

**UNIVERSIDAD DE CONCEPCIÓN-CHILE
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**Optimización de Operaciones de Fiscalización
Pesquera con predicción de posición mediante
Machine Learning**

por

Julio A. Espinoza Llanos

Profesor Guía:

Rodrigo De la Fuente G., Ph. D.

Concepción, Julio de 2021

Tesis presentada a la

**DIRECCIÓN DE POSTGRADO
UNIVERSIDAD DE CONCEPCIÓN**



Para optar al grado de

MAGÍSTER EN INGENIERÍA INDUSTRIAL

A mi familia por la paciencia y apoyo para hacer lo que me propongo.

RESUMEN

Optimización de Operaciones de Fiscalización Pesquera con predicción de posición mediante Machine Learning

Julio Andrés Espinoza Llanos

Concepción, Junio de 2021

PROFESOR GUÍA: Rodrigo De la Fuente Gallegos, Ph. D.

PROGRAMA: Magíster en Ingeniería Industrial

La fiscalización Oceánica corresponde a una de las tareas de monitoreo y control de la actividad pesquera que la Armada de Chile y otras instituciones realizan con el objeto de combatir la pesca no declarada no reglamentada (INDNR) y con ello salvaguardar la conservación de las especies marinas, contribuyendo a su vez con la sustentabilidad de la actividad pesquera de la cual depende el país. Aunque las actividades de fiscalización son intensas, la posición de las embarcaciones es dinámica, lo que dificulta su localización. Es por esto que es primordial para Instituciones como la Armada maximizar el número de buques fiscalizados al menor costo posible. Este estudio integra dos técnicas, consistentes en un modelo de Machine learning (ML) y otro de optimización. Con el objetivo de determinar la localización más probable de un grupo de embarcaciones pesqueras y a su vez, obtener la ruta más corta que maximice la cantidad de fiscalizaciones, el modelo de ML se compone de una red Neuronal de tipo Long short-term memory (LSTM). Éste tiene la capacidad de entregar regiones probables de localización de acuerdo a sus datos históricos, simulando la incertidumbre existente. Por otro lado, el modelo de optimización está basado en el problema clustered orienteering (COP). Los resultados muestran una desviación aceptable de la predicción de localización de flotas pesqueras de 6 MN, la que puede ser sustituida mediante capacidades de detección con radares marítimos y a su vez, se logró positividad frente a la obtención de un alto número de embarcaciones pesqueras fiscalizadas.

Palabras clave: Machine Learning, LSTM, COP.

ABSTRACT

Optimization of Fisheries Surveillance Operations with position prediction using Machine Learning

Julio Andrés Espinoza Llanos

Concepción, June 2021

THESIS SUPERVISOR: Rodrigo De la Fuente Gallegos, Ph. D.

PROGRAM: Master in Industrial Engineering

The Oceanic surveillance corresponds to one of the monitoring and control tasks of the fishing activity that the Chilean Navy and other institutions carry out in order to combat unreported unregulated fishing (IUU) and thereby safeguard the conservation of marine species, contributing in turn to the sustainability of the fishing activity on which the country depends. Although the surveillance activities are intense, the position of the vessels is dynamic, which makes their location difficult. This is why it is essential for Institutions such as the Navy to maximize the number of fiscal vessels at the lowest possible cost. This study integrates two techniques, consisting of a machine learning (ML) model and an optimization model. In order to determine the most probable location of a group of fishing vessels and, in turn, obtain the shortest route that maximizes the number of inspections, the ML model is made up of a Neural network of the Long short-term memory type (LSTM). It has the ability to deliver probable location regions according to its historical data, simulating the existing uncertainty. On the other hand, the optimization model is based on the clustered orientation problem (COP). The results show an acceptable deviation of the prediction of location of fishing fleets of 6 MN, which can be replaced by detection capabilities with maritime radars and, in turn, positiveness was achieved compared to obtaining a high number of fishing vessels audited.

Key Words: Machine Learning, LSTM, COP.

Índice general

Índice de figuras	iii
Índice de tablas	iv
1 Introducción	1
1.1 Introducción	1
1.2 Descripción del problema	2
1.3 Objetivos	3
1.3.1 Objetivo general:	3
1.3.2 Objetivos específicos:	3
1.4 Contribución	4
2 Revisión de la literatura	5
2.1 Operaciones de fiscalización pesquera	5
2.2 Trabajos previos	7
2.2.1 Machine learning para fiscalización pesquera	7
2.2.2 Modelos TSP para fiscalización pesquera	8
2.3 Resumen de literatura revisada	10
3 Metodología	12
3.1 Modelo LSTM para predicción de posición de embarcaciones	12
3.1.1 Recolección y preparación de datos	13
3.1.2 Entrenamiento	15
3.1.3 Predicción	17
3.2 Modelo COP fiscalización pesquera	18
3.3 Integración de modelos	21

4	Caso de estudio	23
4.1	Recolección y preparación de datos	23
4.1.1	Implementación del modelo LSTM	26
4.1.2	Generación de instancias	26
4.1.3	Implementación del modelo COP	29
5	Resultados y discusión	30
5.1	Resultados computacionales y discusión	30
6	Conclusiones	36
7	Referencias	41

Índice de figuras

2.1	Generalizaciones del TSP.	9
2.2	Solución factible de un SGTSP tour.	10
3.1	Cell state LSTM.	13
3.2	Modelo LSTM.	16
3.3	Integración de modelos.	22
4.1	Visualización de los datos filtrados.	25
4.2	Arco tangente coordenadas.	26
4.3	Visualización de tracks generados.	27
4.4	Método sliding window con 1 timeStep.	28
4.5	Predicción de posición de embarcación pesquera.	29
5.1	Rutas embarcaciones fiscalizadoras.	35

Índice de tablas

1.1	Distribución de zonas navales.	3
2.1	Resumen revisión bibliográfica	11
4.1	Comparación de datos pre y post procesado.	24
4.2	Generación de instancias para el modelo COP.	28
5.1	Resultados modelo COP.	31
5.2	Solución de rutas buques fiscalizadores.	32
5.3	Distancia por ruta de buque fiscalizador.	33

Capítulo 1

Introducción

En este capítulo se presenta la contextualización de las actividades de fiscalización pesquera y la problemática detectada. Asimismo, se exponen los objetivos del estudio y su contribución.

1.1 Introducción

La fiscalización Oceánica de embarcaciones corresponde a las tareas de monitorización y control de la actividad pesquera por parte de la Armada de Chile y otras instituciones, dispuestas en la ley general de pesca y acuicultura (Chile, 1991). Ésta abarca tanto aguas nacionales como en alta mar adyacente a la zona exclusiva y tiene la misión de combatir la pesca no declarada no reglamentada (INDNR), tal como lo describe la Dirección General del Territorio Marítimo y de Marina Mercante en su cuenta pública (DGTM, 2019). En relación a esto, la INDNR es una actividad ilegítima que amenaza directamente a la conservación de las poblaciones de peces, generando efectos adversos sobre los ecosistemas y la biodiversidad marina, así como a los medios de vida y actividades económicas de pescadores que viven legítimamente de las pesquerías, generando efectos severos a la sustentabilidad de toda la actividad pesquera (DGTM, 2019). Dentro de este marco, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) señala que la pesca INDNR supone una pérdida de 26 millones de toneladas anuales, equivalentes a alrededor de una quinta parte de la captura mundial, socavando los esfuerzos para garantizar una pesca sostenible en el tiempo y de conservación de ecosistemas marinos y biodiversidad (DGTM, 2019). Es por ello la importancia de realizar este tipo de tareas fiscalizadoras.

En cuanto al espacio marítimo nacional, cubre un área 4 veces mayor que la superficie continental chilena. Esto pone de manifiesto el difícil trabajo de fiscalización de la pesca INDNR por parte de las instituciones estatales, considerando los elevados costos asociados para el cumplimiento de dichas tareas y los escasos recursos disponibles para su realización. En

este sentido, el auge de los sistemas automáticos de posicionamiento de buques (AIS) para las tareas de monitoreo han permitido contar con una fuente de datos cooperativa de movimientos de buques. Estos datos son generados en tiempo real, recepcionando señales de velocidad, rumbo, latitud y longitud, entre otras informaciones estáticas de las embarcaciones. Además, pueden ser descargadas desde distintas fuentes online. No obstante, el posicionamiento de las embarcaciones es dinámico, por lo que la planificación de dichas tareas debe ser proyectada dinámicamente a una posición futura de las embarcaciones, trayendo consigo una dificultad adicional. Así, surge la necesidad de contar con un sistema que optimice la toma de decisiones en cuanto a la ruta de las unidades fiscalizadoras y efectividad en la búsqueda de embarcaciones.

1.2 Descripción del problema

El objetivo principal en las operaciones de fiscalización pesquera está relacionado con el fortalecimiento de la cooperación internacional y local con el fin de gestionar la pesca sostenible y a su vez mantener el estatus de los ecosistemas y biodiversidad marina. Así, es primordial para la Armada Nacional maximizar la cantidad de buques fiscalizados con el menor costo posible asociado a lo largo de todo el territorio marítimo nacional y sujeto a escasos recursos de fiscalización. En términos detallados, el área marítima de Chile está comprendida hasta las 200 millas marinas, representando aproximadamente $3.150.739 \text{ km}^2$, es decir, 4 veces más que la superficie terrestre chilena, la que se subdivide en 3 zonas: el Mar Territorial (hasta 12 millas desde la línea base de costa), Zona Contigua (hasta 24 millas) y la Zona Económica Exclusiva (hasta las 200 millas). Asimismo, la Armada de Chile subdivide el territorio nacional en cinco zonas navales, tal como se muestra en la Tabla 1.1. En cada una de estas zonas existen recursos fiscalizadores marítimos como aeromarítimos, conformando un conjunto polivalente de medios navales con capacidades y autonomías distintas.

Ahora, en base a los registros de la cuenta pública de la DGTM (2019), en el primer operativo efectuado entre abril y mayo del 2019 por una fragata de la Escuadra Nacional, se recorrieron 48.799 MN^2 , lo que solo en combustible, en promedio, equivalen a cerca de 4.500 m^3 o USD 3M, logrando pesquisar solo 1 pesquero nacional, 1 pesquero extranjero mayor y 1 pesquero extranjero menor. En la práctica, esto generaliza a la mayoría de los operativos. Por otro lado, la posición de las flotas pesqueras generalmente es dinámica y se mueven en grupos, por lo que se requiere predecir su posición futura para direccionar las unidades fiscalizadoras a zonas con mayor probabilidad de localización de flotas pesqueras. Entonces, para resolver la problemática, se modela un problema de fiscalización pesquera para apoyar las labores de

Tabla 1.1: Distribución de zonas navales.

Zona Naval	Territorio
I	Caldera
	Coquimbo
	Hanga Roa
	Valparaíso
	San Antonio
II	Valdivia
	Talcahuano
III	Punta arenas
	Puerto Williams
	Antártica Chilena
IV	Arica
	Iquique
	Antofagasta
V	Puerto Montt
	Castro
	Aysén

Fuente: (DGTM, 2019)

fiscalización que realiza la Armada de Chile, con el objetivo de contribuir a la disminución de costos de operación y maximización de fiscalizaciones. Se espera que esto traiga consigo, en el largo plazo, una disminución del impacto en los ecosistemas marinos y biodiversidad garantizando una pesca sostenible en el tiempo.

1.3 Objetivos

A continuación, se presentan los objetivos generales y específicos del estudio.

1.3.1 Objetivo general:

Modelar problema de fiscalización pesquera para apoyar labores de fiscalización que realiza la Armada de Chile.

1.3.2 Objetivos específicos:

- Buscar y preprocesar datasets de posición de buques pesqueros.
- Estudiar uso de RNN para predicción de posición futura.

- Implementar computacionalmente predicción de posición de buques pesqueros basado en RNN.
- Generar las instancias de prueba para el modelo.
- Estudiar problemas TSP con sincronización en clusters.
- Proponer modelo basado en TSP con datos de posición fija.
- Implementar computacionalmente el modelo basado en TSP.
- Analizar resultados obtenidos con instancias reales y de prueba de la literatura.

1.4 Contribución

Las principales contribuciones de este trabajo son:

- Presentar un primer acercamiento a la optimización de labores de fiscalización pesquera nacional, complementando técnicas de machine learning y métodos de optimización.
- Disminución de costos para la ejecución de las operaciones de fiscalización pesquera.
- Disminuir la contaminación producida por embarcaciones fiscalizadoras como un efecto secundario.
- Efectuar transferencia tecnológica a Instituciones como la Armada de Chile en cuanto al asesoramiento de las operaciones de fiscalización pesquera tanto Antártico como continental.

Capítulo 2

Revisión de la literatura

En este capítulo se presentan diversos estudios relacionados con la fiscalización pesquera, donde el área de machine learning (ML) fue incluido como método complementario a los modelos de optimización de rutas para la resolución del problema planteado.

2.1 Operaciones de fiscalización pesquera

Previo a revisar la conformación de las operaciones de fiscalización pesquera, se debe precisar sobre los sistemas AIS, dado que en general, los estudios de ML relacionados con fiscalización pesquera trabajan con este tipo de datos. En efecto, el estudio de Corvalán and Abarca (2019) describe los Automatic Identification System (AIS) como dispositivos que permiten la transmisión automática de la posición geográfica de las embarcaciones, mientras está efectuando sus labores de pesca. De esta manera pueden ser monitoreados y controlados en caso de cometer infracciones a la ley de pesca. Su importancia radica en que estos sistemas facilitan la tarea de monitorear embarcaciones en amplios espacios de territorio marítimo, incluso mundialmente. Sin embargo, no es una técnica infalible por lo que debe ser complementada con tareas de fiscalización por parte de las instituciones de cada país. El sistema, por su parte, se compone de transmisores, antenas receptoras, estaciones de procesamiento de información y satélites propiamente tal. El elemento más importante del sistema es el dispositivo de posicionamiento automático (GPS), que permite obtener la latitud, longitud, rumbo y velocidad de las embarcaciones siendo transferida a un transceptor instalado en la embarcación que transmite la señal a un satélite. Una vez captada la señal en el satélite, se envía a una estación receptora en tierra, la cual es redirigida de forma encriptada a los servidores de la Dirección General del Territorio Marítimo y Marina Mercante (DIRECTEMAR), quienes decodifican y reenvían al CMC (Centro de control y monitoreo de SERNAPESCA) para ser analizada. Asimismo, la emisión de localización se realiza mediante reportes automáticos en lapsos de tiempo constantes definidos por SERNAPESCA y son aplicados a naves artesanales,

industriales, transporte y flota extranjera. Lo anterior, sin perjuicio de los requisitos generales existentes en la ley de Pesca Nacional como lo señala Corvalán and Abarca (2019).

En relación a la actividad pesquera y su fiscalización, Dunn et al. (2018) señala que los esfuerzos en la actividad pesquera actualmente tienen una tendencia lineal al alza. Esto pone de manifiesto la importancia y los posibles problemas que conlleva el control y monitoreo de la captura. Teniendo en cuenta esto, se han generado una serie de instrumentos o acuerdos internacionales para el monitoreo y control de la actividad pesquera a nivel mundial. Según la DGTM (2019), parte del marco regulatorio al cual está suscrito nuestro país corresponde a la ratificación del texto de la Convención de la Organización regional de Pesca del Pacífico Sur en el año 2012 (Nacional, 2012). Luego, se adhiere al acuerdo de Nueva York en el año 2015 y recientemente ingresa como estado cooperante a la Comisión Interamericana del Atún Tropical (CIAT). Esto demanda, entre otras cosas, el monitoreo y control de las actividades de pesca desarrolladas en nuestras costas. A pesar de lo anteriormente expuesto, al no cumplir satisfactoriamente dichos acuerdos se da espacio para la realización de actividades como la pesca ilegal no declarada no reglamentada (INDNR), lo cual socava los esfuerzos internacionales, regionales y mundiales por gestionar la pesca sostenible y a su vez mantener el status de los ecosistemas y biodiversidad marina. Para esto, tal como lo señala la DGTM (2019), la conducción de las operaciones de fiscalización pesquera en la Armada de Chile son planificadas desde el Comando de Operaciones Navales, asesorado por la Dirección de Intereses Marítimos y Medio Ambiente Acuático (DIRINMAR) y la Dirección de Seguridad y Operaciones Marítimas (DIRSOMAR), dependientes de la Dirección General del Territorio Marítimo y Marina Mercante (DIRECTEMAR). Mediante información de medios satelitales y electrónicos (AIS), informes meteorológicos y portales de bases de datos web marítimos adscritos por membresía o pagada, se contrastan con información de aviones de aeroexploración marítima, los que centralizados conforman un panorama de superficie marítimo que mediante el análisis interpretativo, se convierte en un escenario marítimo reconocido. Asimismo, las exploraciones aeromarítimas (EAM), actividades de vuelo sobre el espacio marítimo, permiten llegar a áreas de interés y son llevadas a cabo por cada una de las Zonas Navales mediante una planificación previa, donde se define el área a fiscalizar, cuándo realizarlo y por cuánto tiempo. Lo anterior, se respalda por antecedentes oceanográficos, comportamiento de los recursos marinos o de pesca, de la flota e historia de la pesquería. Sin embargo, también existen actividades de EAM que se llevan a cabo sin una planificación previa y que dependen principalmente de denuncias efectuadas por el sector pesquero.

2.2 Trabajos previos

La problemática presentada ha sido abordada por Krisnafi et al. (2017) relacionado con la asignación de buques fiscalizadores en la República de Indonesia que puede considerarse como un primer acercamiento a la optimización de operaciones de fiscalización pesquera. Sin embargo, existen estudios de optimización de rutas con nodos clusterizados que se asemejan al comportamiento de los operativos de fiscalización como generalizaciones del modelo Travelling salesman problem (TSP).

2.2.1 Machine learning para fiscalización pesquera

Existen algunos estudios relacionados con la predicción del comportamiento de buques como en Chuaysi and Kiattisin (2020), que se introducen en la temática del combate de la pesca ilegal. Establecen el uso de datos de posición de buques pesqueros y utilizando técnicas de clasificación mediante machine learning de embarcaciones de “Pesca” y “No pesca”. Los autores obtienen resultados robustos con un 97% de precisión. Además, introducen el concepto de uso de datos AIS o VMS (Vessel Monitoring Systems), como una tecnología que permite la obtención de posicionamiento global de las embarcaciones. Suman también el cálculo de la velocidad y los ángulos de giro que permiten describir y analizar patrones de movimiento. Si bien, el estudio de Chuaysi and Kiattisin (2020) no incorpora predicción de posicionamiento ni mejoras en la optimización de rutas, en efecto detalla el tipo de datos disponibles de localización de embarcaciones de pesca que pueden ser utilizados, su composición y preprocesamiento. Relacionado con el preprocesamiento de los datos, es importante señalar que los datos abiertos AIS disponibles para ser utilizados en este tipo de problemas es parcializado. Es decir, no son trayectorias completas, incluso algunos puntos pueden representar solo situaciones de fondeo y deriva, lo cual puede influir en la predicción de posición, sea cual sea el modelo. Es por esta razón que los datos deben ser debidamente preparados, siendo un ejemplo de ello el trabajo de Duca et al. (2017) en donde se intenta obtener o reconstruir trayectorias a partir de datos AIS mediante el algoritmo K-Nearest Neighbor (KNN) como clasificador. Éste agrupa las coordenadas por track o camino de cada embarcación, llegando a obtener un accuracy de 93%. En relación a la predicción de trayectorias, de Souza et al. (2016) y Ayhan and Samet (2016) introducen los modelos de Cadena de Markov Ocultas (HMM). El primero, para mejorar la detección de patrones de pesca de embarcaciones con el objetivo de controlar la actividad humana y su impacto en los recursos marinos. El segundo, para la predicción de tracks de aeronaves sobre una dimensión 4-D, no obstante, este método tiene mejor rendimiento con datos de corto plazo. Igualmente, en el trabajo de González-Andrés et al. (2020) se analizan 8 años de

registros de fiscalización pesquera con el objeto de determinar la distribución espacio-temporal de la pesca ilegal, áreas propensas a ésta y zonas más rentables para la actividad pesquera. Introducen a su vez, un Modelo Aditivo Generalizado (GAM) en el que no solo se predice un punto de coordenadas de un buque o una flota, sino más bien regiones mediante mapas de densidad. Esto permite obtener un rango de decisión más amplio de posicionamiento. El territorio marítimo nacional posee una gran dimensión, por lo que se prevé que la necesidad de contar con cierto grado de incertidumbre cobra una importancia relevante que se debe tomar en consideración. Bajo la misma óptica de las predicciones de posición de vehículos, Tampakis et al. (2020) afirman que las redes neuronales de tipo Long-Short Term Memory (LSTM), redes con la capacidad de procesar datos históricos, emergen como una técnica efectiva para problemas con datos de series temporales a largo plazo. Esto es totalmente adecuado a los datos AIS históricos que se utilizan en el monitoreo de la actividad pesquera mundial. Lo anterior, permite predecir una latitud y longitud futura con distintas ventanas de tiempo.

Teniendo en cuenta lo anteriormente señalado, en el trabajo de Tang et al. (2019) se introduce un modelo de predicción de trayectorias de embarcaciones que utiliza redes LSTM para obtener inferencia de posiciones como una distribución de densidad, similar a González-Andrés et al. (2020). Es aplicado a la navegación autónoma de embarcaciones para evitar colisiones, permitiendo que este modelo prediga parámetros de una distribución normal bivariada asumida de forma premeditada, obteniendo así regiones de densidad. En este sentido, las redes LSTM aparecen como una solución a la falta de contexto en las redes neuronales recurrentes, es decir, mantienen una dependencia entre los datos históricos para la inferencia a largo plazo. Este trabajo tiene una relevancia fundamental para la problemática de aplicación, debido a que se busca incertidumbre y predicciones que establezcan cierto contexto. Esto solo se puede realizar con redes LSTM representando uno de los componentes necesarios para la problemática abordada en el presente estudio.

2.2.2 Modelos TSP para fiscalización pesquera

Cabe destacar que el único trabajo detectado relacionado con la vigilancia pesquera a nivel internacional fue el de Krisnafi et al. (2017). Quienes mediante algoritmos genéticos, logran resolver el problema de asignación de buques de vigilancia pesquera. Con este trabajo, ellos contribuyen con una herramienta para el soporte de decisiones de altos cargos de las fuerzas Armadas de Indonesia. Sin embargo, una técnica no relacionada con la fiscalización pesquera es el clásico modelo “Travelling Salesman Problem” (TSP) el que tiene características inequívocas relacionadas con el problema planteado. Ésta intenta encontrar la ruta mas corta (o menos costosa) que un agente viajero puede tomar al visitar una cierta cantidad de nodos. Donde cada

nodo debe ser visitada una sola vez y el agente debe retornar al mismo punto de partida. Sin embargo, el comportamiento de las flotas pesqueras, en general, se manifiesta en grupos o flotas y por ello se requiere obtener un modelo de fiscalización que actúe de forma clusterizada. Frente a estas características, Zhang et al. (2020) plantea un problema de orientación con cubrimiento clusterizado (CCOP). Este problema es una generalización del modelo de orientación, orientado al monitoreo de la calidad de reservas de agua. Mediante la recolección de muestras con vehículos autónomos, consideran la cobertura de áreas solapadas. Si bien el problema es aplicable a la problemática del estudio en cuestión, se trata con datos estáticos, requiriendo alguna técnica de predicción de posición futura para ingresar datos a éste. Bajo esta misma perspectiva, Angelelli et al. (2014) señala que algunas generalizaciones del modelo TSP son los problemas de ruteo con ganancias o Profits (TSPP, Traveling Salesman Problems with Profits) asociadas a la visita de cada nodo y los Problemas de Orientación (OP) o también llamados Selective Traveling Salesman Problem, Prize Collecting TSP (PCTSP) y Profitable Tour Problem (PTP). En este tipo de problemas la función objetivo es encontrar la ruta de vehículo que maximiza la ganancia total recolectada de tal forma que su duración no exceda un umbral dado. Además, los mismos autores señalan que este tipo de problema es el que más ha recibido la atención en la literatura. Sin embargo, los estudios mas recientes se han enfocado en variantes o generalizaciones del OP, especialmente en múltiples vehículos llamados Team Orienteering Problem (TOP) y Clustered Orienteering Problem (COP), donde los clientes están agrupados en clusters y deben ser servidos. Un diagrama resumen de las generalizaciones del TSP se muestra en la Figura 2.1.

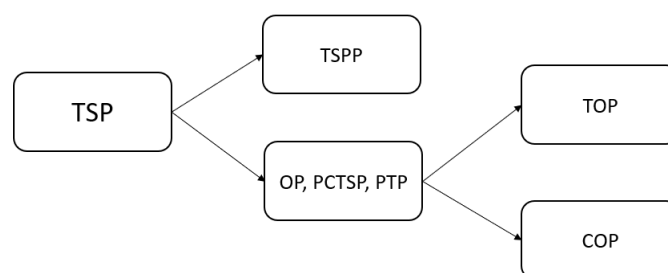


Figure 2.1: Generalizaciones del TSP.

Fuente: Elaboración propia

Por su parte, Yahiaoui et al. (2019) y Angelelli et al. (2014) abordan otra perspectiva de la generalización del problema de orientación. Ellos plantean un algoritmo que maximiza el profit recolectado con respecto a un límite de tiempo de viaje y utiliza una serie de clusters que deben ser visitados al mínimo costo. No obstante, el profit obtenido solo se adquiere al visitar

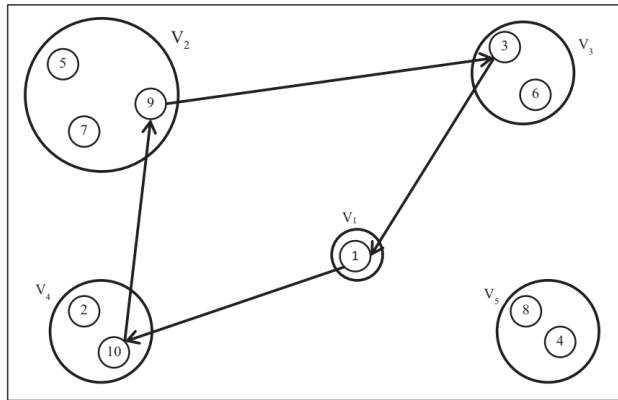


Figure 2.2: Solución factible de un SGTSP tour.
(Derya et al., 2020)

la totalidad de los nodos de un cluster, lo cual no siempre se presenta en la realidad al efectuar fiscalizaciones pesqueras. Esto porque se debe contar con flexibilidad en las operaciones para efectuar cambios de decisión y dirigirse a otros clusters. Ambos trabajos obtienen soluciones robustas mediante métodos exactos y Tabu Search, respectivamente. Finalmente, en el trabajo de Derya et al. (2020), se aplica la generalización del problema de orientación clusterizado (SGTSP). Sin embargo, bajo una nueva condición, en la que no requiere que todos los clusters sean visitados para obtener el profit, ni tampoco se necesite recorrer la totalidad de nodos de un cluster (ver Figura 2.2), manifestando ser adaptable al problema que se plantea en este trabajo.

2.3 Resumen de literatura revisada

A continuación se muestra la Tabla 2.1 correspondiente al resumen de la literatura revisada en esta investigación. Cabe destacar que la revisión de la literatura se efectúa en los siguientes ámbitos: optimización de rutas (OPT), predicción de posición de embarcaciones con técnicas de machine learning (ML) y fiscalización pesquera (FP) propiamente tal.

La literatura relacionada con la optimización de fiscalizaciones pesqueras es escasa y por el mismo motivo, la brecha de conocimiento detectada es amplia. En primer lugar, no existe una vinculación de problemas de predicción de posición futura de vehículos y optimización de rutas en un mismo modelo, considerando el comportamiento dinámico de las flotas pesqueras y el alto costo de las operaciones de fiscalización. En segundo lugar, los modelos de generalización del TSP como el COP pueden ser adaptados a este tipo de problemática. Consiguen que las rutas lleguen y recorran de manera particular a grupos de vehículos o clusters mediante restricciones

Tabla 2.1: Resumen revisión bibliográfica

Autor	Año	Problema Abordado	Tipo de Problema			Técnica
			ML	OPT	FP	
Chuaysi and Kiattisin	2020	Trazabilidad de alimentos marinos.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	KNN, NN.
Tampakis et al.	2020	Predicción de pesca en tiempo real.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	LSTM
Tang et al.	2019	Navegación autónoma, evitar colisiones.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	LSTM
Salman et al.	2018	Predicción de Clima con redes LSTM.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	LSTM
Zhong et al.	2019	Restauración de trayectorias de buques.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	RNN
Duca et al.	2017	Predicción de trayectorias de buques.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	KNN
Varsamopoulos et al.	2018	Diseño de Redes neuronales para corrección de errores.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	LSTM
González-Andrés et al.	2020	Distribución espacio temporal de pesca ilegal.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	GAM
Ayhan and Samet	2016	Predicción trayectorias de Aviones.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	HMM
de Souza et al.	2016	Mejora de detección patrones de pesca.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	HMM
Zhang et al.	2019	Monitoreo de calidad de agua.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	CCOP
Angelelli et al.	2014	Generalización COP.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Tabu search, branch and cut.
Yahiaoui et al.	2019	Generalización COP.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Cutting planes.
Derya et al.	2020	Generalización COP.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Met. Exactos.
Krisnafi et al.	2017	Asignación de buques de vigilancia pesquera.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Algoritmos Genéticos.
Chisman	1975	TSP Clusterizado.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Met. exactos.
Kitjacharoenchai et al.	2019	mTSP con drones.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Combinación de heurísticas.
Corvalán and Abarca	2019	Sist. de monitoreo Fiscalizaciones.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	-
DGTM	2019	Cuenta Pública 2019 Fiscalización Pesquera.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	-
Dunn et al.	2018	Monitoreo satelital de buques en alta mar.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	-

Fuente: Elaboración propia

de autonomía dentro de éstos y obtienen con esto una disminución de costos de viaje. En tercer lugar, falta de introducción de un factor de incertidumbre en la predicción de la posición futura de embarcaciones pesqueras. Esto es una importante consideración a la hora de conducir modelos de optimización de rutas, ya que factores externos como la corriente, climáticos, entre otros, impactan en la predicción exacta de la posición.

En este trabajo se abordan estas brechas de conocimiento con el fin de modelar el problema de fiscalización pesquera para resolver el problema de optimización de fiscalizaciones pesqueras para la Armada de Chile.

Capítulo 3

Metodología

En este capítulo se explica la metodología de implementación del modelo de optimización para el problema de fiscalización pesquera. Para abordar este problema se describe un modelo de orientación clusterizado (COP, por sus siglas en inglés) con el objeto de establecer una herramienta útil para la planificación de operaciones de fiscalización pesquera de la Armada de Chile. También, se describe un modelo de machine learning basado en redes neuronales recurrentes de tipo LSTM para la predicción de la posición futura de cada embarcación, lo que sirve para la generación de instancias para el modelo COP.

3.1 Modelo LSTM para predicción de posición de embarcaciones

Para predecir la posición futura de las embarcaciones a fiscalizar, la red se basa en un modelo LSTM, ya que como menciona Salman et al. (2018), permiten almacenar información a largo plazo, compuesta por una memoria llamada cell state. Para así almacenar u olvidar información relevante desde pasos de tiempo anteriores y por mecanismos de control a través de compuertas que permiten la operación anterior, tal como se ve en la Figura 3.1. Los datos de localización de las embarcaciones pesqueras que se utilizan con este modelo, se obtienen desde bases de datos AIS, las cuales incluyen coordenadas y fechas de posicionamiento. Asimismo, como una forma de complementar las variables de entrenamiento del modelo, se calcula el rumbo respectivo de cada par de coordenadas a partir de los datos señalados anteriormente, logrando también con ello entregar mayor información de contexto a la red.

Similar al procedimiento que realiza Tang et al. (2019), este trabajo puede ser formulado de la siguiente manera: dado un conjunto de estados $S = \{(Lat_j^t, Long_j^t, Rumbo_j^t)\}$ para las embarcaciones pesqueras $j = 1, 2, \dots, N$, en los pasos de tiempo $t = k - n, \dots, k$, se predice una posición futura $\{(\hat{x}_j^{k+m}, \hat{y}_j^{k+m})\}$, donde j corresponde a cada uno de las embarcaciones objetivo,

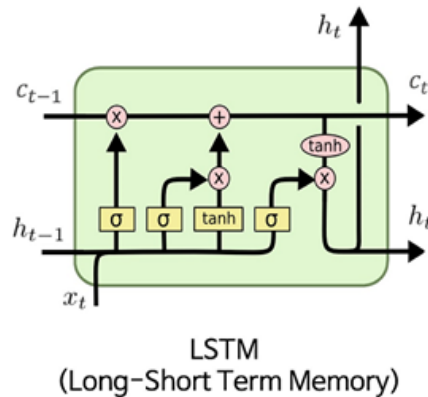


Figure 3.1: Cell state LSTM.
(Varsamopoulos et al., 2018)

N es el número total de embarcaciones objetivo, n es el primer momento de observación, k es el paso actual y m es el momento futuro a predecir. La secuencia de entrada puede ser observada como un problema de series de tiempo, donde dicha secuencia está representada por el estado de los primeros n momentos observados, y la secuencia de salida representa la localización de los buques a predecir mediante la construcción de trayectorias mediante el modelo LSTM. Esta fase del trabajo se realiza en los siguientes pasos: recolección y preparación de los datos, entrenamiento y finalmente predicción de la posición de embarcaciones pesqueras. Posteriormente, el entrenamiento utiliza los datos obtenidos en la primera etapa para ajustar las matrices de pesos, los que finalmente desencadenan en la etapa final, donde se ejecuta la predicción de posición.

3.1.1 Recolección y preparación de datos

Primero se debe visualizar los datos obtenidos, con el objetivo de tener una imagen clara de los elementos recolectados. Para esta etapa se puede recurrir a herramientas como la API de Google Maps, QGIS o simplemente la librería matplotlib de Python. Por otro lado, se deben eliminar los errores existentes en los datos y estructurarlos de tal manera que puedan ser utilizados posteriormente en el modelo. Como generalmente los datos están agrupados de forma diaria y los tramos de tiempo son bastante grandes para la generación de trayectorias, se debe realizar un procedimiento de interpolación. El procedimiento de preparación de datos se muestra en el pseudocódigo del Algoritmo 1.

Durante el proceso de preparación de los datos se considera sólo aquellos pertenecientes a las 200 millas chilenas, correspondientes a aproximadamente entre la latitud 18.7° S - 66.19°

Algorithm 1 Preparación de datos

Input: Base de datos AIS de Global Fishing Watch (GFW)

Output: Trayectorias de embarcaciones pesqueras

- 1: Concatenar archivos *csv* con datos AIS diarios.
 - 2: Leer archivo *csv* concatenado.
 - 3: Filtrar datos solo a embarcaciones dentro de las 200 millas Chilenas.
 - 4: Indexar datos por día.
 - 5: Calcular distancias y rumbos.
 - 6: Agrupar por MMSI únicos.
 - 7: **for** MMSI **do**
 - 8: **if** Tracks no son estacionarios **then**
 - 9: Interpolar Tracks por minuto.
 - 10: **else**
 - 11: **break**
 - 12: **for** MMSI **do**
 - 13: Guardar Tracks
 - 14: **return** Tracks agrupados por MMSI
-

S y longitud 67.45° W - 79.19° W. Para el cálculo de distancias en el mar, no es posible utilizar el método Euclidiano, ya que por razones obvias, las embarcaciones no se pueden mover hacia el interior del mar, al ser la tierra esférica. Es por esto que el cálculo de distancias se realiza mediante el método geodésico (Bessel et al., 2010). Éste considera la curvatura de la tierra para su medición exacta en la superficie, de otra manera, los errores de medición son elevados.

Los datos de posición de las embarcaciones como series de tiempo, son almacenados, por lo general, sin un orden específico. Entonces, es necesario agruparlos por un identificador de embarcación, denominado en los sistemas AIS como MMSI (Maritime Mobile Service Identity) o número único de identificación. Permitiendo agrupar trayectorias completas de cada embarcación, para posteriormente discriminar si éstos se encuentran en estado estacionario o en navegación, tomando como referencia un radio de 0.3 millas náuticas (1.6 km) de distancia. Se considera que un buque detenido no debería moverse mas allá de señalada distancia debido al peligro de varada que reviste sobretodo si se encuentra en cercanías a costa. En otras palabras, si la embarcación registra un movimiento mayor a las 0.3 millas náuticas de distancia, es considerado como una posición que conforma parte de un track o ruta de viaje, de otro modo se considera que es una posición fija, no perteneciente a un track. Finalmente, se realiza una interpolación lineal para la obtención de los datos en minutos para cada día, con el objetivo de visualizar la trayectoria con una resolución más detallada. Es importante señalar que el procedimiento descrito provoca una disminución del número de registros recopilados. Por lo tanto, se debe asumir esta pérdida de datos históricos y asegurar un número aceptable de datos

para la etapa de entrenamiento que permita obtener predicciones con mínimo error.

3.1.2 Entrenamiento

La etapa de entrenamiento consiste en entrenar la red LSTM con los datos ya preparados, ajustando los pesos y comparando el error entre la salida esperada y la predicha. La red de capas ocultas recurrentes es alimentada por una entrada de vectores de estado $S = (s_1, \dots, s_t)$, obteniendo en la salida los estados ocultos $h^M = (h_1^M, \dots, h_t^M)$, mediante las Ecuaciones (3.1)-(3.5).

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (3.1)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (3.2)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3.3)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (3.4)$$

$$h_t = o_t \tanh(c_t) \quad (3.5)$$

Donde:

- W_{xi} : Matriz de pesos entre entrada y capa oculta.
- W_{xf} : Matriz de pesos entre la entrada y la *forget gate*.
- W_{ci} : Matriz de pesos entre *cell state* y entrada.
- i_t, f_t, c_t, o_t : Compuertas LSTM.

En la Figura 3.2 se puede observar la configuración gráfica planteada para este estudio, tal como lo hizo (Tang et al., 2019). Posterior al cálculo de las capas ocultas, se obtienen los valores de salida con las Ecuaciones (3.6) y (3.7).

$$o_{t+1} = \sum_{l=1}^M W_{h^l o} h_t^l + b_o \quad (3.6)$$

$$\hat{p}_{t+1} = Y(o_{t+1}) \quad (3.7)$$

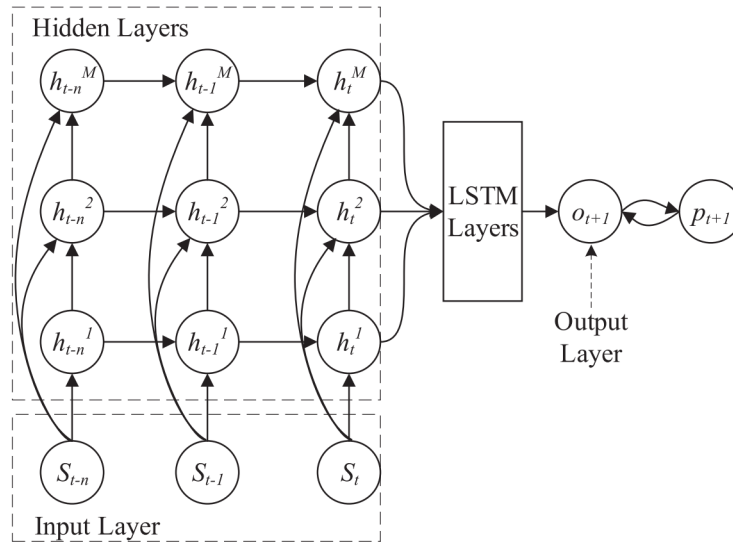


Figure 3.2: Modelo LSTM.
(Tang et al., 2019)

Donde:

- o_{t+1} : Salida antes de ser activada.
- $W_{h^l o}$: Matriz de pesos entre capas ocultas.
- b_o : Vector bias de la capa de salida.
- M : Número de capas ocultas.
- \hat{p}_{t+1} : Predicción.
- $Y(\cdot)$: Función de capa de salida.

El método de entrenamiento del modelo a emplear es de aprendizaje supervisado, en el cual se ingresan datos de entrada etiquetados con su salida respectiva, para que el modelo ajuste sus parámetros o pesos.

En cuanto a la predicción \hat{p}_{t+1} , se asume deliberadamente una distribución bivariada Gaussiana y puede ser expresada como $(\mu_{t+1}, \sigma_{t+1}, \rho_{t+1})$, parametrizadas por:

- $\mu_{t+1} = (\mu_x^{t+1}, \mu_y^{t+1})$: Media.

- $\sigma_{t+1} = (\sigma_x^{t+1}, \sigma_y^{t+1})$: Desviación estándar.
- ρ : coeficiente de correlación.

Los parámetros de distribución pueden ser finalmente obtenidos con la salida de la red:

$$(\mu_{t+1}, \sigma_{t+1}, \rho_{t+1}) = W_p \hat{p}_{t+1} \quad (3.8)$$

Es así que una vez que se obtiene la salida, los pesos de la red se ajustan mediante la minimización de la Ecuación (3.9), correspondiente a la función de pérdida con distribución Gaussiana diseñada por Tang et al. (2019) y optimizada mediante gradiente descendente.

$$L(S) = - \sum_{t=1}^m \log \Pr(p_{t+1} | \hat{p}_{t+1}) + \frac{\lambda}{2n} \sum_w w^2 \quad (3.9)$$

$$\Pr(p_{t+1} | \hat{p}_{t+1}) = \mathbf{N}(p_{t+1} | \mu_{t+1}, \sigma_{t+1}, \rho_{t+1}) \quad (3.10)$$

$$\mathbf{N}(p_{t+1} | \mu_{t+1}, \sigma_{t+1}, \rho_{t+1}) = \frac{1}{2\pi\sigma_x^{t+1}\sigma_y^{t+1}\sqrt{1-\rho_{t+1}^2}} \exp\left[\frac{-Z}{2(1-\rho_{t+1}^2)}\right] \quad (3.11)$$

$$Z = \frac{(x_{t+1} - \mu_x^{t+1})^2}{\sigma_x^2} + \frac{(y_{t+1} - \mu_y^{t+1})^2}{\sigma_y^2} \quad (3.12)$$

Donde $\frac{\lambda}{2n} \sum_w w^2$ es la normalización L_2 y $\Pr(p_{t+1} | \hat{p}_{t+1})$ es la probabilidad de p_{t+1} dado un vector de entrada S con su respectiva salida \hat{p}_{t+1} .

3.1.3 Predicción

Una vez completado el entrenamiento se da inicio a la predicción en la que el modelo recibe los primeros n tiempos de estado $S_{t-n}, \dots, S_{t-1}, S_t$ y predice la ubicación de las embarcaciones después de m pasos de tiempo como $(\hat{x}_{t+m}, \hat{y}_{t+m})$. Se asume que la posición se describe por medio de una distribución normal bivariada con parámetros dados por la Ecuación

(3.8). Finalmente, se obtiene la predicción de ubicación de las embarcaciones mediante una distribución Gaussiana $(\hat{x}, \hat{y}) \sim N(\mu, \sigma, \rho)$.

3.2 Modelo COP fiscalización pesquera

La metodología de fiscalización pesquera en Chile, según la DGTM (2019), se realiza mediante buques de la Armada de Chile que poseen distintas capacidades o autonomía, distintos gastos o consumos como por ejemplo: aceite, petróleo, agua, horas de servicio de maquinaria o mantenimiento, entre otros. Adicionalmente, estas embarcaciones fiscalizadoras se encuentran distribuidas en las distintas Zonas Navales, actuando bajo una programación de fiscalización que se realiza con un tiempo previo determinado. Para luego instintivamente, las embarcaciones fiscalizadoras arriben a las áreas que contengan la mayor cantidad de embarcaciones pesqueras, con el objeto de inspeccionar el máximo número posible.

Similar al trabajo de Derya et al. (2020), relacionado con el modelo COP, los nodos pertenecen a diferentes clusters y sin ser necesario servir todos los nodos de un mismo cluster, se maximiza la cantidad de nodos total visitados, todo esto bajo un cierto límite de distancia. Además, se cuenta con las siguientes características operativas para una mejor fiscalización:

- Los conjuntos de embarcaciones o flotas se tratan como clusters.
- Al ingresar al cluster existe una limitación en cuanto a la distancia máxima que puede recorrer el vehículo dentro de éste, con el objeto de que exista mayor diversidad en las visitas de cada cluster.
- Se consideran múltiples buques fiscalizadores.
- Al asignar una ruta, debe ser hacia una posición futura del pesquero.

El último punto de las consideraciones es obtenido mediante un modelo de ML que se explica en la siguiente sección.

El modelo de fiscalización pesquera se formula de la siguiente manera:

- Conjuntos:

$N = \{v_0, v_1, \dots, v_n\}$ Es un conjunto de nodos o embarcaciones pesqueras, donde v_0 es el puerto de origen de embarcaciones fiscalizadoras.

$Clusters = \{CL_1, CL_2, \dots, CL_k\}$ Es un conjunto de Clusters de embarcaciones

pesqueras.

- Parámetros:

P = Número Suficientemente grande para penalización.

$c_{i,j}$ = Costo o distancia, $\forall i, j \in N$.

b = Número de embarcaciones fiscalizadoras.

D = Distancia máxima de la ruta completa(Kms).

DC = Distancia máxima dentro del cluster (Kms).

C_m = Nodo perteneciente a un Cluster, $\forall m \in Clusters$.

- Variables:

$x_{i,j}$: Variable binaria, 1 si arco $\langle i, j \rangle$ es atravesado, e.o.c. 0., $\forall i, j \in N$.

z_j : Variable binaria, 1 si nodo j es fiscalizado, e.o.c. 0., $\forall j \in N$

h_j : Variable binaria, 1 si el primero nodo de un cluster es visitado., $\forall j \in N$.

u_i : Distancia acumulada hasta el arribo al nodo i . (Puerto de origen es el primer vértice), $\forall i \in N$.

a_m : Distancia acumulada hasta el arribo al cluster, $\forall m \in Clusters$.

- Consideraciones:

1. El costo $c_{i,j}$ está asociado a cada arco (i, j) .
2. Existen múltiples vehículos disponibles y un límite de autonomía de toda la flota fiscalizadora D es impuesto.
3. No es necesario que todos los vértices pertenecientes a un cluster sean visitados consecutivamente, sino que puede visitar a distintos clusters de manera libre.
4. El objetivo es encontrar la ruta que maximice el número de fiscalizaciones realizadas.

- Formulación:

$$\max \sum_{i \in N \setminus \{v_0\}} z_i \quad (3.13)$$

Sujeto a:

$$\sum_{j \in N, i \neq j} x_{j,i} = z_i \quad \forall i \in N \setminus \{v_0\} \quad (3.14)$$

$$\sum_{j \in N, i \neq j} x_{i,j} = z_i \quad \forall i \in N \setminus \{v_0\} \quad (3.15)$$

$$\sum_{j \in N, j \neq 1} x_{1,j} = b \quad (3.16)$$

$$\sum_{j \in N, j \neq 1} x_{j,1} = b \quad (3.17)$$

$$u_i \leq Dz_i \quad \forall i \in N \setminus \{v_0\} \quad (3.18)$$

$$\sum_{i \in C_m} h_i \geq \sum_{i \in C_m} \frac{z_i}{|C_m|} \quad \forall m \in Clusters \quad (3.19)$$

$$u_j \geq u_i + c_{i,j} - P(1 - x_{i,j}) \quad \forall i \in N, j \in N \setminus \{v_0\}, i \neq j \quad (3.20)$$

$$a_m \leq u_i + P(1 - z_i) \quad \forall m \in Clusters, i \in C_m \quad (3.21)$$

$$u_i \leq a_m + DC + P(1 - z_i) \quad \forall i \in N \setminus \{v_0\}, m \in Clusters \quad (3.22)$$

$$a_m \leq u_i + P(1 - h_i) \quad \forall m \in Clusters, i \in C_m \quad (3.23)$$

$$h_i \leq z_i \quad \forall m \in Clusters, i \in C_m \quad (3.24)$$

$$\sum_{i \in N} \sum_{j \in N} x_{i,j} c_{i,j} \leq D \quad (3.25)$$

$$u_1 = 0 \quad (3.26)$$

$$x_{i,j} \in \{0, 1\} \quad \forall v_i \in N \quad (3.27)$$

$$z_j \in \{0, 1\} \quad \forall v_j \in N \quad (3.28)$$

$$h_j \in \mathbb{R} \quad \forall v_j \in N \quad (3.29)$$

$$u_j \in \mathbb{R} \quad \forall v_j \in N \quad (3.30)$$

$$a_m \in \mathbb{Z} \quad \forall m \in Clusters \quad (3.31)$$

La función objetivo (3.13) permite maximizar la cantidad de embarcaciones fiscalizadas. El conjunto de restricciones (3.14)-(3.15) imponen el número de arcos que entran y salen del nodo fiscalizado. Mientras que el conjunto de restricciones (3.16)-(3.17) establecen el número de embarcaciones fiscalizadoras que entran y salen del puerto de origen. Las restricciones (3.18) imponen el límite máximo de distancia de la ruta junto a las restricciones (3.25). El conjunto de restricciones (3.19) fuerzan a que el primer nodo visitado del cluster sea uno solo. Las restricciones (3.20) son restricciones de eliminación de subtours MTZ (Sawik, 2016). El conjunto (3.21) son restricciones dentro del cluster que imponen que el arribo al cluster sea menor que la variable de orden del tour general. El conjunto de restricciones (3.22) restringen la distancia recorrida dentro del cluster. Las restricciones (3.23) imponen el orden de tiempo entre arribo al cluster y arribo al primer nodo fiscalizado dentro del cluster. El conjunto de restricciones (3.24) aseguran que el primer nodo de cluster visitado tenga menor distancia que el nodo visitado. Las restricciones (3.25) imponen a la flota fiscalizadora una distancia máxima a recorrer. Las restricciones (3.26) fuerzan a que el nodo 1 comience en la distancia 0. Finalmente, el conjunto de restricciones (3.27) - (3.31) representan el dominio de las variables.

3.3 Integración de modelos

La integración de los modelos de ML y optimización se realiza de acuerdo al diagrama de la Figura 3.3. Inicialmente, existe un ingreso de las posiciones históricas de las embarcaciones pesqueras obtenidas por medios externos con sistemas AIS. Posteriormente, el modelo LSTM predice una región de probabilidad de localización de éstos, la cual debe ser cruzada con información actualizada que es obtenida desde factores externos cualitativos, como por ejemplo el clima o corrientes circundantes. Una vez efectuada dicha tarea, se selecciona un conjunto de coordenadas, el cual es clusterizado mediante el método K-means (Na et al., 2010). Además, se calcula la distancia entre cada coordenada para generar una matriz con ellas, las que son ingresadas al modelo COP. A su vez, este modelo entrega una solución para obtener la ruta más corta en recorrer el máximo de embarcaciones pesqueras llegando en forma sincronizada a cada

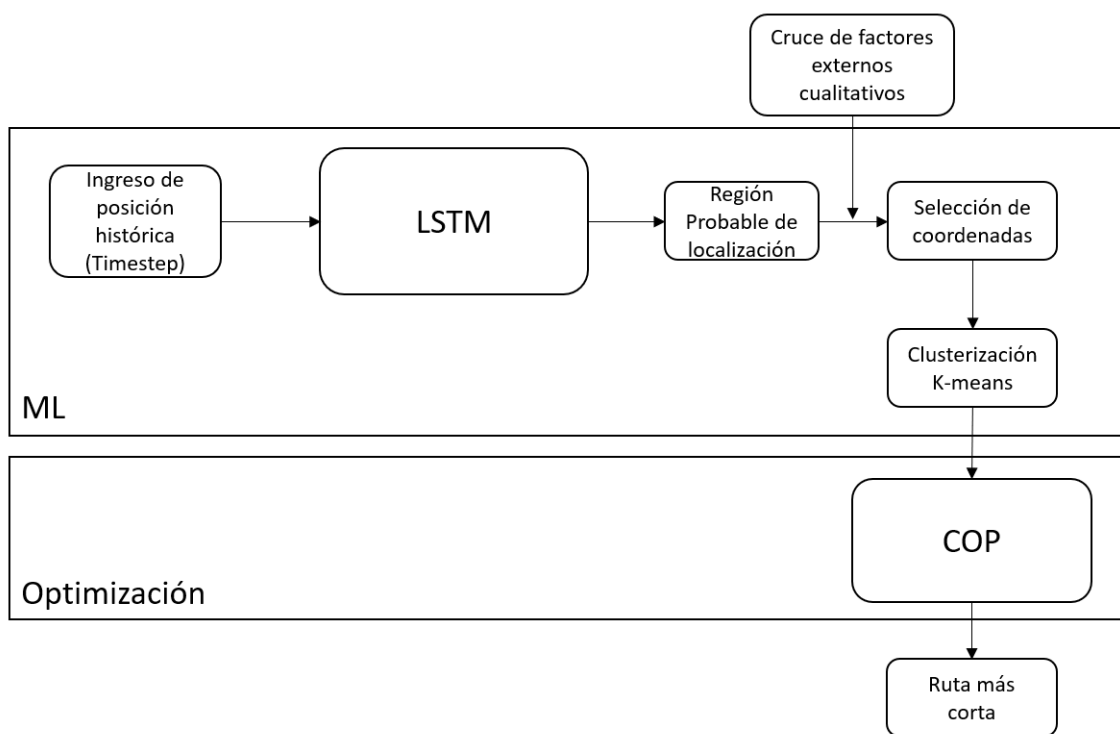


Figure 3.3: Integración de modelos.

Fuente: Elaboración propia

cluster, con el objeto de evitar el escape de embarcaciones. Vale destacar que para la medición de distancias se utiliza el método de Haversine (Mahmoud & Akkari, 2016), ya que es el más adecuado por su exactitud en este estudio.

Capítulo 4

Caso de estudio

En este capítulo se aplica la metodología propuesta anteriormente sobre un conjunto de datos AIS recopilados desde el repositorio del website Global Fishing Watch. Lo anterior, tiene la finalidad de testear el desempeño del modelo COP y ML de forma complementaria.

4.1 Recolección y preparación de datos

Con el objetivo de mostrar la efectividad de la metodología propuesta, primero se realiza la recolección y preparación de datos correspondiente a la descarga de base de datos desde la página Web Global Fishing Watch (GFW). Ésta es una organización no gubernamental sin fines de lucro formada con el objeto de colaborar con distintos estados en la sustentabilidad de las actividades extractivas en los océanos, la transparencia de información y el apoyo a la conservación de los océanos y sus recursos. Esta organización utiliza un sistema de identificación automática (AIS) sobre toda la superficie oceánica para la detección y seguimiento de todas las grandes embarcaciones que navegan en ella, recopilando información sobre el rumbo de la embarcación, posición geográfica y velocidad. Debido a la complejidad en que están almacenados los datos de GFW (esfuerzo pesquero) se deben preprocesar, limpiar y darles la forma correspondiente para ingresarlos a un modelo LSTM de forma entrada-etiqueta. En detalle, la base de datos extraída se encuentra en múltiples archivos, todos en formato *csv*, y comprenden los registros mundiales diarios por embarcación desde el año 2012 al 2018, conteniendo datos de identificación (MMSI), horas de pesca, latitud, longitud y fecha. Los datos de longitud y latitud se encuentran en una resolución de 100° , es decir, están agrupados en grillas de 0.01 por lado. Posteriormente, se realiza un consolidado de todos los archivos entre los años mencionados, hasta obtener un solo archivo *csv* mediante la librería *Dask* de Python. Una vez que los datos son filtrados y limitados entre las latitudes: 18.7°S - 67.45°S y longitudes 66.19°W - 79.19°W (considerando el límite marítimo chileno), se obtiene la Figura 4.1. El resumen del procesado de los datos se puede apreciar en la Tabla 4.1, mostrando la gran cantidad de

datos descartada solo en filtrado. No obstante son suficientes una vez que se interpolen y se generen múltiples rutas para cada MMSI. Por otra parte, también es necesario señalar que las embarcaciones pueden tener uno o varias trayectorias de viaje llamados tracks, registradas para distintas fechas.

Tabla 4.1: Comparación de datos pre y post procesado.

Procesos	Datasets	
	Total de coordenadas	Total de embarcaciones
Antes de Procesar	10.000	173
Después de Procesar	272.000	173

Fuente: Elaboración propia

Las coordenadas geográficas son transformadas a puntos geométricos en formato de proyección de la European Petroleum Survey Group o *EPSG* : 9155. Para la creación de los tracks se agrupan los datos en MMSI y fechas, aplicando interpolación cúbica a minutos. Además, para la creación de los tracks se utiliza un máximo de 25 puntos en cada uno. Por otro lado, se agregan 2 columnas adicionales: rumbo y velocidad, con el objeto de contar con más features para la predicción. Para esto, la velocidad en nudos se obtiene calculando en base a las distancias entre coordenadas y el tiempo entre coordenadas. Para la obtención del rumbo o ángulo de movimiento de las embarcaciones, se utiliza la Ecuación (4.1). (Ver Figura 4.2.)

$$\theta = \arctan \frac{[LON2 - LON1]}{[LAT2 - LAT1]} \quad (4.1)$$

Posteriormente, se generan las rutas o tracks de cada embarcación identificada por su MMSI único. Para esto se considera que la embarcación se encuentra en régimen estacionario si la variación de distancia entre coordenadas sucesivas es menor a 0.3 millas. Luego, se filtran los datos nuevamente para conformación de los tracks, los que finalmente son integrados al modelo LSTM. Así, se logra crear 369 tracks, como los que se muestran en la Figura 4.3.

Para terminar, los tracks son muestreados mediante el método sliding window de 10 minutos. En la Figura 4.4 se muestra un ejemplo de funcionamiento del proceso de sliding window (Vafaeipour et al., 2014), donde existen samples diarios, con un solo timestep y dos features correspondientes a la latitud y longitud. Para esto es necesario establecer un largo de track de 25 puntos horarios con su respectiva interpolación, objeto estandarizar el largo de la secuencia y conseguir con ello una mayor cantidad de tracks para el entrenamiento.

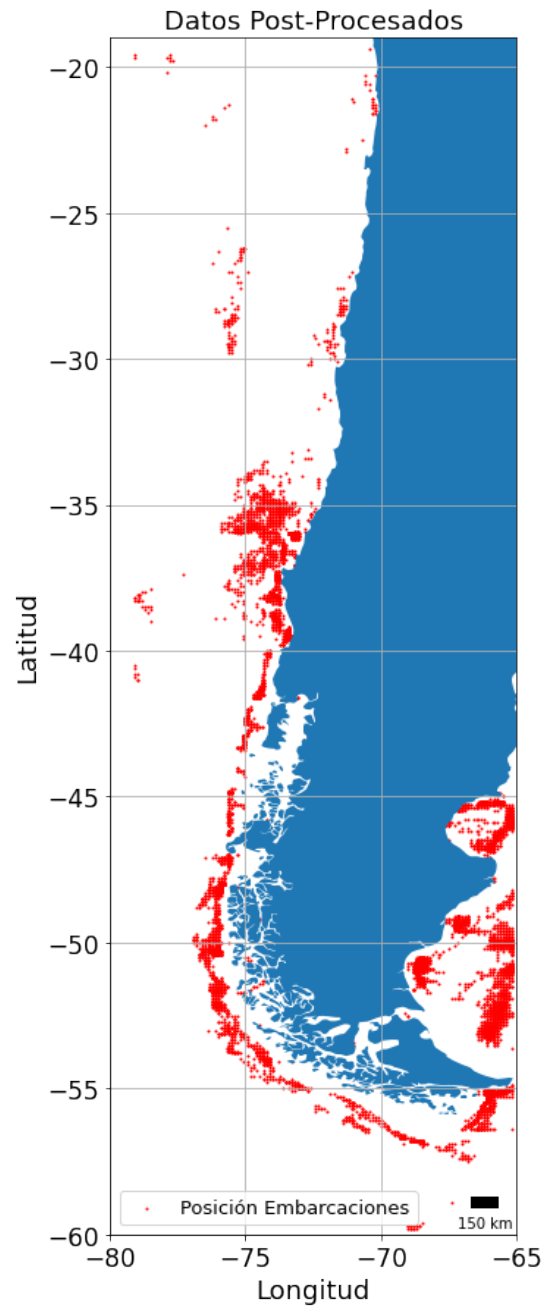


Figure 4.1: Visualización de los datos filtrados.
Fuente: Elaboración propia

La localización de los 10 minutos previos son consideradas como timestep para generar la localización probable 10 minutos después. Posteriormente, los datos son escalados entre [0,1] para acelerar el proceso de aprendizaje.

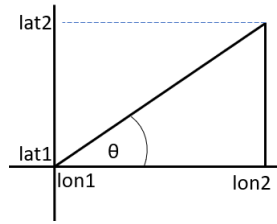


Figure 4.2: Arco tangente coordenadas.

Fuente: Elaboración propia

4.1.1 Implementación del modelo LSTM

A continuación, se realiza la implementación del modelo LSTM mediante el Framework Keras de Python. Los datos son divididos en dos sets: entrenamiento y pruebas, con una proporción 70 – 30% respectivamente. El set de entrenamiento se utiliza para entrenar el modelo y el set de prueba para obtener las predicciones finales. Su implementación se lleva a cabo de acuerdo a lo señalado en la metodología expuesta en este trabajo, donde se utiliza una red LSTM para predecir los parámetros de una distribución normal bivariada. Además, se utilizan dos capas ocultas con 128 unidades en el modelo. El tamaño del batch size utilizado en el entrenamiento es de 1024 con 100 épocas. Al mismo tiempo, se utiliza regularización L2 con parámetro $\lambda = 0.5$, para evitar el sobreajuste del aprendizaje. El algoritmo de optimización utilizado es Adam con una tasa de aprendizaje de 0.001 (learning rate). Para evaluar el desempeño del algoritmo se utiliza error cuadrático medio (RMSE). Cabe destacar que estos parámetros fueron utilizados debido al buen desempeño señalado en el estudio de Tang et al., 2019 para la predicción de trayectorias de embarcaciones.

La distribución de posición de las embarcaciones pesqueras predichas por el modelo LSTM son visualizadas cada 20 minutos, es decir, se observan los primeros 10 minutos de posición de una embarcación. Además, se predice la próxima posición 10 minutos después, es decir, el 20vo minuto. Dicha distribución se utiliza en la generación de instancias para el modelo COP.

4.1.2 Generación de instancias

Para generar las instancias que se utilizan en el modelo COP, se emplea el dataset de prueba definido previamente como el 30% del dataset total para la predicción. En particular, se elige una instancia aproximada a las coordenadas 36.1° S y 73.1 W, cercano a las costas de Chile para la predicción de la próxima posición. Las coordenadas de embarcaciones que se utilizan para la ejecución del modelo COP se encuentran definidos dentro del área de probabilidad generado por

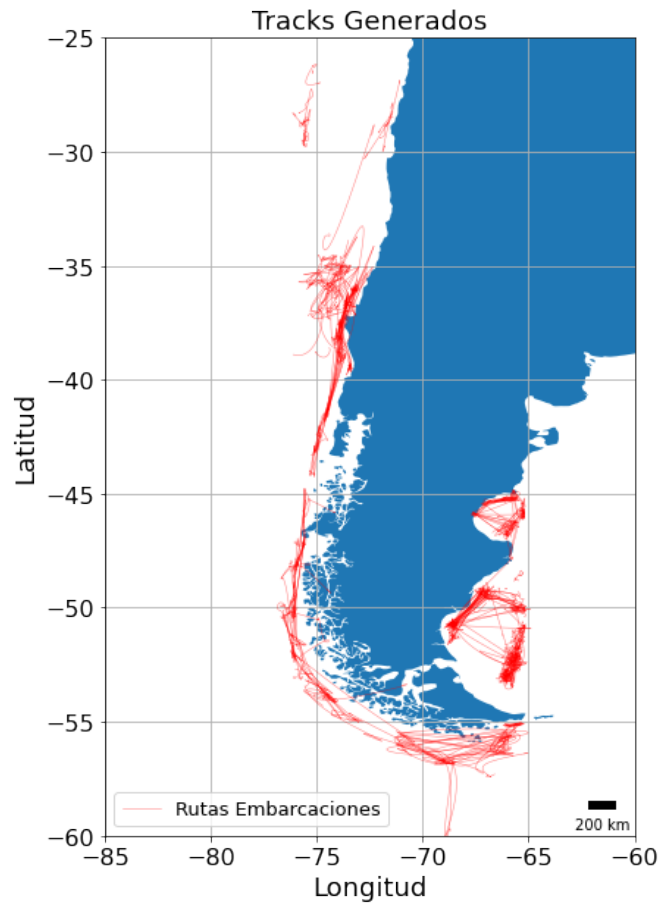


Figure 4.3: Visualización de tracks generados.
Fuente: Elaboración propia

el modelo LSTM, tal como se muestra en la Figura 4.5. Donde el símbolo "x" indica la etiqueta o posición real que debería tener la embarcación. Los puntos verdes son los datos de posición durante los 10 minutos previos de navegación. La zona de color mas claro significa que se tiene un máximo de probabilidad de contener la posición predicha dentro de los próximos 10 minutos. Para implementar el modelo COP se genera un grupo de puntos aleatorios al interior de la zona predicha, simulando un grupo de embarcaciones dentro de la zona probable de localización, se agrupan en clusters mediante un algoritmo de K-means e integran al modelo COP. En total, se generan 6 instancias, comenzando con un bajo número de nodos o embarcaciones para observar el comportamiento general del modelo, para posteriormente, efectuar un aumento progresivo de éstos. Asimismo, se aumenta el número de clusters y buques fiscalizadores hasta un máximo de cuatro, ya que es una cantidad razonable para realizar operaciones de este tipo y no se justificaría utilizar una mayor cantidad en la práctica, por los costos asociados.

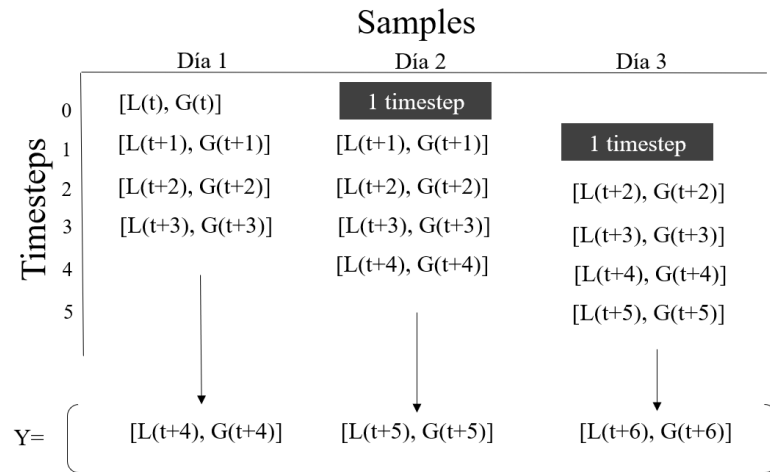


Figure 4.4: Método sliding window con 1 timeStep.

Fuente: Elaboración propia

En la Tabla 4.2 se muestra el resumen de las instancias generadas con el número de nodos (embarcaciones pesqueras) involucrados, N° de clusters y N° de buques fiscalizadores que realizarán la operación.

Tabla 4.2: Generación de instancias para el modelo COP.

Instancias	N° Nodos	N° Clusters	N° Buques Fiscalizadores	DC	D
1	5	2	1	5	5000
2	5	2	2	5	5000
3	10	2	1	5	5000
4	10	2	2	5	5000
5	10	4	1	5	5000
6	10	4	2	5	5000
7	10	4	4	5	5000
8	20	2	1	5	5000
9	20	2	2	5	5000
10	20	4	1	5	5000
11	20	4	2	5	5000
12	20	4	3	5	5000
13	20	4	4	5	5000

Fuente: Elaboración propia

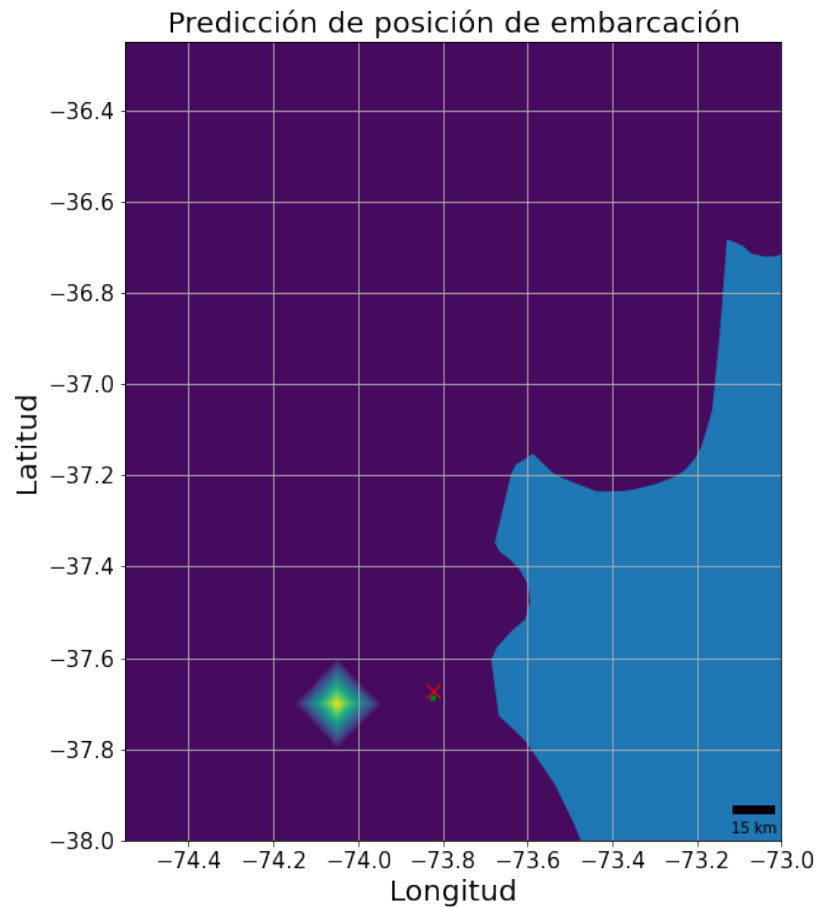


Figure 4.5: Predicción de posición de embarcación pesquera.

Fuente: Elaboración propia

4.1.3 Implementación del modelo COP

Para la implementación del modelo COP se utilizan matrices de distancias creadas bajo el método de Haversine y se establecen los parámetros señalados en la Tabla 4.2 para cada una de las instancias. Los parámetros de distancia máxima de la ruta general D para cada una de las instancias y la distancia máxima de recorrido al interior de cada cluster DC , se definen de acuerdo a consideraciones del establecimiento de la autonomía de petróleo de los buques fiscalizadores a un máximo de $200 m^3$. Lo que se traduce en una capacidad de navegación de aproximadamente 10 días seguidos a una velocidad de 12 nudos. Esto implicaría que la distancia máxima que podría recorrer el buque fiscalizador es en promedio de 5000 kms. Es así que se determina utilizar un parámetro $D = 5000$ y $DC = 5$.

Capítulo 5

Resultados y discusión

En este capítulo son presentados y discutidos los resultados obtenidos en el estudio. Las instancias utilizadas fueron originadas en la sección previa, hasta un máximo de 20 embarcaciones pesqueras y un máximo de 4 buques fiscalizadores con método computacional.

5.1 Resultados computacionales y discusión

Los resultados de la distribución normal bivariada obtenidos mediante el modelo LSTM arrojan un error RMSE de 2.87 respecto a la coordenada real para el parámetro μ . Asimismo, la desviación de los resultados para el caso de estudio es de 0.1° , lo que representa en la práctica 6 millas náuticas. Si bien la desviación es levemente alta, para efectos del caso práctico propuesto y considerando que se involucra una característica de incertidumbre en la predicción, esta desviación puede ser perfectamente sustituida mediante el alcance de detección de un radar marítimo. Las predicciones entregan un error de 0.15° de latitud y 0.07° de longitud, correspondientes a 9 y 4 millas náuticas respectivamente, tal como se muestra previamente en la Figura 4.5.

Respecto al modelo COP, en la Tabla 5.1 se detallan los parámetros de cada instancia y se resumen los resultados obtenidos como la “solución óptima” y el “tiempo de ejecución”, donde el “N° Nodos” corresponde a la cantidad de embarcaciones pesqueras de la instancia y la “solución óptima” concierne al número de embarcaciones fiscalizadas. El tiempo de ejecución está vinculado con la obtención de una solución óptima en cada instancia. No se considera tiempo límite en las iteraciones, ya que en todas las instancias se logra obtener solución óptima.

Los experimentos computacionales para el modelo COP se llevan a cabo mediante el software comercial IBM ILOG CPLEX Optimization Studio v12.9.0. sobre un sistema operativo Windows 10. El hardware utilizado fue un notebook Huawei modelo D14, memoria RAM de 8 GB y procesador AMD Ryzen 5-3500U de 2.10 GHz. Para los modelos de predicción

Tabla 5.1: Resultados modelo COP.

Instancias	N° Nodos	N° Clusters	N° Buques Fiscalizadores	Solución Óptima	Tiempo ejecución (seg.)
1	5	2	1	2	1
2	10	2	1	8	1
3	10	2	2	9	1
4	10	4	1	9	1
5	10	4	2	9	1
6	10	4	4	9	1
7	20	2	1	9	405
8	20	2	2	12	1098
9	20	2	3	16	3506
10	20	4	1	8	309
11	20	4	2	12	383
12	20	4	3	16	1751
13	20	4	4	18	6414

Fuente: Elaboración propia

de posicionamiento de embarcaciones se emplea el mismo hardware, la plataforma de Google Colab, Python 2.7 y el framework Keras para la implementación de la red.

Los resultados indican que el modelo alcanza a fiscalizar sobre un 50 % de los nodos de cada instancia, a excepción de las instancias 7 y 10. Ésta se debe al bajo número de embarcaciones fiscalizadoras disponibles, siendo lógico desde el punto de vista práctico, ya que una sola embarcación fiscalizadora no sería capaz de abarcar tantos nodos bajo limitaciones de autonomía. Además, los resultados sugieren que el modelo propuesto alcanza el máximo de embarcaciones pesqueras de acuerdo a la autonomía disponible. Por otra parte, se puede observar que a partir de los 20 nodos el modelo requiere mayor tiempo de ejecución, no obstante, se mantiene bajo un margen aceptable desde el punto de vista del tiempo necesario para su ejecución, previo al despliegue de las fuerzas navales de fiscalización. En la Tabla 5.2 se muestran las rutas obtenidas para cada buque fiscalizador correspondiente a cada instancia.

Con el objeto de visualizar el comportamiento particular de las rutas generadas para cada Buque fiscalizador (BF) con 5, 10, y 20 nodos en la Figura 5.1 se muestra de forma gráfica las rutas generadas para las instancias 1, 5 y 12 sobre un conjunto de nodos clasificados en colores para distinguir su pertenencia en cada cluster. En éstos se puede identificar el correcto comportamiento del modelo, es decir, el o los buques fiscalizadores intentan recorrer un cluster

Tabla 5.2: Solución de rutas buques fiscalizadores.

Instancia	Buques Fiscalizadores			
	BF 1	BF 2	BF 3	BF 4
1	1-5-3-2-1	-	-	-
2	1-9-5-10-6-3-1	-	-	-
3	1-4-6-3-1	1-9-5-10-8-7-1	-	-
4	1-9-5-10-6-3-1	-	-	-
5	1-4-6-3-1	1-7-9-5-10-2-1	-	-
6	1-7-1	1-8-2-1	1-9-5-4-1	1-10-6-3-1
7	1-5-20-19-6-14-18-1	-	-	-
8	1-5-20-8-17-2-9-1	1-10-12-18-19-7-1	-	-
9	1-5-20-8-17-2-9-1	1-10-12-18-19-7-1	1-16-4-6-14-1	-
10	1-5-20-18-19-6-14-1	-	-	-
11	1-5-20-8-17-2-9-1	1-10-12-18-19-7-1	-	-
12	1-5-20-8-17-2-9-1	1-10-12-18-19-7-1	1-16-4-6-14-1	-
13	1-2-9-15-11-1	1-5-20-6-14-3-1	1-8-18-19-17-1	1-10-12-16-4-1

Fuente: Elaboración propia

completo bajo una restricción o límite de distancia al interior de éste. Al percibir que la capacidad disponible para recorrer la distancia remanente al interior del cluster no es suficiente para terminarlo, se redirige a un nodo de otro cluster, todo esto sin sobrepasar la distancia total disponible de autonomía de los buques fiscalizadores. Por otra parte, después de recorrer los clusters, si el buque fiscalizador mantiene disponible aún autonomía remanente, se redirige nuevamente hacia alguno de los clusters con nodos pendientes de fiscalizar hasta completar su autonomía. Se puede apreciar también que los buques fiscalizadores no necesariamente visitan todos los nodos de cada cluster, por lo que el comportamiento del modelo es adecuado y acorde a lo propuesto.

La Tabla 5.3 muestra el comportamiento desde el punto de vista de los costos asociados a cada ruta, cumpliendo con la distancia total asignada. Asimismo, se puede apreciar que a medida que la cantidad de embarcaciones fiscalizadoras aumenta, la distancia a recorrer en total aumenta de forma similar para distintos números de clusters. Es decir, la variación de distancia entre un set de nodos que contenga 2 y 4 clusters para un número dado de embarcaciones fiscalizadoras se comportan de manera similar. Sin embargo, la variación al aumento de número de embarcaciones fiscalizadoras es notablemente mayor. De la misma tabla se observa que las rutas de las instancias 2 y 4 son idénticas, a pesar de tener distintos números de clusters (2 y 4 respectivamente). Asimismo, las instancias 9 y 12 obtienen la misma ruta para 2 y

Tabla 5.3: Distancia por ruta de buque fiscalizador.

Distancias (Costo)					
Instancia	Ruta 1	Ruta 2	Ruta 3	Ruta 4	Distancia Total
1	4092	-	-	-	4092
2	1069	-	-	-	1069
3	1056	1584	-	-	2640
4	1069	-	-	-	1069
5	1056	1051	-	-	2107
6	1022	1043	1048	1057	4172
7	1079	-	-	-	1079
8	1081	1065	-	-	2147
9	1081	1065	1072	-	3219
10	1072	-	-	-	1072
11	1081	1065	-	-	2147
12	1081	1065	1072	-	3219
13	1053	1059	1063	1058	4234

Fuente: Elaboración propia

4 clusters respectivamente, ambos con 3 vehículos. Las mayores distancias se logran en las instancias con 4 buques fiscalizadores y 4 clusters. Contrariamente, las menores distancias se obtienen cuando las instancias consisten solo de 1 buque fiscalizador, independiente del número de clusters involucrados. Los resultados arrojan una inclinación al aumento de las distancias recorridas cuando se utiliza una mayor cantidad de BF. De acuerdo a esto, se puede inferir que con este modelo existe una mayor diversificación o cobertura de fiscalizaciones al utilizar una mayor cantidad de BF, lo que se traduce en un uso de recursos eficiente al evitar rutas innecesarias o con mayor costo.

Considerando que los datos históricos utilizados en el modelo LSTM no tienen una calidad adecuada, deben ser pre-procesados y adaptados, lo cual puede provocar que el rango de incertidumbre obtenido mediante el error señalado al inicio de este capítulo sea más alto que el deseado. Sin embargo, es un error aceptable que puede ser cubierto por detecciones de radar como fue comentado anteriormente. Por otra parte, el entrenamiento del modelo LSTM demora aproximadamente 6 horas en su ejecución, por lo que este tiempo depende de la cantidad de datos históricos utilizados. Esto podría llegar a ser una limitante computacional al utilizar una mayor cantidad de datos. No obstante, las predicciones de posición de las embarcaciones son obtenidas en un tiempo mínimo. En la segunda fase de la integración, el modelo COP adaptado a la fiscalización de flotas pesqueras, cumple con el comportamiento deseado, alcanzando a fiscalizar sobre un 50 % de los nodos o embarcaciones con un bajo

costo o distancia, bajo un tiempo de ejecución máximo de 1 hora con 46 minutos, lo cual es aceptable para una planificación de operaciones de fiscalización, debido a que previo a los despliegues de Unidades fiscalizadoras existe un periodo de preparación tanto de logística como de personal, ligeramente superior. Finalmente, las rutas generadas por el modelo COP se cruzan visualmente, lo que se puede deber a la escala de visualización o simplemente al comportamiento del modelo, lo que debería ser mejorado. También es evidente que el BF intenta llegar a cada cluster, no siendo mandatorio su recorrido completo, lo cual es uno de los comportamientos buscados por el modelo. Para concluir este capítulo, algunas limitaciones del modelo propuesto son el número de embarcaciones pesqueras, ya que a contar de los 20 buques el tiempo de ejecución aumenta de manera exponencial. Por su parte, el error de predicción puede variar tanto positiva como negativamente, de acuerdo a la calidad y pre-procesamiento de los datos de entrenamiento del modelo LSTM. Por último, la integración de los modelos planteados representa una aproximación a la planificación de operaciones de fiscalización pesquera, que puede ser perfeccionado en orden a mejorar las limitaciones existentes.

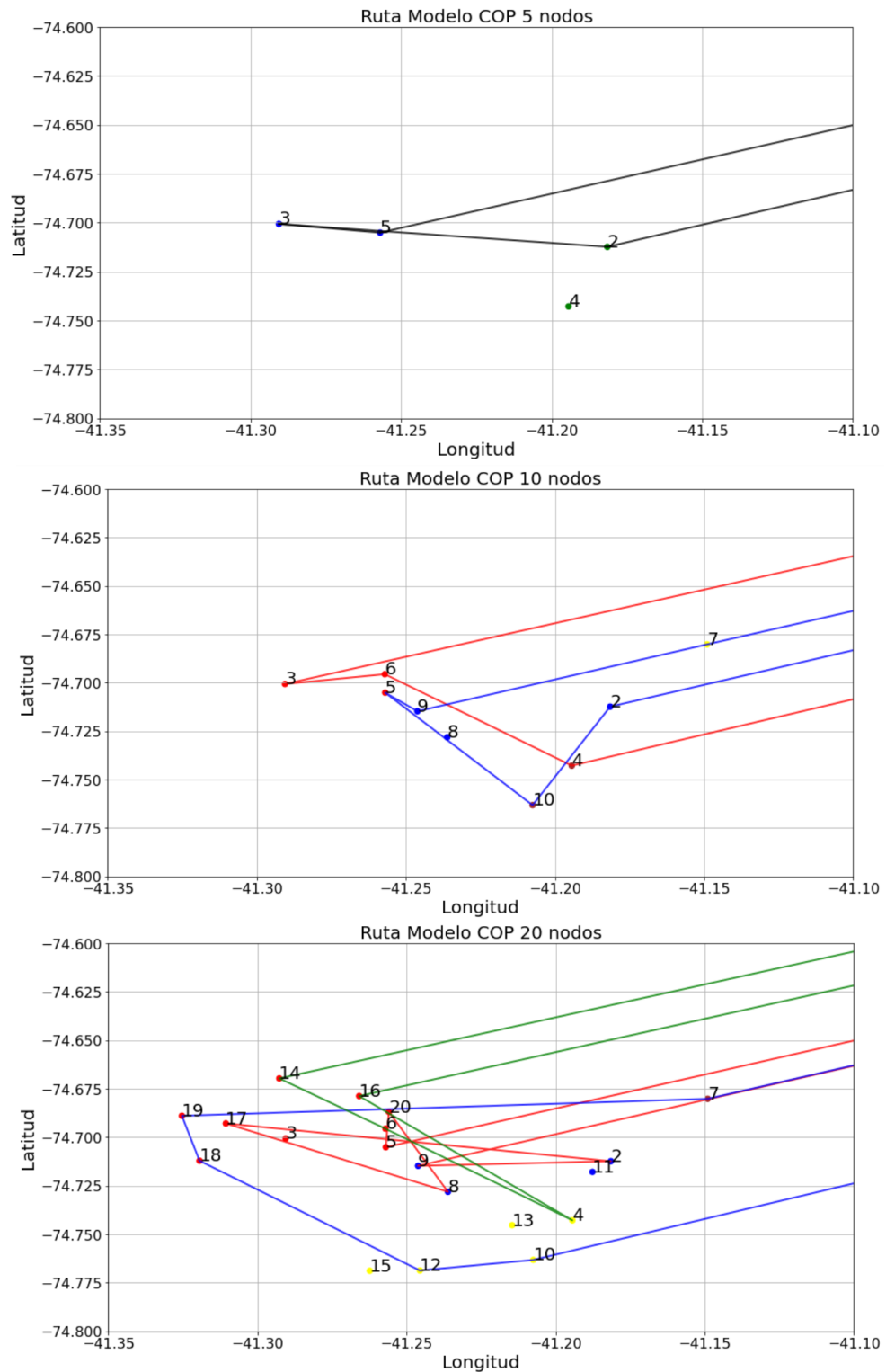


Figure 5.1: Rutas embarcaciones fiscalizadoras.

Fuente: Elaboración propia

Capítulo 6

Conclusiones

En este estudio se presenta un caso de estudio aplicado de integración de dos modelos, para optimizar operaciones de fiscalización pesquera que realiza la Armada de Chile. El primer modelo corresponde a una red neuronal recurrente de tipo LSTM utilizada para predecir la posición de una embarcación pesquera en los próximos 10 minutos, a partir de la observación de los primeros 10 minutos, en cercanías de las costas de Chile. La red LSTM está conformada por dos capas y es entrenada-testeada con un dataset abierto de Global Fishing Watch, para predecir una distribución de probabilidad de localización de embarcaciones pesqueras. A su vez, el modelo de red neuronal se integra a un modelo COP adaptado a las operaciones de fiscalización, siendo testeado con 13 instancias.

Mediante comparación experimental, se puede decir que el modelo LSTM desempeña un trabajo de predicción de posición de embarcaciones pesqueras con un margen de error aceptable, considerando que puede ser sustituido con las capacidades de radares marítimos y que puede ser mejorado con datasets de mejor calidad. La integración de modelos logra obtener un alto número de embarcaciones pesqueras fiscalizadas contemplando su localización en distintos clusters. lo realiza con un mínimo costo o distancia y siempre manteniendo un límite de autonomía de los buques fiscalizadores, cumpliendo con el comportamiento de las operaciones de fiscalización perseguido. Los resultados arrojan una inclinación al aumento de las distancias recorridas cuando se utiliza una mayor cantidad de BF. De acuerdo a esto, se puede inferir que con este modelo existe una mayor diversificación o cobertura de fiscalizaciones al utilizar una mayor cantidad de BF, lo que se traduce en un uso de recursos eficiente al evitar rutas innecesarias o con mayor costo.

La integración de modelos propuesta en este trabajo representa un acercamiento inicial al problema de optimización de las operaciones de fiscalización pesquera. Futuras extensiones propuestas para este trabajo son: (i) el complemento del modelo con técnicas de simulación, con el objeto de predecir dinámicamente los cambios de posición de las embarcaciones pesqueras

considerando la posición actual de un BF; (ii) la disminución del error de predicción de la red LSTM con datasets de mejor calidad y aplicación de otras técnicas de ML para su comparación; (iii) incluir capacidades a las flotas fiscalizadoras como una extensión de un modelo Vehicle Routing Problem (VRP).

Referencias

- Angelelli, E., Archetti, C., & Vindigni, M. (2014). The clustered orienteering problem. *European Journal of Operational Research*, 238(2), 404–414. <https://doi.org/10.1016/j.ejor.2014.04.006>
- Ayhan, S., & Samet, H. (2016). Aircraft trajectory prediction made easy with predictive analytics. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 21–30. <https://doi.org/10.1145/2939672.2939694>
- Bessel, F. W., Karney, C. F. F., & Deakin, R. E. (2010). The calculation of longitude and latitude from geodesic measurements. *Astronomische Nachrichten*, 331(8), 852–861. <https://doi.org/10.1002/asna.201011352>
- Chile, G. d. (1991). LEY GENERAL DE PESCA Y ACUICULTURA N° 18.892, 203. <https://www.subpesca.cl/portal/615/w3-article-88020.html>
- Chisman, J. A. (1975). The clustered traveling salesman problem. *Computers & Operations Research*, 2(2), 115–119. [https://doi.org/10.1016/0305-0548\(75\)90015-5](https://doi.org/10.1016/0305-0548(75)90015-5)
- Chuaysi, B., & Kiattisin, S. (2020). Fishing vessels behavior identification for combating IUU fishing: Enable traceability at sea. *Wireless Personal Communications*, 115(4), 2971–2993. <https://doi.org/10.1007/s11277-020-07200-w>
- Corvalán, I. A., & Abarca, L. D. (2019). DISPOSITIVOS TECNOLÓGICOS COMO MEDIO DE PRUEBA EN FISCALIZACIONES PESQUERAS: DRONES Y SISTEMA DE MONITOREO SATELITAL. *Repositorio Universidad de Chile*, 1(1), 100. <http://repositorio.uchile.cl/handle/2250/176149>
- Derya, T., Dinler, E., & Keçeci, B. (2020). Selective generalized travelling salesman problem. *Mathematical and Computer Modelling of Dynamical Systems*, 26(1), 80–118. <https://doi.org/10.1080/13873954.2019.1705496>
- de Souza, E. N., Boerder, K., Matwin, S., & Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning (A. C. Tsikliras, Ed.). *PLOS ONE*, 11(7), e0158248. <https://doi.org/10.1371/journal.pone.0158248>

- DGTM. (2019). Cuenta pública 2019. <https://www.directemar.cl/directemar/depto-pesca-acuicultura-y-rrmm/cuentas-publicas>
- Duca, A. L., Bacciu, C., & Marchetti, A. (2017). A k-nearest neighbor classifier for ship route prediction. *OCEANS 2017 - Aberdeen*, 1–6. <https://doi.org/10.1109/OCEANSE.2017.8084635>
- Dunn, D. C., Jablonicky, C., Crespo, G. O., McCauley, D. J., Kroodsma, D. A., Boerder, K., Gjerde, K. M., & Halpin, P. N. (2018). Empowering high seas governance with satellite vessel tracking data. *Fish and Fisheries*, *19*(4), 729–739. <https://doi.org/10.1111/faf.12285>
- González-Andrés, C., Sánchez-Lizaso, J. L., Cortés, J., & Pennino, M. G. (2020). Illegal fishing in isla del coco national park: Spatial-temporal distribution and the economic trade-offs. *Marine Policy*, *119*, 104023. <https://doi.org/10.1016/j.marpol.2020.104023>
- Kitjacharoenchai, P., Ventresca, M., Moshref-Javadi, M., Lee, S., Tanchoco, J. M., & Brunese, P. A. (2019). Multiple traveling salesman problem with drones: Mathematical model and heuristic approach. *Computers & Industrial Engineering*, *129*, 14–30. <https://doi.org/10.1016/j.cie.2019.01.020>
- Krisnafi, Y., Iskandar, B. H., & Wisudo, S. H. (2017). Optimization of fisheries surveillance vessel deployment in indonesia using genetic algorithm (case study: Fisheries management area 711, republic of indonesia). *IO*(3), 14.
- Mahmoud, H., & Akkari, N. (2016). Shortest path calculation: A comparative study for location-based recommender system. *2016 World Symposium on Computer Applications & Research (WSCAR)*, 1–5. <https://doi.org/10.1109/WSCAR.2016.16>
- Na, S., Xumin, L., & Yong, G. (2010). Research on k-means clustering algorithm: An improved k-means clustering algorithm. *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, 63–67. <https://doi.org/10.1109/IITSI.2010.74>
- Nacional, B. (2012). Convención sobre la conservación y ordenamiento de los recursos pesqueros en alta mar en el océano pacifico sur. <https://obtienearchivo.bcn.cl/obtienearchivo?id=recursolegales/10221.3/43476/1/Dto89.pdf>
- Salman, A. G., Heryadi, Y., Abdurahman, E., & Suparta, W. (2018). Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting. *Procedia Computer Science*, *135*, 89–98. <https://doi.org/10.1016/j.procs.2018.08.153>
- Sawik, T. (2016). A note on the miller-tucker-zemlin model for the asymmetric traveling salesman problem. *Bulletin of the Polish Academy of Sciences Technical Sciences*, *64*(3), 517–520. <https://doi.org/10.1515/bpasts-2016-0057>

- Tampakis, P., Chondrodima, E., Pikrakis, A., Theodoridis, Y., Pristouris, K., Nakos, H., Petra, E., Dalamagas, T., Kandiros, A., Markakis, G., Maina, I., & Kavadas, S. (2020). Sea area monitoring and analysis of fishing vessels activity: The i4sea big data platform. *2020 21st IEEE International Conference on Mobile Data Management (MDM)*, 275–280. <https://doi.org/10.1109/MDM48529.2020.00063>
- Tang, H., Yin, Y., & Shen, H. (2019). A model for vessel trajectory prediction based on long short-term memory neural network. *Journal of Marine Engineering & Technology*, 1–10. <https://doi.org/10.1080/20464177.2019.1665258>
- Vafaeipour, M., Rahbari, O., Rosen, M. A., Fazelpour, F., & Ansarirad, P. (2014). Application of sliding window technique for prediction of wind velocity time series. *International Journal of Energy and Environmental Engineering*, 5(2), 105. <https://doi.org/10.1007/s40095-014-0105-5>
- Varsamopoulos, S., Bertels, K., & Almudever, C. G. (2018). Designing neural network based decoders for surface codes. *Springer*, 13. <https://doi.org/10.1109/TC.2019.2948612>
- Yahiaoui, A.-E., Moukrim, A., & Serairi, M. (2019). The clustered team orienteering problem. *Computers & Operations Research*, 111, 386–399. <https://doi.org/10.1016/j.cor.2019.07.008>
- Zhang, W., Wang, K., Wang, S., & Laporte, G. (2020). Clustered coverage orienteering problem of unmanned surface vehicles for water sampling. *Naval Research Logistics (NRL)*, 67(5), 353–367. <https://doi.org/10.1002/nav.21906>
- Zhong, C., Jiang, Z., Chu, X., & Liu, L. (2019). Inland ship trajectory restoration by recurrent neural network. *Journal of Navigation*, 72(6), 1359–1377. <https://doi.org/10.1017/S0373463319000316>

UNIVERSIDAD DE CONCEPCIÓN - FACULTAD DE INGENIERÍA

RESUMEN DE TESIS MAGÍSTER EN INGENIERÍA INDUSTRIAL

Departamento		Ingeniería Industrial	
Título		Optimización de Operaciones de Fiscalización Pesquera con predicción de posición mediante Machine Learning.	
Nombre Memorista		Julio Espinoza Llanos	
Modalidad	Investigación	Profesor(es) Patrocinante(s)	
Concepto		Rodrigo De la Fuente, Ph.D.	
Calificación			
Fecha		Ingeniero Supervisor	Institución
Comisión (Nombre y Firma)			
Resumen			
<p>La fiscalización Oceánica corresponde a una de las tareas de monitoreo y control de la actividad pesquera que la Armada de Chile y otras instituciones realizan con el objeto de combatir la pesca no declarada no reglamentada (INDNR) y con ello salvaguardar la conservación de las especies marinas, contribuyendo a su vez con la sustentabilidad de la actividad pesquera de la cual depende el país. Aunque las actividades de fiscalización son intensas, la posición de las embarcaciones es dinámica, lo que dificulta su localización. Es por esto que es primordial para Instituciones como la Armada maximizar el número de buques fiscalizados al menor costo posible. Este estudio integra dos técnicas, consistentes en un modelo de Machine learning (ML) y otro de optimización. Con el objetivo de determinar la localización más probable de un grupo de embarcaciones pesqueras y a su vez, obtener la ruta más corta que maximice la cantidad de fiscalizaciones, el modelo de ML se compone de una red Neuronal de tipo Long short-term memory (LSTM). Éste tiene la capacidad de entregar regiones probables de localización de acuerdo a sus datos históricos, simulando la incertidumbre existente. Por otro lado, el modelo de optimización está basado en el problema clustered orienteering (COP). Los resultados muestran una desviación aceptable de la predicción de localización de flotas pesqueras de 6 MN, la que puede ser sustituida mediante capacidades de detección con radares marítimos y a su vez, se logró positividad frente a la obtención de un alto número de embarcaciones pesqueras fiscalizadas.</p>			