อัลกอริทึมเอ็มอาร์เอ็นสำหรับปัญหาการจำแนกประเภท

MRN Algorithm for Classification Problems

อนุสรณ์ เจริญนาน
ANUSORN CHARLEONNAN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2558

KMITL-2015-SC-M-002-054

อัลกอริทึมเอ็มอาร์เอ็นสำหรับปัญหาการจำแนกประเภท

MRN Algorithm for Classification Problems

อนุสรณ์ เจริญนาน

ANUSORN CHARLEONNAN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2558

KMITL-2015-SC-M-002-054

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MRN Algorithm for Classification Problems

ANUSORN CHARLEONNAN

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF SCIENCE IN COMPUTER SCIENCE

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2015

KMITL-2015-SC-M-002-054

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2015

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ "อัลกอริทึมเอ็มอาร์เอ็นสำหรับปัญหาการจำแนกประเภท"
"MRN ALGORITHM FOR CLASSIFICATION PROBLEMS"

ชื่อนักศึกษา นายอนุสรณ์ เจริญนาน

รหัสประจำตัว 53650806

ปริญญา วิทยาศาสตรมหาบัณฑิต (สาขาวิชาวิทยาการคอมพิวเตอร์)

ภาควิชา วิทยาการคอมพิวเตอร์

อาจารย์ที่ปรึกษาวิทยานิพนธ์ ดร.สายชล ใจเย็น

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม -

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
ผศ.ดร.ศรัณย์ อินทโกสุม ประธานกรรมการ	
ผศ.ดร.อนันตพร หรรษคุณาฒัย อาจารย์บัณฑิตประจำ (ในสาขาวิชาที่เกี่ยวข้อง)	
ผศ.ดร.ศุภกานต์ พิมลธเรศ ผู้ทรงคุณวุฒิจากภายนอกสถาบันฯ	
ดร.สายชล ใจเย็น อาจารย์ที่ปรึกษาวิทยานิพนธ์	

วัน/ เดือน/ ปี ที่สอบ 23 กรกฎาคม พ.ศ.2558 เวลา 13.00-16.00 น.

สถานที่สอบ ห้อง 306 อาคารปฏิบัติการใหม่ ชั้น 3

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

คณะวิทยาศาสตร์รับรองแล้ว

(รองศาสตราจารย์ ดร.ดนัย ธนะบริพัฒน์)
คณบดีคณะวิทยาศาสตร์
วันที่ 28 เดือน ก.ค. พ.ศ. 58

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	อัลกอริทึมเอ็มอาร์เอ็นสำหรับปัญหาการจำแนกประเภท
นักศึกษา	นายอนุสรณ์ เจริญนาน
รหัสประจำตัว	53650806
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	วิทยาการคอมพิวเตอร์
พ.ศ.	2558
อาจารย์ผู้ควบคุมวิทยานิพนธ์	อาจารย์ ดร.สายชล ใจเย็น

บทคัดย่อ

ปัจจุบันนี้ มีการนำแนวคิดของโครงข่ายประสาทเทียมในหลากหลายรูปแบบมาใช้ในการแก้ปัญหาการจำแนกประเภทข้อมูล โดยส่วนมาก นิยมใช้การส่งอยู่ 2 ประเภทคือ การส่งแบบเชิงเส้น และ การส่งแบบไม่เชิงเส้น เพื่อส่งข้อมูลจากปริภูมิข้อมูลนำเข้าไปยังปริภูมิคุณลักษณะ นอกจากนี้ การทำงานโครงข่ายประสาทเทียมในบางรูปแบบ ยังมีการใช้แนวคิดมาจากทฤษฎีความน่าจะเป็น ในทางปฏิบัติ แต่ละโมเดลจะมีความเหมาะสมกับข้อมูลแต่ละประเภท ขึ้นอยู่กับการกระจายตัวของข้อมูลนั้น ๆ ดังนั้น ข้อมูลบางประเภทจึงมีความเหมาะสมกับการทำงานโดยใช้การส่งแบบเชิงเส้น การส่งแบบไม่เชิงเส้น หรือทฤษฎีความน่าจะเป็น เนื่องจากการกระจายตัวของข้อมูลในปัญหาการจำแนกกลุ่มข้อมูลมีความหลากหลาย งานวิจัยชิ้นนี้จึงนำเสนออัลกอริทึมแบบใหม่ ซึ่งใช้การรวมกลุ่มของ โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และตัวจำแนกแบบเบย์อย่างง่าย สำหรับการแก้ปัญหาการจำแนกประเภทข้อมูล จากผลการทดลองที่ได้ แสดงให้เห็นว่าอัลกอริทึมแบบใหม่นี้สามารถให้ผลความถูกต้องของการจำแนกประเภทข้อมูลได้อย่างมีประสิทธิภาพมากขึ้น ในชุดข้อมูลที่ได้ทำการทดสอบ

คำสำคัญ : การเรียนรู้แบบรวมกลุ่ม เอดาบูสท์ โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น ตัวจำแนกแบบเบย์อย่างง่าย โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis	MRN Algorithm for Classification Problems
Student	Mr.anusorn Charleonnan
StudentID	53650806
Degree	Master of Science
Program	Computer Science
Year	2015
Thesis Advisor	Dr. Saichon Jaiyen

ABSTRACT

Currently, various perspectives of artificial neural networks are proposed for solving classification problems. They are mainly based on two types of mapping functions, namely, linear and nonlinear, for mapping an input space into a feature space. In addition, some neural networks are based on probability theory. Practically, each model is appropriated for each data type, depending on a distribution of the data. Therefore, some data are appropriated for linear mapping, nonlinear mapping, or probabilistic models. Since the data distribution in classification problems is various, we propose the new ensemble model based on Multilayer Perceptron Neural Network, Radial Basis Function Neural Network, and Naive Bayes Classifier for classification problems. According to the experimental results, they have shown that our proposed model can improve the accuracy of classification on the testing data sets.

Keywords : Ensemble Learning, AdaBoost, Multilayer Perceptron Neural Network, Naive Bayes Classifier, Radial Basis Function Neural Network

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยคำแนะนำ คำปรึกษา ความรู้ ความดูแลเอาใจใส่อย่างดียิ่งและหนังสือต่างๆ จาก ท่านอาจารย์ ดร.สายชล ใจเย็น ผู้เป็นอาจารย์ที่ปรึกษา ซึ่งท่านได้สละเวลาให้กับข้าพเจ้าอย่างเต็มที่ ข้าพเจ้ารู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงใคร่ขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ ผศ.ดร.ศรัณย์ อินทโกสุม และ ผศ.ดร. อนันตพร หรรษคุณาฒัย สำหรับคำแนะนำ และคำปรึกษาต่างๆ ในการสอบหัวข้อวิทยานิพนธ์ด้วยดีตลอดมา

ขอกราบขอบพระคุณ ผศ.ดร.ศุภกานต์ พิมลธเรศ ซึ่งให้ความกรุณามาเป็นกรรมการผู้ทรงคุณวุฒิจากภายนอก ทำให้ได้รับคำปรึกษา และคำแนะนำต่างๆ ในการสอบวิทยานิพนธ์เป็นอย่างดี

ขอกราบขอบพระคุณ ผศ.ดร.นันทิกา เบญจเทพานันท์ ซึ่งให้ความกรุณา ให้คำแนะนำปรึกษาด้านต่างๆ ในการทำวิทยานิพนธ์เป็นอย่างดี

ขอกราบขอบพระคุณบิดา มารดา พี่จักษุ พี่ณัฏฐาพร ที่ได้ให้การสนับสนุนในด้านต่างๆ และเป็นกำลังใจในระหว่างการศึกษาเล่าเรียนเป็นอย่างดียิ่ง

ขอขอบคุณน้องๆ ว่าที่ร้อยตรีเสฏฐนันท์ ทองสุวรรณ์ นายนิพัทธ์ คล้ายโพธิ์ นายวริษฐ์ และทุกๆ คน ที่ได้ช่วยกันฝ่าฟันปัญหาต่างๆ ด้วยกันตลอดระยะเวลาการทำวิทยานิพนธ์

ขอขอบคุณเจ้าหน้าที่ของคณะวิทยาศาสตร์ทุกท่านที่ให้การช่วยเหลือ อนุเคราะห์ ในการดำเนินการต่างๆ ให้สำเร็จลุล่วงไปด้วยดี

ขอกราบขอบพระคุณท่านอาจารย์ในภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ได้ให้ความรู้ทางวิชาการ หากการจัดทำวิทยานิพนธ์ครั้งนี้มีคุณประโยชน์ใดๆ ผู้จัดทำของมอบเป็นเครื่องสักการะคุณในคุณงามความดีของบิดา มารดา ครูบาอาจารย์ทุกท่าน

อนุสรณ์ เจริญนาน

กรกฎาคม 2558

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

บทคัดย่อภาษาไทย ..... ก

บทคัดย่อภาษาอังกฤษ ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฉ

สารบัญรูป ..... ช

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา 2

1.3 สมมติฐานของการศึกษา 2

1.4 ขอบเขตการวิจัย 2

1.5 ขั้นตอนการศึกษาและดำเนินการวิจัย 3

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4

2.1 การจำแนกประเภทข้อมูล 4

2.2 โครงข่ายประสาทเทียม 5

2.2.1 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า 7

2.2.2 โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอน 8

2.2.3 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนชั้นเดียว 9

2.2.4 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น 9

2.2.5 การเรียนรู้แบบแพร่กระจายย้อนกลับ 11

2.3 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน 13

2.4 ตัวจำแนกแบบเบย์อย่างง่าย 15

2.5 การเรียนรู้แบบรวมกลุ่ม 17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

2.6 อัลกอริทึมเอดาบูสต์ 19

2.7 เทคนิค *K*-Fold Cross-Validation 21

บทที่ 3 วิธีการดำเนินงานวิจัย 22

3.1 ข้อมูลที่ใช้ในการทดลอง 22

3.2 อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น 22

3.3 ขั้นตอนการสร้างโมเดลการเรียนรู้แบบรวมกลุ่ม 26

3.4 การคำนวณผลลัพธ์ของโมเดลการเรียนรู้แบบรวมกลุ่ม 27

3.5 การวัดประสิทธิภาพ 29

บทที่ 4 ผลการทดลอง 31

บทที่ 5 สรุปและข้อเสนอแนะ 35

5.1 สรุป 35

5.2 ข้อเสนอแนะ 36

เอกสารอ้างอิง 37

ผลงานที่ได้รับการตีพิมพ์ 39

ประวัติผู้เขียน 45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ 2.1 การเปรียบเทียบระหว่างเซลล์ประสาทในสมองมนุษย์และเซลล์ประสาทเทียม 7
ตารางที่ 3.1 คุณลักษณะของชุดข้อมูลที่ใช้ในการทดลอง 22
ตารางที่ 4.1 โมเดลการเรียนรู้แบบรวมกลุ่มในแต่ละรูปแบบที่ใช้ในการทดลองนี้ 31
ตารางที่ 4.2 ค่าความถูกต้องของโมเดลการเรียนรู้แบบเดี่ยว และโมเดลการเรียนรู้แบบรวมกลุ่มซึ่งมีการสลับที่ของตัวจำแนกประเภทต่างชนิดกัน 32
ตารางที่ 4.3 ค่าความถูกต้องของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้ตัวจำแนกประเภทแบบเดียวกันทั้งหมด และแบบที่ใช้ตัวจำแนกประเภทต่างชนิดกัน 33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ 2.1 เซลล์ประสาทในสมองมนุษย์ 5
รูปที่ 2.2 การทำงานของเซลล์ประสาทเทียม 6
รูปที่ 2.3 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า 8
รูปที่ 2.4 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนชั้นเดียว 9
รูปที่ 2.5 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น 10
รูปที่ 2.6 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน 13
รูปที่ 2.7 โมเดลการเรียนรู้แบบรวมกลุ่ม 18
รูปที่ 2.8 เทคนิค k-fold cross-validation 21
รูปที่ 3.1 อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น 24
รูปที่ 3.2 การเตรียมข้อมูลชุดฝึกสอนและข้อมูลทดสอบ 29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันข้อมูลมีขนาดใหญ่ขึ้นและมีความซับซ้อนมากขึ้นประกอบกับมีความต้องการในการสกัดความรู้เพื่อนำไปใช้ประโยชน์ในด้านต่างๆ เช่น การตลาด การตัดสินใจลงทุน การแพทย์ เป็นต้น จึงจำเป็นที่จะต้องคิดค้นและพัฒนาเทคนิควิธีการที่จะนำเอาองค์ความรู้จากข้อมูลเหล่านั้นไปใช้ประโยชน์ โดยความรู้ที่ได้มานั้นมีความถูกต้องมากที่สุดและใช้เวลาน้อยที่สุด จึงมีการนำเอาความรู้ทางด้านโครงข่ายประสาทเทียม ตัวจำแนกแบบเบย์อย่างง่าย มาใช้เป็นเครื่องมือที่ทรงประสิทธิภาพในการสกัดองค์ความรู้เหล่านั้น เนื่องจากข้อมูลในปัจจุบันมีหลากหลายรูปแบบจึงทำให้เกิดปัญหาในการจำแนกประเภทข้อมูลหรือการวิเคราะห์ข้อมูล บางครั้งทำให้เกิดความล่าช้าหรือเกิดความผิดพลาดสูง และข้อมูลบางรูปแบบก็จะเหมาะกับบางวิธีการ จึงเป็นสาเหตุให้ผู้วิจัยคิดค้นเทคนิคใหม่ขึ้นมาเพื่อเพิ่มประสิทธิภาพความถูกต้อง โดยแนวคิดของวิธีการใหม่นี้จะใช้โครงข่ายประสาทเทียม ตัวจำแนกแบบเบย์อย่างง่าย และอัลกอริทึมเอดาบูสต์มาประยุกต์ใช้ในการสร้างโมเดลการเรียนรู้แบบรวมกลุ่มแบบใหม่ที่เพิ่มความถูกต้องแม่นยำในการจำแนกประเภทข้อมูลได้ดียิ่งขึ้น โดยเทคนิคการจำแนกประเภทข้อมูลนี้สามารถถูกนำไปประยุกต์ใช้ในหลายๆ ด้าน เช่น การจัดกลุ่มลูกค้าเพื่อผลลัพธ์ทางการตลาดที่ดีขึ้น การตรวจสอบขบวนการอนุมัติเงินกู้ของธนาคาร การจำแนกชนิดของสิ่งมีชีวิต และการวินิจฉัยโรคทางการแพทย์ เป็นต้น

เทคนิคการจำแนกประเภทข้อมูลที่เรานิยมใช้กันในปัจจุบันมีหลายวิธี ได้แก่ ตัวจำแนกแบบเบย์อย่างง่าย (Naive Bayes Classifier) ซึ่งเป็นโมเดลการจำแนกประเภทที่อยู่บนพื้นฐานของทฤษฎีบทของเบย์ (Bayes' Theorem) และสมมติฐานของการเกิดเหตุการณ์ต่างๆ ที่เป็นอิสระต่อกัน (Independence) โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial Basis Function) และ โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multilayer Layer Perceptron) โดยแต่ละตัวจำแนกประเภทเหล่านี้ก็จะเหมาะกับข้อมูลบางประเภทแตกต่างกันไป จึงส่งผลให้การทำงานกับข้อมูลบางประเภทมีค่าความถูกต้องต่ำ ดังนั้นงานวิจัยนี้จึงมุ่งเน้นไปที่การเพิ่มประสิทธิภาพของการจำแนกประเภทข้อมูลโดยการนำวิธีการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำแนกประเภทข้อมูลหลายๆ แบบมาทำงานร่วมกันแบบรวมกลุ่ม [1] เพื่อเพิ่มประสิทธิภาพความถูกต้องให้สูงยิ่งขึ้น

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

เพื่อพัฒนาตัวจำแนกประเภทแบบรวมกลุ่มแบบใหม่โดยใช้ตัวจำแนกประเภทแบบที่ใช้การส่งเชิงเส้น การส่งไม่เชิงเส้น และทฤษฎีความน่าจะเป็น สำหรับปัญหาการจำแนกประเภทข้อมูล

1.3 สมมติฐานของการศึกษา

ข้อมูลในปัจจุบันมีความหลากหลายมาก และอัลกอริทึมที่ใช้ในการจำแนกประเภทข้อมูล (Data Classification) ก็มีความหลากหลายเช่นเดียวกัน อัลกอริทึมรูปแบบหนึ่งก็อาจจะทำงานได้ดีกับข้อมูลบางประเภท แต่อาจจะให้ความถูกต้องไม่ดีกับข้อมูลบางประเภท อัลกอริทึมรูปแบบหนึ่งก็อาจจะเหมาะกับข้อมูลอีกรูปแบบหนึ่ง หากเราสร้างอัลกอริทึมขึ้นมาใหม่โดยใช้อัลกอริทึมพื้นฐานที่นิยมใช้งานมาทำงานร่วมกัน จะทำให้สามารถประยุกต์ใช้กับข้อมูลได้หลากหลายมากขึ้น และให้ค่าความถูกต้องที่ดีขึ้นกว่าเดิมได้

1.4 ขอบเขตการวิจัย

ในการทำวิทยานิพนธ์นี้มีขอบเขตของการวิจัยดังนี้

1. งานวิจัยนี้เสนอวิธีการใหม่ในการจำแนกประเภทข้อมูล โดยใช้โครงข่ายประสาทเทียมได้แก่ โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน ตัวจำแนกแบบเบย์อย่างง่าย และอัลกอริทึมเอดาบูสต์ (Adaboost) [1] โดยที่ตัวจำแนกประเภทเหล่านี้จะทำงานร่วมกันเป็นการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) และเปรียบเทียบประสิทธิภาพกับตัวจำแนกประเภทแบบเดี่ยว (Single Classifier)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การทดสอบประสิทธิภาพจะทำกับชุดข้อมูล (Dataset) ที่มีความหลากหลาย ทั้งที่เป็นแบบสองคลาส (Binary Class) และแบบหลายคลาส (Multi Class) โดยทำการนำชุดข้อมูลมาจัดรูปแบบและนำมาทดสอบกับอัลกอริทึมแบบเดิม จากนั้นจะบันทึกค่าความถูกต้องไว้ เมื่อทำการทดสอบขั้นแรกเสร็จก็นำข้อมูลชุดเดียวกันมาทดสอบกับอัลกอริทึมแบบใหม่ตามที่ได้ตั้งสมมติฐานไว้ จากนั้นบันทึกผลการทดลอง โดยการทดลองทั้งสองแบบจะทำการปรับค่าพารามิเตอร์ให้เหมือนกัน โดยจะทำการทดสอบซ้ำเป็นจำนวนรอบที่เท่ากันแล้วเอาค่าความถูกต้องที่ดีที่สุดมาเปรียบเทียบกัน

1.5 ขั้นตอนการศึกษาและดำเนินการวิจัย

วิทยานิพนธ์นี้มีขั้นตอนการดำเนินงานวิจัยดังนี้

1. ศึกษาการทำงานของอัลกอริทึมแบบต่างๆที่ใช้ในการจำแนกประเภทข้อมูล (Data Classification) หลายๆ รูปแบบว่ามีข้อดีข้อเสียอย่างไร รวมถึงการใช้วิธีการของโครงข่ายประสาทเทียมที่เกี่ยวข้องกับการจำแนกประเภทข้อมูล

2. เลือกชุดข้อมูลมาทดสอบกับอัลกอริทึม โดยศึกษาถึงความเหมาะสมกับข้อมูลที่ได้ทดสอบแล้วบันทึกผลการทดลอง

3. ศึกษาการออกแบบอัลกอริทึมแบบใหม่ โดยใช้อัลกอริทึมหลายๆ รูปแบบ

4. นำชุดข้อมูลที่ได้เตรียมไว้มาทำการทดสอบกับอัลกอริทึมแบบใหม่นี้ จากนั้นบันทึกค่าความถูกต้องของแต่ละชุดข้อมูล

5. ประเมินผลการทดลอง และสรุปผล

6. เขียนวิทยานิพนธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและความรู้ที่เกี่ยวข้อง วิธีการแก้ปัญหาแบบต่างๆ เนื้อหาที่นำเสนอจะประกอบไปด้วย การจำแนกประเภทข้อมูล โครงข่ายประสาทเทียม และอัลกอริทึมที่ใช้ในการตัดสินใจ

2.1 การจำแนกประเภทข้อมูล

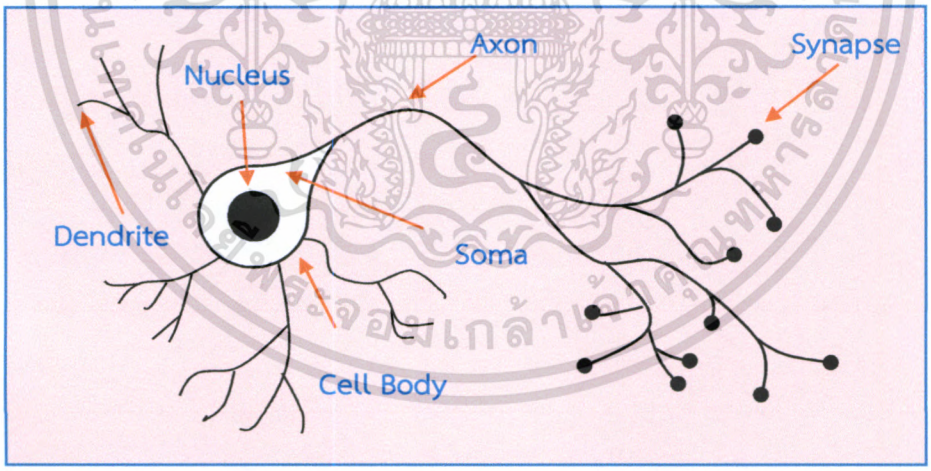
การจำแนกประเภทข้อมูล (Classification) คือ กระบวนการสร้างโมเดลจำแนกประเภทข้อมูล (Data Classification Model) เพื่อทำนายกลุ่มของข้อมูลใหม่อย่างมีหลักการ ตัวอย่างของกลุ่มข้อมูล เช่น กลุ่มของลูกค้าชั้นดีที่ใช้บัตรเครดิต กลุ่มของสินค้ายอดขายสูง ปานกลาง หรือต่ำ กลุ่มของชาวนาที่รายได้สูง หรือต่ำ ข้อมูลของลักษณะเฉพาะของโรคทางการแพทย์ เป็นต้น ในที่นี้คำว่ากลุ่มจะเรียกว่าคลาส (class) โดยในคลาสเดียวกันนั้นจะต้องมีข้อมูลที่มีความเหมือนหรือคล้ายคลึงกันมากกว่าข้อมูลที่อยู่ต่างคลาสกัน การสร้างโมเดลจำแนกประเภทข้อมูลจะเกิดมาจากการหาความสัมพันธ์ของข้อมูลในฐานข้อมูล โดยข้อมูลทั้งหมดจะมีการแบ่งออกเป็น 2 ชุด คือ ชุดข้อมูลฝึกสอน (training set) ซึ่งเป็นชุดข้อมูลที่มีหน้าที่ในการสร้างโมเดลเพื่อจำแนกประเภทข้อมูลออกมา และชุดข้อมูลทดสอบ (test set) เป็นชุดข้อมูลที่ใช้ประเมินความถูกต้องของโมเดลจำแนกประเภทข้อมูล

โมเดลจำแนกประเภทข้อมูลได้ถูกนำมาประยุกต์ใช้งานหลายๆ ด้าน ไม่ว่าจะเป็นการวิเคราะห์หุ้นเพื่อหาว่าหุ้นแต่ละบริษัทมีคุณภาพเป็นอย่างไรเมื่อมีปัจจัยที่เกี่ยวข้อง ไม่ว่าจะเป็นการเติบโตของรายได้ ความสามารถในการควบคุมต้นทุน ความผันผวนของรายได้และกำไร การพยากรณ์อากาศ การจัดสรรกฎหมายที่เหมาะสมในการพิจารณาคดีความ การจัดการความสัมพันธ์ของลูกค้า การทำนายลักษณะของโรคทางการแพทย์ และอื่นๆ [2] เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 โครงข่ายประสาทเทียม

ในปัจจุบันระบบประสาทของมนุษย์ได้ถูกนำมาจำลองและพัฒนาเป็นโครงสร้างของโครงข่ายประสาทเทียม (ANN : Artificial Neural Network) โดยระบบนี้คอมพิวเตอร์สามารถทำงานเลียนแบบการทำงานของสมองมนุษย์ได้ ช่วยให้คอมพิวเตอร์สามารถจดจำและเรียนรู้จากประสบการณ์ได้มากขึ้น ส่งผลให้คอมพิวเตอร์สามารถเชื่อมโยงข้อเท็จจริงจำนวนมากเข้าด้วยกันเพื่อหาข้อสรุปและดำเนินการปรับปรุงองค์ความรู้ที่มีอยู่ได้อย่างมีประสิทธิภาพ การทำงานของโครงข่ายประสาทเทียมเป็นการทำงานที่เลียนแบบการทำงานของระบบประสาทในสมองมนุษย์ซึ่งประกอบไปด้วยเซลล์ประสาท (Neural) จำนวนมากเชื่อมโยงกันเป็นโครงข่าย โดยเซลล์ประสาทเหล่านี้อยู่เรียงกันเป็นกลุ่ม หากมีชนิดเดียวกันจะอยู่ในกลุ่มเดียวกัน โดยแต่ละกลุ่มจะเชื่อมโยงเซลล์ประสาทจำนวนมากเข้าด้วยกันเป็นโครงข่าย โครงข่ายการทำงานแบบนี้เป็นจุดเริ่มต้นหรือเป็นต้นแบบในการสร้างระบบประสาทเสมือนให้กับคอมพิวเตอร์ โดยกำหนดให้หนึ่งหน่วยของเซลล์ประสาทเรียกว่า โหนด (Node) โดยหลายๆ โหนดจะเชื่อมโยงกันเป็นโครงข่ายคล้ายกับกลุ่มเซลล์ประสาทในสมองมนุษย์ที่เชื่อมโยงกันเป็นโครงข่าย โหนดต่างๆ ในโครงข่ายจะถูกแบ่งออกเป็นชั้น (Layer) และแต่ละชั้นจะมีหน้าที่แตกต่างกัน โครงสร้างของเซลล์ประสาทในสมองของมนุษย์ประกอบด้วยส่วนประกอบหลายส่วน ดังแสดงในรูปที่ 2.1

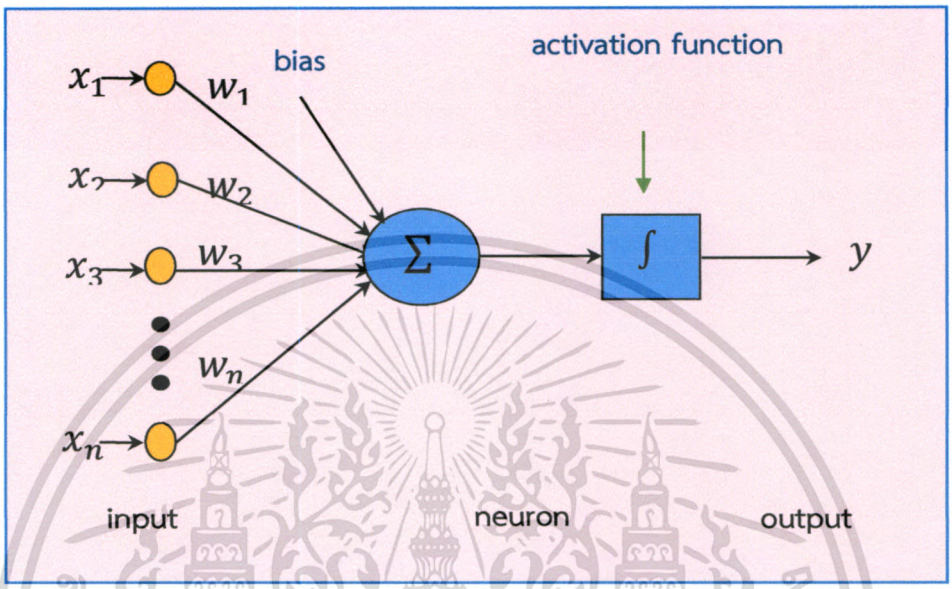


รูปที่ 2.1 เซลล์ประสาทในสมองมนุษย์

โครงข่ายประสาทเทียมมีคุณสมบัติคล้ายกับการส่งผ่านสัญญาณประสาทในสมองมนุษย์ โดยมีความสามารถในการรวบรวมความรู้ (Knowledge) โดยผ่านกระบวนการเรียนรู้ (Learning Process) และความรู้เหล่านี้จะจัดเก็บอยู่ในโครงข่ายในรูปแบบค่าน้ำหนัก (Weight) ซึ่งสามารถปรับเปลี่ยนค่าได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อมีการเรียนรู้สิ่งใหม่เข้าไป ค่าน้ำหนักทำหน้าที่เปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ปัญหาเฉพาะอย่างของมนุษย์ [3] ดังแสดงในรูปที่ 2.2



รูปที่ 2.2 การทำงานของเซลล์ประสาท

การทำงานของโครงข่ายประสาทเทียม (Artificial Neural Network) เป็นการจำลองการทำงานของโครงข่ายใยประสาทในสมองมนุษย์ (Biological Neural Network) โดยประกอบด้วยส่วนต่างๆ ดังนี้

1. ข้อมูลนำเข้า (Input) คือ ชุดของข้อมูลที่ต้องการประมวลผลโดยจะถูกจำแนกตามลักษณะของข้อมูล (Attribute) โดยข้อมูลเหล่านี้อาจจะอยู่ในรูปแบบข้อมูลตัวเลข ตัวอักษร ภาพ หรือเสียง โดยจะถูกแปลงให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถนำไปประมวลผลได้
2. ค่าน้ำหนัก (Weight) คือ ค่าความสัมพันธ์ระหว่างข้อมูลนำเข้ากับข้อมูลส่งออกซึ่งเป็นส่วนประกอบที่มีความสำคัญมากในโครงข่ายประสาทเทียมเนื่องจากเป็นค่าเฉพาะที่กำหนดน้ำหนักให้กับข้อมูลนำเข้าแต่ละตัวเพื่อใช้ในการจำแนกความแตกต่างของข้อมูลนำเข้าเหล่านั้น ค่าน้ำหนักที่แตกต่างกันจะสามารถแสดงระดับความสำคัญของข้อมูลนำเข้าที่แตกต่างกัน จึงทำให้สามารถจำแนกประเภทข้อมูลที่ตรงกับลักษณะของปัญหาได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ฟังก์ชันกระตุ้น (Activation Function) คือ ฟังก์ชันที่ใช้สำหรับบีบค่าที่กำหนดให้อยู่ในช่วงที่เราต้องการ อย่างเช่น ฟังก์ชัน Linear และ Sigmoid เป็นต้น

การเรียนรู้ของโครงข่ายประสาทเทียมสามารถแบ่งได้เป็น 3 ประเภท คือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และการเรียนรู้แบบเสริมแรง (Reinforcement Learning)

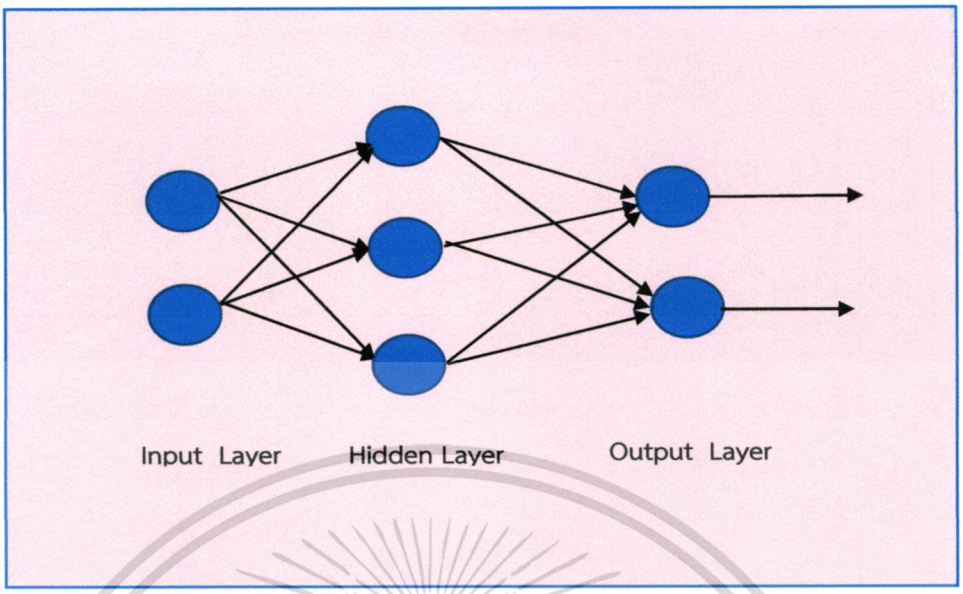
ตารางที่ 2.1 การเปรียบเทียบระหว่างเซลล์ประสาทในสมองมนุษย์และเซลล์ประสาทเทียม

Biological Neural Network	Neural Network
Soma	Neuron
Dendrite	Input
Axon	Output
Synapse	Weight

2.2.1 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า

โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า (Feed Forward Neural Network) คือ โครงข่ายประสาทเทียมที่จะดำเนินการส่งข้อมูลไปในทิศทางเดียว จากข้อมูลนำเข้า ไปยังข้อมูลส่งออก โดยไม่ย้อนกลับไปทำงานในลักษณะวนซ้ำ ทำให้ข้อมูลต่างๆ ไม่มีการนำมาคำนวณซ้ำอีก ดังแสดงในรูปที่ 2.3 เมื่อผ่านการทำงานของส่วนใดมาแล้วจะต้องดำเนินการต่อไปจนได้ผลลัพธ์ในที่สุด โครงข่ายประสาทเทียมลักษณะนี้ ได้แก่ เพอร์เซ็ปตรอน (Perceptron) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 โครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า

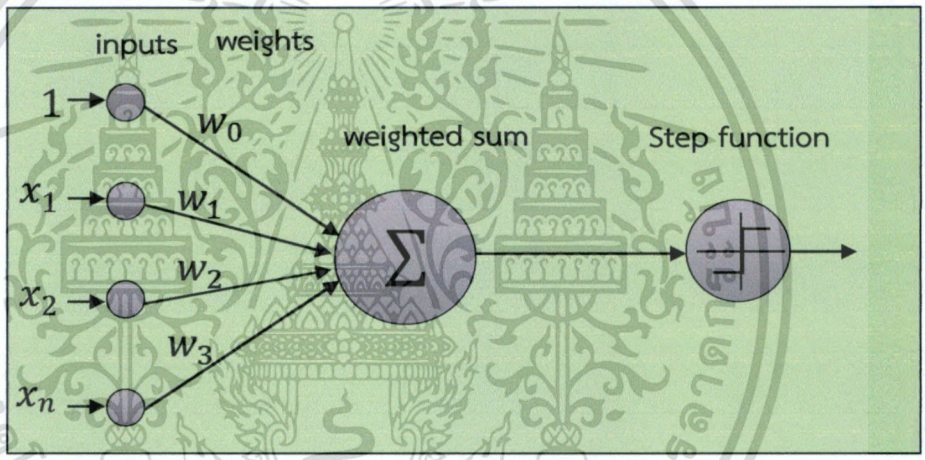
2.2.2 โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอน

โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอน (Perceptron) [4] เป็นโครงข่ายประสาทเทียมที่มีลักษณะการคำนวณแบบแพร่ไปข้างหน้า ซึ่งแต่ละส่วนจะเชื่อมโยงกันด้วยค่าน้ำหนักมากกว่าหนึ่งค่า นั่นคือ เป็นโครงข่ายที่มีการเชื่อมโยงกันมากกว่าหนึ่งจุด โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนจะประกอบด้วยเซลล์ประสาทเทียมหรือที่เรียกว่า นิวรอน (Neuron) อย่างน้อยหนึ่งชั้น โดยข้อมูลนำเข้าจะถูกส่งตรงไปยังชั้นข้อมูลส่งออกโดยผ่านค่าน้ำหนัก เนื่องจากการทำงานในลักษณะแบบนี้จึงสามารถพิจารณาว่าเป็นโครงข่ายแบบแพร่ไปข้างหน้า (Feed Forward Neural Network) โดยนิวรอนแต่ละตัวจะคำนวณผลรวมค่าถ่วงน้ำหนักของข้อมูลนำเข้า นั่นคือผลรวมของผลคูณระหว่างค่าข้อมูลนำเข้ากับค่าน้ำหนักของข้อมูลนำเข้าแต่ละตัว ถ้าค่าที่ได้จากการคำนวณนี้มากกว่าค่าคงที่ที่กำหนด (Threshold) นิวรอนก็จะให้ค่าผลลัพธ์เท่ากับ 1 แต่ถ้ามีค่าน้อยกว่าก็จะให้ค่าผลลัพธ์เท่ากับ -1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนชั้นเดียว

โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนชั้นเดียว (Single Layer Perceptron) เป็นโครงข่ายประสาทเทียมที่มีโครงสร้างไม่ซับซ้อน ถือเป็นโครงสร้างขั้นพื้นฐานและง่ายต่อการคำนวณ โดยประกอบไปด้วยชั้นข้อมูลนำเข้า และชั้นข้อมูลส่งออก ความเข้มของสัญญาณหรือปริมาณข้อมูลที่นำเข้าสู่โหนดในชั้นข้อมูลส่งออกจะขึ้นอยู่กับค่าน้ำหนักที่อยู่บนเส้นเชื่อมโยง โดยโหนดในชั้นข้อมูลส่งออกจะทำการคำนวณผลรวมเชิงเส้นของผลคูณระหว่างข้อมูลนำเข้าและค่าน้ำหนักของแต่ละจุดที่เชื่อมโยงกันแล้วใช้ฟังก์ชันทางคณิตศาสตร์ที่เรียกว่า ฟังก์ชันกระตุ้น (Activation Function) ในการบีบค่าผลรวมเชิงเส้นให้อยู่ในช่วงที่ต้องการแล้วส่งผลลัพธ์ที่ได้ออกมาเป็นค่าผลลัพธ์ของชั้นข้อมูลส่งออก ดังรูปที่ 2.4



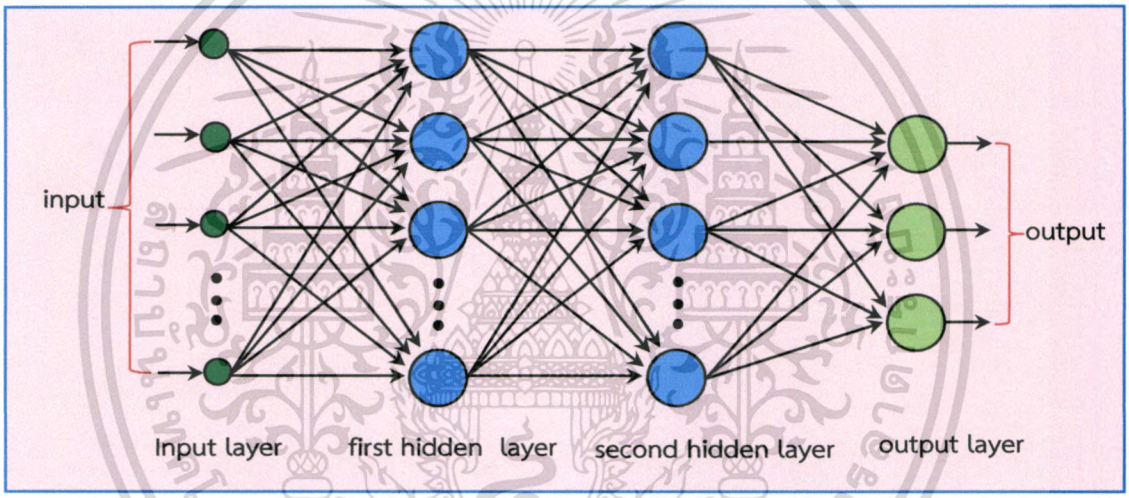
รูปที่ 2.4 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนชั้นเดียว

2.2.4 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น

โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multi Layer Perceptron) [4,5] เป็นโครงข่ายประสาทเทียมที่มีมากกว่าหนึ่งชั้น การทำงานของโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น จะมีโครงสร้างซ้อนกันหลายๆชั้น ใช้สำหรับงานที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ โดยมีกระบวนการเรียนรู้เป็นแบบมีผู้สอน และใช้อัลกอริทึมแบบแพร่กระจายย้อนกลับ (Backpropagation Algorithm) สำหรับการฝึกสอน โดยการทำงานจะประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Forward Pass) และการส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้า ข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลนำเข้า และจะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลส่งออก ส่วนการส่งผ่านย้อนกลับ ค่าน้ำหนักของเส้นเชื่อมโยงจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการเรียนรู้แบบแก้ไขข้อผิดพลาด (Error-Correction Learning Rule) โดยใช้ผลต่างของผลลัพธ์ที่แท้จริง (Actual Response) กับผลลัพธ์เป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด (Error Signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลลัพธ์ของโครงข่ายเข้าใกล้กับผลลัพธ์เป้าหมาย



รูปที่ 2.5 โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น

โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้นประกอบด้วย ชั้นข้อมูลนำเข้า (Input Layer) ชั้นข้อมูลส่งออก (Output Layer) และชั้นซ่อน (Hidden Layer) โดยชั้นซ่อนอาจจะมีหลายชั้นก็ได้ ในแต่ละชั้นของชั้นซ่อน (hidden layer) จะมีฟังก์ชันสำหรับคำนวณเมื่อได้รับสัญญาณเอาต์พุตจากโหนดในชั้นก่อนหน้านี้ เรียกว่าฟังก์ชันกระตุ้น โดยในแต่ละชั้นไม่จำเป็นต้องเป็นฟังก์ชันเดียวกันก็ได้ ชั้นซ่อนนั้นมีหน้าที่สำคัญคือ แปลงข้อมูลที่เข้ามาในชั้นปัจจุบัน ให้สามารถแยกแยะความแตกต่างแบบเชิง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เส้นได้ (Linearly Separable) และก่อนที่ข้อมูลจะถูกส่งไปถึงชั้นข้อมูลส่งออก ในบางครั้งอาจจำเป็นต้องใช้ชั้นซ่อนมากกว่า 1 ชั้นในการแปลงข้อมูลให้อยู่ในรูปแบบเชิงเส้น

2.2.5 การเรียนรู้แบบแพร่กระจายย้อนกลับ

การเรียนรู้แบบแพร่กระจายย้อนกลับ (Backpropagation Learning) [1] มีลักษณะการทำงานประกอบด้วยการทำงานสองส่วนย่อย คือ การส่งผ่านไปข้างหน้า (Feed Forward) และ การส่งผ่านย้อนกลับ (Backward Pass) กระบวนการทำงานของการเรียนรู้แบบแพร่กระจายย้อนกลับ [1] สามารถอธิบายได้ดังนี้

การทำงานแบบส่งผ่านไปข้างหน้า

1. ทำการสุ่มค่าน้ำหนัก (weights) และค่าไบแอส (bias)
2. รับข้อมูลจากชุดของข้อมูลฝึกสอน $D = \{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ โดยที่ $\mathbf{x}_i$ คือ ข้อมูลนำเข้า และ d_i คือ ผลลัพธ์ที่ต้องการ
3. ทำการคำนวณค่าผลลัพธ์ (Output) ของแต่ละโหนดในชั้นซ่อนโดยใช้สมการ 2.1

$$y_j = \varphi_j(v_j) \tag{2.1}$$

โดยที่ $\varphi_j(v_j) = \frac{1}{1+e^{-v_j}}$

$v_j = \sum_{i=1}^{N} x_i w_{ij}$

w_{ij} คือ ค่าน้ำหนักของเส้นเชื่อมจากโหนด i ไปยังโหนด j

4. ทำการคำนวณค่าผลลัพธ์ (Output) ของแต่ละโหนดในชั้นข้อมูลส่งออก โดยใช้สมการ 2.2

$$y_k = \varphi_k(v_k) \tag{2.2}$$

โดยที่ $v_k = \sum_{j=1}^{N} y_j w_{jk}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานแบบส่งผ่านย้อนกลับ

5. ทำการคำนวณค่า Error Gradient ของแต่ละโหนดข้อมูลในชั้นข้อมูลส่งออกโดยใช้สมการ 2.3

$$\delta_k = (d_k - y_k)y_k(1 - y_k) \quad (2.3)$$

โดยที่ y_k คือ ค่าผลลัพธ์ของโหนดที่ k ในชั้นข้อมูลส่งออก

6. ทำการคำนวณค่า Error Gradient ของแต่ละโหนดในชั้นซ่อนโดยใช้สมการ 2.4

$$\delta_j = y_j(1 - y_i)\sum_k \delta_k w_{jk} \quad (2.4)$$

โดยที่ y_j คือ ค่าผลลัพธ์ของโหนดที่ j ในชั้นซ่อน

7. ทำการปรับค่าน้ำหนักของทุกเส้นเชื่อมโดยใช้สมการที่ 2.5

$$w_{ij} = w_{ij} + \Delta w_{ij} \quad (2.5)$$

โดยที่ $\Delta w_{ij} = \eta\delta_j x_i$

η คือ อัตราการเรียนรู้ (learning rate)

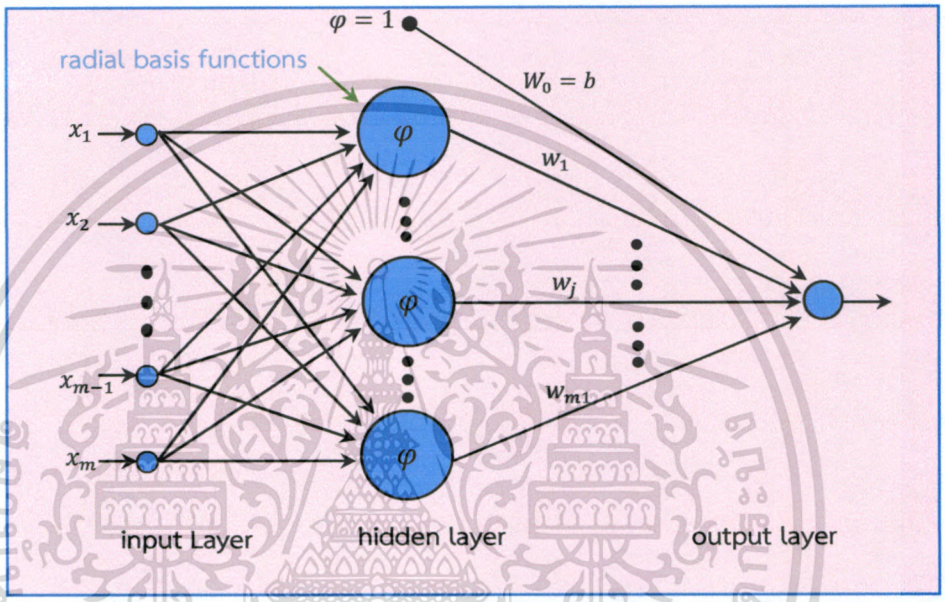
การปรับเปลี่ยนค่าน้ำหนักจะเริ่มจากชั้นข้อมูลส่งออกแพร่ย้อนกลับไปจนถึงชั้นข้อมูลนำเข้า ทำซ้ำขั้นตอนที่ 2 ถึง 7 จนกระทั่งโครงข่ายประสาทเทียมสามารถจำแนกข้อมูลตัวอย่างได้อย่างถูกต้อง

การเรียนรู้แบบแพร่กระจายย้อนกลับเป็นวิธีการหนึ่งของโครงข่ายประสาทเทียมที่ง่ายต่อการเข้าใจ เนื่องจากขบวนการเรียนรู้และปรับปรุงแก้ไขนั้นเป็นไปโดยอัตโนมัติ หากโครงข่ายประสาทเทียมให้คำตอบที่ผิด ค่าน้ำหนักจะถูกปรับจนกว่าค่าความผิดพลาดจะน้อยลงหรืออยู่ในเกณฑ์ที่ยอมรับได้ นั่นคือ ค่าที่ได้ในครั้งถัดไปจะมีความถูกต้องมากยิ่งขึ้น ด้วยโครงสร้างที่มีลักษณะเป็นชั้น แต่ละชั้นเชื่อมโยงถึงกันอย่างทั่วถึง เมื่อโครงข่ายประสาทเทียมได้รับข้อมูลนำเข้า จะคำนวณค่าน้ำหนักของหน่วยรับข้อมูลนำเข้าไปยังชั้นซ่อน และจากชั้นซ่อนไปยังชั้นส่งข้อมูลออก เมื่อเกิดผลต่างระหว่างค่าผลลัพธ์จริงกับค่าผลลัพธ์เป้าหมาย โครงข่ายประสาทเทียมจะปรับค่าความผิดพลาดจากชั้นข้อมูลส่งออก และแพร่ย้อนกลับไปยังชั้นซ่อนจากนั้นจึงแพร่ย้อนกลับไปยังชั้นรับข้อมูลนำเข้าตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน

โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial Basis Function) [6,7,8,9] เป็นโครงข่ายประสาทเทียมแบบแพร่ไปข้างหน้า (Feed Forward Networks) ประกอบด้วยชั้นข้อมูลนำเข้า (Input Layer) ชั้นข้อมูลส่งออก (Output Layer) และชั้นซ่อน (Hidden Layer) โดยที่แต่ละโหนดในชั้นซ่อนจะประกอบด้วย เรเดียลเบสิสฟังก์ชัน (Radial Basis Function)



รูปที่ 2.6 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน

โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน จะใช้เรเดียลเบสิสฟังก์ชันในชั้นซ่อนและในส่วนของชั้นข้อมูลส่งออกจะเป็นการรวมค่าแบบเชิงเส้นจากชั้นซ่อน ในส่วนของข้อมูลนำเข้าที่เข้าสู่โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันอาจจะเป็นแบบไม่เชิงเส้น แต่ในส่วนของผลลัพธ์จะเป็นลักษณะเชิงเส้น เมื่อต้องการที่จะใช้งานโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันจะต้องมีการกำหนดชนิดของเรเดียลเบสิสฟังก์ชันที่จะใช้ และกำหนดจำนวนของโหนดในชั้นซ่อนที่จะทำการประมวลผล และกำหนดอัลกอริทึมที่ใช้ในการเรียนรู้ (Training Algorithm) โดยการเรียนรู้ของโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันจะเป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) จึงต้องมีชุดข้อมูลฝึกสอน (Training Set) พร้อมกับผลลัพธ์ชุดหนึ่งที่จะใช้สอนให้โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันเรียนรู้ และผลของการเรียนรู้ของโครงข่ายประสาทเทียมนี้ คือ การหาค่าน้ำหนัก (Weight) ของโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เรเดียลเบสิสฟังก์ชันที่จะใช้ในการจำแนกประเภทข้อมูลที่เป็นส่วนของชุดข้อมูลทดสอบ (Testing Set) โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน มีขั้นตอนการเรียนรู้ ดังนี้

1. รับข้อมูลฝึกสอน $D = \{(\mathbf{x}_i, d_i)\}_{i=1}^{N}$ โดยที่ $\mathbf{x}_i$ คือ ข้อมูลนำเข้า และ d_i คือ ผลลัพธ์ที่ต้องการ
2. คำนวณค่าผลลัพธ์ในชั้นซ่อนโดยใช้ สมการที่ 2.6

$$F(\mathbf{x}) = \sum_{i=1}^{N} w_i \varphi(\| \mathbf{x} - \mathbf{x}_i \|) \tag{2.6}$$

โดยที่ $\varphi(r) = EXP\left(-\frac{r^2}{2\sigma^2}\right)$

$\|.\|$ คือ นอร์มแบบยุคลิด (Euclidean Norm)

$\mathbf{x}_i$ คือ จุดศูนย์กลางของเรเดียลเบสิสฟังก์ชัน (Radial Basis Function)

จากสมการที่ 2.6 จะได้เมทริกซ์ของข้อมูลที่ผ่าน Gaussian Function ดังสมการที่ 2.7

$$\Phi = [\varphi_{ij}]_{N \times N} \tag{2.7}$$

เรียกเมทริกซ์ Φ ว่า Interpolation Matrix ที่มีประกอบด้วยสมาชิก φ_{ij} โดยเมทริกซ์นี้จะถูกนำมาคูณกับเวกเตอร์น้ำหนักเพื่อหาคำตอบดังสมการที่ 2.8

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_N \end{bmatrix} \tag{2.8}$$

โดยที่ $\varphi_{ji} = \varphi(\| \mathbf{x}_i - \mathbf{x}_j \|), (i,j) = 1, 2, \ldots N$ (2.9)

3. คำนวนหาค่าน้ำหนักของแต่ละเส้นเชื่อมระหว่างชั้นซ่อนและชั้นข้อมูลส่งออก โดยใช้สมการที่ 2.9 ซึ่งเราสามารถเขียนให้อยู่ในรูปแบบ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Phi \mathbf{w} = \mathbf{d} \tag{2.10}$$

โดยที่ $\mathbf{d} = [d_1, d_2, \ldots, d_N]^T$ เป็นเวกเตอร์ผลลัพธ์ที่ต้องการ

$\mathbf{w} = [w_1, w_2, \ldots, w_N]^T$ เป็นเวกเตอร์น้ำหนัก

N คือ ขนาดของชุดข้อมูลฝึกสอน

ดังนั้น เวกเตอร์น้ำหนัก $\mathbf{w}$ สามารถคำนวณได้โดยใช้สมการที่ 2.11

$$\mathbf{w} = \Phi^{+}\mathbf{d} \tag{2.11}$$

โดยที่ Φ^{+} คือ Pseudo-inverse ของเมทริกซ์ Φ

ผลลัพธ์ของโครงข่ายประสาทเทียมแรเดียลเบสิกฟังก์ชันสามารถคำนวณได้จากสมการที่ 2.12

$$F(\mathbf{x}) = \sum_{i=1}^{N} w_i \varphi(\| \mathbf{x} - \mathbf{x}_i \|) \tag{2.12}$$

2.4 ตัวจำแนกแบบเบย์อย่างง่าย

ตัวจำแนกแบบเบย์อย่างง่าย (Naïve Bayes Classifier) [3,10] เป็นเทคนิคในการจำแนกประเภทข้อมูลที่ได้รับความนิยมสูงวิธีการหนึ่ง โดยอาศัยหลักการทางคณิตศาสตร์และทฤษฎีทางสถิติมาใช้ในการสร้างทฤษฎีบทของเบย์ (Theorem bayes) โดยมีหลักการดังนี้

$$P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A) \tag{2.13}$$

จากสมการ 2.13 สามารถนำมาสร้างทฤษฎีบทของเบย์ ได้ดังนี้

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \tag{2.14}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากทฤษฎีบทของเบย์หากนำมาใช้ในงานการจำแนกประเภทข้อมูลสามารถทำให้เข้าใจได้ง่ายขึ้น โดยการเปลี่ยนสัญลักษณ์ A และ B ให้เป็น A และ C ตามลำดับ โดยที่ A คือ แอตทริบิวต์ (Attribute) และ C คือ คลาส (Class) ดังสมการที่ 2.15

$$P(C|A) = \frac{P(A|C)P(C)}{P(A)} \quad (2.15)$$

โดยที่

P(C|A) คือ ความน่าจะเป็นของ C เมื่อกำหนด A

P(A|C) คือ ความน่าจะเป็นของ A เมื่อกำหนด C

P(C) คือ ค่าความน่าจะเป็นของคลาส C

จากหลักการที่แต่ละแอตทริบิวต์เป็นอิสระต่อกัน หรือไม่ขึ้นต่อกัน (Independent) ทำให้สามารถเปลี่ยนสมการ P(A|C) ได้เป็น

$$P(A|C) = P(A_1|C) \times P(A_2|C) \times \dots \times P(A_n|C) \quad (2.16)$$

การประยุกต์ใช้ทฤษฎีแบบเบย์สามารถปรับให้สอดคล้องกับการจำแนกประเภทข้อมูลได้ดังสมการที่ 2.17

$$P(Class|Attribute) = \frac{P(Class)P(Attribute|Class)}{P(Attribute)} \quad (2.17)$$

โดยที่ *P(Class|Attribute)* คือ ความน่าจะเป็นในการเกิดคลาส (Class) เมื่อมีคุณลักษณะ (Attribute) เกิดขึ้น

P(Class) คือ ความน่าจะเป็นในการเกิดคลาส

P(Attribute) คือ ความน่าจะเป็นในการเกิดคุณลักษณะ

จากสมการที่ 2.17 เมื่อข้อมูลแต่ละตัวมีจำนวนคุณลักษณะมากกว่าหนึ่งตัวจะได้สมการ ดังนี้

$$P(Class|A_1, A_2, \dots, A_n) = \frac{P(Class)P(A_1, A_2, \dots, A_n|Class)}{P(A_1, A_2, \dots, A_n)} \quad (2.18)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถ้าการเกิดคุณลักษณะ A_i ใดๆ เป็นอิสระต่อกันสามารถคำนวณ $P(Class|A_1, A_2, \dots, A_n)$ ได้ ดังสมการที่ 2.19

$$P(A_1, A_2, \dots, A_n|Class) = \prod_{i=1}^{n} P(A_i|Class) \quad (2.19)$$

โดยที่ $$\prod_{i=1}^{n} P(A_i|Class) = P(A_1|Class)P(A_2|Class) \dots P(A_n|Class) \quad (2.20)$$

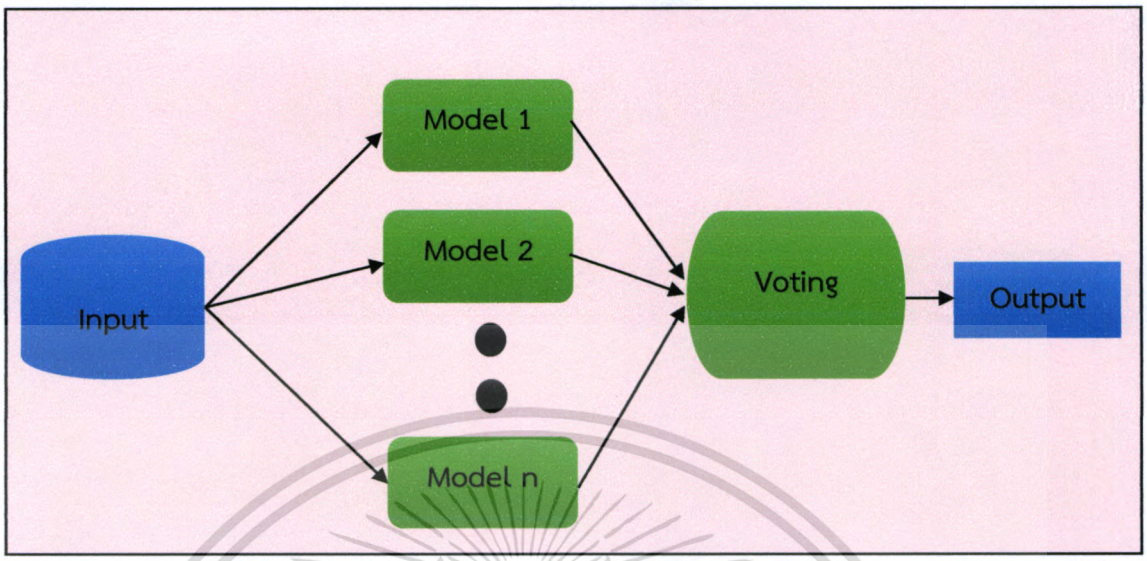
ตัวจำแนกแบบเบย์อย่างง่ายจะคำนวณหาค่าความน่าจะเป็น $P(A_1, A_2, \dots, A_n|Class)$ โดยจะขึ้นอยู่กับประเภทของข้อมูลที่มีการเรียนรู้ในชุดข้อมูลนั้น กำหนดให้ C_j เป็นสมาชิกของคลาสต่างๆ ซึ่งผลลัพธ์ของการจำแนกประเภทจะกำหนดคลาสที่มีค่าความน่าจะเป็นมากที่สุดให้กับตัวอย่างที่ต้องการทำนาย ดังสมการที่ 2.21

$$y = argmax_{C_j \in C} P(C_j) \prod_{i=1}^{n} P(A_i|C_j) \quad (2.21)$$

2.5 การเรียนรู้แบบรวมกลุ่ม

การเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) [1,11,12,13] คือ การเรียนรู้ที่เป็นการรวมกลุ่มของตัวจำแนกประเภทแบบพื้นฐานหลายๆ ตัวให้ทำงานร่วมกันเพื่อเพิ่มประสิทธิภาพในการเรียนรู้ซึ่งจะส่งผลให้เกิดความถูกต้องมากขึ้น และมีประสิทธิภาพมากกว่าโมเดลแบบเดี่ยว งานวิจัยนี้มุ่งเน้นไปที่การประยุกต์ใช้ โครงข่ายประสาทเทียมทำงานร่วมกับการใช้ตัวจำแนกแบบเบย์อย่างง่ายในการสร้างโมเดลการเรียนรู้แบบรวมกลุ่มโดยใช้อัลกอริทึมเอดาบูสต์ เพื่อให้สามารถจำแนกประเภทข้อมูลในหลายๆ รูปแบบได้อย่างมีประสิทธิภาพเมื่อเปรียบเทียบกับการใช้โมเดลแบบเดี่ยว ดังนั้นการเรียนรู้แบบรวมกลุ่มจะช่วยให้การเรียนรู้มีประสิทธิภาพสูงขึ้น และสามารถแก้ไขปัญหาที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ [14,15]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 โมเดลการเรียนรู้แบบรวมกลุ่ม

จากรูปที่ 2.7 การเรียนรู้แบบรวมกลุ่มเป็นเทคนิคการนำโมเดลจำแนกประเภทหลายโมเดลมาทำงานร่วมกันในการจำแนกประเภทข้อมูลเพื่อเพิ่มประสิทธิภาพของการจำแนกประเภท หลักเกณฑ์ในการคัดเลือกโมเดลที่จะนำมารวมกลุ่มประกอบด้วยปัจจัยสามอย่าง คือ ความแม่นยำของโมเดล ความหลากหลายของโมเดล และประสิทธิภาพโดยรวมของโมเดล การเรียนรู้แบบรวมกลุ่มประกอบด้วย 4 ส่วน คือ ข้อมูลนำเข้า (Input) โมเดลการจำแนกประเภท (Base Classifier) อัลกอริทึมที่ใช้ในการเลือกคำตอบ (Voting) และข้อมูลส่งออก (Output) อธิบายได้ดังนี้

1. ข้อมูลนำเข้า คือ ส่วนของข้อมูลสำหรับใช้ในการสอนและทดสอบระบบ
2. โมเดลการจำแนกประเภทพื้นฐาน คือ ส่วนของอัลกอริทึมการจำแนกประเภทที่นำมาใช้ในการเรียนรู้แบบรวมกลุ่ม โดยจะทำการสร้างโมเดลเดียวกันจำนวนหลายๆ โมเดลเข้ามาช่วยในการเรียนรู้เพื่อเพิ่มประสิทธิภาพให้ผลออกมาดียิ่งขึ้น
3. การเลือกคำตอบ คือ อัลกอริทึมที่ในการเลือกคำตอบโดยจะทำการรวบรวมผลลัพธ์ที่ได้จากโมเดลการจำแนกประเภทพื้นฐานทั้งหมดมาทำการโหวตผลลัพธ์
4. ข้อมูลส่งออก คือ ผลลัพธ์ที่ได้ออกมาจากการเรียนรู้แบบรวมกลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานของการเรียนรู้แบบรวมกลุ่ม เริ่มจากการเลือกหรือกำหนดโมเดลการจำแนกประเภทพื้นฐาน (Base Classifier) จากนั้นกำหนดจำนวนโมเดลการเรียนรู้ว่าจะใช้จำนวนเท่าไหร่ในการเรียนรู้เพื่อที่จะวัดผลออกมาได้อย่างมีประสิทธิภาพ เมื่อทำการกำหนดทุกอย่างเรียบร้อยแล้ว นำข้อมูลนำเข้ามาเพื่อทำการสอนระบบ จากนั้นอัลกอริทึมที่ใช้สร้างโมเดลแบบรวมกลุ่มจะนำผลลัพธ์ที่ได้จากโมเดลการเรียนรู้มาทำการปรับปรุงค่าการกระจายตัวของข้อมูลนำเข้าและดำเนินการจนครบจำนวนโมเดลที่กำหนดแล้วจึงนำไปใช้ทำนายผลข้อมูลสำหรับทดสอบ ในวิทยานิพนธ์นี้ประยุกต์ใช้อัลกอริทึมเอดาบูสต์ในการสร้างโมเดลการเรียนรู้แบบรวมกลุ่มเพื่อจำแนกประเภทข้อมูล

2.6 อัลกอริทึมเอดาบูสต์

อัลกอริทึมเอดาบูสต์ (Adaboost algorithm) [1] คือ อัลกอริทึมที่พัฒนาขึ้นมาสำหรับการให้น้ำหนักกับตัวจำแนกประเภทโดยให้ค่าน้ำหนักจากความผิดพลาดของตัวจำแนกประเภท ซึ่งมีลำดับขั้นตอนดังนี้

อัลกอริทึม : Adaboost(M1)

นำเข้า: Training set $S = \{x_i, y_i\}_{i=1}^{N},\ i = 1, \ldots, N;\ and\ y_i \in C, C = \{c_1, \ldots., c_m\}$

T : Number of iteration

I : Weak learner

ขั้นตอนการทำงาน :

1. $D_1(i) \leftarrow \frac{1}{N}\ for\ i = 1, \ldots, N$ (2.22)
2. For t =1 to T do
3. $h_t \leftarrow I\ (S, D_t)$
4. $\epsilon_t \leftarrow \sum_{i=1}^{N} D_t(i) \mathrm{I}[h_t(x_i) \neq y_i]$ (2.23)
5. $if\ \epsilon_t > 0.5\ \ then$
6. $T \leftarrow t - 1$
7. return

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. End if

9. $\beta_t = \epsilon_t/(1-\epsilon_t)$. (2.24)

10. $D_{t+1}(i) = D_t(i) \times \beta_t^{1-\mathrm{I}[h_t(\mathrm{x_i}) \neq y_i]} \; for \; i = 1,..,n$ (2.25)

11. Normalize $D_t(i)$ to be a proper distribution.

12. End for

ผลลัพธ์:

$$H(\mathrm{x}) = \arg\max_{y \in \mathrm{c}} \sum_{t=1}^{T} \ln(\frac{1}{\beta_t}) \, \mathrm{I}[h_t(\mathrm{x}) = y] \quad (2.26)$$

โดยที่ h_t คือผลลัพธ์ของตัวจำแนกที่ t และ β_t คือ สัมประสิทธิ์ความผิดพลาด

หลักการทำงานของอัลกอริทึมเอดาบูสต์ มีดังนี้

ขั้นตอนที่ 1 กำหนดค่าการแจกแจงหรือค่าน้ำหนัก (Distribution) ของแต่ละข้อมูล โดยที่ค่า N คือ จำนวนข้อมูลทั้งหมด

ขั้นตอนที่ 2 กำหนดค่าเริ่มต้นการวนซ้ำ หรือจำนวนรอบ เท่ากับ T

ขั้นตอนที่ 3 เรียกใช้งาน Weak Learner ที่ต้องการ โดยค่า h_t คือ ผลลัพธ์การเรียนรู้ของโมเดลจำแนกประเภทแบบพื้นฐาน (Base Classifier)

ขั้นตอนที่ 4 คำนวณหาค่าความผิดพลาด โดยค่า ϵ_t คือ ค่าความผิดพลาด (Error Rate)

ขั้นตอนที่ 5 ถ้าค่าความผิดพลาด ϵ_t มีค่าเกิน 0.5 ให้หยุดการทำงานจากลูป โดยให้ดำเนินการตามขั้นตอนที่ 6 และ 7

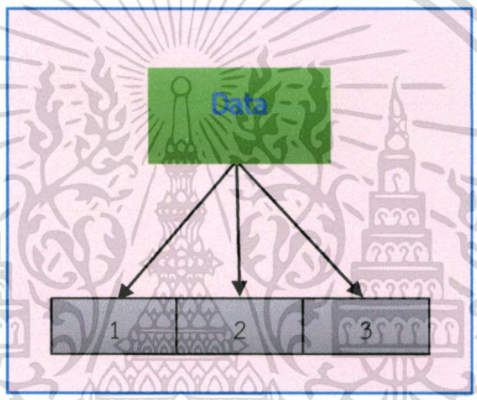
ขั้นตอนที่ 9 ทำการคำนวณหาค่าสัมประสิทธิ์ความผิดพลาด β_t

ขั้นตอนที่ 10-11 ทำการคำนวณและปรับค่าการแจกแจง (Distribution)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 เทคนิค *K*-Fold Cross-Validation

เทคนิค *K*-Fold Cross-Validation เป็นวิธีการที่ใช้วัดประสิทธิภาพในการทำนายตัวอย่างของโมเดลการจำแนกประเภท โดยพื้นฐานของเทคนิคนี้คือ การสุ่มตัวอย่าง (Sampling) โดยเริ่มจากแบ่งชุดข้อมูลออกเป็นส่วนๆที่เรียกว่า fold และนำบางส่วนจากชุดข้อมูลนั้นมาทดสอบผลลัพธ์จากการทำนายข้อมูลทดสอบของโมเดล ในกรณีการเลือกสุ่มข้อมูลแบบมีความเที่ยงตรง *K* กลุ่ม เราจะแบ่งข้อมูลออกเป็น k ชุดเท่าๆกัน และทำการคำนวณค่าความแม่นยำจากการทำนาย *K* รอบ โดยแต่ละรอบจะมีการสร้างโมเดลจำแนกประเภทหนึ่งตัวจากข้อมูลเรียนรู้ *K* -1 ชุด และที่เหลืออีก 1 ชุด จะใช้ในการทดสอบ



รูปที่ 2.8 เทคนิค k-fold cross-validation

จากรูปที่ 2.8 ในการทำงานรอบแรก ข้อมูลในชุดที่ 1 จะใช้เป็นข้อมูลทดสอบ ส่วนข้อมูลในชุดที่ 2 และ 3 จะนำมาใช้เป็นชุดข้อมูลสำหรับการเรียนรู้ ซึ่งจะได้โมเดลจำแนกประเภท 1 ตัว รอบที่สอง ข้อมูลในชุดที่ 2 จะใช้เป็นข้อมูลทดสอบ ส่วนข้อมูลในชุดที่ 1 และ 3 จะนำมาใช้เป็นชุดข้อมูลสำหรับการเรียนรู้ และรอบสุดท้าย ข้อมูลในชุดที่ 3 จะใช้เป็นข้อมูลทดสอบ ส่วนข้อมูลในชุดที่ 1 และ 2 จะนำมาใช้เป็นชุดข้อมูลสำหรับการเรียนรู้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงานวิจัย

ในบทนี้จะเป็นการกล่าวถึงขั้นตอน วิธีการดำเนินการวิจัย โดยจะขอกล่าวถึงวิธีการและอัลกอริทึมที่ใช้รวมถึงแนวความคิดการออกแบบอัลกอริทึมใหม่ เพื่อใช้ในการทดสอบแนวคิดและสมมติฐานที่ตั้งขึ้นมา

3.1 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองนำมาจากเว็บไซต์ UCI (https://archive.ics.uci.edu/ml/index.html) ซึ่งประกอบด้วยข้อมูลหลากหลายประเภทและหลากหลายคุณลักษณะ โดยได้เลือกข้อมูลที่เหมาะสมกับปัญหาการจำแนกประเภทข้อมูล (Classification) จำนวนทั้งหมด 7 ชุดข้อมูล ดังตารางที่ 3.1

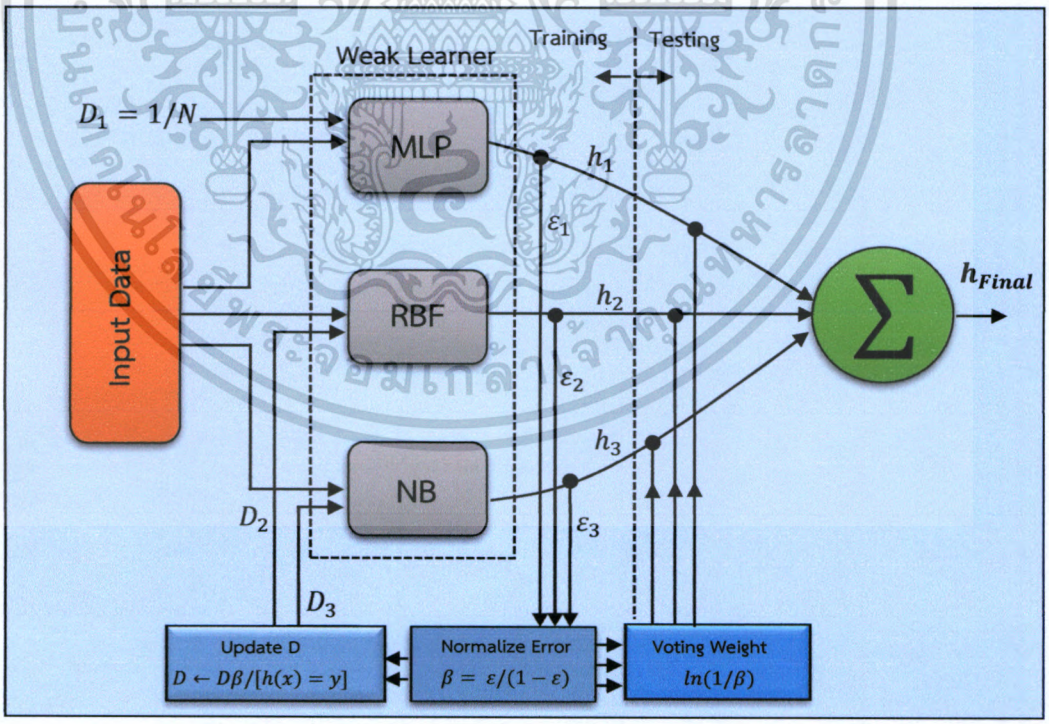
ตารางที่ 3.1 คุณลักษณะของชุดข้อมูลที่ใช้ในการทดลอง

ชุดข้อมูล	คุณลักษณะ					
	ลักษณะชุดข้อมูล	ชนิดข้อมูล	ประเภท	จำนวนข้อมูล	จำนวนคุณลักษณะ	จำนวนกลุ่มข้อมูล
iris	ข้อมูลหลายตัวแปร	จำนวนจริง	การจำแนกประเภท	150	4	3
wine	ข้อมูลหลายตัวแปร	จำนวนเต็ม จำนวนจริง	การจำแนกประเภท	178	13	3
Balance Scale	ข้อมูลหลายตัวแปร	จำนวนเต็ม	การจำแนกประเภท	625	4	3
Lung Cancer	ข้อมูลหลายตัวแปร	จำนวนเต็ม	การจำแนกประเภท	32	54	3
Qualitative Bankruptcy	ข้อมูลหลายตัวแปร	จำนวนเต็ม	การจำแนกประเภท	250	6	2
Banknote authentication Set	ข้อมูลหลายตัวแปร	จำนวนเต็ม	การจำแนกประเภท	569	29	2
Breast Cancer Wisconsin (Diagnostic)	ข้อมูลหลายตัวแปร	จำนวนจริง	การจำแนกประเภท	569	29	2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น

อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น (MRN Learning Algorithm) เป็นโครงสร้างที่นำเอาโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multi Layer Proceptron) โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial Basis Function) และตัวจำแนกแบบเบย์อย่างง่าย (Naïve Bayes Classifier) มาทำงานร่วมกันในลักษณะของการเรียนรู้แบบรวมกลุ่ม โดยใช้หลักการทำงานตามแนวความคิดในการนำโครงข่ายประสาทเทียม และตัวจำแนกแบบเบย์อย่างง่ายมาทำงานร่วมกันเพื่อเพิ่มประสิทธิภาพของการเรียนรู้แทนที่จะใช้อัลกอริทึมเพียงชนิดเดียว แต่หลักการของอัลกอริทึมที่ถูกสร้างขึ้นมาใหม่นั้นจะใช้อัลกอริทึมทั้งสองแบบทำงานร่วมกัน โดยหลังจากที่สร้างโมเดลแล้วก็จะได้ค่าน้ำหนักของแต่ละตัวจำแนกประเภทเพื่อนำมาใช้ในการพิจารณาการจำแนกประเภทข้อมูล ในการนำโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้นไว้เป็นส่วนแรกเพราะจากการศึกษาพบว่าเครือข่ายประสาทเทียมชนิดนี้มีประสิทธิภาพสูงมากในการจำแนกประเภทข้อมูล อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น มีหลักการทำงาน ดังนี้



รูปที่ 3.1 อัลกอริทึมการเรียนรู้เอ็มอาร์เอ็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึม : MRN learning algorithm

ข้อมูลนำเข้า:

ชุดข้อมูลเรียนรู้ S = $\{\mathbf{x}_1, \ldots\ldots, \mathbf{x}_n\}$ โดยที่ $\mathbf{x}_i \in X$

โดยที่ $y_i \in \Omega, \Omega = \{y_1, \ldots\ldots, y_C\}$

เอาต์พุต:

$$H(\mathbf{x}) = \arg\max_{y\in\Omega}\left\{\ln\left(\frac{1}{\beta_a}\right)\mathrm{I}[h_a(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_b}\right)I[h_b(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_c}\right)\mathrm{I}[h_c(\mathbf{x}) = y]\right\} \quad (3.1)$$

โดยที่ $I[x] = \begin{cases} 1\ if\ x\ is\ true \\ 0\ otherwise \end{cases}$

วิธีการทำงาน:

1. กำหนดค่าเริ่มต้นให้กับข้อมูล $D_1(i)$ = 1/*n*, *i*=1,2,...,*n*.
2. สร้าง Base Classifier โดยใช้โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (MLP Weak Learner) $h_a : X \rightarrow \Omega$.
3. คำนวณค่าความผิดพลาดของ h_a:

$$\varepsilon_a = \sum_{i=1}^{n} I[h_a(\mathbf{x_i}) \neq y_i] \cdot D_i \quad (3.2)$$

โดยที่ ε_a เป็นค่า Error ของ h_a และ y_i เป็นผลลัพธ์ที่ต้องการ

4. ถ้า ε_a> 0.5 ให้กลับไปทำขั้นตอนที่ 2
5. คำนวณค่าถ่วงน้ำหนักของตัวจำแนกประเภท β_a จากสูตร

$$\beta_a = \varepsilon_a/(1 - \varepsilon_a) \quad (3.3)$$

6. ปรับปรุงค่าการกระจาย (Distribution) D_a :

$$D_{a+1}(i) = \frac{D_a(i)}{z_a} \times \begin{cases} \beta_a\ if\ h_a(\mathbf{x_i}) = y_i \\ 1\ otherwise \end{cases} \quad (3.4)$$

โดยที่ z_a เป็นค่าคงที่สำหรับการนอร์มัลไลซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. สร้าง Base Classifier โดยใช้โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (RBF Weak Learner) $h_b : X \rightarrow \Omega$.

8. คำนวณค่าความผิดพลาดของ h_b

$$\varepsilon_b = \sum_{i=1}^{n} I[h_b(\mathbf{x}_i) \neq y_i] \cdot D_i \quad (3.5)$$

โดยที่ ε_b เป็นค่า Error ของ h_b และ y_i ผลลัพธ์ที่ต้องการ

9. ถ้า ε_b> 0.5 ให้กลับไปทำขั้นตอนที่ 7.

10. คำนวณค่าถ่วงน้ำหนักของตัวจำแนกประเภท β_b จากสูตร

$$\beta_b = \varepsilon_b/(1-\varepsilon_b) \quad (3.6)$$

11. ปรับปรุงค่าการกระจาย D_i :

$$D_{b+1}(i) = \frac{D_b(i)}{z_b} \times \begin{cases} \beta_b \; if \; h_b(\mathbf{x}_i) = y_i \\ 1 \; otherwise \end{cases} \quad (3.7)$$

โดยที่ z_b เป็นค่าคงที่สำหรับการนอร์มัลไลซ์

12. สร้าง base classifier โดยใช้ตัวจำแนกแบบเบย์อย่างง่าย (Naïve Bayes Classifier Weak Learner) $h_c : X \rightarrow \Omega$.

$$\varepsilon_c = \sum_{i=1}^{n} I[h_c(\mathbf{x}_i) \neq y_i] \cdot D_i \quad (3.8)$$

โดยที่ ε_c เป็นค่า Error ของ h_c และ y_i ผลลัพธ์ที่ต้องการ

13. ถ้า ε_c> 0.5 ให้กลับไปทำขั้นตอนที่ 12.

14. คำนวณค่าถ่วงน้ำหนักของตัวจำแนกประเภท β_c จากสูตร

$$\beta_c = \varepsilon_c/(1-\varepsilon_c) \quad (3.9)$$

15. ปรับปรุงค่าการกระจาย D_c :

$$D_{c+1}(i) = \frac{D_c(i)}{z_c} \times \begin{cases} \beta_c \; if \; h_c(\mathbf{x}_i) = y_i \\ 1 \; otherwise \end{cases} \quad (3.10)$$

โดยที่ z_c เป็นค่าคงที่สำหรับการนอร์มัลไลซ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับข้อมูลทดสอบ เราสามารถระบุกลุ่มของข้อมูลได้โดยใช้สมการ ดังนี้

$$H(\mathbf{x}) = \arg\max_{y\in\Omega}\left\{\ln\left(\frac{1}{\beta_a}\right)\mathrm{I}[h_a(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_b}\right)I[h_b(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_c}\right)\mathrm{I}[h_c(\mathbf{x}) = y]\right\} \quad (3.11)$$

3.3 ขั้นตอนการสร้างโมเดลการเรียนรู้แบบรวมกลุ่ม

เมื่อได้สร้างอัลกอริทึมตามสมมุติฐานแล้ว และจัดการเตรียมข้อมูลสำหรับทดสอบโมเดลเรียบร้อยแล้ว ขั้นตอนต่อไปคือการสร้างโมเดลการเรียนรู้แบบรวมกลุ่ม โดยจะทำการโหลดข้อมูล (Data Set) ที่เลือกมาทดสอบและทำการแบ่งข้อมูลออกเป็น 3 ชุด (K=3) โดยใช้เทคนิค K-Fold Cross-Validation จากนั้นจะนำข้อมูลจำนวน 2 ชุด มาเป็นชุดข้อมูลสำหรับฝึกสอน (Training Set) และข้อมูลชุดที่เหลือจะใช้ในการทดสอบโมเดล โดยกำหนดค่าการแจกแจงเริ่มต้นให้กับแต่ละข้อมูล $D_1(i) = 1/n$, i=1,2,...,n ในขั้นตอนนี้เป็นการกำหนดน้ำหนักเริ่มต้นให้กับข้อมูลในชุดฝึกสอน จากนั้นในขั้นตอนของการฝึกสอนจะนำข้อมูลชุดฝึกสอนไปใช้ฝึกสอนการเรียนรู้ในส่วนที่เรียกว่า ตัวจำแนกประเภทพื้นฐาน (Base Classifier) โดยขั้นตอนของการเรียนรู้จะมีอยู่ 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1

ในขั้นตอนนี้จะเริ่มจาก ตัวจำแนกประเภทแบบโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (MLP Weak Learner) ในขั้นตอนที่ 2 จากนั้นคำนวณหาค่าความผิดพลาดของ h_a นั่นคือ ε_a จากสูตรในขั้นตอนที่ 3 อัลกอริทึมจะทำงานวนซ้ำในขณะที่ค่าความผิดพลาด ε_a มีค่าเกิน 0.5 จากนั้นเมื่อจบการเรียนรู้ส่วนนี้แล้วจะคำนวณหาค่าถ่วงน้ำหนักของตัวจำแนกประเภท คือ β_a จากสูตรในขั้นตอนที่ 5 จากนั้นจะทำการหาค่าและทำการปรับค่าการแจกแจงจากสูตรในขั้นตอนที่ 6

ขั้นตอนที่ 2

ในขั้นตอนนี้จะเริ่มการเรียนรู้ในส่วนที่ 2 โดยใช้ตัวจำแนกประเภทแบบโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (RBF Weak Learner) และคำนวณหาค่าความผิดพลาดของ h_b จากสูตรในขั้นตอนที่ 8 อัลกอริทึมจะทำงานวนซ้ำในขณะที่ค่าความผิดพลาด ε_b มีค่าเกิน 0.5 จากนั้นเมื่อจบการเรียนรู้ส่วนนี้แล้วจะคำนวณหาค่าถ่วงน้ำหนักของตัวจำแนกประเภทคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

β_b จากสูตรในขั้นตอนที่ 10 จากนั้นจะทำการหาค่าและทำการปรับค่าการกระจาย จากสูตรในขั้นตอนที่ 11

ขั้นตอนที่ 3

ในขั้นตอนนี้ จะทำการเริ่มการเรียนรู้ในส่วนที่ 3 โดยใช้ตัวจำแนกประเภทแบบเบย์อย่างง่าย (Naïve Bayes Classifier Weak Learner) จากนั้นคำนวณหาค่าความผิดพลาดของ h_c จากสูตรในขั้นตอนที่ 12 อัลกอริทึมจะทำงานวนซ้ำในขณะที่ค่าความผิดพลาด ε_c มีค่าเกิน 0.5 จากนั้นเมื่อจบการเรียนรู้ส่วนนี้แล้วจะคำนวณหาค่าถ่วงน้ำหนักของตัวจำแนกประเภทคือ β_c ของตัวจำแนกประเภทการเรียนรู้แบบเบย์อย่างง่าย จากสูตรในขั้นตอนที่ 14 จากนั้นจะทำการหาค่าและทำการปรับค่าการกระจายจากสูตรในขั้นตอนที่ 15

ขั้นตอนที่ 4

ในขั้นตอนสุดท้ายของของอัลกอริทึม MRN จะใช้ค่าถ่วงน้ำหนัก β_a , β_b และ β_c มาหาค่าผลลัพธ์โดยจะทำการนำค่าถ่วงน้ำหนักของแต่ละตัวจำแนกประเภท มาคำนวณแล้วเลือกค่าน้ำหนักรวมที่สูงที่สุดมาทำนายผลลัพธ์ของชุดข้อมูลทดสอบ ดังสมการที่ 3.11

3.4 การคำนวณผลลัพธ์ของโมเดลการเรียนรู้แบบรวมกลุ่ม

ในหัวข้อนี้จะเป็นการแสดงการหาคำตอบจากอัลกอริทึมการเรียนรู้เอ็มอาร์เอ็นโดยใช้ค่าถ่วงน้ำหนักของตัวจำแนกประเภทแต่ละตัว ตามสมการ 3.11 ดังนี้

$$\mathrm{H}(\mathbf{x}) = \arg\max_{y \in \Omega} \left\{ \ln\left(\frac{1}{\beta_a}\right) I[h_a(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_b}\right) I[h_b(\mathbf{x}) = y] + \ln\left(\frac{1}{\beta_c}\right) \mathrm{I}[h_c(\mathbf{x}) = y] \right\}$$

โดยที่ $I[x] = \begin{cases} 1 \; if \;\; x \; is \; true \\ 0 \;\; otherwise \end{cases}$

จากสมการ y เป็นคลาสที่เป็นไปได้สำหรับปัญหานี้ เมื่อแต่ละตัวจำแนกประเภทได้ทำนายคำตอบมาแล้วก็จะทำการหาคำตอบดังตัวอย่างต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 1

สำหรับข้อมูลทดสอบใดๆ สมมติให้ y เป็นสมาชิกของเซตคำตอบ โดยที่ y = {1,2,3 }

ขั้นตอนที่ 2

หลังจากสร้างโมเดลการเรียนรู้แบบรวมกลุ่มเสร็จแล้ว จะได้ค่า $\ln\left(\frac{1}{\beta_a}\right)$, $\ln\left(\frac{1}{\beta_b}\right)$, และ $\ln\left(\frac{1}{\beta_c}\right)$ ตามลำดับ หลังจากนั้นใส่ข้อมูลทดสอบเข้าไปในโมเดลการเรียนรู้แบบรวมกลุ่ม โดยแต่ละตัวจำแนกประเภทจะทำนายคลาสคำตอบของข้อมูลทดสอบออกมา สมมติว่าตัวจำแนกประเภทแบบโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้นทำนายว่าเป็นคลาสที่ 1 และมีค่าน้ำหนัก $\ln\left(\frac{1}{\beta_a}\right)$ = 0.5 ตัวจำแนกประเภทแบบโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันทำนายว่าเป็นคลาสที่ 2 และมีค่าน้ำหนัก $\ln\left(\frac{1}{\beta_b}\right)$ = 0.2 และตัวจำแนกประเภทแบบเบย์อย่างง่ายทำนายว่าเป็นคลาสที่ 2 และมีค่าน้ำหนัก $\ln\left(\frac{1}{\beta_c}\right)$ = 0.1

ขั้นตอนที่ 3

คำนวณค่าผลลัพธ์ของอัลกอริทึมเอ็มอาร์เอ็น ดังนี้

กรณี คลาส y=1 จะได้ผลรวมของค่าน้ำหนัก คือ

(0.5) (1) + (0.2)(0) + (0.1)(0) = 0.5

กรณี คลาส y=2 จะได้ผลรวมของค่าน้ำหนัก คือ

(0.5) (0) + (0.2)(1) + (0.1)(1) = 0.3

กรณี คลาส y=3 จะได้ผลรวมของค่าน้ำหนัก คือ

(0.5) (0) + (0.2)(0) + (0.1)(0) = 0.0

เพราะฉะนั้นคลาสที่น่าจะเป็นไปได้สำหรับข้อมูลนี้ คือ คลาส y=1 เพราะมีค่าน้ำหนักรวมมากที่สุด แต่หากเป็นวิธีการ voting คำตอบที่เป็นไปได้สำหรับข้อมูลนี้ คือ คลาส y=2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การวัดประสิทธิภาพของตัวจำแนกประเภทในการทดลองจะใช้ ค่าความถูกต้อง (Accuracy) ของตัวจำแนกประเภท ซึ่งสามารถคำนวณได้จากสมการที่ 3.12

$$\text{Accuracy}_i = \frac{Correctly\ Classification\ Instances}{Total\ Number\ of\ Instances} \times 100 \qquad (3.12)$$

โดยที่ $Correctly\ Classification\ Instances$ คือ จำนวนข้อมูลที่ทำนายได้อย่างถูกต้อง

$Total\ Number\ of\ Instances$ คือ จำนวนข้อมูลทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

เนื่องจากในการทดลองนี้ ผู้วิจัยได้ออกแบบการทดลองโดยสลับที่ตำแหน่งของตัวจำแนกประเภทพื้นฐานแต่ละแบบในโมเดลการเรียนรู้แบบรวมกลุ่ม เพื่อให้เกิดการวัดประสิทธิภาพของโมเดลการเรียนรู้แบบรวมกลุ่มอย่างชัดเจนและครอบคลุมในทุกรูปแบบ ดังนี้

ตารางที่ 4.1 โมเดลการเรียนรู้แบบรวมกลุ่มในแต่ละรูปแบบที่ใช้ในการทดลองนี้

ชื่อ	ชนิดตัวจำแนกประเภทพื้นฐาน		
	ลำดับที่ 1	ลำดับที่ 2	ลำดับที่ 3
MRN	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	ตัวจำแนกแบบเบย์อย่างง่าย
MNR	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	ตัวจำแนกแบบเบย์อย่างง่าย	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน
NRM	ตัวจำแนกแบบเบย์อย่างง่าย	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น
NMR	ตัวจำแนกแบบเบย์อย่างง่าย	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน
RNM	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	ตัวจำแนกแบบเบย์อย่างง่าย	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น
RMN	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	ตัวจำแนกแบบเบย์อย่างง่าย
MMM	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น	โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น
RRR	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน	โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน
NNN	ตัวจำแนกแบบเบย์อย่างง่าย	ตัวจำแนกแบบเบย์อย่างง่าย	ตัวจำแนกแบบเบย์อย่างง่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 1

เป็นการทดสอบประสิทธิภาพการทำงานของโมเดลการเรียนรู้แบบเดี่ยว คือ MLP RBF NB และโมเดลการเรียนรู้แบบรวมกลุ่มในแต่ละรูปแบบ ซึ่งมีการสลับที่ของตัวจำแนกประเภทในรูปแบบต่างๆ ได้แก่ MRN NRM RNM RMN MNR และ NMR โดยในการทดลองนี้ ผู้วิจัยได้ทดสอบประสิทธิภาพของโมเดลการเรียนรู้ในแต่ละรูปแบบเป็นจำนวน 10 รอบและบันทึกค่าความถูกต้องที่ให้ค่าสูงที่สุดของแต่ละรูปแบบในทุกชุดข้อมูล เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการเรียนรู้แต่ละรูปแบบ ดังแสดงในตารางที่ 4.2 นี้

ตารางที่ 4.2 ค่าความถูกต้องของโมเดลการเรียนรู้แบบเดี่ยว และโมเดลการเรียนรู้แบบรวมกลุ่ม ซึ่งมีการสลับที่ของตัวจำแนกประเภทต่างชนิดกัน

ชุดข้อมูล	ค่าความถูกต้อง								
	MLP	RBF	NB	MRN	NRM	RNM	RMN	MNR	NMR
Iris	96.67	96.67	96.66	**97.33**	96.60	97.33	97.30	96.62	96.70
Wine	97.75	77.53	98.30	**98.87**	98.30	97.10	97.74	97.74	98.31
Lung Cancer	53.30	56.60	66.06	**71.80**	65.50	59.39	59.44	63.03	55.75
Balance scale	91.04	87.50	88.35	**91.20**	90.30	90.00	89.28	91.20	90.23
Qualitative Bankruptcy	99.60	99.21	76.40	**99.98**	98.80	99.60	99.20	93.50	99.20
Breast Cancer Wisconsin	97.10	92.79	93.32	**97.89**	93.30	96.30	95.77	93.32	97.53
Liver	68.40	64.05	60.84	**70.40**	61.10	66.60	69.44	61.10	69.40
ค่าเฉลี่ย	86.27	82.05	82.85	**89.64**	86.27	86.62	86.88	85.22	86.73

จากผลการทดลอง จะเห็นได้ว่า อัลกอริทึม เอ็มอาร์เอ็น (MRN) ให้ค่าความถูกต้องที่สูงที่สุดในทุกชุดข้อมูลที่ใช้ในการทดสอบ ทั้งแบบสองคลาสและแบบหลายคลาส โดยให้ค่าความถูกต้องที่สูงกว่าโมเดลการเรียนรู้แบบเดี่ยวและโมเดลการเรียนรู้แบบรวมกลุ่มในรูปแบบที่เหลือ นอกจากนี้ เมื่อพิจารณาจากค่าความถูกต้องเฉลี่ยโดยรวม ยังพบอีกว่า ผลของโมเดลการเรียนรู้แบบรวมกลุ่มทุกรูปแบบ ยกเว้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึม เอ็มเอ็นอาร์ จะให้ค่าถูกต้องที่สูงกว่า โมเดลการเรียนรู้แบบเดี่ยว ซึ่งแสดงโดยนัยว่า การนำตัวจำแนกประเภทต่างชนิดกันมาสร้างเป็นโมเดลการเรียนรู้แบบรวมกลุ่มมีแนวโน้มที่จะให้ประสิทธิภาพในการจำแนกประเภทกลุ่มข้อมูล ได้ดีกว่า โมเดลการเรียนรู้แบบเดี่ยว

ขั้นตอนที่ 2

เป็นการทดสอบประสิทธิภาพการทำงานของโมเดลการเรียนรู้แบบรวมกลุ่ม โดย MMM เป็นตัวแทนของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้การส่งแบบเชิงเส้น RRR เป็นตัวแทนของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้การส่งแบบไม่เชิงเส้น NNN เป็นตัวแทนของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้ทฤษฎีความน่าจะเป็น และ MRN เป็นตัวแทนของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้รูปแบบการส่งทั้งแบบเชิงเส้น ไม่เชิงเส้น และทฤษฎีความน่าจะเป็น โดยในการทดลองนี้ ผู้วิจัยได้ทดสอบประสิทธิภาพของโมเดลการเรียนรู้ในแต่ละโมเดลเป็นจำนวน 10 รอบและบันทึกค่าความถูกต้องที่ให้ค่าสูงที่สุดของแต่ละโมเดลในทุกชุดข้อมูล เพื่อเปรียบเทียบประสิทธิภาพของโมเดลการเรียนรู้แต่ละรูปแบบ ดังแสดงในตารางที่ 4.3

ตารางที่ 4.3 ค่าความถูกต้องของโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้ตัวจำแนกประเภทแบบเดียวกันทั้งหมด และแบบที่ใช้ตัวจำแนกประเภทต่างชนิดกัน

ชุดข้อมูล	ชนิดของโมเดลการเรียนรู้แบบรวมกลุ่ม			
	MMM	RRR	NNN	MRN
Iris	97.31	97.31	96.67	**97.33**
Wine	98.32	79.76	98.31	**98.87**
Lung Cancer	56.06	62.22	68.48	**71.80**
Balance scale	90.55	87.19	88.47	**91.20**
Qualitative Bankruptcy	99.61	99.20	84.69	**99.98**
Breast Cancer Wisconsin	97.71	92.96	93.49	**97.89**
Liver	69.44	66.66	62.50	**70.40**
ค่าเฉลี่ย	87.00	83.61	84.66	**89.64**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.3 แสดงให้เห็นว่า อัลกอริทึม MRN ซึ่งเป็นโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้รูปแบบการส่งทั้งแบบเชิงเส้น แบบไม่เชิงเส้น และทฤษฎีความน่าจะเป็น ให้ค่าความถูกต้องที่สูงที่สุดในทุกชุดข้อมูลที่ใช้ในการทดสอบ ทั้งแบบสองคลาสและแบบหลายคลาส นอกจากนี้ เมื่อพิจารณาเฉพาะค่าความถูกต้องเฉลี่ยโดยรวมของการจำแนกประเภทกลุ่มข้อมูล ยังพบอีกว่า อัลกอริทึม MRN ให้ค่าความถูกต้องสูงที่สุด คือ 89.64% โดยอัลกอริทึม MMM ให้ค่าความถูกต้องที่สูงรองลงมาคือ 87.00% และอัลกอริทึม RRR และ NNN ให้ค่าความถูกต้องที่ใกล้เคียงกัน คือ 83.61% และ 84.66% ตามลำดับ จากผลการทดลองนี้ แสดงให้เห็นว่า โมเดลการเรียนรู้แบบรวมกลุ่ม ซึ่งใช้ตัวจำแนกประเภทบนพื้นฐานของการส่งรูปแบบเดียวกัน คือ MMM, RRR และ NNN ให้ประสิทธิภาพการจำแนกประเภทข้อมูลที่ต่ำกว่าโมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้ตัวจำแนกประเภทบนพื้นฐานของการส่งต่างชนิดกัน คือ อัลกอริทึม MRN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปและข้อเสนอแนะ

5.1 สรุป

งานวิจัยนี้เป็นการนำเสนอการสร้างอัลกอริทึมการจำแนกประเภทแบบใหม่ คือ อัลกอริทึม เอ็มอาร์เอ็น ซึ่งใช้เทคนิคของการเรียนรู้แบบรวมกลุ่ม เพื่อให้เกิดการทำงานร่วมกันของ โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และตัวจำแนกแบบเบย์อย่างง่าย โดยในการทดลองนี้ได้เลือกใช้อัลกอริทึม เอดาบูสต์ ในการสร้างโมเดลการเรียนรู้แบบรวมกลุ่ม และการคำนวณค่าถ่วงน้ำหนักสำหรับตัวจำแนกประเภทแต่ละตัว จากผลการทดลองการเปรียบเทียบประสิทธิภาพของตัวจำแนกประเภทข้อมูล ของโมเดลการเรียนรู้แบบเดี่ยว และโมเดลการเรียนรู้แบบรวมกลุ่มซึ่งใช้ตัวจำแนกประเภทแบบเดียวกัน และต่างชนิดกัน ในชุดข้อมูลทั้งแบบสองคลาส และแบบหลายคลาส สามารถแสดงถึงข้อสรุปได้ ดังนี้

ประเด็นที่หนึ่ง คือ การสลับที่ของตัวจำแนกประเภทการเรียนรู้ ในโมเดลการเรียนรู้แบบรวมกลุ่มนั้น มีผลต่อประสิทธิภาพการเรียนรู้ และจากผลการทดลองของค่าความถูกต้องเฉลี่ยโดยรวม พบว่า โมเดลการเรียนรู้แบบรวมกลุ่มมีแนวโน้มที่จะให้ประสิทธิภาพการจำแนกประเภทข้อมูล ได้ดีกว่าโมเดลการเรียนรู้แบบเดี่ยว

ประเด็นที่สอง คือ โมเดลการเรียนรู้แบบรวมกลุ่ม ที่ใช้ตัวจำแนกประเภทที่มีรูปแบบการส่งแบบเดียวกันทั้งหมดและแบบที่ใช้รูปแบบการส่งต่างชนิดกัน พบว่า โมเดลการเรียนรู้แบบรวมกลุ่มที่ใช้ตัวจำแนกประเภทที่มีรูปแบบการส่งต่างชนิดกัน คือ อัลกอริทึม เอ็มอาร์เอ็น ซึ่งเกิดจากการทำงานร่วมกันของตัวจำแนกประเภทที่ใช้รูปแบบการส่งทั้งแบบเชิงเส้น แบบไม่เชิงเส้น และทฤษฎีความน่าจะเป็นนั้น สามารถให้ประสิทธิภาพในการจำแนกประเภทกลุ่มข้อมูลได้ดีกว่าโมเดลที่ใช้ตัวจำแนกประเภทที่มีรูปแบบการส่งแบบเดียวกันทั้งหมด นอกจากนี้ ยังพบว่า อัลกอริทึมแบบ เอ็มเอ็มเอ็ม ซึ่งเป็นโมเดลที่ใช้ตัวจำแนกประเภทแบบเดียวกันทั้งหมด นั่นคือ การใช้โครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น มีค่าความถูกต้องที่ใกล้เคียงกับโมเดลการเรียนรู้แบบรวมกลุ่มซึ่งใช้ตัวจำแนกประเภทต่างชนิดกัน นั่นแสดงให้เห็นว่า ลักษณะการทำงานของโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้นมีประสิทธภาพสูงในการจำแนกประเภทของข้อมูลดีกว่าแบบอื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากข้อสังเกตที่กล่าวมาข้างต้น การใช้โมเดลการเรียนรู้แบบรวมกลุ่มสามารถเพิ่มประสิทธิภาพการเรียนรู้ในการจำแนกประเภทกลุ่มของข้อมูลได้ และในการทดลองนี้ พบว่า อัลกอริทึม เอ็มอาร์เอ็น ให้ค่าความถูกต้องที่สูงที่สุด ในทุกชุดข้อมูลที่นำมาใช้ทดสอบ ทั้งแบบสองคลาสและแบบหลายคลาส ซึ่งแสดงให้เห็นว่า อัลกอริทึมแบบใหม่ที่นำเสนอมานี้ มีความเหมาะสมที่จะใช้ในการจำแนกประเภทของข้อมูลได้อย่างมีประสิทธิภาพ

5.2 ข้อเสนอแนะ

ในการทดลองกับข้อมูลขนาดใหญ่ เมื่อมีข้อมูลชุดใหม่ที่ต้องการเรียนรู้เพิ่มเติม อัลกอริทึมแบบใหม่นี้จะต้องทำการเอาข้อมูลใหม่มารวมกับข้อมูลเก่าแล้วทำการเรียนรู้ใหม่อีกครั้งเพื่อจำแนกประเภทของข้อมูลชุดใหม่ ทำให้เสียเวลาในการเรียนรู้ ดังนั้น จึงควรมีการศึกษาและพัฒนาอัลกอริทึมนี้ ให้สามารถเรียนรู้แบบเพิ่มเติมได้ โดยไม่ต้องเสียเวลาในการเรียนรู้ใหม่ หรือที่เรียกว่า Incremental Learning

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] R. Polikar, "**Bootstrap-Inspired Techniques in Computation Intelligence**", IEEE Trans. Signal Processing Magazine, vol.24, pp. 59 – 72, July 2007.

[2] J. Han, M. Kamber and J. Pei, "**Data Mining Concepts and Techniques**", Third Edition, Elsevier, 2012.

[3] A. Zagorecki, "**Feature Selection for Naive Bayesian Network Ensemble using Evolutionary Algorithms**", IEEE Computer Science and Information Systems, pp. 381 – 385, Sept. 2014.

[4] T. Windeatt, "**Accuracy/Diversity and Ensemble MLP Classifier Design**", IEEE Trans. Neural Networks, vol.17, pp. 1194 – 1211, Sept 2006.

[5] T. Windeatt, R. Duangsoithong and R. Smith, "**Embedded Feature Ranking for Ensemble MLP Classifiers**", IEEE trans. Neural Networks, vol.22, pp. 988 – 994, May 2011.

[6] L. Pan and X. Li, "**Improved RBF Neural Network Ensemble Prediction Model for PMI**", IEEE Computer Science & Service System, pp. 2194 – 2196, Aug. 2012.

[7] R. Ade and P.R. Deshmukh, "**An incremental ensemble of classifiers as a technique for prediction of student's career choice**", IEEE Networks & Soft Computing, pp. 384 – 387, Aug. 2014.

[8] M. Luckner and K. Szyszko, "**RBF ensemble based on reduction of DAG structure**", IEEE Computer Science and Information Systems, pp. 99 – 105, Sept. 2013.

[9] L. Wang and J. Wu, "**Application of Hybrid RBF Neural Network Ensemble Model Based on Wavelet Support Vector Machine Regression in Rainfall Time Series Forecasting**" IEEE Computational Sciences and Optimization, pp. 867 – 871, June 2012.

[10] S. Nikolic, M. Knezevic, V. Ivancevic and I. Lukovic, "**Building an ensemble from a single Naive Bayes classifier in the analysis of key risk factors for Polish State Fire Service**", IEEE Computer Science and Information Systems, pp. 361 – 367, Sep. 2014.

[11] T. Suwan, S. Jaiyen and R. Wiangsripanawan, "**Edited audio detection using ensemble learning**", IEEE Knowledge and Smart Technology, pp. 71 – 74, Jan. 2015.

[12] A. Margoosian and J. Abouei, "**Ensemble-based classifiers for cancer classification using human tumor microarray data**", IEEE Electrical Engineering, pp. 1 – 6, May. 2013.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[13] R. Diao and Q. Shen, "**Fuzzy-rough classifier ensemble selection**" IEEE Fuzzy Systems, pp. 1516 – 1522, June 2011.

[14] H. Woo and C. Park, "**Semi-supervised Ensemble Learning Using Label Propagation**", IEEE Computer and Information Technology, pp. 421 – 426, Oct. 2012.

[15] U. Bhowan, M. Johnston, M. Zhang and X. Yao, "**Reusing Genetic Programming for Ensemble Selection in Classification of Unbalanced Data**", IEEE Trans. Evolutionary Computation, vol.18, pp. 893 – 908, Nov. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ผลงานที่ได้รับการตีพิมพ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

JCSSE 2015

2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE)

"Shaping the Future with Convergence"

22 – 24 July 2015

Prince of Songkla University, Hat Yai, Songkhla, Thailand

Workshop on e-Science and High Performance Computing (eHPC 2015)

IEEE Catalog Number : CFP1532P-USB

ISBN : 978-1-4799-1965-9

SIPA

IEEE

IEEE THAILAND SECTION

Technical Co-Sponsor ECTI Association

SIFT

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเอกสารไปเพื่อการศึกษาหรือการใช้งานเชิงพาณิชย์ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A New Ensemble Model based on Linear Mapping, Nonlinear Mapping, and Probability Theory for Classification Problems

Anusorn Charleonnan and Saichon Jaiyen
Department of Computer Science, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Thailand
s3650806@kmitl.ac.th, kjsaicho@kmitl.ac.th

***Abstract*—Currently, various perspectives of neural networks are proposed for solving classification problems. Some of them are based on two types of mapping functions, namely, linear and nonlinear, for mapping an input space into a feature space. In addition, some neural networks are proposed based on probability theory. Since some models are appropriated for some kinds of data, depending on a distribution of the data, some data are appropriated for linear mapping, some is for nonlinear mapping, and some is for probabilistic models. Due to the fact that the data distribution in classification problems are various, we propose the new ensemble model based on linear mapping, nonlinear mapping, and probability theory for classification problems. According to the experimental results, they have shown that our proposed model can improve the accuracy of classification on the testing data sets.**

Keywords-Ensemble; AdaBoost; Multilayer Perceptron Neural Network (MLP); Naive Bayes; Radial Basis Function Neural Network (RBF).

I. INTRODUCTION

An ensemble learning is a combination of classifiers used to improve the performance of classification models. Windeatt [3] proposed a measurement for achieving optimal performance of ensemble MLP classifier. Suwan, Jaiyen, and Wiangsripanawan [4] applied ensemble learning for detecting edited audios by using the ensemble of decision trees based on AdaBoost algorithm. From the results of this method, it can detect edited audios with very high accuracy. Margoosian and Abouei [5] presented Ensemble-based classifiers for classifying cancer on multi-category microarray data sets. This ensemble learning used Naïve Bayes and k-nearest neighbor classifiers for creating the ensemble model. Diao and Shen [6] proposed a new approach which was ensemble classifier based on fuzzy-rough feature selection and harmony search. The fuzzy-rough was used for measuring the dependency of features in order to select a minimal subset of features in the data set. Woo and Park [7] presented Semi-supervised ensemble learning using label propagation in the data samples for unlabeled data set. Nikolic, Knezevic and Ivancevic [8] suggested the ensemble of Naive Bayes classifiers for predicting incidents related to firefighters. This method gives high prediction performance, which is more than 90% of accuracy. Bhowan, Johnston, Zhang and Yao [9] proposed classification algorithms evolving ensemble classifier using genetic programming (GP) to automatically find the best individuals for solving unbalanced data. The performance of this algorithm is higher than that of other bagging and boosting approaches. Windeatt, Duangsoithong and Smith [10] suggested the feature ranking method for ensemble multilayer perceptron (MLP) to solve multiclass classification problems with feature ranking. Pan and Li [11] proposed method for improving the ensemble of RBF Neural Networks. This method was used to improve prediction accuracy of PMI. Ade and Deshmukh [12] presented an incremental ensemble of classifiers for predicting the career of students. These ensemble algorithms consist of three classifiers, including Naïve Bayes, K-Star and SVM classifiers. The ensemble model used the voting scheme for improving the prediction accuracy of this data. Luckner and Szyszko [13] proposed an ensemble structure of RBF classifiers. Under the structure of this method, the number of classifiers can be reduced. The distances between classes was used to measure the similarity between classes. Wang and Wu [14] suggested hybrid RBF neural network ensemble model based on Wavelet Support Vector Machine Regression in rainfall time series forecasting. For creating the ensemble, it used Bagging and Boosting technique. Zagorecki [15] proposed feature selection method for Naive Bayesian network ensemble using evolutionary algorithms. This method used a genetic algorithm for identifying an optimal set of variables used in the classifiers.

Ensemble learning is a popular method for solving classification problems. The ensemble model includes many weak classifiers in order to improve the accuracy of classifiers. In real data set, it may not appropriate for only one type of weak classifier. In this paper, we propose a new ensemble model based on linear and nonlinear mapping and Probability theory for classification problems. The proposed ensemble model consists of three types of weak classifiers, including Multilayer Perceptron (MLP) and Radial Basis Function (RBF) neural networks, and Naive Bayes. Due to the fact that MLP neural network can perform the linear mapping from input space to feature space, RBF neural network can perform nonlinear mapping, and Naïve Bay apply the probability theory to build the model, the new ensemble model is proposed based

978-1-4799-1966-6/15/$31.00 ©2015 IEEE

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

on these three models. In addition, AdaBoost.M1 algorithm is adopted as the method for creating the ensemble model. For performance evaluation, the experimental results of the proposed model is compared with the individual Multilayer Perceptron Neural Network, Radial Basis Function Neural Network and Naive Bayes Classifier.

II. The Related Methods

A. Multilayer Perceptron Neural Network.

Artificial neural network is one of the most important disciplines in machine learning. The artificial neural network is a computational model that simulates a mechanism of human brain in order to make computers learn from a given data for solving a specific problem. Multilayer perceptron (MLP) neural network is a widely used neural network in data mining due to the reliable and efficient ability in classification. The Back propagation learning algorithm is used to train MLP neural network. The process of Back-propagation learning algorithm is described in Algorithm 1.

Algorithm 1: Back-propagation algorithm

1. Input the training data (x_i, y_i) to the MLP neural network and compute the network outputs.
2. For each output neuron j, calculate the local gradient δ_j defined by
$$\delta_j = e_j \varphi'(V_j)$$
where $\varphi(.)$ is the activation function, $v_j = \sum_i w_{ji} y_i$ and $y_i = \varphi(v_j)$.
3. For each hidden neuron j, calculate the local gradient $\delta_j$defined by
$$\delta_j = \varphi'(v_j) \sum_k \delta_k(w_{kj})$$
4. Update each weight $w_{ji}$of the network by
$$w_{ji}^{new} = w_{ji}^{old} + \Delta w_{ji}$$
where $\Delta w_{ji} = \eta \delta_j x_i$ and η is the learning rate.

B. Radial Basis Function Neural Network.

Radial basis function (RBF) neural network is designed to perform a nonlinear mapping from the input space to the hidden space and a linear mapping from the hidden space to the output space. The structure of RBF neural network consists of an input layer, a hidden layer, and an output layer. The input layer consists of source nodes whose number of nodes is equal to the number of data dimensions. The hidden layer consists of hidden neurons that use the basis function as its kernel function. The output layer gives the response of the network for input patterns applied to the input layer. The basis function can be defined as

$$\emptyset(r) = e^{-(\varepsilon r)^2}$$

The response of RBF neural network can be calculated as

$$f(x) = \sum_{i=1}^{n} w_i \emptyset(||X - X_i||)$$

where $f(x)$ is the output of RBF neural network and w_i is the weight of RBF neural network that can be estimated using the matrix methods.

C. Naive Bayes Classifier.

Naïve Bayes (NB) classifier is based on Bayes theorem. To classify an unseen data, Naïve Bayes classifier computes the posterior probability for each class C_i as:

$$P(C_i|X) = \frac{P(C_i) \prod_{k=1}^{n} P(X_k|C_i)}{P(X)}$$

where C_i is the class label and X is an instance to be classified. Finally, it assigns the class label that has the highest probability to the unseen data.

III. The Proposed Method

AdaBoost.M1 algorithm is used to build the ensemble of classifiers in order to solve the multiclass classification problems [2]. In this paper, we propose the new ensemble model based on linear and nonlinear mapping, and probability theory for classification problems. Therefore, we create the ensemble of MLP, RBF, and Naïve Bayes classifiers, as the proposed MRN, by the AdaBoost.M1 algorithm. The proposed ensemble learning algorithm can be described in Algorithm 2.

Algorithm 2: MRN learning algorithm

Input:

Training data S = $\{x_1, \ldots, x_n\}$, $x_i \in X$ with correct class $y_i \in \Omega, \Omega = \{y_1, \ldots, y_C\}$

Out put:

$$H(x) = \arg\max_{y \in \Omega} \sum_{t=1}^{T} \ln\left(\frac{1}{\beta_t}\right) [h_t(x) = y]$$

Method:

1. Initialize the distribution $D_1(i) = 1/n$, i=1,2,…,n.
2. Train MLP weak learner with training set to receive the hypothesis $h_a : X \rightarrow \Omega$.
3. Compute the error of h_a:
$$\varepsilon_a = \sum_{i=1}^{n} I[h_a(x_i) \neq y_i] \cdot D_i$$
where ε_a is error of h_a and$y_i$is the desired class.
4. If $\varepsilon_a > 0.5$ then go back to step 2.
5. Calculate the normalized error
$$\beta_a = \varepsilon_a/(1 - \varepsilon_a)$$
6. Update the distribution D_a :
$$D_{a+1}(i) = \frac{D_a(i)}{z_a} \times \begin{cases} \beta_a if h_a(x_i) = y_i \\ 1 \ otherwise \end{cases}$$
where z_a is a normalization constant.
7. Train weak learner RBF with training set to receive the hypothesis $h_b : X \rightarrow \Omega$.
8. Calculate the normalized error

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\varepsilon_b = \sum_{i=1}^{n} I[h_b(x_i) \neq y_i] \cdot D_i$$

where ε_b is error of h_b and y_i is the desired class.

9. If $\varepsilon_b > 0.5$ then go back to step 7.
10. Calculate the normalized error
$$\beta_b = \varepsilon_b/(1-\varepsilon_b)$$
11. Update distribution D_i :
$$D_{b+1}(i) = \frac{D_b(i)}{z_b} \times \begin{cases} \beta_b if\, h_b(x_i) = y_i \\ 1\ otherwise \end{cases}$$
where z_b is a normalization constant.
12. Train weak learner Naïve Bayes with training set to receive the hypothesis $h_c : X \rightarrow \Omega$.
$$\varepsilon_c = \sum_{i=1}^{n} I[h_c(x_i) \neq y_i] \cdot D_i$$
where ε_c is error of h_c and y_i is the desired class.
13. If $\varepsilon_c > 0.5$ then go back to step 12.
14. Calculate the normalized error
$$\beta_c = \varepsilon_c/(1-\varepsilon_c)$$
15. Update distribution D_c :
$$D_{c+1}(i) = \frac{D_c(i)}{z_c} \times \begin{cases} \beta_c if\, h_c(x_i) = y_i \\ 1\ otherwise \end{cases}$$
where z_c is a normalization constant.

For the incoming unseen data, the class label of them can be calculated by the following equation.

$$H(x) = \arg\max_{y \in \Omega} \left\{ \ln\left(\frac{1}{\beta_a}\right)[h_a(x) = y] + \ln\left(\frac{1}{\beta_b}\right)[h_b(x) = y] + \ln\left(\frac{1}{\beta_c}\right)[h_c(x) = y] \right\}$$

IV. Expermental Results

The data set used in the experiments is collected from Machine Learning Repository (UCI) web site. Seven data sets are chosen for evaluating the performance of the proposed model. The details of these data set are revealed in Table I.

TABLE I. THE DATA SETS USED IN THE EXPERIMENT

Data set	Instance	Attribute	Class
Qualitative Bankruptcy	250	6	2
Banknote authentication Set	1,372	4	2
Breast Cancer Wisconsin (Diagnostic)	569	29	2
Iris	150	4	3
Wine	178	13	3
Balance Scale	625	4	3
Lung Cancer	32	54	3

In this paper, the performance of the proposed MRN ensemble model is evaluated and compared with MLP neural network, RBF neural network, and Naïve Bay classifier using the data set shown in Table I.

A. Multiclass Classification Problems

In this experiment, we apply the proposed MRN model for solving the multiclass classification problems using Iris data set, Wine data set, Balance Scale data set, and Lung Cancer data set. For MLP weak learner in MRN, the number of epochs are set as 1000. The number of hidden neurons is set as 10. For RBF week learner, the spread distance of radial basis function is set as 0.5. This experiment employs 3-fold cross validation to train and test all classifier models. The data set is divided into three subsets. Then, two subsets are used as a training set and another subset is used as the testing set. For multiclass classification problems, the comparative results between the proposed model, and the individual MLP neural network, RBF neural network, and Naïve Bay classifier are shown in Figure 1, 2, 3, and 4, respectively.

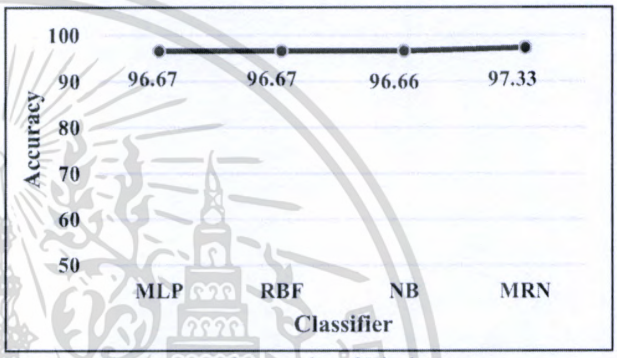


Figure 1. The comparative results of Iris data set

Figure1 shows the accuracies of all methods. From the experimental result in Figure 1, it can be seen that the accuracy of the proposed MRN model is slightly higher than that of MLP, RBF and NB classifiers. The MRN model can achieve the accuracy of 97.33% whereas both MLP and RBF achieve the accuracy of 96.67% and NB classifier achieves the accuracy of 96.66%.

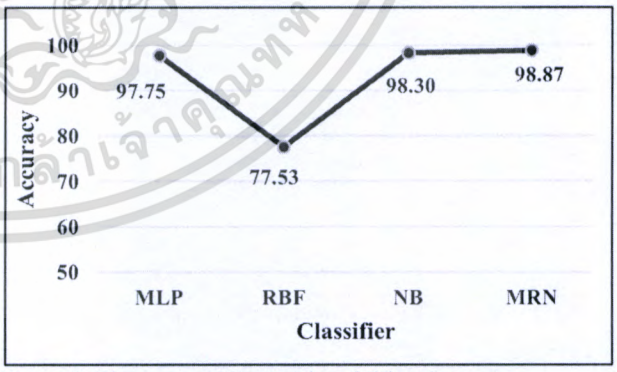


Figure 2. The comparative results of Wine data set

Figure 2 demonstrates the experimental results of the proposed MRN model under Wine data set. The results of the propose model is compared with MLP, RBF and NB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

classifiers. It has shown that the MRN model can achieve the highest accuracy at 98.87% whereas MLP, RBF, and NB classifiers can achieve the accuracies at 97.75%, 77.53 %, and 98.30%, respectively. It can be noted that the accuracy of the proposed MRN model is slightly higher than the accuracy of MLP and NB classifiers but very higher than the accuracy of RBF classifier.

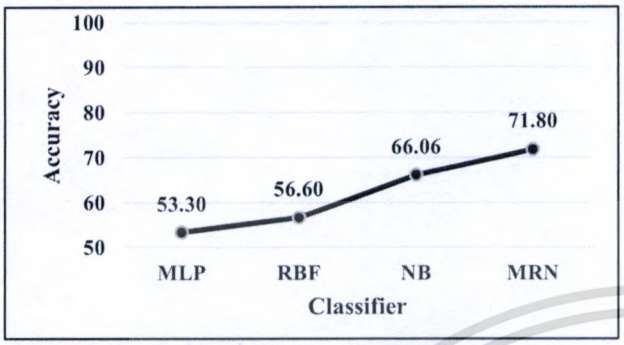


Figure 3. The comparative results of Lung Cancer data set

Figure 3 demonstrates the experimental results of the proposed MRN model under Lung Cancer data set. The result of the propose model is compared with MLP, RBF and NB classifiers. The accuracy of the MRN Model can achieve the highest accuracy at 71.80% whereas MLP, RBF, and NB classifiers can achieve the accuracies at 53.30%, 56.60%, and 66.06%, respectively.

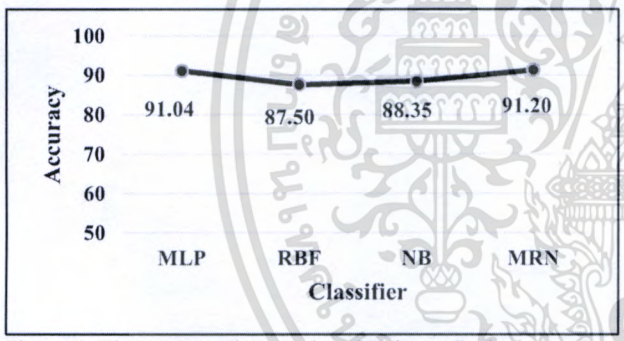


Figure 4. The comparative results of Balance Scale data set

Figure 4 shows the accuracies of all methods under Balance Scale data set. From the experimental results in Figure 4, it can be seen that the accuracy of the proposed MRN model is slightly higher than the accuracy of MLP classifiers but higher than RBF and NB classifiers. Accordingly, the MRN model can achieve the accuracy of 91.20% whereas MLP achieves the accuracy of 91.04%, RBF achieves the accuracy of 87.50% and NB classifier achieves the accuracy of 88.35%.

B. Two Class Classification

In addition, the two-class classification problems under Qualitative Bankruptcy, Cancer Wisconsin (Diagnostic), and Liver data sets are chosen for evaluating the performance of the proposed MRN model. The details of setting up parameters is the following. For MLP weak classifier, the number of epochs is set as 1000, the hidden neurons are set as 10. For RBF weak learner, the spread distance of radial basis function is set as 0.5. The comparative results are demonstrated in Figure 5, 6, and 7, respectively.

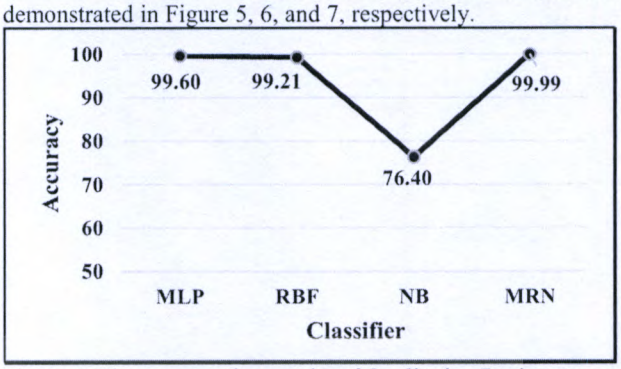


Figure 5. The comparative results of Qualitative Bankruptcy data set

Figure 5 demonstrates the experimental results of the proposed MRN model under Qualitative Bankruptcy data set. The performance of the proposed model is compared with the individual MLP, RBF and NB classifiers. The accuracy of the MRN model can achieve the highest accuracy at 99.99% whereas MLP, RBF, and NB classifiers can achieve the accuracies at 99.60%, 99.21%, and 76.40%, respectively. It can be noted that the accuracy of the proposed MRN model is slightly higher than the accuracy of MLP and RBF classifiers but very higher than the accuracy of NB classifier.

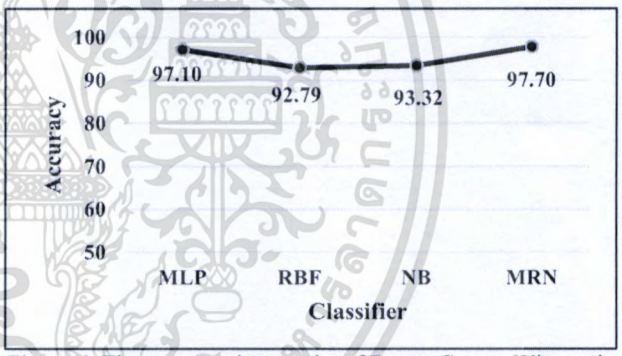


Figure 6. The comparative results of Breast Cancer Wisconsin (Diagnostic) data set.

Figure 6 shows the accuracies of all methods under Breast Cancer Wisconsin data set. From the experimental results in Figure 6, it can be seen that the accuracy of the proposed MRN model is slightly higher than the accuracy of MLP classifiers but higher than RBF and NB classifiers. The MRN model can achieve the accuracy of 97.70% whereas MLP achieves the accuracy of 97.10%, RBF achieves the accuracy of 92.79%, and NB classifier achieves the accuracy of 93.32%, respectively.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

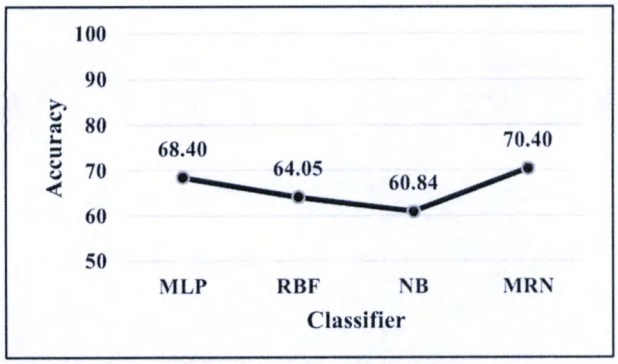


Figure 7. The comparative results of Liver data set

Figure 7 demonstrates the experimental results of the proposed MRN model under Liver data set. The performance of the propose model is compared with MLP, RBF and NB classifiers. The accuracy of the MRN model can achieve the highest accuracy at 70.40% whereas MLP, RBF, and NB classifiers can achieve the accuracies at 68.40%, 64.05%, and 60.84%, respectively.

V. Conclusion.

In this paper, we propose a new ensemble model based on linear, nonlinear mapping functions, and probability theory. Therefore, MLP and RBF are adopted for performing linear and nonlinear, respectively, mapping from input space into the feature space, while Naïve Bayes classifier is adopted for performing probabilistic theory. Consequently, we create the ensemble of MLP classifier, RBF classifier, and Naïve Bayes classifier (MRN) by using Adaboost.M1 algorithm for classification problems. For performance evaluation, the proposed MRN model is compared with the individual classifiers, including MLP, RBF, and Naïve Bayes classifiers, under two class and multiclass classification problems. From the experimental results, it reveals that the proposed MRN ensemble model can achieve the best accuracy in all data set.

References

[1] J. Han, M. Kamber and J. Pei, "Data Mining Concepts and Techniques", Third Edition, Elsevier, 2012.

[2] R. Polikar, "Bootstrap-Inspired Techniques in Computation Intelligence", IEEE Trans. Signal Processing Magazine, vol.24, pp. 59 – 72, July 2007.

[3] T. Windeatt, "Accuracy/Diversity and Ensemble MLP Classifier Design", IEEE Trans. Neural Networks, vol.17, pp. 1194 – 1211, Sept 2006.

[4] T. Suwan, S. Jaiyen and R. Wiangsripanawan, "Edited audio detection using ensemble learning", IEEE Knowledge and Smart Technology, pp. 71 – 74, Jan. 2015.

[5] A. Margoosian and J. Abouei, "Ensemble-based classifiers for cancer classification using human tumor microarray data", IEEE Electrical Engineering, pp. 1 – 6, May. 2013.

[6] R. Diao and Q. Shen, "Fuzzy-rough classifier ensemble selection" IEEE Fuzzy Systems, pp. 1516 – 1522, June 2011.

[7] H. Woo and C. Park, "Semi-supervised Ensemble Learning Using Label Propagation", IEEE Computer and Information Technology, pp. 421 – 426, Oct. 2012.

[8] S. Nikolic, M. Knezevic, V. Ivancevic and I. Lukovic, "Building an ensemble from a single Naive Bayes classifier in the analysis of key risk factors for Polish State Fire Service", IEEE Computer Science and Information Systems, pp. 361 – 367, Sep. 2014.

[9] U. Bhowan, M. Johnston, M. Zhang and X. Yao, "Reusing Genetic Programming for Ensemble Selection in Classification of Unbalanced Data", IEEE Trans. Evolutionary Computation, vol.18, pp. 893 – 908, Nov. 2014.

[10] T. Windeatt, R. Duangsoithong and R. Smith, "Embedded Feature Ranking for Ensemble MLP Classifiers", IEEE trans. Neural Networks, vol.22, pp. 988 – 994, May 2011.

[11] L. Pan and X. Li, "Improved RBF Neural Network Ensemble Prediction Model for PMI", IEEE Computer Science & Service System, pp. 2194 – 2196, Aug. 2012.

[12] R. Ade and P.R. Deshmukh, "An incremental ensemble of classifiers as a technique for prediction of student's career choice", IEEE Networks & Soft Computing, pp. 384 – 387, Aug. 2014.

[13] M. Luckner and K. Szyszko, "RBF ensemble based on reduction of DAG structure", IEEE Computer Science and Information Systems, pp. 99 – 105, Sept. 2013.

[14] L. Wang and J. Wu, "Application of Hybrid RBF Neural Network Ensemble Model Based on Wavelet Support Vector Machine Regression in Rainfall Time Series Forecasting" IEEE Computational Sciences and Optimization, pp. 867 – 871, June 2012.

[15] A. Zagorecki, "Feature Selection for Naive Bayesian Network Ensemble using Evolutionary Algorithms", IEEE Computer Science and Information Systems, pp. 381 – 385, Sept. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ – สกุล	นายอนุสรณ์ เจริญนาน
วัน เดือน ปีเกิด	3 มิถุนายน 2517
ที่อยู่	78/1 เทศบาลสาย 3 ต.ขลุง อ.ขลุง จ.จันทบุรี 22110
ประวัติการศึกษา	พ.ศ. 2540 จบการศึกษาปริญญาตรีวิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันราชภัฏรำไพพรรณี
ผลงานทางวิชาการ	1. A New Ensemble Model based on Linear Mapping, Nonlinear Mapping, and Probability Theory for Classification Problems

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้