Paper 6_evizal v.3

by Evizal Abdul Kadir

Submission date: 11-Sep-2022 10:21PM (UTC+0700)
Submission ID: 1896987446
File name: Paper2_Evizal_v.1.docx (689.18K)
Word count: 3085
Character count: 16430

Fire Hotspots Mapping and Forecasting in Indonesia Using Deep Learning Algorithm

Sri Listia Rosa
Department of Informatics Engineering
Universitas Islam Riau
Pekanbaru, 28284 Indonesia
srilistiarosa@eng.uir.ac.id

Evizal Abdul Kadir
Department of Informatics Engineering
Universitas Islam Riau
Pekanbaru, 28284 Indonesia
evizal@eng.uir.ac.id

Abdul Syukur
Computer Science and Information Engineering, National Taiwan University Science and Technology, Taipei,Taiwan
d10815810@gapps.ntust.edu.tw

Hitoshi Irie
Centre for Environmental Remote Sensing, Chiba University
Chiba, 263-0022 Japan
hitoshi.irie@chiba-u.jp

Rizky Wandri
Department of Informatics Engineering
Universitas Islam Riau
Pekanbaru, 28284 Indonesia
rizkywandri@eng.uir.ac.id

Muhammad Fikri Evizal
Computer Science and Information Engineering, National Dong Hwa University, Hualien, Taiwan
muhammadfikrievizal@gmail.com

***Abstract* —Indonesia is one of the countries in South East Asia has significant forest fire with dangerous impact to neighboring countries of the emission of haze and carbon. In this research aims to do plotting and mapping location with high number fire hotspot then forecasting potential number of hotspots in future time based on previous of history data collected. The forecasting data achieve is very important and beneficial for the authorities as one of references for preventive action and avoid scattering of forest fire. Long Short-Term Memory (LSTM) algorithm implemented in this research for analysis and forecasting of fire hotspot number while Python programming used to plot hotspot point. The source of fire hotspot dataset is referred to The National Aeronautics and Space Administration (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) recorded from year 2021 with total number is about 100,000 hotspots in Indonesia region. Results show the distribution of fire hotspot concentration most in Sumatra and Kalimantan Island because the typical of land which peat that potential for getting fire. Forecasting of number hotspot for the year 2022 has achieve with good results with error less than 5% which only 4.56%.**

***Keywords*— Fire Hotspots, Mapping, Forecasting, LSTM Algorithm**

I. Introduction

Forest fire is a common issue globally, most of region with high intensity of wild and forest potential for getting fire. Beside that behavior of community living around the forest area is another source of getting fire for example free land cleaning and de-forestry. Indonesia is a country located in South East Asia which tropical region that has only two season which dry or summer and raining. Contour and typical land of Indonesia with forestry and peat land especially for Sumatra and Kalimantan Islands high potentially contribute fire in dry season. Most of every year forest fire happened in summer because of typical land and behavior of community in rural area with lack of education the cleaning the land by deforestation. Furthermore, impact of the forest fire to the community is very dangerous for example air pollution due to haze and carbon emit from the fire to the human health as well as for the ecosystem surrounding the forest area. Children and human has difficulty fir breath and respiratory issue because of bad oxygen as well as for flora and fauna.

Research in this topic has been done by several researchers in prediction of forest fire with various of approach and objectives. In [1-4] discussed method of data analysis in prediction of fire hotspot used machine learning with several kinds of model. Climate change has influence of the characteristic and typical source of environmental and data aspect. The prediction process of fire hotspot and accuracy related to the meteorological data as well as other factor in environmental changes. The other research discussed on the data analysis with comprehensive study of fire hotspot by measure the size of spot and fire concentration in common country potentially getting fire in tropical region. The analysis of fire hotspot also considers the color and size of fire with prediction potential spreading scales as elaborated in [5-7]. The occurrence of fire hotspot determined of how high potential and the impact due to forest fire.

In the first stage of analysis by identify amount of smoke spread on the air from forest fire is one of the techniques discussed by [8-11] to check how much the potential the area getting fire. LSTM algorithm applied to identify of model and pattern of the previous collected fire hotspot data. The prediction only covers in the small dedicated area to proof that proposed algorithm is working fine. Wild and forest fire investigation and forecasting by consider fire dataset by computerized reasoning as discussed is one of method to do a prediction of fire spreading. Recurrent Neural Network (RNN) is an algorithm which capable to integrated in prediction and propagation fire spreading likely better in data analysis by a time series data. Wireless sensor network (WSN) is a technology used to detect wild and forest fire in low level or ground sensing system as elaborate in [12-15]. The rising of ground temperature due to forest fire has impact to many of ecosystem on earth, efforts have been taken to minimize and overcome the fire by prediction the event. WSN system applied for ground sensing with advantages to collect fire data directly with high accuracy and minimum noise, but in limited cover area due to short sensing range. Another discussed results as mention in [16] which a method to do predication of forest fore hotspot used machine learning algorithm.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. Fire Data and Monitoring

There are many kinds of natural disaster worldwide, wild and forest fire is one of the disaster due to natural phenomena for example extreme hot land due to summer or dry season. Indonesia is a country located in South East Asia with typical forestry because of in tropical region. Forest fire one of the disasters that frequency happen in almost every year during summer because of typical peatland that easy getting fire when the ground is dry. Many efforts implement by government, industries and community to prevent case of forest fire but not significantly solve the issue, several research have been conducted as well to find the source of fires. In this research use deep learning algorithm and LSTM model to do a prediction and plot forecasting results in order to achieve distribution and scattering area of fire hotspot in Indonesia region. Collected fire dataset by NASA is very useful to analysis the data and plot in a map, then how the hotspot distribution be able to analysis in detail.

The method used which is LSTM deep learning algorithm model has ability to handle the problems with long-term dependencies of RNN. While in some other algorithms unable to do in accurate result because of long history and variety of data. Big numbers of data with variety in which the conventional algorithm unable to do in precise of the information stored in the long-term memory but can gives more accurate prediction from the recent information. LSTM can apply by default to retain the data in a long-term period of time. Normally used for predicting, processing, and classifying on the basis of time-series data [17]. Prediction model to calculate future number of fire hotspot used LSTM algorithm require to justify the error, several model can be used to calculate the forecasting such as Mean Square Error (MSE) to calculate error in square. While Mean Average Error (MAE) to calculate average error in a prediction of dataset. A method called R^2 shows a proportion the number of variants in the prediction of dataset, the calculation of all those methods can be find as equation (1), (2), and (3) for *MSE*, *MAE*, and R^2respectively.

$$MSE = \frac{\sum_1^N (yi - y_i')^2}{N} \quad (1)$$

$$MAE = \frac{\sum_1^N |yi - y_i'|}{N} \quad (2)$$

$$R^2 = 1 - \frac{\sum_1^N (yi - y_i')^2}{\sum_1^N (yi - y_{avg})^2} \quad (3)$$

where y_i is actual number of fire hotspot in a time of i, y_i' is a number of prediction hotspot in a time of i^{th}, y_{avg} is number of sample dataset as training data, while error number use of metric regression model. All the model used to calculate error assisting to check how the performance of forecasting dataset in future time as well as calculate mean error in a simulation.

The source of fire hotspot dataset refers to MODIS data available in NASA earth data, there are 15 indicators in the dataset as shows. MODIS dataset based on satellite imaging to detect the active fire hotspot from the sky, in some case the image received from satellite may in low quality then effected to the decision which number of hotspot and location. Strategy to improve data quality and value that related to the number of fire hotspot which preliminary analysis by filtering incomplete or missing of dataset. Only complete and valuable dataset process by the application to determine location and number of fire hotspot. Total number of MODIS dataset collected from NASA more than 100,000 data as for the year 2021 and forecasting for year 2022 [18]. All the fire dataset only in Indonesia territory, data normalization is applied for the first step of processing to avoid noise and incomplete data in decision. Table 1 shows a set of data after filtering and select only used data for analysis which only four parameters in year 2010 to 2021.

TABLE I
NUMBER OF DAILY HOTSPOT DATA YEAR 2010 TO 2021

No	Latitude	Longitude	Date	Total
0	0.02110	116.87390	2010-01-01	42
1	0.48080	116.08060	2010-01-01	66
11	2.15090	117.49680	2010-01-01	0
10	-8.10890	118.07430	2010-01-01	0
8	-8.15960	117.58570	2010-01-01	43
..	..	..	..	..
14208	-7.22331	110.42920	2021-12-31	67
14209	-6.96059	110.45844	2021-12-31	55
14210	-5.80178	139.61118	2021-12-31	65
12211	-4.51654	136.84802	2021-12-31	56
14213	-4.54666	136.77507	2021-12-31	52

The distribution of fire data refer to the table 1 based on daily event which in single day may have many occurrence of fire hotspot detected. In order to create daily event of fire graph then the data has to scale or group in every single day that consist numbers of hotspot. Table 2 shows fires dataset has been group into a single date of fire occurrence, the total number of data from year 2010 to 2021which is 12 years more than 4000 hotspot. Those number of data used for data training and testing to achieve forecasting in future year. Grouping of the data to single day to optimize in the analysis that forecasting of fire hotspot refer to single data. The distribution of the data plotted in the line graph to check maximum and minimum versus time in every year. Scalability applied to do analysis in detail based on monthly then number of hotspots be able to check in detail compare to yearly. The minimize and analysis only the use of data is good to reduce processing time and memory then fast processing.

TABLE II
HOTSPOT DATA YEAR 2010 TO 2021

No	Date	Total
0	2010-01-01	12
1	2010-01-02	12
2	2010-01-03	5
3	2010-01-04	14
4	2010-01-05	36
..	..	..
4360	2021-12-27	7
4361	2021-12-28	6
4362	2021-12-29	4
4363	2021-12-30	30
4364	2021-12-31	7

A special RNN type of network which is LSTM algorithm working based on a memory that in long data with short-term network has ability to learn in long-term connections. One of the advantages of the LSTM model is to do a forecasting of event in series of time data and wide range of issues can be handle. Thus, many case to solve the problem use this algorithm and model to do analysis of data as well as in prediction of data in future time. Furthermore, LSTM algorithm ability to organize model in order to form of a chain structure and has four interacting layers with a unique method of communication each other's in data processing. Figure 1 shows an analysis block diagram how the forecasting process of the fires hotspot in the future times. Data training dan testing in separate block and running in each process before LSTM algorithm do the analysis and forecasting.

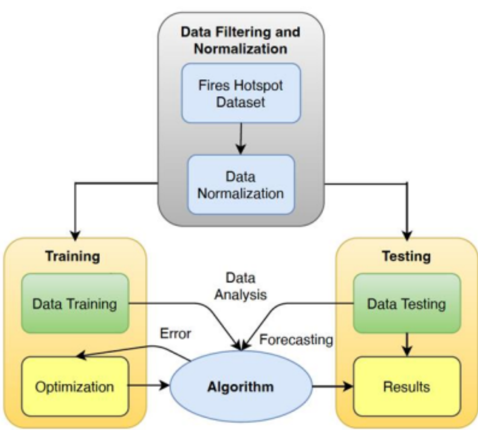

Fig. 1. The analysis of block in data forecasting

The algorithm in this process of data mapping and forecasting is refer to LSTM and the pseudo code as shows in the figure 2, where *X* is the input data of fire hotspot in time series from year 2010 to 2021 and the *O* is output forecasting data for the future year and in this case prediction for one year which is 2022.

Algorithm FORE-LSTM.
Input: data stream $X = \{x_1, x_2, \dots, x_i, \dots\}, x_i \in R$, epoch *I*, number of iterations *K*, error parameters σ, cycle index *N*, Number of decompositions *m*, White noise data *W*.
Output: Forecasting Result *O*.

$X \leftarrow x_1, x_2, \dots, x_i, \dots, x_i \in R$ // Input data
$O \leftarrow \{\}$ // Output data
for each $i \in N$ **do**
  $H_i^+ \leftarrow x + W_i^+$; $H_i^- \leftarrow x + W_i^-$
  for $j \in I$ **do**
    $H_{ij}^+ \leftarrow FORE\ (H_i^+); H_{ij}^- \leftarrow FORE\ (H_i^-)$
  end for
end for
for $j \in m$ **do**
  $O \leftarrow \sum\left(LSTM(IMF_j)\right)$
end for

Fig. 2. The proposed deep learning algorithm

III. RESULTS AND DISCUSSION

The results of mapping and forecasting of fire hotspot dataset refer to MODIS data consist from several year with parameter coordinate of hotspot, date and time, confidence level which represent the probability of occurrence of fire, brightness level, etc. as shown in previous data. In this research and analysis, the available dataset only used four indicator that to process further which are acquisition date (acq_date), coordinate (latitude and longitude), and confidence level. The distribution of level hotspot classified into five categories or level of confidence, this method applied in order to check which hotspot is very potential getting fire for forest and which data is just low level that less occurrence that not much impact. Figure 3 shows a complete map of Indonesia region with number of fire hotspot plotting base on the location and confidence level.

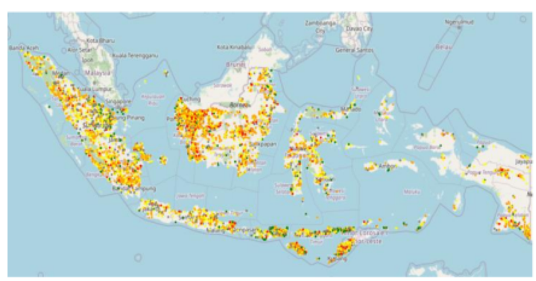

Fig. 3. Mapping of fire hotspot in Indonesia region

The number of fire hotspot in Indonesia region refer to the collected data and confidence level which classified in five levels, then total number hotspot detected in every year with thousands hotspot. In this analysis data presentation by plotted the graph yearly indicate every level from January to December in year 2021. Figure 4 shows the distribution of hotspot every month for every level and the black line is the total number of hotspots for the year 2021.

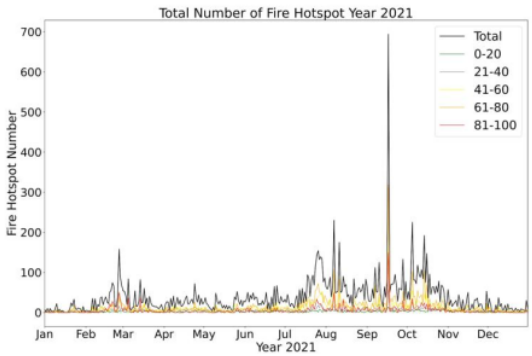

Fig. 4. Fire graph of hotspot event in montly for the year 2021

Finally, forecasting result of fire hotspot achieved for the year 2022, the distribution of training data more than 4000 dataset and 20% testing then plotted into a forecasting graph. Fig. 9 shows the actual data year 2020 and 2021 and forecasting data for the 2022 which similar trend and pattern as well as the distribution of fire hotspot and rise in end of year.

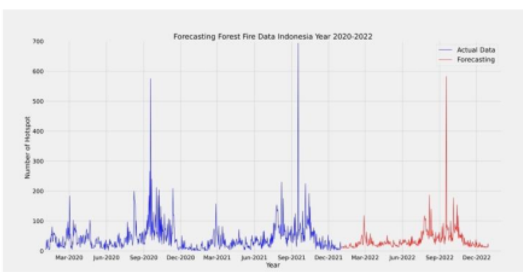

Fig. 5. Forecasting results of the hotspot for the year 2022

IV. CONCLUSION

Forecasting model used LSTM model have been done for the year 2022 with training data 80% and 20% testing data as results shows with performance accurate more than 95% with error 4.56%, the successful of data forecasting indicate by the pattern of data very similar as well as the trend for hotspot in every month. Future work has plan to minimize or scale down the mapping and forecasting in small zone based on state or area may improve the accuracy to achieve detail results.

ACKNOWLEDGEMENT

We would like to express our gratitude to the Ministry of Education, Culture, Research and Technology of Indonesia for funding the research. Chiba University Japan for the overseas joint research program, National Taiwan University of Science and Technology and Universitas Islam Riau, Indonesia for research facilities.

REFERENCES

[1] H. Liang, M. Zhang, and H. Wang, "A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors," *IEEE Access,* vol. 7, pp. 176746-176755, 2019, doi: 10.1109/ACCESS.2019.2957837.

[2] N. Omar, A. Al-zebari, and A. Sengur, "Deep Learning Approach to Predict Forest Fires Using Meteorological Measurements," in *2021 2nd International Informatics and Software Engineering Conference (IISEC)*, 16-17 Dec. 2021 2021, pp. 1-4, doi: 10.1109/IISEC54230.2021.9672446.

[3] T. Lamjiak, R. Kaewthongrach, B. Sirinaovakul, P. Hanpattanakit, A. Chithaisong, and J. Polvichai, "Characterizing and forecasting the responses of tropical forest leaf phenology to El Nino by machine learning algorithms," *PLOS ONE,* vol. 16, no. 8, p. e0255962, 2021, doi: 10.1371/journal.pone.0255962.

[4] E. Abdul Kadir, S. Listia Rosa, A. Syukur, M. Othman, and H. Daud, "Forest fire spreading and carbon concentration identification in tropical region Indonesia," *Alexandria Engineering Journal,* vol. 61, no. 2, pp. 1551-1561, 2022/02/01/ 2022, doi: https://doi.org/10.1016/j.aej.2021.06.064.

[5] N. Khabarov, E. Moltchanova, and M. Obersteiner, "Valuing Weather Observation Systems For Forest Fire Management," *IEEE Systems Journal,* vol. 2, no. 3, pp. 349-357, 2008, doi: 10.1109/JSYST.2008.925979.

[6] J. Han, G. Kim, C. Lee, Y. Han, U. Hwang, and S. Kim, "Predictive Models of Fire via Deep learning Exploiting Colorific Variation," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 11-13 Feb. 2019 2019, pp. 579-581, doi: 10.1109/ICAIIC.2019.8669042.

[7] E. A. Kadir, H. Irie, and S. L. Rosa, "Modeling of Wireless Sensor Networks for Detection Land and Forest Fire Hotspot," in *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, 22-25 Jan. 2019 2019, pp. 1-5, doi: 10.23919/ELINFOCOM.2019.8706364.

[8] X. Li *et al.*, "Prediction of Forest Fire Spread Rate Using UAV Images and an LSTM Model Considering the Interaction between Fire and Wind," *Remote Sensing,* vol. 13, no. 21, p. 4325, 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/21/4325.

[9] Y. Cao, F. Yang, Q. Tang, and X. Lu, "An Attention Enhanced Bidirectional LSTM for Early Forest Fire Smoke Recognition," *IEEE Access,* vol. 7, pp. 154732-154742, 2019, doi: 10.1109/ACCESS.2019.2946712.

[10] S. Natekar, S. Patil, A. Nair, and S. Roychowdhury, "Forest Fire Prediction using LSTM," in *2021 2nd International Conference for Emerging Technology (INCET)*, 21-23 May 2021 2021, pp. 1-5, doi: 10.1109/INCET51464.2021.9456113.

[11] M. S. Akther and Q. K. Hassan, "Remote Sensing-Based Assessment of Fire Danger Conditions Over Boreal Forest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 4, no. 4, pp. 992-999, 2011, doi: 10.1109/JSTARS.2011.2165940.

[12] E. A. Kadir, M. Othman, and S. L. Rosa, "Smart Sensor System for Detection and Forecasting Forest Fire Hotspot in Riau Province Indonesia," in *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 4-5 July 2021 2021, pp. 1-6, doi: 10.1109/ICOTEN52080.2021.9493535.

[13] A. H. Zaji, H. Bonakdari, and B. Gharabaghi, "Remote Sensing Satellite Data Preparation for Simulating and Forecasting River Discharge," *IEEE Transactions on Geoscience and Remote Sensing,*

vol. 56, no. 6, pp. 3432-3441, 2018, doi: 10.1109/TGRS.2018.2799901.
[14] R. L. C. Costa, E. Miranda, P. Dias, and J. Moreira, "Experience: Quality Assessment and Improvement on a Forest Fire Dataset," *J. Data and Information Quality,* vol. 13, no. 1, p. Article 5, 2021, doi: 10.1145/3428155.
[15] E. A. Kadir, S. L. Rosa, and R. A. Ramadhan3, "Detection of Forest Fire Used Multi Sensors System for Peatland Area in Riau Province," in *6th International Conference on Industrial Mechanical Electrical Chemical Engineering 2020*, Surakarta, Indonesia, W. Sutopo, Ed., 2020: AIP Conference Proceedings.
[16] W. Benzekri, A. E. Moussati, O. Moussaoui, and M. Berrajaa, "Early Forest Fire Detection System using Wireless Sensor Network and Deep Learning," *International Journal of Advanced Computer Science and Applications (IJACSA)* vol. 11, no. 5, pp. 496-503, 2020.
[17] N. F. Rahim, M. Othman, R. Sokkalingam, and E. Abdul Kadir, "Forecasting Crude Palm Oil Prices Using Fuzzy Rule-Based Time Series Method." IEEE Access Vol. 6, 2018, pp. 32216-32224.
[18] NASA. Fires Hotspot Information [Online] Available: https://firms.modaps.eosdis.nasa.gov/download/

Paper 6_evizal v.3

ORIGINALITY REPORT

15%	4%	14%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Evizal Abdul Kadir, Hsiang Tsung Kung, Sri Listia Rosa, Andrew Sabot, Mahmod Othman, Mark Ting. "Forecasting of Fires Hotspot in Tropical Region Using LSTM Algorithm Based on Satellite Data", 2022 IEEE Region 10 Symposium (TENSYMP), 2022 Publication	11%
2	repository.uir.ac.id Internet Source	2%
3	epdf.pub Internet Source	1%
4	Www.naturalnews.com Internet Source	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		