การเปรียบเทียบผลการพยากรณ์ราคาบิทคอยน์โดยใช้วิธีการวิเคราะห์การถดถอย โครงข่ายประสาทเทียม และการโปรแกรมเชิงพันธุกรรม

เอ ดีหลี

งานนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา
มกราคม 2560
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา

COMPARISON OF BITCOIN FORECASTING ALGORITHMS: REGRESSION ANALYSIS, NEURAL NETWORKS AND GENETIC PROGRAMMING

A DEELEE

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT
FOR THE MASTER DEGREE OF SCIENCE IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATICS BURAPHA UNIVERSITY
JANUARY 2017
COPYRIGHT OF BURAPHA UNIVERSITY

คณะกรรมการควบคุมงานนิพนธ์และคณะกรรมการสอบงานนิพนธ์ ได้พิจารณางานนิพนธ์ของ เอ ดีหลี ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมงานนิพนธ์

..... อาจารย์ที่ปรึกษาหลัก
(ผู้ช่วยศาสตราจารย์ ดร.สุนิสา ริมเจริญ)

คณะกรรมการสอบงานนิพนธ์

..... ประธาน
(ดร.ปัทมา เจริญพร)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.สุนิสา ริมเจริญ)

..... กรรมการ
(ดร.คนึงนิจ กุโบลา)

คณะวิทยาการสารสนเทศ อนุมัติให้รับงานนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศของมหาวิทยาลัยบูรพา

..... คณบดีคณะวิทยาการสารสนเทศ
(ผู้ช่วยศาสตราจารย์ ดร.กฤษณะ ชินสาร)
วันที่ 20 เดือน กรกฎาคม พ.ศ. 2560

กิตติกรรมประกาศ

งานนิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยความอนุเคราะห์จาก ผู้ช่วยศาสตราจารย์ ดร.สุนิสา ริมเจริญ อาจารย์ที่ปรึกษา ที่ได้กรุณาให้ข้อเสนอแนะต่าง ๆ อีกทั้งยังคอยช่วยเหลือแนะนำ แก้ไขข้อบกพร่องด้วยความเอาใจใส่ดูแลอย่างดียิ่งมาโดยตลอด ผู้วิจัยขอขอบพระคุณท่านเป็นอย่างมาก

ขอกราบขอบพระคุณมารดา บิดา ที่ได้ให้การสนับสนุนและการศึกษาที่ดีมาโดยตลอด ขอขอบคุณครอบครัว พี่ ๆ เพื่อน ๆน้อง ๆ ที่ร่วมงาน และเพื่อน ๆ สาขาวิชาเทคโนโลยีสารสนเทศ รุ่นที่ 8 คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพาทุกคนที่คอยให้ความช่วยเหลือ ให้คำปรึกษา ให้กำลังใจ และสนับสนุนผู้วิจัยเสมอมา

คุณค่าและประโยชน์ของวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นกตัญญูกตเวทิตาแด่ บุพการี บูรพาจารย์ และผู้มีพระคุณทุกท่านที่ทำให้ผู้วิจัยเป็นผู้มีความรู้ มีอาชีพการงานที่ดีมาจนถึงปัจจุบัน ขอสิ่งศักดิ์สิทธิ์โปรดคลบันดาลให้ทุกท่านมีสุขภาพกาย สุขภาพใจที่แข็งแรงด้วยเทอญ

เอ ดีหลี

55920329: สาขาวิชา: เทคโนโลยีสารสนเทศ; วท.ม. (เทคโนโลยีสารสนเทศ)
คำสำคัญ: การพยากรณ์/ บิทคอยน์/ วิธีวิเคราะห์สหสัมพันธ์ / วิธีคัดเลือกคุณสมบัติ / วิธีวิเคราะห์การถดถอย/ วิธีโครงข่ายประสาทเทียม/ วิธีการโปรแกรมเชิงพันธุกรรม

เอ ดีหลี: การเปรียบเทียบผลการพยากรณ์ราคาบิทคอยน์ โดยใช้วิธีการวิเคราะห์การถดถอย วิธีการโครงข่ายประสาทเทียม และวิธีการโปรแกรมเชิงพันธุกรรม คณะกรรมการควบคุมงานนิพนธ์: สุนิสา ริมเจริญ, Ph.D., 102 หน้า. ปี พ.ศ. 2559.

งานวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์หาปัจจัยพื้นฐาน เพื่อมาทดสอบการพยากรณ์ของราคาบิทคอยน์ โดยใช้วิธีวิเคราะห์การถดถอย วิธีโครงข่ายประสาทเทียม และวิธีการโปรแกรมเชิงพันธุกรรม

ผู้วิจัยได้เลือกข้อมูลปัจจัยที่คาดว่าจะมีผลต่อราคาบิทคอยน์ ประกอบด้วยข้อมูล 2 ส่วน ได้แก่ 1.อัตราแลกเปลี่ยนเงิน 2.ข้อมูลหุ้น 3.ทองคำ 4.พลังงาน ส่วนที่ 2 ได้แก่ ราคาบิทคอยน์ย้อนหลัง 1 วัน และราคาบิทคอยน์ ย้อนหลัง 2 วัน โดยใช้ข้อมูลย้อนหลัง ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึงวันที่ 31 มีนาคม 2559 รวม 640 วัน โดยจะใช้ข้อมูลราคาปิดของแต่ละวัน

จากนั้นนำข้อมูลปัจจัยทั้งหมดมาคัดเลือกข้อมูล เพื่อหาข้อมูลปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีคัดเลือกคุณสมบัติ เมื่อได้ข้อมูลปัจจัยที่ผ่านการคัดเลือกแล้ว นำปัจจัยที่ได้มาทดลองพยากรณ์ราคาบิทคอยน์ ด้วยเทคนิคทั้ง 3 วิธี ด้วยการแบ่งอัตราข้อมูลที่ 80:20 และ 90:10โดยใช้วิธีวัดค่าความคลาดเคลื่อน คือ RMSE, RAE และ MAPE เป็นตัวเปรียบเทียบ

จากการทดลองพบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีวิเคราะห์การถดถอย โดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1831.89 RAE เท่ากับ 1.04 และ MAPE เท่ากับ 9.10 วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 90:10 โมเดลที่ใช้ learningRate 0.3 และ trainingTime 5000 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 และ MAPE เท่ากับ 9.06 และวิธีโปรแกรมพันธุกรรม โดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 80:20 โมเดลที่ใช้มีค่า elitism 200 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.10 RAE เท่ากับ 1.07 และ MAPE เท่ากับ 9.44 ซึ่งวิธีการพยากรณ์ราคาบิทคอยน์ ทั้ง 3 วิธีได้ผลที่ใกล้เคียงกันต่างกันเพียงเล็กน้อย

55920329: MAJOR: INFORMATION TECHNOLOGY; M.Sc,.
KEYWORDS: FORECASTING/ BITCOIN/ CORRELATION/ FEATURE SELECTION/ REGRESSION ANALYSIS/ NEURAL NETWORKS/ GENETIC PROGRAMMING

A DEELEE: COMPARISON OF BITCOIN FORECASTING ALGORITHMS: REGRESSION ANALYSIS, NEURAL NETWORKS AND GENETIC PROGRAMMING. ADVISORY COMMITTEE: SUNISA RIMCHAROEN, Ph.D., 102 p. 2016.

This research aims to determine the underlying factors to forecast of Bitcoin prices by using Regression Analysis, Neural Networks and Genetic Programming.

The factors that affect the fluctuations of the Bitcoin prices are divided into two groups. The first group is underlying information that involves 1) currency exchange rates 2) Stock exchange 3) Gold prices 4) Oil prices, the second factor is Bitcoin prices of yesterday and Bitcoin prices of two last days ago. The data used in the experiments are 21 months starting from July 1, 2557 to March 31, 2559, totally 640 days. Noted that we use only the closing price in each day.

The factors that are associated with Bitcoin prices are selected using correlation analysis and feature selection method. The selected are used as input to forecast the price of Bitcoin with 3 methods: Regression analysis, Neural networks and Genetic programming. The training and testing data are divided with the ratio 80:20 and 90:10. The forecasting results are measured in terms of Root Mean Square Error (RMSE), Relative Absolute Error (RAE) and Mean Absolute Percent Error (MAPE).

The results show that the forecasting of Bitcoin prices by using Regression analysis yields the RMSE of 1831.89, RAE of 1.04 and MAPE of 9.10. By using Neural networks with the ratio of training and testing 90:10, learningRate 0.3 and trainingTime 5000, they yield RMSE of 1750.96, MAPE of 9.06 and RAE of 1.03. Genetic programming with the ratio of training and testing of 80:20, elitism 200, it yields RMSE of 1909.10, the RAE of 1.07 and the MAPE of 9.44. To summarize, the forecasting of Bitcoin prices using the three methods yield quite similar results, they have a bit slightly difference of error rates.

สารบัญ

หน้า

บทคัดย่อภาษาไทย…….. จ

บทคัดย่อภาษาอังกฤษ…….. ฉ

สารบัญ……..…….. ช

สารบัญตาราง.....…….. ฌ

สารบัญภาพ.....…….. ฐ

บทที่

1 บทนำ…..……..... 1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์ของการวิจัย.....2

1.3 ขอบเขตของงานวิจัย.....2

1.4 เครื่องมือที่ใช้ในการทดลอง.....3

1.5 ประโยชน์ที่คาดว่าจะได้รับ.....3

1.6 ขั้นตอนวิธีการดำเนินงาน..... 3

2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 4

2.1 ความรู้เกี่ยวกับ บิทคอยน์ (Bitcoin).....5

2.2 ข้อมูลที่วิเคราะห์และวิธีการคัดเลือกข้อมูล..... 6

2.2.1 การวิเคราะห์สหสัมพันธ์ (Correlation)……..... 7

2.2.2 การคัดเลือกคุณสมบัติ (Feature Selection)…….. 8

2.3 ความรู้เกี่ยวกับการพยากรณ์และเทคนิคที่ใช้ในการพยากรณ์..... 13

2.3.1 การวิเคราะห์การถดถอย (Regression Analysis)……..... 14

2.3.2 โครงข่ายประสาทเทียม (Neural Networks)……..…. 17

2.3.3 การโปรแกรมเชิงพันธุกรรม (Genetic Programming)……... 22

2.4 ความรู้เกี่ยวกับค่าความคลาดเคลื่อนของค่าพยากรณ์..... 25

2.4.1 Root Mean Squared Error (RMSE)……... 25

2.4.2 Relative Absolute Error (RAE)……..……... 25

2.4.3 Mean Absolute Percentage Error (MAPE)……... 25

2.5 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง..... 26

สารบัญ (ต่อ)

บทที่ หน้า

3 วิธีดำเนินงานวิจัย..... 30

3.1 การเตรียมข้อมูลที่จะนำไปใช้ในการทดลอง..... 31

3.2 การหาความสัมพันธ์ของข้อมูลด้วยการวิเคราะห์สหสัมพันธ์ (Correlation)…….. 32

3.3 การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติหรือคุณลักษณะ (Feature Selection)..... 36

3.4 การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis)..... 40

3.5 การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks)..... 42

3.6 การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming)..... 49

3.7 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์..... 53

3.7.1 การหาค่าความคลาดเคลื่อนด้วย Root Mean Square Error (RMSE)……. 53

3.7.2 การหาค่าความคลาดเคลื่อนด้วย Relative Absolute Error (RAE)……. 53

3.7.3 การหาค่าความคลาดเคลื่อนด้วย Mean Absolute Percent Error (MAPE)…. 53

4 ผลการดำเนินงาน..... 55

4.1 การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis)..... 55

4.2 การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks)..... 58

4.3 การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming)..... 71

4.4 สรุปผลการทดลอง..... 87

5 สรุปผลและข้อเสนอแนะ..... 90

5.1 สรุปผล..... 90

5.2 ข้อเสนอแนะ..... 91

บรรณานุกรม 92

ภาคผนวก 95

ประวัติย่อผู้วิจัย..... 102

สารบัญตาราง

ตารางที่		หน้า
2-1	ลำดับการค้นหาแบบลึกก่อน.....	12
2-2	ลำดับการค้นหาแบบกว้างก่อน.....	14
3-1	ตัวอย่างข้อมูลปัจจัยที่ใช้ในการพยากรณ์ราคาบิทคอยน์.....	36
3-2	ชื่อข้อมูลปัจจัยที่ใช้ในการพยากรณ์ราคาบิทคอยน์.....	37
3-3	ตัวอย่างข้อมูลตามปัจจัยที่ได้จากวิธีการวิเคราะห์สหสัมพันธ์.....	44
3-4	ตัวอย่างข้อมูลตามปัจจัยที่ได้จากวิธีการคัดเลือกคุณสมบัติ.....	44
3-5	ค่าพารามิเตอร์ที่ใช้ในการทดลองสำหรับ Weka.....	49
3-6	ค่าพารามิเตอร์ที่ใช้ในการทดลองสำหรับ GPdotNET.....	57
4-1	ผลลัพธ์ที่ได้จากการทดลอง (Train) ของวิธีการวิเคราะห์สหสัมพันธ์.....	61
4-2	แสดงผลลัพธ์ที่ได้จากการทดลอง (Train) ของวิธีการคัดเลือกคุณสมบัติ.....	61
4-3	ผลลัพธ์ที่ได้จากวิธีการวิเคราะห์การถดถอย.....	62
4-4	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.3.....	64
4-5	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.2.....	65
4-6	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.1.....	66
4-7	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20.....	66
4-8	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.3.....	67
4-9	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.2.....	68
4-10	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.1.....	68
4-11	สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20.....	69

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-12	ผลการทดลองของข้อมูลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมในอัตราส่วน 80 ต่อ 20	70
4-13	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.3	70
4-14	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.2	71
4-15	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.1	71
4-16	สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10	72
4-17	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.3	73
4-18	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.2	74
4-19	ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.1	74
4-20	สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10	75
4-21	ผลการทดลองของข้อมูลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมในอัตราส่วน 90 ต่อ 10	76
4-22	สรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียม	76
4-23	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 1	78
4-24	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 100	78
4-25	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 200	79

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-26	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 300.....	80
4-27	สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20.....	80
4-28	ผลผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 1….....	81
4-29	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 100.....	82
4-30	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 200.....	82
4-31	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 300.....	83
4-32	สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20.....	84
4-33	สรุปผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 80 ต่อ 20.....	84
4-34	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 1.....	85
4-35	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 100.....	86
4-36	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 200.....	86
4-37	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 300.....	87
4-38	สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10.....	88

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
4-39	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 1	88
4-40	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 100	89
4-41	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 200	90
4-42	ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 300	91
4-43	สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10	91
4-44	สรุปผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 90 ต่อ 10	92
4-45	สรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรม	92
4-46	สรุปผลการพยากรณ์ราคาบิทคอยน์	93

สารบัญภาพ

ภาพที่		หน้า
1-1	กราฟราคาปิดของ บิทคอยน์ ในแต่ละวัน	2
2-1	วิธีการได้มาของบิทคอยน์	6
2-2	ลำดับการเดินทางบนโหนดของการค้นหาแบบลึกก่อน	11
2-3	โครงสร้างข้อมูลแบบกราฟ	12
2-4	ลำดับการค้นหาแบบกว้างก่อนบนโครงสร้างต้นไม้	13
2-5	ขั้นตอนของการค้นหาแบบดีที่สุดก่อน	15
2-6	กรอบแนวคิดการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ……	17
2-7	โครงสร้างระบบประสาท	21
2-8	โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว	23
2-9	โครงข่ายประสาทเทียมแบบหลายชั้น	23
2-10	ขั้นตอนการทำงานของการโปรแกรมเชิงพันธุกรรม	26
2-11	ผลที่ได้จากการสร้างประชากรโดย (ก) ระเบียบวิธีต้นไม้แบบไม่เต็มต้น (ข) ระเบียบวิธีต้นไม้แบบเต็มต้น	27
2-12	การข้ามสายพันธุ์	28
2-13	การกลายพันธุ์	28
3-1	ข้อมูลในรูปแบบไฟล์ .csv ที่เตรียมไว้	38
3-2	เครื่องมือชื่อ Data Analysis	38
3-3	เครื่องมือชื่อ Correlation	39
3-4	ภายในเครื่องมือชื่อ Correlation	39
3-5	ผลลัพธ์ที่ได้จากการหาค่าการวิเคราะห์สหสัมพันธ์	40
3-6	การเลือกไฟล์ csv ที่เตรียมไว้	41
3-7	ไฟล์ .arff ที่ปรับให้อยู่ในรูปแบบโปรแกรม Weka แล้ว	41
3-8	หน้าต่างเลือกอัลกอริทึมในการหาความสัมพันธ์	42
3-9	หน้าต่างแสดงผลลัพธ์	42
3-10	เครื่องมือชื่อ Regression	45
3-11	ภายในเครื่องมือชื่อ Regression	45
3-12	ผลลัพธ์จากไฟล์ทดลองที่ได้จากการหาค่าการวิเคราะห์การถดถอย	46

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
3-13	ผลลัพธ์จากไฟล์ทดสอบที่ได้จากการหาค่าการวิเคราะห์การถดถอย….....	47
3-14	การเลือกไฟล์ .csv ที่เตรียมไว้.....	48
3-15	หน้าต่างเลือกอัลกอริทึมในการทดลองพยากรณ์.....	49
3-16	การกำหนดค่าพารามิเตอร์.....	50
3-17	ผลลัพธ์ที่ได้จากการทดลอง.....	50
3-18	หน้าการ Save model.....	51
3-19	หน้าการ Load model.....	51
3-20	หน้าการเลือกไฟล์ที่เตรียมไว้ใช้ในการทดสอบ.....	52
3-21	เลือกเมนู Re-evaluate model on current test set เพื่อใช้โมเดลในการทำนายข้อมูล.....	52
3-22	เลือกเมนู Visualize classifier errors.....	53
3-23	หน้าต่าง Weka Classifier Visualize.....	53
3-24	ข้อมูลสำหรับโปรแกรม GPdotNET.....	54
3-25	หน้าต่างเริ่มโปรแกรม GPdotNET.....	55
3-26	หน้านำข้อมูลเข้าสู่โปรแกรม.....	55
3-27	หน้าต่างเมื่อทำการเลือกข้อมูลเสร็จแล้ว.....	56
3-28	หน้าตั้งค่าพารามิเตอร์สำหรับโปรแกรม GPdotNET.....	56
3-29	หน้าผลลัพธ์ของ GPdotNET.....	57
4-1	โปรแกรม Weka.....	63
4-2	ข้อมูลที่ได้จากไฟล์ .arff.....	64
4-3	ข้อมูลที่ได้จาก GPdotNET.....	77
4-4	กราฟสรุปผลการพยากรณ์ราคาบิทคอยน์.....	95

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

บิทคอยน์ (Bitcoin) เป็นสกุลเงินดิจิตอล ที่ถูกคิดค้นขึ้นมา เพื่อใช้ในการจ่ายเงินแลกเปลี่ยนสินค้าหรือบริการชนิดใหม่ แทนที่เงินตราสกุลดั้งเดิมต่างๆที่เราเคยใช้กันมาเป็นเวลานานแล้ว เช่น USD, JPY, EUR, หรือ THB ซึ่งปกติแล้ว หน้าที่หลักของ “เงิน” คือการเป็นสื่อกลางสำหรับแลกเปลี่ยนสินค้าหรือบริการ ซึ่งเงินตราในรูปแบบใหม่ในสกุล บิทคอยน์ นี้ ก็ทำหน้าที่เช่นเดียวกันกับสิ่งที่เงินตราสกุลต่างๆ เคยทำหน้าที่มา เพียงแต่จะมีความแตกต่างบางอย่างที่แตกต่างไปจากระบบเดิมอย่างชัดเจน ซึ่ง บิทคอยน์ ได้ถูกออกแบบมา เพื่อที่จะให้เป็นระบบการเงินในโลกออนไลน์ โดยที่จะต้องมีคุณสมบัติเช่นเดียวกับระบบเงินตราในรูปแบบดั้งเดิม แต่อีกหนึ่งคุณสมบัติที่ทำให้ บิทคอยน์ ต่างไปจาก ระบบเงินตราอื่น ๆ คือ บิทคอยน์ เป็นระบบที่ไม่ได้อยู่ภายใต้ความควบคุมดูแลโดยธนาคารและรัฐบาลใด ๆ แต่อยู่ได้ด้วยข้อตกลงของผู้ใช้และความซับซ้อนของคณิตศาสตร์เพียงเท่านั้น ซึ่งจากคุณสมบัติของ บิทคอยน์ ที่ไม่ได้บังคับตัวเองให้ต้องอยู่ภายใต้การควบคุมขององค์กรธนาคารขนาดใหญ่หรือรัฐ ทำให้ค่าใช้จ่ายในการบริหารปริมาณเงินนี้รวมไปถึงค่าใช้จ่ายในการทำธุรกรรมลดน้อยลงจากการใช้เงินตราสกุลดั้งเดิมเป็นอย่างมาก

ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาราคาซื้อขายของบิทคอยน์ที่มีอัตราขึ้นลงสูงมากในแต่ละวันดังตัวอย่างในภาพที่ 1-1 ซึ่งผู้วิจัยจะศึกษาและวิเคราะห์ข้อมูลจากองค์ประกอบหลาย ๆ ด้าน เช่น อัตราแลกเปลี่ยนเงินตรา ราคาน้ำมัน หุ้น ซึ่งเป็นปัจจัยพื้นฐานนำมาวิเคราะห์ว่ามีผลต่อการขึ้นลงของราคาบิทคอยน์ เพื่อที่จะนำไปพยากรณ์ราคาบิทคอยน์ ซึ่งจะใช้เครื่องมือ หลาย ๆ ตัวมาใช้เพื่อให้ได้ผลลัพธ์ที่แม่นยำขึ้น

ภาพที่ 1-1 กราฟราคาปิดของบิทคอยน์ ในแต่ละวัน (ที่มา © bitcoincharts.com)

ในช่วงหลายปีที่ผ่านมานี้งานวิจัยทางด้านการพยากรณ์ได้รับความสนใจมากขึ้น และมีการศึกษากันอย่างแพร่หลายอีกทั้งยังมีการพัฒนาวิธีที่มีความซับซ้อนมากขึ้นตามลำดับเรื่อย ๆ ซึ่งการทดลองในครั้งนี้ทางผู้วิจัยได้เลือกใช้วิธีการพยากรณ์ทั้งหมด 3 วิธี คือ 1. วิธีการวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีทางคณิตศาสตร์ที่ใช้หาความสัมพันธ์ระหว่างตัวแปรเชิงปริมาณตั้งแต่สองตัวขึ้นไปเพื่อใช้ในการพยากรณ์ค่าของตัวแปรหนึ่งจากตัวแปรอื่น ๆ ด้วยโปรแกรม Microsoft Excel 2. โครงข่ายประสาทเทียม (Artificial Neural Networks) หรือที่มักจะเรียกสั้น ๆ ว่า โครงข่ายประสาท (Neural Networks หรือ Neural Net) เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล (Data Mining) ด้วยโปรแกรม WEKA และ 3. การโปรแกรมเชิงพันธุกรรม (Genetic Programming) เป็นการสร้างโปรแกรมคอมพิวเตอร์ให้ทำงานตามคำสั่งโดยอัตโนมัติด้วยโปรแกรม GPdotNET แล้วนำผลการทดลองของทั้ง 3 วิธีมาเปรียบเทียบว่าวิธีไหนได้ผลที่มีความแม่นยำที่สุด เพื่อเป็นแนวทางสำหรับนักลงทุนและผู้ที่สนใจเรื่องการพยากรณ์ราคาบิทคอยน์ต่อไป

1.2 วัตถุประสงค์ของการวิจัย

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อวิเคราะห์หาปัจจัยพื้นฐานเพื่อมาทดสอบการพยากรณ์ของราคาบิทคอยน์ ให้ใกล้เคียงและแม่นยำที่สุดด้วยวิธีวิเคราะห์การถดถอย วิธีโครงข่ายประสาทเทียม และวิธีการโปรแกรมเชิงพันธุกรรม

1.3 ขอบเขตของงานวิจัย

1. ศึกษาข้อมูลอัตราการซื้อขายราคาบิทคอยน์ โดยใช้ข้อมูล ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึง วันที่ 31 มีนาคม 2559 จำนวน 21 เดือน

2. ศึกษาข้อมูลปัจจัยพื้นฐานที่คาดว่ามีผลต่อการขึ้นลงของราคาบิทคอยน์ โดยเลือกปัจจัยดังนี้ อัตราแลกเปลี่ยนเงินตรา ราคาน้ำมัน หุ้น และราคาทองแท่ง ทองรูปพรรณ ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึงวันที่ 31 มีนาคม 2559 จำนวน 21 เดือน

3. ใช้ราคาปิดในแต่ละวันมาทำการทดลอง

4. ใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation) และวิธีการคัดเลือกคุณสมบัติ (Feature Selection) เพื่อคัดกรองข้อมูลและหาความสัมพันธ์ของปัจจัยพื้นฐานที่จะนำไปใช้ในการทดลอง

5. ใช้วิธีการวิเคราะห์การถดถอย ด้วยโปรแกรม Microsoft Excel วิธีการโครงข่ายประสาทเทียมด้วยโปรแกรม WEKA และวิธีการโปรแกรมเชิงพันธุกรรมด้วยโปรแกรม GPdotNET

6. หาค่าความคลาดเคลื่อนของค่าพยากรณ์ ด้วย Root Mean Square Error (RMSE), Relative Absolute Error (RAE) และ Mean Absolute Percent Error (MAPE)

1.4 เครื่องมือที่ใช้ในการทดลอง

1. ระบบปฏิบัติการ Windows 8.0
2. โปรแกรม Microsoft Excel 2013
3. โปรแกรม Weka
4. โปรแกรม GPdotNET

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำวิธีการวิเคราะห์การถดถอยด้วย โปรแกรม Microsoft Excel และโครงข่ายประสาทเทียมด้วย โปรแกรม WEKA และวิธีการวิเคราะห์การโปรแกรมเชิงพันธุกรรม ด้วย โปรแกรม GPdotNET มาพยากรณ์ราคาบิทคอยน์ ได้

2. สามารถนำวิธีการวิจัยไปประยุกต์ใช้ในการพยากรณ์ข้อมูลอย่างอื่นได้

1.6 ขั้นตอนวิธีการดำเนินงาน

1. ปรึกษาเรื่องหัวข้อในการทำงานวิจัยกับอาจารย์ที่ปรึกษา
2. รวบรวมข้อมูลที่เกี่ยวข้อง
3. ดำเนินการศึกษาข้อมูลเพื่อวางแผนการทำงานและกำหนดขอบเขตของงานวิจัย
4. ศึกษาค้นคว้าทฤษฎีที่เกี่ยวข้องรวมถึงวิธีการต่างๆ ที่จำเป็นต้องใช้ในการทำงานวิจัย
5. เตรียมข้อมูลปัจจัยพื้นฐานที่จะใช้ในการทดลอง
6. เก็บข้อมูลจากการทดลองพยากรณ์
7. ทดสอบข้อมูลที่ได้จากการพยาการณ์
8. สรุปผลการดำเนินงาน

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยการพยากรณ์ราคาบิทคอยน์ (Bitcoin) โดยใช้ข้อมูลปัจจัยพื้นฐานต่างๆ ผ่านวิธีการคัดเลือกปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ เพื่อนำมาวิเคราะห์ให้ได้ค่าพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพนั้น ผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางการวิจัย ซึ่งมีสาระสำคัญสรุปได้ดังนี้

2.1 ความรู้เกี่ยวกับ บิทคอยน์ (Bitcoin)

2.2 ข้อมูลที่วิเคราะห์และวิธีการคัดเลือกข้อมูล

- 2.2.1 การวิเคราะห์สหสัมพันธ์ (Correlation)
- 2.2.2 การคัดเลือกคุณสมบัติ (Feature Selection)

2.3 ความรู้เกี่ยวกับการพยากรณ์และเทคนิคที่ใช้ในการพยากรณ์

- 2.3.1 การวิเคราะห์การถดถอย (Regression Analysis)
- 2.3.2 โครงข่ายประสาทเทียม (Neural Networks)
- 2.3.3 การโปรแกรมเชิงพันธุกรรม (Genetic Programming)

2.4 ความรู้เกี่ยวกับค่าความคลาดเคลื่อนของค่าพยากรณ์

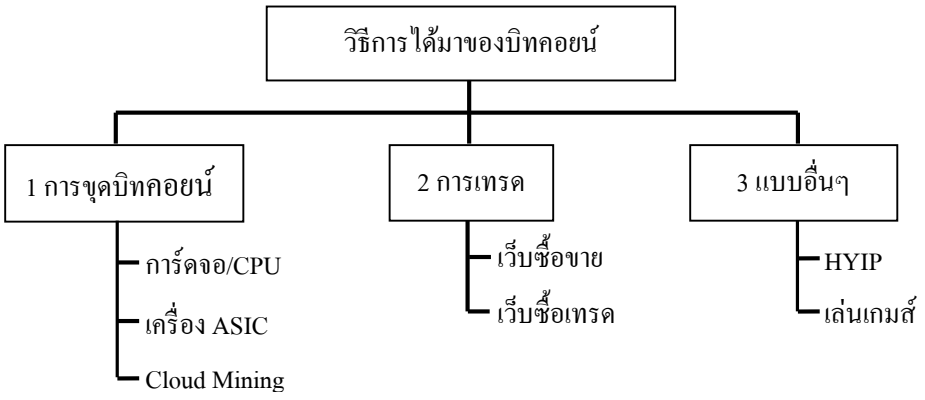
- 2.4.1 Root Mean Square Error (RMSE)
- 2.4.2 Relative Absolute Error (RAE)
- 2.4.3 Mean Absolute Percent Error (MAPE)

2.5 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความรู้เกี่ยวกับ บิทคอยน์ (Bitcoin)

บิทคอยน์ คือ เงินดิจิตอลสกุลหนึ่งที่ใช้ซื้อสินค้าและบริการบนโลกออนไลน์ ซึ่งเป็นสื่อกลางในการแลกเปลี่ยนระหว่างผู้ใช้งานทั้งสองฝ่ายโดยไม่มีการควบคุมและกำหนดมูลค่าจากหน่วยงานใด บิทคอยน์ถูกคิดค้นขึ้นมาตั้งแต่ปี 2009 ภายใต้แนวคิดระบบการเงินแบบไร้ศูนย์กลางที่ไม่ต้องถูกควบคุมและกำหนดมูลค่าจากหน่วยงานใดๆ ทั้งจากเจ้าของเว็บไซต์ สถาบันการเงินหรือหน่วยงานกลางของรัฐเหมือนกับเงินสกุลอื่น ทำให้ค่าเงินของบิทคอยน์จะไม่ถูกแทรกแซงจากการดำเนินนโยบายการเงิน ด้วยคุณสมบัติดังกล่าวบิทคอยน์จึงเริ่มได้รับความนิยมมากขึ้น

ปัจจุบันสามารถแลกเปลี่ยนบิทคอยน์กับเงินสกุลต่างๆ ได้ทั่วโลกผ่านตัวกลางในการแลกเปลี่ยนบนอินเทอร์เน็ต การใช้งานในระบบบิทคอยน์ผู้ใช้งานจะต้องลงโปรแกรม Bitcoin Wallet ในคอมพิวเตอร์เพื่อเป็นกระเป๋าเงินออนไลน์ และสร้างเลขที่บัญชี (address) เพื่อใช้งาน เมื่อมีการทำธุรกรรมต่างๆ ระหว่างผู้ใช้งานทั้งสองฝั่งจะเป็นการแลกเปลี่ยนบิทคอยน์กันโดยตรงผ่านเครื่องคอมพิวเตอร์ทั้งสองโดยไม่ต้องผ่านตัวกลางใดๆ เพียงแต่จะมีการแจ้งสถานะ การทำธุรกรรมให้เครื่องอื่นๆ ที่อยู่ในระบบได้รับทราบเพื่อเป็นการยืนยันความถูกต้องเท่านั้น สำหรับผู้ใช้งานทั่วไปสามารถจัดหาเงินบิทคอยน์ได้จากการแลกเปลี่ยนด้วยเงินสกุลต่างๆ มากกว่า 20 สกุลเงิน เช่น USD, EUR, CNY, JPY, HKD รวมถึง THB โดยผ่านตัวกลางในการแลกเปลี่ยน (Trader) บนอินเทอร์เน็ต ได้แก่ Bitstamp, btc-e.com และLocalBitcoins.com ซึ่งในประเทศไทยก็มีเว็บที่ใช้แลกเปลี่ยนที่น่าเชื่อถือได้แก่ www.exchangercoin.com, www.bitcoinaec.com, www.bx.in.th, www.coins.co.th, และ bitcoin.co.th ซึ่งการได้มาของบิทคอยน์นั้นสามารถแบ่งออกได้เป็น 3 แบบ ดังนี้คือ 1.ได้มาโดยการขุด 2.ได้มาโดยการเทรด 3.ได้มาโดยรูปแบบอื่นๆ เช่น เล่นเกมส์ใช้งานเว็บไซต์ต่างๆ หรือลงทุนเป็นต้น ซึ่งทั้ง 3 วิธีนั้นมีความคุ้มค่าแตกต่างกันออกไป ดังภาพที่ 2-1



ภาพที่ 2-1 วิธีการได้มาของบิทคอยน์

ในงานวิจัยนี้จะกล่าวในเรื่องของการเทรดและการแลกเปลี่ยนเพียงอย่างเดียว คือการที่เราเอาเงินบาทไปซื้อเหรียญบิทคอยน์มาโดยตรงและเก็บเข้ากระเป๋าของเรา เมื่อช่วงใดก็ตามที่ราคาของบิทคอยน์มีราคาสูงขึ้น ก็สามารถขายออกมาทำกำไรได้ ซึ่งต้องสมัครและลงทะเบียนตามเว็บต่างๆ ที่รับแลกเปลี่ยนบิทคอยน์

บิทคอยน์เริ่มได้รับความนิยมจากจุดเด่นด้านการเข้าถึงง่าย สะดวกต่อการใช้งาน มีต้นทุนในการทำธุรกรรมที่ต่ำ และไม่มีการควบคุมใดๆ แต่บิทคอยน์ก็มีปัญหาในการใช้งานจากความผันผวนของค่าเงิน และการนำไปใช้ซื้อสินค้าที่ผิดกฎหมาย บิทคอยน์ได้รับความนิยมจากจุดเด่นที่สำคัญ คือ 1) การเป็นสกุลเงินดิจิตอลที่เข้าถึงได้ง่ายและสะดวกต่อการใช้งานบนโลกออนไลน์ 2) มีต้นทุนในการทำธุรกรรมที่ต่ำ ไม่ต้องใช้เอกสารที่ยุ่งยาก ทำให้ขีดจำกัดในการโอนเงินระหว่างประเทศลดน้อยลงไป ช่วยให้การค้าระหว่างประเทศเกิดการขยายตัวเพิ่มขึ้น 3) การเป็นสกุลเงินที่ไม่มีการควบคุมใดๆ ทำให้เหมาะกับการทำธุรกรรมในประเทศที่กฎหมายการเงินค่อนข้างปิดกั้น หรือประเทศในกลุ่มที่โดนคว่ำบาตรในระดับนานาชาติ แต่บิทคอยน์ก็มีปัญหาในการใช้งานจากความผันผวนของค่าเงินที่มีสูงทำให้ผู้ขายสินค้าและบริการจะมีความเสี่ยงสูงในด้านอัตราแลกเปลี่ยนเมื่อแปลงค่ากลับมาอยู่ในรูปของเงินสกุลจริง ประกอบกับคุณสมบัติของระบบบิทคอยน์ที่ผู้ใช้งานไม่จำเป็นต้องเปิดเผยตัวตนในการทำธุรกรรมทางการเงิน จึงกลายเป็นแรงดึงดูดให้บิทคอยน์ได้รับความนิยมมากในกลุ่มผู้ที่ต้องการซื้อสินค้าผิดกฎหมาย เช่น การสั่งซื้อยาเสพติดประเภทต่างๆ รวมถึงการเล่นพนันออนไลน์ การเป็นแหล่งเลี่ยงภาษี และการฟอกเงิน เป็นต้น

หลายประเทศเริ่มให้ความสนใจศึกษากลไกบิทคอยน์เพื่อออกแบบกฎระเบียบควบคุมให้เหมาะสม อส่วนประเทศไทยได้ประกาศห้ามทำธุรกรรมเกี่ยวกับบิทคอยน์ ตามมาตรา 9 ของ พ.ร.บ.เงินตรา กำหนดไว้ว่า “ห้ามมิให้ผู้ใดทำ จำหน่าย ใช้ หรือนำออกใช้ ซึ่งวัตถุหรือเครื่องหมายใดๆ แทนเงินตรา เว้นแต่จะได้รับอนุญาตจากรัฐมนตรี” โดยธนาคารแห่งประเทศไทยให้เหตุผล 3 ข้อหลัก คือ 1) กฎหมายปัจจุบันยังไม่มีการรองรับเงินดิจิตอลที่มีลักษณะการทำงานอย่างบิทคอยน์ 2) ไทยยังไม่มีมาตรการควบคุมการเคลื่อนย้ายเงินทุนที่ครอบคลุมเพียงพอ 3) ความกังวลเรื่องการใช้เงินสกุลบิทคอยน์ในการซื้อสินค้าและบริการที่ผิดกฎหมาย รวมถึงการใช้เป็นเครื่องมือในการเก็งกำไรค่าเงินบาท

2.2 ข้อมูลที่วิเคราะห์และวิธีการคัดเลือกข้อมูล

ข้อมูลที่นำมาวิเคราะห์ในงานนิพนธ์นี้จะเป็นข้อมูลที่มาจาก 2 ส่วนคือ

ส่วนที่ 1 ได้แก่ ข้อมูลอัตราการซื้อขายบิทคอยน์ โดยใช้ข้อมูลอัตราการซื้อขายบิทคอยน์ ณ วันที่ปัจจุบัน, ข้อมูลอัตราการซื้อขายบิทคอยน์ย้อนหลัง 1 วัน และข้อมูลอัตราการซื้อขายบิท

คอยน์ย้อนหลัง 2 วัน โดยให้ชื่อแทนปัจจัยดังกล่าว คือ BTC, BTC1 และ BTC2 ดังตารางที่ 3-2 ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึง วันที่ 31 มีนาคม 2559 จำนวน 640 วัน 21 เดือน

ส่วนที่ 2 ได้แก่ ข้อมูลปัจจัยพื้นฐานที่ผู้วิจัยคาดว่ามีความสัมพันธ์และมีผลต่อการขึ้นลงของราคาบิทคอยน์ โดยเลือกปัจจัยดังนี้ อัตราแลกเปลี่ยนเงินตรา ราคาน้ำมัน หุ้น และราคาทองแท่งทองรูปพรรณ ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึงวันที่ 31 มีนาคม 2559 จำนวน 640 วัน 21 เดือน ซึ่งจะนำมารวมกับข้อมูลในส่วนที่ 1 เพื่อพยากรณ์ราคาบิทคอยน์ คำอธิบายตัวย่อในแต่ละคอลัมน์แสดงในตารางที่ 3-2 เมื่อได้ข้อมูลทั้งหมดมาแล้วจะนำมาคัดเลือกข้อมูลที่มีความสัมพันธ์กับราคาบิทคอยน์ ด้วยวิธีการดังต่อไปนี้

2.2.1 การวิเคราะห์สหสัมพันธ์ (Correlation)

การวิเคราะห์สหสัมพันธ์ เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป (หรือข้อมูล 2 ชุดขึ้นไป) ตัวอย่างการศึกษาความสัมพันธ์ เช่น การหาความสัมพันธ์ระหว่างอายุและความดันโลหิต ความสัมพันธ์ระหว่างส่วนสูงกับน้ำหนัก ความสัมพันธ์ระหว่างระดับการศึกษากับพฤติกรรมการดูแลตนเอง ความสัมพันธ์ระหว่างพฤติกรรมของเด็กกับวิธีการอบรมเลี้ยงดูเด็ก เป็นต้น ในการพิจารณาความสัมพันธ์ระหว่างตัวแปรว่ามีมากน้อยเพียงใดนั้น จะใช้ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) เป็นค่าที่วัดความสัมพันธ์ ซึ่งโดยวิธีการทางสถิติมีอยู่หลายวิธี การใช้สถิติตัวใดขึ้นอยู่กับลักษณะของตัวแปรหรือระดับของการวัดในตัวแปรนั้นๆ ในการวัดความสัมพันธ์แต่ละแบบจะต้องมีการทดสอบนัยสำคัญก่อน จึงจะสรุปได้ว่าตัวแปรคู่ใดมีความสัมพันธ์กันจริงหรือไม่ มากน้อยเพียงใด สำหรับการแปรผลจะมองในแง่ของความเกี่ยวพัน ความสอดคล้อง การแปรผันร่วมกัน หรือไม่สัมพันธ์กัน แต่ไม่ได้หมายความว่าตัวแปรหนึ่งเป็นเหตุและอีกตัวแปรเป็นผล (หรือไม่สามารถระบุได้ว่าตัวแปรไหนเป็นตัวแปรต้นหรือตัวแปรตาม) เช่น ศึกษาความสัมพันธ์ระหว่างส่วนสูงกับน้ำหนัก เราไม่สามารถบอกได้ว่าส่วนสูงหรือน้ำหนักตัวใดเป็นเหตุและตัวใดเป็นผล บอกได้เพียงว่ามีความสัมพันธ์กันหรือไม่ และมีขนาดของความสัมพันธ์กันมากน้อยเพียงใด

ค่าสัมประสิทธิ์สหสัมพันธ์ จะใช้สัญลักษณ์ r แทนสัมประสิทธิ์สหสัมพันธ์ของกลุ่มตัวอย่าง และ ρ แทนสัมประสิทธิ์สหสัมพันธ์ของประชากร ค่าสัมประสิทธิ์สหสัมพันธ์ที่ใช้วัดขนาดของความสัมพันธ์กันระหว่างตัวแปรมี 2 ลักษณะ คือ $-1 \leq r \leq 1$ และ $0 \leq r \leq 1$

การบอกระดับหรือขนาดของความสัมพันธ์ จะใช้ตัวเลขของค่าสัมประสิทธิ์สหสัมพันธ์ หากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเข้าใกล้ -1 หรือ 1 แสดงถึงการมีความสัมพันธ์กันในระดับสูง แต่หากมีค่าเข้าใกล้ 0 แสดงถึงการมีความสัมพันธ์กันในระดับน้อย หรือไม่มีเลย สำหรับการพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ โดยทั่วไปอาจใช้เกณฑ์ดังนี้ (Hinkle D.E. 1998, p.118)

ค่า r	ระดับของความสัมพันธ์
.90 – 1.00	มีความสัมพันธ์กันสูงมาก
.70 – .90	มีความสัมพันธ์กันในระดับสูง
.50 – .70	มีความสัมพันธ์กันในระดับปานกลาง
.30 – .50	มีความสัมพันธ์กันในระดับต่ำ
.00 – .30	มีความสัมพันธ์กันในระดับต่ำมาก

เครื่องหมาย +, - หน้าตัวเลขสัมประสิทธิ์สหสัมพันธ์ จะบอกถึงทิศทางของความสัมพันธ์โดยที่หาก

r มีเครื่องหมาย + หมายถึง การมีความสัมพันธ์กันไปในทิศทางเดียวกัน (ตัวแปรหนึ่งมีค่าสูง อีกตัวหนึ่งจะมีค่าสูงไปด้วย)

r มีเครื่องหมาย - หมายถึง การมีความสัมพันธ์กันไปในทิศทางตรงกันข้าม (ตัวแปรหนึ่งมีค่าสูง อีกตัวหนึ่งจะมีค่าต่ำ)

ยกเว้นค่าสัมประสิทธิ์สหสัมพันธ์บางชนิดที่มีลักษณะ $0 \leq r \leq 1$ ซึ่งจะบอกได้เพียงขนาดหรือระดับของความสัมพันธ์เท่านั้น ไม่สามารถบอกทิศทางของความสัมพันธ์ได้

ค่าสัมประสิทธิ์สหสัมพันธ์จะใช้ได้อย่างเหมาะสมกับข้อมูลที่มีความสัมพันธ์เชิงเส้นเท่านั้น ดังนั้น ในการคำนวณหากพบว่าค่า $r = 0$ การตีความหมายว่าข้อมูลไม่มีความสัมพันธ์กันอาจไม่ถูกต้อง เนื่องจากอาจเป็นไปได้ว่าข้อมูลมีความสัมพันธ์กันในลักษณะอื่นที่ไม่ใช่เชิงเส้น (เช่น เส้นโค้ง ฯลฯ) ดังนั้น ในการสรุปจึงควรสรุปว่าข้อมูล 2 ชุดไม่มีความสัมพันธ์เชิงเส้น จึงจะถูกต้องชัดเจนกว่า สมการของสัมประสิทธิ์สหสัมพันธ์คือ

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

เมื่อ x_i, y_i คือ ค่าใดๆ ของแต่ละตัวแปรที่เป็นคู่กัน

$\bar{x}, \bar{y}$ คือ ค่าเฉลี่ยของแต่ละตัวแปร

2.2.2 การคัดเลือกคุณสมบัติ (Feature Selection)

การคัดเลือกคุณสมบัติเป็นเทคนิคที่ช่วยลดจำนวนตัวแปรที่จะใช้ในตัวแบบพยากรณ์ อาจกระทำเพื่อเลือกตัวแปรที่ดีที่สุดเพียงตัวเดียว หรือเลือกกลุ่มของตัวแปรที่มีความสำคัญต่อการพยากรณ์ กระบวนการคัดเลือกคุณสมบัติเป็นกระบวนการที่สำคัญในการเตรียมข้อมูลของการทำเหมืองข้อมูล เพื่อทำให้การสร้างตัวแบบพยากรณ์มีประสิทธิภาพ เพราะจะช่วยลดมิติของข้อมูล และอาจช่วยให้การเรียนรู้วิธีการพยากรณ์ดำเนินการได้เร็วขึ้นและมีประสิทธิภาพมากขึ้น ใน

งานวิจัยนี้ จะใช้วิธีการคัดเลือกคุณลักษณะแบบ CFS Subset Evaluable และวิธีการค้นหาแบบ Best-first search

การคัดเลือกคุณสมบัติแบบ (Correlation-based Feature Selection) เป็นการคัดเลือกกลุ่มคุณสมบัติอย่างง่าย ใช้หลักการคำนวณค่าความสัมพันธ์ระหว่างคุณสมบัติย่อยต่อค่าพยากรณ์ ซึ่งอาจใช้คำนวณด้วยค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson's correlation) และมีการจัดอันดับตามค่าความสัมพันธ์เพื่อประเมินค่าความสามารถในการพยากรณ์ของแต่ละคุณสมบัติ นอกจากนั้นยังพิจารณาคัดเลือกกลุ่มของคุณสมบัติที่มีความสัมพันธ์ภายในระหว่างคุณสมบัติย่อยกันเองต่ำเพื่อลดความซ้ำซ้อนของอิทธิพลการพยากรณ์ (Hall and Smith. 1998) ได้ดังสมการนี้

$$M_s = \frac{\overline{kr_{of}}}{\sqrt{k + k\,(k-1)\,\overline{r_{ff}}}}$$

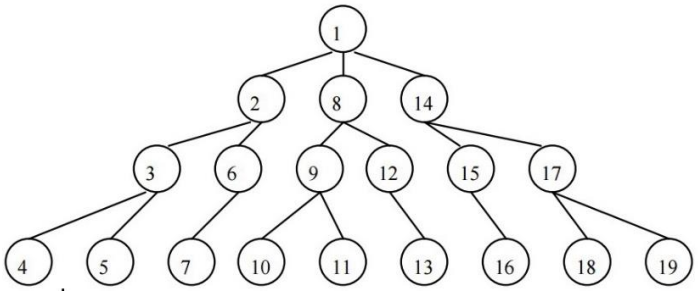
โดย M_i คือ ค่าที่ค้นหาได้มิติข้อมูลกลุ่มย่อย S ซึ่งประกอบด้วยมิติข้อมูล k

$\overline{r_{of}}$ คือ ค่าเฉลี่ยความสัมพันธ์ของตัวแปรกับคลาส $(f \in S)$

$\overline{r_{ff}}$ คือ ค่าเฉลี่ยความสัมพันธ์ระหว่างมิติของข้อมูล

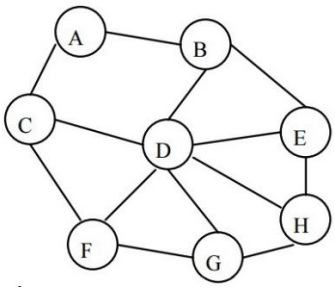
การค้นหาดีสุดก่อน (Best-first search) เป็นกระบวนการค้นหาข้อมูลที่ได้นำเอาข้อดีของทั้งการค้นหาแบบลึกก่อน (Depth first search) และการค้นหาแบบกว้างก่อน (Breadth first search) มารวมกันเป็นวิธีการเดียว โดยที่แต่ละขั้นของการค้นหาในโหนดลูกนั้น การค้นหาแบบดีที่ดีก่อนจะเลือกเอาโหนดที่ดีที่สุด (most promising) และการที่จะทราบว่าโหนดใดดีที่สุดนี้สามารถทำได้โดยอาศัยฮิวริสติกฟังก์ชัน ซึ่งฮิวริสติก ฟังก์ชันนี้จะทำหน้าที่เหมือนตัววัดผล และให้ผลของการวัดนี้ออกมาเป็นคะแนน

การค้นหาแบบลึกก่อน (Depth first search) เป็นการค้นหาที่กำหนดทิศทางจากรูปของโครงสร้างต้นไม้ ที่เริ่มต้นจากโหนดราก (Root node) ที่อยู่บนสุด แล้วเดินลงมาให้ลึกที่สุด เมื่อถึงโหนดล่างสุด (Terminal node) ให้ย้อนขึ้นมาที่จุดสูงสุดของกิ่งเดี่ยวกันที่มีกิ่งแยกและยังไม่ได้เดินผ่าน แล้วเริ่มเดินลงจนถึงโหนดลึกสุดอีก ทำเช่นนี้สลับไปเรื่องจนพบโหนดที่ต้องการหาหรือสำรวจครบทุกโหนดแล้วตามภาพที่ 2-2 การค้นหาแบบลึกก่อนจะมีลำดับการเดินตามโหนดดังตัวเลขที่กำกับไว้ในแต่ละโหนด



ภาพที่ 2-2 ลำดับการเดินทางบนโหนดของการค้นหาแบบลึกก่อน

โครงสร้างข้อมูลที่ใช้สำหรับการค้นหานี้สามารถใช้กับโครงสร้างกราฟได้ด้วย โดยอาศัยหลักการเดียวกัน แต่สำหรับการเดินทางบนกราฟนั้นจะไม่มีโหนดลึกสุด ดังนั้นการเดินทางบนกราฟจะต้องปรับวิธีการเป็นดังนี้ โดยเริ่มจากโหนดเริ่มต้น จากนั้นให้นำโหนดที่อยู่ติดกับโหนดที่กำลังสำรวจอยู่ (ที่ยังไม่ได้ทำการสำรวจและยังไม่ได้อยู่ในแสต็กมาใส่แสต็ก) มาเก็บไว้ในแสต็กเมื่อสำรวจโหนดนั้นเสร็จให้พอพ (pop) ตัวบนสุดของโหนดออกมาทำการสำรวจ แล้วนำโหนดข้างเคียงทั้งหมดที่ยังไม่ได้สำรวจมาต่อท้ายแสต็ก แล้วพอพตัวบนสุดออกมาสำรวจ ทำเช่นนี้เรื่อยๆ จนกระทั่งพบโหนดที่ต้องการ หรือสำรวจครบทุกโหนด



ภาพที่ 2-3 โครงสร้างข้อมูลแบบกราฟ

การสำรวจจะเริ่มต้นที่ A และนำโหนดข้างเคียง B และ C มาเก็บไว้ในแสต็ก เมื่อสำรวจ A เสร็จแล้วพอพข้อมูลจากแสต็กออกมาได้ C ทำการสำรวจ C แล้วนำโหนดข้างเคียงกับ C ที่ยังไม่ได้ทำการสำรวจแล้วยังไม่ได้อยู่ในแสต็กมาใส่แสต็ก D และ F พุช (Push) ใส่แสต็ก ดังนั้นในแสต็กตอนนี้มี B D F อยู่ เมื่อสำรวจ C เสร็จ พอพ F ออกมาทำการสำรวจ แล้วนำโหนดข้างเคียงที่ยังไม่ได้สำรวจและยังไม่ได้อยู่ในแสต็กมาใส่แสต็ก ซึ่งก็คือ G ดังนั้น ข้อมูลใน

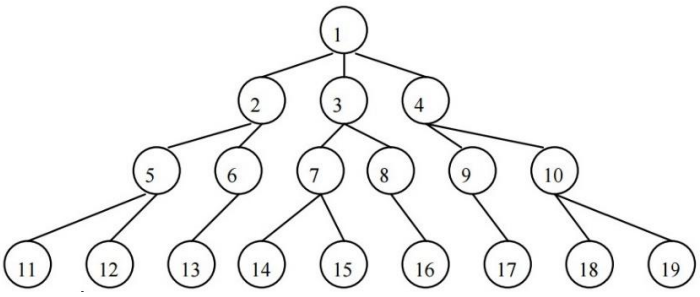
แสต็กจะเป็น B D G ทำเช่นนี้ไปเรื่อยๆ จนจบการทำงานก็จะได้ลำดับการสำรวจคือ (A C F G H E D B) ตามตารางที่ 2-1 ดังต่อไปนี้

ตารางที่ 2-1 ลำดับการค้นหาแบบลึกก่อน

โหนดที่สำรวจ	แสต็ก
A	B C
C	B D F
F	B D G
G	B D H
H	B D E
E	B D
D	B
B	

ในการค้นหาข้อมูลแบบนี้บนโครงสร้างของกราฟ มีข้อที่น่าสังเกตคือ โหนดที่เริ่มต้นการสำรวจจะต้องมีการกำหนดมาให้ว่าโหนดใดเป็นโหนดเริ่มต้น และข้อสังเกตอีกประการหนึ่งคือวิธีการค้นหาแบบลึกก่อนที่ใช้สำหรับโครงสร้างข้อมูลแบบกราฟ สามารถใช้กับโครงสร้างข้อมูลแบบต้นไม้ได้ด้วย

การค้นหาแบบกว้างก่อน (Breadth first search) เป็นการกำหนดทิศทางการค้นหาแบบทีละระดับของโครงสร้างต้นไม้ โดยเริ่มจากโหนดราก (ระดับที่ 0) แล้วลงมาระดับที่ 1 จากซ้ายไปขวา เมื่อเสร็จระดับที่ 1 ไประดับที่ 2 จากซ้ายไปขวาเช่นกัน ทำเช่นนี้เรื่อยๆ จนพบโหนดที่ต้องการตามภาพที่ 2-4 ลำดับการเดินทางของโหนดเป็นไปตามหมายเลขที่กำกับไว้บนโหนด



ภาพที่ 2-4 ลำดับการค้นหาแบบกว้างก่อนบนโครงสร้างต้นไม้

สำหรับการค้นหาแบบกว้างก่อนบนโครงสร้างต้นไม้ จะอาศัยโครงสร้างข้อมูลแบบคิว (Queue) มาช่วย และด้วยวิธีการเช่นเดียวกับการค้นหาแบบลึกก่อนคือให้เริ่มต้นสำรวจที่โหนดเริ่มต้น แล้วนำโหนดข้างเคียงเก็บไว้ในคิว เมื่อสำรวจโหนดเริ่มต้นเสร็จให้นำข้อมูลในคิวออกมาสำรวจ แล้วนำโหนดข้างเคียงที่ยังไม่ได้สำรวจและไม่ได้อยู่ในคิวใส่คิวไว้ ทำเช่นนี้ไปเรื่อยๆ จนพบโหนดที่ต้องการ หรือเมื่อสำรวจครบทุกโหนด

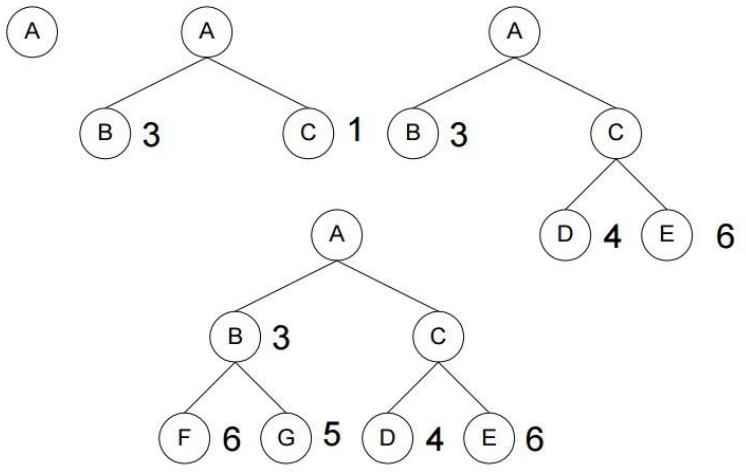
จากภาพที่ 2-3 โครงสร้างข้อมูลแบบกราฟ การสำรวจเริ่มต้นที่ A นำโหนดข้างเคียง B C ไว้ในคิว เมื่อสำรวจ A เสร็จ นำข้อมูลในคิว คือ B ออกมาสำรวจ แล้วนำข้อมูลข้างเคียงคือ D E ใส่คิว ตอนนี้คิวจะมี B D E อยู่ แล้วนำ B ออกมาสำรวจ ทำเช่นนี้เรื่อยๆ จะได้ลำดับการสำรวจข้อมูลคือ (A B C D E F G H) ตามตารางที่ 2-2 ดังต่อไปนี้

ตารางที่ 2-2 ลำดับการค้นหาแบบกว้างก่อน

โหนดที่สำรวจ	คิว
A	B C
B	C D E
C	D E F
D	E F G H
E	F G H
F	G H
G	H
H	

ฮิวริสติกฟังก์ชัน (heuristic function) หมายถึงฟังก์ชันที่ทำหน้าที่ในการวัดขนาดของความเป็นไปได้ในการแก้ปัญหาซึ่งจะแสดงด้วยตัวเลข วิธีการดังกล่าวจะกระทำได้โดยการพิจารณาถึงวิธีการ (aspects) ต่างๆ ที่ใช้ในการแก้ปัญหา ณ สถานะหนึ่งว่าจะสามารถแก้ปัญหาได้ตามที่ต้องการหรือไม่ โดยกำหนดเป็นน้ำหนักที่ให้กับการแก้ปัญหาของแต่ละวิธี น้ำหนักเหล่านี้จะถูกแสดงด้วยตัวเลขที่กำกับไว้กับโหนดต่างๆ ในกระบวนการค้นหา และค่าเหล่านี้จะเป็นตัวที่ใช้ในการประมาณความเป็นไปได้ว่าเส้นทางที่ผ่านโหนดนั้นจะมีความเป็นไปได้ในการนำไปสู่หนทางการแก้ปัญหาได้มากน้อยแค่ไหน ซึ่งจุดประสงค์ของฮิวริสติกฟังก์ชันก็คือ การกำกับทิศทางของกระบวนการค้นหาเพื่อให้อยู่ในทิศทางที่ได้ประโยชน์สูงสุด โดยการบอกว่าเราควรเลือกเดินเส้นทางไหนก่อน ในกรณีที่มีเส้นทางมากกว่าหนึ่งเส้นทางต้องเลือก

ตัวอย่างของการค้นหาแบบดีที่สุดก่อน ขั้นตอนนี้เริ่มจากตอน 1 สร้างโหนดราก (root node) ในขั้นตอน 2 สร้างโหนดลูก B และC แล้วตรวจสอบโหนด B และC ด้วยฮิวริสติกฟังก์ชัน ได้ผลออกมาเป็นคะแนนคือ 3 และ1 ตามลำดับ จากนั้นให้เลือกโหนด C เป็นโหนดต่อไปที่เราสนใจ เพราะมีค่าน้อยกว่า (หมายเหตุ ในการเลือกนี้จะเลือกค่ามากสุด หรือน้อยสุดก็ได้ขึ้นอยู่กับลักษณะของปัญหา) แล้วสร้างโหนดลูกให้กับโหนด C ในขั้นตอน 3 ได้โหนด D และE แล้วตรวจสอบคะแนนได้ 4 และ 6 ตามลำดับ จากนั้นทำการเปรียบเทียบค่าของโหนดท้ายสุด หรือเทอร์มินอล โหนด (terminal node) ทุกโหนดว่าโหนด ใดมีค่าดีที่สุด ในที่นี้จะต้องเลือกโหนด B เพราะมีคะแนนเพียง 3 (เลือกคะแนนต่ำสุด) แล้วสร้างโหนด ลูกตามขั้นตอน 4 ได้ F และG แล้วตรวจ สอบคะแนนได้ 6 และ5 คะแนนตามลำดับ ทำเช่นนี้เรื่อยๆ จนพบคำตอบหรือจนไม่สามารถสร้างโหนดต่อไปได้อีก ดังภาพที่ 2-5



ภาพที่ 2-5 ขั้นตอนของการค้นหาแบบดีที่สุดก่อน

2.3 ความรู้เกี่ยวกับการพยากรณ์และเทคนิคที่ใช้ในการพยากรณ์

การพยากรณ์ คือ การคาดการณ์ถึงสิ่งใดสิ่งหนึ่งที่จะเกิดขึ้นในช่วงเวลาในอนาคต และนำค่าพยากรณ์ที่ได้นั้นมาใช้ประโยชน์ เพื่อการตัดสินใจใดๆ หรืออาจหมายถึง การทำนายเหตุการณ์ในอนาคต อาจนำข้อมูลในอดีตมาพยากรณ์อนาคตโดยใช้หลักคณิตศาสตร์ ใช้ดุลยพินิจของผู้พยากรณ์ หรืออาจใช้หลายๆ วิธีมารวมกัน เพื่อให้การพยากรณ์มีความแม่นยำมากที่สุด ซึ่งมีหลักการดังนี้

การพยากรณ์แบ่งได้ 2 ประเภท คือ การพยากรณ์เชิงคุณภาพ (Qualitative methods) เป็นการพยากรณ์โดยอาศัยข้อมูลหรือตัวเลขจากอดีต (เช่น ยอดขาย กำลังการผลิต) มาสร้างตัวแบบ และ

พยากรณ์ไปในอนาคต และการพยากรณ์เชิงปริมาณ (Quantitative methods) เป็นกลุ่มของวิธีการพยากรณ์ที่อาศัยข้อมูลและวิธีการเชิงคุณภาพ ใช้กับลักษณะของปัญหาที่ไม่มีข้อมูลย้อนหลังหรือมีข้อมูลไม่มากพอมาใช้สร้างตัวแบบ ซึ่งในงานวิจัยนี้จะใช้การพยากรณ์เชิงปริมาณที่ใช้ข้อมูลเชิงปริมาณ (ตัวเลข) ในอดีตเพื่อนำมาพยากรณ์ค่าในอนาคต โดยสร้างตัวแบบทางคณิตศาสตร์ การพยากรณ์ประเภทนี้แบ่งออกเป็น 2 เทคนิคย่อย คือ

1) การพยากรณ์ความสัมพันธ์ (Casual Forecasting) เป็นเทคนิคที่ใช้ปัจจัยที่คาดว่าจะมีความสัมพันธ์กับตัวแปรที่จะพยากรณ์ เช่น ถ้าต้องการพยากรณ์ยอดขาย จะพิจารณาหาความสัมพันธ์ระหว่างยอดขายกับค่าโฆษณา รายได้ของประชากร สภาพสินค้า ฯลฯ การหาความสัมพันธ์ดังกล่าวจะใช้เทคนิคที่เรียกว่า การวิเคราะห์ความถดถอย และสหสัมพันธ์

2) การพยากรณ์อนุกรมเวลา (Time series forecasting) เป็นเทคนิคที่ใช้เฉพาะข้อมูลในอดีตของตัวแปรที่ต้องการพยากรณ์ เพื่อพยากรณ์ค่าของตัวแปรนั้นในอนาคต เช่น ใช้ข้อมูลยอดขายปี 2550-2555 เพื่อพยากรณ์ยอดขาย ปี 2556

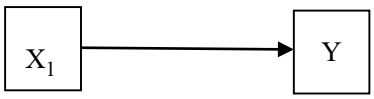
2.3.1 การวิเคราะห์การถดถอย (Regression Analysis)

การวิเคราะห์ความถดถอย เป็นการศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะประมาณหรือพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์ โดยจะต้องมีการกำหนดหรือทราบค่าของตัวแปรอื่นๆ ล่วงหน้า เช่น ถ้าทราบความสัมพันธ์ระหว่างยอดขายกับค่าโฆษณาแล้ว จะทำให้สามารถประมาณ / พยากรณ์ยอดขายเมื่อกำหนด หรือทราบงบประมาณในการโฆษณา และจะศึกษาถึงการเปลี่ยนแปลงของยอดขายเมื่องบประมาณในการโฆษณาเปลี่ยนแปลงไป โดยอาศัยหลักการของการวิเคราะห์ความถดถอย การวิเคราะห์ความถดถอยมี 2 ประเภท คือ การวิเคราะห์ความถดถอยอย่างง่าย และการวิเคราะห์ความถดถอยเชิงซ้อน หรือ เชิงพหุ

การวิเคราะห์การถดถอย เป็นเทคนิคการพยากรณ์ที่นิยมใช้กันอย่างกว้างขวางมากที่สุดในสาขาวิชาต่าง ๆโดยใช้ตัวแปรอิสระ (Independent Variable) หรือตัวแปรพยากรณ์ (Predictor) อย่างน้อยหนึ่งตัวไปพยากรณ์หรืออธิบายความผันแปรของตัวแปรตาม (Dependent Variable) ถ้าการวิเคราะห์การถดถอยประกอบด้วยตัวแปรอิสระ 1 ตัว ใช้สำหรับพยากรณ์ตัวแปรตาม 1 ตัว ซึ่งวัดเป็นค่าในเชิงปริมาณ เรียกว่า การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression) แต่ถ้าใช้ตัวแปรอิสระมากกว่า 1 ตัว มาพยากรณ์ตัวแปรตามที่วัดในเชิงปริมาณ 1 ตัว เรียกว่า การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression)

กรอบแนวคิดในการวิเคราะห์การถดถอย จะอาศัยความสัมพันธ์ระหว่างตัวแปร เป็นหลักและคำตอบที่ต้องการคือ มีตัวแปรใดบ้างที่สามารถพยากรณ์ตัวแปรตามที่เราสนใจศึกษา และตัว

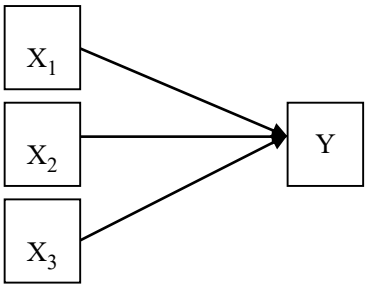
แปรใดพยากรณ์ได้มากน้อยกว่ากัน รวมทั้งส่งผลในทางบวกหรือทางลบ สามารถเขียนเป็นสมการกรอบแนวคิดดังภาพที่ 2-6



X = 1 ตัว Y = 1 ตัว

การวิเคราะห์การถดถอยอย่างง่าย

Simple Regression



X > 1 ตัว Y = 1 ตัว

การวิเคราะห์การถดถอยพหุคูณ

Multiple Regression

ภาพที่ 2-6 กรอบแนวคิดการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ

ในการวิเคราะห์การถดถอยพหุคูณ จะช่วยให้ได้สมการพยากรณ์เชิงเส้นตรงในรูปคะแนนดิบ ดังนี้

$$Y' = a+b_1X_1+b_2X_2+\ldots+b_kX_k$$

เมื่อ Y' แทน คะแนนพยากรณ์ของตัวแปรตาม (ตัวเกณฑ์)

a แทน ค่าคงที่ของสมการพยากรณ์ในรูปแบบคะแนนดิบ

$b_1, b_2 \ldots b_k$ แทน ค่าน้ำหนักคะแนนหรือสัมประสิทธิ์การถดถอยของตัวแปรอิสระ (ตัวพยากรณ์) ตัวที่ 1 ถึงตัวที่ k ตามลำดับ

$X_1, X_2 \ldots X_k$ แทน คะแนนของตัวแปรอิสระ (ตัวพยากรณ์) ตัวที่ 1 ถึงตัวที่ k ตามลำดับ

k แทน จำนวนตัวแปรอิสระ (ตัวพยากรณ์)

การเขียนสมการในรูปคะแนนดิบจะต้องทราบค่า a และ b เพื่อนำมาแทนค่าในสมการ ค่า a จากสูตร

$$a = \overline{Y} - b_1 \overline{X}_1 - b_2 \overline{X}_2 - \ldots - b_k \overline{X}_k$$

เมื่อ a แทน ค่าคงที่สำหรับสมการพยากรณ์ในรูปคะแนนดิบ

$\overline{Y}$ แทน ค่าเฉลี่ยสำหรับตัวแปรตาม

$\overline{X}_1, \overline{X}_2, \overline{X}_k$ แทน ค่าเฉลี่ยของตัวแปรอิสระ (ตัวแปรพยากรณ์) ตัวที่ 1 ถึง k ตามลำดับ

b_1, b_2, b_k แทน ค่าน้ำหนักของตัวแปรอิสระ (ตัวแปรพยากรณ์) ตัวที่ 1 ถึง k ตามลำดับ

k แทน จำนวนตัวแปรอิสระ (ตัวแปรพยากรณ์)

ค่า b หาจากสูตร

$$b_j = \beta_j \frac{S_y}{S_j}$$

เมื่อ b_j แทน ค่าน้ำหนักคะแนนหรือสัมประสิทธิ์การถดถอยของตัวแปรอิสระ (ตัวพยากรณ์) ตัวที่ j ที่ต้องการหาค่าน้ำหนัก

β_1 แทน ค่าน้ำหนักเบต้าของตัวแปรอิสระ (ตัวพยากรณ์) ตัวที่ j

S_y แทน ส่วนเบี่ยงเบนมาตรฐานของตัวแปรตาม (ตัวเกณฑ์)

S_j แทน ส่วนเบี่ยงเบนมาตรฐานของตัวแปรอิสระ (ตัวแปรพยากรณ์)

สัมประสิทธิ์การถดถอย (b) เป็นค่าที่ชี้ถึงว่า เมื่อตัวแปรอิสระ (ตัวพยากรณ์) (X) ตัวนั้นเปลี่ยนแปลงไป 1 หน่วยจะทำให้ตัวแปรตรา (ตัวเกณฑ์) (คะแนนพยากรณ์ของตัวแปรตาม) เปลี่ยนแปลงไป b หน่วย

การวิเคราะห์การถดถอยพหุคูณ จะต้องคำนวณหาค่า a และ $b_1\ b_2 \ldots b_k$ เพื่อนำมาแทนค่าลงในสมการ โดยถือหลักการที่ว่า ค่า b ทุกตัวต้องเป็นค่าที่ทำให้สมการพยากรณ์มีความคลาดเคลื่อนในการพยากรณ์น้อยที่สุด นอกจากจะหาค่า a และ b แต่ละตัวแล้ว ควรทดสอบความนัยสำคัญของค่า b แต่ละตัวด้วย

จุดประสงค์หลักของการวิเคราะห์การถดถอย คือ เพื่อพยากรณ์ตัวแปรหนึ่ง โดยใช้ค่าที่ทราบของตัวแปรอีกตัวหนึ่ง การพยากรณ์นี้อาศัยหลักการ เช่น $Y = a + bX$ ซึ่งมีไว้เพื่อประมาณค่าของตัวแปรที่ไม่ทราบค่า Y เมื่อทราบค่าของตัวแปร X เรียกสมการนี้ว่า สมการการถดถอย เมื่อทราบสมการการถดถอย เราสามารถพยากรณ์ Y จากค่า X ที่กำหนดให้ สมการการถดถอยไม่เหมือนสมการทางคณิตศาสตร์ทั่ว ๆ ไป ในแง่ที่เราไม่สามารถจะมั่นใจกับค่าของ Y ที่ได้จากสมการการถดถอย เนื่องจากค่านี้มีการคลาดเคลื่อนและเป็นเพียงค่าประมาณของค่าที่แท้จริง (Y) เท่านั้น

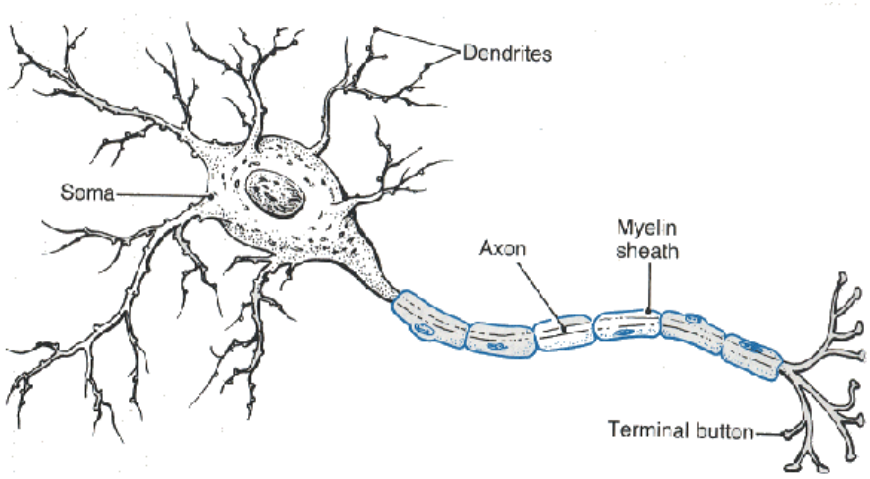
2.3.2 โครงข่ายประสาทเทียม (Neural Networks)

โครงข่ายประสาทเทียม เป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่เน้นด้านการคำนวณ และได้รับแรงบันดาลใจจากโมเดลทางชีวภาพ โดยมีความพยายามที่จะลอกเลียนความสามารถของสมองมนุษย์ยกตัวอย่างเช่น พัฒนาการด้านสมองของเด็กตั้งแต่แรกเกิดจนถึงสามปีที่ในระยะเริ่มต้นของวัยนี้ เด็กจะเรียนรู้ว่าผู้ใดคือพ่อแม่ หรือคนแปลกหน้าในระยะต่อมาจะเริ่มตอบสนองเมื่อถูกเรียกชื่อ รู้จักแยกแยะสีและสิ่งของต่างๆ เริ่มเปล่งเสียงพูดเป็นคำ และเริ่มแสดงความคิดเห็น เป็นต้น จากลักษณะดังกล่าวจะเห็นว่าความสามารถในการเรียนรู้ของสมองมนุษย์มีการพัฒนาขึ้นตามลำดับ และทำงานได้อย่างมีประสิทธิภาพ แต่เมื่อเปรียบเทียบกับการทำงานของเครื่องคอมพิวเตอร์จะพบว่ามีการทำงานตามคำสั่งอย่างตรงไปตรงมาเท่านั้น ยกตัวอย่างเช่น เปิดแฟ้มข้อมูล การใช้อินเตอร์เน็ต และสืบค้นข้อมูล เป็นต้น การทำงานดังกล่าวไม่มีประสิทธิภาพพอที่จะให้เครื่องแสดงความคิดเห็น หรือ วิพากษ์วิจารณ์เกี่ยวกับข้อมูลต่างๆ ได้ ดังนั้น จะเห็นได้ว่าการทำงานของสมองมนุษย์มีความ ซับซ้อนเป็นอย่างมาก ซึ่งนักประสาทวิทยาตั้งแต่อดีตถึงปัจจุบันได้พยายามศึกษาค้นคว้ามาอย่างต่อเนื่องที่จะสร้างโปรแกรมคอมพิวเตอร์ที่มีลักษณะการประมวลผลคล้ายการทำงานของสมองมนุษย์และผลการศึกษาค้นคว้าที่ผ่านมา ส่วนหนึ่งก็สามารถนำไปประยุกต์ใช้ในสาขาวิชาต่างๆ ได้อย่างเป็นโยชน์อย่างมากมาย รูปแบบของโครงข่ายประสาทเทียมเป็นความก้าวหน้าอย่างหนึ่งของความพยายามที่จะพัฒนาโปรแกรมคอมพิวเตอร์ในปัจจุบันให้สามารถทำนายสิ่งต่างๆ ได้อย่างกว้างขวาง เนื่องจากโครงข่ายประสารทเทียมสามารถหาความสัมพันธ์ของข้อมูลนำเข้า (input) และข้อมูลส่งออก (output) ได้โดยไม่จำเป็นต้องรู้ความสัมพันธ์ทางคณิตศาสตร์ของข้อมูลที่เป็นข้อมูลนำเข้าและข้อมูลส่งออกมาก่อน

หลักการของโครงข่ายประสาทเทียม คือแบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองการทำงานของโครงข่ายประสาทในสมองมนุษย์ โดยที่โครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาท (Neuron) และจุดประสานประสาทหรือไซแนปส์ (Synapses) โดยโครงสร้างของการส่งสัญญาณประสาทประกอบจากการเชื่อมต่อระหว่างเซลล์ประสาทหลายพันล้านเซลล์ เซลล์ประสาทแต่ละเซลล์ประกอบด้วยแขนงรับสัญญาณประสาท ซึ่งเป็นเสมือนหน่วยรับข้อมูลป้อนเข้า เรียกว่า “เดนไดรท์” (dendrites) และส่วนปลายของเซลล์ประสาทในการส่งสัญญาณประสาท ซึ่งเป็นเสมือนหน่วยส่งข้อมูลออกของเซลล์ เรียกว่า “แอคซอน” (axon) โดยการส่งสัญญาณประสาทดังกล่าวอาจทำให้เกิดได้ทั้งการกระตุ้นและยับยั้ง ทั้งนี้ นอกจากลักษณะดังกล่าวแล้ว วิธีการประมวลผลภายในเซลล์ประสาทแต่ละเซลล์ยังมีการขยายหรือลดขนาดของสัญญาณอีกด้วย โดยสัญญาณจากเดนไดรท์ต่างๆ จะรวมกันเข้าสู่เซลล์ประสาท และหากสัญญาณ

รวมมีความแรงเกินค่าระดับ (threshold) ของเซลล์ประสาทนั้นๆ เซลล์ประสาทก็จะส่งสัญญาณออกทางแอคซอนต่อไป

กระบวนการเรียนรู้ในสิ่งมีชีวิตจะมีผลให้เกิดการสร้างไซแนปส์ระหว่างเซลล์ประสาทขึ้นมาใหม่ หรือทำให้เกิดการเปลี่ยนสภาพของไซแนปส์ต่างๆ ในโครงข่ายของเซลล์ประสาท โครงข่ายประสาทของสิ่งมีชีวิตจึงไม่ได้ทำงานแบบเป็นลำดับขั้นตอน (sequential) แต่เพียงอย่างเดียวแสดงดังภาพที่ 2-7



ภาพที่ 2-7 โครงสร้างระบบประสาท (www.mindcreators.com/NeuronBasics.htm)

โครงสร้างประสาทเทียมมีคุณลักษณะคล้ายกับการส่งผ่านสัญญาณประสาทในสมองของมนุษย์กล่าวคือ มีความสามารถในการรวบรวมความรู้ (knowledge) โดยผ่านกระบวนการเรียนรู้ (learning process) และความรู้เหล่านี้จะจัดเก็บอยู่ในโครงข่ายในรูปแบบค่าน้ำหนัก (weight) ซึ่งสามารถปรับเปลี่ยนค่าได้เมื่อมีการเรียนรู้สิ่งใหม่ๆ เข้าไป ค่าน้ำหนักทำหน้าที่เปรียบเสมือนความรู้ที่รวบรวมไว้เพื่อใช้ในการแก้ปัญหาเฉพาะอย่างของมนุษย์ การประมวลผลต่างๆ เกิดขึ้นในหน่วยประมวลผลย่อย เรียกว่า โหนด (node) ซึ่งโหนดเป็นการจำลองลักษณะการทำงานมาจากเซลล์การส่งสัญญาณ (signal) ระหว่างโหนดที่เชื่อมต่อกัน (connection) จำลองมาจากการเชื่อมต่อของเดนไดร์ทและแอคซอนในระบบประสาทของมนุษย์ ภายในโหนดจะมีฟังก์ชันกระตุ้น (activationfunction) หรือฟังก์ชันการแปลง (transfer function) ซึ่งทำหน้าที่เปรียบเสมือนกระบวนการทำงานในเซลล์ โครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบดังนี้

1. ข้อมูลป้อนเข้า (input) เป็นข้อมูลที่เป็นตัวเลข หากเป็นข้อมูลเชิงคุณภาพต้องแปลงให้อยู่ในรูปเชิงปริมาณที่โครงข่ายประสาทเทียมยอมรับได้

2. ข้อมูลส่งออก (output) คือ ผลลัพธ์ที่เกิดขึ้นจริง (actual output) จากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม

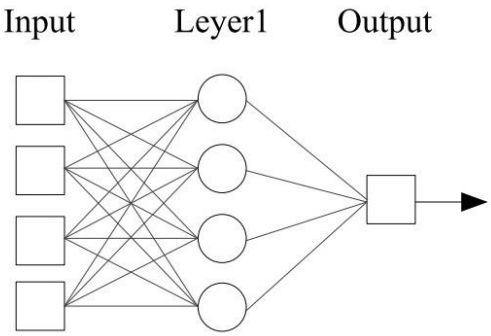
3. ค่าน้ำหนัก (weights) คือ สิ่งที่ได้จากการเรียนรู้ของโครงข่ายประสาทเทียม หรือเรียกอีกอย่างหนึ่งว่า ค่าความรู้ (knowledge) ค่านั้นจะถูกเก็บเป็นทักษะเพื่อใช้ในการจดจำข้อมูลอื่นๆ ที่อยู่ในรูปแบบเดียวกัน

4. ฟังก์ชันผลรวม (Summation function: S) เป็นผลรวมของข้อมูลป้อนเข้า (input) และค่าน้ำหนัก (wi)

5. ฟังก์ชันการแปลง (transfer function) เป็นการคำนวณการจำลองการทำงานของโครงข่ายประสาทเทียม เช่น ซิกมอยด์ฟังก์ชัน (sigmoid function) ฟังก์ชันไฮเปอร์โบลิกแทนเจนต์ (hyperbolic tangent function) เป็นต้น

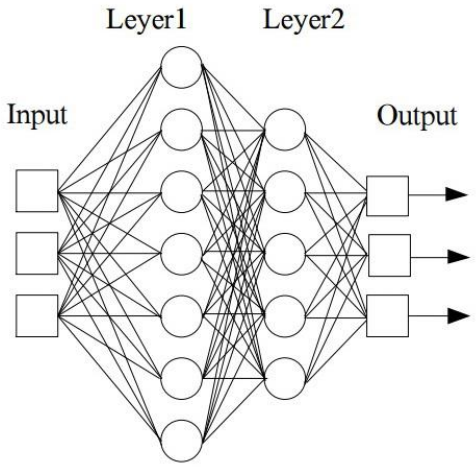
ลักษณะของโครงข่ายประสาทเทียมประกอบด้วยเซลล์ประสาทเทียม หรือโหนดจำนวนมากเชื่อมต่อกัน ซึ่งการเชื่อมต่อแบ่งออกเป็นกลุ่มย่อยเรียกว่า ชั้น (layer) ชั้นแรกเป็นชั้นนำข้อมูลเข้าเรียกว่า ชั้นรับข้อมูลป้อนเข้า (input layer) ส่วนชั้นสุดท้ายเรียกว่า ชั้นส่งข้อมูลออก (output layer) และชั้นที่อยู่ระหว่างชั้นรับข้อมูลป้อนเข้า และชั้นส่งข้อมูลออกเรียกว่า ชั้นแอบแฝง (hidden layer) ซึ่งโดยทั่วไปชั้นแอบแฝงอาจมีมากกว่า 1 ชั้นก็ได้ ด้วยเหตุนี้จึงสามารถแบ่งประเภทของโครงข่ายประสาทเทียมตามจำนวนชั้นของโครงข่ายแบบกว้างๆ ได้ 2 แบบ ได้แก่ โครงข่ายแบบชั้นเดียว (single layer) และโครงข่ายแบบหลายชั้น (multi-layer)

1. โครงข่ายแบบชั้นเดียวเป็นโครงข่ายประสาทเทียมอย่างง่ายที่มีเพียงชั้นรับข้อมูลป้อนเข้าและชั้นส่งข้อมูลออกเท่านั้น โหนดในชั้นรับข้อมูลป้อนเข้าทำหน้าที่รับข้อมูลเข้า (input value) แล้วส่งข้อมูลผ่านเส้นเชื่อมโยงต่างๆ ไปให้โหนดในชั้นส่งข้อมูลออก ความเข้มของสัญญาณ หรือปริมาณข้อมูลที่นำเข้าสู่โหนดในชั้นส่งข้อมูลออกจะขึ้นอยู่กับค่าน้ำหนักที่อยู่บนเส้นเชื่อมโยง โหนดในชั้นส่งข้อมูลออกจะนำข้อมูลที่ได้รับมาคำนวณโดยใช้ฟังก์ชันทางคณิตศาสตร์ที่เรียกว่า ฟังก์ชันการแปลง (transfer function) ที่เหมาะสมกับปัญหาแล้วส่งผลลัพธ์ที่ได้ออกมาเป็นข้อมูลส่งออก เช่น โครงข่ายแบบชั้นเดียวแบบเพอเซบตรอนอย่างง่าย (simple perceptron) และโครงข่ายโฮบฟิลด์ (Hopfield networks) ลักษณะโครงข่ายแบบชั้นเดียวแสดงดังภาพที่ 2-8



ภาพที่ 2-8 โครงสร้างโครงข่ายประสาทเทียมแบบชั้นเดียว

2. โครงข่ายแบบหลายชั้นเป็นโครงข่ายที่มีชั้นแอบแฝงตั้งแต่ 1 ชั้นขึ้นไป โครงข่ายแบบหลายชั้นจะใช้ในกรณีที่ปัญหามีความซับซ้อน ซึ่งโครงข่ายแบบชั้นเดียวไม่สามารถแก้ปัญหาได้ จึงเพิ่มจำนวนโหนดที่มีการคำนวณหรือชั้นแอบแฝงให้กับโครงข่าย ตัวอย่างของโครงข่ายแบบหลายชั้น เช่น การแพร่ย้อนกลับ (back propagation) เซลฟ์ ออร์แกนไนซิ่งแมปซ์ (self-organizing maps) และเคาน์เตอร์พรอพะเกชัน (counter propagation) เป็นต้น ลักษณะโครงสร้างโครงข่ายแบบหลายชั้น แสดงดังภาพที่ 2-9



ภาพที่ 2-9 โครงข่ายประสาทเทียมแบบหลายชั้น

ประเภทของการเรียนรู้ของโครงข่ายประสาทเทียม

1. การเรียนรู้แบบมีผู้สอน (Supervised learning) ข้อมูลจะประกอบด้วยตัวอย่างข้อมูลที่ต้องการสอน และผลลัพธ์ที่ต้องการให้โครงข่ายสร้าง เมื่อมีการนำข้อมูลในลักษณะเดียวกันมาเป็นข้อมูลป้อนเข้า โครงข่ายจะกำหนดค่าผลลัพธ์ที่เป็นเป้าหมายให้กับข้อมูลป้อนเข้าแต่ตัว โครงข่ายจะนำค่าผิดพลาดระหว่างค่าเป้าหมายกับค่าผลลัพธ์ที่ได้มาใช้ในการปรับค่าน้ำหนักเพื่อให้ค่าผลลัพธ์ที่ใกล้เคียงกับเป้าหมายที่มากที่สุด ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกันการสอนนักเรียนโดยมีครูผู้สอนคอยให้คำแนะนำ ตัวอย่างแบบจำลองนี้ ได้แก่ การแพร่ย้อนกลับ และเพอเซพตรอน (perceptron) เป็นต้น

2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) การเรียนรู้แบบนี้จะสอนโครงข่ายโดยการนำข้อมูลป้อนเข้าอย่างต่อเนื่องเพียงอย่างเดียวไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลป้อนเข้าแต่ละตัว การปรับน้ำหนักจะใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่า โดยค่าน้ำหนักจะปรับตามกลุ่มที่ข้อมูลป้อนเข้าที่มีรูปแบบคล้ายคลึงกัน ถ้าหากเปรียบเทียบกับมนุษย์จะเหมือนกับการที่เราสามารถแยกแยะพันธุ์พืช พันธุ์สัตว์ตามลักษณะรูปร่างของมันได้ด้วยตนเอง ตัวอย่างแบบจำลองนี้ได้แก่ เคาน์เตอร์เพรอพะเกชัน (counter propagation: CPN) แบบจำลองอะแดปทีฟรีโซแนนซ์เทียรี่ (Adaptive Resonance Theory neural networks: ART) เป็นต้น

โครงข่ายประสาทเทียมเป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ที่ประยุกต์ใช้ความรู้จากหลากหลายสาขามารวมเข้าด้วยกัน นับว่าเป็นศาสตร์ที่กำลังมีบทบาทอย่างยิ่งในปัจจุบัน กล่าวโดยสรุป โครงข่ายประสาทเทียมสามารถนำมาประยุกต์ใช้กับงานด้านต่างๆ ได้ดังนี้

1. การจำแนกรูปแบบ (pattern recognition) เช่น การมอบเห็นวัตถุหรือการวิเคราะห์เสียงพูดเพื่อแปลความหมาย

2. การทำนาย (prediction) หรือการพยากรณ์ (forecasting) เช่น การทำนายราคาหุ้นของตลาดหลักทรัพย์ การพยากรณ์อัตราการไหลของน้ำ การพยากรณ์ราคาสินค้า เป็นต้น

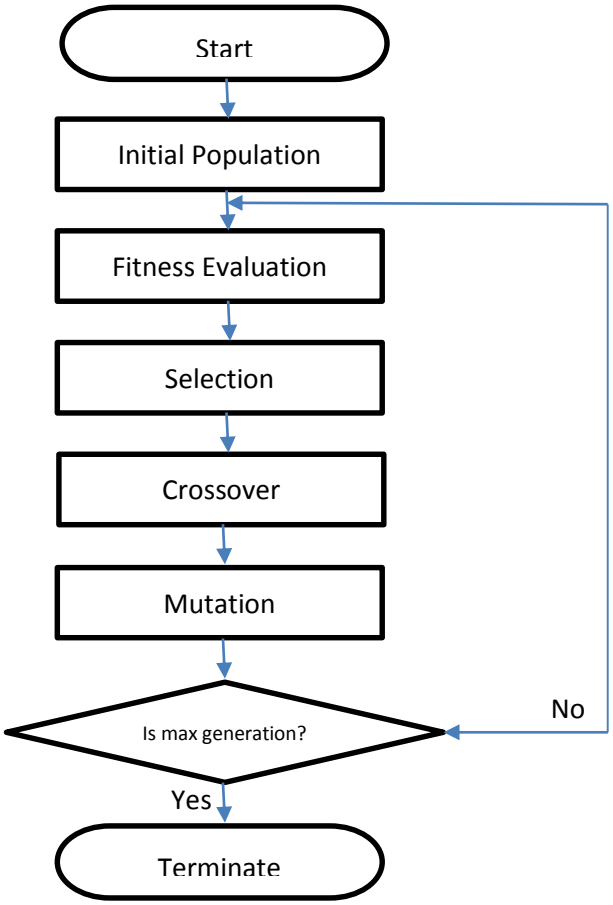
3. การควบคุม (control) เช่น การควบคุมระบบของเครื่องปรับอากาศ การควบคุมระบบเครื่องยนต์ และการควบคุมหุ่นยนต์ เป็นต้น

4. การหาความเหมาะสม (optimization) เช่น การเลือกระยะทางที่ใกล้หรือสั้นที่สุดในการเดินทาง (shortest path)

5. การจัดกลุ่ม (clustering) และการจัดหมู่ (categorization) เช่น การวิเคราะห์ข้อมูลจากภาพถ่ายดาวเทียม หรือภาพถ่ายทางอากาศ

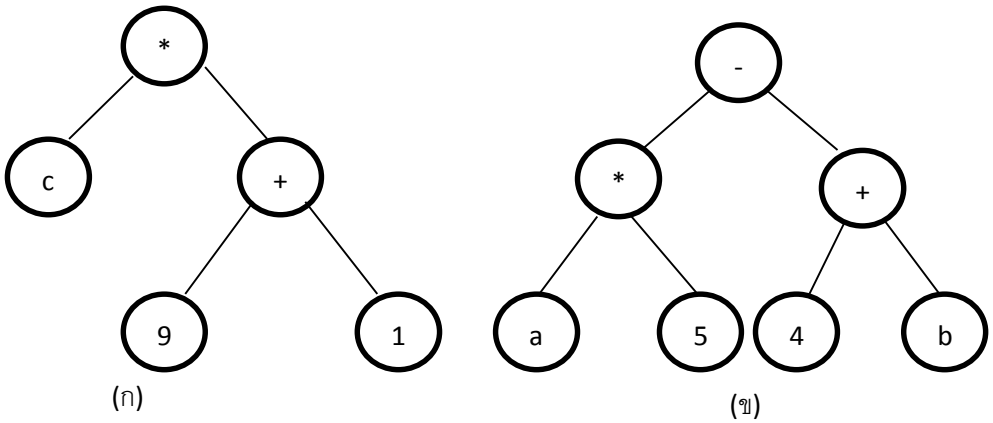
2.3.3 การโปรแกรมเชิงพันธุกรรม (Genetic Programming)

การโปรแกรมเชิงพันธุกรรม เป็นเทคนิคการคำนวณซึ่งมีรากฐานมาจากกระบวนการทางชีวภาพ โดยที่กลไกการทำงานอิงมาจากทฤษฎีวิวัฒนาการ การโปรแกรมเชิงพันธุกรรมพัฒนาจากชั้นตอนวิธีเชิงพันธุกรรม แต่เปลี่ยนรูปแบบของผลเฉลยให้อยู่ในรูปของต้นไม้ที่ประกอบด้วยโหนดจากเซตฟังก์ชัน และเซตปลายฟังก์ชันคือเซตของตัวดำเนินการ เช่น ตัวดำเนินการทางคณิตศาสตร์ (+, -) ฟังก์ชันทางคณิตศาสตร์ (sin, esp) ตัวดำเนินการทางตรรกศาสตร์ (if –then – else) และตัวดำเนินการซ้ำ (for, do while) เป็นต้น เซตปลายคือเซตของตัวถูกดำเนินการหรืออาร์กิวเมนต์ของฟังก์ชัน เช่น ตัวแปรคงที่ จากรูปแบบของผลเฉลยสรุปได้ว่าการโปรแกรมเชิงพันธุกรรมคือเทคนิคการค้นหาโปรแกรมคอมพิวเตอร์สำหรับแก้ปัญหาที่ต้องการ ขั้นตอนการทำงานของการโปรแกรมเชิงพันธุกรรมแสดงดังภาพที่ 2-10 และรายละเอียดการทำงานแต่ละขั้นตอนเป็นดังนี้



ภาพที่ 2-10 ขั้นตอนการทำงานของการโปรแกรมเชิงพันธุกรรม

การสร้างประชากรเริ่มต้น โครงสร้างอย่างง่ายของประชากรจะอยู่ในรูปต้นไม้ โดยปกติการสร้างประชากรมีอยู่ 2 วิธี คือ ระเบียบวิธีต้นไม้แบบไม่เต็มต้น (Grow Method) และระเบียบวิธีต้นไม้แบบเต็มต้น (Full Method)

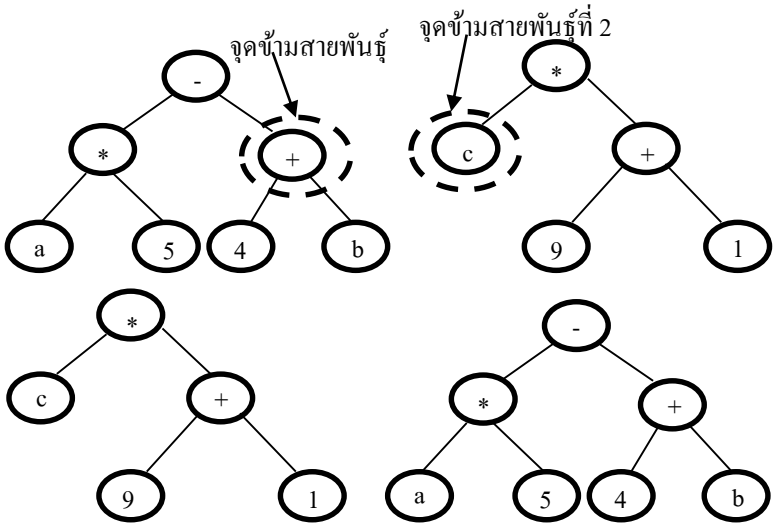


ภาพที่ 2-11 ผลที่ได้จากการสร้างประชากรโดย (ก) ระเบียบวิธีต้นไม้แบบไม่เต็มต้น
(ข) ระเบียบวิธีต้นไม้แบบเต็มต้น

การคำนวณค่าความแข็งแรง คือ ค่าที่บอกถึงคุณลักษณะของประชากร ถ้ามีค่ามากแสดงว่าเป็นประชากรที่ดี แต่ถ้ามีค่าน้อยแสดงว่าเป็นประชากรที่ไม่ดี โดยปกตินิยามความแข็งแรงอิงค่าความผิดพลาดของการแก้ปัญหาด้วยสมาชิกของประชากร

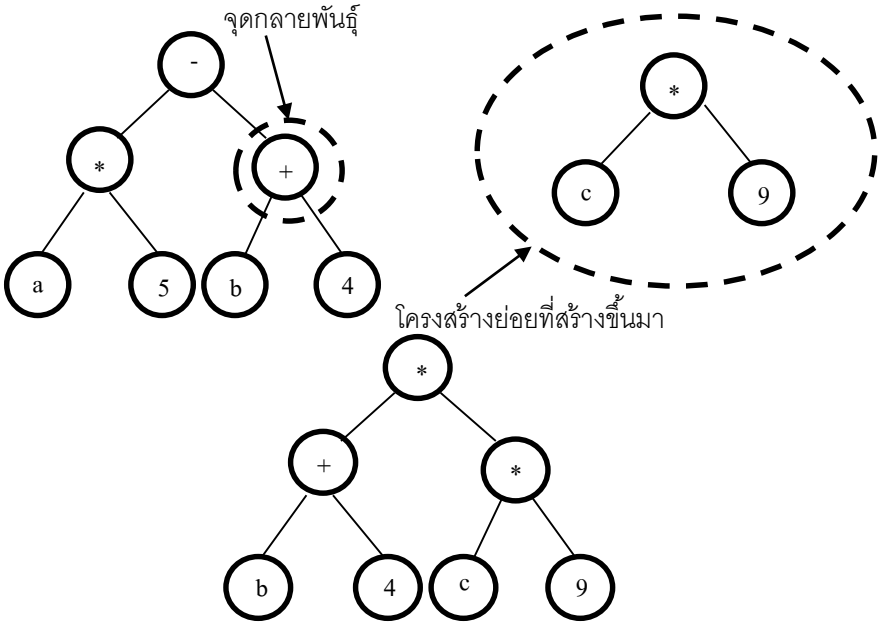
การคัดเลือกจะสร้างประชากรใหม่ โดยแพร่พันธุ์สมาชิกของประชากรรุ่นปัจจุบันไปยังประชากรรุ่นใหม่ โดยใช้ค่าความแข็งแรงเป็นตัวกำหนดอัตราการแพร่พันธุ์ สมาชิกของประชากรที่มีค่าความแข็งแรงมากจะมีโอกาสในการแพร่พันธุ์สูงกว่าสมาชิกของประชากรที่มีค่าความแข็งแรงน้อย

การข้ามสายพันธุ์จะสุ่มเลือกสมาชิกของประชากรครั้งละ 2 ตัวจากประชากรที่ได้หลังจากระบวนการคัดเลือกและเรียกว่า สมาชิกรุ่นพ่อแม่ จากนั้นยืนสมาชิกรุ่นพ่อแม่ทั้งสองตัวนี้จะถูกสุ่มเลือกจะในการข้ามสายพันธุ์ และทำการข้ามสายพันธุ์ส่งผ่านไปยังสมาชิกรุ่นลูก จำนวน 2 ตัวเช่นกัน ดังภาพที่ 2-12



ภาพที่ 2-12 การข้ามสายพันธุ์

การกลายพันธุ์เป็นอีกกระบวนการหนึ่งในการสร้างความหลากหลายของประชากรจากการสุ่มจุดในการกลายพันธุ์และแทนที่จุดนั้นด้วยต้นไม้ย่อยต้นใหม่



ภาพที่ 2-13 การกลายพันธุ์

2.4 ความรู้เกี่ยวกับค่าความคลาดเคลื่อนของค่าพยากรณ์

ตัววัดความคลาดเคลื่อนที่ใช้ในงานวิจัย ประกอบด้วย

2.4.1 Root Mean Squared Error (RMSE)

รากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง คือ การวัดค่าความแตกต่างระหว่างค่าจริงและค่าที่ประมาณได้จากแบบจำลอง หาก RMSE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง ดังนั้นหากนี้มีค่าเท่ากับศูนย์แล้ว จะหมายความว่าไม่เกิดความคลาดเคลื่อนในแบบจำลองนี้ RMSE คำนวณได้ดังสมการนี้

$$\text{RMSE} = \sqrt{\frac{\sum(a_i - p_i)^2}{n}}$$

โดย a_i คือ ค่าจริง

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

2.4.2 Relative Absolute Error (RAE)

ค่าความคลาดเคลื่อนสัมบูรณ์ คือ ค่าปริมาณความแตกต่างระหว่างค่าความจริงกับค่าที่วัดได้ มีสูตรคำนวณได้ดังนี้

$$\text{RAE} = \frac{|P_1 - a_1| + \cdots + |P_n - a_n|}{|a_1 - \bar{a}| + \cdots + |a_n - \bar{a}|}$$

โดย a_i คือ ค่าจริง

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

$\bar{a}$ คือ ค่าเฉลี่ยค่าจริง

2.4.3 Mean Absolute Percentage Error (MAPE)

ค่าเฉลี่ยของค่าสัมบูรณ์ของเปอร์เซ็นต์ของความคลาดเคลื่อน คือ วิธีวัดความแม่นยำโดยคำนวณเปอร์เซ็นต์ความผิดพลาดในการพยากรณ์ โดยไม่คำนึงถึงเครื่องหมาย ค่าที่ได้ต่ำ มีความแม่นยำสูง เช่น ถ้า MAPE = 4% แสดงว่าวิธีที่เลือกมีความคลาดเคลื่อนในการพยากรณ์เฉลี่ยอยู่ที่ร้อยละ 4 โดยมีสูตรในการคำนวณ ดังนี้

$$\text{MAPE} = \left(\frac{100}{n}\right) \sum \left|\frac{a_i - p_i}{a_i}\right|$$

โดย a_i คือ ค่าจริง

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

2.5 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เกียรติศักดิ์ จันทร์แก้ว (การประยุกต์ใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้นเพื่อการพยากรณ์ปริมาณการส่งออกข้าวจ้าวนาปี) ได้สร้างโมเดลการพยากรณ์ปริมาณการส่งออกข้าวจ้าวนาปีของประเทศโดยใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Multilayer Perceptron Polynomial Artificial Neural network) ทำการเปรียบเทียบกับโครงข่ายประสาทแบบธรรมดา ใช้ข้อมูลอนุกรมเวลาของกรมสถิติแห่งชาติ ตั้งแต่ปี พ.ศ. 2544 – 2550 ผู้วิจัยได้ทำการแบ่งข้อมูลออกเป็น 2 กระบวนการ คือ การเรียนรู้ชุดข้อมูล (Training) และการทดสอบ (Testing) ทำการสอนให้เกิดการเรียนรู้ด้วยเทคนิคการแพร่กระจายย้อนกลับ ฝึกสอนด้วยวิธี Levenberg-Maquardt algorithm (trainlm) และฟังก์ชันการปรับการเรียนรู้ (Learning Function) แบบ LearnD (Grad Descent) โดยใช้ค่าเฉลี่ยกำลังสองสัมบูรณ์ (Mean Square Error: MSE) ของผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมทั้งสองชนิดเป็นตัวชี้วัด ผลการวิจัยพบว่าโครงสร้างที่เหมาะสมที่สุดที่ใช้ในการพยากรณ์ปริมาณการส่งออกข้าวจ้าวนาปีคือ 10-10-1 ฟังก์ชันกระตุ้นคือ logsig tansig purelin โครงข่ายประสาทเทียมให้ค่าความผิดพลาดในการเรียนรู้ชุดข้อมูลเท่ากับ 0.0033 และค่าความผิดพลาดในการทดสอบ (Testing) เท่ากับ 0.0297 และโครงข่ายประสาทเทียมแบบพหุนามหลายชั้นให้ค่าความผิดพลาดในการเรียนรู้ชุดข้อมูลเท่ากับ 0.0010 และค่าความผิดพลาดในการทดสอบ (Testing) เท่ากับ 0.0259 ซึ่งแสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Multilayer Perceptron Polynomial Artificial Neural Network) เป็นโมเดลการพยากรณ์ที่ให้ค่าความผิดพลาดต่ำกว่า เมื่อทดสอบกับชุดข้อมูลจริง และเหมาะกับการนำไปใช้เพื่อการพยากรณ์ปริมาณการส่งออกข้าวจ้าวนาปีได้อย่างมีประสิทธิภาพ

วชิราภรณ์ แก้วมาตย์ (2556) : การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อหาตัวแบบโครงข่ายประสาทเทียมของดัชนีราคาหลักทรัพย์ของ 9 ประเทศหลัก ได้แก่ ไทย สิงคโปร์ มาเลเซีย อินโดนีเซีย ฟิลิปปินส์ สหรัฐอเมริกา อังกฤษ ญี่ปุ่น และฮ่องกง โดยใช้ข้อมูลราคาปิดรายวันของดัชนีราคาหลักทรัพย์ ระหว่างวันที่ 1 มกราคม พ.ศ.2550 ถึงวันที่ 29 มิถุนายน พ.ศ.2555 รวมจำนวนข้อมูลทั้งสิ้น 1,435 วัน เป็นข้อมูลชุดแรกสำหรับสร้างแบบจำลองโครงข่ายประสาทเทียมและข้อมูลชุดที่สองคือ วันที่ 2 กรกฎาคม พ.ศ.2555 ถึงวันที่ 30 พฤศจิกายน พ.ศ.2555 รวมประมาณ 100 วัน เพื่อทดสอบประสิทธิภาพการพยากรณ์ แล้วนำข้อมูลนี้มาสร้างรูปแบบพยากรณ์ด้วยโครงข่ายประสาทเทียมแบบหลายชั้น ได้แก่ 1 ชั้นอินพุต (input layer) 1 ชั้นซ่อนเร้น (hidden layer) และ 1 ชั้นเอาท์พุต (output layer) และให้การเรียนรู้ของโครงข่ายประสาทเทียมเป็นแบบแพร่ย้อนกลับ

ผลการศึกษาพบว่าโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับให้ผลการพยากรณ์ที่แม่นยำโดยมีค่า MAPE ของข้อมูลชุดทดสอบ (test) ของประเทศต่างๆ ได้แก่ ไทย สิงคโปร์ มาเลเซีย อินโดนีเซีย ฟิลิปปินส์ สหรัฐอเมริกา อังกฤษ ญี่ปุ่น และฮ่องกง มีค่า MAPE 0.198, 0.2367, 0.1983, 0.4191, 0.9812, 0.7045×10^{-4}, 0.5206, 0.5157×10^{-3} และ 0.6634×10^{-5} ตามลำดับ

คมสัน สุริยะ (2548) เสนอการพยากรณ์ราคาน้ำมันดิบด้วยแบบจำลองเครือข่ายประสาทแบบแพร่ย้อนกลับ กับราคาน้ำมันดิบ Brent รายวันและทดสอบการพยากรณ์ไปข้างหน้า 34 วันทำการ ทดสอบโดยใช้แบบจำลองจำนวน 5 แบบ คือ แบบจำลองที่มีจำนวนเครือข่ายประสาทในชั้นซ่อนเร้นจำนวน 100 ตัว 157 ตัว 200 ตัว 231 ตัว และ 500 ตัว และวัดความแม่นยำของการพยากรณ์ด้วย Mean Absolute Percentage Error (MAPE) ซึ่งผลการทดลองที่ได้แสดงให้เห็นว่าการใช้จำนวนเครือข่ายประสาทในชั้นซ่อนเร้น จำนวน 200 นิวรอล ให้ผลพยากรณ์ราคาน้ำมันดิบได้ดีที่สุด มีความคลาดเคลื่อนโดยเฉลี่ย ±1.89% จากราคาจริง

สิทธิชัย เทพไพฑูรย์ เสนอแบบจำลองการทำนายแบบใหม่ที่อาศัยเจเนติคอัลกอริธึมร่วมกับโครงข่ายประสาทเทียม ซึ่งอยู่บนพื้นฐานของงานวิจัยและจากการทดลองที่ว่าในตัวโครงข่ายประสาทเทียมนั้นคุณภาพของชุดฝึกสอนมีผลต่อการทำนายมากกว่าสถาปัตยกรรม จำนวน หรือชนิดของตัวแปรที่ใส่เข้าไป ด้วยเหตุนี้เจเนติคอัลกอริธึมจึงทำหน้าที่เลือกชุดฝึกสอนที่เหมาะสมให้กับโครงข่ายประสาทเทียม

ผลการทดลองแสดงให้เห็นว่าการใช้เจเนติคอัลกอริธึมเป็นวิธีที่มีเสถียรภาพมากกว่าและทำกำไรได้มากกว่า เนื่องจากมีความแม่นยำในการทำนายของแต่ละช่วงเวลาที่การกระจายตัวที่ดีกว่า

บุษรา ลิ้มพิพัฒนางกูร ใช้วิธีการพยากรณ์โอกาสในการสำเร็จการศึกษาโดยใช้เทคนิคเจเนติคอัลกอริธึมเพื่อเลือกลักษณะ (Feature) ที่เหมาะสมก่อนนำไปใช้พยากรณ์ด้วยเครือข่ายประสาท (GA/ANN) โดยทดสอบทั้งเจเนติคอัลกอริธึมแบบวัตถุประสงค์เดียว (Single Objective Genetic Algorithm: SGA) และเจเนติคอัลกอริธึมแบบหลายวัตถุประสงค์ (Multi Objectives Genetic Algorithm: MGA) เปรียบเทียบกับเครือข่ายประสาทแบบทั่วไป (Artificial Neural Network: ANN) จากการทดลองพบว่า (GA/ANN) นั้นสามารถลดจำนวนลักษณะที่ใช้ในโมเดลได้อย่างชัดเจน โดย (SGA/ANN) สามารถลดจำนวน Feature ลดลงได้ 42% และ(MGA/ANN) สามารถลดจำนวน Feature ลดลง ได้ 45% และเทคนิคที่เสนอนั้นยังให้ประสิทธิภาพการทำนายที่ใกล้เคียงกับการใช้เครือข่ายประสาทเพียงอย่างเดียว เมื่อเปรียบเทียบระหว่างเทคนิค ANN กับ SGA/ANN และเทคนิค ANN กับ MGA/ANN โดยใช้ t-Test Paired Two Sample for Means โดย

มีค่า P-Value=0.64 และ0.98 ตามลำดับ อย่างไรก็ตามถ้าพิจารณาเปรียบเทียบในแง่ของความผันผวนของผลทำนายจะเห็นว่าเทคนิค ANN เพียงอย่างเดียวและเทคนิค MGA/ANN มีความคงตัวมากกว่าเทคนิค SGA/ANN

เมตตา คงคากูล ศึกษาเทคนิคการพยากรณ์ผลสัมฤทธิ์ทางการศึกษาของนักศึกษา จากพฤติกรรมการใช้อินเตอร์เน็ตโดยวิธีโครงข่ายประสาทเทียม ซึ่งได้นำหลักการและทฤษฎีโครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Perceptron) ใช้เทคนิควิธีการสอนด้วยอัลกอริธึมของ Levenberg-Marquardt มาประยุกต์ใช้ในการพยากรณ์ โดยกรณีศึกษาเป็นนักศึกษาระดับปริญญาตรี มหาวิทยาราชภัฏหมู่บ้านจอมบึง ซึ่งมีการเตรียมข้อมูลก่อนทำการวิเคราะห์ด้วยการตรวจสอบค่าความเชื่อมั่นของเครื่องมือด้วย Cornbach's Alpha และทำการวิเคราะห์ปัจจัยด้วยเทคนิค Factor analysis จากนั้นนำมาทำการ Normalization ด้วยฟังก์ชัน Sigmoid เพื่อให้ได้ค่านำเข้าตัวแปรอยู่ในช่วงที่ต้องการ แล้วไปวิเคราะห์ในระบบโครงข่ายประสาทเทียมแบบหลายชั้นด้วยวิธี Backpropagation ผลจากการศึกษาพบว่าโมเดลของโครงข่ายใยประสาทเทียมที่เหมาะสมที่สุดเป็นโครงข่ายใยประสาทเทียมที่มีโครงสร้างประกอบด้วย จำนวนหน่วยในชั้น รับข้อมูล - จำนวนหน่วยในชั้นแฝง - จำนวนหน่วยในชั้นแสดงผลเป็น 12-2-1 ระบบสามารถพยากรณ์ผลสัมฤทธิ์ทางการเรียนจากพฤติกรรมการใช้อินเตอร์เน็ต โดยใช้วิธีโครงข่ายประสาทเทียม พบว่าโครงข่ายใยประสาทเทียมดังกล่าวให้ค่าประสิทธิภาพการพยากรณ์เท่ากับ 58% แสดงให้เห็นว่า วิธีการโครงข่ายประสาทเทียม สามารถนำไปประยุกต์ใช้เพื่อการพยากรณ์ผลสัมฤทธิ์ทางการเรียนได้

ภรัณยา อำมฤครัตน นำเสนอการเปรียบเทียบประสิทธิภาพของแบบจำลองในการลดมิติข้อมูลและจำแนกข้อมูล ซึ่งใช้ข้อมูลตัวอย่างจากฐานข้อมูล UCI Machine Learning Database Repository ได้แก่ Ozone, Ionosphere และSonar นำมาวิเคราะห์ทำการเปรียบเทียบประสิทธิภาพโดยใช้วิธีการลดมิติข้อมูล (Dimensionality Reduction) แบบ PCA (Principal Components Analysis) และCFS (Correlation-based Feature Selection) ร่วมกับวิธี การจำแนกข้อมูลแบบโครงข่ายประสาทเทียมแบบ Multi-Layer Perceptron (MLP) เปรียบเทียบกับ ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines : SVM) การวัดประสิทธิภาพสามารถวัดได้จากความถูกต้องของการจำแนกประเภทของข้อมูลโดยนับจากค่า ความถูกต้องของการจำแนกประเภทข้อมูลที่วัดได้ ซึ่งการทดสอบแบบจำลองที่ได้จะทำการทดสอบผลบนพื้นฐานวิธี 5 - fold Cross Validation โดยผลการทดลองที่ได้ พบว่า วิธีการลดมิติข้อมูลแบบ CFS ร่วมกับวิธีการจำแนกข้อมูลแบบ MLP เข้ามาใช้ในการจำแนกข้อมูลนั้นจะมีประสิทธิภาพที่ดีกว่าการใช้โมเดลแบบอื่นๆ

เกียรติศักดิ์ จันทร์แก้ว ได้สร้างโมเดลพยากรณ์โดยใช้โครงข่ายประสาทเทียมและโครงข่ายประสาทเทียมฟังก์ชันพหุนาม ผู้วิจัยได้ทำการแบ่งข้อมูลออกเป็น *2* ขบวนการคือ การเรียนรู้ชุดข้อมูล (Training) และการทดสอบ (Testing) ทำการสอนให้เกิดการเรียนรู้ด้วยเทคนิคการแพร่กระจายย้อนกลับ ฝึกสอนด้วยวิธี Levenberg-Maquardt algorithm (trainlm) และฟังก์ชันการปรับการเรียนรู้ (Learning Function) แบบ LearnD (Grad Descent) โดยใช้ค่าเฉลี่ยกำลังสองสัมบูรณ์ (Mean Square Error: MSE) ของผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมทั้งสองชนิดเป็นตัวชี้วัด ผลการวิจัยพบว่าโครงสร้างที่ เหมาะสมที่สุดของโมเดลการพยากรณ์อนุกรมเวลาราคาปาล์มน้ำมัน ผลการวิจัยพบว่าเทคนิคโครงข่ายประสาทเทียม ให้ค่าความผิดพลาดในการเรียนรู้ (Training) มีค่าเท่ากับ 0.0330 และค่าความผิดพลาดในการทดสอบ (Testing) มีค่าเท่ากับ 0.3750 และโครงข่ายประสาทเทียมฟังก์ชันพหุนาม ให้ค่าความผิดพลาดในการเรียนรู้ (Training) มีค่าเท่ากับ 0.0190 และค่าความผิดพลาดในการทดสอบ (Testing) มี ค่าเท่ากับ 0.3690 ซึ่งแสดงให้เห็นว่าโครงข่ายประสาทเทียมฟังก์ชันพหุนามเป็นโมเดลการพยากรณ์ที่ เหมาะสมที่สุดสำหรับพยากรณ์อนุกรมเวลาราคาปาล์มน้ำมัน

บทที่ 3
วิธีดำเนินงานวิจัย

งานวิจัยนี้เป็นงานวิจัยเพื่อทดสอบความสามารถในการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์การถดถอย วิธีโครงข่ายประสาทเทียม และวิธีการโปรแกรมเชิงพันธุกรรม โดยใช้ข้อมูลปัจจัยในด้านต่าง ๆ ที่คาดว่าจะมีผลต่อราคาบิทคอยน์ ซึ่งข้อมูลที่ผู้วิจัยเลือกใช้ประกอบไปด้วย ข้อมูลอัตราการซื้อขายเงินแต่ละประเทศ ข้อมูลราคาหุ้นต่าง ๆ ข้อมูลราคาทองคำ ข้อมูลด้านพลังงานต่าง ๆ และข้อมูลราคาบิทคอยน์ย้อนหลัง ซึ่งข้อมูลที่ได้จะต้องนำมาปรับให้สามารถนำข้อมูลไปดำเนินการ เพื่อหาข้อมูลที่มีความสัมพันธ์กับราคาบิทคอยน์ ก่อนที่จะนำไปใช้พยากรณ์ราคาบิทคอยน์ เพื่อให้ได้ค่าพยากรณ์ราคาบิทคอยน์ ที่แม่นยำและใกล้เคียงที่สุด และเมื่อได้ผลลัพธ์แล้วจะต้องทำการวิเคราะห์เพื่อหาค่าความคลาดเคลื่อนที่ได้จากการพยากรณ์ราคาบิทคอยน์ โดยมีขั้นตอนในการดำเนินการวิจัยดังต่อไปนี้

3.1 การเตรียมข้อมูลที่จะนำไปใช้ในการทดลอง

3.2 การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ (Correlation)

3.3 การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติ (Feature Selection)

3.4 การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis)

3.5 การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks)

3.6 การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming)

3.7 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์

3.7.1 Root Mean Square Error (RMSE)

3.7.2 Relative Absolute Error (RAE)

3.7.3 Mean Absolute Percent Error (MAPE)

3.1 การเตรียมข้อมูลที่จะนำไปใช้ในการทดลอง

ข้อมูลที่ใช้ในงานวิจัยนี้ผู้วิจัยได้เลือกข้อมูลปัจจัยที่คาดว่าจะมีผลต่อราคาขึ้นลงของราคาบิทคอยน์ ประกอบด้วยข้อมูล 2 ส่วน คือ ข้อมูลปัจจัยด้านต่าง ๆ ได้แก่ 1. อัตราแลกเปลี่ยนเงินตราต่างประเทศ มีค่าเงินดอลลาร์สหรัฐของอเมริกา (USD) ค่าเงินยูโรของประเทศในยุโรป (EUR) ค่าเงินเยนของญี่ปุ่น (JPY) ค่าเงินดอลล่าร์ฮ่องกงของฮ่องกง (HKD) ค่าเงินหยวนของจีน และค่าเงินวอนของเกาหลี 2. ข้อมูลการซื้อขายหุ้น มี NASDAQ, INDEX, NIKKEI, DOWJONE และ HANGSENG 3. ราคาทอง มี ราคาทองแท่ง และราคาทองรูปพรรณ 4. ข้อมูลน้ำมัน มี น้ำมันดีเซล, แก๊สโซฮอล์ 91, แก๊สโซฮอล์ e20 และแก๊ส NGV ส่วนที่ 2 ได้แก่ ราคาบิทคอยน์ ย้อนหลัง 1 วัน และราคาบิทคอยน์ ย้อนหลัง 2 วัน โดยจะใช้ข้อมูลทั้งหมดย้อนหลังจำนวน 21 เดือน ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึง วันที่ 31 มีนาคม 2559 รวม 640 วัน โดยจะใช้ข้อมูลราคาปิดของแต่ล่ะวัน ซึ่งข้อมูลจะอยู่ในรูปแบบตารางคำนวณ Excel โดยมีตัวอย่างข้อมูลดังตารางที่ 3-1 และคำอธิบายตัวย่อในแต่ละคอลัมน์แสดงในตารางที่ 3-2

ตารางที่ 3-1 ตัวอย่างข้อมูลปัจจัยที่ใช้ในการพยากรณ์ราคาบิทคอยน์

Date	USD	EUR	NASDAQ	INDEX	GOLD	NGV	BTC2	BTC1	BTC
DAY1	32.73	44.78	4,458.65	1,485.75	20,350.00	10.50	18,392.50	18,891.55	21,888.60
DAY 2	32.64	44.77	4,457.73	1,491.81	20,350.00	10.50	18,891.55	21,888.60	20,236.79
DAY 3	32.63	44.66	4,485.93	1,493.21	20,300.00	10.50	21,888.60	20,236.79	22,009.22
DAY…	…..	…..	…..	…..	…..	…..	….	…..	…..
DAY549	36.28	40.07	5,007.41	1,288.02	18,300.00	13.50	17,640.57	14,665.91	12,921.35

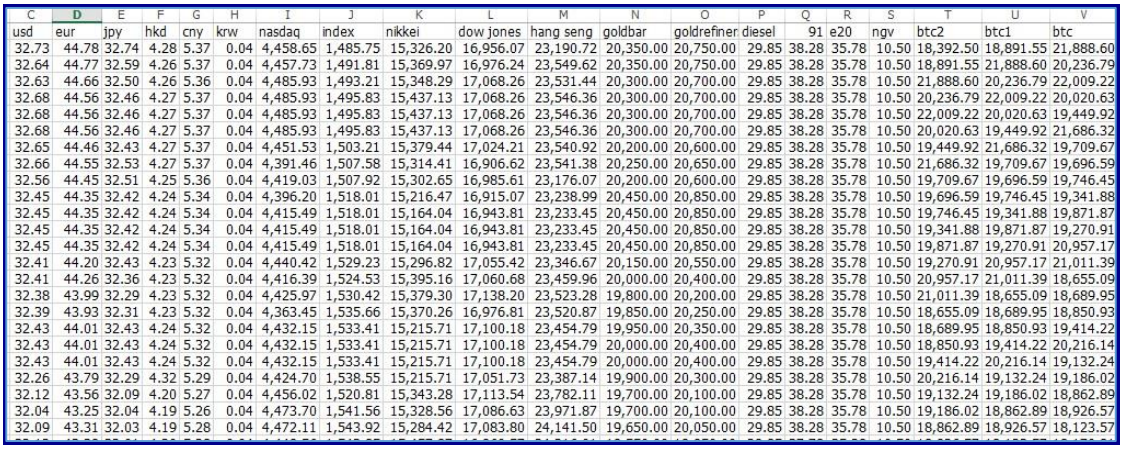
จากข้อมูลในตารางที่ 3-1 ในช่องบนสุดจะเป็นชื่อปัจจัยที่จะนำมาใช้ในการพยากรณ์ราคาบิทคอยน์ มีทั้งหมด 19 ปัจจัยหรือ 19 Attrubutes ดังที่แสดงในตารางที่ 3-2 เมื่อได้ข้อมูลมาทั้งหมดแล้วบันทึกไฟล์ในรูปแบบ .csv (Comma delimited)

ตารางที่ 3-2 ชื่อข้อมูลปัจจัยที่ใช้ในการพยากรณ์ราคาบิทคอยน์

ปัจจัย	ชื่อ
ค่าเงินดอลลาร์สหรัฐ	USD
ค่าเงินยูโร	EUR
ค่าเงินเยน	JPY
ค่าเงินฮ่องกง	HKD
ค่าเงินหยวน	CNY
ค่าเงินวอน	KRW
หุ้น NASDAQ	NASDAQ
หุ้น INDEX	INDEX
หุ้น NIKKEI	NIKKEI
หุ้น DOWNJONE	DOWNJONE
หุ้น HANGSENG	HANGSENG
ทองแท่ง	GOLDBAR
ทองรูปพรรณ	PENDANTS
น้ำมันดีเซล	DIESEL
แก๊สโซฮอล์ 91	91
แก๊สโซฮอล์ e20	E20
แก๊ส NGV	NGV
ราคา บิทคอยน์ ย้อนหลัง 1 วัน	BTC1
ราคา บิทคอยน์ ย้อนหลัง 2 วัน	BTC2

3.2 การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ (Correlation)

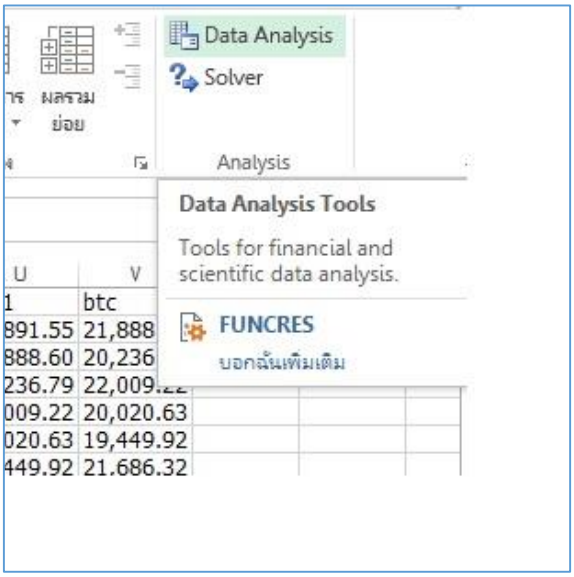
เมื่อได้ข้อมูลมาทั้งหมดแล้วนำมาปรับข้อมูลให้อยู่ในรูปแบบไฟล์ .csv แล้ว จะทำการคัดกรองข้อมูลจากข้อมูลปัจจัยทั้งหมดที่เลือกมาว่าข้อมูลปัจจัยใดมีผลหรือมีค่าเป็นไปในทิศทางเดียวกับราคาบิทคอยน์ ที่จะทำการพยากรณ์ ด้วยวิธีการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ (Correlation) โดยทางผู้วิจัยได้ใช้โปรแกรม Microsoft Excel ในการหาค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ ในครั้งนี้ซึ่งขั้นตอนจะอธิบายตามภาพที่ 3-1 ถึง 3-5



C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
usd	eur	jpy	hkd	cny	krw	nasdaq	index	nikkei	dow jones	hang seng	goldbar	goldrefiner	diesel	91	e20	ngv	btc2	btc1	btc
32.73	44.78	32.74	4.28	5.37	0.04	4,458.65	1,485.75	15,326.20	16,956.07	23,190.72	20,350.00	20,750.00	29.85	38.28	35.78	10.50	18,392.50	18,891.55	21,888.60
32.64	44.77	32.59	4.26	5.37	0.04	4,457.73	1,491.81	15,369.97	16,976.24	23,549.62	20,350.00	20,750.00	29.85	38.28	35.78	10.50	18,891.55	21,888.60	20,236.79
32.63	44.66	32.50	4.26	5.36	0.04	4,485.93	1,493.21	15,348.29	17,068.26	23,531.44	20,300.00	20,700.00	29.85	38.28	35.78	10.50	21,888.60	20,236.79	22,009.22
32.68	44.56	32.46	4.27	5.37	0.04	4,485.93	1,495.83	15,437.13	17,068.26	23,546.36	20,300.00	20,700.00	29.85	38.28	35.78	10.50	20,236.79	22,009.22	20,020.63
32.68	44.56	32.46	4.27	5.37	0.04	4,485.93	1,495.83	15,437.13	17,068.26	23,546.36	20,300.00	20,700.00	29.85	38.28	35.78	10.50	22,009.22	20,020.63	19,449.92
32.68	44.56	32.46	4.27	5.37	0.04	4,485.93	1,495.83	15,437.13	17,068.26	23,546.36	20,300.00	20,700.00	29.85	38.28	35.78	10.50	20,020.63	19,449.92	21,686.32
32.65	44.46	32.43	4.27	5.37	0.04	4,451.53	1,503.21	15,379.44	17,024.21	23,540.92	20,200.00	20,600.00	29.85	38.28	35.78	10.50	19,449.92	21,686.32	19,709.67
32.66	44.55	32.53	4.27	5.37	0.04	4,391.46	1,507.58	15,314.41	16,906.62	23,541.38	20,250.00	20,650.00	29.85	38.28	35.78	10.50	21,686.32	19,709.67	19,696.59
32.56	44.45	32.51	4.25	5.36	0.04	4,419.03	1,507.92	15,302.65	16,985.61	23,176.07	20,200.00	20,600.00	29.85	38.28	35.78	10.50	19,709.67	19,696.59	19,746.45
32.45	44.35	32.42	4.24	5.34	0.04	4,396.20	1,518.01	15,216.47	16,915.07	23,238.99	20,450.00	20,850.00	29.85	38.28	35.78	10.50	19,696.59	19,746.45	19,341.88
32.45	44.35	32.42	4.24	5.34	0.04	4,415.49	1,518.01	15,164.04	16,943.81	23,233.45	20,450.00	20,850.00	29.85	38.28	35.78	10.50	19,746.45	19,341.88	19,871.87
32.45	44.35	32.42	4.24	5.34	0.04	4,415.49	1,518.01	15,164.04	16,943.81	23,233.45	20,450.00	20,850.00	29.85	38.28	35.78	10.50	19,341.88	19,871.87	19,270.91
32.45	44.35	32.42	4.24	5.34	0.04	4,415.49	1,518.01	15,164.04	16,943.81	23,233.45	20,450.00	20,850.00	29.85	38.28	35.78	10.50	19,871.87	19,270.91	20,957.17
32.41	44.20	32.43	4.23	5.32	0.04	4,440.42	1,529.23	15,296.82	17,055.42	23,346.67	20,150.00	20,550.00	29.85	38.28	35.78	10.50	19,270.91	20,957.17	21,011.39
32.41	44.26	32.36	4.23	5.32	0.04	4,416.39	1,524.53	15,395.16	17,060.68	23,459.96	20,000.00	20,400.00	29.85	38.28	35.78	10.50	20,957.17	21,011.39	18,655.09
32.38	43.99	32.29	4.23	5.32	0.04	4,425.97	1,530.42	15,379.30	17,138.20	23,523.28	19,800.00	20,200.00	29.85	38.28	35.78	10.50	21,011.39	18,655.09	18,689.95
32.39	43.93	32.31	4.23	5.32	0.04	4,363.45	1,535.66	15,370.26	16,976.81	23,520.87	19,850.00	20,250.00	29.85	38.28	35.78	10.50	18,655.09	18,689.95	18,850.93
32.43	44.01	32.43	4.24	5.32	0.04	4,432.15	1,533.41	15,215.71	17,100.18	23,454.79	19,950.00	20,350.00	29.85	38.28	35.78	10.50	18,689.95	18,850.93	19,414.22
32.43	44.01	32.43	4.24	5.32	0.04	4,432.15	1,533.41	15,215.71	17,100.18	23,454.79	20,000.00	20,400.00	29.85	38.28	35.78	10.50	18,850.93	19,414.22	20,216.14
32.43	44.01	32.43	4.24	5.32	0.04	4,432.15	1,533.41	15,215.71	17,100.18	23,454.79	20,000.00	20,400.00	29.85	38.28	35.78	10.50	19,414.22	20,216.14	19,132.24
32.26	43.79	32.29	4.32	5.29	0.04	4,424.70	1,538.55	15,215.71	17,051.73	23,387.14	19,900.00	20,300.00	29.85	38.28	35.78	10.50	20,216.14	19,132.24	19,186.02
32.12	43.56	32.09	4.20	5.27	0.04	4,456.02	1,520.81	15,343.28	17,113.54	23,782.11	19,700.00	20,100.00	29.85	38.28	35.78	10.50	19,132.24	19,186.02	18,862.89
32.04	43.25	32.04	4.19	5.26	0.04	4,473.70	1,541.56	15,328.56	17,086.63	23,971.87	19,700.00	20,100.00	29.85	38.28	35.78	10.50	19,186.02	18,862.89	18,926.57
32.09	43.31	32.03	4.19	5.28	0.04	4,472.11	1,543.92	15,284.42	17,083.80	24,141.50	19,650.00	20,050.00	29.85	38.28	35.78	10.50	18,862.89	18,926.57	18,123.57

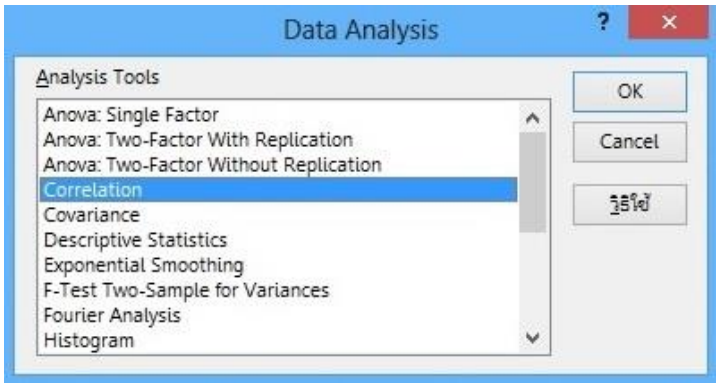
ภาพที่ 3-1 ข้อมูลในรูปแบบไฟล์ .csv ที่เตรียมไว้

เมื่อเปิดไฟล์ข้อมูลที่เตรียมไว้ทั้งหมดในรูปแบบไฟล์ .csv แล้วก็จะทำการหาค่า การวิเคราะห์สหสัมพันธ์ ด้วยโปรแกรม Microsoft Excel โดยการเลือกที่เมนู ข้อมูล จากนั้นเลือกเครื่องมือที่ชื่อ Data Analysis ดังภาพที่ 3-2



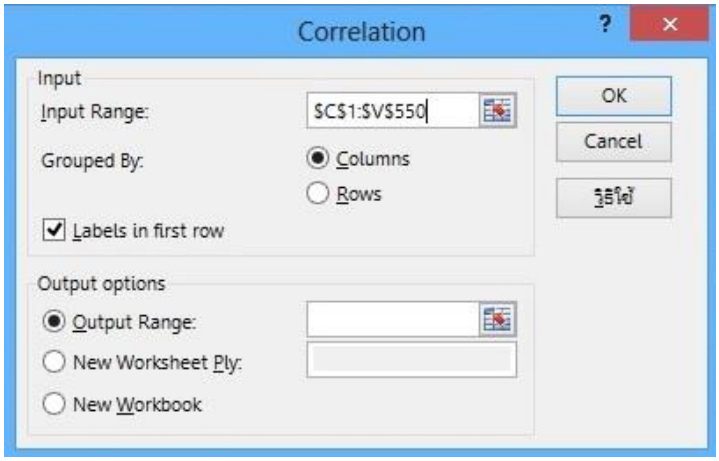
ภาพที่ 3-2 เครื่องมือชื่อ Data Analysis

เมื่อเข้ามาในเครื่องมือ Data Analysis แล้ว ทำการเลือกเครื่องมือที่จะใช้ในการวิเคราะห์ในงานวิจัยนี้จะใช้เครื่องมือที่ชื่อ Correlation ดังภาพที่ 3-3



ภาพที่ 3-3 เครื่องมือชื่อ Correlation

เมื่อเลือกเครื่องมือ Correlation แล้วจะขึ้นหน้าตาภายใน ดังภาพที่ 3-4



ภาพที่ 3-4 ภายในเครื่องมือชื่อ Correlation

ภายในเครื่องมือ Correlation ให้กำหนดค่าในช่องต่างๆ ดังนี้

- เลือกข้อมูลที่ต้องการในช่อง Input Range ในงานวิจัยนี้คือข้อมูลที่ใช้ทั้งหมด จำนวน 640 วัน 19 Attributes หรือ 19 ปัจจัย
- กำหนดให้แถวแรกเป็นชื่อใส่เครื่องหมดถูกในช่อง Labels in first row
- เลือกการแสดงผลในช่อง Output options

หลังจากกำหนดค่าเสร็จแล้วกดปุ่ม OK โปรแกรมจะทำการคำนวณหาค่าการวิเคราะห์สหสัมพันธ์ ตามค่าที่ตั้งไว้และจะปรากฏผลออกมาดังภาพที่ 3-5

	usd	eur	jpy	hkd	cny	krw	nasdaq	index	nikkei	dow jones	hang seng	goldbar	goldrefiner	diesel	91	e20	ngv	btc2	btc1	btc
usd	1																			
eur	-0.03098	1																		
jpy	0.08571	0.68176	1																	
hkd	0.94205	-0.05627	-0.06263	1																
cny	0.97947	0.01444	0.09916	0.91825	1															
krw	-0.06776	0.65757	0.65056	-0.08343	-0.04399	1														
nasdaq	0.5128	-0.69021	-0.51609	0.49979	0.48334	-0.40945	1													
index	-0.91759	-0.06836	-0.21978	-0.86159	-0.86474	-0.06603	-0.38188	1												
nikkei	0.51707	-0.73192	-0.54758	0.50165	0.49854	-0.51467	0.9453	-0.39061	1											
dow jones	-0.15285	-0.67314	-0.64254	-0.12772	-0.17486	-0.38514	0.70377	0.25639	0.60597	1										
hang seng	-0.50628	-0.44096	-0.38687	-0.47286	-0.47215	-0.06057	0.27157	0.5328	0.30222	0.56691	1									
goldbar	-0.07903	0.36496	0.3131	-0.0729	-0.09011	0.458	-0.37693	0.0658	-0.37193	-0.34048	-0.07511	1								
goldrefiners	-0.07905	0.36622	0.31332	-0.07292	-0.09007	0.45865	-0.37709	0.06613	-0.37222	-0.34029	-0.07474	0.99977	1							
diesel	-0.86514	0.41774	0.17945	-0.82272	-0.79676	0.29129	-0.70775	0.80611	-0.72079	-0.12818	0.25351	0.12694	0.1274	1						
91	-0.67759	0.63369	0.41342	-0.6543	-0.61821	0.47956	-0.75459	0.55946	-0.76495	-0.35677	0.07025	0.18205	0.18265	0.88892	1					
e20	-0.73149	0.61561	0.37851	-0.70613	-0.6607	0.467	-0.75629	0.62387	-0.76011	-0.31166	0.15307	0.17672	0.1774	0.94125	0.96433	1				
ngv	0.70214	-0.65364	-0.4473	0.68447	0.64971	-0.52968	0.79608	-0.56995	0.81926	0.3666	-0.0729	-0.20919	-0.2098	-0.88457	-0.93539	-0.95346	1			
btc2	-0.13516	0.66867	0.58593	-0.1404	-0.16098	0.63789	-0.45573	-0.08994	-0.56095	-0.33343	-0.30212	0.08362	0.08391	0.35219	0.57293	0.52534	-0.6105	1		
btc1	-0.12808	0.66384	0.58536	-0.13989	-0.15405	0.63894	-0.44769	-0.09275	-0.55578	-0.32889	-0.29841	0.08319	0.08368	0.34548	0.56805	0.51858	-0.60374	0.94162	1	
btc	**-0.1213**	**0.6636**	**0.5992**	**-0.1333**	**-0.147**	**0.6383**	**-0.4435**	**-0.0974**	**-0.5533**	**-0.3305**	**-0.3006**	**0.0846**	**0.0854**	**0.3422**	**0.5645**	**0.5148**	**-0.5973**	**0.9229**	**0.9411**	**1**

ภาพที่ 3-5 ผลลัพธ์ที่ได้จากการหาค่าการวิเคราะห์สหสัมพันธ์

จากภาพที่ 3-5 จะได้ค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลทั้งหมดโดยจะทำการเปรียบเทียบกับค่า ราคาบิทคอยน์ ปัจจุบัน ในงานวิจัยนี้คือ ค่า BTC โดยเปรียบเทียบกับข้อมูลหรือปัจจัยตัวอื่น ซึ่งการบอกระดับหรือขนาดของความสัมพันธ์ จะใช้ตัวเลขของค่าสัมประสิทธิ์สหสัมพันธ์ หากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเข้าใกล้ 1 แสดงถึงการมีความสัมพันธ์กันในระดับสูง แต่หากมีค่าเข้าใกล้ 0 แสดงถึงการมีความสัมพันธ์กันในระดับน้อยมาก หรือไม่มีเลย สำหรับการพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์โดยทั่วไปอาจใช้เกณฑ์ดังนี้ (Hinkle D. E. 1998, p.118)

0.90 - 1.00 มีความสัมพันธ์กันสูงมาก

0.70 - .90 มีความสัมพันธ์กันในระดับสูง

0.50 - .70 มีความสัมพันธ์กันในระดับปานกลาง

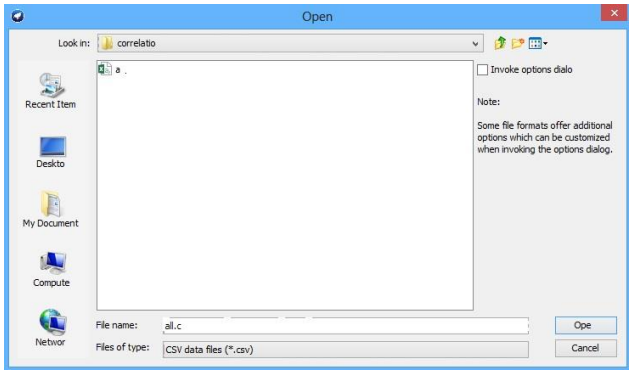
0.30 - .50 มีความสัมพันธ์กันในระดับต่ำ

0.00 - .30 มีความสัมพันธ์กันในระดับต่ำมาก

จากเกณฑ์เราจะเลือกข้อมูลที่มีค่าอยู่ใน (0.50) – (1.00) และ(-0.50) – (-1.00) ซึ่งจะได้ข้อมูลที่มีความสัมพันธ์กับข้อมูล BTC คือ EUR, JPY, KRW, NIKKIE, 91, E20, NGV, BTC1 และ BTC2 จำนวน 9 ปัจจัยหรือ 9 Attributes จากนั้นนำมาสร้างข้อมูลสำหรับใช้ในการพยากรณ์ในรูปแบบไฟล์ .csv ต่อไป

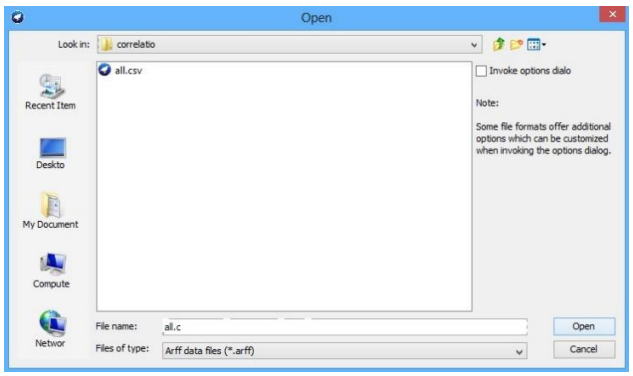
3.3 การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติ หรือ คุณลักษณะ (Feature Selection)

จากเทคนิคการหาค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ ที่ได้ข้อมูลและปัจจัยที่มีผลต่อการพยากรณ์ราคาบิทคอยน์แล้ว ทางผู้วิจัยได้ทำการหาเทคนิคการหาค่าความสัมพันธ์ของข้อมูลอีก 1 เทคนิค เพื่อใช้ในการเปรียบเทียบเพื่อให้ได้ข้อมูลการพยากรณ์ราคาบิทคอยน์ที่ใกล้เคียงที่สุด คือการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติ (Feature Selection) โดยเทคนิคนี้ทางผู้วิจัยจะหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติด้วย โปรแกรม Weka ในครั้งนี้ ดังภาพที่ 3-6



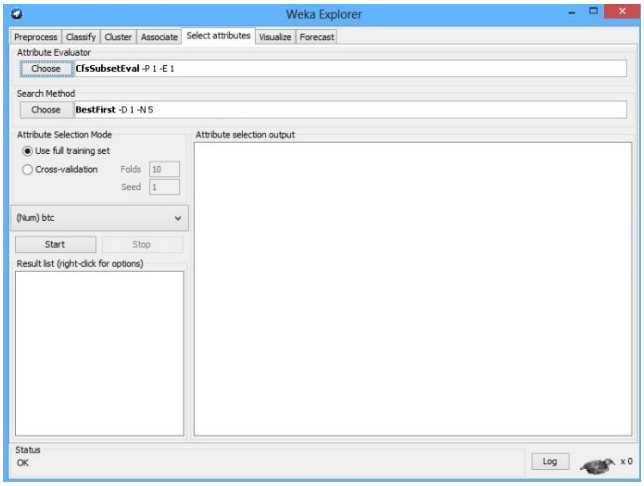
ภาพที่ 3-6 การเลือกไฟล์ csv ที่เตรียมไว้

จากนั้นกดปุ่ม Save… เพื่อบันทึกข้อมูลที่นำเข้าให้อยู่ในรูปแบบของโปรแกรม Weka ซึ่งมีหน้าสกุลไฟล์เป็น .arff โดยชื่อไฟล์จะตั้งชื่อไฟล์ให้ตรงกับไฟล์ csv ดังแสดงในภาพที่ 3-7



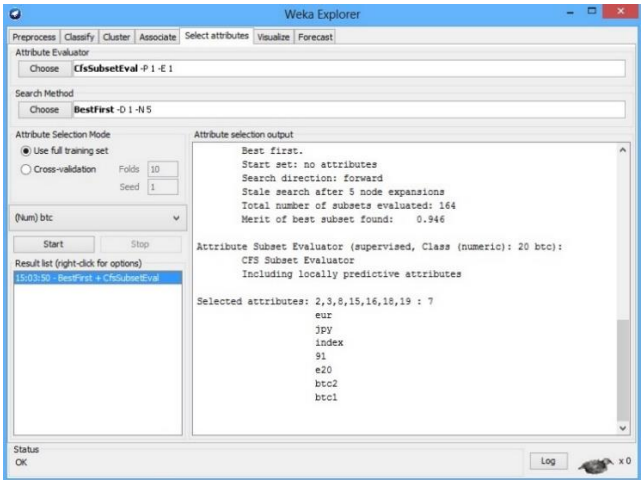
ภาพที่ 3-7 ไฟล์ .arff ที่ปรับให้อยู่ในรูปแบบโปรแกรม Weka แล้ว

เมื่อเราได้ไฟล์ .arff ที่จะนำมาใช้ในการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติด้วยโปรแกรม Weka จากนั้นให้เลือกแท็บ Select Attributes แล้วกดปุ่ม Choose ตรง Attribute Evaluator แล้วคลิกเลือกอัลกอริทึม CfsSubsetEval และกดปุ่ม Choose ตรง Search Method แล้วเลือกวิธีการค้นหา BestFirst ดังแสดงในภาพที่ 3-8



ภาพที่ 3-8 หน้าต่างเลือกอัลกอริทึมในการหาความสัมพันธ์

จากนั้นเลือกอัลกอริทึม CfsSubsetEval และเลือกวิธีการค้นหา BestFirst เสร็จแล้วให้ทำการกดปุ่ม Start โปรแกรมจะรันข้อมูลจนได้ผลลัพธ์ ดังแสดงในภาพที่ 3-9



ภาพที่ 3-9 หน้าต่างแสดงผลลัพธ์

จากผลลัพธ์ที่ได้จากการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติด้วยโปรแกรม Weka ซึ่งทางผู้วิจัยต้องการทราบว่า ปัจจัยที่ผู้วิจัยนำมาใช้ในการทดลองทั้งหมด มีปัจจัยอะไรบ้างที่มีความสัมพันธ์กับ ข้อมูลราคาบิทคอยน์ ซึ่งผู้วิจัยจะใช้ตัวย่อว่า BTC ซึ่งผลลัพธ์ที่ได้ หรือ ปัจจัยข้อมูล คือ EUR, JPY, INDEX, 91, E20, BTC1 และBTC2 จำนวน 7 ปัจจัยหรือ 7 Attributes จากนั้นนำมาสร้างข้อมูลสำหรับใช้ในการพยากรณ์ในรูปแบบไฟล์ .csv ต่อไป

เมื่อได้ข้อมูลที่ผ่านการทดลองจาก การหาความสัมพันธ์ของข้อมูลด้วยเทคนิค Correlation และการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติเราจะได้ข้อมูลที่จะนำไปใช้ในการทดลองจำนวน 2 ชุดจากที่กล่าวมาแล้วข้างต้น เพื่อนำมาใช้พยากรณ์ราคาบิทคอยน์ โดยจะมีการแบ่งข้อมูลเป็น 2 ส่วน คือ ส่วนที่ใช้ในการทดลอง ใช้สร้างโมเดล จำนวน 549 วันหรือ 18 เดือน แบ่งข้อมูลเป็นอัตราส่วนคือ 80 ต่อ 20 และ90 ต่อ 10 จากนั้นทำการสร้างชุดข้อมูลในรูปแบบไฟล์ .csv ซึ่งจะได้ไฟล์ข้อมูลจำนวน 8 ชุด สำหรับวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ เพื่อใช้เป็นอัตราการเรียนรู้และการตรวจสอบค่าพยากรณ์ราคาบิทคอยน์ ข้อมูลส่วนที่ 2 คือส่วนที่จะนำมาใช้ทดสอบการพยากรณ์ราคาบิทคอยน์ ว่ามีประสิทธิภาพแค่ไหน โดยจะใช้ข้อมูลจำนวน 91 วัน หรือ 3 เดือน สำหรับ วิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ เมื่อเตรียมข้อมูลเสร็จแล้วก็จะนำข้อมูลมาทำการทดลองการพยากรณ์ราคาบิทคอยน์ ด้วย การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks) และการพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming) ส่วนวิธีการวิเคราะห์การถดถอย (Regression Analysis) จะใช้ข้อมูลจำนวน 549 วัน หรือ 18 เดือน สำหรับ วิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ เพื่อใช้ทดลองและจะใช้ข้อมูลจำนวน 91 วัน หรือ 3 เดือน เพื่อทดสอบการพยากรณ์ราคาบิทคอยน์ ดังตารางที่ 3-3 และ3-4

ตารางที่ 3-3 ตัวอย่างข้อมูลตามปัจจัยที่ได้จากวิธีการวิเคราะห์สหสัมพันธ์

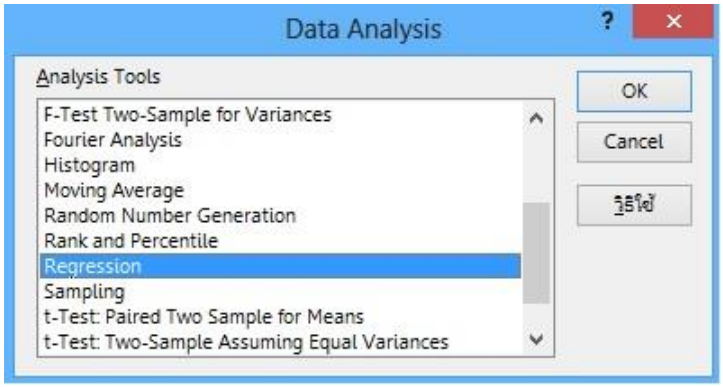
eur	jpy	krw	nikkei	91	e20	ngv	btc2	btc1	btc
44.78	32.74	0.04	15326.2	38.28	35.78	10.5	18392.5	18891.55	21888.6
44.77	32.59	0.04	15369.97	38.28	35.78	10.5	18891.55	21888.6	20236.79
44.66	32.5	0.04	15348.29	38.28	35.78	10.5	21888.6	20236.79	22009.22
44.56	32.46	0.04	15437.13	38.28	35.78	10.5	20236.79	22009.22	20020.63
44.56	32.46	0.04	15437.13	38.28	35.78	10.5	22009.22	20020.63	19449.92
44.56	32.46	0.04	15437.13	38.28	35.78	10.5	20020.63	19449.92	21686.32
44.46	32.43	0.04	15379.44	38.28	35.78	10.5	19449.92	21686.32	19709.67
44.55	32.53	0.04	15314.41	38.28	35.78	10.5	21686.32	19709.67	19696.59
44.45	32.51	0.04	15302.65	38.28	35.78	10.5	19709.67	19696.59	19746.45

ตารางที่ 3-4 ตัวอย่างข้อมูลตามปัจจัยที่ได้จากวิธีการคัดเลือกคุณสมบัติ

eur	jpy	index	91	e20	btc2	btc1	btc
44.78	32.74	1485.75	38.28	35.78	18392.5	18891.55	21888.6
44.77	32.59	1491.81	38.28	35.78	18891.55	21888.6	20236.79
44.66	32.5	1493.21	38.28	35.78	21888.6	20236.79	22009.22
44.56	32.46	1495.83	38.28	35.78	20236.79	22009.22	20020.63
44.56	32.46	1495.83	38.28	35.78	22009.22	20020.63	19449.92
44.56	32.46	1495.83	38.28	35.78	20020.63	19449.92	21686.32
44.46	32.43	1503.21	38.28	35.78	19449.92	21686.32	19709.67
44.55	32.53	1507.58	38.28	35.78	21686.32	19709.67	19696.59
44.45	32.51	1507.92	38.28	35.78	19709.67	19696.59	19746.45

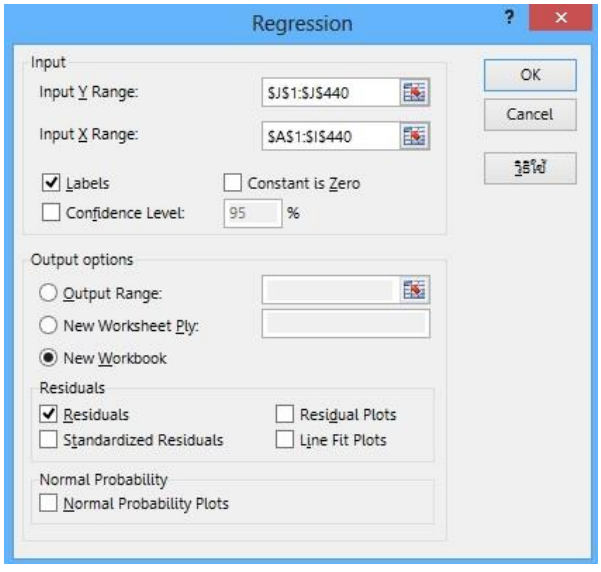
3.4 การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis)

เมื่อได้ข้อมูลแล้วก็จะเริ่มทำการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis) โดยวิธีการนี้ทางผู้วิจัยได้ใช้เครื่องมือ คือ โปรแกรม Microsoft Excel ในการทดลอง จากนั้นเปิดข้อมูลไฟล์ที่เตรียมไว้สำหรับการพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย ขึ้นมาก่อน ด้วยโปรแกรม Microsoft Excel และเลือกที่เมนู ข้อมูล จากนั้นเลือกเครื่องมือที่ชื่อ Data Analysis และเลือกเครื่องมือที่ชื่อ Regression ตามภาพที่ 3-10



ภาพที่ 3-10 เครื่องมือชื่อ Regression

เมื่อเลือกเครื่องมือ Regression แล้วจะขึ้นหน้าตาภายใน ดังภาพที่ 3-11



ภาพที่ 3-11 ภายในเครื่องมือชื่อ Regression

ภายในเครื่องมือ Regression ให้กำหนดค่าในช่องต่างๆ ดังนี้

- เลือกข้อมูลที่ต้องการพยากรณ์ในช่อง Input Y Range ในงานวิจัยนี้คือข้อมูลราคาบิทคอยน์ โดยผู้วิจัยจะใช้แทนด้วยชื่อ BTC
- เลือกข้อมูลที่ต้องการในช่อง Input X Range ในงานวิจัยนี้คือข้อมูลที่มีความสัมพันธ์กับราคาบิทคอยน์ ทั้งหมด
- กำหนดให้แถวแรกเป็นชื่อใส่เครื่องหมดถูกในช่อง Labels
- เลือกการแสดงผลในช่อง Output options
- เลือก Residuals เพื่อให้แสดงข้อมูลที่ได้จากการพยากรณ์ราคาบิทคอยน์

หลังจากกำหนดค่าเสร็จแล้วกดปุ่ม OK โปรแกรมจะทำการคำนวณหาค่า Regression ตามค่าที่ตั้งไว้และจะปรากฏผลออกมาดังภาพที่ 3-12

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.96516404							
5	R Square	0.93154163							
6	Adjusted R Squ	0.93042978							
7	Standard Error	942.305842							
8	Observations	439							
9									
10	ANOVA								
11		df	SS	MS	F	Significance F			
12	Regression	7	5207589610	7.44E+08	837.8281	1.90E-246			
13	Residual	431	382702269.4	887940.3					
14	Total	438	5590291880						
15									
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
17	Intercept	-1695.906	2214.530179	-0.76581	0.444209	-6048.5281	2656.71617	-6048.52809	2656.716171
18	eur	-8.2121867	36.18091678	-0.22698	0.82055	-79.325175	62.9008017	-79.3251752	62.90080171
19	jpy	77.7665657	34.77902266	2.236019	0.025862	9.40897676	146.124155	9.40897676	146.1241546
20	index	-1.3233003	1.068137752	-1.23889	0.216062	-3.4227072	0.77610666	-3.42270721	0.77610666
21	91	33.0124691	35.85152409	0.920811	0.357664	-37.453103	103.478042	-37.4531035	103.4780417
22	e20	92.5946468	50.13126373	1.847044	0.065426	-5.9375159	191.12681	-5.93751591	191.1268095
23	btc2	0.36730487	0.044333803	8.284984	1.51E-15	0.28016751	0.45444222	0.28016751	0.454442215
24	btc1	0.45612887	0.044230982	10.31243	1.96E-22	0.36919362	0.54306413	0.36919362	0.543064131
25									

ภาพที่ 3-12 ผลลัพธ์จากไฟล์ทดลองที่ได้จากการหาค่าการวิเคราะห์การถดถอย

จากภาพที่ 3-12 เราจะได้ค่าต่างๆจากการทดลองของ ไฟล์ที่เตรียมไว้สำหรับการพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย พร้อมผลลัพธ์การพยากรณ์ราคาบิทคอยน์ ซึ่งเราจะนำมาหาค่าการพยากรณ์ราคาบิทคอยน์ จากไฟล์ทดสอบที่เตรียมไว้ โดยให้เรียนรู้จากไฟล์ทดลองด้วยสมการถดถอยดังนี้

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$$

อธิบายสมการ

Y คือ ค่าที่เราจะพยากรณ์

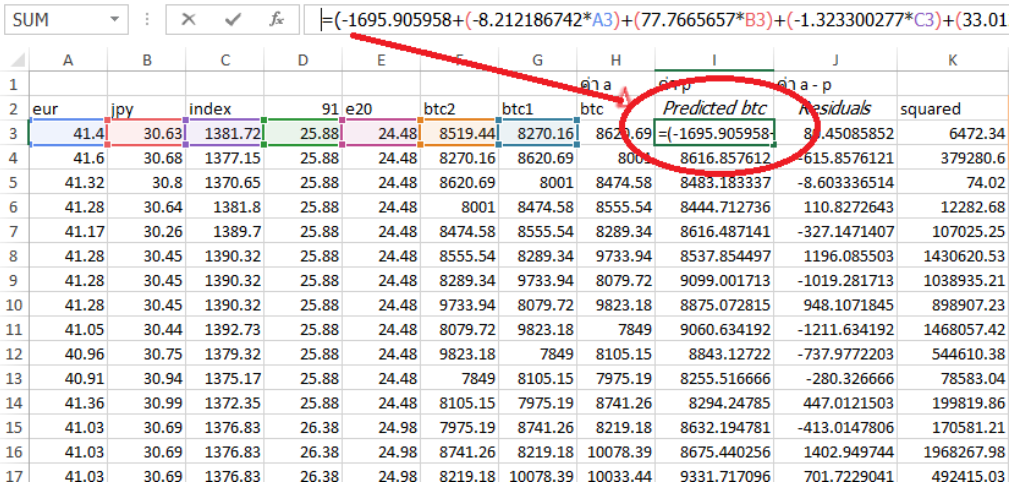
β_0 คือ ค่าคงที่

β_i คือ สัมประสิทธิ์ของแต่ละปัจจัย

x_i คือ ค่าของแต่ละปัจจัย

n คือ จำนวนปัจจัย

โดยแทนค่าเพื่อหาค่าพยากรณ์ราคาบิทคอยน์ ของ ไฟล์ทดลองจากภาพที่ 3-12 ดังนี้ Y = (-1695.905958+ (-8.212186742*A3) + (77.7665657*B3) + (-1.323300277*C3) + (33.01246914*D3) + (92.59464677*E3) + (0.367304865*F3) + (0.456128874*G3)) ของแต่ละวัน ก็จะได้ค่าพยากรณ์ราคาบิทคอยน์ ของวันนั้นดังภาพที่ 3-13



SUM | =(-1695.905958+(-8.212186742*A3)+(77.7665657*B3)+(-1.323300277*C3)+(33.01

	A	B	C	D	E	F	G	H	I	J	K
1								ค่า a	ค่า p	ค่า a - p	
2	eur	jpy	index	91	e20	btc2	btc1	btc	*Predicted btc*	*Residuals*	squared
3	41.4	30.63	1381.72	25.88	24.48	8519.44	8270.16	8620.69	=(-1695.905958+	4.45085852	6472.34
4	41.6	30.68	1377.15	25.88	24.48	8270.16	8620.69	8001	8616.857612	-615.8576121	379280.6
5	41.32	30.8	1370.65	25.88	24.48	8620.69	8001	8474.58	8483.183337	-8.603336514	74.02
6	41.28	30.64	1381.8	25.88	24.48	8001	8474.58	8555.54	8444.712736	110.8272643	12282.68
7	41.17	30.26	1389.7	25.88	24.48	8474.58	8555.54	8289.34	8616.487141	-327.1471407	107025.25
8	41.28	30.45	1390.32	25.88	24.48	8555.54	8289.34	9733.94	8537.854497	1196.085503	1430620.53
9	41.28	30.45	1390.32	25.88	24.48	8289.34	9733.94	8079.72	9099.001713	-1019.281713	1038935.21
10	41.28	30.45	1390.32	25.88	24.48	9733.94	8079.72	9823.18	8875.072815	948.1071845	898907.23
11	41.05	30.44	1392.73	25.88	24.48	8079.72	9823.18	7849	9060.634192	-1211.634192	1468057.42
12	40.96	30.75	1379.32	25.88	24.48	9823.18	7849	8105.15	8843.12722	-737.9772203	544610.38
13	40.91	30.94	1375.17	25.88	24.48	7849	8105.15	7975.19	8255.516666	-280.326666	78583.04
14	41.36	30.99	1372.35	25.88	24.48	8105.15	7975.19	8741.26	8294.24785	447.0121503	199819.86
15	41.03	30.69	1376.83	26.38	24.98	7975.19	8741.26	8219.18	8632.194781	-413.0147806	170581.21
16	41.03	30.69	1376.83	26.38	24.98	8741.26	8219.18	10078.39	8675.440256	1402.949744	1968267.98
17	41.03	30.69	1376.83	26.38	24.98	8219.18	10078.39	10033.44	9331.717096	701.7229041	492415.03

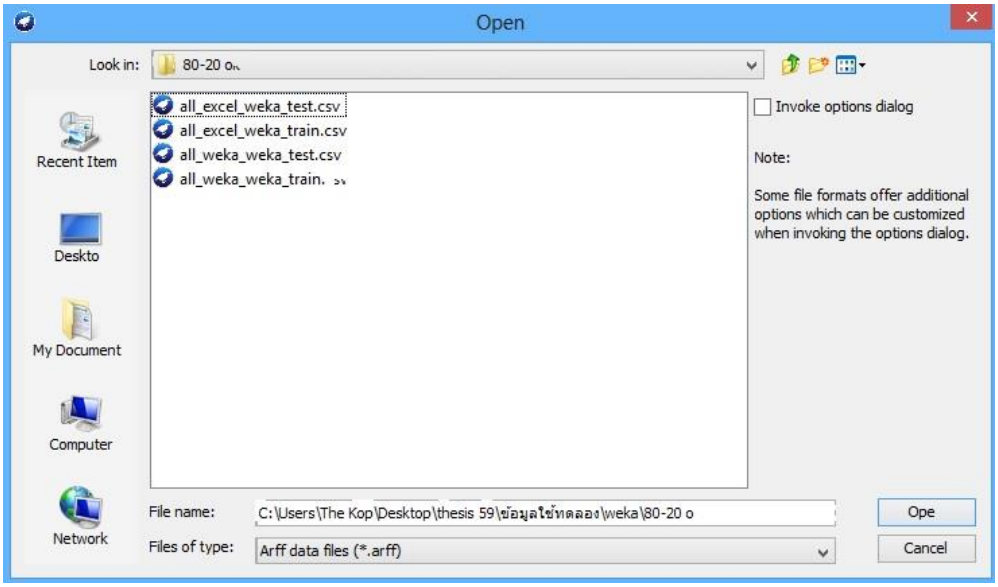
ภาพที่ 3-13 ผลลัพธ์จากไฟล์ทดสอบที่ได้จากการหาค่าการวิเคราะห์การถดถอย

3.5 การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks)

การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks) เทคนิคนี้ทางผู้วิจัยได้ใช้โปรแกรม Weka ในการทดลองด้วยเทคนิคนี้ การเตรียมข้อมูลสำหรับใช้ทดลองด้วยโปรแกรม Weka ต้องเป็นไฟล์ .csv ซึ่งผู้วิจัยได้เตรียมไว้แล้ว เมื่อได้ข้อมูลมาแล้ว ก็จะนำไฟล์เข้าสู่

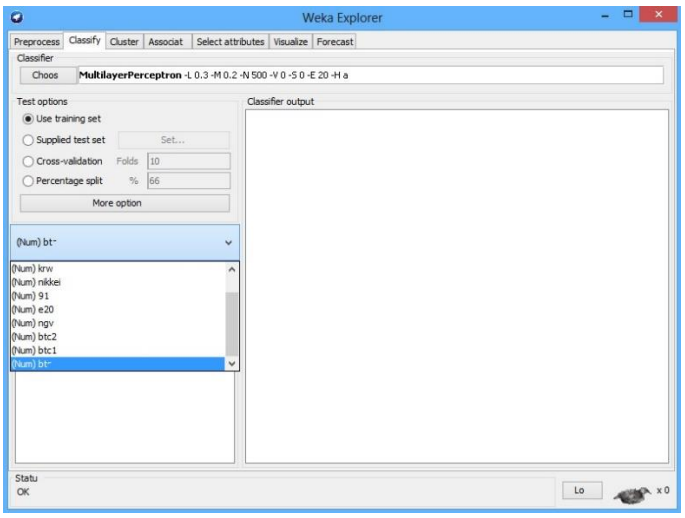
โปรแกรม Weka ซึ่งการทดลองนี้จะแบ่งข้อมูลทั้งหมดออกเป็น 2 ส่วน คือ ส่วนทดลอง และส่วนทดสอบ โดยส่วนทดลอง จะแบ่งข้อมูลออกเป็น 2 ชุด ตามอัตราส่วนของข้อมูล 80 ต่อ 20 ของข้อมูลส่วนที่ 1 และ90 ต่อ 10 ของข้อมูลส่วนที่ 1 เหมือนกัน ส่วนข้อมูลส่วนที่ 2 จะใช้สำหรับทดสอบ ซึ่งมีขั้นตอนดังนี้

1. เปิดโปรแกรม Weka แล้วเลือกปุ่ม Explorer จะปรากฏหน้าต่าง Weka Explorer จากนั้นคลิกไปที่ปุ่ม Open file … จะมีหน้าต่างเพื่อเลือกไฟล์ ให้เลือกชนิดไฟล์ที่ต้องการเป็น .csv และทำการเลือกไฟล์ที่ได้จัดเตรียมข้อมูลไว้ก่อนแล้ว ดังภาพที่ 3-14



ภาพที่ 3-14 การเลือกไฟล์ .csv ที่เตรียมไว้

2. ให้เลือกไฟล์สำหรับทดลอง ซึ่งจะมีไฟล์ .csv จำนวน 2 ไฟล์ คือ ไฟล์ Train และTest ให้เราเปิด ไฟล์ Train ขึ้นมาก่อนเพื่อให้เรียนรู้อัลกอริทึมโครงข่ายประสาทเทียม จากนั้นไปที่แถบเมนูชื่อ Classify แล้วกดปุ่ม Choose แล้วคลิกเลือกอัลกอริทึม MultilayerPerceptron ดังภาพที่ 3-15



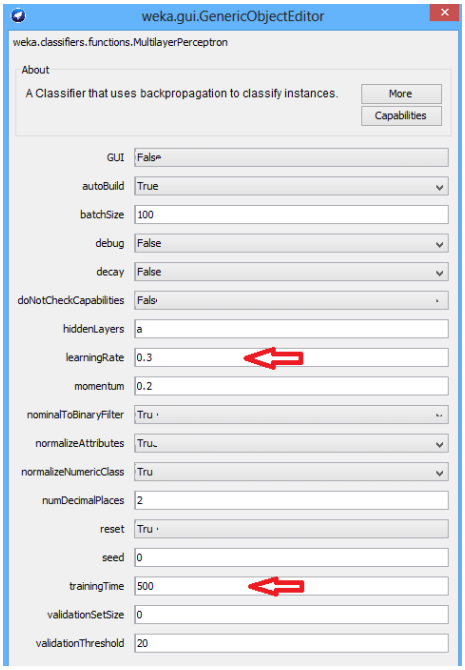
ภาพที่ 3-15 หน้าต่างเลือกอัลกอริทึมในการทดลองพยากรณ์

3. เมื่อเลือกอัลกอริทึม MultilayerPerceptron ขึ้นมาแล้ว จากไฟล์ Train ในการปรับค่าพารามิเตอร์เพื่อใช้ในการทดลองจะกำหนดค่าดังตารางที่ 3-5

ตารางที่ 3-5 ค่าพารามิเตอร์ที่ใช้ในการทดลองสำหรับ Weka

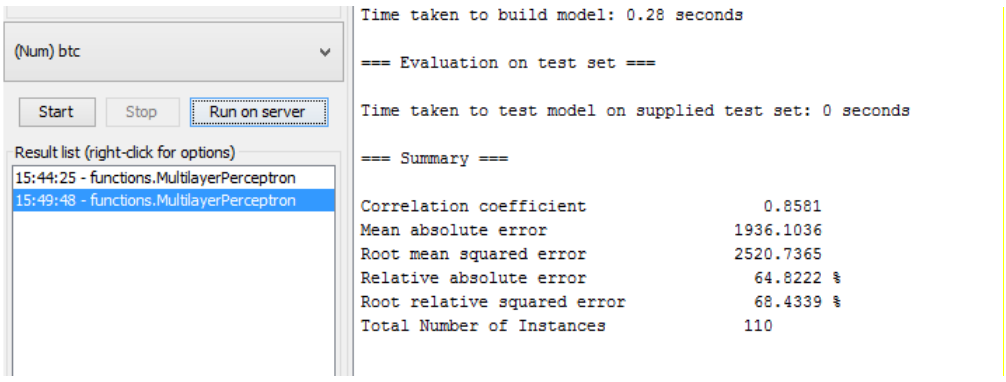
ชื่อพารามิเตอร์	learningRate	trainingTime
ค่าที่ใช้ในการทดลอง	0.3	500
		1000
		2000
		5000
	0.2	500
		1000
		2000
		5000
	0.1	500
		1000
		2000
		5000

4. บริเวณช่องด้านหลังปุ่ม Choose จะมีข้อมูลชื่ออัลกอริทึมและการกำหนดค่าเริ่มต้นให้คลิกที่ช่องดังกล่าว จะปรากฏหน้าต่างเพื่อกำหนดค่า learningRate และtrainingTime แสดงไว้ในภาพที่ 3-16



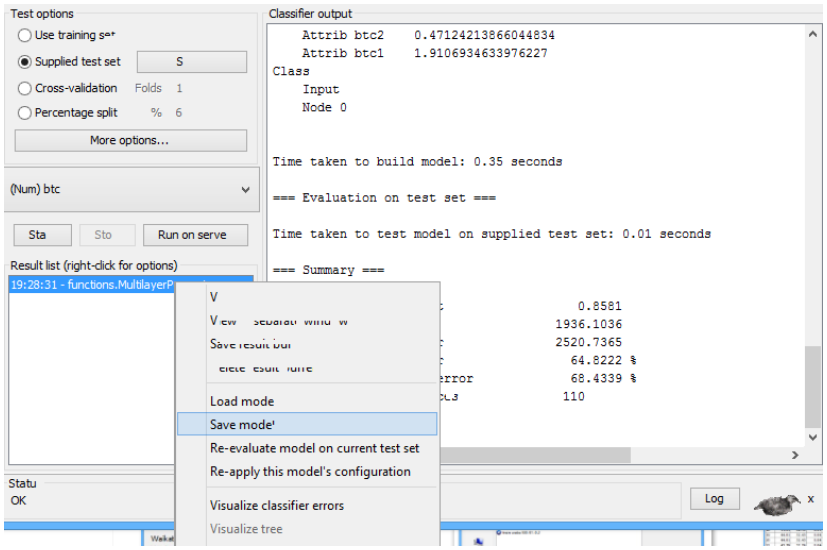
ภาพที่ 3-16 การกำหนดค่าพารามิเตอร์

5. หลังจากกำหนดค่าพารามิเตอร์เสร็จแล้วตรง Test Option ให้เลือกที่ Use training set เพื่อหาค่ากดปุ่ม OK จากนั้นคลิกที่ปุ่ม Start โปรแกรมจะทำการรันตามค่าที่ตั้งไว้และจะปรากฏผลออกมาดังภาพที่ 3-17

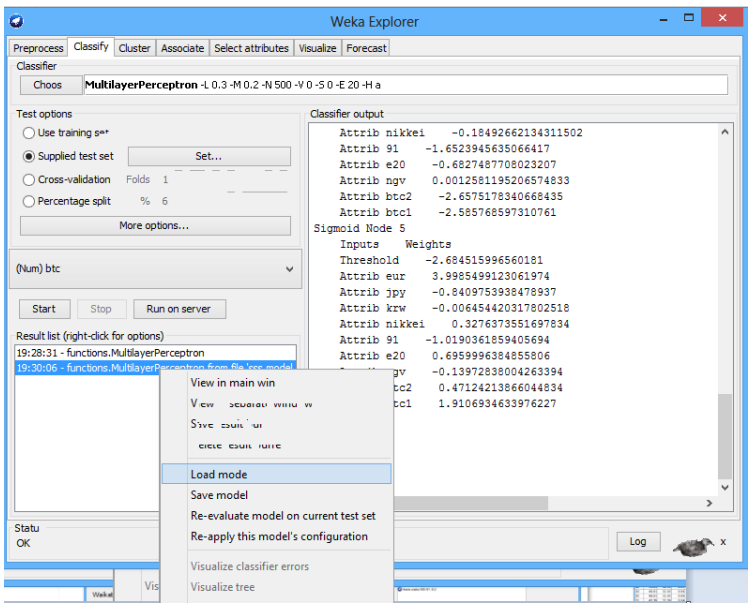


ภาพที่ 3-17 ผลลัพธ์ที่ได้จากการทดลอง

เมื่อได้ทดลองไฟล์ข้อมูลแล้วให้ทำการ Save model ดังภาพที่ 3-18 เพื่อที่เราจะนำโมเดลที่ทำการทดลองจากไฟล์ Train มาให้ไฟล์ Test ทำการเรียนรู้โมเดลเพื่อเป็นการทดสอบเพื่อให้ได้ค่าพยากรณ์ วิธีการโหลดไฟล์โมเดลเข้ามาใช้งาน โดยการคลิกขวาในส่วนของ Result list และเลือกเมนู load model ดังภาพที่ 3-19

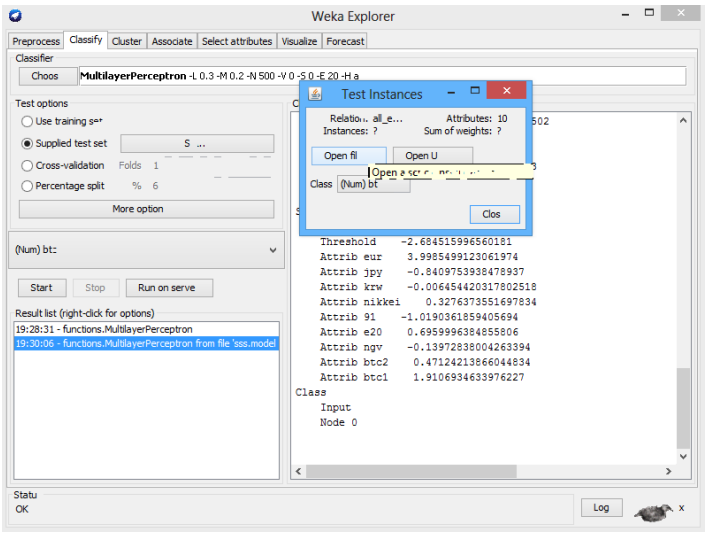


ภาพที่ 3-18 หน้าการ Save model



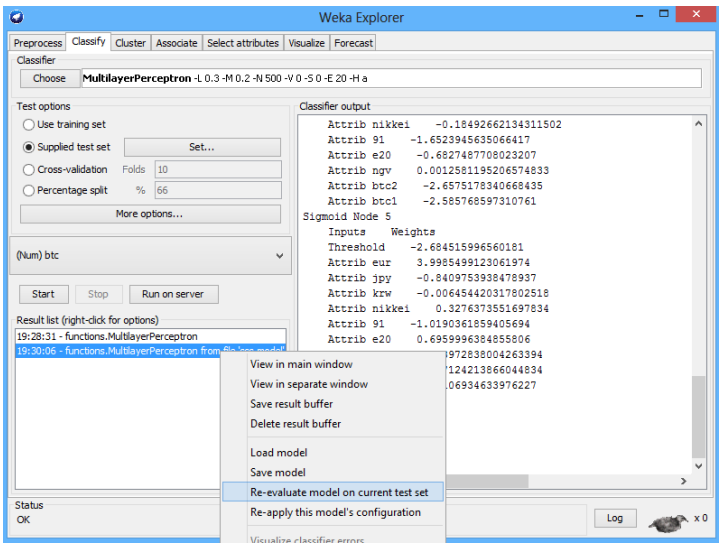
ภาพที่ 3-19 หน้าการ Load model

หลังจาก Load model เรียบร้อยแล้วก็มาเลือกที่เมนู Supplied test set และกดปุ่ม Set… จะปรากฏหน้าจอดังภาพที่ 3-20 เพื่อให้เลือกไฟล์ในส่วนที่ 2 ที่เตรียมไว้ใช้ในการทดสอบ



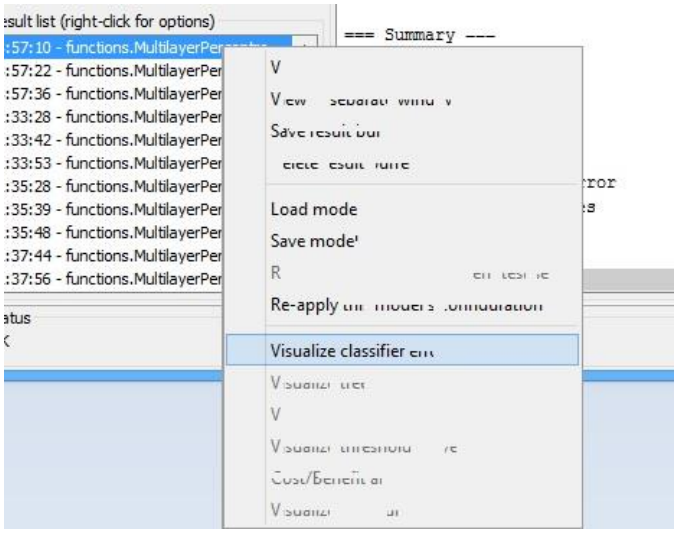
ภาพที่ 3-20 หน้าการเลือกไฟล์ที่เตรียมไว้ใช้ในการทดสอบ

จากนั้นให้คลิกขวาที่โมเดลในส่วน Result list ที่ต้องการใช้งานและเลือกเมนู Re-evaluate model on current test set เพื่อใช้โมเดลทำนายคลาสให้กับข้อมูลที่โหลดเข้ามาดังในภาพที่ 3-21

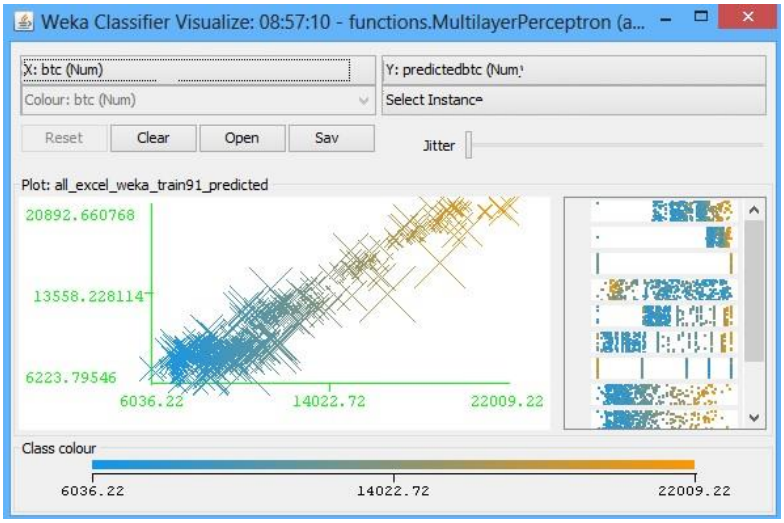


ภาพที่ 3-21 เลือกเมนู Re-evaluate model on current test set เพื่อใช้โมเดลในการทำนายข้อมูล

เมื่อได้ผลลัพธ์แล้วก็จะทำการนำข้อมูลมาหาค่าต่อไปซึ่งมีวิธีการนำข้อมูลไปใช้ ให้คลิกขวาที่โมเดลในส่วน Result list ที่ต้องการใช้งานและเลือกเมนู Visualize classifier errors ดังภาพที่ 3-22 และจะมีหน้าต่าง Weka Classifier Visualize ให้ทำการกด SAVE ไฟล์ผลลัพธ์ดังภาพที่ 3-23



ภาพที่ 3-22 เลือกเมนู Visualize classifier errors



ภาพที่ 3-23 หน้าต่าง Weka Classifier Visualize

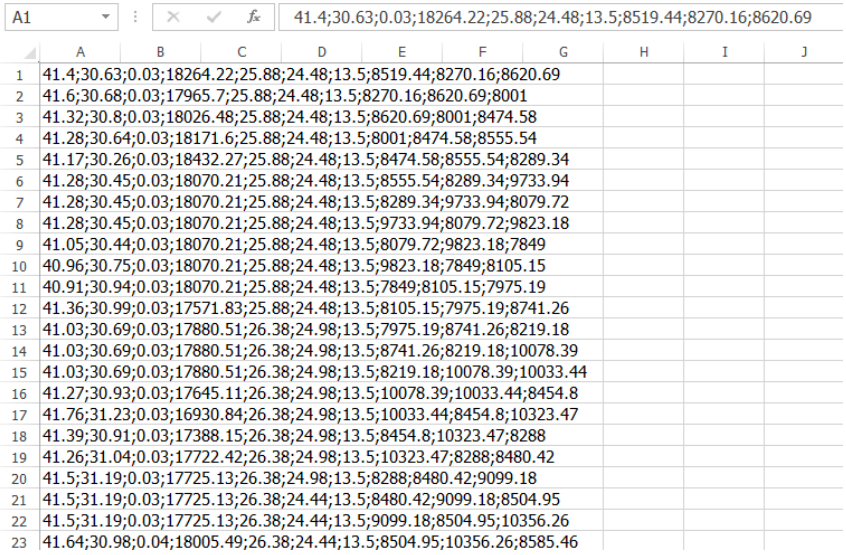
เมื่อทำการ SAVE ผลลัพธ์แล้ว ก็จะได้ผลการทดสอบหรือค่าพยากรณ์ราคาบิทคอยน์ตามที่เราทดลองจนครบทุกค่าที่ปรับค่าพารามิเตอร์ของเทคนิคการพยากรณ์ด้วยอัลกอริทึมโครงข่ายประสาทเทียม แล้วก็จะนำผลลัพธ์มาดำเนินการทดลองต่อไป

3.6 การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming)

ในการทดลองการพยากรณ์ราคาบิทคอยน์ ในงานวิจัยนี้จะใช้เทคนิคในการทดลอง 3 เทคนิคเพื่อใช้เปรียบเทียบกันว่าเทคนิคใดมีการพยากรณ์ราคาบิทคอยน์ ได้ประสิทธิหรือ ดีที่สุดจาก 3 เทคนิคนี้ที่ผู้วิจัยเลือกมาใช้

เมื่อทำการทดลองการพยากรณ์ราคาบิทคอยน์ จากเทคนิคการพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis) และเทคนิคการพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks) ไปแล้วข้างต้น ต่อไปจะเป็นวิธีการพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming) ด้วยโปรแกรม GPdotNET ซึ่งมีขั้นตอนดังนี้

1. การเตรียมข้อมูลสำหรับการทดลองด้วยโปรแกรม GPdotNET ต้องเป็นข้อมูลในรูปแบบไฟล์ .csv และเมื่อเปิดไฟล์ .csv ขึ้นมาแล้ว ข้อมูลจะต้องอยู่ในเซลเดียวกันแล้วคั่นด้วยเครื่องหมายเซมิโคลอน (“;”) ระหว่างข้อมูล ดังภาพที่ 3-24

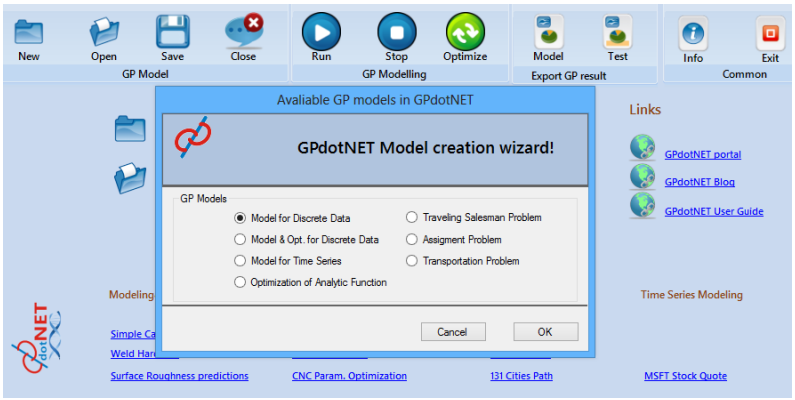


	A
1	41.4;30.63;0.03;18264.22;25.88;24.48;13.5;8519.44;8270.16;8620.69
2	41.6;30.68;0.03;17965.7;25.88;24.48;13.5;8270.16;8620.69;8001
3	41.32;30.8;0.03;18026.48;25.88;24.48;13.5;8620.69;8001;8474.58
4	41.28;30.64;0.03;18171.6;25.88;24.48;13.5;8001;8474.58;8555.54
5	41.17;30.26;0.03;18432.27;25.88;24.48;13.5;8474.58;8555.54;8289.34
6	41.28;30.45;0.03;18070.21;25.88;24.48;13.5;8555.54;8289.34;9733.94
7	41.28;30.45;0.03;18070.21;25.88;24.48;13.5;8289.34;9733.94;8079.72
8	41.28;30.45;0.03;18070.21;25.88;24.48;13.5;9733.94;8079.72;9823.18
9	41.05;30.44;0.03;18070.21;25.88;24.48;13.5;8079.72;9823.18;7849
10	40.96;30.75;0.03;18070.21;25.88;24.48;13.5;9823.18;7849;8105.15
11	40.91;30.94;0.03;18070.21;25.88;24.48;13.5;7849;8105.15;7975.19
12	41.36;30.99;0.03;17571.83;25.88;24.48;13.5;8105.15;7975.19;8741.26
13	41.03;30.69;0.03;17880.51;26.38;24.98;13.5;7975.19;8741.26;8219.18
14	41.03;30.69;0.03;17880.51;26.38;24.98;13.5;8741.26;8219.18;10078.39
15	41.03;30.69;0.03;17880.51;26.38;24.98;13.5;8219.18;10078.39;10033.44
16	41.27;30.93;0.03;17645.11;26.38;24.98;13.5;10078.39;10033.44;8454.8
17	41.76;31.23;0.03;16930.84;26.38;24.98;13.5;10033.44;8454.8;10323.47
18	41.39;30.91;0.03;17388.15;26.38;24.98;13.5;8454.8;10323.47;8288
19	41.26;31.04;0.03;17722.42;26.38;24.98;13.5;10323.47;8288;8480.42
20	41.5;31.19;0.03;17725.13;26.38;24.98;13.5;8288;8480.42;9099.18
21	41.5;31.19;0.03;17725.13;26.38;24.44;13.5;8480.42;9099.18;8504.95
22	41.5;31.19;0.03;17725.13;26.38;24.44;13.5;9099.18;8504.95;10356.26
23	41.64;30.98;0.04;18005.49;26.38;24.44;13.5;8504.95;10356.26;8585.46

ภาพที่ 3-24 ข้อมูลสำหรับโปรแกรม GPdotNET

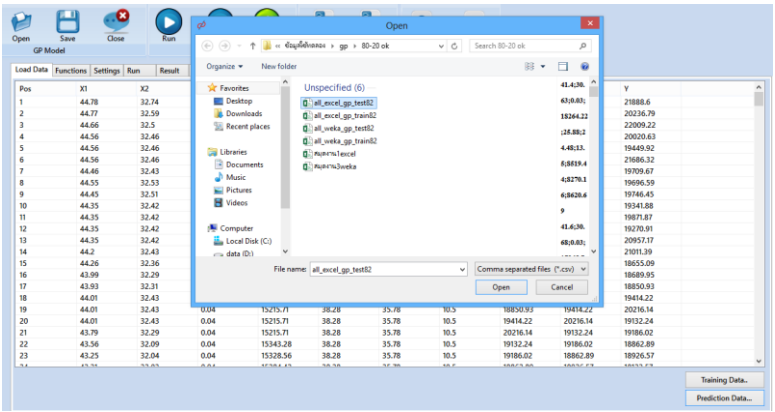
2. เมื่อได้ไฟล์ .csv สำหรับใช้ทดลองกับโปรแกรม GPdotNET แล้ว ก็จะเตรียม ไฟล์ .csv เพื่อใช้ในการทดลองซึ่งจะใช้ข้อมูลทั้งหมดจำนวน 640 วัน หรือ 21 เดือน โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือ ส่วนที่ใช้สำหรับทดลอง และส่วนที่ใช้สำหรับทดสอบ โดยข้อมูลในส่วนที่ใช้ทดลองนั้นก็จะแบ่งข้อมูลออกเป็น 2 ชุด ตามอัตราส่วนของข้อมูล 80 ต่อ 20 และ90 ต่อ 10

ซึ่งจะได้ ไฟล์สำหรับส่วนที่ 1 ใช้สำหรับทดลอง 8 ไฟล์ และไฟล์สำหรับส่วนที่ 2 อีก 2 ไฟล์ เมื่อได้ไฟล์ทั้งหมดแล้วก็จะนำเข้าข้อมูลสำหรับโปรแกรม GPdotNET โดยเปิดโปรแกรม GPdotNET ขึ้นมาให้เลือกเมนู NEW จะได้หน้าต่างขึ้นดังภาพที่ 3-25



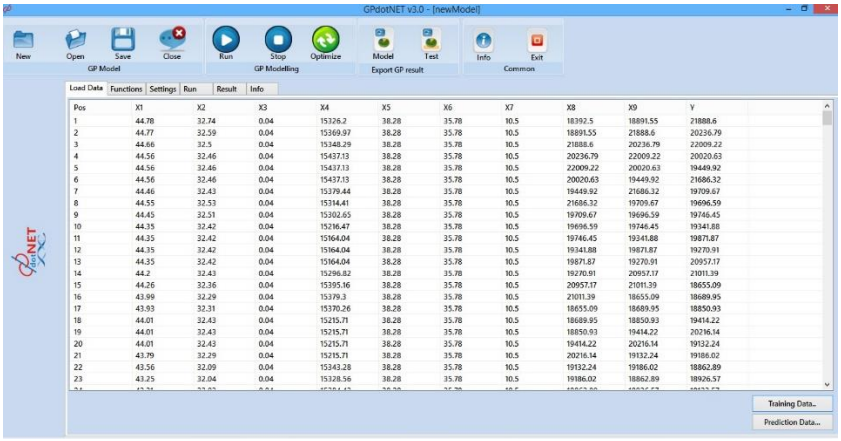
ภาพที่ 3-25 หน้าต่างเริ่มโปรแกรม GPdotNET

เมื่อขึ้นหน้าต่าง NEW ขึ้นมาแล้ว ตรง GP Models ให้เลือกหัวข้อ Model for Discrete Data แล้วกด OK ก็จะเข้าหน้าการนำเข้าข้อมูลที่เราเตรียมไว้สู้โปรแกรม ดังภาพที่ 3-26



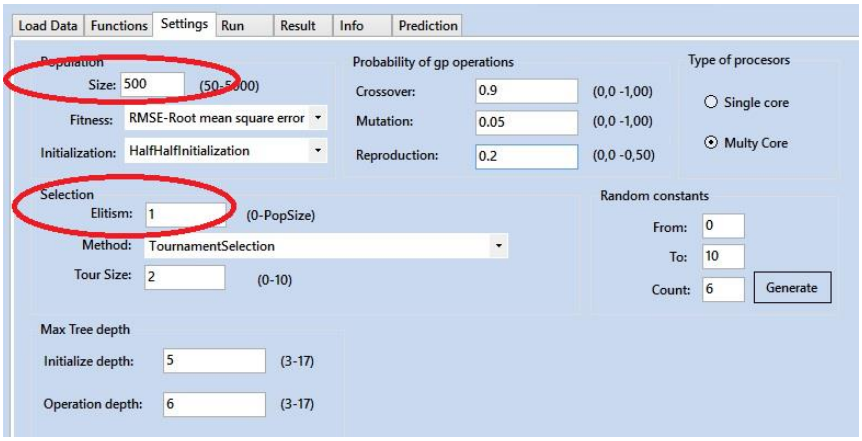
ภาพที่ 3-26 หน้านำข้อมูลเข้าสู่โปรแกรม

เมื่อเข้ามาแล้วตรงแท็บ New Data จะให้โหลดไฟล์ที่เราเตรียมไว้ในส่วนที่ 1 จำนวน 2 ชุด ใน 1 ชุดจะประกอบด้วย ไฟล์จำนวน 2 ไฟล์ คือ ใช้สำหรับ Train และสำหรับ Test ก็ให้ใส่ข้อมูลตามนี้ คือ จากภาพที่ 2-26 ตรงมุมล่างขวา ให้เลือกไฟล์ Train ใน Training Data… และเลือกไฟล์ Test ที่ Prediction Data… เมื่อเลือกข้อมูลทั้งหมดแล้วก็จะได้ดังภาพที่ 2-27



ภาพที่ 3-27 หน้าต่างเมื่อทำการเลือกข้อมูลเสร็จแล้ว

ที่แท็บ Settings จะปรากฏหน้าการตั้งค่าพารามิเตอร์ดังภาพที่ 3-28 และสำหรับงานวิจัยนี้ผู้วิจัยได้ทำการปรับค่าที่ Population Size และSelection Elitism โดยมีการปรับค่าพารามิเตอร์ที่ใช้ในการทดลองดังตารางที่ 3-6

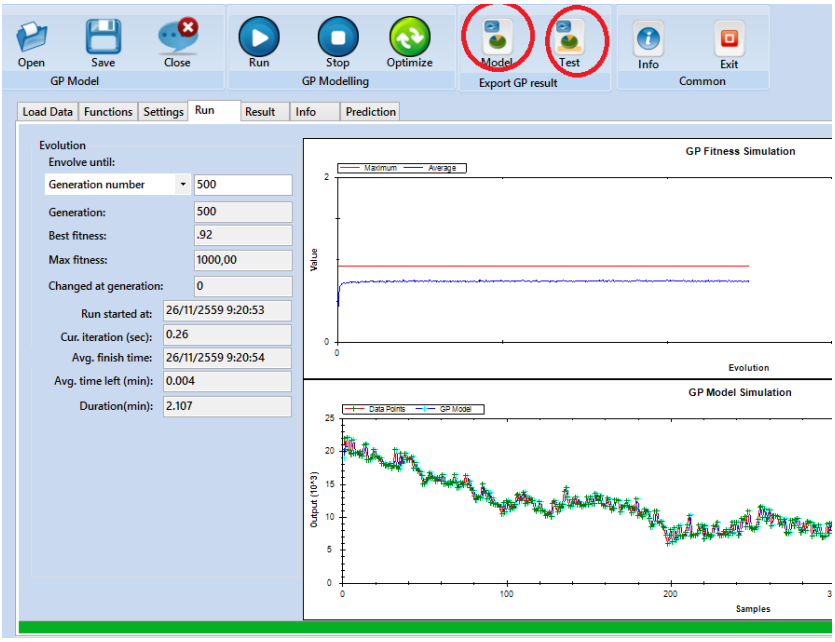


ภาพที่ 3-28 หน้าตั้งค่าพารามิเตอร์สำหรับโปรแกรม GPdotNET

ตารางที่ 3-6 ค่าพารามิเตอร์ที่ใช้ในการทดลองสำหรับ GPdotNET

Population Size	Selection Elitism
500	1
	100
	200
	300

เมื่อกำหนดพารามิเตอร์และโหลดข้อมูลเสร็จแล้วที่แท็บ Run หรือ รูปปุ่ม Run ให้กดรันข้อมูล ตามค่าที่เรากำหนดไว้ เมื่อรันเสร็จเราจะได้ข้อมูลเพื่อนำมาทำการทดลองต่อไปโดยในส่วนของ Export GP Result ก็จะได้ผลลัพธ์ที่ได้จากการทดลอง ส่วนของ Model ก็จะได้ผลลัพธ์ของไฟล์ Train และTest ก็จะเป็นผลลัพธ์ของ ไฟล์ Test ให้ทำการ SAVE ไฟล์ผลลัพธ์ทั้ง 2 เก็บไว้เพื่อที่จะนำมาใช้ทดลองต่อไปดังภาพที่ 3-29



ภาพที่ 3-29 หน้าผลลัพธ์ของ GPdotNET

เมื่อเราได้ผลลัพธ์จากทั้ง 3 เทคนิคแล้ว ก็จะนำผลลัพธ์ทั้งหมดมาวัดความคลาดเคลื่อนของข้อมูลว่าข้อมูลที่ได้มีประสิทธิภาพมากน้อยแค่ไหน ดังจะกล่าวต่อไป

3.7 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์

เมื่อได้ผลลัพธ์จากการทดลองพยากรณ์ราคาบิทคอยน์ จากวิธีการวิเคราะห์การถดถอย (Regression Analysis) เทคนิคการพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks) และเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming) ทั้ง 3 เทคนิคแล้ว เมื่อเปิดไฟล์ที่ได้จากการทดลอง ค่าที่เราจะนำมาใช้ในการหาค่าความคลาดเคลื่อนของค่าพยากรณ์ คือ ค่าจริงของราคาบิทคอยน์ หรือ ชื่อที่ผู้วิจัยตั้งไว้ คือ BTC และค่าที่เราทำการทดลองหรือค่าที่เราพยากรณ์ทางผู้วิจัยให้ชื่อ Predicted มาใช้หาค่าความคลาดเคลื่อน

3.7.1 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์ด้วย Root Mean Square Error (RMSE)

สามารถหาได้จากสมการนี้ $\text{RMSE} = \sqrt{\frac{\sum(a_i - p_i)^2}{n}}$

a_i คือ ค่าจริงของบิทคอยน์

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

3.7.2 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์ด้วย Relative Absolute Error (RAE)

สามารถหาได้จากสมการนี้ $\text{RAE} = \frac{|p_1 - a_1| + |p_2 - a_2| + \ldots + |p_n - a_n|}{|a_1 - \bar{a}| + |a_2 - \bar{a}| + \ldots + |a_n - \bar{a}|}$

a_i คือ ค่าจริงของบิทคอยน์

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

$\bar{a}$ คือ ค่าเฉลี่ยค่าจริงของบิทคอยน์

3.7.3 การหาค่าความคลาดเคลื่อนของค่าพยากรณ์ด้วย Mean Absolute Percent Error (MAPE)

สามารถหาได้จากสมการนี้ $\text{MAPE} = \left(\frac{100}{n}\right) \sum \left|\frac{a_i - p_i}{a_i}\right|$

a_i คือ ค่าจริงของบิทคอยน์

p_i คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวัน

เมื่อเอาข้อมูลที่ได้มาแทนค่าในสมการก็จะได้ค่า RMSE, RAE และMAPE ของแต่ละเทคนิคจากนั้นก็เอาผลลัพธ์ที่ได้มาเปรียบเทียบกันของแต่ละเทคนิค และดูว่าค่า RMSE, RAE และ MAPE มีค่าที่น้อยที่สุดก็จะได้เทคนิคที่มีการพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพที่สุดจากทั้ง 3 เทคนิค

บทที่ 4
ผลการดำเนินงาน

ข้อมูลราคาบิทคอยน์ที่ใช้ในการทดลองนี้ ได้รวบรวมมาจากราคาปิดของการซื้อขายในแต่ละวัน ส่วนข้อมูลของปัจจัยอื่น ๆ ที่เกี่ยวข้องรวบรวมมาจาก 4 แหล่ง คือ (1) ราคาทองคำ แบ่งข้อมูลเป็น ราคาทองคำแท่ง และราคาทองคำรูปพรรณ ได้ข้อมูลจากเว็บไซต์ทองคำราคา.com (2) ราคาพลังงาน แบ่งข้อมูลเป็น ราคาน้ำมันดีเซล ราคาแก๊สโซฮอล 91 ราคาแก๊สโซฮอล E20 และราคาแก๊ส NGV ได้ข้อมูลจาก เว็บไซต์ของบริษัท บางจากปิโตรเลียม จำกัด (มหาชน) (3) ราคาอัตราแลกเปลี่ยนเงินตราต่างประเทศเทียบกับเงินบาทไทย แบ่งข้อมูลเป็น USD คือ เงินดอลล่าร์สหรัฐ JPY คือ เงินเยนญี่ปุ่น EUR คือ เงินยูโร และHKD คือ เงินดอลล่าร์ฮ่องกง ได้ข้อมูลจากเว็บไซต์ธนาคารกรุงเทพ (4) ราคาหุ้นในตลาดหลักทรัพย์ แบ่งข้อมูลเป็น SET Index คือ ตลาดหลักทรัพย์แห่งประเทศไทย HANG SENG Index คือ ตลาดหลักทรัพย์แห่งประเทศฮ่องกง NIKKEI Index คือ ตลาดหลักทรัพย์แห่งประเทศญี่ปุ่น DOW JONES Index คือ ตลาดหลักทรัพย์ของสหรัฐอเมริกา และNASDAQ Index คือ ตลาดหลักทรัพย์ของสหรัฐอเมริกา ซึ่งมีการเก็บรวบรวมตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึง วันที่ 31 มีนาคม 2559 จำนวน 21 เดือน 640 วัน นำมาคัดเลือกเพื่อหาความสัมพันธ์ของข้อมูลเหล่านี้ที่มีความสัมพันธ์กับราคาบิทคอยน์ โดยใช้วิธีการคัดเลือก 2 วิธี คือ การหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการวิเคราะห์สหสัมพันธ์ (Correlation) และการหาความสัมพันธ์ของข้อมูลด้วยเทคนิคการคัดเลือกคุณสมบัติ (Feature Selection) ผลการหาความสัมพันธ์ของทั้ง 2 วิธีได้ผลดังตารางที่ 3-3 และ3-4 จากนั้นเมื่อเราได้ ปัจจัยมาทั้ง 2 แบบแล้วจะทำการแบ่งข้อมูล จากทั้ง 640 วัน สำหรับใช้ทดลองกับเทคนิคที่ทางผู้วิจัยเลือกมา คือ 1. การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis) 2. การพยากรณ์ด้วยอัลกอริทึมโครงข่ายประสาทเทียม (Neural Networks) และ3. การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming) โดยจะเตรียมชุดข้อมูลสำหรับเทคนิคการพยากรณ์ต่าง ๆ ดังนี้

4.1 การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย (Regression Analysis)

การพยากรณ์ด้วยวิธีการวิเคราะห์การถดถอย ข้อมูลที่จะเตรียมในการทดลองแบ่งเป็น 2 ส่วน คือ ส่วนที่ใช้ทดลอง (Train) และส่วนที่ใช้ทดสอบ (Test) ฉะนั้น ข้อมูลที่เราจะได้จะมี ข้อมูลของวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ วิธีละ 2 ไฟล์ โดยใช้ข้อมูลส่วนที่ใช้

ทดลอง จำนวน 549 วัน และข้อมูลส่วนที่ใช้ทดสอบจำนวน 91 วัน จะได้ ไฟล์ .csv สำหรับใช้ทดลองทั้งหมด 4 ไฟล์ ซึ่งได้ผลลัพธ์ดังตารางที่ 4-1 และตารางที่ 4-2

ตารางที่ 4-1 ผลลัพธ์ที่ได้จากการทดลอง (Train) ของวิธีการวิเคราะห์สหสัมพันธ์

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-3078.10	3680.65	-0.84	0.40	-10308.27	4152.08	-10308.27	4152.08
eur	28.36	37.68	0.75	0.45	-45.64	102.37	-45.64	102.37
jpy	96.11	41.32	2.33	0.02	14.93	177.28	14.93	177.28
krw	13177.97	15751.67	0.84	0.40	-17764.22	44120.16	-17764.22	44120.16
nikkei	0.03	0.06	0.53	0.59	-0.08	0.14	-0.08	0.14
91	44.46	43.11	1.03	0.30	-40.22	129.14	-40.22	129.14
e20	-44.37	57.38	-0.77	0.44	-157.09	68.35	-157.09	68.35
ngv	-60.80	193.00	-0.32	0.75	-439.92	318.32	-439.92	318.32
btc2	0.29	0.04	7.11	0.00	0.21	0.37	0.21	0.37
btc1	0.60	0.04	14.68	0.00	0.52	0.68	0.52	0.68

ตารางที่ 4-2 แสดงผลลัพธ์ที่ได้จากการทดลอง (Train) ของวิธีการคัดเลือกคุณสมบัติ

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-93.43	2505.15	-0.04	0.97	-5014.44	4827.58	-5014.44	4827.58
eur	-5.40	40.40	-0.13	0.89	-84.76	73.97	-84.76	73.97
jpy	87.18	41.05	2.12	0.03	6.53	167.82	6.53	167.82
index	-1.66	1.23	-1.35	0.18	-4.06	0.75	-4.06	0.75
91	44.80	42.80	1.05	0.30	-39.28	128.88	-39.28	128.88
e20	3.12	52.43	0.06	0.95	-99.87	106.10	-99.87	106.10
btc2	0.29	0.04	7.06	0.00	0.21	0.37	0.21	0.37
btc1	0.60	0.04	14.76	0.00	0.52	0.68	0.52	0.68

เมื่อเราได้ผลลัพธ์จาก ไฟล์ Train แล้ว ก็จะนำค่าที่ได้มาหาค่า Test จากวิธีการวิเคราะห์การถดถอยแล้วนำไปหาค่าความคลาดเคลื่อน RMSE, RAE และMAPE ได้ผลลัพธ์ดังนี้

ตารางที่ 4-3 ผลลัพธ์ที่ได้จากวิธีการวิเคราะห์การถดถอย

	RMSE		RAE		MAPE	
	train	test	train	test	train	test
วิธีการวิเคราะห์สหสัมพันธ์	1130.15	1832.63	0.3	1.04	8.14	9.13
วิธีการคัดเลือกคุณสมบัติ	1129.41	1831.89	0.3	1.04	8.12	9.10

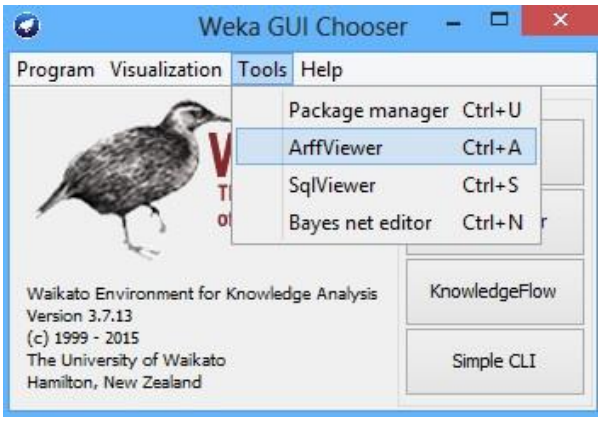
จากตารางที่ 4-3 จะสรุปได้ว่า การทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์การถดถอย และได้ค่าความสัมพันธ์จากวิธีการวิเคราะห์สหสัมพันธ์ มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1130.15, RAE เท่ากับ 0.3 และMAPE เท่ากับ 8.14 และข้อมูลที่ได้จากการหาความสัมพันธ์ด้วยวิธีการคัดเลือกคุณสมบัติมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1129.41, RAE เท่ากับ 0.3 และ MAPE เท่ากับ 8.12 เมื่อได้ค่า train มาแล้ว ก็นำค่า train ที่ได้จากโมเดลมาทอสอบเพื่อให้ได้ค่า test และนำไปหาค่าความคลาดเคลื่อนผลลัพธ์ที่ได้ คือ ข้อมูลวิธีการวิเคราะห์สหสัมพันธ์ จากค่า test มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1832.63, RAE เท่ากับ 1.04 และMAPE เท่ากับ 9.13 และข้อมูลวิธีการคัดเลือกคุณสมบัติ จากค่า test มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1831.89, RAE เท่ากับ 1.04 และMAPE เท่ากับ 9.10

เมื่อได้ข้อมูลที่ได้จากค่าความคาดเคลื่อน ของวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติแล้ว เราจะพิจารณาจากค่าที่ น้อยที่สุด หรือ ค่าที่ ใกล้ศูนย์ที่สุด จะเป็นค่าที่ได้ค่าพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพที่สุดโดยดูจากค่าความคลาดเคลื่อนทั้ง 3 วิธี จะสรุปได้ดังนี้โดยพิจารณาค่า test เทคนิคการวิเคราะห์การถดถอย ที่ใช้ วิธีการคัดเลือกคุณสมบัติมีค่า RMSE เท่ากับ 1831.89 น้อยกว่า ข้อมูลที่ใช้วิธีการวิเคราะห์สหสัมพันธ์ ซึ่งมีค่าเท่ากับ 1832.63 พิจารณาค่า RAE ข้อมูลวิธีการคัดเลือกคุณสมบัติและวิธีการวิเคราะห์สหสัมพันธ์มีค่าเท่ากันที่ 1.04 และเมื่อพิจารณา ค่า MAPE ข้อมูลวิธีการคัดเลือกคุณสมบัติมีค่าเท่ากับ 9.10 ซึ่งน้อยกว่า ข้อมูลที่ใช้วิธีการวิเคราะห์สหสัมพันธ์ ซึ่งมีค่าเท่ากับ 9.13 จะเห็นได้ว่า เทคนิคการวิเคราะห์การถดถอย ที่ใช้วิธีการหาความสัมพันธ์ของข้อมูลด้วยวิธีการคัดเลือกคุณสมบัติจะมีค่าความคลาดเคลื่อนที่มีประสิทธิภาพใกล้เคียงกับข้อมูลที่ใช้วิธีการหาความสัมพันธ์ด้วยวิธีการวิเคราะห์สหสัมพันธ์

4.2 การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks)

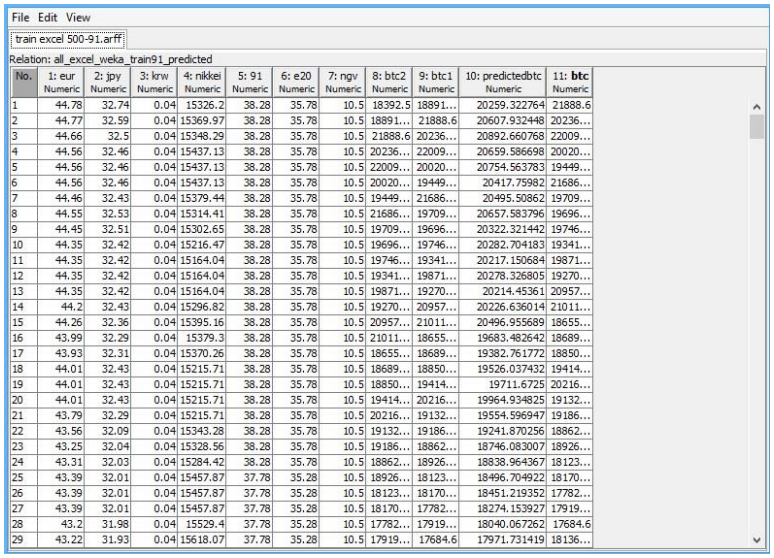
การพยากรณ์ด้วยอัลกอริทึม โครงข่ายประสาทเทียม ข้อมูลที่ใช้ทั้งหมด 640 วัน แบ่งเป็น 2 ส่วน คือ ส่วนที่ใช้ทดลอง (Train) และส่วนที่ใช้ทดสอบ (Test) ตามผลการหาความสัมพันธ์ด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ โดยข้อมูลที่จะใช้ทดลองจะแบ่งเป็น 2 ชุด ตามอัตราส่วน 80 ต่อ 20 และ90 ต่อ 10 จะใช้ข้อมูลจำนวน 549 วัน และส่วนที่ใช้ทดสอบข้อมูลจะใช้ข้อมูลตามผลการหาความสัมพันธ์ด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ แต่จะใช้ข้อมูลชุดเดียว จำนวน 91 วัน ซึ่งจะได้ไฟล์ .csv ทั้งหมด 12 ไฟล์

เมื่อได้ไฟล์ข้อมูลที่จะใช้ทำการทดลองทั้งหมดแล้ว ก็เข้าสู่การทดลองการพยากรณ์ราคาบิทคอยน์ ด้วย อัลกอริทึม โครงข่ายประสาทเทียม เมื่อได้ผลการทดลองจาก โปรแกรม Weka ที่กล่าวมาในบทที่ 3 แล้ว เราจะได้ ไฟล์ผลลัพธ์เป็น ไฟล์ .arff ซึ่งมีวิธีเปิดโดยเปิดโปรแกรม Weka ขึ้นมาดังภาพที่ 4-1



ภาพที่ 4-1 โปรแกรม Weka

เมื่อเปิดโปรแกรม Weka ขึ้นมาแล้ว ไปที่เมนู Tools เลือก ArffViewer จะขึ้นหน้าต่างให้เราเลือกไฟล์ .arff ที่เราต้องการเปิดขึ้นมาเมื่อเลือกไฟล์ .arff ที่เราต้องการแล้ว ให้กด OK จะได้ข้อมูลการพยากรณ์ราคาบิทคอยน์ ด้วยวิธี โครงข่ายประสาทเทียม ดังภาพที่ 4-2



File Edit View

train excel 500-91.arff

Relation: all_excel_weka_train91_predicted

No.	1: eur Numeric	2: jpy Numeric	3: krw Numeric	4: nikkei Numeric	5: 91 Numeric	6: e20 Numeric	7: ngv Numeric	8: btc2 Numeric	9: btc1 Numeric	10: predictedbtc Numeric	11: btc Numeric
1	44.78	32.74	0.04	15326.2	38.28	35.78	10.5	18392.5	18891...	20259.322764	21888.6
2	44.77	32.59	0.04	15369.97	38.28	35.78	10.5	18891...	21888.6	20607.932448	20236...
3	44.66	32.5	0.04	15348.29	38.28	35.78	10.5	21888.6	20236...	20892.660768	22009...
4	44.56	32.46	0.04	15437.13	38.28	35.78	10.5	20236...	22009...	20659.586698	20020...
5	44.56	32.46	0.04	15437.13	38.28	35.78	10.5	22009...	20020...	20754.563783	19449...
6	44.56	32.46	0.04	15437.13	38.28	35.78	10.5	20020...	19449...	20417.75982	21686...
7	44.46	32.43	0.04	15379.44	38.28	35.78	10.5	19449...	21686...	20495.50862	19709...
8	44.55	32.53	0.04	15314.41	38.28	35.78	10.5	21686...	19709...	20657.583796	19696...
9	44.45	32.51	0.04	15302.65	38.28	35.78	10.5	19709...	19696...	20322.321442	19746...
10	44.35	32.42	0.04	15216.47	38.28	35.78	10.5	19696...	19746...	20282.704183	19341...
11	44.35	32.42	0.04	15164.04	38.28	35.78	10.5	19746...	19341...	20217.150684	19871...
12	44.35	32.42	0.04	15164.04	38.28	35.78	10.5	19341...	19871...	20278.326805	19270...
13	44.35	32.42	0.04	15164.04	38.28	35.78	10.5	19871...	19270...	20214.45361	20957...
14	44.2	32.43	0.04	15296.82	38.28	35.78	10.5	19270...	20957...	20226.636014	21011...
15	44.26	32.36	0.04	15395.16	38.28	35.78	10.5	20957...	21011...	20496.955689	18655...
16	43.99	32.29	0.04	15379.3	38.28	35.78	10.5	21011...	18655...	19683.482642	18689...
17	43.93	32.31	0.04	15370.26	38.28	35.78	10.5	18655...	18689...	19382.761772	18850...
18	44.01	32.43	0.04	15215.71	38.28	35.78	10.5	18689...	18850...	19526.037432	19414...
19	44.01	32.43	0.04	15215.71	38.28	35.78	10.5	18850...	19414...	19711.6725	20216...
20	44.01	32.43	0.04	15215.71	38.28	35.78	10.5	19414...	20216...	19964.934825	19132...
21	43.79	32.29	0.04	15215.71	38.28	35.78	10.5	20216...	19132...	19554.596947	19186...
22	43.56	32.09	0.04	15343.28	38.28	35.78	10.5	19132...	19186...	19241.870256	18862...
23	43.25	32.04	0.04	15328.56	38.28	35.78	10.5	19186...	18862...	18746.083007	18926...
24	43.31	32.03	0.04	15284.42	38.28	35.78	10.5	18862...	18926...	18838.964367	18123...
25	43.39	32.01	0.04	15457.87	37.78	35.28	10.5	18926...	18123...	18496.704922	18170...
26	43.39	32.01	0.04	15457.87	37.78	35.28	10.5	18123...	18170...	18451.219352	17782...
27	43.39	32.01	0.04	15457.87	37.78	35.28	10.5	18170...	17782...	18274.153927	17919...
28	43.2	31.98	0.04	15529.4	37.78	35.28	10.5	17782...	17919...	18040.067262	17684.6
29	43.22	31.93	0.04	15618.07	37.78	35.28	10.5	17919...	17684.6	17971.731419	18136...

ภาพที่ 4-2 ข้อมูลที่ได้จากไฟล์ .arff

เมื่อเปิดไฟล์ .arff แล้วให้ทำการ SAVE เป็น ไฟล์ .csv เพื่อที่จะได้นำข้อมูลไปหาค่าความคลาดเคลื่อน RMSE, RAE และMAPE ซึ่งผลลัพธ์ที่ได้ทำการทดลองจากไฟล์ข้อมูลทั้งหมดได้ผลลัพธ์ดังนี้

ตารางที่ 4-4 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.3

learningRate	0.3								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	1033.18	2520.74	4066.7	8.9	13.91	23.16	0.3	0.67	2.67
1000	1023.74	3320.81		8.78	18.96		0.29	0.91	
2000	1010.5	4315.92		8.67	24.49		0.29	1.2	
5000	1036.2	6842.67		9	33.72		0.3	1.73	

จากตารางที่ 4-4 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-4 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบใน

การทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 2520.74 ค่า RAE เท่ากับ 0.67 และค่า MAPE เท่ากับ 13.91 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 4066.7 ค่า RAE เท่ากับ 2.67 และค่า MAPE เท่ากับ 23.16

ตารางที่ 4-5 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.2

learningRate	0.2								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	1023.07	1852.42	**3254.43**	8.82	11.5	**17.36**	0.3	0.5	**2.03**
1000	1023.08	1990.32		8.86	11.81		0.3	0.54	
2000	1015.85	2580.05		8.78	14.31		0.29	0.69	
5000	989.61	7187.89		8.45	37.32		0.28	1.89	

จากตารางที่ 4-5 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-5 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1852.42 ค่า RAE เท่ากับ 0.5 และค่า MAPE เท่ากับ 11.5 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3254.43 ค่า RAE เท่ากับ 2.03 และค่า MAPE เท่ากับ 17.36

ตารางที่ 4-6 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.1

learningRate	0.1								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	1058.69	1776.47	**3029.2**	9.14	11.6	**15.83**	0.31	0.5	**1.86**
1000	1025.76	2111.7		8.87	12.23		0.3	0.56	
2000	1014.79	2377.54		8.81	13.35		0.3	0.63	
5000	1004.23	3221.28		8.66	17.45		0.29	0.86	

จากตารางที่ 4-6 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-6 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1776.47 ค่า RAE เท่ากับ 0.5และค่า MAPE เท่ากับ 11.6 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3029.2 ค่า RAE เท่ากับ 1.86 และค่า MAPE เท่ากับ 15.83

สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ดังตารางที่ 4-7

ตารางที่ 4-7 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20

learningRate	RMSE	RAE	MAPE
0.3	4066.65	2.67	23.16
0.2	3254.43	2.03	17.36
0.1	3029.2	1.86	15.83

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-7 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์

สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 คือ โมเดลที่มีค่า learningRate 0.1 และtrainingTime 500 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3029.2 ค่า RAE เท่ากับ 1.86 และค่า MAPE เท่ากับ 15.83

ตารางที่ 4-8 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.3

learningRate	0.3								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	983.1	5198.5	**9406.32**	8.4	27.93	**59.45**	0.28	1.4	**6.65**
1000	981.22	8314.49		8.39	43.28		0.28	2.19	
2000	993.57	8983.04		8.59	49.47		0.29	2.47	
5000	1008.85	5258.41		8.76	29.64		0.29	1.47	

จากตารางที่ 4-8 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-8 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 5189.5 ค่า RAE เท่ากับ 1.4 และค่า MAPE เท่ากับ 27.93 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 9406.32 ค่า RAE เท่ากับ 6.65 และค่า MAPE เท่ากับ 59.45

ตารางที่ 4-9 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.2

learningRate	0.2								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	1028.72	2882.94	5903.76	8.82	15.62	36.03	0.3	0.76	4.08
1000	978.43	8514.53		8.42	44.74		0.28	2.27	
2000	968.5	10708.07		8.36	57.76		0.28	2.9	
5000	974.3	6485.42		8.41	36.49		0.28	1.81	

จากตารางที่ 4-9 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-9 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 2882.94 ค่า RAE เท่ากับ 0.76 และค่า MAPE เท่ากับ 15.62 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 5903.76 ค่า RAE เท่ากับ 4.08 และค่า MAPE เท่ากับ 36.03

ตารางที่ 4-10 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ learningRate 0.1

learningRate	0.1								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	1036.11	1790.86	3879.72	8.9	11.1	21.91	0.3	0.49	2.53
1000	1012.28	2437.92		8.68	13.81		0.29	0.65	
2000	973.02	5746.43		8.28	32.15		0.28	1.6	
5000	977.07	6583.33		8.41	38.01		0.29	1.87	

จากตารางที่ 4-10 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-10 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบ ในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1790.86 ค่า RAE เท่ากับ 0.49 และค่า MAPE เท่ากับ 11.1 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3979.72 ค่า RAE เท่ากับ 2.53 และค่า MAPE เท่ากับ 21.91

สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ดังตารางที่ 4-11

ตารางที่ 4-11 สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20

learningRate	RMSE	RAE	MAPE
0.3	9406.32	6.65	59.45
0.2	5903.76	4.08	36.03
0.1	3879.72	2.53	21.91

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-11 โมเดลที่พยากรณ์ราคา บิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 คือ โมเดลที่มีค่า learningRate 0.1 และtrainingTime 500 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3879.72ค่า RAE เท่ากับ 2.53 และค่า MAPE เท่ากับ 21.91

จากผลการทดลองของข้อมูลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมในอัตราส่วน 80 ต่อ 20 เมื่อพิจารณาตารางสรุปผล ตารางที่ 4-7 และตารางที่ 4-11 จากวิธีการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีการคัดเลือกคุณสมบัติ พบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ในการเลือกปัจจัยได้ผลที่ใกล้เคียงและมีประสิทธิกว่าการเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติดังตารางที่ 4-12

ตารางที่ 4-12 ผลการทดลองของข้อมูลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมในอัตราส่วน 80 ต่อ 20

การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
learningRate	RMSE	RAE	MAPE	RMSE	RAE	MAPE
0.3	4066.65	2.67	23.16	9406.32	6.65	59.45
0.2	3254.43	2.03	17.36	5903.76	4.08	36.03
0.1	3029.2	1.86	15.83	3879.72	2.53	21.91

ต่อไปจะเป็นการหาผลลัพธ์การพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียมกับข้อมูลที่แบ่งอัตราส่วน 90 ต่อ 10 ได้ค่าความคลาดเคลื่อนดังนี้

ตารางที่ 4-13 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.3

learningRate	0.3								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	953.87	4173.84	**5760.13**	7.33	24.32	**34.95**	0.26	2.46	**3.96**
1000	919.94	5454.27		7.21	30.98		0.26	3.45	
2000	898.93	6162.47		7.08	34.65		0.25	4.22	
5000	883.17	7254.08		6.98	40.71		0.25	4.21	

จากตารางที่ 4-13 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-13 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบ ในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 4173.84 ค่า RAE เท่ากับ 2.46 และค่า MAPE เท่ากับ 24.32 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 5760.13 ค่า RAE เท่ากับ 3.96 และค่า MAPE เท่ากับ 34.95

ตารางที่ 4-14 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.2

learningRate	0.2								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	948.66	4324.33	**5752.46**	7.16	25.34	**34.78**	0.26	2.54	**3.95**
1000	925.4	5284.11		7.04	30.41		0.26	2.91	
2000	903.35	6449.54		6.95	36.4		0.26	3.74	
5000	884.75	7869.01		6.83	44.27		0.25	4	

จากตารางที่ 4-14 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-14 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบ ในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 4324.33 ค่า RAE เท่ากับ 2.54 และค่า MAPE เท่ากับ 25.34 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 5752.46 ค่า RAE เท่ากับ 3.95 และค่า MAPE เท่ากับ 34.78

ตารางที่ 4-15 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.1

learningRate	0.1								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	962.81	2437.81	**3190.77**	7.27	14.09	**17**	0.27	1.01	**1.99**
1000	955.54	3119.15		7.2	17.53		0.27	1.58	
2000	942.51	7510.57		7.08	42.3		0.27	3.15	
5000	876.8	9825.82		6.72	54.55		0.25	3.17	

จากตารางที่ 4-15 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-15 คือ โมเดลที่ใช้ trainingTime ที่ 500 รอบ ในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 2437.81 ค่า RAE เท่ากับ 1.01 และค่า MAPE เท่ากับ 14.09 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3190.77 ค่า RAE เท่ากับ 1.99 และค่า MAPE เท่ากับ 17.00

สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ดังตารางที่ 4-16

ตารางที่ 4-16 สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10

learningRate	RMSE	RAE	MAPE
0.3	5760.13	3.96	34.95
0.2	5752.46	3.95	34.78
0.1	3190.77	1.99	17.00

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-16 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 คือ โมเดลที่มีค่า learningRate 0.1 และtrainingTime 500 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 3190.77 ค่า RAE เท่ากับ 1.99 และค่า MAPE เท่ากับ 17.00

ตารางที่ 4-17 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.3

learningRate	0.3								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	981.02	1951.16	**1750.96**	7.62	10.34	**9.06**	0.28	0.69	**1.03**
1000	983.93	1940.21		7.63	10.78		0.28	0.7	
2000	985.02	1915.19		7.65	10.68		0.28	0.71	
5000	936.66	1836.53		7.26	10.27		0.27	0.78	

จากตารางที่ 4-17 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-17 ได้ผลลัพธ์ ค่า RMSE โมเดลที่ใช้ trainingTime ที่ 5000 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1836.53 ค่า RAE โมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน เท่ากับ 0.69 และค่า MAPE โมเดลที่ใช้ trainingTime ที่ 5000 รอบในการทดลองมีค่าความคลาดเคลื่อน เท่ากับ 10.27 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 ค่า RAE เท่ากับ 1.03 และค่า MAPE เท่ากับ 9.06

ตารางที่ 4-18 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.2

learningRate	0.2								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	952.09	1941.46	**2862.14**	7.57	10.87	**14.83**	0.29	1.09	**1.74**
1000	938.56	1915.88		7.48	10.8		0.28	1.07	
2000	935.94	1874.11		7.23	10.8		0.28	1.03	
5000	973.26	1787.73		7.38	10.19		0.29	0.98	

จากตารางที่ 4-18 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-18 คือ โมเดลที่ใช้ trainingTime ที่ 5000 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1787.73 ค่า RAE เท่ากับ 0.98 และค่า MAPE เท่ากับ 10.19 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2862.14 ค่า RAE เท่ากับ 1.74 และค่า MAPE เท่ากับ 14.83

ตารางที่ 4-19 ผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ learningRate 0.1

learningRate	0.1								
trainingTime	การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)								
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
500	962.74	1982.81	**2306.65**	7.35	11.5	**11.43**	0.28	0.89	**1.36**
1000	952.98	1926.6		7.29	11.13		0.28	0.9	
2000	948.28	1910.83		7.25	10.96		0.28	0.95	
5000	977.34	1827.07		7.31	10.56		0.28	0.94	

จากตารางที่ 4-19 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-19 คือ โมเดลที่ใช้ trainingTime ที่ 5000 รอบในการทดลองมีค่าความคลาดเคลื่อน RMSE เท่ากับ 1827.07 และค่า MAPE เท่ากับ 10.56 กับโมเดลที่ใช้ trainingTime ที่ 500 รอบในการทดลองมีค่าความคลาดเคลื่อน RAE เท่ากับ 0.89 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2306.65 ค่า RAE เท่ากับ 1.36 และค่า MAPE เท่ากับ 11.43

สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10 ดังตารางที่ 4-20

ตารางที่ 4-20 สรุปผลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10

learningRate	RMSE	RAE	MAPE
0.3	1750.96	1.03	9.06
0.2	2862.14	1.74	14.83
0.1	2306.65	1.36	11.43

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-20 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิที่สุดของการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกคุณสมบัติข้อมูลอัตราส่วน 90 ต่อ 10 คือ โมเดลที่มีค่า learningRate 0.3 และtrainingTime 5000 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 และค่า MAPE เท่ากับ 9.06 ส่วนค่า RAE เท่ากับ 1.03 ได้จากโมเดลที่มีค่า learningRate 0.1 และtrainingTime 500

จากผลการทดลองการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมของข้อมูลในอัตราส่วน 90 ต่อ 10 เมื่อพิจารณาตารางสรุปผล ตารางที่ 4-16 และตารางที่ 4-20 จากวิธีการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติ พบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ ได้ผลที่ใกล้เคียงและมีประสิทธิภาพกว่า วิธีการวิเคราะห์สหสัมพันธ์ดังตารางที่ 4-21

ตารางที่ 4-21 ผลการทดลองของข้อมูลการพยากรณ์โดยใช้วิธีโครงข่ายประสาทเทียมในอัตราส่วน 90 ต่อ 10

การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
learningRate	RMSE	RAE	MAPE	RMSE	RAE	MAPE
0.3	5760.13	3.96	34.95	1750.96	1.03	9.06
0.2	5752.46	3.95	34.78	1978.76	1.13	9.61
0.1	3190.77	1.99	17.00	2480.46	1.46	12.39

เราจะสรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยอัลกอริทึม โครงข่ายประสาทเทียม (Neural Networks) จากตารางที่ 4-22 ได้ผลลัพธ์ดังนี้

ตารางที่ 4-22 สรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียม

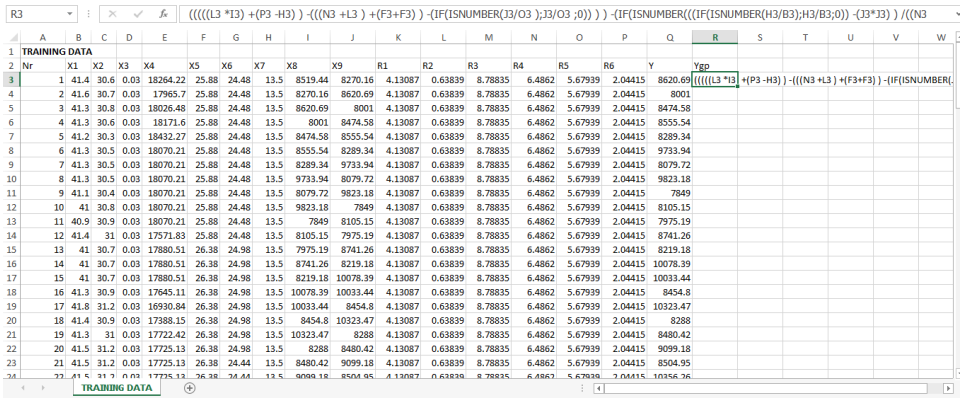
การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
ERRORS	RMSE	RAE	MAPE	RMSE	RAE	MAPE
80/20	3029.2	1.86	15.83	3879.72	2.53	21.91
90/10	3190.77	1.99	17.00	1750.96	1.03	9.06

เมื่อพิจารณาตารางที่ 4-22 จะสรุปผลได้ว่า วิธีการเลือกปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ เพื่อมาพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการคัดเลือกคุณสมบัติ ให้ค่าความคลาดเคลื่อนที่ใกล้เคียงและมีประสิทธิภาพกว่าวิธีการเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และการแบ่งอัตราส่วนข้อมูลที่ใช้ในการทดลอง ถ้าแบ่งอัตราส่วนข้อมูลที่ 90 ต่อ 10 จะได้ค่าความคลาดเคลื่อนที่ดีกว่าการแบ่งอัตราส่วนข้อมูลที่ 80 ต่อ 20 ซึ่งการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียม ที่ใกล้เคียงและมีประสิทธิที่สุดคือ โมเดลที่ใช้ learningRate 0.3 และtrainingTime 5000 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 และค่า MAPE เท่ากับ 9.06 ส่วนค่า RAE เท่ากับ 1.03 ได้จากโมเดลที่มีค่า learningRate 0.1 และtrainingTime 500

4.3 การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม (Genetic Programming)

การพยากรณ์ด้วยเทคนิคการโปรแกรมเชิงพันธุกรรม ข้อมูลที่ใช้ทั้งหมด 640 วัน แบ่งเป็น 2 ส่วน คือ ส่วนที่ใช้ทดลอง (Train) และส่วนที่ใช้ทดสอบ (Test) ตามผลการหาความสัมพันธ์ด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติโดยข้อมูลที่จะใช้ทดลองจะแบ่งเป็น 2 ชุด ตามอัตราส่วน 80 ต่อ 20 และ90 ต่อ 10 จะใช้ข้อมูลจำนวน 549 วัน และส่วนที่ใช้ทดสอบข้อมูลจะใช้ข้อมูลตามผลการหาความสัมพันธ์ด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติ แต่จะใช้ข้อมูลชุดเดียว จำนวน 91 วัน ซึ่งจะได้ไฟล์ .csv ทั้งหมด 12 ไฟล์ ซึ่งจะเหมือนกับการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีโครงข่ายประสาทเทียม

เมื่อได้ไฟล์ข้อมูลที่จะใช้ทำการทดลองทั้งหมดแล้ว ก็เข้าสู่การทดลองการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรม เมื่อได้ผลการทดลองจาก โปรแกรม GPdotNET ที่กล่าวมาในบทที่ 3 แล้ว เราจะได้ ไฟล์ผลลัพธ์มา เมื่อเปิดขึ้นมาจะได้ข้อมูลการทดลอง ดังภาพที่ 4-3



ภาพที่ 4-3 ข้อมูลที่ได้จาก GPdotNET

เมื่อเปิดไฟล์ที่ได้จากการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรมด้วยโปรแกรม GPdotNET แล้ว จากนั้นนำข้อมูลไปหาค่าความคลาดเคลื่อน RMSE, RAE และ MAPE จะได้ผลการทดลองจากไฟล์ข้อมูลทั้งหมดได้ผลลัพธ์ดังนี้

ตารางที่ 4-23 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 1

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	962.10	1800.28	10120.38	7.57	10.76	9.45	0.27	0.48	7.29
รอบที่ 2	991.14	1709.82		7.77	10.48		0.28	0.46	
รอบที่ 3	971.26	1690.90		7.60	10.29		0.27	0.46	
รอบที่ 4	972.90	1933.67		7.65	11.18		0.27	0.50	
รอบที่ 5	1068.03	1658.05		8.31	10.57		0.30	0.45	

จากตารางที่ 4-23 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-23 คือ โมเดลที่รันในรอบที่ 5 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1658.05 ค่า RAE เท่ากับ 0.45 กับโมเดลที่รันในรอบที่ 3 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความ

คลาดเคลื่อน MAPE เท่ากับ 10.29 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 10120.38 ค่า RAE เท่ากับ 7.29 และค่า MAPE เท่ากับ 9.45

ตารางที่ 4-24 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 100

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	973.07	1652.28	1994.78	7.53	10.28	9.88	0.27	0.45	1.11
รอบที่ 2	1089.16	1683.80		8.45	10.75		0.30	0.46	
รอบที่ 3	970.66	1747.64		7.62	10.71		0.27	0.47	
รอบที่ 4	1089.25	1682.85		8.46	10.74		0.30	0.46	
รอบที่ 5	971.74	1697.95		7.66	10.60		0.27	0.46	

จากตารางที่ 4-24 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-24 คือ โมเดลที่รันในรอบที่ 1 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1652.28 ค่า RAE เท่ากับ 0.45 และค่า MAPE เท่ากับ 10.28 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1994.78 ค่า RAE เท่ากับ 1.11 และค่า MAPE เท่ากับ 9.88

ตารางที่ 4-25 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 200

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validataion	test	train	validataion	test	train	validataion	test
รอบที่ 1	971.08	1742.5	2090	7.63	10.73	10.15	0.27	0.47	1.16
รอบที่ 2	1023.4	4437.9		8.03	24.36		0.29	1.22	
รอบที่ 3	1068.4	1662.9		8.28	10.65		0.3	0.45	
รอบที่ 4	971.27	1768.86		7.68	10.73		0.27	0.47	
รอบที่ 5	972.74	1702.83		7.61	10.64		0.27	0.47	

จากตารางที่ 4-25 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-25 คือ โมเดลที่รันในรอบที่ 3 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1662.9 ค่า RAE เท่ากับ 0.45 กับโมเดลที่รันในรอบที่ 5 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน MAPE เท่ากับ 10.64 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2090.00 ค่า RAE เท่ากับ 1.16 และค่า MAPE เท่ากับ 10.15

ตารางที่ 4-26 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 300

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validataion	test	train	validataion	test	train	validataion	test
รอบที่ 1	973.65	1774.88	1925.99	7.65	10.86	15.79	0.27	0.48	1.85
รอบที่ 2	971.38	1663.62		7.57	10.35		0.27	0.46	
รอบที่ 3	964.27	1666.62		7.47	10.17		0.27	0.45	
รอบที่ 4	973.01	1705.10		7.64	10.64		0.27	0.47	
รอบที่ 5	971.67	1725.81		7.66	10.65		0.27	0.47	

จากตารางที่ 4-26 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-26 คือ โมเดลที่รันในรอบที่ 2 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1663.62 กับโมเดลที่รันในรอบที่ 3 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อนค่า RAE เท่ากับ 0.45 ค่า MAPE เท่ากับ 10.17 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1925.99 ค่า RAE เท่ากับ 1.85 และค่า MAPE เท่ากับ 15.79

สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 ดังตารางที่ 4-27

ตารางที่ 4-27 สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20

elitism	RMSE	RAE	MAPE
1	10120.38	7.29	9.45
100	1994.78	1.11	9.88
200	2089.7	1.16	10.15
300	1925.99	1.85	15.79

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-27 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 80 ต่อ 20 คือ โมเดลที่มีค่า elitism 300 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1925.99 โมเดลที่มีค่า elitism 100 ได้ค่าความคลาดเคลื่อน ค่า RAE เท่ากับ 1.11 และโมเดลที่มีค่า elitism 1 ค่า MAPE เท่ากับ 9.45

ตารางที่ 4-28 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 1

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1078.98	1669.54	2208.18	8.34	10.71	10.15	0.3	0.46	1.29
รอบที่ 2	973.45	1720.98		7.69	10.64		0.28	0.47	
รอบที่ 3	970.98	1714.60		7.63	10.67		0.27	0.47	
รอบที่ 4	972.60	1763.81		7.52	10.73		0.27	0.48	
รอบที่ 5	972.98	1790.76		7.66	10.83		0.27	0.48	

จากตารางที่ 4-28 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-29 คือ โมเดลที่รันในรอบที่ 1 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1669.54 ค่า RAE เท่ากับ 0.46 กับโมเดลที่รันในรอบที่ 2 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน MAPE เท่ากับ 10.64 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2208.18 ค่า RAE เท่ากับ 1.29 และค่า MAPE เท่ากับ 10.15

ตารางที่ 4-29 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 100

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	974.42	1784.81	1987.04	7.63	10.76	9.78	0.27	0.48	1.12
รอบที่ 2	1041.18	2381.13		8.04	14.07		0.29	0.66	
รอบที่ 3	985.75	1672.52		7.70	10.56		0.28	0.46	
รอบที่ 4	970.37	1670.96		7.53	10.28		0.27	0.45	
รอบที่ 5	980.83	1685.53		7.73	10.61		0.28	0.46	

จากตารางที่ 4-29 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-29 คือ โมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1670.96 ค่า RAE เท่ากับ 0.45 และค่า MAPE เท่ากับ 10.28 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1987.04 ค่า RAE เท่ากับ 1.12 และค่า MAPE เท่ากับ 9.78

ตารางที่ 4-30 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 200

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	969.21	1771.69	1909.1	7.54	10.63	9.44	0.27	0.48	1.07
รอบที่ 2	973.42	1700.13		7.65	10.52		0.27	0.46	
รอบที่ 3	963.34	1748.34		7.49	10.61		0.27	0.47	
รอบที่ 4	970.31	1735.95		7.64	10.68		0.27	0.47	
รอบที่ 5	971.73	1790.87		7.53	10.77		0.27	0.48	

จากตารางที่ 4-30 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-30 คือ โมเดลที่รันในรอบที่ 2 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1700.13 ค่า RAE เท่ากับ 0.46 และค่า MAPE เท่ากับ 10.52 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.1 ค่า RAE เท่ากับ 1.07 และค่า MAPE เท่ากับ 9.44

ตารางที่ 4-31 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ที่ elitism เท่ากับ 300

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	973.76	1708.95	80273.64	7.72	10.65	519.73	0.28	0.47	58.25
รอบที่ 2	1068.23	1680.77		8.21	10.73		0.30	0.46	
รอบที่ 3	972.91	1764.80		7.50	10.73		0.27	0.48	
รอบที่ 4	1078.49	1691.93		8.33	10.78		0.30	0.46	
รอบที่ 5	1014.08	1642.58		7.91	10.44		0.28	0.45	

จากตารางที่ 4-31 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-31 คือ โมเดลที่รันในรอบที่ 5 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1642.58 ค่า RAE เท่ากับ 0.45 และค่า MAPE เท่ากับ 10.44 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 80273.64 ค่า RAE เท่ากับ 58.25 และค่า MAPE เท่ากับ 519.73

สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 ดังตารางที่ 4-32

ตารางที่ 4-32 สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20

elitism	RMSE	RAE	MAPE
1	2208.18	1.29	10.15
100	1987.04	1.12	9.78
200	1909.1	1.07	9.44
300	80273.64	58.25	519.73

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-32 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 80 ต่อ 20 คือ โมเดลที่มีค่า elitism 200 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.10 ค่า RAE เท่ากับ 1.07 และค่า MAPE เท่ากับ 9.44

จากผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 80 ต่อ 20 เมื่อพิจารณา ตารางสรุปผล ตารางที่ 4-27 และตารางที่ 4-32 จากวิธีการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติ พบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ ได้ผลที่ใกล้เคียงและมีประสิทธิภาพกว่า วิธีการวิเคราะห์สหสัมพันธ์ ดังตารางที่ 4-33

ตารางที่ 4-33 สรุปผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 80 ต่อ 20

การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
elitism	RMSE	RAE	MAPE	RMSE	RAE	MAPE
1	10120.38	7.29	9.45	2208.18	1.29	10.15
100	1994.78	1.11	9.88	1987.04	1.12	9.78
200	2089.7	1.16	10.15	1909.1	1.07	9.44
300	1925.99	1.85	15.79	80273.64	58.25	519.73

ต่อไปจะเป็นการหาผลลัพธ์การพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรมกับข้อมูลที่แบ่งอัตราส่วน 90 ต่อ 10 ได้ค่าความคลาดเคลื่อนดังนี้

ตารางที่ 4-34 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 1

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1137.68	1867.68	2060.64	8.82	10.02	10.43	0.33	0.69	1.16
รอบที่ 2	1000.90	2093.75		7.77	12.06		0.29	0.84	
รอบที่ 3	1012.64	1947.22		7.94	11.08		0.30	0.77	
รอบที่ 4	1109.96	1849.10		8.52	10.07		0.32	0.71	
รอบที่ 5	1003.59	1957.11		7.74	11.17		0.29	0.78	

จากตารางที่ 4-34 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-34 คือ โมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1849.10 กับโมเดลที่รันในรอบที่ 1 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อนค่า RAE เท่ากับ 0.69 ค่า MAPE เท่ากับ 10.02 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2060.64 ค่า RAE เท่ากับ 1.16 และค่า MAPE เท่ากับ 10.43

ตารางที่ 4-35 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 100

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1137.10	1865.98	2110.22	8.79	10.02	10.43	0.33	0.70	1.17
รอบที่ 2	1007.52	1969.03		7.80	11.27		0.29	0.78	
รอบที่ 3	1091.81	1903.52		8.43	10.35		0.31	0.72	
รอบที่ 4	1011.98	2036.86		7.90	11.75		0.30	0.81	
รอบที่ 5	1010.69	2085.27		7.92	12.05		0.30	0.83	

จากตารางที่ 4-35 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-35 คือ โมเดลที่รันในรอบที่ 1 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1865.98 ค่า RAE เท่ากับ 0.70 และค่า MAPE เท่ากับ 10.02 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2110.22 ค่า RAE เท่ากับ 1.17 และค่า MAPE เท่ากับ 10.43

ตารางที่ 4-36 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 200

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1013.92	2022.50	2109.86	7.89	11.65	10.43	0.30	0.81	1.17
รอบที่ 2	1009.10	2086.86		7.81	12.05		0.29	0.83	
รอบที่ 3	1012.93	2026.56		7.88	11.68		0.30	0.81	
รอบที่ 4	1137.45	1866.12		8.80	10.02		0.33	0.70	
รอบที่ 5	1013.26	2153.20		7.78	12.43		0.29	0.86	

จากตารางที่ 4-36 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-36 คือ โมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1866.12 ค่า RAE เท่ากับ 0.70 และค่า MAPE เท่ากับ 10.02 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2109.86 ค่า RAE เท่ากับ 1.17 และค่า MAPE เท่ากับ 10.43

ตารางที่ 4-37 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 300

การเลือกปัจจัยโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ (Correlation)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1011.17	2060.98	2110.15	7.97	11.92	10.43	0.30	0.82	1.16
รอบที่ 2	1044.80	2313.65		7.81	13.29		0.29	0.92	
รอบที่ 3	1137.07	1866.12		8.80	10.02		0.33	0.70	
รอบที่ 4	1140.86	1871.62		8.89	10.03		0.33	0.69	
รอบที่ 5	1013.49	2086.97		7.95	12.06		0.30	0.83	

จากตารางที่ 4-37 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจากในตารางที่ 4-37 คือ โมเดลที่รันในรอบที่ 3 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1866.12 ค่า MAPE เท่ากับ 10.02 กับโมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RAE เท่ากับ 0.69 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2110.15 ค่า RAE เท่ากับ 1.16 และค่า MAPE เท่ากับ 10.43

สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 ดังตารางที่ 4-38

ตารางที่ 4-38 สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10

elitism	RMSE	RAE	MAPE
1	2060.64	1.16	10.43
100	2110.22	1.17	10.43
200	2109.86	1.17	10.43
300	2110.15	1.16	10.43

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-38 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการวิเคราะห์สหสัมพันธ์ ข้อมูลอัตราส่วน 90 ต่อ 10 คือ โมเดลที่มีค่า elitism 1 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2060.64 ค่าความคลาดเคลื่อน RAE เท่ากับ 1.16 มีโมเดลที่มีค่า elitism เท่ากันอยู่ 2 โมเดล คือ โมเดลที่มีค่า elitism 1 และ3 และค่า MAPE มีค่าเท่ากับ 10.43 ซึ่งเท่ากันทุกโมเดล

ตารางที่ 4-39 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 1

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validation	test	train	validation	test	train	validation	test
รอบที่ 1	1010.63	2116.24		7.77	12.22		0.29	0.85	
รอบที่ 2	1042.91	2173.53		8.30	12.56		0.31	0.86	
รอบที่ 3	1014.08	2056.86	7114.70	7.68	11.38	44.81	0.29	0.82	4.92
รอบที่ 4	1026.90	1934.78		8.10	11.07		0.30	0.77	
รอบที่ 5	1003.92	2090.91		7.76	12.08		0.29	0.84	

จากตารางที่ 4-39 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-39 คือ โมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1934.78 ค่า RAE เท่ากับ 0.77 และค่า MAPE เท่ากับ 11.07 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 7114.70 ค่า RAE เท่ากับ 4.92 และค่า MAPE เท่ากับ 44.81

ตารางที่ 4-40 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 100

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validataion	test	train	validataion	test	train	validataion	test
รอบที่ 1	1067.87	1875.35	2822.61	8.14	10.34	15.79	0.31	0.73	1.71
รอบที่ 2	1040.59	2134.24		7.95	12.32		0.30	0.85	
รอบที่ 3	1003.34	1961.63		7.76	11.21		0.29	0.78	
รอบที่ 4	1009.79	2060.33		7.85	11.90		0.29	0.82	
รอบที่ 5	1002.68	1962.62		7.76	11.24		0.29	0.78	

จากตารางที่ 4-40 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-40 คือ โมเดลที่รันในรอบที่ 1 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1875.35 ค่า RAE เท่ากับ 0.73 และค่า MAPE เท่ากับ 10.34 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2822.61 ค่า RAE เท่ากับ 1.71 และค่า MAPE เท่ากับ 15.79

ตารางที่ 4-41 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 200

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validataion	test	train	validataion	test	train	validataion	test
รอบที่ 1	1012.98	2053.40	2105.77	7.80	11.84	10.41	0.29	0.83	1.17
รอบที่ 2	1009.77	2114.12		7.83	12.21		0.29	0.84	
รอบที่ 3	1009.76	2120.17		7.80	12.24		0.29	0.85	
รอบที่ 4	1100.74	1861.62		8.45	10.13		0.32	0.71	
รอบที่ 5	1009.13	2092.04		7.82	12.08		0.29	0.83	

จากตารางที่ 4-41 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-41 คือ โมเดลที่รันในรอบที่ 4 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1861.62 ค่า RAE เท่ากับ 0.71 และค่า MAPE เท่ากับ 10.13 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2105.77 ค่า RAE เท่ากับ 1.17 และค่า MAPE เท่ากับ 10.41

ตารางที่ 4-42 ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ที่ elitism เท่ากับ 300

การเลือกปัจจัยโดยใช้วิธีการคัดเลือกคุณสมบัติ (Feature Selection)									
errors	RMSE			MAPE			RAE		
	train	validataion	test	train	validataion	test	train	validataion	test
รอบที่ 1	1037.99	1942.80		8.08	11.03		0.30	0.77	
รอบที่ 2	1019.63	1989.53		7.96	11.43		0.30	0.79	
รอบที่ 3	1011.06	2100.58	2472.06	7.77	12.14	13.36	0.29	0.84	1.45
รอบที่ 4	1010.74	2048.01		7.85	11.82		0.29	0.82	
รอบที่ 5	1031.94	1921.08		8.14	10.95		0.30	0.76	

จากตารางที่ 4-42 พิจารณาค่าความคลาดเคลื่อนช่อง validation ที่มีค่าความคลาดเคลื่อนน้อยที่สุดจาก RMSE, RAE และMAPE ในตารางที่ 4-44 คือ โมเดลที่รันในรอบที่ 5 ได้ผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรม มีค่าความคลาดเคลื่อน RMSE เท่ากับ 1921.08 ค่า RAE เท่ากับ 0.76 และค่า MAPE เท่ากับ 10.95 เมื่อได้โมเดลที่มีค่าน้อยที่สุดแล้ว ให้เอาโมเดลนั้นไปหาค่า test ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2472.06 ค่า RAE เท่ากับ 1.45 และค่า MAPE เท่ากับ 13.36

สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 ดังตารางที่ 4-43

ตารางที่ 4-43 สรุปผลการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10

elitism	RMSE	RAE	MAPE
1	7114.70	4.92	44.81
100	2822.61	1.71	15.79
200	2105.77	1.17	10.41
300	2472.06	1.45	13.36

พิจารณาค่าที่น้อยที่สุดจากตารางที่ 4-43 โมเดลที่พยากรณ์ราคาบิทคอยน์ ได้ใกล้เคียงและมีประสิทธิภาพที่สุดของการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกคุณสมบัติ ข้อมูลอัตราส่วน 90 ต่อ 10 คือ โมเดลที่มีค่า elitism 200 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 2105.77 ค่า RAE เท่ากับ 1.17 และค่า MAPE เท่ากับ 10.41

จากผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 90 ต่อ 10 เมื่อพิจารณาตารางสรุปผล ตารางที่ 4-38 และตารางที่ 4-43 จากวิธีการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติ พบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรมโดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติได้ผลที่ใกล้เคียงและมีประสิทธิภาพใกล้เคียงกับ วิธีการวิเคราะห์สหสัมพันธ์ต่างกันเล็กน้อย ดังตารางที่ 4-44

ตารางที่ 4-44 สรุปผลการทดลองการพยากรณ์โดยใช้วิธีการโปรแกรมเชิงพันธุกรรมของข้อมูลในอัตราส่วน 90 ต่อ 10

การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
elitism	RMSE	RAE	MAPE	RMSE	RAE	MAPE
1	2060.64	1.16	10.43	7114.7	4.92	44.81
100	2110.22	1.17	10.43	2822.61	1.71	15.79
200	2109.86	1.17	10.43	2105.77	1.17	10.41
300	2110.15	1.16	10.43	2472.06	1.45	13.36

เราจะสรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรม จากตารางที่ 4-45 ได้ผลลัพธ์ดังนี้

ตารางที่ 4-45 สรุปผลการทดลองพยากรณ์ราคาบิทคอยน์ ด้วยวิธีการโปรแกรมเชิงพันธุกรรม

การเลือกปัจจัย	วิธีการวิเคราะห์สหสัมพันธ์			วิธีการคัดเลือกคุณสมบัติ		
ERRORS	RMSE	RAE	MAPE	RMSE	RAE	MAPE
80/20	1994.78	1.11	9.88	1909.1	1.07	9.44
90/10	2060.64	1.16	10.43	2105.77	1.17	10.41

เมื่อพิจารณาตารางที่ 4-45 จะสรุปผลได้ว่า วิธีการเลือกปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ เพื่อมาพยากรณ์ราคาบิทคอยน์ จากการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ ให้ค่าความคลาดเคลื่อนที่ใกล้เคียงและมีประสิทธิภาพกว่าวิธีการเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และการแบ่งอัตราส่วนข้อมูลที่ใช้ในการทดลอง ถ้าแบ่งอัตราส่วนข้อมูลที่ 80 ต่อ 20 จะได้ค่าความคลาดเคลื่อนที่ดีกว่าการแบ่งอัตราส่วนข้อมูลที่ 90 ต่อ 10 ซึ่งการพยากรณ์ราคาบิทคอยน์ด้วยวิธีการโปรแกรมเชิงพันธุกรรมที่ใกล้เคียงและมีประสิทธิที่สุดกับราคาบิทคอยน์ คือ โมเดลที่ใช้การคัดเลือกด้วยวิธีการคัดเลือกคุณสมบัติ มีค่า elitism 200 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.10 ค่า RAE เท่ากับ 1.07 และค่า MAPE เท่ากับ 9.44

4.4 สรุปผลการทดลอง

เมื่อได้ผลสรุปจากการทดลองการพยากรณ์ราคาบิทคอยน์ จากเทคนิควิธีการวิเคราะห์การถดถอย วิธีการโครงข่ายประสาทเทียม และวิธีการโปรแกรมเชิงพันธุกรรม ซึ่งมีวิธีการคัดเลือกปัจจัยโดยวิธีการวิเคราะห์สหสัมพันธ์และวิธีการคัดเลือกคุณสมบัติ โดยแบ่งอัตราส่วนของข้อมูลเป็น 80 ต่อ 20 และ90 ต่อ 10 เสร็จแล้ว นำผลสรุปจากค่าความคลาดเคลื่อน RMSE, RAE และ MAPE ที่น้อยที่สุดจากตารางที่ 4-3, 4-22 และ4-45 ได้ผลสรุปดังตารางที่ 4-46

ตารางที่ 4-46 สรุปผลการพยากรณ์ราคาบิทคอยน์

	RMSE	RAE	MAPE
วิธีวิเคราะห์การถดถอย	1831.89	1.04	9.10
วิธีโครงข่ายประสาทเทียม	1750.96	1.03	9.06
วิธีโปรแกรมพันธุกรรม	1909.10	1.07	9.44

จากตารางที่ 4-46 สรุปผลการพยากรณ์ราคาบิทคอยน์ ได้ดังนี้

4.4.1 วิธีวิเคราะห์การถดถอย ได้ผลการทดลองที่ดีที่สุดจากวิธีการหาความสัมพันธ์ของข้อมูลด้วยวิธีการคัดเลือกคุณสมบัติได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1831.89 RAE เท่ากับ 1.04 และMAPE เท่ากับ 9.10 โดยนำผลลัพธ์จากตารางที่ 4-2 และข้อมูลปัจจัยที่ได้จากการหาความสัมพันธ์ของข้อมูลด้วยวิธีการคัดเลือกคุณสมบัติ จากตารางที่ 3-4 มาแทนค่าในสมการถดถอยดังนี้

Y = (-93.43 + (-5.40*eur) + (87.18*jpy) + (-1.66*index) + (44.80*91) + (3.12*e20) + (0.29*btc2) + (0.60*btc1))

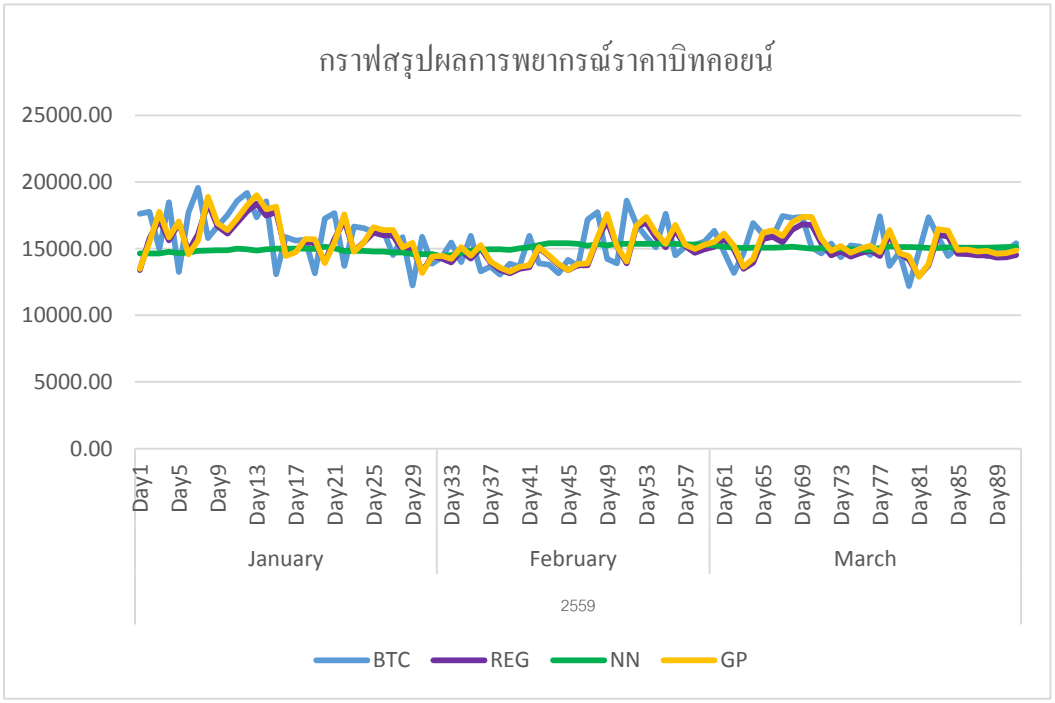
4.4.2 วิธีการโครงข่ายประสาทเทียม ได้ผลการทดลองที่ดีที่สุดจากวิธีการหาความสัมพันธ์ของข้อมูลด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 90 ต่อ 10 ได้โมเดลที่ใช้ learningRate 0.3 และtrainingTime 5000 ให้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 และค่า MAPE เท่ากับ 9.06 ส่วนค่า RAE เท่ากับ 1.03 ได้จากโมเดลที่มีค่า learningRate 0.1 และ trainingTime 500

4.4.3 วิธีการโปรแกรมเชิงพันธุกรรม ได้ผลการทดลองที่ดีที่สุดจากวิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 80 ต่อ 20 โมเดลที่ใช้มีค่า elitism 200 รันในรอบที่ 2 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.10 ค่า RAE เท่ากับ 1.07 และค่า MAPE เท่ากับ 9.44 โดยนำสมการจาก ตารางที่ 4-30 ได้สมการมาดังนี้

Y = (((IF (ISNUMBER (((IF (ISNUMBER (R6 /91), R6 /91, 0)) + (R3 +R4)) / ((INDEX*91) - (JPY-EUR))), ((IF (ISNUMBER (R6 /91), R6 /91, 0)) + (R3 +R4)) / ((INDEX*91) - (JPY-EUR)), 0)) - ((IF (ISNUMBER (E20/EUR), E20/EUR, 0)) - ((E20+BTC1) - (R4 -R5)))) + (IF (ISNUMBER ((IF (ISNUMBER (((IF (ISNUMBER ((IF (ISNUMBER (JPY/E20), JPY/ E20, 0))/JPY), (IF (ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY, 0)) + ((R3 -JPY) +BTC1)) /((IF(ISNUMBER(INDEX/(JPY-R5)),INDEX/(JPY-R5),0)) - ((IF(ISNUMBER(BTC2/EUR) ,BTC2/EUR,0)) * (R1 +INDEX)))),((IF(ISNUMBER((IF(ISNUMBER(JPY/E20),JPY/E20,0)) /JPY), (IF (ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY, 0)) + ((R3 -JPY) +BTC1)) /((IF(ISNUMBER(INDEX/(JPY-R5)),INDEX/(JPY-R5),0)) -((IF(ISNUMBER(BTC2/EUR),

BTC2/EUR, 0)) *(R1 +INDEX))), 0)) / (IF (ISNUMBER (((E20-JPY) + (JPY+91)) / ((R2 - (E20*91)) *((R2 *R2) - (BTC1-BTC2)))), ((E20-JPY) + (JPY+91)) / ((R2 - (E20*91)) *((R2 *R2) -(BTC1-BTC2))),0))),(IF(ISNUMBER(((IF(ISNUMBER((IF(ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY), (IF (ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY, 0)) + ((R3 -JPY) +BTC1)) / ((IF (ISNUMBER (INDEX/ (JPY-R5)), INDEX/ (JPY-R5), 0)) - ((IF (ISNUMBER (BTC2/EUR), BTC2/EUR, 0)) *(R1 +INDEX)))), ((IF (ISNUMBER ((IF (ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY), (IF (ISNUMBER (JPY/E20), JPY/E20, 0)) /JPY,0)) +((R3 -JPY) +BTC1)) /((IF(ISNUMBER(INDEX/(JPY-R5)),INDEX/(JPY-R5),0)) - ((IF (ISNUMBER (BTC2/EUR), BTC2/EUR, 0)) *(R1 +INDEX))), 0)) / (IF (ISNUMBER (((E20-JPY) + (JPY+91)) / ((R2 - (E20*91)) *((R2 *R2) - (BTC1-BTC2)))) ,((E20-JPY) +(JPY+91)) /((R2 -(E20*91)) *((R2 *R2) -(BTC1-BTC2))),0)),0)))

เมื่อนำผลการทดลองจาก ตารางที่ 4-46 มาพล็อตกราฟ ดังภาพที่ 4-4 จะเห็นว่า วิธีโครงข่ายประสาทเทียมมีเส้นกราฟที่วิ่งเป็นแนวนอนไม่สัมพันธ์กับราคาบิทคอยน์ แต่ได้ค่าพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพที่สุดเมื่อดูจากค่าความคลาดเคลื่อนในตารางที่ 4-46 ส่วน วิธีวิเคราะห์การถดถอย และวิธีการโปรแกรมเชิงพันธุกรรมเส้นกราฟวิ่งไปในทิศทางที่สัมพันธ์กับราคาบิทคอยน์ จากผลการทดลองทั้งหมดสรุปได้ว่าวิธีการพยากรณ์ราคาบิทคอยน์ ทั้ง 3 วิธีได้ผลที่ใกล้เคียงกันต่างกันเพียงเล็กน้อยขึ้นอยู่กับช่วงเวลาของราคาจริงที่มีการผันผวนมากหรือน้อยในแต่ละวัน



ภาพที่ 4-4 กราฟสรุปผลการพยากรณ์ราคาบิทคอยน์

บทที่ 5
สรุปผลและข้อเสนอแนะ

งานนิพนธ์นี้นำเสนอการพยากรณ์ราคาบิทคอยน์ จากข้อมูลปัจจัยต่างๆที่ผ่านการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกปัจจัยด้วยวิธีการวิเคราะห์สหสัมพันธ์และวิธีคัดเลือกคุณสมบัติก่อนจะนำปัจจัยที่ได้มาทดลองพยากรณ์ราคาบิทคอยน์ ด้วยเทคนิคทั้ง 3 วิธี คือ 1. วิธีวิเคราะห์การถดถอย 2. วิธีโครงข่ายประสาทเทียม และ 3. วิธีการโปรแกรมเชิงพันธุกรรม ด้วยการแบ่งอัตราข้อมูลที่ 80 ต่อ 20 และ90 ต่อ 10 เพื่อให้ได้ผลการพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพที่สุด ซึ่งจากการทดสอบสามารถสรุปผลดังนี้

5.1 สรุปผล

ข้อมูลที่ใช้ในงานวิจัยนี้ผู้วิจัยได้เลือกข้อมูลปัจจัยที่คาดว่าจะมีผลต่อราคาขึ้นลงของราคาบิทคอยน์ ประกอบด้วยข้อมูล 2 ส่วน คือ ข้อมูลปัจจัยด้านต่างๆ ได้แก่ 1. อัตราแลกเปลี่ยนเงินประกอบด้วย ค่าเงินดอลลาร์สหรัฐของอเมริกา (USD) ค่าเงินยูโรของประเทศในยุโรป (EUR) ค่าเงินเยนของญี่ปุ่น (JPY) ค่าเงินดอลล่าร์ฮ่องกงของฮ่องกง (HKD) ค่าเงินหยวนของจีน และค่าเงินวอนของเกาหลี 2. ข้อมูลการซื้อขายหุ้น มี NASDAQ, INDEX, NIKKEI, DOWJONE และ HANGSENG 3. ราคาทองแท่ง และราคาทองรูปพรรณ 4. ข้อมูลน้ำมัน มี น้ำมันดีเซล, แก๊สโซฮอล์ 91, แก๊สโซฮอล์ e20 และแก๊ส NGV ส่วนที่ 2 ได้แก่ ราคาบิทคอยน์ ย้อนหลัง 1 วัน และราคาบิทคอยน์ ย้อนหลัง 2 วัน โดยจะใช้ข้อมูล ตั้งแต่วันที่ 1 กรกฎาคม 2557 ถึง วันที่ 31 มีนาคม 2559 รวม 640 วัน โดยจะใช้ข้อมูลราคาปิดของแต่ละวันโดยจะให้แทนชื่อปัจจัยต่างๆ ดังตารางที่ 3-2

จากนั้นนำข้อมูลปัจจัยทั้งหมดมาคัดเลือกข้อมูลเพื่อหาข้อมูลปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีคัดเลือกคุณสมบัติ ซึ่งวิธีการวิเคราะห์สหสัมพันธ์ได้ข้อมูลที่มีความสัมพันธ์กับข้อมูล BTC คือ EUR, JPY, KRW, NIKKIE, 91, E20, NGV, BTC1 และBTC2 จำนวน 9 ปัจจัยหรือ 9 Attributes และวิธีคัดเลือกคุณสมบัติ ได้ข้อมูลที่มีความสัมพันธ์กับข้อมูล BTC คือ EUR, JPY, INDEX, 91, E20, BTC1 และBTC2 จำนวน 7 ปัจจัยหรือ 7 Attributes ดังตารางที่ 3-3 และ3-4

เมื่อได้ข้อมูลที่มีความสัมพันธ์กับราคาบิทคอยน์ จากการคัดเลือกปัจจัยที่มีความสัมพันธ์กับราคาบิทคอยน์ ด้วยวิธีการวิเคราะห์สหสัมพันธ์ และวิธีคัดเลือกคุณสมบัติ แล้วนำปัจจัยที่ได้มาทดลองพยากรณ์ราคาบิทคอยน์ ด้วยเทคนิคทั้ง 3 วิธี คือ 1. วิธีวิเคราะห์การถดถอย 2. วิธีโครงข่าย

ประสาทเทียม และ 3. วิธีการโปรแกรมเชิงพันธุกรรม ด้วยการแบ่งอัตราข้อมูลที่ 80 ต่อ 20 และ90 ต่อ 10

จากผลการทดลองพบว่าการพยากรณ์ราคาบิทคอยน์ ด้วยวิธี 1.วิธีวิเคราะห์การถดถอย โดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ ได้ค่าความคลาดเคลื่อนที่น้อยที่สุดที่ ค่า RMSE เท่ากับ 1831.89 ค่า RAE เท่ากับ 1.04 และค่า MAPE เท่ากับ 9.10 2.วิธีโครงข่ายประสาทเทียมโดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 90 ต่อ 10 โมเดลที่ใช้ learningRate 0.3 และtrainingTime 5000 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1750.96 และค่า MAPE เท่ากับ 9.06 ส่วนค่า RAE เท่ากับ 1.03 ได้จากโมเดลที่มีค่า learningRate 0.1 และtrainingTime 500 และ 3.วิธีโปรแกรมพันธุกรรม โดยใช้วิธีการคัดเลือกปัจจัยด้วยวิธีการคัดเลือกคุณสมบัติ โดยใช้อัตราส่วนข้อมูล 80 ต่อ 20 โมเดลที่ใช้มีค่า elitism 200 ได้ค่าความคลาดเคลื่อน RMSE เท่ากับ 1909.10 ค่า RAE เท่ากับ 1.07 และค่า MAPE เท่ากับ 9.44 ซึ่งวิธีการพยากรณ์ราคาบิทคอยน์ ทั้ง 3 วิธีได้ผลที่ใกล้เคียงกันต่างกันเพียงเล็กน้อย

เมื่อเราต้องการซื้อหรือขายราคาบิทคอยน์ เราสามารถพิจารณาจากปัจจัยที่มีอิทธิพลต่อราคาบิทคอยน์ มากที่สุด จาก สมการที่ได้จากการพยากรณ์ราคาบิทคอยน์ ด้วยวิธีวิเคราะห์การถดถอยได้ โดยใช้วิธีการคัดเลือกข้อมูลด้วยวิธีการคัดเลือกคุณสมบัติ

Y = (-93.43 + (-5.40*eur) + (87.18*jpy) + (-1.66*index) + (44.80*91) + (3.12*e20) + (0.29*btc2) + (0.60*btc1))

โดยการเอาค่าสัมประสิทธิ์สหสัมพันธ์ของแต่ละปัจจัยคูณกับค่าปัจจัย จะได้ปัจจัยที่มีอิทธิพลต่อราคาบิทคอยน์ มากที่สุด จากสมการนี้คือ btc1 (ราคาบิทคอยน์ ย้อนหลัง 1 วัน) เมื่อได้ปัจจัยที่มีอิทธิพลต่อราคาบิทคอยน์ แล้ว ก็สามารถคาดเดาแนวโน้มการขึ้นลงของราคาบิทคอยน์ได้

5.2 ข้อเสนอแนะ

เพื่อให้ได้การพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพที่สุดผู้วิจัยเห็นว่าควรมีการพัฒนาวิธีการดังต่อไปนี้

5.2.1 ศึกษาข้อมูลปัจจัยอื่นๆ ที่น่าจะมีความสัมพันธ์กับราคาบิทคอยน์ มากกว่านี้เพื่อให้ได้ข้อมูลที่มีความสัมพันธ์และมีผลกับราคาบิทคอยน์

5.2.2 หาอัลกอริทึมใหม่ที่มีประสิทธิภาพ มาใช้ในการในการพยากรณ์ราคาบิทคอยน์ เพื่อให้ได้ผลพยากรณ์ราคาบิทคอยน์ ที่ใกล้เคียงและมีประสิทธิภาพมายิ่งขึ้น

5.2.3 ลองใช้โปรแกรมพยากรณ์อื่นๆ เพื่อพยากรณ์ราคาบิทคอยน์

บรรณานุกรม

การซื้อขายแลกเปลี่ยน Bitcoins ในประเทศไทย. (2557). วันที่ค้นข้อมูล 12 กรกฎาคม 2557, เข้าถึงได้ที่ https://bitcoin.co.th/

การวิเคราะห์การถดถอย. “Regression.” (2557). วันที่ค้นข้อมูล 10 ตุลาคม 2557, เข้าถึงได้ที่ http://www.excel-easy.com/examples/regression.html

เกียรติศักดิ์ จันทร์แก้ว. (2554). *การประยุกต์ใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น เพื่อการพยากรณ์ปริมาณการส่งออกข้าวจ้าวนาปี*. การประชุมและนำเสนอผลงานวิชาการเทคโนโลยีภาคใต้วิจัย “สหวิทยาการเพื่อการพัฒนาอย่างยั่งยืน” 21 มกราคม 2554

ข้อมูลราคาบิทคอยน์ย้อนหลัง. “Bitcoin charts.” (2557). วันที่ค้นข้อมูล 13 กรกฎาคม 2557, เข้าถึงได้ที่ http://bitcoincharts.com/

ข้อมูลการซื้อขายบิทคอยน์ย้อนหลัง. “Bitcoin Trading Charts.” (2557). วันที่ค้นข้อมูล 19 ตุลาคม 2557, เข้าถึงได้ที่ https://bx.in.th/

ข้อมูลราคาทองคำ. (2557). วันที่ค้นข้อมูล 21 ตุลาคม 2557, เข้าถึงได้ที่ http://www.goldpricethai.com/th/archive

ข้อมูลย้อนหลังของหุ้น. “NASDAQ Composite.” (2557). วันที่ค้นข้อมูล 24 ตุลาคม 2557 เข้าถึงได้ที่ http://th.investing.com/indices/nasdaq-composite-historical-data

ข้อมูลย้อนหลังของหุ้น. “SET INDEX.” (2557). วันที่ค้นข้อมูล 26 ตุลาคม 2557 เข้าถึงได้ที่ http://www.settrade.com/

ข้อมูลย้อนหลังของราคาน้ำมัน. (2557). วันที่ค้นข้อมูล 27 ตุลาคม 2557 เข้าถึงได้ที่ http://www.bangchak.co.th/

คมสัน สุริยะ. (2548). *การพยากรณ์ราคาน้ำมันดิบด้วยแบบจำลองนิวรอลเน็ตเวิร์ค*. ศูนย์การวิเคราะห์เชิงปริมาณ, คณะเศรษฐศาสตร์, มหาวิทยาลัยเชียงใหม่.

ชวลิต ทับสีรัก. (2555). *ตัวแปรเชิงคุณภาพกับการวิเคราะห์การถดถอย*. วารสารการวัดผลการศึกษา, 17 กรกฎาคม 2555. มหาวิทยาลัยมหาสารคาม, หน้า 31 – 42.

ชาญชัย ชัยสุขโกศล, บุญเสริม แก้วกำเนิดพงษ์ และอรวรรณ จันทร์เพ็ญ. (2541). *การแก้ปัญหาอย่างง่ายของหุ่นยนต์แบบอิงพฤติกรรมด้วยโปรแกรมพันธุการ*. วิทยานิพนธ์วิศวกรรมศาสตร์บัณฑิต, สาขาวิศวกรรมคอมพิวเตอร์, บัณฑิตวิทยาลัย, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี.

บรรณานุกรม (ต่อ)

เทิดศิลป์ โสมูล. (2555). *ระบบควบคุมการเลี้ยงด้วยภาพแบบอัตโนมัติโดยใช้เครือข่ายประสาทเทียม*. หลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต, สาขาวิชาวิศวกรรมไฟฟ้า, มหาวิทยาลัยเทคโนโลยีสุรนารี.

ไท้ ปริญญา. (2549). GENETIC PROGRAMIMNG. วันที่ค้นข้อมูล 25 ธันวาคม 2557 เข้าถึงได้ที่ https://www.l3nr.org/posts/157040

นันทชัย กานตานันทะ. (2555). วารสารวิศวกรรมศาสตร์ ปีที่ 4 ฉบับที่ 1 จุฬาลงกรณ์มหาวิทยาลัย.

นิภาพร ชนะมาร. (2557). *การวิเคราะห์ปัจจัยการเรียนรู้ด้วยการคัดเลือกคุณสมบัติและการพยากรณ์*. ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ, คณะวิทยาศาสตร์, มหาวิทยาลัยนเรศวร.

บุษรา ลิ้มพิพัฒนางกูร. (2549). *การพยากรณ์โอกาสการสำเร็จการศึกษาของนักศึกษา โดยใช้จีนีติกอัลกอริทึมแบบหลายวัตถุประสงค์ เปรียบเทียบกับวิธีทางระบบโครงข่ายประสาทเทียม: กรณีศึกษานักศึกษาภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ*. สาขาวิชาวิทยาการคอมพิวเตอร์, ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ, บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ผศ.ดร.กรุง สินอภิรมย์สราญ. (2557). *การจัดจำแนกประเภทโดยเครือข่ายประสาท Weka*. วันที่ค้นข้อมูล 22 พฤศจิกายน 2557 เข้าถึงได้ที่ http://pioneer.netserv.chula.ac.th/~skrung/csc662/lab10_weka_neuralNetwork_4up.pdf

ภรัณยา อำมฤครัตน. (2553). *การเปรียบเทียบประสิทธิภาพการลดมิติข้อมูลและจำแนกข้อมูลโดยวิธีการทางเครือข่ายประสาทเทียม*. การประชุมทางวิชาการเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 11 มหาวิทยาลัยขอนแก่น.

เมตตา คงคากูล. (2550). *การพยากรณ์ผลสัมฤทธิ์ทางการเรียน จากพฤติกรรมการใช้อินเตอร์เน็ตโดยวิธีโครงข่ายประสาทเทียม กรณีศึกษา: นักศึกษาปริญญาตรี มหาวิทยาลัยราชภัฏหมู่บ้านจอมบึง*. สาขาวิชา วิทยาการคอมพิวเตอร์, ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ, บัณฑิตวิทยาลัย, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

วชิราภรณ์ แก้วมาตย์. (2544). *การพยากรณ์ดัชนีตลาดหลักทรัพย์ของกลุ่มประเทศอาเซียนโดยใช้โครงข่ายประสาทเทียม*. วิทยานิพนธ์ปริญญาเศรษฐศาสตรมหาบัณฑิต, มหาวิทยาลัยขอนแก่น.

บรรณานุกรม (ต่อ)

สิทธิชัย เทพไพฑูรย์. *การทำนายสัญญาณการซื้อขายหลักทรัพย์ด้วยการใช้เจเนติคอัลกอริธึม ร่วมกับ โครงข่ายประสาทเทียม*. ภาควิชา วิศวกรรมคอมพิวเตอร์, คณะวิศวกรรมศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

อดิศร ศิริจอมแปง. (2559). *การค้นหาในแนวกว้าง, แนวลึก และ ดีที่สุดก่อน (Breadth – first search)*. วันที่ค้นข้อมูล 25 พฤศจิกายน 2557 เข้าถึงได้ที่ https://www.l3nr.org/posts/157040

Jason Brownlee. (2558). Feature *Selection to Improve Accuracy and Decrease Training Time*. Retrieved February 2, 2016, from http://machinelearningmastery.com/feature-selection-to-improve-accuracy-and-decrease-training-time/

ภาคผนวก

ผลการวิเคราะห์สหสัมพันธ์ (Correlation) ด้วยโปรแกรม Excel

usd	*eur*	*jpy*	*hkd*	*cny*	*krw*	*nasdaq*	*index*	*nikkei*	*dow jones*	*hang seng*	*goldbar*	*goldrefiners*	*diesel*	*91*	*e20*	*ngv*	*btc2*	*btc1*	*btc*
1.00																			
-0.03	1.00																		
0.09	0.68	1.00																	
0.94	-0.06	-0.06	1.00																
0.98	0.01	0.10	0.92	1.00															
-0.07	0.66	0.65	-0.08	-0.04	1.00														
0.51	-0.69	-0.52	0.50	0.48	-0.41	1.00													
-0.92	-0.07	-0.22	-0.86	-0.86	-0.07	-0.38	1.00												
0.52	-0.73	-0.55	0.50	0.50	-0.51	0.95	-0.39	1.00											
-0.15	-0.67	-0.64	-0.13	-0.17	-0.39	0.70	0.26	0.61	1.00										
-0.51	-0.44	-0.39	-0.47	-0.47	-0.06	0.27	0.53	0.30	0.57	1.00									
-0.08	0.36	0.31	-0.07	-0.09	0.46	-0.38	0.07	-0.37	-0.34	-0.08	1.00								
-0.08	0.37	0.31	-0.07	-0.09	0.46	-0.38	0.07	-0.37	-0.34	-0.07	1.00	1.00							
-0.87	0.42	0.18	-0.82	-0.80	0.29	-0.71	0.81	-0.72	-0.13	0.25	0.13	0.13	1.00						
-0.68	0.63	0.41	-0.65	-0.62	0.48	-0.75	0.56	-0.76	-0.36	0.07	0.18	0.18	0.89	1.00					
-0.73	0.62	0.38	-0.71	-0.66	0.47	-0.76	0.62	-0.76	-0.31	0.15	0.18	0.18	0.94	0.96	1.00				
0.70	-0.65	-0.45	0.68	0.65	-0.53	0.80	-0.57	0.82	0.37	-0.07	-0.21	-0.21	-0.88	-0.94	-0.95	1.00			
-0.14	0.67	0.59	-0.14	-0.16	0.64	-0.46	-0.09	-0.56	-0.33	-0.30	0.08	0.08	0.35	0.57	0.53	-0.61	1.00		
-0.13	0.66	0.59	-0.14	-0.15	0.64	-0.45	-0.09	-0.56	-0.33	-0.30	0.08	0.08	0.35	0.57	0.52	-0.60	0.94	1.00	
-0.12	**0.66**	**0.60**	**-0.13**	**-0.15**	**0.64**	**-0.44**	**-0.10**	**-0.55**	**-0.33**	**-0.30**	**0.08**	**0.09**	**0.34**	**0.56**	**0.51**	**-0.60**	**0.92**	**0.94**	**1.00**

ผลการคัดเลือกคุณสมบัติ (Feature Selection) ด้วยโปรแกรม Weka

=== Run information ===

Evaluator: weka.attributeSelection.CfsSubsetEval -P 1 -E 1

Search: weka.attributeSelection.BestFirst -D 1 -N 5

Relation: all

Instances: 549

Attributes: 20

usd

eur

jpy

hkd

cny

krw

nasdaq

index

nikkei

dow jones

hang seng

goldbar

goldrefiners

diesel

91

e20

ngv

btc2

btc1

btc

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:

Best first.

Start set: no attributes

Search direction: forward

Stale search after 5 node expansions

Total number of subsets evaluated: 164

Merit of best subset found: 0.946

Attribute Subset Evaluator (supervised, Class (numeric): 20 btc):

CFS Subset Evaluator

Including locally predictive attributes

Selected attributes: 2,3,8,15,16,18,19 : 7

eur

jpy

index

91

e20

btc2

btc1

ตัวอย่างผลการวิเคราะห์การถดถอย (Regression Analysis) ด้วยโปรแกรม Excel

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.965248804
R Square	0.931705253
Adjusted R Square	0.930272496
Standard Error	943.370403
Observations	439

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	9	5208504309	578722701	650.2884269	9.5338E-244
Residual	429	381787570.7	889947.7172		
Total	438	5590291880			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-7048.463482	3913.508039	-1.801060177	0.07239552	-14740.49924	643.5722727	-14740.49924	643.5722727
eur	26.90726438	33.99620787	0.791478405	0.429102253	-39.91259233	93.72712109	-39.91259233	93.72712109
jpy	82.41608364	35.08072916	2.349326414	0.019260032	13.46459042	151.3675768	13.46459042	151.3675768
krw	29057.81201	22965.31131	1.265291449	0.206453675	-16080.71687	74196.3409	-16080.71687	74196.3409
nikkei	0.032111169	0.050266232	0.638821877	0.523279723	-0.066687569	0.130909906	-0.066687569	0.130909906
91	41.61696166	36.65866429	1.135255811	0.256902013	-30.43597744	113.6699008	-30.43597744	113.6699008
e20	68.07394005	51.28924041	1.327255766	0.185130001	-32.73552989	168.88341	-32.73552989	168.88341
ngv	68.93098253	207.1919092	0.332691478	0.739529694	-338.3066053	476.1685703	-338.3066053	476.1685703
btc2	0.36196624	0.044873655	8.066341823	7.31988E-15	0.273766662	0.450165819	0.273766662	0.450165819
btc1	0.452254464	0.044568664	10.14736426	7.89621E-22	0.364654349	0.539854579	0.364654349	0.539854579

ตัวอย่างผลการโครงข่ายประสาทเทียม (Neural Networks) ด้วยโปรแกรม Weka

No.	eur	jpy	krw	nikkei	91	e20	ngv	btc2	btc1	predictedbtc	btc
1	40.07	30.64	0.03	19,033.71	23.38	21.49	13.55	14,665.91	12,921.35	12288.98	17636.68
2	40.07	30.64	0.03	19,033.71	23.38	21.49	13.55	12,921.35	17636.68	12288.98	17761.99
3	40.07	30.64	0.03	19,033.71	23.38	21.49	13.55	17636.68	17761.99	12288.98	15000
4	40.09	31.06	0.0342	18450.98	23.38	21.49	13.55	17761.99	15000	13157.34	18489.98
5	39.75	30.95	0.0342	18374	23.38	21.49	13.55	15000	18489.98	13083.74	13245.55
6	39.65	31.26	0.0341	18,191.32	23.38	21.49	13.55	18489.98	13245.55	13039.69	17730.5
7	39.97	31.54	0.034	17,767.34	23.38	21.49	13.55	13245.55	17730.5	13084.86	19569.47
8	40.11	31.39	0.034	17,697.96	23.43	21.49	13.55	17730.5	19569.47	13111.83	15799.79
9	40.11	31.39	0.034	17,697.96	23.13	21.19	13.55	19569.47	15799.79	13076.09	16686.53
10	40.11	31.39	0.034	17,697.96	23.13	21.19	13.55	15799.79	16686.53	13076.09	17535.55
11	40.37	31.71	0.0338	17,697.96	23.13	21.19	13.55	16686.53	17535.55	13082.79	18587.36
12	40.17	31.64	0.0337	17,218.96	23.13	21.19	13.55	17535.55	18587.36	13014.79	19193.86
13	39.92	31.46	0.0337	17,715.63	23.13	21.19	13.55	18587.36	19193.86	12960.64	17364.53
14	40.21	31.64	0.0337	17,240.95	23.13	21.19	13.55	19193.86	17364.53	13023.26	18556.7
15	40.29	31.69	0.0336	17,147.11	23.13	21.19	13.55	17364.53	18556.7	13015.62	13076.23
16	40.29	31.69	0.0336	17,147.11	23.13	21.19	13.55	18556.7	13076.23	13015.62	15873.02
17	40.29	31.69	0.0336	17,147.11	23.13	21.19	13.55	13076.23	15873.02	13015.62	15601.1
18	40.27	31.71	0.0336	16,955.57	23.13	21.19	13.55	15873.02	15601.1	13011.53	15680.13
19	40.16	31.58	0.0336	17,048.37	22.73	20.79	13.55	15601.1	15680.13	12939.68	13159.86
20	40.44	32.03	0.0335	16,416.19	22.73	20.79	13.55	15680.13	13159.86	12976.14	17261.22
21	40.19	31.79	0.0335	16,017.26	22.73	20.79	13.55	13159.86	17261.22	12921.94	17667.84
22	39.76	31.32	0.0338	16,958.53	22.73	20.79	13.55	17261.22	17667.84	12904.18	13699.94
23	39.76	31.32	0.0338	16,958.53	22.73	20.79	13.55	17667.84	13699.94	12904.18	16666.67
24	39.76	31.32	0.0338	16,958.53	22.73	20.79	13.55	13699.94	16666.67	12904.18	16534.39
25	39.68	31.17	0.0339	17,110.91	22.73	20.79	13.55	16666.67	16534.39	12911.92	16260.16
26	39.71	31.18	0.0336	16,708.90	22.73	20.79	13.55	16534.39	16260.16	12841.85	16393.44
27	39.56	31.02	0.0335	17,163.92	22.73	20.79	13.55	16260.16	16393.44	12784.01	14534.88
28	39.7	30.92	0.0333	17,041.45	22.73	20.79	13.55	16393.44	14534.88	12763.13	15873.02
29	39.67	30.46	0.0335	17,518.30	22.73	20.79	13.55	14534.88	15873.02	12802.98	12232.53
30	39.67	30.46	0.0335	17,518.30	22.73	20.79	13.55	15873.02	12232.53	12802.98	15906.68
31	39.67	30.46	0.0335	17,518.30	22.73	20.79	13.55	12232.53	15906.68	12802.98	13873.16
32	39.27	30.2	0.0333	17,865.23	22.73	20.79	13.55	15906.68	13873.16	12667.11	14347.2
33	39.49	30.35	0.0333	17,750.68	22.73	20.79	13.55	13873.16	14347.2	12714.46	15467.9

ตัวอย่างผลการโปรแกรมพันธุ์กรรม (Genetic Programming) ด้วยโปรแกรม GPdotNET

TRAINING

DATA

Nr	X1	X2	X3	X4	X5	X6	X7	X8	X9	R1	R2	R3	R4	R5	R6	Y	Ygp
1	41.4	30.63	0.03	18264.22	25.88	24.48	13.5	8519.44	8270.16	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8620.69	8397.541821
2	41.6	30.68	0.03	17965.7	25.88	24.48	13.5	8270.16	8620.69	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8001	8427.126396
3	41.32	30.8	0.03	18026.48	25.88	24.48	13.5	8620.69	8001	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8474.58	8328.659754
4	41.28	30.64	0.03	18171.6	25.88	24.48	13.5	8001	8474.58	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8555.54	8223.238249
5	41.17	30.26	0.03	18432.27	25.88	24.48	13.5	8474.58	8555.54	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8289.34	8502.307559
6	41.28	30.45	0.03	18070.21	25.88	24.48	13.5	8555.54	8289.34	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	9733.94	8424.537804
7	41.28	30.45	0.03	18070.21	25.88	24.48	13.5	8289.34	9733.94	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8079.72	8935.967856
8	41.28	30.45	0.03	18070.21	25.88	24.48	13.5	9733.94	8079.72	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	9823.18	8934.836779
9	41.05	30.44	0.03	18070.21	25.88	24.48	13.5	8079.72	9823.18	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	7849	8868.265047
10	40.96	30.75	0.03	18070.21	25.88	24.48	13.5	9823.18	7849	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8105.15	8876.797086
11	40.91	30.94	0.03	18070.21	25.88	24.48	13.5	7849	8105.15	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	7975.19	7979.21776
12	41.36	30.99	0.03	17571.83	25.88	24.48	13.5	8105.15	7975.19	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8741.26	8052.738465
13	41.03	30.69	0.03	17880.51	26.38	24.98	13.5	7975.19	8741.26	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8219.18	8348.405853
14	41.03	30.69	0.03	17880.51	26.38	24.98	13.5	8741.26	8219.18	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	10078.39	8493.443977
15	41.03	30.69	0.03	17880.51	26.38	24.98	13.5	8219.18	10078.39	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	10033.44	9085.518586
16	41.27	30.93	0.03	17645.11	26.38	24.98	13.5	10078.39	10033.44	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	8454.8	9999.400229
17	41.76	31.23	0.03	16930.84	26.38	24.98	13.5	10033.44	8454.8	7.25829	6.49389	6.85987	9.14874	9.48616	4.00323	10323.47	9251.809855

ประวัติย่อของผู้วิจัย

ชื่อ – สกุล	นายเอ ดีหลี
วัน เดือน ปีเกิด	29 สิงหาคม 2521
สถานที่เกิด	จังหวัดกรุงเทพมหานคร
สถานที่อยู่ปัจจุบัน	22/4 หมู่ 5 ต.คลองนครเนื่องเขต อ.เมือง จ.ฉะเชิงเทรา รหัสไปรษณีย์ 24000
ประวัติการทำงาน	
พ.ศ.2547 – ปัจจุบัน	เจ้าหน้าที่งานเทคโนโลยีสารสนเทศ ฝ่ายสำนักผู้อำนวยการ โรงเรียนเซนต์หลุยส์ ฉะเชิงเทรา
ประวัติการศึกษา	
พ.ศ.2536	มัธยมศึกษาปีที่ 3 โรงเรียนโพธิสัมพันธ์พิทยาคาร จ.ชลบุรี
พ.ศ.2539	มัธยมศึกษาปีที่ 6 โรงเรียนโพธิสัมพันธ์พิทยาคาร จ.ชลบุรี
พ.ศ.2545	วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยราชภัฏราชนครินทร์ จ.ฉะเชิงเทรา
พ.ศ.2552	ประกาศนียบัตรบัณฑิตวิชาชีพครู มหาวิทยาลัยราชภัฏราชนครินทร์ จ.ฉะเชิงเทรา