Analysis of 911 Calls

Sai Ganesh Nindra

Abstract—Quick emergency response times is necessary as It ensures public safety and helps the emergency services to serve and protect its people quicker and faster. There are various factors which affect commute like weather, traffic and transport alerts. Using historical data and leveraging the current days machine learning techniques, These techniques can help the authorities in rectifying and adapting to the situation to provide the services sooner. Emergency services have the requirement of reaching on scene as soon as possible. Using all the different data that affect the commute between the center and the scene we hope to demonstrate an optimized time which the dispatch centers should have taken. We can predict response times for any future cases and this opens up doors for future developments in resource optimization.

Index Terms— Dispatch Response times, Machine Learning, Weather

1 INTRODUCTION

This project focuses on predicting emergency response times (TIM_MIN) using machine learning models. The dataset comprises geospatial, temporal, and environmental features, such as Latitude, Longitude, hour, temperature_2m (°C), and rain (mm), which are necessary for understanding the dynamics of emergency service delays. By applying and comparing models like Ridge Regression, Random Forest, XGBoost, and Stacking, the project explores both linear and non-linear relationships within the data. The methodology involves Data Preprocessing, Feature Engineering and Modelling. This project depicts the application of supervised learning in a real-world context, showcasing how feature selection, preprocessing, and model optimization are tailored to a specific problem domain.

911 Dispatch calls data is one of the most important thing to be optimized, this generation where crime is becoming more common we need dispatch teams to act faster than ever. Coordination between complainant, precinct call receiver and patrol unit is of utmost importance to make this process go through faster. There are definitely other factors like traffic and weather that affect the time taken for patrol units to arrive on scene.

From the article 'The Evolution of Emergency Calls'[1] we learn that it was not until 1999 that the force could retrieve the call location, which speeds up things as there are cases where the complainant/caller does not know their location. However since reporting is still done through call, it supposedly is not the most efficient method because there are a lot Crime in Progress situations where sometimes the caller cannot speak as it will compromise their safety.

From a New York times article '911 Systems Disrupted in at Least 3 States'[2] we gain insight on the fact that the 911 centers even though they faced outage due to the CrowdStrike situation they were up and running with reporting happening in written form.

Sometimes being even 30 seconds earlier to a scene can help the authorities alleviate the situation and save lives or prevent said crime from happening. It is all a matter of seconds. Response time is a crucial factor in dispatch units reaching the scene as soon as possible and analysis will help us figure out.

This article clearly tells us how important response times are, its headline is 'As Vallejo police force shrinks, 911 response times soar'[3] Vallejo might be one case but any force/department must have enough units to respond to every call and respond as soon as possible to prevent crime. If data shows a certain amount of calls, departments must be equipped with sufficient manpower to respond to the call and dispatch units for it. Lives are lost and are turned upside down within the blink of an eye, every 911 call is important and must be attended to. The crucial factors affecting response times are usually, total number of dispatch units available to call sent to dispatch ratio, weather and traffic.

The inferences made from this project can help us in understanding various aspects of a 911 call. It can let us know which factor is more interlinked to response time, which model should be used to predict response times in a more accurate and sound manner and which areas in a geographical range take the most amount of time and require more optimization of resource allocation.

The main research objectives of this project are :

RO1:- To describe the trends within the boroughs of NYC to visualize crime statistics

RO1:- To describe the trends within Category of crime on how what incident is reported more

RO1:- To describe trends within Latitude and other quantitative features to see if there are more incidents reported in that specific latitude with respect to the longitude, temperature, rainfall and response time.

RO1:- To describe trends within Longitude and other quantitative features to see if there are more incidents re ported in that specific Longitude with respect to the temperature and rainfall.

RO1:- To describe trends within temperature and other quantitative features to see if there are more incidents reported in that specific temperature with respect to the temperature and response time.

RO2:- (Regression) To predict time taken by the NYPD to reach the scene based on factors like latitude, longitude and weather

RO3:- To defend the model for predicting the time taken by dispatch to arrive on scene in RO2 by evaluating different metrics and validating the model.

RO4:- To evaluate the causal relationships implied by the RO2 model

1.1 Related Work

The article, "A Review of Incident Prediction, Resource Allocation, and Dispatch Models for Emergency Management"[4] is on emergency response management. Its prime focus is on incident prediction, detection, resource allocation and dispatching. It highlights problems in using spatiotemporal decision-making. The use of data driven models and predictive algorithms aligns closely with my work. The key takeaways being the way of incorporating real time incident data and improving dispatch efficiency in their way.

The article published by Zhou, Chen, Li, Liu and Yu[5] explores optimizing emergency medical services altogether, this employs a new approach by using spatiotemporal data for relocating the vehicles to make response times faster. Instead of provided data to optimize their work, their paper is suggesting to start from the location of dispatch vehicles. They are using real time traffic data and relocation strategies. Future implementation of my project could enhance the outcome of the models by using big data analysis and dynamic relocation optimization techniques.

The report published by Vera Institute of Justice[6] focuses on the overreliance on police for non-violent and non -criminal situations. There needs to be better resource allocation to serve matters with higher degree of risk. They have performed analysis as well on 911 calls and their aim is to improve response systems. One big takeaway for me was maybe categorizing incidents based on criminal and non-criminal for refined analysis but again it would be a future implementation as you will see in the coming sections.

The paper " Classifications of events by dispatchers and officers"[7] explores the accuracy of police call classification. To create a model that optimizes the response time we need to have all the accurate details about a 911 call. This paper helps highlight the challenges of incident misclassification by dispatchers. It helps me understand how important data quality is and its effectiveness. Integrating more refined data as well could further optimize predictions for emergency responses

The document published by Jack, Keller and Stephen from the esteemed Northeastern University[8], performs a big data analysis on crime data and the meaning of these 911 calls and its data. The same data which I have used in my analysis and prediction. They dive into the theoretical aspects of analyzing the data like how the crimes are reported, errors in reporting, how to interpret and operationalize the data and how accurate the classification of crimes are which has been problem in this project as well.

The paper authored by Ketki, Shruthi and Sukhada[9] has been a huge inspiration and been helpful as they are also performing an analysis and prediction on 911 calls. They use geospatial data and have employed regression techniques for modelling. But at the same they have employed different tools of Big Data and have used Frequent Pattern mining as well.

Despite the extensive research done by all the esteemed authors above, a clear gap exists between response time prediction with dynamic resource allocation optimization. Many studies focus on either one of them, they lack a framework that does the prediction using factors like traffic data, weather data and dispatcher accuracy. This project addresses this gap by developing models for predicting response times factoring in the weather data. Unlike the existing work we can have it incorporated with the correct dispatcher data and traffic data if and when available. None of these explicitly predict the response times on any given input. This project bridges that gap in which a model is giving a time limit which holds the department up to a standard. But remains the first milestone in a long journey to optimize and make this field better in the interest of the public

2 Exploratory Data Analysis

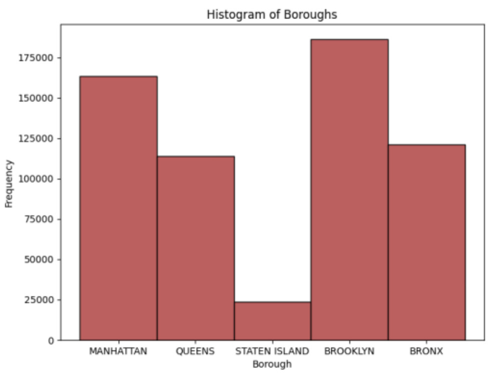
To perform analysis of 911 calls, the data was sourced from a government website of the city of New York, https://data.cityofnewyork.us/. This data is available for public access, it is updated at regular intervals and the data provided is of recent time. It was downloaded as a comma separated value file. Weather data was sourced from a open source website Open-Meteo which consists of historical data dating all the way back to early 20th century.

The dataset used is indicative of the real world. The extremes of weather data also exist where there has been no rainfall and a downpour on other days. The file consists of ever 911 call attended by the dispatch center which provides us with data like when the call was accepted, when a unit was dispatched and when it arrived on the scene. The time taken by the officials has also been variable, which indicates that it covers all possible scenarios such as extreme traffic, bad weather conditions, hostage and any other emergency situations which require different protocols. It consists data like latitude and longitude as well, but only that of New York which is the main area of concern. It covers all the boroughs that make up New York. The data that was chosen for analysis was of the month June and the weather data is sound as well, no snowfall recorded and variable levels of rainfall.

The records are easily interpretable given the easy-to-read column names. Taking a record into account, it can be easily interpreted as having an ID 104281501 which was created on 30th June 2024. The incident time was noted as 23:59:59, reported to the 7th Precinct. The incident took place in Manhattan and the patrol unit dispatched was of South Manhattan. It was classified as a non-CIP job which means it is not a crime in progress. The dispatch time was the very instant the call was registered and the arrival time was the same as well. This can sometimes be indicative of human error or a right place right time situation where. The GEO_CD_X refers to the X coordinate of where the call came from and GEO_CD_Y is the Y coordinate of where the call came from. The latitude and longitude correspond to the segment of the road where the incident occurred. The description tells us about the situation at the site of incident here, meeting the complainant inside.

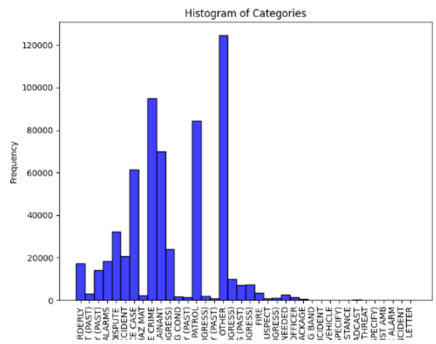
The data was transformed using different techniques to harvest the necessary information for future use. Given the large size of the dataset (over 3 million rows), it was partitioned into different chunks of a certain size and then merged with the weather data with respect to the date of the incident. The standard check for null values and imputation was done. Redundant columns were dropped, and new columns were constructed to extract meaningful and necessary data for visualization.

Upon visualizing the qualitative data of borough names and the number of incidents in each borough, we could gather insights such as the borough that receives the greatest number of calls as well as least number of calls. This is a histogram plot which provides us a frequency distribution plot. [RO1]

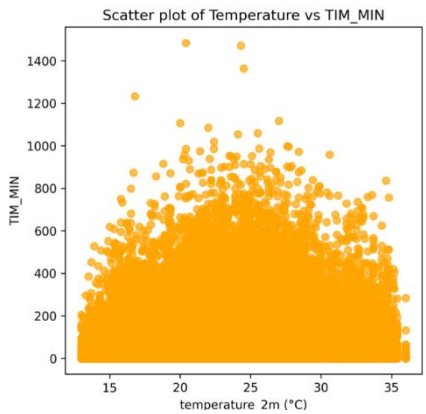


As it can be seen from the above graph that the crime is soaring in the borough of Brooklyn and comparatively within the limits of New York it is very low in Staten Island. But it is possible that since Staten Island is relatively much smaller in size, the crimes reported are also much lesser in number. [RO1]

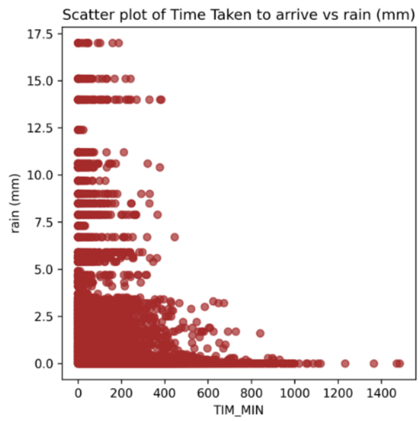
Another frequency distribution plot helps us analyze the type of crimes reported in the entire city of New York. The plot doesn't seem to be the most useful one, as there are multiple categories of crimes which are reported so less compared to the others. It is hard to figure out the range in which they have been reported. The category 'OTHER' has been reported far too many times that the histogram is highly skewed. It is possible that due to many crimes being reported in a vague manner it might not have been classified properly at the time of inception of the dispatch call. The second category has been 'Investigate Possible Crime' which are mostly non-CIP jobs (Non Crime in Progress jobs). [RO1]



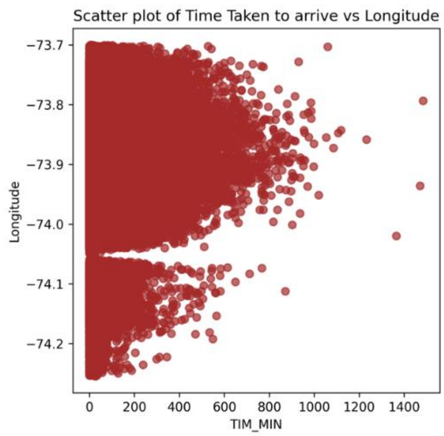
Another major factor when it comes to response times is the weather. To see the effects of temperature and response times. From the scatter plot we can deduce that when the temperature has been on both the extreme ends that is in and around 15 C and 35 C the response time has been less comparatively. This may be due to less traffic in cases of extreme weather which hasn't been accounted for in this project. [RO1]



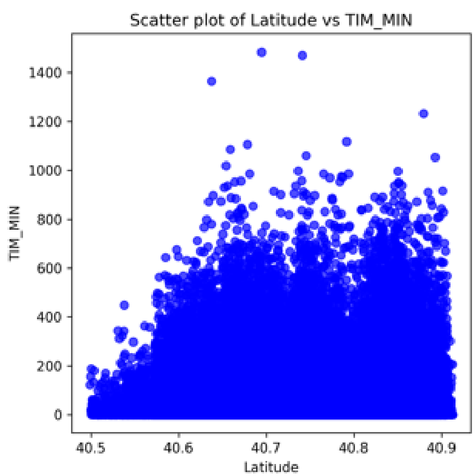
In the scatter plot below, we see a plot of time taken to arrive at scene in minutes vs Rain in millimeters. We can infer that irrespective of the rain the dispatch units have reached quickly. Dispatch units have taken hours together to arrive at scene at times even when there was no rain and in cases with heavy rain they have still managed to reach faster. [RO1]



In the second plot we see that regions lying between longitude of -74.05 and -74.25 have faster response time than those in the region with longitude -73.7 to -74.05. One more inference is the region at and around -74.05 has had dispatch units arrive in little to no time, which is a commendable achievement by dispatch units in that area. [RO1]



The scatter plot for most of the quantitative features relative to latitude and longitude don't give any specific or useful information except for the one below. [RO1]



The above plot is a latitude vs time taken to arrive on scene in minutes scatter plot. This indicates that regions within 40.5 and 40.6 latitude have far lesser response times than the others, which of course is a commendable job by the emergency services that are providing their services in that region. [RO1]

3 Methodology

The first machine learning algorithm applied was XGBoost, as a regression model to predict time taken by the dispatch units to reach the scene. This algorithm was chosen particularly for its ability to handle complex non-linear relationships. The dataset after going through pre-processing which included feature scaling, one hot encoding and outlier analysis was fed to the algorithm. XGBoost can handle both numerical and categorical variables in the dataset. Scaling isn't critical for XGBoost, and it might learn relationships between rain and hour with respect to time taken. The model was trained with fine-tuned hyperparameters such as number of boosting rounds, learning rate and maximum tree depth. For further analysis, its feature importance scores also can be derived to jot down the features that affect response time predictions making it a valuable tool. [RO2]

Alternatively, three other approaches were implemented as well to analyze the results and to see how they vary.

The first alternative approach was implementing Ridge Regression. In some sense it does serve the purpose of that of a baseline model. It captures linear relationships well and gives us a sense of direction whether implementing linear models is the way to go or complex non-linear models need to be implemented like XGBoost and Random Forest. With respect to this dataset, the ridge model would not be able to capture ay relation or correlation between temperature and rain with respect to the time taken. This approach is very inexpensive and a perfect way to test if the data consists of linear relationships or not. It was chosen for its ability to handle multicollinearity and since it prevents overfitting by using the L2 regularization technique. The model was fine-tuned with hyperparameters such as regularization strength to have a balance in bias and variance. [RO2]

The second alternative approach was Random Forest Algorithm which was applied to predict the time taken as well by following the same procedures of preprocessing as done in XGBoost. The model was trained with fine-tuned hyperparameters such as number of trees and number of CPU cores usage. By using the number of CPU cores as a parameter, it ensure efficient resource utilization since we are dealing with very large datasets. It can capture relationships like how precinct number and hour affect time taken. Scaling of features wasn't necessary for random forest as well. It is considered to be a simpler and less computation heavy approach compared to XGBoost. Unlike XGBoost it doesn't depend on boosting and relies on bagging. It is less prone to overfitting as well. In this approach it averages the predictions from many trees which in turn reduces the variance. It is a faster and easier to implement model. [RO2]

The third and final approach was to implement stacking. In this approach the aim is to improve the prediction on time taken by stacking multiple models on top of each other to extract the best that each model has to offer. It combines the strengths of XGBoost, Random Forest as well as Ridge Regression to produce a more accurate and robust model. This approach provides more diversity in handling various types of data. It reduces bias and variance and handles the weakness of certain models when implemented individually. As this dataset has mainly non-linear relations, whatever the meta model captures and is underperforming, the stacking model automatically relies more on the base learners. [RO2]

The approaches used cover three pivotal topics such as bagging (Random Forest), boosting (gradient boosting in XGBoost) as well as stacking.

4 RESULTS AND DISCUSSION

The models gave varying and useful results. The Ridge Regression model was the worst one among the lot with high mean absolute and root mean square errors. Moreover, it has a very poor R squared value which almost deems the model not useful for this data. It explains very little variance and doesn't seem to have gotten along with the complex nature of the target variable.

On the other hand, the random forest model outperformed XGBoost in all error metrics I have chosen which implies that it has captured the essence and complex nature of the target variable well. It shows more variance as well and can understand more than XGBoost (the relationship and patterns) [RO3]

The stacking model as stated before harnesses the best of both the models and gives much better mean absolute error, root mean squared error and r squared values. This implies that it is good at penalizing large errors as well as captures the relationships and trends very well. This is indicative of high predictive power from the model. Despite having a higher Mean Absolute Error than Random Forest, I would say this is the best model due to its Lower Root Mean Squared Error and Higher R squared value. It combines both Random Forest as well as XGBoost and uses Ridge as a meta model and increased the variance that has been explained to the extent that the complexities of predicting time taken have been understood. It captures large errors as well as has better performance. [RO3]

Model(Evaluation)	MAE	RMSE	R^2
XGBoost (Model)	3.8217	2.3006	0.0959
RandomForest (RF)	3.6232	2.3021	0.0936
Ridge Regression	4.0079	2.3273	0.0532
Stacking - Model	3.6529	2.2793	0.1289

The errors are quite significant and the reason as to why this might have happened is due to the large variability in the target variable. In a real-world scenario, there are cases where the police would take less than 5 minutes and more than 30 minutes to arrive at the scene, and the data being used is of the real world and these extremes in the values of time taken might be the reason. The model might not have caught some complex relationships between the features.

Using custom input and vary few values it was found that the precinct number[RO4] highly influences the time taken which is sound, as the further the precinct which responds is, the longer it takes for them to arrive at the scene. The second feature that influences the target variable is rain[RO4] in millimeters. It is a known fact that weather conditions do make commute difficult and faster in few cases. It is found that there is an increase in the time when we increase the rainfall level. With the increase in rainfall the time increases with a small increment. The model, however, didn't capture anything at an extreme since the dataset used is of the month July and there hasn't been any storms or heavy rainfall and hence failed to catch high and abnormal values. The maximum it increased was by 25%.

This might not be a feasible method as having such high errors and less correlated factors is not an ideal case to implement in the real world. Emergency response time must be as optimized as possible as it is a matter of public safety. But with more factors like traffic data and more geospatial analysis we could achieve more accurate and useful results.

5 CONCLUSION

In this project, machine learning models, including Ridge Regression, Random Forest, XGBoost, and Stacking, were applied to predict response times (TIM_MIN) based on various geospatial and environmental features such as Latitude, Longitude, temperature_2m (°C), rain (mm), and others. Each model contributed to the task at hand, with Stacking combining their strengths to improve accuracy. The work highlights the importance of leveraging both numerical and categorical data, as well as effective feature engineering and scaling, to capture complex patterns influencing response times.

With access to additional geospatial data, such as real-time traffic conditions, road network density, or incident locations mapped against urban layouts, this work could be extended to significantly improve predictions. Incorporating interactive maps to visualize predicted response times and dynamically adjust inputs (like weather changes or traffic jams) would add more practical value.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my professor, Ruben Acuna, for their guidance, encouragement, and insights throughout this project. Their expertise and feedback were vital in shaping the direction of this work. I am also deeply thankful to my teaching assistant, Mitchell Buckner, for their consistent support, timely clarifications, and technical assistance. Additionally, I extend my appreciation to the authors of related research whose work provided a strong foundation and inspiration for this study.

REFERENCES

[1] The Evolution of Emergency Calls https://www.tylertech.com/resources/blog-articles/the-evolution-of-emergency-calls

[2] '911 Systems Disrupted in at Least 3 States https://www.nytimes.com/2024/07/19/business/emergency-911-calls-tech-outage.html

[3] As Vallejo police force shrinks, 911 response times soar https://openvallejo.org/2024/09/23/as-vallejo-police-force-shrinks-911-response-times-soar

[4] A Review of Emergency Incident Prediction, Resource Allocation, and Dispatch Models by Ayan Mukhopadhyay , Geoffrey Pettet , Sayyed Vazirizade , Yevgeniy Vorobeychik , Mykel Kochenderfer, Abhishek Dubey https:doi.org/10.1016/j.ap.2021.106501

[5] Spatiotemporal Data-Driven Multiperiod Relocation Optimization of Emergency Medical Services by by Xinxin Zhou, Yujie Chen, Yingying Li, Bingjie Liu and Zhaoyuan Yu **https://doi.org/10.3390/ijgi12070269**

[6] 911 Analysis: Call Data Shows We Can Rely Less on Police https://vera-institute.files.svdcdn.com/production/downloads/publications/911-analysis-we-can-rely-less-on-police.pdf

[7] Re-assessing measurement error in police calls for service: Classifications of events by dispatchers and officers by Rylan Simpson, Carlena Orosco doi: 10.1371/journal.pone.0260365

[8] Big Data Analysis Examining the Meaning of Boston's 911 Call Data and Implications for Public Policy by Dr. Jack McDevitt, Keller Sheppard, and Stephen Douglas

[9] Analysis and Prediction of 911 Calls based on Location using Spark Big Data Platform by Ketki Deshpande, Shruti Pandey, Sukhada Deshpande https://hpcf-files.umbc.edu/research/papers/IS789_Project_Report-HPCF-2019-29.pdf