ILKOM Jurnal Ilmiah Vol. 17, No. 1, April 2025, pp. 27-33
Accredited 2nd by RISTEKBRIN No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Research Article **Open Access (CC–BY-SA)**

Evaluation of Multi-Class Classification Performance Lung Cancer Through K-NN and SVM Approach

Muh. Indra Endriartono Saputra Troy [a,1]; Sitti Rahmah Jabir [a,2,*]; Siska Anraeni [a,3]
[a] *Informatics Engineering, Faculty of Computer Science, Universitas Muslim Indonesia Makassar, Indonesia*
[1] *13020210298@umi.ac.id;* [2] *rahmahjabir@umi.ac.id;* [3] *siska.anraeni@umi.ac.id*

Article history: Received November 07, 2024; Revised December 23, 2024; Accepted March 11, 2025; Available online April 20, 2025

Abstract

Lung cancer is one of the deadliest diseases in the world with a mortality rate of 25% of all cancer-related deaths in 2021. Lung cancer is a lung disease caused by genetic changes in respiratory epithelial cells, resulting in uncontrolled cell proliferation. In an effort to improve diagnosis and treatment, this study proposes an approach for multiclass performance evaluation using K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms based on 2024 data. in this study KNN is implemented conventionally while SVM applies 2 kernel processes, namely Linear and Polynominal. The data used is 1000 rows and uses 24 variables with a ratio of 70% training data and 30% testing data, the data in this study includes important information such as medical history, diagnostic test results, and clinical characteristics of patients. this study aims to determine which algorithm has the best performance by looking at the final results based on accuracy in identifying lung cancer data. Based on the research and discussion of SVM and KNN performance evaluation, the SVM algorithm produces an accuracy of 98.28%, surpassing the accuracy of the KNN algorithm of 97.25%. Therefore, the results show that the SVM algorithm is superior to the KNN algorithm. The KNN and SVM methods were implemented for multi-class classification of lung cancer, allowing identification of various subtypes of lung cancer with optimal accuracy.

Keywords: K-Nearest Neighbors; Lung Cancer; Support Vector Machine.

Introduction

The lungs are one of the most important organs that humans have. Lung health is essential for survival, but there are still many people who don't care about their lung health. This has led to the level of public awareness to live clean, and the minimal uncertainty of lung symptoms makes the rate of lung disease even higher [1], [2] . Smoking is one of the causes of disease and does not only affect active smokers, as smoking has become one of the causes of lung disease. Detection of lung cancer can be done conventionally, for example through consultation with an internal medicine specialist (*Pulmonology*) and laboratory tests, so the cost will be quite large [3], [4].

According to the latest GLOBACON statistics collected in 2018 show that there were approximately 2,094,000 cases of lung cancer. Asymptomatic lung development is the most important component in making this disease so deadly. Making lung cancer the leading cancer incidence worldwide. With an estimated 1,369,000 cases, lung cancer is the second most common cancer in men, after prostate cancer, and the second most common cancer in women, after breast cancer, with 725,000 cases [5]. Symptoms of lung cancer generally include hemoptysis, hoarseness, coughing, and chest pain which can be treated with chemotherapy, radiotherapy and surgical therapy. Many people know that other conditions along with lung cancer cause lung X-rays to be inaccurate. As lung cancer can spread quickly, early diagnosis is crucial. Imaging technologies such as low-dose CT today allow lung cancer to be diagnosed at an early stage [6], [7].

Data mining is a series of activities that use existing data to obtain analytical material for technology development, decision making, and so on [8]. Data mining is a series of activities that use existing data to obtain analytical material for technology development, decision making, and so on. There are various methods in classification, such as Naïve Bayes Classifier, Support Vector Machine and others. This research uses KNN and SVM classification methods to process existing data [9], [10]. KNN and SVM methods, can be used to predict various information from data based on diagnosis, for example, classifying lung cancer based on data experienced by patients, which is used to forecast the likelihood of developing the disease and generate a new classification of the disease data [11], [12].

Previous studies have implemented the K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) algorithms in lung cancer classification. One of them is a study conducted by Ravi Kumar Sachdeva, who used a

http://dx.doi.org/10.33096/ilkom.v17i1.2464.27-33

dataset consisting of 309 rows with 15 variables. In the study, the KNN algorithm managed to achieve an accuracy rate of more than 91.93%, while SVM showed better performance with an accuracy exceeding 96.77%. This finding indicates that SVM is more effective in processing lung cancer datasets, mainly due to its ability to determine the optimal hyperplane to distinguish classes. On the other hand, the accuracy of KNN is highly dependent on the number of nearest neighbors, which can be affected if the dataset has a high dimensionality. Therefore, the selection of the algorithm should be tailored to the characteristics of the dataset and the parameters used in the classification process. Level as outcome criteria. The sample of the dataset used is presented [13].

This study aims to evaluate two machine learning algorithms with simulated split data comparison of 70% testing data and 30% training data using lung cancer data consisting of 1000 rows with 26 data variables through the KNN and SVM algorithm approaches on lung cancer data. This study aims to determine which algorithm has the best accuracy results in identifying lung cancer data, so that it can help to draw a more accurate conclusion. The KNN method is implemented conventionally, while SVM is applied with 2 kernels namely Linear and Polymonial. To evaluate the performance of the two algorithms, classification is performed by comparing the accuracy of the two algorithms.

Method

The series of processes that will be carried out in the research of multiclass performance evaluation in lung cancer using KNN and SVM algorithms are shown in **Figure 1**.

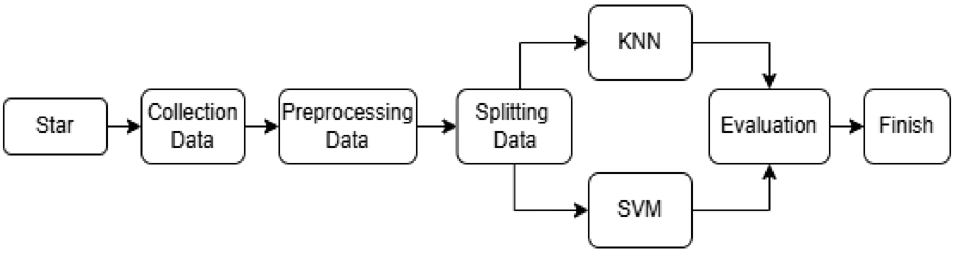

Figure 1. Series of Research Processes

Based on **Figure 1**, the stages carried out in this study are data collection, data preprocessing, feature selection, splitting data separation between training and testing data, and evaluating models from the KNN and SVM methods. The hardware used in this research, namely the Windows 10 64 bit system operating system, Intel Core i5-5200U @ 2.20GHz processor, 8GB RAM, 512 memory, and the graphics used Intel HD Graphics 5500. The software used in this research is Google Collaboratory with python programming language.

A. Data Description

Data obtained from the Kaggle website as a reference data in this study and the data will be processed. The data obtained has 1000 rows and 26 variables (Gender, Air Pollution, Alcohol use, Dust Allergy, OccuPational Hazards, Genetic Risk, chronic Lung Disease, Balanced Diet, Obesity, Smoking, Passive Smoker, Chest Pain, Coughing of Blood, Fatigue, Weight Loss, Shortness of Breath, Wheezing, Swallowing Difficulty, Clubbing of Finger Nails, Frequent Cold, Dry Cough, Snoring, Level). The data will be divided into two parts, namely Testing Data and Training Data**Error! Reference source not found.Error! Reference source not found.**. The data used is in **Table 1**.

Table 1. Metadata

Index	Patient Id	Age	…	Dry Cough	Snoring	Lavel
0	P1	33	…	3	4	Low
1	P10	17	…	7	2	Medium
2	P100	35	…	7	2	High
3	P1000	37	…	7	5	High
…	…	…	…	…	…	…
996	P996	37	…	1	4	High
997	P997	25	…	7	2	High
998	P998	18	…	2	3	High
999	P999	47	…	7	2	High

B. Data Prerocessing

Data preprocessing is a fairly important stage in the classification process because preprocessing includes a series of steps taken to make raw data or data collected from sources more ready for analysis or modeling [14]. Data processing aims to clean, change and prepare data so that it is easier and more accurate in carrying out the analysis process. In this research, data processing includes processes, such as data cleaning and data transformation. Data preprocessing has an important role in data mining to create consistent and high-quality datasets [15].

C. Python

Python is a programming language developed by Guido Van Rossum specifically for scripting and web development purposes. Python is the programming language used in the creation of KNN and SVM algorithms [16]. There are several libraries that can be used for KNN and SVM, such as NumPy, matplotlib.pyplot and pandas. NumPy is a library used for numerical operations and multidimensional arrays. Matplotlib is a library used in data visualization such as making graph plots Pandas is a library that focuses on performing data manipulation, data preparation, and data cleaning [17].

Phyton can be used for various software development purposes or machine learning purposes and phyton is a high-level language program that is easy to learn and understand which can make it easier for users to find a mistake [18]. The use of python can be accessed through Goggle Collab or Google Colaboratory by providing free access to users such as GPU Google Collab which can be accessed easily and flexibly [19].

D. K-Nearest Neighbor

KNN is a method to classify objects based on the learning data of the object's nearest neighbors. Near or far neighbors are usually calculated based on Euclidean distance [20], [21]. As a system capable of searching data, a classification system is required. A supervised learning algorithm is used for KNN, which classifies new query results according to the majority of KNN categories [22]. This space is divided into sections based on the classification of the training sample, a point in this space is marked with a class, if a class is its classification it is most often found in the point's nearest neighbor. Near or far neighbors are usually calculated based on Euclidean Distance [23], [24]. general formula used in the KNN method:

$$d_i = \sqrt{(X_{KI} + X_{KJ})^2 + (Y_{KI} + Y_{KJ})^2 + (Z_{KI} + Z_{KJ})^2} \tag{1}$$

E. Support Vector Machine

SVM is a learning system that uses linear function theory in features that are trained using algorithms based on optimal theory. The basic theory of SVM is derived from a mixture of pre-existing computational theories [25], [26]. SVM are powerful algorithms for classification and regression problems and are widely applied to real-world applications. SVM builds its solution in terms of training input subsets, which gives the SVM framework the same advantages as parametric techniques in terms of reduced computation time for testing and storage requirements [27]. The basic principle of SVM is the development of linear classifiers to process non-linear problems. This basic principle uses the kernel trick method on high-dimensional features [28]. The accuracy of the data generated by the SVM algorithm is determined by the hyperplane calculation formula in SVM as follows:

$$f(x) = W \,.\, X + \mathrm{b} \tag{3}$$

and the kernel functions that can be used are shown in **Table 2** [29].

Table 2. Kernel Table

Kernel Name	Kernel Fungsion
Linear (Dot)	$\boldsymbol{G(X_1, X_2) = x_1 \times x_2}$
Polynomial	$\boldsymbol{G(X_1, X_2) = \gamma x_1 \times c}$

F. Performance Value Results (Accuracy, Precision, Recall dan F1-score)

The results of the performance value present the prediction by comparing the results of the original value with the predicted value to produce accuracy, precision, recall and f1-score values in lung cancer disease [30].

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{4}$$

$$\text{Precision} = \frac{(TP)}{(TP+FP)} \tag{5}$$

$$\text{Recall} = \frac{(TP)}{(TP+FN)} \tag{6}$$

$$\text{F1-Score} = 2\,\frac{Precision \times Recall}{Precision \times Recall} \tag{7}$$

Results and Discussion

To facilitate the testing of the proposed method in this study, the authors used Goggle Collab tools to shorten the testing time of the User Knowlage Modeling Dataset to see the accuracy results of the multiclass classification comparison of KNN and SVM methods. This research was conducted with a lung cancer dataset consisting of 1000 rows with a split data method of 70% training data and 30% testing data.

A. Data Preprocessing

At this stage the data will be preprocessed so that the data can produce good accuracy. Preprocessing data processing is carried out to remove unimportant data variables and change inconsistent data characters from Letters to Numeric in order to speed up the training process and improve the quality of the model which can affect the accuracy results. The data columns that will be removed in Preprocessing are Index and Patient Id, and the data columns that are changed in the Preprocessing stage are Level and age.

B. KNN Algorithm Implementation

KNN classification calculation at this stage is calculating KNN conventionally without performing feature selection. In the KNN algorithm, experiments were carried out with a K value of 10, namely 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 by looking at the results of the average value based on precision, recall and f1-score accuracy for 10 times can be seen in Table 3.

Table 3. Conventional KNN result table

K	Target class	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
1	*Low*	*100*	*95*	*97*	*98,57*
	Medium	*96*	*100*	*98*	
	High	*100*	*100*	*100*	
2	*Low*	*100*	*95*	*97*	*98,57*
	Medium	*96*	*100*	*98*	
	High	*100*	*100*	*100*	
3	*Low*	*100*	*95*	*97*	*98,57*
	Medium	*96*	*100*	*98*	
	High	*100*	*100*	*100*	
4	*Low*	*100*	*95*	*97*	*96.28*
	Medium	*90*	*100*	*95*	
	High	*100*	*94*	*97*	
5	*Low*	*100*	*95*	*97*	*96.28*
	Medium	*90*	*100*	*95*	
	High	*100*	*94*	*97*	
6	*Low*	*100*	*95*	*97*	*96.28*
	Medium	*90*	*100*	*95*	
	High	*100*	*94*	*97*	
7	*Low*	*100*	*95*	*97*	*97.57*
	Medium	*96*	*97*	*97*	
	High	*97*	*100*	*99*	
8	*Low*	*100*	*95*	*97*	*97.57*
	Medium	*96*	*97*	*97*	
	High	*97*	*100*	*99*	
9	*Low*	*100*	*92*	*96*	*96.42*
	Medium	*94*	*96*	*95*	
	High	*97*	*100*	*98*	
10	*Low*	*100*	*92*	*96*	*96.42*
	Medium	*94*	*96*	*95*	
	High	*97*	*100*	*98*	
Average Accuracy					*97.25*

Based on Table 3 above, it can be seen that the KNN method is able to provide a high enough value with an average accuracy value of 97.25%. awhere the highest accuracy value of the KNN method is obtained when the K value is 1, 2 and 3, which is 98.57%, while the lowest accuracy value is obtained when the K value is 9 and 10 with an accuracy value of 96.16%. based on tests that have been carried out from the lung cancer disease data set, it can be seen that the KNN algorithm can provide a good accuracy value in classification.

C. SVM Algorithm Implementation

The results of the SVM algorithm use two kernels, namely Linear and Polynominal with split data 70%: 30%. The accuracy obtained in the SVM algorithm is held by the Linear kernel with a percentage of 98.28%. then, the precision value obtained with an average value of 98%, the recall value obtained with an average value of 98.3% and the f1-score value obtained an average value of 98%. While the polynominal kernel obtained an accuracy value of 89.71%. then, the precision value obtained with an average value of 89.3%, the recall value obtained an average value of 89.6% and the f1-score value obtained an average value of 89.3%. accuracy results from accuracy, precision, recall and f1-score can be seen in Table 4.

Table 4. Results of Accuracy, Precision, Recall and F1-Score on SVM algorithm

Kernel	Target class	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Linear	*Low*	*98*	*98*	*98*	*98.28*
	Medium	*98*	*97*	*97*	
	High	*98*	*100*	*99*	
Polynominal	*Low*	*83*	*91*	*87*	*89.71*
	Medium	*90*	*78*	*83*	
	High	*95*	*100*	*98*	

D. Comparison of the results of algorithms

In this section, we discuss the test results that have been obtained to prove the performance of the two algorithms that have been applied. Both algorithms have good performance in classifying lung cancer. For a clearer comparison between the conventional KNN algorithm and SVM with a linear kernel can be seen in Table 5.

Table 5. Comparison value of Accuracy, Precision, Recall and F1-score results

Algorithm	Target class	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
KNN	*Low*	*100*	*94.4*	*96.8*	*97.25*
	Medium	*93.8*	*98.6*	*96.3*	
	High	*98.8*	*98.2*	*98.5*	
SVM	*Low*	*98*	*98*	*98*	*98.28*
	Medium	*98*	*97*	*97*	
	High	*98*	*100*	*99*	

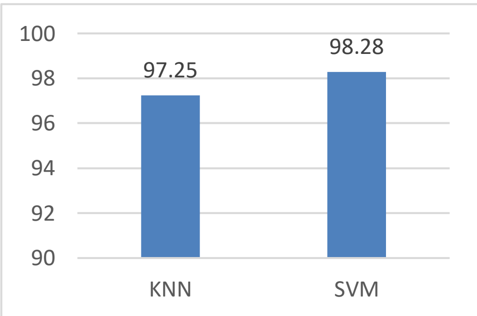

Figure 2. Comparison value of KNN and SVM algorithm accuracy results

Based on the Table 5 and Figure 2 it can be seen that the SVM method is able to provide a good accuracy value with an accuracy value of 98.28%. by looking at the average value based on precision, recall and f1-score the SVM algorithm with a linear kernel has a higher value than the conventional KNN algorithm. Based on the final results of the two algorithms, SVM and KNN can properly classify lung cancer and by producing accuracy values not too far away. This shows that both SVM and KNN algorithms are effective in classifying lung cancer. However, in the classification of lung cancer, the SVM algorithm is better than the KNN algorithm. In this research, there is a need for improvement to conduct classification experiments, especially in implementing code in the phyton language for the KNN and SVM algorithms to work optimally. And also in this study can also test the KNN algorithm using other testing methods and also SVM using other kernel methods, such as RBF (Radial Basis Function) and Sigmoid.

Conclusion

Based on the results of this study, it can be concluded that the KNN and SVM algorithms can produce accuracy values above 95%, which shows that both algorithms have good performance in classifying lung cancer diseases. In the split data simulation dataset 70% training data and 30% testing data where the SVM algorithm is superior to the KNN algorithm by producing accuracy of 98.28%, precision 98%, recall 98.3%, and f1-score 98%. While the KNN algorithm produces an accuracy value of 97.25%, precision 97.5%, recall 97.2% and f1-score 97.2%. This is influenced because SVM functions well in a high feature space with many relevant features. This allows it to handle multiclass problems well, especially in a large feature space.

References

[1] J. Alam, S. Alam, and A. Hossan, "Multi-stage lung cancer detection and prediction using multi-class svm classifie," in *2018 International conference on computer, communication, chemical, material and electronic engineering (IC4ME2)*, IEEE, 2018, pp. 1–4. doi: 10.1109/IC4ME2.2018.8465593.

[2] S. Kumar, H. Kumar, G. Kumar, S. P. Singh, A. Bijalwan, and M. Diwakar, "A methodical exploration of imaging modalities from dataset to detection through machine learning paradigms in prominent lung disease diagnosis: a review," *BMC Med. Imaging*, vol. 24, no. 1, pp. 1–42, 2024, doi: 10.1186/s12880-024-01192-w.

[3] E. Taucher, I. Mykoliuk, J. Lindenmann, and F. M. Smolle-Juettner, "Implications of the Immune Landscape in COPD and Lung Cancer: Smoking Versus Other Causes," *Front. Immunol.*, vol. 13, no. March, pp. 1–18, 2022, doi: 10.3389/fimmu.2022.846605.

[4] R. Nooreldeen and H. Bach, "Current and future development in lung cancer diagnosis," *Int. J. Mol. Sci.*, vol. 22, no. 16, pp. 2163–2169, 2021, doi: 10.3390/ijms22168661.

[5] K. C. Thandra, A. Barsouk, K. Saginala, J. S. Aluru, and A. Barsouk, "Epidemiology of lung cancer," *Wspolczesna Onkologia*, vol. 25, no. 1. Termedia Publishing House Ltd., pp. 45–52, 2021. doi: 10.5114/wo.2021.103829.

[6] H. S. Choi, B. K. Jeong, H. Jeong, I. B. Ha, and K. M. Kang, "Role of radiotherapy in the management of malignant airway obstruction," *Thorac. Cancer*, vol. 11, no. 8, pp. 2163–2169, 2020, doi: 10.1111/1759-7714.13523.

[7] P. A. Grenier, A. L. Brun, and F. Mellot, "The Potential Role of Artificial Intelligence in Lung Cancer Screening Using Low-Dose Computed Tomography," *Diagnostics*, vol. 12, no. 10, pp. 1–13, 2022, doi: 10.3390/diagnostics12102435.

[8] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, 2022, doi: 10.1186/s40561-022-00192-z.

[9] H. Mydyti, "Data Mining Approach Improving Decision-Making Competency along the Business Digital Transformation Journey: A Case Study – Home Appliances after Sales Service," *SEEU Rev.*, vol. 16, no. 1, pp. 45–65, 2021, doi: 10.2478/seeur-2021-0008.

[10] M. Bansal, A. Goyal, and A. Choudhary, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning," *Decis. Anal. J.*, vol. 3, no. May, p. 100071, 2022, doi: 10.1016/j.dajour.2022.100071.

[11] K. B. C and N. K. B, "Image Processing-based Performance Evaluation of KNN and SVM Classifiers for Lung Cancer Diagnosis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 5, 2024, doi: 10.14569/IJACSA.2024.0150546.

[12] P. Chaturvedi, A. Jhamb, M. Vanani, and V. Nemade, "Prediction and Classification of Lung Cancer Using Machine Learning Techniques," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1099, no. 1, p. 012059, Mar. 2021, doi: 10.1088/1757-899x/1099/1/012059.

[13] R. K. Sachdeva, P. Bathla, P. Rani, R. Lamba, G. S. P. Ghantasala, and I. F. Nassar, "A novel K-nearest neighbor classifier for lung cancer disease diagnosis," *Neural Comput. Appl.*, vol. 36, no. 9, 2024, doi: 10.1007/s00521-024-10235-w.

[14] V. Çetin and O. Yıldız, "A Comprehensive Review On Data Preprocessing Techniques In Data Analysis," *Pamukkale Univ. J. Eng. Sci.*, vol. 28, no. 2, pp. 299–312, 2022, doi: 10.5505/pajes.2021.62687.

[15] K. Maharana, S. Mondal, and B. Nemade, "A Review: Data Pre-Processing and Data Augmentation Techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022, doi: 10.1016/j.gltp.2022.04.020.

[16] A. Jaiswal and A. Dwivedi, "Python: The Versatile Language-Recent Trends in Programming Languages Recent Trends in Programming Languages Python: The Versatile Language," *STM Journals*, vol. 8, no. 1, p. 2021, 2021, doi: 10.37591/RTPL.

[17] A. Ramesh, B. Dileep, A. Naveen, Y. Manoj, and V. Vijay, "Prediction At Early-Stages Using Machine Learning," *Int. J. Progress. Res. Eng. Manag. Sci.*, vol. 04, no. 04, pp. 308–316, 2024.

[18] S. Saabith, T. Vinothraj, and M. Fareez, "A review on Python libraries and Ides for Data Science," *Int. J. Res. Eng. Sci*, vol. 9, no. 11, pp. 36–53, 2021.

[19] A. A. S. Arip, K. Sazali, K Norazlianie, and N. Jamaludin, T Ahmad, M Faiz, Ab. Razak, "Object Detection for Safety Attire Using YOLO (You Only Look Once)," *J. Adv. Res. Appl. Mech.*, vol. 113, no. 1, pp. 37–51, 2024, doi: 10.37934/aram.113.1.3751.

[20] A. Aisyah and S. Anraeni, "Analisis Penerapan Metode K-Nearest Neighbor (K-NN) pada Dataset Citra Penyakit Malaria," *Indones. J. Data Sci.*, vol. 3, no. 1, pp. 17–29, 2022, doi: 10.56705/ijodas.v3i1.22.

[21] S. Anraeni, E. R. Melani, and H. Herman, "Ripeness identification of chayote fruits using HSI and LBP feature extraction with KNN classification," *Ilk. J. Ilm.*, vol. 14, no. 2, pp. 150–159, 2022, doi: 10.33096/ilkom.v14i2.1153.150-159.

[22] P. M. Arabi, "Computer Aided Classification of Lung Cancer, Ground Glass Lung and Pulmonary Fibrosis Using Machine Learning and KNN Classifier," 2024. doi: 10.14569/IJACSA.2024.01507111.

[23] N. Maleki, Y. Zeinali, and S. T. A. Niaki, "A k-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection," *Expert Syst. Appl.*, vol. 164, Feb. 2021, doi: 10.1016/j.eswa.2020.113981.

[24] A. R. Isnain, J. Supriyanto, and M. P. Kharisma, "Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 2, p. 121, Apr. 2021, doi: 10.22146/ijccs.65176.

[25] Amaliah Faradibah, Dewi Widyawati, A Ulfah Tenripada Syahar, and Sitti Rahmah Jabir, "Comparison Analysis of Random Forest Classifier, Support Vector Machine, and Artificial Neural Network Performance in Multiclass Brain Tumor Classification," *Indones. J. Data Sci.*, vol. 4, no. 2, pp. 54–63, Jul. 2023, doi: 10.56705/ijodas.v4i2.73.

[26] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, 2023, doi: 10.3390/computation11030052.

[27] S. R. Jabir, Purnawansyah, H. Darwis, H. Lahuddin, A. Faradibah, and A. W. M. Gaffar, "Evaluation of Tourism Object Rating Using Naïve Bayes, Support Vector Machine, and K-Means for Business Intelligence Application in Indonesia Tourism," *Proc. 2024 18th Int. Conf. Ubiquitous Inf. Manag. Commun. IMCOM 2024*, 2024, doi: 10.1109/IMCOM60618.2024.10418390.

[28] K. N. S Khan, U. Samra, H. Aznaoui, C. B. Şahin, and Ö. B. Dinler, "Generalization of Linear and Non-Linear Support Vector Machine In Multiple Fields: A Review," *Comput. Sci. Inf. Technol.*, vol. 4, no. 3, pp. 226–239, 2023, doi: 10.11591/csit.v4i3.pp226-239.

[29] U. Amelia, J. Indra, and A. F. N. Masruriyah, "The Classification Of Stroke Prediction Using The Support Vector Machine (SVM) Method," *Sci. Student J. Information, Technol. Sci.*, vol. III, no. 2, pp. 254–259, 2022.

[30] A. Tasnim, M. Saiduzzaman, M. A. Rahman, J. Akhter, and A. S. M. M. Rahaman, "Performance Evaluation of Multiple Classifiers for Predicting Fake News," *J. Comput. Commun.*, vol. 10, no. 09, pp. 1–21, 2022, doi: 10.4236/jcc.2022.109001.