

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

TRẦN THỊ DIỆU THÚY

**PHƯƠNG PHÁP NGOẠI SUY VÀ ỨNG DỤNG
TRONG DỰ BÁO Ở TRƯỜNG TRUNG HỌC PHỔ
THÔNG**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN - 2022

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

TRẦN THỊ DIỆU THÚY

**PHƯƠNG PHÁP NGOẠI SUY VÀ ỨNG DỤNG
TRONG DỰ BÁO Ở TRƯỜNG TRUNG HỌC PHỔ
THÔNG**

Chuyên ngành: KHOA HỌC MÁY TÍNH

Mã số: 84 80 101

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. VŨ MẠNH XUÂN

THÁI NGUYÊN - 2022

LỜI CAM ĐOAN

Em xin cam đoan Luận văn “Phương pháp ngoại suy và ứng dụng trong dự báo ở trường trung học phổ thông” đã được thực hiện theo đúng mục tiêu đề ra dưới sự hướng dẫn của TS. Vũ Mạnh Xuân. Kết quả đạt được trong luận văn là sản phẩm của cá nhân em. Trong toàn bộ luận văn, những điều đã 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.

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

Thái Nguyên, ngày 5 tháng 10 năm 2022

Người cam đoan

Trần Thị Diệu Thúy

LỜI CẢM ƠN

Em xin bày tỏ lời cảm ơn chân thành tới tập thể 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 em học tập và nghiên cứu trong hai năm học tại trường.

Em xin chân thành gửi lời cảm ơn sâu sắc tới thầy giáo TS. Vũ Mạnh Xuân đã cho em nhiều sự chỉ bảo quý báu, đã tận tình hướng dẫn và tạo điều kiện cho em hoàn thành luận văn tốt nghiệp này.

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

Em xin chân thành cảm ơn!

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	iv
MỤC LỤC	v
DANH MỤC CÁC BẢNG	ix
DANH MỤC CÁC HÌNH	x
MỞ ĐẦU	1
CHƯƠNG 1. BÀI TOÁN DỰ BÁO	4
1.1. Tổng quan chung về dự báo	4
1.1.1. Tổng quan về dự báo	4
1.1.2. Khái niệm dự báo	5
1.1.3. Đặc điểm của dự báo	6
1.1.4. Các loại dự báo	6
1.1.4.1. Căn cứ vào thời đoạn dự báo	6
1.1.4.2. Căn cứ vào nội dung công việc cần dự báo	6
1.1.5. Tình hình các phương pháp/mô hình dự báo ở Việt nam	7
1.1.5.1. Phương pháp ngoại suy	7
1.1.5.2. Phương pháp chuyên gia	7
1.1.5.3. Phương pháp mô hình hoá	8
1.1.6. Quy trình dự báo	8
1.2. Một số bài toán dự báo ở trường THPT	10
1.2.1. Thực trạng	10
1.2.2. Một số bài toán	12
1.3. Kết luận chương	13
CHƯƠNG II. PHƯƠNG PHÁP NGOẠI SUY	14
2.1. Khái niệm ngoại suy	14
2.1.1. Khái niệm	14
2.1.2. Sử dụng ngoại suy trong dự báo	14
2.1.3. Ưu, nhược điểm của phương pháp ngoại suy	15

2.1.4. Tính chính xác của phương pháp ngoại suy.....	15
2.2. Phương pháp ngoại suy dựa vào mô hình hồi quy.....	17
2.2.1. Hồi quy tuyến tính đơn.....	18
2.2.1.1. Sắp xếp số liệu	18
2.2.1.2. Mô hình hồi quy tuyến tính đơn.....	18
2.2.1.3. Phương pháp bình phương bé nhất	19
2.2.1.4. Trường hợp X không phải biến ngẫu nhiên	20
2.2.2. Hồi quy phi tuyến tính.....	22
2.2.3. Hồi quy bội tuyến tính.....	23
2.2.4. Hồi quy đa thức	24
2.3. Phương pháp ngoại suy dựa vào mô hình ARIMA.....	25
2.3.1. Hàm tự tương quan ACF.....	25
2.3.2. Mô hình AR(p).....	26
2.3.3. Mô hình MA(q)	27
2.3.4. Sai phân I(d).....	28
2.3.5. Mô hình ARIMA.....	28
2.3.6. Các bước phát triển mô hình ARIMA.....	29
2.4. Phương pháp ngoại suy với mạng nơ ron (neural network).....	30
2.4.1. Các ký hiệu và khái niệm.....	31
2.4.1.1. Layers	31
2.4.1.2. Units	32
2.4.1.3. Weights và Biases	32
2.4.2. Activation functions	33
2.4.2.1. Hàm sigmoid	33
2.4.2.2. Hàm tanh	34
2.4.2.3. Hàm ReLU	35
2.4.2.4. Hàm Leaky ReLU	36
2.4.3. Thuật toán backpropagation.....	36
2.5. Kết luận chương 2	37

CHƯƠNG III. ỨNG DỤNG PHƯƠNG PHÁP NGOẠI SUY VÀO BÀI TOÁN DỰ BÁO Ở TRƯỜNG THPT	38
3.1. Bài toán dự báo số lượng tuyển sinh.....	38
3.1.1. Phân tích dữ liệu.....	39
3.1.2. Kết quả thử nghiệm	41
3.1.2.1. Ngoại suy bằng phương pháp hồi quy tuyến tính đơn	41
3.1.2.2. Ngoại suy bằng phương pháp hồi quy đa thức	43
3.1.2.3. Ngoại suy bằng phương pháp hồi quy phi tuyến tính	44
3.1.2.4. Ngoại suy bằng mô hình ARIMA	46
3.1.2.4. Ngoại suy bằng phương pháp mạng nơ ron	49
3.1.3. Nhận xét	50
3.2. Bài toán dự báo điểm thi tốt nghiệp.....	50
3.2.1. Phân tích dữ liệu.....	52
3.2.2. Kết quả thử nghiệm	53
3.2.2.1. Ngoại suy bằng phương pháp hồi quy bội tuyến tính	53
3.2.2.2. Ngoại suy bằng phương pháp mạng nơ ron	54
3.2.3. Nhận xét	55
3.3. Kết luận chương	55
KẾT LUẬN	57
TÀI LIỆU THAM KHẢO	58
PHỤ LỤC	59

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Viết tắt	Viết đầy đủ
ACF	Auto Correlation Function
AR	Auto Regression
ARIMA	AutoRegressive Intergrated Moving Average
GDCD	Giáo dục công dân
MA	Moving Average
MAD	Mean Absolute Deviation
MSE	Mean Squared Error
THCS	Trung học cơ sở
THPT	Trung học phổ thông
TN	Tốt nghiệp

DANH MỤC CÁC BẢNG

Bảng 2.1: Bảng phân tích phương sai (1).....	22
Bảng 2.2: Bảng phân tích phương sai (2).....	24
Bảng 3.1: Dữ liệu từ năm 2006 đến năm 2021	38
Bảng 3.2: Kết quả và đánh giá sai số của hồi quy tuyến tính đơn	43
Bảng 3.3: Bảng kết quả và đánh giá sai số hồi quy đa thức với cách chia dữ liệu 1	43
Bảng 3.4: Bảng kết quả và đánh giá sai số hồi quy đa thức với cách chia dữ liệu 2	44
Bảng 3.5: Bảng kết quả và đánh giá sai số với hồi quy phi tuyến tính	45
Bảng 3.6: So sánh giá trị thống kê (Test Statistic) và giá trị tới hạn (Critical Value)	46
Bảng 3.7: So sánh giá trị thống kê (Test Statistic) và giá trị tới hạn (Critical Value) sau khi tính sai phân	47
Bảng 3.8: Bảng kết quả các tham số	49
Bảng 3.9: Bảng kết quả và đánh giá với mô hình ARIMA.....	49
Bảng 3.10: Bảng kết quả và đánh giá sai số với mạng nơ ron	50
Bảng 3.11: Bảng điểm năm 2020 tổng gồm 10 cột và 531 hàng dữ liệu	51
Bảng 3.12: Bảng điểm năm 2021 gồm 10 cột và 544 hàng dữ liệu	51
Bảng 3.13: Bảng điểm năm 2022 gồm 9 cột và 525 hàng dữ liệu	52
Bảng 3.14: Bảng đánh giá sai số với hồi quy bội tuyến tính	533
Bảng 3.15: Bảng kết quả dự báo điểm của học sinh với hồi quy bội tuyến tính	54
Bảng 3.16: Bảng đánh giá sai số với mạng nơ ron	534
Bảng 3.17: Bảng kết quả dự báo điểm của học sinh với mạng nơ ron	54
Bảng 3.18: Bảng phụ dự báo điểm thi tốt nghiệp môn Toán (hồi quy tuyến tính)	63
Bảng 3.19: Bảng phụ dự báo điểm thi tốt nghiệp môn Toán bằng mạng nơ ron	64

DANH MỤC CÁC HÌNH

Hình 2.1: Mô hình tuyến tính.....	19
Hình 2.2: Hoạt động của các nơ ron.....	30
Hình 2.3: Mô hình mạng nơ ron với 2 hidden layers.....	31
Hình 2.4: Các ký hiệu sử dụng trong mạng nơ ron.....	32
Hình 2.5: Đồ thị hàm sigmoid.....	33
Hình 2.6: Đồ thị hàm tanh.....	34
Hình 2.7: Đồ thị hàm ReLU.....	35
Hình 2.8: Đồ thị hàm Leaky ReLU.....	36
Hình 3.1: Biểu đồ thể hiện chuỗi dữ liệu từ bảng.....	39
Hình 3.2: Minh họa đánh giá kết quả dự báo bằng hồi quy.....	39
Hình 3.3: Biểu đồ kiểm tra tính ổn định của dữ liệu.....	46
Hình 3.4: Biểu đồ kiểm tra tính ổn định của dữ liệu sau khi tính sai phân.....	47
Hình 3.5: Xác định giá trị ACF.....	47
Hình 3.6: Đồ thị của model MA(2).....	47