Vessel ETA Prediction

Author:
Ruby Pel *(5853419)*

Supervisors:
Prof. dr. A.P.J.M. (Arno) Siebes
Wilte Falkena (external: Pon)
Second corrector:
dr. Ing. habil G. (Georg) Krempl

Master Thesis - 14 ECTS
Applied Data Science
Graduate School of Natural Sciences
Utrecht University
July 2022

Contents

Chapter 1

Introduction

A major challenge in the maritime industry is knowing where and when vessels will arrive. The captain of each ship probably knows, however, this data is not shared neither accurately or consistently. The principal of this research is Pon, the number one car importer in the Netherlands, with a strong position in the United States. They are also in the top 5 bicycle manufacturers and are well established in the world of marine solutions, excavation, energy supply, flow control (valves and circuit breakers) and industry services. Pon has multiple companies across the world which independently order various goods from all over the world. These goods are shipped on many vessels using even more containers. Knowing where and when ships will arrive is therefore valuable information for companies like Pon.

Ships enter and leave ports multiple times a day. Their number is constantly increasing as the number of products being transported increases, but a port's capacity is limited. The amount of available space is a major constraint due to two factors: the number of available docks for loading and unloading [1] and the number of vessels that can be in a port's canal at the same time.

When a vessel enter or leaves a port that will be signalised by an enter or leave action respectively. Upon arrival into a port a ship can either stay for a period of time to load or unload cargo, or pass through. The next port it will visit is not directly known. The only way to tell what the next port will be is to check for the next chronological date and time of that same ship. Thus, comes into discussion: what is the final destination of a ship and how long will it take to reach it?

An inaccurate prediction of the estimated time of arrival (ETA) will result in delays in time that will lead to a multitude of negative consequences. From delays in loading or unloading the cargo in the ships to more manpower being deployed in order to compensate for the delay. On the other hand, being ahead of schedule can also result in the ship waiting for the dock to be available or prepared that can result in block the waterway for other ships. Ultimately, an accurate ETA prediction will result in minimum delay and inconveniences in the port and more ships will be able to conduct their business without much trouble.

This research is part of a larger project, together with two other researches (that of A. Khan and M. Lupulescu), which has the objective to build a DSS which predicts the travel time and destination of any vessel on the globe. This research in particular is regarding the prediction of the ETA of a vessel, specifically by using a neural network and the data that Pon has acquired. The research question therefore is:

> "Can a neural network predict the travel duration from the current vessel location to its

destination?"

1.1 Literature overview

There is a plethora of studies in the field of ETA prediction, specifically on using machine learning models to do so. Parolas [1] and Flapper [2] did the most similar researches. Parolas focused his research on ships that are in a port or on the verge of entering. He used GPS data combined with weather predictions by using data from a vessel traffic service system. Furthemore, the time frame of the data was much smaller so it had to be more accurate. Flapper's research was extremely similar, but on a bigger time scale and three different datasets which made the predictions less accurate.

The most popular machine learning models used for ETA predictions were: Support Vector Machines (SVM), Gradient Boosting, Kalman Filtering and Long Short Term Memory (LSTM) [2, 3, 4, 5, 6]. Kalman Filtering was mostly used in combination with other models. The combined machine learning models seemed to outperform the single method ones [3, 5, 6]. The combined models were split into learning from a current situation and past data. By doing so, the model was capable to apply past situation to present ones and adjust accordingly. This acts as a fail-safe for unforeseen situations has as extreme weather conditions, accidents or blockages occurring along the sailing route. Another popular approach was the use of Neural Networks, which displayed much better performance with the trade off of needing more data fed into them as opposed to some of the others [3, 4]. Hardij showed that a feed forward neural network outperformed an LSTM, which means that taking previous data into account does not necessarily make a model on vessel ETA prediction perform better [7] .

While the above researches did have promising results, they do slightly differ from the objective of the current project. Either by the means of transportation (e.g. by bus or walking, instead of ships) or in that the ETA was predicted for one specific route. In this research we want to know the ETA for a vessel to a specific destination, regardless of where it started. Another difference between this study and those that predict a vessel's ETA [1] [7], is that the reported ETA by a vessel's crew was available in their data. This was not the case for the dataset from Pon. This data could be used to determine whether the model was able improve this ETA, and it turned out to be an important model feature in Parolas' research [1].

Chapter 2

Data

The data provided in this paper consisted of three data sets which originate from an Automatic Identification System (AIS) that transmits a ship's position so that other ships are aware of it position. This information is picked up by ground stations and satellites in order to make vessels trackable even in the most remote areas of the ocean. The provided data sets are the following: Vessels information, Ports information and Port visits.

The Vessels information data set contains 4962 unique vessels located globally. The attributes present can be found in Table 2.1. The Ports information data set contains 3575 unique ports. The attributes present can be found in Table 2.2. The Port visits data set contains 9269452 rows, with information regarding 83729 unique ships and 2103 ports. This data set is a pre-processed version of 9 billion of the raw GPS vessel coordinates and consists of more than 2.5 years of historical data, ranging from 2019-04-01 until 2022-01-04. It states the enter and exit activity (action) per port, with the associated port information. Each row represents a vessel-port activity, where the vessel either enters or exits a port. The attributes present can be found in Table 2.3.

2.1 General data exploration and preparation

In order to get a better grasp on the data, the Port visits and Vessels information data was first merged based on the unique mmsi attribute, and then merged with the Ports information data. This dataset contained 784152 observations with 40 attributes.

Then, the number of outliers for each attribute was observed. First, an attribute was considered

Attribute	**Data type**	**Description**
imo	integer	Unique ship code
ship_type	string	Type of ship
mmsi	integer	Unique ship code
length	float	Length of the ship
speed	string	Speed category of the ship (low, medium, high)
depth	float	Depth of the ship

Table 2.1: Vessels information.

Attribute	Data Type	Description
port_index	float	Index of port
portname	string	Name of the port
code	string	Code of the port
prttype	string	Type of the port (sea/river)
prtsize	string	Relative size of the port
status	string	If the port is open/closed (or unknown)
maxdepth	float	Max depth for vessels
maxlength	float	Max length for vessels
annualcapa	string	-
humuse	string	-
locprecisi	string	-
latitude	float	Latitude of the port
longitude	float	Longitude of the port
iso3	string	Short description of country name
iso3_op	string	Short description of country name
country	string	Full country name
lastcheckd	string	-
remarks	string	-
url_lca	string	-
source	string	-
createdate	string	Date of creation
updatedate	string	Last updated time
geonameid	float	-
gdb_geomat	float	-
country_name	string	Full country name
code_2	string	Alternative country code
code_3	string	Alternative country code
country_code	float	Country code
iso_3166_2	string	-
continent	string	Full continent name
sub_region	string	Full region name
region_code	float	Region code
sub_region_code	float	Sub-region code

Table 2.2: Ports information

Attribute	Data type	Description
mmsi	int	Unique identification number of the vessel
port_name	string	Name of the port
port_lat	float	Latitude of the port
port_long	float	Longitude of the port
datetime	string	Date and time of the action
action	string	If the vessel enters or exits the port
distance_to_port	float	Distance to the absolute port center coordinates
stay_duration	int	The number of hours that the vessel was inside the port
visit_uuid	string	Unique identifier to map the enter action to the exit action
imo	int	Unique identifier for ships
latest_know_port	int	1 if this is the latest known location in the dataset, 0 otherwise
port_index	int	Unique id of the port

Table 2.3: Port visits.

an outlier if it fell outside of the 0.01 and 0.99 range. The 'stay_duration' attribute has 4864 outliers. Reducing the thresholds to a maximum of 0.1 and 0.9 would mostly show outliers for unique identification attributes, such as 'mmsi', 'imo' or 'geoname id', which were not relevant. Given the changes the data set would need to undergo, the outliers were all kept in the data set.

The correlation between attributes was also checked, and it was found that the following attributes are highly correlated: latitude and port_lat with a coefficient of 0.999235; longitude and port_long with a coefficient of 0.997628; depth and length with a coefficient of 0.971337. The first two cases are essentially the same data contained in two data sets before merging, whereas the depth and length are shown in a similar manner. Thus, the first two pairs of correlated attributes were merged into one attribute.

The combined data set was analysed on missing values for each attribute and those with more than 90% of the data missing were removed. The attributes with their respective ratio of missing values can be seen in Appendix A.

Lastly, it was found that the 'ship_type' attribute contains five types of ships. This can be seen in Table 2.4. Pon is a family business involved in mobility products, services, and solutions globally. Since the data set of this research is concerned only with the delivery of cars and car parts with the use of vessels, the tanker ships were removed since they are not related to car or car parts delivery.

Ship type	Number of ships
Chemical/Oil Tanker	192866
Container Ship	281622
Crude Oil Tanker	43054
General Cargo Ship	264402
Tanker	2208

Table 2.4: Ship types.

2.2 ETA feature generation

Before data preparation, it was important to consider what the relevant attributes were for ETA prediction and what additional features had to be extracted from the data.

From the Port visits data set and Vessels data set combined (i.e. the AIS data), the following attributes were used for generating the final data set:

- Mmsi
- Date and time
- Port latitude and longitude (which is also considered a vessel's location)
- Port index
- Distance to the port
- Stay duration
- Ship's characteristics: ship type, length and depth

The following features have been generated from the above list and added:

- Trips number
- Speed since last exit (in km/h)
- Distance from last exit (in km)
- Minutes since last exit
- Actual Time of Arrival (in minutes)
- Distance to destination (in km)

The distance was calculated using Haversine formula, to account for the earth being a sphere. The formula can be found in Appendix B.

2.3 Trip definition

A downside to AIS data, is that there is no clear start and end point for a vessel. This means that this had to be defined. To reduce the problem at hand, a single port was chosen as the destination: Europort. This port lies in the harbour of Rotterdam; an interesting area for Pon. Also, it is closest to where most general cargo and container ships arrive, according to Port of Rotterdam [8], and had the most entries in our dataset out of all the ports in the area, namely 5724 visits. The starting point, however, could be any port that occurs prior to Europort for a specific vessel.

To further specify a trip, it must meet the following criteria:

(a) The stay duration (i.e. how long a ship stays in a port) for each port visit is no longer than 72 hours (three days);

(b) The speed between to consecutive port visits is between 5 and 60 km/h or equal to 0 km/h (which means the ship stayed in the same port);

(c) After filtering on the above criteria, the trip should at least contain two ports before arriving in Europort.

A reason for capping the stay duration with criterion (a) is that it is likely that the ports visited before a long stay duration are not relevant for the trip after said stay. The Review of Maritime Transport reports that container ships spent a median of 0.71 days in a port during 2020 (considering all ports in the world)[9]. In the Netherlands this number was 0.8 days, while in Belgium and Germany this number was 1.04 and 0.98 days respectively. These figures gave an idea of what a 'normal' stay duration is around Europort. Eventually, a maximum of 72 hours was chosen (and a median of 0.63 days), because this would still include 90% of the entries. This meant that we did not lose much data with this criterion, but did remove some extreme large values that could influence the training.

Not considering very slow and very fast trips between two ports with criterion (b) excludes unrealistic values. According to Rodrigue, containerships travel at speeds between 22.2 to 56.3 km/h [10] However, this is when they are at sea. By taking a lower minimum value it could also be taken into account, that a ship might slow down when approaching a port or when not at open sea. Keeping the entries where the speed is 0 (when a vessel has stayed in the same port as before) is important; otherwise important information would have been lost about the time it takes a vessel to travel between two ports.

Criterion (c) was applied to the data, because the very first entry and last entry of a trip could not be used for the prediction. The first entry does not have any information on the speed, because the port prior to it is unknown. The last entry, so that of the destination, is irrelevant because the vessel has already arrived at the destination. This meant that we needed at least one more port visit besides the start and end for that trip to be relevant, which is how three port visits in total was determined.

2.4 Data preparation

The merged and slightly cleaned data set from section 2.1 was sorted by vessel 'mmsi' and then by the 'datetime' attribute for convenience. For each vessel the enter and exit actions are then displayed in chronological order.

Based on the desired features and the trip definition, the filtering of the merged data can be summarised by the following steps:

1. Group the data by 'mmsi'
2. Calculate the time in minutes and the distance in km between consecutive rows, then add the speed in km/h as an attribute;
3. Add an attribute that keeps track of when a trip ends, according to the criteria in section 2.3;
4. Label the trips with a unique number;
5. Calculate the distance and the ATA to Europort for each row and add them as separate attributes;
6. Remove any rows that have an ATA of more than 7 days;
7. Remove all rows where the 'action' is 'enter'.

This resulted in a dataframe with 2046 trips, consisting of 4861 port visits in total.

Chapter 3

Methods

This section describes the methods used to answer the research questions and reach the goal of this project.

3.1 Translation of the research question to a data science question

To make an accurate prediction on a vessel's estimated time of arrival (ETA) at a specific port (in this case this is Europort), the features extracted from the dataset as in chapter 2 can be used. Along with the AIS data, the target variable can be predicted using a neural network.

To answer the research question, there will be looked at what network parameters give the best result. Furthermore, since the true ETA is unknown, it cannot be compared with the results of the neural network. Therefore, it was determined that the neural network would be sufficient in estimating the ETA if it has a mean absolute error (MAE) of 12 hours or less when predicting a maximum of seven days ahead.

3.2 Motivated selection of methods for analysis

Many different machine learning methods could be used to tackle the problem at hand. One of the simpler methods, linear regression, however was not applicable to this case because of the non-linear relationship between the input features and the output. An SVM and neural network can actually fit this non-linear relationship, but an SVM needs a lot of data. Therefore, a feedforward neural network was chosen. Although these also need quite some data, in this case it was not necessary to create a very complex neural network (which would need a lot more data). Also, a widely used rule-of-thumb is that the sample size must at least be a factor 10 times the network's total number of weights [11]. In this case, the total number of weights was 120, which meant that at least 120 x 10 = 1200 data points were necessary for training. As will later be explained in detail, the number of training points was around 2700.

3.3 Model input

The input vector for the model contained the following features: weekday, month, time, port latitude, port longitude, distance to port, stay duration, distance to previous port, distance, minutes since previous port visit, ship type, speed in km/h, length and depth of a ship. The datetime object was separated into the number of the day in the week (0 to 6), the number of the month (0 to 11), and minutes after midnight. All the categorical features were also given a number, to replace the strings.

Then, the data was split into train and test data. Since the predictions had to be made based on data that lies in the future, the data was not split randomly. The data was ordered based on the arrival datetime in Europort, after which the order of trips was extracted. The data was then ordered according to this trips order and split into 80% data for training, and 20% for testing. The training data was later split in the same manner and with the ratio: 70% for training and 30% for validating. It was also ensured that all the data from a particular trip is either in training or testing data. Because of the difference in value ranges across all of the features in the input data, the data was then normalized on the scale [0,2]. This ensured that the values were all on the same scale, without distorting variations in the value ranges.

3.4 Model development

This section will describe the architecture of the neural network for the prediction of the travel duration from the current vessel location to Europort.

First of all, the number of input nodes of the network is determined by the number of features, which were 14 in this case. The number of output nodes is determined by the target, which is 1; the travel duration to Europort. The number of hidden nodes is not determined by the data or the task at hand, but can be chosen. For most problems, one hidden layer is sufficient [12]. In the experiments on the training set, adding more hidden layers did not improve the performance. Probably because the size of the training data was not large enough for the model to perform well when having to learn a lot of weights, relatively. A rectified linear activation function (ReLU) was used for this hidden layer, because it learns fast and offers better generalization and performance compared to sigmoid and tanh [13]. After the hidden layer, a dropout layer was added to reduce potential overfitting of the model. For the output layer, a linear activation function was used, because this can output continuous negative and positive values.

Various combinations of features for the input data were tested to determine what configuration of the data leads to the best results. In addition, the number of epochs, batch size, dropout, and portion of train data for validation were considered with the hyperparameter tuning. This resulted in the following configuration:

- Epochs: 20
- Batch size: 5
- Dropout: 0.2
- 30% of training data for validation

3.5 Model evaluation metrics

The model was evaluated using mean absolute error (MAE), together with the root mean square error (RMSE). The MAE gives insight into how far off the model's preditions are on average, while the RMSE gives insight into the variance of the predictions. The goal of the model is to reduce the MAE and RMSE, so it will fit the true data better.

Chapter 4

Results

This chapter presents the analysis results of the selected neural network model's performance on the test data.

The overall MAE of the model was 2244 minutes when the maximum actual time of arrival was less than 7 days. The RMSE was 0.3021. For figure 4.1 and 4.1 the MAE (in minutes) was estimated on twelve hour intervals of the hours before a ship would arrive at Europort. It shows that the model performed best at predicting the ETA of a vessel when it had around 60 hours left. Between 48 and 72 hours, the model was able to have an MAE of less than twelve hours. When the actual time of arrival was lower than 48 hours, the model's MAE was higher. This is also the case for when the time of arrival was higher than 72 hours; the MAE gradually increases.

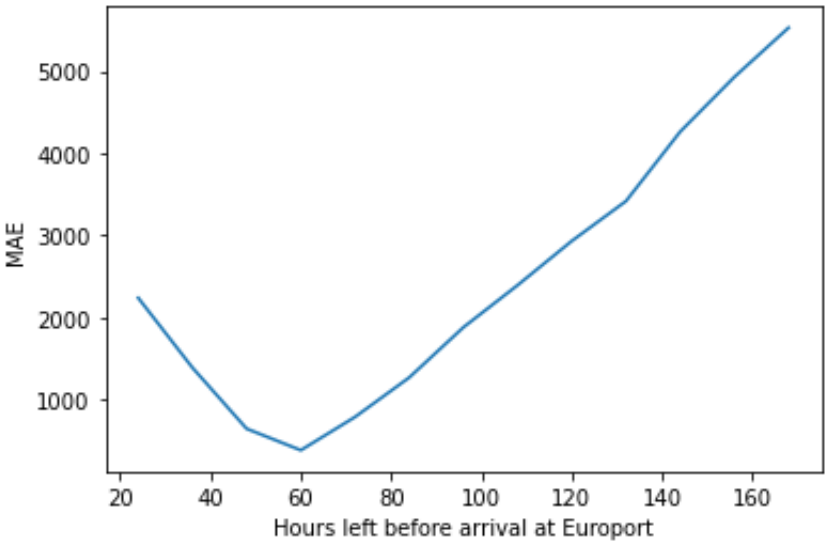

Figure 4.1: (MAE in minutes)

	MAE	RMSE
12-24 hours	2236	3237
24-36 hours	1386	1478
36-48 hours	642	799
48-60 hours	380	500
60-72 hours	785	903
72-84 hours	1256	1397
84-96 hours	1878	1996
96-108 hours	2391	2511
108-120 hours	2930	3036
120-132 hours	3416	3475
132-144 hours	4261	4316
144-156 hours	4929	4991
156-168 hours	5528	5625

Table 4.1: The MAE and RMSE in minutes for twelve hour time intervals of time left until the actual time of arrival.

A neural network is often seen as a 'black box', because what exactly happens in the hidden layers remains unknown. However, by looking at the weights the network gave to the edges between the input and hidden layer, one can get a sense on what features were important (and which were not). Features with a relatively higher set of weights than the others are the stay duration, the distance to Europort, and the depth.

	neuron 1	**neuron 2**	**neuron 3**	**neuron 4**	**neuron 5**	**neuron 6**	**neuron 7**	**neuron 8**
bias	0.01	-0.04	-0.04	-0.06	-0.11	0.05	-0.01	-0.10
week-day	-0.40	-0.34	0.02	0.16	-0.23	-0.02	0.36	0.21
month	-0.1	0.25	-0.5	0.11	0.28	0.04	0.04	0.10
time	0.04	0.45	0.44	0.12	-0.15	-0.26	-0.13	0.32
port_lat	0.4	0.16	-0.30	-0.12	-0.18	-0.44	0.31	-0.43
port_long	-0.35	-0.50	-0.46	-0.15	0.28	0.52	-0.29	-0.52
distance_to_port	0.07	0.13	-0.48	-0.48	-0.32	-0.20	0.17	-0.00
stay_duration	-0.01	0.14	-0.15	0.02	-0.20	0.56	0.26	0.15
distance_prev	-0.01	-0.166	-0.00	0.27	0.17	0.12	-0.37	0.22
distance_ep	-0.50	0.40	-0.19	-0.13	0.22	0.33	0.62	0.20
minutes_prev	0.22	0.46	-0.13	-0.18	-0.19	0.13	-0.06	-0.62
ship_type	-0.30	0.09	0.23	0.32	-0.38	0.31	0.03	-0.06
speed_kmh	-0.5	0.37	0.20	-0.57	-0.21	-0.28	-0.32	0.41
length	-0.26	-0.07	0.02	-0.12	-0.55	0.15	-0.33	-0.09
depth	0.18	0.38	-0.09	0.09	0.24	0.30	0.48	0.31

Table 4.2: The weights assigned to each feature by the model, between the input layer and the hidden layer.

Chapter 5

Conclusion and discussion

5.1 Answering the research question

The research question was: "Can a neural network predict the travel duration from the current vessel location to its destination?". To answer this question, the MAE was determined, and had to be 12 hours (720 minutes) or less when predicting a maximum of seven days ahead. Overall, the model's MAE was 2244 minutes (37.4 hours), which means that the neural network was not successful. When the actual time of arrival was between 48 and 60 hours, the model was however succesful. Between 60 and 72 hours the model scored an MAE of 13 hours, which was close to the desired result.

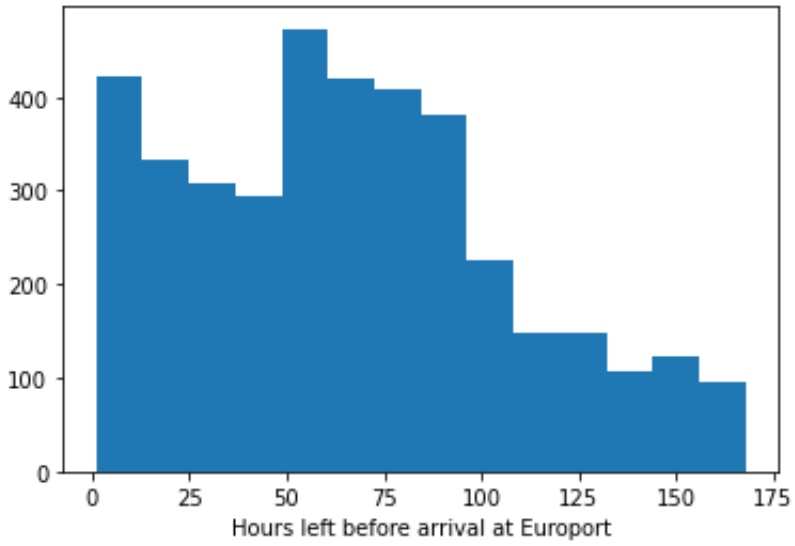

Figure 5.1: Total of training data per 12 hour intervals.

Surprisingly, the port longitude and latitude features were relatively weighted quite low by the

model. One would expect that the location of a vessel is important in predicting the ETA at its destination. The next section will cover more on this feature. The stay duration was weighted relatively high. This can be expected, because together with the time travelled between two ports, it covers the total time between entering the previous port and entering the current port in which a vessel is located. Furthermore, the distance to Europort was weighted relatively high. This seems very reasonable; that the time to the destination depends on distance that a vessel removed from this destination. Lastly, the depth of a vessel was weighted relatively heavily. This could be due to the fact that a ship that lies deeper in the water, has more counterforce from the water, which might make it slower.

5.2 Limitations and recommendations

There are multiple possible causes for the lack of performance of the trained model. Since previous research has shown that a good performing neural network model can be constructed, there have to be reasons why it is not the case for this research.

Firstly, many choices had to be made to define the trips, which have influenced the training and testing set available. Figure 5.1 shows the total of training data per 12 hour intervals. The most data was available at the intervals where the model actually performed best. This could indicate that when there is slightly more training data, the model performs better. Pon could, for example, group the trips to all of the Rotterdam ports to then have more data to train the model on.

Secondly, the normalization of the variables might have changed their relationship or made it hard to find that relationship for the model. This concerns the port latitude and longitude, and the time features. In future research, it could be useful to compare the performance when these variables are transformed to another scale.

Thirdly, previous research shows that the reported ETA and the direction of a vessel are important features in predicting the ETA [1]. This is data that is also available in AIS, but was not available in the dataset for this research. For future generation of the dataset, it might be helpful to include those attributes as well.

Chapter 6

Appendix

6.1 Appendix A: Missing data attributes

Attribute	Number of missing values	Ratio of missing values
type	*784152*	*1.000000*
length	3546	0.004522
depth	512	0.000653
portname	214	0.000273
code	34684	0.044231
prttype	1020	0.001301
prtsize	1020	0.001301
status	968	0.001234
maxdepth	*783884*	*0.999658*
maxlength	*784152*	*1.000000*
annualcapa	*784152*	*1.000000*
humuse	7532	0.009605
locprecisi	1094	0.001395
latitude	214	0.000273
longitude	214	0.000273
iso3	4304	0.005489
iso3_op	15288	0.019496
country	12128	0.015466
lastcheckd	*768114*	*0.979547*
remarks	*781186*	*0.996218*
url_lca	*784152*	*1.000000*
source	*768896*	*0.980545*
createdate	244	0.000311
updatedate	244	0.000311
geonameid	244	0.000311
gdb_geomat	*784152*	*1.000000*
country_name	4856	0.006193
code_2	4856	0.006193
code_3	4856	0.006193
country_code	4856	0.006193
iso_3166_2	4856	0.006193
continent	4856	0.006193
sub_region	4856	0.006193
region_code	4856	0.006193
sub_region_code	4856	0.006193

Table 6.1: Attributes missing data

6.2 Appendix B: Formulas

6.2.1 Haversine formula code

```
def haversine(lat1, lon1, lat2, lon2):
```

```
    lon1, lat1, lon2, lat2 = map(np.radians, [lon1, lat1, lon2, lat2])

    dlon = lon2 - lon1
    dlat = lat2 - lat1

    a = np.sin(dlat/2.0)**2 + np.cos(lat1) * np.cos(lat2) * np.sin(dlon/2.0)**2

    c = 2 * np.arcsin(np.sqrt(a))
    km = 6367 * c
    return km
```

6.2.2 MAE formula

$$\frac{1}{N}\Sigma_{i=1}^{N}|actual_i - predicted_i|$$

6.2.3 RMSE formula

$$\text{RMSE} = \sqrt{\frac{\Sigma_{i=1}^{N}(predicted_i - actual_i)^2}{N}}$$

Bibliography

[1] Ioannis Parolas. Eta prediction for containerships at the port of rotterdam using machine learning techniques. 2016.

[2] Edwin Flapper. Eta prediction for vessels using machine learning. 2020.

[3] Vivek Kumar, B. Anil Kumar, and Lelitha Devi Vanajakshi. Comparison of model based and machine learning approaches for. 2014.

[4] Avigdor Gal, Avishai Mandelbaum, François Schnitzler, Arik Senderovich, and Matthias Weidlich. Traveling time prediction in scheduled transportation with journey segments. *Information Systems*, 64:266–280, March 2017. doi: 10.1016/j.is.2015.12.001. URL `https://doi.org/10.1016/j.is.2015.12.001`.

[5] Zhengyi Wang, Man Liang, and Daniel Delahaye. A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area. *Transportation Research Part C: Emerging Technologies*, 95:280–294, October 2018. doi: 10.1016/j.trc.2018.07.019. URL `https://doi.org/10.1016/j.trc.2018.07.019`.

[6] Bin Yu, Zhong-Zhen Yang, Kang Chen, and Bo Yu. Hybrid model for prediction of bus arrival times at next station. *Journal of Advanced Transportation*, 44(3):193–204, July 2010. doi: 10.1002/atr.136. URL `https://doi.org/10.1002/atr.136`.

[7] Raymond Hardij. Predicting arrival times for tankers ships using recurrent neural networks. 2018.

[8] Port of Rotterdam. Feiten en cijfers, 2022. URL `https://www.portofrotterdam.com/nl/online-beleven/feiten-en-cijfers`.

[9] United Nations Conference on Trade and Development. *Review of maritime transport 2021*. Review of Maritime Transport. United Nations, New York, NY, March 2022.

[10] Jean-Paul Rodrigue. *The Geography of Transport Systems*. Routledge, May 2020. doi: 10.4324/9780429346323. URL `https://doi.org/10.4324/9780429346323`.

[11] Yaser S. Abu-Mostafa. Hints. *Neural Computation*, 7(4):639–671, July 1995. doi: 10.1162/neco.1995.7.4.639. URL `https://doi.org/10.1162/neco.1995.7.4.639`.

[12] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. doi: 10.1038/nature14539. URL `https://doi.org/10.1038/nature14539`.

[13] Chigozie Nwankpa, Winifred Ijomah, Anthony Gachagan, and Stephen Marshall. Activation functions: Comparison of trends in practice and research for deep learning, 2018. URL `https://arxiv.org/abs/1811.03378`.