Using Remote Sensing Data and Machine Learning Methods to Predict Wildfire Severity

by

Geneveive Dorrell

A THESIS

Presented to the Computer Science department of the University of Oregon
in partial fulfillment of the requirements for
departmental honors upon receiving a
Bachelors of Science degree

THESIS APPROVAL PAGE

Student: Genevieve Dorrell

Title: Using Remote Sensing Data and Machine Learning Methods to Predict Wildfire Severity

This thesis has been accepted and approved in partial fulfillment of the requirements for departmental honors of bachelors of science degree in the Computer science department by:

Thanh Nguyen Chairperson

Degree awarded Bachelors of Science

THESIS ABSTRACT

Genevieve Dorrell

Bachelors of Science

Computer Science Department

June 2021

Title: Using Remote Sensing Data and Machine Learning Methods to Predict Wildfire Severity

The danger of forest fires has significantly risen over the past decade due to climate change and improper forest management. Wildfires have a severe effect on social and ecological systems. Being able to predict the severity of forest fires would be valuable knowledge to have fighting fires or when allocating resources for forest management. In this work, I attempt to apply machine learning techniques to accomplish this task. Although machine learning algorithms have been shown to be powerful in many other fields, these methods have been underutilized in the prediction of wildfire severity. First, I built the wildfire dataset which consists of various domain features, ranging from land management and logging practices, including logging data, forest composition data, stream locations, stand age index, and satellite imaging. Second, I employed multiple machine learning techniques to predict the severity a wildfire would have on an unburned forest using publicly available datasets and remote sensing data. My models predict the severity class pixel by pixel of the satellite image which can be used to provide fire severity prediction maps. The most promising machine learning techniques were a random forest model, a deep neural network and a 1 dimensional convolutional neural network with the respective accuracies of 66.7%, 61.40 %, and 69.67 %. These results could offer a way to better control and reduce forest fires by helping firefighters fight fires and by predicting future fire severity so that we can target locations for better land management practices.

ACKNOWLEDGMENTS

Thank you Professor Thanh Nguyen for being my advisor during such an unprecedented time. You helped me through the ends and outs of machine learning and computer science. You are a great mentor. Thank you Professor Lucas for helping me narrow my focus and land on the topic and Thank you for being so supportive and welcoming.

TABLE OF CONTENTS

Introduction

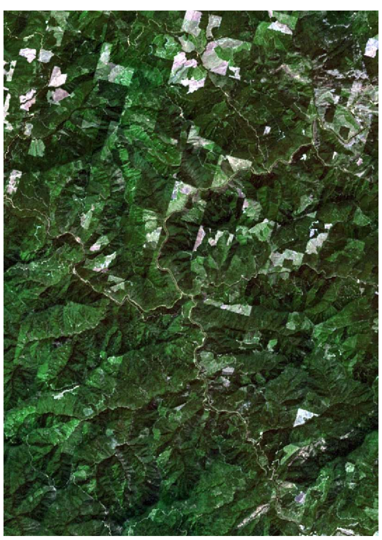
It is widely accepted that forest fires have become more prevalent on the west coast of North America as the global climate has changed over the past several decades (Abatzoglou et. al. 2016). Since 2000, on average 7.0 million Acres burn each season. The number of fires over the last thirty years, while variable, has decreased while the acres burned and has more than doubled since the 1990s (Hoover & Hanson 2021). This is because the severity and danger of large-scale intense forest fires has increased. The fire season on the west coast of North America has become a threat to local ecosystems and a humanitarian crisis. Every year wildfires decimate forests, they cover the western side of the United states in smoke, stress local ecosystems, and they burn the houses of thousands. Being able to predict what forests will burn uncontrollably, would be a valuable tool for land management and fire prevention practices. Machine Learning approaches have been used to create severity maps for fires that have already burned since the 1970s, however Machine Learning approaches have not yet been used in depth to attempt to predict the severity an unburned forest would have (Jain et al. 2020). A 2020 review of machine learning applications in wildlife science and management "found only one study that used ML to predict fire behavior related to fire severity" and that paper was more focused on using a random forest to determine factors that were predictive of severity, instead of trying to use different machine learning methods to predict wildfire severity. When constructing my data set to apply different machine learning techniques to, I chose to only use publicly available datasets that were available for the entire west coast of the United States, so that the prediction techniques used here can be applied to different geographic locations. The United States has been tracking fires since the 1950s through the monitoring trends in burn severity project (MTBS). For every major fire in the United States they create perimeter and severity maps (Eidenshink et al 2007). I used the fire severity map available for the 2013 Oregon Douglas Complex fire that burned 19,760 ha of forestland as the severity labels for my training date set (Zald et al. 2018). For the input training data, it is important to understand how land management practices have changed and the long-term effects of land management on fire severity. Land management practices like logging have already been shown to be strong predictors of forest fire severity in the Pacific Northwest and specifically in the Douglas Complex fire (Zald et al. 2018). The Douglas Complex fire burned a lot of 30-50 years timber rotations (Zald et al. 2018). Because of these timber extraction techniques other papers have found that land ownership is one of the biggest predictors of fire severity (Zald et al. 2018). It has also been identified that past fires showed higher burn rates in Oregon's most managed forest (stone et al. 2004). In relation to the Douglas Complex fire, the BLM (Bureau of Land Management) and Willamette Forest Service land burned at a lower severity compared to plantation forests, this supports that timber plantations are more susceptible to fire (Zald et al. 2018). In an attempt to train the model on land management practices I included data on the year the forest was last logged, forest composition data from the year 2000, and the stand age index data in the training set. Land management variables will be important data for the model to predict forest fire severity accurately, because of previous literature and the intense ecological effects of clear cutting (Zald 2018). I also included publicly available data on steam locations elevation, aspect and slope. I did not include data on climate because I was predicting severity pixel by pixel so climate would have been the same for all data points for the same fire, but weather would be interesting descriptive data to add to the models. That is why my models are currently very region specific.

Methods

Study site

July 26, 2013 the Douglas Complex fire burned 79,209 ha of forestland in southwestern Oregon ((Eidenshink et al 2007). The fire burned through BML tree plantations and the Willamette national forest (Zald et al 2018). The fires depicted in figure 1 below were started from multiple lightning ignition points, that is why Douglas Complex fire has three entries in the Monitoring Trends in Burn Severity project (MTBS) despite the ignition date for all the fires being the same (Eidenshink et al 2007). I chose this study site because research on the biggest predictors of fire severity have already been published for two sections of this site, and many datasets have data up to the years 2012 and 2013 so right before the Douglas Complex fire burned.

Before:



After:

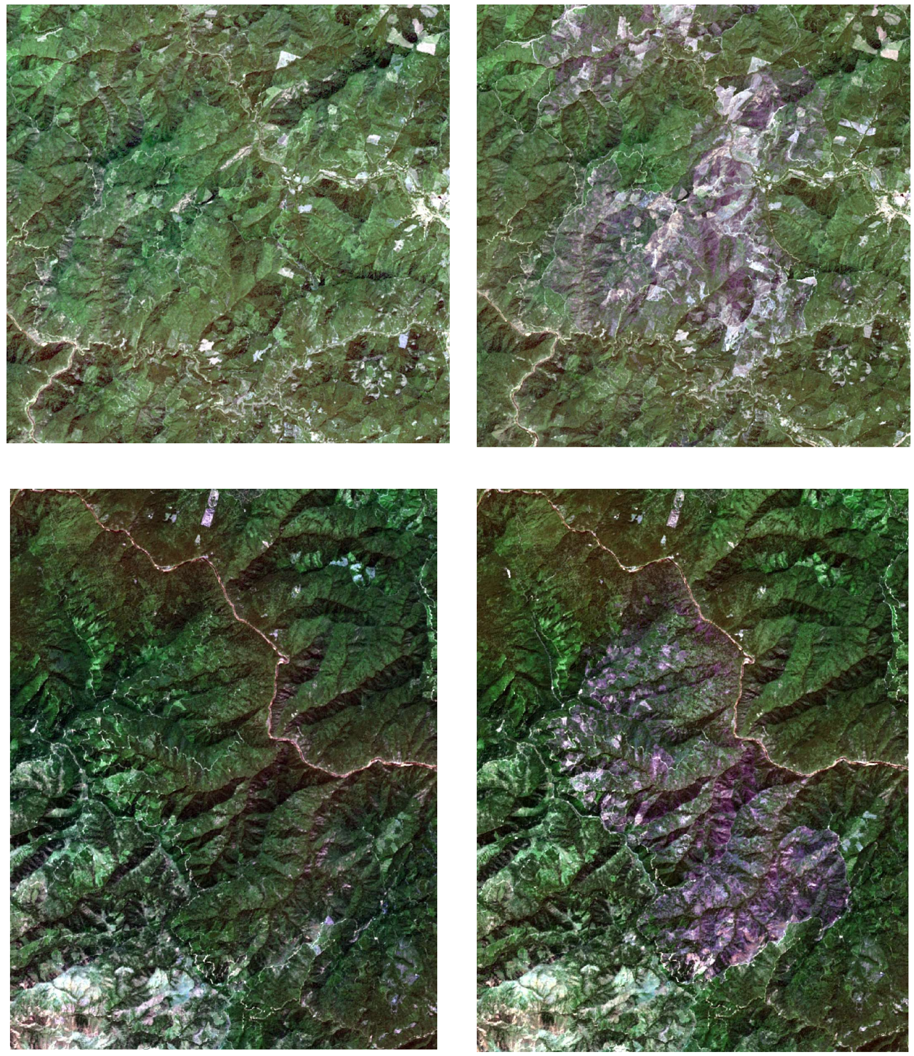


Figure 1. satellite pictures of the burned areas of the study site before and after the fire (Eidenshink et al 2007)

Data sources

All of the data I used to train and test my model is freely available, and with the exception of the stand age data set, the data is regionally available for any area in the United states. I chose to use publicly available datasets that cover a large range of land so that this methodology could be applied to other geographic locations.

Severity Labels

The severity labels I used to classify the data came from the Monitoring Trends in Burn Severity project (Eidenshink et al 2007). The United States monitors all the large wildfires that occur. The Monitoring Trends in Burn Severity project produces a shape file for each fire perimeters and uses Landsat satellite images of before and after the fire to create fire severity maps. The fire severity maps are calculated using the difference between the Normalized Burn Ratio (dNBR) of pre and post fire Landsat images. Specifically for my site the dNBR was calculated using the before and after images depicted in figure 1 (Miller et al. 2007). Lastly to create the severity map, the dNBR is then classified into one of five severity classes: unburned to low, low, moderate, high, and increased greenness. This produces burn severity maps like the two severity maps for my study site in fig 2 below. I used the severity classification of the 30 x 30 m pixels as the labels to train my neural network on. I trained the network on four of the five severity classes, leaving out the pixels with increased greenness due to a lack of data in that category.

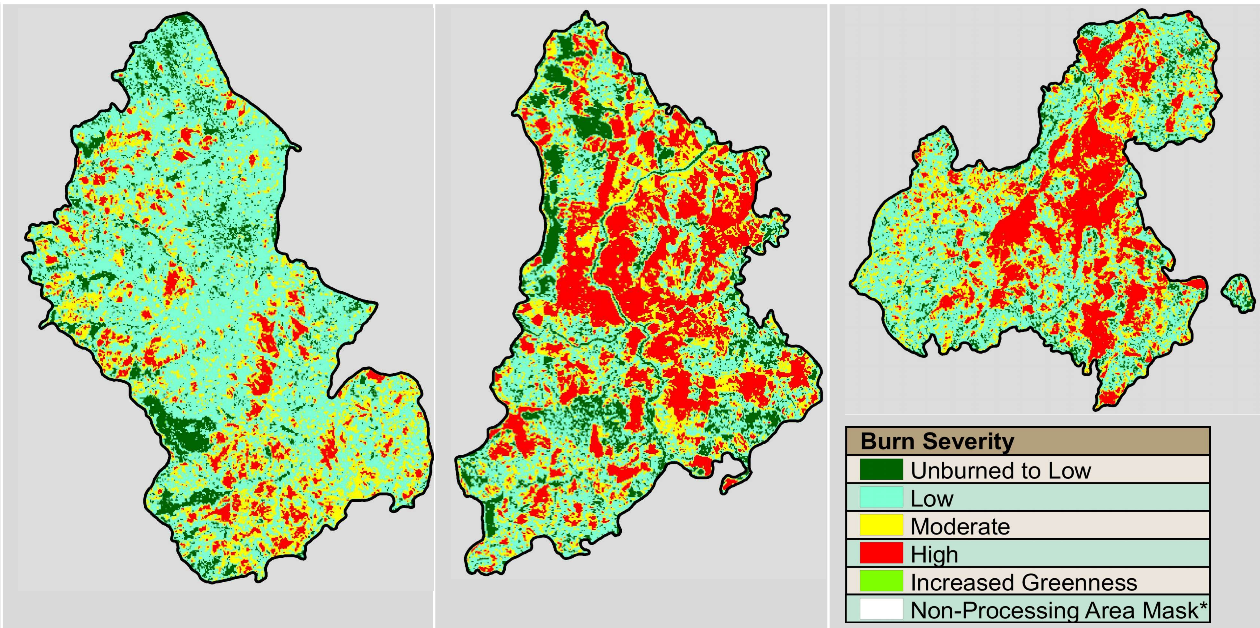


Figure 2. The burn severity maps of the Douglas Complex Fire, and the severity classes of the pixels (Eidenshink et al 2007)

Descriptive data

Because my neural network is classifying each pixel of a 30 x 30 Landsat image I needed more data than just the other Landsat channel values from the pre fire landsat geo tiff. I included forest gain and loss data that is available globally from the Hansen et al 2013 paper, to inform the model on logging practices and other tree loss events. Specifically, I included data on the year there were tree loss events since the year 200 and the overall tree coverage in the year 2000 as depicted in figure 3 below. The loss year since the year 2000 is supposed to infore the model about loss events but because the loss events only go back to the year 200 there is a lot of black space indicating no data. The addition of the data on the tree coverage defined as canopy closure in the year 2000 is an attempt to fill in those gaps with tree coverage data. Both data sets have a resolution of about 30 x 30 meters for each pixel.

a). b).

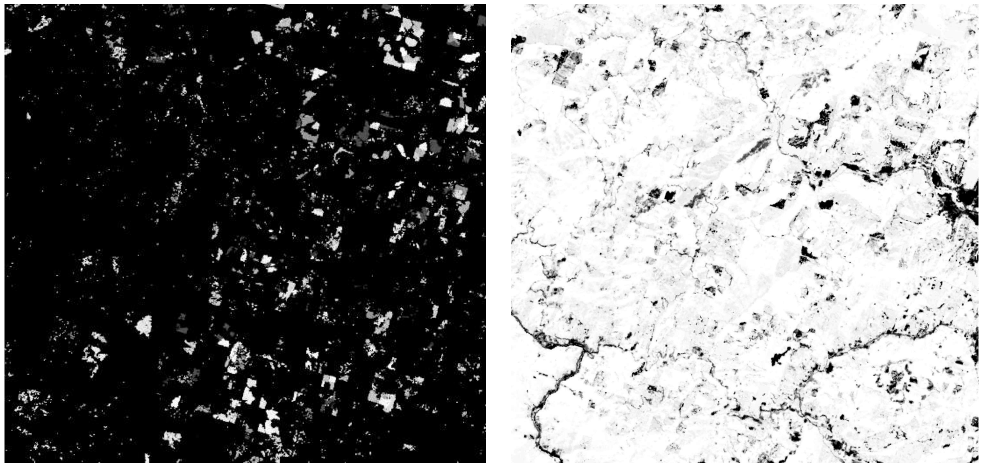


Figure 3. a). Data on the loss year. Black means no loss since the year 2000 and the respective data range from 1-13 a lighter color being a higher number and more recent loss. b). The overall tree canopy coverage in the year 2000 for trees defined as canopy closure for all vegetation taller than 5m in height, with values representing a percentage coverage pre 30 x 30 m pixel with range 0–100 lighter meaning more coverage. (Hansen et al 2013)

I got elevation data from the US government's 3D Elevation Program (Lukas & Baez. 2021) as shown in figure 4 below. I then used the elevation data and the slope tool available in arc gis pro to create a percent rise raster using the geodesic method with meters as the unit of measurement, as shown in figure 4 below. I also used the elevation data to calculate the aspect. I used the arcgis aspect tool and the planar method to calculate the aspect for each pixel. I included the slope and elevation value at each pixel in the data set, as shown in figure 4 below. The slope is useful for determining tree closeness and the aspect, the direction of the slope gives the model information on whether a pixel is on the north or south face of a mountain. I included the elevation slope and aspect geo tiff data in my data. All the geo tiff rasters hava resolution of 5 x 5 m.

a). b). c).

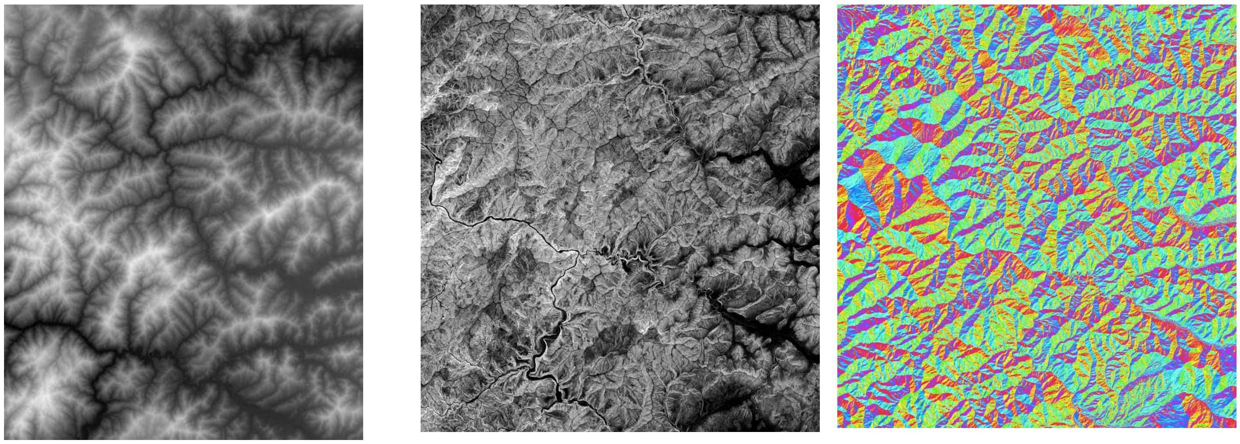


Figure 4. a). elevation geo tiff, lighter means a higher elevation (Lukas & Baez. 2021). b). slope geo tif lighter means a steeper slope. c). aspect 0 - 360 degree direction of slope. Each degree is depicted by a color in a gradient order.

I also included data on each pixel's distance to the closest stream. The stream location data was obtained from the U.S. Forest Service databases in the format of a geodatabase of lines (ALP 2021). I calculated each pixel's distance from to a stream using the Euclidean distance tool available in the ArcGis pro to create an output raster with pixel resolution of 5 x 5 meters, as shown in figure 5 below.

a). b).

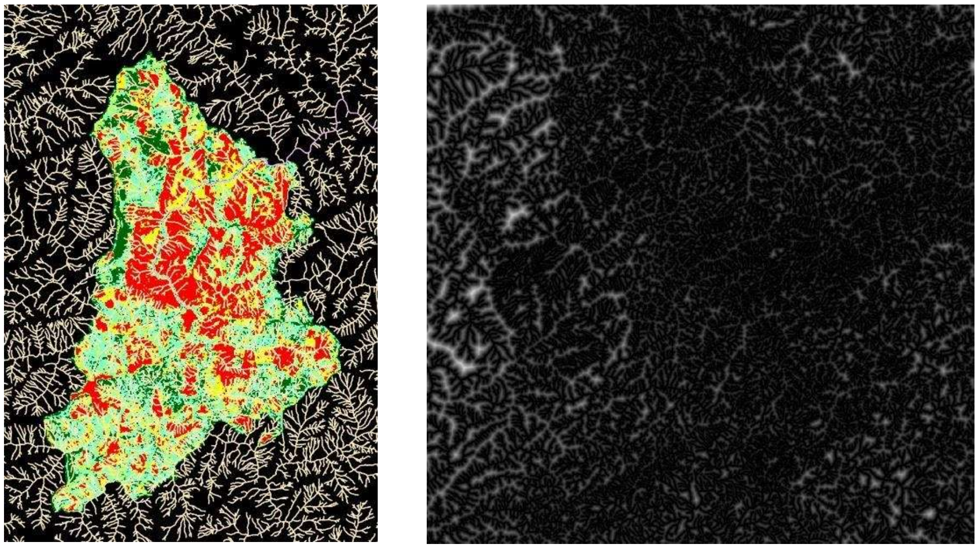


Figure 5. a). Stream lines obtained from the geo database imposed over one of the study site severity maps (Eidenshink et al 2007, ALP 2021). b. The euclidean distance calculated using those stream lines. Darker color means a closer distance to a stream.

I also included data on stand age using the Old-Growth Structure Index dataset made publicly available by the northwest forest plan (Davis, el al. 2013). The Old-Growth Structure Index is a sum of various characteristics that are indicative of old growth forests (Davis, el al. 2013). Specifically, they looked at the density of large live trees, the diversity of live-tree sizes, the density of large snags, and the percentage cover of down woody material. I included the general old-growth structure index data set, the 80 year threshold dataset and 200 year threshold dataset dataset well. All the old-growth structure index data sets had a resolution of 30 x 30 m.

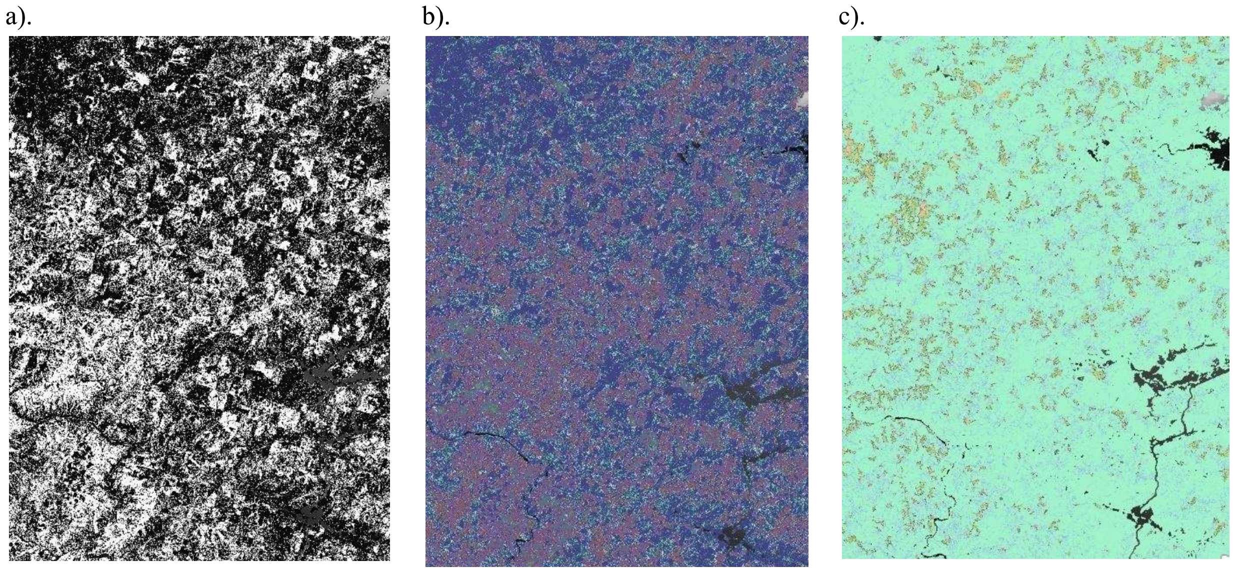


Figure 6). a). The Old-growth structure index sum of ecological features with a range between 0 - 100. lighter indicates an older forest. b). The old-growth structure index for tree stands that are estimated to be at least 80 years old with the ecological features color coded. c). The old-growth structure index for tree stands that are estimated to be at least 80 years old with the ecological features color coded. (Davis, el al. 2013)

Dataset design

The dataset is designed so that each data point represents a 30 x 30 m pixel from a Landsat image. The labels are the severity that each pixel was classified as by the Monitoring Trends in Burn Severity project when they created the severity maps. Each pixel in my data set has the 8 channels of the Landsat image from before the fire burned and then the values of the other descriptive factors I included. The values of the descriptive factors were calculated using the center of each 30 x 30 m Landsat pixel. Each pixel also contains the landsat channels and other descriptive factors of its surrounding pixels because the surrounding pixel's values do influence the fire severity of the central pixel, as illustrated in figure 7 below. The final data set contained a 7 x 7 grid of pixel data points with the central pixel being the pixel that is being classified. This means that I have compiled a data set with pixels that have other descriptive factors as well as the surrounding areas values as input for my models. I used 6,379 test cases randomly selected per class, so the test suite had 25516 data points total. I chose to use this many test cases because that number of test cases consisted of 20% of the total data set. The training dataset is balanced among the classes and has 102,080 data points total.

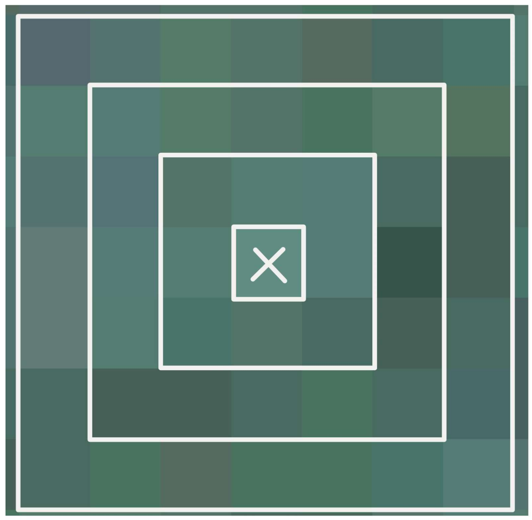


Figure 7. Depicts the two layers of surrounding pixels included as descriptive data for each data point.

Data analysis

I then trained multiple machine learning models on the training data set and used the test data set to determine the different models' predictive power.

Linear regression

Using the Skleanr python package I implemented a Linear regression model. A linear regression model attempts to classify the inputs by separating the different classes using linear lines, planes or in this case hyper planes. The training of a linear regression model involves calibrating the weights associated with the linear lines, planes of hyper planes, by minimizing the residual sum of squares between the labeled data points, and the targets predicted by the linear regression approximation (Galton 1886).

Decision Tree

Using the Skleanr python package I implemented a decision tree model. Decision trees are a non-parametric supervised learning method used for classification and regression. A decision tree is a sequence of binary decisions that split the data into different branches. Branches terminate in leaves when a data point is classified. The decision tree trains by choosing splits that cost the least amount of accuracy (Fürnkranz J. 2011).

Random Forest

Using the Skleanr python package I also implemented a random forest. A random forest is a grouping of Decisions trees. To introduce randomness and differences among the trees a random forest splits nodes based off of the best feature in a random subset of the total features when learning the training data. A random forest comes up with a classification by running a data point through all the decision trees and classifying it into the class that the most decision trees classified it as (Breiman 2001).

Neural networks

Neural Networks are inspired by the human brain. The most basic component of a neural network is a node or a neuron. Each neuron has a set of inputs that are combined linearly by multiplying by their associated weights. The output is then put through an activation function to introduce non-linearity.

Neural Net 1

For my first neural network I used the rectified linear unit (ReLU) activation function. Neural networks learn and remember patterns by optimizing the weights associated with each feature input. For my first Deep Neural network (DNN) I implemented a deep/multilayer feedforward model or multilayer perceptron model, because my neural network had more than 3 hidden layers (LeCun et al. 2015). The neural networks learning rate used was 0.001 and the momentum was 0.9. The loss was calculated using the cross-entropy loss function. The last layer of my neural network has 4 nodes because that is the number of severity classes I am classifying pixels into, and it is put through the softmax function so the output of each node's output represents the probability the pixel is in each of the 4 severity classes as shown in figure 8 below that depicts the architecture of my DNN.

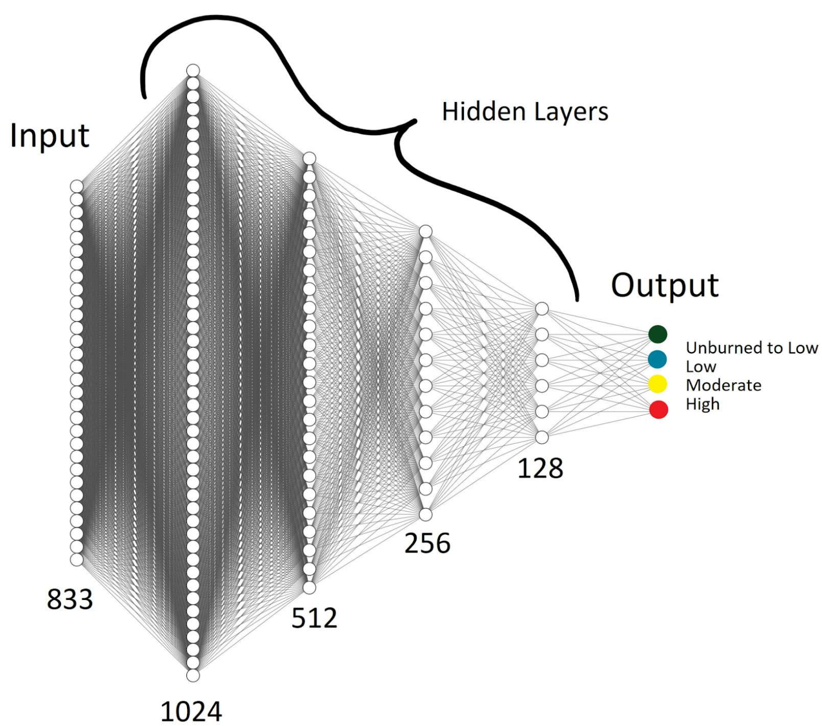


Figure 8. Architecture of Deep pytorch Neural network. Each circle represents a node or neuron. The actual number of nodes per layer used inthe model I developed are written below.

Convolutional Neural network

The convolutional neural network is a type of deep machine learning model inspired by the visual cortex in humans. Convolutional neural networks are better at picking up patterns in the data temporally and spatially. They also are good at reducing the amount of data but not losing predictive power. That is why they are so powerful when it comes to image classification. A convolutional network has a kernel matrix that is multiplied by sections of the input to extract high level features. The output is then put through an activation function which in the case of my 1 dimensional convolutional neural network is the ReLU function again. The convolution layers are then put through a max pooling layer to reduce noise and extract important features. As shown in figure 9 below after the convolutional layers the model then has seven linear neural layers like the layers seen above in the DNN. The last 4 nodes are again put through the softmax function so the output of each node's output is the probability the pixel is in that class.

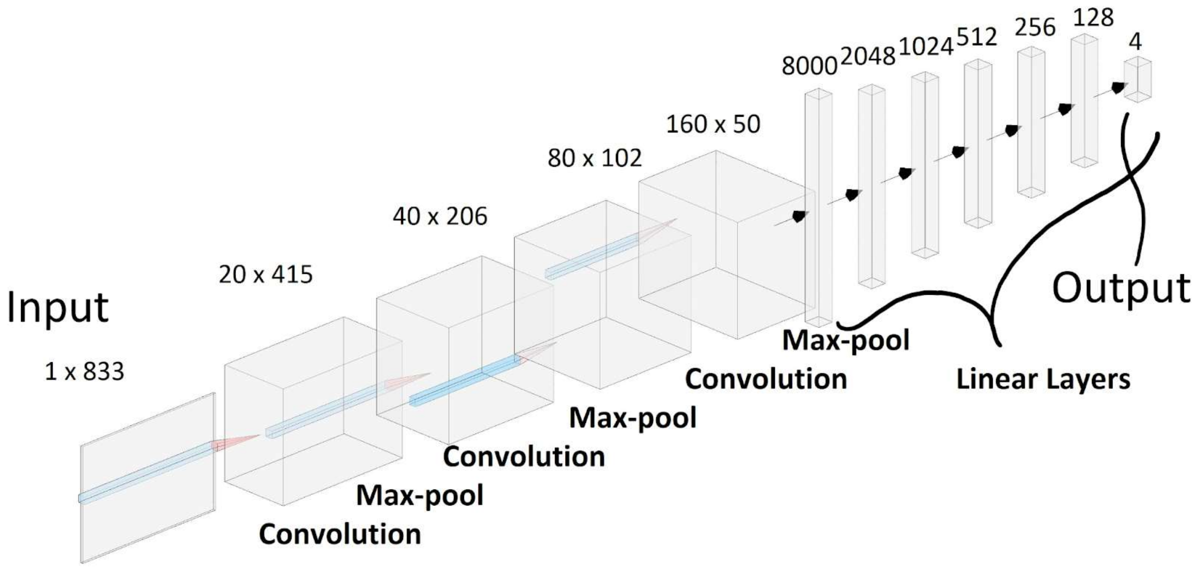


Figure 9. The architecture of my 1D convolutional neural network with layer sizes drawn above.

Results

The logistic regression model had an accuracy of 44.04%. The decision tree model had an accuracy of 44.88%.

Random Forest

The random forest had an overall accuracy of 66.67%, an accuracy of 71% for the lowest severity class, 73% accuracy for the low severity class, 51% accuracy for the moderate severity class, and 72% accuracy for the high severity class. As shown in figure 10 below, if the random forest incorrectly classifies a pixel, it is more likely to predict the pixel is in a severity class close to the correct class. The random forest, out of all the models, was the second most accurate and the best at classifying pixels in the low severity class.

Model's Severity Predictions

Pixel Severity Classification

	Unburned to Low	Low	Moderate	High
Unburned to Low	71%	16%	7%	7%
Low	6%	73%	13%	7%
Moderate	6%	23%	51%	19%
High	5%	10%	13%	72%

Figure 10. The error matrix for the random forest model. The boxes with green background are the accuracy of the model at classifying pixels in that class.

Neural Net 1

The first deep neural network (DNN) had an accuracy of 61.4 % overall, and an accuracy of 74 % for the unburned to low severity class, an accuracy of 66 % for the low burned class, an accuracy 39 % for the moderate severity class and an accuracy of 68 % for the high severity class. As shown in figure 11 below The model overfit to the training data a little bit but overall the loss and accuracy curves show the model learning the training dataset to the best of its capability.

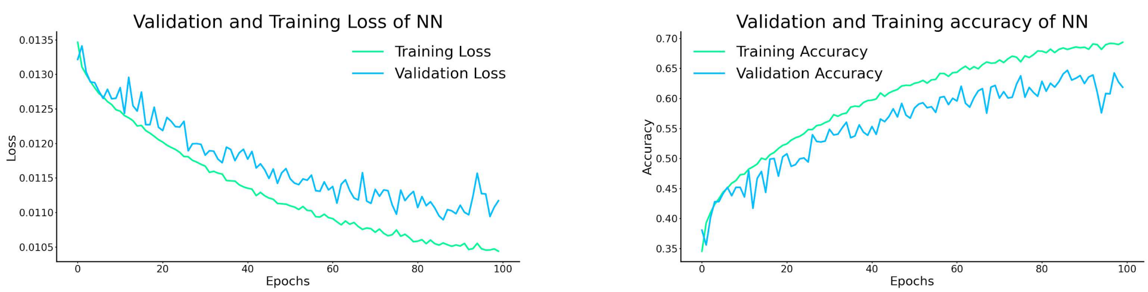


Figure 11. Accuracy and loss averages for the DNN model over epochs during training

As shown in figure 12 below, if the DNN incorrectly classifies a pixel, it is more likely to predict the pixel is in a severity class close to the correct class. The DNN was also more likely to overestimate the severity level than under estimate it. The DNN was best at classifying pixels in the lowest severity class.

Model's Severity Predictions

Pixel Severity Classification	Unburned to Low	Low	Moderate	High
Unburned to Low	74%	15%	4%	7%
Low	17%	66%	9%	8%
Moderate	14%	27%	39%	21%
High	10%	10%	12%	68%

Figure 12. The error matrix for the first DNN. The boxes with green background are the accuracy of the model at classifying pixels in that class.

Convolutional Neural Network

The 1 dimensional convolutional neural network (1D CNN) had an overall accuracy of 69.67%. As shown in figure 13 below the model is overfitting to the training data set meaning the model is too complex for the problem or I need to add more training data.

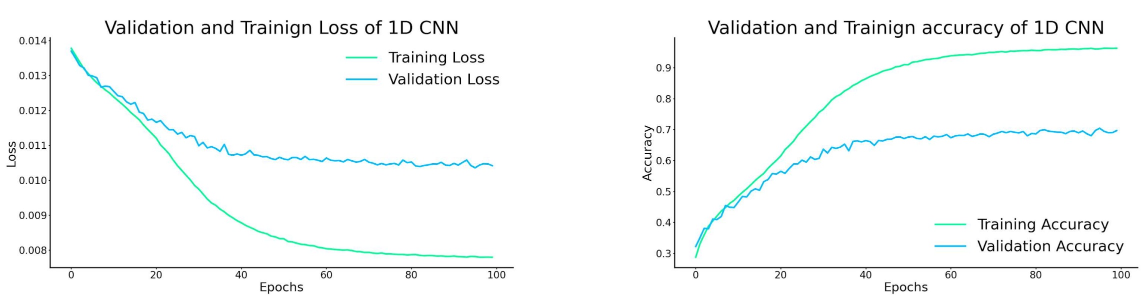


Figure 13. Accuracy and loss averages for the 1D CNN model over epochs during training

The 1DCNN had an accuracy of 75% for the unburned to low severity class, an accuracy of 70% for the low severity class, an accuracy of 58% for the moderate severity class, and an accuracy of 76% for the high severity class. As shown in figure 14 below, if the 1D CNN incorrectly classifies a pixel, it is more likely to predict the pixel is in a severity class close to the correct class. The 1D CNN out of all the models was the most accurate overall and the best at classifying pixels in the unburned to low, moderate and high severity classes.

Model's Severity Predictions

Pixel Severity Classification	Unburned to Low	Low	Moderate	High
Unburned to Low	75%	12%	8%	5%
Low	9%	70%	17%	5%
Moderate	5%	19%	58%	17%
High	3%	5%	16%	76%

Figure 14. The error matrix for the 1D CNN. The boxes with green background are the accuracy of the model at classifying pixels in that class.

Overall, multiple machine learning methods have predictive power when it comes to predicting wildfire severity on a small scale but the CNN was the most successful as shown in figure 15 below.

Model	Accuracy
Linear Regression	44.04 %
Decision Tree	44.88 %
Random Forest	66.67 %
DNN	61.40 %
1D CNN	69.67 %

figure 15. The overall accuracies achieved by all the machine learning methods

Discussion

The other paper found by the 2020 review that attempted to predict wildfire severity examined two of the three fires started on july 26 2013 that I looked at. The paper looked at using a random forest model to predict fire severity and achieved an accuracy of 31% for the entire study site (Zald et al. 2018). That is why it is surprising the random Forest model I implemented was so successful, especially because I included no climate or weather variables. The other paper in the area found that the daily weather was the greatest predictor of severity. The success of the random forest and the other classical machining learning metrics demonstrate that this is a problem that can be modeled using the training data I chose and machine learning techniques.

The DNN model was more accurate than the logistic regression and decision tree machine learning methods. This was expected because DNNs are a method that usually has more predictive power for complicated problems, but it was surprising that the random forest out performed the DNN.

The 1D CNN on the other hand was the most successful and accurate model overall It was the best at predicting the lowest and highest severity classes with respective accuracies of 75% and 76%. This is important because being able to detect forests that are at risk of a high severity fire, would allow for land management intervention to reduce fire risk. On the other hand, being able to accurately predict unburned to low severity areas is important for determining where wildlife refuge may be during or after a wildfire event. Also while actively fighting fires, having a predicted wildfire severity mapr would be useful for the ecologists that help firefighters fight fires. The over-fitting of the model also suggests that it could be trained on a more generalized data set and still have predictive power meaning this model could be the best model to use when scaling up in predictive power regionally.

Limitations

This model was trained on two fires that are in the same geographic region and that were started on the same date. This means the model does not yet scale up to other regions or possibly other weather patterns. However, the data that the model trained on is available along the entire west coast of the United States so this methodology could be used to create other regional models. The model is also only as good as the severity data it was trained on. This model is predicting severity at a very fine scale 30 x 30 meters and based off of the severity classes created by the MBTS projects. The MBTS severity maps are based off of the dNBR, however the classification of the severity classes is variable and subjective (Kolden et al. 2015). However the methodology for calculating the MBTS severity maps is always improving (Picotte et al. 2020).

Future directions

Previous work done on this site found that fire weather was the most important predictor of fire severity so expanding my dataset to include Prism data on weather during the fire would be an interesting topic to investigate for further research (Zald et al 2018). Also looking into how the methodology scales up regionally by using a dataset that includes more fires or attempting to predict the severity of fires the model was not trained on would be very useful.

Conclusion

Overall, the 1D CNN model shows the most promise for scaling up regionally but further research is needed. This research provides a valuable methodology for using machine learning techniques to predict the behavior of future forest fires, by providing insight into how to control the severity of future fires by highlighting high risk areas in forests. This research is built off of other research that suggests controlling logging practices reduces forest fire severity and indirectly promotes healthier forest management. The implications of reducing forest fires has huge economic and social implications. Just as we have seen this last year, large scale forest fires have the capability to cover the entire western coast of the United States in smoke, decimate food production and destroy people's businesses and homes. This research provides a methodology to predict forest fire severity and can hopefully be scaled up and provide some insight on how to manage the increasingly devastating fire season.

References

C. A. Kolden, A. M. S. Smith, J. T. Abatzoglou, Limitations and utilisation of Monitoring Trends in Burn Severity products for assessing wildfire severity in the USA. *Int. J. Wildland Fire* **24**, 1023 (2015).

J. J. Picotte, *et al.*, Changes to the Monitoring Trends in Burn Severity program mapping production procedures and data products. *Fire Ecology* **16**, 16 (2020).

V. Lukas, V. Baez, "3D Elevation Program—Federal best practices" (US Geological Survey, 2021).

R. J. Davis, *et al.*, Northwest Forest Plan--the first 20 years (1994-2013): status and trends of late-successional and old-growth forests. *Gen. Tech. Rep. PNW-GTR-911. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 112 p.* **911** (2015).

F. Pedregosa, *et al.*, Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* **12**, 2825–2830 (2011).

M. C. Hansen, *et al.*, High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).

J. D. Miller, A. E. Thode, Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sens. Environ.* **109**, 66–80 (2007).

K. Hoover, Wildfire Statistics Congressional Research Service (2018).

C. Stone, A. T. Hudak, P. Morgan, Forest Harvest Can Increase Subsequent Forest Fire Severity (2004) (June 10, 2021).

H. S. J. Zald, C. J. Dunn, Severe fire weather and intensive forest management increase fire severity in a multi-ownership landscape. *Ecol. Appl.* **28**, 1068–1080 (2018).

F. Galton, Regression towards mediocrity in hereditary stature. *j. Anthropol. Inst. G. B. Irel.* **15**, 246 (1886).

J. Fürnkranz, Decision Tree (2010).

J. Eidenshink, *et al.*, A Project for Monitoring Trends in Burn Severity. *Fire Ecology* **3**, 3–21 (2007).

L. Breiman, Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

J. T. Abatzoglou, A. P. Williams, Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11770–11775 (2016).

P. Jain, *et al.*, A review of machine learning applications in wildfire science and management. *Environmental Reviews* (2020) https:/doi.org/10.1139/er-2020-0019 (June 10, 2021).