Jurnal Teknologi Riset Terapan (JATRA)
Vo 2 , No 2, 2024, 115-122 https://doi.org/10.35912/jatra.v2i2.5011

Classification of Rare Mussaenda Species in Indonesia's Tropical Forests Using the CNN Algorithm

H. F. Muchammad Raja[1*], Meizano Ardhi Muhammad[2], Martinus[3], W. Pandu[4], A. Muhkito[4,] A. Muhammad[5]
University of Lampung, Bandar Lampung, Indonesia[1,2,3,4,5]
muchammad.raja.haikal@students.ac.id[1,2,3,4,5]

Riwayat Artikel
Diterima pada 25 Juni 2024
Revisi 1 pada 1 Juli 2024
Revisi 2 pada 9 Juli 2024
Revisi 3 pada 15 Juli 2024
Disetujui pada 28 Juli 2024

Abstract
Purpose: *Mussaenda frondosa* is a rare plant species native to Indonesia’s tropical forests, with limited research focused on its classification and identification, particularly using machine learning. This study aims to develop a classification model for *Mussaenda* species using a Convolutional Neural Network (CNN) approach to support the advancement of automated plant identification systems.
Methodology/approach: The dataset used consists of 650 labeled images, categorized into six primary parts of the plant: leaves, stems, twigs, fruits, flowers, and trees. A CNN model was developed and trained over 200 epochs to classify the images according to these categories. Preprocessing techniques such as resizing, normalization, and data augmentation were applied to enhance model performance.
Results/findings: The trained CNN model achieved an accuracy of 80%, demonstrating its ability to classify *Mussaenda frondosa* components despite the relatively small dataset. Visual inspection of prediction outputs showed consistent identification across several categories, particularly leaves and flowers.
Conclusion: The results suggest that CNN can be effectively used to classify rare plant species like *Mussaenda frondosa*. The model's performance also indicates that even a limited dataset, when properly processed, can yield promising classification results.
Limitations: The main limitation of this research is the small dataset size, which may restrict the model's generalizability to broader plant species or more diverse environmental conditions..
Contribution: This study contributes to the field of plant classification by providing a foundation dataset and a validated CNN model for rare tropical species. It opens pathways for further research in biodiversity monitoring and conservation using AI.

Keywords: *Cicarimanah Villag, Digital Marketing, Product marketing, TikTok, UMKM.*

How to Cite: Raja,, H. F. M., Muhammad, M. A., Martinus, M., Pandu, W., Muhkito, A., Muhammad, A. (2024). Classification of Rare Mussaenda Species in Indonesia's Tropical Forests Using the CNN Algorithm. *Jurnal Teknologi Riset dan Terapan (Jatra)*, 2(2), 115-122.

1. Pendahuluan

Mussaenda frondosa is one of Indonesia's rare plant species found in tropical forests (Shanthi & Radha, 2020). Despite its significant potential for botanical research and conservation, classification of this species has rarely been conducted using automated methods such as machine learning (Faeni, Puspitaningtyas, & Safitra, 2021). One of the main challenges in classifying this species is the limited availability of datasets (Dewi, Kesiman, Sunarya, Indradewi, & Andika, 2024). Traditional methods for plant identification often require extensive expertise and considerable time, hindering efforts to expedite the identification process. Therefore, an innovative approach is needed to classify this rare species more efficiently and accurately. Research that utilizes deep learning technology, particularly the Convolutional Neural Network (CNN) algorithm, presents a promising solution to improve the accuracy and efficiency of *Mussaenda frondosa* classification (Larese et al., 2014; Tugrul, Elfatimi, & Eryigit, 2022).

The development of artificial intelligence technology, especially in deep learning, has achieved significant breakthroughs in image processing (Dung, 2024). CNN is one of the most effective algorithms for image classification due to its ability to recognize complex patterns (Chompookham & Surinta, 2021; Hajam, Arif, Khanday, & Neshat, 2023). CNN has been applied in various studies for plant classification, including tropical forest species and medicinal plants, with significant results in improving identification accuracy (Arrofiqoh & Harintaka, 2018). Some studies have even achieved accuracy rates of up to 99.12% in leaf image classification (Biswas et al., 2022; Diwedi, Misra, & Tiwari, 2024). This demonstrates CNN's high efficiency in processing diverse image data, particularly in large and heterogeneous datasets (Hu, Chen, Yang, Zhang, & Cui, 2018; Lee, Lim, Song, & Alqahtani, 2023).

However, the application of CNN for classifying rare species like *Mussaenda frondosa* remains limited, despite this technology's proven accuracy in automatically identifying various plant species (Anisman, 2021). This creates a significant opportunity for further research, especially in building visual datasets that support the development of precise automated models (Albakia & Saputra, 2023). Such advancements not only pave the way for innovation in plant classification but also make substantial contributions to digital conservation of rare species through machine learning technology (Boulent, Foucher, Théau, & St-Charles, 2019; Larese et al., 2014).

2. Methodology

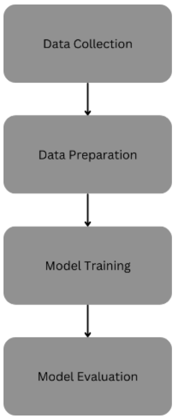

Figure 1: Research Flow

2.1 Data Collection

The subject of this study is *Mussaenda frondosa*, a rare tropical forest species in Indonesia. Image data were collected specifically from various parts of the plant, including leaves, stems, twigs, fruits, flowers, and the main tree structure. A total of 650 high-resolution images were obtained under controlled lighting conditions to ensure quality consistency, enabling precise classification.

2.2 Dataset Preparation

Collected images were standardized and preprocessed. The dataset was split into training (454 images), validation (65 images), and testing (131 images) sets. Data augmentation techniques were applied, including image rotation, cropping, and brightness adjustments, to enhance model generalization and mitigate overfitting risks. Each image was normalized to a fixed size and color scale compatible with the CNN model input requirements.

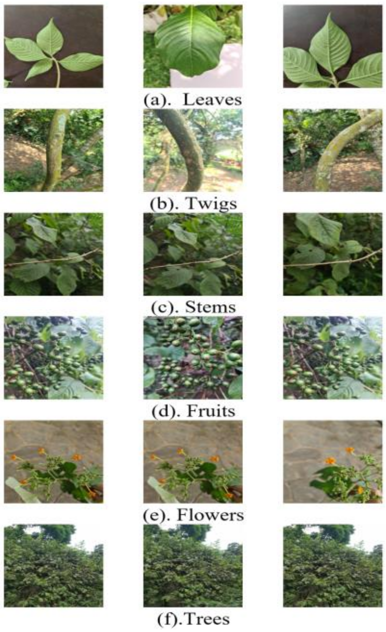

Figure 2: Sample Images of *Mussaenda frondosa* Plants

2.3 CNN Architecture Design

The Convolutional Neural Network (CNN) architecture applied in this study is organized sequentially, designed to process *Mussaenda* frondosa images through various transformation layers. The CNN model is designed with a total of 44,399,174 trainable parameters as can be. This design is aimed at capturing complex visual patterns from plant images, with special attention to computational efficiency and overfitting prevention through the use of dropout. Each layer has a specific function in thefeature extraction process from the input images, progressively capturing visual patterns with increasing levels of complexity. Key layers included:

1. Convolutional Layers: Initial layers with 32 filters and a 3x3 kernel size to extract edge and texture features. The ReLU activation function was used to prevent vanishing gradients.

$$f(x)=max(0,x) \qquad (1)$$

2. Pooling Layers: MaxPooling with a 2x2 pool size, reducing spatial dimensions for computational efficiency while retaining essential features.
3. Flatten and Dropout Layers: Flattening converts the feature map into a one-dimensional array, with a dropout rate of 0.5 to prevent overfitting by randomly deactivating neurons.
4. Output Layer: Comprising six neurons for classification into the six defined classes (leaves, stems, twigs, fruits, flowers, and tree) with a softmax activation function to generate probabilistic outputs.

Table 1: CNN Model Design

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_1 (Conv2D)	(None, 109, 109, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_2 (Conv2D)	(None, 52, 52, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 128)	0
flatten (Flatten)	(None, 86528)	0
dropout (Dropout)	(None, 86528)	0
dense (Dense)	(None, 512)	44302848
dense_1 (Dense)	(None, 6)	3078

Total params: 44399174 (169.37 MB)
Trainable params: 44399174 (169.37 MB)
Non-trainable params: 0 (0.00 Byte)

2.4 Model Training

Model training was conducted using the Adam optimization algorithm, leveraging its adaptive learning rate capabilities to improve convergence speed and handle vanishing gradient issues. Training was set for 200 epochs with a batch size of 32. Each epoch refined model weights to optimize the classification accuracy based on training and validation data performance.

2.5 Model Evaluation

Model performance was evaluated using several metrics:

1. Accuracy: The percentage of correctly classified samples out of the total.
2. Loss: Measurement of prediction error during training.
3. Precision, Recall, and F1-Score: Detailed metrics assessing model sensitivity, specificity, and balanced accuracy, especially useful for imbalanced class data.

3. Result and Discussion

The CNN model trained over 200 epochs showed a clear increase in both training and validation accuracy. Initially, training accuracy was approximately 50%, indicating the early phase of learning visual patterns. Within the first 50 epochs, training accuracy rose sharply to 85%, reaching 95% by epoch 100, and nearly 100% by epoch 200. This progression indicates effective pattern recognition and high classification accuracy on the training data.

Conversely, validation accuracy increased more gradually. Starting at 55% in the first epoch, it rose to around 80% by epoch 200. This slower increase suggests overfitting on the training data, where the model began memorizing patterns specific to the training set, reducing generalization. Additionally, the limited size and variation in the validation dataset required extended epochs for the model to generalize effectively. The gradual rise in validation accuracy indicates the model's adjustment to new patterns present only in validation data, optimizing weights for better accuracy on unseen data. Figure 3 illustrates the trajectory of training and validation accuracy over epochs (Suciati, Simamora, Panusunan, & Fauzan, 2023).

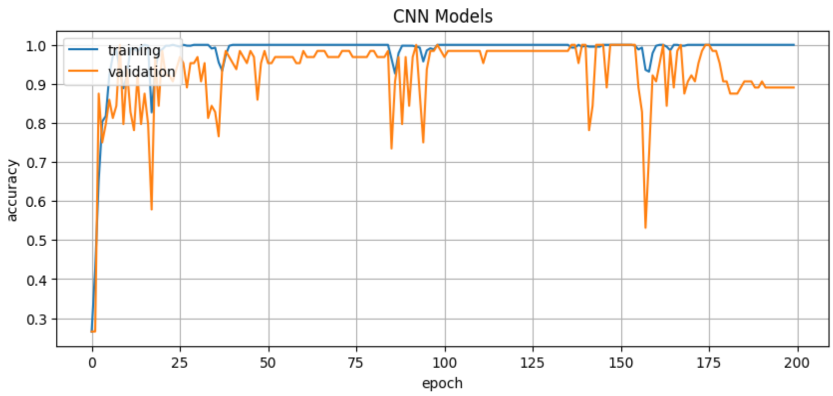

Figure 3: Accuracy Graph

In addition to accuracy, the loss value was also evaluated during training to observe how well the model minimized prediction errors. At the beginning of training, the training loss started at a value of 5, indicating that the model made many errors in classifying images at this early stage. As the number of epochs increased, the training loss significantly decreased, reaching a value of 0.15 by epoch 200 (Makur, Karta, & Oktaviani, 2023). The validation loss showed a similar downward trend, though it was slightly more fluctuating than the training loss. At the first epoch, the validation loss was around 1.80, but it steadily decreased, reaching 0.22 by the final epoch. Small fluctuations in validation loss may be attributed to variations in the validation data, yet the overall downward trend indicates the model's capacity to generalize well to new data (Antika, Rusmana, & Widianingsih, 2023).

It has been proven that both training and validation losses decrease as epochs increase, indicating that the model became increasingly efficient in learning visual patterns from the dataset and correcting classification errors. The slight fluctuations in validation loss also suggest that the model maintained good generalization as can be seen in figure 4.

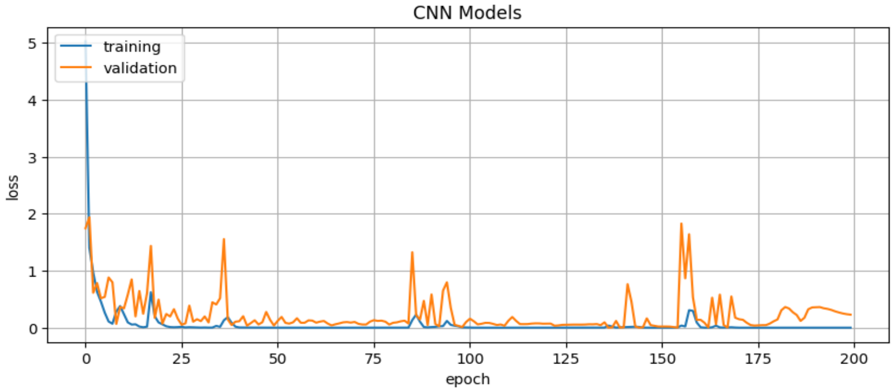

Figure 4: Loss Graph

The increase in accuracy and the decrease in loss indicate the model's improved ability to make accurate predictions. This process enables the model to refine its understanding of data patterns and significantly reduce prediction errors, resulting in a final loss value of 0.22 and an accuracy of 89.23%. The CNN model was trained over 200 epochs with the *Mussaenda frondosa* image dataset, and the results of model training and evaluation are presented in the form of a confusion matrix and performance graphs, including accuracy and loss during training and validation (Fatchurrohman & Saputri, 2023).

The evaluation of CNN model performance in image classification using the confusion matrix illustrates how well the model's predictions align with the true image labels from the test data. Each value on the matrix's diagonal represents correctly classified images, while values outside the diagonal indicate classification errors between certain classes. Darker blue colors signify higher accuracy, meaning more images were correctly classified, while lighter blue areas highlight where misclassifications were more frequent. The test results show that the model successfully classified most classes with high accuracy, especially in the leaf and stem classes, where the model could recognize the specific visual features of these classes. This is reflected in the high diagonal values for these two classes in the confusion matrix (Sutrisno et al., 2023).

In Figure 5, some classification errors are observed in the fruit classes. All fruit samples were classified as trees. This error can be explained by conditions during dataset collection. During image capture, *Mussaenda frondosa* was in its fruiting season, resulting in many tree images being obscured by fruit. This made it difficult for the model to distinguish between tree and fruit accurately, as the fruits visually dominated the tree, leading to the model interpreting them as part of the tree itself. This misclassification is shown by non-diagonal values in the confusion matrix, where some fruit images are detected as trees.

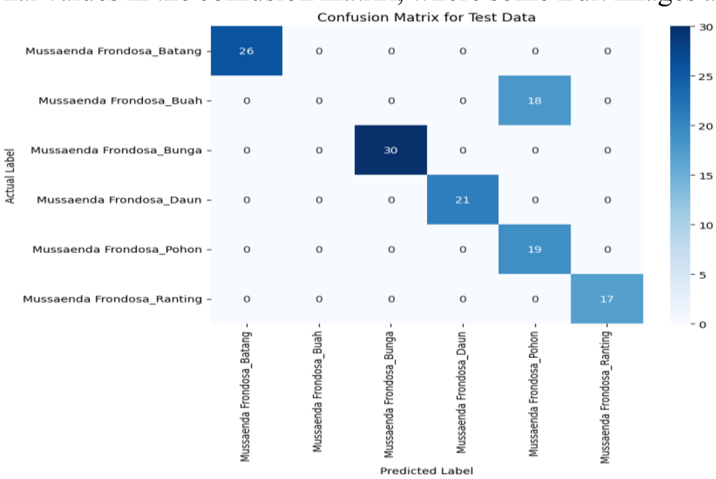

Figure 5: Confusion Matrix

In Table III, the model's performance in classifying several classes, such as Stem, Flower, Leaf, and Twig, shows excellent results, with perfect precision, recall, and F1-score values in certain classes. The model accurately identified each image from these classes without error (Faisol, Paujiah, Russel, & Ramelan, 2022). Conversely, the model demonstrated weakness in classifying images from the Fruit class, with precision and recall values of 0.00, indicating that it could not correctly classify any images from the Fruit class. This misclassification was due to the visual similarity between fruits and trees, especially as the dataset was collected during the fruiting season, with trees obscured by fruit. The Tree class had a lower precision of 0.51 but achieved a recall of 1.00, showing that the model recognized all images in the Tree class, even though half of its predictions were incorrect (Safitri, Husniati, & Permadhy, 2021).

Overall, the model achieved an accuracy of 86%, meaning it correctly classified 86% of the test dataset images. The macro-average values for precision, recall, and F1-score are 0.75, 0.83, and 0.78, respectively, indicating varied performance across classes. The weighted average for precision, recall, and F1-score are 0.79, 0.86, and 0.82, reflecting the model's overall performance after accounting for the data distribution in each class. Although the model performed well in classifying most classes, challenges remain in distinguishing between fruit and tree images.

Table 2: Model Performance

	precision	recall	f1-score	support
Mussaenda Frondosa_Batang	1.00	1.00	1.00	26
Mussaenda Frondosa_Buah	0.00	0.00	0.00	18
Mussaenda Frondosa_Bunga	1.00	1.00	1.00	30
Mussaenda Frondosa_Daun	1.00	1.00	1.00	21
Mussaenda Frondosa_Pohon	0.51	1.00	0.68	19
Mussaenda Frondosa_Ranting	1.00	1.00	1.00	17
Accuracy			0.86	131
macro avg	0.75	0.83	0.78	131
weighted avg	0.79	0.86	0.82	131

4. Conclusion

A Convolutional Neural Network (CNN)-based classification model for the species *Mussaenda frondosa* was successfully developed and implemented using a dataset of 650 images, divided into six classes: leaves, stems, twigs, fruits, flowers, and trees. The model achieved an accuracy of 80% after 200 training epochs, demonstrating that the dataset, despite its limited size, was adequate for effective classification of the species. Through data augmentation and careful architectural design, the model was able to capture distinguishing features across six categories: leaves, stems, twigs, fruits, flowers, and trees. However, one anomaly was detected in the fruit class, where images of fruits were classified as trees. This issue arose due to the conditions during dataset collection, as *Mussaenda frondosa* was in its fruiting season, resulting in many tree images being obscured by fruit, thus impacting classification accuracy. To ensure better dataset quality and reduce the risk of similar misclassifications, it is recommended for further dataset expansion, that dataset collection be aligned with the plant's seasonal stages so that all parts of the plant are clearly visible and well-documented.

The findings confirm that CNNs hold significant promise for the automated classification of rare plant species, contributing to digital conservation and biodiversity efforts. By refining the dataset and implementing additional enhancements, future research could improve model robustness, ultimately supporting conservation initiatives in tropical forest ecosystems. This study thus underscores the value of leveraging deep learning in environmental conservation, offering a feasible path towards efficient and accurate identification of rare species in Indonesia's forests.

Acknowledgment
The authors would like to thank Gita Persada butterfly park and LPPM University of Lampung for their support in dataset collection, as well as the funding agencies that facilitated this research.

References

Albakia, S. A. E., & Saputra, R. A. (2023). Identifikasi Jenis Daun Tanaman Obat Menggunakan Metode Convolutional Neural Network (CNN) Dengan Model VGG16. *Jurnal Informatika Polinema, 9*(4), 451-460. Https://doi.org/10.33795/jip.v9i4.1420

Anisman, H. B. (2021). Analisis Faktor – Faktor yang Memengaruhi Kinerja Keuangan pada Pusat Pendapatan Pemerintah Daerah Kabupaten Tulang Bawang. *Reviu Akuntansi, Manajemen, dan Bisnis, 1*(2), 77-90. doi:10.35912/rambis.v1i2.408

Antika, Z. R., Rusmana, O., & Widianingsih, R. (2023). Analisis Determinasi Minat dan Penggunaan Financial Technology Payment Menggunakan Theory of Planned Behavior: Studi pada Mahasiswa Unsoed. *Jurnal Ilmu Siber dan Teknologi Digital, 1*(2), 111-124. doi:10.35912/jisted.v1i2.2097

Arrofiqoh, E. N., & Harintaka, H. (2018). Implementasi metode convolutional neural network untuk klasifikasi tanaman pada citra resolusi tinggi. *Geomatika, 24*(2), 61. 10.24895/JIG.2018.24-2.810

Biswas, P., Anand, U., Saha, S. C., Kant, N., Mishra, T., Masih, H., . . . Majumder, M. (2022). Betelvine (Piper betle L.): A comprehensive insight into its ethnopharmacology, phytochemistry, and pharmacological, biomedical and therapeutic attributes. *Journal of cellular and molecular medicine, 26*(11), 3083-3119. DOI: 10.1111/jcmm.17323

Boulent, J., Foucher, S., Théau, J., & St-Charles, P.-L. (2019). Convolutional neural networks for the automatic identification of plant diseases. *Frontiers in plant science, 10*, 941. Https://doi.org/10.3389/fpls.2019.00941

Chompookham, T., & Surinta, O. (2021). Ensemble methods with deep convolutional neural networks for plant leaf recognition. *ICIC Express Letters, 15*(6), 553-565. DOI:10.24507/icicel.15.06.553

Dewi, N. P. D. A. S., Kesiman, M. W. A., Sunarya, I. M. G., Indradewi, G. A. A. D., & Andika, I. G. (2024). Klasifikasi Jenis Daun Tumbuhan Herbal Berdasarkan Lontar Usada Taru Pramana Menggunakan CNN. *Techno. Com, 23*(1), 271-283. Https://doi.org/10.62411/tc.v23i1.9510

Diwedi, H. K., Misra, A., & Tiwari, A. K. (2024). CNN-based medicinal plant identification and classification using optimized SVM. *Multimedia Tools and Applications, 83*(11), 33823-33853. DOI:10.1007/s11042-023-16733-8

Dung, K. D. (2024). Leadership, proactive behavior and innovative work behaviors of teachers in Barkin-Ladi. *Annals of Management and Organization Research, 6*(1), 13-24. doi:10.35912/amor.v6i1.1867

Faeni, D. P., Puspitaningtyas, R., & Safitra, R. (2021). Work Life Balance, Peningkatan Karir dan Tekanan Kerja terhadap Produktivitas: Kasus pada Lembaga Sertifikasi Profesi P3 Pembangun Penyuluh Integritas Bangsa. *Studi Akuntansi, Keuangan, dan Manajemen, 1*(1), 45-57. doi:10.35912/sakman.v1i1.602

Faisol, A., Paujiah, S., Russel, E., & Ramelan, M. R. (2022). Pelatihan dan Pendampingan Penggunaan Aplikasi Digital dalam Perencanaan Bisnis dan Keuangan BUMDes. *Jurnal Abdimas Multidisiplin, 1*(1), 35-40. doi:10.35912/jamu.v1i1.1438

Fatchurrohman, M., & Saputri, P. L. (2023). Limitation of Non-Halal Income (Interest) in The Criteria of Sharia Securities List in Indonesia Stock Exchange. *Bukhori: Kajian Ekonomi dan Keuangan Islam, 2*(1), 29-38. doi:10.35912/bukhori.v2i1.1708

Hajam, M. A., Arif, T., Khanday, A. M. U. D., & Neshat, M. (2023). An effective ensemble convolutional learning model with fine-tuning for medicinal plant leaf identification. *Information, 14*(11), 618. Https://doi.org/10.3390/info14110618

Hu, J., Chen, Z., Yang, M., Zhang, R., & Cui, Y. (2018). A multiscale fusion convolutional neural network for plant leaf recognition. *IEEE Signal Processing Letters, 25*(6), 853-857. DOI:10.1109/LSP.2018.2809688

Larese, M. G., Namías, R., Craviotto, R. M., Arango, M. R., Gallo, C., & Granitto, P. M. (2014). Automatic classification of legumes using leaf vein image features. *Pattern Recognition, 47*(1), 158-168. DOI:10.1016/j.patcog.2013.06.012

Lee, C. P., Lim, K. M., Song, Y. X., & Alqahtani, A. (2023). Plant-CNN-ViT: plant classification with ensemble of convolutional neural networks and vision transformer. *Plants, 12*(14), 2642. Https://doi.org/10.3390/plants12142642

Makur, B., Karta, N. L. P. A., & Oktaviani, L. (2023). Pengaruh Electronic Word of Mouth terhadap Kepercayaan dan Keputusan Pembelian pada Aplikasi Shopee Mahasiswa Universitas Triatma Mulya. *Jurnal Bisnis Dan Pemasaran Digital, 2*(1), 25-38. doi:10.35912/jbpd.v2i1.2255

Safitri, L. I., Husniati, R., & Permadhy, Y. T. (2021). Pengaruh Teamwork, Disiplin Kerja, dan Iklim Organisasi terhadap Kinerja Karyawan: Studi di Rumah Sakit X Jakarta Selatan. *Studi Ilmu Manajemen Dan Organisasi, 2*(2), 125-137. doi:10.35912/simo.v2i2.806

Shanthi, S., & Radha, R. (2020). Anti-microbial and phytochemical studies of Mussaenda frondosa Linn. Leaves. *Pharmacognosy Journal, 12*(3). DOI:10.5530/pj.2020.12.94

Suciati, H., Simamora, A. W., Panusunan, P., & Fauzan, F. (2023). Analisa Campuran CPHMA terhadap Penambahan Variasi Aspal Penetrasi 60/70 pada Karakteristik Marshall. *Jurnal Teknologi Riset Terapan, 1*(2), 75-86. doi:10.35912/jatra.v1i2.2294

Sutrisno, P., Debora, D., Destriana, N., Putri, A. T. K. P. S., Marlinah, A., Wijaya, N., & Lekok, W. (2023). Pendampingan Pelatihan Software Akuntansi Accurate dalam Membantu Guru & Siswa-Siswi Smk untuk Meningkatkan Kompetensi dan Profesionalisme. *Jurnal Pemberdayaan Ekonomi, 2*(1), 29-37. doi:10.35912/jpe.v2i1.716

Tugrul, B., Elfatimi, E., & Eryigit, R. (2022). Convolutional neural networks in detection of plant leaf diseases: A review. *Agriculture, 12*(8), 1192. Https://doi.org/10.3390/agriculture12081192