

Global Report

아시아 컨테이너 선박의 리니어 운송경로 선택을 위한 딥 러닝 기반 연료소비모델

1. 개요
2. 2위상(two-phase) 데이터 수집 및 필터링
3. 딥 러닝(Deep learning) 및 ANN 아키텍처
4. 딥 러닝 ANN의 예측결과_Hyper 파라미터 선택
5. 라이너 운송경로 최적화를 위한 ATSP 알고리즘
6. 결론

1. 개요

선박 운영비용(operating costs)은 매우 다양한 구성요소(승무원의 인건비, 수리 및 유지보수 비용, 해양 보험 및 P&I(Protection and Indemnity : 보호 및 배상), 창고보관 비용, 연료비용 등)로 구성된다. 특히 이 중 60% 이상을 차지하는 연료비용의 경우, 불안정 적이고 상승하는 연료가격은 해운회사에 심각한 경영난을 초래하고 있다[1][2]. 이러한 해상운송 연비를 최소화하기 위해서는 해운회사의 라이너 선박 운용 시 가장 효과적인 기본 솔루션인 최적의 경로선택이 무엇보다 중요하다. 이를 위해 실제 운송경로에 대한 데이터를 학습하여 고도화된 방법으로 연료비용을 예측하여 연비예측 오차를 최소화할 수 있는 방안이 필요하다. 그간 해상운송 경로최적화를 위한 관련 연구로 2002년 Kim,J.U 등이 제시한 “다중기간 다중정지 운송계획 문제에 대한 추론 알고리즘(Heuristic algorithms for a multi-period multi-stop transportation planning problem)”가 있다[3]. 그러나 이 연구는 가장 빠르고 효과적인 방법 중 하나인 ATSP 알고리즘과 하이브리드 기계학습(hybrid machine-learning) 방법을 적용한 연구와는 거리가 있다.

이 연구에서는 2위상 방법(two-phase method)을 적용한 딥 러닝(deep-learning) 기반의 ANN(Artificial Neural Network : 인공신경망)을 생성하여 라이너 운송경로에 대한 ATSP(Asymmetric Traveling Salesman Problem) 알고리즘을 통해 해상운송 연비를 최소화할 수 있는 연료소비 추정 모델을 제시한다. 제안 모델은 실제 상황을 반영할 수 있으며, 보다 더 다양한 관측에 적용될 수도 있다.

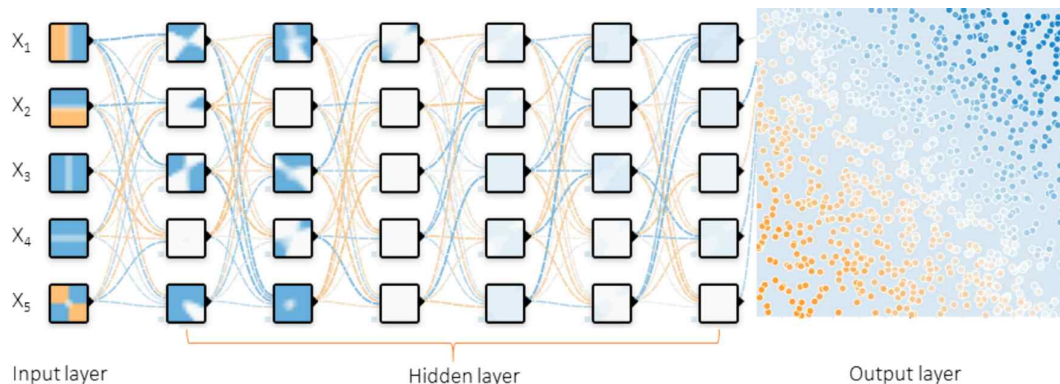
2. 2위상(two-phase) 데이터 수집 및 필터링

이 연구에서 제시한 2위상 방법을 적용한 딥 러닝 기반의 ANN을 생성하여 라이너 운송경로에 대한 ATSP 알고리즘을 적용한 해상운송 연비 최소화 추정 모델을 개발하기 위해 2017.02~2019.01월까지 2년간 세계 유수 해운회사의 베트남 지사로부터 항해 히스토리 데이터를 제공받았다. 데이터집합 정보에는 선박의 이름, 위치(경도와 위도), 출발 및 도착 항구, 출발 및 도착시간, 운항거리(Nm), 출발-도착 항구 간 운항속도(노트), 항해시간(시간), 연료소비량(Mt) 등이다. 추가 정보(선박의 선적용량(TEU), 평균 풍속(Beaufort scale에로 변환한 속도), 바람의 방향 등)는 보조 데이터수집 단계를 통해 수집되었다. 특히 선박의 용량은 2019년 기준 "Vesseltracking"에서 수집하였다. CMNE(Copernicus Marine Environment : 코페르니쿠스 해양 환경) 모니터링 서비스를 통해 데이터 패키지에서 바람(풍향/풍속)과 관련된 기상조건에 대한 세부정보를 수집하였다. 모든 데이터집합의 운용영역은 아시아 지역으로 국한하였으며, 항해 중 바람의 방향에 따른 정보를 수치데이터화 하여 향후 분석 수행시 활용토록 하였다. 끝으로 모든 수집데이터에 포함된 노이즈를 제거하여 필터링하였다. 아울러 다양한 데이터집합 정보의 근사값을 얻기 위해 1969년 Grubbs가 제시한 Z-score를 적용하였으며[4], 1930년 Mahalanobis가 제시한 디노이징 기술(denoising techniques)을 적용하여 데이터집합 정보의 근사값을 추출하였다[5]. 10,853개의 원본 데이터집합을 필터링하여 105개의 노이즈와, 비정상적인 287개의 데이터를 제거하고 10,461개의 데이터집합 정보의 근사값을 추가로 분석하였다.

3. 딥 러닝(Deep learning) 및 ANN 아키텍처

이론적으로 신경망은 동물의 신경(neuron)을 기반으로 하는 간단한 처리요소 또는 노드 간 상호 연결된 어셈블리이다. 네트워크의 처리기능을 표현하는 노드 또는 가중치 간 연결능력은 패턴 집합에 대한 일련의 과거의 트레이닝 패턴에 대한 적응학습 프로세스에 의해 저장되는 형태로 나타난다. 데이터 모델링을 위한 ANN 모델의 사용범위는 재래적 방법인 표준 비선형 회귀분류, 또는 클러스터링 기술에 비해 선호되는 기술이다. 따라서 이들은 예측과 연관된 문제에서 주로 이용된다. 신경망에는 여러 개의 뉴런을 포함하는 각 계층이 네트워크로 서로 연결된 레이어 집합을 포함한다. 여기에는 하나의 입력 및 출력 레이어, 여러 개의 숨겨진 레이어가 있다. 모든 입출력 레이어 노드 수는 각각의 입출력 변수 수와

동일하다. 입력 계층에서 시작하여 각 레이어의 노드는 ANN 프로그래밍이 각 학습 시간을 설정하는 서로 다른 가중치로 다음 레이어의 노드에 연결되도록 설계되어 있다. 이러한 프로세스는 출력 값이 관측 값에 비해 가장 낮은 손실을 기록할 때까지 지속적으로 반복된다. ANN의 시각화 사례를 [그림 1]에 나타낸다.



* 자료 : Linh Bui-Duy, Ngoc Vu-Thi-Minh, “Utilization of a deep learning-based fuel consumption model inchoosing a liner shipping route for container ships in Asia”, The Asian Journal of Shipping and Logistics xxx (2020) xxx-xxx, p.4.

[그림 1] ANN의 시각화 사례

숨겨진 단일 레이어 피드포워드 ANN은 얇은 또는 깊은 MLP(Multi-Layer Perceptron) ANN과 같은 데이터 마이닝에 적합한 여러 유형의 ANN이 존재한다. 특히 MLP는 데이터가 단일 ANN보다 더 많은 숨겨진 레이어를 통과하기 때문에 단일 계층 피드포워드 형태의 ANN보다 더 정확한 결과를 제공한다. 10개의 숨겨진 레이어가 있는 깊은 MLP-ANN을 선택하여 시행착오와 정확도-수렴시간을 기반으로 ANN 유형에 대한 테스트를 설정하였다. “임의의 포리스트 기술(random forest technique)”로 수행된 분류결과에 기초하여 5개의 변수(평균속도/수송용량/운송시간/풍속/풍향)를 포함하는 입력노드로 설정하였다. 출력 레이어에 있는 1개의 노드는 연료소비를 추정하는 노드이다. 이러한 입출력 차이로 인한 변수를 최소/최대 특성배율을 사용하여 0~1 사이의 모든 입출력 값을 정규화하고 이들을 상호 비교할 수 있도록 (식 3)과 같이 고유한 축척을 만들었다.

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \dots\dots\dots (\text{식 1})$$

여기서, X_n 은 변수 X 의 정규화 된 값(실제 값은 변수), $X_{\max}$ 과 $X_{\min}$ 은 각각 X 의 최대값과 최소값을 나타낸다. 이 연구에서 깊은 ANN 아키텍처는 10개의 숨겨진 레이어로 구성되며, ANN의 추가 파라미터인 바이어스 x_0 은 뉴런에 대한 입력의 가중합계와 함께 출력을 조정하는 데 사용된다. N 개의 입력을 가진 뉴런의 경우 $\{w_i\}$ 는 연결의 가중치, $\{x_i\}$ 는 뉴런에 대한 입출력 값을 의미한다. 뉴런은 연결되는 또 다른 뉴런으로부터 입력 값을 수신하고 뉴런의 활성화한다[6].

4. 딥 러닝 ANN의 예측결과_Hyper 파라미터 선택

Hyper-parameter의 선택은 모델의 예측결과에 매우 중요한 영향을 미친다. 주요 hyper-parameter에는 임시값 수, CLR(Cyclical Learning Rate), 배치 크기 및 각 숨겨진 레이어의 노드 수가 포함된다. 너무 작은 CLR은 오버 피팅의 위험을 증가시킬 수 있다. 그간 여러 연구에서 시행착오법에 따른 특정 숨겨진 노드와 CLR은 100개 노드당 0.001 CLR로 설정하였다[7]. 숨겨진 레이어의 수를 테스트하기 위한 테스트는 1~15개로 적용하였다. 이를 바탕으로 100, 200, 500개의 레이어를 대상으로 프로그래밍 하

여 3단계에서 200개의 레이어를 선택하였다. CLR의 경우, 작은 값으로 시작하면 네트워크가 수렴되기 시작하면서 학습속도가 증가함에 따라 테스트/유효성검사 손실이 증가할 수 있어 수렴이 발생할 때까지 정확도를 점차 증가시키면서 hyper-parameter를 선택하였다. 시행착오에 따른 hyper-parameter 조정 프로세스를 <표 1>에 나타낸다. CLR은 0.001, 0.005 수준으로 테스트되었으며, 유효성검사 결과는 0.005의 CLR에서 상당히 수렴되었음을 알 수 있다.

<표 1> 시행착오법에 따른 hyper-parameter 조정 프로세스

단계	숨겨진 레이어	숨겨진 노드	에포크	CLR	실행시간 [초]	MAPE 학습	MAPE 검증	MAPE 테스트	MASE 테스트	MASE 유효성검사	MASE 테스트
	1				23.56	11.4	12.1	11.5	3.14	3.22	10
	2				39.16	11	11.8	11.2	3.01	2.89	3.02
	4				60.24	9.2	9.5	9.5	2.56	2.78	2.59
	6				100.98	8.5	8.8	8.7	1.14	1.37	1.2
1	100	100	0.001	0.001	189.53	7.9	8.5	8.0	1.08	1.15	1.13
2	10	50	100	0.001	256.14	7.1	8.1	7.7	0.95	1.04	0.99
3	10	150	100	0.001	323.15	6.5	6.9	6.6	0.74	0.86	0.8
4	10	150	200	0.005	488.62	6.1	6.6	6.2	0.69	0.80	0.77

* 자료 : Linh Bui-Duy, Ngoc Vu-Thi-Minh, “Utilization of a deep learning-based fuel consumption model inchoosing a liner shipping route for container ships in Asia”, The Asian Journal of Shipping and Logistics xxx (2020) xxx-xxx, p.4.

<표 1>의 조정 프로세스를 토대로 선택한 hyper-parameter들의 정규화된 값을 <표 2>에 나타낸다.

<표 2> Selected hyper-parameters for deep ANN

Hyper-parameters	Value
숨겨진 레이어 수	10
에포크 수	200
CLR	0.005
배치(Batch) 크기	256
활성화 기능	ReLU
학습 알고리즘	BP, dropout
성능 기준	MAPE; MASE

* 자료 : Linh Bui-Duy, Ngoc Vu-Thi-Minh, “Utilization of a deep learning-based fuel consumption model inchoosing a liner shipping route for container ships in Asia”, The Asian Journal of Shipping and Logistics xxx (2020) xxx-xxx, p.5.

5. 라이너 운송경로 최적화를 위한 ATSP 알고리즘

해운회사의 실질적인 라이너 운송문제는 왕복 라이너 운송노선을 운영하는 것에서 기인한다. 이러한 문제해결을 위해서는 선박의 선적용량에 대한 정보, 운영 경로상에 있는 모든 항구의 이름, 각 항구 간 거리 등에 대한 정보가 필수적으로 요구된다. 예를 들면, “이 배는 경로상의 모든 항구를 한 번 경유한 다음 첫 번째 출발 항구로 돌아가야 합니다. 경로의 항만 도착순서는 연비를 최소화할 수 있는 운영비용 규칙에 따라 정의됩니다.” 등이다. 이는 이 연구에서 가정할 수 있으며, 실제에서는 다른 추가적인 비용에 따라 달라질 수도 있다고 가정할 수 있다. 3장에서 설명한 딥 러닝 ANN 모델을 활용하여 예측된 연료소비량을 도출하려면 최적의 운송경로를 정의하여 이 문제를 해결할 수 있다. 간단히 요약하면 다음과 같다:

- 첫 번째 단계는 딥 러닝 ANN 및 TSP 알고리즘에 필요한 모든 파라미터를 준비하는 것이다. 이를 위해 CME(Copernicus Marine Environment) 모니터링 서비스에서 기상과 관련된 다양한 조건들을 획득해야 한다. 이를 토대로 선박의 선적용량, 항만 간 거리, 운행속도 및 평균속도 등을 결정해야 한다.
- 두 번째 단계는 최적의 ANN 모델을 사용하여 1단계에서 정의한 파라미터를 이용하여 연료소비량을 예측해야 한다.

이러한 단계를 거쳐 항해거리와 이에 따른 항해시간을 계산한다. 이를 토대로 각 출발 항구에서 연료를 구매한다고 가정하고 연비(USD/ton)는 (식2)와 같이 대칭거리 행렬로 계산할 수 있다.

$$[D_{ij}]_{n \times n} = \begin{bmatrix} 0 & D_{12} & \dots & D_{1n} \\ D_{21} & 0 & \dots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \dots & 0 \end{bmatrix} \dots\dots\dots (\text{식 2})$$

여기서, n은 운송경로상의 총 항구 수, D_{ij} 는 항구 i에서 항구 j까지의 거리이다.

6. 결론

이 연구는 총 연료비용을 최소화하기 위해 최적의 운송경로를 설계하는 문제를 다루었다. 이를 위해 하이브리드 기계학습 기술을 이용하여 아시아 지역의 상선에 대한 운송정보를 학습하여 컨테이너 선박의 연료소비량을 예측하는 데 활용하였다. 아울러 딥 머신 러닝 방법을 적용한 운송경로에 따른 연료소비 모델을 추정할 수 있는 ATSP(Asymmetric Traveling Salesman Problem) 알고리즘을 사용하여 연료비용을 최소화하기 위한 최적의 운송경로 선택 플랫폼을 제시하였다. 라이너 운송경로의 가정으로 예측 결과를 적용하고 이를 ATSP 알고리즘과 결합하여 연비를 최소화할 수 있는 최적의 경로를 설계하였다. 이는 연비가 선박의 총 운영비용의 60% 이상을 차지하기 때문에 실질적인 의미가 있다. 이 연구결과는 국내 컨테이너항의 최적 운송경로 설정을 위한 정책에 벤치마킹 자료로 활용할 수 있을 것이다.

[참고문헌]

- [1] Golias, M. M., Saharidis, G. K., Boile, M., Theofanis, S., & Ierapetritou, M. G. (2009). The berth allocation problem: Optimizing vessel arrival time. Maritime Economics and Logistics. <<http://dx.doi.org/10.1057/mel.2009.12Grubbs>>
- [2] Bal Bes, ikc, i, E., Arslan, O., Turan, O., & Ölc, er, A. I. (2016). An artificial neural network based decision support system for energy efficient ship operations. Computers and Operations Research. <<http://dx.doi.org/10.1016/j.cor.2015.04.004>>
- [3] Kim, J. U., Kim, Y. D., & Shim, S. O. (2002). Heuristic algorithms for a multi-period multi-stop transportation planning problem. Journal of the Operational Research Society. <<http://dx.doi.org/10.1057/palgrave.jors.2601386>>
- [4] Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. Technometrics. <<http://dx.doi.org/10.1080/00401706.1969.10490657>>

- [5] Mahalanobis, P. C. (1930). On tests and measures of group divergence. Journal and Proceedings, Asiatic Society of Bengal.
- [6] Parkes, A. I., Sobey, A. J., & Hudson, D. A. (2018). Physics-based shaft power prediction for large merchant ships using neural networks. Ocean Engineering.
<<http://dx.doi.org/10.1016/j.oceaneng.2018.07.060>>
- [7] Choi, B., Lee, J. H., & Kim, D. H. (2008). Solving local minima problem with large number of hidden nodes on two-layered feed-forward artificial neural networks. Neurocomputing.
<<http://dx.doi.org/10.1016/j.neucom.2008.04.004>>

<원문제목> Utilization of a deep learning-based fuel consumption model in choosing a liner shipping route for container ships in Asia

<원문출처> The Asian Journal of Shipping and Logistics xxx (2020) xxx-xxx