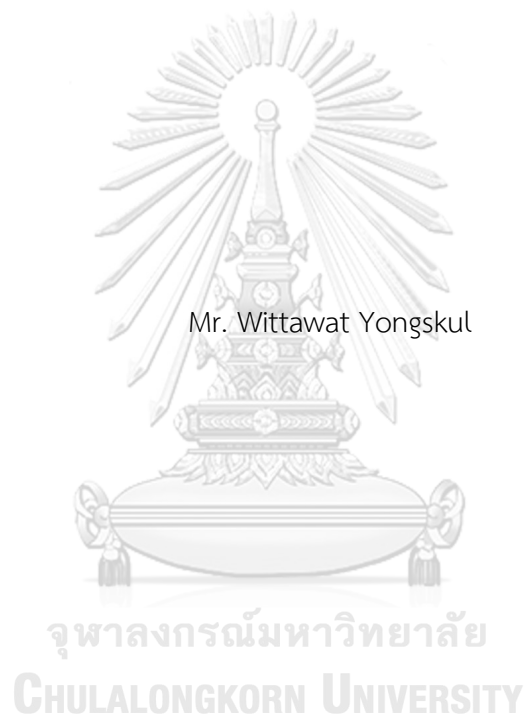


ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบ
เสริมกำลัง



สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศทางธุรกิจ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

DECISION SUPPORT SYSTEM FOR EQUITY TRADING IN THAI STOCKS MARKET USING
REINFORCEMENT LEARNING TECHNIQUE



An Independent Study Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Information Technology in Business

FACULTY OF COMMERCE AND ACCOUNTANCY

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อสารนิพนธ์	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุน ในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
โดย	นายวิทวัส ยงสกุล
สาขาวิชา	เทคโนโลยีสารสนเทศทางธุรกิจ
อาจารย์ที่ปรึกษาหลัก	รองศาสตราจารย์ ดร.จันทร์เจ้า มงคลนาวิน

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับสารนิพนธ์ฉบับนี้เป็น
ส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

คณะกรรมการสอบสารนิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ หม่อมหลวงสราวิกา อุณหนันท์)

..... อาจารย์ที่ปรึกษาหลัก
(รองศาสตราจารย์ ดร.จันทร์เจ้า มงคลนาวิน)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ธารทัศน์ โมกขมรรคกุล)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทวัส ยงสกุล : ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง. (DECISION SUPPORT SYSTEM FOR EQUITY TRADING IN THAI STOCKS MARKET USING REINFORCEMENT LEARNING TECHNIQUE) อ.ที่ปรึกษาหลัก : รศ. ดร.จันทร์เจ้า มงคลนาวิน

ณ ปัจจุบัน ผลกระทบจากปัจจัยด้านเศรษฐศาสตร์และการเงิน ทั้งในระดับประเทศและระดับโลก ทำให้การออมเงินโดยการฝากธนาคารพาณิชย์ ซึ่งถือเป็นหนึ่งในรูปแบบการลงทุนที่มีความเสี่ยงต่ำที่สุดนั้น มีความน่าสนใจน้อยลงอย่างมากเมื่อเทียบกับช่วงเวลาในอดีต และเมื่อประกอบกับปัจจัยในเรื่องของอัตราเงินเฟ้อแล้ว ทำให้ผู้คนในทุกระดับต่างต้องหาวิธีการออมเงินหรือการลงทุนที่จะทำให้เงินออมหรือเงินลงทุนของตนมีผลตอบแทนที่สูงกว่าอัตราเงินเฟ้อ ทั้งนี้เพื่อให้เงินออมหรือเงินลงทุนนั้นสามารถรักษามูลค่าไว้ได้ตามกาลเวลา แม้ว่าวิธีการลงทุนดังกล่าวจะหมายถึงการต้องยอมรับความเสี่ยงในการลงทุนที่มากขึ้นเมื่อเทียบกับการฝากเงินในธนาคารพาณิชย์

ในช่วงเวลาที่ผ่านมาความก้าวหน้าทางเทคโนโลยีและความทั่วถึงของระบบอินเทอร์เน็ตทำให้การเรียนรู้ของเครื่อง (Machine Learning) และ ปัญญาประดิษฐ์ (Artificial Intelligence) หรือ AI ถูกนำมาประยุกต์ในหลายอุตสาหกรรม ซึ่งรวมไปถึงตลาดเงิน-ตลาดทุน และ การลงทุนในตลาดหลักทรัพย์ด้วย

โครงการ "ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย" ถูกพัฒนาขึ้นโดยนำแนวคิดเรื่อง การเรียนรู้ของเครื่อง ประเภทการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) มาประยุกต์ในการสร้างระบบที่สามารถเรียนรู้จากข้อมูลซื้อขายหลักทรัพย์ในอดีตและนำตัวแบบที่ได้มาใช้แนะนำการซื้อขายหลักทรัพย์ที่สนใจได้ ระบบดังกล่าวประกอบไปด้วย 3 ระบบย่อย ได้แก่ ระบบนำเข้าข้อมูลราคาหลักทรัพย์ ระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง และระบบแนะนำการซื้อขายหลักทรัพย์ โดยใช้ระบบจัดการฐานข้อมูลแบบ NoSQL และใช้ Goole Colab เป็นเครื่องมือในการพัฒนา สาขาวิชา เทคโนโลยีสารสนเทศทางธุรกิจ ลายมือชื่อนิสิต

ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก

6380316326 : MAJOR INFORMATION TECHNOLOGY IN BUSINESS

KEYWORD:

Wittawat Yongskul : DECISION SUPPORT SYSTEM FOR EQUITY TRADING IN
THAI STOCKS MARKET USING REINFORCEMENT LEARNING TECHNIQUE.

Advisor: Assoc. Prof. JANJAO MONGKOLNAVIN, Ph.D.

At present, economic and financial factors at regional and global levels make savings in commercial banks, the lowest-risk investment, much less attractive compared to the past. Moreover, by considering an inflation rate, people at all levels look for saving and investment approaches that yield higher returns than the inflation rate. To maintain wealth over time by adopting these new investment methods, they have to accept higher risk than savings accounts.

Recently, the advancement of technologies and the broad internet coverage enabled the application of machine learning and artificial intelligence in various industries, including financial and stock market investment.

This project uses reinforcement learning, one area of machine learning techniques, to create a system to learn past stock trading data and use the learned model to make stock trading advice. This system consists of 3 subsystems: Stock Price Data Importing system, Reinforcement Learning Modeling system, and Stock Trading Advising system. The system uses NoSQL and Google Colab as database management and development tools.

Field of Study: Information Technology in Business Student's Signature

Academic Year: 2022 Advisor's Signature

กิตติกรรมประกาศ

โครงการพิเศษเรื่อง "ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง" ถูกพัฒนาขึ้นได้อย่างสำเร็จ ด้วยความช่วยเหลือ การให้คำแนะนำ และสนับสนุนจากหลายฝ่าย โดยเฉพาะอย่างยิ่ง ผู้ช่วยศาสตราจารย์ ดร.จันท์เจ้า มงคลนาวิน ผู้ซึ่งเป็นอาจารย์ที่ปรึกษาโครงการ โดยได้สละเวลาอันมีค่าในการให้คำแนะนำ ให้คำปรึกษา คอยตรวจสอบและแก้ไขจุดบกพร่องต่างๆ จนโครงการนี้เสร็จสมบูรณ์

ขอขอบพระคุณบิดาและมารดา ที่ให้การสนับสนุนทางด้านการเงินและคอยให้กำลังใจเสมอ ตลอดจนการเรียนและการพัฒนาโครงการ

ขอขอบพระคุณคณาจารย์ทุกท่าน ในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจ ที่ได้ประสิทธิ์ประสาทวิชาความรู้ต่างๆ ตลอดระยะเวลาที่ศึกษาในหลักสูตรนี้

สุดท้ายนี้หากโครงการพิเศษนี้มีข้อผิดพลาดหรือบกพร่องประการใด ผู้จัดทำขออภัยมา ณ ที่นี้ และหวังเป็นอย่างยิ่งว่าโครงการพิเศษนี้จะประโยชน์และแรงบันดาลใจแก่ผู้ที่สนใจต่อไป

วิทวัส ยงสกุล



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฌ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มาของโครงการ.....	1
1.2 วัตถุประสงค์ของโครงการ	3
1.3 ขอบเขตของโครงการ	4
1.4 วิธีการดำเนินงานโครงการ	6
1.4.1 การศึกษาและวิเคราะห์ระบบ (System Analysis)	6
1.4.2 การออกแบบระบบ (System Design).....	6
1.4.3 การพัฒนาระบบ (System Development).....	7
1.4.4 การทดสอบระบบ (System Testing).....	7
1.4.5 การจัดทำคู่มือการใช้งาน (User Document).....	7
1.5 เทคโนโลยีที่ใช้ในการพัฒนาระบบ.....	8
บทที่ 2 แนวคิดและเหตุผล.....	9
2.1 แนวคิดเกี่ยวกับการเรียนรู้แบบเสริมกำลัง.....	9
2.2 การเรียนรู้แบบเสริมกำลังในบริบทของการตัดสินใจซื้อขายหลักทรัพย์.....	12
บทที่ 3 โครงสร้างขององค์กรและการดำเนินการ	13

3.1 ประวัติองค์กร	13
3.2 โครงสร้างองค์กร.....	14
3.3 ลักษณะการดำเนินงานขององค์กร	15
3.4 ปัญหาที่เกิดขึ้นในปัจจุบัน	16
บทที่ 4 การพัฒนาระบบ	17
4.1 รายละเอียดของระบบ.....	17
4.2 การออกแบบระบบ	37
4.3 การออกแบบผลลัพธ์ (Output Design).....	40
4.4 การออกแบบข้อมูลนำเข้า (Input Design).....	41
4.5 การออกแบบส่วนติดต่อผู้ใช้งาน (User Interface Design).....	43
4.6 การออกแบบระบบรักษาความปลอดภัย.....	50
4.7 การติดตั้งและพัฒนาระบบ.....	50
บทที่ 5 บทสรุป ปัญหา และข้อเสนอแนะ	54
5.1 บทสรุป.....	54
5.2 ปัญหา.....	55
5.3 ข้อจำกัดของระบบ.....	56
5.4 ข้อเสนอแนะ	56
บรรณานุกรม.....	58
ภาคผนวก ก.....	61
พจนานุกรมข้อมูล.....	61
ภาคผนวก ข.....	62
เมนูการทำงานของระบบ.....	62
ภาคผนวก ค.....	67
ตัวอย่างรายงาน.....	67



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญตาราง

	หน้า
ตารางที่ 1: เทคโนโลยีที่ใช้ในการพัฒนาระบบ.....	8
ตารางที่ 2: Meta Data ของข้อมูลที่น่าเข้าสู่ระบบ	18
ตารางที่ 3: Use Case1 การเข้าสู่ระบบ (Login).....	22
ตารางที่ 4: Use Case2 การออกจากระบบ (Logout).....	24
ตารางที่ 5: Use Case3 การลงทะเบียนผู้ใช้งาน (Register)	25
ตารางที่ 6: Use Case4 การสร้างตัวแบบ (Create Model).....	26
ตารางที่ 7: Use Case5 การแสดงประวัติการสร้างตัวแบบ (View Model History).....	28
ตารางที่ 8: Use Case6 การแก้ไขตัวแบบ (Edit Model)	29
ตารางที่ 9: Use Case7 การลบตัวแบบ (Delete Model).....	31
ตารางที่ 10: Use Case8 การเรียกดูคำแนะนำ (Generate Advice)	32
ตารางที่ 11: User Table	61
ตารางที่ 12: Model Table.....	61

สารบัญรูปภาพ

	หน้า
รูปที่ 1: กระบวนการเรียนรู้แบบเสริมกำลัง	9
รูปที่ 2: โครงสร้างองค์กรของบริษัทต้นแบบ	14
รูปที่ 3: ตัวอย่างข้อมูลราคาหลักทรัพย์.....	18
รูปที่ 4: Use Case Diagram	21
รูปที่ 5: Sequence Diagram 1 การเข้าสู่ระบบ (Login)	33
รูปที่ 6: Sequence Diagram 2 การออกจากระบบ (Logout).....	33
รูปที่ 7: Sequence Diagram 3 การลงทะเบียนผู้ใช้งาน (Register)	34
รูปที่ 8: Sequence Diagram 4 การสร้างตัวแบบ (Create Model).....	34
รูปที่ 9: Sequence Diagram 5 การแสดงประวัติการสร้างตัวแบบ (View Model History)	35
รูปที่ 10: Sequence Diagram 6 การแก้ไขตัวแบบ (Edit Model).....	35
รูปที่ 11: Sequence Diagram 7 การลบตัวแบบ (Delete Model)	36
รูปที่ 12: Sequence Diagram 8 การเรียกดูคำแนะนำ (Generate Advice)	36
รูปที่ 13: Design Model ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง	37
รูปที่ 14: Entities Relationship Diagram	38
รูปที่ 15: สถาปัตยกรรมของระบบ System Architecture	39
รูปที่ 16: ตัวอย่างการแสดงผลในรูปแบบกราฟ	40
รูปที่ 17: ตัวอย่างการแสดงผลในรูปแบบข้อความ Logs	40
รูปที่ 18: Input Design ของการสมัครสมาชิก (Register).....	41
รูปที่ 19: Input Design ของการสร้างตัวแบบ (Create Model).....	42
รูปที่ 20: ตัวอย่างหน้าจอเข้าสู่ระบบ (Login).....	43

รูปที่ 21: ตัวอย่างหน้าจอลงทะเบียนผู้ใช้งาน (Register).....	43
รูปที่ 22: ตัวอย่างหน้าจอเลือกหลักทรัพย์และขอบเขตของข้อมูลราคาย้อนหลัง	44
รูปที่ 23: ตัวอย่างหน้าจอแบ่งข้อมูลสำหรับการฝึกและทดสอบตัวแบบ	44
รูปที่ 24: ตัวอย่างหน้าจอตั้งค่าตัวแปรต่าง ๆ ของตัวแบบ	45
รูปที่ 25: ตัวอย่างหน้าจอแสดงประวัติการสร้างตัวแบบ	46
รูปที่ 26: ตัวอย่างหน้าจอแก้ไขตัวแบบ	47
รูปที่ 27: ตัวอย่างหน้าจอเรียกดูคำแนะนำ	48
รูปที่ 28: ตัวอย่างหน้าจอลบตัวแบบ	49
รูปที่ 29: หน้าจอสมัครสมาชิก Deta.sh.....	50
รูปที่ 30: ชุดเครื่องมือของ Deta.sh	51
รูปที่ 31: หน้าจอ Google Colab.....	52
รูปที่ 32: ตัวอย่างการพัฒนาในระบบใน Google Colab.....	52
รูปที่ 33: Streamlit Library.....	53
รูปที่ 34: Login Menu	62
รูปที่ 35: Logout Menu	62
รูปที่ 36: Register Menu.....	63
รูปที่ 37: Select stock and time period.....	63
รูปที่ 38: Split Dataset	64
รูปที่ 39: Set Parameters	64
รูปที่ 40: View Model History Menu.....	65
รูปที่ 41: Edit Parameters.....	65
รูปที่ 42: Delete Model Menu.....	66
รูปที่ 43: Generate Advice Menu.....	66
รูปที่ 44: Training Logs	67

รูปที่ 45: Reward History Chart 67

รูปที่ 46: Test Result 68

รูปที่ 47: Generate Advice and historical actions..... 68



บทที่ 1

บทนำ

ในบทนี้กล่าวถึงความสำคัญและที่มาของโครงการ วัตถุประสงค์ของโครงการ ขอบเขตของโครงการ วิธีการดำเนินงานโครงการ เทคโนโลยีที่ใช้ในการพัฒนาระบบ ตลอดจนประโยชน์ที่จะได้รับจากโครงการนี้

1.1 ความสำคัญและที่มาของโครงการ

ณ ปัจจุบัน อัตราดอกเบี้ยเงินฝากธนาคารตกต่ำลงอย่างต่อเนื่อง โดยเป็นผลมาจากปัจจัยด้านเศรษฐศาสตร์และการเงิน ทั้งในระดับประเทศและระดับโลก ทำให้การออมเงินโดยการฝากธนาคารพาณิชย์ ซึ่งถือเป็นหนึ่งในรูปแบบการลงทุนที่มีความเสี่ยงต่ำที่สุดนั้น มีความน่าสนใจน้อยลงอย่างมากเมื่อเทียบกับช่วงเวลาในอดีต และเมื่อประกอบกับปัจจัยในเรื่องของอัตราเงินเฟ้อแล้ว ทำให้ผู้คนในทุกระดับต่างต้องหาวิธีการออมเงินหรือการลงทุนที่จะทำให้เงินออมหรือเงินลงทุนของตนมีผลตอบแทนที่สูงกว่าอัตราเงินเฟ้อ ทั้งนี้เพื่อให้เงินออมหรือเงินลงทุนนั้นสามารถรักษามูลค่าไว้ได้ตามกาลเวลา แม้ว่าวิธีการลงทุนดังกล่าวจะหมายถึงการต้องยอมรับความเสี่ยงในการลงทุนที่มากขึ้นเมื่อเทียบกับการฝากเงินในธนาคารพาณิชย์

การลงทุนในตราสารทุน ประเภทหุ้นสามัญ ของบริษัทที่จดทะเบียนในตลาดหลักทรัพย์นั้น เป็นหนึ่งในตัวเลือกทางการลงทุนที่มีมาอย่างยาวนานและเป็นที่รู้จักอย่างแพร่หลาย โดยมีแนวคิดที่สำคัญคือ การซื้อหุ้นสามัญ ทำให้ผู้ซื้อ มีสถานะเป็นเจ้าของบริษัทนั้นๆ รวมทั้งมีส่วนได้ส่วนเสียหรือสิทธิในทรัพย์สินหรือผลกำไรของบริษัท โดยผู้ที่ซื้อหุ้นสามัญในตลาดหลักทรัพย์นั้นจะมีโอกาสรับผลตอบแทนได้ 2 วิธีคือ

1. รับเงินปันผลจากบริษัทที่ถือหุ้น (Dividend) ซึ่งเป็นส่วนแบ่งในผลกำไรของบริษัทนั้นๆ ที่ได้มาจากการดำเนินงาน โดยส่วนแบ่งดังกล่าวจะจัดสรรให้ผู้ถือหุ้นทุกคนตามสัดส่วนของการถือหุ้น แต่มูลค่าที่ได้จะมากหรือน้อยนั้น ขึ้นอยู่กับผลประกอบการของบริษัท และ ดุลยพินิจของคณะกรรมการบริหารของบริษัท

2. ผลกำไรจากส่วนต่างของราคาหุ้น (Capital Gain) ซึ่งเป็นผลกำไรที่เกิดขึ้นเมื่อหุ้นสามัญดังกล่าวถูกขายในตลาดหลักทรัพย์ ณ ราคาที่สูงกว่า ราคาที่ผู้ลงทุนได้ซื้อเข้ามาในตอนแรก โดย

ผลตอบแทนในส่วนนี้เองที่ทำให้ การลงทุนในหุ้นสามัญในตลาดหลักทรัพย์นั้น ได้รับความนิยมเป็นอย่างมาก

ในระยะเวลากว่า 40 ปี ที่ตลาดหลักทรัพย์ไทยได้ถูกจัดตั้งขึ้น และเปิดให้ประชาชนและนิติบุคคลทุกภาคส่วนเข้ามามีส่วนร่วมในการลงทุน ซื้อ-ขาย หุ้นสามัญนั้น องค์ความรู้ใหม่ๆ ด้านการลงทุนในหุ้นสามัญและตลาดหลักทรัพย์ ก็ได้เกิดขึ้นมาอย่างต่อเนื่อง เช่น แนวคิดการเป็นนักลงทุนแบบเน้นคุณค่า (Value Investor) ที่เน้นวิเคราะห์ปัจจัยต่างๆ ที่กระทบต่อบริษัท ทั้งในระดับมหภาคและจุลภาค ไปจนถึงประเมินมูลค่าที่แท้จริงของบริษัทนั้นๆ ก่อนลงทุน หรือ แนวคิดนักลงทุนแบบเทคนิค (Technical) ที่เน้นศึกษาพฤติกรรมราคาของหุ้นที่สนใจ เพื่อทำนายทิศทางของราคาหุ้นในอนาคต แต่ด้วยความก้าวหน้าทางเทคโนโลยีและความทั่วถึงของระบบอินเทอร์เน็ต ทำให้แนวคิดเรื่อง การเรียนรู้ของเครื่อง (Machine Learning) และ ปัญญาประดิษฐ์ (Artificial Intelligence) หรือ AI ถูกนำมาประยุกต์ใช้ในหลายอุตสาหกรรม ซึ่งรวมไปถึงอุตสาหกรรมตลาดเงิน-ตลาดทุน และการลงทุนในตลาดหลักทรัพย์ด้วย

การเรียนรู้ของเครื่อง หรือ Machine Learning คือ การทำให้ระบบคอมพิวเตอร์สามารถประมวลผล คาดการณ์ และตัดสินใจปัญหาต่างๆ ด้วยตนเองผ่านการเรียนรู้จากชุดข้อมูลที่ถูกป้อนเข้าไป โดยการเรียนรู้นั้นจะเป็นไปตามขั้นตอนวิธี หรือ อัลกอริทึม (Algorithm) ที่ตั้งค่าไว้

อัลกอริทึมในการสร้างการเรียนรู้ของเครื่องสามารถแบ่งได้เป็น 3 กลุ่ม¹คือ

1. Supervised Learning คือการเรียนรู้แบบมีผู้สอน เช่น ในการพัฒนาระบบจำแนกระหว่างภาพรถยนต์และภาพรถมอเตอร์ไซด์ ข้อมูลภาพแต่ละภาพที่ถูกส่งเข้าไปในระบบเพื่อให้ระบบเรียนรู้นั้น จะถูกระบุไว้ตั้งแต่ต้นว่าภาพดังกล่าวนั้น คือภาพของรถยนต์หรือรถมอเตอร์ไซด์ จากนั้นให้ระบบหาความสัมพันธ์ระหว่างภาพที่วิเคราะห์กับชนิดของภาพที่ระบุไว้ จนเกิดเป็นโมเดลที่สามารถประมวลผลภาพที่ยังไม่เคยเห็น และคาดการณ์ได้ว่าภาพดังกล่าวเป็นภาพรถยนต์หรือรถมอเตอร์ไซด์

2. Unsupervised Learning คือการเรียนรู้แบบไม่มีผู้สอน เช่น ในการพัฒนาระบบจำแนกระหว่างภาพรถยนต์และภาพรถมอเตอร์ไซด์ ข้อมูลภาพแต่ละภาพที่ถูกส่งเข้าไปในระบบจะไม่ได้รับ

¹ Bhatt, S. (2018). "Reinforcement Learning 101." from <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>.

การระบุไว้ตั้งแต่ต้นว่าภาพดังกล่าวนี้ คือภาพของรถยนต์หรือรถมอเตอร์ไซค์ โดยระบบจะเรียนรู้ภาพต่างๆ และแบ่งภาพออกเป็นกลุ่ม โดยให้ภาพในแต่ละกลุ่มมีคุณลักษณะที่ใกล้เคียงกันมากที่สุด

3. Reinforcement Learning คือการเรียนรู้แบบลองผิดลองถูก ซึ่งไปประกอบไปด้วย สภาพแวดล้อม (Environment), สถานการณ์ (State) และ ระบบรางวัล (Reward) กล่าวคือ ระบบจะกระทำ (Action) ตัวเลือกต่างๆ ในสภาพแวดล้อมนั้นๆ โดยจะกระทำตามสถานการณ์ที่ถูกป้อนเข้าไปให้ระบบรับทราบ ซึ่งถ้าระบบตัดสินใจถูกต้องในสถานการณ์นั้นๆ ก็จะได้รางวัล ซึ่งการมองหารางวัลที่สูงที่สุด ทำให้ระบบมีการพัฒนาตนเองอย่างต่อเนื่อง

ด้วยลักษณะของการลงทุน ซื้อ-ขาย หุ้นสามัญในตลาดหลักทรัพย์ ที่มีปัจจัยด้านกำไรและขาดทุนเป็นเป้าหมายสำคัญ ประกอบกับมีข้อมูลในอดีตเพื่อใช้ในการเรียนรู้และตัดสินใจได้ว่า ในแต่ละสถานการณ์จะ ซื้อ (Buy) ถือ (Hold) หรือ ขาย (Sell) ทำให้เทคนิค การเรียนรู้ของเครื่อง ในกลุ่ม Reinforcement Learning เหมาะสมสำหรับนำไปประยุกต์ใช้ในการวิเคราะห์การลงทุน ซื้อ-ขาย หุ้นสามัญ ในตลาดหลักทรัพย์ จึงเป็นที่มาของการพัฒนาโครงการ ระบบสนับสนุนการตัดสินใจ สำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง (Decision Support System for Equity Trading in Thai Stocks Market using Reinforcement Learning Technique) เพื่อเป็นตัวช่วยผู้ลงทุนในการตัดสินใจ ซื้อ (Buy) ถือ (Hold) หรือ ขาย (Sell) หุ้นสามัญที่ตนสนใจ โดยใช้ข้อมูลราคาและข้อมูลประกอบอื่นๆ เช่น ปริมาณการซื้อขาย (Volume) ในการประมวลผลและให้คำแนะนำ

1.2 วัตถุประสงค์ของโครงการ

โครงการ “ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง” จัดทำขึ้นโดยมีวัตถุประสงค์ดังนี้

1. เพื่อสร้างระบบในการให้คำแนะนำ เพื่อสนับสนุนการตัดสินใจลงทุนในหลักทรัพย์
2. เพื่อประยุกต์ใช้เทคนิคการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ในการพัฒนาระบบสนับสนุนการตัดสินใจสำหรับซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย

1.3 ขอบเขตของโครงการ

โครงการ “ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง” เป็นการศึกษาและวิเคราะห์ข้อมูล โดยข้อมูลที่ใช้นั้นจะเป็นข้อมูลประวัติราคาย้อนหลังของหลักทรัพย์ที่สนใจเป็นหลัก โดยมี Yahoo Finance เป็นแหล่งข้อมูล ซึ่งข้อมูลดังกล่าวสามารถนำเข้าสู่ระบบที่พัฒนาได้โดยนำเข้าข้อมูลผ่าน Application Programming Interface (API) ที่ Yahoo Finance ได้พัฒนาไว้แล้ว โดยขอบเขตของโครงการมีดังนี้

1. การนำเข้าข้อมูลประวัติราคาหลักทรัพย์ที่จดทะเบียนในตลาดหลักทรัพย์ไทย จากฐานข้อมูล Yahoo Finance ซึ่งจะประกอบไปด้วยข้อมูล ราคาเปิด (Open), ราคาสูงสุด (High), ราคาต่ำสุด (Low), ราคาปิด (Close) โดยเป็นข้อมูลรายวัน
2. การสร้างตัวแบบในการวิเคราะห์ประวัติราคาหลักทรัพย์ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง โดยผู้ใช้ระบบสามารถกำหนดการตั้งค่าของพารามิเตอร์ (Parameter) ที่ส่งผลกับพฤติกรรมการเรียนรู้ของตัวแบบได้ เช่น
 - ค่าแกมมา (Gamma) หรือเรียกได้อีกชื่อหนึ่งว่า ค่า Discount Rate โดยจะมีค่าเป็นตัวเลขระหว่าง 0 ถึง 1 โดยหากค่าแกมมามีค่าสูง ตัวแบบจะให้ความสำคัญกับ รางวัล (Reward) ในระยะยาวมากกว่า และในทางตรงกันข้าม หากค่าแกมมามีค่าต่ำ ตัวแบบจะให้ความสำคัญกับ รางวัล ในระยะสั้น
 - ค่าแกมมามีบทบาทที่สำคัญที่สุดใน ฟังก์ชันคุณค่า (value function) ซึ่งเป็นฟังก์ชันสำหรับคำนวณผลรางวัลที่จะเกิดขึ้นในระยะยาว
 - ค่าเอปไซลอน (Epsilon) มีค่าเป็นตัวเลขระหว่าง 0 ถึง 1 โดยหากมีค่าเท่ากับ 1 ตัวแบบจะเลือกกระทำตัวเลือกต่างๆ เช่น การซื้อ (Buy), การถือ (Hold), การขาย (Sell) โดยการสุ่มอย่างสมบูรณ์โดยไม่สนใจรางวัลที่จะได้รับ (Absolute Exploration) ในขณะที่ถ้าค่าเอปไซลอน มีค่าเป็น 0 ตัวแบบจะเลือกการกระทำที่จะให้รางวัลมากที่สุดโดยสมบูรณ์ (Absolute Exploitation)²

² AlgoAddict (2019). "[Reinforcement Learning 101] ตะลุยคาสีในด้วย RL: Multi-Armed Bandit (1)- Epsilon Greedy." from <https://algoaddict.wordpress.com/2019/06/22/reinforcement-learning-101-beating-the-casino-multi-armed-bandit-problem-explore-exploit-delima-epsilon-greedy-and-sample-mean/>.

ค่าเอปไซลอน ถูกนำมาใช้ในการเรียนรู้แบบเสริมกำลังเพื่อแก้ปัญหา exploration vs exploitation dilemma กล่าวคือ เมื่อตัวกระทำ (Agent) ต้องตัดสินใจเลือกการกระทำ (Action) ในสถานะ (State) หนึ่ง ๆ นั้น ผู้กระทำควรเลือกตัดสินใจแบบสุ่มเพื่อสำรวจองค์ความรู้ให้มากขึ้น (exploration) หรือควรเลือกตัดสินใจเลือกการกระทำที่จะให้ผลรางวัลสูงสุดตามการเรียนรู้ที่เกิดขึ้นไปแล้ว (exploitation) ซึ่งทั้งสองทางเลือก มีผลดีและผลเสียที่ต่างกัน ค่าเอปไซลอนจะเข้ามาแก้ไขปัญหาดังกล่าวโดยทำหน้าที่เป็นความน่าจะเป็นที่ตัวกระทำจะเลือกตัดสินใจแบบสุ่ม ซึ่งจะทำให้ตัวกระทำมีโอกาสที่จะตัดสินใจในรูปแบบต่าง ๆ ดังนี้

- ตัวกระทำ มีความน่าจะเป็นเท่ากับ ค่าเอปไซลอน (ϵ) ที่จะตัดสินใจแบบสุ่มเพื่อ explore
 - ตัวกระทำ มีความน่าจะเป็นเท่ากับ $1 - \epsilon$ ที่จะตัดสินใจแบบ exploit
 - ค่าอัตราการลดลงของเอปไซลอน (Epsilon Decay)³ มีค่าเป็นตัวเลขระหว่าง 0 ถึง 1 มีหน้าที่ทำให้ค่าเอปไซลอน (Epsilon) ของตัวแบบมีค่าลดลงในทุก ๆ รอบของการเรียนรู้ข้อมูล ทั้งนี้เพื่อให้ตัวแบบเน้นเรียนรู้แบบ Exploration ในการเรียนรู้รอบแรก ๆ และเปลี่ยนนโยบายการเรียนรู้เป็น Exploitation ในการเรียนรู้รอบหลัง ๆ
 - จำนวนรอบการเรียนรู้ (Episode) คือจำนวนรอบที่ตัวแบบจะเรียนรู้จากข้อมูลชุดฝึก (Train Set) เช่น เมื่อ จำนวนรอบของการเรียนรู้ถูกกำหนดไว้ที่ 20 รอบ และข้อมูลชุดฝึกประกอบไปด้วยข้อมูลราคาหลักทรัพย์จำนวน 100 วัน แปลว่า ตัวกระทำ (Agent) ของระบบจะได้รับข้อมูลสถานะ (State) และตัดสินใจเลือกการกระทำ (Action) รอบละ 100 ครั้ง จำนวน 20 รอบติดต่อกัน
3. ผู้ใช้งานสามารถประเมินประสิทธิภาพของตัวแบบได้ รวมทั้งใช้ตัวแบบในการวิเคราะห์ และสร้างคำแนะนำในการซื้อขายหลักทรัพย์ที่สนใจ

³ Noparat, N. (2020). "มาทำความรู้จักกับ การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)." from <https://bigdata.go.th/big-data-101/introduction-to-reinforcement-learning/>.

1.4 วิธีการดำเนินงานโครงการ

โครงการ “ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง” มีขั้นตอนในการดำเนินโครงการดังนี้

1.4.1 การศึกษาและวิเคราะห์ระบบ (System Analysis)

1. ศึกษาลักษณะภาพรวมและการดำเนินงานของธุรกิจ รวมถึงวิเคราะห์ปัญหาที่เกิดขึ้นในการดำเนินธุรกิจ โดยสัมภาษณ์บุคลากรจากฝ่ายงานที่เกี่ยวข้อง และเพื่อรับรู้ความต้องการของผู้ใช้งาน
2. รวบรวมปัญหาที่เกิดขึ้นและสรุปความต้องการ เพื่อกำหนดขอบเขตในการพัฒนาระบบได้อย่างเหมาะสม
3. ศึกษาเทคโนโลยีต่างๆที่จะสามารถนำมาใช้ในการพัฒนาระบบได้ รวมทั้งคัดเลือกเทคโนโลยีที่เหมาะสมในการนำมาพัฒนาระบบสนับสนุนการตัดสินใจ

1.4.2 การออกแบบระบบ (System Design)

1. ออกแบบกระบวนการทำงานของระบบย่อยต่างๆ เช่น ระบบนำเข้าข้อมูลราคาหลักทรัพย์, ระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง และ ระบบแนะนำการซื้อขายหลักทรัพย์
2. ออกแบบหน้าจอ User Interface และออกแบบส่วนต่อประสานระหว่าง ระบบ Reinforcement Learning และ ระบบ User Interface
3. ออกแบบกระบวนการในการจัดการข้อมูล (Data Manipulation) ที่ครอบคลุมการนำเข้าข้อมูลจาก Yahoo Finance ผ่านส่วนต่อประสาน API และ ออกแบบกระบวนการประมวลผลข้อมูล (Data Preprocessing) เพื่อให้ข้อมูลอยู่ในรูปแบบที่พร้อมใช้สำหรับการวิเคราะห์บนระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง และ ระบบแนะนำการซื้อขายหลักทรัพย์

1.4.3 การพัฒนาระบบ (System Development)

1. พัฒนาระบบนำเข้าข้อมูลราคาหลักทรัพย์ ที่มีหน้าที่นำข้อมูล ประวัติราคาหลักทรัพย์ที่ผู้ใช้สนใจเข้าสู่ระบบ
2. พัฒนาระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
3. พัฒนาระบบแนะนำการซื้อขายหลักทรัพย์
4. พัฒนาระบบในส่วนของ User Interface ที่มีหน้าที่รับ Input จากผู้ใช้งานและแสดงผลลัพธ์ของระบบ (Output) ให้แก่ผู้ใช้งาน

1.4.4 การทดสอบระบบ (System Testing)

1. ทดสอบระบบ User Interface บนการทำงานในแต่ละ Use Case
2. ตรวจสอบความถูกต้องและความสมบูรณ์ของระบบ รวมทั้งสรุปรายการข้อผิดพลาดที่เกิดขึ้น
3. ปรับปรุงและแก้ไขข้อผิดพลาดที่เกิดขึ้นให้ถูกต้องสมบูรณ์

1.4.5 การจัดทำคู่มือการใช้งาน (User Document)

จัดทำคู่มือสำหรับการใช้งาน (User Manual) ซึ่งเป็นเอกสารที่อธิบายขั้นตอนการใช้งานของระบบที่พัฒนาขึ้น โดยนำเสนอในรูปแบบของหน้าจอแสดงผลพร้อมคำอธิบายประกอบรูปภาพ

1.5 เทคโนโลยีที่ใช้ในการพัฒนาระบบ

เทคโนโลยีที่ใช้ในการพัฒนาระบบสำหรับโครงการ “ระบบสนับสนุนการตัดสินใจ สำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง” มีรายละเอียด ดังนี้

ตารางที่ 1: เทคโนโลยีที่ใช้ในการพัฒนาระบบ

ด้าน Software:	
ระบบปฏิบัติการ	Windows 10 Education
ระบบจัดการฐานข้อมูล	NoSQL
เครื่องมือที่ใช้ในการออกแบบระบบ	diagrams.net
เครื่องมือที่ใช้ในการพัฒนาระบบ	Google Colab, Google Cloud Platform
เครื่องมือที่ใช้ในการแสดงผลระบบ	Microsoft Edge, Google Chrome
ภาษาที่ใช้ในการพัฒนาระบบ	Python, Tensorflow, Streamlit
ด้าน Hardware:	
หน่วยประมวลผลกลาง (CPU)	Intel(R) Core(TM) i7-3540M CPU @ 3.00GHz 3.00 GHz
หน่วยความจำ (Memory)	4.00 GB
หน่วยเก็บข้อมูล (Storage)	180 GB

บทที่ 2

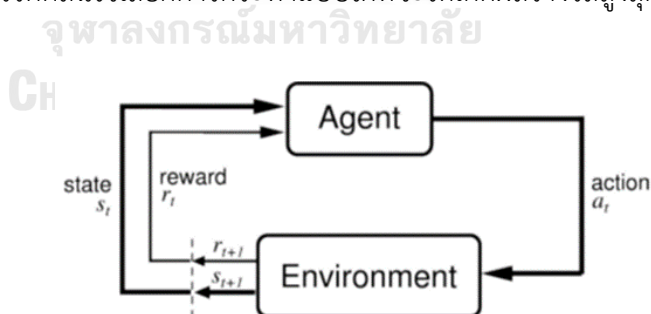
แนวคิดและเหตุผล

ใบบทนี้จะกล่าวถึงแนวคิดต่างๆ ที่ถูกนำมาประยุกต์ใช้ในการพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง

2.1 แนวคิดเกี่ยวกับการเรียนรู้แบบเสริมกำลัง

การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) คือการเรียนรู้ที่มีจุดมุ่งหมายโดยตรง และมุ่งพยายามให้การกระทำต่าง ๆ บรรลุตามจุดมุ่งหมายนั้นมากที่สุด โดยการตัดสินใจจะเกิดขึ้นในสภาพแวดล้อม (Environment) ที่จะมีการระบุและกำหนด สถานะ (State), การกระทำ (Action) และผลรางวัล (Reward) เอาไว้ เพื่อนำไปใช้ในกระบวนการเรียนรู้

กระบวนการเรียนรู้แบบเสริมกำลังนั้น เริ่มต้นที่ ตัวกระทำ (Agent) ที่ทำหน้าที่เป็นผู้ตัดสินใจของระบบ ได้รับ สถานะ (State) จากสภาพแวดล้อม (Environment) โดย ตัวกระทำ (Agent) ก็จะเลือก การกระทำ (Action) หนึ่ง ๆ ที่เป็นไปได้ จากนั้น ตัวกระทำ (Agent) จะได้รับ ข้อมูล สถานะ (State) ถัดไป พร้อมทั้ง ผลรางวัล (Reward) ของการตัดสินใจที่ผ่านมา โดย กระบวนการดังกล่าวจะดำเนินไปเรื่อย ๆ จนกระทั่ง ตัวกระทำ (Agent) เกิดการเรียนรู้ว่า ในสถานะหนึ่ง ๆ ที่เกิดขึ้นนั้น ควรตัดสินใจเลือกการกระทำแบบใดที่จะให้เกิดผลรางวัลสูงสุดในระยะยาว



รูปที่ 1: กระบวนการเรียนรู้แบบเสริมกำลัง

กระบวนการเรียนรู้แบบเสริมกำลังสามารถอธิบายได้ตามรูปที่ 1 ซึ่งมีองค์ประกอบต่าง ๆ ได้แก่

- state s_t คือ สถานะของสิ่งแวดล้อม ณ เวลา t
- Agent คือ ตัวกระทำในระบบ
- action a_t คือ การกระทำที่ตัวกระทำตัดสินใจเลือกกระทำ ณ เวลา t
- Environment คือ สภาพแวดล้อมของการเรียนรู้
- reward r_t คือ ผลรางวัลที่ตัวกระทำได้รับจากการกระทำ ณ เวลา t
- reward r_{t+1} คือ ผลรางวัลที่ตัวกระทำได้รับจากการกระทำ ณ เวลา $t+1$
- state s_{t+1} คือ สถานะของสิ่งแวดล้อม ณ เวลา $t+1$

ในการเรียนรู้แบบเสริมกำลังนั้น มี 4 องค์ประกอบย่อย⁴ที่ทำให้การเรียนรู้เกิดขึ้นได้ คือ

1. นโยบาย (Policy) คือหลักการตัดสินใจเลือกการกระทำ (Action) ถัดไปที่จะทำให้ได้รับผลรางวัล (Reward) สูงที่สุดโดยเฉลี่ยในระยะยาว โดยแสดงได้ตามสมการดังนี้

$$\pi(s) = \operatorname{argmax} g(s,a)$$

- $\pi(s)$ คือ policy เมื่อเกิดสถานะ s
- $g(s,a)$ คือ ผลเฉลี่ยรางวัลสะสม เมื่อเลือกการกระทำ a บนสถานะ s
- argmax คือการเลือกค่าที่สูงที่สุด

2. ฟังก์ชันรางวัล (reward function) คือฟังก์ชันสำหรับคำนวณผลรางวัลที่จะเกิดขึ้นเมื่อตัวกระทำ (Agent) ตัดสินใจเลือกการกระทำหนึ่ง ๆ (Action) ในสถานะใด ๆ (State) หมายความว่า การกระทำที่เหมือนกันบนสถานะที่ต่างกัน อาจนำไปสู่การได้รับผลรางวัลที่ต่างกันด้วย โดยหากการตัดสินใจนั้นถูกต้อง ก็จะได้รับผลรางวัลในระดับสูง แต่ถ้าหากการตัดสินใจนั้นถูกต้องน้อยกว่าหรือเป็นการตัดสินใจที่ผิด ก็จะได้รับผลรางวัลในระดับที่ต่ำลงหรือระดับติดลบ ทั้งนี้ขึ้นอยู่กับฟังก์ชันรางวัลที่ออกแบบไว้ว่าจะทำงานอย่างไร

ผลรวมของรางวัลตลอดการเรียนรู้สามารถสรุปได้ตามสมการดังนี้

$$R_{t:T} = g(s_t, a_t) + g(s_{t+1}, a_{t+1}) + g(s_{t+2}, a_{t+2}) + \dots + g(s_T, a_T)$$

- $R_{t:T}$ คือ ผลรวมของรางวัลตั้งแต่เวลา t ถึง เวลา T

⁴ Richard S Sutton, A. G. B. (2018). *Reinforcement Learning: An Introduction*, 2nd edition, MIT press.

- $g(s_t, a_t)$ คือ ฟังก์ชันรางวัลเมื่อเกิดการกระทำ a บนสถานะ s ณ เวลา t
- $g(s_{t+1}, a_{t+1})$ คือ ฟังก์ชันรางวัลเมื่อเกิดการกระทำ a บนสถานะ s ณ เวลา $t+1$
- $g(s_{t+2}, a_{t+2})$ คือ ฟังก์ชันรางวัลเมื่อเกิดการกระทำ a บนสถานะ s ณ เวลา $t+2$
- $g(s_T, a_T)$ คือ ฟังก์ชันรางวัลเมื่อเกิดการกระทำ a บนสถานะ s ณ เวลา T

3. ฟังก์ชันคุณค่า (value function)⁵ คือฟังก์ชันคำนวณผลรางวัลที่จะเกิดขึ้นในระยะยาว โดยมีความแตกต่างกับฟังก์ชันรางวัลคือ ฟังก์ชันรางวัลนั้นจะเกี่ยวข้องกับการกระทำหนึ่ง ๆ บนสถานะหนึ่ง ๆ เท่านั้น ในขณะที่ฟังก์ชันคุณค่า จะให้ความสำคัญกับการกระทำต่าง ๆ และ สถานะต่าง ๆ ที่เกิดขึ้นตลอดการเรียนรู้

ในฟังก์ชันคุณค่าจะมีการประยุกต์ใช้ค่าแกมมา (Gamma) เข้ามาทำหน้าที่เป็น อัตราการลดทอน (discount rate) สำหรับผลรางวัลที่จะเกิดขึ้นในช่วงเวลาต่าง ๆ ซึ่งค่าแกมมา (Gamma) จะมีค่าอยู่ในขอบเขตระหว่าง 0 ถึง 1 โดยการตั้งค่าแกมมาไว้ในระดับสูงจะหมายถึงการให้ความสำคัญกับ ผลรางวัล (Reward) ในระยะยาว ในขณะที่การตั้งค่าแกมมาไว้ในระดับที่ต่ำ จะหมายถึงการให้ความสำคัญกับผลรางวัลในระยะสั้น

ฟังก์ชันคุณค่าสามารถแสดงได้ตามสมการดังนี้

$$V^\pi(s) = \mathbb{E} \left[\sum_{t=1}^T \gamma^{t-1} r_t \right]$$

- $V^\pi(s)$ คือ ฟังก์ชันคุณค่าของนโยบาย π ที่เริ่มตั้งแต่สถานะ s
- γ^{t-1} คือ ค่าแกมมา ยกกำลังด้วย เวลา $t - 1$
- r_t คือ ผลรางวัล ณ เวลา t
- $\mathbb{E} \left[\sum_{t=1}^T \gamma^{t-1} r_t \right]$ คือ ค่าที่คาดหวังของ ผลรวมรางวัลที่เกิดขึ้นตั้งแต่ เวลา t จึงถึง T ซึ่งผ่านการ discount ด้วย ค่าแกมมา ยกกำลังด้วย เวลา $t - 1$ แล้ว

4. แบบจำลองของสภาพแวดล้อม (model of environment) ซึ่งจะทำหน้าที่สร้างสถานะ เพื่อใช้สำหรับการตัดสินใจ และเปลี่ยนหรือสร้างสถานะถัดไปเรื่อย ๆ ตลอดกระบวนการเรียนรู้

⁵ Marukatat, S. (2017). "Reinforcement Learning." from <https://medium.com/nectec/reinforcement-learning-43ea03c2e00e>.

2.2 การเรียนรู้แบบเสริมกำลังในบริบทของการตัดสินใจซื้อขายหลักทรัพย์

การตัดสินใจซื้อขายหลักทรัพย์นั้น มีลักษณะเหมือนกับการเรียนรู้แบบเสริมกำลังอยู่แล้ว กล่าวคือ นักลงทุน ซึ่งเปรียบเสมือน ผู้กระทำ (Agent) ในระบบการเรียนรู้แบบเสริมกำลัง จะต้องตัดสินใจเลือกการกระทำ (Action) ที่เป็นไปได้ ซึ่งได้แก่ การซื้อ (Buy), การถือ (Hold) และการขาย (Sell) หลักทรัพย์ โดยนักลงทุนจะตัดสินใจโดยพิจารณาจาก ประวัติราคาของหลักทรัพย์, ประมาณเงินสดที่มีอยู่ และตัวแปรอื่นๆ ซึ่งเปรียบเสมือน สถานะ (State) ที่นักลงทุนจะรับรู้ ณ เวลาหนึ่งๆ ซึ่งสุดท้ายแล้วสิ่งที่นักลงทุนต้องการคือผลกำไรที่สูงที่สุดจากการลงทุน เปรียบเสมือนการได้รับผลรางวัล (Reward) ที่สูงที่สุด

องค์ประกอบต่างๆ ของการเรียนรู้แบบเสริมกำลัง ในบริบทของการตัดสินใจซื้อขายหลักทรัพย์ สามารถสรุปได้ดังนี้

- State คือสถานะปัจจุบันของ สภาพแวดล้อม (Environment) เช่น ราคาหลักทรัพย์ ปัจจุบัน, ปริมาณการซื้อขายของหลักทรัพย์, ประมาณเงินสดที่มีอยู่
- Agent คือ นักลงทุนที่จะทำหน้าที่ตัดสินใจเลือกการกระทำที่เป็นไปได้ ตาม State ที่รับรู้จากสภาพแวดล้อม
- Action คือ การกระทำที่เป็นไปได้ในการตัดสินใจซื้อขายหลักทรัพย์ ซึ่งได้แก่ การซื้อ (Buy), การถือ (Hold) และการขาย (Sell)
- Reward คือ ผลกำไรจากการลงทุน ซึ่งจะมีค่าเป็นบวกเมื่อได้กำไร และมีค่าเป็นลบเมื่อเกิดผลขาดทุน จากการลงทุน

บทที่ 3

โครงสร้างขององค์กรและการดำเนินการ

ใบบทนี้จะกล่าวถึงประวัติองค์กรหรือบริษัทที่นำมาเป็นต้นแบบในการพัฒนาระบบสนับสนุน การตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง และ ยังกล่าวถึง โครงสร้างองค์กร ลักษณะการดำเนินงานขององค์กร และ ปัญหาที่เกิดขึ้น

3.1 ประวัติองค์กร

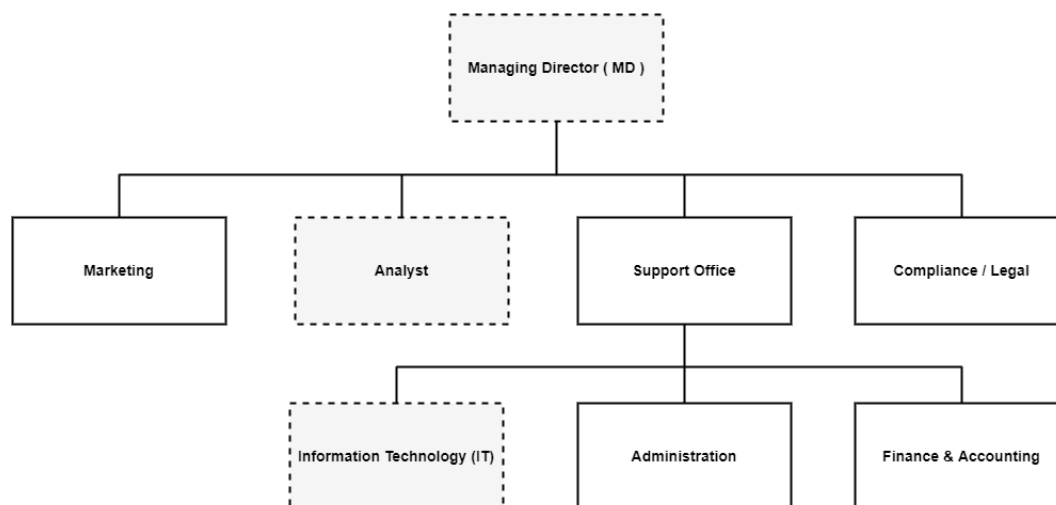
บริษัทที่เป็นต้นแบบ ดำเนินธุรกิจเกี่ยวกับการให้คำแนะนำด้านการลงทุนในตลาดหลักทรัพย์ แห่งประเทศไทย (Stocks Exchange of Thailand) โดยมีกิจกรรมทางธุรกิจที่สำคัญคือการผลิต งานวิจัยหรือรายงานวิเคราะห์ หลักทรัพย์ต่างๆ ที่ได้จดทะเบียนให้เป็นสินค้าเพื่อการซื้อขายในตลาด หลักทรัพย์ และเพื่อให้คำแนะนำในการลงทุนแก่ลูกค้าของบริษัท ซึ่งมีการสร้างรายได้หลากหลาย รูปแบบ เช่น


1. บริษัทได้รับค่าตอบแทนจากลูกค้าโดยการผลิตและส่งมอบงานวิจัยหรือรายงานวิเคราะห์ หลักทรัพย์ เพียงอย่างเดียว
2. บริษัทได้รับส่วนแบ่งกำไรจากการลงทุนจากลูกค้า เมื่อลูกค้าลงทุนตามคำแนะนำของ บริษัท เพียงอย่างเดียว
3. บริษัทได้รับทั้งค่าตอบแทนในการผลิตและส่งมอบงานวิจัยหรือรายงานวิเคราะห์ หลักทรัพย์ และ ส่วนแบ่งกำไรจากการลงทุนจากลูกค้า เมื่อลูกค้าลงทุนตามคำแนะนำของบริษัท

ทั้งนี้มูลค่าของค่าตอบแทนในการผลิตและส่งมอบงานวิจัยหรือรายงานวิเคราะห์หลักทรัพย์ และ อัตราส่วนแบ่งกำไรจากการลงทุน จะขึ้นอยู่กับผลการเจรจาต่อรองกับลูกค้าเป็นรายๆไป

ด้วยความก้าวหน้าของเทคโนโลยีด้านการเรียนรู้ของเครื่องและปัญญาประดิษฐ์ในปัจจุบัน บริษัทที่เป็นต้นแบบของโครงการนี้ จึงต้องการนำเทคโนโลยีดังกล่าวมาประยุกต์ใช้ในธุรกิจ โดยมีใช้ เพียงเพื่อสร้างมูลค่าเพิ่มให้กับสินค้าและบริการเดิม แต่ยังเป็นการเพิ่มศักยภาพในการเสนอผลิตภัณฑ์ ใหม่ๆ ให้กับลูกค้าของบริษัทด้วย

3.2 โครงสร้างองค์กร



 ฝ่ายที่เกี่ยวข้องกับโครงการพิเศษที่พัฒนา

รูปที่ 2: โครงสร้างองค์กรของบริษัทต้นแบบ

โครงสร้างองค์กรของบริษัทต้นแบบแสดงได้ดังรูปที่ 2 โดย Managing Director หรือ MD เป็นผู้มีอำนาจในการบริหารและตัดสินใจสูงสุด โดยแบ่งการดำเนินงานของบริษัทออกเป็น 4 กลุ่มงาน ได้แก่

1. กลุ่มงานด้าน Marketing

มีหน้าที่วางแผนและดำเนินการด้านการตลาด เพื่อให้กลุ่มลูกค้าเป้าหมายของบริษัท ทั้งที่เป็นบุคคล และ นิติบุคคล รับรู้และเข้าถึงสินค้าและบริการของบริษัท

นอกจากนี้กลุ่มงานดังกล่าวยังมีหน้าที่เป็นช่องทางหลักในการสื่อสารกับลูกค้า ซึ่งรวมถึงการนัดหมายกับลูกค้าเพื่อส่งมอบงานวิจัยและ/หรือรายงานการวิเคราะห์หลักทรัพย์ ด้วยเช่นกัน

2. กลุ่มงานด้าน Analyst

มีหน้าที่ผลิตงานวิจัยและ/หรือรายงานการวิเคราะห์หลักทรัพย์ ตามความต้องการของลูกค้า หรือ ตามที่หน่วยงานภายในร้องขอ

3. กลุ่มงานด้าน Support Office

มีหน้าที่สนับสนุนการดำเนินงานโดยรวมของบริษัท โดยแบ่งออกเป็น 3 กลุ่มงานย่อย ได้แก่

3.1 งานด้าน Administration

มีหน้าที่สนับสนุนการดำเนินงานโดยรวมของบริษัทในด้านงานธุรการและสวัสดิการของบุคลากรในบริษัท

3.2 งานด้าน Information Technology (IT)

มีหน้าที่สนับสนุนการดำเนินงานโดยรวมของบริษัทในด้านเทคโนโลยีสารสนเทศ

3.3 งานด้าน Finance & Accounting

มีหน้าที่บริหารจัดการด้านการเงินและการจัดทำบัญชีของบริษัท รวมไปถึงการจัดทำรายงานทางบัญชีของบริษัท

4) กลุ่มงานด้าน Compliance / Legal

มีหน้าที่ดำเนินการในงานด้านกฎหมายของบริษัท รวมทั้งให้คำปรึกษาด้านกฎหมายแก่บุคคลหรือหน่วยงานต่างๆ ภายในบริษัท

3.3 ลักษณะการดำเนินงานขององค์กร

บริษัทต้นแบบมีกิจกรรมทางธุรกิจที่สำคัญคือการผลิตงานวิจัยหรือรายงานวิเคราะห์หลักทรัพย์ และการให้คำแนะนำในการลงทุนแก่ลูกค้า โดยมีรายได้จากการดำเนินงานหลายรูปแบบ ดังนี้

1. บริษัทได้รับค่าตอบแทนจากลูกค้าโดยการผลิตและส่งมอบงานวิจัยหรือรายงานวิเคราะห์หลักทรัพย์ เพียงอย่างเดียว โดยงานวิจัยหรือรายงานวิเคราะห์นั้นมีขอบเขตตั้งแต่ขนาดเล็ก เช่น การวิจัยหรือวิเคราะห์หลักทรัพย์แบบรายตัว ไปจนถึงขอบเขตขนาดใหญ่ เช่น การวิจัยหรือวิเคราะห์อุตสาหกรรมหนึ่งๆ เป็นต้น ซึ่งขอบเขตของการวิจัยหรือวิเคราะห์นั้น ขึ้นอยู่กับความต้องการของลูกค้า และ การเจรจาตกลงกันระหว่างลูกค้าและบริษัท

2. บริษัทได้รับส่วนแบ่งกำไรจากการลงทุนจากลูกค้า เมื่อลูกค้าลงทุนตามคำแนะนำของบริษัท เพียงอย่างเดียว โดยอัตราส่วนแบ่งกำไรดังกล่าวจะขึ้นอยู่กับผลการเจรจาตกลงกันระหว่างลูกค้าและบริษัท

เนื่องจากลูกค้าของบริษัทมิได้ลงทุนกับบริษัทโดยตรง ดังนั้นเพื่อความเป็นธรรมและโปร่งใสสำหรับทั้งลูกค้าและบริษัท จะต้องมีการสร้างค่าดัชนี (Index) หนึ่งๆ ขึ้นมาเพื่อเป็นตัวแทนของคำแนะนำด้านการลงทุนที่บริษัทได้นำเสนอให้แก่ลูกค้า โดยดัชนีดังกล่าวจะเป็นเครื่องมือชี้วัดว่าคำแนะนำด้านการลงทุนของบริษัททำให้ลูกค้าเกิดผลกำไรได้หรือไม่

3. บริษัทได้รับทั้งค่าตอบแทนในการผลิตและส่งมอบงานวิจัยหรือรายงานวิเคราะห์หลักทรัพย์ และ ส่วนแบ่งกำไรจากการลงทุนจากลูกค้า เมื่อลูกค้าลงทุนตามคำแนะนำของบริษัท

3.4 ปัญหาที่เกิดขึ้นในปัจจุบัน

เนื่องจากการวิเคราะห์หลักทรัพย์นั้นจะกระทำโดยนักวิเคราะห์เป็นหลัก ซึ่งอาจมีข้อบกพร่องตามธรรมชาติของมนุษย์เช่น อคติในการวิเคราะห์ที่เกิดจากการยึดติดกับทฤษฎีเดิมๆ ที่อาจสามารถอธิบายพฤติกรรมราคาของหลักทรัพย์ในอดีตได้ดี แต่ในปัจจุบันทฤษฎีดังกล่าวอาจมีประสิทธิภาพหรือความสามารถในการอธิบายพฤติกรรมราคาของหลักทรัพย์นั้นๆ ที่น้อยลง

นอกจากนี้ ด้วยข้อจำกัดในการทำ Peer Review โดยใช้บุคลากรในองค์กรเท่านั้น ทำให้ขอบเขตขององค์ความรู้ที่ใช้ในการวิเคราะห์และวิจารณ์ทฤษฎีหรือคำแนะนำในการลงทุนนั้น มีขอบเขตที่จำกัด ซึ่งนำไปสู่ความไม่หลากหลายในการวิเคราะห์หรือวิจัยหลักทรัพย์

บทที่ 4

การพัฒนาระบบ

ในบทนี้จะกล่าวถึงรายละเอียดของระบบ การออกแบบระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง รวมถึงการออกแบบระบบคลังข้อมูลเพื่อสนับสนุนระบบดังกล่าว

4.1 รายละเอียดของระบบ

1. ขอบเขตการทำงาน (System Scope)

“ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง” ประกอบไปด้วย 3 ระบบย่อยคือ

1.1 ระบบนำเข้าข้อมูลราคาหลักทรัพย์

ระบบนี้เป็นระบบสำหรับนำเข้าข้อมูลประวัติราคาของหลักทรัพย์ที่ผู้ใช้สนใจเข้าสู่ระบบ ซึ่งจะดำเนินการโดยใช้ส่วนต่อประสาน yfinance ที่ถูกพัฒนาไว้แล้วโดย Yahoo Finance ซึ่งส่วนต่อประสานดังกล่าวจะช่วยให้ระบบที่กำลังพัฒนานี้สามารถเข้าถึงข้อมูลประวัติราคาหลักทรัพย์ที่ลงทะเบียนในตลาดหลักทรัพย์ไทยได้ทุกรายการ และ ยังรวมถึงการปรับปรุงข้อมูลที่ได้นำมาสู่ระบบ ให้อยู่ในรูปแบบที่เหมาะสม เช่น การปรับเปลี่ยนชนิดของข้อมูล (Data Types), การลบข้อมูลหรือฟิลด์ (Field) ของข้อมูลที่ไม่เกี่ยวข้อง, การเพิ่มฟิลด์ของข้อมูลใหม่ที่คำนวณมาจากฟิลด์เดิม (Calculated Field)

ข้อมูล (Data) ที่จะถูกนำเข้ามาสู่ระบบคือข้อมูลประวัติราคาของหลักทรัพย์หนึ่ง ๆ ซึ่งประกอบไปด้วยฟิลด์ของข้อมูลจำนวน 7 คอลัมน์ (Column) ได้แก่ วันที่ (Date), ราคาเปิด (Open), ราคาสูงสุด (High), ราคาต่ำสุด (Low), ราคาปิด (Close), ราคาปิดที่ปรับปรุงแล้ว (Adj Close) และ ปริมาณการซื้อขาย (Volume) โดยข้อมูล 1 รายการ (Record) เทียบเท่ากับข้อมูลของราคาหลักทรัพย์ 1 วัน โดยมีตัวอย่างแสดงดังรูปที่ 3

	Open	High	Low	Close	Adj Close	Volume
Date						
2022-01-04	230.0	232.0	227.0	228.0	223.972000	5137200
2022-01-05	227.0	227.0	224.0	226.0	222.007339	5786600
2022-01-06	223.0	224.0	221.0	221.0	217.095673	8504200
2022-01-07	221.0	222.0	215.0	221.0	217.095673	12073900
2022-01-10	221.0	221.0	218.0	219.0	215.131012	6136300
2022-01-11	221.0	226.0	221.0	224.0	220.042679	8143300
2022-01-12	225.0	226.0	221.0	222.0	218.078003	3214800
2022-01-13	223.0	224.0	218.0	220.0	216.113342	4652800
2022-01-14	220.0	222.0	219.0	220.0	216.113342	6093900
2022-01-17	219.0	220.0	217.0	218.0	214.148666	5103000

รูปที่ 3: ตัวอย่างข้อมูลราคาหลักทรัพย์

จากรูปที่ 3 ซึ่งเป็นตัวอย่างของข้อมูลที่จะถูกนำเข้าสู่ระบบนั้น แสดงให้เห็นถึงข้อมูลประวัติราคาของหลักทรัพย์หนึ่ง ๆ จำนวน 10 วัน โดยมีฟิลด์ของข้อมูลจำนวน 7 คอลัมน์ และมีฟิลด์ วันที่ (Date) ทำหน้าที่เป็น Index Column ของชุดข้อมูล

ตารางที่ 2: Meta Data ของข้อมูลที่น่าเข้าสู่ระบบ

ชื่อฟิลด์	ชนิดของข้อมูล	รูปแบบของข้อมูล	คำอธิบาย	หน่วย
Date	datetime	YYYY-MM-DD	วันที่บันทึกข้อมูลราคา	
Open	float		ราคาเปิดของหลักทรัพย์	บาท
High	float		ราคาสูงสุดที่เกิดการซื้อขายในวัน ของหลักทรัพย์	บาท
Low	float		ราคาต่ำสุดที่เกิดการซื้อขายในวัน ของหลักทรัพย์	บาท
Close	float		ราคาปิดของหลักทรัพย์	บาท
Adj Close	float		ราคาปิดของหลักทรัพย์ที่ปรับปรุงแล้ว ตาม Corporate Action ที่เกิดขึ้น (เช่นการปันผล, การแตกพาร์ เป็นต้น)	บาท
Volume	int		ปริมาณการซื้อขายในวัน ของหลักทรัพย์	หุ้น

1.2 ระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง

ระบบนี้เป็นระบบสำหรับสร้างตัวแบบที่จะนำมาแนะนำการซื้อขายหลักทรัพย์ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง โดยตัวแบบดังกล่าวจะได้รับการตั้งค่าพารามิเตอร์ (Parameters) ต่าง ๆ โดยผู้ใช้งาน และถูกฝึก (Train) ด้วยข้อมูลประวัติราคาของหลักทรัพย์ ที่ถูกนำเข้ามาโดย ระบบนำเข้าข้อมูลราคาหลักทรัพย์

ระบบนี้ยังครอบคลุมถึงการทดสอบตัวแบบดังกล่าวบนข้อมูลชุดทดสอบ (Test Set) ที่ผู้ใช้สามารถกำหนดเองได้ เพื่อให้ผู้ใช้สามารถประเมินประสิทธิภาพและความแม่นยำของตัวแบบได้ รวมทั้งครอบคลุมถึงการบันทึกตัวแบบไว้ใช้ในภายหลังและการลบตัวแบบที่ไม่ต้องการออกจากระบบด้วย

ระบบนี้ทำงานโดยรับ input จากผู้ใช้งานในเรื่องการตั้งค่าพารามิเตอร์ต่าง ๆ ในการเรียนรู้แบบเสริมกำลัง (เช่น ค่าเกมมา, ค่าเอปไซลอน, จำนวนรอบของการเรียนรู้) และ ข้อมูลประวัติราคาย้อนหลังของหลักทรัพย์ในช่วงเวลาที่ผู้ใช้งานเลือก ที่มีจุดประสงค์เพื่อใช้เป็นข้อมูลสำหรับฝึก (Train Set)

จากนั้น ตัวแบบจะถูกสร้างขึ้นตามค่าพารามิเตอร์ที่ได้รับมา ส่วนข้อมูลสำหรับฝึก (Train Set) จะถูกนำไปใช้ในการเรียนรู้ โดยทำหน้าที่เป็น สถานะ (State) ที่สภาพแวดล้อม (Environment) จะทยอยป้อนให้ตัวแบบ เพื่อให้ตัวแบบเลือกตัดสินใจไปที่สถานะ หรือก็คือการที่ตัวแบบเลือกตัดสินใจต่อราคาหลักทรัพย์ในแต่ละวันในข้อมูลชุดฝึก โดยในแต่ละการตัดสินใจก็จะเกิดผลรางวัลแก่ตัวแบบ ตามผลกำไรหรือขาดทุนที่เกิดขึ้น

เมื่อข้อมูลรายการสุดท้ายของข้อมูลชุดฝึกถูกป้อนให้กับตัวแบบแล้ว สภาพแวดล้อมจะป้อนข้อมูลรายการแรกของชุดฝึกให้กับตัวแบบใหม่อีกครั้ง เพื่อเป็นการเริ่มการเรียนรู้รอบถัดไป

กระบวนการดังกล่าวจะเกิดขึ้นแบบซ้ำ ๆ จนกว่าจำนวนรอบการเรียนรู้จะครบตามการตั้งค่าที่ผู้ใช้งานได้ตั้งค่าไว้ ซึ่งเมื่อกระบวนการเรียนรู้เสร็จสิ้นแล้ว จะได้มาซึ่ง output ของระบบคือ ตัวแบบที่ผ่านการฝึกแล้วและพร้อมนำไปใช้งานเพื่อตัดสินใจเลือก action ที่เหมาะสม บนข้อมูลราคาในช่วงเวลาอื่น ๆ ของหลักทรัพย์เดียวกันที่ถูกนำมาใช้ฝึกตัวแบบโดยข้อมูลราคาดังกล่าวจะอยู่นอกเหนือจากข้อมูลชุดฝึก

1.3 ระบบแนะนำการซื้อขายหลักทรัพย์

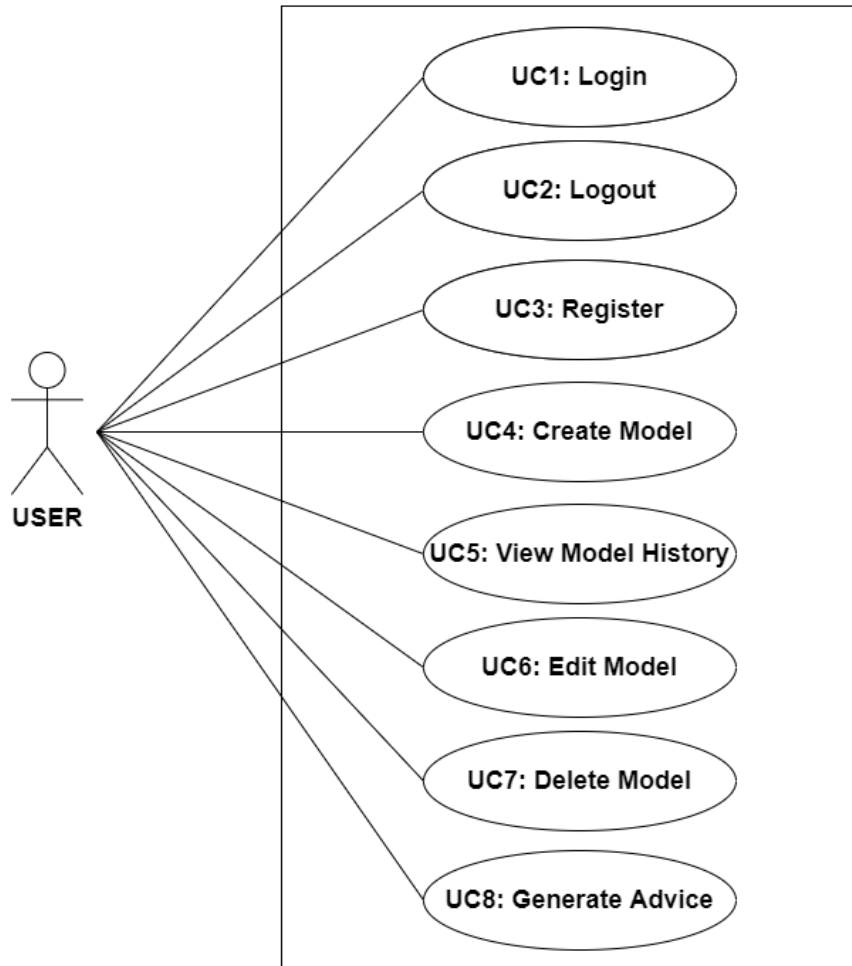
ระบบนี้เป็นระบบสำหรับนำตัวแบบที่ถูกสร้างและฝึกแล้วใน ระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มาให้คำแนะนำการซื้อขายหลักทรัพย์เดียวกันที่ได้นำข้อมูลประวัติราคามาใช้ในการฝึกตัวแบบ

ระบบแนะนำการซื้อขายหลักทรัพย์ ทำงานโดยรับ input เป็นข้อมูลประวัติราคาของหลักทรัพย์ในช่วงเวลาที่ผู้ใช้งานต้องการทราบคำแนะนำในการลงทุน โดยข้อมูลดังกล่าวจะถูกป้อนเข้าไปยังตัวแบบที่ได้ถูกสร้างและฝึกไว้แล้วใน ระบบสร้างตัวแบบด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง ข้อมูลที่ถูกป้อนเข้าไปนี้จะทำหน้าที่เป็น สถานะ (State) ใน สภาพแวดล้อม (Environment) ของการเรียนรู้แบบเสริมกำลัง

จากนั้น ตัวแบบดังกล่าวจะแสดงผล เป็น การกระทำ (Action) ที่ตัวแบบประเมินว่าจะได้ ผลรางวัล (Reward) ที่สูงที่สุดใน State นั้นๆ เช่น ควรซื้อ (Buy), ควรถือ (Hold) หรือ ควรขาย (Sell) โดยอาศัยการคำนวณที่ฝึกมาแล้วด้วยข้อมูลชุด Train Set ซึ่งการกระทำที่ตัวแบบตัดสินใจเลือกนี้ ถือเป็น output ของ ระบบแนะนำการซื้อขายหลักทรัพย์ ด้วยเช่นกัน

2. แผนภาพแสดงการทำงานของผู้ใช้ระบบ (Use Case Diagram)

ระบบสนับสนุนการตัดสินใจซื้อขายตราสารทุนในตลาดหุ้นไทย
ด้วยเทคนิคอินฟอร์สเมนต์การเรียนรู้



รูปที่ 4: Use Case Diagram

3. คำอธิบายแผนภาพแสดงการทำงานของผู้ใช้งาน

ตารางที่ 3: Use Case1 การเข้าสู่ระบบ (Login)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการเข้าใช้งานระบบ
Precondition	มีข้อมูลของผู้ใช้งานอยู่ในระบบ
Success Guarantee	ผู้ใช้งานเข้าสู่ระบบได้
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งานกรอกข้อมูล ชื่อผู้ใช้ (Username) และ รหัสผ่าน (Password) เพื่อเข้าสู่ระบบ 2. ระบบค้นหาข้อมูล ชื่อผู้ใช้ (Username) และ ตรวจสอบความถูกต้องของ รหัสผ่าน (Password) 3. ระบบแสดงผลการลงชื่อเข้าใช้ระบบสำเร็จ

Extensions	<p>2a. กรณีที่ไม่มี ชื่อผู้ใช้ (Username) ในระบบ</p> <ol style="list-style-type: none"> 1. ระบบแสดงการแจ้งเตือนว่าข้อมูลผู้ใช้งานไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน 3. ระบบแสดงหน้าจอให้ผู้ใช้งานกรอกข้อมูล ชื่อผู้ใช้ (Username) และ รหัสผ่าน (Password) เพื่อลงชื่อเข้าใช้อีกครั้ง <p>2b. กรณีที่รหัสผ่าน (Password) ไม่ถูกต้อง</p> <ol style="list-style-type: none"> 1. ระบบแสดงการแจ้งเตือนว่าข้อมูลผู้ใช้งานไม่ถูกต้องผู้ใช้งานรับทราบการแจ้งเตือน 2. ระบบแสดงหน้าจอให้ผู้ใช้งานกรอกข้อมูล ชื่อผู้ใช้ (Username) และ รหัสผ่าน (Password) เพื่อลงชื่อเข้าใช้อีกครั้ง <p>2c. กรณีผู้ใช้งานกรอกข้อมูลในการเข้าสู่ระบบไม่ครบถ้วน</p> <ol style="list-style-type: none"> 1. ระบบแสดงการแจ้งเตือนข้อมูลที่ยังไม่ได้กรอก 2. ผู้ใช้งานกรอกข้อมูลที่ยังกรอกไม่ครบตามการแจ้งเตือนของระบบ
------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ตารางที่ 4: Use Case2 การออกจากระบบ (Logout)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการออกจากระบบหลังการใช้งานเสร็จสิ้น
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ
Success Guarantee	ผู้ใช้งานออกจากระบบได้สำเร็จ
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งานแจ้งความต้องการออกจากระบบ 2. ระบบแสดงผลการออกจากระบบสำเร็จ

ตารางที่ 5: Use Case3 การลงทะเบียนผู้ใช้งาน (Register)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการให้มีชื่อผู้ใช้งานที่ต้องการอยู่ในระบบ
Precondition	ไม่มีข้อมูลผู้ใช้งานอยู่ในระบบ
Success Guarantee	มีข้อมูลผู้ใช้งานในระบบ
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งานระบบแจ้งความต้องการสมัครสมาชิก 2. ผู้ใช้งานกรอกข้อมูล ชื่อผู้ใช้ (Username), รหัสผ่าน (Password) และชื่อ (name) ที่ต้องการ 3. ระบบตรวจสอบข้อมูล ชื่อผู้ใช้ (Username) เพื่อป้องกันการมีชื่อผู้ใช้งาน (Username) ซ้ำซ้อนกับที่มีอยู่ในระบบอยู่แล้ว 4. ระบบแสดงผลการลงทะเบียนสำเร็จ
Extensions	<p>3a. ชื่อผู้ใช้งาน (Username) ซ้ำซ้อนกับข้อมูลที่มีอยู่ในระบบ</p> <ol style="list-style-type: none"> 1. ระบบแสดงการแจ้งเตือน ชื่อผู้ใช้งาน (Username) ซ้ำซ้อน 2. ผู้ใช้งานรับทราบการแจ้งเตือน 3. ระบบแสดงหน้าจอให้ผู้ใช้งานกรอกข้อมูล ชื่อผู้ใช้ (Username) และ รหัสผ่าน (Password) เพื่อลงทะเบียนผู้ใช้งานอีกครั้ง

ตารางที่ 6: Use Case4 การสร้างตัวแบบ (Create Model)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการสร้างตัวแบบในการวิเคราะห์หลักทรัพย์
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ
Success Guarantee	ตัวแบบถูกสร้างและแสดงอยู่ในระบบ
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งานกรอกข้อมูล ชื่อหลักทรัพย์ (stockQuote), วันที่เริ่ม (startDate) และ วันที่สิ้นสุด (endDate) ของข้อมูลที่ต้องการนำมาใช้สร้างตัวแบบ 2. ระบบแสดงข้อมูลประวัติราคาของหลักทรัพย์นั้นๆ ตั้งแต่ วันที่เริ่ม (startDate) จนถึง วันที่สิ้นสุด (endDate) 3. ผู้ใช้งานกำหนดค่าการแบ่งช่วงเวลาออกเป็น Train Set และ Test Set 4. ผู้ใช้งานกรอก ชื่อโมเดล (modelName) และ กำหนดค่า gamma, epsilon, epsilon_min, epsilon_decay, episode_count, initial_balance และ commission_fee 5. ผู้ใช้งานยืนยันการสร้างตัวแบบ 6. ระบบสร้างตัวแบบและแสดงผล

Extensions	<p>2a. กรณีชื่อหลักทรัพย์ (Stock Quote) ไม่ถูกต้อง</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ชื่อหลักทรัพย์ (Stock Quote) ไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูล ชื่อหลักทรัพย์ (Stock Quote) ใหม่ <p>2b. กรณีที่ผู้ใช้กรอกข้อมูลวันที่ผิดพลาด (เช่น กรอกวันที่ในอนาคต, กรอกวันที่เริ่มสลับกับวันที่สิ้นสุด)</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ข้อมูลวันที่ไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูล วันที่ ใหม่ <p>4a. กรณีผู้ใช้กรอกค่า γ, ϵ, ϵ_{\min} หรือ ϵ_{decay} ไม่อยู่ในขอบเขต 0 ถึง 1</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ค่าที่กรอกอยู่ในขอบเขตที่ไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูลใหม่ <p>4b. กรณีผู้ใช้กรอกค่า initial_balance น้อยกว่า 0</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน initial_balance ต้องมีค่ามากกว่า 0 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูล initial_balance ใหม่ <p>4c. กรณีผู้ใช้กรอกค่า commission_fee ไม่อยู่ในขอบเขต 0 ถึง 1</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ค่าที่กรอกอยู่ในขอบเขตที่ไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูลใหม่
------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ตารางที่ 7: Use Case5 การแสดงประวัติการสร้างตัวแบบ (View Model History)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการดูประวัติการสร้างตัวแบบ
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ
Success Guarantee	ตัวแบบ, ข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบถูกแสดง
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งาน แจ้งความต้องการดูประวัติการสร้างตัวแบบ 2. ระบบแสดงข้อมูล ประวัติการสร้างตัวแบบ 3. ผู้ใช้งานเลือกตัวแบบที่ต้องการ 4. ระบบแสดงข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบที่ผู้ใช้งานเลือก
Extensions	<p>2a. กรณีไม่มีประวัติการสร้างตัวแบบ</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือนการไม่พบข้อมูลย้อนหลัง 2. ผู้ใช้งานรับทราบการแจ้งเตือน

ตารางที่ 8: Use Case6 การแก้ไขตัวแบบ (Edit Model)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการแก้ไขการตั้งค่าของตัวแบบที่ถูกสร้างไว้แล้ว
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ มีตัวแบบถูกสร้างไว้แล้วในระบบ
Success Guarantee	ตัวแบบได้รับการตั้งค่าใหม่ตามที่ต้องการ
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งาน แจ้งความต้องการแก้ไขตัวแบบ 2. ระบบแสดงข้อมูล ประวัติการสร้างตัวแบบ 3. ผู้ใช้งานเลือกตัวแบบที่ต้องการแก้ไข 4. ระบบแสดงข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบที่ผู้ใช้งานเลือก 5. ผู้ใช้งานกรอกข้อมูลการตั้งค่าตัวแบบใหม่ ที่ต้องการ เช่น ชื่อ โมเดล (modelName), gamma, epsilon, epsilon_min, epsilon_decay, episode_count, initial_balance หรือ commission_fee 6. ผู้ใช้งานยืนยันการแก้ไขตัวแบบ 7. ระบบสร้างตัวแบบตามค่าพารามิเตอร์ที่ผู้ใช้กำหนดและแสดงผล
Extensions	<p>2a. กรณีไม่มีประวัติการสร้างตัวแบบ</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือนการไม่พบข้อมูลย้อนหลัง 2. ผู้ใช้งานรับทราบการแจ้งเตือน <p>5a. กรณีผู้ใช้กรอกค่า gamma, epsilon, epsilon_min หรือ</p>

	<p>epsilon_decay ไม่อยู่ในขอบเขต 0 ถึง 1</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ค่าที่กรอกอยู่ในขอบเขตที่ไม่ถูกต้อง 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูลใหม่ <p>5b. กรณีผู้ใช้กรอกค่า initial_balance น้อยกว่า 0</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน initial_balance ต้องมีค่ามากกว่า 0 2. ผู้ใช้งานรับทราบการแจ้งเตือน และ กรอกข้อมูล initial_balance ใหม่
--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



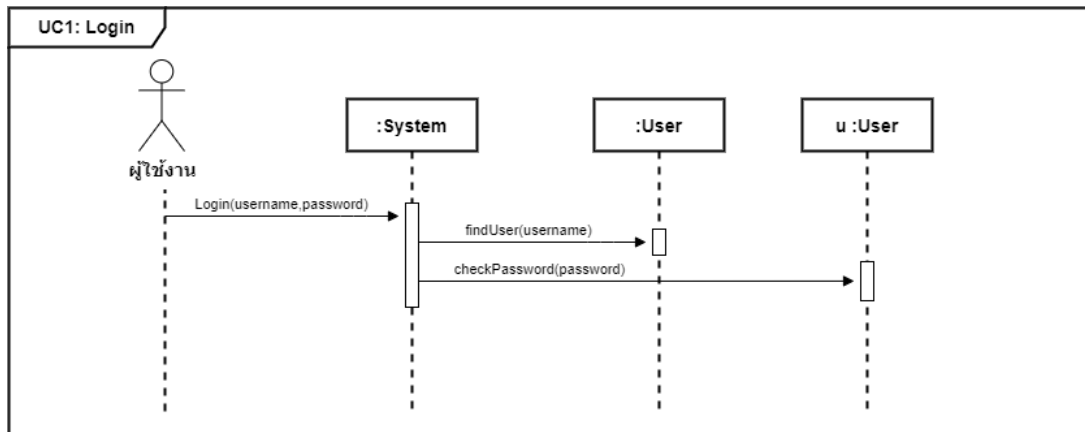
ตารางที่ 9: Use Case7 การลบตัวแบบ (Delete Model)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการลบตัวแบบออกจากระบบ
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ
Success Guarantee	ตัวแบบที่ผู้ใช้งานเลือก ถูกลบออกจากระบบ
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งาน แจ้งความต้องการลบตัวแบบ 2. ระบบแสดงข้อมูล ประวัติการสร้างตัวแบบ 3. ผู้ใช้งานเลือกตัวแบบที่ต้องการลบ 4. ระบบแสดงข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบที่ผู้ใช้งานเลือก 5. ผู้ใช้งานยืนยันลบตัวแบบออกจากระบบ 6. ระบบลบตัวแบบที่ผู้ใช้งานเลือก ออกจากระบบ
Extensions	<p>2a. กรณีไม่มีประวัติการสร้างตัวแบบ</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือนการไม่พบข้อมูลย้อนหลัง 2. ผู้ใช้งานรับทราบการแจ้งเตือน <p>5a. กรณีผู้ใช้งานไม่ยืนยันหรือยกเลิกการลบตัวแบบ</p> <ol style="list-style-type: none"> 1. ระบบแสดงข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบที่ผู้ใช้งานเลือก

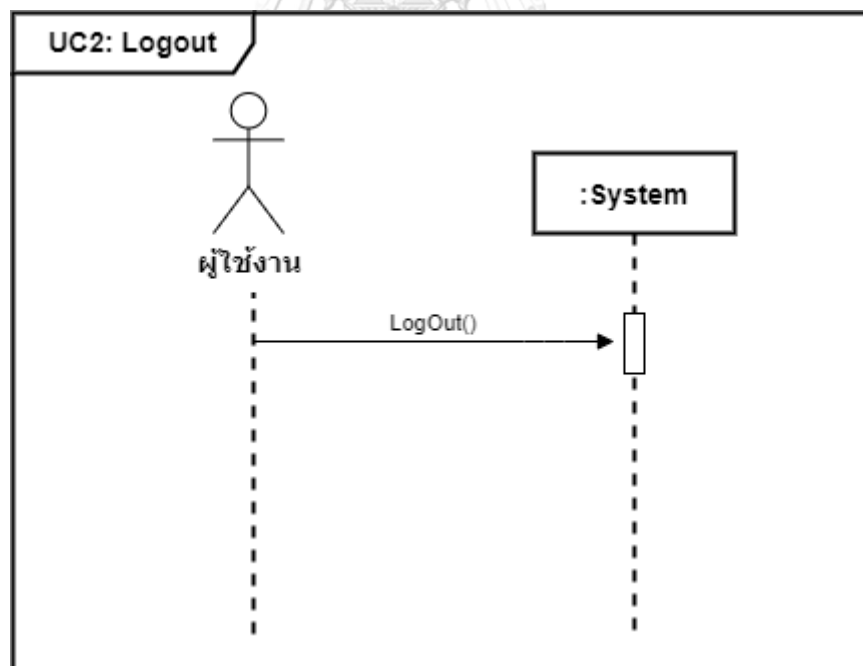
ตารางที่ 10: Use Case8 การเรียกดูคำแนะนำ (Generate Advice)

Scope	ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง
Level	User-goal
Primary Actor	ผู้ใช้งาน (User)
Stakeholder and their interest	ผู้ใช้งาน : ต้องการดูคำแนะนำการซื้อขายหลักทรัพย์ที่เกิดจากการวิเคราะห์ของตัวแบบ
Precondition	ผู้ใช้งานเข้าใช้งานอยู่ในระบบ
Success Guarantee	ผู้ใช้งานได้รับคำแนะนำการซื้อขายหลักทรัพย์
Main Success Scenario	<ol style="list-style-type: none"> 1. ผู้ใช้งาน แจ้งความต้องการเรียกดูคำแนะนำ 2. ระบบแสดงข้อมูล ประวัติการสร้างตัวแบบ 3. ผู้ใช้งานเลือกตัวแบบที่ต้องการใช้ในการสร้างคำแนะนำ 4. ระบบแสดงข้อมูลการตั้งค่า และ ผลการทำงาน ของตัวแบบที่ผู้ใช้งานเลือก 5. ผู้ใช้งานกรอกข้อมูลช่วงเวลา ที่ต้องการให้ตัวแบบให้คำแนะนำ 6. ผู้ใช้งานยืนยันการเรียกดูคำแนะนำ 7. ระบบแสดงคำแนะนำการซื้อขายหลักทรัพย์ที่ได้จากตัวแบบ
Extensions	<p>2a. กรณีไม่มีประวัติการสร้างตัวแบบ</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือนการไม่พบข้อมูลย้อนหลัง 2. ผู้ใช้งานรับทราบการแจ้งเตือน <p>5a. กรณีไม่พบข้อมูลประวัติราคาของหลักทรัพย์ในช่วงเวลาที่ผู้ใช้งานกรอก</p> <ol style="list-style-type: none"> 1. ระบบแจ้งเตือน ไม่พบข้อมูลประวัติราคา

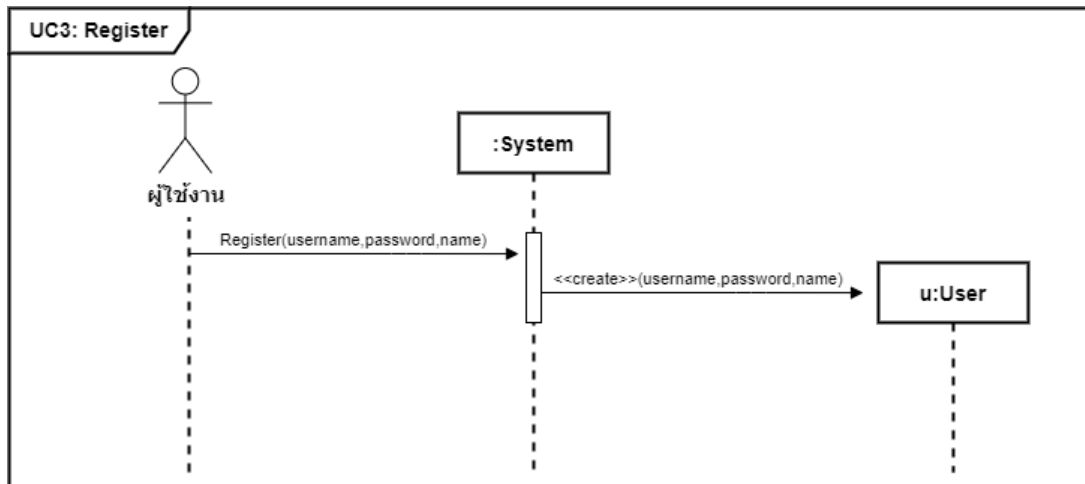
4. แผนภาพแสดงกระบวนการทำงานของระบบ (Sequence Diagram)



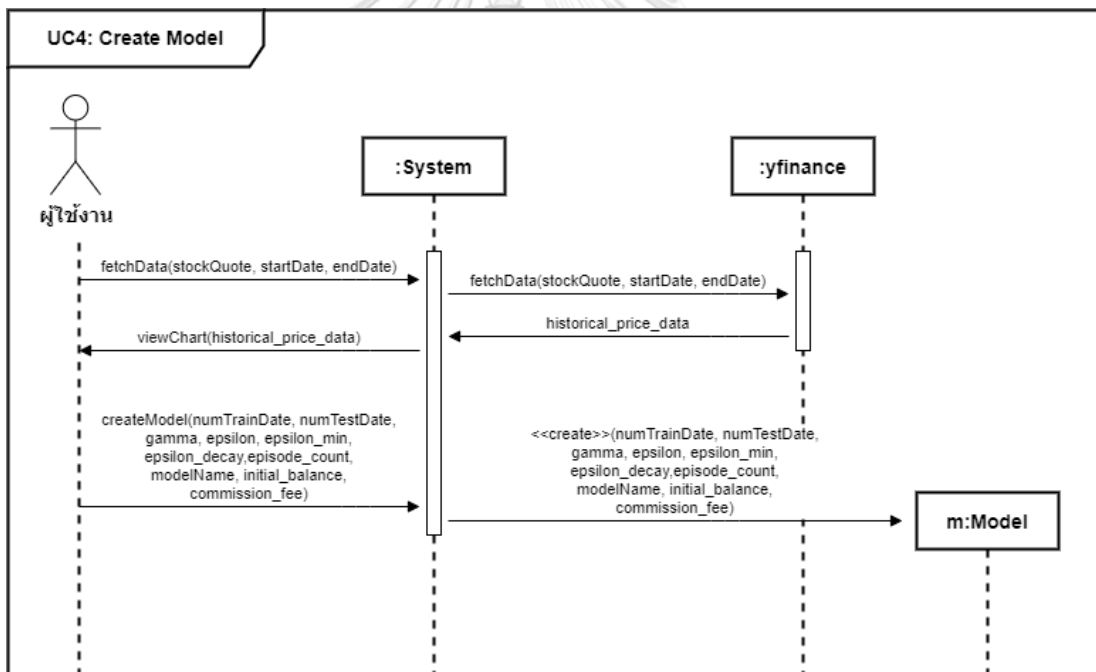
รูปที่ 5: Sequence Diagram 1 การเข้าสู่ระบบ (Login)



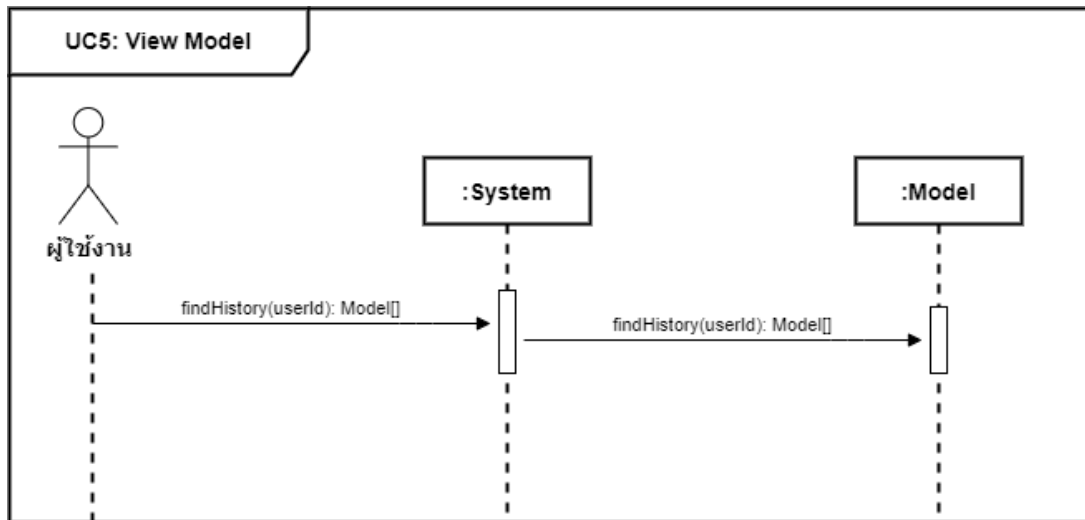
รูปที่ 6: Sequence Diagram 2 การออกจากระบบ (Logout)



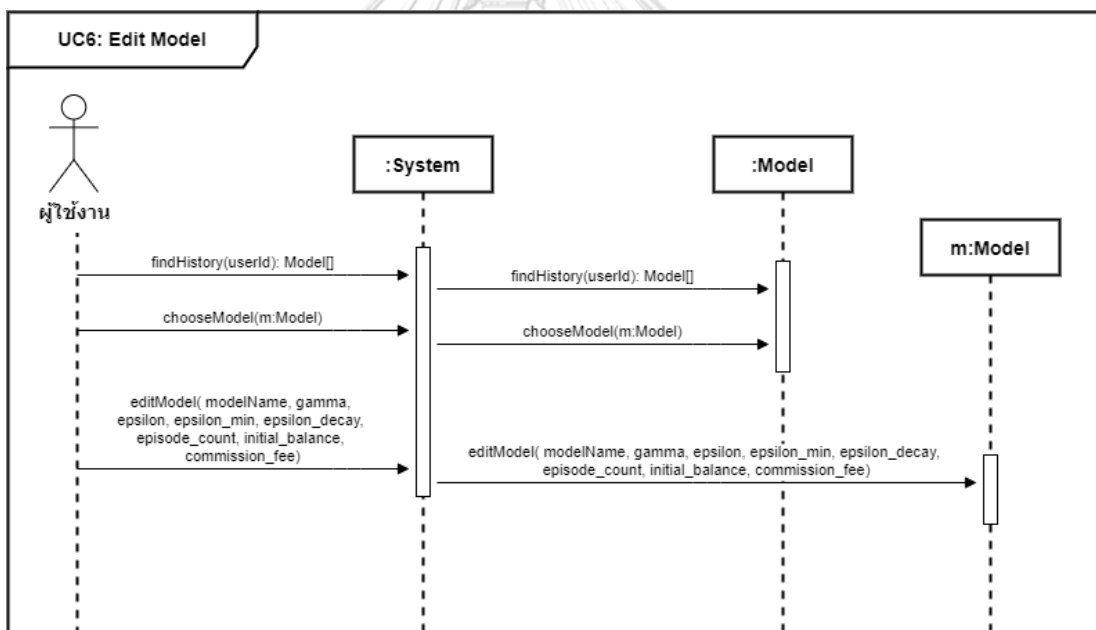
รูปที่ 7: Sequence Diagram 3 การลงทะเบียนผู้ใช้งาน (Register)



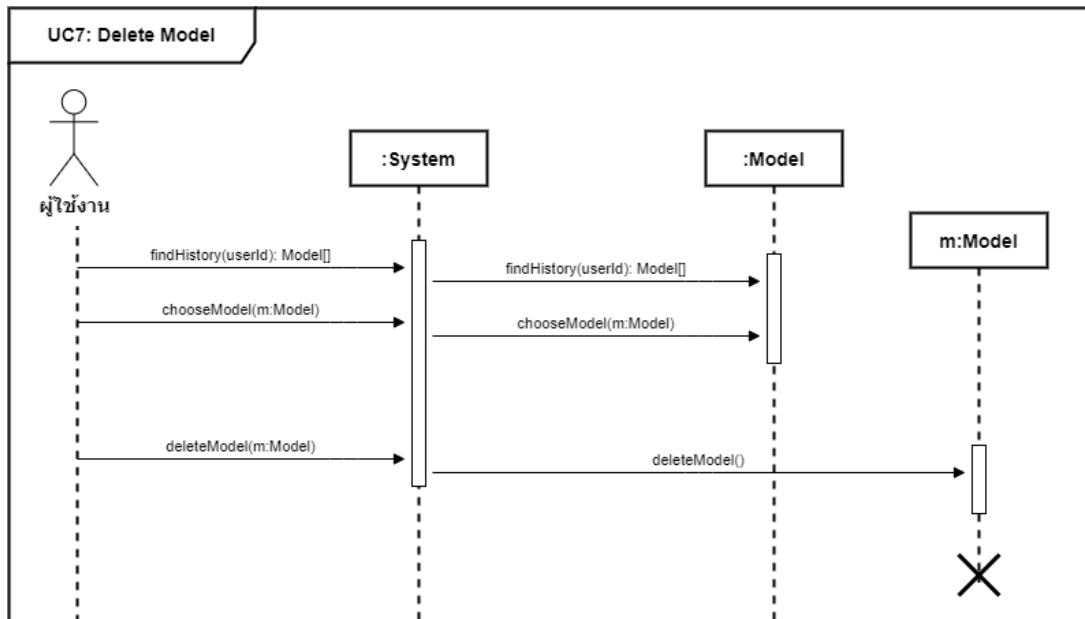
รูปที่ 8: Sequence Diagram 4 การสร้างตัวแบบ (Create Model)



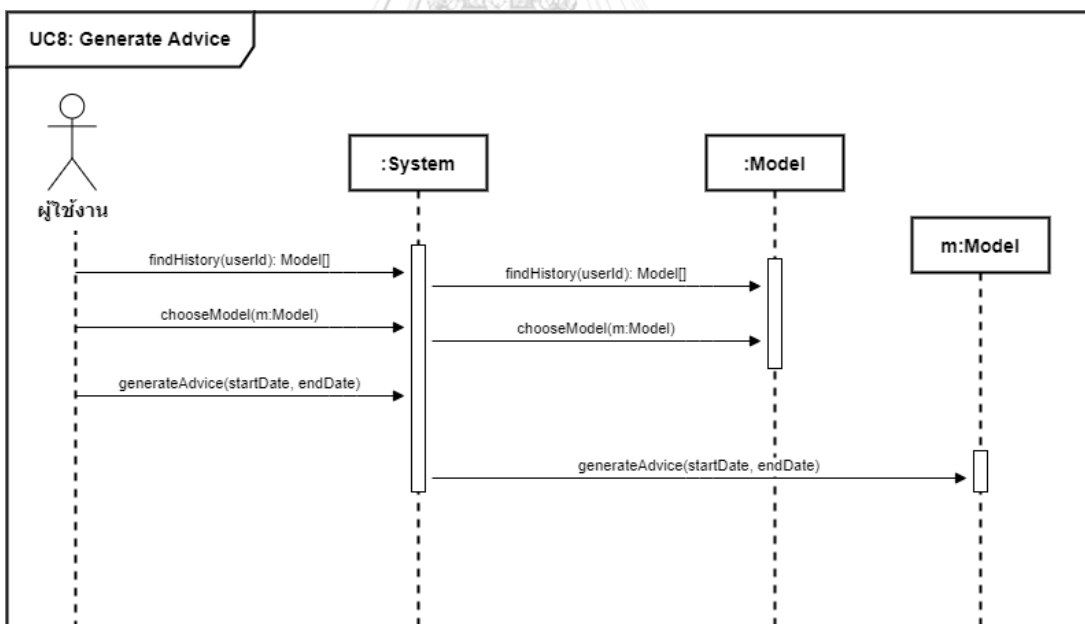
รูปที่ 9: Sequence Diagram 5 การแสดงประวัติการสร้างตัวแบบ (View Model History)



รูปที่ 10: Sequence Diagram 6 การแก้ไขตัวแบบ (Edit Model)



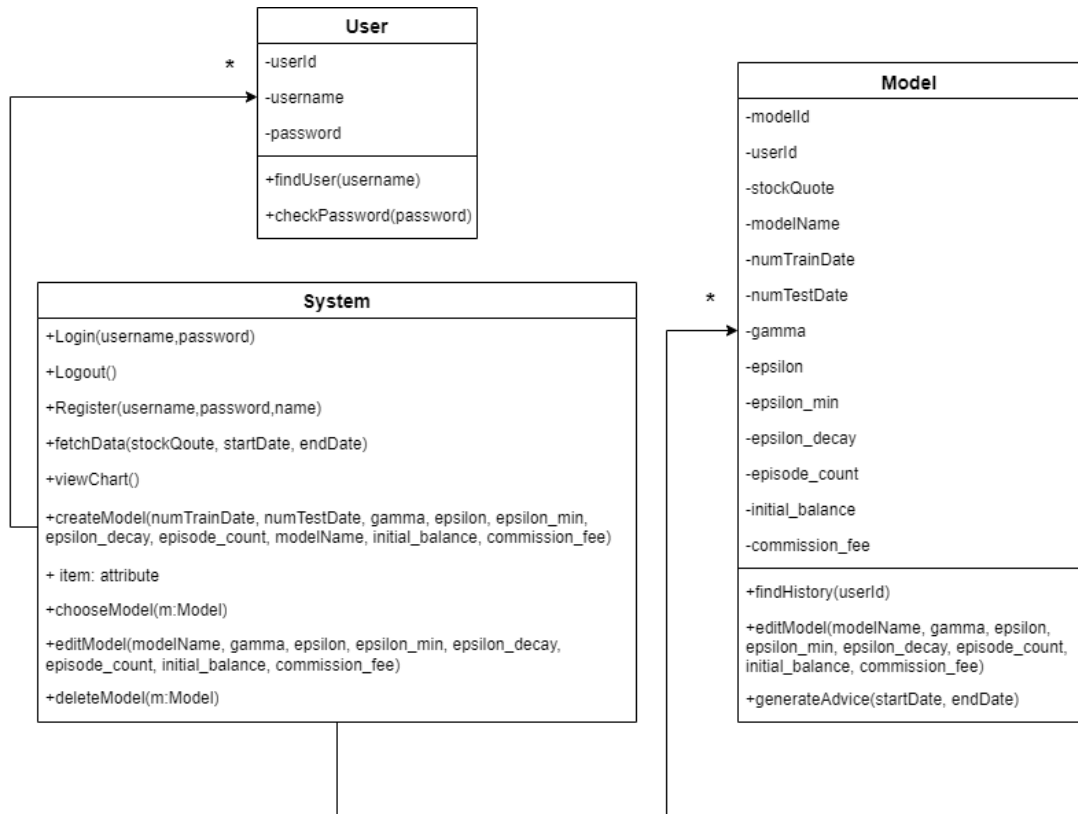
รูปที่ 11: Sequence Diagram 7 การลบตัวแบบ (Delete Model)



รูปที่ 12: Sequence Diagram 8 การเรียกดูคำแนะนำ (Generate Advice)

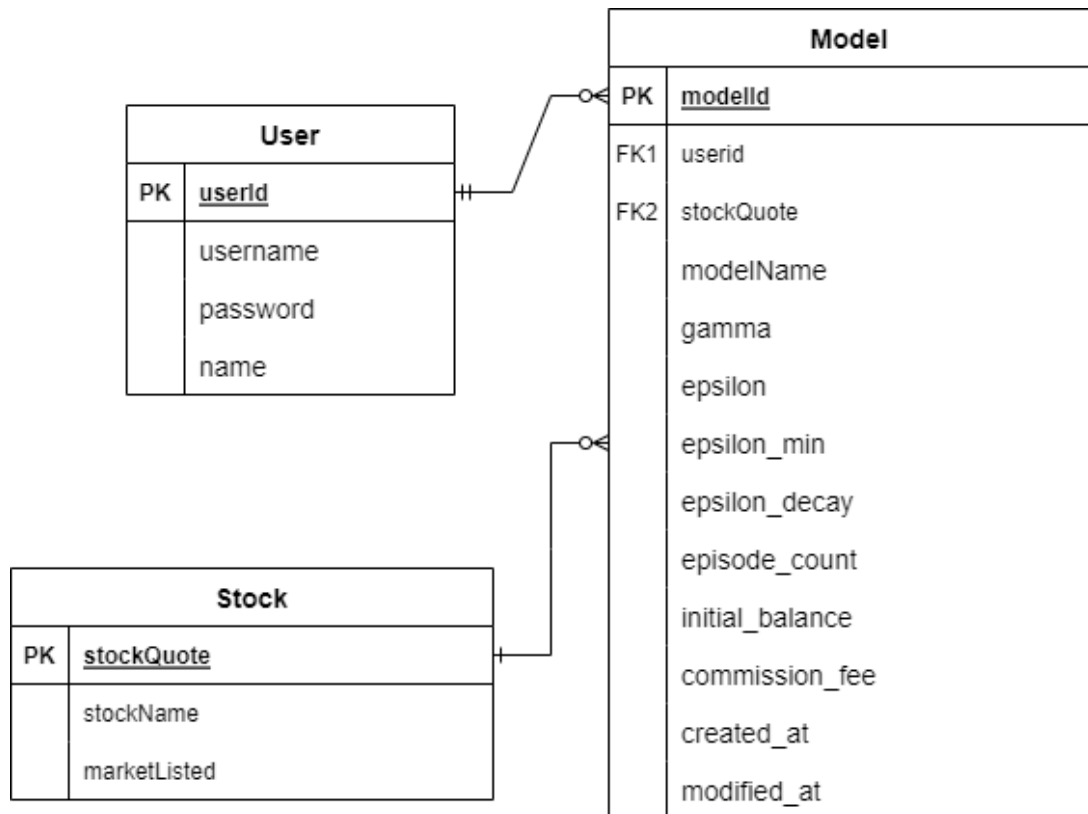
4.2 การออกแบบระบบ

1. โครงสร้างของระบบ Design Model



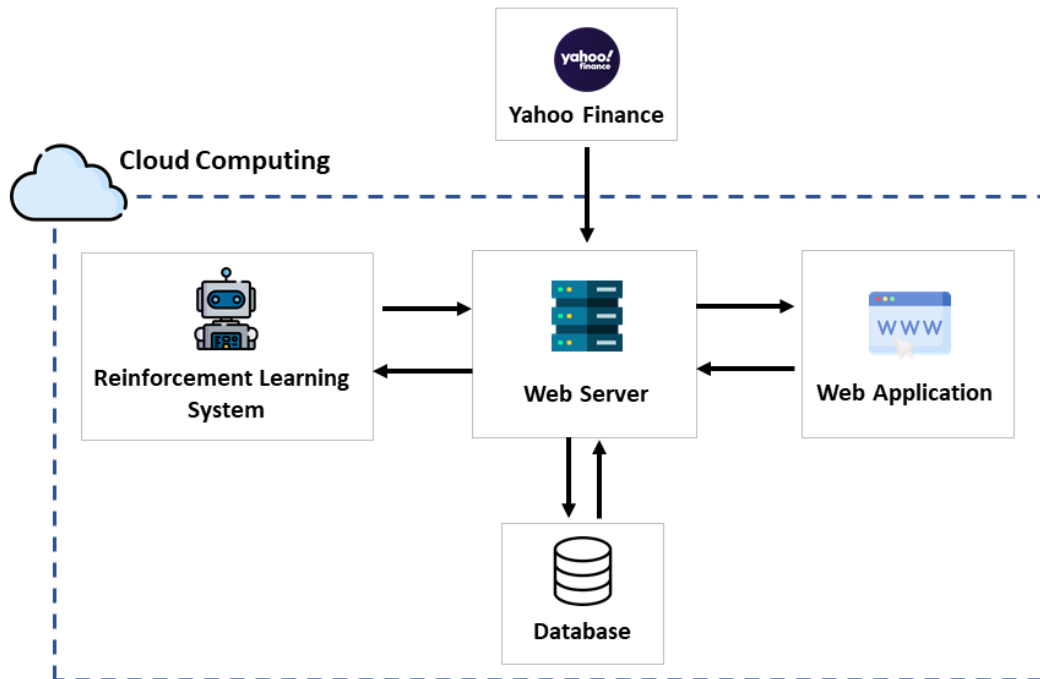
รูปที่ 13: Design Model ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย
ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง

2. Entities Relationship Diagram: ERD



รูปที่ 14: Entities Relationship Diagram

3. สถาปัตยกรรมของระบบ System Architecture



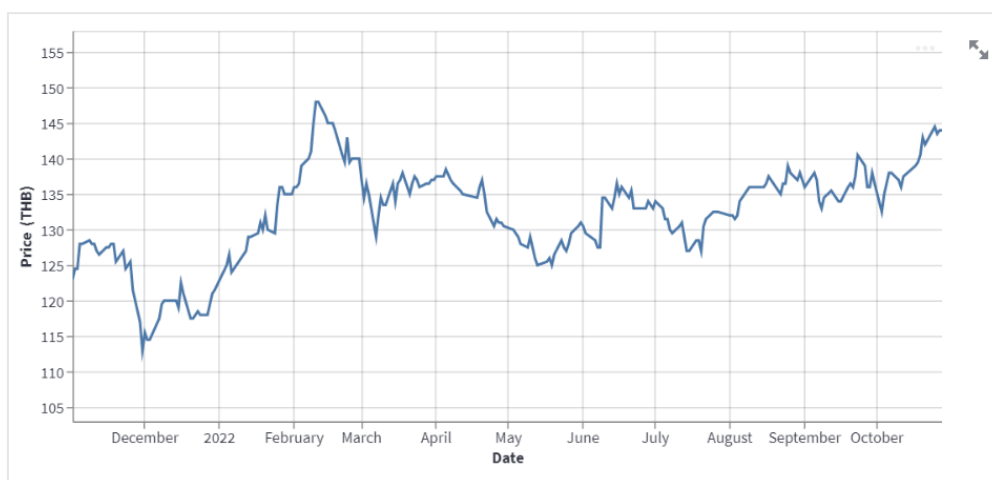
รูปที่ 15: สถาปัตยกรรมของระบบ System Architecture

4.3 การออกแบบผลลัพธ์ (Output Design)

การออกแบบผลลัพธ์ ที่ได้จากการพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง

1. การแสดงผลในรูปแบบกราฟ

เป็นรูปแบบการแสดงผลในเชิงรูปภาพที่ช่วยให้ผู้ใช้งานสามารถมองเห็นภาพรวม และทำความเข้าใจข้อมูลนั้น ๆ ได้ดียิ่งขึ้น



รูปที่ 16: ตัวอย่างการแสดงผลในรูปแบบกราฟ

2. การแสดงผลในรูปแบบข้อความ Logs

เป็นรูปแบบการแสดงผลที่ช่วยให้ผู้ใช้งานทราบถึงการดำเนินงานของระบบ

```

--- Episode 1 / 3 ---
---Episode 1 of 3 done...
---Total Reward: -13.50 | Account_Balance: 192366.49
--- Episode 2 / 3 ---
---Episode 2 of 3 done...
---Total Reward: -5.50 | Account_Balance: 194470.38
--- Episode 3 / 3 ---
---Episode 3 of 3 done...
---Total Reward: -26.50 | Account_Balance: 189347.26

```


รูปที่ 17: ตัวอย่างการแสดงผลในรูปแบบข้อความ Logs


4.4 การออกแบบข้อมูลนำเข้า (Input Design)

ข้อมูลที่ถูกนำเข้าสู่ระบบจะถูกจัดเก็บไว้ในฐานข้อมูลกลาง ซึ่งทำหน้าที่เป็นแหล่งเก็บข้อมูลหลักของระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง

การออกแบบข้อมูลนำเข้าจะเป็นไปตามรูปแบบและข้อมูลที่เกี่ยวข้องกับการจัดเก็บข้อมูลในกลุ่มต่างๆ โดยกลุ่มของข้อมูลดังกล่าวสามารถแบ่งออกได้เป็น 2 กลุ่ม ดังนี้


1. ข้อมูลผู้ใช้งาน (User) โดยมีการนำเข้าข้อมูลเมื่อผู้ใช้งานสมัครสมาชิก (Register)



Register 

Name
eg. Michael Burry

Username
eg. mikeburr1971

Password
your password 

Register

รูปที่ 18: Input Design ของการสมัครสมาชิก (Register)

2. ข้อมูลตัวแบบ (Model) โดยมีการนำเข้าสู่ข้อมูลเมื่อผู้ใช้งานดำเนินการสร้างตัวแบบ (Create Model)

Model parameters

Model name:
model_01

Gamma:
0.00 0.90 1.00

Starting epsilon (random walk probability):
0.00 1.00

Epsilon decline rate (random walk probability decline):
0.001 0.001 0.01

Minimum epsilon:
0.01 0.01 0.10

Learning rate:
0.001 0.001 0.01

Trading parameters

Initial account balance (THB):
1000000 - +

Trading size as a percentage of initial account balance (%):
0 10 100

Commission fee (%):
0.157 - +

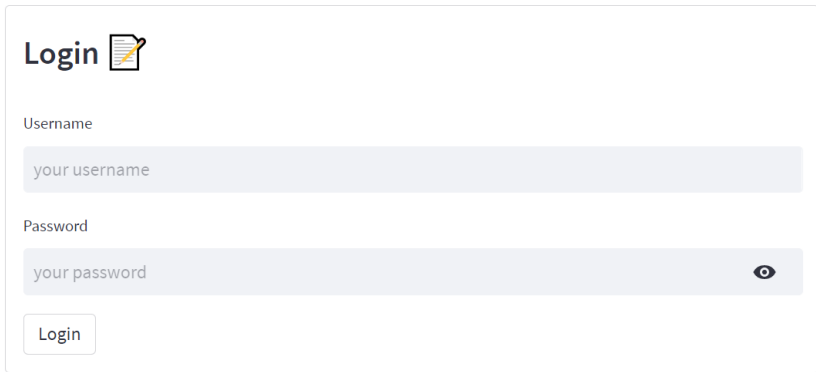
รูปที่ 19: Input Design ของการสร้างตัวแบบ (Create Model)

4.5 การออกแบบส่วนติดต่อผู้ใช้งาน (User Interface Design)

การออกแบบส่วนติดต่อผู้ใช้งาน คือการออกแบบที่เกี่ยวข้องกับองค์ประกอบต่าง ๆ ที่ผู้ใช้งานสามารถใช้เพื่อสื่อสารและสั่งการระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง โดยมีส่วนติดต่อผู้ใช้งานสำหรับการดำเนินกิจกรรมต่าง ๆ ของระบบดังนี้

1. การเข้าสู่ระบบ (Login)


มีหน้าที่รับข้อมูล ชื่อผู้ใช้งาน (Username) และ รหัสผ่าน (Password) จากผู้ใช้งาน เพื่อตรวจสอบสิทธิ์ในการเข้าใช้งานระบบ และเข้าถึงข้อมูลต่าง ๆ ของผู้ใช้งานฐานข้อมูล



รูปที่ 20: ตัวอย่างหน้าจอเข้าสู่ระบบ (Login)

2. การลงทะเบียนผู้ใช้งาน (Register)

มีหน้าที่รับข้อมูล ชื่อผู้ใช้งาน (Username) และ รหัสผ่าน (Password) จากผู้ใช้งาน เพื่อเพิ่มสิทธิ์ในการเข้าใช้งานระบบให้แก่ผู้ใช้งาน



รูปที่ 21: ตัวอย่างหน้าจอลงทะเบียนผู้ใช้งาน (Register)

3. การสร้างตัวแบบ (Create Model)

มีหน้าที่รับข้อมูล ชื่อหลักทรัพย์ และขอบเขตของข้อมูลราคาย้อนหลังของหลักทรัพย์นั้น ๆ ที่ผู้ใช้งานต้องการใช้ในการฝึกและทดสอบตัวแบบ รวมทั้งยังครอบคลุมถึงการกำหนดและตั้งค่าตัวแปรต่าง ๆ ที่เกี่ยวข้องกับการเรียนรู้ของตัวแบบ

Select stock and time period

Select your Stock

BBL


BANGKOK BANK PUBLIC COMPANY LIMITED --- Market: SET --- Industry: Financials --- Sector: Banking

Select start date:

2021/10/30

Select end date:

2022/10/30

View Dataset 

รูปที่ 22: ตัวอย่างหน้าจอเลือกหลักทรัพย์และขอบเขตของข้อมูลราคาย้อนหลัง



รูปที่ 23: ตัวอย่างหน้าจอแบ่งข้อมูลสำหรับการฝึกและทดสอบตัวแบบ

Set parameters for your trading model

Model parameters

Model name:

Gamma:

0.90

0.00

1.00

Starting epsilon (random walk probability):

1.00

0.00

1.00

Epsilon decline rate (random walk probability decline):

0.001

0.001

0.01

Minimum epsilon:

0.01

0.01

0.10

Learning rate:

0.001

0.001

0.01

Trading parameters

Initial account balance (THB):

 - +

Trading size as a percentage of initial account balance (%):

10

0

100

Commission fee (%):

 - +

รูปที่ 24: ตัวอย่างหน้าจอตั้งค่าตัวแปรต่าง ๆ ของตัวแบบ

4. การแสดงประวัติการสร้างตัวแบบ (View Model History)

มีหน้าที่แสดงผลข้อมูลของตัวแบบทั้งหมด ที่ผู้ใช้งานเคยสร้างไว้แล้ว

model_name	gamma	learning_rate	initial_balance	trading_size
<input checked="" type="checkbox"/> bbl_01	0.9	0.001	1000000	0.1
<input type="checkbox"/> bbl_02	0.8	0.002	1200000	0.25
<input type="checkbox"/> ppt_05	0.85	0.005	1980000	0.15
<input type="checkbox"/> scg_111	0.75	0.04	2550000	0.3
<input type="checkbox"/> mint_01	0.95	0.099	3390000	0.5

More model information: ▼

รูปที่ 25: ตัวอย่างหน้าจอแสดงประวัติการสร้างตัวแบบ



5. การแก้ไขตัวแบบ (Edit Model)

มีหน้าที่รับข้อมูลการตั้งค่าตัวแปรต่าง ๆ ที่เกี่ยวข้องกับการเรียนรู้ของตัวแบบ เพื่อเปลี่ยนแปลงการตั้งค่าของตัวแบบเดิมที่ผู้ใช้งานได้สร้างไว้แล้ว

Model parameters

Model name:

Gamma:

0.00 0.60 1.00

Starting epsilon (random walk probability):

0.00 0.90 1.00

Epsilon decline rate (random walk probability decline):

0.001 0.002 0.01

Minimum epsilon:

0.01 0.02 0.10

Learning rate:

0.001 0.002 0.01

Trading parameters

Initial account balance (THB):

 - +

Trading size as a percentage of initial account balance (%):

0 25 100

Commission fee (%):

 - +

รูปที่ 26: ตัวอย่างหน้าจอแก้ไขตัวแบบ

6. การเรียกดูคำแนะนำ (Generate Advice)

มีหน้าที่แสดงผลคำแนะนำที่ถูกประมวลโดยตัวแบบที่ผู้ใช้งานเลือก



รูปที่ 27: ตัวอย่างหน้าจอเรียกดูคำแนะนำ



7. การลบตัวแบบ (Delete Model)

มีหน้าที่แสดงรายการตัวแบบทั้งหมดที่ผู้ใช้งานเคยสร้างไว้ เพื่อให้ผู้ใช้งานได้เลือกตัวแบบที่ต้องการลบออกจากระบบ

model_name	gamma	learning_rate	initial_balance	trading_size
<input type="checkbox"/> bbl_01	0.9	0.001	1000000	0.1
<input type="checkbox"/> bbl_02	0.8	0.002	1200000	0.25
<input type="checkbox"/> ppt_05	0.85	0.005	1980000	0.15
<input checked="" type="checkbox"/> scg_111	0.75	0.04	2550000	0.3
<input type="checkbox"/> mint_01	0.95	0.099	3390000	0.5

Edit Delete

Are you sure?

Please confirm your choice:

No

Yes

Confirm

Model scg_111 has been successfully deleted

รูปที่ 28: ตัวอย่างหน้าจอลบตัวแบบ



4.6 การออกแบบระบบรักษาความปลอดภัย

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีระบบควบคุมความปลอดภัยโดย ผู้ใช้งานต้องยืนยันสิทธิ์ในการเข้าถึงและใช้งานระบบโดยผ่านกระบวนการ การเข้าสู่ระบบ (Login) ซึ่งผู้ใช้งานจะต้องกรอกข้อมูล ชื่อผู้ใช้งาน (Username) และ รหัสผ่าน (Password) โดยผู้ใช้งานแต่ละคนจะสามารถเข้าถึงได้เฉพาะข้อมูลของตนเองเท่านั้น และไม่สามารถเข้าถึงข้อมูลของผู้ใช้งานคนอื่นได้ แม้ว่าข้อมูลต่างๆจะถูกเก็บไว้บนฐานข้อมูลกลางเดียวกัน

4.7 การติดตั้งและพัฒนาระบบ

การพัฒนาระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีขั้นตอนในการพัฒนาระบบในส่วนต่าง ๆ ดังนี้

1. ระบบฐานข้อมูล (Database)

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง ใช้ระบบฐานข้อมูลแบบ NoSQL ซึ่งให้บริการแบบออนไลน์โดย Deta.sh ซึ่งมีขั้นตอนในการสร้างฐานข้อมูลดังนี้

1.1 สมัครใช้งาน Deta.sh

The image shows a 'Sign Up' form on the Deta.sh website. It features three input fields: 'Username' with a person icon, 'Password' with a lock icon and a toggle for visibility, and 'Email' with an envelope icon. Below the fields is a text line: 'By signing up, you agree to our [terms](#) and [privacy policy](#).' followed by a blue button labeled 'Sign Up ->'. At the bottom, there is a link: 'Have an account? [Sign In](#)'.

รูปที่ 29: หน้าจอสมัครสมาชิก Deta.sh

1.2 สร้างฐานข้อมูลผ่านชุดเครื่องมือ (Software Development Kit)

Put

`put` is the fastest way to store an item in the database.

If an item already exists under a given key, `put` will replace this item.

In the case you do not provide us with a key, we will auto generate a 12 char long string as a key.

JavaScript **Python** Go

```
put(
  data: typing.Union[dict, list, str, int, float, bool],
  key: str = None,
  *,
  expire_in: int = None,
  expire_at: typing.Union[int, float, datetime.datetime] = None
)
```

Parameters

- **data** (required) - Accepts: `dict`, `str`, `int`, `float`, `bool` and `list`.
 - Description: The data to be stored.
- **key** (optional) - Accepts: `str` and `None`
 - Description: the key (aka ID) to store the data under. Will be auto generated if not provided.
- **expire_in** (optional) - Accepts: `int` and `None`
 - Description: seconds after which the item will expire in, see also [expiring items](#)
- **expire_at** (optional) - Accepts: `int`, `float`, `datetime.datetime` and `None`
 - Description: time at which the item will expire in, can provide the timestamp directly(`int` or `float`) or a `datetime.datetime` object, see also [expiring items](#)

รูปที่ 30: ชุดเครื่องมือของ Deta.sh

2. ระบบ Back-end

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีระบบ Back-end ที่ทำหน้าที่ดึงข้อมูลราคาหลักทรัพย์ย้อนหลัง สร้างตัวแบบขึ้นมาเพื่อเรียนรู้ข้อมูลราคาดังกล่าว ไปจนถึงสื่อสารกับระบบฐานข้อมูลเพื่อเรียกดู บันทึก แก้ไข และลบข้อมูล ทั้งนี้เพื่อประมวลผลข้อมูลต่าง ๆ ให้เสร็จสิ้นและพร้อมสำหรับแสดงผลต่อไป โดยระบบ Back-end เหล่านี้ จะถูกพัฒนาบน Google Colab ซึ่งเป็นแพลตฟอร์มออนไลน์สำหรับการพัฒนาโปรแกรมคอมพิวเตอร์

ยินดีต้อนรับสู่ Colaboratory

ไฟล์ แก้ไข มุมมอง แทรก รันใหม่ เครื่องมือ ความช่วยเหลือ

ยินดีต้อนรับสู่ Colab

หากคุณเคยเคยกับ Colab อยู่แล้ว ลองดูวิดีโอนี้เพื่อให้อ่านเกี่ยวกับตารางแบบอินเทอร์แอคทีฟ มุมมองประวัติโค้ดที่เรียกใช้ และหาเคล็ดลับ

3 Cool Google Colab Features

Colab คืออะไร

Colab หรือ "Colaboratory" ช่วยให้คุณสามารถเขียนและเรียกใช้ Python ในเบราว์เซอร์ได้ด้วย

- ไม่ต้องกำหนดค่าใด
- เข้าถึง GPU โดยไม่มีค่าใช้จ่าย
- แชร์ได้ง่ายๆ

ไม่ว่าคุณจะเป็นนักเรียน/นักศึกษา นักวิทยาศาสตร์ข้อมูล หรือนักวิจัย AI ระบบ Colab ก็ช่วยให้งานของคุณง่ายขึ้นได้ รับชม [ข้อมูลเบื้องต้นเกี่ยวกับ Colab](#) เพื่อดูข้อมูลเพิ่มเติม หรือจะเริ่มต้นใช้งานที่ด้านล่างเลยก็ได้

▼ เริ่มต้นใช้งาน

เอกสารที่ควูล่าถึงอ่านไม่ใช่น่าอ่านแบบคงที่ แต่เป็นสภาพแวดล้อมแบบอินเทอร์แอคทีฟที่เรียกว่า **สมุดบันทึก Colab** ซึ่งให้คุณเขียนและเรียกใช้โค้ดได้ ลองดูตัวอย่างนี้ นี่คือการ **เซลล์โค้ด** ที่มีสคริปต์ Python สั้นๆ ซึ่งจะคำนวณค่าค่าหนึ่ง เก็บไว้ในตัวแปร แล้วพิมพ์ผลลัพธ์ออกมา

```
[ ] seconds_in_a_day = 24 * 60 * 60
seconds_in_a_day

86400
```

รูปที่ 31: หน้าจอ Google Colab

▼ Import library

```
[ ] pip install yfinance
```

```
[ ] import numpy as np
import pandas as pd
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.models import load_model

import yfinance as yf
```

▼ Create: ReplayBuffer object

```
[ ] class ReplayBuffer():
    def __init__(self, max_size, input_dims):
        self.mem_size = max_size
        self.mem_cntr = 0

        self.state_memory = np.zeros((self.mem_size, input_dims), dtype=np.float32)
        self.new_state_memory = np.zeros((self.mem_size, input_dims), dtype=np.float32)
        self.action_memory = np.zeros(self.mem_size, dtype=np.int32)
        self.reward_memory = np.zeros(self.mem_size, dtype=np.float32)
        self.terminal_memory = np.zeros(self.mem_size, dtype=np.int32)

    def store_transition(self, state, action, reward, state_, done): # to store experience into memory
        index = self.mem_cntr % self.mem_size
        self.state_memory[index] = state
        self.new_state_memory[index] = state_
        self.reward_memory[index] = reward
        self.action_memory[index] = action
```

รูปที่ 32: ตัวอย่างการพัฒนาในระบบใน Google Colab

3. ระบบ Front-end

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีระบบ Front-end ที่ทำหน้าที่แสดงผลที่ประมวลเสร็จสิ้นแล้วจากระบบ Back-end ผ่านหน้าจอให้กับผู้ใช้งาน ซึ่งพัฒนาโดยใช้ Streamlit Library

```

This is a summary of the docs, as of Streamlit v1.14.0.

Install & Import
streamlit run first_app.py

# Import convention
>>> import streamlit as st

Command line
streamlit --help
streamlit run your_script.py
streamlit hello
streamlit config show
streamlit cache clear
streamlit docs
streamlit --version

Pre-release features
pip uninstall streamlit
pip install streamlit-nightly --u

Learn more about experimental features

Magic commands
# Magic commands implicitly
# call st.write().
'_This_ is some **Markdown***'
my_variable
'dataframe:', my_data_frame

Display text
st.text('Fixed width text')
st.markdown('_Markdown_') # see *
st.latex(r'' e^{i\pi} + 1 = 0 ''
st.write('Most objects') # df, en
st.write(['st', 'is <', 3]) # see
st.title('My title')
st.header('My header')
st.subheader('My sub')
st.code('for i in range(8): foo()
* optional kwarg unsafe_allow_html:

Display data
st.dataframe(my_dataframe)
st.table(data.iloc[0:10])

Display interactive widgets
st.button('Click me')
st.checkbox('I agree')
st.radio('Pick one', ['cats', 'dog
st.selectbox('Pick one', ['cats',
st.multiselect('Buy', ['milk', 'a
st.slider('Pick a number', 0, 100
st.select_slider('Pick a size', [
st.text_input('First name')
st.number_input('Pick a number', 1
st.text_area('Text to translate')
st.date_input('Your birthday')
st.time_input('Meeting time')
st.file_uploader('Upload a CSV')
st.download_button('Download file
st.camera_input("Take a picture")
st.color_picker('Pick a color')

# Use widgets' returned values in
>>> for i in range(int(st.number_
>>> foo()
>>> if st.sidebar.selectbox('I:',|

Optimize performance
Legacy caching
>>> @st.cache
... def foo(bar):
... # Do something expensive in
... return data
>>> # Executes foo
>>> d1 = foo(ref1)
>>> # Does not execute foo
>>> # Returns cached item by refer
>>> d2 = foo(ref1)
>>> # Different arg, so function ;
>>> d3 = foo(ref2)

Cache data objects
# E.g. Dataframe computation, sto
>>> @st.experimental_memo
... def foo(bar):
... # Do something expensive an
... return data
# Executes foo
>>> d1 = foo(ref1)

```

รูปที่ 33: Streamlit Library

บทที่ 5

บทสรุป ปัญหา และข้อเสนอแนะ

บทนี้จะกล่าวถึงบทสรุป ปัญหา และข้อเสนอแนะของการพัฒนาระบบสนับสนุนการตัดสินใจ สำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง เพื่อใช้เป็นแนวทางในการพัฒนาระบบสารสนเทศอื่น ๆ ต่อไป

5.1 บทสรุป

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง เริ่มต้นจากการวิเคราะห์วิธีการดำเนินงาน ความต้องการ และปัญหาที่เกิดขึ้นของบริษัทต้นแบบ เพื่อนำมาใช้พัฒนาระบบที่สามารถตอบสนองความต้องการของผู้ใช้งานได้ โดยเฉพาะอย่างยิ่งการช่วยสนับสนุนผู้ใช้งานในการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ซึ่งระบบนี้สามารถตอบสนองวัตถุประสงค์ของโครงการ ดังนี้

1. เพื่อสร้างระบบในการให้คำแนะนำ เพื่อสนับสนุนการตัดสินใจลงทุนในหลักทรัพย์
2. เพื่อประยุกต์ใช้เทคนิคการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ในการพัฒนาระบบสนับสนุนการตัดสินใจสำหรับซื้อขายหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง ประกอบด้วย 8 การทำงานย่อย คือ

1. การเข้าสู่ระบบ (Login)
2. การออกจากระบบ (Logout)
3. การลงทะเบียนผู้ใช้งาน (Register)
4. การสร้างตัวแบบ (Create Model)
5. การแสดงประวัติการสร้างตัวแบบ (View Model History)
6. การแก้ไขตัวแบบ (Edit Model)
7. การลบตัวแบบ (Delete Model)
8. การเรียกดูคำแนะนำ (Generate Advice)

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง ถูกพัฒนาโดยใช้ ภาษา Python เป็นหลัก โดยมี Tensorflow และ yfinance

เป็น Library ที่สำคัญสำหรับระบบ Back-end ที่เกี่ยวข้องกับการดึงข้อมูลราคาหลักทรัพย์ย้อนหลัง การสร้างตัวแบบ การสร้างกระบวนการเรียนรู้แบบเสริมกำลัง และการประมวลผลคำแนะนำเพื่อสนับสนุนการตัดสินใจซื้อขายหลักทรัพย์ ในขณะที่ระบบ Front-end นั้น จะถูกพัฒนาโดยใช้ Streamlit

5.2 ปัญหา

ปัญหาที่ผู้จัดทำ ค้นพบระหว่างการพัฒนาาระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีดังนี้

1. Tensorflow ยังไม่มี ไลบรารีที่ถูกพัฒนามาเพื่อการเรียนรู้แบบเสริมกำลังโดยตรง

ปัญหา: แม้ว่า Tensorflow จะเป็นไลบรารีที่พร้อมรองรับการนำไปประยุกต์ใช้ใน การเรียนรู้ของเครื่อง (Machine Learning) หลายหลายรูปแบบ แต่การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ยังเป็นการเรียนรู้ที่ยังไม่มีฟังก์ชันสำเร็จรูปที่สามารถนำไปประยุกต์ใช้ได้ทันที

แนวทางแก้ไข: ศึกษาทฤษฎีของการเรียนรู้แบบเสริมกำลัง แล้วนำไลบรารี Tensorflow มาปรับเปลี่ยนและต่อยอดให้สามารถประยุกต์ใช้กับการเรียนรู้แบบเสริมกำลังได้

2. กระบวนการเรียนรู้แบบเสริมกำลังบนข้อมูลขนาดใหญ่ใช้ทรัพยากรคอมพิวเตอร์สูง

ปัญหา: เมื่อข้อมูลราคาย้อนหลังของหลักทรัพย์มีขนาดใหญ่ขึ้น เวลาที่ต้องใช้ในกระบวนการเรียนรู้แบบเสริมกำลังก็จะยิ่งมากขึ้นตามไปด้วย เนื่องจากทรัพยากรของระบบคอมพิวเตอร์นั้นมีจำกัด โดยในบางกรณีนั้นระบบต้องใช้เวลาดำเนินการหลายชั่วโมง เพื่อให้ได้มาซึ่งตัวแบบที่มีประสิทธิภาพเพียงพอ

แนวทางแก้ไข: ปรับปรุงกระบวนการทำงานของการเรียนรู้แบบเสริมกำลัง โดยส่งกระบวนการที่ใช้ทรัพยากรสูง ไปทำงานบน Google Cloud Platform แทน ซึ่งมีบริการให้ยืมใช้ทรัพยากรทางคอมพิวเตอร์กำลังสูง

5.3 ข้อจำกัดของระบบ

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง จำลองการซื้อขายหลักทรัพย์แบบซื้อสลับกับขาย กล่าวคือเมื่อตัวแบบตัดสินใจซื้อหลักทรัพย์แล้ว ตัวแบบจะไม่สามารถซื้อหลักทรัพย์เพิ่มได้จนกว่าจะขายหลักทรัพย์เก่าออกไปเสียก่อน และการซื้อหลักทรัพย์ในแต่ละครั้งจะมีมูลค่าเท่ากันตลอดการจำลองตามที่ใช้ระบุ

การจำลองดังกล่าวอาจไม่สะท้อนการซื้อขายหลักทรัพย์ในโลกแห่งความเป็นจริงได้ทั้งหมด เนื่องจากในความเป็นจริงนั้น นักลงทุนไม่จำเป็นต้องซื้อสลับกับขายและการซื้อขายแต่ละครั้งอาจมีมูลค่าที่แตกต่างกันได้ตามกลยุทธ์ส่วนตัวของนักลงทุน

5.4 ข้อเสนอแนะ

ระหว่างการพัฒนาาระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลังนั้น ผู้จัดทำได้พบปัญหาต่างๆ ตามที่ได้กล่าวไว้ ผู้จัดทำจึงขอเสนอแนะแนวทางในการพัฒนาระบบหรือโครงการที่มีลักษณะเดียวกันกับระบบนี้ โดยแบ่งเป็น 2 กลุ่มข้อเสนอแนะ ดังนี้

1. ข้อเสนอแนะในกรณีที่มีการพัฒนาระบบเพิ่มเติม

1.1 เพิ่มตัวเลือกในการนำข้อมูลราคาย้อนหลังของหลักทรัพย์ที่สนใจ ในกรอบเวลาที่เล็กกว่ารายวัน เช่น รายชั่วโมง หรือ รายนาที ซึ่งจะเหมาะสมกับการซื้อขายหลักทรัพย์แบบวันต่อวัน (Day Trading) และ การซื้อขายด้วยความถี่สูง (High Frequency Trading) มากกว่า

1.2 เพิ่มข้อมูลตัวชี้วัดทางเทคนิค (Technical Indicator) ซึ่งเป็นเครื่องมือที่ถูกใช้โดยทั่วไปอยู่แล้วโดยนักลงทุนทั้งรายใหญ่และรายย่อย เช่น MACD และ RSI เป็นต้น

1.3 เพิ่มความสามารถในการปรับตัวแปรการตั้งค่าของตัวแบบการเรียนรู้ได้แบบอัตโนมัติ (Automated Hyperparameter Tuning) เพื่อให้ตัวแบบมีประสิทธิภาพผลสูงสุด

1.4 เพิ่มความสามารถในการเข้าถึงและใช้งานตัวแบบหนึ่งๆ ได้โดยผู้ใช้งานหลายคน ที่อาจทำงานอยู่ในองค์กรหรือฝ่ายเดียวกัน เพื่อให้ตอบโจทย์การใช้งานระบบในระดับองค์กร ที่เป็นการใช้ตัวแบบร่วมกันระหว่างหลายบุคคลหรือหลายฝ่ายภายในองค์กร

2. ข้อเสนอแนะในกรณีที่จะมีการพัฒนาระบบนี้ขึ้นอีกครั้ง

2.1 ศึกษาเทคนิคด้านการเรียนรู้แบบเสริมกำลังในแขนงอื่นเพิ่มเติม ที่นอกเหนือจาก Q-Learning เช่น Policy Optimization เป็นต้น

2.2 ศึกษาเครื่องมือสำหรับพัฒนาระบบ Front-end แบบ No-Code หรือ Low-Code ที่มีให้บริการอยู่แล้วในปัจจุบัน เพื่อเพิ่มขีดความสามารถของระบบ Front-end

2.3 ศึกษาเครื่องมือด้าน Machine Learning อื่นๆ ที่เปิดให้บริการอยู่แล้วโดยบริษัทด้านเทคโนโลยีต่างๆ เช่น Microsoft, Amazon และ Google เพื่อเพิ่มความเป็นไปได้ด้านขีดความสามารถของการเรียนรู้แบบเสริมกำลัง



บรรณานุกรม



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

AlgoAddict (2019). "[Reinforcement Learning 101] ตะลุยคาสีโนด้วย RL: Muti-Armed Bandit (1)- Epsilon Greedy." from <https://algoaddict.wordpress.com/2019/06/22/reinforcement-learning-101-beating-the-casino-multi-armed-bandit-problem-explore-exploit-delima-epsilon-greddy-and-sample-mean/>.

Bhatt, S. (2018). "Reinforcement Learning 101." from <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>.

Marukatat, S. (2017). "Reinforcement Learning." from <https://medium.com/hectec/reinforcement-learning-43ea03c2e00e>.

Noparat, N. (2020). "มาทำความรู้จักกับ การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)." from <https://bigdata.go.th/big-data-101/introduction-to-reinforcement-learning/>.

Richard S Sutton, A. G. B. (2018). Reinforcement Learning: An Introduction, 2nd edition, MIT press.



ภาคผนวก ก

พจนานุกรมข้อมูล

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีตารางสำหรับจัดเก็บข้อมูล ดังนี้

1. User Table

ตารางที่ 11: User Table

Name	Data Type	Primary Key	Description
username	String	Yes	ชื่อผู้ใช้งาน
name	String	No	ชื่อ-นามสกุล ผู้ใช้
password	String	No	รหัสผ่าน

2. Model Table

ตารางที่ 12: Model Table

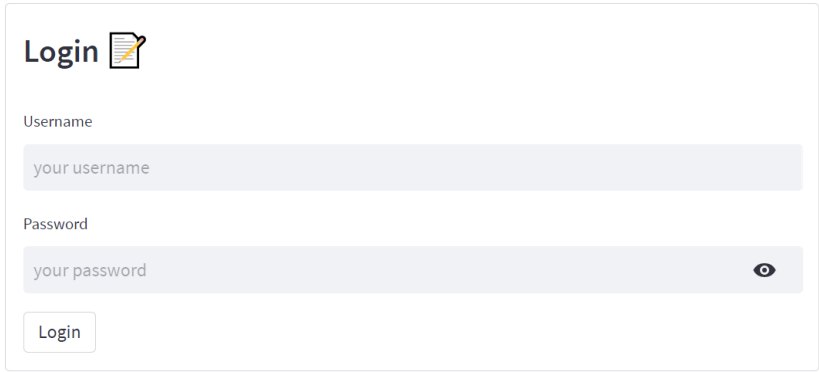
Name	Data Type	Primary Key	Description
key	String	Yes	รหัสตัวแบบ
username	String	No	ชื่อผู้ใช้งาน
model_name	String	No	ชื่อตัวแบบ
gamma	Float	No	ค่าแกมมา
epsilon	Float	No	ค่าเอปไซลอน
epsilon_dec	Float	No	อัตราการลดค่าเอปไซลอน
epsilon_min	Float	No	ค่าเอปไซลอนต่ำสุด
learning_rate	Float	No	ค่า learning rate
account_balance	Integer	No	มูลค่าบัญชีเริ่มต้น
trading_size_pct	Integer	No	ร้อยละของบัญชีในการซื้อขาย
commission_fee	Float	No	ค่าคอมมิชชั่นในการซื้อขาย
trained_status	Boolean	No	สถานะการฝึกของตัวแบบ
trained_episode	Integer	No	จำนวนรอบการฝึกตัวแบบ
saved_status	Boolean	No	สถานะการบันทึกตัวแบบ

ภาคผนวก ข

เมนูการทำงานของระบบ

1. เมนูเข้าสู่ระบบ (Login)

มีหน้าที่รับข้อมูล ชื่อผู้ใช้งาน (Username) และ รหัสผ่าน (Password) จากผู้ใช้งาน เพื่อตรวจสอบสิทธิ์ในการเข้าใช้งานระบบ และเข้าถึงข้อมูลต่างๆของผู้ใช้ในฐานข้อมูล



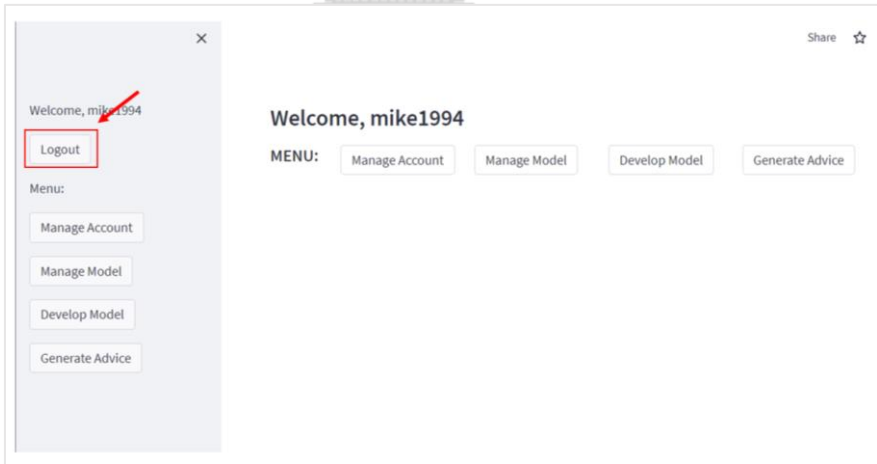
The screenshot shows a login form with the following elements:

- Title: Login
- Username field: contains the text "your username"
- Password field: contains the text "your password" and has a visibility toggle icon (an eye) on the right side.
- Login button: located below the password field.

รูปที่ 34: Login Menu

2. เมนูออกจากระบบ (Logout)

มีหน้าที่นำผู้ใช้งานออกจากระบบ



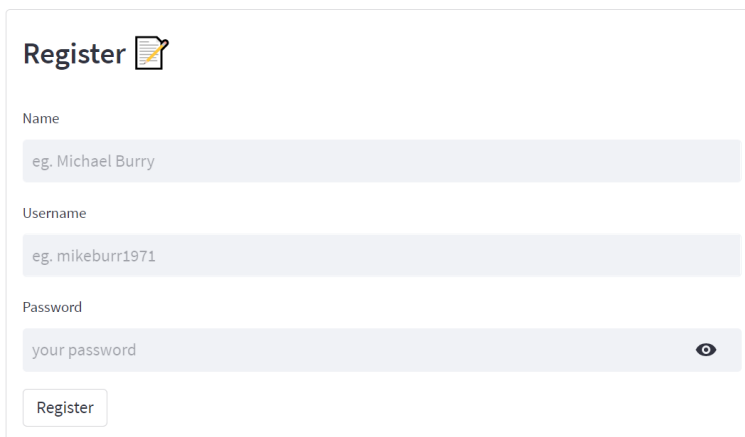
The screenshot shows a user interface with the following elements:

- Header: Welcome, mike1994
- Navigation Menu (left sidebar):
 - Logout (highlighted with a red box and a red arrow)
 - Manage Account
 - Manage Model
 - Develop Model
 - Generate Advice
- Main Content Area:
 - Header: Welcome, mike1994
 - MENU: Manage Account, Manage Model, Develop Model, Generate Advice

รูปที่ 35: Logout Menu

3. เมนูลงทะเบียนผู้ใช้งาน (Register)

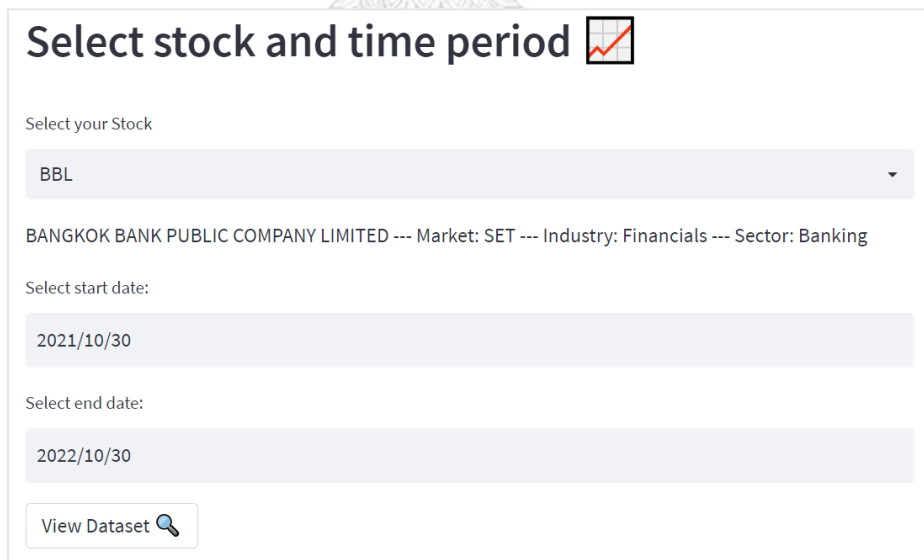
มีหน้าที่รับข้อมูล ชื่อผู้ใช้งาน (Username) และ รหัสผ่าน (Password) จากผู้ใช้งาน เพื่อเพิ่มสิทธิ์ในการเข้าใช้งานระบบให้แก่ผู้ใช้งาน



รูปที่ 36: Register Menu

4. เมนูสร้างตัวแบบ (Create Model)

มีหน้าที่รับข้อมูล ชื่อหลักทรัพย์ และขอบเขตของข้อมูลราคาย้อนหลังของหลักทรัพย์นั้นๆ ที่ผู้ใช้งานต้องการใช้ในการฝึกและทดสอบตัวแบบ รวมทั้งยังครอบคลุมถึงการกำหนดและตั้งค่าตัวแปรต่างๆที่เกี่ยวข้องกับการเรียนรู้ของตัวแบบ



รูปที่ 37: Select stock and time period



รูปที่ 38: Split Dataset

Set parameters for your trading model 💡

Model parameters

Model name:
BBL_271022

Gamma:
0.00 0.90 1.00

Starting epsilon (random walk probability):
0.00 1.00

Epsilon decline rate (random walk probability decline):
0.001 0.01

Minimum epsilon:
0.01 0.10

Learning rate:
0.001 0.01

Trading parameters

Initial account balance (THB):
1000000 - +

Trading size as a percentage of initial account balance (%):
0 10 100

Commission fee (%):
0.157 - +

Set Parameters

รูปที่ 39: Set Parameters

5. เมนูแสดงประวัติการสร้างตัวแบบ (View Model History)

มีหน้าที่แสดงผลข้อมูลของตัวแบบทั้งหมด ที่ผู้ใช้งานเคยสร้างไว้แล้ว

model_name	gamma	learning_rate	initial_balance	trading_size
<input checked="" type="checkbox"/> bbl_01	0.9	0.001	1000000	0.1
<input type="checkbox"/> bbl_02	0.8	0.002	1200000	0.25
<input type="checkbox"/> ppt_05	0.85	0.005	1980000	0.15
<input type="checkbox"/> scg_111	0.75	0.04	2550000	0.3
<input type="checkbox"/> mint_01	0.95	0.099	3390000	0.5

More model information: ▼

รูปที่ 40: View Model History Menu

6. เมนูแก้ไขตัวแบบ (Edit Model)

มีหน้าที่รับข้อมูลการตั้งค่าตัวแปรต่างๆที่เกี่ยวข้องกับการเรียนรู้ของตัวแบบ เพื่อเปลี่ยนแปลงการตั้งค่าของตัวแบบเดิมที่ผู้ใช้งานได้สร้างไว้แล้ว

Model parameters

Model name:
PTT_101022

Gamma:
0.00 0.50 1.00

Starting epsilon (random walk probability):
0.00 0.90 1.00

Epsilon decline rate (random walk probability decline):
0.001 0.002 0.01

Minimum epsilon:
0.01 0.02 0.10

Learning rate:
0.001 0.002 0.01

Trading parameters

Initial account balance (THB):
2500000 - +

Trading size as a percentage of initial account balance (%):
0 25 100

Commission fee (%):
0.157 - +

Set Parameters

รูปที่ 41: Edit Parameters

7. เมนูลบตัวแบบ (Delete Model)

มีหน้าที่แสดงรายการตัวแบบทั้งหมดที่ผู้ใช้งานเคยสร้างไว้ เพื่อให้ผู้ใช้งานได้เลือกตัวแบบที่ต้องการลบออกจากระบบ

model_name	gamma	learning_rate	initial_balance	trading_size
<input type="checkbox"/> bbl_01	0.9	0.001	1000000	0.1
<input type="checkbox"/> bbl_02	0.8	0.002	1200000	0.25
<input type="checkbox"/> ppt_05	0.85	0.005	1980000	0.15
<input checked="" type="checkbox"/> scg_111	0.75	0.04	2550000	0.3
<input type="checkbox"/> mint_01	0.95	0.099	3390000	0.5

Edit Delete

Are you sure?

Please confirm your choice:

No

Yes

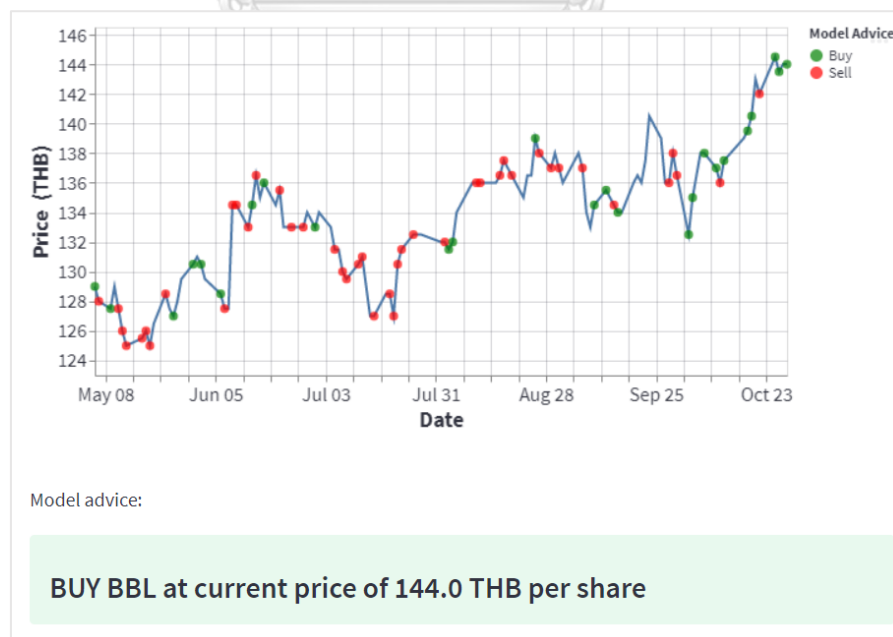
Confirm

Model scg_111 has been successfully deleted

รูปที่ 42: Delete Model Menu

8. เมนูเรียกดูคำแนะนำ (Generate Advice)

มีหน้าที่แสดงผลคำแนะนำที่ถูกประมวลโดยตัวแบบที่ผู้ใช้งานเลือก



รูปที่ 43: Generate Advice Menu

ภาคผนวก ค

ตัวอย่างรายงาน

ระบบสนับสนุนการตัดสินใจสำหรับการซื้อขายตราสารทุนในตลาดหุ้นไทย ด้วยเทคนิคการเรียนรู้แบบเสริมกำลัง มีตัวอย่างรายงาน ดังนี้

1. รายงานผลการฝึกตัวแบบ

```

--- Episode 1 / 4 ---
---Episode 1 of 4 done...
---Total Reward: -3.00 | Account_Balance: 996451.82

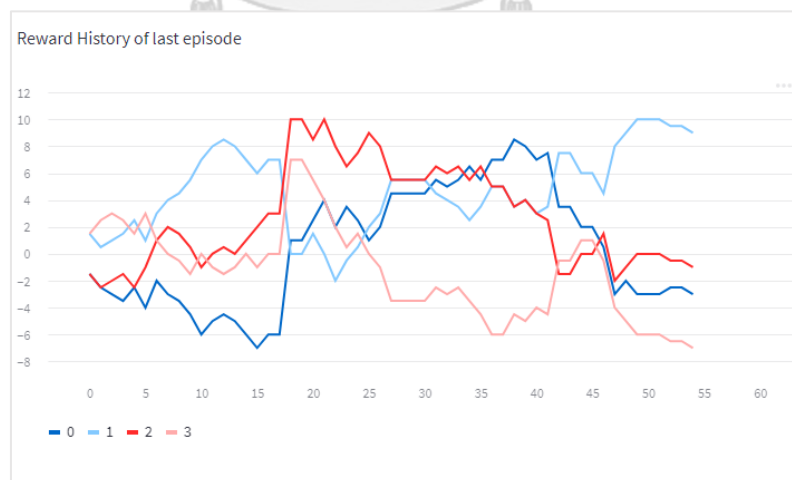
--- Episode 2 / 4 ---
---Episode 2 of 4 done...
---Total Reward: 9.00 | Account_Balance: 1001813.80

--- Episode 3 / 4 ---
---Episode 3 of 4 done...
---Total Reward: -1.00 | Account_Balance: 997803.45

--- Episode 4 / 4 ---
---Episode 4 of 4 done...
---Total Reward: -7.00 | Account_Balance: 995573.72

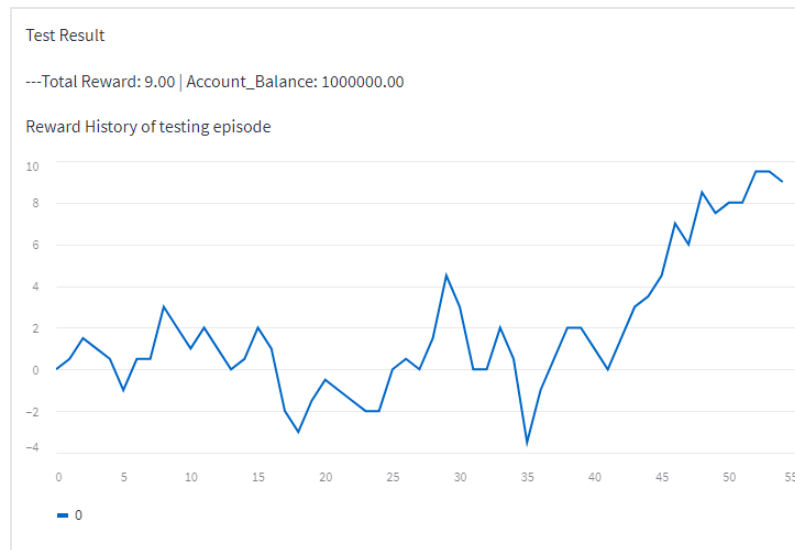
```

รูปที่ 44: Training Logs



รูปที่ 45: Reward History Chart

2. รายงานผลการทดสอบตัวแบบ



รูปที่ 46: Test Result

3. รายงานคำแนะนำของตัวแบบ



รูปที่ 47: Generate Advice and historical actions

ประวัติผู้เขียน

ชื่อ-สกุล	วิทวัส ยงสกุล
วัน เดือน ปี เกิด	13 มิถุนายน 2537
สถานที่เกิด	สุราษฎร์ธานี
วุฒิการศึกษา	จุฬาลงกรณ์มหาวิทยาลัย
ที่อยู่ปัจจุบัน	902/270, ซอย สุขุมวิท 38, ถนน สุขุมวิท, แขวง พระโขนง, เขต คลองเตย, กรุงเทพมหานคร



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY