บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

2.1 การพยากรณ์ความต้องการไฟฟ้า

การพยากรณ์ความต้องการไฟฟ้า หมายถึง การหาค่าประมาณของพลังงานไฟฟ้าหรือพลังไฟฟ้าที่จะต้องจ่ายให้แก่ผู้ใช้ไฟฟ้า ในอนาคต

2.1.1 จุดมุ่งหมายของการพยากรณ์

การพยากรณ์ความต้องการไฟฟ้า สามารถแบ่งได้เป็น 3 อย่าง คือ การพยากรณ์ความต้องการพลังงานไฟฟ้า (energy forecast) การพยากรณ์ความต้องการไฟฟ้าสูงสุด (peak demand forecast) และการประมาณค่าตัวประกอบโหลด (load factor) ซึ่งจุดมุ่งหมายของการพยากรณ์ เป็นดังนี้

ก. การพยากรณ์ความต้องการพลังงานไฟฟ้า เพื่อวางแผนการจัดหาเชื้อเพลิง การวิเคราะห์ด้านการเงิน และการศึกษาอัตราค่าพลังงาน

ข. การพยากรณ์ความต้องการไฟฟ้าสูงสุด เพื่อการก่อสร้างโรงไฟฟ้าใหม่ การจัดตารางการบำรุงรักษา และการศึกษาอัตราค่าพลังงาน

ค. การประมาณค่าตัวประกอบโหลด เพื่อใช้ในการกำหนดชนิดของโรงไฟฟ้าที่เกี่ยวข้องกับการใช้ไฟฟ้า

2.1.2 การแบ่งประเภทโหลด

การแบ่งประเภทของโหลด เป็นไปตามกลุ่มของผู้ใช้ไฟฟ้า ที่พิจารณาจากลักษณะสมบัติของโหลด (load characteristics) ดังนี้

ก. กลุ่มที่พักอาศัย (residential) จะประกอบด้วย ครอบครัวเดี่ยว ครอบครัวขนาดใหญ่ อพาร์ตเมนท์ และอาคารพักอาศัย การใช้ไฟฟ้าในกลุ่มนี้ มีความสัมพันธ์กับปัจจัยทางเศรษฐกิจน้อยที่สุด

ข. กลุ่มธุรกิจและการค้า (commercial) ได้แก่ ศูนย์การค้า โรงพยาบาล อาคารสำนักงาน โรงเรียน และธุรกิจบริการต่างๆ การใช้ไฟฟ้าจะมีความเกี่ยวพันค่อนข้างมากกับการเปลี่ยนแปลงทางเศรษฐกิจ

ค. กลุ่มอุตสาหกรรม (industrial) ได้แก่ ภาคการผลิต เหมืองแร่ โรงงาน

อุตสาหกรรมต่างๆ การเปลี่ยนแปลงทางเศรษฐกิจจะมีผลกระทบต่อการใช้ไฟฟ้า เช่นกัน

สำหรับในประเทศไทยนั้น กลุ่มธุรกิจการค้า ยังแบ่งตามขนาด เป็น ธุรกิจขนาดเล็ก ขนาดกลาง ขนาดใหญ่ และธุรกิจพิเศษ เช่น โรงแรม นอกจากนี้ ยังมีผู้ใช้ไฟฟ้าที่เป็นหน่วยงานของรัฐบาล (government) และอื่นๆ เช่นไฟฟ้าแสงสว่างสาธารณะ (public lightings) อีกด้วย ซึ่งผู้ใช้ไฟแต่ละประเภทนี้ จะมีค่าตัวประกอบโหลดแตกต่างกัน คือ ธุรกิจพิเศษมีตัวประกอบโหลดสูงที่สุด (ประมาณ 90 %)ในขณะที่ อาคารพักอาศัย มีตัวประกอบโหลดต่ำที่สุด (ประมาณ 60 %) ซึ่งแสดงถึง ความสม่ำเสมอของการใช้ไฟฟ้าในคาบเวลาหนึ่งๆ

2.1.3 ปัจจัยที่มีผลต่อการพยากรณ์

ปัจจัยที่มีผลต่อความต้องการใช้ไฟฟ้า ที่ต้องนำมาพิจารณาเพื่อหารูปแบบการใช้ไฟฟ้าซึ่งมีผลต่อการพยากรณ์ความต้องการไฟฟ้าในอนาคต นั้น ได้แก่

ก. ปัจจัยทางเศรษฐกิจ (economics) คือ ค่าผลผลิตมวลรวมในประเทศ (gross domestic product) หรือ GDP อันเป็นผลมาจาก การผลิตสินค้าอุตสาหกรรม การบริการ การเกษตร การก่อสร้าง และการส่งออกและการนำเข้าสินค้า ปัจจัยนี้มีผลมากที่สุด (เกือบ 100 %) ต่อการพยากรณ์ความต้องการไฟฟ้าล่วงหน้าตั้งแต่ 3 ปี ขึ้นไป ในขณะที่มีผลน้อยลงสำหรับ การพยากรณ์ความต้องการไฟฟ้ารายปี (yearly) ในคาบไตรมาส (quarter) และรายเดือน (monthly) ตามลำดับ ส่วนการพยากรณ์ความต้องการไฟฟ้าในคาบเวลาตั้งแต่ หนึ่งสัปดาห์ลงมา ปัจจัยนี้จะไม่นำมาคิด

ข. ปัจจัยสภาพอากาศ (weather) คือ ค่าความชื้นสัมพัทธ์ อุณหภูมิ ฤดูกาล และความเร็วลม ปัจจัยเหล่านี้ จะมีผลกระทบมากที่สุด ต่อการพยากรณ์ความต้องการไฟฟ้าล่วงหน้า ในคาบเวลาไม่เกินหนึ่งสัปดาห์ และมีผลน้อยมาก ต่อการพยากรณ์ความต้องการไฟฟ้าล่วงหน้าตั้งแต่ 3 ปีขึ้นไป จึงมักไม่นำไปคิด

2.1.4 เทคนิคในการพยากรณ์

เทคนิคต่างๆ ที่ถูกนำมาใช้ในการสร้างแบบจำลอง เพื่อการพยากรณ์ความต้องการไฟฟ้าล่วงหน้า ได้แก่

ก. การสร้างกราฟประมาณค่า (extrapolation trend curve) ส่วนใหญ่ใช้ในการพยากรณ์ความต้องการไฟฟ้าระยะปานกลาง (medium-term) ซึ่งรูปแบบของกราฟที่ประมาณค่านี้ มีหลายแบบ เช่น กราฟเส้นตรง (straight line) กราฟพาราโบลา (parabola) กฎการยกกำลัง

(power law) กราฟเส้นโค้งเอกโปเนนเชียล(exponential) แบบปกติ แบบกอมเปิตส์ (Gompertz) แบบลอกจิส ติก (logistic) แบบไฮเปอร์โบลิก (hyperbolic) เป็นต้น

ข. วิธีบ๊อก-เจนกินส์ (Box-Jenkins) เหมาะสมกับการพยากรณ์ความต้องการไฟฟ้าระยะสั้น (short-term) และระยะปานกลาง เช่น แบบจำลอง ARMA (auto-regressive moving average model)

ค. การวิเคราะห์ถดถอย (regressive analysis) สามารถใช้ได้กับการพยากรณ์ความต้องการไฟฟ้าระยะสั้น และระยะปานกลาง ซึ่งจะนิยมเมื่อมีหลายตัวแปร (multivariable)

ง. เครือข่ายประสาท (neural network) ใช้หลักการเรียนรู้เลียนแบบระบบประสาทของมนุษย์ เหมาะสมกับ การพยากรณ์ความต้องการไฟฟ้าระยะสั้น

จ. แบบจำลองภาคผู้ใช้ไฟ (end-used model) ใช้กับการพยากรณ์ความต้องการไฟฟ้าระยะยาว (long-term) ซึ่งจะทำการพยากรณ์แต่ละภาคผู้ใช้ไฟ จากการพยากรณ์ของอัตราการเปลี่ยนแปลงและสถิติ (statistical and changing forecast)

ฉ. เทคนิค stock market ใช้สำหรับการพยากรณ์ความต้องการไฟฟ้าระยะสั้น ได้แก่ วิธี pattern recognition และ stochastic process

2.1.5 การพยากรณ์ความต้องการไฟฟ้าในประเทศไทย

การพยากรณ์ความต้องการไฟฟ้าในประเทศไทย นอกเหนือจากแต่ละหน่วยงานที่เกี่ยวข้องกับ การผลิต การส่ง และการจำหน่ายไฟฟ้า ดำเนินการเพื่อใช้เป็นข้อมูลของหน่วยงานเอง แล้ว ยังมีคณะอนุกรรมการ ประกอบด้วย สำนักงานคณะกรรมการนโยบายพลังงานแห่งชาติ สำนักงานคณะกรรมการพัฒนาสังคมและเศรษฐกิจแห่งชาติ สถาบันวิจัยและพัฒนาแห่งประเทศไทย การไฟฟ้าฝ่ายผลิตแห่งประเทศไทย (กฟผ.) การไฟฟ้านครหลวง (กฟน.) และการไฟฟ้าส่วนภูมิภาค (กฟภ.) เป็นคณะทำงานพยากรณ์ความต้องการไฟฟ้าของประเทศ เพื่อใช้ในการวางแผนพัฒนากำลังไฟฟ้า และแผนการพัฒนาแหล่งผลิตก๊าซธรรมชาติ ในระยะยาว วิธีการพยากรณ์จะแยกกระทำในแต่ละภาคผู้ใช้ไฟฟ้า คือ ที่อยู่อาศัย ธุรกิจ อุตสาหกรรม และอื่นๆ ซึ่งแต่ละภาคผู้ใช้ไฟฟ้า จะอาศัยเทคนิคการพยากรณ์แตกต่างกันไป ขึ้นกับคาบเวลาการพยากรณ์ที่ต้องการ และข้อมูลที่มี โดยเฉพาะอย่างยิ่ง ปัจจัยจาก GDP จากนั้นจึงนำมารวมกัน ค่าพยากรณ์จะมีทั้ง ความต้องการพลังงานไฟฟ้า ความต้องการไฟฟ้าสูงสุด และค่าตัวประกอบโหลด

สำหรับการพยากรณ์ความต้องการไฟฟ้าระยะปานกลาง และ ระยะสั้น จะมีการดำเนินการโดย กฟผ. กฟน. และ กฟภ. โดยค่าความต้องการไฟฟ้าระยะปานกลาง จะใช้ในการวางแผนจัดหาเชื้อเพลิงเพื่อการผลิตไฟฟ้า การวิเคราะห์ทางการเงิน การจัดตารางการซ่อมบำรุง และการจัด

ลำดับการทำงานร่วมกัน ในขณะที่ ค่าความต้องการไฟฟ้าระยะสั้น ใช้ในการวางแผนการผลิตไฟฟ้ารายวัน เช่น การทำยูนิตคอมมิตเมนต์ (unit commitment) และการประเมินค่าทางเศรษฐศาสตร์ของการจัดสรรการผลิตไฟฟ้า (economic dispatch) ด้วย วิธีการโดยทั่วไปที่ใช้ สำหรับการพยากรณ์ความต้องการไฟฟ้าระยะสั้น คือ เทคนิคเครือข่ายประสาท โดยไม่คิดปัจจัย GDP แต่จะใช้ข้อมูลอุณหภูมิ จากสถานีไฟฟ้าย่อย และองค์การของรัฐ จากนั้น ผลลัพธ์ที่ได้ จะถูกปรับแต่งอีกครั้ง ด้วยประสบการณ์ ของผู้ทำการพยากรณ์ และค่าพยากรณ์อากาศของหน่วยงานรัฐบาล

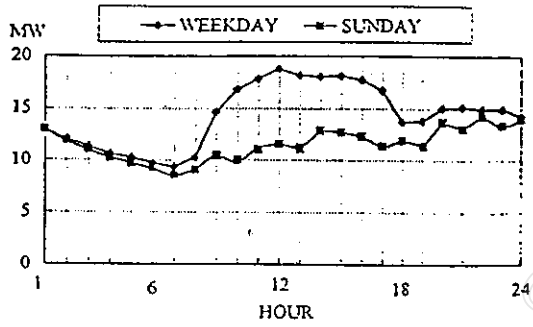
2.2 ความต้องการไฟฟ้าของสถานีไฟฟ้าย่อย

สถานีไฟฟ้าย่อยจำหน่าย เป็นสถานีจ่ายไฟฟ้าที่อยู่ใกล้กับกลุ่มของผู้ใช้ไฟฟ้ามากที่สุด โดยทั่วไปจะจำแนกตามประเภทของผู้ใช้ไฟฟ้าเป็น อุตสาหกรรม ที่อยู่อาศัย และอาคารพาณิชย์ โดยขึ้นกับเปอร์เซ็นต์ส่วนใหญ่ของความต้องการไฟฟ้าที่จ่ายจากสถานีไฟฟ้าย่อยนั้น ไปยังผู้ใช้ไฟแต่ละประเภท นอกจากนี้ รูปแบบของการใช้ไฟฟ้าแต่ละช่วงเวลายังแตกต่างกันไป กล่าวคือ ผู้ใช้ไฟประเภทอุตสาหกรรม มักจะมีการใช้ไฟฟ้าสูงสม่ำเสมอในเวลากลางวัน ของวันทำงาน ส่วนวันหยุดนั้น จะมีความต้องการไฟฟ้าต่ำตลอดเวลา เพราะไม่มีการผลิต สำหรับผู้ใช้ไฟประเภทที่อยู่อาศัยจะมีความต้องการไฟฟ้าสูงสุดในเวลาเย็นและค่ำ ยกเว้นวันหยุดอาจจะมีการใช้ไฟฟ้าในเวลากลางวันสูงกว่าปกติอันเนื่องมาจากการ ใช้เครื่องใช้ไฟฟ้าและเปิดเครื่องปรับอากาศในบ้านอยู่อาศัย ขณะที่ประเภทอาคารพาณิชย์ ย่านการค้านั้น จะใช้ไฟฟ้าสูงสุดช่วงเวลาบ่าย ถึงค่ำ โดยเฉพาะในวันหยุดจะสูงกว่าวันทำงาน อันเนื่องจากโหลดประเภทเครื่องปรับอากาศ ตัวอย่างรูปแบบของโหลดโดยทั่วไป ของผู้ใช้ไฟฟ้าแต่ละประเภท ในฤดูร้อน แสดงในรูปที่ 2.1

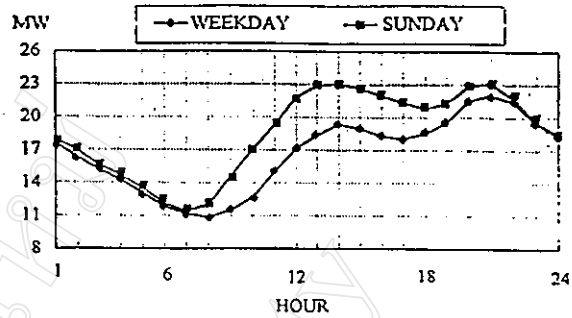
2.3 เทคนิคการพยากรณ์ความต้องการไฟฟ้าระยะสั้น

การพยากรณ์ความต้องการไฟฟ้าระยะสั้น ได้รับความสนใจ และศึกษามาเป็นระยะเวลาพอสมควร โดยมีการพัฒนาวิธีการมาเป็นลำดับ ซึ่งปัจจัยที่มีผลกระทบจะเกี่ยวข้องกับ วัน เวลา ลักษณะกลุ่มผู้ใช้ไฟฟ้า สภาพอากาศ หรือเหตุการณ์สำคัญต่างๆ ซึ่งปัจจัยเหล่านี้ สามารถคาดการณ์ได้ล่วงหน้า วิธีการต่างๆที่ใช้ ในการพยากรณ์ความต้องการไฟฟ้า สามารถแบ่งกว้างๆ ได้เป็นสองกลุ่ม คือ การใช้เทคนิคทางสถิติ (statistic-based techniques) และการใช้เทคนิคปัญญาประดิษฐ์ (artificial intelligence techniques)

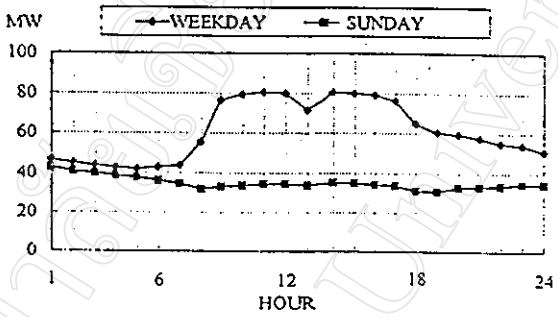
เทคนิคทางสถิติที่นำมาใช้ ได้แก่ วิธีกฎเกณฑ์การตัดสินใจ (deterministic) ซึ่งเป็นวิธีการเก่าแก่ที่สุด ใช้การสร้างกราฟด้วยการประมาณค่า (extrapolation) และวิธีกฎเกณฑ์ความน่าจะเป็น (probabilistic) ซึ่งพยายามจำลองโหลดของระบบ โดยอาศัยกระบวนวิธีสโตแคสติก



(ก) ที่อยู่อาศัย



(ข) ธุรกิจ พาณิชย์



(ค) อุตสาหกรรม

รูปที่ 2.1 ตัวอย่างรูปแบบของโหลดโดยทั่วไปของสถานีไฟฟ้าย่อยที่จ่ายไฟฟ้าแก่ผู้ใช้ไฟฟ้า 3 ประเภท ในฤดูร้อน [3]

(stochastic) วิธีการเหล่านี้ ล้วนแต่ใช้หลักการวิเคราะห์ความไวต่อสภาพอากาศ (weather sensitive analysis) [21] นอกจากนี้ยังมีวิธีการวิเคราะห์การถดถอยหลายตัวแปร (multi-variable regression) [11] วิธีสเตตสเปซ (state space methods) และวิธีอนุกรมเวลา (time series methods) [22, 2] วิธีการที่กล่าวถึงนี้ สามารถนำไปใช้พยากรณ์ความต้องการไฟฟ้าระยะสั้นได้ผลที่มีความแม่นยำแตกต่างกัน แม้ว่าจะบางวิธีการจะเป็นที่น่าพอใจ แต่ยังมีข้อจำกัด คือ

ก. ต้องใช้ข้อมูลในอดีตจำนวนมากในการหาค่าสหสัมพันธ์ (correlation) ซึ่งข้อมูลที่นำมาใช้อาจจะล้าสมัย

ข. ความสัมพันธ์ของตัวแปรอินพุต (input) และเอาต์พุต (output) มีความไม่เป็นเชิงเส้นสูง (nonlinearity) จึงยากที่จะหาได้อย่างถูกต้อง

ค. ปัญหาของการที่ตัวแปรอินพุตอยู่บนแนวเส้นตรงเดียวกัน (collinearity problem) ทำให้สามารถนำตัวแปรอินพุตไปใช้ในแบบจำลองได้จำนวนจำกัด

ง. แบบจำลองไม่มีความยืดหยุ่นต่อกรณีที่โหลดมีการเปลี่ยนแปลงอย่างฉับพลัน

เทคนิคปัญญาประดิษฐ์ ได้แก่ ฟัซซีลอจิก (fuzzy logic) เครือข่ายประสาท (neural

network) และระบบผู้เชี่ยวชาญ (expert systems) หรือการผสมเทคนิคเข้าด้วยกัน (hybrid) เช่น นิวโรฟัซซี (neuro-fuzzy) หรือเครือข่ายฟัซซีนิวรอล (fuzzy neural networks) [13] และการใช้หลักวิธีการถ่ายทอดพันธุกรรม (genetic algorithms) [7, 14] จะสามารถแก้ปัญหาข้อจำกัดดังกล่าวข้างต้นได้ ซึ่งการนำมาใช้เป็นเครื่องมือช่วยสร้างแบบจำลอง เพื่อพยากรณ์ความต้องการไฟฟ้าระยะสั้นนั้น ส่วนมากจะกระทำสำหรับระบบไฟฟ้าขนาดใหญ่ ซึ่งมีข้อจำกัดเกี่ยวกับการใช้ข้อมูล อินพุต ที่ไม่ละเอียด เช่น ใช้ข้อมูลสภาพอากาศของจังหวัดใดจังหวัดหนึ่ง แทนสภาพอากาศของภูมิภาคที่จังหวัดนั้นตั้งอยู่ หรือมีข้อมูลไม่ครบถ้วนสมบูรณ์ นอกจากนี้ เหตุการณ์สำคัญบางอย่างที่เกิดขึ้นในท้องถิ่น เช่น ที่สถานีไฟฟ้าย่อย ไม่ได้ถูกนำมาคิดด้วย เป็นต้น

โดยทั่วไปการพยากรณ์ความต้องการไฟฟ้าเป็นรายชั่วโมง ของสถานีไฟฟ้าย่อย จะใช้วิธีการทางสถิติ และความน่าจะเป็น ซึ่งหลังจากได้ผลลัพธ์แล้ว จะนำมาให้ผู้เชี่ยวชาญพิจารณาปรับปรุงอีกครั้งหนึ่ง แม้ว่าผลลัพธ์จะเป็นที่ยอมรับได้ แต่ยังขึ้นกับความรู้สึก และประสบการณ์ของบุคคล บางครั้งจึงมีความคลาดเคลื่อนมาก ทั้งนี้แนวคิดในการพยากรณ์ความต้องการไฟฟ้าของสถานีไฟฟ้าย่อย คือ การศึกษาความสัมพันธ์ ระหว่างโหลดและปัจจัยที่มีอิทธิพล เช่น วัน เวลา และข้อมูลสภาพอากาศของวันที่จะพยากรณ์ ซึ่งค่อนข้างจะซับซ้อน จึงเป็นการยากที่จะใช้เทคนิคธรรมดา มาหาความเกี่ยวพันระหว่างกัน การนำเทคนิคเครือข่ายประสาท และหลักวิธีการถ่ายทอดพันธุกรรม มาประยุกต์ใช้ ในการหาความเกี่ยวพันระหว่างกัน ของสภาพอากาศ และการเปลี่ยนแปลงของโหลด เพื่อพยากรณ์ความต้องการไฟฟ้าสูงสุด ของแต่ละวัน และความต้องการไฟฟ้ารายชั่วโมง มีข้อดีเหนือกว่าวิธีการทางสถิติและความน่าจะเป็น ในแง่ความสามารถในการจำลองปัญหาหลายตัวแปร โดยไม่ต้องตั้งสมมติฐานเชิงซ้อน ในกลุ่มของตัวแปรอินพุต และไม่ขึ้นกับความเชี่ยวชาญของบุคคล นอกจากนี้ ยังไม่ต้องคำนึงถึง ความสัมพันธ์ที่ไม่เป็นเชิงเส้นของกลุ่มตัวแปรอินพุต ด้วยการเรียนรู้ จากข้อมูลที่ทำการสอน และหาค่าแบบจำลองที่เหมาะสมที่สุด ด้วยวิธีการเลียนแบบการวิวัฒนาการของสิ่งมีชีวิต

ปัจจัยที่มีอิทธิพล ต่อลักษณะสมบัติของโหลด (load characteristic) ในประเทศไทย ได้แก่วันในสัปดาห์ และสภาพอากาศ ซึ่งจากการศึกษาใน [10] พบว่า อุณหภูมิเป็นตัวแปรสภาพอากาศที่สำคัญที่สุด เนื่องจากมีความเกี่ยวพันอย่างมากกับค่าของโหลด ความเร็วลมมีผลต่อโหลดบ้างเล็กน้อย ในขณะที่ปริมาณเมฆที่ปกคลุม ไม่มีความสัมพันธ์กับโหลดแต่อย่างใด นอกจากนี้ ในฤดูฝน ถ้าฝนตกชุก จะทำให้อุณหภูมิลดลงบ้าง แต่ค่อนข้างยากต่อการคาดคะเนล่วงหน้าว่า ฝนจะตกในเวลาใด และเป็นปริมาณมากน้อยเพียงใด ในทางปฏิบัติ ข้อมูลต่างๆเหล่านี้ จะมีจำกัด จึงมักจะไม่นำมาคิด ในบางกรณี มีการนำค่าความชื้นสัมพัทธ์มาพิจารณาด้วย เนื่องจากเกี่ยวข้องกับความ

รู้สึกสบาย (comfortability) ของมนุษย์ ซึ่งจะต้องใช้เครื่องปรับอากาศ หรือ เครื่องทำความร้อน มาปรับสภาพแวดล้อม

แบบจำลองที่จะใช้ในการพยากรณ์ความต้องการไฟฟ้าของแต่ละระบบ จะมีความแตกต่างกัน เพราะปัจจัยต่างๆ ไม่เหมือนกัน โดยทั่วไป จะทำการพยากรณ์ ความต้องการไฟฟ้า ล่วงหน้าหนึ่งวัน ถึง หนึ่งสัปดาห์ และอาจจะแยกแบบจำลองออกเป็นหลายๆ รูปแบบขึ้นกับ ลักษณะรูปแบบการเปลี่ยนแปลงของโหลด เช่น วันหยุดสุดสัปดาห์ (weekends) คือ วันอาทิตย์ และ/หรือวันเสาร์ วันทำงาน (weekdays) คือ วันจันทร์ ถึง วันศุกร์ หรือแยกวันจันทร์ ออกมาต่างหาก เป็นต้น นอกจากนี้ วันหยุดพิเศษอื่นๆ (public holidays) เช่น วันสงกรานต์ หรือวันหยุดชดเชย ทั้งนี้ จะต้องทำการวิเคราะห์ และพิจารณาตัดสินใจ ว่าจะสร้างแบบจำลองลักษณะใด

สำหรับแนวทางการพยากรณ์นั้น อาจจะกระทำได้หลายแบบ เช่น การพยากรณ์โหลดทีละชั่วโมง ต่อเนื่องกันแบบอนุกรม การพยากรณ์โหลดพร้อมกันทั้ง 24 ชั่วโมง หรือ การพยากรณ์ค่าโหลดสูงสุด และค่าโหลดต่ำสุด แล้วจึงมาเลือกรูปแบบของโหลดที่เหมาะสมกับวันที่พยากรณ์ ทำการหาค่าโหลดแต่ละชั่วโมง ก็ได้ ทั้งนี้ ขึ้นกับเทคนิค หรือ วิธีการที่เลือกใช้ และข้อมูลที่มี ว่าจะเหมาะสมกับแนวทางใด

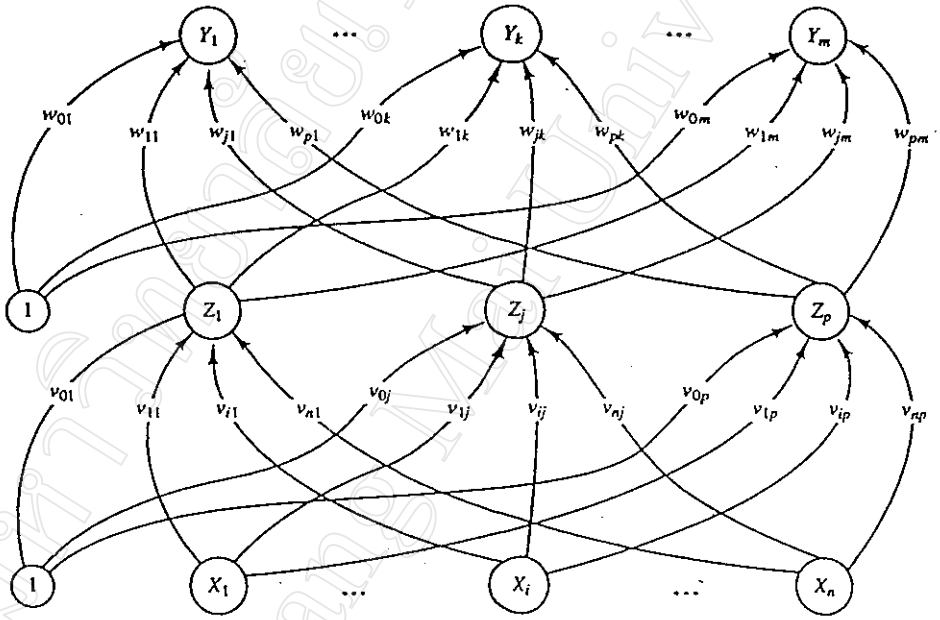
2.4 เทคนิคเครือข่ายประสาทแบบแพร่กลับ

เทคนิคเครือข่ายประสาท แบ่งเป็นสองหลักวิธี (algorithm) คือ หลักวิธีการสอนและเรียนรู้ (training and learning phase) กับ หลักวิธีการหาคำตอบ (recall phase) การสอนและการเรียนรู้ของเครือข่ายประสาท เป็นการคำนวณที่มีรูปแบบอยู่บนพื้นฐานตามประสบการณ์ ไม่ใช่เป็นความเข้าใจในเรื่องของการเชื่อมต่อระหว่างอินพุตและเอาต์พุต หรือไม่ใช่เป็นการพิจารณาจากเหตุและผล (cause and effect) ส่วนหลักวิธีการหาคำตอบ จะเป็นการใช้ในการหาคำตอบ เมื่อป้อนอินพุตที่ต้องการเข้าไป แล้วเครือข่ายประสาทให้เอาต์พุตออกมา

2.4.1 สถาปัตยกรรม

เครือข่ายประสาทแบบแพร่กลับ (neural network with backpropagation of error) หรือ BPN เป็นเครือข่ายประสาทชนิดหลายชั้น (multilayer) ภายในเครือข่ายประกอบด้วย ชั้นอินพุต (input layer) ชั้นซ่อน (hidden layer) และชั้นเอาต์พุต (output layer) โดยอาจจะมีมากกว่าหนึ่งชั้นซ่อนก็ได้ หน่วยประสาท หรือ นิวรอน (neuron) ในชั้นอินพุต ให้เป็น **X** ส่วนนิวรอนในชั้นเอาต์พุต ให้เป็น **Y** กรณีเครือข่ายประสาทแบบหนึ่งชั้นซ่อน จะกำหนดให้นิวรอนในชั้นซ่อนเป็น **Z** โดยชั้นอินพุตมีจำนวน n นิวรอน ชั้นซ่อนมี p นิวรอน และชั้นเอาต์พุตมี m นิวรอน

ค่าน้ำหนักที่เชื่อมระหว่างชั้นอินพุตกับชั้นซ่อน และชั้นซ่อนกับชั้นเอาต์พุต ให้เป็น v_{ij} และ w_{ij} ตามลำดับ ดังรูปที่ 2.2 นอกจากนี้ ยังมีค่าน้ำหนักเพิ่มเติมจากค่าน้ำหนักทั่วไป คือ ค่าไบอัส (bias) ของชั้นซ่อนและชั้นเอาต์พุต แทนด้วย v_{oj} และ w_{oj} ตามลำดับ ค่าไบอัสนี้ กำหนดให้สัญญาณอินพุตเป็นค่าคงที่เท่ากับหนึ่ง การสร้างโปรแกรมสำหรับทำงานเป็น BPN นี้ เป็นที่นิยมกันอย่างแพร่หลาย ในการประยุกต์ใช้งานด้านต่างๆ ลักษณะเด่นของ BPN คือ มีการเรียนรู้โดยนำเอาค่าผิดพลาดในการเรียนรู้แต่ละครั้งมาปรับค่าน้ำหนัก โดยทำการปรับน้ำหนักแบบแพร่ย้อนกลับ คือ ปรับค่าน้ำหนักระหว่างชั้นซ่อนกับชั้นเอาต์พุตก่อน แล้วจึงทำการปรับค่าน้ำหนักระหว่างชั้นอินพุตกับชั้นซ่อน



รูปที่ 2.2 สถาปัตยกรรมเครือข่ายประสาทแบบแพร่กลับชนิดหนึ่งชั้นซ่อน [6]

2.4.2 หลักวิธีการสอนและการเรียนรู้

กระบวนการสอน (training) เป็นแบบป้อนไปข้างหน้า (feedforward) กล่าวคือ แต่ละนิวรอนอินพุต $\mathbf{X}_i$ รับสัญญาณ (ข้อมูล) อินพุต x_i และส่งกระจายสัญญาณนี้ไปยังแต่ละนิวรอน $\mathbf{Z}_1, \mathbf{Z}_2, ..., \mathbf{Z}_j, ..., \mathbf{Z}_p$ ในชั้นซ่อน ซึ่งแต่ละนิวรอนในชั้นซ่อนจะคำนวณค่ากระตุ้น (activation) ของมัน แล้วส่งสัญญาณเอาต์พุต z_j ไปยังแต่ละนิวรอน $\mathbf{Y}_k$ ในชั้นเอาต์พุต จากนั้นนิวรอนเอาต์พุตจะคำนวณค่ากระตุ้นของมัน ได้สัญญาณเอาต์พุต y_k ออกมา

ในระหว่างช่วงเวลาการสอน แต่ละนิวรอนเอาต์พุตจะทำการเปรียบเทียบค่ากระตุ้นที่คำนวณได้ y_k กับค่าเป้าหมาย (target value) t_k ของมัน แล้วตรวจสอบค่าผิดพลาด

(error) ค่าผิดพลาดนี้จะถูกนำไปคำนวณ ตัวประกอบ δ_k ($k = 1,\ldots, m$) ซึ่งตัวประกอบนี้จะใช้ในการกระจายค่าผิดพลาดจากนิวรอนเอาต์พุต $\mathbf{Y}_k$ ย้อนกลับไปยังทุกนิวรอนในชั้นซ่อนที่อยู่ก่อนหน้าชั้นเอาต์พุตนี้ และยังจะใช้ในการปรับเปลี่ยนค่าน้ำหนักระหว่างชั้นซ่อนกับชั้นเอาต์พุตต่อไป อีกด้วย ในทำนองเดียวกัน จะมีการคำนวณค่าตัวประกอบ δ_j ($j = 1, \ldots, p$) ที่แต่ละนิวรอนของชั้นซ่อนด้วย ซึ่งตัวประกอบนี้จะใช้ในการปรับเปลี่ยนค่าน้ำหนักระหว่างชั้นอินพุตกับชั้นซ่อน

หลังจากที่ได้มีการหาค่าตัวประกอบ δ ครบหมดแล้ว เครือข่ายประสาทจะเรียนรู้(learning) ความผิดพลาดที่แพร่กลับมา แล้วทำการปรับค่าน้ำหนักของทุกๆ ชั้น พร้อมกันหมด ค่าน้ำหนักใหม่นี้ จะใช้ในการสอนและเรียนรู้สำหรับรูปแบบอินพุต (input pattern) ถัดไป ซึ่งกระบวนการสอนและเรียนรู้จะทำซ้ำเช่นเดิม จนกว่า เครือข่ายประสาทจะเรียนรู้ครบทุกรูปแบบอินพุตที่ป้อนให้ ค่าผิดพลาดระหว่างค่าเอาต์พุตกับค่าเป้าหมายของแต่ละนิวรอนเอาต์พุต ที่คำนวณได้หลังจาก การเรียนรู้หนึ่งครั้ง (epoch) นี้ จะถูกนำไปคำนวณหา ค่าเฉลี่ยของความผิดพลาดกำลังสอง (mean square error) หรือ MSE จากนั้น จะนำค่า MSE นี้ไปเปรียบเทียบกับค่าที่กำหนดไว้ หากว่า MSE ที่คำนวณได้ยังมากกว่า ค่าที่กำหนด เครือข่ายประสาทจะถูกสอนและเรียนรู้ซ้ำอีกครั้ง กระทำเช่นนี้เรื่อยไป จนกว่าจะได้ MSE สุดท้าย น้อยกว่า ค่าที่กำหนด จึงจะหยุดกระบวนการสอนและเรียนรู้ ค่าน้ำหนักที่ปรับเปลี่ยนได้ในครั้งสุดท้าย จะถูกนำไปใช้ในการหาคำตอบของปัญหาเมื่อนำเครือข่ายนี้ไปใช้ประโยชน์

2.4.3 ค่าน้ำหนักเริ่มต้น

ก่อนที่จะเริ่มทำการสอนให้แก่เครือข่ายประสาท จะต้องทำการกำหนดค่าน้ำหนักเริ่มต้นให้กับค่าน้ำหนักที่เชื่อมโยงระหว่างนิวรอนแต่ละชั้นให้ครบ โดยค่านี้จะเป็นเลขจำนวนจริงที่มีค่าน้อยๆ ที่ได้มาจากการสุ่มค่า (random number) ซึ่งจะมีผลต่อเวลาที่ใช้ในการสอน และอาจจะส่งผลถึงค่าน้ำหนักที่ได้ภายหลังจากการสอนและการเรียนรู้สิ้นสุดว่าจะมีศักยภาพมากน้อยเพียงใด ในการที่จะแปลงข้อมูลอินพุตไปเป็นเอาต์พุต การกำหนดค่าน้ำหนักเริ่มต้นโดยทั่วไปจะสุ่มค่าในช่วงระหว่าง –0.5 และ 0.5 หรือช่วงระหว่าง –1 และ 1 หรืออาจจะอยู่ในช่วงระหว่างค่าใดๆที่เหมาะสมตามคุณสมบัติการออกแบบเครือข่ายนั้นๆ ต่อมาในปี ค.ศ. 1990 งิวเย็นและวิดโรว์ (Nguyen and Widrow) ได้พัฒนาการสุ่มค่าน้ำหนักแบบธรรมดา เพื่อให้การเรียนรู้ของเครือข่ายเร็วขึ้น โดยวิธีการนี้มีพื้นฐานมาจากการวิเคราะห์เชิงเรขาคณิตของการตอบสนองต่อสัญญาณอินพุตที่เข้ามายังนิวรอนในชั้นซ่อน โดยใช้การแปลงฟูเรียร์ (Fourier Transform) ซึ่งจะกำหนดค่าน้ำหนักที่เชื่อมต่อระหว่างชั้นอินพุตและชั้นซ่อน ดังสมการที่ (2.1)

$$v_{ij}\ (\text{new}) = \beta\ v_{ij}\ (\text{old}) / \| v_{ij}(\text{old}) \| \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

โดยที่ β คือ สเกลแฟคเตอร์ ซึ่งมีค่า เท่ากับ $0.7\ (p)^{1/n}$

n คือจำนวนของอินพุต และ p คือ จำนวนของนิวรอนในชั้นซ่อน

v_{ij} (old) เป็นค่าน้ำหนักระหว่างชั้นอินพุตกับชั้นซ่อน ที่สุ่มระหว่าง –1 กับ 1 หรือ ระหว่าง $-\gamma$ กับ γ เมื่อ γ เป็นค่าคงที่ใดๆ

ส่วนการกำหนดค่าไบแอส จะสุ่มค่า ระหว่าง $-\beta$ กับ β

การกำหนดค่าน้ำหนักเริ่มต้นของงิวเย็นและวิดโรว์นี้ ได้รับความนิยมใช้งานอย่างแพร่หลาย โดยทดสอบบนเครือข่ายประสาทที่ใช้ฟังก์ชันการกระตุ้นแบบไบโพลาซิกมอยด์ ซึ่งจะกล่าวถึงต่อไป สำหรับการกำหนดค่าน้ำหนักเริ่มต้นแบบอื่น ก็มีผู้ได้นำเสนอไว้ในงานวิจัยต่างๆ แต่มุ่งเน้นที่จะใช้สำหรับเครือข่ายประสาทในงานวิจัยนั้น จึงไม่ขอกล่าวในที่นี้

2.4.4 ฟังก์ชันการกระตุ้น

ฟังก์ชันการกระตุ้น (activation function) สำหรับเครือข่ายประสาทแบบแพร่ความผิดพลาดย้อนกลับ จะต้องมีคุณลักษณะที่เป็นฟังก์ชันต่อเนื่อง สามารถหาค่าอนุพันธุ์ได้ และยังมีค่าเพิ่มขึ้นอย่างเดียว (monotonically non-decreasing) และควรจะหาค่าอนุพันธุ์ได้ง่าย โดยทั่วไปค่าอนุพันธุ์ที่ได้สามารถเขียนในรูปของค่าของฟังก์ชัน ส่วนมากจะนิยมใช้ฟังก์ชันแบบไบนารีซิกมอยด์ (binary sigmoid) ที่มีย่านค่าของฟังก์ชัน เป็น (0, 1) ซึ่งนิยามด้วยสมการที่ (2.2)

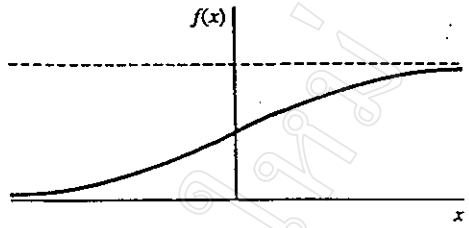
$$f_1(x) = 1 / (1 + \exp(-x)) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.2)$$

ดังแสดงให้เห็นในรูปที่ 2.3

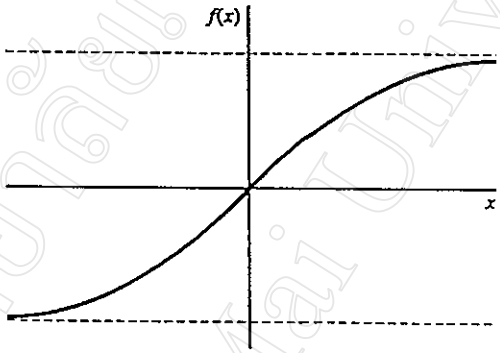
นอกจากนี้ ยังนิยมใช้เป็นฟังก์ชันแบบไบโพลาซิกมอยด์ (bipolar sigmoid) ที่มีย่านค่าของฟังก์ชันเป็น (-1, 1) และนิยามด้วยสมการที่ (2.3)

$$f_2(x) = [2 / (1 + \exp(-x))] - 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.3)$$

ดังแสดงให้เห็นรูปของฟังก์ชัน ในรูปที่ 2.4



รูปที่ 2.3 ฟังก์ชันไบนารีซิกมอยด์



รูปที่ 2.4 ฟังก์ชันไบโพลาซิกมอยด์

2.4.5 ค่าโมเมนตัมและอัตราการเรียนรู้

การเรียนรู้แบบแพร่กลับความผิดพลาด จะกำหนดอัตราการเรียนรู้ (learning rate) ที่มีค่าน้อยๆ โดยทั่วไปจะอยู่ระหว่าง 0.05 ถึง 0.25 เพื่อให้การปรับค่าน้ำหนัก มีความละเอียดมาก ทำให้การปรับค่าน้ำหนักภายในเครือข่ายต้องกระทำหลายครั้ง ส่งผลให้ต้องใช้เวลานานในการเรียนรู้ แต่ถ้ากำหนดให้อัตราการเรียนรู้มีค่ามากจะทำให้การปรับค่าน้ำหนักหยาบ ซึ่งการเรียนรู้อาจจะไม่เกิดการลู่เข้า (converge) สู่ค่าความผิดพลาดต่ำสุดที่กำหนด ในบางกรณีอาจจะให้อัตราการเรียนรู้มีการปรับค่าได้ เช่น กำหนดให้อัตราการเรียนรู้แปรผันตามจำนวนรอบที่สอน หรือตามค่าความผิดพลาดเฉลี่ยของระบบ เป็นต้น ก็จะสามารถช่วยเพิ่มอัตราเร็วในการเรียนรู้ได้ด้วย

นอกจากนี้ ยังอาจจะเลือกการใช้อัตราการเรียนรู้ที่มีค่าน้อย แต่เพิ่มพารามิเตอร์ที่เรียกว่า ค่าโมเมนตัม (momentum) เข้าไปในพจน์ (term) ที่ใช้ในการปรับค่าน้ำหนัก หรือ Weight Correction Term เพื่อเพิ่มความเร็วในการลู่เข้าหาคำตอบ ค่าโมเมนตัมที่ใช้ จะอยู่ในช่วง 0 กับ 1

2.4.6 จำนวนชั้นซ่อนและนิวรอนในชั้นซ่อน

ยังไม่มีวิธีการใดๆ ที่จะสามารถกำหนดจำนวนชั้นซ่อน และจำนวนนิวรอนในชั้นซ่อน ให้มีค่าแน่นอนได้ แต่ก็มีทฤษฎีที่กล่าวไว้ว่า ชั้นซ่อนเพียงชั้นเดียวก็เพียงพอแล้ว สำหรับการนำไปใช้กับปัญหาที่ไม่เป็นเชิงเส้นใดๆ แต่ในบางกรณีการเพิ่มชั้นซ่อนก็สามารถช่วยให้การเรียนรู้ง่ายขึ้น รวดเร็วขึ้น สำหรับจำนวนนิวรอนในชั้นซ่อนนั้นไม่สามารถกำหนดเป็นกฎเกณฑ์ตายตัว เนื่องจากจำนวนนิวรอนในชั้นซ่อนนี้ ขึ้นอยู่กับความซับซ้อนของปัญหา โดยมีการนำเสนอเทคนิคต่างๆ เพื่อใช้หาจำนวนนิวรอนที่เหมาะสมในชั้นซ่อน ซึ่งวิธีที่นับว่าแพร่หลายพอสมควร คือ ใช้จำนวนนิวรอนให้มีจำนวนมากไว้ก่อน แล้วค่อยๆ ลดจำนวนลงมา จนได้จำนวนที่น้อยที่สุดที่เครือข่ายประสาทจะยังคงเรียนรู้ได้ หรือ เรียกว่า เทคนิคการพรุนนิ่ง (pruning) นอกจากนี้ ยังอาจจะใช้วิธีการกำหนดจำนวนนิวรอนในชั้นซ่อนจาก ค่าเฉลี่ยทางเรขาคณิต (geometric mean) ของจำนวนนิวรอนในชั้นอินพุตกับชั้นเอาต์พุต ก็ได้

2.5 หลักวิธีถ่ายทอดพันธุกรรม

หลักวิธีการถ่ายทอดพันธุกรรม (genetic algorithm) หรือ GA นับเป็นหลักวิธีการค้นหา (search) แบบหนึ่ง ที่มีพื้นฐานเลียนแบบกลไกของการคัดเลือกพันธุ์ตามธรรมชาติของสิ่งมีชีวิต โดยที่พันธุ์ดีและเหมาะสมกว่า จะถูกคัดเลือกให้คงอยู่ หลักวิธี GA จึงมักถูกนำไปใช้ ในการแก้ปัญหาออปติไมเซชัน (optimization problem) เพื่อหาคำตอบที่ดีที่สุด ภายใต้เงื่อนไข (constraint) และฟังก์ชันของจุดประสงค์ (objective function) ที่กำหนด โดยเริ่มต้นจากการสุ่มค่าคำตอบขึ้นมาจำนวนหนึ่ง เรียกว่าเป็นประชากร (population) โดยที่แต่ละประชากร เรียกว่า โครโมโซม (chromosome) ในแต่ละโครโมโซมประกอบด้วยยีน (gene) หลายๆยีนเรียงต่อกัน โดยมากจะแทนแต่ละยีนด้วยเลขไบนารี (binary) จากนั้นจึงพิจารณาดูประชากรทั้งหมดว่ามีความเหมาะสมเพียงใด โดยการนำไปแทนในฟังก์ชันจุดประสงค์ ซึ่งผลลัพธ์แสดงถึงค่าความเหมาะสม (fitness value) ของประชากรนั้น เมื่อทราบค่าความเหมาะสมแล้วจึงทำการคัดเลือกพ่อพันธุ์และแม่พันธุ์ (parent selection) โดยอาจจะใช้วิธีการที่เรียกว่า วงล้อรูเล็ต (roulette-wheel method) ซึ่งจะเป็นการนำโครโมโซมรุ่นพ่อและรุ่นแม่ ถ่ายทอดไปยังรุ่นลูก(offspring) โดยตรง หรืออาจใช้วิธีการอื่นๆ ก็ได้

หลังจากที่ได้พ่อพันธุ์และแม่พันธุ์แล้ว สุ่มเลือกพ่อพันธุ์และแม่พันธุ์ มาผลิตรุ่นลูกเป็นประชากรรุ่นใหม่ ที่เกิดขึ้นมาด้วยวิธีการกลายพันธุ์หรือมิวเตชัน(mutation) หรือ การครอสโอเวอร์(crossover)

ก. การมิวเตชัน เป็นการเปลี่ยนค่าของยีนในโครโมโซมตามตำแหน่งที่สุ่มขึ้นมา โดยที่ถ้าค่าของยีน(แบบไบนารี)ในตำแหน่งที่สุ่มขึ้นมานั้น เป็น 1 จะถูกเปลี่ยนเป็น 0

ข. การครอสโอเวอร์ เป็นการสลับยีน หรือชุดของยีนระหว่างโครโมโซม ตามตำแหน่งที่สุ่มขึ้นมา จะได้โครโมโซมใหม่ที่เป็นรุ่นลูก มีทั้งยีนรุ่นพ่อและรุ่นแม่ผสมกัน

เมื่อผ่านกระบวนเหล่านี้แล้วจะได้ประชากรรุ่นใหม่ ซึ่งเกิดจากการนำโครโมโซมที่ดีไปแทนที่รุ่นพ่อรุ่นแม่ที่ด้อยกว่า ทำกระบวนการเช่นนี้ไปเรื่อยๆ จนถึงยุคสุดท้ายซึ่งโครโมโซมที่ได้จะเป็นคำตอบที่ดีที่สุดของปัญหานั่นเอง อย่างไรก็ตาม ถ้าหากการลู่เข้าหาคำตอบเกิดขึ้นรวดเร็วเกินไป (premature convergent) คำตอบที่ได้อาจจะเป็นความเหมาะสมเฉพาะที่ (local minima) ซึ่งหมายถึงว่ามีคำตอบอื่นที่ดีกว่า แต่กระบวนการค้นหาได้ข้ามคำตอบนั้นไป อาจจะเกิดขึ้นในขั้นตอนการสุ่มหาด้วยเทคนิคความน่าจะเป็น จึงจะต้องอาศัยเทคนิคอื่นๆเข้ามาช่วยในกระบวนการค้นหาคำตอบด้วย เช่น การคัดเอาโครโมโซมที่ดีที่สุดในแต่ละรุ่นมารวมกัน (elitist selection) หรือการปรับแต่งค่าความเหมาะสม (fitness value) ของทุกโครโมโซมเสียใหม่ เพื่อไม่ให้มีรุ่นลูกเกิดขึ้นซ้ำมากเกินไป หรือที่เรียกว่า การสเกลเชิงเส้น (linear scaling) เป็นต้น

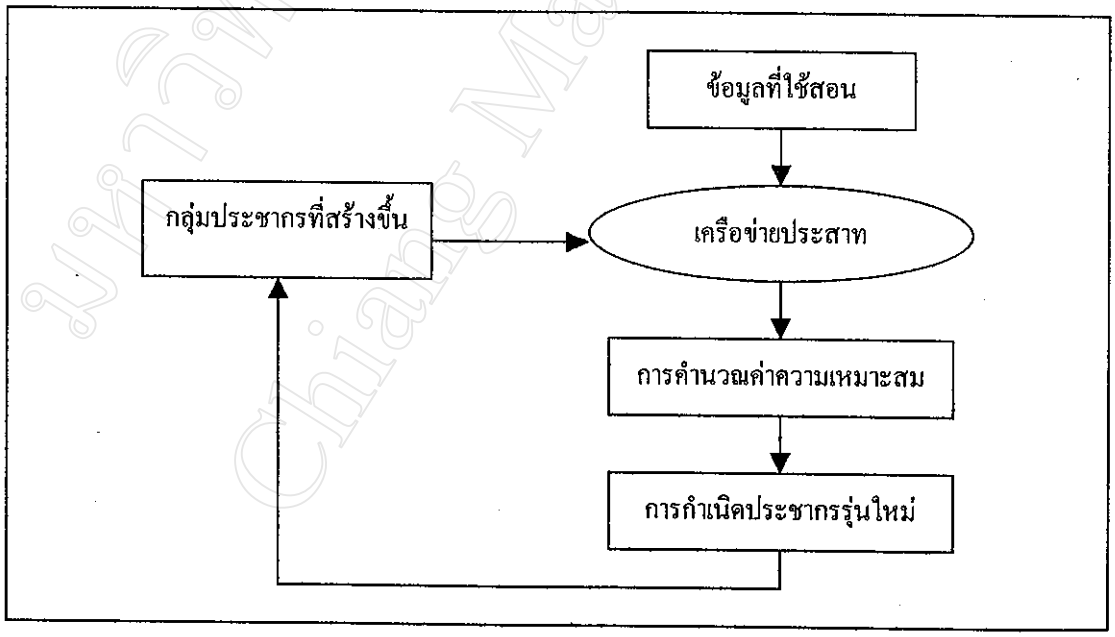
2.6 การใช้เทคนิคเครือข่ายประสาทร่วมกับหลักวิธีถ่ายทอดพันธุกรรม

เนื่องจากในขั้นตอนของการที่จะทำให้ความผิดพลาด ระหว่างค่าเอาต์พุตและค่าเป้าหมายที่ต้องการ (desired target) ของเครือข่ายประสาท มีค่าน้อยที่สุด โดยทั่วไปจะใช้ข้อมูลเกรเดียนท์ (gradient information) ของฟังก์ชันความผิดพลาด (error function) ซึ่งจะทำให้ได้ค่าความผิดพลาดน้อยที่สุดเฉพาะที่ (local minima) การนำ GA ซึ่งมีความสามารถในการค้นหาคำตอบที่ครอบคลุม (global search) มาใช้ร่วมกับเครือข่ายประสาท อาจจะกระทำด้วยวิธีการใด วิธีหนึ่ง เช่น การใช้หารูปแบบ (topology) สถาปัตยกรรมที่เหมาะสมของเครือข่ายประสาท หรือการใช้หาค่าน้ำหนักเริ่มต้น (initial weight) ของเครือข่ายประสาท เป็นต้น

ในกรณีของการใช้ GA มาช่วยในการหาค่าน้ำหนักเริ่มต้นของเครือข่ายประสาทนั้น จะทำให้ไม่จำเป็นต้องใช้ข้อมูลเกรเดียนท์ของความความผิดพลาดอีกต่อไป ซึ่งส่งผลให้การสอนและการเรียนรู้ของเครือข่ายประสาท หลุดพ้นจากค่าเหมาะสมที่สุด (optimize) เฉพาะที่ได้ หลักวิธีที่ใช้ ได้แก่ การกำหนดค่าน้ำหนักที่เชื่อมต่อระหว่างนิวรอนของชั้นในเครือข่ายประสาททั้งหมด ให้อยู่ในรูปแบบโครโมโซม โดยที่แต่ละโครโมโซมประกอบค่าน้ำหนักที่เรียงต่อกันไป เป็นอันดับ ซึ่งค่าน้ำหนักเสมือนเป็นยีนส์ในโครโมโซม การกำหนดรูปแบบของยีนส์ อาจจะใช้เข้ารหัส (encode) เป็นแบบสตริงไบนารี (binary string) แบบจำนวนจริง (real number coding) หรือแบบจำนวนเต็ม (integer coding) ก็ได้ ถ้าหากกำหนดเป็นสตริงไบนารี จะต้องมีการถอดรหัส (decode)

ส่วนถ้าหากว่า เป็นจำนวนจริงหรือจำนวนเต็ม ก็ไม่จำเป็นต้องถอดรหัสอีก จากนั้นจะนำค่าของโครโมโซมที่ถอดรหัสแล้วไปทำการหาค่าความเหมาะสมของมันในการที่จะอยู่รอด

สำหรับการนำ GA ใช้ร่วมกับเครือข่ายประสาทนั้น แสดงในรูปที่ 2.5 โดยจะทำการกำหนดโครโมโซมแบบสุ่มขึ้นมาจำนวนหนึ่ง แล้วนำไปเข้ากระบวนการสอนของเครือข่ายประสาท เอาต์พุตที่ได้จะถูกนำไปหาค่าฟังก์ชันความเหมาะสม ซึ่งจะเป็นดัชนีชี้ว่า โครโมโซมนั้นมีโอกาสอยู่รอดหรือไม่นอกจากนี้ บางโครโมโซมยังถูกสุ่มเลือกมาทำกระบวนการครอสโอเวอร์และ มิวเตชัน แล้วนำไปเข้ากระบวนการสอนของเครือข่ายประสาท ตรวจสอบความเหมาะสม ร่วมกันกับโครโมโซมเริ่มต้น คัดเลือกพันธุ์ ด้วยวิธีการใดวิธีการหนึ่ง จนได้ประชากรของโครโมโซมที่มีค่าความเหมาะสมตามต้องการ นำไปใช้ในกระบวนการ GA ในรุ่นถัดไป ทำเช่นนี้จนกว่า จะถึงจุดที่กำหนด เช่น กำหนดจำนวนรุ่นของวิวัฒนาการ หรือ กำหนดค่าขอบเขตของความเหมาะสม ในทางปฏิบัติจะใช้ฟังก์ชันค่าความเหมาะสม เป็นส่วนกลับของค่าความผิดพลาด หรืออาจจะใช้เป็นฟังก์ชันเอกโพเนนเชียล (exponential function) คือ $e^{\lambda P}$ เมื่อ λ คือ ค่าความคลาดเคลื่อนของเอาต์พุตของเครือข่าย และ P คือ ค่าจำนวนเต็มลบใดๆ [14]



รูปที่ 2.5 การนำหลักวิธีการถ่ายทอดพันธุกรรมมาใช้ร่วมกับเครือข่ายประสาท