

충남대학교 컴퓨터융합학부

종합설계1 프로젝트 최종보고서

프로젝트명:

해상교통 빅데이터 분석 및 머신러닝을
활용한 선박 이동 예측 모델 개발

작성일: 2025. 06. 10.

참여인원 명단:

지도교수	김경섭 교수님 (서명)
산업체 멘토 (이름/소속)	이은규/삼성중공업
학생1	김연범
학생2	이한을
학생3	이동원(종합설계 기수강자)

1. 연구개발의 배경 및 필요성

1. 연구개발 동기

해상교통은 전 세계 물류 운송량의 90% 이상을 차지하는 핵심 수단으로 그 안정성과 효율성 확보는 국가 경제와 글로벌 무역의 지속가능성을 위해 필수적이다. 최근 해운 산업 전반에 걸쳐 디지털 전환이 가속화 되면서, AIS(Automatic Identification System) 데이터를 활용한 해상교통 빅데이터 분석이 활발히 이루어 지고 있다. 그러나 기존 선박 운항 및 경로 예측 시스템은 주로 정적 모델이나 단순 규칙 기반 접근법에 의존하고 있어 바다 위 실시간 교통 상황의 복잡성과 변동성을 충분히 반영하지 못한다. 이에 본 연구는 이러한 한계를 극복하기 위해 대규모 AIS 데이터를 기반으로 한 머신러닝 기법을 도입하여 실시간 선박 이동 예측 및 충돌 방지 모델을 개발하고자 한다. 이를 통해 기존 시스템의 예측 정확도를 개선하고 동적 해상 환경에서도 신속한 의사결정 지원이 가능하도록 하는 것이 연구의 핵심 동기이다.

2. 연구개발의 중요성

선박 이동경로를 사전에 예측하는 것은 해양 사고의 주요 원인인 충돌을 미리 감지해 인명·재산 피해를 줄이는 가장 효과적인 수단이다. AIS 빅데이터는 세계 물류의 90 % 이상을 담당하는 해상교통의 움직임을 고해상도로 기록하므로 예측 모델이 위험 수역과 혼잡 패턴을 조기에 식별할 정보 기반을 제공한다. 선박의 미래경로를 정확하게 예측함으로써 관제 시스템은 충돌 위험 선박을 실시간 경고하고 회피 지시를 내려 해양 안전성을 획기적으로 높일 수 있다. 더 나아가 예상 도착 시간(ETA)과 혼잡 정보를 바탕으로 항만은 정박지·예인선·하역 장비를 선제적으로 배치해 운영 효율을 극대화하고 불필요한 체선료를 줄일 수 있다.

3. 국내외 관련 연구개발 현황 및 문제점

국외에서는 AIS 데이터를 활용한 해상 교통 분석 및 예측 기술이 활발히 연구되고 있으며 항로 최적화, 충돌 방지, 해양사고 예방 등을 위한 다양한 모델들이 개발되고 있다. 특히, 유럽과 북미를 중심으로 해상 디지털 전환이 빠르게 이루어지며 딥러닝 및 강화학습 기반의 실시간 예측 시스템이 상용화 단계에 근접하고 있다. 반면, 국내에서는 해양수산부, 한국해양교통안전공단(KOMSA) 등에서 관련 연구가 수행되고 있으나 대부분이 통계 기반 분석이나 규칙 중심의 예측 방식에 머무르고 있어 실시간성이나 동적 환경 반영 측면에서 한계가 뚜렷하다. 항만 관제사의 경우 AIS 데이터 기반 예측의 신뢰도가 낮아 실시간 대응에 어려움을 겪고 있으며 선박 운항자는 실시간 항로 변경 및 위험 예측이 어렵다는 문제를 안고 있다. 현재 국내의 선박 경로 예측 기술은 복잡한 교통 상황과 다양한 변수들을 종합적으로 고려하지 못하는 구조이며 머신러닝 및 인공지능 기반의 실시간 분석 및 예측 기술 개발이 시급한 상황이다.

4. 본 프로젝트의 해결 방안

1) 목표:

어텐션 매커니즘 기반 딥러닝 모델을 활용하여 선박의 현재 상태와 목적지를 바탕으로 향후 이동 경로를 실시간으로 예측하는 시스템의 구축을 목표로 한다.

2) 핵심 구성 요소

- 트랜스포머 기반 경로 예측 모델: 시계열 예측에 특화된 트랜스포머 구조를 활용하여 선박의 이동 패턴 학습
- 어텐션 매커니즘: 과거 항해 이력의 중요도를 동적으로 반영하여 예측 정확도 향상

- 자가회귀 방식 예측: 한 스텝씩 미래 좌표를 예측하고 그 결과를 다음 입력으로 사용하는 auto-regressive 구조 적용

3) 기대 효과

- 선박 경로의 실시간 예측 정확도 개선
- 기존 규칙 기반 모델의 한계를 극복
- 해상 교통의 디지털 전환 및 자동화 기반 마련

2. 연구개발의 목표 및 범위

1. 연구개발의 전체 목표

본 연구개발의 최종 목표는 AIS 데이터를 활용한 실시간 선박 경로 예측 시스템을 구축하는 것이다. 기존의 통계 기반 예측이나 단순 회귀 모델이 가지는 정적 예측 한계를 극복하고 트랜스포머와 어텐션 메커니즘을 기반으로 한 딥러닝 모델을 통해 선박의 이동 경로를 정밀하고 동적으로 예측함으로써 해상 교통의 안전성과 운항 효율성을 개선하는 것을 목표로한다.

2. 연구개발의 내용 및 범위

1) 데이터 기반 구축

- AIS 빅데이터를 활용한 전처리 및 학습 데이터셋 구축
- 위도, 경도, SOG, COG 목적지 등 핵심 특성(feature) 추출

2) 딥러닝 기반 예측 모델 개발

- 트랜스포머 구조를 기반으로 한 시계열 예측 모델 설계
- 어텐션 메커니즘을 적용하여 과거 항해 정보의 가중치 반영
- 자가회귀(auto-regressive) 방식으로 다단계 미래 위치 예측

3) 예측 시스템 구현 및 검증

- 실시간 입력 데이터를 활용한 선박 경로 예측 시스템 구현
- 시각화 및 도착시간(ETA) 추정 기능 포함
- 실제 항로 데이터와 비교를 통한 성능 평가 및 정확도 검증

4) 활용 및 확장 가능성 제시

- 해양 관제, 자율 운항 시스템, 해상 물류 최적화 분야와의 연계 가능성 탐색
- 향후 충돌 방지 및 경로 자동 보정 기능 추가를 위한 기반 마련

3. 연구개발 수행 내용 및 결과

1. 연구개발 수행 내용

본 프로젝트는 AIS 데이터를 기반으로 어텐션 중심의 딥러닝 모델을 활용하여 선박의 이동 경로를 예측하는 시스템을 개발하는 것으로, 다음과 같은 세부 절차를 통해 추진된다.

가) AIS 데이터 수집 및 구성

- (1) 데이터 구성: MMSI, 일시, 위도, 경도, SOG, COG, 선수 방향(Heading)
- (2) 전처리 절차:
 - (a) 30초 간격 시계열 리샘플링: mean-interpolation 활용
 - (b) 이상치 제거: SOG, COG, 위도, 경도별 이상치 제거
 - (c) 출발지-목적지 쌍 기반 그룹화

- 출발 및 목적지 기반으로 동일 항로 그룹 생성
- 각 MMSI에 대해 동일 OD 쌍이 반복된 경로를 추출해 모델 학습용 시계열 데이터 구성

나) 입력/출력 구조 정의 및 학습 데이터 구성

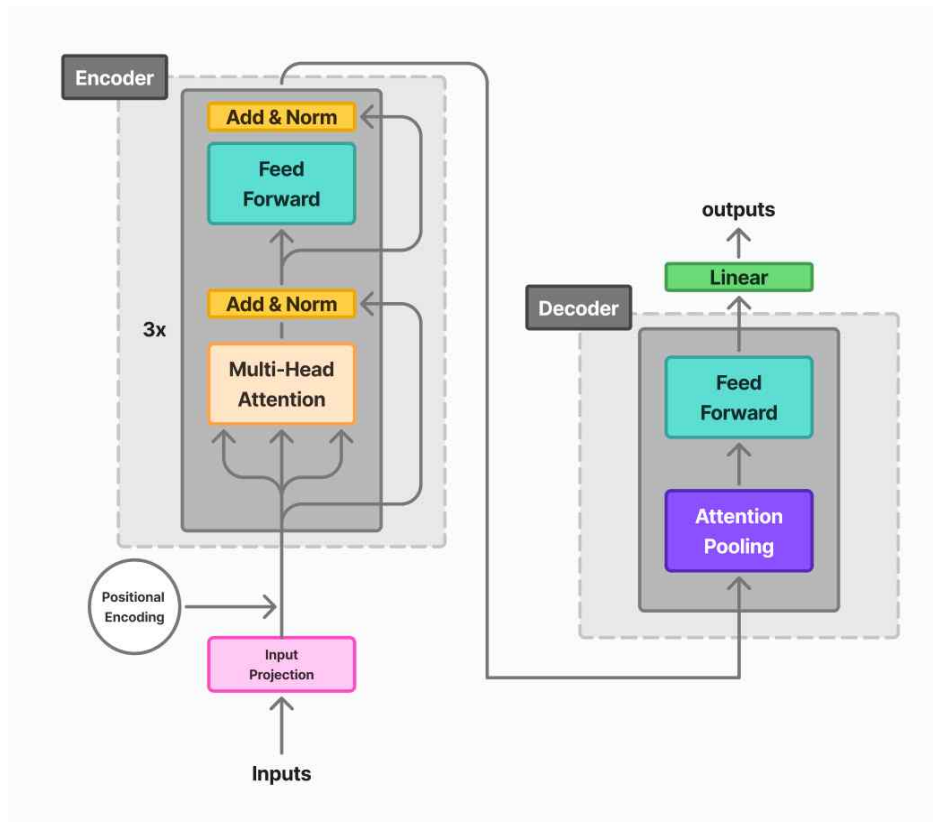
(1) 입력값:

- (a) 과거 20분 간격의 데이터(30초 단위, 총 40개), 각 시간에 대해 다음을 포함:
- 위도(latitude), 경도(longitude), SOG, COG -> 4차원
 - 목적지와 현재 위치 간 유클리디안 거리 -> 1차원

(2) 출력값:

- (a) 미래 30초 뒤 시점의 위도, 경도 SOG, COG 예측 -> 4차원 단일 벡터

다) 모델 아키텍처 설계



(1) 인코더 구조(Transformer 기반):

- (a) Input Projection을 통해 입력 시계열 데이터를 Transformer가 처리할 수 있는 고정 차원의 벡터로 임베딩
- (b) 포지셔널 인코딩(Positional Encoding)을 적용하여 시계열 순서 정보 반영
- (c) 멀티 헤드 어텐션(Multi-head Self-Attention)을 적용해 시점 간 상호관계 학습
- (d) Residual Connection과 Layer Normalization을 적용해 학습 안정성 및 정보 손실 방지

(2) 디코더 구조(MLP 기반)

- (a) 어텐션 풀링(Attention Pooling)을 적용해 전체 시퀀스의 핵심 정보를 요약하여 고정 길이 벡터

추출

(b) Fully Connected Layer를 통해 위도, 경도, SOG, COG를 예측

(c) 비선형 활성화 함수로 GELU(Gaussian Error Liner Unit)를 사용해 모델의 표현력 향상

라) Loss Function 및 학습 전략

(1) 복합 Loss 구성:

```
# 위치 오차 (위경도 기준)
loc_loss = F.mse_loss(pred_lat, true_lat) + F.mse_loss(pred_lon, true_lon)

# SOG, COG 오차
motion_loss = (
    (1-cog_weight) * F.mse_loss(pred_sog, true_sog) +
    cog_weight * F.mse_loss(pred_cog, true_cog)

# 최종 Loss: 위치 오차 + SOG, COG 오차
total_loss = (
    self.weight_main * loc_loss +
    (1 - self.weight_main) * motion_loss
)
```

(a) loc_loss: 위도, 경도에 대한 MSE

(b) motion_loss: SOG, COG에 대한 MSE

(c) weight_main, cog_weight 하이퍼 파라미터를 통해 가중치 조절

(2) 훈련 방식:

(a) 단일 스텝 예측 학습: 20분 입력 -> 30초 예측

(b) 배치 단위의 시퀀스 훈련, weight decay를 적용하는 AdamW 옵티마이저를 사용해 더 나은 일반화 성능 확보

(c) Cosine-Annealing 스케줄링을 통해 학습률을 코사인 곡선 형태로 조절하여, 빠른 초기 학습과 안정적인 후반 수렴 동시 달성

(d) Gradient Norm Clipping 기법을 적용해 Gradient Exploding 방지

(e) Noise Injection 기법을 활용해 차후 Auto-regressive Inference 과정에서의 강건성 확보

마) 자가회귀 기반 경로 예측 및 생성

(1) Auto-regressive Inference:

(a) 20분 입력 -> 30초 출력 -> 이 출력을 다음 입력의 가장 최근 시점으로 추가

(b) 반복적으로 예측을 수행하여 최종 목적지 도달까지 전체 경로를 생성

(c) 예측 시마다 목적지 정보는 고정 입력으로 유지

(d) Monte-Carlo Dropout(MC Dropout) 기법을 활용하여, 매 step 마다 dropout을 적용, 이들의 평균을 사용함으로써 error accumulation 현상을 완화

(2) 예측 종료 조건:

(a) 목적지 좌표와의 거리가 일정 기준 이하일 경우 종료

(b) 또는 최대 반복 횟수 도달 시 종료

바) 지도 기반 예측 경로 시각화

(1) 시각화 도구:

(a) folium 라이브러리 기반 지도에 예측 경로(자가 회귀 결과) 시각화

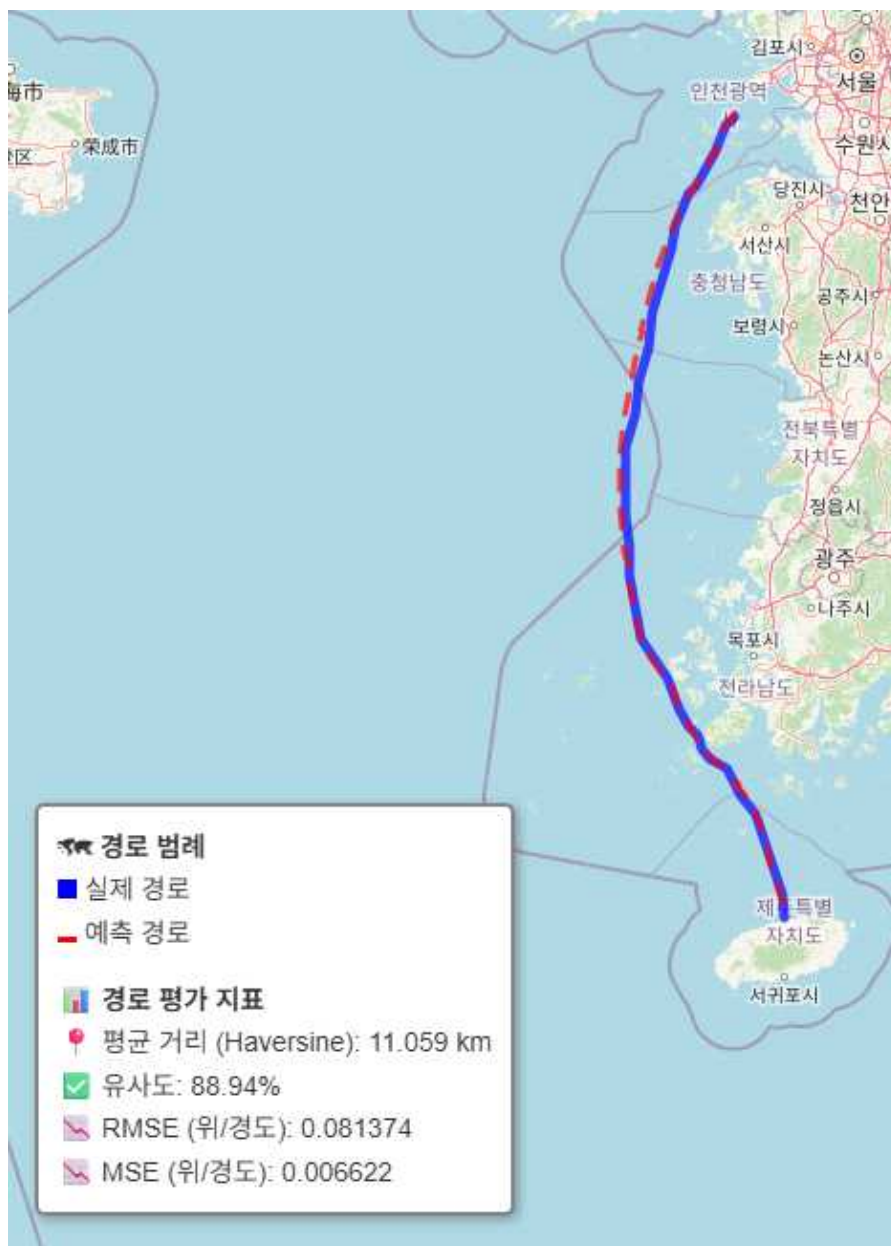
(b) 실제 경로와의 비교를 통해 오차 확인 및 성능 평가 가능

2. 연구개발 결과

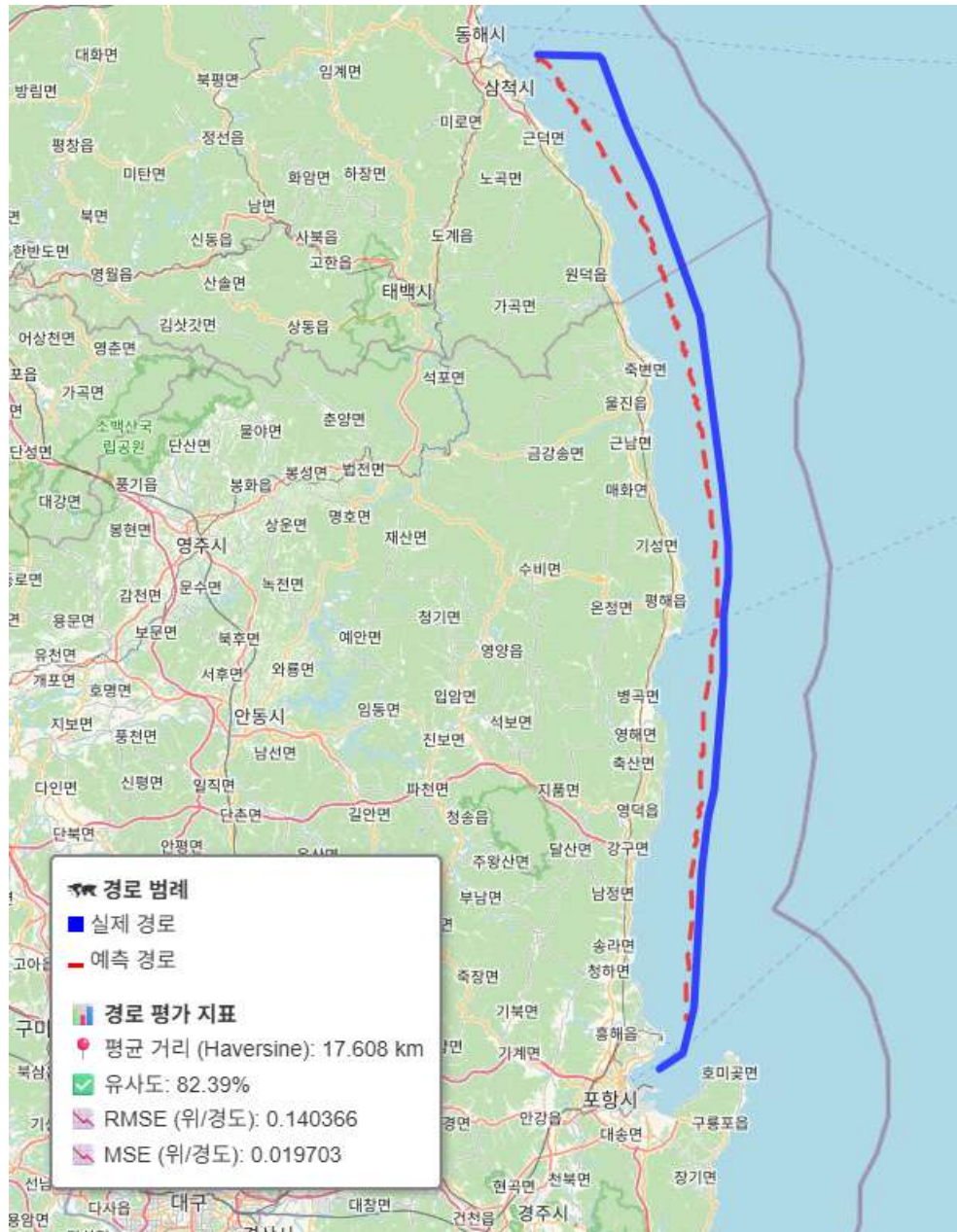
가) 모델 성능 테스트

아래 지도는 각 항로 사례에 대한 예측 경로와 실제 경로를 비교하여 모델의 성능을 정량적/시각적으로 평가한 결과이다.

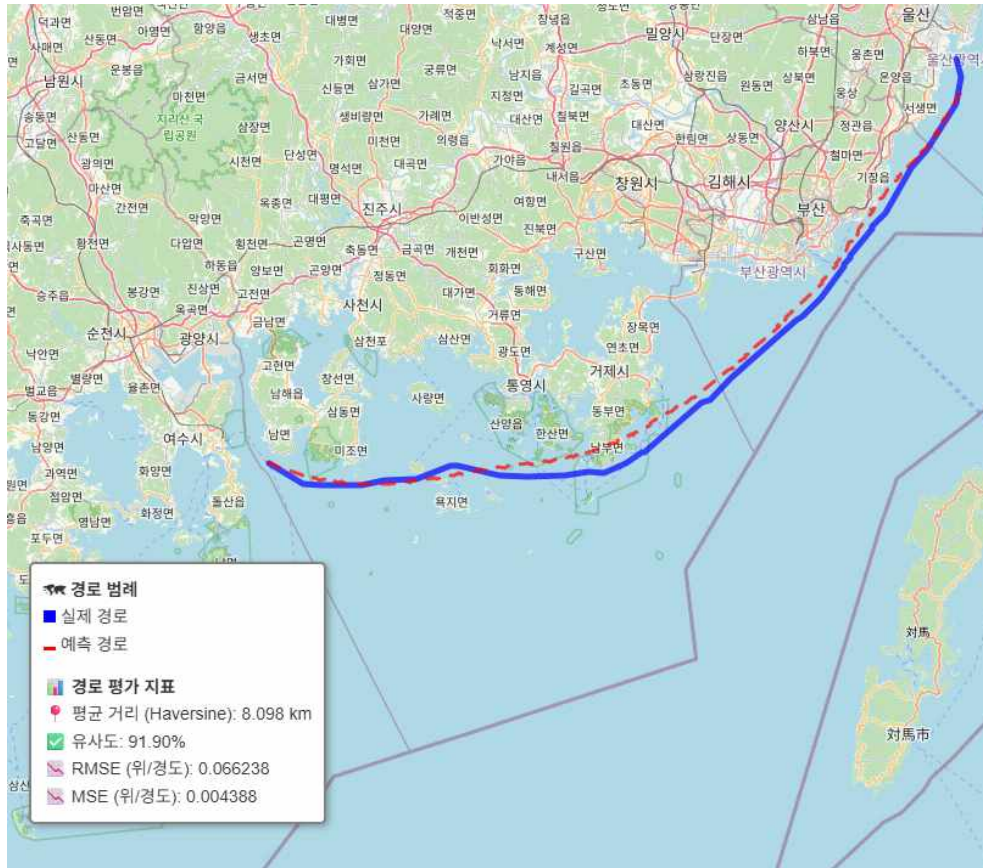
(1) 인천항 -> 제주항



(2) 동해항 -> 포항항



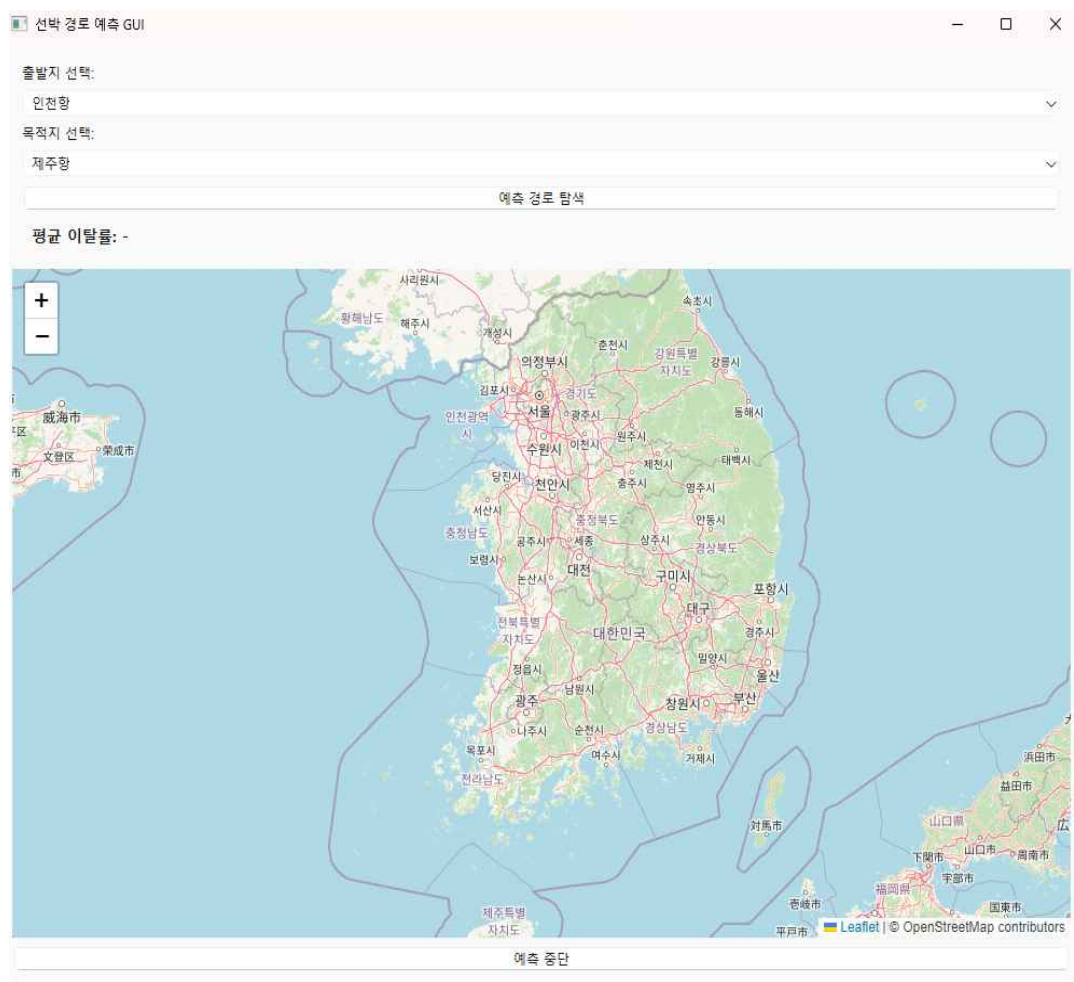
(3) 여수항 -> 울산항



나) 데모 테스트

다음은 PyQt6 기반 GUI와 Folium 시각화를 결합하고, CSR(Client-Side Rendering) 방식으로 동적 화면 최적화를 구현한 데모 버전이다.

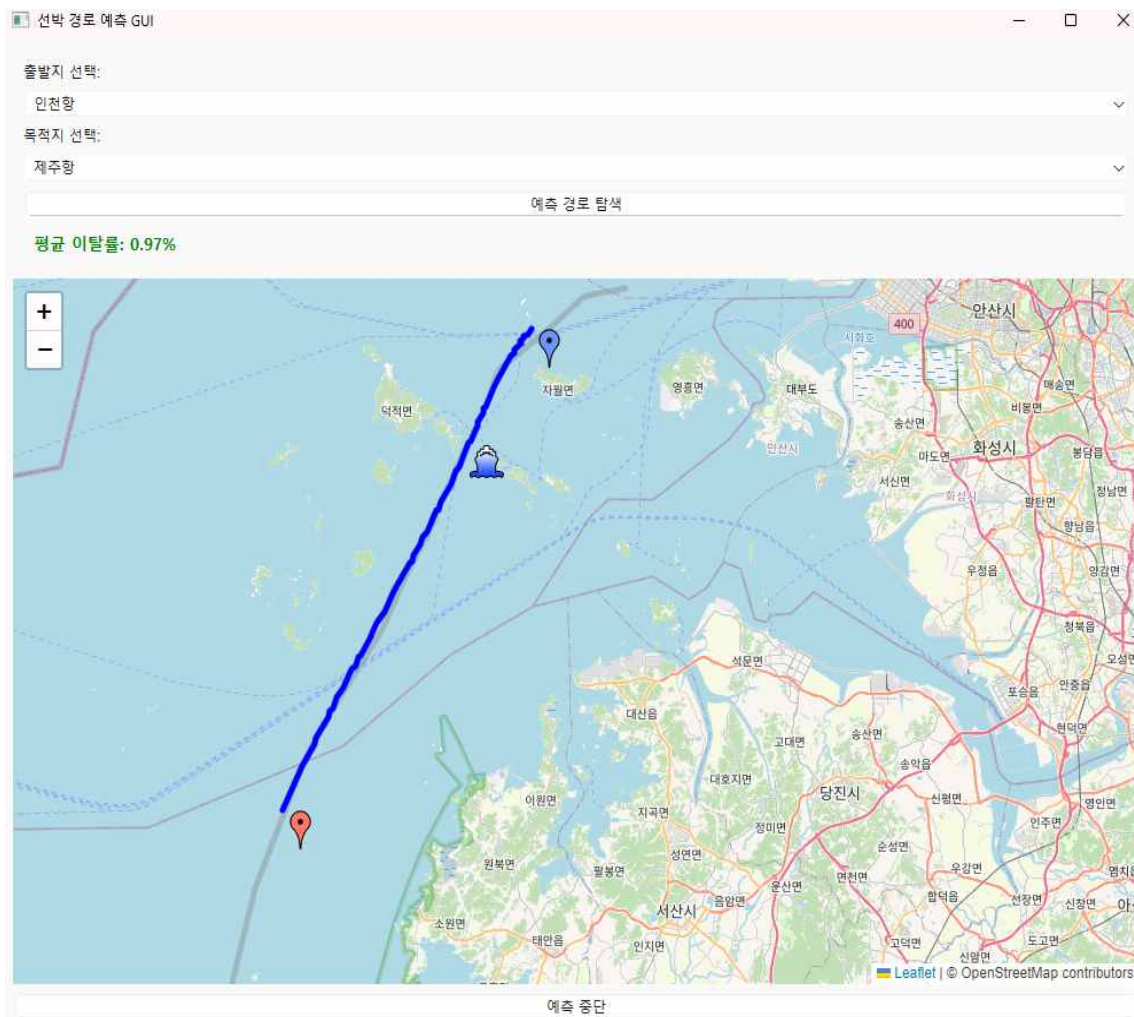
(1) 초기 화면



- 출발지 선택: 출발 항구를 선택하는 리스트 박스
- 목적지 선택: 도착 항구를 선택하는 리스트 박스
- 예측 경로 탐색: 선택한 출발지/목적지를 기준으로 항로 예측을 시작하는 버튼
- 평균 이탈률: 출발지에서 목적지까지 가장 많이 활용된 항로와 모델이 예측한 항로와의 차이 백분율
- 지도: 예상 경로, 선박의 위치 등을 실시간으로 표현하는 지도
- 예측 중단: 모델의 항로 예측 중단 버튼

(2) 초기 실행 화면

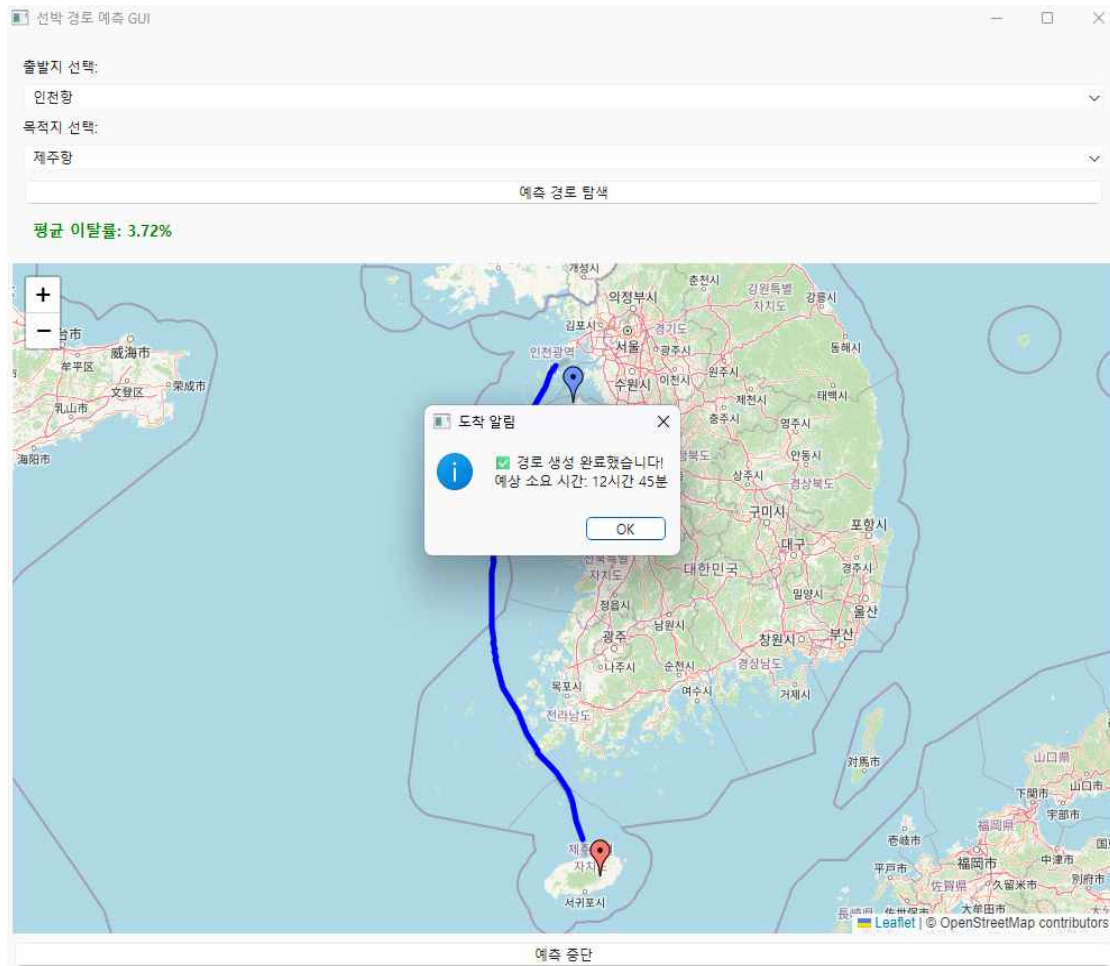
예측 경로 탐색 버튼 클릭 시 발생하는 화면으로, 매 초마다 모델이 예상한 경로가 지도상에 동적으로 업데이트된다. 그와 동시에, 최빈 경로와 모델이 예측한 경로와의 차이를 기반으로 평균 이탈률이 계산돼 표시된다.



- 파란색 마크: 선박의 출발 위치
- 빨간색 마크: 선박의 예상 도착 위치(목적지)
- 파란색 선박: 선박의 현재 위치(모델의 예상 경로를 따라간다고 가정)
- 파란색 경로: 모델이 예측한 선박의 예상 경로
- 회색 경로: 빅데이터 분석을 통해 얻은, 출발지에서 목적지까지의 최빈 경로(이탈률 계산에 활용)

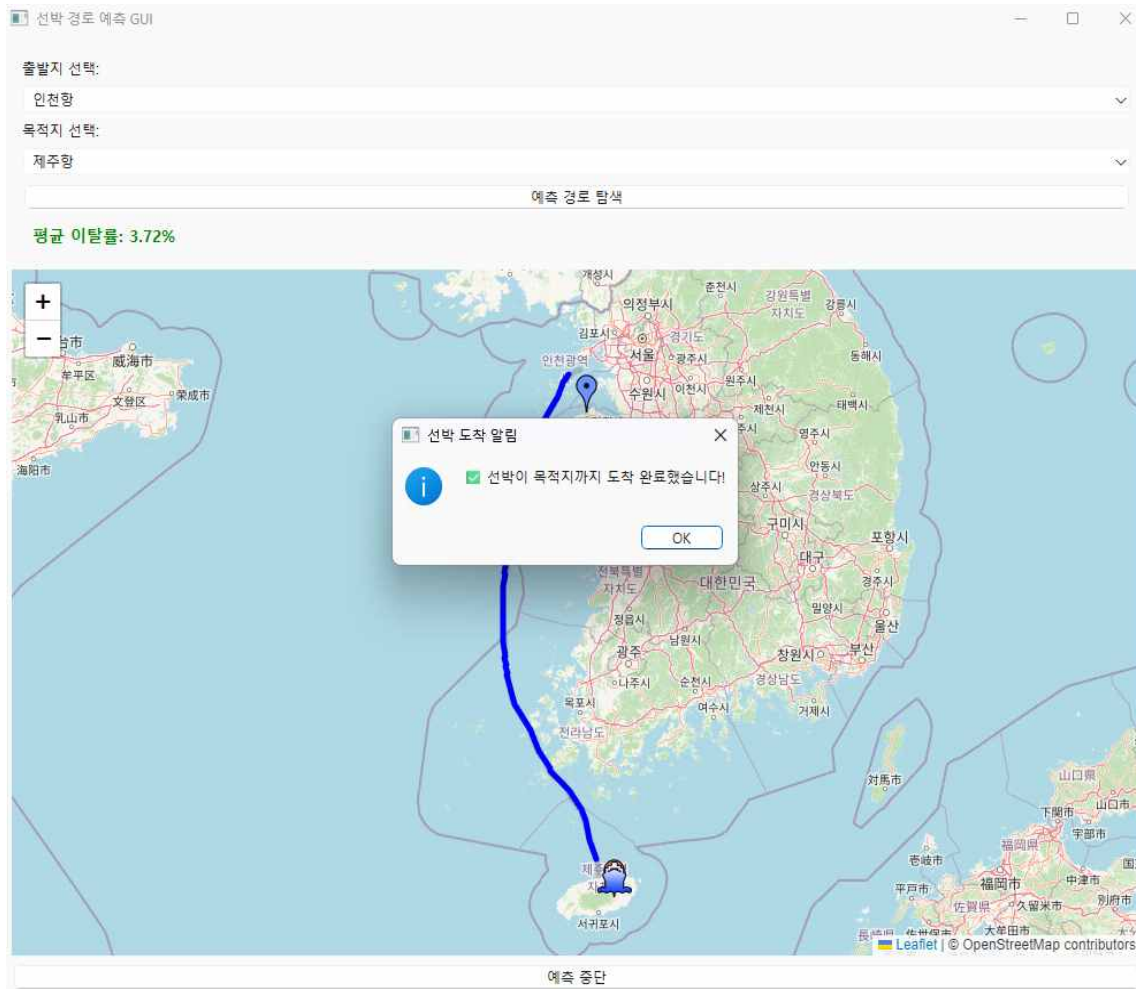
(3) 경로 생성 완료 화면

모델이 예상한 경로(파란색 선)가 목적지에 도달했을 경우 나타나는 화면으로, 목적지까지의 경로 생성 여부와 선박의 예상 도착 시간(ETA)이 표시된다.



(4) 선박 도착 완료 화면

선박이 예상 경로를 따라 이동하여 최종적으로 목적지까지 도착 완료했을 경우 나타나는 화면이다.



다) 깃허브 주소: <https://github.com/yeonbum/BigDataAnalysis>

3. 결과에 대한 평가

가) 모델 성능 테스트

총 3개의 경로에 대해서 테스트를 진행, 각 경로별 평가 지표의 평균을 계산함.

정량적 평가 지표	모델 성능
평균 거리 오차	12.255km
유사도	87.8%
RMSE	0.09599
MSE	0.01023

- 평균 거리 오차: 실제 경로와 예측 경로 간 평균 Haversine 거리 오차
- 유사도: 실제 경로와 예측 경로 간 경로 유사도로, 평균 거리 오차 지표를 기준으로 측정
- RMSE: 실제 경로와 예측 경로 간 위도, 경도 RMSE
- MSE: 실제 경로와 예측 경로 간 위도, 경도 MSE

나) 데모 테스트

실시간 선박 이동 경로 예측 시스템을 성능과 UI 측면에서 평가함.

- 자가 회귀 단계를 통해 생성되는 예측 경로를 시각화하는 과정에서, 초 단위(1000ms)로 html 파일을 매번 새롭게 렌더링하는 과정 때문에 시간이 오래 걸려, 이를 해결하기 위해 CSR(Client-Side Rendering) 기법을 적용해 소요 시간을 절반 이상으로 단축시켰다.
- 모델이 예측/생성하는 경로의 신뢰도를 높이기 위해, 정량적 지표인 평균 이탈률을 매 초마다 업데이트 및 시각화를 진행했다.

4. 기대효과 및 활용방안

본 프로젝트를 통해 개발되는 선박 이동 경로 예측 모델은 다음과 같은 효과를 기대할 수 있다.

1. 해상 물류 운영 효율 향상: 선박의 이동 경로와 도착 시간을 사전에 예측함으로써 항만 운영 및 물류 일정 계획의 정밀도를 높이고, 대기 시간 및 연료 낭비를 최소화할 수 있다.
2. 항로 안정성 강화: 예상 경로를 기준으로 교통 혼잡 구간이나 비정상 항로 이탈 등을 사전에 탐지하여, 해양 사고 예방에 기여할 수 있다.
3. 자율 운항 및 스마트 해양 교통 기술 개발 기반 확보: 향후 자율 운항 선박(MASS) 및 스마트 항만 시스템의 핵심 기술 중 하나로 활용될 수 있는 기반 모델로서 기능할 수 있다.

5. 향후 과제 및 추진 일정

1. 향후 과제

- (1) 모델의 예측 정확도를 더 개선해야 한다.
- (2) 더 많은 항구를 기준으로 예측 경로가 생성될 수 있도록 모델을 개선해야 한다.

2. 추진 일정 (2025-2학기 예정)

- (1) 입력 데이터에 지형, 기상 정보 등 다양한 Feature 들을 추가해, 보다 정밀한 예측 경로를 생성하도록 할 예정이다.
- (2) AIS 데이터로부터 더 많은 항로를 추출하여 모델의 학습에 사용할 예정이다.

6. 팀 구성 및 역할

- | |
|-----------------------------------|
| 1. 김연범: 데이터베이스 구축, 모델 설계 및 GUI 구현 |
| 2. 이한을: 데이터베이스 구축, 모델 설계 및 GUI 구현 |
| 3. 이동원: 참고 문헌 조사, front-end 구현 |

7. 참고문헌

- [1] 황선일, 박상원, 이경미 그리고 박민정, AIS 데이터 기반 해상교통 안전 평가모델 개발 연구, *한국해양수산개발원*, 2024.
- [2] 김상우, 임연희, 조현희 그리고 이길중, AIS 데이터 분석을 통한 선박 안전항로 딥러닝 모듈 개발, *ACK 2024 학술발표대회 논문집*, vol. 31, no. 2, pp. 1012-1013, 2024.
- [3] Samuele Capobianco and Leonardo M. Millefiori etc, Deep Learning Methods for Vessel Trajectory Prediction based on Recurrent Neural Networks, *arXivpreprint*, arXiv: 2101.02486, 2021.
- [4] 김재원 그리고 정민, AIS 데이터 신뢰성 개선에 관한 기초 연구:해상교통관제사를 중심으로, *한국해양경찰학회보*, vol. 11, no. 4, pp. 49-68, 2021.