AJEST

Al-Salam Journal for Engineering and Technology
Journal Homepage: http://journal.alsalam.edu.iq/index.php/ajest
ISSN: 2790-4822

AL-SALAM UNIVERSITY COLLEGE

Classification of Large-Scale Datasets of Landsat-8 Satellite Image Based on LIBLINEAR Library

Ehsan Ali Al-Zubaidi[1,2*], Furkan Rabee[1], Ahmed H. Al-Sulttani [2]

[1]Department of Computer Science, Faculty of Computer Science and Mathematics, University of Kufa, Najaf, IRAQ

[2]Department of Environmental Planning, Faculty of Physical Planning, University of Kufa, Najaf, IRAQ

*Corresponding Author: Ehsan Ali Al-Zubaidi

DOI: https://doi.org/10.55145/ajest.2022.01.02.003
Received April 2022; Accepted May2022; Available online July 2022

ABSTRACT: One of the most common machine learning approaches is linear classification. Multithreading used in LIBLINEAR Library has become an important research topic for processing large datasets. Working in a multithreaded environment accelerates the training of large datasets and increases classification accuracy. This study examines multiple linear support vector machine implementations for binary classification of Landsat-8 satellite image datasets to determine the optimal model. Four datasets have been created for different scenes in Iraq. Each dataset contains millions of pixels to be classified optimally. This article selects an appropriate algorithm for such a large dataset of current types in the LIBLINEAR library optimal for organizing large datasets using the linear SVM algorithm. Each dataset containing about 4.5 million samples was used to test the performance. The seven linear SVM techniques have statically examined the most effective SVM implementation method. Accuracy, F-Score, and Kappa are used as model assessment measures to evaluate and rank the models' performances. Based on the results, the LibLINEAR library with type 4 (LINLINEAR4) classifier was the best classifier for satellite image classification in remote sensing when applied to large datasets. The accuracy, Kappa, and f-score of the LIBLINEAR 4 classifier are as follows: (92.89 %), (73.53 %), and (77.77%) with dataset1, (96.25 %), (59.97 %), and (61.88%) with dataset2, (96.65%), (68.33%), and (70.03%), and (94.72%), (53.67%), and (56.25%), respectively, with dataset4.

Keywords: Linear Classification, SVM, LibLINEAR, Large Dataset, Satellite Image.

1. INTRODUCTION

In recent years, the availability of data collecting has increased continuously. There are further difficulties to consider, such as processing large amounts of data and its storage. Artificial intelligence algorithms make it feasible for humans to interpret the data in a certain way that is unable to manage [1, 2]. Support Vector Machine (SVM) is a machine learning approach in data mining that identifies an optimal separation that correctly divides and maximizes the margins between two or more classes [3]. In recent years, support vector machine (SVM) models have been presented as one of the most the solutions to classification problems in mathematical programming. Since the traditional soft margin was first implemented, many different types of issues may be solved with the help of SVM algorithms. Face recognition technology, machine vision software, and bioinformatics [4, 5].

LibLINEAR can classify data that can be linearly separated via a hyperplane while maintaining maximum distance from the support vectors. On the other hand, this is something that may be done by using a variety of SVM types, such as L2-regularized L2-loss support vector classification (dual) and L2-regularized L2-loss support vector classification (primal) [6]. [dual] and [primal] refer to the same thing. It is critical to developing answers to issues of large-scale classification in a variety of applications, including text classification. The linear type, which has evolved over the years, is now considered to be one of the most effective learning techniques for vast sparse data with a large number of occurrences and variables [7, 8]. The LIBLINEAR algorithm is beneficial for training on large-scale issues. For instance, training only takes seconds to train a text classification issue using the Reuters Corpus Volume 1 (rcv1), which contains more than 600,000 samples. The same job would take a standard SVM solver like LIBSVM many hours to do. In addition, LIBLINEAR is on par with or even faster than other state-of-the-art linear classifiers like Pegasos [9, 10].

*Corresponding author: ihsana.kareem@uokufa.edu.iq

Kumar and Sree [6] SVM implementations resulting from 17 different combinations of SVM classifier libraries, SVM types, and Kernel techniques were statistically examined to find the most effective SVM implementation technique. Chiang, et al. [11] proposed parallelizing the dual coordinate descent algorithm using a new framework for large-scale linear classification. Lee, et al. [12] parallel sparse matrix-vector multiplications in a Newton technique for significant logistic regression. Various implementations, ranging from simple to complex, are compared and evaluated. Results show that great acceleration may be achieved under optimal conditions. Cervantes, et al. [13] the normal SVM is unsuitable for classifying big datasets since its training difficulty is significantly dependent on the size of the training set, proposed a novel SVM classification technique for large datasets.

The objectives of the current study are as the followings:

1.Apply a multithreaded library to choose the optimal model for classifying large datasets.

2.Calculating the training execution time for millions of recodes.

3.Creating new datasets for various desertification-affected Iraq areas to perform future studies on these areas.

2. MATERIALS AND METHODS

Figure 1 illustrates the stages that have been followed in this study through a flow diagram. This diagram includes the steps of data preparation, classification, accuracy assessments, and the results required to solve the research problem. The approach consists of three major stages: data preparation for training, constructing training classification models using seven models of the LIBLINEAR library and evaluating the classification accuracy of models using accuracy measures (Overall accuracy and F-score). The raw images from the Landsat-8 satellite were obtained from the United States Geological Survey (USGS) (https://earthexplorer.usgs.gov) and downloaded in the GeoTIFF file format. Satellite images of different scenes from various places in Iraq were do

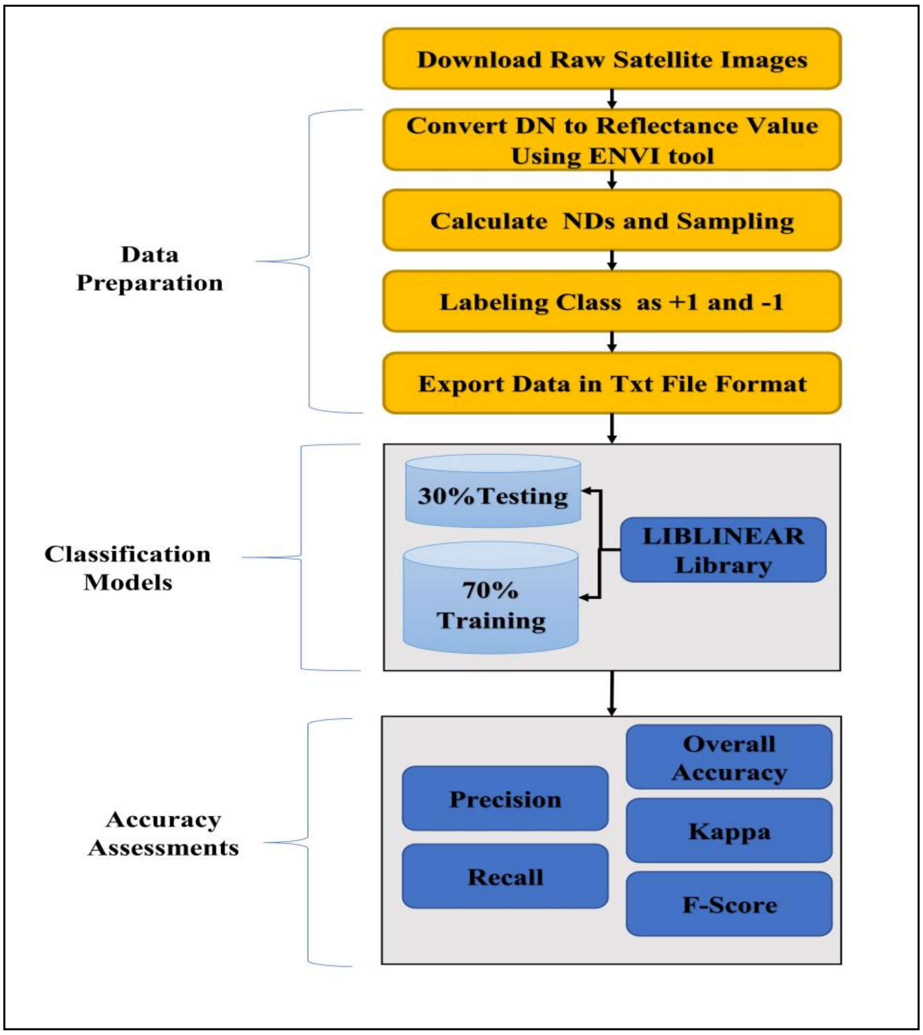
Download Raw Satellite Images
Convert DN to Reflectance Value Using ENVI tool
Calculate NDs and Sampling
Labeling Class as +1 and -1
Export Data in Txt File Format
Data Preparation
30%Testing
70% Training
LIBLINEAR Library
Classification Models
Overall Accuracy
Precision
Kappa
Recall
F-Score
Accuracy Assessments

FIGURE 1. Diagram of Proposed Methodology

Images were converted from digital numbers to reflectance values using the ENVI software. The raw Landsat 8 satellite images will be calibrated. The image acquired from the Landsat satellite-8 has dimensions of around (7881 * 7881) pixels, which means that it must be processed using approximately 62 million pixels. Thus, samples have been selected from each of the chosen scenes for Iraq. Class labels were created using the ArcGIS program, supervised learning methods, and direct observation of the studied phenomenon. Pixels that were part of the phenomenon were given a value of +1, while pixels with a value of -1 were not part of the phenomenon. Finally, all pixel values have been exported to a Txt file using the extract multi values to points tool in ArcGIS.

2.1 MATERIALS

Four datasets used in this study were created using the (ArcGIS 10.6) tool to collect Landsat 8 satellite images. They are selecting specific areas of Iraq and downloading images of them from the Landsat 8 satellite, as these images were used in other studies requiring the detection of certain phenomena using machine learning and remote sensing methods. Table 1 below contains the details of different datasets used for this study.

Table 1. Datasets were used in this study

ID	Datasets	# Rows	Class Label		Download Dataset Links
			-1	+1	
1	Dataset1	4,522,880	3790073	732807	https://www.kaggle.com/datasets/ehsanalikareem/baiji-dataset
2	Dataset2	4,527,159	4266788	260371	https://www.kaggle.com/datasets/ehsanalikareem/amara-dataset
3	Dataset3	4,525,248	4223634	301614	https://www.kaggle.com/datasets/ehsanalikareem/najaf-dataset
4	Dataset4	4,524,790	4174625	350165	https://www.kaggle.com/datasets/ehsanalikareem/samawah-dataset

All these datasets consist of 15 columns, and each column was generated using the calculation (Normalize Differences) between one band and the other band for Landsat-8 satellite images. An SVM is trained using 15 normalization differences (NDs) generated from six Landsat-8 bands (B2, B3, B4, B5, B6, and B7), comprising three visible bands (Visible, Near-infrared, and SWIR) and two SWIR bands (SWIR). Four places in Iraq with sand dunes were chosen to assess the threat of desertification. One of the most significant aspects of climate change is this phenomenon. Satellite images and remote sensing must be used to monitor and track the development of dunes over the next several years. There were four separate regions of Iraq where the data was collected. Sand dunes may be found at Baiji, Najaf, Samawah, and Amarah, among other locations.

2.2 METHODS

Support vector machines (SVMs) are a non-parametric supervised statistical learning approach. Because of this, there are no presumptions made on the distribution of the underlying data. At the beginning of the process, a collection of labeled data instances is presented to the procedure. The SVM training approach attempts to find a hyperplane that divides the dataset into a certain number of predefined classes in a manner that is consistent with the examples used for training [14]. The decision boundary that reduces the number of incorrect classifications made throughout the training process is the optimal separation hyperplane. Learning is an iterative process that involves identifying a classifier with an adequate decision boundary to separate training patterns (in a potentially high-dimensional space) and then separating simulation data using the same parameters. This process is known as "classifier searching" (dimensions) [15].

Models are constructed using LibLINEAR for each of the four datasets processed throughout the training process. LibLINEAR is now supports (1) L1-regularized L2-loss SVC, (2) L1-regularized L2-loss SVC, (3) L1-loss SVC, (4) L2-loss SVC, (5) L2-regularized LR, (6) L1-regularized LR (7) L2-regularized L2-loss SVR [17-19].

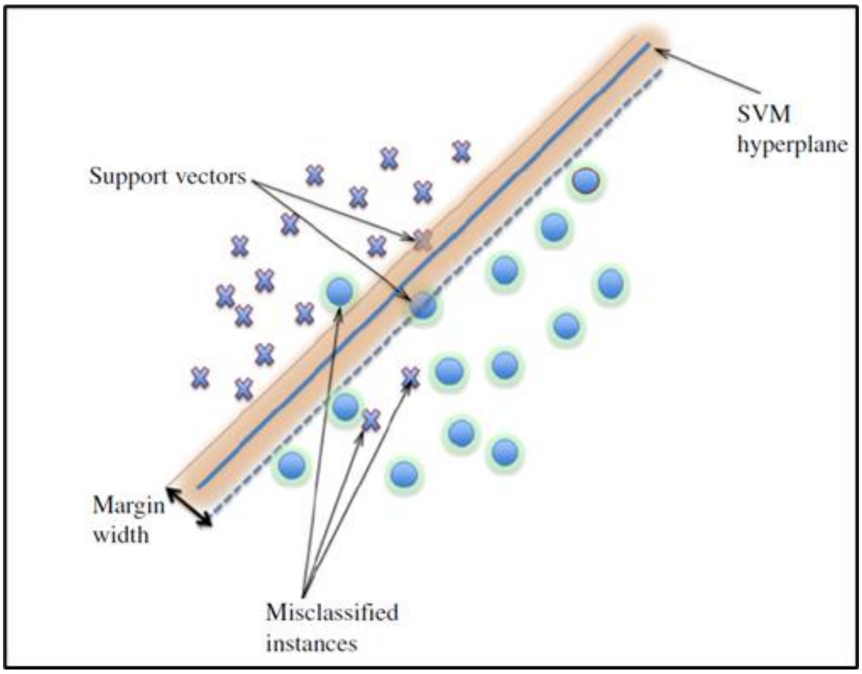

FIGURE 2. **Linear SVM [16]**

2.2.1 LIBLINEAR LIBRARY

LibLINEAR is a collection of linear SVM classifiers for large-scale linear classification that supports logistic regression and linear SVM. The LibLINEAR library helps a set of instance label pair instances is provided [Xi, Yi], i=1,⋯.,l , xi ∈ Rn ,yi ∈[-1,+1], in two groups. The response vector y ∈ Rl such that yi =[1,-1], a linear classifier function generates the weight vector was the model. The decision function is as sgn (wT x). The following points will show seven types used in this study [12, 20].

1. **(LibLINEAR 1)** L2-regularized L2-loss support vector classification (dual).

   min_alpha 0.5(alpha^T (Q + I/2/C) alpha) − e^T alpha (1)

   subject to 0 <= alpha_i

2. **(LibLINEAR 2)** L2-regularized L2-loss support vector classification (primal).

   $$\min_{w} \frac{1}{2} w^T w + C \sum_{i=1}^{l} (\max(0,1 - yiw^T xi))^2 \qquad (2)$$

3. **(LibLINEAR 3)** L2-regularized L1-loss support vector classification (dual) [21, 22].

   $$\min_{w} \frac{1}{2} \alpha^T Q\alpha - e^T \alpha \qquad (3)$$

4. **(LibLINEAR 4)** Support vector classification by Crammer and Singer.

Support Vector Classification by Crammer and Singer, Let S= [x1, y1,….,(xm,ym)] be a collection of m training cases. If xi is an example taken from the domain of x ⊑ Rn , then each label yi must be an integer from the set Y=1..., K. This is our assumption. A framework was employed based on classifiers of this kind in this approach. Where M is a k * n-dimensional matrix, and M_r is the rows of M. The inner-product of the row of M and the instance x is sometimes referred to as the trust or similarity score for the r class. As an outcome of the above definition, the predicted label is the index of the row that has the highest score of similarity to x. It is also possible to use linear binary classifiers to predict whether an instance x is labelled 1 or 2. Whether w.x is greater than zero or less than zero (as seen in the notation above). Using a 2 * n matrix, M1 = w and M2 = - w, a classifier can be built [23].

5. **(LibLINEAR 5)** L1-regularized L2-loss support vector classification.

$$\min_{w} \quad \|\mathrm{w}\| 1 + C \sum_{i=1}^{l} (\max\,(0{,}1 - yiw^{T}xi))^{2} \qquad (4)$$

6. **(LibLINEAR 6)** L1-regularized logistic regression.

$$\min_{w} \quad \|\mathrm{w}\| 1 + C \sum_{i=1}^{l} \log(1 + e^{-yiw^{T}xi}) \qquad (5)$$

7. **(LibLINEAR 7)** L2-regularized logistic regression (dual).

$$\min_{w} \frac{1}{2} \alpha^{T} Q \alpha + \sum_{i:\alpha i>0} \alpha i \log \alpha i + \sum_{i:\alpha i<c} (C - \alpha i) \log(C - \alpha i) - \sum_{i=1}^{l} C \log C \quad (6)$$

Subject to $0 \leq \alpha i \leq C$,i=1,…..,l. [20, 24].

2.2.2 ACCURACY ASSESSMENTS METRICS

Models are constructed using the R programming language, with 70 % of data used for training and 30 % for testing to evaluate the performance of the models. The first step is to determine the accuracy of all algorithms using R software (caret package). Accuracy is determined by dividing the number of accurately predicted pixels by the total number of pixels [25-27]. Accuracy is given by the formula shown in Eq. (7).

$$\text{Overall Accuracy (OA)} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \qquad (7)$$

In equations (7) and (8), the words "true positives," "true negatives," "false positives," and "false negatives," respectively, are denoted by the symbols "TP," "TN," "FP," and "FN," respectively [26]. The F-Score is a measurement of accuracy that considers the statistics recall r and precision p [27, 28]. The term "precision" refers to the ratio of "true positives," or "TP," to "all predicted positives," or "TP + FP." The ratio of true positives to all real positives (true positives plus false positives) is a recall. The formula for the F score is stated below in Eq. (8) [28, 29].

$$\mathrm{F} - \mathrm{Score} = \frac{2 * \mathrm{P} * \mathrm{R}}{(\mathrm{P} + \mathrm{R})} \text{ where } \mathrm{P} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})} \quad \text{and } \mathrm{R} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})} \qquad (8)$$

The Kappa statistic assesses the accuracy of a dataset's predicted classifications compared to its actual classifications while also considering the possibility that the two may match. On the other hand, it does not consider the costs much like the straightforward success rate. Better models will have a Kappa value that is closer to 1 [29, 30]. The Kappa metric may be calculated using the Eq. (9) shown below:

$$\mathrm{Kappa} = \frac{\mathrm{Po} - \mathrm{Pe}}{(1 - \mathrm{Pe})} \text{ where } \mathrm{Po} = \frac{(\mathrm{TP} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})} \quad \text{and } \mathrm{Pe} = \mathrm{P}(+1) + \mathrm{P}(-1)$$

$$\mathrm{P}(+1) = \frac{\mathrm{TP} + \mathrm{FN}}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})} * \frac{\mathrm{TP} + \mathrm{FP}}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})}$$

$$\mathrm{P}(-1) = \frac{\mathrm{FP} + \mathrm{TN}}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})} * \frac{\mathrm{FN} + \mathrm{TN}}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})} \qquad (9)$$

3. RESULTS AND DISCUSSION

The results will be presented in graphical and tabular forms to facilitate comparisons between the most effective classifiers for millions of rows of data. This part describes the models created during the experiments, the measurements obtained, and the various conclusions from analyzing the data taken during the experimentations. The parameters setting for all datasets are (type =1 to 7, cost = 1 and 100, bias = 1, verbose = TRUE). Tables 2,3,4, and 5 show the accuracies, Kappa, and F-scores of the models constructed and evaluated with dataset1, dataset2, dataset3, and dataset4 using R software (caret package). Furthermore, the training time for each technique with each dataset was computed to show the training time required with each dataset.

Table 2. Overall Accuracy, Kappa, and F-score for Dataset1

Classifiers	Training Time	OA%	Kappa%	Precision%	Recall%	F-score%
LibLINEAR1	2.362814 mins	92.78	72.35	80.49	73.08	76.61
LibLINEAR2	3.399906 secs	91.62	66.69	79.46	65.06	71.54
LibLINEAR3	1.133383 mins	92.86	73.43	78.64	76.74	77.68
LibLINEAR4	**57.28854 secs**	**92.89**	**73.53**	**78.75**	**76.81**	**77.77**
LibLINEAR5	4.910605 mins	92.69	72.06	80.14	72.94	76.37
LibLINEAR6	26.37953 secs	92.44	71.07	79.31	72.10	75.53
LibLINEAR7	1.08847 mins	92.74	72.31	80.04	73.44	76.60

Table 3. Overall Accuracy, Kappa, and F-score for Dataset2

Classifiers	Training Time	OA%	Kappa%	Precision%	Recall%	F-score%
LibLINEAR1	2.40823 mins	95.97	53.20	77.24	42.85	55.12
LibLINEAR2	3.607034 secs	94.79	24.55	71.94	15.96	26.13
LibLINEAR3	34.28934 secs	96.14	58.18	74.34	50.48	60.13
LibLINEAR4	**1.250147 mins**	**96.25**	**59.97**	**74.82**	**52.75**	**61.88**
LibLINEAR5	4.522531 mins	95.17	36.42	72.57	26.07	38.36
LibLINEAR6	18.90121 secs	95.90	53.69	73.90	44.69	55.70
LibLINEAR7	57.66224 secs	95.93	54.11	74.17	45.12	56.11

Table 4. Overall Accuracy, Kappa, and F-score for Dataset3

Classifiers	Training Time	OA%	Kappa%	Precision%	Recall%	F-score%
LibLINEAR1	4.519993 mins	96.61	68.47	84.71	59.95	70.21
LibLINEAR2	5.138106 mins	95.88	57.35	86.42	45.14	59.30
LibLINEAR3	1.050985 mins	96.60	67.73	86.26	58.14	69.46
LibLINEAR4	**1.23546 mins**	**96.65**	**68.33**	**86.42**	**58.87**	**70.03**
LibLINEAR5	3.763984 mins	96.27	64.19	83.72	54.57	66.07
LibLINEAR6	9.612394 mins	91.34	65.32	83.13	54.37	65.87
LibLINEAR7	1.04557 hours	96.35	66.33	81.28	58.77	68.21

Table 5. Overall Accuracy, Kappa, and F-score for Dataset4

Classifiers	Training Time	OA%	Kappa%	Precision%	Recall%	F-score%
LibLINEAR1	55.38465 mins	93.74	34.29	84.03	23.30	36.49
LibLINEAR2	6.88804 secs	92.00	0.00	0.00	0.00	0.00
LibLINEAR3	21.59737 mins	94.64	51.89	79.19	41.48	54.44
LibLINEAR4	**19.77738 mins**	**94.72**	**53.67**	**78.01**	**43.98**	**56.25**
LibLINEAR5	3.802544 mins	92.00	0.00	0.00	0.00	0.00
LibLINEAR6	2.337529 hours	91.97	0.00	0.04	0.00	0.00
LibLINEAR7	1.014323 hours	94.55	51.16	77.90	41.04	53.76

The accuracy rate for the fourth type of library work (LibLINEAR) was the best, according to Tables 2, 3, and 5. According to these results, this classifier (LibLINEAR4) is suitable for datasets with millions of records. This type was invented by researchers (Crammer and Singer) to deal with big data. Figure 3 shows the overall accuracy, Kappa, and F-score for each classifier using the four datasets. Figure 4 shows the time taken to train 1 datasets1. The most significant drawbacks are that the majority of the studies that use remote sensing data, which is large data and requires advanced methods of processing and analysis, are working to classify it using the Support Vector Machine (SVM) algorithm in the library (LIBSVM), which is considered ineffective with large data. As a result, in this article we present results showing that the use of a library (LIBLINEAR) is better in terms of efficiency and speed when working with remote sensing data.

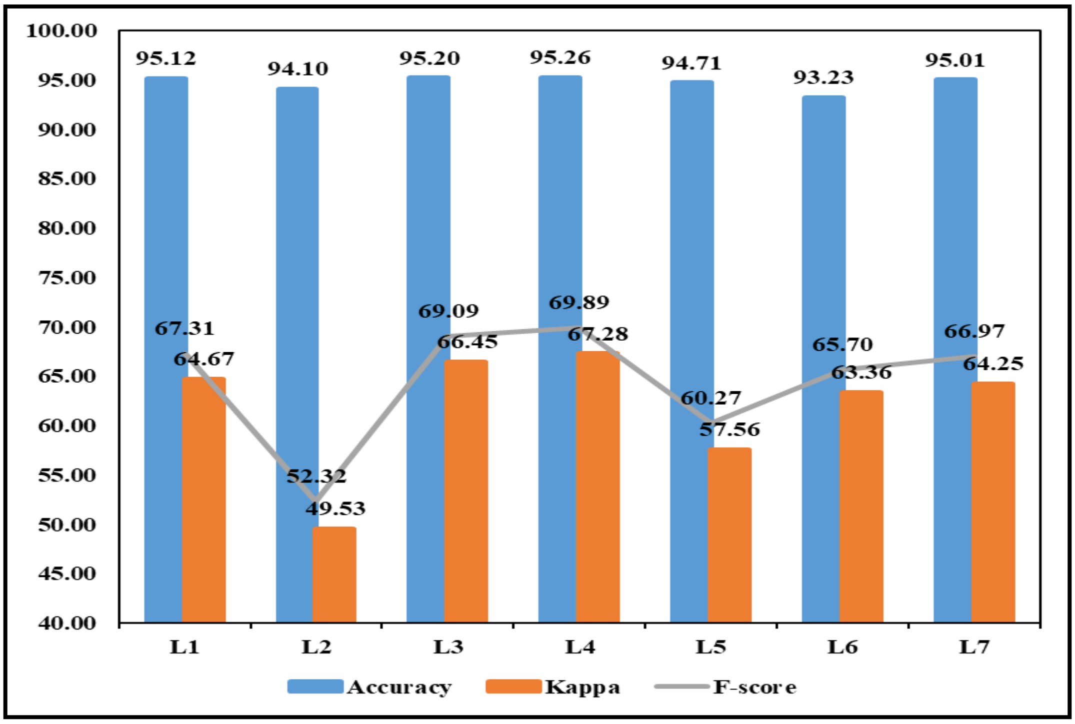

FIGURE 3. Average of Accuracy, Kappa, and F-score for Datasets 1,2, and 3

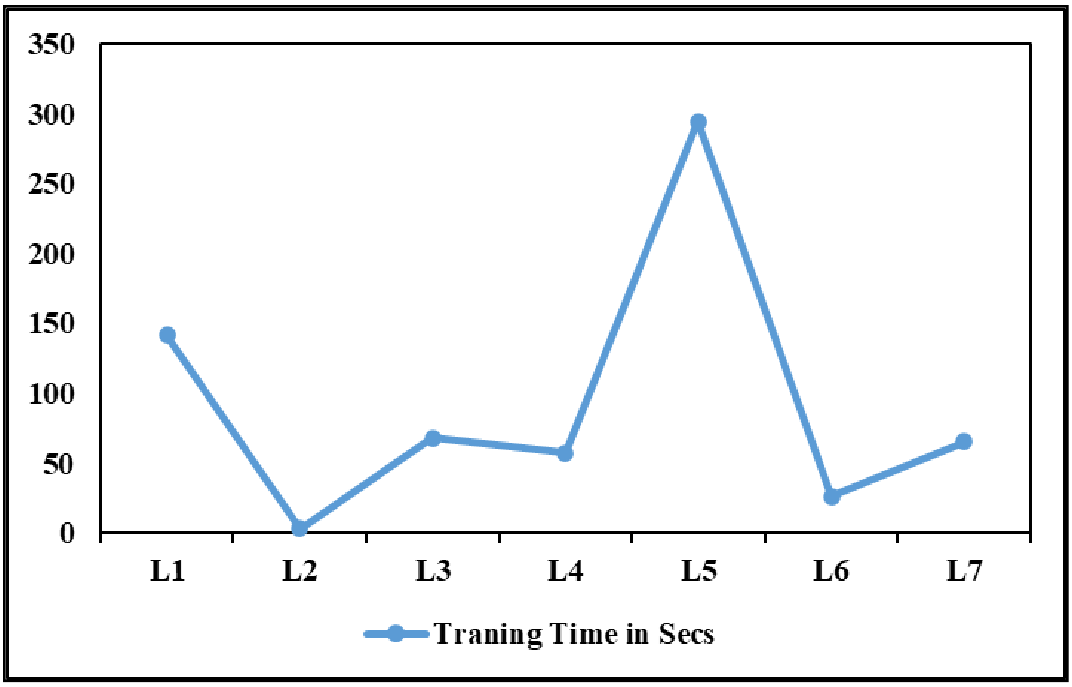

FIGURE 4. Training Time with Dataset1

4. CONCLUSIONS

LIBLINEAR is an open-source library or package suitable for large dataset linear classification. Investigations and analysis indicate that solvers in LIBLINEAR have high theoretical and practical performance. This article used many linear SVM implementations on large datasets to select the best SVM classifier. The results indicated the LIBLINEAR4 classifier, a fast and effective classification approach for big datasets. The results of the experiments indicate that the level of accuracy achieved by using the LibLINEAR4 type is comparable to that achieved by other SVM implementations such as (e1071), despite the fact that the amount of time spent training is greatly reduced. In addition, the LIBLINEAR method is scalable, meaning that it may be used to huge datasets while still maintaining a high level of classification accuracy.

REFERENCES

[1] D. Chicco, M. J. Warrens, and G. Jurman, "The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment," *IEEE Access*, vol. 9, pp. 78368-78381, 2021, doi: 10.1109/access.2021.3084050.

[2] E. A. Al-Zubaidi and M. M. Mijwil, "Medical Image Classification for Coronavirus Disease (COVID-19) Using Convolutional Neural Networks," *Iraqi Journal of Science*, pp. 2740-2747, 2021/08/31 2021, doi: 10.24996/ijs.2021.62.8.27.

[3] E. Moghadas, J. Rezazadeh, and R. Farahbakhsh, "An IoT patient monitoring based on fog computing and data mining: Cardiac arrhythmia usecase," *Internet of Things*, vol. 11, p. 100251, 2020/09 2020, doi: 10.1016/j.iot.2020.100251.

[4] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189-215, 2020/09 2020, doi: 10.1016/j.neucom.2019.10.118.

[5] P. Pławiak and U. R. Acharya, "Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals," *Neural Computing and Applications*, vol. 32, no. 15, pp. 11137-11161, 2019/01/05 2019, doi: 10.1007/s00521-018-03980-2.

[6] Y. Zhuang, Y. Juan, G.-X. Yuan, and C.-J. Lin, "Naive Parallelization of Coordinate Descent Methods and an Application on Multi-core L1-regularized Classification," presented at the Proceedings of the *27th ACM International Conference on Information and Knowledge Management*, 2018/10/17, 2018. [Online]. Available: http://dx.doi.org/10.1145/3269206.3271687.

[7] B. Schölkopf and A. J. Smola, "Learning with Kernels," ed: The MIT Press, 2018.

[8] M. Hammad, A. Maher, K. Wang, F. Jiang, and M. Amrani, "Detection of abnormal heart conditions based on characteristics of ECG signals," *Measurement*, vol. 125, pp. 634-644, 2018/09 2018, doi: 10.1016/j.measurement.2018.05.033.

[9] Z. Ao, Y. Su, W. Li, Q. Guo, and J. Zhang, "One-Class Classification of Airborne LiDAR Data in Urban Areas Using a Presence and Background Learning Algorithm," *Remote Sensing*, vol. 9, no. 10, p. 1001, 2017/09/27 2017, doi: 10.3390/rs9101001.

[10] W.-L. Chiang, M.-C. Lee, and C.-J. Lin, "Parallel Dual Coordinate Descent Method for Large-scale Linear Classification in Multi-core Environments," *presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016/08/13, 2016. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939826.

[11] W.-L. Chiang, M.-C. Lee, and C.-J. Lin, "Parallel Dual Coordinate Descent Method for Large-scale Linear Classification in Multi-core Environments," *presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016/08/13, 2016. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939826.

[12] M.-C. Lee, W.-L. Chiang, and C.-J. Lin, "Fast Matrix-Vector Multiplications for Large-Scale Logistic Regression on Shared-Memory Systems," *presented at the 2015 IEEE International Conference on Data Mining*, 2015/11, 2015. [Online]. Available: http://dx.doi.org/10.1109/icdm.2015.75.

[13] S. K. C and R. S. R. J, "COMPARISON OF PERFORMANCES OF DIFFERENT SVM IMPLEMENTATIONS WHEN USED FOR AUTOMATED EVALUATION OF DESCRIPTIVE ANSWERS," *ICTACT Journal on Soft Computing*, vol. 05, no. 03, pp. 971-978, 2015/04/01 2015, doi: 10.21917/ijsc.2015.0136.

[14] A. Abdiansah and R. Wardoyo, "Time Complexity Analysis of Support Vector Machines (SVM) in LibSVM," *International Journal of Computer Applications*, vol. 128, no. 3, pp. 28-34, 2015/10/15 2015, doi: 10.5120/ijca2015906480.

[15] M. Z. Islam, "EXPLORE: A Novel Decision Tree Classification Algorithm," *in Data Security and Security Data*, ed: Springer Berlin Heidelberg, 2012, pp. 55-71.

[16] P. Danenas and G. Garsva, "Credit risk evaluation modeling using evolutionary linear SVM classifiers and sliding window approach," *Procedia Computer Science*, vol. 9, pp. 1324-1333, 2012, doi: 10.1016/j.procs.2012.04.145.

[17] C.-C. Chang and C.-J. Lin, "LIBSVM," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 1-27, 2011/04 2011, doi: 10.1145/1961189.1961199.

[18] S. Shalev-Shwartz, Y. Singer, N. Srebro, and A. Cotter, "Pegasos: primal estimated sub-gradient solver for SVM," *Mathematical Programming*, vol. 127, no. 1, pp. 3-30, 2010/10/16 2010, doi: 10.1007/s10107-010-0420-4.

[19] S. S. Keerthi, S. Sundararajan, K.-W. Chang, C.-J. Hsieh, and C.-J. Lin, "A sequential dual method for large scale multi-class linear svms," *presented at the Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08*, 2008. [Online]. Available: http://dx.doi.org/10.1145/1401890.1401942.

[20] C.-J. Hsieh, K.-W. Chang, C.-J. Lin, S. S. Keerthi, and S. Sundararajan, "A dual coordinate descent method for large-scale linear SVM," *presented at the Proceedings of the 25th international conference on Machine learning - ICML* '08, 2008. [Online]. Available: http://dx.doi.org/10.1145/1390156.1390208.

[21] C.-J. Hsieh, K.-W. Chang, C.-J. Lin, S. S. Keerthi, and S. Sundararajan, "*A dual coordinate descent method for large-scale linear SVM," presented at the Proceedings of the 25th international conference on Machine learning - ICML* '08, 2008. [Online]. Available: http://dx.doi.org/10.1145/1390156.1390208.

[22] J. Cervantes, X. Li, W. Yu, and K. Li, "Support vector machine classification for large data sets via minimum enclosing ball clustering," *Neurocomputing*, vol. 71, no. 4-6, pp. 611-619, 2008/01 2008, doi: 10.1016/j.neucom.2007.07.028.

[23] V. Vapnik, "The Problem of Estimating Dependences from Empirical Data," in Estimation of Dependences Based on Empirical Data, ed: Springer New York, 2006, pp. 1-26.

[24] V. Vapnik, "The Problem of Estimating Dependences from Empirical Data," in Estimation of Dependences Based on Empirical Data, ed: Springer New York, 2006, pp. 1-26.

[25] T. Joachims, "Training linear SVMs in linear time," *presented at the Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* - KDD '06, 2006. [Online]. Available: http://dx.doi.org/10.1145/1150402.1150429.

[26] K. Crammer and Y. Singer, "Ultraconservative Online Algorithms for Multiclass Problems," in Lecture Notes in Computer Science, ed: Springer Berlin Heidelberg, 2001, pp. 99-115.

[27] V. N. Vapnik, "Direct Methods in Statistical Learning Theory," in *The Nature of Statistical Learning Theory*, ed: Springer New York, 2000, pp. 225-265.

[28] C. J. C. Burges, Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121-167, 1998, doi: 10.1023/a:1009715923555.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995/09 1995, doi: 10.1007/bf00994018.

[30] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," *presented at the Proceedings of the fifth annual workshop on Computational learning theory - COLT* '92, 1992. [Online]. Available: http://dx.doi.org/10.1145/130385.130401.