|  |  |
| --- | --- |
|  | ENGG 680 – Introduction to Digital Engineering |
| *Project Report*  *Wildfire Assessment and Predictive Modelling*  Group 5 – Fall 2024  Vikesh Dharmeshkumar Patel  30255939  Boya Douho  30261119  Kazi Zarin Tasnim Rafa  30233941  Ashkan Einiaghdam  30270232  Ray Pan  30265201  Socretes Saha  30264159  Chunsheng Xiao  30066914 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | *Title of Project* | Wildfire Assessment and Predictive Modelling | | *Group Number* | 5 |   We, the undersigned, certify that this is our own work, which has been done expressly for this course, either without the assistance of any other party or where appropriate we have acknowledged the work of others. Further, we have read and understood the section in the university calendar on plagiarism/cheating/other academic misconduct and we are aware of the implications thereof. We request that the total mark for this assignment be distributed as follows among group members:   |  |  | | --- | --- | | *Your Name* | Vikesh Dharmeshkumar Patel | | *Student ID* | 30255939 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |  |  |  | | --- | --- | | *Your Name* | Boya Douho | | *Student ID* | 30261119 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |  |  |  | | --- | --- | | *Your Name* | Kazi Zarin Tasnim Rafa | | *Student ID* | 30233941 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |  |  |  | | --- | --- | | *Your Name* | Ashkan Einiaghdam | | *Student ID* | 30270232 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  | |  |  | | *Your Name* | Ray Pan | | *Student ID* | 30265201 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |  |  |  | | --- | --- | | *Your Name* | Socretes Saha | | *Student ID* | 30264159 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |  |  |  | | --- | --- | | *Your Name* | Chunsheng Xiao | | *Student ID* | 30066914 | | *Contribution (%) and Hours* | 14.285% , 5 | | *Signature and Date* |  |   *\***Contribution total should be 100%.* |

Table of Contents

1. **Abstract1**
2. **Introduction**3
3. **Problem Statement 4**
4. **Literature Review**5
5. **Methodology**6
6. Data Collection and Preprocessing…………………………………………………………………………………………………………...
7. Exploratory Data Analysis…………………………………………………………………………………………………………………………
8. Target Definition and Features………………………………………………………………………………………………………………..
9. Model Training and Prediction………………………………………………………………………………………………………………..
10. Evaluation Metrics………………………………………………………………………………………………………………………………….
11. Visualization of Predictions…………………………………………………………………………………………………………………….
12. **Results and Discussion……………………………………………………………………………………………………………………….**
13. **Model Limitation……………………………………………………………………………………………………………………………….**
14. **Conclusion and Future Work……………………………………………………………………………………………………………..**
15. **References…………………………………………………………………………………………………………………………………………**

# Abstract

The environmental and economic hazards posed by wildfires are severe, particularly in regions such as Alberta, and Canada, where they are relatively common and cause significant harm. In this research, a predictive model will be developed to analyze and predict the dangers associated with wildfires, based on essential environmental factors that include temperature, humidity, and wind speed. The proposed model uses a hybrid approach by integrating the strengths of CNN, Google Earth Engine API, Random Forests, NDVI, and Support Vector Regression methods to enhance predictive performance. Data sources will include Sentinel satellite imagery and meteorological data from Google Earth Engine and the National Weather Service API, respectively, processed in Python using libraries like Rasterio and Pandas. The aim is to give an effective and location-specific tool to assist in wildfire prevention and response strategy, having wider applications in environmental management, infrastructure protection, and disaster preparedness. This project underlines leveraging engineering techniques to address emergent complex natural disasters with data-driven and real-time solutions.

# Introduction

Wildfires are the most disastrous natural events that have posed a threat to ecosystems, communities, and infrastructure. Alberta, Canada, in recent years, has been facing its share of growing incidents of wildfires due to climatic changes and regional environmental reasons. Such incidents, arising out of unpredictability and intensity, therefore call for the development of strong predictive systems to analyze the risk of wildfires and their timely warnings. It addresses this need by designing an advanced Wildfire Prediction Model using the power of machine learning algorithms. We will include broad data sources such as Sentinel satellite images and historical weather data in a combined model of CNN, Random Forests, and Support Vector Regression to provide a forecast of the likelihood of wildfires with a high degree of accuracy. This predictive model will be a tool in the management of wildfires for mitigating the impacts of such disasters and improving preparedness by evolving data-driven insights. In addition, the project discusses the general use of remote sensing and machine learning in engineering over a wide range of environmental monitoring and uses, and puts forth a solution scalable to regional and global efforts on wildfire prevention.

# Problem Statement

Wildfires have always been a recurring threat to the natural wilderness and wildlife habitat in Alberta, and to human life and communities. Over time, the frequency of these disasters has increased and causes immense loss both to the environment as well as economically each year. Traditional approaches to wildfire prediction depend on the same dataset type and single model approach that could never represent the complex nature of environmental variables affecting the origin and development process of wildfires. This project tries to fill these gaps by constructing one comprehensive predictive model for Alberta specifically. The approach being undertaken in this present study forms an ensemble of different machine learning models: Convolutional Neural Networks, Random Forest, and Support Vector Regression. It does so in a manner that gives it a superior predictive power, leveraging a lot of different inputs-mostly satellite images and historical weather conditions. In integrating these datasets and techniques, the project aims to develop a robust system that can help determine the risks of wildfires at various environmental scenarios, with the aim of strengthening Alberta's preparedness and response against such hazards. This initiative tends to show how data-driven tools can minimize the impact of a wildfire, which is important for engineering innovations in environmental protection and disaster responsiveness.

Literature Review

Prediction of wildfires has been an active area of research in recent years. Many machine learning methods have shown promising performance that contributes to enhancing the accuracy and reliability of wildfire occurrence prediction. Early-stage risk assessment has been performed using conventional methods such as logistic regression, which shows 84.4% success rates using parameters like temperature and humidity (Nikova & Deliyski, 2023). Advanced methods, including Decision Trees and Random Forests, enhance interpretability and robustness, making them suitable for complex, non-linear environmental relationships (Collins et al., 2018). Deep learning models like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have emerged as effective tools for spatial and temporal data analysis, achieving wildfire detection accuracies of up to 98.47% (Guo et al., 2022). Despite these developments, significant gaps persist particularly in integrating diverse data sources and the interoperability of diverse regions; most the studies isolate models involving small datasets, thus rendering their solutions inapplicable to heterogeneous environments such as Alberta. These limitations are overcome by this project, which fuses CNN, Random Forest, and Support Vector Regression into the hybrid model, making use of diverse datasets of Sentinel satellite imagery and historic meteorological data to improve predictive accuracy and increase regional adaptability.

# Results

*.*

## ***RandomForestClassifier***

# Discussion

As expected, and can be seen in figure 1, the accuracy of the numerical integration improves when integration step size is reduced.

# Conclusion

Briefly describe your work and results.

# Appendix

*Present your code (****must also submit .ipynb files)****! You can copy and paste from Colab (looks best)*

## Code – Question 1

# prompt: Hello World!

print("Hello World!")