**BÁO CÁO BÀI TOÁN DỰ ĐOÁN XU HƯỚNG CỦA CHỨNG KHOÁN**

**--o--**

1. **Giới thiệu tổng quan**
2. **Bài toán dự đoán xu hướng giá chứng khoán**

Ngày nay, sự phát triển mạnh mẽ của của các thị trường chứng khoán đã đem lại lợi nhuận to lớn cho nhiều nhà đầu tư. Việc nhận định đúng xu hướng của thị trường là một vấn đề cốt yếu đem lại thu nhập khổng lồ cho họ từ những khoản đầu tư nhỏ. Thế nhưng, thị trường chứng khoán tuy không thay đổi ngẫu nhiên nhưng để xác định đúng xu hướng hiện tại hoặc những điểm thay đổi xu hướng là một quá trình trải qua kinh nghiệm lâu dài đối với bất kỳ nhà đầu tư nào. Vì vậy, yêu cầu đặt ra là nhà đầu tư cần sự hỗ trợ tốt nhất cho những quyết định với việc giải quyết hiệu quả bài toán dự đoán xu hướng giá chứng khoán

1. **Một số hướng tiếp cận**

## Tình hình trong nước

Không có nhiều những phương pháp tốt được áp dụng cho thị trường chứng khoán Việt Nam. Một nguyên nhân đầu tiên có thể dễ dàng nhận thấy đó là sự non trẻ của thị trường dẫn đến việc một số nghiên cứu áp dụng một vài vấn đề lý thuyết trong chứng khoán vào thị trường Việt Nam không chính xác. Hơn nữa, các phương pháp hầu hết là áp dụng khai thác dữ liệu để đưa ra các luật tác động trong nội tại thị trường. Chính điều này làm cho các phương pháp đó mang lại hiệu quả không cao.

## Tình hình thế giới

Nhận thấy có 2 xu hướng chính: mô hình dự đoán giá hoặc xu hướng giá trong tương lai và mô hình hỗ trợ ra quyết định mua bán.

1. *Mô hình dự đoán giá hoặc xu hướng giá*

Thông thường, có rất nhiều đề tài nghiên cứu hướng về việc sử dụng mạng nơ-ron nhân tạo (ANN). Gần đây nhất, các nghiên cứu hướng về Support vector machine (SVM-SVR) vì những khuyết điểm được khắc phục từ ANN. Bên cạnh đó, cũng có một số phương pháo khác như sử dụng các mô hình Grey model, hay mô hình Markov. Đơn cử một vài trường hợp còn sử dụng phương pháp Text mining.

* Phương pháp ANN

Đặc điểm của phương pháp này cần một bộ dữ liệu đầu vào lớn, sự quyết định mô hình các tầng của ANN và cách thức để tối ưu các tham số. Thường các nghiên cứu gần đây áp dụng lai nhiều phương pháp:

* ANN + Fuzzy (tối ưu hóa kết quả dự đoán)
* ANN + GA (tối ưu hóa việc chọn tham số)
* Recurrent ANN (tối ưu phương pháp huấn luyện nhằm tránh tối ưu cục bộ)
* *Ưu và nhược điểm*
* Áp dụng cho bất cứ bộ dữ liệu mà không cần các luật hay những quy ước về cấu trúc.
* Thích nghi khá tốt với sự hỗn loạn của thị trường
* Một số phương pháp lai cho những kết quả tốt nhờ biết khai thác tối đa đặc trưng dữ liệu
* Cần phải thực nghiệm nhiều lần mới mong có được bộ tham số tối ưu cho mô hình
* Nếu dữ liệu quá biến động, thường không tổng quát được, dẫn đến tình trạng quá khớp.
* Phương pháp SVM - SVR

Phương pháp này có đặc trưng là khả năng tổng quát hóa cao, đặc biệt quyết định khi số chiều dữ liệu là lớn(vừa phải). Nó có thể khắc phục được những nhược điểm của ANN trong việc tránh nguy cơ quá khớp và dữ liệu training cũng không cần quá lớn như ANN.

Thông thường để tăng tính động của SVM, một số nghiên cứu gần đây áp dụng sự cải tiến trong quá trình phản hồi trong khi học với mô hình recurrent SVM. Ngoài ra một số phương pháp lai cũng là tăng khả năng của mô hình này

* SVM + Fuzzy
* SVM + các phương pháp chọn dữ liệu đầu vào ( tối ưu hóa các biên dữ liệu)
* SVM + GA
* SVM + Box theory (nâng cao khả năng dự đoán)
* *Ưu và nhược điểm*
* Cũng như ANN, nó mềm dẻo với mối quan hệ phi tuyến của thị trường.
* Dữ liệu được tổng quát hóa cao hơn.
* Với việc áp dụng các phương pháp lai, mô hình thích nghi với cơ chế động của thị trường, khả năng chọn những tham số tối ưu là rất tốt.
* Tuy nhiên, sự quyết định của mô hình nằm ở kernel function nên cần sự xem xét kỹ lưỡng cho phần này.
* Dễ dàng dẫn đến mô hình phức tạp nếu không cân nhắc trong việc quyết định số chiều của không gian đầu vào.

1. *Mô hình hỗ trợ ra quyết định mua bán*

Thông thường, có rất nhiều nghiên cứu thiên về việc xây dựng cây quyết định, luật kết hợp, hay các thuật toán phân lớp. Do đặc trưng mô hình này không liên quan nhiều đến bài toán đặt ra nên khảo sát này tạm thời bỏ qua.

* *Ưu và nhược điểm*

Có quan tâm đến ý nghĩa giữa các thành phần trong ảnhđể giúp việc nhận diện tốt hơn. Tuy nhiên, với hình ảnh chỉ là một logo kích thước không lớn, việc phân biệt những thành phần chính đôi khi không mang lại nhiều hiệu quả và thậm chí tỷ lệ so khớp thành công có phần thấp hơn cả GD.

1. **Các phương pháp nghiên cứu**

Như đã giới thiệu ở trên, bài toán dự đoán xu hướng có hai hướng tiếp cận chính là ANN và SVM-SVR. Với sự phân tích và những kết quả thực nghiệm của những nghiên cứu, dễ dàng nhận thấy khả năng giải quyết tốt bài toán của phương pháp SVM là rất lớn. Hơn nữa, với đặc trưng thị trường Việt Nam, dữ liệu không quá lớn, và một số kết quả khẳng định rằng khả năng dự đoán đối với những thị trường mới nổi là khả quan thì việc chọn mô hình SVM làm nhân cho giải pháp hứa hẹn sẽ chứng tỏ hiệu quả của nó. Sau đây là các ý tưởng cho phương pháp nghiên cứu:

* Với khả năng tổng quát hóa cao, SVM sẽ được chọn để làm nhân cho phương pháp này.
* Để tránh nhiễu khi áp dụng các tham biến (chỉ số), giá đóng cửa vẫn là đầu vào chính. Tuy nhiên, nhóm sẽ nghiên cứu để có được một vài yếu tố kỹ thuật (có thể liên quan đến khối lượng giao dịch hoặc vài chỉ số kỹ thuật) nhằm nâng cao khả năng phối hợp của dữ liệu.
* Áp dụng thêm các phương pháp (như GA, Fuzzy, tùy vào tình hình) để cải thiện việc chọn tham số cũng như tính động và hiệu quả của mô hình.
* Kiểm thử với các mô hình khác, dự tính là một phương pháp ANN đã cải tiến và phương pháp SVM truyền thống.

1. **Nội dung và các vấn đề đi sâu nghiên cứu**
2. **Phương pháp ANN và các cải tiến**

Các lý thuyết và mô hình ANN và những cải tiến đã mang lại hiệu quả trong việc giải quyết bài toán dự đoán xu hướng sẽ được tìm hiểu kỹ để xây dựng nên mô hình so sánh. Xoáy sâu vào những nhược điểm khi cài đặt thực tế để đưa ra được mô hình khắc phục hiệu quả cho việc nghiên cứu.

1. **Phương pháp SVM và các cải tiến**

Cài đặt mô hình SVM truyền thống, kèm theo việc phát hiện ra những hạn chế. Từ đó, đưa ra thêm một phương pháp lai để khắc phục những hạn chế đó.

1. **Các vấn đề kỹ thuật trong chứng khoán**

Vấn đề các biến đầu vào cho mô hình sẽ được nghiên cứu kỹ lưỡng. Trong phần này, nhóm sẽ cố gắng đưa ra những cấu trúc cũng như những biến đầu vào liên quan phù hợp nhất cho mô hình