

## 장단기 기억 신경망을 활용한 선박교통 해양사고 패턴 분석 및 예측

장다운\* · 김주성\*\*†

\* 목포해양대학교 해상운송시스템학과 박사과정, \*\* 목포해양대학교 항해학부 교수

## Analysis and Prediction Methods of Marine Accident Patterns related to Vessel Traffic using Long Short-Term Memory Networks

Da-Un Jang\* · Joo-Sung Kim\*\*†

\* PhD Candidate, Graduate School of Maritime Transportation System, Mokpo National Maritime University, Mokpo 58628, Korea

\*\* Professor, Division of Navigation Science, Mokpo National Maritime University, Mokpo 58628, Korea

**요 약 :** 해양사고 예방을 위해서는 사고의 원인과 결과에 대한 분석 및 진단뿐만 아니라, 사고의 발생 패턴과 변화 추이를 예측함으로써 정량적 위험도를 제시할 필요성이 있다. 선박교통과 관련된 해양사고 예측은 선박의 충돌위험도 분석 및 항해 경로 탐색 등 선박교통의 흐름에 관한 연구가 주로 수행되었으며, 해양사고의 발생 패턴에 대한 분석은 전통적인 통계 분석에 따라 제시되었다. 본 연구에서는 해양사고 통계 자료 중 선박교통관련 사고의 월별, 시간대별 발생 현황 데이터를 활용하여 해양사고 발생 예측 모델을 제시하고자 한다. 국내 해양사고 발생 현황 중 월별, 시간대별 데이터 집계 가능한 1998년부터 2021년까지의 통계자료 중 선박교통 관련 데이터를 분류하여 정형 시계열 데이터로 변환하였으며, 대표적인 인공지능 모델인 순환 신경망 기반 장단기 기억 신경망을 통하여 예측 모델을 구축하였다. 검증데이터를 통하여 모델의 성능을 검증한 결과 RMSE는 초기 신경망 모델에서 월별 52.5471, 시간대별 126.5893으로 나타났으며, 관측값으로 신경망 모델을 업데이트한 결과 RMSE는 월별 31.3680, 시간대별 36.3967로 개선되었다. 본 연구에서 제안한 신경망 모델을 기반으로 다양한 해양사고의 특징 데이터를 학습하여 해양사고 발생 패턴을 예측할 수 있을 것이다. 향후 해양사고 발생 위험의 정량적 제시와 지역기반의 위험지도 개발 등에 관한 추가 연구가 필요하다.

**핵심용어 :** 선박교통, 장단기 기억 신경망, 순환신경망, 해양사고, 패턴인식

**Abstract :** Quantitative risk levels must be presented by analyzing the causes and consequences of accidents and predicting the occurrence patterns of the accidents. For the analysis of marine accidents related to vessel traffic, research on the traffic such as collision risk analysis and navigational path finding has been mainly conducted. The analysis of the occurrence pattern of marine accidents has been presented according to the traditional statistical analysis. This study intends to present a marine accident prediction model using the statistics on marine accidents related to vessel traffic. Statistical data from 1998 to 2021, which can be accumulated by month and hourly data among the Korean domestic marine accidents, were converted into structured time series data. The predictive model was built using a long short-term memory network, which is a representative artificial intelligence model. As a result of verifying the performance of the proposed model through the validation data, the RMSEs were noted to be 52.5471 and 126.5893 in the initial neural network model, and as a result of the updated model with observed datasets, the RMSEs were improved to 31.3680 and 36.3967, respectively. Based on the proposed model, the occurrence pattern of marine accidents could be predicted by learning the features of various marine accidents. In further research, a quantitative presentation of the risk of marine accidents and the development of region-based hazard maps are required.

**Key Words :** Vessel traffic, Long short-term memory network, Recurrent neural network, Marine accident, Pattern recognition

\* First Author : dujang@mmu.ac.kr, 061-240-7452

† Corresponding Author : jskim@mmu.ac.kr, 061-240-7193

## 1. 서론

중앙해양안전심판원과 국가통계포털의 통계 자료에 따르면, 2012년부터 2021년까지 최근 10년간 대한민국 영해 내에서 발생한 해양사고는 총 22,504건으로 연평균 2,250건이며, 매 5년간 평균 해양사고 발생 건수는 1,681건(2012년에서 2016년)에서 2,820건(2017년에서 2021년)으로 그 증가율은 67.78%로 크게 증가하였다(Korean Maritime Safety Tribunal, 2022; Korean Statistical Information Service, 2022). 이 기간 충돌, 접촉, 좌초 등 선박교통 관련 해양사고는 2,270건으로 전체 해양사고의 10.09%를 차지한다. 한편, 2017년부터 2021년까지 최근 5년간 선박교통 관련 해양사고의 연간 추이를 살펴보면, 2017년 432건으로 전체 해양사고의 16.73%, 2018년 412건(15.42%), 2019년 422건(14.20%), 2020년 514건(16.29%), 2021년 421건(15.48%)으로 매년 전체 해양사고의 14~16%를 차지하며 지속적으로 발생함을 알 수 있다. 해양사고는 인명과 재산의 손실을 야기할 뿐만 아니라 심각한 해양환경 오염을 유발할 수 있으며, 일반적으로 해양사고의 유발 요인을 교통 환경의 변화와 선박교통의 밀집에 따른 해양사고 위해 요소의 증가를 그 주요 요인으로 지목하고 있다(Kim, 2014). e-navigation의 도입과 자율운항선박으로 대표되는 첨단 의사결정 시스템의 도입, 그리고 해양사고의 예방과 잠재적 위험 요인의 효율적인 관리를 위한 선박교통관제(Vessel Traffic Services, VTS) 제도의 운영에도 불구하고 이러한 해양에서의 위협은 지속되고 있다(International Maritime Organization, 2020; Lee and Kim, 2022). 해양사고의 직간접적 요인은 선박과 운항자, 기상과 운항환경 등 여러 요인으로 분류되며, 내외부의 영향에 따라 직관적인 예측이 어렵고, 그 위험성을 정량적으로 표현하는 것 또한 도전적인 연구과제로 남아있다. 이러한 해양에서의 사고 위험의 모호성으로 인하여 선박 교통의 안전 및 효율성 제고를 통한 재산과 환경의 보호를 목적으로 VTS가 운영되고 있으며, 여러 연구에서 그 가치와 중요성을 인정받고 있다(Lee et al., 2022; Shin and Ji, 2020; Jeong and Jung, 2012). 해상에서의 선박 운항 및 선박 교통과 관련된 해양사고는 운항자 측면뿐만 아니라 관찰자 측면인 선박 교통관제사(Vessel Traffic Services Operator, VTSO)의 측면에서도 정량적 예측이 어려우며, 선박 교통의 특수성으로 인한 적시에 다량의 정보처리, 제한된 인적·물적 자원의 활용, 광범위한 대상 해역에서 다중 선박의 감시 및 관리, 제한된 해양기상 및 항해환경 등과 같은 여러 제약 사항이 존재하는 것으로 알려져 있다(Kim, 2013). 교통관련 사고의 예측과 패턴 분석은 육상 교통에서 주로 활발하게 이루어지고 있으며, 육상에서뿐만 아니라 해상에서도 교통관련 사고를 사전에 예방하기 위해서는 미래에 발생할 사고의 빈도와 추세를

정확히 예측하고, 이를 경감시킬 수 있는 제도적 장치와 기술적 접근이 요구 된다(Roh and Bac, 2021). 본 연구에서는 해양사고의 종류와 발생 시각을 집계하여 공표한 지난 1998년부터 2021년까지 24년 동안의 해양사고 통계 자료 중 선박교통관련 사고의 월별, 시간대별 발생 현황 데이터를 수집하고, 이를 통하여 해상 및 육상의 항해 의사결정 지원을 위한 시스템 설계와 기능적 요구에 충족하는 해양사고 발생 예측 모델을 제시하고자 한다.

## 2. 선행 연구

해양사고 분석에 관한 연구는 크게 해양 안전 제도 정비를 위한 정책적 연구와 해양사고의 원인 규명을 위한 내용적 분석으로 분류할 수 있으며, 해양사고의 사례 분석, 사고의 종류 및 발생 현황의 통계적 분석과 기술적 분석 등이 그 방법론으로 제시되어 왔다. Lee et al.(2019)은 관련 연구에서 Cross join 분석 방법을 적용하여 선종 및 사고유형에 따른 관계 분석과 M-SHELL Model 및 FTA기법을 활용한 선박 침몰사고의 사고 원인을 분석하여 해양사고의 원인 조사 분석 방법에 대한 개선 방안을 제시하였다. 한편, 통계적 분석을 통한 해양사고 원인 분석에 관한 연구로는 어선의 해양사고 빈도 분석과 해양사고의 종류별, 원인별 분석을 통한 안전성 제고 방안에 관한 연구가 수행되었으며(Jung, 2018; Jung et al., 2017), 분산분석 기법을 통한 해양사고의 발생 시간대와 집단, 속력, 상대 거리의 유의성 분석을 통한 해양사고 예방 대책을 제언한 연구도 수행되었다(Park and Ahn, 2007). 한편, 사례 분석 중심의 연구로는 한국과 일본의 해양오염 사고 사례를 중심으로 지방정부의 재난 대응 체계를 비교한 사례 연구(Hyun et al., 2009)와 여객선 사고를 중심으로 해양사고의 사례 분석과 해양경찰의 대응방안에 대한 정책적 제도 개선에 관한 연구 등이 있다(Roh, 2014). 그러나 이러한 사례 분석과 통계적 기법을 활용한 연구에서는 과거의 발생 빈도에 따른 추정을 근거로 수행되기 때문에 모집단의 데이터가 전체 현상을 대표하지 못할 수도 있다는 한계가 있으며, 연구자의 의도에 따라 설계된 절차를 통해 분석되어 고정된 모형의 형태로 수행된다는 한계가 있다. 따라서, 데이터 분석에 있어 기계학습을 통한 학습대상의 선정과 지속적인 반복 학습을 통한 성능 개선을 특징으로 하는 인공지능 기법이 도입되었다. 인공지능 기법을 활용한 해양사고 분석 및 예측 모형 개발에 관한 연구로는 해양사고의 기술적 분석과 이를 통한 Grid based Data-set 구축과 인공지능망을 적용한 연안 해역의 해양사고 예측 모형 개발에 관한 연구가 있으며(Shin et al., 2021), 이와 유사한 육상 교통 발생 패턴 예측에 관한 연구로는 장단기 기억 신경망을 활용한 운전자

연령대별 교통사고 발생 빈도와 사고 시간에 따른 발생 빈도 예측 모델 개발에 대한 연구가 있다(Roh and Bae, 2021). 이러한 인공지능 기반의 시계열 데이터 예측에 관한 연구에서는 예측 성능이 검증된 장단기 기억 신경망을 주로 활용하였으며, 이러한 방법론은 재화에 대한 가격지수 예측과 전력량 수요에 대한 예측, 화재 등 재난에 대한 발생 예측 등으로 연구 활용 범위가 확장되었다(Bae and Yu, 2018; Sun et al., 2021; Lee and Song, 2021; Yun and Lee, 2022). 또한, 최근에는 2019년 발생한 Covid-19에 따른 확진자 변화 및 예측 모델 개발에 관한 연구도 수행된 바 있다(Kim and Kim, 2021). 따라서, 본 연구에서는 기존의 사례 분석과 통계적 분석 방법의 제한점을 극복하고, 정확한 시계열 데이터 예측을 위하여 기존의 머신러닝 기법 중 순환 신경망 기법의 하나인 장단기 기억 신경망을 통한 선박교통 관련 해양사고 예측 모델 개발에 활용하고자 하며, 예측 모델의 성능 개선을 위하여 알고리즘 구성에 있어 초기 신경망 모델을 관측값으로 업데이트하여 모델의 최종 성능을 개선하고자 한다.

### 3. 예측 모델 개발 및 시뮬레이션

#### 3.1 장단기 기억 신경망

네트워크의 반복 구조를 통하여 이전 단계에서 얻은 정보가 지속되도록 고안된 순환 신경망(Recurrent Neural Network, RNN) 기법 중 하나인 장단기 기억(Long Short-Term Memory, LSTM) 모델은 1997년 Hochreiter and Schmidhuber에 의하여 제안되었다(Hochreiter and Schmidhuber, 1997). Fig. 1과 같이 일반적인 RNN의 구조 A는 하나의 네트워크를 반복하여 입력값  $x_t$ 에 대하여 출력값  $h_t$ 를 입력으로 하는 루프를 형성하여 현재의 상태가 다음 상태에 영향을 미치도록 하는 구조이다. 여기서 일반적인 RNN은 back-propagation 과정에서 급격한 기울기 값의 감소 문제가 발생하며, 이는 셀 간격이

멀어질수록 성능이 저하되는 장기 의존성 문제가 발생하는 한계가 있었다(Shin et al., 2017). LSTM은 이러한 장기 의존성 문제를 해결함으로써 시계열 데이터 분석 및 자연어 처리 등에 널리 사용되어 왔다(Lee and Song, 2021). Fig. 2는 simple RNN과 LSTM의 구조를 도식화한 그림이다. RNN은 네트워크 모듈을 반복하는 체인의 형태를 가지고 있다. LSTM은 RNN과 같은 체인 구조를 가지고 있지만, 각 모듈의 구조가 RNN과는 상이한 형태를 가지고 있다. LSTM 순환 신경망의 구조는 Fig. 3과 같고, 각 계층에서의 동작은 식(1)~(4)와 같이 표현할 수 있으며, 각 계층의 설명은 다음과 같다(Olah, 2015). 여기서,  $x$ 는 입력값,  $h$ 는 은닉층(hidden state),  $C$ 는 셀 상태(cell state),  $b$ 는 편향(bias),  $W$ 는 가중치(weight),  $\sigma$ 는 시그모이드(sigmoid),  $o$ 는 출력 gate 결과를 의미한다. 첫 번째 cell 구조에서 forget gate layer는 어떤 정보를 버릴 것인지를 결정하는 것으로서  $h_{t-1}$ 과  $x_t$ 를 받아 0과 1사이의 값을 sigmoid 함수를 거쳐 반환한다. 다음으로 input gate layer는 sigmoid 및 tangent layer를 거쳐 cell state에 저장할 값을 결정한다. 다음으로 update gate layer에서는 기존의 cell state 값에 갱신된 값을 더하여 도출한다. 마지막으로 출력값을 결정하는 output gate layer에서는 sigmoid와 tangent layer의 출력값을 곱하여 output으로 전달한다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \end{aligned} \tag{2}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{3}$$

$$\begin{aligned} o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \tag{4}$$

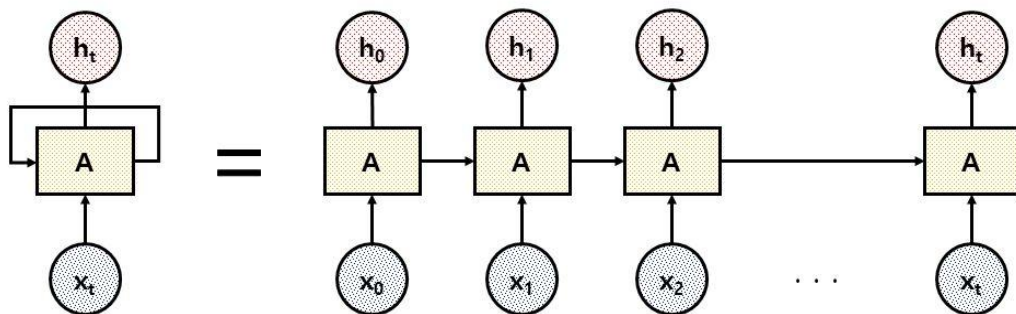


Fig. 1. RNN structure.

장단기 기억 신경망을 활용한 선박교통 해양사고 패턴 분석 및 예측

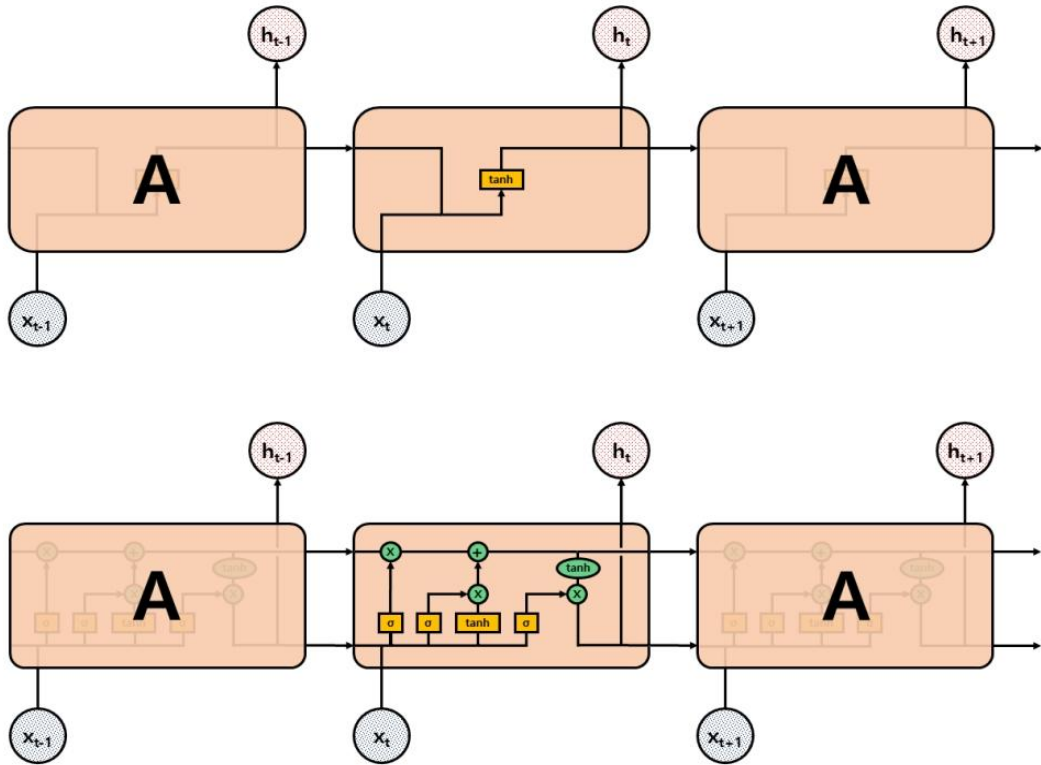


Fig. 2. Comparison of simple RNN and LSTM network.

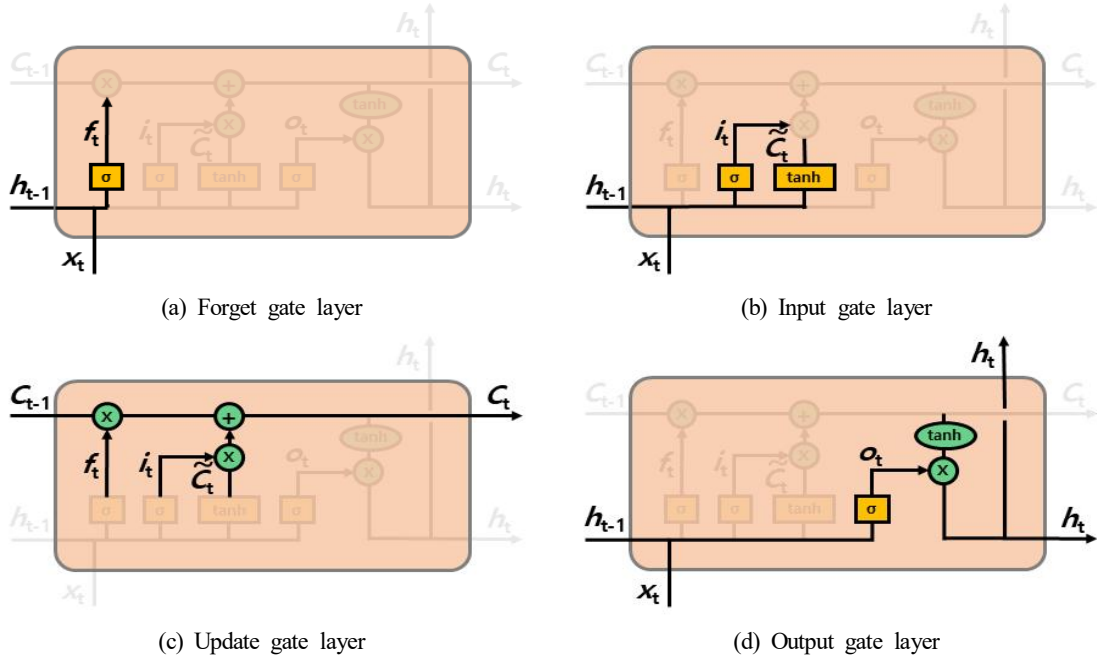


Fig. 3. LSTM gate structure.

해양사고 예측 모형은 과거 발생한 사고 통계 결과에서 획득한 데이터를 참조하여 미래에 발생 가능한 해양사고를 예측하는 것이다. 따라서 과거 통계 데이터와 새롭게 입력되는 데이터를 hidden layer에 얼마나 반영할 것인가의 세부 설정이 본 알고리즘을 통하여 업데이트되는 장점을 가지고 있어 본 연구에서는 LSTM 순환신경망을 주요 추정 알고리즘으로 활용하였다. 해양사고 발생의 추세적 변화량을 회귀 분석으로 도출할 경우 여러 특징 데이터가 추가되었을 때 모수가 적어지거나 전체 해양사고에 대한 추정치를 산출할 경우, 일부 특징 데이터에 대한 추정이 어려울 수 있다는 문제가 존재하지만, LSTM 순환신경망을 활용한 모델에서는 bias와 weight를 업데이트하기 때문에 이러한 문제에서 비교적 자유로울 수 있다는 장점도 있다(Lee and Song, 2021).

### 3.2 예측 모델 개발

LSTM을 활용한 선박교통 해양사고 시계열 데이터 예측 방법은 Fig. 4와 같다. LSTM 신경망은 시간 스텝에 따라 하나의 시간 스텝을 예측한 후 신경망의 상태를 업데이트 한다. 정형 데이터는 정해진 규칙이 없는 비정형 데이터와 달리 일정한 형태를 나타내므로 발생 건수에 따른 시계열 데이터의 벡터 형태로 전처리한다. 여기서, 머신러닝 모델을 위한 학습 데이터의 형태는 학습 전 적합한 피팅과 특징 변수의 편향, 최적화를 위하여 정규화 작업을 수행한다. 정규화 작업에서는 일반적으로 식(5)의 Min-Max Normalization 과 식(6)의 Z-Score Normalization 방법을 사용하며, Min-Max

Normalization의 경우 이상치에 대하여 취약하다는 단점이 있으므로 본 연구에서는 평균이 0, 분산이 1이 되도록 변환하는 Z-Score Normalization 방법을 사용하였다. 여기서,  $x$ 는 학습에 사용되는 가공된 데이터 세트의 모든 변수를 의미하며,  $\sigma$ 는 표준편차,  $\mu$ 는 평균을 의미한다.

$$\tilde{x}_i = \frac{x_i - \min x}{\max x - \min x} \quad (5)$$

$$Z_i = \frac{x_i - \mu}{\sigma}$$

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (6)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2}$$

데이터 정규화 이후 훈련 데이터와 테스트 데이터로 데이터셋을 분류하며, 본 연구에서는 90%의 훈련 데이터와 10%의 테스트 데이터로 분류하여 학습을 진행하였다. 다음으로 LSTM 신경망 모델은 최적의 성능을 도출하기 위해 알고리즘에 사용되는 적절한 hyper-parameter의 선정이 요구된다. 튜닝과정에서 시행 착오법을 통해 parameter를 결정하였으며, LSTM 신경망 모델을 정의하는 과정에서 hidden layer 내의 hidden unit의 수, activation function, gradient threshold, optimizer, epoch 등을 선정하여야 한다.

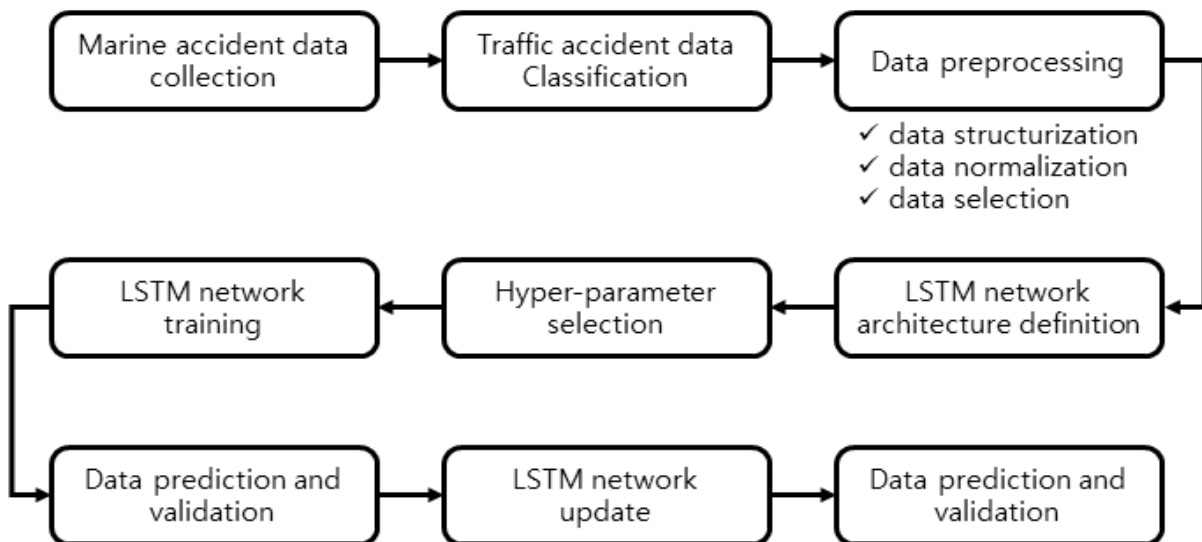


Fig. 4. Model design and simulation process.

## 장단기 기억 신경망을 활용한 선박교통 해양사고 패턴 분석 및 예측

본 연구에서는 월별 데이터의 경우 hidden unit을 72, 시간대별 데이터의 경우 54로 선정하였으며, solver는 adam으로 지정하여 150 epoch에 대하여 훈련을 진행하였다. 여기서 기울기의 임계값은 1로 설정하였으며, initial learning rate를 0.005로 하고, 100회 반복 후 학습률을 저하시키는 방법을 사용하였다. 설계된 LSTM 모델의 parameter 조절에 따른 오차검증 및 성능평가 방법은 평균 제곱근 오차(Root Mean Squared Error, RMSE)를 사용하였다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

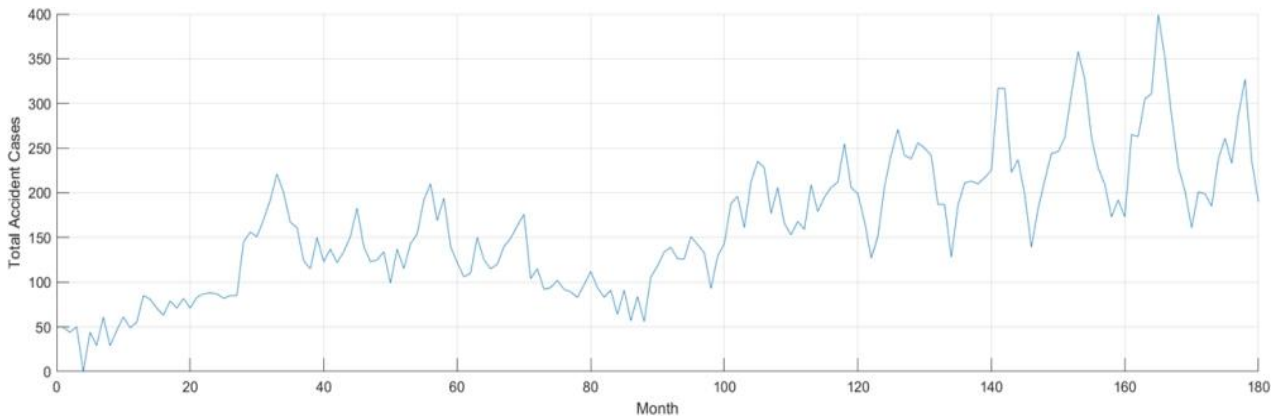
$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (8)$$

RMSE는 관찰값과 예측값의 차를 제공하여 평균화하는 식(7)의 평균 제곱 오차(Mean Squared Error, MSE)에 루트를 취한 값으로 MSE는 오차의 제곱을 구하기 때문에 실제 평균보다 커지는 특성이 있어 왜곡을 줄여 잔차(residual)들을

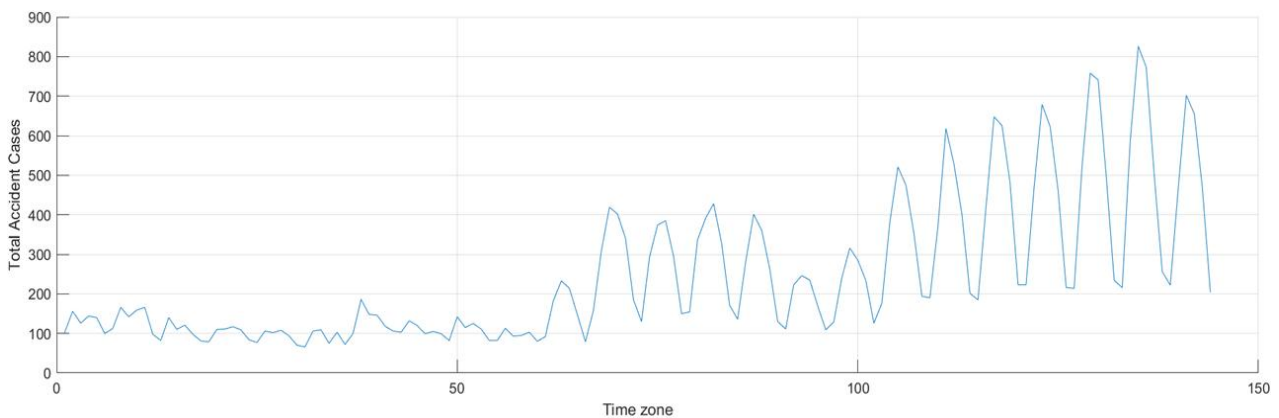
하나의 값으로 종합할 때 식(8)의 RMSE를 사용한다. 여기서,  $y$ 는 관찰값을  $\hat{y}$ 은 예측값을 의미한다.

### 3.3 시뮬레이션

해양사고 통계 자료는 중앙해양안전심판원과 국가통계포털에서 제공하는 사고 종류별, 월별, 시간대별 데이터를 활용하였다. 두 기관에서는 월별 통계 데이터의 경우 2007년 1월부터 2021년 12월까지 180개월의 데이터를 제공하며, 시간대별 통계 데이터의 경우 1998년부터 2021년까지 24년 동안의 데이터를 제공한다. 원본 데이터는 필드를 구분한 텍스트 파일인 CSV(Comma-Separated Values) 형식의 데이터를 정형화된 형태로 활용할 수 있다. Fig. 4에 따라 수집한 데이터를 전처리하여 시각화한 결과는 Fig. 5(a) 월별 선박교통 관련 해양사고 통계 및 시간대별 사고 발생 통계로 Fig. 5(b)와 같다. Fig. 5에서 선박교통 해양사고 건수는 점차 증가하는 것을 확인할 수 있으며, Fig. 6(a) 및 (b)와 같이 8월에서 10월, 08시에서 16시 사이에 비교적 해양사고 건수가 많이 발생하는 것을 확인할 수 있으나, 특정한 패턴을 확인할 수는 없었다.

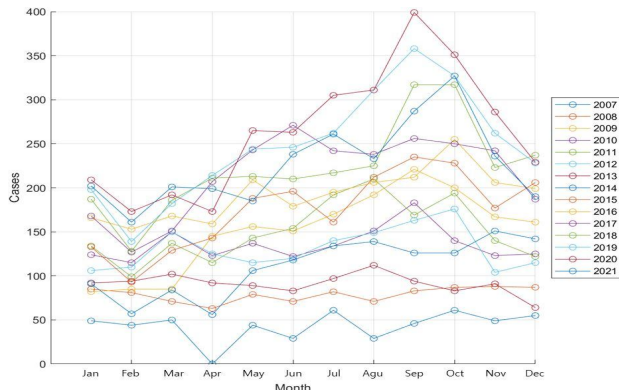


(a) Number of accidents per month

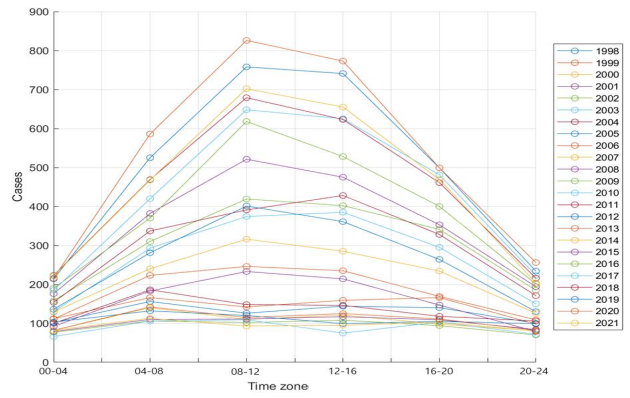


(b) Number of accidents per time zone

Fig. 5. Time series data of marine accidents.

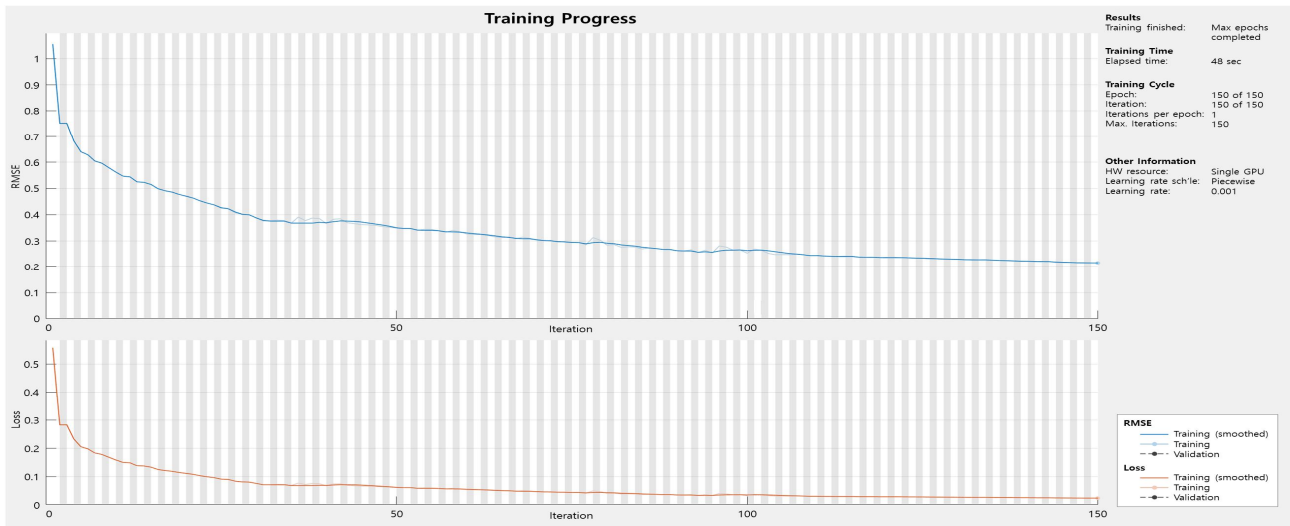


(a) Number of accidents per month

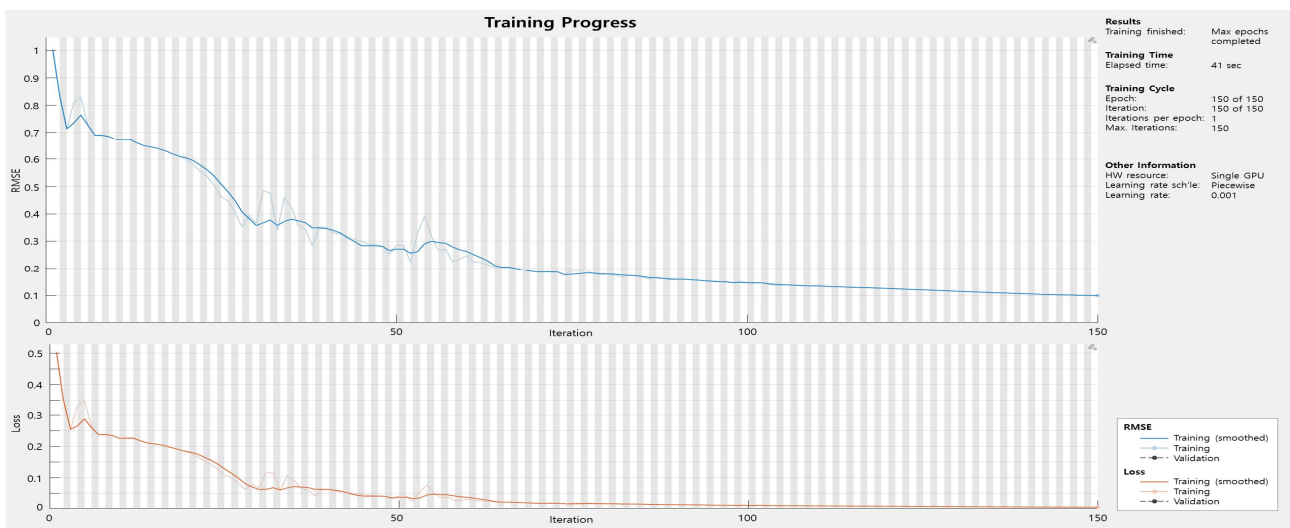


(b) Number of accidents per time zone

Fig. 6. Time series data of marine accidents.



(a) Accidents per month



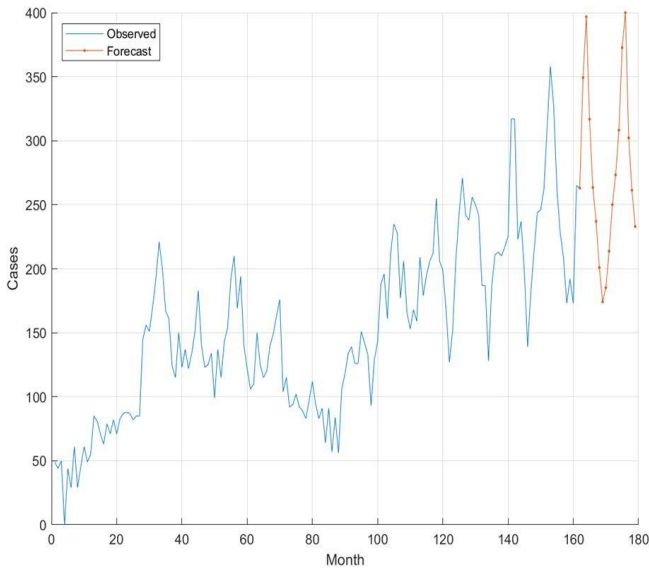
(b) Accidents per time zone

Fig. 7. Data training results.

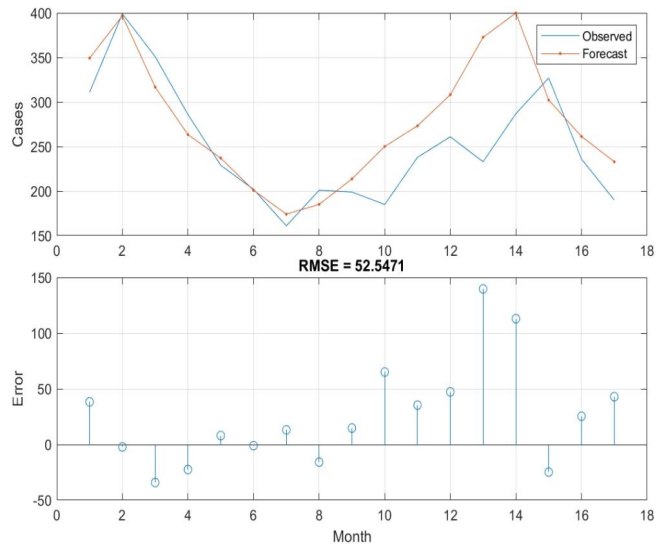
## 장단기 기억 신경망을 활용한 선박교통 해양사고 패턴 분석 및 예측

Fig. 7은 data training 결과를 시각화한 것으로 (a) 월별 데이터의 경우 hidden unit을 72, (b) 시간대별 데이터의 경우 54로 구성하였다. Fig. 7(a) 및 (b) 모두 solver는 adam, epoch은 150으로 설정하였으며, 발산을 방지하기 위하여 기울기 임계값을 1로 설정하고, initial learning rate를 0.005, 100 epoch 이후 학습률을 저하시키는 방법을 사용하였다. LSTM 모델의 parameter 조정에 따른 오차검증 및 성능평가 결과는 Fig. 8

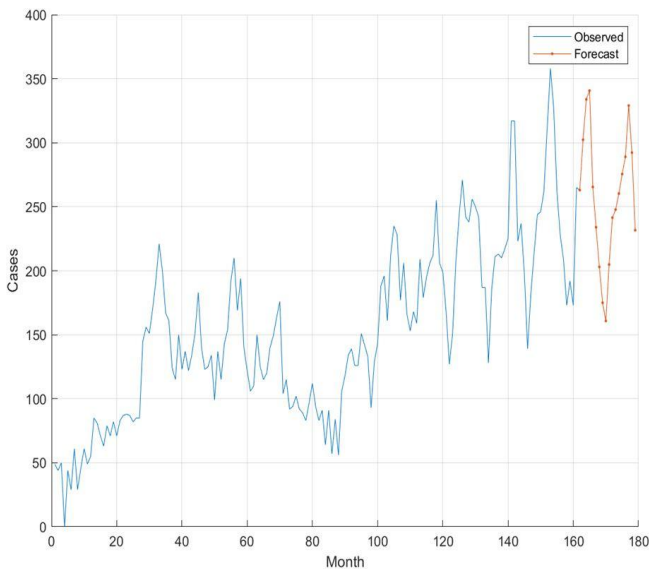
및 Fig. 9와 같다. 검증데이터를 통하여 모델의 성능을 검증한 결과 RMSE는 초기 신경망 모델에서 Fig. 8과 같이 월별 52.5471로 나타났으며, 관측값으로 신경망 모델을 업데이트한 결과 RMSE는 31.3680로 개선되었다. 한편, 시간대별 데이터의 경우 Fig. 9와 같이 초기 신경망 모델에서 RMSE는 126.5893으로 나타났으며, 관측값으로 신경망 모델을 업데이트한 결과 RMSE는 36.3967로 개선되었다.



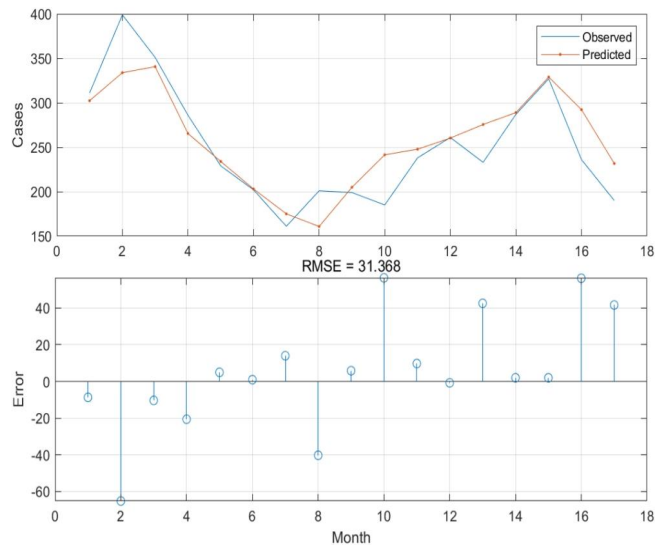
(a) Initial LSTM Network Model Prediction Results



(b) Initial LSTM Network Model Verification Results



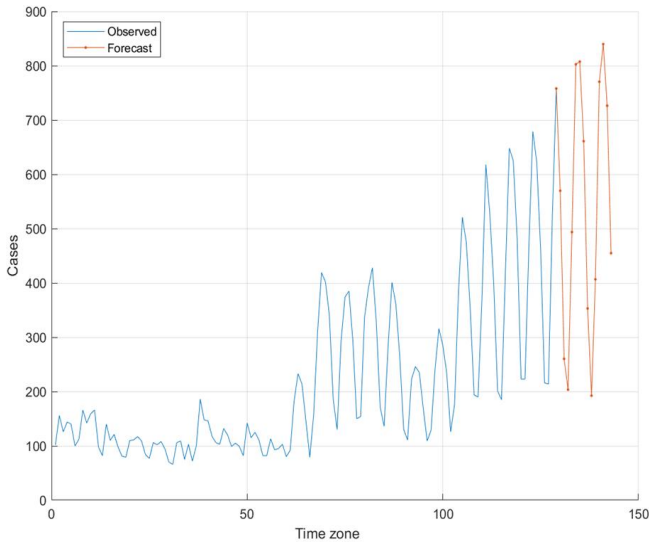
(c) Updated LSTM Network Model Prediction Results



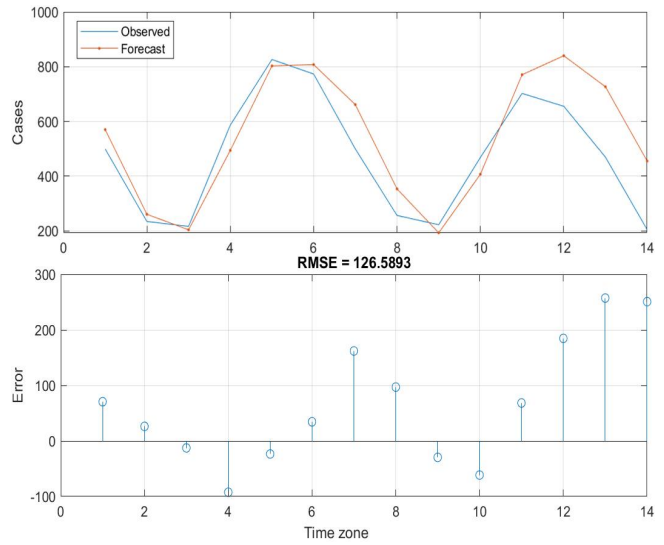
(d) Updated LSTM Network Model Verification Results

Fig. 8. Simulation results of accident per month.

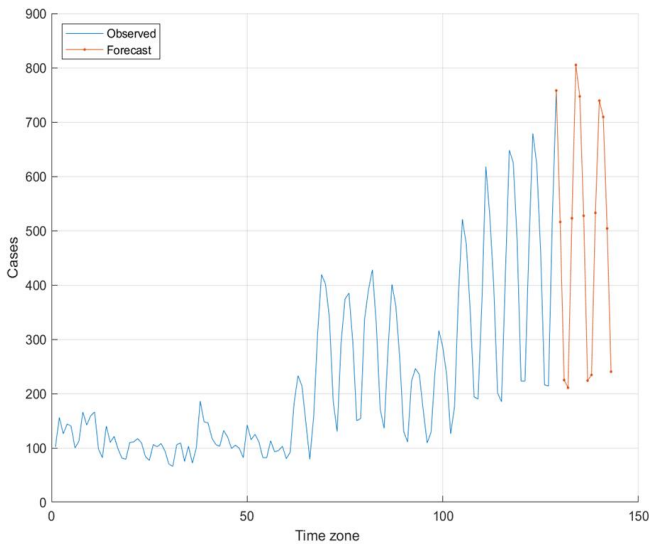




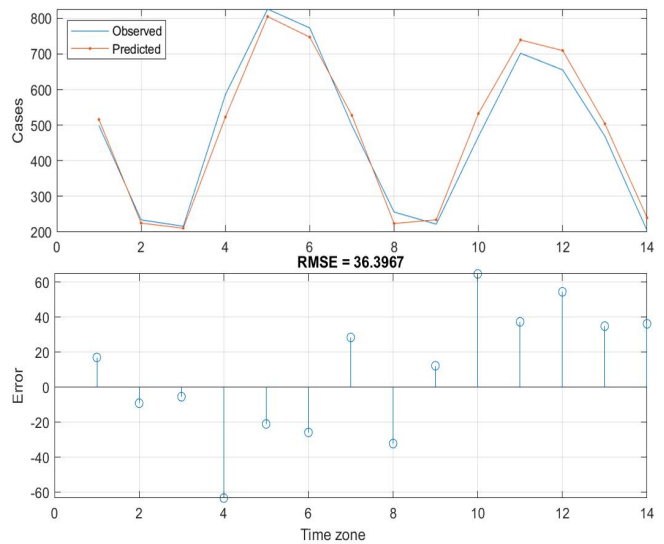
(a) Initial LSTM Network Model Prediction Results



(b) Initial LSTM Network Model Verification Results



(c) Updated LSTM Network Model Prediction Results



(d) Updated LSTM Network Model Verification Results

Fig. 9. Simulation results of accident per time zone.

#### 4. 결론

첨단 항해통신 장비의 도입과 VTS에서의 첨단 감시 체계 도입 등 해양사고 예방을 위한 민관의 노력에도 불구하고 해양사고는 지속적으로 발생하고 있다. 해양사고의 예방은 항해통신 장비의 기술적 발전과 운용인력의 전문화를 위한 교육훈련 제도의 개선, 해양사고의 주요 원인과 결과에 대한 분석 등 여러 측면에서의 연구·개발이 요구된다. 특히, 해양사고 분석과 관련하여 사고의 종류별 발생 패턴과 변화량을 예측함으로써 사고 원인과 결과에 대한 정량적 위험도를 제시하고 이를 예방하기 위한 기술적, 제도적 개선이 필

요하다. 본 연구에서는 선박교통과 관련된 해양사고 데이터를 수집하고 학습함으로써 해양사고 발생 패턴을 예측하고 이를 활용하기 위한 인공신경망 모델을 제안하였다. 기존의 전통적 분석 기법인 사례 분석 및 통계적 분석 방법의 제한점을 보완하고, 보다 정확한 시계열 데이터의 예측을 위하여 대표적인 인공지능 기법인 RNN기반의 LSTM을 활용한 선박교통 관련 해양사고 예측 모델을 구축하였다. 제안하는 모델의 성능 개선을 위하여 초기 신경망 모델을 관측값으로 업데이트하여 최종 모델 성능을 개선하도록 학습 모델의 알고리즘 구성하였으며, 검증데이터를 통하여 모델 성능을 확인한 결과 모델 업데이트를 통한 성능 개선을 확인하였다.

한편, 본 연구에서는 중앙해양안전심판원과 국가통계포털에서 제공하는 해양사고 종류별, 월별, 시간별 데이터를 활용하였으며, 사고 종류에 따라 월별 및 시간대별 통계 집계 이루어진 2007년 1월부터 2021년 12월까지 180개월 및 1998년부터 2021년까지 24년의 제한된 데이터를 활용하였다는 한계가 있다. 또한, 제한된 데이터셋의 절대량으로 인하여 해양사고의 세부 요인에 따른 분류 및 학습을 진행하지 못하였다는 한계가 있다. 향후 장기간의 데이터 수집과 세부 특징 데이터에 따른 해양사고 요인 분류를 통하여 보다 정확한 예측 모델을 구축하여야 할 것이다. 본 연구에서 제안한 LSTM 신경망 모델을 기반으로 다양한 해양사고의 특징 데이터를 학습하여 해양사고 발생 패턴을 예측할 수 있을 것으로 기대하며, 해양사고 특징별 예측 데이터베이스 구축과 정량적 사고 위험도 기반의 위험지도 개발 등을 통하여 클라우드 기반 차세대 VTS 시스템과 같은 클라우드 컴퓨팅 기반의 항해 종합 데이터 집적 감시 시스템과 항해 의사결정 지원 시스템의 연계 데이터로 활용할 수 있을 것으로 기대한다.

## 사 사

This research was a part of the project titled ‘Development of cloud-based next-generation VTS Integration platform’, funded by the Korea Coast Guard.

## References

- [1] Bae, S. W. and J. S. Yu(2018), Predicting the Real Estate Price Index Using Machine Learning Methods and Time Series Analysis Model, *Housing Studies Review*, Vol. 26, No. 1, pp. 107-133.
- [2] Hochreiter, S. and J. Schmidhuber(1997), Long Short-Term Memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780.
- [3] Hyun, S., B. K. Lee, G. W. Kim, and B. J. Chu(2009), A Comparative Study on Disaster Response Systems of Local Governments: Focusing on the Cases of Marine Pollution Accidents in Korea and Japan, *Korean Public Administration Review*, 43(3), pp. 273-306.
- [4] International Maritime Organization(2020), *International Convention for the Safety of Life at Sea 1974 Amended 2020*; International Maritime Organization: London, UK.
- [5] Jeong, J. Y. and C. Y. Jung(2012), Empirical Study on the Performance Analysis and Function of Jindo Coastal Vessel Traffic Service, *Journal of the Korean Society of Marine Environment & Safety*, Vol. 18, No. 4, pp. 308-315.
- [6] Jung, C. H.(2018), A Study on the Improvement of Safety by Accidents Analysis of Fishing Vessels, *Journal of Fisheries and Marine Sciences Education*, Vol. 30, No. 1, pp. 176-186.
- [7] Jung, C. H., H. Y. Kim, and C. H. Lee(2017), A Proposal of Marine Accident Prevention Policy through the Statistical Analysis, *Journal of Korean Maritime Police Science*, Vol. 7, No. 3, pp. 159-181.
- [8] Kim, J. H. and J. Y. Kim(2021), Covid19 trends predictions using time series data, *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 25, No. 7, pp. 884-889.
- [9] Kim, J. S.(2013), A Basic Study on the VTS Operator’s Minimum Safe Distance, *Journal of the Korean Society of Marine Environment & Safety*, Vol. 19, No. 5, pp. 476-482.
- [10] Kim, J. S.(2014), A study on Cognitive Work Analysis of VTS operation and effective utilization, M.S. thesis, Mokpo maritime university.
- [11] Korean Maritime Safety Tribunal(2022), Status of Marine Accidents, <https://www.kmst.go.kr/web/stcAnnualReport.do?menuIdx=126> (Accessed July 2022).
- [12] Korean Statistical Information Service(2022), Annual Marine Accident Statistics, [https://kosis.kr/statHtml/statHtml.do?orgId=146&tblId=DT\\_MLTM\\_1348](https://kosis.kr/statHtml/statHtml.do?orgId=146&tblId=DT_MLTM_1348) (Accessed July 2022).
- [13] Lee, J. and J. Song(2021), An Empirical Study on Prediction of the Art Price using Multivariate Long Short Term Memory Recurrent Neural Network Deep Learning Model, *The Journal of the Korea Contents Association*, Vol. 21, No. 6, pp. 552-560.
- [14] Lee, L. N. and J. S. Kim(2022), Development of Priority Index for Intelligent Vessel Traffic Monitoring System in Vessel Traffic Service Areas, *Applied Sciences*, Vol. 12, No. 8, pp. 3807.
- [15] Lee, L. N., J. S. Kim, H. H. Lee, J. S. Lee, and H. Namgung(2022), Analysis of User Requirements for Development of Vessel Traffic Services Cloud System, *Journal of the Korean Society of Marine Environment & Safety*, Vol. 28, No. 2, pp. 314-323.
- [16] Lee, M. K., Y. S. Park, E. B. Lee, and S. J. Na(2019), A Study on the Improvement plan of Marine Accident Investigation Method, *Journal of Korean Maritime Police Science*, Vol. 9, No. 1, pp. 111-129.
- [17] Olah, C.(2015), Understanding LSTM Networks, <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (Accessed July 2022).

- [18] Park, B. S. and Y. S. Ahn(2007), Statistical Analysis of Marine Accidents by ANOVA, Journal of the Korean Society of Marine Environment & Safety, Vol. 13, No. 3, pp. 191-198.
- [19] Roh, H. R.(2014), Cases Analysis of Maritime Accidents and Countermeasure of Korea Coast Guard: Focus on the Passenger Ship Accidents, Korean Journal of Public Safety and Criminal Justice, Vol. 23, No. 4, pp. 127-160.
- [20] Roh, J. Y. and S. H. Bae(2021), Forecasting of Traffic Accident Occurrence Pattern Using LSTM, The Journal of The Korea Institute of Intelligent Transportation Systems, Vol. 20, No. 3, pp. 59-73.
- [21] Roh, Y. J. and S. H. Bae(2021), Forecasting of Traffic Accident Occurrence Pattern Using LSTM, The journal of the Korea Institute of Intelligent Transportation Systems, Vol. 20, No. 3, pp. 59-73.
- [22] Shin, D. H. and S. H. Ji(2020), A Critical Review of the Act on Vessel Traffic Services, Journal of the Korean Society of Marine Environment & Safety, Vol. 26, No. 4, pp. 336-345.
- [23] Shin, D. H., K. H. Choi, and C. B. Kim(2017), Deep Learning Model for Prediction Rate Improvement of Stock Price Using RNN and LSTM, The Journal of Korean Institute of Information Technology, Vol. 15, No. 10, pp. 9-16.
- [24] Shin, D. S., S. H. Im, C. W. Choi, J. O. Lee, and H. C. Park(2021). Marine Accident Prediction Model Based on Artificial Neural Network. Transportation Technology and Policy, Vol. 18, No. 2, pp. 43-50.
- [25] Sun, R., J. Yoo, J. Bang, J. H. Song, and B. M. So(2021), LSTM Model Design for Short-Term Prediction of Wind Power Generation Using Actual Data from a Wind Condition Measuring Instrument, The Journal of Next-generation Convergence Technology Association, Vol. 5, No. 6, pp. 1018-1026.
- [26] Youn J. and T. J. Lee(2022), LSTM-based Fire and Odor Prediction Model for Edge System, KIPS Transactions on Computer and Communication Systems, Vol. 11, No. 2, pp. 67-72.

---

Received : 2022. 07. 19.

Revised : 2022. 08. 11. (1st)

: 2022. 08. 25. (2nd)

Accepted : 2022. 08. 29.