**รางายสรุปหัวข้อการค้นคว้า**

**Forex exchange rate forecasting**

**using deep recurrent neural networks**

**จัดทำโดย**

**นายปัณณวิชญ์ พันธ์วงศ์ 600610752**

**เสนอ**

**อ.ดร.เกษมสิทธิ์ ตียพันธ์**

**รายงานเล่มนี้เป็นส่วนหนึ่งของวิชา**

**261499**

**ภาคเรียนที่ 1 ปีการศึกษา 2563**

# **สารบัญ**

[สารบัญ 1](#_Toc55044616)

[Overview 3](#_Toc55044617)

[1. Neural network architectures 4](#_Toc55044618)

[1.1 Neural network architectures 4](#_Toc55044619)

[**1.1.1.** **Feedforward Neural Networks (FNN)** 4](#_Toc55044620)

[**1.1.2.** **Recurrent Neural Networks (RNN)** 4](#_Toc55044621)

[**1.1.3.** **Long Short-Term Memory networks (LSTM)** 5](#_Toc55044622)

[**1.1.4.** **Gated Recurrent Units (GRU)** 6](#_Toc55044623)

[2. Experimental design 7](#_Toc55044624)

[2.1 Data 7](#_Toc55044625)

[2.2 Data preprocessing 7](#_Toc55044626)

[**2.2.1** **Length of the input sequences Training and trading window** 7](#_Toc55044627)

[**2.2.2** **Feature** 8](#_Toc55044628)

[**2.2.3** **Targets** 8](#_Toc55044629)

[**2.2.4** **Loss function** 8](#_Toc55044630)

[**2.2.5** **Activation function** 9](#_Toc55044631)

[**2.2.6** **Regularization** 9](#_Toc55044632)

[**2.2.7** **Regularization** 9](#_Toc55044633)

[3. Model architecture 11](#_Toc55044634)

[3.1 Train data 11](#_Toc55044635)

[3.2 Test data 12](#_Toc55044636)

[3.3 Architecture 12](#_Toc55044637)

[4. Evaluation 13](#_Toc55044638)

[4.1 การวัดความแม่นยำของการทำนาย 13](#_Toc55044639)

[**4.1.1** **Logarithmic loss (Log loss)** 13](#_Toc55044640)

[**4.1.2** **Predictive accuracy (Acc.)** 13](#_Toc55044641)

[**4.1.3** **Area under the receiver operator characteristic curve (AUC)** 13](#_Toc55044642)

[4.2 การวัดความเสี่ยงในการลงทุน 13](#_Toc55044643)

[**4.2.1** **Sharpe ratio (SR)** 13](#_Toc55044644)

[**4.2.2** **Standard deviation (SD)** 13](#_Toc55044645)

[4.3 การวัดผลตอบแทนสุทธิ 13](#_Toc55044646)

[4.4 ผลการทดลอง 14](#_Toc55044647)

[5. Conclusion 16](#_Toc55044648)

[Reference 17](#_Toc55044649)

# **Overview**

ในงานวิจัยนี้ได้นำเสนอเกี่ยวกับการใช้ Artificial Neural Network (ANN) ซึ่งประกอบไปด้วย Feedforward neural Networks (FNN), Recurrent Neural Networks (RNN), Long Short-Term Memory Networks (LSTM) และ Gated Recurrent Units (GRU) ในการเปรียบเทียบการทำนาย Classification ราคาของตลาดฟอเร็กซ์ ใน 4 คู่สกุลเงิน ประกอบไปด้วย EUR/USD ,GBP/USD ,USD/JPY และ USD/CHF โดยได้สรุปผลในรูปแบบ 2 มาตราวัดผลของแต่ละ Artificial Neural Network (ANN) ได้แก่ มาตราวัดผลความแม่นยำของผลทำนาย และ มาตราวัดผลในการลงทุน ซึ่งงานวิจัยนี้ยังได้เสนอแนวทางในการพัฒนาและปัญหาที่พบในขณะทำการทดลองเพื่อให้งานวิจัยในอนาคตสามารถนำไปพัฒนาต่อได้อีกด้วย

# **Neural network architectures**

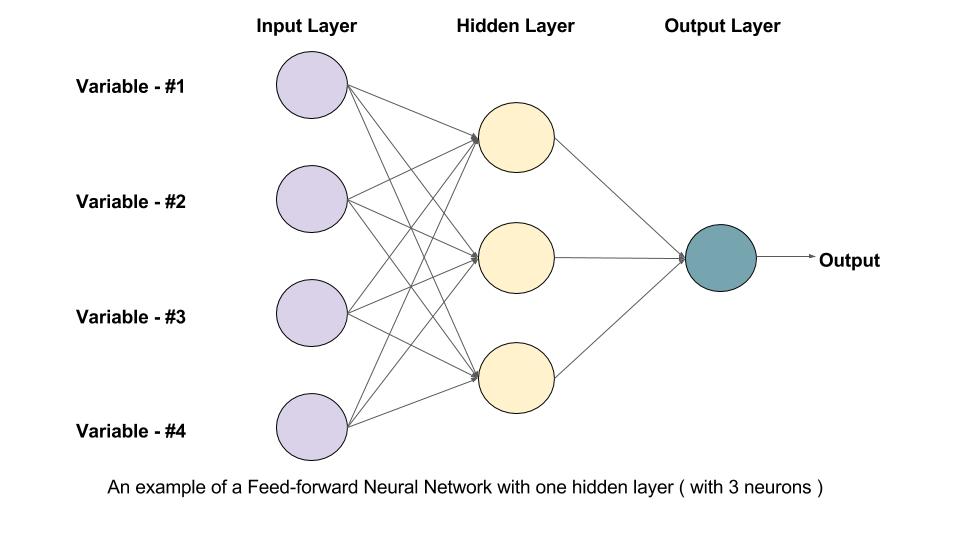
ในหัวข้อนี้จะกล่าวถึงโครงสร้างของ Neural network ที่ใช้สำหรับงานวิจัยนี้ซึ่งประกอบไปด้วย Feedforward neural Networks (FNN), Recurrent Neural Networks (RNN), Long Short-Term Memory Networks (LSTM) และ Gated Recurrent Units (GRU)

## **Neural network architectures**

ในงานวิจัยนี้จะใช้โครงสร้างของ Artificial Neural Network ในการทดลองทั้งหมด 4 โครงสร้าง

### **Feedforward Neural Networks (FNN)**

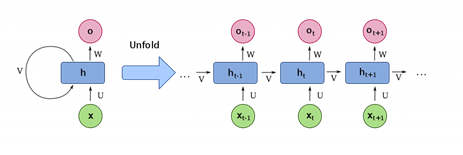
คือโครงข่ายประสามเทียมที่เป็นการจำลองมาจากสมองของเรา แต่ในส่วนคอมพิวเตอร์นั้นไม่ได้มีโครงข่ายที่ซับซ้อนเหมือนกับสมองของเรา มันมีหน้าที่แค่รันโปรแกรมตามคำสั่งของเราเท่านั้น ดังนั้นเมื่อเราจะให้มันทำการเรียนรู้อะไรซักอย่างจึงเป็นเรื่องยากในรูปแบบปกติ จึงเกิดการจำลองแนวทางการเรียนรู้ของคนไปสู่คอมพิวเตอร์ด้วย Neural Network นั่นเอง

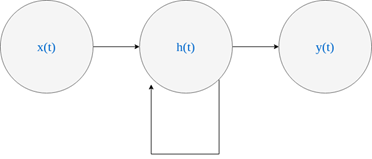


รูปที่ 1 แสดงตัวอย่างโครงสร้างของ FNN  
(<https://www.learnopencv.com/)>

### **Recurrent Neural Networks (RNN)**

คือ Artificial Neural Network แบบหนึ่งที่ออกแบบมาแก้ปัญหาสำหรับงานที่ข้อมูลมีลำดับ Sequence ชื่อ Recurrent มันก็แปลตรงตัวแล้วว่าจะมี “การเกิดซ้ำ” กล่าวคือ มีการเรียก hidden state จาก state ก่อนๆมา เช่น เรียกข้อมูลจาก Hidden state h(t-1) ที่ได้จากการคำนวณจากข้อมูล x(t-1) มารวมคำนวณกับ Hidden state h(t) ที่ได้จากากรคำนวณจากข้อมูล x(t) ด้วยนั่นเอง

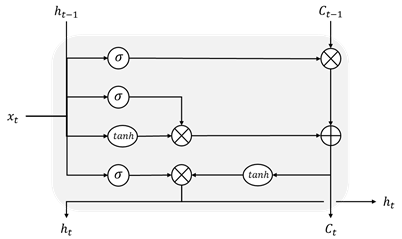




รูปที่ 2 แสดงโครงสร้างการทำงานของ RNN  
(<https://www.bualabs.com/>)

### **Long Short-Term Memory networks (LSTM)**

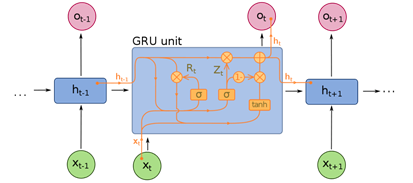
LSTM เป็น Neural Network ถูกพัฒนามาจาก RNNs หลักการของ LSTMS คือการเก็บ ‘สถานะ’ ของแต่ละ node เอาไว้ด้วย เผื่อว่าตอนย้อนกลับมาจะได้รู้ว่าค่านี้แท้จริงแล้วเป็นอะไรมาก่อน สิ่งที่ทำให้ LSTMs โดดเด่นขึ้นมานั้นก็คือการที่มันสามารถเลือกได้ว่า ข้อมูลไหนที่ควรจะจดจำ ข้อมูลไหนที่ควรจะกำจัดทิ้งออกไป ผ่านการ ‘ลืม’ ของสถานะใน node นั้น ๆ Forget gate คือประตูที่จะใช้เป็นตัวกำหนดว่า ข้อมูลที่เข้ามานั้นสมควรที่จะได้ออกไปทั้งหมดหรือไม่.

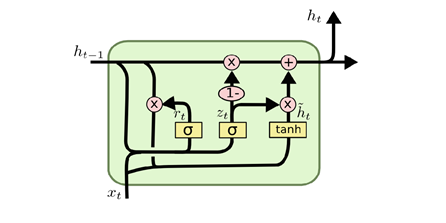


รูปที่ 3 แสดงโครงสร้าง Forget gate  
(<https://medium.com/>)

### **Gated Recurrent Units (GRU)**

อัพเกรดมาจาก RNNs เป็นกลไลปิดเปิดการอัพเดทสถานะภายใน Recurrent Neural Network ที่คล้ายกับ Long Short-Term Memory (LSTM) ที่จะมี Forget Gate แต่มี Parameter น้อยกว่า LSTM เนื่องจากไม่มี Output Gate





รูปที่ 4 แสดงโครงสร้างของ GRU  
(<https://medium.com/>)

# **Experimental design**

ในหัวข้อนี้จะกล่าวถึงกระบวนการทดลองของงานวิจัยนี้ซึ่งประกอบไปด้วย คำอธิบายข้อมูล, วิธีการเตรียมข้อมูล รวมถึงการกำหนดค่าไฮเปอร์พาริมเตอร์

## **Data**

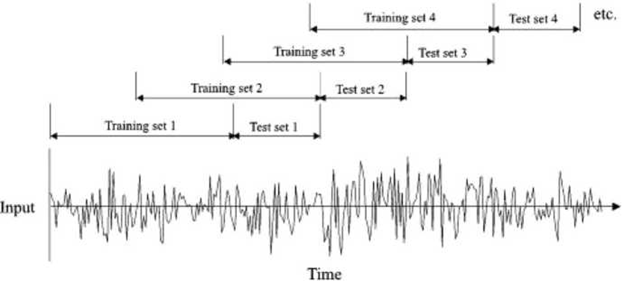
เป็นชุดข้อมูลของ 4 คู่สกุลเงินหลัก ได้แก่ EUR/USD ,GBP/USD ,USD/JPY และ USD/CHF ข้อมูลของแต่ละคู่สกุลเงินมี 12,710 แถว ที่เป็นข้อมูลกรอบเวลา 1 วัน อยู่ในช่วงเวลา 4 ม.ค. 1971 จนถึง 25 ส.ค. 2017 ที่ใข้งานวิจัยก่อนหน้า (Kuan and Liu 1995; Tenti 1996; Giles et al. 2001; Kiani and Kastens 2008; Hussain et al. 2008)

## **Data preprocessing**

### **Length of the input sequences Training and trading window**

ทำการแบ่งข้อมูลสำหรับการเทรน และสำหรับทดสอบ โดย sequences length สำหรับอินพุตมีความยาว τ = 240 และทำการแบ่งข้อมูลเป็นทั้งหมด 4 ชุด (van de Gucht et al. 1996)

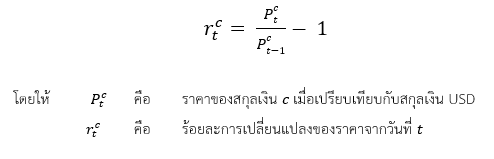
โดยชุดที่ 1 จะนำข้อมูลบางส่วนจากชุดที่ 2 มาใช้เทรนและทดสอบ , ชุดที่ 2 จะนำข้อมูลบางส่วนในชุดที่ 3 มาใช้เทรน และทดสอบ ชุดต่อไปก็จะทำแบบเดียวกันกับชุด 1 และ 2 จนไปถึงชุดที่ 4 ซึ่งกระบวนการนี้เรียกว่า the window-based cross-validation



รูปที่ 5 แสดงการแบ่งข้อมูล

### **Feature**

ใช้เพียงแค่ 1 feature เป็นอัตราแลกเปลี่ยนที่แสดงถึงราคาเมื่อเปรียบเทียบกับสกุลเงิน USD โดยคำนวณจากร้อยละการเปลี่ยนแปลงของราคาจากวันที่ t



โดยอินพุตที่จะเข้าเทรนในโมเดล จะประกอบไปด้วย ลำดับของร้อยละการเปลี่ยนแปลงของราคาจากเวลาที่ t ของตั้งแต่วันที่ จนวันที่ t



### **Targets**

เป็นปัญหา binary classification problem ซึ่ง target จะเป็นร้อยละการเปลี่ยนแปลงของราคาวันถัดไป โดยอ้างอิงงานวิจัย Fischer and Krauss (2018) จะกำหนดเป็นดังนี้



โดยถ้าค่า r ถ้ามากว่าหรือเท่ากับ 0 จะกำหนดให้เเป็น 1 และกรณีอื่นๆให้เป็น 0

### **Loss function**

ซึ่งในการเทรนจะใช้ cross-entropy ระหว่าง ค่าที่ทำนายและค่าจริง



### **Activation function**

ซึ่งแต่ละนิวรอนเน็ตเวิร์คใช้ activation function ดังนี้

|  |  |  |
| --- | --- | --- |
| **โครงสร้าง** | **Activation ภายใน** | **Activation output** |
| FNN | ReLu | Sigmoid |
| RNN | Tanh |
| LSTM | Sigmoid,Tanh |
| GRU |

โดย FNN ใช้ Activation function เป็น ReLu เนื่องจากต้องการปรับปรุงการไหลของ Gradient และ ลดระยะเวลาในการเทรน

ในเอาต์พุตของแต่ละโครงสร้างนั้นใช้ Activation function เป็น Sigmoid เนื่องจาก ค่าอยู่ในช่วง [ 0 , 1 ] ซึ่งเป็นค่าของความน่าจะเป็น

### **Regularization**

ใช้วิธีการ Dropout เป็นการสุ่มถอดบาง neurons ออก ในระหว่างการเทรน

### **Regularization**

วิจัยนี้ได้ทำการทดลองเปลี่ยนค่าตามตารางดังกล่าว และได้สรุปผลในการหาโครงสร้างที่ดีที่สุดของแต่ละโดเมลในหัวข้อถัดไป

|  |  |
| --- | --- |
| **ชื่อ** | **ปริมาณ** |
| Hidden Layer | 1,2,3,4 |
| Neurons per Hidden Layer | 25,50,100,200,400,800,1600 |
| Dropout | 0 ถึง 60% , เพิ่มทีละ 10% |
| Optimizer and learning rate | Adam และ RMSprop |
| Batch size | 16,32,64,128,256 |

# **Model architecture**

## **Train data**

โดยรายละเอียดของการเทรนของแต่ละโครงสร้างนั้นมีรายละเอียดดังนี้

|  |  |
| --- | --- |
| **รายการ** | **รายละเอียด** |
| ปริมาณข้อมูล | 750 |
| ความยาวอินพุต | 240 (ข้อ 2.2.1)  \*สำหรับ FNN ความยาว คือ 1 |
| ระยะข้อมูล | -1 ถึง 1 |
| Mini batch size | 32 |
| Optimizer | Adam(default parameters) |
| Epochs | 100 epochs โดยหยุดก่อนกำหนดหลังจาก 10 ช่วงเวลาโดยไม่ต้องปรับปรุง validation loss 20% |

## **Test data**

ใช้ข้อมูลสำหรับทดสอบ 250 ข้อมูล โดยทำการ Preprocessing ตามข้อ 2.2.3

## **Architecture**

สำหรับทุกๆโครงสร้างมีรายละเอียดดังนี้

|  |  |
| --- | --- |
| Hidden Layer | 3 |
|  |  |
| Neurons Per Hidden Layer | 50 |
| Dropout | 25% หลังแต่ละ Hidden Layer |

# **Evaluation**

สำหรับการวัดผลของงานวิจัยนี้ ได้ทำการวัดผลในทั้งหมด 2 มาตราวัดได้แก่ การวัดความแม่นยำของระบบ และ การวัดผลในการลงทุน

## **การวัดความแม่นยำของการทำนาย**

ใช้วิธีการวัดความแม่นยำทั้งหมด 3 รูปแบบ ดังนี้

### **Logarithmic loss (Log loss)**

เป็น loss function สำหรับ minimize ขณะกำลังเทรน

### **Predictive accuracy (Acc.)**

  ใช้จำแนกประสิทธิภาพในการแบ่ง class

### **Area under the receiver operator characteristic curve (AUC)**

ซึ่งถ้าหากพื้นที่ใต้โค้ง ROC มาก จะหมายถึงมีความถูกต้องมากตาม

## **การวัดความเสี่ยงในการลงทุน**

### **Sharpe ratio (SR)**

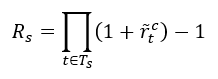
ใช้สำหรับเป็นมาตราวัดความเสี่ยงในการลงทุน

### **Standard deviation (SD)**

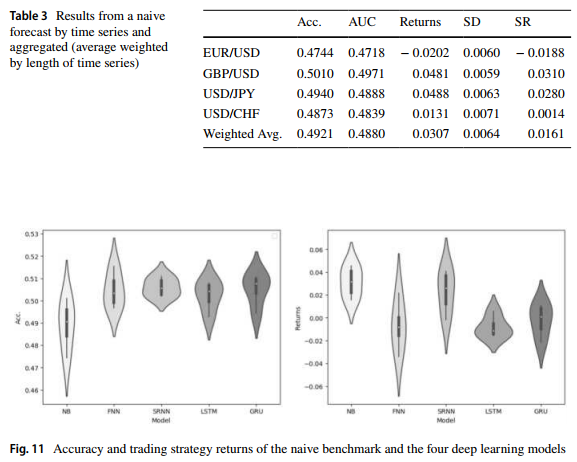
ค่าเบี่ยงเบนมาตราฐานใช้วิเคราะห์ผลตอบแทนการลงทุน

## **การวัดผลตอบแทนสุทธิ**

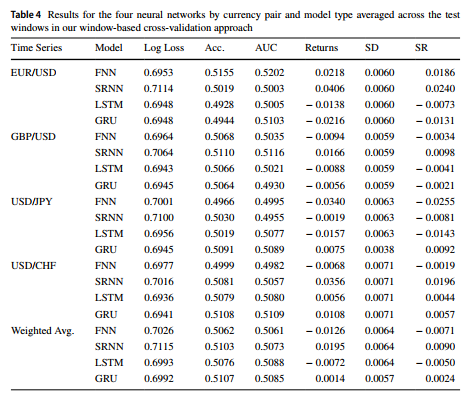
ในงานวิจัยนี้จะกล่าวถึง ผลตอบแทนที่ได้(Rs) (ระยะเวลาประมาณ 1 ปี) จากชุดทดสอบประกอบไปด้วยข้อมูล 240 วันซื้อขาย โดยใช้สมการดังนี้

. 

## **ผลการทดลอง**



จากตารางที่ 3 เป็นผลลัพธ์การทำนายจากงานวิจัย พบว่า คู่สกุลเงิน GBP/USD มีค่าความแม่นยำเป็นอันดับหนึ่งเมื่อเปรียบเทียบกับคู่สกุลเงินอื่นๆ และอันดับสุดท้ายคือ คู่สกุลเงิน EUR/USD ที่มีความแม่นยำต่ำสุด



ตารางที่ 4 แสดงค่าเฉลี่ยของการทดสอบในรูปแบบ multiple test windows ดังที่กล่าวในหัวข้อ **2.2.1 Length of the input sequences, Training and trading window** ซึ่งสรุปได้ 3 ประการดังนี้

* **Training loss (Log Loss)**

ในส่วนของ LSTM และ GRU ทำงานได้ดีกว่า FNN และ RNN ในแต่ละคู่สกุลเงิน เพียงเล็กน้อย เนื่องจาก LSTM และ GRU เหมาะสมกับข้อมูลที่มีลำดับ sequences ยาวๆ

* **Economic measures**

พบว่าไม่มีโมเดลไหนเลยที่สามารถให้ผลตอบแทนในทางที่ดีได้ ทั้งในแง่ของ ผลตอบแทนสุทธิ และความเสี่ยงในการลงทุน ซึ่งจากตารางที่ 4 จะเห็นได้ว่า RNN มีประสิทธิภาพเทียบเท่ากับ LSTM และ GRU

* **AUC**

ในส่วนของพื้นที่ใต้รูปโค้งพบว่า มีค่าต่างกันไม่มากนัก ซึ่งในแต่ละคู่สกุลเงิน  
ก็จะมีโมเดลที่ทำได้ดีที่แตกต่างกันไป

# **Conclusion**

งานวิจัยนี้ทำการทดลอง Forex classification โดย Target ที่กำหนดไว้ คือ ร้อยละการเปลี่ยนแปลงของราคาวันถัดไป ซึ่งเป็นประเภท Binary classification ซึ่งมีการทดลองทำทั้งหมด 4 โมเดล ได้แก่ FNN,RNN,LSTM และ GRU โดยมี feature คือ ร้อยละการเปลี่ยนแปลงของราคาในเวลา t ก่อนหน้า ซึ่งมาตราวัดผล หรือ การวัดควาแม่นยำของระบบที่จะ 2 กลุ่มหลักๆ คือ มาตราวัดความแม่นยำของระบบ และมาตราวัดในทางการเงิน โดยผลสรุปที่ได้ในการทดลองพบว่า ทุกๆโมเดลมีความแม่นยำในการทำนาย(Acc) อยู่ที่ประมาณ 50% และมีความเสี่ยงในการลงทุนค่อนข้างสูง ซึ่งข้อเสนอแนะสำหรับงานวิจัยในอนาคตมีดังนี้

* **Feature**

feature ที่ใช้มีเพียงแค่ค่าของร้อยละการเปลี่ยนแปลงของราคา ซึ่งอาจจะเป็น feature ที่น้อยไปหรือไม่เหมาะสมกับการทำนายตลาดฟอเร็กซ์

* **Forex market**

เนื่องจากตลาด Forex มีความผันผวนค่อนข้างสูงกว่าตลาดการเงินอื่นๆ ทำให้ความแม่นยำของระบบทำนายไม่สามารถคาดคะเนได้

* **Amounts of data**

ปริมาณข้อมูลอาจจะมีผลต่อการเรียนรู้ของระบบซึ่งหากน้อยเกินไปก็จะทำให้ผลของการทำนายไม่ดีนัก

# **Reference**

[ 1 ] Kuan, C. M., & Liu, T. (1995). Forecasting exchange rates using feedforward and recurrent neural networks. Journal of Applied Econometrics, 10(4), 347–364

[ 2 ] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929–1958.

[ 3 ] Giles, C. L., Lawrence, S., & Tsoi, A. C. (2001). Noisy time series prediction using recurrent neural networks and grammatical inference. Machine Learning, 44(1), 161–183.

[ 4 ] Kiani, K. M., & Kastens, T. L. (2008). Testing forecast accuracy of foreign exchange rates: Predictions from feed forward and various recurrent neural network architectures. Computational Economics, 32(4), 383–406.

[ 5 ] Hussain, A. J., Knowles, A., Lisboa, P. J. G., & El-Deredy, W. (2008). Financial time series prediction using polynomial pipelined neural networks. Expert Systems with Applications, 35(3), 1186–1199

[ 6 ] van de Gucht, L. M., Dekimpe, M. G., & Kwok, C. C. Y. (1996). Persistence in foreign exchange rates.

Journal of International Money and Finance, 15(2), 191–220.

[ 7 ] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.

[ 8 ] Alexander Jakob Dautel, Wolfgang Karl Härdle, Stefan Lessmann, Hsin‑Vonn Seow . (2020). Forex exchange rate forecasting using deep recurrent neural networks.