

BỘ GIAO THÔNG VẬN TẢI BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC HÀNG HẢI VIỆT NAM



HOÀNG HỒNG GIANG

**NGHIÊN CỨU ỨNG DỤNG TRÍ TUỆ NHÂN TẠO
VÀ DỮ LIỆU LỚN TRONG QUẢN LÝ
GIAO THÔNG HÀNG HẢI TRÊN TUYẾN LUỒNG
VŨNG TÀU - SÀI GÒN**

LUẬN ÁN TIẾN SĨ KỸ THUẬT

HẢI PHÒNG - 2024

BỘ GIAO THÔNG VẬN TẢI

BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC HÀNG HẢI VIỆT NAM

HOÀNG HỒNG GIANG

NGHIÊN CỨU ỨNG DỤNG TRÍ TUỆ NHÂN TẠO
VÀ DỮ LIỆU LỚN TRONG QUẢN LÝ
GIAO THÔNG HÀNG HẢI TRÊN TUYẾN LUỒNG
VŨNG TÀU - SÀI GÒN

LUẬN ÁN TIẾN SĨ KỸ THUẬT

NGÀNH: KHOA HỌC HÀNG HẢI;

MÃ SỐ: 9840106

NGÀNH: KHOA HỌC HÀNG HẢI

Người hướng dẫn khoa học: 1. PGS.TS. Phạm Văn Thuận

2. PGS.TS. Trần Văn Lượng

HẢI PHÒNG - 2024

LỜI CAM ĐOAN

Tôi xin cam đoan đề tài: *Nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu lớn trong quản lý giao thông hàng hải trên tuyến luồng Vũng Tàu - Sài Gòn* là công trình nghiên cứu của riêng tôi, không có phần nội dung nào được sao chép một cách bất hợp pháp từ công trình nghiên cứu của tác giả khác.

Kết quả nghiên cứu, nguồn số liệu trích dẫn, tài liệu tham khảo và kế thừa nêu trong luận án đều được chỉ rõ về nguồn gốc, xuất xứ và đảm bảo tính trung thực.

Hải Phòng, ngày tháng năm 2024

NGHIÊN CỨU SINH

Hoàng Hồng Giang

LỜI CẢM ƠN

Tôi xin chân thành cảm ơn Trường Đại học Hàng hải Việt Nam đã tạo điều kiện cho tôi học tập, nghiên cứu và thực hiện Luận án này.

Tôi xin trân trọng bày tỏ lòng tri ân sâu sắc đến thầy giáo hướng dẫn trực tiếp PGS.TS. Phạm Văn Thuận, Trường Đại học Hàng hải Việt Nam đã tận tình chỉ bảo, hướng dẫn trong suốt quá trình học tập, nghiên cứu và thực hiện Luận án tiến sĩ.

Tôi cũng xin chân thành cảm ơn các thầy giáo, các nhà khoa học, Khoa Hàng hải, Viện Đào tạo Sau đại học, Trường Đại học Hàng hải Việt Nam đã tạo điều kiện thuận lợi cho tôi trong suốt quá trình học tập và nghiên cứu thực hiện luận án.

Cuối cùng, tôi xin bày tỏ lòng biết ơn đến gia đình, bạn bè, đồng nghiệp đã luôn động viên, khuyến khích, giúp đỡ để nghiên cứu sinh hoàn thành Luận án này.

Hải Phòng, ngày tháng năm 2024

NGHIÊN CỨU SINH

Hoàng Hồng Giang

MỤC LỤC

LỜI CAM ĐOAN.....	i
LỜI CẢM ƠN	ii
MỤC LỤC.....	iii
DANH MỤC CÁC CHỮ VIẾT TẮT VÀ KÝ HIỆU	vi
DANH MỤC CÁC BẢNG.....	viii
DANH MỤC CÁC HÌNH.....	ix
MỞ ĐẦU	1
CHƯƠNG 1. NGHIÊN CỨU CƠ SỞ LÝ THUYẾT NHẬN DẠNG, ĐÁNH GIÁ NGUY CƠ ĐÂM VÀ GIỮA CÁC THUYỀN TRÊN TUYẾN LUỒNG VŨNG TÀU - SÀI GÒN.....	7
1.1. Tổng quan về vấn đề nghiên cứu	7
1.2. Đánh giá nguy cơ đâm va trong hàng hải và trên tuyến luồng Vũng Tàu - Sài Gòn.....	13
1.3. Nghiên cứu xây dựng thuật toán xác định nguy cơ đâm va trong hàng hải giữa 2 tàu mục tiêu được quan sát từ radar của trạm VTS tại khu vực biển rộng	22
1.4. Nghiên cứu nguyên tắc xác định nguy cơ đâm va giữa các mục tiêu áp dụng trong khu vực luồng lạch hẹp.....	37
1.5. Kết luận chương 1	39
CHƯƠNG 2. TỔNG QUAN VỀ ĐÁNH GIÁ NGUY CƠ ĐÂM VÀ GIỮA CÁC MỤC TIÊU TRÊN VÙNG BIỂN RỘNG VÀ TRONG LUỒNG CHẠY TÀU	40
2.1. Ứng dụng hệ thống mô phỏng của Trường Đại học Hàng hải Việt Nam trong việc kiểm tra thuật toán xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát	40

2.2. Xây dựng bài tập mô phỏng kiểm tra thuật toán xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát	45
2.3. Đánh giá kết quả thực nghiệm xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát từ trạm VTS	46
2.4. Xác định điểm nóng giao thông trên khu vực luồng Vũng Tàu - Sài Gòn ..	51
2.5. Xây dựng công thức tính toán xác định nguy cơ đâm va theo <i>ETA</i> tới các điểm nóng giao thông trên tuyến luồng Vũng Tàu - Sài Gòn	60
2.6. Kết luận chương 2	64
CHƯƠNG 3. SỬ DỤNG AI ĐÁNH GIÁ NGUY CƠ ĐÂM VA GIỮA HAI TÀU THUYỀN TẠI KHU VỰC ĐẦU LUỒNG VŨNG TÀU - SÀI GÒN.....	
3.1. Dữ liệu tính toán xác định nguy cơ đâm va trên vùng biển rộng.....	65
3.2. Đánh giá bộ dữ liệu phục vụ huấn luyện mô hình trên vùng biển rộng.....	66
3.3. Xây dựng mô hình mạng nơ-ron để huấn luyện	68
3.4. Đánh giá kết quả huấn luyện mô hình mạng nơ ron nhân tạo trên vùng biển rộng.....	71
3.5. Kết luận chương 3	73
CHƯƠNG 4. SỬ DỤNG AI ĐÁNH GIÁ NGUY CƠ ĐÂM VA TẠI KHU VỰC CHẬT HẸP CỦA LUỒNG VŨNG TÀU - SÀI GÒN.....	
4.1. Dữ liệu về nguy cơ đâm va trên luồng lạch hẹp đưa vào mô hình tính toán	74
4.2. Phân tích dữ liệu.....	75
4.2.1. Trực quan hóa dữ liệu	75
4.2.2. Tiền xử lý dữ liệu	76
4.2.3. Độ đo sử dụng	77
4.2.4. Sự quá khớp (Overfitting)	79
4.3. Lựa chọn và thực hành huấn luyện mô hình AI cho luồng lạch hẹp	80
4.3.1. Mô hình hồi quy đa thức kết hợp Lasso (Polynomial kết hợp với Lasso)	80
4.3.2. Phân tích và đánh giá	91

4.3.3. Đánh giá độ tin cậy của mô hình đề xuất trên tập dữ liệu lớn	94
4.4. Kết luận chương 4	96
KẾT LUẬN VÀ KIẾN NGHỊ.....	98
TÀI LIỆU THAM KHẢO.....	101
PHỤ LỤC 1: CÁC THUẬT TOÁN AI ĐƯỢC THỬ NGHIỆM ĐỐI VỚI TẬP DỮ LIỆU	113
PHỤ LỤC 2: CHI TIẾT QUÁ TRÌNH HUẤN LUYỆN MẠNG NƠ-RON NHÂN TẠO VỚI DỮ LIỆU CHẠY TÀU TRÊN VÙNG BIỂN RỘNG	127

DANH MỤC CÁC CHỮ VIẾT TẮT VÀ KÝ HIỆU

<i>Chữ viết tắt</i>	<i>Giải thích</i>
AIS	Automatic Identification System <i>Hệ thống tự động nhận dạng tàu biển</i>
ARPA	Automatic Radar Plotting Aids <i>Thiết bị đồ giải radar tự động</i>
CCTV	Closed Circuit Television <i>Hệ thống giám sát hình ảnh</i>
COLREG 1972	International Regulations for Preventing Collisions at Sea <i>Quy tắc quốc tế về phòng ngừa đâm va tàu biển năm 1972</i>
CPA	Closest Point of Approach <i>Cận điểm</i>
CRI	Collision Risk Index <i>Chỉ số nguy cơ đâm va</i>
DBSCAN	Density - Based Spatial Clustering of Applications with Noise <i>Thuật toán phân cụm dựa trên mật độ không gian với các dạng dữ liệu có nhiễu</i>
DCPA	Distance to Closest Point of Approach <i>Khoảng cách đến cận điểm</i>
DCR	Dynamic Collision Risk <i>Nguy cơ đâm va động</i>
DNV	Det Norske Veritas
ECDIS	Electronic Chart Display and Information System <i>Hệ thống hiển thị thông tin hải đồ điện tử</i>
ETA	Estimated Time of Arrival <i>Thời gian dự kiến tàu đến điểm nóng giao thông</i>
FN	False Negative <i>Sai lầm loại 1</i>

FP	False Positive <i>Sai lầm loại 2</i>
GMDSS	Global Maritime Distress and Safety System <i>Hệ thống thông tin cứu nạn và an toàn hàng hải toàn cầu</i>
MAE	Mean Absolute Error <i>Phương pháp đo lường sự khác biệt giữa hai biến liên tục</i>
MMSI	Maritime Mobile Service Identity <i>Mã nhận dạng dịch vụ di động hàng hải</i>
MSE	Mean Squared Error <i>Sai số bình phương trung bình</i>
OS	Owner Ship <i>Tàu chủ</i>
SPCT	Saigon Premier Container Terminal – SPCT <i>Cảng container trung tâm Sài Gòn</i>
TCPA	Time to Closest Point of Approach <i>Thời gian đến cận điểm</i>
TN	True Negative <i>Tổng số trường hợp dự báo khớp mẫu âm tính</i>
TP	True Positive <i>Tổng số trường hợp dự báo khớp mẫu dương tính</i>
TP HCM	Thành phố Hồ Chí Minh
TS	Target Ship <i>Tàu mục tiêu</i>
VHF	Very High Frequency <i>Tần số sóng cao</i>
VLCC	Very Large Crude oil Carrier <i>Tàu chở dầu thô cỡ lớn</i>
VTS	Vessel Traffic Service <i>Hệ thống giám sát và điều phối giao thông hàng hải</i>

DANH MỤC CÁC BẢNG

<i>Số bảng</i>	<i>Tên bảng</i>	<i>Trang</i>
2.1	Vị trí các trạm VTS ảo	46
2.2	Kết quả thực nghiệm	47
2.3	Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS1	48
2.4	Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS2	48
2.5	Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS3	49
2.6	Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS4	49
2.7	So sánh giá trị tính toán và giá trị ARPA hiển thị	50
3.1	Dữ liệu ban đầu	66
3.2	Một số mẫu dữ liệu sau khi chuẩn hóa	67
3.3	Giá trị trung bình và độ lệch chuẩn sau khi đã chuẩn hóa dữ liệu	68

DANH MỤC CÁC HÌNH

<i>Số hình</i>	<i>Tên hình</i>	<i>Trang</i>
1.1	Tồn tại nguy cơ đâm va khi phương vị không đổi, khoảng cách giảm dần	14
1.2	Tồn tại nguy cơ đâm va khi phương vị thay đổi không rõ rệt, khoảng cách giảm dần	15
1.3	Trường hợp phương vị không đổi, khoảng cách giảm dần	16
1.4	Trường hợp phương vị thay đổi, khoảng cách giảm dần	17
1.5	Trường hợp tàu mục tiêu ở cung phần tư thứ nhất so với trạm VTS	24
1.6	Trường hợp tàu mục tiêu ở cung phần tư thứ hai so với trạm VTS	24
1.7	Trường hợp tàu mục tiêu ở cung phần tư thứ ba so với trạm VTS	25
1.8	Trường hợp tàu mục tiêu ở cung phần tư thứ tư so với trạm VTS	26
1.9	Trường hợp tàu mục tiêu 2 ở cung phần tư thứ nhất so với tàu 1	26
1.10	Trường hợp tàu mục tiêu 2 ở cung phần tư thứ hai so với tàu 1	27
1.11	Trường hợp tàu mục tiêu 2 ở cung phần tư thứ ba so với tàu 1	27
1.12	Trường hợp tàu mục tiêu 2 ở cung phần tư thứ tư so với tàu 1	28
1.13	Trường hợp phương vị không đổi, khoảng cách giảm dần	29
1.14	Trường hợp phương vị không đổi, khoảng cách tăng dần	31
1.15	Trường hợp phương vị thay đổi, khoảng cách giảm dần	31
1.16	Trường hợp phương vị thay đổi, khoảng cách tăng dần	33
2.1	Hệ thống mô phỏng buồng lái tại Trường Đại học Hàng hải Việt Nam	40
2.2	Chứng chỉ của DNV cấp cho hệ thống mô phỏng	41

2.3	Bố trí tàu và các trạm VTS ảo trong bài tập thực nghiệm	45
2.4	Màn hình radar của tàu chủ	47
2.5	Câu hỏi trong phiếu khảo sát về đánh giá những điểm có nguy cơ mất an toàn giao thông trên tuyến luồng Vũng Tàu - Sài Gòn	52
2.6	Kết quả đánh giá của hoa tiêu với 10 điểm có nguy cơ đâm va trên tuyến luồng Vũng Tàu - Sài Gòn	52
2.7	Câu hỏi thăm dò các phương pháp xác định nguy cơ đâm va trên luồng	54
2.8	Kết quả thăm dò về đánh giá nguy cơ đâm va của hoa tiêu	54
2.9	Ý kiến xác định nguy cơ đâm va theo độ chênh ETA đến 1 điểm	56
2.10	Khúc cua Hải Lý	56
2.11	Khúc cua Mũi Đèn Đỏ	57
2.12	Ngã ba Nhà Bè	57
2.13	Khúc cua mũi Ô Rơ	58
2.14	Khúc cua Đá Hàn	58
2.15	Khúc cua An Thạnh	59
2.16	Luồng Vũng Tàu - Sài Gòn	60
2.17	Trường hợp tàu điểm 2 ở cung phần tư thứ nhất so với điểm 1	61
2.18	Trường hợp điểm 2 ở cung phần tư thứ hai so với điểm 1	61
2.19	Trường hợp điểm 2 ở cung phần tư thứ ba so với điểm 1	62
2.20	Trường hợp điểm 2 ở cung phần tư thứ tư so với điểm 1	63
3.1	Độ chênh lệch giữa các output	67
3.2	Sơ đồ mô hình mạng	69
3.3	Thông số chi tiết của mô hình	70
3.4	Độ chính xác của mô hình trên tập training và validation	71
3.5	Hàm loss trên tập training và validation	72
3.6	Normalized confusion matrix trên tập test set	72
4.1	Hình ảnh dữ liệu nhiễu	75

4.2	Hình ảnh phân bố lớp của từng WP	76
4.3	Ví dụ về Good Fitting và Over Fitting	79
4.4	Minh họa Polynomial	81
4.5	Minh họa sự ảnh hưởng của alpha tới hệ số ước lượng	83
4.6	Hình ảnh giữa chuẩn hóa L1 và L2	84
4.7	F1-Score trung bình cho các WP	94
4.8	Phân bố dữ liệu giữa các WP trên tập dữ liệu lớn	95
4.9	Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với dữ liệu huấn luyện trên tập dữ liệu lớn	95
4.10	Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với dữ liệu kiểm thử trên tập dữ liệu lớn	96

MỞ ĐẦU

1. Tính cấp thiết của luận án

Hội nghị lần thứ VIII Ban Chấp hành Trung ương Đảng khóa XII đã ban hành Nghị quyết số 36-NQ/TW, ngày 22-10-2018, về chiến lược phát triển bền vững kinh tế biển Việt Nam đến năm 2030, tầm nhìn 2045 trong đó xác định Việt Nam phải trở thành quốc gia mạnh về biển, giàu từ biển, phát triển bền vững, thịnh vượng, an ninh và an toàn.

Phát triển kinh tế biển gắn liền với việc bảo đảm an toàn cho hoạt động giao thông vận tải trong đó có giao thông hàng hải, giao thông thủy nội địa. Để làm được việc đó, hoạt động quản lý, giám sát tàu biển phải được tăng cường nhằm bảo đảm an toàn hàng hải cho mỗi con tàu và đảm bảo lưu thông thông suốt qua các tuyến vận tải. Kinh tế càng phát triển, năng lực vận tải càng cần phải tăng cường. Đi cùng với nó là việc tăng cường số lượng phương tiện tham gia giao thông đặc biệt là trên các tuyến vận tải nối với các trung tâm kinh tế. Số lượng phương tiện gia tăng dẫn đến việc theo dõi, đánh giá nguy cơ đâm va, hướng dẫn di chuyển của các tàu thuyền nhằm đảm bảo an toàn càng trở nên khó khăn. Vấn đề này gây áp lực lớn đối với không chỉ các Điều hành viên VTS và Giám sát viên VTS, những người quản lý giao thông hàng hải mà còn với cả các hoa tiêu, các thuyền trưởng đang tham gia giao thông trên các tuyến luồng. Dưới áp lực cao của công việc, rất dễ xảy ra các sai sót trong quản lý điều hành, trong việc phát hiện và đánh giá nguy cơ đâm va giữa các tàu thuyền, từ đó có thể dẫn đến các tai nạn hàng hải.

Việc quản lý an toàn giao thông hàng hải được giao cho các cán bộ cảng vụ hàng hải. Ở những nơi có trạm VTS, công việc này được giao cho các Điều hành viên VTS và Giám sát viên VTS. Các cán bộ này thực thi hoạt động quản lý, điều tiết giao thông hàng hải trên các tuyến luồng ra vào các cảng biển đảm bảo an toàn hàng hải. Trong thực thi các hoạt động này, những người này phải xác định được các nguy cơ xảy ra tai nạn đâm va, các tình huống giao thông phức

tạp có thể xảy ra giữa các tàu thuyền để triển khai các hoạt động điều tiết nhằm đảm bảo an toàn cho các phương tiện, các tuyến luồng giao thông.

Để điều tiết giao thông hàng hải, đa phần các Điều hành viên VTS và Giám sát viên VTS điều tiết giao thông cần thu thập thông tin giao thông sử dụng các thiết bị như radar, camera, AIS. Căn cứ các thông tin này, các quyết định hướng dẫn, điều tiết giao thông được đưa ra nhằm đảm bảo an toàn cho giao thông hàng hải. Đây là một công việc có khối lượng rất lớn đặc biệt là ở những khu vực quản lý có mật độ tàu thuyền qua lại đông, nhiều giao cắt phức tạp. Người sỹ quan điều tiết giao thông phải nắm được sự di chuyển không chỉ của các tàu thuyền với nhau mà còn sự di chuyển qua các điểm cần thiết trên các đoạn luồng, sự di chuyển qua các bãi cạn, khu vực neo và thậm chí cả tàu thuyền đang neo đậu có bị trôi dạt hay không cũng phải nằm trong tầm kiểm soát. Mặc dù một số nơi có trang bị trạm VTS với nhiều tính năng hỗ trợ trong việc theo dõi chuyển động của các mục tiêu nhưng những tính năng này đòi hỏi phải thực hiện nhiều thao tác theo dõi, bám sát đồng thời cũng có nhiều mục tiêu không cung cấp thông tin về tuyến đường hành trình. Tuy vậy, đối với các tuyến luồng cần quản lý dài, mật độ phương tiện và giao cắt giao thông phức tạp, các vấn đề về phát hiện, xử lý thông tin giao thông sẽ càng phức tạp, gây áp lực rất lớn cho các sỹ quan quản lý giao thông. Từ đó, phát sinh nhu cầu cần có một công cụ hỗ trợ cho các sỹ quan quản lý và điều hành giao thông hàng hải nhằm đảm bảo hoạt động giao thông được diễn ra thông suốt.

Đối với các công cụ hỗ trợ hoạt động quản lý giao thông hàng hải, như đã đề cập ở trên, các thiết bị đánh giá nguy cơ mất an toàn hàng hải đều đã có. Người sỹ quan quản lý giao thông hàng hải có thể sử dụng các trang thiết bị như hệ thống radar, AIS, camera, hải đồ điện tử hoặc thậm chí hệ thống VTS hoàn chỉnh để theo dõi, giám sát và điều tiết giao thông. Mặc dù vậy, khi cần dữ liệu của phương tiện nào, đánh giá các nguy cơ đối với phương tiện đó thì đòi hỏi phải tương tác với từng phương tiện đó. Thường thì chỉ có thể theo dõi đồng thời 2 mục tiêu, muốn thêm thông tin khác thì phải chuyển sang mục tiêu khác. Một

số chức năng tự động báo động lệch đường, báo động qua điểm, báo động đi vào vùng nguy hiểm của hệ thống VTS có thể làm giảm đi áp lực công việc của người sỹ quan nhưng không có khả năng theo dõi đồng thời toàn vùng, toàn bộ các tình huống giao thông để từ đó đưa ra các cảnh báo sớm cho phương tiện. Chính vì thế, cần thiết có một công cụ nào đó có thể hỗ trợ người vận hành quản lý, nắm bắt tình hình giao thông trên toàn tuyến luồng và đưa ra khuyến cáo, báo động với người điều hành, giúp giảm tải công việc đồng thời đảm bảo an toàn khai thác tuyến luồng giao thông. Cùng với sự phát triển của khoa học công nghệ, các ứng dụng của trí tuệ nhân tạo có thể đưa vào để xử lý các bài toán giao thông này. Các tình huống giao thông đặc biệt sẽ được phân tích, định nghĩa để đưa vào cho hệ thống tự học, từ thực tiễn học tập của máy sẽ đưa ra các phương án đánh giá tình huống giao thông, phương án xử lý giao thông tối ưu trong thực tiễn. Kết quả học tập liên tục sẽ được sử dụng làm cơ sở để khuyến cáo cho người sỹ quan trong công tác điều hành, quản lý giao thông đảm bảo an toàn.

Thông thường, thông tin của các tàu thuyền do VTS hay các tàu thuyền khác thu thập được gắn liền với thông tin về vị trí của các tàu thuyền trong khu vực kiểm soát, thông tin về hệ thống luồng lạch tại khu vực và 1 số thông tin về điều kiện khí tượng thủy văn khác. Như vậy, đánh giá nguy cơ đâm va ở các tuyến luồng khác nhau sẽ khác nhau. Do vậy, cần lựa chọn khu vực để triển khai nghiên cứu cho phù hợp vừa đảm bảo tính cấp thiết mà thực tiễn đòi hỏi vừa có thể được sử dụng cho các khu vực khác sau này.

Thành phố Hồ Chí Minh là đầu tàu kinh tế của cả nước, Cục Thống kê Thành phố Hồ Chí Minh cho biết tổng sản phẩm trên địa bàn GRDP của thành phố năm 2023 ước đạt 1.621.191 tỷ đồng, tăng 5,81% so với cùng kỳ năm 2022. Để có được kết quả nêu trên, có một phần không nhỏ là có sự đóng góp của giao thông vận tải biển, vận tải thủy trên tuyến luồng Vũng Tàu - Sài Gòn. Hơn thế nữa, tuyến luồng này còn phục vụ cho các hoạt động vận tải của các địa bàn lân cận. Trong 8 năm qua, số lượng tàu thuyền (bao gồm tàu biển, phương tiện thủy

nội địa, tàu, phương tiện pha sông biển) vào, rời khu vực vùng nước cảng biển Thành phố Hồ Chí Minh như sau:

Năm	Tàu thuyền (lượt)	GT	DWT (tấn)
Năm 2016	115.762	298.566.216	375.381.388
Năm 2017	124.508	257.247.201	392.141.160
Năm 2018	115.711	273.177.659	414.174.223
Năm 2019	113.639	298.669.076	450.386.747
Năm 2020	112.695	321.462.048	471.990.196
Năm 2021	102.162	296.158.037	438.582.734
Năm 2022	113.685	315.721.967	461.007.043
Năm 2023	112.569	347.001.101	497.423.606

Với lượng hàng hóa thông qua cảng:

Năm	Hàng container (TEU)	Hàng lỏng (tấn)	Hàng khô (tấn)
Năm 2016	5.716.639	10.884.594	36.180.004
Năm 2017	5.956.810	12.143.645	36.924.903
Năm 2018	6.313.417	12.863.711	41.897.036
Năm 2019	6.587.070	12.405.673	51.152.892
Năm 2020	7.781.681	11.401.740	49.881.894
Năm 2021	7.891.104	10.966.175	54.267.537
Năm 2022	8.040.670	12.276.317	49.804.452
Năm 2023	7.908.855	12.741.317	54.155.580

(Nguồn: Cảng vụ Hàng hải Hồ Chí Minh)

Số liệu thống kê cho thấy tầm quan trọng của tuyến luồng Sài Gòn – Vũng Tàu trong việc đảm bảo kết nối giao thông của đầu tàu kinh tế với các khu vực khác đồng thời cũng cho thấy rõ sự gia tăng áp lực trong đảm bảo an toàn hàng hải trên tuyến luồng này. Ngoài ra, tuyến luồng Vũng Tàu - Sài Gòn là một tuyến luồng dài, địa hình giao thông phức tạp, kết quả nghiên cứu với tuyến đường này sẽ có nhiều khả năng áp dụng cho các tuyến luồng khác.

Xuất phát từ thực tiễn nêu trên, nghiên cứu sinh đề xuất nghiên cứu thực hiện đề tài: *Nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu lớn trong quản lý giao thông hàng hải trên tuyến luồng Vũng Tàu - Sài Gòn.*

2. Mục đích nghiên cứu của luận án

Luận án được thực hiện nhằm mục đích sau đây:

- Nghiên cứu phương pháp đánh giá nguy cơ đâm va trên vùng biển rộng và trên luồng lạch hẹp;
- Xây dựng cơ sở dữ liệu đánh giá nguy cơ đâm va giữa các tàu thuyền trên tuyến luồng Vũng Tàu - Sài Gòn;
- Ứng dụng AI trong đánh giá nguy cơ đâm va giữa các tàu thuyền trên tuyến luồng Vũng Tàu - Sài Gòn.

3. Đối tượng và phạm vi nghiên cứu của luận án

Để đạt được mục đích của luận án, tập trung nghiên cứu các đối tượng sau:

- Các tình huống giao thông hàng hải và các nguy cơ dẫn đến hoặc được đánh giá là có thể dẫn đến tai nạn hàng hải;
- Đánh giá của hoa tiêu dẫn tàu đối với các tình huống, chỉ số mà căn cứ vào đó hoa tiêu xác định tồn tại nguy cơ đâm va giữa các tàu thuyền chạy trên tuyến luồng Vũng Tàu - Sài Gòn;
- Hoạt động điều tiết giao thông của sỹ quan hàng hải tại Việt Nam.

Để triển khai các hoạt động nghiên cứu, đề tài tập trung triển khai nghiên cứu các đối tượng liên quan đến trạm VTS của Việt Nam do Cảng vụ Hàng hải Thành phố Hồ Chí Minh quản lý.

4. Phương pháp nghiên cứu của luận án

Nghiên cứu sinh sử dụng các phương pháp nghiên cứu sau đây:

- Phương pháp phân tích số liệu để đánh giá các tình huống giao thông;
- Phương pháp chuyên gia phục vụ nghiên cứu đánh giá các tình huống mất an toàn giao thông, phương pháp đánh giá nguy cơ mất an toàn và xử lý các tình huống giao thông;

- Phương pháp thực nghiệm học máy với cơ sở dữ liệu tình huống giao thông nhằm đánh giá nguy cơ đâm va giữa các tàu thuyền đang hành trình trên tuyến luồng Vũng Tàu - Sài Gòn; tự động phát hiện nguy cơ đâm va và cảnh báo cho hoạt động quản lý giao thông hàng hải.

5. Ý nghĩa khoa học, ý nghĩa thực tiễn của luận án

Đề tài hoàn thành sẽ xây dựng nên cơ sở lý thuyết nhận diện được các tình huống xảy ra các nguy cơ tai nạn đâm va giữa các tàu thuyền trong hàng hải. Cơ sở này có thể sử dụng trong các nghiên cứu khác liên quan đến xử lý nguy cơ đâm va đảm bảo an toàn giao thông hàng hải không chỉ trên biển rộng mà cả trong phạm vi giới hạn của luồng lạch hẹp.

Việc xây dựng thành công hệ học máy sẽ làm tăng cường khả năng phát hiện cảnh báo nguy cơ xảy ra tai nạn trong hàng hải liên quan đến đâm va giữa các tàu và giữa tàu với các mục tiêu cố định, phát hiện và cảnh báo các điểm nóng về giao thông góp phần giảm tải công việc cho sỹ quan quản lý giao thông hàng hải, nâng cao khả năng an toàn trong khai thác cảng biển, góp phần đảm bảo thực hiện các mục tiêu kinh tế, xã hội.

6. Các kết quả mới đạt được của luận án

Đã xây dựng được phương án đánh giá nguy cơ đâm va giữa các tàu thuyền hoạt động trong luồng lạch hẹp nói chung và luồng Vũng Tàu - Sài Gòn nói riêng theo ETA của các tàu thuyền đến các điểm nóng giao thông.

Kiểm định kết quả nghiên cứu thông qua huấn luyện mô hình AI cho thấy độ chính xác của mô hình cao. Khẳng định phương pháp nghiên cứu và thuật toán đề xuất hoàn toàn phù hợp với thực tiễn và có ý nghĩa khoa học.

CHƯƠNG 1. NGHIÊN CỨU CƠ SỞ LÝ THUYẾT NHẬN DẠNG, ĐÁNH GIÁ NGUY CƠ ĐÂM VA GIỮA CÁC THUYỀN TRÊN TUYẾN LUỒNG VĨNG TÀU - SÀI GÒN

1.1. Tổng quan về vấn đề nghiên cứu

Hội nghị lần thứ VIII Ban Chấp hành Trung ương Đảng khóa XII đã ban hành Nghị quyết số 36-NQ/TW, ngày 22-10-2018 về chiến lược phát triển bền vững kinh tế biển Việt Nam đến năm 2030, tầm nhìn 2045 trong đó xác định Việt Nam phải trở thành quốc gia mạnh về biển, giàu từ biển, phát triển bền vững, thịnh vượng, an ninh và an toàn. Để làm được việc đó, hoạt động quản lý, giám sát tàu biển phải được tăng cường nhằm bảo đảm an toàn hàng hải cho mỗi con tàu và đảm bảo lưu thông thông suốt qua các tuyến vận tải [1].

Việc quản lý an toàn giao thông hàng hải được giao cho các cán bộ cảng vụ hàng hải. Các cán bộ này thực thi hoạt động quản lý, điều tiết giao thông hàng hải trên các tuyến luồng ra vào các cảng biển đảm bảo an toàn hàng hải. Trong thực thi các hoạt động này, người cán bộ phải xác định được các nguy cơ xảy ra tai nạn đâm va, các tình huống giao thông phức tạp có thể xảy ra giữa các tàu thuyền để triển khai các hoạt động điều tiết nhằm đảm bảo an toàn cho các phương tiện, các tuyến luồng giao thông.

Để điều tiết giao thông hàng hải, đa phần các sỹ quan điều tiết giao thông cần thu thập thông tin giao thông sử dụng các thiết bị như radar, camera, AIS. Căn cứ các thông tin này, các quyết định hướng dẫn, điều tiết giao thông được đưa ra nhằm đảm bảo an toàn cho giao thông hàng hải. Đây là một công việc có khối lượng rất lớn đặc biệt là ở những khu vực quản lý có mật độ tàu thuyền qua lại đông, nhiều giao cắt phức tạp. Người sỹ quan điều tiết giao thông phải nắm được sự di chuyển không chỉ của các tàu thuyền với nhau mà còn sự di chuyển qua các điểm cần thiết trên các đoạn luồng, sự di chuyển qua các bãi cạn, khu vực neo và thậm chí cả tàu thuyền đang neo đậu có bị trôi dạt hay không cũng phải nằm trong tầm kiểm soát. Mặc dù một số nơi có trang bị trạm VTS [2], [3]

với nhiều tính năng hỗ trợ trong việc theo dõi chuyển động của các mục tiêu nhưng những tính năng này đòi hỏi phải thực hiện nhiều thao tác theo dõi, bám sát đồng thời cũng có nhiều mục tiêu không cung cấp thông tin về tuyến đường hành trình. Tuy vậy, đối với các tuyến luồng cần quản lý dài, mật độ phương tiện và giao cắt giao thông phức tạp, các vấn đề về phát hiện, xử lý thông tin giao thông sẽ càng phức tạp, gây áp lực rất lớn cho các sỹ quan quản lý giao thông. Từ đó, phát sinh nhu cầu cần có một công cụ hỗ trợ cho các sỹ quan quản lý và điều hành giao thông hàng hải nhằm đảm bảo hoạt động giao thông được diễn ra thông suốt.

Đối với các công cụ hỗ trợ hoạt động quản lý giao thông hàng hải, như đã đề cập ở trên, các thiết bị đánh giá nguy cơ mất an toàn hàng hải đều đã có. Người sỹ quan quản lý giao thông hàng hải có thể sử dụng các trang thiết bị như hệ thống radar, AIS, camera, hải đồ điện tử hoặc thậm chí hệ thống VTS hoàn chỉnh để theo dõi, giám sát và điều tiết giao thông. Mặc dù vậy, khi cần dữ liệu của phương tiện nào, đánh giá các nguy cơ đối với phương tiện đó thì đòi hỏi phải tương tác với từng phương tiện đó. Thường thì chỉ có thể theo dõi đồng thời 2 mục tiêu, muốn thêm thông tin khác thì phải chuyển sang mục tiêu khác. Một số chức năng tự động báo động lệch đường, báo động qua điểm, báo động đi vào vùng nguy hiểm của hệ thống VTS có thể làm giảm đi áp lực công việc của người sỹ quan nhưng không có khả năng theo dõi đồng thời toàn vùng, toàn bộ các tình huống giao thông để từ đó đưa ra các cảnh báo sớm cho phương tiện. Chính vì thế, cần thiết có một công cụ nào đó có thể hỗ trợ người vận hành quản lý, nắm bắt tình hình giao thông trên toàn tuyến luồng và đưa ra khuyến cáo, báo động với người điều hành, giúp giảm tải công việc đồng thời đảm bảo an toàn khai thác tuyến luồng giao thông.

Điều tiết giao thông trên một tuyến luồng dựa trên các thông tin chuyển động của các tàu thuyền và đưa ra kết quả đánh giá về nguy cơ đâm va dựa trên kinh nghiệm của sỹ quan VTS. Đánh giá này phần nhiều mang theo ý kiến chủ quan của người sỹ quan. Do đó, kết quả thu được sẽ khác nhau giữa các sỹ quan

VTS và khác với những hoa tiêu, thuyền trưởng của tàu thuyền đang di chuyển trên tuyến luồng. Xuất phát từ các đánh giá khác nhau đó, kết quả xác định nguy cơ đâm va sẽ khác nhau. Để có thể đưa ra đánh giá tốt nhất, theo xu thế chung nhất về nguy cơ đâm va, cần có một cơ sở dữ liệu đủ lớn để kiểm tra, đánh giá mối liên hệ giữa chuyển động của các tàu thuyền và nguy cơ đâm va giữa chúng. Trên cơ sở dữ liệu lớn và quan điểm về nguy cơ đâm va giữa các tàu thuyền của các chuyên gia dẫn tàu trên tuyến luồng, mối quan hệ giữa sự di chuyển của các tàu thuyền trên luồng và nguy cơ đâm va giữa chúng sẽ được làm rõ.

Cùng với sự phát triển của khoa học công nghệ, trí tuệ nhân tạo nói chung [12-16], và học máy đã được chứng minh là các tiếp cận hiệu quả [39-112] cho các bài toán giao thông đường thủy này. Các tình huống giao thông đặc biệt sẽ được phân tích, định nghĩa để đưa vào cho hệ thống tự học, từ thực tiễn học tập của máy sẽ đưa ra các phương án đánh giá tình huống giao thông, phương án xử lý giao thông tối ưu trong thực tiễn. Kết quả học tập liên tục sẽ được sử dụng làm cơ sở để khuyến cáo cho người sỹ quan trong công tác điều hành, quản lý giao thông đảm bảo an toàn.

Xuất phát từ thực tiễn nêu trên, nghiên cứu sinh đề xuất nghiên cứu thực hiện đề tài: *Nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu lớn trong quản lý giao thông hàng hải trên tuyến luồng Vũng Tàu - Sài Gòn.*

Liên quan đến nội dung nghiên cứu, qua phân tích cho thấy đây là một lĩnh vực tương đối mới. Xem xét các công trình nghiên cứu đã công bố hiện nay ở Việt Nam và trên thế giới về lĩnh vực này ta có thể thấy lĩnh vực này, có thể thấy những vấn đề mà đề tài dự kiến thực hiện, chưa từng có công trình nào tương tự như vậy. Cụ thể như sau:

Nghiên cứu của tác giả Phạm Tú Anh: “*Bố trí VTS cho hệ thống cảng thành phố Hồ Chí Minh trong tương lai*” [2] chuyên ngành Công trình thủy thực hiện năm 2003. Đề tài nghiên cứu và phân tích các vấn đề bao gồm: giới thiệu về chức năng, nhiệm vụ và nguyên lý hoạt động của trung tâm VTS; giới thiệu về các trang thiết bị máy móc được lắp đặt tại các trạm kiểm soát giao thông tại khu

vực luồng vào cảng thành phố Hồ Chí Minh; nghiên cứu lắp đặt các trang thiết bị của hệ thống. Trong phần này dựa vào điều kiện địa hình, các cơ sở vật chất có sẵn tại cảng và khu vực luồng ra vào mà tác giả nghiên cứu việc bố trí lắp đặt các trang thiết bị của hệ thống như cột anten Radar, các Camera theo dõi hay các khu vực lắp đặt hệ thống theo dõi thủy triều... Đề tài chưa đề cập đến vấn đề tự động đánh giá nguy cơ đâm va.

Đối với các nghiên cứu đánh giá nguy cơ đâm va thì đa phần là các nghiên cứu đánh giá nguy cơ đâm va trên quan điểm giữa một tàu chủ với 1 tàu mục tiêu chứ không trên quan điểm đánh giá nguy cơ đâm va của 2 đối tượng tàu biển của trạm VTS. Các công trình này thường vận dụng cơ sở lý thuyết là phương pháp đồ giải tránh va tương đối (hình 4) để đánh giá nguy cơ đâm va giữa tàu chủ với tàu mục tiêu. Theo phương pháp này, cơ sở để đánh giá nguy cơ đâm va là chuyển động tương đối của tàu mục tiêu so với tàu chủ. Có thể liệt kê một vài công trình liên quan đến những nghiên cứu dạng này như:

Đề tài nghiên cứu khoa học cấp Bộ “*Nghiên cứu xây dựng mô hình mô phỏng tránh đâm va tàu thuyền trên biển để cảnh báo, trợ giúp tránh va tàu thuyền trên biển áp dụng cho tàu thuyền nhỏ*” [4] năm 2011 và đề tài “*Nghiên cứu chế tạo thiết bị hỗ trợ tự động đồ giải các thông số mục tiêu (có kết nối radar) phòng tránh đâm va tàu thuyền trên biển dùng cho các tàu thuyền nhỏ*” [5] năm 2013 do PGS.TS. Phạm Văn Cương là các đề tài sử dụng nguyên lý đồ giải tránh va tương đối để giải quyết vấn đề xác định nguy cơ đâm va giữa tàu mục tiêu với tàu chủ.

Luận án tiến sĩ “*Xây dựng mô hình đánh giá rủi ro tai nạn đâm va tại cảng biển Vũng Tàu, ứng dụng phòng ngừa tai nạn và hỗ trợ công tác cải tạo, nâng cấp luồng hàng hải*” [17] năm 2022 do NCS Lê Văn Thức nghiên cứu đã xây dựng và sử dụng Phần mềm mô phỏng nhanh (fast time simulation) phục vụ mô phỏng, đánh giá trực quan nguy cơ đâm va và nhằm ứng dụng trong công tác đánh giá rủi ro, phòng ngừa tai nạn đâm va và hỗ trợ công tác nghiên cứu cải tạo, nâng cấp luồng hàng hải tại cảng biển Vũng Tàu. Phương pháp và kết quả

nghiên cứu của Luận án có thể ứng dụng để tiếp tục nghiên cứu, phát triển mô hình thích hợp cho các cảng biển khác tại Việt Nam. Tuy nhiên, các kết quả cần đáp ứng theo thời gian thực và cần có hệ thống can thiệp, sử lý dữ liệu lớn do hệ thống VTS lưu trữ.

Liên quan đến các ứng dụng của nguyên lý đồ giải tránh va tương đối chúng ta còn có thể kể đến một thiết bị rất quan trọng trong hàng hải: đó là thiết bị Radar/ARPA. Thiết bị tự động đồ giải tránh va radar đã từ lâu xuất hiện trên thị trường trang thiết bị Hàng hải. Từ những năm 70 đến 80 của thế kỷ trước, với sự phát triển của vi xử lý và kỹ thuật máy tính người ta đã ứng dụng các tiến bộ này vào việc nâng cao hiệu quả khai thác hoạt động của radar. Chiếc ARPA thương mại đầu tiên do Norcontrol (nay là một phần của hãng Kongsberg Maritime) đã được bàn giao cho tàu chở hàng chuyên tuyến Taimyr năm 1969. Cho đến hiện nay, ARPA đã xuất hiện trên rất nhiều phương tiện vận tải thủy.

Là một thiết bị yêu cầu bắt buộc phải trang bị trên các tàu theo qui định của SOLAS, ARPA được IMO yêu cầu có chức năng được sử dụng để: cải thiện tiêu chuẩn tránh va trên biển bằng cách giảm tải công việc của người sử dụng bằng cách cho phép họ có thể tự động thu thập các thông tin với nhiều mục tiêu thông qua việc lựa chọn đồ giải bằng tay từng mục tiêu đơn lẻ. Căn cứ trên yêu cầu đó cộng với sự phát triển về mặt công nghệ của riêng mình, các hãng sản xuất thiết bị Hàng hải đã chế tạo thành công và đưa sản phẩm của họ lên thị trường. Khi đề cập đến thiết bị này, người ta thường nói đến những cái tên quen thuộc như Kongsberg, Furuno, Koden, Icom, JRC... Đối với hệ thống radar của các trạm VTS hiện nay cũng sử dụng nguyên lý đồ giải tương đối như vậy nhằm hỗ trợ đánh giá nguy cơ đâm va.

Đối với hệ thống VTS hiện hành, các đánh giá nguy cơ đâm va có thể dựa vào các quan sát trực quan với các vector chuyển động của các mục tiêu từ đó người sỹ quan đưa ra các phán đoán về khả năng gặp nhau của các tàu thuyền tại các khu vực có khả năng dẫn đến nguy cơ đâm va, khả năng đi lệch đường của các mục tiêu hay khả năng mục tiêu di chuyển vào các khu vực nguy hiểm. Việc

đánh giá trên các căn cứ này có độ chính xác không cao. Nếu có thể xây dựng được phương pháp đánh giá nguy cơ đâm va trực tiếp giữa 2 mục tiêu thì sẽ tốt hơn phương pháp hiện nay.

Bên cạnh đánh giá nguy cơ đâm va giữa các tàu thuyền sử dụng thông tin do radar cung cấp, cũng có nhiều nghiên cứu sử dụng thông tin từ AIS [2-33]. Tuy vậy, việc đánh giá nguy cơ đâm va từ thông tin vị trí tàu do AIS cung cấp sẽ bỏ lọt các tàu thuyền nhỏ không trang bị AIS. Ngoài ra, còn một số nghiên cứu sử dụng thuật toán phân cụm và lý thuyết hàm tin cậy để đánh giá rủi ro xảy ra đâm va trên các khu vực cố định; một số tập trung đánh giá rủi ro đâm va theo vùng bao quanh các tàu thuyền. Một số nghiên cứu còn bổ sung thêm đánh giá nguy cơ đâm va theo các điều kiện khí tượng thủy văn khác... Những nghiên cứu này đã đề xuất các lý thuyết đánh giá nguy cơ đâm va và được kiểm nghiệm trên kết quả mô phỏng. Tuy vậy, việc triển khai trong thực tiễn vẫn còn những hạn chế do người sử dụng cần xử lý nhiều thông tin trước khi có thể đưa ra đánh giá về nguy cơ đâm va. Trong thực tiễn dẫn tàu, các hoa tiêu và thuyền trưởng có các phương pháp đánh giá nguy cơ đâm va khác đơn giản và hiệu quả hơn.

Đối với các nghiên cứu sử dụng trí tuệ nhân tạo trong phòng tránh nguy cơ đâm va hiện không có nhiều công trình nghiên cứu. Năm 2019, hãng Fujitsu đã công bố những thông tin ban đầu về nghiên cứu tính toán nguy cơ đâm va và dự đoán các điểm nóng trong kiểm soát giao thông hàng hải. Trong các thông báo của mình, Fujitsu giới thiệu ứng dụng công nghệ có tên Human Centric AI Zinrai [21] trong xác định nguy cơ đâm va và dự đoán các vùng mà có nguy cơ đâm va tập trung. Tuy nhiên, các thông tin này chỉ chung chung và chưa công bố thông tin về cơ sở lý thuyết để có thể ứng dụng trong các công trình nghiên cứu khác. Đặc biệt, đối với việc kiểm soát giao thông hàng hải ở Việt Nam thì mảng nghiên cứu này vẫn còn bỏ ngỏ.

Các nghiên cứu đánh giá nguy cơ đâm va, rủi ro đâm va phần nhiều được giới hạn trên các vùng biển rộng, ít có các nghiên cứu trong các khu vực luồng lạch hẹp. Bên cạnh đó, việc khó được tiếp cận với thông tin di chuyển của tàu

thuyền tại các trạm VTS, đánh giá nguy cơ đâm va trên hệ thống luồng lạch tại Việt Nam phức tạp, hạn chế về cơ sở vật chất phục vụ nghiên cứu... cũng là rào cản không nhỏ đối với các nhà nghiên cứu. Do đó, việc nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu lớn trong xử lý, xác định nguy cơ đâm va giữa các tàu thuyền để từ đó hỗ trợ các sỹ quan quản lý giao thông hàng hải hiện nay trong việc duy trì, đảm bảo an toàn giao thông hàng hải là một hướng đi mới, mang tính thực tiễn cao tại Việt Nam.

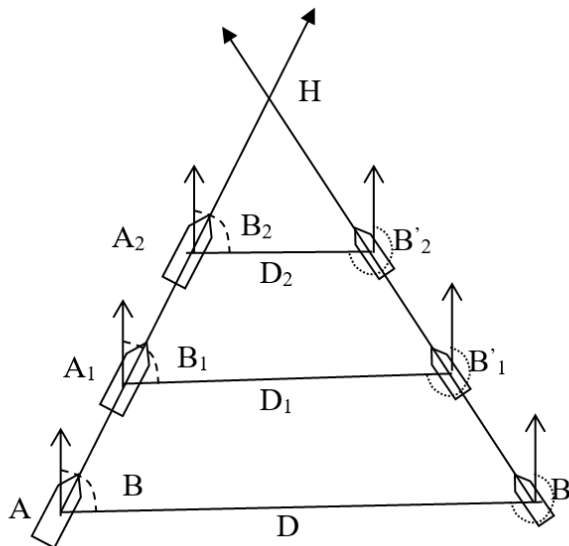
1.2. Đánh giá nguy cơ đâm va trong hàng hải và trên tuyến luồng Vũng Tàu - Sài Gòn

Theo khoản d Điều 7 Colreg72, đánh giá nguy cơ đâm va giữa các tàu thuyền, căn cứ trên 2 yếu tố phương vị và khoảng cách [7], [8], [19]:

“i. Có nguy cơ đâm va, khi phương vị la bàn của tàu thuyền đang đến gần không thay đổi rõ rệt.

ii. Đôi khi nguy cơ đâm va vẫn có thể xảy ra ngay cả khi quan sát thấy phương vị thay đổi rõ rệt, đặc biệt là khi đến gần một tàu rất lớn hoặc một đoàn tàu lai hay một tàu thuyền khác ở khoảng cách gần”.

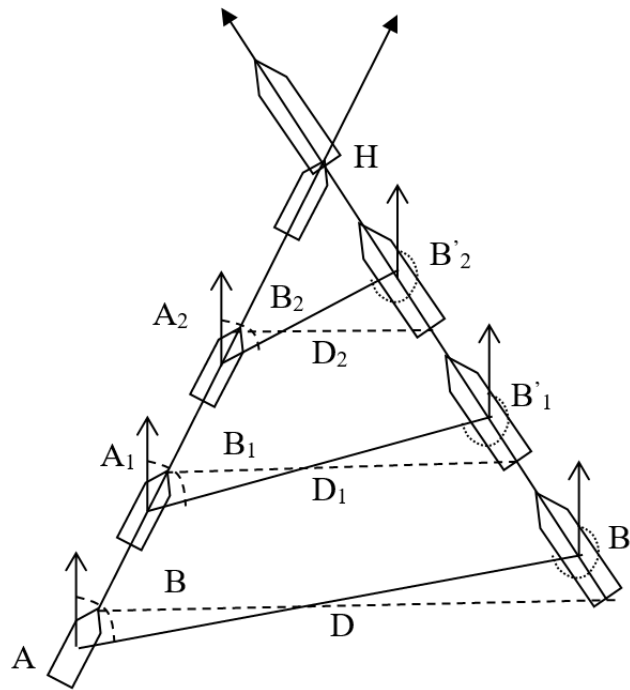
Theo khoản d Điều 7, nguy cơ đâm va giữa 2 tàu xảy ra khi khoảng cách giảm dần và phương vị không đổi hoặc thay đổi không rõ rệt hay có thể thay đổi khi đến gần một tàu lớn, một đoàn lai. Hai điều kiện trên phải diễn ra đồng thời thì mới xảy ra nguy cơ đâm va. Để minh họa cho nguy cơ đâm va giữa các tàu thuyền chúng ta xem xét các hình vẽ dưới đây [7]:



Hình 1.1. Tồn tại nguy cơ đâm va khi phương vị không đổi, khoảng cách giảm dần

Theo hình 1.1, khi phương vị giữa 2 tàu không đổi ($B = B_1 = B_2$), khoảng cách giữa 2 tàu giảm dần ($D > D_1 > D_2$) thì hai điểm đo phương vị, khoảng cách sẽ gặp nhau tại điểm H. Hai tàu sẽ đâm va với nhau. Trong đó B, B₁, B₂ là phương vị của tàu B quan sát tại tàu A; B', B'₁, B'₂, là phương vị của tàu A khi quan sát tại tàu B tương ứng với thời điểm quan sát lần 1, 2, 3.

Theo hình 1.2, khi phương vị giữa 2 tàu thay đổi không rõ rệt ($B \approx B_1 \approx B_2$), khoảng cách giữa 2 tàu giảm dần ($D > D_1 > D_2$) thì hai điểm đo phương vị, khoảng cách sẽ không gặp nhau tại điểm H. Tuy nhiên, trong trường hợp này, nếu đo từ điểm khác trên tàu tới điểm khác của tàu mục tiêu (ví dụ đo từ mũi tàu A đến lái tàu B) thì phương vị sẽ không đổi, khoảng cách giảm dần. Thỏa mãn điều kiện như mục ii, khoản d Điều 7 đề cập. Mũi tàu A sẽ gặp lái tàu B tại điểm H – tồn tại nguy cơ đâm va giữa hai tàu.

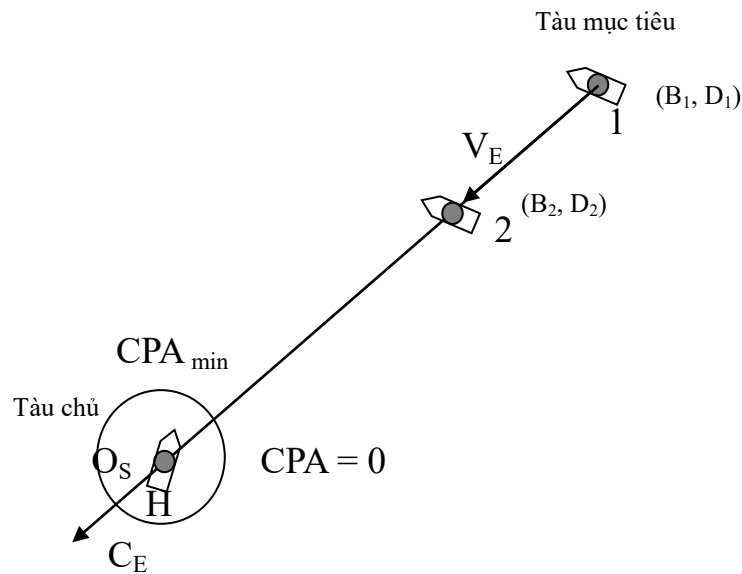


Hình 1.2. Tôn tại nguy cơ đâm va khi phương vị thay đổi không rõ rệt, khoảng cách giảm dần

Trong thực tiễn, hai tàu không thể tránh nhau ở khoảng cách quá gần, cần có khoảng cách an toàn giữa chúng để đề phòng các trường hợp sai số trong xác định khoảng cách tiếp cận, sai lệch do điều khiển, tác động tương hỗ giữa các tàu, giữa tàu với bờ hay các chướng ngại hàng hải khác, tác động của các điều kiện ngoại cảnh đến việc điều khiển con tàu, khả năng điều khiển của con tàu ... Khoảng cách an toàn đó được gọi là CPA_{min} hay khoảng cách tiếp cận gần nhất tối thiểu của tàu thuyền mục tiêu. Khi hai tàu thuyền tiến đến gần nhau ở khoảng cách nhỏ hơn CPA_{min} thì được đánh giá có tồn tại nguy cơ đâm va giữa các tàu đó.

Đánh giá nguy cơ đâm va theo hình 1.1 và 1.2 là phương pháp đánh giá theo chuyển động thực của hai tàu. Trong quá trình chuyển động, người ta liên tục đo phương vị và khoảng cách tới mục tiêu. Nếu xem xét thấy kết quả đo thỏa mãn điều kiện đã nêu thì đánh giá là có nguy cơ đâm va giữa các tàu thuyền.

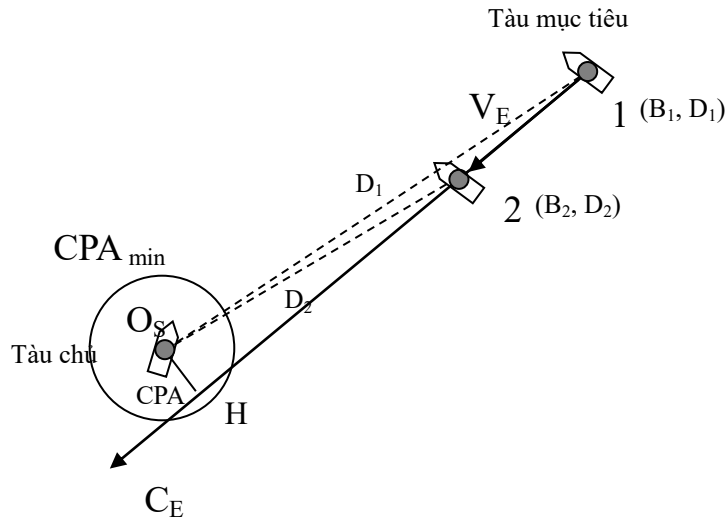
Bên cạnh sử dụng phương pháp đánh giá nguy cơ đâm va theo chuyển động thực người ta đánh giá theo chuyển động tương đối giữa các tàu. Phương pháp này thường được sử dụng trong đồ giải tránh va radar.



Hình 1.3. Trường hợp phương vị không đổi, khoảng cách giảm dần

Theo phương pháp này, người ta phân định ra tàu chủ và tàu mục tiêu. Tàu chủ là tàu thực hiện công tác quan sát đo phương vị, khoảng cách tới tàu mục tiêu. Trong chuyển động tương đối, tàu chủ được đánh giá đứng yên còn tàu mục tiêu chuyển động tương đối so với nó. Chuyển động tương đối là sự kết hợp chuyển động của tàu mục tiêu với tàu chủ. Hình 1.3 và 1.4 biểu thị mối tương quan giữa 2 tàu trong 2 trường hợp phương vị không đổi và phương vị có thay đổi kết hợp với khoảng cách giảm dần.

Sau 2 lần quan sát người ta xác định được phương vị và khoảng cách tương ứng là (B_1, D_1) và (B_2, D_2) tại các thời điểm tương ứng là t_1 và t_2 . Lấy điểm O (tàu chủ) làm tâm, thao tác các đường phương vị B_1 và B_2 , trên đó lấy các điểm 1 và 2 ứng với các khoảng cách D_1 và D_2 . Đường 1-2 kéo dài sẽ là đường chuyển động tương đối C_E của tàu mục tiêu. Từ O hạ đường vuông góc với C_E tại điểm H. OH là khoảng cách ngắn nhất các tàu đi qua nhau hay còn gọi là khoảng cách tiếp cận gần nhất CPA.



Hình 1.4. Trường hợp phương vị thay đổi, khoảng cách giảm dần

Ta thấy rằng sự thay đổi về phương vị quan sát dẫn đến sự thay đổi của giá trị CPA. Khi phương vị không đổi, CPA = 0. Khi phương vị thay đổi chậm, CPA nhỏ. Khi phương vị thay đổi nhanh, CPA sẽ lớn.

Nếu $CPA \geq CPA_{min}$, hai tàu đi qua nhau ở khoảng cách ngoài khoảng cách an toàn. Không tồn tại nguy cơ đâm va.

$CPA < CPA_{min}$ và $D_2 < D_1$ hai tàu đi qua nhau ở khoảng cách nhỏ hơn khoảng cách an toàn. Có thể xảy ra nguy cơ đâm va.

Như vậy, yếu tố phương vị không thay đổi hoặc thay đổi không rõ ràng theo như khoản d Điều 7 sẽ tương ứng với yếu tố là $CPA < CPA_{min}$ khi đề giải tránh va tương đối.

Khi phương vị không đổi, điểm H trùng với O hay nói cách khác, C_E đi qua vị trí tàu chủ (hình 1.3). Khi phương vị có thay đổi, điểm H không trùng với điểm O (hình 1.4). Trên đường C_E mục tiêu sẽ chuyển động với tốc độ tương đối V_E . Nếu tàu mục tiêu:

$$V_E = \frac{AB}{t_2 - t_1} \quad (1.1)$$

Thời gian mục tiêu đến điểm tiếp cận gần nhất được gọi là T_{CPA} . T_{CPA} được qui ước tính từ thời điểm quan sát lần cuối đến khi tàu mục tiêu qua cận điểm.

$$T_{CPA} = \frac{BH}{V_E} \quad (1.2)$$

Nếu tàu chưa đi qua cận điểm, khoảng cách giữa hai tàu giảm xuống, tàu mục tiêu di chuyển với tốc độ tương đối hướng tới cận điểm. V_E mang dấu dương. Ngược lại, V_E mang dấu âm. Tương ứng với điều đó, nếu hai tàu tiến lại gần nhau, T_{CPA} mang dấu dương. Nếu 2 tàu đi xa nhau, T_{CPA} mang dấu âm. Điều kiện về dấu của T_{CPA} sẽ tương ứng với điều kiện khoảng cách giữa 2 tàu.

Nói tóm lại, khi sử dụng radar với phương pháp đồ giải tương đối, nguy cơ đâm va sẽ xảy ra khi đồng thời thỏa mãn 2 điều kiện: $CPA < CPA_{min}$ và $T_{CPA} > 0$.

Khi các tàu tiến đến gần nhau trên các đoạn luồng lạch hẹp mà độ rộng không đủ để có thể đi qua nhau ở khoảng cách an toàn ($CPA < CPA_{min}$) hoặc tại khu vực vận động khó khăn dẫn đến khó có khả năng điều khiển chúng qua nhau ở khoảng cách an toàn thì nguy cơ đâm va sẽ xảy ra.

Về nguyên tắc có thể dựa vào 2 thông số CPA và T_{CPA} để đánh giá nguy cơ đâm va trên các đoạn luồng hẹp này. Tuy nhiên, để tàu thuyền gặp nhau ở những đoạn hẹp như vậy thì cho dù có đánh giá được nguy cơ đâm va nhưng cũng không thể thực hiện hành động tránh đâm va như mong muốn vì không có đủ không gian an toàn để điều động. Đó là chưa kể khi hành trình trên các đoạn luồng lạch hẹp, các tàu luôn thay đổi hướng đi, tốc độ theo điều kiện luồng lạch, việc đánh giá, phát hiện nguy cơ đâm va theo cách đồ giải radar sẽ không thực sự tin cậy để thực hiện hành động tránh đâm va có hiệu quả. Việc loại trừ nguy cơ đâm va thay bằng các đánh giá theo phương pháp đồ giải thông thường được mà sử dụng các biện pháp khác. Thông thường, đối với các tàu lớn chỉ có thể hành hải an toàn trên phạm vi giới hạn của luồng lạch hẹp, người ta thường xác định các khu vực hạn chế các tàu thuyền gặp nhau và thực hiện hành động đánh giá nguy cơ đâm va thông qua giá trị ETA (Estimated Time of Arrival: dự kiến thời gian tàu đến điểm nóng giao thông) tới các khu vực này. Nếu ETA của 2 tàu đến khu vực nguy hiểm gần với nhau thì người ta sẽ thực hiện các biện pháp phòng ngừa bằng cách điều chỉnh ETA của các tàu đến các khu vực đó để cho 2 tàu không thể gặp nhau lúc đi qua các khu vực hạn chế đó. Đối với hành trình của các tàu lớn, người ta có thể phát thông báo ETA qua các điểm đặc trưng

thuộc luồng lạch hẹp hoặc hệ thống phân luồng thông qua các bản tin an toàn hàng hải để các tàu thuyền trong khu vực biết mà hành trình thận trọng.

Tuyến luồng Vũng Tàu - Sài Gòn là một tuyến giao thông vận tải huyết mạch nối Thành phố Hồ Chí Minh với các khu vực lân cận. Hiện nay, tuyến luồng Vũng Tàu - Sài Gòn được sử dụng khai thác cho các tàu container ra, vào cảng Tân Cảng Cát Lái có chiều dài lên đến 222 m, mớn nước đến 11,0 m và các tàu hàng rời có trọng tải toàn phần đến 60.000 DWT (giảm tải).

Luồng Vũng Tàu - Sài Gòn bắt đầu từ phao số “0” Vịnh Gành Rái dẫn vào cảng Sài Gòn qua các sông: Sông Ngã Bảy, sông Lòng Tàu, sông Nhà bè và sông Sài Gòn. Các thông số kỹ thuật cơ bản của luồng như sau:

- Cao độ đáy -8,5m (hệ Hải đồ);
- Chiều rộng luồng 150m;
- Chiều dài toàn tuyến 84.5 km;
- Bán kính cong nhỏ nhất khoảng $R_{\min} = 410\text{m}$ tại khúc cua Mũi Đèn Đỏ.

Vị trí các đoạn cạn:

- Đoạn cạn vịnh Gành Rái: Từ hạ lưu phao « 4B » -1100m đến thượng lưu phao « 6A » +700m;
- Đoạn cạn Dàn Xây: Từ hạ lưu phao « 29 » -400m đến thượng lưu phao « 26 » +400m;
- Đoạn cạn Kervella: Đoạn từ hạ lưu phao « 28 » -400m đến phao « 33 »;
- Đoạn cạn Mũi L'est: Đoạn từ hạ lưu phao « 34 » -1100m đến phao « 37A »;
- Đoạn cạn Prôpntis: Từ phao « 37A » đến thượng lưu phao « 44 » +750m;
- Đoạn cạn Đá Hàn: Từ hạ lưu phao « 46 » -100m đến thượng lưu phao « 43 » +300m;
- Đoạn cạn Navioil: Từ phao « 67 » đến thượng lưu phao « 69 » +850m;
- Đoạn cạn 71 – 83: Từ hạ lưu phao « 71 » -850m đến thượng lưu đăng tiêu « 83 » +1300m.

Liên quan đến tốc độ chạy tàu, theo Điều 14 Nội quy Cảng biển Thành phố Hồ Chí Minh quy định tàu thuyền hành trình với tốc độ giới hạn ở một số đoạn luồng cụ thể:

- Khu vực Thiêng Liêng: nếu có tàu neo đậu, buộc phao tốc độ không vượt quá 12 hải lý/giờ.

- Sông Lòng Tàu từ ngã ba sông Đồng Tranh đến mũi Pha Mi: tốc độ không vượt quá 12 hải lý/giờ.

- Sông Nhà Bè, sông Đồng Nai từ ngã ba Bình Khánh đến ngã ba Tắc Thầy Bẫy: tốc độ không vượt quá 10 hải lý/giờ.

- Sông Sài Gòn từ mũi Đèn Đỏ đến hạ lưu cầu K15C bên cảng Bến Nghé: tốc độ không vượt quá 07 hải lý/giờ.

- Sông Sài Gòn từ cầu K15C – cảng Bến Nghé đến rạch Thị Nghé: tốc độ không vượt quá 06 hải lý/giờ.

- Sông Soài Rạp từ khu vực quay trở cảng SPCT đến ngã ba Bình Khánh: tốc độ không vượt quá 10 hải lý/giờ.

- Sông Soài Rạp từ cảng SPCT đến ngã ba Vàm Cỏ: tốc độ không vượt quá 12 hải lý/giờ.

Hạn chế tốc độ quy định tại Điều này không áp dụng đối với các tàu công vụ, tàu cứu hỏa, tàu cứu nạn đang làm nhiệm vụ và các tàu thuyền cao tốc khác có thiết kế đặc biệt chống tạo sóng.

Liên quan đến chế độ hoa tiêu, theo Quyết định số 548/QĐ-CHHVN ngày 28/4/2017 của Cục Hàng hải Việt Nam về việc công bố tuyến dẫn tàu trong vùng hoa tiêu hàng hải bắt buộc của Việt Nam và quy định số lượng tối thiểu hoa tiêu các hạng, phương tiện đưa, đón hoa tiêu đối với từng tuyến dẫn tàu quy định:

Vùng hoa tiêu bắt buộc thuộc địa phận các tỉnh Đồng Nai, Bình Dương, Long An và Thành phố Hồ Chí Minh sử dụng 02 tuyến luồng hàng hải (Vũng Tàu - Sài Gòn và Soài Rạp) với 03 tuyến dẫn tàu như sau như sau:

- Tuyến dẫn tàu Vũng Tàu - Sài Gòn, Đồng Nai từ vùng đón trả hoa tiêu đến các bến cảng khu vực Thành phố Hồ Chí Minh, các tỉnh Đồng Nai, Bình Dương, Long An (trừ bến cảng Tân Cảng Cát Lái).

- Tuyến dẫn tàu sông Soài Rạp Từ vùng đón trả hoa tiêu đến các bến cảng trên sông Soài Rạp (ngang ngã ba Hiệp Phước Nhà Bè).

- Tuyến dẫn tàu bến cảng Tân Cảng Cát Lái Từ vùng đón trả hoa tiêu đến bến cảng Tân Cảng Cát Lái.

Để quản lý giao thông trên tuyến luồng này, từ tháng 05/2012 cơ quan quản lý nhà nước về hàng hải đã chính thức đưa vào vận hành khai thác hệ thống VTS được lắp đặt tại khu vực luồng Vũng Tàu - Sài Gòn phục vụ cho công tác điều phối giám sát giao thông của các phương tiện thủy hoạt động ra vào tại khu vực này nhằm tăng cường công tác đảm bảo an toàn, an ninh hàng hải, phòng chống cháy nổ và phòng ngừa ô nhiễm môi trường.

Hệ thống quản lý hành hải luồng Vũng Tàu - Sài Gòn (sau đây gọi tắt là hệ thống VTS) ban đầu bao gồm 03 trạm Radar chuyên dụng đặt ở Núi Lớn – Vũng Tàu, Cần Giờ và Quận 7 - TP. Hồ Chí Minh, 01 trạm AIS đặt ở Núi Lớn và 04 Camera đặt ở khu vực có mật độ lưu thông lớn nhất khu vực là từ Nhà Bè đến cảng Sài Gòn. Thấy được hiệu quả thiết thực của hệ thống, Cục Hàng hải Việt Nam đã kiến nghị với Bộ Giao thông vận tải trình Chính phủ để đầu tư, lắp đặt và hoàn thiện dự án VTS luồng Cái Mép – Thị Vải, tích hợp với hệ thống VTS Vũng Tàu - Sài Gòn. Mục đích để bao phủ toàn bộ khu vực tuyến luồng Cái Mép – Thị Vải cũng như tăng cường phát hiện các mục tiêu nhỏ hoạt động trong khu vực VTS. Hệ thống VTS khu vực cảng biển Vũng Tàu – TP. Hồ Chí Minh – Đồng Nai – Mỹ Tho hiện nay bao gồm 06 Radar TERMA chuyên dùng, 07 Camera, 02 trạm thu AIS, 03 thiết bị đo gió.

Hệ thống VTS dựa trên nền hải đồ điện tử được tích hợp các thiết bị: Radar, AIS, Camera và VHF. Ngoài ra, hệ thống còn có ngân hàng dữ liệu Database. Tất cả các tín hiệu này được truyền về các trung tâm, qua thiết bị phần mềm xử lý, các tín hiệu mục tiêu được hiển thị trên màn hình điều khiển. Trong ca trực

giám sát viên/điều hành viên sẽ chịu trách nhiệm theo dõi mục tiêu liên tục trong vùng của mình phụ trách. Với trang thiết bị hiện có hệ thống VTS có thể phát hiện mục tiêu ở những khu vực từ ngoài hải đăng Vũng Tàu 15 hải lý trở vào bao trùm toàn bộ tuyến luồng sông Sài Gòn, Soài Rạp, sông Dinh, sông Cái Mép – Thị Vải (khu vực VTS). Với trang thiết bị của hệ thống như trên, hệ thống có hầu hết các thông tin của tàu trong khu vực VTS như vị trí tàu, cảng đến, thông tin hàng hóa trên tàu... để cảnh báo hỗ trợ kịp thời cho Hoa tiêu, Thuyền trưởng đưa ra các quyết định điều động tàu thuyền nhằm đảm bảo an toàn, an ninh hàng hải và phòng ngừa ô nhiễm môi trường.

Ngoài nhiệm vụ chính là đảm bảo an toàn hàng hải, hệ thống VTS còn có thể hỗ trợ công tác đảm bảo an ninh hàng hải, phòng ngừa ô nhiễm môi trường. Với các thông tin về vị trí tàu, số nhận dạng, cảng đến và hàng hoá trên tàu. Hệ thống VTS duy trì và giám sát lưu lượng tàu truy cập gần như tất cả các tàu trong khu vực VTS góp phần ngăn chặn các vụ khủng bố, cướp biển, bảo đảm an ninh hàng hải.

Luồng Vũng Tàu - Sài Gòn có thể phân ra 2 khu vực với cách đánh giá nguy cơ đâm va giữa các tàu thuyền với nhau. Tại khu vực đầu luồng từ biển vào, không gian vận động giữa các tàu rộng, nguy cơ đâm va có thể xác định theo CPA và T_{CPA} . Trong khu vực luồng còn lại, việc đánh giá nguy cơ đâm va sẽ được trình bày tại các phần sau.

1.3. Nghiên cứu xây dựng thuật toán xác định nguy cơ đâm va trong hàng hải giữa 2 tàu mục tiêu được quan sát từ radar của trạm VTS tại khu vực biển rộng

Ở phần trước, việc đánh giá nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng đã được đề cập. Việc đánh giá đó dựa trên dữ liệu đo đạc từ 2 tàu thuyền với nhau. Tuy nhiên, trong quản lý an toàn giao thông hàng hải, số liệu quan sát lại từ một bên thứ 3 đó là trạm VTS. Để có thể đánh giá nguy cơ đâm va giữa 2 tàu theo CPA và T_{CPA} nêu trên, ta cần phải chuyển thông tin thu nhận từ trạm VTS về dữ liệu theo CPA và T_{CPA} đối với từng cặp tàu. Chi tiết chuyển đổi dữ liệu như sau:

Giả sử từ trạm VTS có vị trí (φ_0, λ_0) , ta quan sát được n tàu mục tiêu.

Tàu mục tiêu 1 có phương vị PT_{S1} và khoảng cách D_{S1} tính từ trạm VTS.

Tàu mục tiêu 2 có phương vị PT_{S2} và khoảng cách D_{S2} tính từ trạm VTS.

.....

Tàu mục tiêu n có phương vị PT_{Sn} và khoảng cách D_{Sn} tính từ trạm VTS.

Để xác định được khoảng cách và phương vị giữa các tàu mục tiêu, cần phải tính toán được vị trí các tàu mục tiêu từ trạm VTS.

Công thức tính vị trí tàu mục tiêu được xây dựng như sau:

- Xây dựng hệ trục tọa độ OXY , trục hoành X tương ứng với giá trị kinh độ, trục tung Y tương ứng với giá trị vĩ độ. Tuy nhiên do tỉ lệ khoảng cách giữa các kinh tuyến và vĩ tuyến không bằng nhau nên vị trí kinh độ, vĩ độ của trạm VTS và các tàu mục tiêu sẽ được chuyển sang hệ OXY bằng công thức sau:

$$\begin{cases} X = R\lambda \cos \varphi_{TG} \\ Y = R\varphi \end{cases} \quad (1.3)$$

Trong đó:

- R là bán kính trái đất (hải lý);

- λ là kinh độ (rad);

- φ là vĩ độ (rad);

- φ_{TG} là vĩ độ trung gian tại khu vực đang xét (trong trường hợp này có thể chọn vĩ độ trung gian là φ_0 (rad)).

Dữ liệu kinh vĩ độ được sử dụng được tính toán theo hệ tọa độ VN2000.

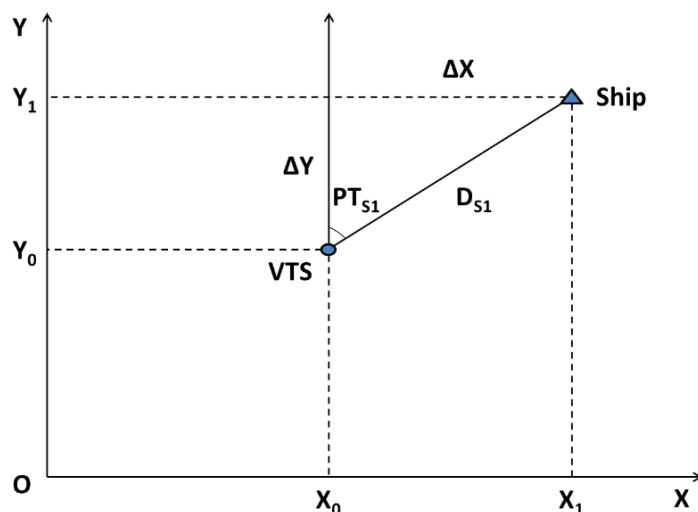
Như vậy, sau khi đổi, ta có vị trí trạm VTS là (X_0, Y_0) .

Vị trí các tàu mục tiêu cần tính là $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

Để tính vị trí tàu mục tiêu 1 (X_1, Y_1) , ta làm như sau (các tàu khác làm tương tự):

Xét vị trí tương quan của tàu mục tiêu so với trạm VTS, tàu có thể nằm ở 1 trong 4 cung phần tư có gốc là trạm VTS. Như vậy, ta có thể chia làm 4 trường hợp như sau:

Nếu $0^\circ \leq PT_{S1} \leq 90^\circ$ (hình 1.5)



Hình 1.5. Trường hợp tàu mục tiêu ở cung phần tư thứ nhất so với trạm VTS

Khi đó:

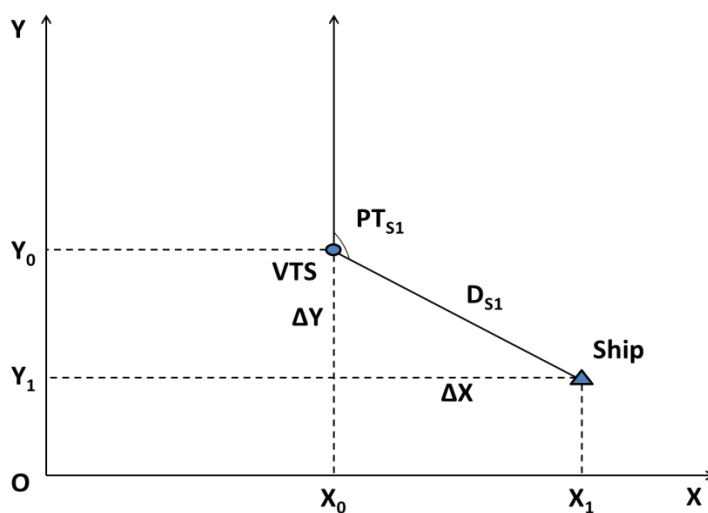
$$X_1 > X_0 \rightarrow \Delta X = X_1 - X_0 = D_{S1} \sin PT_{S1} \quad (1.4)$$

$$Y_1 > Y_0 \rightarrow \Delta Y = Y_1 - Y_0 = D_{S1} \cos PT_{S1} \quad (1.5)$$

$$\rightarrow X_1 = D_{S1} \sin PT_{S1} + X_0 \quad (1.6)$$

$$Y_1 = D_{S1} \cos PT_{S1} + Y_0 \quad (1.7)$$

Nếu $90^\circ < PT_{S1} \leq 180^\circ$ (hình 1.6)



Hình 1.6. Trường hợp tàu mục tiêu ở cung phần tư thứ hai so với trạm VTS

Khi đó:

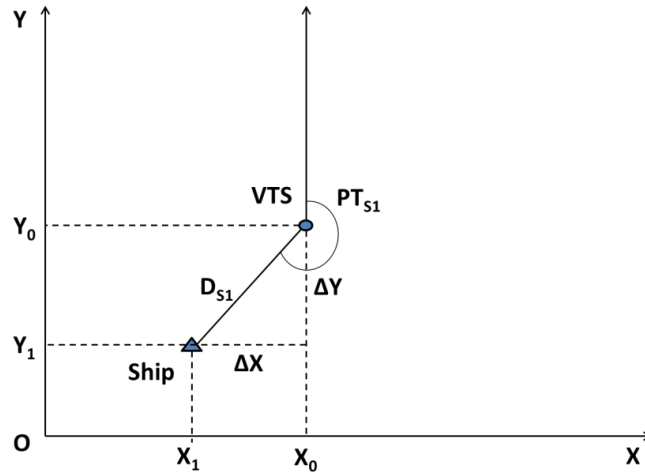
$$X_1 > X_0 \rightarrow \Delta X = X_1 - X_0 = D_{S1} \sin(180^\circ - PT_{S1}) = D_{S1} \sin PT_{S1} \quad (1.8)$$

$$Y_0 > Y_1 \rightarrow \Delta Y = Y_0 - Y_1 = D_{S1} \cos(180^\circ - PT_{S1}) = -D_{S1} \cos PT_{S1} \quad (1.9)$$

$$\rightarrow X_1 = D_{S1} \sin PT_{S1} + X_0 \quad (1.10)$$

$$Y_1 = D_{S1} \cos PT_{S1} + Y_0 \quad (1.11)$$

Nếu $180^\circ < PT_{S1} \leq 270^\circ$ (hình 1.7)



Hình 1.7. Trường hợp tàu mục tiêu ở cung phần tư thứ ba so với trạm VTS

Khi đó:

$$X_0 > X_1 \rightarrow \Delta X = X_0 - X_1 = D_{S1} \sin(PT_{S1} - 180^\circ) = -D_{S1} \sin PT_{S1} \quad (1.12)$$

$$Y_0 > Y_1 \rightarrow \Delta Y = Y_0 - Y_1 = D_{S1} \cos(PT_{S1} - 180^\circ) = -D_{S1} \cos PT_{S1} \quad (1.13)$$

$$\rightarrow X_1 = D_{S1} \sin PT_{S1} + X_0 \quad (1.14)$$

$$Y_1 = D_{S1} \cos PT_{S1} + Y_0 \quad (1.15)$$

Nếu $270^\circ < PT_{S1} \leq 360^\circ$ (hình 1.8)

Khi đó:

$$X_0 > X_1 \rightarrow \Delta X = X_0 - X_1 = D_{S1} \sin(360^\circ - PT_{S1}) = -D_{S1} \sin PT_{S1} \quad (1.16)$$

$$Y_1 > Y_0 \rightarrow \Delta Y = Y_1 - Y_0 = D_{S1} \cos(360^\circ - PT_{S1}) = D_{S1} \cos PT_{S1} \quad (1.17)$$

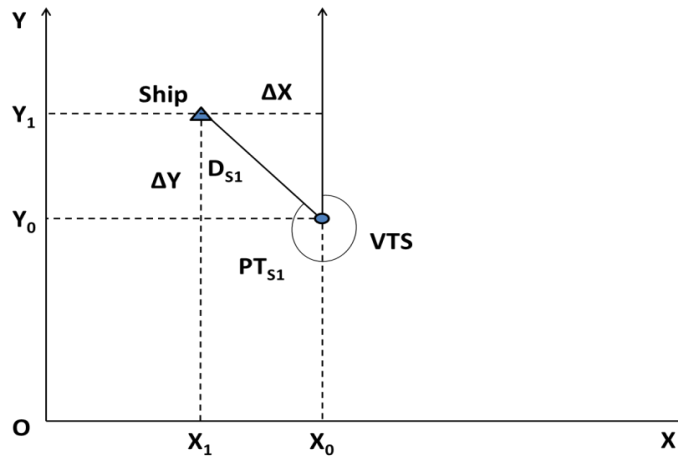
$$\rightarrow X_1 = D_{S1} \sin PT_{S1} + X_0 \quad (1.18)$$

$$Y_1 = D_{S1} \cos PT_{S1} + Y_0 \quad (1.19)$$

Như vậy, vị trí của các tàu mục tiêu có thể được tính theo hệ tọa độ OXY như trên.

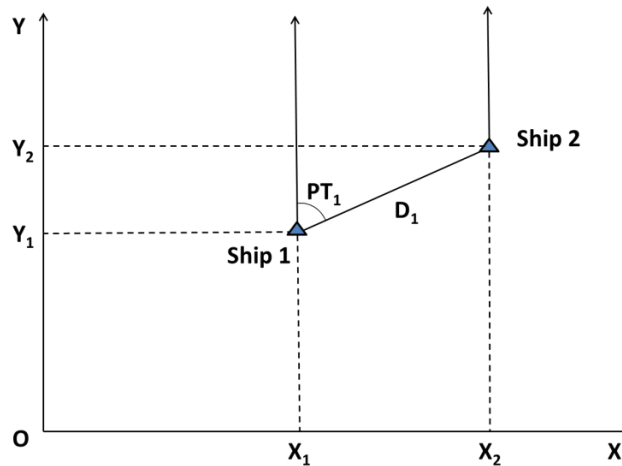
Phương vị và khoảng cách từ tàu mục tiêu 1 đến tàu mục tiêu 2 có thể được tính như sau:

Khi đã có được vị trí của các tàu mục tiêu, phương vị và khoảng cách giữa các tàu này cũng được tính toán dựa trên 4 trường hợp khi tàu mục tiêu 2 nằm ở 1 trong 4 cung phần tư có gốc là tàu mục tiêu 1.



Hình 1.8. Trường hợp tàu mục tiêu ở cung phần tư thứ tư so với trạm VTS

Nếu $X_2 > X_1, Y_2 > Y_1$ (hình 1.9)



Hình 1.9. Trường hợp tàu mục tiêu 2 ở cung phần tư thứ nhất so với tàu 1

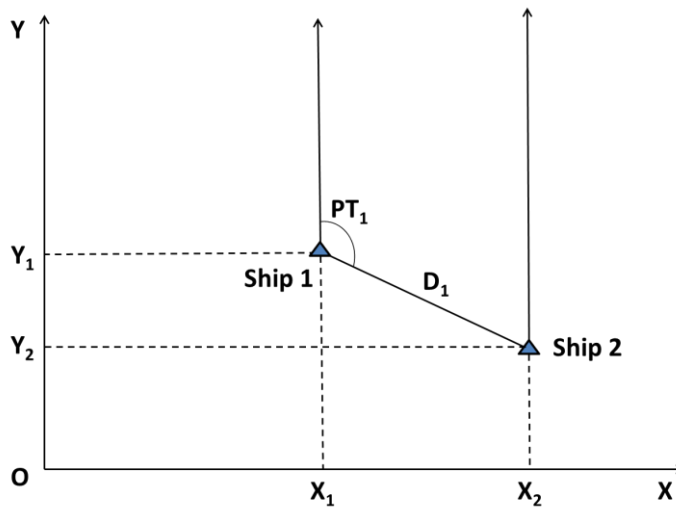
Khoảng cách từ tàu 1 đến tàu 2:

$$D_1 = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (1.20)$$

Phương vị từ tàu 1 đến tàu 2 trong trường hợp này phải thỏa mãn $0^\circ \leq PT_1 \leq 90^\circ$:

$$PT_1 = \arctan \left(\frac{X_2 - X_1}{Y_2 - Y_1} \right) \quad (1.21)$$

Nếu $X_2 > X_1, Y_2 < Y_1$ (hình 1.10)



Hình 1.10. Trường hợp tàu mục tiêu 2 ở cung phần tư thứ hai so với tàu 1

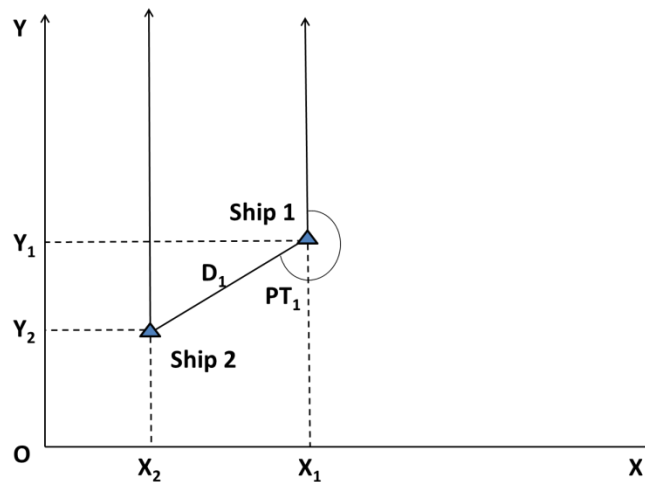
Khoảng cách từ tàu 1 đến tàu 2:

$$D_1 = \sqrt{(X_2 - X_1)^2 + (Y_1 - Y_2)^2} \quad (1.22)$$

Phương vị từ tàu 1 đến tàu 2 trong trường hợp này phải thỏa mãn $90^\circ < PT_1 \leq 180^\circ$:

$$PT_1 = 180^\circ - \arctan\left(\frac{X_2 - X_1}{Y_1 - Y_2}\right) \quad (1.23)$$

Nếu $X_2 < X_1, Y_2 < Y_1$ (hình 1.11)



Hình 1.11. Trường hợp tàu mục tiêu 2 ở cung phần tư thứ ba so với tàu 1

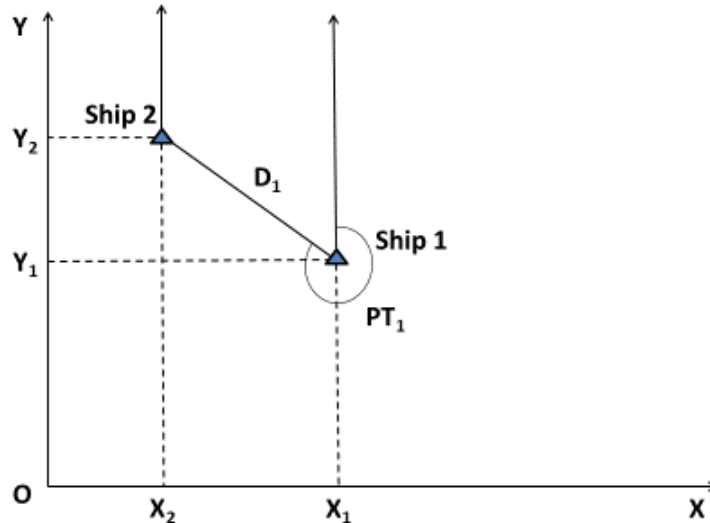
Khoảng cách từ tàu 1 đến tàu 2:

$$D_1 = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \quad (1.24)$$

Phương vị từ tàu 1 đến tàu 2 trong trường hợp này phải thỏa mãn $180^\circ < PT_1 \leq 270^\circ$:

$$PT_1 = 180^\circ + \arctan\left(\frac{X_1 - X_2}{Y_1 - Y_2}\right) \quad (1.25)$$

Nếu $X_2 < X_1, Y_2 > Y_1$ (hình 1.12)



Hình 1.12. Trường hợp tàu mục tiêu 2 ở cung phần tư thứ tư so với tàu 1

Khoảng cách từ tàu 1 đến tàu 2:

$$D_1 = \sqrt{(X_1 - X_2)^2 + (Y_2 - Y_1)^2} \quad (1.26)$$

Phương vị từ tàu 1 đến tàu 2 trong trường hợp này phải thỏa mãn $270^\circ < PT_1 < 360^\circ$:

$$PT_1 = 360^\circ - \arctan\left(\frac{X_1 - X_2}{Y_2 - Y_1}\right) \quad (1.27)$$

Để tính toán các thông số CPA, T_{CPA} phục vụ đánh giá nguy cơ đâm va ta xây dựng công thức tính như sau [11, 12]:

a. Nếu phương vị giữa 2 lần quan sát mục tiêu không đổi $PT_1 = PT_2$ thì $CPA = 0$

* Nếu $D_2 = D_1$ có nghĩa là vị trí tương đối giữa 2 tàu không thay đổi. Điều này có nghĩa là tàu mục tiêu và tàu chủ chuyển động cùng một hướng đi và một

giá trị vận tốc. Do 2 tàu giữ nguyên trạng thái tương đối so với nhau cho nên coi CPA và T_{CPA} không xác định.

Hai tàu không có nguy cơ đâm va.

* Nếu hai tàu đi lại gần nhau $D_2 < D_1$ (hình 1.13) thì ta có:

Ảnh của mục tiêu di chuyển qua tâm màn ảnh tương ứng với nó là khoảng cách tiếp cận gần nhất $CPA = 0$.

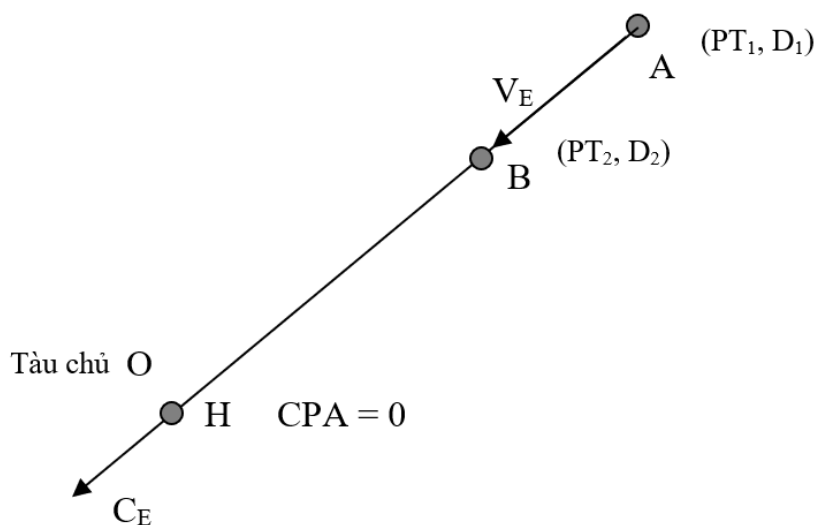
$$V_E = \frac{D_1 - D_2}{t_2 - t_1} \quad (1.28)$$

$$C_E = (PT_2 + 180^\circ) \quad (1.29)$$

Do giá trị hướng đi biến thiên từ 0° đến 360° nên ta phải giới hạn bằng công thức.

Nếu $C_E \geq 360^\circ$ thì $C_E = C_E - 360^\circ$; nếu khác giữ nguyên C_E .

$$T_{CPA} = \frac{D_2}{V_E} \quad (1.30)$$



Hình 1.13. Trường hợp phương vị không đổi, khoảng cách giảm dần

Với phương vị không đổi, khoảng cách giảm dần tương ứng với $CPA = 0$, $T_{CPA} > 0$, tồn tại nguy cơ đâm va giữa 2 tàu.

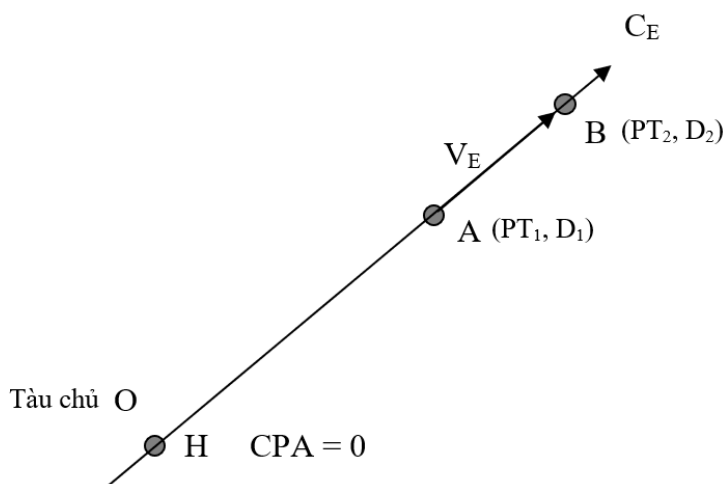
* Nếu hai tàu đi xa nhau (hình 1.14): $D_2 > D_1$ thì:

Trường hợp này vệt đi kéo dài của ảnh mục tiêu qua tâm màn ảnh (vị trí tàu chủ) cho nên điểm tiếp cận gần nhất chính là vị trí tàu chủ.

$$CPA = 0 \quad (1.30)$$

$$V_E = \frac{D_1 - D_2}{t_2 - t_1} \quad (V_E < 0 \text{ do tàu mục tiêu đi qua cận điểm}) \quad (1.31)$$

$$C_E = (PT_2 + 180^\circ) \quad (1.32)$$



Hình 1.14. Trường hợp phương vị không đổi, khoảng cách tăng dần

Do giá trị hướng đi biến thiên từ 0° đến 360° nên ta phải giới hạn bằng công thức

Nếu $C_E \geq 360^\circ$ thì $C_E = C_E - 360^\circ$; nếu khác giữ nguyên C_E .

$$T_{CPA} = \frac{D_2}{V_E} \quad (T_{CPA} < 0) \quad (1.33)$$

Hai tàu đang đi xa nhau, $CPA = 0$ nhưng $T_{CPA} < 0$ cho nên không tồn tại nguy cơ đâm va giữa các tàu.

b. Nếu phương vị giữa hai lần quan sát thay đổi $PT_1 \neq PT_2$

* Nếu khoảng cách giữa 2 tàu giảm dần $D_2 < D_1$ (hình 1.15)

Tính các thông số:

$$AB = \sqrt{D_1^2 + D_2^2 - 2D_1D_2 \cos(PT_1 - PT_2)} \quad (1.34)$$

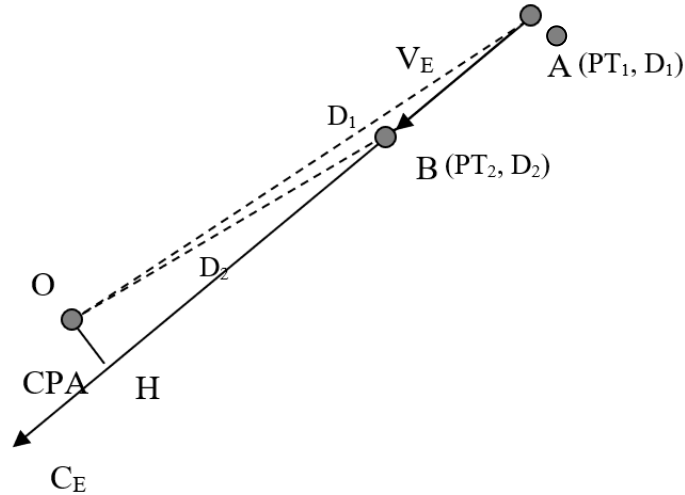
$$\cos OAH = \frac{D_1^2 + AB^2 - D_2^2}{2AB \cdot D_1} = X \quad (1.35)$$

$$OAH = \arccos(X) = Y \quad (1.36)$$

$$CPA = D_1 \sin(Y) \quad (\text{hải lý}) \quad (1.37)$$

$$T_{CPA} = \frac{D_1 \cos(Y) - AB}{V_E} \text{ (phút, giây)} \quad (1.38)$$

$$OBH = \arcsin\left(\frac{CPA}{D_2}\right) \text{ (độ)} \quad (1.39)$$



Hình 1.15. Trường hợp phương vị thay đổi, khoảng cách giảm dần

Phương vị nghịch của các lần quan sát được tính như sau:

Nếu $(PT_1 + 180^0) < 360^0$ thì $PTN_1 = (PT_1 + 180^0)$

Nếu $(PT_1 + 180^0) > 360^0$ thì $PTN_1 = (PT_1 + 180^0) - 360^0$

Nếu $(PT_2 + 180^0) < 360^0$ thì $PTN_2 = (PT_2 + 180^0)$

Nếu $(PT_2 + 180^0) > 360^0$ thì $PTN_2 = (PT_2 + 180^0) - 360^0$

Góc mạn của tàu mục tiêu (GM) được tính theo công thức:

Nếu $0 \leq (PT_1 - C_0) \leq 180^0$ thì $GM_1 = (PT_1 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_1 - C_0) \leq -180^0$ thì $GM_1 = 360^0 - (PT_1 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_1 - C_0) > 180^0$ thì $GM_1 = (PT_1 - C_0) - 360^0 \rightarrow$ góc mạn trái

Nếu $-180^0 < (PT_1 - C_0) < 0$ thì $GM_1 = (PT_1 - C_0) \rightarrow$ góc mạn trái

Nếu $0 \leq (PT_2 - C_0) \leq 180^0$ thì $GM_2 = (PT_2 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_2 - C_0) \leq -180^0$ thì $GM_2 = 360^0 - (PT_2 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_2 - C_0) > 180^0$ thì $GM_2 = (PT_2 - C_0) - 360^0 \rightarrow$ góc mạn trái

Nếu $-180^0 < (PT_2 - C_0) < 0$ thì $GM_2 = (PT_2 - C_0) \rightarrow$ góc mạn trái

Tốc độ chuyển động của ảnh mục tiêu V_E (tốc độ tương đối):

$$V_E = \frac{AB}{t_2 - t_1} \quad (1.40)$$

Hướng chuyển động của ảnh mục tiêu C_E (hướng tương đối)

Nếu $0 < GM_1, GM_2 < 180^0$ (tàu mục tiêu bên mạn phải)

Nếu $GM_1 > GM_2$ thì tàu mục tiêu qua phía trước mũi.

Ta có: $C_E = PTN_2 + OBH$ (độ)

Nếu $GM_1 < GM_2$ thì tàu mục tiêu qua phía sau lái.

Ta có: $C_E = PTN_2 - OBH$ (độ)

Nếu $-180^0 < GM_1, GM_2 < 0$ (tàu mục tiêu bên mạn trái)

Nếu $GM_1 > GM_2$ thì tàu mục tiêu qua phía sau lái.

Ta có: $C_E = PTN_2 + OBH$ (độ)

Nếu $GM_1 < GM_2$ thì tàu mục tiêu qua phía trước mũi.

Ta có: $C_E = PTN_2 - OBH$ (độ)

Nếu $-90^0 \leq GM_1 \leq 0; 0 \leq GM_2 \leq 90^0$ thì tàu mục tiêu qua phía trước mũi.

Ta có: $C_E = PTN_2 - OBH$ (độ)

Nếu $0 < GM_1 \leq 90^0; -90^0 \leq GM_2 < 0$ thì tàu mục tiêu qua phía trước mũi.

Ta có: $C_E = PTN_2 + OBH$ (độ)

Nếu $90 < GM_1 \leq 180^0; -180^0 \leq GM_2 < -90^0$ thì tàu mục tiêu qua phía sau lái.

Ta có: $C_E = PTN_2 + OBH$ (độ)

Nếu $-180^0 \leq GM_1 < -90^0; 90 < GM_2 \leq 180^0$ thì tàu mục tiêu qua phía sau lái.

Ta có: $C_E = PTN_2 - OBH$ (độ)

Công thức hạn chế giá trị của C_E trong khoảng 0^0 đến 360^0 :

Nếu $C_E \geq 360^0$ thì $C_E = C_E - 360^0$; nếu khác giữ nguyên C_E .

Trong trường hợp này hai tàu tiến lại gần nhau $T_{CPA} > 0$, để đánh giá nguy cơ đâm va chúng ta cần so sánh giá trị CPA với CPA_{min} . Nếu $CPA < CPA_{min}$ thì kết luận tồn tại nguy cơ đâm va giữa 2 tàu.

* Nếu khoảng cách giữa 2 tàu tăng dần $D_2 > D_1$ (hình 1.16)

Tính các thông số:

$$AB = \sqrt{D_1^2 + D_2^2 - 2D_1D_2 \cos(PT_1 - PT_2)} \quad (1.41)$$

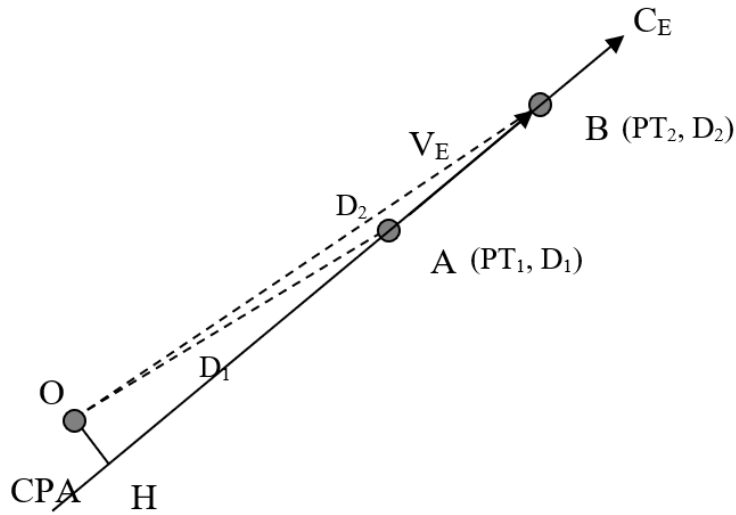
$$\cos OAH = \frac{D_1^2 + AB^2 - D_2^2}{2AB.D_1} = X \quad (1.42)$$

$$OAH = \arccos(X) = Y \quad (1.43)$$

$$CPA = D_1 \sin(Y) \text{ (hải lý)} \quad (1.44)$$

$$T_{CPA} = \frac{D_1 \cos(Y) - AB}{V_E} \text{ (phút, giây)} \quad (1.45)$$

$$OBH = \arcsin\left(\frac{CPA}{D_2}\right) \text{ (độ)} \quad (1.46)$$



Hình 1.16. Trường hợp phương vị thay đổi, khoảng cách tăng dần

Phương vị nghịch của các lần quan sát được tính như sau:

Nếu $(PT_1 + 180^0) < 360^0$ thì $PTN_1 = (PT_1 + 180^0)$

Nếu $(PT_1 + 180^0) > 360^0$ thì $PTN_1 = (PT_1 + 180^0) - 360^0$

Nếu $(PT_2 + 180^0) < 360^0$ thì $PTN_2 = (PT_2 + 180^0)$

Nếu $(PT_2 + 180^0) > 360^0$ thì $PTN_2 = (PT_2 + 180^0) - 360^0$

Góc mạn của tàu mục tiêu (GM) được tính theo công thức:

Nếu $0 \leq (PT_1 - C_0) \leq 180^0$ thì $GM_1 = (PT_1 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_1 - C_0) \leq -180^0$ thì $GM_1 = 360^0 - (PT_1 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_1 - C_0) > 180^0$ thì $GM_1 = (PT_1 - C_0) - 360^0 \rightarrow$ góc mạn trái

Nếu $-180^0 < (PT_1 - C_0) < 0$ thì $GM_1 = (PT_1 - C_0) \rightarrow$ góc mạn trái

Nếu $0 \leq (PT_2 - C_0) \leq 180^0$ thì $GM_2 = (PT_2 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_2 - C_0) \leq -180^0$ thì $GM_2 = 360 - (PT_2 - C_0) \rightarrow$ góc mạn phải

Nếu $(PT_2 - C_0) > 180^0$ thì $GM_2 = (PT_2 - C_0) - 360^0 \rightarrow$ góc mạn trái

Nếu $-180^0 < (PT_2 - C_0) < 0$ thì $GM_2 = (PT_2 - C_0) \rightarrow$ góc mạn trái

Tốc độ chuyển động của ảnh mục tiêu V_E (tốc độ tương đối)

$$V_E = -\frac{AB}{t_2 - t_1} \quad (1.47)$$

Hướng chuyển động của ảnh mục tiêu C_E (hướng tương đối)

Nếu $0 < GM_1, GM_2 < 180^0$ (tàu mục tiêu bên mạn phải)

Nếu $GM_1 > GM_2$ thì tàu mục tiêu đi ra xa phía trước mũi.

$$\text{Ta có: } C_E = PTN_2 + 180^0 - OBH \quad (\text{độ}) \quad (1.48)$$

Nếu $GM_1 < GM_2$ thì tàu mục tiêu ra xa phía sau lái.

$$\text{Ta có: } C_E = PTN_2 - 180^0 + OBH \quad (\text{độ}) \quad (1.49)$$

Nếu $-180^0 < GM_1, GM_2 < 0$ (tàu mục tiêu bên mạn trái)

Nếu $GM_1 > GM_2$ thì tàu mục tiêu ra xa phía sau lái.

$$\text{Ta có: } C_E = PTN_2 + 180^0 + OBH \quad (\text{độ}) \quad (1.50)$$

Nếu $GM_1 < GM_2$ thì tàu mục tiêu ra xa phía trước mũi.

$$\text{Ta có: } C_E = PTN_2 - 180^0 - OBH \quad (\text{độ}) \quad (1.51)$$

Nếu $-90^0 \leq GM_1 \leq 0; 0 \leq GM_2 \leq 90^0$ thì tàu mục tiêu ra xa phía trước mũi.

$$\text{Ta có: } C_E = PTN_2 - 180^0 + OBH \quad (\text{độ}) \quad (1.52)$$

Nếu $0 < GM_1 \leq 90^0; -90^0 \leq GM_2 < 0$ thì tàu mục tiêu ra xa phía trước mũi.

$$\text{Ta có: } C_E = PTN_2 + 180^0 - OBH \quad (\text{độ}) \quad (1.53)$$

Nếu $90 < GM_1 \leq 180^0; -180^0 \leq GM_2 < -90^0$ thì tàu mục tiêu ra xa phía sau lái.

$$\text{Ta có: } C_E = PTN_2 - 180^0 + OBH \quad (\text{độ}) \quad (1.54)$$

Nếu $-180^0 \leq GM_1 < -90^0; 90 < GM_2 \leq 180^0$ thì tàu mục tiêu ra xa phía sau lái.

$$\text{Ta có: } C_E = PTN_2 + 180^0 - OBH \quad (\text{độ}) \quad (1.55)$$

Công thức hạn chế giá trị của C_E trong khoảng 0^0 đến 360^0 :

Nếu $C_E \geq 360^0$ thì $C_E = C_E - 360^0$; nếu khác giữ nguyên C_E .

Trong trường hợp này 2 tàu đi xa nhau cho nên không tồn tại nguy cơ đâm va.

Như vậy, các hệ công thức giúp đánh giá nguy cơ đâm va giữa 2 tàu mục tiêu dựa trên thông tin quan sát từ trạm VTS đã được xây dựng trong phần này. Đây là cơ sở quan trọng để có thể xây dựng phần mềm đánh giá nguy cơ đâm va giữa 2 tàu mục tiêu sử dụng dữ liệu quan sát từ bên thứ 3 giúp các trạm VTS có thể đánh giá nguy cơ đâm va một cách định lượng. Các công thức trên cũng có thể được sử dụng để giúp AI nhận dạng tình huống có nguy cơ đâm va giữa 2 tàu thuyền mục tiêu.

Trong một khu vực chạy tàu, thường có rất nhiều phương tiện tham gia giao thông đặc biệt là ở các khu vực luồng lạch hẹp, đầu các tuyến luồng. Như vậy, 1 con tàu khi tham gia giao thông muốn đảm bảo an toàn thì không chỉ đánh giá nguy cơ đâm va với chỉ 1 tàu mục tiêu khác mà còn với nhiều tàu trong khu vực lân cận. Việc đánh giá nguy cơ đâm va với nhiều mục tiêu tại cùng một thời điểm và cùng đưa ra thông tin cảnh báo có thể dẫn tới quá tải trong tiếp nhận thông tin của người điều khiển phương tiện chứ chưa nói đến người sỹ quan VTS khi cần phải theo dõi, phát hiện nguy cơ đâm va giữa nhiều mục tiêu trong khu vực trạm VTS. Vì vậy, cần đề ra những nguyên tắc nhất định trong xác định và cảnh báo nguy cơ đâm va.

Như ta đều biết, việc phát hiện, đánh giá nguy cơ đâm va giữa các tàu thuyền phải đảm bảo làm sao có đủ thời gian, không gian để thực hiện hành động tránh đâm va một cách hiệu quả. Nếu không đảm bảo được yêu cầu đó thì mục tiêu của việc phòng ngừa tai nạn đâm va của các tàu thuyền không thể thực hiện được. Vì vậy, việc xây dựng các nguyên tắc xác định nguy cơ đâm va giữa nhiều tàu thuyền cũng phải đảm bảo làm sao có thể phòng ngừa được đâm va giữa chúng.

Nguyên tắc trước tiên có lẽ căn cứ trên khoảng cách giữa các tàu mục tiêu. Khoảng cách càng lớn thì thời gian dự trữ cho hành động tránh đâm va càng lớn, càng có khả năng đảm bảo an toàn. Theo khoản (a) điều 8 [9, 10] có qui định mở “*nếu hoàn cảnh cho phép*” thì hành động tránh va phải được tiến hành “*một cách dứt khoát, kịp thời và phù hợp với kinh nghiệm của người đi biển lành nghề*”. Hành động tránh va được cho là kịp thời khi nó được tiến hành trong thời gian đủ sớm để tránh nguy cơ đâm va với tàu thuyền đang đến gần. Trong thực tế, đã có rất nhiều vụ va chạm xảy ra vì đã không đưa ra hành động tránh va kịp thời. Việc đưa ra hành động kịp thời không những là yêu cầu đối với các tàu có nghĩa vụ phải nhường đường, phải hành động mà còn cả đối với tàu thuyền được nhường đường khi xét thấy hành động đơn thuần của tàu thuyền phải nhường đường không đủ để tránh một nguy cơ đâm va trước mắt.

Việc xác định tính kịp thời của hành động không được nêu trong quy tắc nào chú dẫn nào về chúng và tất nhiên là không thể ấn định trước một cách chính xác được là khoảng cách nào tới tàu kia thì phải bắt đầu điều động để tránh nó trong một điều kiện cụ thể. Trong những trường hợp riêng biệt, tùy thuộc vào hoàn cảnh xung quanh, tốc độ tiếp cận giữa các tàu và khả năng điều động của các tàu, khoảng cách này có thể là khác nhau. Theo thực tiễn đi biển, ta có thể lựa chọn khoảng cách tại đó đưa ra hành động tránh va như sau:

- + Khi có nhiều tàu mục tiêu cùng có nguy cơ đâm va thì tránh các tàu ở gần trước, tàu ở xa sau;
- + Khi tốc độ tiếp cận giữa hai tàu lớn thì cần phải hành động sớm hơn khi tốc độ tiếp cận nhỏ;
- + Các tàu có khả năng điều động tốt thường tránh va ở khoảng cách gần hơn các tàu khó điều khiển, vận động khó khăn;
- + Vị trí của tàu ở gần các chướng ngại hàng hải cũng là cơ sở để lựa chọn tránh sớm ở khoảng cách xa hay tránh muộn.

1.4. Nghiên cứu nguyên tắc xác định nguy cơ đâm va giữa các mục tiêu áp dụng trong khu vực luồng lạch hẹp

Như đã trình bày trong phần 1.2. trong khu vực luồng lạch hẹp, phương pháp đánh giá nguy cơ đâm va theo CPA và T_{CPA} không còn phù hợp do các yếu tố:

- Không gian chật hẹp khó có thể áp dụng CPA_{min} một cách hợp lý;
- Trạng thái chuyển động của tàu luôn thay đổi theo hình dạng luồng và các yêu cầu về tốc độ chạy trên luồng (nếu có). Trạng thái chuyển động của các tàu luôn thay đổi dẫn đến các đánh giá về nguy cơ đâm va trong từng thời điểm sẽ khác nhau.

Để có thể đánh giá nguy cơ đâm va giữa các tàu thuyền trên luồng lạch hẹp, người ta thường lựa chọn giải pháp khác để đánh giá nguy cơ đâm va giữa các tàu thuyền. Khi hành trình trong các tuyến luồng chật hẹp và nông cạn, mật độ giao thông đông, nguy cơ đâm va cao thì việc tăng cường nguồn nhân lực buồng lái cũng được các tàu thuyền tham gia giao thông áp dụng. Người ta bố trí hoa tiêu, thuyền trưởng tăng cường sỹ quan, thủy thủ trong vận hành các con tàu, các trang thiết bị liên quan cũng sẵn sàng để có thể thay đổi nhanh chóng trạng thái chuyển động của con tàu. Nhờ được bổ sung nhân sự, sẵn sàng các trang thiết bị, các tàu sẽ được điều khiển tốt hơn, nguy cơ đâm va được phát hiện kịp thời hơn, từ đó đưa ra được các giải pháp phòng ngừa va chạm tàu thuyền trên luồng lạch hẹp tốt và kịp thời hơn. Các tàu có thể đi qua nhau ở các khoảng cách nhỏ hơn mà vẫn đảm bảo an toàn.

Thông thường, khi tham gia giao thông trên hệ thống luồng lạch hẹp, các tàu kích thước nhỏ nhờ khả năng vận động linh hoạt, chiếm dụng ít không gian giao thông ít hơn các tàu lớn. Các tàu lớn thường bị khống chế bởi các điều kiện chật hẹp và nông cạn cho nên vận động khó khăn. Các tàu thuyền khác thường phải có trách nhiệm không được làm cản trở sự đi lại an toàn của tàu thuyền bị mớn nước khống chế này theo quy định tại Điều 9 Colreg72. Nếu có 2 tàu lớn đều chịu các ảnh hưởng bởi luồng lạch hẹp cùng đi qua một khu vực hạn chế

nào đó tại cùng một thời điểm thì sẽ dẫn đến nguy cơ mất an toàn giao thông. Trong quản lý an toàn giao thông, người ta thường tránh để các tàu lớn gặp nhau tại những điểm mất an toàn giao thông (còn gọi là điểm nóng giao thông) nhằm đảm bảo an toàn.

Để làm được nhiệm vụ tránh các tàu lớn gặp nhau tại các điểm nóng giao thông, thông tin về thời gian dự kiến đến các điểm đó *ETA* của từng tàu cần được kiểm soát, cập nhật thường xuyên. Để xác định được giá trị *ETA* đến các điểm nóng giao thông, thông tin về tuyến hành trình, vị trí của từng tàu, vị trí của các điểm nóng giao thông, tốc độ của tàu cần được thu thập và tính toán.

Trước tiên, các thông tin về vị trí các điểm nóng giao thông cần được khảo sát. Đối tượng khảo sát được xác định là từ cơ quan quản lý tuyến luồng, các hoa tiêu, thuyền trưởng đã có kinh nghiệm dẫn tàu trên tuyến luồng đó. Căn cứ theo kết quả đánh giá của họ, số lượng và vị trí của điểm nóng giao thông được xác định.

Từ các thông tin về vị trí của tàu theo thời gian, đối chiếu với tuyến hành trình trên luồng, quãng đường, vận tốc, thời gian đến các điểm nóng giao thông được xác định. Cách tính vận tốc của tàu, quãng đường từ tàu đến các điểm chuyển hướng trên tuyến hành trình, đến điểm nóng giao thông được xác định theo các công thức tính toán như trong mục 1.3 ở trên. Từ quãng đường đến điểm nóng giao thông, vận tốc của tàu vừa tính toán được, thời gian dự kiến tàu đến điểm nóng giao thông (*ETA*) được xác định làm cơ sở để so sánh với *ETA* của các tàu thuyền khác.

Sau khi tính được *ETA* của các tàu đến từng điểm nóng giao thông, giá trị chênh lệch giữa các *ETA* là bao nhiêu để kết luận tồn tại nguy cơ đâm va giữa 2 tàu thuyền được khảo sát. Từ kết quả khảo sát, độ chênh *ETA* coi là tồn tại nguy cơ đâm va giữa 2 tàu được xác định một cách định lượng. Giá trị này được sử dụng làm căn cứ khi xây dựng cơ sở dữ liệu đánh giá nguy cơ đâm va phục vụ huấn luyện AI.

1.5. Kết luận chương 1

Trong chương 1, cơ sở lý thuyết về đánh giá nguy cơ đâm va trên biển rộng và trong hệ thống luồng lạch hẹp đã được xây dựng. Đối với luồng Vũng Tàu - Sài Gòn, có thể đánh giá nguy cơ đâm va giữa các tàu thuyền trên biển theo các thông số CPA và T_{CPA} trước khi các tàu nhập luồng. Đoạn còn lại, cần xác định các điểm nóng giao thông của luồng và đánh giá nguy cơ đâm va theo ETA của các tàu đến các điểm nóng giao thông đó đảm bảo 2 tàu không đồng thời xuất hiện tại các điểm nóng về giao thông gây mất an toàn.

Chi tiết việc đánh giá nguy cơ đâm va sẽ được triển khai tại các phần sau.

CHƯƠNG 2. TỔNG QUAN VỀ ĐÁNH GIÁ NGUY CƠ ĐÂM VÀ GIỮA CÁC MỤC TIÊU TRÊN VÙNG BIỂN RỘNG VÀ TRONG LUỒNG CHẠY TÀU

2.1. Ứng dụng hệ thống mô phỏng của Trường Đại học Hàng hải Việt Nam trong việc kiểm tra thuật toán xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát

Trong phần 1.2, nghiên cứu sinh đã triển khai xây dựng công thức đánh giá nguy cơ đâm va giữa 2 tàu thuyền mục tiêu dựa trên kết quả quan sát của bên thứ 3 (trạm VTS). Để kiểm chứng công thức này, cần đánh giá độ chính xác của công thức trong thực tiễn. Để làm việc này, hệ thống mô phỏng buồng lái của Trường Đại học Hàng hải Việt Nam được sử dụng.



Hình 2.1. Hệ thống mô phỏng buồng lái tại Trường Đại học Hàng hải Việt Nam

Hệ thống mô phỏng buồng lái Navi Trainer Professional 5000 (NT PRO 5000) tại Trường Đại học Hàng hải Việt Nam là một hệ thống có khả năng phục vụ đào tạo, huấn luyện và nghiên cứu khoa học do Transas xây dựng và phát triển. Hệ thống này được cấp chứng chỉ công nhận của Đăng kiểm Na Uy (DNV)

và Cục Hàng hải Liên Bang Nga. Đây là một trong những cơ sở quan trọng để có thể sử dụng hệ thống trong triển khai các hoạt động đào tạo, huấn luyện và nghiên cứu khoa học.



Hình 2.2. Chứng chỉ của DNV cấp cho hệ thống mô phỏng

Hệ thống bao gồm 01 phòng huấn luyện viên và 01 phòng mô phỏng buồng lái. Phòng huấn luyện viên được trang bị các thiết bị sau:

- 01 máy chủ NT PRO SERVER;
- 02 máy Instructor Navi Trainer;
- 01 máy Instructor GMDSS TGS 5000;
- 01 máy Model Wizard;
- 07 máy truy theo 07 máy chiếu trong buồng mô phỏng buồng lái chính;
- 01 máy CCTV.

Các máy này được liên kết mạng lan với máy chủ NT PRO SERVER.

Phòng mô phỏng buồng lái được trang bị:

- Hệ thống thông tin liên lạc GMDSS;

- 02 Radar/ARPA;
- 01 trạm Conning display;
- 01 màn hình IBID;
- 01 màn hình Bearing;
- 01 hệ thống thông tin và hiển thị hải đồ điện tử ECDIS (Navi Sailor 4000);
- Các trang thiết bị nghi khí Hàng hải khác và bàn thao tác hải đồ.

Hệ thống mô phỏng buồng lái có góc nhìn 240⁰ cho phép thuyền viên quan sát trên một góc nhìn rộng.

Hệ thống có thể cho phép xây dựng các tình huống giao thông trên biển với nhiều loại tàu chủ và tàu mục tiêu khác nhau. Các thiết bị buồng lái cho phép tương tác như trên buồng lái của các tàu biển thông thường. Hệ thống cho phép mô phỏng 09 loại tàu chủ khác nhau. Cụ thể:

- Tàu hàng rời

Lượng giãn nước	26343 tấn
Điều kiện xếp hàng	Không hàng
Mới nước mũi	6.57 m
Mớn nước lái	6.66 m
LOA	199.99 m
B	23.76 m
Full SeaAhead	15.2 kts (<i>tới hết máy khẩn cấp</i>)
Full Ahead	11.2 kts (<i>tới hết máy</i>)

- Tàu hàng rời Panamax

Lượng giãn nước	69580 tấn
Điều kiện xếp hàng	Đầy tải
Mới nước mũi	12 m
Mớn nước lái	12 m
LOA	230 m
B	32 m

Full SeaAhead	14.8 kts
Full Ahead	10.3 kts
- Tàu chở ô tô	
Lượng giãn nước	39282 tấn
Điều kiện xếp hàng	Đầy tải
Mới nước mũi	11.2 m
Mớn nước lái	11.2 m
LOA	199.1 m
B	32.36 m
Full SeaAhead	19.6 kts
Full Ahead	12.5 kts
- Tàu cảnh sát biển	
Lượng giãn nước	834 tấn
Điều kiện xếp hàng	Không hàng
Mới nước mũi	3.64 m
Mớn nước lái	3.64 m
LOA	65.9 m
B	10.7 m
Full SeaAhead	21.5 kts
Full Ahead	19.6 kts
- Tàu container	
Lượng giãn nước	66700 tấn
Điều kiện xếp hàng	Đầy tải
Mới nước mũi	12 m
Mớn nước lái	12 m
LOA	289 m
B	32.2 m
Full SeaAhead	23.5 kts
Full Ahead	16.9 kts

- Tàu chở khí tự nhiên hóa lỏng (LNG)

Lượng giãn nước	109623 tấn
Điều kiện xếp hàng	Đầy tải
Mới nước mũi	11.15 m
Mớn nước lái	11.15 m
LOA	294.6 m
B	46.66 m
Full SeaAhead	19.5 kts
Full Ahead	11.9 kts

- Tàu cứu hộ

Lượng giãn nước	30.6 tấn
Điều kiện xếp hàng	Không hàng
Mới nước mũi	0.92 m
Mớn nước lái	0.92 m
LOA	18.94 m
B	5.65 m
Full SeaAhead	29.9 kts
Full Ahead	21.5 kts

- Tàu lai

Lượng giãn nước	5291 tấn
Điều kiện xếp hàng	Đầy tải
Mới nước mũi	6.52 m
Mớn nước lái	6.6 m
LOA	80.4 m
B	18 m
Full SeaAhead	16.2 kts
Full Ahead	15 kts

- Tàu VLCC

Lượng giãn nước	189405.99 tấn
-----------------	---------------

Điều kiện xếp hàng	Đầy tải
Mới nước mũi	16.62 m
Mớn nước lái	16.62 m
LOA	280.5 m
B	50 m
Full SeaAhead	16.5 kts
Full Ahead	15 kts

Các tàu này có thể sử dụng làm cả tàu chủ và tàu mục tiêu trong các tình huống mô phỏng lại chuyển động của tàu thuyền.

2.2. Xây dựng bài tập mô phỏng kiểm tra thuật toán xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát

Thuật toán đánh giá nguy cơ đâm va được nêu trong mục 1.2 cần được kiểm tra độ chính xác. Để kiểm tra, một tình huống chuyển động giữa 2 tàu được xây dựng và triển khai thực hiện trên hệ thống mô phỏng NT PRO 5000 của Transas tại Trường Đại học Hàng hải Việt Nam. Trong tình huống này, 2 tàu được bố trí chạy cắt hướng nhau và thực hiện quan sát đánh giá nguy cơ đâm va bằng thiết bị ARPA của tàu chủ so với tàu mục tiêu đồng thời so sánh với kết quả tính toán theo công thức tại phần 1.2 để đánh giá nguy cơ đâm va từ 04 trạm VTS ảo như trong hình 2.3.



Hình 2.3. Bố trí tàu và các trạm VTS ảo trong bài tập thực nghiệm

Bảng 2.1. Vị trí các trạm VTS ảo

VTS 1	20.71074N	107.0212E
VTS 2	20.60344N	106.9787E
VTS 3	20.59573N	106.7909E
VTS 4	20.66689N	106.8162E

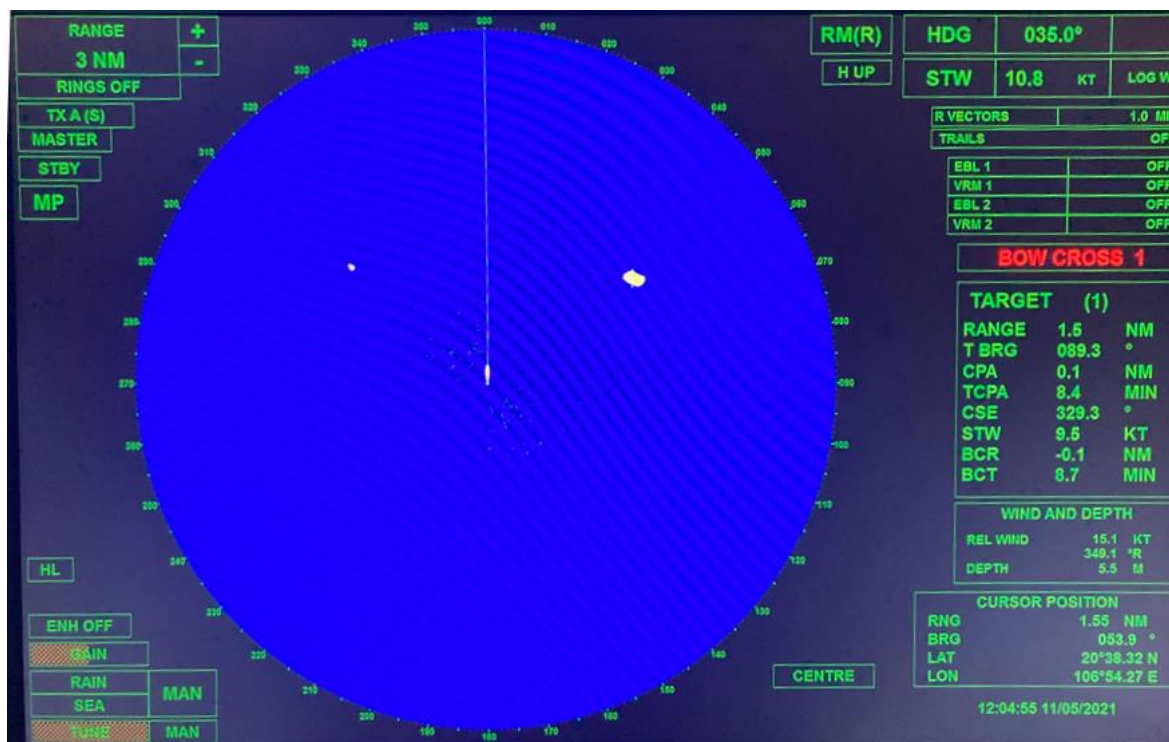
Hai tàu được chọn là 2 tàu chạy với tốc độ không đổi (tàu chủ 10.7kts, tàu hàng 9.5kts) và duy trì hướng đi để thuận tiện đánh giá nguy cơ đâm va trong quá trình thực nghiệm. Bốn trạm VTS ảo được đặt ở 4 cung phần tư so với tàu chủ và tàu mục tiêu đảm bảo tính toán được các tình huống tương quan giữa vị trí tàu chủ và tàu mục tiêu với trạm VTS như đã nêu trong phần 1.2.

2.3. Đánh giá kết quả thực nghiệm xác định nguy cơ đâm va giữa 2 tàu thuyền trên vùng biển rộng dựa trên kết quả quan sát từ trạm VTS

Từ bài tập mô phỏng được xây dựng ở phần 2.2, công tác thực nghiệm được triển khai để đối chiếu, so sánh kết quả tính toán lý thuyết ở phần 1.2 với kết quả tính toán của hệ thống ARPA của Transas.

Từ vị trí của tàu chủ và tàu mục tiêu so với 4 trạm VTS ảo được ghi lại trong quá trình thực nghiệm, các thông tin quan sát từ trạm VTS sẽ được tính toán để đánh giá nguy cơ đâm va.

Đồng thời với việc ghi lại thông tin vị trí của tàu mục tiêu, tàu chủ, các thông tin CPA , T_{CPA} của tàu mục tiêu trên màn hình radar (hình 2.4) cũng được ghi lại phục vụ đánh giá độ chính xác của công thức lý thuyết trong phần 1.2. Các thông tin thực nghiệm ghi lại thể hiện trong bảng 2.2.



Hình 2.4. Màn hình radar của tàu chủ

Bảng 2.2. Kết quả thực nghiệm

OS				TS				CPA	TCPA	Time	
Latitude		Longitude		Latitude		Longitude					
20	37.946	106	52.359	20	38.05	106	54.57	0.1	10	12:02:33	t1
20	38.142	106	52.507	20	38.17	106	54.37	0.1	9.4	12:03:49	t2
20	38.307	106	52.633	20	38.32	106	54.27	0.1	8.4	12:04:55	t3
20	38.449	106	52.739	20	38.48	106	54.18	0.1	7.4	12:05:56	t4
20	38.6	106	52.853	20	38.61	106	54.11	0.1	6.4	12:06:56	t5
20	38.745	106	52.963	20	38.74	106	54.01	0.1	5.4	12:07:58	t6
20	38.912	106	53.09	20	38.89	106	53.95	0.1	4.3	12:09:03	t7
20	39.026	106	53.176	20	39.01	106	53.86	0.1	3.5	12:09:52	t8
20	39.144	106	53.266	20	39.13	106	53.8	0.1	2.6	12:10:43	t9
20	39.283	106	53.373	20	39.25	106	53.72	0	1.7	12:11:41	t10
20	39.378	106	53.448	20	39.34	106	53.67	0	1	12:12:20	t11
20	39.492	106	53.535	20	39.44	106	53.61	0	0.3	12:13:03	t12

Từ kết quả thực nghiệm, thực hiện tính toán các số liệu theo công thức nêu ở phần 1.2 ta có kết quả tính toán chuyển đổi phương vị và khoảng cách giữa 2 tàu mục tiêu với từng trạm VTS 1, 2, 3, 4 như trong các bảng 2.3, 2.4, 2.5, 2.6.

Bảng 2.3. Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS1

Time		OS				TS			
		D(OS)	PT(OS)	Lat	Long	D(TS)	PT(TS)	Lat	Long
t1	12:02:33	17.739	240.63	20.63243	106.8727	14.407	233.8	20.63417	106.9095
t2	12:03:49	17.338	241.26	20.6357	106.8751	14.531	235.25	20.63617	106.9065
t3	12:04:55	17	241.81	20.63845	106.8772	14.504	236.53	20.63867	106.9049
t4	12:05:56	16.714	242.31	20.64082	106.879	14.51	237.69	20.64133	106.9034
t5	12:06:56	16.41	242.85	20.64333	106.8809	14.493	238.88	20.6435	106.902
t6	12:07:58	16.118	243.39	20.64575	106.8827	14.505	240.06	20.64567	106.9004
t7	12:09:03	15.783	244.04	20.64853	106.8848	14.491	241.36	20.64817	106.899
t8	12:09:52	15.557	244.49	20.65043	106.8863	14.515	242.31	20.65017	106.8977
t9	12:10:43	15.323	244.98	20.6524	106.8878	14.546	243.29	20.65217	106.8963
t10	12:11:41	15.046	245.57	20.65472	106.8896	14.564	244.41	20.65417	106.895
t11	12:12:20	14.855	245.98	20.6563	106.8908	14.583	245.18	20.65567	106.894
t12	12:13:03	14.632	246.49	20.6582	106.8923	14.581	246.13	20.65733	106.8931

Bảng 2.4. Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS2

Time		OS				TS			
		D(OS)	PT(OS)	Lat	Long	D(TS)	PT(TS)	Lat	Long
t1	12:02:33	11.493	286.31	20.63243	106.8727	7.966	295.41	20.63417	106.9095
t2	12:03:49	11.356	288.43	20.6357	106.8751	8.349	295.89	20.63617	106.9065
t3	12:04:55	11.251	290.26	20.63845	106.8772	8.618	297.13	20.63867	106.9049
t4	12:05:56	11.173	291.85	20.64082	106.879	8.875	298.03	20.64133	106.9034
t5	12:06:56	11.099	293.57	20.64333	106.8809	9.13	299.08	20.6435	106.902
t6	12:07:58	11.036	295.25	20.64575	106.8827	9.393	299.93	20.64567	106.9004
t7	12:09:03	10.975	297.2	20.64853	106.8848	9.672	301	20.64817	106.899
t8	12:09:52	10.942	298.54	20.65043	106.8863	9.889	301.6	20.65017	106.8977
t9	12:10:43	10.912	299.94	20.6524	106.8878	10.118	302.17	20.65217	106.8963
t10	12:11:41	10.884	301.6	20.65472	106.8896	10.371	302.92	20.65417	106.895
t11	12:12:20	10.868	302.75	20.6563	106.8908	10.548	303.39	20.65567	106.894
t12	12:13:03	10.858	304.12	20.6582	106.8923	10.753	304.09	20.65733	106.8931

Bảng 2.5. Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS3

Time		OS				TS			
		D(OS)	PT(OS)	Lat	Long	D(TS)	PT(TS)	Lat	Long
t1	12:02:33	9.439	64.37	20.63243	106.8727	13.064	70.89	20.63417	106.9095
t2	12:03:49	9.83	63.11	20.6357	106.8751	12.845	69.47	20.63617	106.9065
t3	12:04:55	10.164	62.13	20.63845	106.8772	12.799	68.03	20.63867	106.9049
t4	12:05:56	10.451	61.32	20.64082	106.879	12.74	66.75	20.64133	106.9034
t5	12:06:56	10.759	60.52	20.64333	106.8809	12.716	65.38	20.6435	106.902
t6	12:07:58	11.058	59.79	20.64575	106.8827	12.676	64.06	20.64567	106.9004
t7	12:09:03	11.405	59	20.64853	106.8848	12.674	62.56	20.64817	106.899
t8	12:09:52	11.642	58.49	20.65043	106.8863	12.65	61.48	20.65017	106.8977
t9	12:10:43	11.89	57.98	20.6524	106.8878	12.627	60.35	20.65217	106.8963
t10	12:11:41	12.185	57.42	20.65472	106.8896	12.629	59.05	20.65417	106.895
t11	12:12:20	12.389	57.06	20.6563	106.8908	12.631	58.16	20.65567	106.894
t12	12:13:03	12.631	56.62	20.6582	106.8923	12.666	57.07	20.65733	106.8931

Bảng 2.6. Tính toán phương vị và khoảng cách của mục tiêu từ trạm VTS4

Time		OS				TS			
		D(OS)	PT(OS)	Lat	Long	D(TS)	PT(TS)	Lat	Long
t1	12:02:33	7.014	123.1	20.63243	106.8727	10.368	110.53	20.63417	106.9095
t2	12:03:49	7.044	119.48	20.6357	106.8751	9.994	109.94	20.63617	106.9065
t3	12:04:55	7.094	116.46	20.63845	106.8772	9.749	108.69	20.63867	106.9049
t4	12:05:56	7.148	113.91	20.64082	106.879	9.518	107.63	20.64133	106.9034
t5	12:06:56	7.223	111.25	20.64333	106.8809	9.302	106.34	20.6435	106.902
t6	12:07:58	7.31	108.74	20.64575	106.8827	9.078	105.12	20.64567	106.9004
t7	12:09:03	7.428	105.94	20.64853	106.8848	8.86	103.53	20.64817	106.899
t8	12:09:52	7.517	104.07	20.65043	106.8863	8.685	102.46	20.65017	106.8977
t9	12:10:43	7.619	102.19	20.6524	106.8878	8.505	101.3	20.65217	106.8963
t10	12:11:41	7.752	100.04	20.65472	106.8896	8.319	99.82	20.65417	106.895
t11	12:12:20	7.851	98.61	20.6563	106.8908	8.193	98.78	20.65567	106.894
t12	12:13:03	7.792	96.95	20.6582	106.8923	8.063	97.32	20.65733	106.8931

Từ vị trí tàu chủ (OS) và tàu mục tiêu (TS) chúng ta xác định được các phương vị và khoảng cách tính từ các trạm VTS 1, 2, 3, 4 đến các mục tiêu như các giá trị trong các bảng trên. Từ các giá trị phương vị và khoảng cách này, áp dụng các công thức trong phần 1.2, thực hiện tính toán CPA và T_{CPA} để đánh giá nguy cơ đâm va.

Bảng 2.7. So sánh giá trị tính toán và giá trị ARPA hiển thị

Time	Δt	PT ₁₂ (^o)	PT ₂₁ (^o)	V _E	Tính toán		ARPA		
					TCPA (m)	CPA (NM)	TCP A (m)	CPA (NM)	
t1	12:02:33		88.8	267.13		0.1	10	0.1	
t2	12:03:49	0.021111	89	269.09	11.27368	9.38	0.1	9.4	0.1
t3	12:04:55	0.018333	89.3	269.52	11.12727	8.40	0.1	8.4	0.1
t4	12:05:56	0.016944	89.7	268.69	11.09508	7.41	0.1	7.4	0.1
t5	12:06:56	0.016667	90	269.52	11.1	6.41	0.1	6.4	0.1
t6	12:07:58	0.017222	90.6	270.3	11.03226	5.41	0.1	5.4	0.1
t7	12:09:03	0.018056	91.3	271.57	11.07692	4.31	0.1	4.3	0.1
t8	12:09:52	0.013611	92.2	271.43	11.16735	3.45	0.1	3.5	0.1
t9	12:10:43	0.014167	93.7	271.61	11.29412	2.57	0.1	2.6	0.1
t10	12:11:41	0.016111	96.8	275.8	10.92414	1.69	0	1.7	0
t11	12:12:20	0.010833	102.6	280.37	11.26154	0.99	0	1	0
t12	12:13:03	0.011944	127.1	306.54	10.71628	0.32	0	0.3	0

Do giá trị vị trí tính toán của 2 tàu không phụ thuộc vào vị trí trạm VTS, các giá trị hướng đi, tốc độ của 2 tàu không thay đổi nên chúng ta chỉ cần tính cho 1 trường hợp. Giá trị tính toán và giá trị quan sát trên ARPA của hệ thống mô phỏng được thể hiện trong các bảng 2.7.

Từ so sánh giá trị CPA , T_{CPA} trong bảng 2.7, chúng ta thấy giá trị tính toán và giá trị do ARPA của Transas hiển thị có sai lệch không lớn. Giá trị của T_{CPA} khác biệt lớn nhất là 0.03 phút, giá trị CPA không khác biệt. Giá trị khác biệt này được cho là hiển thị của ARPA đối với giá trị T_{CPA} đến 0.1 phút. Việc đối sánh giá trị tính toán không làm tròn và giá trị hiển thị của ARPA được làm tròn gây ra sự khác biệt trên.

Tóm lại, chúng ta có thể khẳng định rằng các công thức tính toán được xây dựng ở phần 2.3 cho phép đánh giá nguy cơ đâm va theo CPA và T_{CPA} thỏa mãn các yêu cầu của thiết bị ARPA hiện hành.

2.4. Xác định điểm nóng giao thông trên khu vực luồng Vũng Tàu - Sài Gòn

Như đã nêu ở phần đặt vấn đề, điểm nóng giao thông là nơi có nguy cơ cao xảy ra tai nạn đâm va trên tuyến luồng. Để đảm bảo an toàn, các điểm nóng này cần xác định để làm cơ sở cho việc đánh giá các nguy cơ đâm va giữa các tàu thuyền khi gặp nhau tại các điểm này. Việc xác định điểm nóng giao thông được thực hiện qua việc hỏi ý kiến chuyên gia - các hoa tiêu thường ngày làm việc trên tuyến luồng này.

Xuất phát từ hải đồ của Hải quân nhân dân Việt Nam tỷ lệ 1:100000 tại vĩ tuyến 16^0 về sơ đồ, vị trí các điểm có nguy cơ mất an toàn giao thông trên tuyến luồng Vũng Tàu - Sài Gòn, nghiên cứu sinh đã xác định 10 vị trí được đánh giá là có nguy cơ trên tuyến luồng này. Chi tiết 10 điểm như sau:

- Khu vực phao số 0
- Khu vực phao GR
- Khúc cua An Thạnh (Tiêu 31)
- Khúc cua Coude Lest (Tiêu 38)
- Khúc cua Houlede Lest (Phao 40)
- Khúc cua Đá Hàn (Phao 48)
- Mũi Ô rơ (Phao 55)
- Ngã ba Nhà Bè (Phao 58)
- Khúc cua Mũi đèn đỏ
- Khúc cua Hải lý 3 (Phao 68)

Để đánh giá chi tiết cụ thể hơn, nghiên cứu sinh đã xây dựng bảng hỏi các hoa tiêu hoạt động trên tuyến luồng này về đánh giá nguy cơ đối với 10 điểm nêu trên, đồng thời bổ sung các điểm khác (*nếu có*). Chi tiết như sau:

1. Xin vui lòng đánh giá nguy cơ đâm va giữa các tàu thuyền tại các địa điểm sau:

Điểm có nguy cơ mất an toàn giao thông	Nguy cơ đâm va		
	Cao	Trung bình	Thấp
Khu vực phao số 0			
Khu vực phao GR			
Khúc cua An Thạnh (Tiêu 31)			
Khúc cua Coude Lest (Tiêu 38)			
Khúc cua Houle Lest (Phao 40)			
Khúc cua Đá Hàn (Phao 48)			
Mũi Ô rơ (Phao 55)			
Ngã ba Nhà Bè (Phao 58)			
Khúc cua Mũi đèn đỏ			
Khúc cua Hải lý 3 (Phao 68)			

2. Xin vui lòng bổ sung các điểm có nguy cơ đâm va giữa các tàu thuyền khác (nếu có):

Điểm có nguy cơ mất an toàn giao thông	Nguy cơ đâm va		
	Cao	Trung bình	Thấp

Hình 2.5. Câu hỏi trong phiếu khảo sát về đánh giá những điểm có nguy cơ mất an toàn giao thông trên tuyến luồng Vũng Tàu - Sài Gòn

Kết quả khảo sát 60 hoa tiêu khu vực I và hoa tiêu Cát Lái như sau:

Điểm có nguy cơ mất an toàn giao thông	Nguy cơ đâm va		
	Cao	Trung bình	Thấp
Khu vực phao số 0	20	29	11
Khu vực phao GR	8	34	17
Khúc cua An Thạnh (Tiêu 31)	21	29	10
Khúc cua Coude Lest (Tiêu 38)	32	26	3
Khúc cua Houle Lest (Phao 40)	36	19	5
Khúc cua Đá Hàn (Phao 48)	33	24	5
Mũi Ô rơ (Phao 55)	14	35	11
Ngã ba Nhà Bè (Phao 58)	22	28	10
Khúc cua Mũi đèn đỏ	23	17	6
Khúc cua Hải lý 3 (Phao 68)	19	21	4

Hình 2.6. Kết quả đánh giá của hoa tiêu với 10 điểm có nguy cơ đâm va trên tuyến luồng Vũng Tàu - Sài Gòn

Theo kết quả như trên, mặc dù đánh giá theo từng điểm có khác nhau, nhưng đa số các hoa tiêu đều xác định 10 điểm nêu trên đều là điểm có nguy cơ đâm va. Bên cạnh đó, các hoa tiêu còn đánh giá thêm 25 điểm có nguy cơ đâm va. Danh sách các điểm bổ sung như sau:

- Bến phà Bình Khánh
- Cầu Phước Khánh
- Ngã 3 Vàm Láng
- Khu vực phao 49 (luồng Soài Rạp)
- Khu vực sông Đông Tranh
- Khu vực bến phà Cát Lái
- Bến phà Phước Khánh (luồng Soài Rạp)
- Khúc cua sông 4 mặt
- Cầu Bình Khánh
- Thiêng Liêng (HL32)
- Dàn Xây (HL24)
- Khu vực cầu Phú Mỹ
- Khu vực cầu K15-k13
- Khu vực vòng quay trở VQ1 và VQ2
- Cảng Calofic
- Cảng Sài Gòn Hiệp Phước
- Vùng neo Soài Rạp
- Phao đỏ 28
- Đầu đáy hạ phao xanh
- Khúc cua Lombard
- Khu vực đón trả hoa tiêu 1, 2
- Mom hải lý 29 (gần ngã tư)
- Propotis
- Phao 34
- Phao 14

Tuy nhiên, việc bổ sung điểm này không đạt được sự đồng thuận cao như 10 điểm nêu trên. Không điểm nào trong số 25 điểm bổ sung có được 25% ý kiến đồng thuận. Do đó, 25 điểm này không được xác định là điểm nóng giao thông trên luồng. Các điểm nóng giao thông được xác định là 10 điểm ban đầu. Tàu thuyền nên tránh gặp nhau tại 10 điểm này. Nếu 2 tàu gặp nhau tại 10 điểm nóng giao thông này, có khả năng xảy ra nguy cơ đâm va giữa chúng.

Để có thể đánh giá khả năng gặp nhau tại các điểm nóng giao thông, cần đối chiếu, so sánh dự kiến thời gian tàu đến các điểm nóng đó. Đối chiếu so sánh *ETA* của 2 tàu với cùng một điểm nóng nhằm đánh giá khả năng 2 tàu gặp nhau ở điểm đó hay không.

Đánh giá nguy cơ đâm va giữa các tàu thuyền, ngoài việc sử dụng thông tin *CPA* và T_{CPA} , *ETA*, khi dẫn tàu người ta có thể sử dụng các biện pháp khác. Để làm rõ nội dung này, trong Phiếu thăm dò gửi các hoa tiêu khu vực Vũng Tàu - Sài Gòn, 3 câu hỏi được thiết kế để kiểm tra. Cụ thể như sau:

2. Theo ông, thông tin nào giúp đánh giá sớm nguy cơ đâm va trên luồng?

A. <i>ETA</i> của các tàu đến các điểm có nguy cơ đâm va nêu ở phần II	
B. <i>CPA</i> và T_{CPA} hoặc phương vị, khoảng cách của tàu mục tiêu	
C. Thông tin khác (đề nghị mô tả):	
.....	
.....	

Hình 2.7. Câu hỏi thăm dò các phương pháp xác định nguy cơ đâm va trên luồng

Kết quả phân hồi nhận được như sau:

THÔNG TIN GIÚP ĐÁNH GIÁ NCDV	
A. <i>ETA</i> của các tàu đến các điểm có nguy cơ đâm va nêu ở phần II	46
B. <i>CPA</i> và T_{CPA} hoặc phương vị, khoảng cách của tàu mục tiêu	43
C. Thông tin khác	4

Hình 2.8. Kết quả thăm dò về đánh giá nguy cơ đâm va của hoa tiêu

Trong khi trả lời Phiếu thăm dò, có một số hoa tiêu không trả lời các câu hỏi này. Đa phần các hoa tiêu vẫn sử dụng *ETA* (46/60), *CPA*, T_{CPA} (43/60) là các

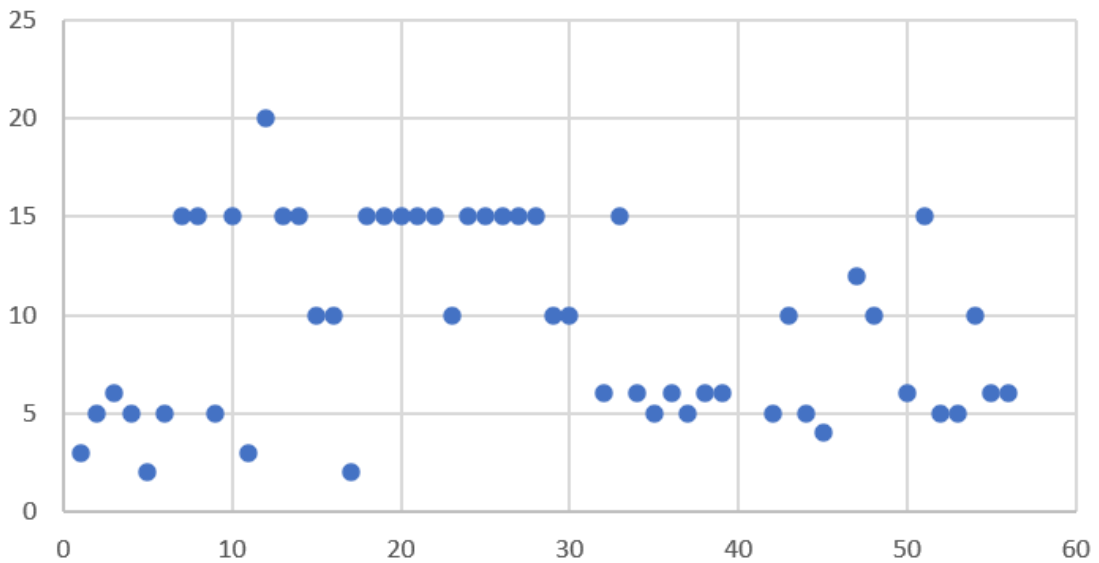
phương pháp đánh giá nguy cơ đâm va của họ khi dẫn tàu trên luồng. Có 04 ý kiến mô tả thêm thông tin giúp đánh giá nguy cơ đâm va trên luồng Vũng Tàu - Sài Gòn dựa vào thông tin về sự di chuyển của các tàu đặc biệt (tàu hạn chế khả năng điều động, tàu không chế bởi mớn nước) hoặc thông tin về điểm gặp nhau của các tàu trên thiết bị chuyên dụng hoặc thông tin do radar cung cấp. 04 ý kiến bổ sung này thực chất vẫn là xác định, đối sánh *ETA* của các tàu trên tuyến luồng. Vì vậy, có thể khẳng định việc xác định nguy cơ đâm va của hoa tiêu sử dụng theo 2 phương pháp:

- Sử dụng *ETA* của tàu thuyền đến các điểm nóng giao thông;
- Sử dụng các thông tin đồ giải radar: *CPA* và *T_{CPA}*.

Đối với việc đánh giá nguy cơ đâm va sử dụng *CPA* và *T_{CPA}*, đây là phương pháp phổ biến của người đi biển. Các tiêu chí đánh giá đã được định lượng rõ ràng. Đối với việc đánh giá nguy cơ đâm va bằng *ETA*, thông tin này mặc dù được sử dụng nhưng chưa thống nhất ý kiến. Để làm rõ nội dung này, phương pháp chuyên gia tiếp tục được sử dụng với câu hỏi: “*Theo ông, giá trị ETA của 2 tàu đến các điểm có nguy cơ đâm va nêu ở phần II chênh lệch là bao nhiêu phút thì có thể đánh giá là có nguy cơ đâm va?*”. Kết quả thu được như sau:

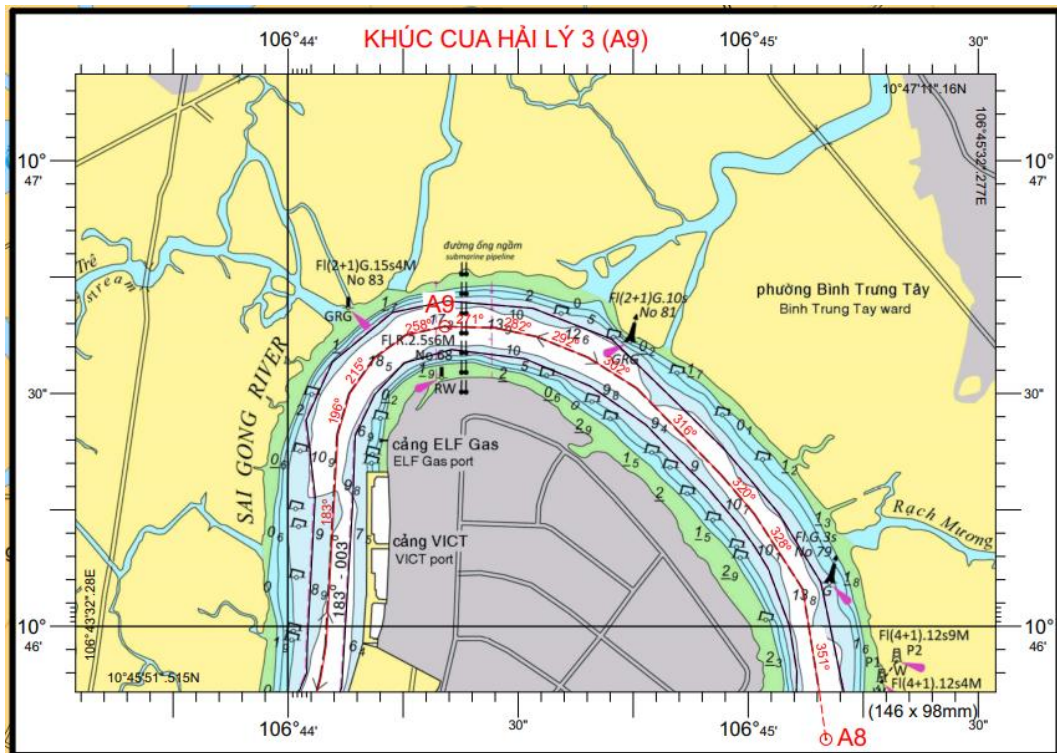
- Có 51/60 người được hỏi trả lời câu hỏi này.
- Giá trị độ chênh *ETA* của 2 tàu thuyền đến 1 điểm tương đối phân tán giữa các hoa tiêu. Có 54 hoa tiêu có ý kiến trả lời về nội dung này. Trong khi đa phần cho ý kiến độ chênh *ETA* của 2 tàu thuyền đến điểm nóng giao thông rơi vào khoảng 5 đến 6 phút tương ứng với khoảng cách tầm 1NM là giá trị hạn mức để xác định có nguy cơ đâm va thì một số ý kiến cho rằng độ chênh *ETA* lớn hơn nữa đã có thể xác định là có nguy cơ đâm va. Giá trị trung bình tầm 9,4 phút.

Nguy cơ đâm va theo chênh ETA

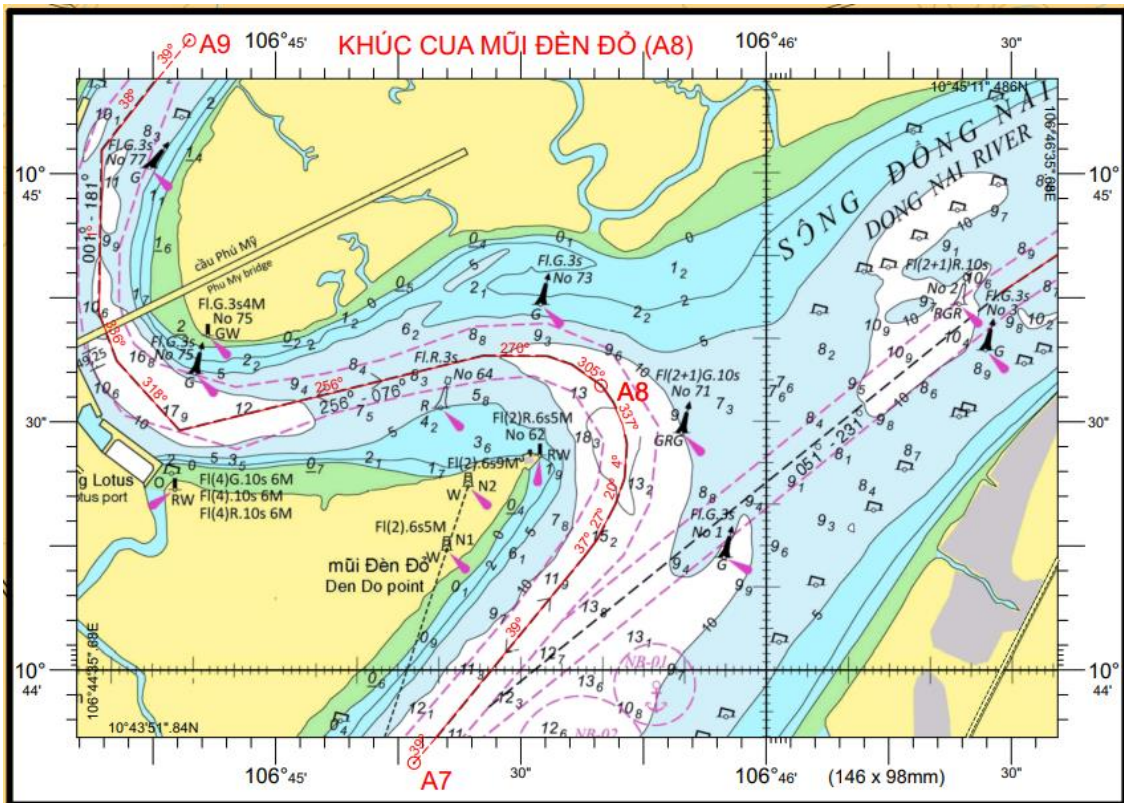


Hình 2.9. Ý kiến xác định nguy cơ đâm va theo độ chênh ETA đến 1 điểm

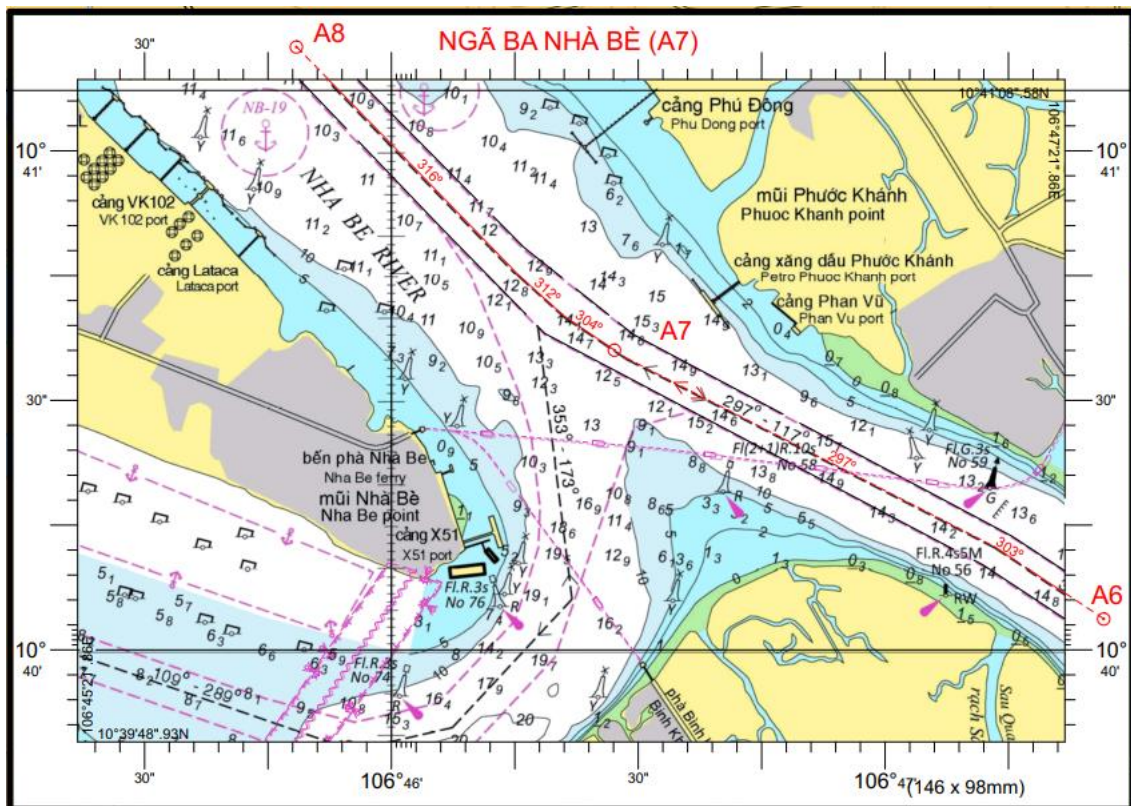
Nếu tính theo số phút chênh nhau của *ETA* giữa 2 tàu, thông tin này khó đưa ra được cảm nhận trực quan như khi so sánh khoảng cách giữa các tàu tại thời điểm 1 tàu đã đến điểm nóng giao thông, 1 tàu còn cách điểm đó một khoảng tương ứng với độ chênh *ETA*. Để đánh giá kỹ hơn, chúng ta cần xem xét địa hình một số điểm nóng sau đây:



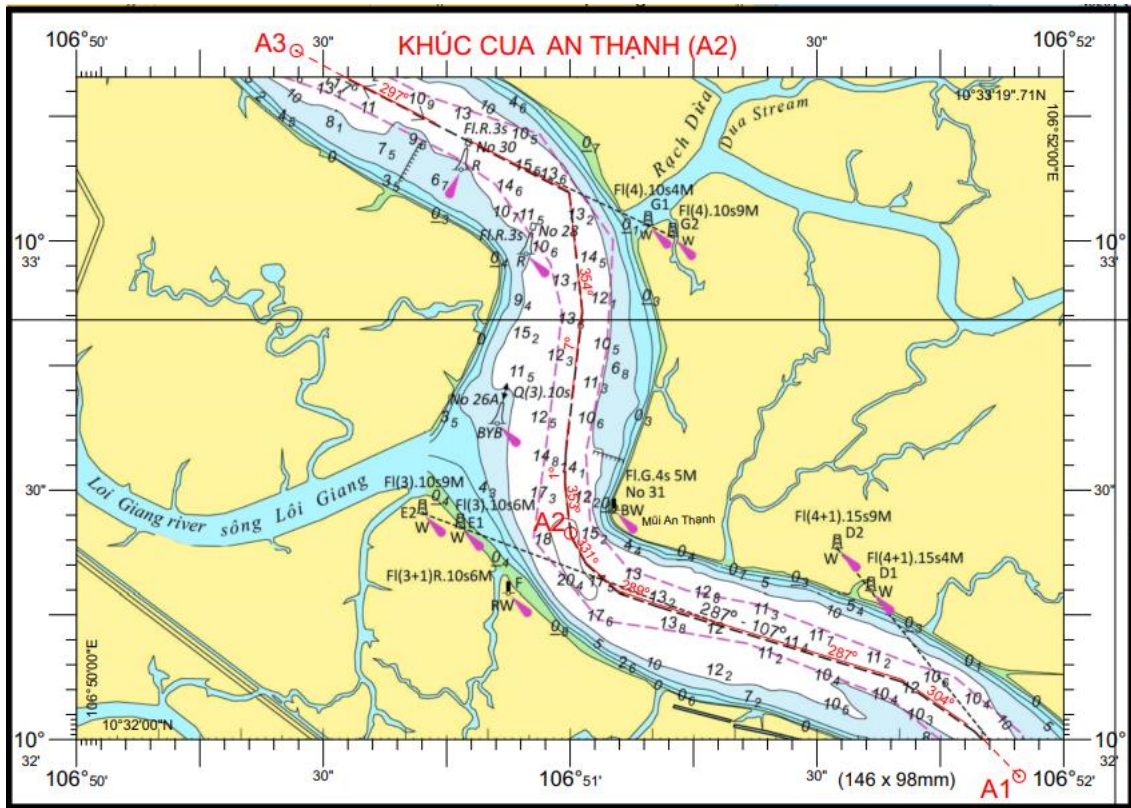
Hình 2.10. Khúc của Hải Lý



Hình 2.11. Khúc của Mũi Đền Đỏ



Hình 2.12. Ngã ba Nhà Bè



Hình 2.15. Khúc cua An Thạnh

Theo dõi các khúc cua trên thực tế ta thấy, khi độ chênh *ETA* khoảng 5 phút tương ứng với khoảng cách 1NM giữa 2 tàu thì vẫn khả năng 2 tàu thuyền gặp nhau trên đoạn luồng khó khăn cho việc tránh va. Trong khi đó, độ chênh khoảng cách 2NM (ứng với độ chênh *ETA* khoảng 10 phút) có thể cho phép 2 tàu không gặp nhau trên các khúc cua này. Nếu lựa chọn độ chênh *ETA* lớn hơn thì nguy cơ đâm va sẽ được đánh giá sớm hơn đồng nghĩa với việc cảnh báo nhiều hơn. Nếu cảnh báo nhiều quá sẽ có thể dẫn tới quá tải cho hệ thống VTS. Vì vậy, chọn độ chênh *ETA* khoảng 10 phút làm giới hạn đánh giá nguy cơ đâm va (cũng gần với giá trị trung bình mà các hoa tiêu đưa ra) là một lựa chọn hợp lý.

2.5. Xây dựng công thức tính toán xác định nguy cơ đâm va theo *ETA* tới các điểm nóng giao thông trên tuyến luồng Vũng Tàu - Sài Gòn



Hình 2.16. Luồng Vũng Tàu - Sài Gòn

Tuyến luồng Vũng Tàu - Sài Gòn là tuyến luồng dài, quanh co. Các tàu thuyền phải di chuyển theo hình dạng của tuyến luồng là tổng hợp của nhiều đoạn di chuyển ngắn qua nhiều điểm chuyển hướng khác nhau. Khoảng cách và hướng di chuyển giữa 2 điểm được tính theo công thức dưới đây.

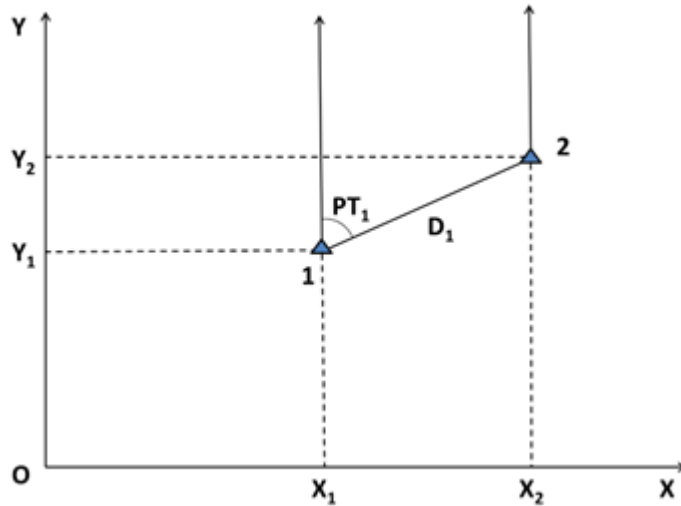
Khi quan sát dữ liệu vị trí điểm 1 và 2 ta thấy chúng có các tọa độ như sau:

Điểm 1 ($X_1; Y_1$)

Điểm 2 ($X_2; Y_2$)

- $X_1; X_2$ là các giá trị vĩ độ của 2 điểm;
- $Y_1; Y_2$ là các giá trị kinh độ của 2 điểm.

Nếu $X_2 > X_1, Y_2 > Y_1$ (hình 2.17)



Hình 2.17. Trường hợp tàu đi từ điểm 2 ở cung phần tư thứ nhất so với điểm 1

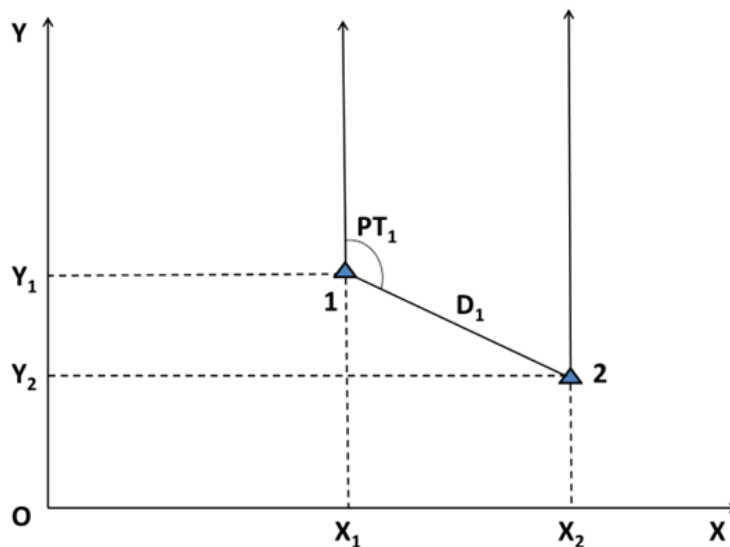
Khoảng cách từ điểm 1 đến điểm 2:

$$D_1 = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (2.1)$$

Hướng từ điểm 1 đến điểm 2 trong trường hợp này phải thỏa mãn $0^\circ \leq PT_1 \leq 90^\circ$:

$$PT_1 = \arctan\left(\frac{X_2 - X_1}{Y_2 - Y_1}\right) \quad (2.2)$$

Nếu $X_2 > X_1, Y_2 < Y_1$ (hình 2.18)



Hình 2.18. Trường hợp điểm 2 ở cung phần tư thứ hai so với điểm 1

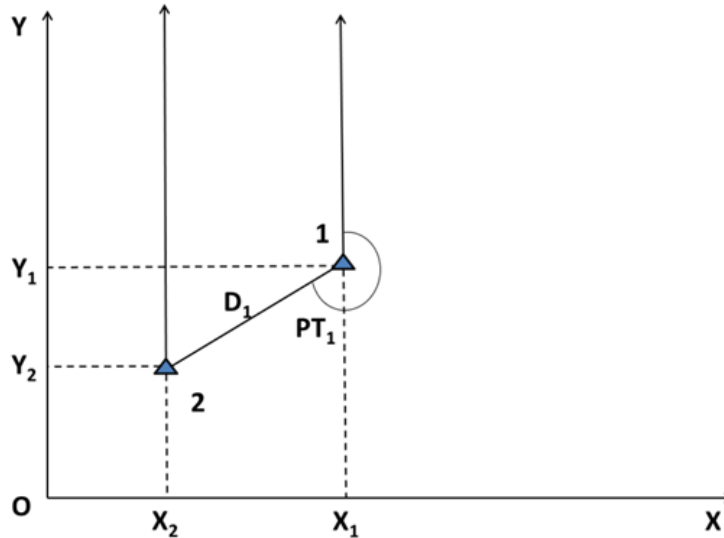
Khoảng cách từ điểm 1 đến điểm 2:

$$D_1 = \sqrt{(X_2 - X_1)^2 + (Y_1 - Y_2)^2} \quad (2.3)$$

Hướng từ điểm 1 đến điểm 2 trong trường hợp này phải thỏa mãn $90^\circ < PT_1 \leq 180^\circ$:

$$PT_1 = 180^\circ - \arctan\left(\frac{X_2 - X_1}{Y_1 - Y_2}\right) \quad (2.4)$$

Nếu $X_2 < X_1, Y_2 < Y_1$ (hình 2.19)



Hình 2.19. Trường hợp điểm 2 ở cung phần tư thứ ba so với điểm 1

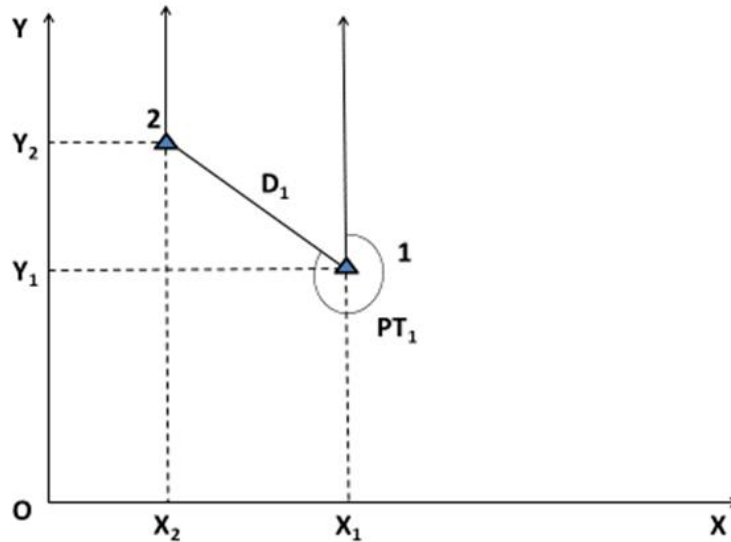
Khoảng cách từ điểm 1 đến điểm 2:

$$D_1 = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \quad (2.5)$$

Hướng từ điểm 1 đến điểm 2 trong trường hợp này phải thỏa mãn $180^\circ < PT_1 \leq 270^\circ$:

$$PT_1 = 180^\circ + \arctan\left(\frac{X_1 - X_2}{Y_1 - Y_2}\right) \quad (2.6)$$

Nếu $X_2 < X_1, Y_2 > Y_1$ (hình 2.20)



Hình 2.20. Trường hợp điểm 2 ở cung phần tư thứ tư so với điểm 1

Khoảng cách từ điểm 1 đến điểm 2:

$$D_1 = \sqrt{(X_1 - X_2)^2 + (Y_2 - Y_1)^2} \quad (2.7)$$

Hướng từ điểm 1 đến điểm 2 trong trường hợp này phải thỏa mãn $270^\circ < PT_1 < 360^\circ$:

$$PT_1 = 360^\circ - \arctan\left(\frac{X_1 - X_2}{Y_2 - Y_1}\right) \quad (2.8)$$

Từ công thức trên ta tính được giá trị hướng và khoảng cách giữa 2 điểm. Áp dụng công thức này ta có thể tính được hướng đi, quãng đường di chuyển của tàu thuyền sau 2 lần quan sát, từ đó tính được tốc độ tàu theo công thức:

$$v = \frac{D_1}{t_2 - t_1} \quad (2.9)$$

Công thức nêu trên cũng cho phép ta tính toán được hướng đi và quãng đường giữa 2 điểm chuyển hướng. Tổng hợp lại ta tính toán được quãng đường phải di chuyển từ vị trí tàu đến điểm cần đến:

$$S = \sum D_i \quad (2.10)$$

Thời gian dự kiến đến điểm cần đến được tính như sau:

$$ETA = \frac{S}{v} \quad (2.11)$$

Với các hệ công thức nêu trên ta tính toán được ETA của các tàu đến từng điểm nóng giao thông. So sánh ETA của các tàu tới các điểm nóng giao thông ta xây dựng được cơ sở dữ liệu để đánh giá nguy cơ đâm va giữa các tàu thuyền trên luồng Vũng Tàu - Sài Gòn phục vụ cho huấn luyện AI.

2.6. Kết luận chương 2

Trong chương 2, nguy cơ đâm va giữa các tàu thuyền đang di chuyển trên luồng Vũng Tàu - Sài Gòn cũng như tại khu vực đầu luồng được xác định theo 2 điều kiện: một là tính toán theo CPA và T_{CPA} ; hai là theo ETA đến các điểm nóng giao thông. Các công thức tính toán đã được xác định qua việc giải các bài toán hình học và kiểm tra so sánh với kết quả mô phỏng tại hệ thống mô phỏng buồng lái của Trường Đại học Hàng hải Việt Nam. Đối sánh cho thấy các hệ công thức đảm bảo để xây dựng cơ sở dữ liệu về nguy cơ đâm va trên tuyến luồng Vũng Tàu - Sài Gòn phục vụ huấn luyện các mô hình học máy.

- Làn quan sát t1:
 - Vị trí tàu 1: pa1, la1
 - Vị trí tàu 2: pb1, lb1
- Làn quan sát 2:
 - Vị trí tàu 1: pa2, la2
 - Vị trí tàu 2: pb2, lb2
- y (nguy cơ đâm va - ncdv): 0/1
 - 0 - Hai tàu không có nguy cơ đâm va
 - 1 - Hai tàu có nguy cơ đâm va

Miền xác định:

- $\Delta t = t_2 - t_1$: [1, 5]
- pa1, pb1, pa2, pb2: [10, 12] - vĩ độ giới hạn khu vực Sài Gòn - Vũng Tàu
- la1, lb1, la2, lb2: [106, 108] - kinh độ giới hạn khu vực Sài Gòn - Vũng Tàu
- ncdv: 0, 1

Trên cơ sở xác định dữ liệu đầu vào và đầu ra của hệ thống, giới hạn miền xác định ta tiến hành xây dựng bộ cơ sở dữ liệu của các tình huống có/không có nguy cơ đâm va giữa các tàu thuyền.

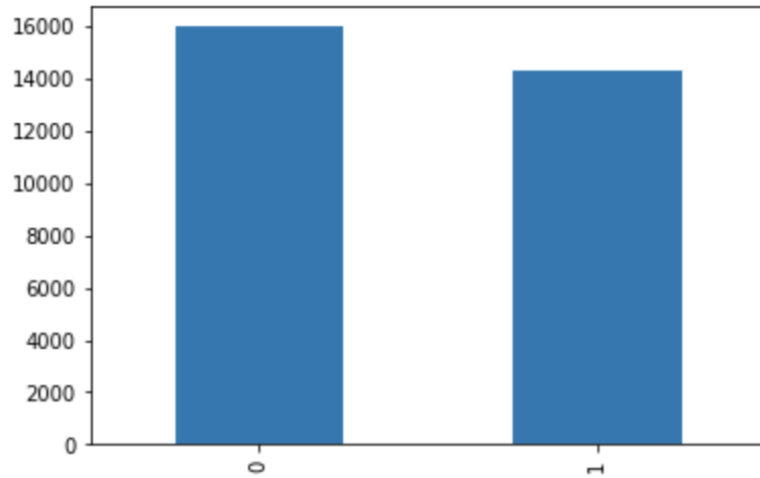
3.2. Đánh giá bộ dữ liệu phục vụ huấn luyện mô hình trên vùng biển rộng

Trước khi huấn luyện mô hình, bước phân tích và tiền xử lý dữ liệu được tiến hành. Dữ liệu huấn luyện gồm 30.240 mẫu, Bảng 3.1 thể hiện 5 mẫu dữ liệu đầu.

Bảng 3.1. Dữ liệu ban đầu

t1	pa1	la1	pb1	lb1	t2	pa2	la2	pb2	lb2	ncdv
0	11°30'N	106°30'E	11°50'N	106°30'E	1	11°30.2'N	106°30'E	11°49.767'N	106°30'E	0
0	11°30'N	106°30'E	11°50'N	106°30'E	2	11°30.433'N	106°30'E	11°49.533'N	106°30'E	0
0	11°30'N	106°30'E	11°50'N	106°30'E	3	11°30.7'N	106°30'E	11°49.3'N	106°30'E	1
0	11°30'N	106°30'E	11°50'N	106°30'E	4	11°31'N	106°30'E	11°49.067'N	106°30'E	1
0	11°30'N	106°30'E	11°50'N	106°30'E	5	11°31.333'N	106°30'E	11°48.833'N	106°30'E	1

Độ chênh lệch giữa các output được thống kê trong Hình 3.1 thì số lượng các tình huống không có nguy cơ đâm va nhiều hơn các tình huống có nguy cơ đâm va.



Hình 3.1. Độ chênh lệch giữa các output

Chuyển các dữ liệu kinh độ, vĩ độ sang dạng thập phân:

$$x = \text{độ} + \text{phút} / 60$$

Chuẩn hóa dữ liệu:

$$z = \frac{x - u}{s}$$

trong đó: u, s là giá trị trung bình và độ lệch chuẩn.

Bảng 3.2. Một số mẫu dữ liệu sau khi chuẩn hóa

t1	pa1	la1	pb1	lb1	t2	pa2	la2	pb2	lb2	ncdv
0.0	1.35	-1.4	2.1	-1.6	-1.41	1.37	-1.43	2.09	-1.4	0
0.0	1.35	-1.4	2.1	-1.6	-0.71	1.38	-1.43	2.08	-1.4	0
0.0	1.35	-1.4	2.1	-1.6	0.0	1.39	-1.43	2.07	-1.4	1
0.0	1.35	-1.4	2.1	-1.6	0.71	1.4	-1.43	2.06	-1.4	1
0.0	1.35	-1.4	2.1	-1.6	1.41	1.42	-1.43	2.05	-1.4	1

Bảng 3.2 thể hiện 5 mẫu dữ liệu đầu sau khi chuẩn hoá (đã làm tròn với độ chính xác hai chữ số sau dấu phẩy). Giá trị trung bình và độ lệch chuẩn sau khi chuẩn hoá dữ liệu được thể hiện trong Bảng 3.3.

Bảng 3.3. Giá trị trung bình và độ lệch chuẩn sau khi đã chuẩn hóa dữ liệu

	t1	pa1	la1	pb1	lb1	t2	pa2	la2	pb2	lb2
mean	0.0	- 6.3685324 28338907e -14	1.50907556 5860808e- 13	- 8.5056631 05057557e -15	- 1.19106370 85663809e- 14	0.0	- 9.1697143 3381837e- 14	- 3.82640751 82914356e- 14	- 6.22976005 73194164e- 15	1.5090755 65860808e -13
std	0.0	1.0000165 34801755	1.00001653 48017875	1.0000165 348016243	1.00001653 48016203	1.0000 165348 016248	1.0000165 348014514	1.00001653 48016228	1.00001653 48015708	1.0000165 348017875

Để huấn luyện mô hình, ta chia dữ liệu thành 3 tập: training set, validation set và test set theo tỉ lệ 3:1:1 (đây là tỷ lệ chia theo kinh nghiệm thông thường khi huấn luyện model khi bộ cơ sở dữ liệu không lớn lắm).

- Training set: (18.144, 10), (18.144)
- Validation set: (6.048, 10), (6.048)
- Test set: (6.048, 10), (6.048)

3.3. Xây dựng mô hình mạng nơ-ron để huấn luyện

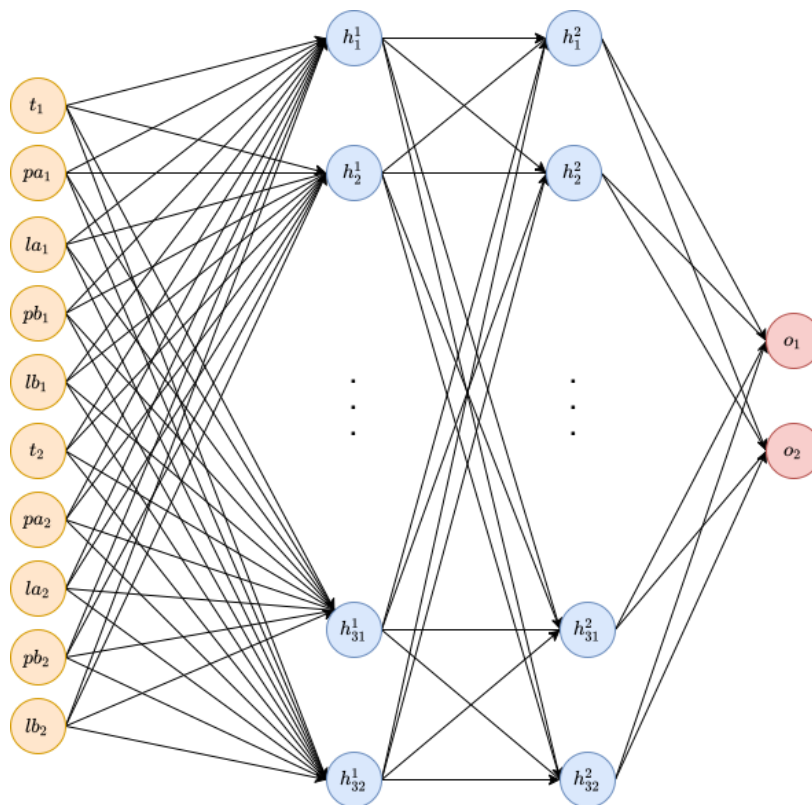
Mạng nơ-ron nhân tạo (Artificial Neural Networks) , đây là một chuỗi những thuật toán được đưa ra để tìm kiếm các mối quan hệ cơ bản trong tập hợp các dữ liệu. Thông qua việc bắt bước cách thức hoạt động từ não bộ con người. Nói cách khác, mạng nơ-ron nhân tạo được xem là hệ thống của các tế bào thần kinh nhân tạo (neuron). Đây thường có thể là hữu cơ hoặc nhân tạo về bản chất. Mạng Neural Network là sự kết hợp của những tầng perceptron hay còn gọi là perceptron đa tầng (Multilayer Perceptron - MLP). Mỗi tầng bao gồm một tập hợp các neuron và các kết nối đến các neuron của tầng tiếp theo. Mỗi một mạng Neural Network thường bao gồm 3 kiểu tầng là:

- Tầng input layer (tầng vào): Tầng này nằm bên trái cùng của mạng, thể hiện cho các đầu vào của mạng. Số lượng neuron tầng input thường bằng với kích thước của dữ liệu đầu vào (ví dụ, số phần tử trong véc tơ đầu vào)
- Tầng output layer (tầng ra): Là tầng bên phải cùng và nó thể hiện cho những đầu ra của mạng. Số lượng neuron của tầng đầu ra thường

bằng với số lớp mà mạng cần dự đoán (prediction) hoặc phân loại (classification). Trong bài toán phát hiện tàu va chạm thì có thể có hai nơ ron đầu ra: một nơ ron biểu diễn xác suất nguy cơ va chạm, nơ ron đầu ra còn lại biểu diễn xác suất không có nguy cơ va chạm.

- Tầng hidden layer (tầng ẩn): Tầng này nằm giữa tầng vào và tầng ra nó thể hiện cho quá trình suy luận logic của mạng. Thường bao gồm các trọng số (learnable parameters) được điều chỉnh thông qua quá trình huấn luyện mạng.

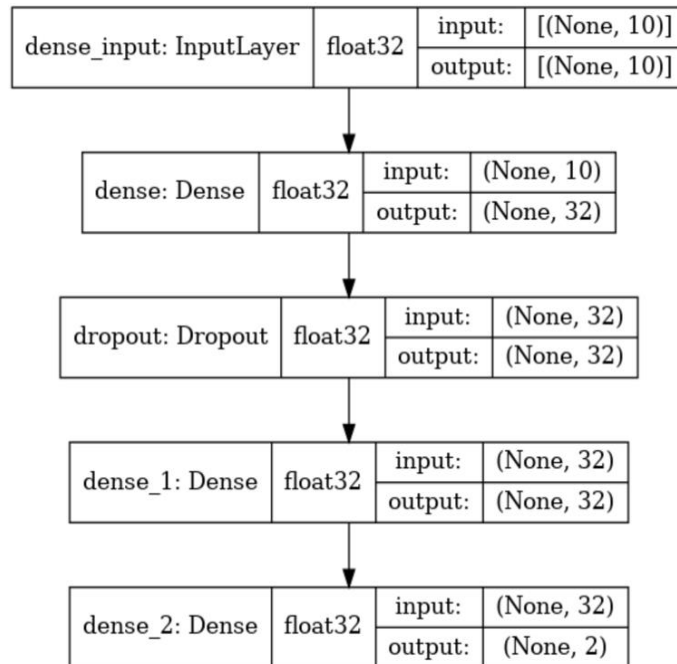
Trong bài toán xác định nguy cơ đâm va tàu thuyền trên vùng biển rộng, cụ thể là vùng đầu luồng tại cửa Vũng Tàu. Qua một số thử nghiệm đơn giản (pilot study) trên tập dữ liệu này. Nghiên cứu sinh lựa chọn kiến trúc mạng nơ-ron nhân tạo đơn giản nhưng khá hiệu quả. Kiến trúc mạng nơ-ron được thể hiện như hình 3.2 dưới đây:



Hình 3.2. Kiến trúc mô hình mạng nơ-ron

Do dữ liệu có số lượng đặc trưng (input features) và số lượng mẫu (training samples) ít nên mạng multi-layer feed-forward network được lựa chọn để huấn luyện, sơ đồ mạng được thể hiện trong hình 3.2.

Do dữ liệu có số lượng đặc trưng (input features) và số lượng mẫu (training samples) ít nên mạng multi-layer feed-forward network [5] được lựa chọn để huấn luyện, sơ đồ mạng được thể hiện trong hình 3.3.



Hình 3.3. Thông số chi tiết của mô hình

Cấu trúc mạng bao gồm 4 layers:

- 1 input layer: gồm 10 nơ ron, biểu diễn véc tơ gồm 10 đặc trưng của dữ liệu đầu vào.
- 2 tầng ẩn (hidden layers): mỗi tầng gồm 32 nơ ron, biểu diễn các đặc trưng học từ dữ liệu dưới dạng các trọng số (weight). Kích thước hai lớp hidden layers được chọn dựa trên thực nghiệm trong khoảng 16 - 64. Thực nghiệm cho thấy huấn luyện mạng với kích thước hai lớp hidden layers 16, mô hình không đạt được độ chính xác cao (xảy ra hiện tượng underfitting), với kích thước [32, 64) bắt đầu xảy ra hiện tượng overfitting). Mạng đạt kết quả tốt nhất với kích thước hai lớp hidden layers là 32. Thông số chi tiết của mô hình sau khi tối ưu được mô tả chi tiết trong hình 3.3. Các trọng số trong 2 lớp ẩn này chứa các thông tin từ chi tiết cho đến khái quát về dữ liệu trong quá trình huấn luyện. Các nhiều dữ liệu được đưa vào huấn luyện thì các trọng số này càng chứa các thông tin có tính khái quát cao (generalization).

- Lớp đầu ra (output layer), gồm hai nơ ron đầu ra: một nơ ron biểu diễn xác suất nguy cơ va chạm, nơ ron đầu ra còn lại biểu diễn xác suất không có nguy cơ va chạm.

Giữa hai hidden layer sử dụng dropout [6] để giảm chi phí tính toán đồng thời làm tăng khả năng tổng quát hoá của mô hình và tránh hiện tượng overfitting [7]. Mô hình được tối ưu sử dụng thuật toán Adam [8], và hàm mục tiêu (objective function/ loss function): Binary Cross Entropy (BCE):

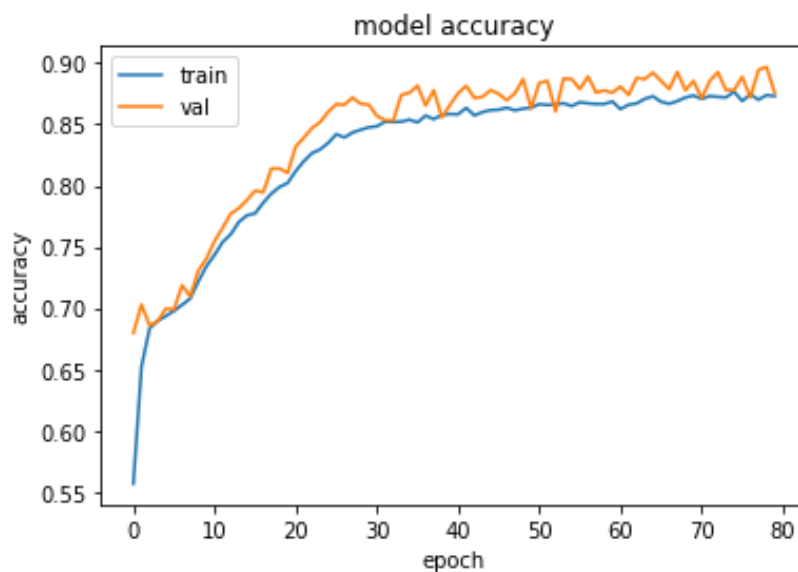
$$L_{BCE} = -\frac{1}{n} \sum_{i=1}^n (y_i * \log y'_i + (1 - y_i) * \log(1 - y'_i))$$

trong đó: n - là số lượng mẫu, y_i là đầu ra thực tế (y_true), y'_i là đầu ra của mô hình (y_predicted). Quá trình huấn luyện sẽ ước lượng các tham số của mô hình nhằm mục tiêu tối thiểu hóa hàm loss.

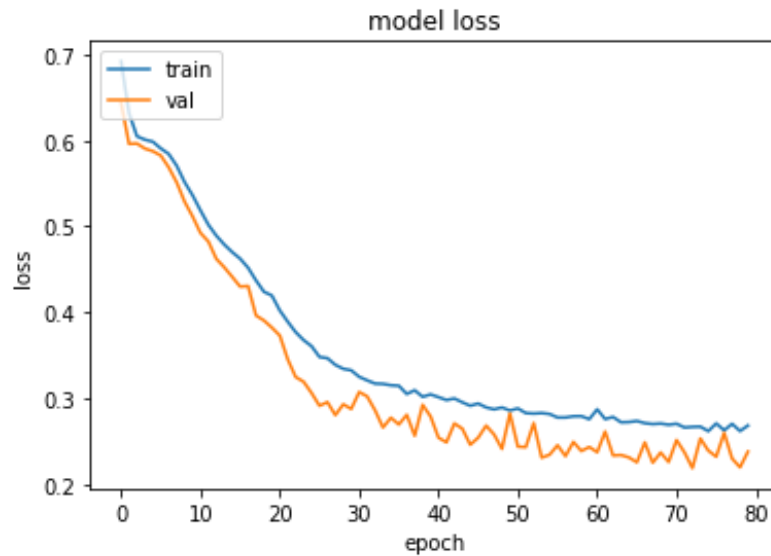
Như vậy tổng số tham số của mạng sẽ là: $10 \times 32 \times 32 \times 2 = 20.480$ tham số.

Chi tiết quá trình huấn luyện được thể hiện trong Phụ lục 2. Thực nghiệm chạy mô hình cho thấy không cần chạy đến epoch thứ 80 thì mạng hội tụ và dừng quá trình huấn luyện.

3.4. Đánh giá kết quả huấn luyện mô hình mạng nơ ron nhân tạo trên vùng biên rộng

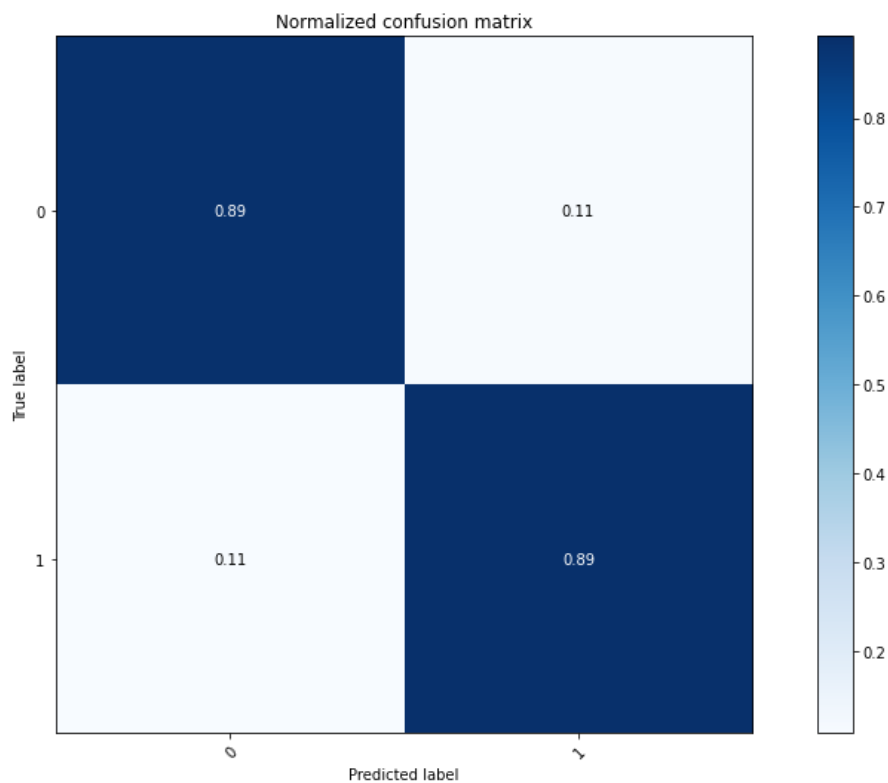


Hình 3.4. Độ chính xác của mô hình trên tập training và validation



Hình 3.5. Hàm loss trên tập training và validation

Sau khi training mô hình cho kết quả tốt trên tập validation set với accuracy ~ 90%. Chi tiết thông tin về Accuracy và Loss được ghi lại trong hình 3.4 và 3.5. Thử nghiệm mô hình trên tập test set cho kết quả tốt với Accuracy ~ 85%.



Hình 3.6. Normalized confusion matrix trên tập test set

Kết quả trên test set:

189/189 [=====] - 0s 1ms/step - loss: 0.3516 - accuracy: 0.8515

Precision, Recall và F1 (macro) trên tập test set: 0.8883345859055505, 0.8895823013058037, 0.8888411682892907

Confusion matrix trên tập test set như trong hình 3.6.

Phân tích dữ liệu đánh giá nguy cơ đâm va trên tập test set sau khi được chuẩn hóa nêu trên ta thấy giá trị dự đoán nguy cơ đâm va/không có nguy cơ đâm va có độ chính xác đến gần 90%. Với bộ dữ liệu hiện hành, số lượng hơn 30000 dữ liệu được coi là tương đối nhỏ so với các cơ sở dữ liệu huấn luyện AI. Kết quả trên là tương đối tốt với các dự đoán của máy tính. Hơn thế nữa, đánh giá nguy cơ đâm va không chỉ qua 2 lần quan sát, nó là một chuỗi các lần quan sát liên tục cho nên khả năng dự đoán nguy cơ đâm va sẽ càng tăng cường lên sau nhiều lần quan sát và qua đó tránh được các hiện tượng bỏ lọt nguy cơ đâm va giữa các tàu thuyền. Độ chính xác như trên của mô hình dự đoán nguy cơ đâm va hoàn toàn có thể chấp nhận được.

3.5. Kết luận chương 3

Trong phần này, việc nghiên cứu đánh giá nguy cơ đâm va dựa trên giá trị quan sát vị trí của 2 tàu thuyền trên biển sử dụng mạng nơ ron đơn giản gồm 1 lớp đầu vào (gồm 10 nơ ron); 2 lớp ẩn (kích thước 32 nơ ron mỗi tầng) và 2 lớp đầu ra được huấn luyện trên tập dữ liệu hơn 30 ngàn mẫu đã đạt được kết quả *F1-score* đạt hơn 88% trên tập kiểm chứng. Kết quả nghiên cứu này cho thấy tiềm năng ứng dụng trí tuệ nhân tạo trong đánh giá nguy cơ đâm va giữa các tàu thuyền trên biển nhằm phát hiện một cách tự động nguy cơ đâm va và đưa ra cảnh báo đối với người điều khiển phương tiện phục vụ quản lý an toàn giao thông hàng hải là khả thi.

Từ các kết quả trên cũng cho ta thấy khả năng mở rộng để giải quyết bài toán nguy cơ đâm va giữa các tàu thuyền trên các vùng biển hoặc sông hẹp hơn với dữ liệu phức tạp hơn vẫn có thể đem lại triển vọng tốt. Chính vì vậy, từ các kết quả của các nghiên cứu này ta cũng có thể nghiên cứu đề xuất các mô hình học máy hiệu quả hơn, có khả năng giải quyết các bài toán phát hiện nguy cơ đâm va phức tạp hơn, ví dụ như trên vùng biển, sông hẹp. Điều này sẽ được trình bày tại chương kế tiếp.

CHƯƠNG 4. SỬ DỤNG AI ĐÁNH GIÁ NGUY CƠ ĐÂM VÀ TẠI KHU VỰC CHẶT HẸP CỦA LUỒNG VỮNG TÀU - SÀI GÒN

4.1. Dữ liệu về nguy cơ đâm va trên luồng lạch hẹp đưa vào mô hình tính toán

Dữ liệu thực tế trích xuất được tại trạm VTS Cảng vụ Hàng hải Thành phố Hồ Chí Minh và Cảng vụ Hàng hải Vũng Tàu là các tập dữ liệu vô cùng lớn lưu trữ các thông số về thông tin, vị trí, tuyến hành trình theo thời gian thực của hàng trăm lượt tàu mỗi ngày. Đây là các tập dữ liệu lớn, phức tạp đến mức khó có thể xử lý được bằng các phương pháp truyền thống. Đối với mục đích nghiên cứu khoa học và tính toán nguy cơ đâm va như đã đề xuất ở các chương trước, tác giả đề xuất sử dụng dữ liệu giả lập, đây là một dạng dữ liệu tinh chỉnh dựa trên yếu tố tuyến luồng hàng hải được giới hạn an toàn, các thông số tốc độ, vị trí, hướng chạy tàu,... hoàn toàn nằm trong phạm vi giới hạn của các số liệu thực tiễn.

Dữ liệu đưa vào chương trình tính toán là dữ liệu giả lập được căn cứ vào vị trí giữa 2 cặp tàu khi tiến tới cùng một điểm nóng giao thông, các thông số chính gồm vị trí tàu, tốc độ, hướng di chuyển.

- Dữ liệu gồm có các sheet name: "WP01", "WP05", "WP06", "WP55", "WP64", "WP67", "WP79", "WP94", "WP98", "WP110", "WP125"

- Các giá trị ban đầu gồm có: t_1 , t_2 , $xa1$, $ya1$, $xb1$, $yb1$, $xa2$, $ya2$, $xb2$, $yb2$, W_{px} , W_{py}

- Đầu ra là ETA_{Wpa} , ETA_{Wpb} , ETA

Gọi tàu thứ nhất là A, tàu thứ hai là B. Trong đó:

- t_1 , t_2 là thời gian tại 2 thời điểm
- $(xa1, ya1)$, $(xa2, ya2)$ là tọa độ của tàu A tại thời điểm t_1 , t_2
- $(xb1, yb1)$, $(xb2, yb2)$ là tọa độ của tàu B tại thời điểm t_1 , t_2
- W_{px} , W_{py} là tọa độ của điểm nóng WP (có 11 điểm nóng)
- ETA_{Wpa} , ETA_{Wpb} là thời gian ước tính đến điểm nóng

- *ETA* là thời gian ước tính va chạm của 2 tàu A, B được tính bằng trị tuyệt đối của hiệu giữa *ETA_Wpa* và *ETA_Wpb*:

$$ETA = | ETA_Wpa - ETA_Wpb |$$

Nếu *ETA* > 10 thì kết luận 2 tàu có nguy cơ va chạm, ngược lại thì kết luận 2 tàu không có nguy cơ va chạm.

- Chia dữ liệu thành 2 phần:

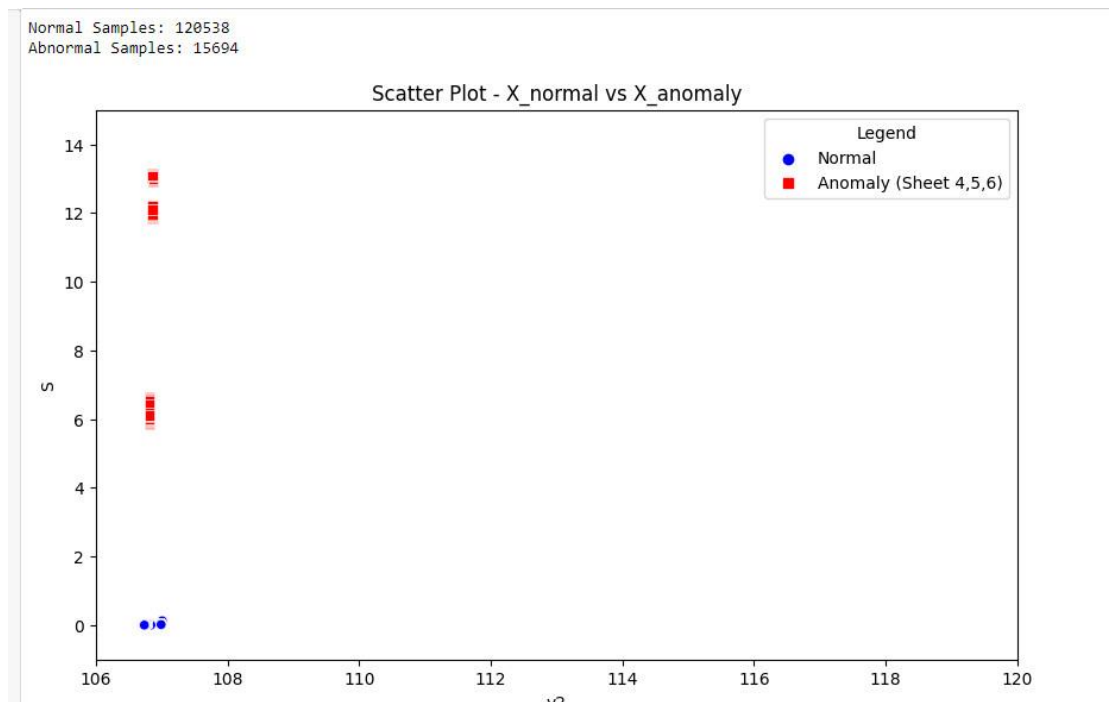
Tập huấn luyện gồm có 80488 dữ liệu (9 điểm nóng sau: "WP01", "WP05", "WP06", "WP55", "WP64", "WP67", "WP79", "WP94", "WP98")

Tập kiểm tra gồm có 55744 dữ liệu (2 điểm nóng sau: "WP110", "WP125")

4.2. Phân tích dữ liệu

4.2.1. Trực quan hóa dữ liệu

- Xem xét khái quát dữ liệu

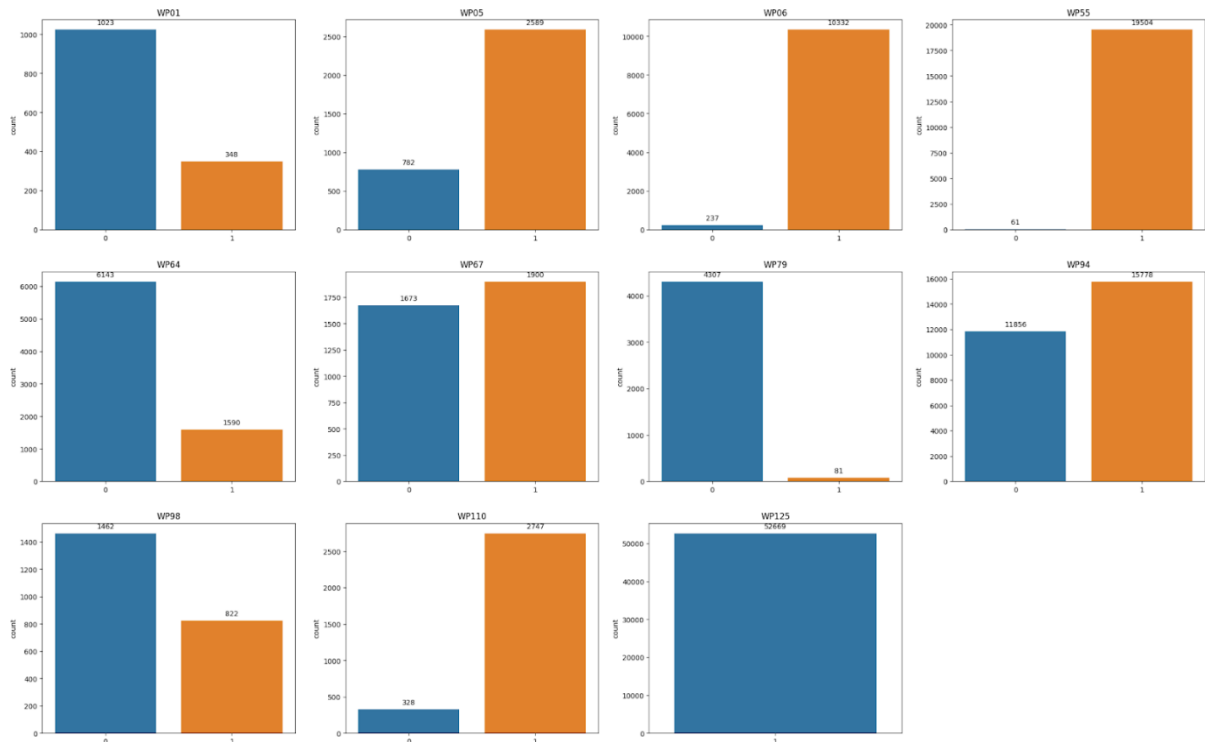


Hình 4.1. Hình ảnh dữ liệu nhiễu

Trên đây là hình ảnh được trực quan hóa của toàn bộ dữ liệu, chấm xanh là các dữ liệu có sự phân bố giống nhau và chấm đỏ các dữ liệu có sự phân bố rất khác ở các sheet thứ 4, 5, 6. Có thể thấy, dữ liệu nhiễu là các giá trị rất khác

so với tổng thể của dữ liệu nhưng lại chiếm số lượng rất lớn. Điều này dẫn đến khi xây dựng mô hình, khả năng tổng quát hóa của mô hình sẽ kém đi.

- Xem xét khái quát các lớp (0 và 1)



Hình 4.2. Hình ảnh phân bố lớp của từng WP

Các lớp 0 và 1 đang bị mất cân bằng. Lớp 1 (tức có nguy cơ va chạm) chiếm đa số so với lớp 0 (Không có nguy cơ va chạm).

- + Ở các WP01, WP64, WP79, WP98 có lớp 1 nhiều hơn lớp 0
- + Ở các WP05, WP67, WP94, WP110 có lớp 0 nhiều hơn lớp 1
- + Ở các WP06, WP55 gần như chỉ có lớp 1
- + Ở WP 125 chỉ có lớp 1

Dữ liệu mất cân bằng lớp làm cho mô hình bị bias (tức là thiên vị về 1 lớp). Các mô hình có thể dự đoán hầu như luôn là lớp đa số mà không quan tâm đến lớp thiểu số, vì điều này có thể giúp chúng đạt được độ chính xác cao hơn mà không cần phải học đối với lớp thiểu số.

4.2.2. Tiền xử lý dữ liệu

Tiền xử lý dữ liệu là quá trình chuẩn bị, biến đổi và làm sạch dữ liệu để chuẩn bị cho việc huấn luyện mô hình hoặc phân tích dữ liệu.

- Đảm bảo dữ liệu đầu vào không bị thiếu.

- Các dữ liệu thuộc 1 đặc trưng cùng đơn vị.
- Loại bỏ các dữ liệu nhiễu trên tập huấn luyện.
- Không sử dụng thời gian làm đầu vào để tránh bị thiên vị (bias).
- Đa dạng hóa đặc trưng để mô hình học được nhiều đặc trưng hơn của dữ liệu.

+ Tính vận tốc và hướng của tàu A và tàu B:

Gọi V_a, C_a, V_b, C_b là vận tốc và hướng của tàu A, B

$$V_a = \text{acos}(\sin(\text{radians}(xa1)) * \sin(\text{radians}(xa2)) + \cos(\text{radians}(xa1)) * \cos(\text{radians}(xa2)) * \cos(\text{radians}(ya2) - \text{radians}(ya1))) * 6371000 / 1852 * 60 / 5 \quad (4.1)$$

Trong đó: acos là hàm arccosine, radians chuyển đổi giá trị từ độ sang radian.

$$C_a = \text{degrees}(\text{atan2}(\cos(\text{radians}(xa1)) * \sin(\text{radians}(xa2)) - \sin(\text{radians}(xa1)) * \cos(\text{radians}(xa2)) * \cos(\text{radians}(ya2) - \text{radians}(ya1)), \sin(\text{radians}(ya2) - \text{radians}(ya1)) * \cos(\text{radians}(xa2)))) + 360 \quad (4.2)$$

Trong đó: degrees chuyển đổi giá trị từ radian sang độ, atan2 là hàm arctang của hai biến.

+ Tính khoảng cách Euclid tọa độ giữa vị trí của tàu với điểm nóng (vị trí của tàu A - điểm nóng và vị trí của tàu B - điểm nóng)

Công thức Euclid:

$$\text{Distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (4.3)$$

4.2.3. Độ đo sử dụng

Để đảm bảo việc đánh giá chính xác các mô hình sau khi huấn luyện, cần sử dụng độ đo phù hợp dưới đây.

Có 4 trường hợp của dự đoán có thể xảy ra:

Đặt Positive là 1 (có nguy cơ va chạm) và Negative là 0 (không có nguy cơ va chạm)

- True Positive (TP): thực tế là có nguy cơ va chạm và mô hình dự đoán đúng là có nguy cơ va chạm.

- True Negative (TN): thực tế là không có nguy cơ va chạm và mô hình dự đoán đúng là không có nguy cơ va chạm

- False Positive (FP): thực tế là không có nguy cơ va chạm, mô hình dự đoán sai là có nguy cơ va chạm

- False Negative (FN): thực tế là có nguy cơ va chạm, mô hình dự đoán sai là không có nguy cơ va chạm.

Dựa vào các dự đoán trên, sử dụng các độ đo sau cho bài toán phân loại có nguy cơ va chạm hay không:

- *Accuracy* được định nghĩa là tỷ lệ phần trăm dự đoán đúng cho dữ liệu thử nghiệm. Nó được tính bằng cách chia số lần dự đoán đúng cho tổng số lần dự đoán :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.4)$$

- *Precision* sẽ cho biết thực sự có bao nhiêu dự đoán Positive thật sự là đúng so với thực tế (True):

$$Precision = \frac{TP}{TP + FP} \quad (4.5)$$

- *Recall* đo lường tỷ lệ dự báo chính xác các trường hợp Positive trên toàn bộ các mẫu thuộc nhóm Positive:

$$Recall = \frac{TP}{TP + FN} \quad (4.6)$$

- *F1-score*, là kỳ vọng harmonic (harmonic mean) của *Precision* và *Recall*. *F1-score* lớn khi cả 2 giá trị *Precision* và *Recall* đều lớn. Ngược lại, chỉ cần 1 giá trị nhỏ sẽ làm cho *F1-Score* nhỏ. *F1-Score* càng lớn càng tốt. Khi lý tưởng nhất thì *F1-score* = 1 (khi *Recall* = *Precision* = 1).

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4.7)$$

Ngoài ra, có thể sử dụng các độ đo sai số sau:

- **Mean Squared Error (MSE)** được định nghĩa là trung bình tổng bình phương sai số giữa đầu ra dự đoán và kết quả thực.

$$MSE = \frac{1}{n} \sum_1^n (y_i - \hat{y}_i)^2 \quad (4.8)$$

- **Mean Absolute Error (MAE)** được định nghĩa là trung bình tổng trị tuyệt đối sai số giữa đầu ra dự đoán và kết quả thực:

$$MAE = \frac{1}{n} \sum_1^n |y_i - \hat{y}_i| \quad (4.9)$$

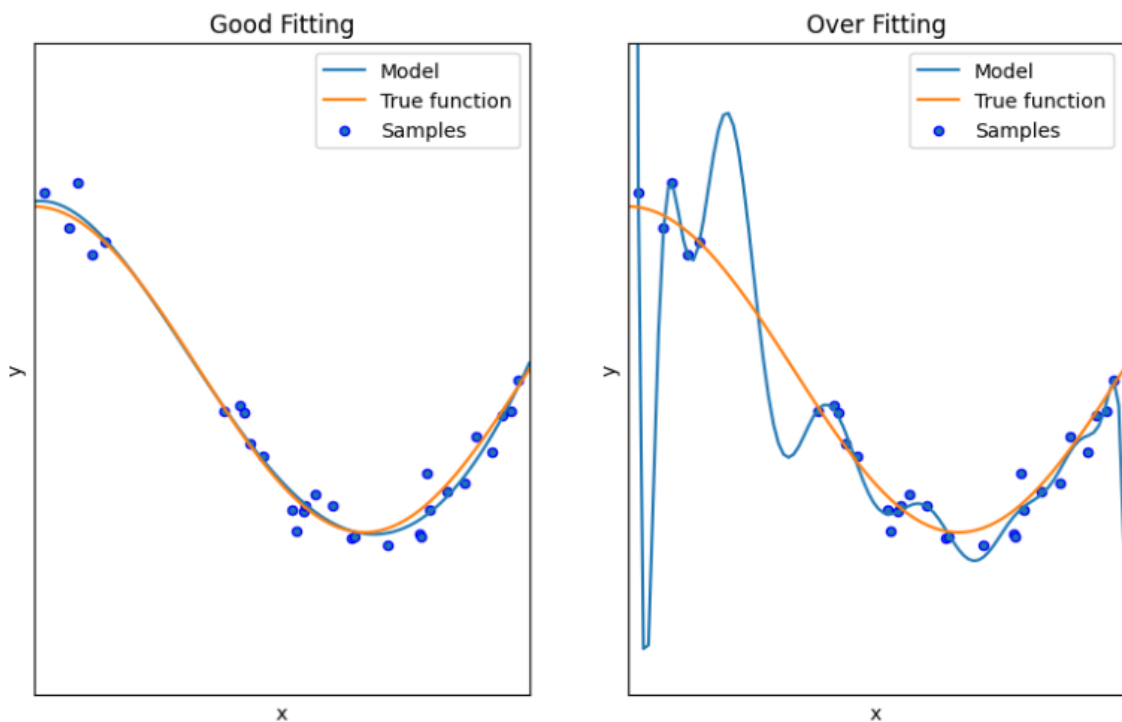
MSE và MAE có miền giá trị từ $[0, +\infty]$. Trên cùng tập dữ liệu, MSE và MAE càng nhỏ thì có độ chính xác càng cao.

4.2.4. Sự quá khớp (Overfitting)

Overfitting là hiện tượng khi một mô hình học quá mức từ dữ liệu huấn luyện, đến mức mô hình "nhớ" các chi tiết và nhiễu trong dữ liệu huấn luyện mà không học được quy luật tổng quát. Khi một mô hình đã bị overfitting, hiệu suất của nó trên dữ liệu mới (dữ liệu chưa từng thấy) có thể giảm đi đáng kể.

Dấu hiệu của overfitting bao gồm:

- Hiệu suất tốt trên tập huấn luyện nhưng kém trên tập kiểm tra.
- Mô hình có độ phức tạp cao.



Hình 4.3. Ví dụ về Good Fitting và Over Fitting

Hình bên trái (Good Fitting) cho thấy mô hình được huấn luyện rất giống với mô hình thực tế. Hình bên phải (Over Fitting) cho thấy mô hình được huấn luyện khác so với mô hình thực tế, vì mô hình này cố gắng đi qua tất cả các điểm dữ liệu (kể cả dữ liệu nhiễu).

Do dữ liệu của chúng ta đang bị nhiễu lớn nên rất dễ xảy ra hiện tượng overfitting này. Dưới đây, chúng ta sẽ đi thử nghiệm các mô hình và xem xét mô hình nào dễ bị overfitting và cách để tìm được mô hình tốt nhất

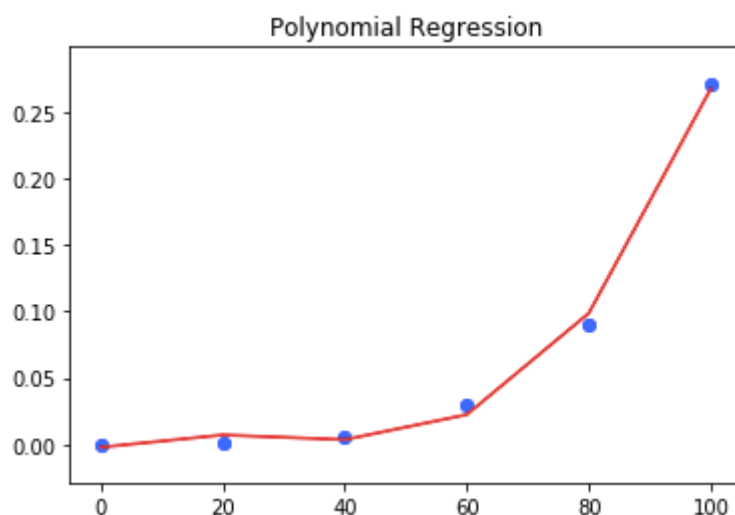
4.3. Lựa chọn và thực hành huấn luyện mô hình AI cho luồng lạch hẹp

Có nhiều thuật toán khác nhau trong mô hình huấn luyện AI đối với tập dữ liệu lớn được nêu ở trên, nghiên cứu sinh đã chạy thử các thuật toán như: Mạng nơ-ron, Thuật toán hồi quy tuyến tính (Linear Regression, cho kết quả F1-score = 50%), Mô hình hồi quy Logistics (Logistic Regression, F1-score =49%), Mô hình cây quyết định (Decision Tree Regressor, F1-score =47%). Tuy nhiên các kết quả thu được không thỏa mãn bài toán đặt ra, thường bị quá khớp (overfitting), không thỏa mãn tập kiểm tra, khó kiểm soát quá trình huấn luyện (Phụ lục 1). Trong các mô hình đó, chỉ có mô hình hồi quy đa thức kết hợp Lasso là mang lại kết quả tối ưu, đáp ứng được vấn đề đặt ra là tính toán khả năng đâm va tại điểm nóng giao thông. Dưới đây là chi tiết các bước thực hiện chạy mô hình học máy thông qua các thông số điều chuẩn và tham số kiểm soát mức độ của thành phần điều chuẩn

4.3.1. Mô hình hồi quy đa thức kết hợp Lasso (Polynomial kết hợp với Lasso)

4.3.1.1. Polynomial

Polynomial (Hồi quy đa thức) là một thuật toán học máy có giám sát (*Supervised learning*). Đây là một phương pháp để dự đoán biến phụ thuộc y dựa trên giá trị của biến X theo một hàm đa thức bậc (Polynomial degree) xác định. Nó có thể được sử dụng cho các trường hợp chúng ta muốn dự đoán các loại dữ liệu theo chuỗi thời gian liên tục (continuous time series); mà bản chất dữ liệu cảm biến sử dụng cho bài toán phát hiện nguy cơ va chạm là dạng dữ liệu chuỗi thời gian.



Hình 4.4. Minh họa Polynomial

Giả sử:

Đầu vào (input) $x = [x_1, x_2, x_3, \dots, x_n]$ là một vector chứa đặc trưng của dữ liệu cảm biến.

Đầu ra (output) $y = f(x)$ là một số vô hướng được tính toán từ hàm $f(x)$ dựa trên các giá trị của véc tơ X.

Degree là số được xác định trước thể hiện bậc của đa thức.

Hệ số Degree quyết định mức độ phức tạp của mô hình. Mức độ cao hơn có thể dẫn đến mô hình phức tạp hơn, có thể dẫn đến hiệu số tốt trên dữ liệu huấn luyện nhưng có thể dẫn đến hiện tượng overfitting trên dữ liệu mới.

Hàm đa thức có thể mô tả mối quan hệ giữa *output* và *input* là:

Degree = 1 tương đương đa thức bậc 1 (giống với Linear Regression)

Degree = 2 tương đương đa thức bậc 2

Giả sử ban đầu miền feature là $x = [x_1, x_2]$ thì sau khi mở rộng với bậc(degree) bằng 2 sẽ trở thành:

$$x' = [x_1, x_2, x_1x_2, x_1^2, x_2^2]$$

Tương tự với các bậc còn lại.

4.3.1.2. Lasso

Lasso là một phương pháp phân tích hồi quy thực hiện cả lựa chọn biến và chính quy hóa (regularization). Thuật toán như sau:

$x = [x_1, x_2, x_3, \dots, x_n]$ là một vector chứa thông tin *input*.

$y = [y_1, y_2, y_3, \dots, y_n] = f(x)$ là một số vô hướng biểu diễn *output*.

Một hàm số đơn giản nhất có thể mô tả mối quan hệ giữa *output* và *input* là:

$$y \approx f(x) = \hat{y} \quad (4.10)$$

Ý tưởng cơ bản của Lasso là thêm một thành phần regularization vào hàm mất mát của mô hình hồi quy tuyến tính thông thường.

Ta có hàm tổn thất (Loss function) của Hồi quy tuyến tính thông thường như sau:

$$L_0(w) = \frac{1}{2} \sum_{i=1}^n (y_i - \bar{x}_i w)^2 \quad (4.11)$$

Lasso sử dụng regularization $L1$, tức là thêm tổng giá trị tuyệt đối của các hệ số vào hàm mất mát, điều này có tác dụng tạo ra các hệ số nhỏ, thậm chí có thể đưa một số hệ số về 0, từ đó thực hiện việc chọn lọc đặc trưng.

$$L(w) = \sum_{i=1}^n |y_i - \bar{x}_i w|^2 + \alpha \|w\|_1 \quad (4.12)$$

Hàm tối ưu w của Lasso là:

$$w^* = \arg \min_w \sum_{i=1}^n |y_i - \bar{x}_i w|^2 + \alpha \|w\|_1 \quad (4.13)$$

Trong phương trình trên: $\alpha \|w\|_1$ đại diện cho *thành phần điều chuẩn*.

Bài toán tối ưu hàm tổn thất $L(w)$ về bản chất là tối ưu song song hai thành phần bao gồm tổng bình phương phần dư và *thành phần điều chuẩn*.

Giải bài toán tối ưu hàm mục tiêu của Lasso theo đạo hàm của w :

$$\frac{\partial L(w)}{\partial w} = \frac{1}{N} \frac{\partial |y_i - \bar{x}_i w|^2}{\partial w} + \alpha \frac{\partial \|w\|_1}{\partial w} \quad (4.14)$$

- Điều chuẩn $L1$ ($L1$ regularization):

Điều chuẩn (regularization) giúp kiểm soát độ lớn của các hệ số (weights) trong mô hình bằng cách thêm vào hàm loss 1 đại lượng $\alpha \|w\|_1$.

Việc tối ưu hóa model cũng đồng nghĩa với việc làm giảm hàm tổn thất (loss function) dẫn đến giảm weight làm cho mô hình bớt phức tạp. Từ đó giúp

mô hình tránh overfitting.

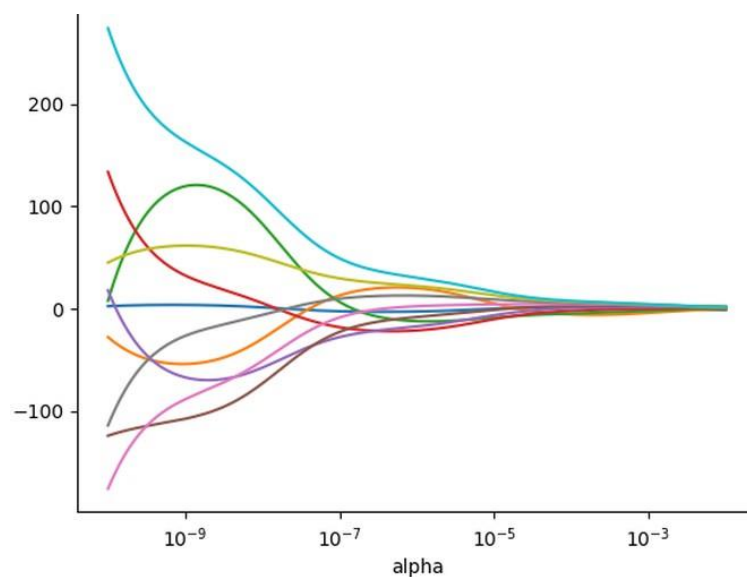
- **Phân tích hệ số alpha:** Tham số alpha kiểm soát mức độ của *thành phần điều chuẩn* (regularization).

+ Nếu $\alpha = 0$, mô hình tương đương với mô hình hồi quy tuyến tính thông thường mà không có *thành phần điều chuẩn*.

+ Nếu α nhỏ thì vai trò của *thành phần điều chuẩn* trở nên ít quan trọng. Mức độ kiểm soát *quá khớp* của mô hình sẽ trở nên kém hơn.

+ Nếu α lớn chúng ta muốn gia tăng mức độ kiểm soát lên độ lớn của các hệ số ước lượng.

Điều chuẩn (regularization) trong Lasso giúp kiểm soát overfitting trong mô hình hồi quy tuyến tính. Việc giảm thiểu các hệ số có thể dẫn đến một mô hình đơn giản hơn và ít phức tạp hơn, giảm nguy cơ overfitting.



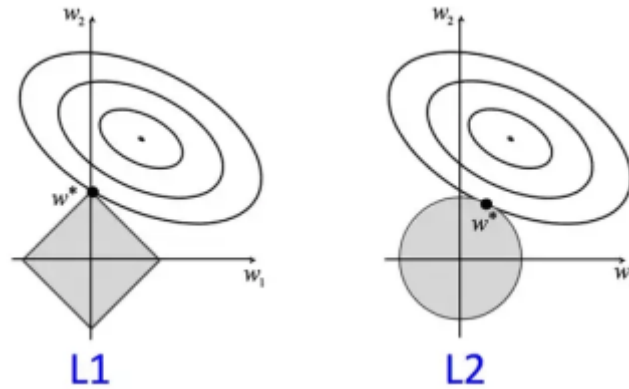
Hình 4.5. Minh họa sự ảnh hưởng của alpha tới hệ số ước lượng

Ngoài ra, việc chọn giá trị alpha phù hợp có thể là một thách thức. Giá trị alpha quá cao có thể làm mất mát thông tin quan trọng, trong khi giá trị alpha quá thấp có thể làm mô hình trở nên quá phức tạp và dễ bị overfitting.

Các kiểu hiệu chỉnh khác nhau sẽ tạo ra các miền khác nhau cho w .

Lasso thường tạo ra nghiệm thưa, tức là nhiều thành phần của w có giá trị là 0.

Vì thế Lasso thực hiện đồng thời việc hạn chế và lựa chọn đặc trưng.



Hình 4.6. Hình ảnh giữa chuẩn hóa L1 và L2

Miền xác định của hồi quy Lasso trên đồ thị miền xác định này là một vùng hình thoi màu xám nằm bên trái $|w_1 + w_2| \leq t$. Hình bên phải có miền xác định là một hình tròn $|w_1^2 + w_2^2| \leq C$.

Đồ thị hàm mục tiêu $L(w)$ được thể hiện qua 1 tập hợp các đường đồng mức hình ellipse. Mỗi một đường đồng mức sẽ trả về cùng một giá trị hàm mục tiêu. Các đường đồng mức ở gần tâm (chấm tròn đen ở giữa) thì càng có giá trị nhỏ hơn. Khi mở rộng dần đường đồng mức cho tới khi tiệm cận miền xác định chúng ta sẽ thu được nghiệm của bài toán.

Đối với hồi quy Lasso thì tập nghiệm thường chạm đỉnh của hình thoi (L1- norm). Đây là những điểm tương ứng với một chiều bằng 0. Trong khi đó, trong khi đó hình bên phải (L2-norm) thì miền xác định là một hình tròn nên điểm tiếp xúc sẽ thường có tọa độ khác 0.

Chính vì thế Lasso thực hiện đồng thời việc hạn chế sự quá khớp (overfitting) và lựa chọn đặc trưng.

Áp dụng

- Dữ liệu huấn luyện và dữ liệu kiểm tra có input và output như sau:
Input: xa1, ya1, xb1, yb1, xa2, ya2, xb2, yb2, Wpx, Wpy
Output: ETA_Wpa, ETA_Wpb
- **Mô hình:** Decision Tree sử dụng tiêu chí phân tách là Gini index
- **Kết quả:**

Lasso với alpha = 1

+ Độ đo sai số như sau:

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	3.68055	34.42609
tàu B	7.27996	16.12679

Ta có bảng độ đo chính xác như sau:

+ Tập huấn luyện

	Precision	Recall	F1-score	Support
0	0.83	0.76	0.79	15421
1	0.93	0.95	0.94	49373
Accuracy			0.90	64794
Macro avg	0.88	0.86	0.87	64794

+ Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	328
1	0.99	1.00	1.00	55416
Accuracy			0.99	55744
Macro avg	0.50	0.50	0.50	55744

Với alpha = 1, trên tập huấn luyện mô hình dự đoán tốt với cả 2 lớp dữ liệu với F1-score trung bình (macro avg) của 2 lớp là 87% nhưng không dự đoán đúng lớp 0 nào trên tập kiểm tra => Mô hình không tốt, chỉ đạt F1-score trung bình (macro avg) của 2 lớp trên tập kiểm tra là 50%.

Lasso với alpha = 2

+ Độ đo sai số như sau:

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	12.1357	18.2547
tàu B	11.0462	7.1724

Ta có bảng độ đo chính xác như sau:

+ Tập huấn luyện

	Precision	Recall	F1-score	Support
0	1.00	0.09	0.17	15421
1	0.78	1.00	0.88	49373
Accuracy			0.78	64794
Macro avg	0.89	0.55	0.53	64794

+ Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	328
1	0.99	1.00	1.00	55416
Accuracy			0.99	55744
Macro avg	0.50	0.50	0.50	55744

Với $\alpha = 2$, độ đo sai số trên tập huấn luyện và tập kiểm tra không quá cao và tương đương nhau, thậm chí độ đo sai số trên tập kiểm tra còn thấp hơn so với $\alpha = 1$. Tuy nhiên, trên tập huấn luyện mô hình chỉ dự đoán tốt với dữ liệu thuộc lớp 1; với lớp 0, mô hình dự đoán sai rất nhiều và chỉ đạt $F1$ -score là 17%. Trên tập kiểm tra, mô hình dự đoán không dự đoán đúng lớp 0 nào. Kết quả này cho thấy mô hình không thực sự tốt, có $F1$ -score trung bình (macro avg) của 2 lớp là 53% trên tập huấn luyện và 50% trên tập kiểm tra.

Lasso với alpha = 0.5

+ Độ đo sai số như sau

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	7.0945	1.0742
tàu B	6.8453	0.2814

Ta có bảng độ đo chính xác như sau:

- Tập huấn luyện

	Precision	Recall	F1-score	Support
0	0.96	0.49	0.64	15421
1	0.86	0.99	0.92	49373
Accuracy			0.87	64794
Macro avg	0.91	0.74	0.78	64794

- Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	328
1	0.99	1.00	1.00	55416
Accuracy			0.99	55744
Macro avg	0.50	0.50	0.50	55744

Với alpha = 0.5: Mặc dù độ đo sai số của tàu A và tàu B đạt kết quả nhỏ nhất so với hai lần thử nghiệm trên (với alpha = 1 và alpha = 2) nhưng độ đo sự chính xác lại kém. Trên tập huấn luyện, dự đoán nhãn 0 chỉ đạt 64% và trên tập kiểm tra không dự đoán đúng nhãn 0 nào, tức mô hình không thực sự tốt. Kết quả với *F1-score* trung bình (macro avg) của 2 lớp là 78% trên tập huấn luyện và 50% trên tập kiểm tra.

+ Kết luận: Mô hình Lasso đạt kết quả tốt nhất với $\alpha = 1$.

Với $\alpha = 2$ hay $\alpha = 0.5$ thì ta thấy các mô hình có dấu hiệu quá khớp (overfitting). Trong đó chỉ đạt kết quả tương đối tập trên huấn luyện nhưng trên tập kiểm tra đối với lớp 0, mô hình không dự đoán đúng bất kỳ mẫu nào thuộc lớp 0.

Do đó, ta nên kết hợp với hồi quy đa thức (Polynomial regression)

Dưới đây là kết quả thử nghiệm mô hình kết hợp Polynomial với Lasso ($\alpha = 1$) với các bậc Degree khác nhau, lần lượt là 2, 3, 4.

Polynomial (Degree = 3) kết hợp với Lasso ($\alpha = 1$)

+ Độ đo sai số như sau:

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	5.58716	2.27946
tàu B	5.12589	3.24671

Ta có bảng độ đo chính xác như sau:

+ Tập huấn luyện

	Precision	Recall	F1-score	Support
0	0.95	0.62	0.75	15421
1	0.89	0.99	0.94	49373
Accuracy			0.90	64794
Macro avg	0.92	0.81	0.85	64794

+ Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.76	0.89	0.82	328
1	1.00	1.00	1.00	55416
Accuracy			1.00	55744
Macro avg	0.88	0.95	0.91	55744

Mặc dù độ đo sai số của tàu A và tàu B chưa đạt kết quả nhỏ nhất so với các thử nghiệm trên nhưng độ đo sự chính xác lại cho kết quả khả quan. Trên tập huấn luyện, dự đoán nhãn 0 đạt 75% *F1*-score và nhãn 1 đạt 94%. Đồng thời *F1*-score trên tập kiểm tra dự đoán đúng nhãn 0 nào tức là 82% và nhãn 1 lên tới 100% *F1*-score. Kết quả chung với *F1*-score trung bình (macro avg) của 2 lớp là 85% trên tập huấn luyện và 91% trên tập kiểm tra cho thấy mô hình học khá tốt từ dữ liệu và kết quả hứa hẹn.

Polynomial (Degree = 2) kết hợp với Lasso (alpha = 1)

+ Độ đo sai số như sau

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	6.71096	0.17897
tàu B	6.32211	1.66528

Ta có bảng độ đo chính xác như sau:

+ Tập huấn luyện

	Precision	Recall	F1-score	Support
0	0.90	0.67	0.77	15421
1	0.90	0.98	0.94	49373
Accuracy			0.90	64794
Macro avg	0.90	0.83	0.86	64794

+ Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.00	0.00	0.00	328
1	0.99	1.00	1.00	55416
Accuracy			0.99	55744
Macro avg	0.50	0.50	0.50	55744

Mặc dù độ đo sai số trên tập kiểm tra của tàu A và tàu B cho kết quả khá tốt song độ chính xác lại cho kết quả chưa cao. Trên tập huấn luyện, dự đoán nhãn 0 đạt 77% $F1$ -score và nhãn 1 đạt 94% với $F1$ -score trung bình là 86%. Ngược lại, $F1$ -score trên tập kiểm tra dự đoán đúng nhãn 0 lại là 0% và nhãn 1 lên tới 100% $F1$ -score. Kết quả chung với $F1$ -score trung bình (macro avg) của 2 lớp là 86% trên tập huấn luyện và 50% trên tập kiểm tra cho thấy mô hình bị overfitting nên đa thức hồi qui bậc 2 không hiệu quả cho bài toán phát hiện nguy cơ va chạm.

Polynomial (Degree = 4) kết hợp với Lasso (alpha = 1)

+ Độ đo sai số như sau

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	5.22354	6.53838
tàu B	4.40549	9.57899

Ta có bảng độ đo chính xác như sau:

+ Tập huấn luyện

	Precision	Recall	F1-score	Support
0	0.93	0.69	0.79	15421
1	0.91	0.98	0.94	49373
Accuracy			0.91	64794
Macro avg	0.92	0.84	0.87	64794

+ Tập kiểm tra

	Precision	Recall	F1-score	Support
0	0.45	0.99	0.61	328
1	1.00	0.99	1.00	55416
Accuracy			0.99	55744
Macro avg	0.73	0.99	0.81	55744

Độ đo sai số trên tập kiểm tra của tàu A và tàu B cho kết quả không tốt bằng đa thức hồi quy bậc 3 và bậc 2. Trên tập huấn luyện, dự đoán nhãn 0 đạt 79% f1-score và nhãn 1 đạt 94% với $F1$ -score trung bình là 87%. Đồng thời $F1$ -score trên tập kiểm tra dự đoán đúng nhãn 0 lại là 61% và nhãn 1 lên tới 100% $F1$ -score. Kết quả chung với $F1$ -score trung bình (macro avg) của 2 lớp là 87% trên tập huấn luyện và 81% trên tập kiểm tra cho thấy mô hình tương đối tốt nhưng chưa tốt bằng đa thức hồi qui bậc 3.

Polynomial kết hợp với Lasso ($\alpha = 1$) đạt kết quả tốt nhất khi bậc của Polynomial là 3. Mô hình đạt $F1$ -score trung bình (macro avg) của 2 lớp là 85% trên tập huấn luyện và 91% trên tập kiểm tra và tốt với cả 2 lớp 0 và 1. Trong đó với lớp 1 mô hình dự đoán gần như hoàn toàn chính xác. Ngoài ra, độ đo sai số trên cả tập huấn luyện và tập kiểm tra đều tương đối nhỏ và cân bằng nhau mặc dù đây chưa phải là độ đo sai số nhỏ nhất trong 3 lần thử nghiệm mô hình kết hợp nhưng lại đạt được kết quả tốt nhất.

4.3.2. Phân tích và đánh giá

4.3.2.1. Nhận xét dữ liệu

- Huấn luyện mô hình trên 9 WP đầu tiên và kiểm thử trên 2 WP cuối cùng để đảm bảo độ tổng quát hóa và có khả năng dự đoán trên tập dữ liệu mới, chứ không chỉ dữ liệu được huấn luyện.
- Dữ liệu có số lượng nhiều lớn làm cho mô hình huấn luyện dễ bị quá khớp (overfitting). Khi cố gắng tạo ra mô hình phù hợp với tất cả các điểm dữ liệu thì mô hình sẽ trở nên quá phức tạp đồng thời giảm tính khái quát hóa (generalization) của mô hình.
 - + Cụ thể, các dữ liệu có sự phân bố rất khác của các WP55, WP64, WP67) dữ liệu chứa khá nhiều điểm ngoại lai (outliers) và mất cân bằng.
 - + Có thể giải quyết vấn đề này bằng cách loại bỏ các dữ liệu nhiễu trên hoặc áp dụng các mô hình không bị ảnh hưởng bởi dữ liệu nhiễu.
 - + Trong bài toán này, chúng ta tiếp cận hướng tối ưu và hiệu quả hơn là loại bỏ dữ liệu nhiễu.

- Dữ liệu mất cân đối giữa 2 nhãn dẫn đến mô hình sẽ học được nhiều hơn ở 1 lớp và có xu hướng dự đoán nhiều về lớp đó.
 - + Ở các WP01, WP64, WP79, WP98 có lớp 1 nhiều hơn lớp 0
 - + Ở các WP05, WP67, WP94, WP110 có lớp 0 nhiều hơn lớp 1
 - + Ở các WP06, WP55 gần như chỉ có lớp 1
 - + Ở WP 125 chỉ có duy nhất lớp 1

Lớp 1 (tức có nguy cơ va chạm) chiếm đa số so với lớp 0 (Không có nguy cơ va chạm). Điều này có thể được giải thích do dữ liệu được sinh ra, chưa phải dữ liệu thực tế nên có thể không chính xác. Ngoài ra, có thể tồn tại các điểm nóng WP, mà ở đó thường xuyên diễn ra sự va chạm hoặc không diễn ra sự va chạm nào cũng là nguyên nhân dẫn đến sự mất cân bằng dữ liệu

- Các thuộc tính không nhiều, cụ thể chỉ có 6 thuộc tính đầu vào (tọa độ (x, y) của 2 tàu tại 2 thời điểm và tọa độ WP). Chính vì vậy phương pháp phân loại thông thường thường không phát huy hiệu quả trong các trường hợp này.

Giải pháp đề xuất:

- + Sử dụng các mô hình hồi quy (regression).
- + Mở rộng các thuộc tính bằng cách tạo thêm các giá trị thuộc tính (như tính V, C theo một công thức xác định; tính thêm khoảng cách giữa mỗi tàu so với WP) để mô hình học được đa dạng hơn.

4.3.2.2. Đề xuất mô hình

- Áp dụng mô hình cũ: Mô hình Neural Network (2 layer) có thể phù hợp với dữ liệu trên tập huấn luyện. Tuy nhiên mô hình không có khả năng tổng quát hóa nên đã đạt kết quả kém trên tập kiểm tra. Ngoài ra, mô hình khó kiểm soát các trọng số trong quá trình huấn luyện. Mô hình này phù hợp với việc xác định nguy cơ đâm va tại khu vực biển rộng, cụ thể là khu vực đầu vào luồng Vũng Tàu - Sài Gòn, nơi có mật độ tàu thuyền cao.

- Áp dụng các mô hình mới: Linear Regression, Logistic Regression, Decision Tree Regressor, Lasso cũng đạt kết quả tương đối ổn định trên tập huấn luyện, nhưng rất dễ bị overfitting dẫn đến đạt kết quả kém trên tập kiểm tra.

- + Trong các mô hình trên, có những mô hình có độ sai số rất nhỏ (chẳng hạn Decision Tree Regressor có độ đo sai số trên tập huấn luyện bằng 0) nhưng lại rất tệ trên tập kiểm tra. Như vậy, không phải mô hình nào đạt độ sai số nhỏ nhất sẽ cho ra kết quả tốt nhất. Vì độ sai số bằng 0 có thể là dấu hiệu của Overfitting.

- + Lý do: Dữ liệu khi ở miền không gian gốc có thể không tuyến tính. Cần mở rộng miền không gian gốc lên số chiều cao hơn sao cho trên miền không gian mới dữ liệu này là tuyến tính (Có thể xấp xỉ được 1 hyperspace đi qua các điểm dữ liệu đó với sai số nhỏ).

- + Để giải quyết vấn đề này, cần áp dụng phương pháp Polynomial Features (là 1 kỹ thuật để mở rộng Feature Space). Tuy nhiên, việc mở rộng thêm các chiều không gian mới dựa trên các chiều không gian cũ có thể gây ra overfitting và một số chiều trong không gian mới có thể không quan trọng trong việc hỗ trợ mô hình dự đoán.

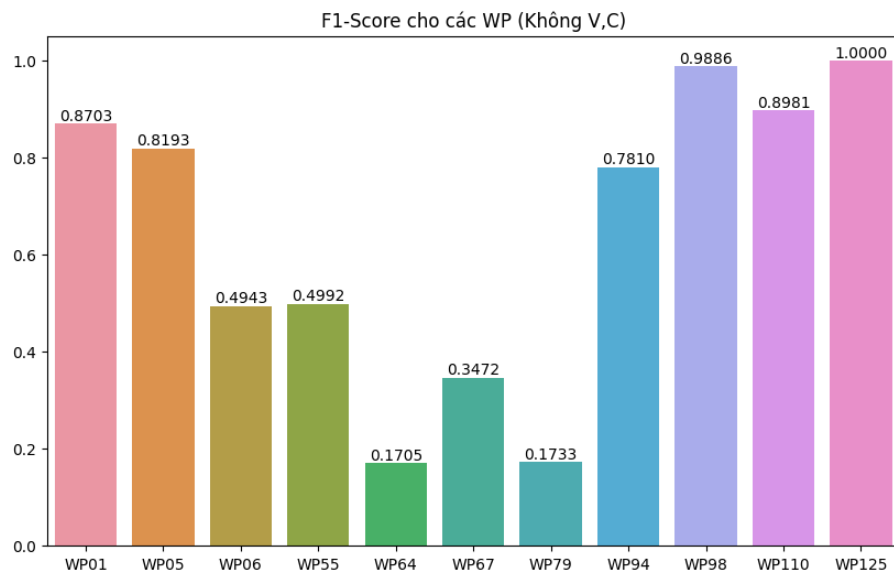
- + Trong các mô hình trên, mô hình Lasso có khả năng loại bỏ một số đặc trưng không quan trọng đó (trong lý thuyết đã nêu).

Từ những phân tích trên, đề xuất sử dụng mô hình kết hợp giữa Polynomial và Lasso. Kết quả tốt nhất là **mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$)** với F1-score trung bình (macro avg) của 2 lớp trên tập huấn luyện đạt 85% và trên tập kiểm tra đạt 91%.

- + Mô hình kết hợp này đã giải quyết được vấn đề overfitting ở hầu hết các mô hình thử nghiệm trước đó.

- + Mô hình kết hợp này có độ đo sai số tốt nhất trong tất cả các mô hình thử nghiệm trước đó, cân bằng ở cả 2 lớp trên tập kiểm tra và tập huấn luyện.

- + Mô hình kết hợp này dự đoán được lớp 0 trên tập kiểm tra trong khi các mô hình trước đó gần như không dự đoán đúng lớp 0 nào trên tập kiểm tra.
- Các độ đo cần được xem xét (nêu trong phần 3 mục II). Trong đó cần chú ý đến macro avg (là $F1$ -score trung bình (macro avg) của 2 lớp) để đảm bảo có nhận xét tổng quát nhất về kết quả.

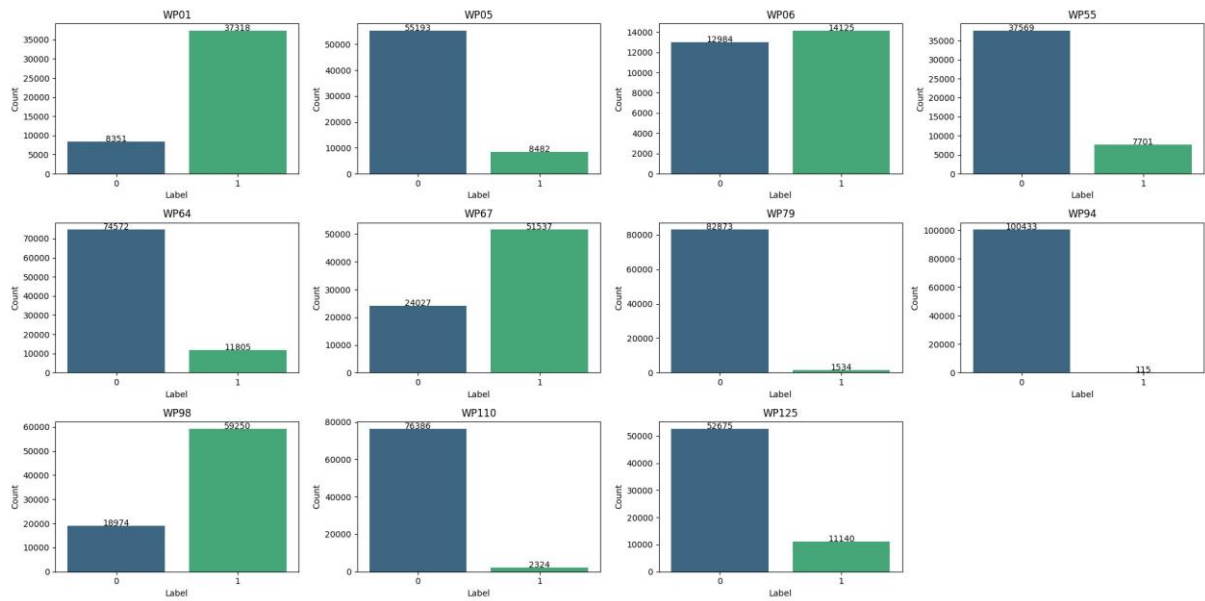


Hình 4.7. F1-Score trung bình cho các WP

- Mô hình cho kết quả dự đoán chưa tốt ở các WP06, WP55, WP64, WP67, WP79 (dưới 60% $F1$ -score) do tính chất mất cân bằng của dữ liệu.

4.3.3. Đánh giá độ tin cậy của mô hình đề xuất trên tập dữ liệu lớn

Để xác nhận độ tin cậy (reliability), khả mở (scalability) và tính khái quát (generalization) đối với **mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$)**, nghiên cứu sinh đã thu thập thêm 1 tập dữ liệu lớn với 746.379 mẫu dữ liệu (lớn hơn 20 lần so với bộ dữ liệu đã thử nghiệm ở trên) với dữ liệu đa dạng, phức tạp hơn. Sự phân bố dữ liệu được trình bày trong hình 4.8 dưới đây:



Hình 4.8. Phân bố dữ liệu giữa các WP trên tập dữ liệu lớn

Phân bố dữ liệu đã được trình bày trên hình 4.8, như đã thấy, dữ liệu khá mất cân bằng, đặc biệt là các WP74, WP79 và WP110 (số lượng dữ liệu nguy cơ đâm va ít hơn rất đáng kể so với số lượng dữ liệu không có nguy cơ đâm va). Điều này phản ánh tập dữ liệu lớn cũng tương đối phù hợp so với các tình huống trên thực tế.

Training Result				
	precision	recall	f1-score	support
0	0.94	0.95	0.94	185302
1	0.94	0.94	0.94	175195
accuracy			0.94	360497
macro avg	0.94	0.94	0.94	360497
weighted avg	0.94	0.94	0.94	360497

Hình 4.9. Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với dữ liệu huấn luyện trên tập dữ liệu lớn

Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với bộ dữ liệu lớn được trình bày trong hình 4.9 cho thấy độ chính xác (precision), độ nhạy (recall), và $F1$ -score đều đạt 94% trên tập huấn luyện (training). Tức là hiệu suất phát hiện nguy cơ đâm va tàu cao hơn so với tập có ít dữ liệu hơn. Điều này chứng tỏ khi mô hình được huấn luyện với nhiều dữ liệu

hơn đáng kể thì vẫn chưa bị quá khớp, và vì vậy mô hình có độ tin cậy cũng như tính khái quát tốt.

Testing Result				
	precision	recall	f1-score	support
0	0.91	0.73	0.81	15478
1	0.97	0.99	0.98	127047
accuracy			0.96	142525
macro avg	0.94	0.86	0.89	142525
weighted avg	0.96	0.96	0.96	142525

Hình 4.10. Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với dữ liệu kiểm thử trên tập dữ liệu lớn

Kết quả thử nghiệm mô hình đa thức hồi quy bậc 3 kết hợp với Lasso ($\alpha = 1$) với dữ liệu kiểm thử trên tập dữ liệu lớn trên hình 4.10 cho kết quả độ chính xác, độ nhạy, và $F1$ -score đều đạt 96%. Kết quả này cho thấy khi tăng dữ liệu huấn luyện thì độ chính xác phát hiện nguy cơ đâm và tàu cũng được cải thiện. Mô hình đề xuất phù hợp với bài toán phát hiện nguy cơ đâm va trong vùng hẹp.

Mô hình hồi quy đa thức bậc 3 kết hợp với Lasso ($\alpha = 1$) cũng cho kết quả tốt nhất với $F1$ -score trung bình (weight avg) của 2 lớp lên tới 94% trên tập huấn luyện và 96% trên tập kiểm tra và tốt đều trên cả 2 lớp 0 và 1. Ngoài ra, độ đo sai số trên tập huấn luyện là 2.03 với tàu A, và 1.94 với tàu B và tập kiểm tra đều cho kết quả RMSE nhỏ hơn 0.6. Điều này chứng tỏ mô hình hồi quy đa thức bậc 3 kết hợp Lasso cho kết quả đáng tin cậy trên các tập dữ liệu có quy mô khác nhau.

4.4. Kết luận chương 4

Trong phần này, một số mô hình học máy hồi quy được khảo sát và thử nghiệm; kết quả các thử nghiệm cho thấy mô hình hồi quy đa thức bậc 3 kết hợp với Lasso cho kết quả với $F1$ -score trung bình (macro avg) của 2 lớp trên tập huấn luyện đạt 85% và trên tập kiểm tra đạt 91%. Các tham số của mô hình hồi

quy đa thức kết hợp Lasso đã giải quyết hiệu quả bài toán xác định nguy cơ đâm va tại điểm nóng đối với các cặp tàu mục tiêu bất kỳ, trên cơ sở dữ liệu thực được cung cấp theo thời gian thực. Các dữ liệu được giả lập, rút gọn hơn, thuận lợi hơn cho quá trình tính toán so với kho dữ liệu được lưu trữ trong hệ thống VTS của luồng hàng hải.

Ngoài ra, nghiên cứu sinh cũng tiến hành kiểm chứng độ tin cậy của mô hình hồi quy đa thức bậc 3 kết hợp với Lasso ($\alpha=1$) bằng việc thử nghiệm trên một tập dữ liệu lớn (gồm hơn 746 nghìn mẫu). Kết quả thử nghiệm cho thấy mô hình đã cải thiện đáng kể độ chính xác, độ nhạy, $F1$ -score, các sai số trên tập dữ liệu lớn. Điều này chứng tỏ mô hình đề xuất không bị quá khớp và vẫn học được các đặc trưng từ dữ liệu để cải thiện hiệu quả của mô hình. Như vậy, mô hình đề xuất có độ tin cậy và tính khái quát cao, phù hợp và hiệu quả bài toán xác định nguy cơ đâm va tại điểm nóng đối với các cặp tàu trong nhiều tình huống khác nhau.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Luận án đã đạt được mục tiêu ban đầu về nghiên cứu ứng dụng trí tuệ nhân tạo và dữ liệu lớn trong quản lý giao thông hàng hải trên tuyến luồng Vũng Tàu - Sài Gòn. Các dữ liệu về đánh giá nguy cơ đâm va giữa các tàu thuyền trên vùng biển rộng, đánh giá nguy cơ đâm va giữa các tàu thuyền trên luồng lạch hẹp đã được nghiên cứu, tạo lập cơ sở dữ liệu, kiểm tra dữ liệu và sử dụng cơ sở dữ liệu để huấn luyện mô hình AI nhằm mục tiêu sử dụng trí tuệ nhân tạo trong phát hiện sớm nguy cơ đâm va giữa các tàu thuyền, phục vụ tốt hơn công tác quản lý an toàn giao thông trên tuyến luồng Vũng Tàu - Sài Gòn.

Bên cạnh đánh giá nguy cơ đâm va trên khu vực biển rộng (phù hợp với khu đầu luồng Vũng Tàu theo CPA và T_{CPA} , luận án đã đề xuất và khảo sát các hoa tiêu làm việc tại khu vực này để xây dựng phương án đánh giá nguy cơ đâm va giữa các tàu thuyền hoạt động trong luồng lạch hẹp nói chung và luồng Vũng Tàu - Sài Gòn nói riêng theo ETA của các tàu thuyền đến các điểm nóng giao thông. Kết quả khảo sát cũng xác định được 10 điểm nóng giao thông (vị trí có nguy cơ đâm va cao giữa các tàu thuyền trên tuyến luồng Vũng Tàu - Sài Gòn).

Kết quả phát hiện nguy cơ đâm va giữa hai tàu bất kỳ của mô hình AI cho thấy độ chính xác của mô hình đạt khoảng 96%. Giá trị này có thể chấp nhận được trong dự đoán nguy cơ đâm va và vì với việc liên tục dự đoán theo thời gian, vị trí các tàu thay đổi liên tục, kết quả dự đoán cũng sẽ thay đổi, sẽ xóa được các điểm đen nơi không thể dự báo được chính xác. Hơn thế nữa, nếu đưa vào thực tế, các thông tin dữ liệu liên tục được bổ sung sẽ kéo theo mô hình học máy sẽ được huấn luyện liên tục và trở thành công cụ phát hiện đâm va và tàu với độ chính xác ngày càng được cải thiện.

2. Kiến nghị

Kết quả nghiên cứu cho thấy việc huấn luyện AI đã được thực hiện tốt với dữ liệu giả lập. Để có thể triển khai áp dụng trong thực tế, cơ sở dữ liệu cần

được kết nối với thông tin về sự di chuyển thực tế trên tuyến luồng Vũng Tàu - Sài Gòn. Việc làm này vừa để kiểm chứng thực tế, vừa tạo thêm cơ sở dữ liệu để huấn luyện AI. Kiến nghị các cơ quan quản lý tạo điều kiện cho nghiên cứu sinh tiếp cận, kết nối với hệ thống VTS để tiến hành phát triển ứng dụng của kết quả nghiên cứu vào thực tiễn.

Kết quả nghiên cứu có thể được áp dụng không chỉ nâng cao hiệu quả quản lý an toàn hàng hải tại các trạm VTS mà còn có thể áp dụng cho các hoa tiêu, thuyền trưởng dẫn tàu trên tuyến luồng này. Một khi nắm rõ tình hình giao thông tại khu vực, họ có thể điều chỉnh hoạt động của con tàu đang điều khiển, tránh phải chờ VTS thông tin mới hành động.

DANH MỤC CÁC BÀI BÁO KHOA HỌC ĐÃ CÔNG BỐ CÓ LIÊN QUAN ĐẾN LUẬN ÁN

I. Các công bố trong nước

[1]. Hoàng Hồng Giang, Phạm Văn Thuận, Trần Văn Lượng, Lương Tú Nam. (2024). *Ứng dụng mạng nơ-ron dự đoán nguy cơ đâm va giữa các tàu thuyền trên biển*, Tạp chí Giao thông vận tải, số 729 tháng 5/2023, ISSN 2354-0818, tr. 153 - 155.

[2]. Hoàng Hồng Giang, Phạm Văn Thuận, Trần Văn Lượng (2024). *Nghiên cứu xác định nguy cơ đâm va giữa các tàu thuyền khu vực luồng Sài Gòn - Vũng Tàu*. Tạp chí Khoa học Công nghệ Hàng hải, số tháng 78-04/2024, tr. 27 - 30.

II. Các công bố quốc tế

[1]. Hoang Hong Giang; Pham Van Thuan; Tran Van Luong; Luong Tu Nam (2021). *Evaluation of the Risk of Collision Between Two Target Ships Based On Observation Data From A Third Party*. Turkish Online Journal of Qualitative Inquiry (TOJQI), Volume 12, Issue 7, July, 2021, p. 1413 - 1426.

[2]. Hoang Hong Giang; Pham Van Thuan; Tran Van Luong; Luong Tu Nam (2022). *An evaluation model of the risk of collision between two target ships in a pollution management strategy in coastal waters*. International Journal of Knowledge-Based Development (IJKBD), Vol. 12, No. 3/4, 2022. **SCOPUS**, p. 195 - 216.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1]. Ban Chấp hành Trung ương, *Nghị quyết số 36-NQ/TW ngày 22-10-2018, về chiến lược phát triển bền vững kinh tế biển Việt Nam đến năm 2030, tầm nhìn 2045.*
- [2]. Phạm Tú Anh, *Bố trí VTS cho hệ thống cảng thành phố Hồ Chí Minh trong tương lai*, luận văn thạc sĩ, 2003.
- [3]. Lê Tuấn Anh, *Nghiên cứu ứng dụng hệ thống VTS tại cảng Hải Phòng*, luận văn thạc sĩ, 2010.
- [4]. PGS.TS. Phạm Văn Cương, *Nghiên cứu xây dựng mô hình mô phỏng tránh đâm va tàu thuyền trên biển để cảnh báo, trợ giúp tránh va tàu thuyền trên biển áp dụng cho tàu thuyền nhỏ*, đề tài NCKH cấp Bộ, 2011.
- [5]. PGS.TS. Phạm Văn Cương, *Nghiên cứu chế tạo thiết bị hỗ trợ tự động đồ giải các thông số mục tiêu (có kết nối radar) phòng tránh đâm va tàu thuyền trên biển dùng cho các tàu thuyền nhỏ*, đề tài NCKH cấp Bộ, 2013.
- [6]. Hoàng Văn Hùng, *Trung tâm dịch vụ giao thông tàu biển VTS và thực tiễn tại Việt Nam*, đồ án tốt nghiệp, 2009.
- [7]. Trần Đắc Sửu, Nguyễn Văn Phòng, *Hướng dẫn giải thích qui tắc phòng ngừa đâm va tàu thuyền trên biển 1972*, Nhà xuất bản Giao thông vận tải, Hà Nội, 2000.
- [8]. Phạm Văn Thuận, *Qui tắc phòng ngừa đâm va tàu thuyền trên biển 1972*, Nhà xuất bản Khoa học và Kỹ thuật, Hà Nội, 2012.
- [9]. Phạm Văn Thuận, *Nghiên cứu xây dựng bộ bài tập bổ sung cho thiết bị mô phỏng huấn luyện Radar/ARPA*, luận văn Thạc sĩ kỹ thuật, 2004.
- [10]. Phạm Văn Thuận, “*Xây dựng hệ công thức tính toán các thông số chuyển động của tàu mục tiêu từ kết quả quan sát*”, Tạp chí Khoa học Công nghệ Hàng hải số 28, 2011.

- [11].Luu Việt Hùng, Trần Văn Lượng, Nguyễn Việt Thành, *Sổ tay đảm bảo an toàn hàng hải khu vực ven biển Việt Nam*, Nhà xuất bản Hàng hải, 2017.
- [12].Bùi Xuân Toại, Trương Gia Việt (Biên dịch), *Trí tuệ nhân tạo - Các cấu trúc và chiến lược giải quyết vấn đề*, NXB Thống kê, 2000.
- [13].Bạch Hưng Khang, Hoàng Kiếm, *Trí tuệ nhân tạo - Các phương pháp và ứng dụng*, NXB Khoa học kỹ thuật, 1989.
- [14].PGS.TS. Nguyễn Thanh Thủy, *Trí tuệ nhân tạo - Các phương pháp giải quyết vấn đề và kỹ thuật xử lý tri thức*, NXB Giáo dục, 1995.
- [15].Nguyễn Trung Tuấn, *Trí tuệ nhân tạo (tài liệu dùng cho sinh viên, kỹ sư, cử nhân ngành CNTT)*, NXB Giáo dục, 1998.
- [16].Lê Thị Nhung, *Dữ liệu lớn và một số thách thức đối với phân tích dữ liệu lớn*, Học viện Nông nghiệp Việt Nam, 2024.
- [17].Lê Văn Thức, *Xây dựng mô hình đánh giá rủi ro tai nạn đâm va tại cảng biển Vũng Tàu, ứng dụng phòng ngừa tai nạn và hỗ trợ công tác cải tạo, nâng cấp luồng hàng hải''*, 2022, Luận án tiến sĩ.

Tiếng Anh

- [18]. IMO, *Resolution A.1158 (32) Guidelines for Vessel Traffic Services*, 2022.
- [19]. IALA, *VTS manual*, Edition 8, 2021.
- [20]. IMO, *International regulation for preventing collision at sea 1972 and amendments*.
- [21]. Krzysztof Marcjan, Lucjan Gucma, Kotkowska, *The collision risk Management for ships navigating on coastal waters based on ship domain and near-miss concept*, European Research study journal, Volum XXIV, Issue 4, 2021.
- [22]. Yuzuru Yamakage, Seishi Okamoto, *Toward AI for human beings: Human centric AI Zinzai*, Fujitsu Science Technology journal, vol. 53, No.1, 2017.

- [23]. George F. Luger, William A. Stubblefield, *Artificial Intelligence - Structure and Strategies for Complex Problem Solving(3rd edition)*, Wesley Publishing Company, 1997.
- [24]. Elaine Rich, Kevin Knight, *Artificial Intelligence (Second Edition)*, McGraw-Hill, 1991.
- [25]. Tom M. Mitchell, *Machine Learning*, McGraw Hill, Inc, 1997.
- [26]. Jean-Louis Laurière, *Problem solving methods and Artificial Intelligence*, Prentice Hall, 1987, 1990.
- [27]. Judea Pearl, *Heuristics - Intelligent Search Strategies for Computer Problem Solving*, Addison-Wesley Publishing company, 1984.
- [28]. Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, Second edition, 2002.
- [29]. Pham Van Thuan, *Study on handling characteristics concerning ship's loading conditions*, doctoral dissertation, 2009.
- [30]. Van Luong Tran, Namkyun IM, *A study on ship automatic berthing with assistance of auxiliary devices*, International Journal of Naval Architecture and Ocean Engineering, 2012
- [31]. Tran Van Luong, *A study on automatic ship control based on artificial neural networks and free running model ship experiment*, doctoral dissertation, 2012.
- [32]. Tran Van Luong, *Ship Automatic Control Based on Intelligent Systems*, ISBN: 978-604-937-146-2, NXB Hàng hải, 2017.
- [33]. Van Luong Tran, Namkyun IM, *Automatic berthing control of ship using artificial neural networks with assistance of side thruster and tugboat*, MARSIM Panama, 2009.
- [34]. Yosra Hajjaji, Wadii Boulila , Imed, *Big data and IoT-based applications in smart environments: A systematic review*, Computer science review, volume 39, 2021.

- [35]. Ayodele Adebisi, Olawande Daramola, *Big data stream analysis: a systematic literature review*, Journal of big data, 2019.
- [36]. Sun D, Zhang G, Zheng W, Li K. *Key technologies for big data stream computing*. In: Li K, Jiang H, Yang LT, Guzzocrea A, editors. Big data algorithms, analytics and applications. New York: Chapman and Hall/CRC; 2015. p. 193–214. ISBN 978-1-4822-4055-9.
- [37]. Chung D, Shi H. *Big data analytics: a literature review*. J Manag Anal. 2015;2(3):175–201.
- [38]. Lu J, Li D. *Bias correction in a small sample from big data*. IEEE Trans Knowl Data Eng. 2013;25(11):2658–63.
- [39]. Wienhofen LW, Mathisen BM, Roman D. *Empirical big data research: a systematic literature mapping*. CoRR, abs/1509.03045. 2015.
- [39]. X. C. L. Z. S. a. Z. M. X. P. Tong, "Vessel trajectory prediction in curving channel of inland river," *International Conference on Transportation Information and Safety, Wuhan, China, 2015*.
- [40] Y. Y. Q. J. Y. Y. H. Cheng, "An improved Markov method for prediction of user mobility," *International Conference on Network and Service Management, 2016*.
- [41] V. W. A. J. M. Schreier, "An Integrated Approach to Maneuver-Based Trajectory Prediction and Criticality Assessment in Arbitrary Road Environments," *IEEE Trans. Intell. Transp. Syst.*, 2016.
- [42] K. J. N. H. C. T. G. L. A. G. S.J. Qiao, "Trajectory Prediction Algorithm Based on Gaussian Mixture Model," *J. Softw*, 2015.
- [43] A. T. C. G. S. H. Rong, "Ship trajectory uncertainty prediction based on a Gaussian Process model," *Ocean Eng*, 2019.
- [44] V. A. M. V. F. Mazzarella, "Knowledge-based vessel position prediction using historical AIS data," *Proceedings of the Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE, 2015*.

- [45] B. Ristic, B. La Scala, M. Morelande and N. Gordon, "Statistical analysis of motion patterns in AIS data: Anomaly detection and motion prediction," *In Proceedings of the 2008 11th International Conference on Information Fusion, Cologne, Germany, 30 June–3 July 2008; IEEE: Piscataway, NJ, USA, 2008.*
- [46] A. F. B.-O. E. E. B. S. Hexeberg, "AIS-based vessel trajectory prediction," *Proceedings of the 20th International Conference on Information Fusion (FUSION), IEEE, 2017.*
- [47] S. H. A. F. B.-O. E. E. B. B.R. Dalsnes, "The neighbor course distribution method with gaussian mixture models for AIS-based vessel trajectory prediction," *Proceedings of the 21st International Conference on Information Fusion (FUSION), IEEE , 2018.*
- [48] K. H. a. T. O. D. Kim, "Distributed stochastic search algorithm for multi-ship encounter situations," *Journal of Navigation*, vol. 70, 2017.
- [49] P. O. C. G. S. L. P. Perera, "Maritime Traffic Monitoring Based on Vessel Detection, Tracking, State Estimation, and Trajectory Prediction," *IEEE Trans. Intell. Transp. Syst.*, 2012.
- [50] F. C. Q. H. C. Y. T. Xu, "Research on estimation of AIS vessel trajectory data and Kalman filter algorithm," *Model Electron. Tech*, 2014.
- [51] J. M. Y. J. M. J.H. Park, "Trajectory estimation of a moving object using Kalman filter and Kohonen networks," *Robotica*, 2007.
- [52] P. Passenier, "An Adaptive Track Predictor for Ships," *Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands*, 1989.
- [53] A. Czapiewska and J. Sadowski, "Algorithms for Ship Movement Prediction for Location Data Compression," *TransNav Int. J. Mar. Navig. Saf. Sea Transp.*, 2015.
- [54] C. Schöller, V. Aravantinos, F. Lay and A. Knoll, "What the constant velocity model can teach us about pedestrian motion prediction.," *IEEE*

Robot. Autom. Lett., 2020.

- [55] T. Johansen, T. Perez and A. Cristofaro, "Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment.," *IEEE Trans. Intell. Transp. Syst.*, 2016.
- [56] P. Last, C. Bahlke, M. Hering-Bertram and L. Linsen, "Comprehensive Analysis of Automatic Identification System (AIS) Data in Regard to Vessel Movement Prediction," *J. Navig.*, 2014.
- [57] L. Sang, X. Yan, A. Wall, J. Wang and Z. Mao, "CPA calculation method based on AIS position prediction," *J. Navig.*, 2016.
- [58] G. Y. S. K. Z.] J. Liu, "Vessel Trajectory Prediction Model Based on AIS Sensor Data and Adaptive Chaos Differential Evolution Support Vector Regression (ACDE-SVR)," *Appl. Sci.*, 2019.
- [59] D. Rumelhart, G. Hinton and R. Williams, "Learning representations by back-propagating errors.," *Nature* , 1986.
- [60] H. Tang, Y. Yin and H. Shen, "A model for vessel trajectory prediction based on long short-term memory neural network," *J. Mar. Eng. Technol. Proc. Inst. Mar. Eng. Sci. Technol.*, 2022.
- [61] M. Gao, G. Shi and S. Li, "Online Prediction of Ship Behavior with Automatic Identification System Sensor Data Using Bidirectional Long Short-Term Memory Recurrent Neural Network," *Sensors*, 2018.
- [62] C. Wang and Y. Fu, "Ship Trajectory Prediction Based on Attention in Bidirectional Recurrent Neural Networks," *In Proceedings of the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), Shenyang, China, 13–15 November 2020.*
- [63] S. Mehri, A. Alesheikh and A. Basiri, "A Contextual Hybrid Model for Vessel Movement Prediction," *IEEE Access*, 2021.
- [64] M. Ding, W. Su, Y. Liu, J. Zhang and J. Wu, "A Novel Approach on Vessel

- Trajectory Prediction Based on Variational LSTM," *In Proceedings of the 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 27–29 June 2020.*
- [65] C. Wang, H. Ren and H. Li, "Vessel trajectory prediction based on AIS data and bidirectional GRU," *In Proceedings of the 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), Chongqing, China, 10–12 July 2020.*
- [66] L. Z. Y. Z. L. T. Wang, "Trajectory Predictor by Using Recurrent Neural Networks in Visual Tracking," *IEEE T. Cybern*, 2017.
- [67] G. Y. S. S. L. M. Gao, "Online Prediction of Ship Behavior with Automatic Identification System Sensor Data Using Bidirectional Long Short-Term Memory Recurrent Neural Network," *Sensors 2018*.
- [68] L. H. H. G. C. P. Payeur, "Trajectory prediction for moving objects using artificial neural networks," *IEEE Trans. Ind. Electron*, 1995.
- [69] D. Nguyen, L. Chan and M. Ali, "Vessel Trajectory Prediction using Sequence-to-Sequence Models over Spatial Grid," *In Proceedings of the the 12th ACM International Conference, Hamilton, New Zealand, 25–29 June 2018*.
- [70] N. Forti, L. Millefiori, P. Braca and P. Willett, "Prediction of vessel trajectories from ais data via sequence-to-sequence recurrent neural networks," *In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020.*
- [71] J. Sekhon and C. A. Fleming, "A Spatially and Temporally Attentive Joint Trajectory Prediction Framework for Modeling Vessel Intent," *Learn. Dyn. Control*, 2020.
- [72] L. You, S. Xiao, Q. Peng, C. Claramunt and J. Zhang, "ST-Seq2Seq: A Spatio-Temporal Feature-Optimized Seq2Seq Model for Short-Term

- Vessel Trajectory Prediction," *IEEE Access*, 2020.
- [73] S. Capobianco, L. Millefiori, N. Forti, P. Braca and P. Willett, "Deep Learning Methods for Vessel Trajectory Prediction based on Recurrent Neural Networks," *IEEE Trans. Aerosp. Electron. Syst.*, 2021.
- [74] S. Zhang, L. Wang, M. Zhu, S. Chen and Z. Zeng, "A Bi-directional LSTM Ship Trajectory Prediction Method based on Attention Mechanism," *In Proceedings of the 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 March 2021*.
- [75] S. Wang and Z. He, "A prediction model of vessel trajectory based on generative adversarial network," *J. Navig.*, 2021.
- [76] I. Goodfellow, J. Pouget-Abadie, M. Mirza, X. Bing, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, "Generative adversarial nets," *In Proceedings of the Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014*..
- [77] D. Nguyen and R. Fablet, "TrAISformer-A generative transformer for AIS trajectory prediction," *arXiv*, 2021.
- [78] K. Q. M. E. C. C. Abdulllah Mohamed, "Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction," *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [79] G. C. H. X. S. G. H. Feng, "IS-STGCNN: an Improved Social spatial-temporal graph convolutional neural network for ship trajectory prediction," *Ocean. Eng.*, 2022.
- [80] Y. Z. N. L. L. M. W. L. J. S. J. Jiang, "Prediction of ship trajectory in nearby port waters based on attention mechanism model," *J. Mar. Sci. Eng.*, 2023.
- [81] M. V. K. B. G. Pallotta, "Vessel pattern knowledge discovery from AIS

- data: a framework for anomaly detection and route prediction," *Entropy*, 15 (12), 2013.
- [82] S. H. P. B. K. B. G. Pallotta, "Context-Enhanced Vessel Prediction Based On Ornstein-Uhlenbeck Processes Using Historical AIS Traffic Patterns : Real-World Experimental Results," *Proceedings of the 17th International Conference on Information Fusion (FUSION)*, IEEE, 2014.
- [83] R. Laxhammar, "Anomaly detection for sea surveillance," *In Proceedings of the 2008 11th International Conference on Information Fusion, Cologne, Germany, 30 June–3 July 2008; IEEE: Piscataway, NJ, USA, 2008.*
- [84] L. Zhao and G. Shi, "Maritime anomaly detection using density-based clustering and recurrent neural network," *J. Navig.*, 2019.
- [85] M. Riveiro, G. Pallotta and M. Vespe, "Maritime anomaly detection," *A review. Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 2018.
- [86] R. McLane and J. Wolf, "Symbolic and Pictorial Displays for Submarine Control," *IEEE Trans. Hum. Factors Electron*, 1967.
- [87] Q. M. T. F. F. L. Zhang, "Big AIS data based spatial- temporal analyses of ship traffic in Singapore port waters," *Transp. Res. Part E: Logist. Transp. Rev.*, 2017.
- [88] Z. W. Z. Z. Z. Liu, "A novel framework for regional collision risk identification based on AIS data," *Appl. Ocean Res.*, 89, 2019.
- [89] Z. S. C. Z. C. X. Q. C. D. H. Y. Z. Y. Wen, "Automatic ship route design between two ports: a data-driven method," *Appl. Ocean Res.*, 96, 2020.
- [90] G. Z. L. R. E. R. G.-B. H. E. Tu, "Exploiting AIS data for intelligent maritime navigation: a comprehensive survey from data to methodology," *IEEE Trans. Intell. Transp. Syst.*, 2017.
- [91] T. M. K.G. Aarsæther, "Estimating navigaion patterns from AIS," *J. Navig.*, 62 (04), 2009.

- [92] G. F. E. S. R. Laxhammar, "Anomaly detection in sea traffic - a comparison of the gaussian mixture model and the kernel density estimator," *Proceedings of the 12th International Conference on Information Fusion, IEEE*, 2009.
- [93] B. S. M. M. N. G. B. Ristic, "Statistical analysis of motion patterns in AIS Data: Anomaly detection and motion prediction," *Proceedings of the 11th International Conference on Information Fusion, IEEE*, 2008.
- [94] L. van Breda and P. Passenier, "Effect of path prediction on navigational performance.," *J. Navig.*, 1998.
- [95] K. Aarsæther and T. Moan, "Estimating navigation patterns from AIS," *J. Navig.* , 2009.
- [96] H. Tang, L. Wei, Y. Yin, H. Shen and Y. Qi, "Detection of abnormal vessel behaviour based on probabilistic directed graph model," *J. Navig.*, 2020.
- [97] M. Łacki, "Intelligent prediction of ship maneuvering," *TransNav Int. J. Mar. Navig. Saf. Sea Transp.*, 2016.
- [98] T. Xu, X. Liu and X. Yang, "Ship Trajectory online prediction based on BP neural network algorithm," *In Proceedings of the 2011 International Conference of Information Technology, Computer Engineering and Management Sciences, Nanjing, China, 24–25 September 2011; IEEE: Piscataway, NJ, USA, 2011*, vol. 1.
- [99] H. Zhou, Y. Chen and S. Zhang, "Ship trajectory prediction based on BP neural network," *J. Artif. Intell.*, 2019.
- [100] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput*, 1997.
- [101] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv*, 2014.
- [102] K. Cho, B. van Merriënboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H.

- Schwenk and Y. Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv*, 2014.
- [103] S. apobianco, N. Forti, L. Millefiori, P. Braca and P. Willett, "Recurrent Encoder-Decoder Networks for Vessel Trajectory Prediction with Uncertainty Estimation," *IEEE Trans. Aerosp. Electron. Syst.*, 2022.
- [104] G. Y. S. K. Z.] J. Liu, "Vessel Trajectory Prediction Model Based on AIS Sensor Data and Adaptive Chaos Differential Evolution Support Vector Regression (ACDE-SVR)," *Appl. Sci.*, 2019.
- [105] L. Z. H. a. J. H. Xu, "A Spatio-Temporal Convolutional Neural Network for Skeletal Action Recognition," *International Conference on Neural Information Processing*, 2017.
- [106] M. C. T. Baklacioglu, "Aero-propulsive modelling for climb and descent trajectory prediction of transport aircraft using genetic algorithms," *Aeronaut. J.*, 2014.
- [107] T. W. Y. Z. F. D. S. Y. Yang, "Particle filter based on simulated annealing for target tracking. J. Optoelectron," *Laser*, 2011.
- [108] T. P. a. A. C. T. A. Johansen, "Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, 2016.
- [109] M. T. W. D. Z. a. T. L. H. Y. Kuwata, "Safe maritime autonomous navigation with COLREGS, using velocity obstacles," *IEEE Journal of Oceanic Engineering*, vol. 39, 2013.
- [110] A. Lazarowska, "Ant colony optimization based navigational decision support system," *Procedia Computer Science*, vol. 35, 2014.
- [111] N. V. D. V. T. P. A. S. a. E. H.-V. H. D. Nguyen, "A method for knowledge representation to design intelligent problems solver in mathematics based on rela-ops model," *IEEE Access*, vol. 8, 2020.

[112] C. T. a. R. Bucknall, "Path-planning algorithm for ships in close-range encounters," *Journal of Marine Science and Technology*, 2010.

PHỤ LỤC 1: CÁC THUẬT TOÁN AI ĐƯỢC THỬ NGHIỆM ĐỐI VỚI TẬP DỮ LIỆU

1. Mô hình học máy (Machine Learning Model)

1.1. Mạng Neural Network

- Input: $x_{a1}, y_{a1}, x_{b1}, y_{b1}, x_{a2}, y_{a2}, x_{b2}, y_{b2}, W_{px}, W_{py}$
- Output: ETA_{Wpa}, ETA_{Wpb}
- **Model:** Neural Network 2 lớp
- **Kết quả:**

Ta có bảng độ đo chính xác như sau:

- Tập huấn luyện

	precision	recall	f1-score	support
0	0.94	0.79	0.86	15421
1	0.94	0.98	0.96	49373
accuracy			0.94	64794
macro avg	0.94	0.89	0.91	64794

- Tập kiểm tra

	precision	recall	f1-score	support
0	0.00	0.00	0.00	328
1	0.99	1.00	1.00	55416
accuracy			0.99	55744
macro avg	0.50	0.50	0.50	55744

Kết luận: Mô hình đạt kết quả tốt trên tập huấn luyện với F1-score trung bình (macro avg) của 2 lớp là 91% và dự đoán tốt với cả 2 lớp 0 và 1.

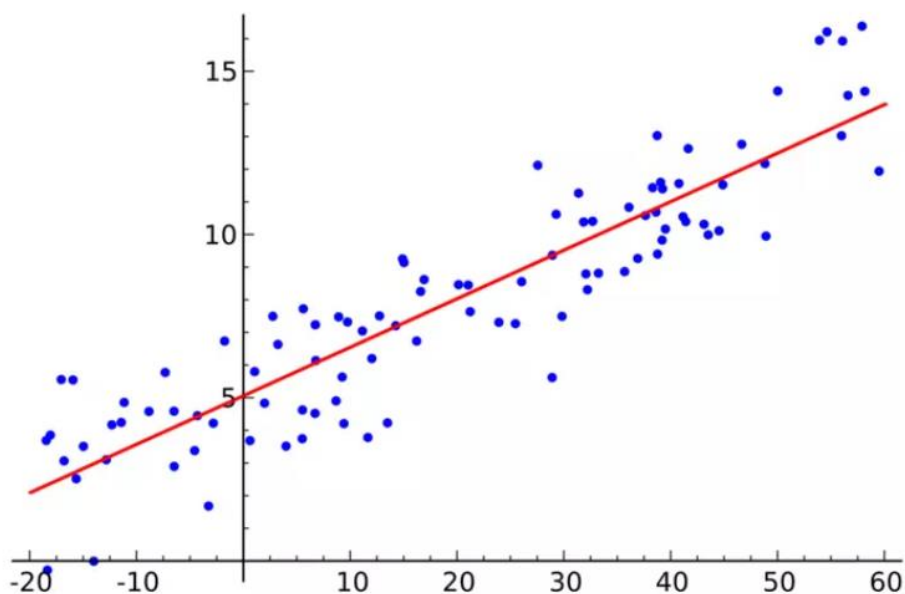
Như vậy, chỉ qua 2 layer, trên tập kiểm tra mô hình dự đoán 100% dữ liệu với lớp 1 nhưng không dự đoán đúng bất kỳ nhãn 0 nào. Mô hình học sâu chỉ có F1-score trung bình (macro avg) của 2 lớp trên tập kiểm tra là 50% => Mô hình đang hoạt động không tốt khó kiểm soát trong quá trình training và dễ bị overfitting.

Cần thử nghiệm thêm với một số mô hình Machine Learning khác.

1.2. Thuật toán hồi quy tuyến tính (*Linear Regression*)

Linear Regression (Hồi quy tuyến tính) là một thuật toán *Supervised learning*. Đây là một phương pháp để dự đoán biến phụ thuộc y dựa trên giá trị của biến độc lập X . Thuật toán được sử dụng cho các trường hợp chúng ta muốn dự đoán một số lượng liên tục.

Mục tiêu là để làm sao một đường thẳng $f(x)$ có thể tạo được sự phân bố gần nhất với hầu hết các điểm X . Do đó làm giảm khoảng cách (sai số) của các điểm dữ liệu cho đến đường đó.



Hình PL1.1. Minh họa về *Linear Regression*

Lý thuyết

$x = [x_1, x_2, x_3, \dots, x_n]$ là một vector chứa thông tin *input*.

$y = [y_1, y_2, y_3, \dots, y_n] = f(x)$ là một số vô hướng biểu diễn *output*.

Một hàm số đơn giản nhất có thể mô tả mối quan hệ giữa *output* và *input* là:

$$y \approx f(x) = \hat{y}$$

$$f(x) = w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_nx_n + w_0 \quad (\text{PL1.1})$$

Trong đó $w_1, w_2, w_3, \dots, w_n$ là các hằng số, w_0 còn được gọi là bias.

Mối quan hệ $y \approx f(x)$ bên trên là một mối quan hệ tuyến tính (linear).

Mục tiêu: Tìm các hệ số tối ưu $\{w_1, w_2, w_3, \dots, w_n, w_0\}$ sao cho giá trị dự đoán $f(x)$ gần với giá trị thực y nhất có thể.

Hàm số $L(w)$ được gọi là hàm tổn thất (loss function) của bài toán Linear Regression, biểu thị sai số giữa giá trị dự đoán và giá trị thực:

$$L(w) = \frac{1}{2} \sum_{i=1}^n (y_i - \bar{x}_i w)^2 \quad (\text{PL1.2})$$

Giá trị của w là giá trị cần tìm sao cho hàm mất mát đạt giá trị nhỏ nhất được gọi là *điểm tối ưu* (optimal point), ký hiệu:

$$w^* = \underset{w}{\text{arg min}} L(w) \quad (\text{PL1.3})$$

Đạo hàm theo w của hàm mất mát:

$$\frac{\partial L(w)}{\partial w} = \bar{X}^T (\bar{X}w - y) \quad (\text{PL1.4})$$

w là nghiệm của bài toán tối ưu. Do đó cần giải phương trình đạo hàm (gradient) Loss function $L(w)$ (theo tham số w) bằng 0. Các giá trị w sẽ được cập nhật lại trong quá trình training và dừng lại cho đến khi hàm loss đủ nhỏ thỏa mãn.

Áp dụng

- Input: $x_{a1}, y_{a1}, x_{b1}, y_{b1}, x_{a2}, y_{a2}, x_{b2}, y_{b2}, W_{px}, W_{py}$
- Output: ETA_{Wpa}, ETA_{Wpb}
- **Model:** Linear Regression
- **Kết quả:**

Dùng độ đo sai số bình phương trung bình MSE (Mean Square Error)

+ Độ đo sai số như sau:

<i>MSE</i>	Tập huấn luyện	Tập kiểm tra
tàu A	3.0176	98.7165
tàu B	10.5549	59.7410

Ta có bảng độ đo chính xác như sau:

• Tập huấn luyện

	precision	recall	f1-score	support
0	0.78	0.80	0.79	15421
1	0.89	0.99	0.94	49373
accuracy			0.90	64794
macro avg	0.86	0.86	0.86	64794

• Tập kiểm tra

	precision	recall	f1-score	support
0	0.00	0.00	0.00	328
1	1.00	1.00	1.00	55416
accuracy			1.00	55744
macro avg	0.50	0.50	0.50	55744

Kết luận:

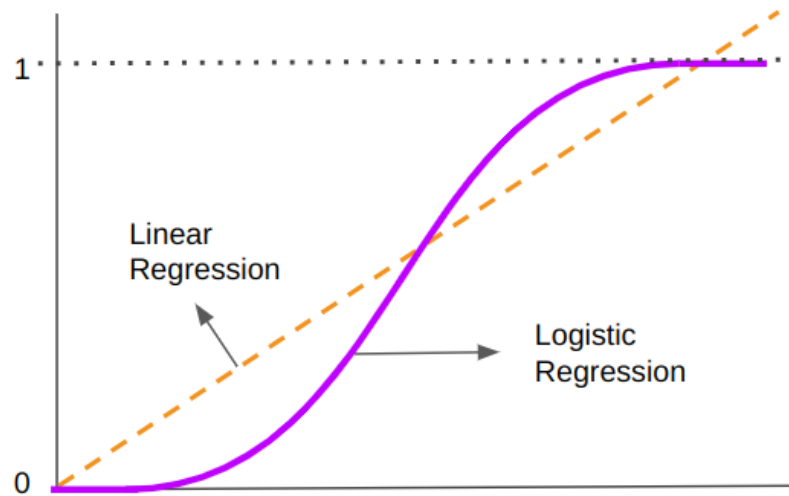
+ Trên tập huấn luyện: Độ đo sai số của tàu B tương đối cao và cao hơn so với tàu A => Mô hình dự đoán ETA của tàu A tốt hơn. Mô hình có F1-score trung bình (macro avg) của 2 lớp là 86% và cao ở cả hai lớp (0 và 1).

+ Trên tập kiểm tra: Độ đo sai số của tàu A và B đều rất cao và cao

hơn so với tập huấn luyện => Mô hình có dấu hiệu bị overfitting. Mô hình có độ chính xác là 100% đối với lớp 1, nhưng mô hình không dự đoán đúng bất kỳ mẫu nào thuộc lớp 0 => Mô hình không tốt, có F1-score trung bình (macro avg) của 2 lớp chỉ là 50%.

1.3. Mô hình hồi quy Logistics (*Logistic Regression*)

Logistic Regression là một thuật toán *Supervised learning*. Đây là phương pháp cho bài toán phân loại được dùng để gán các đối tượng cho 1 tập hợp giá trị rời rạc (như 0, 1, 2, ...).



Hình PL1.2. Minh họa Logistic Regression

Lý thuyết

$x = [x_1, x_2, x_3, \dots, x_n]$ là một vector chứa thông tin *input*.

$y = [y_1, y_2, y_3, \dots, y_n]$ là tập hợp các giá trị rời rạc chứa thông tin *output*.

Đầu ra dự đoán của Logistic Regression thường được viết chung dưới dạng:

$$f(x) = \theta(\bar{X}w) = \hat{y} \quad (\text{PL1.5})$$

w là các hằng số, x là giá trị input, $f(x)$ là đầu ra dự đoán

Trong đó θ được gọi là *logistic function*. Một số activation cho mô hình tuyến tính là sigmoid, tanh, ...

Ví dụ:

+ Hàm Sigmoid: được sử dụng nhiều nhất, vì nó bị chặn trong khoảng (0,1) và có đạo hàm tại mọi điểm nên có thể dùng gradient descent:

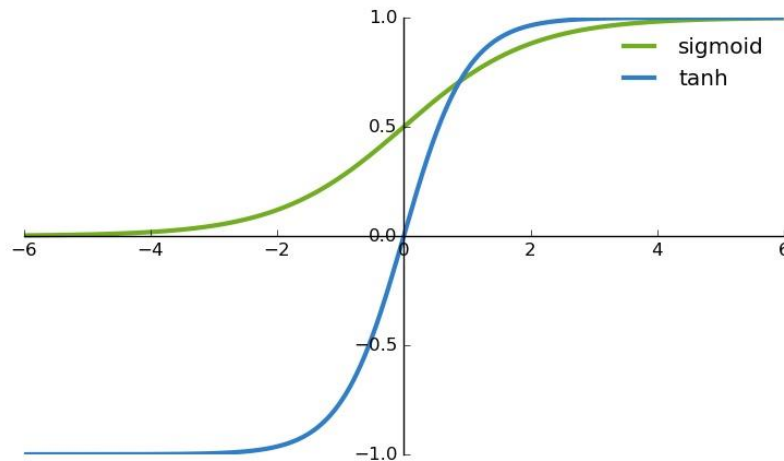
$$f(x) = \frac{1}{1 + e^{-x}} \quad (\text{PL1.6})$$

Đặc biệt công thức đạo hàm đơn giản:

$$\sigma' = \frac{e^{-x}}{(1 + e^{-x})^2} = \frac{1}{1 + e^{-x}} \frac{e^{-x}}{1 + e^{-x}} = \sigma(x)(1 - \sigma(x)) \quad (\text{PL1.7})$$

+ Hàm *tanh*: nhận giá trị trong khoảng (-1,1)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (\text{PL1.8})$$



Hình PL1.3. Hàm Sigmoid và hàm Tanh

Khi áp dụng hàm *activation Sigmoid* ta có công thức của *Logistic regression* là

$$\hat{y} = \sigma(\bar{X}w)$$

$$\hat{y} = \frac{1}{1 + e^{-(\bar{X}w)}} \quad (\text{PL1.9})$$

Cần tìm w để biểu thức sau đây đạt giá trị lớn nhất. Đây được hiểu là xác suất xảy ra sự kiện đầu ra y khi biết tham số w và dữ liệu đầu vào X .

$$P(y|X; w) \quad (\text{PL1.10})$$

Giá trị của w là giá trị cần tìm:

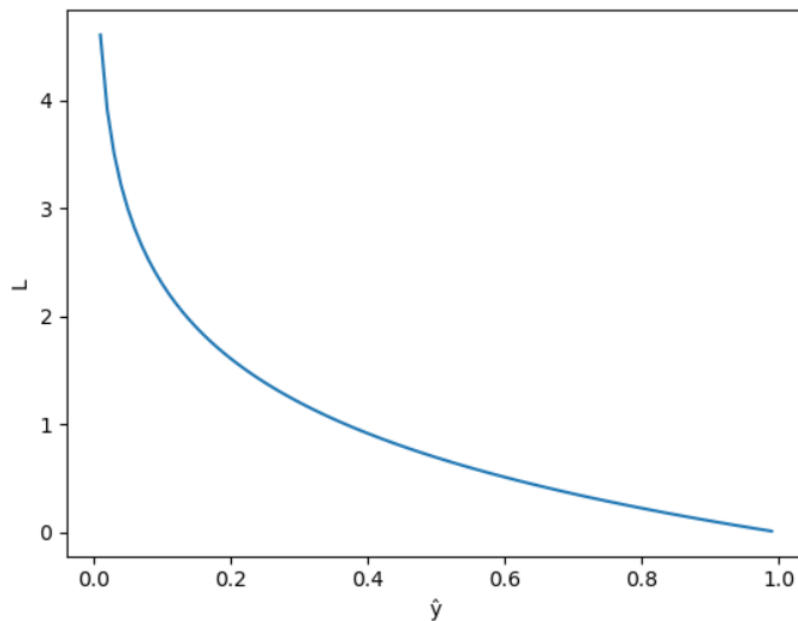
$$w^* = \underset{w}{\operatorname{arg\,max}} P(y | X; w) \quad (\text{PL1.11})$$

Hàm số $J(w)$ được gọi là hàm tổn thất (*loss function*) của bài toán Logistic Regression. Một phương pháp thường được sử dụng đó là lấy logarit tự nhiên (cơ số e) của *likelihood function* biến phép nhân thành phép cộng và để tránh việc số quá nhỏ. Sau đó lấy ngược dấu để được một hàm và coi nó là hàm mất mát. Lúc này bài toán tìm giá trị lớn nhất (*maximum likelihood*) trở thành bài toán tìm giá trị nhỏ nhất của hàm tổn thất (hàm này còn được gọi là *negative log likelihood*):

$$J(w) = -\log P(y|X; w) = -\sum_{i=1}^N (y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)) \quad (\text{PL1.12})$$

Trong đó \hat{y}_i là giá trị dự đoán. Về phải có tên gọi là *cross entropy*, thường được sử dụng để đo *khoảng cách* giữa hai phân phối (distributions).

+ Nếu $y_i = 1 \Rightarrow J(w) = -\sum_{i=1}^N (\log \hat{y}_i)$

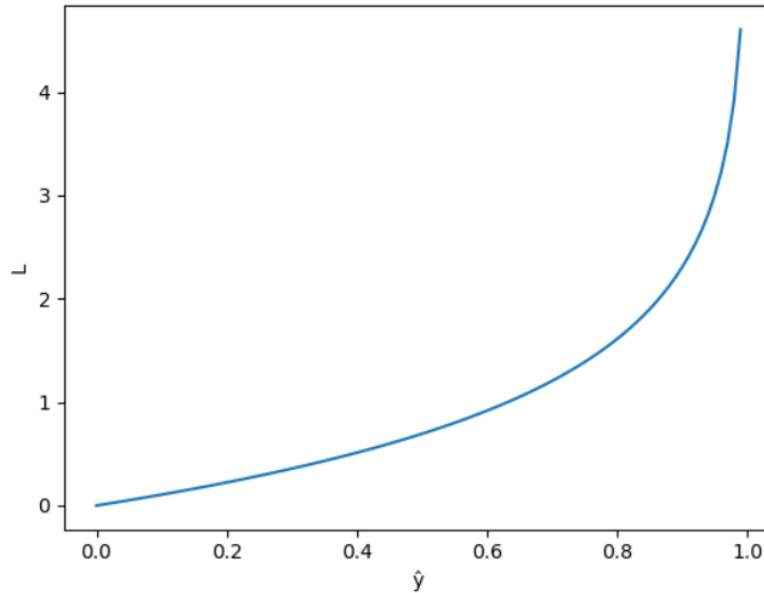


Hình PL1.3. Đồ thị hàm loss function trong trường hợp $y_i = 1$

- Hàm loss giảm dần từ 0 đến 1.
- Khi model dự đoán \hat{y}_i gần 1, tức giá trị dự đoán gần với giá trị thật \hat{y}_i thì loss nhỏ, xấp xỉ 0.

- Khi model dự đoán \hat{y}_i gần 0, tức giá trị dự đoán ngược lại giá trị thật y_i thì loss rất lớn.

+ Nếu $y_i = 0 \Rightarrow J(w) = -\sum_{i=1}^N (\log(1 - \hat{y}_i))$



Hình PL1.4. Đồ thị hàm loss function trong trường hợp $y_i = 0$

- Hàm loss tăng dần từ 0 đến 1.
- Khi model dự đoán \hat{y}_i gần 0, tức giá trị dự đoán gần với giá trị thật y_i thì loss nhỏ, xấp xỉ 0.
- Khi model dự đoán \hat{y}_i gần 1, tức giá trị dự đoán ngược lại giá trị thật y_i thì L rất lớn.

\Rightarrow Hàm loss nhỏ khi giá trị model dự đoán gần với giá trị thật và rất lớn khi model dự đoán sai, hay nói cách khác loss càng nhỏ thì model dự đoán càng gần với giá trị thật.

Để tìm giá trị của w là giá trị cần tìm sao cho hàm mất mát đạt giá trị nhỏ nhất, ta tính đạo hàm theo w của hàm mất mát:

$$\frac{\partial J(w; x_i, y_i)}{\partial w} = -\left(\frac{y_i}{\hat{y}_i} - \frac{1 - y_i}{1 - \hat{y}_i} \right) \frac{\partial \hat{y}_i}{\partial w} = \frac{\hat{y}_i - y_i}{\hat{y}_i(1 - \hat{y}_i)} \frac{\partial \hat{y}_i}{\partial w} \quad (\text{PL1.13})$$

Áp dụng công thức chuỗi đạo hàm để tìm được nghiệm của bài toán dạng như sau:

$$\frac{\partial \hat{y}_i}{\partial w} = \frac{\partial \hat{y}_i}{\partial s} \frac{\partial s}{\partial w} \quad (\text{PL1.14})$$

w là nghiệm của bài toán tối ưu. Do đó cần giải phương trình đạo hàm (gradient) Loss function $J(w)$ (theo w) bằng 0. Các giá trị w sẽ được cập nhật lại trong quá trình training và dừng lại cho đến khi hàm loss đủ nhỏ thỏa mãn.

Áp dụng

- Input: xa1, ya1, xb1, yb1, xa2, ya2, xb2, yb2, Wpx, Wpy
- Output: 0 (không có nguy cơ va chạm); 1 (có nguy cơ va chạm)
- **Mô hình:** Logistic Regression
- **Kết quả:**

Ta có bảng độ đo chính xác như sau:

- Tập huấn luyện

	precision	recall	f1-score	support
0	0.85	0.68	0.75	15421
1	0.91	0.96	0.93	49373
accuracy			0.89	64794
macro avg	0.88	0.82	0.84	64794

- Tập kiểm tra

	precision	recall	f1-score	support
0	0.00	0.00	0.00	328
1	0.99	0.96	0.98	55416
accuracy			0.96	55744
macro avg	0.50	0.48	0.49	55744

Kết luận:

+ Trên tập huấn luyện: Mô hình có F1-score trung bình (macro avg) của 2 lớp là 84% và cao ở cả hai lớp (0 và 1).

+ Trên tập kiểm tra: Mô hình có độ chính xác là 98% với dữ liệu thuộc lớp 1, nhưng mô hình không dự đoán đúng bất kỳ mẫu nào thuộc lớp 0 => Mô hình rất kém với F1-score trung bình (macro avg) của 2 lớp là 49%.

1.4. Mô hình cây quyết định (Decision Tree Regressor)

Decision Tree là thuật toán *supervised learning*, có thể giải quyết cả bài toán regression và classification. Cây quyết định nhận đầu ra là các giá trị liên tục (số thực) được gọi là cây hồi quy.

Tổng quan:

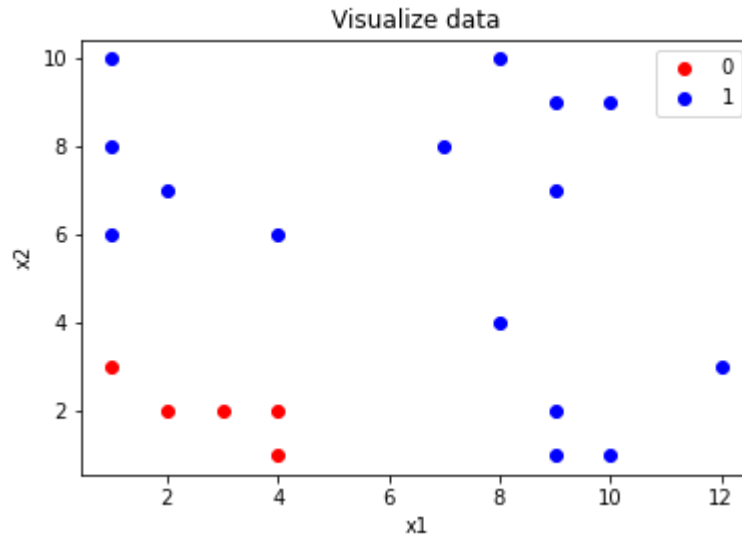
Một cây quyết định là một cấu trúc với mỗi một nút (node) là một phép so sánh, nhằm thực hiện thử một đặc trưng nào đó. Mỗi một nhánh (branch, edge, etc.) thể hiện kết quả của phép so sánh đó, và mỗi một ngọn, tức nút cuối (leaf node), biểu diễn kết quả dự đoán của thuật toán, sau khi tính. Đường đi từ gốc đến ngọn thể hiện luật hồi quy.

Mục tiêu chính là xây dựng được một mô hình có khả năng dự đoán giá trị của đầu ra dựa vào một vài biến đầu vào.

Lý thuyết

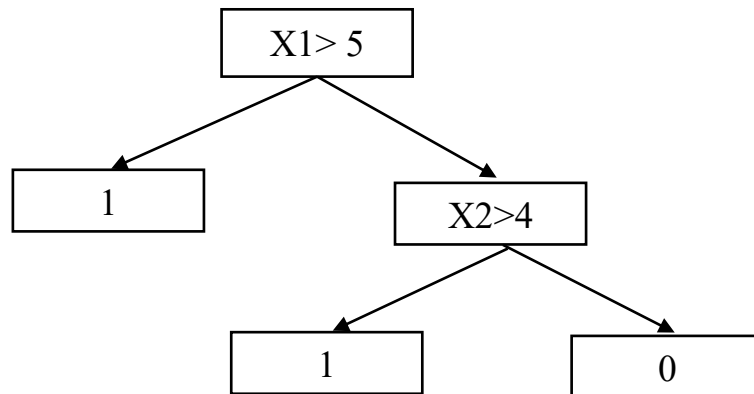
Bước huấn luyện ở thuật toán Decision Tree sẽ xây dựng ra một cây quyết định.

Giả sử có bài toán phân loại 2 lớp và mỗi dữ liệu có 2 thuộc tính là x_1 và x_2 . Dữ liệu khi vẽ biểu đồ scatter lên sẽ như sau:



Hình PL1.5. Minh họa dữ liệu

Tạo cây quyết định như sau:



Hình PL1.6. Minh họa cây quyết định

Để tạo cây quyết định phù hợp cần có những tiêu chí chọn các nhánh của cây quyết định dựa vào các chỉ số đánh giá như Gini Index, Độ tăng thông tin.

Gini index là độ đo tần suất một phần tử được chọn ngẫu nhiên từ tập hợp bị dán nhãn sai nếu nó được dán nhãn ngẫu nhiên theo phân phối nhãn trong tập dữ liệu con. Gini index được dùng để đánh giá xem việc phân chia ở node điều

kiện có tốt hay không. Để tính Gini index, trước hết mình sẽ tính chỉ số Gini, chỉ số Gini tính ở từng node.

$$Gini = 1 - \sum_{i=1}^c (p_i)^2 \quad (PL1.15)$$

Trong đó C là số lớp cần phân loại, $p_i = \frac{n_i}{N}$, n_i là số lượng phần tử ở lớp thứ i .

N là tổng số lượng phần tử ở node đó:

$$N = \sum_{i=1}^n n_i \Rightarrow \sum_{i=1}^N p_i = 1 \quad (PL1.16)$$

Chỉ số gini thấp nhất (bằng 0) khi node đó chỉ chứa dữ liệu của 1 lớp duy nhất, chỉ số gini cao nhất khi dữ liệu các lớp ở trong node đó cân bằng.

=> Ta mong muốn khi phân tách thì chỉ số gini ở những lớp con thấp.

Sau khi tính được chỉ số gini ở node cha và 2 node con được tính, ta tính được chỉ số gini index:

$$gini_{index} = gini(p) - \sum_{i=1}^K \frac{m_k}{M} gini(c_k) \quad (PL1.17)$$

Trong đó $gini(p)$ là chỉ số gini ở node cha, K là số node con được tách ra, $gini(c_k)$ là chỉ số gini ở node con thứ k . M là số phần tử ở node p , m_i là số phần tử ở node con thứ i ,

$$\sum_{i=1}^K m_i = M$$

Thực ra gini index tính độ lệch gini của node cha với tổng các giá trị gini có đánh trọng số của các node con.

Độ tăng thông tin (Information Gain) dựa vào khái niệm entropy và khái niệm nội dung thông tin trong lý thuyết thông tin.

Định nghĩa của entropy:

$$H(T) = - \sum_{i=1}^J p_i \log p_i \quad (PL1.18)$$

Trong đó p_i là xác suất dữ liệu rơi vào từng lớp tại nút con xuất phát từ phân nhánh của cây.

$$IG(T, a) = H(T) - H(T|a) \quad (PL1.19)$$

Trong đó $IG(T, a)$ là độ tăng thông tin (Information Gain), $H(T)$ là entropy (parent), $H(T|a)$ là tổng entropy (children).

Lượng thông tin tăng lên chính là tiêu chí để ta quyết định cây phân nhánh dựa trên đặc trưng nào. Tại mỗi điểm phân nhánh, ta nên chọn cách phân nhánh để có thể tạo ra các nút con nhất quán nhất. Để xây dựng cây, ta cần tính độ tăng thông tin của các cách phân nhánh có thể xảy ra. Cách phân nhánh tốt nhất chính là cách tối ưu độ tăng thông tin nhất. Quá trình này được lặp lại nhiều lần với các nút không lá cho đến khi cây hoàn thiện.

Gini index và độ tăng thông tin là hai độ đo thông tin để quyết định phân nhánh. Thông thường, sử dụng một trong hai độ đo này sẽ không tạo ra hiệu suất quá khác biệt. Thường, độ đo Gini sẽ nhanh hơn, nên được sử dụng nhiều hơn. Cây tạo ra bởi 2 độ đo này cũng tương tự nhau. Gini index sẽ cho ta cây với các nhánh cô lập theo phân phối dữ liệu, trong khi đó thì độ tăng thông tin cho ta cây cân bằng hơn.

Áp dụng

- Input: $xa1, ya1, xb1, yb1, xa2, ya2, xb2, yb2, Wpx, Wpy$
- Output: ETA_{Wpa}, ETA_{Wpb}
- **Mô hình:** Decision Tree sử dụng tiêu chí phân tách là Gini index.
- **Kết quả:**
- + Độ đo sai số như sau:

MSE	Tập huấn luyện	Tập kiểm tra
tàu A	0.0	113.7419
tàu B	0.0	5.5270

Ta có bảng độ đo chính xác như sau:

- Tập huấn luyện

	precision	recall	f1-score	support
0	1.00	1.00	1.00	15421

1	1.00	1.00	1.00	49373
accuracy			1.00	64794
macro avg	1.00	1.00	1.00	64794

• Tập kiểm tra

	precision	recall	f1-score	support
0	0.00	0.00	0.00	328
1	0.99	0.89	0.94	55416
accuracy			0.89	55744
macro avg	0.50	0.45	0.47	55744

Kết luận:

+ Trên tập huấn luyện: Độ đo sai số bằng 0 cho cả hai tàu => Mô hình đã học được hoàn toàn dữ liệu của tập huấn luyện. Mô hình có F1-score trung bình (macro avg) của 2 lớp là 100%, có thể là dấu hiệu của overfitting.

+ Trên tập kiểm tra: Độ đo sai số trên tàu A rất lớn, gấp nhiều lần tàu B. Kết quả mô hình có độ chính xác là 89% với dữ liệu thuộc lớp 1, nhưng mô hình không dự đoán đúng bất kỳ mẫu nào thuộc lớp 0 => Mô hình rất kém, có F1-score trung bình (macro avg) của 2 lớp là 47%

+ Có thể thấy, dù mô hình có độ sai số bằng 0 trên tập dữ liệu nhưng không phải là một mô hình tốt khi thử nghiệm trên tập kiểm tra.

PHỤ LỤC 2: CHI TIẾT QUÁ TRÌNH HUẤN LUYỆN MẠNG NƠ-RON NHÂN TẠO VỚI DỮ LIỆU CHẠY TÀU TRÊN VÙNG BIỂN RỘNG

Epoch 1/100

567/567 [=====] - 2s 3ms/step - loss: 0.6920
- accuracy: 0.5567 - val_loss: 0.6433 - val_accuracy: 0.6796

Epoch 2/100

567/567 [=====] - 1s 2ms/step - loss: 0.6317
- accuracy: 0.6521 - val_loss: 0.5960 - val_accuracy: 0.7029

Epoch 3/100

567/567 [=====] - 2s 3ms/step - loss: 0.6048
- accuracy: 0.6835 - val_loss: 0.5964 - val_accuracy: 0.6857

Epoch 4/100

567/567 [=====] - 1s 2ms/step - loss: 0.6009
- accuracy: 0.6895 - val_loss: 0.5906 - val_accuracy: 0.6893

Epoch 5/100

567/567 [=====] - 1s 2ms/step - loss: 0.5985
- accuracy: 0.6935 - val_loss: 0.5875 - val_accuracy: 0.6996

Epoch 6/100

567/567 [=====] - 1s 2ms/step - loss: 0.5907
- accuracy: 0.6976 - val_loss: 0.5824 - val_accuracy: 0.6987

Epoch 7/100

567/567 [=====] - 1s 2ms/step - loss: 0.5843
- accuracy: 0.7026 - val_loss: 0.5687 - val_accuracy: 0.7186

Epoch 8/100

567/567 [=====] - 1s 2ms/step - loss: 0.5706
- accuracy: 0.7077 - val_loss: 0.5513 - val_accuracy: 0.7095

Epoch 9/100

567/567 [=====] - 1s 2ms/step - loss: 0.5516
- accuracy: 0.7217 - val_loss: 0.5287 - val_accuracy: 0.7305

Epoch 10/100

567/567 [=====] - 1s 2ms/step - loss: 0.5363
- accuracy: 0.7341 - val_loss: 0.5117 - val_accuracy: 0.7404
Epoch 11/100
567/567 [=====] - 1s 2ms/step - loss: 0.5187
- accuracy: 0.7434 - val_loss: 0.4924 - val_accuracy: 0.7546
Epoch 12/100
567/567 [=====] - 1s 2ms/step - loss: 0.5019
- accuracy: 0.7539 - val_loss: 0.4824 - val_accuracy: 0.7655
Epoch 13/100
567/567 [=====] - 1s 2ms/step - loss: 0.4892
- accuracy: 0.7603 - val_loss: 0.4622 - val_accuracy: 0.7771
Epoch 14/100
567/567 [=====] - 1s 2ms/step - loss: 0.4789
- accuracy: 0.7703 - val_loss: 0.4524 - val_accuracy: 0.7816
Epoch 15/100
567/567 [=====] - 1s 2ms/step - loss: 0.4701
- accuracy: 0.7757 - val_loss: 0.4413 - val_accuracy: 0.7882
Epoch 16/100
567/567 [=====] - 1s 3ms/step - loss: 0.4625
- accuracy: 0.7773 - val_loss: 0.4300 - val_accuracy: 0.7958
Epoch 17/100
567/567 [=====] - 1s 2ms/step - loss: 0.4519
- accuracy: 0.7860 - val_loss: 0.4306 - val_accuracy: 0.7945
Epoch 18/100
567/567 [=====] - 1s 2ms/step - loss: 0.4371
- accuracy: 0.7934 - val_loss: 0.3963 - val_accuracy: 0.8138
Epoch 19/100
567/567 [=====] - 1s 2ms/step - loss: 0.4240
- accuracy: 0.7988 - val_loss: 0.3905 - val_accuracy: 0.8140
Epoch 20/100
567/567 [=====] - 1s 2ms/step - loss: 0.4196
- accuracy: 0.8020 - val_loss: 0.3823 - val_accuracy: 0.8104

Epoch 21/100

567/567 [=====] - 1s 2ms/step - loss: 0.4022
- accuracy: 0.8118 - val_loss: 0.3730 - val_accuracy: 0.8318

Epoch 22/100

567/567 [=====] - 1s 2ms/step - loss: 0.3891
- accuracy: 0.8199 - val_loss: 0.3453 - val_accuracy: 0.8391

Epoch 23/100

567/567 [=====] - 1s 2ms/step - loss: 0.3770
- accuracy: 0.8264 - val_loss: 0.3248 - val_accuracy: 0.8469

Epoch 24/100

567/567 [=====] - 1s 2ms/step - loss: 0.3677
- accuracy: 0.8295 - val_loss: 0.3193 - val_accuracy: 0.8519

Epoch 25/100

567/567 [=====] - 1s 2ms/step - loss: 0.3606
- accuracy: 0.8348 - val_loss: 0.3062 - val_accuracy: 0.8601

Epoch 26/100

567/567 [=====] - 1s 2ms/step - loss: 0.3481
- accuracy: 0.8419 - val_loss: 0.2913 - val_accuracy: 0.8664

Epoch 27/100

567/567 [=====] - 1s 2ms/step - loss: 0.3467
- accuracy: 0.8393 - val_loss: 0.2956 - val_accuracy: 0.8659

Epoch 28/100

567/567 [=====] - 1s 2ms/step - loss: 0.3389
- accuracy: 0.8433 - val_loss: 0.2804 - val_accuracy: 0.8719

Epoch 29/100

567/567 [=====] - 1s 2ms/step - loss: 0.3343
- accuracy: 0.8456 - val_loss: 0.2930 - val_accuracy: 0.8667

Epoch 30/100

567/567 [=====] - 1s 2ms/step - loss: 0.3326
- accuracy: 0.8476 - val_loss: 0.2876 - val_accuracy: 0.8661

Epoch 31/100

567/567 [=====] - 1s 2ms/step - loss: 0.3250
- accuracy: 0.8484 - val_loss: 0.3073 - val_accuracy: 0.8568
Epoch 32/100
567/567 [=====] - 1s 2ms/step - loss: 0.3209
- accuracy: 0.8521 - val_loss: 0.3023 - val_accuracy: 0.8533
Epoch 33/100
567/567 [=====] - 1s 2ms/step - loss: 0.3171
- accuracy: 0.8519 - val_loss: 0.2861 - val_accuracy: 0.8530
Epoch 34/100
567/567 [=====] - 1s 2ms/step - loss: 0.3168
- accuracy: 0.8522 - val_loss: 0.2660 - val_accuracy: 0.8740
Epoch 35/100
567/567 [=====] - 1s 2ms/step - loss: 0.3152
- accuracy: 0.8538 - val_loss: 0.2774 - val_accuracy: 0.8757
Epoch 36/100
567/567 [=====] - 1s 2ms/step - loss: 0.3146
- accuracy: 0.8516 - val_loss: 0.2695 - val_accuracy: 0.8816
Epoch 37/100
567/567 [=====] - 1s 2ms/step - loss: 0.3051
- accuracy: 0.8571 - val_loss: 0.2808 - val_accuracy: 0.8654
Epoch 38/100
567/567 [=====] - 1s 2ms/step - loss: 0.3093
- accuracy: 0.8542 - val_loss: 0.2565 - val_accuracy: 0.8778
Epoch 39/100
567/567 [=====] - 1s 2ms/step - loss: 0.3017
- accuracy: 0.8576 - val_loss: 0.2924 - val_accuracy: 0.8557
Epoch 40/100
567/567 [=====] - 1s 2ms/step - loss: 0.3045
- accuracy: 0.8584 - val_loss: 0.2789 - val_accuracy: 0.8657
Epoch 41/100
567/567 [=====] - 1s 2ms/step - loss: 0.3013
- accuracy: 0.8582 - val_loss: 0.2542 - val_accuracy: 0.8753

Epoch 42/100

567/567 [=====] - 1s 2ms/step - loss: 0.2980
- accuracy: 0.8634 - val_loss: 0.2488 - val_accuracy: 0.8814

Epoch 43/100

567/567 [=====] - 1s 2ms/step - loss: 0.2998
- accuracy: 0.8573 - val_loss: 0.2707 - val_accuracy: 0.8715

Epoch 44/100

567/567 [=====] - 1s 2ms/step - loss: 0.2953
- accuracy: 0.8600 - val_loss: 0.2647 - val_accuracy: 0.8724

Epoch 45/100

567/567 [=====] - 1s 2ms/step - loss: 0.2914
- accuracy: 0.8616 - val_loss: 0.2460 - val_accuracy: 0.8780

Epoch 46/100

567/567 [=====] - 1s 2ms/step - loss: 0.2939
- accuracy: 0.8619 - val_loss: 0.2542 - val_accuracy: 0.8748

Epoch 47/100

567/567 [=====] - 1s 2ms/step - loss: 0.2894
- accuracy: 0.8635 - val_loss: 0.2682 - val_accuracy: 0.8697

Epoch 48/100

567/567 [=====] - 1s 2ms/step - loss: 0.2871
- accuracy: 0.8614 - val_loss: 0.2580 - val_accuracy: 0.8753

Epoch 49/100

567/567 [=====] - 1s 2ms/step - loss: 0.2893
- accuracy: 0.8628 - val_loss: 0.2412 - val_accuracy: 0.8869

Epoch 50/100

567/567 [=====] - 1s 2ms/step - loss: 0.2858
- accuracy: 0.8635 - val_loss: 0.2825 - val_accuracy: 0.8631

Epoch 51/100

567/567 [=====] - 1s 2ms/step - loss: 0.2882
- accuracy: 0.8665 - val_loss: 0.2438 - val_accuracy: 0.8838

Epoch 52/100

567/567 [=====] - 1s 2ms/step - loss: 0.2828
- accuracy: 0.8658 - val_loss: 0.2432 - val_accuracy: 0.8854
Epoch 53/100
567/567 [=====] - 1s 2ms/step - loss: 0.2823
- accuracy: 0.8662 - val_loss: 0.2712 - val_accuracy: 0.8606
Epoch 54/100
567/567 [=====] - 1s 2ms/step - loss: 0.2828
- accuracy: 0.8672 - val_loss: 0.2309 - val_accuracy: 0.8874
Epoch 55/100
567/567 [=====] - 1s 2ms/step - loss: 0.2817
- accuracy: 0.8649 - val_loss: 0.2344 - val_accuracy: 0.8869
Epoch 56/100
567/567 [=====] - 1s 2ms/step - loss: 0.2777
- accuracy: 0.8680 - val_loss: 0.2454 - val_accuracy: 0.8788
Epoch 57/100
567/567 [=====] - 1s 2ms/step - loss: 0.2777
- accuracy: 0.8672 - val_loss: 0.2327 - val_accuracy: 0.8891
Epoch 58/100
567/567 [=====] - 1s 2ms/step - loss: 0.2791
- accuracy: 0.8666 - val_loss: 0.2493 - val_accuracy: 0.8760
Epoch 59/100
567/567 [=====] - 1s 2ms/step - loss: 0.2792
- accuracy: 0.8666 - val_loss: 0.2386 - val_accuracy: 0.8776
Epoch 60/100
567/567 [=====] - 1s 2ms/step - loss: 0.2757
- accuracy: 0.8687 - val_loss: 0.2434 - val_accuracy: 0.8758
Epoch 61/100
567/567 [=====] - 1s 2ms/step - loss: 0.2872
- accuracy: 0.8624 - val_loss: 0.2373 - val_accuracy: 0.8808
Epoch 62/100
567/567 [=====] - 1s 2ms/step - loss: 0.2758
- accuracy: 0.8660 - val_loss: 0.2613 - val_accuracy: 0.8740

Epoch 63/100

567/567 [=====] - 1s 2ms/step - loss: 0.2782
- accuracy: 0.8671 - val_loss: 0.2335 - val_accuracy: 0.8879

Epoch 64/100

567/567 [=====] - 1s 2ms/step - loss: 0.2721
- accuracy: 0.8708 - val_loss: 0.2340 - val_accuracy: 0.8869

Epoch 65/100

567/567 [=====] - 1s 2ms/step - loss: 0.2724
- accuracy: 0.8729 - val_loss: 0.2312 - val_accuracy: 0.8922

Epoch 66/100

567/567 [=====] - 1s 2ms/step - loss: 0.2736
- accuracy: 0.8684 - val_loss: 0.2255 - val_accuracy: 0.8854

Epoch 67/100

567/567 [=====] - 1s 2ms/step - loss: 0.2711
- accuracy: 0.8667 - val_loss: 0.2487 - val_accuracy: 0.8788

Epoch 68/100

567/567 [=====] - 1s 2ms/step - loss: 0.2700
- accuracy: 0.8689 - val_loss: 0.2251 - val_accuracy: 0.8927

Epoch 69/100

567/567 [=====] - 1s 2ms/step - loss: 0.2706
- accuracy: 0.8719 - val_loss: 0.2369 - val_accuracy: 0.8778

Epoch 70/100

567/567 [=====] - 1s 2ms/step - loss: 0.2692
- accuracy: 0.8735 - val_loss: 0.2259 - val_accuracy: 0.8854

Epoch 71/100

567/567 [=====] - 1s 2ms/step - loss: 0.2704
- accuracy: 0.8707 - val_loss: 0.2514 - val_accuracy: 0.8712

Epoch 72/100

567/567 [=====] - 1s 2ms/step - loss: 0.2660
- accuracy: 0.8731 - val_loss: 0.2371 - val_accuracy: 0.8854

Epoch 73/100

567/567 [=====] - 1s 2ms/step - loss: 0.2667
- accuracy: 0.8724 - val_loss: 0.2188 - val_accuracy: 0.8925
Epoch 74/100
567/567 [=====] - 1s 2ms/step - loss: 0.2670
- accuracy: 0.8720 - val_loss: 0.2532 - val_accuracy: 0.8780
Epoch 75/100
567/567 [=====] - 1s 2ms/step - loss: 0.2618
- accuracy: 0.8767 - val_loss: 0.2392 - val_accuracy: 0.8780
Epoch 76/100
567/567 [=====] - 1s 2ms/step - loss: 0.2707
- accuracy: 0.8690 - val_loss: 0.2321 - val_accuracy: 0.8887
Epoch 77/100
567/567 [=====] - 1s 2ms/step - loss: 0.2627
- accuracy: 0.8744 - val_loss: 0.2597 - val_accuracy: 0.8720
Epoch 78/100
567/567 [=====] - 1s 2ms/step - loss: 0.2703
- accuracy: 0.8701 - val_loss: 0.2299 - val_accuracy: 0.8945
Epoch 79/100
567/567 [=====] - 1s 3ms/step - loss: 0.2620
- accuracy: 0.8738 - val_loss: 0.2197 - val_accuracy: 0.8968
Epoch 80/100
567/567 [=====] - 1s 2ms/step - loss: 0.2683
- accuracy: 0.8730 - val_loss: 0.2379 - val_accuracy: 0.8757