Universiteit Leiden
The Netherlands

Bachelor Informatica

Learning From Lifeguard Katwijk's Incident Reports:
Finding Patterns and Building a Predictive Model

Marit Kraijenoord

Supervisors:
Dr. Matthijs van Leeuwen & Dr. Francesco Bariatti

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

15/01/2026

Abstract

Safety along the beach of Katwijk in the Netherlands is one of Lifeguard Katwijk's main goals, and over the years, they have modernised it as much as possible. Unfortunately, the weather circumstances seem to make every summer day unpredictable, but in other areas of the world, such as France and China, researchers have effectively analysed external circumstances to find patterns among accidents on the beach, making it easier to take preventive measures.

In this thesis, we have used Lifeguard Katwijk's incident reports and the daily weather data to analyse the circumstances of an incident. Our main research question asked if there are patterns present in the data and if we can use those patterns to predict the number and type of incident. By using exploratory data analysis we wanted to find out if we can discover relations between the circumstances and the incident aspects. Additionally, we have used the data to learn a machine learning model that can predict the number of incidents in a day and the type of incident that has happened.

The results show that there are relations present between the number of incidents and the weather circumstances, such as the temperature and wind. Furthermore, we found that the predictive model could predict the number of incidents, although there are various steps for improvement of its effectiveness.

Keywords: beach incidents, incident prediction, predictive modelling, exploratory data analysis, data visualisation, regression model, classification model

Contents

1 Introduction

Lifeguard Katwijk is a Lifeguard Station located on the beach of Katwijk, in the west of the Netherlands. During the summer months, the lifeguards are available every day to give help to whoever needs it, and over the years, safety on Katwijk's beach has modernised. Vehicles and vessels are more robust, have more features and contain multiple tools the lifeguards can use to achieve their goal: keeping the people on the beach as safe as possible. Whenever an incident takes place, the lifeguards do their best to help them, and after the incident they record it to learn from it for the future. These incident reports contain relevant information surrounding the circumstances of the incident and the victim(s). Over the years, recording these incidents creates a dataset that contains all kinds of interesting aspects of the incidents.
For this thesis, we dive deeper into the incident reports to find if there are any patterns present that the lifeguards can use to take safety on Katwijk's beach to the next level. Similar studies have been done in France [TSGJ+22] and in China [LTZ+24], and there, researchers found that analysing external circumstances lead to the ability to predict under which circumstances an accident could likely happen. We want to find out if we can do a similar experiment with the data from Katwijk's beach, as it might behave differently from a beach in France or China. Finding positive results in Katwijk could not only help the lifeguards from Lifeguard Katwijk, but could also provide a basis for similar experiments along other beaches in the Netherlands.
To find out if we can discover associations between external circumstances and aspects of the incidents, we use exploratory data analysis. By visualising different aspects of the data and calculating statistical significance if applicable, we can find out if relations between variables are merely a coincidence or if there is correlation. Furthermore, we use machine learning models to learn a predictive model that can give an estimate of the number of incidents that will happen in a day, and also what type of incident it will likely be. By being aware of the dangers of certain circumstances, the lifeguards can take measures to prevent such incidents from happening, resulting in fewer victims overall.

1.1 Thesis overview

In this section we give an overview of what this thesis contains and where to find certain aspects. This section contains the introduction and an overview of the thesis; Section 2 covers the background and Section 3 provides insight into related work. Section 4 elaborates on the main research question and the sub-questions, while Section 5 provides information on the data processing aspect. Section 6 covers the research methods used and Section 7 elaborates on the experiments and their outcome. Section 8 concludes this thesis and provides possibilities for further research.
This Bachelor Thesis is a mandatory part of the Bachelor Informatica. It is done under the LIACS (Leiden Institute of Advanced Computer Science) department and main supervision is provided by Dr. Matthijs van Leeuwen, with Dr. Francesco Bariatti as the second supervisor.

2 Background

Katwijk is a small coast town located in the west of the Netherlands. Since 1923, Lifeguard Katwijk has been a steady presence on Katwijk's beach, trying their best to preserve safety on the beach. Lifeguard Katwijk started as a small association to teach people how to swim but since then, it has become more professional and extensive. Over time, beach surveillance has modernised to ensure the safety of the people visiting the beach, residents and tourists alike. Currently, Lifeguard Katwijk operates out of two stations, one on the north end of the beach and one on the south end of the beach. These stations contain everything that the lifeguards may need to do their job. During the summer months, their job includes manning the station daily from 9:30 to 18:00 and keeping an eye out for the visitors. This means patrolling the beach and the sea, either with a jeep, a boat, a rescue water craft, or on foot. The patrols do their best to keep people out of trouble but are also available when the public asks for their help.

Lifeguard Katwijk does not work alone. They are in close contact with other emergency services, such as the coast guard, police, ambulance, fire department and the Royal Dutch Rescue Company (KNRM). Furthermore, when an accident happens at the beach and bystanders call the emergency number, emergency services dispatch will alarm Lifeguard Katwijk. Lifeguard Katwijk's vehicles are well-suited to driving on sand, meaning they can more easily access places that might be difficult to reach with an ambulance, for example.

The Lifeguard stations are manned during the weekends from the end of May until the beginning of September, with the stations being manned daily from mid June until the end of August. Each day, the lifeguards come in at 9:30 and start the day by preparing, which includes checking the expected weather circumstances. Since the weather circumstances can change a lot from day to day, every day is different. There are a lot of factors that affect how many people will visit the beach that day. This includes the weather, such as the expected temperature, wind, rain and cloud cover, but it also depends on the time of the year. The beach is much more crowded during the holidays, as the local residents do not have to go to school or work, and tourists go on holidays to Katwijk.

During the day, many different types of incidents can happen. Each station has a first aid room, where people can go if they have a small wound or are not feeling well. In cases where the victim needs expert treatment, the lifeguards can send them to the doctor or call an ambulance if it is an emergency.

Not only first aid incidents happen, but there is also a possibility of other incidents happening, such as water sports enthusiasts getting into trouble because they underestimated the conditions, or because their gear breaks. Furthermore, during very crowded days, it might happen that parents lose their children, in which case the lifeguards start a search and rescue to find the child. In short, a lot can happen on the beach and the lifeguards are trained to respond to each situation appropriately.

To keep track of important data, Lifeguard Katwijk has created their own web-based system, which can bee seen in Figure 1. In this system, they can track the daily patrols and the number of small first aid incidents, but also the larger incidents. Smaller first aid incidents are situations where the lifeguards can solve a problem by applying first aid, without needing additional assistance. Any incident that requires additional assistance, either from lifeguards or from emergency services, is classified as a larger incident and requires an incident report to be made.

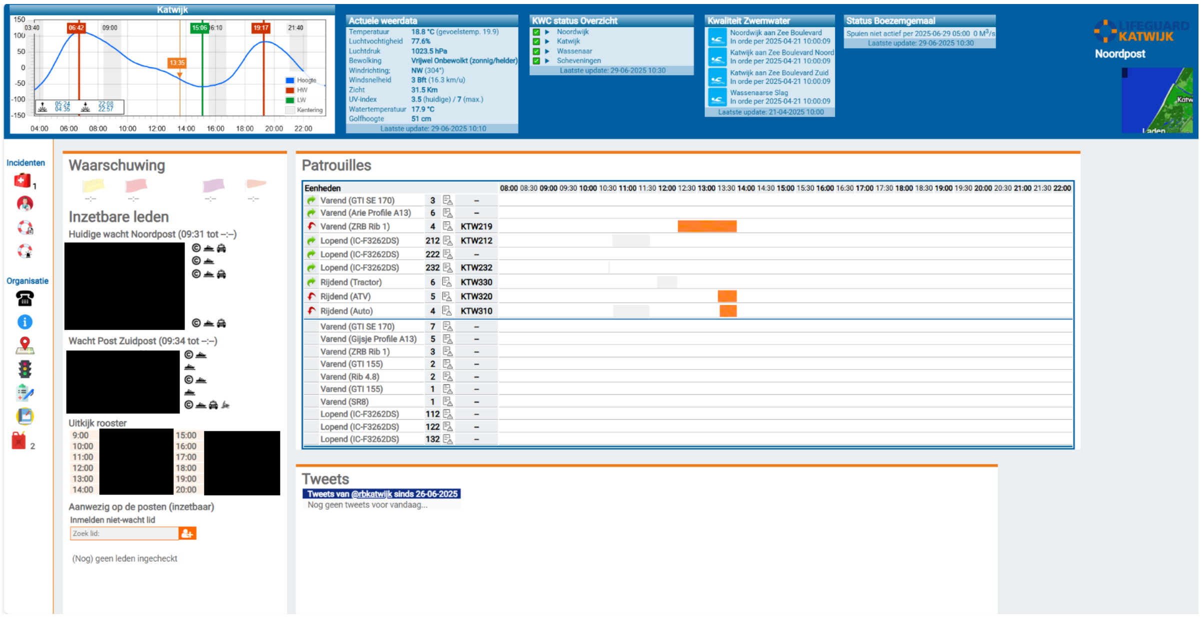


Figure 1: Lifeguard Katwijk Watch Report (Screenshot taken June 29th 2025, 13:35)

Some important aspects in this Watch Report, such as the tide (top left graph), patrols (center, “Patrouilles”) and current weather circumstances (top center, “Actuele Weerdata”), are updated throughout the day. Each time an incident takes place, the lifeguards fill out an incident report and write down any important information regarding the incident. This ranges from the weather conditions to the initial notification and the cause of the incident, to the result and the victim’s information. Additionally, the lifeguards present at the time of the incident are noted. In case of a very serious incident, aftercare is provided to the people involved in the incident if they need it. Further information regarding the specifics of the incident report can be found in Section 5.
For this thesis, we analyse the data from these incident reports to find patterns and develop a predictive model. This information can help the lifeguards when preparing for their day appropriately and take more preventive measures to ensure the safety of the visitors at the beach.

3 Related Work

The data used for this thesis is quite unique and therefore there is little related work to be found on data in the Netherlands. However, De Korte et al. (2021) [dKCT21] use a Bayesian network (BN) approach to model and predict shore-break-related injuries and rip-current drowning incidents based on external circumstances such as the weather, wave, tide, and beach morphology. Along the Gironde coast in south-west France, researchers analysed 442 drownings caused by rip currents and 715 injuries caused by shore-break waves. They used the environmental conditions at the time to train two separate Bayesian networks and found that the Bayesian network for predicting shore-break-related injuries systematically performed better than the rip current Bayesian network. Furthermore, they found that more surf zone injuries were observed on warm sunny days with light winds and long-period waves. There were more shore-break-related injuries at high tide with steep beach profiles, whereas more rip-current drowning happened at low tide with near-shore-normal wave incidence. As said in the paper: "Although the shore-break and rip-current BNs improve prior estimates, they still have a large percentage of wrong but confident predictions". This shows that there are still improvements to be made. Therefore, De Korte et al. (2021) [dKCT21] advocate to keep developing such Bayesian networks to gain a better understanding of hazard, exposure and life risk.

Tellier et al. (2022) [TSGJ+22] used data on rescues and drownings from 2011 to 2017 along the Gironde surf coast in France to predict the risk of drowning events based on seasonality, holidays, weekends, weather and oceanic conditions. They found that air temperature, wave parameters, seasonality and holidays were associated with drownings. They tested the effectiveness of their predictive daily model with 1, 2 and 3 days before the incident happened and found that they all yielded similar results. As mentioned in the paper: "The daily model had areas under the curves (AUCs) of 0.88 (95% CI 0.84–0.91) for 2011–2013 and 0.82 (95% CI 0.78–0.86) for 2015–2017". The daily model that is mentioned refers to the model that predicts the chance of drowning for a specific day, three days in advance. An AUC between 0.8 and 0.9 is considered very good and therefore they could conclude that drownings along the Gironde surf coast can be anticipated up to 3 days in advance, giving them the ability to take additional preventive measures. Tellier et al. (2022) [TSGJ+22] reveal that using external circumstances to predict the likelihood of drownings taking place was effective and can aid in preventing drownings.

Li et al. (2024) [LTZ+24] performed a comprehensive analysis of characteristics and factors for beach accidents to develop a predictive model in China. They focused on characteristics of the circumstances such as the age, gender and activity of beachgoers. Some of the potential factors they took into account were aspects such as meteorology, waves, tide, and beach morphology. They found that beach accidents occur mainly in summer with most accidents happening in the afternoon and evening. Furthermore, they found that "90% of accidents occur when the beach is at a high-risk level for rip currents". They trained three machine learning models to predict beach accidents, namely Support Vector Machines, Back Propagation Neural Network, and Random Forest Algorithms. They found that their Support Vector Machine model had an accuracy of approximately 78% in predicting "safe" and "dangerous" classes, whereas the Back Propagation Neural Network had an accuracy of 65% and the Random Forest Algorithms an accuracy of 57%. Li et al. (2024) [LTZ+24] show the effectiveness of using external circumstances to predict beach accidents.

Each of these papers reveals that using external circumstances to predict beach accidents can lead to useful discoveries and effective models. They find that the weather circumstances, such as air temperature, wind, tide, or beach morphology, can help in analysing and predicting incidents along the beach. We hope to find similarly positive results for Lifeguard Katwijk's incident reports and weather circumstances.

4 Research Questions

The focus of this thesis is to discover if there are patterns in the incident reports and if we can use those patterns to build a predictive model. Thus, the main research question of this thesis is:

What patterns can be found in the data from the incident reports from Lifeguard Katwijk, and how can we learn from those patterns to predict the number and type of incidents in a day?

This research question is quite broad and covers a lot of aspects of the data. Therefore, the main research question has been split into multiple sub-questions that cover the important aspects within the main research question.
First of all, we would like to find out if there are connections to be found that relate the external circumstances to the number of incidents in a day. External circumstances such as temperature, tide and wind might have different effects on the number of incidents. Therefore, the first sub-question is:

RQ1. Is there a relation between a single variable among temperature, tide and wind, and the number of incidents on a particular day?

Furthermore, we are interested in finding out if there are certain locations along the beach of Katwijk where more incidents have happened than in other locations. Thus, sub-question two is:

RQ2. Are there locations along Katwijk's beach where incidents happen more often than in other locations?

In addition to sub-questions that can be answered by exploratory data analysis, we also have sub-questions that are more likely to be answered by analysing a predictive model. Firstly, we are interested in knowing how the temperature, wind and rain affect the effectiveness of the predictive model. Thus, the third sub-question is:

RQ3. How do the temperature, wind and rain affect the effectiveness of a predictive model that predicts the number of incidents in a day?

As there are a lot of different types of incidents, we are curious to find out if we can predict the type of incident based on the weather circumstances. Thus, the last sub-question is:

RQ4. How effective is a predictive model in predicting what type of incident has happened, based on the weather circumstances?

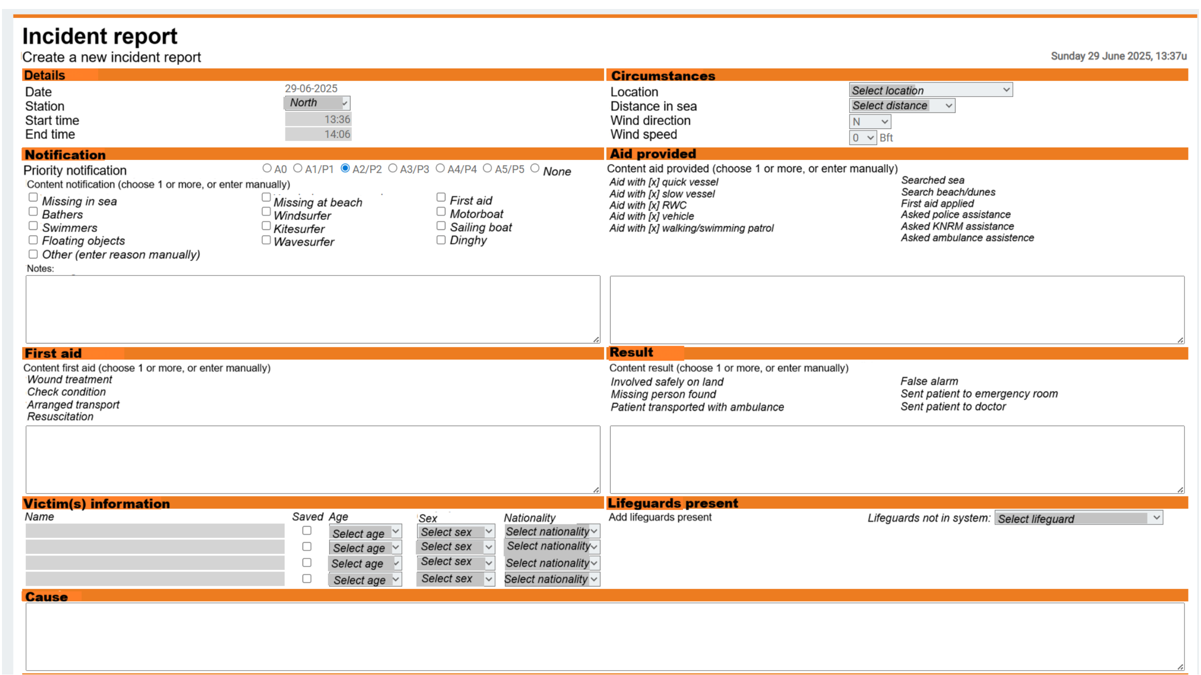
These sub-questions split the main research question into smaller parts that are easier to answer. The combined findings of the sub-questions gives us the information needed to answer the main research question.

5 Data

To answer the research questions, we have to prepare the data such that it can be used for the experiments. Since the incident report data is quite unique, we explain in Section 5.1 what the data looks like and what can be expected when working with it. Furthermore, Section 5.2 elaborates on the data that was needed in addition to the incident reports. Section 5.3 explains how we have cleaned the data to make it usable for the experiments. Lastly, Section 5.4 shows some visualisations to give an idea of what the data contains.

5.1 Data Explanation

The most important data in this project is the incident report data. Since May 1st 2005, the lifeguards have had to fill in an incident report whenever an incident happened that required additional assistance. This form has changed a lot over the years, but the current form can be found in Figure 2. The aspects of this form have been translated to English, but the original form is in Dutch. The data runs from May 1st 2005 and for this project, the data up until September 1st 2025 is used. This means that there are 1869 incident reports to work with.



Incident report

Create a new incident report

Sunday 29 June 2025, 13:37u

Details

Date 29-06-2025
Station North
Start time 13:36
End time 14:06

Circumstances

Location Select location
Distance in sea Select distance
Wind direction N
Wind speed 0 Bft

Notification

Priority notification A0 A1/P1 A2/P2 A3/P3 A4/P4 A5/P5 None

Content notification (choose 1 or more, or enter manually)

Missing in sea, Bathers, Swimmers, Floating objects, Other (enter reason manually), Missing at beach, Windsurfer, Kitesurfer, Wavesurfer, First aid, Motorboat, Sailing boat, Dinghy

Notes:

Aid provided

Content aid provided (choose 1 or more, or enter manually)

Aid with [x] quick vessel
Aid with [x] slow vessel
Aid with [x] RWC
Aid with [x] vehicle
Aid with [x] walking/swimming patrol
Searched sea
Search beach/dunes
First aid applied
Asked police assistance
Asked KNRM assistance
Asked ambulance assistence

First aid

Content first aid (choose 1 or more, or enter manually)

Wound treatment
Check condition
Arranged transport
Resuscitation

Result

Content result (choose 1 or more, or enter manually)

Involved safely on land
Missing person found
Patient transported with ambulance
False alarm
Sent patient to emergency room
Sent patient to doctor

Victim(s) information

Name	Saved	Age	Sex	Nationality
		Select age	Select sex	Select nationality
		Select age	Select sex	Select nationality
		Select age	Select sex	Select nationality
		Select age	Select sex	Select nationality

Lifeguards present

Add lifeguards present

Lifeguards not in system: Select lifeguard

Cause

Figure 2: Lifeguard Katwijk incident report (Screenshot taken June 29th 2025, 13:37)

The most important aspects of each incident are recorded in the incident report. This includes the date, the start time, the end time, and from which station the incident was handled, which can be either the North station or the South station. Additionally, the circumstances are noted, which includes the location, distance in sea (if applicable), wind direction and wind speed. Some of these aspects are filled in automatically, as they are taken from the Watch Report seen in Figure 1, such as the wind direction and wind speed. The location is a drop-down menu where they can choose the

location based on the buildings on the beach. This includes the restaurants, the Lifeguard stations and the surf schools, but also the option to say the incident happened somewhere else in Katwijk, like the coastal road, or outside of Katwijk, either to the north or to the south.
After the circumstances are noted, the notification, response, first aid applied and result of the incident are written down. The notification is the initial notification that was provided, which can be a visitor that called for help, or one of the lifeguards seeing something happen from the station. This does not have to be the same as what actually has happened. For example, the initial notification can be a person missing in sea, but if the person was found on land, the notification stays the same. The response contains the help the lifeguards have provided. This ranges from their own vehicles and/or vessels, or the help of police, ambulance or KNRM. The first aid that was applied is put in a separate section so that the lifeguards are able to elaborate if needed. Lastly, the result notes what happened to the victim(s) after the help was given, such as them being found if they were lost, or them being taken to the hospital by the ambulance.
In addition to these aspects from the incidents, any important information from the victim(s) is stored and the lifeguards that were present are noted. Lastly, if applicable, the cause and any additional information is stored. The cause might range from a rip current to underlying conditions in the victim, although there is not always a specific cause for an incident. The extensiveness of the information noted depends on the severity of the incident. As can be seen in Figure 2, the lifeguards have the ability to select common notifications or responses to an incident, but also to elaborate on certain situations. For example, sometimes a bystander might have helped a victim initially, which is also noted.
The current incident report tool is quite automated and easy to use, but twenty years ago it was a little different. Back then, the lifeguards had to fill in everything manually and none of the aspects had a list of possible situations or a drop-down menu. This means that the incident reports from then contain very human language, which means there were also spelling and grammar mistakes. This makes it easy to read, but harder to use it when doing the experiments. To be able to work with the data, it had to be structured properly.

5.2 Data Gathering

The incident reports contain a lot of relevant data, but not everything is present in the reports. Some weather circumstances are missing, which means we had to gather those to get a more complete view of each incident.
The first missing aspect is the temperature. It is present in the Watch Report (Figure 1), but it is not automatically added to the incident reports. Luckily, the Royal Dutch Meteorological Institute (KNMI) gathers this data and on their website [KNMc] we can find the data we need. Unfortunately, the closest station to Katwijk, station 210 (Valkenburg Zh), does not provide any data after May 3rd 2016. This means that part of the air temperature data (01-05-2005 until 02-05-2016) is from station 210 (Valkenburg Zh), and part of it (03-05-2005 - 01-09-2025) is from station 215 (Voorschoten), which is the next closest station. The air temperature is not only relevant for each incident report, but also for the experiments. The daily air temperature is also needed when no incidents took place.
In addition to the air temperature, we need the wind (speed and direction) and rain data from that period. The wind data can be found on the KNMI website as well [KNMb], where we take the daily average data of the wind speed and wind direction. The KNMI website also provides the rain

data [KNMc], where we take the average rain fall per hour.
Lastly, we want to add information about the tide at the moment of the incident to each incident report. In this case, both the tide (high, low, or changing) and the direction (north, south, or changing) is important. This data we can find on the Rijkswaterstaat website, under "waterinfo" [Rij]. The closest station that measured for the period needed is the one in Scheveningen, thus we gathered the data from that station.
As the air temperature, wind, rain, and tide complete the relevant weather circumstances, we now have all the data needed te get a clearer overview of each incident, but also on the days surrounding it. In Section 5.3, the process to structure and clean the data is explained more thoroughly.

5.3 Data Preprocessing

After all the data has been gathered, it needs to be cleaned and structured appropriately. Some data is irrelevant to this project, and other aspects need to be more structured to be able to use it. The incident report data contains all the information that is entered when writing an incident report. However, some of these entries contain private data. Therefore, some manual changes have been made to the original incident report data, so that it can be used for the experiments in this thesis. First of all, there are some irrelevant columns that can be dropped immediately. Examples of these columns are: "priority", "press information", "details", "lifeguards present", "report written by", and more. These columns either contain information that is not applicable for this thesis, or privacy sensitive information. Then, each column is assigned the equivalent English name and the correct data type. Lastly, some columns undergo changes to their values so that the data is more structured and easier to use.
The following sub-sections will explain the changes made to the relevant columns.

5.3.1 Distance in Sea

The first column undergoing changes is the "distance in sea" column. This column contains integers, but depending on the initial value, this might be the actual distance in sea, or the already pre-assigned value. For uniformity, we have created a binning that assigns each distance in sea to a pre-defined range, so that the distance in sea is an ordinal variable.

5.3.2 Notification, Location, Result and First Aid

The notification column contains the initial notification as it was first provided. In most cases, this means that it is written down in very human language, which makes it hard to categorise. Therefore, a file is provided that has already categorised all notifications in the incident reports. When preparing the data, each incident is categorised based on the corresponding category in the provided file. The same is done for the location, result and the first aid that was provided. After this conversion is done, the incident reports contain a short text with the relevant information for that incident. Furthermore, each category is assigned a corresponding numerical value. The tables with categories for each column can be found in the Appendix (Section 8). The categories for the notification column can be found in Table 11. The locations have all been changed to the current location names, these can be found in Table 12. The categories for the result can be found in Table 13. The first aid that was provided during the incident is also categorised, this can be found in Table 14.

5.3.3 Aid Provided

To convert the column "aid provided", several columns are added. These columns are named after the vehicles, vessels or emergency instances that can provide aid. These columns and the original "aid provided" column are assigned a value, which is either 0 or 1. When the assigned value is 1, it means that the instance that the column is named after was present at the incident. If at least one instance is present, the "aid provided" column is assigned a 1 as well. Multiple vehicles and emergency services can be present at the same incident. This conversion makes it easier to see which instances were present at the incident.

5.3.4 Weather Circumstances

Lastly, the weather circumstances that are missing from each incident have to be added. Therefore, part of the data preparation includes structuring the weather data that was gathered and then adding it to each incident.
Firstly, the temperature has to be added. This is done by comparing the date and hour of the start time of the incident with the date and hour in the air temperature data. Then, we add the tide. This is done by calculating the time from the last high or low tide until the start time of the incident. High and low tide do not take the same time to complete and therefore the calculation differs depending on the previous tide. The duration of each tide can be found in Table 1.

Time	Tide	Tide direction
4 hours	High tide (highest point at start)	North
1 hour	Changing tide	Changing
5 hours	Low tide (lowest point after 2 hours)	South
1 hour	Changing tide	Changing
1.5 hours	High tide (highest point at end)	North

Table 1: Duration of different stages of the tide

By calculating the time since the last tide, we can find the tide and direction of the tide at the start time of the incident.
After adding both tide and temperature, we also add the volume of rain that fell during the hour of the start time of the incident in millimetres. Lastly, we change the wind direction. We want each incident to contain the wind direction as the degrees corresponding to the direction (e.g. north changes from N to 360°). Additionally, we add the wind direction as a sinus- and cosinus-component based on the degree value of the wind direction. This is done so that the proximity of the wind directions is correctly shown. For example, north (360°) and north-east (45°) are very close to each other, but their degrees do not reflect that. Adding the sinus- and cosinus-component helps in preserving that proximity.
After making all these changes, the incident reports are more complete and the weather circumstances data is structured so it is easy to work with.

To give an idea of how the incident report data looks before and after running the data preparation, the Appendix (Section 8) contains Figure 14, Figure 15 and Figure 16. These Figures show the initial incident reports and the final versions.

5.4 Data Visualisation

This section shows some data visualisations to give an idea of what the data contains. Figure 3 shows the number of incidents per year. We can see that there are some years where the number of incidents is higher or lower than in other years, but the average number of incidents in a year is 89 incidents. We can also see that since 2022, there has been a steady decrease in the number of incidents per year. This could be due to the lifeguards getting better at preventing incidents from happening, but there might be other reasons contributing to this decrease in incidents.

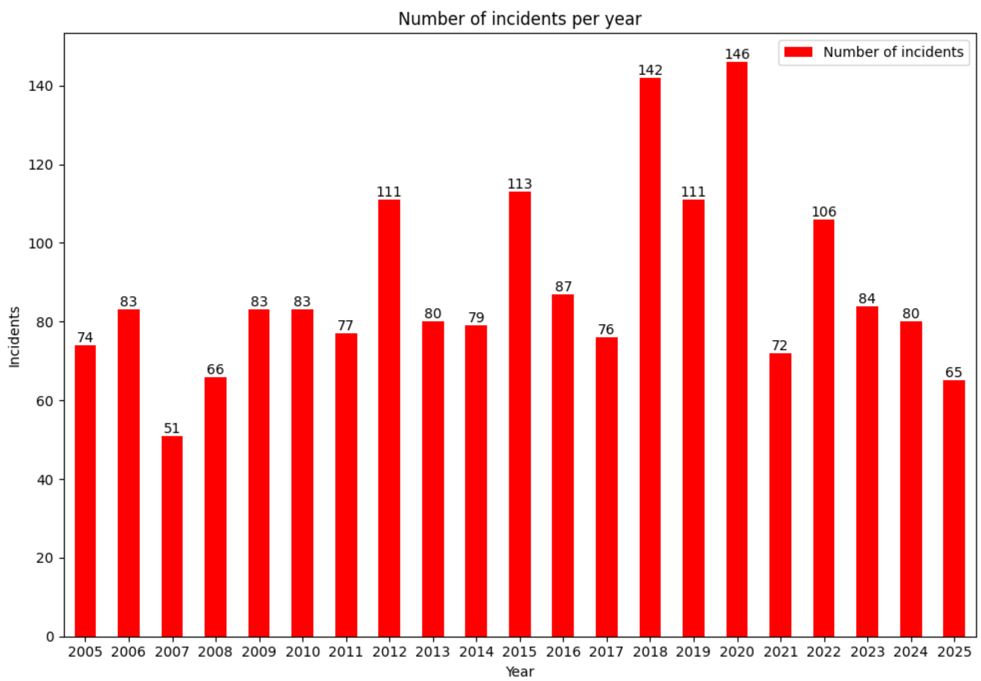


Figure 3: Number of incidents per year in the period 2005-2025

Figure 4 shows the number of incidents that have happened in each month, over the period 2005 until 2025. We can see that the months June, July and August have the most incidents, which is to be expected, as most people visit the beach in those months, as the temperature tends to be higher.

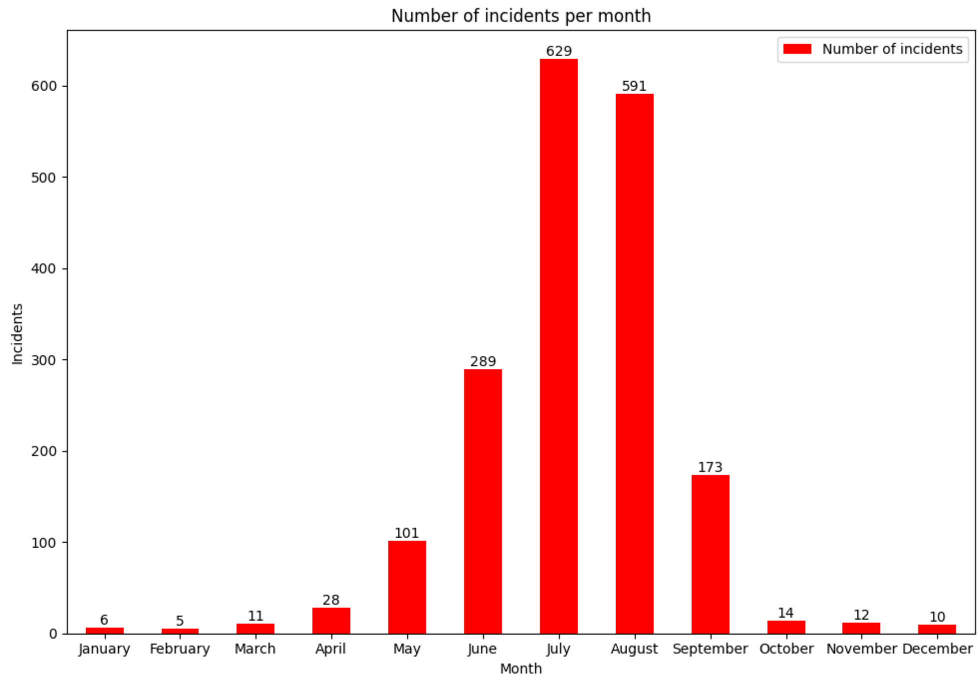


Figure 4: Number of incidents per month in the period 2005-2025

6 Methods

This section elaborates on the methods used for both the Exploratory Data Analysis (EDA) (Section 6.1) and the Machine Learning Models (Section 6.2). These two aspects combined will aid in answering the research questions. RQ1-2 can be answered by using EDA, while RQ3-4 can be answered with the ML Models.

6.1 Exploratory Data Analysis

We start by using exploratory data analysis (EDA) as a way of visualising the data. This will help in understanding the data contents and if there are any associations between external factors and aspects of the incidents. Furthermore, by creating plots from the data, we can answer some of the research questions.

An example of an association that we expect to see in the data would be the number of incidents compared to the maximum temperature in a day. If there would be no relation, we would expect the average to be the same for each temperature. Another example would be the percentage of incidents that took place in sea compared to the wind speed. If there was no association between the two, we would expect the percentage of incidents in sea to be the same for each wind speed.

To decide if there is actually a relation between two variables, we calculate the statistical significance with $p < 0.05$. This means that if the p-value of the relation is less than 0.05, there is a more than 95% chance that it is because there is a relation between the two variables and thus a less than 5% chance that those values are a coincidence. To calculate the statistical significance, we use the Chi-Square test for a 2 by 2 contingency table and the Chi-Square test for Goodness of Fit.

By using exploratory data analysis to investigate the data, visualise possible associations and calculate statistical significance, we have a clearer overview when developing the predictive models. Furthermore, we answer RQ1 and RQ2 using exploratory data analysis, as these questions focus on associations within the data.

6.2 Machine Learning Models

In addition to doing exploratory data analysis, we use predictive machine learning models to find out if we can effectively predict several aspects of incidents based on external circumstances. The models that will be used are Linear Regression, XGBoost, Regression Tree and the Decision Tree Model. These models will be developed using the python scikit-learn library [PVG+11].
For each model, we only take into account the data from the months June, July and August. We do this so that there will be a more representative distribution of incidents, as most incidents take place during those months (see Section 5.4). Including the other months means that there will be a lot of days without any incidents, which can cause overfitting on the training set. Furthermore, the lifeguards are present at the stations every day in that period, meaning people are far more likely to ask for their aid if they need it.

6.2.1 Linear Regression Model

A Regression Model takes a set of one or more input features and aims to predict a continuous output variable. Thus, a Linear Regression Model [MPV21] takes the input features and tries to find a linear relationship between those and the output variable. In this thesis, we use the Least Squares Linear Regression Model to predict the number of incidents in a specific day, based on weather circumstances. To be able to do that, we have gathered all days in the months June, July and August and determined the number of incidents on each date in those months for the time period 2005-2025. Then, we set different combinations of the external circumstances (temperature, wind, rain) as the input features, meaning we develop multiple Linear Regression models, one for each set of input features. This means we have the ability to compare the results of each combination of input features, giving us the ability to compare which (combination of) input features provides the best results.
We have chosen the Linear Regression Model because it is a very simple and easy-to-implement machine learning model. The other Regression Models used are more complex, but do not necessarily have to perform better. The Linear Regression Model is included to serve as a baseline for the other models and to find out if there is a linear relationship between the input features and the prediction variable. Each of the Regression Models will be compared to find out which model performs best. These results will be used for answering RQ3.

6.2.2 XGBoost Regression Model

The XGBoost Regression Model [Wad20] is a Regression Model that takes a set of input variables and uses those to predict the pre-defined outcome variable. It differs from the Linear Regression Model in that it keeps improving itself by trying to minimise a loss function on the train set. In this thesis we have used the Root Mean Squared Error (RMSE) as the loss function that the XGBoost model tries to minimise as we also analyse the RMSE in the experiments. By optimising the RMSE,

we can compare it to the other models.
The XGBoost model uses the same dataset and train/test split as the Linear Regression Model, so that we can fairly compare them. Furthermore, the XGBoost model needs additional hyperparameters. The objective is to minimise the RMSE and the tree method is the histogram. Additionally, we train the XGBoost model for a maximum of 100 rounds, with an intermediate score shown every 10 rounds. If the best score is found before the 100 rounds are over, the model stops earlier. These hyperparameters are the exact same for each variation of the XGBoost Regression Model.
As with the Linear Regression Model, we use external circumstances as input features, with each variation providing different results that we can compare.

6.2.3 Regression Tree Model

The Regression Tree Model [Fla12] is also a Regression Model that takes a set of input variables to predict a pre-defined output variable. It is a Decision Tree that can predict continuous numerical values by recursively splitting the data on the input features. After a pre-defined level has been reached, it takes the mean of the output variable of all instances in a certain leaf. By visualising the Decision Tree we can see what features and values result in the best split, leading to insights on how a prediction came to be.
As with the other Regression models, we use a dataset from the months June, July and August in the time period 2005-2025. In the dataset, each day is included with the weather circumstances and the number of incidents for that day. Additionally, this model needs a set maximum depth of the tree, which was set to 4 for these experiments. For the input features we use different combinations of the weather circumstances included (temperature, wind and rain). After developing all three Regression models, we can compare them.

6.2.4 Classification Decision Tree

A Decision Tree [Fla12] is a classification model that works similarly to the Regression Tree Model. A classification model takes a set of input features and then classifies an instance into 2 or multiple classes. It does not predict a continuous variable, but rather a categorical variable. At each level, it decides on the feature and value that results in the best split of the dataset. At the pre-defined level, each leaf will have several instances left. A perfect Decision Tree would have each leaf only contain instances of the same category. However, this is very unlikely and thus each leaf gets assigned the category that has the most instances.
In this thesis we use a Decision Tree to predict the type of incident for each instance, to answer RQ4. Since there are 32 types of incidents and most types do not contain more than 100 incidents, we focus on two types. We create a Decision Tree that predicts whether an incident is a first aid incident or not, and a model that predicts whether an incident happened in sea or not. These categories are quite prevalent in the data set, with 853 out of 1869 (45.6%) incidents being first aid incidents, and 766 out of 1869 (41.0%) incidents having happened in sea.
As with the Regression Tree Model, we can visualise the Decision Tree so that we can analyse the decisions it has made. Furthermore, we can see what features provided the best split.
We use the same train/test split as with the previous models. The maximum depth is 3 and each leaf must have a minimum of 5 samples. We train each Decision Tree twice, once on Gini Impurity and once on Entropy [SSBD14]. Gini Impurity is calculated by how often a random sample would

be mislabelled if it was assigned by class probability. The lower the Gini Impurity, the better the model, with a Gini Impurity of 0 being a perfect score. Gini Impurity can be biased towards classes with more instances during split selection, which is why we also train a model on Entropy. Entropy calculates the uncertainty in a node's class distribution. A perfect score for Entropy is also 0. By training a model on each of these impurity measures, we can analyse if one of the measures favours a certain class more heavily than the other.

Section 6.2.6 will cover all evaluation metrics used in each model.

6.2.5 Input/Output Features

For each of the models, both the Regression models and the Classification model, we use the same input features. These features are the temperature (in degrees Celsius (°C)), the rain (in millimetres (mm)) and the wind. The wind is split into wind speed and direction. The wind speed is measured on the Beaufort scale [KNMa], which ranges from 0 (calm) to 12 (hurricane). The wind direction is split into its sinus- and its cosinus-component.
For each model we use the same split of the data, where 80% of the data is in the train set, and 20% is in the test set. For each Regression Model, we do an experiment where we randomise the order of the data using the same random seed. Additionally, for the Linear Regression Model we do one experiment without randomising the order of the dataset. We do this so we can see if using the first 80% of incidents is effective for predicting the last 20% of incidents. This information is important as it shows whether it would be effective to use the models in the future where the full dataset would be used as the train set to predict outcome variables in the future.

6.2.6 Evaluation Metrics

To test the effectiveness of each model, we use different evaluation metrics, depending on the model's task. For each of the Regression Models we compute the following metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and the R^2-score. The MSE is the average of all squared errors, which is the difference between the actual value and the predicted value. It is easy to compute but is very sensitive to outliers. Therefore, we also calculate the RMSE, which is the square root of the MSE. Like the MSE, it penalises large errors, but as it is on the same scale as the original data, it provides a more intuitive measure of the magnitude of the error. Both the MSE and the RMSE should have a score as close to 0 as possible. Finally, we also calculate the R^2-score, which is the proportion of the variance in the output variable that can be explained by the model. It assesses the overall model fit and shows how well the input features explain variation in the output variable. We can see it as a percentage, where it represents the percentage of variance that can be explained. The R^2-score should be as close to 1 as possible.
For the Regression Models we show all three of these scores, for both the train set and the test set. We include the train set to see if the model is not overfitting on the train set, which would show through much better scores on the train set than on the test set. We do not expect the models to perform better on the test set than on the train set, but the differences should not be very large. We evaluate the Regression models by comparing the R^2-scores for each model and combination of input features. From there, we can see which of the three Regression Models performed best.
The Decision Tree Classification Model is evaluated on different metrics than the Regression Models,

because its goal is different. For this model, we use Precision, Recall, the F1-score, Accuracy and Area Under Curve (AUC). Precision is the number of instances that was assigned a specific class (e.g. 1) and belongs to that class, divided by the number of instances that was assigned that specific class. In short, it is the number of True Positives divided by the sum of the number of False Positives and the number of True Positives. A higher Precision means a higher proportion of instances assigned to that class actually belong to that class. It does not take into account the number of instances that belong to that class but were not assigned. Recall is calculated by retrieving the True Positives (assigned and belongs to class) and dividing it by the sum of the number False Negatives (not assigned but does belong to class) and the number of True Positives. A higher Recall means that a higher proportion of instances was correctly assigned to that class, and fewer positive instances were missed. By combining Precision and Recall, we get a more complete overview of the effectiveness of the model. The F1-score is a combination of Precision and Recall and is therefore more balanced than using only of the two. The Accuracy is the number of correctly assigned instances, taken over all instances. Lastly, we calculate the Area Under the Curve (AUC), which measures a model's ability to distinguish between the classes. It is defined by the calculation of the area under the Receiver Operating Characteristic (ROC), with the ROC showing the model's relationship between its True Positive Rate and False Positive Rate at different thresholds. Using a threshold of 0.5, any AUC value higher than that means the model performs better than simply guessing. A value close to 1 is best for the AUC, but also for the Precision, Recall, the F1-score and Accuracy.

By explaining all Machine Learning models, their input and output, hyperparameters, and evaluation metrics used, we aim to make the experiments reproducible.

7 Experiments

To answer the research questions we visualise parts of the data in Section 7.1. Section 7.2 contains Machine Learning models experiments.

7.1 Exploratory data analysis

Research sub-questions 1 and 2 can be answered by visualising the data and these visualisations are discussed in this sub-section.

7.1.1 Temperature and Number of Incidents

RQ1 asks if there is a relation between external circumstances and the number of incidents that happens per day. This section covers the maximum temperature in a day and if it affects the number of incidents that day.

Figure 5 shows the number of incidents per year in the period 2005 until 2025, the number of days where the temperature was higher than 20.0°C at one point and the average maximum temperature in a day during the summer months (June, July and August). We can see that on average 72 incidents happen during those months and the average summer has approximately 50 days where the temperature is higher than 20.0°C at some point during the day. The average maximum temperature in a day during the summer months is represented by the red line. We can see that in some years,

there is a clear correlation between the average maximum temperature and the number of incidents. For example, 2006, 2018 and 2022 all experience a warmer summer on average, and also had more incidents than the neighbouring years. Additionally, 2007, 2008, 2011, 2021 and 2024 experienced a low in average temperature, but also had fewer incidents than the neighbouring years.
Another interesting aspect is the two peaks in incidents in 2018 and 2020. 2018 had 110 incidents happen during the months June until August and also had one of the warmest summers on average in that time period. However, 2020 had fewer warm days, but a higher number of incidents. A possible reason for this peak in the number of incidents could be the COVID-19 virus that was present that summer. People were not allowed to meet others inside, but outside was permitted. Even though the days might not have been as warm as in other years, it is likely that people went to the beach to meet up with friends and family. Because there were more people on the beach, there is a higher likelihood of an incident happening.

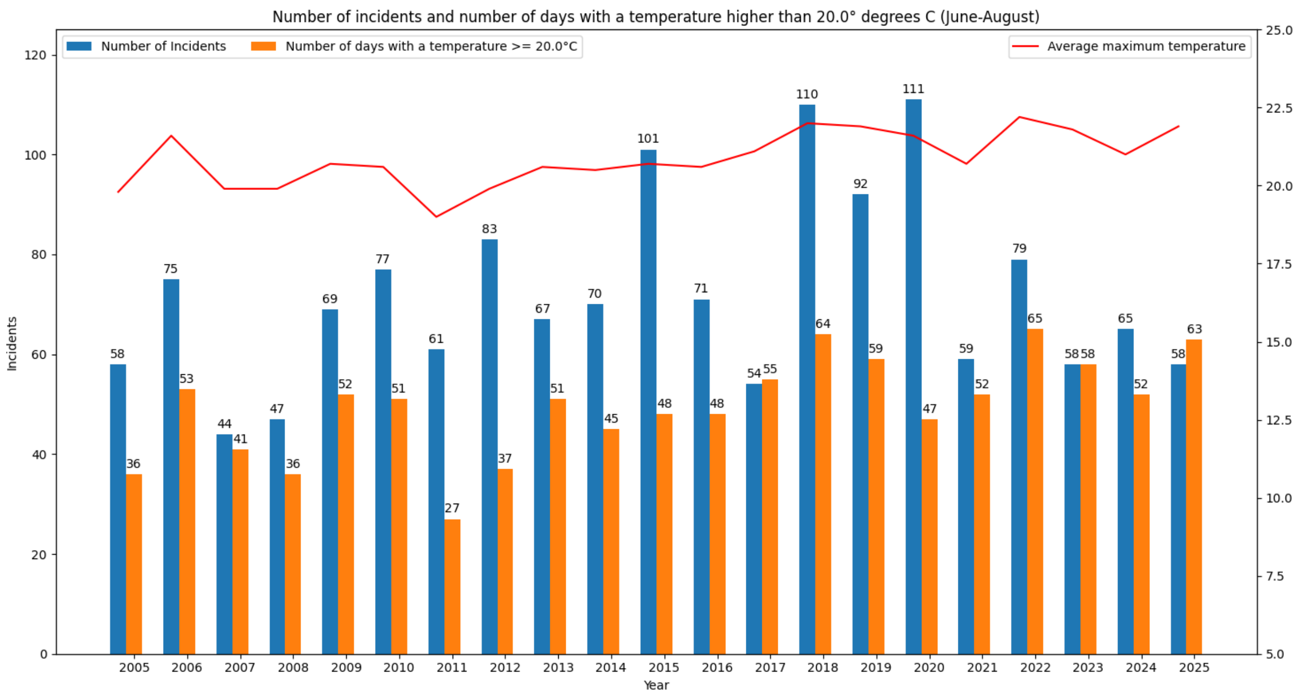


Figure 5: Number of incidents, number of warm days (temperature higher than 20.0°C) and average temperature per year in the period 2005-2025 (June, July and August)

RQ1 asks if there is a relation between the maximum temperature in a day and the number of incidents that happen. Figure 6 shows the average number of incidents (the red line), and the number of days (blue bars) per maximum temperature. To create this graph, the highest temperature and the number of incidents was determined per day in the time period 2005-2025, for the months June, July and August. Then, each day was assigned to a temperature range as seen on the y-axis in Figure 6. For example, temperature 16.0°C contains all days that had a maximum temperature ranging from 15.6°-16.5°C. After this, the number of incidents over all days with the same temperature is aggregated and divided over the total amount of days with the same temperature. Figure 6 shows that there is a clear relation between the maximum temperature in a day and the average number of incidents in that day. The higher the temperature, the higher the number of incidents in a day likely will be, based on the results in Figure 6.

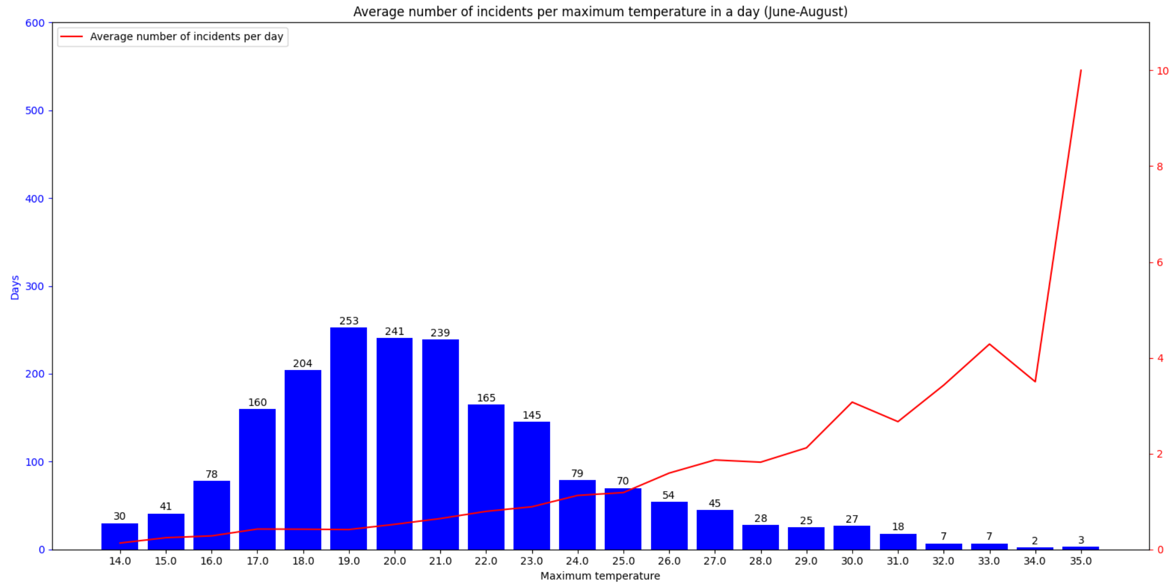


Figure 6: Average number of incidents for a maximum temperature in a day (June, July and August)

7.1.2 Tide and Number of Incidents in Sea

RQ1 asks how the tide affects the number of incidents happening in sea. Table 2 contains the number of incidents at each tide and also splits them into incidents that happened in sea and incidents that happened on land. An important thing to note is that high and low tide do not take the same time, high tide completes in 5.5 hours while low tide takes 5 hours. Therefore, we have included two columns where the number of incidents is divided by the time it takes to complete the tide, so we can fairly compare them.

If we look at all incidents, we can see that the number of incidents per hour in each tide is almost equal. From this we can conclude that tide likely has little impact on the number of incidents in general. However, we can see a difference in both the incidents that took place in sea and the incidents that did not. To find out if there is a significant difference, we performed a Chi Square test on Table 2 and found that the Chi Square statistic is 8.3116. The p-value is 0.003939, which means it is significant at $p < 0.05$. Therefore, we can say with a certainty of more than 95% that there is a relation between the number of incidents that happen in sea and the tide, with high tide causing more incidents in sea.

7.1.3 Wind and Number of Incidents in Sea

For RQ1 we want to find out if there is a relation between the wind speed and the number of incidents that happens in sea. Table 3 shows the number of incidents per wind speed and the number of incidents per wind speed that took place in sea. The third row shows the percentage of

	High tide	Low tide	High tide (per hour)	Low tide (per hour)
All incidents	829	753	150.7	150.6
Incidents (land)	475	486	86.4	97.2
Incidents (sea)	354	267	64.4	53.4

Table 2: Number of incidents per tide, split between sea and land and calculated for the number of incidents per hour of tide

incidents that took place in sea per wind speed.
We expect that people need help in sea more often when the wind speed is higher. In case of an eastern wind, people on floating objects get pushed into sea faster when the wind speed is higher. When the wind blows from the west, waves tend to be higher and stronger, causing people to underestimate their ability to get back to land safely. Additionally, when there is either a northern or a southern wind, the current along the coast tends to be stronger. This is especially the case when the wind and the underlying current (dependent on the tide) point in the same direction. When there is a northern wind (which blows towards the south) and the tide is low, that means the direction of the current is towards the south. This combination causes a very strong current, which can not be seen from the land. In short, for each direction, we expect a higher wind speed to cause more problems. As we expect a higher number of incidents with a higher wind speed for each wind direction, we combine them for this experiment. In addition to the effects the wind direction has on the sea, there is another reason why we expect more incidents to happen in sea when the wind speed is higher. Water sports enthusiasts often go out to sea when the waves are high enough to practise. The height of the waves is dependent on the wind speed, where waves tend to be higher when the wind speed is higher. In short, for all wind directions, we expect the number of incidents in sea to be higher when the wind speed is higher.

	0	1	2	3	4	5	6	7	8
All incidents	44	87	310	673	448	228	63	13	3
Incidents (in sea)	12	20	100	227	194	123	43	8	2
% of total in sea	27.3%	23.0%	32.3%	33.7%	43.3%	53.9%	68.3%	61.5%	66.6%

Table 3: Number of incidents per wind speed, split for total and incidents located in sea

Table 3 shows the number of incidents per wind speed, in total and for all incidents in sea. The third row shows the percentage of incidents for that wind speed that took place in sea. We can see that there is clear difference between the lower wind speeds and the higher wind speeds. With these results, we have used a Chi Square Goodness of Fit test to calculate if there is a statistical difference between the number of incidents in sea depending on the wind speed. If there is no relation, we expect the percentages of incidents in sea to be approximately equal for each wind speed. The Chi Square Goodness of fit test resulted in a p-value of < 0.001, which is smaller than $p < 0.05$, meaning these differences are significant. Therefore, we can conclude that there is a less than 5% chance that these numbers are a coincidence, and that it is very likely that there is a correlation between the number of incidents that happen in sea and the wind speed, where higher wind speeds cause more accidents to happen in sea.

7.1.4 Location and Number of Incidents

RQ2 asks whether there are locations along Katwijk's beach where incidents happen more often than in other locations. To answer this sub-question, we have created Figure 7. This figure shows the number of incidents, grouped by the location they were assigned to, in the order the locations appear on the beach. Here, Airtime is located most north and Skuytevaert most south of all. There are two significant peaks, at "Noordpost" and "Zuidpost". These are the North and South stations respectively, so a peak in the number of incidents there would be expected. For any first aid incident where an expert's opinion is needed, the lifeguards have to fill in an incident report. This can mean calling an ambulance for the victim, or sending them to the doctor. Since each station has their own first aid room, people can come to the station if they need first aid. The peaks at "Noordpost" and "Zuidpost" are likely there because a lot of those incidents were first aid incidents where an expert's opinion was needed.

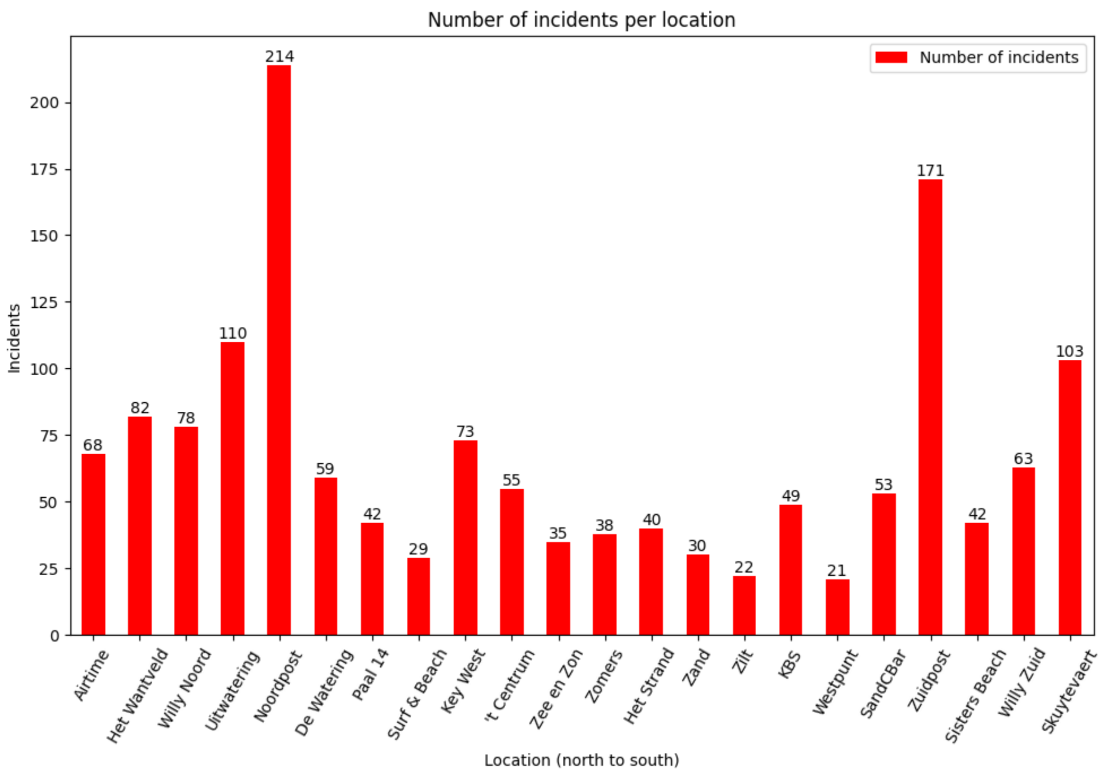


Figure 7: Number of incidents grouped per location, order as it is on the beach, with Airtime located most north and Skuytevaert located most south

The next highest peak is at "Uitwatering". The "Uitwatering" is not a beach restaurant or a surf school, but it is where the old Rhine flowed in to sea. The river water flowing into the sea might cause some unexpected currents, causing the high number of incidents happening there.

In Figure 7, there are three surf school locations. These are "Airtime", "KBS" and "Skuytevaert". In these three locations, there are often more water sports enthusiasts than in other locations. This might explain the peak at "Skuytevaert", but there are no peaks at "Airtime" and "KBS".

Looking at this figure, there seems to be two locations where incidents happen significantly more often, but when diving deeper, these peaks can be explained. Therefore, from Figure 7, we can not conclude that there are locations where incidents happen more often than in other locations due to external circumstances.

7.2 Predictive modelling

In addition to using graphs to find patterns in the data, we have created Machine Learning models to help with answering the research questions.
For RQ3, we want to find out if external circumstances can accurately predict the number of incidents in a day. This research question will by answered by using Regression Models. Section 7.2.1 covers the Linear Regression Model, Section 7.2.2 the XGBoost Model and Section 7.2.3 the Regression Tree Model. RQ4 asks if we can predict the type of incident based on external circumstances. As this is a classification problem, we use a Decision Tree, which is covered in Section 7.2.4.

7.2.1 Linear Regression Model

As mentioned in Section 6, we create two Linear Regression Models, one where the data order is randomised, and one where it is not. The results for the Linear Regression Model with the data order not randomised can be found in Table 4. Here and in further tables, temp. is short for temperature. In this table, we can see the MSE, RMSE and R^2-score for both the train and the test set. The left-most column shows the different combinations of input features that have been used.

Input	**MSE train**	**MSE test**	**RMSE train**	**RMSE test**	**R^2 train**	**R^2 test**
Temp.	1.313	1.315	1.146	1.147	0.230	0.107
Wind	1.579	1.438	1.257	1.199	0.074	0.024
Rain	1.681	1.455	1.297	1.206	0.014	0.012
Temp., Wind	1.280	1.353	1.131	1.163	0.249	0.081
Temp., Rain	1.308	1.316	1.144	1.147	0.233	0.106
Wind, Rain	1.555	1.425	1.247	1.194	0.088	0.032
Temp., Wind, Rain	1.274	1.358	1.129	1.165	0.253	0.078

Table 4: Linear Regression Model results (data order not randomised)

If we look at Table 4, we can see that in most cases, the MSE, RMSE and R^2-score values are quite similar for both the train and test set. It is only not the case for the R^2-scores of the temperature/wind and temperature/wind/rain combination. There, the R^2-score is more than 3 times better on the train set than on the test set, while that is not the case for the other models. This might be due to overfitting on the train set in those two models.
Furthermore, we can conclude that the model performs best when temperature is taken into account, as those models have the lowest MSE and RMSE on the test set, and their R^2-score is higher. We can not say that a Linear Regression Model is very effective in this case, because the average

number of incidents in a day is 0.781, with a standard deviation of 1.289. All models have an RMSE of the test set that is lower than the standard deviation of the mean, but in most cases the differences are not very large. This means that these models perform better than estimating the mean, but not by much.

Input	**MSE train**	**MSE test**	**RMSE train**	**RMSE test**	**R^2 train**	**R^2 test**
Temp.	1.386	1.006	1.177	1.003	0.210	0.215
Wind	1.643	1.182	1.282	1.087	0.064	0.077
Rain	1.731	1.455	1.316	1.206	0.013	0.012
Temp., Wind	1.359	1.013	1.166	1.007	0.225	0.209
Temp., Rain	1.382	1.316	1.176	1.147	0.212	0.106
Wind Rain	1.619	1.425	1.272	1.194	0.077	0.032
Temp., Wind, Rain	1.355	1.358	1.164	1.165	0.228	0.078

Table 5: Linear Regression Model results (data order randomised)

Table 5 shows the results of the same model, but the only difference is that the data has been randomised. In this Table, we can see that the model seems to perform similarly or a little better for the test set than for the train set when looking at the MSE and RMSE. This is unexpected, but not entirely impossible. It is likely that the specific random seed creates a split where the model performs very well on the test set. As in Table 4, we can see that the R^2-score of the test set of the temperature/wind/rain model is a lot worse than the R^2-score of the train set, but that large difference is not present in the temperature/wind model as it is in Table 4. Like before, it is likely that the temperature/wind model is overfitted on that train set, causing a large difference in R^2-scores between the train and test set.
From Table 5 we can also conclude that the models where temperature is included perform best. The temperature, temperature/wind and temperature/rain models perform best with R^2-scores of 0.215, 0.209 and 0.106 respectively. These R^2-scores are not very good but better than the other Linear Regression models that are included.

7.2.2 XGBoost Regression Model

The XGBoost Regression Model tries to predict a continuous numerical value by optimising a loss function which is the RMSE in this case. Table 6 shows the results of the XGBoost Model. As with the Linear Regression Model, the MSE, RMSE and R^2-score is calculated for the train and test set for each model.
There are multiple things to note in the results in Table 6. First of all, there are some models where the MSE and/or RMSE values of the train set dip below 1, something we did not see for the Linear Regression models. This suggests that the XGBoost model performs better compared to the Linear Regression models when looking at the MSE/RMSE values. Furthermore, the model that takes the temperature and rain into account has MSE and RMSE values of the test set that are below

one. Additionally, most MSE and RMSE scores are quite similar between the train and the test set. When we look at the R^2-scores, we can see that each model performs better on the train set than on the test set. A likely explanation for this is that the models are all overfitted on the train set. If we compare the R^2-scores of the different models, we can see that the models that take temperature into account perform best. In this case, the temperature/rain model had the highest R^2-score on the test set, with the temperature, temperature/wind and temperature/wind/rain models following behind.

Input	**MSE train**	**MSE test**	**RMSE train**	**RMSE test**	**R^2 train**	**R^2 test**
Temp.	1.105	1.028	1.051	1.014	0.370	0.197
Wind	1.362	1.231	1.167	1.109	0.224	0.039
Rain	1.649	1.216	1.284	1.103	0.060	0.050
Temp., Wind	0.741	1.030	0.861	1.015	0.578	0.196
Temp., Rain	1.016	0.973	1.008	0.987	0.421	0.240
Wind Rain	1.195	1.10	1.093	1.077	0.319	0.094
Temp., Wind, Rain	0.968	1.045	0.984	1.022	0.448	0.184

Table 6: XGBoost Model results (data order randomised)

7.2.3 Regression Tree Model

The last Regression Model we have made is the Regression Tree Model. A Regression Tree Model predicts continuous numerical values using a tree structure to divide the data into smaller groups based on the provided features.

The results for the Regression Tree Model can be found in Table 7. Since we have used the same combinations of external circumstances and the same random seed as in Section 7.2.1 and Section 7.2.2, we can directly compare their results. Table 7 shows the MSE, RMSE and R^2-score for both the train and the test set, where we can see that the model performs better on the test set than on the train set for the MSE and RMSE values. However, the R^2-scores for the test set are not better in most cases.

As with the previous two Regression Models, we can see that the models that take the temperature into account perform best, with the temperature/rain model having the highest R^2-score test score and the temperature, temperature/wind/rain, and temperature/wind models following behind in that order.

Input	MSE train	MSE test	RMSE train	RMSE test	R^2 train	R^2 test
Temp.	1.159	1.003	1.076	1.002	0.339	0.217
Wind	1.539	1.216	1.240	1.103	0.123	0.051
Rain	1.686	1.209	1.298	1.099	0.039	0.056
Temp. Wind	1.155	1.012	1.074	1.006	0.342	0.210
Temp. Rain	1.144	0.992	1.070	0.996	0.348	0.225
Wind Rain	1.447	1.201	1.203	1.096	0.175	0.062
Temp. Wind Rain	1.138	1.010	1.067	1.005	0.351	0.211

Table 7: Regression Tree Model results (data order randomised)

To discover what features are most important when splitting, we can visualise the decision tree. The Regression Tree for the model that includes temperature, wind and rain can be found in Figure 8. The Regression Tree shows that both the root and the first level split on temperature, which is in line with our expectations. This means that dividing on temperature results in the best split in the data for the first two splits. On the second and third level we can see that there is no single feature that provides the best split. On those levels, the best feature to split on is dependent on the results of the previous splits. From Figure 8 we can conclude that temperature results in the best split for the data. This is likely due to temperature having the largest influence on the number of incidents that happen in a day.

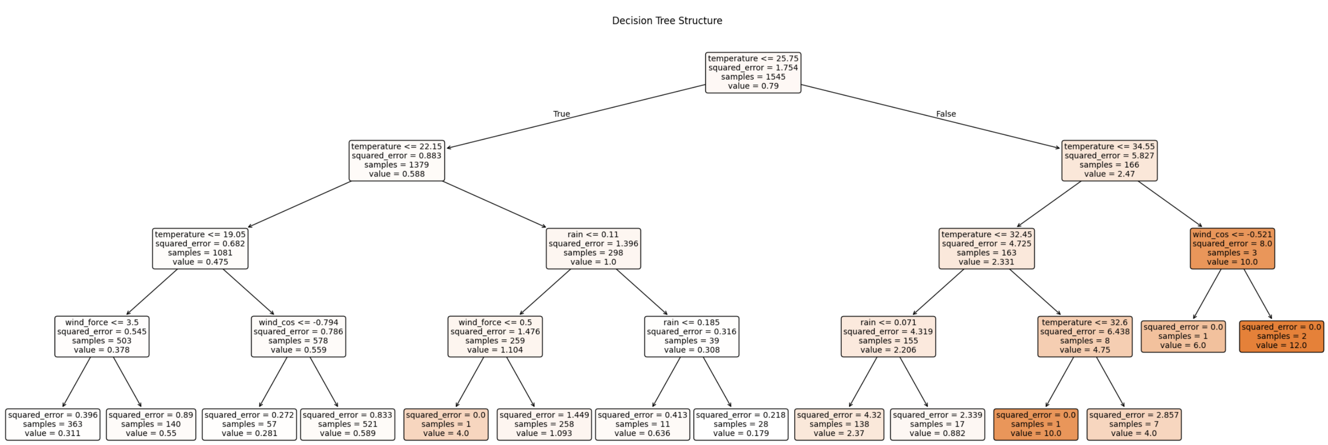


Figure 8: Regression Decision Tree, includes temperature, wind and rain

As we have now analysed each Regression Model, we can combine their results to compare them. Table 8 shows the R^2-scores on the test set for each Regression Model used, with the right-most column showing the average for that set of input features. The last row shows the averages per

Regression Model used. If we compare the averages for each model, we can see that the Regression Tree Model had the highest average R^2-score, followed by the XGBoost Model, with the Linear Regression Model having the lowest R^2-score. Therefore, we can conclude that for these experiments, the Regression Tree Model performed best. When we compare the averages per set of input features, we can see that the temperature model has the highest average R^2-score, followed by the temperature/wind model and the temperature/rain model. The models that do not take temperature into account perform worse than those that do.

Input	Linear Regression	XGBoost	Regression Tree	***Average***
Temp.	<u>0.215</u>	0.197	0.217	*0.210*
Wind	0.077	0.039	0.051	*0.056*
Rain	0.012	0.050	0.056	*0.039*
Temp. **Wind**	0.209	0.196	0.210	*0.205*
Temp. **Rain**	0.106	<u>0.240</u>	<u>0.225</u>	*0.190*
Wind **Rain**	0.032	0.094	0.062	*0.063*
Temp. **Wind** **Rain**	0.078	0.184	0.211	*0.158*
Average	*0.104*	*0.143*	*0.147*	

Table 8: Regression Models R^2-scores on the test set per type of model and combination of input features (data order randomised)

RQ3 asked how the temperature, wind and rain affect the effectiveness of a predictive model that predicts the number of incidents in a day. Although the R^2-scores in Table 8 are not very good in general, from these results we can conclude that including temperature in the set of input features likely causes the models to perform better than when it is not included. These results suggest that including air temperature as an input feature in any future beach incident prediction models would be beneficial.

Another way of comparing all three Regression Models, is by comparing their predictions on a smaller scale. Figure 9 shows the predictions of each Regression Model and the actual number of incidents per week. For this experiment, all incidents from 2005 to 2023 are included in the train set, with all incidents in 2024 and 2025 being in the test set. In contrast to earlier experiments, we take a smaller part of the data set as the test set, so it is easier to compare the results. There are several things to note in Figure 9. For example, week 15/07/24 and week 05/08/24 have very high numbers of incidents in a week, and in both cases, the XGBoost Model was able to come close to the actual values, with the Regression Tree Model also coming close to the actual value in week 05/08/24. Furthermore, there are multiple weeks where all three models predict a much higher number of incidents than there actually were, such as week 24/06/24 and 25/08/25. This is likely due to the circumstances being in line with other weeks where there were a lot of incidents. Possible explanations for the low actual value of incidents could be more or better supervision by Lifeguard Katwijk, or it not being a holiday meaning people have less time to visit the beach (for

example week 24/06/24).
Figure 9 gives us the ability to analyse specific predictions and compare results from the different Regression Models we have used. These insights can help when improving incident prediction models, which can lead to better performing models.

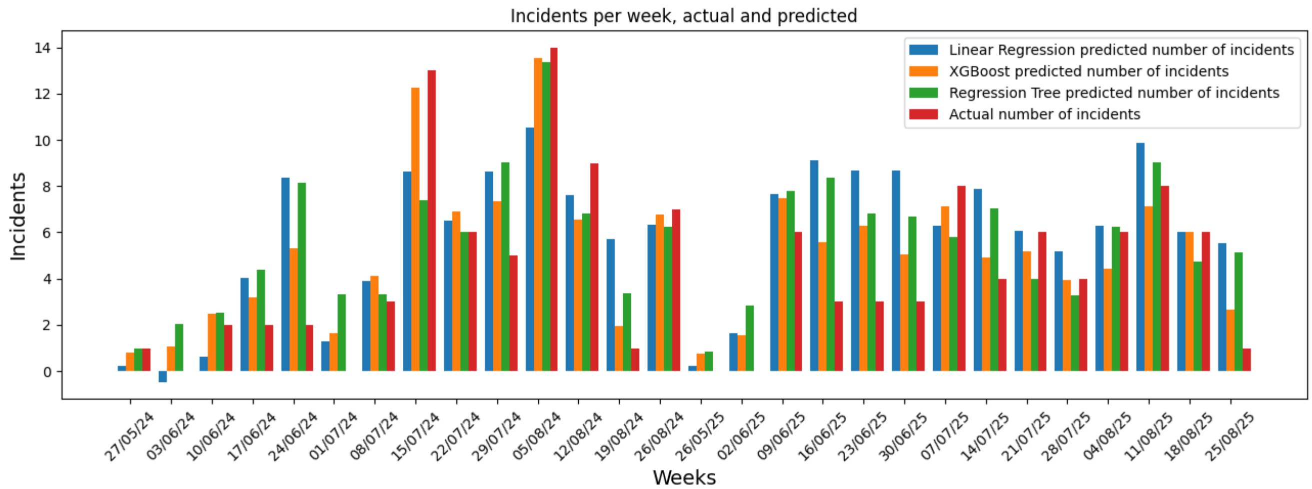


Figure 9: Predicted and Actual number of incidents per week

7.2.4 Decision Tree

A Decision Tree is not a Regression Model, but rather a Classification Model. It is a model that predicts the class rather than a continuous numerical value. Therefore, we use it to create a model that will use the external circumstances to predict the type of incident. For this model we build two models, both with the same external factors as input, but with a different goal. The first one will try to predict if an incident was categorised as a first-aid incident or not. The second one will try to predict if an incident was located in sea or not. In both cases we optimised the model on gini and on entropy, but since they both resulted in the exact same results for both classifiers, we have combined their results per classifier.
For predicting if an incident is a first aid incident or not, we would expect the temperature to have a large effect on the splits. First aid incidents are quite common, with 853 out of 1869 (45.6%) incidents being a first aid incident. When the temperature is higher, there are more people present on the beach, which means there is a higher likelihood of first aid incidents happening.

Class	Precision	Recall	F1-score	Support
Not first aid	0.626	0.629	0.627	202
First aid	0.561	0.558	0.560	172

Table 9: Decision Tree results for classifying as first aid or not results (gini & entropy)

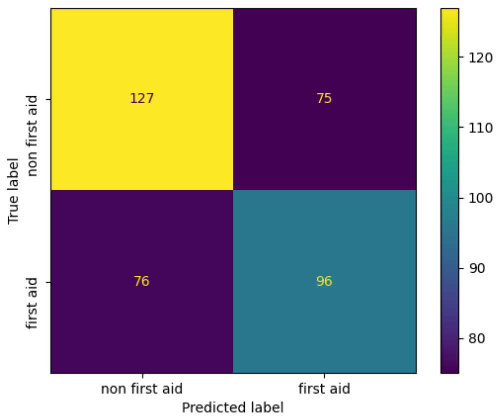


Figure 10: Confusion Matrix for Decision Tree incidents classified as first aid or not

Table 9 shows the results of the Decision Tree that both the model trained on entropy and the model trained on gini produced. We can see that the model is more effective when categorising incidents as not a first aid incident, as precision, recall and f1-score are all closer to one when compared to the first aid category. Additionally, the accuracy for this model is 0.596, meaning 59.6% of instances were classified correctly. Furthermore, we have calculated the AUC, which is 0.607 for the model trained on gini and 0.612 for the model trained on entropy. With a threshold of 0.5 (guessing the category based on class distribution), both of these models perform better than guessing the category.

Additionally, we can see the Confusion Matrix of both the model trained on gini and the one trained on entropy in Figure 10. We can see that although the models are quite accurate in predicting the right category, there are also a lot of misclassified incidents in there.

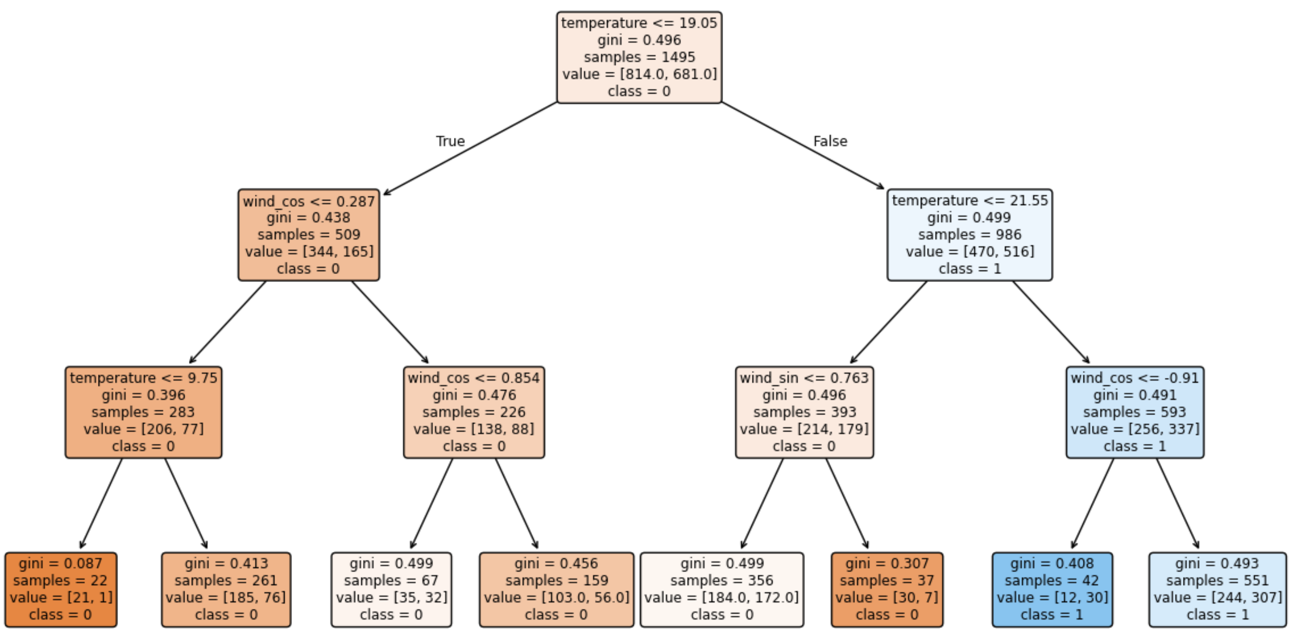


Figure 11: Decision Tree for classifying on first aid or not, trained on gini

Figure 11 shows the Decision Tree for the model trained on gini. In the Appendix (Section 8), Figure 17 shows the Decision Tree for the model trained on entropy. Although both the Decision Trees trained on gini and on entropy produce the exact same results in Table 9, there were different routes to those answers. When analysing the final leaves of both trees, we can see that only the two right-most leaves have assigned class 1 (first aid incident). As the maximum depth of the tree was set to 3, the best splits in most cases still resulted in a larger portion of incidents in each leaf being non-first aid incidents. Therefore, there are only two relevant splits that result in different majority classes, namely the first one (temperature $\leq$ 19.05) and the second split on the right (temperature $\leq$ 21.55). Since a temperature of 19.05 or lower is included in a temperature of 21.55 or lower, we can conclude that for classifying an instant as first aid or not, the temperature at the time is the only relevant input for this model. Any incident with a temperature of 21.55 or lower is classified as a non-first aid incident, and any incident with a temperature higher than 21.55 classified as a first aid incident. From these Decision Trees, we can conclude that the temperature provides the best split when classifying between an instance being a first aid incident or not.

Class	Precision	Recall	F1-score	Support
Not in sea	0.655	0.833	0.734	228
In sea	0.548	0.315	0.400	146

Table 10: Decision Tree results for classifying if incident location is in sea or not (gini & entropy)

For the second Decision Tree, we have created a model that categorises incidents based on whether it is a sea notification or not. There are 766 out of 1869 (41.0%) incidents that are considered sea incidents. In this case, we expect the wind speed and the wind direction to play a larger part in deciding on the split than the temperature.

Table 10 shows the Decision Tree results for the Decision Tree that classifies incidents as a sea incident or not. From these results we can see that the recall is quite high for the category not in sea. This means that little incidents that did not take place in sea, are miscategorised. On the other hand, the recall is quite low for the category incidents in sea, which means that there are a lot of incidents categorised as not taking place in sea, when they actually did take place in sea. This is also reflected in Figure 12, which shows the Confusion Matrix for both models. The accuracy for both models is 63.1%. Furthermore, the AUC value for both the model trained on gini and the model trained on entropy is 0.590. With a threshold of 0.5, these models perform better than just guessing, but the difference is not large. This suggests that there are possibilities for improving.

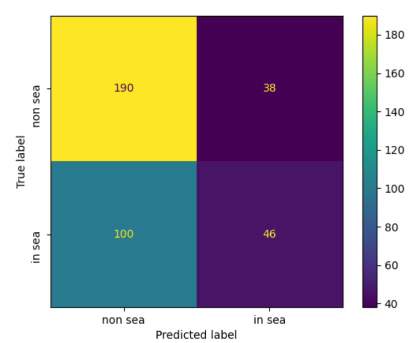


Figure 12: Confusion Matrix for Decision Tree incidents located in sea or not

As can be seen from Table 10, the category for incidents not located in sea is most accurate. Furthermore, we can see that there are a lot of incidents that took place in sea, but were not classified as such. This Decision Tree does not seem as effective in classifying incidents on whether they took place in sea or not.

Figure 13 shows the Decision Tree for the model trained on gini, whereas Figure 18 (Appendix, Section 8) shows the Decision Tree for the model trained on entropy. We can clearly see that they split on the same features each time, with the same values. This also explains the equal values in Table 10. An interesting thing to note here is that both Decision Trees start by splitting on

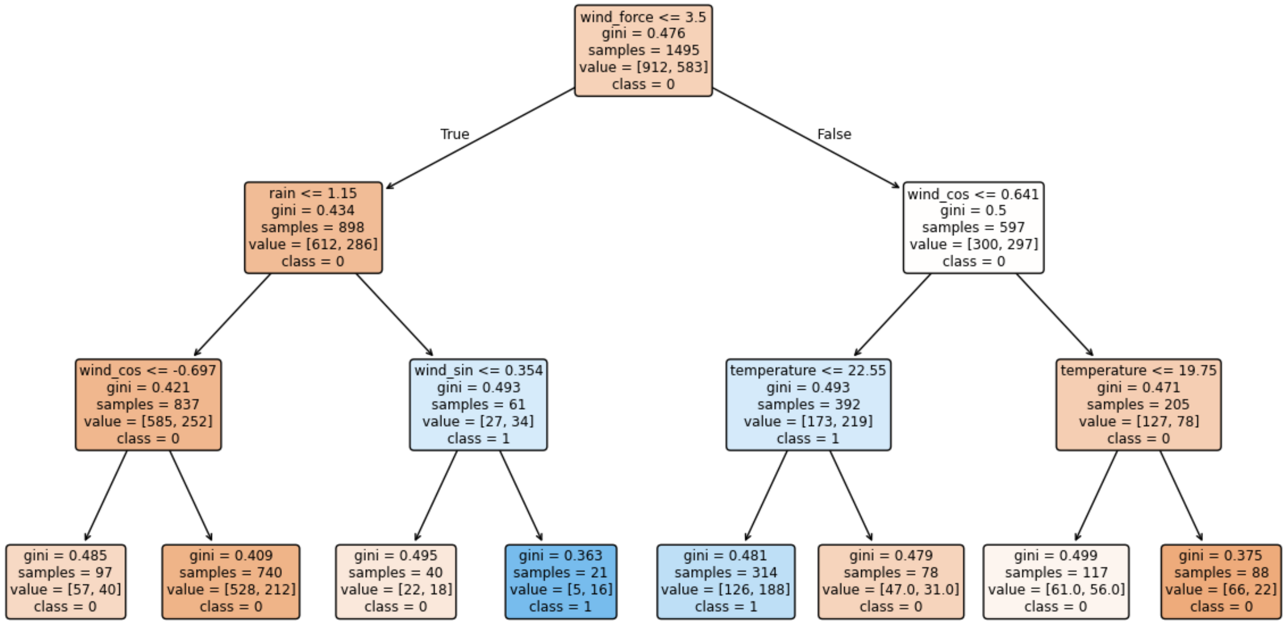


Figure 13: Decision Tree for incident located in sea or not, trained on gini

wind speed, and after that on rain and part of the wind direction. Only on the third level do both Decision Trees start splitting on temperature. Thus, we can conclude that for deciding if an incident took place in sea or not, the wind and rain provide the best split.

RQ4 asks whether we can build a predictive model that can predict the category of an incident accurately. From the results we have gathered we can see that it is possible to build such a model, but the effectiveness varies based on the type of incident it tries to predict. This can also be due to the distribution of the types of incidents in the dataset. There are a lot of first aid incidents, and also a lot of incidents that take place in sea. We can speculate that trying to build a Decision Tree that categorises incidents that appear not as much as others will likely give a less effective model.

8 Conclusions and Further Research

This thesis focused on finding patterns and learning from Lifeguard Katwijk's incident reports. We wanted to find out if there are relations present in the data between external circumstances and aspects of the incident data. Furthermore, we wanted to investigate if we can build a predictive model that can predict different aspects of the incident reports.

First of all, from the results of the research we can conclude that the incident reports contain relations between external cirucmstances and aspects of the incidents. There is a clear connection between the maximum temperature in a day and the average number of incidents that happens. Furthermore, incidents that happen in sea are more likely to happen at high tide and with a higher wind speed. The results do not seem to show a clear correlation between a location and the number of incidents that have happened there.

Additionally, the results from Section 7.2 show that we can use Regression Models such as the Linear Regression, XGBoost or Regression Tree Model to estimate the number of incidents that happen in a day based on external circumstances. The R^2-scores from each model and each combination of input features showed that models that included temperature tended to work better, although the best performing model still only had an R^2-score of 0.225, which is not considered to be very good. A Decision Tree can be used to approximate the type of incident, although it was not as effective as expected. It does show which external circumstances resulted in the best split in the data set for approximating the type of incident, which could be useful for future improvements.

In short, from the results we can conclude that there are multiple patterns to be found and that we can use the incident report data to build a predictive model. However, there are some improvements possible. To expand the incident report data and to be able to dive deeper into possible connections, it would be interesting to add more relevant data. Possible additions could be the wave height, the speed of the current, the number of lifeguards present at the station or whether there was a patrol in the area of the incident when it happened. These aspect might give new insights into the data and aid in taking more preventive measures.

In addition to expanding the dataset, there are other interesting options for further research. By combining the predictive models and creating a user interface, it could be added to the Lifeguard Watch Report. That way, it can help the lifeguards in preparing for their day. Furthermore, there are some aspects of the incident report form that could benefit from an update to make each future incident report more expansive. Lastly, it would be interesting to ask other Lifeguard stations in the Netherlands if they have a way of storing information about their incidents, as it might be interesting to combine them to get a more robust dataset.

In conclusion, this Bachelor Thesis about learning from Lifeguard Katwijk's incident reports and finding patterns provided us with interesting insights into the data and paved the way for further improvements in safety along Katwijk's beach.

References

[dKCT21] E. de Korte, B. Castelle, and E. Tellier. A bayesian network approach to modelling rip-current drownings and shore-break wave injuries. *Natural Hazards and Earth System Sciences*, 21(7):2075–2091, 2021.

[Fla12] Peter Flach. *Machine Learning, The Art and Science of Algorithms that Make Sense of Data*, chapter 5. Cambridge University Press, 2012.

[KNMa] KNMI. Beaufortscaal. https://cdn.knmi.nl/system/ckeditor_assets/attachments/60/beaufort.pdf.

[KNMb] KNMI. Daggegevens van weerstations. https://daggegevens.knmi.nl/klimatologie/daggegevens. [Online; accessed 11.11.2025].

[KNMc] KNMI. Uurwaarden van weerstations. https://daggegevens.knmi.nl/klimatologie/uurgegevens. [Online; accessed 14.10.2025].

[LTZ+24] Yuan Li, Jialin Tang, Chi Zhang, Qinyi Li, Shanhang Chi, Yao Zhang, Hongshuai Qi, and Chuang Zhang. Analysis and machine-learning-based prediction of beach accidents on a recreational beach in china. *Anthropocene Coasts*, 7(29), 2024.

[MPV21] Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining. *Introduction to Linear Regression Analysis*, chapter 2. Wiley, sixth edition, 2021.

[PVG+11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[Rij] Rijkswaterstaat. Rijkswaterstaat waterinfo. https://waterinfo.rws.nl/thema/Waterbeheer. [Online; accessed 14.10.2025].

[SSBD14] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*, chapter 18.2. Cambridge University Press, 2014.

[TSGJ+22] Éric Tellier, Bruno Simonnet, Cédric Gil-Jardiné, Marion Lerouge-Bailhache, Bruno Castelle, and Rachid Salmi. Predicting drowning from sea and weather forecasts: development and validation of a model on surf beaches of southwestern france. *Injury Prevention*, 28(1):16–22, 2022.

[Wad20] Corey Wade. *Hands-On Gradient Boosting with XGBoost and scikit-learn*. Packt Publishing Limited, 2020.

Appendix

Notification value	Notification category
0	First aid
1	Missing on the beach/in sea
2	Missing on the beach
3	Missing in sea
4	Missing (other)
5	Bather needing help
6	Swimmer needing help
7	People on supboard/floating object needing help
8	People on dinghy/rowing boat needing help
9	People on motorboat needing help
10	People on sailing boat needing help
11	Wing-/windsurfer needing help
12	Wave surfer needing help
13	Kiter needing help
14	People on catamaran needing help
15	(Possible) suicide attempt
16	Assisting ambulance
17	Assisting police
18	Assisting KNRM
19	Assisting EHBZ
20	Vehicle needing assistance
21	Fire
22	Child found (parents missing)
23	Assisting with an event
24	Drunk person causes disturbance
25	Dangerous situation (dangerous liquid/object)
26	Distress signal spotted
27	Animal needing help
28	Confused person
29	Unattended items found
30	Floating object spotted in sea
31	Other

Table 11: Notification categories and numerical value assigned

Location value	Location
0	Other
1	North of Airtime
2	Airtime
3	Het Wantveld
4	Willy Noord
5	Uitwatering
6	Noordpost
7	De Watering
8	Paal 14
9	Surf & Beach
10	Key West
11	t́ Centrum
12	Zee en Zon
13	Zomers
14	Het Strand
15	Zand
16	Zilt
17	KBS
18	Westpunt
19	SandCBar
20	Zuidpost
21	Sisters Beach
22	Willy Zuid
23	Skuytevaert
24	South of Skuytevaert

Table 12: Current locations and numerical value assigned

Result value	Result category
0	Other
1	First Aid
2	Send patient to doctor
3	Ambulance called
4	Arranged Transport
5	Assisted by other emergency services
6	Dangerous situation/object under control
7	Assisted vessel (sea)
8	Assisted vehicle
9	Assisted animal
10	Those involved brought to safety
11	Missing person found
12	Reunited with family/friends
13	Materials reunited with owner
14	False alarm
15	SAR stopped

Table 13: Result categories and numerical value assigned

First aid value	First aid category
0	Not applicable
1	Check condition
2	Wound treatment
3	Sprained limb treated
4	Cooled affected area
5	Raised (body) temperature
6	Arranged transport
7	Called ambulance
8	Send to doctor
9	Resuscitation

Table 14: First aid categories and numerical value assigned

incident_p	datum	post_fk	actie_start	actie_eind	actie_locatie	afstandinz	windrichtir	windkrach	melding	actie	ehbo	resultaat
672	8-7-2013	2	14:00:00	14:21:00	KRB noordpo	300	NE	3	Zeilboot in	Hulp met 1 SR8(en)		Doorgestuurd naar ...Skuytevaert
673	8-7-2013	2	18:00:00	18:18:00	Noordpost	150	ENE	4	(Kite)surfer in	Hulp met 1 SR8(en)	Controle toestand	Betrokkene(n) op de kant
674	13-7-2013	2	17:25:00	17:53:00	Zomer	-1	N	3	Vermissing aan de water	Hulp met autoHulp met	n.v.t.	Vermiste(n) gevonden, werd aangetroffen bij Paal 14
675	15-7-2013	2	15:20:00	15:26:00	Noordpost	-1	NW	3	Persoon heeft skimboar	Op de post	Wondbehandeling	Doorgestuurd naar dokterspost
676	15-7-2013	1	20:59:00	21:25:00	Benny\'s Bea	0	NW	0	Knieschijf verdraaid	Hulp met 1 LP(s) en met k	Controle	Doorgestuurd naar LUMC
677	15-7-2013	1	21:30:00	21:35:00	Het strand	0	NNW	2	Vermissing op strand	Hulp met auto	N.V.T.	Loos alarm
678	17-7-2013	2	19:01:00	19:18:00	Willy Noord	200	N	3	Groep zwemmers in prol	Hulp met autoHulp met	N.v.t.	Loos alarm
679	18-7-2013	1	23:54:00	01:30:00	Onbekend	-1	NNE	2	Mogelijke zelfdoding	Hulp met 2 SR8(en)Hulp	Geen	Na overleg met Meldkamer en kustwacht, einde actie zonder resultaat
680	18-7-2013	1	16:45:00	17:05:00	Paal 14	-1	N	4	10-jarige surfer vermist i	Hulp met 2 SR8, 1 LP		Betrokkene op de kant

Figure 14: Incident reports before data preprocessing

incident_i	date	post	incident_s	incident_e	incident_locat	distance_i	wind_f	wind_dire	temperat	rain	tide	tide_(	notification	result	first_aid	aid_provid	Jeep	ATV	Boat_40h	Boat_60h	RWC	WP	KNRM	Ambulanc	Police
672	8-7-2013	2	14:00:00	14:21:00	Noordpost	0	3	45.0	24.0	0.0	C	C	people on sailing boat	Assisted vehicle	Not applicable	1	0	0	0	1	0	0	0	0	0
673	8-7-2013	2	18:00:00	18:18:00	Noordpost	0	4	90.0	21.0	0.0	H	N	kiter needing help	Those involved brought to	Check condition	1	0	0	0	1	0	0	0	0	0
674	13-7-2013	2	17:25:00	17:53:00	Zomers	0	3	360.0	17.7	0.0	H	N	missing on the beach	Missing person found	Not applicable	1	1	0	0	1	0	1	0	0	0
675	15-7-2013	2	15:20:00	15:26:00	Noordpost	0	3	305.0	20.7	0.0	L	S	first aid	Send patient to doctor	Wound treatment	1	0	0	0	0	0	0	0	0	0
676	15-7-2013	1	20:59:00	21:25:00	Sisters Beach	0-50	0	305.0	16.6	0.0	H	N	first aid	Send patient to doctor	Check condition	1	0	0	0	0	0	1	0	0	0
677	15-7-2013	1	21:30:00	21:35:00	Het Strand	0-50	2	360.0	15.6	0.0	H	N	missing on the beach	False alarm	Not applicable	1	1	0	0	0	0	0	0	0	0
678	17-7-2013	2	19:01:00	19:18:00	Willy Noord	0	3	360.0	21.4	0.0	L	S	swimmer needing help	False alarm	Not applicable	1	1	0	0	1	0	0	0	0	0
679	18-7-2013	1	23:54:00	01:30:00	Other	0	2	360.0	16.7	0.0	H	N	(possible) suicide atte	SAR stopped	Not applicable	1	1	0	0	1	0	1	0	0	0
680	18-7-2013	1	16:45:00	17:05:00	Paal 14	0	4	360.0	23.5	0.0	L	S	missing in sea	Those involved brought to	Not applicable	1	0	0	0	1	0	1	0	0	0

Figure 15: Incident reports after data preprocessing

incident_i	date	post	incident_s	incident_e	incident_l	distance_i	wind_f	wind_sin	wind_cos	temperatu	rain	tide	tide_d	notificati	result	first_aid	aid_provic	Jeep	ATV	Boat_40hp	Boat_60hp	RWC	WP	KNRM	Ambulance	Police
672	8-7-2013	2	14:00:00	14:21:00	6	300	3	0.707	0.707	24.0	0.0	C	C	10	8	0	1	0	0	0	1	0	0	0	0	0
673	8-7-2013	2	18:00:00	18:18:00	6	150	4	1.0	0.0	21.0	0.0	H	N	13	10	1	1	0	0	0	1	0	0	0	0	0
674	13-7-2013	2	17:25:00	17:53:00	13	-1	3	-0.0	1.0	17.7	0.0	H	N	2	11	0	1	1	0	0	1	0	1	0	0	0
675	15-7-2013	2	15:20:00	15:26:00	6	-1	3	-0.819	0.574	20.7	0.0	L	S	0	2	2	1	0	0	0	0	0	0	0	0	0
676	15-7-2013	1	20:59:00	21:25:00	21	0	0	-0.819	0.574	16.6	0.0	H	N	0	2	1	1	0	0	0	0	0	1	0	0	0
677	15-7-2013	1	21:30:00	21:35:00	14	0	2	-0.0	1.0	15.6	0.0	H	N	2	14	0	1	1	0	0	0	0	0	0	0	0
678	17-7-2013	2	19:01:00	19:18:00	4	200	3	-0.0	1.0	21.4	0.0	L	S	6	14	0	1	1	0	0	1	0	0	0	0	0
679	18-7-2013	1	23:54:00	01:30:00	0	-1	2	-0.0	1.0	16.7	0.0	H	N	15	15	0	1	1	0	0	1	0	1	0	0	0
680	18-7-2013	1	16:45:00	17:05:00	8	-1	4	-0.0	1.0	23.5	0.0	L	S	3	10	0	1	0	0	0	1	0	1	0	0	0

Figure 16: Incident reports after data preprocessing, file with numerical values

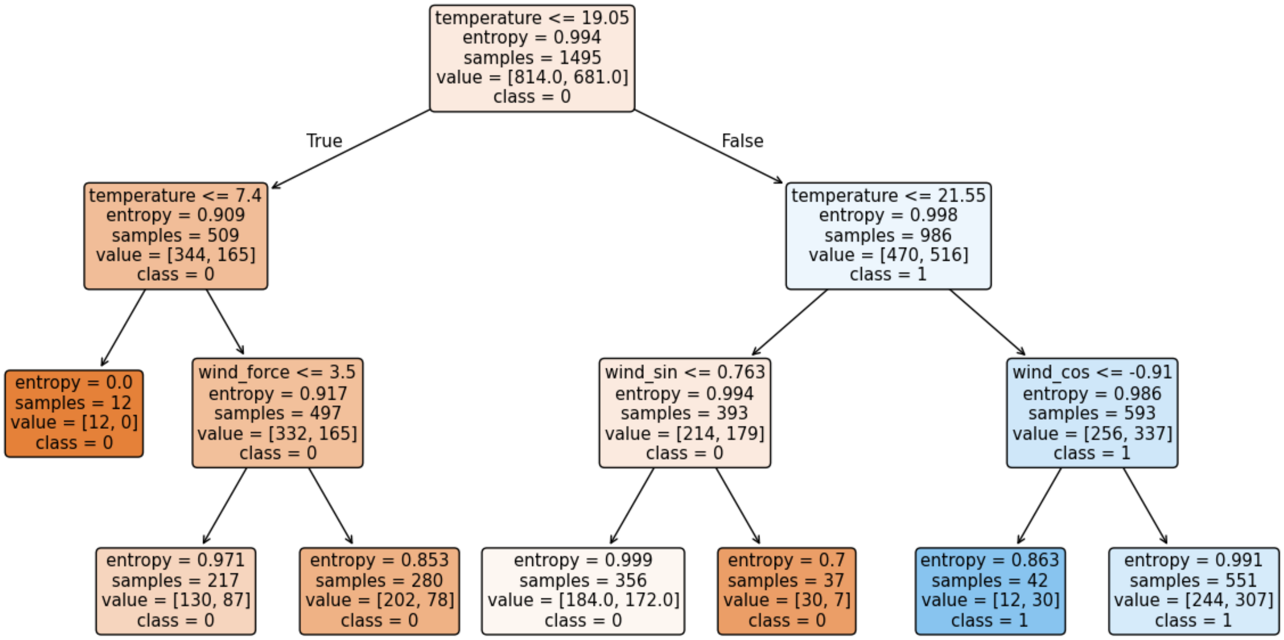


Figure 17: Decision Tree for classifying as first aid or not, trained on entropy

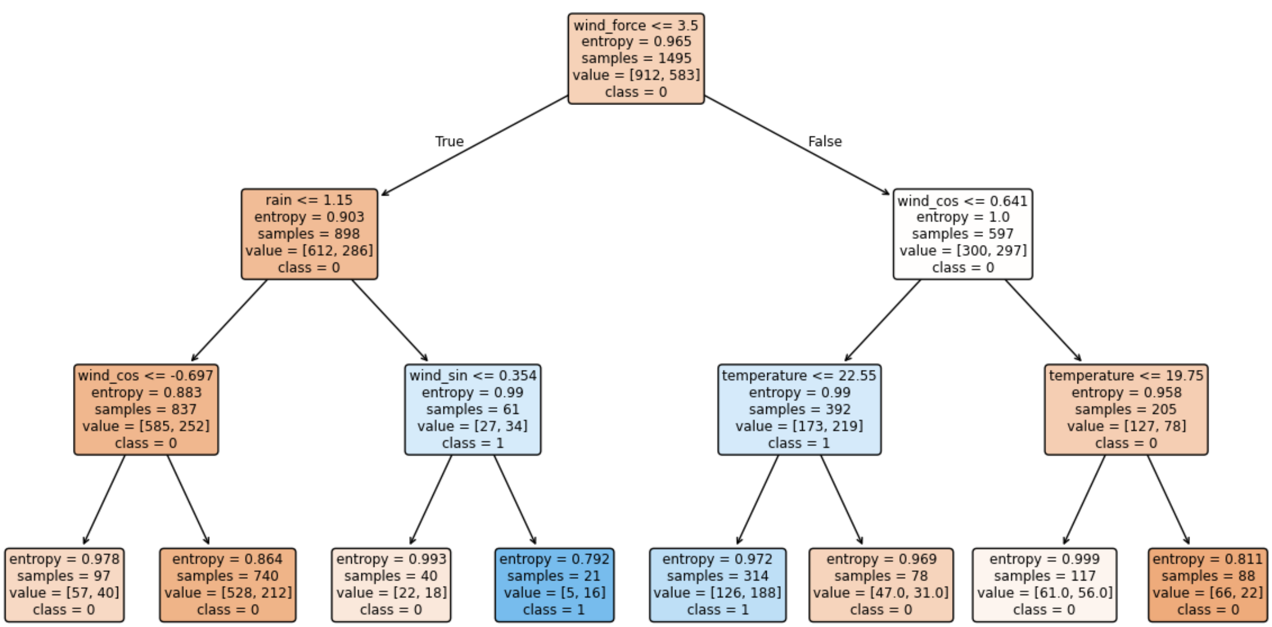


Figure 18: Decision Tree for incident location in sea or not, trained on entropy