Fire Analysis Report

Ezgi BAKIR, Graduate School of Informatics, METU, Ankara, Turkey, ezgi.bakir@metu.edu.tr

Abstract— **This study investigates two key questions related to fire incidents in İzmir, Turkey. The first question examines whether there is a significant difference in average firefighter arrival times between urban and rural areas. The second question explores the possibility of predicting fire outcomes using early incident data such as fire type, cause, arrival time, and materials used. The dataset includes 12986 fire records from 2023, provided by İzmir Büyükşehir Belediyesi. Preprocessing steps included handling missing values, encoding categorical variables, generating new features, and managing outliers. Statistical analysis showed that rural areas have significantly longer response times, supported by both the Mann-Whitney U test and visual evidence. For the second question, a Random Forest Classifier was used for prediction. The model was tuned using RandomizedSearchCV and evaluated with F1 score due to class imbalance. Feature importance analysis revealed that fire type is the strongest predictor of fire outcomes. Additional analysis showed that some fire types are more likely to end in complete loss. The model outperformed a baseline classifier and proved useful in identifying high-risk cases. These findings can support fire prevention strategies, emergency planning, and better resource allocation.**

Keywords—fire, arrival time, firefighter, rural area, urban area, fire type, fire cause. fire response time, rural vs urban, fire outcome prediction, Random Forest, feature engineering, emergency services, İzmir fire data

I. INTRODUCTION

Fire incidents are serious threats for communities because they endanger lives and cause damage to properties. They also put the future safety and growth of cities at risk [1]. Understanding the key factors that affect fires is crucial for developing effective prevention and response strategies. One of the most important factors in controlling fires is how quickly firefighters arrive at the scene. Faster arrival can lead to reduce damage and save lives [2].

However, the infrastructure, population density, and resources accessible in urban and rural locations are different [3]. They may therefore respond to emergencies at different times. Furthermore, if it were possible to predict fire outcomes from early data, then emergency planning, risk assessment, and resource allocation would all be more effective.

As cities grow and both rural and urban areas expand, understanding these differences in fire response between urban and rural regions, along with better prediction methods, becomes even more important for keeping communities safe and resilient.

This research aims to address two fundamental questions concerning fire incidents. These are:

1) Is there a significant difference in average arrival time between urban and rural fires?

2) Can we predict the outcome of a fire using features like fire type, cause, arrival time, and materials used?

To support this research, we first need to look at previous studies on fire response times and fire outcome predictions. Studies consistently show that fire response times are shorter in urban areas compared to rural ones. A research study analyzed more than 6000 structure fire responses in New Zealand, and they found that urban regions benefit from closer fire stations and better road infrastructure and hence they have faster arrival times. In contrast, rural areas often face longer delays because of geographic isolation and lower service coverage [1].
Similarly, a systematic review of 37 studies comparing emergency medical services in urban and rural areas also concluded that rural response times were much longer. 93% of the studies that this research had reviewed reported longer arrival times for rural areas [4]. This is mostly because of fewer stations, longer distances, and road network limitations.

There are two research studies from Sweden. One of them [2] looked at real vs. estimated fire response times and they found out that both real and estimated response times are longer for rural areas. Similarly, the other study [5] showed that emergency services respond more slowly in rural areas. They had an average response time of 19 minutes in rural areas compared to 8 minutes in urban areas.

As a result, these findings proved that people living in rural areas face unequal fire respond service compared to the ones who live in urban areas and this inequality might have some safety and survival risks.

Another issue this paper is trying to examine is whether there can be some factors that help us to predict the outcome of a fire incident. In a study, the researchers developed a machine learning model with 34 features, including building type, land use, and demographics. The model sorted buildings into five fire risk levels, and it found that the top 22% of highly risk buildings were responsible for 54% of real fire incidents. This shows strong predictive power when many variables are included [6].

Another study used Random Forest and ARIMA models to predict urban fire incidents. They used fire incident data from 2009 to 2018 in Austin, Texas. They found that both fire type and neighborhood factors (like socioeconomic status) affected the prediction accuracy. The models were especially useful for planning resource allocation and understanding patterns of fire risk in different city areas [1].

These studies show that machine learning can successfully predict fire risks and outcomes, especially when it uses multiple features like fire type, response time, materials used, and environment. Such models can support fire prevention, emergency response, and urban safety planning. They can also help decision makers use resources more efficiently and reduce damage in areas that have a higher risk for fire incidents.

II. DATASET AND FEATURES

The dataset used in this study is a collection of 12,986 fire incidents happened in İzmir that is collected and maintained by İzmir Büyükşehir Belediyesi. The fire incidents collection ranges from January 2023 to December 2023 and contains many other information such as the date of the incident, fire type (trash, vehicle, building, ship or boat, etc), fire outcomes (extinguished at the beginning, saved after partially burnt, and completely burnt), structure type (reinforced concrete, steel, masonry, etc.), number of male, female, firefighter deaths or injuries, number of animal deaths, arrival time of firefighters, amount of materials (water, foam, dry chemical powder) used, team departure time from the station, and the district and address region that the incident happened.

The dataset includes both numerical and categorical features.

Numerical columns are:
male_deaths, female_deaths, firefighter_deaths, male_injured, female_injured, firefighter_injured, large_livestock_deaths, small_livestock_deaths, poultry_deaths, pet_deaths, other_animal, arrival_time_minute, foam_used_kg, water_used_m3, and dry_chemical_powder_used_kg.

Categorical columns include:
fire_type, fire_cause, fire_outcome, structure_type, team_departure_time, district, and address_region.

This separation was necessary for selecting appropriate preprocessing techniques, such as encoding for categorical variables and statistical analysis for numerical ones.

III. PREPROCESSING

Some preprocessing operations are done on the dataset in order to clean the data, handle missing values, and prepare it for analysis and modeling. These are as follows:

- *arrival_time_min* feature had a data type as object. It was converted to float.
- For *arrival_time_minute* at first, 618 missing values were detected, which was unusual. After checking the Excel file, it was found that some entries used the format "00:0x" to represent x minutes. To fix this, all "00:" were replaced with "0" in Excel. This solved the issue and only 11 real missing values remained.
- 11 missing values in *arrival_time_minute* and the 3 in *water_used_kg*, are filled with their respective medians to preserve central tendency without skewing the data.
- There were also other missing values in other features. *structure_type* has 9870 missing values. However, because this feature will not be used in any analysis, the missing values are ignored here.
- There was only one duplicate row, and it was dropped.
- The categorical columns *structure_type* and *fire_district* were not encoded, as they were not needed for the research questions.
- For the columns *fire_outcome* and *address_region*, OneHotEncoder was used. These features have low cardinality and there was no natural order between classes. Also, for *address_region* feature, 'İL DIŞI' class is left out because there were only 6 of them out of 12986 entries and they were irrelevant to the research questions.
- For *fire_type* and *fire_cause*, LabelEncoder was applied. The decision was made based on the machine learning model planned for the second research question. The Random Forest Classifier was chosen, and since it is a tree-based model, it can handle numeric labels without one-hot encoding.

IV. FEATURE GENERATION

For tree-based models like Random Forest, feature generation can improve the model's ability to find useful split points. That's why a new feature called *total_material_used* was created. This feature combines the amounts of foam, water, and chemical powder used to stop the fire.
This feature was added to help answer the second research question. Instead of using foam, water, and powder separately, the model can split on a single feature.

V. SCALING, NORMALIZATION, DISCRETIZATION

Throughout the project, scaling, normalization, or discretization were not applied. These steps were not needed for the planned analysis and model. For the first research question, average arrival times between urban and rural fires were compared using raw values. This method works without scaling. For the second question, a Random Forest model was used. Since it is a tree-based algorithm, it is not affected by the scale or range of features. It splits data based on thresholds, not distances. Continuous features can be handled directly, so discretization was also not required. These transformations would not improve the results and might even reduce data quality.

VI. EXPLORATORY DATA ANALYSIS

Statistical summary of numerical data is examined. The data shows that most fire incidents in 2023 were minor, causing no deaths or injuries to people or animals. Median values for all human and animal impact columns are zero, indicating that in the majority of cases, no one was hurt or killed. Male injuries were slightly more common than female or firefighter injuries, though all occurred at very low rates. Animal deaths were rare, but poultry deaths showed a few extreme cases, with only one incident involving up to 17500 losses. This might be due to an entry error but when dataset is examined, it was realized that the burning site was a chicken farm and it was completely burnt down. So, it might be correct but it's still an outlier. Another important finding here is the average arrival time of firefighters. There were also some extreme values in the *arrival_time_minute* variable. While the median was 5 minutes, the maximum was 1445 minutes. According to the IQR method, values above 10.5 minutes are outliers. These values may be data entry errors as well. Outliers can affect hypothesis test results because these tests are sensitive to distribution. So, it was handled with winsorization before analysis. This method reduced the extreme values without removing data as seen in Table 1. It helped protect the validity of hypothesis tests and made the analysis more reliable.

Metric	**Value**
Count	12,985
Mean	5.64
Std Dev	3.75
Min	0.00
25%	3.00
Median	5.00
75%	6.00
Max	23.00

Table 1. Summary statistics for arrival_time_minute after winsorization.

Apart from that, firefighting resources like foam and water were used occasionally, with foam usage reaching up to 15,000 kg in some cases. Most fires require little or no chemical intervention. Overall, while severe incidents occurred, the vast majority of fires had a minimal impact.

Some graphical representations regarding the distribution of some important features are as follows:

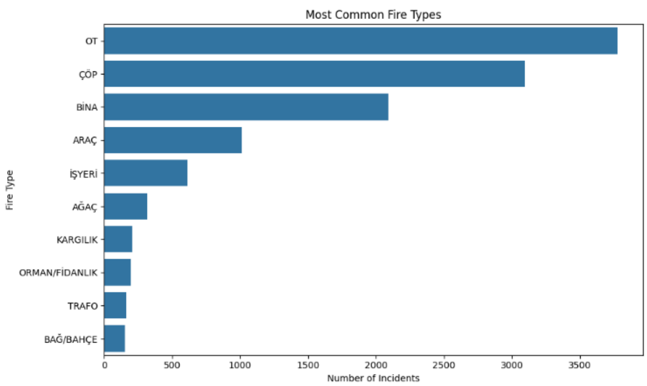

Fig. 1. Most common fire types.

Figure 1 shows the most common fire types recorded in 2023. The majority of incidents were related to grass (OT), trash (ÇÖP), and building (BİNA) fires. Grass fires were the most frequent, with over 3,700 cases.

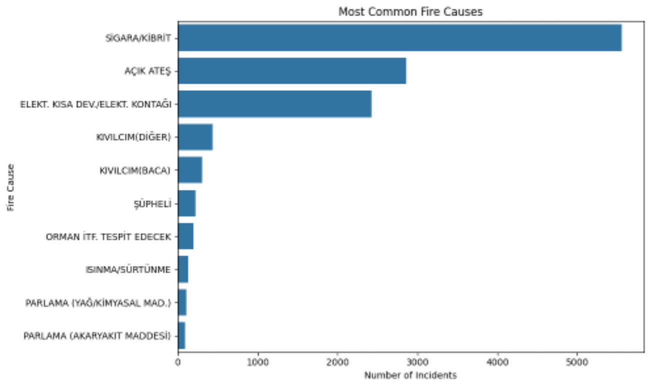

Fig. 2. Most common fire causes.

Figure 2 shows the most common causes of fire incidents. The leading cause was smoking-related materials such as cigarettes or matches (SİGARA/KİBRİT), followed by open flames (AÇIK ATEŞ) and electrical failures (ELEKT. KISA DEV./ELEKT. KONTAĞI). These top three causes account for the majority of recorded fires.

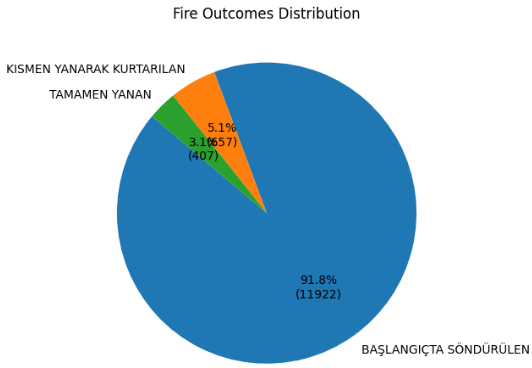

Fig. 3. Distribution of fire outcomes.

Most fires (91.8%) were extinguished at the beginning. A smaller portion (5.1%) were partially saved, and only 3.1% were completely burned.

VII. HYPOTHESIS TESTING

To investigate whether there is a significant difference in firefighter arrival times between urban and rural areas, some statistical tests were applied. Before using any test, three assumptions were checked:
Independence: Each fire record was separate and valid.
Normality: Shapiro-Wilk test was used.
Equality of Variance: Levene's test was applied.

The results were as follows:
Shapiro-Wilk test p-values: 0.0000 (both urban and rural)
Levene's test p-value: 0.0000

These results show that the normality assumption and equal variance assumption were violated ($p < 0.05$). Even after log transformation, Q-Q plots confirmed that the data did not follow a normal distribution (see Fig. 4 and Fig. 5).

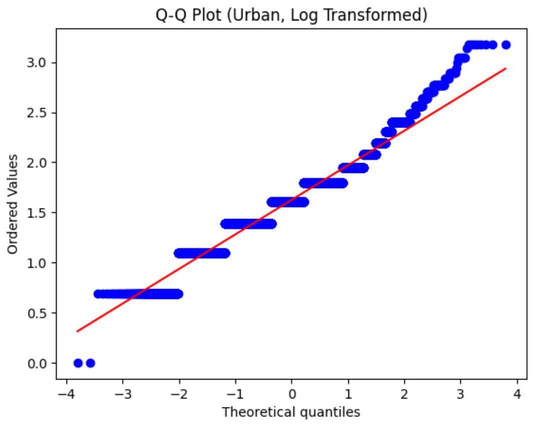

Fig. 4. Normality check using Q-Q plot for urban area (log-transformed arrival times).

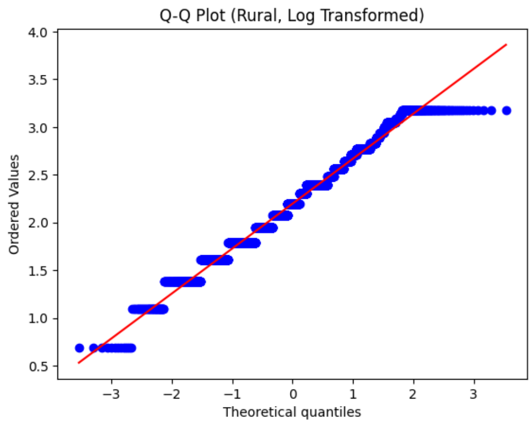

Fig. 5. Normality check using Q-Q plot for rural area (log-transformed arrival times).

Because of the large sample sizes, the Central Limit Theorem could allow a t-test, but a non-parametric test was chosen for more reliable results. Therefore, the Mann-Whitney U test was applied.

Hypotheses:

- H_0: There is no difference in arrival times between urban and rural areas.
- H_1: There is a significant difference in arrival times.

With a significance level of $\alpha = 0.05$, the p-value was 0.0000. Since $p < 0.05$, the null hypothesis was rejected. The result shows a statistically significant difference. In addition, the U statistics supported the finding that rural response times were longer. This conclusion is also confirmed by the boxplot in Fig. 6, which clearly shows:

- Rural areas have a higher median arrival time.
- Rural data is more spread out and contains more outliers.
- Urban response times are shorter and more concentrated.

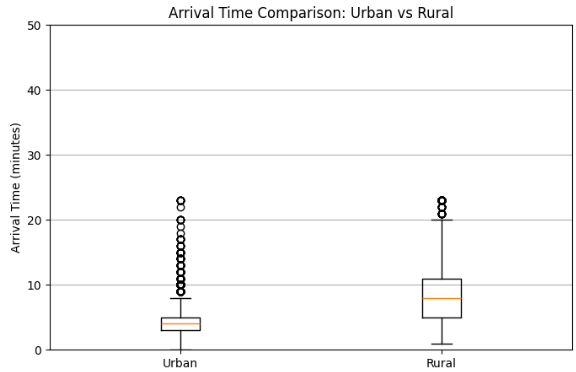

Fig. 6. Boxplot comparing firefighter arrival times in urban and rural areas.

These findings confirm that rural areas experience slower fire response times, which is consistent with the literature and may reflect geographic or infrastructure disadvantages.

VIII. MACHINE LEARNING ALGORITHM

To predict whether a fire would be fully burned, a Random Forest Classifier was used. The target variable (fire_outcome_TAMAMEN YANAN) was imbalanced, so the model was built with class_weight='balanced' to give more weight to the minority class. The dataset was split into three parts: Training set (7787 samples), validation set (2596 samples) and test set (2596 samples). Stratified splitting was used to preserve class distribution. For hyperparameter tuning, RandomizedSearchCV was used with 3-fold cross-validation. Ten combinations of parameters were tested. This method is faster than full grid search but still finds effective settings. The best parameters were: n_estimators = 200, max_depth = 10, min_samples_split = 2 and min_samples_leaf = 2. The best model achieved an F1 score of 0.44 on the validation set, which is a suitable metric for imbalanced data as it balances precision and recall.

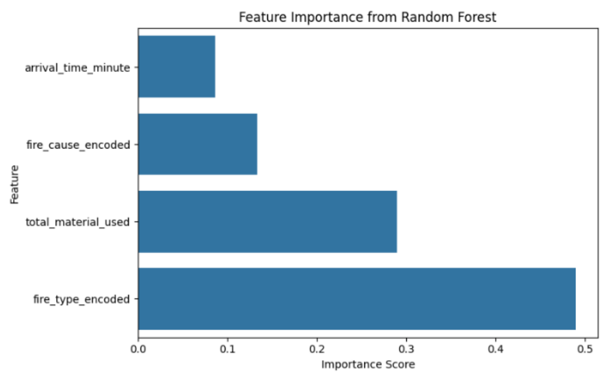

Fig. 7. Feature importance scores of predictors

To understand which features influenced the predictions the most, feature importance scores were calculated (Fig. 7).

- fire_type_encoded was the most important feature.
- total_material_used was also highly important.
- arrival_time_minute had the lowest importance.

These results mean that the type of fire is more informative than how fast firefighters arrive. It also suggests that specific fire types may carry higher risk.

The model was compared with a naive baseline classifier that always predicts the majority class that is fires not fully burned (Table 2).

Model	Accuracy	Precision	Recall	F1 Score
Baseline (Most Frequent)	0.9688	0.0000	0.0000	0.0000
Random Forest	0.9456	0.3052	0.5802	0.4000

Table 2: Comparison of baseline and Random Forest model performance

Although the baseline has high accuracy, it completely fails to detect the minority class. The Random Forest model has lower accuracy but performs much better in F1 score. This shows that it is much more effective at identifying rare but critical cases like completely burned fires.

IX. IMPLICATIONS and DISCUSSIONS

This study tried to answer two important questions. For the first question, our analysis showed that firefighters take longer to arrive in rural areas. This supports what other studies also found. Rural areas often have fewer fire stations and longer distances. This makes response times slower. This can be a safety risk for people living in those areas. For the second question, a Random Forest model is used to predict if a fire will be fully burned. We used features that are available early in an incident. The model gave better results than a baseline model. It showed that fire type and total material used are strong predictors. This means we can make useful predictions with limited information. It also means fire services can prepare better and act faster.

There are also some limitations in this study. We only used data from one city and one year. We also used only four features for prediction. If we had more data and more features (like weather, building type, or region), the model might be even better.

One advantage of our study is that we used real, detailed fire records. We also used a model that works well with mixed data types. The findings can help with emergency planning and fair resource allocation.

REFERENCES

[1] Y. Yuan and A. G. Wylie, "Comparing machine learning and time series approaches in predictive modeling of urban fire incidents: a case study of Austin, Texas," ISPRS International Journal of Geo-Information, vol. 13, no. 5, p. 149, Apr. 2024, doi: 10.3390/ijgi13050149.

[2] J. Hassler, T. A. Granberg, and V. Ceccato, "Socio-spatial inequities of fire and rescue services in Sweden: an analysis of real and estimated response times," Fire Technology, vol. 60, no. 1, pp. 193–212, Oct. 2023, doi: 10.1007/s10694-023-01496-3.

[3] E. Claridge and M. Spearpoint, "New Zealand fire Service Response Times to Structure Fires," Procedia Engineering, vol. 62, pp. 1063–1072, doi: 10.1016/j.proeng.2013.08.162.

[4] A. Alruwaili and A. R. M. Alanazy, "Prehospital time interval for urban and rural Emergency Medical services: A Systematic Literature review," Healthcare, vol. 10, no. 12, p. 2391, Nov. 2022, doi: 10.3390/healthcare10122391.

[5] A. Svensson, B. Nilsson, E. Lantz, A. Bremer, K. Årestedt, and J. Israelsson, "Response times in rural areas for emergency medical services, fire and rescue services and voluntary first responders during out-of-hospital cardiac arrests," Resuscitation Plus, vol. 17, p. 100548, Jan. 2024, doi: 10.1016/j.resplu.2023.100548.

[6] S. Ahn, J. Won, J. Lee, and C. Choi, "Comprehensive building fire risk prediction using machine learning and stacking ensemble methods," Fire, vol. 7, no. 10, p. 336, Sep. 2024, doi: 10.3390/fire7100336.