Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018)
IEEE Xplore Compliant Part Number: CFP18K74-ART; ISBN:978-1-5386-2842-3

Data mining approach to predict forest fire using fog computing

Aakash Rajagopal S
B.Tech, Computer Science and Engineering Dept., SASTRA Deemed University, Thanjavur – 613401, India
aakashrajagopal07@gmail.com

Moses Nishanth
B.Tech, Computer Science and Engineering Dept., SASTRA Deemed University, Thanjavur – 613401, India
dmosesn@gmai.com

Rajageethan R
B.Tech, Computer Science and Engineering Dept., SASTRA Deemed University, Thanjavur – 613401, India
Chrisdante15@gmail.com

Ramachandra Rao
B.Tech, Computer Science and Engineering Dept., SASTRA Deemed University, Thanjavur – 613401, India
purohithram@gmail.com

Ezhilarasie R
Assistant Professor-II School of Computing SASTRA Deemed University, Thanjavur – 613401, India
ezhil@cse.sastra.edu

***Abstract*—Fires caused in forests are known to be one of the most hazardous environmental issues which cannot be neglected. Fast and quick prediction is the only way by which we can at least face it with readily available fire extinguishing resources. To accomplish this , one of the best ways is to use automatic tools based on locally placed sensors, such as rain, thermal readings, relative dampness and wind. In this work, we inquire a Data Mining (DM) approach to predict the area prone to forest fire. SVM (Support Vector Machine) algorithm is used for live sensor data. SVM algorithm utilizes four sensor inputs related to weather info data (i.e. thermal data, relative dampness, rain and wind velocity) and it is feasible to predict the burned land zone due to small fires, which are periodic or frequent. Awareness about those zones are utilized for developing and deploying fire fighting resource management . The data received from the sensor devices are processed in individual fog nodes and the cumulative data is collected upon and is used for further analysis. The data transfer happens wirelessly via Zig-Bee tool. The results obtained is utilized to predict the areas which are prone to and will be affected by sudden outburst of forest fire.**

Keywords*—Internet of Things (IoT), Fog based computing, Canadian Forest Fire Index, Support Vector Machine (SVM).

I. INTRODUCTION

Fires caused in forest areas are known to be one of the most problematic environmental hazards, creating damage in both economical and ecological way. The rate of occurrence of forest fires have been increased dramatically over the recent years. So, prediction of forest fires is a vital factor in controlling them. The arguments and values analyzed in determining forest fire are rainfall, humidity of that area, wind speed and temperature variations.

Several predictions methodologies were proposed earlier and some of them were even effective by implementing Support Vector Machine (SVM). It is a developed learning model paralleled with learning algorithms examining the given input data and uses it for grouping and regression data analysis on the same.

Fast detection and notifying the concerned person is the key element in a successful fire fighting process. Since traditional human surveillance is expensive and gets affected by many subjective factors, demand for an automated solution got high. Satellites have procurement costs and the determination is not adequate and suitable for all cases. Moreover, scanners are quite costly to use. Weather conditions, such as temperature, mist and humidity are known to affect fire occurrence. Since programmed meteorological stations are frequently accessible, such information can be gathered progressively, with least cost.

Improvement of Internet of things (IOT) [1] has the component size of wearable and brilliant home/city applications. The emotional augmentation in the number of web depended or connected devices, the increased necessary of real-time, highly responsive services are describing to be challenging for the cloud computing framework. For cost-proficient provisioning constrained assets, we together research undertaking circulation furthermore virtual machine arrangement towards MIST [1]. Results demonstrate that MIST based conspire beats conventional distributed computing as the quantity of utilizations requesting ongoing administration increments.

Fog Computing[6][7] is a computing concept that employs at least one community end-client customers to complete an imperative measure of capacity (as opposed to put away fundamentally in cloud server farms), communication, control, design, estimation and administration. Otherwise called as Fog computing or Edge Computing, it facilitates the operations like computing, storage, and networking services between end devices and cloud computing data centers. While edge computing is traditionally referred to the location where services are performed, fog computing refers to the essential counts utilizing look-into tables with readings from four meteorological perceptions (i.e. temperature, relative dampness, wind speed and precipitation). Fog computing[8][9][10] is a medium weight and intermediate level of computing power which can be placed exactly between the levels of cloud computing and device level computing.

Mist computing[1], a light weight and with an essential type of figuring (computation) power which uses microcomputers or microcontrollers to feed into Fog Computing nodes and subsequently feed into the Cloud computing nodes later.

Fog computing is used to reduce the computational time and to increase the bandwidth to achieve low response time. Forest fire spreads at a larger rate and hence it is necessary to use fog computing and in order to anticipate timberland fire using fog computing is very much necessary.

II. RELATED WORKS

In the past, meteorological data has been incorporated into number based values, which are used for preventing any dangers due to fire (e.g. warning the public of a fire danger) and to support fire management decisions (e.g. state of readiness, prioritizing and sorting targets or evaluating protocols for safe fire fighting). In particular, the Canadian forest Fire Weather Index (FWI) system was designed in the 1970s when computers were not common, thus it demanded essential counts utilizing look-into tables with readings from four meteorological perceptions (i.e. temperature, relative dampness, wind speed and precipitation) that could be received from weather stations. Nevertheless, this way is highly used in several countries around the world nowadays.

On the other hand, the interest in Data Mining (DM)[11][12], also known as Knowledge Discovery in Databases (KDD), came into action due to the advancements in Information Technology, leading to an dramatic growth of business, scientific and engineering databases. Every one of these information holds crucial informations, which incorporates examples and patterns, which can be utilized to enhance decision making standards.

A few DM strategies have been connected to the fire location space. For example, we use Neural Networks (NN) to predict human caused wildfire situations. Infrared scanners and NN were jointly used in to reduce forest fire false alarms with a 90% success. A spatial grouping in image mining was embraced by Hsu et al to recognize timberland fire spots in satellite pictures. In 2005, satellite pictures from North America woods fires were encouraged into a Support Vector Machine (SVM)[3][4][5], which got 75% precision at discovering smoke at the 1.1-km pixel level. Apply Logistic Regression models, Random Forest (RF) and Decision Trees (DT) to detect the presence of fire in the Slovenian timberland, using the detailed inputs from satellite-data and also from meteorological data. Best model was produced by a bagging DT, with an overall 80% accuracy.

They have significantly done a remarkable work in implementing a prediction model[2] for forest fire using thermal reading, relative dampness, wind velocity and Rain. Several experiments were carried by considering five DM techniques (namely multiple regression, SVM, NN, RF and DT) and four feature selection setups (i.e. using spatial or temporal data, the FWI system and meteorological data). The recommended solution involves only four weather variables (i.e. rain, wind, temperature and humidity)[13][14][15] in association with a SVM and it is fit for anticipating the consumed territory of negligible flames, which constitute the large amount of the fire occurrences. Such knowledge is specifically useful for fire management decision support (e.g. planning the resources under safety protocols).

We implement SVM by collecting the data using sensors and fog computing is implemented. Fog nodes are connected in a specially appointed network system and if an area is found to be prone for forest fire we alert the forest authority.

III. ALGORITHM

A. FOREST FIRE DATA

SVM can be utilized in anticipating forest fire with more accuracy with these parameters: rain, wind speed, relative humidity and temperature. The data are collected using arduino board using thermal relative dampness and rain sensors. Wind speed is an important factor and keeps varying over time however for a particular area depending upon if its city or rural it proves to be a constant. Hence a constant value is taken based on that. All these data are collected inside our University in places with high concentration of trees.

Support Vector Machines are unmistakable class of calculations, utilizing kernels, local minima's absence, meagre condition of the arrangement and limit control obtained by acting on the edges, numerous support vectors, etc. SVM can be connected to both characterization and relapse issues. The importance in utilizing SVM is, enormous amount of computational advantages can be obtained from a small subset of training points. The intensive loss function epsilon helps us to ensure global minimum and also optimization of reliable generalization bound.

In SVM, the info is mapped onto a n-dimensional feature space using non-linear mapping, based on which a linear model will be constructed in the feature space. Depending on the application, specific kernel function can be selected. The necessary data are collected from different sensors, which holds value about Temperature, Relative Humidity,Wind Speed and precipitation.

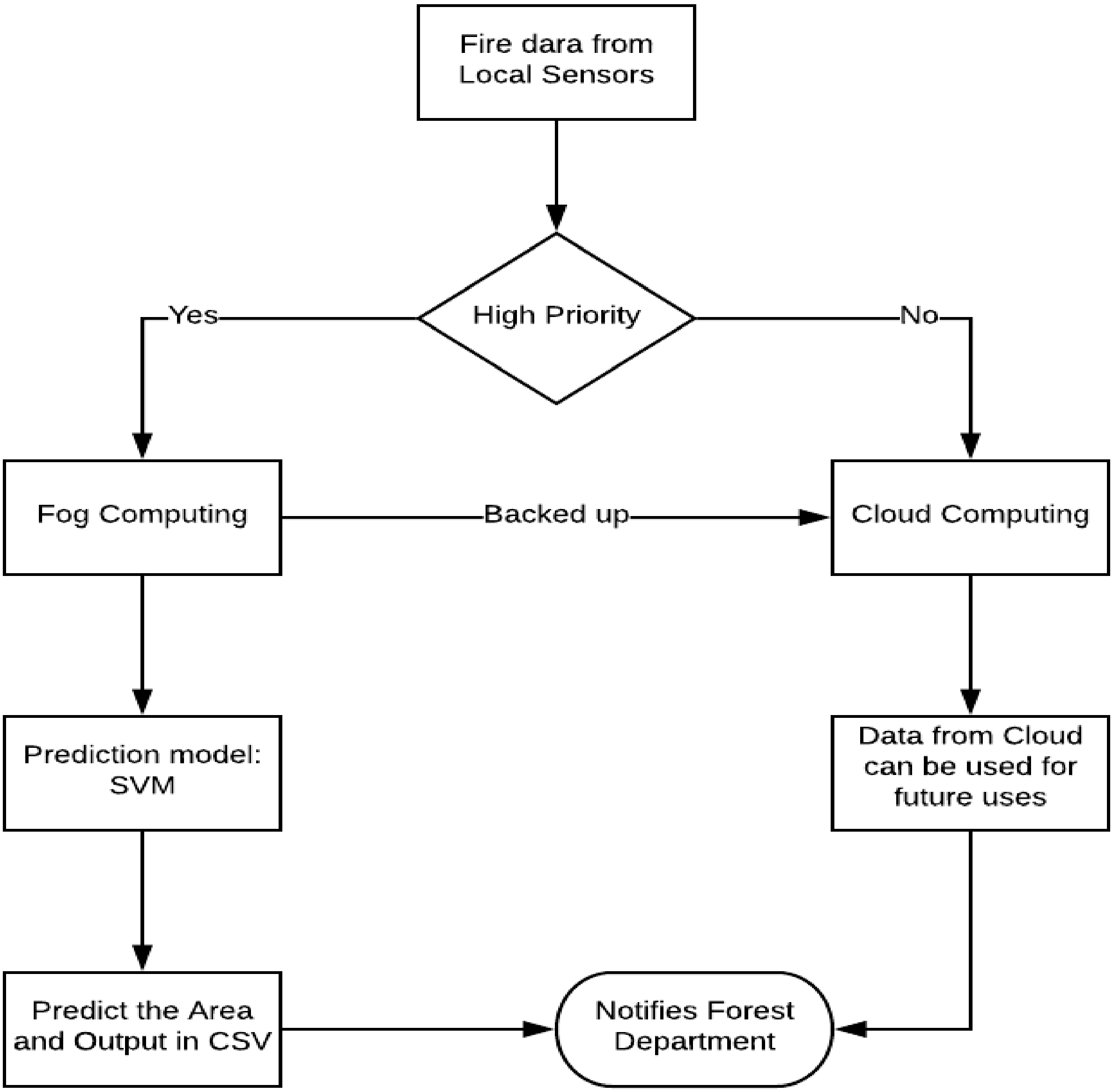

Fig. 1. Overview of Forest fire prediction using fog

B. METHODOLOGY

The raw data collected from different sensors are then transmitted to a nearest fog node, where in the initial computation takes place.On applying logarithm over the dataset we arrive at the resultant dataset which can be used in the training of SVM algorithm. Here the Support Vector Machine (SVM) is configured with linear kernel and epsilon of value 0.329. .To minimize the error rate in raining dataset, an exponential function (e^x-1) is applied.The predict function uses the provided model on the collected dataset in prediction. The correctness of training dataset can be evaluated by using Mean Absolute Deviation (MAD) and Root Mean Square Error (RMSE). Finally, the algorithm is fed with the sensor data obtained from each node. The results obtained can be utilized in the prediction of the areas vulnerable to forest fire.

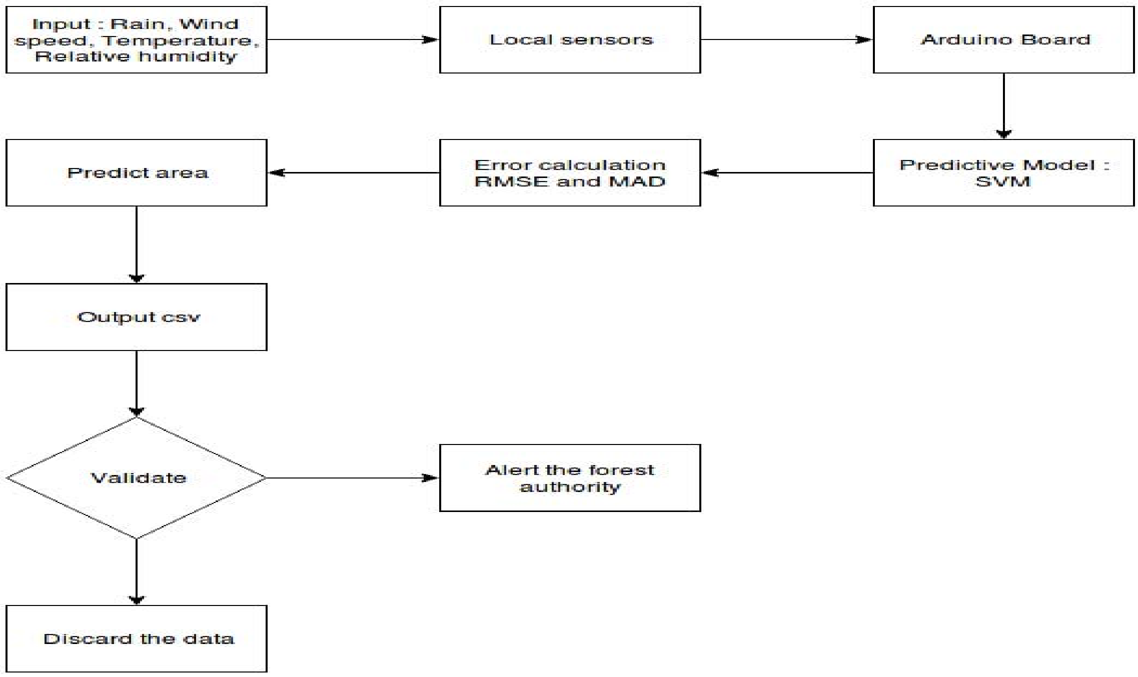

Fig. 2. Data Flow in Forest fire prediction

C. PSEUDOCODE

1. a = read(sensor data)
2. b = read(training data)
3. c = b
4. for(j in 1: nrow(b))
5.   c[j,13] = log(c[j,13]+1)
6. End for
7. Ds = subset(c, select = c (temp, RH, wind, rain, area))
8. traindata = ds[1:517,]
9. linear = svm(area~data = traindata, type="eps-regression", kernel="linear", gamma = 1000, cost = 3, epsilon = 0.329, cross = 517)
10. predictlinear = predict(linear,traindata)
11. for(k: the number of rows in trainingdata)
12.   predictlinear[k] = exp(predictlinear[k])-1
13. End for
14. tc = training data
15. for(d: the number of rows in training data)
16.   tc[d,5] = exp(tc[d,5])-1
17. End for
18. rmse(tc$area, predictlinear)
19. MAD = 0.0
20. for(k: number of rows in training data)
21.   MAD = MAD + abs(tc[k,5] – predictlinear[k])
22. End for
23. MAD = MAD/517
24. sensor_data = predict(linear, pre-processed sensor data)
25. for(e: the number of rows in sensor data)
26.   sensor_data[e] = exp(q[e])-1
27.   write(sensor_data to file)
28. End for

IV. RESULTS AND DISCUSSIONS

The algorithms mentioned above were implemented and analyzed in R language in the local server based FOG computing methodology. The MAD value were found to be in the range 12.80-12.83 and the RMSE value in the range 64.78-64.81.This trained data helped in predicting the forest fire area in hectares with an accuracy of 46%. Then the forest authority is to be alerted if the conditions of prediction were severe.

TABLE I. PREDICTED FOREST FIRE AREA

TEMP	RH	WS	PREC	AREA
32	27	15	0	0.594285177
33	26	17	0	0.624201548
33	26	20	0	0.655005365
33	26	22	0	0.675865084
33	25	19	0	0.642754844
33	25	21	0	0.663460188
33	25	17	0	0.622307223
33	25	14	0	0.59211203
32	26	13	0	0.572604584
33	25	19	0	0.642754844
33	25	23	0	0.684426502
33	25	18	0	0.632499035
33	25	20	0	0.653075114

V. CONCLUSION

The forest fire detection using fog computing methodology has been very effective and fog computing proved to be as efficient as it is was expected to perform. Processing data in fog level with the SVM algorithm to predict the area of fire using the weather parameters(rainfall, humidity, wind speed, temperature) has helped us achieve low latency and fast results which is the key factor in the case of real time emergency situations. Thus fog computing can be used for many real time - emergency decision making situations and simultaneously the data can be fed into the cloud later for further processing and data recovery.

REFERENCES

[1] Hamid Reza Arkian, Abolfazl Diyanat, Atefe Pourkhalili. “MIST: Fog-based data analytics scheme with cost-efficient resource provisioning for IoT crowdsensing applications.”
[2] Paulo Cortez and Anibal Morais. “A Data Mining Approach to Predict Forest Fires using Meteorological Data.”
[3] Rupan Panja, Nikhil R. Pal."MS-SVM: Minimally Spanned Support Vector Machine."
[4] Hong-tao Zhang, Ming-xu Gao."The Application of Support Vector Machine (SVM) Regression Method in Tunnel Fires."
[5] Yong Liu, Shizhong Liao."Granularity selection for cross-validation of SVM."
[6] Redowan Mahmud, Satish Narayana Srirama, Kotagiri Ramamohanarao, Rajkumar Buyya." Quality of Experience (QoE)-aware placement of applications in Fog computing environments."
[7] Pengfei Hu, Sahraoui Dhelim, Huansheng Ning, Tie Qiu."Survey on fog computing: architecture, key technologies, applications and open issues."

[8] Roby Lynn, Ethan Wescoat, Dongmin Han, Thomas Kurfess. "Embedded fog computing for high-frequency MTConnect data analytics."

[9] Yan Sun, Nan Zhang. "A resource-sharing model based on a repeated game in fog computing."

[10] Kang Kai, Wang Cong, Luo Tao."Fog computing for vehicular Ad-hoc networks: paradigms, scenarios, and issues."

[11] Juergen Lenz, Thorsten Wuest, Engelbert Westkämper."Holistic approach to machine tool data analytics."

[12] Kaleem Razzaq Malik, Yacine Sam, Majid Hussain, Abdelrahman Abuarqoub."A methodology for real-time data sustainability in smart city: Towards inferencing and analytics for big-data."

[13] Wei Fang, Xinhong Song, Xiaojun Wu, Jun Sun, Mengqi Hu."Novel efficient deployment schemes for sensor coverage in mobile wireless sensor networks."

[14] Xu Xiaoli, Wang Huan."The Wireless Sensor Network Construction of the Photovoltaic Power System Zig Bee."

[15] M.I. Benakila, L. George, S. Femmam."A Beacon-Aware Device For The Interconnection of Zig Bee Networks."