



การพัฒนาขั้นตอนวิธีแบบผสมผสานสำหรับพยากรณ์ผลสัมฤทธิ์ทางการเรียน

Development of Hybrid Algorithm Model for Students' Achievement

Forecasting

สุวรรณณี รูปจิ้น* มณเฑียร รัตนศิริวงศ์วุฒิ¹ และ มนต์ชัย เทียนทอง²

¹ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
1518 ถนนประชาราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800

²ภาควิชาคอมพิวเตอร์ศึกษา คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
1518 ถนนประชาราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800

*E-mail: thoobjeen@gmail.com

บทคัดย่อ

งานวิจัยนี้นำเสนอการพัฒนาขั้นตอนวิธีแบบผสมผสานวิธีการพยากรณ์เชิงสาเหตุ (Hybrid system) ซึ่งประกอบไปด้วยขั้นตอนวิธีทางพันธุกรรมร่วมกับโครงข่ายประสาทเทียมและฟัซซีลอจิก สำหรับใช้ในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียน โดยใช้ข้อมูลผลสัมฤทธิ์ตั้งแต่ปี 2556 – 2558 เป็นจำนวน 1,320 คน แบ่งเป็น 660 คนสำหรับชุดข้อมูลเรียนรู้ (Training data) และ 660 คนสำหรับชุดข้อมูลทดสอบ (Testing data) จากผลการทดลองเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์พบว่าสามารถพยากรณ์ค่าที่มีความใกล้เคียงกับค่าเป็นจริงมากที่สุดทั้งชุดข้อมูลเรียนรู้ (Training data) และชุดข้อมูลทดสอบ (Testing data) และสร้างตัวแบบพยากรณ์กฎฟัซซีจำนวน 4 กฎ มีค่าผิดพลาด RMSE = 0.024 สำหรับชุดข้อมูลเรียนรู้ และมีค่าผิดพลาด RMSE = 0.087 สำหรับชุดข้อมูลทดสอบ

คำสำคัญ: ระบบผสม การพยากรณ์เชิงสาเหตุ ขั้นตอนวิธีทางพันธุกรรม ฟัซซีลอจิก โครงข่ายประสาทเทียม

Received: February 25, 2016

Revised: June 10, 2016

Accepted: June 24, 2016

Abstract

In this research, development of hybrid algorithm model by combination of causal forecasting methods using genetic algorithm and neural network and fuzzy logic is proposed for students' achievement forecasting. The students' learning achievement data set between the years 2013-2015 was used. The author selected 1,320 students' grade: 660 for training and 660 for testing. From the results, the developed model can predict the most possible values that compare to the real values of both training data and testing data. A fuzzy rules system with 4 rules was generated by hybrid algorithm and had RMSE = 0.024 in the training data. The resulted fuzzy rules system produced RMSE = 0.087 on the testing data.

Keywords: hybrid system, causal forecasting, genetic algorithm, fuzzy logic, neural network

1. บทนำ

การประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียนสามารถถูกพิจารณาได้ว่าเป็นปัญหาการจัดกลุ่มข้อมูล (Clustering problem) โดยที่จำนวนของกลุ่ม (Class) จะเกิดขึ้นบนพื้นฐานของกระบวนการวัดและประเมินผลทางด้านสติปัญญา ความรู้ความสามารถของผู้เรียน [1] ข้อมูลผลสัมฤทธิ์ส่วนใหญ่เกิดจากการวัดด้านปัจจัยเชิงพุทธิศาสตร์แล้วนำมาเปรียบเทียบกับเกณฑ์มาตรฐานที่ตั้งไว้ ทำให้สามารถแยกผู้เรียนออกเป็นกลุ่ม ได้แก่ A (Excellent), B+ (Very good), B (Good), C+ (Fair Good), C (Fair), D+ (Poor), D (Very poor) และ F (Fail) ปัญหาการแบ่งกลุ่มผู้เรียนออกเป็นกลุ่มตามระดับสติปัญญาความรู้ความสามารถของผู้เรียนที่มีลักษณะคล้ายคลึงกัน (homogenous grouping) ให้อยู่ในกลุ่มเดียวกันนั้นเป็นเรื่องที่กระทำได้ยากและเป็นปัญหาที่ผู้สอนต้องใช้วิจารณญาณความเหมาะสมในการตัดสินใจ โดยทั่วไปแล้วนั้นผู้สอนส่วนใหญ่มักจะทำการแบ่งกลุ่มผู้เรียนตามคะแนนรวมที่ได้มาจากการประเมินผลการเรียนหลายวิธีเข้าด้วยกัน ซึ่งคะแนนรวมที่ได้มานั้นอาจจะไม่สามารถบ่งชี้ถึงขนาดหรือปริมาณความรู้ความสามารถของ

ผู้เรียนได้อย่างแท้จริง ดังตัวอย่าง เช่น เมื่อทำการพิจารณาจากคะแนนสอบของผู้เรียนคนแรก 50 60 และ 70 กับคะแนนสอบของผู้เรียนคนที่สอง 70 60 และ 50 ซึ่งในกรณีนี้ถ้าใช้วิธีทางคณิตศาสตร์ที่นิยมใช้ในการแบ่งกลุ่มผู้เรียน โดยทั่วไปจะพบว่าผู้เรียนสองคนนี้มีระดับความรู้ความสามารถอยู่ในกลุ่มเดียวกัน แต่ในความเป็นจริงแล้วอาจจะไม่ได้เป็นเช่นนั้น เพราะเมื่อได้ลองพิจารณาจากสัดส่วนคะแนนสอบของผู้เรียนคนแรกนั้นจะพบว่ามีการพัฒนาขึ้น ซึ่งอาจเกิดจากการเปลี่ยนแปลงและประสบการณ์การเรียนรู้ของผู้เรียน แต่ในทางตรงกันข้ามผู้เรียนคนที่สองกลับมีสัดส่วนผลคะแนนสอบที่ต่ำลง จึงเป็นไปได้ยากที่จะจัดผู้เรียนตามระดับความรู้ความสามารถให้อยู่ในกลุ่มเดียวกัน ดังนั้นจึงต้องหาวิธีการประเมินผลสัมฤทธิ์ในแง่มุมใหม่ที่จะสามารถสะท้อนถึงความรู้ความสามารถของผู้เรียนแต่ละคนได้อย่างแท้จริง

งานวิจัยที่ [2-6] ได้นำเสนอการประยุกต์ใช้ระบบตรรกศาสตร์คลุมเครือ (Fuzzy logic) ที่มีคุณสมบัติช่วยในการตัดสินใจภายใต้ความคลุมเครือ ไม่ชัดเจน สามารถตีความและให้เหตุผลคล้ายกับตรรกะวิธีการคิดของมนุษย์มาใช้ใน

กระบวนการประเมินผลสัมฤทธิ์ของผู้เรียน โดยการ ใช้ฟังก์ชันความเป็นสมาชิกและกฎฟัซซีมาพิจารณา ถึงปัจจัยที่ส่งผลต่อผลสัมฤทธิ์อันเกิดจาก คุณลักษณะของคำถาม ได้แก่ ความยากของคำถาม (Difficulty) ความสำคัญของคำถาม (Importance) และความคลุมเครือของคำถาม (Vagueness) ในแต่ละ ข้อ ซึ่งวิธีการนี้มีประโยชน์ในแง่การประเมิน ผลสัมฤทธิ์จัดลำดับผู้เรียนที่มีผลคะแนนสอบซ้ำกัน งานวิจัยที่ [7] นำเสนอการประยุกต์ใช้ตรรกศาสตร์ คลุมเครือประเมินผลผู้เรียนจากปัจจัยที่ส่งผลต่อ ผลสัมฤทธิ์ที่เกิดจากลักษณะการตรวจข้อสอบ อัตนัยในกรณีที่มีผู้ตรวจหลายคน ซึ่งมีแนวคิด ว่า ผู้ตรวจแต่ละคนจะมีลักษณะการตรวจข้อสอบที่ แตกต่างกันดังเช่น ผู้ตรวจที่เข้มงวดและผู้ตรวจที่ ผ่อนปรน โดยระบบตรรกศาสตร์คลุมสามารถทำ การแยกลักษณะของผู้ตรวจแต่ละคน (Examiner type) จากกระดาษคำตอบของผู้เรียนออกเป็น Extremely strict, Medium strict, Normal, Medium Lenient และ Extremely Lenient งานวิจัยที่ [8] เสนอวิธีการเปรียบเทียบรูปแบบการประเมิน ผลสัมฤทธิ์ทางการเรียนระหว่างแบบ Classical method กับ Fuzzy logic method การประเมินผล แบบเก่านั้นจะนำคะแนนรวมที่ได้จากการสอบใน แต่ละครั้งมาทำการประเมินผลผู้เรียน แต่การนำ คะแนนรวมมาใช้ในการประเมินอาจจะไม่สื่อ ความหมายถึงความสามารถของผู้เรียนที่แท้จริง จึง ได้นำระบบตรรกศาสตร์คลุมเครือมาประยุกต์ใช้ ประเมินผู้เรียนและผลจากการวิจัยพบว่าระบบ ตรรกศาสตร์คลุมเครือสามารถประเมินผลผู้เรียนได้ ตรงตามลักษณะความสามารถของผู้เรียนและ สามารถเพิ่มโอกาสให้ผู้เรียนมีผลการประเมินที่ดี ขึ้นกว่าแบบเก่า งานวิจัยที่ [9] นำเสนอการประยุกต์ใช้ ขั้นตอนวิธีทางพันธุกรรม (Genetic algorithm : GA)

ในการแก้ปัญหาการบริหารจัดการชั้นเรียนให้กับ กลุ่มนักศึกษาใหม่ ซึ่งอาศัยหลักการของทฤษฎี วิวัฒนาการจากชีววิทยาและการคัดเลือกตาม ธรรมชาติในการแก้ปัญหาการหาค่าที่เหมาะสม ที่สุด นอกจากนี้ยังสามารถนำขั้นตอนวิธีทาง พันธุกรรมมาผสมผสานการทำงานร่วมกับระบบ ตรรกศาสตร์คลุมเครือ (Genetic fuzzy hybrid) [10] ในการแก้ปัญหาการกำหนดจำนวนกฎฟัซซีที่ เหมาะสมให้กับระบบ งานวิจัยที่ [11] เสนอการ ประยุกต์ใช้โครงข่ายประสาทเทียม (Neural network) ในการสร้างรูปแบบการทำนายผลการ เรียนของนักศึกษา ซึ่งหลักการสำคัญของโครงข่าย ประสาทเทียม คือ มีความพยายามที่จะลอกเลียนแบบ การทำงานของเซลล์ประสาทในสมองมนุษย์ เพื่อให้ทำงานได้อย่างมีประสิทธิภาพ และงานวิจัยที่ [12] ได้นำเสนอการเปรียบเทียบวิธีการประเมินผล สัมฤทธิ์ของผู้เรียนด้วยเทคนิคการประมวลผลแบบ เบบ (Soft computing techniques) ซึ่งประกอบไปด้วย การประยุกต์ใช้วิธีการจัดกลุ่ม K-means การจัดกลุ่ม Fuzzy C-mean (FCM) และวิธีการจัดกลุ่ม Subtractive Clustering (SC) รวมถึงได้ผสมผสานวิธีการทำงาน ระหว่างการจัดกลุ่ม Subtractive clustering กับ Fuzzy C-means (SC-FCM) และวิธีการจัดกลุ่ม Subtractive clustering กับระบบ Adaptive neuro fuzzy inference system (SC-ANFIS) ผลจากการ ทดลองเปรียบเทียบของแต่ละวิธีการพบว่าวิธีการ ผสมผสาน SC-ANFIS สามารถสร้างรูปแบบพยากรณ์ ผลสัมฤทธิ์ที่มีประสิทธิภาพมากที่สุด โดยสามารถ ทำการเปรียบเทียบได้จากค่าความ คลาดเคลื่อนของรากกำลังสองเฉลี่ย (Root mean squared error) ซึ่งค่าเฉลี่ย RMSE น้อยที่สุดอยู่ที่ 0.028 สำหรับชุดข้อมูลเรียนรู้ (Learning data) และ 0.087 สำหรับชุดข้อมูลทดสอบ (Testing Data)

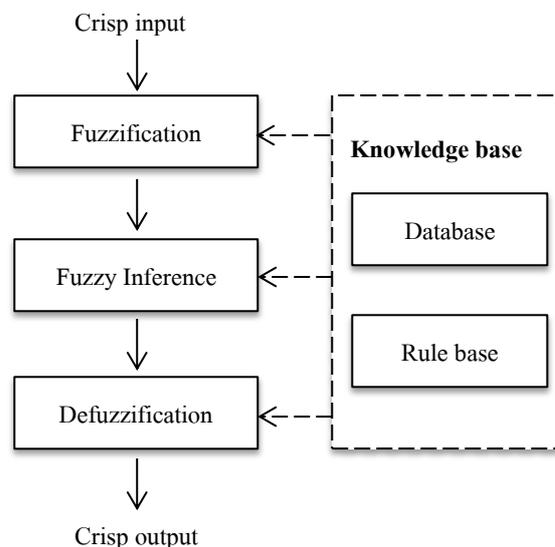
งานวิจัยนี้จึงมีแนวคิดที่จะศึกษาการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียนจากปัจจัยเชิงพุทธิศาสตร์โดยใช้วิธีการผสมผสานขั้นตอนการทำงานระหว่างขั้นตอนวิธีทางพันธุกรรมร่วมกับโครงข่ายประสาทเทียมและระบบตรรกศาสตร์คลุมเครือ รวมทั้งมีการเปรียบเทียบและคัดเลือกวิธีการที่สามารถทำการเรียนรู้จากข้อมูลที่มีประสิทธิภาพมากที่สุด เพื่อที่ผู้สอนสามารถนำไปประยุกต์ใช้สร้างตัวแบบพยากรณ์ในการวิเคราะห์และจำแนกผู้เรียนออกเป็นกลุ่มเก่ง กลุ่มปานกลาง และกลุ่มอ่อน ซึ่งผู้สอนสามารถนำไปใช้เป็นแนวทางในการพัฒนาปรับปรุงการเรียนการสอนให้ตรงกับความรู้ความสามารถของผู้เรียนในแต่ละกลุ่ม

2. ทฤษฎีที่เกี่ยวข้อง

2.1 ฟัซซีลอจิก (Fuzzy logic)

ฟัซซีลอจิกหรือตรรกศาสตร์คลุมเครือเป็นวิธีทางคณิตศาสตร์ที่ช่วยในการตัดสินใจภายใต้ความคลุมเครือหรือไม่ชัดเจน คล้ายกับตรรกะทางความคิดของมนุษย์ คิดค้นโดย L. A. Zadeh ในปี ค.ศ. 1965 ที่อาศัยฟัซซีเซต (Fuzzy Set) เพื่อสื่อถึงความไม่แน่นอน [13] โดยฟัซซีเซตยอมให้มีการกำหนดค่าระดับความเป็นสมาชิกในเซต (Degree of Membership) อยู่ระหว่าง 0 และ 1 ซึ่งแตกต่างจากเซตแบบฉบับ (Classical Set) ที่มีการกำหนดค่าความเป็นสมาชิกเพียงสองค่าเท่านั้น คือ 0 หมายถึงไม่เป็นสมาชิกในเซตและ 1 หมายถึงเป็นสมาชิกในเซต การกำหนดค่าระดับความเป็นสมาชิกของตัวแปรที่สนใจนั้นต้องอาศัยฟังก์ชันความเป็นสมาชิก (Membership Function) ซึ่งมีอยู่หลายชนิด เช่น ฟังก์ชันสามเหลี่ยม ฟังก์ชันสี่เหลี่ยมคางหมู ฟังก์ชันซิกมอยด์ ฟังก์ชันเกาส์เซียน เป็นต้น การเลือกใช้ฟังก์ชันความเป็นสมาชิกชนิดใดจะขึ้นอยู่กับความ

เหมาะสมของข้อมูลตัวแปรนั้น ๆ นอกจากนี้ฟัซซีเซตยังใช้ได้กับตัวแปรเชิงภาษา (Linguistic Variables) เพื่อใช้แสดงคุณภาพหรือปริมาณ [14]



รูปที่ 1 ขั้นตอนการทำงานของระบบอนุมานฟัซซี

ขั้นตอนการทำงานของระบบอนุมานฟัซซีสามารถแบ่งเป็น 3 ขั้นตอนหลัก ๆ ดังแสดงในรูปที่ 1 ขั้นตอนที่หนึ่งคือการแปลงค่าของข้อมูลนำเข้าเป็นค่าฟัซซีนำเข้า (Fuzzification) เป็นการคำนวณค่าความเป็นสมาชิกของข้อมูลนำเข้า (Crisp Input) โดยใช้ฟังก์ชันความเป็นสมาชิก ขั้นตอนที่สองเป็นการอนุมานหรือตีความ (Fuzzy Inferencing) โดยการนำค่าฟัซซีนำเข้าไปตีความหรืออนุมานผ่านกฎฟัซซีที่กำหนดขึ้นภายในระบบ และจะได้ผลลัพธ์เป็นค่าฟัซซีส่งออก โดยกฎฟัซซีที่นิยมใช้คือกฎฟัซซีแบบ ถ้า-แล้ว (Fuzzy If-Then Rule) ที่อาศัยหลักการของเหตุและผล และในขั้นตอนสุดท้ายคือการนำค่าฟัซซีให้เป็นค่าปกติ (Defuzzification) โดยการนำค่าฟัซซีส่งออกมาแปลงเป็นค่าเป็นค่าปกติ (Crisp Output) ซึ่งมีหลายวิธี เช่น วิธีถ่วงเฉลี่ยถ่วงน้ำหนัก วิธีการหาจุดศูนย์กลาง [14]

2.2 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic algorithm)

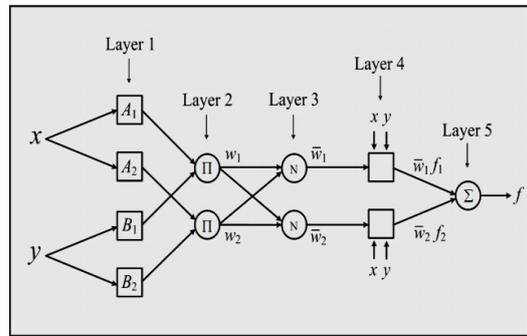
ขั้นตอนวิธีเชิงพันธุกรรมเป็นการค้นหาคำตอบโดยประมาณของปัญหาซึ่งอาศัยหลักของทฤษฎีวิวัฒนาการจากชีววิทยาและการคัดเลือกตามธรรมชาติ เพื่อแก้ปัญหาการหาค่าที่เหมาะสมที่สุด โดยการแทนคำตอบที่มีอยู่ในลักษณะของโครโมโซม (Chromosome)

ขั้นตอนของวิธีเชิงพันธุกรรมจะประกอบไปด้วย 4 ส่วนหลัก ๆ คือ 1) ประชากรของกลุ่มคำตอบแต่ละชุด ซึ่งแต่ละชุดจะแสดงค่าที่เป็นไปได้ของคำตอบที่ต้องการ 2) ค่าที่เหมาะสม (Fitness function) ซึ่งเป็นค่าที่ได้จากการประเมินของกลุ่มคำตอบแต่ละกลุ่มว่าเป็นค่าที่ต้องการหรือเป็นค่าที่ไม่ต้องการ 3) วิธีการคัดเลือก (Selection function) คือวิธีการตัดสินใจว่าจะทำการเลือกกลุ่มคำตอบที่ดีจากกลุ่มประชากรในปัจจุบันเพื่อจะใช้ในการสร้างกลุ่มประชากรใหม่ และ 4) ตัวปฏิบัติการทางพันธุกรรม (Genetic operator) เช่น การไขว้ข้าม (Crossover) และการกลายพันธุ์ (Mutation) เป็นกระบวนการสำหรับการสร้างกลุ่มประชากรใหม่ [15]

2.3 ระบบนิเวศฟัซซี (Adaptive neural fuzzy inference system : ANFIS)

ระบบอนุมานฟัซซีบนฐานโครงข่ายที่ปรับตัวได้หรือระบบนิเวศฟัซซีเป็นการประยุกต์ใช้ระบบโครงข่ายประสาทเทียม (Artificial Neural Networks หรือ ANN) ที่มีจุดเด่นด้านการเรียนรู้จากข้อมูลมาใช้ร่วมกับตรรกศาสตร์คลุมเครือหรือฟัซซีลอจิก (Fuzzy logic) ที่มีจุดเด่นสามารถตีความให้อยู่ในรูปของกฎ If-Then โดยโครงข่ายประสาทเทียมจะทำหน้าที่ช่วยสร้างกฎเกณฑ์ฟัซซีให้กับข้อมูลนำเข้าและข้อมูลนำออก ทำให้การสร้างระบบ

ฟัซซีนั้นง่ายขึ้น และยังช่วยให้การวิเคราะห์ข้อมูลของระบบมีความแม่นยำขึ้น



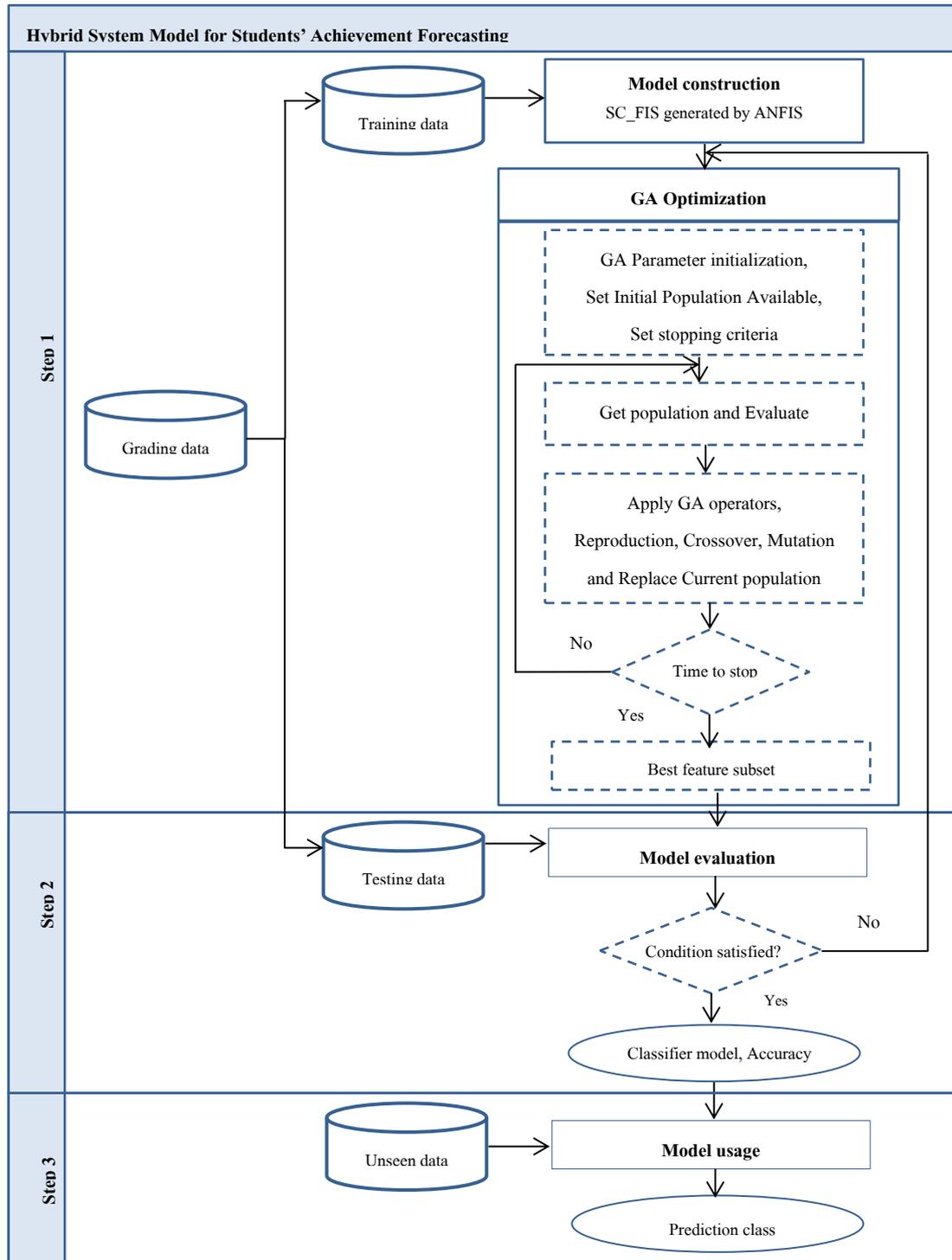
รูปที่ 2 โครงสร้างของ ANFIS

ระบบผสมนิเวศฟัซซีที่มีชื่อเสียงเป็นที่รู้จักกันมากได้แก่ ANFIS ดังแสดงโครงสร้างในรูปที่ 2 ที่ใช้โครงข่ายประสาทเทียมมาช่วยในการเรียนรู้ข้อมูล ซึ่งเป็นเทคโนโลยีที่มาจากงานวิจัยด้านปัญญาประดิษฐ์ (Artificial intelligence : AI) เพื่อใช้ในการคำนวณหาค่าฟังก์ชันจากกลุ่มข้อมูลวิธีการของระบบโครงข่ายประสาทเทียมเป็นวิธีการที่ให้ระบบเรียนรู้จากตัวอย่างต้นแบบแล้วทำการสอน (Training) ให้ระบบได้รู้จักคิดแก้ปัญหาโครงสร้างของระบบโครงข่ายประสาทเทียมจะประกอบไปด้วยโหนดข้อมูลนำเข้าและข้อมูลนำออก และการประมวลผลมีลักษณะกระจายอยู่ในโครงสร้างเป็นชั้น ๆ ได้แก่ ชั้นข้อมูลนำเข้า ชั้นข้อมูลนำออก และชั้นซ่อน ซึ่งการประมวลผลของนิเวศฟัซซีเน็ตเวิร์กจะอาศัยการส่งการทำงานผ่านโหนดต่าง ๆ ในชั้นเหล่านี้ [16]

3. วิธีดำเนินงานวิจัย

งานวิจัยนี้ผู้วิจัยได้เลือกใช้ข้อมูลเกี่ยวกับผลสัมฤทธิ์ทางการเรียนของผู้เรียน เพื่อนำมาใช้เป็นชุดข้อมูลเรียนรู้ (Training data) และชุดข้อมูลทดสอบ (Testing data) สำหรับสร้างและทดสอบตัวแบบผสมผสานวิธีการพยากรณ์ จากวิชาที่จัดให้มี

การเรียนการสอนขึ้นในระดับชั้นปริญญาตรีของ การศึกษา
 กลุ่มวิชาศึกษาทั่วไปตลอดระยะเวลา 6 ภาค



รูปที่ 3 ขั้นตอนการสร้างตัวแบบพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียน

(1/2556 – 2/2558) มีจำนวนผู้เรียนทั้งสิ้น 1,320 คน แบ่งเป็นผลการสอบ Midterm Final Assignment1 และ Assignment2 โดยใช้ระดับแบ่งผลสัมฤทธิ์ตามเกณฑ์มาตรฐานของสถาบันอุดมศึกษาในประเทศไทยแบ่งเป็น 8 ระดับ คือ A (Excellent), B+ (Very good), B (Good), C+ (Fair Good), C (Fair), D+ (Poor), D (Very poor) และ F (Fail) ข้อมูลดังกล่าวนั้นจะถูกนำมาใช้สร้างและวิเคราะห์ประสิทธิภาพของตัวแบบพยากรณ์ ดังแสดงข้อมูลในตารางที่ 1

ตารางที่ 1 ข้อมูลจำนวนผู้เรียนและคะแนนเฉลี่ยแบ่งตามเกรด

Grade	Student	Percent	Average
A	72	6.0%	85.9
B+	280	23.3%	76.7
B	386	32.2%	72.2
C+	250	20.8%	67.2
C	182	15.2%	62.3
D+	72	6.0%	57.7
D	46	3.8%	52.1
F	32	2.7%	46.1
All	1,320	100%	

ขั้นตอนการสร้างตัวแบบพยากรณ์ผลสัมฤทธิ์ของผู้เรียน เพื่อให้ได้ค่าพยากรณ์ที่

ใกล้เคียงกับค่าตามความเป็นจริงมากที่สุดและมีค่าความผิดพลาด RMSE น้อยที่สุดนั้น ผู้วิจัยจึงได้นำเสนอกรอบการทำงานระบบผสมผสานขั้นตอนวิธีทางพันธุกรรมร่วมกับโครงข่ายประสาทเทียมและฟิวชันลอจิกคังแสดงในรูปที่ 2 โดยเริ่มต้นจากการสร้างระบบฟิวชันจากกลุ่มตัวอย่างข้อมูลชุดสอนและใช้ ANFIS ทำการประเมินค่าความเหมาะสมของระบบ จากนั้นใช้ GA สร้างประชากรพารามิเตอร์ฟังก์ชันความเป็นสมาชิกของจำนวนกฎฟิวชันและทำการประเมินค่าความเหมาะสม หากยังไม่ได้ตามเงื่อนไขการหยุดค้นหา ก็จะสร้างประชากรของพารามิเตอร์ชุดใหม่โดยใช้ตัวดำเนินการทางพันธุกรรมและทำการประเมินค่าความเหมาะสมใหม่อีกครั้ง ทำซ้ำเช่นนี้ไปเรื่อย ๆ จนกว่าจะตรงกับเงื่อนไขการหยุดค้นหา ซึ่งสามารถอธิบายเป็นลำดับขั้นตอนได้ 6 ขั้นตอน ดังนี้

1) ขั้นตอนการเตรียมข้อมูลที่ใช้ในการสร้างตัวแบบพยากรณ์ ซึ่งจะใช้วิธีการแบ่งข้อมูลแบบ Split test โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลเรียนรู้ (Training data) สำหรับสร้างตัวแบบและชุดข้อมูลทดสอบ (Testing data) สำหรับใช้ทดสอบตัวแบบพยากรณ์ ดังแสดงรายละเอียดข้อมูลในตารางที่ 2

ตารางที่ 2 Training data and Testing data

Level of Achievement	Training data and Testing data			
	Grade	Marks	Training data	Testing data
Cluster-1 (Excellence)	A	80.00 – 100.00	36	36
Cluster-2 (Very good)	B+	75.00 – 79.99	140	140
Cluster-3 (Good)	B	70.00 – 74.99	193	193
Cluster-4 (Fairly good)	C+	65.00 – 69.99	125	125
Cluster-5 (Fair)	C	60.00 – 64.99	91	91
Cluster-6 (Poor)	D+	55.00 – 59.99	36	36
Cluster-7 (Very poor)	D	50.00 – 54.99	23	23
Cluster-8 (Fail)	F	1.00 – 49.99	16	16
All			660	660

2) ทำการ โหลดชุดข้อมูลเรียนรู้ (Training data) สำหรับนำมาสร้างระบบกฎฟัซซีแบบ Takagi-sugeno [17] โดยใช้วิธีการจัดกลุ่มแบบลบออก (Subtractive Clustering) [18] มาช่วยสร้างจำนวนกฎฟัซซีที่เหมาะสมให้กับโครงสร้างระบบอนุมานฟัซซี ดังต่อไปนี้

ขั้นตอนที่ 1 นำข้อมูลทุกจุดมาคำนวณหา ค่าความหนาแน่นจากสมการ

$$D_i = \sum_{j=i}^n \exp\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right) \quad (1)$$

โดยที่ x_i คือข้อมูลจุดที่คำนวณหาค่าความหนาแน่น

x_j คือข้อมูลจุดอื่น ๆ ในชุดข้อมูล

r_a คือค่ารัศมีระหว่างกลุ่ม

$\| \cdot \|$ คือการวัดระยะทางแบบยูคลิเดียน

การกำหนดค่ารัศมีระหว่างกลุ่มมีผลต่อการจัดกลุ่ม โดยค่ารัศมีของกลุ่มน้อยจะทำให้จำนวนกลุ่มเพิ่มมากขึ้น ซึ่งในงานวิจัยนี้ได้กำหนดค่ารัศมีในการจัดกลุ่มของข้อมูลนำเข้าในแต่ละมิติเป็น 0.30 0.25

0.30 0.40 และ 0.25 ตามลำดับ เมื่อคำนวณค่าความหนาแน่นครบทุกจุดแล้ว จุดใดที่มีค่าความหนาแน่นสูงสุดจุดนั้นจะกลายเป็นจุดศูนย์กลางของกลุ่ม

ขั้นตอนที่ 2 เมื่อได้จุดศูนย์กลางของกลุ่มแรกแล้วให้ทำการคำนวณหาจุดศูนย์กลางของกลุ่มถัดไปจากสมการ

$$D_i = D_i - D_k^* \exp\left(-\frac{\|x_i - x_k^*\|^2}{(r_b/2)^2}\right) \quad (2)$$

โดยที่ x_k^* คือจุดศูนย์กลางใหม่ที่อาจเป็นไปได้

x_j คือจุดศูนย์กลางล่าสุดที่หาได้

r_b คือค่ารัศมีระหว่างกลุ่ม

η ค่าใช้ลดปัญหาระยะขอบของกลุ่มที่อยู่ใกล้กัน

ขั้นตอนที่ 3 คำนวณหาจุดศูนย์กลางของกลุ่มที่อาจเป็นไปได้จุดใหม่ ซึ่งจะเป็จุดศูนย์กลางของกลุ่มถัดไปก็ต่อเมื่อ

$$\frac{d_{\min}}{r_a} + \frac{D_k^*}{D_i} \geq 1 \quad (3)$$

โดยที่ d_{min} คือระยะทางที่น้อยที่สุดระหว่าง x_k^* กับ x_i ถ้าจุดศูนย์กลางใหม่ไม่เป็นตามเงื่อนไขของสมการที่ 3 จุดศูนย์กลางใหม่นั้นจะถูกปฏิเสธ และย้อนกลับไปคำนวณหาจุดศูนย์กลางใหม่ในขั้นตอนที่สอง โดยเงื่อนไขในการกำหนดจุดสิ้นสุดของการจัดกลุ่มคือ

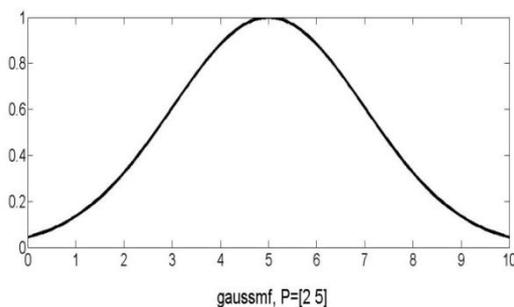
$$D_k^* > \epsilon D_i \tag{4}$$

โดยที่ ϵ คือค่าสัดส่วนในการปฏิเสธ

ขั้นตอนที่ 4 เมื่อสิ้นสุดกระบวนการจัดกลุ่มแล้วสามารถคำนวณค่าความเป็นสมาชิกของข้อมูลในแต่ละกลุ่มได้จากสมการ

$$\mu_{ik} = \exp\left(\frac{\|x_i - x_k\|^2}{(r_a/2)^2}\right) \tag{5}$$

หลังจากได้จำนวนกลุ่มของข้อมูล ทำให้ได้โครงสร้างระบบอนุमानฟัซซีเริ่มต้น โดยใช้ฟังก์ชันความเป็นสมาชิกของข้อมูลนำเข้า คือ เกาซ์เซียน มีพารามิเตอร์ 2 ค่า คือ $\{m, \sigma\}$ ซึ่ง m หมายถึงค่าเฉลี่ย และ σ หมายถึง ค่าเบี่ยงเบนมาตรฐาน ดังสมการที่ 6 และฟังก์ชันความเป็นสมาชิกของข้อมูลนำเข้าคือ linear จากนั้นใช้ ANFIS ทำการประเมินค่าความเหมาะสมพารามิเตอร์ของกฎฟัซซีในระบบ



รูปที่ 4 ฟังก์ชันสมาชิกเกาซ์เซียน

$$f(x; m, \sigma) = \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right) \tag{6}$$

3) ใช้ขั้นตอนวิธีทางพันธุกรรมฝึกให้ตัวแบบเรียนรู้ โดยการสร้างประชากรพารามิเตอร์ฟังก์ชันความเป็นสมาชิกของกฎฟัซซีและทำการประเมินค่าความเหมาะสม หากยังไม่ได้ตามเงื่อนไขการหยุดค้นหาที่จะสร้างประชากรของพารามิเตอร์ชุดใหม่โดยใช้ตัวดำเนินการทางพันธุกรรมและทำการประเมินค่าความเหมาะสมใหม่อีกครั้ง ทำซ้ำเช่นนี้ไป จนกว่าจะตรงกับเงื่อนไขการหยุดค้นหาในตารางที่ 3 แสดงการกำหนดพารามิเตอร์ให้กับตัวดำเนินการ GA

ตารางที่ 3 การกำหนดพารามิเตอร์ให้กับตัวดำเนินการ GA

Setting Type	Value
Encoding scheme	Value Encoding
Population size	25
Evolution generation	50
Selection	Roulette Wheel
Crossover	One point
Mutation	Uniform
PC	0.4
PM	0.05
Elitism	Yes
Generations	50

สำหรับโครงสร้างของโครโมโซมที่ใช้ในขั้นตอนวิธีทางพันธุกรรมจะประกอบไปด้วยค่าพารามิเตอร์ของฟังก์ชัน ข้อมูลนำเข้าและข้อมูลนำเข้า ดังแสดงในรูปที่ 5

$$\left[\begin{array}{cc} \underbrace{m_{11}, \sigma_{11}; m_{12}, \sigma_{12}}_{\text{Parameter of input1}}; \underbrace{m_{21}, \sigma_{21}; m_{22}, \sigma_{22}}_{\text{Parameter of input2}} \\ \underbrace{m_{11}, \sigma_{11}; m_{12}, \sigma_{12}}_{\text{Parameter of output1}}; \underbrace{m_{21}, \sigma_{21}; m_{22}, \sigma_{22}}_{\text{Parameter of output2}} \end{array} \right]$$

รูปที่ 5 โครงสร้างโครง โมโซมฟังก์ชันสมาชิกเกาส์เซียน

4) ทำการตรวจสอบความแม่นยำของโครง โมโซมโดยใช้ชุดข้อมูลทดสอบ ด้วยวิธีการคำนวณ หาค่าความคลาดเคลื่อนของรากกำลังสองเฉลี่ย (Root mean squared error : RMSE) ดังสมการที่ 6

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{obs} - x_{model})^2}{n}} \quad (7)$$

โดยที่ x_{obs} คือค่า Output data ที่ได้จากการพัฒนา ระบบ และ x_{model} คือค่า Output data ที่ได้จาก โมเดลต้นแบบ

5) การเลือกตัวแบบจะเลือกตัวแบบที่มีค่าพยากรณ์แม่นยำที่สุดและมีค่า RMSE น้อยที่สุดจากการทดลองสร้างตัวแบบพยากรณ์หลาย ๆ ตัวแบบที่แตกต่างกัน ซึ่งแต่ละตัวแบบจะต่างกันที่ค่าพารามิเตอร์ฟังก์ชันความเป็นสมาชิกกฎฟัซซีของข้อมูลนำเข้าและข้อมูลนำออก

6) เป็นขั้นตอนการนำตัวแบบพยากรณ์ที่สร้างขึ้นมาไปใช้กับข้อมูลที่ไม่เคยเห็นมาก่อน (Unseen data) เพื่อทำการพยากรณ์ให้กับข้อมูลชุดนั้น

4. ผลการทดลอง

จากการนำข้อมูลผลสัมฤทธิ์ทางการเรียนของผู้เรียนจำนวน 1,320 คน ไปทดสอบกับตัวแบบพยากรณ์ทั้ง 3 ตัวแบบ ซึ่งได้แก่ ตัวแบบที่หนึ่งเป็น

การผสมผสานขั้นตอนวิธีการจัดกลุ่ม Subtractive clustering กับวิธี Fuzzy logic (SC-FL) ตัวแบบที่สองเป็นการผสมผสานขั้นตอนวิธี Subtractive clustering กับวิธี Adaptive neuro fuzzy inference system (SC-ANFIS) และตัวแบบที่สามเป็นการผสมผสานขั้นตอนวิธี Genetic algorithm กับวิธี Adaptive neuro fuzzy inference system (GA-ANFIS) ซึ่งสามารถทำการสรุปผลการทดลองได้ดังนี้ สำหรับชุดข้อมูลเรียนรู้ ดังแสดงในรูปที่ 6 พบว่า ตัวแบบ SC-FL สามารถทำการพยากรณ์โดยมีเกรด F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ 99% เกรด A และ D อยู่ในระดับร้อยละ 98% ตามลำดับ ซึ่งเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 93% ตัวแบบ SC-ANFIS สามารถทำการพยากรณ์โดยมีเกรด A C D+ D และ F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ 100% เกรด C+ อยู่ในระดับร้อยละ 99% ตามลำดับ ซึ่งเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 98.75% และตัวแบบ GA-ANFIS สามารถทำการพยากรณ์โดยมีเกรด A B+ C+ C D+ D และ F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ 100% เกรด B อยู่ในระดับร้อยละ 99% ตามลำดับ ซึ่งเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 99.88% และสำหรับในชุดข้อมูลทดสอบดังแสดงในรูปที่ 7 พบว่า ตัวแบบ SC-FL สามารถทำการพยากรณ์โดยมีเกรด F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ 98% เกรด D อยู่ในระดับร้อยละ 97% ตามลำดับ ซึ่งเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 92.88% ตัวแบบ SC-ANFIS สามารถทำการพยากรณ์โดยมีเกรด A D+ D และ F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ

100% เกรด C อยู่ในระดับร้อยละ 98% ตามลำดับ ซึ่งเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 98.63% และตัวแบบ GA-ANFIS สามารถทำการพยากรณ์โดยมีเกรด A C+ C D+ D และ F มีค่าความถูกต้องในการพยากรณ์มากที่สุดอยู่ในระดับร้อยละ 100% เกรด B และ B+ อยู่ในระดับร้อยละ 99% ตามลำดับเมื่อพิจารณาประสิทธิภาพโดยรวมแล้วมีความถูกต้องอยู่ที่ร้อยละ 99.75%

นอกจากนี้เมื่อทำการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error : RMSE) พบว่าสำหรับชุดข้อมูลเรียนรู้ ตัวแบบพยากรณ์ SC-FL ให้ค่า RMSE น้อยที่สุดเท่ากับ 0.871 ตัวแบบ SC-ANFIS เท่ากับ 0.173 และตัวแบบ GA-ANFIS เท่ากับ 0.024 และสำหรับชุดข้อมูลทดสอบ ตัวแบบพยากรณ์ SC-LF ให้ค่า RMSE น้อยที่สุดเท่ากับ 1.429 ตัวแบบ SC-ANFIS เท่ากับ 0.237 และตัวแบบ GA-ANFIS เท่ากับ 0.087 ดังแสดงในตารางที่ 4

ตารางที่ 4 แสดงค่าความผิดพลาดของแต่ละตัวแบบ

Forecasting methods	RMSE	
	Training data	Testing data
SC-FL	0.871	1.429
SC-ANFIS	0.173	0.237
GA-ANFIS	0.024	0.087

5. บทสรุป

งานวิจัยนี้ได้นำความรู้ทางด้านวิทยาการคอมพิวเตอร์มาใช้ในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียน โดยวิธีการผสมผสานขั้นตอนการทำงานซึ่งประกอบไปด้วยวิธีการจัดกลุ่มแบบลบออก (Subtractive algorithm) วิธีโครงข่ายประสาทเทียม (Neural network) วิธีระบบตรรกศาสตร์คลุมเครือ (Fuzzy logic) และขั้นตอนวิธีทางพันธุกรรม

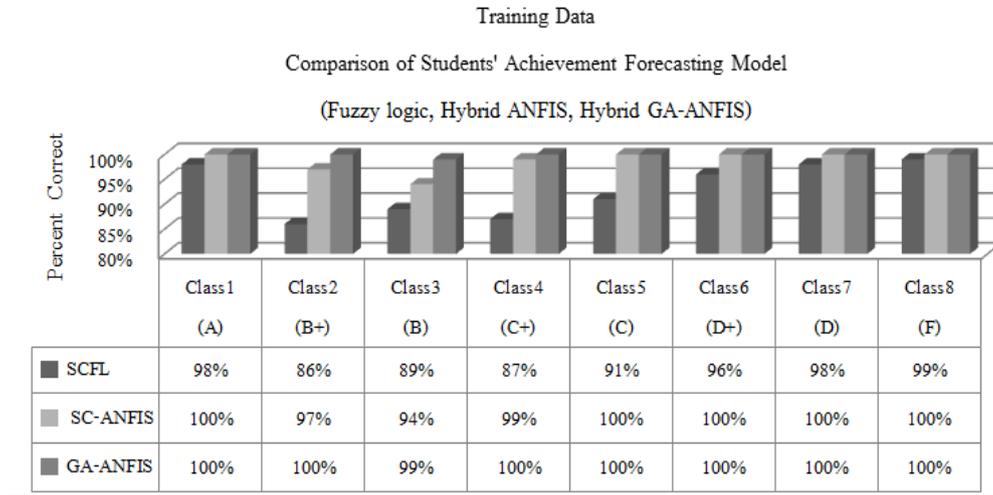
(Genetic algorithm) มาทำการสร้างตัวแบบพยากรณ์ผลสัมฤทธิ์ทางการเรียน จากนั้นจึงทำการเปรียบเทียบหาประสิทธิภาพความแม่นยำในการพยากรณ์เพื่อหาตัวแบบพยากรณ์ที่ให้ความแม่นยำสูงสุดที่เหมาะสมกับการนำไปใช้ในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียนจากปัจจัยเชิงพุทธพิสัย

ผลการทดสอบพบว่าตัวแบบพยากรณ์ GA-ANFIS สามารถสร้างตัวแบบพยากรณ์ระบบกฎฟuzzy ที่มีจำนวนกฎที่เหมาะสมจำนวน 4 กฎ ดังแสดงในรูปที่ 8 และให้ผลการประเมินค่าความแม่นยำในการพยากรณ์ผลสัมฤทธิ์ของผู้เรียนมีร้อยละความถูกต้องมากที่สุด โดยให้ค่าความแม่นยำโดยรวมมากที่สุดสำหรับในชุดข้อมูลเรียนรู้ร้อยละ 99.88% และสำหรับชุดข้อมูลทดสอบร้อยละ 99.75% นอกจากนี้เมื่อทำการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย พบว่าตัวแบบพยากรณ์ GA-ANFIS ให้ค่า RMSE น้อยที่สุดเมื่อเทียบกับตัวแบบอื่นอยู่ที่ 0.024 สำหรับชุดข้อมูลเรียนรู้ และ 0.087 สำหรับชุดข้อมูลทดสอบ

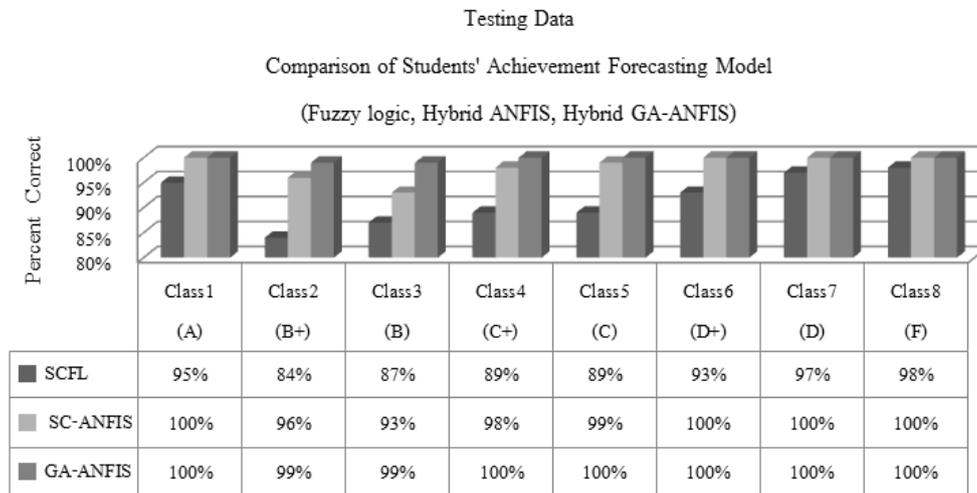
จากผลการวิจัยพบว่าตัวแบบพยากรณ์ GA-ANFIS มีค่าความแม่นยำในการพยากรณ์มากกว่าตัวแบบ SCFL และตัวแบบ SC-ANFIS อาจเนื่องมาจากการผสมผสานขั้นตอนการทำงานนั้นได้นำข้อดีของแต่ละวิธีการมาทำงานร่วมกัน โดยเริ่มจากการใช้วิธีการจัดกลุ่ม Subtractive clustering มาทำการกำหนดโครงสร้างของระบบฟuzzy เริ่มต้นให้กับ ANFIS และใช้ขั้นตอนวิธีทางพันธุกรรม (GA) แก้ปัญหาการหาค่าที่เหมาะสมที่สุดให้กับพารามิเตอร์ฟังก์ชันความเป็นสมาชิกใน ANFIS ซึ่งสอดคล้องกับงานวิจัยที่ [19] ที่ได้ประยุกต์ใช้ตัวแบบ ANFIS เปรียบเทียบกับตัวแบบระบบผสมผสาน GA-ANFIS ในการพยากรณ์ค่าความเข้มข้นของสาร

นิโคตินในปอดของผู้ที่สูบบุหรี่ ซึ่งผลจากการวิจัย พบว่าตัวแบบพยากรณ์ GA-ANFIS ให้ค่าผิดพลาด

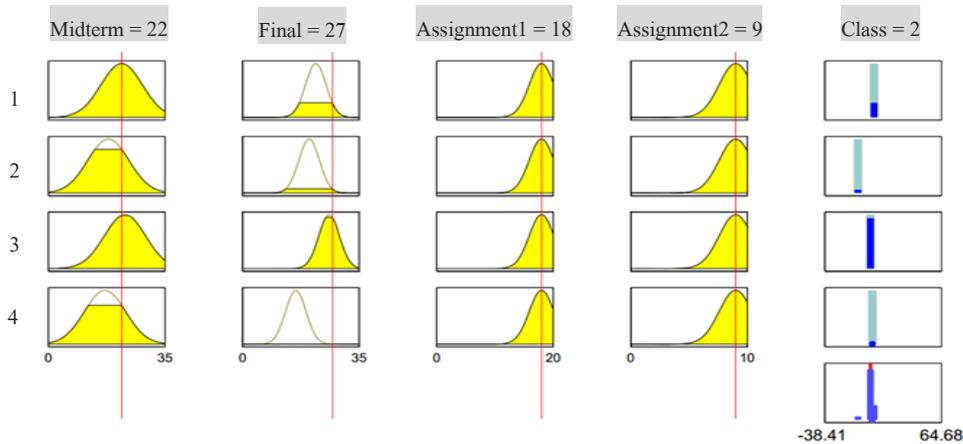
ในการพยากรณ์น้อยกว่าตัวแบบ ANFIS โดยที่มีค่า RMSE เท่ากับ 0.001 และ 0.013 ตามลำดับ



รูปที่ 6 ค่าประสิทธิภาพในการพยากรณ์ของชุดข้อมูลเรียนรู้



รูปที่ 7 ค่าประสิทธิภาพในการพยากรณ์ของชุดข้อมูลทดสอบ



รูปที่ 8 ระบบกฎฟัซซี่ที่ได้จากขั้นตอนวิธีผสมผสานตัวแบบ GA-ANFIS

6. เอกสารอ้างอิง

[1] อ. อภิชาติบรรลือ. เอกสารประกอบการสอน รายวิชาการวัดและประเมินผลทางการศึกษา. กรุงเทพมหานคร : ภาควิชาครุศาสตร์ เครื่องกล คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ. 2553

[2] Saleh, I. and S.-i. Kim, A fuzzy system for evaluating students' learning achievement. Expert Systems with Applications, 2009. 36(3, Part 2): p. 6236-6243.

[3] Shyi-Ming, C. and L. Ting-Kuei. A new method to evaluate students' learning achievement by automatically generating the importance degrees of attributes of questions. in Machine Learning and Cybernetics (ICMLC), 2010 International Conference on. 2010.

[4] Hameed, I.A., Using Gaussian membership functions for improving the reliability and robustness of students' evaluation systems. Expert Systems with Applications, 2011. 38(6): p. 7135-7142.

[5] Ingoley, S.N. and P.D. J.W. Bakal, Evaluating Students' Performance using Fuzzy Logic. International Conference in Recent Trends in Information Technology and Computer Science (ICRTITCS - 2012). Proceedings published in International Journal of Computer Applications® (IJCA) (0975 – 8887).

[6] Ingoley, S.N. and J.W. Bakal. Evaluating students' performance using four-node fuzzy controller. in Engineering (NUiCONE), 2013 Nirma University International Conference on. 2013.

[7] Ingoley, S.N. and J.W. Bakal. A novel method for inferring strict and lenient marks into normal marks using fuzzy logic. in Engineering (NUiCONE), 2013 Nirma University International Conference on. 2013.

- [8] Sakthivel, E., K.S. Kannan, and S. Arumugam, Optimized Evaluation of Students Performances Using Fuzzy Logic. *International Journal of Scientific & Engineering Research*, 2013. 4(9): p. 6.
- [9] Zukhri, Z. and Omar, K., Solving new student allocation problem with genetic algorithm: a hard problem for partition based approach. *Int. J. Soft Comput. Appl.*, 2008, 2, 6–15.
- [10] Mankad, K., Sajja, P. S. and Akerkar, R., Evolving rules using genetic fuzzy approach: an educational case study. *Int. J. Soft Comput.*, 2011, 2(1), 35–46.
- [11] Arsad, P.M., N. Buniyamin, and J.I.A. Manan. Neural network model to predict electrical students' academic performance. in *Engineering Education (ICEED)*, 2012 4th International Congress on. 2012.
- [12] R.-S. Yadav, P. Ahmed, A.-K. Som, and S. Pal, “Academic performance evaluation using soft computing techniques” *CURRENT SCIENCE*, vol. 106, no. 11, pp. 1505-1517, 2014.
- [13] L. A. Sadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338-353, June 1965.
- [14] พ. มีสัจ, ระบบฟัซซีและโครงข่ายประสาทเทียม. เอกสารประกอบการสอน, คณะเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2551.
- [15] O. yildiz, A. Bal, S. Gulsecen and F. D. Kentli, "A Genetic-Fuzzy Based Mathematical Model to Evaluate The Distance Education Students' Academic Performance". *Procedia - Social and Behavioral Sciences*, 2012. 55(0): p. 409-418.
- [16] Yuill, W., et al. Application of Adaptive Neuro Fuzzy Inference System (ANFIS) based short term load forecasting in South African power networks. in *Universities Power Engineering Conference (UPEC)*, 2010 45th International. 2010.
- [17] K. Yan, C. and T.S. Dillon. Traffic flow prediction using orthogonal arrays and Takagi-Sugeno neural fuzzy models. in *Neural Networks (IJCNN)*, 2014 International Joint Conference on. 2014.
- [18] Zhou, H., X.H. Wu, and X.G. Li. An ANFIS model of electricity price forecasting based on subtractive clustering. in *Power and Energy Society General Meeting*, 2011 IEEE. 2011.
- [19] Fazlic, L.B. and A. Avdagic. Prediction of lung nicotine concentration based on novel GA-ANFIS system approach. in *Information, Communication and Automation Technologies (ICAT)*, 2015 XXV International Conference on. 2015.