Model Classification of Fire Weather Index using the SVM-FF Method on Forest Fire in North Sumatra, Indonesia

Darwis Robinson Manalu, Opim Salim Sitompul, Herman Mawengkang, Muhammad Zarlis
Doctoral Study Program (S3) in Computer Science,
Faculty of Computer Science and Information Technology,
Universitas Sumatera Utara, Medan, North Sumatera, Indonesia

***Abstract*—As a tropical country, Indonesia is situated in Southeast Asia nation has vast forests. Forest fire occur busy vary due to land conditions and forest conditions in drought season. The indicator used mitigated potential forest fire is to study the indicator behavior of the fire weather index (FWI). The data is gathered from the observation station in north Sumatra province, computation and estimation FWI by Canadian Forest Fire Weather Index based on the data gathered. It is found that there is gathered outlier data. to hope will it, it is necessary to conduct classification and predict this of the dataset by machine learning approach using Support Vector Machine Forest Fire (SVM-FF), which is a further development of the previous models, known as the c-SVM and v-SVM. This method includes a balancing parameter by determining the lower and upper limits of a support vector. Furthermore, it allowed the balancing parameter value to be negative. The results showed that the classification of FWI was at low, medium, high, and extreme levels. The low FWI value has an average of 0.5 which is in the 0 to 1 interval. There was an increase in the model's accuracy and performance from its predecessor, which include the c-SVM and v-SVM with respective values of 0.96 and 0.89. Meanwhile, it was observed that with the SVM-FF model, the accuracy was quite better with a value of 0.99, indicating that it is useful as an alternative to classify and predict forest fires.**

Keywords—Fire weather index; forest fire; support vector machine; SVM-FF model

I. Introduction

Forest fire mitigation has been a priority for everyone in order to avoid increasing damage to nature [1]. Several methods applied by forest managers in this process include creating awareness on the prohibition of forest burning and monitoring of fire-prone areas, both of which are man-made and natural factors [2]. One of the data sources used to determine forest fires was the distribution of hotspots [3][4].

A previous research on machine learning that employed a random forest model [5] includes performance measurement of Forest Fire Prediction [6][7]. Another one was the application of machine learning for classification and prediction using fire weather index data[8]. Classification by using multiple variables is a frequently encountered realted to data mining problem[9]. The components of the fire weather index in the meteorological data [10] are temperature, rain, wind, air humidity, and other supporting elements for calculating the Fire Weather Index (FWI) [11] daily in forest areas having fire outbreak potential.

This data source was processed to determine the distribution of FWI in North Sumatra Province. However, the observation data used still needs to be pre-processed to avoid missing value or outlier data [12]. It is important to note that one of the characteristics of meteorological and weather data is the outlier, which are the emergence of extreme parameter values [13]. This makes it to have an impact on model misclassification, biases in parameter estimations, incorrect results, and imperfect forecasts [12] [14], which are classified as a loss function using the Support Vector Machine (SVM) method in machine learning [15][16][17].

The research objective was to classify and predict forest fires in North Sumatra using the fire weather index behavior[18]. The benefit of this research is to provide information on fire weather index values as a reference in predicting the potential for forest fires in an area [19], especially in North Sumatra, and early information on disaster mitigation in the forest fire sector. Research urgency is that the classification of potential forest fires [20] should not be based on the distribution of hotspots alone but needs to be reviewed from the behavior of the fire weather index with the Support Vector Machine Forest Fire (SVM-FF) model approach in reducing and mitigating forest fire disasters in the province of North Sumatra.

II. Related Works

This research dealing with forest fire classification covers the way how to optimize parameters at SVM to reduce misclassification [21]. With the Weather Research and Forecasting (WRF) mesoscale model method, FWI classification and mapping were carried out in Greece [22].

Further research uses a multi-factor forest fire prediction model with a machine learning approach using the random forest model method which aims to determine the highest incidence in China [23]. The determination of hotspot points that are also used the classification algorithm are C5.0 and Random Forest producing rules-based model[24][25] namely Forest Fire Detection carried out by using an application to determine hotspots in forested areas using Landsat 8 and Band ten satellite images or thermal bands having information on temperature [26][27], Predicting Rainfall from Weather

Observations using SVM Approach to identify The Parameter of Fuel Moisture as Fire Weather Index [28], Predicting fire-prone areas with machine learning techniques [29], and Researching into forest risk assessment system in China [30]. The study presents the classification and prediction of forest fires.

III. Fire Weather Index

Obtaining the value of Fire Weather Index (FWI) requires a process that includes meteorological elements. The data collection process starts from observation to data processing [31]. The steps taken include calculating the value of Fire Weather Observations, Fire Behaviors Indices, Fine Fuel Moisture Code (FFMC) [32], Duff Moisture Code (DMC), Drought code (DC), Initial Spread Index (ISI), Buildup Index (BUI), and Fire Weather Index (FWI) [29].

This becomes the basis for processing observation data to produce a fire weather index value in numerical form, which is based on the ISI and BUI, utilized as a general fire hazard index for forest areas [33]. The following equations are employed in the calculation of the FWI value.

$$f(D) = 0.626U^{0.0809} + 2, \quad U \leq 80 \qquad (1)$$

$$f(D) = \frac{1000}{(25+108.64\, e^{-0.023U})} + 2, \quad U > 80 \quad (2)$$

$$B = 0.1\, R\, f(D) \qquad (3)$$

$$\ln S = 2{,}72\, (0.434 \ln B)^{0.647}, \quad B > 1 \; (4)$$

$$S = B,\ Bz \qquad (5)$$

Where:

f(D) is the Function of drought (*drought*)

U represents the BUI value

R denotes the ISI value

B the FWI (*intermediate form*)

S represents the FWI (*final form*)

The calculations of these components are based on daily meteorological data observations such as temperature, relative humidity, wind speed, and 24-hour rainfall [32]. The three components of the FWI system provide a numerical rating of the forest's relative fire potential, including Fire Weather Observations, Fuel Moisture Code, and Fire Behaviors Indices. Fig. 1 shows the components of the FWI System.

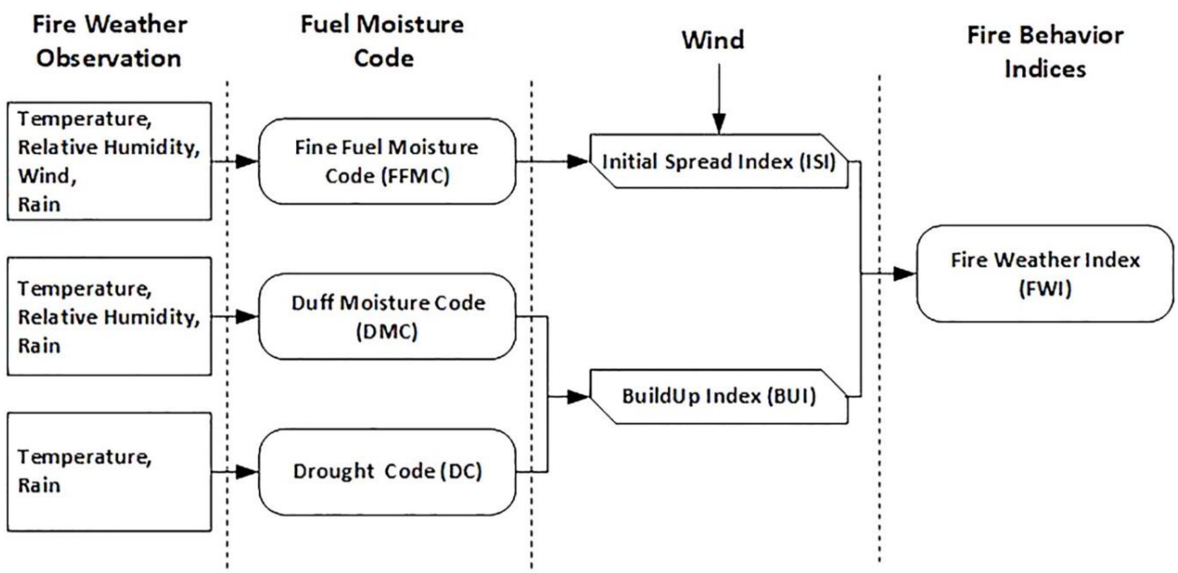

Fig. 1. FWI schematic [34].

One of the classification tasks is building a prediction engine with a good degree of accuracy capable of generalizing [35]. To achieve this purpose, SVM is used to minimize the objective function that contains a loss function [36], with the aim of finding the function f(x) as a hyper-plane and making the error (ε) as minor possible [21]. It is important to note that the current studies on SVM generally is focused on improving and formulating loss functions that eventually produce many variants. For example, the C-SVM [37] applied the surrogate function formulation by adding parameters $C \in (0, \infty)$ to the loss function in Eq. (6).

$$min_{w,b}\ \frac{C}{2}\|w\|^2 + \sum_{i\in I}[-y_i\,(w^T x_i + b) + 1]^+ \qquad (6)$$

The v-SVM model in [38] added a parameter to increase the cardinality of the data as expressed in Eq. (7).

$$min_{w,b,p}\ v + \frac{1}{|I|}\sum_{i\in I}[-y_i\,(w^T x_i + b) + 1]^+ \qquad (7)$$

In many studies, both C-SVM and v-SVM produced optimal solutions because the parameter v was able to determine the lower and upper bounds of a support vector.

IV. Methodology

The meteorological data method was used in this research and processing was performed before pre-processing the observation data from all Meteorological, Climatological, and Geophysical Agency (BMKG) stations [39]. The details of the method are shown in Fig. 2.

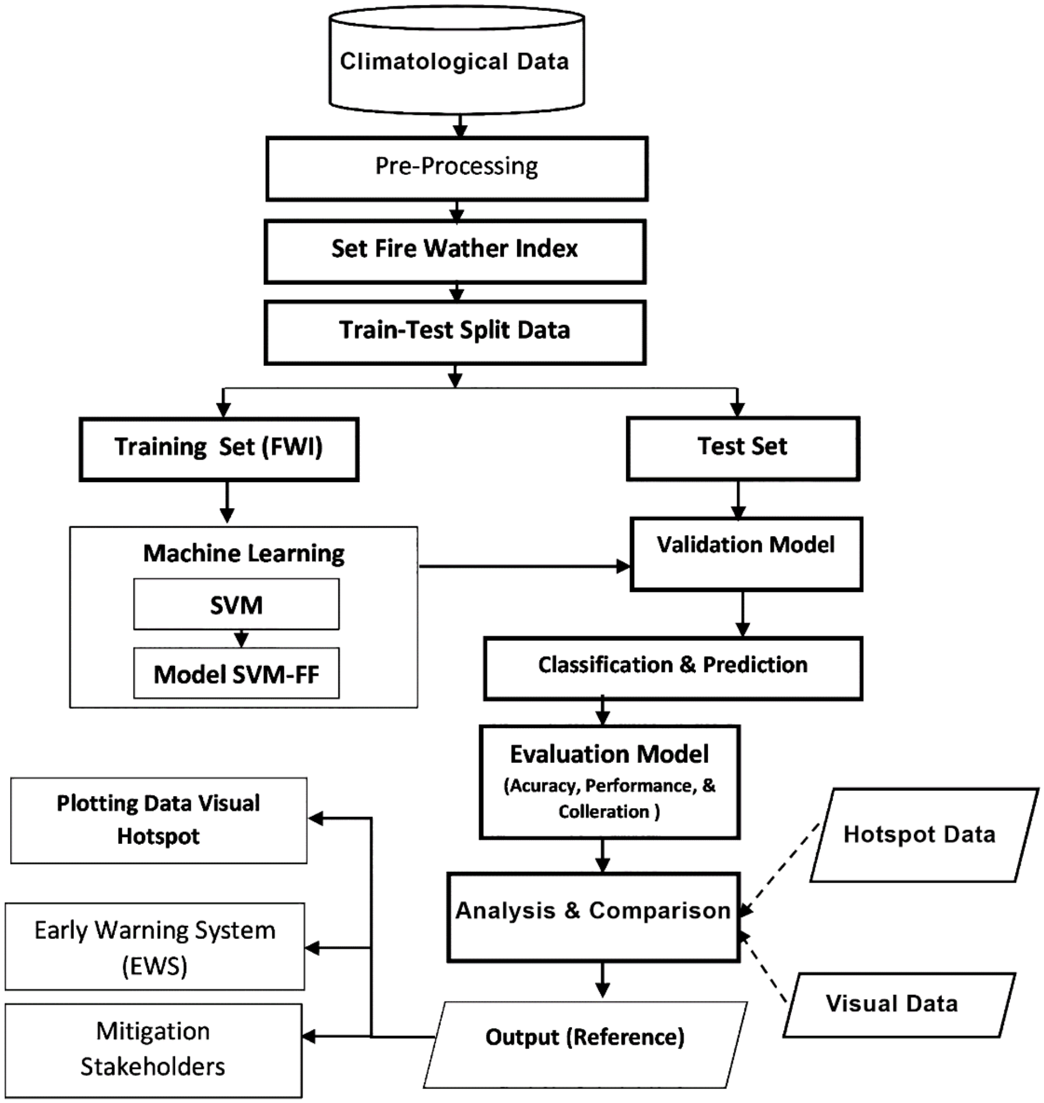

Fig. 2. Research framework.

The SVM-FF focused on increasing negative outlier data in order to get better performance, but when the parameter v in the v-SVM shown in Eq. (7) has an extreme value, the classification model produced poor performance. First, it assigned a large value to the parameter v, which encouraged overfitting [39] and makes the model look too good but not optimal to generalize from the dataset.

Second, it produced a small value of w=0 and b=0 in Eq. (3) to (5) above, which results in a poor solution. It was observed that the c-SVM and v-SVM models produced optimal solutions because the parameter v was able to determine the lower and upper bounds of a support vector. Based on the above considerations, the proposed solution to force the value of the parameter w in Eq. (7) was to set the value of the parameter p to be negative. This is consistent with [39] that applied optimization to produce negative values as expressed in Eq. (8).

$$C = min_{w,b,p}\,\{-vp + \frac{1}{|I|}\sum_{i\in I}[-y_i\,(w^T x_i + b) + 1]^+\} \quad (8)$$

w: weight

b: biases always positive

p: balancing parameter

y: target/class/category

x: feature data

t: time epoch

i: index

A. SVM-FF Algorithm

The SVM-FF model is based on Eq. (7) with the following steps:

Algorithm 1: SVM-FF model

Initiation value $\nu \in (0, 1)$, value p, w, b
Minimizing the loss function with equation (8):

$$C = min_{w,b,p} \{-\nu p + \frac{1}{|I|}\sum_{i \in I}[-y_i\,(w^T x_i\; + b) + 1]^+\}$$

If the value is infinite, the algorithm stops.
Otherwise, the optimal solution is (w_i, b_i)

If value $C_{i+1} = C_i$ the algorithm stops.
Determine as much C_{i+1} sample data from data x_i

Set index to $I_i + 1$
$i = i + 1$

B. Model Validation

The performance test of the SVM-FF model was based on several criteria, such as the average error value using the Receiver Over Characteristic (ROC) curve function [40].

C. Error Plot

An error plot is a graphical representation that shows the range of the actual value to those predicted by the model. When there are many N data observed, then the average of the data is described in Eq. (9) [39].

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (9)$$

According to Eq. (9), the error value is defined as the square root of the difference between the data x_i and the average data x as expressed in Eq. (10).

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (10)$$

D. Receiver Operating Characteristic

The most recommended test criterion for multi-class classification was Receiver Operating Characteristic (ROC) analysis [41]. The ROC curve is capable of displaying a sensitivity plot in the y-axis, and 1-specificity in the x-axis. Fig. 3 shows a hypothetical ROC curve that represents diagnostic accuracy. In line A, the curve showed 100% accuracy or 1.0, while in line B, it represents 85% accuracy or 0.85, and curve C depicted 50% accuracy or 0.5.

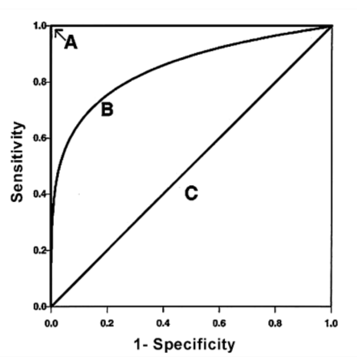

Fig. 3. ROC curve.

V. Discussion

A. BMKG Dataset Exploration

The FWI is a multi-dimensional dataset that consists of 12,897 raw data with ten features and one class. The description of each feature is described in Table I below.

TABLE I. FWI DATASET FEATURES

Feature	Description	Value (max, min, average)
Tn	Real data, minimum temperature	33, 1, 23.25
Tx	Real data, maximum temperature	233, 3, 31.16
Tavg	Real data, mean temperature	31.6, 18.6, 26.9
RH_avg	Real data, average humidity	121, 7, 85.62
RR	Real data, rainfall	8888, 0, 10.23
ss	Real data, duration of sunshine	23, 0, 4.50
ff_x	Real data, maximum wind speed	45, 0, 4.40
ddd_x	Real data, wind direction at maximum speed	2860, 0, 191.62
ff_avg	Real data, average wind speed	15, 0, 1.38
S	Nominal data, dataset class target value	(1,2,3,4)

B. Data Preparation

Climatological dataset contains about 28% of missing value. Data pre-processing involves a series of data preparation process used to handle missing value. Columns in the dataset which are having missing values replaced with the mean of remaining values in the column [42].

C. Condition of Data Outliers

One of the characteristics of disaster, weather, and climate is outlier data. It has been observed that the existence of this data increased as a result of extreme parameters, such as temperature, wind speed, etc. The BMKG data applied in this research was not exempted from the outlier problems. The existence of outliers in data features are shown in Fig. 4 and 5.

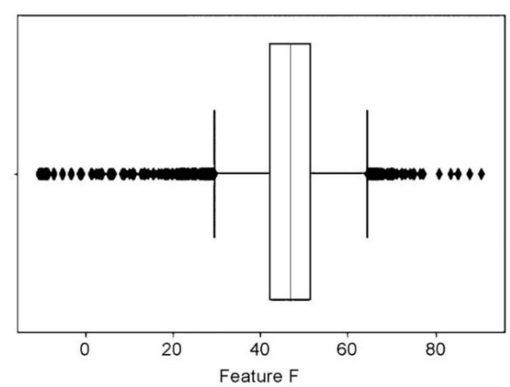

Fig. 4. Outliers on the RH_avg feature (Average Humidity).

Fig. 4 shows the condition of the dataset outliers, specifically for the RH_avg feature or mean humidity. It was observed that some values deviate significantly from the median or the middle value of the data (blue). Furthermore, the outlier values move to the right and left of the median.

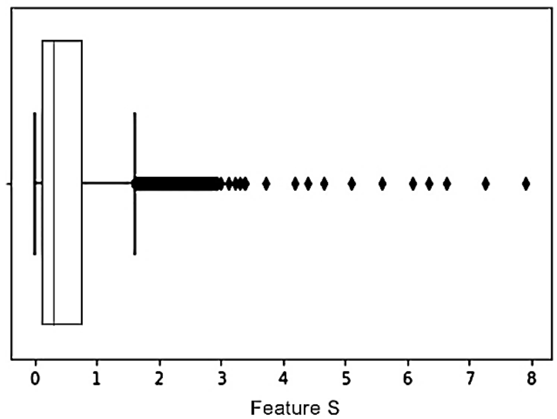

Fig. 5. Outliers on the ff_avg Feature (Average Wind Speed).

The condition of the outlier dataset in Fig. 5 for other features was ff_avg or average wind speed. Furthermore, the outliers are to the right of the data or values above the median. It is important to note that the first step in the classification process with SVM-FF are parameters initiation that include the conventional SVM, namely the kernel type parameter, *w* value, gamma value, and *r* coefficient. Other parameters adapted from the SVM-FF model are the values of *v* and *p* which have been described in Eq. (8). The description of each feature is described in Table II below.

TABLE II. INITIALIZATION OF SVM-FF PARAMETERS

No	Parameter	Default value	Model Type
1	Kernel	RBF	SVM
2	w	1	SVM
3	Gamma	1	SVM
4	r Coefficient	0	SVM
5	v	(0,1)	SVM-FF
6	p	(0,1)	SVM-FF

The pre-processed FWI data consists of 12,800 raw data, which consist of low, medium, high, and extreme classes. Furthermore, the number of raw data per class was 6,126, 3,394, 3,150 and 130 data, respectively and the graph of these distributions is shown in Fig. 6.

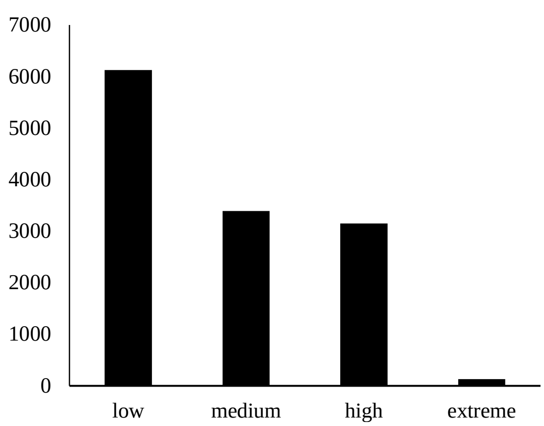

Fig. 6. FWI dataset class distribution graph.

It is important to note that the data validation scheme is based on the k-Fold cross-validation method with k=10. This method randomly divided the dataset into ten sub-datasets. In each fold, the training data was nine sub-datasets, while the testing data was one sub-dataset. The distribution scheme of these datasets is shown in Table III below, is showing that in the first fold, the amount of training data was 11,520 and the testing data was 1,280. The same number of data applies from the Second fold through to the tenth fold.

In the testing phase, the SVM-FF model was applied to classify the dataset based on the attribute values of the FWI dataset. The dataset classification results in the first fold (fold-0) are shown in Table III. A total of 25 raw data are taken in both upper data (1 to 25) and lower data (1,256 to 1,280) out of 1.280 raw data in fold-0, respectively.

The data with serial number 1 has an actual class of "medium", and the classification result is "high", making, the classification status was misclassified. Furthermore, the data with sequence number 2 has a “low” actual class and classification result of "low" respectively, which means a correct classification. It can be concluded that in the fold-0, out of 1,280 raw data, 141 data are misclassified and 1,139 with correct status, indicating that the accuracy of data classification in fold-0 is 0.890 globally.

TABLE III. RESULTS OF DATA CLASSIFICATION USING THE SVM-FF MODEL

No	Actual	Classification	Status
1	medium	high	misclassified
2	low	low	correct
3	high	high	correct
4	medium	medium	correct
5	low	low	correct
6	low	low	correct
7	low	low	misclassified
8	medium	low	correct
9	high	high	correct
10	medium	medium	correct
1,272	low	low	correct
1,273	high	high	correct
1,274	extreme	extreme	correct
1,275	high	high	correct
1,276	extreme	high	misclassified
1,277	high	high	correct
1,278	high	high	correct
1,279	extreme	extreme	correct
1,280	medium	low	correct

Similarly, the accuracy of each fold in both the testing and training stages is described in Table IV. It was observed that in fold-0, the training and testing accuracies were 0.978 and 0.890, respectively. When the overall fold was tested with 10 folds, the average training and testing accuracies were 0.983 and 0.906, respectively.

TABLE IV. AVERAGE ACCURACY OF TRAINING AND TESTING

Fold	Testing	Training
Fold 0	0.890	0.978
Fold 1	0.919	0.978
Fold 2	0.957	0.981
Fold 3	0.906	0.987
Fold 4	0.919	0.982
Fold 5	0.894	0.981
Fold 6	0.904	0.982
Fold 7	0.904	0.988
Fold 8	0.879	0.980
Fold 9	0.879	0.988
Correctly classified	0.906	0.983
Global Classification Error	0.154	0.064
Stddev Global Classification Error	0.020	0.003

In addition to the analysis related to the testing phase, further evaluation was performed to determine the performance of each fold classification per class in the training and testing phases. The results are shown in Table V and Table VI, in which the SVM-FF model has the highest accuracy for the "high" class classification and the lowest results for the "medium".

TABLE V. RESULTS OF PERFORMANCE TEST PER FOLD PER CLASS IN THE TRAINING PHASE

	Low	Medium	High	Extreme
Fold0	1.000	0.887	1.000	0.843
Fold1	1.000	0.867	1.000	0.874
Fold2	0.994	0.908	1.000	0.859
Fold3	1.000	0.887	0.995	0.864
Fold4	0.994	0.882	1.000	0.853
Fold5	0.994	0.892	0.995	0.858
Fold6	0.994	0.893	1.000	0.859
Fold7	1.000	0.908	1.000	0.859
Fold8	1.000	0.872	1.000	0.864
Fold9	1.000	0.897	1.000	0.869
Average	0.998	0.889	0.999	0.860

TABLE VI. RESULTS OF PERFORMANCE TEST PER FOLD PER CLASS IN THE TESTING PHASE

	Low	Medium	High	Extreme
Fold-0	0.880	0.838	0.965	0.692
Fold-1	0.950	0.727	1.000	0.905
Fold-2	0.950	0.818	1.000	0.619
Fold-3	0.950	0.864	0.909	0.714
Fold-4	1.000	0.727	0.952	0.682
Fold-5	0.950	0.636	0.952	0.818
Fold-6	0.950	0.714	1.000	0.714
Fold-7	0.950	0.810	0.955	0.667
Fold-8	0.950	0.714	0.955	0.667
Fold-9	1.000	0.545	0.955	0.810
Average	0.953	0.739	0.964	0.729

D. SVMFF Model Performance

The performance test of the SVM-FF model is based on several criteria, which includes the average value of the error and the Receiver Over Characteristic (ROC) Curve.

E. Error Graph

The error plot of a graph property represents variations in the data, which is due to the errors or uncertainty of a model in the form of visualization of vertical lines in the error point. The benchmark for determining the error plot was Standard Deviation, and the graph for that of SVM-FF model was visualized in Fig. 7 and Fig. 8.

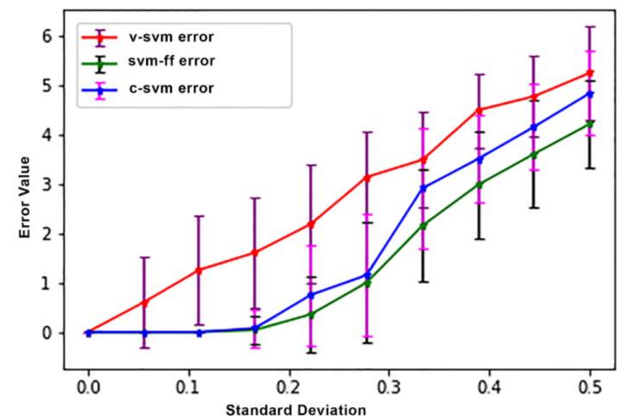

Fig. 7. C-SVM, v-SVM, and SVM-FF model error graphs.

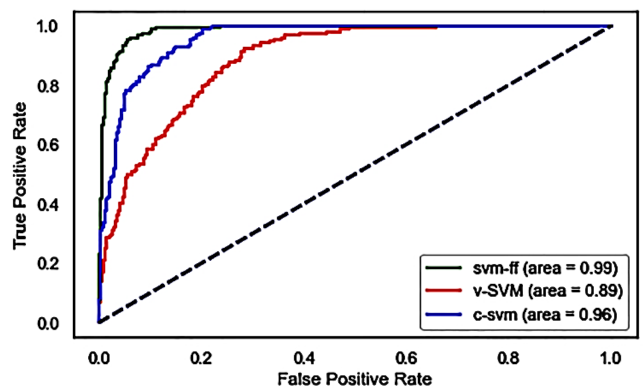

Fig. 8. ROC Comparison graph between SVM-FF, v-SVM, and c-SVM.

It was observed that the error graph or error plot for the SVM-FF model was lower or better compared to the other two models, which include c-SVM and v-SVM.

F. ROC Curve

The ROC curve shows the probability or accuracy of the model in graphical form. According to the test results, the SVMFF model has a better model accuracy in comparison with the other two SVM.

The ROC graph above shows the accuracy and performance of the model, in which the c-SVM and v-SVM models produced accuracies of 0.96 0.89, respectively. When the SVM-FF model was used, the accuracy was improved with a value of 0.99.

G. Decision Region Boundary (DRB)

A two-dimensional graph represents the correspondence of multiple data feature to classes. It is important to note that the SVM-FF method has three types of kernels, which include linear, sigmoid, and Radial Basis Function (RBF). The application of these three kernels affected the shape of the SVM-FF hyperplane. The SVM-FF DRB model is shown in Fig. 9, 10, and 11.

There are four classes of fire weather index datasets, comprising low, medium, high, and extreme respectively indicated with blue, orange, green, and red. The classification was performed in the SVM-FF by placing data features in different regions. It was observed that when the linear kernels were used, there were still deviations in the classification areas, particularly in the extreme classes, which are marked with red circles.

The data area also appears to have been classified, but some data points still occupied different positions. For example, the maximum margin is still very close to the hyperplane in the form of a straight line. Meanwhile, in the data section with a yellow triangle symbol, some still occupy the blue and green areas. This shows that the accuracy with the kernel liner was not optimal as expected.

Furthermore, with the use of the sigmoid kernel, the position of the hyperplane was observed after the data points. The sigmoid line shape in the green round category is seen in Fig. 11. Based on the three figures and the classification results using the kernel, the RBF type presented a better class distinction when compared to linear or sigmoid kernels.

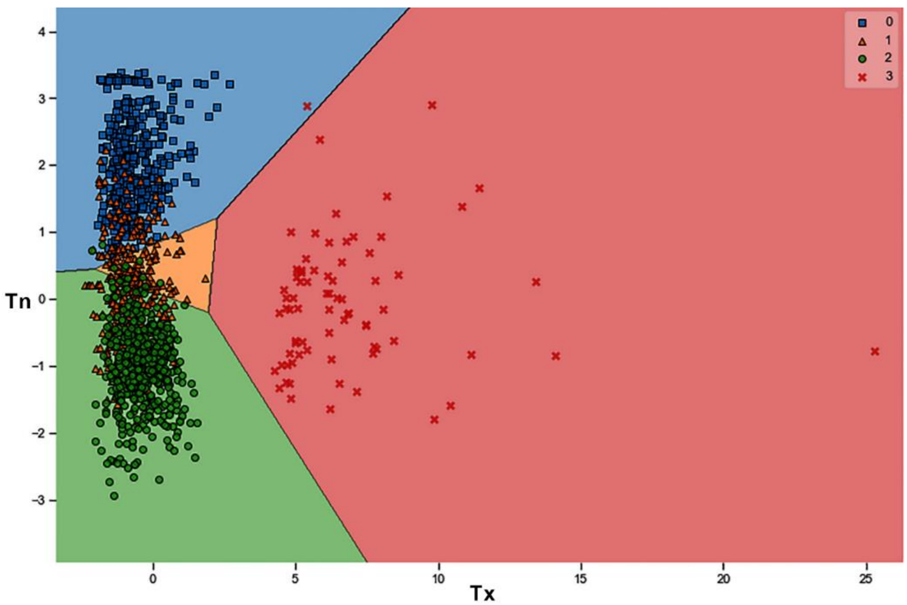

Fig. 9. DRB with linear kernel.

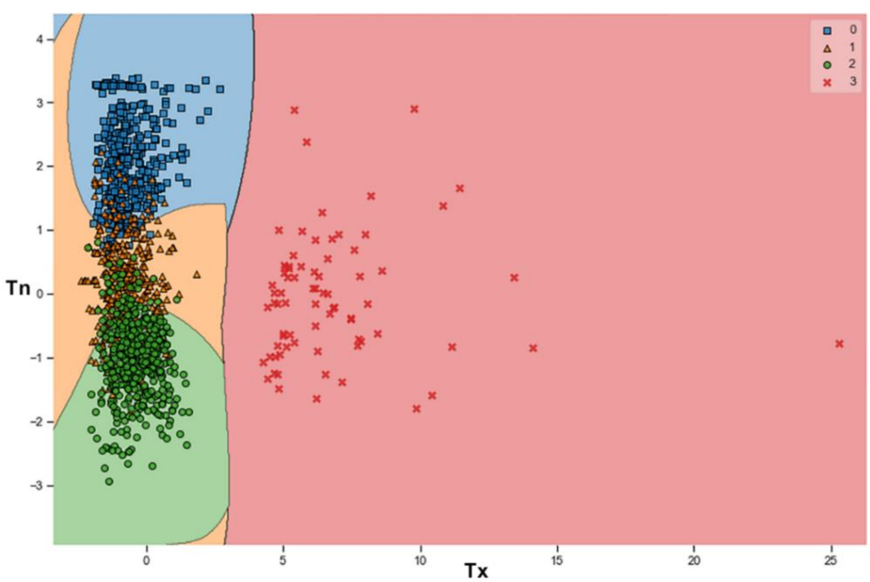

Fig. 10. DRB with RBF kernel.

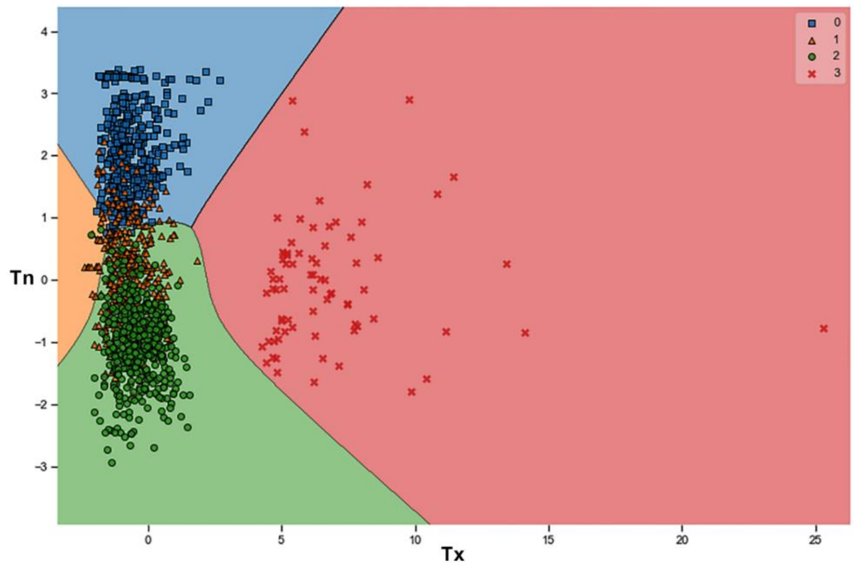

Fig. 11. DRB with sigmoid kernel.

H. Model Significance Test

A statistical significance test approach was used to test the significance of the models as shown in Table VII. The Wilcoxon test was applied at a significance level of 95% or an error rate of 5% (0.05), and was declared to be significantly different when the p-value is < 0.05.

TABLE VII. THE WILCOXON TEST RESULTS OF SVM-FF MODEL

VS	R^+	R^-	Exact P-value	Asymptotic P-value
v-svm	55.0	0.0	0.00195	0.004317
c-svm	55.0	0.0	0.00195	0.004317

According to Table VII, the p-value of SVM-FF VS v-SVM was 0.00195 < 0.05, indicating that the accuracy of SVM-FF differs significantly from the v-SVM. When the p-value of SVM-FF VS c-SVM was 0.00195 < 0.05, the SVM-FF accuracy was observed to differ significantly from the c-SVM.

VI. CONCLUSION

In conclusion, the testing results when the SVM-FF model was used for FWI conditions in North Sumatra Province provided four classifications, namely low, medium, high, and extreme, indicated as blue, orange, green, and red zones, respectively. From 2017 to 2020, FWI conditions have been classified as a low-level/blue zone with an average value of 0.5 per day, which is on a scale of 0 and 1.

This classification was conducted using all observational datasets by entering outlier data to determine the performance optimization of the SVM-FF model and was compared to the result of its predecessor. It was observed that the c-SVM and v-SVM models produced accuracies of 0.96 and 0.89, respectively. Meanwhile, the SVM model -FF was able to improve the accuracy with a value of 0.99, indicating that it was able to work more optimally.

The model's performance when several kernel functions were utilized showed that the Radial Basis Function kernel provided classification results with a better hyper-plane position compared to the maximum position of the data margin.

REFERENCES

[1] M. B. R. Prayoga and R. H. Koestoer, "Improving Forest Fire Mitigation in Indonesia: A Lesson from Canada," J. Wil. dan Lingkung., vol. 9, no. 3, pp. 293–305, 2021, doi: 10.14710/jwl.9.3.293-305.

[2] M. Batista, B. Oliveira, P. Chaves, J. C. Ferreira, and T. Brandao, "Improved Real-time Wildfire Detection using a Surveillance System," Lect. Notes Eng. Comput. Sci., vol. 2240, pp. 520–526, 2019.

[3] Z. Muchamad and M. Emanuel, "Classification of Hotspots Causing Forest and Land Fires Using the Naive Bayes Algorithm," pp. 555–567, 2022.

[4] M. Zainul and E. Minggu, "Classification of Hotspots Causing Forest and Land Fires Using the Naive Bayes Algorithm," Interdiscip. Soc. Stud., vol. 1, no. 5, pp. 555–567, 2022, doi: 10.55324/iss.v1i5.62.

[5] A. Apostolakis, S. Girtsou, C. Kontoes, I. Papoutsis, and M. Tsoutsos, Implementation of a Random Forest Classifier to Examine Wildfire Predictive Modelling in Greece Using Diachronically Collected Fire Occurrence and Fire Mapping Data, vol. 12573 LNCS, no. March. Springer International Publishing, 2021. doi: 10.1007/978-3-030-67835-7_27.

[6] Singh, K. P. Neethu, K. Madhurekaa, A. Harita, and P. Mohan, "Parallel SVM model for forest fire prediction," Soft Comput. Lett., vol. 3, no. July, p. 100014, 2021, doi: 10.1016/j.socl.2021.100014.

[7] H. Y. Lin, "Effective feature selection for multi-class classification models," Lect. Notes Eng. Comput. Sci., vol. 3 LNECS, pp. 1474–1479, 2013.

[8] H. Miyajima, H. Miyajima, and N. Shiratori, "Fast and secure edge-computing algorithms for classification problems," IAENG Int. J. Comput. Sci., vol. 46, no. 4, pp. 1–6, 2019.

[9] H. Y. Lin and Y. H. Lai, "Construction and evaluation of a robust classification model for multi-objective problems," IMECS 2011 - Int. MultiConference Eng. Comput. Sci. 2011, vol. 1, pp. 346–350, 2011.

[10] BMKG, "Bddan Metereologi, Klimatologi dan Geofisika," BMKG, 2021. https://www.bmkg.go.id/cuaca/prakiraan-cuaca-indonesia.bmkg?Prov=34&NamaProv=Sumatera Utara (accessed Jun. 17, 2021).

[11] CWFIS, "Background Information Canadian Forest Fire Weather Index (FWI) System," 2017. [Online]. Available: https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi

[12] L. Sunitha, D. M. BalRaju, Kiran, and J. Sasi, "Detection and Analysis of Outliers and Applying Data Mining Methods on Weather Data of Bhanur

Village Detection and Analysis of Outliers and Applying Data Mining Methods on Weather Data of Bhanur Village Abstract :," no. January, 2021.

[13] C. an Hsiao and H. Chen, "On classification from the view of outliers," IAENG Int. J. Comput. Sci., vol. 37, no. 4, 2010.

[14] D. R. Manalu, M. Zarlis, H. Mawengkang, and O. S. Sitompul, "Forest Fire Prediction in Northern Sumatera using Support Vector Machine Based on the Fire Weather Index," AIRCC Publ. Corp., vol. 10, no. 19, pp. 187–196, 2020, doi: 10.5121/csit.2020.101915.

[15] V. Kecman, "Support Vector Machines – An Introduction," no. May 2005, pp. 1–47, 2005, doi: 10.1007/10984697_1.

[16] S. S. Mehta and N. S. Lingayat, "Support vector machine for cardiac beat detection in single lead electrocardiogram," Lect. Notes Eng. Comput. Sci., vol. 2, no. May, pp. 1630–1635, 2007.

[17] Z. J. Lv, Q. Xiang, and J. G. Yang, "Application of Genetic Algorithm-support vector machine for prediction of spinning quality," Proc. World Congr. Eng. 2011, WCE 2011, vol. 2, pp. 1033–1038, 2011.

[18] N. Members, "Fire Weather Index (FWI) System," National Wildfire Coordinating Group, 2021. https://www.nwcg.gov/publications/pms437/cffdrs/fire-weather-index-system (accessed Apr. 21, 2021).

[19] N. Coordinating Group Wilfire, "Fire Weather Index (FWI) System," NWCG, 2022. https://www.nwcg.gov/publications/pms437/cffdrs/fire-weather-index-system

[20] M. Yandouzi et al., "Forest Fires Detection using Deep Transfer Learning," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 8, pp. 268–275, 2022, doi: 10.14569/IJACSA.2022.0130832.

[21] A. Tharwat, A. E. Hassanien, and B. E. Elnaghi, "A BA-based algorithm for parameter optimization of Support Vector Machine," Pattern Recognit. Lett., vol. 93, pp. 13–22, 2017, doi: https://doi.org/10.1016/j.patrec.2016.10.007.

[22] Varela, "Fire Weather Index (FWI) classification for fire danger assessment applied in Greece," Tethys, J. Weather Clim. West. Mediterrania, no. 1994, pp. 31–40, 2015, doi: 10.3369/tethys.2018.15.03.

[23] L. Yudong, F. Zhongke, Z. Ziyu, C. Shilin, and Z. Hanyue, "Research on Multi-Factor Forest Fire Prediction Model Using Machine Learning Method in China," 2020.

[24] G. P. Siknun and I. S. Sitanggang, "Web-based classification application for forest fire data using the shiny framework and the C5 . 0 algorithm," Procedia Environ. Sci., vol. 33, pp. 332–339, 2016, doi: 10.1016/j.proenv.2016.03.084.

[25] G. ElSharkawy, Y. Helmy, and E. Yehia, "Employability Prediction of Information Technology Graduates using Machine Learning Algorithms," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 10, pp. 359–367, 2022, doi: 10.14569/IJACSA.2022.0131043.

[26] B. Tien, Dieu, H. N. Duc, and S. Pijush, "Spatial pattern analysis and prediction of forest fire using new machine learning approach of Multivariate Adaptive Regression Splines and Differential Flower Pollination optimization: A case study at Lao Cai province (Viet Nam)," J. Environ. Manage., vol. 237, no. January, pp. 476–487, 2019, doi: 10.1016/j.jenvman.2019.01.108.

[27] A. Srivastava, S. Umrao, S. Biswas, and I. Zafar, "FCCC : Forest Cover Change Calculator User Interface for Identifying Fire Incidents in Forest Region using Satellite Data," vol. 14, no. 7, 2023.

[28] Wijayanto, S. O, K. N D, and H. Y, "Classification Model for Forest Fire Hotspot Occurrences Prediction Using ANFIS Algorithm," IOP Conf. Ser. Earth Environ. Sci. 54 012059, no. January, 2017, doi: 10.1088/1755-1315/54/1/012059.

[29] A. M. Elshewey, "Forest Fires Detection Using Machine Learning Techniques," vol. XII, no. Ix, pp. 510–517, 2020, [Online]. Available: https://www.xajzkjdx.cn/gallery/54-sep2020.pdf

[30] X. Chen, T. Li, L. Ruan, K. Xu, J. Huang, and Y. Xiong, "Research and application of fire risk assessment based on satellite remote sensing for transmission line," Lect. Notes Eng. Comput. Sci., vol. 2219, pp. 284–287, 2015.

[31] D. R. Manalu, M. Zarlis, H. Mawengkang, and O. S. Sitompul, "Predicting rainfall from weather observations using SVM approach for identify the parameter of fuel moisture code as fire weather index," J. Theor. Appl. Inf. Technol., vol. 99, no. 16, pp. 4090–4097, 2021.

[32] Canada.ca, "Natural Resources Canada," Government of Canada, 2020. https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi

[33] M. Noor, Kebakaran Lahan Gambut. Universitas Lambung Mangkurat, 2019. [Online]. Available: http://eprints.ulm.ac.id/9594/1/2. Kebakaran Lahan Gambut-Faktor Penyebab dan Mitigasinya.pdf

[34] G. of Canada, "Canadian Wildland Fire Information System | Canadian Forest Fire Weather Index (FWI) System," https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi, 2020. https://cwfis.cfs.nrcan.gc.ca/background/summary/fwi (accessed Oct. 19, 2020).

[35] S. D. Jena, J. Kaur, Rani, and Rajneesh, "A Review of Prediction of Software Defect by Using Machine Learning Algorithms BT - Recent Innovations in Computing," 2022, pp. 61–70.

[36] M. Tanveer, A. Sharma, and P. N. Suganthan, "General twin support vector machine with pinball loss function," Inf. Sci. (Ny)., vol. 494, pp. 311–327, 2019, doi: https://doi.org/10.1016/j.ins.2019.04.032.

[37] I. Ibrahim, R. Silva, M. H. Mohammadi, V. Ghorbanian, and D. A. Lowther, "Surrogate-Based Acoustic Noise Prediction of Electric Motors," IEEE Trans. Magn., vol. 56, no. 2, pp. 1–4, 2020, doi: 10.1109/TMAG.2019.2945407.

[38] Y. Wang, X. Tang, H. Chen, T. Yuan, Y. Chen, and H. Li, "Sparse additive machine with pinball loss," Neurocomputing, vol. 439, pp. 281–293, 2021, doi: https://doi.org/10.1016/j.neucom.2020.12.129.

[39] BMKG, "BMKG, Data Online Pusat Database," 2020. https://dataonline.bmkg.go.id/home

[40] P. A. Emelia Akashah, S. K. Sugathan, and A. T. S. Ho, "Receiver operating characteristic (ROC) graph to determine the most suitable pairs analysis threshold value," Proc. - Adv. Electr. Electron. Eng. - IAENG Spec. Ed. World Congr. Eng. Comput. Sci. 2008, WCECS 2008, no. November, pp. 224–230, 2008, doi: 10.1109/WCECS.2008.35.

[41] K. H. Zou, O'Malley, A. James, and L. Mauri, "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models," Circulation, vol. 115, no. 5, pp. 654–657, 2007, doi: 10.1161/CIRCULATIONAHA.105.594929.

[42] N. Mamat, S. Fatin, and M. Razali, "Comparisons of Various Imputation Methods for Incomplete Water Quality Data : A Case Study of The Langat River , Malaysia," vol. 35, no. 1, pp. 191–201, 2023.