MỤC LỤC

[Định hướng thiết kế mô hình 2](#_Toc273655342)

[1.1 Lựa chọn dữ liệu: 2](#_Toc273655343)

[1.2 Tiền xử lý dữ liệu: 2](#_Toc273655344)

[1.3 Chuẩn hóa dữ liệu: 2](#_Toc273655345)

[1.4 Cấu trúc mạng Nơron: 2](#_Toc273655346)

[1.5 Các độ đo để đánh giá mô hình: 3](#_Toc273655347)

# Định hướng thiết kế mô hình

## Lựa chọn dữ liệu:

* Dữ liệu được sử dụng cho quá trình dự đoán trong mạng noron được áp dụng ở đây là dự đoán giá cả của những ngày sắp tới. Nên dữ liệu áp dụng vào mô hình chính là giá đóng cửa của tất cả các cổ phiếu trên thị trường từ trước đến nay.

## Tiền xử lý dữ liệu:

* Trước khi áp dụng vào mô hình thì ta cần tiến hành tiền xử lý dữ liệu như khử nhiễu bằng phương pháp trung bình chuyển động (mô hình chuẩn), hoặc phương pháp Independent component analysis (Cải tiến) giúp mô hình có khả năng dự đoán giá chính xác hơn. Đối với các nước phát triển và ổn định thì tỉ lệ nhiễu thường thấp hơn so với các nước đang phát triển như Việt Nam. Nên việc khử nhiễu này sẽ đóng một vai trò quan trọng trong mô hình.

## Chuẩn hóa dữ liệu:

* Để đưa dữ liệu vào mô hình, ta cần tiến hành chuẩn hóa các giá trị đầu vào thành các giá trị [0,1] thông qua công thức sau:
* Đây là công thức dựa trên mô hình ARIMA phát triển.

## Cấu trúc mạng Nơron:

* Ta sử dụng mạng 3 lớp lan truyền ngược với luật học tổng quát delta. Ở đây là lớp xuất chỉ có một nơron là giá của ngày dự đoán. Số node ở lớp đầu vào và lớp ẩn sẽ qua thực nghiệm mà lựa chọn phù hợp.
* Một số công thức áp dụng cho mô hình:
* Trong đó:
* *c,p là nơ ron lớp hiện hành, lớp trước đó.*
* *ic,p là giá trị đầu vào*
* *wc,p là trọng số kết nối*
* Ở lớp xuất nơron đầu ra được tính theo công thức:

## Các độ đo để đánh giá mô hình:

* Ta sẽ đánh giá kết quả theo 3 độ đo sau: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), đây là 3 phương pháp được dùng nhiều nhất trong các mô hình dự đoán giá chứng khoán.
* Ta có công thức như sau:
* Trong đó:
* *Ai là giá trị thực*
* *Pi là giá trị dự đoán*
* *N là chiều dài chuỗi giá.*