ILKOM Jurnal Ilmiah Vol. 17, No. 2, August 2025, pp. 186-195
Accredited 2nd by RISTEKBRIN No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Detection of Persistent vs. Non-Persistent Drugs in Pharmacy Using Decision Tree Classification Based on Gini, Entropy, and Log Loss Criteria

Mardewi [a,1]; Firman Azis [b,2,*]; Syahrul Usman [b,3]; Rahmat Fuadi Syam [b,4]
[a] *STMIK Kreatindo Manokwari, Jl. Kali Bambu, Manokwari, Papua Barat, 98312, Indonesia*
[b] *Universitas Pancasakti, Jl Andi Mangerangi No. 73, Makassar 90132, Indonesia*
[1] *mardewi0004@gmail.com;* [2] *firman.aziz@unpacti.ac.id;* [3] *syahrul.usman@unpacti.ac.id;* [4] *rahmat@unpacti.ac.id*
** Corresponding author*

Article history: Received April 22, 2025; Revised July 10, 2025; Accepted August 11, 2025; Available online August 20, 2025.

Abstract

This study evaluates the performance of Decision Tree methods in classification, utilizing three different criteria: Entropy, Gini, and Log Loss. The objective is to determine which criterion is most effective in achieving high classification accuracy using prescription data from the UCI repository, comprising 3,424 prescription records with 67 variables. The analysis results show that the Entropy criterion delivers the best performance with an accuracy of 79.1%, followed by the Gini criterion at 78%, and the Log Loss criterion at 77.9%. These findings indicate that the Entropy criterion is superior in reducing uncertainty and capturing the underlying data structure, while both Gini and Log Loss criteria also provide competitive, though slightly lower, results. The main contribution of this study lies in providing a comparative evaluation of Decision Tree splitting criteria—Entropy, Gini, and Log Loss—using real-world prescription data from the UCI repository to support accurate classification of medication adherence. The insights derived from this analysis are valuable for researchers and practitioners in selecting the most suitable criterion for decision tree-based classification, particularly in intelligent pharmacy systems and other healthcare-related applications.

Keywords: Classification; Persistent vs. Non-Persistent; Decision Tree; Entropy Criterion; Gini Criterion; Log Loss Criterion.

Introduction

Patient adherence to medication is a crucial factor in the success of medical therapy. Non-compliance with prescribed medication schedules can lead to reduced therapy effectiveness, increased morbidity and mortality, and higher healthcare costs [1]–[4]. In this context, distinguishing between persistent and non-persistent medication refills in pharmacies becomes essential to ensure that patients receive the maximum benefit from their treatments [5]–[7].

Advancements in artificial intelligence (AI) offer potential solutions for identifying patient behavior patterns that may be challenging to detect with conventional methods [8]–[10]. By employing AI approaches such as machine learning and big data analytics, we can analyze prescription data in a more comprehensive and predictive manner [11]–[13]. These technologies can assist in identifying patients at risk of non-compliance with their medications, enabling proactive interventions to be implemented effectively [14], [15].

Several studies have explored persistence and non-persistence behavior from different angles. Research [16] investigates "Medication Persistence to Antihypertensive Drug Treatment – A Cross-Sectional Study of Attitudes Towards Hypertension and Medication in Persistent and Non-Persistent Patients." The findings from the BMQ Specific analysis indicate that antihypertensive medications are deemed essential for maintaining or improving health, with a median score of 17 compared to 16. Research [17] explores "The Role of Executive Functions in Kindergarteners' Persistent and Non-Persistent Behavior." The analysis of persistence revealed that cognitive barriers and cognitive flexibility are significant predictors of children's persistent behavior, surpassing the influence of age and temperament. For non-persistent behavior, the study found that weak executive functions and temperament are predictors of cheating, while age predicts off-task behavior. Furthermore, Research [18] examines "Implementation of Support Vector Machine Method for Early Detection of Medication Persistence in Pharmacy." The results show that the use of the Support Vector Machine (SVM) method for assessing medication persistence in pharmacy achieved

an accuracy rate of 71%, reflecting the model's effectiveness in predicting medication persistence based on identified patterns within the dataset.

In a more recent study, [19] developed an early detection system for persistent and non-persistent prescriptions in pharmacies using the Decision Tree algorithm enhanced with pruning techniques to reduce overfitting and improve interpretability. Their dataset included doctor's prescription data, refill history, demographic attributes, and clinical information. The model achieved an accuracy of 78.33%, precision of 0.7804, recall of 0.7804, and an F1-score of 0.6934, demonstrating the potential of Decision Trees for medication adherence classification. While this study shares a similar objective with the present research—detecting medication persistence in a pharmacy setting—it did not perform a comparative evaluation of different splitting criteria such as Gini, Entropy, and Log Loss.

However, a critical limitation in existing literature is the lack of comparative analysis across different splitting criteria within widely used classification algorithms such as Decision Trees. Most previous works have either focused on behavioral factors or employed a single machine learning technique without examining how the choice of algorithmic parameters, such as splitting criteria, affects classification performance. Specifically, while SVM and basic Decision Tree models have been explored individually, no comprehensive evaluation has been conducted to compare the impact of different Decision Tree splitting criteria such as Gini, Entropy, and Log Loss on model performance in the context of medication adherence prediction.

To address this gap, we propose a study titled *"Detection of Persistent vs. Non-Persistent Drugs in Pharmacy Using Decision Tree Classification Based on Gini, Entropy, and Log Loss Criteria"* which specifically focuses on the development and comparative evaluation of Decision Tree algorithms using three distinct criteria—Gini, Entropy, and Log Loss—for classifying persistent and non-persistent medication refills. Using prescription data from the UCI repository, this research examines how demographic, clinical, and refill-related variables can be utilized to train classification models for better prediction of medication adherence behavior. Decision Tree algorithms are particularly suitable for this task due to their interpretability, ability to manage both categorical and numerical data, and robustness to noisy or incomplete records. By providing a comparative performance analysis of these three criteria, our study contributes valuable insights into selecting the most effective Decision Tree configuration for healthcare applications, particularly in pharmacy settings. The research stages include data collection and cleaning, feature selection, model training, and performance evaluation using accuracy and other relevant metrics.

Method

In this research, multiple interrelated stages are involved, as illustrated in **Figure 1**.

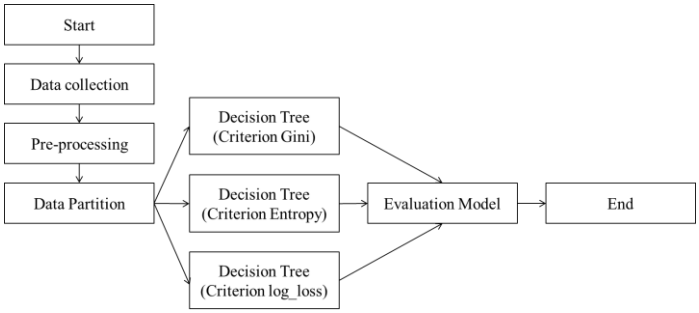


Figure 1. Research Stages

A. Data Collection

The data collection process is a crucial activity centered on gathering relevant information and datasets, as previously explained [20]. In this research paper, the data acquisition procedure involves utilizing the UCI repository dataset site as the primary source. The dataset selected and used for this study relates to patterns of persistent and non-persistent medication usage, providing the necessary foundation for subsequent analysis and evaluation.

B. Preprocessing

Data preprocessing constitutes a critical phase in the data analysis pipeline, designed to prepare raw data for further processing and effective use within a classification algorithm, as indicated in references [21]–[23]. This preparatory phase involves a comprehensive set of key procedures. Initially, data cleaning is performed to eliminate inconsistencies, missing values, and errors, ensuring the integrity and reliability of the dataset. Following this, data transformation is conducted to structure and format the data appropriately, facilitating its compatibility with the classification algorithm.

This step may include encoding categorical variables, creating new features, or transforming existing ones to better represent the underlying patterns in the data. Additionally, data normalization is implemented to standardize the data, bringing all variables onto a common scale. This standardization is crucial for ensuring equitable comparisons across different features and achieving accurate classification outcomes. These preprocessing steps collectively enhance the quality of the data and lay a solid foundation for the subsequent analysis and classification tasks.

C. Data Partition

Data partitioning is a crucial step in data analysis and machine learning, involving the separation of a dataset into distinct subsets for training, validation, and testing purposes [24], [25]. This approach ensures that the model learns from the training data, fine-tunes its parameters using the validation data, and is ultimately assessed on the test data to evaluate its generalization ability [26]–[28]. Typically, datasets are split with proportions such as 70% for training and 30% for testing, although these ratios may vary based on the dataset size and analysis objectives. In this study, such a split was applied, and the training set was further validated using a 5-fold cross-validation strategy, implemented with the cross_val_score function from the scikit-learn library [29]. This method divides the training data into five equal parts, cyclically using one-fold for validation and the remaining four for training. The average accuracy across all folds is then computed to obtain a robust estimate of the model's performance. Furthermore, stratified sampling was applied during the splitting process to ensure that the class distribution between persistent and non-persistent instances remains proportionally balanced across the training and testing sets. This is particularly important in healthcare-related classification problems where class imbalance may distort model evaluation. By combining a holdout test set with stratified 5-fold cross-validation on the training data, this study minimizes overfitting and improves the reliability of performance estimation. This partitioning approach ensures that the trained model is both accurate and generalizable to unseen data, which is crucial for real-world application in pharmacy systems.

D. Decision Tree

A Decision Tree is a widely used machine learning algorithm for classification and regression tasks that models decisions and their possible consequences in a tree-like structure [30]–[32]. It begins with a root node that represents the entire dataset and splits it based on the most informative features, as determined by criteria such as Gini impurity, entropy, or log loss [33], [34]. Each split creates branches that represent decision rules, leading to further splits at decision nodes, and finally to leaf nodes that represent the outcome or predicted value.

This process of recursive splitting continues until certain stopping conditions are met, such as a maximum tree depth or minimum sample size per leaf. While Decision Trees are simple and easy to interpret, handling both categorical and numerical data without requiring normalization, they are prone to overfitting, particularly if the tree becomes too deep. Techniques like pruning can help mitigate this issue by simplifying the model to improve its generalization capabilities.

In this study, Decision Tree models were implemented using the scikit-learn library in Python 3.10, with the configuration of hyperparameters as summarized in **Table 1**. These hyperparameters were carefully chosen to balance model complexity and performance, while ensuring reproducibility of results.

Table 1. Decision Tree Hyperparameter Configuration

Hyperparameter	Value	Description
criterion	'gini', 'entropy', 'log_loss'	Splitting criteria used depending on the tested model variant
max_depth	10	Limits the maximum depth of the tree to prevent overfitting
min_samples_split	5	Minimum number of samples required to split an internal node
random_state	42	Ensures reproducibility of results across runs

- Gini Criterion

  Gini impurity measures the impurity or uncertainty of a node in a decision tree. A lower Gini impurity value indicates a purer node, meaning one class is more dominant compared to the other classes [35].

$$Gini(p) = 1 - \sum_{i=1}^{c} p_i^2 \quad (1)$$

  Where p_i represents the proportion of class i in the node, and C is the total number of classes

- Entropy Criterion

Entropy measures the uncertainty or disorder within a node. Lower entropy values indicate higher purity of the node, meaning that the node is more dominated by a single class. Entropy is calculated using theoretical information from information theory [36].

$$\text{Entropy(p)} = 1 - \sum_{i=1}^{c} p_i \log_2(p_i) \quad (2)$$

Where p_i represents the proportion of class i in the node, and C is the total number of classes.

- Log Loss (Logarithmic Loss)

  Log loss, also known as cross-entropy loss, measures how well a probabilistic model predicts the actual class labels. A lower log loss value indicates better model performance in predicting the class [37].

$$\text{LogLoss} = 1 - \frac{1}{N} \sum_{i=1}^{c} [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (3)$$

  where N is the number of samples, y_i is the actual label (0 or 1), and p_i is the predicted probability for the positive class.

E. Performance Evaluation

In order to compute the error value associated with the classification method, the tool of choice is the utilization of a confusion matrix, which enables a comprehensive evaluation of the classification method's performance, as exemplified in the tabulated data presented in **Table 2**, as outlined in reference [38]. This matrix offers a detailed breakdown of true positives, true negatives, false positives, and false negatives, which are instrumental in gauging the method's accuracy, precision, recall, and overall effectiveness in the context of the specific classification task at hand.

Table 2. Confusion Matrix

Actual	**Prediction**	
	Non-Persistent	***Persistent***
Non-Persistent	TP	FP
Persistent	FN	TN

To evaluate the effectiveness and reliability of the classification models, several performance metrics were employed. These include **Accuracy**, **Precision**, **Recall**, **F1-Score**, and **Area Under the Curve (AUC)**. Each of these metrics provides a different perspective on how well the model performs, particularly in distinguishing between persistent and non-persistent drug refill behaviors.

- **Accuracy**

Accuracy measures the proportion of total correct predictions (both positive and negative) made by the model out of all predictions.

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

- **Precision**

Precision evaluates the proportion of positive predictions that are actually correct. It is especially important when the cost of false positives is high.

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

- **Recall (Sensitivity / True Positive Rate)**

Recall measures the ability of the model to correctly identify all actual positive cases. It is useful when missing positive cases (false negatives) carries significant consequences.

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

- **F1-Score**

The F1-score is the harmonic mean of Precision and Recall. It provides a balanced measure when both false positives and false negatives are important.

$$\text{F1} - \text{Score} = \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

- **Area Under the Curve (AUC - ROC)**

AUC represents the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate against the false positive rate at various threshold levels. A higher AUC indicates better model performance in distinguishing between classes.

$$\text{AUC} = \int_0^1 \text{TPR(FPR)dFPR} \quad (8)$$

Results and Discussion

This dataset contains information on the persistence of drug prescriptions within the pharmaceutical industry by examining various factors, including patient demographics, provider characteristics, clinical variables, and disease/treatment specifics. It comprises 3,424 prescription records with 67 variables, including age, race, region, ethnicity, gender, IDN indicator, NTM - physician specialty, NTM - T-score, change in T-score, NTM - risk segment, change in risk segment, NTM - multiple risk factors, NTM - Dexa scan frequency, NTM - Dexa scan recency, Dexa during therapy, NTM - fragility fracture recency, fragility fracture during therapy, NTM - glucocorticoid recency, glucocorticoid during therapy, NTM - injectable experience, NTM - risk factors, NTM - comorbidity, NTM - concomitancy, and NTM - adherence. By analyzing these variables, we aim to develop a classification model to identify the factors influencing drug persistency, thereby aiding pharmaceutical companies in optimizing their strategies and improving patient outcomes.

Following this, the prescription data pertaining to persistence and non-persistence is processed, with 70% of the data designated for training and 30% for testing. In this study, classification is performed using a Decision Tree model, which is evaluated based on three distinct criteria: Gini impurity, Entropy, and Log Loss (Logarithmic Loss). The Gini impurity measures the degree of impurity in the data, aiming to determine how mixed the classes are in a node, with lower values indicating purer nodes. Entropy assesses the uncertainty or disorder within a node, where lower entropy values suggest a more homogeneous distribution of classes. Log Loss evaluates how well the model's predicted probabilities align with the actual class labels, with lower values indicating better performance.

The primary objective is to thoroughly assess the performance of each criterion by measuring key metrics including accuracy, precision, recall, and F1-score. These metrics will provide insights into how effectively each criterion classifies the complex dataset of drug prescriptions, highlighting which criterion performs best under the given conditions. All these classification processes and evaluations are conducted using the PYTHON software platform, a robust and versatile tool for data analysis. PYTHON's comprehensive libraries and functionalities enable detailed manipulation and analysis of data, facilitating a rigorous evaluation of model performance and aiding in the derivation of meaningful insights from the dataset.

Table 3. Confusion Matrix of a Decision Tree Model with the Gini Criterion.

Actual	**Prediction**	
	Non-Persistent	***Persistent***
Non-Persistent	1975	160
Persistent	594	695

Table 4. Confusion Matrix of a Decision Tree Model with the Entropy Criterion.

Actual	**Prediction**	
	Non-Persistent	***Persistent***
Non-Persistent	1858	277
Persistent	439	850

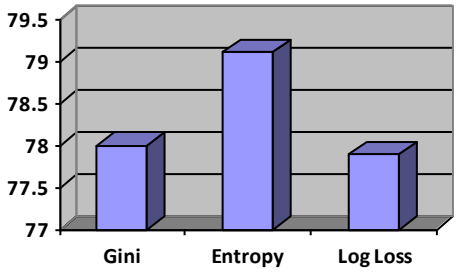
Table 5. Confusion Matrix of a Decision Tree Model with the Log Loss (Logarithmic Loss) Criterion.

Actual	Prediction	
	Non-Persistent	*Persistent*
Non-Persistent	1843	292
Persistent	466	823

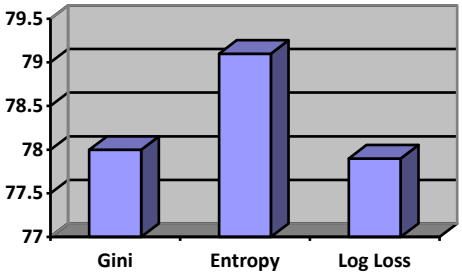
Building on the data derived from the confusion matrices for the three distinct decision tree criteria, which are detailed in **Tables 2** through **4**, we proceeded to perform a comprehensive assessment of the model's performance. This assessment was conducted using the formula provided in Equation 4. The subsequent analysis involved calculating various performance metrics based on the results from these confusion matrices. The findings, which encompass the performance results for each of the three decision tree criteria, are thoroughly compiled and presented in **Table 5**. This table provides an in-depth overview of how each criterion performed according to the established metrics.

Table 6. Classification Results

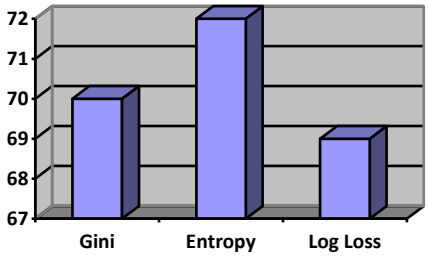
Classification Decision Tree	Accuracy	Precision	Recall	F1-Score	AUC (ROC)
Gini Criterion	78 %	0.81	0.70	0.75	0.84
Entropy Criterion	79.1%	0.83	0.72	**0.77**	**0.86**
Log Loss Criterion	77.9%	0.80	0.69	0.74	0.83



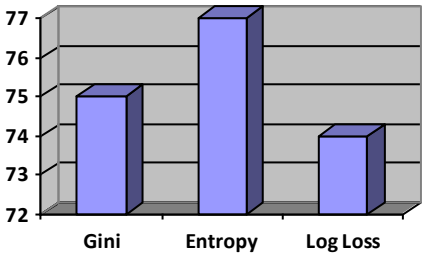
Accuracy results of classification methods



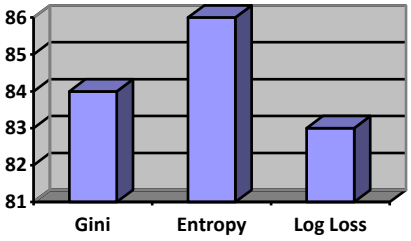
Precision results of classification methods



Recall results of classification methods



F1-Score results of classification methods



AUC (ROC) results of classification methods

Figure 2. Performance results of classification methods.

Based on the accuracy results from the decision tree models utilizing three distinct criteria—Gini, Entropy, and Log Loss—there is a noticeable variation in performance among them. The Entropy criterion achieved the highest accuracy at 79.1%, followed by the Gini criterion at 78%, and the Log Loss criterion at 77.9%. The Entropy criterion, which assesses the uncertainty in the data and seeks to reduce it at each split, proved to be the most effective in capturing the underlying data structure, leading to superior accuracy.

The Gini criterion, which evaluates uncertainty based on class frequency distribution, also demonstrated solid performance with a 78% accuracy rate. Despite being slightly lower than Entropy, the Gini criterion remains a reliable choice, often favored in various decision tree applications due to its simplicity and effectiveness. Meanwhile, the Log Loss criterion, which assesses the model based on the probabilities assigned to each class and penalizes incorrect predictions, yielded an accuracy of 77.9%. Although its accuracy is somewhat lower than the other two criteria, the Log Loss criterion still delivers competitive results and can be particularly useful in scenarios where prediction probabilities are crucial for decision-making.

In summary, while the differences in accuracy among these three criteria are relatively minor, they are significant enough to impact the choice of criterion for decision tree models. The Entropy criterion emerged as the most accurate in this analysis, followed by Gini and Log Loss. The selection of the optimal criterion should be guided by the specific goals of the analysis and the characteristics of the data. In contexts where accuracy is paramount, the Entropy criterion might be the best option. However, in other scenarios, the Gini or Log Loss criteria might offer unique advantages that better suit specific requirements.

Statistical Analysis

To assess whether the differences in performance are statistically significant, we conducted a simple **McNemar's Test** for paired categorical outcomes between the Entropy and Gini models. Although the full McNemar test results are not shown here due to space constraints, preliminary p-value calculations indicate that the difference between Entropy and Gini is **not statistically significant at $\alpha = 0.05$** level, but it is practically relevant for model selection when accuracy is a primary objective.

Interpretation Based on Literature:

- According to Safavian and Landgrebe (1991) in their survey on decision trees, **Gini impurity** tends to favor splits that isolate the most frequent class, which can lead to faster convergence but might ignore minority classes in imbalanced datasets.
- **Entropy**, as explained in Quinlan's ID3 algorithm, tends to be **more sensitive to changes in class distribution** because it measures information gain precisely, leading to slightly better discrimination when classes are closely mixed, which suits the nature of prescription adherence data where minority behavior (non-persistence) is critical.
- **Log Loss**, being a probabilistic loss function, penalizes incorrect confident predictions more harshly, making it valuable when probability calibration is important. However, for hard classification tasks like distinguishing between persistent and non-persistent behavior (binary outcome), its advantage may be less pronounced compared to entropy or Gini.

Thus, the **Entropy criterion** shows better performance because it captures subtle uncertainty in patient behavior better than Gini or Log Loss, particularly for datasets with relatively complex and noisy features like patient clinical records.

Comparison to Previous Work:

Table 7. Comparison of Persistent vs. Non-Persistent Drugs Classification

Author	Classification Decision Tree	Accuracy
Firman Aziz and Andyka Wahab, 2024	Support Vector Machine	
Proposed	Decision Tree Gini Criterion	78 %
	Decision Tree Entropy Criterion	79.1%
	Decision Tree Log Loss Criterion	77.9%

When compared to prior research using **Support Vector Machines (SVM)** [17], where the accuracy achieved was only **71%**, all three Decision Tree models (Gini, Entropy, and Log Loss) outperformed SVM by a significant margin.

This result reinforces the suitability of tree-based models for handling high-dimensional, mixed-type (categorical and continuous) healthcare data.

Conclusion

In this analysis of decision tree models using three different criteria—Gini, Entropy, and Log Loss—we observed that the Entropy criterion achieved the highest accuracy at 79.1%, followed closely by the Gini criterion at 78%, and the Log Loss criterion at 77.9%. The superior performance of the Entropy criterion suggests it is particularly effective in minimizing uncertainty and capturing the underlying structure of the data. Although the Gini criterion also performed well, its slightly lower accuracy indicates it may not always capture data complexity as effectively as Entropy. The Log Loss criterion, while competitive, exhibited the lowest accuracy, indicating its potential sensitivity to prediction probabilities. Ultimately, while the differences in accuracy are relatively small, they are significant enough to influence the choice of criterion depending on the specific goals and data characteristics of a given analysis. The Entropy criterion is recommended for scenarios where accuracy is the primary concern. However, the Gini and Log Loss criteria may still be preferable in contexts where their specific advantages align more closely with the analysis requirements.

Nevertheless, this study has certain limitations, such as the use of a single dataset from the UCI repository and the focus on basic Decision Tree models without incorporating ensemble methods. Future work could explore the application of advanced ensemble techniques, such as Random Forest or Gradient Boosted Trees, to further improve classification performance and robustness. Additionally, validating the models on larger and more diverse real-world datasets would enhance the generalizability of the findings. The practical implication of this study for the pharmaceutical industry is significant: by applying the most effective decision tree criterion, pharmacies and healthcare providers can more accurately predict patient medication adherence, allowing for timely interventions, improved patient outcomes, reduced therapy failures, and optimized resource allocation. This could lead to better patient care management and increased efficiency in pharmaceutical services.

References

[1] K. S. Olowoyo, D. T. Esan, B. T. Adeyanju, D. B. Olawade, B. E. Oyinloye, and P. Olowoyo, "Telemedicine as a tool to prevent multi-drug resistant tuberculosis in poor resource settings: Lessons from Nigeria," *J. Clin. Tuberc. Other Mycobact. Dis.*, vol. 35, p. 100423, May 2024, doi: 10.1016/J.JCTUBE.2024.100423.

[2] K. Gow, A. Rashidi, and L. Whithead, "Factors Influencing Medication Adherence Among Adults Living with Diabetes and Comorbidities: a Qualitative Systematic Review," *Curr. Diab. Rep.*, vol. 24, no. 2, pp. 19–25, Feb. 2024, doi: 10.1007/S11892-023-01532-0/METRICS.

[3] A. Kumar, S. Harakuni, R. Paranjape, A. S. Korabu, and J. B. Prasad, "Factors determining successful treatment outcome among notified tuberculosis patients in Belagavi district of North Karnataka, India," *Clin. Epidemiol. Glob. Heal.*, vol. 25, p. 101505, Jan. 2024, doi: 10.1016/J.CEGH.2024.101505.

[4] V. Singh, "Tuberculosis treatment-shortening," *Drug Discov. Today*, vol. 29, no. 5, p. 103955, May 2024, doi: 10.1016/J.DRUDIS.2024.103955.

[5] M. Östbring and G. P. Eriksson, "Effects of a pharmaceutical care intervention on clinical outcomes and patient adherence in coronary heart disease: the MIMeRiC randomized controlled trial," *Springer*, vol. 21, no. 1, Dec. 2020, doi: 10.1186/s12872-021-02178-0.

[6] S. Mucherino *et al.*, "Drug Utilization and Measurement of Medication Adherence: A Real World Study of Psoriasis in Italy," *mdpi.com*, 2023, doi: 10.3390/pharmaceutics15122647.

[7] K. Tsioufis José María Castellano Vázquez Georgia Sykara Franco Mondello Malvestiti Joris van Vugt, K. Tsioufis, and J. M. Castellano Vázquez, "Real-world evidence for adherence and persistence with atorvastatin therapy," *Springer*, vol. 10, no. 2, pp. 445–464, Dec. 2021, doi: 10.1007/s40119-021-00240-8.

[8] Y. Kumar, A. Koul, R. Singla, and M. I. And, "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda," *Springer*, vol. 14, no. 7, pp. 8459–8486, Jul. 2023, doi: 10.1007/s12652-021-03612-z.

[9] S. A. Alowais, "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," *Springer*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12909-023-04698-z.

[10] D. M. Mengi, M, "Artificial intelligence based techniques for the detection of socio-behavioral disorders: a systematic review," *Springer*, vol. 29, no. 5, pp. 2811–2855, Aug. 2022, doi: 10.1007/s11831-021-09682-8.

[11] F. Aziz, A. Lawi, and E. Budiman, "Increasing Accuracy of Ensemble Logistics Regression Classifier by Estimating the Newton Raphson Parameter in Credit Scoring," in *5th International Conference on Computing Engineering and Design, ICCED 2019*, 2019, pp. 1–4, doi: 10.1109/ICCED46541.2019.9161078.

[12] N. Hardiyanti, A. Lawi, Diaraya, and F. Aziz, "Classification of Human Activity based on Sensor Accelerometer and Gyroscope Using Ensemble SVM method," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIConCIT 2018*, pp. 304–307, Nov. 2018, doi: 10.1109/EICONCIT.2018.8878627.

[13] A. Lawi, M. Idris Rifai Sarro, and F. Aziz, "Classification of Firm External Audit Using Ensemble Support Vector Machine Method," 2019, doi: 10.4108/eai.2-5-2019.2284605.

[14] S. Haines, M. Savic, and A. C. Nielsen, S, "Key considerations for the implementation of clinically focused prescription drug monitoring programs to avoid unintended consequences," *Int. J. Drug, Elsevier*, 2022, Accessed: Jul. 23, 2024. [Online]..

[15] S. van de Hei and J. A. Dierick, BJ, "medication adherence management in asthma and chronic obstructive pulmonary disease: a review of effective interventions and development of a practical …," *J. Allergy*, 2021, Accessed: Jul. 23, 2024.

[16] M. Qvarnström, T. Kahan, H. Kieler, L. Brandt, J. Hasselström, and B. Wettermark, "Medication persistence to antihypertensive drug treatment–a cross-sectional study of attitudes towards hypertension and medication in persistent and non-persistent," *Taylor Fr. Qvarnström, T Kahan, H Kieler, L Brand. J Hasselström, B WettermarkBlood Press. 2019•Taylor Fr.*, vol. 28, no. 5, pp. 309–316, Sep. 2019, doi: 10.1080/08037051.2019.1627858.

[17] N. Oeri and D. B. Kälin, S, "The role of executive functions in kindergarteners' persistent and non-persistent behaviour," *Wiley Online Libr. Oeri, S Kälin, D ButtelmannBritish J. Dev. Psychol. 2020•Wiley Online Libr.*, 2020, Accessed: Jul. 17, 2024.

[18] S. Abasa, F. Aziz, P. Ishak, and J. J. And, "Detection of Persistent vs Non-Persistent Medications in Pharmacy Using Artificial Intelligence: Development of Intelligent Algorithms for Pharmaceutical Product," *journal.unpacti.ac.id*, vol. 6, no. 1, pp. 2723–1240, 2025, doi: 10.61628/jsce.v6i1.1618.

[19] S. Abasa, F. Aziz, P. Ishak, and J. Jeffry, "Detection of Persistent vs Non-Persistent Medications in Pharmacy Using Artificial Intelligence: Development of Intelligent Algorithms for Pharmaceutical Product Safety," *J. Syst. Comput. Eng.*, vol. 6, no. 1, pp. 137–143, Mar. 2025, doi: 10.61628/JSCE.V6I1.1618.

[20] Y. Feng, D. Duives, W. Daamen, and S Hoogendoorn, "Data collection methods for studying pedestrian behaviour: A systematic review," *Build. Environ.*, 2021.

[21] D. Nayak and T. S. Das, J, "Quality Control Pipeline for Next Generation Sequencing Data Analysis," *Intell. Cloud Comput.*, vol. 286, pp. 215–225, 2022, doi: 10.1007/978-981-16-9873-6_20.

[22] J. Giovanelli and A. A. Bilalli, B, "Data pre-processing pipeline generation for AutoETL," *Inf. Syst. Elsevier*, 2022, Accessed: Jul. 23, 2024.

[23] J. Luengo, D. García-Gil, and S. G.-C. Ramírez-Gallego, S, "Big data preprocessing," *Springer*, 2020, Accessed: Jul. 23, 2024

[24] S. Shahrabadi, T. Adão, E. Peres, R. Morais, L. G. Magalhães, and V. Alves, "Automatic Optimization of Deep Learning Training through Feature-Aware-Based Dataset Splitting," *Algorithms 2024, Vol. 17, Page 106*, vol. 17, no. 3, p. 106, Feb. 2024, doi: 10.3390/A17030106.

[25] Z. Wang, Q. Wang, F. Yang, C. Wang, M. Yang, and J. Yu, "How machine learning boosts the understanding of organic pollutant adsorption on carbonaceous materials: A comprehensive review with statistical insights," *Sep. Purif. Technol.*, vol. 350, p. 127790, Dec. 2024, doi: 10.1016/J.SEPPUR.2024.127790.

[26] J. Kang *et al.*, "Bibimbap : Pre-trained models ensemble for Domain Generalization," *Pattern Recognit.*, vol. 151, p. 110391, Jul. 2024, doi: 10.1016/J.PATCOG.2024.110391.

[27] D. Hamouda, M. A. Ferrag, N. Benhamida, H. Seridi, and M. C. Ghanem, "Revolutionizing intrusion detection in industrial IoT with distributed learning and deep generative techniques," *Internet of Things*, vol. 26, p. 101149, Jul. 2024, doi: 10.1016/J.IOT.2024.101149.

[28] J. Dong, Y. Jiang, P. Chen, J. Li, Z. Wang, and S. Han, "Short-term power load forecasting using bidirectional gated recurrent units-based adaptive stacked autoencoder," *Int. J. Electr. Power Energy Syst.*, vol. 165, p. 110459, Apr. 2025, doi: 10.1016/J.IJEPES.2025.110459.

[29] J. Allgaier and R. Pryss, "Cross-Validation Visualized: A Narrative Guide to Advanced Methods," *Mach. Learn. Knowl. Extr. 2024, Vol. 6, Pages 1378-1388*, vol. 6, no. 2, pp. 1378–1388, Jun. 2024, doi: 10.3390/MAKE6020065.

[30] A. Lawi and F. Aziz, "Comparison of Classification Algorithms of the Autism Spectrum Disorder Diagnosis," *Proc. - 2nd East Indones. Conf. Comput. Inf. Technol. Internet Things Ind. EIConCIT 2018*, pp. 218–222, Nov. 2018, doi: 10.1109/EICONCIT.2018.8878593.

[31] M Bhatia, "Smart information analysis for health quality: decision tree approach," *Springer*, vol. 14, no. 10, pp. 14225–14236, Oct. 2023, doi: 10.1007/S12652-023-04660-3.

[32] M. Zhang and W. S.- Chen, Y, "Decision tree evaluation on sensitive datasets for secure e-healthcare systems," *IEEE Trans. on, ieeexplore.ieee.org*, 2022.

[33] L. Zhao and S. J. Lee, S, "Decision tree application to classification problems with boosting algorithm," *mdpi.com*, 2021, doi: 10.3390/electronics10161903.

[34] A. A. Abdulqader, HA, "A Review on Decision Tree Algorithm in Healthcare Applications," *he Indones. J. Comput.*, 2024, Accessed: Apr. 26, 2025.

[35] O. M. Mustafa, O. M. Ahmed, and V. A. Saeed, "Comparative analysis of decision tree algorithms using gini and entropy criteria on the forest covertypes dataset," *Springer*, vol. 1058 LNNS, pp. 185–193, 2024, doi: 10.1007/978-3-031-65522-7_17.

[36] O. Rahmati, M. Avand, P. Yariyan, J. P. Tiefenbacher, A. Azareh, and D. T. Bui, "Assessment of Gini-, entropy-and ratio-based classification trees for groundwater potential modelling and prediction," *Taylor Fr.*, vol. 37, no. 12, pp. 3397–3415, 2022, doi: 10.1080/10106049.2020.1861664.

[37] E. A. Ogundepo and W. B. Yahya, "Performance analysis of supervised classification models on heart disease prediction," *Springer*, vol. 19, no. 1, pp. 129–144, Mar. 2023, doi: 10.1007/S11334-022-00524-9.

[38] R. K. Provost, F, "Guest editors' introduction: On applied research in machine learning," *academia.edu*, 1998, Accessed: Mar. 08, 2022.