Exploring Seasonal Patterns in Wildfire Features for Enhanced Prediction

1st Yanting Zheng
Khoury College of Computing Science
Northeastern University
Vancouver, Canada
zheng.yant@northeastern.edu

1st Huizi Wang
Khoury College of Computing Science)
Northeastern University)
Vancouver, Canada
alinawang.ubc@gmail.com

1st Jinda Zhang
Khoury College of Computing Science
Northeastern University
Vancouver, Canada
jindaz.work@outlook.com

***Abstract*—This study explores the seasonal variations in wildfire prediction by analyzing the importance of different features like temperature, wind speed, and the Normalized Difference Vegetation Index (NDVI). Employing machine learning and deep learning techniques, the research identifies key seasonal factors that influence wildfire occurrences. The findings aim to enhance the precision of wildfire prediction models and advocate for adaptable, season-specific management strategies. This approach has significant implications for effective wildfire risk mitigation, leveraging the potential of remote sensing and advanced computational models in environmental studies.**

***Index Terms*—Machine Learning, Deep Learning, Wildfire Prediction, Seasonal Variability, Feature Importance Analysis, Environmental Risk Management, Data Analysis in Environmental Science**

I. INTRODUCTION

The increasing frequency and severity of wildfires globally, and particularly in Canada, underscore the urgent need for improved detection, prediction, and management strategies.

Table 1 provides an overview of the scale and impact of wildfires in Canada and Globally:

Metric	**Canada**	**Global**
Average annual number of wildfires	7,000 [30]	Over 300,000 [14]
Area burned annually (hectares)	Over 2.5 million [30]	Approximately 350 million [13]
Economic Impact (Annually)	CAD 1 billion [9]	USD 100 billion [41]
Number of fatalities (Annually)	Varies, but can be significant especially in bad seasons	Approximately 340,000 [39]
Impact on air quality	Significant, especially in populated areas affected by wildfire smoke [12]	Significant in many parts of the world, especially in regions with frequent wildfires [29]

TABLE I
IMPACT OF WILDFIRES IN CANADA AND GLOBALLY

According to a study conducted by Henderson et al., [17], even short-term exposure to wildfire smoke can exacerbate asthma and other respiratory conditions and has been linked to increased hospital admissions for respiratory issues. Children, the elderly, and people who have breathing problems are even more vulnerable under this circumstance. Wildfires not only affect forests, and homes but also our daily health and well-being.

Moreover, wildfires present a considerable risk to global forests and ecosystems, resulting in substantial ecological and economic damage. In the context of climate change, increasing global temperatures, and rising CO2 emissions, accurately predicting and understanding wildfires is crucial for reducing ecological and socio-economic impacts.

Wildfire occurrence is considered influenced by four main factors: weather or climate, fuels, ignition agents, and humans [16]. Of all these factors, most existing research focuses on environmental and anthropogenic aspects due to the availability of relevant data. This paper concentrates on identifying key features that significantly contribute to wildfire occurrences across different seasons. By pinpointing these critical factors, we aim to enhance wildfire prediction accuracy and inform more effective management strategies. The research specifically examines the variances in feature importance seasonally, offering insights into tailored wildfire prevention measures for each season.

Our study is driven by the objective to identify and analyze the seasonal variance in key features influencing wildfire occurrences. This insight is vital for tailoring prediction models to different seasonal dynamics. By clarifying the seasonal impact on key factors, we address the research question: 'How do specific features influence wildfire occurrences in various seasons, and how can this knowledge enhance prediction and management strategies?' This approach offers a refined perspective on wildfire prediction, moving beyond the common summer-centric view to include spring and fall.

Our findings will not only be invaluable for the academic community but also for forest management authorities, policymakers, and emergency response teams. For example, forest management authorities can develop season-tailored strategies. Policymakers can design policies to allocate resources more efficiently during high-risk seasons. Emergency response teams can prepare for wildfire outbreaks on seasonally relevant predictors.

This paper focuses on feature importance differences in wildfire prediction. We start with an introduction to the severity of wildfires in the world. Then we look at previous research works in wildfire prediction. Afterward, we explain our research scope, methodology, and metrics. Then we present the result according to the metrics. In the end, we discuss our limitations and future work.

II. Related Works

A. Data Collection

Most research includes natural predictors, which can be divided further into climatic, topographic, and land cover. Some research also includes anthropogenic factors which may include unemployment rate [31], proximity to urban and roads. As for climatic factors, most research considers temperature, precipitation, wind speed and direction, soil moisture and humidity. As for topographic factors, slope and aspect [31] [27] [22] are shared by most research predictors. Land cover as the fuel to wildfire is widely adopted in existing research. These factors can be retrieved directly from satellite maps. Anthropogenic factors are considered less frequently than environmental ones, given that the data can be harder to define and process. However, it is found that both lighting and human-caused fires occur with greater probability when closer to roads and populated places in Canada [16].

More specifically, V. L. S. Arruda et al. [3] use the Deep Neural Network (DNN) model for detecting and mapping burned areas in the Cerrado biome in Brazil. The input variable this paper uses includes Google Earth Engine data, such as NDVI (Normalized Difference Vegetation Index), NBR (Normalized Burn Ratio), and delta NBR (Difference Normalized Burned Index). The output variable of this paper is Landsat spectral bands. G. Charizanos and H. Demirhan [10] use NDVI, humidity, highest temperature, and mean wind speed as the predictor. S. Tavakkoli Piralilou et al [21] use remote sensing data at different spatial resolutions, such as Landsat 8, Sentinel-2, ALOS, and SRTM datasets. The output data of this study are the wildfire susceptibility prediction (WSP) maps. These maps indicate the areas at risk of wildfires in the Guilan Province, Iran.

B. Methods

There are two mainstream machine learning methods used to predict wildfire occurrences: classical Machine Learning algorithms [31] [27] [37] [21] [33] and Deep Neural Networks [22] [43]. Classical Machine Learning algorithms have the advantage of easy interpretation and are mostly used on tabular data and small datasets, while Deep Neural Network is usually much more complex and considered a "black box", in the meanwhile is versatile enough to handle all kinds of data and large datasets.

1) Classical Machine Learning Methods: Some research papers used classical Machine Learning algorithms such as Random Forest and SVM. S. Tavakkoli Piralilou et al [21] employ machine learning algorithms such as artificial neural network (ANN), support vector machines (SVM), and random forest (RF) to train and test the models. This paper addresses the class imbalance in wildfire susceptibility prediction by using the Synthetic Minority Over-sampling Technique (SMOTE) algorithm. SMOTE is used to generate synthetic samples of the minority class (i.e., wildfire locations) to balance the class distribution and improve the performance of the machine learning models.

Another recent research discovered fire susceptibility using ML methods and the Google Earth Engine dataset in Gangwon-do, Korea [33]. The authors constructed forest fire susceptibility mapping using classification and regression trees (CART), boosted regression trees (BRTs), and random forest (RF) algorithms. The evaluation metrics used in this paper is ROC and AUC. Input data is the distance to urban areas, rainfall amount, annual average temperature, drain density, normalized difference vegetation index (NDVI), topographic wetness index, aspect, slope, distance to rivers, distance to roads, and elevation. Output data is forest fire susceptibility maps (FFSMs). This paper also emphasizes the importance of human factors in wildfires.

2) Deep Neural Networks(DNN): Besides image segmentation methods for wildfire prediction, there were also papers using image classification methods for detection, e.g. using CNN-based models, such as YOLO, as well as vision transformers(VIT).

Furthermore, a deep learning approach for early wildfire detection from hyperspectral satellite imagery, although not fully extracted, suggests a focus on advanced machine learning techniques for wildfire detection using sophisticated satellite data [28].

While existing studies have explored various wildfire predictors, there is a notable gap in understanding which of these plays a more crucial role across different seasons. Our research seeks to fill this gap by analyzing and comparing the significance of various features in predicting wildfires seasonally. This understanding is vital for developing season-specific wildfire management and prediction models

As for the metrics to evaluate the model performance, no matter which Machine Learning method they used, most research mentioned recall, F1 Score, accuracy, and specificity.

C. Challenges

One of the challenges that most research mentioned is the imbalanced nature of wildfire datasets. Various methodologies are applied under certain circumstances to address this issue, such as oversampling and class weights.

D. Research Gap

Recent advancements in machine learning and deep learning have significantly good results in wildfire prediction. Despite these developments, focus on how the relevance of predictors changes with seasons lacks.

Our research aims to bridge this gap by investigating how different predictors' effectiveness varies across seasons. We plan to use different models for seasonal analysis, such as Machine Learning and Deep Learning techniques. Additionally, we seek to explore the year-to-year variance in predictor relevance.

III. Data source and workflow

A. Data Source

We collected all data from Google Earth Engine (GEE) since it is more readily accessible than downloading data from different websites. However, the limitation of using GEE is that some data isn't listed in its catalogue. All available features are listed in fig. 3. Based on what we learned from the literature review, we put our features into these categories: climate, topography and vegetation.

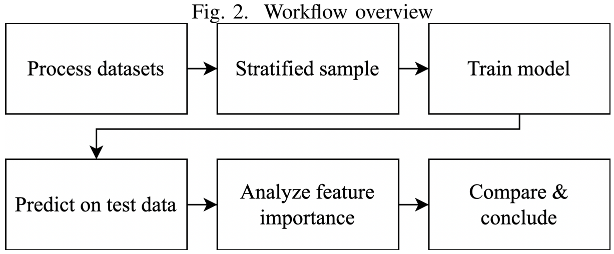
Fig. 1. Available features

climate	precipitation
	evapotranspiration
	soil moisture
	temperature
	wind speed
topology	elevation
vegetation	NDVI
	landcover

B. Workflow

Our workflow can be summarized with fig. 2. We will describe the details of of the first two stages in this section. The rest of the stages will be addressed in the following sections.

Fig. 2. Workflow overview



Data Processing

- **Unify temporal resolutions** The data maps we collected from GEE have different temporal resolutions while what we need is monthly data. So we made the temporal resolution for sampling monthly.
- **Convert label data to class** We defined our problem as a classification problem. What we directly sampled from GEE are float numbers. So what we did is to select a threshold value. All values below it are put into 'no fire' class and those greater or equal to it are 'fire' class. The threshold value is Kelvin Temperature and we chose 300K as the boundary.
- **Normalize input features** Some of the models we experimented with are feature-scaling sensitive models, such as DNN models. We normalized each of the input features with a fixed range. The selection of ranges is based on the observation of value distribution in the region of interest.

Startified Sample

Wildfire problem is known to suffer from the imbalance issue. This is because fire pixels are sparsely distributed and occupy a much smaller portion on the map compared to no-fire pixels. We handled this issue with 'stratifiedSample' API of GEE. In this way, we ensured the two classes are equal in amount so that our model is trained on a diverse and representative dataset.

C. Metrics

1) Recall: Given the high threat of wildfires, we will use the recall (true positive rate) as an evaluation metric. Recall measures the model's ability to correctly identify actual positive instances, which is identifying the occurrence of wildfires. Recall is important in cases where missing actual wildfire occurrences (false negatives) can have severe consequences. Minimizing false negative cases can help policymakers better understand the occurrence of wildfire and prevent it.

The formula for recall is given by:

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Where:

- True Positives (TP) represents the number of correctly predicted wildfires. False Negatives (FN) represents the number of actual wildfires that were not predicted by the model.

2) F1 Score: The F1 score combines both precision and recall into a single metric.

The formula for F1 Score is given by:

$$\text{F1 Score} = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

3) Validation Strategy: Our research will be validated using rigorous cross-validation techniques to ensure the reliability and applicability of our findings. The 5-fold cross-validation technique will be used to ensure that each fold is a good representation of the entire dataset given that the dataset itself is imbalanced. (Will dive deeper once finish analyzing our dataset using different algorithms, To be continued) Performance metrics such as recall, precision, F1 score, and area under the ROC curve will be used to evaluate the model's predictive capabilities.

D. Challenges

The challenges we anticipate are span data collection, model training, and validation.

1) Data Collection: Due to the limited amount of maps on the GEE platform, we have to abandon predictors that are not accessible on GEE, such as road density and unemployment rate, etc. We also need to try filtering data that is temporally overlapped especially for factors that can change dramatically over time.

2) Imbalanced Data: Wildfire occurrences, when compared to non-occurrences, present an imbalanced dataset. This imbalance can bias the models towards predicting the majority class, leading to reduced sensitivity in detecting actual wildfire events.

IV. ALGORITHM

In this section, we will discuss the fundamental principles behind the feature importance analysis. Then each of us will describe our choice of models.

1) Feature Importance Analysis: Permutation feature importance is defined to be the decrease in a model score when a single feature value is randomly shuffled [25]. The idea behind this method is, that the more important a feature is, the larger impact it has on the result. Shuffling such features will break the link between features and labels, thus causing errors to increase.

The steps of my analysis can be described as below:

- Calculate the error between prediction and actual labels as baseline error.
- For each input feature, shuffle all test records of it and maintain the correct order for other features.
- Predict with shuffled test data and measure the error after the shuffling.
- Sort the error from large to small, then we get the order of each feature's importance.

A. Justin

XGBoost: XGBoost, also known as Extreme Gradient Boosting, is a powerful ensemble learning algorithm with exceptional performance in classification and regression tasks. Developed by Tianqi Chen and Carlos Guestrin [**?**]. XGBoost creates an ensemble of decision trees, excellent at both classification and regression tasks, making it an ideal choice for our wildfire prediction model.

The key components and benefits of XGBoost include:

- **Gradient Boosting:** XGBoost is based on the concept of gradient boosting, which sequentially builds an ensemble of decision trees, each correcting the errors made by the previous ones.
- **Regularization:** XGBoost uses L1 (Lasso) and L2 (Ridge) regularization techniques to prevent overfitting.
- **Tree Pruning:** XGBoost uses tree pruning to control the depth of individual trees, preventing excessive branching and reducing overfitting.

I use XGBoost to train a binary classification model that distinguishes between wildfire occurrences and non-occurrences based on a variety of features we pre-defined for all 3 seassons, for 3 years from 2019 to 2022. With preprocessing pipelines specified in section III.B workflow, we make sure dataset is balanced, and then use model.feature_importance API from xgboost library to perform feature importance analysis and plot the visualized results for better understanding.

Kernel SVM, LSTM, KNN, TabNet, Logistic Regression, Naive Bayes, 1D-CNN: Alongside XGBoost, I also explore various models for wildfire prediction. Kernel SVM leverages support vector machines with kernel tricks for classification. LSTM, as shwon in 5, is a type of recurrent neural network, captures temporal dependencies is first invented by Hochreiter and Schmidhuber [18]. KNN utilizes nearest neighbors for classification. TabNet [2] employs attention mechanisms for tabular data as shown in 3, which is first invented by Vaswani et al [38]. The architecture can be show in 4 .Logistic Regression and Naive Bayes offer simpler probabilistic models. 1D-CNN applies convolutional layers for feature extraction from sequential data.

Fig. 3. Attention is All you Need

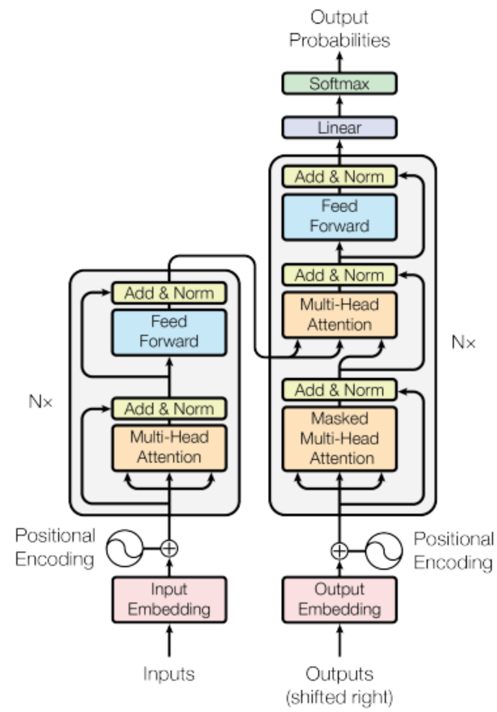


Fig. 4. TabNet Architecture

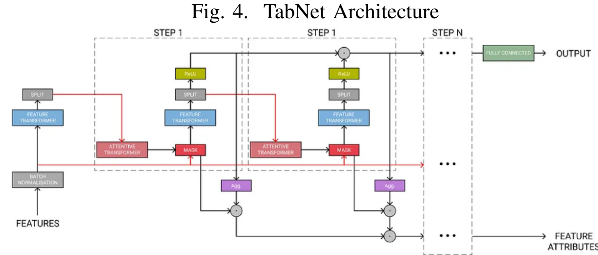
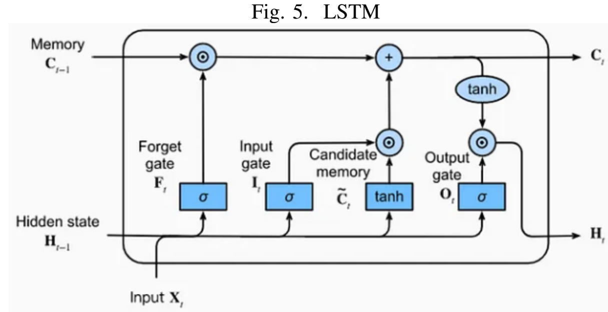


Fig. 5. LSTM



In our wildfire prediction task, each model is trained and evaluated based on predefined features across multiple seasons and years. We balance the dataset using preprocessing pipelines and conduct feature importance analysis, visualizing the results for comprehensive insights.

1) Randomized Search: I use Randomized Search Cross-Validation(CV) proposed by J. Bergstra and Y. Bengio in 2012, which is considered more efficient than Grid Search with Cross-Validation for hyperparameter tuning [**?**]. Randomized Search explores a random set of hyperparameter combinations within defined ranges, which can find optimal hyperparameters more efficiently. This approach is particularly advantageous when there is a large search space and limited computational resources. While training wildfire prediction models, Randomized Search allows us to efficiently discover hyperparameters that yield the best results without exhaustive computation.

2) Feature Importance Analysis by Season: Feature importance analysis is a crucial step in understanding the predictive power of the features in our model. It helps us identify which features contribute most to the model's prediction decision, this can provide valuable insights into the factors that are most influential in predicting wildfire occurrences.

In XGBoost, feature importance is calculated as the average gain of the feature when it is used in trees. Here are the steps we follow for feature importance analysis:

Train the model: We first train our XGBoost model using the optimal hyperparameters obtained from the Randomized Search.

Calculate feature importance: We then use the model.feature_importance API from the xgboost library to calculate the importance of each feature. This gives us a score for each feature in our dataset, indicating how useful or valuable each feature was in the construction of the boosted decision trees within the model.

Visualize the results: To better understand and interpret the feature importance, we plot the scores in a bar chart. The y-axis represents the features, and the x-axis represents the importance scores. Features are sorted by their scores, allowing us to easily identify which features are most important in our model.

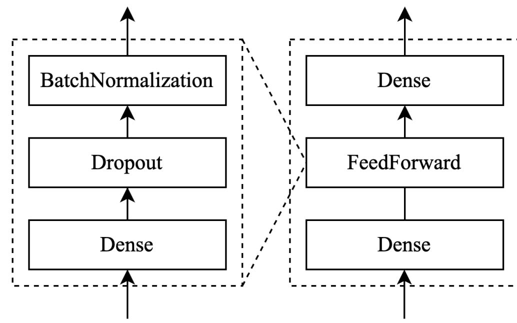
Interpret the results: The higher the score, the more important the feature is. Features with low scores may not be necessary for our model and could potentially be dropped to simplify our model.

B. Mia

1) Artificial Neural Networks: Artificial Neural Networks (ANN) based on convolutions, recurrent mechanisms, or transformers seem to outperform traditional machine learning models in a multitude of domains. However, they still face challenges when tackling tabular data [6]. It remains unclear why they cannot achieve the same level of predictive quality as in computer vision or natural language processing, and they even cannot compete with traditional methods, especially ensemble methods such as XGBoost or random forest. The hypothesis focuses on the differences between tabular data and image or language data, including heterogeneous traits in tabular data compared to homogeneous data (image, text, or audio) and the weaker feature correlation of tabular data than the spatial or semantic relationships exhibited in image or speech data. Hence A.Kadra et al. [1] called tabular datasets the "last unconquered castle" for ANN models.

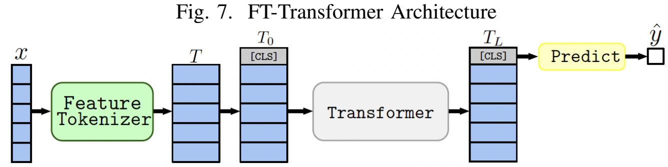
In my experiment, I designed a baseline ANN model with 2 hidden layers and a deeper ANN model with 5 hidden layers, batch normalization layers, and dropout layers. Fig. 6 shows the structure of the deeper ANN model. This feed-forward unit is common in ANN architecture and general enough for most problems.

Fig. 6. Deeper ANN Architecture



2) FT-Transformer: Recently, Transformer, believed to be the state-of-the-art deep architecture, has shown comparative performance on tabular data problems [6] [19] [20] [24] [35]. Yury et al. designed FT-Tranformer and proved that it can outperform the traditionally dominant gradient-boosted decision trees (GBDT) methods in some tabular problems [42]. I used an implementation of FT-Transformer using TensorFlow from https://github.com/aruberts/TabTransformerTF and modified it to make sure the architecture is consistent with that in the original paper. Fig. 7 demonstrates an overview of the architecture of FT-Transformer and fig. 8 shows the structure of the transformer block.

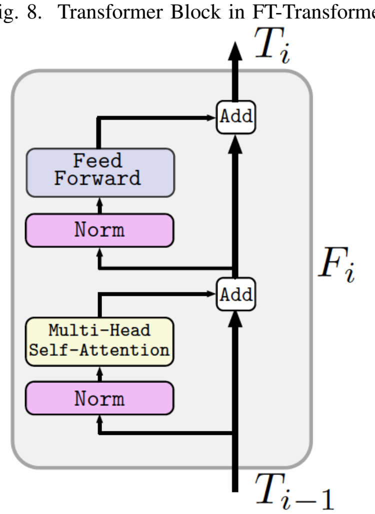
Fig. 7. FT-Transformer Architecture



As stated in the original paper, FT-Transformer uses pre-normalization, which differs from post-normalization in that the input will be added with the output of multi-head self-attention (MHSA) directly without normalization. Another detail worth noting is that the first Transformer layer will skip the first normalization given that the experiment results conducted by the original authors showed that this could improve the performance of FT-Transformer significantly.

The feature importance of FT-Transformer is different from the permutation method mentioned at the beginning of this section. Transformer architecture uses an attention mechanism and the attention maps are used for evaluating feature importance. To make sure this method is consistent with the results derived from more general analysis methods, the original paper used permutation tests (PT) and got consistent rank order results. In this way I make sure the different feature importance

Fig. 8. Transformer Block in FT-Transformer



analysis method won't impact the result so that the result from FT-Transformer is comparable with other models that use PT.

3) Fine-tuning: As for fine-tuning, my focus is on tuning the model architecture as well as the training process. Even though the ANN models and FT-Transformer have quite different structures and thus have different hyperparameters, the tuning principles behind them are the same. All the hyperparameters can be categorized as

- **Regularization** dropout rate, batch normalization, early stop, etc.
- **Architecture** layer number, units, and activation function.
- **Training process** epoch number, batch size, learning rate.

C. Huizi

1) Random Forest: Random Forest (RF) as we can see in Fig. 9 is an ensemble-based method that builds upon the foundational principles of decision trees. By introducing randomness at multiple stages, RF ensures diversity among its constituent trees, thereby enhancing overall prediction accuracy [7]. The ensemble approach ensures that the individual biases and potential errors of constituent trees are averaged out, leading to a model with enhanced robustness and generalizability.

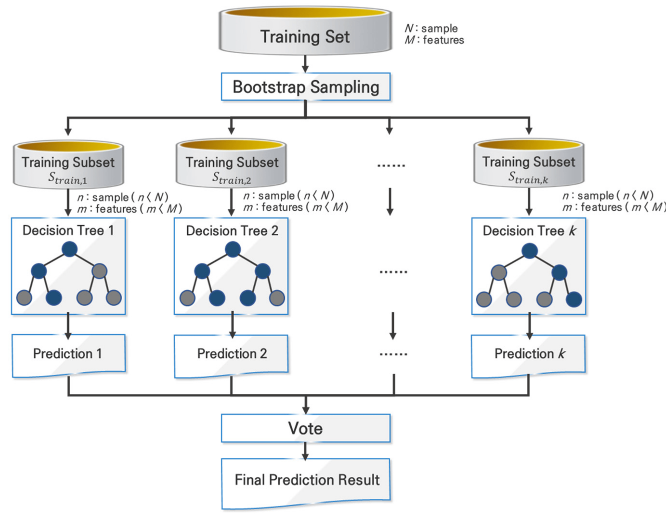
This method operates by constructing multiple decision trees during training and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. A key concept in Random Forest is 'bagging' or bootstrap aggregating, which helps in reducing variance and avoiding overfitting. Mathematically, if we have N trees and $Y_i(x)$ is the prediction of the i^{th} tree, the final output is given by

$$Y(x) = \frac{1}{N}\sum_{i=1}^{N} Y_i(x)$$

2) Key Components and Benefits:

- **Ensemble of Decision Trees:** RF creates a collective of decision trees, where each tree's errors are corrected by subsequent trees.
- **Bagging or Bootstrap Aggregating:** This technique reduces variance and enhances model robustness against overfitting by randomly sampling data points with replacement for tree construction.
- **Tree Pruning and Depth Control:** RF regulates tree depth to avoid excessive branching, ensuring a balanced model complexity.

Fig. 9. Random Forest Architecture



3) Implementation in Wildfire Prediction: RF is used to differentiate between wildfire occurrences and non-occurrences, considering various seasonal features from 2019 to 2022.

- **Data Preprocessing:** The dataset is preprocessed for balance and consistency, involving steps like normalization, handling missing values, and categorical data encoding, as detailed in Section III.B.
- **Feature Importance Analysis:** Utilizing the `model.feature_importance` attribute from the RF library, feature importance is calculated and visualized for insightful interpretations.

4) Hyperparameter Tuning: For hyperparameter tuning, Randomized Search CV is employed to efficiently explore various hyperparameter combinations, enhancing model performance while managing computational resources.

- **Number of Trees (n_estimators):** This parameter determines the number of decision trees. A higher number often results in better performance but at the cost of computational time. [34]
- **Maximum Depth of Trees (max_depth):** It affects the depth of individual trees, influencing the model's complexity and its ability to capture intricate patterns [26]
- **Minimum Samples for Split (min_samples_split):** A critical parameter that affects the decision of when a node should split, with implications for model granularity.
- **Bootstrap Sampling:** Employs a sampling technique, where data points are sampled with or without replace-

ment. This introduces variability among trees, a cornerstone for RF's ensemble nature [7]

5) Feature Importance: An essential step in understanding the predictive power of different features. In RF, importance is gauged by the frequency and depth of a feature's use across all trees. Visualization of feature importance is done through bar charts, helping identify key predictors in wildfire occurrences.

6) Model Architecture: RF was chosen for its effectiveness in handling large, complex datasets, and its robustness against overfitting. Its ensemble nature allows for a comprehensive analysis of various factors influencing wildfires.

V. Simulation Plan

A. Build Dataset

To investigate how different predictors' effectiveness varies across seasons, we define Spring as March to May (3-5), Summer as June to September (6-9) and Fall as October to December (10-12). We sampled 400 points each month for every season across Canada from 2018 to 2021. Among the samples, we selected samples from 2018 to 2020 as training and validation dataset, and samples from 2021 as test dataset. We split training and validation dataset with ratio 4:1.

To summarize, see the description below:

- **Training dataset:** From 2018 to 2020, with 5760 samples for each season.
- **Validation dataset:** From 2018 to 2020, with 1440 samples for each season.
- **Test dataset:** In 2021, with 2400 samples for each season.

B. Choose Model Parameters

1) Mia: As mentioned before, I trained 3 kinds of models: baseline DNN, deeper DNN and FT-Transformer. I fine-tuned both deeper DNN and FT-Transformer with the objective of maximizing both accuracy and recall metrics. The tuner library I used is KerasTuner [23].

For both baseline DNN and deeper DNN, the preliminary training takes 10 epochs. The tuning for deeper DNN takes 10 epochs and 10 trials with hyperparameters including unit number, layer number, activation function, dropout rate and learning rate.

For FT-Transformer, the preliminary training takes 200 epochs with early stopping callback that monitors the validation accuracy. According to the training curve plotted, the model converged at around 10 to 20 epochs. So I set 20 epochs and 30 trials for tuning.

2) Justin: While tuning hyperparameter and analyze feature importance, For each season, the following steps is performed:

Preprocessing: Convert the categorical variable 'labels' into dummy/indicator variables, for both the training and evaluation dataframes. We then separate the features (X) and the target variable (Y) for both datasets.

Model Initialization: Initialize an classifier model.

Hyperparameter Tuning: Perform Randomized Search Cross-Validation with 10 iterations and 3-fold cross-validation. The scoring metric used is 'accuracy'. The best model from the randomized search is then stored in best model list.

Model Training: train the XGBoost model on the training dataset.

Model Evaluation: Use the trained model to make predictions on the evaluation dataset. I then calculate the accuracy and recall of the model, and print these metrics. I also compute and print the confusion matrix.

Feature Importance: Calculate the feature importance using the feature importances attribute of the model.

This process is repeated for each season's data. At the end of the simulation, we have a list of the best models for each season and a dictionary containing the feature importance for each model. These can be used for further analysis and interpretation.

XGBoost: A randomized search was conducted for the XGBoost model with parameters including `n_estimators`, `max_depth`, and `learning_rate`. Three different values for each parameter were tested through 10 iterations and 3-fold cross-validation per seasonal dataset. The best model configurations, chosen based on accuracy, were saved for future use.

Kernel SVM: The Kernel SVM model underwent training with parameters such as `C`, `kernel`, `gamma`, and `degree`. A randomized search was conducted using 10 iterations and 3-fold cross-validation per seasonal dataset to determine the optimal hyperparameters. The best models, selected based on accuracy, were saved for subsequent analysis.

LSTM: The LSTM model's hyperparameters, including `hidden_units`, `lstm_layers`, `dense_units`, and `dropout_rate`, were optimized through random search. Each model was trained for 10 epochs with a batch size of 32. The best-performing LSTM models across different seasons were stored for further analysis.

KNN: Utilizing the K-Nearest Neighbors algorithm, a randomized search was conducted for parameters such as `n_neighbors`, `weights`, and `p`. This search included 10 iterations and 5-fold cross-validation per seasonal dataset to optimize accuracy. The best models, identified through this process, were saved for future use.

TabNet: The TabNet model, as shown in 4, used a randomized search for parameters including `n_d`, `n_a`, `n_steps`, `gamma`, and `n_independent`, 5-Fold cross-validation to employed to evaluate each model's accuracy. The best models, selected based on performance, were saved for further examination.

1D-CNN, Logistic Regression, Naive Bayes: Similar procedures were followed for the 1D-CNN, Logistic Regression, and Naive Bayes models. Each model have list of specific parameter tuning and cross-validation techniques to identify the best configurations based on accuracy for each season's dataset.

3) Huizi: Random Forest algorithm simulation plan.

Objective: To improve the accuracy and reliability of wildfire predictions across different seasons using the Random Forest algorithm.

Data Preparation and Sampling: Gather and label seasonal data over multiple years, preparing separate datasets for Spring, Summer, and Fall.

Model Training and Hyperparameter Tuning: Train the Random Forest model on each seasonal dataset. Optimize hyperparameters using RandomizedSearchCV to enhance model performance for each specific season.

Cross-Validation and Feature Importance Analysis: Implement cross-validation for model stability and generalizability. Conduct feature importance analysis to identify key predictors for each season.

Performance Evaluation: Evaluate the model's performance using Precision, Recall, and F1-Score. Employ confusion matrices for detailed analysis and to understand the model's predictive capabilities in different seasonal contexts.

Parameter Comparison: Compare the performance of models with default settings against those with optimized parameters, emphasizing improvements achieved through tuning.

Outcome Compilation: Repeat the process for each season's data. Compile the best-performing models for each season and create a dictionary outlining feature importance for each model.

Expected Outcomes: Enhanced predictive accuracy across various seasons, with insights into the most significant predictors for different seasonal contexts. Overall improvement in the model's robustness and adaptability, leading to more effective wildfire prediction and management strategies

VI. Results

A. Model Selection

1) training and evaluation: Figure 11 and 10 illustrates the recall and accuracy for validation across seasons for various models. For Spring, we choose finetuned Artificial Neural Network have high recall(0.95), with accuracy(0.75). For Summer, we choose Random Forest for its consistent high accuracy and recall(0.86). For Fall, we choose XGBoost for its consistent high accuracy and recall(0.89).

Fig. 10. Accuracy across Seasons for Different Models

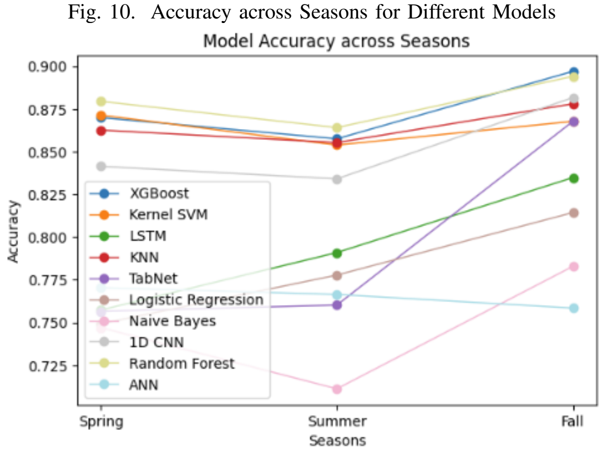
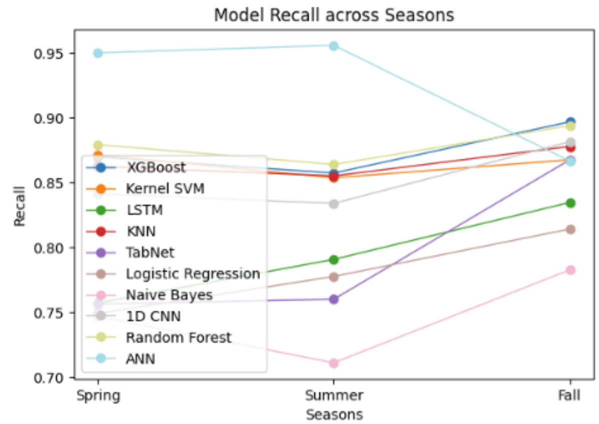


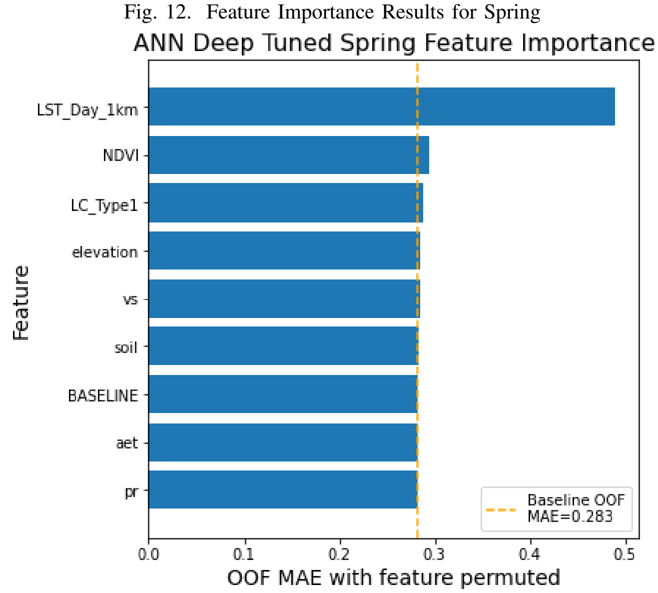
Fig. 11. Recall across Seasons for Different Models



B. Feature Importance Analysis

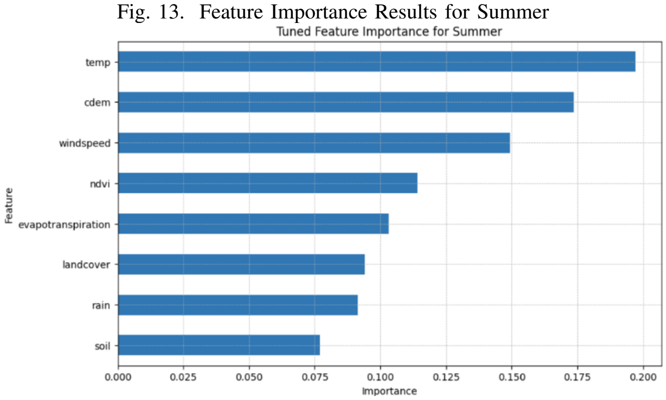
1) Spring: As we can see in 12, the most important feature for Spring is Land Surface Temperature. This finding aligns with our understanding of wildfires, where temperature plays a importance role in influencing vegetation flammability and fire spread.

Fig. 12. Feature Importance Results for Spring



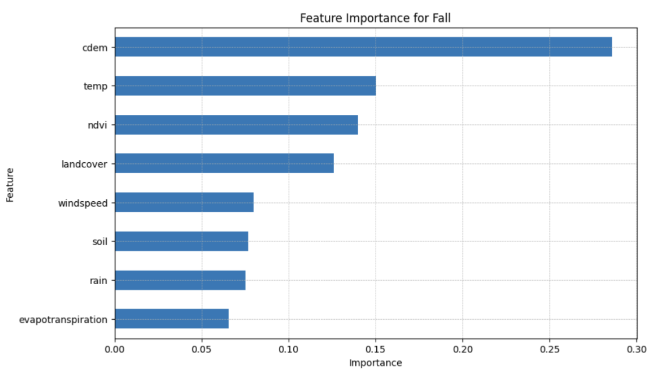
2) Summer: The top3 importance feature for summer wildfire prediction are temperature, wind speed, and elevation 13. High temperatures contribute to the drying of vegetation, making it more susceptible to ignition. In hot conditions, the moisture content in plants decreases, creating favorable conditions for the rapid spread of wildfires. Strong winds can carry embers over long distances, potentially igniting new fires. Wind also influences the speed at which a fire spreads through vegetation. Additionally, different elevations may have varying vegetation types, each with its own fire behavior characteristics.

Fig. 13. Feature Importance Results for Summer



3) Fall: As in 14, the top 3 important feature for Fall is elevation, temperature and Normalized Difference Vegetation Index(NDVI). During the Fall season, elevation plays a crucial role in influencing temperature, humidity, and atmospheric conditions. Higher elevations tend to have cooler temperatures, which can affect vegetation and wildfire behavior. Lower NDVI values may indicate decreased vegetation cover and increased vulnerability to wildfires.

Fig. 14. Feature Importance Results for Fall



4) Feature Analysis: The feature analysis across the three seasons provides insights for key factors influencing wildfire predictions. In Spring, Land Surface Temperature emerges as the most critical feature, aligning with the understanding of how temperature impacts vegetation flammability and fire spread.

Moving into Summer, temperature, wind speed, and elevation stand out as the top three influential factors. The combination of high temperatures, strong winds, and varying elevations contributes to conditions for rapid wildfire propagation.

In Fall, top features are elevation, temperature, and Normalized Difference Vegetation Index (NDVI). Lower NDVI values indicate decreased vegetation cover and heightened vulnerability to wildfires.

VII. Limitation and Future Work

There are several limitations in our work:

- **Lack certain features** In some literature for wildfire prediction, researchers found that anthropological factors also play an important role in accurately predicting wildfire occurrences. We didn't include such features because they are inaccessible from GEE.
- **Lack of exploration for seasonal performance differences** For certain models, the performance can vary significantly across the three seasons. We didn't find enough information that can explain the differences. More training data may be needed to ensure the models can learn enough patterns for the poorly-performed seasons.

In the future, besides including more possibly relevant features in our study, and increasing training data to address the limitations above, we will focus on improving our results and potential for applications with the following experiments:

- **Excluding less important features** A potential future research direction is conducting experiments where features with low importance are excluded.
- **Exploring beyond the default architectures** In this study, we borrowed models from existing research without applying any changes to the architectures. For example, in FT-Transformer, there are other options to embed the numeric features other than the piecewise linear method we used, such as periodic encoding.
- **Application** Getting decent results is just the first step. What makes the results meaningful is figuring out what we can do in wildfire management. To achieve this goal, more studies will be needed outside the realm of machine learning.

VIII. Conclusion

In conclusion, our study has elucidated the dynamic nature of wildfire predictability across different seasons. We demonstrated that factors such as the Digital Elevation Model (CDEM) and Normalized Difference Vegetation Index (NDVI) hold varying degrees of importance depending on the season, which underscores the complexity of wildfire prediction. Spring's susceptibility to wildfires is largely dictated by Land Surface Temperature, while summer's risk is heightened by a trio of temperature, wind speed, and elevation. Come fall, elevation, temperature, and NDVI values, become more significant. Our research faces limitations due to data constraints, notably the exclusion of certain variables that could potentially enhance model accuracy. Future research should focus on expanding the dataset and incorporating more comprehensive feature sets, including anthropogenic factors, to refine the predictive models further. Ultimately, our findings advocate for the development of adaptable, season-specific wildfire management strategies, potentially aiding policymakers and emergency response teams in mitigating wildfire risks effectively. This study lays the groundwork for future exploration in the realm of precision forecasting, moving towards a more proactive and targeted approach in wildfire management.

REFERENCES

[1] A. Kadra, M. Lindauer, F. Hutter, and J. Grabocka, "Well-tuned simple nets excel on tabular datasets," in Proc. Adv. Neural Inf. Process. Syst., 2021, pp. 1–14.

[2] S. O. Arik and T. Pfister, "TabNet: Attentive Interpretable Tabular Learning." arXiv, Dec. 09, 2020. Accessed: Nov. 30, 2023. [Online]. Available: http://arxiv.org/abs/1908.07442

[3] V. L. S. Arruda, V. J. Piontekowski, A. Alencar, R. S. Pereira, and E. A. T. Matricardi, "An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna," Remote Sensing Applications: Society and Environment, vol. 22, p. 100472, Apr. 2021, doi: 10.1016/j.rsase.2021.100472.

[4] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, "Vision Transformers for Remote Sensing Image Classification," Remote Sensing, vol. 13, no. 3, Art. no. 3, Jan. 2021, doi: 10.3390/rs13030516.

[5] J. Bergstra and Y. Bengio, "Random Search for Hyper-Parameter Optimization".

[6] Borisov, Vadim, et al. "Deep neural networks and tabular data: A survey." IEEE Transactions on Neural Networks and Learning Systems (2022).

[7] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[8] N. R. Canada, "Canadian Wildland Fire Information System — Canadian Forest Fire Weather Index (FWI) System." Accessed: Sep. 26, 2023. [Online]. Available: https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi

[9] Canadian Broadcasting Corporation (CBC). "The cost of wildfires in Canada."

[10] G. Charizanos and H. Demirhan, "Bayesian prediction of wildfire event probability using normalized difference vegetation index data from an Australian forest," Ecological Informatics, vol. 73, p. 101899, Mar. 2023, doi: 10.1016/j.ecoinf.2022.101899.

[11] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[12] Environment and Climate Change Canada. "Wildfires and Air Quality."

[13] European Commission. "Forest Fires in Europe, Middle East and North Africa."

[14] Global Fire Monitoring Center (GFMC). "Global Wildland Fire Statistics."

[15] R. Ghali, M. A. Akhloufi, M. Jmal, W. Souidene Mseddi, and R. Attia, "Wildfire Segmentation Using Deep Vision Transformers," Remote Sensing, vol. 13, no. 17, Art. no. 17, Jan. 2021, doi: 10.3390/rs13173527.

[16] Gralewicz, Nicholas J., et al. "Spatial and Temporal Patterns of Wildfire Ignitions in Canada from 1980 to 2006." International Journal of Wildland Fire, vol. 21, no. 3, 2012, p. 230, https://doi.org/10.1071/wf10095. Accessed 14 Apr. 2020.

[17] Henderson, S. B., Brauer, M., MacNab, Y. C., Kennedy, S. M. (2011). Three measures of forest fire smoke exposure and their associations with respiratory and cardiovascular health outcomes in a population-based cohort. Environmental Health Perspectives, 119(9), 1266-1271.

[18] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997 Nov 15;9(8):1735-80. doi: 10.1162/neco.1997.9.8.1735. PMID: 9377276.

[19] Hollmann, Noah, et al. "TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second." ArXiv.org, 7 May 2023, arxiv.org/abs/2207.01848.

[20] Huang, Xin, et al. "TabTransformer: Tabular Data Modeling Using Contextual Embeddings." ArXiv:2012.06678 [Cs], 11 Dec. 2020, arxiv.org/abs/2012.06678.

[21] F. Huot et al., "Deep Learning Models for Predicting Wildfires from Historical Remote-Sensing Data." arXiv, Feb. 10, 2021. doi: 10.48550/arXiv.2010.07445.

[22] Jaafari, E.K. Zenner, M. Panahi, H. Shahabi "Hybrid Artificial Intelligence Models Based on a Neuro-Fuzzy System and Metaheuristic Optimization Algorithms for Spatial Prediction of Wildfire Probability." Agricultural and Forest Meteorology, vol. 266-267, 15 Mar. 2019, pp. 198–207, www.sciencedirect.com/science/article/pii/S0168192318304088, https://doi.org/10.1016/j.agrformet.2018.12.015.

[23] Team, Keras. "Keras Documentation: KerasTuner." Keras.io, keras.io/keras_tuner/.

[24] Kossen, Jannik, et al. "Self-Attention between Datapoints: Going beyond Individual Input-Output Pairs in Deep Learning." ArXiv.org, 1 Feb. 2022, arxiv.org/abs/2106.02584. Accessed 4 Dec. 2023.

[25] L. Breiman, "Random Forests", Machine Learning, 45(1), 5-32, 2001.

[26] Louppe, G., Wehenkel, L., Sutera, A., Geurts, P. (2013). Understanding variable importances in forests of randomized trees. Advances in neural information processing systems, 431-439.

[27] Malik, Ashima, et al. "Data-Driven Wildfire Risk Prediction in Northern California." Atmosphere, vol. 12, no. 1, 1 Jan. 2021, p. 109, www.mdpi.com/2073-4433/12/1/109/htm, https://doi.org/10.3390/atmos12010109.

[28] N.T.Toan, P.T. Cong, N.Q.V.Hung and J.Jo, "A deep learning approach for early wildfire detection from hyperspectral satellite images", IEEE Xplore, 2021. DOI: 10.1109/RITAPP.2019.8932740.

[29] NASA Earth Observatory. "World of Change: Global Temperatures and Wildfires."

[30] Natural Resources Canada. "Canadian Wildland Fire Information System."

[31] Oliveira, Sandra, et al. "Modeling Spatial Patterns of Fire Occurrence in Mediterranean Europe Using Multiple Regression and Random Forest." Forest Ecology and Management, vol. 275, July 2012, pp. 117–129, https://doi.org/10.1016/j.foreco.2012.03.003. Accessed 12 May 2020.

[32] Permutation feature importance scikit-learn. Accessed: Oct. 11, 2023. [Online]. Available: https://scikit-learn/stable/modules/permutation_importance.html

[33] Y. Piao, D. Lee, S. Park, H. G. Kim, and Y. Jin, "Forest fire susceptibility assessment using google earth engine in Gangwon-do, Republic of Korea," Geomatics, Natural Hazards and Risk, vol. 13, no. 1, pp. 432–450, Dec. 2022, doi: 10.1080/19475705.2022.2030808.

[34] Probst, P., Wright, M. N., Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301.

[35] Somepalli, Gowthami, et al. "SAINT: Improved Neural Networks for Tabular Data via Row Attention and Contrastive Pre-Training." ArXiv:2106.01342 [Cs, Stat], 2 June 2021, arxiv.org/abs/2106.01342.

[36] S. Tavakkoli Piralilou et al., "A Google Earth Engine Approach for Wildfire Susceptibility Prediction Fusion with Remote Sensing Data of Different Spatial Resolutions," Remote Sensing, vol. 14, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/rs14030672.

[37] Trucchia, Andrea, et al. "Defining Wildfire Susceptibility Maps in Italy for Understanding Seasonal Wildfire Regimes at the National Level." Fire, vol. 5, no. 1, 21 Feb. 2022, p. 30, https://doi.org/10.3390/fire5010030. Accessed 22 Mar. 2022.

[38] A. Vaswani et al., "Attention Is All You Need." arXiv, Aug. 01, 2023. Accessed: Nov. 09, 2023. [Online]. Available: http://arxiv.org/abs/1706.03762

[39] World Health Organization (WHO). "Health impacts of wildfires."

[40] Wildfire Season Summary - Province of British Columbia. (2017). Gov.bc.ca.372

[41] World Bank. "The Economic Cost of Wildfires."

[42] Yury Gorishniy, et al. "Revisiting Deep Learning Models for Tabular Data." ArXiv (Cornell University), 22 June 2021, https://doi.org/10.48550/arxiv.2106.11959. Accessed 20 Nov. 2023.

[43] Zhang, Guoli, et al. "Deep Neural Networks for Global Wildfire Susceptibility Modelling." Ecological Indicators, vol. 127, Aug. 2021, p. 107735, https://doi.org/10.1016/j.ecolind.2021.107735. Accessed 9 July 2021. history/wildfire-season-summary https://www2.gov.bc.ca/gov/content/safety/wildfire-status/about-bcws/wildfire-373