

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM



NGUYỄN ĐỨC KHOAN

**PHÁT TRIỂN MỘT HỆ THỐNG XÁC ĐỊNH
GIÁ TỰ ĐỘNG CHO VÉ MÁY BAY**

LUẬN VĂN THẠC SĨ

Chuyên ngành: Công nghệ thông tin

Mã số ngành: 60480201

TP. HỒ CHÍ MINH, tháng 01 năm 2016

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**



NGUYỄN ĐỨC KHOAN

**PHÁT TRIỂN MỘT HỆ THỐNG XÁC ĐỊNH
GIÁ TỰ ĐỘNG CHO VÉ MÁY BAY**

LUẬN VĂN THẠC SĨ

Chuyên ngành: Công nghệ thông tin

Mã số ngành: 60480201

CÁN BỘ HƯỚNG DẪN KHOA HỌC: PGS. TS QUẢN THÀNH THƠ

TP. HỒ CHÍ MINH, tháng 01 năm 2016

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH TẠI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**

Cán bộ hướng dẫn khoa học: PGS. TS Quản Thành Thơ

Luận văn Thạc sĩ được bảo vệ tại Trường Đại học Công nghệ TP. HCM
ngày 20 tháng 03 năm 2016

Thành phần Hội đồng đánh giá Luận văn Thạc sĩ gồm:

TTT	Họ và tên	Chức danh Hội đồng
1	PGS.TS. Võ Đình Bảy	Chủ tịch
2	GS.TSKH. Hoàng Văn Kiếm	Phản biện 1
3	TS. Lê Văn Quốc Anh	Phản biện 2
4	TS. Lê Tuấn Anh	Ủy viên
5	TS. Nguyễn Thị Thúy Loan	Ủy viên, Thư ký

Xác nhận của Chủ tịch Hội đồng đánh giá Luận văn sau khi Luận văn đã được
sửa chữa (nếu có).

Chủ tịch Hội đồng đánh giá LV

TRƯỜNG ĐH CÔNG NGHỆ TP. HCM
PHÒNG QLKH – ĐTSĐH

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

TP. HCM, ngày -- tháng 01 năm 2016

NHIỆM VỤ LUẬN VĂN THẠC SĨ

Họ tên học viên: NGUYỄN ĐỨC KHOAN

Giới tính: NAM

Ngày, tháng, năm sinh: 17/11/1983

Nơi sinh: NINH THUẬN

Chuyên ngành: Công nghệ thông tin

MSHV: 1441860014

I- Tên đề tài:

PHÁT TRIỂN MỘT HỆ THỐNG XÁC ĐỊNH GIÁ TỰ ĐỘNG CHO VÉ MÁY BAY.

II- Nhiệm vụ và nội dung:

Xây dựng mô hình để biểu diễn việc xác định giá vé máy bay tự động dựa trên các kỹ thuật machine learning và data mining.

III- Ngày giao nhiệm vụ: 20/08/2015

IV- Ngày hoàn thành nhiệm vụ: 15/01/2016

V- Cán bộ hướng dẫn: PGS. TS. QUẢN THÀNH THƠ

CÁN BỘ HƯỚNG DẪN

KHOA QUẢN LÝ CHUYÊN NGÀNH

PGS. TS. QUẢN THÀNH THƠ

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả nêu trong Luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tôi xin cam đoan rằng mọi sự giúp đỡ cho việc thực hiện Luận văn này đã được cảm ơn và các thông tin trích dẫn trong Luận văn đã được chỉ rõ nguồn gốc.

Học viên thực hiện Luận văn

Nguyễn Đức Khoan

LỜI CẢM ƠN

Luận văn được thực hiện tại Khoa Công nghệ thông tin - Trường Đại học Công nghệ TP HCM, dưới sự hướng dẫn khoa học của PGS. TS. Quản Thành Thơ.

Trước tiên tôi xin bày tỏ lòng biết ơn sâu sắc tới PGS. TS. Quản Thành Thơ đã đưa tôi đến với lĩnh vực nghiên cứu này. Thầy đã tận tình giảng dạy, hướng dẫn giúp tôi tiếp cận và đạt được thành công trong công việc nghiên cứu của mình. Thầy đã luôn tận tâm động viên, khuyến khích và chỉ dẫn giúp tôi hoàn thành được luận văn này.

Tôi cũng xin cảm ơn các bạn trong nhóm Airlink Đại Học Bách Khoa TP. HCM đã góp phần hỗ trợ tôi hoàn thành luận văn này.

Tôi xin bày tỏ lòng biết ơn tới các Thầy Cô thuộc Khoa Công nghệ thông tin và cán bộ phòng Quản lý khoa học và đào tạo sau đại học - Trường Đại học Công nghệ TP HCM, đã tạo mọi điều kiện thuận lợi giúp đỡ tôi trong quá trình học tập và nghiên cứu tại trường.

Tôi xin chân thành cảm ơn các Thầy Cô trong Hội đồng đánh giá luận văn Thạc Sĩ đã đóng góp ý kiến quý báu giúp tôi hoàn thiện bản luận án.

Học viên thực hiện Luận văn

Nguyễn Đức Khoan

TÓM TẮT

Trong lĩnh vực thương mại điện tử hiện nay, việc tối ưu hóa doanh thu bằng các chiến lược bán hàng khác nhau đang là vấn đề ưu tiên hàng đầu. Trong đó chiến lược thay đổi giá bán tùy vào mỗi điều kiện bán hàng khác nhau cho cùng sản phẩm, dịch vụ để thu hút khách hàng và bán được nhiều hàng hóa hơn đang là chiến lược được đa số nhà bán hàng áp dụng. Cùng với sự phát triển của internet, ngành hàng không đang phổ biến chiến lược bán hàng theo giá động thông qua các đại lý bán hàng trực tuyến. Yêu cầu đặt ra là phát triển được một hệ thống có thể xác định giá vé máy bay tự động dựa trên nhu cầu khách hàng.

Luận văn nghiên cứu áp dụng được mô hình Markov ẩn để phát triển mô hình xác định giá vé tự động. Đồng thời, xây dựng được công thức tính mức tăng giảm giá vé cho phù hợp với dữ liệu quan sát. Bởi vì mô hình Markov ẩn chỉ hỗ trợ ra quyết định tăng hoặc giảm hoặc giữ nguyên giá chứ không đề xuất mức tăng hoặc giảm giá. Luận văn cũng tìm ra được phương pháp xác định sự phân phối giá vé trong trường hợp có sự kiện diễn ra và kết hợp sự phân phối này với mô hình Markov ẩn để xác định giá vé trong trường hợp có sự kiện diễn ra cho phù hợp với dữ liệu quan sát.

Tác giả đã xây dựng được một hệ thống mô phỏng qui trình bán vé tự động, từ đó đưa ra các số liệu đánh giá mô hình đã xây dựng dựa trên doanh thu của việc bán vé.

Kết quả thí nghiệm cho thấy mô hình xác định giá vé tự động dựa trên mô hình Markov ẩn trong tất cả các trường hợp đánh giá đều đem lại doanh thu cao hơn so với chiến lược bán vé theo mô hình giá tĩnh.

ABSTRACT

Nowadays, the optimization of revenue by using the difference strategies in the area of e-commerce is a recent trend. The strategy of dynamic pricing based on many sale conditions for same product, service in order to attract customers and sell more goods are strategies being most companies apply. Along with the advance of the internet, the airline industry is a most successfully industry which is applying the by dynamic pricing strategies through the online sales agencies. The issue is currently the sales agencies are updated the airline ticket price manually based on the experienced sellers. They expect to develop a system that has ability to determine the price of airline ticket automatically based on customer demands.

The thesis has researched and applied the Hidden Markov Model to develop the auto pricing model which determines the price for airline ticket automatically. Moreover, the thesis has developed formulas to calculate the degree of increasing or decreasing ticket price according to the observation data. Because the Hidden Markov Model only supports for decision making such as increase, decrease or unchanged, but not for degree of the decisions. The thesis has studied the approach to determine the distribution of ticket prices in the case of event, and combine this distribution with the Hidden Markov Model to adjust price according to the observation data in the case of event.

Author has been developed a simulation system to simulate the auto sale process for airline ticket and output statistical data which supports researcher evaluates the developed model based on the revenue.

The experimental results show that the developed auto pricing model based on Hidden Markov Model in all cases of evaluation has generated revenues higher than the traditional static pricing model.

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
TÓM TẮT	iii
ABSTRACT	iv
DANH MỤC TỪ VIẾT TẮT.....	viii
DANH MỤC CÁC BẢNG.....	ix
DANH MỤC HÌNH VẼ, ĐỒ THỊ	x
CHƯƠNG 1 GIỚI THIỆU	1
1.1. ĐẶT VẤN ĐỀ.....	1
1.2. MỤC TIÊU CỦA ĐỀ TÀI.....	4
1.2.1. Mục tiêu của đề tài.....	4
1.2.2. Phạm vi nghiên cứu	5
1.2.3. Nội dung nghiên cứu.....	5
1.3. CẤU TRÚC CỦA LUẬN VĂN.....	6
CHƯƠNG 2 CÁC NGHIÊN CỨU LIÊN QUAN	7
2.1. CÁC NGHIÊN CỨU HIỆN CÓ.....	7
2.2. CÁC LÝ THUYẾT VỀ GIÁ ĐỘNG (DYNAMIC PRICING).....	8
2.2.1. Định nghĩa.....	8
2.2.2. Phân loại giá động.....	9

2.2.3. Các mô hình đang được dùng trong việc xác định giá động	9
CHƯƠNG 3 CƠ SỞ LÝ THUYẾT	13
3.1. CÁC YẾU TỐ ẢNH HƯỞNG ĐẾN GIÁ VÉ MÁY BAY	13
3.2. MÔ HÌNH MARKOV ẨN	15
3.2.1. Giới thiệu	15
3.2.2. Định nghĩa.....	17
3.2.3. Các bước xác định trạng thái kế tiếp dựa vào trên mô hình HMM	18
3.2.4. Ví dụ minh họa mô hình HMM	19
CHƯƠNG 4 HỆ THỐNG ĐỀ NGHỊ.....	23
4.1. XÂY DỰNG MÔ HÌNH HỆ THỐNG.....	23
4.1.1. Các thông số của mô hình HMM.....	23
4.1.2. Xác định dữ liệu huấn luyện.....	23
4.2. XÁC ĐỊNH HOẠT ĐỘNG THAY ĐỔI GIÁ VÉ	25
4.3. XÁC ĐỊNH MỨC TĂNG GIẢM GIÁ VÉ.....	28
4.3.1. Trường hợp tăng giá	29
4.3.2. Trường hợp giảm giá	31
4.4. XÁC ĐỊNH GIÁ VÉ TRONG TRƯỜNG HỢP CÓ SỰ KIẾN.....	34
4.4.1. Cách tính giá vé trong trường hợp có sự kiện.....	34
4.4.2. Áp dụng mô hình HMM trong trường hợp có sự kiện	37
CHƯƠNG 5 HIỆN THỰC VÀ THỬ NGHIỆM	39
5.1. XÂY DỰNG MÔ HÌNH MÔ PHỎNG.....	39

5.1.1. Mục tiêu và phạm vi áp dụng của hệ thống mô phỏng.....	39
5.1.2. Mô phỏng actor.....	40
5.1.3. Hệ thống mô phỏng.....	42
5.1.4. Hiện thực chương trình mô phỏng.....	44
5.2. THỬ NGHIỆM VÀ ĐÁNH GIÁ MÔ HÌNH.....	47
5.2.1. Phương pháp đánh giá	47
5.2.2. Kết quả thử nghiệm.....	51
5.2.3. Đánh giá kết quả thử nghiệm.....	63
CHƯƠNG 6 KẾT LUẬN.....	66
6.1. KẾT QUẢ ĐẠT ĐƯỢC.....	66
6.2. HẠN CHẾ CỦA HỆ THỐNG.....	67
6.3. HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI	67
TÀI LIỆU THAM KHẢO.....	69

DANH MỤC TỪ VIẾT TẮT

Ký hiệu	Tiếng Anh	Tiếng Việt
HMM	Hidden Markov Model	Mô hình Markov Ẩn
Actor	Simulated customer	Khách hàng mô phỏng
DS	Daily Search	Số lượng tìm kiếm hàng ngày
CR	Conversion Rate	Tỉ lệ chuyển đổi giữa số lượng đặt vé và số lượng tìm kiếm

DANH MỤC CÁC BẢNG

Bảng 3.1. Xác suất chuyển đổi các trạng thái	20
Bảng 3.2. Xác suất quyết định công việc dựa vào thời tiết.....	20
Bảng 3.3. Bảng kết quả xác suất cho từng trạng thái	21
Bảng 4.4. Bảng dữ liệu thống kê xác suất chuyển trạng thái	24
Bảng 4.5. Bảng dữ liệu thống kê xác suất phân phối dữ liệu quan sát	24
Bảng 4.6. Bảng ví dụ dữ liệu quan sát	25
Bảng 4.7 Kết quả tính xác suất từng trạng thái	26
Bảng 4.8. Bảng dữ liệu kết quả ra quyết định	27
Bảng 5.1. Bảng xác suất chuyển trạng thái được dùng trong thử nghiệm	49
Bảng 5.2. Bảng xác suất phân phối dữ liệu quan sát được dùng trong thử nghiệm.....	49

DANH MỤC HÌNH VẼ, ĐỒ THỊ

Hình 1.1 Giá tĩnh.....	3
Hình 1.2 Giá động.....	3
Hình 1.3 Mô hình nghiên cứu	6
Hình 3.1 Mô hình các yếu tố ảnh hưởng đến giá vé	15
Hình 3.2. Mô hình Markov 3 trạng thái	16
Hình 3.3. Mô hình Markov ẩn.....	16
Hình 3.4. Ví dụ một mô hình dự đoán của HMM.....	19
Hình 3.5. Biểu diễn cách xác định trạng thái ẩn	21
Hình 4.1. Biểu đồ phân phối giá vé theo thời gian	23
Hình 4.2. Mô hình dữ liệu huấn luyện cho HMM	24
Hình 4.3. Biểu đồ biểu diễn đường phân phối Gaussian	35
Hình 4.4. Biểu đồ giá vé theo phân phối Gaussian	36
Hình 4.5. Biểu đồ so sánh giá được tính theo Gaussian và theo HMM.....	38
Hình 5.1. Giao diện chương trình mô phỏng	44
Hình 5.2. Giao diện màn hình thay đổi các xác suất học máy.....	47
Hình 5.3. Màn hình thống kê dữ liệu quá khứ	47
Hình 5.4. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	51
Hình 5.5. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động....	52
Hình 5.6. Biểu đồ so sánh doanh thu trong 2 trường hợp	52
Hình 5.7. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	53

Hình 5.8. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động....	54
Hình 5.9. Biểu đồ so sánh doanh thu trong 2 trường hợp	54
Hình 5.10. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	55
Hình 5.11. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động..	56
Hình 5.12. Biểu đồ so sánh doanh thu trong 2 trường hợp	56
Hình 5.13. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	57
Hình 5.14. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động..	58
Hình 5.15. Biểu đồ so sánh doanh thu trong 2 trường hợp	58
Hình 5.16. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	59
Hình 5.17. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động..	60
Hình 5.18. Biểu đồ so sánh doanh thu trong 2 trường hợp	60
Hình 5.19. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh.....	61
Hình 5.20. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động..	62
Hình 5.21. Biểu đồ so sánh doanh thu trong 2 trường hợp	62
Hình 5.22. Biểu đồ tổng quan so sánh các trường hợp giá tĩnh và giá động	63
Hình 5.23. Biểu đồ tổng quan so sánh các trường hợp giá tĩnh và giá động có sự kiện	64

CHƯƠNG 1 GIỚI THIỆU

1.1. ĐẶT VẤN ĐỀ

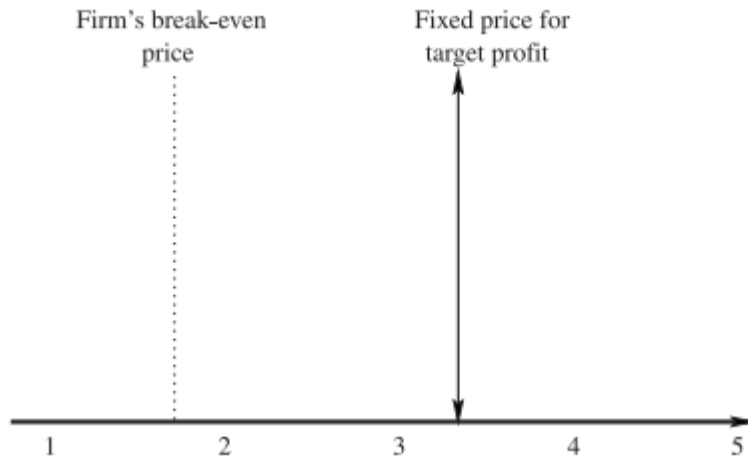
Cùng với sự tăng trưởng mạnh của internet và các dịch vụ online, các công ty kinh doanh trực tuyến hiện đang vật lộn với các tác vụ phức tạp để xác định đúng giá sản phẩm, dịch vụ đến đúng khách hàng, đúng thời điểm. Các tác vụ này yêu cầu các công ty không chỉ dựa vào chi phí hoạt động, chi phí sản xuất mà còn dựa trên sự hiểu biết về khách hàng và các nhu cầu trong tương lai. Các thuận lợi về công nghệ internet, thương mại điện tử, sự tăng trưởng nhanh chóng khối lượng thông tin thu thập được từ khách hàng đã giúp các nhà bán hàng thuận lợi hơn trong việc hiểu nhu cầu khách hàng, thị trường để thay đổi giá bán cho phù hợp, đó cũng là tiền đề cho các nghiên cứu về giá động (dynamic pricing).

Cũng do các thuận lợi về mạng internet, người mua hiện nay có thể dễ dàng so sánh được giá cả của các sản phẩm, dịch vụ và dễ dàng đưa ra các thương lượng về giá cả. Cùng lúc đó, công nghệ cho phép người bán thu thập các thông tin chi tiết về người mua như thói quen, sở thích, khả năng chi tiêu,... từ đó đưa ra mức giá phù hợp. Vì vậy, người mua và người bán có thể tương tác ngay trong thế giới điện tử, kết quả cho ra giá động phản ánh đúng nhất giá trị trên thị trường của các sản phẩm, dịch vụ.

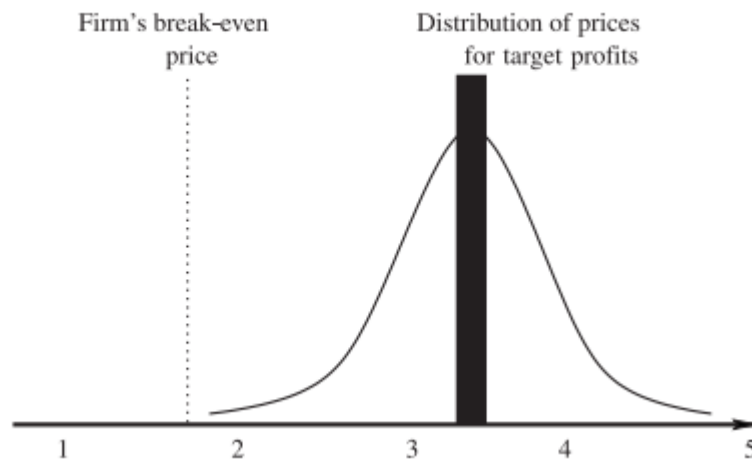
Đối với ngành hàng không, vào giữa thập niên 1970, các quy định nghiêm ngặt về giá của ngành hàng không bị bãi bỏ. Nhiều hãng hàng không đã thay đổi bằng cách cung cấp nhiều loại giá khác nhau, thay đổi triết lý hoạt động. Một số trong đó đã phát triển mạnh. Các hãng hàng không khác khác thì phá sản. Các quy định trước đây đã giới hạn tính đa dạng của các sản phẩm, dịch vụ cho người tiêu dùng. Ngành hàng không là ngành chịu ảnh hưởng nặng nề nhất. Sau gần 40 năm hoạt động với các quy định nghiêm ngặt, các hãng hàng không đã có thể tự do chọn đường bay để phục vụ và tự quyết định về giá cả. Các nghiên cứu đã chỉ ra sự thay đổi bản chất của thị trường

cạnh tranh. Ví dụ như trước đây, các hãng hàng không càng lớn, càng nhiều chuyến bay trên thị trường thì càng được nhiều người biết đến hơn, các hãng bay gần nơi người có nhu cầu di chuyển thường được ưa chuộng hơn vì người di chuyển có thể giảm được thời gian sắp xếp chuyến đi, vì giá cả đã được quy định như nhau. Tuy nhiên, sau khi bãi bỏ quy định, sự cạnh tranh về lịch trình đã giảm tầm quan trọng, các hãng hàng không bắt đầu cạnh tranh dựa trên giá cả và chiến lược cạnh tranh dựa trên việc phát triển một hệ thống mạng kết nối các đường bay. Nhờ đó, các hãng bay nhỏ có thể cạnh tranh được với những hãng lớn nhờ chi phí thấp hơn cùng các gói giảm giá. Bây giờ đã có sự cân bằng giữa cạnh tranh giá cả và cạnh tranh về lịch bay. Từ đây, xuất hiện yêu cầu thay đổi về giá (dynamic pricing) dựa trên nhu cầu khách hàng của các hãng hàng không để cạnh tranh trên thị trường [2].

Một ví dụ cụ thể trong ngành hàng không để chứng minh lợi ích của việc thay đổi giá bán để tăng lợi nhuận [15]. Như kịch bản trong Hình 1.1, hãng bán vé với giá cố định. Theo kịch bản này, chỉ có khách hàng 4 và 5 bằng lòng trả tiền mua vé với giá hãng đưa ra. Kịch bản thứ 2 như trong Hình 1.2, giá vé được phân phối động, hãng có thể bán được vé cho cả khách hàng 2, 3, 4 và 5. Như vậy lợi nhuận của hãng sẽ tăng lên trong kịch bản thứ 2. Lý do kịch bản 2 có thể bán được cho nhiều đối tượng khách hàng hơn dựa theo đánh giá khả năng và sự hài lòng của khách hàng. Khách hàng 2, 3 bằng lòng chi trả cho giá vé thấp hơn với các điều khoản ràng buộc chặt chẽ hơn như loại ghế, chính sách trả vé, thời gian đặt vé (đặt càng sớm giá càng thấp),...



Hình 1.1 Giá tĩnh [15]



Hình 1.2 Giá động [15]

Có thể nói, ngành hàng không là một ngành đặc trưng cho việc thay đổi giá vé theo nhu cầu khách hàng. Tính phức tạp của việc thay đổi giá vé trong ngành hàng không cũng nhiều hơn các ngành thương mại điện tử khác. Giá vé máy bay có thể thay đổi hàng ngày thậm chí hàng giờ trên thị trường.

Trong thời đại internet ngày càng phát triển mạnh mẽ, việc bán hàng trực tuyến đang là xu hướng của nhiều ngành nghề, trong đó cũng có ngành hàng không. Hiện nay

hình thức bán vé máy bay qua các kênh trực tuyến đang là hình thức phổ biến nhất, thông qua các đại lý vé máy bay trực tuyến. Mọi giao dịch mua bán đều được thực hiện qua phương thức bán hàng trực tuyến.

Như vậy, ngành hàng không đang phổ biến chiến lược bán hàng theo giá động, vừa thông qua các đại lý bán vé trực tuyến. Tuy nhiên, việc thay đổi giá bán đang được cập nhật một cách thủ công thủ công. Việc cập nhật thủ công này cần người có đủ kinh nghiệm bán hàng trong lĩnh vực hàng không nhưng thời gian cập nhật giá vé không được ngay tức thời và còn nhiều sai sót do chủ quan của con người. Vì vậy, yêu cầu đặt ra là nghiên cứu xây dựng một hệ thống có thể xác định giá vé máy bay một cách tự động nhằm tăng độ chính xác và phản hồi gần như ngay lập tức so với nhu cầu thị trường.

1.2. MỤC TIÊU CỦA ĐỀ TÀI

1.2.1. Mục tiêu của đề tài

1.2.1.1. Mục tiêu tổng quát

Xây dựng mô hình để biểu diễn việc xác định giá vé máy bay tự động dựa trên các kỹ thuật machine learning và data mining.

1.2.1.2. Mục tiêu cụ thể

- Xác định được các yếu tố ảnh hưởng đến giá vé và mức độ ảnh hưởng của các yếu tố đó đến giá vé.
- Phát triển các thuật toán học máy để học dữ liệu xác định giá vé dựa trên các yếu tố ảnh hưởng đến giá vé.
- Xây dựng mô hình xác định giá vé tự động dựa trên các yếu tố ảnh hưởng đến giá vé và các thuật toán đã phát triển được.

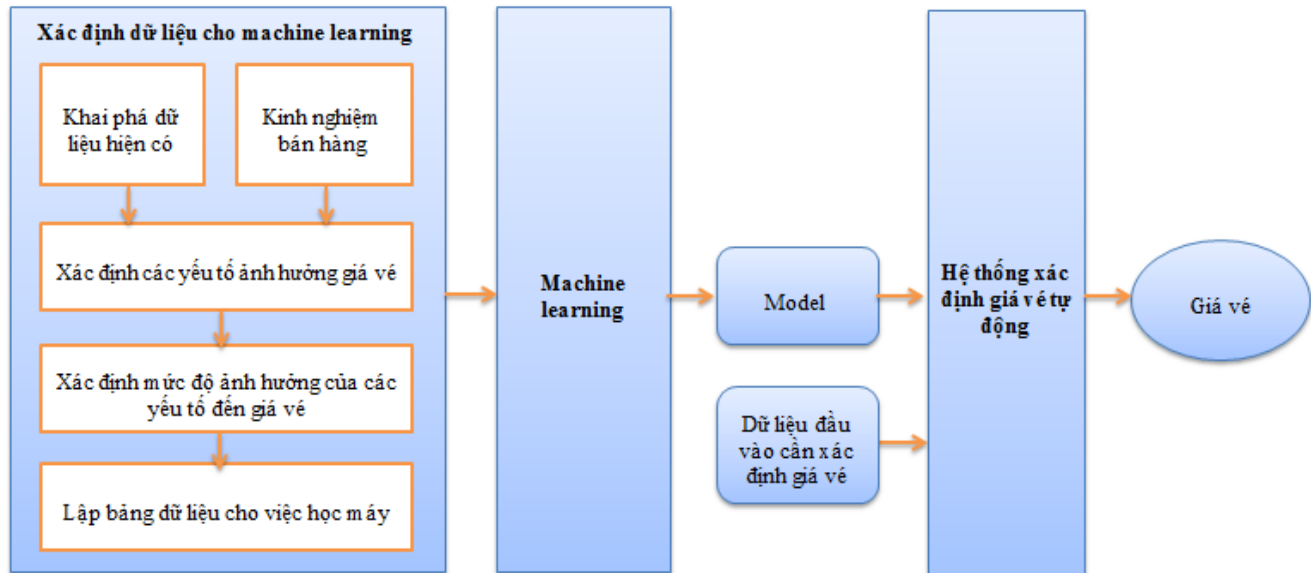
1.2.2. Phạm vi nghiên cứu

- Nghiên cứu được áp dụng cho các đại lý vé máy bay với mô hình bán vé thông qua các trang web bán vé máy bay trực tuyến.

1.2.3. Nội dung nghiên cứu

- Nghiên cứu các lý thuyết về dynamic pricing
- Nghiên cứu hệ thống bán vé máy bay của các đại lý hiện tại
 - o Nghiên cứu quy trình bán vé và dữ liệu hiện có của các đại lý vé máy bay
 - o Xác định các yếu tố ảnh hưởng đến giá vé
 - o Xác định mức độ ảnh hưởng của các yếu tố đến giá vé dựa trên kinh nghiệm người bán hàng
- Nghiên cứu xây dựng mô hình xác định giá tự động
 - o Nghiên cứu phát triển các thuật toán tự động điều chỉnh giá vé dựa trên các kỹ thuật machine learning và data mining
 - o Xây dựng mô hình xác định giá vé tự động dựa trên các yếu tố ảnh hưởng đến giá vé và các thuật toán nêu trên
- Kiểm tra đánh giá mô hình
- Kết luận và đưa ra hướng phát triển trong tương lai

Các nội dung chính cần nghiên cứu được mô hình hóa như trong Hình 1.3:



Hình 1.3 Mô hình nghiên cứu

1.3. CẤU TRÚC CỦA LUẬN VĂN

Đầu tiên, Chương 1 giới thiệu tổng quát về đề tài, lý do chọn đề tài, mục tiêu nghiên cứu, nội dung nghiên cứu, phạm vi nghiên cứu. Chương 2 trình bày các nghiên cứu liên quan đến nội dung đề tài. Chương 3 trình bày cơ sở lý thuyết của đề tài. Chương 4 trình bày về việc xây dựng mô hình hệ thống. Chương 5 trình bày việc xây dựng ứng dụng thử nghiệm, thử nghiệm và kết quả thử nghiệm. Cuối cùng, Chương 6 trình bày kết luận của tác giả về đề tài nghiên cứu.

CHƯƠNG 2 CÁC NGHIÊN CỨU LIÊN QUAN

2.1. CÁC NGHIÊN CỨU HIỆN CÓ

Trên thế giới hiện nay đã có một số công trình nghiên cứu về vấn đề thay đổi giá bán. Đặc biệt có nhiều sáng chế tại Mỹ liên quan đến việc du lịch/đi lại (travel) và giá động (dynamic pricing) như:

- Các hệ thống hỗ trợ người dùng đưa ra các tiêu chí cho việc du lịch, di chuyển và đưa ra các đề nghị cùng với mức giá phù hợp như hệ thống của: Brent J. Garback [3], Richard S. Tagawa [4] David M. DeLorme, Keith A. Gray, T. Angus Ferguson [7], Alicia Gonzalez Acebo, Christopher David Frawley [8].
- Giá thay đổi dựa trên việc tích lũy điểm thưởng cho khách hàng quen thuộc Scott B. Heintzeman, Thomas W. Storey, Barbara Monson, Steven J. Medina, Gregory A. Malark [5].
- Phương pháp tự động quản lý các thông tin dịch vụ du lịch Wayne L. Flake, Krishna Kambhampaty, Stephen B. Molsberry, Gregory B. Clem [6].
- Hệ thống trao đổi hàng hóa hỗ trợ thương lượng giá Eugene August Fusz, Christopher Ames Kline [9],
- Hệ thống hỗ trợ bán hàng tự động Lawrence B. Lockwood [10].
- Hệ thống hỗ trợ mua bán sản phẩm dựa trên mối liên quan giữa người mua và người bán Bernard R. Baker III [1]
- Hệ thống du lịch dựa trên thẻ hành khách Richard P. Sehr [11].
- Hệ thống quản lý chào mua (CPO) cho việc nhận và xử lý CPO cho các gói của các thành phần sản phẩm hoặc dịch vụ Jay S. Walker, Daniel E. Tedesco, Andrew S. VanLuchene, James A. Jorasch, T. Scott Case [12].

- Hệ thống tự sắp xếp các mặt hàng tồn kho vào giao dịch còn trống trên một hệ thống giao dịch có kết nối với nhau, đặc biệt cho các dịch vụ vận tải, hàng không luân chuyển quan nhiều trạm trung gian trên một chuyến đi Stanley Dalnekoff, Michael L. Schneider [13].
- Hệ thống kết nối người du lịch với các đại lý du lịch và các hãng hàng không John R. Shoolery, Guy E. Pasela, Diane L. De La Torre, Kenneth M. Leung, Kathleen A. Morris [14].

Các nghiên cứu và hệ thống trên chủ yếu áp dụng cho lĩnh vực chung như vận tải, du lịch và tác giả chưa tìm thấy nghiên cứu nào cho vấn đề thay đổi giá tự động cho vé máy bay.

2.2. CÁC LÝ THUYẾT VỀ GIÁ ĐỘNG (DYNAMIC PRICING)

2.2.1. Định nghĩa

Giá động là việc thay đổi động giá bán đến người tiêu dùng dựa trên giá trị mà các khách hàng này định cho một sản phẩm hoặc dịch vụ [15]. Nhiều khái niệm khác cũng được sử dụng để mô tả giá động. Giá động bao gồm 2 khía cạnh: (1) sự phân tán giá (price dispersion) và (2) sự phân biệt giá (price discrimination).

Đối với sự phân tán giá, có thể về mặt không gian hoặc thời gian. Ví dụ: về mặt không gian giá có thể khác nhau cho cùng một sản phẩm ở các quốc gia hoặc khu vực khác nhau. Về mặt thời gian, giá có thể thay đổi theo thời gian cho cùng một sản phẩm.

Đối với sự phân biệt giá, có thể có sự khác biệt về giá đối với người tiêu dùng khác nhau cho cùng một sản phẩm, được phân làm 3 mức độ:

- Mức độ 1: nhà sản xuất đưa ra mức giá khác nhau cho các đơn vị sản phẩm khác nhau, mức giá này lại có thể khác nhau từ người này sang người khác, dựa trên khả năng chi trả tối đa mà người mua vui lòng chi trả. Trong mức

độ này, nhà sản xuất phải có đủ thông tin để xác định mức tối đa mà người mua vui lòng chi trả.

- Mức độ 2: còn được gọi là giá phi tuyến (nonlinear pricing), nhà sản xuất bán theo số lượng đơn vị sản phẩm. Cùng số lượng đơn vị sản phẩm sẽ cho cùng mức giá với mọi người tiêu dùng.
- Mức độ 3: nhà sản xuất bán sản phẩm đến đối tượng người dùng khác nhau có thể cho mức giá khác nhau, nhưng mỗi đơn vị sản phẩm bán cho một đối tượng người dùng cụ thể thì có giá giống nhau. Ví dụ: giá khác nhau cho đối tượng sinh viên, công nhân, người thu nhập cao,... (nhưng trong cùng một đối tượng thì giá giống nhau).

2.2.2. Phân loại giá động

Elmaghraby and Keskinocak phân loại giá động (dynamic pricing) thành 2 loại như sau [15]:

- Cơ chế giá định trước (posted price mechanisms): một sản phẩm hoặc dịch vụ được bán theo giá được xác định trước bởi người bán (không có sự thương lượng về sau).
- Cơ chế giá định sau (price discovery mechanisms): giá được xác định thông qua một qui trình thương thảo (ví dụ như trong các buổi đấu giá).

2.2.3. Các mô hình đang được dùng trong việc xác định giá động

2.2.3.1. Các mô hình dựa trên hàng tồn kho

Các tác giả đã tham gia nghiên cứu với mô hình này gồm có Varian and Salop and Stiglitz. Elmaghraby and Keskinocak, Swann, and Chan, Shen, Simchi-Levi, và Swann [15].

Với mô hình này thì giá được quyết định chủ yếu dựa vào mức tồn kho và mức độ của các dịch vụ khách hàng. Có 3 đặc điểm của thị trường có thể ảnh hưởng đến giá động trong mô hình này như sau:

- Replenishment vs no-replenishment of inventory (R/NR): có thể hay không thể bổ sung hàng vào kho.
- Dependent vs independent demand over time: phụ thuộc hoặc không phụ thuộc vào nhu cầu theo thời gian. Ví dụ: nhu cầu đối với một số sản phẩm như thuốc chữa bệnh hay thực phẩm cốt yếu thì ít khi thay đổi.
- Myopic vs strategic customers (M/S): khách hàng quyết định mua sản phẩm bằng cách có tính toán chiến lược, tầm nhìn hay chỉ dựa vào cảm tính khi thấy giá sản phẩm thấp hơn giá mong muốn.

2.2.3.2. Các mô hình hướng dữ liệu

Mô hình này sử dụng các phương pháp thống kê hoặc các kỹ thuật tương tự để khai dữ liệu có sẵn về sở thích và các mẫu giao dịch để tính toán giá tối ưu.

Các công ty sản xuất và dịch vụ lớn thường tích lũy dữ liệu lớn về khách hàng và họ có thể tận dụng để cải thiện doanh thu và lợi nhuận bằng cách phát triển các giải thuật khai thác dữ liệu. Ngành hàng không và ngành công nghiệp khách sạn được đề cử như là các ví dụ về thành công trong việc áp dụng mô hình này. Một ví dụ cụ thể: Morris và các đồng nghiệp [15] đã xem xét các chiến lược về giá động của ngành công nghiệp hàng không bằng cách phát hiện mẫu trong sở thích của khách hàng. Hai chiến lược thay đổi giá được đưa ra để đánh giá dựa trên hệ thống mô phỏng thị trường Arena. Thứ nhất là việc thay đổi giá dựa trên số lượng ghế đã bán và thay đổi số lượng ghế bán ra tùy theo nhu cầu khách hàng. Kết quả cho thấy việc thay đổi giá bán dựa trên số lượng ghế đã bán ra có thể theo dõi được nhu cầu khách hàng và tăng doanh thu tốt hơn việc thay đổi số vé bán theo nhu cầu khách hàng.

2.2.3.3. Các mô hình dựa trên đấu giá

Đấu giá tạo thành một mô hình tự nhiên để định giá động. Kết quả của một cuộc đấu giá được xác định bởi đặc điểm cung cầu và do đó giá được xác định bởi một cuộc đấu giá thực sự dựa trên điều kiện thị trường, cung cấp cho các nhà thầu biết giá trị thực sự của chúng.

2.2.3.4. Các mô hình thuyết trò chơi

Trong một kịch bản có nhiều người bán có thể cạnh tranh trong cùng một vùng khách hàng và điều này có tạo ra một trò chơi giá động giữa những người bán. Mô hình lý thuyết trò chơi dẫn đến những cách thú vị của việc tính giá tối ưu trong một số tình huống. Ví dụ: Bernstein and Federgruen đưa ra mô hình trong đó cho nhiều cửa hàng bán lẻ tham gia để bán cùng một sản phẩm đến cùng một vùng đối tượng khách hàng, với một số điều kiện cho trước như lượng tồn kho, số đơn hàng tồn, nhu cầu khách hàng,... Sau đó cho họ tự quyết định giá bán lẻ theo một hàng số nào đó dựa trên giá cơ bản để có thể bán được hàng và cạnh tranh với nhau. Sau đó, tập hợp các hàng số này từ các nhà bán lẻ vào một hệ thống dữ liệu trung tâm để thống kê để điều chỉnh và cho ra một mức giá phù hợp cả với nhà cung cấp và các cửa hàng bán lẻ [15].

2.2.3.5. Các mô hình học máy

Học máy gần đây đã nổi lên như là một công cụ mô hình hóa phổ biến để định giá động trong thương mại điện tử. Trong một thị trường điển hình, môi trường liên tục thay đổi với nhu cầu và nguồn cung cấp dao động với nhiều cách. Trong một kịch bản như vậy, thường không thể lường trước được tất cả các diễn biến có thể có của hệ thống. Lượng thông tin có sẵn cũng được hạn chế (ví dụ, một người bán không có thông tin đầy đủ về giá cả của những người bán hàng cạnh tranh). Với mô hình học máy, ta có thể đặt tất cả các dữ liệu có sẵn vào quan điểm và thay đổi chiến lược giá thích nghi tốt nhất với môi trường. Với các mô hình hướng dữ liệu, phương pháp tiếp

cận để định giá động (như đã đề cập ở trên) có thể sử dụng kỹ thuật học máy để xác định giá động.

2.2.3.6. Các mô hình mô phỏng

Mô hình mô phỏng luôn có thể dùng trong bất kỳ vấn đề ra quyết định nào. Một mô hình mô phỏng cho giá động có thể sử dụng các mô hình đã nêu trên hoặc sử dụng hệ thống mẫu hoặc bất kỳ các cách nào khác mà có thể bắt chước hoạt động của hệ thống.

Trong các mô hình trên thì *mô hình dựa trên hàng tồn kho*, *mô hình hướng dữ liệu*, *mô hình học máy*, *mô hình mô phỏng* là các mô hình có thể phù hợp để áp dụng cho việc xác định giá động trong ngành hàng không.

CHƯƠNG 3

CƠ SỞ LÝ THUYẾT

3.1. CÁC YẾU TỐ ẢNH HƯỞNG ĐẾN GIÁ VÉ MÁY BAY

Để xác định được giá vé máy bay, cần phải xác định được các yếu tố ảnh hưởng đến giá vé, để dựa trên các yếu tố đó mà đề xuất tăng giá hoặc giảm giá.

Giá vé máy bay thường bị ảnh hưởng bởi rất nhiều yếu tố như sau:

1. Địa điểm đến
2. Các mùa trong năm, mùa du lịch
3. Bay thẳng hoặc qua các trạm trung gian
4. Dịch vụ khách hàng
5. Giờ khởi hành
6. Loại ghế ngồi
7. Tình hình thời tiết, thiên tai, an ninh
8. Các sự kiện đang và sắp diễn ra
9. Nguồn đặt vé
10. Thời gian đặt vé
11. Đối thủ cạnh tranh
12. Các điều kiện ràng buộc: như không được phép hoàn vé, đổi tên người đi, đổi ngày bay; hay vé đã đặt mua là phải thanh toán ngay chứ không có thời gian giữ chỗ; không được chọn chỗ ngồi
13. Và một số yếu tố khác

Giá vé máy bay một phần được quy định bởi các hãng hàng không dựa vào một số yếu tố cơ bản như: địa điểm đến, bay thẳng hoặc qua các trạm trung gian, dịch vụ khách hàng, giờ khởi hành, mùa du lịch,... giá này được gọi là giá cơ bản (base price).

Tuy nhiên, các đại lý vé máy bay có thể tăng giảm giá vé để phù hợp với chiến lược kinh doanh của mình để bán được vé cho nhiều loại khách hàng nhằm tăng doanh thu.

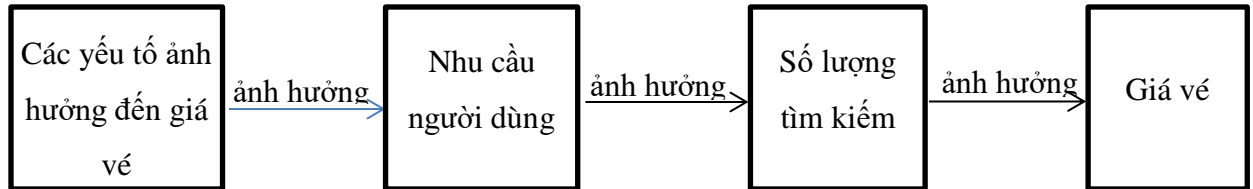
Thông thường, các hãng hàng không cũng có các dịch vụ bán vé riêng. Tuy nhiên, giá vé hãng tự bán thường phải đảm bảo cao hơn giá bán cho các đại lý.

Các đại lý vé máy bay thường dựa vào nhu cầu thay đổi của khách hàng để tăng hoặc giảm giá vé cho phù hợp. Ví dụ: khi khách hàng có nhu cầu đặt vé cao thì có thể tăng giá vé, ngược lại khi nhu cầu đặt vé thấp thì giảm giá bán. Vì vậy, giá vé có thể thay đổi từng giờ, từng ngày. Việc tăng giảm giá vé phải đảm bảo 2 yêu cầu là không giảm thấp hơn giá cơ bản và không tăng cao hơn mức giá trần (là mức giá cao nhất các đại lý quy định cho từng loại vé).

Với cùng một loại vé cho cùng một chuyến bay, cùng giờ khởi hành thì các yếu tố ảnh hưởng đến giá vé thường là các thông tin về thời tiết, an ninh, mùa cao điểm du lịch, các sự kiện đang và sắp diễn ra tại điểm đến,... Tuy nhiên, nếu dựa vào các yếu tố này để xây dựng mô hình tăng giảm giá vé thì người dùng vẫn phải theo dõi và nhập rất nhiều thông tin phức tạp vào hệ thống nên sẽ mất nhiều thời gian. Khi quan sát kỹ, tác giả nhận thấy rằng các yếu tố trên ảnh hưởng đến giá vé máy bay thông qua một yếu tố trung gian duy nhất đó là *nhu cầu người dùng*. Ví dụ như tại một địa điểm tình hình an ninh không tốt, có xung đột diễn ra thì nhu cầu di chuyển đến địa điểm ấy sẽ giảm đi. Hoặc khi có một sự kiện sắp diễn ra ví dụ như Thế vận hội thì nhu cầu di chuyển đến các địa điểm diễn ra Thế vận hội sẽ tăng lên nhanh chóng. Hoặc trong dịp lễ tết thì nhu cầu di chuyển của người dân trong nước là rất cao, kể cả những chuyến bay từ nước ngoài về nước. Qua đó, thay vì dựa vào nhiều yếu tố đã nêu trên, ta có thể tìm cách xác định nhu cầu người dùng và dựa vào yếu tố này để tăng giảm giá vé.

Quan sát quy trình bán vé của các đại lý vé máy bay trực tuyến tác giả nhận ra rằng: nhu cầu di chuyển của khách hàng thể hiện thông qua số lượng tìm kiếm của người dùng trên các trang web bán vé trực tuyến. Khi nhu cầu người dùng cho một chuyến bay tăng, số lượng tìm kiếm cho chuyến bay đó cũng tăng và ngược lại. Từ đó, ta có

thể sử dụng số lượng tìm kiếm như một dữ liệu quan sát để đưa ra quyết định tăng giảm giá vé. Hình 3.1 biểu diễn các yếu tố ảnh hưởng đến giá vé:



Hình 3.1 Mô hình các yếu tố ảnh hưởng đến giá vé

Vấn đề đặt ra tiếp theo là khi đã xác định được nhu cầu đặt vé, hệ thống sẽ quyết định tăng hay giảm giá vé bao nhiêu thì phù hợp. Giả sử việc tăng giá là quá cao và rất ít người chịu mua vé với giá đó thì không hiệu quả, còn trong trường hợp giảm quá nhiều thì có thể số lượng đặt vé tăng vọt nhưng doanh thu không đạt như mong muốn. Trong thực tế, người bán vé thường sẽ thử tăng hoặc giảm giá vé tùy theo nhu cầu người dùng, sau đó theo dõi tình hình kết quả đặt vé để điều chỉnh lại giá vé cho phù hợp. Kết quả phản hồi chính là tỉ lệ chuyển đổi (conversion rate) là tỉ lệ giữa số lượng đặt vé và số lượng tìm kiếm.

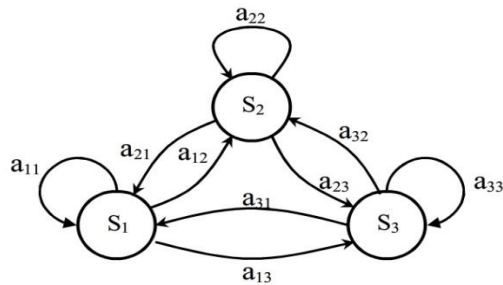
Yêu cầu đặt ra là phải xây dựng được một mô hình dựa trên dữ liệu quan sát để đưa ra quyết định thay đổi giá vé, sau đó dựa vào kết quả phản hồi của lần thay đổi giá vé kế trước trước cùng với dữ liệu quan sát tại thời điểm ra quyết định để đưa ra quyết định thay đổi giá vé tiếp theo.

3.2. MÔ HÌNH MARKOV ẨN

3.2.1. Giới thiệu

Mô hình Markov ẩn được mở rộng từ mô hình Markov [16]. Trong đó, các trạng thái của mô hình Markov không quan sát được (ẩn) nhưng có thể quan sát được các sự kiện phát sinh từ trạng thái ẩn.

Mô hình Markov ẩn phải thỏa 2 điều kiện. Đầu tiên, tại thời điểm t , các quan sát Y_t được sinh ra từ một số qui trình của trạng thái S_t mà trạng thái S_t là trạng thái ẩn. Thứ hai, các trạng thái của các qui trình ẩn thỏa mãn “thuộc tính Markov”, nghĩa là trạng thái S_t chỉ phụ thuộc vào trạng thái S_{t-1} và độc lập với tất cả các trạng thái trước thời điểm $t-1$ như ví dụ trong Hình 3.2. Mô hình Markov 3 trạng thái



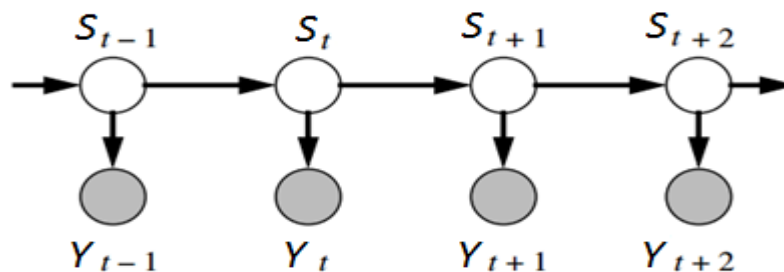
Hình 3.2. Mô hình Markov 3 trạng thái

Có thể biểu diễn các thuộc tính Markov đã mô tả ở trên bằng sự kết hợp chuỗi các trạng thái và quan sát như sau [16]:

$$P(S_{1:T}, Y_{1:T}) = P(S_1)P(Y_1|S_1) \prod_{t=2}^T P(S_t|S_{t-1})P(Y_t|S_t) \quad (3.1)$$

Trong đó, $X_{1:T}$ nghĩa là X_1, \dots, X_T

Ví dụ: Mô hình Markov ẩn



Hình 3.3. Mô hình Markov ẩn

3.2.2. Định nghĩa

Theo định nghĩa [18], mỗi mô hình Markov ẩn được định nghĩa bởi 5 thành phần: tập các trạng thái, xác suất xảy ra trạng thái, xác suất chuyển trạng thái, xác suất phân phối quan sát, và xác suất khởi điểm.

Để hiểu một cách hoàn chỉnh HMM, 5 thành phần trên được định nghĩa như sau:

1. Tập N trạng thái được định nghĩa như sau:

$$S = \{S_1, \dots, S_N\}$$

2. Tập M quan sát được định nghĩa như sau:

$$V = \{v_1, \dots, v_M\}$$

3. Bảng xác suất chuyển trạng thái $A = \{a_{ij}\}$, a_{ij} là xác suất mà trạng thái tại thời điểm $t+1$ là S_j , khi trạng thái tại thời điểm t là S_i . Trong trường hợp $a_{ij}=0$ nghĩa là trong quá trình huấn luyện không có trường hợp nào mà S_i chuyển thành S_j .

$$a_{ij} = p\{q_{t+1} = j | q_t = i\}, \quad 1 \leq i, j \leq N$$

Trong đó q_t là trạng thái hiện tại. Xác suất chuyển trạng thái phải thỏa ràng buộc:

$$a_{ij} \geq 0, \quad 1 \leq i, j \leq N, \quad \sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N$$

4. Bảng xác suất phân phối dữ liệu quan sát lên mỗi trạng thái là $B = \{b_j(k)\}$, trong đó $b_j(k)$ là xác suất phân phối của quan sát k lên trạng thái S_j

$$b_j(k) = p\{o_t = v_k | q_t = j\}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M$$

Trong đó, v_k là quan sát thứ k , o_t là vector tham số hiện tại và phải thỏa ràng buộc sau:

$$b_j(k) \geq 0, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M, \quad \sum_{k=1}^M b_j(k) = 1, \quad 1 \leq j \leq N$$

5. Bảng xác suất ban đầu $\pi = \{\pi_i\}$, trong đó π_i là xác suất của trạng thái S_i tại thời điểm ban đầu $t=0$:

$$\pi_i = p\{q_1 = i\} \text{ và } 1 \leq i \leq N$$

Từ các định nghĩa trên, mô hình HMM thường được biểu diễn như sau:

$$\lambda = (A, B, \pi)$$

3.2.3. Các bước xác định trạng thái kế tiếp dựa vào trên mô hình HMM

Yêu cầu: Xác định trạng thái S_t tại thời điểm t

Bước 1: Xác định các thông số của mô hình HMM gồm: tập trạng thái, tập dữ liệu quan sát, xác suất chuyển trạng thái, xác suất phân phối quan sát.

Bước 2: Xác định xác suất chuyển trạng thái a_{ij} từ trạng thái S_{t-1} đến tất cả các trạng thái $\{S_i\}$ trong mô hình Markov ẩn.

Bước 3: Xác định xác suất phân phối của dữ liệu quan sát $b_i(k)$ đối với tất cả các trạng thái $\{S_i\}$ trong mô hình Markov ẩn.

Bước 4: Tính xác suất cho từng trạng thái

Giả sử ta có chuỗi n trạng thái $\{S_i\}$, tại thời điểm t , ta có:

$$P(S_i) = a_{ij} * b_i(k) \quad (3.7)$$

Trong đó:

$P(S_i)$: xác suất xảy ra trạng thái S_i

a_{ij} : xác suất chuyển từ trạng thái S_j sang trạng thái S_i , với $j = i-1$

$b_i(k)$: xác suất phân phối quan sát K lên trạng thái thứ S_i

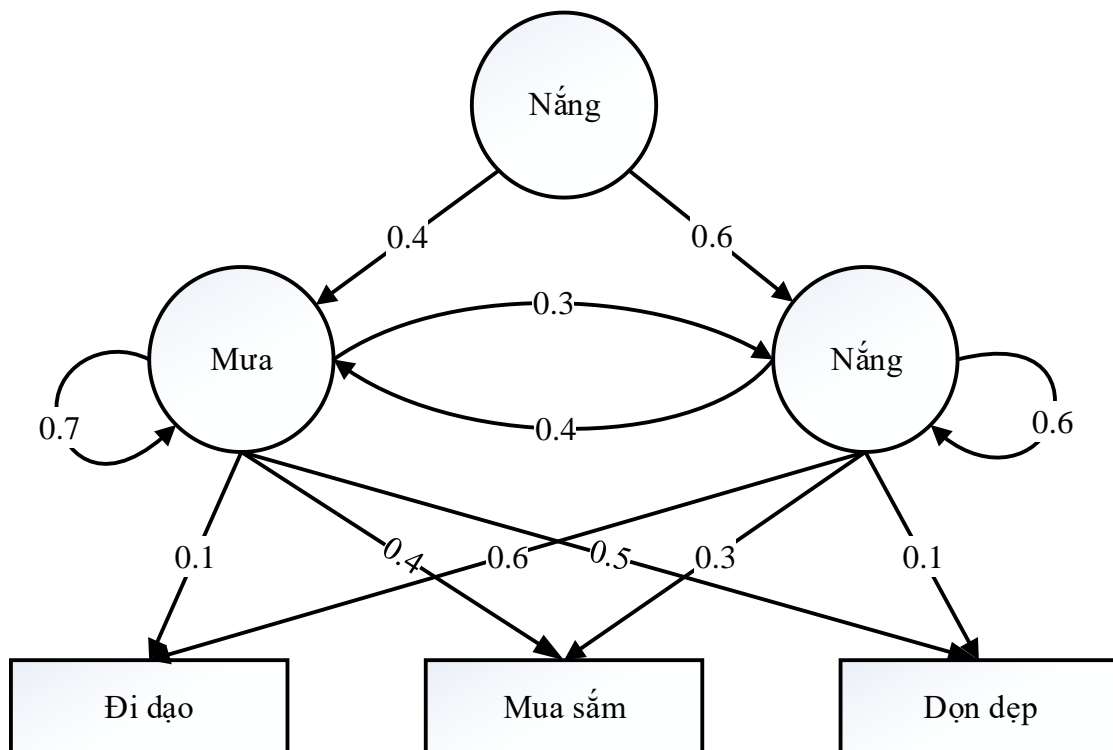
Bước 5: Chọn trạng thái có kết quả tính xác suất cao nhất làm trạng thái quyết định

3.2.4. Ví dụ minh họa mô hình HMM

Bài toán: Mai có một người bạn tên Thành sống ở rất xa. Hàng ngày họ gọi điện thoại cho nhau và Thành kể cho Mai nghe anh ta đã làm gì trong ngày. Thành chỉ có 3 công việc mà anh thích làm hàng ngày là 1) đi dạo, 2) mua sắm và 3) dọn dẹp. Sự lựa chọn phải làm gì của Thành thì phụ thuộc trực tiếp vào thời tiết hôm đấy thế nào. Như vậy, Mai không nhận được thông tin cụ thể về thời tiết nơi Thành đang sống nhưng Mai lại biết về *xu hướng chung* của thời tiết tại khu vực đó. Dựa vào lời kể của *công việc hàng ngày* của Thành, Mai có thể đưa ra dự đoán về thời tiết tại khu vực thành sống hôm đó. Mai biết hôm qua tại khu vực Thành sống thời tiết ở trạng thái **Nắng** và hôm nay Thành thích **Mua sắm**.

Hướng giải quyết bằng mô hình HMM

Mô hình dự đoán trong ví dụ này được mô tả trong Hình 3.4. Ví dụ một mô hình dự đoán của HMM:



Hình 3.4. Ví dụ một mô hình dự đoán của HMM

Bước 1: Xác định các thông số cho mô hình HMM

- Các trạng thái trong mô hình Markov gồm có: Mưa, Nắng
- Trạng thái khởi tạo: Nắng (hôm qua trời nắng)
- Các xác suất chuyển tiếp giữa các trạng thái

Trong ví dụ này, xu hướng chuyển đổi trạng thái thời tiết tại khu vực của Thành mà Mai biết như trong Bảng 3.1:

Bảng 3.1. Xác suất chuyển đổi các trạng thái

	Mưa	Nắng
Mưa	0.7	0.3
Nắng	0.4	0.6

Có nghĩa là nếu hôm nay trời mưa thì 30% ngày mai trời sẽ nắng, nếu hôm nay trời nắng thì 40% ngày mai trời sẽ mưa.

- Các xác phân phối của dữ liệu quan sát

Xác suất Thành quyết định làm một công việc nào đó phụ thuộc vào thời tiết được phân phối như trong Bảng 3.2:

Bảng 3.2. Xác suất quyết định công việc dựa vào thời tiết

	Đi dạo	Mua sắm	Dọn dẹp
Mưa	0.1	0.4	0.5
Nắng	0.6	0.3	0.1

- Các dữ liệu quan sát

Thông tin công việc Thành thích làm trong ngày (Mua sắm)

Bước 2: Xác định xác suất chuyển đổi từ trạng thái Nắng sang các trạng thái Mưa và Nắng

Dựa vào Bảng 3.1, xác suất chuyển đổi từ trạng thái nắng sang mưa hoặc vẫn nắng như sau:

	Mưa	Nắng
Nắng	0.4	0.6

Bước 3: Xác định xác suất phân phối của việc Mua sắm đối với 2 trạng thái Mưa và Nắng

Dựa vào Bảng 3.2, xác suất Thành quyết định đi mua sắm dựa trên thời tiết mưa hay nắng như sau:

	Mưa	Nắng
Mua sắm	0.4	0.3

Bước 4: Tính xác suất cho từng trạng thái

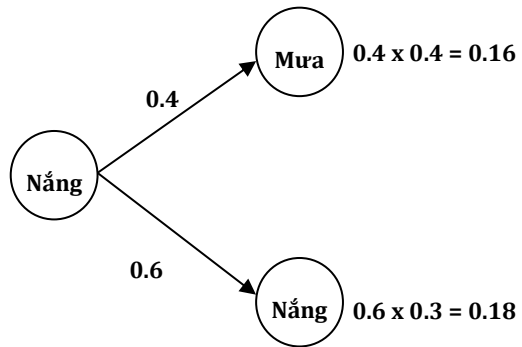
Dựa vào các tham số trên, ta xác định xác suất (P) cho từng trạng thái dựa trên công thức:

$P = \text{Xác suất chuyển trạng thái (bước 2)} * \text{xác suất quyết định công việc (bước 3)}$

Cách tính được biểu diễn dạng bảng và sơ đồ như Bảng 3.3 và Hình 3.5:

Bảng 3.3. Bảng kết quả xác suất cho từng trạng thái

	Mưa	Nắng
Nắng	0.4	0.6
Mua sắm	0.4	0.3
P	0.16	0.18



Hình 3.5. Biểu diễn cách xác định trạng thái ẩn

Bước 5: Kết luận

Theo cách tính trên, khả năng trời mưa là 0.16, khả năng trời nắng là 0.18, Mai có thể dự đoán thời tiết hôm nay tại khu vực Thành sồng là trời **nắng**.

CHƯƠNG 4 HỆ THỐNG ĐỀ NGHỊ

4.1. XÂY DỰNG MÔ HÌNH HỆ THỐNG

Xây dựng mô hình giá vé tự động dựa trên mô hình Markov ẩn (HMM).

4.1.1. Các thông số của mô hình HMM

6. Trạng thái ẩn của mô hình như sau:

- Giá tăng (Price+)
- Giá giảm (Price-)
- Giá không đổi (Price=)

7. Loại dữ liệu cần quan sát gồm có:

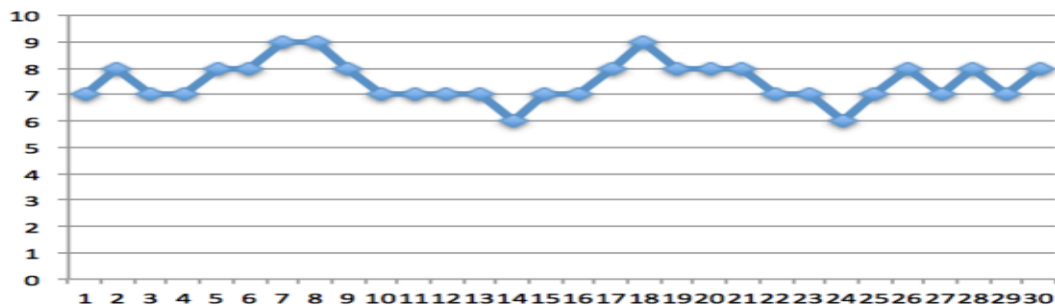
- Số lượng tìm kiếm hàng ngày (daily-search: DS)
- Tỷ lệ chuyển đổi trung bình (conversion rate: CR)

Từ 2 loại dữ liệu trên, ta có 4 trường hợp quan sát như sau:

- DS tăng, CR tăng: DS+CR+
- DS tăng, CR giảm: DS+CR-
- DS giảm, CR tăng: DS-CR+
- DS giảm, CR giảm: DS-CR-

4.1.2. Xác định dữ liệu huấn luyện

Giả sử ta có biểu đồ dữ liệu phân phối giá vé theo thời gian như Hình 4.1:



Hình 4.1. Biểu đồ phân phối giá vé theo thời gian

Dùng phương pháp thống kê từ dữ liệu phân phối giá vé trong quá khứ, giả sử bảng dữ liệu xác suất chuyển trạng thái thu được như Bảng 4.4 (Phần A):

Bảng 4.4. Bảng dữ liệu thống kê xác suất chuyển trạng thái

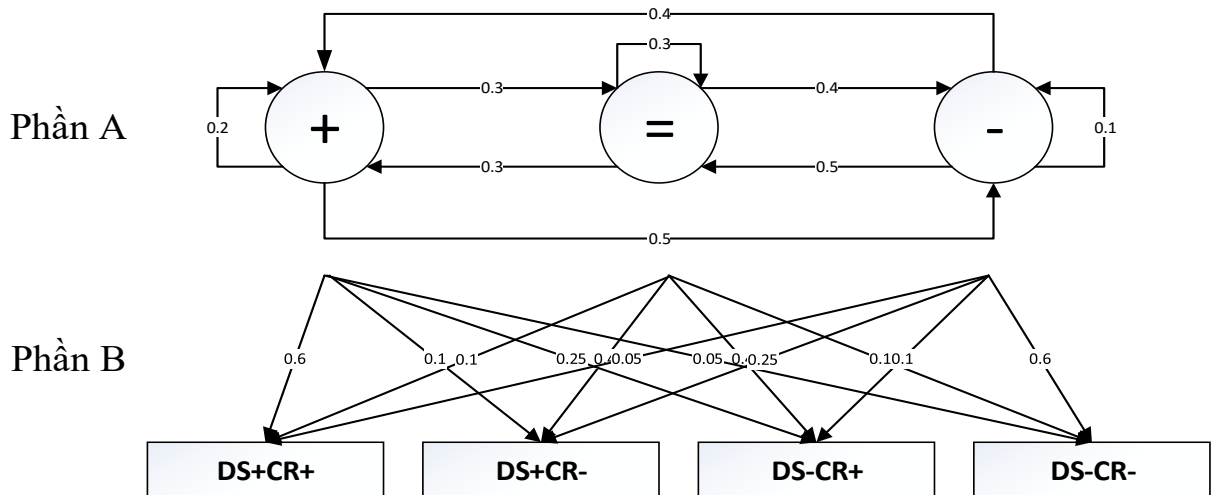
	Price+	Price=	Price-
Price+	0.2	0.3	0.5
Price=	0.3	0.3	0.4
Price-	0.4	0.5	0.1

Dùng phương pháp thống kê từ dữ liệu đặt vé trong quá khứ, giả sử bảng dữ liệu xác suất phân phối của quan sát thu được như Bảng 4.5 (Phần B):

Bảng 4.5. Bảng dữ liệu thống kê xác suất phân phối dữ liệu quan sát

	DS+CR+	DS+CR-	DS-CR+	DS-CR-
Price+	0.6	0.1	0.25	0.05
Price=	0.1	0.4	0.4	0.1
Price-	0.05	0.25	0.1	0.6

Mô hình hóa các bảng dữ liệu xác suất trên như Hình 4.2:



Hình 4.2. Mô hình dữ liệu huấn luyện cho HMM

Hai bảng dữ liệu trên được sử dụng như dữ liệu huấn luyện cho mô hình xác định giá vé tự động. Dữ liệu này có thể được cập nhật theo thời gian để cho ra được quyết định chính xác nhất theo mong muốn người bán hàng.

4.2. XÁC ĐỊNH HOẠT ĐỘNG THAY ĐỔI GIÁ VÉ

Phương pháp tổng quát để xác định trạng thái quyết định tại thời điểm thứ n :

- Xác định xác suất chuyển trạng thái từ trạng thái tại thời điểm thứ $(n-1)$ đến tất cả các trạng thái có thể có tại thời điểm thứ n (gồm 3 trạng thái Price+, Price=, Price-)
- Ứng với trường hợp quan sát cụ thể tại thời điểm n , tìm xác suất phân phối của trường hợp quan sát này tới tất cả các trạng thái của mô hình.
- Kết hợp xác suất của 2 bước trên để xác định xác suất từng trạng thái ẩn, sau đó chọn trạng thái có xác suất cao nhất làm trạng thái quyết định.

Ví dụ cho trường hợp cụ thể:

Giả sử ta có các xác suất chuyển trạng thái như Bảng 4.4, xác suất phân phối dữ liệu quan sát như Bảng 4.5 và bảng dữ liệu quan sát như Bảng 4.6:

Bảng 4.6. Bảng ví dụ dữ liệu quan sát

Date	DS	CR	Decision
1/12	216	0.57	Price=
2/12	215	0.53	
3/12	215	0.52	
4/12	212	0.49	
5/12	206	0.5	
6/12	209	0.44	
7/12	207	0.36	
8/12	198	0.42	
9/12	198	0.39	
10/12	207	0.46	

11/12	197	0.49	
12/12	188	0.42	
13/12	190	0.4	
14/12	193	0.43	
15/12	197	0.48	
16/12	193	0.48	
17/12	199	0.44	

Để xác định quyết định tăng hay giảm giá tại thời điểm ngày 2/12, ta làm như sau:

- Trạng thái kế trước là [Price=], xác định xác suất chuyển trạng thái từ [Price=] sang các trạng thái [Price+], [Price=], [Price-] dựa vào Bảng 4.4, ta được bảng dữ liệu sau:

	Price+	Price=	Price-
Price=	0.3	0.3	0.4

- Nhìn vào dữ liệu quan sát ta thấy số lượng tìm kiếm giảm, tỉ lệ chuyển đổi giảm. Suy ra dữ liệu quan sát được là: DS-CR-.
Xác định xác suất phân phối của DS-CR- đến tất cả các trạng thái dựa vào Bảng 4.5, ta được bảng dữ liệu sau:

	Price+	Price=	Price-
DS-CR-	0.05	0.1	0.6

- Nhân xác suất chuyển trạng thái với xác suất phân phối dữ liệu quan sát tương ứng với trường hợp DS-DR-, ta được như Bảng 4.7:

Bảng 4.7 Kết quả tính xác suất từng trạng thái

	Price+	Price=	Price-
Price=	0.3	0.3	0.4
DS-CR-	0.05	0.1	0.6
P	$0.05 \times 0.3 = 0.015$	$0.1 \times 0.3 = 0.03$	$0.6 \times 0.4 = 0.24$

- Dựa vào kết quả tính toán trên, ta xác định trạng thái tại ngày 2/12 là [Price-], vì xác suất xảy ra trạng thái [Price-] là lớn nhất.

Thực hiện cách tính tương tự cho các ngày còn lại, ta được kết quả như Bảng 4.8:

Bảng 4.8. Bảng dữ liệu kết quả ra quyết định

Date	DS	CR	Obs	Price+	Price=	Price-	Decision
1/12	216	0.57					Price=
2/12	215	0.53	DS-CR-	0.015	0.03	0.24	Price-
3/12	215	0.52	DS+CR-	0.04	0.2	0.025	Price=
4/12	212	0.49	DS-CR-	0.015	0.03	0.24	Price-
5/12	206	0.5	DS-CR+	0.1	0.2	0.01	Price=
6/12	209	0.44	DS+CR-	0.03	0.12	0.1	Price=
7/12	207	0.36	DS-CR-	0.015	0.03	0.24	Price-
8/12	198	0.42	DS-CR+	0.1	0.2	0.01	Price=
9/12	198	0.39	DS+CR-	0.03	0.12	0.1	Price=
10/12	207	0.46	DS+CR+	0.18	0.03	0.02	Price+
11/12	197	0.49	DS-CR+	0.05	0.12	0.05	Price=
12/12	188	0.42	DS-CR-	0.015	0.03	0.24	Price-
13/12	190	0.4	DS+CR-	0.04	0.2	0.025	Price=
14/12	193	0.43	DS+CR+	0.18	0.03	0.02	Price+
15/12	197	0.48	DS+CR+	0.12	0.03	0.025	Price+
16/12	193	0.48	DS-CR+	0.05	0.12	0.05	Price=
17/12	199	0.44	DS+CR-	0.03	0.12	0.1	Price=

Đến bước này, ta đã xây dựng được mô hình đưa ra được quyết định tăng hay giảm giá dựa trên dữ liệu quan sát. Bước tiếp theo cần nghiên cứu để mô hình xác định được mức tăng giảm cho từng trường hợp cụ thể là bao nhiêu. Khi đã xác định được mức tăng giảm cụ thể cho từng trường hợp, giá vé được đề xuất theo mô hình HMM sẽ có công thức như sau:

Trường hợp tăng giá:

$$\text{Giá vé} = \text{Baseline} * (1 + H_t)$$

(4.1)

Trong đó:

Baseline: giá cơ bản của chuyến bay mà đại lý có thể bán không lời cũng không lỗ

H_t : hệ số tăng giá

Trường hợp giảm giá:

$$\text{Giá vé} = \text{Baseline} * (1 - H_g)$$

(4.2)

Trong đó:

Baseline: giá cơ bản của chuyến bay mà đại lý có thể bán không lời cũng không lỗ

H_g : hệ số giảm giá

Ghi chú:

- Thông thường, trong cơ sở dữ liệu giá vé của các đại lý đã lưu trữ giá vé cơ bản cho từng thời điểm bay của từng chuyến bay, từng loại vé cụ thể. Vì thế baseline là con số có thể xác định được dựa vào cơ sở dữ liệu giá vé của đại lý.
- Handling fee là mức phí được quy định riêng của các đại lý để đảm bảo giá vé bán ra có thể đem lại lợi nhuận cho đại lý.

4.3. XÁC ĐỊNH MỨC TĂNG GIẢM GIÁ VÉ

Khi tăng hoặc giảm giá cần dựa trên dữ liệu quan sát để đưa ra mức tăng giảm giá cho phù hợp. Ví dụ trong trường hợp quan sát thấy số lượng tìm kiếm của khách hàng càng nhiều chứng tỏ nhu cầu tăng cao thì mức tăng càng cao và mức giảm càng nhỏ. Hoặc khi số lượng đặt vé càng gần số lượng tìm kiếm (chứng tỏ nhu cầu càng cao hoặc mức giá hiện tại còn khá thấp so với mức mà khách hàng hài lòng để bỏ tiền ra mua vé) thì mức tăng càng cao, mức giảm càng nhỏ và ngược lại.

4.3.1. Trường hợp tăng giá

4.3.1.1. Các yêu cầu cho hệ số tăng giá

- Yêu cầu T1: Nếu cùng số lượng tìm kiếm hoặc cùng số lượng đặt vé thì khi tỉ lệ chuyển đổi (Conversion rate) càng lớn thì hệ số tăng càng lớn; ngược lại, tỉ lệ chuyển đổi càng nhỏ thì hệ số tăng càng nhỏ. *Lý do là nếu tình hình đặt vé càng tốt thì mức tăng càng cao.*
- Yêu cầu T2: Nếu cùng tỉ lệ chuyển đổi thì khi số lượng tìm kiếm càng lớn thì hệ số tăng càng lớn, ngược lại, số lượng tìm kiếm càng nhỏ thì hệ số tăng càng nhỏ. *Lý do là nếu số lượng tìm kiếm tăng thì thể hiện nhu cầu khách hàng tăng nên mức tăng càng cao.*

4.3.1.2. Xây dựng công thức tính hệ số tăng giá

Công thức tính hệ số tăng giá đề nghị ban đầu:

$$H_t = \log_{S+1} \left(\frac{S+1}{|S-B|+1} \right) \quad (4.3)$$

Trong đó:

S là số lượng tìm kiếm

B là số lượng đặt vé ($B \leq S$)

H_t là hệ số tăng giá

Giải thích công thức:

- Nếu cùng số lượng tìm kiếm, khi khoảng cách giữa số lượng tìm kiếm và số lượng book $|S-B|$ càng nhỏ (tỉ lệ chuyển đổi càng nhỏ) thì $\left(\frac{S}{|S-B|}\right)$ càng lớn (*thỏa yêu cầu T1*)
- Mẫu số có cộng thêm 1 đơn vị để trường hợp $B = S$ thì mẫu số sẽ khác 0. Vì lý do đó, thay vì dùng S ta sẽ dùng $S+1$.

- Để hệ số tăng có miền giá trị trong khoảng $[0, 1]$, ta dùng hàm logarit với cơ số là số lượng tìm kiếm là $(S+1)$.

Ví dụ:

Trường hợp $S=1500, B=500$, ta có:

$$H_t = \log_{1501}(1501/(|1500-500|+1)) = 0.05$$

Nếu S giữ nguyên ($S = 1500$), B tăng lên ($B=1000$), ta có:

$$H_t = \log_{1501}(1501/(|1500-1000|+1)) = 0.15 \text{ (thỏa yêu cầu T1)}$$

\Rightarrow khoảng cách giữa S và B càng nhỏ thì hệ số tăng càng nhỏ, thỏa yêu cầu T1

Nếu số S giảm ($S=1000$), B giữ nguyên ($B=500$), ta có:

$$H_t = \log_{1001}(1001/(|1000-500|+1)) = 0.1$$

\Rightarrow khoảng cách giữa S và B càng nhỏ thì hệ số tăng càng lớn, thỏa yêu cầu T1

- Nếu $S = B$ thì mức tăng sẽ cao nhất, khi đó

$$H_t = \log_{S+1}(S+1) = 1$$

- Nếu $B = 0$ thì mức tăng sẽ thấp nhất, khi đó

$$H_t = \log_{S+1}((S+1)/(S+1)) = 0$$

- Tuy nhiên, công thức này có hạn chế trong trường hợp số lượng tìm kiếm khác nhau, nhưng tỉ lệ chuyển đổi (CR) tương đương thì mức tăng không còn thỏa yêu cầu T2

o Ví dụ: trường hợp $S=10, B=5$, ta có:

- $H_t = \log_{11}(11/(10-5+1)) = 0.25$

- Mức tăng này lớn hơn trường hợp $S=1000, B=500$ ($H_t = 0.1$, mặc dù CR đều = 0.5) thì không đạt yêu cầu T2

- Vì lý do trên, đề xuất thêm một hệ số giúp hệ số tăng tỉ lệ thuận với số lượng tìm kiếm, công thức đề xuất:

$$H_t = \log(S+1) * \log_{S+1}\left(\frac{S+1}{|S-B|+1}\right) \quad (4.4)$$

Ví dụ:

Trường hợp $S=1000$, $B=500$ ($CR = 0.5$), ta có:

$$H_t = \log 1001 * \log_{1001}(1001/(1000-500+1)) = 3 * 0.1 = 0.3$$

Trường hợp $S=10$, $B=5$ ($CR=0.5$), ta có:

$$H_t = \log 11 * \log_{11}(11/(10-5+1)) = 1.04 * 0.25 = 0.26$$

\Rightarrow Tỷ lệ chuyển đổi bằng nhau trong trường hợp $S=10$, $B=5$ và trường hợp $S=1000$, $B=500$ nhưng H nhỏ hơn khi S nhỏ, thỏa yêu cầu T2

Để phù hợp thực tế, đề xuất cho thêm một hệ số điều chỉnh là hệ số mà người dùng có thể thay đổi để kiểm soát mức tăng cho phù hợp với từng trường hợp trong thực tế.

Công thức tính hệ số tăng giá được đề xuất sau cùng là:

$$H_t = A_t * \log(S+1) * \log_{S+1}\left(\frac{S+1}{|S-B|+1}\right) \quad (4.5)$$

H_t : là hệ số tăng giảm

A_t : là hệ số điều chỉnh do người dùng quyết định để kiểm soát mức tăng

S : là số lượng tìm kiếm

B : là số lượng đặt vé

4.3.2. Trường hợp giảm giá

4.3.2.1. Các yêu cầu cho hệ số giảm giá

- Yêu cầu G1: Nếu cùng số lượng tìm kiếm hoặc cùng số lượng đặt vé thì khi tỷ lệ chuyển đổi (Conversion rate) càng lớn thì hệ số giảm càng

nhỏ; ngược lại, tỉ lệ chuyển đổi càng nhỏ thì hệ số giảm càng lớn. Lý do là nếu tình hình đặt vé càng tốt thì mức giảm càng nhỏ.

- Yêu cầu G2: Nếu cùng tỉ lệ chuyển đổi thì khi số lượng tìm kiếm càng lớn thì hệ số tăng càng nhỏ; ngược lại, số lượng tìm kiếm càng nhỏ thì hệ số tăng càng lớn. Lý do là nếu nhu cầu khách hàng giảm thì mức giảm càng cao để kích thích nhu cầu.

4.3.2.2. Xây dựng công thức tính hệ số giảm giá

Công thức tính hệ số tăng giá đề nghị ban đầu:

$$H_g = \log_{S+1} \left(\frac{S+1}{B+1} \right) \quad (4.6)$$

Trong đó:

S là số lượng tìm kiếm

B là số lượng đặt vé ($B \leq S$)

H_g là hệ số tăng giảm

Giải thích công thức:

- Cùng số lượng tìm kiếm S, tỉ lệ chuyển đổi càng nhỏ (B càng nhỏ) thì biểu thức $\left(\frac{S}{B}\right)$ càng lớn, nên hệ số H càng lớn và ngược lại (thỏa yêu cầu G1)
- Nếu cùng tỉ lệ chuyển đổi thì $\left(\frac{S}{B}\right)$ không đổi, nhưng khi S tăng thì $\log_S\left(\frac{S}{B}\right)$ sẽ giảm và ngược lại (thỏa yêu cầu G2)
- Mẫu số có cộng thêm 1 đơn vị để trường hợp số đặt vé $B = 0$ thì mẫu số sẽ khác 0. Vì lý do đó, thay vì dùng S ta sẽ dùng S+1.
- Để hệ số giảm có miền giá trị trong khoảng $[0, 1]$, ta dùng hàm logarit với cơ số là số lượng tìm kiếm (S+1).

Ví dụ:

- Trường hợp ban đầu $S=200$, $B=50$, ta có:

$$H_g = \log_{201}(201/51) = 0.26$$

- Nếu S giữ nguyên (S = 200), B giảm (B=25), ta có:

$$H_g = \log_{201}(201/26) = 0.39$$

=> *Tỉ lệ chuyển đổi càng nhỏ, hệ số giảm càng lớn thỏa yêu cầu G1*

- Nếu số S giảm (S=100), B giữ nguyên (B=25), ta có:

$$H_g = \log_{101}(101/26) = 0.29$$

=> *Tỉ lệ chuyển đổi bằng nhau trong trường hợp B=200, S=50 và B=100, S=25, nhưng số lượng tìm kiếm càng nhỏ mức giảm càng lớn, thỏa yêu cầu G2*

- Nếu B = 0 thì mức tăng sẽ cao nhất, khi đó

$$H_g = \log_{S+1}(S+1) = 1$$

- Nếu B = S thì mức tăng sẽ thấp nhất, khi đó

$$H_g = \log_{S+1}((S+1)/(B+1)) = 0$$

Để phù hợp thực tế, đề xuất cho thêm một hệ số điều chỉnh là hệ số mà người dùng có thể thay đổi để kiểm soát mức giảm cho phù hợp với từng trường hợp trong thực tế.

Công thức tính hệ số giảm giá được đề xuất sau cùng là:

$$H_g = A_g * \log_{S+1}\left(\frac{S+1}{B+1}\right)$$

(4.7)

H_g: là hệ số tăng giảm

A_g: là hệ số điều chỉnh do người dùng quyết định để kiểm soát mức giảm

S: là số lượng tìm kiếm

B: là số lượng đặt vé

4.4. XÁC ĐỊNH GIÁ VÉ TRONG TRƯỜNG HỢP CÓ SỰ KIỆN

Trong trường hợp có sự kiện diễn ra, giá vé được phân phối tuân theo đường phân phối chuẩn hay còn gọi là phân phối Gaussian. Tại thời điểm cao điểm sự kiện thì giá vé là cao nhất và giá vé thấp dần vào những ngày càng xa thời điểm cao điểm sự kiện. Với các ngày trước khi diễn ra sự kiện là những ngày hành khách có nhu cầu mua vé đến tham gia sự kiện. Với các ngày sau khi diễn ra sự kiện là những ngày hành khách có nhu cầu mua vé quay về.

Thời điểm cao điểm sự kiện được định nghĩa là thời điểm mà nhu cầu di chuyển của hành khách là cao nhất.

4.4.1. Cách tính giá vé trong trường hợp có sự kiện

Giả sử giá cơ bản (baseline) đang là 2,000USD.

Người bán hàng quy định rằng tại thời điểm cao điểm, giá sẽ là gấp 2.5 lần (tức là 5,000USD).

Thời gian ảnh hưởng của luật này là 10 ngày (5 ngày trước và 5 ngày sau sự kiện).

Tức là ngày thứ 6 trước sự kiện và ngày thứ 6 sau sự kiện, giá sẽ là 2,000USD, giá ngay thời điểm sự kiện là 5,000USD.

Dựa vào các yếu tố này, ta có thể tính các thông số cho đường phân phối Gaussian tương ứng như sau:

Vì giá vé tuân theo phân phối Gaussian nên hàm mật độ giá vé với biến ngẫu nhiên x (ngày) sẽ có dạng sau:

$$f(x) = Ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (A \in \mathbb{R}) \quad (4.8)$$

Trong đó:

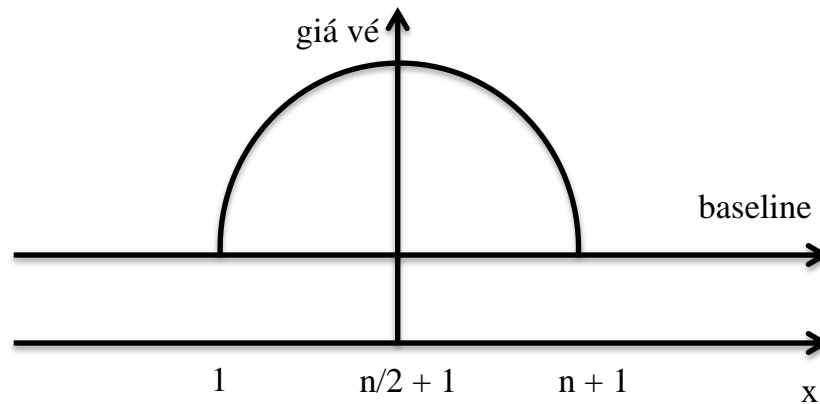
A: giá vé cao nhất tại thời điểm cao điểm của sự kiện

x : ngày (tại thời điểm bắt đầu bị ảnh hưởng bởi sự kiện $x = 1$)

μ : ngày cao điểm của sự kiện

σ : độ lệch chuẩn

Ta sử dụng hệ trục tọa độ như Hình 4.3:



Hình 4.3. Biểu đồ biểu diễn đường phân phối Gaussian

Thông số đầu vào của bài toán là:

- n là số ngày ảnh hưởng của sự kiện đến giá vé (n chẵn)
- baseline là giá cơ bản
- k là số lần gấp của giá tại thời điểm cao điểm sự kiện so với baseline

Ta có:

$$(1) f(1) = Ae^{\frac{-(1-\mu)^2}{2\sigma^2}} = \text{baseline}$$

$$(2) f(n+1) = Ae^{\frac{-(n+1-\mu)^2}{2\sigma^2}} = \text{baseline}$$

$$(3) f(n/2+1) = Ae^{\frac{-(\frac{n}{2}+1-\mu)^2}{2\sigma^2}} = \text{baseline} * k$$

Từ (1) và (2), suy ra $1 - \mu = -(n + 1 - \mu)$ hay $\mu = n/2 + 1$, thay vào (3), suy ra $A = \text{baseline} * k$

$$\text{Từ (1) ta có } \sigma = \sqrt{\frac{(1-\mu)^2}{2\ln(\frac{A}{\text{baseline}})}} \quad (4.9)$$

Áp dụng cho một trường hợp ví dụ

Với dữ liệu $n = 20$, $\text{baseline} = 2000$, $k = 2.5$ ta có:

$$A = \text{baseline} * k = 5000$$

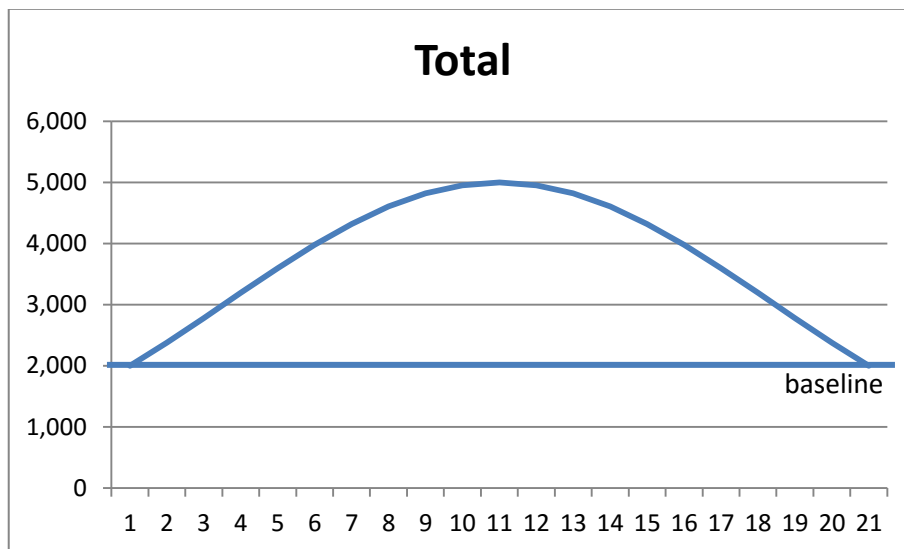
$$\mu = n / 2 + 1 = 11$$

$$\sigma = \sqrt{\frac{(1-\mu)^2}{2\ln(\frac{A}{\text{baseline}})}} = 7.387004359$$

$$\text{Giá vé của ngày thứ 2 là } f(2) = 5000e^{\frac{-(2-6)^2}{2\sigma^2}} = 2380$$

$$\text{Giá vé của ngày thứ 3 là } f(3) = 5000e^{\frac{-(3-6)^2}{2\sigma^2}} = 2781$$

Sau khi tính giá vé cho tất cả các ngày chịu ảnh hưởng bởi sự kiện với số liệu như trên, ta có biểu đồ phân phối giá vé qua các ngày như Hình 4.4:



Hình 4.4. Biểu đồ giá vé theo phân phối Gaussian

Tuy nhiên, để việc tăng giảm giá vé trong trường hợp có sự kiện được hợp lý hơn, cần thiết đưa yếu tố nhu cầu người dùng vào trong việc xem xét thay đổi giá vé. Để làm điều đó, ta lại áp dụng mô hình HMM đã xây dựng kết hợp với phân phối Gaussian cho trường hợp có sự kiện để điều chỉnh lại giá vé cho phù hợp nhu cầu người dùng.

4.4.2. Áp dụng mô hình HMM trong trường hợp có sự kiện

Để áp dụng HMM cho trường hợp có sự kiện, xin nhắc lại công thức tính giá vé được đề xuất bởi mô hình HMM như sau:

$$\text{Giá vé} = \text{Baseline} * (1 \pm H)$$

(4.10)

Trong đó:

Baseline: giá cơ bản của chuyến bay (là giá hãng bay đã cộng thêm handling fee)

H: hệ số tăng/giảm giá được đề xuất dựa trên mô hình HMM (các tính được đề xuất ở mục 4.3)

Trong trường hợp có sự kiện, tại các ngày có ảnh hưởng bởi sự kiện, baseline trong công thức trên sẽ là giá được tính dựa trên phân phối Gaussian. Vì vậy, công thức tính giá vé trong trường hợp có sự kiện kết hợp mô hình HMM là:

$$\text{Giá vé} = A e^{\frac{-(x-\mu)^2}{2\sigma^2}} (1 \pm H)$$

(4.11)

Trong đó:

A: giá vé tại thời điểm cao điểm của sự kiện

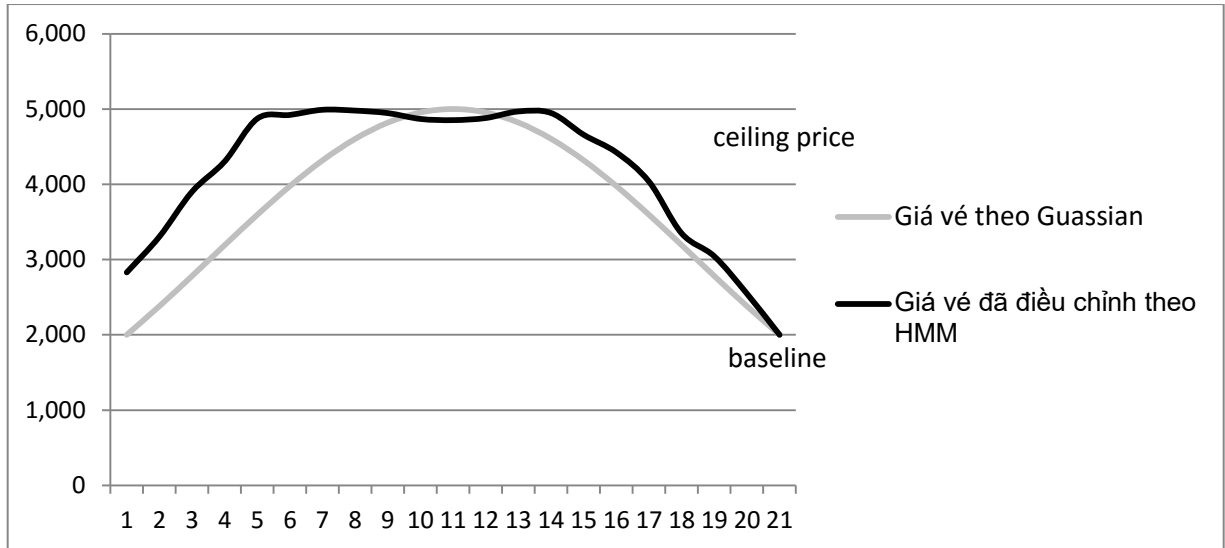
x: ngày (tại thời điểm bắt đầu bị ảnh hưởng bởi sự kiện $x = 1$)

μ : ngày cao điểm của sự kiện

σ : độ lệch chuẩn

H: là hệ số tăng/giảm giá được đề xuất dựa trên mô hình HMM.

Ví dụ: so sánh trường hợp giá được đề xuất bởi phân phối Gaussian và giá được điều chỉnh lại theo mô hình HMM được biểu diễn như biểu đồ Hình 4.5:



Hình 4.5. Biểu đồ so sánh giá được tính theo Gaussian và theo HMM

Trong ví dụ này: giá vé được đề xuất theo mô hình HMM không được nhỏ hơn baseline (2000) và không được lớn hơn ceiling price (5000).

CHƯƠNG 5

HIỆN THỰC VÀ THỬ NGHIỆM

5.1. XÂY DỰNG MÔ HÌNH MÔ PHÒNG

Để đánh giá mô hình, cần đánh giá hiệu quả của việc tăng giảm giá tự động được đề xuất từ mô hình HMM dựa trên doanh thu của việc bán vé. Vì mô hình HMM đưa ra quyết định kế tiếp dựa trên việc quan sát số lượng tìm kiếm và hiệu quả của quyết định tăng giảm giá trước đó thông qua tỉ lệ chuyển đổi, nên không thể dùng dữ liệu quá khứ để đánh giá mô hình. Vì dữ liệu quá khứ thì số lượng tìm kiếm và tỉ lệ chuyển đổi là cố định và ta không thể xác định được cụ thể nhu cầu khách hàng tại thời điểm quá khứ để xác định khi giá thay đổi, khách hàng có quyết định đặt vé không. Trong phần lý thuyết về giá động có nhắc tới mô hình mô phỏng trong việc xác định giá động. Đây cũng chính là mô hình phù hợp nhất dùng để đánh giá mô hình giá động dựa trên HMM đã được xây dựng. Yêu cầu đặt ra là cần xây dựng một chương trình mô phỏng việc bán vé trực tuyến gồm có 2 phần chính:

- Mô phỏng hệ thống bán vé với việc thay đổi giá vé tự động theo đề xuất của mô hình HMM.
- Mô phỏng người dùng có nhu cầu đặt vé máy bay (gọi tắt là actor)

5.1.1. Mục tiêu và phạm vi áp dụng của hệ thống mô phỏng

5.1.1.1. Mục tiêu

- Xây dựng công cụ dùng để mô phỏng việc khách tìm kiếm, đặt vé và việc bán vé của đại lý vé máy bay
- Cho ra kết quả đánh giá mô hình giá động áp dụng HMM

5.1.1.2. Phạm vi áp dụng

- Chỉ áp dụng cho một chuyến bay cụ thể.
- Chuyến bay được định nghĩa là một chặng bay từ sân bay đầu đến sân bay đích trong một khoảng thời gian nào đó thuộc về một hãng bay nào đó.
- Chỉ áp dụng cho chuyến bay một chiều.
- Chưa xét yếu tố hủy vé, giảm giá theo số lượng, khuyến mãi,...
- Trong trường hợp mô phỏng có sự kiện, chỉ áp dụng cho trường hợp có một sự kiện trong khoảng thời gian mô phỏng.

5.1.2. Mô phỏng actor

5.1.2.1. Mô tả actor

- Được miêu tả như những người có nhu cầu tìm kiếm và đặt vé máy bay cho cùng một chuyến bay nào đó.
- Nhu cầu đặt vé có thể cao và có thể thấp. Nếu nhu cầu cao, thì actor sẽ đặt vé sau khi tìm kiếm vé dù giá vé cao hơn mong muốn.
- Có những mẫu người có thể bằng lòng mua vé giá cao mặc dù nhu cầu thấp (người rộng rãi), cũng có mẫu người chỉ mua vé giá rẻ (người hà tiện).

5.1.2.2. Các thuộc tính của actor

- Ngày bắt đầu tìm kiếm vé.
- Ngày khởi hành.
- Số lượng vé cần đặt vé.
- Giá mong muốn.
- Trọng số nhu cầu (w): việc đặt vé này có thể là bắt buộc do nhu cầu cao ($w \geq 0.9$) hoặc có thể chỉ cần đặt một phần số lượng vé cần đặt ban đầu khi giá cao hơn giá mong muốn và nhu cầu lớn hơn mức trung bình ($0.9 > w \geq 0.5$), hoặc có thể sẽ không đặt vé nếu giá cao hơn giá mong muốn và nhu

cầu thấp hơn mức trung bình ($w < 0.5$). Trong trường hợp có sự kiện và ngày khởi hành của actor nằm trong khoảng ngày ảnh hưởng bởi sự kiện thì hệ thống sẽ tạo hệ số nhu cầu cao hơn bình thường.

- Loại người mua: người rộng rãi: sẵn sàng mua vé dù giá vé cao hơn giá mong muốn, người hà tiện: nếu nhu cầu không cao thì chỉ mua vé trong trường hợp giá vé thấp hơn hoặc bằng giá mong muốn.

5.1.2.3. Cách hoạt động của actor

- Đến thời gian cần đặt vé, actor vào hệ thống để tìm kiếm vé.
- Khi tìm được vé, xem xét giá:
 - o Nếu trong khoảng mong muốn thì đặt tất cả vé cần đặt
 - o Nếu ngoài khoảng mong muốn thì xét trọng số nhu cầu (w):
 - Nếu trọng số $w \geq 0.9$ (nhu cầu cao) hoặc là người rộng rãi thì sẽ đặt hết vé cần đặt.
 - Nếu trọng số $0.9 > w > 0.6$ (nhu cầu hơn mức trung bình) thì đặt theo tỉ lệ w .
 - Nếu còn trong khoảng thời gian đợi thì có thể đợi đến ngày/giờ tiếp theo để vào và tiếp tục tìm kiếm và xem giá (quay lại bước 2).
 - Nếu trọng số $w < 0.6$ thì không đặt,
 - Nếu còn trong khoảng thời gian đợi (chưa đến ngày khởi hành) thì có thể đợi đến ngày/giờ tiếp theo để vào và tiếp tục tìm kiếm.
- Trong trường hợp không còn đủ vé thì xem như số lượng vé mong muốn bằng số lượng vé còn lại.

5.1.3. Hệ thống mô phỏng

5.1.3.1. Mô tả hệ thống mô phỏng

- Cung cấp chức năng tìm kiếm và đặt vé cho các actor (mô phỏng 1 dịch vụ bán vé của các đại lý vé máy bay trực tuyến).
- Cho phép điều chỉnh các thông số mô phỏng (các thông số của Gaussian, HMM,...).
- Ra quyết định tăng giảm giá với mức giá cụ thể.
- Ghi nhận hàng ngày số lượng tìm kiếm, số lượng đặt vé, số lượng vé bán, giá vé, hệ số tăng giảm, doanh thu,...
- Đưa ra các số liệu thống kê sau khi mô phỏng.
- Hỗ trợ tạo tự động các actor theo các thông số mô phỏng.

5.1.3.2. Thông số đầu vào của thử nghiệm

Hệ thống mô phỏng cho phép cấu hình các thông số sau đây:

- Cho phép mô phỏng trong trường hợp giá tĩnh hoặc giá động.
- Cho phép mô phỏng trong trường hợp có sự kiện hay không.
- Mô phỏng trường hợp tăng giá theo từng ngày.
- Khoảng thời gian xem xét tăng giảm giá: ví dụ mô phỏng một ngày bằng 3 giây,...
- Số ngày cần mô phỏng.
- Giá cơ bản của chuyến bay: giá tối thiểu có thể giảm.
- Giá trần: giá cao nhất có thể tăng.
- Hệ số tăng/giảm giá ban đầu: ví dụ giá cơ bản là 1000 USD, hệ số ban đầu là 1.5 thì giá vé ngày đầu tiên của mô phỏng sẽ là 1500 USD.
- Số lượng vé có thể đặt mỗi ngày của chuyến bay mô phỏng.
- Số lần mô phỏng.

- Ngày bắt đầu sự kiện.
- Mức tăng cao nhất tại thời điểm diễn ra sự kiện.
- Số ngày ảnh hưởng bởi sự kiện: ví dụ thời gian sự kiện ảnh hưởng lên giá là 20 ngày, thì giá sẽ bắt đầu tăng theo sự kiện kể từ ngày thứ 10 trước ngày cao điểm sự kiện diễn ra và kết thúc tăng giá theo sự kiện sau 9 kể từ ngày cao điểm sự kiện diễn ra.
- Số lượng actors tham gia trong thời gian mô phỏng.
- Khoảng thời gian actor bắt đầu tìm kiếm cho đến khi khởi hành: ví dụ, ngày actor bắt đầu tìm kiếm vé cho đến ngày khởi hành là 10 ngày.
- Số lượng vé tối đa mỗi actor cần đặt.

5.1.3.3. Cách hoạt động của hệ thống

- Khởi tạo các actor dựa trên các thông số mô phỏng dành cho actor từ tập tin có sẵn hoặc hệ thống tự tạo ngẫu nhiên sau đó cho phép lưu lại thông tin các actor vào CSV để có thể tùy chỉnh và áp dụng cho các lần sau.
- Đọc các thông số cấu hình cho mô hình HMM.
- Cứ mỗi khoảng thời gian định trước, xét số lượng tìm kiếm của các actor và tỉ lệ chuyển đổi dựa trên đặt vé của các actor để đưa ra quyết định tăng/giảm giá dựa trên một số luật như:
 - o Giá được đề xuất dựa trên mô hình HMM (xem cách tính ở mục 4.3).
 - o Giá cơ bản của các ngày bị ảnh hưởng bởi sự kiện được tính theo công thức Gaussian.
 - o Giá không được cao hơn giá trần.
 - o Giá không được thấp hơn giá cơ bản.
- Ghi nhận lại các thông số thống kê như: số lượng tìm kiếm, số lượng đặt vé, số lượng vé bán, giá vé, hệ số tăng giảm, doanh thu,...
- Hết khoảng thời gian của mô phỏng thì dừng và xuất ra báo cáo cuối cùng.

- Cho phép xuất ra tập tin CSV các Kết quả thử nghiệm.

5.1.4. Hiện thực chương trình mô phỏng

- Ngôn ngữ lập trình: C#.NET
- Môi trường lập trình: Microsoft Visual Studio 2015.
- Hình 5.1 là giao diện chính của chương trình mô phỏng bán vé.

Hình 5.1. Giao diện chương trình mô phỏng

Giải thích các thông tin trên giao diện

❖ Các thông số mô phỏng:

- Price type: danh sách loại mô phỏng (giá tĩnh, giá động, giá tĩnh có sự kiện, giá động có sự kiện).
- Duration: số ngày cần mô phỏng.
- Base price: giá cơ bản đã bao gồm handling fee.

- Max price: mức giá tối đa có thể tăng (giá trần).
- Starting factor: hệ số tăng giá cho ngày đầu tiên mô phỏng. Giá vé của ngày đầu tiên sẽ được tính là $\text{Base price} * \text{Starting factor}$.
- Available ticket per day: số lượng vé có thể bán mỗi ngày.
- Pricing change interval: mô phỏng một ngày bằng bao nhiêu giây.
- No. of simulation: số lần mô phỏng.
- Number of actors: số lượng actor cần phát sinh tự động để mô phỏng.
- Maximum actor dur.: số ngày tối đa cho hoạt động tìm kiếm và đặt vé của mỗi actor. Ví dụ: Maximum actor dur = 10 và ngày actor bắt đầu tìm kiếm vé là 1/12 thì ngày khởi hành của actor sẽ trong khoảng từ ngày 1/12 đến ngày 10/12.
- Maximum actor tickets: số vé tối đa mỗi actor cần đặt.
- Control factor: là hệ số điều chỉnh A trong công thức tính hệ số tăng/giảm giá.
- Event date: ngày cao điểm của sự kiện. Thông số này cần dùng để áp dụng công thức Gaussian.
- Event effective days: số ngày mà giá vé bị ảnh hưởng bởi sự kiện. Ví dụ: sự kiện diễn ra ngày 10/12 và event effective days = 20 thì ngày bắt đầu ảnh hưởng bởi sự kiện sẽ là ngày 01/12 và kết thúc ảnh hưởng vào ngày 20/12. Thông số này cần dùng để áp dụng công thức Gaussian.
- Event max rate: hệ số tăng giá vào ngày cao điểm của sự kiện, cho biết vào ngày cao điểm của sự kiện giá vé sẽ tăng tối đa bao nhiêu lần so với giá cơ bản. Ví dụ: giá cơ bản là 100, ngày cao điểm của sự kiện là 10/12, hệ số tăng giá vào ngày cao điểm của sự kiện 5 thì giá vé vào ngày cao điểm của sự kiện sẽ là 500. Thông số này cần dùng để áp dụng công thức Gaussian.
- Start date: Ngày đầu tiên của mô phỏng.

❖ Các nút lệnh:

- Start: bắt đầu mô phỏng.
- Stop: dừng mô phỏng.
- Generate actors: phát sinh tự động thông tin các actor dựa vào các thông số mô phỏng ở trên.
- Load actors: tải thông tin các actor từ tập tin CSV.
- Save actors: lưu lại thông tin của actor vào tập tin CSV.
- Save results: lưu lại Kết quả thử nghiệm.

❖ Các lưới hiển thị thông tin

- Actor information: hiển thị các actor cùng tất cả các thuộc tính của actor
- Price change history: hiển thị lịch sử tăng giảm giá và một số thông tin khác trong quá trình mô phỏng.
- Activity log: lưu lại nhật ký hoạt động của các actor trong quá trình mô phỏng.
- Statistic: lưu lại kết quả thống kê của mỗi lần mô phỏng.

5.1.4.1. Các màn hình khác

- ❖ Màn hình cho phép thay đổi cấu hình các xác suất học máy cho mô hình HMM như Hình 5.2

The screenshot shows a window titled "Auto Pricing Simulation" with a "Simulator" tab and a "Probability config" sub-tab. It contains two tables for configuring probabilities.

States transition probability

State	Price+	Price=	Price-
Price+	0.2	0.5	0.3
Price=	0.5	0.2	0.3
Price-	0.4	0.5	0.1

States Observations distribution

State	DS+CR+	DS+CR-	DS-CR+	DS-CR-
Price+	0.6	0.2	0.15	0.05
Price=	0.2	0.4	0.3	0.1
Price-	0.05	0.2	0.15	0.6

At the bottom of the window, there are two buttons: "Save configuration" and "Statistics from log".

Hình 5.2. Giao diện màn hình thay đổi các xác suất học máy

Màn hình này cho phép thay đổi các xác suất chuyển trạng thái cũng như các xác suất phân phối dữ liệu quan sát làm dữ liệu huấn luyện cho mô hình HMM.

❖ Màn hình thống kê dữ liệu quá khứ như Hình 5.3

State	DS+CR+	DS+CR-	DS-CR+	DS-CR-
Increase	0.123	0.494	0.175	0.208
NotChange	0.333	0.167	0.333	0.167
Decrease	0.34	0.14	0.445	0.076

Hình 5.3. Màn hình thống kê dữ liệu quá khứ

Màn hình này cho phép đưa dữ liệu giá vé và dữ liệu đặt vé ở quá khứ để thống kê cho ra các xác suất chuyển trạng thái và xác suất phân phối của dữ liệu quan sát.

5.2. THỬ NGHIỆM VÀ ĐÁNH GIÁ MÔ HÌNH

5.2.1. Phương pháp đánh giá

- Đánh giá bằng phương pháp mô phỏng, sử dụng phần mềm mô phỏng được mô tả ở mục 5.
- Đánh giá dựa trên doanh thu trong 2 chiến lược:

- Bán vé theo giá cố định (giá tĩnh)
 - Trong trường hợp không có sự kiện (gọi tắt là giá tĩnh).
 - Trong trường hợp không sự kiện (gọi tắt là giá tĩnh có sự kiện).
- Bán vé theo giá động được đề xuất bởi mô hình HMM (giá động)
 - Trong trường hợp không có sự kiện (gọi tắt là giá động).
 - Trong trường hợp không sự kiện (gọi tắt là giá động có sự kiện).
- Vì doanh thu của của việc bán vé theo giá tĩnh có thể khác nhau trong các trường hợp có mức giá bán khác nhau, nên để việc đánh giá được khách quan hơn, cần đánh giá 2 chiến lược trên trong 3 trường hợp:
 - Giá khởi điểm thấp.
 - Giá khởi điểm trung bình.
 - Giá khởi điểm cao.
- Vậy có tất cả 6 trường hợp cần đánh giá:
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm thấp.
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm trung bình.
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm cao.
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm thấp.
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm trung bình.
 - Đánh giá doanh thu trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm cao.

- Đối với trường hợp mô phỏng cho giá động thì với mỗi lần mô phỏng có thể cho ra doanh thu khác nhau trong cùng một điều kiện đầu vào do có một số yếu tố ngẫu nhiên trong quá trình mô phỏng, nên trong mỗi trường hợp mô phỏng của giá động cần thực hiện mô phỏng nhiều lần và lấy trung bình doanh thu làm giá trị so sánh.
- Đối với trường hợp mô phỏng cho giá tĩnh thì mỗi trường hợp chỉ cần mô phỏng 1 lần và cùng thông tin actor được dùng cho trường hợp giá động. Vì trường hợp giá tĩnh thì kết quả thử nghiệm sẽ không có sự khác biệt nhiều khi thử nghiệm trong cùng một điều kiện đầu vào.
- Các thông số đầu vào cho mỗi thử nghiệm trong cùng một trường đánh giá là như nhau.
- Bảng 3.1 là xác suất chuyển trạng thái được sử dụng cho tất cả các thử nghiệm như sau:

Bảng 5.1. Bảng xác suất chuyển trạng thái được dùng trong thử nghiệm

	Price+	Price=	Price-
Price+	0.2	0.5	0.3
Price=	0.5	0.2	0.3
Price-	0.4	0.5	0.1

- Bảng 5.2 là xác suất phân phối giữa các trạng thái và dữ liệu qua sát được sử dụng cho tất cả các thử nghiệm như sau:

Bảng 5.2. Bảng xác suất phân phối dữ liệu quan sát được dùng trong thử nghiệm

	DS+CR+	DS+CR-	DS-CR+	DS-CR-
Price+	0.6	0.2	0.15	0.05
Price=	0.2	0.4	0.3	0.1
Price-	0.05	0.2	0.15	0.6

- **Kết quả thử nghiệm mong muốn:** trong 6 trường hợp, trung bình doanh thu của mô hình giá động được đề xuất bởi mô hình HMM đều cao hơn so với doanh thu của mô hình bán vé theo giá tĩnh.

5.2.2. Kết quả thử nghiệm

5.2.2.1. Thử nghiệm trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm thấp

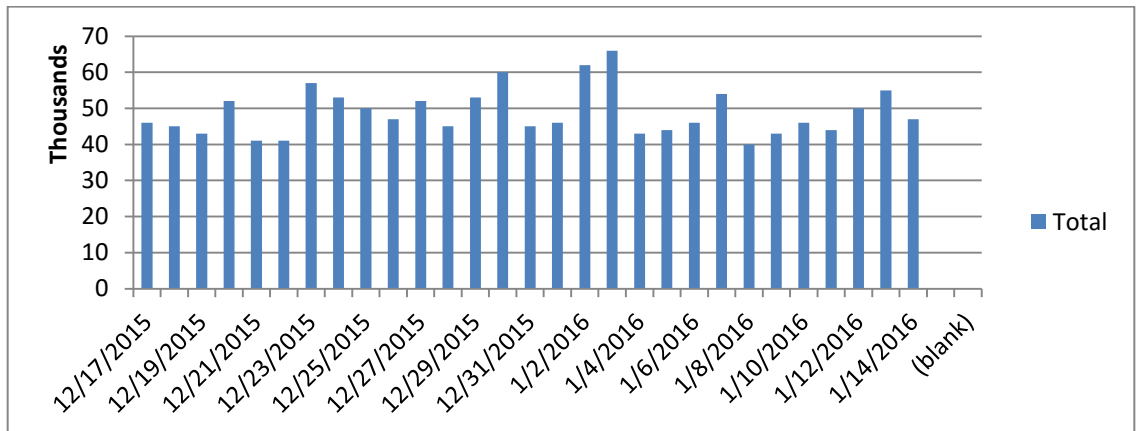
5.2.2.1.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 3,000 USD
- Hệ số tăng giá khởi điểm: 1
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt vé trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5

5.2.2.1.2. Kết quả thử nghiệm

❖ Trường hợp giá tĩnh

- Kết quả thử nghiệm từng ngày như Hình 5.4

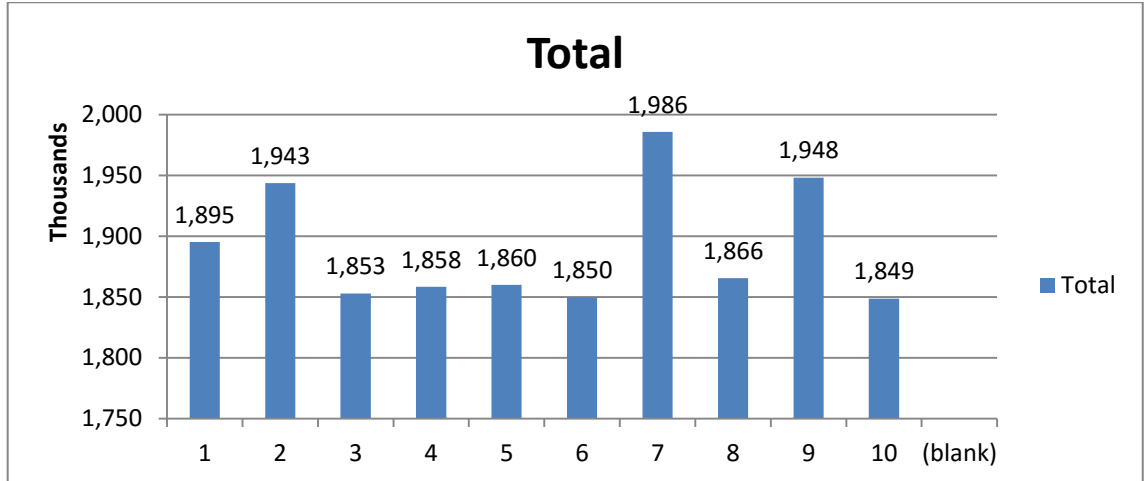


Hình 5.4. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 1,416,000 USD

❖ Trường hợp giá động

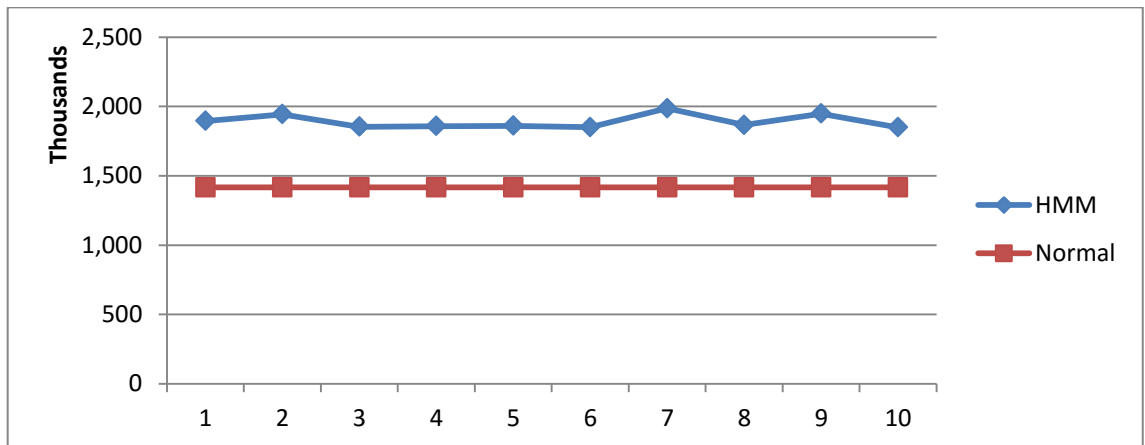
- Kết quả tổng quan doanh thu của 10 lần thử nghiệm như Hình 5.5



Hình 5.5. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 1,890,750 USD

❖ Biểu đồ so sánh 2 trường hợp như Hình 5.6



Hình 5.6. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh: 1,416,000 USD

Trung bình doanh thu của trường hợp giá động: 1,890,750 USD

Tỉ lệ chênh lệch: $(1,890,750 - 1,416,000) / 1,416,000 = 0.33528$

5.2.2.2. Thử nghiệm trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm trung bình

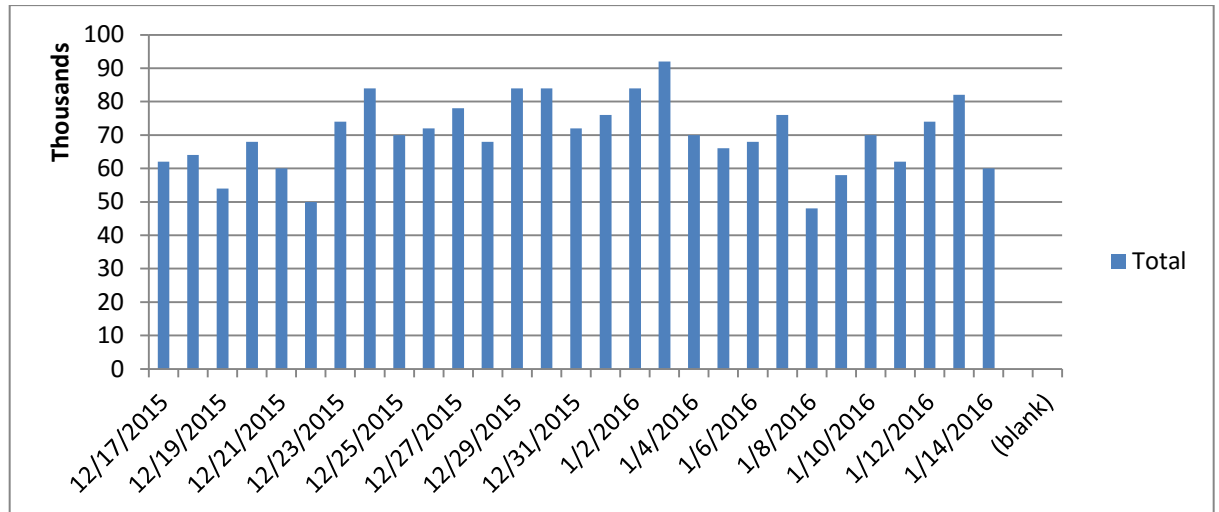
5.2.2.2.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 3,000 USD
- Hệ số tăng giá khởi điểm: 2
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5

5.2.2.2.2. Kết quả thử nghiệm

❖ Trường hợp giá tĩnh

- Kết quả thử nghiệm từng ngày như Hình 5.7

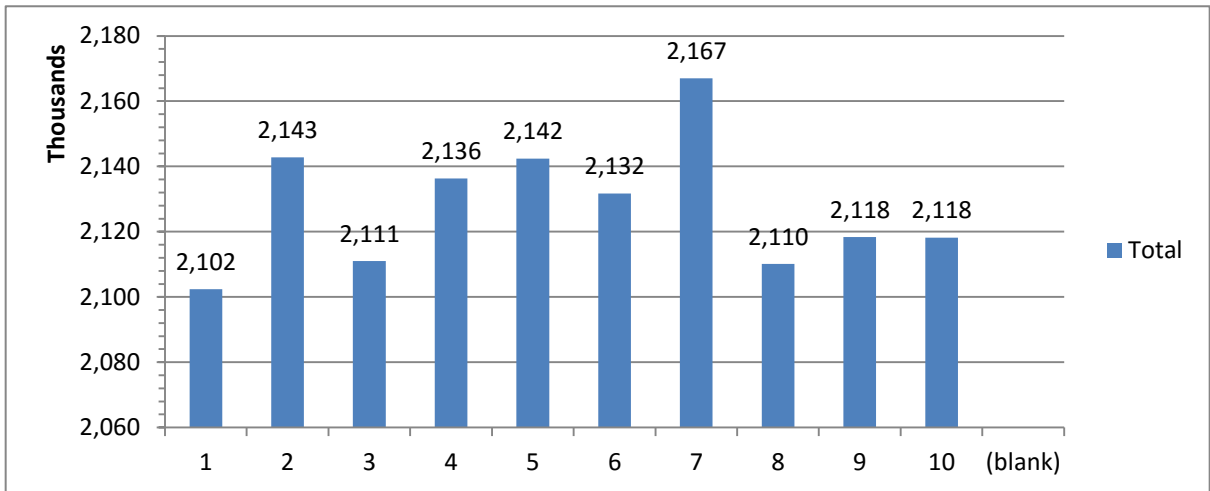


Hình 5.7. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 2,030,000 USD

❖ Trường hợp giá động

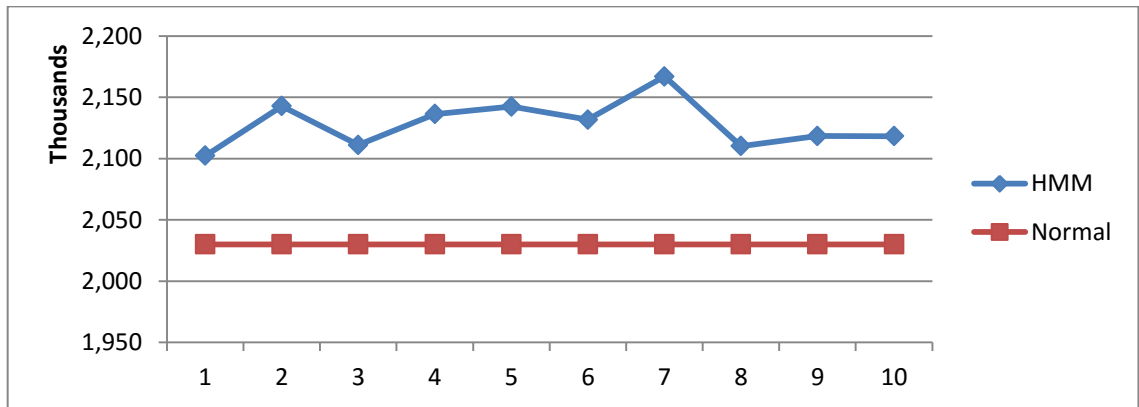
- Kết quả doanh thu của 10 lần thử nghiệm như Hình 5.8



Hình 5.8. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 2,128,026 USD

❖ Biểu đồ so sánh doanh thu trong 2 trường hợp như Hình 5.9



Hình 5.9. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh: 2,030,000 USD

Trung bình doanh thu của trường hợp giá động: 2,128,026 USD

Tỉ lệ chênh lệch: $(2,128,026 - 2,030,000) / 2,030,000 = 0.04829$

5.2.2.3. Thử nghiệm trong trường hợp bán vé theo giá tĩnh và theo giá động với giá khởi điểm cao

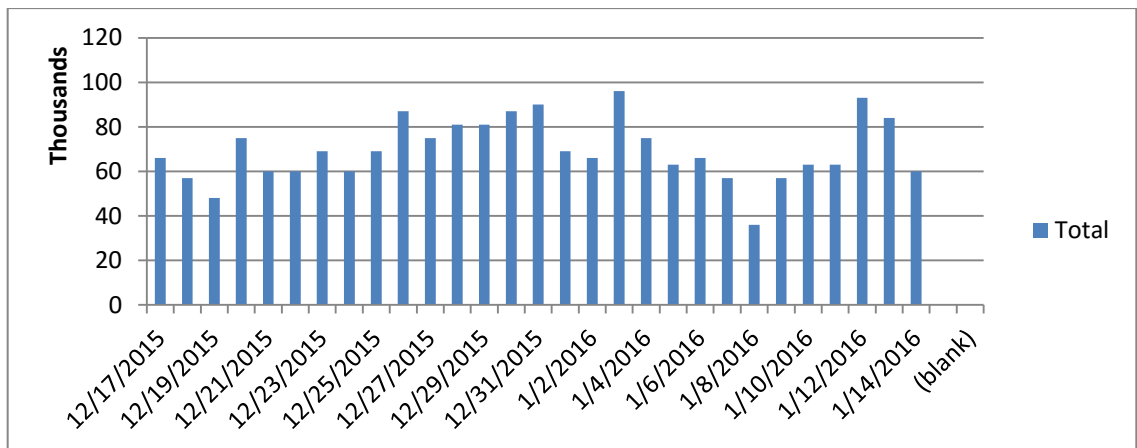
5.2.2.3.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 3,000 USD
- Hệ số tăng giá khởi điểm: 3
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5

5.2.2.3.2. Kết quả thử nghiệm

❖ Trường hợp giá tĩnh

- Kết quả thử nghiệm từng ngày như Hình 5.10

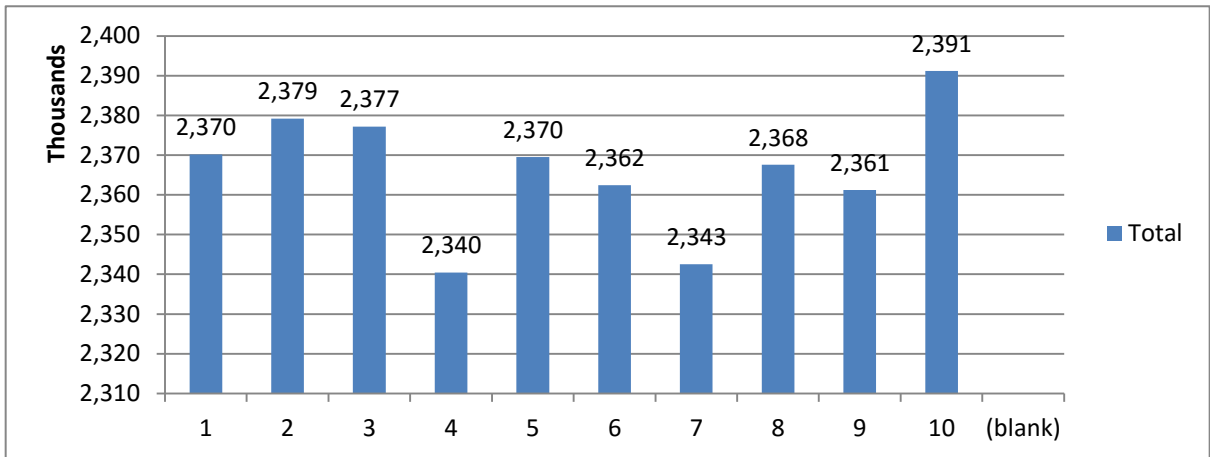


Hình 5.10. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 2,013,000 USD

❖ Trường hợp giá động

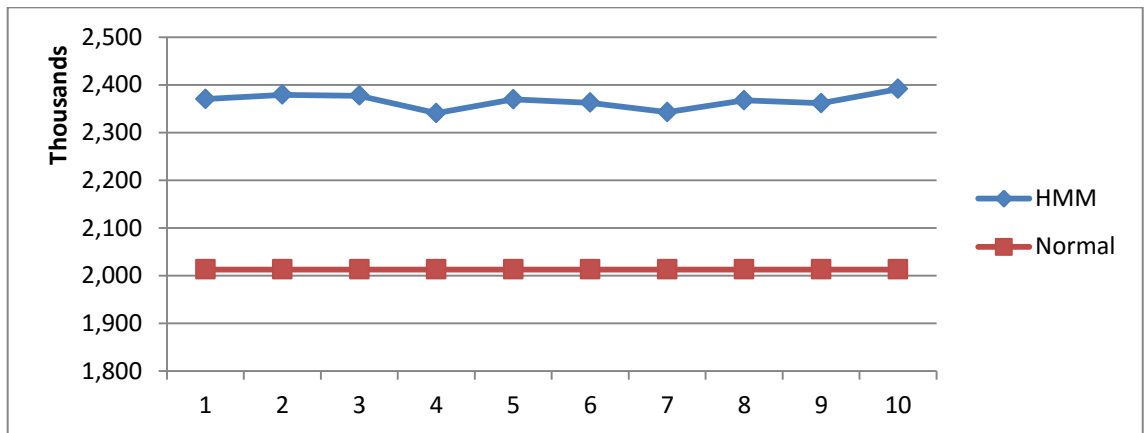
- Kết quả doanh thu của 10 lần thử nghiệm như Hình 5.11



Hình 5.11. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 2,366,154 USD

❖ Biểu đồ so sánh doanh thu trong 2 trường hợp như Hình 5.12



Hình 5.12. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh: 2,013,000 USD

Trung bình doanh thu của trường hợp giá động: 2,366,154 USD

Tỉ lệ chênh lệch: $(2,366,154 - 2,013,000) / 2,013,000 = 0.17544$

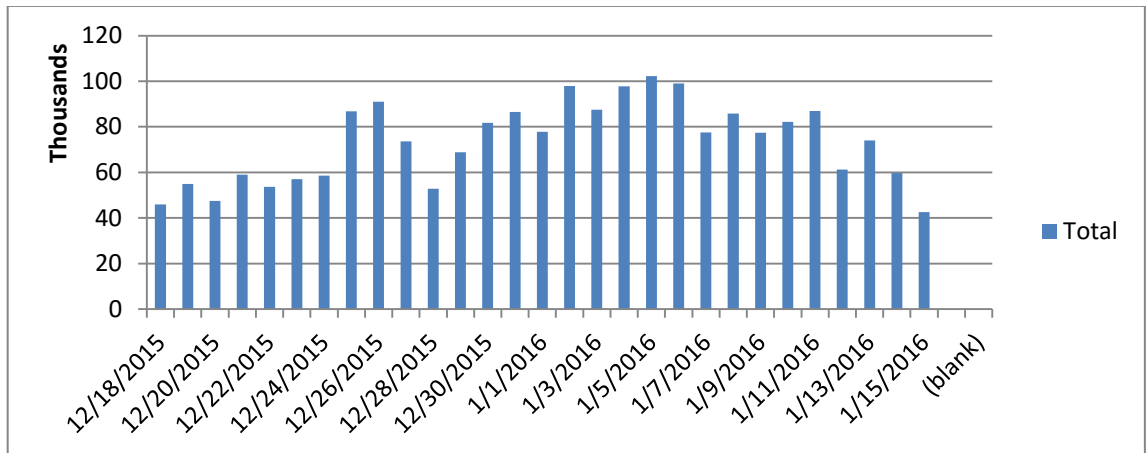
5.2.2.4. Thử nghiệm trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm thấp

5.2.2.4.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 5,000 USD
- Hệ số tăng giá khởi điểm: 1
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5
- Số ngày giá vé bị ảnh hưởng bởi sự kiện: 20 ngày
- Mức tăng lớn nhất tại thời điểm cao điểm của sự kiện: 2.5

5.2.2.4.2. Kết quả thử nghiệm

- ❖ Trường hợp giá tĩnh có sự kiện như Hình 5.13

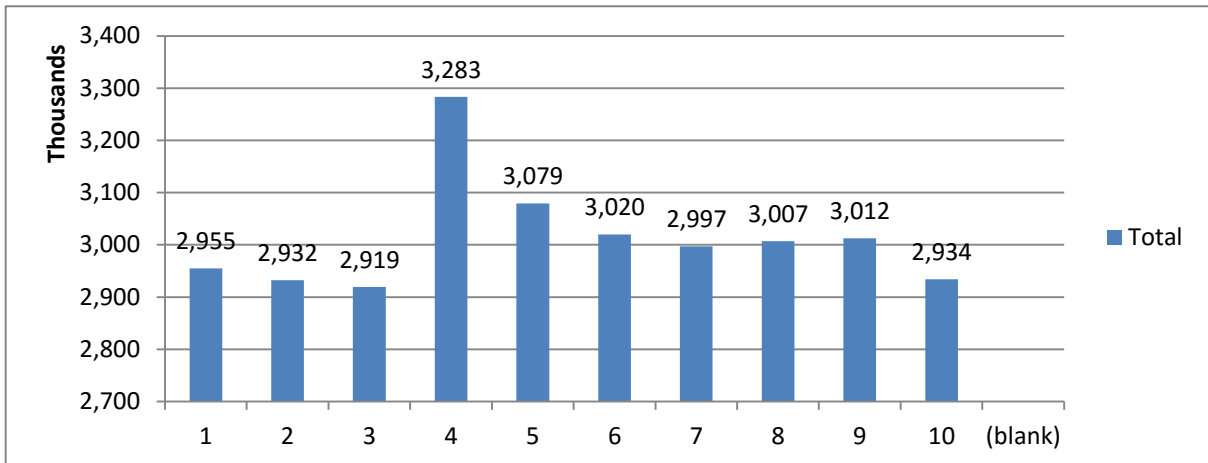


Hình 5.13. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 2,127,561 USD

❖ Trường hợp giá động có sự kiện

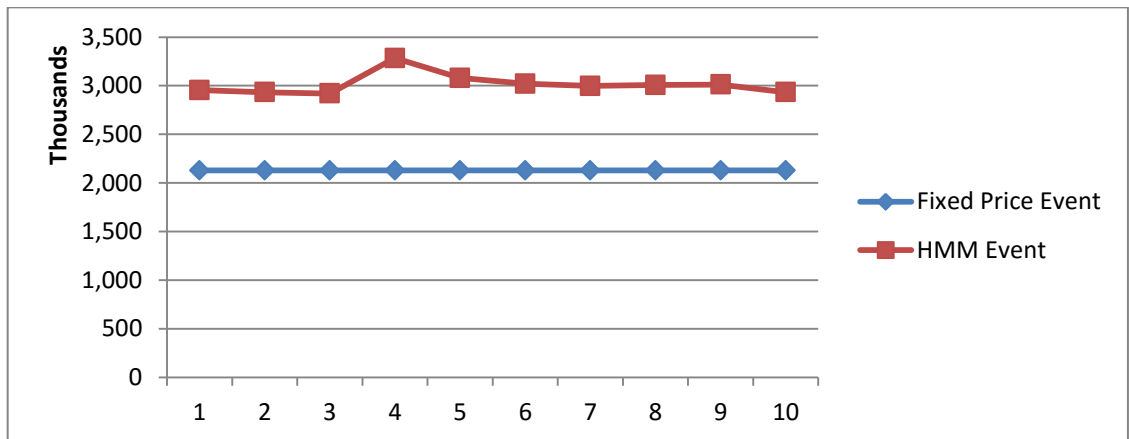
- Kết quả doanh thu của 10 lần thử nghiệm như Hình 5.14



Hình 5.14. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 3,013,928 USD

❖ Biểu đồ so sánh doanh thu trong 2 trường hợp như Hình 5.15



Hình 5.15. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh có sự kiện: 2,127,561 USD

Trung bình doanh thu của trường hợp giá động có sự kiện: 3,013,928 USD

Tỉ lệ chênh lệch: $(3,013,928 - 2,127,561) / = 0.41661$

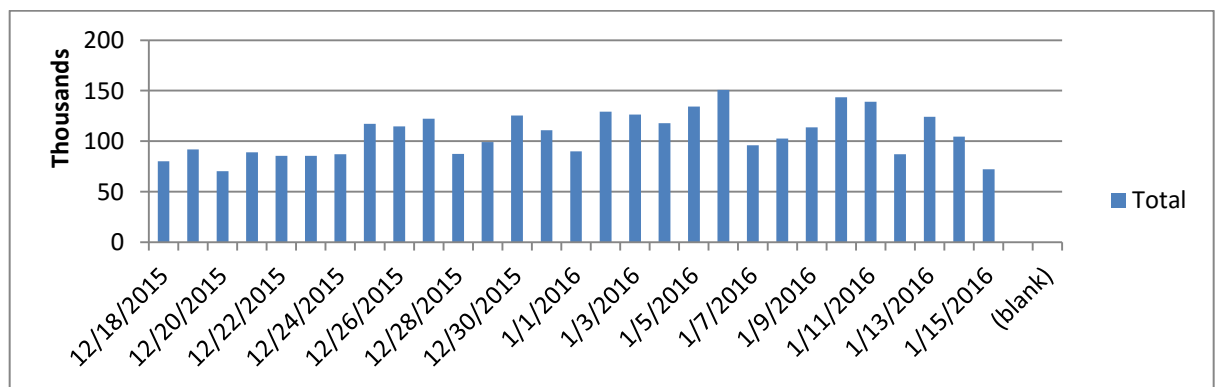
5.2.2.5. Thử nghiệm trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm trung bình

5.2.2.5.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 5,000 USD
- Hệ số tăng giá khởi điểm: 2
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5
- Số ngày giá vé bị ảnh hưởng bởi sự kiện: 20 ngày
- Mức tăng tối đa tại thời điểm cao điểm của sự kiện: 2.5

5.2.2.5.2. Kết quả thử nghiệm

- ❖ Trường hợp giá tĩnh có sự kiện
- Kết quả thống kê theo ngày như Hình 5.16

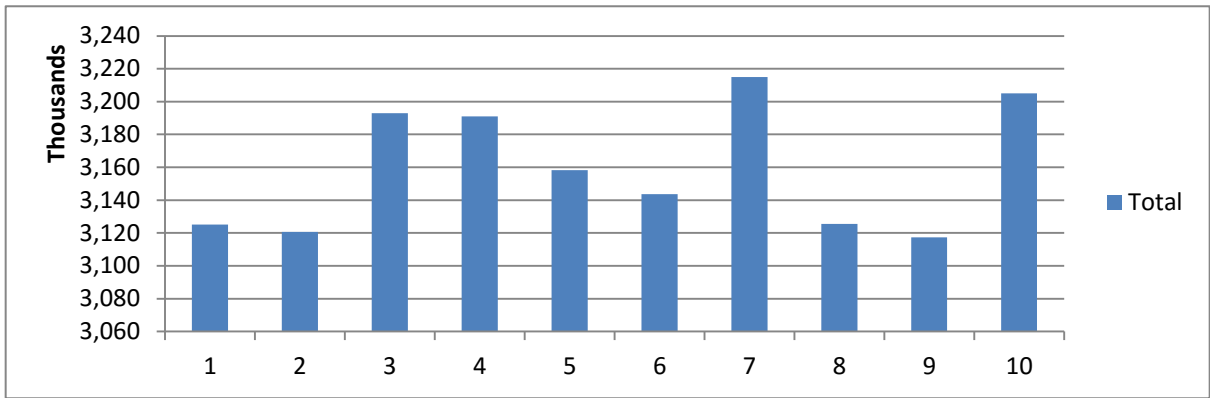


Hình 5.16. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 3,097,714 USD

❖ Trường hợp giá động có sự kiện

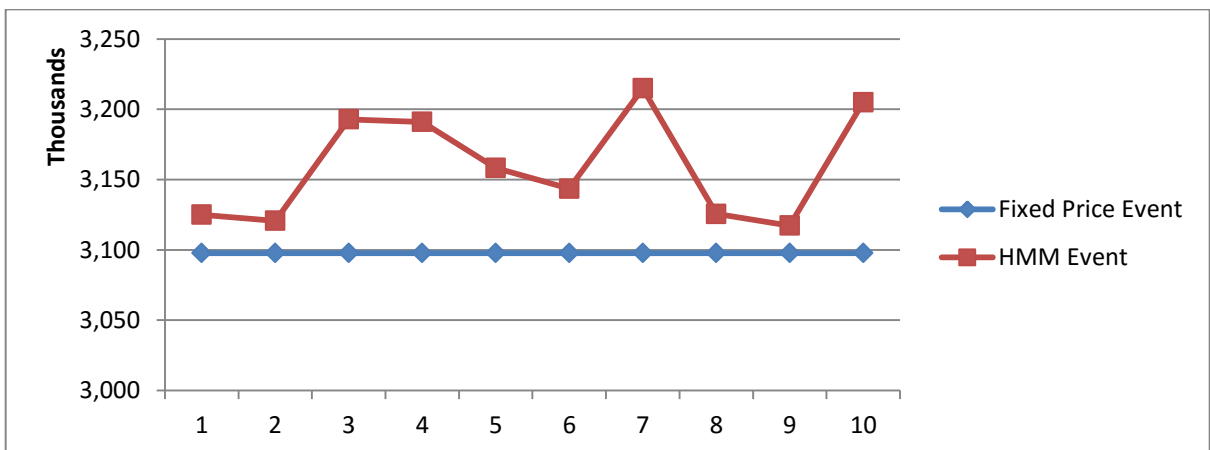
- Kết quả doanh thu của 10 lần thử nghiệm như Hình 5.17



Hình 5.17. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 3,159,438 USD

❖ Biểu đồ so sánh doanh thu trong 2 trường hợp như Hình 5.18



Hình 5.18. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh có sự kiện: 3,097,714 USD

Trung bình doanh thu của trường hợp giá động có sự kiện: 3,159,438 USD

Tỉ lệ chênh lệch: $(3,159,438 - 3,097,714) / 3,097,714 = 0.01993$

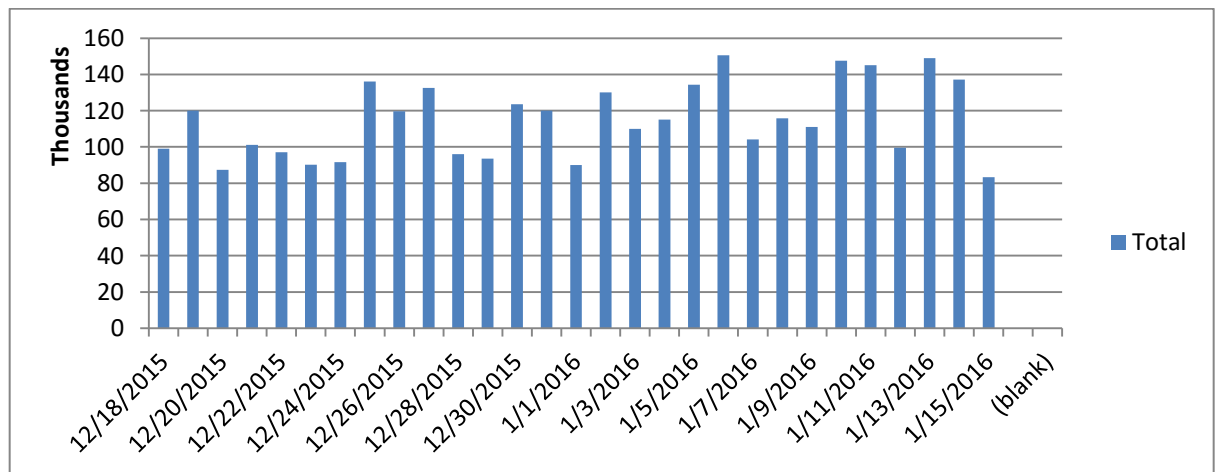
5.2.2.6. Thử nghiệm trong trường hợp bán vé theo giá tĩnh có sự kiện và theo giá động có sự kiện với giá khởi điểm cao

5.2.2.6.1. Thông số đầu vào của thử nghiệm

- Thời gian mô phỏng: 30 ngày
- Giá cơ bản chuyến bay: 1,000 USD
- Giá tối đa: 5,000 USD
- Hệ số tăng giá khởi điểm: 3
- Tổng số người có nhu cầu tìm kiếm và đặt vé: 1000 người
- Khoảng thời gian cho phép đặt trước ngày khởi hành: 10 ngày
- Số lượng vé tối đa mỗi người có nhu cầu đặt: 3 vé
- Hệ số điều chỉnh: 0.5
- Số ngày giá vé bị ảnh hưởng bởi sự kiện: 20 ngày
- Mức tăng tối đa tại thời điểm cao điểm của sự kiện: 2.5

5.2.2.6.2. Kết quả thử nghiệm

- ❖ Trường hợp giá tĩnh có sự kiện
- Kết quả thống kê theo ngày như Hình 5.19

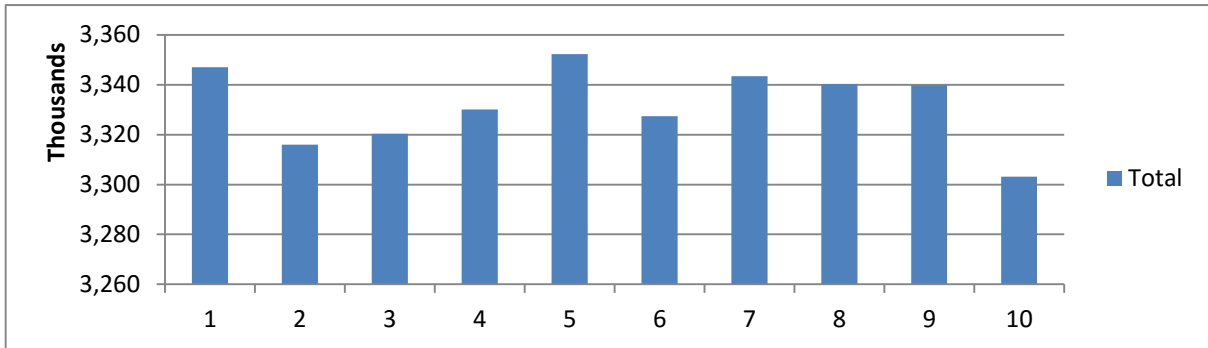


Hình 5.19. Biểu đồ biểu diễn doanh thu theo ngày trường hợp giá tĩnh

Doanh thu: 3,330,548 USD

❖ Trường hợp giá động có sự kiện

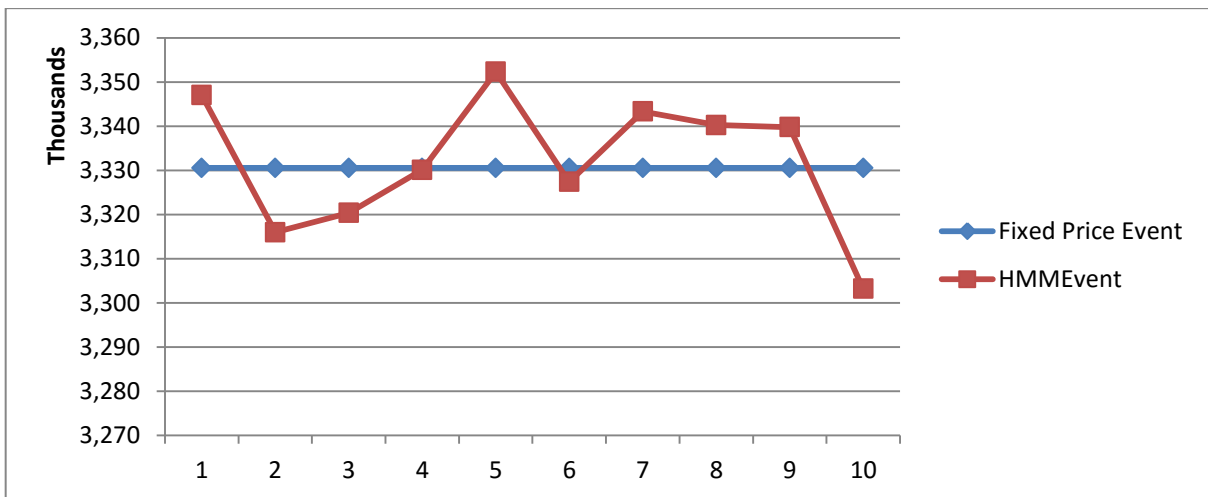
- Kết quả doanh thu của 10 lần thử nghiệm như Hình 5.20



Hình 5.20. Biểu đồ biểu diễn doanh thu qua 10 lần thử nghiệm trường hợp giá động

Trung bình doanh thu: 3,331,961 USD

❖ Biểu đồ so sánh doanh thu trong 2 trường hợp như Hình 5.21



Hình 5.21. Biểu đồ so sánh doanh thu trong 2 trường hợp

Doanh thu trường hợp giá tĩnh có sự kiện: 3,330,548 USD

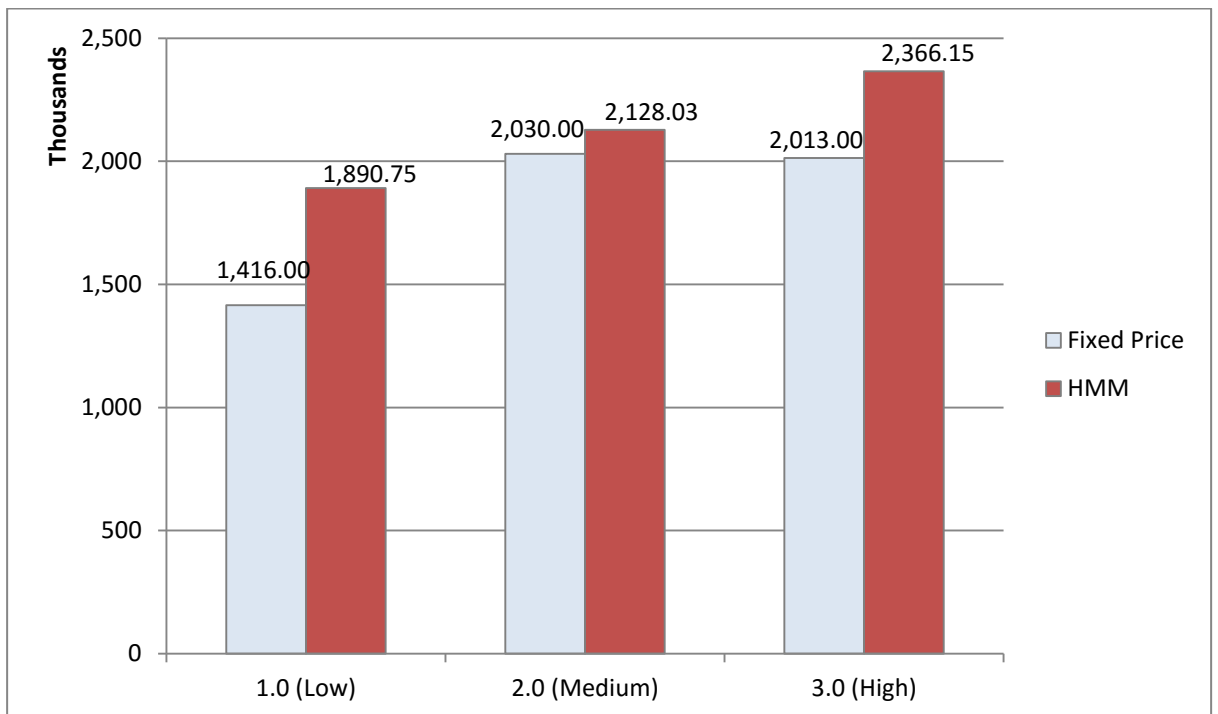
Trung bình doanh thu của trường hợp giá động có sự kiện: 3,331,961 USD

Tỉ lệ chênh lệch: $(3,331,961 - 3,330,548) / 3,330,548 = 0.00042$

5.2.3. Đánh giá kết quả thử nghiệm

5.2.3.1. Tổng quan kết quả thử nghiệm

Kết quả tổng quan được biểu diễn như biểu đồ trong hình Hình 5.22 và Hình 5.23



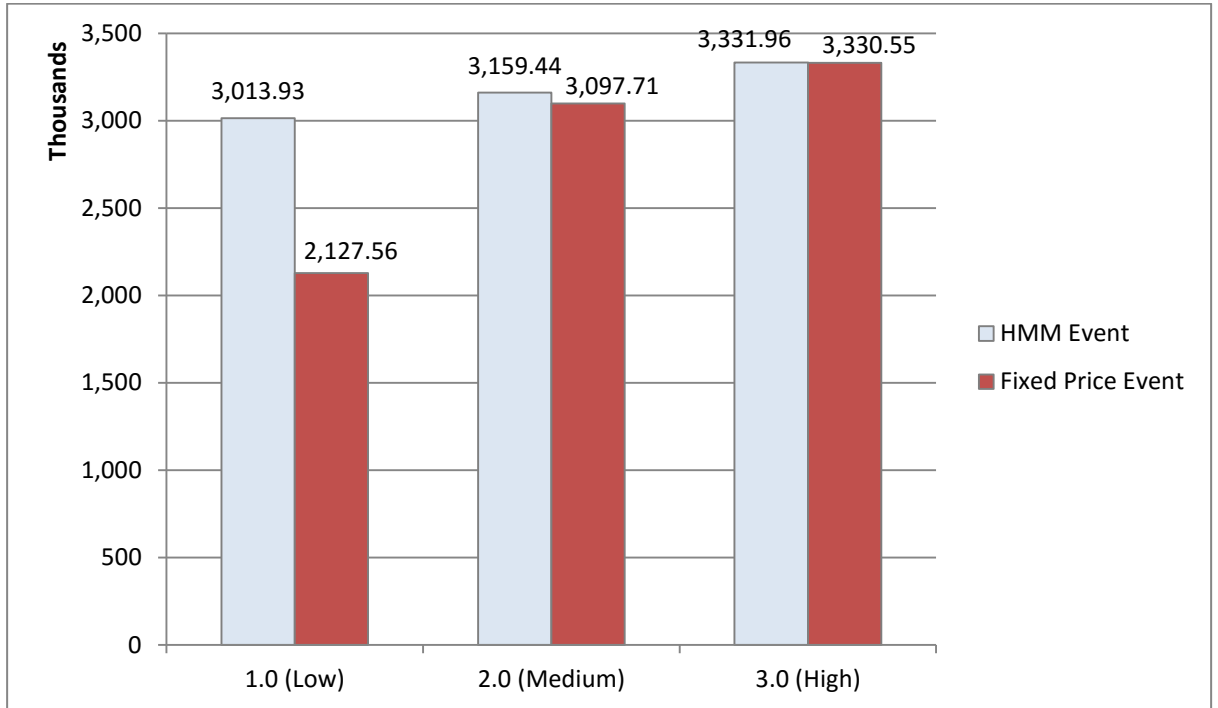
Hình 5.22. Biểu đồ tổng quan so sánh các trường hợp giá tĩnh và giá động

Chú thích:

1.0 (Low): Trường hợp giá khởi điểm thấp

2.0 (Medium): Trường hợp giá khởi điểm trung bình

3.0 (High): Trường hợp giá khởi điểm cao



Hình 5.23. Biểu đồ tổng quan so sánh các trường hợp giá tĩnh và giá động có sự kiện

Chú thích:

1.0 (Low): Trường hợp giá khởi điểm thấp

2.0 (Medium): Trường hợp giá khởi điểm trung bình

3.0 (High): Trường hợp giá khởi điểm cao

5.2.3.2. Đánh giá và nhận xét

Từ các kết quả so sánh tổng quan trên, ta có thể nhận thấy trong 6 trường hợp cần đánh giá thì tất cả đều cho kết quả như mong muốn. Đó là trường hợp bán vé theo giá động được đề xuất bởi mô hình HMM đều cho kết quả doanh thu cao hơn trường hợp bán vé theo giá tĩnh hoặc giá tĩnh có sự kiện. Tuy nhiên, trong trường hợp có sự kiện và bán với giá khởi điểm cao thì trường hợp giá động theo HMM cho tỉ lệ chênh lệch còn khá thấp so với các trường hợp còn lại. Vì trong trường hợp có sự kiện thì nhu cầu

di chuyển cao, nên trong nhiều trường hợp dù giá vé cao người ta vẫn sẵn sàng bỏ tiền mua. Vì vậy, trong kịch bản giá khởi điểm đã đưa lên mức cao (3000USD) hơn mức giá tại thời điểm cao điểm sự kiện (2500USD) thì mô hình HMM không phát huy được hiệu quả cao như các trường hợp khác.

CHƯƠNG 6

KẾT LUẬN

Mô hình HMM là mô hình phù hợp để áp dụng cho bài toán xác định giá vé máy bay tự động vì nó vừa có mô hình học từ dữ liệu quá khứ, vừa có thể đưa ra quyết định dựa vào dữ liệu quan sát và kết quả phản hồi của quyết định kể trước. Mô hình xác định giá vé tự động được xây dựng dựa trên mô hình HMM cùng với đường phân phối Gaussian đã cho kết quả tốt hơn mô hình giá tĩnh thông thường. Tuy nhiên, kết quả còn có thể tốt hơn nếu xác định được bộ dữ liệu học máy chính xác hơn. Mô hình này cho phép người sử dụng có thể thay đổi bộ dữ liệu này theo kinh nghiệm qua thời gian để tối ưu hóa doanh thu hoặc tối ưu hóa về số lượng vé tùy theo nhu cầu cụ thể. Đến thời điểm hoàn thành luận văn, tác giả chưa tìm thấy nghiên cứu liên quan nào tương tự được áp dụng cho việc xác định giá vé máy bay tự động.

6.1. KẾT QUẢ ĐẠT ĐƯỢC

Qua việc nghiên cứu thực hiện đề tài, tác giả đã thu được những kết quả sau:

- Hiểu qui trình bán vé máy bay của các đại lý.
- Tìm hiểu được các yếu tố ảnh hưởng đến giá vé máy bay.
- Hiểu cách xác định giá vé máy bay của các đại lý.
- Hiểu và áp dụng được mô hình HMM vào bài toán của luận văn.
- Hiểu và áp dụng được lý thuyết về đường phân phối chuẩn hay còn gọi là đường phân phối Gaussian vào bài toán của luận văn.
- Xây dựng được công thức tính mức độ tăng giảm giá vé dựa vào dữ liệu quan sát.
- Xây dựng được hệ thống mô phỏng hệ thống thay đổi giá vé tự động, quy trình bán vé và đặt vé.
- Chứng minh được mô hình giá động đã xây dựng đem lại kết quả tốt hơn so với mô hình giá tĩnh.

6.2. HẠN CHẾ CỦA HỆ THỐNG

- Chưa áp dụng được mô hình cho trường hợp nhiều sự kiện cùng xảy ra tại một điểm đến.
- Chưa phát sinh actor theo hàm phân phối.
- Chưa áp dụng được cho loại vé hai chiều (có khứ hồi).
- Chưa xét yếu tố hủy vé, giảm giá theo số lượng, khuyến mãi,...
- Chưa xem xét đến yếu tố đối thủ cạnh tranh.
- Dữ liệu quan sát chỉ mới dừng ở 2 yếu tố quan sát là số lượng tìm kiếm và tỉ lệ chuyển đổi.
- Chưa xét trường trường hợp số lượng tìm kiếm và tỉ lệ chuyển đổi bằng nhau qua các lần quyết định (như $DS=CR=$, $DS=CR+$, $DS=CR-$, $DS+CR=$, $DS-CR=$)
-

6.3. HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI

- Mở rộng dữ liệu quan sát để thêm nhiều yếu tố quan sát thay vì chỉ có số lượng tìm kiếm và tỉ lệ chuyển đổi. Ví dụ như yếu tố tìm kiếm theo từng giờ, tỉ lệ chuyển đổi theo từng giờ hoặc các yếu tố có ảnh hưởng bởi giá vé máy bay.
- Cải tiến hệ thống mô phỏng để phát sinh actor theo hàm phân phối chuẩn.
- Mở rộng cho trường hợp có nhiều sự kiện cùng diễn ra tại một điểm đến nào đó và có chồng chéo ảnh hưởng lên nhau.
- Đề xuất giá vé cho trường hợp vé hai chiều.
- Có thể mở rộng áp dụng cho trường hợp xác định giá theo giờ.
- Hệ thống có thể xác định được giá của đối thủ cạnh tranh để giá đề xuất có thể cạnh tranh được trên thị trường. Khi xem xét giá đối thủ cạnh tranh có thể lấy dữ liệu từ các trung tâm dữ liệu giá vé như Skype Scanner hoặc crawl trực tiếp dữ liệu từ các trang web bán vé máy bay trực tuyến.

- Mở rộng mô hình cho các loại sản phẩm, dịch vụ trực tuyến khác phù hợp với chiến lược kinh doanh theo mô hình giá động.

TÀI LIỆU THAM KHẢO

- [1] B.R. Baker III. "Benefits tracking and correlation system for use with third-party enabling organization". U.S. Patent No. 5,864,822. 26 Jan. 1999.
- [2] E.E. Bailey, D.R. Graham, and D.P. Kaplan. "Deregulating the Airlines". MIT press, Vol. 10. 1985.
- [3] B.J. Garback. "Computer travel planning system". U.S. Patent No. 5,237,499. 17 Aug. 1993.
- [4] R.S. Tagawa. "Self-service system for selling travel-related services or products". U.S. Patent No. 5,732,398. 24 Mar. 1998.
- [5] S.B. Heintzeman, T.W. Storey, B. Monson, S.J. Medina, and G.A. Malark. "System for awarding credits to persons who book travel-related reservations". U.S. Patent No. 5,483,444. 9 Jan. 1996.
- [6] W.L. Flake, K. Kambhampaty, S.B. Molsberry, and G.B. Clem. "Automated travel service management information system". U.S. Patent No. 5,832,451. 3 Nov. 1998.
- [7] D.M. DeLorme, K.A. Gray, and A.T. Ferguson. "Travel reservation information and planning system". U.S. Patent No. 5,948,040. 7 Sep. 1999.
- [8] A.G. Acebo, and C.D. Frawley. "Pre- and post-ticketed travel reservation information management system". U.S. Patent No. 6,023,679. 8 Feb. 2000.
- [9] E.A. Fusz, and C.A. Kline. "Online product exchange system". U.S. Patent No. 7,103,568. 5 Sep. 2006.
- [10] L.B. Lockwood. "Automated sales system". U.S. Patent No. 5,309,355. 3 May 1994.
- [11] R.P. Sehr. "Travel system and methods utilizing multi-application passenger cards". U.S. Patent No. 6,085,976. 11 Jul. 2000.

- [12] J.S. Walker, D.E. Tedesco, A.S. VanLuchene, J.A. Jorasch, and T.S. Case. "Conditional purchase offer (CPO) management system for packages". U.S. Patent No. 6,553,346. 22 Apr. 2003.
- [13] S. Dalnekoff, and M.L. Schneider. "Computerized system with means to automatically clear and sell wait-listed customer reservations ". U.S. Patent No. 4,931,932. 5 Jun. 1990.
- [14] J.R. Shoolery, G.E. Pasela, D.L. De La Torre, K.M. Leung, and K.A. Morris. "Corporate travel controller". U.S. Patent No. 5,570,283. 29 Oct. 1996.
- [15] Y. Narahari and et al. "Dynamic pricing models for electronic business". In Sadhana (Academy Proceedings in Engineering Sciences). Indian Academy of Sciences. Vol. 30. No. 2: 3. 2005.
- [16] L.R. Rabiner, and B.H Juang. "An introduction to hidden Markov models." ASSP Magazine, IEEE 3. No. 1: 4-16. 1986.
- [17] G.L Kouemou. "History and theoretical basics of hidden Markov Models". INTECH Open Access Publisher. 2011.