Contents

Team3-Business Problem

Forecasting Neighborhood-Level Fire Risk in Montreal

Team:

- *Hammed Akande,*
- *Osama Wanis,*
- *Pak Lo,*
- *Ekue Afanou*

Introduction

Globally, structural fires continue to be a serious public safety concern. Recent research estimates that fires cause approximately 401,000 global deaths annually, with urbanization and congestion increasing the risk (Zhang, 2023). In Canada, fire departments responded to over 39,000 fire incidents in 2021, of which 42% were structural and most were residential (Statistics, 2023). In the United States, the NFPA reports that fire departments responded to 1.39 million fires in 2023, leading to 3,670 civilian deaths and $23 billion in property damage. Montreal is exposed to similar risks due to its dense residential areas, commercial districts, and industrial zones. Currently, the city's fire prevention strategy is largely reactive: officials respond to fires as they occur and carry out routine inspections without a predictive framework. This project proposes a data-driven approach to forecast neighborhood-level fire risk monthly. By aggregating fire incidents by neighborhood and month, and integrating weather, demographic, and built-environment data, we will classify each neighborhood 's risk (low, medium, or high) for the coming month. This monthly horizon avoids the unrealistic task of predicting the exact time of fires, yet aligns with inspection schedules and community outreach programs.

Business Problem

Montreal's Fire Safety Service compiles detailed records of every fire intervention, including date, location and type of call. However, this data has not been translated into forward-looking risk assessments. Fire prevention resources are limited: inspectors cannot visit all properties frequently, and crews often prioritize emergency responses. Neighborhoods differ widely in building stock, density and socioeconomic characteristics, producing heterogeneous fire risk profiles. The challenge is to transform historical incident data into actionable forecasts of where fires are most likely to occur in forthcoming months. We will operationalize risk as the expected number of fires per neighborhood per month and classify the predicted counts into low, medium or high tiers. This classification will help allocate inspection crews, plan public education campaigns and pre-deploy firefighting resources.

Rationale and Justification

The rationale for a predictive fire-risk model rests on several grounds. Firstly, the public safety benefit is substantial. By identifying high-risk neighborhoods in advance, officials can prioritize preventive measures such as smoke alarm distribution, code enforcement, and community training, potentially averting fires and saving lives. Additionally, resource allocation can be made far more efficient. New York City's Fire Department (FDNY) inspects only about ten percent of the city's 330,000 buildings annually (Crowley et al., 2023). To focus on the most hazardous properties, the FDNY developed the FireCast system, which combines data from multiple agencies and uses an algorithm to rank buildings for inspection (Roth & Tarleton, 2014). This risk-based inspection system ensures limited inspection hours are directed at the highest-risk locations. Furthermore, socioeconomic equity considerations underscore the need for targeted prevention. A study of the Shreveport Fire Department found that fire deaths, injuries and property damage were disproportionately concentrated in socioeconomically depressed areas and recommended community-specific interventions (Hall, 2024). By incorporating other data sets like demographic and income variables, we can identify vulnerable communities and inform equitable resource allocation. Finally, predictive analytics aligns with broader smart-city initiatives: insurers, emergency services and urban planners all seek granular risk forecasts to support resilience planning. Montreal's investment in this project will place it alongside cities like New York and Pittsburgh, which have adopted advanced analytics to improve fire safety.

Literature Review

Existing work on urban fire-risk prediction spans municipal applications and academic research. FDNY's FireCast uses dozens of variables—building age, use, violation history and behavioural cues—to assign risk scores to buildings and generate ranked inspection lists. Other U.S. cities, such as Pittsburgh, Atlanta and New Orleans, have partnered with universities or non-profit organizations to develop predictive models that identify high-risk commercial properties or areas lacking smoke alarms. Academic models treat fire prediction as a spatio-temporal machine-learning problem. Wang et al. (2019) developed the CityGuard (NeuroFire) model, which combines recurrent neural networks for temporal trends with spatial dependency components, demonstrating improved performance over baseline methods. Oh and Jeong (2022) used extreme gradient boosting to classify grid cells in a Korean city and achieved high accuracy. Socioeconomic and demographic variables are recognized as important risk factors. Research on Edmonton's emergency events found that combining incident history with socioeconomic indicators improved monthly and weekly predictions (Sharma et al., 2024). The National Fire Academy report on "Fire and the Poor" documented that poor neighborhoods suffer disproportionate fire deaths and injuries and advocated for targeted prevention (Crowley et al. 2023). Overall, several literatures emphasize integrating multi-source data—incident history, building attributes, weather and socioeconomic conditions—and using ensemble or deep learning models to generate risk scores for defined spatial units.

Proposed Approach and Data Sources

We will compile a comprehensive dataset for Montreal and develop a supervised learning model to forecast monthly fire risk at the neighborhood level. Key data sources include:

• Fire incidents: historical records from Montreal's open-data portal will provide the target variable. We will aggregate counts by neighborhood and month and use recent incident counts as autoregressive features.

• Fire stations: locations of fire stations will allow us to compute coverage indicators such as the distance from each neighborhood to the nearest station, which may relate to response times.

• Weather: monthly averages of temperature, precipitation and wind speed from Environment Canada will capture seasonal effects on fire risk.

• Demographics and socioeconomic indicators: census variables (population density, median income, percentage of low-income households, unemployment rate) will describe social vulnerability and exposure.

• Built environment and land use: data on building age, density and land-use mix (residential, commercial, industrial) will be used to infer structural vulnerability. If available, data on code violations or prior inspections will provide behavioral signs.

All data sets will be spatially joined to official neighborhood boundaries and temporally aggregated to a monthly resolution, creating a panel of neighborhood–month observations with associated predictors.

Methodology

Our modelling pipeline will be as follows:

1. Pre-processing and temporal aggregation: clean and filter incident data to include genuine fires, geocode locations to neighborhoods and aggregate counts by month. We will create lagged variables for recent incident counts (e.g., one- and three-month lags) to capture autoregressive trends. Continuous variables will be normalized, and missing values will be imputed using domain-appropriate methods.

2. Feature engineering and spatial context: construct predictors for each neighborhood–month, including seasonal indicators, weather variables, socioeconomic metrics, building age and density, land-use mix and distance to the nearest fire station. To capture spatial dependencies, we will compute adjacency-based features (e.g., average fire incidence in neighboring areas) and incorporate census-tract covariates that reflect poverty, age structure, and housing quality. Temporal features such as month-of-year will model seasonality.

3. Model selection and hybrid pipeline: explore a range of predictive models. Ensemble tree methods, such as random forests, will establish baseline performance for our model. Ensemble methods are expected to perform well in modelling nonlinear relationships and handling mixed data types. We will evaluate spatio-temporal deep-learning models (e.g., long short-term memory networks) to capture complex dependencies across time and space and compare performance with our baseline model. A key consideration will be model interpretability. Tree-based models naturally provide feature-importance scores, and we will apply SHapley Additive Explanations (SHAP) to quantify the contribution of each predictor to individual predictions (Yang, et al. 2025). This will help managers understand which neighborhood attributes drive elevated risk and support transparent decision-making.

4. Evaluation and validation: split the data into training and testing sets, reserving the most recent year for out-of-sample evaluation. To account for temporal autocorrelation, we will use rolling time-series cross-validation to tune hyperparameters and assess predictive stability. Spatial cross-validation (e.g., leave-one-neighborhood-out) will evaluate how models generalize to unseen areas. Performance metrics will include accuracy, precision, recall and F1-score for the high-risk class, along with the area under the ROC curve. Predictions will be benchmarked against naive baselines such as using the previous year's counts or random classification.

5. Risk classification and human-in-the-loop deployment: translate predicted counts or probabilities into classification categories using empirically chosen thresholds. Visualize results on interactive maps and produce ranked inspection lists. Use SHAP-based analyses to highlight the most influential factors for each neighborhood, enabling targeted interventions. Recognizing that algorithmic outputs should inform rather than replace professional judgment, we will design the system as a human-in-the-loop tool: fire-prevention officials can review the model's recommendations, adjust thresholds based on operational considerations, and provide feedback to refine the model. Attention will be paid to socioeconomic equity to ensure that preventive actions do not disproportionately affect disadvantaged communities.

Expected Outcomes

• Interactive risk maps and tables that visualize predicted risk levels and help decision-makers direct resources. *(Must Have)*
• A predictive model that classifies each neighborhood's fire risk (low, medium or high) for the upcoming month. *(Must Have)*
• Insights into the most influential risk factors in Montreal, informing longer-term strategies such as upgrading older housing stock or targeting fire-safety education. *(Must Have)*
• Improved resource allocation by focusing inspections and prevention efforts on neighborhoods predicted to be high risk. *(Could Have)*

• A methodological framework and dataset that future researchers can build on for multi-step forecasting or property-level risk assessments. (*Could Have)*

Conclusion

Fire incidents impose severe human and economic costs worldwide, with tens of thousands of deaths and billions in property damage each year. Montreal's dense urban environment means it shares this risk. By aggregating fire incident data and integrating weather, demographic and built-environment indicators, we can anticipate where fires are most likely to occur monthly. This project proposes a data-driven model to classify neighborhoods into fire-risk tiers and inform proactive inspection and prevention activities. Evidence from other cities and research studies indicates that risk-based approaches enhance safety outcomes and optimize the use of limited resources. The proposed model will support Montreal's fire service in transitioning from reactive response to proactive prevention, ultimately making the city safer and more resilient.

References

Crowley C. (2023), Increasing damages from wildfires warrant investment in wildland fire management. U.S. Department of the Interior - Office of Policy Analysis. [*https://www.doi.gov/sites/doi.gov/files/ppa-report-wildland-fire-econ-review-2023-05-25.pdf*](https://www.doi.gov/sites/doi.gov/files/ppa-report-wildland-fire-econ-review-2023-05-25.pdf)

Hall, S. (2024). Fire Loss in the United States. National Fire Protection Association Research. [*https://www.nfpa.org/education-and-research/research/nfpa-research/fire-statistical-reports/fire-loss-in-the-united-states*](https://www.nfpa.org/education-and-research/research/nfpa-research/fire-statistical-reports/fire-loss-in-the-united-states)

Oh, H. Y. & Jeong M. H. (2022) Grid-based Urban Fire Prediction Using Extreme Gradient Boosting (XGBoost). Sensors and Materials. [*https://doi.org/10.18494/SAM4194*](https://doi.org/10.18494/SAM4194)

Roth, J. & Tarleton, J. (2014)- Profiles in Public Service - The Analytics of Fire. The Analytics of Fire. Urban Omnibus. [*https://urbanomnibus.net/2014/06/the-analytics-of-fire/#:~:text=2,a%20building%20or%20people%E2%80%99s%20behavior*](https://urbanomnibus.net/2014/06/the-analytics-of-fire/#:~:text=2,a%20building%20or%20people%E2%80%99s%20behavior)

Sharma, D. P., et al. (2024). Statistical and Machine Learning Models for Predicting Fire and Other Emergency Events. [*https://arxiv.org/abs/2402.09553*](https://arxiv.org/abs/2402.09553)

Statistics Canada (2023), [*https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510019201*](https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510019201)

Wang Q., et al. (2019) CityGuard: Citywide Fire Risk Forecasting Using A Machine Learning Approach. Association for Computing Machinery. [*https://doi.org/10.1145/3369814*](https://doi.org/10.1145/3369814)

Yang, X., et al. (2025) Explainable Artificial Intelligence (XAI) framework using XGBoost and SHAP for assessing urban fire risk based on spatial distribution features. International Journal of Disaster Risk Reduction, Volume 129 [*https://www.sciencedirect.com/science/article/abs/pii/S2212420925006223*](https://www.sciencedirect.com/science/article/abs/pii/S2212420925006223)

Zhang, C. (2023). Review of Structural Fire Hazards, Challenges, and Prevention Strategies. Fire 2023, 6(4), 137. [*https://doi.org/10.3390/fire6040137*](https://doi.org/10.3390/fire6040137)