International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Fire Detection Using Support Vector Machines (SVM)

Vigneshwaran SR[1], S. S. Shanthakumari[2], Vinodhini Ranganathan[3]

[1, 3]Student, Indian Institute of Management, Bangalore, Karnataka – 560076, India

[2]Director, School of Management, Karpagam University, Coimbatore, TamilNadu – 641032, India

Abstract: ***Earth is an intrinsically flammable planet owing to its cover of carbon-rich vegetation, seasonally dry climates, widespread lightning and volcano ignitions. Therefore, something is always burning on this earth. Wildfires can be natural or man-made (accidental or deliberate).These fires release greenhouse gases and huge volumes of smoke thereby polluting the air and the environment which in turn cause severe degradation to the ecosystems. Based on observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra and Aqua satellites, several algorithms have been designed and many approaches have been proposed to detect wildfire. While the question of accuracy over the existing methods of fire detection remains, this paper proposes a more robust and accurate approach: using SVM (Support Vector Machines) for detecting wildfire. And also will compare SVM with logistic regression which is the most common used algorithm for solving industry problems.***

Keywords: SVM; wildfire; MODIS

1. Introduction

Wildfire is a large sweeping and destructive conflagration especially in a wilderness or a rural area where there is abundant vegetation. The primary natural causes of wildfire, as identified by NASA, are lightning, spontaneous combustion, heat waves caused by climate change and volcanic eruptions. The most common direct human causes of wildfire ignition include deliberate measures such as arson, animal husbandry, and land-conversion burning. Accidental causes include cast-away cigarettes, power-line arcs, campfires, fireworks and sparks from equipment. Wildfires are difficult to control, since, by the time they are detected, a large area of vegetation is already lost. Some fires burn for days and weeks spreading to the nearby living areas, causing severe damage to the ecosystem and to the people.

1.1 Why is it Important

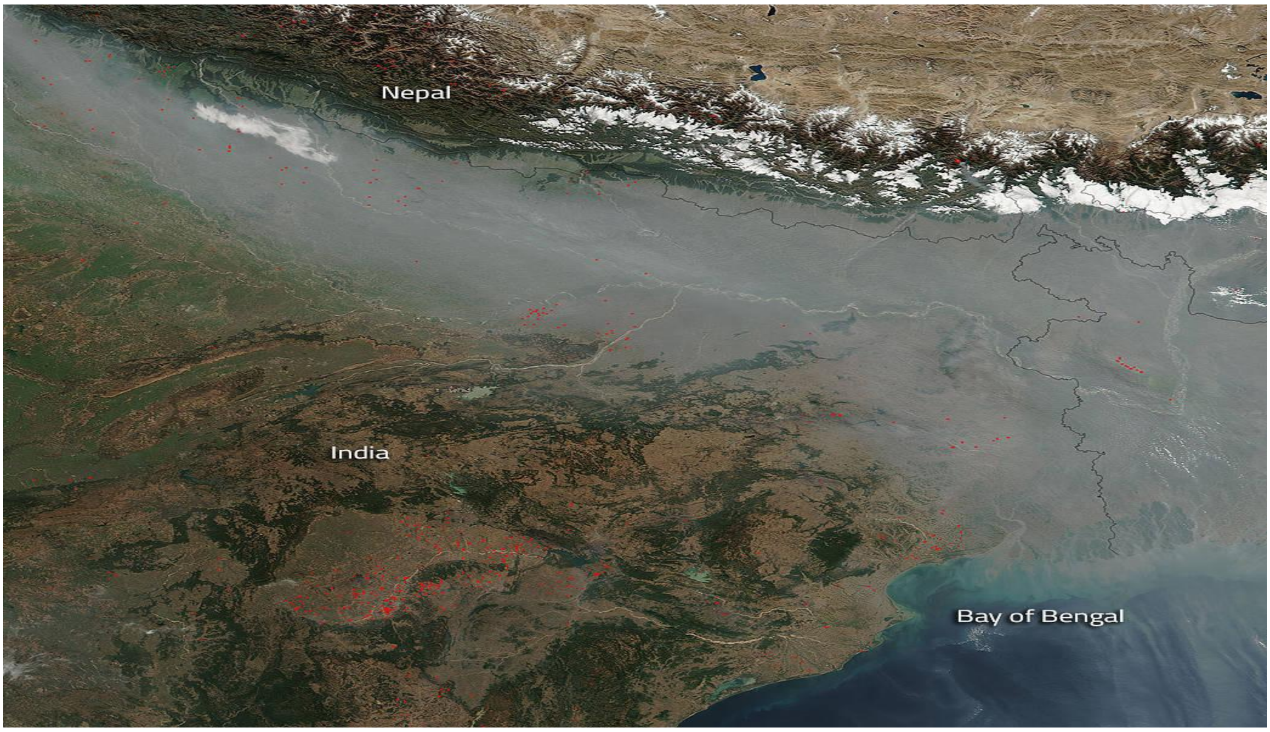
Wildfire is becoming a major concern since rising temperatures are transforming many landscapes. We're likely to see more wildfires in more places than just the forest, in the future.

On May 1, 2016, a wildfire began in Fort McMurray, Alberta, Canada. On May 3, it swept through the community and spread across approximately 590,000 hectares before it was declared to be under control on July 5, 2016. It continues to smolder till date, and may not be fully extinguished until the spring of 2017.

Similarly, On May 2, 2016, forest fires destroyed 3,500 hectares (8,600 acres) of forest land in Uttarakhand, India. These fires started in the pine forests and produced clouds of smoke for many days after the incident. Soaring temperatures reignited the forest fires on 18 May 2016 destroying nearly 180 hectares of green land. Parliamentary Standing Committee on Science and Technology in India reported that the number of incidents of forest fire has increased by 55%.

According to the U.S. Department of Agriculture, the wildfire season in the US is, on an average, 78 days longer than it was in 1970. The US had an early start to the wildfire season in 2016 and the total number of wildfires rose to 29000 and approximately 2.6 million acres of land was destroyed.Such heavy wildfire was due to the fact that the US faced the driest season in 14 years. The major causes of these wildfires in the US include drought, hot weather, and accumulated Chir Pine (inflammable due to their high-resin content which then becomes a source of fuel for these wildfires). El Niño was also a major contributing factor for the climate change.

A paper published by Alicia M. Kinoshita et al in the Anthropocene Journal (2016), explains in detail about wildfire and its adverse effects on the environment.



1.2 How does a wildfire

There are lots of ways due to which wildfires get started. But take a look below at the most common ways in which wildfires are started.

a) Campfires:
Camping is common thing that occurs almost in all hill stations. People of all generations spend time in the woods to enjoy at their outdoors. So camping needs fires for various activities and these fires can start wildfires if not put on and off properly.

b) Smoking:
Smoking is one of the most common things that we can in see day to day life. Even though government started imposing more tax to reduce the consumption of cigarettes, there is yet no sign in reduction of smoking. However when people smoke, sometimes the buds are not properly extinguished before its thrown away. And this not properly extinguished bud will end up in fire.

c) Lighting :
Lighting causes most of the wildfires. It's not easy to believe but an investigator has evidence for it. When lighting strikes it produces a spark. For an example when an lightning strikes on rocks, power cables on street and tress fire just set them off.

d) Burning debris:
Most of place they burn Junk and yard waste. And it's the most common practice way for disposing waste. Because we don't have proper system for disposing waste. And at times they may get out of hand and start fire.

e) Accidents or equipment failure:
The industries equipment's with non-proper maintenance and sometimes vehicle accidents on roads and highways will set fire when they go wrong. This might create major problem if they are not stopped at earlier stages. And this is whyit's mandatory for setting up safety measures at industries and giving proper knowledge to everyone on fire prevention.

f) FireWork
Many places it's been banned from making fireworks and in some places it's allowed with proper safety measures. Fireworks are banned because of explosive nature and chance of higher potential to start fire. If fireworks are not made up in proper places they might end up in fires. And firework is used in many occasions for celebrations and widely used across world.

g) Arson
Arsonist is a person who set fire with an intention of causing damage to piece of land or his property. And 30% of wildfires are accounted by arsonists. For example people set fire to hisland to claim insurance.

1.3 Factors that make wildfire burn more

There are ingredients which are capable to burn and there are few factors exist. When you combine these factors and ingredients it will help to burn more quickly. Let's looks some of them.

a) Wind
Wind provides fresh supply of oxygen which is a key ingredient of fire. Wind might lead to change, direction of fire to new areas and also wind is a major cause of spread of wildfire.

b) Slope
Uphill or downhill, wildfires usually move faster. Wildfires burn or spread faster when they have slope which is steeper.

This is because wind action will be more at uphill and also steeper slopes has lot of fuels in close proximity.

c) Temperature

When the temperature is higher it conducts more fire because they absorb more moisture from fuels. And this is the reason the areas with higher temperature tend to set fire.

d)Humidity

It's known that fuels with locations where there is high humidity and rainfall tend to be moist and damp. The amount of water vapor in air is known as humidity. The higher moisture in the fuels, tend to have higher water vapor in air and thus they are less likely to catch fire.

e) Times and Seasons

There are places where summer extents and it registers to lot of fires. It's because the heat during summer which makes fuels drier and that enables to provide richer oxygen than in winter seasons. For example the Harmattan winds from sahara desert on dry season in west Africa is the reason for more fire.

f) Fuels

Fuel composition is also responsible for ease of wildfire spread. Vegetation and Trees with more moisture tend to slow down fire whereas dry vegetation such as small trees, dead leaves and grasses tend to fasten fire. On the other end there are vegetation composite of high oils and resins makes fire burn more with ease.

g)Space between fuels

Fuels with close proximity tend to burn more and spread faster. If fuels are patchy or sparsely distributed the fire tend to slow down. And one of available common method for avoiding fire that's been widely in practice is to create a ring of space around.

1.4 Effects of wildfires

There are many effects of wildfires but most common are Economic cost, Soils and organic matter and Watershed.

a) Economic cost

In TV you might see firefighters battling a wildfire. That might give you a rough idea on how important is the wildfire because firefighters risk their own life and also damage it caused to wildlife and vegetation. Not only damaging vegetation it might also destroy houses and anything in its way. On the other end the government spends millions of dollars to logistics, aircraft, and chemicals to avoid fires. So wildfire leads to huge economic loss.

b) **Soils and organic matter**

Forest soils are very rich in nutrients, decaying debris and also they are composed of features which support organic activities and myriad of life forms. These wildfires might raise temperatures of soils to over 900°C and this might clears or destroy all organic value of the soil.

c) **Watershed**

One of the major effects caused by wildfire is watershed. The natural layering of the soils will be affected because of burned organic matter in the soil. And this might negatively affect percolation, making the surface of soil water repellent and infiltration. And water on the surface therefore unable to drain and this might causes erosion.

Many researchers say that not all forest fires are bad. There are few argue that there are benefits too. They also say that birds and animals either they suffer or sometimes they die. But most of them have routes to escape from wildfires.

After fires many plants can quickly grow back. Some plants have their seeds and they are exposed to ash soils. A Serotinus cone from jackpine tree is an example; these trees opened up their seeds and get exposed to ash enriched soils. Yellow birch and white pine species are also among the species which are benefited from forest fires.

1.5 Fighting Wildfires

Not all the fires are fought in the same manner. The big idea behind is same but they are fought differently. The common idea behind is to deprive the fire of its fuel and let it go out by itself. And this can achieved my many ways.

a) Firelines or Firebreaks:

Create a ring around the fire are and get rid of all fuel in the fire path using bull dozers and land equipment. Once the fire nears ring, it can no longer spreads.

b) Firing out:

Fire fighters burn out areas or fields before the fire actually gets there. They look for boundary or natural edge such as plain field, plain, road and they do a controlled burn of all the fuel between the fire and the barrier

c) Aircrafts

Every government has special aircrafts called air tankers which fly over fire and dump water. There are air tankers uses fire retardant which are called jellu-o, and chemicals such as ammonium phosphate. At times buckets carry 100-2000 gallons of water and sprinkle over fires. Some advanced countries do developed pilot less air tankers which are automated and it can be controlled with the help of computers.

d) Technology

These days the technologies are getting more advanced. With the help of satellites, computers, digital equipment's they can monitor fires. Or even they can forecast wind directions and take better decisions to avoid or control fires. And there are ways by which fire fighters can optimize plan to control fires.

e) Firemen

irefighters who are well trained are key to control wildfires. Even though there are technologies exist, there is need of firemen who will get in to action. They wear oxygen mask and fire proof clothing before getting in to action. Nomex is one fire proof material used for clothing. With special tents they protect themselves from extreme heat in case they get trapped.

2. Review of Literature

Several algorithms have been developed since the early 90's, for detecting wildfires.

Boyd M. Harnden et al (1973) designed a simulation model, which could be easily adapted for analysis of any forest fire detection. While F. Andreucc et al (1993) developed a mathematical model to calculate the properties of smoke plume produced by a forest fire and used it to detect fire systems automatically, based on infra-red passive sensors and active optical sensors (lidar).

In the last decade, most of the research on fire detection was done using sensors and images (Image processing). BiswajeetPradhan et al (2007) using remote sensing and geographical information systems (GIS) predicted forest fire using Frequency ratio model. JieHou et al (2010) proposed a new algorithm (NAUTEA), which uses target extraction and recognition for fire detection. MajaStula et al (2012) published a paper titled iForestFire, an automatic fire detection algorithm which uses various complex image processing algorithms. Interestingly S.Movaghati et al (2012) used MODIS data to detect forest fire using an agent-based algorithm. But his paper was restricted only to Iran.

In last two years lot of study happening on fire detection. Like Yang Jia et al (2015) used adaptive flame segmentation and recognition algorithm, to detect fire detection in spacious buildings. In 2016 C. Emmy Prema et al developed an image processing approach for detecting smoke in videos. And Gong Cheng et al (2016) did a survey on object detection in optical remote sensing images, his paper provides a review of the recent progress in this field.

In this paper, a popular machine learning algorithm, Support Vector Machine (SVM) is used. Proposed by Vapnik in 1998, SVM is a supervised learning model used for solving classification problems. It is widely used in various object detection applications, such as Ship detection (Bi et al., 2012; Zhu et al., 2010), Airport detection (Tao et al., 2011), Airplane detection (Sun et al., 2012; Zhang et al., 2014b, 2015b), Road extraction (Das et al., 2011; Huang and Zhang., 2009, Song and Civco.,2004), change detection (Bovolo et al.,2008; De Morsier et al., 2013), weakly supervised object detection (Han et al.,2015; Zhang et al., 2015a).

Using SVM as an approach for fire detection has not been explored so far. This paper aims to provide an accurate and reliable SVM model for fire detection.

3. Methodology

The data used for analysis comes from Earth Observing System Data and Information System (EOSDIS), a wing of NASA's Earth Observing System (EOS). EOSDIS offers earth system science data products and associated services. The data from EOSDIS is accessible via FTP and HTTP.

The fire maps show the locations of actively burning fires around the world on a monthly basis, based on observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite. For this paper, we have used data from MODIS for the months of April 2014 and April 2015. Data collected in April 2014 is used for the training model while April 2015 is used for validating the model.

In this paper we will be comparing the Logistic regression and Support Vector machines and also will try to show why SVM is better.

3.1 Logistic Regression Vs SVM: Part I

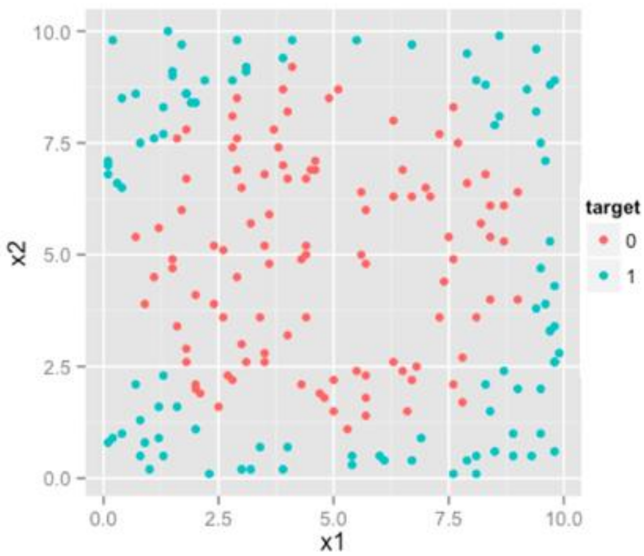
Classifications are one of the major problems that many researchers face while working on standard business problems across industries. In this paper will be comparing major two techniques out of many, Logistic Regression and Support Vector Machines [SVM].

Of all the algorithms used for classification. Time and again I have seen analyst asking which to choose for their particular problem. Classical and the most correct but least satisfying response to that question is "it depends!". So here will explain some light on it depends on What ? .

It's a very simplified 2-D explanation and responsibility of extrapolating this understanding to higher dimensional data, painfully lies in the reader's hand.

Let's start with the most important question: what are we exactly trying to do in classification? Are we are trying to classify. If yes, In order to classify, we need to get a decision boundary or a curve, which separates two classes in our feature space.

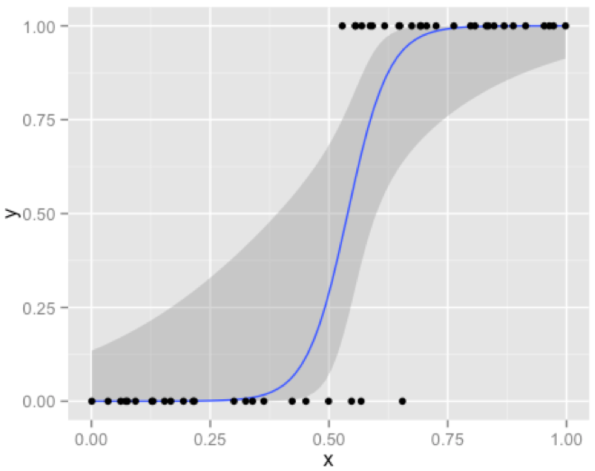
Feature space would be confusing to many researchers who haven't encountered it before but sounds like a very fancy word. To clarify this, let me show you an example. Let's have a sample data with 3 variables; x1, x2 and target. Target is a binary variable which takes two values 0 and 1, depending on values taken by predictor variables x1 and x2. The plot if this data is shown below.



Feature space is right here. In our example, we have two predictors/features; hence feature space is shown in 2D. You can see the target two classes shown using different colors. All we wanted from algorithm is to give us curve/line which can separate two classes.

We can see that an ideal decision boundary (or separating curve) would be circular. The difference between algorithms lies in the shape formed by decision boundary.

Here we start with logistic regression. Researchers still not clear about the shape of decision boundary given by a logistic regression. This clarity lacks because looking at S shaped curve many times in the context of logistic regression.



In the above figure the blue curve shown is not decision boundary. It's a way where transformed response from binary response by which logistic regression is modelled. Logistic regression decision boundary is always a line/ hyper plane. Showing logistic regression equation is the alternative way you will get convinced by this approach.

$$\log\left(\frac{p}{1+p}\right) = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 \ldots.$$

Let's for simplification we assume, F is a linear

combination of all the predictors .

$$F = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 \ldots.$$

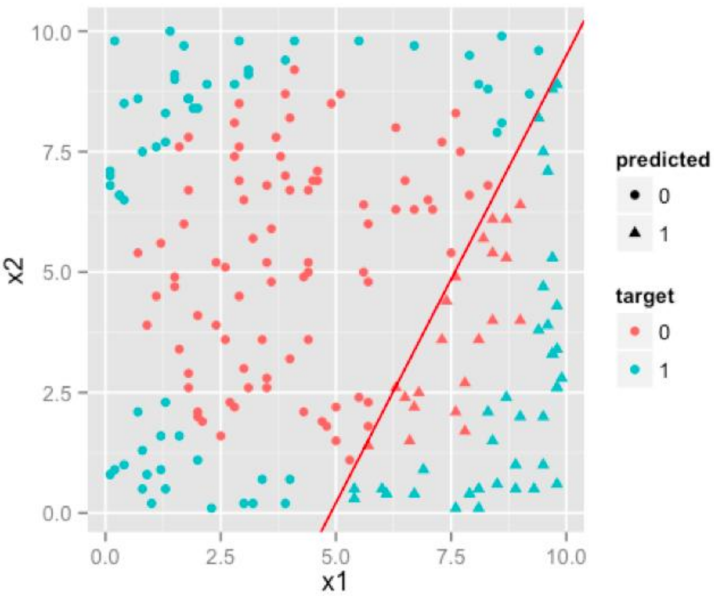
The above equation can also be written as:

$$p = \frac{1}{1+e^{-F}}$$

For predicting in logistic regression, you should decide cutoff probabilities, above which your prediction will be 0 or 1 otherwise. Let's assume c is cutoff value. Then decision

process will be like this:

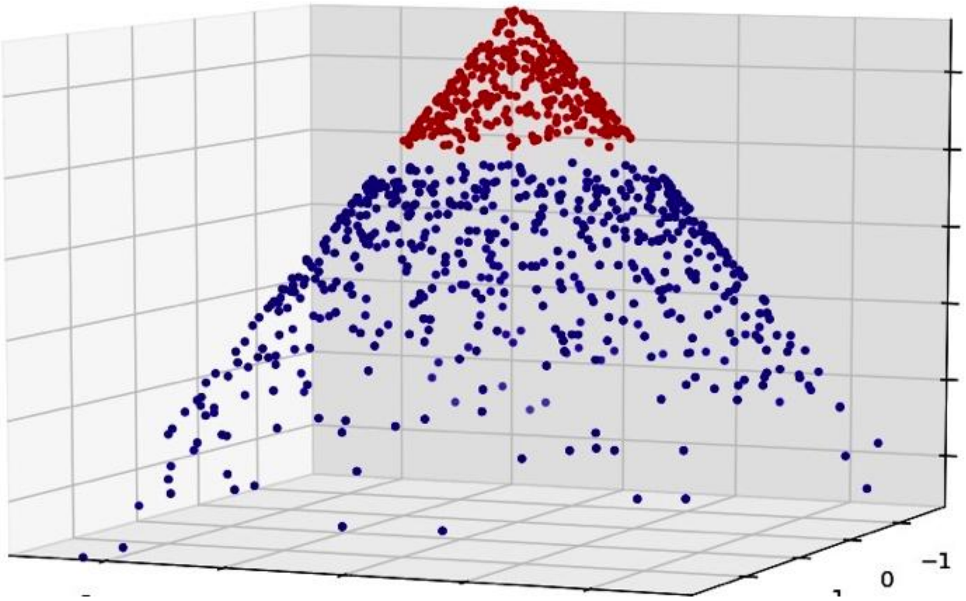
Y=1 if c<p, otherwise 0, Which defines decision boundary F>constant.
F>constant, here is a linear decision boundary. Then logistic regression for our data will be like image below.



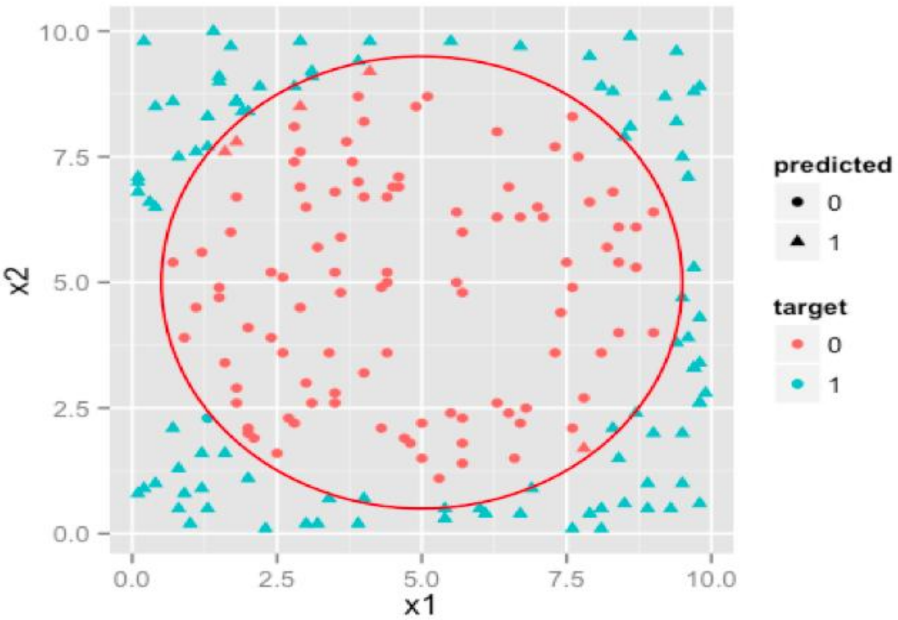
The outcome shows, decision boundary will always be linear produced by logistic regression. So it cannot emulate a decision boundary which is circular. So, logistic regression will work for classification problems where classes can be linearly separable.

3.1.1 Support Vector Machines

SVM functions by projecting feature space into kernel space and making classes linearly separable. Or to explain still simpler, SVM adds additional dimension to feature space in a way that makes classes linear separable. And this planar boundary when projected to original feature space emulate decision boundary which is non linear. The image below explains might give you better explanation.



From above image you can see that with third dimension is added to data, we can separate two classes with a linear separator, and when projected back to original 2D feature space, it becomes circular boundary. The below image shows, how SVM performs for our sample data.



Now the difference might be clear. In SVM the decision boundary can be circular. But still the basic researcher question remains unanswered. That is when and which algorithm should be used when dealing with multi-dimensional data. And it will be explained in following section.

3.2 Logistic Regression Vs SVM: Part II

In this part will brief how to choose between Support Vector Machines and Logistic Regression. As we already discussed most popular answer will be it depends. Let's will shed some light on it depends on what. Almost every technique will have properties which is inherent by their design. Let's explain some of them to provide few insights on selection.

Let's start with Logistic Regression, the most common used algorithm for solving industry problems. But stilllogistic regression is losing to other techniques with efficiency and ease of implementation to practice.

Logistic Regression is very convenient and useful in a way that it doesn't give outright classes and discrete output as output. Instead we get probabilities for each observation. You can apply many custom performance metrics and standard to this probability score,use a cut off and in turn which classify output. Finance industry scorecard is a very popular application, where cutoff value is adjusted to get different result for classification from the same model developed. There are few algorithms provide direct result. Instead other algorithms provide discrete direct classifications as output. Also with logistic regression it's pretty efficient in terms of time and memory requirement. And also logistic regression has algorithm to handle large data and can be applied on distributed data.

Further logistic regression algorithm is robust and not much affected by multicollinearity for mild cases. By implementing logistic regression with L2 regularization severe cases of multi collinearity can be handled. L2 regularization is not the best choice if parsimonious model is needed, because it keeps all features in model.

Logistic regression starts to fails when you have large good chunk of missing data and having large numbers of features. Having many categorical variables will also be a problem for logistic regression. Another issue with logistic regression is that it uses all data for coming up with scores. Although this is not a major issue, but it can argued that most cases which lie at extreme end of scores should not really be concern when it comes to developing a separation curve. It should depend on boundary cases, some might argue. Also you have to depend on transformations when some of features are nonlinear, which might become a hassle as size of feature space increases. And to summarize logistic regression, we have given few important pros and cons below

Logistic Regression Pros
- Each observation will have probability score
- Implementation will be efficient across tools
- With L2 regularization is countered to an extent, multi collinearity is not an issue

Logistic Regression Cons
- When feature space is too large doesn't perform well
- Large number of categorical features/variable can't be handled
- Relies on entire data and transformations for non-linear features

3.2.1 Support Vector Machines

Here comes Support Vector Machines. The most important thing about support vector machines is that they rely on boundary cases for building separating curve. They can also handle decision boundaries which is nonlinear. Reliance on boundary cases helps them to handle missing data. Large feature space can be handled with support vector machines and which makes SVM most used algorithms in text analysis. Whereas logistic regression is not a choice because it fails to handle large features. Result of SVM is not easier as logistic regression for layman. SVM is very costly to train huge data because of nonlinear kernel. To summarize pros and cons of SVM

SVM Pros:
- Large feature space can be handled
- Non-linear feature interactions can be handled
- Does not rely on entire data

SVM Cons:
- With large number of observations it's not efficient
- It not easy to find appropriate kernel

3.3 Other algorithms

Letassume Y Variable could be numeric or qualitative, and also in the same way let's assumeX variable can also be numeric, qualitative or combination of them. Based on the variables nature, we get combination of various scenarios, which indeed require specific type of Analytics technique. Image below shows the different technique which can be used for different combination of variables.

		X or Predictor Variables			Pattern Finding (No Y Variable)
		Continuous	Categorical	Mix	
Y or Response Variable	Continuous	• Linear Regression • Regression Tree • K Nearest Neighbor (KNN) • Neural Networks	• Linear Regression • Regression Tree • Neural Networks	• Linear Regression • Regression Tree • Neural Networks	Clustering • Hierarchical • Non-hierarchical Sequence Analysis Association Rules
	Categorical	• Logistic Regression • Decision Tree • K Nearest Neighbor (KNN) • Neural Networks • Discriminant Analysis	• Naive Bayes • Logistic Regression • Decision Tree • Neural Networks	• Logistic Regression • Decision Tree • Neural Networks	

As you can see there are various techniques like Logistic Regression, Decision Trees, K Nearest Neighbor (KNN), Neural Networks and Discriminant Analysis which can be used for classifications. But in this paper will be using only Logistic Regression and Support Vector Machines.

As already discussed Support Vector Machine (SVM) is a machine learning algorithm majorly used for classification problems or regression challenges. In this paper, we developed a model using R programming package 'e1071', and tuning the parameters radial to kernel, and type to C-Classification.

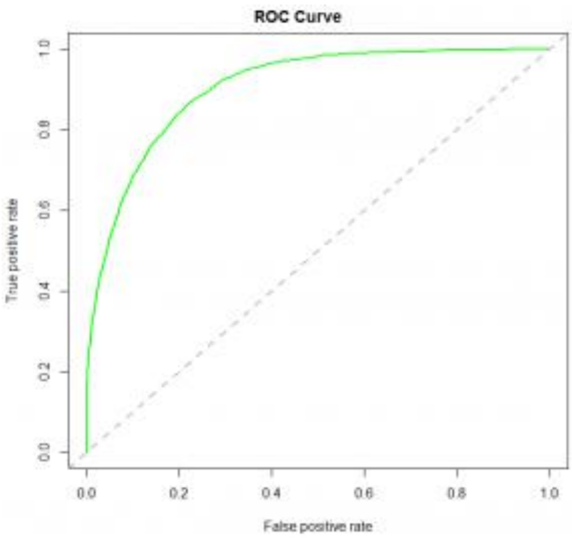
And similarly for logistic regression we have a function generalized linear models glm() in R and its popularly used for practice and it gives AIC, confusion matrix and Roc as output by default to evaluate the performance.

AIC (Akaike Information Criteria) is a metric of adjustedR^2. This metric is the measure of fit which penalizes model for the having number of coefficients. So Analyst always looks for minimum AIC value.

Null Deviance and Residual Deviance is an intercept which indicates the response predicted by model. Analysts look for the lower value. Lower the value, the better the model. Whereas residual deviance shows the response predicted by a model. Lower the value, better the model.

Confusion Matrixis a table shows actual vs predicted values. It's one of the easiest ways to find accuracy and also it helps to avoid over fitting.

Roc Curve (Receiver Operating Curve) will summarize the performance of model by evaluating tradeoff between sensitivity and specificity. Always consider p>0.5 while plotting ROC because we are concerned about success rate. The area under curve (AUC), mostly referred as index of accuracy (A) or concordance index is a ROC curve perfect performance metric. The model prediction will be better if higher the area under the curve. Image below shows ROC curve. The curve touches top left corner when the model accurately predict all observations.



And in this paper will use confusion matrix to compare model we developed by logistic regression and Support vector machines.

4. Analysis & Results

After Data cleaning, SVM model was developed using the training data. The data from MODIS gives us the following information.

Attribute	Long Description
Latitude	Center of 1km fire pixel but not necessarily the actual location of the fire as one or more fires can be detected within the 1km pixel.
Longitude	Center of 1km fire pixel but not necessarily the actual location of the fire as one or more fires can be detected within the 1km pixel.
Brightness	Channel 21/22 brightness temperature of the fire pixel measured in Kelvin.
Scan	The algorithm produces 1km fire pixels but MODIS pixels get bigger toward the edge of scan. Scan and track reflect actual pixel size.
Track	The algorithm produces 1km fire pixels but MODIS pixels get bigger toward the edge of scan. Scan and track reflect actual pixel size.
Acq_Date	Date of MODIS acquisition.
Acq_Time	Time of acquisition/overpass of the satellite (in UTC).
Satellite	A = Aqua and T = Terra.
Confidence	It is intended to help users gauge the quality of individual hotspot/fire pixels. Confidence estimates range between 0 and 100%.
Bright_T31	Channel 31 brightness temperature of the fire pixel measured in Kelvin.
FRP	Depicts the pixel-integrated fire radiative power in MW (megawatts).
DayNight	D = Daytime, N = Nighttime

For our analysis,we have used only 3 of the below data attributes
(1) Brightness
(2) FRP
(3) Confidence. (AConfidence level greater than 90% is coded as '1' [fire] and '0' [no fire] otherwise)
A model was built using the new data set with the above mentioned data attributes.

4.1 Prediction of FIRE using Logistic Regression

Initially model is developed with the variables mentioned above with family as binominal using the training data. And the summary of the model is shown below.

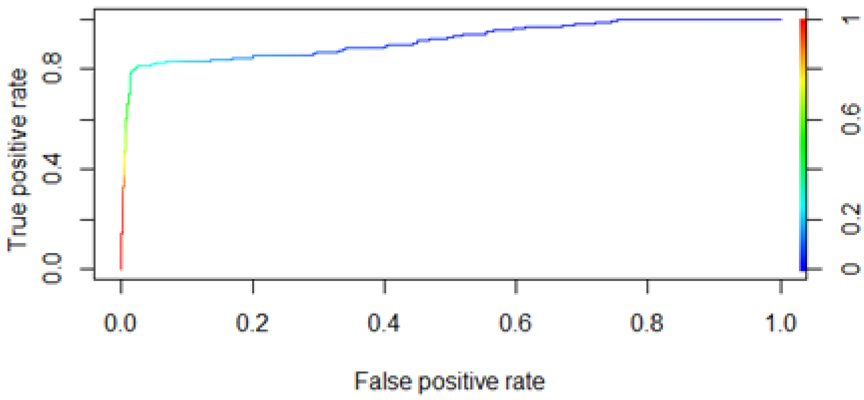
```
>model<- glm (class ~ ., data=trainData, family = binomial)
>summary(model)

Call:
glm(formula = class ~ ., family = binomial, data = trainData)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.9844  -0.3873  -0.1944  -0.0932   3.0644

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -60.180233   5.056659  -11.90   <2e-16 ***
brightness    0.175626   0.015189   11.56   <2e-16 ***
frp          -0.001904   0.001942   -0.98    0.327
---
Signif.codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 889.22  on 999  degrees of freedom
Residual deviance: 449.83  on 997  degrees of freedom
AIC: 455.83

Number of Fisher Scoring iterations: 6
```

The model is fit and variables are significant. And developed logistic model is tested with test data and accuracy is measured. And the confusion matrix for test data is given below.

		Actual	
		0	1
Predicted	0	833	4
	1	108	55

With the confusion matrix we can calculate the accuracy of the model. And the formula to calculate is given below.
= (True positive +True Negative) / (True positive +True Negative +False positive + False Negative)

For the model we developed we get an accuracy of 88%. That is we can predict fire 88% of the times. But still can other model can give better solutions will be question running on mind. And ROC curve for the logistic regression model is shown below.



The ROC curve is yet to touch left corner which indicates the model is good but not the best. It can be improved.

4.2Prediction of FIRE using Support Vector Machines

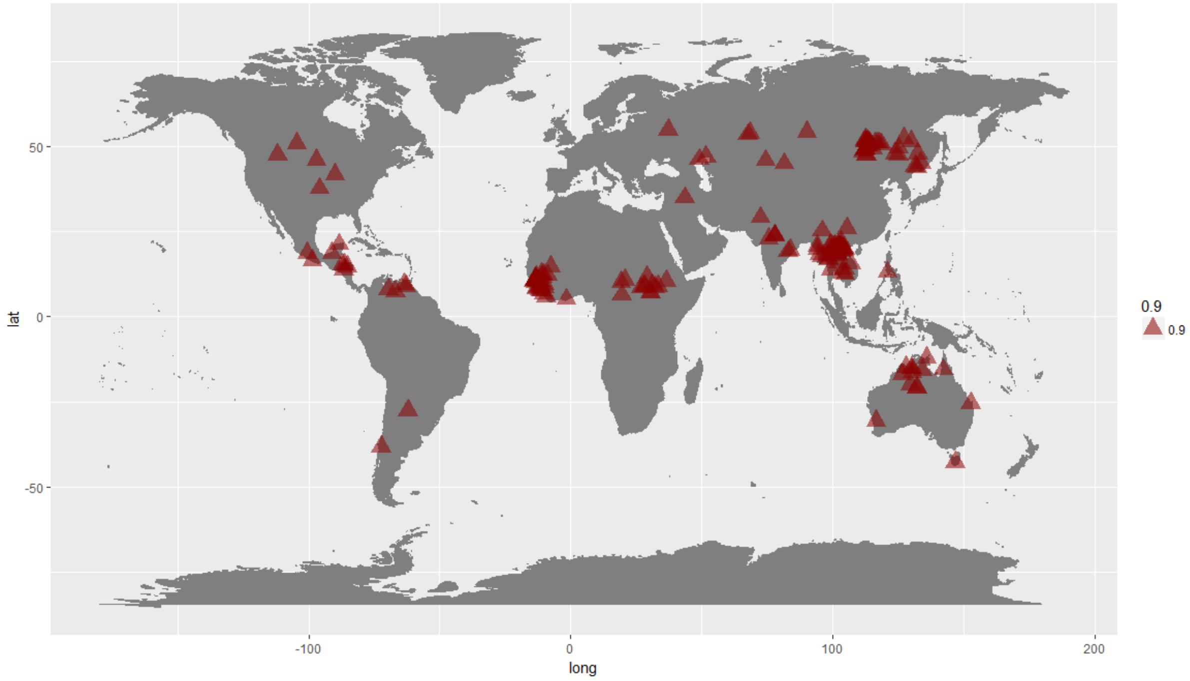
```
>svm_model<- svm(class ~ ., data=trainData,kernel='radial',type='C-classification')
>summary(svm_model)

Call:
svm(formula = class ~ ., data = trainData, kernel = "radial", type = "C-classification")

Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
cost:  1
gamma:  0.5

Number of Support Vectors:  171
( 84 87 )

Number of Classes:  2

Levels:
 0 1
```

With the same variables used for logistic regression, Support Vector Machines model is developed using training data with adjusting the kernel for radial. And the developed model is tested with test data. And the confusion matrix for test data is given below.

		Actual	
		0	1
Predicted	0	804	52
	1	11	133

The model accuracy was 94% that's far better than logistic regression where accuracy was only 88%. And for note, not always SVM is better than Logistic Regression. For our problem SVM performs better than logistic regression. But there will be problem where the outcome will be otherwise.

In order to make result attractive using the Latitudinal and Longitudinal in data, the fire regions were plotted on the map below. The map was developed using R.



5. Conclusion

There are many approaches for fire detection. While some models are complex and others are simple, and some are usually restricted to a particular demography. In this paper, we have developed a simple approach of using SVM and Logistic Regression for wildfire detection across the world. The Logistic Regressionmethodachieves an average detection rate of 88.0% whereas SVM method achieves an average detection rate of 94%. This proves that the SVM classifier has the best stability and the accuracy compared to most of the other approaches proposed till date. The result also shows that both the accuracy and robustness of the classification has improved using SVM and therefore SVM is appropriate for wildfire detection.

Towards end, remember clean data will beat any algorithm. You should remember touse domain knowledge to engineer, and should try various iterations of ideas for better result. And with efficient computing availability should try to ensembles multiple models for better accuracy. And ensembles multiple methodsis one the limitations of paper and will be addressed in our next paper.

References

Journal

[1] Alicia M. Kinoshita et al (2016). Wild fire, water, and society: Toward integrative research in the "Anthropocene". Anthropocene 16 (2016) 16–27.
[2] Andreucci, F., Arbolinno, M.V.,1993. A study of Forest Fire Automatic Detection Systems. IL NuovoCimento, Vol.16 C, N.1.
[3] Bai, X., Zhang, H., Zhou, J., 2014. VHR object detection based on structural feature extraction and query expansion. IEEE Trans. Geosci. Remote Sens. 52, 6508–6520.
[4] Bi, F., Zhu, B., Gao, L., Bian, M., 2012. A visual search inspired computational model for ship detection in optical satellite images. IEEE Geosci. Remote Sens. Lett. 9, 749–753.
[5] BiswajeetPradhan., MohdDiniHairi Bin Suliman.,MohamadArshad Bin Awang., 2007. Forest fire susceptibility and risk mapping using remote sensing and Geographical Information Systems(GIS). Emerald Group Publishing Limited. Disaster Prevention and Management, Vol. 16 No. 3, pp. 344-352.
[6] Bovolo, F., Bruzzone, L., Marconcini, M., 2008. A novel approach to unsupervised change detection based on a semisupervised SVM and a similarity measure. IEEE Trans. Geosci. Remote Sens. 46, 2070–2082.
[7] Boyd M. Harnden., Michael Maher, P., Gregory A. Martin., 1973. Forest Fire Detection Systems Design.Jstor. Management Science, Vol. 20, No. 4, Application Series, Part 2, Management Science in Canada, pp. 617-628.
[8] Cheng, G., Han, J., Zhou, P., Guo, L., 2014a. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. ISPRS J. Photogramm. Remote Sens. 98, 119–132.
[9] Cheng, G., Han, J., Zhou, P., Guo, L., 2014b. Scalable multi-class geospatial object detection in high-spatial-resolution remote sensing images. In: Proc. IEEE Int. Geosci. Remote Sens. Sympos., pp. 2479–2482.
[10] Daniel Chongo., RyotaNagasawa ., AhmedouOuldCherif Ahmed ., Mst Farida Perveen., 2007. Fire monitoring in Savanna ecosystems using MODIS data: a case study of Kruger National Park, South Africa.International Consortium of Landscape and Ecological Engineering and Springer. 3:79–88.
[11] Das, S., Mirnalinee, T., Varghese, K., 2011. Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. IEEE Trans. Geosci. Remote Sens. 49, 3906–3931 .
[12] De Morsier, F., Tuia, D., Borgeaud, M., Gass, V., Thiran, J.-P., 2013. Semi-supervised novelty detection using svm entire solution path. IEEE Trans. Geosci. Remote Sens. 51, 1939–1950.
[13] Emmy Prema, C., Vinsley, S.S., Suresh, S., 2016. Multi Feature Analysis of Smoke in YUV Color Space for Early Forest Fire Detection.Springer Science.Fire Technology. 52, 1319–1342.
[14] Gong Cheng.,Junwei Han., 2016.A survey on object detection in optical remote sensing images. Elsevier. ISPRS Journal of Photogrammetry and Remote Sensing 117,11–28.
[15] Han, J., Zhang, D., Cheng, G., Guo, L., Ren, J., 2015. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. IEEE Trans. Geosci. Remote Sens. 53, 3325–3337.
[16] Huang, X., Zhang, L., 2009. Road centerline extraction from high-resolution imagery based on multiscale structural features and support vector machines. Int. J. Remote Sens. 30, 1977–1987.
[17] Inglada, J., 2007. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. ISPRS J. Photogramm. Remote Sens. 62, 236–248.
[18] JieHou., JiaruQian ., Weijing Zhang., Zuozhou Zhao., Peng Pan., 2011. Fire detection algorithms for video images of large space structures.Springer Science.Multimed Tools Appl 52:45–63.
[19] Li, Z., Itti, L., 2011. Saliency and gist features for target detection in satellite images. IEEE Trans. Image Process. 20, 2017–2029.
[20] MajaStula ., DamirKrstinic ., LjiljanaSeric ., 2012. Intelligent forest fire monitoring system.Springer Science.InfSyst Front 14:725–739.
[21] Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 66, 247–259.
[22] Song, M., Civco, D., 2004. Road extraction using SVM and image segmentation.Photogramm. Eng. Remote Sens. 70, 1365–1371 .
[23] Sun, X., Wang, H., Fu, K., 2010. Automatic detection of geospatial objects using taxonomic semantics. IEEE Geosci. Remote Sens. Lett. 7, 23–27
[24] Tao, C., Tan, Y., Cai, H., Tian, J., 2011. Airport detection from large IKONOS images using clustered SIFTkeypoints and region information. IEEE Geosci. Remote Sens. Lett. 8, 128–132.
[25] Yang Jia., Gaohua Lin., Jinjun Wang., Jun Fang1., Yongming Zhang., 2016. Light Condition Estimation Based o n Video Fire Detection in Spacious Buildings. Springer. Arab J SciEng 41:1031–1041.
[26] Zhang, W., Sun, X., Fu, K., Wang, C., Wang, H., 2014b. Object detection in highresolution remote sensing images using rotation invariant parts based model. IEEE Geosci. Remote Sens. Lett. 11, 74–78.
[27] Zhang, W., Sun, X., Wang, H., Fu, K., 2015b. A generic discriminative part-based model for geospatial object detection in optical remote sensing images. ISPRS J. Photogramm. Remote Sens. 99, 30–44.
[28] Zhu, C., Zhou, H., Wang, R., Guo, J., 2010. A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features.IEEE Trans. Geosci. Remote Sens. 48, 3446–3456.

Book References

[29] Vapnik, V.N., Vapnik, V., 1998. Statistical Learning Theory.Wiley, New York.

Website References:

[30] Forest Fire Numbers Rise in India in 2016.https://www.nasa.gov/image feature/goddard/2016/forest-fire-numbers-rise-in-india-in-2016
[31] About NASA’s EOS. http://earthobservatory.nasa.gov/
[32] Download Near Real-Time Data. https://earthdata.nasa.gov/earth-observation-data/near-real-time/download-nrt-data
[33] What's New in Collection 6 MODIS Active Fire / Hotspot Data?.https://earthdata.nasa.gov/what-is-new-collection-6-modis-active-fire-data
[34] Images for SVM and Logistic Regression. http://www.edvancer.in/logistic-regression-vs-decision-trees-vs-svm-part1
[35] Basics of Logistic Regression. https://www.analyticsvidhya.com/blog/2015/11/beginners-guide-on-logistic-regression-in-r/
[36] Basics of wildfires. http://eschooltoday.com/natural-disasters/wildfires/how-to-fight-wildfires.html