

ỨNG DỤNG MẠNG TRÍ TUỆ NHÂN TẠO KHAI PHÁ DỮ LIỆU TRONG GIAO DỊCH NGOẠI HỐI

Nguyễn Thị Thu Thủy

Trường Đại học Thương mại

Email: NguyenTithuthuy@gmail.com; Thuynguyenthithu@vcu.edu.vn

Đặng Xuân Vương

Trường Phổ thông Trung học chuyên Sơn La

Ngày nhận: 15/12/2016

Ngày nhận bản sửa: 18/01/2017

Ngày duyệt đăng: 25/8/2017

Tóm tắt:

Thị trường giao dịch ngoại hối (FoReign Exchange Market) được biết đến như một thị trường giao dịch tài chính sôi động và hoạt động liên tục từ 21 giờ (GMT) Chủ nhật tới 21 giờ (GMT) Thứ sáu hàng tuần. Bài toán giao dịch ngoại hối là việc dự báo chiến lược mua vào hoặc bán ra của các cặp tỷ giá khác nhau trên thị trường căn cứ vào tập dữ liệu quá khứ đủ lớn. Mạng trí tuệ nhân tạo và đặc biệt là mô hình máy véc tơ hỗ trợ (Support Vector Machine – SVM) đã cải thiện đáng kể các kết quả hỗ trợ cho việc dự báo. Ví dụ như hệ số sinh lời tăng từ 0.77 (không áp dụng mô hình SVM) lên 3.03 (có áp dụng mô hình SVM).

Từ khóa: Thị trường ngoại hối, giao dịch, Support Vector Machine (SVM), hỗ trợ, dự báo.

Applying neural networks for mining foreign exchange data

Abstract:

Foreign Exchange Market is a dynamic financial market in the world. It has been worked every week from 9PM on Sunday until 9PM on Friday. The main question is to predict the buying/selling exchange rate between alternative money such as USD vs. EUR, which is based on the numerous historical data from exchange market. Neural Network, in particular Support Vector Machine, can improve the predicted results. For example, profitability can be increased from 0.77 (without SVM model) up to 3.03 (with SVM model).

Keywords: Foreign exchange market; exchange; support vector machine (SVM); support; prediction.

1. Giới thiệu

“Forex” là cụm được bắt nguồn từ “FoReign Exchange” (trao đổi ngoại tệ). Đây cũng là tên gọi của thị trường tiền tệ quốc tế với các giao dịch tài chính hàng ngày rất lớn. Chính vì khối lượng giao dịch hàng ngày lớn nên nó tạo ra một thị trường có tính thanh khoản rất cao. Theo Levinson (2005) và Lê Văn Tư (2006), thị trường ngoại hối là nơi diễn ra các hoạt động mua bán và trao đổi ngoại tệ thông qua quan hệ cung cầu. Việc trao đổi bao gồm việc mua một đồng tiền này và đồng thời bán một đồng tiền khác. Như vậy, các đồng tiền được trao đổi từng

cặp với nhau.

Các mô hình về mạng trí tuệ nhân tạo được xây dựng cho việc tìm kiếm các tri thức trong tập dữ liệu mà không cần sự tác động của con người tạo ra (Baasher & Fakhr, 2011; Punniyamoorthy & Thoppan, 2013). Hệ thống tự huấn luyện để sinh ra các tập luật, các chiến lược quyết định dựa trên dữ liệu đầu vào và đầu ra. Các mô hình này có thể được xây dựng như là một nhà đầu tư thực sự trên thị trường giao dịch với đầy đủ các tri thức và kỹ năng để phân tích thị trường. Theo Carney & Cunningham (1996), mô hình mạng trí tuệ nhân tạo có những ưu điểm trong

việc phân tích dữ liệu. Đó là, khi mạng được huấn luyện với tập dữ liệu không đủ lớn thì hiệu quả của việc phân lớp không giảm chất lượng đi nhiều.

Trong bài báo này, mạng trí tuệ nhân tạo và cụ thể là mô hình máy véc tơ hỗ trợ (Support Vector Machine – SVM) được sử dụng để khai phá các dữ liệu giao dịch trên thị trường ngoại hối. Thông qua việc dự báo từ mô hình máy véc tơ hỗ trợ, các chiến lược mua vào/bán ra của các cặp tỷ giá sẽ được xây dựng. Nội dung bài báo được chia thành các mục sau. Mục 2 sẽ giới thiệu chung về giao dịch ngoại hối và bài toán giao dịch ngoại hối. Mục 3 sẽ đề cập đến ứng dụng của mạng trí tuệ nhân tạo trong các bài toán giao dịch ngoại hối, và đề xuất các bước cụ thể để giải quyết bài toán giao dịch ngoại hối. Phần thực nghiệm sẽ được trình bày ở mục 4, và kết luận chung sẽ được trình bày ở mục 5.

2. Bài toán giao dịch ngoại hối

Thị trường ngoại hối là một thị trường đặc biệt trong thị trường tài chính tiền tệ. Nó có một số đặc điểm cơ bản như: Có lợi nhuận cao nhờ giao dịch ngoại hối thực tế cao hơn so với tỷ lệ lạm phát; hoặc nếu có một đồng tiền giảm giá thì có nghĩa là có một đồng tiền nào đó sẽ tăng giá; bảo toàn vốn nếu có biến động về kinh tế; và thực hiện liên tục và không phụ thuộc vào vị trí địa lý.

Thành phần tham gia thị trường ngoại hối cũng khá đa dạng như: Các ngân hàng thương mại; các ngân hàng trung ương; các cá nhân hay các nhà kinh doanh (khách hàng mua bán lẻ); các nhà môi giới ngoại hối; hoặc các doanh nghiệp. Các doanh nghiệp tham gia vào thị trường ngoại hối sẽ chủ yếu là các doanh nghiệp có chức năng kinh doanh xuất nhập khẩu. Các doanh nghiệp này vừa là chủ thể có nhu cầu về ngoại tệ để thanh toán các hợp đồng thương mại quốc tế, vừa là chủ thể cung ngoại tệ khi có các khoản thu về việc xuất khẩu hàng hoá dịch vụ. Vì vậy, các doanh nghiệp này được xem như là chủ thể hình thành nên khối lượng mua và bán ngoại hối lớn nhất trên thị trường ngoại hối.

Do đặc thù của thị trường này là ngoại hối được giao dịch qua các ngân hàng, nhà môi giới, chuyên viên ngoại hối, các tổ chức tài chính và các cá nhân trên toàn thế giới nên ngoại hối được giao dịch liên tục từ 21 giờ (GMT) Chủ nhật tới 21 giờ (GMT) Thứ sáu hàng tuần. Cũng do múi giờ khác nhau trên thế giới, nên tại mỗi thời điểm sẽ có một số trung tâm hoạt động và một số trung tâm không hoạt động. Các trung tâm này hoạt động gối nhau như Hình 1.

Nền tảng của thị trường giao dịch ngoại hối là Internet. Các nhà đầu tư (nhà giao dịch) sử dụng nền tảng này để tham gia thị trường. Họ có thể mua hoặc bán một cặp tiền tệ cụ thể để hưởng chênh lệch tỷ giá giữa các cặp tiền này. Ví dụ, cặp tiền tệ có thể là EUR/USD (Euro so với US Dollar), USD/JPY (US Dollar Vs Japanese Yen), GBP/USD (Great Britain Pound vs US Dollar) hoặc nhiều sự kết hợp các cặp tiền tệ khác.

Tỷ giá ngoại hối thường bị ảnh hưởng bởi một số các yếu tố như các sự kiện kinh tế hoặc chính trị, hay tâm lý của nhà đầu tư. Chính các điều này làm cho thị trường giao dịch không ổn định và phức tạp. Nhà đầu tư tham gia vào thị trường ngoại hối luôn tìm mọi cách để giải thích cho sự biến động của tỷ giá, và họ hy vọng việc dự báo của mình là chính xác. Dự báo chính xác tỷ giá và đúng thời điểm cho ra lợi nhuận tối đa. Ngược lại, giao dịch dựa trên tính toán sai có thể dẫn đến nguy cơ mất lợi nhuận.

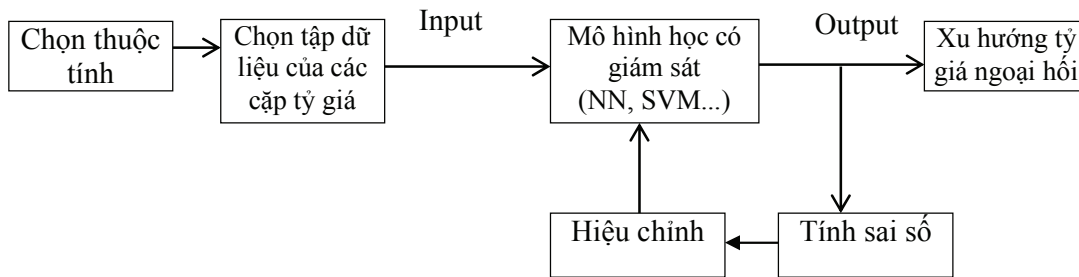
Bài toán giao dịch trên thị trường ngoại hối là bài toán dự báo xu hướng theo ngày của tỷ giá quy đổi của các cặp tiền tệ theo chuỗi thời gian biến động phức tạp. Đầu vào của bài toán là lịch sử tỷ giá trao đổi của các cặp tiền tệ, số liệu về các sự kiện kinh tế đã xảy ra. Đầu ra của bài toán chính là xu hướng tăng hay giảm của các cặp tiền tệ cần xét. Một số lượng lớn các thuộc tính được lấy từ dữ liệu chuỗi thời gian sinh ra từ lịch sử giao dịch giữa tỷ giá các cặp tiền tệ sẽ được xem xét. Việc thu thập này nhằm mục đích chọn ra các tập dữ liệu phù hợp cho bài toán phân lớp.

Hình 1. Lịch hoạt động của các trung tâm giao dịch ngoại hối

12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5	6	7	8	9	10	11

Nguồn: Tác giả tổng hợp

Hình 2. Sơ đồ xây dựng mô hình học có giám sát dự báo xu hướng tỷ giá ngoại hối



Trên góc độ về khoa học máy tính, bài toán giao dịch ngoại hối sẽ được chuyển về bài toán phân lớp nhị phân là dự đoán xu hướng tăng hay giảm để hỗ trợ đưa ra các quyết định giao dịch như: hoạt động mua, bán, chốt lời, dừng lỗ. Các mô hình học máy, ví dụ như mô hình máy véc tơ hỗ trợ SVM áp dụng cho việc dự báo (Kim, 2003), đã được sử dụng với mục đích là dự báo xu hướng tỷ giá của các cặp tiền tệ trên thị trường ngoại hối.

3. Mô hình mạng trí tuệ nhân tạo trong dự báo xu hướng tỷ giá ngoại hối

Việc dự báo tỷ giá sử dụng mô hình dự báo chuỗi thời gian được biết đến với việc sử dụng phổ biến mô hình Regressive Integrated Moving Average (ARIMA) (Box & Jenkin, 1970). Tuy nhiên, ARIMA được phát triển dựa trên giả thiết rằng chuỗi thời gian dự báo là tuyến tính và không biến động. Thời gian gần đây, mô hình dự báo chứng khoán và ngoại hối được sử dụng phổ biến là GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) và các nghiên cứu chỉ ra rằng nó tốt hơn việc sử dụng mô hình ARIMA (Kumar, 2006; Yu, 2002).

Ngoài các mô hình mạng Nơ ron (NN) (xem chi tiết ở các nghiên cứu của Gan & Ng, 2010; và Nemeş & Butoi, 2013), mô hình SVM cũng được biết đến như là các mô hình được sử dụng để dự báo tỷ giá ngoại hối. Mô hình SVM được Cortes & Vapnik (1995) đề xuất lần đầu tiên trong các bài toán kỹ thuật. Tuy nhiên, Kim (2003) đã áp dụng mô hình SVM này cho dự báo tỷ giá ngoại hối. Sau đó, mô hình SVM cũng đã được nhiều nhà nghiên cứu khác như Liu & Wang (2008), Baasher & Fakhr (2011), và Sercan & cộng sự (2013) áp dụng.

Ở nghiên cứu của Kim (2003), Liu & Wang (2008), và Sercan & cộng sự (2013), mô hình SVM đã được sử dụng cho việc dự báo thông qua việc đề xuất tham số của biên trên và hàm nhân Kernel. Trong nghiên cứu này, mô hình SVM với hàm nhân

Kernel trong nghiên cứu của Kim (2003), và Sercan & cộng sự (2013) sẽ được vận dụng cho dữ liệu thực nghiệm. Tuy nhiên, các tham số của mô hình sẽ được điều chỉnh trong quá trình thực nghiệm để chọn ra mô hình với kết quả tốt hơn.

Căn cứ vào những phân tích ở trên, rõ ràng mô hình học có giám sát được sử dụng rất nhiều trong bài toán giao dịch ngoại hối. Mô hình học có giám sát dự báo xu hướng tỷ giá ngoại hối có thể thực hiện theo các bước như Hình 2. Các bước đề xuất này được xây dựng căn cứ trên quy trình khai phá dữ liệu của Fayyad & cộng sự (1996). Trong các bước này, việc tính sai số và hiệu chỉnh mô hình được lặp đi lặp lại để chọn ra mô hình với các tham số tốt nhất cho việc dự báo xu hướng tỷ giá ngoại hối cuối cùng.

Các bước được tiến hành như sau:

Bước 1: Chọn tập thuộc tính. Các thuộc tính trong tập dữ liệu thô thường bao gồm cả những thuộc tính không cần thiết cho bài toán dự báo. Việc chọn ra các thuộc tính có liên quan đến bài toán dự báo chính là thu hẹp phạm vi và hướng đến mục đích của bài toán giao dịch ngoại hối do người dùng đặt ra.

Bước 2: Xác định tập dữ liệu. Chi phí và chất lượng của dữ liệu cũng nên được chú ý ở bước này. Một số vấn đề khác cần lưu ý gồm: phương pháp tính toán, dữ liệu không bị sửa chữa trước đó, có một độ trễ phù hợp với dữ liệu, và sự đảm bảo rằng nguồn dữ liệu này vẫn còn có thể sử dụng trong tương lai. Đây là những điểm lưu ý dựa trên kinh nghiệm thực nghiệm nhiều lần trước đó. Ví dụ, trong nghiên cứu này, việc thực nghiệm đã được thực hiện nhiều lần đối với nhiều tập dữ liệu khác nhau nhằm mục đích rút ra được tập dữ liệu cuối cùng phù hợp cho việc cài đặt robot giao dịch ở giai đoạn sau.

Bước 3: Xử lý dữ liệu. Dữ liệu được chia thành các tập huấn luyện, kiểm thử và kiểm chứng (Training, Testing, và Validation) và đưa vào mô hình học có

giám sát. Các giá trị của thuộc tính cần được biến đổi vào đoạn [0,1]. Theo Beleites & cộng sự (2013), kích thước của tập kiểm thử thường bằng từ 10 đến 30% kích thước của tập huấn luyện. Tỷ lệ này nhằm mục đích tạo ra được mô hình học với kết quả tốt. Trong nghiên cứu này, tập kiểm thử, và tập huấn luyện cũng được lấy theo tỷ lệ trên. Trong đó, dữ liệu của tập kiểm thử sẽ được lấy là các mẫu dữ liệu ở những khoảng thời gian gần đây nhất.

Bước 4: Thiết lập cấu trúc của mô hình học có giám sát. Ở bước này, nghiên cứu sẽ chọn mô hình, các tham số cho mô hình tương ứng.

Bước 5: Đánh giá. Việc đánh giá được thực hiện thông qua việc sử dụng độ lệch (hay còn gọi là lỗi) giữa output dự báo và output trong tập huấn luyện. Phương pháp chủ yếu được dùng với bài toán giao dịch ngoại hối là MSE (Mean Square Error- sai số toàn phương trung bình). Việc đánh giá này còn kiểm tra xem có rơi vào tình trạng Over-Training (huấn luyện quá) không để có thể phải thay đổi số lần lặp hoặc các tham số cho phù hợp. Tình trạng Over-Training xảy ra là do các mẫu kiểm thử bị rơi vào các trường hợp đặc biệt của tập huấn luyện. Điều này dẫn đến mô hình sẽ cho ra kết quả tốt. Tuy nhiên, nếu thay đổi tập kiểm thử bằng các mẫu kiểm thử khác kết quả thì điều đó không còn đúng nữa.

Đầu ra cuối cùng (output) của mô hình chính là việc xác định được xu hướng tỷ giá ngoại hối của cặp tiền tệ mong muốn. Mô hình nói trên sẽ được áp dụng vào cho robot giao dịch ngoại hối, xu hướng tỷ giá ngoại hối của mô hình sẽ được dùng để là quyết định thực hiện các giao dịch của robot.

4. Thực nghiệm

4.1. Lựa chọn tập thuộc tính

Việc lựa chọn tập thuộc tính làm không gian đầu vào cho mô hình là rất quan trọng, nó có tính chất quyết định đến hiệu quả huấn luyện mô hình cũng như hiệu quả khi đưa vào sử dụng. Việc lựa chọn tập thuộc tính dựa phần lớn vào kinh nghiệm của người xây dựng mô hình. Trong bài báo này, việc đề xuất xây dựng mô hình phân lớp nhị phân SVM và cài đặt cho robot giao dịch trên phần mềm giao dịch MetaTrader 4 với thông số sau: Sử dụng mục tiêu lợi nhuận khoảng 10 – 15 Pips cho mỗi giao dịch; Các dữ liệu phân tích của giá mở cửa, đóng cửa, cao, thấp cho một số cây nến gần nhất trên các khung thời gian M1 (1 phút), M5 (5 phút), M15 (15 phút), H1 (1 giờ) và D1 (hàng ngày).

Ngoài ra, nghiên cứu còn sử dụng thêm các chỉ báo kỹ thuật như RSI (Relative Strength Index – Chỉ số sức mạnh tương đối), các đường MA (Moving Average - đường trung bình), Giải băng Bollinger

Bảng 1. Mô tả lựa chọn tập thuộc tính

Feature# (số thuộc tính)	Dữ liệu	Số lượng	Ghi chú
#1-12	O, H, L, C trên Timeframe M1	12	4 giá trị x 3 nến gần nhất
#13-24	O, H, L, C trên Timeframe M5	12	4 giá trị x 3 nến gần nhất
#24-32	O, H, L, C trên Timeframe M15	8	4 giá trị x 2 nến gần nhất
#33-40	O, H, L, C trên Timeframe H1	8	4 giá trị x 2 nến gần nhất
#41-52	O, H, L, C trên Timeframe D1	12	4 giá trị x 3 nến gần nhất
#53-57	Dữ liệu về thời gian	5	Các thuộc tính thời gian
#58-59	RSI(7) trên Timeframe M5, M15	2	1 giá trị x 2 Timeframe
#60-61	RSI(14) trên Timeframe M5, M15	2	1 giá trị x 2 Timeframe
#62-69	MA(9), MA(12)	8	4 giá trị x 2 Timeframe
#70-75	Custom Indicator	6	2 PAX + 4 MKC
#76-79	Bollinger Bands	4	2 Bands x 2 Timeframe
#80-83	Average True Range	4	2 giá trị x 2 Timeframe
#84-91	Highest, Lowest M1, M5	8	4 giá trị x 2 Timeframe

(Bollinger Bands) và thêm một vài các chỉ báo kỹ thuật tùy biến (Custom Indicators). Các thuộc tính lựa chọn cho mô hình được thể hiện ở Bảng 1.

4.2. Thu thập và tiền xử lý dữ liệu

Việc thu thập và tiền xử lý dữ liệu được tiến hành ngay trên phần mềm MetaTrader 4 bằng một Script (một đoạn code nhỏ) để tải (download) và bổ sung dữ liệu cho phần mềm MetaTrader 4. Việc này đảm bảo dữ liệu đầy đủ (từ ngày 01/01/2013 đến ngày 30/9/2016) và đáng tin cậy. Trong quá trình thu thập và xử lý dữ liệu trên phần mềm MetaTrader 4, cũng cần phải chú ý đến vấn đề chênh lệch múi giờ giữa các Server (máy chủ) của các Broker (nhà đầu tư) đặt ở các khu vực khác nhau, để đảm bảo đồng bộ trong quá trình huấn luyện và sử dụng mô hình. Vì vậy, cần chuyển đổi sang giờ GMT (múi giờ 0) bằng cách đưa vào thêm các tham số:

- Server_Offset: Múi giờ của của khu vực đặt Server;

- DST (Daylight Saving Time): Khu vực đặt Server của Broker có điều chỉnh thời gian tiết kiệm ánh sáng không?

- DSTStart: Ngày bắt đầu điều chỉnh trong năm (Chỉ dùng khi DST = true);

- DSTEnd: Ngày thôi điều chỉnh DST (Chỉ dùng khi DST = true).

Với các tham số trên thì ta có thể quy đổi thời gian trên Server về thời gian GMT như sau:

- Nếu DST = true: $Time_{GMT} = TIME_{Server} - (Server_{offset} + 1)$ (giờ);

- Nếu DST = false: $Time_{GMT} = TIME_{Server} - Server_{offset}$ (giờ).

Quá trình thu thập dữ liệu cũng cần đảm bảo số lượng mẫu dương và số lượng mẫu âm trong tập dữ liệu không quá chênh lệch. Điều này sẽ được đảm bảo một cách tương đối nếu ta chọn hai giá trị SL (Stop Loss – Dừng lỗ) và TP (Take Profit – Chốt lời) xấp xỉ nhau.

4.3. Lựa chọn tham số cho mô hình

Quá trình thực nghiệm đã được thực hiện thông qua việc lựa chọn hai hàm nhân cho mô hình là hàm Gaussian RBF (RsGBF), và hàm nhân đa thức (Polynomial) của mô hình SVM (Support Vector Machine). Để lựa chọn các tham số, việc tiến hành thử nghiệm được lặp đi lặp lại nhiều lần để so sánh, đánh giá, điều chỉnh với mô hình. Các tham số thử nghiệm được lựa chọn với các hàm nhân trên như sau:

$$\text{Hàm nhân GsRBF: } k(x,y) = e^{-\frac{\|x-y\|^2}{2\sigma}} = e^{-\gamma\|x-y\|^2}$$

đã được thử nghiệm một số giá trị $\gamma \in (0;5]$ và thông thường, $\gamma \in [1;2]$ cho kết quả chấp nhận được.

Hàm nhân Polynomial: $k(x,y) = (x \cdot y + \theta)^d$ trong đó $\theta \in \mathbb{R}$, $d \in \mathbb{N}^*$ đã được thử nghiệm với $d=2,3,4$ và $\theta \in [0;1]$. Quá trình thử nghiệm cho thấy $d=3$ và $\theta=1$ cho kết quả khá tốt.

Tham số C đã được thử nghiệm với một số giá trị $C \in [1;10]$ và thông thường $C \in [1;2]$ cho kết quả chấp nhận được.

4.4. Dữ liệu huấn luyện

Cặp tiền tệ EURUSD (Euro vs US Dollar) đã được lựa chọn để thử nghiệm với mô hình do tính phổ biến của chúng trên thị trường giao dịch ngoại hối. Việc thu thập dữ liệu được chia thành 2 tập không giao nhau:

Bảng 2. Bảng lựa chọn tham số cho mô hình

Mô hình Tham số	Poly1	Poly2	Poly3	GsRBF1	GsRBF2	GSRBF3
Tham số C	2.0	1.0	1.0	1.0	2.0	1.0
Hàm nhân	Đa thức	Đa thức	Đa thức	GsRBF	GsRBF	GsRBF
Tham số	Bậc 2	Bậc 3	Bậc 3	$\gamma=2.0$	$\gamma=2.0$	$\gamma=1.0$
Số vectors hỗ trợ	1000	1500	1890	2500	3000	3000

- Tập D gồm các mẫu thu được trong giai đoạn 01/01/2013 đến 31/12/2015.

- Tập D' gồm các mẫu thu được trong giai đoạn 01/01/2016 đến 30/9/2016.

Tập D sẽ sử dụng trong quá trình lựa chọn tham số cho mô hình bằng các phương pháp hold-out (splitting)¹ và Cross Validation (k-fold)². Sau đó toàn bộ tập D sẽ được dùng để huấn luyện mô hình và đánh giá, so sánh giữa các mô hình bằng tập để chọn mô hình tốt nhất.

4.5. Kết quả thực nghiệm

Với việc lựa chọn tập thuộc tính gồm 91 thuộc tính ở Bảng 1, sau khi lựa chọn các tham số chấp nhận được, chúng tôi đã lựa chọn ra 6 mô hình với các tham số như tại Bảng 2.

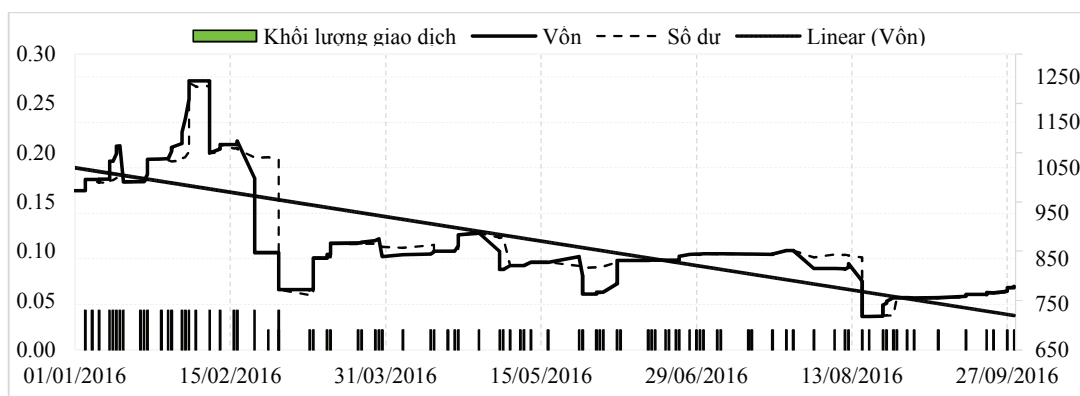
Việc thực nghiệm được tiến hành với 6 mô hình nói trên. Kết quả của các mô hình với hàm Kernel là Polynominal (Poly1,2,3 trong Bảng 2) cho kết quả tốt hơn các mô hình sử dụng hàm Gaussian RBF

(GsRBF 1,2,3) (độ chính xác tương ứng khoảng 70% so với 65%). Việc cài đặt robot với mô hình Poly2 được thực hiện vì độ chính xác của mô hình này cao hơn một lượng nhỏ so với các mô hình Poly 1,3 còn lại.

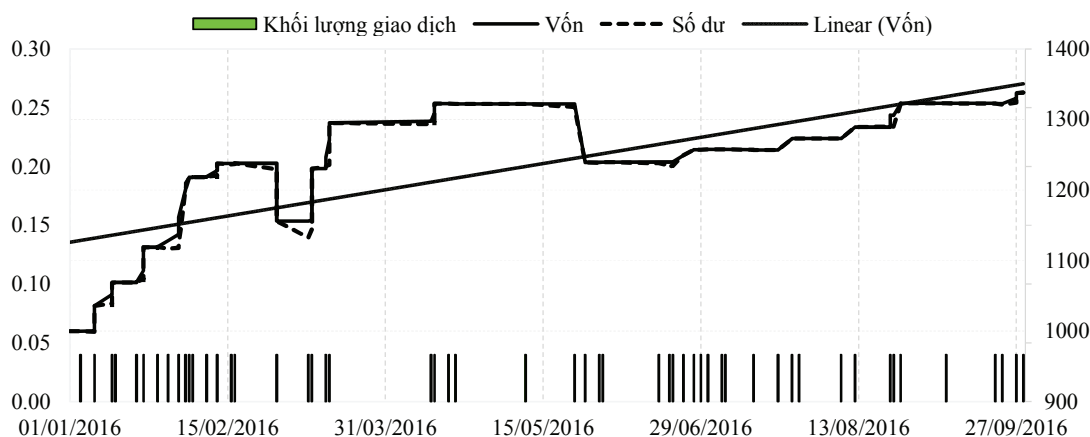
4.5.1. Cài đặt trên Robot giao dịch ngoại hối

Robot giao dịch ngoại hối (Expert Advisor) chạy trên phần mềm MetaTrader 4 là một chương trình được viết bằng ngôn ngữ MQL4, biên dịch bởi MetaEditor và chạy như một “phần bổ trợ” (Plugin) cho MetaTrader 4. Thực nghiệm được tiến hành với Broker giả định: ICMarkets và cặp tiền tệ giao dịch: EUR/USD (Euro vs US Dollar). Mức ký quỹ ban đầu: 1000USD, và đòn bẩy (Margin Level): 1:500. Kiểu tài khoản là ECN (Electronic Communication Network) vì kiểu tài khoản này có tốc độ thực hiện giao dịch nhanh, thường có spread (chênh lệch giá mua và bán) thấp nhưng có phí hoa hồng cho nhà môi giới khoảng 7 USD cho 1.0 lot EUR/USD giao dịch. Spread giả định (chênh lệch giá mua và bán):

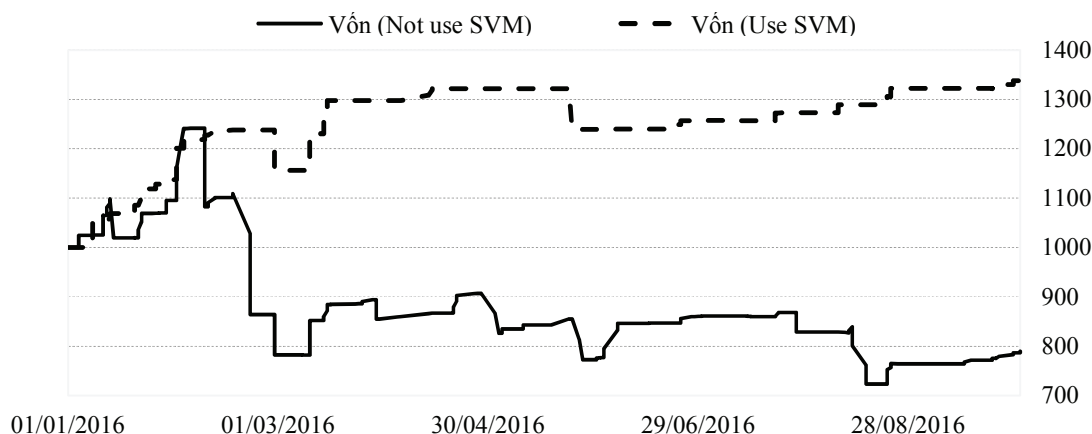
Hình 4. Kết quả giao dịch của robot trong thử nghiệm 1 (không sử dụng SVM)



Hình 5. Kết quả giao dịch của robot (có sử dụng SVM)



Hình 6. Tương quan vốn tích lũy của hai giao dịch



Bảng 3. So sánh tương quan chỉ số của 2 giao dịch trong thử nghiệm

STT	Chỉ số	Không dùng SVM		Có dùng SVM	
1	Lãi ròng		-210.46		337.94
2	Lãi gộp		718.37		504.7
3	Lỗ gộp		-928.83		-166.77
4	Hệ số sinh lời		0.77		3.03
5	Tăng trưởng		-21.0%		33.8%
6	Tăng trưởng bình quân (tháng)		-2.59%		3.29%
7	Drawdown tuyệt đối		371.72		151.5
8	Drawdown tương đối	645.36	50.67%	450.67	34.69%
9	Tổng số giao dịch		189		96
10	Giao dịch sinh lời	147	77.8%	85	88.5%
11	Giao dịch thua lỗ	42	22.2%	11	11.5%
12	Giao dịch bán sinh lời	111	91.0%	71	100.0%
13	Giao dịch mua sinh lời	78	59.0%	25	56.0%

8 – 10 point (0.8 – 1.0 pip).

4.5.2. Phân tích các kết quả thử nghiệm

Với điều kiện cài đặt ở trên, nghiên cứu đã tiến hành thử nghiệm với robot qua 2 thử nghiệm: Có sử dụng mô hình SVM và không sử dụng mô hình SVM. Các kết quả được trình bày qua các hình 4, 5, và 6.

Để có cái nhìn khái quát về kết quả của hai giao dịch trên, các chỉ số được thể hiện việc thử nghiệm với mô hình SVM và không sử dụng mô hình SVM ở Bảng 3.

Qua một số chỉ số quan trọng như Lãi ròng, Hệ số sinh lời, Tỷ lệ tăng trưởng bình quân tháng, Bảng 3 cho biết việc giao dịch có sử dụng kết hợp mô hình SVM cho kết quả tốt hơn hẳn. Tuy nhiên, với

mức Drawdown (tỷ lệ giảm sút của tài khoản) ~35% cho thấy giao dịch này vẫn tiềm ẩn nguy cơ rút vốn vẫn còn cao hơn mức chấp nhận được (khoảng dưới 30%).

5. Kết luận

Thông qua việc nghiên cứu các lý thuyết về các mô hình học máy, chúng ta thấy được ý nghĩa và khả năng áp dụng các mô hình này là rất lớn, để thay thế con người giải quyết các vấn đề trong cuộc sống. Việc thực nghiệm xây dựng mô hình SVM, cài đặt cho robot, thử nghiệm, so sánh và đánh giá kết quả đã thể hiện việc có thể vận dụng các lý thuyết học máy vào giải quyết các bài toán thực tế, và cụ thể ở đây là vận dụng mô hình SVM cho bài toán giao dịch ngoại hối. Tuy nhiên, bản thân thị trường ngoại

hồi luôn rất khó lường, luôn có những thời điểm biến động không theo một quy luật nào đã có trước đó. Vì vậy, không thể có một mô hình nào có thể dự báo được một cách chính xác tuyệt đối xu hướng những biến động trong tương lai.

Hơn nữa, trong số 6 mô hình SVM với các tham

số đề xuất, độ chính xác (khoảng 70%) của mô hình cài đặt cho robot còn chưa như kỳ vọng của các nhà đầu tư. Với việc cải tiến mô hình, lựa chọn các tập thuộc tính đủ lớn hay thay đổi tham số của mô hình sẽ hy vọng cải thiện được kết quả của các robot giao dịch trong tương lai.

Phần mềm sử dụng

Metatrader 4, (2016). Retrieved on January 15th 2017, from < <http://www.metaquotes.net/>>

Strategy Quant Tick Downloader. Retrieved on January 15th 2017, from <<http://www.strategyquant.com/tickdownloader>>

Tickstory Lite. Retrieved on January 15th 2017, from < <https://tickstory.com>.>

Ghi chú:

1. Hold-out là phương pháp chia tập dữ liệu thành hai tập huấn luyện và kiểm tra.
2. Cross Validation là phương pháp chia tập dữ liệu ngẫu nhiên thành k tập con, quá trình huấn luyện và kiểm tra sẽ được thực hiện k lần.

Tài liệu tham khảo

Baasher, A.A., Fakhr, M.W. (2011), '*FOREX Trend Classification using Machine Learning Techniques*', World Scientific and Engineering Academy and Society (WSEAS) Proceedings, Recent Researches in Applied Informatics and Remote Sensing, USA.

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C. Popp, J. (2013), '*Sample size planning for classification models*', Anal Chim Acta, 760, 25-33.

Box, G.E.P. & Jenkins, G. M. (1970), '*Time Series Analysis: Forecasting and Control*', Holden- Day conference Proceedings, San Francisco, CA.

Carney, J.G. & Cunningham, P. (1996), '*Neural Networks and Currency Exchange Rate Prediction*', Foresight Business, Retrieved on January 15th 2017, from <<http://www.maths.tcd.ie/pub/fbj/forex4.html>>.

Cortes C. & Vapnik V. (1995), '*Support Vector Networks*', Machine Learning, 20, 273-297.

Fayyad, U., Shapiro, G.P. & Smyth, P. (1996), '*From Data Mining to Knowledge Discovery in Databases*'. American Association for Artificial Intelligence, 37-54.

Fxstreet (2017), Ngày truy cập: 15/01/2017 từ trang <http://www.fxstreet.com.vn>

Gan, W-S. & Ng, K-H. (2010), '*Multivariate FOREX: Forecasting using Artificial Neural Networks*', IEEE Xplore, DOI: 10.1109/ICNN.1995.487560.

Kim, K. (2003), '*Financial time series forecasting using support vector machines*', Neurocomputing, 55, 307 – 319.

Kumar, S.S.S. (2006), '*Forecasting Volatility – Evidence from Indian Stock and Forex Market*', Indian Institute of Management Kozhikode, Retrieved on January 15th 2017, from <<http://hdl.handle.net/123456789/775>>.

Lê Văn Tư (2006), '*Ngân hàng và thị trường tài chính*', Nhà xuất bản Tài chính, Hà Nội

Levinson, M. (2005), '*The Economist – Guide to the Financial Markets*', 4th Edition, Profile Books Publisher, USA.

Liu, L. & Wang, W. (2008), '*Exchange Rates Forecasting with Least Squares SVM*', Computer Science and Software Engineering Proceedings, IEEE.

Nemeş, M.D. & Butoi, A. (2013), '*Data Mining on Romanian Stock Market Using Neural Networks for Price Prediction*', Informatica Economică, 17 (3), 125-136.

Punniyamorthy, M., & Thoppan, J. J. (2013), '*ANN-GA based model for stock market surveillance*', Financial Crime, 20 (1), 52-66.

Sercan, A., Burc, E. & Adam, G. (2013), '*Supervised classification-based stock prediction and portfolio optimization*', CS 229 Project Milestone Report, retrieved on January, 15th 2017, from <<http://cs229.stanford.edu/proj2013>>.

Yu, J. (2002), '*Forecasting Volatility in the New Zealand Stock Market*', Applied Financial Economics, 12, 193-202.