PREDICTING
THE USE OF
FIRE DEPARTMENT MATERIALS
AT AN INCIDENT

A NUMBER OF TECHNIQUES PRESENTED TO ANALYZE AND PREDICT INCIDENTS

WRITTEN BY

OMESH DEBIPERSAD

Vrije Universiteit
Brandweer Amsterdam-Amstelland
Research paper MSc. Business Analytics
August 2017

Omesh Debipersad:
Predicting the use of fire brigade materials at an incident

Supervisors:
Prof. R.D. van der Mei (VU)
G. Legemaate, MSc. (company)

Location:
Vrije Universiteit Amsterdam
Faculty of Sciences
Business Analytics
De Boelelaan 1081a
1081 HV Amsterdam

Research company:
Brandweer Amsterdam-Amstelland
Karspeldreef 16
1090 AD Amsterdam

Preface

This research is conducted as a part of the curriculum of the Master Business Analytics. The research paper consists of 6 ECTS, which is the number of credits awarded for a month full time of research. The purpose of this report is to develop research skills, writing skills and presentation skills.

Professor R.D. van der Mei introduced me to Guido Legemaate. Guido Legemaate is a Phd student of Professor van der Mei and also one of the few data specialists at the fire department nationwide. In cooperation with Rob and Guido the following research question emerged:

Can we make a good forecast of the number of trucks needed at each incident? Furthermore, can we do this using the combination of incident data, weather data and demographic data?

The question was very interesting to research and fitted perfectly to my interests. To conduct this research I was able to work with Guido Legemaate at the head office of the fire department Amsterdam-Amstelland. Working together with Guido was very pleasant and I gained quite some knowledge. Furthermore, I did not have to spend much time to preprocess the data, because the provided data was already clean.

Finally, I would like to thank Professor van der Mei for his time. The sometimes eye-opening questions helped to improve my research. Talking to him is also a very pleasant experience.

Abstract

Goal & approach
The goal is to predict the most important truck types needed at an incident given incident, weather and demographic data. The truck types we tried to predict are water tender (TS), safety vehicle (RV), assistance vehicle (HV), officer at service (OVD) and water incident vehicle (WO). We made a forecast for incidents where at most two trucks were needed of each type, since by the definition of the fire department an incident is called large when an incident gets three or more trucks. It has been proven that large incidents occur randomly and do not have predictive power.

Methods
Data analysis has shown that data of new year's eve have to be predicted separately, because when we look at the number of incidents we see that for each year there is a peak in number of incidents. The incidents with two or less trucks per type will then be modeled using three algorithms. These algorithms are random forest, gradient boosting and artificial neural network. The two former ones are machine learning algorithms based on decision trees and the latter one is a deep learning algorithm.

Results
The first fitting of the data with random forest and gradient boosting gave some interesting insights about the predictability of the different truck types. HV, OVD and WO cannot predict more than one truck. This is because most of the times when more than one truck is needed, a large incident occurred. Therefore, we only focused on TS and RV trucks in the remainder of the result gathering. In general the following features were most important: **type of incident, location specific information and function of building.**

To improve the model we did feature selection. We took the ten and five most important features for TS and RV truck types. We saw that in general both the fit and mean squared error improved, except for random forest when predicting RV trucks. The fit and mean squared error were worse when applying feature selection. We suspect that overfitting sneaked into the fitting of the algorithm. Gradient boosting performed best with the five most important features resulting in a mean squared error of 0.0378 and R^2 of 0.5526 for TS trucks and a mean squared error of 0.0387 and R^2 of -0.2036 for RV trucks. The mean squared error seems positive for RV trucks, but the fit is poor making it a bad prediction.
Finally we added combinations of weather variables to the data, but that made the predictions for the number of TS and RV trucks worse.

Recommendations
For further research, it would be useful to find the relationship between the important features and truck usage per incident. We have identified the important features, but not the explicit relationships. Furthermore, it could be interesting to further explore neural networks as we think that it could improve our findings. Also one could try to mathematically prove randomness of the occurrence of large incidents according to the definition of the fire department. Finally some data

about house specific information could be interesting to improve the performance of the models.

Contents

1 Introduction

1.1 Fire department Amsterdam-Amstelland

Fire department Amsterdam-Amstelland exists of six municipalities, namely Aalsmeer, Amstelveen, Amsterdam, Diemen, Ouder-Amstel en Uithoorn. The region has about one million inhabitants and about two million tourists coming to the region on a yearly basis. There are about 1100 people working for this fire department. Not only do they prevent regular incidents, such as fires, but they also intervene in large crisissituations. The fire department has nineteen stations which consist of at least one water tender. Furthermore, there are many other firefighting vehicles available, such as vehicles with longer ladders or boats. The stations are staffed with people for 24 hours, so that any incident could be approached at any time.

1.2 Problem statement

The fire department has about 13000 incidents per year. That is about 35 in its region per day. This is quite a large number of incidents. Therefore is it necessary to gain insight in these incidents. As a result the following research question has emerged:
Can we make a good forecast of the number of trucks per type needed at each incident? Furthermore, can we do this using the combination of incident data, weather data and demographic data?
This is what we will try to research and come up with answer for. It is for example very useful to know where the most incidents are occurring, so that people or materials could be relocated. Maybe even a new station is needed. What we will try to do is using the combination of incident, weather and demographic data analyze what factors are important in the occurrence of an incident. Using this information the number of trucks per type will be predicted.

1.3 Previous research

Ab Boersema and Jeffrey de Deijn have done similar researches as this one, only using different techniques and different data for the fire department.
Ab Boersema (2013) has researched whether it was possible to predict the number of incidents. He used a multiplicative model to predict the number of incidents per neighborhood. His conclusions were that the predictions were better when looked at the data as a whole, rather than per neighborhood. This is because some neighborhoods have less incidents.
Jeffrey de Deijn (2016) researched whether is was possible to predict the number of trucks needed at each station and the number of incidents that each station has to handle. His conclusions were that the generalized linear model was better at predicting based on his performance measures, however random forest predicted busier days the best. To get the best of both worlds he combined both models.

1.4 Structure of the report

The report will begin with some exploratory data analysis in section 2. In this section we will try to find some interesting observations and more importantly get some insight in the data. This will be followed by the modeling of the data using the algorithms in section 3. Using some insight from the data analysis and previous research we will model the data. Also some theoretical background about the algorithms will be given. Next, the results of the models will be presented in section 4. This part of the paper will provide insights in how well the model performed using some performance metrics and visual support. Finally, some conclusions and recommendations for further research will be given in section 5 and 6. From a business point of view the most interesting sections would be 2 and 4, since the insights from the data will be provided in section 2 and the results will be shown in section 4. Finally, the research was in a combination of Postgresql and Python.

2 The data

We started with data from three sources:

- The fire department provided us with incident and data of material sent out. This data ranges from 1-1-2008 to 1-5-2017;
- Open source weather data from the KNMI. This data ranges from 1-1-2008 to 1-5-2017;
- Demographic data from CBS on neighborhood and area level. This data is from 9-9-2013.

For the purpose of a more detailed research we added data about coordinates per neighborhood and coordinates per fire station.

The provided incident data consist of on scene data and characteristics per incident, such as street name, postal code, type of incident, date, time, function of the building (except for an outdoor fire) and many more.
The data of the material sent provides us the data needed to specify the number of trucks per incident and it consists of truck types, time between a call and leaving, date, time, called station, order of arriving etc.
The KNMI data is detailed weather data specified on a daily level. It gives us data about the temperature, windspeed, sunshine on a day, rain, and airpressure. The data was measured at the Schiphol KNMI station. The reason why we chose to use the data from this station is because is it the closest station to the operating area of the fire department, so this data will give us the most reliable weather conditions to analyze.
Finally, the area and neighborhood data from CBS both have the same characteristics. The neighborhood data is only more detailed, since the areas are smaller. The data consist of number of people living, the area, the number of males, the number of females, the number of non western people, Surinamese people, Moroccan people, Antillean people, married people, widowed people, divorced people etc.

2.1 Exploratory data analysis

In this section we are exploring what kind of data we are dealing with. Different kind of plots, tables and statistics will be shown, in order to gain interesting insights.

2.1.1 The incidents

Figure 1 shows the incident distribution over each incident type. It is interesting that OMS incidents are most frequently occurring. OMS incidents are incidents in which notifications automatically generated. When such an incident happens a water tender has to go to the incident. These are generated even when the smallest amount of smoke is being developed, so that is why these incidents are occurring so frequently.

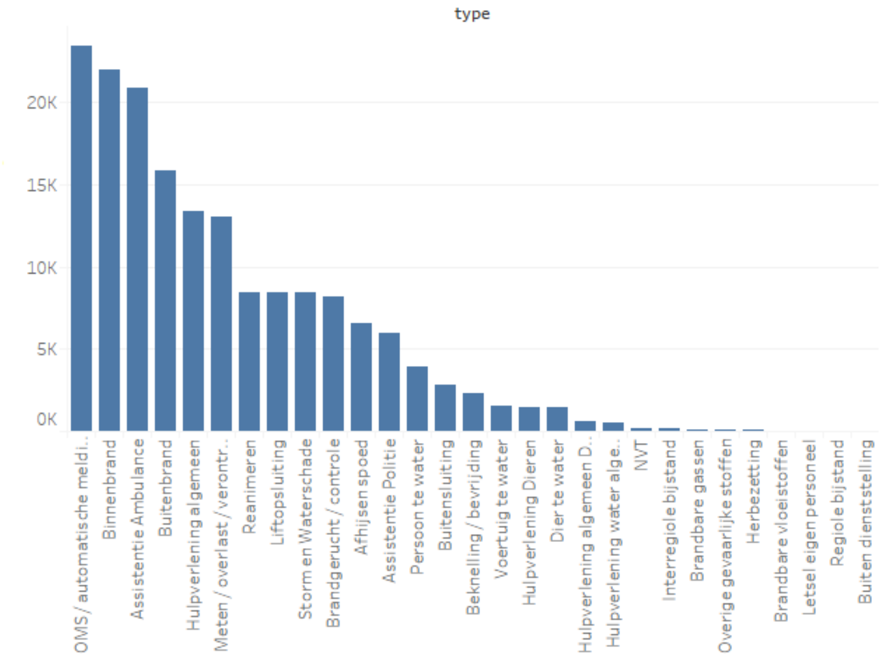

Figure 1: Number of incidents at for each incident type

Figure 2 shows the number of trucks used at each incident type per truck type. In Figure 1 we saw that OMS incidents are most frequently occurring, however when figure 2 is observed we see that indoor fires needed most of the material. Another interesting observation is that TS, which is a water tender, is most frequently send to an incident. This is because a water tender is the first truck to be sent to an incident.

When we look at the TS trucks only we see that these type of trucks were most used at OMS incidents. The reason behind this lies in figure 1, because OMS incidents are occurred the most. Many times these OMS incidents are not significant, but still a TS was sent. At both an inside and outside fire the amount of safety vehicles with higher ladders (RV) are also used quite a few times.

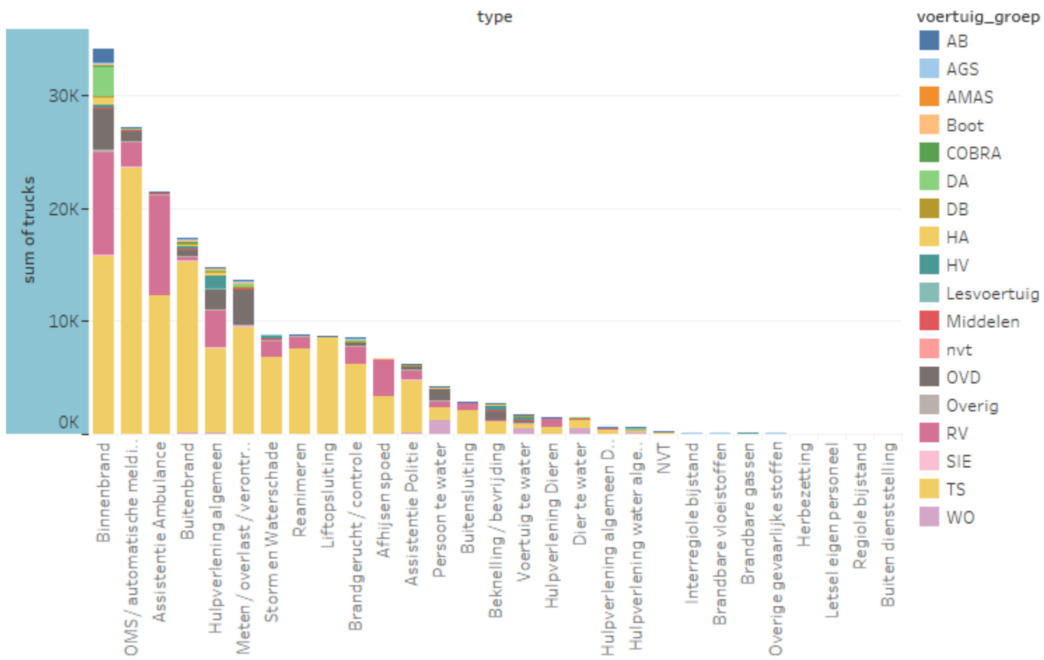

Figure 2: The trucks that were used at each incident type per truck type

Figure 3 shows us the number of trucks used over days. This is quite an interesting plot, since it shows some patterns. For example, when we look at new year's evening we see a returning peak. Furthermore, there are also some random days with peaks in the number of trucks used. This could be an indication of a big incident on such days. What this plot shows us, is that we need to consider new year's eve separately in our model.

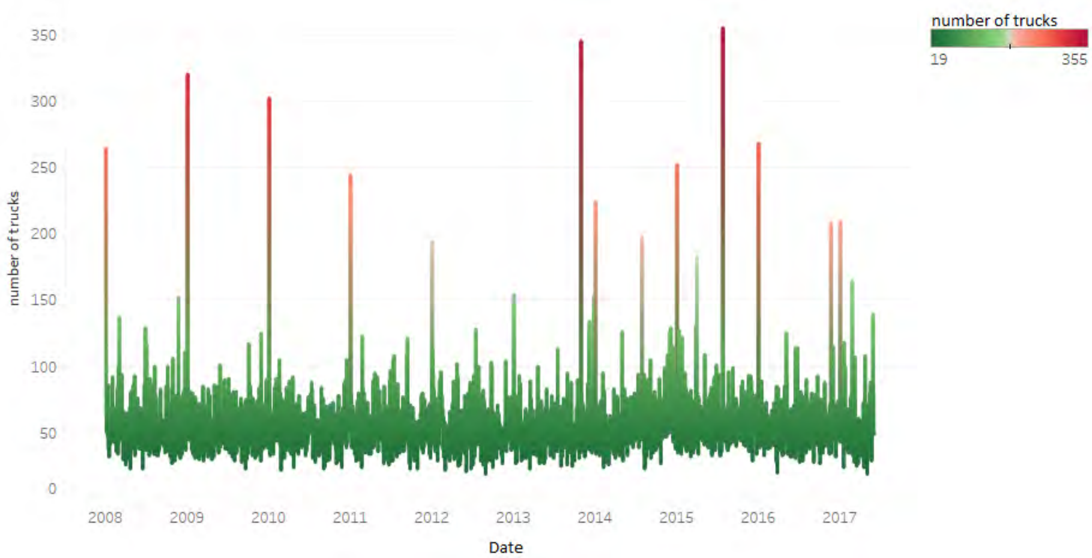

Figure 3: The number of trucks used per day

2.1.2 The stations

In this section we will present some analyses about the fire stations and the incidents.

Figure 4 shows us the distribution between the number of stations and the number of incidents. The distribution of the number of trucks used within a station is specified in the color distinction. It is no surprise anymore that the TS and RV trucks are used the most after the last section. Once again TS trucks are the first to be sent to an incident.

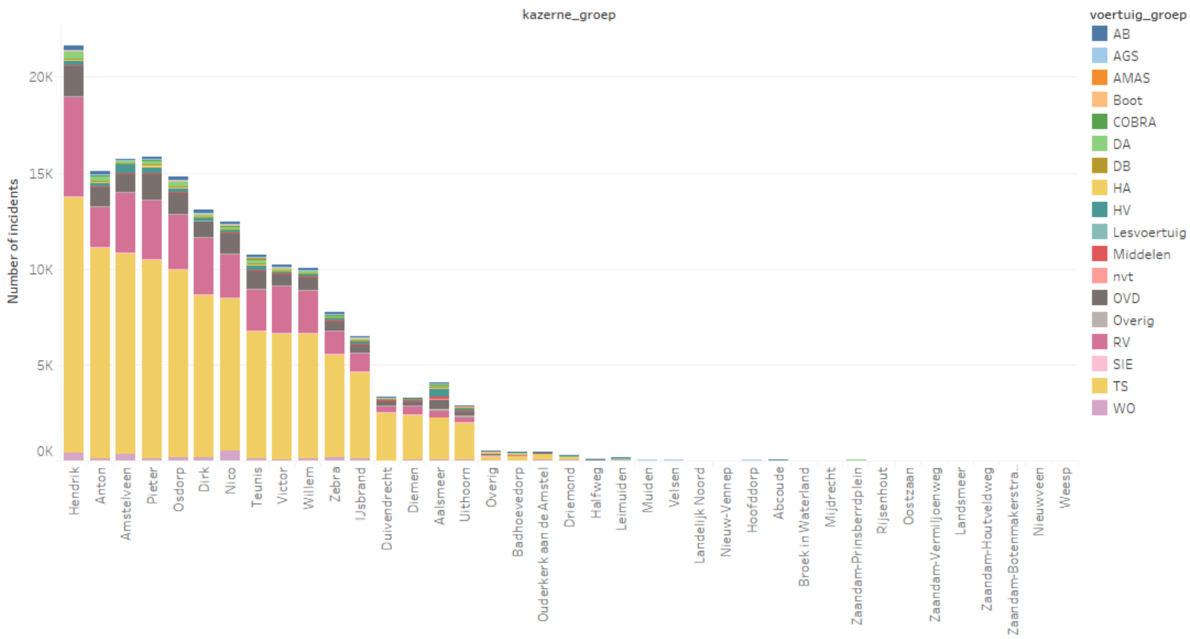

Figure 4: The stations together with the truck distribution vs. the number incidents

Figure 5 is a plot of the locations of the stations in the upper plot and the stations arriving first at an incident in the lower plot. The first arrivals seem to be distributed in parts of areas. During the research it has come to our attention that each station has its own working region. Based on this figure we suspect that these regions are about the same as the distribution in this plot. However, we do not have evidence to prove this. This plot makes it reasonable to assume that the station closest to an incident is sent first.

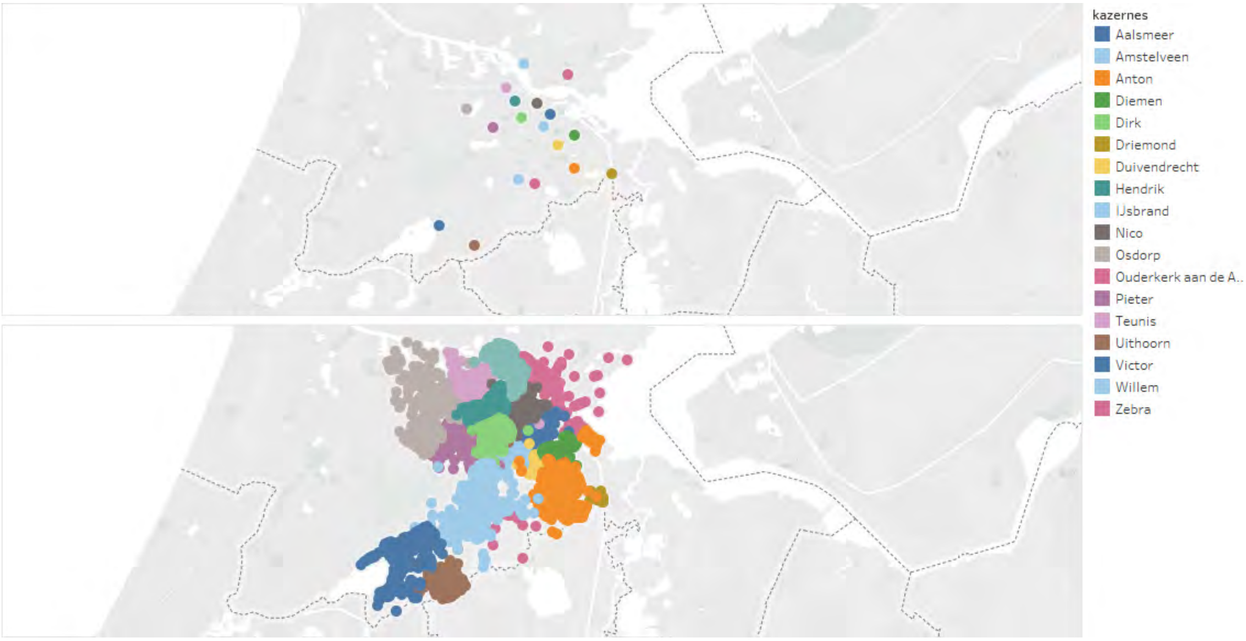

Figure 5: The stations arriving first at an incident

Another interesting observation we made is that of the time to genuine large incidents. Their was a suspicion that the driving times to big incidents is longer than to regular incidents. Since they wanted to know the driving time to genuine large incidents we defined a genuine large incident as one that has had more than ten TS trucks involved. What we did is find the average time from the call to arriving at the incident scene. There are 35 incidents which are genuinely large using our definition. In table 1 the comparison between the average time, average distance in meters as the crow flies between the large incidents and all data, but the large incidents.

Table 1: The average times to an incident and distances in meters

Time in seconds	distance in meters	data
1727.9322	3449.4696	Large incidents
1414.8799	1809.4007	All data

And indeed, both the time to an incident and the distance in meters are larger for the large incidents compared to all the data, but the large incidents. In table 20 the average times and distances to the neighborhoods could be found.

2.1.3 Neighborhood

This section will provide us information about neighborhoods and areas in relation to incidents.

Figure 7 gives us insight in which district encounter what percentage of incidents. Only the eleven largest districts are kept. This is because the percentage of incidents is negligible in the districts not taken into account. Stadsdeel Zuid has the largest amount of incidents, followed by Stadsdeel Centrum. This has a relation with the people living in these districts and neighborhoods which will be shown next.

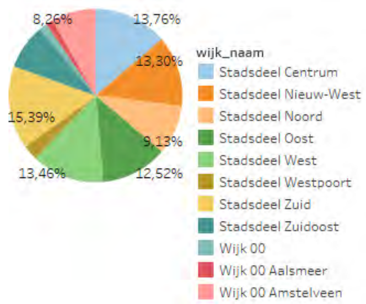

Figure 6: People living in neighborhood vs. number of incidents

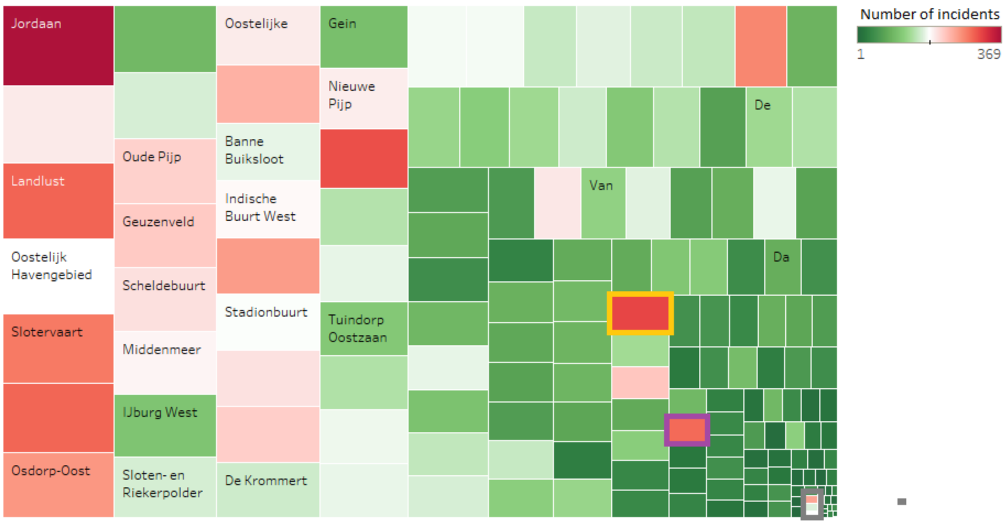

Figure 7: People living in neighborhood vs. number of incidents

What Figure 7 shows is a relation between the number of people living in a neighborhood and the number of incidents. The number of people living determine the size of a square and the color determines the number of incidents. Dark red means there were many incidents in that neighborhood, whereas green means there were only a few incidents in that area. Interesting to see is that the neighborhood Jordaan has the most people living and the most incidents. Even more interesting is to see area the neighborhoods where there are many incidents, but lesser people living. The neighborhood with the yellow box is Nieuwmarkt. The purple box is Sloterdijk. This is a industrial area, so these industries could be the reason that there are so many incidents in this neighborhood. The most interesting neighborhoods are the ones in the gray area in the bottom right. These are also industrial neighborhoods like the port of Amsterdam. Also the ratio incidents/people living in the gray area is around one, since there are relatively few people living here, but still the amount of incidents is quite significantly. In table 20 you can see the average time and distance to a neighborhood. In table 21 the genuine largest incidents are shown. One could notice that the neighborhood

Westelijk Havengebied and Oostelijk Havengebied (which are part of the port of Amsterdam) have the most of these types of incidents.

We also calculated the Pearson product-moment correlation coefficients for several relationships between the CBS data and the number of incidents.

Table 2: the Pearson product-moment correlation coefficients

	Number of incidents
Number of inhabitants	0.763652778
Number of male inhabitants	0.754069277
Number of female inhabitants	0.771738066
Population density	0.131327935
Western foreigners	0.176412095
Non-western foreigners	0.17641231

It is interesting to observe that the number inhabitants has a strong correlation with the number of incidents. This holds also for the number of male and female inhabitants. On the other hand, population density and foreigners seems to be almost independent of the number of incidents.

2.1.4 The weather

In this section we will try to analyze some weather components in relation to the number of incidents.

Figure 12 shows us the relationship between minimum visibility and the number of incidents. The x-axis here means: 0: <100 m, 1:100-200 m, 2:200-300 m,..., 49:4900-5000 m, 50:5-6 km, 56:6-7 km, 57:7-8 km,..., 79:29-30 km, 80:30-35 km, 81:35-40 km,..., 89: >70 km. Now, the interesting part here is that there is a peak when the visibility is extremely low. However, between 50 and 70 there are also some peaks. That does not mean that it is reasonable to say that under these visibility conditions there is a higher probability of incidents. It may just as well be that these condition are normal Dutch visibility conditions.

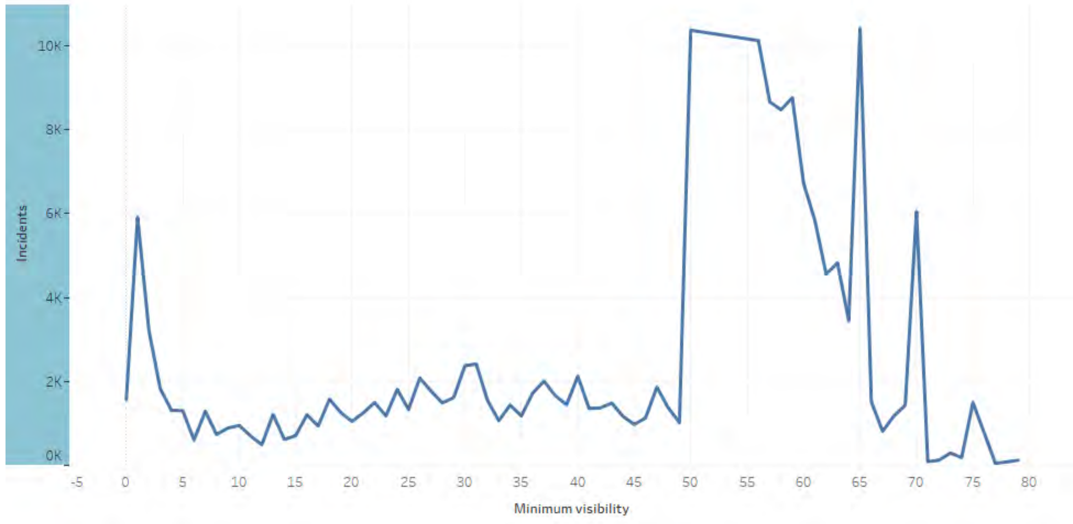

Figure 8: Minimum visibility vs. number of incidents

Another example is Figure 9. This figure shows the relationship between the number of incidents and the temperature in 0.1 degrees Celsius. Here we see that the number of incidents are concentrated between -40 and 240, meaning most incidents happen between -4 degrees Celsius and 24 degrees Celsius. Once again, these are temperatures which are common in the Netherlands. As a conclusion it is not safe to say that these weather conditions have a relationship with the number of incidents.

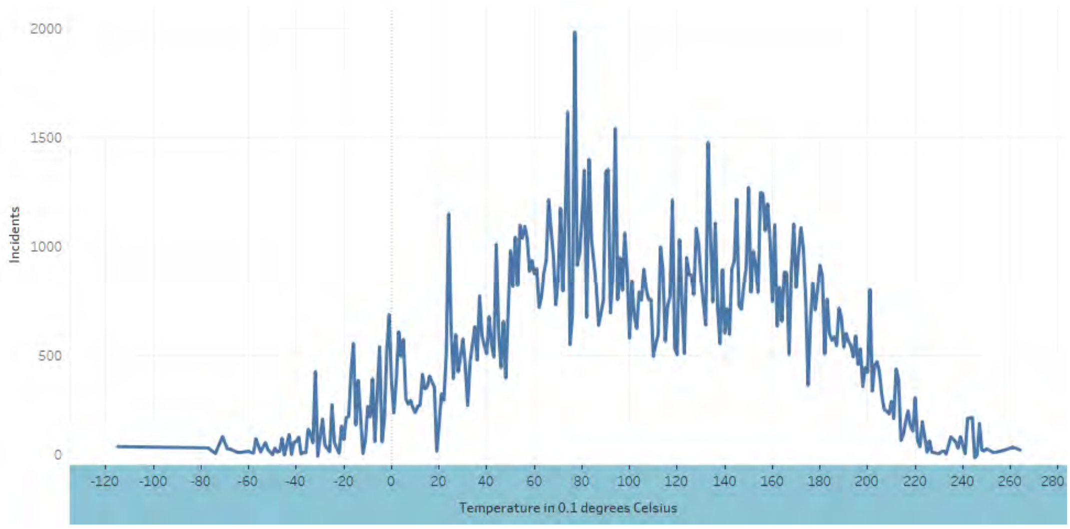

Figure 9: Temperature in 0.1 degrees Celsius vs. number of incidents

These conclusions could also be made for other weather variables. So, we are going to insert these variables into the algorithms we are going to use and let the algorithms decide whether they are important or not for incidents.

2.2 Large incidents

In section 2.1.2 we talked about genuinely large incidents when determining the time and distance to these incidents. There are also other incidents who are entitled to the definition large, but are not genuinely large. There are two definitions available and both will be used in this research. De Deijn (2017) defined a large incident as an incident that was handled by 6 or more trucks, regardless of truck type. He proved using an inhomogeneous Poisson process that these incidents occur randomly. The second definition is that of the fire department itself. A fire is called large when three or more distinct trucks go to an incident. De Deijn's definition also counted the same trucks who went multiple days to the same incident as one truck extra.

For the data analysis we will use de Deijn's definition of large incidents. For the modeling part the fire department's definition will be used.

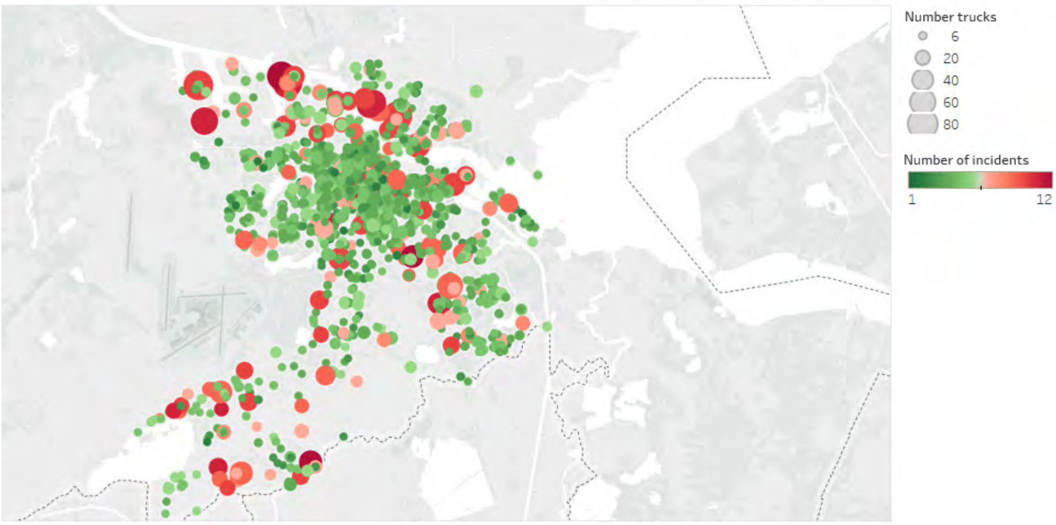

Figure 10: Number of big incidents and trucks per postal code

Figure 10 shows the distribution of large incidents per postal code. The circles are larger when there went more trucks to the same postal code. The color shows the number of incidents there have been on that postal code. A remarkable observation is that around the river IJ there were many large incidents. When we look at the table 21 in the appendix we see that there was quite a number of genuinely large incidents at the Westelijk havengebied and Oostelijk havengebied, which is also at the IJ.

2.3 New year's eve

Figure 3 showed us that there was a returning pattern for new year's eve. That is the reason why we are going to explore the data on these days separately. The pattern showed us yearly peaks at new year's eve.

Figure 11 shows us the number of trucks used per district, where the colors show the type of trucks. This plot only includes the trucks which were used at new year's from 2008 until 2017. It is interesting to see that the large districts of Amsterdam needed most of the trucks, thus encountered most of the incidents. Amstelveen is quite a large area, hence needed many trucks as well. Concerning the distribution of trucks, most of the trucks used are of type TS. This is the truck sent first when an incident occurs.

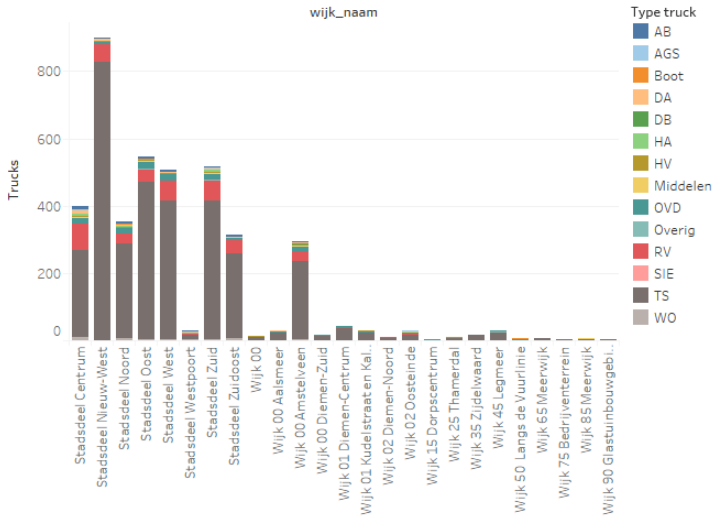

Figure 11: Number of trucks used at new year's eve per district including the distribution of truck types

3 Methods

In this section we will explain how we modeled the data in order to predict the number of trucks needed at an incident. Also how the process of preprocessing the data happened. Furthermore we will explain which algorithms were used and tell something about the theoretical background of these algorithms.

3.1 Preprocessing data

In order to prevent overfitting we had to reduce the data. There were many double features and features that all had the information about the same district or neighborhood. We obviously removed all identification columns, since they are unique for each incident. We removed the features containing information about postal codes and street names, because they are very specific features and almost unique for incidents making them prone to overfitting. We have the information about neighborhoods and districts, so information about locations would induce overfitting. Finally, we also removed features containing the exact times of the incidents. From those features we already removed necessary information like day of the week, month of year, day part and hour of the day.

3.2 Missing data

Unfortunately we had to deal with missing data and the libraries in Python simply cannot deal with the value NaN. However, when analyzing these missing data we encountered a pattern. When examining the missing data in combination with incident type we found the following. A huge part of these missing data was due to outdoor incidents. When an outdoor incident happens there is no data about address, postal code or even city. Patterns like this were not clear when examining other types of incidents. See table 22 in the appendix for the numbers of missing neighborhood-values per incident type. These numbers are comparable to other similar location specific features.

What we did to deal with these missing values is for the types where we found the above pattern we created a new category. We filled the missing feature with the type of incident. For example, if a neighborhood values was missing and the type of incident was outdoor fire, that neighborhood got the name "outdoor fire". This happened for all the missing values for that incident. The reason behind this is that the algorithms could recognize this is as a separate category and "know" that something different is happening with this category. The algorithm will still know that it belongs to the incident type. For types where we did not find any pattern we simply created the category "unknown".

3.3 Defining the predictor

In section 2.2 we talked about two definitions of large incidents. We used the one of de Deijn (2017) for data analysis purposes, but for modeling purposes we will use the definition of the fire department. This definition was that an incident is large if three or more distinct fire trucks go to an incident. For these kind of incidents we have too less data available to do a decent prediction. As for the de Deijn (2017) conclusions regarding large incidents, they occur randomly and cannot be predicted. Therefore we will use incidents which got two or less trucks per truck type.

How shall we predict the number of trucks used then? What we did is join the number of trucks used from the materials sent data to the preprocessed incident data. Using this join we got for every incident the number of truck types used per incident. Unfortunately, this was not the distinct number of trucks used at an incident. If there was a large incident and one truck had to go on multiple days to an incident it was counted as another truck. To get the distinct number of trucks we filtered the trucks based on truck type per station and day.

Finally we join the weather data to the preprocessed truck and incident data. The "algorithm-ready" looked like this:

Distinct truck type per incident	Count of truck type per incident
Weather data	Incident data

We wanted to know if we can predict the number of trucks used at an incident per truck type. In Figure 2 we saw the number of trucks used at each type incident. From this plot we see that not every truck type is used many times, meaning there not every truck type has sufficient data to do a decent prediction. Based on data available per truck type we decided to try and predict the following truck types:

1. Water tender (TS);
2. Safety truck (RV);
3. Water incidents truck (WO);
4. Service officer (OVD);
5. Assistance truck (HV)

3.4 Algorithms

In this section we will explain some theoretical background about the algorithms used to predict the number of trucks used. We will also explain the model parameters and why we chose these parameters.

3.4.1 Random Forest

Random Forest is an ensemble method. Ensembles are a divide-and-conquer method to improve performance. The main principle behind ensemble methods is that a group of weak learners can come together to form one strong learner. The weak learner in the case of random forest is a decision tree. Through bagging, a decision tree is made multiple times and combined into a strong learner. The result may either be the average or the weighted average of all the terminal nodes that are reached. It tries to find the best combination of features to predict the score. Furthermore it handles missing values well. This data set has many missing values as the exploratory data analysis showed. (L. Breiman, 2001)
The parameters chosen for this algorithm are as follows: the number of trees to build is 500; the minimum sample leaf size is 50; the maximum number of features allowed to try is the square root of the number of features and the maximum depth of the tree is 7 nodes. The reason for this parameter setting is so that the decision trees will not overfit or underfit. The trees are deep enough to make a decent prediction and have enough features to split on. Without this parameter setting the algorithm performed significantly worse.

3.4.2 Gradient Boosting

Gradient Boosting is also an ensemble method; it also combines a group of multiple weak learners into one strong learner. Once again the weak learner is a decision tree. The difference with random forest however, is that for each tree it tries to compensate the shortcomings of the earlier trees using gradient descent, by focusing on the entities that are responsible for the remaining error. This model is also excellent for classification problems. Because the algorithm tries to minimize the error and the error, for obvious reasons, has to be small, this algorithm is promising. (J.H. Friedman, 2001)
The parameters chosen for this algorithm are as follows: the number of threes to build is 500; the minimum number of observations in a node to be considered for a split is 50; the maximum depth of the trees is 6 nodes and the maximum number of features to be considered for a split is the square root of the number of features. This parameter setting has the same goal as the random forest parameter setting, namely obtaining a good fit.

3.4.3 Artificial Neural Networks

Dr. Robert Hecht-Nielsen came up with one of the simplest definitions of a neural network: *"...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.* in In "Neural Network Primer: Part I" by Maureen Caudill, AI Expert, Feb. 1989. Artificial neural networks(ANNs) are modeled after a human brain but on a much smaller scale. Complex ANNs can have hundreds or even thousands of nodes, whereas a brain has billions of neurons. Although the mathematics are not a trivial matter, it can be understood when the details are left out.

Basics of neural networks
Neural networks are typically organized in layers where each layer contains nodes. The nodes are interconnected and contain so called "activation functions". The input data is presented to the network through an input layer and the output is presented in an output layer. A neural network learn sort of in the way people learn. A child learns that a banana is a banana through pictures of bananas. A neural network learns to forecast using historical data. The learning rule which is most widely used and the one which we used as well is the delta rule. The delta rule is often utilized by the most common class of ANNs called 'backpropagational neural networks' (BPNNs). Backpropagation is an abbreviation for the backwards propagation of error. With the delta rule, learning is a supervised process which occurs which each epoch (is when the network gets an new input pattern) through a forward activation flow of outputs, and the backwards error propagation of weight adjustments. Simpler said, the network first makes a guess what the output should be and then sees how far it was from its answer and finally makes adjustments to its connection weights.

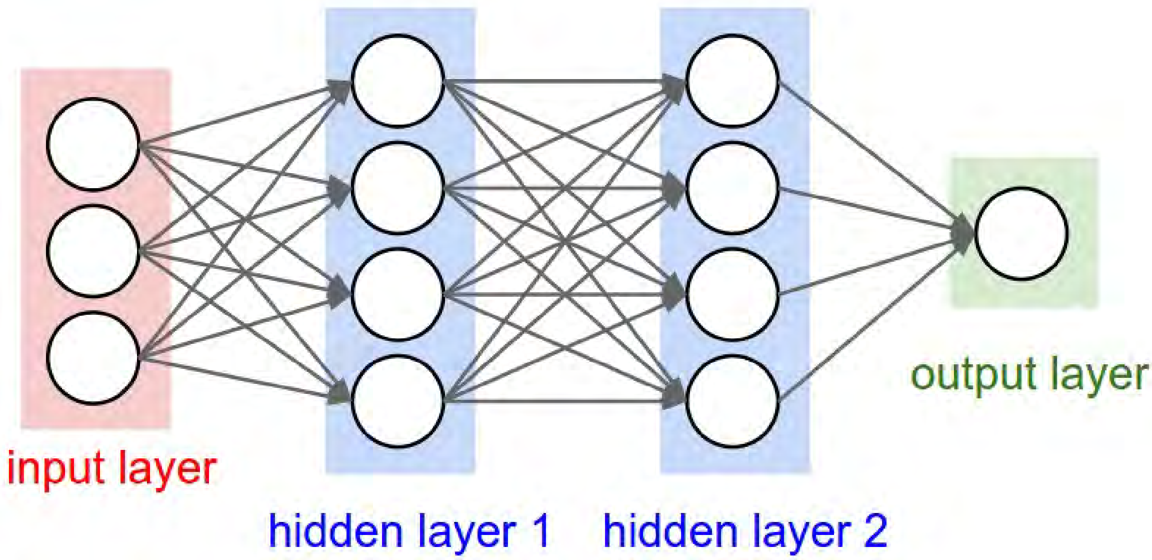

Figure 12: An example of a neural network

Advantages
In general you could expect your model to train well given your training data. This applies to problems where your data is dynamic and non-linear. ANNs are able to capture many kinds of relationships, which conventional techniques like random forests cannot do.

Limitations

- Backpropagation neural networks are the ultimate black boxes. It is not possible to analyze what happens inside the networks. There are two things the user can define or analyze and those are defining the network and analyzing the weights of the connection nodes.
- The running time is very long. This is because the machine's CPU must calculate the function of each node and connection separately, which can be

problematic when processing large amounts of data. This can be solved by calculating it on a parallel machine.

Our network
We have defined a model of one input, 6 hidden and one output layer. The input layer contains the number of input features plus 15 nodes and so does the first hidden layer. The second hidden layer contains the number of input variables plus 10 nodes, the third hidden layer number of input variables plus 5 nodes, the fourth hidden layer has the just the number of input variables as nodes, the fifth hidden layer has the number of input layers minus 5 nodes and the final hidden layer has the number of input layers minus 15. To prevent overfitting we then randomly removed 30% of the hidden layers using dropout. Finally, the output layer contained one node only, since we predicted just one variable, the number of trucks used.

The input layer and hidden layers all had the same activation function called the rectifier which is defined as $f(x) = x^+ = max(0,x)$ and the output layer had the linear activation function which is $f(x) = x$, so that it returns the predicted value even if negative.

Finally we had a gradient descent optimization function, which tried to minimize the mean absolute error in the backpropagation epochs.(2017)

3.5 Validation

To train and validate we split the data in to 70% training data and 30% validation data. The training data is used to fit the algorithms. Using this data the algorithms will try to find the optimal combinations to forecast the output variable, in this case the number of truck types needed at an incident. When the fitting has been completed, we will need to use the validation data to predict the output variable. The validation data has not been seen by the training data. Once the output variable has been predicted they have to be compared to the original output variables. The comparison, or better said validation, will be done using the **mean squared error** and R^2.

3.5.1 Mean squared error

The mean squared error is an estimator that measures the average difference between original output values and predicted output values. It is a measurement for the prediction error and the closer it is to zero the smaller the error is. The formula for the mean squared error looks like this:

$$\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$$

Where Y_i is the original value and $\hat{Y}_i$ is the predicted value. (F. Bijma, et al. 2013 p. 44)

3.5.2 R squared

The R^2 value is also called the coefficient of determination and gives information about the fit of a regression problem. What it does is the combination of several statistics. Firstly the sum of squares for Y (SSY) are needed, this is the sum of squares without any predicted variables:

$$SSY = \frac{1}{n}\sum_{i=1}^{n} Y_i - \overline{Y}$$

With $\overline{Y}$ being the the average of all original output variables.
This is followed by the residual sum of squared (RSS), which is the squared difference between original values and predicted values:

$$RSS = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)$$

With $\hat{Y}_i$ being the predicted values and Y_i the original values.
Finally we get the regression sum of squares (SS_{reg}) which is the difference between SSY and RSS:

$$SS_{reg} = SSY - RSS$$

Using all these statistics we can put them together and compose the R squared value:

$$R^2 = \frac{SS_{reg}}{SSY} = \frac{SSY - RSS}{SSY} = 1 - \frac{RSS}{SSY}$$

It holds that the closer the R^2 gets to one the better the fit of the model and if it gets negative the prediction does not follow the trend of the data. (F. Bijma, et al. 2015 p. 110)

4 Results

In this section we will present the results. Also the results of the different algorithms will be compared.

4.1 New year

To predict the number of trucks for new year we will only use the TS trucks used at incidents. This is because the other truck types do not have enough data to make a good prediction. We will predict the number of TS trucks with the random forest and gradient boosting algorithms.

4.1.1 Random Forest

The important features using random forest are given in Figure 13. When we analyze the meaning of this plot we see that the type of incident is important for the prediction of the TS trucks. Furthermore, city name and district are important when we want to know how many TS trucks are needed at an incident. The function of an building and priority of an incident also influence the prediction of the number of TS trucks.

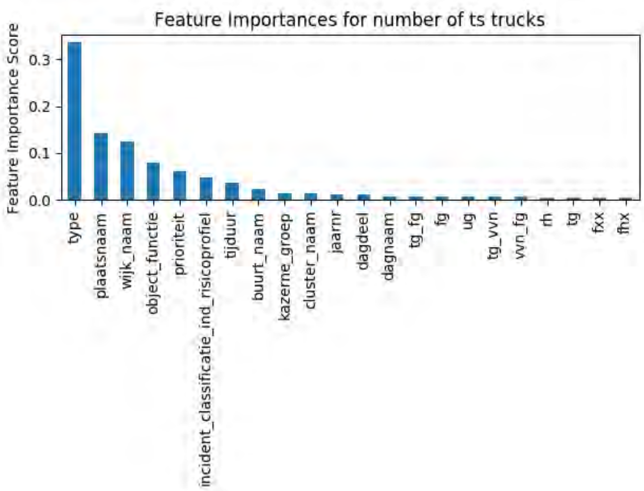

Figure 13: The feature importance of TS truck prediction on new year's eve

In figure 14 is the actual number of trucks plotted in orange versus the number of trucks predicted in blue. When we look at the prediction plot of the random forest algorithm for TS trucks, we can conclude that the prediction itself is not good. It predicts that all incidents need only one TS truck even when there are two trucks needed. The mean squared error for this case is 0.0750, which looks small, but looking at Figure 14 is not good.

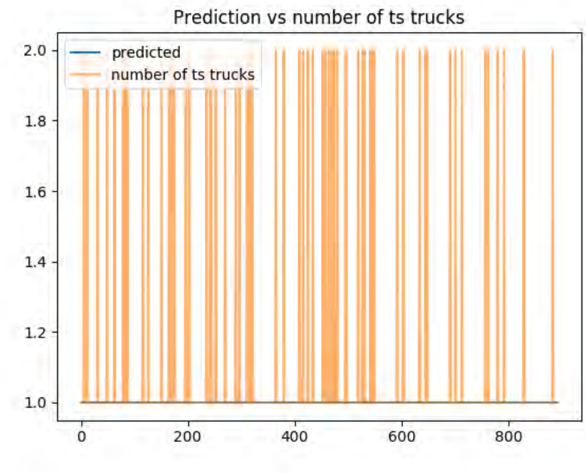

Figure 14: Number of big incidents and trucks per postal code

4.1.2 Gradient boosting

Figure 15 shows us the feature importance of the gradient boosting algorithm. When we analyze this plot we see that type of incident is important in predicting the number of TS trucks. The function of a building, city name and district are also decisive when obtaining a good prediction of TS trucks.

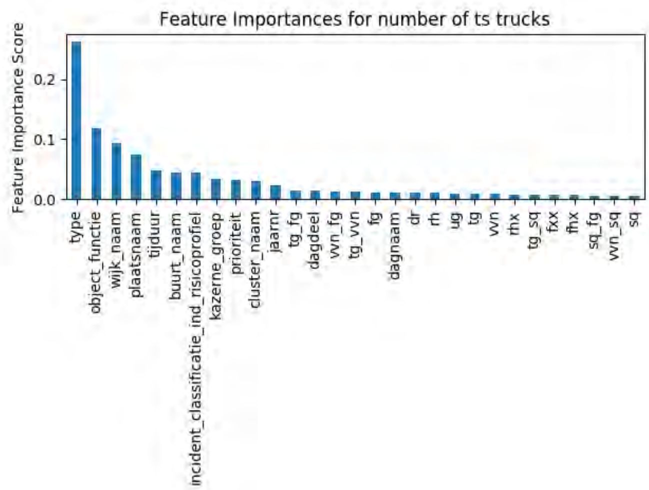

Figure 15: The feature importance of TS truck prediction on new year's eve

Figure 16 shows us the prediction plot of the number of TS trucks. This plot has a better fit when it comes to predicting two TS trucks at an incident. It should come as no surprise that the mean squared error here is quite a bit smaller than the one at random forest. The mean squared error is 0.0560.

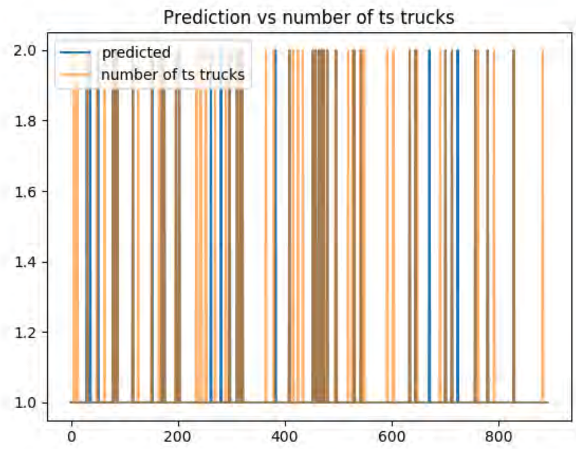

Figure 16: Number of big incidents and trucks per postal code

4.1.3 Conclusion

When looking at the feature importances of both algorithms we notice an overlap. We see that the type of an incident, function of an building and location specific features are important. Based on these features the number of TS trucks could be predicted.

When looking at the performance of both algorithms gradient boosting outperforms random forest. The mean squared error is better and the fit of the prediction against the real data is better.

4.2 Regular days

In this section we discuss the results of all the days other than new year's eve. We will compare the performance of the algorithms and discuss the feature importances.

4.2.1 TS trucks

Random forest Figure 17 shows us the feature importances for predicting the TS trucks. Looking at the plot, the type of incident is most important followed by the building function. Interesting to see is that the fire station from where the truck is coming is also important according to the random forest algorithm. Looking at the performance the mean squared error is 0.0489 and the R^2 value is 0.2643.

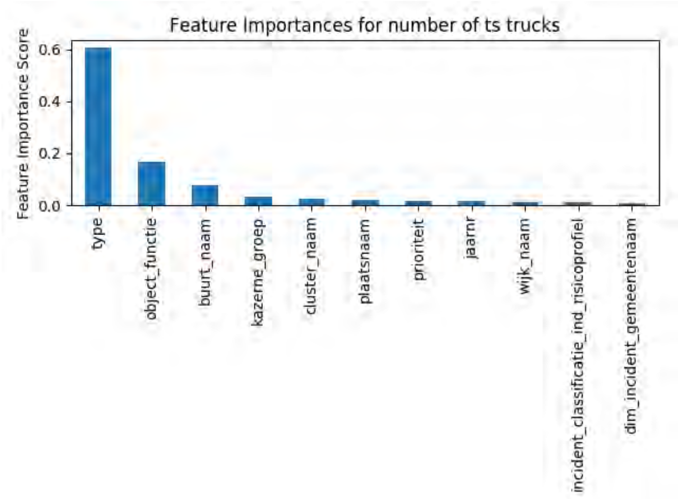

Figure 17: Feature importances random forest for prediction of TS trucks

Gradient boosting According to Figure 18 gradient boosting the most important feature is once again type of incident. The function of a building, neighborhood and station are also important. If the performance of gradient boosting is evaluated the mean squared error is 0.0444 and the R^2 is 0.3815.

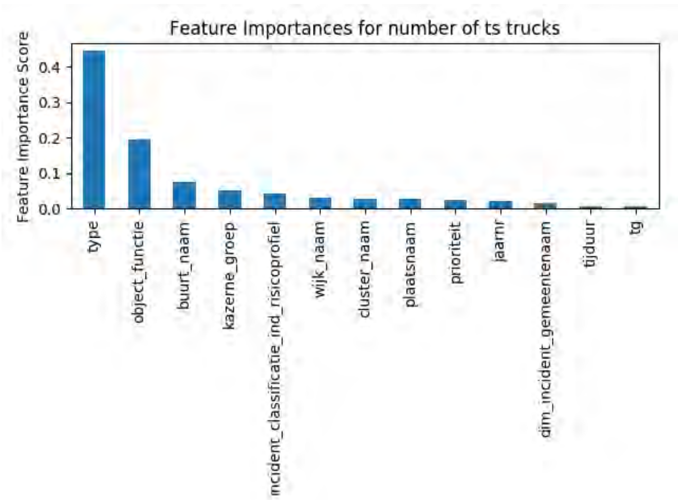

Figure 18: Feature importances gradient boosting for prediction of TS trucks

4.2.2 RV trucks

Random forest In Figure 19 we can conclude that the type of incident and district are by far the most important to determine the number of RV trucks needed at an incident. This is also something we saw in the previous section. To quantify the performance of this algorithm we have a mean squared error value of 0.03918 and an R^2 value of -0.2316. Although the performance looks decent with a small error, the fit is poor.

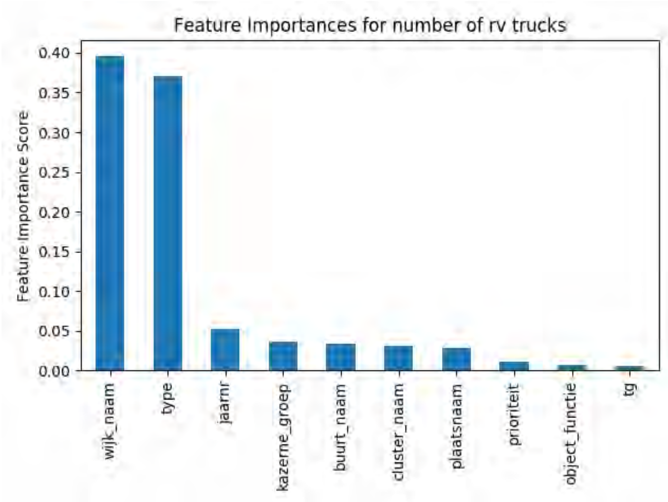

Figure 19: Feature importances random forest for prediction of RV trucks

Gradient boosting In Figure 20 we once again see that district and type of incident are most important to determine the number of RV trucks used. The performance of this algorithm is as follows: 0.0410 as mean squared error value and -0.4957 as R^2 value.

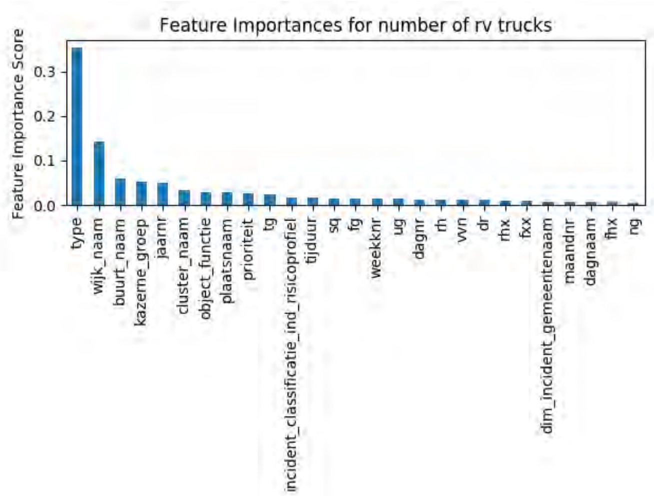

Figure 20: Feature importances gradient boosting for prediction of RV trucks

4.2.3 HV trucks

Random forest In Figure 21 the most important features are given to determine the number of HV trucks used. Not only are district and function of a building important to decide the number of HV trucks at an incident, but there are some weather variables important as well. Fg, tg and sq mean the average rainfall, temperature and sunshine of a day, so apparently they influence the decision of how many HV trucks to use. The mean squared error and the R^2 are 0.0501 and 0 respectively.

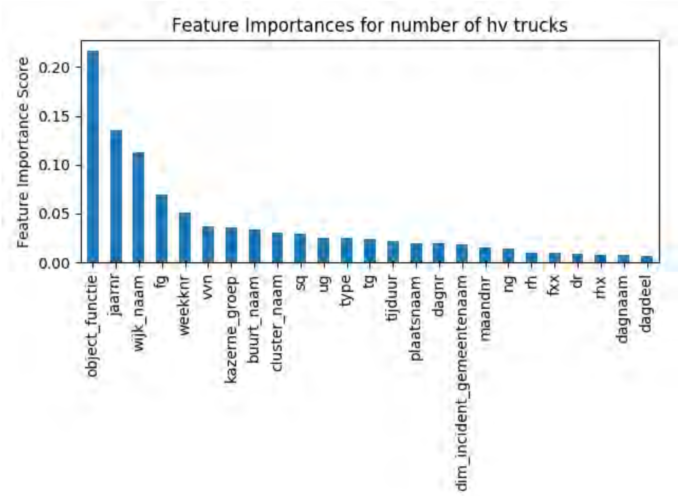

Figure 21: Feature importances random forest for prediction of HV trucks

Gradient boosting Figure 22 shows that apart form date, incident and location specific features, the vvn (windspeed) is important. The mean squared error and the R^2 are 0.0501 and 0 respectively.

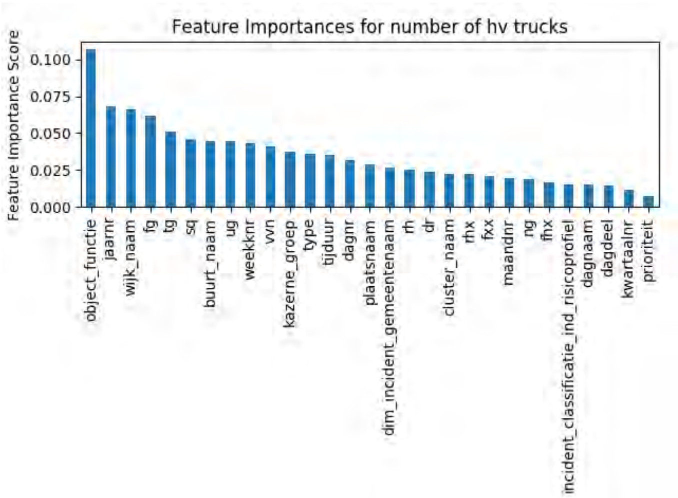

Figure 22: Feature importances gradient boosting for prediction of HV trucks

Something remarkable is happening here, both the mean squared error and R^2 are the same. Figure 23 has the answer to this observation. The algorithms only predict that one HV truck is needed for every incident. This is because most of the time when there is more than one HV truck needed a large incident has occurred. These occur randomly and therefore cannot be predicted.

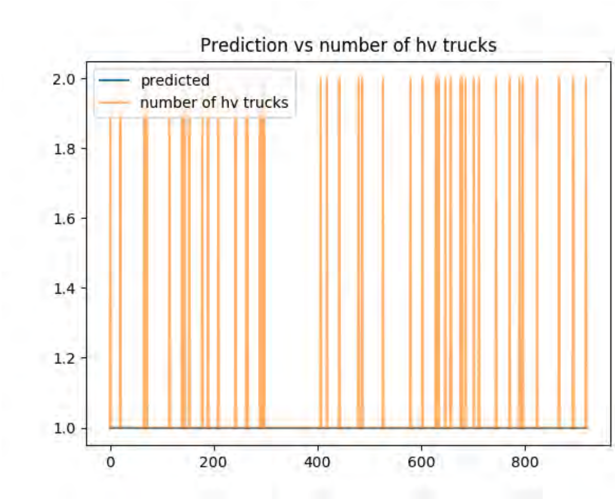

Figure 23: Predictions vs. originals of the HV trucks

4.2.4 OVD

Random forest By now it is evident that type of incident and neighborhood factors are the most important features that decide the number of trucks needed. Figure 24 also gives us that insight. Temperature and windspeed are also of importance in this algorithm.

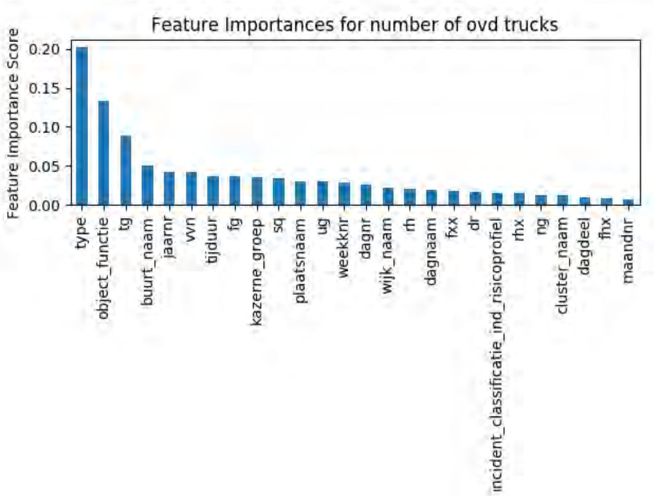

Figure 24: Feature importances random forest for prediction of OVD's

Gradient boosting In Figure 25 we see that temperature and sunshine on the day are important, whereas windspeed and rainfall also have significant importance. The other important features are evident based on previous analyses of the features.

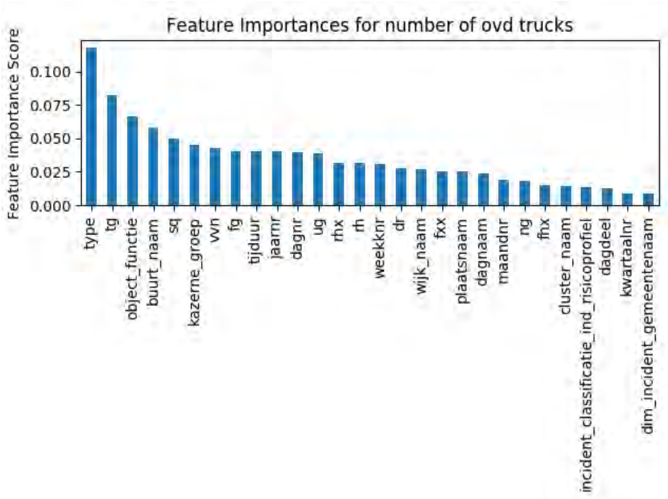

Figure 25: Feature importances gradient boosting for prediction of OVD's

The metrics for both models are mean squared error with a value of 0.0225 and an R^2 value of 0. This means that once again only the value one has been predicted. Figure 26 is its understatement. This has to do with the fact that multiple OVD's are most of the time only needed at a large incident.

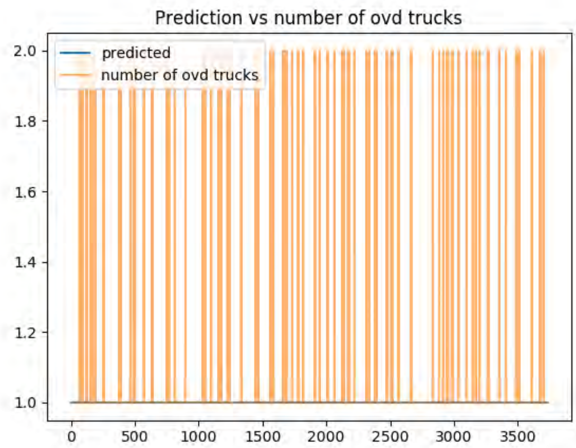

Figure 26: Predictions vs. originals of the OVD's

4.2.5 WO trucks

Random forest Figure 27 had an interesting feature which is most important. The fire department gives a priority to each incident, based on the information received from a call and that priority is most important to decide the number of WO trucks. The station could be important, because not all stations have WO trucks. Furthermore we see the evident features of type and location.

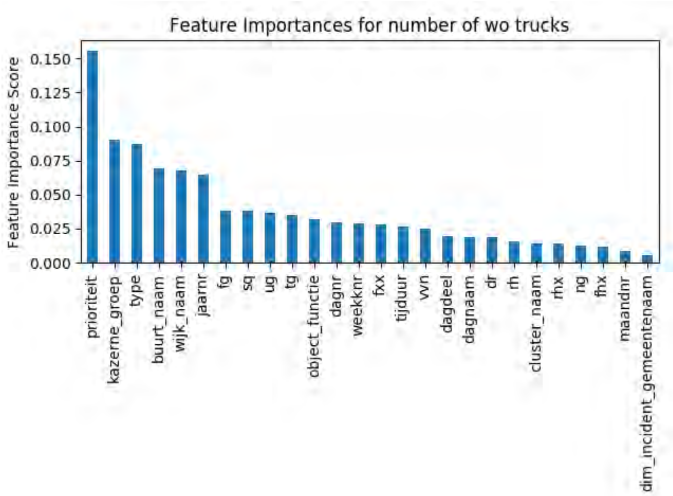

Figure 27: Feature importances random forest for prediction of WO trucks

Gradient boosting Figure 28 shows us the evident features and weather features as important features. Rainfall, temperature, sunshine and humidity.

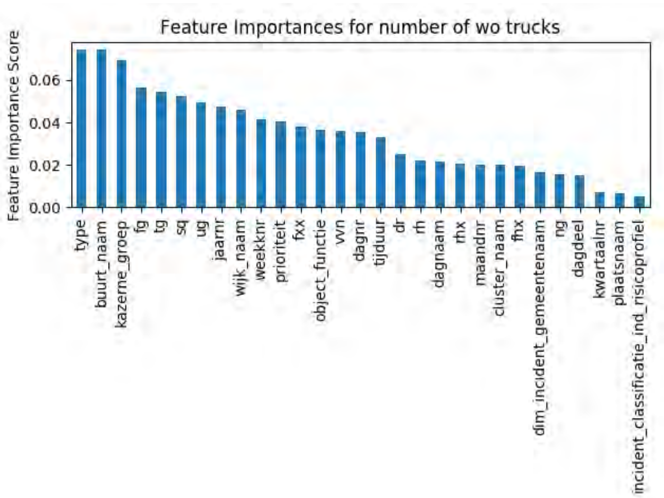

Figure 28: Feature importances random forest for prediction of WO trucks

The performance of these algorithms returned a mean squared error value of 0.0767 and an R^2 value of 0. This is again due to the fact that only ones have been predicted as number of WO trucks needed, which is shown in Figure 29.

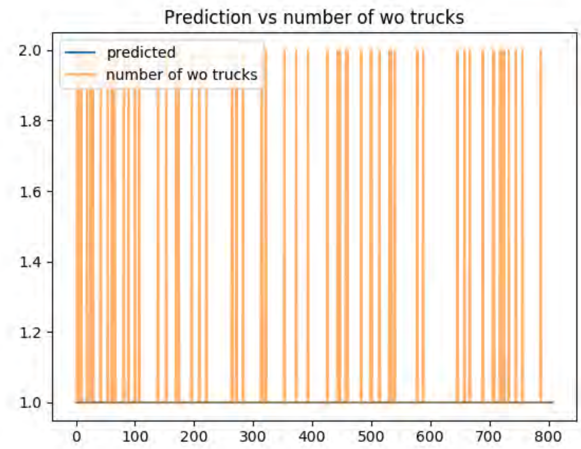

Figure 29: Predictions vs. originals of the WO trucks

4.3 Feature selection

The algorithms showed us which features are important and which are poor. When all features are given to the algorithms, they will try to make decision trees with all those features. Therefore, some trees will return a higher error, which ultimately results in poorer general performance. In order to improve the performance of the algorithms we have chosen do some feature selection, so that the features will make decision trees with better performance.

In the previous section we saw that OVD's, HV and WO trucks cannot be predicted. Therefore, we will only focus on TS an RV trucks. We will see and compare what happens to the performance of the models. We will select the top ten and top five features from the feature importance figures.

4.3.1 TS trucks

Random forest

Table 3: The mean squared error performance of TS with feature selection for random forest

All features	Top 10	Top 5
0.0489	0.0433	0.0420

Table 4: The R^2 performance of TS with feature selection for random forest

All features	Top 10	Top 5
0.264318420166	0.440209151076	0.453919010307

In Table 3 and 4 we see that indeed with feature selection the performance improves. The mean squared error becomes smaller and the R^2 gets closer to one.

Gradient boosting

Table 5: The mean squared error performance of TS with feature selection for gradient boosting

All features	Top 10	Top 5
0.0444	0.0388	0.0378

Table 6: The R^2 performance of TS with feature selection for gradient boosting

All features	Top 10	Top 5
0.3815	0.4589	0.5526

The gradient boosting algorithm with feature selection performs better than the random forest algorithm. It has both larger R^2 values and smaller mean squared error values.

4.3.2 RV trucks

Random forest

Table 7: The mean squared error performance of RV with feature selection for random forest

All features	Top 10	Top 5
0.0392	0.04023	0.04013

Table 8: The R^2 performance of RV with feature selection for random forest

All features	Top 10	Top 5
-0.2316	-0.4032	-0.3852

In Tables 7 and 8 we see that feature selection does not improve the performance. Feature selection between the top 10 and top 5 features makes a difference, but not compared to all the features. We suspect that some overfitting sneaked into the algorithm.

Gradient boosting

Table 9: The mean squared error performance of RV with feature selection for gradient boosting

All features	Top 10	Top 5
0.0410	0.0388	0.0387

Table 10: The R^2 performance of RV with feature selection for gradient boosting

All features	Top 10	Top 5
-0.4957	-0.2184	-0.2036

We see that feature selection has tremendously improved the performance of gradient boosting. Unfortunately the fit is still bad.

4.4 Combination of weather variables

We saw in section 4.2 that weather features are important features when it comes to predicting certain kind of truck types. Following these observations, we want to observe what combinations of features will do to the model. We will look at TS and RV trucks only, because the other truck types do not have predictive power. Furthermore, we will add the combinations to the top 5 features only. The reason behind this choice is that in general the performance improved with feature selection. The only exception in this case is for the RV trucks with random forest, but as we stated we suspect that there was overfitting.

The combinations of weather variables we added are:

- Temperature and hours of sunshine (tg_sq);
- Temperature and visibility (tg_vvn);
- Temperature and rainfall (tg_fg);
- Visibility and hours of sunshine (vvn_sq);
- Visibility and rainfall (vvn_fg);
- Hours of sunshine and rainfall (sq_fg)

4.4.1 TS trucks

Random forest Figure 30 shows the feature importances of the random forest algorithm. The combined weather features are not really important in deciding the number of TS trucks necessary.

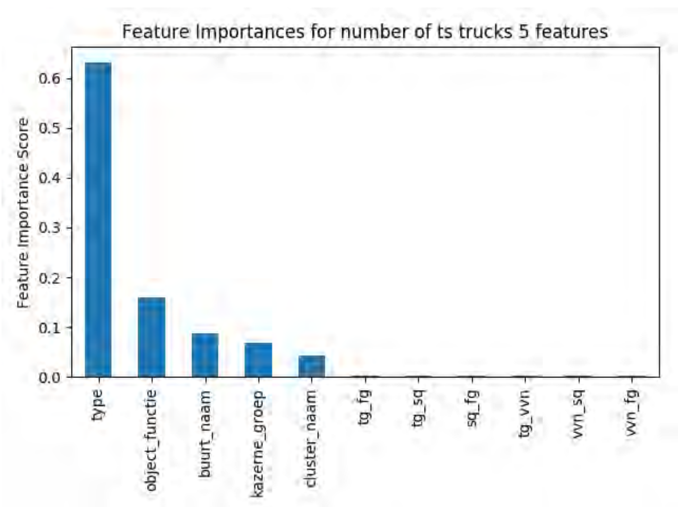

Figure 30: Feature importances random forest for prediction of TS trucks

When comparing the metrics in tables 11 and 12 we see not only that those features are not important, they also decrease the performance significantly.

Table 11: The mean squared error performance of TS with weather combination for random forest

Top 5	Top 5 + combined features
0.0420	0.0531

Table 12: The R^2 performance of TS with weather combination for random forest

Top 5	Top 5 + combined features
0.4539	0.1705

Gradient boosting Figure 31 also shows us that combined weather features do not add much to the algorithm.

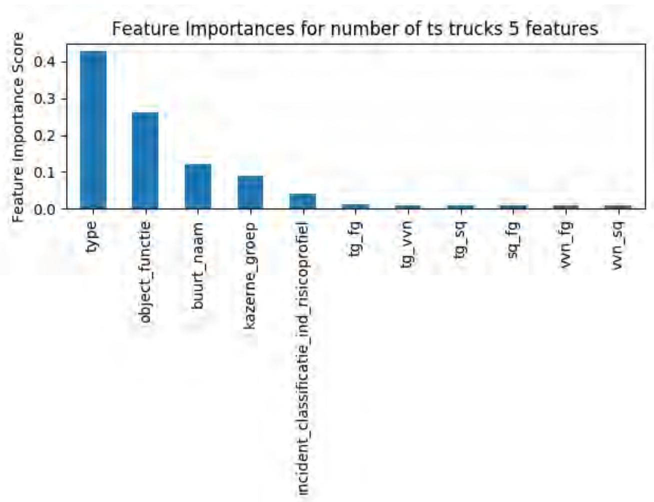

Figure 31: Feature importances gradient boosting for prediction of TS trucks

Tables 13 and 14 show that the added features just make the model perform worse. Both the mean squared error and R^2 values worse when evaluating the model with the added features.

Table 13: The mean squared error performance of TS with weather combination for gradient boosting

Top 5	Top 5 + combined features
0.0377	0.0419

Table 14: The R^2 performance of TS with weather combination for gradient boosting

Top 5	Top 5 + combined features
0.5526	0.4738

4.4.2 RV trucks

Random forest Figure 32 prove that the combined weather variables do not add much to the algorithm. In tables 15 and 16 we see that the performance has not improved. Adding the features actually has decreased the performance.

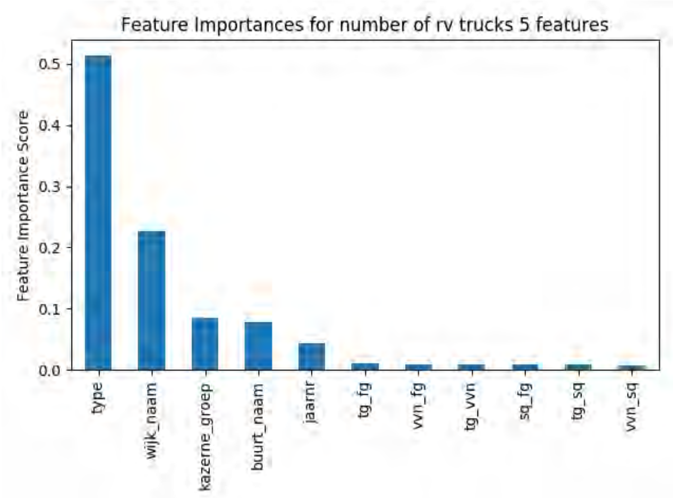

Figure 32: Feature importances random forest for prediction of RV trucks

Table 15: The mean squared error performance of RV with weather combination for random forest

Top 5	Top 5 + combined features
0.04013	0.0436

Table 16: The R^2 of RV with weather combination for random forest

Top 5	Top 5 + combined features
-0.3852	-0.8981

Gradient boosting Figure 33 also shows that adding combined features really does not create added value. As for the performance, tables 17 and 19 show that it has not improved. As a matter of fact it has performed worse.

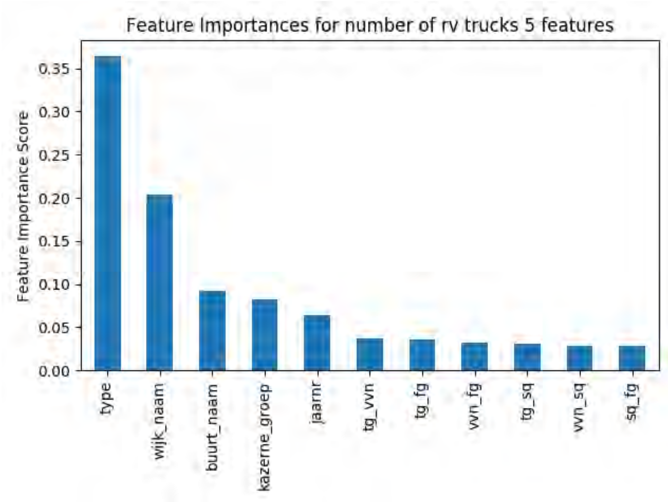

Figure 33: Feature importances gradient boosting for prediction of RV trucks

Table 17: The mean squared error performance of RV with weather combination for gradient boosting

Top 5	Top 5 + combined features
0.0387	0.0400

Table 18: The R^2 performance of RV with weather combination for gradient boosting

Top 5	Top 5 + combined features
-0.2036	-0.3585

4.5 Neural Networks

In order to improve the performance of random forest and gradient boosting we tried to fit a neural network to predict. We focused on TS and RV trucks. As described in in section 3.4.3 we stated that a neural network is extremely difficult to analyze as to what happens inside the structure. Unfortunately, we could not say anything about which features are important. We are only limited to evaluating the performance of the outcome.

Regarding the performance of the model for predicting TS and RV trucks we can say that is poor. It has performed no where near as good as the previous algorithms. The mean squared error and R^2 are both extremely bad compared to the previous algorithms. Unfortunately, we were not able to properly tune the algorithm. This was because of the horrible running time. The algorithm was running for 45 minutes to fit the TS trucks.

Table 19: The performance of neural networks for TS and RV trucks

TS		RV	
mean squared error	R^2	mean squared error	R^2
0.0953	-6.5373	0.0588	-2.9828

5 Conclusions and recommendations

5.1 Conclusion

Regarding the features importances we saw that there are some specific features which appeared as important for every truck type. These are type of incident, station and location specific features like city, district and neighborhood and function of the building. Furthermore there were some specific features for each truck type and algorithm:

- TS: riskprofile (gradient boosting)
- HV: visibility (random forest)
- HV: rainfall, temperature, sunshine (gradient boosting)
- OVD: temperature and visibility (random forest)
- OVD: temperature, sunshine, visibility and rainfall (gradient boosting)
- WO: priority (random forest)
- WO: rainfall, temperature, sunshine, humidity (gradient boosting)

When we look at performance and predictive power we see that OVD's, HV and WO trucks cannot be predicted. This is because most of the time when there is more than one truck needed of these types it is for a large incident. We saw that large incidents occur randomly and cannot be predicted. The algorithm returns exactly the same as what mathematics proved. Regarding TS and RV trucks we saw that gradient boosting performed better for TS trucks (resulting in better mean squared error and R^2 values) and worse for RV trucks (resulting in worse mean squared error and R^2) than random forest.

Then we looked at feature selection for TS and RV trucks. This showed in general great improvement. If the top 10 and top 5 features were chosen, the performance increased. The only case where this was not the case was for the RV trucks. In this case the performance actually decreased. We suspect that overfitting has sneaked into the fitting of the data with the algorithm. If we decrease the features from 10 to 5, the performance does increase.

To properly examine the weather variables we created combinations weather variables to explicitly give the algorithms the option to chose these combinations. Unfortunately, these combination were not only unimportant, they actually made the algorithms perform worse. We did this for TS and RV trucks, because the other truck types do not have predicting power.

Finally, we tried fitting the data on a artificial neural network. We can be brief about this attempt. It did not work out very well. The results were poor. Much more poor than the worst performances of random forest and gradient boosting.

We were not able to improve the model, because of the horrible running time. This was many times larger than for random forest and gradient boosting.

To return to our research question: *Can we make a good forecast of the number of trucks per type needed at each incident? Furthermore, can we do this using the combination of incident data, weather data and demographic data?*
We would say that we succeeded for a part only. We were unable to make a decent prediction for the OVD's, HV trucks and WO trucks. We managed to make a forecast for TS and RV trucks. However, the fit for the RV trucks was also pretty poor. The only truck type we actually managed to do make a decent forecast with was TS truck type. When looked at the feature importances of the data we see that the prediction mostly got its performance from incident features. Weather features only decreased the performance and demographic data was only used for data analysis, since we only had data from 2013.

Regarding the importance of this research for the fire department we can say that we have found some interesting insights during the data analysis phase of this research. Furthermore, we have made clear what features are important for assessing the number of trucks needed at an incident.

5.2 Recommendations

We have found the features which are important in determining the number of truck types needed. However, we did not look at the explicit relationship between these variables and the number of trucks needed at an incident. So, given these important features it would be useful to research what the influence of each feature is on the truck usage.

Unfortunately we did not manage to get good results with a neural network. We tried only one neural network structure and did not improve this structure because of lack of time. Furthermore, the running time of the algorithm was long compared to the decision tree algorithms. An opportunity to improve the forecasting of the number of truck types necessary is to further research neural networks. There are also different scopes of neural networks to be tested, like convolutional neural networks or LSTMs. Apart from these scopes different structures of artificial neural networks should be tried as well.

Since the function of a building is important in determining the number of truck types necessary at an incident, it would be interesting to have some building specific data. This data could be about the year of build, type of building etc. We think that data like this could improve the performance of the model and will definitely give insight in the data analysis phase.

Our final recommendation is to give a mathematical prove of randomness of the occurrence of large incidents using the definition of the fire department. This was that an incident is called large when three or more trucks go to the incident location. We assumed in combination with de Deijn's prove that it is reasonable to say that they are randomly occurring. The first prediction showed the same for HV, OVD and WO trucks. Mostly, more than one truck of this type is used when an incident is large. The prediction showed that if there were more than one truck used, it was not able to predict them. However, this is not a mathematical prove.

6 Appendix

Table 20: The 25 longest average time and distance to an incident per neighborhood

Time in seconds	Distance in meters	neighborhood
6274.4362606232294618	1148.72231923599	Kinkerbuurt
4364.7456546929316338	775.906007668428	Frankendael
4323.9630522088353414	2790.98225675215	Buiksloterham
3692.7720046529662660	5265.14018734812	Westelijk Havengebied
3495.8333333333333333	2297.0356224675	Sloterdijk
3436.3021316033364226	743.866635290159	Transvaalbuurt
3227.4428364688856729	1281.84005068562	IJselbuurt
3191.4647887323943662	3635.34589316379	Meerwijk
3174.9094397544128933	1249.15932866426	Indische Buurt Oost
3151.3375262054507338	1118.93447871086	Vondelbuurt
2823.1610738255033557	4439.37466488651	Greenpark
2769.3025974025974026	1755.16876103468	Stadionbuurt
2689.4192495921696574	1521.95676590717	Hoofddorppleinbuurt
2680.8452758402029169	1761.11731664903	Nieuwendam-Noord
2525.9399161620866325	954.401176069621	Oude Pijp
2472.1123139377537212	1648.09467218618	Willemspark
2468.1635514018691589	9739.03641540902	Bedrijventerrein
2461.2153518123667377	1850.41934637553	Haarlemmerbuurt
2436.7941952506596306	766.629842428863	Da Costabuurt
2398.8926553672316384	673.856905059549	Duivelseiland
2373.1396648044692737	2471.69895311998	Banne Buiksloot
2262.1523178807947020	3666.32773783907	Zeeburgereiland/Nieuwe Diep
2253.4747409326424870	1333.16308327115	Grachtengordel-West
2234.9252669039145907	2925.19771587751	Nieuwendammerham
2229.7931488801054018	813.693201520077	Indische Buurt West

Table 21: All genuinely large incidents

yyyymmdd	straatnaam	type	voertuig_groep	number_of_trucks	kazerne_groep	buurt_naam	wijk_naam
2008-06-26	Klimopweg	Binnenbrand	HA	12	IJsbrand	Buiksloterham	Stadsdeel Noord
2008-06-26	Klimopweg	Binnenbrand	TS	19	IJsbrand	Buiksloterham	Stadsdeel Noord
2008-07-17	Kortvoort	Binnenbrand	TS	11	Anton	Bijlmer-Oost (E, G, K)	Stadsdeel Zuidoost
2008-10-27	Drechtdijk	Binnenbrand	TS	12	Uithoorn	Meerwijk	Wijk 85 Meerwijk
2008-12-21	Aalsmeerderweg	Binnenbrand	HA	11	Aalsmeer	Greenpark	Wijk 02 Oosteinde
2008-12-31		Buitenbrand	TS	10	Pieter	Sloten- en Riekerpolder	Stadsdeel Nieuw-West
2009-01-01		Buitenbrand	TS	12	Pieter	Sloten- en Riekerpolder	Stadsdeel Nieuw-West
2009-06-15	Cruquiusweg	Binnenbrand	TS	16	Victor	Oostelijk Havengebied	Stadsdeel Oost
2009-08-30	Orion	Binnenbrand	TS	10	Amstelveen	Middenhoven	Wijk 00 Amstelveen
2009-10-08	Katrijpstraat	Binnenbrand	TS	11	Zebra	Nieuwendam-Noord	Stadsdeel Noord
2009-11-22	Amstelveenseweg	Binnenbrand	TS	17	Pieter	Schinkelbuurt	Stadsdeel Zuid
2009-11-26	Werfstraat	Binnenbrand	TS	15	IJsbrand	Buiksloterham	Stadsdeel Noord
2010-12-14	Kerkstraat	Binnenbrand	RV	10	Dirk	Grachtengordel-Zuid	Stadsdeel Centrum
2010-12-14	Kerkstraat	Binnenbrand	TS	18	Dirk	Grachtengordel-Zuid	Stadsdeel Centrum
2010-12-31		Buitenbrand	TS	11	Pieter	Sloten- en Riekerpolder	Stadsdeel Nieuw-West
2011-02-22	Latexweg	Binnenbrand	TS	19	Osdorp	Westelijk Havengebied	Stadsdeel Westpoort
2011-07-13	Keizersgracht	Binnenbrand	TS	12	Hendrik	Grachtengordel-West	Stadsdeel Centrum
2012-04-18	Rapenburg	Hulpverlening algemeen	TS	10	Nico	Nieuwmarkt/Lastage	Stadsdeel Centrum
2013-01-08	Anne Frankstraat	Binnenbrand	TS	10	Nico	Nieuwmarkt/Lastage	Stadsdeel Centrum
2013-03-15	Havenstraat	Binnenbrand	TS	10	Pieter	Schinkelbuurt	Stadsdeel Zuid
2013-07-16	Drostenburg	Binnenbrand	TS	17	Anton	Bijlmer-Centrum (D, F, H)	Stadsdeel Zuidoost
2013-08-11	Valkenburgerstraat	Binnenbrand	TS	11	Nico	Nieuwmarkt/Lastage	Stadsdeel Centrum
2013-12-03	Heining	Binnenbrand	TS	17	Osdorp	Bedrijventerrein Sloterdijk	Stadsdeel Westpoort
2014-08-07	Ondernemingsweg	Binnenbrand	TS	14	Uithoorn	Bedrijventerrein	Wijk 75 Bedrijventerrein
2014-09-04	Beukenhorst	Binnenbrand	TS	10	Diemen	Beukenhorst	Wijk 00 Diemen-Zuid
2015-09-08	De Boelelaan	Storm en Waterschade	TS	10	Pieter	Buitenveldert-West	Stadsdeel Zuid
2016-01-01	Govert Flinckstraat	Binnenbrand	TS	12	Dirk	Oude Pijp	Stadsdeel Zuid
2016-05-04	Entrepothof	Binnenbrand	TS	10	Victor	Oostelijk Havengebied	Stadsdeel Oost
2016-05-07	Hoofddorpplein	Binnenbrand	TS	11	Pieter	Hoofddorppleinbuurt	Stadsdeel Zuid
2016-07-08	Apollolaan	Binnenbrand	TS	10	Dirk	Apollobuurt	Stadsdeel Zuid
2016-10-23	Ankerweg	Binnenbrand	TS	10	Teunis	Westelijk Havengebied	Stadsdeel Westpoort
2017-01-05	Mainhavenweg	Hulpverlening algemeen	TS	11	Teunis	Bedrijventerrein Sloterdijk	Stadsdeel Westpoort
2017-03-07	Schaafstraat	Binnenbrand	TS	13	Nico	Nieuwendammerham	Stadsdeel Noord
2017-04-22	Westhavenweg	Binnenbrand	TS	19	Teunis	Westelijk Havengebied	Stadsdeel Westpoort
2017-05-27	Kajuitweg	Binnenbrand	TS	20	Teunis	Westelijk Havengebied	Stadsdeel Westpoort

Table 22: Missing neighborhood-values per incident type

Sum of missing values	Type of incident
1	
50	Dier te water
0	Buiten dienststelling
757	Assistentie Ambulance
12	Brandbare vloeistoffen
3	Afhijsen spoed
0	Brandbare gassen
39	Reanimeren
0	Letsel eigen personeel
4	Buitensluiting
82	Persoon te water
0	Overige gevaarlijke stoffen
301	Buitenbrand
1209	Hulpverlening algemeen
42	Assistentie Politie
6	Hulpverlening Dieren
41	Storm en Waterschade
74	Voertuig te water
9	Regionale bijstand
117	Beknelling / bevrijding
15	Liftopsluiting
33	Hulpverlening water algemeen
170	NVT
21	Brandgerucht / controle
183	OMS / automatische melding
154	Binnenbrand
55	Meten / overlast / verontreiniging
104	Interregiole bijstand
2	Hulpverlening algemeen Dieren
0	Herbezetting

7 Bibliography

http://pages.cs.wisc.edu/ bolo/shipyard/neural/local.html, A Basic Introduction to Neural Networks
L. Breiman. Random forests. In Machine learning, pp. 5–32, 2001.
F. Bijma, Marianne Jonker, Aad van der Aart, Inleiding in de statistiek, 2013 p.44
F. Bijma, M.C.M. de Gunst, Statistical Data Analysis, 2015 p. 110
A. Boersema, Het voorspellen van incidenten bij de brandweer, 2013
J. de Deijn, Predicting fire brigade incidents, 2016
J. H. Friedman. Greedy function approximation: a gradient boosting machine. In Annals of Statistics, pp. 1189–1232, 2001