International Journal of Artificial Intelligence in Medical Issues

Volume 1 Issue 1
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v1i1.83

Research Article

Performance Analysis of the Decision Tree Classification Algorithm on the Pneumonia Dataset

Ahmad Naswin[1,*], Adityo Permana Wibowo[2]
[1] Universitas Megarezky, Makassar, Indonesia, ahmadnaswin@unimerz.ac.id
[2] Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia, adityopw@uty.ac.id

Correspondence should be addressed to Ahmad Naswin; ahmadnaswin@unimerz.ac.id

Received 10 January 2023; Revised 20 February 2023; Accepted 1 April 2023; Published 31 May 2023

Abstract:
The rapid advancements in machine learning have paved the way for innovative approaches in medical imaging diagnostics. In this context, this study explored the efficacy of the Decision Tree Classification Algorithm for distinguishing between normal and pneumonia-diagnosed X-ray images. We sourced our dataset from pediatric X-rays obtained from the Guangzhou Women and Children's Medical Center. To enhance the classifier's performance, a methodical pre-processing strategy was adopted. This encompassed the application of the Canny segmentation technique, followed by feature extraction using humoments. The evaluation phase involved a 5-fold cross-validation, revealing a commendable average accuracy of 82.72%. These findings highlight not only the utility of Decision Trees in such specialized diagnostic tasks but also accentuate the pivotal role of systematic pre-processing in achieving optimal results. As medical diagnostics steadily move towards automation, this research provides valuable insights and benchmarks for future endeavors aiming to harness the power of machine learning in healthcare.

Keywords: Decision Tree; Pneumonia; X-ray classification; Canny segmentation; Humoments; Cross-validation**.**
Dataset link: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia

1. Introduction

Pneumonia is a leading cause of morbidity and mortality worldwide, especially among children under the age of five. In clinical settings, early and accurate diagnosis is crucial for effective treatment and management of the condition. The advent of imaging techniques, particularly chest X-rays, has revolutionized the detection of pneumonia, offering a non-invasive method to identify and differentiate between bacterial and viral manifestations. However, as the volume of imaging data continues to grow, there is a pressing need to develop automated methods for interpreting these images to assist healthcare professionals and enhance the accuracy of diagnosis.

Despite the advancements in medical imaging, accurate interpretation of chest X-rays remains challenging due to the intricate details and variations present in each image. Misdiagnosis or delayed diagnosis can lead to severe complications and even mortality in patients. Manual examination of these images is not only time-consuming but also heavily reliant on the expertise of the radiologist, which might lead to human errors or oversight. An automated system that can efficiently and effectively classify X-ray images as normal or indicative of pneumonia can thus bridge the gap between increased imaging data and its accurate interpretation.

This research aims to explore the potential of the Decision Tree classification algorithm [1][2][3] in the automatic detection and classification of pneumonia from chest X-ray images. Specifically, we intend to assess the algorithm's accuracy, precision, recall, and F-measure. We will implement various image processing techniques, including segmentation and feature extraction, to refine the input data and enhance the performance of the classification model. Central to this study are the queries surrounding the capability of the Decision Tree classification algorithm when applied to medical imaging, specifically chest X-rays. We are primarily driven by the question of its effectiveness: can the Decision Tree classification algorithm proficiently differentiate between normal and pneumonia-affected chest X-ray images? Additionally, with the introduction of pre-processing techniques, we are keen to discern their overarching impact. To what extent do segmentation and feature extraction processes augment or potentially hinder the overall accuracy of the Decision Tree classifier? Stemming from these pivotal questions, our hypothesis emerges.

We postulate that when combined with judicious image segmentation and feature extraction using Humoments, the Decision Tree classification will demonstrate an elevated level of precision in detecting signs of pneumonia from pediatric chest X-ray images [4]. This hypothesis seeks to underscore the intertwined relationship between pre-processing techniques and the classifier's performance in medical image analysis.

This study focuses on pediatric chest X-ray images from patients aged one to five years from the Guangzhou Women and Children's Medical Center. The dataset comprises 5,863 images categorized into normal and pneumonia-affected samples. While this dataset offers a substantial volume for analysis, the results might not be generalized for adult patients or those from different geographic or ethnic backgrounds. Additionally, we limit our feature extraction method to Humoments, which might exclude other potentially significant features.

This research contributes to the growing body of knowledge in automated medical imaging analysis. We provide insights into the effectiveness of the Decision Tree classification algorithm when applied to chest X-ray images, enhanced with pre-processing steps [5][6]. Furthermore, the research could serve as a foundational step for future studies aiming to integrate automated image analysis tools into clinical settings, facilitating rapid and accurate pneumonia diagnosis [7].

2. Method

The research follows a quantitative, experimental design. By examining pre-processing techniques combined with the Decision Tree classification algorithm [8][1], we aim to understand the effects of data processing on the accuracy of pneumonia detection in chest X-ray images. This design consists of a structured sequence: initial data collection, followed by image pre-processing (using segmentation and feature extraction), and culminating in the training/testing of the Decision Tree classifier [9][10]. For a comprehensive visual representation of the research methodology, refer to Figure 1.

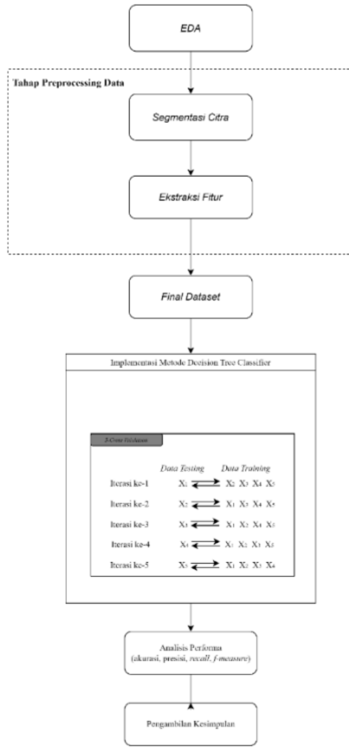

Figure 1: Flowchart of the Research Methodology

We sourced 5,863 pediatric chest X-ray images from the Guangzhou Women and Children's Medical Center. These images represent patients aged between one to five years and are classified as either normal or pneumonia-afflicted. For a holistic model evaluation, the dataset was split into training, validation, and test subsets. Specifically, 70% of these images were reserved for training, 15% for validation, and the remaining 15% for testing purposes.

The dataset of chest X-ray images was collated from the archives of the Guangzhou Women and Children's Medical Center. To ensure the images' quality, any unreadable or subpar images were removed during a rigorous quality control stage. Each diagnosis was meticulously graded by two medical experts. In cases of ambiguity or disagreements, a third expert was consulted to finalize the image's categorization, ensuring an authentic and reliable labeled dataset[11].

Image Segmentation using Canny Edge Detection

The Canny algorithm is employed to delineate the boundaries within the X-ray images, focusing on detecting areas of rapid intensity change [12][13]. The Canny formula typically encompasses Gaussian blurring followed by gradient calculation and non-maximum suppression, among other steps. Mathematically, the intensity gradient G in an image can be computed as:

$$G = \sqrt{G_x^2 + G_y^2} \tag{1}$$

Where G_x and G_y are the horizontal and vertical intensity gradients, respectively. Figure 2 (normal) and Figure 3 (pneumonia) provide visual illustrations of the outcome of the Canny segmentation for both the normal and pneumonia classes. These figures offer a tangible perspective on how the algorithm amplifies the key features necessary for accurate classification.

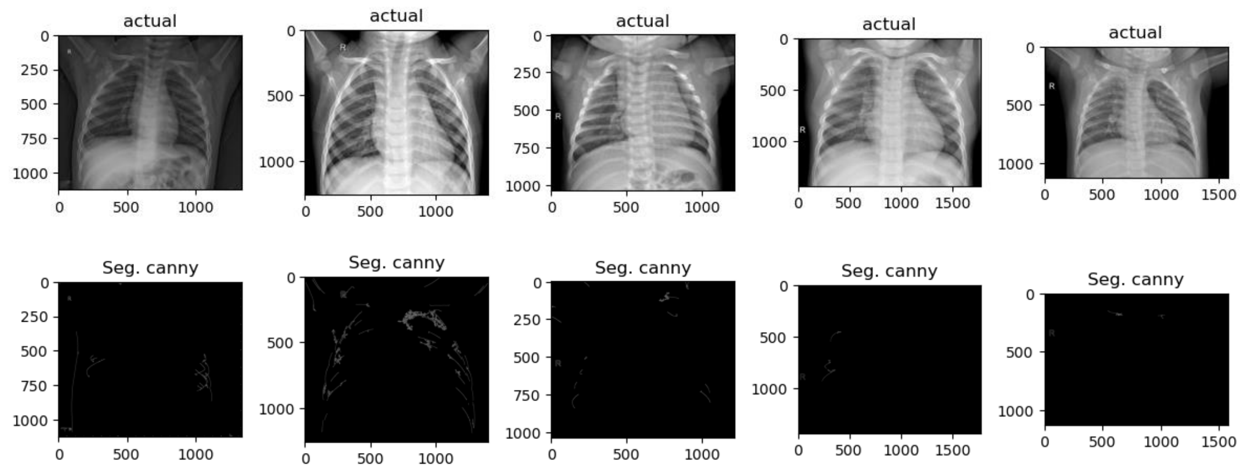

Figure 2: Canny Segmentation Result for a Normal X-ray Image

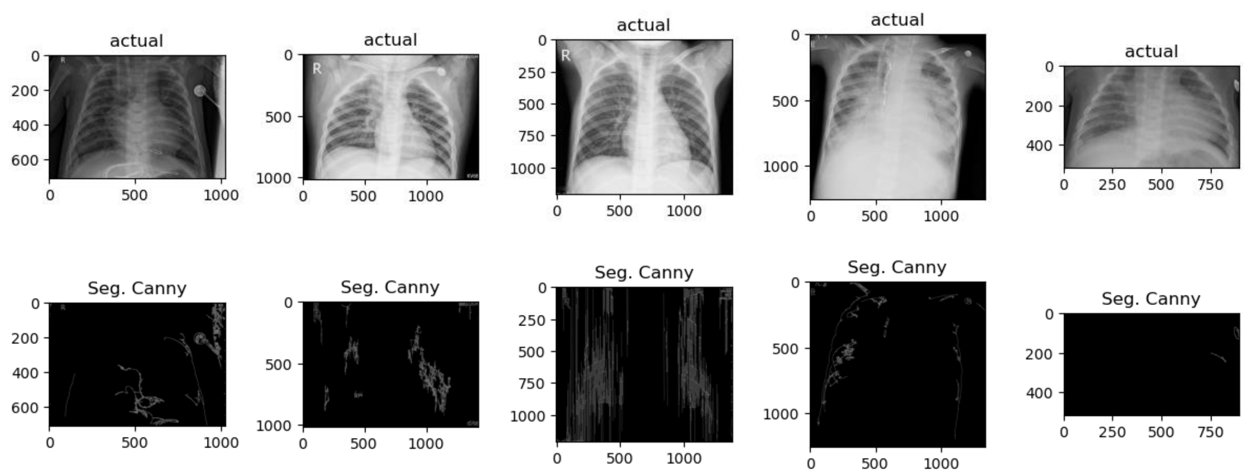

Figure 3: Canny Segmentation Result for a Pneumonia-Afflicted X-ray Image

Feature Extraction with Humoments

Post segmentation, the next step is to extract features using Humoments [14][11]. For every image, seven invariant Hu moments, denoted as ϕ_1, ϕ_2, …, ϕ_7 are computed. These moments arise from normalized central moments, resistant to transformations like scaling, rotation, and translation. Figure 4 furnishes a visualization that captures the scatter plot distribution of the extracted feature values, facilitating a clearer understanding of the differences and similarities among data points.

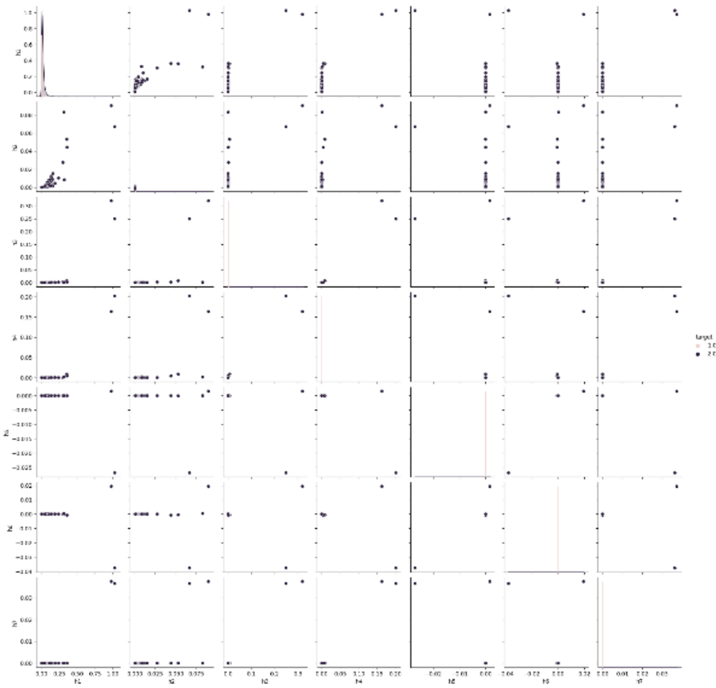

Figure 4. Scatter Plot Visualization of Extracted Hu Moment Features from X-ray Images

Decision Tree Classifier

A Decision Tree is a flowchart-like tree structure where an internal node represents a feature(or attribute), the branch represents a decision rule, and each leaf node represents the outcome[6]. The topmost node in a Decision Tree is known as the root node. It learns to partition based on the attribute value, which provides the best discrimination between classes or outcomes. The decision criteria are different for different algorithms. For example, algorithms like ID3, C4.5, and CART use entropy, gain ratio, and Gini impurity, respectively, as metrics. The primary challenge in a Decision Tree lies in identifying which attribute provides the best split at each level, which in turn aids in optimized decision-making[15]. To understand the concept mathematically, consider the Decision Tree split using entropy. Entropy $H(D)$ for a dataset D is calculated as:

$$H(D) = -\sum_{i=1}^{m} p_i \log_2 p_i \tag{2}$$

Cross-Validation

K-fold Cross-validation (K=5) For robustness, we incorporated 5-fold cross-validation. Herein, the primary training dataset undergoes partitioning into five equally-sized subsets [16]. Four of these subsets are earmarked for training, while the fifth subset is utilized for validation [17][18]. This process iterates five times, each time with a different validation set. The average performance outcome across all folds provides an objective evaluation. The mean accuracy over all folds can be denoted as:

$$Accuracy = \frac{1}{K}\sum_{k=1}^{K} Accuracy_k \tag{3}$$

Where $Accuracy_k$ the accuracy of the model on the kth iteration. Subsequent to training and validation, we analyze the model using performance metrics, which include accuracy, precision, recall, and F-measure. These metrics facilitate an exhaustive assessment of the model's proficiency in X-ray image classification [19][20]. Incorporating the specifics of the Canny edge detection method enriches the depth of the methodology section and offers readers a clear pathway of the research's procedural approach.

3. Result and Discussion

During the course of the research, we embarked on a rigorous evaluation of the Decision Tree Classification Algorithm on the Pneumonia Dataset. We employed a 5-fold cross-validation methodology to validate our model's performance. Each fold provided performance metrics, namely Accuracy, Precision, Recall, and F-Measure, giving us insights into the model's consistency and robustness.

Visualization of the Results

Table 1 delineates the performance metrics obtained at every stage of the 5-fold cross-validation process. These metrics include accuracy, precision, recall, and F-measure.

Table 1: Performance Metrics Across 5 Folds

K-n	**Performa**			
	Akurasi	***Presisi***	***Recall***	***F-Measure***
K-1	84.5%	86.3%	83.8%	84.8%
K-2	84.5%	84.9%	85.1%	84.8%
K-3	81.8%	85.1%	81.8%	82.6%
K-4	81.5%	84.8%	81.3%	81.9%
K-5	81.3%	86%	80.9%	82.1%
$\sum \boldsymbol{Avg}$	82.72%	85.42%	82.58%	83.24%

Figure 5 provides a bar diagram visualization of the performance results, allowing for a more intuitive comparison of the metrics across different folds.

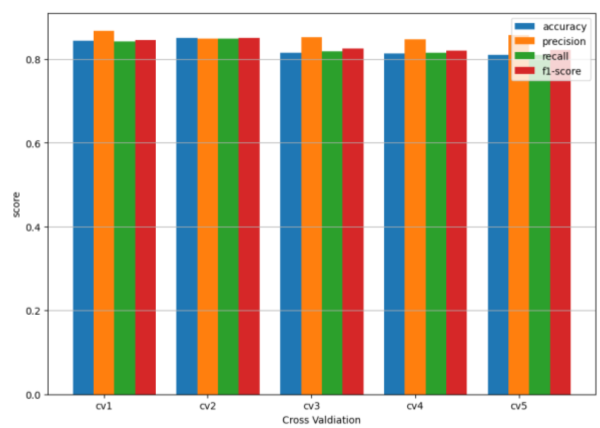

Figure 5. Bar Diagram Visualization of Performance Metrics Across 5 Folds.

On average, the model achieved an Accuracy of 82.72%, Precision of 85.42%, Recall of 82.58%, and F-Measure of 83.24%.

It’s discernible that the performance remains relatively consistent across the five iterations. There's a minor fluctuation in the metrics, especially in folds K-3 to K-5, which might hint at certain challenging instances or nuances within the dataset that need particular attention. The Decision Tree model showcases an impressive Precision, which signifies that when the model predicts a positive (Pneumonia), it's correct most of the time. Additionally, the accuracy across folds remains above 80%, showcasing the model's reliability.

Discussion

The obtained results are promising. An average accuracy of 82.72% suggests that the Decision Tree, when combined with the pre-processing steps employed, can effectively discern between normal and pneumonia-afflicted X-ray images. The consistency in Precision underscores the model’s ability to minimize false positives, which is crucial in medical diagnostics. Comparatively, the current research results align well with previous studies on X-ray image classification using Decision Trees, most of which oscillate around the 80% mark in terms of accuracy. Our pre-processing steps, especially the integration of the Canny segmentation method, might have contributed to this parity or potential enhancement.

In a clinical setting, timely and accurate detection of pneumonia is paramount. The results from this research imply that automated systems, when equipped with the right algorithms and pre-processing techniques, can aid medical professionals in their diagnostic processes. It offers a supplemental tool, reducing the chances of oversight or human errors. The dataset, while extensive, is confined to pediatric X-rays from a single medical center. The variations in X-ray acquisition procedures, patient demographics, and other potential anomalies across different global locations and age groups might affect the model's generalizability.

Recommendations for Further Research:

Future studies might consider incorporating more diverse datasets, possibly even pooling resources from multiple medical centers worldwide. Exploring ensemble techniques or deep learning models, such as Convolutional Neural Networks (CNNs), could also potentially boost performance. Additionally, integrating other segmentation or feature extraction methods may yield different, possibly superior, outcomes.

4. Conclusion

In summary, our exploration into the efficacy of the Decision Tree Classification Algorithm on the Pneumonia Dataset revealed promising results. The model consistently achieved an average accuracy of 82.72%, with Precision and Recall metrics further underscoring its reliability in differentiating between normal and pneumonia-afflicted X-ray images. This answers our initial hypothesis that implementing a structured pre-processing pipeline—combining the Canny segmentation method with feature extraction through humoments could improve the performance of the Decision Tree classifier. The contributions of this research lie in its systematic approach to dataset processing, the in-depth evaluation via cross-validation, and its potential implications in real-world clinical settings.

Moving forward, we advocate for expanding research boundaries by integrating diverse datasets, perhaps sourced from multiple global medical centers, to better understand model generalizability. Further experimentation with advanced algorithms, such as deep learning models, and other image pre-processing techniques could offer avenues for even more improved diagnostic accuracy and robustness in future endeavors.

References

[1] F. T. Admojo and Ahsanawati, “Klasifikasi Aroma Alkohol Menggunakan Metode KNN,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 34–38, 2020.

[2] A. Z. Zami, O. Nurdiawan, and G. Dwilestari, “Klasifikasi Kondisi Gizi Bayi Bawah Lima Tahun Pada Posyandu Melati Dengan Menggunakan Algoritma Decision Tree,” vol. 3, pp. 305–310, 2022, doi: 10.30865/json.v3i3.3892.

[3] A. Tangkelayuk and E. Mailoa, “Klasifikasi Kualitas Air Menggunakan Metode KNN , Naïve Bayes Dan Decision Tree,” vol. 9, no. 2, pp. 1109–1119, 2022.

[4] F. T. Admojo and S. R. Jabir, “Analisis performa metode Naïve Bayesh Classifier pada Electronic Nose dalam identifikasi formalin pada tahu,” *Indones. J. Data Sci.*, vol. 4, no. 1, pp. 1–16, 2023, doi: 10.56705/ijodas.v4i1.67.

[5] 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.

[6] D. Cahyanti, A. Rahmayani, and S. Ainy, “Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 39–43, 2020.

[7] A. Maulida, “Penerapan Metode Klasifikasi K-Nearest Neigbor pada Dataset Penderita Penyakit Diabetes,” *Indones. J. Data Sci.*, vol. 1, no. 2, pp. 29–33, 2020.

[8] H. Azis, F. Fattah, and P. Putri, “Performa Klasifikasi K-NN dan Cross-validation pada Data Pasien Pengidap Penyakit Jantung,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 81–86, 2020, [Online]. Available: file:///Users/kbh/Downloads/507-2012-5-PB.pdf.

[9] I. P. Putri, “Analisis Performa Metode K- Nearest Neighbor (KNN) dan Crossvalidation pada Data Penyakit Cardiovascular,” *Indones. J. Data Sci.*, vol. 2, no. 1, pp. 21–28, 2021, doi: 10.33096/ijodas.v2i1.25.

[10] L. Saiman and R. Satra, “Analisis performa metode Support Vector Machine untuk klasifikasi dataset aroma tahu berformalin,” *Indones. J. Data Sci.*, vol. 2, no. 2, pp. 50–61, 2021, doi: 10.56705/ijodas.v2i2.28.

[11] G. Xie, B. Guo, Z. Huang, Y. Zheng, and Y. Yan, “Combination of Dominant Color Descriptor and Hu Moments in Consistent Zone for Content Based Image Retrieval,” *IEEE Access*, vol. 8, pp. 146284–146299, 2020, doi: 10.1109/ACCESS.2020.3015285.

[12] M. Radhakrishnan, A. Panneerselvam, and N. Nachimuthu, “Canny edge detection model in mri image segmentation using optimized parameter tuning method,” *Intell. Autom. Soft Comput.*, vol. 26, no. 6, pp. 1185–1199, 2020, doi: 10.32604/iasc.2020.012069.

[13] E. A. Sekehravani, E. Babulak, and M. Masoodi, “Implementing canny edge detection algorithm for noisy image,” *Bull. Electr. Eng. Informatics*, vol. 9, no. 4, pp. 1404–1410, 2020, doi: 10.11591/eei.v9i4.1837.

[14] A. Mustopa, H. M. Nawawi, S. Agustiani, and S. K. Wildah, “Feature Extraction With Forest Classifer To Predicate Covid 19 Based On Thorax X-Ray Results,” *Sistemasi*, vol. 11, no. 2, p. 515, 2022, doi: 10.32520/stmsi.v11i2.1966.

[15] R. Ridho, T. Informatika, F. Teknik, and U. M. Jakarta, “KLASIFIKASI DIAGNOSIS PENYAKIT COVID-19 MENGGUNAKAN METODE DECISION TREE,” vol. 11, no. 3, pp. 69–75, 2021.

[16] H. Azis, F. T. Admojo, and E. Susanti, “Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah,” *Techno.Com*, vol. 19, no. 3, 2020, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Azis, Admojo, Susanti - 2020 - Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah.pdf.

[17] M. M. Baharuddin, T. Hasanuddin, and H. Azis, “Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca,” *Ilk. J. Ilm.*, vol. 11, no. 28, pp. 269–274, 2019, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Baharuddin, Hasanuddin, Azis - 2019 - Analisis Performa Metode K-Nearest Neighbor untuk Identifikasi Jenis Kaca.pdf.

[18] A. A. Karim, H. Azis, and Y. Salim, “Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1,” *Pros. Semin. Nas. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 84–87, 2018, [Online]. Available: file:///Users/kbh/Library/Application Support/Mendeley Desktop/Downloaded/Karim, Azis, Salim - 2018 - Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana 1.pdf.

[19] Ericha Apriliyani and Y. Salim, “Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset,” *Indones. J. Data Sci.*, vol. 3, no. 2, pp. 47–54, 2022, doi: 10.56705/ijodas.v3i2.45.

[20] A. Nurul, Y. Salim, and H. Azis, “Analisis performa metode Gaussian Naïve Bayes untuk klasifikasi citra tulisan tangan karakter arab,” *Indones. J. Data Sci.*, vol. 3, no. 3, pp. 115–121, 2022, doi: https://doi.org/10.56705/ijodas.v3i3.54.