How to build an optimal and operational knowledge base to predict firefighters' interventions

Christophe Guyeux, Abdallah Makhoul, and Jacques M. Bahi

Femto-ST Institute, UMR 6174 CNRS, University of Bourgogne Franche-Comté
christophe.guyeux@univ-fcomte.fr

Abstract. It has recently been shown that pre-emergency transport, whether performed by firefighters or private ambulances, has a predictive character due to the fact that rescue is directly related to human activity, which is itself predictable. XGBoost has emerged as the best tool to predict the number of interventions by type, but how to design an optimal and operational knowledge base has not been discussed so far. We propose to explain how to make such a base with a content that is both relevant and can be continuously updated, making possible the industrialization of the process, and thus a better operational response of the concerned services. We show that three feature selection tools custom-built for XGBoost are mature enough to allow the optimization of such a database, and a good accuracy in predictions. We also show what these tools can bring in terms of business knowledge, and discuss the organizational and efficiency consequences that such an optimized predictive model could bring.

Keywords: Feature selection, Predictive models, XGBoost, Fire department intervention

1 Introduction

In a number of countries such as France, the fire department does not only deal with fires, but also with pre-sanitary emergency transport. Their ambulances go out for accidents, suicides, drownings, etc., and this kind of personal assistance, whether it is urgent or not, represents up to 80% of their interventions. However, if fires are difficult to predict, this is less true in the case of personal assistance. Indeed, accidents involving humans are directly related to human activity: for example, people do not usually drown in the middle of the night because they are sleeping. And variables such as the weather, the length of daylight, the presence of snow or ice, or the quality of the air, condition human activity and its accidents to a certain extent.

Being able to operationally predict the number of interventions by type is very interesting for fire and rescue centers, because it allows for example to size the staff, or to adapt the skills according to the expected risk. But being able to make such predictions requires a good knowledge of the explanatory variables impacting such or such type of intervention: we know that putting many more features than necessary, with many redundancies and correlations, only complicates the learning task, and degrades the accuracy of predictions. On the other hand, these variables must be continuously accessible, so that the predictions can also be made continuously: without this, these predictions have no operational interest.

For some time now, XGBoost has been one of the gold standards among the predictive models available in the literature [16]. This optimized distributed gradient boosting library designed to be highly efficient, flexible and portable, is frequently the tool used by finalists of Kaggle-type competitions. The regressions produced are often of very high quality, provided that the knowledge base is also of high quality. The features must be chosen with discernment, avoiding both the exclusion of variables rich in information, and the inflation of variables that are often repetitive or correlated, which will only confuse the signal and unnecessarily complicate the problem.

The selection of explanatory variables is obviously important, and can be done in various ways. Many solutions have existed in the literature for a long time, and this research area is now mature. But recently, three feature selection tools have been custom-built for XGBoost, and it has been shown that feature selection is better than with traditional techniques, when the regressor is XGBoost. The first method is directly integrated into the XGBoost library, while the other two are XGBfir [7] and Shapley [25]. The objective of this research work is to apply these three tools in the case of regressions by type of intervention of the firemen, and to draw from it on the one hand business knowledge, and on the other hand the way to make good predictions in this framework, with state of the art tools and in an operational context.

The contributions of this article are threefold. First, we explain in a detailed way how to set up, at an operational level, a predictive model allowing to accurately predict the number of firefighters interventions, by hour and by type of intervention. We then qualitatively compare three state of the art methods to detect a posteriori the importance of explanatory variables, by comparing what they produce with business knowledge and expectations, when they are obvious (e.g., river water level for flooding type interventions). Finally, and conversely, we use these three tools to extract or validate business knowledge for various types of interventions, such as the determination of features impacting childbirth outcomes.

The remainder of this article is organized as follows. In the next section, related works in the field of intervention prediction are presented. The design of the basis of knowledge is detailed in Section 3. Experimental protocol and obtained results are then presented and discussed in Section 4. This article ends by a conclusion section, in which the contributions are summarized and intended future work is outlined.

2 Related work

Among the works reviewed and related to the optimization of fire departments responses to incidents, we mainly found contributions for the prediction of interventions [13, 30] and fires [27].

Fernandes, P.A.M predicted the fire spread in a flat terrain in Shrubland in Portugal on a series of experiments and prescribed fires in four different shrub fuel types considering weather, fuel conditions and fire spread rates up to 20 minutes [26]. Moreover, in their work, Pirklbauer, K. and Findling, R.D. proposed an approach for predicting the fire departments' deployment category based on time, weather, and location information. The performance of the model using multiple machine learning algorithms was then compared [31]. Lian, X. et al. applied distributed computing and machine learning algorithms (Linear Regression, Decision Tree Regression, and Random Forest Regression) to predict the emergency response time for San Francisco Fire Department [23].

Furthermore, Bradstock, R.A. et al. explored large fire ignition days probability in Sydney, Australia, using a Bayesian logistic regression influenced by the ambient and drought weather components of the Forest Fire Danger Index [10]. Similarly, Coffield, S.R. et al. used decision trees to classify the final size of fire at the time of ignition in Alaskan boreal forests into small, medium, and large. Fang, H. et al. implemented a machine learning based approach to identify automatically the stages of fire development in residential fires from a collection of fireground information using Gaussian Mixture Models and Hidden Markov Models [20]. In addition, O'Connor et al. arised a boosted logistic regression model to classify final fire locations using a dataset that includes topographic features, fuel types, and natural barriers to fire spread in southern Idaho and northern Nevada [29].

Likewise, in certain parts of the world firefighters are part of emergency medical service, since they also provide ambulances. In this way, predictions of traffic accidents [33], ambulance response time and resource allocation are also included [11, 15, 18]. Furthermore, we can find works related to the predictions of rare events such as earthquakes [32] and hurricanes [24], which would allow firefighters to identify a specific location for damage assessment and develop better strategies when succouring the population.

Additionally, few studies on time series forecasting using skills in Artificial Intelligence have considered the topic of our research about fire department operations. Nahuis, S.L.C. et al. used Long Short-Term Memory and demonstrated the possibility to build a neural network from scratch and the ability to predict the number of deployments in 2017 from those from 2012 to 2016 [28]. Couchot, J.F. et al. [17] worked on a learning process based on real but anonymous data and conducted the study using the Extreme Gradient Boosting technique. In [22], Guyeux C. et al. applied an ad-hoc Multi-Layer Perceptron in which hyper-parameters were selected using a supercomputer, and their work gave a reasonable prediction of firefighters' operations. Furthermore, S. Cerna et al. [12] compared between XGBoost, Gradient Boosting, and AdaBoost, as these techniques are considered very effective in modeling nonlinear systems. XGBoost has appeared as the best predictive model for firemen interventions. Finally, Arcolezi, H.H. et al. [9] focused on local-differential privacy-based data in their study. Their approach was to predict the number of firefighters in specific locations by applying differential privacy along with XGBoost techniques.

None of these articles have investigated the basis of knowledge, in terms of predictive power and operational character.

3 Database design

Let's start by presenting in detail how we built our knowledge base, its features and its targets.

3.1 The targets

The listing produced by the Doubs fire department consists of an Excel file containing 250,344 interventions from January 1st 2015 to December 31st 2021. This corresponds to 35,800 interventions per year, or a little more than 4 interventions per hour. Note that the number of interventions has been steadily increasing over these 7 years, for various reasons related to population growth, aging of the population, closure of small hospitals, and Covid-19. For each intervention, we have its unique identifier, the start time, the end time, the location of the intervention, and finally the reason for discharge.

The original set contained 520 different exit reasons, with often overlapping headings, misspellings, reasons that were too specific and never used, or conversely with headings that were too crude to be useful. For example, there were 17 different discharge reasons for "attempted suicide," including headings such as "attempted suicide by gas," "threatened defenestration," and "attempted suicide by defenestration." However, while the ways of doing this differ, the cause of the discharge is always the same, namely attempted suicide. It has been known since Durkheim that suicides are partly explained by social factors and have a seasonal character [19], which makes it possible to hope for some prediction of these cases, at least at the margin. However, it is illusory to imagine being able to predict the victims' choices, the way they commit suicide, which can hardly be explained by external variables. Moreover, dividing this reason for exit into 17 subsets leads to small knowledge bases, on which it is not possible to do quality machine learning due to lack of cases.

Such a finding can be made, at least to a lesser extent, for the 5 reasons related to childbirth, or for the 29 reasons for exits related to heating fires. While it is obvious that the latter have a certain predictable character (they occur, for example, in winter), going as far as the type of dwelling in the title of the reason (house, semi-detached house, etc.) does not make sense for the purpose we propose. Therefore, we grouped the entire set of exit reasons into 14 coherent, non-empty intersection subsets.

These 14 families of reasons are respectively: fires, road accidents, suicides, drownings, deliveries, fires on public roads, floods, fires due to heating, emer-

Brawl or violence with a firearm,
Brawl or violence with a firearm or a knife,
Brawl or violence with a knife,
Brawl or violence without a weapon,
Wound by firearm,
Wounding with a knife.

Table 1. Outings for brawl-type reasons

Destruction of hymenopteres,
Destruction of hymenopteres with payment,
Destruction of insects,
Waiting for a beekeeper,
Waiting beekeeper.

Table 2. Outings for wasp-type reasons

gency personal assistance, non-emergency personal assistance, accidents on public roads, brawls, discomfort, and interventions due to wasps and other hymenoptera. The reasons for these overlapping groupings are either the consistency of the component headings (see Tables 1 and 2 for the content of the smallest subsets) or operational importance (such as emergency personal assistance).

These 14 subsets of exit reasons will lead to 14 independent regressions, in which we will try to determine the importance of the features described below on the evolution of these targets. More precisely, for each intervention carried out by the Doubs firefighters, we know the reason for leaving, the day and time of departure for the intervention, and those of the return to the fire station. We are therefore able to establish, for each time slot of the 2015-2021 period, the number of interventions per hour and per type. This is exactly the number we will be trying to predict.

3.2 The features

The second step consists of listing the potential causalities that can, at least in part, explain the number of interventions associated with each of the 14 families of types previously introduced. These features should not only have an impact on at least one type of intervention, but they should also be accessible (by calculation, on the internet...), otherwise they would have no practical value.

The most directly accessible and potentially most impactful features are those of the calendar or ephemeris type. Indeed, the main activity of the fire department in France is the rescue of people, emergency or not. These accidents are directly proportional to the human activity, which is much more important during the day than at night. It also changes in configuration between weekdays and weekends, leading to two seasonal patterns (daily and weekly) with an obvious predictive character. The daily profile is also clear (see Fig. 1), with a continuous increase in the number of interventions during the day, except for a dip during the lunch break, followed by a decrease in the early evening. This decrease, in turn, occurs later in the week-end.

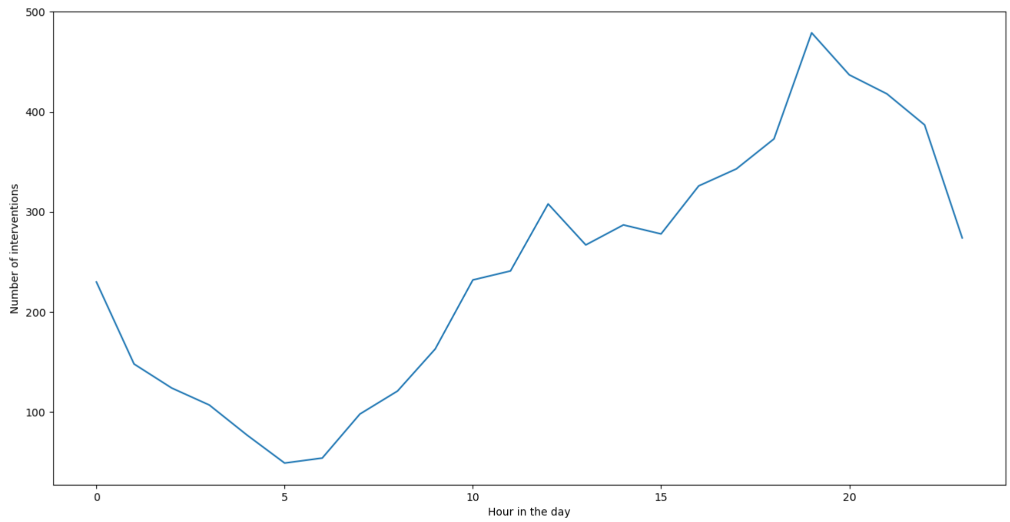

Fig. 1. Number of interventions per hour

This dependence on the time of day and day of the week is not unique to emergency and non-emergency rescue targets. It also occurs in most other types of human activity. Thus, suicides are rare in the middle of the night on weekdays, but are relatively frequent on weekends, in the evening. In the same way, it is rare to drown at 3 o'clock in the morning, and one can affirm in the same way that the road accidents and the brawls have appreciably the same daily profile, because they are also directly related to the human activity. Conversely, although childbirth is part of life-saving assistance, there is no reason a priori to assume a daily or weekly seasonality: there is no reason why there should be more childbirths on Saturdays than, say, on Thursdays. And at the opposite end of the spectrum from these types with a seasonality derived from the rhythm of human activity, we find floods, for which there is no reason to have a daily or weekly profile.

Other calendar or ephemeral factors may impact the number of interventions per type and per hour. Holidays and vacations have a clear impact on first aid, but also on road accidents and drownings: it is not when children go to school that they drown in their private pool. In the same way, and in a coupled way, the period in the year (for example, the month) impacts the number of interventions: chimney fires occur in winter when drownings occur in summer. Wasps are not present all year round. The beginning and end of Ramadan is a time of celebration, when this period is prone to discomfort. The sun does not rise and set at the same time throughout the year, which modifies human activity and the resulting accidents. For example, lumberjacks only work when it is daylight, and they can only have accidents during those times. Finally, the phase of the moon has been added, because there is a belief in the world of firefighters that it has an impact on deliveries: this belief needs to be questioned.

One of the most easily accessible reasons that comes next, and likely to have an impact on some of these targets is meteorology. On the MétéoFrance website [4], we have access to the following data, renewed every 3 hours for about fifty stations spread out in France: temperature, pressure, pressure variation, barometric tendency, humidity, dew point, quantity of rain fallen during the last hour, during the last three hours, wind direction and speed during the last 10 minutes, and horizontal visibility. It is obvious that weather has an impact on road accidents, whether it is snow or ice, and that motorcyclists skid when it rains. Conversely, traffic is less dense than in good weather. Personal assistance is also impacted: for example, icy conditions also lead to pedestrian falls, and seasonal flu and gastroenteritis epidemics occur when it is cold. Finally, these variables are correlated with most of the calendar and ephemeris data previously mentioned, as for example the temperature is lower at night than during the day.

These quantitative data on the weather can be completed from the bulletins of vigilance provided by MétéoFrance [6], which for a collection of risks (high winds, rain and floods, thunderstorms, snow and ice ...) and for a given period, assigns a color code of risk: green, orange and red These data are obviously correlated to the previous ones, but they have the double interest of being more qualitative than quantitative, and of being the result of a human expertise based on the raw data (it is thus enriched data).

A second category of variables that are easily accessible and distinct from meteorological data consists of the epidemiological information produced by the Sentinel network [5]. Through this site, it is possible to access the weekly incidence of influenza syndromes, acute diarrhea, chickenpox, and since 2020, acute respiratory infections. These illnesses account for a certain share of personal assistance during the periods of the year when these epidemics are rife. But the weekly frequency is far too coarse-grained compared to our ambition to make hourly predictions, and to the dynamics and rapid evolution of the number of cases during epidemics. Therefore, the search for more operational indicators led us to turn to regional google search trends [2]. These regional google trends are interesting in many ways. First of all, they are hourly trends, and they allow to anticipate the appearance of epidemic peaks. But they are also a way to continuously monitor the appearance of specific, non-seasonal events such as demonstrations, sales, championships, urban violence, attacks, industrial accidents, heat waves, festivals... all of which have a probable impact on different families of interventions.

These features are complemented by the following data, which can be retrieved from the internet, and which have a high probability of conditioning the number of interventions for certain types. First of all, we have access to water levels and flows at about twenty points distributed over all the rivers in our region, and these data produced by Hydroréel [3] are directly linked to flooding type interventions. In the same way, we can know the air quality in different urban places thanks to the site of the association Atmo-BFC [1], which has a rich network of stations allowing to measure for example the PM10 and PM2.5 particles, the ozone, the sulfur dioxide... and these data probably have an impact on the discharges for malaise.

The above features have the double advantage of presenting a sufficiently long history at the right frequency (mainly hourly, or every 3 hours), and also providing a real-time measurement and display. The history makes learning possible, while the real time measurements then allow continuous prediction, and thus have an operational utility. But the deployment, the industrialization of such a solution is complex if based on all these features, and such complexity would not be justified if such features would bring little more than a simple time series approach, basing its predictions only on past interventions, and trends and seasonality detected in the series of interventions, in the absence of any consideration of external features. Also, the features detailed in this section will be "competed" with the number of interventions at time $t-1$, $t-2$..., with rolling averages or simple linear regressions, in short with a predictive model considering only the time series (ARIMA, etc.)

4 Experimental results

4.1 Experimental protocol

A four-week history was incorporated into the features, leading to the addition of 672 variables. The knowledge base was then separated into a test set (20% of the data, randomly drawn) and the remaining 80% was divided into a training set (80%, randomly drawn) and a validation set. The quantitative variables were standardized, while the qualitative variables were target encoded using the category encoders library provided by Will McGinnis.

A learning was then performed with the Poisson regressor of XGBoost [16], with a maximum tree depth for base learner fixed at 6. This gradient boosting library was chosen following the comparative study published in [13,14], and the value of the hyperparameters was set as advised, with the max_depth variable being the most important to choose. The training stop was determined by an early stopping rounds criterion set at 10 (on the evaluation set). This regressor was then used through three different methods to determine the importance of each feature on the prediction quality. These three methods are respectively :

1. The feature_importance method from XGBoost [16, 21]. This omportance is computed for a single decision tree by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. The feature importances are then averaged across all of the the decision trees within the model.
2. The TreeExplainer method from the SHAP library a game theoretic approach to explain the output of any machine learning model. As stated in

their website: “SHAP connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions” [25].

3. The use of Xgbfir provided by Facebook [7], a XGBoost model dump parser, which ranks features as well as feature interactions by different metrics.

4.2 Obtained results

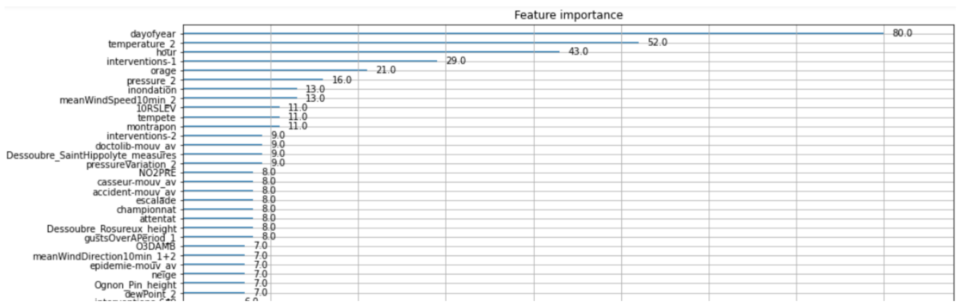

Fig. 2. XGBoost feature imporance for flooding events

The case of flooding in detail Our first focus was on flooding. As can be seen in 2, the most important feature according to XGBoost is the Day of year. This is not surprising as flooding events only occur at certain times of the year (there is a strong annual seasonality). Weather-type features are also obviously important, accounting for 3 of the 10 most important variables. Somewhat strangely, time of day comes third, which can be explained by the fact that rescuing people or vehicles when the water is overflowing implies that there is human activity: such rescues are more likely to occur during the day than at night. Interventions that took place in the last hour come next, which is easily explained by the fact that these types of interventions are generally quite long. Then, in fifth and seventh position, we find reassuringly the local google searches corresponding to the terms ”storm” and ”flood”. Finally, we see appearing in these 30 most important features, 3 measures of flow or height of watercourse. This makes sense, even though one might have expected a greater representativeness of these hydrolic stations. Finally, we quickly see features appear that have nothing to do with flooding interventions, which shows that, if XGBoost is a very good regression tool, it is not very good for feature selection.

These elements are to a large extent reflected in the importance of features according to Shapley, cf. Fig 4. We find indeed the strong importance of the day in the year, the hour, and the interventions in the past hour, followed to a lesser extent by the google searches for ”storm” and ”storm”. Two of the twenty features selected are about weather data, and two others about river water levels. We also find, as before, a certain number of less important features, whose link

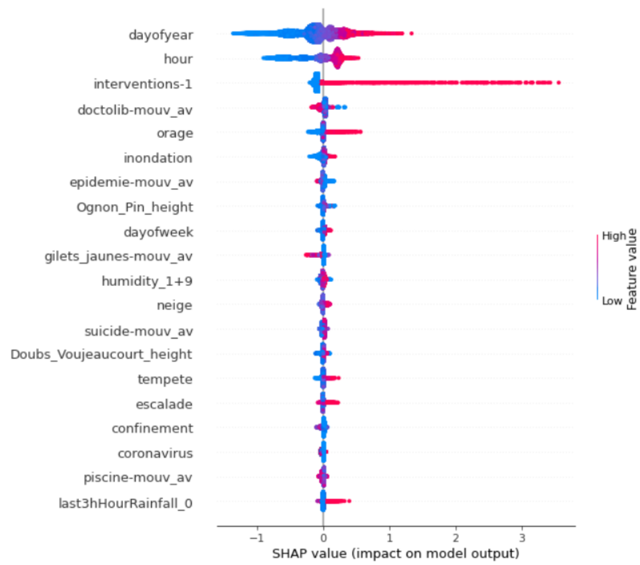

Fig. 3. Shapley feature imporance for flooding events

with floods is not obvious, but can in some cases still be justified. First of all, since the floods take place in winter in the Doubs, we understand that these events occur at the same time as the seasonal flu. And therefore, when there are floods, these are periods when Internet users are looking for doctolib, epidemic, or snow. Conversely, in winter, people don't swim in outdoor pools or go climbing, hence the negative impact of searches for these keywords on the occurrence of floods. The coronavirus pandemic, with its periods of confinement and curfews, implies that there are far fewer people outside, and therefore fewer people to rescue in case of flooding. The remaining variables (suicide...) cannot be explained.

Fig. 4. Xgbfir feature imporance for flooding events

Xgbfir, on the other hand, mixes some interesting results with some disappointing ones. First of all, we are surprised by the strong importance given to features related to air quality (O3DAMB for the Ozone measurements at Dambière...), to the point of being found in 3 of the 4 most important features. On the other hand, we see a good representation of the weather variables, with co-occurrences that make sense, such as lastHourRainfall_0—lastHourRainfall_0+9, corresponding to the amount of water that fell during the last hour and nine hours ago, for the weather station "0", i.e. the closest one. This co-occurrence can be translated as "it rained a lot, and for a long time". The co-occurrence humidity_1+8—storming also makes sense, for similar reasons, when Loue_Vuillafans_heigth—interventions-8 can be understood as: the water is rising at the Loue Vuillafans station, and there were many interventions 8 hours ago. We can think that a certain number of these interventions are still going on, and that they are related to flooding.

From the above, a number of lessons can be drawn. First, all three methods produce sensible results, but with noise. Therefore, one should not use only one technique, but all three, and cross-check the information by looking for points of convergence. Second, some variables are clearly strongly correlated with the number of flood interventions, making a predictive model possible. If some of these features are natural, this approach also makes it possible to highlight other features that we do not necessarily think about. These methods also allow us to know which indicators to monitor in order to measure the risk of flooding, and which features to choose for learning followed by predictions. Finally, they allow us to assess how many variables of each type should be considered, showing for example that it is not necessary to integrate many weather features, nor many river heights, probably because within each of these groups of variables, there are many correlations, and considering all of them would introduce a useless or even harmful redundancy.

The other types of intervention in summary We applied this method to all types of interventions. However, due to space constraints, only some of the points of interest will be discussed below.

At the level of deliveries, the day in the year stands out in first position for XGBoost and XGBfir. It is known that some periods (e.g., New Year's Eve) are more favorable than others (e.g., Lent and Ramadan) for conception, and that births are not evenly distributed throughout the year. Weather conditions also stand out, but to a lesser extent, which may be explained by the fact that weather is correlated with the day of the year. Covid is also a feature to consider, and it is known that this pandemic also had an impact on births. On the other hand, and contrary to popular belief, the phases of the Moon have no impact on interventions for childbirth. The reasons put forward to justify this false link are generally that we see the effect of the moon on the tides, and that there must be such an effect on the pregnant woman's belly because the amniotic fluid would be sensitive to this attraction. As we can see, this is not the case. Moreover, the tides are very weak on the Mediterranean Sea and null on the lakes, because a great volume is necessary for there to be effect. On the other hand, the variation of the attraction of the Moon is not related to its phase, but to its variable distance from the center of the Earth: it is not when there is a full moon that the attraction is stronger.

With regard to all types of fires, we note first of all the importance of the day of the year and the weather conditions: there are more fires when it is hot and dry than when it is wet and cold. Then the time of day: fires with a human origin occur during the active period of the human being (daytime). Since a fire takes time to be extinguished, nearby history variables are also important. Finally, these fires affect air quality variables, which could be a good way to detect fire starts earlier. Again, XGBfir is less trustworthy than the other two variable selection tools.

If we focus on fires on public roads (cars catching fire, etc.), the importance of the features changes relatively. Air quality is less important, certainly because these fires are much more contained than forest fires. The variables of day of the year and time of day remain very important, due to the seasonality of human activity and weather conditions. The keyword "accident" appears in the 6 most important features for XGBoost and also appears for Shapley, which would also give us a head start for the intervention: if we see this google search becoming more important, we should expect to receive calls reporting an accident with fire very quickly. It should be noted that the interventions for road accidents have approximately the same profile of important features: accidents with fire are proportional to the accidents as a whole, and they have approximately the same causes.

As far as the risk of chimney fires is concerned, we can see the importance of the time of year (winter), and the time of day is also more important than in previous cases of fire. This is understandable because, for example, in the middle of the night, fires go out, and the risk of fire with it. The weather variables are also important, because we heat when it is cold, and the air quality variables reappear, showing the pollution of chimney fires. Finally, some keywords appear in the searches of the Internet users, such as "snow" and "fireplace".

In terms of suicides, there is mainly an annual and weekly seasonality [19,35]. Thus, in the Doubs, there is a greater tendency to commit suicide on weekends in the early evening, when the summer solstice is reputed to be the day with the most suicides in the year [34]. The weather also has a slight impact, as does the internet search for "suicide" [8]. As for drownings, they occur during the day when the weather is good, hence the importance of the features of time of day, day of year, temperature and rain.

In terms of discomfort, the importance of time of day and day of year, temperature and pressure, air quality, and "doctolib" research appear. The important features for emergency and non-emergency rescue and for accidents on public roads can be deduced directly from the above, without bringing much new information. On the contrary, the importance tends to be smoothed out, blurring the information contained in more specific types of intervention. Finally, nothing obvious emerges for brawls, which is not surprising, nor for the risk related to hymenoptera, which may seem a priori surprising (one would think that google searches for "bee" or "guepes" would be important). In fact, these interventions have dropped since the fire department made them chargeable.

5 Conclusion

In this paper, we investigated the quality of three feature importance measures and their ability to explain the target variable. 14 targets were considered in the context of firefighters' interventions, related either to climatic incidents or to human activity. In each case, a regression with a large number of features and XGBoost was performed, then Shapley, XGBfir and XGBoost feature importance were applied. We showed that it was better to consider these three tools together, which allowed to better separate the important features from the noise. We also saw that XGBfir produced overall worse results. The analysis of the features that emerge allows us to extract business knowledge. They also allow to reduce the learning time per type of intervention, by drastically reducing the features to the strict necessary ones.

We would like to extend the number of feature selection tools, and to deduce the best subset to consider. We will also look at the extent to which the choice of hyperparameters, or even the regression method, impacts the importance of the features and the explainability resulting from the learning. Finally, we will try to extend the study to other departments, with other forms of risk: Cevennes events, industrial risk, coastal risk, urban risk, etc.

References

1. ATMO-BFC, howpublished = https://www.atmo-bfc.org/accueil. Accessed: 2022-01-10.
2. Google trends. https://trends.google.fr/trends/?geo=FR.
3. Hydroréel, howpublished = https://www.rdbrmc.com/hydroreel2/index.php. Accessed: 2022-01-10.
4. Météo-France public data. https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32. Accessed: 2021-02-01.
5. Réseau sentinelles. https://www.sentiweb.fr/france/fr/? Accessed: 2022-01-10.
6. Vigilance Météo-France public data. https://vigilance.meteofrance.fr/fr. Accessed: 2022-01-10.
7. Xgbfir. https://github.com/limexp/xgbfir. Accessed: 2022-01-10.
8. Türkan Akkaya-Kalayci, Benjamin Vyssoki, Dietmar Winkler, Matthaeus Willeit, Nestor D Kapusta, Georg Dorffner, and Zeliha Özlü-Erkilic. The effect of seasonal changes and climatic factors on suicide attempts of young people. *BMC psychiatry*, 17(1):1–7, 2017.
9. Heber H Arcolezi, Jean-Francois Couchot, Selene Cerna, Christophe Guyeux, Guillaume Royer, Bechara Al Bouna, and Xiaokui Xiao. Forecasting the number of firefighter interventions per region with local-differential-privacy-based data. *Computers & Security*, 96:101888, 2020.
10. Ross A. Bradstock, Janet S. Cohn, A. M. Gill, Michael Bedward, and Christopher Lucas. Prediction of the probability of large fires in the sydney region of south-eastern australia using components of fire weather. 2017.

11. A.S. Carvalho, M.E. Captivo, and I. Marques. Integrating the ambulance dispatching and relocation problems to maximize system's preparedness. *European Journal of Operational Research*, 283(3):1064–1080, June 2020.
12. Selene Cerna, Christophe Guyeux, Héber H Arcolezi, Raphaël Couturier, and Guillaume Royer. A comparison of lstm and xgboost for predicting firemen interventions. In *World Conference on Information Systems and Technologies*, pages 424–434. Springer, 2020.
13. Selene Cerna, Christophe Guyeux, Héber H. Arcolezi, Raphaël Couturier, and Guillaume Royer. A comparison of LSTM and XGBoost for predicting firemen interventions. In *Trends and Innovations in Information Systems and Technologies. WorldCIST 2020*, volume 1160, pages 424–434. Springer International Publishing, 2020.
14. Selene Cerna, Christophe Guyeux, Guillaume Royer, Céline Chevallier, and Guillaume Plumerel. Predicting fire brigades operational breakdowns: A real case study. *Mathematics*, 8(8):1383, August 2020.
15. Albert Y. Chen, Tsung-Yu Lu, Matthew Huei-Ming Ma, and Wei-Zen Sun. Demand forecast using data analytics for the preallocation of ambulances. *IEEE Journal of Biomedical and Health Informatics*, 20(4):1178–1187, July 2016.
16. Tianqi Chen and Carlos Guestrin. Xgboost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug 2016.
17. Jean-François Couchot, Christophe Guyeux, and Guillaume Royer. Anonymously forecasting the number and nature of firefighting operations. In *Proceedings of the 23rd International Database Applications & Engineering Symposium*, pages 1–8, 2019.
18. Idalia Flores de la Mota, Esther Segura Perez, and Alexander Vindel Garduno. Optimization and simulation of an ambulance location problem. In *2017 Winter Simulation Conference (WSC)*. IEEE, December 2017.
19. Emile Durkheim. Suicide: A study in sociology (ja spaulding & g. simpson, trans.). *Glencoe, IL: Free Press.(Original work published 1897)*, 1951.
20. Hongqiang Fang, S.M. Lo, Yunjie Zhang, and Yixin Shen. Development of a machine-learning approach for identifying the stages of fire development in residential room fires. *Fire Safety Journal*, 126:103469, 2021.
21. James Franklin. The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85, 2005.
22. Christophe Guyeux, Jean-Marc Nicod, Christophe Varnier, Zeina Al Masry, Nourredine Zerhouny, Nabil Omri, and Guillaume Royer. Firemen prediction by using neural networks: A real case study. In *Proceedings of SAI Intelligent Systems Conference*, pages 541–552. Springer, 2019.
23. Xu Lian, Sarah Melancon, Jon-Ross Presta, Adam Reevesman, Brian Spiering, and Diane Woodbridge. Scalable real-time prediction and analysis of san francisco fire department response times. In *2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 694–699, 2019.
24. Xiaoyu Sean Lu, MengChu Zhou, and Liang Qi. Analyzing temporal-spatial evolution of rare events by using social media data. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, October 2017.
25. Scott M. Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. From

local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, 2(1):2522–5839, 2020.

26. Paulo A. Martins Fernandes. Fire spread prediction in shrub fuels in portugal. *Forest Ecology and Management*, 144(1):67–74, 2001.
27. Thiago Fonseca Morello, Rossano Marchetti Ramos, Liana O. Anderson, Nathan Owen, Thais Michele Rosan, and Lara Steil. Predicting fires for policy making: Improving accuracy of fire brigade allocation in the brazilian amazon. *Ecological Economics*, 169:106501, March 2020.
28. Selene Leya Cerna Nahuis, Christophe Guyeux, Héber Hwang Arcolezi, Raphaël Couturier, Guillaume Royer, and Anna Diva Plasencia Lotufo. Long short-term memory for predicting firemen interventions. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 1132–1137. Ieee, 2019.
29. Christopher D O'Connor, David E Calkin, and Matthew P Thompson. An empirical machine learning method for predicting potential fire control locations for pre-fire planning and operational fire management. *International journal of wildland fire*, 26(7):587–597, 2017.
30. Kevin Pirklbauer and Rainhard Dieter Findling. Predicting the category of fire department operations. In *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*. ACM, December 2019.
31. Kevin Pirklbauer and Rainhard Dieter Findling. Predicting the category of fire department operations. In *Proceedings of the 21st International Conference on Information Integration and Web-Based Applications Services*, iiWAS2019, page 659–663, New York, NY, USA, 2019. Association for Computing Machinery.
32. Risul Islam Rasel, Nasrin Sultana, G.M. Azharul Islam, Mahfuzul Islam, and Phayung Meesad. Spatio-temporal seismic data analysis for predicting earthquake: Bangladesh perspective. In *2019 Research, Invention, and Innovation Congress (RI2C)*. IEEE, December 2019.
33. S. Hari Sankar, K. Jayadev, B. Suraj, and P. Aparna. A comprehensive solution to road traffic accident detection and ambulance management. In *2016 International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEES)*. IEEE, November 2016.
34. Cornelis AJ Van Houwelingen and Domien GM Beersma. Seasonal changes in 24-h patterns of suicide rates: a study on train suicides in the netherlands. *Journal of affective disorders*, 66(2-3):215–223, 2001.
35. J. Yu, D. Yang, Y. Kim, M. Hashizume, A. Gasparrini, B. Armstrong, Y. Honda, A. Tobias, F. Sera, A. M. Vicedo-Cabrera, and et al. Seasonality of suicide: a multi-country multi-community observational study. *Epidemiology and Psychiatric Sciences*, 29:e163, 2020.