Predicting Forest Fires in Madagascar

Jessica Edwards[1] , **Manana Hakobyan**[1] , **Alexander Lin**[1] and **Christopher Golden**[2]
[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[2]T. H. Chan School of Public Health, Harvard University, Boston, MA, USA
jedwards@college.harvard.edu, mananahakobyan@g.harvard.edu, alin@seas.harvard.edu, golden@hsph.harvard.edu

Abstract

The negative externalities of large-scale forest fires were shown to outweigh the benefits of natural forest fires leading to worsened air quality, erosion, landslides and increased risk of respiratory and cardiovascular diseases. For decades the island of Madagascar has been struggling to manage its forest fires while receiving little attention from the international research community. The current fire management resources and prediction models are based on domain-specific, hand-crafted features which require a separate installment of sensors and adjustment to country-specific needs. These can turn out to be very costly and hard to adapt for the country. Our approach aims to create a cost-efficient, AI-driven approach for predicting forest fires in Madagascar a month in advance. We use existing open-source data from Google Earth Engine to train a neural network for fire detection. Our final model achieves 78% balanced accuracy and 83% fire detection accuracy (recall) outperforming our baseline model.[1]

1 Introduction

The increasing research on forest fires and its externalities in the past decade elicited their hazardous impact on human health and environment. Even though the fires are a natural phenomenon and are often beneficial for the vegetation purposes, it is shown that the secondary effects of the forest fires such as erosion, landslides, introduction of invasive species, and changes in water quality are more harmful than the fire itself [**?**]. Additionally, the smoke and other pollutants worsen the air quality, increasing the risk of respiratory and cardiovascular diseases [Liu *et al.*, 2015]. These negative externalities of forest fires arouse a serious concern for thoughtful fire management, posing a challenge for researchers and policy makers.

Over centuries Madagascar has been affected by wildfires and agricultural fires, dealing with substantial issues concerning fire management in the country. In 2015, NASA published a report stating that "the island of Madagascar is on fire". In 2019, the Malagasy government called on the international community to aid its fire-fighting efforts as the protected forest lands were burning. Despite all the fire events and the obvious need for research and policy aid, Madagascar, unfortunately, has not received enough attention from AI researchers. Focusing on Madagascar, we aim to bridge this gap and contribute to the country in need.

The existing fire detection resources such as the ones provided by Global Forest Watch (GFW) do not allow enough time for preparation and management [Pro *et al.*, 2014]. The GFW's interactive Fire Map allows one to track the fires in real time as well as access the dates and locations of the previous fires, but does not have a prediction feature. On the other hand, the predominant prediction approaches like the Canadian Forest Fire Danger Rating System are modeling-based and dependent on country-specific features [Wotton, 2009]. The lack of flexibility is a major drawback of these systems since they deploy country and climate-specific handcrafted features. To make it suitable for Madagascar the system would require a separate installment process of various sensors which is very inefficient in terms of monetary cost and time.

This paper utilizes state-of-the-art machine learning models combined with Google Earth Engine's imagery for creating a cost-effective flexible prediction algorithm for fires specifically in Madagascar. The training dataset is taken from Landsat 7 satellite and is further preprocessed to be in the form of histogram bins. It is important to note that the imagery is filtered according to a polygon bounding the entirety of forestland on Madagascar's east coast, hence, narrowing down our predictions to areas most susceptible to forest fire. To be more precise with the predictions, we further divided the polygon into small rectangular regions and considered them as our units for location. The response variable is matched from the Fire Information for Resource Management System (FIRMS) database, which has an archive of all the fire events starting from 2012. Since our goal is to make a prediction model which will allow some time for fire management efforts, the events were matched with satellite imagery with a 5-weeks lag. We train a neural network that performs binary classification task for identifying the fire events for each rectangular region a month in advance. The final model achieves 78% balanced and 83% fire prediction accuracy, ex-

[1]Our code is publicly available at https://github.com/al5250/mdg-fire.

ceeding the results of a baseline model. Balanced accuracy is chosen as a main metric of evaluation because of the existing class imbalance in the data. Our approach is the first to predict the fires in Madagascar one month in advance, potentially allowing time for fire management efforts.

2 Related Work

Fire prediction is not a new idea and it has extensively been researched in the past. The majority of existing fire prediction approaches such as Nepstad *et al.* [1998] are feature-based, heavily rely on domain knowledge, and do not utilize modern AI tools. Despite early fire prediction systems, models like the one described by Cheney *et al.* [1998] try to predict the spread of the fire based on wind speed and other environmental indicators, and contribute with alternative ways to fire management efforts. The system currently in use for fire management in Canada, New Zealand and some European Union countries is a variation of Canadian Forest Fire Danger Rating System (CFFDRS). Canadian Forest Fire Danger Rating system's inputs are human or lightning-made risks, weather, topography and fuels, and the outputs are digit rankings of the fire risk [Wotton, 2009]. While it is very important to create models relying on domain expertise, these models are difficult and costly to adapt and, hence, lack practicality.

Alonso-Betanzos *et al.* [2003] is one of the earliest works to apply machine learning for fire prediction. The authors predict 4 risk categories using neural networks and achieve 78% raw accuracy. There are other papers which deploy neural networks in the domain of fire management, however, they either use environmental variables similar to the earlier model-based approaches or do not use them for fire prediction specifically. No AI-driven fire prediction models have been made for Madagascar. Our main inspiration is taken from Yang *et al.* [2021], who use remote sensing data combined with neural networks for predicting fires in Indonesia. The choice of using remote sensing data makes the approach very flexible and low-cost since the satellite imagery data is open-sourced and available for anyone to use. Our model attempts to synthesize previous wisdom from the domain-based and AI-driven approaches to create a new prediction algorithm for Madagascar.

3 Data

3.1 Landsat 7 Satellite Images

Landsat 7 (U.S. Geological Survey 2019) is part of the Landsat program under NASA and U.S. Geological Survey and has an orbit period of 16 days. Satellite images from Landsat 7 obtained through Google Earth Engine are used as the input source for our hotspot prediction model[2]. In order to obtain these images, a grid was created along the east coast of Madagascar to obtain 8 km $\times$ 8 km images of areas along this region over a one-year span. Using Google Earth Engine, a grid with 2054 areas measuring 8 km by 8 km was generated (Figure 1). We attempt to predict fires for each square in this grid corresponding to a distinct place p in Madagascar.

[2]https://developers.google.com/earth-engine/datasets/catalog/landsat-7

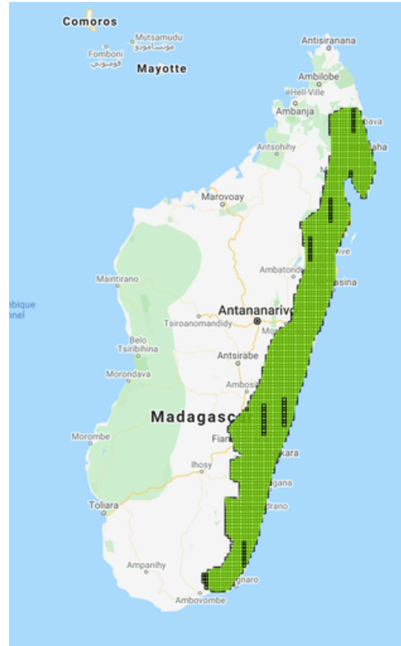

Figure 1: Google Earth Engine 8 km by 8 km grid along the east coast of Madagascar.

Enhanced Thematic Mapper Plus (ETM+), the imaging instrument on Landsat 7 satellite, produces imagery of Earth as eight spectral bands. Following prior work, we found that seven out of the eight bands are sampled to 30m in the Landsat 7 data available on Google Earth Engine, so we focus on the resolution on these bands [Yang *et al.*, 2021].

From the satellite imagery, we are able to generate 32-bin histograms of pixel values for each of the seven bands. As a result, we were able to collect satellite images corresponding to one years' worth of data, convert each image to seven histograms of 32 bins, and pair this data with the Fire Information on Resource Management System (FIRMS) data in order to predict "yes" or "no" for fire in a 8 km by 8 km region in Madagascar in one months' time.

The data preprocessing was a highly non-trivial part of this work and we open-source all of our preprocessing code to drive future research in this area.

3.2 Fire Information on Resource Management System (FIRMS) data

In order to train our model, we needed to provide ground truth labels for training and evaluating our predictions. The FIRMS dataset is provided by NASA and used by several fire prediction systems[3]. We used the Fire Information on Resource Management System (FIRMS) archive and preprocessed the data beginning from 2014 in order to predict the fire events roughly a month in advance. In order to do this, the data was matched with the Landsat 7 satellite image dataset with a 5-week lag for each 8 km by 8 km region. Moreover, we allowed a week-long period for the fire event to happen. Due to

[3]https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms

the fact that the response variable (existence of the fire event in each 8 km by 8 km region) in the final dataset is very unbalanced, we introduced "fire weight" in our prediction model to account for this imbalance.

4 Fire Prediction Method

4.1 Task Definition

From our data, we construct training and testing datasets of place-time pairs (p, t). For each place time pair, we create a three-dimensional feature tensor $\boldsymbol{x} \in \mathbb{R}^{T \times C \times B}$ based on our pre-processing of Landsat 7 satellite imagery, as described in Section 3. Here, T is the number of historical time points before t, C is the number of image bands, and B is the number of histogram bins. The Landsat 7 satellite has an orbit of 16 days; gathering approximately a years' worth of image data for each training example leads to $T = 368/16 = 23$ time points. The satellite collects $C = 7$ bands corresponding to different parts of the electromagnetic spectrum. Following prior work, we set the number of histogram bins to $B = 32$ [Yang *et al.*, 2021]. Thus, the total number of features per training example is $D = 23{\cdot}7{\cdot}32 = 5152$. The label $y \in \{0, 1\}$ indicating either "fire" $(y = 1)$ or "no fire" $(y = 0)$ at time $t + 5$ weeks in location p is obtained from the FIRMS dataset. The goal is to train a classifier $f : \mathbb{R}^{T \times C \times B} \to [0, 1]$ such that $f(\boldsymbol{x})$ gives us an accurate *probability* of a fire breaking out (i.e. $f(\boldsymbol{x}) \approx y$).

For our training set, we use data corresponding to all places p and all times t between 2015 and 2018, which gives us a total of $N_{\text{train}} = 158{,}102$ (p, t)-pairs. For our testing set, we use data corresponding to all places p and all times t in the year 2019, which gives us a total of $N_{\text{test}} = 41{,}575$ (p, t)-pairs. Note that this is in contrast to Yang *et al.* [2021], who only evaluate their models on specific months. Evaluating on a full year gives us a more complete picture of model performance by summarizing across the seasonal variations of different months.

4.2 Models

We train two classifiers – a logistic regression baseline f_{LR} and a neural network ("AI system") f_{NN} to tackle the fire classification task.

Logistic Regression

The logistic regression classifier learns a feature weight $\boldsymbol{w} \in \mathbb{R}^D$ and bias $b \in \mathbb{R}$ such that the prediction for a *flattened* input $\boldsymbol{x} \in \mathbb{R}^D$ is

$$f_{\text{LR}}(\boldsymbol{x}) = \sigma(\boldsymbol{w}^\top \boldsymbol{x} + b), \tag{1}$$

where $\sigma(z) = 1/(1 + \exp(-z))$ is the sigmoid function.

Neural Network

We design the architecture of f_{NN} following prior work [Yang *et al.*, 2021]. Given $\boldsymbol{x}$, we first apply a 2D convolutional layer [LeCun *et al.*, 1995] to integrate data across the image bands and pixel bins dimensions for each time point. We use $H = 64$ filters of size 3×3 and 2D max pooling to summarize across each filter. Then, to integrate data across the time dimension, we feed the $(T \times H)$-dimensional output to a one-layer long short-term memory (LSTM) [Hochreiter and Schmidhuber, 1997] with hidden dimension H. We apply dropout [Srivastava *et al.*, 2014] with rate 0.1 to the H-dimensional representation at the last time point and feed it through three fully connected layers with output sizes 256, 32, 1 and rectified linear unit (ReLU) activations. The final output is passed through the sigmoid function $\sigma(\cdot)$ to obtain a predicted probability.

In comparison to logistic regression, we hope that the neural network can better synthesize spatial and temporal information due to the inductive biases of its compositional layers. Furthermore, the neural network should be able to learn more complicated and highly non-linear decision boundaries that the logistic regression baseline cannot represent.

Training

For fairness of comparison, we keep all training details (other than the aforementioned model architectures) the same for these two classifiers. Given a training example $(\boldsymbol{x}, y)$, we use the binary cross entropy loss function as our training objective, a standard choice in classification problems [Murphy, 2012]:

$$\ell_f(\boldsymbol{x}, y) = y \log f(\boldsymbol{x}) + (1 - y) \log(1 - f(\boldsymbol{x})). \tag{2}$$

To optimize Equation (2) for over all training examples, we use mini-batch gradient descent with batch size 256 for 20 epochs. We employ the Adam optimizer [Kingma and Ba, 2014] with learning rate 0.001 and default hyperparameter settings.

Due to class imbalance in the FIRMS dataset (i.e. "fires" are much less common than "no fires"), we *oversample* the "fire" events by factor P during training. More specfically, when adding an example to the mini-batch, we uniformly select a random "fire" event with probability $P/(P + 1)$ and uniformly select a random "no fire" event with probability $1/(P + 1)$.

Our models are implemented in PyTorch [Paszke *et al.*, 2019] and trained on a Nvidia T4 Tensor Core GPU. We utilize the FastAI library [Howard and Gugger, 2020] for model training, Hydra [Yadan, 2019] for hyperparameter management, and TensorBoard [Girija, 2016] for visualization. We open-source all of our code to help facilitate future research.

4.3 Evaluation

After training, we can obtain a binary prediction $\hat{y} \in \{0, 1\}$ for each example $\boldsymbol{x}$ on the testing set by letting

$$\hat{y} = \begin{cases} 0, & f(\boldsymbol{x}) < 0.5, \\ 1, & f(\boldsymbol{x}) \geq 0.5. \end{cases} \tag{3}$$

Due to class imbalance, it would be inappropriate to utilize accuracy as the main metric to evaluate our models. Indeed, since the "no fire" event occurs 90% of the time, a trivial classifier that always predicts "no fire" obtains 90% accuracy despite being terrible at fire detection. Instead, we choose to focus on *balanced accuracy*, a popular choice for imbalanced binary prediction tasks. This metric is defined as

$$\frac{1}{2} \cdot \frac{\sum_{i=1}^{N_{\text{test}}} \mathbb{I}[\hat{y}_i = y_i = 1]}{\sum_{i=1}^{N_{\text{test}}} \mathbb{I}[y_i = 1]} + \frac{1}{2} \cdot \frac{\sum_{i=1}^{N_{\text{test}}} \mathbb{I}[\hat{y}_i = y_i = 0]}{\sum_{i=1}^{N_{\text{test}}} \mathbb{I}[y_i = 0]}, \tag{4}$$

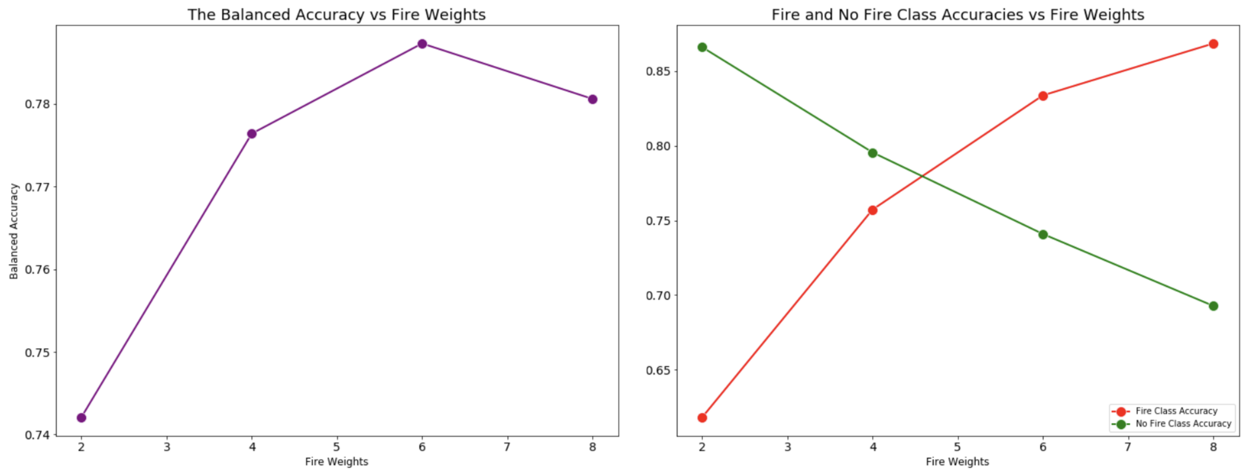

Figure 2: Balanced accuracy (left) and class accuracies (right) as functions of different ”fire” class weights.

Table 1: Evaluated metrics on the testing set of the Madagascar fire prediction task, along with model training time.

Model	**Balanced Acc**	**“Fire” Acc (Recall)**	**“No Fire” Acc**	**Precision**	**Training Time**
Logistic Regression	74.91%	76.15%	73.67%	32.15%	14.1 min
Neural Network	78.73%	83.37%	74.09%	32.47%	21.1 min

where the notation $\mathbb{I}[e]$ denotes 1 if event e is true and 0 otherwise. Intuitively, balanced accuracy is the average over the accuracy for each class. The aforementioned trivial classifier would obtain a low score of 50% on balanced accuracy, indicating that this metric is a much better way to evaluate our problem.

5 Results

We found that for both models, balanced accuracy is optimized at oversampling factor $P = 6$. Table 1 provides the main results comparing the performances of f_{LR} and f_{NN} on the testing set. In addition to balanced accuracy, we also provide other common metrics, such as precision and recall. In terms of the precision-recall tradeoff, our models typically have much higher recall than precision. This phenomenon is preferable for our application, because a false negative (i.e. failing to detect a true fire) is much worse than a false positive (i.e. triggering a false warning when no fire will actually occur).

From Table 1, we see that the neural network outperforms the logistic regression baseline on all reported metrics. In terms of balanced accuracy, the gap in model performance is almost 4%. Furthermore, in terms of *fire detection* accuracy (also known as recall), the neural network outperforms the baseline by over 7%. Thus, there seems to be a notable performance advantage in using the more complex model. In addition, the neural network only requires $1.5\times$ more time to train on our GPU.

Figure 2 presents the effect of changing the “fire” oversampling factor P on balanced accuracy for f_{NN}. As shown in Equation (4), balanced accuracy is comprised of two components – accuracy on the “fire” class and accuracy on the “no fire” class. As P increases, accuracy on the “fire” class increases, since these examples are sampled more during training. On the other hand, as P increases, accuracy on the “no fire” class decreases since these examples are sampled less. Balanced accuracy – which assumes a 50%-50% weighting between the class accuracies – is optimized at $P = 6$. However, depending on the needs of a particular application, P may be tuned to find the best tradeoff between accuracy at fire detection and not having too many false alarms.

6 Conclusion

This work aims to use machine learning models to predict forest fires a month in advance along the east coast of Madagascar, a country that has dealt with substantial issues regarding the management of wildfires and agriculture fires. Using state of the art machine learning models, Landsat 7 satellite imagery via Google Earth Engine, and FIRMS data, we showed that our neural network model achieved a balance accuracy of 78% and a fire-detection accuracy of 83%.

Future work and improvements can be made on this project by incorporating soil moisture data as features within our model to build a more robust neural network that could better predict fires. Additionally, many wildfires in Madagascar are manmade for agricultural purposes, so it would be useful to incorporate human activity data to predict when these agricultural fires may occur.

7 Broader Impact

There are several benefits to continuing further research on our work aimed to predict forest fires in Madagascar. As stated in Section 1, harmful effects such as erosion, landslides, introduction of invasive species, changes in water quality, and decreased air quality which increases the risk of respiratory and cardiovascular diseases are several reasons why policy makers should be well-equipped to manage forest fires [Threat, 2006; Woo *et al.*, 2020].

However, we must be aware that there are also economic and agricultural benefits to fires in the region, and that people in Madagascar may want fires to occur to increase their own prosperity. As a result, a model predicting forest fires with the intent of decreasing the amount of fires that occur in the future could negatively impact communities that rely on agricultural fires for economic and social gain, and could potentially harm impoverished communities. Therefore, this model must be used responsibly by AI researchers and policy makers alike in order to find the appropriate balance between managing negative environmental, social, and health effects and allowing agriculturalists to continue to reap the benefits of fires in regards to vegetation purposes.

References

Amparo Alonso-Betanzos, Oscar Fontenla-Romero, Bertha Guijarro-Berdiñas, Elena Hernández-Pereira, Marıa Inmaculada Paz Andrade, Eulogio Jiménez, Jose Luis Legido Soto, and Tarsy Carballas. An intelligent system for forest fire risk prediction and fire fighting management in galicia. *Expert systems with applications*, 25(4):545–554, 2003.

NP Cheney, JS Gould, and Wr R Catchpole. Prediction of fire spread in grasslands. *International Journal of Wildland Fire*, 8(1):1–13, 1998.

Sanjay Surendranath Girija. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *Software available from tensorflow. org*, 39(9), 2016.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

Jeremy Howard and Sylvain Gugger. Fastai: A layered api for deep learning. *Information*, 11(2):108, 2020.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

Jia C Liu, Gavin Pereira, Sarah A Uhl, Mercedes A Bravo, and Michelle L Bell. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. *Environmental research*, 136:120–132, 2015.

Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

Daniel Nepstad, Adriana Moreira, Adalberto Veríssimo, Paul Lefebvre, Peter Schlesinger, Christopher Potter, Carlos Nobre, Alberto Setzer, Thelma Krug, Ana Cristina Barros, et al. Forest fire prediction and prevention in the brazilian amazon. *Conservation Biology*, 12(5):951–953, 1998.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*, 2019.

Global Forest Watch Pro, Forest Watcher, and Forest Atlases. Global forest watch. 2014.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

A Mounting Threat. Wildfire hazards—a national threat. 2006.

Seung Hyun Lucia Woo, Jia Coco Liu, Xu Yue, Loretta J Mickley, and Michelle L Bell. Air pollution from wildfires and human health vulnerability in alaskan communities under climate change. *Environmental Research Letters*, 15(9):094019, 2020.

B Mike Wotton. Interpreting and using outputs from the canadian forest fire danger rating system in research applications. *Environmental and Ecological Statistics*, 16(2):107–131, 2009.

Omry Yadan. Hydra - a framework for elegantly configuring complex applications. Github, 2019.

Suwei Yang, Massimo Lupascu, and Kuldeep S Meel. Predicting forest fire using remote sensing data and machine learning. *arXiv preprint arXiv:2101.01975*, 2021.