

บทที่ 2

ทฤษฎี แนวคิด และผลงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและผลงานวิจัยที่เกี่ยวข้อง โดยเนื้อหาจะแบ่งเป็นส่วนๆ ซึ่งในส่วนแรกจะกล่าวถึงทฤษฎีโครงข่ายประสาทเทียม (Artificial Neural Networks, ANNs) และทฤษฎีโครงข่ายประสาทเทียมแบบแพร่กลับ (Back-Propagation Network, BPN) จากนั้นจะกล่าวถึงการออกแบบการทดลอง (Design of Experiment) ทฤษฎีสมบัติของแม่เหล็ก (Magnetic Properties) การจำลองสถานการณ์มอนติคาร์โล (Monte Carlo Simulation) การวิเคราะห์การถดถอย (Regression analysis) และสรุปสาระสำคัญจากเอกสารที่เกี่ยวข้องตามลำดับ

2.1 ทฤษฎีโครงข่ายประสาทเทียม

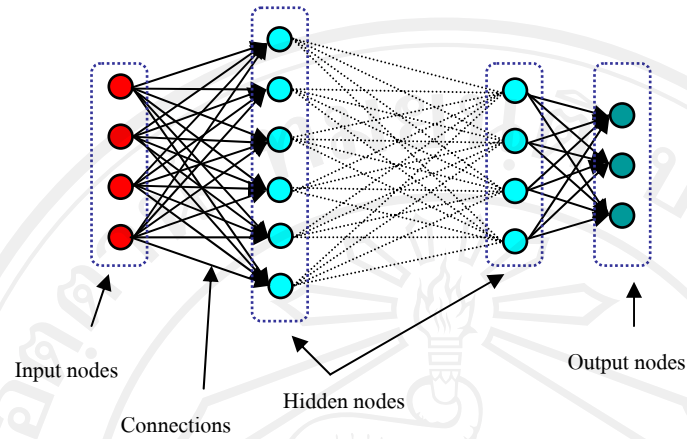
โครงข่ายประสาทเทียม คือ รูปแบบการคำนวณในทางคณิตศาสตร์ที่เลียนแบบตามระบบการทำงานของกระแสประสาทของมนุษย์ และได้มีผู้นิยามความหมายของโครงข่ายประสาทเทียมไว้เป็นจำนวนมาก ตัวอย่างเช่น โครงข่ายประสาทเทียม คือ กระบวนการประมวลผลข้อมูลในรูปแบบที่มีการเชื่อมต่อระหว่างโหนด (Node) โดยที่ความสามารถในการคำนวณของโครงข่ายจะขึ้นกับความแข็งแรง (Strengths) หรือค่าน้ำหนัก (Weight) ภายในที่เชื่อมต่อระหว่างโหนดแต่ละจุด ซึ่งค่าน้ำหนักจะได้มาจากการปรับค่า หรือการเรียนรู้ โดยการฝึกจากรูปแบบที่กำหนด (Kevin, 1997) หรือ

โครงข่ายประสาทเทียม เป็นระบบการประมวลผลแบบขนานในรูปแบบการคำนวณที่ไม่เป็นเชิงเส้นของแต่ละเซลล์ประสาทหรือนิวรอน (Neurons) ซึ่งมีรูปแบบการทำงานที่คล้ายกับระบบกระแสประสาทของมนุษย์ โดยภายในโครงข่ายจะมีการคำนวณความแข็งแรงของการเชื่อมโยง (Strengths of connection) เพื่อให้ได้ผลลัพธ์ออกมา (Satish, 2005)

โดยทั่วไปโครงสร้างของโครงข่ายประสาทเทียมจะประกอบด้วย

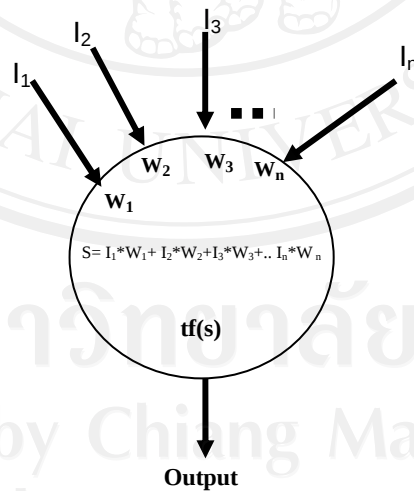
1. ชั้นอินพุต (Input layer) เป็นชั้นรับข้อมูลเข้าสู่ระบบเพื่อนำมาคำนวณหาค่าผลลัพธ์
2. ชั้นซ่อน (Hidden layer(s)) เป็นชั้นที่มีส่วนทำการประมวลผลอยู่ภายในและโครงข่ายประสาทเทียมสามารถมีชั้นซ่อนได้หลายชั้น
3. ชั้นเอาต์พุต (Output layer) เป็นชั้นที่แสดงค่าผลลัพธ์สุดท้ายที่ได้จากการคำนวณโดยระบบโครงข่ายประสาทเทียม

ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 โครงสร้างของระบบโครงข่ายประสาทเทียมแบบ Multi-Layer-Perceptron
(ที่มา: นวกัก.2006, online)

โดยปกติแล้วในการทำงาน ระบบจะมีการเชื่อมโยงระหว่างโหนดของแต่ละชั้นและการกำหนดค่าน้ำหนักของอินพุตแต่ละตัวเมื่อมีอินพุตเข้ามาระบบก็จะเอาค่าอินพุตมาคูณกับค่าน้ำหนักของแต่ละขาแล้วแปลงผ่านฟังก์ชันถ่ายโอนข้อมูล จะได้ค่าผลลัพธ์ซึ่งจะเป็นอินพุตของชั้นถัดไป และจนกว่าจะได้ค่าผลลัพธ์สุดท้าย ดังรูปที่ 2.2



รูปที่ 2.2 การคูณระหว่างอินพุตและค่าน้ำหนัก

(source: Qnet manual, 2000)

ฟังก์ชันถ่ายโอนข้อมูล (Transfer function)

เป็นฟังก์ชันที่ใช้ควบคุมสัญญาณที่จะส่งไปยังโหนดเอาต์พุต โดยจะปรับค่าผลคูณหว่างอินพุตกับค่าน้ำหนักของข้อมูลให้อยู่ในช่วง 0 ถึง 1 หรือ -1 ถึง 1 ซึ่งฟังก์ชันถ่ายโอนข้อมูลมีหลายลักษณะดังในรูป 2.3 และแต่ละแบบจะมีวิธีการคำนวณดังนี้

1. ฟังก์ชันซิกมอยด์ (Sigmoid function) หรือ S-shape curve

$$f(x) = \frac{1}{1 + \exp(-x)}$$

ซึ่งค่าอนุพันธ์ฟังก์ชันคือ

$$f'(x) = f(x)[1 - f(x)]$$

2. ฟังก์ชันเกาส์เซียน (Gaussian function)

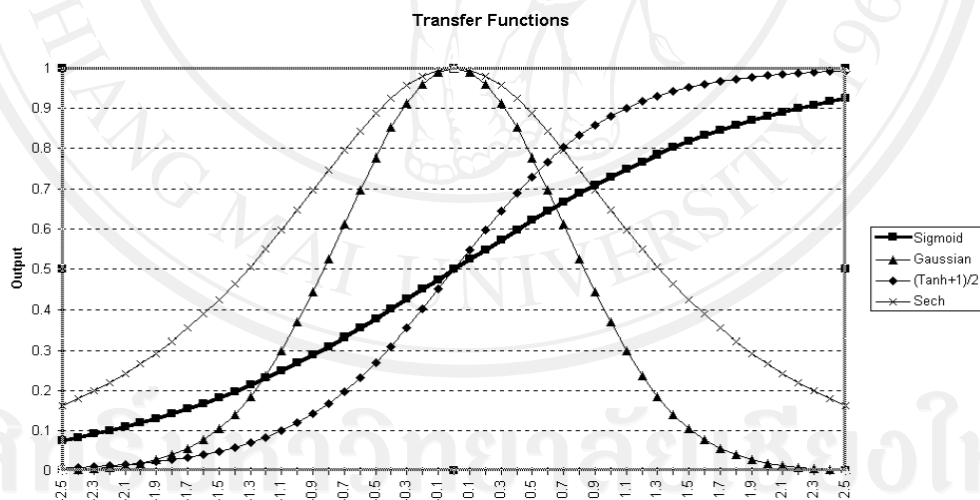
$$g(x) = \exp(-x^2)$$

3. ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (Hyperbolic tangent function)

$$h(x) = \frac{\tanh(x) + 1}{2}$$

4. ฟังก์ชันไฮเพอร์โบลิกซีแคนต์ (Hyperbolic secant function)

$$f(x) = \text{sech}(x)$$



รูปที่ 2.3 ฟังก์ชันถ่ายโอนข้อมูลแบบต่างๆ

(source: Qnet 2000 manual)

ฟังก์ชันถ่ายโอนข้อมูลมีหลายแบบดังที่กล่าวมา และในการทดลองนี้จะใช้แบบซิกมอยด์ เพราะเป็นฟังก์ชันที่นิยมใช้กับการเรียนรู้แบบแพร่กลับ ซึ่งความสัมพันธ์ระหว่างค่าของฟังก์ชันที่จุดใด ๆ กับค่าอนุพันธ์ของฟังก์ชันที่จุดใด ๆ จะมีค่าลดลงในช่วงของการเรียนรู้ (อนุสรณ์ , 2543)

จำนวนของชั้นซ่อน (Number of hidden layers)

การกำหนดจำนวนของชั้นซ่อน ยังไม่มีวิธีการใดที่สามารถกำหนดจำนวนของชั้นซ่อนให้มีค่าแน่นอนได้ แต่ก็มีทฤษฎีกล่าวไว้ว่า ชั้นซ่อนเพียงชั้นเดียวก็เพียงพอแล้ว สำหรับการนำไปใช้กับปัญหาที่ไม่เป็นเชิงเส้นใดๆ แต่ในบางกรณีการเพิ่มชั้นซ่อนก็สามารถช่วยให้การเรียนรู้ง่ายขึ้น และรวดเร็วขึ้นและอีกปัญหาหนึ่งในชั้นซ่อน นั่นคือควรจะมีจำนวนนิวรอนกี่นิวรอนซึ่งก็ยังไม่มีความชัดเจนตายตัวใดๆ กำหนดให้อย่างแน่นอนได้ (สุรยุทธ์, 2541) การกำหนดจำนวนโหนดของชั้นซ่อนน้อยเกินไป จะจำกัดความสามารถในการเจเนอราไลซ์เซชัน (Generalization) ของโครงข่ายฯ ส่วนการกำหนดจำนวนโหนดมากเกินไปจะนำไปสู่การเกิด “Over training” หรือ “Memorization” ดังนั้นจึงต้องกำหนดโดยวิธีการลองผิดลองถูก (Su and Hsieh, 1998)

อัตราการเรียนรู้ (Learning rate)

เป็นค่าสัดส่วนคงที่ของการเปลี่ยนแปลงค่าน้ำหนัก อัตราการเรียนรู้ค่ามากจะส่งผลให้โครงข่ายฯมีการเรียนรู้เร็วซึ่งอาจนำไปสู่ภาวะ “local minima” หรือเกิดการแกว่งของระบบได้

อัตราการเรียนรู้แบบแปรกลับ ที่มีอัตราการเรียนรู้ค่าน้อยๆ จะทำให้การเรียนรู้สำเร็จได้ช้า แต่ถ้าให้อัตราการเรียนรู้ค่ามากๆ จะทำให้การเรียนรู้ใช้เวลานานไม่สามารถหาคำตอบได้จึงนิยมใช้ค่าน้อยๆก่อน เพราะหาคำตอบได้แน่นอน การปรับค่าอัตราการเรียนรู้เป็นวิธีจะช่วยให้การเรียนรู้เกิดความรวดเร็วมากขึ้น คือ การเปลี่ยนแปลงค่าอัตราการเรียนรู้ในระหว่างการเรียนรู้ เรียกว่า กฎ Delta-Bar-Delta โดยที่จะให้ค่าน้ำหนักแต่ละตัวมีค่าอัตราการเรียนรู้ของตัวเอง ถ้าค่าน้ำหนักเปลี่ยนแปลงไปในทิศทางเดียวกันติดต่อกันหลายครั้งอาจจะเพิ่มขึ้นเหมือนกันหรืออาจจะลดลงเหมือนกัน ค่าอัตราการเรียนรู้ของน้ำหนักนั้นควรจะเพิ่มขึ้นด้วย แต่ถ้าทิศทางของการเปลี่ยนแปลงค่าน้ำหนักเปลี่ยนไป คือ เพิ่มขึ้นหรือลดลงสลับกันค่าอัตราการเรียนรู้ควรจะมีค่าลดลง (อนุสรณ์, 2543)

โมเมนตัม (Momentum)

ค่าโมเมนตัมเป็นการคำนวณค่าสัดส่วนของค่าน้ำหนักสุดท้ายที่เปลี่ยนแปลงซึ่งจะถูกนำมาเพิ่มลงในค่าน้ำหนักใหม่ โดยทั่วไปค่าโมเมนตัมอาจจะใช้หรือไม่ใช้ก็ได้ แต่ค่านี้อาจมีคุณสมบัติช่วยป้องกันการแกว่ง (Oscillate) ของระบบโดยค่านี้อาจสัมพันธ์กับค่าอัตราเร็วในการเรียนรู้ คือถ้าอัตราเร็วในการเรียนรู้มีค่ามากแต่ค่าโมเมนตัมมีค่าน้อยจะทำให้โครงข่ายฯไม่เกิดการแกว่ง แต่จะมีผลต่อเวลาที่ใช้เรียนรู้

หลักการของค่าโมเมนต์อาศัยเหตุผลที่ว่าในกรณีที่ค่าของข้อมูลที่ใช้เรียนรู้นั้นค่ามีความแตกต่างกันมากเมื่อเทียบกับค่าข้อมูลส่วนใหญ่ เพื่อหลีกเลี่ยงการกระจายตัวของข้อมูลเช่นนี้ จึงต้องใช้ค่าอัตราการเรียนรู้ที่มีค่าน้อยๆ แต่การทำเช่นนี้จะส่งผลให้ใช้เวลานานในการเรียนรู้ เพราะค่าน้ำหนักที่ถูกต้อง (Weight correction term) มีความละเอียดมาก โครงข่ายจึงต้องมีการปรับค่าน้ำหนักหลายครั้ง ดังนั้นเพื่อให้กระบวนการเรียนรู้เกิดการลู่เข้า (Convergence) ได้เร็วขึ้นจึงต้องมีการเพิ่มค่าโมเมนต์เข้าไปในสมการของการปรับค่าน้ำหนัก (อนุสรณ์ , 2543)

ค่ารากที่สองของค่าผิดพลาดกำลังสองเฉลี่ย (Root Mean Square Error, RMS error)

เป็นค่าผิดพลาดที่ใช้ตรวจสอบการเรียนรู้ของโครงข่ายประสาทเทียมในขณะฝึกสอนซึ่งจะใช้ในการวัดความเที่ยงตรงของแบบจำลองและใช้เป็นเงื่อนไขในการลู่เข้าของโครงข่ายฯ สำหรับการคำนวณค่ารากที่สองของค่าผิดพลาดกำลังสองเฉลี่ย สำหรับชุดข้อมูล N ค่าซึ่งมีค่าผิดพลาดเท่ากับ $\{x_1, x_2, x_3, \dots, x_N\}$ เช่น

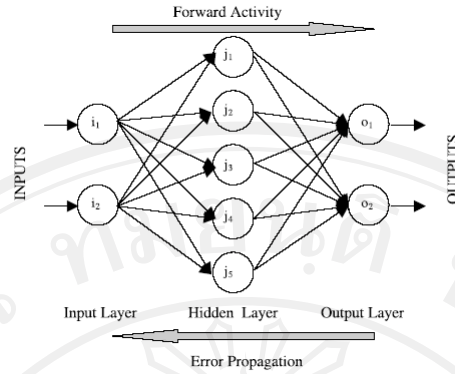
$$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_N^2}{N}}$$

2.2 ทฤษฎีโครงข่ายประสาทเทียมแบบแพร่กลับ

ทฤษฎีโครงข่ายประสาทเทียมแบบแพร่กลับเป็นทฤษฎีที่ใช้ในการฝึกโครงข่ายฯ โดยเรียนรู้แบบมีผู้สอนซึ่งจะมีการเปรียบเทียบค่าความแตกต่างระหว่างค่าผลลัพธ์ที่เอาต์พุตกับค่าผลลัพธ์เป้าหมายแต่ละคู่ จากนั้นจึงนำไปปรับค่าน้ำหนักในทุกๆ คู่ของอินพุต-เอาต์พุต ซึ่งการฝึกโครงข่ายฯ จะวนซ้ำไปหลายรอบจนกว่าจะได้สมรรถนะของเน็ตเวิร์คตามที่ต้องการ

อัลกอริทึมในการเรียนรู้ของโครงข่ายประสาทเทียมแบบแพร่กลับจะเริ่มจากการป้อนสัญญาณ คำนวณไปข้างหน้าแล้วตามด้วยการป้อนสัญญาณแบบย้อนกลับ ในครั้งแรกจะป้อนสัญญาณไปข้างหน้าโดยนำข้อมูลเข้าสู่แต่ละชั้นของโครงข่ายฯ แล้วคำนวณค่าระดับกิจกรรม (Activation level) ส่งผ่านไปขึ้นชั้นช่อนจากนั้นชั้นช่อนแต่ละชั้นจะรวมค่าผลลัพธ์ของข้อมูลแล้วคำนวณผ่านฟังก์ชันซิกมอยด์ เพื่อคำนวณค่าผลลัพธ์ของโครงข่ายฯ ต่อไป

ในขั้นตอนการแพร่กลับ โครงข่ายฯ จะทำการเปรียบเทียบค่าความแตกต่างระหว่างค่าผลลัพธ์จากโครงข่ายฯ และค่าผลลัพธ์เป้าหมายหรือค่า “ผิดพลาด” แล้วนำไปคำนวณเพื่อปรับค่าน้ำหนักผ่านชั้นช่อนแต่ละชั้นย้อนกลับไปดังในรูปที่ 2.4 ซึ่งจะมีผลทำให้ค่าน้ำหนักถูกต้องมากขึ้น และค่าผิดพลาดลดลง



รูปที่ 2.4 โครงสร้างของโครงข่ายประสาทเทียมแบบแพร่กลับ

(source: Güven and Kara, 2006)

และค่าผิดพลาดจะวัดโดยการคำนวณค่าผลรวมของค่าผิดพลาดกำลังสองทั้งหมดในแต่ละรอบการฝึก (Total Pattern Sum of Squares Error, PSS) เริ่มจากสมการที่ (2.1)

$$PSS = \sum_{j=1}^O (T_j - A_j)^2 \quad (2.1)$$

เมื่อ O คือ จำนวนหน่วยของผลลัพธ์, T_i คือ ค่าผลลัพธ์ที่ได้จากการฝึกโครงข่ายฯ และ A_j คือ ค่าผลลัพธ์จริง ในการฝึกจะทำการเข้าไปเรื่อยๆ จนกระทั่งเมื่อค่าผิดมาถึงค่าที่ผิดพลาดกำหนด (Tolerable error level) คือ $E \geq TSS$ ถ้า E คือ เงื่อนไขค่าผิดพลาดที่ระบุ หรือระบบจะหยุดการฝึกเมื่อผลรวมของค่าผิดพลาดกำลังสอง (Total Sum of Squares Error, TSS) ทั้งหมดมีค่าต่ำกว่าค่า E ดังในสมการที่ (2.2)

$$E \geq TSS = \sum_{j=1}^P PSS_i \quad (2.2)$$

เมื่อ P คือ จำนวนรูปแบบในหนึ่งชุดของการทดลอง (Batch)

ในการฝึกโครงข่ายฯ ด้วยทฤษฎีแบบแพร่กลับจะเปรียบเทียบคู่ผลลัพธ์แล้วคำนวณค่าความแตกต่างและค่าอนุพันธ์ฟังก์ชันซิกมอยด์ เพื่อหาค่าผิดพลาด ดังสมการที่ (2.3)

$$\Delta_{oi} = \frac{\partial E}{\partial W_{ij}} = (T_i - A_{oi}) A_{hi} (1 - A_{hi}) \quad (2.3)$$

เมื่อ Δ_{oi} คือ ค่าผิดพลาดที่คำนวณได้จากโหนดเอาต์พุตที่ i , W_{ij} หมายถึง ค่าน้ำหนักที่ i จากโหนดที่ j , E คือ ค่าผิดพลาดของโครงข่ายฯ, T คือ ค่าผลลัพธ์เป้าหมายของโหนดที่ i ,

A_{oi} คือ ค่าผลลัพธ์จากโครงข่ายโหนดที่ i , $A_{hi}(1-A_{hi})$ คือ อนุพันธ์ของฟังก์ชันซิกมอยด์ และ A_{hj} คือ ค่าการประมวลผลของโหนดที่ j ที่ชั้นซ่อนก่อนหน้า และจากค่าผิดพลาดที่คำนวณได้จะนำมาคำนวณค่าน้ำหนักที่เปลี่ยนไปดังสมการที่ (2.4)

$$\Delta W_{ij} = -k\Delta_{oi}A_{hj} \quad (2.4)$$

เมื่อ k คือ ค่าคงที่ ที่เรียกว่า อัตราการเรียนรู้ และเทอมของค่าผิดพลาดดังกล่าวจะถูกส่งย้อนกลับไปยังชั้นซ่อนซึ่งแต่ละโหนดจะหาค่าผลรวมและคำนวณค่าผิดพลาดได้ดังสมการที่ (2.5)

$$\Delta_{ji} = A_i(1-A_i)\sum k(\Delta_{ok}W_{jk}) \quad (2.5)$$

เมื่อ A_i คือ ค่าการประมวลผลของโหนดที่อินพุต, Δ_{ok} คือ ค่าผิดพลาดของโหนด k ที่ชั้นซ่อน, W_{jk} คือ ค่าน้ำหนักที่ j จากโหนดที่ k ค่าน้ำหนักที่เปลี่ยนไปจะถูกรวบรวมดังในสมการที่ (2.6)

$$\Delta V_{ji} = -k\Delta_{ji}A_i \quad (2.6)$$

ส่วนค่าน้ำหนักระหว่าง Output-unit และ Hidden-unit จะถูกปรับดังในสมการที่ (2.7)

$$W_{ij(new)} = W_{ij(old)} + \Delta W_{ij} \quad (2.7)$$

และค่าน้ำหนักระหว่าง Input-unit และ Hidden-unit จะถูกปรับดังในสมการที่ (2.8)

$$V_{ji(new)} = V_{ji(old)} + \Delta V_{ji} \quad (2.8)$$

จากนั้นโครงข่ายจะทดสอบเงื่อนไข ถ้าผลของค่าผิดพลาดมีค่ามากกว่าค่าที่กำหนด ก็จะกลับไปทำตามสมการที่ (2.3) ใหม่

และมีการเพิ่มค่าโมเมนตัมลงไปเพื่อเร่งอัตราการลู่เข้าและป้องกันการเข้าสู่สภาวะของค่าผิดพลาดต่ำสุดเฉพาะที่ (local minima) ซึ่งไม่ใช่ค่าผิดพลาดต่ำสุดที่แท้จริง ค่าโมเมนตัมจะช่วยเร่งให้โครงข่ายฯเข้าสู่สภาวะของค่าผิดพลาดโดยรวมต่ำสุด (global) ได้เร็วขึ้นในขณะประมวลผลข้อมูล ดังในสมการที่ (2.9)

$$\Delta W_{ij}(t+1) = \eta \sum (\Delta_j O_i) + \alpha \Delta W_{ij}(t) \quad (2.9)$$

เมื่อ α คือ อัตราการเปลี่ยนแปลงค่าโมเมนตัม

และถ้าค่าผิดพลาดของโครงข่ายเพิ่มขึ้นจากเดิม เช่น 1 เปอร์เซนต์ อัตราการเปลี่ยนแปลงค่าโมเมนตัมหรือ α จะถูกปรับเป็นค่าศูนย์จนกระทั่งค่าผิดพลาดของระบบเริ่มลดลง จากนั้นค่า α จะถูกกำหนดกลับไปหาค่าเริ่มต้นเดิม ซึ่งโมเมนตัมเป็นค่าที่จะกำหนดสัดส่วนของการเปลี่ยนแปลงค่าน้ำหนักสุดท้ายที่จะเพิ่มเข้าสู่ค่าน้ำหนักใหม่ เป็นต้น

โดยทั่วไปมักจะมีการนำทฤษฎีโครงข่ายฯแบบแพร่กลับไปใช้กับงานในด้านการจัดกลุ่ม (Classification) หรือการพยากรณ์ต่างๆ เช่น การทำนายค่าตลาดหุ้น การสังเคราะห์คำพูด (Speech Synthesis) การแยกประเภทรูปภาพ (Image classification) การแยกประเภทเสียง Sonar Return การเข้ารหัสข้อมูล (Information encoding) เป็นต้น

ทฤษฎีโครงข่ายประสาทเทียมจะถูกนำไปใช้ในการเรียนรู้ความสัมพันธ์ระหว่างคู่ของข้อมูลนำเข้าและค่าผลลัพธ์ที่มีความซับซ้อนซึ่งยากอธิบายโดยใช้สมการหรือแบบจำลองทั่วไป เนื่องจากโครงข่ายประสาทเทียมเป็นวิธีที่มีประสิทธิภาพและข้อดีดังที่ได้กล่าวมาในตอนต้น

การทดสอบ (Testing)

การทดสอบชุดข้อมูลเป็นการตรวจสอบสมรรถนะของโครงข่ายฯ จากค่าน้ำหนักที่ได้มาในระบบซึ่งเป็นการวัดความสามารถในการแยกแยะข้อมูลที่โครงข่ายฯไม่เคยเรียนรู้มาก่อน โดยการแบ่งข้อมูลที่ใช้ฝึกออกเป็นสองชุด คือ ชุดที่ใช้ฝึกและชุดทดสอบ จากนั้นชุดทดสอบข้อมูลจะถูกป้อนเข้าสู่ระบบที่ได้รับการฝึกเรียบร้อยแล้วเพื่อทำการคำนวณค่าแตกต่างระหว่างค่าผลลัพธ์จากโครงข่ายฯกับค่าจริง ยิ่งได้ค่าผิดพลาดน้อยโครงข่ายฯก็จะมีสมรรถนะมากขึ้นเท่านั้น

2.3 การออกแบบการทดลอง

การออกแบบการทดลองทางสถิติเป็นกระบวนการในการวางแผนการทดลองเพื่อให้ได้มาซึ่งข้อมูลที่เหมาะสม โดยวิธีการทางสถิติจะมีการรวบรวมข้อมูล การวิเคราะห์ การให้ผลการทดลองที่มีเหตุผลและการสรุปผล แนวคิดทางสถิติในการออกแบบการทดลองจึงจำเป็นถ้าจะต้องมี

การกล่าวสรุปผลจากข้อมูลทีวเคราะห์และวิธีการทางสถิติจะเป็นเครื่องมือวิเคราะห์ที่เป็นรูปธรรม ดังนั้นปัญหาของการทำการทดลองจะมีอยู่สองประเด็นที่สัมพันธ์กัน คือ การออกแบบการทดลอง และการวิเคราะห์ข้อมูลทางสถิติ เนื่องจากวิธีการวิเคราะห์จะขึ้นอยู่กับวิธีการออกแบบการทดลอง หลักการพื้นฐานของการออกแบบการทดลอง 3 ประการ คือ

การทดลองซ้ำ (Replication) เป็นการเพิ่มความเที่ยงตรงแม่นยำของข้อมูลมากขึ้นซึ่งจะส่งผลถึงการวิเคราะห์สรุปผลการทดลองและยังทำให้ผู้ทดลองสามารถประมาณค่าความผิดพลาดในการทดลองได้

การสุ่ม (Randomization) วิธีการวิเคราะห์ข้อมูลเชิงสถิติมักกำหนดว่าข้อมูลหรือความผิดพลาดจะต้องเป็นตัวแปรสุ่ม การสุ่มเป็นการช่วยสมดุลและกระจายความผิดพลาดในการทดลองที่ไม่สามารถหลีกเลี่ยงได้ไปสู่ทุกๆการทดลองด้วยโอกาสและขนาดเท่าๆกัน

การควบคุม (Blocking) เป็นการควบคุมสภาพในการทดลองให้มีสภาพใกล้เคียงกันมากที่สุด ดังเช่น วัสดุที่ใช้ทดลองควรมีความเป็นอันหนึ่งอันเดียวกัน ใช้เครื่องจักรเดียวกัน ผู้ทดลองคนเดียวกัน ใช้วิธีทดลองวิธีเดียวกัน ช่วงเวลาทดลองใกล้เคียงกัน โดยเปลี่ยนแปลงเฉพาะเงื่อนไขของปัจจัยที่สนใจศึกษาเท่านั้น

ขั้นตอนแนะนำในการออกแบบการทดลอง

การใช้แนวคิดทางสถิติในการออกแบบและวิเคราะห์การทดลองนั้นผู้ทดลองจะต้องชัดเจนในความคิดว่าจะศึกษาอะไร จะเก็บรวบรวมข้อมูลอย่างไร และจะวิเคราะห์ข้อมูลอย่างไร

1. การเข้าใจและการกล่าวถึงปัญหา
2. การเลือกตัวแปรผลตอบ
3. เลือกปัจจัย ระดับ และช่วงที่จะทดลอง
4. เลือกวิธีการออกแบบการทดลอง
5. ทำการทดลอง
6. วิเคราะห์ข้อมูลทางสถิติ
7. สรุปผลการทดลองและให้ข้อเสนอแนะ

(Montgomery, 2005)

การออกแบบเชิงแฟคทอเรียลแบบ 2^k

การออกแบบเชิงแฟคทอเรียลใช้ในการทดลองที่เกี่ยวกับหลายปัจจัย กรณีที่มีปัจจัย k ปัจจัย แต่ละปัจจัยประกอบด้วย 2 ระดับ คือ สูง และต่ำ ในหนึ่งเรพลิเคตที่บริบูรณ์จะประกอบด้วย

ข้อมูลทั้งสิ้น $2 \times 2 \times 2 \times 2 \times \dots \times 2 = 2^k$ ข้อมูล และลักษณะการออกแบบเชิงแฟคทอเรียลแบบ 2^k จะอยู่บนข้อสมมติที่ว่า 1) ปัจจัยทั้งหมดมีค่าตายตัว 2) การออกแบบเป็นเชิงสุ่มสมบูรณ์ 3) สมมติฐานเกี่ยวกับความเป็นปกติที่ยอมรับได้ การออกแบบ 2^k มีประโยชน์มากต่องานทดลอง ในช่วงเริ่มแรก เมื่อมีปัจจัยจำนวนมากที่เราต้องการที่จะตรวจสอบ เพราะจะทำให้เกิดการทดลอง จำนวนน้อยที่สุดที่สามารถทำได้เพื่อศึกษาผลของปัจจัยทั้ง k ชนิด

ซึ่งการวิเคราะห์จะใช้วิธีที่เรียกว่า การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) โดยใช้หลักการแบ่งแยกความแปรปรวน (ในรูปของค่าผลรวมกำลังสอง) ออกเป็นส่วนๆ ของแหล่งที่มาหรือสาเหตุของความแปรปรวนนั้น แล้วทดสอบโดยใช้ตัวสถิติ F เป็นตัวทดสอบ ความแตกต่างของแหล่งที่มาดังกล่าวว่ามีนัยสำคัญหรือไม่ โดยกำหนดสมมติฐานทดสอบได้แก่

H_0 : อิทธิพลจากทุกปัจจัยมีค่าเท่ากับศูนย์

H_1 : มีอย่างน้อยหนึ่งค่าที่ไม่เท่ากับศูนย์

ตารางที่ 2.1 การวิเคราะห์ความแปรปรวนของการออกแบบเชิงแฟคทอเรียลแบบ 2^4

แหล่งของความ ผันแปร (Source of variation)	องศาความเป็นอิสระ (Degree of freedom : df)	ค่าผลรวมกำลังสอง (Sum of squares : SS)	ค่าเฉลี่ยกำลังสอง (Mean squares : MS)	F_0
A	1	SS_A	MS_A	MS_A/MS_E
B	1	SS_B	MS_B	MS_B/MS_E
C	1	SS_C	MS_C	MS_C/MS_E
D	1	SS_D	MS_D	MS_D/MS_E
AB	1	SS_{AB}	MS_{AB}	MS_{AB}/MS_E
AC	1	SS_{AC}	MS_{AC}	MS_{AC}/MS_E
AD	1	SS_{AD}	MS_{AD}	MS_{AD}/MS_E
BC	1	SS_{BC}	MS_{BC}	MS_{BC}/MS_E
BD	1	SS_{BD}	MS_{BD}	MS_{BD}/MS_E
CD	1	SS_{CD}	MS_{CD}	MS_{CD}/MS_E
ABC	1	SS_{ABC}	MS_{ABC}	MS_{ABC}/MS_E
ABD	1	SS_{ABD}	MS_{ABD}	MS_{ABD}/MS_E
ACD	1	SS_{ACD}	MS_{ACD}	MS_{ACD}/MS_E
BCD	1	SS_{BCD}	MS_{BCD}	MS_{BCD}/MS_E
ABCD	1	SS_{ABCD}	MS_{ABCD}	MS_{ABCD}/MS_E
Error	$abcd(n-1)$	SS_E	MS_E	
Total	$nabcd-1$	SS_T	MS_T	

ในการวิเคราะห์ความแปรปรวนจะพิจารณา SST (Total Sum of Squares), $SStr$ (Treatment sum of squares) และ SSE (Error Sum of Squares) โดยที่

$SST = SStr + SSE$ ซึ่ง $SStr$ จะประกอบด้วยค่าผลรวมกำลังสองจากปัจจัยต่างๆ ดังในตารางที่ 2.1

การทดลองแบบแฟกทอเรียลนั้นมีวัตถุประสงค์เพื่อศึกษาอิทธิพลหลัก และอิทธิพลร่วมของปัจจัย จึงต้องทำการคำนวณเพื่อประมาณค่าอิทธิพลดังกล่าว โดยจะเริ่มจากการหาค่า contrast ก่อน แล้วจึงคำนวณค่าของผลรวมกำลังสอง และในการหาค่า contrast จะใช้วิธีการกำหนดในสูตรต่อไปนี้

$$Contrast_{AB...K} = (a \pm 1)(b \pm 1) \dots (k \pm 1)$$

ในกรณี 2^3 ต้องการหา AB contrast จะเขียนผลคูณเพื่อกระจายหา contrast ได้ดังนี้

$$\begin{aligned} \text{เช่น } AB \text{ contrast} &= (a-1)(b-1)(c+1) \\ &= abc + ab + c + (1) - ac - bc - a - b \end{aligned}$$

Contrast ที่ได้จะเกิดจากการกระจายสูตรผลคูณทางด้านขวามือ โดยค่า "1" จะหมายถึง (1)

ใน treatment combination และจะให้เครื่องหมายบวกเมื่อไม่มีปัจจัยอยู่ใน contrast ส่วนเครื่องหมายลบจะให้ เมื่อมีปัจจัยนั้นอยู่ใน contrast

จาก contrast ทั้งหมดสำหรับอิทธิพลต่างๆ ที่คำนวณได้แล้ว จึงนำมาคำนวณค่าอิทธิพลหลัก และอิทธิพลร่วม พร้อมทั้งค่าผลรวมกำลังสองของแต่ละอิทธิพล ได้ดังนี้

ค่าอิทธิพลของปัจจัย AB...K เท่ากับ

$$AB...K = \frac{1}{n \cdot 2^{k-1}} (\text{contrast}_{AB...K})$$

และค่าผลรวมกำลังสอง สำหรับอิทธิพลใดๆ สามารถคำนวณได้จากสูตร

$$SS_{AB...K} = \frac{1}{n \cdot 2^k} (\text{contrast}_{AB...K})^2$$

ดังนั้น

$$SS_A = \frac{(\text{contrast}_A)^2}{2 \cdot 2^4} = \frac{(a-1)(b+1)(c+1)(d+1)(e+1)}{32}$$

$$SS_B = \frac{(\text{contrast}_B)^2}{2 \cdot 2^4} = \frac{(a+1)(b-1)(c+1)(d+1)(e+1)}{32}$$

$$SS_C = \frac{(\text{contrast}_C)^2}{2 \cdot 2^4} = \frac{(a+1)(b+1)(c-1)(d+1)(e+1)}{32}$$

$$SS_D = \frac{(contrast_D)^2}{2.2^4} = \frac{(a+1)(b+1)(c+1)(d-1)(e+1)}{32}$$

$$SS_{AB} = \frac{(contrast_{AB})^2}{2.2^4} = \frac{(a-1)(b-1)(c+1)(d+1)(e+1)}{32}$$

$$SS_{AC} = \frac{(contrast_{AC})^2}{2.2^4} = \frac{(a-1)(b+1)(c-1)(d+1)(e+1)}{32}$$

$$SS_{AD} = \frac{(contrast_{AD})^2}{2.2^4} = \frac{(a-1)(b+1)(c+1)(d-1)(e+1)}{32}$$

$$SS_{BC} = \frac{(contrast_{BC})^2}{2.2^4} = \frac{(a+1)(b-1)(c-1)(d+1)(e+1)}{32}$$

$$SS_{BD} = \frac{(contrast_{BD})^2}{2.2^4} = \frac{(a+1)(b-1)(c+1)(d-1)(e+1)}{32}$$

$$SS_{CD} = \frac{(contrast_{CD})^2}{2.2^4} = \frac{(a+1)(b+1)(c-1)(d-1)(e+1)}{32}$$

$$SS_{ABC} = \frac{(contrast_{ABC})^2}{2.2^4} = \frac{(a-1)(b-1)(c-1)(d+1)(e+1)}{32}$$

$$SS_{ABD} = \frac{(contrast_{ABD})^2}{2.2^4} = \frac{(a-1)(b-1)(c+1)(d-1)(e+1)}{32}$$

$$SS_{ACD} = \frac{(contrast_{ACD})^2}{2.2^4} = \frac{(a-1)(b-1)(c-1)(d+1)(e+1)}{32}$$

$$SS_{BCD} = \frac{(contrast_{BCD})^2}{2.2^4} = \frac{(a+1)(b-1)(c-1)(d-1)(e+1)}{32}$$

$$SS_{ABCD} = \frac{(contrast_{ABCD})^2}{2.2^4} = \frac{(a-1)(b-1)(c-1)(d-1)(e+1)}{32}$$

$$SSE = SST + SS_{tr}$$

$$SST = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \sum_{l=1}^2 \sum_{m=1}^2 Y_{ijklm}^2 - \frac{Y^2}{n2^4}$$

$$SSE = SST - SS_A - SS_B - SS_C - SS_D - SS_{AB} - SS_{AC} - SS_{AD} - SS_{BC} - SS_{BD} - SS_{CD} - SS_{ABC} - SS_{ABD} - SS_{ACD} - SS_{BCD} - SS_{ABCD}$$

งานวิจัยนี้ได้มีการประยุกต์ใช้โครงข่ายประสาทเทียมร่วมกับเทคนิคการออกแบบการทดลอง เพื่อออกแบบโครงสร้างของโครงข่ายที่เหมาะสมที่สุดในการคำนวณค่าอุณหภูมิของวัสดุแม่เหล็กเฟอร์ไรต์ในระดับโมเลกุลซึ่งจะเกี่ยวข้องกับการศึกษาปัจจัยของการทดลอง เช่น จำนวนโหนดในแต่ละชั้น จำนวนชั้นซ่อน อัตราการเรียนรู้ และค่าโมเมนตัม เป็นต้น

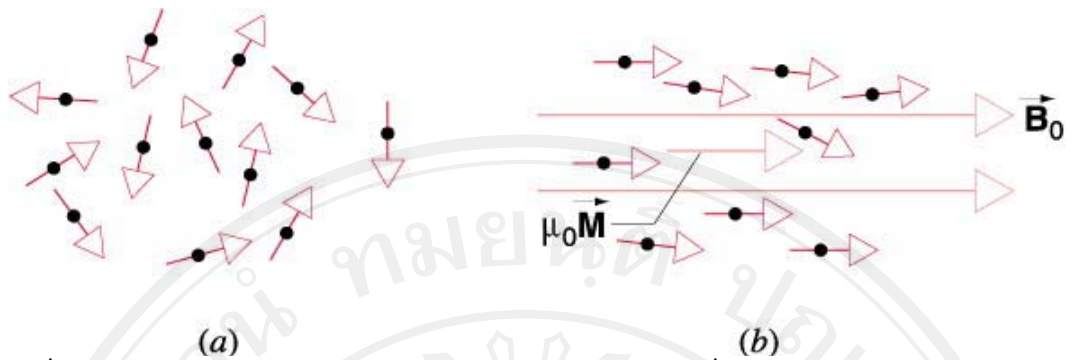
2.4 ทฤษฎีสัมบัติของแม่เหล็ก

ความเป็นแม่เหล็ก (Magnetism) เป็นปรากฏการณ์ที่วัสดุแม่เหล็กแสดงแรงดูดหรือผลักกับวัสดุแม่เหล็กอื่น สมบัติทางแม่เหล็กในระดับมหภาคเป็นผลสืบเนื่องจากโมเมนต์แม่เหล็ก ซึ่งเกิดจากอิเล็กตรอนแต่ละตัวในอะตอมจะมีโมเมนต์แม่เหล็กซึ่งจะเกิดจากแหล่งกำเนิด 2 แหล่ง แหล่งแรกมาจากการเคลื่อนที่แบบโคจรรอบนิวเคลียสของอิเล็กตรอนนั้น ซึ่งเทียบได้กับการเคลื่อนที่ของกระแสไฟฟ้าไหลวนเป็นวงเล็กๆ ทำให้เกิดสนามแม่เหล็กขนาดเล็กมากและมีโมเมนต์แม่เหล็กตลอดแนวแกนของการหมุนและอิเล็กตรอนแต่ละตัวยังเสมือนมีการหมุนรอบตัวเอง (Spinning) อีกด้วย ซึ่งจะทำให้เกิดโมเมนต์แม่เหล็กที่มีทิศทางตามแกนการหมุนรอบตัวเองในทิศ “ขึ้น” หรือ “ลง” ในแนวที่ขนานกันแต่มีทิศทางตรงกันข้ามเท่านั้น โมเมนต์แม่เหล็กที่เป็นพื้นฐานที่เล็กที่สุดเรียกว่า หน่วยโมเมนต์แม่เหล็กของบอร์ (Bohr Magneton), μ_B มีขนาดเท่ากับ $9.27 \times 10^{-24} \text{ A}\cdot\text{m}^2$ สำหรับอิเล็กตรอนแต่ละตัวมีค่าโมเมนต์แม่เหล็กจากการหมุนรอบตัวเองเป็น $\pm \mu_B$ (ค่าบวกสำหรับทิศขึ้นและค่าลบสำหรับทิศลง) ในแต่ละอะตอมโมเมนต์แม่เหล็กจากการโคจรของอิเล็กตรอนบางคู่จะหักล้างกันเองและกรณีโมเมนต์จากการหมุนรอบตัวเองก็เช่นกัน ดังนั้นค่าโมเมนต์แม่เหล็กสุทธิของอะตอมหนึ่งจะเท่ากับผลรวมของโมเมนต์แม่เหล็กของอิเล็กตรอนที่เกี่ยวข้องทั้งหมด

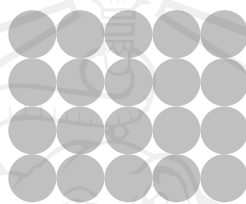
สภาพความเป็นแม่เหล็กแบ่งออกเป็นแบบไดอะ (Diamagnetism) พารา (Paramagnetic) และเฟอร์โร (Ferromagnetism) และวัสดุจะแสดงสภาพความเป็นแม่เหล็กอย่างใดอย่างหนึ่งในที่กล่าวมานี้ ขึ้นอยู่กับลักษณะการตอบสนองของขั้วแม่เหล็กจากอิเล็กตรอนและอะตอมของวัสดุนั้น ที่มีต่อสนามแม่เหล็กที่มากระทำจากภายนอก

2.4.1 สภาพแม่เหล็กแบบพารา (Paramagnetic)

ในวัสดุของแข็งบางชนิดอะตอมแต่ละอะตอมมีโมเมนต์ของไดโพลแบบถาวรซึ่งเป็นผลมาจากการที่โมเมนต์แม่เหล็กจากการโคจรและ/หรือหมุนรอบตัวเองของอิเล็กตรอนหักล้างกันไม่หมด ในสภาพที่ไม่มีสนามแม่เหล็กภายนอกการเรียงตัวของไดโพลของโมเมนต์แม่เหล็กระดับอะตอมเหล่านี้จะเป็นไปอย่างสุ่มทำให้โดยรวมแล้วชิ้นวัสดุนั้นไม่มีสภาพแม่เหล็กให้เห็นในระดับมหภาค แต่ไดโพลของอะตอมเหล่านี้สามารถหมุนได้อย่างอิสระทำให้เกิดสภาพแม่เหล็กแบบพาราขึ้น ดังรูปที่ 2.5 วัสดุที่มีสภาพแม่เหล็กแบบไดอะและแบบพาราต่างถือว่าเป็น “แม่เหล็กไม่ถาวร” เพราะจะแสดงสภาพแม่เหล็กก็ต่อเมื่อมีสนามแม่เหล็กภายนอกมากระทำเท่านั้น



รูปที่ 2.5 สภาพแม่เหล็ก Paramagnetism (a) การจัดเรียงตัวของอะตอมเมื่อไม่มีสนามแม่เหล็กภายนอกกระทำ
(b) อะตอมเมื่อได้รับอิทธิพลจากสนามแม่เหล็กภายนอก

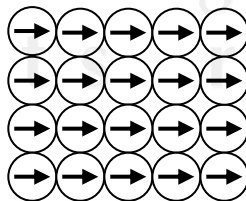


(c) สภาพแม่เหล็ก Diamagnetism ที่ไม่มี Magnetic Moment

(source : Pinpan , 2006)

2.4.2 สภาพแม่เหล็กแบบเฟอร์โร (Ferromagnetism)

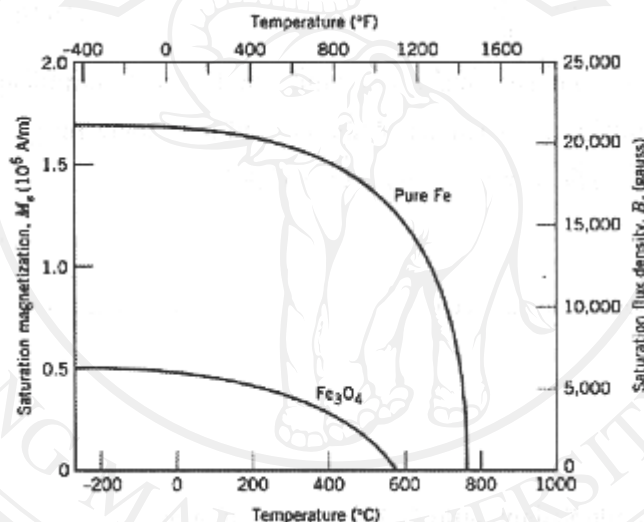
เป็นสภาพแม่เหล็กที่มีโมเมนต์แม่เหล็กแบบถาวรอยู่ในตัวถึงแม้ว่าจะไม่ได้อยู่ภายใต้สนามแม่เหล็กก็ตาม ทำให้สามารถแสดงสภาพแม่เหล็กที่ถาวรและแรงได้พบในพวกโลหะทรานซิชัน เช่น เหล็ก โคบอลต์ นิกเกิล และโลหะหายากเช่น กาลีโดเนียม โมเมนต์แม่เหล็กแบบถาวรในวัสดุที่มีสภาพแม่เหล็กแบบเฟอร์โรเป็นผลมาจากโมเมนต์แม่เหล็กที่เกิดการหมุนรอบตัวเองของอิเล็กตรอน ทั้งนี้มีการหักล้างกันของโมเมนต์ดังกล่าวไม่หมด ซึ่งจะมีผลต่อโมเมนต์โดยรวมเช่นกัน นอกจากนี้การเกิดปฏิสัมพันธ์แบบคู่ควบ (Coupling interaction) ทำให้โมเมนต์แม่เหล็กจากการหมุนรอบตัวเองของอะตอมที่อยู่ติดกันวางตัวในทิศที่สอดคล้องกันแม้จะไม่มีสนามแม่เหล็กภายนอกอยู่เลย ดังรูปที่ 2.6



รูปที่ 2.6 การเรียงตัวอย่างเป็นระเบียบของไดโพลในวัสดุแม่เหล็กเฟอร์โรที่
สถานะพื้นและปราศจากสนามแม่เหล็กภายนอก

2.4.3 อิทธิพลของอุณหภูมิที่มีผลต่อพฤติกรรมทางแม่เหล็ก

เมื่ออุณหภูมิสูงขึ้นแอมพลิจูดของการสั่นสะเทือนของอะตอมในของแข็งเพิ่มขึ้นทำให้ทิศทางของโมเมนต์แม่เหล็กที่อาจเรียงตัวดีแล้วเกิดการจลน์เรียงตัวผิดทิศทางไปจากเดิมและเนื่องจากโมเมนต์แม่เหล็กในอะตอมมีอิสระที่จะหมุนทิศทางได้ ดังนั้นเมื่ออุณหภูมิสูงขึ้น การสั่นของอะตอมเพิ่มมากขึ้นเนื่องจากความร้อน ผลก็คือทำให้ค่าสภาพแม่เหล็กของวัสดุลดต่ำลงและลดเหลือศูนย์อย่างรวดเร็ว ณ อุณหภูมิที่เรียกว่า อุณหภูมิคูรี, T_c หรืออุณหภูมิวิกฤต ซึ่งสภาวะดังกล่าวจะส่งผลกระทบต่อความสามารถในการทำงานที่ผิดปกติในชิ้นส่วนอุปกรณ์ต่างๆ เช่น มอเตอร์ เทปบันทึกเสียง ฮาร์ดดิสก์ไดรฟ์ ค่าของอุณหภูมิคูรีจะแตกต่างกันไปในวัสดุแต่ละชนิดเช่น ในเหล็ก โคบอลต์ นิกเกิล ค่าดังกล่าวเป็น 768 1120 335 °C ตามลำดับ



รูปที่ 2.7 สภาพแม่เหล็กอิ่มตัวเป็นฟังก์ชันของอุณหภูมิสำหรับเหล็กและ Fe_3O_4
(ที่มา: สุวันชัยและคณะ.2548)

2.4.4 การเก็บข้อมูลด้วยแม่เหล็ก

วัสดุแม่เหล็กได้ทวีความสำคัญขึ้นมามาก ในการเก็บบันทึกข้อมูลสารสนเทศต่างๆ กล่าวคือ เป็นการจัดเก็บข้อมูลในรูปแบบอิเล็กทรอนิกส์ทั้งในเทปบันทึกเสียง ฮาร์ดดิสก์ไดรฟ์ บัตรเครดิตและอื่นๆ อันที่จริงแล้วข้อมูลในหน่วยไบท์ของคอมพิวเตอร์เสียงหรือภาพที่อยู่ในรูปแบบของสัญญาณไฟฟ้าถูกบันทึกในส่วนเล็กๆ ของสื่อบันทึกด้วยแม่เหล็ก การถ่ายข้อมูลลงและอ่านข้อมูลจากเทปหรือดิสก์ทำโดยใช้หัวอ่าน/บันทึกแบบเหนี่ยวนำด้วยกระแสไฟฟ้าและทำให้เกิดสภาพแม่เหล็กขึ้นในบริเวณเล็กๆ ของส่วนที่อยู่ใกล้กับหัวอ่าน/บันทึกนั้น เมื่อปลดสนามแม่เหล็กออกแล้ว สภาพแม่เหล็กนี้ยังคงอยู่ กล่าวคือสัญญาณได้ถูกบันทึกไปแล้วนั่นเอง สื่อแม่เหล็กที่ใช้

บันทึกจะมีอยู่สองชนิดหลักได้แก่ แบบอนุภาค (Particulate) และแบบฟิล์มบาง (Thin film) สื่อแบบอนุภาคปกติจะเป็นเฟอร์ไรต์ หรือ CrO_2 แต่เทคโนโลยีการเก็บข้อมูลด้วยฟิล์มบางจะให้ความจุในการจัดเก็บข้อมูลสูงกว่าและต้นทุนต่ำลง โดยหลักแล้วจะใช้กับโครฟที่ใช้ดิสก์แข็งประกอบด้วยโครงสร้างแบบหลายชั้น (Multilayered structure) ในสื่อเหล่านี้ชั้นฟิล์มบางที่เป็นสารแม่เหล็กเท่านั้นนับเป็นส่วนประกอบในการบันทึกข้อมูลจริงๆ ฟิล์มนี้จะเป็นโลหะผสม CoPtCr หรือ CoCrTa และมีความหนา 10-50 nm ส่วนชั้นพื้นที่ข้างใต้ฟิล์มนี้เป็นโครเมียมบริสุทธิ์หรือโลหะผสมของโครเมียมสำหรับดิสก์เก็บข้อมูลด้วยแม่เหล็กซึ่งใช้ฟิล์มบางดังกล่าว ทิศทางเชิงผลึกวิทยา (Crystallographic direction) ของทิศที่เกิดการเหนี่ยวนำสภาพแม่เหล็กได้ง่าย ในแต่ละเกรนจะถูกจัดเรียงในทิศทางเดียวกับการเคลื่อนที่ของแผ่นดิสก์ หรือตรงข้ามกลไกของการบันทึกข้อมูลทางแม่เหล็กที่เกิดขึ้นในโดเมนเดี่ยวแต่ละโดเมนนี้เหมือนกับในแบบอนุภาค กล่าวคือ จะเกิดสภาพสนามแม่เหล็กได้เพียงสองแบบขึ้นอยู่กับทิศการเหนี่ยวนำสภาพแม่เหล็กในทิศทางหนึ่งและในทิศตรงข้าม ความหนาแน่นในการเก็บข้อมูลของฟิล์มบางสูงกว่าแบบอนุภาคเพราะประสิทธิภาพในการอัดตัว (Packing) ของโดเมนในฟิล์มบางสูงกว่านั่นเอง

2.5 การจำลองสถานการณ์มอนติคาร์โล

วิธีการจำลองสถานการณ์มอนติคาร์โล เป็นวิธีหนึ่งที่ได้รับการยอมรับเป็นอย่างดีว่าสามารถนำมาใช้จำลองหาค่าอุณหภูมิคูรีของสารแม่เหล็กเฟอร์ไรต์ที่มีความสัมพันธ์กับโครงสร้างหรือโครงสร้างผลึกของสารแม่เหล็กนั้นๆ ในการจำลองด้วยวิธีมอนติคาร์โลจะเริ่มจากการกำหนดระบบที่มีโครงสร้างผลึกดังที่ต้องการและวิธีการคำนวณจะขึ้นอยู่กับชนิดของอัลกอริทึม จากนั้นระบบจะทำการสุ่มทิศของสปินแม่เหล็กของแต่ละอะตอมในค่าอุณหภูมิที่กำหนดไปเรื่อยๆ โดยทิศทางที่สุ่มเลือกขึ้นมาจะถูกกำหนดโดยความน่าจะเป็นที่คำนวณได้จากพลังงานแม่เหล็กของระบบ ณ ที่อุณหภูมินั้น ตัวอย่างเช่น ถ้าพิจารณาอัลกอริทึมของเมโทรโพลิซิส (Metropolis *et al.*, 1953) ทิศทางสปินของระบบแม่เหล็กสปินหนึ่งจะถูกกำหนดใหม่โดยการสุ่ม จากนั้นคำนวณความแตกต่างของพลังงานที่เกิดขึ้นจากการสุ่มและค่าความน่าจะเป็นในการยอมรับการสุ่มทิศทางของสปินแม่เหล็กจากสมการที่ (2.10)

$$p = \exp(-\Delta E / k_B T) \quad (2.10)$$

เมื่อ p คือ ค่าความน่าจะเป็นในการยอมรับทิศสปินที่สุ่มใหม่

ΔE คือ ความแตกต่างของค่าพลังงานเนื่องจากการเปลี่ยนทิศของสปิน

k_B คือ ค่าคงที่ของโบลทซ์มาน

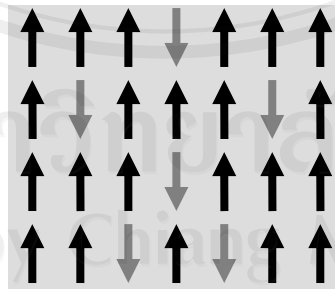
T คือ ค่าอุณหภูมิที่ทำการจำลองสถานการณ์

โดยในขณะที่ทำการจำลองสถานการณ์ค่าตัวเลขสุม r จะถูกสุ่มขึ้นมาในช่วง 0 ถึง 1 จากนั้นระบบจะพิจารณาว่าถ้าค่าสุม r นี้เท่ากับหรือน้อยกว่าค่าความน่าจะเป็น p ก็จะยอมรับทิศสปินของแม่เหล็กที่สุ่มขึ้นมา และขั้นตอนในการทำงานจะวนซ้ำจนกว่าจะสิ้นสุดการจำลองสถานการณ์

โดยทั่วไปก่อนเริ่มทำการจำลองสถานการณ์มอนติคาร์โล จะต้องเลือกแบบจำลองที่จะศึกษา ซึ่งในการศึกษานี้เลือกแบบจำลองไอซิงค์ (Ising Model)

2.5.1 แบบจำลองไอซิงค์

แบบจำลองไอซิงค์เป็นแบบจำลองซึ่งได้รับการยอมรับว่าเหมาะสมกับการทำนายพฤติกรรมแม่เหล็กฟิล์มบาง แบบจำลองไอซิงค์ถูกเสนอขึ้นครั้งแรกในปี ค.ศ. 1924 โดย Ernst Ising ให้สมมติว่าสปินของอนุภาคหรือขั้วแม่เหล็กจะชี้ได้ในทิศ $+z$ และ $-z$ เท่านั้น ซึ่งจะมีค่าเป็น 1 เมื่อสปินมีทิศชี้ขึ้นและมีค่าเป็น -1 เมื่อสปินมีทิศชี้ลง ดังรูป 2.8 จึงส่งผลให้แบบจำลองนี้เหมาะสมในการนำมาใช้จำลองสปินของแม่เหล็กฟิล์มบางในสื่อบันทึกข้อมูล เนื่องจากข้อมูลที่บันทึกนั้นจะอยู่ในลักษณะของเลขฐานสอง เมื่อพิจารณาพลังงานของระบบจะพบว่า ประกอบขึ้นจากสองส่วนคือ ส่วนที่สปินของอนุภาคมีอันตรกิริยากับสนามแม่เหล็กภายนอกกับส่วนที่สปินของอนุภาคมีอันตรกิริยากับสปินของอนุภาคข้างเคียง หรือเรียกว่า อันตรกิริยาแลกเปลี่ยนระหว่างสปิน (exchange interaction, J) และอันตรกิริยาแลกเปลี่ยนระหว่างสปินจะมีความแรงสุดระหว่างสปินที่อยู่ใกล้กันที่สุด เท่านั้น



รูปที่ 2.8 การจัดเรียงตัวของสปินในสองมิติแบบแลตทิซจตุรัส (square lattice)

ซึ่งสามารถเขียนแฮมิลโทเนียนของระบบได้ดังสมการ (2.11)

$$H = -\sum_{\langle i,j \rangle} J_{ij} s_i s_j - h(t) \sum_i s_i \quad (2.11)$$

เมื่อ

H คือ พลังงานของระบบ

J_{ij} คือ ค่าอันตรกิริยาแลกเปลี่ยน(พลังงาน)ระหว่างอะตอมสปิน

$S_{i(j)}$ คือ ค่าสปินของอะตอมที่ตำแหน่ง i ในโครงผลึกของสาร

$h(t)$ คือ สนามแม่เหล็กภายนอกที่มากระทำ

$\langle i, j \rangle$ แสดงถึงตำแหน่งของอะตอมคู่ที่อยู่ใกล้กันมากที่สุด (nearest-neighbor pairs) ซึ่งจากรูปก็คือ สปินที่ติดจากตำแหน่ง i ไปทางด้านบน ล่าง ซ้าย และขวานั้นเอง (พิจารณาเฉพาะ 2 มิติเท่านั้น) ทั้งนี้ในการคำนวณจะกำหนดให้ค่าสปินของ $S_{i(j)}$ มีค่าเท่ากับ ± 1 ที่ตำแหน่งของแลตทิซ $i(j)$ โดยขนาดของโมเมนต์แม่เหล็กจะคูณรวมกับ J

ในแบบจำลองนี้สปินที่อยู่ใกล้กันที่สุดจะมีเครื่องหมายของอันตรกิริยาแลกเปลี่ยน ซึ่งขึ้นกับการวางตัวของสปิน ถ้า $J > 0$ ค่าอันตรกิริยาแลกเปลี่ยนจะมีค่าเป็นบวกเมื่อสปินสวนทางกันและมีค่าเป็นลบเมื่อสปินวางตัวในทิศเดียวกัน โดยสปินของคู่แม่เหล็กที่อยู่ใกล้กัน และเหนี่ยวนำกันได้เมื่อชี้ในทิศเดียวกันจะมีผลลดพลังงานภายในระบบให้ต่ำลง ดังนั้นสำหรับระบบแม่เหล็กเฟอร์โรสถานะพลังงานต่ำสุดของระบบจะเกิดเมื่อทุกสปินวางตัวในทิศเดียวกัน และถ้า $J < 0$ จะเป็นอันตรกิริยาสำหรับสถานะแม่เหล็กแบบแอนติเฟอร์โรและสถานะพลังงานต่ำสุดของระบบจะเกิดเมื่อทุกสปินวางตัวสวนทางกัน ในการศึกษานี้ใช้หน่วยของพลังงานและอุณหภูมิคือ J และ J/k_B

สำหรับการคำนวณในหน่วยของเคลวิน (K) สามารถคำนวณโดยใช้ค่า J ของสารแต่ละชนิดดังเช่น ถ้าค่า J ของโลหะชนิดหนึ่ง คือ Y Joules จะสามารถคำนวณหาค่าอุณหภูมิครีได้ คือ

$$T_c = 2.2692 \times \frac{Y \text{ Joules}}{1.38 \times 10^{-23} \text{ Joules/K}} = \frac{2.2692Y}{1.38 \times 10^{-23}} \text{ K}$$

โดยปกติแล้วค่าอุณหภูมิวิกฤตมาตรฐาน (Critical temperature) ที่ได้จากแบบจำลองสปินแบบไอซิงค์ 2 มิติ มีค่าเท่ากับ $T_{\text{crit}} = \frac{2}{\ln(1+\sqrt{2})} \approx 2.2692 \text{ J/k}_B$ (McCoy and Wu, 1973)

และวิธีมอนติคาร์โลจะทำการสุ่มกลับทิศของสปินแม่เหล็กเหล่านี้จนกระทั่งระบบมีพลังงานต่ำที่สุดและกวัดแกว่งรอบๆพลังงานนั้นในแต่ละอุณหภูมิที่พิจารณา(ระบบแม่เหล็กอยู่ในสถานะสมดุล) แล้วทำการวัดหาค่าของคุณสมบัติแม่เหล็กที่ต้องการเช่น สภาพแม่เหล็กหรือแมกเนไทเซชัน (magnetization, m) ซึ่งเป็นผลรวมของสปินแม่เหล็กต่อปริมาตร ค่าสภาพรับไว้ได้ทาง

แม่เหล็ก (magnetic susceptibility, χ) หรือ ค่าความผันผวนที่อันดับที่สี่ของสภาพแม่เหล็ก (4th order cumulant of magnetization, U_L) ซึ่งในทางทฤษฎีหากทำการจำลองสถานการณ์ที่อุณหภูมิคูรีนี้ ค่าความผันผวนที่อันดับที่สี่จะมีค่าเท่ากันเสมอไม่ว่าระบบจะมีขนาด L (หมายถึงจำนวนอะตอมที่มีการจัดเรียงตัวในลักษณะโครงผลึกแบบ 2 มิติซึ่งจะมีทิศทางไปตามแนวแกน x และแกน y) เท่าไรก็ตาม (Binder and Heermann, 1992)

ดังนั้นค่าอุณหภูมิคูรีจะสามารถหาได้จากจุดตัดค่าความผันผวนที่อันดับที่สี่ของสภาพแม่เหล็กจาก L ค่าต่างๆ โดย U_L นี้มีรูปดังสมการ (2.12)

$$U_L = 1 - \langle m^4 \rangle / 3 \langle m^2 \rangle^2 \quad (2.12)$$

เมื่อ m คือ ค่าสภาพแม่เหล็กที่คำนวณได้จากค่าสปินเฉลี่ย ดังสมการ (2.13)

$$m = \sum_i S_i / N \quad (2.13)$$

ส่วน $\langle \dots \rangle$ คือค่าเฉลี่ยคาดหวังระหว่างการจัดเรียงตัวอะตอมแม่เหล็กในทิศทางต่างๆและสามารถคำนวณได้ดังสมการ (2.14)

$$\langle m^2 \rangle = \sum_{t=1}^{t_{\max}} m_t^2 / t_{\max} \quad (2.14)$$

เมื่อ m_t คือ ค่าสภาพแม่เหล็ก ณ เวลา t
 $t_{\max}$ คือ จำนวนครั้งที่ทำการวัดค่าทั้งหมด

ซึ่งในการวัดแต่ละครั้งจะใช้เวลาในหนึ่งช่วงคือ 1 ขั้นตอนมอนติคาร์โลต่อจุดสปิน (Monte Carlo step per site, mcs) ซึ่งเทียบได้กับจำนวนครั้งการสุ่มทิศทางสปินเท่ากับจำนวนสปินทั้งหมดในระบบ คือ N ซึ่งระบบจะปรับเปลี่ยนสปินไปเรื่อยๆ และจะทำการวัดค่าทั้งหมดหลังจากที่ระบบเข้าสู่สภาวะสมดุลทางความร้อน

ณ อุณหภูมิคูรีค่าความผันผวนจะเป็นอิสระจากค่า L ดังนั้น สำหรับ L และ L' ที่มีขนาดแตกต่างกันมาก จะมีค่าของ $(U_L / U_{L'})_{T=T_c} = 1$ (Binder, 1981; Binder and Heermann, 1992) แต่เนื่องจากการมีขนาดระบบที่จำกัด ทำให้กราฟของค่าความผันผวนที่ซึ่งได้รับจากระบบขนาดต่างๆ อาจไม่ตัดกัน

ที่อุณหภูมิค่าใดค่าหนึ่ง ดังนั้นในการหาค่าอุณหภูมิวิกฤตจึงต้องทำการประมาณแบบเชิงเส้นจาก $T_c(b = L/L') \rightarrow 0$ (Biner, 1981; Binder and Heermann, 1992)

อย่างไรก็ตามในการคำนวณค่าอุณหภูมิคูรีด้วยวิธีนี้ จะต้องใช้ทรัพยากรคอมพิวเตอร์และเวลาสูงมาก เนื่องจากต้องทำการแปรค่าขนาดระบบ (L) หลายๆค่า และทำการแปรค่าอุณหภูมิไปเรื่อยๆจนกว่าจะพบค่าคิวมูแลนต์เท่ากัน นอกจากนั้น เนื่องจากค่าคิวมูแลนต์คำนวณจากค่าเฉลี่ยของสภาพแม่เหล็กกำลังสองและกำลังสี่ ดังนั้นหากจำนวนข้อมูลที่ใช้ในการหาค่าเฉลี่ยมีค่าไม่มาก จะทำให้ได้ค่าคิวมูแลนต์ที่มีความแม่นยำต่ำซึ่งจะส่งผลต่อการหาค่าอุณหภูมิคูรีได้ วิธีมอนติคาร์โล ยังคงมีข้อจำกัด ดังที่กล่าวมา

ดังนั้นในงานวิจัยนี้จึงสนใจที่จะใช้เทคนิคโครงข่ายประสาทเทียมมาช่วยคำนวณค่าจุดตัดของคิวมูแลนต์จากข้อมูลที่จำลองสถานการณ์ด้วยวิธีการมอนติคาร์โลโดยใช้อัลกอริทึมเมโทรโพลิติส ซึ่งเทคนิคดังกล่าวมีความแม่นยำสูงและลดการสิ้นเปลืองทรัพยากร ทั้งที่ข้อมูลที่ใช้ในการศึกษาได้มาจากระบบไอซิงค์แบบแลตทิซจตุรัสในสองมิติที่มีจำนวนอะตอมแม่เหล็กวางตัวที่จุดแลตทิซเป็นจำนวน $N = L \times L$ เมื่อ L คือ จำนวนอะตอมซึ่งจัดเรียงตัวในลักษณะโครงผลึกแบบสองมิติซึ่งมีตำแหน่งไปตามทิศทางในแนวแกน X และแกน Y เป็นต้น

2.6 การวิเคราะห์การถดถอยและสหสัมพันธ์

การศึกษาความสัมพันธ์ระหว่างตัวแปรใช้สำหรับการหารูปแบบของความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ เพื่อใช้ในการทำนายหรือประมาณค่าตัวแปรตามที่สนใจศึกษา เมื่อกำหนดตัวแปรอิสระตัวอื่นๆมาให้ค่าของตัวแปรตัวหนึ่ง จะเปลี่ยนแปลงไปตามค่าของตัวแปรอีกตัวหนึ่ง เรียกว่า การวิเคราะห์การถดถอย อีกกรณีหนึ่งจะเป็นการศึกษาโครงสร้างความสัมพันธ์ระหว่างตัวแปรว่ามีความสัมพันธ์กันในทิศทางใด และมีความสัมพันธ์กันมากน้อยเพียงใด โดยตัวแปรหนึ่งไม่ได้เป็นสาเหตุที่ทำให้เกิดการเปลี่ยนแปลงในตัวแปรอีกตัวแปรหนึ่ง ขณะเดียวกันค่าของตัวแปรเหล่านั้น ไม่ได้ถูกกำหนดไว้ล่วงหน้าแต่เป็นไปโดยสุ่ม การศึกษาความสัมพันธ์เหล่านั้น จะเรียกว่า การวิเคราะห์สหสัมพันธ์ (Correlation analysis)

การวิเคราะห์การถดถอย เป็นการศึกษาที่ข้อมูลประกอบด้วย ตัวแปรสองประเภทคือ ตัวแปรอิสระ (Independent variable) ซึ่งเป็นตัวแปรที่สามารถกำหนดค่าได้แน่นอน หรือสามารถวัดค่าได้ โดยไม่มีความคลาดเคลื่อนและตัวแปรตาม (Dependent variable) เป็นตัวแปรที่เกิดขึ้นโดยสุ่ม

วัตถุประสงค์หลักของการวิเคราะห์การถดถอยประกอบด้วย

- (1) ต้องการทราบว่าตัวแปรอิสระมีอิทธิพลต่อตัวแปรตามหรือไม่ กล่าวคือ ถ้าค่าของตัวแปรอิสระเปลี่ยนแปลงไป จะทำให้ค่าของตัวแปรตามเปลี่ยนแปลงไปหรือไม่
- (2) ต้องการสร้างฟังก์ชันความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามหรือสร้างสมการที่ใช้ในการทำนายค่าของตัวแปรตาม เมื่อทราบค่าตัวแปรอิสระที่เรียกว่าสมการถดถอย (Regression equation) หรือ สมการทำนาย

ในกรณีที่ต้องการศึกษาอิทธิพลของตัวแปรอิสระเพียงตัวเดียวที่มีต่อตัวแปรตาม จะเรียกว่า การวิเคราะห์การถดถอยแบบง่าย (Simple regression analysis) ในกรณีที่ต้องการศึกษาอิทธิพลของตัวแปรอิสระหลายตัวที่มีต่อตัวแปรตาม จะเรียกว่า การวิเคราะห์การถดถอยแบบหลายชั้น (Multiple regression analysis) ในขณะเดียวกันความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามอาจเป็นความสัมพันธ์เชิงเส้นตรงหรืออาจมีความสัมพันธ์ที่ไม่เป็นเชิงเส้นตรง (Nonlinear หรือ Curvilinear regression)

ดังนั้น เทคนิคการวิเคราะห์การถดถอยจึงมีทั้งการวิเคราะห์การถดถอยเชิงเส้นตรงแบบง่าย การวิเคราะห์การถดถอยเชิงเส้นตรงแบบหลายชั้น การวิเคราะห์การถดถอยไม่เป็นเชิงเส้นตรงแบบง่าย (Simple nonlinear regression) การวิเคราะห์การถดถอยไม่เป็นเชิงเส้นตรงแบบหลายชั้น (Multiple nonlinear regression)

ในงานวิจัยนี้จะใช้เฉพาะการวิเคราะห์การถดถอยเชิงเส้นตรงแบบง่าย ซึ่งมีรายละเอียดดังนี้

2.6.1 การวิเคราะห์การถดถอยเชิงเส้นตรงแบบง่าย

เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหรือ ตัวแปรทำนาย (Predictor variable) เพียงตัวเดียวกับตัวแปรตามโดยลักษณะความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามเป็นแบบเชิงเส้นตรง ในทางทฤษฎีถือว่า ตัวแปรอิสระเป็นสาเหตุของความเปลี่ยนแปลงของตัวแปรตาม หรือกล่าวว่าตัวแปรอิสระมีอิทธิพลต่อการเปลี่ยนแปลงของตัวแปรตาม

โดยทั่วไปนิยมให้ X เป็นตัวแปรอิสระและ Y เป็นตัวแปรตาม ดังนั้นการวิเคราะห์การถดถอยจึงเป็นการศึกษาอิทธิพลของตัวแปร X ที่มีต่อตัวแปร Y

การถดถอยเชิงเส้นตรงแบบง่ายจะมีแบบจำลองการถดถอย (Regression model) ดังสมการที่ (2.15)

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.15)$$

เมื่อ

Y_i = ค่าสังเกตตัวที่ i ของตัวแปรตาม Y

x_i = ค่าสังเกตตัวที่ i ของตัวแปรตาม x

β_0 = จุดที่เส้นถดถอยตัดแกน Y (Y - intercept)

β_1 = สัมประสิทธิ์การถดถอย (Regression coefficient) ซึ่งเป็นค่าที่บอกถึงปริมาณความเปลี่ยนแปลงของ Y เมื่อ x เปลี่ยนไป 1 หน่วย

ε_i = ความคลาดเคลื่อนสุ่ม (Random error)

ข้อสมมติ (Assumptions) ของการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอยเชิงเส้นตรงแบบง่ายมีข้อสมมติ ดังนี้

- 1) ค่าของตัวแปรอิสระ x เป็นค่าที่กำหนดได้แน่นอน หรือเป็นค่าที่วัดได้โดยไม่มี ความคลาดเคลื่อน
- 2) ที่แต่ละค่าของตัวแปร x จะมีประชากรของตัวแปร Y ที่มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับ $\mu_{Y.X} = \beta_0 + \beta_1 x$ และความแปรปรวนเท่ากัน เท่ากับ $\sigma_{Y.X}^2$
- 3) ความคลาดเคลื่อน ε_i เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ อิสระกันโดยมีค่าเฉลี่ย เท่ากับ 0 และความแปรปรวนเท่ากับ $\sigma_{Y.X}^2$

การวิเคราะห์การถดถอย จัดได้ว่าเป็นเทคนิคทางสถิติที่มีความแกร่ง (Robustness) ค่อนข้างสูง ดังนั้น ถ้าข้อมูลที่ใช้ศึกษา ไม่ได้มีการฝ่าฝืนข้อสมมติอย่างรุนแรง จะไม่ส่งผลกระทบต่อผลสรุปของการวิเคราะห์มากนัก

2.6.2 การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์ β_0 และ β_1 ในตัวแบบการถดถอย จะอาศัยวิธีกำลังสองน้อยที่สุด (Least squares method) ซึ่งมีวิธีการดังสมการที่ (2.16)

ถ้ากำหนดให้

$$E = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 \quad (2.16)$$

โดยการหาอนุพันธ์ของสมการข้างต้น เทียบกับ β_0 และ β_1 แล้วให้ผลลัพธ์เท่ากับ 0 จะได้สมการที่ (2.17) และ (2.18)

$$\frac{\partial E}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0 \quad (2.17)$$

$$\frac{\partial E}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0 \quad (2.18)$$

ถ้าให้ b_0 และ b_1 เป็นตัวประมาณ (Estimators) ของพารามิเตอร์ β_0 และ β_1 ตามลำดับ จะได้สมการที่เรียกว่า สมการปกติ (Normal equations) ดังสมการที่ (2.19) และสมการที่ (2.20)

$$nb_0 + b_1 \sum x_i = \sum y_i \quad (2.19)$$

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum y_i x_i \quad (2.20)$$

การหาค่าของ b_0 และ b_1 อาจใช้วิธีแก้สมการปกติ หรืออาจสร้างเป็นสูตรในการคำนวณได้ ดังสมการที่ (2.21) และสมการที่ (2.22)

$$\begin{aligned} b_1 &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\ &= \frac{\sum x_i y_i - (\sum x_i)(\sum y_i)/n}{\sum x_i^2 - (\sum x_i)^2/n} \end{aligned} \quad (2.21)$$

$$b_0 = \bar{y} - b_1 \bar{x} \quad (2.22)$$

จะได้สมการถดถอย (Regression equation) หรือ สมการทำนาย (Prediction equation) ดังสมการที่ (2.23)

$$\hat{y}_i = b_0 + b_1 x_i \quad (2.23)$$

ค่า $\hat{y}_i$ ที่แต่ละค่าของ x_i จะถูกเรียกว่า predicted value จากสมการถดถอยจะสามารถลากเส้นถดถอย (Regression line) ได้ โดยเส้นถดถอยจะมีคุณสมบัติดังนี้

- 1) เส้นถดถอยที่ได้จะผ่านจุดตัดของแกน $\bar{x}$ และ $\bar{y}$
- 2) ผลจากการใช้วิธีกำลังสองน้อยที่สุด จะทำให้ค่า $\sum (y_i - \hat{y}_i)^2$ มีค่าน้อยที่สุด

การแบ่งแยกความผันแปรในตัวแปร Y

ส่วนเบี่ยงเบนของค่าสังเกต Y_i ใดๆ จากค่าเฉลี่ย ($\bar{Y}$) จะสามารถแยกออกเป็น 2 ส่วนดังนี้

$$Y_i - \bar{Y} = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i)$$

ในที่นี้

$$\sum (Y_i - \bar{Y})^2 \quad \text{คือ ค่า Total SS}$$

$$\sum (\hat{Y}_i - \bar{Y})^2 \quad \text{คือ ค่า Regression SS (หรือ SS due to X)}$$

$$\sum (Y_i - \hat{Y}_i)^2 \text{ คือ ค่า Error หรือ Residual SS}$$

ดังนั้นความผันแปรใน Y สามารถแยกออกเป็นส่วนๆ ในเทอมของ SS ได้ดังนี้

$$\text{Total SS} = \text{Regression SS} + \text{Error SS}$$

ซึ่งในทางปฏิบัติสามารถคำนวณ ได้ดังนี้

$$\begin{aligned} \text{Total SS} &= \sum Y_i^2 - (\sum Y_i)^2 / n, df = n-1 \\ \text{Regression SS} &= b_1 [\sum X_i Y_i - (\sum X_i)(\sum Y_i)/n] = b_1^2 \sum (X_i - \bar{X})^2, df = 1 \\ \text{Error SS} &= \text{Total SS} - \text{Regression SS}, df = n-2 \end{aligned}$$

ตารางที่ 2.2 ความคลาดเคลื่อนมาตรฐาน

ค่าสถิติ	ความคลาดเคลื่อนมาตรฐาน
b_0	$S_{b_0} = \sqrt{S_{Y.X}^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right]}$
b_1	$S_{b_1} = \sqrt{\frac{S_{Y.X}^2}{\sum (X_i - \bar{X})^2}}$
$\hat{Y}_i$ ที่ $X=X_0$	$S_{\hat{Y}_i} = \sqrt{S_{Y.X}^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]}$
Y_i ที่ $X=X_0$	$S_{Y_i} = \sqrt{S_{Y.X}^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]}$

$$\text{โดย } S_{Y.X}^2 = \text{Error MS} = \frac{\text{Error SS}}{n-2} = \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2}$$

ซึ่งเป็นตัวประมาณของ $\sigma_{Y.X}^2$

2.6.3 การทดสอบสมมติฐาน

2.6.3.1 การทดสอบค่าพารามิเตอร์ β_0 กำหนดสมมติฐาน ดังนี้

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

สำหรับทดสอบว่าเส้นถดถอยผ่านจุดกำเนิดหรือไม่และใช้สถิติทดสอบคือ

$$t = \frac{b_0 - \beta_0}{S_{b_0}}, \text{ df} = n-2$$

2.6.3.2 การทดสอบค่าพารามิเตอร์ β_1 กำหนดสมมติฐาน ดังนี้

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

เป็นการทดสอบว่า $\beta_1 = 0$ หรือไม่ คือ การทดสอบว่า ตัวแปร X มีอิทธิพลต่อตัวแปร Y หรือไม่

(1) ใช้หลักการวิเคราะห์ความแปรปรวน (F test)

ตารางที่ 2.3 การวิเคราะห์ความแปรปรวน (F test)

source	df	SS	MS	F
Regression	1	$\sum (\hat{Y}_i - \bar{Y})^2$	$\sum (\hat{Y}_i - \bar{Y})^2 / 1$	$\frac{MSR}{MSE}$
Error	n-2	$\sum (Y_i - \hat{Y}_i)^2$	$\sum (Y_i - \hat{Y}_i)^2 / (n-2)$	
Total	n-1	$\sum (Y_i - \bar{Y})^2$		

(2) ใช้การทดสอบแบบ t

$$t = \frac{b_1 - \beta_1}{S_{b_1}}, \text{ df} = n-2$$

ซึ่ง $F = t^2$

สัมประสิทธิ์ตัวกำหนด (coefficient of determination) : R^2

สัมประสิทธิ์ตัวกำหนด หรือสัมประสิทธิ์การตัดสินใจ เป็นค่าสถิติที่บอกให้ทราบว่าความผันแปรทั้งหมดในตัวแปร Y ถูกกำหนดโดยตัวแปรอิสระ X มากน้อยเพียงใด

$$R^2 = \frac{\text{Regression}}{\text{Total SS}}$$

ค่า R^2 จะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้า R^2 มีค่ามาก เข้าใกล้ 1 แสดงว่าตัวแปรอิสระ X สามารถอธิบายความผันแปรในตัวแปร Y ได้สูง

และถ้านำ R^2 มาหาค่ารากที่สอง จะได้ค่าที่เรียกว่า สัมประสิทธิ์สหสัมพันธ์ (Coefficient of correlation): R เขียนอยู่ในรูปของ

$$R = \pm\sqrt{R^2}$$

$$\text{โดยที่ } -1 \leq R \leq 1$$

ค่า R ที่เป็นลบ หมายถึงตัวแปรอิสระ X และตัวแปรตาม Y มีความสัมพันธ์กันในทิศทางตรงกันข้ามส่วนค่าบวกหมายถึงตัวแปรทั้งสองมีความสัมพันธ์ในทิศทางเดียวกัน

2.6.4 สมการถดถอยรูปพหุนาม

สมการถดถอยรูปพหุนาม (Polynomial Form) เป็นการวิเคราะห์การถดถอยที่มีลักษณะของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามในเชิงเส้นโค้ง ซึ่งแบบจำลองของสมการจะอยู่ในรูปแบบของตัวแปรอิสระที่มีการยกกำลังตั้งแต่ 2 ไปเรื่อยๆ โดยมีรูปแบบของสมการทั่วไปแบบยกกำลัง k ดังสมการที่ (2.24)

$$Y = A + B_1X + B_2X^2 + B_3X^3 + \dots + B_kX^k + E \quad (2.24)$$

เมื่อ $A, B_1, B_2, B_3, \dots, B_k$ คือ ค่าพารามิเตอร์ที่ได้จากการประมาณค่าและ E คือ ค่าคลาดเคลื่อนมาตรฐาน

ถ้า k มีค่าเป็น 2 แบบจำลองของสมการพหุนามจะเป็นสมการแบบ Quadratic ซึ่งเมื่อนำมาเขียนกราฟจะได้กราฟในรูปของพาราโบลา (Parabola)

และสามารถหาค่าสัมประสิทธิ์การถดถอยได้ตามสูตรดังต่อไปนี้

$$A = \frac{n * (\sum y) * (\sum x^4) - (\sum x^2 * y) * (\sum x^2)}{n * (\sum x^4) - (\sum x^2)^2}$$

$$B_1 = (\sum xy) / \sum x^2$$

$$B_2 = \frac{n * (\sum x^2 y) - (\sum x^2) * (\sum y)}{n * (\sum x^4) - (\sum x^2)^2}$$

2.6.5 การตรวจสอบแบบจำลองการถดถอย

เมื่อใช้แบบจำลองการถดถอยในการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ในขั้นตอนแรกในการตัดสินใจว่าตัวแปรทั้งสองมีความสัมพันธ์กันในรูปแบบใดคือการเขียนแผนภาพการกระจายต่อจากนั้นจึงทำการประมาณค่าพารามิเตอร์ของแบบจำลอง ตามที่เราคาดว่าตัว

แปรมีความสัมพันธ์กันและทำการตรวจสอบความเหมาะสมของแบบจำลองว่าเป็นไปตามข้อสมมติต่างๆที่ กำหนดไว้ในทฤษฎีการถดถอยหรือไม่ เช่น ลักษณะการเป็นเชิงเส้น ค่าความคลาดเคลื่อนมีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวนคงที่เท่ากับ σ^2 เป็นต้น การที่แบบจำลองการถดถอยมีลักษณะต่างๆ ที่ไม่เป็นไปตามข้อสมมติจะส่งผลให้การนำผลการถดถอยไปใช้ในการพยากรณ์ลดความแม่นยำลงได้ ดังนั้นจึงต้องมีการตรวจสอบแบบจำลองดังกล่าวว่า เป็นไปตามข้อสมมติของลักษณะเชิงเส้นหรือไม่ ดังนี้

1. ตรวจสอบคุณสมบัติของค่าเรซิดวล
2. ตรวจสอบการแจกแจงของค่าเรซิดวลมาตรฐาน
3. ตรวจสอบค่าเรซิดวลโดยพล็อตระหว่างค่าเรซิดวลเทียบกับค่า X
4. ตรวจสอบค่าเรซิดวลโดยพล็อตระหว่างค่าเรซิดวลเทียบกับค่า Y จากสมการพยากรณ์

ค่าเรซิดวลและค่าเรซิดวลมาตรฐาน (Residual and Standardized Residuals)

ค่าเรซิดวล : e_i เป็นค่าที่แสดงความแตกต่างระหว่างค่าสังเกตกับค่าทำนาย
 ดังสมการที่ (2.25)

$$e_i = Y_i - \hat{Y}_i \quad (2.25)$$

ถ้าแบบจำลองการถดถอยที่เราพิจารณามีความเหมาะสมสำหรับข้อมูลที่เก็บรวบรวมมาได้ แล้วค่า e_i ควรจะมีคุณสมบัติเป็นไปตามข้อสมมติสำหรับค่า ε_i นั่นคือมีค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวนคงที่เท่ากับ σ^2 ซึ่งจากข้อตกลงดังกล่าวนี้เอง การวิเคราะห์เรซิดวลจึงมีความเหมาะสมอย่างมากในการนำมาใช้ตรวจสอบความเหมาะสมของแบบจำลอง

คุณสมบัติของค่าเรซิดวล

1. ค่าเฉลี่ยของค่าเรซิดวลในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายมีค่าเท่ากับ 0

$$\bar{e} = \frac{\sum_{i=1}^n e_i}{n} = 0$$

2. ค่าความแปรปรวนของ $e_i = \frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-2} = \frac{\sum_{i=1}^n e_i^2}{n-2}$ มีค่าเท่ากับ $\frac{SSE}{n-2}$ ซึ่งเป็น

ค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อน (Mean Square Error): MSE นั่นเอง

จากที่กล่าวมาเกี่ยวกับค่าเรซิดวลและคุณสมบัติต่างๆ เราจึงสามารถนำค่าเรซิดวลมาใช้ในการตรวจสอบความเหมาะสมของแบบจำลองการถดถอยเชิงเส้นอย่างง่าย ในหลายลักษณะดัง

เช่น ตรวจสอบความเป็นเชิงเส้นของแบบจำลอง ค่าความแปรปรวนของความคลาดเคลื่อน ข้อมูลที่มีความผิดปกติ การแจกแจงของความคลาดเคลื่อน เป็นต้น

2.6.6 การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างกับค่ามาตรฐาน

ในการเปรียบเทียบค่าเฉลี่ยของประชากรกับค่ามาตรฐาน ในกรณีที่ทราบความแปรปรวน (σ^2) และจำนวนตัวอย่างมีขนาดเล็ก จะใช้สถิติทดสอบ t โดยสมมติว่าข้อมูลมีการแจกแจงแบบปกติหรือเบี่ยงเบนไม่มากนัก ซึ่งใช้ค่า S^2 เพื่อประมาณ σ^2

สมมติฐาน

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

สถิติทดสอบ

$$t_0 = \frac{\bar{y} - \mu_0}{S / \sqrt{n}}$$

ปฏิเสธสมมติฐานหลักเมื่อ $|t_0| > t_{\alpha/2, n-1}$ หรือมีค่าวิกฤตน้อยกว่าระดับนัยสำคัญ (P value < α)

2.7 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง

จากการสำรวจงานวิจัยที่เกี่ยวข้องเพื่อใช้เป็นแนวทางในการทำวิจัยพบว่า ได้มีการนำโครงข่ายประสาทเทียมไปประยุกต์ใช้กับการศึกษาด้านคุณสมบัติทางไฟฟ้าของวัสดุแม่เหล็ก เช่น Yamazaki *et al.* (1996) ได้ประยุกต์ใช้โครงข่ายประสาทเทียมในระบบการออกแบบรูปทรงแม่เหล็กถาวรและการคัดเลือกวัสดุแม่เหล็กที่ได้คุณภาพโดยใช้คอมพิวเตอร์ช่วยซึ่งจะอยู่บนพื้นฐานของการคำนวณทางด้านทฤษฎีแม่เหล็กไฟฟ้า ได้แก่ การคำนวณความสัมพันธ์ระหว่างค่าสภาพแม่เหล็ก (Magnetization, M) และความหนาแน่นของฟลักซ์แม่เหล็ก (Magnetic flux density, B) ผลที่ได้พบว่าวิธีการโครงข่ายประสาทเทียมค่อนข้างจะมีประสิทธิภาพ ต่อมา Rudnick *et al.* (1998) ได้ศึกษาแบบจำลองโครงข่ายประสาทเทียมในการจำลองค่ารูปแบบพัลส์ของกระแสไฟฟ้าและอุณหภูมิที่เพิ่มขึ้นเนื่องจากพลังงานที่สูญเสียในรูปความร้อนในการทำใหวัสดุเป็นแม่เหล็ก ซึ่งสิ่งที่ได้สามารถนำไปใช้ควบคุมพลวัตของแม่เหล็กช่วยประหยัดเวลาในการคำนวณหาค่ารูปแบบพัลส์กระแสที่ดีที่สุดในแต่ละอุณหภูมิที่ทำให้สภาพแม่เหล็กในวัสดุแม่เหล็กถึงจุดอิ่มตัวแบบมีประสิทธิภาพสูงสุด และ Seguin *et al.* (1999) ได้รายงานวิธีการใช้โครงข่ายประสาทเทียมแบบป้อนสัญญาณไปข้างหน้าในการศึกษาสมรรถนะที่ดีที่สุดของระบบอุปกรณ์แม่เหล็กไฟฟ้าและสร้างแบบจำลอง

จากวงจรแม่เหล็กโดยหาค่าพารามิเตอร์ที่มีคุณภาพตามข้อกำหนดซึ่งมีปัจจัยจำกัดที่ต้องพิจารณา คือ การใช้พลังงานและต้นทุน

และยังมีการนำโครงข่ายประสาทเทียมไปประยุกต์ใช้กับการศึกษาทางด้านวงฮิสเทอรีซิส เช่น Kuczmann and Iványi (2002) ได้ศึกษาการจำลองแบบทางพฤติกรรมของวัสดุแม่เหล็กโดยใช้โครงข่ายประสาทเทียมจำลองลักษณะของวงฮิสเทอรีซิสที่เกี่ยวข้องกับขนาด (Scalar hysteresis) และคุณสมบัติในด้านต่างๆของวัสดุแม่เหล็ก พบว่าคุณสมบัติหลากหลายของวัสดุแม่เหล็กสามารถจำลองได้โดยใช้อัลกอริทึมอย่างง่ายและค่าผลลัพธ์ที่ได้จากการทำนายลักษณะของวงฮิสเทอรีซิสจะนำไปเปรียบเทียบกับผลที่ได้จากแบบจำลอง Preisach เป็นต้น ต่อมา Kuczmann and Iványi (2003) ได้ใช้โครงข่ายประสาทเทียมแบบป้อนสัญญาณไปข้างหน้าประมาณค่าของความสัมพันธ์ของฟังก์ชันแบบไม่เป็นเชิงเส้น ที่ต่อเนื่องสร้างแบบจำลองเพื่อประมาณคุณลักษณะของทิศทางของวงฮิสเทอรีซิส (Vector hysteresis) ในการอธิบาย isotropic และ anisotropic ในวัสดุแม่เหล็กในรูปแบบของแบบจำลองสองมิติและสามมิติ และในปีเดียวกัน Makaveev *et al.* (2003) ได้สร้างแบบจำลองในการคำนวณทางพลวัตของวงฮิสเทอรีซิสในวัสดุแม่เหล็กเฟอร์ไรต์เป็นชั้นฟิล์ม SiFe โดยใช้โครงข่ายประสาทเทียมแบบป้อนสัญญาณไปข้างหน้าเพื่อนำไปประยุกต์ใช้ในมอเตอร์ไฟฟ้า และพบว่าแบบจำลองสามารถพยากรณ์ค่าความสัมพันธ์ของวงจรรสนามแม่เหล็กและฟลักซ์แม่เหล็ก (BH loop) ทั้งคลื่นวงโคจร (Arbitrary waveforms) และความถี่เหมือนกับในมอเตอร์ไฟฟ้า ในปี 2006 Kucuk (2006) ได้ใช้โครงข่ายประสาทเทียมและ genetic algorithm พยากรณ์ค่าทางพลวัตของวงฮิสเทอรีซิสของแกนแม่เหล็กที่ทำจาก 3% SiFe ที่ความหนาขนาด 0.27 มม. 0.1 มม. และความกว้างของแถบเกจ 0.08 มม. โดยวัดที่เหนือช่วงความถี่ 50–1000 เฮิร์ตซ์ โดยใช้ค่าอินพุตพารามิเตอร์ ได้แก่ ขนาดมิติของแกนลวด ค่าการเหนี่ยวนำของสนามแม่เหล็ก ความหนาของแถบโลหะและความถี่ของสนามแม่เหล็ก และเมื่อเปรียบเทียบการประมาณค่าวงฮิสเทอรีซิสที่ได้จากโครงข่ายประสาทเทียมและแบบจำลองทางพลวัต Preisach และ Energetic พบว่าการพยากรณ์ค่าวงฮิสเทอรีซิสของแกนลวดที่ได้จากการฝึกโครงข่ายประสาทเทียมแบบ genetic algorithm สามารถยอมรับได้ Kuczmann and Iványi (2006) ได้ศึกษาลักษณะเทคนิคการจำลองสถานการณ์บนพื้นฐานโครงข่ายประสาทเทียมเพื่อจำลองทิศทางของวงฮิสเทอรีซิสทั้งกรณี Isotropic และ Anisotropic ของวัสดุเฟอร์ไรต์แมกเนติก โดยเปรียบเทียบผลได้กับฟังก์ชันแบบ Everett ในแบบ 2D และ 3D เป็นต้น

นอกจากนี้ยังมีการนำโครงข่ายประสาทเทียมมาประยุกต์ใช้กับการออกแบบหาค่าพารามิเตอร์ในกระบวนการต่างๆเช่น Cherigui *et al.* (2005) ได้ใช้โครงข่ายประสาทเทียมออกแบบค่าพารามิเตอร์ในกระบวนการเคลือบแบบ high-velocity oxy-fuel (HVOF) ที่มีผลต่อคุณสมบัติทางแม่เหล็กของวัสดุ FeNb โดยในขั้นตอนแรกจะใช้โครงข่ายประสาทเทียมมาหาค่าความสัมพันธ์ของ

ตัวแปรระบบที่ดีที่สุดในระดับการเคลื่อนของโครงสร้างที่เป็นแบบจุลภาคและมีการประมาณค่าผลกระทบต่างๆที่มีต่อคุณสมบัติทางแม่เหล็กและได้ใช้โครงข่ายพยากรณ์ค่าคุณสมบัติของแม่เหล็กที่มีความสัมพันธ์กับโครงสร้างในระดับจุลภาคพบว่า มีความเป็นไปได้ที่จะบ่งชี้ถึงการปรับปรุงค่าสนามลบถ่วงแม่เหล็ก (coercivity) ในโครงสร้างระดับจุลภาค ค่าสภาพแม่เหล็กอิ่มตัว (Saturation magnetization) และค่าสภาพแม่เหล็กตกค้าง (Remanent magnetization) เป็นต้น ส่วน Hamzaoui *et al.* (2005) ได้ใช้วิธีการคำนวณของโครงข่ายประสาทเทียมพยากรณ์ค่าคุณสมบัติของแม่เหล็กของวัสดุผสม alloyed Fe-10%Ni and Fe-20%Ni nanocrystalline ซึ่งฝึกโครงข่ายประสาทเทียมในกระบวนการเพื่อหาค่าที่ดีที่สุด จากนั้นทำการทดสอบแล้วจึงทำนายค่าคุณสมบัติทางแม่เหล็กนอกช่วงของค่าพารามิเตอร์จากกระบวนการ พบว่าค่าผลลัพธ์ที่ได้จากการพยากรณ์และค่าจากการทดลองมีความสอดคล้องกัน โดยมีค่าส่วนเบี่ยงเบนมาตรฐานที่แตกต่างกันน้อยกว่า 6%

ส่วนการนำโครงข่ายมาประยุกต์ใช้กับงานด้านอื่นๆ เช่น การคำนวณเกี่ยวกับพลังงานแม่เหล็ก Kucuk (2006) ได้ใช้โครงข่ายประสาทเทียมแบบหลายชั้นซ่อนคำนวณการสูญเสียพลังงานเฉพาะในแกนแม่เหล็ก โดยฝึกโครงข่ายด้วยทฤษฎีโครงข่ายประสาทเทียมแบบแพร่กลับซึ่งผสมผสานกับอัลกอริทึมในการเรียนรู้แบบ delta-bar-delta และค่าผลลัพธ์ที่ได้จากวิธีโครงข่ายประสาทเทียมจะเปรียบเทียบกับวิธีการแบบดั้งเดิม พบว่าแบบจำลองจากโครงข่ายประสาทเทียมสามารถประมาณค่าการสูญเสียได้สอดคล้องกับวิธีการดั้งเดิม

จะเห็นว่าการนำโครงข่ายประสาทเทียมไปประยุกต์ใช้กับงานวิจัยที่เกี่ยวข้องกับคุณสมบัติทางแม่เหล็กในด้านต่างๆ ได้แก่ การสร้างแบบจำลองเพื่อหาค่าพารามิเตอร์ที่เหมาะสมสำหรับกระบวนการคำนวณหาความสัมพันธ์ระหว่างสนามแม่เหล็กและฟลักซ์ การพยากรณ์ทิศทางของวงฮิสเทอรีซิสและการศึกษาในด้านพลังงานของแกนแม่เหล็ก เป็นต้น

สำหรับงานวิจัยนี้จะขอนำเสนอการประยุกต์ใช้โครงข่ายประสาทเทียมในการคำนวณค่าอุณหภูมิคูรีซึ่งมีผลกระทบต่อคุณสมบัติในด้านสภาพแม่เหล็กของวัสดุแม่เหล็กเฟอร์โร ดังจะกล่าวในบทต่อไป