

# Mining Toronto Fire Services Incident Data

By Geoffrey Clark

July 28th, 2018

Completed for Dr. Ceni Babaoglu

CKME136: Data Analytics Capstone Course

Data Science Lab, Ryerson University

## ABSTRACT

The goal of this project was to explore viability of 'classical' machine learning algorithms on a fire incident data set. An open source data set on Toronto Fire Services Incidents was obtained, cleaned and the subject of some exploratory analysis. A binary class attribute was assigned to each incident based on post-incident features such as damage and number of responding units. Three machine learning algorithms—Logistic Regression, Naive Bayes and Random Forest—were then used to predict the class attribute using features whose values would be known at the time of incident call. The results are presented and discussed.

## Overview

- Introduction
- Literature Review
- Dataset
- Approach
- Exploratory Analysis
- Modeling
- Results
- Conclusions

## Introduction

Emergency Services in general, and the Fire Brigade in particular, are in constant demand and often in situations where they find themselves at risk. Any additional information that can be provided to the Fire Services at the time of an incident may prove invaluable in improving performance and success. Computer Aided Dispatch systems, already in use (Bebbee, D. 2012), provide a huge opportunity: both in their logging of data and their integration within the emergency services systems. An effective

machine learning model, appropriately trained, may be able to provide first responders with insights about the incidents that they are responding to before they arrive and thus help reduce cost, time and risk (Asgary, A., & Sadeghi-Naini, A. 2012). This project explores the viability of such a model and compares three well known machine learning algorithms: Logistic Regression, Naive Bayes, and Random Forest.

## Literature Review

There is a wide array of fire incident research and evidence of agreement to its inherent value (Sekizawa, A., 2012). Research exists on training Neural Networks to predict fire incident characteristics such as risk (A., Naini, AS., & Levy, J. 2012), number of firefighters responding (Asgary, A., & Sadeghi-Naini, A., 2013), loss and no-loss fire incidents (Asgary, A., Sadeghi-Naini, A., & Kong, A., 2009), and critical incidents (Asgary, A., Sadeghi-Naini, A., & Levy, J., 2009). These approaches were all relatively successful in their goals, however, complex models such as Deep Learning and Artificial Neural Networks often have associated costs such as additional complexity and may be more difficult to interpret as a result (G. Casella, S. Fienberg, I. Olkin, 2013, p. 25). It is prudent to explore traditional methods before adopting a more complex solution while maintaining a constant awareness of the trade-offs and costs associated with the more complex solution.

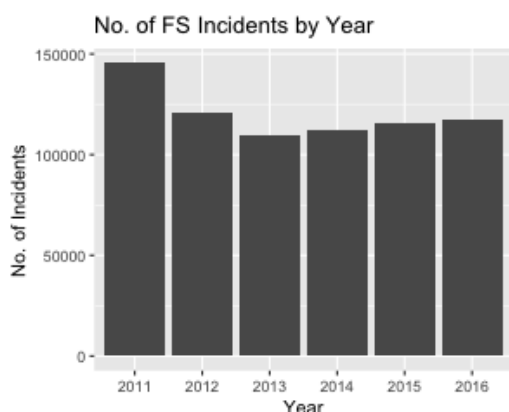
Fire services incident data has been the subject of other types of research also, including spacial and temporal analysis in both Toronto (Asgary, A., Ghaffari, A., Levy, J., 2009) and Ankara (Ceyhan, E., Ertügay, K., & Düzgün, Ş., 2013). International scope is not limited to Japan, Canada and Turkey, however, as research has been published on the implementation of fire incident reporting systems and their value in the United States as well (Bergen, G., Frattaroli, S., Ballesteros M.F., Ta, V.M., Beach, C., & Gielen, A.C., 2008).

## Dataset

The dataset used in this project was concurrently a great resource and large challenge. Retrieved in May, 2018, from the Toronto Open Data Catalogue the dataset contains 720,370 records of Toronto Fire Services Incidents from the years 2011 through 2016. Up to 100 features are provided from multiple

sources including the Computer Aided Dispatch System and Office of the Fire Marshall Incident Reports.

Each observation in the data set corresponds to an incident such as fire, medical call, or police assist and has a corresponding unique identifier. Some features, such as incident number, event type and alarm source appear to have been available at the time of incident call and are most likely inputted by the Computer Aided Dispatch (CAD) system.



**Figure 1:** A yearly breakdown of incident quantities

This characteristic of being inputted automatically has led to these features being relatively reliable and complete compared to others in the dataset.

| Feature                              | Description   |
|--------------------------------------|---|
| EVENT_TYPE                           | Type of event: Medical, fire, etc                     |
| INCIDENT_DATE                        | Date of initial call                                  |
| RESPONDING_UNITS                     | Number of units who responded (total)                 |
| RESPONSE_TYPE                        | Type of response: Rescue, medical, etc                |
| ALARM_TO_FD                          | Source of call  |
| EST_KM                               | Estimated Distance to Incident                        |
| INITIAL_UNIT_PERSONNEL               | Number of personnel responding on first apparatus     |
| EST_LOSS                             | Estimated loss due to incident                        |
| EST_VALUE_AT_RISK                    | Estimated value at risk                               |
| FF_INJURIES                          | Firefighter Injuries                                  |
| FF_FATALITIES                        | Firefighter Fatalities                                |
| CIVILIAN_INJURIES                    | Civilian Injuries                                     |
| CIVILIAN_FATALITIES                  | Civilian Fatalities                                   |
| RESCUES                              | Number of rescues                                     |
| OFM_INVESTIGATIONS_CONTACTED         | Office of Fire Marshall Investigations unit contacted |
| INCIDENT_DAY, MONTH, HOUR, MIN, YEAR | Extracted from INCIDENT_DATE                          |

**Table 1:** Some features used in this project, with descriptions. the first year of the dataset, 2011.

Also present in the dataset are features that contain values which would be entered, most likely manually, following an incident and include such features as estimated damage, number of firefighter and civilian injuries and casualties, number of rescues, and whether or not the Office of the Fire Marshall (OFM) Investigations Department was contacted. The features used in the analysis and modeling sections of this project were chosen largely for their completeness time of input: before or after the incident. Some features, such as highest alarm level reached and which fire department station dispatched, were dropped due to lack of availability or obfuscated values.

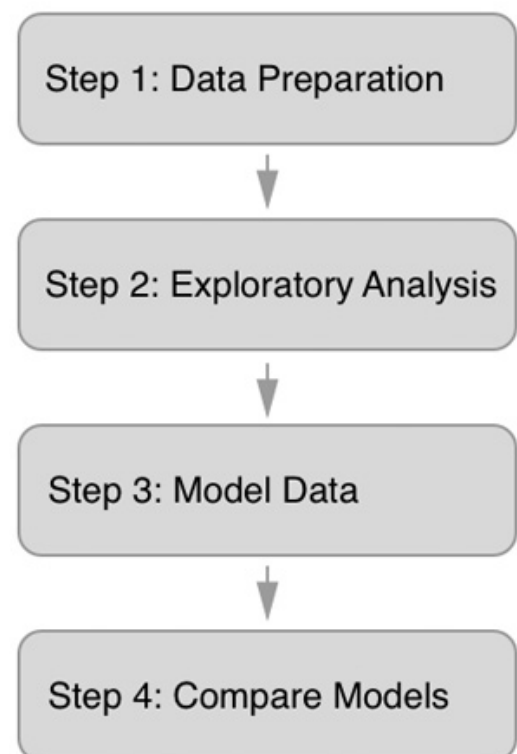
The breakdown of number of incidents by year in **Figure 1** reveals a relatively constant trend of incidents with a slightly higher amount in

## Approach

The approach of this project was to first import, clean and prepare the dataset, perform an exploratory analysis, create a model to predict a dependent variable of choice, and lastly to evaluate the performance. A graphical outline is provided in **Figure 2**.

### Step 1: Data Preparation

Published as separate XML files the data was initially in a difficult structure for analysis and thus converted into tabular structure with each row representing an observation (incident) and each column representing an attribute (feature). Following conversion each feature was inspected quantitatively, for NA values, summary statistics, and qualitatively, for source, description, and time of input relative to the incident.



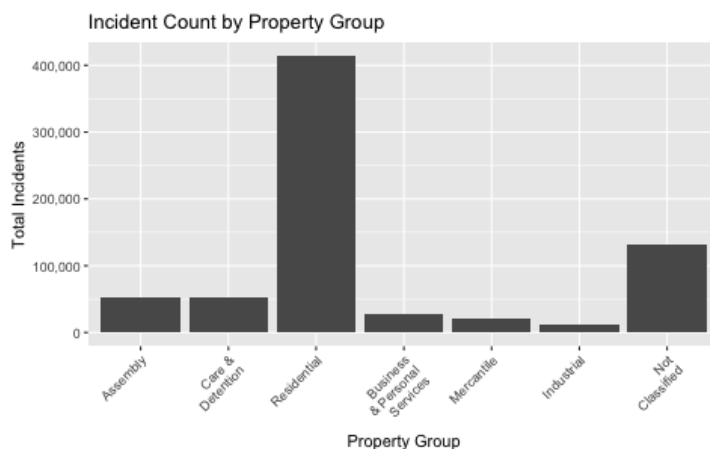
**Figure 2:** A block diagram of this project.

## Step 2: Exploratory Analysis

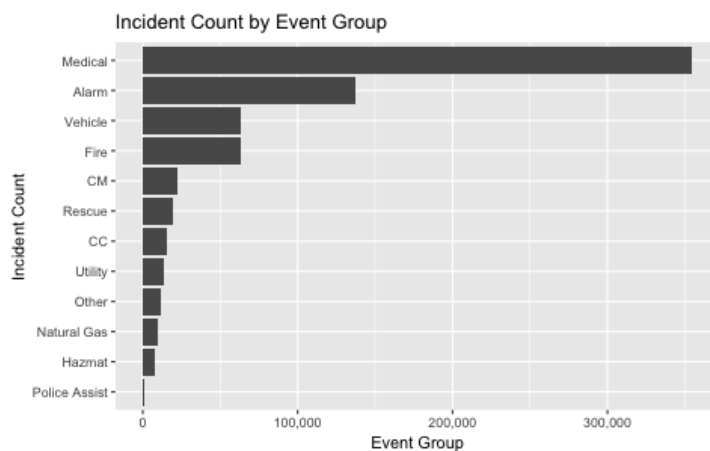
Following the conversion and initial preparation stage an exploratory analysis is conducted where features and relationships are examined in more detail. The primary concern of this step is to improve understanding of the data. Univariate, bivariate and multivariate relationships among features are examined. It is during this exploratory analysis that features are examined for their suspected relevance in defining or predicting the class variable.

## Step 3: Model Data

Following the initial analysis three models are used to predict the class attribute: Logistic Regression, Naive Bayes and Random Forest. The models are first used to classify the class attribute on a subset containing only fire incidents. During this process the models are trained on 10 randomly selected training sets and then tested on the corresponding test sets. This procedure, similar to k-fold cross validation, is to determine the presence of high-variance (overfitting) in the model. After the models are run on the fire incident subset they are then trained and run on the entire data set in order to assess generalizability.



**Figure 3:** Total incident count for each property type (grouped for similarity): Residential accounts for over 50% of all incidents



**Figure 4:** Total incident account for each event type (grouped for similarity): Medical accounts for nearly 50% of all incidents.

## Step 4: Compare Models

After training and running the models the results are accumulated and compared using four metrics: Accuracy, Recall, Precision and F<sub>1</sub>-Score. F<sub>1</sub>-Score is weighted heavier than accuracy due to the former's tendency to be low in the absence of any True Positives, a situation that can result in the presence of a strong class imbalance.

## Exploratory Analysis

### Event & Property Type

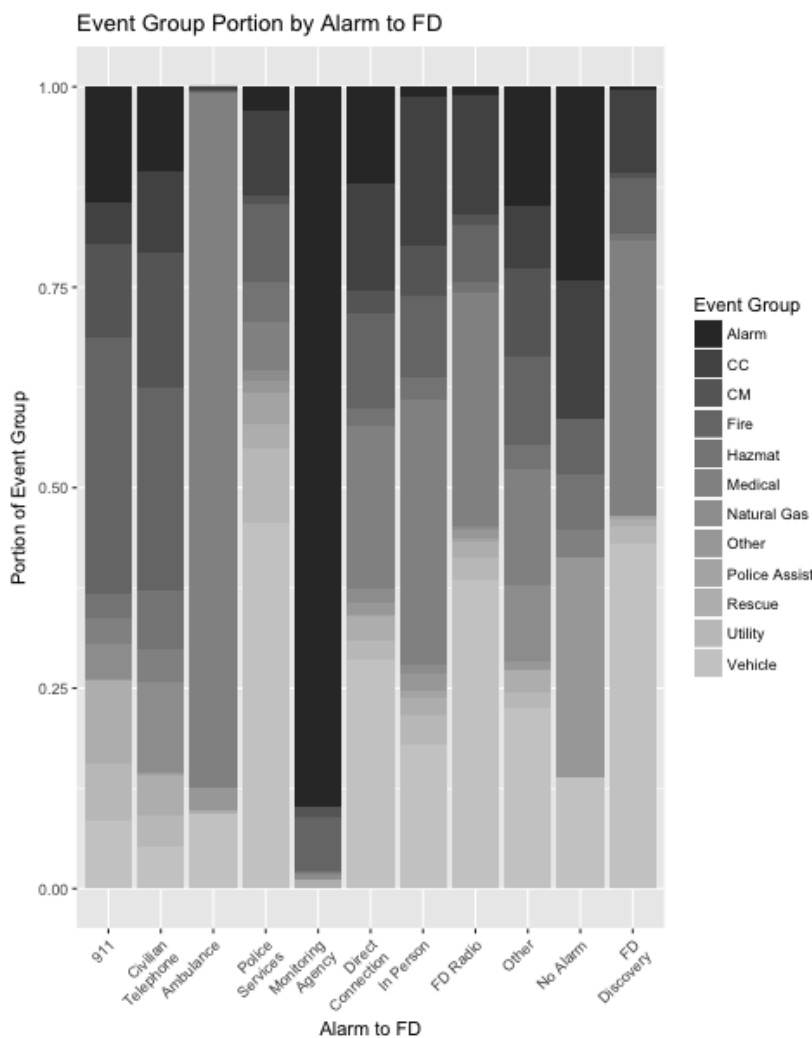
Considering every incident an event that happens at some type of property facilitates the types of breakdown shown in **Figure 3** and **Figure 4**; total incidents by property group and event group, respectively. Both groups have been combined by combining many similar types of levels into one; otherwise the features themselves would have far too many features (115 in the case of event type). It is clear from these figures that the most numerous occurring property type

is by far Residential while Medical events occur for nearly half of all incidents (approximately 350,000/720,000).

Regarding the source of data, event type is likely available before the incident as the feature EVENT\_TYPE\_CD (Event Type Call Dispatched) contains event type information that is dispatched before the arrival of first responders. Property type, however, is sourced from the OFM Standard Incident Codes List and is thus likely inputted following an event.

### Alarm to Fire Department, Event Group

The Alarm to Fire Department feature, ALARM\_TO\_FD, describes from where the incident was reported. EVENT\_GROUP is a grouping of different types of event. The features are related in the sense that ALARM\_TO\_FD is the input source (911, Police, etc) whereas EVENT\_GROUP describes what type of incident the call turned out to be. Perhaps a reasonable expectation would be that certain types of calls correspond to certain types of events. **Figure 5** shows the portions of Event Type group by Call source.



**Figure 5:** The portion of Event Types (grouped for similarity) for each call source (Alarm to FD).

The visual in Figure 5 displays the portion of Event Type Group by Call Source (Alarm to Fire Department) and suggests that the source of call will impact which type of call is received. For example, if an incident is reported by a Monitoring Agency there is greater than a 75% chance that the Event Type will be an Alarm.

### Temporal Feature Analysis

This dataset includes multiple temporal features including Incident and Dispatch Date as well as corresponding times for when the call was first received, when first responders initially arrived on the scene, and when the incident was controlled.

A breakdown of incident quantities by hour and day is provided in **Figure 6**. In general, the trend of number of incident calls received increases from a low around 4 am where they begin to increase until the max which occurs at approximately

6pm. After 6pm the quantity of incidents per hour tends to decrease again until 4 am. This cycle is

pretty standard for Monday through Thursday, however, Friday through Sunday show increased relative call volumes in the evening when compared to weekdays. There is also a noticeable increase in call volumes early Sunday morning where activity from Saturday night would be carried over.

## Association Rules

Association rules were run on all observations of the following features: EVENT\_GROUP, ALARM\_TO\_FD, RESPONSE\_GROUP, PROPERTY\_GROUP, MONTH, DAY, HOUR, EVENT\_TYPE\_CD. No rules were found with the three temporal features (MONTH, DAY, HOUR), however, other interesting patterns were discovered. Three of these rules are summarized below:

1. {Event group = Alarm, Property Group = Residential} => {Alarm to Fire Department = From Monitoring Agency}

This rule describes incidents where the Fire Services responded to Residential Alarms. Of these alarms, 82% (confidence) were received from a monitoring agency.

2. {Event group = Alarm} => {Response Group = False Fire Calls}

This simple statistic, also with high lift (3.9), shows that 81% of the alarms that the fire services respond to are false fire calls. This is an incredibly high percentage and represents approximately 15% of all incidents in this dataset.

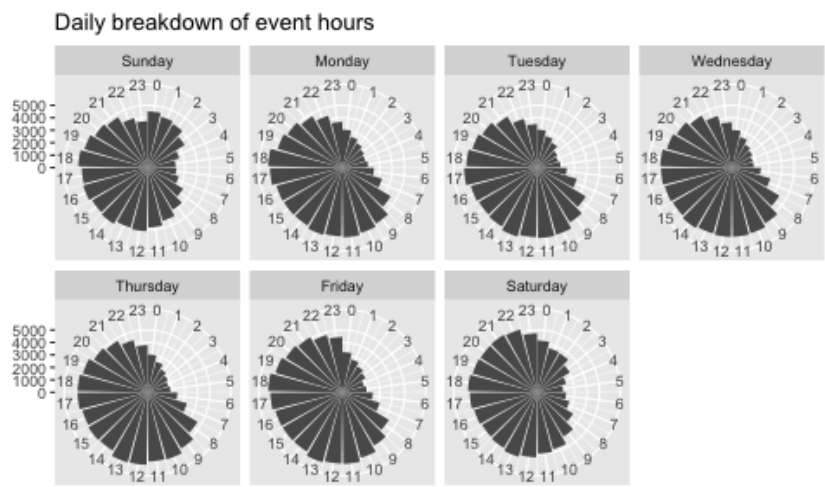
3. {Event group = Alarm, Alarm to Fire Department = From Monitoring Agency} => {Response Group = False Fire Calls}

This rule elaborates on rule #2 by showing that a high percentage (approximately 82%) of aforementioned false alarm calls are received from a monitoring agency.

4. {Response group = Medical/resuscitator call, Event type call as dispatched = Medical} => {Event group => Medical}

This rule perhaps seems obvious but is worth noting: with a confidence of 100%, all incidents responded to as medical calls, which were dispatched as medical calls, ended up being classified as medical calls. This data set has a high proportion of medical calls. Note that only 138,000 of the some 350,000 medical incidents in this data set fit that pattern. This suggests that there is some variability among these features of medical calls.

5. {Alarm to Fire Department = From Ambulance, Event type call as dispatched = Medical} => {Event group => Medical}



**Figure 6:** Incident counts grouped by hour and day. There is increased incident quantities on Weekend evenings (Friday, Saturday, Sunday morning) versus the weekdays.

As with rule #4 this is perhaps a sanity check (despite a lift value of 2) and may seem as expected: Every incident received from an Ambulance is dispatched as Medical with an event type categorized as Medical. This stipulates that ambulances make no other calls to the fire services besides medical calls.

## Modeling

### Model Selection

To proceed with modeling fire risk was chosen as a dependent variable and risk was first categorized according to criteria provided in prior research and secondly by extending the classification thresholds in order to reduce a significant class imbalance. After assigning the dependent attributes three classifiers were run to predict the class variable: Logistic Regression, Naive Bayes and Random Forest. These models were chosen for their ability to handle categorical inputs, of which the majority of predictors are, as well as for their differences in the linearity and feature independence assumptions. Logistic Regression is unique among the three in its incorporation of continuous predictors, for Naive Bayes and Random Forest the continuous predictors (EST\_KM, INITIAL\_UNIT\_PERSONNEL) were grouped into discrete buckets.

The models were trained on the same features, with similar random selection of training and test sets, in order to facilitate performance comparison. Each classifier was trained and run ten times with the results being recorded.

### Fire Classification: Multi-class

In order to predict the risk of structural fire incidents, Asgari et al. (2012) classify each incident into five distinct categories, ranging from very low risk to very high risk, based on the values of three features:

- **Very Low:** 0 Fatalities, 0 Injuries, Damage < \$5000
- **Low:** 0 Fatalities, 0 Injuries, \$5000 <= Damage < \$10,000
- **Moderate:** 0 Fatalities, 0 Injuries, \$10,000 <= Damage < \$50,000 *or* 0 Fatalities, 1 Injury, Damage < \$50,000
- **High:** 0 Fatalities, Injuries <= 1, \$50,000 <= Damage < \$100,000 *or* Injuries > 1, Damage < \$100,000
- **Very High:** Fatalities > 0 *or* Damage >= \$100,000

**Table 2** shows a cross table of Incident Class by Incident Type, after assigning classes based on the above criteria, and makes two things apparently clear: that there is a significant class imbalance when using this method (*Very Low* accounts for 99% of observations!) and most incidents that aren't classified as *Very Low* are likely from one of two incident types: Fire or Vehicle, which respectively contain 73% and 23% of all *Non-Very Low* incidents.

## Fire Classification: Binary Class

In order to help combat the significant class imbalance an alternative classification method was applied to the data set: by slightly changing the above thresholds, and examining a few post-incident features, a binary class was created to identify critical incidents was created. The characteristics chosen to declare incidents critical is as follows:

1. All incidents that were previously classified as *Low, Medium, High* or *Very High* risk are now assigned *critical*.
2. Damage (EST\_LOSS): There are a total of 10,480 incidents with Damage > 0. All of these incidents are assigned *critical*.
3. Responding Units: The 99th percentile of RESPONDING\_UNITS is 7 units. Any incident with seven or more responding units is assigned as *critical*.
4. Rescues: There are 7907 incidents where RESCUES > 0.
5. Office of the Fire Marshall Investigations: there are 1144 observations in the data set where OFM\_INVESTIGATIONS\_CONTACTED is TRUE.
6. Estimated value at risk: This data set also contains a feature EST\_VALUE\_AT\_RISK which is categorical and contains levels pertaining to dollar value estimates. There are 9565 observations with non-zero value estimates for this feature. All incidents with a non-zero value for this feature are assigned to the *critical* class.

After reducing the classification levels and choosing a wider threshold in which to assign observations to the critical class the class imbalance is reduced significantly: from approximately 99% to 96% for the whole dataset (as displayed in **Table 3**).

As with the multi-class classifier, the critical incidents are not distributed evenly across event types: this breakdown is shown in **Figure 7**. To help further balance the *critical* incidents relative to *non-critical* the decision was made to model only incidents whose event group is Fire. This further reduced the class imbalance to a 79/21 percent split between *non-critical* and *critical*, respectively.

|               | VL     | L    | M    | H    | VH   | % non-VL |
|---------------|--------|------|------|------|------|----------|
| Alarm         | 137691 | 24   | 42   | 5    | 4    | 1.35     |
| CC            | 15243  | 11   | 10   | 0    | 2    | 0.41     |
| CM            | 22639  | 0    | 2    | 0    | 0    | 0.03     |
| Fire          | 59153  | 848  | 1888 | 598  | 747  | 73.71    |
| Hazmat        | 7866   | 2    | 2    | 0    | 1    | 0.09     |
| Medical       | 354719 | 0    | 22   | 1    | 0    | 0.41     |
| Natural Gas   | 9574   | 2    | 3    | 0    | 0    | 0.09     |
| Other         | 11895  | 1    | 5    | 3    | 2    | 0.19     |
| Police Assist | 1126   | 0    | 0    | 0    | 0    | 0.00     |
| Rescue        | 19238  | 0    | 3    | 0    | 1    | 0.07     |
| Utility       | 13541  | 0    | 3    | 0    | 0    | 0.05     |
| Vehicle       | 62069  | 474  | 705  | 78   | 47   | 23.55    |
| Total         | 714754 | 1362 | 2685 | 685  | 804  | 0.007    |
| %             | 99.23  | 0.18 | 0.37 | 0.09 | 0.11 |          |

**Table 2:** A cross-table of Fire Incident Types (Grouped) vs Fire Risk Classification as outlined by Asgari et al. (2012). Note the high class imbalance with regards to VL incidents: over 99%!

|       | Critical | Non-critical |
|-------|----------|--------------|
| Count | 29782    | 690588       |
| %     | 4.1      | 95.9         |

**Table 3:** A reduction in the class imbalance after loosening the threshold for critical incidents.

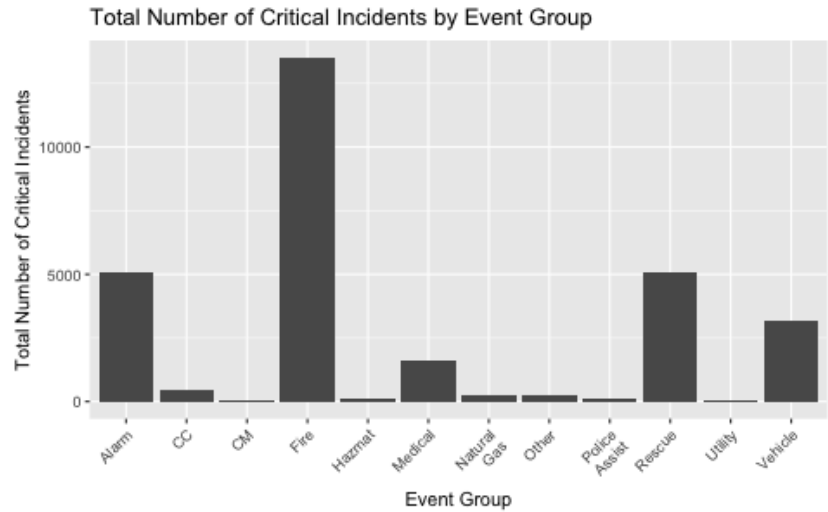


## Fire Incidents: Binary Class with Logistic Regression

Logistic Regression was the first classifier used. As the class feature, CRITICAL, was created using features known after the incident (INJURIES, FATALITIES, EST\_LOSS, etc), it was then regressed onto features which are apparently available at call time in an effort to contribute to the research question, "What further information, if any, can be provided to first responders at the time of incident call?"

The logistic model was initially trained on the Fire Incidents without any interaction:

CRITICAL ~ EVENT\_TYPE\_CD + ALARM\_TO\_FD + RESPONSE\_TYPE + EST\_KM + INITIAL\_UNIT\_PERSONNEL + INCIDENT\_DAY + INCIDENT\_MONTH + INITIAL\_CALL\_HOUR + INITIAL\_CALL\_MIN



**Figure 7:** The above figure shows the unequal distribution of critical incidents to various event types. Most critical incidents occur during an alarm, fire, rescue or vehicle event.

|                 |   | Actual Class |      |
|-----------------|---|--------------|------|
|                 |   | 0            | 1    |
| Predicted Class | 0 | 19423        | 1175 |
|                 | 1 | 383          | 3717 |

**Table 4:** Confusion Matrix of Logistic Regression classifier on FIRE incidents. 0 and 1 represent critical and non-critical incidents, respectively.

| Statistic           | Value (%) |
|---------------------|-----------|
| Accuracy            | 91.6      |
| False-Negative Rate | 31.5      |
| Precision           | 90.6      |
| Recall              | 68.4      |
| F1-Score            | 77.9      |

**Table 5:** Five comparative metrics for the above confusion matrix

After training the model on the training set, probabilities were assigned to observations in the test set. Observations with a probability of critical greater than 0.5 were classified as *critical* and all other observations were classified as *non-critical*. The confusion matrix of this classification is presented in **Table 4** and contains favourable results.

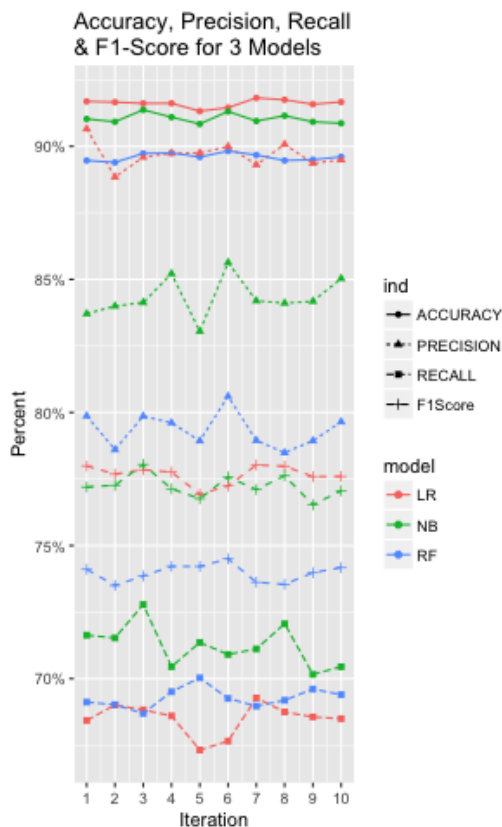
**Table 5** contains the calculated evaluation metrics for the associated confusion matrix. Accuracy of 92% is acceptable for early attempts but this value is likely to be able to be increased with more fine tuning of the model. Note that by assigning Non-critical to every class you would receive an accuracy of 79% due to the class imbalance. This is why Precision, Recall and F<sub>1</sub>-Score are also calculated as they depend on the presence of True Positives (correctly predicting the presence of the class attribute) and thus are well suited for class imbalances.

## Fire Class: Binary Class with Naive Bayes & Random Forest

In order to benchmark the performance of the Logistic Regression classifier the critical incidents were also predicted using Naive Bayes and Random Forest classifiers. The general formula and features presented above were used for all models with some minor adjustments where necessary: Naive Bayes

and Random Forest can only handle categorical input features with Random Forest having an upper limit on factor levels. To meet these requirements some adjustments were required.

Each model, including Logistic Regression, was trained and tested on ten different random samples with the resulting evaluation metrics recorded. Only fire incidents were used: the training sets comprised 60% of the fire incident subset with the test sets accounting for the remaining 40%. Accuracy, Precision, Recall and the F<sub>1</sub>-Score are plotted for each iteration in **Figure 8** while the averages for the corresponding metrics, across the 10 iterations, is presented in **Table 6**.



**Figure 8:** Accuracy, Precision, Recall & F<sub>1</sub>-Score plotted for each iteration of the 3 models

| Model               | Accuracy | Precision | Recall | F <sub>1</sub> -Score |
|---------------------|----------|-----------|--------|-----------------------|
| Logistic Regression | 91.61    | 89.68     | 68.49  | 77.66                 |
| Naive Bayes         | 91.04    | 84.32     | 71.24  | 77.22                 |
| Random Forest       | 89.60    | 79.35     | 69.27  | 73.97                 |

**Table 6:** Average evaluation metrics for the 3 models run in this project

accurately labeled (Precision). Conversely, however, it identified on average the lowest percentage of actual critical incidents (Recall), although in this case the average difference between the other models is significantly less. This relationship brings the F<sub>1</sub>-Score for the three models much closer together,

## Model Evaluation

Accuracy is the metric with the least variance between models and between iterations. This might partly be attributed to the class imbalance: a null model would achieve accuracy of 79% by simply predicting every incident to be non-critical! That only leaves a 21% margin for improvement. With this in mind the models improved the accuracy of the null model by between 11 and 14 percent which is a non-trivial amount.

The remaining three metrics; precision, recall, and F<sub>1</sub>-Score are dependent on the quantity of True Positives and thus would be 0 in the case of the null model. Precision is a percentage of predicted *critical* incidents that actually were (TP/TP+FP) whereas recall represents the percentage of *critical* incidents that were accurately identified (TP/TP+FN). F1-Score is the harmonic average of these two metrics (2TP/2TP+FP+FN). Using metrics with the quantity of True Positives as the numerator ensures that they are sensitive to models that aren't effective in the presence of a class imbalance: the null model would be 0 for all metrics!

Depending on a specific application a model may be trained to maximize a certain metric: perhaps some situations demand high recall, others high precision, and some still where a balance between the two is most desirable. Among the three models trained and tested in this project Linear Regression had the highest percentage of predicted critical incidents be

especially between Linear Regression and Naive Bayes which alternate between maximizing precision and recall, respectively.

## ANOVA

A One-way Analysis of Variance was run on the ten Accuracy and  $F_1$ -Score metrics produced from each model's test set predictions. The corresponding ANOVA tables are provided for Accuracy and  $F_1$ -Score in **Table 7** and **Table 8**, respectively.

The p-value represents the probability that the averages for the corresponding metrics are from the same population and are both incredibly low for each metric tested. Thus it is safe to conclude, with 99% confidence, that the evaluation metrics were sampled from different populations and thus the choice of model will impact the resulting predictive performance. Not shown, however, are the associated coefficients and p-values for differences between the models  $F_1$ -Score: perhaps as expected by the very close averages for  $F_1$ -Score between Logistic Regression and Naive Bayes there is a much lower p-value for a difference between these two models. Although the difference is still significant, it is much less so with a p-value of 0.0152, which corresponds to only a rejection of the averages being equal at the 95% confidence level. This means that although the specific metrics themselves show a significant difference between the three models, there is much less significant difference between the metric averages of the two models Naive Bayes and Logistic Regression.

|           | Df | Sum Sq   | Mean Sq  | F value | Pr(>F)    |
|-----------|----|----------|----------|---------|-----------|
| model     | 2  | 2.15e-03 | 1.07e-03 | 435.32  | < 2.2e-16 |
| Residuals | 27 | 6.69e-05 | 2.48e-06 |         |           |

**Table 7:** One way ANOVA: Accuracy vs. Model

|           | Df | Sum Sq  | Mean Sq | F value | Pr(>F)    |
|-----------|----|---------|---------|---------|-----------|
| model     | 2  | 0.00814 | 0.00407 | 288.81  | < 2.2e-16 |
| Residuals | 27 | 0.00038 | 0.00001 |         |           |

**Table 8:** One way ANOVA:  $F_1$ -Score vs. Model

## Conclusion

A fire incident dataset was obtained from an open data catalogue, converted to tabular structure, and subjected to preparation followed by an exploratory data analysis. Trends and relationships between features were explored before a binary class attribute was assigned based on features available following an incident. Three predictive models were then used to predict the class attribute using a different set of features which would be available at the time of incident call. Each model was trained on 10 randomly selected training sets, comprising 60% of a subset of the data, and then used to make predictions on the corresponding test set which accounted for the other 40%. Four evaluation criteria were then selected, three of which due to their sensitivity to the imbalance that existed in the class attribute. Generally speaking Logistic Regression and Naive Bayes were the two most successful model, with the highest precision and recall respectively. As they alternated between highest performance on these two metrics the average between them,  $F_1$ -Score, was much closer but still significant. Thus in closing the choice of model should be decided upon based on specific criteria outlined by the researcher.

# References

- Asgary, A., Naini, A.S., & Levy, J. (2012). Modeling the risk of structural fire incidents using a self-organizing map. *Fire Safety Journal*, 49, 1-9.doi:10.1016/j.firesaf.2011.12.007
- Asgary, A., & Sadeghi-Naini, A. (2013) Modeling number of firefighters responding to an incident using artificial neural networks. *International Journal of Emergency Services*, 2(2), 104-18.doi:10.1108/IJES-03-2012-0001
- Asgary, A., Sadeghi-Naini, A., & Levy, J. (2009). Intelligent Security Systems Engineering for Modeling Fire Critical Incidents: Towards Sustainable Security. *Journal of Systems Science and Systems Engineering*, 18(4), 477-488.doi:10.1007/s11518-009-5121-2
- Asgary, A., Sadeghi-Naini, A., & Kong, A. (2009). Modeling loss and no-loss fire incidents using artificial neural network: Case of Toronto. *Science and Technology for Humanity*, 159-163.doi:10.1109/TIC-STH.2009.5444513
- Ceyhan, E., Ertügay, K., & Düzgün, Ş. (2013). Exploratory and inferential methods for spatio-temporal analysis of residential fire clustering in urban areas. *Fire Safety Journal*, 58, 226-249.doi:10.1016/j.firesaf.2013.01.024
- Asgary, A., Ghaffari, A., Levy, J. (2009). Spatial and temporal analyses of structural fire incidents and their causes: A case of Toronto, Canada. *Fire Safety Journal*, 45, 44-57.doi:10.1016/j.firesaf.2009.10.002
- Sekizawa, A. (2012). Necessity of Fire Statistics and Analysis Using Fire Incident Database - Japanese Case -. *Fire Science and Technology*, 31(3), 67-75.doi:10.3210/fst.31.67
- Bergen, G., Frattaroli, S., Ballesteros M.F., Ta, V.M., Beach, C., & Gielen, A.C. (2008) *J Community Health*, 33, 103-109.doi:10.1007/s10900-007-9070-8
- Walmsley, P. (2007) *XQuery*. Sebastopol, CA: O'Reilly Media, Inc.
- Ripley, B.D., & Hornik, K. (2001) Date-Time Classes. *R News*, 1(2), 8-11. Retrieved from [http://cran.r-project.org/doc/Rnews/Rnews\\_2001-2.pdf](http://cran.r-project.org/doc/Rnews/Rnews_2001-2.pdf)
- Bebee, D., (2012) New dispatch system helps fire department reduce response times: *The Record*. Retrieved from <http://www.oafc.on.ca/article/new-dispatch-system-helps-fire-department-reduce-response-times>
- G. Casella, S. Fienberg, I. Olkin, (2013). *An introduction to statistical learning: with applications in R*. New York, NY: Springer Science+Business Media
- Fire Services Incident Data: *City of Toronto Open Data Catalogue* (2018). Retrieved from <https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#e3d443bb-2593-2615-4972-20e24c0ab876> under the Open Data License: <https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-licence/>