Mining Toronto Fire Services Incident Data

CKME 136: Data Analytics Capstone
by Geoffrey Clark
2018

Research Question

- What additional information, if any, can be provided to first responders at the time of incident call?

Dataset

- Open Source: Toronto Open Data Catalogue
- XML
- 720,340 Observations, 100 Features
- 2011 - 2016

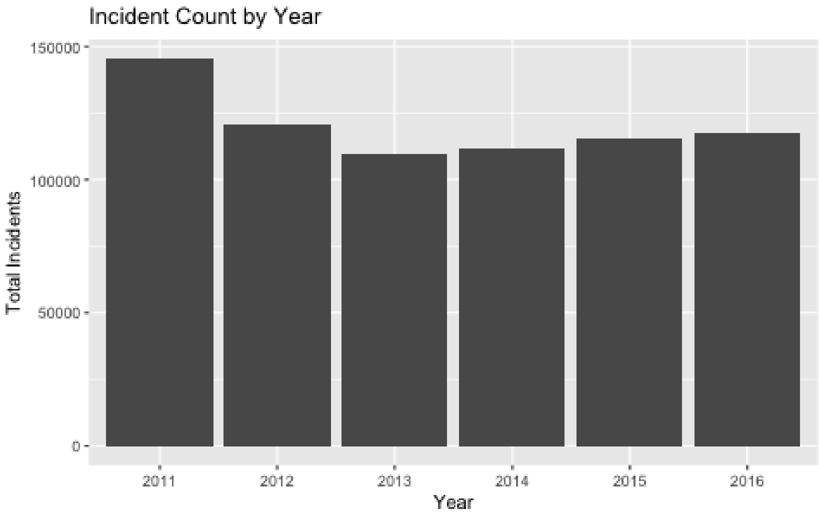
YSICAL_CONDITION_1><PHYSICAL_CONDITION_2></PHYSICAL_CONDITION_2><PHYSICAL_CONDITION_3>
</CIV_EVACUATION_REASON_1><CIV_EVACUATION_REASON_2></CIV_EVACUATION_REASON_2><OPP> </O
CIPAL_BUILDING_OFFICE> </MUNICIPAL_BUILDING_OFFICE><MUNICIPAL_HEALTH_OFFICE> </MUNICIP
FIRE></EXTENT_FIRE><EXTENT_SMOKE></EXTENT_SMOKE><EST_NUM_PERSONS_DISPLACED>0</EST_NUM_
C_TYPE></OCC_TYPE><BLD_STATUS></BLD_STATUS><BLD_HEIGHT>0</BLD_HEIGHT><LEVEL_OF_ORIGIN>
ESENCE_AND_OPERATION_MAIN_FLOOR><SMOKE_ALARM_FAILURE_TO_OPERATE></SMOKE_ALARM_FAILURE_
SMOKE_ALARM_ON_ALL_FLOORS></SMOKE_ALARM_ON_ALL_FLOORS><SMOKE_ALARM_IMPACT_ON_EVAC></SM
></FIRE_ALARM_SYSTEM_PRESENCE><FIRE_ALARM_SYSTEM_OPERATION></FIRE_ALARM_SYSTEM_OPERATI
LER_SYSTEM_ACTIVATION></SPRINKLER_SYSTEM_ACTIVATION><RespondingUnits><CAD_UNIT_ID>A135
IME>2011-01-02 09:31:51</CLEAR_TIME></RespondingUnits><RespondingUnits><CAD_UNIT_ID>P1
AR_TIME>2011-01-02 09:35:45</CLEAR_TIME></RespondingUnits></INCIDENT><INCIDENT><INCIDE
><ARRIVE_DATE>2011-01-02 09:00:29</ARRIVE_DATE><EVENT_TYPE_CD>MECP</EVENT_TYPE_CD><MAI
1000665000</FD_STATION><OFM_INVESTIGATIONS_CONTACTED>0</OFM_INVESTIGATIONS_CONTACTED><
IAL_CALL_HOUR><INITIAL_CALL_MIN>55</INITIAL_CALL_MIN><INITIAL_CALL_SEC>31</INITIAL_CAL
NSCENE_HOUR><ONSCENE_MIN>0</ONSCENE_MIN><ONSCENE_SEC>29</ONSCENE_SEC><INITIAL_UNIT_PER
TO_FD><RESPONSE_TYPE>89</RESPONSE_TYPE><RESCUES>0</RESCUES><FF_INJURIES>0</FF_INJURIES
T_APP_SEC><CONTROL_DATE></CONTROL_DATE><CONTROL_HOUR></CONTROL_HOUR><CONTROL_MIN></CON

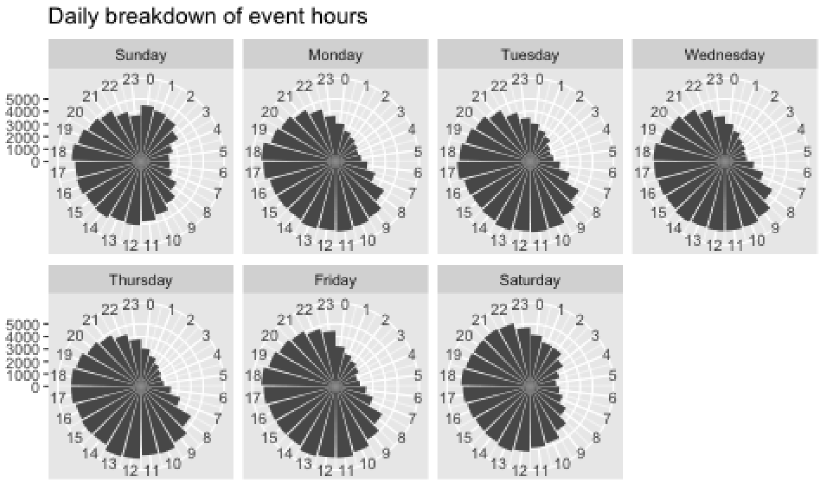
Geoffrey Clark

Features

- Temporal: Date & Time of Incident call, dispatch, arrival (on-scene) & control
- Types of Incidents: Fire, Alarm, Vehicle, Rescue, Medical, …
- Location: Property Type, FSA, Main & Cross Street
- Call Source: Agency, Ambulance, 911...

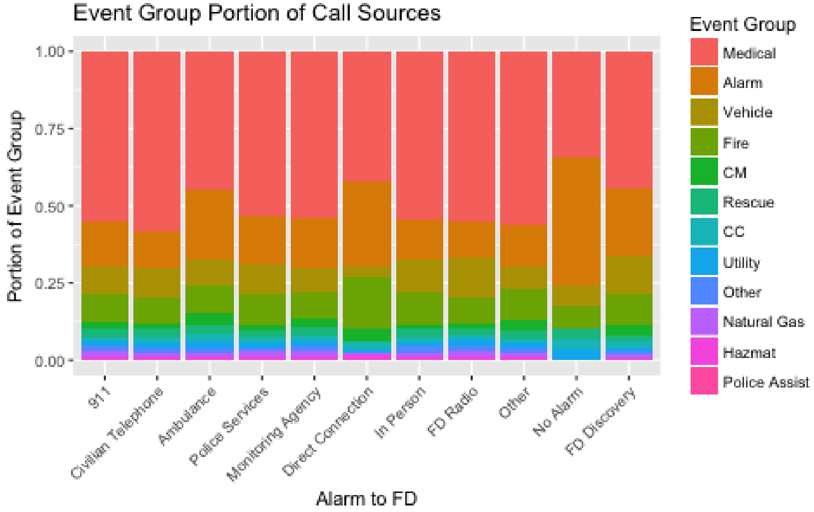
Some Patterns

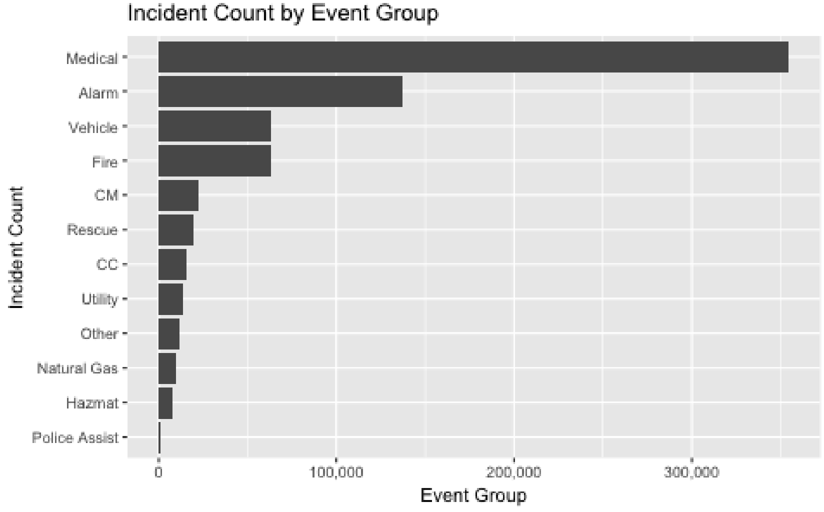




Geoffrey Clark

Some Patterns





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

81% Confidence! 112,659 Incidents

Geoffrey Clark

Modeling

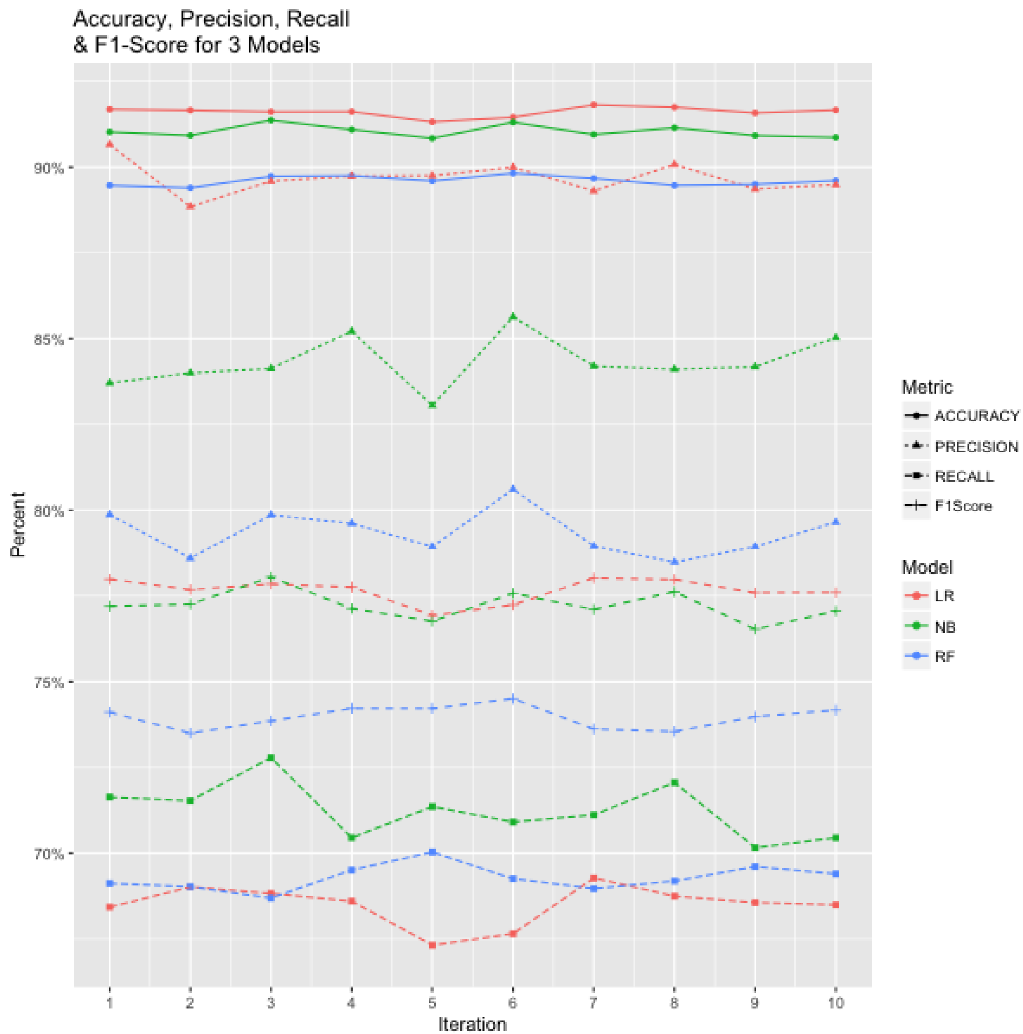
- Logistic Regression, Naive Bayes, Random Forest
- Train & Predict “Critical Incidents”
- Randomly selected training set (60%), test set (40%)
- “Before” and “After” features
- Accuracy, Precision, Recall & F_1-Score

Supervised Learning

- Critical Incidents:
  - Class Label: Damage, Responding Units, OFM Investigations, Injuries, Fatalities, Rescues
  - Predict: Event type, call source, estimated km, month, day, hour..
- Only Fire Incidents Used

	Non-Critical	Critical
Whole Dataset	690,588 (96%)	29,782 (4%)
Fire Incidents	49,722 (79%)	13,512 (21%)

Modeling Results



		Predicted	
		Non-Critical	Critical
Actual	Non-Critical	19423	1175
	Critical	383	3717

Accuracy	91.6
False-Negative Rate	31.5
Precision	90.6
Recall	68.4
F_1-Score	77.9

Geoffrey Clark

Conclusion

- What additional information, if any, can be provided to first responders at the time of incident call?
- Logistic Regression: Accuracy, Precision
- Naive Bayes: Recall
- Generalizability
  - Entire Dataset
- Limitations
  - Time, Domain knowledge, Computation Power, Dataset

Questions & Thanks