Enhancing Wildfire Behavior Prediction With Long Short-Term Memory

장단기 기억 신경망 (LSTM)을 통한 산불 양상 예측

2024.06.18

서울과학기술대학교 데이터사이언스학과 석사과정

22512087 Yaowamal Luetrakulset

Contents

1. Introduction

1.1. Background

- Wildfires are a **serious global problem** that occurs annually in many countries.
- They pose a significant threat to **human life**.
- They cause the loss of property, environmental damage, and economic impacts.
- There are **many cases of wildfires** in the following countries.

1. Introduction

1.2. Wildfire cases

No.	Country	Date	Burned area	No. of deaths	Cause
1.	Chile	The first week of February 2024	**43,000 hectares** (70% of the size of Seoul)	**123 people,** more than 300 people missing	El Niño and Climate change
2.	Maui, U.S. state of Hawaii	August 8th, 2023	**890 hectares** (1.5% of the size of Seoul)	**98 people**	Sparks from power lines, El Niño, and Hurricane Dora
3.	Thailand (Mae Hong Son Province)	March 7th, 2024	**40,000 hectares** (67% of the size of Seoul)	**1 person killed**, 2 injured	The hot, dry, and windy environment in the region.

1. Introduction

1.3. Objectives

- To develop **a new and more comprehensive wildfire dataset** (Thailand) for using in wildfire prediction models.
- **Added variables**: slope direction, wind direction, day of the week, weekend.
- To **develop wildfire prediction models** to predict wildfire occurrence, wildfire size, wildfire speed, wildfire duration, and wildfire direction.
- To develop **LSTM wildfire prediction models** and evaluate their performance against other machine learning (logistic regression, random forest, XGBoost, and ANN).

1. Introduction

1.4. Contribution

- This study aims to create **a more comprehensive dataset** that can improve predictive modeling of wildfires.
- This new dataset **added more variables**, which are slope direction, wind direction, day of the week, and whether it was a weekend or not.
- This study developed LSTM model, which is **the first study using LSTM model to predict wildfire in Thailand**.
- This study developed models to predict various aspects of wildfire behavior, such as **wildfire occurrence, wildfire size, wildfire speed, wildfire duration, and wildfire direction**.
- This study recognized the relationship between slope direction and wind direction and proposed **an interaction variable** to be considered as a significant factor in wildfire behavior prediction models.

2. Related works

2.1. Predictive modeling of wildfires: A new dataset and machine learning approach ***(Sayad et al., 2019)***

- Predicting wildfire occurrence in specific areas of Canadian forests.
- **Region:** Canada
- **Period:** From 2013 to 2014.
- **Data:**
  - Dependent variable:
    - Wildfire occurrence
  - Independent variables:
    - Plant condition (NDVI: Normalized Difference Vegetation Index)
    - Soil temperature (LST: Land Surface Temperature)
- **Models:** Artificial neural networks and support vector machines.

2. Related works

2.2. A Comparison of Forecasting Methods for Seasonal time series with Many Zeros ***(Sawetsuthipan & Charnsethikul, 2021)***

- This study analyzes the time series of daily wildfire areas in Chiang Mai.
- **Region:** Thailand
- **Period:** From 2009 to 2019.
- **Data:**
    - Dependent variable:
        - Number of wildfires
    - Independent variables:
        - Temperature, dew point, humidity, wind speed, air pressure.
- **Models:**
    - Truncated Fourier series (mathematical technique)
    - Holt-Winters' additive method (statistical forecasting technique)
    - Box-Jenkins' SARIMAX method (statistical modelling and time series analysis technique)
    - Multiple regression with categorical variables
    - Multivariate Polynomial Regression
    - Artificial Neural Network (ANN):
        - Input layer = 5
        - Hidden layer nodes = 5, 10, 15, 20, 25, 30, 35, 40.
        - Learning rates = 0.01, 0.05, 0.1, 0.2, 0.4, 0.8, 1.

2. Related works

2.3. Global Wildfire Susceptibility Mapping Based on Machine Learning Models *(Shmuel & Heifetz, 2022)*

- This study predicts wildfire occurrence and burned areas based on meteorological data, fuel characteristics, topography, anthropogenic factors, and regional wildfire history.
- **Region:** Global
- **Period:** From 2003 to 2005
- **Data:**
  - Dependent variable:
    - Monthly wildfire area
  - Independent variables:
    - Temperature, humidity, wind speed, precipitation, mean relative humidity of the previous month, mean precipitation of the previous month, mean relative humidity of the previous year, mean precipitation of the previous year, proportion of burned area, central burned area, damaged area, latitude, longitude, month, drought index, Duff dryness index, fine fuel moisture index, fire danger index.
- **Models:** Random Forest, XGBoost, MLP, and logistic regression.

3. Data

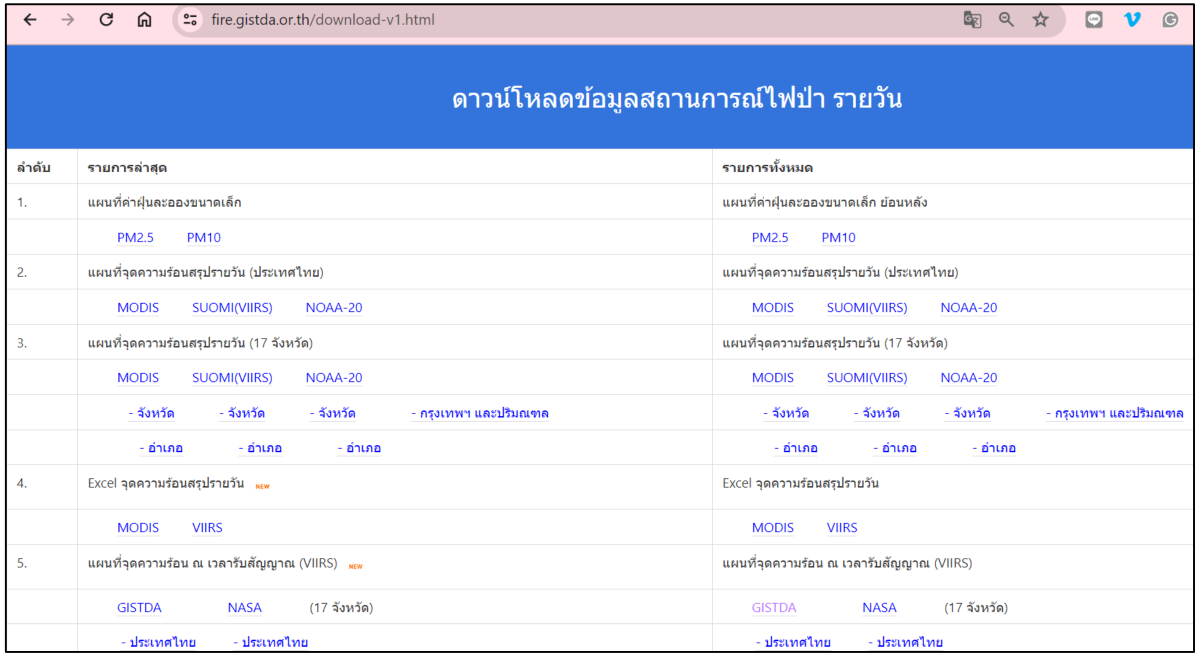
3.1. Data variables

References	Data variables
[21, Radke et al., 2019]	Size, Temperature, Dew_Point, Humidity, Wind_Speed, Pressure, Precipitation, Wind_direction, DEM
[23, Bayat & Ydz, 2022]	DC, DMC, FFMC, Month, Day_of_a_week
[28, Shmuel & Heifetz, 2022]	Status, Temperature_yesterday, Dew_Point_yesterday, Humidity_yesterday, Wind_Speed_yesterday, Pressure_yesterday, Precipitation_yesterday, Temperature_two_days_before, Dew_Point_two_days_before, Humidity_two_days_before, Wind_Speed_two_days_before, Pressure_two_days_before, Precipitation_two_days_before, Temperature_three_days_before, Dew_Point_three_days_before, Humidity_three_days_before, Wind_Speed_three_days_before, Pressure_three_days_before, Precipitation_three_days_before, Slope, Population_density,
[54, Khanmohammadi et al., 2022]	Speed
[55, Xiao, 2023]	Duration
[56, Srivas et al., 2016]	Direction
[57, Hall, 2019]	Slope_direction
[58, Bergado et al., 2021]	Landcover

3. Data

3.2. Fire data (Y1 = Wildfire occurrence)

- **Source:** Thailand Wildfire Monitoring System
- **Provider:** GISTDA (Geo-Informatics and Space Technology Development Agency)
- **Data range:** From 2014 to 2021
- **Collection technique:** Direct download of Excel files from the website (https://fire.gistda.or.th/download-v1.html).



Source: https://fire.gistda.or.th/download-v1.html

3. Data

3.3. Fire data (Y2 = Wildfire size, Y3 = Wildfire speed, Y4 = Wildfire duration, Y5 = Wildfire direction)

- **Source:** Global Fire Atlas dataset
- **Provider:** Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC).
- **Data range:** From 2003 to 2016.
- **Collection technique:** Download directly from the website (https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1642) (.dbf file) and use of latitude and longitude to determine the location of countries and cities. Latitude and longitude are used as key values (Geopy library).

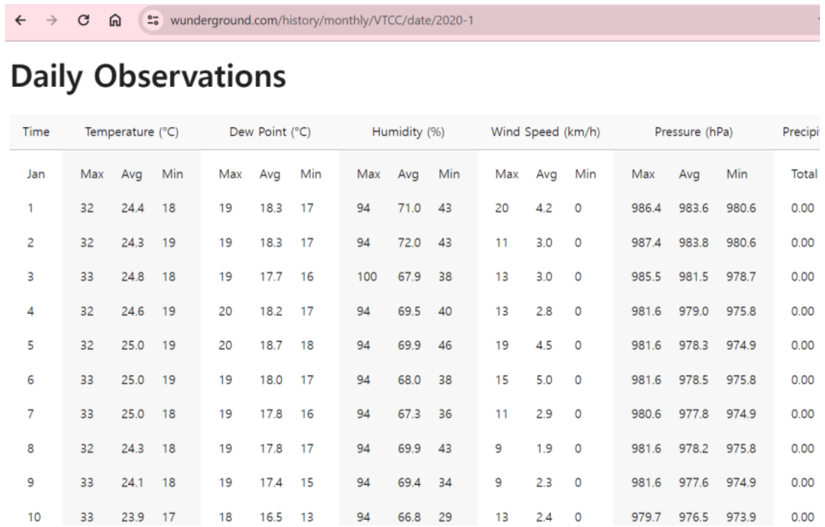
	A	B	C	D	E	F	G	H	I	J
1	fire_ID	latitude	longitude	size	perimeter	start_date	start_DOY	end_date	end_DOY	duration
2	1	21.5313	−158.2130	6.43	12.04	2003-07-20	201	2003-07-30	211	11
3	2	20.6521	−156.2600	0.21	1.85	2003-06-27	178	2003-06-27	178	1
4	3	20.6396	−156.2470	9.00	16.67	2003-07-08	189	2003-07-13	194	6
5	4	20.5396	−156.6700	1.50	7.41	2003-05-02	122	2003-05-11	131	10
6	5	19.4230	−155.4220	6.22	12.96	2003-02-25	56	2003-03-01	60	5
7	6	19.3771	−155.1350	5.14	19.45	2003-05-27	147	2003-05-30	150	4
8	7	19.3146	−155.1150	3.64	10.19	2003-02-07	38	2003-02-23	54	17
9	8	19.3271	−155.0520	0.43	3.70	2003-02-03	34	2003-02-07	38	5
10	9	19.1646	−155.4720	6.43	16.67	2003-02-23	54	2003-02-28	59	6
11	10	19.0063	−155.7730	0.21	1.85	2003-02-21	52	2003-02-21	52	1
12	11	18.9563	−155.6380	##	26.85	2003-07-18	199	2003-07-20	201	3
13	12	43.3479	−123.7710	0.43	2.78	2003-11-01	305	2003-11-02	306	2
14	13	43.3354	−123.7970	1.07	7.41	2003-11-01	305	2003-11-05	309	5
15	14	43.3021	−123.8660	1.07	4.63	2003-11-02	306	2003-11-03	307	2
16	15	43.3021	−123.6890	0.21	1.85	2003-08-03	215	2003-08-03	215	1
17	16	43.2979	−123.8180	0.64	4.63	2003-11-01	305	2003-11-03	307	3
18	17	43.2896	−123.8700	0.21	1.85	2003-11-03	307	2003-11-03	307	1
19	18	43.1771	−123.7440	0.43	2.78	2003-08-01	213	2003-08-01	213	1
20	19	43.0854	−123.8100	0.21	1.85	2003-08-02	214	2003-08-02	214	1

Source: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1642

3. Data

3.4. Meteorological data

- Temperature, dew point, wind speed, wind direction, precipitation, and air pressure.
- **Data range:** Today, yesterday, 2 days before. Data is daily average (Mean). Wind direction data uses the most frequent occurring value in a day (Mode).
- **Source:** Weather Underground
- **Collection technique:** Using city airport codes (XXXX) to collect information for Thailand (https://api.weather.com/v1/location/XXXX:9:TH/observations/historical.json?apiKey=e1f10a1e78da46f5b10a1e78da96f525&units=e&).


wunderground.com/history/monthly/VTCC/date/2020-1

Daily Observations

Time	Temperature (°C)			Dew Point (°C)			Humidity (%)			Wind Speed (km/h)			Pressure (hPa)			Precipi
Jan	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Total
1	32	24.4	18	19	18.3	17	94	71.0	43	20	4.2	0	986.4	983.6	980.6	0.00
2	32	24.3	19	19	18.3	17	94	72.0	43	11	3.0	0	987.4	983.8	980.6	0.00
3	33	24.8	18	19	17.7	16	100	67.9	38	13	3.0	0	985.5	981.5	978.7	0.00
4	32	24.6	19	20	18.2	17	94	69.5	40	13	2.8	0	981.6	979.0	975.8	0.00
5	32	25.0	19	20	18.7	18	94	69.9	46	19	4.5	0	981.6	978.3	974.9	0.00
6	33	25.0	19	19	18.0	17	94	68.0	38	15	5.0	0	981.6	978.5	975.8	0.00
7	33	25.0	18	19	17.8	16	94	67.3	36	11	2.9	0	980.6	977.8	974.9	0.00
8	32	24.3	18	19	17.8	17	94	69.9	43	9	1.9	0	981.6	978.2	975.8	0.00
9	33	24.1	18	19	17.4	15	94	69.4	34	9	2.3	0	981.6	977.6	974.9	0.00
10	33	23.9	17	18	16.5	13	94	66.8	29	13	2.4	0	979.7	976.5	973.9	0.00

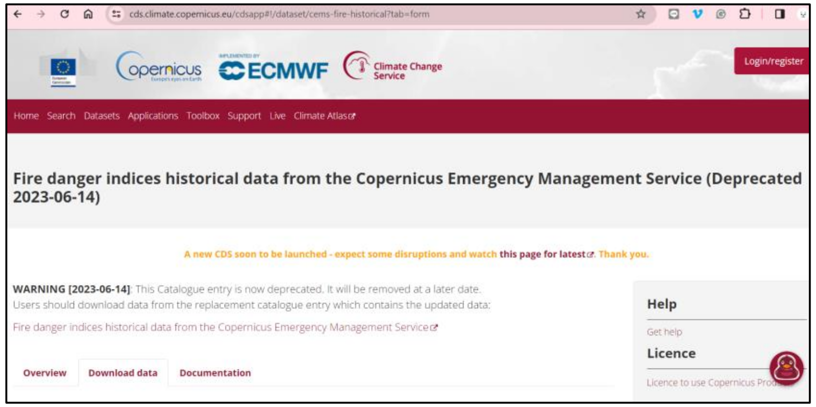
Source: https://www.wunderground.com/history/monthly/VTCC/date/2020-1

3. Data

3.5. Moisture data

3.5.1. Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), and Drought Code (DC)

- **Description:** FFMC represents the dryness of the soil at a depth of approximately 1~2 cm. DMC represents the dryness of the soil at a depth of 2~10 cm. DC represents the dryness of the soil at a depth of 10~20 cm.
- **Source:** Fire Risk Index Record data from Copernicus Emergency Management Service (currently provided until 2023-06-14).
- **Collection technique:** Download NC files from the website (https://cds.climate.copernicus.eu/cdsapp#!/dataset/cems-fire-historical?tab=form) and convert to Excel file using the program Panoply.



Fuel characteristics	Data range
FFMC	0 to 101
DMC	0 to 1000
DC	0 to 1000

Source: https://cds.climate.copernicus.eu/cdsapp#!/dataset/cems-fire-historical?tab=form

3. Data

3.5. Moisture data

3.5.2. Land cover

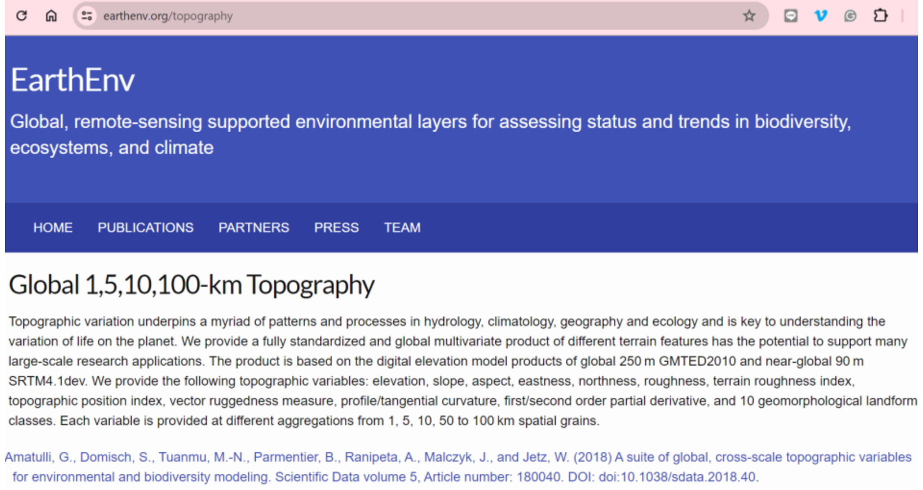
- **Description:** Land cover refers to forests and vegetation types in wildfire ignition areas.
- **Source:** Global Fire Atlas dataset
- **Collection technique:** Download directly from the website (https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1642) (.dbf file) and using latitude and longitude to determine the location of countries and cities. Latitude and longitude are used as key values (geopy library).

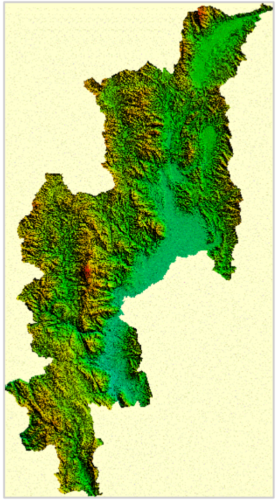
Land cover	Description
Evergreen broadleaf forest	Most of the trees and bushes in this area retain their green foliage throughout the year.
Deciduous forest	It consists of a community of broad-leaved trees that have an annual cycle of leaf sprouting and leaf falling.
Mixed forest	It consists of a mixture or mosaic of tree communities from the other four forest types.
Closed shrub	Land with trees less than 2 meters in height and with a shrub canopy cover of more than 60%.
Open shrub	Land with trees less than 2 meters in height and with a shrub canopy cover of between 10 and 60%.
Woody savannah	Land with herbaceous and other understory systems and between 30 and 60% forest cover.
Savannah	Land with herbaceous and other understory systems and between 10 and 30% forest cover.

3. Data

3.6. Topographical data

- Slope, Slope direction, and Digital Elevation Model (DEM)
- **Description:** Slope refers to the steepness of a surface. The slope direction is the direction of the inclined plane. DEM represents the elevation of the terrain above sea level.
- **Source:** EarthEnv project collection
- **Collection technique:** Download the picture.tiff file from the website (https://www.earthenv.org/topography) and use Python code with the rasterio library to extract values from the file using latitude and longitude coordinates as keys.



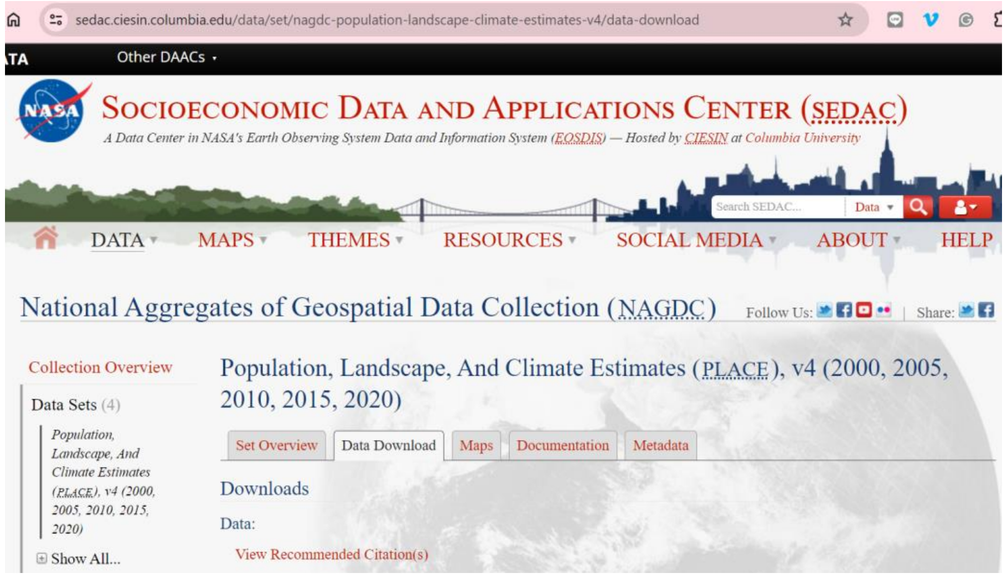


Source: https://www.earthenv.org/topography

3. Data

3.7. Population Density

- **Description:** Number of people per square kilometer.
- **Source:** National Geospatial Data Collection Aggregation
- **Collection technique:** Download the picture.tiff file from the website (https://sedac.ciesin.columbia.edu/data/set/nagdc-population-landscape-climate-estimates-v4/data-download) and use Python code along with the rasterio library to extract specific numeric values from the file using latitude and longitude coordinates as keys.



Source: https://sedac.ciesin.columbia.edu/data/set/nagdc-population-landscape-climate-estimates-v4/data-download

4. Experiment

4.1. Dataset

Dataset 1: (Y1 = Wildfire occurrence)

- **Region:** Chiang Mai (Thailand)
- **Period:** From 2014.01.03 to 2021.12.31
- **Training : Testing =** 80 : 20
- **Total number of observations:** 3 million
- **Number of wildfire observations:** 10 thousand
- The dataset is **oversampled** to balance the dataset.

Dataset 2: (Y2 = Wildfire size, Y3 = Wildfire speed, Y4 = Wildfire duration, Y5 = Wildfire direction)

- **Region:** All regions of Thailand
- **Period:** From 2003.01.13 to 2016.06.11
- **Training : Testing =** 80 : 20
- **Number of observations:** 20 thousand

4. Experiment

4.2. Model

- Each model is run in total 20 times to prevent the randomness of the model.

Logistic Regression

- Class weight (class_weight) = balanced

Random Forest

- Maximum depth (max_depth) = 10, 20, and 30
- Number of trees (n_estimators) = 10, 50, and 100

XGBoost

- Maximum depth (max_depth) = 10, 20, and 30
- Number of trees (n_estimators) = 10, 50, and 100

ANN

- 3 Layers
- Number of input nodes = 32, 64, and 128
- Number of epochs = 10, 50, and 100
- Optimizer = Adam

LSTM

- Number of epochs = 10, 50, and 100
- Window size = 3
- Optimizer = Adam

4. Experiment

4.3. Model result

- **Wildfire occurrence**
- XGBoost has the highest performance.

Model	Accuracy	AUC	F1
Logistic Regression	0.8605 (0.0010)	0.8605 (0.0020)	0.8597 (0.0010)
Random Forest	0.8954 (0.0011)	0.8954 (0.0011)	0.8948 (0.0011)
XGBoost	**0.9403 (0.0007)**	**0.9403 (0.0007)**	**0.9401 (0.0007)**
ANN	0.7010 (0.0034)	0.7170 (0.0023)	0.7002 (0.0003)
LSTM	0.8123 (0.0130)	0.8133 (0.0137)	0.8120 (0.0132)

* Parenthesis shows the standard deviation of each measure of 20 times experiments.

4. Experiment

4.3. Model result

- **Wildfire size**
- LSTM has the highest performance.

Classification

Model	Accuracy	AUC	F1
Logistic Regression	0.7702 (0.0057)	0.6207 (0.0115)	0.7717 (0.0057)
Random Forest	0.7715 (0.0047)	0.6132 (0.0028)	0.7318 (0.0063)
XGBoost	0.7522 (0.0073)	0.6255 (0.0039)	0.7261 (0.0089)
ANN	0.7758 (0.0069)	0.6933 (0.0031)	0.7137 (0.0094)
LSTM	**0.8468 (0.0055)**	**0.8271 (0.0063)**	**0.8467 (0.0055)**

Regression

Model	MAE	RMSE	MSE
Logistic Regression	1.5616 (0.0408)	3.7992 (0.4008)	10.5872 (1.8206)
Random Forest	1.4962 (0.0589)	3.8739 (0.4111)	11.1679 (1.6113)
XGBoost	1.4858 (0.0305)	3.7522 (0.2769)	10.1525 (1.4732)
ANN	1.2013 (0.0498)	3.9055 (0.3596)	11.3765 (1.2272)
LSTM	**0.0307 (0.0091)**	**0.0430 (0.0055)**	**0.0018 (0.0004)**

4. Experiment

4.3. Model result

- **Wildfire speed**
- LSTM has the highest performance.

Classification

Model	Accuracy	AUC	F1
Logistic Regression	0.8733 (0.0056)	0.7582 (0.0161)	0.8733 (0.0056)
Random Forest	0.8741 (0.0040)	0.7513 (0.0028)	0.8200 (0.0062)
XGBoost	0.8683 (0.0041)	0.7551 (0.0019)	0.8216 (0.0056)
ANN	0.8761 (0.0054)	0.7669 (0.0031)	0.8221 (0.0079)
LSTM	**0.9183 (0.0013)**	**0.9081 (0.0014)**	**0.9183 (0.0013)**

Regression

Model	MAE	RMSE	MSE
Logistic Regression	0.2669 (0.0032)	0.4244 (0.0120)	0.0032 (0.0101)
Random Forest	0.2594 (0.0052)	0.4217 (0.0173)	0.1781 (0.0146)
XGBoost	0.2562 (0.0056)	0.4237 (0.0165)	0.1798 (0.0139)
ANN	0.2514 (0.0036)	0.4275 (0.0104)	0.0036 (0.0089)
LSTM	**0.0340 (0.0070)**	**0.0613 (0.0027)**	**0.0038 (0.0003)**

4. Experiment

4.3. Model result

- **Wildfire duration**
- LSTM has the highest performance.

Model	MAE	RMSE	MSE
Logistic Regression	1.6791 (0.0219)	2.6159 (0.0670)	6.8473 (0.3521)
Random Forest	1.6314 (0.0281)	2.5429 (0.0706)	6.4712 (0.3611)
XGBoost	1.6916 (0.0203)	2.6330 (0.0547)	6.9359 (0.2882)
ANN	2.2917 (0.0479)	4.6318 (0.0912)	8.4618 (0.8486)
LSTM	**0.0673** **(0.0037)**	**0.1024** **(0.0014)**	**0.0104** **(0.0002)**

4. Experiment

4.3. Model result

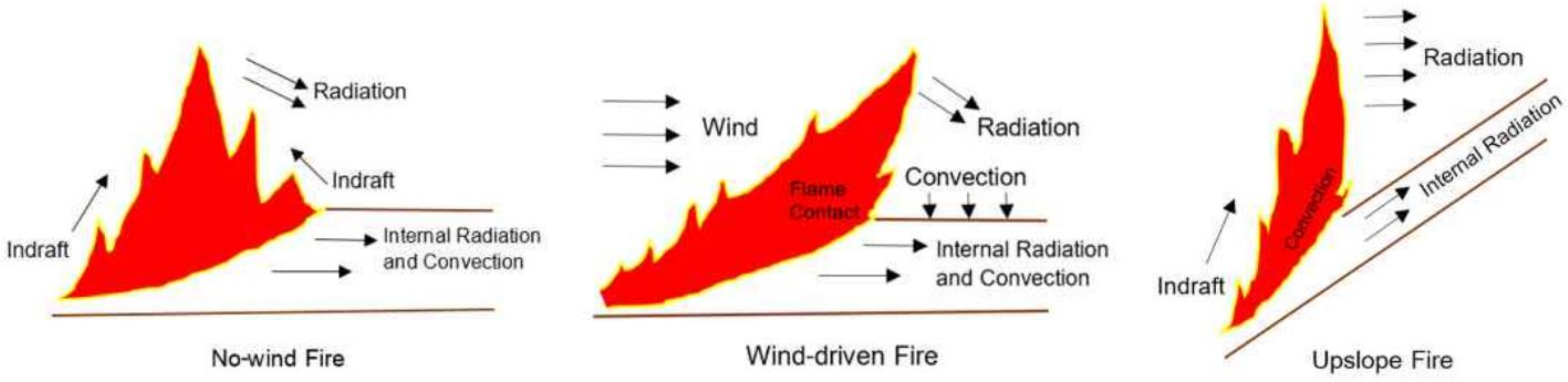
- **Wildfire direction**
- LSTM has the highest performance.

Model	Accuracy	AUC	F1
Logistic Regression	0.6914 (0.0059)	0.7265 (0.0148)	0.6688 (0.0051)
Random Forest	0.6210 (0.0065)	0.6399 (0.0017)	0.6440 (0.0082)
XGBoost	0.6330 (0.0063)	0.6498 (0.0084)	0.6661 (0.0084)
ANN	0.6595 (0.0051)	0.6459 (0.0083)	0.6887 (0.0058)
LSTM	**0.9285** **(0.0017)**	**0.7547** **(0.0109)**	**0.9229** **(0.0025)**

5. Additional analysis

5.1. Interaction variable

- It represents the relationship between **slope direction and wind direction**.
- It could significantly affect wildfire behaviors.
- It was calculated by finding the angle between the wind direction and slope direction using the **cosine function**.



Source: https://www.researchgate.net/figure/3-Differences-in-Fire-Behavior-Caused-by-Wind-and-Slope_fig3_292405151

5. Additional analysis

5.1.1. Interaction variable result

- **Wildfire size**
- LSTM (Interaction) has a better performance than LSTM.

Classification

Model	Accuracy	AUC	F1
LSTM	0.8468 (0.0055)	**0.8271 (0.0063)**	0.8467 (0.0055)
LSTM (Interaction)	**0.8489 (0.0047)**	0.8269 (0.0054)	**0.8467 (0.0047)**

Regression

Model	MAE	RMSE	MSE
LSTM	0.0307 (0.0091)	0.0430 (0.0055)	**0.0018 (0.0004)**
LSTM (Interaction)	**0.0297 (0.0109)**	**0.0428 (0.0067)**	0.0018 (0.0006)

5. Additional analysis

5.1.1. Interaction variable result

- **Wildfire speed**
- LSTM (Interaction) has a better performance than LSTM.

Classification

Model	Accuracy	AUC	F1
LSTM	0.9183 (0.0013)	0.9081 (0.0014)	0.9183 (0.0013)
LSTM (Interaction)	**0.9196 (0.0006)**	**0.9085 (0.0007)**	**0.9186 (0.0006)**

Regression

Model	MAE	RMSE	MSE
LSTM	0.0340 (0.0070)	0.0613 (0.0027)	0.0038 (0.0003)
LSTM (Interaction)	**0.0323 (0.0048)**	**0.0610 (0.0033)**	**0.0037 (0.0004)**

5. Additional analysis

5.1.1. Interaction variable result

- **Wildfire duration**
- LSTM (Interaction) has a better performance than LSTM.

Model	MAE	RMSE	MSE
LSTM	0.0673 (0.0037)	**0.1024 (0.0014)**	**0.0104 (0.0002)**
LSTM (Interaction)	**0.0654 (0.0025)**	0.1033 (0.0016)	0.0106 (0.0003)

5. Additional analysis

5.1.1. Interaction variable result

- **Wildfire direction**
- LSTM (Interaction) has a better performance than LSTM.

Model	Accuracy	AUC	F1
LSTM	0.9285 (0.0017)	0.7547 (0.0109)	0.9229 (0.0025)
LSTM (Interaction)	**0.9288 (0.0020)**	**0.7578 (0.0124)**	**0.9235 (0.0029)**

5. Additional analysis

5.2. Prediction using previous day factors

- Wildfires can occur at any time of day - morning, afternoon, or evening.
- If a fire starts in the morning, the **current day's factors would not be applicable**.
- People may want to know the likelihood of a wildfire of the next day, especially in the evening, so they can be alert when the likelihood is high.
- This model was proposed that uses **factors only from previous days, not the current day**.

$$Y_t = F(X_{1,t-1}, X_{1,t-2}, X_{1,t-3}, \ldots, X_{n,t-1}, X_{n,t-2}, X_{n,t-3})$$

5. Additional analysis

5.2.1. Prediction using previous day factors result

- **Wildfire occurrence**
- XGBoost has the highest performance.
- LR, RF, XGBoost, and ANN had a slightly lower accuracy for this additional morning fire prediction compared to the original midday fire prediction model.
- LSTM model had higher accuracy when predicting morning fires using only prior days' data.

Model	Accuracy	AUC	F1
Logistic Regression	0.8430 (0.0023)	0.8575 (0.0020)	0.8415 (0.0024)
Random Forest	0.8888 (0.0066)	0.8903 (0.0033)	0.8887 (0.0053)
XGBoost	**0.9370 (0.0030)**	**0.9391 (0.0021)**	**0.9371 (0.0025)**
ANN	0.7004 (0.0123)	0.7064 (0.0230)	0.7074 (0.0134)
LSTM	0.8239 (0.0193)	0.8244 (0.0210)	0.8236 (0.0198)

6. Conclusion

- This study conducted developing **wildfire prediction models to predict wildfire occurrence, wildfire size, wildfire speed, wildfire duration, and wildfire direction.**
- The performance of **wildfire behavior prediction models** (wildfire size, wildfire speed, wildfire duration, and wildfire direction) showed a **significant improvement with the LSTM model**.
- The LSTM's sequential modeling capabilities made it suitable for time-series forecasting.
- The prediction of **wildfire occurrence** in **XGBoost model is the best** among the other models.
- XGBoost outperforms LSTMs because it is better suited to the nature of the data and the specific requirements of the prediction task.
- This study creates new dataset by **adding more variables**, which are slope direction, wind direction, day of the week, and whether it was a weekend or not.
- This study developed LSTM model, which is **the first study using LSTM model to predict wildfire in Thailand**.

THANK YOU

Appendix

- **Dependent variables**
- **Y1 = Wildfire occurrence, Y2 = Wildfire size, Y3 = Wildfire speed, Y4 = Wildfire duration, Y5 = Wildfire direction**

Attribute	Description	Units/Format	Type
Status	Wildfire or no wildfire status (0 = no wildfire, 1 = wildfire)		Binary
Size	Total burned size per wildfire (0.21 – 97.75)	Km^2	Continous
Status (Size)	burned size per wildfire (0 = small wildfire (<1), 1 = medium wildfire (1 to 3), 2 = large wildfire (>3))		Categorical
Speed	Average wildfire speed (0.25-6.6)	Km/day	Continous
Status (Speed)	Speed per wildfire (0 = low speed (<1), 1 = medium speed (1-3), 2 = high speed (>3))		Categorical
Duration	Wildfire Duration (1-38)	Days	Continous
Direction	The predominant direction of wildfire spread is provided only for wildfires that have occurred several days. (0) indicates no data. (North, Northeast, East, Southeast, South, Southwest, West, Northwest)	Cardinal direction	Categorical

Appendix

- **Independent variables**

Attribute	Description	Units/Format	Type
Temperature	Today's average temperature (9.9-41)	°C	Continous
Dew_Point	Today's average dew point (1-27)	°C	Continous
Humidity	Today's average humidity (11.7-100)	%	Continous
Wind_Speed	Today's average wind speed (0.9-33.2)	km/h	Continous
Pressure	Today's average pressure (951.7-1020.1)	hPa	Continous
Precipitation	Today's average precipitation (0-51.1)	mm	Continous
Temperature_yesterday	Yesterday's average temperature (12.2-42)	°C	Continous
Dew_Point_yesterday	Yesterday's average dew point (1-27)	°C	Continous
Humidity_yesterday	Yesterday's average humidity (11-94.7)	%	Continous
Wind_Speed_yesterday	yesterday's average wind speed (1-61.2)	km/h	Continous
Pressure_yesterday	Yesterday's average pressure (955.2-1020.6)	hPa	Continous
Precipitation_yesterday	Yesterday's average precipitation (0-47)	mm	Continous

Appendix

- **Independent variables**

Attribute	Description	Units/Format	Type
Temperature_two_days_before	Average temperature two days ago (13.2-41)	°C	Continous
Dew_Point_two_days_before	Average dew point two days ago (0.4-26.7)	°C	Continous
Humidity_two_days_before	Average humidity two days ago (11-100)	%	Continous
Wind_Speed_two_days_before	Average wind speed two days ago (0.5-41.3)	km/h	Continous
Pressure_two_days_before	Average pressure two days ago (957.1-1019)	hPa	Continous
Precipitation_two_days_before	Average precipitation two days ago (0-52.8)	mm	Continous
Temperature_three_days_before	Average temperature three days ago (13.2-41)	°C	Continous
Dew_Point_three_days_before	Average dew point three days ago (0.4-26.7)	°C	Continous
Humidity_three_days_before	Average humidity three days ago (11-100)	%	Continous
Wind_Speed_three_days_before	Average wind speed three days ago (0.5-41.3)	km/h	Continous

Appendix

- **Independent variables**

Attribute	Description	Units/Format	Type
Pressure_three_days_before	Average pressure three days ago (957.1-1019)	hPa	Continous
Precipitation_three_days_before	Average precipitation three days ago (0-52.8)	mm	Continous
Wind_direction	Average wind direction (CALM, VAR, N, NNE, NE, ENE, E, ESE, SE, SSE, S, SSW, SW, WSW, W, WNW, NW, NNW)	Cardinal direction	Categorical
DEM	Digital elevation model (2.25-2343)	Meter	Continous
DC	Drought code (5-163)		Continous
DMC	Duff moisture code (0.25-372)		Continous
FFMC	Fine fuel moisture code (16-96)		Continous
Slope	Incline of the location (0-40)	Degree	Continous
Slope_direction	Incline direction of the location (N, NE, E, SE, S, SW, W, NW)	Cardinal direction	Categorical
Population_density	Population density of the area (0-2040)	Number of people per square kilometer	Continous

Appendix

- **Independent variables**

Attribute	Description	Units/Format	Type
Interaction	Relationship between wind direction and slope direction ((-1) – 1)		Continous
LU_CODE	Land cover type in the wildfire occurrence dataset. (Protected forest, Reserve forest)		Categorical
Landcover	Land cover type of wildfire size, wildfire rate, wildfire duration, and fire direction dataset (Evergreen forest, Deciduous forest, Mixed forest, Closed shrub, Open shrub, Woody savannah, Savannah)		Categorical
Month	Months in a year (January, February, March, April, May, June, July, August, September, October, November, December)		Categorical
Day_of_a_week	Days (Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday)		Categorical
Is_weekend	Weekend (Yes, No)		Categorical