Jurnal CoreIT, Vol.X, No.X, Month Year
ISSN 2460-738X (Print)
ISSN 2599-3321 (Online)

Performance Comparasion of Adaboost and PSO Algorithms for Cervical Cancer Classification Using KNN Algorithm

Romdan Muhamad Ubaidilah[1], Sutedi[2*]
[12] Faculty of Computer Science, Institute of Informatics and Business Darmajaya, Lampung, Indonesia
[1]romdan.2221210029@mail.darmajaya.ac.id, [2]sutedi@darmajaya.ac.id

Abstract. Cervical most cancers affects women's reproductive organs and is the second maximum generally identified infection amongst girls international. According to the World Health Organization (WHO), over 600,000 women are diagnosed with cervical cancer annually, resulting in more than 300,000 deaths from the disease. Lack of foresight and early cervical cancer identification results in many deaths. To find out if a patient has cancer cells in her cervix, four screening techniques can be used: Hinselmann, Schiller, Cytology, and Biopsy. The KNN method will be used in this study to assess patient health history data, and the Adaboost and PSO algorithms will then be used to optimize the results. The two optimization methods will be compared to find the most accurate model in identifying patterns in cervical cancer patients and predicting patient screening results, whether positive or negative regarding cervical cancer. This research uses the RapidMiner tool. The final results show that the KNN algorithm is able to carry out multilabel classification analysis effectively, and the classification results optimized with PSO produce an increase in the level of accuracy.

Keywords: Cervical Cancer, Classification, KNN, Adaboost, PSO

Received July 2024 / **Revised** July 2024 / **Accepted** September 2024

INTRODUCTION

Cervical cancer is a serious contamination affecting woman reproductive health. this feminine reproductive organ is critical to the human reproductive system, and cervical cancer can be fatal for a girl if now not diagnosed and dealt with early. Cervical most cancers has come to be a first-rate fear for the international society, with occurrence growing over time. in this paragraph, we can pass over the information of cervical cancer, which includes its devastating impact on women health and the tries performed to conquer it.

Cervical maximum cancers is a shape of most cancers that develops within the cervix, the lowermost part of the uterus. This cervix has an crucial position in the woman reproductive technique, cervical maximum cancers typically develops slowly from exceptional mobile changes to cancer cells[1]. In step with information obtained from the World Health Organization (WHO), cervical cancer is one of the leading causes of dying in women worldwide. In 2022, are about more than 600,000 newly diagnosed instances and more than 300,000 deaths from cervical cancer [2]. Those mortality expenses are alarming, specifically as maximum of the sufferers come from countries with low and medium Human development Index (HDI), which often do no longer have adequate access to healthcare offerings and information on early detection and treatment of cervical most cancers[3]. Regular with GLOBOCAN 2022, in Indonesia there are greater than 36,000 cases and 20,000 deaths because of cervical most cancers. In Indonesia, cervical cancers is the 1/two main purpose of maximum cancers demise after breast most cancers[4].

The primary chance element for cervical most cancers is infection via the Human Papilloma Virus (HPV). This virus can cause adjustments in the cells of the cervix, triggering tumour increase. Cervical Intraepithelial Neoplasia (CIN), a pre-cancerous stage, is regularly the start of the cervical most cancers technique. but, it ought to be cited that now not all HPV infections will cause cervical cancer, and maximum infections will depart on their own with out causing critical health issues [5]. Cervical cancer typically takes years or even decades to develop from the early tiers to turn out to be malignant [6]. To lessen the chance of growing the sickness, early detection and right treatment are important.

DOI: 10.24014/coreit.v10i2.31711

Preventive measures are key in decreasing the burden of cervical maximum cancers. the World Health Organization (WHO) has emphasised the importance of vaccination in the direction of Human Papilloma Virus (HPV) as a key approach to defend girls from infections that could doubtlessly motive cervical most cancers. HPV vaccines have been proven to be powerful in stopping HPV infection, and consequently, decreasing the risk of developing cervical cancer [7]. Supplying HPV vaccination to prone populations, including adolescent ladies earlier than they're uncovered to the virus, can help lessen the incidence of cervical cancer inside the future. further, vast vaccination programmes can also assist lessen health disparities between unique financial companies, by way of supplying extra get right of entry to to powerful health protection [8].

In addition to HPV vaccination, early detection is also an important step in cervical most cancers prevention. Screening methods inclusive of Pap smear take a look at and through (Acetic Acid visible Inspection) take a look at have become standardised in early detection of cervical cancer. The Pap smear check, which involves sampling uterine cells for evaluation, has been proven to be effective in detecting adjustments in cervical cells which can cause cancer [9]. But, the manual Pap smear analysis manner has boundaries, together with a bent in the direction of procedural errors and the time required to finish the evaluation [10]. Consequently, studies continues to expand automatic analysis methods the usage of laptop technology and information mining. those techniques allow for faster and extra accurate interpretation of Pap smear information, consequently enabling extra effective early diagnosis and well timed remedy.

Research associated with cervical cancer has been a prime awareness inside the quest to enhance early detection, diagnosis and treatment of the sickness. numerous studies were conducted by using scientists and researchers in an try to increase new strategies which are greater powerful in addressing the challenges related to cervical cancer. underneath, a number of the great research associated with this topic will be mentioned in more element:

1. **Life Expectancy Prediction Study of Lung Cancer Patients using Boosted KNN**
   This take a look at objectives to increase a prediction version for the existence expectancy of lung cancer sufferers after present process thoracic surgical procedure. This study uses the enhanced okNearest Neighbor (KNN) set of rules (Boosted KNN) to are expecting lifestyles expectancy consequences. The results of this observe show that the advanced model has a high prediction accuracy charge, reaching 85.11% [11]. This studies offers crucial insights into the potential of the KNN algorithm inside the prediction of health results of sufferers with cancer.

2. **Optimising diabetes classification with a machine learning-based framework**
   This journal discusses various optimisation techniques, including Adaboost, to improve diabetes classification accuracy. This research shows that using data balancing techniques and more sophisticated algorithms can improve classification results on unbalanced diabetes datasets [12].

3. **Comparison of Bagging and Adaboost Methods on C4.5 Algorithm for Stroke Prediction**
   The results showed that both ensemble techniques significantly improved model accuracy compared to the use of the standard C4.5 algorithm. However, Adaboost tends to provide better results in some evaluation metrics compared to Bagging. These findings highlight the importance of using ensemble techniques in the development of more effective and reliable medical prediction systems [13].

4. **Implementation of Adaboost Method to Optimise Diabetes Disease Classification with Naïve Bayes Algorithm**
   This observe applies the Adaboost method to enhance the accuracy of diabetes sickness type the usage of the Naïve Bayes algorithm. With a 60/forty facts cut up, the Naïve Bayes algorithm produces an accuracy of zero.7608. After the software of Adaboost, the accuracy improved to 0.7694. This indicates that Adaboost is effective in improving the overall performance of Naïve Bayes set of rules in diabetes class [14]

5. **Chronic Kidney Disease Prediction using PSO-based KNN Algorithm**
   This study explores the use of Particle Swarm Optimisation (PSO)-based totally KNN set of rules for persistent kidney disease prediction. The consequences of this look at show that the KNN algorithm

optimised with PSO has a higher stage of accuracy in persistent kidney disorder prediction. This studies makes an essential contribution to the development of persistent disorder prediction strategies the usage of optimisation strategies [15].

6. **Use of Particle Swarm Optimisation with Naïve Bayes Method to Predict Heart Disease**
   This studies introduces experiments involving the Naïve Bayes algorithm with move-validation, wherein the second one experiment makes use of Particle Swarm Optimisation (PSO) function optimisation. The accuracy of the primary experiment reached eighty five.12%, whilst the second experiment reached eighty four.16%, indicating that PSO presents a mild development in prediction accuracy. version assessment using location below the Curve (AUC) confirmed appropriate overall performance from both experiments [16].

7. **Comparison of Support Vector Machine Algorithm and AdaBoost for Prediction of Student Graduation Time**
   This research compares the performance of support Vector device (SVM) and AdaBoost algorithms in predicting scholar graduation time. AdaBoost showed higher accuracy (zero.73) compared to SVM (0.sixty two), highlighting the effectiveness of AdaBoost in predicting commencement time. This observe emphasises the importance of utilising device learning strategies to examine student facts for correct predictions [17].

8. **Prediction of Customer Churn Rate in Telecommunication Company with AdaBoost Algorithm**
   This examine uses ensemble strategies which include AdaBoost to predict consumer churn inside the telecommunications enterprise. AdaBoost confirmed the very best accuracy of 80%, highlighting its effectiveness in predicting purchaser churn. This check offers perception into the significance of records mining techniques to analyse customer behaviour and take proactive measures in maintaining customers [18].

9. **Anaemia Prediction using Particle Swarm Optimisation (PSO) and Naïve Bayes Algorithm**
   This studies combines Naïve Bayes algorithm with particle swarm optimisation (PSO) to are expecting anaemia. The Naïve Bayes model completed an accuracy of ninety three.88%, which improved barely to ninety four.02% after integration with PSO. This result indicates that PSO can enhance the performance of Naïve Bayes prediction in detecting anaemia [19].

10. **Feature Selection using Information Gain in K-Nearest Neighbor (KNN) and Modified KNearest** Neighbor (MKNN) Methods for Chronic Kidney Disease Classification
    This studies specializes in feature selection using the facts benefit method for continual kidney disease type. MKNN performed 100% accuracy for 20 capabilities with a okay=5 fee, while KNN finished the highest accuracy for five functions with a k=three price. This study highlights the importance of function selection in enhancing the performance of category algorithms for continual kidney disease analysis [20].

Several researchers have focused on developing prediction models and improving the classification efficiency of medical data using various algorithms, including KNN, Adaboost, and other ensemble techniques. However, research specifically addressing algorithm optimisation for early detection of cervical cancer is limited. Therefore, this study intends to fill this gap by developing a more accurate and efficient data analysis method for cervical cancer early detection using a combination of KNN, Adaboost, and PSO algorithms. In this observe, optimization became performed the use of the Adaboost and PSO algorithms in classifying pap smear statistics for cervical most cancers sufferers the use of the KNN set of rules to determine whether or not there was an boom in accuracy from the use of the 2 algorithms. Experiments in this look at used RapidMiner tool.

RESEARCH METHODOLOGY

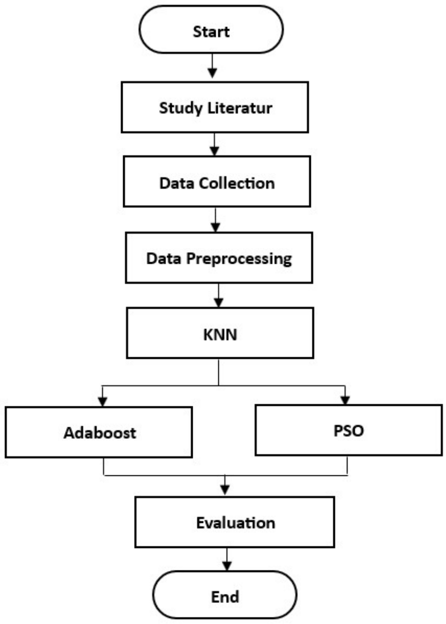

Figure 1. Research Stages

Data Collection

The first step in this research is data collection. The data used is the Cervical Cancer Risk Classification data which contains a list of risk factors for cervical cancer. This dataset is obtained from the website www.kaggle.com. This dataset consists of 858 records covering 36 attributes or features, with 4 target variables (Biopsy, Cytology, Hinselmann, and Schiler). These target variables will be used in training the model to classify cervical cancer. After training and testing the dataset using the KNN algorithm, as well as optimising it using the Adaboost and PSO algorithms, an analysis was conducted regarding the performance of each algorithm [21], [22].

Data Preprocessing

At this stage, data preprocessing is carried out. It starts by analysing the data and replacing each empty value in the attribute with the average function in RapidMiner so that all data is filled and ready to use [5].

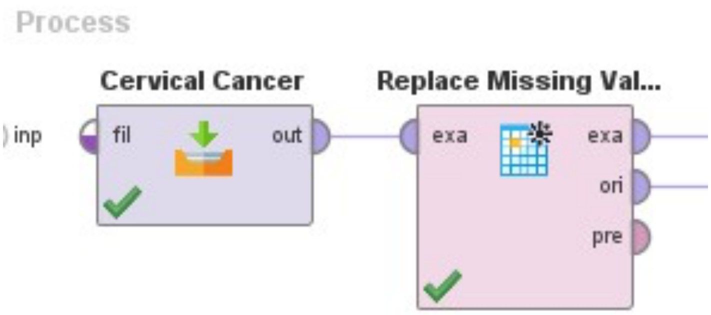

Figure 2. Replace Missing Value

K-Nearest Neighbour

After the statistics processing is entire, this studies maintains with the aid of making use of the k-Nearest Neighbor (KNN) algorithm. The k-Nearest Neighbor (KNN) algorithm is one of the similarity-based classification techniques, in which the prediction of latest information labels is primarily based on most of the people of its nearest neighbour information labels within the characteristic area. in spite of its simplicity, KNN is able to deal with classification issues properly, mainly in the case of facts that isn't always linear or has complex styles. inside the context of this research, KNN is used to categorise cervical cancer primarily based at the functions in the dataset.

KNN uses the concept of similarity between training data and test data to determine the class label of the new test data. There are several commonly used similarity metrics, such as Euclidean Distance and Cosine Similarity. Euclidean Distance measures the distance between two points in the feature space, while Cosine Similarity measures the directional similarity between two vectors in the feature space. These two metrics are used to calculate the level of similarity between the training and test data in this study [6], [10].

In the implementation of KNN in this study, four target variables will be used in the dataset, namely Biopsy, Cytology, Hinselmann, and Schiler. These four variables will be the classification targets in the KNN model training. The main purpose of using KNN is to classify patients based on the risk of developing cervical cancer based on their features.

The KNN model training process will involve the following steps:

1. Selecting the value of k, which is the number of nearest neighbours that will be used to determine the class label of the test data.
2. Calculating the distance between the test data and each training data in the feature space using the chosen similarity metric.
3. Selects the k training data with the closest distance to the test data.
4. Calculates the majority of the class labels of the k nearest neighbours.
5. Set the majority class label as the predicted class label for the test data.

By implementing the KNN algorithm, it is expected to obtain a cervical cancer classification model that can classify patients with a high level of accuracy. The results of this classification can be used to support efforts for early detection and treatment of this disease, as well as improve the quality of life of patients affected by the disease [6], [11].

Adaboost

Adaboost, or Adaptive Boosting, is a boosting method that iteratively combines several weak models into one strong model. In the context of this research, Adaboost will be applied to improve the accuracy of the KNN model in classifying cervical cancer.

The first step in the implementation of Adaboost is the initialisation of weights for each instance in the dataset. The initial weights are assigned to each instance with the same value. Then, the KNN model is trained on the weighted dataset. After obtaining the initial model, the weights on incorrectly classified instances are increased, while the weights on correctly classified instances are decreased. This is done to give more focus to the hard-to-classify instances in the next iteration.

The process continues by combining the results of several weighted KNN models to form a stronger final model. By using the boosting approach, it is expected that the resulting model will be able to improve performance in classifying cervical cancer, especially in handling imbalanced data [23].

Particle Swarm Optimisation (PSO)

Particle Swarm Optimisation (PSO) is an optimisation technique inspired by the social behaviour of flocks of birds or fish. In this research, PSO will be used to optimise the KNN parameters, such as the K value and other parameters that affect the performance of the model.

The first step in PSO is particle initialisation, where each particle represents a potential solution with different parameter values. Next, a health evaluation is performed to measure the performance of each particle based on the classification accuracy of the KNN model. The process continues by updating the particle's speed and position based on the particle's personal and social experience.
Optimal search is performed by repeating the update process until convergence to the optimal solution that maximises model accuracy. By applying PSO, it is expected to find a combination of parameters that provides the best accuracy for cervical cancer classification [15].

Evaluation

Evaluation is done by measuring the level of accuracy to ensure that the results obtained achieve the expected quality standards. In the context of this research, the evaluation is carried out using the cross fold validation method. This approach is used to evaluate the performance of a model or algorithm by dividing the dataset into two parts: training data and test data. Cross fold validation was chosen because it has been proven to be effective in reducing bias in sampling. This technique consistently divides the dataset into training and test data parts, giving each data a chance to be tested. The value of K in this context refers to the number of folds used for the division between training and test data. In this study, K=10 was chosen as the optimal number of folds [24].

RESULTS AND DISCUSSION

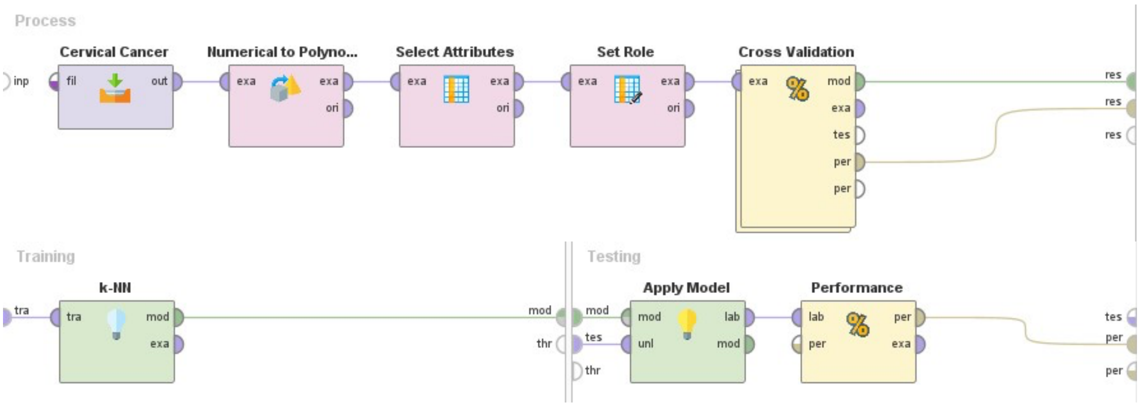

Figure 3. Testing Model using KNN algorithm

The process in RapidMiner began with the use of the "Cervical Cancer" dataset consisting of 858 data entries. This dataset has a wide array of attributes that include demographic information and medical history of the patient, such as age, number of sexual partners, smoking history, and history of hormonal contraceptive use, among others. The first step was to convert the numerical attributes into polynomials using the "Numerical to Polynomial" operator by setting the attribute filter type to "all". This aims to expand the feature representation and allow the model to capture more complex relationships between attributes.

After that, the most relevant attributes were selected for classification purposes. Attribute selection is done by selecting the attributes "biopsy", "citology", and "schiller" which are considered to have a significant correlation with cervical cancer diagnosis. Then, the attribute "Hinselmann" was set as the label to be predicted by the model, using the "Set Role" operator which sets the role of the attribute.

Next, a cross-validation process is performed to evaluate the performance of the model. In this process, the dataset was divided into 10 folds to enable more objective testing of the model's performance. The next stage is model training using the K-Nearest Neighbors (KNN) algorithm with parameter k=5. The trained model was then tested using separate test data by applying the previously generated model.

This study tests the performance of the K-Nearest Neighbors (KNN) algorithm in classifying the risk of cervical cancer based on patient clinical data. The evaluation process was conducted after the dataset went through a series of preprocessing processes. The cross-validation method with 10 folds is used to ensure the validity of the evaluation results, and performance measurement is performed using Confusion Matrix. Table 1 shows the performance evaluation results of the KNN algorithm at various values of k (number of nearest neighbours).

Table 1. KNN Testing Results

	Accuracy			
K	Hinselmann	Schiller	Citology	Biopsy
1	91.61%	86.95%	91.50%	90.44%
3	95.11%	89.98%	94.17%	92.43%
5	95.81%	91.26%	94.64%	93.01%

The test results show that the highest accuracy rate is achieved when k = 5. This accuracy rate is 95.81% for Hinselmann, 91.26% for Schiller, 94.64% for Citology, and 93.01% for Biopsy. These results demonstrate the ability of the KNN model to provide satisfactory predictions in classifying cervical cancer risk. This high accuracy reflects how the KNN model, with the right parameters, can be an effective tool in supporting data-driven medical diagnostics. The combination of proper preprocessing and the use of robust validation methods ensures that the results obtained have a high degree of reliability.

This research also explores the optimisation of the KNN algorithm using two different techniques: Adaboost and Particle Swarm Optimisation (PSO).

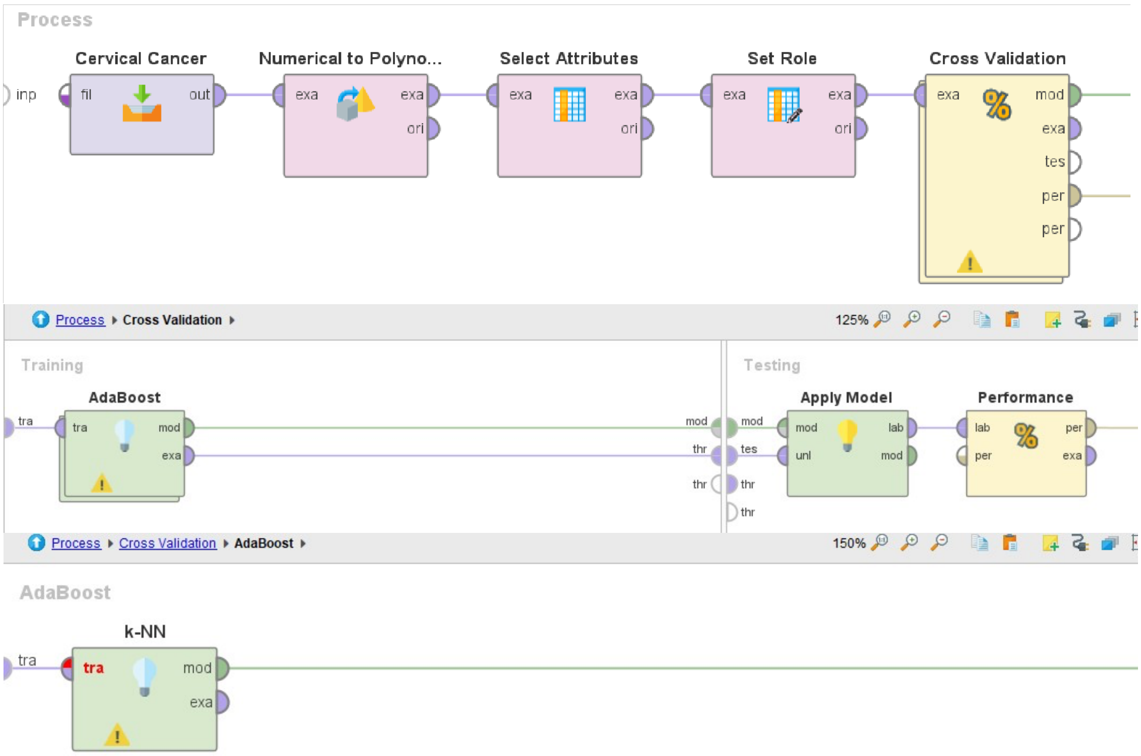

Figure 4. Testing Model using KNN algorithm optimised with Adaboost

KNN (k-Nearest acquaintances) has most important drawbacks in phrases of sensitivity to outliers and inability to handle unbalanced facts properly. to triumph over those troubles, optimization with Adaboost and PSO (Particle Swarm Optimization) is important. Adaboost is a boosting approach that could improve the accuracy of the model through giving more weight to statistics this is difficult to categorise, as a consequence assisting KNN in detecting greater complicated styles [11]. PSO, as a swarm-based totally optimization algorithm, is used to optimize attribute weights in KNN. This facilitates in choosing the most applicable attributes for classification, improving the general precision and accuracy of the KNN version [15]. The aggregate of those two strategies is predicted to conquer the constraints of KNN and produce an improved and correct version.

First, Adaboost is used to improve the performance of the KNN algorithm. This ensemble learning technique improves the model's capabilities by way of giving better weights to statistics this is difficult to classify with the aid of KNN. This process includes using 10-fold move-validation to ensure accurate evaluation. The consequences display that the combination of KNN with Adaboost does now not increase or decrease accuracy as compared to traditional KNN. table 2 shows the accuracy effects with Adaboost.

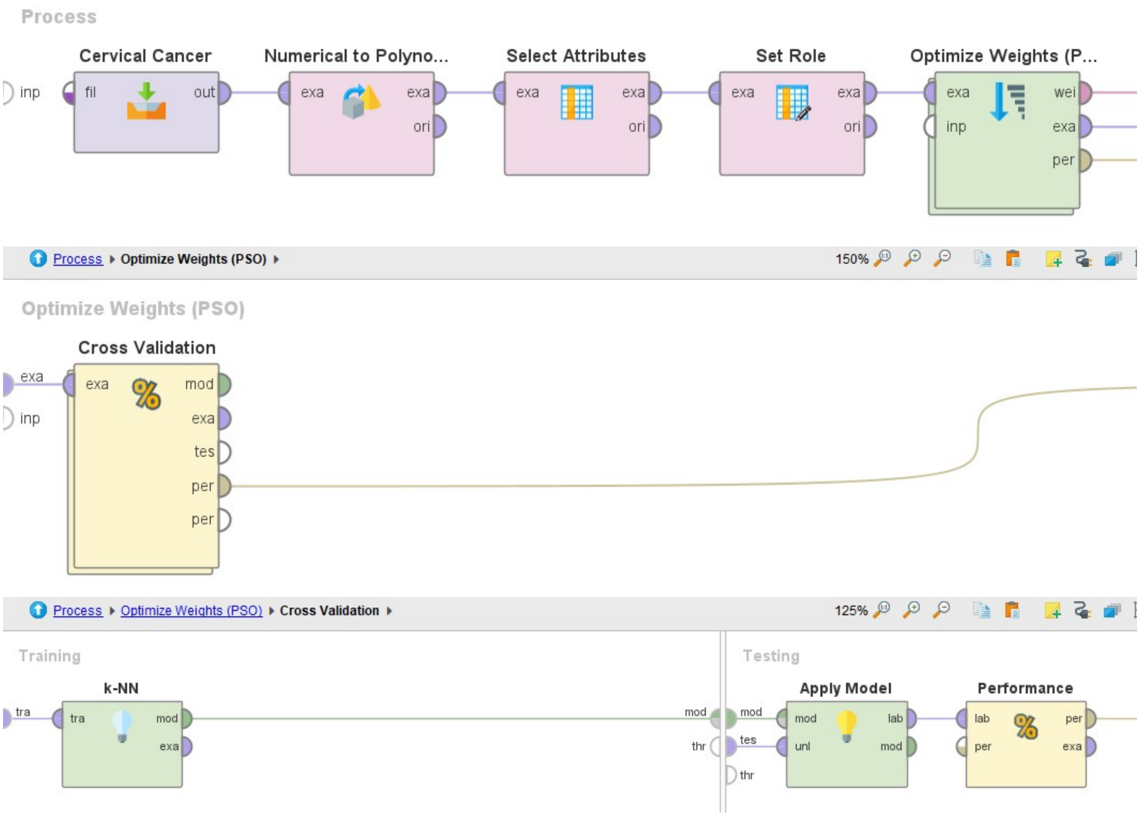

Figure 5. Testing Model using KNN algorithm optimized with PSO

Second, PSO was used to optimise the attribute weights in the KNN algorithm. PSO is an optimisation algorithm that mimics the behaviour of a flock of birds or fish to find the optimal solution. This optimisation process aims to determine the best weight for each attribute, thus improving the performance of the KNN model. The test results show that the combination of KNN with PSO provides an increase in accuracy, although not significant, especially on the Hinselmann target which increased from 95.81% to 95.92%.

Furthermore, the KNN algorithm was retested with k = 5 optimised using Adaboost and PSO techniques. The results of these tests are documented in Figure 6.

Table 2. Comparison of Accuracy Levels

	Accuracy			
	Hinselmann	**Schiller**	**Citology**	**Biopsy**
KNN	95.81%	91.26%	94.64%	93.01%
KNN + Adaboost	95.81%	91.26%	94.64%	93.01%
KNN + PSO	95.92%	91.49%	94.88%	92.82%

From the test results on model performance, it can be seen that the KNN algorithm is able to perform multilabel classification analysis well. Based on the results of the tested model accuracy performance, the KNN algorithm does not experience an increase or decrease in accuracy when combined with Adaboost.

In contrast, when combined with PSO, the KNN algorithm experienced an increase in accuracy, although not very significant. It can be seen in the Hinselmann target variable which experienced an increase in accuracy from 95.81% to 95.92%. Experiments on this model spent an average training time of about two minutes. Experiments show that cervical cancer diagnosis can be done through analysing a person's medical record data.

The findings of this study show that optimising the KNN algorithm with PSO gives better results compared to Adaboost. However, the accuracy improvement achieved is still limited. This research can be extended by testing the model using the latest observation data for further development, as well as exploring other optimisation techniques to improve model performance in cervical cancer classification.

CONCLUSION

From the results of this study, it can be concluded that the KNN algorithm has good potential in performing multilabel classification analysis for cervical cancer classification based on clinical data. Although the use of Adaboost does not provide a significant change in performance, the combination of KNN with Adaboost still shows a consistent ability to maintain model accuracy. However, optimisation with PSO managed to improve the accuracy rate, especially on the Hinselmann target. Based on the test results, it was found that the accuracy of the KNN model was 95.81%, while after being optimised with PSO, the accuracy increased to 95.92%.

These findings have important implications in the development of a more effective and accurate cervical cancer early diagnosis system. Future research can expand the scope by testing the model using the latest observation data for further development. In addition, exploration of other optimisation techniques can also be done to improve the performance of the KNN model. Thus, the results of this study can be a valuable foundation for further research in the field of cervical cancer classification.

REFERENCES

[1] N. Cancer Institute, “Understanding Cervical Changes: A Health Guide,” 2021. Accessed: Jul. 06, 2024. [Online]. Available: https://www.cancer.gov/types/cervical/screening/abnormal-hpv-paptest-results/understanding-cervical-changes.pdf

[2] Sung H *et al.*, “Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” 2022. Accessed: Jul. 06, 2024. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf

[3] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA Cancer J Clin*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/caac.21660.

[4] C. M. S. I. P. D. P. M. Z. A. B. F. Ferlay J, “Statistics at a glance, 2022 Top 5 most frequent cancers Number of new cases 408 661 Number of deaths 242 988 Number of prevalent cases (5-year),” 2021. Accessed: Jul. 06, 2024. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/360-indonesia-fact-sheet.pdf

[5] H. Akbar, “Classification of Cervical Cancer Using Convolutional Neural Network Model (AlexNet),” Journal of Informatics and Computer, Accredited by KEMENRISTEKDIKTI, vol. 4, no. 1, 2021, doi: 10.33387/jiko.

[6] R. Kurniawan, D. E. K. Sasmito, and F. Suryani, “Classification of Cervical Cells Using Nuclei Feature Analysis on Pap Smear Images,” in National Seminar on Medical Informatics (SNIMed), 2013.

[7] S. Rathod, J. Potdar, A. Gupta, N. Sethi, and A. Dande, “Empowering Women’s Health: Insights Into HPV Vaccination and the Prevention of Invasive Cervical Cancer,” *Cureus*, Nov. 2023, doi: 10.7759/cureus.49523.

[8] J. C. Spencer, N. T. Brewer, T. Coyne-Beasley, J. G. Trogdon, M. Weinberger, and S. B. Wheeler, “Reducing poverty-related disparities in cervical cancer: The role of HPV vaccination,” *Cancer Epidemiology Biomarkers and Prevention*, vol. 30, no. 10, pp. 1895–1903, Oct. 2021, doi: 10.1158/1055-9965.EPI-21-0307.

[9] A. Gates *et al.*, "Screening for the prevention and early detection of cervical cancer: protocol for systematic reviews to inform Canadian recommendations," *Syst Rev*, vol. 10, no. 1, Dec. 2021, doi: 10.1186/s13643-020-01538-9.

[10] A. Kim, K. C. Chung, C. Keir, and D. L. Patrick, "Patient-reported outcomes associated with cancer screening: a systematic review," *BMC Cancer*, vol. 22, no. 1, Dec. 2022, doi: 10.1186/s12885-02209261-5.

[11] R. T. Prasetio and S. Susanti, "Prediction of Life Expectancy of Lung Cancer Patients After Thoracic Surgery Using Boosted k-Nearest Neighbor," JURNAL RESPONSIF, vol. 1, no. 1, pp. 64–69, 2019, [Online]. Available: http://ejurnal.univbsi.id/index.php/jti

[12] X. Feng, Y. Cai, and R. Xin, "Optimizing diabetes classification with a machine learning-based framework," *BMC Bioinformatics*, vol. 24, no. 1, Dec. 2023, doi: 10.1186/s12859-023-05467-x.

[13] N. D. Saputri, K. Khalid, and D. Rolliawati, "SISTEMASI: Journal of Information Systems Comparison of Bagging and Adaboost Methods on C4.5 Algorithm for Stroke Prediction." [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[14] L. Pebrianti, F. Aulia, H. Nisa, and K. Saputra, "Journal of Information Systems and Technology Implementation of Adaboost Method to Optimize Diabetes Classification Using Naïve Bayes Algorithm," vol. 7, no. 2, 2022, [Online]. Available: http://jurnal.unmuhjember.ac.id/index.php/JUSTINDO

[15] W. Yunus, "K-Nearest Neighbor Algorithm Based on Particle Swarm Optimization for Chronic Kidney Disease Prediction," Jurnal Teknik Elektro CosPhi, vol. 2, no. 2, pp. 2597–9329, 2018.

[16] C. C. C. C. Eko Justino Wahyu, "The Application Of Particle Swarm Optimization Using Naive Bayes Method For Predicting Heart Disease," Bandar Lampung, 2022. Accessed: Jul. 13, 2024. [Online]. Available: https://jurnal.darmajaya.ac.id/index.php/icitb/article/view/3395

[17] H. Sabita and S. Trisnawati, "Comparison of Support Vector Machine and AdaBoost Algorithms in Predicting Student Graduation Time," Teknika, vol. 17, no. 2, pp. 1–5, 2023, doi: https://doi.org/10.5281/zenodo.8220872.

[18] A. S. W. G. Iqbal Muhammad Latief, "Prediction of Customer Churn Rate in Telecommunications Companies Using AdaBoost Algorithm," JURNAL INFORMATIKA, vol. 21, no. 1, 2021, doi: https://doi.org/10.30873/ji.v21i1.2867.

[19] S. Tri Utami, S. Lestari, and H. Widi Nugroho, "PREDICTION OF ANEMIA USING THE PARTICLE SWARM OPTIMIZATION (PSO) AND NAÏVE BAYES ALGORITHM," *Jurnal CoreIT*, vol. X, No.X, 2024, doi: 10.24014/coreit.v10i1.28428.

[20] A. Ramadhan, E. Budianita, F. Syafria, and S. Ramadhani, "Feature Selection using Information Gain on the K-Nearest Neighbor (KNN) and Modified K-Nearest Neighbor (MKNN) Methods for Chronic Kidney Disease Classification," *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, vol. 9, no. 2, p. 17, Dec. 2023, doi: 10.24014/coreit.v9i2.26834.

[21] H. Lin, Y. Hu, S. Chen, J. Yao, and L. Zhang, "Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks," *IEEE Access*, vol. 7, pp. 71541–71549, 2019, doi: 10.1109/ACCESS.2019.2919390.

[22] M. M. Rahaman *et al.*, "A survey for cervical cytopathology image analysis using deep learning," 2020, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2020.2983186.

[23] P. Kumar *et al.*, "Futuristic Trends in Network and Communication Technologies Communications in Computer and Information Science 1395," 2020. [Online]. Available: http://www.springer.com/series/7899

[24] S. Lestari, Yulmaini, Aswin, S. Y. Ma'ruf, Sulyono, and R. R. N. Fikri, "Alleviating cold start and sparsity problems in the micro, small, and medium enterprises marketplace using clustering and imputation techniques," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 3, pp. 3220–3229, Jun. 2024, doi: 10.11591/ijece.v14i3.pp3220-3229.