Applicability of Artificial Intelligence and Machine Learning in the Maritime Industry and Port Management.

Pavel Skournik

Managing Director, Tidalis Americas LTD.

Author Note

Pavel Skournik is now Managing Director of Tidalis Americas LTD, formerly a Saab Group company.

This paper expresses the personal views of the author, with extensive acknowledgment and gratitude to Tidalis (formerly Saab Group, formerly Klein System Group) clients for their partnership and progressive commitment to advancing technology towards a safer, more efficient and effective industry. The author would like to note that Tidalis products today contain some elements discussed in this paper.

Correspondence may be addressed to: Pavel Skournik, Tidalis Americas LTD, 300, 3500 Gilmore Way, Burnaby, BC Canada V5G 0B8 Email: pavel.skournik@tidalis.com.

Abstract

Artificial Intelligence (AI) and Machine Learning (ML) have moved beyond a curiosity, into an experimental stage in seemingly all aspects of life, including the maritime industry. The maritime industry is one of the oldest and most traditional industries that still relies as much on human intuition as on data. This is due to the sheer size of the global network and therefore planning problems, and since there still are very different approaches, rules and habits by port, by region, by carrier and truly by any player in the process. On the other hand, the maritime industry has been slowly working towards adopting technology and optimizing processes for years and AI and ML provide a unique opportunity to achieve that goal. These technologies are slowly being adopted to assist with operational optimization, safety improvements, and to increase efficiency in general in the maritime industry and port management. Early proponents suggest the use of AI and ML in port management can improve the accuracy and speed of decision-making, which can in turn result in increased productivity and cost savings. If technology continues to evolve and improve, it is likely that the maritime industry will continue to find value in using AI and ML in even more areas, potentially leading to linked improvements and optimizations that were hard to even chase before. **Objective:** The objective of this paper is to introduce the concepts of AI and Machine Learning and specifically to their potential application in Port Operations, Planning, Scheduling, and Traffic Safety. The goal is to make a complicated and mathematically heavy subject easier to understand for a non-technical audience. There is a lot of information (and misinformation) on the internet, as well as from various vendors offering their solutions, and the goal of the paper is to help maritime professionals to make a more informed evaluation of the applicability and validity of the information or products in their own situation. The focus is the maritime industry in general and ports specifically, with applicability to anyone in the port, board directors, port CEO's, harbormasters and IT leaders.

Keywords artificial intelligence, AI, machine learning, ML, maritime technology, port management

Applicability of Artificial Intelligence and Machine Learning in the Maritime Industry and Port Management

Overview of AI and Machine Learning - History and Differences

Artificial Intelligence (AI) and Machine Learning (ML) are not new concepts, even though they are a vogue curiosity in everyday life. In fact, AI and ML have had decades of research and seen much evolution. AI is a broad term that simply includes work to develop computer systems that can execute specifically detailed tasks that would typically require human intelligence, such as visual perception, speech recognition, and decision-making. Machine Learning is a type of AI that involves the creation of algorithms that allow systems to learn and improve from experience without being explicitly programmed.

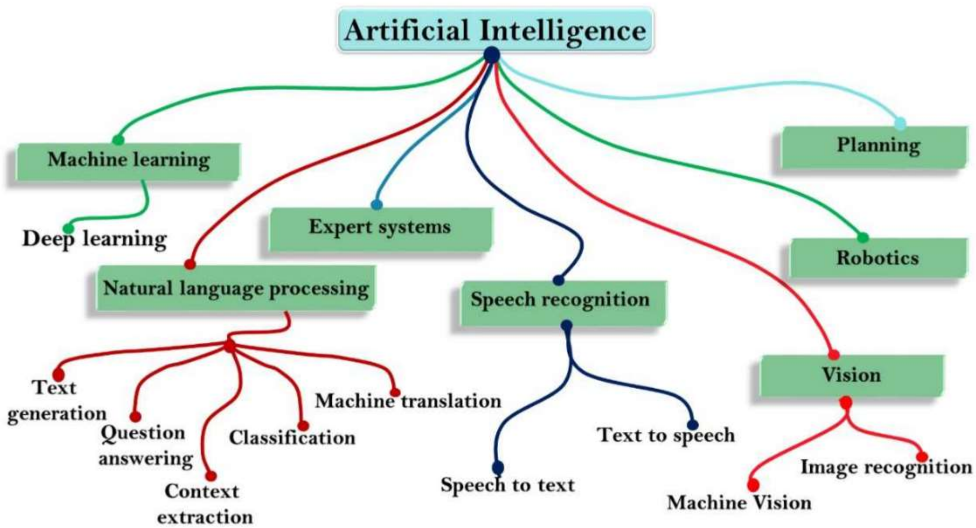
More technically, artificial intelligence is the science and engineering of teaching computers to behave in ways that in the past required human intelligence. AI includes a broad group of methods for computers to predict an outcome with a certain level of probability based on the available information and algorithms. Machine Learning is one of the methods we can use to achieve AI. The main benefits of AI in general, and ML specifically, are that computers can process enormous amounts of data exponentially faster than humans could ever achieve. The capabilities of ML algorithms combined with real human industry experience might create a great advantage for making better decisions or faster decisions.

The history of AI dates back to the 1950s (Wikipedia, 2023) when researchers first began exploring the concept of machines being able to perform tasks that would typically require human intelligence. AI started with the first algorithms for calculations and gaming software. It is still inconclusive which game is considered the first example of using AI. "OXO" (also known as "Tic-Tac-Toe") was developed in 1952 at the University of Cambridge; "OXO" was a computer program that allowed users to play the game against the computer. While the game itself is simple, the fact that the computer was able to play the game at all was a significant achievement at the time. It was considered by many as the first example of artificial intelligence. "Pong" was developed in 1972 by Atari; "Pong" included an actual simple AI algorithm to control the movement of the computer-controlled paddle. In the early 1980s, the first machine learning algorithms were introduced to fight against bank fraud and spam emails. Over the years, AI has evolved and expanded, enabled by the increased computing capacity of machines, and the advent of new technologies and algorithms, such as deep learning, which have significantly improved the performance of AI systems.

In the 2000s, deep learning started to evolve. It uses self-learning algorithms built on neural networks to improve performance and accuracy. A neural network is a type of machine-learning model inspired by the structure and function of the human brain. It consists of a large number of interconnected processing nodes called neurons, organized into layers. The neurons in the input layer receive raw data, such as images or text, and each neuron performs a simple mathematical operation on the input. The output from the input layer is then passed through one or more hidden layers, where computations that are more complex are performed. In the end, the output layer produces the final prediction or classification result.

Figure 1

Artificial Intelligence Family Tree



Neural networks should be trained in order to understand how to process data. Training is a process of assigning weights to each neuron. With time that leads to minimizing the difference between the actual output from the ML algorithm and the desired output. The process also uses an optimization algorithm to improve the overall network's performance on the training data through iterative enhancements. Neural networks can be very useful for many applications. They are performing great in image and speech recognition, NLP, and predictive analytics.

Machine learning (ML) become very popular in recent years, mainly because of the availability of large amounts of data and improved available computational power. With access to more data and computational power, models can be trained on larger and more complex datasets to provide better accuracy and faster performance.

For example, a neural network trained on a large dataset of images can attain higher accuracy in image recognition tasks than it was possible on the network trained on a smaller dataset.

Access to more data also leads to the development of more sophisticated machine learning models, such as deep neural networks with many layers or models that use reinforcement learning.

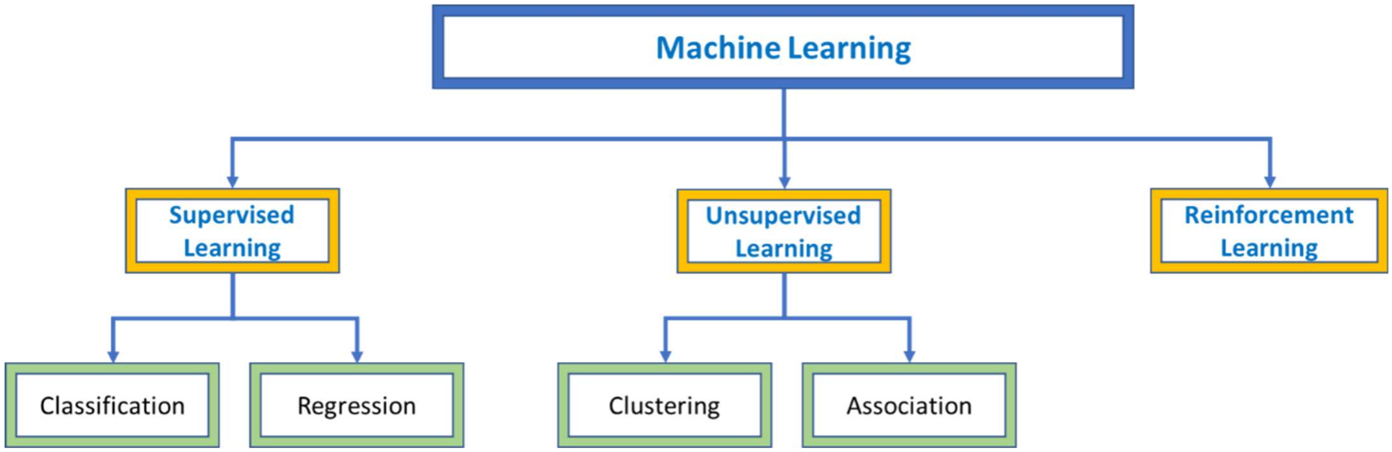
The primary goal of Machine learning algorithms is to learn from data and make predictions or decisions without being specifically programmed. You can see examples of that in areas such as natural language processing (ChatGTP is one of the most famous results) or computer vision recognizing license plates of moving cars.

Overview of different methods of Machine Learning

Machine learning in essence is the creation of algorithms enabling systems to learn from previous executions and analyze outcomes to improve. An algorithm can be as simple as the steps to follow a recipe, or a bit more relevant, the 'BEDMAS' rules that are the accepted order of operations to solve a complex mathematical question (brackets, exponents, division, multiplication, add, then subtract). Over the years, ML has evolved and expanded, with the advent of new technologies and algorithms that have significantly improved potential performance. Machine learning offers a range of techniques for building intelligent systems that can learn and improve from experience. The choice of method will depend on the nature of the problem, the amount of data available, and the desired outcome. By understanding the different methods of machine learning, the most appropriate approach for a given problem can be chosen to achieve better results.

Figure 2

Machine Learning Family Tree



Some of these methods (Castanon, 2023) are:

Supervised Learning: This is the most common form of machine learning, where the system is trained on a labelled data set to predict the output for new, unseen data. The system is set up with identified known data that is referred to as 'labelled' data. The labelled data is a base and can hopefully provide patterns or relationships. The patterns can be used as a template and applied to new data to make predictions. For example, a financial institution may use logistic regression or decision trees to detect fraudulent transactions based on input features such as transaction amount, location, and customer behavior.

Unsupervised Learning: In this type of machine learning, the system is expected to identify patterns and relationships in the data based on the unlabeled dataset. The system does not have a specific target output and focuses on discovering the structure and relationship of the data. A cybersecurity company may use anomaly detection algorithms, such as Gaussian mixture models or auto encoders, to detect unusual network activity or behaviour that could indicate a cyber-attack. The algorithms can learn normal patterns and identify differences that deviate from those patterns. A retail company could use unsupervised learning algorithms to analyse customer transaction data and identify items that are frequently purchased together. This may help the company improve product placement and recommendation systems.

Reinforcement Learning: The system learns by receiving "feedback". The feedback is provided in the form of rewards or penalties. The feedback is then used by the system to improve its predictions over time. Reinforcement learning is often used in situations when the system must make quick decisions to achieve a specific goal, for example, in computer games.

Semi-Supervised Learning: This type of machine learning combines elements of both supervised and unsupervised learning. A system is trained on a mixture of labelled and unlabeled data and uses the labelled data to make predictions but also is trying to make sense of the structure of the data. This method is often used when there is a shortage of labelled data.

Transfer Learning: This is an interesting type of machine learning since the system uses the results from one task to improve its performance on a new task. Transfer learning has been particularly useful in computer vision work, where the system can use the result it has learned from a previous image classification task to improve its performance on a new image classification task.

Considerations in selecting and implementing Machine Learning solutions.

There are risks with any technology project, and ML projects are no different. In order to be sure that in the end the new system is dependable, trustworthy and able to achieve the goals set out, the ML project should be reviewed for potential risks and mitigation strategies.

1.1. One of the risks associated with ML systems is the potential for biased outcomes. Since ML systems need to be trained on data, and if that data contains biases, the system may carry those biases into its predictions and decisions and create a completely different outcome without a user realizing it. It is important to review the training data to ensure that it is representative and unbiased to reduce this risk. It is also important to monitor the output of the system to detect any biases that may arise in the future.

1.2. Another point to consider is the lack of interpretability and transparency in ML systems. You can consider an ML system as a "black box" due to its complex algorithms and internal workings. It is often difficult to understand how it makes decisions. If the human users do not understand the inner workings of the system or cannot test enough to feel comfortable, this could mean a lack of trust and make it harder to find errors. To mitigate this risk, it is important to develop ML systems that are interpretable and transparent, allowing port personnel to understand how the system is making decisions.

1.3. Data privacy and security are also important considerations in implementing ML systems. (Koerner, 2022). Machine learning systems typically handle large amounts of sensitive data because they require access to data to learn from it and make predictions. The more volume the better. The more data that a machine learning system has access to, the better it can learn and make accurate predictions. However, this also makes machine learning systems vulnerable to data breaches. If an attacker could access a system, they could potentially steal or manipulate sensitive data, compromise the integrity of the system, or even use the system to launch attacks on other systems. In fact, machine-learning systems are vulnerable to certain types of attacks that are specific to machine learning, such as adversarial attacks. These are attacks that involve manipulating input data to cause a machine learning system to make incorrect predictions, and they can be difficult to detect and prevent.

To mitigate this risk, it is important to implement strong security measures and data protection policies, such as encryption and access controls, to protect the data and prevent unauthorized access. This is no different from normal good practice data protection for any sensitive data, however recognizing that the added level of security protocols may be necessary if the data is linked, pooled or aggregated. It is also important to use best practices for ML, such as training ML models on anonymized data whenever possible and implementing protocols to detect and prevent adversarial attacks.

1.4. One of the major risks is the difference between correlation and causation that is hard to understand for AI algorithms. Causation is a relationship where a change in one variable directly causes a change in the other variable. Correlation is a relationship between two variables, where a change in one variable may look like it could be caused by or related to a change in the other variable, but those variables are independent of each other.

The difference between correlation and causation is important because these systems are often trained on large datasets to identify patterns and relationships. If the system identifies a correlation between two variables, it may not necessarily mean that there is a causal relationship between those variables. For example, there is a strong correlation between ice cream sales and the number of shark attacks, but it does not mean that ice cream causes shark attacks. Both variables are changing in a similar way due to a third variable - the temperature. Usually, both numbers rise in summer and fall in colder times of the year.

If an algorithm mistakes correlation for a causation it will build the wrong predictive system that will provide inaccurate results. For example, if a correlation is identified between a vessel's speed and fuel consumption, it does not necessarily mean that reducing speed will lead to lower fuel consumption. There are other variables, such as weather conditions or vessel design, that may also play a role in the relationship between speed and fuel consumption but are not included in the prediction calculation.

1.5. The next point to consider is “accuracy”. (Munim etal, 2020). When looking into integrating ML systems into the workflow of the port the accuracy of the predictions or the calculations is critical. To achieve acceptable accuracy levels good quality and large amounts of training data should be used./ Also, it is important to validate that the calculation or prediction algorithms are properly designed and implemented.

An important first step to ensuring accuracy is the quality and quantity of data used for the training and validation of the algorithms. Such algorithms require large amounts of accurate and relevant data to make accurate predictions and decisions. In the maritime industry, there is a lot of data. The problem is that it is not always accurate or structured data. The data is often generated from multiple sources, like shipboard sensors, satellite imagery, and vessel performance data. However, this data may be incomplete, inconsistent, or outdated. Inconsistency of data will impact the accuracy of the algorithms. Some ways to improve accuracy are to collect additional data, clean and pre-process existing data, or use alternative data sources.

A second step to ensuring accuracy in ML in the maritime industry is understanding the complexity of the underlying processes and relationships being modeled. The maritime world involves complex and interrelated factors, such as weather conditions, vessel performance, and operational constraints that can affect the accuracy of ML algorithms. In order to make accurate predictions and decisions, ML algorithms must be designed and trained to consider these factors and their interactions. After design and training, the validation of the design must test to ensure the model captures the varying underlying complexities and especially the relationships between variables.

1.6. Another big challenge of ML and AI is that algorithms being designed often assume that the world does not change. It is a significant obstacle in their practical application and result in poor performance and incorrect predictions in real-world applications where the environment is constantly evolving. To overcome this challenge, it is important to regularly update and retrain the models with fresh data, and to incorporate data from multiple sources. By addressing this issue, ports can ensure that their ML and AI algorithms remain accurate and relevant as natural factors keep moving.

In many cases, ML and AI algorithms are trained on historical data that represents a snapshot of the world at a particular point in time. This training data is used to build models that make predictions about future events or outcomes. However, if the world changes significantly between the time the data was collected and the time the predictions are made, the models may no longer be accurate. For example, in the maritime industry, ML and AI algorithms are now often being used to make predictions about traffic congestions, traffic patterns, terminal operation time and other similar variables. However, these algorithms can become quickly outdated if there are significant shifts in the global economy, geopolitical events, or global health situations such as was particularly significant with the COVID pandemic in 2020.

To mitigate the problem of ML and AI algorithms becoming outdated, it is important to regularly update and retrain the models with fresh data. This can help ensure that the algorithms remain useful and current in a constantly changing world. Additionally, it is important to incorporate data from multiple sources, including data from real-time data feeds, to improve the accuracy and adaptability of the models.

1.7. The implementation of AI and ML raises important ethical and cultural issues that need to be carefully considered. These issues need to be reflected on in a contextual nuanced way to guide the design and development of AI and ML systems and then provide ‘rules of engagement’ for their use. To ensure that AI and ML systems are used in a responsible and ethical manner, it is important to have clear policies and regulations in place, and to consider the cultural attitudes and values of the populations that will be impacted by these systems.

One of the key ethical issues in AI and ML is bias. AI and ML algorithms are only as unbiased as the data they are trained on, and if the data contains biases, the algorithms will reflect those biases in their predictions and decisions. For example, if an AI system is trained on data that reflects racial or gender biases, it may make discriminatory predictions or decisions.

Another important ethical issue is privacy. AI and ML algorithms often require access to large amounts of personal data (for example cargo manifest contains commodities linked to consignee personal information) which can raise concerns about data privacy and security. In order to ensure that AI and ML systems are used in a responsible and ethical manner, it is important to have clear policies and regulations in place to protect personal data.

In addition to ethical issues, there are also cultural issues that need to be considered in the implementation of AI and ML. For example, different groups of people may have different attitudes towards technology and automation, which can affect the adoption and use of AI and ML systems. , It is very similar to the early days of any automation initiatives, with unions and workforces in general opposing reduction in jobs. The use of AI and ML tools will not necessarily reduce the need for humans, as some use cases are quite unrelated to reducing human effort, as in example of estimated time of arrival calculations. Additionally, different people may have different values and norms that need to be considered in the design of AI and ML algorithms that will provide the desired result.

Examples of applicability of Machine Learning to Port Management and Port Operations

On the surface, port operations seem like a natural fit for AI and ML for two key reasons. First, if a port has any historical records that are already digital or can be digitized, the historical data can be used to train the AI/ML algorithm. This may be previous vessel performance data, weather conditions, voyage record details, or any other information. This historical data creates a base picture as a starting point. Then, this base can be enhanced by adding current real-time information, such as the current weather, vessel speed etc. The more data volume and data types the better, as the work of an AI/ML algorithm is improved by more data, then feedback on that data, and repeating this in a repetitive learning cycle. In that way the AI/ML model is continually retraining by incorporating new information and as long as the data is relevant, this should result in predictions that are more and more accurate.

The second reason AI and ML are a natural fit for work with port operations is that AI/ML models are very good at identifying patterns and anomalies. Since port operations do typically follow a regular cadence, though every port is different and has their own, the rhythm of a port will usually follow a pattern of weekly/monthly visits by many of the same vessels, same cargo, seasonality, etc. Once routines are mapped by the AI/ML model, differences can be detected to whatever tolerance level that makes sense, and adjusted over time to be useful. For instance a model may aim to find vessels consistently delayed, relate that to weather conditions or operational issues that would otherwise be missed by a human-only review of large volumes of data. The benefit of the AI/ML may be only in the identifying of the issue that needs more human analysis and plan for improvement, but the flagging could be critical on its own. The work may otherwise have been impossible, or not done at all if it required too much time to sift through information by port operations, and the human time is better spent making improvements. It can be thought of in a similar way as the broad concept of ‘automating’ anything in the workplace, to put human effort into the things automation cannot do, and especially to create capacity of time and clarity of situation to make continuous improvements. There are use cases already with the aim to reduce crew and passenger accidents, improve fuel economy, and some ports around the world (for instance Rotterdam) use ML-based systems to predict vessels’ time of arrival.

Using both of these AI/ML natural strengths, large volumes of data and pattern identification, there are areas in the maritime industry testing and progressing with AI/ML application. One of these is cargo management. Shipping companies can now find tools to optimize their cargo routes, reduce transit times and minimize the risk of loss or damage to cargo. This is starting to mean that shipping companies and their customers can choose between routes for efficiency, reduce shipping costs, increase profits, and provide better service to their end user customers.

Another area with some already available AI/ML tools is the prediction of weather conditions. This is commonly available even to the general public in severe weather modelling such as the North American hurricane watch by NOAA. Furthering the value of some of that information, it is now possible for shipping companies to have tools that take in weather variables and make predictions and recommendations for route adjustments to minimize delays and risks such as damage or loss of cargo.

A third area with some currently available tools is in vessel traffic management. AI/ML is starting to be incorporated into the monitoring of vessel traffic, prediction of potential incidents and related allocation of resources. This may be with the aim to reduce congestion, minimize delays or increase the overall safety and efficiency.

These are still early days in the exploration of using AI and ML in the maritime industry and specifically in port operations, however there are solid and specific use cases that are showing promise for value. The next evolutions of these tools will determine if they can be trusted in a programmatic accepted way. Now, let’s dive a bit deeper.

Several ports around the world are already using AI and ML algorithms to improve operational performance. Port of Rotterdam, Port of Singapore, Poer of Antwerp, Port of

Shanghai are using various algorithms to calculate more accurate vessel ETA, optimize vessel operations and predict potential bottlenecks in the supply chain. Historical vessel tracking data, real-time sensor information, and weather data are used in combination to make predictions of future situations that are more accurate.

Here is a closer look at how they are using these technologies:

- Predictive modeling: The Port is using predictive modelling to predict vessel ETA and improve vessel scheduling. The combination of historical visit information, real-time sensor data, historical and current weather data are used to make more accurate predictions of vessel arrival time. This helps the port to better plan its water and shore operations.
- Vessel traffic management: the port is using AI to optimize vessel traffic management and reduce vessel wait times. The port is using AI to predict and schedule vessel movements through the port area, identify potential traffic chock points, and better plan the allocation of resources. This helps to reduce port congestion and improve the safety of the navigation in the port.
- Supply chain optimization: Port is using AI and ML algorithms to optimize supply chain operations and improve the flow of cargo through the port. The port is using AI to create better schedules and identify potential bottlenecks in the container handling and movement.
- Real-time monitoring: Port is using real-time monitoring to keep track of vessel movements and improve navigation safety. The port is using real-time information, such as vessel speed and weather conditions, to make predictions that are more accurate and advise avoidance maneuvers when needed.

ETA Prediction calculations (Just in Time arrivals)

Just in Time (JIT) is a supply chain management philosophy that aims to minimize the amount of inventory held by the supplier and increase efficiency by delivering goods to the customer exactly when they are needed. In the context of maritime transportation, JIT arrivals refer to vessels that arrive at the port exactly when the cargo is ready to be loaded or unloaded. In 2020, the International Maritime Organization (IMO) released a guideline advising ports how to facilitate a JIT arrivals process. A Just in Time arrivals process is recognized as a means of increasing port efficiency and port call optimization. Successful implementation can have a significant environmental impact through reduced greenhouse gas emissions from optimizing the ships' speed and route to arrive just in time for the required operations. The ultimate goal is for the ship to maintain an optimal operating speed to arrive at the Pilot Boarding Ground when the availability of berth, fairway, marine services (like pilots, tugs, linesmen) is assured. Just in Time arrivals processes are also designed to reduce the time at anchorage and therefore reduce congestion in the port area increasing the safety and reducing the emissions.

To achieve JIT arrivals, accurate ETA calculations are critical. The ETA is used to predict the time of arrival of the vessel at the port, which helps the port operations, terminal and the customer to plan the activities accordingly. By accurately predicting the ETA, the port can coordinate the cargo operations efficiently and minimize any delays. This, in turn, helps to achieve JIT arrivals, where the vessel arrives at the port exactly when the cargo is ready to be loaded or unloaded. Artificial Intelligence (AI) and Machine Learning (ML) are increasingly being used to make more accurate ETA predictions.

There are several methods of AI and ML algorithms being used for ETA predictions in the maritime industry. HMM-based prediction models - Hidden Markov Model (HMM) is a statistical model that can be used to predict future states based on historical data. Deep learning models (such as neural networks), Decision trees and random forest models, Support vector machines (SVM) can be trained on large amounts of historical vessel data to make ETA predictions.

These models can take into account a wide range of factors, including vessel performance, weather conditions, and route information, to provide predictions that are more accurate. These algorithms can be customized and optimized for specific use cases, such as predicting ETA for different types of vessels or for different routes.

Let us look at the specific steps involved in estimated time of arrival (ETA) prediction in ports. To make an accurate ETA calculation involves multiple factors, including vessel speed, cargo loading and unloading time, vessel routing and routing time, weather conditions, and port congestion, among others. To calculate the ETA of a vessel, these factors are taken into account and combined using mathematical models and algorithms.

Figure 3

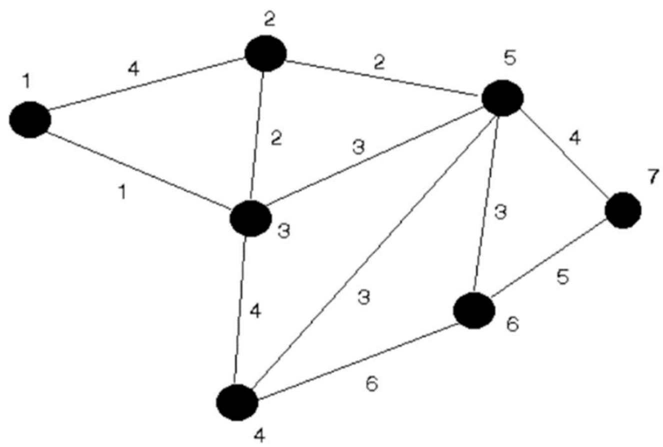
ETA Calculation Inputs and Data Requirements

Vessel Information: The first step is to gather information about the vessel, such as its size, draft, speed, and current location. This information is used in the future steps to estimate the vessel's future position and speed.

Route Planning: The next step is to plan the vessel's route. That means determining the waypoints to be visited and the estimated time of arrival at each waypoint. The route will also be optimized to minimize the total voyage time. The route-planning step in ETA prediction is a crucial piece that involves determining the optimal route for a vessel to travel from its current location to its destination. It starts with the selection of the waypoints that the vessel will pass along its route. This involves identifying the areas that the vessel will need to pass through or avoid during its voyage. A routing algorithm is then used to determine the optimal route based on various factors, such as vessel speed, fuel consumption, and expected weather conditions at various waypoints. The routing algorithm takes into account the vessel's size, draft, and other characteristics. The calculated route is then optimized to minimize the total voyage time. This involves selecting the waypoints, speeds, and fuel consumption to find the optimal route. Note that the optimal route is not always a shortest route. The planned route is assessed for any potential risks, such as shallow waters, narrow or busy channels, or areas prone to piracy. Depending on the external factors like weather and other potential safety risks, the algorithm can decide to route the vessel through different waypoints. Therefore, if any risks are identified, the route may be adjusted to avoid these areas. The final route is then approved and communicated to the vessel's crew, who will use it to guide their voyage.

Figure 4

Route Planning Waypoint Illustration



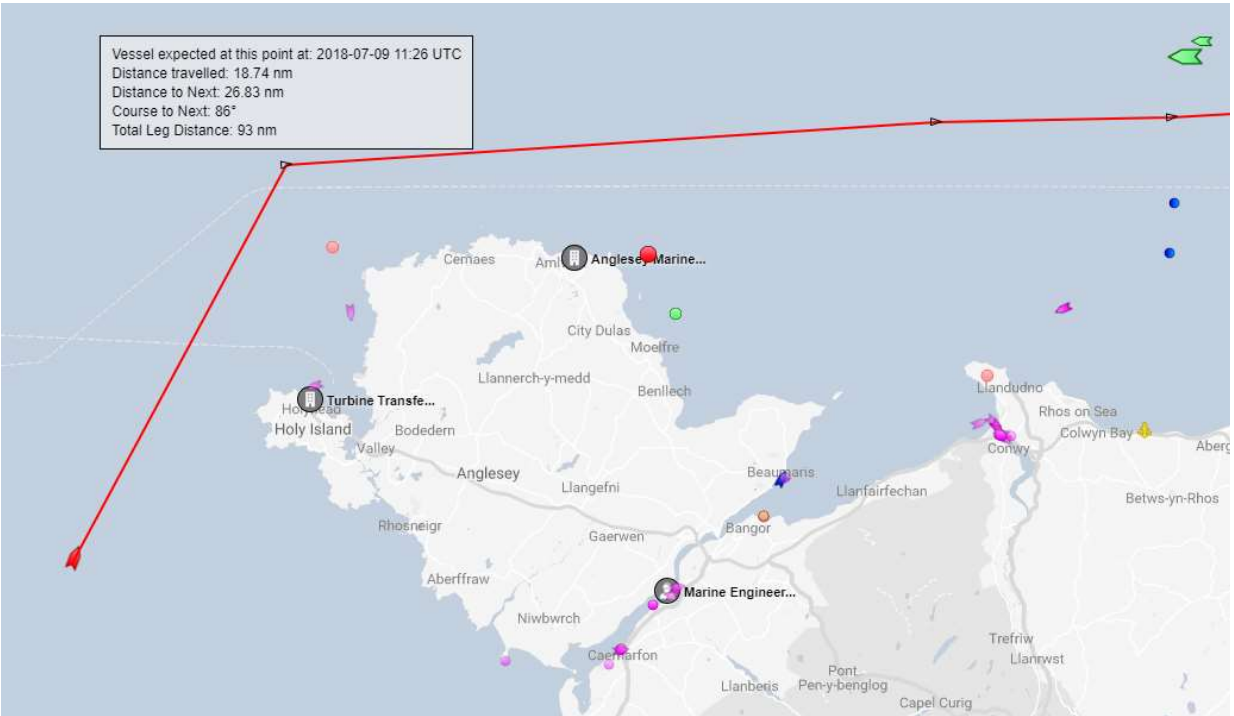
Weather Forecasting: After the route is planned, the weather conditions, such as wind and waves, are taken into consideration because they can significantly influence a vessel's speed. Weather forecasts are used to estimate the impact of weather on the vessel's voyage.

ETA Calculation: Finally, all of the information gathered and estimated in the previous steps is combined to calculate the ETA of the vessel. A simulation model is used to estimate the vessel's position and speed based on the information gathered about the vessel, its route, and the expected weather conditions at each waypoint. The final ETA is calculated by combining the results of the simulation by estimating the total voyage time.

After the ETA along the planned route is calculated the algorithm looks at the port congestion situation and potentially may alter the selected route (and, as a result, the estimated time of arrival) to reduce the wait time at anchorage (JIT arrivals).

Figure 5

ETA and Route Simulation



Port Operations and Port Congestion: Port operations, such as cargo loading and unloading, and availability of marine services also have an impact on a vessel's optimal ETA. The resource availability and the time required for these operations are estimated and included in the ETA calculation. Port congestion, caused by a large number of vessels waiting to enter or exit the port, can also impact a vessel's ETA. Information about port congestion is used to estimate the time duration required for a vessel to enter or exit the port. Port congestion is an important factor to consider in ETA. When there is a large number of vessels waiting to enter or exit the port, the time required to complete the port operations, such as cargo loading and unloading, can increase dramatically. This, in turn, can cause delays in the voyage and result in a longer passage times.

There are several reasons why port congestion occurs. When there is a high demand for port services, such as cargo loading and unloading, the port can become congested as more and more vessels arrive. Some ports have limited capacity, and when this capacity is reached, congestion can occur. This is particularly common in smaller ports with limited infrastructure and resources. Inefficient port management can also contribute to congestion. This can include poor coordination of cargo operations, limited resources, and insufficient staffing. Weather conditions, such as storms or high winds, can also affect port operations and lead to congestion. By considering port congestion in the ETA calculations, vessel operators can make informed decisions about voyage planning and make adjustments to minimize the impact of congestion on their ETA. For example, they may choose to schedule their arrival during less busy times. By accurately predicting the ETA, vessel operators can minimize the impact of port congestion on their voyage and ensure that they arrive at their destination on time.

In this use case of Just in Time arrival, the use of AI and ML for ETA predictions can help reduce costs and improve efficiency for shipping companies. Equally important, by providing more accurate ETA predictions, ports can better plan their operations, optimize vessel utilization, and reduce the need for costly contingency plans. It also begins a circle of increased efficiency, as the vessel arrival prediction is more accurate, so then the port work can be planned better, which feeds back to the vessel prediction being better, in a continuum.

Cargo Manifest pre-processing data cleanup

Efficient and accurate processing of cargo manifest data is critical of port operations. Historically, it was a labour-intensive process with people reading manifest books and typing the information into the system. With the progress of technology, this process can be automated for the most part. The process of automated import of cargo manifest requires several stages. The main goal is to enable an automated system that can read and parse the cargo manifest file, convert/understand textual information – such as commodity type, consignee, shipper, and receiver – into a format understandable and used the system. Subsequently, the system must assess the validity of this information and undertake data processing with minimal human intervention. The ultimate aspiration of this entire process is to seamlessly import the cargo manifest file can be done without manual interference.

Cargo manifests, typically compiled by the sender, often arrive filled with irregularities, errors, or inaccuracies due to a multitude of factors. These include the human element of manual data entry, the assortment of distinct formats originating from diverse sources, and the potential language variations encountered in international shipments. For instance, a recipient might be marked as “CanTire,” “C Tire,” “CanT,” or “CT,” necessitating the system's ability to discern that these textual variations correspond to “Canadian Tire.”

Fully relying on an unsupervised, entirely automatic processing approach proves untenable, as a Machine-Learning algorithm, while always delivering a response, might not always present the correct one. The process of importing cargo manifests mandates absolute accuracy, allowing for no room for error. Addressing this challenge requires for the implementation of human-assisted Machine Learning, specifically utilizing an "active learning" strategy. In this context, the model is designed to recognize instances where matches are uncertain, and where a guaranteed accurate prediction cannot be made. In such instances, human intervention is required, facilitating the model to learn from these corrections, and thus, optimizing its learning process. With time the number of such instances will be reduced, as the model will learn most of the variations found in manifests received by the port. From the experience of actual running of such process by various ports, the prediction of over 90% is achieved within 6 months and after 9-10 month of processing, only a few lines of the manifest would require human intervention.

Following the import of data into the system, the next phase involves a validation of the acquired information. This entails a series of checks aimed at ensuring the integrity and correctness of the data. For instance, the system can cross-verify if a container's weight falls within the maximum permissible limit or discern if a container declared as empty indeed contains cargo. Moreover, using historical data, the system can ascertain the accuracy of declarations such as a container declared as a transhipment having actually arrived and remained at the terminal. To accommodate the unique requirements of individual ports, specific rules are configured to satisfy the specific needs.

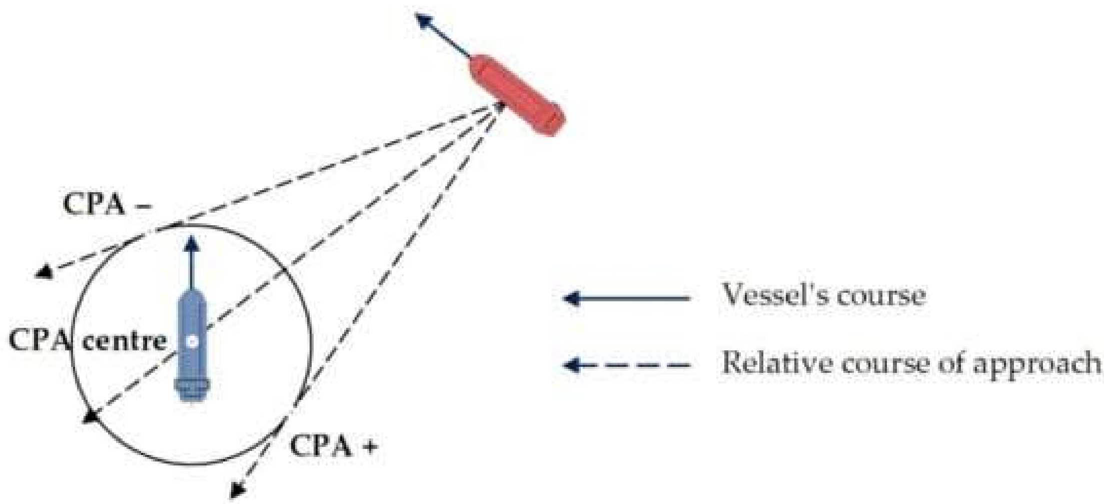
Closest Point of Approach and collision avoidance

Decision-making and responsiveness are the primary activities in avoiding collisions at sea. There is always a limited number of personnel on the bridge during the voyage and the amount of information required per person is constantly increasing. This creates a burden on the decision-making process. The decisions are also influenced by the traffic situation, weather conditions, and the personnel experience. Collision avoidance has an additional problem as the crew members have extensive knowledge of their own vessel but limited information of the vessels in the vicinity. That means that decisions are made in an uncertain environment. The use of AI and Machine Learning in the calculation of Closest Point of Approach (CPA) and collision avoidance provides a significant benefit in such an environment.

In maritime, ships constantly communicate with each other to exchange information about their positions and movements. CPA is a crucial calculation in determining the risk of collision between two ships, and it involves estimating the closest distance at which two ships would come if their present courses were to continue. AI and ML algorithms can be used to analyze the ship's data and predict the risk of collision by estimating the CPA and suggesting avoidance maneuvers. Avoidance maneuvers require planning and analysis of the navigation situation in an appropriate timeframe. Vessel movement prediction is an important part of planning, with extrapolation of the vessel’s trajectory within a time delay. On the radar the navigation situation is presented from the own vessel perspective. The observed vessel is considered a target vessel. The usual algorithm calculates the decision using two steps. In step one, it collects input data of own and any target vessels to assess the collision risk. To predict the target vessel position at the time of collision avoidance maneuver, time delay calculation is used to obtain the new relative position of a target vessel. At this point four parameters (Position, Speed, Course and Time) are calculated for further processing as input variables. The second step is calculation of the optimal decision (course alteration to reduce the risk of the collision).

Figure 6

Closest Point of Approach Deviation Calculation

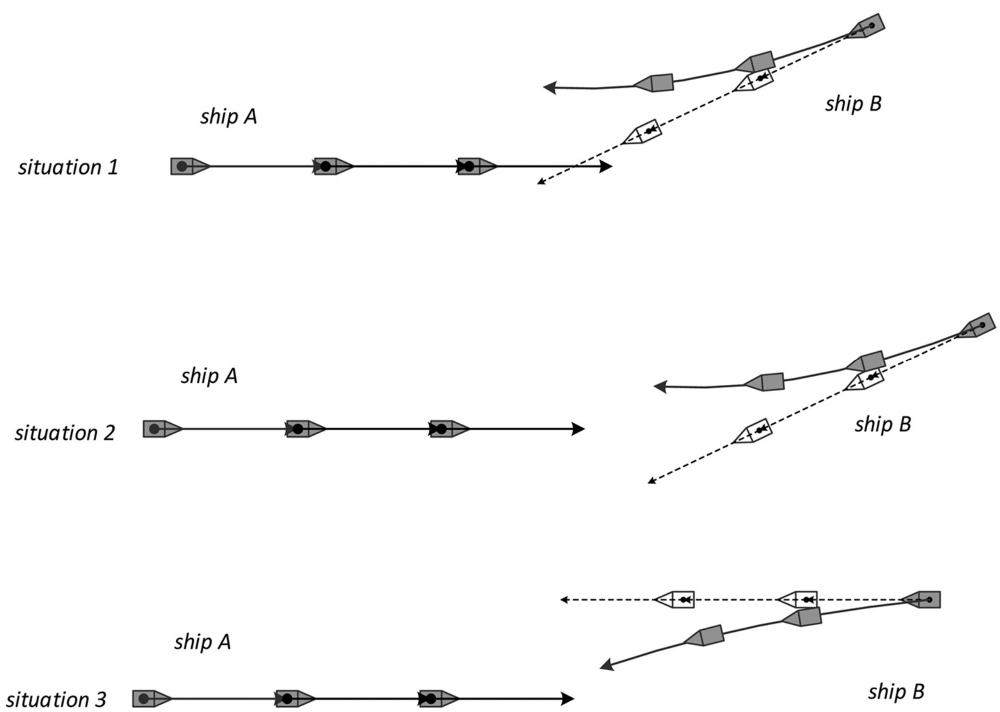


There are various methods to predict the position of the target vessels in relation to own vessel. Various algorithms are used to depending on the distance between vessels. The outcome of the calculations is used to provide real-time warnings to the ship's crew and suggest corrective actions to avoid a collision. Predictive analytics' algorithms analyze large amounts of data about ship movements, weather conditions, and ocean currents to predict the risk and recommend the avoidance maneuver. The uncertainty of the position, speed, and course measurements is also taken into account when calculating CPA. Uncertainty can arise from measurement errors, environmental conditions, or other factors. A collision risk threshold, also known as the collision avoidance criteria, is set to determine the level of risk that is acceptable before taking corrective actions to avoid a collision.

Most long-distance algorithms are based on analysis of Automatic Identification System (AIS) data, which is a system that provides real-time information about the position, speed, and heading of ships. AIS messages contain data that can be used to calculate the CPA between multiple ships and predict the risk of collision. CPA is usually calculated based on the speed and direction of the approaching ship. To improve the CPA calculation, additional parameters to standard AIS information containing the Speed Over Ground (SOG), Course Over Ground (COG) like Change of Speed (COS) and Rate of Turn (ROT) are calculated and used as well.

Figure 7

Collision Avoidance Options



On shorter distances, there is use of computer vision algorithms that analyze the video feeds from ship-borne cameras to detect other ships in real-time. These algorithms can then use this information to calculate the CPA between the ships and predict the risk of collision.

AI and ML algorithms can significantly improve the accuracy and speed of CPA calculations, reducing the risk of collision and increasing overall safety. However, it is essential to note that AI and ML are not a substitute for human judgment and decision-making, and their output should be carefully evaluated and validated before being used to make operational decisions.

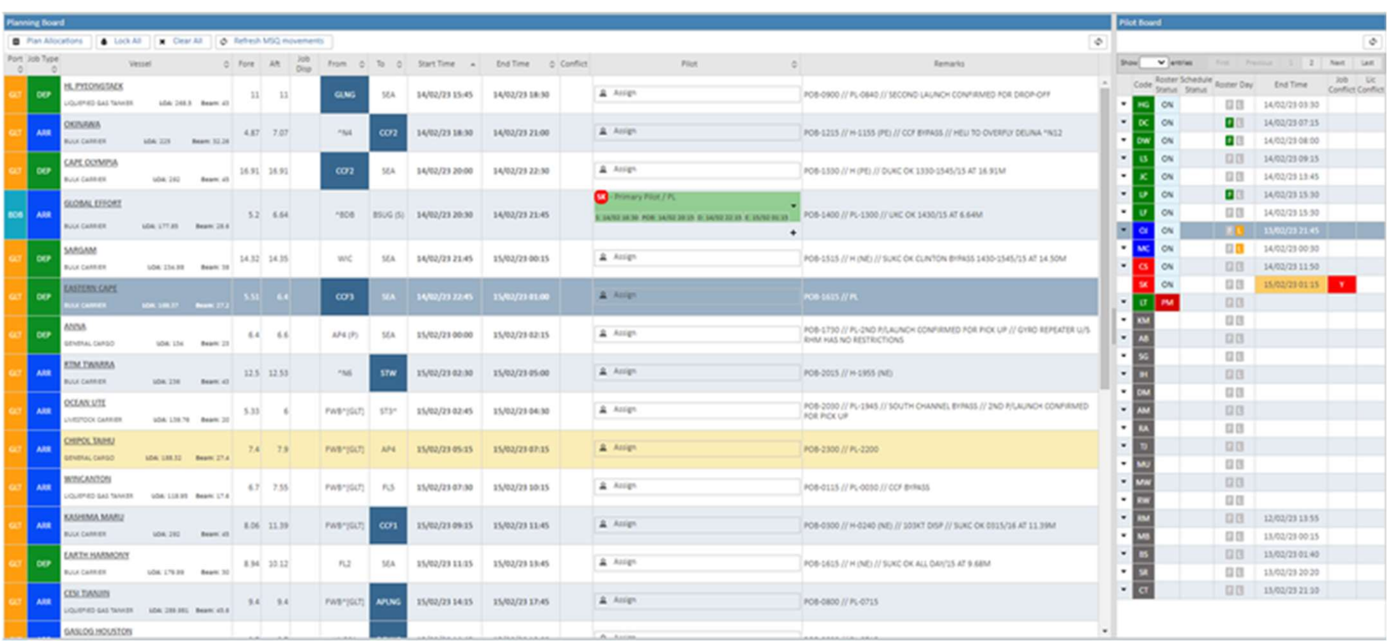
Pilot scheduling and allocation

The scheduling of maritime pilots can be a complex task, as it requires the coordination of numerous variables such as vessel size, arrival times, departure times, port conditions, weather forecasts, pilot availability, and more. In recent years, AI has been increasingly utilized to optimize and streamline this process.

One of the key benefits of using AI for scheduling marine pilots is that it can process and analyze large amounts of data quickly and accurately, providing insights and recommendations that might not be immediately apparent to a human scheduler. This helps reducing errors, increasing efficiency, and improving overall scheduling outcomes.

Figure 8

Pilot Resource Scheduling Example



In recent years, Machine-Learning algorithms have also been used to develop predictive models for scheduling maritime pilots. These models use historical data to learn patterns and relationships in the data and can be used to make predictions about pilot demand and availability. The models can also be used to optimize the allocation of pilot resources based on real-time data, allowing for more efficient and effective scheduling. Proper scheduling of the pilots requires looking at two different sides of the operation. One part of the system is using Machine Learning algorithms that learn from historical data and make predictions based on patterns and trends. The system looks at historical vessel traffic patterns and predicts how many and what type of vessels are likely to require pilot services at a given time, allowing schedulers to proactively allocate pilot resources. The second part is using optimization algorithms, which analyze various scheduling scenarios and determine the most efficient and effective way to allocate pilot resources.

When selecting which pilot can be assigned to which vessel a number of factors are taken into consideration in order to determine the most appropriate pilot for a given job. Some of the key factors include the experience and qualifications of the pilots, the type and size of the vessel, the prevailing weather and sea conditions, and the location and timing of the vessel's arrival and departure.

To select the most appropriate pilot for a given vessel, the algorithm uses a range of techniques and has to go through a number of stages, including:

Pilot availability: The algorithm will first consider which pilots are available to work at the time the vessel is scheduled to arrive or depart. This will depend on factors such as the pilot's work schedule, any regulations or agreements governing pilot availability, and other logistical considerations. The system should be aware of the working rosters and required rest periods between jobs. Determining pilot availability requires tracking the pilots' schedules and ensuring that there are enough pilots on hand to meet the demands of the shipping traffic in the port. The pilot scheduling system will track pilot schedules and availability in real-time, allowing dispatchers to quickly identify available pilots and make assignments as needed. The system may includes tools for managing pilot workloads, monitoring fatigue levels, and ensuring that pilots are given appropriate rest periods. To ensure that there are enough pilots available to meet demand, scheduling systems use forecasting and planning tools. These tools forecast future demand for pilotage services for the next 24 hours, and optimization algorithms that are used to allocate pilot resources more efficiently. The goal of this step is to ensure there are enough available pilots to handle upcoming traffic.

Pilot qualifications: Once the available pilots have been identified, the algorithm needs to consider their qualifications and experience. The level of licensing for maritime pilots is a critical consideration in scheduling process. The license level indicates their level of expertise, training, and experience in navigating vessels generally, and specific types of vessels, through port waters. It is typically regulated by national and international maritime organizations, and the requirements for obtaining and maintaining a maritime pilot license are strictly enforced. In many cases, marine pilot licensing is tiered, with higher levels of licensing indicating more advanced skills and experience. For example, in the United States, there are four levels of marine pilot licensing, with higher levels requiring training that is more extensive, experience, and knowledge. Similarly, in Canada, marine pilot licensing is divided into five levels, each with its own requirements for education, experience, and training. When scheduling marine pilots, it is important to consider the license level required for the specific job that involves the type of vessel and waterway involved. For example, a larger vessel or one that is navigating through a particularly challenging port area may require a pilot with a higher level of licensing (or a special endorsement) and experience. Any algorithm must ensure that the pilots that are assigned to these vessels meet the necessary license requirements to ensure the safe and efficient navigation of the vessel, and to meet insurance and liability criteria.

Safety considerations: The algorithm will also consider safety considerations, such as the prevailing weather and sea conditions, the level of traffic in the port, and any other factors that may affect the safety and efficiency of the pilotage operation.

Pilots fatigue level: Pilots fatigue is an important consideration in the scheduling process for marine pilots. Fatigue is a common problem in the maritime industry, particularly for pilots who may work long hours and irregular schedules, and can have serious consequences for safety and performance. To address the risk of fatigue, most pilotage organizations have implemented regulations or guidelines for pilot scheduling that aim to limit the amount of time pilots can work and ensure that they have adequate time for rest and recovery between shifts. These regulations may vary depending on the country and the type of vessel being piloted, but they aim to limit the amount of time a pilot can be on duty in a given period, how many night assignments a pilot can have in a row and require a minimum amount of rest time between shifts. When scheduling marine pilots, an algorithm must take into account these regulations and guidelines, as well as any additional factors that may affect the risk of fatigue. For example, an algorithm needs to consider the time of day or night that a pilot is scheduled to work, as well as the length and intensity of the pilot's previous shifts for a specific number of days. Sometimes the system may also need to consider individual factors that could affect a pilot's risk of fatigue, such as medical conditions, age, and overall workload. By taking into account these factors, the system can ensure that pilots are assigned shifts that are appropriate for their level of experience and expertise, and those they have adequate time for rest and recovery between shifts, which can help reduce the risk of fatigue-related incidents.

Historical data: The algorithm can also make use of historical data to inform its decisions. For example, if certain pilots have a proven track record of performing well in certain conditions or with certain types of vessels, the algorithm may give them priority when making assignments.

Optimization techniques: finally, the algorithm may use various optimization techniques, such as linear programming or simulated annealing (an optimization technique), to identify the optimal assignment of pilots to vessels, taking into account factors such as pilot availability, vessel traffic patterns, and other logistical constraints. There are various systematic planning algorithms that can be used for scheduling marine pilots. One commonly used algorithm is known as the "Dial-a-Pilot" algorithm:

- Collect data: Gather information about vessel traffic patterns, port conditions, pilot availability, and other relevant factors.
- Define the problem: Determine the specific goals and constraints for the scheduling problem, such as minimizing travel time or maximizing pilot utilization.
- Build the model: Create a mathematical model that represents the scheduling problem, taking into account the various factors and constraints involved. Mathematical models used to optimize the allocation of marine pilots are taking into account factors such as vessel traffic patterns, port conditions, pilot availability, and other variables. These models typically involve a set of mathematical equations that represent the problem at hand, and are used to generate a solution that meets specific goals and constraints. One advanced mathematical model is the simulation model, which involves the use of computer-based simulations to evaluate different pilot scheduling scenarios. In this model, a simulation of the port environment and the behavior of the vessels and pilots is used to generate a set of data that can be used to optimize the allocation of pilot resources. The simulation model can be used to test various scheduling scenarios and evaluate the impact of different factors, such as changes in vessel traffic patterns or changes in pilot availability.

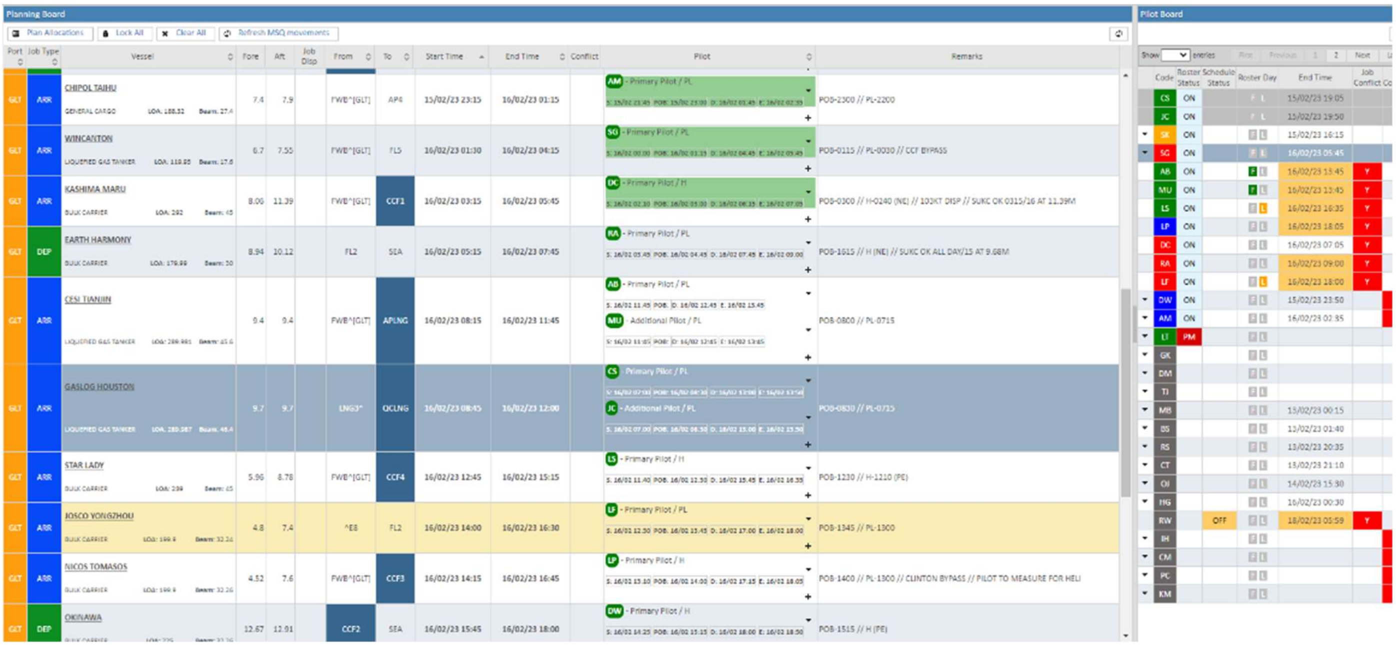
- Solve the model: Use optimization techniques to find the best possible solution to the scheduling problem.
- Evaluate the solution: Evaluate the solution generated by the model to determine whether it meets the goals and constraints defined in step 2.
- Refine the model: If the solution is not optimal, refine the model and re-solve the problem until an acceptable solution is found.

Other algorithms may involve different steps, but the general process typically involves collecting data, defining the problem, building a model, solving the model, evaluating the solution, and refining the model as needed.

Overall, the use of AI in scheduling marine pilots helps improve efficiency, reducing scheduling errors, and increase safety. However, it is important to note that AI is not a panacea, and human oversight and expertise is still essential to ensure that scheduling decisions align with broader port goals and priorities. After the algorithm proposes a scheduling plan, it is presented to the scheduler for the final approval. A proper system allows a scheduler to “lock” individual pilots to specific ships to satisfy factors that are not considered by the optimization algorithms. The next round of optimization will include those decisions and will reschedule other pilots around those constraints.

Figure 9

Pilot Resource Scheduling Factors



It is important to note that scheduling marine pilots can be a complex task, and that the specific steps and algorithms used will depend on the specific needs and constraints of the organization involved. Additionally, the use of AI and ML can also help to automate and streamline the scheduling process, allowing for more efficient and effective allocation of pilot resources.

Conclusion: Do I need AI or ML tools in my port operations?

The question of whether AI or ML tools are a necessity in modern port operations is far from a straightforward one. The answer is not simple but let us break it down.

When thinking about using AI and ML tools for port operations, there are important things to consider. One should start by looking at how things work in the port right now. Check how well different work processes are running and see how they connect to each other. Each process should not be considered in isolation but looked at like a system of systems. Sometimes, trying to automate just one part might cause more problems than it solves. However, if there is an isolated process that relies on a lot of data, and that process can be automated, then you can think about the next step: using AI and ML to make things work better.

AI can handle large volumes of data, make calculations and find patterns that humans might miss. As an example, in the case of the cargo manifest import discussed above, AI can quickly sort through all the details, making sure everything is correct. This means less time spent on checking and fixing mistakes, and a shorter time between receiving cargo manifests and sending the invoice.

Use of AI and ML also comes with challenges. Sometimes, the results might not be right because of missing information or biases in the data. There is also the important question of ethics – we need to make sure AI does not accidentally make unfair decisions.

Before diving in, think about the problem you are trying to solve, and approach it in a business case manner, to assess the best solution. AI/ML could be one potential method to solve a problem, but are not likely the only method, and there is a cost and other considerations. Will they save more time and effort than they cost to set up? Also, consider that connecting AI or ML with your existing systems might need extra work, which can add to the expenses. If your port deals with lots of data, AI might be a great help. On the other hand, if you need human thinking for making complex decisions, human-assisted Machine Learning could be the way to go. The goal should not be to simply have some AI in your port, but instead to solve a problem, and it happens to be by using AI/ML.

Making a decision about AI and ML is not easy. It depends on the specific situation, weighing costs and benefits. Ports are very complex operations, relying on data and operational efficiency. By carefully thinking about how automation, data-smart choices, fairness, and integration between human and machine, there is a course that leads to an efficient solution, but it would be unique in every situation.

Reflections on Learning

While researching the subject of applicability and usage of Artificial Intelligence and Machine Learning for the maritime industry, I have been impressed by the transformative potential these technologies hold and the advancements in the applicability in the last few years. It is still a very fresh subject that I expect will continue to grow significantly in the coming years.

Learning about the different ways and different areas where AI and ML can help maritime operations has shed light on how data-driven decision-making can revolutionize efficiency, safety, and sustainability in this sector. Witnessing the integration of predictive analytics to optimize vessel routing and just-in-time arrivals, AI-powered scheduling, and smart resource allocation has proven the great possibilities for improving operational outcomes.

Furthermore, understanding the significance of real-time data collection, analysis, and response in maritime and port contexts has been eye opening. The ability of AI to process massive datasets from weather patterns to cargo manifest processing, and subsequently provide actionable insights, shows new opportunities for proactive and adaptive management.

However, this research also presented some critical considerations. The necessity for high quality, accurate data to support AI algorithms and calculations solidifies that technological advancements rely fundamentally on data collection, quality and integrity. Moreover, addressing the ethical considerations of AI-driven decision-making, transparency, and potential job displacement underscores the importance of responsible implementation.

While looking at the potential of AI and ML for the maritime industry, a critical topic that emerged is the significance of human-assisted learning. While these technologies show great capabilities, the integration of human expertise and oversight serves as a safeguard to ensure consistently accurate results. Human-assisted learning acts as a corrective mechanism that complements the strengths of AI and ML. Despite their prowess in processing vast amounts of data and recognizing patterns, these technologies can still encounter challenges in specific scenarios or unforeseen circumstances where 100% accuracy is required for safe operation.

Human oversight provides the critical judgment necessary to validate or retrain AI-generated outcomes. Integrating human expertise into the AI-ML loop improves the outcomes and guarantees the results when accuracy and precision are necessary.

In conclusion, my reflection on learning about AI and ML in maritime and port management evokes a sense of excitement tempered by responsibility. The capacity to optimize operations, enhance safety, and reduce environmental impact is enticing. Yet, the need for a holistic approach, encompassing technological prowess, ethical considerations, and stakeholder collaboration, is essential to harness AI and ML's true potential in shaping the future of maritime and port management.

References

Bricko,Tanja, Andrej Andronja, Jure Srse and Renata Boc. (2021). Vessel Multi-Parametric Collision Avoidance Decision Model: Fuzzy Approach.

Brouer, B. D., C. V.Karsten, and D.Pisinger. (2016) Big Data Optimization in Maritime Logistics.

Castanon, Jorge. Towards Data Science (2023) https://towardsdatascience.com/10-machine-learning-methods-that-every-data-scientist-should-know-3cc96e0eeee9

Chin, H.C. and Debnath, A.K. (2009). Modelling perceived collision risk in port water navigation. Safety Science.

Choi, H., Kim, Y., Lee, Y. and Kim, E.T. (2013a). A reactive collision avoidance algorithm for multiple midair unmanned aerial vehicles.

Fabian Pedregosa and others. Scikit-learn: Machine Learning in Python.

Google Team. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems

Hilgert, H. and Baldauf, M. (1997). A common risk model for the assessment of encounter situations on board ships. Ocean Dynamics.

International Conference on Engineering, Technology and Innovation (ICE/ITMC) (2017).

Journal of Navigation publications.

Koerner, Katharina. (2022) https://iapp.org/news/a/privacy-and-responsible-ai/

Marine Digital.

Mirović, M., M. Miličević, and O. Ines. (2018) Big Data in the Maritime Industry.

Munim, Ziaul Haque, Dushenko, Mariia, Jaramillo Jimenez, Veronica, Hassan Shakil, Mohammad, Imset, Marius.(2020). “Big data and artificial intelligence in the maritime industry: a bibliometric review and future research directions” https://www.tandfonline.com/doi/full/10.1080/03088839.2020.1788731

Ocean Engineering publications.

Proceedings of the 12Th ACM International Conference on Distributed and Event-Based Systems.

Projects in which Tidalis Americas LTD was involved as a vendor of software solutions.

Sanchez-Gonzalez, P.-L., D. Díaz-Gutiérrez, T. J. Leo, and L. R. Núñez-Rivas.(2019). “Toward Digitalization of Maritime Transport?”