

การเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ด้วยเครื่องจาก
ตัวแบบถดถอยเชิงเส้นเมื่อข้อมูลมีมิติขั้นสูง

A COMPARISON OF PERFORMANCE OF MACHINE
LEARNING METHODS FROM LINEAR
REGRESSION MODELS USING
HIGH-DIMENSIONAL DATA



ณัฐพล สุมาลัย
ปานวาด เนียมหอม
สุทธราลีณี ศรีอนุชา

ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต (สถิติประยุกต์)

ภาควิชาสถิติ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A COMPARISON OF PERFORMANCE OF MACHINE
LEARNING METHODS FROM LINEAR
REGRESSION MODELS USING
HIGH-DIMENSIONAL DATA



A SPECIAL PROBLEM SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR
THE DEGREE OF BACHELOR OF SCIENCE (APPLIED STATISTICS)
DEPARTMENT OF STATISTIC, SCHOOL OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2023

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปัญหาพิเศษ

การเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ด้วยเครื่องจากตัวแบบถดถอยเชิงเส้นเมื่อข้อมูลมีมิติขั้นสูง

A Comparison of Performance of Machine Learning Methods from Linear Regression Models Using High-Dimensional Data

ชื่อนักศึกษา

นายณัฐพล สุมาลัย รหัสนักศึกษา 63050619

นางสาวปานวาด เนียมหอม รหัสนักศึกษา 63050644

นางสาวสุทธาสินี ศรีอนุชา รหัสนักศึกษา 63050679

ปริญญา

วิทยาศาสตรบัณฑิต (สถิติประยุกต์)

ภาควิชา

สถิติ

ปีการศึกษา

2566

อาจารย์ที่ปรึกษา

รศ.ดร.อัชฌา อระวีพร

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต(สถิติประยุกต์) ประจำปีการศึกษา 2566

คณะกรรมการสอบ	ลายมือชื่อ
ผศ.ดร.สมศรี บัณฑิตวิไล ประธานกรรมการ	
ดร.ธัญรดา ชัยขจรวัฒน์ กรรมการ	
รศ.ดร.อัชฌา อระวีพร กรรมการและอาจารย์ที่ปรึกษา	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปัญหาพิเศษ	การเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ด้วยเครื่องจากตัวแบบถดถอยเชิงเส้นเมื่อข้อมูลมีมิติขั้นสูง	
ชื่อนักศึกษา	นายณัฐพล สุมาลัย	รหัสนักศึกษา 63050619
	นางสาวปานวาด เนียมหอม	รหัสนักศึกษา 63050644
	นางสาวสุทธาสินี ศรีอนุชา	รหัสนักศึกษา 63050679
ปริญญา	วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)	
ภาควิชา	สถิติ	
คณะ	วิทยาศาสตร์	
มหาวิทยาลัย	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.)	
ปีการศึกษา	2566	
อาจารย์ที่ปรึกษา	รศ.ดร.อชฌา อระวีพร	

บทคัดย่อ

ปัญหาพิเศษนี้มีวัตถุประสงค์เพื่อศึกษาการเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ของเครื่องจากตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูง โดยวิธีการเรียนรู้ของเครื่องที่ใช้ในการวิเคราะห์ได้แก่วิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนซ์ โดยทำการเปรียบเทียบค่าสัมประสิทธิ์การถดถอยที่ 1 และ -1 ซึ่งประกอบไปด้วยตัวแปรอิสระสองกลุ่มคือกลุ่มจำนวนตัวแปรอิสระน้อย ได้แก่ 50, 75 และ 100 กลุ่มที่สองคือกลุ่มจำนวนตัวแปรอิสระมาก ได้แก่ 150, 200 และ 250 วิธีดำเนินงานวิจัยจะทำการแบ่งข้อมูลเป็นข้อมูลฝึกฝน 70% และข้อมูลทดสอบ 30% โดยจะใช้โปรแกรมอาร์สตูดิโอในการจำลองและวิเคราะห์ข้อมูลโดยทำซ้ำเป็นจำนวน 1000 รอบและหาวิธีที่ดีที่สุดจากค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุด ผลการวิจัยพบว่าในการเปรียบเทียบค่าสัมประสิทธิ์การถดถอยที่ 1 และ -1 ผลลัพธ์ที่ได้ไม่ได้แตกต่างกันคือโดยส่วนใหญ่วิธีที่ให้ผลลัพธ์ที่ดีที่สุดคือ วิธีซัพพอร์ตเวกเตอร์แมชชีน แต่มีบางกรณีที่มีวิธีอื่นให้ผลลัพธ์ดีกว่าเช่น กรณีที่มีจำนวนตัวแปรอิสระและขนาดตัวอย่างน้อย วิธีที่ให้ผลลัพธ์ดีคือ วิธีค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ และกรณีที่จำนวนตัวแปรอิสระและขนาดตัวอย่างมากวิธีที่ให้ผลลัพธ์ดีคือ วิธีการเรียนรู้เชิงลึก

คำสำคัญ : วิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึก วิธีสโตแคสติกเกรเดียนต์เดสเซนซ์และค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	A Comparison of Performance of Machine Learning Methods from Linear Regression Models Using High-Dimensional Data	
Students	Mr. Nuttapon Sumalai	Student ID 63050619
	Miss Panwad Niamhom	Student ID 63050644
	Miss Sutthasinee Sri-anucha	Student ID 63050679
Degree	Bachelor of Science (Applied Statistics)	
Department	Statistics	
School	Science	
University	King Mongkut's Institute of Technology Ladkrabang (KMITL)	
Academic Year	2023	
Advisor	Assoc. Prof. Dr. Autcha Araveeporn	

Abstract

This special problem aims to study and compare the performance of machine learning from linear regression models using high-dimensional data. Random forest, K-nearest neighbor, support vector machine, deep learning, and stochastic gradient descent analyze these machine learning methods. The comparison of the regression coefficients is considered to be 1 and -1. The independent variables have two groups: a small group of 50, 75, and 100 independent variables and a large group of 150, 200, and 250. The research methods divide the data into a training data set at 70% and a test data set at 30%. The R studio program is used for simulation, and the data is analyzed by repeating 1000 rounds and finding the best method from minimum mean absolute percentage error. The results showed that when comparing the regression coefficients of 1 and -1, the results are similar in most cases. The support vector machine is the best method. In some cases, the other methods give better results, such as the K-Nearest neighbor, which performs well in a few independent variables and has a small sample size. Furthermore, deep learning outperforms in many independent variables and a large sample size.

เอกสารนี้เป็นลิขสิทธิ์ของคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Keywords : Random Forest, k-Nearest-Neighbor, Support Vector Machine, Deep Learning, Stochastic Gradient Descent and Mean absolute percentage error.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปัญหาพิเศษนี้สำเร็จลุล่วงไปได้ด้วยดี เนื่องด้วยได้รับความเมตตากรุณาจาก รศ.ดร.อัชฌา อระวีพร อาจารย์ที่ปรึกษาปัญหาพิเศษ ที่กรุณาสละเวลาให้คำแนะนำ ข้อเสนอแนะ เอื้อเพื่อเอกสารต่าง ๆ อันเป็นประโยชน์ต่อการทำปัญหาพิเศษอย่างยิ่ง อีกทั้งช่วยตรวจทานความถูกต้อง แก้ไขข้อบกพร่อง ตลอดจนติดตามผลงานทุก ๆ ขั้นตอนของการดำเนินงานในการทำปัญหาพิเศษจนกระทั่งเสร็จสมบูรณ์ จึงขอขอบพระคุณด้วยความเคารพอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณ ผศ.ดร.สมศรี บัณฑิตวิไล และ ดร.ธัญรดา ชัยขจรวัฒน์ คณะกรรมการปัญหาพิเศษที่ให้คำปรึกษาและคำแนะนำที่ทำให้ปัญหาพิเศษนี้มีความสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ คณาจารย์ภาควิชาสถิติทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ และช่วยเหลือให้คำแนะนำในเรื่องต่าง ๆ มาโดยตลอด

สุดท้ายนี้ขอขอบพระคุณ บิดามารดาที่คอยสนับสนุน ส่งเสริม และให้กำลังใจผู้ทำปัญหาพิเศษในทุก ๆ ด้านจนทำให้ปัญหาพิเศษนี้สำเร็จลุล่วงได้ตามที่ตั้งใจ ผู้ทำปัญหาพิเศษหวังอย่างยิ่งว่า ปัญหาพิเศษฉบับนี้จะสามารถเป็นประโยชน์และเป็นแนวทางในการพัฒนาต่อไป

ณัฐพล สุมาลัย
ปานวาด เนียมหอม
สุทธาสินี ศรีอนุชา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ง
สารบัญ.....	จ
สารบัญตาราง.....	ช
สารบัญรูป.....	ฉ
บทที่ 1 บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	3
1.3 ขอบเขตของงานวิจัย.....	3
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	5
1.5 นิยามศัพท์เฉพาะ.....	6
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 วิธีป่าสุ่ม (Random Forest).....	7
2.1.1 การถดถอยป่าสุ่ม (Random forest Regression).....	8
2.2 วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ (K Nearest Neighbor : KNN).....	9
2.2.1 ขั้นตอนวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ.....	9
2.2.2 การถดถอยวิธีเพื่อนบ้านใกล้สุด k อันดับ.....	10
2.3 วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM).....	11
2.3.1 ระนาบเกิน (Hyperplane).....	11
2.3.2 ระนาบเกินแบบบัญญัติ (Canonical Hyperplane).....	13
2.3.3 การถดถอยวิธีซัพพอร์ตเวกเตอร์แมชชีน.....	15
2.4 วิธีการเรียนรู้เชิงลึก (Deep Learning).....	16
2.4.1 การถดถอยการเรียนรู้เชิงลึก.....	16
2.4.2 สมการการถดถอยการเรียนรู้เชิงลึก.....	17
2.5 วิธีสโตแคสติกเกรเดียนต์เดสเซนท์ (Stochastic Gradient Descent).....	17
2.6 ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนยกกำลังสอง (Mean Square Error).....	20
2.7 ค่าเฉลี่ยร้อยละความผิดพลาดสัมบูรณ์ (Mean Absolute Percentage Error)....	20
2.8 งานวิจัยที่เกี่ยวข้อง.....	21

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้า
บทที่ 3 วิธีการดำเนินงานวิจัย	22
3.1 การวางแผนงานวิจัย	22
3.2 วิธีดำเนินงานวิจัย.....	23
3.3 ผังงานลำดับขั้นตอนการดำเนินงานวิจัย.....	24
บทที่ 4 ผลการวิจัย	25
4.1 กรณีเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ 1	26
4.1.1 กลุ่มจำนวนตัวแปรอิสระน้อย.....	26
4.1.2 กลุ่มจำนวนตัวแปรอิสระมาก	33
4.2 กรณีเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ -1.....	40
4.2.1 กลุ่มจำนวนตัวแปรอิสระน้อย.....	40
4.2.2 กลุ่มจำนวนตัวแปรอิสระมาก	47
4.3 อภิปรายผล.....	54
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	56
5.1 สรุปผลวิจัย	56
5.2 ข้อเสนอแนะ.....	57
เอกสารอ้างอิง	58
ภาคผนวก.....	60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 ผลการจำแนกประเภทด้วยค่า k ต่างๆ.....	10
4.1 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 1$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	26
4.2 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	28
4.3 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	30
4.4 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนน้อยและค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	32
4.5 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 1$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	33
4.6 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	35
4.7 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	37
4.8 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนมากและค่าสัมประสิทธิ์ การถดถอยเท่ากับ 1	39
4.9 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อย $\mu = 0, \sigma^2 = 1$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	40
4.10 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อย $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง(ต่อ)

ตารางที่	หน้า
4.11 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระน้อย $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	44
4.12 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนน้อยและค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	46
4.13 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมาก $\mu = 0, \sigma^2 = 1$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	47
4.14 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมาก $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	49
4.15 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี จากจำนวนตัวแปรอิสระมาก $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	51
4.16 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนมากและค่าสัมประสิทธิ์ การถดถอยเท่ากับ -1	53
4.17 วิธีที่ดีที่สุดเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ 1	54
4.18 วิธีที่ดีที่สุดเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ -1	54

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
1.1 การแจกแจงปรกติที่มีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 1,5 และ 25.....	4
2.1 หลักการของวิธี Bagging และวิธี Boosting.....	7
2.2 แสดงการแบ่งข้อมูลออกเป็นช่วง	8
2.3 ต้นไม้ตัดสินใจ	8
2.4 ปริภูมิตัวอย่าง.....	9
2.5 ตัวอย่างระนาบเกินในปริภูมิ 1 มิติ.....	11
2.6 ตัวอย่างระนาบเกินในปริภูมิ 2 มิติ.....	12
2.7 ตัวอย่างระนาบเกินในปริภูมิ 3 มิติ.....	12
2.8 ระนาบเกินที่สามารถแบ่งข้อมูล 2 คลาสได้.....	13
2.9 ระนาบเกินที่เกิดปัญหาระนาบเหมาะสมเกินไป	14
2.10 เวกเตอร์ซัพพอร์ต.....	14
2.11 เปรียบเทียบแบบจำลองวิธีซัพพอร์ตเวกเตอร์แมชชีน และวิธีถดถอยซัพพอร์ตเวกเตอร์	15
2.12 โครงข่ายการเรียนรู้เชิงลึก.....	16
2.13 วิธีการทำงานของการเรียนรู้เชิงลึก	17
2.14 วิธีต่างๆของวิธีเกรเดียนเดสเซนท์ทั้ง 3 รูปแบบ	18
2.15 เปรียบเทียบวิธีเกรเดียนเดสเซนท์ (Gradient Descent) และวิธีสโตเคสติก เกรเดียนเดสเซนท์ (Stochastic Gradient Descent).....	19
3.1 แสดงลำดับขั้นตอนการดำเนินงานวิจัย.....	24
4.1 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 ($\mu = 0, \sigma^2 = 1$ และ $\beta_j = 1$)	27
4.2 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 ($\mu = 0, \sigma^2 = 5$ และ $\beta_j = 1$).....	29
4.3 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 ($\mu = 0, \sigma^2 = 25$ และ $\beta_j = 1$).....	31

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.4 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 1$ และ $\beta_j = 1)$	34
4.5 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 5$ และ $\beta_j = 1)$	36
4.6 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 25$ และ $\beta_j = 1)$	38
4.7 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 $(\mu = 0, \sigma^2 = 1$ และ $\beta_j = -1)$	41
4.8 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 $(\mu = 0, \sigma^2 = 5$ และ $\beta_j = -1)$	43
4.9 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 50,75$ และ 100 $n = 20,30$ และ 40 $(\mu = 0, \sigma^2 = 25$ และ $\beta_j = -1)$	45
4.10 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 1$ และ $\beta_j = -1)$	48
4.11 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 5$ และ $\beta_j = -1)$	50
4.12 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่อง ทั้ง 5 วิธี โดยกำหนดให้ $p = 200,250$ และ 300 $n = 50,100$ และ 150 $(\mu = 0, \sigma^2 = 25$ และ $\beta_j = -1)$	52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแปรอิสระ เนื่องจากการสร้างสมการถดถอยเชิงเส้นที่มีการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta} = (X^T X)^{-1} X^T Y$ ซึ่งเมทริกซ์ $(X^T X)^{-1}$ จะเป็นเมทริกซ์เอกฐาน (Singular matrix) จะทำให้ไม่สามารถหาเมทริกซ์ผกผัน (Inverse matrix) ได้ วิฐูรา (2558) ได้เสนอวิธีในการวิเคราะห์ข้อมูลมิติขั้นสูงประกอบด้วยวิธีการถดถอยที่ปรับด้วยฟังก์ชันการลงโทษ (Penalized regression) และ วิธีเบย์ (Bayesian approach) แต่เนื่องจากวิธีดังกล่าวสามารถประยุกต์ใช้ได้กับตัวแปรอิสระที่มีจำนวนไม่มากนักและวิธีการเรียนรู้ของเครื่องเป็นวิธีที่สามารถแก้ไขปัญหาดังกล่าวได้ดี โดยวิธีที่ใช้ในการทำนายโดยการเรียนรู้ของเครื่องประกอบด้วยวิธีป่าสุ่ม (Random Forest) วิธีค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ (k-Nearest-Neighbor) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) วิธีการเรียนรู้เชิงลึก (Deep Learning) และวิธีสโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent)

โสภีและคณะ (2565) ได้ศึกษาตัวแบบการเรียนรู้ของเครื่องที่เหมาะสมสำหรับการทำนายราคาพริกชี้หนูโดยใช้ข้อมูลย้อนหลังของสภาพภูมิอากาศจากสถานีอุตุนิยมวิทยา จังหวัดนครศรีธรรมราชและราคาพริกชี้หนูจากงานพาณิชย์จังหวัดนครศรีธรรมราช พ.ศ. 2553 -2563 โดยใช้ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีการถดถอยเชิงเส้น วิธีค้นหาเพื่อนบ้านใกล้ที่สุด k อันดับ และวิธีป่าสุ่ม ผลการศึกษาพบว่าวิธีป่าสุ่มเป็นตัวแบบที่มีประสิทธิภาพดีที่สุดเนื่องจากให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์น้อยที่สุด จุฬาวัลย์ และศิวา (2563) ได้ศึกษาตัวแบบการเรียนรู้ของเครื่องที่เหมาะสมสำหรับการประมาณค่าความเข้มข้นของตะกอนแขวนลอยในแม่น้ำลำปาวโดยใช้ข้อมูลจากดาวเทียม ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีการถดถอยเชิงเส้น วิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการเรียนรู้เชิงลึก ผลการศึกษาพบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวแบบที่มีประสิทธิภาพดีที่สุด เนื่องจากให้ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองน้อยที่สุด วิรัช และฐิตาภรณ์ (2561) ได้ศึกษาการสร้างแบบจำลองการตั้งจุดการเดินทางเข้าสู่พื้นที่อุทยานโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์โดยใช้ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีวิเคราะห์การถดถอย วิธีโครงข่ายประสาทเทียมและวิธีการเรียนรู้เชิงลึก ผลการศึกษาพบว่าวิธีการเรียนรู้เชิงลึกเป็นตัวแบบที่มีประสิทธิภาพดีที่สุดเนื่องจากให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์น้อยที่สุด Sun, Y.(2021) ทำการตรวจสอบการทำนายราคาบ้านด้วยวิธีเกรเดียนเดสเซนท์แบบต่างๆเช่น วิธีเกรเดียนเดสเซนท์แบบทั่วไป (General Gradient Descent) วิธีสโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent) และวิธีมินิแบชเกรเดียนเดสเซนท์ (Mini-Batch Gradient Descent) จากการเปรียบเทียบค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองน้อยพบว่าวิธีสโตแคสติกเกรเดียนเดสเซนท์เป็นตัวแบบที่มีประสิทธิภาพดีที่สุด Biau et al. (2012) ได้เสนอวิธีค้นหาเพื่อนบ้านใกล้ที่สุด k อันดับ โดยการจำลองชุดข้อมูลขึ้นและใช้วิธีค้นหาเพื่อนบ้านใกล้ที่สุด k อันดับในการประมาณค่าสัมประสิทธิ์การถดถอย ซึ่งวิธีการดังกล่าวสามารถแก้ไขปัญหาที่ต้องทำการเปลี่ยนแปลงข้อมูลก่อนทำการวิเคราะห์การถดถอย

เชิงเส้นส่งผลให้ลดความซับซ้อนในการสร้างตัวแบบการถดถอยเชิงเส้น
เอกสารนี้เป็นลิขสิทธิ์ของนิตยสารวิชาการเพื่อการศึกษาเพื่อให้นักศึกษาได้หันไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับปัญหาพิเศษนี้ผู้วิจัยมีความสนใจที่จะศึกษาการเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ของเครื่องจากตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูง ซึ่งวิธีการจะประกอบไปด้วยวิธีการปาสุ่ม วิธีการค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์ โดยใช้ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสอง (MSE) และค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) เป็นเกณฑ์ในการวัดประสิทธิภาพของแบบจำลองเพื่อหาแบบจำลองที่ดีที่สุดสำหรับข้อมูลมิติขั้นสูง

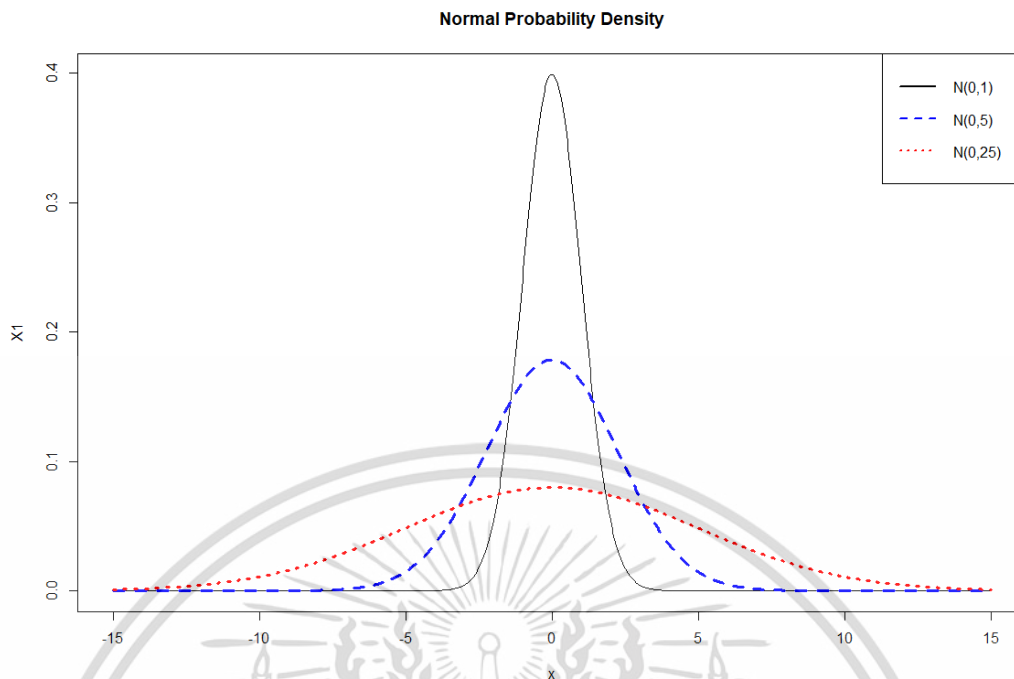
1.2 วัตถุประสงค์

1.2.1 ศึกษาวิธีการทำนายตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่องประกอบไปด้วยวิธีการปาสุ่ม วิธีการค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์

1.2.2 ทำการเปรียบเทียบประสิทธิภาพตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่องประกอบไปด้วยวิธีการปาสุ่ม วิธีการค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์

1.3 ขอบเขตของงานวิจัย

1.3.1 จำลองค่าคลาดเคลื่อน (Error, ε) ของตัวแบบการถดถอยแบบพหุคูณจากข้อมูลที่มีการแจกแจงปรกติ ซึ่งมีค่าเฉลี่ย $E(\varepsilon) = \mu$ และ $Var(\varepsilon) = \sigma^2$ งานวิจัยนี้กำหนดให้ค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1,5 และ 25



รูปที่ 1.1 การแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 1,5 และ 25

1.3.2 กำหนดขนาดตัวอย่าง (n) และจำนวนตัวแปรอิสระ (p) ให้อยู่ในรูปแบบข้อมูลมีมิติ
สูง โดยเป็น 2 กลุ่มประกอบไปด้วยกลุ่มที่มีจำนวนตัวแปรอิสระน้อยและจำนวนตัวแปรอิสระมาก
ดังนี้

กลุ่มจำนวนตัวแปรอิสระน้อย

$$p = 50, n = 20, 30 \text{ และ } 40$$

$$p = 75, n = 20, 30 \text{ และ } 40$$

$$p = 100, n = 20, 30 \text{ และ } 40$$

กลุ่มจำนวนตัวแปรอิสระมาก

$$p = 200, n = 50, 100 \text{ และ } 150$$

$$p = 250, n = 50, 100 \text{ และ } 150$$

$$p = 300, n = 50, 100 \text{ และ } 150$$

1.3.3 สร้างตัวแปรอิสระ (X) จากการแจกแจงปกติจำนวน p

1.3.4 กำหนดค่าค่าคงที่ของสมการถดถอย β_0 เป็น 1000

1.3.5 กำหนดค่าสัมประสิทธิ์การถดถอย β_j ของตัวแบบถดถอยเชิงเส้นพหุเป็น 1 และ -1

โดยที่ $j = 1, 2, \dots, p$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3.6 นำข้อมูลจาก 1.3.1 - 1.3.5 มาสร้างตัวแปรตาม

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i \text{ โดยที่ } i = 1, 2, \dots, n \text{ และ } j = 1, 2, \dots, p$$

เมื่อ Y_i คือ ค่าของตัวแปรตาม

x_{ij} คือ ค่าของตัวแปรอิสระ

β_0 คือ ค่าคงที่ (Constant) ของสมการถดถอย

β_j คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficient) ของตัวแปรอิสระ

ε_i คือ ค่าความคลาดเคลื่อน (Error or Residual)

1.3.7 แบ่งข้อมูลเป็นชุดฝึกฝน (Training Data Set) และชุดทดสอบ (Test Data set) เป็นจำนวน 70 เปอร์เซ็นต์และ 30 เปอร์เซ็นต์ของข้อมูลทั้งหมดตามลำดับ

1.3.8 ชุดข้อมูลฝึกฝนใช้ในการคำนวณค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสอง

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

เมื่อ n คือ ขนาดตัวอย่าง

Y_i คือ ค่าที่สังเกตได้จริงของตัวแปรตาม

\hat{Y}_i คือ ค่าที่ทำนายของตัวแปรตามที่ได้รับจากการเรียนรู้ของเครื่อง

1.3.9 ชุดข้อมูลทดสอบใช้ในการคำนวณค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \right) \times 100$$

เมื่อ n คือ ขนาดตัวอย่าง

Y_i คือ ค่าที่สังเกตได้จริงของตัวแปรตาม

\hat{Y}_i คือ ค่าที่ทำนายของตัวแปรตามที่ได้รับจากการเรียนรู้ของเครื่อง

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 สามารถนำผลการเปรียบเทียบประสิทธิภาพที่ได้ไปใช้เป็นแนวทางในการวิเคราะห์ตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยการเรียนรู้ของเครื่องได้อย่างเหมาะสม

1.4.2 สามารถนำวิธีการวิเคราะห์ตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยการเรียนรู้ของเครื่องจากการศึกษาครั้งนี้ไปประยุกต์ใช้กับปัญหาในชีวิตจริง เช่น ข้อมูลทางการแพทย์ที่มีตัวแปรอิสระเป็นจำนวนมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 นิยามศัพท์เฉพาะ (สำนักงานราชบัณฑิตยสภา, 2561)

1.5.1 การถดถอย (Regression) หมายถึง วิธีการทางสถิติที่ใช้วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระชุดหนึ่งโดยมีจุดมุ่งหมายที่จะอธิบายหรือทำนายตัวแปรตาม ถ้าตัวแปรสุ่ม Y ประกอบด้วยส่วนประกอบ 2 ส่วน ส่วนแรกขึ้นกับตัวแปร X และส่วนที่สองขึ้นกับค่าคลาดเคลื่อนสุ่ม (ε) กล่าวคือ $Y = f(X) + \varepsilon$ แล้ว การถดถอยของ Y บน X คือ สมการที่ได้จากการหาค่าคาดหมายของสมการข้างต้น เมื่อตรึงค่าตัวแปร X ทั้งนี้เรียก X ว่า ตัวแปรอิสระ ตัวแปรอธิบาย ตัวแปรทำนาย หรือตัวแปรถดถอย ส่วนตัวแปรสุ่ม Y คือ ตัวแปรตาม

1.5.2 การเรียนรู้ของเครื่อง (Machine learning) หมายถึง สาขาหนึ่งของปัญญาประดิษฐ์ที่พัฒนามาจากการศึกษาการรู้จำแบบ เกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยโมเดลที่สร้างมาจากชุดข้อมูลตัวอย่างเพื่อการทำนายหรือตัดสินใจโดยไม่ต้องเขียนโปรแกรมไว้อย่างชัดเจน การเรียนรู้ของเครื่องต้องอาศัยวิธีการทางสถิติศาสตร์เป็นอย่างมาก โดยเชื่อมโยงกับองค์ความรู้อื่นที่ต้องการนำไปประยุกต์ เช่น การแพทย์ ชีวสารสนเทศศาสตร์ เคมีสารสนเทศศาสตร์

1.5.3 ข้อมูลมหัต (Big data) หมายถึง เซตข้อมูลขนาดใหญ่ที่มีความซับซ้อนมาก ประกอบด้วยลักษณะอย่างน้อย 3 ประการคือ มีปริมาตร (Volume) มาก มีการเปลี่ยนแปลงอยู่ตลอดเวลาอย่างต่อเนื่อง (Velocity) มีความหลากหลายในโครงสร้างข้อมูล (Variety) นอกจากนี้ อาจมีองค์ประกอบอื่นเพิ่มเติม เช่น มีความแปรผันสูงและมีความซับซ้อนมาก เนื่องจากอยู่ในรูปแบบต่างๆ ทั้งที่มีโครงสร้างที่แน่นอน เช่น เมทริกซ์ข้อมูล หรือไม่มีโครงสร้าง เช่น ข้อความ อีเมล วิดีทัศน์ หรือรวบรวมมาจากหลายแหล่ง

1.5.4 การถดถอยเชิงเส้น (Linear Regression) หมายถึง การถดถอยที่อยู่ในรูปของสมการเชิงเส้นในพจน์ของพารามิเตอร์

1.5.5 ชุดข้อมูลทดสอบ (Test Data Set) หมายถึง ข้อมูลที่สุ่มจากชุดข้อมูลที่ทำการศึกษาวิเคราะห์ซึ่งไม่ได้ใช้ในการพัฒนาตัวแบบและตรวจสอบตัวแบบ แต่นำมาทดสอบตัวแบบที่ผ่านการตรวจสอบจากวิธีต่างๆ เพื่อคัดเลือกตัวแบบที่เหมาะสมไปใช้งาน

1.5.6 ชุดข้อมูลฝึกฝน (Training Data Set) หมายถึง ข้อมูลที่สุ่มหรือเลือกจากชุดข้อมูลที่ทำการศึกษาวิเคราะห์เพื่อใช้ในการพัฒนาตัวแบบ โดยคัดเลือกตัวแปรที่เหมาะสมตามเกณฑ์ที่กำหนดรวมทั้งการประมาณค่าพารามิเตอร์ของตัวแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

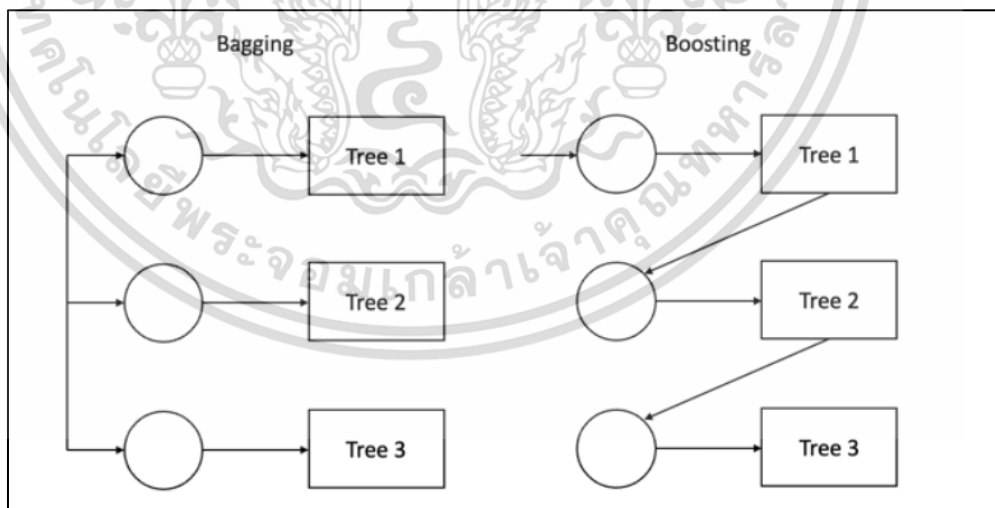
บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

สำหรับปัญหาพิเศษนี้ผู้วิจัยมีความสนใจที่จะศึกษาวิธีการทำนายข้อมูลการถดถอยเชิงเส้น เมื่อข้อมูลมีมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่องประกอบไปด้วย วิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนตเดสเซนท์ ซึ่งรายละเอียดและทฤษฎีและงานวิจัยที่เกี่ยวข้องของแต่ละวิธีมีดังต่อไปนี้

2.1 วิธีป่าสุ่ม (Random Forest)

วิธีป่าสุ่มเป็นวิธีการที่มีรากฐานจากวิธีต้นไม้ตัดสินใจ (Decision Tree) ซึ่งสามารถลดจุดอ่อนของวิธีต้นไม้ตัดสินใจคือวิธีดังกล่าวมีความแปรปรวนที่สูงมาก กล่าวคือเมื่อเปลี่ยนข้อมูลเพียงเล็กน้อย ต้นไม้ที่สร้างขึ้นใหม่อาจต่างจากเดิมโดยสิ้นเชิง ซึ่งวิธีป่าสุ่มสามารถแก้ไขปัญหาดังกล่าวได้ค่อนข้างดี โดยวิธีป่าสุ่มจะแบ่งออกเป็น 2 ประเภทได้แก่ Bootstrap Aggregation (Bagging) ซึ่งเป็นการเรียนรู้ที่แยกกันอย่างสิ้นเชิง หลักการของวิธีดังกล่าวคือจะทำการสุ่มต้นไม้จำนวนหนึ่งและทำการรวมต้นไม้เพื่อให้ได้ผลลัพธ์ที่ดีที่สุดและอีกประเภทคือ Extreme Gradient Boosting (Boosting) ซึ่งวิธีนี้จะเรียนรู้โดยใช้ต้นไม้ต้นก่อนเป็นพื้นฐาน ซึ่งเป็นการเรียนรู้แบบที่จะทำการเรียนรู้และปรับปรุงแบบจำลองไปที่ละขั้นตอนผ่านต้นไม้ต้นก่อนหน้าดังรูปที่ 2.1 (พรทิพย์ และคณะ,2565)

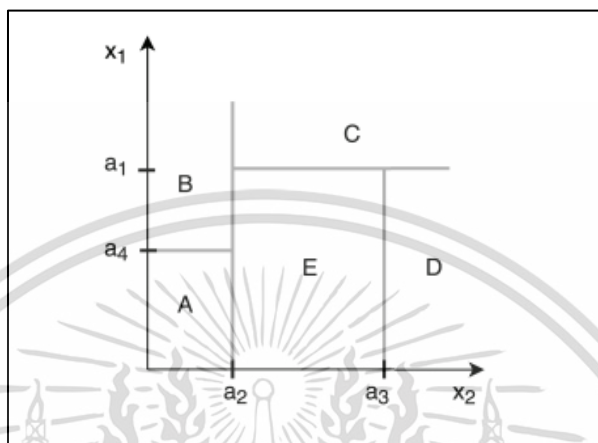


รูปที่ 2.1 หลักการของวิธี Bagging และ วิธี Boosting (พรทิพย์ และคณะ,2565)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

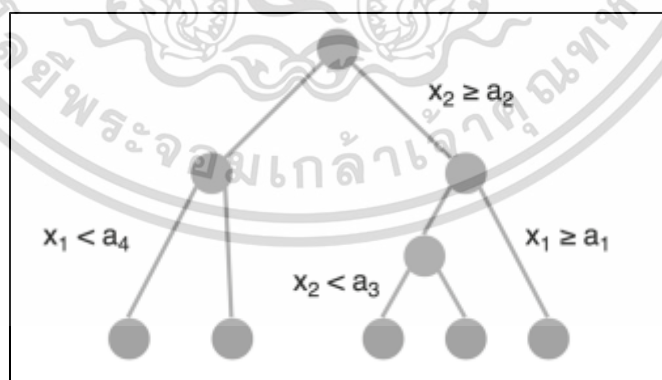
2.1.1 การถดถอยป่าสุ่ม (Random forest Regression)

เนื่องจากว่าวิธีป่าสุ่มมีพื้นฐานมาจากวิธีต้นไม้ตัดสินใจซึ่งโดยปกติในการทำต้นไม้ตัดสินใจค่าของตัวแปรต้องเป็นแบบไม่ต่อเนื่อง (Discrete) แต่ในกรณีที่ค่าของตัวแปรมีค่าต่อเนื่องนั้น จำเป็นจะต้องแบ่งค่าของข้อมูลออกเป็นช่วงๆ ดังรูปที่ 2.2



รูปที่ 2.2 แสดงการแบ่งข้อมูลออกเป็นช่วง

เมื่อทำการแบ่งข้อมูลเป็นช่วงได้แล้วก็จะทำการสร้างต้นไม้ตัดสินใจหลายต้นแบบสุ่มหรือที่เรียกว่าป่าสุ่มและทำการรวมต้นไม้ตัดสินใจเพื่อให้ได้ต้นไม้ที่ดีที่สุด ในที่นี้จะใช้ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองเป็นตัววัดประสิทธิภาพของต้นไม้ตัดสินใจ โดยจะแสดงเป็นต้นไม้ตัดสินใจดังรูปที่ 2.3 (Schonlau and Zou, 2020)



รูปที่ 2.3 ต้นไม้ตัดสินใจ

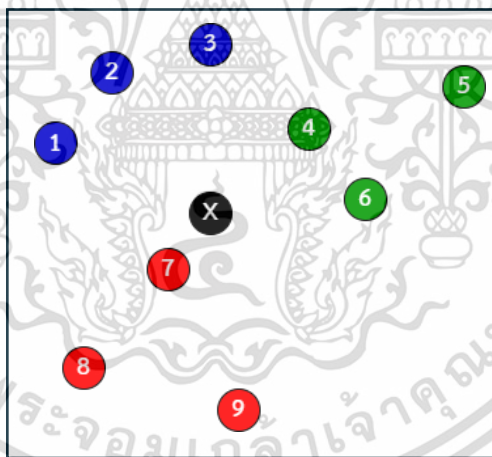
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ (K Nearest Neighbor : KNN)

วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ เป็นวิธีหนึ่งที่ไม่ซับซ้อน มีประสิทธิภาพค่อนข้างสูงและมักจะเป็นวิธีที่ได้รับความนิยมเป็นอันดับแรกในการสร้างแบบจำลองการเรียนรู้ของเครื่อง หลักการสำคัญของขั้นตอนวิธีนี้คือ การค้นหาตัวอย่างจำนวนหนึ่ง (จำนวน k ตัวอย่าง) จากปริภูมิตัวอย่างที่ปรากฏใกล้เคียงที่สุดกับสิ่งที่ต้องการจำแนกประเภทตัวอย่างเช่น การจัดประเภทภาพยนตร์ หากต้องการที่จะระบุว่าภาพยนตร์ประเภทหนึ่งเป็นหนังประเภทความรัก (Romance) หรือประเภทยู (Action) ซึ่งหนังประเภทยูบางครั้งก็มีบทรักเช่นเดียวกับหนังรักซึ่งบางครั้งก็มีฉากบู๊ ด้วยเหตุนี้จึงต้องมีตัววัด (Measure) ปริมาณบางอย่างที่สามารถเปรียบเทียบกันได้ ในปัญหานี้อาจใช้วิธีแก้ปัญหาย่างง่ายคือ ถ้าภาพยนตร์เรื่องนี้มีบทรักมากกว่าบู๊ให้นับเป็นหนังรักและหากมีฉากบู๊มากกว่าบทรักให้เป็นหนังยู ซึ่งวิธีคิดเรียบง่ายนี้เป็นพื้นฐานของวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ โดยมีขั้นตอนดังต่อไปนี้ (ปริญญา, 2562)

2.2.1 ขั้นตอนวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ

กำหนดให้ชุดข้อมูลหนึ่งประกอบด้วยตัวอย่างที่ทราบประเภทจำนวน 9 ตัวอย่าง (1-9) และมีตัวอย่างที่ไม่ทราบประเภท (X) จำนวน 1 ตัวอย่าง มีการกระจายตัวดังรูปที่ 2.4



รูปที่ 2.4 ปริภูมิตัวอย่าง (ปริญญา, 2562)

โดยตัวอย่างที่ 1-3 เป็นประเภท B, ตัวอย่างที่ 4-6 เป็นประเภท G และตัวอย่างที่ 7-9 เป็นประเภท R เมื่อต้องการประเภทของ X ด้วยวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับทำได้ดังต่อไปนี้

ขั้นตอนที่ 1 กำหนดค่า k ซึ่งหมายถึงจำนวนตัวอย่างที่ใกล้เคียงที่สุด k ตัว โดยจะมีค่าตั้งแต่ 1 ถึงขนาดตัวอย่าง (ในที่นี้มีขนาดตัวอย่างเท่ากับ 9)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 กำหนดวิธีการวัดความใกล้เคียงหรือความเหมือน (Similarity) โดยทั่วไปมักวัดจากค่าผกผันของระยะทาง (Distance) ระหว่างตัวอย่างดังสมการ 2.1

$$\text{dist}(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{i,k} - x_{j,k})^2} \quad (2.1)$$

โดยที่ $\text{dist}(x_i, x_j)$ คือ ระยะห่างระหว่างตัวอย่าง x_i กับตัวอย่าง x_j

n คือ จำนวนคุณลักษณะทั้งหมดของตัวอย่าง

$x_{i,k}$ คือ คุณลักษณะที่ k ของตัวอย่าง x_i

$x_{j,k}$ คือ คุณลักษณะที่ k ของตัวอย่าง x_j

ขั้นตอนที่ 3 นับจำนวนตัวอย่างที่ใกล้เคียงที่สุด k ตัว โดยหากจำนวนตัวอย่างประเภทใดมีมากที่สุด k ตัว ให้กำหนดประเภทของตัวอย่างที่ต้องการทราบเป็นประเภทเดียวกับประเภทตัวอย่างที่มีมากที่สุด โดยที่คำตอบจะแตกต่างกันตามค่า k ที่เลือกดังตารางที่ 2.1

ตารางที่ 2.1 ผลการจำแนกประเภทด้วยค่า k ต่างๆ

K	ผลการจำแนกประเภท
1	R
2	R หรือ G
3	G
4	G
5	G หรือ B
6	B
7	B
8	B หรือ R
9	R หรือ G หรือ B

2.2.2 การถดถอยวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ

ที่กล่าวไปข้างต้นจะพบว่าวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ เป็นวิธีที่เหมาะสมสำหรับการจำแนกหรือแยกข้อมูลเป็นกลุ่ม แต่สำหรับข้อมูลที่เป็นข้อมูลการถดถอยนั้นจำเป็นต้องมีการปรับปรุงวิธีการเล็กน้อยเพื่อให้เหมาะสมกับข้อมูลประเภทดังกล่าว โดยวิธีการนั้นจะคล้ายคลึงกันกับวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับจะแตกต่างกันตรงที่เมื่อกำหนดค่า k เรียบร้อยแล้วจะทำการถดถอยค่าของตัวแปรตามโดยใช้ค่าเฉลี่ยของเพื่อนบ้านที่ใกล้ที่สุด k อันดับนั้นเอง

เอกสารนี้เป็นลิขสิทธิ์ของสถาบันวิจัยระบบสาธารณสุข หากมีข้อสงสัยหรือต้องการข้อมูลเพิ่มเติม กรุณาติดต่อฝ่ายประชาสัมพันธ์ โทร. 02-5623.0000 หรือ 02-5623.0001

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การหาค่าเฉลี่ยข้อมูลที่ใกล้ที่สุดจำนวน k ตัวเพื่อพิจารณาว่าข้อมูลดังกล่าวควรมีค่าเป็นเท่าใด ดังสมการที่ 2.2 (Imandoust and Bolandraftar, 2013)

$$\hat{y} = \frac{1}{k} \sum_{i=1}^k y_i \quad (2.2)$$

โดยที่ i คือ ตัวอย่างที่ i
 \hat{y} คือ การทำนาย (ผลลัพธ์)

2.3 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine : SVM)

วิธีซัพพอร์ตเวกเตอร์แมชชีน เป็นเครื่องมือที่ใช้ในการจำแนกประเภทที่นิยมมากในงานที่หลากหลายในปัจจุบัน เช่น การจัดหมวดหมู่เอกสาร การรู้จำใบหน้าและอีกมากมาย ซึ่งมักจะทำให้ประสิทธิภาพที่ดีทั้งนี้เนื่องจากวิธีซัพพอร์ตเวกเตอร์แมชชีนมีการจัดจำข้อมูลที่อยู่บริเวณขอบของแต่ละคลาส หรือที่เรียกว่า “เวกเตอร์ซัพพอร์ต” อีกทั้งวิธีดังกล่าวสามารถใช้กับข้อมูลที่เป็นเชิงเส้นและไม่เชิงเส้น โดยวิธีซัพพอร์ตเวกเตอร์แมชชีนมีรายละเอียดดังต่อไปนี้ (ปริญญา, 2562)

2.3.1 ระนาบเกิน (Hyperplane)

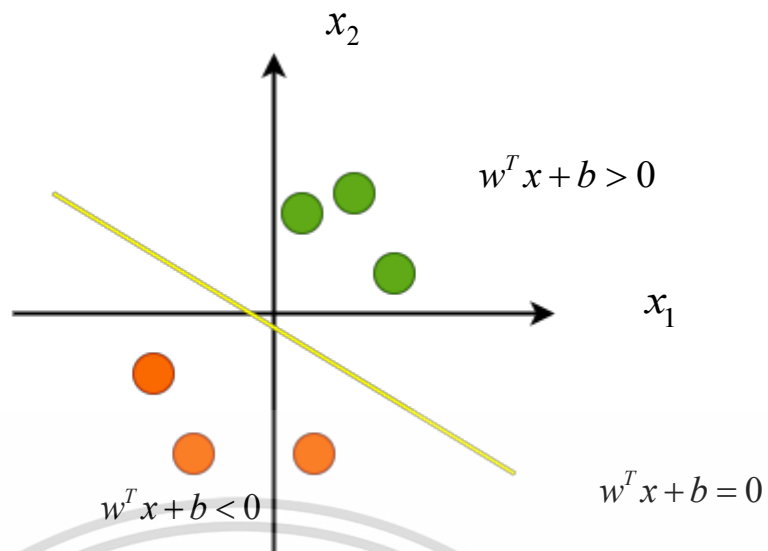
เพื่อความเข้าใจขอยกตัวอย่างระนาบเกิน ในปริภูมิมิติต่างๆให้พิจารณาก่อนดังต่อไปนี้ ในปริภูมิ 1 มิติ สามารถใช้ จุด เพื่อแบ่งข้อมูลออกเป็น 2 คลาส ดังรูปที่ 2.5



รูปที่ 2.5 ตัวอย่างระนาบเกินในปริภูมิ 1 มิติ

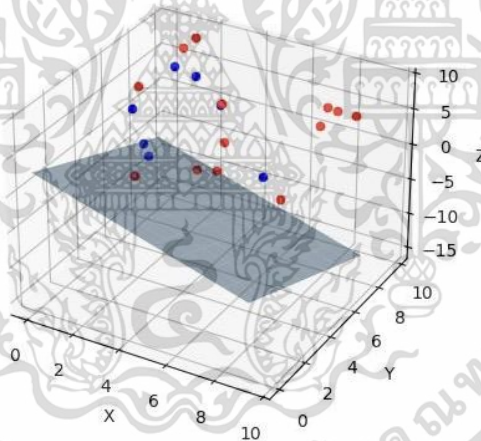
ในปริภูมิ 2 มิติ สามารถใช้ เส้น เพื่อแบ่งข้อมูลออกเป็น 2 คลาส ดังรูปที่ 2.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 ตัวอย่างระนาบเกินในปริภูมิ 2 มิติ

ในปริภูมิ 3 มิติ สามารถใช้ ระนาบ เพื่อแบ่งข้อมูลออกเป็น 2 คลาส ดังรูปที่ 2.7



รูปที่ 2.7 ตัวอย่างระนาบเกินในปริภูมิ 3 มิติ (ปริญญา, 2562)

จากตัวอย่างจะพบว่าระนาบเกินไม่ว่าจะอยู่ในปริภูมิที่มีมิติก็ตาม สามารถเขียนสมการของระนาบเกินในรูปแบบเวกเตอร์ได้เหมือนกัน ดังสมการที่ 2.2

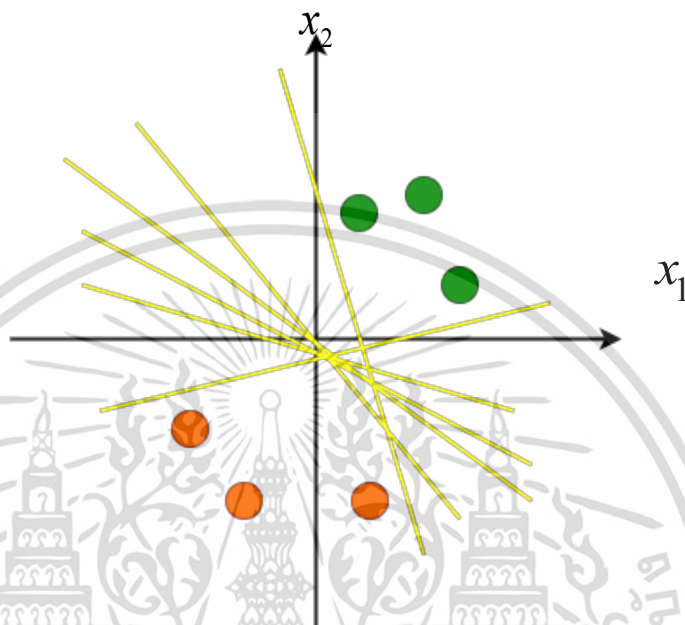
$$w^T x + b = 0 \quad (2.2)$$

เอกสารนี้เป็นเอกสารที่สงวน โดย b คือ ค่าไบแอสหรือระยะเลื่อนขนานจากจุดกำเนิด ไปใช้ประโยชน์ด้านการคำนวณมากกว่ากรณีใดๆ ทั้งสิ้น อีกทั้ง w คือ เวกเตอร์ปกติ (Normal vector) ซึ่งเวกเตอร์นี้จะตั้งฉากกับระนาบเกินและ

มีมิติเท่ากับมิติของปริภูมินั้นๆ

2.3.2 ระนาบเกินแบบบัญญัติ (Canonical Hyperplane)

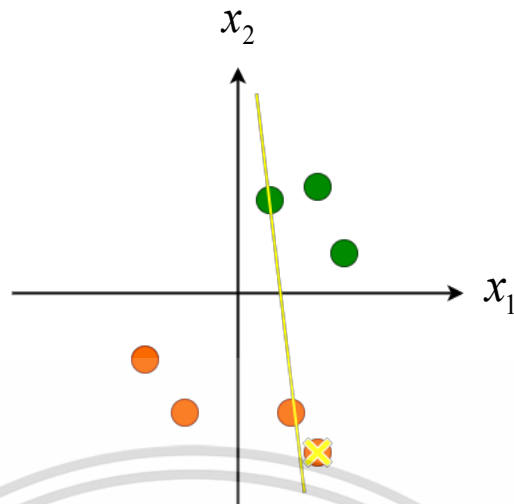
คำถามที่มักเกิดขึ้นในการใช้ระนาบเกินในการแบ่งแยกข้อมูลออกเป็น 2 คลาสก็คือ ระนาบเกินแบบไหนจะดีที่สุด ดูได้จากตัวอย่างดังรูปที่ 2.8



รูปที่ 2.8 ระนาบเกินที่สามารถแบ่งข้อมูล 2 คลาสได้

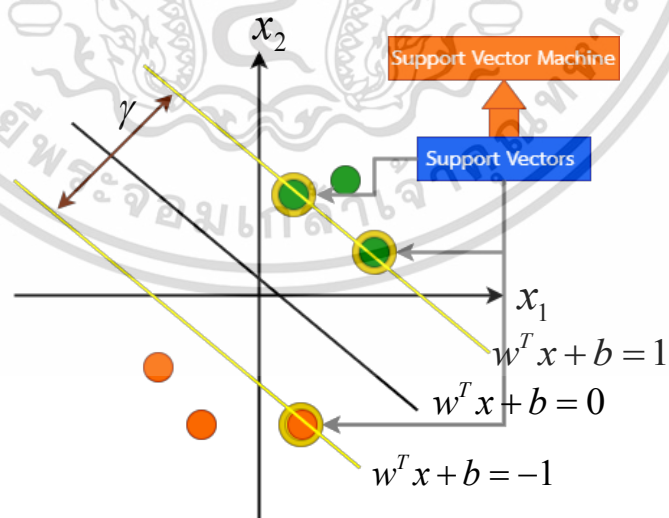
จะพบว่าทุกระนาบเกินในรูปที่ 2.8 สามารถแบ่งข้อมูลชุดนี้ได้ถูกต้องทั้งหมด แต่หากมีข้อมูลเพิ่มเข้ามาใหม่ในกรณีดังรูป 2.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 ระนาบเกินที่เกิดปัญหาระนาบเหมาะสมเกินไป

ซึ่งเรียกระนาบเกินแบบนี้ว่าเกิดปัญหาการเหมาะสมเกินไป (Overfitting) ดังนั้น ระนาบเกินแบบนี้จึงไม่ควรเลือกใช้ กล่าวคือ ระนาบเกินที่ต้องการควรทำให้ใช้ทั่วไปได้ (Generalize) มากกว่านี้ ใน SVM ระยะขอบ (Margin) ที่มากที่สุดระหว่างคลาสหนึ่งกับอีกคลาสหนึ่ง ถูกใช้เป็นบรรทัดฐานในการตัดสินใจเลือกระนาบที่เหมาะสมที่สุด (Optimal Hyperplane) (Vapnik, 1995) และเวกเตอร์ข้อมูลที่อยู่บริเวณของแต่ละคลาสจะเรียกว่า เวกเตอร์ซัพพอร์ต (Support Vector) ดังรูปที่ 2.10



รูปที่ 2.10 เวกเตอร์ซัพพอร์ต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย γ คือระยะขอบ (Margin) และข้อมูลบริเวณขอบแต่ละคลาสที่อยู่ภายในเครื่องหมายวงกลมทั้งหมดคือ เวกเตอร์ซัพพอร์ต และเป็นที่มาของชื่อ “ซัพพอร์ตเวกเตอร์แมชชีน”

2.3.3 การถดถอยวิธีซัพพอร์ตเวกเตอร์แมชชีน

ที่กล่าวไปข้างต้นจะพบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนเป็นวิธีที่เหมาะสมสำหรับการจำแนกหรือแยกข้อมูลเป็นกลุ่มโดยใช้ระนาบเกินในการแบ่งข้อมูลออกจากกัน แต่สำหรับการถดถอยวิธีซัพพอร์ตเวกเตอร์แมชชีนจะใช้ระนาบเกินในการทำนายค่าการถดถอย โดยให้ระนาบเกินทำหน้าที่เหมือนกับขอบเขตของเส้นการถดถอยและให้ข้อมูลทั้งหมดอยู่ภายใต้ขอบเขตของระนาบเกิน ดังสมการและรูปการเปรียบเทียบวิธีซัพพอร์ตเวกเตอร์แมชชีน และวิธีการถดถอยซัพพอร์ตเวกเตอร์ดังต่อไปนี้ (นัทธมน, ชนาพันธุ์และวิกานดา, 2563)

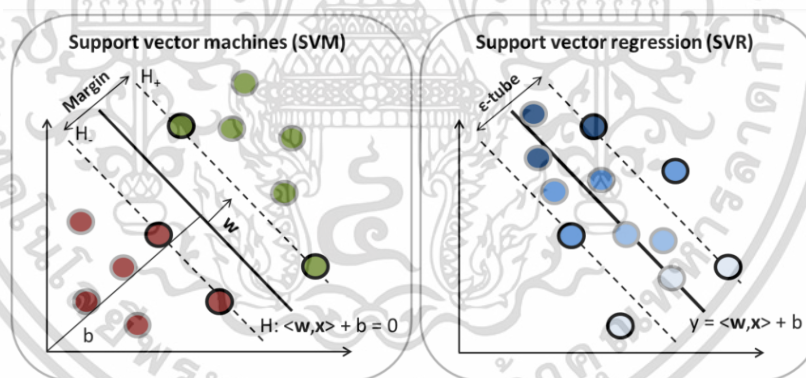
$$\hat{y} = wx + b \quad (2.3)$$

โดยที่ w คือ เวกเตอร์ถ่วงน้ำหนัก (weight)

b คือ ค่าความคลาดเคลื่อน (Bias)

x คือ เวกเตอร์ข้อมูลนำเข้า (Input data) แสดงในรูปของเวกเตอร์คอลัมน์

$(x_1, \dots, x_n)^T$ ขนาดเท่ากับ $n; x_i \in R; i = 1, \dots, n$



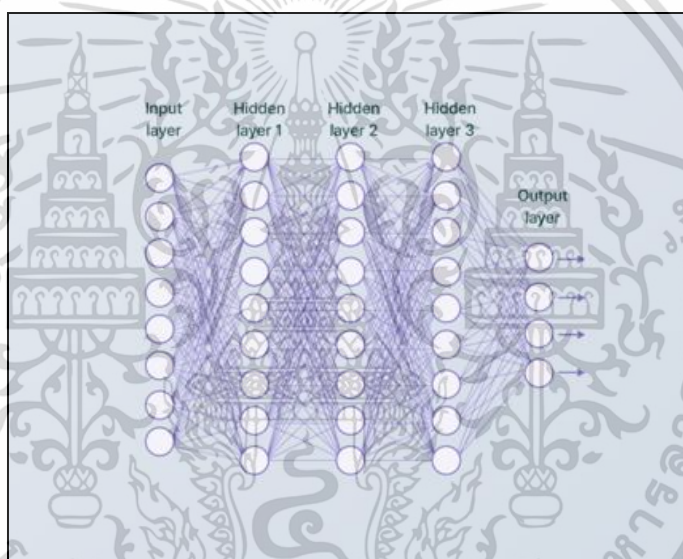
รูปที่ 2.11 เปรียบเทียบแบบจำลองวิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการถดถอยซัพพอร์ตเวกเตอร์ (นัทธมน, ชนาพันธุ์และวิกานดา, 2563)

จากรูปที่ 2.11 แสดงให้เห็นการเปรียบเทียบแบบจำลองวิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการถดถอยซัพพอร์ตเวกเตอร์ ซึ่งแตกต่างกันที่การสร้างเส้นระนาบเกินโดยที่วิธีซัพพอร์ตเวกเตอร์แมชชีนจะสร้างเส้นแบ่งระหว่างข้อมูลและวิธีการถดถอยซัพพอร์ตเวกเตอร์จะสร้างเส้นการถดถอยจากข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 วิธีการเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึกคือ โครงข่ายประสาทเทียมที่มีการเพิ่มขึ้นประมวลผลที่ซ่อนอยู่ (Hidden Layers) เพื่อเพิ่มประสิทธิภาพในการแก้ปัญหาที่ซับซ้อน การเรียนรู้เชิงลึกเริ่มเป็นที่รู้จักกันอย่างแพร่หลายในช่วงที่ผ่านมาและในปัจจุบันจากความสามารถที่สูงในการประยุกต์และใช้งานที่หลากหลาย เช่น ความสามารถในการรับรู้และแยกแยะเสียงในธรรมชาติ (Natural Voice and Speech Recognition) การประยุกต์ใช้งานร่วมกับระบบการมองเห็นของคอมพิวเตอร์ (Computer Vision) และอื่นๆอีกมากมาย ทำให้การเรียนรู้เชิงลึกถูกศึกษาค้นคว้าและพัฒนาอย่างต่อเนื่อง การศึกษาเหล่านี้ทำให้เกิดการเรียนรู้เชิงลึกที่เป็นที่รู้จักมากมายหลายระบบ เช่น GoogleNet, Resnet, Alexnet (ราชทูธา, 2562) ดังรูปที่ 2.12 ที่แสดงลักษณะการเรียนรู้เชิงลึก

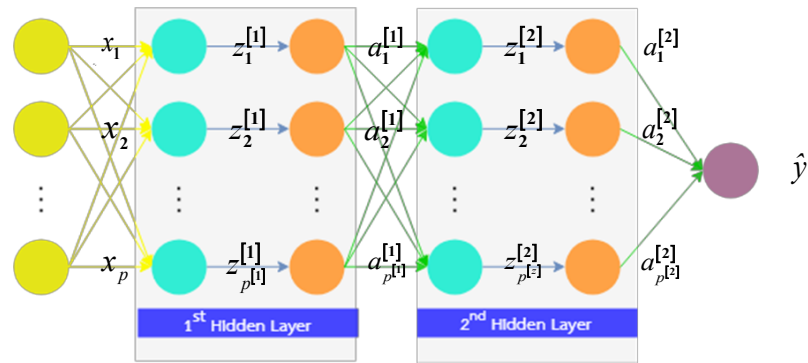


รูปที่ 2.12 โครงข่ายการเรียนรู้เชิงลึก (ราชทูธา, 2562)

2.4.1 การถอดการการเรียนรู้เชิงลึก

การเรียนรู้เชิงลึกนั้นสามารถใช้กับข้อมูลที่เป็นการถอดได้เช่นเดียวกับโครงข่ายประสาทเทียมโดยที่วิธีดังกล่าวจะรวมโหนด (node) ระหว่างชั้นประมวลผลที่ซ่อนอยู่ในทุกๆชั้นจนครบแล้วจึงสร้างสมการสำหรับการทำนายค่าการถอดดังรูป 2.13 (ชฤณ และคณะ, 2566)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 วิธีการทำงานของการเรียนรู้เชิงลึก (ขุณ และคณะ, 2566)

2.4.2 สมการการถอดการการเรียนรู้เชิงลึก

เมื่อรวมทุกโหนดและชั้นประมวลผลที่ซ่อนอยู่จะสามารถเขียนเป็นสมการได้ดังสมการที่ 2.4

$$\hat{y} = b^{[out]} + \sum_{j=1}^n w_j^{[out]} a_j^h \quad (2.4)$$

โดยที่ $b^{[out]}$ คือ ค่าจุดตัดแกน Y

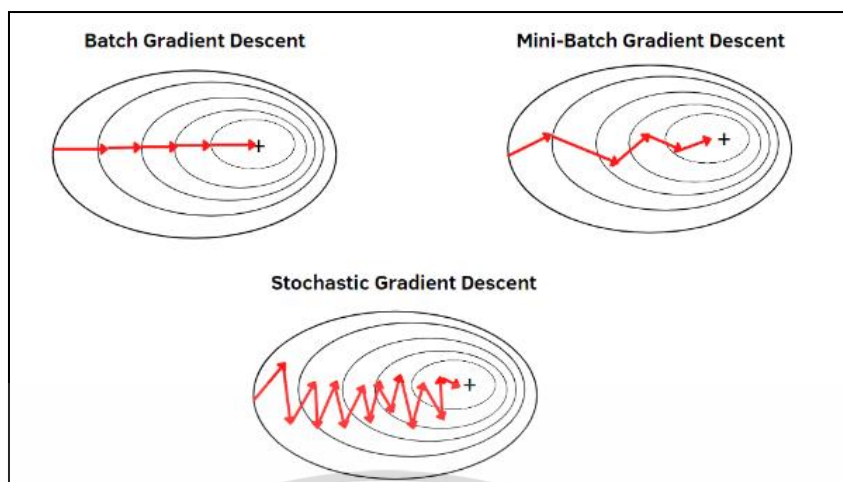
$w_j^{[out]}$ คือ ค่าถ่วงน้ำหนักหรือค่าความชัน (slope)

a_j^h คือ ตัวแปรอิสระที่ถูกเปลี่ยนให้อยู่ในรูปไม่เชิงเส้น (nonlinear)

2.5 วิธีสโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent)

วิธีการปรับให้เหมาะสมลำดับแรก (First-Order Optimization Methods) คือ วิธีเกรเดียนเดสเซนท์ โดยมีตัวแบบของวิธีเกรเดียนเดสเซนท์ (Gradient Descent) 3 รูปแบบ ดังนี้ Batch Gradient Descent (BGD) Mini-batch gradient descent และ Stochastic Gradient Descent (SGD) โดยแต่ละวิธีมีลักษณะดังรูปที่ 2.14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

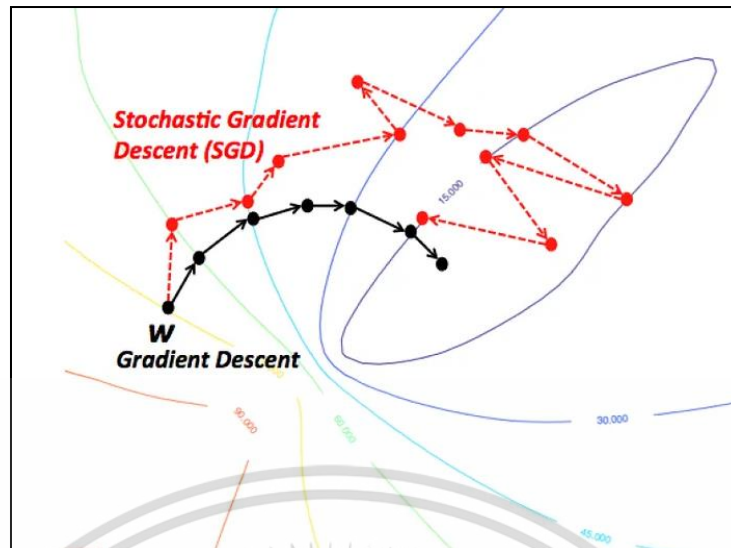


รูปที่ 2.14 วิธีต่างๆของวิธีเกรเดียนเดสเซนส์ทั้ง 3 รูปแบบ (Haji and Abdulazeez. 2021)

โดยที่แต่ละวิธีจะแตกต่างกันไปตามปริมาณของข้อมูล ความถูกต้องของการปรับปรุงพารามิเตอร์และเวลาในการดำเนินการปรับปรุง (Haji and Abdulazeez., 2021) ซึ่งวิธีสโตแคสติกเกรเดียนเดสเซนส์ (Stochastic Gradient Descent) นั้นถูกนำไปใช้กันอย่างแพร่หลายในการทำการศึกษาฝึกฝนในวิธีการเรียนรู้เชิงลึก (Cai, T et al., 2019) และอีกหนึ่งเหตุผลเมื่อเทียบกับวิธีเกรเดียนเดสเซนส์แล้วนั้นจะเห็นว่าปัญหาคือวิธีเกรเดียนเดสเซนส์ต้องการหน่วยความจำค่อนข้างเยอะทำให้ไม่ค่อยได้รับความนิยม จึงเกิดวิธีสโตแคสติกเกรเดียนเดสเซนส์โดยวิธีนี้เป็นวิธีหนึ่งของวิธีเกรเดียนเดสเซนส์ที่ใช้ตัวอย่างข้อมูลแบบสุ่มแทนที่จะเป็นชุดข้อมูลทั้งหมด (Abu-Shaira and Speegle, 2023)

โดยที่วิธีสโตแคสติกเกรเดียนเดสเซนส์จะใช้อัลกอริทึมที่มีการปรับปรุงค่าพารามิเตอร์ในทุกๆชุดข้อมูลฝึกฝนซึ่งเป็นอัลกอริทึมที่สามารถประมวลผลได้ค่อนข้างไว และมีการอัปเดตเพียง 1 ครั้งต่อการฝึกอบรม 1 รอบ (พิศิษฐ์และวรภัทร, 2565) โดยในทุกครั้งที่มีการปรับปรุง ค่าพารามิเตอร์ที่ปรับปรุงจะมีค่าความแปรปรวนสูงและส่งผลกับค่าฟังก์ชันการสูญเสีย (Loss Function) และแปรผันไปตามความหนาแน่นที่แตกต่าง (Different Intensities) โดยขั้นตอนเหล่านี้ช่วยให้ค้นพบค่าที่น้อยที่สุดได้ ดังรูป 2.15

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 เปรียบเทียบวิธีเกรเดียนเตสเซนท์ (Gradient Descent) และวิธีสโตเคสติกเกรเดียนเตสเซนท์ (Stochastic Gradient Descent) (Haji and Abdulazeez. 2021)

ปัญหาการถดถอยเชิงเส้น (Linear Regression) มีจุดประสงค์คือ การทำให้ข้อผิดพลาดเกิดขึ้นลดน้อยลงที่สุดโดยใช้ค่าเฉลี่ยของผลต่างกำลังสอง (MSE) หรือที่เรียกว่าค่าฟังก์ชันการสูญเสีย (loss function) ในการปรับปรุงค่าพารามิเตอร์ต่างๆในสมการดังต่อไปนี้

1. สมการการถดถอยเชิงเส้น

$$\hat{y} = wx + b \quad (2.5)$$

เมื่อ w คือ เวกเตอร์ถ่วงน้ำหนัก (Weight)

b คือ ค่าความคลาดเคลื่อน (Bias)

x คือ เวกเตอร์ข้อมูลนำเข้า แสดงในรูปของเวกเตอร์คอลัมน์ $(x_1, \dots, x_n)^T$ ขนาดเท่ากับ n ; $x_i \in R; i = 1, \dots, n$

2. สมการค่าเฉลี่ยของผลต่างค่าคาดเคลื่อนกำลังสอง (Mean squared error : MSE) ค่าฟังก์ชันการสูญเสีย

$$J(w, b) = \frac{1}{2n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (2.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ที่เมื่อ อีกทั้ง คือ จำนวนตัวอย่างที่ใช้ฝึกฝน อย่างอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

x_i คือ คุณสมบัติของตัวอย่างที่ i

y_i คือ ค่าเป้าหมายที่แท้จริงของตัวอย่างที่ i

กฎเกณฑ์ที่พัฒนาเพื่อวิธีสโตแคสติกเกรเดียนเดสเซนท์ (Stochastic Gradient Descent) มีสมการดังนี้

$$w_{t+1} = w_t - \eta \frac{\partial J}{\partial w} \quad (2.7)$$

$$b_{t+1} = b_t - \eta \frac{\partial J}{\partial b} \quad (2.8)$$

เมื่อ w_t และ b_t คือ พารามิเตอร์ที่ปรับปรุงการเรียนรู้จำนวน t ครั้ง
 η คือ อัตราการเรียนรู้ (Learning rate)

2.6 ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสอง (Mean Square Error : MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

เมื่อ n คือ ขนาดตัวอย่าง

Y_i คือ ค่าที่สังเกตได้จริงของตัวแปรตาม

\hat{Y}_i คือ ค่าที่ทำนายของตัวแปรตามที่ได้รับจากการเรียนรู้ของเครื่อง

2.7 ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (Mean Absolute Percentage Error : MAPE)

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \right) \times 100$$

เมื่อ n คือ ขนาดตัวอย่าง

Y_i คือ ค่าที่สังเกตได้จริงของตัวแปรตาม

\hat{Y}_i คือ ค่าที่ทำนายของตัวแปรตามที่ได้รับจากการเรียนรู้ของเครื่อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8 งานวิจัยที่เกี่ยวข้อง

โสภีและคณะ (2565) ได้ศึกษาตัวแบบการเรียนรู้ของเครื่องที่เหมาะสมสำหรับการทำนายราคาพริกชี้หนูโดยใช้ข้อมูลย้อนหลังของสภาพภูมิอากาศจากสถานีอุตุนิยมวิทยาจังหวัดนครศรีธรรมราชและราคาพริกชี้หนูจากงานพาณิชย์จังหวัดนครศรีธรรมราช พ.ศ. 2553 -2563 โดยใช้ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีการถดถอยเชิงเส้น วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ และวิธีป่าสุ่ม ผลการศึกษาพบว่าวิธีป่าสุ่มเป็นตัวแบบที่มีประสิทธิภาพดีที่สุดเนื่องจากให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์น้อยที่สุด

จุฬาวลัยและศิวา (2563) ได้ศึกษาตัวแบบการเรียนรู้ของเครื่องที่เหมาะสมสำหรับการประมาณค่าความเข้มข้นของตะกอนแขวนลอยในแม่น้ำลำปาวโดยใช้ข้อมูลจากดาวเทียม ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีการถดถอยเชิงเส้น วิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการเรียนรู้เชิงลึก ผลการศึกษาพบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวแบบที่มีประสิทธิภาพดีที่สุดเนื่องจากให้ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองน้อยที่สุด

วิรัชและฐิตาภรณ์ (2561) ได้ศึกษาการสร้างแบบจำลองการติดตามการเดินทางเข้าสู่พื้นที่อุทยานโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์โดยใช้ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิเคราะห์การถดถอย วิเคราะห์ข่ายประสาทเทียมและวิธีการเรียนรู้เชิงลึก ผลการศึกษาพบว่าวิธีการเรียนรู้เชิงลึกเป็นตัวแบบที่มีประสิทธิภาพดีที่สุดเนื่องจากให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์น้อยที่สุด

Sun, Y. (2021) ทำการตรวจสอบการทำนายราคาบ้านด้วยวิธีเกรเดียนต์เดสเซนท์แบบต่างๆ เช่น วิธีเกรเดียนต์เดสเซนท์แบบทั่วไป (General Gradient Descent) วิธีสโตแคสติกเกรเดียนต์เดสเซนท์ (Stochastic Gradient Descent) และวิธีมินิแบชเกรเดียนต์เดสเซนท์ (Mini-Batch Gradient Descent) จากการเปรียบเทียบค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองพบว่าวิธีสโตแคสติกเกรเดียนต์เดสเซนท์เป็นตัวแบบที่มีประสิทธิภาพดีที่สุด

Biau et al. (2012) ได้เสนอวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ โดยการจำลองชุดข้อมูลขึ้นและใช้วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับในการประมาณค่าสัมประสิทธิ์การถดถอย ซึ่งวิธีการดังกล่าวสามารถแก้ไขปัญหาที่ต้องทำการเปลี่ยนแปลงข้อมูลก่อนทำการวิเคราะห์การถดถอยเชิงเส้น ส่งผลให้ลดความซับซ้อนในการสร้างตัวแบบการถดถอยเชิงเส้น

บทที่ 3

วิธีการดำเนินงานวิจัย

การวิจัยครั้งนี้จะทำการศึกษาค่าการเปรียบเทียบประสิทธิภาพการทำนายข้อมูลการถดถอยเชิงเส้นสำหรับข้อมูลมีมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่อง โดยนำข้อมูลที่ได้จากการจำลองมาวิเคราะห์ด้วยวิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์ ในการศึกษาแต่ละวิธีจะใช้โปรแกรมอาร์เข้ามาช่วยในการวิเคราะห์ข้อมูลและดำเนินการตามวัตถุประสงค์ของงานวิจัยต่อไป โดยบทนี้จะกล่าวถึงการวางแผนการวิจัยและวิธีการดำเนินการวิจัยดังนี้

3.1 การวางแผนการวิจัย

ในการทำปัญหาพิเศษครั้งนี้ทำการกำหนดสถานการณ์ต่างๆในการเปรียบเทียบดังนี้

3.1.1 กำหนดขนาดตัวอย่าง (n) และจำนวนตัวแปรอิสระ (p) ให้อยู่ในรูปแบบข้อมูลมีมิติขั้นสูง โดยแบ่งเป็น 2 กลุ่มประกอบไปด้วยกลุ่มที่มีจำนวนตัวแปรอิสระน้อยและจำนวนตัวแปรอิสระมาก ดังนี้

กลุ่มจำนวนตัวแปรอิสระน้อย

$$p = 50, n = 20, 30 \text{ และ } 40$$

$$p = 75, n = 20, 30 \text{ และ } 40$$

$$p = 100, n = 20, 30 \text{ และ } 40$$

กลุ่มจำนวนตัวแปรอิสระมาก

$$p = 200, n = 50, 100 \text{ และ } 150$$

$$p = 250, n = 50, 100 \text{ และ } 150$$

$$p = 300, n = 50, 100 \text{ และ } 150$$

3.1.2 กำหนดค่าคลาดเคลื่อน (Error, ε) ของตัวแบบการถดถอยแบบพหุคูณจากข้อมูลที่มีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ย $E(\varepsilon) = \mu$ และ $Var(\varepsilon) = \sigma^2$ งานวิจัยนี้กำหนดให้ค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1, 5 และ 25

3.1.3 กำหนดค่าคงที่ของสมการถดถอย β_0 เป็น 1000

3.1.4 กำหนดค่าสัมประสิทธิ์การถดถอย β_j ของตัวแบบถดถอยเชิงเส้นพหุเป็น 1 และ -1 โดยที่ $j = 1, 2, \dots, p$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 วิธีดำเนินงานวิจัย

3.2.1 สร้างค่าคลาดเคลื่อนสำหรับสมการการถดถอย

3.2.2 สร้างตัวแปรอิสระจากการแจกแจงปกติจำนวน p ขนาดเท่ากับ n ซึ่งมีค่าเฉลี่ย $E(\varepsilon) = \mu$ และ $Var(\varepsilon) = \sigma^2$ งานวิจัยนี้กำหนดให้ค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

3.2.3 นำข้อมูลจาก 3.1.1 - 3.2.2 มาสร้างตัวแปรตาม

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i \text{ โดยที่ } i = 1, 2, \dots, n \text{ และ } j = 1, 2, \dots, p$$

เมื่อ Y_i คือ ค่าของตัวแปรตาม

x_{ij} คือ ค่าของตัวแปรอิสระ

β_0 คือ ค่าคงที่ (Constant) ของสมการถดถอย

β_j คือ ค่าสัมประสิทธิ์การถดถอย (Regression Coefficient) ของตัวแปรอิสระ

ε_i คือ ค่าความคลาดเคลื่อน (Error or Residual)

3.2.4 แบ่งข้อมูลเป็นชุดฝึกฝน (Training Data Set) และชุดทดสอบ (Test Data set) เป็นจำนวน 70 เปอร์เซ็นต์และ 30 เปอร์เซ็นต์ของข้อมูลทั้งหมดตามลำดับ

3.2.5 นำข้อมูลชุดฝึกฝนเข้าไปฝึกฝนด้วยวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี

3.2.6 นำผลการทำนายจากวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีไปทดสอบกับข้อมูลชุดทดสอบ

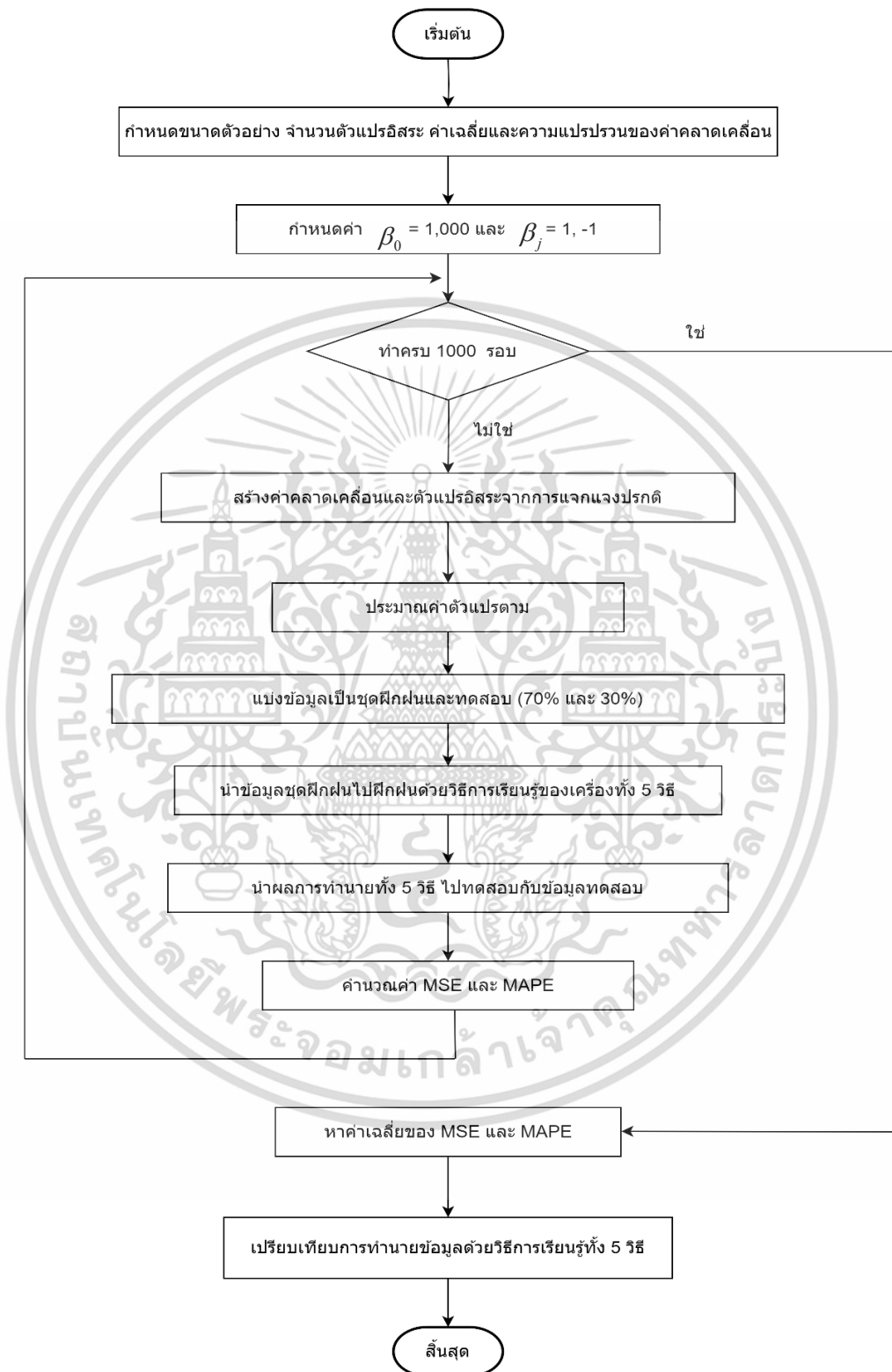
3.2.7 คำนวณค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองและค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์

3.2.8 ทำซ้ำขั้นตอนที่ 3.2.1 – 3.2.7 จำนวน 1000 รอบ

3.2.9 คำนวณค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์

3.2.10 เปรียบเทียบการทำนายข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยวิธีทั้ง 5 วิธี

3.3 ลำดับขั้นตอนการดำเนินงานวิจัย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในห้องปฏิบัติการวิจัยไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการวิจัย

เนื่องจากการวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง เพื่อศึกษาวิธีการทำนายค่าการถดถอยของข้อมูลที่มีมิติขั้นสูงด้วยของเรียนรู้ของเครื่อง ประกอบไปด้วยวิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนเดสเซนส์ เพื่อเปรียบเทียบว่าวิธีใดเป็นวิธีที่ดีที่สุดโดยใช้ค่าเฉลี่ยของผลต่างค่าคลาดเคลื่อนกำลังสองและค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์เป็นเกณฑ์ในการวัดและจะนำเสนอในรูปแบบตารางและรูปภาพ โดยมีสัญลักษณ์แสดงความหมายต่างๆดังนี้

- RF แทน วิธีป่าสุ่ม
- KNN แทน วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ
- SVM แทน วิธีซัพพอร์ตเวกเตอร์แมชชีน
- DL แทน วิธีการเรียนรู้เชิงลึก
- SGD แทน วิธีสโตแคสติกเกรเดียนเดสเซนส์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1 กรณีเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

4.1.1 กลุ่มจำนวนตัวแปรอิสระน้อย

กำหนดค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อน ขนาดตัวอย่าง เป็นไปตามขอบเขตของงานวิจัย แสดงผลดังตารางที่ 4.1-4.4 และรูปที่ 4.1-4.3

ตารางที่ 4.1 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0$, $\sigma^2 = 1$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี	n	p = 50			p = 75			p = 100		
		20	30	40	20	30	40	20	30	40
RF		0.56	0.55	0.54	0.7	0.69	0.67	0.8	0.8	0.78
KNN		0.56	0.54	0.53	0.7	0.68	0.66	0.8	0.79	0.78
SVM		0.57	0.55	0.53	0.7	0.68	0.65	0.8	0.78	0.77
DL		0.61	0.58	0.56	0.75	0.72	0.7	0.88	0.84	0.82
SGD		74.58	63.95	54.35	81.37	80.33	78.2	86.16	83.2	82.4

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

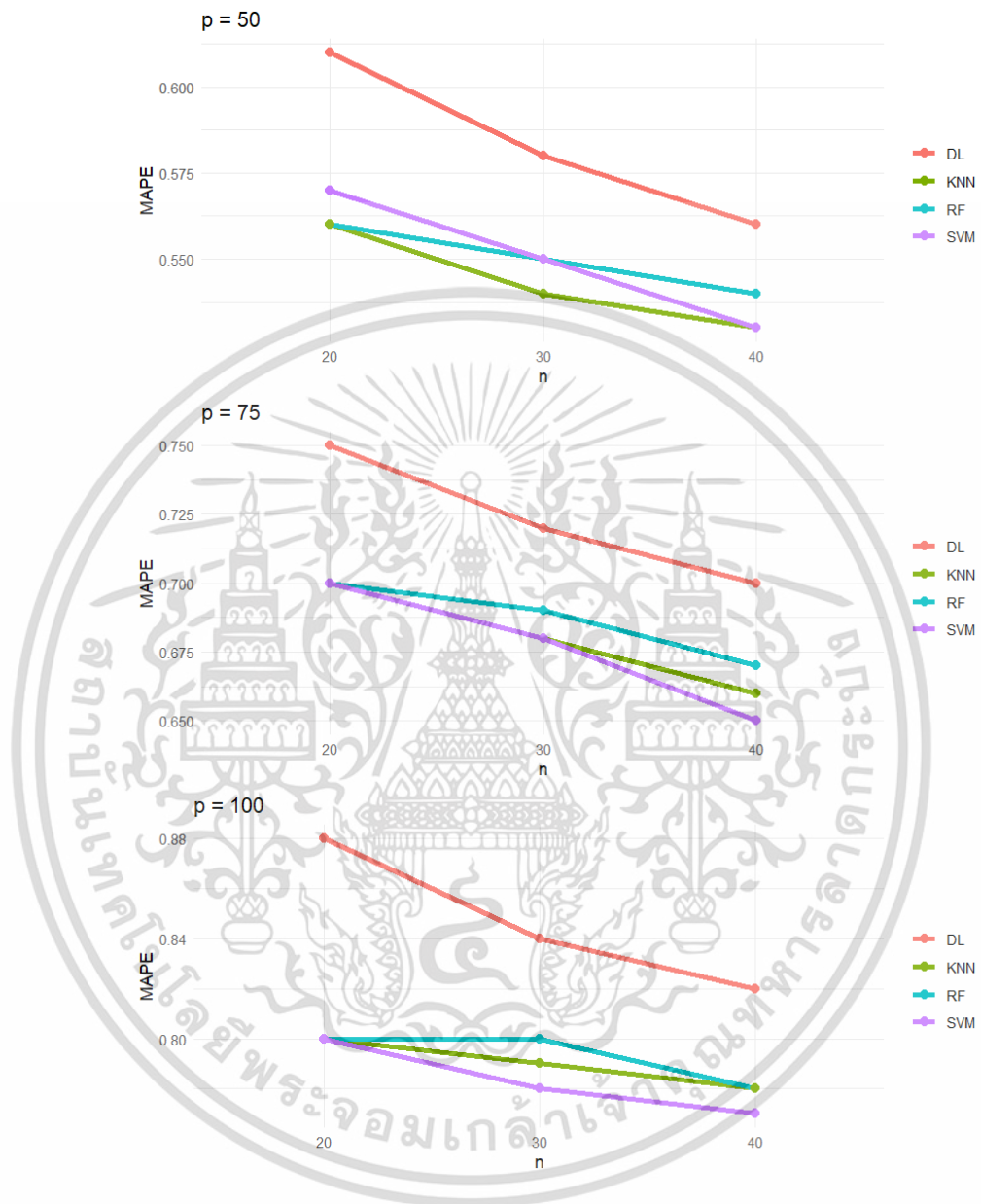
จากตารางที่ 4.1 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.56 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ ร้อยละ 0.54 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.53

เมื่อพิจารณากลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.7 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่ร้อยละ 0.68 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.65

เมื่อพิจารณากลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.8 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.78 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละ

ความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.77 ศึกษานี้ ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.1



รูปที่ 4.1 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0$, $\sigma^2 = 1$, และ $\beta_j = 1$)

จากรูปที่ 4.1 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่น อย่างเป็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี \ n	p = 50			p = 75			p = 100		
	20	30	40	20	30	40	20	30	40
RF	0.59	0.57	0.56	0.72	0.71	0.69	0.82	0.81	0.8
KNN	0.59	0.57	0.56	0.71	0.7	0.69	0.82	0.81	0.8
SVM	0.59	0.57	0.55	0.72	0.69	0.68	0.82	0.8	0.79
DL	0.63	0.6	0.59	0.79	0.74	0.72	0.89	0.86	0.85
SGD	76.6	68.4	65.3	85.24	83.21	79.3	88.4	85.2	83

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

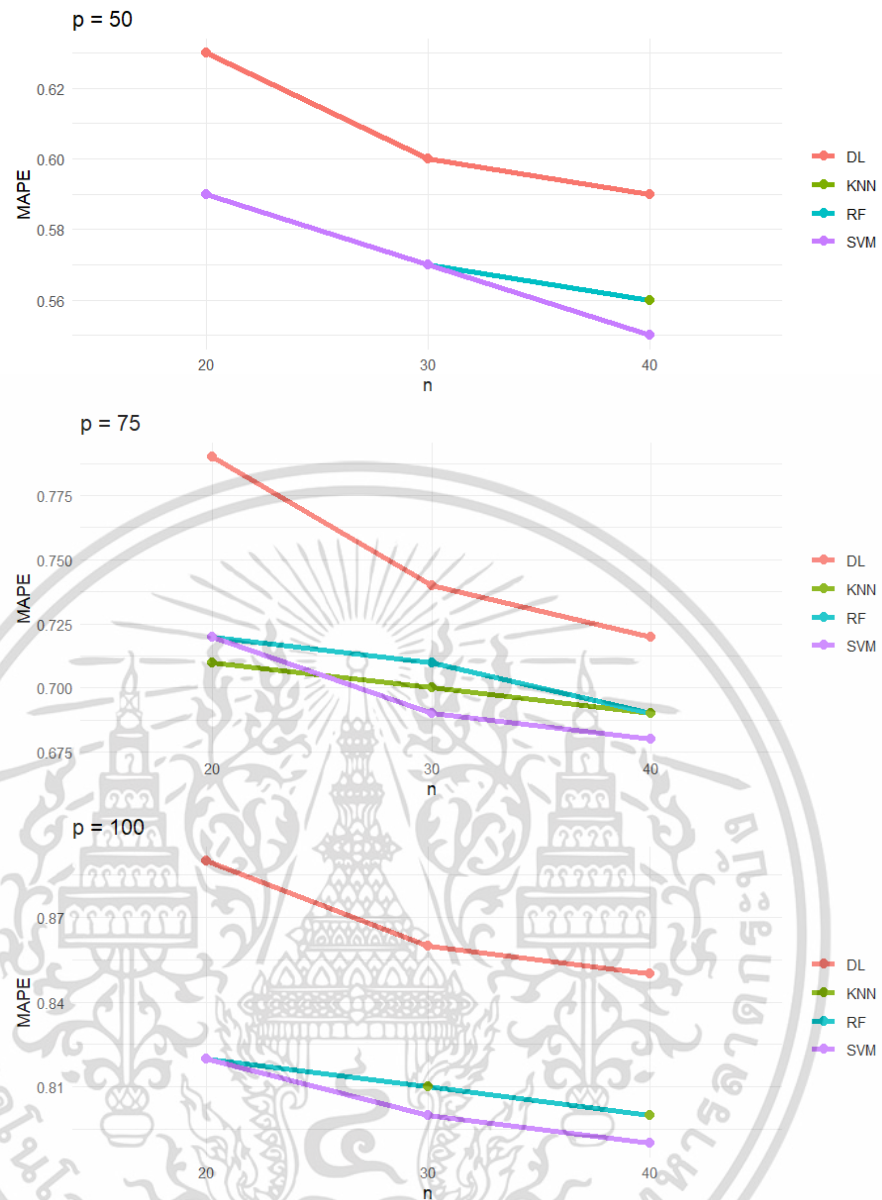
จากตารางที่ 4.2 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.59 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.57 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.55

เมื่อพิจารณากลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.71 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.69 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.68

เมื่อพิจารณากลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.82 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.8 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.79

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0, \sigma^2 = 5$ และ $\beta_j = 1$)

จากรูปที่ 4.2 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี \ n	p = 50			p = 75			p = 100		
	20	30	40	20	30	40	20	30	40
RF	0.7	0.68	0.67	0.8	0.8	0.78	0.89	0.89	0.88
KNN	0.7	0.69	0.68	0.81	0.81	0.79	0.91	0.9	0.89
SVM	0.7	0.68	0.66	0.8	0.79	0.77	0.89	0.88	0.87
DL	0.75	0.72	0.71	0.86	0.85	0.83	0.98	0.95	0.93
SGD	78.4	75.3	68.7	86.4	85.2	83.54	90.3	88.45	85.44

หมายเหตุ : **ตัวหนา** หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

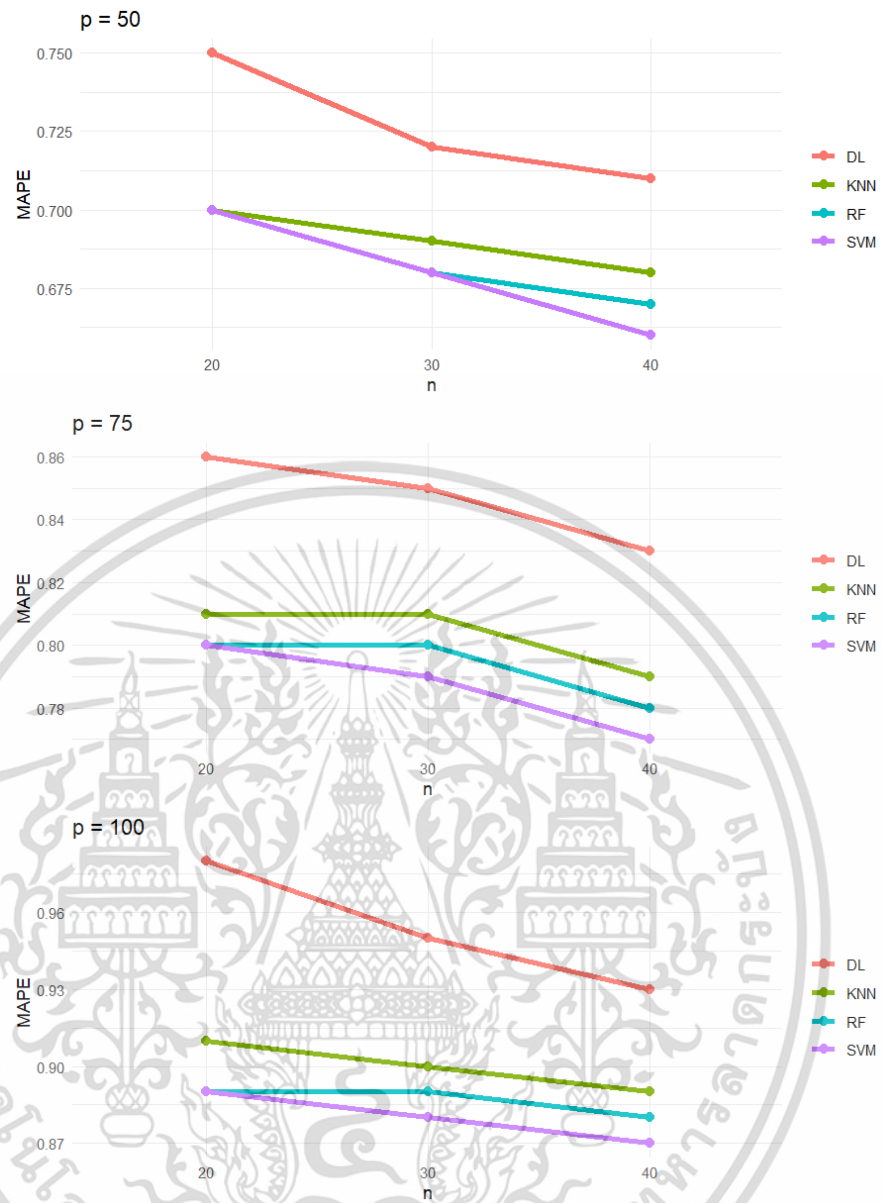
จากตารางที่ 4.3 พบว่าเมื่อพิจารณาในกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.7 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.68 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.66

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.8 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.79 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.77

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.89 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.88 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.87

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0, \sigma^2 = 25, \beta_j = 1$)

จากรูปที่ 4.3 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM โดยจะแสดงผลสรุปวิธีที่ดีที่สุดจากตารางที่ 4.1–4.3 ดังตารางที่ 4.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนน้อยและค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

σ^2	$p = 50$			$p = 75$			$p = 100$		
	$n=20$	$n=30$	$n=40$	$n=20$	$n=30$	$n=40$	$n=20$	$n=30$	$n=40$
1	KNN RF	KNN	KNN SVM	KNN SVM RF	KNN SVM	SVM	KNN SVM RF	SVM	SVM
5	KNN SVM RF	KNN SVM RF	SVM	KNN	SVM	SVM	KNN SVM RF	SVM	SVM
25	KNN SVM RF	SVM RF	SVM	SVM RF	SVM	SVM	SVM RF	SVM	SVM

จากตารางที่ 4.4 จะเห็นว่าในแต่ละจำนวนตัวแปรอิสระและขนาดตัวอย่างบางกลุ่มมีวิธีที่ดีที่สุดมากกว่า 1 วิธี โดยหากพิจารณาจากภาพรวมจะพบว่าวิธีที่ให้ผลลัพธ์ดีในหลายกลุ่มจำนวนตัวแปรอิสระและขนาดตัวอย่างคือ SVM แต่จะมีบางกรณีที่มีวิธีอื่นให้ผลลัพธ์ดีกว่าเช่น ในกรณีที่ปัญหาไม่มีความซับซ้อนมากนักกล่าวคือ ในกรณีที่จำนวนตัวแปรอิสระน้อยและขนาดตัวอย่างน้อย จะเห็นว่าวิธีที่ให้ผลลัพธ์ดีคือ KNN และ SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2 กลุ่มจำนวนตัวแปรอิสระมาก

กำหนดค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อน ขนาดตัวอย่าง เป็นไปตามขอบเขตของงานวิจัย แสดงผลดังตารางที่ 4.5-4.8 และรูปที่ 4.4-4.6

ตารางที่ 4.5 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0$, $\sigma^2 = 1$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี \ n	p = 200			p = 250			p = 300		
	50	100	150	50	100	150	50	100	150
RF	1.12	1.09	1.09	1.26	1.24	1.22	1.38	1.35	1.35
KNN	1.14	1.12	1.11	1.28	1.26	1.25	1.41	1.38	1.38
SVM	1.11	1.07	1.04	1.25	1.22	1.18	1.37	1.33	1.31
DL	1.21	1.12	1.09	1.36	1.28	1.26	1.51	1.42	1.38
SGD	83.09	82.11	80.09	86.3	85.8	84.2	89.09	88.25	87.89

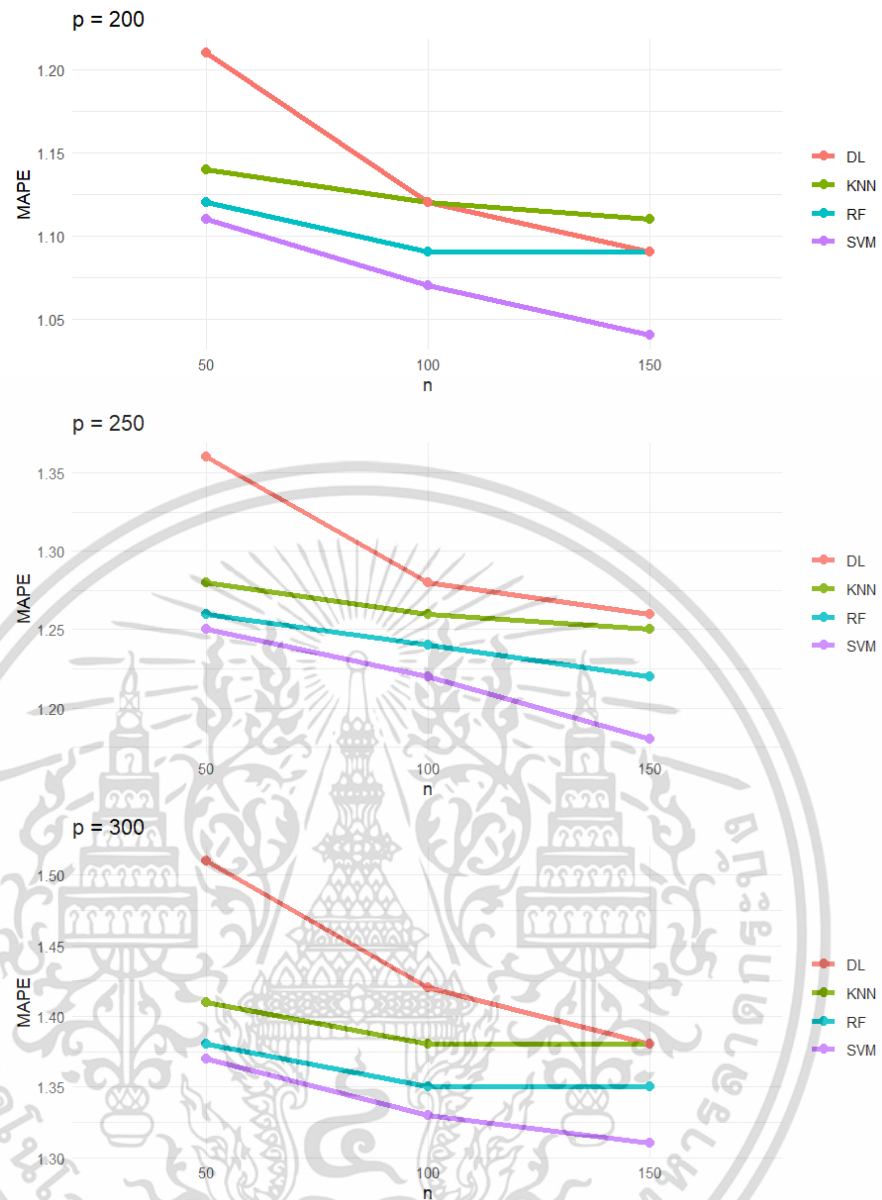
หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

จากตารางที่ 4.5 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.11 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.07 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.04

เมื่อพิจารณากลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.25 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.22 และสุดท้ายกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.18

เมื่อพิจารณากลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.37 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.33 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.31

เอกสารนี้เป็นเอกสารเพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.4 ลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0, \sigma^2 = 1$ และ $\beta_j = 1$)

จากรูปที่ 4.4 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0$, $\sigma^2 = 5$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี \ n	p = 200			p = 250			p = 300		
	50	100	150	50	100	150	50	100	150
RF	1.13	1.11	1.1	1.27	1.25	1.24	1.39	1.36	1.36
KNN	1.15	1.13	1.12	1.29	1.27	1.26	1.42	1.39	1.39
SVM	1.12	1.09	1.06	1.26	1.23	1.19	1.38	1.34	1.32
DL	1.23	1.14	1.1	1.38	1.32	1.28	1.55	1.48	1.41
SGD	84.45	83.45	82.78	87.18	86.78	85.54	90.23	89.56	88.73

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

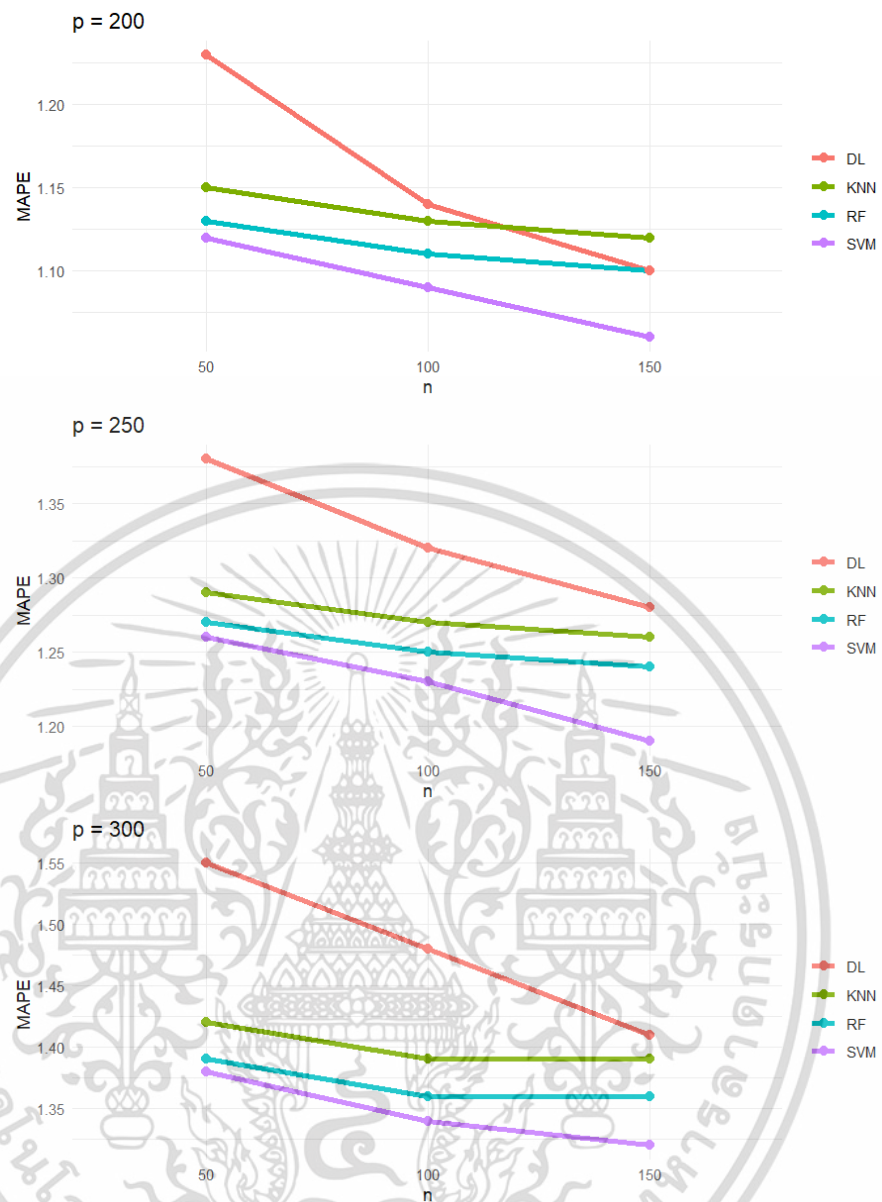
จากตารางที่ 4.6 พบว่าเมื่อพิจารณาในกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.12 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.09 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.06

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.26 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.23 และสุดท้ายกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.19

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.38 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.34 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.32

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0, \sigma^2 = 5$ และ $\beta_j = 1$)

จากรูปที่ 4.5 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

วิธี \ n	p = 200			p = 250			p = 300		
	50	100	150	50	100	150	50	100	150
RF	1.19	1.17	1.16	1.32	1.3	1.29	1.44	1.4	1.41
KNN	1.22	1.2	1.19	1.35	1.33	1.32	1.47	1.45	1.44
SVM	1.18	1.15	1.12	1.32	1.28	1.25	1.43	1.39	1.37
DL	1.24	1.17	1.15	1.41	1.28	1.38	1.57	1.5	1.45
SGD	86.02	84.55	83.42	88.33	87.63	87.63	90.56	90.56	89.4

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

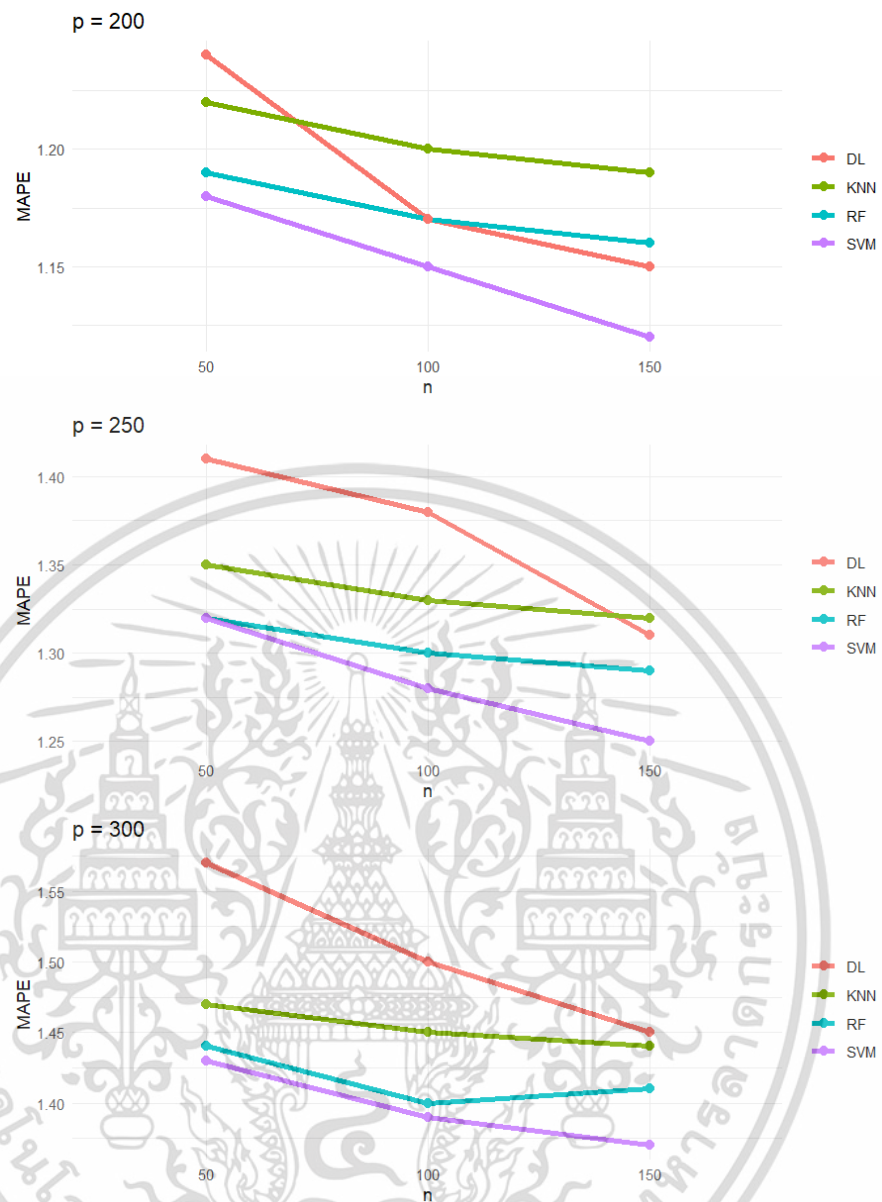
จากตารางที่ 4.7 พบว่าเมื่อพิจารณาในกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.18 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.15 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.12

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.32 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.28 และสุดท้ายกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.25

เมื่อพิจารณาในกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.43 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.39 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.37

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0, \sigma^2 = 25$ และ $\beta_j = 1$)

จากรูปที่ 4.6 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM โดยจะแสดงผลสรุปวิธีที่ดีที่สุดจากตารางที่ 4.5–4.7 ดังตารางที่ 4.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนมากและค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

σ^2	$p = 200$			$p = 250$			$p = 300$		
	$n = 50$	$n = 100$	$n = 150$	$n = 50$	$n = 100$	$n = 150$	$n = 50$	$n = 100$	$n = 150$
1	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM
5	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM
25	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM

จากตารางที่ 4.8 จะเห็นว่าหากพิจารณาจากภาพรวมจะพบว่าวิธีที่ให้ผลลัพธ์ดีในเกือบทุกกลุ่มจำนวนตัวแปรอิสระและขนาดตัวอย่างคือ SVM



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 กรณีเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

4.2.1 กลุ่มจำนวนตัวแปรอิสระน้อย

กำหนดค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อน ขนาดตัวอย่าง เป็นไปตามขอบเขตของงานวิจัย แสดงผลดังตารางที่ 4.9-4.12 และรูปที่ 4.7-4.9

ตารางที่ 4.9 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0$, $\sigma^2 = 1$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี	n	p = 50			p = 75			p = 100		
		20	30	40	20	30	40	20	30	40
RF		0.56	0.55	0.54	0.7	0.69	0.67	0.81	0.8	0.78
KNN		0.56	0.54	0.53	0.7	0.68	0.66	0.81	0.79	0.78
SVM		0.57	0.55	0.53	0.71	0.68	0.66	0.81	0.79	0.77
DL		0.61	0.58	0.56	0.75	0.73	0.70	0.87	0.84	0.83
SGD		78.56	77.4	77.4	81.36	79.3	77.3	86.18	85.25	83.24

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

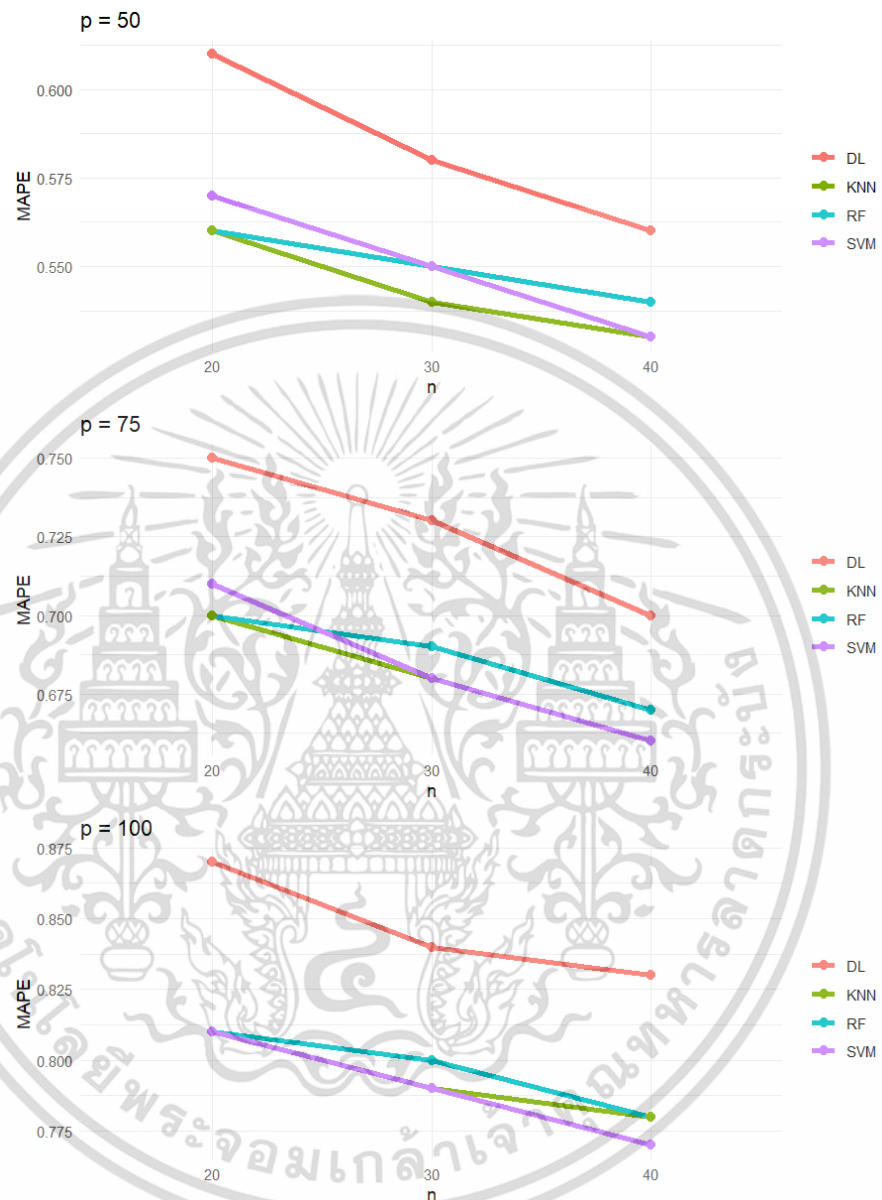
จากตารางที่ 4.9 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.56 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.54 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.53

เมื่อพิจารณากลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี RF และวิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.7 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.68 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.66

เมื่อพิจารณากลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.81 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.79 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้

เอกสารนี้ ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.77 กรุณาอย่าไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.7



รูปที่ 4.7 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0, \sigma^2 = 1$ และ $\beta_j = -1$)

จากรูปที่ 4.7 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด เมื่อพิจารณาไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่มที่ p เท่ากับ 50 วิธีที่ดีที่สุดโดยรวมคือ KNN และหากพิจารณากลุ่ม p เท่ากับ 75 และ 100 จะพบว่าวิธีที่ดีที่สุดโดยรวมคือ SVM

ตารางที่ 4.10 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี	n			$p = 50$			$p = 75$			$p = 100$		
	20	30	40	20	30	40	20	30	40	20	30	40
RF	0.59	0.57	0.56	0.73	0.71	0.69	0.82	0.81	0.8			
KNN	0.59	0.57	0.56	0.72	0.7	0.69	0.82	0.81	0.8			
SVM	0.59	0.57	0.55	0.73	0.7	0.68	0.82	0.8	0.79			
DL	0.63	0.61	0.59	0.78	0.75	0.72	0.89	0.86	0.85			
SGD	79.5	78.1	76.5	82.5	81.77	79.36	87.3	86.44	85.63			

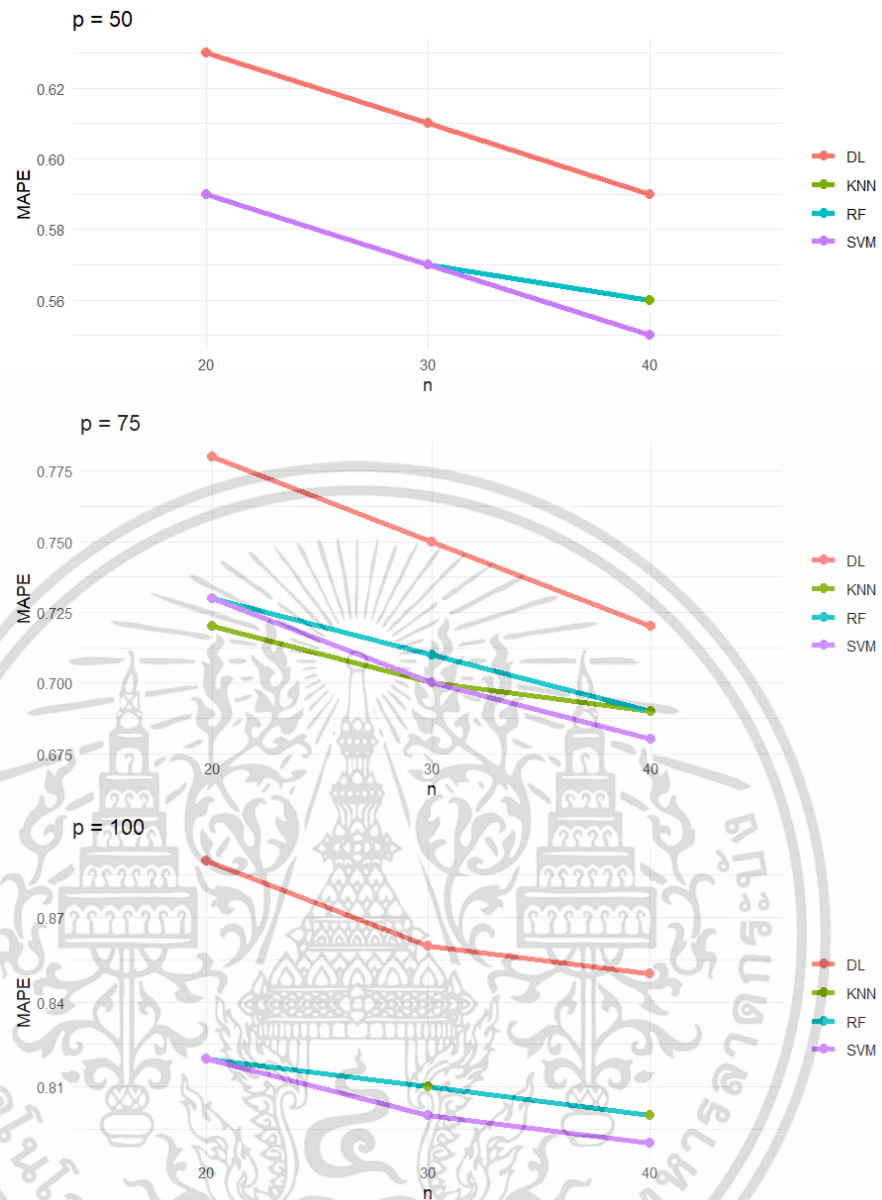
หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

จากตารางที่ 4.10 พบว่าเมื่อพิจารณาจากกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.59 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.57 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.55

เมื่อพิจารณาจากกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี KNN ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.72 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.7 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.68

เมื่อพิจารณาจากกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.82 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.8 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.79

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละเอกสารนี้ความผิดพลาดสัมบูรณ์ดังรูปที่ 4.8 งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0, \sigma^2 = 5$ และ $\beta_j = -1$)

จากรูปที่ 4.8 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.11 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระน้อยเมื่อ $\mu = 0$, $\sigma^2 = 25$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี \ n	p = 50			p = 75			p = 100		
	20	30	40	20	30	40	20	30	40
RF	0.69	0.69	0.67	0.82	0.8	0.79	0.9	0.9	0.88
KNN	0.7	0.69	0.68	0.83	0.8	0.79	0.91	0.9	0.89
SVM	0.69	0.69	0.66	0.82	0.79	0.78	0.91	0.89	0.87
DL	0.75	0.73	0.71	0.88	0.86	0.83	0.99	0.96	0.95
SGD	80.4	79.6	78.5	84.66	82.78	81.33	88.6	87.88	86.98

หมายเหตุ : **ตัวหนา** หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

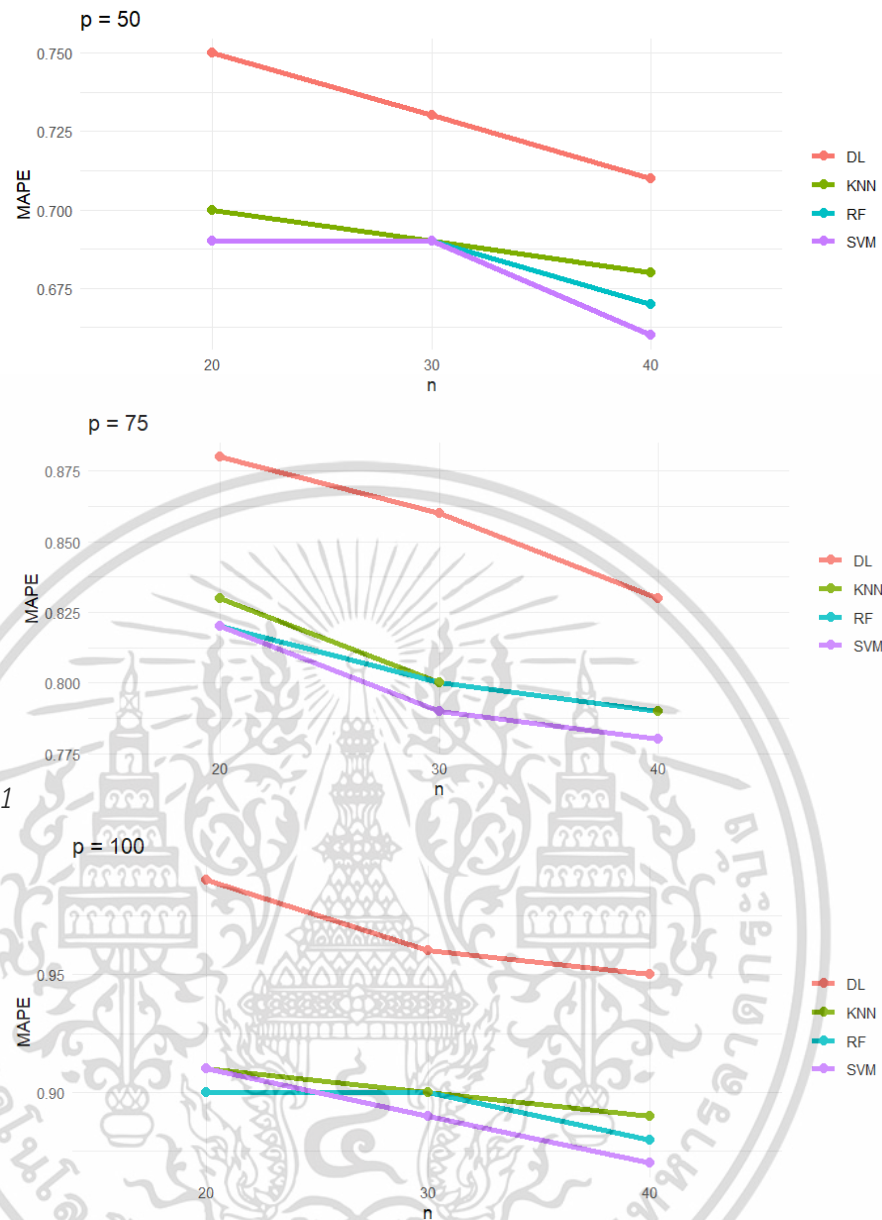
จากตารางที่ 4.11 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 20 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.69 กลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 30 จะเห็นว่าวิธี RF วิธี KNN และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.69 และสุดท้ายกลุ่มที่ p เท่ากับ 50 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.66

เมื่อพิจารณากลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 20 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.82 กลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.79 และสุดท้ายกลุ่มที่ p เท่ากับ 75 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.78

เมื่อพิจารณากลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 20 จะเห็นว่าวิธี RF ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.9 กลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 30 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.89 และสุดท้ายกลุ่มที่ p เท่ากับ 100 และ n เท่ากับ 40 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 0.87

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.9 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 50, 75$ และ 100 $n = 20, 30$ และ 40 ($\mu = 0, \sigma^2 = 25$ และ $\beta_j = -1$)

จากรูปที่ 4.9 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM โดยจะแสดงผลสรุปวิธีที่ดีที่สุดจากตารางที่ 4.9–4.11 ดังตารางที่ 4.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.12 ตารางแสดงผลวิธีที่ดีที่สุดที่สุ่มกลุ่มตัวแปรอิสระจำนวนน้อยและค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

σ^2	$p = 50$			$p = 75$			$p = 100$		
	$n=20$	$n=30$	$n=40$	$n=20$	$n=30$	$n=40$	$n=20$	$n=30$	$n=40$
1	KNN RF	KNN	KNN SVM	KNN RF	KNN SVM	KNN SVM	KNN SVM RF	KNN SVM	SVM
5	KNN SVM RF	KNN SVM RF	SVM	KNN	KNN SVM	SVM	KNN SVM RF	SVM	SVM
25	SVM RF	KNN SVM RF	SVM	SVM RF	SVM	SVM	RF	SVM	SVM

จากตารางที่ 4.12 จะเห็นว่าในแต่ละจำนวนตัวแปรอิสระและขนาดตัวอย่างบางกลุ่มมีวิธีที่ดีที่สุดมากกว่า 1 วิธี โดยหากพิจารณาจากภาพรวมจะพบว่าวิธีที่ให้ผลลัพธ์ดีในหลายกลุ่มจำนวนตัวแปรอิสระและขนาดตัวอย่างคือ SVM แต่จะมีบางกรณีที่มีวิธีอื่นให้ผลลัพธ์ดีกว่าเช่น ในกรณีที่ปัญหาไม่มีความซับซ้อนมากนักกล่าวคือ ในกรณีที่จำนวนตัวแปรอิสระน้อยและขนาดตัวอย่างน้อย จะเห็นว่าวิธีที่ให้ผลลัพธ์ดีคือ KNN และ SVM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 กลุ่มจำนวนตัวแปรอิสระมาก

กำหนดค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อน ขนาดตัวอย่าง เป็นไปตามขอบเขตของงานวิจัย แสดงผลดังตารางที่ 4.13-4.16 และรูปที่ 4.10-4.12

ตารางที่ 4.13 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0$, $\sigma^2 = 1$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี	n	p = 200			p = 250			p = 300		
		50	100	150	50	100	150	50	100	150
RF		1.12	1.09	1.09	1.26	1.23	1.22	1.38	1.35	1.35
KNN		1.13	1.12	1.1	1.28	1.26	1.25	1.41	1.38	1.38
SVM		1.11	1.07	1.04	1.25	1.21	1.18	1.37	1.33	1.31
DL		1.20	1.12	1.05	1.43	1.28	1.19	1.52	1.42	1.24
SGD		83.08	80.5	78.63	86.28	84.53	81.53	89.09	88.44	85.63

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

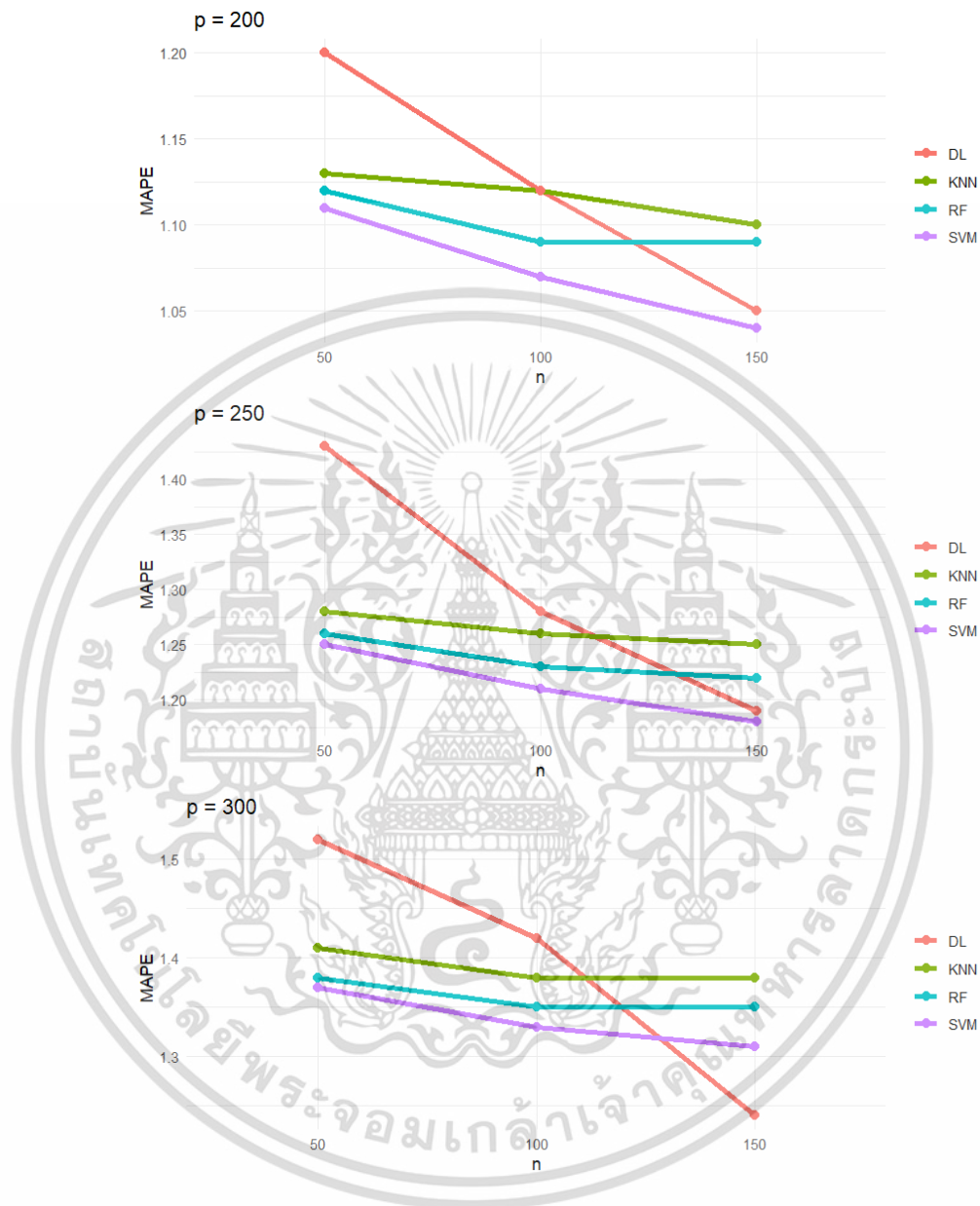
จากตารางที่ 4.13 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.11 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.07 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.04

เมื่อพิจารณากลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.25 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ 1.21 และสุดท้ายกลุ่มที่ p เท่ากับร้อยละ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.18

เมื่อพิจารณากลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.37 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.33 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี DL ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำ

เอกสารนี้ที่สุดที่ร้อยละ 1.24 สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.10



รูปที่ 4.10 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0$, $\sigma^2 = 1$ และ $\beta_j = -1$)

จากรูปที่ 4.10 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM ยกเว้นกรณีที่ $p = 300$ $n = 150$ ที่ วิธี DL ได้ผลดีที่สุดไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.14 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 5$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี \ n	p = 200			p = 250			p = 300		
	50	100	150	50	100	150	50	100	150
RF	1.13	1.11	1.1	1.27	1.24	1.23	1.39	1.36	1.36
KNN	1.14	1.13	1.12	1.29	1.27	1.26	1.42	1.39	1.39
SVM	1.12	1.08	1.05	1.26	1.22	1.19	1.39	1.34	1.32
DL	1.23	1.14	1.11	1.47	1.30	1.26	1.53	1.43	1.28
SGD	84.12	83.42	80.79	87.44	86.45	85.53	90.5	86.56	86.11

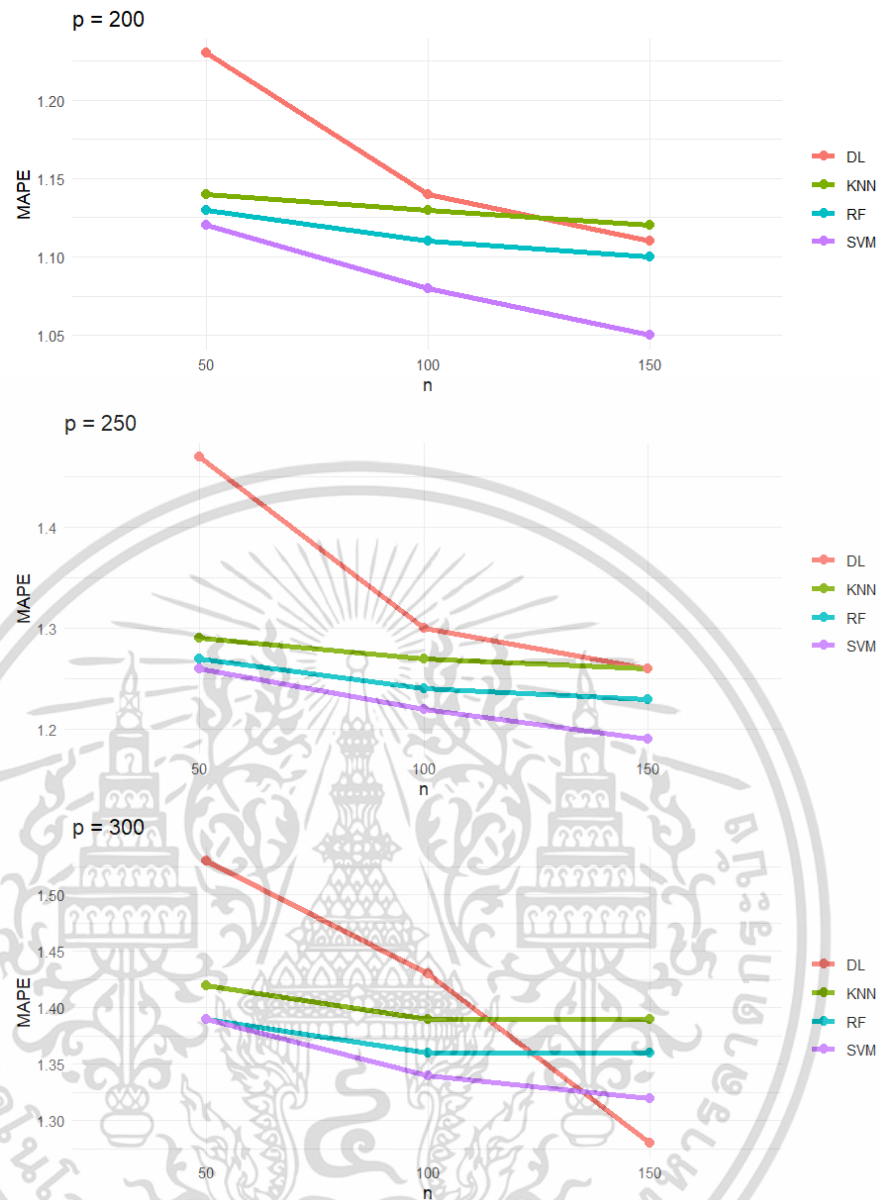
หมายเหตุ : **ตัวหนา** หมายถึง ค่าเฉลี่ยของเปอร์เซ็นต์ความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

จากตารางที่ 4.14 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.12 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.08 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.05

เมื่อพิจารณากลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.26 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.22 และสุดท้ายกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.19

เมื่อพิจารณากลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.39 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.34 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี DL ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.28

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.11



รูปที่ 4.11 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0, \sigma^2 = 5$ และ $\beta_j = -1$)

รูปที่ 4.11 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM ยกเว้นกรณีที่ $p = 300$ $n = 150$ ที่ วิธี DL ได้ผลดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.15 ผลการเปรียบเทียบค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีจากจำนวนตัวแปรอิสระมากเมื่อ $\mu = 0, \sigma^2 = 25$ และค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

วิธี \ n	p = 200			p = 250			p = 300		
	50	100	150	50	100	150	50	100	150
RF	1.18	1.17	1.15	1.32	1.29	1.28	1.44	1.41	1.41
KNN	1.21	1.2	1.18	1.35	1.33	1.32	1.47	1.45	1.44
SVM	1.18	1.14	1.11	1.32	1.27	1.25	1.43	1.39	1.37
DL	1.29	1.21	1.13	1.49	1.46	1.39	1.55	1.48	1.31
SGD	86.44	84.52	82.11	89.63	88.55	86.98	91.3	90.78	87.42

หมายเหตุ : ตัวหนา หมายถึง ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่มีค่าต่ำที่สุด

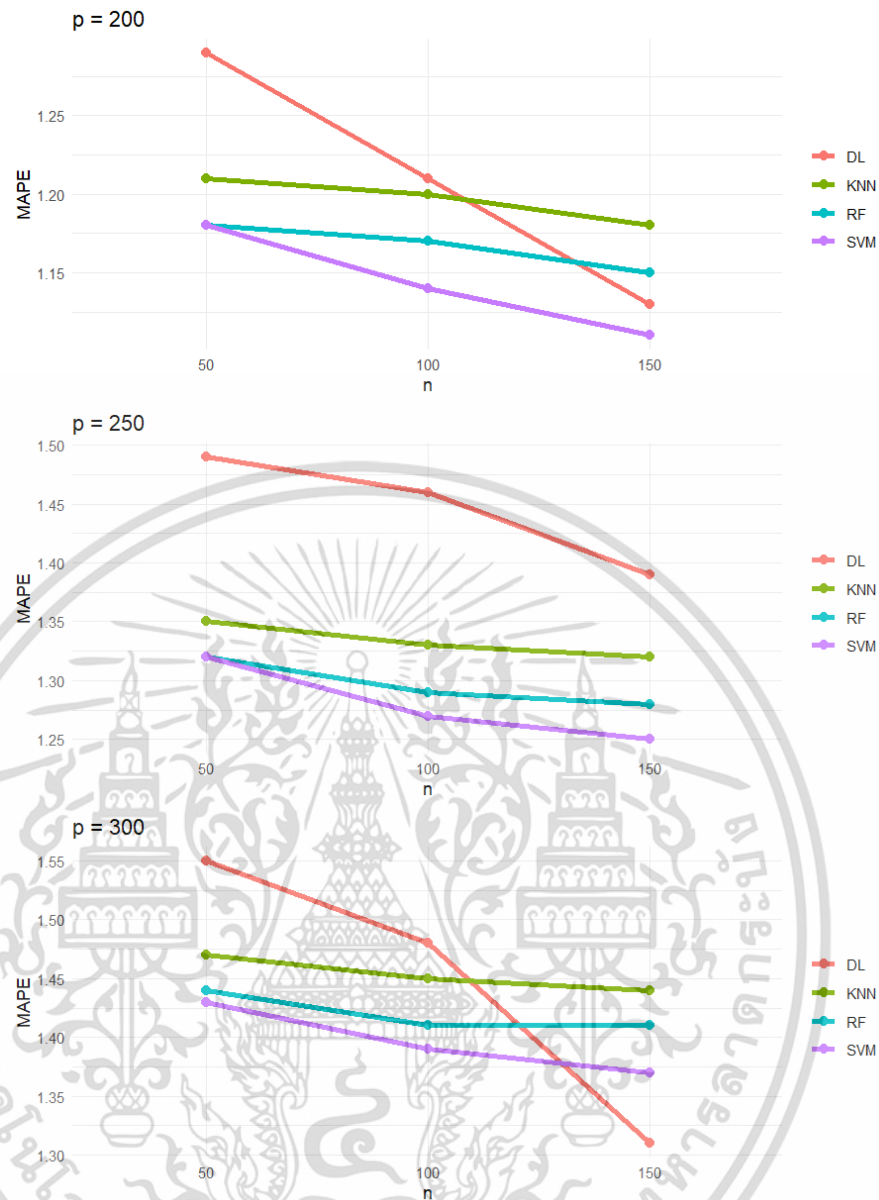
จากตารางที่ 4.15 พบว่าเมื่อพิจารณากลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 50 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.18 กลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.14 และสุดท้ายกลุ่มที่ p เท่ากับ 200 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.11

เมื่อพิจารณากลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 50 จะเห็นว่าวิธี RF และวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.32 กลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.27 และสุดท้ายกลุ่มที่ p เท่ากับ 250 และ n เท่ากับ 150 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.25

เมื่อพิจารณากลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 50 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.43 กลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 100 จะเห็นว่าวิธี SVM ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.39 และสุดท้ายกลุ่มที่ p เท่ากับ 300 และ n เท่ากับ 150 จะเห็นว่าวิธี DL ให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ต่ำที่สุดที่ร้อยละ 1.31

เพื่อให้ง่ายต่อการวิเคราะห์ผลจึงทำการสร้างกราฟสำหรับการแสดงผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ดังรูปที่ 4.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 กราฟแสดงค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ของวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี โดยกำหนดให้ $p = 200, 250$ และ 300 $n = 50, 100$ และ 150 ($\mu = 0, \sigma^2 = 25$ และ $\beta_j = -1$)

จากรูปที่ 4.12 เนื่องจากว่าวิธี SGD มีค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มากกว่าวิธีอื่นอย่างเห็นได้ชัดจึงตัดออกจากการสร้างกราฟแสดงผลและเมื่อพิจารณาจากทั้ง 3 กราฟ เมื่อเพิ่มขนาดตัวอย่างแนวโน้มค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงอย่างเห็นได้ชัด โดยหากมองจากภาพรวมแล้ววิธีที่ดีที่สุดคือวิธี SVM ยกเว้นกรณีที่ $p = 300$ $n = 150$ ที่ วิธี DL ได้ผลดีที่สุด โดยจะแสดงผลสรุปวิธีที่ดีที่สุดจากตารางที่ 4.13–4.15 ดังตารางที่ 4.16

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.16 ตารางแสดงผลวิธีที่ดีที่สุดกรณีกลุ่มตัวแปรอิสระจำนวนมากและค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

σ^2	$p = 200$			$p = 250$			$p = 300$		
	$n=50$	$n=100$	$n=150$	$n=50$	$n=100$	$n=150$	$n=50$	$n=100$	$n=150$
1	SVM	SVM	SVM	SVM	SVM	SVM	SVM	SVM	DL
5	SVM	SVM	SVM	SVM	SVM	SVM	SVM RF	SVM	DL
25	SVM RF	SVM	SVM	SVM RF	SVM	SVM	SVM	SVM	DL

จากตารางที่ 4.16 จะเห็นว่าหากพิจารณาจากภาพรวมจะพบว่าวิธีที่ให้ผลลัพธ์ดีในเกือบทุกกลุ่มจำนวนตัวแปรอิสระและขนาดตัวอย่างคือ SVM จะมีเพียงกรณีที่ข้อมูลมีความซับซ้อนมากแล้วจำนวนตัวแปรอิสระมากและกลุ่มตัวอย่างขนาดใหญ่ที่วิธี DL ให้ผลลัพธ์ดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 อภิปรายผล

จากการศึกษาทำการเปรียบเทียบประสิทธิภาพตัวแบบการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธีประกอบไปด้วยวิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านที่ใกล้ที่สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์ โดยที่ขนาดตัวอย่าง จำนวนตัวแปรอิสระ ค่าสัมประสิทธิ์การถดถอย ค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อนมีค่าที่แตกต่างกันผลการวิจัยพบว่าส่วนใหญ่ วิธีซัพพอร์ตเวกเตอร์แมชชีนเป็นวิธีที่ดีที่สุดซึ่งสอดคล้องกับงานวิจัยของจุฬาลักษณ์ และศิวา (2563) ที่ได้ศึกษาตัวแบบการเรียนรู้ของเครื่องที่เหมาะสมสำหรับการประมาณค่าความเข้มข้นของตะกอนแขวนลอยในแม่น้ำลำปาวโดยใช้ข้อมูลจากดาวเทียม ตัวแบบที่นำมาเปรียบเทียบประกอบด้วย วิธีการถดถอยเชิงเส้น วิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการเรียนรู้เชิงลึก ผลการศึกษาพบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนเป็นตัวแบบที่มีประสิทธิภาพดีที่สุด โดยจะพิจารณาวิธีที่ดีที่สุดจากค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ในแต่ละกลุ่มจำนวนตัวแปรอิสระ(ใช้วิธีการโหวต)จากตารางที่ 4.4 ตารางที่ 4.8 ตารางที่ 4.12 และตารางที่ 4.16 แสดงดังตารางที่ 4.17 และตารางที่ 4.18

ตารางที่ 4.17 วิธีที่ดีที่สุดเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

ความแปรปรวน (σ^2)	จำนวนตัวแปรอิสระ (p)					
	50	75	100	150	200	250
1	KNN	SVM	SVM	SVM	SVM	SVM
5	SVM	SVM	SVM	SVM	SVM	SVM
25	SVM	SVM	SVM	SVM	SVM	SVM

ตารางที่ 4.18 วิธีที่ดีที่สุดเมื่อค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

ความแปรปรวน (σ^2)	จำนวนตัวแปรอิสระ (p)					
	50	75	100	150	200	250
1	KNN	KNN	SVM	SVM	SVM	SVM
5	SVM	SVM	SVM	SVM	SVM	SVM
25	SVM	SVM	SVM	SVM	SVM	SVM

จากตารางที่ 4.17 เมื่อพิจารณาวิธีที่ดีที่สุดจากค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ในแต่ละกลุ่มจำนวนตัวแปรอิสระ พบว่าส่วนใหญ่วิธี SVM ให้ผลดีที่สุดมีเพียงกลุ่มที่ p เท่ากับ 50 σ^2 เท่ากับ 1 ที่ KNN ดีที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.18 เมื่อพิจารณาวิธีที่ดีที่สุดจากค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ในแต่ละกลุ่มจำนวนตัวแปรอิสระ พบว่าส่วนใหญ่วิธี SVM ให้ผลดีที่สุดมีเพียงกลุ่มที่ p เท่ากับ 50 และ $75 \sigma^2$ เท่ากับ 1 ที่ KNN ดีที่สุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้จะทำการศึกษการเปรียบเทียบประสิทธิภาพการทำนายข้อมูลการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่อง โดยนำข้อมูลที่ได้จากการจำลองมาวิเคราะห์ด้วยวิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนเดสเซนท์ เมื่อจำนวนตัวแปรอิสระ ขนาดตัวอย่าง ค่าสัมประสิทธิ์การถดถอย ค่าเฉลี่ยและความแปรปรวนของค่าคลาดเคลื่อนที่ต่างกัน เพื่อหาวิธีที่ดีที่สุดสำหรับข้อมูลมิติขั้นสูง

5.1 สรุปผลวิจัย

จากการศึกษาและเปรียบเทียบประสิทธิภาพการทำนายข้อมูลการถดถอยเชิงเส้นสำหรับข้อมูลมิติขั้นสูงด้วยวิธีการเรียนรู้ของเครื่องทั้ง 5 วิธี จากค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ สามารถสรุปผลการวิเคราะห์ได้ดังนี้

5.1.1 กรณีพิจารณาค่าสัมประสิทธิ์การถดถอยเท่ากับ 1

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยเท่ากับ 1 และพิจารณาค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่ให้ค่าต่ำที่สุด จะพบว่าโดยส่วนใหญ่วิธีซัพพอร์ตเวกเตอร์แมชชีนคือวิธีที่ดีที่สุดแต่มีบางกรณีเช่น กรณีที่กลุ่มจำนวนตัวแปรอิสระน้อยและขนาดตัวอย่างน้อยวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับจะให้ผลลัพธ์ที่ดีที่สุด

5.1.2 กรณีพิจารณาค่าสัมประสิทธิ์การถดถอยเท่ากับ -1

เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยเท่ากับ -1 และพิจารณาค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่ให้ค่าต่ำที่สุด จะพบว่าโดยส่วนใหญ่วิธีซัพพอร์ตเวกเตอร์แมชชีนคือวิธีที่ดีที่สุดแต่มีบางกรณีเช่น กรณีที่กลุ่มจำนวนตัวแปรอิสระน้อยและขนาดตัวอย่างน้อยวิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับจะให้ผลลัพธ์ที่ดีที่สุดและ กรณีที่กลุ่มจำนวนตัวแปรอิสระมากและขนาดตัวอย่างมีขนาดใหญ่วิธีการเรียนรู้เชิงลึกให้ผลลัพธ์ที่ดีที่สุด

5.1.3 ข้อสังเกต

จำนวนตัวแปรอิสระ ขนาดตัวอย่างและความแปรปรวนของค่าคลาดเคลื่อนจะส่งผลต่อค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ที่แตกต่างกันไปดังนี้ ในกรณีที่ขนาดตัวอย่างคงที่การเพิ่มจำนวนตัวแปรอิสระจะส่งผลให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าเพิ่มมากขึ้น ในกรณีที่จำนวนตัวแปรอิสระคงที่การเพิ่มขนาดตัวอย่างจะส่งผลให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าลดลงและสุดท้าย ในการเพิ่มความแปรปรวนของค่าคลาดเคลื่อนจะส่งผลให้ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์มีค่าเพิ่มมากขึ้น

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่ในเชิงพาณิชย์โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น ยกเว้นให้ไม่มีผิดแต่สิ่งใดก็ตาม และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อเสนอแนะ

1. สำหรับผู้ที่สนใจจะนำงานวิจัยไปต่อยอด ควรพิจารณาในการสร้างข้อมูลประเภทอื่นเช่น ข้อมูลการถดถอยไม่เชิงเส้น
2. เนื่องจากข้อมูลสำหรับงานวิจัยนี้เป็นข้อมูลที่ถูกจำลองขึ้นมา ผลสรุปจึงไม่ครอบคลุมข้อมูลทั้งหมด
3. งานวิจัยครั้งนี้ใช้ค่าเฉลี่ยของร้อยละความผิดพลาดสมบูรณ์เป็นเกณฑ์ในการวัดผล แต่ในการประยุกต์ใช้กับธุรกิจอาจต้องพิจารณาปัจจัยทางด้านเวลาเข้ามาเกี่ยวข้องด้วย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- กฤติกา ทิพย์คำมี, อนุพงศ์ สุขประเสริฐ, สุพัตรา กอมจัญ และณัฐกานต์ ชุตติมารังสรรค์. 2566. ประสิทธิภาพของเทคนิคการเรียนรู้ของเครื่องสำหรับการพยากรณ์ฝุ่นละอองขนาดเล็กในอากาศ. วารสารวิทยาศาสตร์ เทคโนโลยี และนวัตกรรม มหาวิทยาลัยกาฬสินธุ์. 1: 58-74.
- ฆฤณ ชินประสาทศักดิ์, ปรีชาพล อินทร์สุข, นันทนัช พุสามปือก, ณัฐพล สนพะเนา. 2566. แก่นแท้ของ Deep Learning : AI ฉบับอธิบายด้วยภาพ. พิมพ์ครั้งที่ 1. เมดบายเอไอ จุฬาลงกรณ์มหาวิทยาลัย และศิลา แก้วปลั่ง. 2563. การประมาณค่าความเข้มข้นของตะกอนแขวนลอยในแม่น้ำโดยใช้ข้อมูลภาพถ่ายดาวเทียม Sentinel-2: กรณีศึกษาลุ่มแม่น้ำลำปาว. วารสารวิทยาศาสตร์และเทคโนโลยีมหาวิทยาลัยมหาสารคาม. 39: 446-456.
- นัทธมน พลายอินทร์, ชนาพันธ์ ชนาเนตรและวิกานดา ผาพันธ์. 2563. การพยากรณ์ค่า Scale ของแผ่น PCB ชนิด Multilayer โดยใช้การเรียนรู้ของเครื่อง. วารสารวิทยาศาสตร์และเทคโนโลยีมหาวิทยาลัยมหาสารคาม. 39: 693-701
- ปริญญา สงวนสัตย์. 2562. Artificial Intelligence with Machine Learning AI สร้างได้ด้วยแมชชีนเลิร์นนิง. พิมพ์ครั้งที่ 1. ไอทีซี พรีเมียร์.
- พรทิพย์ เดชพิชัย, ลัทธพล โขครัตน์ประภา, สุพิชฌ์ ศรีแพทย์, ณรรธคุณ วิรุฬห์ศรี. 2565. การเปรียบเทียบตัวแบบการทำนายระดับและการเปลี่ยนระดับคุณค่าของลูกค้าในระยะยาว. วารสารการจัดการสมัยใหม่. 2: 113-142.
- พิศิษฐ์ บวรเลิศสุธี และ วรภัทร ไพรีเกรง . 2565. ตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์. Journal of Engineering and Digital Technology (JEDT). 10: 71-79
- ราชบัณฑิตยสภา. 2561. พจนานุกรมศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสภา พิมพ์ครั้งที่ 2. กรุงเทพฯ. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย
- ราชนฤธา รัตนวรกานต์. 2562. การหาค่าเหมาะสมที่สุดของการเรียนรู้เชิงลึกโดยใช้อัลกอริทึมเชิงวิวัฒนาการ. วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์. มหาวิทยาลัยศิลปากร
- วิฐุรา พิงพาพงศ์. 2558. บทวิเคราะห์วิธีวิเคราะห์การถดถอยเชิงเส้นสำหรับข้อมูลที่มีมิติสูง. วารสารวิทยาศาสตร์และเทคโนโลยี. 2: 212-223.
- วิรัช หิรัญ และฐิตาภรณ์ พ่อบุตรดี. 2561. แบบจำลองการติดตามการเดินทางโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์และการเรียนรู้เชิงลึก. วารสารศรีประทุมปริทัศน์ ฉบับวิทยาศาสตร์และเทคโนโลยี. 10: 146-157

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- โสภี แก้วชะภา, สมพร เรืองอ่อน และอุทัย คูหาพงศ์. 2565. **เปรียบเทียบประสิทธิภาพของตัวแบบการเรียนรู้ของเครื่องสำหรับทำนายราคาพริกชี้หนู**. วารสารวิชาการชายันต์เทคโนโลยี. 2: 1-11.
- Abu-Shaira, M and Speegle, G. 2023. **Online Linear Regression Based on Weighted Average**. arXiv. Null: 2-14.
- Biau, G, Devroye, L, Dujmović, V and Krzyzak, A. 2012. **An affine invariant k-nearest neighbor regression estimate**. Journal of Multivariate Analysis. 112: 24-34
- Cai, T., Gao, R. Hou, J. Chen, S. Wang, D. He, D. Zhang, Z and Wang, L. 2019. **A Gram-Gauss-Newton Method Learning Overparameterized Deep Neural Networks for Regression Problems**. arXiv.
- Haji, S and Abdulazeez, A. 2021. **Comparison Of Optimization Techniques Based On Gradient Descent Algorithm: A Review**. PalArch's Journal of Archaeology of Egypt / Egyptology (PJAEE). 18(4): 2715-2743
- Imandoust,S and Bolandraftar,M.2013. **Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background**. Journal of Engineering Research and Applications.3:605-610.
- Schonlau,M and Zou,R. 2020. **The random forest algorithm for statistical learning**. The Stata Journal (2020).1:3-29.
- Sun, Y. 2021. **Investigation on House Price Prediction with Various Gradient Descent Method**. Journal of Physics: Conference series. 1827: 1-7.
<https://doi.org/10.1088/1742-6596/1827/1/012186>
- Vapnik, V. N. 1995. **The nature of statistical learning theory**: Springer-Verlag New York, Inc.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำสั่งโปรแกรม R ที่ใช้ในงานวิจัย

วิธีใช้ในการวิเคราะห์ข้อมูลทั้ง 5 วิธี ประกอบไปด้วยวิธีป่าสุ่ม วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการเรียนรู้เชิงลึกและวิธีสโตแคสติกเกรเดียนต์เดสเซนท์

วิธีป่าสุ่ม

```
library(randomForest)
rnd <- 1000
mse_train <- vector('integer',length=rnd )
mape_train <- vector('integer',length=rnd )
mse_test <- vector('integer',length=rnd )
mape_test <- vector('integer',length=rnd )
for (r in 1:rnd)
{
  set.seed(r)
  #set parameter
  cat("\nRound :",r,"\n")
  n <- 100
  P <- 300
  varr <- 5
  error <- rnorm(n, mean=0, sd=sqrt(varr))
  X = vector(mode='list', length=P)
  #B0 is large
  Y = 1000
  beta = 1
  for (i in 1:P)
  {
    X[[i]] = rnorm(n, mean=0, sd=1)
    Y = Y + (beta*X[[i]])
  }
  Y = Y + error
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ขอสงวนสิทธิ์ในสิ่งที่ปรากฏและมีให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

df <- data.frame(matrix(unlist(X), ncol=length(X)))
data <- data.frame(Y,df)

#Divide train test
a = round(0.7*n)
sample_n <- sample(1:nrow(data), a)
train <- data[sample_n,]
test <- data[-sample_n,]

#train Random Forest
col_name <- colnames(data)[2:ncol(data)]
model_rf = randomForest(Y~., train[, c("Y", col_name)])
#train
train_rf <- predict(model_rf, train)
mse_train[r] <- mean((train$Y-train_rf)**2)
mape_train[r] <- mean(abs((train$Y-train_rf)/train$Y)) * 100
cat("MSE train error :",mse_train[r])
cat("\nMAPE train error :",mape_train[r])
cat("\n#####")
#test
test_rf <- predict(model_rf, test)
mse_test[r] <- mean((test$Y-test_rf)**2)
mape_test[r] <- mean(abs((test$Y-test_rf)/test$Y)) * 100
cat("\nMSE test error :",mse_test[r])
cat("\nMAPE test error :",mape_test[r])
cat("\n#####")
}

cat("\n Mean of mse(train) :",mean(mse_train))
cat("\n Mean of mape(train) :",mean(mape_train))
cat("\n#####")
cat("\n Mean of mse(test) :",mean(mse_test))
cat("\n Mean of mape(test) :",mean(mape_test))

```

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์การเชิงงานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีค้นหาเพื่อนบ้านใกล้สุด k อันดับ

```

library(caret)
rnd = 1000
mse_train <- vector('integer',length=rnd )
mape_train <- vector('integer',length=rnd )
mse_test <- vector('integer',length=rnd )
mape_test <- vector('integer',length=rnd )
for (r in 1:rnd)
{
  set.seed(r)
  cat("\nRound :",r,"\n")
  #set parameter
  n <- 150
  P <- 300
  varr <- 25
  error <- rnorm(n, mean=0, sd=sqrt(varr))
  X = vector(mode='list', length=P)
  Y = 1000
  beta = -1
  for (i in 1:P)
  {
    X[[i]] = rnorm(n, mean=0, sd=1)
    Y = Y + (beta*X[[i]])
  }
  Y = Y + error

  #Build data frame
  df <- data.frame(matrix(unlist(X), ncol=length(X)))
  data <- data.frame(Y,df)

  #Divide train test

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งยังมีเหตุให้แจ้งเตือนและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

train <- data[sample_n,]
test <- data[-sample_n,]
train_x <- subset(train,select = -Y)
train_y <- train$Y
test_x <- subset(test,select = -Y)
test_y <- test$Y

#model KNN Note class
model_knn <- knnreg(train_x, train_y)
train_knn <- predict(model_knn, train_x)
#train
mse_train[r] <- mean((train_y-train_knn)**2)
mape_train[r] <- mean(abs((train_y-train_knn)/train$Y)) * 100
cat("MSE train error :",mse_train[r])
cat("\nMAPE train error :",mape_train[r])
cat("\n#####")

#test
test_knn <- predict(model_knn, test_x)
mse_test[r] <- mean((test_y-test_knn)**2)
mape_test[r] <- mean(abs((test_y-test_knn)/test$Y)) * 100
cat("\nMSE test error :",mse_test[r])
cat("\nMAPE test error :",mape_test[r])
cat("\n#####")
}
cat("\n Mean of mse(train) :",mean(mse_train))
cat("\n Mean of mape(train) :",mean(mape_train))
cat("\n#####")
cat("\n Mean of mse(test) :",mean(mse_test))
cat("\n Mean of mape(test) :",mean(mape_test))

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีซัพพอร์ตเวกเตอร์แมชชีน

```

library(e1071)
rnd = 1000
mse_train <- vector('integer',length=rnd )
mape_train <- vector('integer',length=rnd )
mse_test <- vector('integer',length=rnd )
mape_test <- vector('integer',length=rnd )
for (r in 1:rnd)
{
  set.seed(r)
  cat("\nRound :",r,"\n")
  #set parameter
  n <- 150
  P <- 300
  varr <- 1
  error <- rnorm(n, mean=0, sd=sqrt(varr))
  X = vector(mode='list', length=P)
  Y = 1000
  beta = -1
  for (i in 1:P)
  {
    X[[i]] = rnorm(n, mean=0, sd=1)
    Y = Y + (beta*X[[i]])
  }
  Y = Y + error

  #Build data frame
  df <- data.frame(matrix(unlist(X), ncol=length(X)))
  data <- data.frame(Y,df)

  #Divide train test

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งยังมีเหตุแบบเชิงนโยบายและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

train <- data[sample_n,]
test <- data[-sample_n,]

#train support vector machine
col_name <- colnames(data)[2:ncol(data)]
model_svm = svm(Y~., train[, c("Y", col_name)])
#train
train_svm <- predict(model_svm, train)
mse_train[r] <- mean((train$Y-train_svm)**2)
mape_train[r] <- mean(abs((train$Y-train_svm)/train$Y)) * 100
cat("MSE train error :",mse_train[r])
cat("\nMAPE train error :",mape_train[r])
cat("\n#####")

#test
test_svm <- predict(model_svm, test)
mse_test[r] <- mean((test$Y-test_svm)**2)
mape_test[r] <- mean(abs((test$Y-test_svm)/test$Y)) * 100
cat("\nMSE test error :",mse_test[r])
cat("\nMAPE test error :",mape_test[r])
cat("\n#####")
}
cat("\n Mean of mse(train) :",mean(mse_train))
cat("\n Mean of mape(train) :",mean(mape_train))
cat("\n#####")
cat("\n Mean of mse(test) :",mean(mse_test))
cat("\n Mean of mape(test) :",mean(mape_test))

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีการเรียนรู้เชิงลึก

```

library(deepnet)

rnd = 10

mse_train <- vector('integer',length=rnd )
mape_train <- vector('integer',length=rnd )
mse_test <- vector('integer',length=rnd )
mape_test <- vector('integer',length=rnd )

for (r in 1:rnd)
{
  set.seed(r)
  cat("\nRound :",r,"\n")
  #set parameter
  n <- 20
  P <- 50
  varr <- 25
  error <- rnorm(n, mean=0, sd=sqrt(varr))
  X = vector(mode='list', length=P)
  Y = 1000
  #beta
  beta = -1
  for (i in 1:P)
  {
    X[[i]] = rnorm(n, mean=0, sd=1)
    Y = Y + (beta*X[[i]])
  }
  Y = Y + error

  #Build data frame
  df <- data.frame(matrix(unlist(X), ncol=length(X)))

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

#Divide train test
a = round(0.7*n)
sample_n <- sample(1:nrow(data), a)
train <- data[sample_n,]
test <- data[-sample_n,]
train_x <- subset(train,select = -Y)
train_x <- as.matrix(train_x)
train_y <- train$Y
test_x <- subset(test,select = -Y)
test_y <- test$Y

#model Deep learning
model_dl <- dbn.dnn.train(train_x, train_y, hidden=c(2,2,2), activationfun="tanh",
output= "linear" , numepochs=10000)
#train
train_dl <- nn.predict(model_dl, train_x)
mse_train[r] <- mean((train_y-train_dl)**2)
mape_train[r] <- mean(abs((train_y-train_dl)/train$Y)) * 100
cat("MSE train error :",mse_train[r])
cat("\nMAPE train error :",mape_train[r])
cat("\n#####")
#test
test_dl <- nn.predict(model_dl, test_x)
mse_test[r] <- mean((test_y-test_dl)**2)
mape_test[r] <- mean(abs((test_y-test_dl)/test$Y)) * 100
cat("\nMSE test error :",mse_test[r])
cat("\nMAPE test error :",mape_test[r])
cat("\n#####")
}
cat("\n Mean of mse(train) :",mean(mse_train))
cat("\n Mean of mape(train) :",mean(mape_train))
cat("\n#####")
cat("\n Mean of mse(test) :",mean(mse_test))

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การเขียนเพื่อการศึกษาเท่านั้น เมื่อผู้ยืมได้เห็นว่าไปใช้ประโยชน์ด้านการค้า
 "ไม่ว่ากรณีใดๆ" ฟังชั่น อีกฟังก์ชัน ผมเห็นแต่แบบส่งเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
cat("\n Mean of mape(test) :",mean(mape_test))
```

วิธีสโตแคสติกเกรเดียนเตสเซนท์

```
# gradient descent function takes five arguments
```

```
gradDesc <- function(df, x, y, alpha , max_iter){
```

```
  n <- nrow(df)
```

```
  x0 <- rep(1, nrow(df))
```

```
  x1 <- df[,x]
```

```
  x <- as.matrix(cbind(x0, x1))
```

```
  y <- df[[y]]
```

```
  # initialize random weights
```

```
  b <- runif(ncol(x),0,1) #random 0-1
```

```
  yhat <- x %*% b
```

```
  mse <- (1/n) * sum((y - yhat) ** 2)
```

```
  converged <- F
```

```
  iteration <- 0
```

```
  # update weight using GD algorithm
```

```
  while(converged == F){
```

```
    iteration <- iteration + 1
```

```
    #formula update weight
```

```
    b_new <- b - alpha * (1/n * t(x) %*% (yhat - y))
```

```
    b <- b_new
```

```
    yhat <- x %*% b
```

```
    mse_new <- (1/n) * sum((y - yhat) ** 2)
```

```
    #cat(mse_new,"\n")
```

```
  # if iteration hits max_iter, program ends
```

```
  if(iteration == max_iter){
```

```
    converged <- T
```

```
    return(b)
```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น ยกเว้นให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

    }
  }
}

#set parameter
rnd = 1000
mse_train <- vector('integer',length=rnd )
mape_train <- vector('integer',length=rnd )
mse_test <- vector('integer',length=rnd )
mape_test <- vector('integer',length=rnd )
for (r in 1:rnd)
{
  set.seed(r)
  cat("\nRound :",r,"\n")
  n <- 50
  P <- 300
  varr <- 1
  error <- rnorm(n, mean=0, sd=sqrt(varr))
  X = vector(mode='list', length=P)
  Y = 1000
  beta = -1
  for (i in 1:P)
  {
    X[[i]] = rnorm(n, mean=0, sd=1)
    Y = Y + (beta*X[[i]])
  }
  Y = Y + error

  #Build data frame
  df <- data.frame(matrix(unlist(X), ncol=length(X)))
  data <- data.frame(Y,df)

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น ยกเว้นหากมีให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

a = round(0.7*n)
sample_n <- sample(1:nrow(data), a)
train <- data[sample_n,]
test <- data[-sample_n,]

# test function x
col_name <- colnames(data)[2:ncol(data)]
b <- gradDesc(train, x=c(col_name), y = "Y", alpha = 0.01, max_iter = 1000)

# predict train and mse
predict_train <- rep(b[1],length(train$Y))
i <- 1
for(name in col_name)
{
  predict_train <- predict_train + (train[name]*b[i+1])
  i = i+1
}
mse_train[r] <- mean((train$Y-unlist(predict_train))**2)
mape_train[r] <- mean(abs((train$Y-unlist(predict_train))/train$Y)) * 100
cat("MSE train error :",mse_train[r])
cat("\nMAPE train error :",mape_train[r])
cat("\n#####")

#predict test and mape
predict_test <- rep(b[1],length(test$Y))
i <- 1
for(name in col_name)
{
  predict_test <- predict_test + (test[name]*b[i+1])
  i = i+1
}

mse_test[r] <- mean((test$Y-unlist(predict_test))**2)
mape_test[r] <- mean(abs((test$Y-unlist(predict_test))/test$Y)) * 100

```

```

cat("\nMSE test error :",mse_test[r])
cat("\nMAPE test error :",mape_test[r])
cat("\n#####")
}
cat("\n Mean of mse(train) :",mean(mse_train))
cat("\n Mean of mape(train) :",mean(mape_train))
cat("\n#####")
cat("\n Mean of mse(test) :",mean(mse_test))
cat("\n Mean of mape(test) :",mean(mape_test))

library(ggplot2)

data <- data.frame(
  Model = c("RF", "KNN", "SVM", "DL" ),
  n_20 = c(1.44,1.47,1.43,1.55),
  n_30 = c(1.41,1.45,1.39,1.48),
  n_40 = c(1.41,1.44,1.37,1.31)
)

data_long <- reshape2::melt(data, id.vars = "Model", variable.name = "n", value.name
= "MAPE")

ggplot(data_long, aes(x = as.factor(n), y = MAPE, group = Model, color = Model)) +
  geom_line(size=1.5) +
  geom_point(size=2.5) +
  labs(x = "n", y = "MAPE", title = "p = 300") +
  theme_minimal() +
  scale_x_discrete(labels = c("50", "100", "150")) +
  scale_color_discrete(name = "Model")+
  guides(color = guide_legend(title = NULL))

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



งานทะเบียนคณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
คำรับรองเล่มปัญหาพิเศษ

วันที่ 2 เดือน เมษายน พ.ศ 2567

ข้าพเจ้า นายณัฐพล สุมาลัย รหัสประจำตัว 63050619
นางสาวปานวาด เนียมหอม รหัสประจำตัว 63050644
นางสาวสุทธาสินี ศรีอนุชา รหัสประจำตัว 63050679

นักศึกษาหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชา สถิติประยุกต์ ภาควิชา สถิติ
ขอรับรองว่า ปัญหาพิเศษ เรื่อง

การเปรียบเทียบประสิทธิภาพของวิธีการเรียนรู้ด้วยเครื่อง
จากตัวแบบถดถอยเชิงเส้นเมื่อข้อมูลมีมิติขั้นสูง
A Comparison of Performance of Machine Learning
Methods from Linear Regression Models Using
High-Dimensional Data

ปีการศึกษา 2566

เป็นผลงานวิจัยที่ได้คัดลอกหรือละเมิดลิขสิทธิ์ของผู้อื่นและได้ผ่านการตรวจสอบความซ้ำซ้อน
เรียบร้อยแล้ว และได้แนบเอกสารการตรวจสอบการลอกเลียนงานวรรณกรรมที่ตรวจสอบจากเล่ม
ปัญหาพิเศษฉบับสมบูรณ์แล้ว
โปรแกรมอักขราวิสุทธิ์ 0.75%

ลงชื่อ.....ณัฐพล สุมาลัย..... ลงชื่อ.....ปานวาด เนียมหอม..... ลงชื่อ.....สุทธาสินี ศรีอนุชา.....
(นายณัฐพล สุมาลัย) (นางสาวปานวาด เนียมหอม) (นางสาวสุทธาสินี ศรีอนุชา)
นักศึกษา นักศึกษา นักศึกษา

ข้าพเจ้า รศ.ดร. อัชฌา อระวีพร อาจารย์ที่ปรึกษาปัญหาพิเศษ ได้ตรวจสอบปัญหาพิเศษ ของ
นักศึกษาข้างต้นแล้ว ขอรับรองว่าเป็นผลงานวิจัยของนักศึกษาจริงและมีเนื้อหาสมบูรณ์ จึงลงชื่อไว้
เป็นหลักฐาน

ลงชื่อ.....อัชฌา อระวีพร.....
(รศ.ดร. อัชฌา อระวีพร)

อาจารย์ที่ปรึกษา
เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้