Wildfire Prediction Modeling

Jason Chuang

***Abstract*— This paper will explore the potential of machine learning AI model in the prediction of impending wildfires as well as wildfire severity. Using multiple data sources to build the neural network to predict the wildfire severity and sources.**

I. INTRODUCTION

Every year, wildfires ravage California and other dry states. Especially in the summer when the climate is dry and the weather is hot, wildfires become even more rampant. The wild fires in the U.S. in 2018 alone have burned over 51,000 wildfires have spread across California alone destroying over 8.5 million acres of land forcing residents to and animals to relocate and destroying up to. Wildfires are a necessary part of nature to recycle the land, but large fires spreading into residential areas and destroy residential life are dangerous. A method to predict where these wildfires sources as well as the size and severity of the wildfires is necessary. A neural network using data inputs from a variety of sources can be created and utilized for this exact task: predicting and analyzing wildfires. Large amounts of data on previous forest fires such as the range and source of previous forest fires are publicly available. Using this large trove of widely available data, training and data acquisition, a common issue of training neural networks, can be made easy. Through this neural network, wildfire can be preemptively prevented and residents can evacuate from potential wildfire quicker. Not all wildfires can be prevented, but finding the source potential largest wildfires and preventing it would help prevent lives lost and prevent the damage to homes and lives.

II. PROBLEM

A. Wildfires Sources

First, the model built will be created to predict potential wildfire sources. Potential wildfire sources can be described to be as dry arid areas with lots of dead trees. Areas with plentiful dead trees would be more likely to burn. In California, where wildfires are plentiful, there are up to 129 million dead trees which contribute to the plentiful raging wildfires each year. This is partially due to the many droughts in California making the land dry and dead trees plentiful. These dead tree areas are ideal starting places for wildfires. By proactively predicting wildfire sources, preemptive action can be taken to prevent these wildfires. There are many reasons for the source of a wildfire, some natural, some man made. Common wildfire sources are lightning, burning campfires, cigarettes, hot winds, and even the sun. Although it maybe impossible to account for random starts to wildfires, likely wildfire source areas can be found due to the amount of dead dry plants that can easily serve as fuel for these ignition methods. By searching for these fire prone-areas as well as fire prone methods a statistical probability can be found for wildfires starting in a particular location on a particular day. Using this knowledge wildfire sources can be preemptively stopped on days and areas with especially high chance of wildfire starting.

B. Wildfire Spread

Next, the model will be able to predict the spread and range of these wildfires. The spread and range of these wildfires destroy wildlife, force residents to relocate, pollute the air. Wildfire spread is dictated by flammable materials present, weather conditions, and moisture content. With hot dry weather, the wildfire sources are not only more likely to start, but also spread. Fuel for wildfires that contribute to the rapid growth of wildfires are trees, grasses, brush, and even homes. Another popular fuel for wildfires are dead trees, which are extremely common in dry California. Tree mortality in California is publicly available. Using this information we can draw patterns and

conclusions to how a wildfire is likely to spread and to what range. By predicting the spread and range of the wildfires, we can see which fires are potentially the most dangerous in terms of size and heat and can selectively choose which wildfires to impede before damaging large areas polluting air. The spread of wildfire pollutes the air making areas uninhabitable and forcing evacuation from areas. As can be seen from recent wildfires in California, wildfires are a serious threat to the health and safety of

III. Product/Service Overview

The product will utilize machine learning to predict the source and spread of the wildfires. Utilizing public data available through NASA, EPA, and GIS, we have vast amounts of data ready for processing and analyzing to create our model. Through GIS we are able to grab history of of fire detection in the U.S. in the past 17 years in xml format. Using this information we are able train our model to read inputs from history and from this history we are able to see common origins of forest fires as well as the range of these forest fires due to weather conditions and humidity. By using the weather conditions at the time of wildfires in the pat we are able to see when the weather is more likely to lead to the start of a fire. Using machine learning we are able to find patterns between historical data and historical weather conditions and then map them to current weather conditions.

Using a convolutional neural network we are able to train it to detect dry fuel-rich regions looking at satellite images publicly available from NASA. By looking at the specific regions we are able to classify an image using an input layer and multiple hidden layers to give an output layer describing the dryness of the region, which is directly associated to how likely a fire is to start in the region. In the hidden layers we preform convolution and a variety of operations to detect patterns and alter data. Using the output of the convolutional neural network and historical data we then train another model to learn from these in order to predict likeliness of current fires in areas from the images coupled with the historical inputs mentioned above as well as current predicted weather for the next recent months. Using this we are able to find the likely locations for the source of the fire.

Once a prediction is made we can then again apply another neural network using historical data and upcoming weather data to predict the size and range of this potential wildfire. We can utilize the same convolutional neural network mentioned above to look at the surrounding areas of land to detect and look at the amount of dead trees and potential fuel the wildfire can use to spread. Using this information, we can display an output onto a map of the desired region and output onto the graph in colors of areas that are potentially dangerous source of fires. These areas would be denoted by colors of red, orange green. With these color outputs onto the map denoting level of danger as the sources of wildfire. Once the user inputs a data point, the potential spread and range of the point will be given based on factors of surrounding potential fuel, meteorological conditions. The potential spread varies greatly with meteorological data so a potential range can be easily found with meterological data and historical data.

An example of the product can be made in the terms of California, an area rampant with wildfires during certain seasons. Similar to other common artificial intelligence applications today, our application requires large data sets. In order to properly train our data set to a reliable success rate we would require large amounts of data. First we would process all of the information regarding current meteorological data, including weather, wind, temperature, climate, and humidity in areas around California, historical data of fires, and images of tree mortality rate in the previous historical years. Once we process the data into a data format that is able to be easily worked with. Using a convolutional neural network, we are able to train a model for image processing satellite images to accurately detect particularly dry, or forest areas that are more likely to contribute to the spread and source of a wildfire. Using all of these inputs we then create a neural network that will take in these inputs and then using different weights and calculations in the hidden layer to create an output layer to display our results. Our

result would be an outline of the predicted path of the wildfire displayed through a map.

Compared to similar products/services, this product provides a better service with more information than any other available product. In looking at the competitors, it can be found that the competitors are rather simple in their results and are lacking. Similar services provide similar methods, none are holistic and inclusive enough to accurately predict both source and range of the wildfires. I will discuss these comparisons in greater detail in the next section.

IV. Competition

With the prevalence and concern of wildfires, many products have been developed with similar idea: using machine learning to predict wildfire outbreak. One example can be seen in an AI-based computer program being created by University of Alberta wildland fire expert Mike Flannigan and Ryan Lagerqueist who will utilize weather data and pressure fields to predict fire-friendly weather. While an excellent solution to predict wildfire weather, it is severely lacking in regards to more holistic sources of data, such as historical data. This product also does not use the modeling to predict spread and range of the wildfire.

Another similar product can be seen in WIFIRE, an integrated system for fire analysis using satellite imagery and real-time data from cameras and sensors to assemble picture of the fire, properties of the fire, and the fires direction. Although this is extremely similar to the proposed product, WIFIRE only predicts the possible direction and range of the wildfire, not more likely sources and areas of the fire to begin.

Next, a product on using Autonomous Aircraft for Distributed Wildfire Surveillance is also a potential competitor. Using autonomous drones, they are able to use deep learning to accurately track wildfire expansion. While this product may out-preform my own product by using sheer real-time accurate results in its machine learning and surveillance of the wildfire, it is lacking in terms of features when compared to my product. The competitor does not track potential wildfire start sources and therefore is lacking.

Although many products do similar jobs, none are as holistic in predicting source and range of wildfires and utilizing as much data from varied sources as my product. With a more holistic approach and helping predict wildfire range and source, wildfires can be more preemptively stopped.

V. Development Plan

Our plan for development is to have a team of around 12 software engineers working in an Agile methodology. With several sprints, a rough workable prototype can be developed and marketed to potential investors to find additional funding for the application. A rough prototype can be developed in the period of 4 weeks. In this marketing to potential investors, a rough schedule with milestones can be made. From investors we seek a sizeable amount of funding for payment of employees as well as resources such as server hosting. After reliable funding can be found a work schedule can be developed with the milestones in consideration. Working in several sprints in Agile methodology, we seek to further develop the application, we can develop a working application and host the product on Amazon Web Services, Google Cloud, or Microsoft Azure. By utilizing external platforms, we are able to save expenses, prevent extra work, and have external help to configure the server to suit our needs. After doing so, we will undergo severe debugging and testing. Our testing process will be extremely vigorous involving black-box testing as well as white-box testing. We may even release an alpha version and distribute a trial version to potential clients. This will serve to not only test and debug our product but advertise our product to potential clients. Through these many steps of debugging and testing we hope to iron out the bugs of the product and give clients a bug free product. Allowing a trial version to clients allows us to gain insights into improvements of the product as well. In this process the product will continue in the Agile process and will constantly change to suit the needs of the clients and suggestions of the staff and client. Through this intensive testing and training the hope is that the accuracy and reliability of these predictions are high so that our product are easily marketable and can be dependable.

VI. Distribution

Distribution of the product can produce several complications. This product of wildfire detection can be used by governments federal and local around the world. The California wildfires alone have destroyed 2.8 billion dollars worth of damage. In particular the governments that have particularly dangerous and damaging wildfires such as California and China, this product can be proposed to help decrease the damage caused by the wildfire. Our product seeks to market itself to government agencies to help in wildfire prevention and containment, potentially saving billions in damages. We can do this by presenting accuracy prediction of the machine learning product. With a high accuracy prediction of wildfires, our product can be shown to be reliable and viable solution to helping prevent and stifle wildfires. Using this idea, we will market to governments around the world with ample data collection, and advertise our product to potentially saving them billions in fire prevention. A prime candidate this product can be sold to is the U.S.Forest Service, whose fire suppression budget is struggling with the recent rampant wildfires. The fire damage in recent years has caused a huge decrease in available funds due to the resources spent of fire suppression. With a lack of funding, defending the fire prone-California with a lack of funding is extremely difficult. With our product they would be able to not only predict where wildfires will begin, but they will be able to predict the range of and spread of the wildfire. By doing so they may preemptively stop the fires and save millions on wildfire damages throughout the state. Mentioned in the development phase, our product will also given away to several clients to use as a trial to not only test our product but distribute our product in an effective manner so that our more likely potential clients are able to see our product and then wish to purchase it once they see the reliability and effectiveness of our product. In this way we essentially get free exposure and advertising to our clients and distribute our product to potential clients. In this way, our product is distributed to the right audience and therefore more likely to be purchased.

VII. References

References

[1] Albert, Victoria. The Government Is Running Out of Money to Fight California's Wildfires. The Daily Beast, The Daily Beast Company, 9 Aug. 2018, www.thedailybeast.com/the-government-is-running-out-of-money-to-fight-californias-wildfires.

[2] Climate Change: Global Sea Level — NOAA Climate.gov, 16 Nov. 2018, www.climate.gov/news-features/event-tracker/hot-dry-summer-and-slow-start-wet-season-primed-california-november-2018.

[3] CS231n Convolutional Neural Networks for Visual Recognition, cs231n.github.io/convolutional-networks/.xxx

[4] Barron, Laignee, and Mahita Gajanan. Why California's Fires Are Bigger, Deadlier, and More Costly. Time, Time, 17 Oct. 2017, time.com/4985252/california-wildfires-fires-climate-change/.

[5] Corriveau, Kevin. California Wildfires Damage in 2018 Worth over 2.6bn. GCC News — Al Jazeera, Al Jazeera, 7 Sept. 2018, www.aljazeera.com/news/2018/09/wildfires-expected-worsen-climate-change-180907091608388.html.

[6] Facts + Statistics: Wildfires. III, www.iii.org/fact-statistic/facts-statistics-wildfires.

[7] Gaskell, Adi. Using AI To Predict Forest Fires - DZone AI. Dzone.com, 13 July 2018, dzone.com/articles/using-ai-to-predict-forest-fires.

[8] How Big Data Is Helping Predict and Fight Wildfires. Digital Transformation, 24 Oct. 2018, www.delltechnologies.com/en-us/perspectives/how-big-data-is-helping-predict-and-fight-wildfires/.

[9] Julian, et al. Distributed Wildfire Surveillance with Autonomous Aircraft Using Deep Reinforcement Learning. [Astro-Ph/0005112] A Determination of the Hubble Constant from Cepheid Distances and a Model of the Local Peculiar Velocity Field, American Physical Society, 9 Oct. 2018, arxiv.org/abs/1810.04244.

[10] New AI Program Fights Fire with Data. Phys.org - News and Articles on Science and Technology, Phys.org, phys.org/news/2018-06-ai.html.

[11] Wildland Fire. USDA, www.usda.gov/topics/disaster/wildland-fire..

[12] Wildfires. Citizen Corps — Ready.gov, www.ready.gov/wildfires.

[13] Williams, Sky B.T. Wildfire Destruction - A Random Forest Classification of Forest Fires. Towards Data Science, Towards Data Science, 24 Feb. 2018, towardsdatascience.com/wildfire-destruction-a-random-forest-classification-of-forest-fires-e08070230276.