

**ĐẠI HỌC THÁI NGUYÊN**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

---

**NGUYỄN THỊ NGỌC TÚ**

**NGHIÊN CỨU MỘT SỐ KỸ THUẬT NGOẠI SUY VÀ  
ỨNG DỤNG**

**Chuyên ngành: Khoa học máy tính**

Thái Nguyên - 2015

## LỜI CAM ĐOAN

Tôi xin cam đoan Luận văn "*Nghiên cứu một số kỹ thuật ngoại suy và ứng dụng*" đã được thực hiện theo đúng mục tiêu đề ra dưới sự hướng dẫn của TS. Nguyễn Văn Huân. Kết quả đạt được trong luận văn là sản phẩm của cá nhân tôi. Trong toàn bộ luận văn, những điều được trình bày là của cá nhân và được tổng hợp từ nhiều nguồn tài liệu. Tất cả các tài liệu tham khảo đều có xuất xứ rõ ràng và được trích dẫn hợp pháp.

Tôi xin chịu hoàn toàn trách nhiệm và chịu mọi hình thức kỷ luật theo quy định cho lời cam đoan của mình.

*Thái Nguyên, ngày tháng năm 2015*

***Người cam đoan***

**Nguyễn Thị Ngọc Tú**

## LỜI CẢM ƠN

Tôi xin bày tỏ lời cảm ơn chân thành tới tập thể các thầy cô giáo Viện Công nghệ thông tin - Viện Hàn lâm Khoa học và Công nghệ Việt Nam, các thầy cô giáo trường Đại học Công nghệ thông tin và Truyền thông - Đại học Thái Nguyên đã tận tình giảng dạy cũng như tạo mọi điều kiện để tôi học tập và nghiên cứu trong 2 năm học cao học.

Tôi xin chân thành cảm ơn sâu sắc tới thầy giáo TS. Nguyễn Văn Huân đã cho tôi nhiều sự chỉ bảo quý báu, đã tận tình hướng dẫn và tạo điều kiện cho tôi hoàn thành tốt luận văn tốt nghiệp này.

Quá trình thực hiện đề tài không tránh khỏi các thiếu sót, tôi rất mong tiếp tục nhận được sự đóng góp ý kiến của các thầy giáo, cô giáo, các bạn đồng nghiệp đối với đề tài nghiên cứu của tôi để đề tài được hoàn thiện hơn.

*Tôi xin chân thành cảm ơn!*

*Thái Nguyên, ngày tháng năm 201*

*Người cam đoan*

**Nguyễn Thị Ngọc Tú**

## MỤC LỤC

LỜI CAM ĐOAN.....	I
LỜI CẢM ƠN .....	III
DANH MỤC HÌNH VẼ.....	IV
DANH MỤC BẢNG BIỂU.....	VIII
MỤC LỤC .....	IV
MỞ ĐẦU.....	1
Chương 1 GIỚI THIỆU CHUNG VỀ DỮ LIỆU CHUỖI THỜI GIAN.....	5
1.1. Khái niệm Dữ liệu .....	5
1.1.1. Dữ liệu định tính và dữ liệu định lượng.....	5
1.1.2. Dữ liệu sơ cấp và dữ liệu thứ cấp.....	5
1.2. Chuỗi thời gian và Dữ liệu chuỗi thời gian.....	7
1.2.1. Chuỗi thời gian.....	7
1.2.2. Dữ liệu chuỗi thời gian.....	12
1.3. Tổng quan về dự báo.....	16
1.3.1. Đặc điểm của dự báo.....	16
1.3.2. Các loại dự báo.....	17
1.3.3. Các phương pháp dự báo .....	21
Chương 2 MỘT SỐ KỸ THUẬT NGOẠI SUY .....	22
2.1. Khái niệm ngoại suy .....	22
2.1.1. Khái niệm ngoại suy.....	22
2.1.2. Khi nào nên sử dụng phương pháp ngoại suy để dự báo. ....	22
2.1.3. Ưu nhược điểm của phương pháp ngoại suy .....	23
2.1.4. Tính chính xác của dự báo .....	23
2.1.5. Ứng dụng của kỹ thuật ngoại suy vào bài toán dự báo.....	25
2.2. Ngoại suy dựa dựa vào mô hình hồi quy tuyến tính .....	26
2.2.1. Bài toán hồi quy .....	26
2.2.2. Các phương pháp đưa về dạng tuyến tính.....	29

2.2.3. Hồi quy nhiều chiều (Hồi quy bội) .....	30
2.3. Ngoại suy bằng phương pháp trung bình động giản đơn (moving average forecast).....	31
2.4. Ngoại suy bằng phương pháp san bằng mũ (Exponential Smoothing Methods).....	33
2.5. Ngoại suy dựa vào mô hình san mũ Holt- Winters. ....	36
2.5.1. San mũ Holt.....	36
2.5.2. San mũ Holt- Winters .....	37
2.6. Ngoại suy dựa vào mô hình ARIMA (AutoRegressive Integrated Moving Average).....	39
2.6.1. Các công cụ áp dụng trong mô hình .....	40
2.6.2. Mô hình AR(p) (Auto Regression) .....	44
2.6.3. Mô hình MA(q) (Moving Average) .....	45
2.6.4. Sai phân I(d) .....	46
2.6.5. Mô hình ARIMA.....	47
2.6.6. Các bước phát triển mô hình ARIMA. ....	48
2.7. Các bước tiến hành ngoại suy .....	49
<b>Chương 3 ÁP DỤNG KỸ THUẬT NGOẠI SUY VÀO BÀI TOÁN DỰ BÁO.</b>	<b>52</b>
3.1. Mô hình ARIMA cho dự báo tài chính, chứng khoán .....	52
3.1.1. Dữ liệu tài chính.....	52
3.1.2. Mô hình ARIMA cho bài toán dự báo tài chính .....	52
3.1.3. Thiết kế mô hình ARIMA cho dữ liệu.....	53
3.2. Áp dụng .....	55
3.2.1. Môi trường thực nghiệm .....	55
3.2.2. Dữ liệu.....	55
3.2.3. Kiểm tra tính dừng của chuỗi chứng khoán VNM .....	57
3.2.4. Nhận dạng mô hình .....	57
3.2.5. Ước lượng và kiểm định với mô hình ARIMA .....	60

3.2.6 Thực hiện dự báo.....	62
KẾT LUẬN .....	64
TÀI LIỆU THAM KHẢO .....	65

## DANH MỤC HÌNH VẼ

Hình 1.1. Xu hướng giảm theo thời gian .....	13
Hình 1.2. Thành phần mùa.....	14
Hình 1.3. Thành phần chu kỳ .....	15
Hình 3.1: chọn giadongcua làm mục tiêu dự báo .....	56
Hình 3.2: Xác định $d = 0,1,2$ ? .....	57
Hình 3.3: Biểu đồ của SAC và SPAC của chuỗi giadongcua.....	58
Hình 3.4: Biểu đồ của SPAC và SAC ứng với $d=1$ .....	59
Hình 3.5: Ước lượng mô hình ARIMA(2,1,1).....	60
Hình 3.6: Kết quả mô hình ARIMA(2,1,1).....	60
Hình 3.7: Kiểm tra phần dư có nhiều trắng.....	61
Hình 3.8: Dự báo .....	62

**DANH MỤC BẢNG BIỂU**

Bảng 3.1: Dữ liệu đầu vào.....	56
Bảng 3.2: Tiêu chuẩn đánh giá các mô hình ARIMA.....	62
Bảng 3.3: Đánh giá dự báo.....	63



## MỞ ĐẦU

Sự phát triển mạnh mẽ của Công nghệ thông tin và Truyền thông đã kéo theo nhiều lĩnh vực khác phát triển như khoa học máy tính, đồ họa máy tính, khí tượng thủy văn, tự động hóa, thiết kế - xây dựng, kinh tế, tài chính, giáo dục và đào tạo, y tế,... Điều này đã góp phần vào sự nghiệp công nghiệp hóa – hiện đại hóa đất nước. Trong thực tế, đối với mỗi quốc gia, trong đó có Việt Nam muốn phát triển và trở thành các nước công nghiệp thì một trong những vấn đề không thể thiếu là phải nghiên cứu, khảo sát, phân tích tình hình phát triển của đất nước trong lịch sử và hiện tại để làm cơ sở cho việc ngoại suy hay suy luận, xây dựng các kế hoạch phát triển trong tương lai với tất cả các lĩnh vực, ngành nghề,...

Vì ngoại suy hay suy luận là một khoa học và nghệ thuật tiên đoán những sự việc sẽ xảy ra trong tương lai, trên cơ sở phân tích khoa học về các dữ liệu đã thu thập được. Khi tiến hành ngoại suy cần căn cứ vào việc thu thập, xử lý số liệu trong quá khứ và hiện tại để xác định xu hướng vận động của các hiện tượng trong tương lai nhờ vào một số mô hình toán học (Định lượng). Tuy nhiên, ngoại suy cũng có thể là một ngoại suy chủ quan hoặc trực giác về tương lai (Định tính) và để ngoại suy định tính được chính xác hơn, người ta cố loại trừ những tính chủ quan của đối tượng ngoại suy.

Ngoại suy trước hết là một thuộc tính không thể thiếu của tư duy của con người, con người luôn luôn nghĩ đến ngày mai, hướng về tương lai. Trong thời đại công nghệ thông tin và toàn cầu hóa, ngoại suy lại đóng vai trò quan trọng hơn khi nhu cầu về thông tin thị trường, tình hình phát triển tại thời điểm nào đó trong tương lai càng cao. Ngoại suy được sử dụng trong nhiều lĩnh vực khác nhau, mỗi lĩnh vực có một yêu cầu về ngoại suy riêng nên phương pháp ngoại suy được sử dụng cũng khác nhau.

Trong ngoại suy, các đại lượng khảo sát thường không được cho dưới dạng hàm liên tục. Trong khi đó, các phương pháp giải tích toán học thường

tính toán với các hàm cho bởi các công thức, do đó chúng ta không thể áp dụng trực tiếp các hàm cho dưới dạng rời rạc như thế này. Cũng có khi ta biết rằng đại lượng  $y$  là một hàm của đại lượng  $x$ , tức là  $y = f(x)$ , nhưng ta không biết biểu thức hàm  $f(x)$  mà chỉ biết một số giá trị  $y_i$  ứng với các giá trị của  $x$  tại các điểm  $x_i$ .

Thông thường  $x_0 < x_1 < x_2 < \dots < x_n$  và các điểm này có thể phân bố cách đều hoặc không cách đều nhau. Mặc dù ta chỉ biết giá trị của  $y$  tại các điểm mốc  $x_i$  nhưng trong nhiều trường hợp ta cần tính toán với các giá trị  $y$  tại các vị trí khác của  $x$ . Một vấn đề được đặt ra là cho một điểm  $x$  không thuộc các điểm  $x_i$  thì làm thế nào ta có thể tính được giá trị  $y$  tương ứng với nó, sao cho chúng ta tận dụng tối đa các thông tin đã có. Nếu giả định giá trị cần tìm gần đúng của  $y$  tại các điểm  $x$  nằm trong khoảng  $(x_0, x_n)$  thì chúng ta áp dụng các bài toán nội suy.

Nhưng trong thực tế, các bài toán không phải lúc nào các điểm  $x$  cũng nằm trong khoảng  $(x_0, x_n)$ , chẳng hạn như một số các bài toán về ngoại suy:

(1). Trong tương lai các tình huống sẽ tiếp tục diễn ra như đã xảy ra trong quá khứ; (2). Đối tượng ngoại suy quá bi quan hoặc ngược lại quá lạc quan về tình huống cần ngoại suy khi đó kết quả ngoại suy có thể bị ảnh hưởng sai lệch; (3). Đối tượng ngoại suy không biết nhiều lắm về tình huống ngoại suy. Khi đó chúng ta cần giả định rằng diễn biến trong tương lai cũng sẽ tương tự như trong quá khứ. Với các bài toán trên thì  $x$  chắc chắn sẽ nằm ngoài khoảng  $(x_0, x_n)$  và như vậy thì để tìm giá trị gần đúng của  $y$ , chúng ta không thể áp dụng được phương pháp nội suy [1].

Nhằm tìm ra giá trị gần đúng của  $y$ , một trong những phương pháp được sử dụng phổ biến hiện nay là áp dụng phương pháp ngoại suy.