INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 16, Issue: 4, Month: August, Year: 2021
Article Number: 4345, https://doi.org/10.15837/ijccc.2021.4.4345

CCC Publications
AGORA UNIVERSITY PRESS

Novel Machine Learning Approach for Self-Aware Prediction based on the Contextual Reasoning

A. Daranda, G. Dzemyda

Andrius Daranda*
VU Institute of Data Science and Digital Technologies
Vilnius University, Lithuania
*Corresponding author: andrius.daranda@gmail.com

Gintautas Dzemyda
VU Institute of Data Science and Digital Technologies
Vilnius University, Lithuania
gintautas.dzemyda@mif.vu.lt

Abstract

Machine learning is compelling in solving various applied problems. Nevertheless, machine learning methods lack the contextual reasoning capabilities and cannot be fitted to utilize additional information about circumstances, environments, backgrounds, etc. Such information provides essential knowledge about possible reasons for particular actions. This knowledge could not be processed directly by either machine learning methods. This paper presents the context-aware machine learning approach for actor behavior contextual reasoning analysis and context-based prediction for threat assessment.

Moreover, the proposed approach uses context-aware prediction to tackle the interaction between actors. An idea of the technique lies in the cooperative use of two classification methods when one way predicts an actor's behavior. The second method discloses such predicted action (behavior) that is non-typical or unusual. Such integration of two-method allows the actor to make the self-awareness threat assessment based on relations between different actors where some multidimensional numerical data define the connections. This approach predicts the possible further situation and makes its threat assessment without any waiting for future actions. The suggested approach is based on the Decision Tree and Support Vector Method algorithm. Due to the complexity of context, marine traffic data was chosen to demonstrate the proposed approach capability. This technique could deal with the end-to-end approach for safe vessel navigation in maritime traffic with considerable ship congestion.

Keywords: self-awareness, threat assessment, contextual knowledge, contextual reasoning, Decision Tree, Support Vector Machine, OPTICS.

1 Introduction

The main difference between machine learning methods and human is that humans are using contextual information for reasoning, not just statistical data. In general, context supplies information about situations within which something exists or happens, which can help explain it. Sometimes, it is difficult to precisely define the context because it is based on knowledge that is not explicit for a given problem. The context depends directly on the actors and their experience. Moreover, the same situation and context could be differently understood because the actors have a different experience in interpreting the situation. The knowledge that is relevant and could be used to understand the problem and help make a decision on threat assessment in a particular situation. It is defined as *contextual knowledge* [1]. Contextual knowledge may be merely tied to the goal or task. However, if the goal becomes clear, contextual understanding can be proceduralized for decision making [2]. Therefore, the decision-making process has a proceduralized context known by actors and is used for problem-solving. It is crucial because the same actions in different contexts could lead to different outcomes. The human can easily perceive and transfer the contextual knowledge to other humans, but the contextual knowledge is inappropriate for processing on the computer. Therefore, the computer does not operate with context as an additional source of information. Due to this issue, the context could be prepared partly as statistical data of the particular event or object. Nevertheless, the use of contextual knowledge is extremely important to understand and solve a real-world problem. Machine learning methods usually use numerical or visual data to train the model and make a prediction based on the trained model. Many different machine learning methods can predict the outcome based on learned patterns from data. On the other hand, machine learning methods do not utilize the perception of the situation because this understanding is gained from experience. Thus, the machine learning must cover every eventuality of possible situations. Conversely, the human can make a proper threat assessment based on his life experience.

The primary purpose of our research is to extract contextual knowledge from the statistical data set. Also, the proposed approach can make predictions based on extracted contextual knowledge. This approach predicts the possible further situation and makes its threat assessment without any waiting for future actions. Moreover, the proposed approach has the *Self Awareness* ability to make a threat assessment before any further action is completed. This ability allows for distinguishing between normal and unusual activities. Also, it can be adopted from the point of view of any actor. This ability enables predicting and considering the other actors. This "understand "comes before any indications of possible outcomes and predicts the current situation. Moreover, the proposed approach could evaluate its decision on the threat assessment and support system's proposed decision to consider it on the same basis. This *Self Awareness* ability makes an effective decision-making process where the machine learning method could evaluate the threat assessment through its own or other actors' views in different environmental contexts.

Our research is primarily inspired by a complicated and dangerous marine traffic environment. The marine traffic operates in significant density and on large scales. The vessel collisions could lead to enormous losses of human life and property. Therefore, the issue of ensuring safety at sea is the main goal for the mariners. This task could be solved by providing timely, essential navigational information. On the other side, this navigational information could not be enough to avoid collisions or dangerous situations. Therefore, the safe trip with the ship highly depends on the experience of navigators. The experienced navigator can predict a hazardous condition based on his long-term experience. So, the best way to prevent vessel collision is to anticipate all vessel possible routes and actions of the maneuver. At first, the navigators do it in the old fashion way by keeping a visual watch on the ship bridge. In this technology century, navigators do not have a sophisticated way to minimize collision risk. They are using radars, Global Positioning System (GPS), Automatic Identification System (AIS) technologies. Therefore, the threat assessment is obligatory to their own experience and attention. The same issue appears at developing any kind of autonomous vehicle. So, the navigators' main task is to clearly understand the current *context* and make a threat assessment based on their own experience, contextual knowledge, and predictable maneuvers of all actors before it's too late.

In this paper, we create an approach that could predict the possible further situation and makes its threat assessment without any waiting for future actions. It is essential for a safe autonomous navigation level in an unknown situation. The proposed approach consists of two different machine learning approaches. At first, the machine learning algorithm predicts the possible route based on historical data. Then, the other machine learning algorithm evaluates the prediction for an anomaly before any action in the future marine traffic situation. In this case, we get the contextual reasoning, and it's the threat assessment based on contextual knowledge.

2 The context of the marine traffic

Ensuring the safety of the vessel in marine traffic is a challenging task. Vessel safety depends on many different factors. The complexity of safe ship maneuvering relies on the experience of the navigating officer. As shown in Figure 1, the officer in charge should consider many different factors. These factors could influence the decision differently, and it is difficult to determine the significance of these factors. It depends on a particular situation.

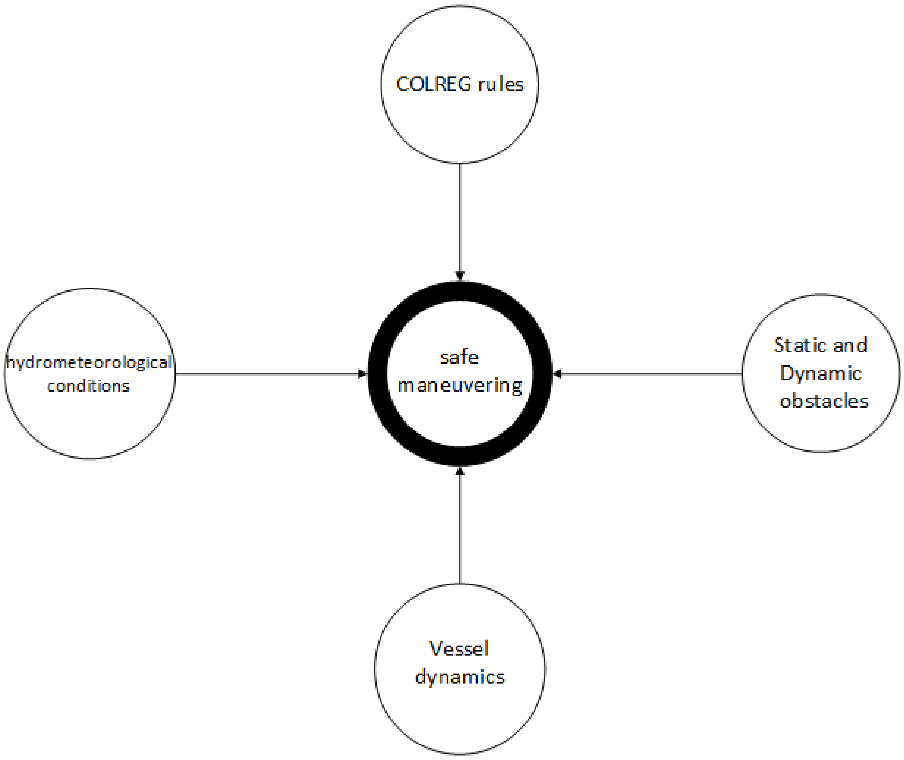

Figure 1: The factors influencing the safe ship maneuvering

The marine traffic rules are described in the Convention on the International Regulations for Preventing Collisions at Sea (COLREG) [3]. All the vessels must navigate according to the COLREG rules to prevent the collision [18]. Also, the officers must take into account the hydro-meteorological conditions and other static or floating obstacles on the route.

The vessels use many different sensors from complex, dynamic, and technical systems to make a full picture of the environment in which vessels are navigating. Nerveless, the many complicated factors influence the assessment of a particular situation.

Despite the progress of technology, the main reason for the vessel collision is human errors. The human must decide on how to navigate and maneuver safely. Therefore, safe navigation depends on the alertness of the navigation officer and his experience. So, the navigation officer's main task is to predict the next movement of the forthcoming vessel and make a decision that depends on the situation and environment. Moreover, the navigation officer's vital duty during the whole voyage could take up to nine months.

3 Recent machine learning methods that ensure the safety of marine traffic

There are developed many systems to ensure the decision on safe marine navigation. Such systems are essential not only for human manned ships but also for future autonomous vessels. These systems are designed to find the best navigation strategy based on marine traffic data. Such an expert system was proposed in [4] to avoid the collision. Another method was proposed in [5]. It generates the optimal multiple interval motion plan. Authors of [6] proposed the combined fuzzy inference system with an expert approach to avoid collisions. In [7], the reasoning system for the ship's turning problems is proposed. This method was applied for ocean navigation and collision. The intelligent decision-making facilities are on the rise to increase marine navigation safety. Also, the radar filtration algorithm for the Vessel Traffic Centre (VTC) is proposed in [8] to raise awareness. Also, the anomaly or dangerous situation detection is important to increase marine navigation safety [9].

Some studies attempted to apply Artificial Neural Networks (ANN). ANN models are learned by navigation officer actions to predict the maneuver actions [10], [11]. Also, [12] tried to predict the effect of the waves on the vessel yaw motion. The big issue for ANN is to predict the vessel collision or route. Haris and Amdahl [13] proposed the simplified analytical method. The authors used the interaction between the deformation on the striking and the struck ships. The ANN was used to train the machine learning model to predict marine traffic instead of developing the mathematical model. [14] used machine learning methods combined with DBSCAN based algorithm to find the marine traffic patterns for ocean-going ships. Kim et al. [15], [16] proposed the method to predict ship trajectory in the harbors. They trained the regression model combined with the dead reckoning model. The trained ANN was used to predict marine traffic in the chosen trajectory cluster. In [17], turning point based clustering was proposed for path prediction. In this model, the multilayered perceptron was used. The ship information such as speed, course, length, and position was taken in training, and output was the next turning point of the ship.

4 Proposed approach

A novel approach for maritime contextual reasoning analysis and context-based threat assessment has been proposed and detailed below. The proposed approach is based on machine learning and provides a complex insight into the context of vessel maneuvering situations and makes a threat assessment without any waiting for future actions. It could ensure an in-depth situation analysis through the same navigational data.

Let us consider a maneuvering situation consisting of two vessels. The first one is its own vessel. The second one is some other vessel that is possibly dangerous for the collision. Decision support aims to make maneuvering safely.

The proposed approach consists of the integral use of such main steps:

1. Prediction of the future turning point of another vessel.
2. Evaluation of the prediction whether it will be a normal action or some unusual action.
3. In the case of predicted unusual action and if the distance between own and other vessels is sufficient for safe maneuvering, update the navigational data, and go to step 1. Otherwise, the decision of the officer in charge of its vessel should be based on the rules described in the Convention on the International Regulations for Preventing Collisions at Sea (COLREG).

Steps 1 and 2 operate with marine traffic data that are obtained in real-time. This functionality is provided by the AIS. All participants of the marine traffic are responsible for sending the correct navigation data. The navigation data is sent automatically to all participants of marine traffic. The AIS data is a mixture of static data (vessel name, maritime mobile service identity (MMSI), vessel type, dimension, the port of destination) and kinematic data (vessel position – latitude & longitude, speed (SOG), course (COG, HDG), rate of turn (ROT)). These data are broadcasting every 2-10 seconds, and it depends on the vessel type and speed. This broadcast generates a large volume of navigation data. For example, the vessels in the Danish coast area generated about 60 GB per month.

To handle such huge data, Steps 1 and 2 may be realized using machine learning methods. However, this data is too large for machine learning and needs some preprocessing that leads to data aggregation. Data aggregation in our case means clustering of turning points of vessels – we reduce the data by using a set of clusters of turning points instead of a whole set. Now, the whole cluster of turn points is represented by a center of the cluster. DBSCAN or OPTICS may be used to cluster such spatial data and to get a set of generalized turn points. A more detailed description of data aggregation is presented in Section 5.

Denote the current position of the other vessel by A. Its nearest aggregated turn point (center of the cluster of turn points) is B. Point B is on the course line from A. Point C must be predicted for threat assessment. An example of the vessel route through points A, B, and C is shown in Figure 2.

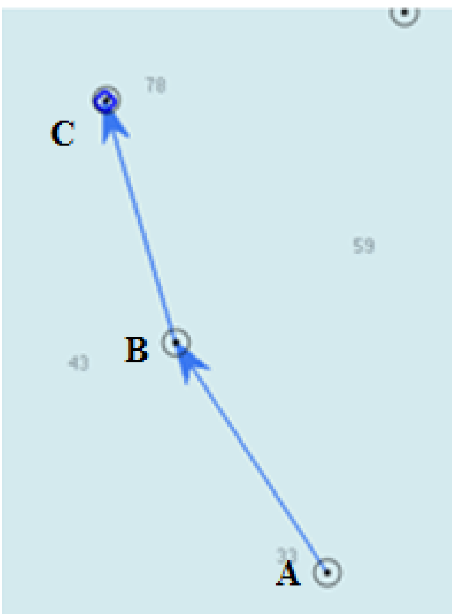

Figure 2: Example of the part of vessel route

The prediction of the turning point (C point) depends on SOG and COG. Due to economic reasons, it is an advantage to make small-angle turns on the route. There are many options for point C because it depends on the vessel port of destination, the situation of maneuvering, weather conditions, etc. To estimate or predict position C, we need some classifier that estimates the most possible aggregated turn point C from historical marine traffic data when the vessel moves from A to B. Decision tree (DT) may serve as such classifier.

The next step is to evaluate possible tracks predicted by DT, i.e. to perform a threat assessment. The result is the answer to whether the predicted route is normal or unusual. Therefore, we need a classifier that separates unusual actions from normal ones. Support Vector Machine (SVM) One-Class classifier may be applied here. The unusual actions could appear in dangerous maneuvering situations. However, unusual actions may be dangerous, too. For example, when vessels are engaged in fishing or sailing vessels to proceed in angular track due to weather or tide conditions. Therefore, unusual situations could not be extracted directly from historical marine traffic. The reason is that the unusual situation does not have any particular pattern, and it could not be abstracted from the whole dataset.

Our approach combines two different machine learning methods to make a threat assessment in the vessel movement. The solution is based on the *context-knowledge*, which was learned from historical marine traffic data. Remarkably, two different machine learning methods are solved the two various *context-based* problems on the same historical dataset.

Consequently, these two different approaches fully complement each other, and it is firmly connected. However, both methods complement each other for better prediction. The proposed approach ensures verification of the DT prediction. It means the *Novelty detection* (One-Class SVM) algorithm could evaluate DT prediction (point C) before the vessel appears in point B. If DT predicts an unusual next turning point C, the system could return to DT with new updated navigational data for the new prediction. This could be repeated as many times as necessary or while it is safe. The navigation situation changes and the navigational data could be updated every several seconds, for example, from the AIS device. This could provide constant surveillance and threat assessment of marine traffic based on contextual reasoning.

The details on possible methods that are combined in the new approach are presented in Sections 5, 6, and 7.

5 Marine traffic data aggregation

Marine traffic generates a big amount of navigation data. Consequently, it is hard to aggregate or process navigational data. The collected data have recorded both normal maneuvering and unusual situations – dangerous maneuvering, not usually behavior or misconduct. However, it is hard to exclude such data because the unusual situation condition does not have a specific pattern.

The first step to relieve the aggregation of the useful navigational data is preprocessed. This allows us to use a big set of navigational data but in a much smaller data volume without losing main patterns. There are various algorithms applied to preprocessing the AIS data. Density-based algorithms are used for spatial data, as a rule. This family's most known algorithm is the Density-based spatial clustering of applications with noise (DBSCAN) [18] algorithm. DBSCAN has some ϵ radius parameters defined in an empirical way. It is a big problem because a small change in this parameter can produce very different results in varying density datasets. Moreover, this is a big issue for real-world, high-dimensional datasets. To overcome these problems, the OPTICS was proposed. The OPTICTS [19] (Ordering Points To Identify the Clustering Structure) is the successor of DBSCAN.

At first, the OPTICS algorithm helps to reduce the size of the historical marine dataset by data aggregation because the result of the OPTICS is a clustered dataset. In addition, the OPTICS filters the meaningful data from the whole dataset. The cluster helps to recreate the generalized route from port A to port B. DBSCAN algorithm has two control parameters: a radius of the neighborhood for some point (ϵ), and the minimum number of neighbors (data points) within radius ϵ (*MinPts*). DBSCAN classifies data points into three groups: core points, reachable points, and noise or outliers. A point p is a core point if at least *MinPts* points are within distance ϵ from it, including this point itself. A part of reachable points is reachable directly. A point q is directly reachable from p if point q is within distance ϵ from core point p. Points of the cluster may be reachable directly from core points, only. A point q is reachable from p if there is a path of points $p_1, ..., p_n$ with $p_1 = p$ and $p_n = q$, where each p_{i+1} is directly reachable from p_i. Note that all points on the path are the core points, with the possible exception of q. All points that are not reachable from any other point are outliers or noise points. Each cluster contains at least one core point.

The OPTICS, like DBSCAN, requires the same two parameters: ϵ and *MinPts*. In contrast to DBSCAN, OPTICS also considers points that are part of a more densely packed cluster. OPTICS has two additional parameters: a *core distance* and *reachability distance*. Each point is assigned a *core distance* that describes the distance to the *MinPts*th closest point. The *reachability distance* of another point o from a point p is either the distance between o and p, or the *core distance* of p, whichever is bigger.

In our case, OPTICS produces the ordered list of the turning points. The turning points are connected by their reachability. OPTICS tries to reduce the reachability of each turning point during processing. Every unprocessed turning point is considered as a new cluster. The neighbors of a new cluster are added to the queue sorted by smallest reachability if their reachability can be reduced.

OPTICS does not create a cluster itself but produces an output list of the turning points with their reachability. The output list of turning points is ordered by the processing queue. The first processed turning point is the first at the output list. Respectively, the last processed turning point is the last in the output list. The cluster is extracted from the output list by identifying valleys in the point reachability. If ϵ is established, it is used to specify a cluster, otherwise, each separated network will constitute a cluster.

In the first step, the historical marine traffic data are preprocessing by the OPTICS algorithm. This aggregated marine traffic data was used in the DTs algorithm and One-Class classification. The clustered dataset extracts the marine traffic information – the turning points on main routes. The extracted and clustered turning points could be considered as contextual knowledge. The clustered data provides insight into vessel maneuverings. These points will be used in further calculations. This contextual knowledge contains the mariner's experience and the generalization of routes.

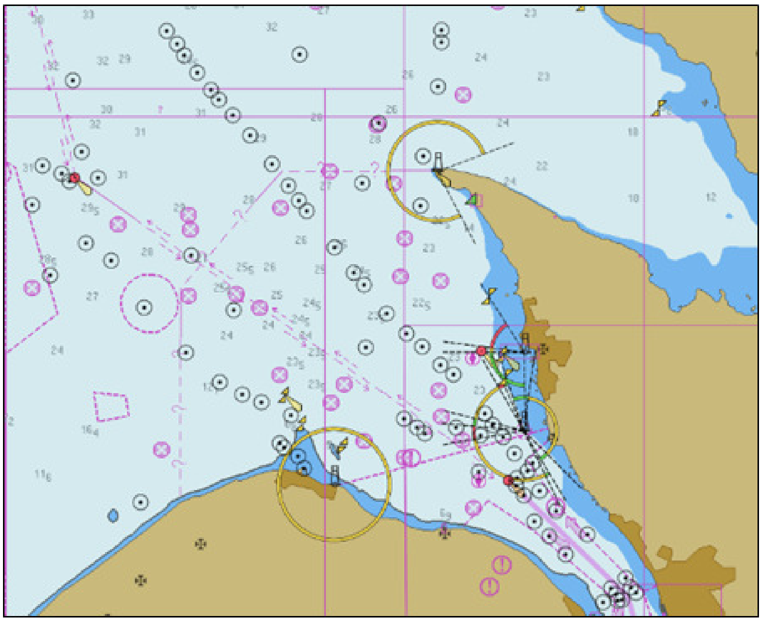

Figure 3: The result of the OPTICS algorithm

As shown in Figure 3 the OPTICS clustering eliminates the mediate waypoints and analyzes the turning points as a reference point for vessel moving behavior prediction. The proposed algorithm does not require every turning point, but it processes the center points of the clusters. Thus, the centers of the marked positions are the calculated geographical center of the clusters. This operation makes it much simpler to process the whole marine traffic dataset. As mentioned before, the marine navigation data accumulates large amounts of various navigational data. The historical marine traffic dataset contains information about vessel position at a particular time and its speed and course.

The particular route from port A to B every time could be slightly different. It could depend on maneuvering situations, hydrometeorology conditions, crew shift, etc. These reasons make the navigational data complicated to extract from the data navigator's experience and further research it. Another significant issue is the required processor time on raw data. The OPTICS algorithm solved these issues. The center points of clusters appear on route waypoints, and it could easily be identified as tracks.

6 Prediction based on Decision Trees algorithm

The DT algorithm is a method for non-parametric supervised machine learning. This method could be used for pattern recognition and classification problems [20], [21], [22]. The DT process the data by looping partitions to achieve the uniform classification of the target value. At each split, the DT reduces the entropy of the target value. This is achieved by adjusting the optimal split of independent variables. The DT has unique abilities to create a sequence of decisions. This sequence makes very fast problem solvers. Moreover, the DT sequences are very easy to understand and interpret. However, DT is hard to optimize. So, creating the DT is the process to find the structure of knowledge.

In our case, DT is used to learn the routes by data aggregated using OPTICS. DT primary purpose is to predict the future turn point (point C in Figure 4). The main problem is that DT has a wide choice of possible turn points, but it should recommend the most probable one (C). The data item for DT training consists of six features: (x1, x2) the current vessel position (Latitude & Longitude) describing the point A, (x3, x4) the kinematic data (SOG – speed over ground, COG – course over ground), (x5, x6) the next position (Latitude & Longitude) describing the point B, the target data (Latitude & Longitude) describing the point C. See Figure 4 for an example of training data.

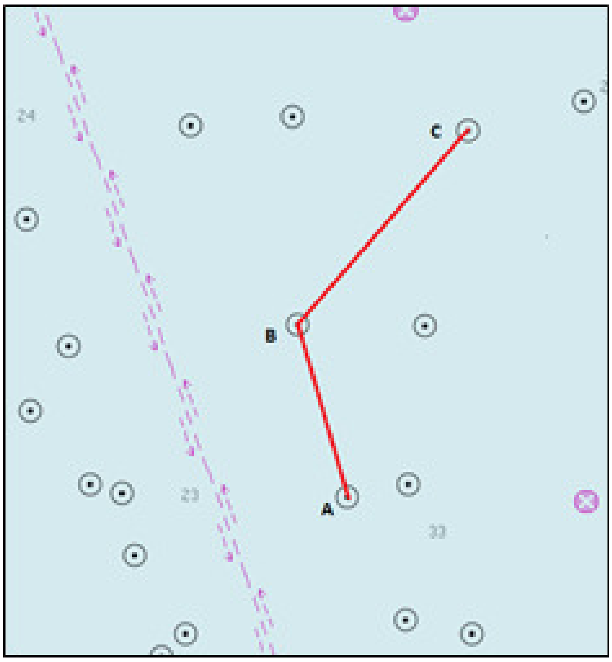

Figure 4: Visual presentation of training data – an example

DT tries to learn the context of marine traffic. The trained model is devoted to a particular geographical location because of the data used for learning.

The DT has a *maxdepth* parameter. The DT with a *maxdepth* size of 3 is shown in Figure 5. However, is not enough often to train the model properly due to data complexity. Moreover, too small *maxdepth* size may decrease the number of features used for decision. DT is realized in the integrated tool scikit-learn [24]. The DT is learning by splitting nodes with samples. It starts from the Root Node (with 393124 samples in our case). The mean squared error (mse) is used to split the node into sub-nodes. DT calculates mse for each divided subset and chooses the result with the smallest mse value. Each sub-node consists of a different number of samples.

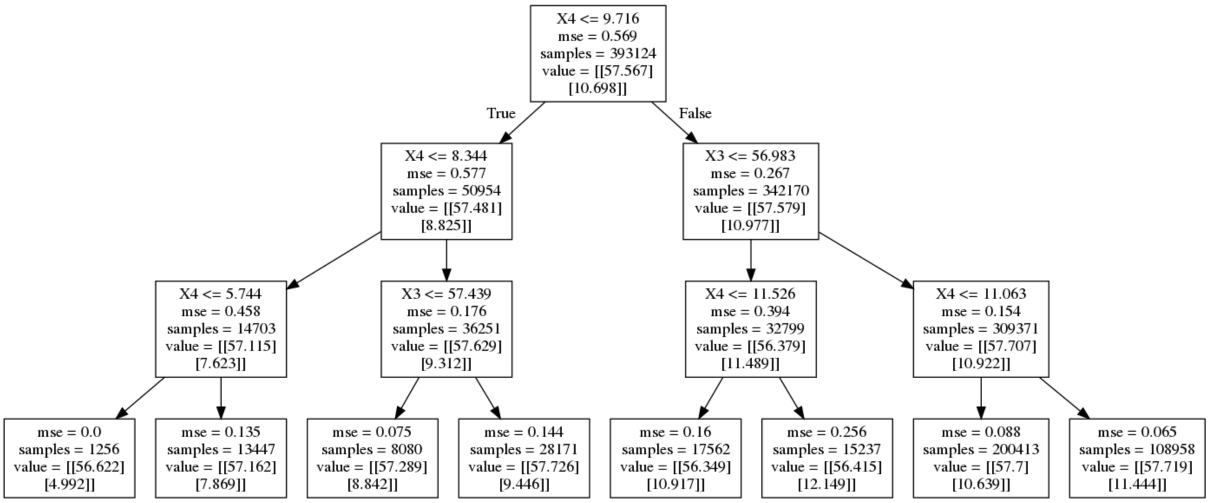

Figure 5: Example of the DT

7 Novelty detection based on SVM

There are developed several different methods that involve the navigation experience to predict the next waypoint. These methods are based on historical marine traffic data. One of the problems is to filter the anomalies, other unusual events, or dangerous situations. The simplest way to filter such data is to search and analyze every accident or dangerous maneuvering. It is challenging for humans because he must find or have prior knowledge about such cases or filter and extract the respective data by hand.

Despite this, the Novelty detection principle was used to tackle this task. Novelty detection is the ability of an intelligent organism to identify an utterly unknown pattern. The pattern should be sufficiently salient or associated with high positive or negative usefulness. This principle came from neurophysiology proposed by E. N. Sokolov [23] in the 1950s. This method is widely used to detect anomalies or outliers. The principles of Novelty detection are given in [24] and [26].

The SVM One-Class algorithm may serve for Novelty detection, i.e. to filter the dangerous or unusual maneuvering situations from navigation data. In general, SVM arranges the dataset on a hyperplane or set of hyper-planes in a high or infinite-dimensional space. This space is used for classification, regression, or other tasks. The best results could be achieved by the hyperplane, which has the most considerable distance to any class's nearest data points (operating margin). The SVM One-Class algorithm learns to classify new data as similar or different from the training set. [26] has suggested a method to solve the One-Class classification problem using SVM.

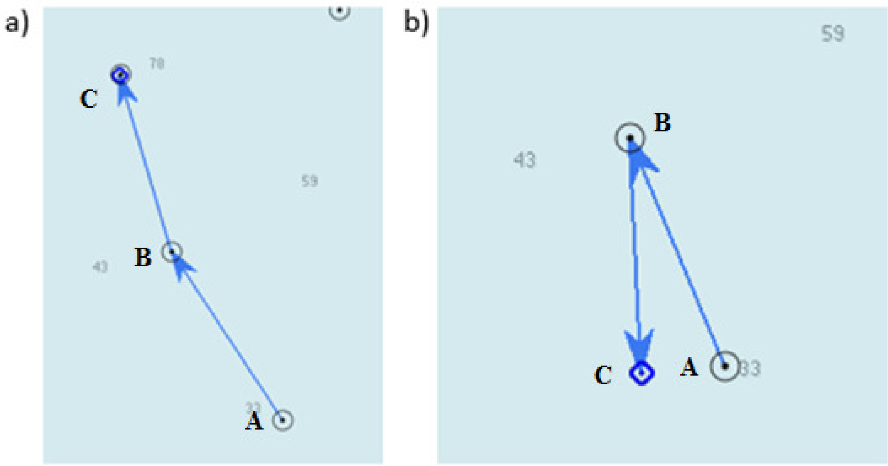

Figure 6: The comparison of normal (a) and unusual (b) vessel track

The typical vessel path should be low-pitched with the lowest possible angles, as shown in Figure 6a. There are many practical reasons for it – fuel economy, big size of vessel, inertia &, etc. Notwithstanding, there could be many reasons for an unusual path, as shown in Figure 6b. For example, it could be the sailing vessel going by changing downwind, the fishing vessel with nets, law enforcement vessel on duty &, etc.

As mentioned before, the DT is trained and prediction is based on the historical vessel routes. The One-Class SVM algorithm was trained for a completely different task. SVM should evaluate the DTs prediction, i.e. the future turning point C predicted by DT should be assessed by SVM.

The data item for SVM training consists of (1) the distance dist1 between points A and B, (2) the distance dist2 between points B and C, (3) the angle ∠ABC. See Figure 7 for an example of training data.

$dist_1$	$dist_2$	∠ABC
6174.52	1287.87	151.20
671.35	121.82	129.92
1526.95	1590.68	57.44
1024.27	24.30	82.41
6029.97	44602.39	163.37
...	...	...
2500.94	310.07	100.69
2762.53	1864.93	144.60
8158.88	24575.09	172.39
4293.26	20709.93	172.25
5300.63	4825.68	160.69

Figure 7: Example of One-Class SVM classification

One-Class SVM [26] is realized in the integrated tool scikit-learn [27].

8 Training and results of the proposed approach

The proposed approach in Section 4 consists of the integral use of two different machine learning methods. We need to train the corresponding classifiers. As it is shown in Figure 8, the first we need is to get historical marine traffic data. The next is the aggregation of a dataset for better training results. The historical marine data are aggregated by OPTICS and, as a result, a dataset of the centers of clusters of turning points is generated. After this, the training process is divided into two main parts:

1. DT is trained to predict the future next point.
2. One-Class SVM classifier is trained to detect unusual patterns. This step needs additional clustered data processing:
   (a) Calculate distances between pairs of turning points (A, B) and (B, C);
   (b) Calculate the angle ∠ABC.
3. The trained models are saved for future usage.

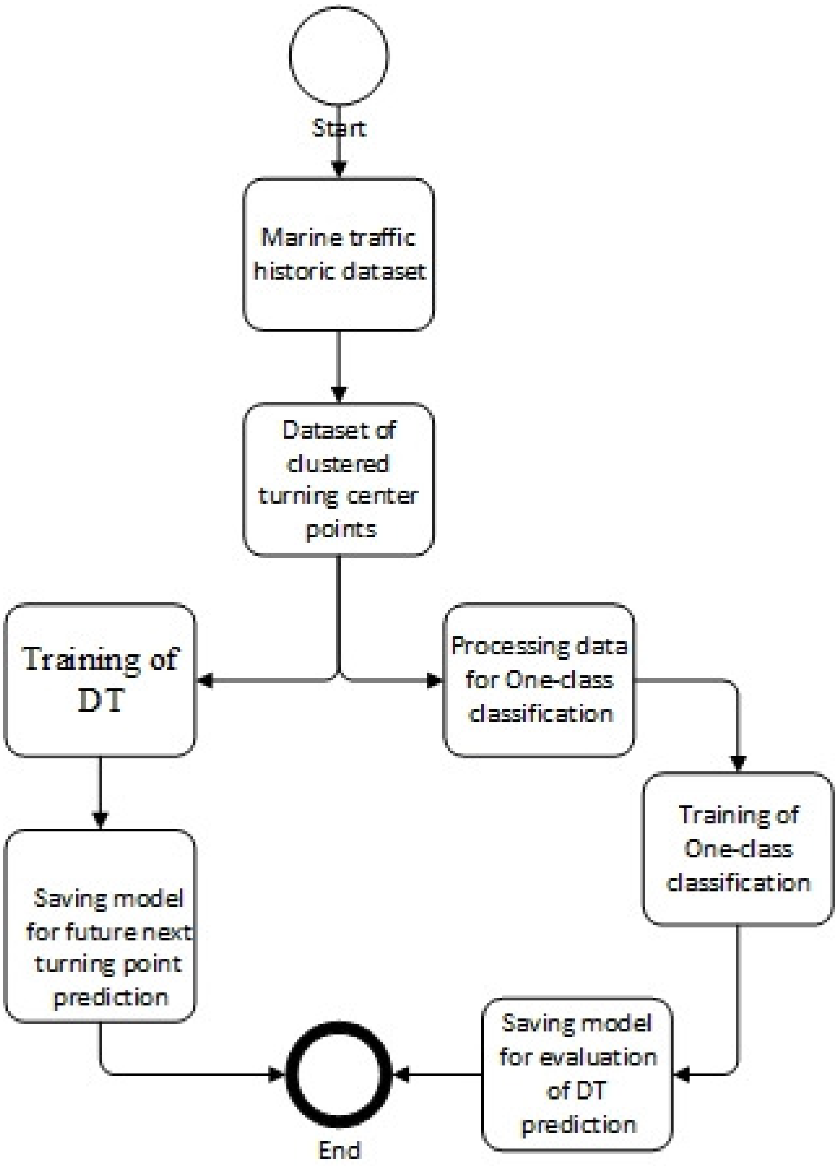

Figure 8: The training process

The main aim of training DT is to learn a particular traffic pattern in a particular geographical area. It is quite a complex problem because the DT method should predict the future next turn point (C) by the current position (A) of the vessel, its course, and speed. The experiments were carried out to evaluate the impact of parameter *maxdepth* on the prediction accuracy. *maxdepth* values were varied from 3 to 37. The results are given in Figure 9. The DT with a *maxdepth* equal to 26 reaches more than 90% accuracy. The maximum accuracy is 94% as *maxdepth* equal to 37.

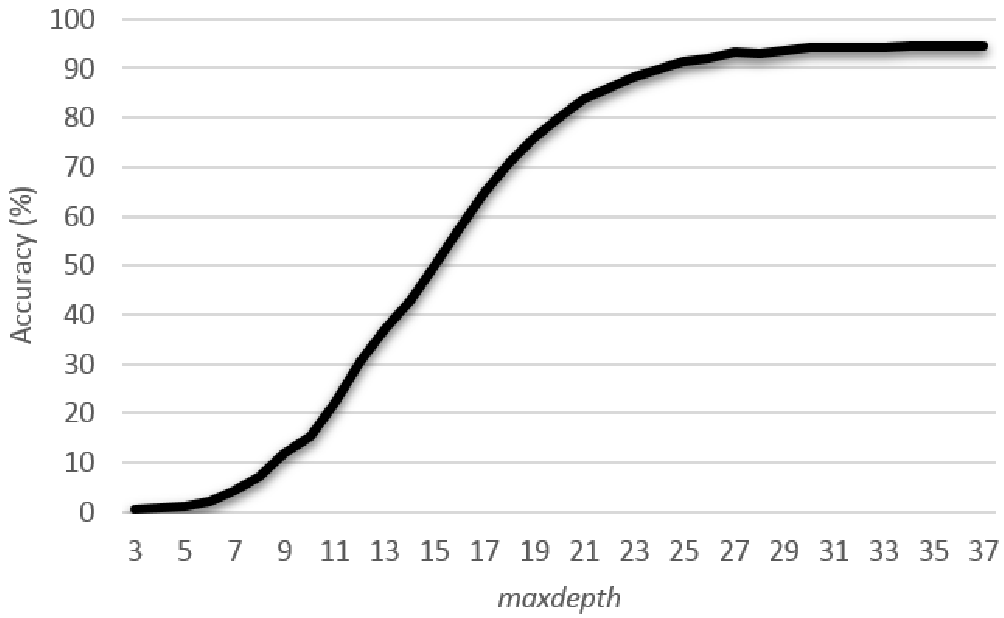

Figure 9: The precision of the DT model with different *maxdepth*.

An example of an anomaly detected by the One-Class SVM classifier is given in Figure 10. In general, the One-Class SVM classifier assigned 5,86% of DT predictions as unusual – 27125 cases from 462499.

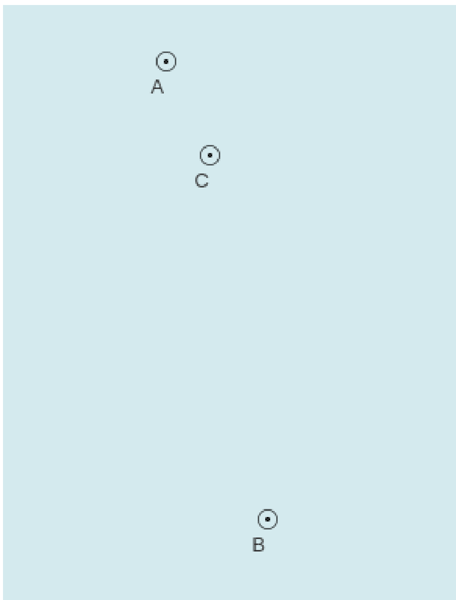

Figure 10: An example of an anomaly detected by One-Class SVM classifier.

9 Discussion

In this paper, we presented a context-based threat assessment approach. It combines two different machine learning methods. This approach predicts the possible further situation and makes its threat assessment without any waiting for future actions. This threat assessment is crucial because it could provide contextual reasoning in maneuvering before any threatening indicator occurs in any particularly maneuvering situation. The approach performs combined contextual reasoning over a holistic assessment of predicted possible future actions. It could be applied to any participants in the marine traffic because it allows analyzing unusual or dangerous actions in a volatile context before any indications of dangerous action occur.

Moreover, the proposed approach could be used by different actors in the same way for threat assessment. It allows us to evaluate possible future actions from different points of view:

1. Own vessel;
2. Another vessel;
3. Outsider – authorities who are responsible for safe marine traffic.

Employment of such different points of view makes the proposed approach suitable for the imitation of some basic cognition processes. The approach allows imitating such cognitive processes as knowing, remembering, judging, and problem-solving. These cognitive processes ensure imagination, perception, and action planning. The approach is an extremely universal model adjusted for marine traffic. In our case, the prediction of a possible future turning point acts as an imagination. It is based on contextual reasoning and provides guessing the possible action of another actor. As we know, perception allows us to take information and utilize it for responding and interacting. Just like that, the One-Class SVM classifier makes a threat assessment on the possible future activities. Thus, the classifier acts as a perception. In the case of a dangerous or unusual situation, the proposed approach could be applied to make a good strategy to avoid it. The strategy could be done by applying the proposed model from other vessel position and make a threat assessment on our self-actions. However, the approach is limited to marine traffic. The safe marine traffic could be achieved by applying the proposed approach in a decision support system that does the threat assessment without any human intervention in the permanent assessment process.

Context-based knowledge and reasoning are essential to make the right decision. Understanding of context is crucial to ensure safe maneuvering and navigation in different and completely unknown situations. This ability is very important for the future autonomous vessel. The unique feature of the proposed approach is the ability to analyze information from different perspectives of actors and it allows to evaluate our planned actions. The traditional machine learning methods do not ensure a complete analysis of the picture in the maneuvering and environment. So, it is incredibly essential to understand and analyze the context information as is. As it was demonstrated, the context-based knowledge provides a full analysis of the particular situation for a better decision.

Author contributions. Conflict of interest

The authors contributed equally to this work. The authors declare no conflict of interest.

References

[1] Rosenfield, I., Clancey, W.J. (1991). Book review "The Invention of Memory: A New View of the Brain", *Artificial Intelligence*, 50, 241–284 (1991).

[2] Brezillon, P., Pomerol, J.-C. (1999). Contextual knowledge and proceduralized context, *In: AAAI Workshop on Modeling Context in AI Applications*, Paris. 16–20 (1999).

[3] The International Regulations for Preventing Collisions at Sea 1972 (COLREGS). *International Maritime Organization (1972)*, 1972.

[4] Fan, Y., Sun, X., Wang, G. (2019). An autonomous dynamic collision avoidance control method for unmanned surface vehicle in unknown ocean environment, *International Journal of Advanced Robotic Systems*, 16 (2019). https://doi.org/10.1177/1729881419831581.

[5] Geng, X., Wang, Y., Wang, P., Zhang, B. (2019). Motion plan of maritime autonomous surface ships by dynamic programming for collision avoidance and speed optimization, *Sensors*, 19 (2019). https://doi.org/10.3390/s19020434.

[6] Ahn, J.H., Rhee, K.P., You, Y.J. (2012). A study on the collision avoidance of a ship using neural networks and fuzzy logic, *Applied Ocean Research*, 37, 162–173 (2012). https://doi.org/10.1016/j.apor.2012.05.008.

[7] Dixena, D., Chakraborty, B., Debnath, N. (2011). Application of Case-based reasoning for ship turning emergency to prevent collision, *In: 2011 9th IEEE International Conference on Industrial Informatics*, Lisbon, IEEE. 654–659. (2011).

[8] Bukhari, A.C., Tusseyeva, I., Lee, B.G., Kim, Y.G. (2013). An intelligent real-time multi-vessel collision risk assessment system from VTS view point based on fuzzy inference system, *Expert Systems with Applications*, 40, 1220–1230 (2013). https://doi.org/10.1016/j.eswa.2012.08.016

[9] Daranda, A., Dzemyda, G. (2020). Navigation decision support: Discover of vessel traffic anomaly according to the historic marine data, *International Journal of Computers, Communications and Control*, 15, 3, 3864 (2020). https://doi.org/10.15837/IJCCC.2020.3.3864

[10] Papenhuijzen, R., Stassen, H.G. (1992). Fuzzy set theory for modelling the navigator's behavior, *IFAC Proceedings Volumes*, 25, 53–59 (1992). https://doi.org/10.1016/S1474-6670(17)50170-6

[11] Westrenen, F. V. (1995). Towards a decision making model of river pilots, *IFAC Proceedings Volumes*, 28, 217–222 (1995). https://doi.org/10.1016/S1474-6670(17)46728-0

[12] Nicolau, V., Aiordachioaie, D., Popa, R. (2004). Neural network prediction of the wave influence on the yaw motion of a ship, *In: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, Budapest, IEEE. pp. 2801–2806 (2004)

[13] Haris, S., Amdahl, J. (2013). Analysis of ship–ship collision damage accounting for bow and side deformation interaction, *Marine Structures*, 32, 18–48 (2013). https://doi.org/10.1016/j.marstruc.2013.02.002

[14] Xiao, Z., Ponnambalam, L., Fu, X., Zhang, W. (2017). Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors, *IEEE Transactions on Intelligent Transportation Systems*, 18, 3122–3134 (2017). https://doi.org/10.1109/TITS.2017.2681810

[15] Kim, J.S. (2017). Vessel target prediction method and Dead Reckoning Position based on SVR seaway model, *International Journal of Fuzzy Logic and Intelligent Systems*, 17, 279–288 (2017). https://doi.org/10.5391/IJFIS.2017.17.4.279

[16] Kim, J.S., Jeong, J.S. (2017). Extraction of reference seaway through machine learning of ship navigational data and trajectory, *International Journal of Fuzzy Logic and Intelligent Systems*, 17, 82–90 (2017). https://doi.org/10.5391/IJFIS.2017.17.2.82

[17] Daranda, A. (2016). Neural network approach to predict marine traffic, *Baltic J. Modern Computing*, 4, 3, 483–495 (2016)

[18] Ester, M., Kriegel, H.-P., Sander, J., Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise, *In: Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, Portland, KDD'96., pp. 226–231 (1996)

[19] Ankerst, M., Breunig, M.M., Kriegel, H.P., Sander, J. (1999). OPTICS: ordering points to identify the clustering structure, *SIGMOD Record*, 28, 49–60 (1999). https://doi.org/10.1145/304181.304187

[20] Quinlan, J.R. (1986). Induction of decision trees, *Machine Learning*, 1, 81–106 (1986). https://doi.org/10.1007/bf00116251

[21] Wang, X., Chen, B., Qian, G., Ye, F. (2000). On the optimization of fuzzy decision trees, *Fuzzy Sets and Systems*, 112, 117–125 (2000). https://doi.org/10.1016/S0165-0114(97)00386-2

[22] Murthy, S.K., Kasif, S., Salzberg, S. (1994). A system for induction of oblique decision trees, *Journal of Artificial Intelligence Research*, 2, 1–32 (1994). https://doi.org/10.1613/jair.63

[23] Sokolov, E.N. (1960). Neuronal models and the orienting reflex, *In: Mary A.B. Brazier (ed.) The Central Nervous System and Behavior*, New York, Macey. pp. 187–276. (1960).

[24] Markou, M., Singh, S. (2003). Novelty detection: A review - Part 1: Statistical approaches, *Signal Processing*, 83, 2481–2497 (2003). https://doi.org/10.1016/j.sigpro.2003.07.018

[25] Markou, M., Singh, S. (2003). Novelty detection: A review - Part 2:: Neural network based approaches, *Signal Processing*, 83, 2499–2521 (2003). https://doi.org/10.1016/j.sigpro.2003.07.019

[26] Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C. (2001). Estimating the support of a high-dimensional distribution, *Neural Computation*, 13, 1443–1471 (2001). https://doi.org/10.1162/089976601750264965

[27] Pedregosa, F., Grisel, O., Weiss, R., Passos, A., Brucher, M., Varoquax, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Brucher, M. (2011). Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* , 12, 2825–2830 (2011).

Cite this paper as:
Daranda A.; Dzemyda G. (2021) Novel Machine Learning Approach for Self-Aware Prediction based on the Contextual Reasoning, *International Journal of Computers Communications* & *Control*, 16(4), 4345, 2021.

https://doi.org/10.15837/ijccc.2021.4.4345