Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal
Vol. 7, No 1, April 2026, pp. 37~52
ISSN: 2722-0001, DOI: DOI : 10.55311/aiocsit.v7i1.392

Monitoring Expiration and Beyond-Use Dates of Non-Compounded Drugs Through Rule-Based and Machine Learning Approaches

Marfuah[1], Masparudin[2], Raymond Erz Saragih[3]
[1]Department of Information Systems, Universitas Universal, Indonesia
[2]Department Software Engineering, Universitas Universal, Indonesia
[3]Department Informatics Engineering, Universitas Universal, Indonesia

ABSTRACT

Despite various educational campaigns, public awareness concerning the safe consumption period of medications remains inadequate. Consumers predominantly depend on the Expiration Date (ED) printed on the original packaging, representing the manufacturer's shelf life. Yet, upon opening, the medication's validity shifts to the Beyond Use Date (BUD). The BUD is dictated by several factors, including the medication category, duration of exposure, physical state, odor, and discoloration. To address this issue, this study implements a hybrid approach combining a Rule-Based Expert System (utilizing IF-THEN logic) with machine learning algorithms. Specifically, the research evaluates the performance of Decision Tree (DT), Gaussian Naive Bayes (GNB), and K-Nearest Neighbor (KNN) models. Experimental results indicate that the DT model achieved the highest accuracy at 95%, followed by GNB (94%) and KNN (81%). These predictive models demonstrate significant potential in tracking medication viability, thereby reducing pharmaceutical waste, improving patient compliance, and mitigating the risk of adverse medication errors.

Keyword : Rule-Based Expert System, Decision Tree, Gaussian Naïve Bayes, K-Nearest Neighbor, Monitoring ED/BUD.

Corresponding Author:
Marfuah,
Department of Information Systems,
Universitas Universal,
Pasir Putih Kelurahan Sadai, Kecamatan Bengkong, Kota Batam, Kepulauan Riau 29444, Indonesia.
Email: marfuah@uvers.ac.id

1. INTRODUCTION

The practice of storing medicines is a common behavior within the community, leading to a relatively high prevalence of drug storage at the household level [1]. The storage of medicines for self-medication is strongly discouraged [2]. Drugs are often stored for emergency purposes, treatment of acute illnesses, or chronic conditions. Non-compounded drugs are frequently perceived as safe to use without a physician's prescription [3] however, this does not imply that their use is entirely free from risk [4]. Improper storage and use of medicines at home may have an impact on public health and the environment, as well as increasing the risk of inappropriate drug use [5].

At present, the public still frequently makes mistakes in obtaining, using, storing, and properly disposing of medicines [6]. This condition may lead to various negative consequences, such as medication errors, suboptimal therapeutic outcomes, and reduced drug effectiveness [7]. Adverse consequences in patient care frequently stem from medication errors, a specific class of medical mistakes primarily driven by the improper administration of drugs [8]. Every year, these pharmacological oversights are responsible for over 7,000 fatalities [9]. To mitigate these risks, the Indonesian Ministry of Health mandates strict Pharmaceutical Service Standards across clinical and community pharmacy settings. This regulation compels pharmacists to demonstrate robust clinical proficiency and proper drug handling capabilities to guarantee therapeutic efficacy and patient security [10].

Despite these regulatory frameworks, a significant number of dispensing pharmacists fail to consistently deliver verbal counseling or affix labels regarding the Beyond Use Date (BUD) [11]. The BUD designates the specific timeframe during which a medication remains chemically stable and safe for consumption once its original seal is broken [11]. Consequently, awareness and comprehension of the BUD concept remain profoundly deficient among the Indonesian general public. According to research on perceptions of expiration dates, approximately 97% of respondents were unaware of BUD, and overall, 100% of informants had never received BUD information from pharmacists [12]. Additionally, drug labelling is often inconsistently applied [13]. Some respondents perceive the Beyond Use Date (BUD) to be the same as the Expiration Date (ED) [12]. The Expiration Date (ED) refers to the shelf life indicated on the primary packaging of a drug [14]. This represents the period during which a drug product may be used after manufacture by the pharmaceutical company [15]. However, few people are aware that once the primary packaging of a drug is opened, the usage period is no longer based on the ED [16]. However, several rules are associated with the concept of BUD. Among these, the rules consider the drug type, form, aroma, and color, with the time period starting from when the drug packaging is opened and generally having a shorter usage limit [17]. If multiple drugs have been opened, it becomes increasingly complex to remember when their usage periods expire. Healthcare teams have made efforts to educate the public regarding this issue [18] to improve knowledge regarding the usage period of drugs [19]. Considering that health quality and status are individual responsibilities [20]. Health is one of the most essential components of life [21].

Previous studies have developed mobile applications to detect drug expiration based on the ED [22]. However, a significant limitation of prior research is its exclusive focus on the ED without comprehensively considering the BUD. The methodologies previously employed typically relied on simple rule-based systems lacking the integration of adaptive artificial intelligence. While various machine learning-based studies have been utilized for data classification, they are generally not combined with rule-based expert systems, which limits their ability to explicitly represent expert knowledge. Furthermore, there is a scarcity of research that combines a Rule-Based Expert System (RBES) approach with machine learning algorithms within a unified hybrid system to enhance accuracy while preserving decision interpretability, particularly in the context of drug monitoring based on ED and BUD.

The novelty of this study lies in incorporating ED and BUD parameters alongside changes in the drug's physical condition—specifically color, shape, and odor—into a singular decision-making framework. This research integrates an RBES with machine learning algorithms, namely DT, GNB, and KNN, to monitor medication usage based on ED and BUD. This hybrid approach is designed to improve accuracy while maintaining the transparency inherent in rule-based decision-making processes. This method is employed due to the presence of a set of rules (IF–THEN) [23] involved in determining drug usage limits. The application of expert systems in the healthcare domain [24], particularly in regulating the usage limits of medicines, can provide patients or users with more objective and timely information. A rule-based expert system requires knowledge representation [25], which can be developed by humans through reasoning or as a result of changes that occur [8]. A rule-based expert system is an artificial intelligence method that relies on a set of logical rules to perform diagnosis or provide recommendations [26]. An expert system is a branch of Artificial Intelligence [27]. Rule-Based Expert Systems (RBES) are used in conjunction with machine learning to achieve better results, thereby creating a hybrid approach (RBES Mixed approach) [28]. This research focuses not only on enhancing predictive accuracy but also contributes to the development of a more transparent, adaptive, and relevant decision support system aimed at improving medication safety within the community.

2. RELATED WORKS

In 2021, a study was conducted on public perceptions of the Beyond Use Date (BUD) of medicines and the role of pharmacists in providing relevant information, particularly in Jakarta, Indonesia, a principal and developing city. Indonesia currently lacks data on the public's awareness of this matter. Based on responses from 60 informants, 97% were unaware of BUD, and 100% had not received this information from pharmacists. Furthermore, 50% of respondents perceived that the expiration date printed on the primary packaging is the same as the BUD, which is attributed to low awareness of BUD[12].

In 2023, a study examined the implementation of expiration date interventions by pharmacists in Indonesia, focusing on pharmacists' obligation to provide BUD information to patients to ensure drug stability and enhance medication safety. The study aimed to evaluate the implementation of this obligation using a cross-sectional design. The sample comprised pharmacists with experience in patient care. The results showed that, among 221 respondents, between 32% and 54% acknowledged not consistently providing BUD labelling and verbal information to patients. This indicates that not all pharmacists communicate BUD information orally to patients[29].

Studies on Expiration Date (ED) have been conducted to develop a mobile application for detecting medication expiration dates. The use of a mobile application is well-suited to daily activities, making this medium an appropriate platform for detecting ED. Expiration date labels on medicines and packaging are often prone to tearing, wrinkling, or fading. The mobile application is designed using QR code technology to facilitate the detection process. This system identifies the expiration dates of medications and alerts patients when medications have expired. Additionally, the application displays other information, such as the medicine name, function, manufacturer, and price[22].

Clinical decision-making has also been conducted using rule-based methods. Decision Trees and Decision Tables are employed to represent knowledge, as this approach is a primary focus in healthcare services. It has been observed that this method can effectively generate accurate and efficient decisions [25].

A rule-based expert system has also been implemented for real-time monitoring of criminal case processing, with several rules based on time intervals. These rules are distributed among three institutions: the police, public prosecutors, and district courts. The application of this rule-based expert system facilitates the monitoring of criminal case resolution based on 17 implemented rules, thereby enabling the accurate delivery of information regarding the case resolution process according to the existing facts[26].

Further research involves developing a diagnostic system to predict the likelihood of diabetes mellitus. To process the empirical data, this research introduced a framework that integrates decision tree algorithms with criteria formulated by domain experts. Ultimately, a Rule-Based Expert System (RBES) was constructed from a total collection of 1,800 records. To evaluate the model, the dataset was partitioned by assigning 75% (1,350 instances) to the training phase, while reserving the remaining 25% to test its predictive performance. The results indicate that the hybrid RBES outperforms traditional RBES, which relies solely on expert rules [28]. In addition, several studies have investigated machine learning approaches or hybrid RBES methods, as summarized in Table 1.

Table 1. Previous Research

Author	Dataset	Approach	Accuracy %	Precision %	Recall %	F1-Score %
Mustafa, et al 2023[30]	Heart disease dataset	Decision Tree (DT)	100	100	100	100
		Gaussian Naive Bayes (GNB)	98	97	100	98
	Breast Cancer Coimbra Dataset	DT	83	88	81	82
		GNB	75	78	75	74
	Early-stage diabetes risk prediction dataset	DT	96	95	96	95
		GNB	94	86	97	91
	WISDM smartphone and smartwatch activity and biometrics dataset	DT	86	84	83	84
		GNB	76	70	69	68
Maduri, et al 2024[31]	Crop yield prediction	DT	98	-	-	-
		GNB	99	-	-	-
		KNN	98	-	-	-
Aldi, et al 2022[32]	Drug/Pharmaceutical dataset	KNN	98	99	96	97

Building on prior studies, this work proposes an innovative approach to monitor medication usage deadlines based on expiration and beyond-use dates by integrating RBES with machine learning techniques (Decision Tree, Gaussian Naive Bayes, and K-Nearest Neighbors), which have proven effective for classifying real-world data. In this work, a rule-based system is used to determine initial constraints. At the same time, machine learning models are employed to improve classification accuracy in the monitoring of safe-to-use and expired medicines. The three methods used in this work are presented as follows:

1. DT is employed to monitor the medication usage deadlines based on expiration and beyond-use dates. Historical data are processed together with rule-based expert systems to improve classification accuracy. This approach ensures transparency in the validation process, which is

essential for ensuring medication safety.
2. GNB is employed to provide a probabilistic approach for monitoring drug usage deadlines based on expiration and beyond-use dates. The model leverages historical data to estimate expiration probabilities based on the recognized patterns within the dataset.
3. KNN The KNN algorithm measures the similarity between current expiration data and existing historical records, thereby assisting in determining the safety and suitability of the medication. KNN can identify patterns in the dataset that may be difficult to discern with alternative methods.

Building upon previous research, this study introduces an innovative approach to pharmaceutical management and enhances medication safety by implementing more adaptive models.

3. METHODOLOGIES

A. Non-Compounded Drugs

Non-compounded medicines are those sold in prepackaged forms, mass-produced by pharmaceutical manufacturers, and do not require a special prescription from a doctor [33]. These medicines already have standardized dosages and have been approved by the relevant drug and food regulatory authorities. Non-compounded medication can be purchased at pharmacies or drugstores without a prescription. However, some specific medicines may still require recommendations or advice from healthcare professionals to ensure proper use in accordance with the individual's health condition. Non-compounded medicines include:
1. Pain relievers
2. Cough medicines
3. Allergy medications
4. Vitamins and supplements
5. Topical medications such as ointments or creams

The advantage of non-compounded medicines lies in their accessibility and convenience, as consumers can purchase and use them without complicated procedures. However, it remains essential to follow the usage instructions on the packaging label or consult a healthcare professional if necessary. In this study, non-compounded medicines refer to those commonly stored at home by the public, with examples provided in Table 2 [17], [34].

Table 2. Drug [35]

No	Drug	Type of Drug	Expired BUD
1	Amoxsan	Dry Syrup (reconstitution)	14 Day
2	Sanmol	Non-reconstituted Syrup	90 Day
3	Omeprazole	Repackaged tablets/capsules	180 Day
4	Betamethason	Cream (ointment)	90 Day
5	Rohto	Multidose eye drops	28 Day

B. Rule-Based Expert System

To emulate human problem-solving capabilities, a rule-based expert system encodes domain knowledge and logical frameworks through conditional IF-THEN statements, mimicking the cognitive processes of a specialist [25]. For the purpose of this research, the core of the system is built upon a comprehensive medication database coupled with expiration guidelines sourced directly from pharmacy professionals. Consequently, specific temporal limits are assigned to each medication category through tailored logical constraints. The architectural design of this proposed system is visually detailed in Fig. 1.

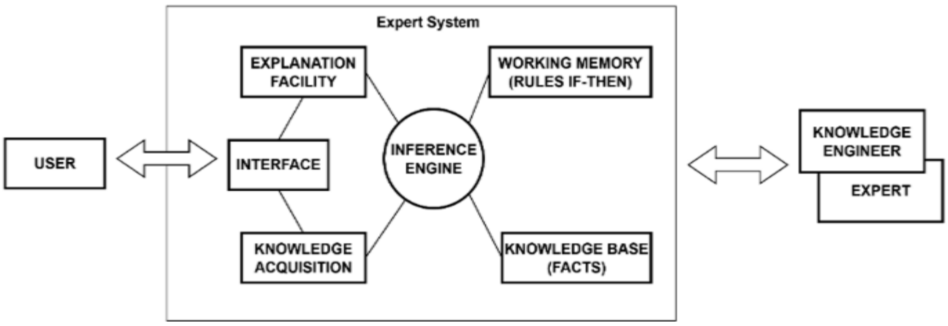


Fig 1. Expert System Structure [36]

Based on Fig. 1, this study focuses on three components: working memory, knowledge base, and inference engine, which are developed by the knowledge engineer in collaboration with the expert. This focus stems from the study's emphasis on rule construction and testing. The three components are described as follows[37]:

1. Working memory acts as a temporary repository, holding the factual data required while the conditional rules are actively processed.
2. Knowledge Base is the database within the expert system that stores facts related to ED and BUD, as well as rules used for reasoning to solve problems, allowing specific conditions to be inferred.
3. The Inference Engine functions as the analytical core, applying predefined logical directives from the knowledge base to deduce final outcomes or formulate recommendations. Forward chaining is applied as the reasoning method.

A working memory is constructed, which is then tested against the knowledge base connected to the inference engine to produce the appropriate output—the knowledge base searches for conformity with the rules established as a guideline for expert reasoning in decision-making[27]. Fig. 2 illustrates the knowledge base.

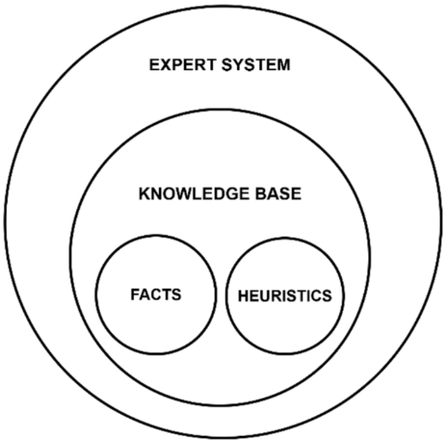


Fig 2. Representation of the knowledge base[36]

The integration of Clinical Decision Support Systems has become highly indispensable, most notably within the medical domain. Furthermore, the exponential daily growth of information necessitates robust administrative frameworks capable of organizing and analyzing these massive datasets to extract actionable insights [25].

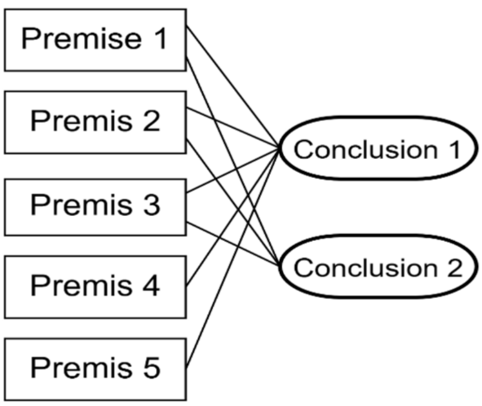


Fig 3. Knowledge graph [38]

In Fig. 3, Premise 1 represents Non-reconstituted Syrup; Premise 2 represents Time <=90 days; Premise 3 represents Color does not change; Premise 4 represents Unchanged Form; Premise 5 represents Smell Doesn't Change; Conclusion 1 represents Safe; and Conclusion 2 represents Expired. The generated sequence of rules is as follows:

IF Premise 1 AND Premise 2 AND Premise 3 AND Premise 4 AND Premise 5 THEN Conclusion 1
IF Premise 1 AND Premise 2 AND Premise 3 THEN Conclusion 2

A premise refers to a fact that must be true before a particular conclusion can be drawn. Rules regarding the usage period of non-compounded medicines are formulated using logical operators, specifically the AND operator.

C. Decision Tree (DT)

Decision Tree (DT) is a machine learning algorithm that falls under the category of supervised learning [39]. The creation of a decision tree is performed through training using a dataset [40]. Architecturally, this classification model originates from a primary root node, which subsequently splits into distinct pathways denoting the target categories for the analyzed problem [25]. Consequently, a decision tree can be regarded as an accumulation of acquired information. The creation begins by identifying the feature with the highest correlation to the class label. The most relevant feature is selected at the root node, followed by iteratively selecting subsequent features based on their relevance. During inference, new instances traverse the tree path according to their feature values to predict the final class label. In this work, the Decision Tree is configured to have a minimum sample size of 1 for leaves and a split of 2. The maximum depth of the tree was not set to make the model more flexible in finding the optimal depth. The Gini function was used to determine the quality of the split. A random state of 1 was set to guarantee the randomness of the generated decision tree.

D. Gaussian Naïve Bayes (GNB)

Serving as a specialized iteration of the standard probabilistic model, Gaussian Naïve Bayes (GNB) is exclusively engineered to accommodate continuous data attributes. It operates on the assumption that features (x_i) within each class (C_k) are independent and follow a normal (Gaussian) distribution[41]. As a probability-based classification algorithm, By analyzing the training instances, this algorithm computes the statistical mean and variance for every distinct attribute across the target categories [42]. The only configuration used in this work regarding the usage of GNB is the variance smoothing, which was set to 1×10^9.

E. K-Nearest Neighbor (KNN)

The K-Nearest Neighbors (KNN) method uses both the training and test datasets to compute pairwise distances between data points. The algorithm operates by measuring spatial proximity between instances, using a specified distance threshold to determine the similarity between the test sample and the training data. If all identified neighbors belong to a single class, the instance is classified into that class according to the established data parameters. The mathematical formulation of the KNN algorithm is as follows [43]:

1. z = (x', y'): where x' represents the attribute vector of the test instance, and y' denotes the unknown class label of the test instance
2. d (x', y'): The current phase entails computing the spatial proximity separating the unclassified sample z from every existing data point, the training dataset, with the results stored in the set D
3. Neighborhood selection: select the subset Dz ⊆ D, consisting of the K samples with the smallest distance to z. The final classification y' is derived from the majority class among the neighbors.

In this work, the number of neighbors used to determine the class of unknown or new data was set to 5, and the metric used to compute distances between data instances was the Minkowski distance with a power parameter of 2, yielding Euclidean distance.

F. Data Collection

Data on the usage period of non-compounded drugs, based on the Expiration Date (ED) and Beyond Use Date (BUD), were collected from various sources, including pharmaceutical literature and experts in pharmacy, to ensure the accuracy and completeness of the information. Based on the collected data, the rules for the Expired Date before the package is opened refer solely to the Expiration Date indicated on the primary packaging. The Expiration Date is translated into the MaximumTime constraint. The system employs two distinct knowledge representations:

IF Time Passes ≤ MaximumTime THEN Safe
IF Time Passes ≥ MaximumTime THEN Expired

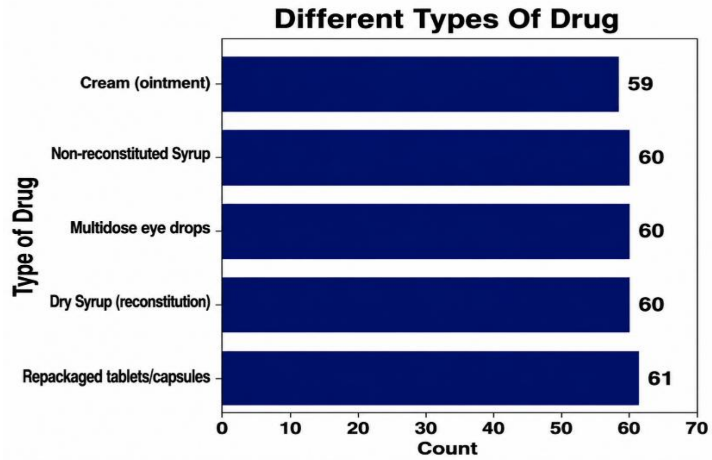
Once the original seal is breached, the safety evaluation shifts to the Beyond Use Date (BUD) framework, wherein the expiration timeline is calculated commencing on the exact day the container is unsealed. Based on the collected data, the rules constructed for Beyond Use Date consist of 16 rules, two conclusions, and 45 types of knowledge representations. The rules and conclusions are presented in table 3.

Table 3. Rule And Conclusion [35]

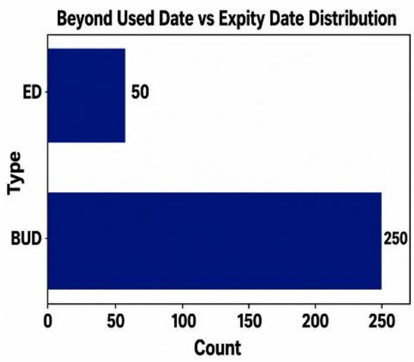
No	Rule	Variable Rule
1	Dry Syrup (reconstitution)	Type Of Drug
2	Non-reconstituted Syrup	
3	Repackaged tablets/capsules	
4	Cream (ointment)	
5	Multidose eye drops	
6	Time ≤ 14 days	Maximum Time
7	Time ≤ 90 days	
8	Time ≤ 180 days	
9	Time ≤ 28 days	
10	Opening Date	Time Passes
11	Production Date	
12	Color does not change	-
13	Unchanged Form	-
14	Smell Doesn’t Change	-
15	Expired Date (ED)	Type
16	Beyond Use Date (BUD)	
No	**Conclusions**	
1	Safe	
2	Expired	

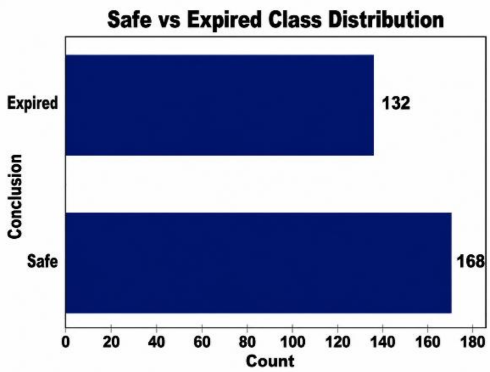
Based on Table 3, the parameters are defined as follows:
1. Type of Drug: Determine the expiration limit, as it may vary according to the specific drug category.
2. MaximumTime: represents the maximum duration for which the medication is considered safe for use.
3. Time Passes: an incremental time metric relative to MaximumTime. It is calculated from the Production Date for Expiration Dates (ED) or the Opening Date for Beyond-Use Dates (BUD). This metric assesses whether the threshold is exceeded and computes the remaining safety duration.
4. Physical Stability (Color, Form, and Odor): Factors such as “Color Does Not Change,” “Unchanged Form,” and “Smell Doesn’t Change” are critical monitoring indicators. If any of these physical properties change, the medication is classified as expired, even if the MaximumTime has not yet been reached.
5. Type: Classified into ED and BUD to distinguish whether the monitoring applies to sealed or opened packaging.

This study utilizes a dataset of 300 records. The collected dataset was divided by allocating 70% of the records to train the algorithms, while the remaining 30% served as an independent holdout sample to validate the model's predictive accuracy on unseen data. This split ensures the model learns from diverse data patterns and is evaluated objectively. The dataset was generated through expert-based rule simulations conducted in this study. The current sample size of 300 instances is deemed sufficient to adequately represent the model's performance. The characteristics of the utilized dataset are illustrated in Fig. 4.



(a)





(b) (c)

Fig 4. Data Characteristics
(a) Type Of Drug (b) ED vs BUD (c) Conclusion

Fig. 4(a) illustrates the distribution of data across the five types of medications used. Fig. 4(b) delineates the characteristics of ED versus BUD; ED applies to medications that rely on the expiration date to determine shelf life, indicating that the primary packaging remains intact or unopened. Conversely, BUD dictates the allowable timeframe—determined by the specific type of medication—used to establish the expiration limit for drugs whose primary packaging has been opened. Finally, Fig. 4(c) displays the data distribution based on the final classifications of "expired" and "safe".

4. RESULTS AND DISCUSSION

RBES is employed to establish the initial constraints for the machine learning models to optimize the resulting outcomes [28]. However, the development of the RBES encountered significant data-acquisition challenges, particularly the difficulty of obtaining expert knowledge due to their limited availability and demanding schedules. This led to a prolonged knowledge representation process. While the RBES defines the foundational boundaries, machine learning models (Decision Trees, GNB, and KNN) are used to improve classification accuracy in monitoring the safety status and expiration of medications.

The machine learning models used in this work are DT, GNB, and KNN. In the DT model, knowledge representation is implemented using rule-based Decision Trees, which are the primary focus in healthcare services. Figure 4 provides a simplified depiction of the knowledge representation, making the information more intuitive [44]. To predict whether a product is expired or safe based on several contributing factors, the Decision Tree classification model is employed, as illustrated in Figure 4. Each rule specifies value ranges for the input variables, which must be taken into account during the expert system's inference process [45]. The root node of this decision tree model is "Color Doesn't Change" (No ≤ 0.5), which performs the initial split into two branches (True and False). The True branch contains 210 samples and leads toward an "Expired" classification with a Gini index of 0.495. This value indicates a higher degree of impurity, suggesting that the data in this branch is non-homogeneous and requires further partitioning. Conversely, the False branch contains 51 samples, all of which are classified as "Safe" with a Gini index of 0, indicating a pure class with zero impurity. The Gini index is a metric for evaluating classification performance; a lower value, such as 0 in this "Safe" class, indicates higher purity and greater accuracy.

The second split occurs within the True branch, where data is further partitioned based on the feature "Smell Doesn't Change" (No ≤ 6.5). The True path at this node yields an "Expired" classification with a Gini index of 0.4, accounting for a significant portion of the data (159 samples). The False path is subdivided into two conditions, both of which result in a "Safe" classification, with Gini values of 0. The Decision Tree continues to branch through conditions such as "Unchanged Form" (No ≤ 8.5), "Time Passed" (≤ 94.5), and "Maximum Time" (≤ 82.5). The majority of instances are classified as "Expired," whereas "Safe" classifications occur only when specific criteria are met. This suggests that certain features[46], [47] particularly elapsed time and drug type strongly correlate with safety status, whereas other factors are associated with expiration. The terminal branches use finer conditions, such as Maximum Time ≤ 98.5 or Time Passed ≤ 184.5, to determine the "Expired" or "Safe" classification of the pharmaceutical products. This can reduce the likelihood of inference errors or incorrect outputs [48].

Fig. 5 presents the Decision Tree model, which is effective for this work. The model demonstrates a robust classification performance, providing strong indications that a product will be classified as "Expired" when specific conditions – primarily those related to temporal factors and drug type – are met. The low Gini index observed across most "Safe" classifications confirms that the tree successfully differentiates between safe and expired products with high precision.

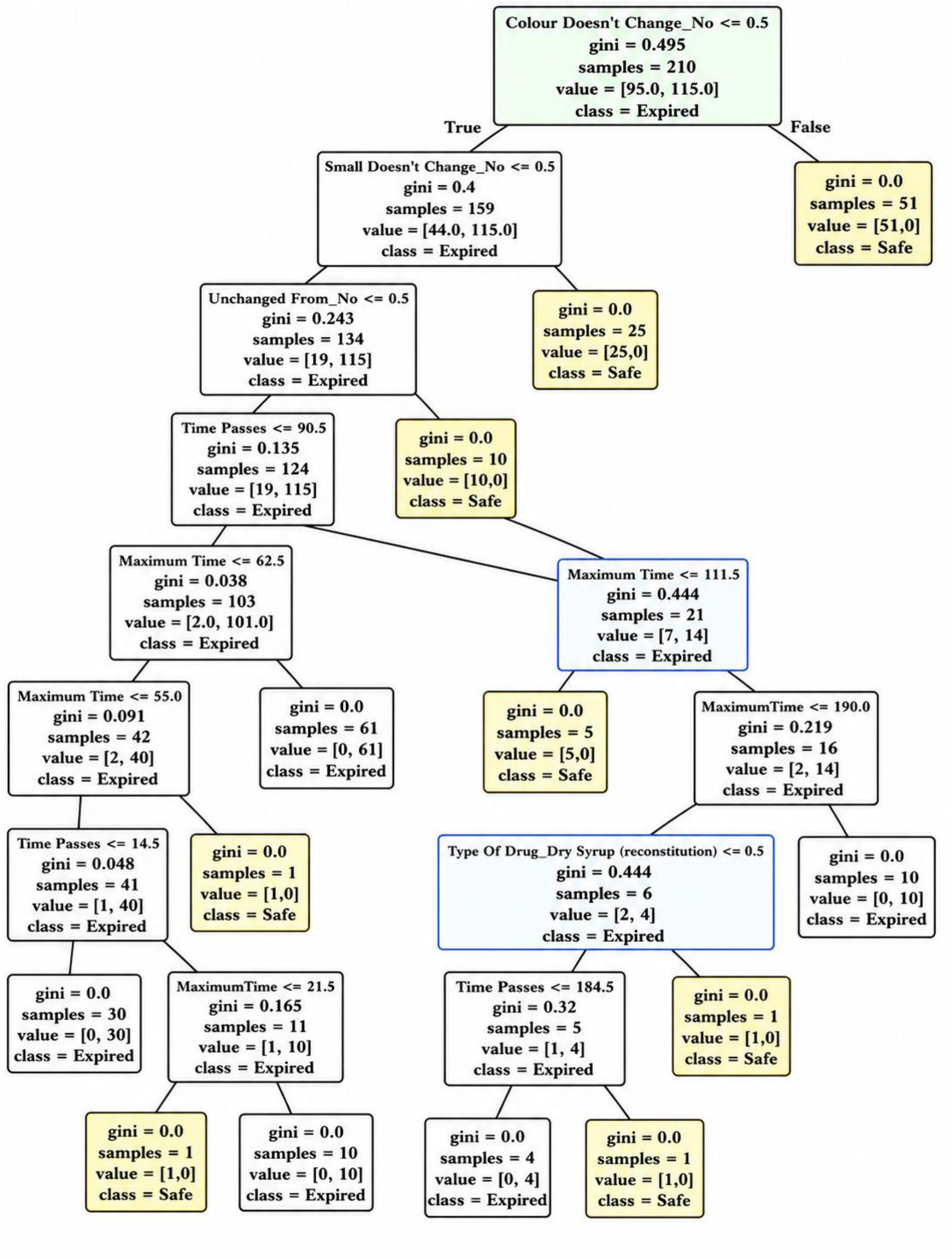
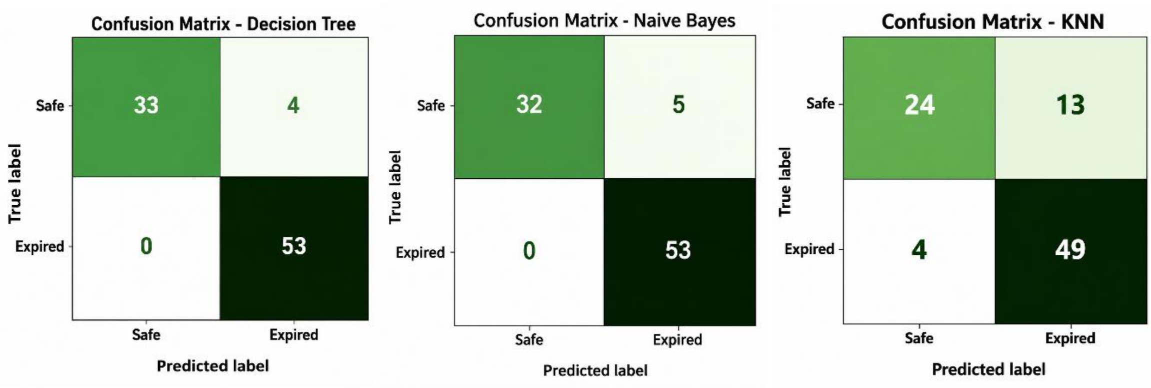


Fig 5. Decision Tree

The Decision Tree model demonstrates an exceptional ability to identify expired medications; specifically, all 53 samples in the "Expired" category were identified perfectly. Regarding the "Safe" category, 33 samples were correctly identified, while the remaining 4 samples were misclassified as "Expired". These findings demonstrate that the classifier exhibits a heightened responsiveness toward the safe usage boundaries. Despite this specific tendency, the system successfully preserves a robust global prediction rate of 95%, as evidenced by the matrix depicted in Fig. 6 (a).



(a) (b) (c)

Fig 6. Confusion Matrix
(a) Decision Tree (b) GNB (c) KNN

As for the GNB model, its performance is comparable to that of the DT model, achieving an accuracy of 94%. All 53 expired medication samples were correctly identified. Within the “Safe” category, 32 samples were correctly classified, while 5 samples were misclassified. These results indicate that the model has a stronger ability to identify expired medications than those within the safe usage timeframe, as illustrated in Fig. 6 (b).

The KNN model demonstrated the lowest performance in identifying both expired medications and those within the safe usage timeframe. Of 53 expired samples, 49 were correctly identified, and four were misclassified. For the “Safe” category, 24 out of 37 samples were correctly identified, with 13 samples misclassified as expired. Consequently, this model achieved an overall accuracy of 81%, as illustrated in Figure 6 (c). Table 4 details the performance of each model in terms of accuracy, which measures the overall correctness of the model's predictions; precision, which evaluates the exactness of positive predictions (e.g., the proportion of medications predicted as expired that are genuinely expired); recall (or sensitivity), which assesses the model's detection capability by determining the proportion of actual expired medications that were correctly identified; and the F1-Score, which provides a balanced measure of the model's performance derived from the precision and recall values. Together, these metrics provide a comprehensive evaluation of the models' predictive capabilities.

Table 4. Results RBES and Machine Learning

Approach	Accuracy	Precision	Recall	F1-Score
DT	95%	96%	95%	95%
GNB	94%	96%	93%	94%
KNN	81%	82%	78%	80%

As shown in Table 4, the DT algorithm achieved superior performance among the three methods across all evaluation metrics. DT achieved an accuracy of 95%, indicating that it correctly classified 95%of the total test instances. GNB performed closely, with 94% accuracy, maintaining good performance. Lastly, KNN achieved an efficacy of 81%, indicating that it was less effective than DT and GNB in this case. High accuracy is imperative for a drug expiration monitoring system; frequent misclassifications would render the model incapable of reliably detecting medication expiration in real-world applications.

In terms of precision, both DT and GNB excelled with a score of 96%. This high precision indicates that 96% of the positive predictions generated were accurate. In contrast, KNN yielded lower precision (82%), indicating a higher rate of false-positive predictions. When a model exhibits a high precision is implemented within this framework, it ensures that the drug expiration monitoring system can generate highly exact and reliable predictions..

Regarding recall, DT recorded the highest value at 95%, followed by GNB at 93%. KNN produced a significantly lower recall of 78%, suggesting a tendency to miss a substantial number of positive instances. Furthermore, a high recall rate ensures that the model deployed within this system can effectively and accurately detect expired medications. The balance between precision and recall is captured by the F1-Score, with DT leading at 95%. Robust model performance signifies an optimal balance and high predictive accuracy. This is important for the proposed system: deploying the best-performing model ensures the drug expiration monitoring system is highly reliable in practice. This confirms that DT is not only precise in its positive predictions but also robust in identifying total positive cases. GNB maintained a strong performance of 94%, while KNN dropped to 80%. A rigorous analysis of model performance is essential to ensure the deployment of the optimal model. This guarantees that the system not only accurately predicts medication expiration but also safeguards users against the adverse consequences of inaccurate predictions, thereby preventing critical errors in medication use. The operational schemes of these three methods are illustrated in Fig. 7.

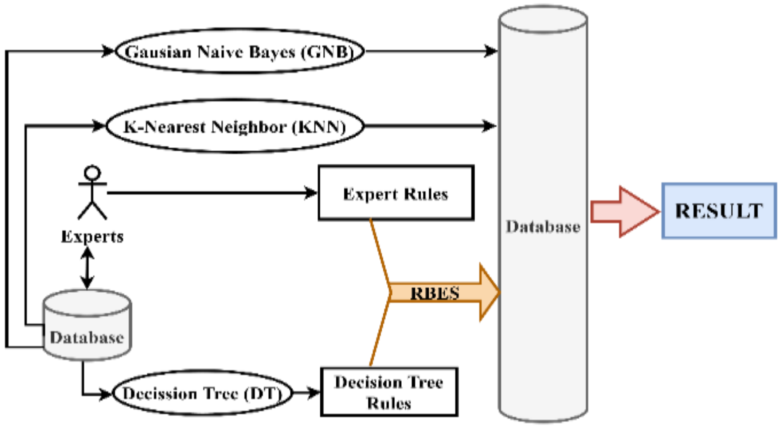


Fig 7. Scheme of the work

DT demonstrated an excellent performance across all evaluation metrics. This indicates that DT is more effective at producing accurate, precise classifications and has a stronger ability to capture true positive cases. This is primarily because the DT algorithm's inherent interpretability offers significant added value: the model's logic is easily explained and understood, making it an ideal choice for applications that demand a transparent decision-making process. Furthermore, DT is a foundational methodology in expert systems because of its hierarchical tree structure. Knowledge representation is seamlessly organized within an IF-THEN rule-based framework, making this method particularly well-suited to the specific data characteristics of this study. In the healthcare field, the 95% accuracy achieved by the DT model in this study is considered highly acceptable.

Compared with the study by Mustafa et al. [30], the Decision Tree classification results in the current research (95% accuracy) fall within a consistent performance range. In their study, the Decision Tree model similarly demonstrated high efficacy, achieving up to 100% accuracy on a heart disease dataset and 96% in diabetes prediction. This corroborates the finding that Decision Trees tend to excel when applied to rule-based datasets or structured features, particularly when integrated with knowledge representations such as an RBES. Consequently, the results of this study confirm that the Decision Tree algorithm remains a highly robust and stable model within the context of rule-based medical decision-making systems.

Furthermore, the 94% accuracy achieved by the Gaussian Naïve Bayes model aligns with the findings of Maduri et al. [31], who demonstrated that GNB could attain up to 99% accuracy in agricultural yield prediction. Although the accuracy is marginally lower in the present study, the GNB algorithm’s performance remains consistent, highlighting its effectiveness as a probabilistic model for handling datasets with specific underlying distributions. This indicates that GNB remains highly relevant for implementation in medication monitoring systems, particularly when the dataset exhibits relatively stable statistical patterns.

Conversely, the K-Nearest Neighbor (KNN) algorithm yielded the lowest performance (81%) among the evaluated methods. This finding is consistent with research by Aldi et al. [32], which demonstrated that KNN's performance is highly dependent on the quality of the data distribution and is highly sensitive to noise. In the present study, KNN proved suboptimal because the characteristics of the ED and BUD data are inherently rule-based and temporal; consequently, they are not entirely compatible with a distance-based learning paradigm.

Parallel to these findings, the Gaussian Naive Bayes algorithm demonstrated highly commendable efficacy. While its statistical outputs fell marginally short of the Decision Tree benchmarks, the model maintained closely competitive results across all evaluated criteria, encompassing the F-measure, recall capabilities, precision rates, and overall correctness.. This model is computationally simpler and is frequently used in scenarios where the independence assumption among features is acceptable. Although GNB is not traditionally used in expert systems, its resulting accuracy is highly commendable. This efficacy stems from GNB's ability to support probabilistic inference by analyzing attributes like a decision tree. The attributes are derived from historical data, and in instances where information is missing for analysis, GNB can leverage expert knowledge to fill those gaps. GNB offers advantages regarding training and prediction speed, making it well-suited for large-scale datasets.

Conversely, KNN exhibited suboptimal performance compared to both DT and GNB. While KNN is recognized for being a simple and intuitive algorithm, its lower accuracy, precision, recall, and F1-Score suggest it is less compatible with the dataset used in this study, particularly in the presence of noise or high data complexity.

This lower accuracy is also attributed to the fact that the processed data does not significantly influence spatial proximity. Consequently, this research data is inherently less suitable for KNN, which relies on inter-data distance analysis rather than explicit rules. Furthermore, KNN suffers from high computational complexity as the volume of data and the number of features increase.

Although the accuracy of GNB and KNN is lower than that of the Decision Tree (DT), their performance remains commendable. These results suggest a significant opportunity for further experimentation within more complex frameworks, specifically in the form of hybrid systems beyond traditional rule-based models. By integrating rule-based expert systems with machine learning algorithms, the accuracy and reliability of the resulting decision-making process are significantly enhanced.

5. CONCLUSION

The initial phase of establishing the Rule-Based Expert System (RBES) was successfully implemented to monitor the usage periods of non-compounded medications. The results indicate that the Decision Tree (DT) machine learning algorithm achieved the highest accuracy of 95%. These findings demonstrate that the integration of RBES and machine learning is highly reliable for providing recommendations on drug-use limits. This study acknowledges certain limitations, primarily the restricted dataset size of 300 instances and its exclusive focus on non-compounded medications across a limited variety of drug categories. Future work will aim to expand upon this research by incorporating compounded medications and utilizing a substantially larger, more complex dataset.

REFERENCES

[1] A. Jafarzadeh, A. Mahboub-Ahari, M. Najafi, M. Yousefi, and K. Dalal, "Medicine storage, wastage, and associated determinants among urban households: a systematic review and meta-analysis of household surveys," *BMC Public Health*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12889-021-11100-4.

[2] A. E. Ikwara and H. Atwijukiire, "Self-medication and medication storage practices among Lira University students in Lira city, Northern Uganda," *Front. Public Health*, vol. 11, pp. 1–10, Nov. 2023, doi: 10.3389/fpubh.2023.1259279.

[3] S. M. Mohamed, "Frequency of use of non-prescribed medication among population sample from Al Madina City and its impact on quality of care in Saudi Arabia," *International Journal of Health Sciences (IJHS)*, vol. 12, no. 5, pp. 3–9, 2018.

[4] H. Chongloi, K. Divya, S. Deepthi, and V. B A, "Non-Prescription Medications Safety And Awareness," *Int. J. of Pharm. Sci*, vol. 2, no. 9, pp. 1230–1238, 2024, doi: 10.5281/zenodo.13834255.

[5] C. T. Chacko, D. Prakash, L. Joseph, A. Shabaraya, and C. Author, “A Review on the Attitude and Practice on Self Medication, Storage and Disposal of Drugs in a Community,” *International Journal of Research and Review (ijrrjournal.com)*, vol. 7, no. 8, pp. 122–129, 2020.

[6] B. KM, M. S. Kumar, R. Reji, A. Kumar, H. Doddayya, and M. S. M, “A Household Survey on Extend of Home Medication Utilization and Storage Habits –A Cross Sectional Survey,” *Saudi Journal of Medical and Pharmaceutical Sciences*, vol. 9, no. 07, pp. 470–476, Jul. 2023, doi: 10.36348/sjmps.2023.v09i07.016.

[7] United States Pharmacopeia, “Pharmaceutical Compounding - Nonsterile Preparations,” 2020, ch. General Chapters 795. Accessed: Aug. 19, 2025. [Online]. Available: https://www.uspnf.com/sites/default/files/usp_pdf/EN/USPNF/revisions/gc-795-rb-notice-20200424.pdf

[8] R. A. Despott, P. Vella Bonanno, and C. Gauci, “Risk Management of Medication Errors: Improving the Quality of Pharmacotherapeutic Practice,” *Pharmacol. Res. Perspect.*, vol. 13, no. 3, pp. 1–12, Jun. 2025, doi: 10.1002/prp2.70093.

[9] D. Ristian Octavia, S. Niken Purwandini, M. Hilaria, and D. P. Pangestu, “Medication Error at the Prescribing Phase,” *Health Information Journal*, vol. 19, no. 1, pp. 12–19, 2021, doi: 10.31965/infokes.Vol.

[10] Ministry of Health Republic of Indonesia, “Regulation of Indonesian Ministry of Health No. 34 on Pharmaceutical Service Standards in Primary Clinics.,” 2021.

[11] F. Cokro, M. A. Chiara, S. T. Arrang, and O. S. Hendra, “Indonesian Pharmacists’ Application of Beyond-Use Date Interventions to Patients and Its Contributing Factors: A Cross-Sectional Study,” *Hosp. Pharm.*, vol. 58, no. 6, pp. 628–633, Dec. 2023, doi: 10.1177/00185787231179833.

[12] F. Cokro, S. T. Arrang, J. A. N. Solang, and P. Sekarsari, “The Beyond-Use Date Perception of Drugs in North Jakarta, Indonesia,” *Indonesian Journal of Clinical Pharmacy*, vol. 10, no. 3, pp. 172–179, Sep. 2021, doi: 10.15416/ijcp.2021.10.3.172.

[13] M. S. Wolf *et al.*, “A Patient-Centered Prescription Drug Label to Promote Appropriate Medication Use and Adherence,” *J. Gen. Intern. Med.*, vol. 31, no. 12, pp. 1482–1489, Dec. 2016, doi: 10.1007/s11606-016-3816-x.

[14] D. Gikonyo, A. Gikonyo, D. Luvayo, and P. Ponoth, “Drug expiry debate: The myth and the reality,” Sep. 01, 2019, *Makerere University, Medical School*. doi: 10.4314/ahs.v19i3.49.

[15] V. Kushwaha, P. Agrawal, B. Pathak, H. Vekaria, and S. Siddiqui, “Effectiveness, Safety and Disposal of Medications Beyond Expiry Dates: Brief Review,” *Int. J. Pharm. Sci. Rev. Res.*, vol. ISSN, no. 4, pp. 105–108, 2025, doi: 10.47583/ijpsrr.2025.v85i04.015.

[16] L. Eltaib, S. A. Alanazi, and S. E. Ali, “Practices and attitudes concerning expiration date, unused, and expired medication disposal,” *Int. J. Med. Sci. Public Health*, no. 0, pp. 431–438, 2020, doi: 10.5455/ijmsph.2020.06099202010082020.

[17] Isnenia and S. Julaiha, “Mother’s Knowledge of Expiration Dates, Beyond-Use Date (BUD), and Storage Conditions for Compounding and Non-Compounding Drugs,” vol. 14, no. 1, pp. 74–83, 2024, doi: 10.22435/jki.v14i1.6634.

[18] W. Alfianna *et al.*, “The Effect of Poster Education on Public Knowledge of Beyond-Use Dates in Bekasi Regency,” *International Journal of Health and Science*, vol. 4, no. 3, pp. 2808–7380, 2024, doi: 10.54543/kesans.v4i3.407.

[19] P. Agrawal, V. Kushwaha, G. Singh Rana, B. K. Shoraisham, and S. Chandra, “The Truth About Expired Drugs: An In-Depth Review,” *Certified Journal │ Agrawal et al. World Journal of Pharmaceutical Research*, vol. 12, pp. 200–214, 2023, doi: 10.20959/wjpr202321-30355.

[20] Z. H. Nodeh, M. Hosseini, M. F. Khoshknab, S. Shirozhan, and H. R. Khankeh, “A scoping review of individual health responsibility: A context-base concept,” *J. Educ. Health Promot.*, vol. 13, no. 1, pp. 1–8, May 2024, doi: 10.4103/jehp.jehp_565_23.

[21] T. Schramme, "Health as Complete Well-Being: The WHO Definition and Beyond," *Public Health Ethics*, vol. 16, no. 3, pp. 210–218, Nov. 2023, doi: 10.1093/phe/phad017.

[22] M. Ramalingam, R. Puviarasi, N. Diyana Afiqah Binti Zakaria, and E. Chinnavan, "Developing Mobile Application For Medicine Expiry Date detection," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 16, pp. 3895–3901, 2018, [Online]. Available: http://www.acadpubl.eu/hub/

[23] E. Deepika, B. R. Devi, Sartitha. B, Begum. B, and K. V. Kumar, "Health Monitoring System Using Rule Based Expert System," *International Journal of Creative Research Thoughts*, vol. 8, pp. 2320–28820, 2020, [Online]. Available: www.ijcrt.org

[24] M. N. O. Sadiku, K. B. Olanrewaju, E. A. Dada, and S. M. Musa, "Expert Systems in Healthcare," *Journal of Scientific and Engineering Research*, vol. 2021, no. 2, pp. 218–222, 2021, [Online]. Available: www.jsaer.com

[25] B. Silva, F. Hak, T. Guimaraes, M. Manuel, and M. F. Santos, "Rule-based System for Effective Clinical Decision Support," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 880–885. doi: 10.1016/j.procs.2023.03.119.

[26] M. Marfuah, "Implementation of Rule-Based Expert Systems for Real-Time Monitoring of Criminal Case Resolution Using Radio Frequency Identification Technology," *Journal Of Business Information Systems*, vol. 7, no. 1, pp. 73–79, May 2017, doi: 10.21456/vol7iss1pp73-79.

[27] B. Reddy and R. Fields, "From Past to Present A Compre hensive Technical Review of Rule-based Expert Systems from 1980-2021," in *Proceedings of the 2022 ACMSE Conference - ACMSE 2022: The Annual ACM Southeast Conference*, Association for Computing Machinery, Inc, Apr. 2022, pp. 167–172. doi: 10.1145/3476883.3520211.

[28] G. Aguilera-Venegas, E. Roanes-Lozano, G. Rojo-Martínez, and J. L. Galán-García, "A proposal of a mixed diagnostic system based on decision trees and probabilistic experts rules," *J. Comput. Appl. Math.*, vol. 427, Aug. 2023, doi: 10.1016/j.cam.2023.115130.

[29] F. Cokro, M. A. Chiara, S. T. Arrang, and O. S. Hendra, "Indonesian Pharmacists' Application of Beyond-Use Date Interventions to Patients and Its Contributing Factors: A Cross-Sectional Study," *Hosp. Pharm.*, vol. 58, no. 6, pp. 628–633, Dec. 2023, doi: 10.1177/00185787231179833.

[30] E. M. Mustafa, M. M. Saad, and L. W. Rizkallah, "Building an enhanced case-based reasoning and rule-based systems for medical diagnosis," *Journal of Engineering and Applied Science*, vol. 70, no. 1, Dec. 2023, doi: 10.1186/s44147-023-00315-4.

[31] P. K. Maduri, P. Dhiman, R. Srivastava, and R. Singh, "Autonomous Crop Prediction System using Machine Learning," 2024.

[32] F. Aldi, I. Nozomi, and S. Soeheri, "Comparison of Drug Type Classification Performance Using KNN Algorithm," *SinkrOn*, vol. 7, no. 3, pp. 1028–1034, Jul. 2022, doi: 10.33395/sinkron.v7i3.11487.

[33] I. Ihsan, Y. Wirasasmita, and Sarni, "Profile Of Waiting Time For Compounded And Non-Combounded Prescription Services At Siko Community Health Center, North Ternate District," *International Jornal Of Multidiciplynary Leaners*, vol. 1, no. 2, pp. 111–15, 2024.

[34] J. Fournier, L. Barret, C. Khouri, F. Naudet, R. Boussageon, and M. Roustit, "The evidence base of the 10 most prescribed drugs in England, France, and the United States: a scoping review," *J. Clin. Epidemiol.*, vol. 174, pp. 1–11, Oct. 2024, doi: 10.1016/j.jclinepi.2024.111478.

[35] Unites States Pharmacopeia, "USP Compounding Standards and Beyond-Use Dates (BUDs)," Jun. 2019. Accessed: Aug. 20, 2025. [Online]. Available: https://www.usp.org/sites/default/files/usp/document/our-work/compounding/usp-bud-factsheet.pdf

[36] D. Janjanam, B. Ganesh, and L. Manjunatha, "Design of an expert system architecture: An overview," in *Journal of Physics: Conference Series*, Purpose-Led Publishing, Feb. 2021, pp. 1–7. doi: 10.1088/1742-6596/1767/1/012036.

[37] L. Sumaryanti, T. Istanto, and S. Pare, "Rule Based Method in Expert System for Detection Pests and Diseases of Corn," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Jul. 2020, pp. 1–9. doi: 10.1088/1742-6596/1569/2/022023.

[38] Z. Zeng, Q. Cheng, and Y. Si, "Logical Rule-Based Knowledge Graph Reasoning: A Comprehensive Survey," *Mathematics*, vol. 11, no. 21, pp. 1–37, Nov. 2023, doi: 10.3390/math11214486.

[39] B. T. Jijo and A. M. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 1, pp. 20–28, Jan. 2021, doi: 10.38094/jastt20165.

[40] A. K. Srivastava, B. K. Singh, and S. Gupta, "Prediction of Tool Wear Using Machine Learning Approaches for Machining on Lathe Machine," 2023.

[41] Z. M. Radeef, S. H. Hashem, and E. K. Gbashi, "New Feature Selection Using Principal Component Analysis," *Journal of Soft Computing and Computer Applications*, vol. 1, no. 2, Dec. 2024, doi: 10.70403/3008-1084.1012.

[42] O. Peretz, M. Koren, and O. Koren, "Naive Bayes classifier – An ensemble procedure for recall and precision enrichment," *Eng. Appl. Artif. Intell.*, vol. 136, Oct. 2024, doi: 10.1016/j.engappai.2024.108972.

[43] L. Rusdiana and M. Marfuah, "The Application of Determining Students' Graduation Status of STMIK Palangkaraya Using K-Nearest Neighbors Method," in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics Publishing, Dec. 2017. doi: 10.1088/1755-1315/97/1/012040.

[44] J. Joosten, A. E. Permanasari, and T. B. Adji, "The use of decision table for reducing complex rules in software testing," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jan. 2020, pp. 1–8. doi: 10.1088/1757-899X/732/1/012086.

[45] J. A. Whittaker, *How to Break Software: A Practical Guide to Testing*, 1st ed. P&C ECS, 2002. [Online]. Available: https://www.researchgate.net/publication/268043176

[46] M. Sholeh, I. Gisfas, Cahiman, and M. A. Fauzi, "Black Box Testing on ukmbantul.com Page with Boundary Value Analysis and Equivalence Partitioning Methods," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1742-6596/1823/1/012029.

[47] X. Guo, H. Okamura, and T. Dohi, "Optimal test case generation for boundary value analysis," *Software Quality Journal*, vol. 32, no. 2, pp. 543–566, Jun. 2024, doi: 10.1007/s11219-023-09659-9.

[48] F. Dobslaw, F. G. D. O. Neto, and R. Feldt, "Boundary Value Exploration for Software Analysis," in *IEEE International Conference on Sofware Testing*, Ferification and Validation Worshops (ICSTW), 2020, pp. 346–353.

[49] E. M. Mustafa, M. M. Saad, and L. W. Rizkallah, "Building an enhanced case-based reasoning and rule-based systems for medical diagnosis," *Journal of Engineering and Applied Science*, vol. 70, no. 1, Dec. 2023, doi: 10.1186/s44147-023-00315-4.