากสมการที่ (19) เมื่อแทนค่า $\delta_{j,m}$ จะได้ค่าเบี่ยงเบนที่ต้องการปรับแก้สำหรับนำไปใช้ในการคำนวณในรอบถัดไป (รอบที่ $n+1$) ในชั้นแสดงผลดังนี้

$$\Delta\theta_{j,m}(n+1) = \eta[o_{j,m} \cdot (1 - o_{j,m}) \cdot (T_j - o_{j,m})] + \alpha\Delta\theta_{j,m}(n) \quad (42)$$

จากสมการที่ (26) เมื่อแทนค่า $\delta_{j,m}$ จะได้ค่าเบี่ยงเบนที่ต้องการปรับแก้สำหรับนำไปใช้ในการคำนวณในรอบถัดไป (รอบที่ $n+1$) ในชั้นแฝงดังนี้

$$\Delta\theta_{j,m}(n+1) = -\eta \left[o_{j,m} (1 - o_{j,m}) \cdot \left[\sum_{k=1}^{n_{m+1}} \delta_{k,m+1} \cdot w_{kj,m+1} \right] \right] + \alpha\Delta\theta_{j,m}(n) \quad (43)$$

ค่าน้ำหนักที่ใช้คำนวณในรอบต่อไปของระบบโครงข่ายประสาทเทียมจะได้จากค่าน้ำหนักในรอบก่อนหน้ารวมกับค่าน้ำหนักที่ต้องการปรับแก้ดังนี้

$$w_{ji,m}(n+1) = w_{ji,m}(n) + \Delta w_{ji,m}(n+1) \quad (44)$$

เมื่อค่า $\Delta w_{ji,m}(n+1)$ เป็นไปตามสมการที่ (29) หรือ (37) และ (30) หรือ (38) สำหรับชั้นแสดงผลและชั้นแฝงตามลำดับ โดยสมการที่ (37) และ (38) เป็นกรณีที่รวมโมเมนตัมเทียม

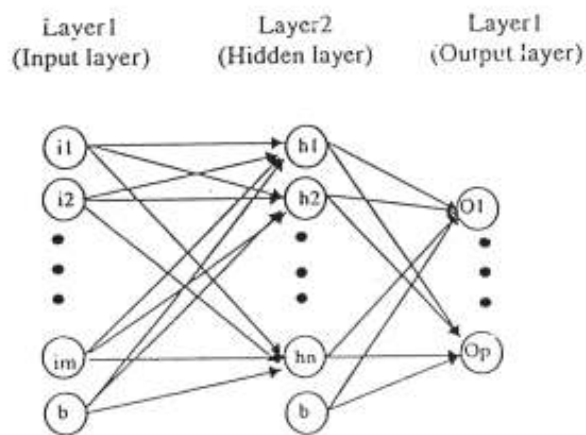
ค่าเบี่ยงเบนที่ใช้คำนวณในรอบต่อไปจะได้ค่าเบี่ยงเบนในรอบก่อนหน้ารวมกับค่าเบี่ยงเบนที่ต้องปรับแก้ดังนี้

$$\theta_{j,m}(n+1) = \theta_{j,m}(n) + \Delta\theta_{j,m}(n+1) \quad (45)$$

เมื่อค่า $\Delta\theta_{j,m}(n+1)$ มีค่าเป็นไปตามสมการที่ (33) หรือ (41) และ (34) หรือ (42) สำหรับชั้นแสดงผลและชั้นแฝงตามลำดับ โดยสมการที่ (41) และ (42) เป็นกรณีที่รวมโมเมนตัมเทียม

2.1 กระบวนการทำงานและขั้นตอนการวิเคราะห์ของระบบ Back Propagation

ในงานวิจัยนี้จะศึกษาแบบจำลองระบบโครงข่ายประสาทเทียมชนิด Multi Layer Feed Forward Network ที่ Rumelhart และคณะ (1986) ได้เสนอไว้ดังแสดงตามภาพที่ 4

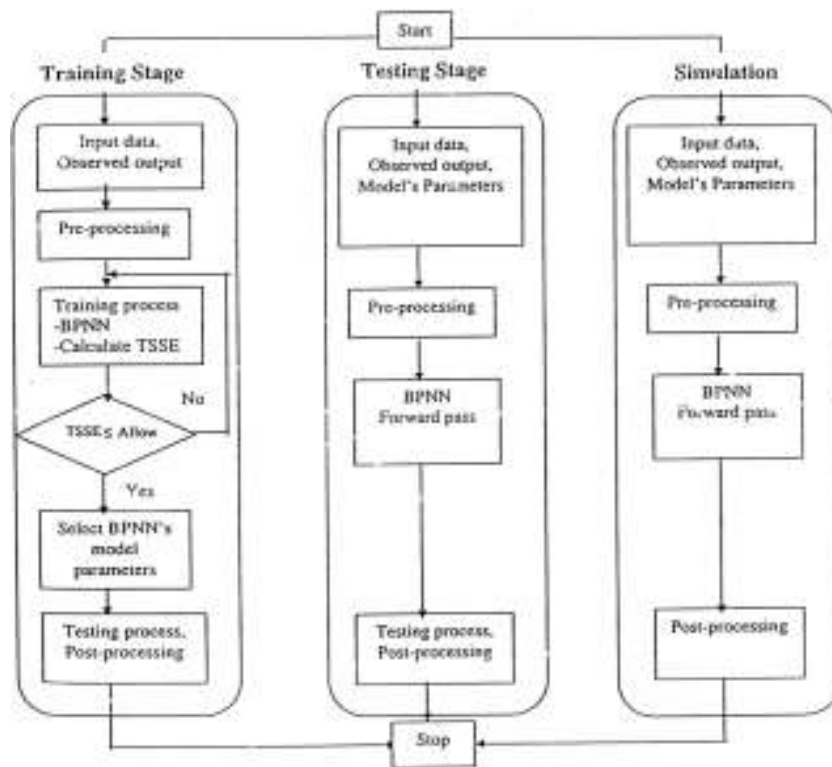


ภาพที่ 4 โครงสร้างของแบบจำลอง BPNN

ที่มา: Rumelhart และคณะ, 1986

จากภาพที่ 4 จะแสดงถึงโครงสร้างของระบบโครงข่ายประสาทเทียมที่มีด้วยกันสามชั้น คือ ชั้นรับข้อมูล (input layer, layer 1) ชั้นแฝง (hidden layer, layer 2) และชั้นแสดงผล (output layer, layer 3) โดยส่วนประกอบของชั้นรับข้อมูลจะประกอบด้วย (i_1, i_2, \dots, i_m) จะแสดงใช้เห็นว่าชั้นรับข้อมูลประกอบด้วยหน่วย (Node) ทั้งหมด m ซึ่งจะขึ้นอยู่กับลักษณะทางอุทกวิทยาที่ใช้กันทั่วไปในการวิเคราะห์หาปริมาณน้ำท่วม เช่น ข้อมูลฝน และข้อมูลของปริมาณน้ำท่าในอดีต เป็นต้น ในส่วนของชั้นแสดงผลจะเป็นผลลัพธ์ที่ต้องการซึ่งในที่นี้จะเป็นปริมาณน้ำท่วมในอนาคต สำหรับในชั้นแฝงโดยทั่วไปแล้วจะมีจำนวนหน่วยเท่ากับ จำนวนหน่วยในชั้นรับข้อมูลหรือมากกว่าก็ได้

ในการใช้งานนั้นจะมีการแยกชุดข้อมูลออกเป็นสองส่วนคือ หนึ่งส่วนของข้อมูลที่ใช้ในการทดสอบ และสองส่วนของข้อมูลที่ใช้ในการตรวจสอบโดย Back Propagation ซึ่งจะคล้ายคลึงกับแบบจำลองอุทกพลศาสตร์ ที่มีการแบ่งข้อมูลออกเป็นสองส่วนเช่นกันคือ ส่วนที่ใช้ในการสอบเทียบ กับส่วนที่ใช้ในการตรวจพิสูจน์ ดังแสดงในภาพที่ 5



ภาพที่ 5 แผนภูมิแสดงกระบวนการคำนวณใน Back Propagation
ที่มา: Manusthiparom C., 2000

2.1.1 ขั้นตอนการเรียนรู้ของแบบจำลองระบบโครงข่ายประสาทเทียม

2.1.1.1 กำหนดโครงสร้างของระบบโครงข่ายประสาทเทียม และการกำหนดค่าเงื่อนไขเริ่มต้นของค่าถ่วงน้ำหนัก และค่าเบี่ยงเบนให้กับโครงสร้างของระบบโครงข่ายประสาทเทียม รวมไปถึงการกำหนดค่าพารามิเตอร์ต่างๆด้วย ได้แก่ Learning Rate (η) Momentum (α) และ Activation Function

2.1.1.2 เริ่มกระบวนการเรียนรู้ของระบบโครงข่ายประสาทเทียมเพื่อการคำนวณหาค่าถ่วงน้ำหนักและค่าเบี่ยงเบน ที่ทำให้ค่าผลลัพธ์ที่ได้จากการคำนวณใกล้เคียงกับผลลัพธ์ที่ได้จากการวัดจริง โดยที่ขั้นตอนการทำงานของ Back Propagation จะสามารถแบ่งออกได้เป็นสองส่วนคือ ส่วนขั้นตอนไปข้างหน้า (Forward Pass) และส่วนขั้นตอนย้อนหลัง (Backward Pass) ซึ่งสามารถแสดงขั้นตอนและรายละเอียดดังต่อไปนี้ และแสดงขั้นตอนเป็นแผนภูมิการทำงานในภาพที่ 6

ส่วนของ Forward Pass

ก. กำหนดค่าของข้อมูลด้านเข้า $(I_1, I_2, \dots, I_{n_0})$ และกำหนดค่าผลลัพธ์ที่ต้องการ $(t_1, t_2, \dots, t_{n_1})$

ข. สำหรับในแต่ละชั้นที่ $m = 1, 2, \dots, L$

คำนวณค่า

$$N_{j,m} = \sum_{i=1}^{n_{m-1}} W_{ji,m} \cdot O_{i,m-1} + \theta_{j,m}$$

โดยที่กำหนดให้

$$O_{i,0} = I_i$$

คำนวณค่าข้อมูลด้านออกในหน่วยที่ j^{th} ในชั้น m ได้ดังนี้

$$O_{j,m} = \frac{1}{1 + e^{-N_{j,m}}} \quad j = 1, 2, \dots, n_m$$

ค. เปรียบเทียบผลลัพธ์ที่ได้จากการคำนวณด้วยระบบโครงข่ายประสาทเทียม $(O_{1,1}, O_{2,1}, \dots, O_{n_1,1})$ กับผลลัพธ์ที่ได้จากการวัดจริง ถ้าค่าความแตกต่างระหว่างสองผลลัพธ์อยู่ในขอบเขตที่ยอมรับได้ จะสิ้นสุดโดยถือว่าระบบเกิดการเรียนรู้ที่แล้ว แต่ในทางกลับกันถ้าผลต่างระหว่างสองผลลัพธ์ยังมีค่ามากไม่สามารถที่จะยอมรับได้ระบบโครงข่ายประสาทเทียมจะมีการกำหนดค่าเริ่มต้นต่างๆใหม่ หรือไม่ก็เปลี่ยนแปลงโครงสร้างของแบบระบบโครงข่ายประสาทเทียม

ส่วนของ Backward Pass

ง. สำหรับชั้น $m = L, L-1, L-2, \dots, 1$

สำหรับ $J = 1, 2, \dots, n_m$ จำนวนค่าดังนี้

$$\delta_{j,m} = O_{j,m} \cdot (1 - O_{j,m}) \cdot (t_j - O_{j,m}) \quad m : \text{output layer}$$

$$\delta_{j,m} = O_{j,m} \cdot (1 - O_{j,m}) \sum_{k=1}^{n_{m+1}} W_{kj,m+1} \cdot \delta_{k,m+1} \quad m : \text{hidden layer}$$

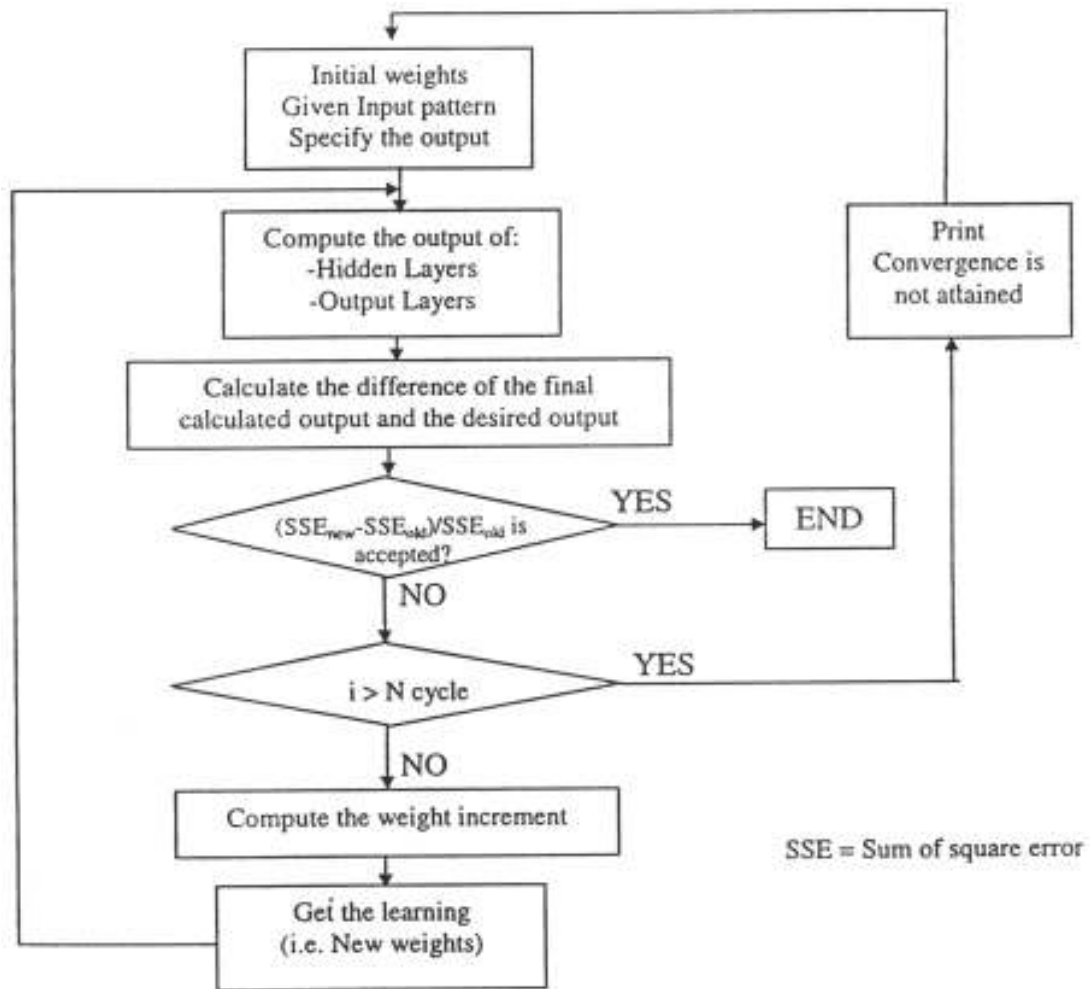
จำนวนการเปลี่ยนแปลงค่าถ่วงน้ำหนักใหม่ดังนี้

$$\Delta W_{ji,m}(n+1) = \eta \cdot \delta_{j,m} \cdot O_{i,m-1} + \alpha \cdot \Delta W_{ji,m}(n)$$

การคำนวณค่าน้ำหนักใหม่ดังนี้

$$W_{ji,m}(n+1) = W_{ji,m}(n) + \Delta W_{ji,m}(n+1)$$

จ. ย้อนกลับไปทำในชั้นที่ ค.



ภาพที่ 6 แผนภูมิแสดงขั้นตอนของวิธีการ Back Propagation

ที่มา : Manusthiparom C., 2000

2.1.2 ขั้นตอนการทดสอบของแบบจำลองระบบโครงข่ายประสาทเทียม

เมื่อแบบจำลองคำนวณค่าพารามิเตอร์ต่างๆ ออกมาแล้ว เช่นค่าถ่วงน้ำหนัก ค่าเบี่ยงเบน ค่าเหล่านี้จะไม่มีการเปลี่ยนแปลงอีกต่อไป โดยค่าพารามิเตอร์นี้จะนำไปใช้ในการทดสอบความถูกต้องของแบบจำลองระบบโครงข่ายประสาทเทียม โดยการคำนวณหาค่าสัมประสิทธิ์ของ Efficiency Index (EI) จาก

$$EI = \frac{\sum_{i=1}^N (e_i - \bar{e})^2 - \sum_{i=1}^N (e_i - F_i)^2}{\sum_{i=1}^N (e_i - \bar{e})^2} \quad (46)$$

เมื่อ

$$Q_i = \text{ค่าปริมาณจริงที่เวลา } i$$

$$\bar{Q} = \text{ค่าเฉลี่ยของค่าปริมาณน้ำจริง}$$

$$F_i = \text{ค่าที่พยากรณ์ได้ที่เวลา } i$$

$$N = \text{จำนวนของข้อมูล}$$

2.1.3 กระบวนการเตรียมชุดข้อมูลของแบบจำลองระบบโครงข่ายประสาทเทียม

ข้อมูลที่นำมาใช้ในการประมวลผลจะถูกแบ่งออกเป็น 2 ส่วน ได้แก่ ส่วนข้อมูลสำหรับการเรียนรู้และส่วนข้อมูลสำหรับการทดสอบ ทั้งนี้ช่วงของข้อมูลในส่วนของการเรียนรู้จะยาวกว่าส่วนของการทดสอบ โดยทั่วไปจะใช้ข้อมูลประมาณ 2/3 ในกระบวนการเรียนรู้ของแบบจำลอง

เนื่องจากข้อมูลมีอยู่หลายรูปแบบ ซึ่งมีค่าแตกต่างกัน ดังนั้นทั้งข้อมูลและผลลัพธ์จะต้องนำมาผ่านกระบวนการ Normalization ก่อนนำไปใช้ในแบบจำลองได้ สำหรับแบบจำลอง BPNN ที่เสนอจะใช้ฟังก์ชันลอกการิทึมในการปรับข้อมูลให้อยู่ในช่วง $[0, 1]$ แต่ทั้งนี้ในการดำเนินการจริงจะใช้ช่วง $[0.05, 0.95]$ แทนช่วง $[0, 1]$ เนื่องจากค่าลอกการิทึมจะเพียงเข้าใกล้ค่า 0 และ 1 การปรับข้อมูลก่อนและหลังการประมวลผลจะสามารถแสดงได้ดังต่อไปนี้

สมมติให้ค่า a และ A เป็นค่าต่ำสุดและสูงสุดของอนุกรมข้อมูลที่จะนำมาใช้ ดังนั้นค่าข้อมูลจริง x_t จะถูกปรับเป็น x'_t ในช่วง $[0.05, 0.95]$ โดยใช้สมการที่ (47)

$$x'_t = \frac{0.9(x_t - a)}{A - a} + 0.05 \quad (47)$$

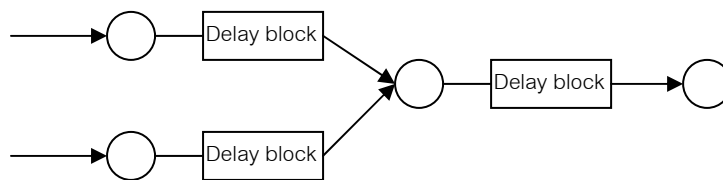
เมื่อได้ผลลัพธ์จากแบบจำลองแล้ว ผลลัพธ์ที่ได้จะถูกปรับให้เป็นค่าข้อมูลจริงเพื่อนำไปใช้งานโดยสมการที่ (48)

$$x_t = \frac{(A - a)(x'_t - 0.05)}{0.9} + a \quad (48)$$

เมื่อ x_t = ค่าข้อมูลจริง
 a = ค่าต่ำสุดของข้อมูล
 A = ค่าสูงสุดของข้อมูล
 x'_t = ค่าที่ถูกรับ

3. โครงข่ายประสาทเทียมชนิด Time Delay Neural Networks (TDNN)

เป็นโครงข่ายที่มีหลักการงานคล้ายกับโครงข่าย Backpropagation แต่แตกต่างกันที่แต่ละหน่วยในชั้นรับข้อมูล และชั้นแอบแฝง จะถูกเชื่อมโยงโดยค่าที่หน่วงเวลา (Time delay) ต่างๆกันใน delay block ดังแสดงในภาพที่ 7 และใน delay block ก็ประกอบไปด้วยค่าถ่วงน้ำหนัก (weights) ของการเชื่อมโยงต่างๆกัน ข้อดีของกระบวนการที่เพิ่มขึ้นมานี้คือช่วยลดขั้นตอนในการทำ lag input data ในไฟล์ข้อมูลก่อนการคำนวณเหมือนกับโครงข่าย Backpropagation ทั้งนี้การหาค่าถ่วงน้ำหนักก็มีวิธีการคล้ายกับโครงข่าย Backpropagation



ภาพที่ 7 โครงข่ายแบบ Time Delay Neural Network

ที่มา : เสรี, 2544

4. โครงข่ายประสาทเทียมชนิด Continuous Adaptive Time Neural Networks (CATNN)

เป็นโครงข่ายที่พัฒนาขึ้นมาจากโครงข่าย TDNN ทั้งนี้ แต่แตกต่างกันตรงที่โครงข่าย TDNN จะใช้ค่าหน่วงเวลา (Time delay) ที่สามารถปรับค่าได้ นั่นคือค่าหน่วงเวลา (Time delay) จะปรับเปลี่ยนค่าเพื่อให้ผลลัพธ์ที่ได้จากการคำนวณด้วยระบบโครงข่ายประสาทเทียมกับผลลัพธ์ที่ได้จากการวัดจริงมีความแตกต่างอยู่ในเกณฑ์ที่ยอมรับได้

การหาค่าเหมาะสมด้วยวิธีทางพันธุศาสตร์ (Genetic Algorithm)

Genetic Algorithm (GA) เป็นเทคนิคการสุ่มค้น (random search algorithm) ปัญหาที่สลับซับซ้อนซึ่งในปัญหาการหาค่าสูงสุดหรือต่ำสุดของฟังก์ชันนั้น Search Space อาจจะมีค่าที่ไม่ต่อเนื่อง (Discontinuities) และบ่อยครั้งที่มีค่าสูงสุดหรือต่ำสุดหลายๆค่า (Sub-Optimum Peaks) ปัญหาดังกล่าวอาจก่อให้เกิดความยุ่งยากมากขึ้นหากใช้วิธีแคลคูลัส หรือใช้การระบุค่าโดยตรงแต่ GA จะค้นหาค่าสูงสุดหรือต่ำสุดของฟังก์ชัน โดยใช้กลุ่มของค่าเริ่มต้นกลุ่มหนึ่งที่สุ่มมาได้แล้วป็นหาค่าสูงสุดหรือต่ำสุดใน Search Space ไปพร้อมกันทั้งกลุ่ม GA จึงมีความเป็นไปได้ที่จะได้เปรียบวิธีที่ Search จากจุดที่หนึ่งไปจุดที่สองและจุดต่อไปเรื่อยๆ ทีละจุดตามวิธีแบบเดิมๆ GA จึงเป็นเทคนิคที่จะให้ค่า local optimal ได้น้อยกว่าวิธีอื่นๆ แม้ไม่เสมอไปที่ GA จะให้ค่าที่ดีที่สุดก็ตาม

Goldberg (1989) อธิบายว่า GA เป็นกระบวนการที่เลียนแบบแนวคิดของทฤษฎีการคัดเลือกตามธรรมชาติ โดยเริ่มจากจำนวนสมาชิกทั้งหมดใน Search Space เป็นการจำลองโคโมโซม ซึ่งเป็นการเลียนแบบส่วนประกอบของเซลล์ในสิ่งมีชีวิตที่เป็นองค์ประกอบของสิ่งมีชีวิต เพื่อใช้ในการหาค่าตอบจาก Search Space และกระบวนการที่ใช้ในการค้นหาค่าตอบก็ได้เลียนแบบมาจากกระบวนการทางธรรมชาติ เช่นการ ซึ่งสามารถกล่าวได้ว่า GA มีลักษณะดังนี้

1. GA เป็นการดำเนินการด้วยการแปลงรูปแบบของตัวพารามิเตอร์ ไม่ได้ใช้ตัวพารามิเตอร์โดยตรงซึ่งมีประโยชน์ในกรณีที่มีตัวแปรมาก หรือตัวแปรมีลักษณะที่ไม่สามารถจะหาค่าความเหมาะสมได้พร้อมกัน
2. GA เป็นการดำเนินการหาค่าตอบจากประชากรของจุดทั้งหมด ไม่ได้ค้นหาค่าตอบจากจุดใดจุดหนึ่ง
3. GA จะใช้ฟังก์ชันเป้าหมายเป็นเพียงแนวทางและตัวชี้นำเท่านั้น ไม่ได้หาค่าตอบจากฟังก์ชันเป้าหมายโดยตรง และไม่ได้ใช้ความรู้อื่นใดเข้ามาช่วยในการหาค่าตอบ จึงเหมาะสมกับฟังก์ชันที่ยุ่งยากซับซ้อน
4. GA จะใช้วิธีการการวิเคราะห์ทางความน่าจะเป็น ไม่ใช่วิธีการวิเคราะห์ทางดิเทอร์มินิสติก

ในทางปฏิบัติ GA ประกอบด้วยกระบวนการต่างๆดังนี้

1. Reproduction เป็นกระบวนการที่ใช้คัดเลือก Gene ในแต่ละ String นั่นคือ Gene ใดที่มีประสิทธิภาพดีก็จะถูกลอกเลียนแบบไปสร้าง String ใหม่ ส่วน Gene ใดที่มีประสิทธิภาพไม่ดีก็จะถูกคัดออก
2. Crossover เป็นกระบวนการแลกเปลี่ยน Gene ของ String ในประชากร ซึ่งทำให้ได้ Gene ในประชากรรุ่นหลังๆมี ประสิทธิภาพที่ดีขึ้นเรื่อยๆ
3. Permutation หรือกระบวนการกลายพันธุ์ ซึ่งเป็นลักษณะการเปลี่ยนแปลง Gene ที่ตำแหน่งใดๆของ String ในประชากรหลังจากกระบวนการ Crossover ขึ้นอย่างสุ่ม ทั้งนี้เพื่อป้องกันการเกิดการหาค่าเหมาะสมเฉพาะแห่ง (local optimal)

การหาโครงสร้างที่เหมาะสมของโครงข่ายประสาทเทียม ในทางปฏิบัติโดยทั่วไปจะใช้วิธีการลองผิดลองถูกเพื่อปรับเปลี่ยนโครงสร้างได้แก่ การปรับเปลี่ยนจำนวนชั้นแอบแฝง และ/หรือ จำนวนหน่วยในชั้นแอบแฝงแต่ละชั้น ซึ่งวิธีการดังกล่าวยุ่งยากและเสียเวลา ดังนั้นเพื่อความสะดวกโปรแกรมที่ใช้ได้รวมวิธีการหาค่าเหมาะสมด้วยวิธีทางพันธุศาสตร์เพื่อช่วยในการหาโครงสร้างของโครงข่ายประสาทเทียมที่เหมาะสมแบบอัตโนมัติ ซึ่งขั้นตอนในการทำงานดังนี้ (เสรี, 2544)

1. กำหนดจำนวนขนาดของประชากรเริ่มต้น (Population size) ที่เป็นตัวแทนโครงสร้างของโครงข่าย
2. กำหนดโครงสร้างของโครงข่ายด้วย Gene
3. เริ่มการคำนวณหา Output ของโครงข่าย
4. เปรียบเทียบและเลือกโครงข่ายที่ให้ผลการคำนวณที่เหมาะสมแล้วเริ่มกระบวนการ Reproduction, Crossover และ Mutation
5. กลับไปดำเนินการตามข้อ 2 จนกระทั่งได้ผลการคำนวณที่ดีที่สุด

การประยุกต์ใช้โครงข่ายประสาทเทียมในงานวิศวกรรมแหล่งน้ำ

ในปัจจุบันนี้ได้มีการประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมสำหรับงานทางด้านแหล่งน้ำอย่างมากในรูปแบบต่างๆกัน เช่น การคาดการณ์ปริมาณน้ำท่า การเคลื่อนตัวของคลื่นน้ำท่วม หาความสัมพันธ์ระหว่างปริมาณฝน และน้ำท่า ดังตัวอย่างการศึกษาที่ผ่านมาในอดีตต่อไปนี้

Siang J.J. (1992) ได้ใช้แบบจำลอง BP Model ชนิด Feed Forward และการเรียนรู้ชนิด Supervised Learning มาใช้ในการคาดการณ์ปริมาณน้ำท่ารายเดือนของแม่น้ำโขงจำนวน 5 สถานี ในการศึกษาครั้งนี้เป็นการคาดการณ์ปริมาณน้ำท่ารายเดือนโดยใช้แบบจำลอง BP Model ขนาดเล็กที่ประกอบด้วย จำนวนชั้นแฝงเพียง 1 ชั้น และจำนวนหน่วยในชั้นแฝงมีค่าตั้งแต่ 1-3 หน่วย เท่านั้น จากการศึกษาแบบจำลอง BP Model ในครั้งนี้พบว่า การเพิ่มจำนวนหน่วยในชั้นแฝงไม่ได้ช่วยให้ผลลัพธ์ที่ได้รับดีขึ้นเท่าที่ควร

Dartus และคณะ (1993) ได้ใช้แบบจำลองโครงข่ายประสาทเทียมในการศึกษาการเคลื่อนตัวของคลื่นน้ำท่วมในทางน้ำเปิด โดยมีวัตถุประสงค์ที่จะแสดงให้เห็นว่าแบบจำลองชนิดนี้สามารถนำมาประยุกต์ใช้ในการจัดการกับระบบบำบัดน้ำเสียได้ ซึ่งผลงานวิจัยได้แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองที่ให้คำตอบที่ใกล้เคียงกับค่าความเป็นจริง โดยต้องผ่านกระบวนการเรียนรู้อย่างเข้มข้น

Gul R. (1993) ได้ประยุกต์ใช้แบบจำลองระบบโครงข่ายประสาทเทียมชนิด Back Propagation ที่มีจำนวนชั้นแฝงจำนวน 1 ชั้น มาใช้คาดการณ์ปริมาณน้ำท่ารายวันล่วงหน้ากับแม่น้ำน่าน ที่อำเภอเมือง และที่บ้านแก่งสร้าง จังหวัดน่าน โดยใช้ข้อมูลปริมาณฝนรายวันของแต่ละสถานี และข้อมูลปริมาณน้ำท่ารายวันของแต่ละสถานีในพื้นที่ลุ่มน้ำย้อนหลัง 5 วัน จากการศึกษาพบว่าผลการคาดการณ์ปริมาณน้ำท่ารายวันล่วงหน้าดีที่ถูกต้องและใกล้เคียงกับค่าที่วัดได้ จะต้องประกอบด้วยตัวแปรนำเข้าเป็นข้อมูลปริมาณฝนรายวัน และข้อมูลปริมาณน้ำท่ารายวันของสถานีที่จะทำการคาดการณ์ย้อนหลัง และค่าปริมาณน้ำท่ารายวันของสถานีทางด้านเหนือลำน้ำ และข้อมูลปริมาณฝนรายวันของสถานีในพื้นที่ทางด้านเหนือลำน้ำของสถานีที่จะทำการคาดการณ์ปริมาณน้ำท่ารายวัน

Karunanithi และคณะ (1994) ได้แสดงให้เห็นว่าแบบจำลองระบบโครงข่ายประสาทเทียมสามารถพยากรณ์ข้อมูลการไหลของแม่น้ำ Huron ในรัฐมิชิแกน ประเทศสหรัฐอเมริกา และได้มีการเปรียบเทียบกับแบบจำลอง Analytic Nonlinear Power Model ซึ่งผลที่ออกมาอยู่ในเกณฑ์ที่น่าพอใจ

Gautam M.R. (1997) ได้ใช้แบบจำลองระบบโครงข่ายประสาทเทียมชนิด Back Propagation มาประยุกต์ใช้กับการคาดการณ์ปริมาณน้ำท่ารายวันล่วงหน้า หนึ่งและสองวันตามลำดับ ที่แม่น้ำป่าสัก อ.หล่มสัก อ.วิเชียรบุรี และ อ.บัวชุม จ.เพชรบูรณ์ โดยจะนำผลการคาดการณ์ปริมาณน้ำท่ารายวันที่ได้ ไปเปรียบเทียบกับแบบจำลองอื่นๆ เช่น แบบจำลอง Tank และ แบบจำลอง NAM ทั้งนี้ได้นำแบบจำลองเชิงคาดการณ์ (Stochastic Model) ชนิด Multiple Linear Regression มาประยุกต์ใช้งานร่วมกับ แบบจำลอง Tank และแบบจำลอง NAM จากผลการศึกษาพบว่า แบบจำลองระบบโครงข่ายประสาทเทียมใช้ผลลัพธ์เป็นที่น่าพอใจเมื่อเปรียบเทียบกับแบบจำลอง Tank และแบบจำลอง NAM ที่ประยุกต์ร่วมกับแบบจำลองเชิงคาดการณ์

Shamseldin A.Y. (1997) ได้ทำการเสนอแนวความคิดและเทคนิคการนำระบบโครงข่ายประสาทเทียมมาประยุกต์ใช้หาความสัมพันธ์ระหว่างปริมาณฝน และน้ำท่าโดยการนำข้อมูลนำเข้าต่างๆกัน 4 รูปแบบมาใช้ โดยรูปแบบแรกข้อมูลนำเข้าจะประกอบด้วย ข้อมูลปริมาณฝนเพียงอย่างเดียว รูปแบบที่สอง ข้อมูลปริมาณฝนเป็นอนุกรมเวลาและข้อมูลปริมาณน้ำท่า รูปแบบที่สาม ข้อมูลนำเข้าประกอบด้วยข้อมูลปริมาณฝนของสถานีนั้น และค่าเฉลี่ยข้อมูลปริมาณฝนที่เป็นอนุกรมเวลา และข้อมูลปริมาณน้ำท่าของสถานีข้างเคียง รูปแบบที่สี่ ข้อมูลนำเข้าประกอบด้วยข้อมูลปริมาณฝน ข้อมูลฝนที่เป็นอนุกรมเวลา ค่าเฉลี่ยของข้อมูลฝนของสถานีใกล้เคียง ค่าปริมาณน้ำท่าของสถานีใกล้เคียง และข้อมูลของสถานีน้ำท่าของสถานีนั้นๆ

Sureerattanan S. และ Phien H.N. (1997) ได้นำแบบจำลองระบบโครงข่ายประสาทเทียมชนิด Back Propagation มาใช้กับปัญหาการคาดการณ์ปริมาณน้ำท่ารายวันเพื่อใช้ในการควบคุมปริมาณน้ำหลาก และควบคุมอุทกภัยที่เกิดขึ้น โดยศึกษาในพื้นที่ลุ่มน้ำแม่กลอง โดยมีข้อมูลนำเข้าได้แก่ ปริมาณฝนและข้อมูลปริมาณน้ำท่ารายวันจากข้อมูลในอดีตของสถานีในกลุ่มน้ำจำนวน 5 สถานี พบว่าผลลัพธ์ที่ได้อยู่ในเกณฑ์ที่ดี

Nam (1998) ได้ประยุกต์ใช้แบบจำลองระบบโครงข่ายประสาทเทียมชนิด Back Propagation ในการพยากรณ์ปริมาณน้ำท่ารายเดือนของแม่น้ำ Black ในประเทศเวียดนาม พบว่าสามารถใช้แบบจำลองชนิด Back Propagation ในการพยากรณ์ปริมาณน้ำท่ารายเดือนได้เป็นอย่างดี โดยมีความละเอียดถูกต้องของผลพยากรณ์เป็นที่น่าพอใจ

Campolo และคณะ (1999) ได้พัฒนาแบบจำลองระบบโครงข่ายประสาทเทียมในการพยากรณ์ปริมาณน้ำท่ารวมในแม่น้ำ Tagliamento ประเทศอิตาลี ในช่วงที่ฝนตกหนัก พบว่าแบบจำลองให้ผลการ

พยากรณ์ที่มีความละเอียดถูกต้องสูง และพบว่าพยากรณ์สามารถทำได้ล่วงหน้าเท่ากับเวลาที่ฝนเริ่มตกจนถึงสถานีที่ตรวจสอบ โดยยังคงมีความถูกต้องเป็นที่น่าพอใจ

Sajikumar และ Thandaveswara (1999) ได้ประยุกต์ใช้แบบจำลองระบบโครงข่ายประสาทเทียมชนิด Temporal Back Propagation กับการวิเคราะห์แบบจำลองน้ำฝน-น้ำท่ารายเดือน และได้ตรวจสอบประสิทธิภาพของแบบจำลองนี้โดยเปรียบเทียบกับแบบจำลอง Functional Series พบว่าแบบจำลองโครงข่ายประสาทเทียมมีประสิทธิภาพสูงกว่า โดยเฉพาะในกรณีที่ข้อมูลมีน้อย

รัชชัย (2543) ได้ประยุกต์ใช้แบบจำลองระบบโครงข่ายประสาทเทียมสำหรับพยากรณ์ปริมาณน้ำท่ารายวัน ในบริเวณลุ่มน้ำน่านตอนบน ที่สถานีวัดน้ำท่า อ.ท่าวังผา โดยอาศัยข้อมูลฝน ข้อมูลการระเหย และข้อมูลน้ำท่าที่วัดได้ในอดีต ในการศึกษาครั้งนี้มีจุดประสงค์หลักคือ การประเมินผลการทำงานของระบบโครงข่ายประสาทเทียม และความแม่นยำ โดยตั้งสมมติฐานว่า ข้อมูลพยากรณ์น้ำฝนเป็นข้อมูลที่ถูกต้องมีค่าเท่ากับข้อมูลที่วัดได้จากสถานี ผลที่ได้ออกมาสั้นให้ผลดี ทั้งในด้านการปรับแต่งแบบจำลองและในด้านการทดสอบ

Anmala และคณะ (2000) ได้ประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมชนิด Feed Forward กับพยากรณ์ปริมาณน้ำท่ารายเดือนโดยเทียบกับแบบจำลองอื่น ๆ พบว่ามีประสิทธิภาพใกล้เคียงกัน และหากเพิ่มกระบวนการ Feedback จะทำให้แบบจำลองมีประสิทธิภาพขึ้น

Coulilyaly และคณะ (2000) ได้พัฒนาแบบจำลองโครงข่ายประสาทเทียมชนิด Back Propagation โดยอาศัยหลักการ Stop Training Approach (STA) ในการพยากรณ์ปริมาณน้ำท่ารายวันโดยเปรียบเทียบกับแบบจำลองเชิงสถิติและแบบจำลองเชิงปฏิบัติการ พบว่าแบบจำลองระบบโครงข่ายประสาทเทียมมีประสิทธิภาพสูงกว่าในด้านความถูกต้องของการพยากรณ์

Gautam และคณะ (2000) ได้ประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมชนิด Back Propagation กับการวิเคราะห์ปริมาณน้ำท่าในพื้นที่ป่าลุ่มน้ำ โดยใช้ข้อมูลด้านกายภาพของกลุ่มน้ำเป็นข้อมูลด้านเข้าของแบบจำลอง พบว่าแบบจำลองนี้สามารถใช้ในการวิเคราะห์ได้โดยมีความถูกต้องเพียงพอ

Manusthiparom C. (2000) ได้ทำการประยุกต์ใช้โครงข่ายประสาทเทียมชนิด Back Propagation ในการพยากรณ์ระดับและปริมาณการไหลของน้ำรายชั่วโมง ที่บริเวณสะพานกรุงเทพโดยจะพยากรณ์ที่ 1 2

และ 3 ชั่วโมงล่วงหน้า โดยระบบโครงข่ายประสาทเทียมจะมีการปรับเทียบ และตรวจพิสูจน์แบบจำลอง จากข้อมูลที่ได้จากการวัดจริง และได้จากการคำนวณโดยแบบจำลอง MIKE 11 โดยแบ่งการตรวจสอบความ ถูกต้องออกเป็น สองกรณีด้วยกันคือ จำนวนข้อมูลด้านออกหนึ่งและสองหน่วย คือ ระดับน้ำและอัตราการ ไหล ในกรณีที่จำนวนหน่วยในชั้นข้อมูลด้านออกมีสองหน่วยประกอบด้วยระดับน้ำและอัตราการไหล จะ ให้ผลลัพธ์ที่มีความถูกต้องน้อยกว่ากรณีที่จำนวนชั้นข้อมูลด้านออกแยกกันคือ ระหว่างระดับน้ำและอัตรา การไหล

วุฒิชัย (2544) ได้ทำการประยุกต์ใช้ระบบโครงข่ายประสาทเทียมชนิด Back Propagation ชนิด Full Link Multilayer Feed Forward มาใช้ในการคาดการณ์ปริมาณน้ำท่ารายวันล่วงหน้า โดยการนำไป ประยุกต์ใช้ในการคาดการณ์ใน แม่น้ำจันทบุรี และแม่น้ำคลองอุตะเถา โดยผลการศึกษพบว่าผลการ คาดการณ์ปริมาณน้ำท่ารายวันล่วงหน้ามีความน่าเชื่อถืออยู่ในเกณฑ์ที่ดีสำหรับการคาดการณ์ปริมาณน้ำท่า รายวัน 1 วันล่วงหน้า และค่าความน่าเชื่อถือมีค่าลดลงสำหรับการคาดการณ์ปริมาณน้ำท่ารายวัน 2 วัน ล่วงหน้า

การประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมสำหรับการประมาณค่าการคายระเหยอ้างอิง เท่าที่พบยังไม่แพร่หลายเช่น

Kumar M. et al.(2002) ได้ประยุกต์ใช้โครงข่ายประสาทเทียมสำหรับการประมาณปริมาณการคาย ระเหยอ้างอิงรายวันที่เมือง Davis มลรัฐ California และทำการเปรียบเทียบประสิทธิภาพของวิธีดังกล่าวกับ วิธีของ Penman-Montieth และวิธี Lysimeter ข้อมูลอุตุนิยมที่ใช้ในครั้งนีเพื่อเป็นข้อมูลในชั้นนำเข้าของ โครงข่ายประสาทเทียมคือ แสงอาทิตย์ อุณหภูมิสูงสุดและต่ำสุด ความชื้นสัมพัทธ์สูงสุดและต่ำสุด ความเร็ว ลม ในการทดสอบครั้งนี้ได้ใช้โครงข่ายประสาทเทียม3ชนิดได้แก่ standard back-propagation โดยใช้ อัตราการเรียนรู้เท่ากับ 0.2 และ 0.8 และ back-propagation ผลการทดสอบพบว่าโครงข่ายประสาทเทียม สามารถประมาณค่าการคายระเหยได้ดีกว่าวิธี Penman-Montieth

Sudheer K.P. et al.(2003) ได้ศึกษาการประมาณค่าการคายระเหยแท้จริงจากการใช้ข้อมูล อุตุนิยมวิทยาที่มีอยู่อย่างจำกัดโดยใช้เทคนิคการคำนวณทางประสาท โดยได้เลือกใช้โครงข่ายประสาทเทียม ชนิด radial-basis function (RBF) สำหรับการคำนวณปริมาณการคายระเหยรายวันที่แท้จริงของข้าว นอกจากนี้ยังได้ศึกษาถึงผลของการเปลี่ยนแปลงจำนวนพารามิเตอร์ทางอุตุนิยมต่างๆที่ใช้เป็นข้อมูลนำเข้า

โครงข่ายประสาทเทียมที่มีผลต่อประสิทธิภาพการคำนวณ ได้แก่ กลุ่มที่ 1 ใช้ข้อมูล อุณหภูมิเฉลี่ยของอากาศ ความชื้นสัมพัทธ์ ความเร็วลม และ แสงอาทิตย์ กลุ่มที่ 2 ใช้ข้อมูล อุณหภูมิเฉลี่ยของอากาศ ความชื้นสัมพัทธ์ และ ความเร็วลม กลุ่มที่ 3 ใช้ข้อมูล อุณหภูมิเฉลี่ยของอากาศอย่างเดียว ตามลำดับ ผลการศึกษาพบว่าการใช้ข้อมูลกลุ่มที่ 3 ให้ประสิทธิภาพการคำนวณดีที่สุดแม้จะใช้จำนวนพารามิเตอร์ทางอุตุนิยมน้อยที่สุด อย่างไรก็ตามการศึกษานี้ยังคงจำกัดเพียงแค่พีชคณิตเดียวเท่านั้นคือค่า

เอกสารอ้างอิง

ธวัชชัย ดิงส์ัญชลี. 2543. การใช้แบบจำลอง Neural Network สำหรับการพยากรณ์น้ำท่า, น. 1-12. ในการประชุมวิชาการเรื่องการพยากรณ์และเตือนภัยน้ำท่วม. จัดโดยชมรมนักอุทกวิทยา และ วิทยาลัยชลประทาน, กรมชลประทาน, อำเภอปากเกร็ด, จังหวัดนนทบุรี.

วุฒิชัย บุญผ่อง. 2544. การประยุกต์ใช้ระบบโครงข่ายประสาทเทียมในการคาดการณ์ปริมาณน้ำท่ารายวัน. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยมหานคร.

เสรี ศุภราทิตย์. 2544. คู่มือการพยากรณ์น้ำโดยโครงข่ายประสาทเทียม. สถาบันพัฒนาการชลประทาน, นนทบุรี

Anmala, J., B. Zhang and R.S. Govindaraju. 2000. Comparison of ANNS and empirical approaches for predicting watershed runoff. *Journal of Water Resources and Management*, Vol.126 (3): 156-166.

Campolo, M., P. Andreussi and A. Soldali. 1999. River flood forecasting with a neural network model. *Water Resources Research*, Vol.35(4): 1191-1197.

Coulibaly, P., F. Anctil and B. Bobee. 2000. Daily reservoir inflow forecasting using artificial neural network with stopped training approach. *Journal of Hydrology*: 244-257

- Dartus, D., J.M. Courivaud and L. Dedeker. 1993. Use of a Neural Net for the study of the flood wave propagation in an open channel. *Journal of Hydraulic Research*, Vol. 31(2), 161-170.
- Gautum, M.R. 1997. *Flood Forecasting Model for Pasak River Basin Thailand*. Thesis, Asian Institute of Technology, Bangkok, Thailand.
- Gautum, M.R., K. Watanake and H. Saegusa. 2000. Runoff analysis in humid forest catchment with artificial neural network. *Journal of Hydrology*, Vol. 235, 117-136.
- Goldberg, D.E. 1989. *Genetic Algorithms in search optimization & machine learning*. Addison-Wesley Reading. Mass, USA.
- Gul, R. 1993. *The Application of Back Propagation Model to Daily Flow Forecasting*. Thesis, Asian Institute of Technology, Bangkok, Thailand.
- Karunanithi, N., J. Grenney, D. Whitley and K. Bovee. 1994. Neural network for river flow prediction. *Journal of Computing in Civil Engineering*, Vol. 8(2), 201-220.
- Manusthiparom, C. 2000. *Neural Network Models for Forecasting Hourly Water Levels and Discharges of Choa Phraya River at Bangkok Memorial Bridge*. Thesis, Asian Institute of Technology, Bangkok, Thailand.
- Nam, L.H. 1998. *Simulation and forecasting of monthly streamflow using a Back-Propagation model a case study: Back river in Vietnam*. Thesis, Asian Institute of Technology, Bangkok, Thailand.

- Rumelhart, D.E. and J.L. McClelland (Eds). 1986. Parallel distributed processing, **Exploration in the microstructures of Cognition**, Vol. 1, The MIT Press.
- Sajikumar, N. and B.S. Thandaveswara. 1999. A non-linear rainfall-runoff model using an artificial neural network. **Journal of Hydrology**, Vol. 216, 31-55.
- Shamseldin, A.Y. 1997. Application of Neural Networks Technique to Rainfall Runoff Modeling, **Journal of Hydrology**, Vol.199, 272-294.
- Siang, J.J. 1992. **Application of Back Propagation Method in Forecasting Problem**. Thesis, Asian Institute of Technology, Bangkok, Thailand.
- Sureeratnan, S. and H.N. Phien. 1997. Back-propagation network for daily streamflow forecasting. **Water Resources Journal**, St/SCAP/SER-C/195, 1-7.
- Tsoukalas, L.H. and R.E. Uhrig. 1997. **Fuzzy and Neural Approaches in Engineering**. First Edition. John Wiley & Sons. New York.