engineering proceedings

MDPI

Proceeding Paper

Ant Colony Optimization Algorithm for Feature Selection in Suspicious Transaction Detection System †

Karina Niyazova [1], Assel Mukasheva [2,*], Gani Balbayev [3], Teodor Iliev [4], Nazym Mirambayeva [1] and Mukhamedali Uzakbayev [1]

[1] Department of Information Technology, Non-Profit JSC "Almaty University of Power Engineering and Telecommunications Named after Gumarbek Daukeyev", Almaty 050013, Kazakhstan; niyazovakarina2000@gmail.com (K.N.); n.mirambayeva@aues.kz (N.M.); m.uzakbayev04@gmail.com (M.U.)
[2] School of Information Technology and Engineering, Kazakh-British Technical University, Almaty 050000, Kazakhstan
[3] Academy of Logistics and Transport Almaty, Almaty 050012, Kazakhstan; g.balbayev@gmail.com
[4] Department of Telecommunication, University of Ruse, 7004 Ruse, Bulgaria; tiliev@uni-ruse.bg
* Correspondence: mukashevascience@gmail.com
† Presented at the 4th International Conference on Communications, Information, Electronic and Energy Systems (CIEES 2023), Plovdiv, Bulgaria, 23–25 November 2023.

Abstract: The fight against financial crimes has become increasingly challenging, and the need for sophisticated systems that can accurately identify suspicious transactions has become more pressing. The goal of the study is to develop a new feature selection method based on swarm intelligence algorithms to improve the quality of data classification. This article is about the development of an information system for the classification of transactions into legal and suspicious in an anti-money laundering sphere. The system utilizes a swarm-algorithm-based feature selection approach, specifically the ant colony optimization algorithm, which was both used and adapted for this purpose The article also presents the system's functional–structural diagram and feature selection algorithm flowchart. The proposed feature selection method can be used to classify data from various subject areas.

Keywords: AML; swarm intelligence; ant colony optimization

Citation: Niyazova, K.; Mukasheva, A.; Balbayev, G.; Iliev, T.; Mirambayeva, N.; Uzakbayev, M. Ant Colony Optimization Algorithm for Feature Selection in Suspicious Transaction Detection System. *Eng. Proc.* **2024**, *60*, 18. https://doi.org/10.3390/engproc2024060018

Academic Editor: Renato Filjar

Published: 11 January 2024

1. Introduction

In 2009, the Republic of Kazakhstan introduced the law "On Counteracting the Legalization (Laundering) of Illegally Obtained Income and the Financing of Terrorism." To comply with financial regulators' legal requirements, financial institutions utilize Anti-Money Laundering (AML) tools. AML is a set of techniques for analyzing money laundering risks, including customer identity verification, transaction monitoring, and regulatory compliance.

The term "money laundering" was first used in the early 20th century to describe the process by which criminals convert illegal funds from illicit activities such as fraud, embezzlement, insider trading, bribery, theft, or tax evasion into legal funds or assets through a series of transactions, ultimately concealing the source of the illegally obtained funds.

In 1989, the Financial Action Task Force on Money Laundering (FATF) was established to combat money laundering and the legalization of proceeds from crime. This organization publishes the recommendations that have been adopted in over 130 countries worldwide [1].

Money laundering is a covert and imperceptible activity, in contrast to fraud. As a result, many individuals may become unwitting victims of the money laundering process without even realizing it. Different assessments suggest that about 2 trillion USD are laundered on an annual basis [2]. The global efforts to combat money laundering do not seem to be making substantial strides, as evidenced by the fact that, this year, the average

risk level has only decreased by a marginal 0.05% to reach 5.25 out of 10 (where 10 is the maximum risk level) [3].

Money laundering usually involves a three-step process: placement, layering, and integration. The first step is placement, which involves withdrawing a large sum of money and depositing it into a financial institution. To avoid suspicion, criminals often use businesses that deal with large amounts of cash, such as restaurants, bars, and car washes [4]. The funds can be broken down into smaller amounts and deposited into bank accounts. Law enforcement agencies have developed ways to make it more difficult to place funds inconspicuously, such as suspicious activity reports and currency transaction reports [4].

The second step, layering, involves numerous transactions, often using front companies and organizations. Criminals may also buy high-value goods and register them under a nominee's name to obscure the money's origin. This makes it harder for law enforcement to track the movement of money, as information on financial transactions may not be available until later [3,4]. Finally, integration is the stage when laundered funds can be used by criminals [4,5].

From the perspective of financial markets, the act of laundering money can take various forms such as smurfing, trading, interbank wire transfers, creating artificial invoices, imitating transactions, and other related activities [5].

Money laundering poses a variety of risks, including operational, reputational, and legal risks, and can also undermine the stability of financial institutions. Developing countries are particularly vulnerable to the impact of money laundering, as they are often unable to be selective about the sources of capital they accept. Criminal tax evasion through money laundering can increase taxes and prices for the general population. By combating money laundering, not only can the number of financial crimes be reduced, but resources can also be freed up to combat more serious crimes [4,5].

The traditional approach to AML remains a manual process. However, since the volume of banking transactions and data is increasing every year, work within the framework of AML should be automated using special systems and tools. Traditional AML software usually relies on pre-defined rules and data, and suffers from various drawbacks, such as inefficient thresholds, high false-positive rates, limited pattern recognition, and inadequate data processing capabilities.

One of the crucial components of AML is transaction monitoring. Due to the high volume and speed of transactions processed by financial institutions daily, it is essential that the transaction monitoring tools used are appropriate for the current conditions and provide reliable control. To ensure the effective detection of suspicious transactions, a robust transaction monitoring process is essential. This process should involve comparing transaction details with identified risks, such as the geographic location of the transaction, the type of products and services being offered, and the type of client involved in the transaction [5]. The process should also consider different typologies for money laundering and other illicit activities to determine if the transaction is unusual or suspicious. By doing so, financial institutions can better identify potentially illicit activities and take appropriate measures to prevent money laundering and other financial crimes.

There are two methods of transaction monitoring: real-time (pre-transaction) monitoring during the transaction, and post-transaction monitoring for identifying patterns and trends [6,7]. The main objective of both approaches is to identify suspicious activities that require further investigation. Real-time monitoring is preferred to reduce the risk of breaching sanctions, while post-transaction monitoring can be useful for identifying recurring criminal activities.

Transaction monitoring is a useful tool that can be utilized not just to detect money laundering, but also to flag indicators of illicit activities like terrorist financing, various types of fraud (such as false insurance claims and chargebacks), identity theft, drug trafficking, corruption, and bribery.

As the volume of digital transactions increases daily, the need for transaction monitoring also grows. The manual monitoring of transactions and identification of suspicious

patterns without an automated system would be nearly impossible. Using an automated transaction monitoring solution can enable companies to effectively monitor and identify unusual financial activities.

There is a main category in transaction monitoring, which is suspicious transaction. If there is reason to believe that the transaction involves the proceeds of crime or is intended for illicit purposes, it may be classified as a suspicious transaction.

To execute appropriate transaction monitoring, financial institutions should keep the following information about the transactions they obtain:

- The name of the customer and/or beneficiary;
- Address;
- Date and nature of the transaction;
- Type and amount of currency involved in the transaction;
- Other relevant information typically recorded by the financial institution [8].

Financial institutions may also keep track of other data points, such as the origin of the funds, the purpose of the transaction, the geographic location of the transaction, and the type of products or services being offered. This information helps to identify and monitor transactions that may be suspicious or unusual.

The concept of incorporating artificial intelligence and machine learning in AML transaction monitoring has been around for some time. Nevertheless, it remains crucial for the monitoring outcomes and even the transactions themselves to undergo manual verification by experts [9]. However, using machine learning has a lot of advantages [10]:

- Scale (machines have no difficulties with dealing with greater amounts of data);
- Diversity (machines can easily adapt to different sets of rules);
- Adaptability (if the models are regularly retrained, the systems keeps up with the latest trends);
- Pattern recognition (machines can see all the data).

Machine learning is a technique used to make predictions based on past data. When the machine learning algorithm is used to classify data into predefined categories based on labeled training data, this is known as classification. Classification is considered a type of supervised learning because it relies on labeled data to make predictions.

Feature selection is a critical aspect of machine learning. This process involves removing redundant or unnecessary data from the input dataset to facilitate their analysis using data mining techniques. There are several reasons why it is essential to reduce the number of features, including [10]:

1. Computational efficiency: With a large number of features, the calculations can become computationally intensive, leading to slower processing times.
2. Overfitting: A large number of input features can increase the risk of overfitting, where the algorithm becomes too specialized to the training data and performs poorly on new data.
3. Memory limitations: In some cases, the input dataset may be too large to fit into the memory, requiring modifications to the classification algorithm to handle out-of-core processing.

Traditional AML software usually relies on pre-defined rules and data, and suffers from various drawbacks, such as inefficient thresholds, high false-positive rates, limited pattern recognition, and inadequate data processing capabilities. The article proposes a novel method for transaction classification using swarm-intelligence-based feature selection. The goal of the study is to develop a new feature selection method, leveraging swarm intelligence algorithms with the aim of enhancing the accuracy and efficiency of classification techniques. The new method is designed to be applied in the field of AML to bolster the detection and prevention of financial crimes.

2. Methods

For the purpose of transaction monitoring, a separate server has to be installed. This server has to be linked to databases such as OFAC and Dow Jones in order to verify whether individuals or entities are on sanction or stop lists. Furthermore, the server should be connected to the core banking system, card processing system and local data storage servers, allowing for it to retrieve transaction data. The infrastructure scheme of the developing system is shown in Figure 1.

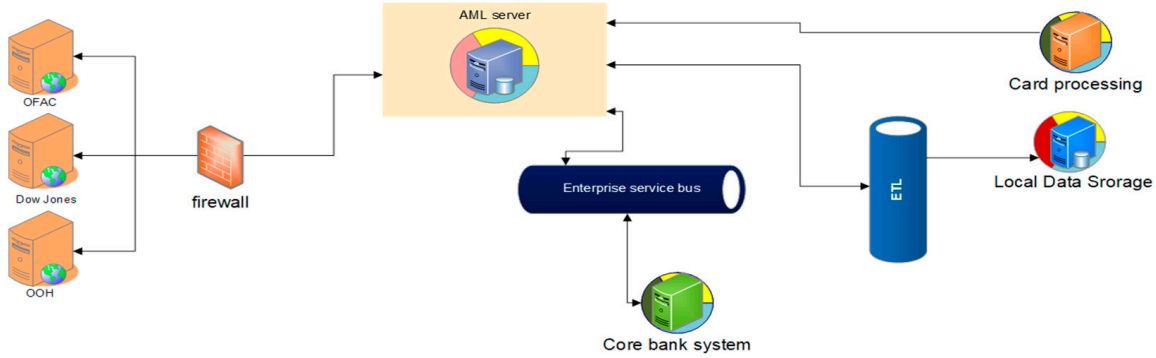

Figure 1. The functional–structural diagram of developed AML system.

In many cases, the prediction outcome is influenced by multiple features or predictors. Therefore, it is crucial to eliminate any excess noise from the data and create a new, reduced dataset. Feature selection involves the projection of the data onto a smaller set of features that preserve as much of the relevant information as possible, improving the efficiency of machine learning algorithms.

Feature selection is essentially an optimization problem, and metaheuristic algorithms are an effective approach to solving it efficiently. These algorithms include evolutionary and swarm-based methods. Swarm intelligence algorithms (or swarm algorithms) are different from evolutionary algorithms as they do not require the generation of new populations, as the number of agents is predetermined. These agents solve simple problems individually, and their group interaction is considered "intelligent".

Swarm intelligence is a population-based, nature-inspired branch of artificial intelligence [11]. It is based on the idea that decentralized and self-organized groups can solve complex problems more efficiently than individuals or centralized systems. Swarm intelligence algorithms often involve the interaction of multiple agents or individuals, who communicate and cooperate with each other to achieve a common goal [12]. Examples of swarm intelligence algorithms include Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Artificial Bee Colony (ABC) Optimization. Swarm intelligence has found applications in various fields, such as optimization, data clustering, and robotics.

Ant Colony Optimization algorithm was introduced by M. Dorigo in 1992. This algorithm mimics the foraging behavior of social ants [13]. The ants have a unique ability to find the shortest path between their nest and food source by utilizing organic substances called 'pheromone trails'. While moving on the ground, each ant releases pheromones that act as clues for other ants, especially when carrying food [14,15]. The nearby ants detect this chemical and follow the path with a higher concentration of pheromones. This enables the ant colony to communicate and cooperate efficiently in finding the optimal route to the food source. The selection of a route is dependent on the number of insects that have previously traveled on that route at a given point in time [16,17]. This probability can be calculated using Equation (1):

$$P = \frac{\tau_{ij}^{\alpha} \cdot \omega_{ij}^{\beta}}{\sum_k \tau_{ij}^{\alpha} \cdot \omega_{ij}^{\beta}}, \quad (1)$$

where $\omega_i \in [0,1]$ is a weight of the *i*-th feature.

Since the basic amount of deposited pheromone is constant value, but every feature has a weight, which influences the amount of pheromones the feature will recieve in the next steps, Equation (2) will be modified in the following way:

$$\tau_{ij} = (1 - \rho) \cdot \tau_{ij} + \sum \Delta \tau_{ij}^{k} (1 + \omega_i). \quad (2)$$

Equation (2) means that the pheromone amount depends on the weight of the feature. The value of the weight can be set manually, depending on the specific task. This paper considers the classification of transactions into legal and illegal. Table 1 lists the possible features of a transaction for which it is necessary to make a selection.

Table 1. The possible set of transaction features.

Operation Features	Sender Data and Beneficiary Data
Operation number	Client type
Branch	Last name, first name
Currency	Birth date
Datetime	Country
Base amount	Economic sector
Currency amount	Client number
EKNP code	
Document number	
Document date	
Document category	
Document type	
Operation status	

Figure 2 illustrates the general approach to feature selection. The primary procedures involve examining the input dataset columns, computing feature weights, performing data selection using the ant colony optimization algorithm, training the model, executing the classifier on a test dataset, and displaying the output results. Features that are not filled or do not significantly affect the classification results may be excluded from the selection process. These features will have a lower weight and may be less likely to be selected. The input data for feature selection are typically in the form of a data table. In order to assign weights to each feature, the data need to be analyzed. The first step is to check each column of the table for empty values in every field. Then, each column is checked to see if it contains empty values in some fields. Finally, columns are checked to determine if they have the same value in every field. These checks help to identify features that can be excluded and assign appropriate weights to the remaining features.

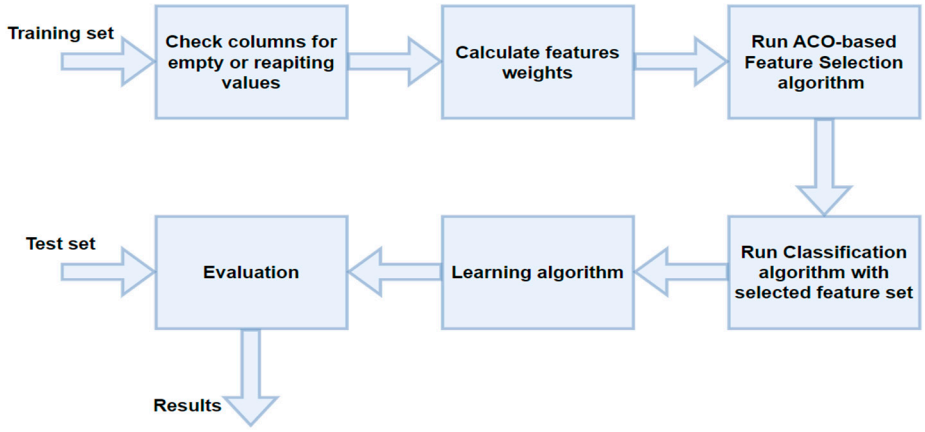

Figure 2. The flowchart of the feature selection approach.

Figure 3 shows the result of such checks in the output log.

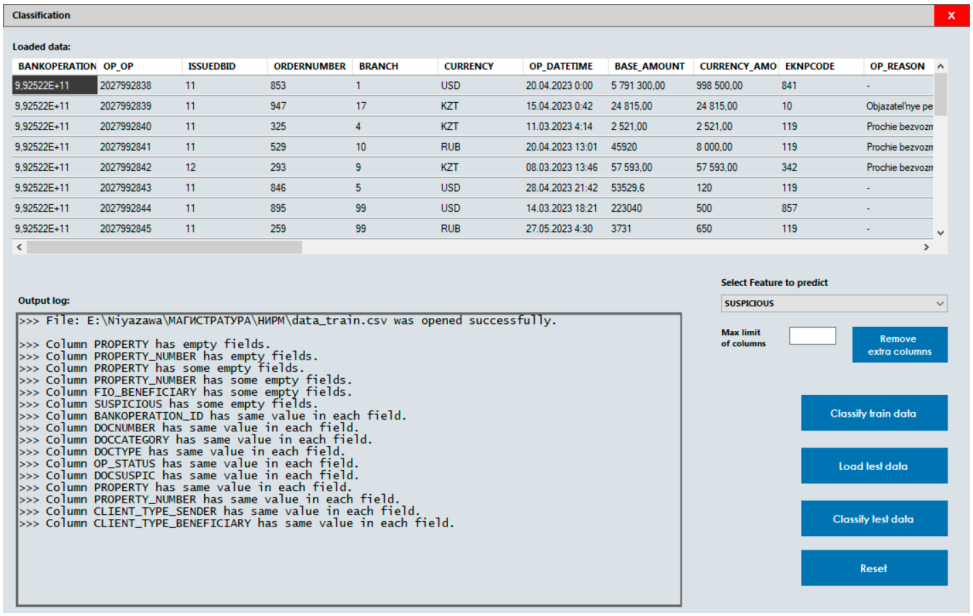

Figure 3. The results of the column checks.

Features that have identical or blank values in all fields are assigned a weight of zero as they have no impact on the final outcome. The weights of other features are determined based on the number of distinct values in each column. The results of these computations are depicted in Figure 4.

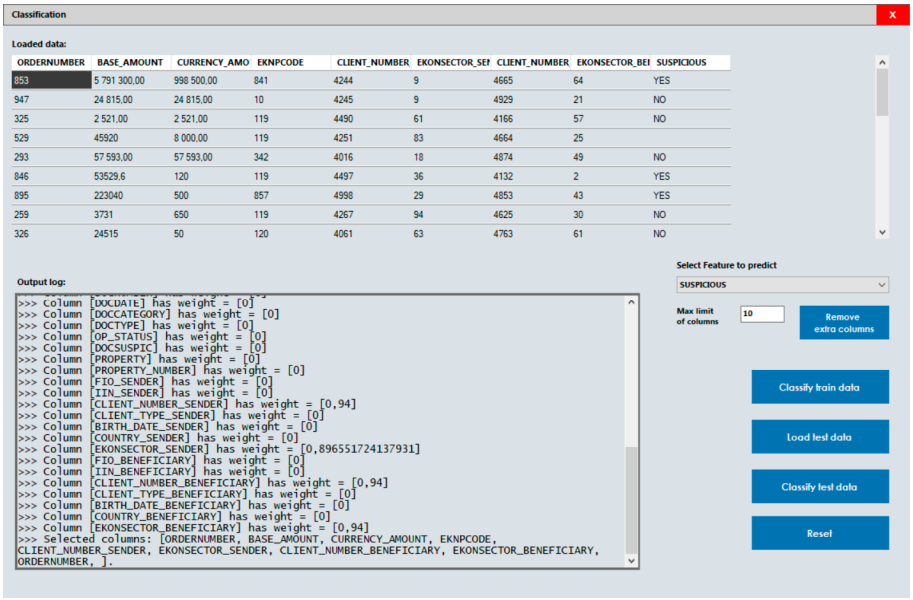

Figure 4. The result of removing extra columns from the input dataset.

The utilized classifier is the Naive Bayes classifier. The Naive Bayes classifier is a type of probabilistic classifier that is commonly used in machine learning for classification tasks. It is based on Bayes' theorem and the assumption of independence between the features. The Naive Bayes classifier calculates the probability of each class given the input features and selects the class with the highest probability as the predicted class for the input. It is called "naive" because it assumes that all the features are independent of each other, which is not necessarily true in practice. Despite this limitation, Naive Bayes classifiers are known for their simplicity, efficiency, and effectiveness, especially in text classification tasks.

The results of the data classification without feature selection are shown in Figure 5. The number displayed in the textbox is significantly larger than the actual number of features, indicating that all features were selected. The prediction error is about 3.5%.

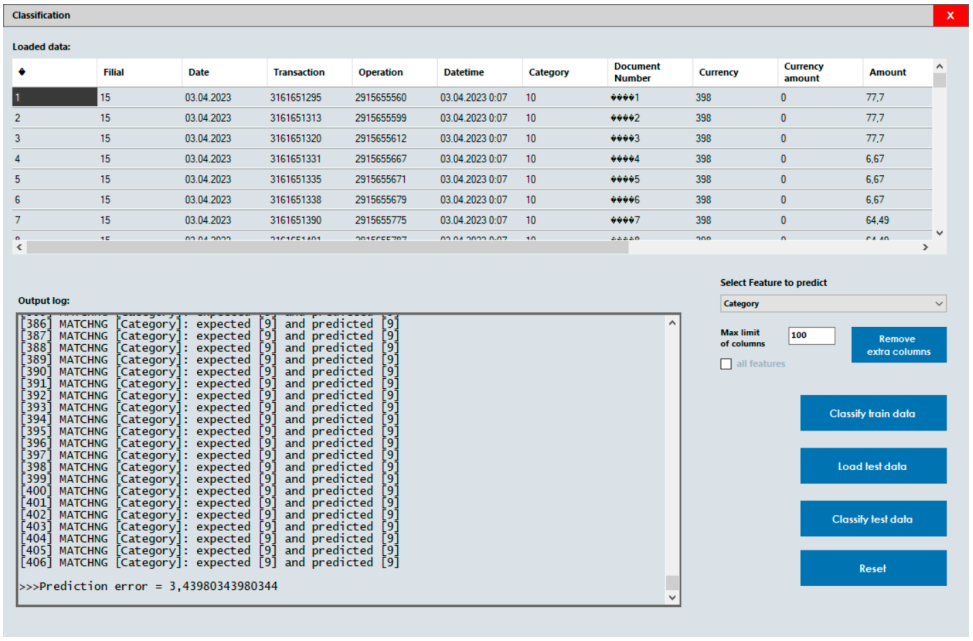

Figure 5. The results of classification without utilizing feature selection.

Figure 6 displays the results for classification with feature selection. The prediction error is about 2%.

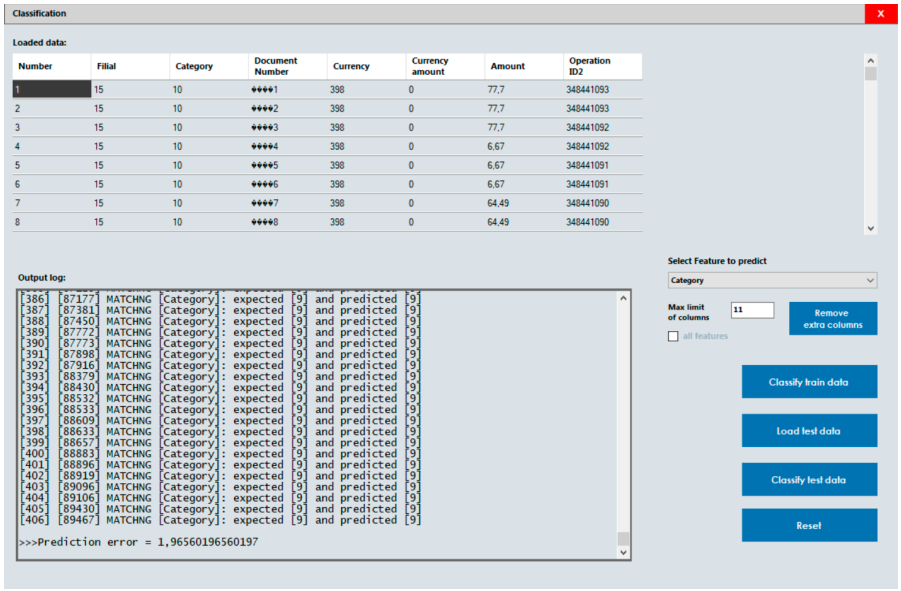

Figure 6. The results of classification after using feature selection.

3. Results

The findings of the research suggest that utilizing a swarm algorithm for feature selection (FS) significantly enhances the outcomes of sample classification. The research encompassed not only datasets relating to banking transactions but also those for iris, Mall Customers, and wine quality.

Classification results comparing the prediction error values of feature selection and those obtained when using the full feature set are shown in Table 2.

Table 2. Comparison of classification results.

Dataset	Selected Feature	Classification Error	Prediction Error without FS	Prediction Error with FS	Classification Error with FS
Transactions	Category	2.54	3.44	1.96	1.97
Iris	Variety	0	6	0	6
Wine	Alcohol	1.36	0	0	1.36
Mall Customers	Genre	36.5	36.5	35.4	35.4

Table 2 shows that, in certain instances, the implementation of the suggested feature selection method resulted in a noteworthy enhancement of the classification outcomes, or the outcome was unchanged and remained equivalent to utilizing all features. Therefore, it can be concluded that the feature selection method has potential benefits.

The results are presented in the form of a graph in Figure 7.

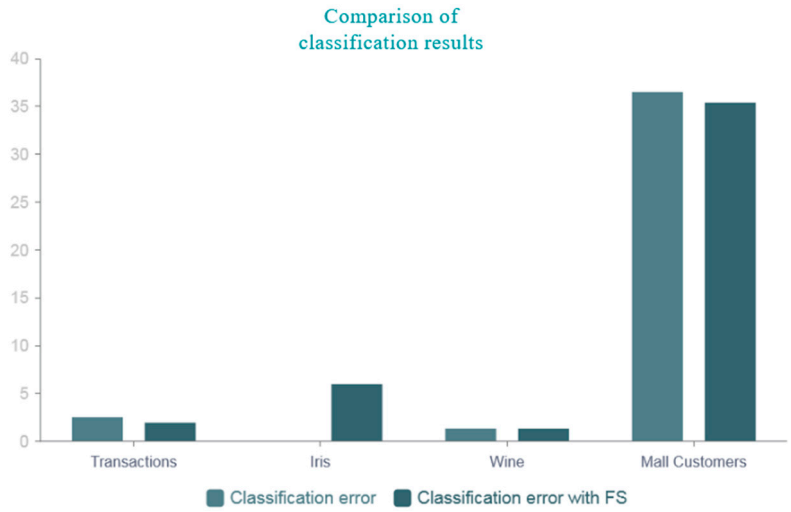

Figure 7. Comparison of classification results on the full dataset and with feature selection on the test set.

4. Discussion

The proposed method implements a modification where feature weight is employed to determine the likelihood of choosing a specific feature, in contrast to the use of a heuristic parameter η, referred to as path visibility. The weight of each feature is computed based on the number of unique values found in the respective column of the input dataset. While this adjustment represents a promising direction in feature selection, it is important to note that our method is not entirely automated, as manual settings are introduced to complement the feature weight calculation process.

In considering the real-world implications of our method, it is essential to highlight the potential benefits it could bring to the field of AML. By enhancing the feature selection process, our approach has the capacity to significantly improve the accuracy and efficiency of classification within AML systems. However, we must also acknowledge the limitations of our method. While it shows promise in various scenarios, there may be specific cases or datasets where its performance could be less optimal. It is crucial to recognize these boundaries to inform practical applications.

To improve the results, other, more accurate methods can be used to calculate feature weights, for example, calculating the level of feature correlation with the target. Moreover, it is important to consider that parameters α and β may also impact the classification results, and further research could also aim to optimize their choice. Future studies should compare the proposed method with existing feature selection methods to evaluate its effectiveness and competitiveness. In future works, it is planned to enhance the algorithm by optimizing

the calculation of feature weights in the input data set and expanding its applicability beyond the current classification task to cover other types of classification.

5. Conclusions

The article describes a method for choosing relevant features, which involves the use of the ant colony optimization algorithm to solve the problem of categorizing transactions into either legal or suspicious, which is necessary for post-transaction monitoring. The modified feature selection method can be applied to a variety of datasets containing many features. The selection of features that have the greatest impact on the result will significantly speed up the work of classification algorithms and improve the quality of the results. In the course of the study, it was found that the selection of features performed using the proposed method, in some cases, leads to better results than using the full set of data. Experiments were carried out on four different data sets.

This study may be useful for specialists involved in swarm intelligence and machine learning, since the feature selection method can be adapted to other swarm algorithms.

Author Contributions: Conceptualization, K.N., A.M., G.B., T.I. and M.U.; methodology, K.N., A.M. and G.B.; software, K.N. and A.M.; validation, A.M., G.B., T.I. and M.U.; formal analysis, A.M. and T.I.; investigation, K.N., A.M., G.B. and M.U.; resources, K.N.; data curation, K.N. and A.M.; writing—original draft preparation, K.N. and A.M.; writing—review and editing, K.N., A.M., G.B. and M.U.; visualization, K.N. and A.M., N.M.; supervision, T.I.; project administration, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Chapman, R. Anti-Money Laundering: A Practical Guide to Reducing Organizational Risk, 2018, Kogan Page. Available online: https://www.perlego.com/book/1015261/antimoney-laundering-a-practical-guide-to-reducing-organizational-risk-pdf (accessed on 1 July 2023).
2. Deloitte Professional Services (DIFC) Limited, Transaction Monitoring Optimisation, Dubai, United Arab Emirates. 2022. Available online: https://www2.deloitte.com/content/dam/Deloitte/xe/Documents/finance/me_transaction-monitor-optimisation.pdf (accessed on 30 April 2023).
3. Basel Institute on Governance, 2022. Basel AML Index 2022: 11th Public Edition–Ranking Money Laundering and Terrorist Financing Risks around the World. Available online: Index.baselgovernance.org (accessed on 1 July 2023).
4. Kebbell, S. *Anti-Money Laundering Compliance and the Legal Profession*; Routledge: Abingdon, UK, 2021; pp. 39–47.
5. Dill, A. *Anti-Money Laundering Regulation and Compliance Key Problems and Practice Areas*; Edward Elgar Publishing: Cheltenham, UK, 2021; pp. 37–49.
6. UK Financial Sanctions: General Guidance for Financial Sanctions under the Sanctions and Anti-Money Laundering Act 2018. Available online: https://www.gov.uk/government/publications/ofsi-guidance-html-documents/uk-financial-sanctions-general-guidance-for-financial-sanctions-under-the-sanctions-and-anti-money-laundering-act-2018 (accessed on 30 April 2023).
7. Financial Intelligence Analysis Unit. Guidance Note: A Look Through the Obligation of Transaction Monitoring. Available online: https://fiaumalta.org/wp-content/uploads/2023/05/Guidance-Note-A-Look-Through-the-Obligation-of-Transaction-Monitoring.pdf (accessed on 30 April 2023).
8. FATF (2012–2023), International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation: The FATF Recommendations. Paris: Financial Action Task Force. 2012. Available online: https://www.fatf-gafi.org/recommendations.html (accessed on 30 April 2023).
9. Chau, D.; Dijck, V. *Anti-Money Laundering Transaction Monitoring Systems Implementation: Finding Anomalies*; John Wiley & Sons: Hoboken, NJ, USA, 2021; pp. 26–29.
10. Saporta, G.; Maraney, S. *Practical Fraud Prevention: Fraud and AML Analytics for Fintech and eCommerce, Using SQL and Python*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022; pp. 79–87.
11. Sharma, A.; Sharma, A.; Pandey, J.K.; Ram, M. *Swarm Intelligence*; CRC Press: Boca Raton, FL, USA, 2022; pp. 109–115.

12. Blum, C.; Merkle, D. *Swarm Intelligence Introduction and Applications*; Natural Computing Series; Springer Science & Business Media: Berlin, Germany, 2008; p. 286.
13. Yang, X.-S. (Ed.) *Nature-Inspired Computation and Swarm Intelligence Algorithms, Theory and Applications*; Elsevier/Academic Press: London, UK, 2020; pp. 8–20.
14. Slowik, A. *Swarm Intelligence Algorithms: Modifications and Applications*; CRC Press: Boca Raton, FL, USA, 2020. [CrossRef]
15. Yedilkhan, D.; Mukasheva, A.; Bissengaliyeva, D.; Suynullayev, Y.D. Performance Analysis of Scaling NoSQL vs SQL: A Comparative Study of MongoDB, Cassandra, and PostgreSQL. In Proceeding of the 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST), Astana, Kazakhstan, 4–6 May 2023; pp. 479–483. [CrossRef]
16. Santos, J.C.G.; Augusto, V. *Swarm Intelligence*; CRC Press: Boca Raton, FL, USA, 2022; pp. 3–21.
17. Mukasheva, A.; Akanov, Z.; Yedilkhan, D. Research of the Regression Analysis Methods for Predicting the Growth of Patients with Diabetes Mellitus. In Proceedings of the 2021 IEEE International Conference on Smart Information Systems and Technologies (SIST), Nur-Sultan, Kazakhstan, 28–30 April 2021; pp. 1–7. [CrossRef]