Transformer-Based Model for Vessel Route Forecasting and Anomaly Detection

Studente
Alessandro Marzilli – 1811092

Relatore
Prof. Danilo Comminiello

Correlatore
Ing. Claudio Greco

Engineering in Computer Science Dipartimento di Ingegneria informatica, automatica e gestionale
Facoltà di Ingegneria dell'informazione, informatica e statistica
Anno Accademico 2022-2023
Sapienza University of Rome

In collaboration with:

Elman s.r.l. is a Italian manufacturer specializing in shipborne and land-based Automatic Identification Systems (AIS) equipment for both military and civil applications.

Table of contents

1. Introduction

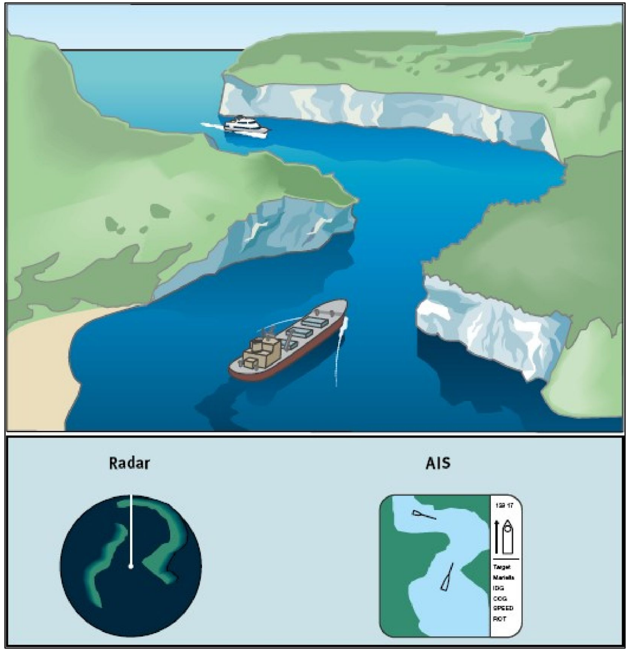
The **maritime industry** has experienced a profound transformation in recent years, driven by the **increased accessibility of data (AIS data)** and the **advancements in deep learning techniques**.

The thesis aims to enhance **safety** and **security** measures by supporting and simplifying the work of Port Authorities, Rescue teams and other parts.

Introduction: AIS Background

AIS is a **tracking system** that uses transponders on ships to transmit and receive data, providing real-time information about vessel movements and characteristics.

AIS enables effective **collision avoidance**, helps in **search and rescue** operations, and assists in the detection of potential **security threats**, such as piracy and smuggling.

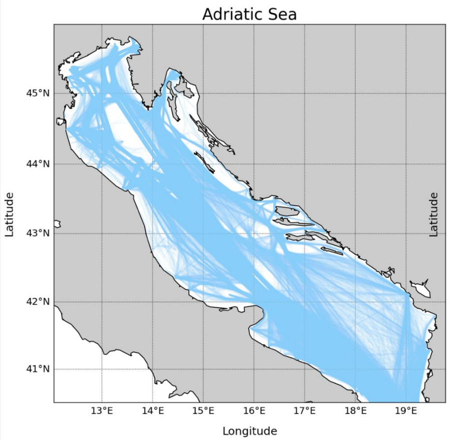


AIS collision avoidance

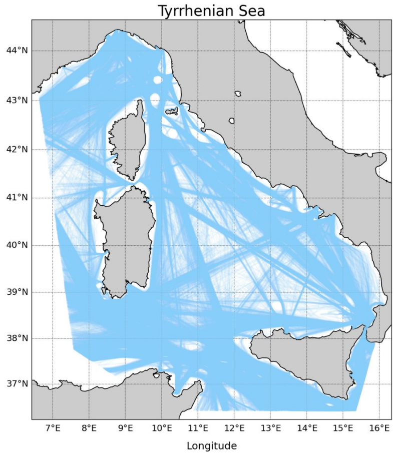
2. AIS Dataset

The AIS dataset is **heterogeneous** and **multimodal**. Each message has multiple attributes: id, type of the ship, SOG, COG, latitude, longitude, destination, ETA, heading, name.

The regions of interest (ROI) studied are from two geographical areas: the **Adriatic Sea** and the **Tyrrhenian Sea**.



From December 2021
To March 2023

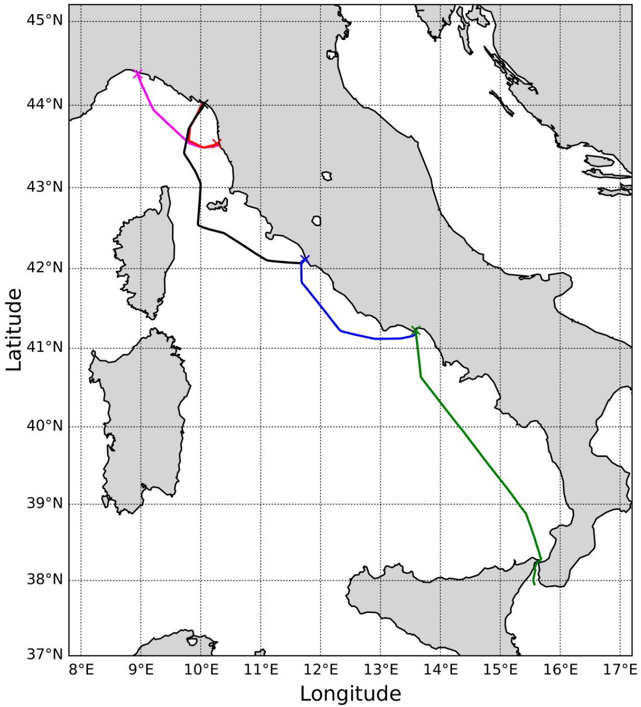


From December 2021
To March 2023

AIS DATASET: Preprocessing Phase

Only two categories are considered: cargo and tanker.

- Remove null values and duplicates
- Remove messages with invalid values (SOG > 30 or COG > 360°)
- Remove moored or at-anchor vessels.
- Delete abnormal messages (empirical speed > 40 knots)
- Divide non-contiguous voyages into contiguous ones
- Reme sub-tracks that are shorter than 20 messages or last less than 4 hours
- Apply interpolation with a down-sampling of 6 minutes to maintain linearity in the tracks
- Further splitting of the tracks greater than 20 hours



Sub-tracks of a single vessel obtained after the pre-processing phase

3. Vessel Route Forecasting

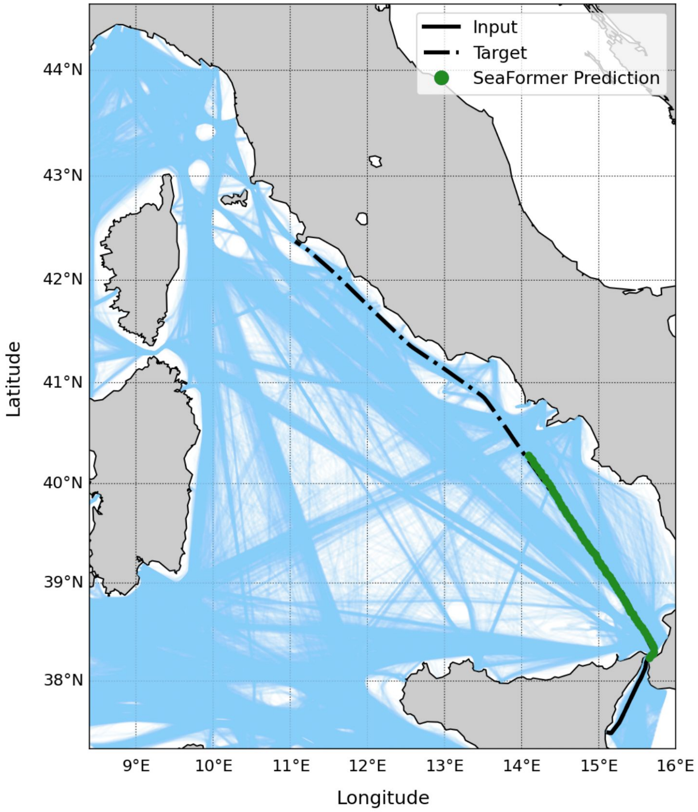
The task focuses on **predicting** the ship's possible route in the following few hours based on its current trajectory.

Formally, given a trajectory of T+1 AIS messages, the vessel route forecasting problem can be formalized as the **estimation of the next L steps ahead** of a ship:

$$p\left(\mathbf{x}_{T+1:T+L} \mid \mathbf{x}_{0:T}\right) = \prod_{l=1}^{L} p\left(\mathbf{x}_{T+l} \mid \mathbf{x}_{0:T+l-1}\right)$$

Vessel Route Forecasting: Proposed Method

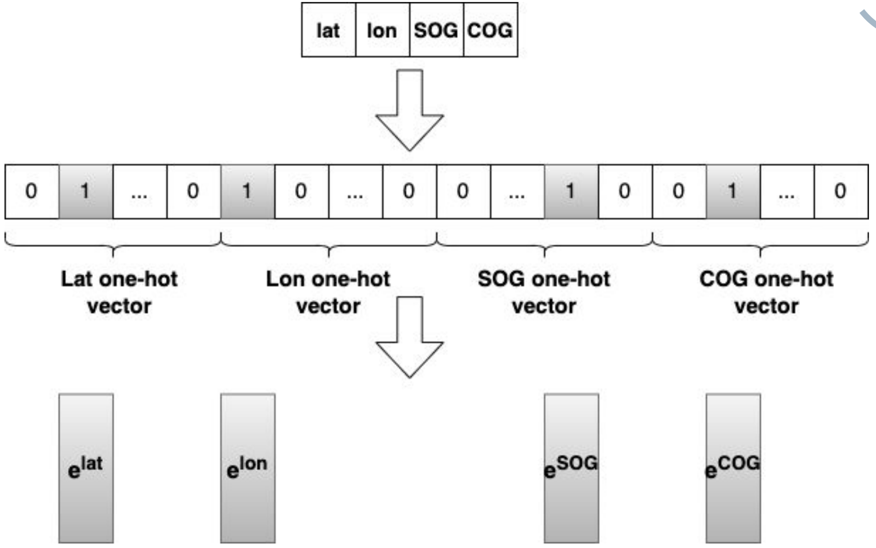
- New representation of AIS data
- New learning approach
- SeaFormer architecture



Proposed Method:
New Representation of AIS Data

A **new representation** of AIS data is introduced to address the issue of heterogeneity.

The key idea is to **expand** the message x_t **into a higher-dimensional embedding vector** e_t, by first mapping it to a **4-hot** vector h_t.



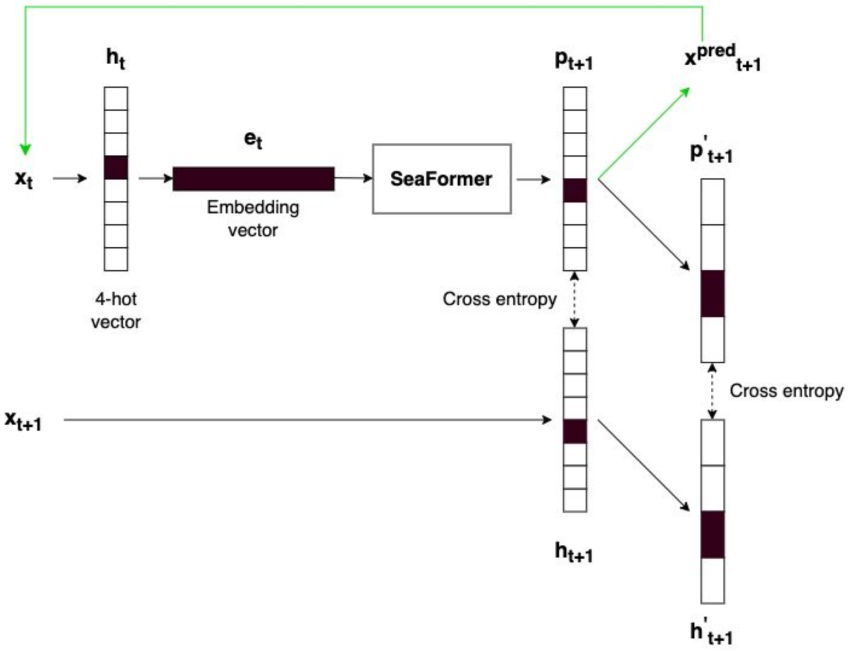
Mapping process of an AIS message

Proposed Method: New Learning Approach

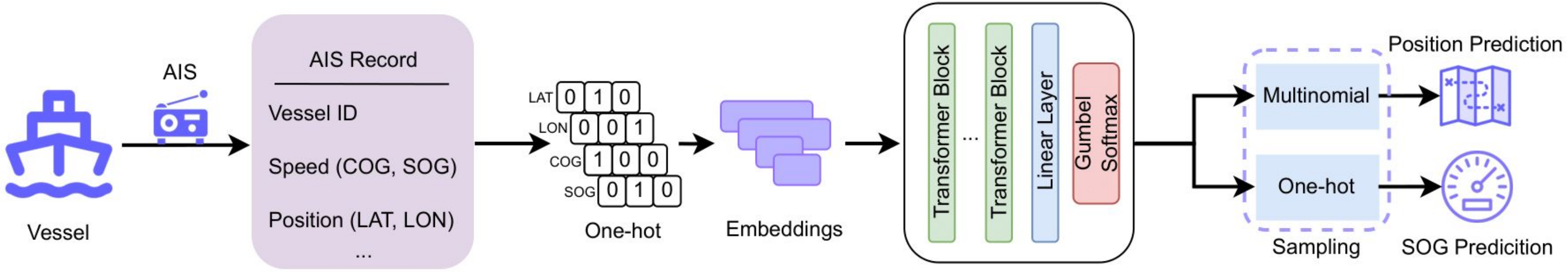
Given the new data representation, the problem can be formulated as a **classification problem**.

Loss function definition:

$$\mathcal{L}_{\mathrm{CE}} = \sum_{l=1}^{L} \underbrace{\mathrm{CE}\left(p_{T+l}, \mathbf{h}_{T+l}\right)}_{\text{high resolution}} + \beta \underbrace{\mathrm{CE}\left(p'_{T+l}, \mathbf{h}'_{T+l}\right)}_{\text{low resolution}}$$



Proposed Method: SeaFormer Architecture



The architecture of the proposed SeaFormer is based on **transformer-like models**, specifically a GPT model.

Key innovations:

- **Gumbel Softmax** as final activation function
- **Different sampling function** for the position and the SOG prediction

Vessel Route Forecasting: Position Prediction

Metrics: A ***best-of-N criterion*** is used to evaluate the model. Then, the **haversine distance** between the true and predicted values is computed for the error on the position.

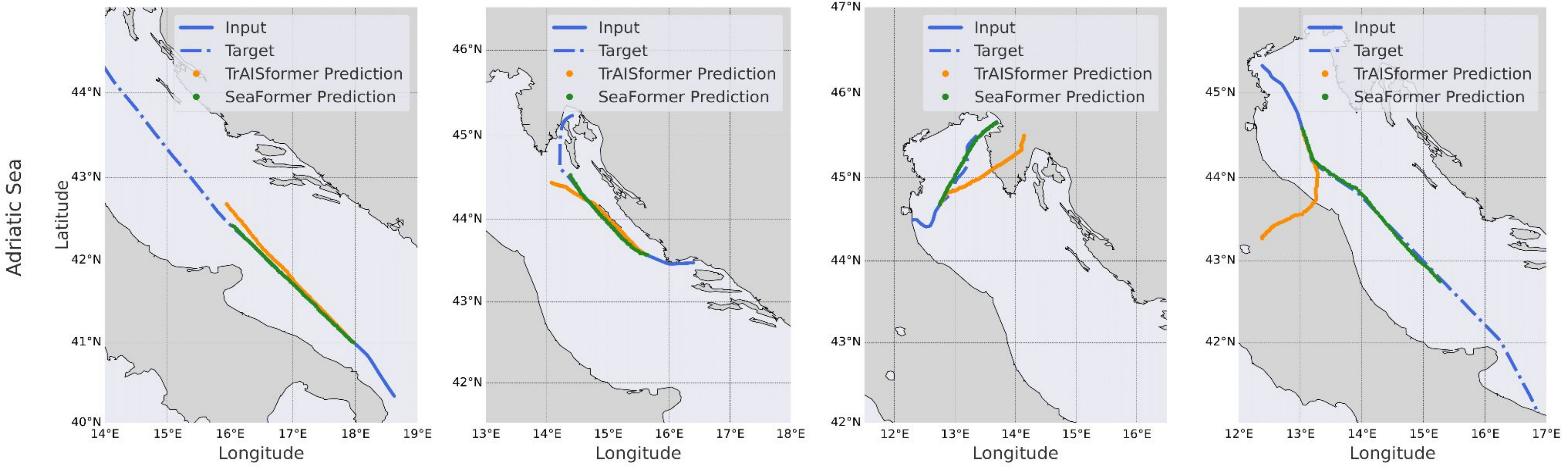
Model	Sampling	1h	2h	3h	4h	5h	6h	7h	8h
Traisformer	Multinomial	0.4050	0.8261	1.3233	1.8707	2.4230	3.0407	3.7505	4.3610
SeaFormer	Multinomial	**0.3942**	**0.7798**	**1.2346**	**1.7605**	**2.2598**	**2.7648**	**3.3026**	**3.8834**

Quantitative comparison on position error in mni (nautical mile) for Adriatic sea

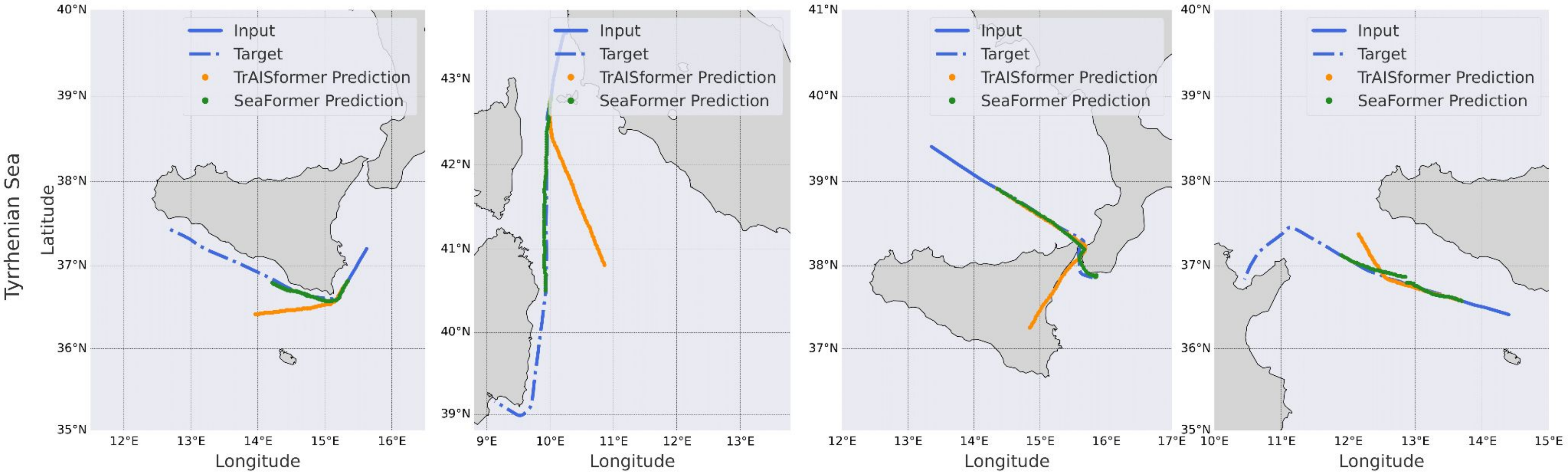
Model	Sampling	1h	2h	3h	4h	5h	6h	7h	8h
Traisformer	Multinomial	0.4420	0.9274	1.5482	2.2582	2.9800	3.7662	4.5669	5.4023
SeaFormer	Multinomial	**0.4149**	**0.8769**	**1.4189**	**2.0214**	**2.6907**	**3.4532**	**4.2613**	**5.1754**

Quantitative comparison on position error in mni (nautical mile) for Tyrrhenian sea

Vessel Route Forecasting: Position Prediction

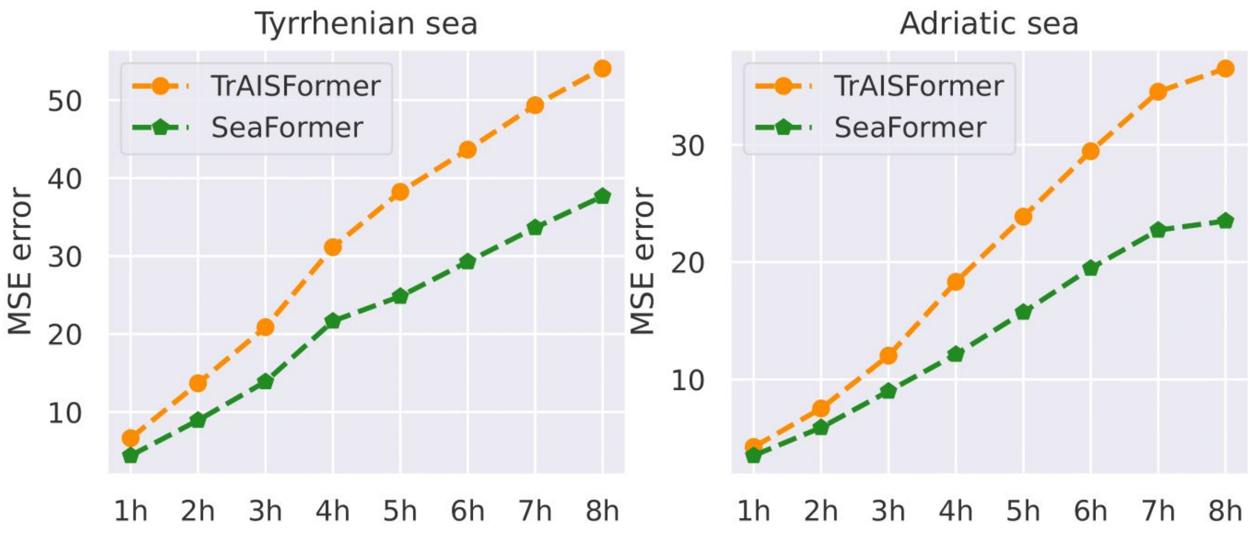


Vessel Route Forecasting: Position Prediction



Vessel Route Forecasting: SOG Prediction

Metrics: the mean squared error (**MSE**) metric is used.

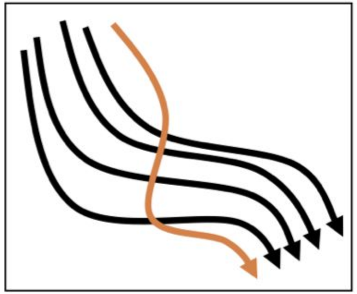


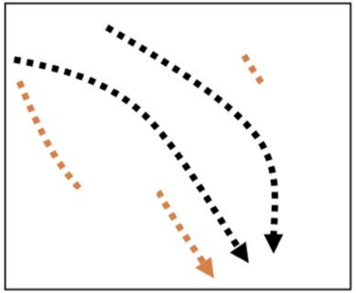
Comparison of MSE error for SOG predictions

4. Anomaly Detection

The task aims to detect **unusual** or **anomalous** trajectories that deviate from expected patterns.

This task helps in identifying potential threats, security breaches, or abnormal vessel activities that require further investigation or intervention.

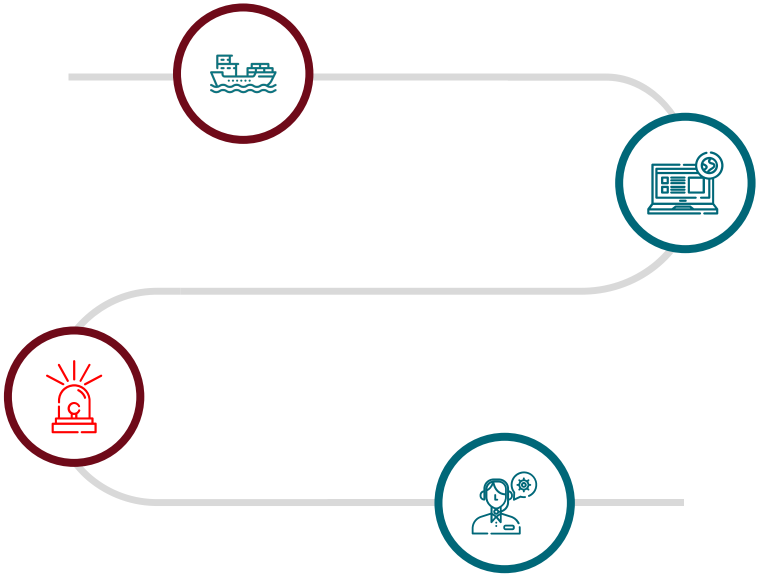

(a) Route Deviation


(b) Unexpected Activity

Anomaly Detection: Proposed Method

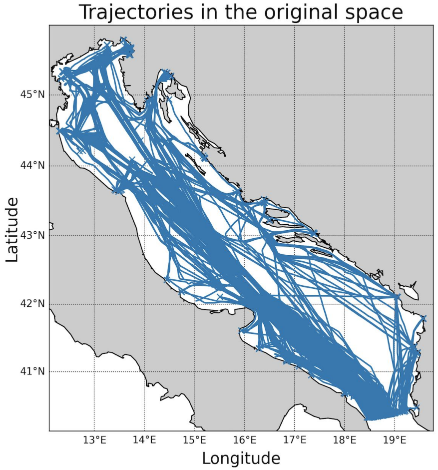
The proposed method is a clustering approach divided in three steps:

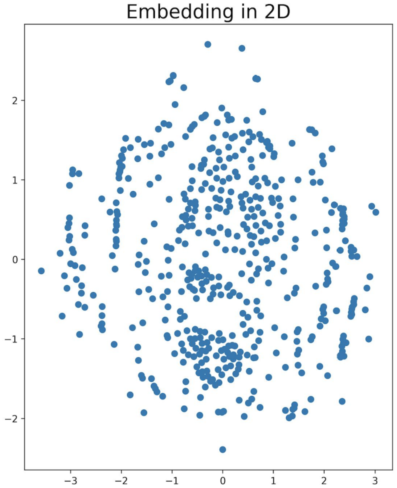
1. The Merge Distance (MD) between trajectories is computed to obtain a distance matrix
2. The trajectories are mapped to a low dimensional space (2D space) using the multidimensional scaling (MDS)
3. The improved DBSCAN algorithm is utilized to perform the clustering

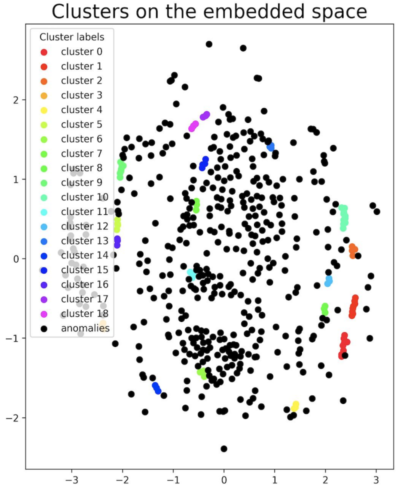


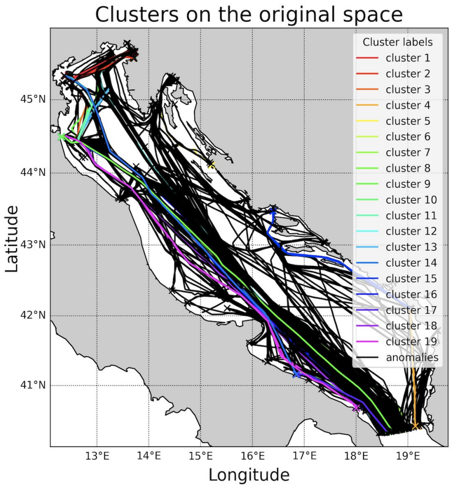
Anomaly Detection: Adriatic Results

For this experiment 500 tracks have been used.



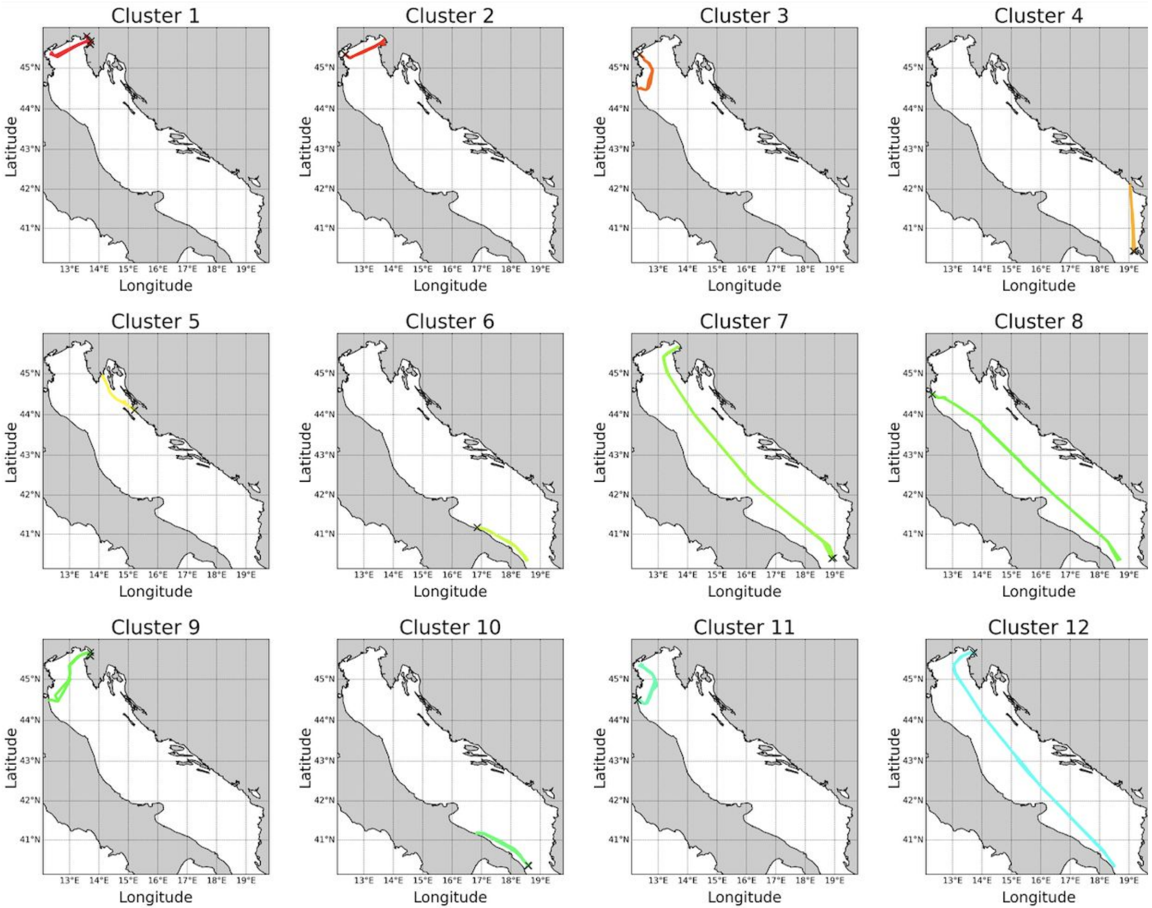






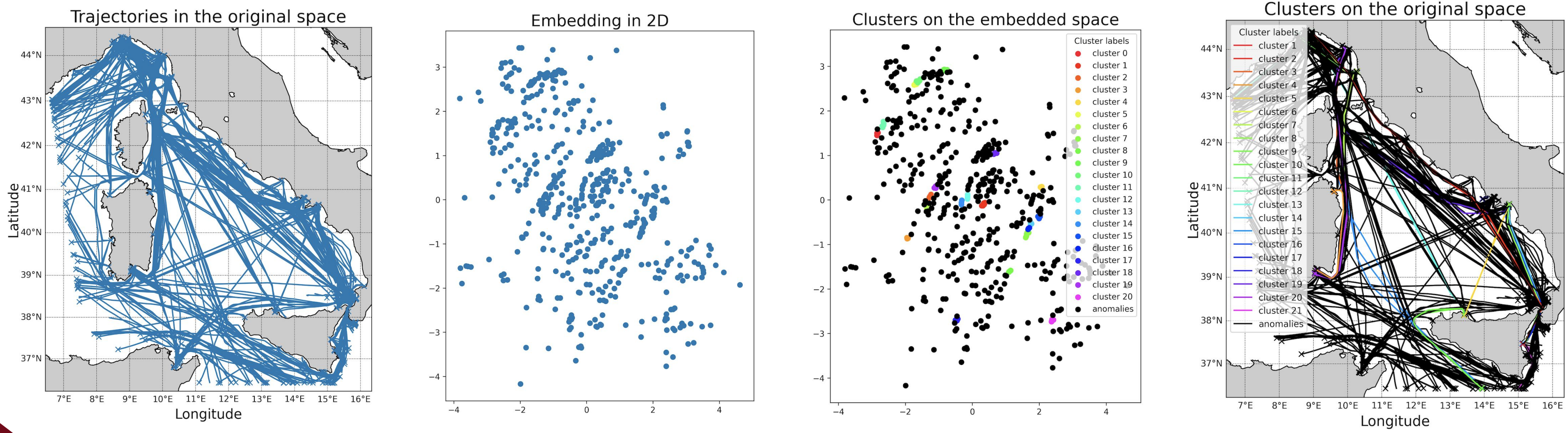
Anomaly Detection: Adriatic Results

The method correctly classifies 19 clusters (a subset of them is shown).



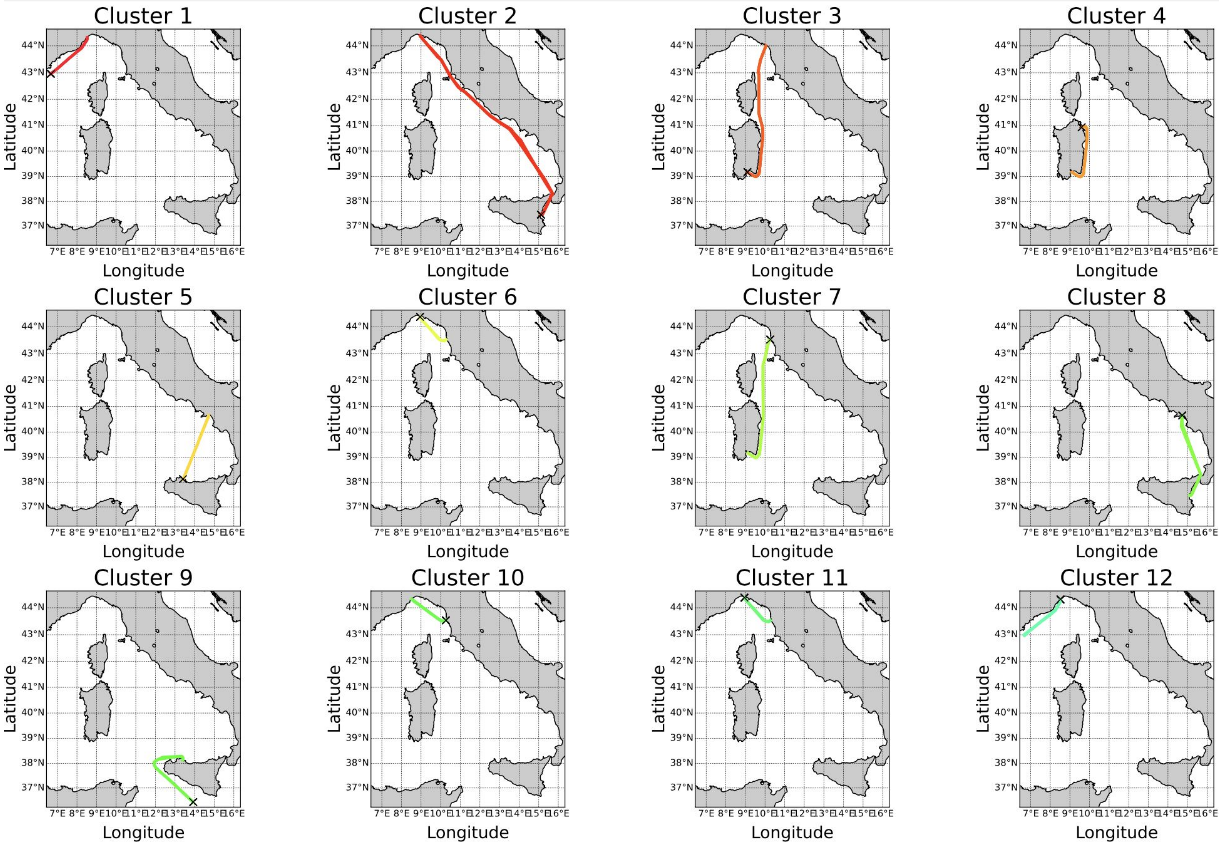
Anomaly Detection: Tyrrhenian Results

For this experiment 500 tracks have been used.



Anomaly Detection: Tyrrhenian Results

The method correctly classifies 21 clusters (a subset of them is shown).

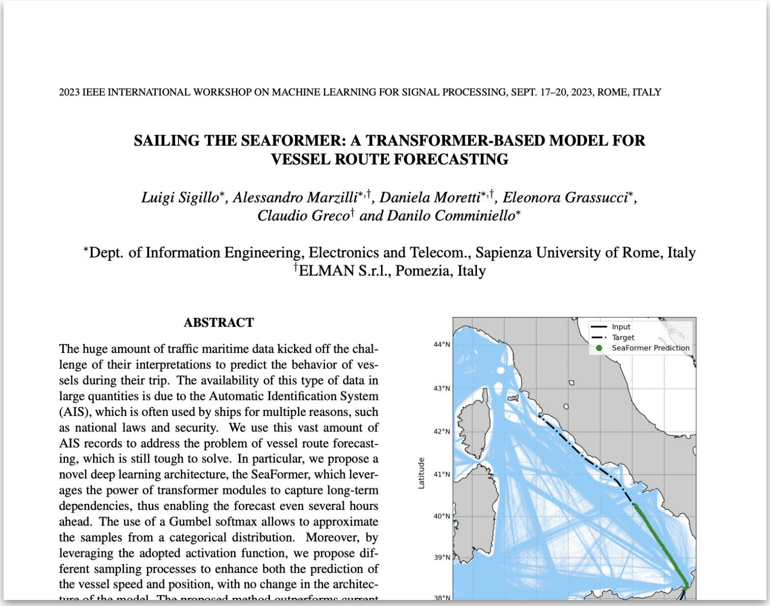


5. Future Works

- Consider a large set of trajectories for the anomaly detection task to improve the promising results already obtained
- Integrate the detection of other anomalies such as: close approach, zone entry and port arrival
- Prediction of the Estimated Time of Arrival (ETA)
- Integrate in the datasets AIS data coming from AIS satellites

Publication

We will also present the paper “Sailing the SeaFormer: a Transformer-based Model for Vessel Route Forecasting” at the “IEEE International Workshop on Machine Learning for Signal Processing”


2023 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 17–20, 2023, ROME, ITALY

SAILING THE SEAFORMER: A TRANSFORMER-BASED MODEL FOR VESSEL ROUTE FORECASTING

Luigi Sigillo[*], *Alessandro Marzilli*[*,†], *Daniela Moretti*[*,†], *Eleonora Grassucci*[*], *Claudio Greco*[†] *and Danilo Comminiello*[*]

[*]Dept. of Information Engineering, Electronics and Telecom., Sapienza University of Rome, Italy
[†]ELMAN S.r.l., Pomezia, Italy

ABSTRACT

The huge amount of traffic maritime data kicked off the challenge of their interpretations to predict the behavior of vessels during their trip. The availability of this type of data in large quantities is due to the Automatic Identification System (AIS), which is often used by ships for multiple reasons, such as national laws and security. We use this vast amount of AIS records to address the problem of vessel route forecasting, which is still tough to solve. In particular, we propose a novel deep learning architecture, the SeaFormer, which leverages the power of transformer modules to capture long-term dependencies, thus enabling the forecast even several hours ahead. The use of a Gumbel softmax allows to approximate the samples from a categorical distribution. Moreover, by leveraging the adopted activation function, we propose different sampling processes to enhance both the prediction of the vessel speed and position, with no change in the architecture of the model.

Thank you for the attention!