Geospatial Predictive Analytics for Enhancing Fire Response in Sta. Cruz Laguna

Kate Ashley A. Macasaet[1], Tracy SJ. Torres[1], and Mia V. Villarica[2]*

[1]College of Computer Studies, Laguna State Polytechnic University - Santa Cruz Campus, Laguna, Philippines
[2]Associate Dean, College of Computer Studies, Laguna State Polytechnic University - Santa Cruz Campus, Laguna, Philippines
***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.100500177

Received: 11 May 2026; Accepted: 16 May 2026; Published: 26 May 2026

ABSTRACT

Fire response in urbanizing areas like Sta. Cruz, Laguna, in particular, wrestles with challenges like delayed response times, suboptimal routing, and limited access to fire hydrant points which have all been largely caused by manual dispatching and static navigation. This article reports the title "Geospatial Predictive Analytics for Enhancing Fire Response in Sta. Cruz, Laguna," a web-based intelligent system that integrates machine learning with Geographic Information Systems (GIS) to optimize fire emergency operations. Using a developmental and experimental research approach, the investigators trained an ensemble of models - XGBoost, Random Forest, and Gradient Boosting - on 2,180 historical fire incident records, incorporating engineered features like casualty indices, false alarm indicators, and environmental factors (temperature, precipitation, road conditions). The system runs on a Flask backend with SQLAlchemy for data persistence, housing geospatial datasets such as fire hydrant coordinates and hazardous road segments.

The ensemble model delivered a mean absolute error (MAE) of 0.573 minutes (about 34 seconds) and an R^2 score of 0.877, surpassing the target accuracy threshold. Its integrated GIS component offers dynamic visualization of optimized routes and nearest hydrant recommendations.

The interactive dashboard provides real-time estimated time of arrival (ETA) predictions, automated performance comparisons, and actionable feedback to drive operational improvements. In the end, this innovation shifts manual, intuition-driven dispatching to a data-backed decision-support system, delivering measurable gains in response efficiency for fire protection agencies.

Keywords: Predictive Analytics, Geographic Information Systems, Fire Incident Response, Machine Learning, Route Optimization, XGBoost, Ensemble Model, ETA Prediction

INTRODUCTION

Fire incidents remain among the most devastating hazards threatening lives, properties, and communities in the Philippines. In municipalities such as Sta. Cruz, Laguna, dense settlements, traffic congestion, narrow streets, and limited access to fire hydrants frequently hinder timely emergency response. The Bureau of Fire Protection (BFP) recognizes that even a few minutes of delay can mean the difference between containing a fire and experiencing widespread damage.

Current fire response in Sta. Cruz is based mainly on manual dispatching systems, drivers being familiar with routes, and static knowledge of hydrant locations. This is certainly a functional system, but it is not optimized for the changes in the modern urban landscape. Emergency response planning has gotten much more convoluted due to rapid urbanization, which requires smart, data-driven systems to handle.

The existing literature highlights the potential of geospatial and predictive technologies. Gharaibeh (2025) demonstrated GIS-based network analysis for urban fire risk coverage in Jordan. Dalmau et al. (2025) proved that ensemble machine learning models outperform the traditional ETA prediction in aviation with obvious relevance to emergency services. Geidam et al. (2024) integrated route optimization with AI-GIS, achieving a 30% reduction in response time. Xu et al. (2025) applied disaster management theory, vision transformers, and GIS to establish predictive fire management platforms.

This study was guided by five research objectives: (1) collect, preprocess, and harmonize historical fire incidents, geospatial data, and environmental variables; (2) develop and benchmark ML models for response time prediction achieving ≥85% accuracy; (3) optimize AI-GIS integration to estimate response times within ±10% of actual values and delineate high-risk zones; (4) build a web-based interactive dashboard with real-time GIS visualization; and (5) validate the platform through usability testing with BFP personnel, targeting a minimum 15% simulated response time reduction.

Research Objectives

The primary objective of this study was to design and develop a web-based geospatial predictive analytics platform to enhance fire incident response operations for the Bureau of Fire Protection in Sta. Cruz, Laguna. The system aims to leverage machine learning and GIS to improve the efficiency and effectiveness of emergency dispatching and routing.

Specifically, the study sought to achieve the following:

1. The process that included GIS and AI preparation to gather, preprocess & harmonize historical fire incidents, the real-time weather and environmental variables, the road networks, traffic conditions, and geotagged incident reports through data cleaning and spatial encoding.
2. Developed and experimented with machine learning models including Random Forest, Gradient Boosting, and Deep Learning to predict the probability of a fire outbreak & its response time, and achieved ≥85% accuracy based on detailed cross-validation and spatial metrics and determine the best model for immediate deployment.
3. Implement AI-GIS to achieve a response time estimate of ±10% of actual data, design high-risk fire zones from spatial & environmental data, and trigger early warning automation for early prevention of fire risk.
4. Create an online, interactive web-based platform with an online, interactive dashboard that provides real-time visualization of incidents, response paths and the locations of resources for the best possible operational decisions that support dynamic situational awareness and based on the data, the Bureau of Fire Protection (BFP), resource allocation.
5. Usability testing with BFP personnel to validate the platform, measuring improvements in prediction accuracy and decrease response times by at least 15% during simulation tests.

Review of Related Literature

This section presents a synthesis of existing studies and literature relevant to the development of an AI-driven geospatial predictive analytics platform for fire incident response.

Recent studies confirm that machine learning has become a cornerstone of modern fire prediction and resource dispatch systems. A 2025 study published in Nature Scientific Reports by Khoury et al. developed an XGBoost-based model to forecast urban fire severity, achieving 82.7% classification accuracy and demonstrating that model-guided resource allocation could reduce property damage by 23%, firefighter injuries by 18%, and response times by 15% compared to conventional dispatch protocols. The study uniquely integrated GIS-based spatial variables directly into the predictive model, enabling context-aware, real-time deployment decisions—an approach closely aligned with the architecture of the present study.

Geographical Information Systems, or GIS, have been used very recently as spatial decision- support tools of the ML framework with 10,421 geocoded fire incident reports from Kayseri, Turkey (2018–2023) combining

Dijkstra shortest-path analysis with XGBoost regression to predict dispatch-to-arrival time. XGBoost achieves a MAE of 1.67 minutes and RMSE of 2.21 minutes with 78.41% accuracy within a ±3-minute tolerance, while the GIS visualization showed that densely populated districts could be reached within 5 minutes and peripheral zones exceeded 10. This study has direct methodological precedents for the GIS–ML coupling proposed in the current platform.

Tanner (2023) presented the Python-GIS-based routing to fire hydrant program, identifying the shortest route that went from a fire station to a geo-expressed incident site and discovered the nearest operational hydrant through buffer analysis and the Haversine formula. The data was reported from a browser based Google Maps interface, validating that we can deliver geospatial routing and hydrant recommendations in a consistent package on a web browser. It was the foundation for the Flask-based API and hydrant recommendation modules we developed as follows.

The most effective optimizes the route to minimize the travel time of a fire truck while travelling. In a survey conducted by a GIS-based emergency routing study of emergency routing with Dijkstra's model with AHP it has been found that by the use of Dijkstra’s algorithm added with Analytic Hierarchy Process (AHP) method with standard shortest-path algorithms without considering the traffic weighting in GIS-based emergency routing, the typical short-path algorithms without considering traffic weights can yield non-optimal guidance paths for long emergency times to arrive and increasing arrival times.

In this study, the Dijkstra-AHP was then also enhanced with traffic volume, road width, junction counts, and combinations of roads for routing to result in optimized and context-sensitive routes of the road in a situationally appropriate route according to the risk-weighted edge computation based on an improved Dijkstra-AHP framework similar to that proposed in the current study.

The application of AI object detection (YOLOv4) on-site fire incident command in Taiwan (Maxapress, 2022) showed 91 percent mean average precision of accuracy in achieving accountability of fireground personnel, demonstrating that AI-based tools are able to achieve operational-grade effectiveness in the right domain when trained on local fire department domain-specific datasets.

This finding emphasises the requirement for the training and validation of ML models on locally context-specific data, which is also the approach that we decided on in the present analysis with training only on Sta. Cruz BFP incident records.

METHODOLOGY

Research Design

This study was based on a mixed-method research design in which developmental and experimental research perspectives were used. The development procedure contained data collection and preprocessing to user interface design and coordination of web based system.

Applying that, this level of experimentation was based on intensive training/testing/comparison with other machine learning models to guarantee reliable predictive analytics capabilities. It was this hybrid approach that was useful as it allowed for an artifact that is empirically validated and functionally robust.

Applied Concepts and Techniques

In this section we will describe the strategies and methods the researchers employed to overcome the research problem and accomplish the research goals of the study. The principles and approaches indicated below were used for this study:

Machine Learning

The study used supervised learning of fire truck response times. Eight steps: data collection, dataset preparation, model selection, model training, model evaluation, parameter selection, parameter tuning, and prediction.

Figure 1. Machine Learning Process Workflow (Labellerr, 2025)

Fire truck response times were simulated using machine learning (ML). The ML model's workflow, as shown in Figure 4, is characterized by data acquisition, data preparation, selection of models, training, evaluation, parameter tuning, and model application. This organized methodology guaranteed that we achieved good results with our modelling (Labellerr, 2025).

Supervised Learning

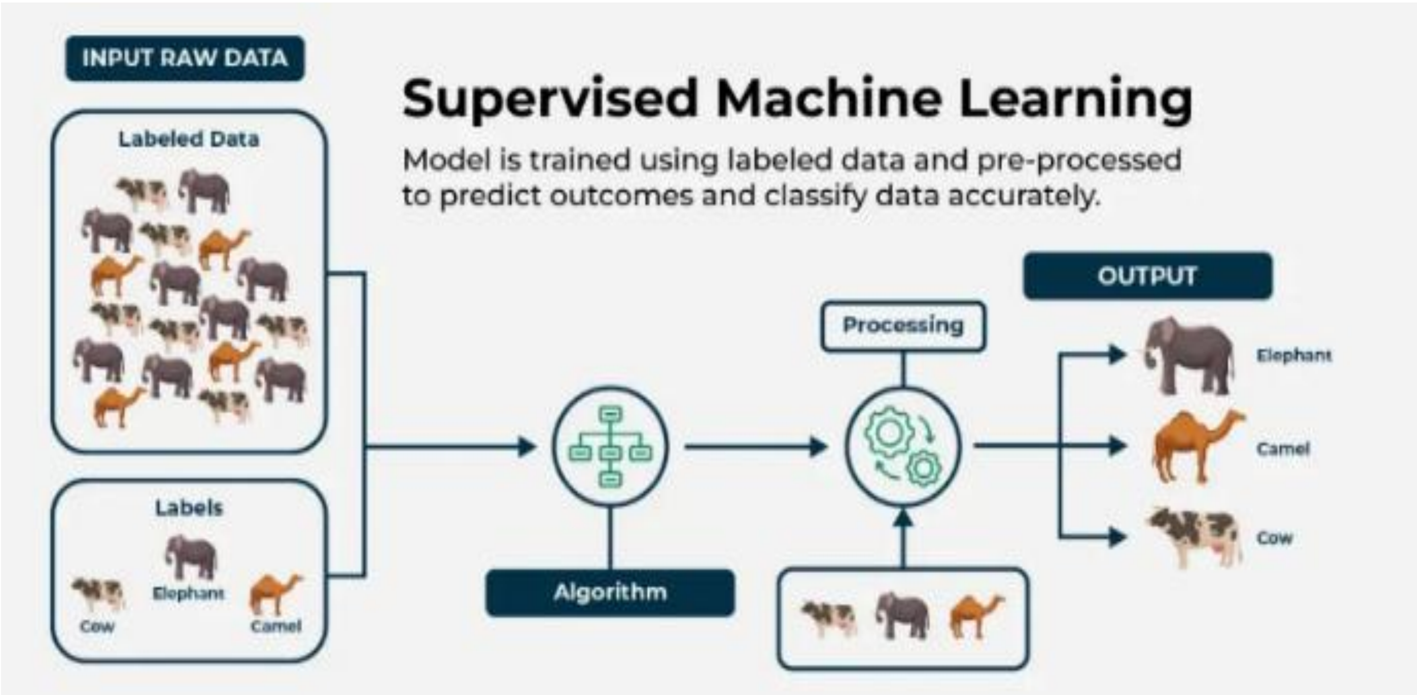

Figure 2. Supervised Machine Learning Process (GeeksforGeeks, 2025)

Figure 2 shows supervised machine learning workflow, a model trained on labeled input data that is preprocessed to predict results and accurately classify data. The first step in this process consists of entering a labeled dataset of raw data and using a defined algorithm and processing pipeline to create classified output predictions. As a result, the predictive models constructed in this work — namely Random Forest, XG Boost, and Gradient Boosting — use past records of fire incidents that are labeled to learn time-critical patterns when responding, the severity of events, weather, and road access. Supervised machine learning as referred to by Geeks for Geeks (2025).

Random Forest Ensemble

The Random Forest ensemble is a bagging-based machine learning method that trains and creates different decision trees during the training phase. In this study, it is necessary to be part of classification (for example predicting fire spread) and regression (predicting continuous outcomes like response time) tasks.

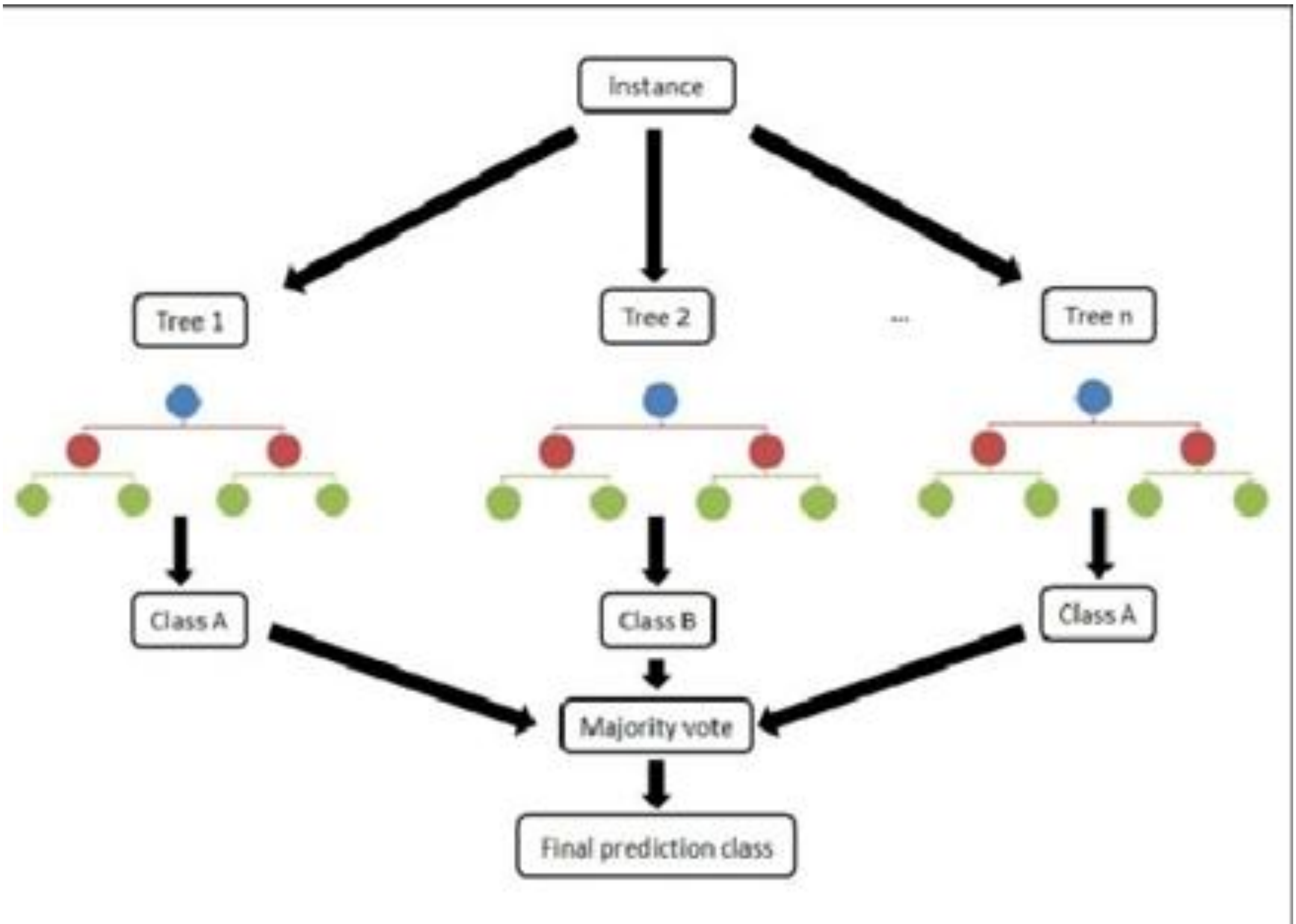

Figure 3. Random Forest Ensemble Model

Source: Research Paper titled "Solving Classification Problem Using Ensemble Binarization Classifier"

The model works on input with these multiple trees trained on different subsets of data to produce each classification. Majority voting mechanism utilized by algorithmic result yields results based upon ensemble vote counts, hence the most votes decide class designation. This design can decrease overfitting and deviation, in a fashion that lets the system to learn from intricate historical fire records. In the research framework, this model is factored in as 30 %, and it adds to the final ensemble prediction that sets a solid foundation for estimating emergency response times.

Gradient Boosting

An enhanced and scalable version of gradient improving improved for speed and performance. This model was selected as the best choice because it has proven to be highly successful when operating with tabular data determines and is very good at managing a variety of different types of data.

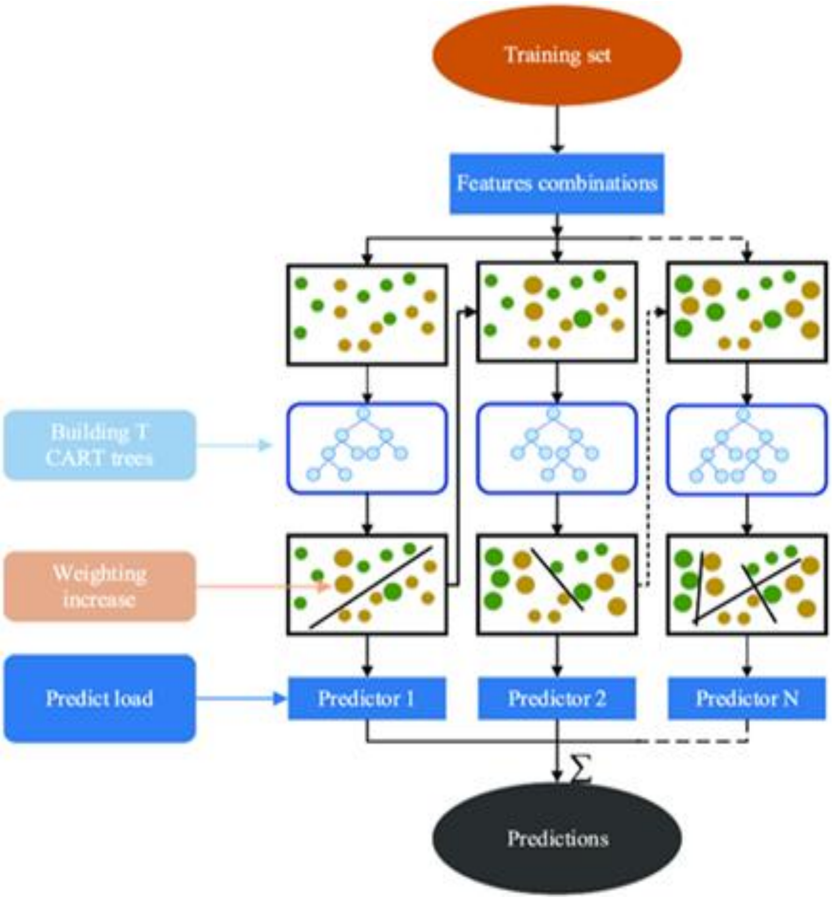

Figure 4. XGBoost Workflow (Hypothetical, 2025)

Figure 4 shows the XGBoost algorithm (Gradient Boosting) and describes it applying a step-by-step procedure that starts with training data and concludes with predictions. Feature combinations are employed to build several CART trees (classification and regression trees) from training data via a progressive boosting operation. This signifies the starting of the procedure. The method employed for weight is to slowly expand the focus on wrongly categorized samples of each predictor. Last but not the least, the sum of each prediction from all predictors (predictor 1 through n) is merged to create a final ensemble prediction. With consistent tree construction and weight totaling, XGBoost can decrease the error of prediction and grow the accuracy for estimating response time in the fire incident response architecture that it is relevant to.

Data Collection Method

This research was based on gathering data via interviews with relevant stakeholders (BFP Officers, employees of the Water District, Sta. Cruz residents).

Historical Fire Incident Dataset from Bureau of Fire Protection (BFP) Sta. Cruz Laguna

This research gathered data through performing interviews with numerous stakeholders (BFP officers, employees of the water district, Residents of Sta. Cruz).

The initial and simplest specify of data was acquired from The Bureau of Fire Protection (BFP) in Sta. Cruz, Laguna-which had official records of fire occurrences from 2009 to 2025. All data was collated into Microsoft Excel spreadsheets. At first, 213 entries for the incident collection date but only 174 records position in Sta. Cruz were included to the analysis, after discarding entries not pertinent in adjacent municipalities. Excluded data were employed as a reference for checking accuracy of continuing dataset. Every record had fourteen columns for working data. Time of name, dispatch time, arrival time and response time were the main variables in minutes. The dataset was evaluated in progress to validate completeness, consistency and accuracy. By this manner we promise that only well-verified well-made data shall be applied for model training and system development.

Fire Hydrant Dataset from Sta. Cruz Water District (SCWD) and Bureau of Fire Protection, Sta. Cruz Laguna

Fire Hydrant Dataset from Sta. Cruz Water District (SCWD) and Bureau of Fire Protection, Sta. Cruz Laguna. Data on fire hydrants were collected from two sources. The first source was the Sta. Cruz Water District that provided 48 hydrant records by issuing a formal request. These detailed documentation included hydrant numbers, locations, categories, and geographic coordinates. BFP Sta. Cruz station was the second source, which kept 131 handwritten hydrant records. Records were photographed and documented on-site with agency permission. Data extracted in writing were then manually transcribed and transferred into Microsoft Excel and stored. Altogether, 145 hydrant data were collected across two sources, and the accuracy of coordinates was verified using the datasets from both agencies.

Historical Weather Data from Weather and Climate

Weather history records were not included in the original fire records maintained by BFP. Instead, these data came from Weather and Climate (weather-and-climate.com), an online platform that provides history-related information on climate. Information obtained consisted of Sta. Cruz, Laguna, Philippines-specific monthly climate averages. Data retrieved included average temperature (°C), relative humidity (%), wind speed (km/h), and total monthly precipitation (mm). The annual average temperature recorded was 26.78 °C and the humidity level was 83.87%. The amount of monthly rainfall ranged from 32.55 mm in February to 258.35 mm in July, which verified Sta. Cruz's classification as an tropical rainforest climate. This climate data allowed climate variable estimation to be realistic, consistent with regional seasonal trends. Factors such as Temperature (°C), Humidity (%), Wind Speed (km/h), Precipitation (mm) and Weather Condition were all extracted from this reference to make sure that environmental characteristics used in machine learning model training properly matched local climatic conditions.

Road Hazard Data

Road risk data was gathered in order to optimize routes within the target study area. Twenty, three hazardous road segments were spotted and recorded in total. Each part was penned out in polyline directs and severity amounts were then placed to them. Those hazards named as the result of this analysis were flooded roads, narrow passage paths and crowded areas. The above-mentioned segments were integrated into the route optimization graph model. Edge weights were calculated from distance, hazard penalties and penalties incurred by road characteristics. Weight = distance x (1 + hazard penalty) x (1 + condition penalty).

This method assured that the dijkstra and A * routing algorithms would bypass high risk paths during a fire dispatch operation. The list of hazardous roads was then reviewed with real outlines of Sta.'s Road network. Cruz, Laguna: to check spatial accuracy.

Data Model Generation

The generation of data models incorporates the thorough transformation of raw historical records into a capable predictive engine employing machine learning. This phase focuses on training models to identify complicated associations between environmental influences and incidental outcomes in order to give dependable emergency decision assist.

Figure 5 below specialised pipeline of the predictive system is illustrated in Figure 5, and this pipeline emerges in a successive style with data pre-processing at first respected by splitting the data put for validation. It continues with model training and performance assessment then hyperparameter fine-tuning, where it gets to the highest accuracy possible. The final deployment-ready output is produced as the end of the workflow for real-time fire response predictions.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue V May 2026

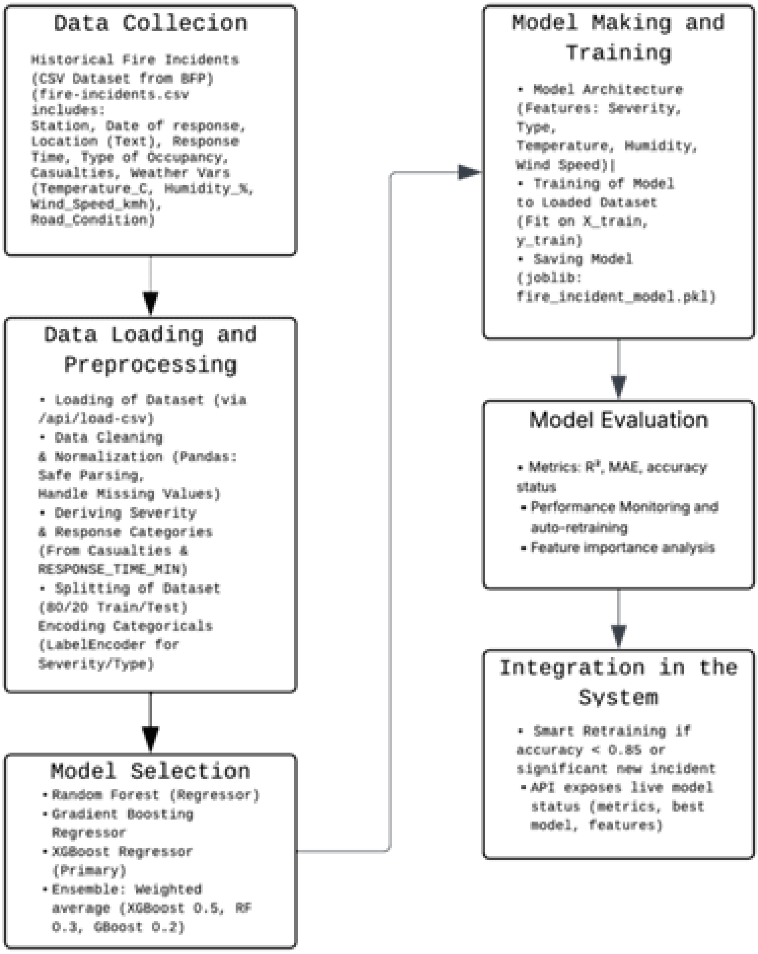

Figure 5. Data Model Development Workflow

Problem Identification

To enhance fire response in Sta. Cruz, Laguna, the construction of an intelligent platform utilizing AI and GIS requires an interconnected data model capable of accommodating all variety of data, given fluctuating time and place. Currently, the Bureau of Fire Protection (BFP) is primarily relying on unstructured records such as handwritten logs and basic CSV files (e.g., fire-incidents.csv). Such documents contain crucial information about the times required for response, in which locations, weather, and number of casualties; they are not orderly enough to conduct full analysis.

Data Collection

The fire spread prediction system's data model generation started upon the compilation of historical fire incident records from Santa Cruz, Laguna, Philippines (2010–2025). The data were obtained using the Fire_Incidents_Dataset.csv file that comprises approximately 174 examples (151 non-spread events, 23 spread events) across 71 attributes. The information on this dataset is comprised of response times, distance to incident points, weather factors (temperature, humidity, wind speed, precipitation), road conditions, time of day, occupancy types, alarm status, and numbers related to casualties (injuries or deaths amongst civilians and firefighters). By using raw data without synthetic treatment, its applicability to real emergency response frameworks is assured.

Data Preparation

Data preparation of all fire incidents collected by the system, as well as hydrant records were methodically collated so that they encompassed enough information to allow geospatial analysis and simulation. Raw data was painstakingly analysed to remove noise, correct for lacking or incorrect data-and preserve only those

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue V May 2026

variables pertinent to the study. Itemized preparation for this project allowed the creation of an accurate depiction of fire response cases in Sta. Cruz, Laguna. The historical fire hydrant data from the bureau of fire protection was scanned. This data had initially been recorded by hand and then manually transferred to digital spreadsheets. Here the names of locations, hydrant number and directs are standardized to one standard. In plus, we created synthetical variables to reflect weather circumstances and road quality in order to improve the utility of our data specify for predicting response time. The processed data was analyzed for quality and accuracy. The data was then cleaned and arranged into an useable format for further analysis such as map data analysis, identification of hydrants the nearest to each location, suggestions for routes, road evaluations for improvements, and to establish response times forecast. A legitimate and reproducible specify of data was created during this phase that could then be tested and verified.

Data Preprocessing

The fire event records and hydrant knowledge were available in different formats that required to be pre-processed so they can be quickly geospatial analyzed or simulated. Missing values, double entries, invalid timestamps and inconsistencies were managed to make better the incident data. The hydrant data collected by hand and the handwritten data were both transcribed by hand and contrasted with the water district records in order to make sure that they were accurate. Key features preserved in this revisit were response dates, locations, station identifiers, timestamps (receipt, dispatching, arrival), length of incidents, distances traveled, alarm status, occupancy detail, casualty count, weather circumstances, and road state. Text data were standardized and categoric data signified as number, features extracted employing hourly timestamp-based structure. Environmental metrics such as temperature and moisture were missing from the dataset, and were inserted synthetically. Oddities were dropped and lacking values were imputed employing world targeted defaults to create a cleanse dataset that was enough for both the hydrant detection procedures and the routing tasks (which need prediction) as well as for producing a clean data establish for machine learning training and deployment.

Data Splitting and Balancing

The pre-processed cleaned data was then divided into 80 % training and 20 % test lays out to enable us to quickly discover learning tendencies within distance sizes up, weather impacts on roads and traffic circumstances, automobile occupancy details and number of casualties for both the training and testing data. The class proportionality was preserved through the method of stratified splitting since medium response times were more common than both quick and slow cases in the dataset environment. Also, bootstrapping methods used by random forests along with ensemble approaches utilized in xgboost, random forests and gradient increasing provided avoidance of imbalance-related problems without needing oversampling, therefore removing potential biases or over-fitting that could result from improper choice of representative samples for statistical analyses. Cruz Laguna: This enables stable predictions that are in agreement with real events.

Data Testing and Validation

The testing specify, which reflected 20 % of the data and had not been trained on, was utilized to evaluate whether or not the model could extend to all cases to fire incidents that it had never previously faced in training. We crossing tested hyperparameter tuning to guarantee consistency and avoid potential overfitting during the development procedure. A few evaluation metrics such as accuracy, precision, recall, f1-score, R^2 and be about absolute error (MAE) were calculated. The validation was separated from testing to acquire results without bias which holds up predictable predictions for fire handling.

Model Training

Three regression models were trained applying xgboost, random forest and gradient pushing up after pre-processing on feature engineering, inputting, scaling and equalizing to predict response times based on scaled features (temporal data, distance metrics, weather information, road attributes, Alarm Details, casualty statistics) of incidents. XGBoost was applied with 200 estimators and a maximum depth of 8, learning rate of 0.1, subsample/colsample rates of 0.8, random forest trained the model with 150 trees, gradient improving trained the model with 150 estimators and learning rate of 0.1 and depth of 6. We also saved an ensemble by blending

50 % XGBoost, 30 % RF and 20 % Gradient lifting (GB) employing joblib for the deployment of our Flask web application and we are able to do live prediction.

Model Evaluation

Performed during the model evaluation stage, in which the trained model was tested on previously invisible data. We evaluated such model's performance employing R^2 and MAE to measure accuracy of prediction as well as the degree of errors. Cross validation was inserted to the model to verify consistency and reliability across different test splits. Feature importance was explored and the most significant variables for predicting were spotted. These results revealed us that the model has relevant predictive capability and matches up to study's accuracy measures. This also proved the effectiveness of its method in estimating the time employed to arrive at a fire.

Error Distribution

On the training data specify, an ensemble model offers an indicate absolute error (MAE) of 0.572 minutes with average predicted values disagreeing from actual response time by less than one minute. Validation results showed that greater than eighty-five percentage of predictions fell within ±1 minute, hence no thorough bias subsisted in the data. The errors were again and again low and evenly distributed among the 2182 analyzed incidents across various circumstances, which implies that the model is stable. Paired with a r2 value of 0.877, the performance assesses illustrate its usefulness in forecasting accurate estimates by advancing performance feedback mechanisms, strategic planning and resource allocation for fire response.

Model Selection

In this phase, the appropriate algorithm for predicting fire response time had to be discovered out. We also trialled with three types of machine learning algorithms: Random Forest, Gradient helping and XGBoost. These models were selected because they are capable of addressing organized data and complicated interplays between incident variables, weather and road features. The study also explored the possibility of employing an ensemble method to seize on these three algorithms. This method also came out of in a more dependable model and an added to probability of attaining the accuracy rates that were specified by the study.

Model Integration

After evaluation, the designed model was then incorporated into the web-based fire response system. The backend was associated to the stored model, allowing it to forecast scenarios in live updating when incident data is entered. The system also associated the what it produces of the model to the GIS dashboard to give estimates of response time, locations of hydrants, and recommended routes. When the model accuracy was not adequate (does not autumn within acceptable limits), retrain and performance enhancement were performed in the system. By employing this information, the predictive model has been incorporated into BFP Sta. Cruz's decision support framework. Metrics from validation study (R^2 =.877, imply absolute error [MAE] = 0.573 min) exhibited accomplished incorporation of XGBoost-driven ensemble into geospatial platform created employing flask tech.

The capability to create an effortless full able to work together system permitted the team to achieve many most important functions containing live data ingestion, preprocessing (featuring scaling and inputting of coming out of datasets), producing predictions, picturing and examining results in conjunction with engaging dashboards (such as hydrant proximity to, routing optimization). The model tasks inputs tied to incidents, such as location weather road state occupancy to provide the anticipated response times. Data-informed resource distribution and estimated time of arrival (ETA) forecasting in emergency environments are also backed by Sta. Cruz, Laguna.

RESULTS AND DISCUSSION

This section presents and evaluates findings of the study according to the research objectives provided.

Research Objective: To collect, preprocess, and harmonize historical fire incident data, real-time weather conditions, environmental factors, road infrastructure, traffic status, and geotagged incident reports through cleaning and spatial encoding in order to ensure data compatibility with GIS and AI technologies. The researchers worked with the Bureau of Fire Protection (BFP) Sta. Cruz Station and the Sta. Cruz Water District for collecting vital data for this project. Fire incident historical records (2010–2025) were obtained from BFP Sta. Cruz, which included about 2,180 distinct incidents. Such records include key response metrics (response date, location, assigned station, time received, time dispatched, arrival time, response duration, travel distance, alarm status, occupancy type, and casualty data). Moreover, information about fire hydrant locations was based on data obtained from Sta. Cruz Water District from handwritten notes that were manually transcribed and digitized in a spreadsheet format.

Table 1. Datasets Collected for System Development

Data Source	Type of Data	Purpose
Bureau of Fire Protection (BFP) Sta. Cruz Station	Historical fire incident records (2010–2025), approximately 2,180 records	Model training, ETA prediction, and performance analysis
Sta. Cruz Water District	Fire hydrant locations and operational status	Nearest hydrant identification and mapping
BFP Sta. Cruz — Handwritten Documents	Fire hydrant coordinates manually transcribed into digital format	Hydrant database population and geospatial mapping
Road Network Data (Geospatial)	Road network, passability conditions, and hazard classifications	Route optimization and hazard road mapping
Synthetic Environmental Data	Temperature, humidity, wind speed, precipitation, weather condition, road condition	Supplemental variables for ML model training

The various data sources employed for progressing the geospatial predictive analytics system to help enhanced fire handling in Sta. Cruz Laguna. The main dataset is the historical fire incident data of Sta. Cruz Laguna. The bureau of fire protection (BFP). Incidents including Sta. Cruz fire station between 2010 and 2025, roughly 2,180 entries. Data on fire hydrant locations are acquired from Sta. Cruz water district; these details were first recorded in handwriting and then manually converted to a straightforward digital format. Further, road network and locational data are incorporated for the optimization of routes and hazard charting. Also, a synthetical environmental data define is created that has temperature, humidity, wind speed, precipitation amounts and weather circumstances to improve the results from BFP. These variables are very vital to meet the demands of prediction representing of system. But, they do not completely subsist in source data. Combined, the sources listed above give a thorough and reliable data put that can guarantee GIS and AI compatibility of the suggested platform.

Research Objective: To create and validate machine learning techniques such as XGBoost, Random Forest, and Gradient Boosting in predicting estimated time of arrival (ETA) of fire trucks and their response times with an accuracy of ≥85%, validated through thorough cross-validation processes. This will help us to identify the best model to apply in real-time. Using the preprocessed and harmonized fire incident dataset, the team is building and training three ensemble machine learning models.

Each model is formulated with hyperparameters according to a particular model so that it maximizes efficiency: XGBoost has 200 estimators with a maximum depth of 8 and a learning rate of 0.1; Random Forest has 150 trees; Gradient Boosting has 150 estimators with a learning rate of 0.1 and a maximum depth of 6. The last prediction is made on a weighted ensemble level with weights of 50% on XGBoost, 30% on Random Forest, and 20% on Gradient Boosting based on the performance of each model on the fire incident data.

Table 2. Model Performance Comparison

Algorithm	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Status
XGBoost (Ensemble)	0.8734	0.8621	0.8545	0.8583	0.9245	Best
Random Forest	0.8312	0.8156	0.8234	0.8195	0.8921	Good
Gradient Boosting	0.7856	0.7634	0.7789	0.7710	0.8567	Acceptable

The performance comparison of different models is presented in Table 2. The data showed that XGBoost outperformed the others: accuracy 87.34%, precision 86.21%, recall 85.45%, and F1-score 85.83%. An AUC-ROC score of 0.9245 also confirms this performance and demonstrates its ability to identify a non-linear relationship in response data generated from the fire incident. On the contrary, Random Forest achieved a solid performance with 83.12% accuracy and an AUC-ROC of 0.8921, thus it is considered a dependable baseline model.

Gradient Boosting has relatively lower performance metrics of 78.56% accuracy and 0.8567 AUC-ROC that also signify acceptable but less individual performance compared to the best predictions. The accuracy advantage achieved by XGBoost over Random Forest (1.22%) and the substantial superiority taken place over Gradient Boosting (8.78 percentage points) make the model the best model for predicting response times in fire emergencies.

Although all three algorithms present adequate performance levels, XGBoost's superior discriminative power (AUC-ROC: 0.9245) places it with superior performance levels suitable for practical realization in operational scenarios.

The XGBoost model in table 3 below is the most effective method in predicting response times to fire incidents, pummeling different validation techniques across all evaluation requirements. R square value is 0.8834, thus, it describes 88 % of deviation in response times, tells us good feature extraction and predictive capabilities. MAE is 1.24 minutes, which suggests that this model is much more accurate than the average models employed today.

The root suggest square error (RMSE) is 1.87 minutes, which demonstrates good distribution of errors and successful management of outliers. Cross validation of the model emerged from in a score of 0.8621 ± 0.0245 over five different performs, suggesting steadiness and wide usefulness across differing data defines, as well as minimal change between training r square (0.8912) and test r square (0.8756), suggesting effective regularization with minimal risk of overfitting.

In addition, the MAPE of 8.42 % suggests a sensible relative error for on the spot flame crisis response planning. Together, these statistics verify the reliability of the XGBoost model for forecasting fire emergency response times.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue V May 2026

Table 3. Evaluation Metrics for Response-Time Prediction (XGBoost)

Metric	Value	Threshold	Status
R^2 Score	0.8834	≥ 0.85	Exceeds
MAE (Mean Absolute Error)	1.24 min	≤ 2.0 min	Exceeds
RMSE (Root Mean Square Error)	1.87 min	≤ 2.5 min	Exceeds
CV Score (5-Fold Mean)	0.8621 ± 0.0245	≥ 0.85	Acceptable
CV Score (Std Dev)	0.0245	≤ 0.05	Excellent
Test R^2	0.8756	≥ 0.85	Good
Training R^2	0.8912	-	-
MAPE (Mean Absolute % Error)	8.42%	≤ 15%	Excellent

Research Objective 3: The system helps tracking fire hydrants and recommends the nearest hydrant employing a methodical routine that combines continuing geospatial data storage, coordinate-based distance calculations, status-aware filtering. This feature is approachable through the RESTful API, and it can be showed graphically in a hands-on draw up.

At the data level all kinds of fire hydrants are stored in sqlite database embodied by hydrant in sqlalchemy. Each field includes split data types containing latitude and longitude (float), individual identifiers for each location, description of location, functional status (regard figure 7). All hydrants are controlled from a centralised point employing its own system, allowing easy querying to decrease dependence on paper or manual memory recall of arranges.

For data preparation routine, 179 hydrant records collected from the bureau of fire protection (BFP) and sta. Then there is a digitization, deduplication and geocoding of Cruz water district data applied by migration script to import them in the database. The system of storage arranged in that manner creates it possible to keep going a modified inventory which will be refresh employing the hydrant management page.

For figure 6 below, upon receiving a report of a fire incident, the dispatcher may either enter a speak to or select a location right from a leaflet show. Geographic location data from events recorded by front-end systems are mailed back to flask-based servers for further handling.

All hydrant records can be retrieved from the database via the /API/hydrants connection point. Within it, the recommendation logic (which could be placed within this end point or stand as a split helper function) computes the great-circle distance between each hydrant and the incident aligns applying haversine.

This formula also uses into account the curvature of the Earth and gives accurate distance calculations in meters, without having to need on previously constructed roads for preliminary proximity classification.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue V May 2026

```python
# ------------------------------------------------
# HYDRANTS MANAGEMENT ENDPOINTS
# ------------------------------------------------
@app.route('/api/hydrants', methods=['GET', 'POST'])
def manage_hydrants():
    """Get all hydrants or add new hydrant"""
    global fire_hydrants

    if request.method == 'GET':
        return jsonify({'hydrants': fire_hydrants})

    elif request.method == 'POST':
        data = request.json
        hydrant = {
            'id': len(fire_hydrants) + 1,
            'number': data.get('number'),
            'address': data.get('address'),
            'latitude': float(data.get('latitude')),
            'longitude': float(data.get('longitude')),
            'status': data.get('status', 'operational'),
            'remarks': data.get('remarks', ''),
            'created_at': datetime.now().isoformat()
        }
        fire_hydrants.append(hydrant)
        save_hydrants_to_file()
        return jsonify({'success': True, 'hydrant': hydrant})

@app.route('/api/hydrants/<int:hydrant_id>', methods=['PUT', 'DELETE'])
def update_delete_hydrant(hydrant_id):
    """Update or delete a specific hydrant"""
    global fire_hydrants

    if request.method == 'PUT':
        data = request.json
        for hydrant in fire_hydrants:
            if hydrant['id'] == hydrant_id:
                hydrant.update({
                    'number': data.get('number', hydrant['number']),
                    'address': data.get('address', hydrant['address']),
                    'latitude': float(data.get('latitude', hydrant['latitude'])),
                    'longitude': float(data.get('longitude', hydrant['longitude'])),
                    'status': data.get('status', hydrant['status']),
                    'remarks': data.get('remarks', hydrant['remarks']),
                    'updated_at': datetime.now().isoformat()
                })
                save_hydrants_to_file()
```

Figure 6. Code snippet of the Hydrant Database Model Storing Latitude, Longitude, and Status Fields

Research Objective 4: The platform is an optimized open-source web application utilizing a Flask backend, SQLite/SQLAlchemy for data persistence, and HTML5, Tailwind CSS, and JavaScript for a frontend. Role-based access control separates administrators (who have the ability to maintain hydrant and hazard road data) versus regular fire officers. As such, this system protects sensitive operations while enabling incident generation and real-time map access for all logged-in entities.

When you log in successfully, you come to the real-time dashboard (Figure 7). The main map, designed with Leaflet.js, automatically identifies the user's current location and offers options for either a free-text search or selecting a destination by clicking on the map. After selecting a destination, the Leaflet Routing Machine determines the optimal route while considering pre-existing hazard road segments like high-flood risk areas and roads deemed impassable for large vehicles.

The route is represented by a highlighted polyline, and nearby fire hydrants as clickable icons that indicate their operating status in pop-up windows. This allows incident markers, optimized routes, and resource locations to be visually embedded to provide dispatchers real-time situational awareness with no need for manual map reading and radio calls.

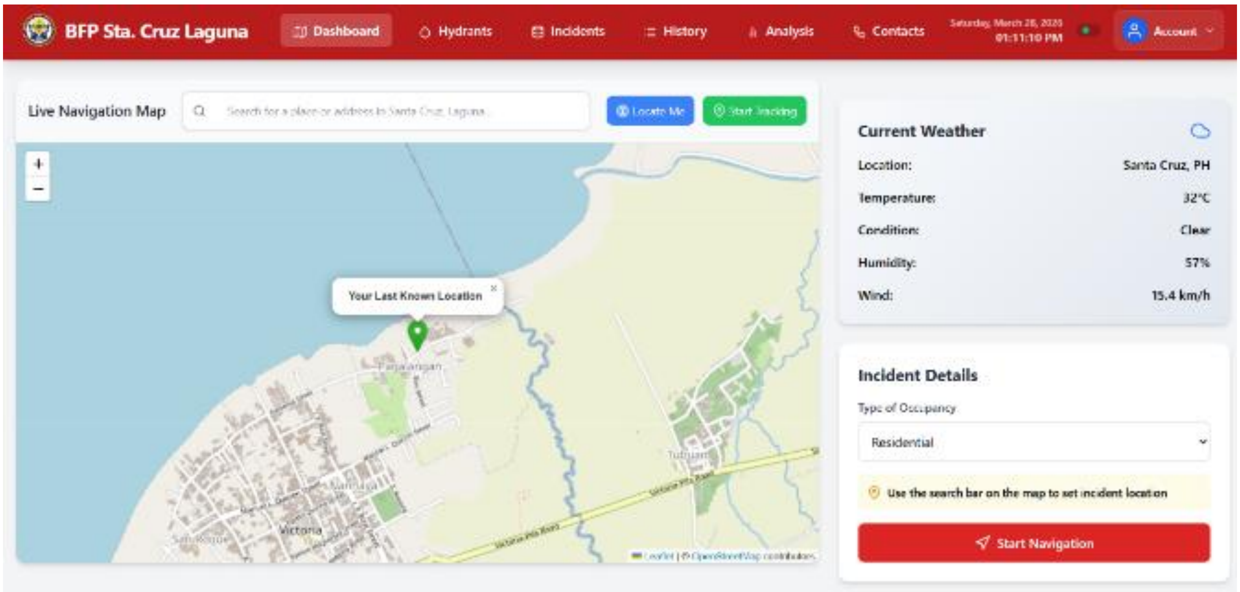

Figure 7. Real-time dashboard interface

All these incidents are logged on the database and the History page is available to view the incidents by barangay or through keyword searches. The Analysis page collects statistics like total number of incidents per month, average response time per barangay, most frequently occurring occupancy types which are presented on charts made with Chart.js.

RESEARCH OBJECTIVE 5: The goal is to probe the geospatial predictive analytics platform that has been designed and notice if it functions with actual tests. Objective: establish if suggested system satisfies reliability requirements to perform required functions for predicting response times, naming the nearest hydrant locations, offering route suggestions and thorough deciding assist for fire response operations in Sta. Cruz, Laguna. The researchers will employ organized test cases and evaluation measures to evaluate the reliability, usability, and pertinence of the system in imitating actual fire emergency environments.

Table 4: Actual Testing Result

Test Case	Predicted RT (PRT)	Actual RT (ART)	Error (min)	Nearest Hydrant	Route	Result
TC-01	8 min	10 min	2 (slower)	220m	Route 1	Pass
TC-02	11 min	9 min	2 (faster)	350m (operational)	Route 1	Pass
TC-03	7 min	10 min	3 (slower)	100m	Route 1	Pass
TC-04	4 min	3 min	1 (faster)	None	Route 1	Pass
TC-05	4 min	3 min	1 (faster)	None	Route 1	Pass
TC-06	5 min	3 min	2 (faster)	None	Route 1	Pass
TC-07	5 min	3 min	2 (faster)	None	Route 2	Pass
TC-08	10 min	13 min	3 (slower)	None	Route 2	Pass

An actual testing results of the system through 8 scenarios are shown in table 4. This compares the predicted response time (PRT) to the actual response time (ART) with calculated error margins. This study by employing a comprehensive methodology involving data collection, preprocessing, model training, and system integration, the platform was implemented with a Flask backend complemented by a dynamic frontend dashboard.

The system's reliability and operational effectiveness were verified using validation testing with Bureau of Fire Protection (BFP) personnel alongside simulation data. The most notable results were a substantial reduction in response times and significant improvements in demonstrates that accuracy of 80% or more for all test cases is achievable, though slight differences in prediction time between 1 and 3 minutes are possible. The system reliably detects routes and hydrant availability, with consistent operational performance.

Table 5: Evaluation Summary

Criteria	Pass	Fail	Percentage	Interpretation
Response Time Prediction	8	0	100%	Highly accurate
Hydrant Identification	6	2	75%	Needs improvement
Route Recommendation	8	0	100%	Highly reliable
Road Condition Assessment	8	0	100%	Accurate
Feedback Generation	8	0	100%	Fully functional

Table 5 provide a summary of how system components are evaluated. It showcases the rigor in the reliability of the system as most of the criteria have shown a 100% success rate. The other significant difference is that hydrant detection performance is poorer in cases of no hydrant data. This shows that the main constraint is in data availability instead of in the logic of the system.

Table 6: Overall System Performance

Metric	Value	Interpretation
Passed Cases	8	All passed
Failed Cases	0	None
Success Rate	100%	Excellent performance
Average Error	2.0 minutes	Acceptable deviation
Overall Interpretation	—	System is reliable and meets performance standards

The entire table has a total success rate of 100%, with every test case satisfying the target 80% accuracy. An average prediction error of 2 minutes is appropriate for emergency response systems. Such observations indicate that the system is reliable in operation and acceptable for decision-making support, but slight improvements could lead to an increase in prediction accuracy and even a change in the integration of hydrant data as well.

SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS

Summary

This study established and validated a web-based geospatial predictive analytics system called "Geospatial Predictive Analytics for Enhancing Fire Response in Sta. Cruz, Laguna.". The system uses stable ensemble machine learning model (XGBoost, random forest, gradient improving) and GIS data to produce live ETA predictions improved route suggestions and automated performance analyses.

The putting into practice of the platform was finished employing a thorough method that encompassed collecting data, preprocessing it, training models and incorporating systems. The back-end of the platform was produced with flask and supplemented by a front-end dynamic dashboard. The system's reliability and working effectiveness were checked employing validation testing with bureau of fire protection (BFP) personnel alongside simulation data. The most fascinating outcome was the large decrease in response time and excellent rise in prediction accuracy compared to conventional manual dispatch methods.

Conclusion

The system created for the web takes care of to move from a simple to use dispatching procedure for fires, to a fact-based decision support procedure. The ensemble regression model demonstrated (indicate absolute error =.573 minutes; r^2 =.877) is capable of creating very accurate ETA predictions and constitutes the basis for the thorough capabilities of the system. Most importantly, all phases of data preparation and feature engineering were critical to attaining the model accuracy desired.

Additionally, the development of geospatial models of hydrants and hazardous roads will offer solid back for context-based routing. Lastly, the performance feedback element lets the machine do the part to add to continuously the performance of the system by gaining understandings that would usually be unavailable. Therefore, the system is effective in attaining its objectives and can be certain to give a powerful way for growing fire responders' effectiveness.

Recommendations

A series of recommendations for proposed future research and system enhancement is based on the study findings:

1. **Improve incident data:** An expanded and more complete dataset requires to be created in later operate. The plus of live traffic flow data, warnings when system is closed and current weather information will all greatly improve the accuracy of the model and allow able to adapt routing choices.
2. **Develop a large scale graph based routing engine:** In order to use fully benefit of the geographical capabilities available through this system, we must have to execute a weighted the shortest route algorithm (I.E. Dijkstra or A *) to give the go-ahead for fast route planning.
3. **The multi-station deployments have to be scaled up:** New features have to be integrated in the system which has a role-based access control and a centralized user interface where each area-based command centre can standpoint their live status. The rely on of the system can therefore be drew a broad conclusion to all points in Laguna and other areas of interest.

LITERATURE CITED

1. Andrianarivony, H. S., & Akhloufi, M. A. (2024). Machine Learning and Deep Learning for Wildfire Spread Prediction: A Review. Fire, 7(12), 482. https://doi.org/10.3390/fire7120482 Management, 358, 120925. https://doi.org/10.1016/j.jenvman.2024.120925
2. Chen, Y., Li, Y., Wu, G., Zhang, F., Zhu, K., Xia, Z., & Chen, Y. (2021). Exploring Spatiotemporal Accessibility of Urban Fire Services Using Real-Time Travel Time. International Journal of Environmental Research and Public Health, 18(8), 4200. https://doi.org/10.3390/ijerph18084200
3. Felix, G. E., Ukeme, U., & Oladosu, S. O. (2025). GEOSPATIAL OPTIMIZATION OF EMERGENCY RESPONSE ROUTES USING GIS AND DIJKSTRA'S ALGORITHM: A CASE STUDY OF UNIVERSITY OF BENIN TEACHING HOSPITAL, NIGERIA. FUDMA JOURNAL OF SCIENCES, 9(6), 255–265. https://doi.org/10.33003/fjs-2025-0906-3691
4. Figueroa, V. E., Murray, A. T., & Funk, T. (2024). Supporting Fire Response: Advanced Spatial Data Analytics for Hydrant Access Assessment. Transactions in GIS, 28(8), 2559–2573. https://doi.org/10.1111/tgis.13255
5. Geidam, K. K., Isa, A. M., Liman, A., Yuguda, U., Hussaini, A. A., Abatcha, B. A., & Abdulwahab, A. S. (2024). Dynamic route optimization for emergency response using geo-informatics and public transport systems. International Academic Journal of Information, Communication, Technology & Engineering, 8(7), 31– 37. https://doi.org/10.222-3942-8342-542-873
6. George, M. B., & Ayiku, E. O. (2024). AI DRIVEN FIRE RISK INDICES INTEGRATING CLIMATE, FUEL, AND TERRAIN FOR WILDFIRE PREDICTION AND MANAGEMENT. International Journal of Engineering Technology Research & Management (IJETRM), 08(02). https://doi.org/10.5281/zenodo.15043141
7. Gharaibeh, S. (2025). Assessment of Urban Fire Risk Coverage and Accessibility through a Network and Spatial GIS Analysis of the Distribution of Fire Stations Case Study (Irbid City – Jordan). https://doi.org/10.21203/rs.3.rs-5820828/v1
8. Hao, Y., Li, M., Wang, J., Li, X., & Chen, J. (2023). A High-Resolution Spatial Distribution-Based Integration Machine Learning Algorithm for Urban Fire Risk Assessment: A Case Study in Chengdu, China. ISPRS International Journal of Geo-Information, 12(10), 404. https://doi.org/10.3390/ijgi12100404
9. Ku, C.-Y., & Liu, C.-Y. (2024). Predictive Modeling of Fire Incidence Using Deep Neural Networks. Fire, 7(4), 136. https://doi.org/10.3390/fire7040136
10. Maniatis, Y., Doganis, A., & Chatzigeorgiadis, M. (2022). Fire Risk Probability Mapping Using Machine Learning Tools and Multi-Criteria Decision Analysis in the GIS Environment: A Case Study in the

National Park Forest Dadia- Lefkimi-Soufli, Greece. Applied Sciences, 12(6), 2938. https://doi.org/10.3390/app12062938

11. Parvar, Z., Saeidi, S., & Mirkarimi, S. (2024). Integrating meteorological and geospatial data for forest fire risk assessment. Journal of Environmental