

LNG 및 LSFO 연료 사용 선박의 CO₂ 배출량 예측을 위한 LightGBM 모델 성능 분석

이상진, 김동건, 최재영, 한명관*

한국조선해양기자재연구원, 한국조선해양기자재연구원, 한국조선해양기자재연구원,
*한국조선해양기자재연구원

sjlee@komeri.re.kr, dgkim@komeri.re.kr, jychoi@komeri.re.kr, *mghan@komeri.re.kr

LightGBM-Based Performance Evaluation for CO₂ Emission Prediction in LNG and LSFO-Fueled Ships

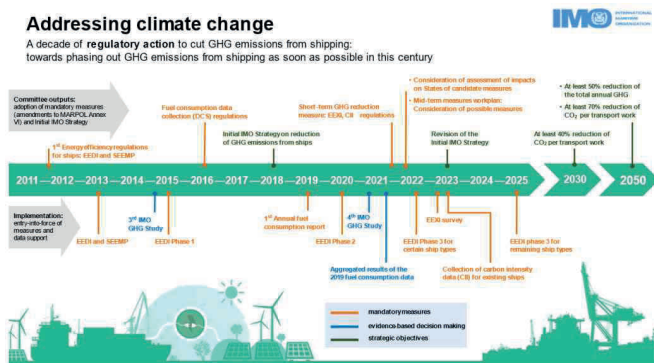
Lee Sang Jin, Kim Dong Geon, Choi Jae Young, Han Myeong Gwan*

Korea Marine Equipment Research Institute, Korea Marine Equipment Research Institute,
Korea Marine Equipment Research Institute, *Korea Marine Equipment Research Institute

요약

본 논문은 LNG 및 LSFO 연료를 사용하는 선박의 에너지 소비와 관련된 이산화탄소(CO₂) 배출량을 예측하기 위해 개발된 머신러닝 기반 모델의 성능을 정량적으로 분석하였다. 예측 모델은 실운항 로그 데이터를 기반으로 LightGBM 알고리즘을 적용하였으며, LNG 및 디젤(LSFO) 연료에 대해 독립적으로 학습되었다. 성능 평가지표로는 결정계수(R²), 평균절대오차(MAE), 평균제곱근오차(RMSE), 평균절대백분율오차(MAPE)를 사용하였다. 테스트 결과 LNG 기반 모델의 평균 MAPE는 2.07%, LSFO 기반 모델은 3.46%로, IMO 환경 규제를 고려한 산업 적용에 충분한 신뢰도를 확보하였다. 본 연구는 선박의 탄소 배출량 예측에 있어 실환경 기반 모델 검증 사례로서의 의미를 가지며, 자율운항선박, 연료전환, 탄소세 대응 등 다양한 응용 가능성을 시사한다.

I. 서론



[그림 1] 선박에서 발생하는 GHG 배출량을 줄이기 위한 IMO 조치

국제해사기구(IMO)의 EEXI(Existing Energy Efficiency Index) 및 CII(Carbon Intensity Indicator) 규제는 해운업계에 있어 탄소배출 감축을 위한 실질적 대응책 마련을 강력히 요구하고 있다. 특히 IMO는 지난 10여 년간 단계적인 온실가스 감축 조치를 추진해 왔으며, 2030년까지 최소 40%, 2050년까지 70% 이상의 탄소배출 감축을 목표로 다양한 규제를 도입하고 있다. 그림 1은 IMO의 GHG 규제 로드맵을 시계열로 정리한 것으로, 단기·중기·장기 전략 목표를 통해 점진적인 감축을 추진하고 있음을 보여준다.

이러한 규제 흐름 속에서 이중연료(Dual-Fuel) 추진 시스템을 적용한 친환경 선박은 연료 다변화 및 에너지 효율화를 실현할 수 있는 핵심 수단으로 주목받고 있다. 그러나 기존의 단일 연료 기반 탄소배출 산정 방식은 이중연료 선박의 연료 운용 특성을 충분히 반영하기 어려우며, 이에 따라 정확한 배출량 산정을 위한 예측 기술 확보가 요구된다.

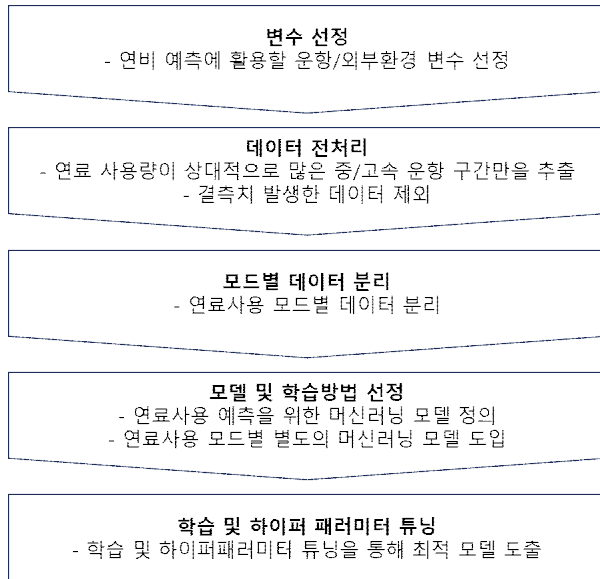
최근에는 빅데이터 기반 머신러닝 기법이 에너지 및 환경 분야의 예측 정밀도를 높이기 위한 유효한 수단으로 각광받고 있다. 특히 선박 운항 로그, 기상 환경 정보, 에너지 운용 이력 등의 복합 데이터를 학습하여 이산화탄소(CO₂) 배출량을 예측하는 알고리즘 개발이 활발히 진행되고 있으며, 해당 기술은 실시간 운항 최적화 및 탄소세 대응 전략 수립에 활용 가능성이 크다.

본 연구는 LNG 및 LSFO 연료 사용 선박의 실운항 데이터를 기반으로 머신러닝 기반 예측 모델을 개발하고, 예측 정확도를 다양한 정량적 지표(R², MAE, RMSE, MAPE 등)를 활용해 평가함으로써, 해당 기술의 산업 적용 가능성과 신뢰도를 검증하는 데 목적이 있다. 이를 통해 국내 친환경 선박 기술의 경쟁력을 제고하고, 국제 환경 규제에 선제적으로 대응할 수 있는 기술 기반을 확보하고자 한다.

II. 본론

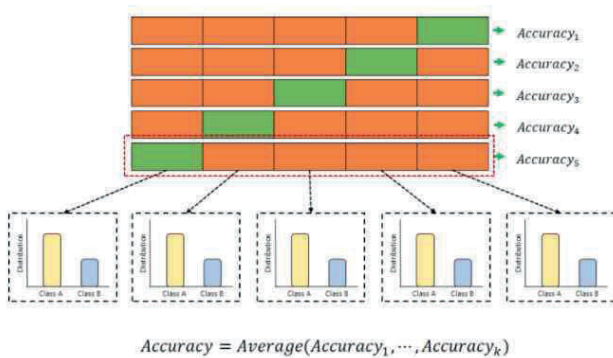
본 연구는 국제해사기구(IMO)의 온실가스 감축 규제에 대응하여, 이중연료(Dual-Fuel) 엔진 기반 선박의 데이터를 바탕으로 탄소배출량을 정밀하게 예측하는 머신러닝 기반 알고리즘을 개발하고 그 성능을 정량적으로 평가하는 데 목적이 있다. 특히 연료 소모량과 이산화탄소 배출량 간의 상관관계를 모델링하고, Carbon Factor를 적용하여 배출량을 환산함으로써 실선 데이터를 활용한 실제 적용 가능성을 검증하였다.

사용된 데이터는 2024년 2월 10일부터 7월 31일까지 약 6개월간 실선에서 10분 간격으로 수집된 로그 데이터로 총 23,815행, 1,055열로 구성되며 전체 셀 수는 약 2,446만 건에 달한다. 데이터 전처리 단계에서는 센서 오류 제거, 연안 및 정박 구간 필터링, 결측치 보간, 순간 급변 값 보정, 파생 변수 생성 등의 작업을 수행하였다.



[그림 2] 머신러닝 기반 연료소모량 예측모델 개발 절차

모델 개발 절차는 그림 2에 제시된 바와 같이 변수 선정부터 모델 학습 및 최적화까지 일련의 단계를 포함한다. 특히 연료 사용 모드(LNG, LSFO)에 따라 데이터를 분리하고 각각에 대해 독립적인 예측 모델을 설계하였다.



[그림 3] 계층별 k-겹 교차 검증(Stratified k-fold cross validation)

훈련 데이터는 8:2 비율로 분할한 후, 모델 성능의 일반화 가능성을 확보하기 위하여 5-Fold 교차검증을 수행하였다. 각 Fold에서는 1개의 검증셋과 4개의 훈련셋으로 반복 학습을 수행하였으며, 그 구조는 그림 3과 같다.

예측 모델로는 LightGBM(Light Gradient Boosting Machine)을 사용하였으며, 모델 성능을 최적화하기 위해 Optuna 라이브러리를 통해 하이퍼 파라미터 튜닝을 수행하였다. 주요 파라미터의 탐색 범위는 표 1에 정리하였다.

[표1] 주요 파라미터의 탐색 범위

지표명	설명	범위
lambda_l1	L1 규제 항	0.0000001~10.0
lambda_l2	L2 규제 항	0.0000001~10.0
num_leaves	트리의 잎 노드 수	2~256
feature_fraction	학습 반복 시 랜덤으로 선택되는 피쳐 비율	0.4~1.0
bagging_fraction	학습 데이터 샘플링 비율	0.4~1.0
min_child_samples	분할 노드 최소 샘플 수	5~100

표 1에 정리된 주요 하이퍼파라미터 탐색 범위를 기반으로 최적화된 LightGBM 모델을 이용하여 LNG 및 LSFO 연료별 CO₂ 배출 예측을 수행하였다. 모델의 테스트 결과는 표 2와 표 3에 요약되어 있다.

[표2] LNG 사용량 기반 CO₂ 배출모델 성능 평가 결과

Metric	Score(%)			
	Iter1	Iter2	Iter3	Mean
R2	0.952	0.942	0.944	0.946
MAE	0.040	0.053	0.049	0.047
RMSE	0.091	0.110	0.102	0.101
MAPE	1.649	2.371	2.191	2.070

표 2는 LNG 연료를 기반으로 한 CO₂ 배출 예측 모델의 3회 반복 테스트 결과를 나타낸다. 결정계수(R²)는 평균 0.946으로 매우 높은 적합도를 나타냈으며, 평균절대백분율오차(MAPE)는 2.07%로, IMO 대응 기준(15% 이하)을 크게 만족하였다.

[표3] LSFO 사용량 기반 CO₂ 배출모델 성능 평가 결과

Metric	Score(%)			
	Iter1	Iter2	Iter3	Mean
R2	0.989	0.992	0.995	0.991
MAE	0.068	0.061	0.053	0.060
RMSE	0.144	0.119	0.100	0.121
MAPE	2.979	2.141	5.253	3.458

표 3은 LSFO 연료를 기반으로 한 CO₂ 배출 예측 모델의 성능을 보여준다. LSFO 모델은 평균 R² 0.991, MAPE 3.46%를 기록하여, LNG 모델 대비 다소 높은 MAPE를 보였으나 여전히 국제 기준을 충분히 만족하는 수준의 예측 정확도를 확보하였다. 특히 LSFO 모델의 경우 반복 학습 과정에서 Iter 3회차의 MAPE가 일시적으로 증가하는 경향이 나타났으나, 평균 성능은 안정적으로 유지되었다.

III. 결론

본 연구는 LNG 및 LSFO 연료를 사용하는 이중연료 추진 선박의 실운항 데이터를 기반으로 이산화탄소 배출량 예측 모델을 개발하고, 예측 정확도를 분석하였다. LightGBM 기반의 머신러닝 모델은 평균 MAPE 2~3% 수준의 성능을 보이며, 실환경 적용에 필요한 신뢰도 기준을 충족하였다. LNG 및 LSFO 연료 각각에 대해 독립적인 모델을 설계함으로써 연료 유형별 최적화를 달성하였고, 이는 탄소세 대응 및 운항 전략 수립을 지원하는 데 유용할 것으로 판단된다. 또한 실데이터 기반 반복 검증을 통해 산업적 적용 가능성을 검증한 이번 연구는 친환경 선박 기술 고도화와 국제 환경 규제 대응력 제고에 기여를 할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부와 한국산업기술진흥원의 “지역혁신클러스터 육성(R&D)(P0025328)”사업의 지원을 받아 수행된 연구결과임.

참 고 문 헌

- [1] Ke, G., Meng, Q., Finley, T., et al. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," Advances in Neural Information Processing Systems (NeurIPS), 2017.
- [2] Soleymani, A., Sharifi, S.M.H., Edalat, P., "Linear Modelling of Marine Vessels Fuel Consumption for Ration of Subsidized Fuel," International Journal of Maritime Technology, vol. 10, 2018.
- [3] 하이브리드 선박 추진을 위한 다중 전력 통합 관리시스템 개발 사업계획서