

Comparison of SVM and Naïve Bayes Algorithms using Binary Grey Wolf Optimizer for Diabetes Mellitus Prediction

Berliana Fajrina, Safina Faradilla Hasibuan, Andi Nugroho

Abstract—Diabetes mellitus (DM) is a metabolic disorder characterized by chronically high blood sugar or glucose levels due to problems with insulin secretion, insulin response, or both. Therefore, an appropriate approach is needed in predicting diabetes to support early diagnosis and more effective prevention efforts. One of the approaches used is the machine learning method, which has proven to be capable of improving prediction accuracy. In this study, the SVM and Naïve Bayes algorithms are applied with feature selection techniques using the Binary Grey Wolf Optimizer (BGWO) to enhance classification performance. Based on the test results, SVM-BGWO showed an accuracy of 73.30% and a precision of 85.74%, while NV-BGWO achieved an accuracy of 72.60% and a precision of 84.88%. These results indicate that SVM-BGWO has superior performance compared to NV-BGWO in terms of accuracy and precision. This research aims to compare the two algorithms in order to find the most accurate model for predicting diabetes mellitus. In addition, it can help medical personnel in providing early diagnosis and enhancing diabetes prevention efforts.

Keywords— Diabetes Mellitus, Binary Grey Wolf Optimizer, SVM, Naïve Bayes.

I. INTRODUCTION

Diabetes mellitus (DM) comprises a collection of metabolic disorders marked by persistent hyperglycemia resulting from deficient insulin secretion, impaired insulin action, or a combination of both [1]. According to the International Diabetes Federation (IDF), 537 billion people worldwide are affected by diabetes, and it is estimated that by 2045 this number will increase to 783 billion. As time progresses, the spread of diseases is becoming more rampant, necessitating faster methods for disease prediction to enable earlier prevention and rescue actions.

The data mining and machine learning approach has made significant progress in translating accessible data into valuable information sequentially to improve the efficiency of the diagnostic process[2]. With the efficiency of the diagnostic process by utilizing machine learning, potential human errors can drastically decrease and optimize the prediction of diabetes mellitus.

The research that has been conducted is the study carried out by Akazue et al. Research has been able to establish that a machine learning approach for survival analysis and prediction can be implemented alongside parametric or non-

parametric tools to model time-to-event data [3]. In the results of the research, the number of diabetes deaths can also be determined. That proves that machine learning can be used in the field of health.

The study "Clustering Urban Roads Using Local Binary Patterns to Enhance the Accuracy of Traffic Flow Prediction" by Priambodo et al. explored a method of improving traffic state predictions by analyzing spatial relationships between road segments. The research applied Local Binary Pattern (LBP) for feature extraction and clustered road segments using the K-means algorithm, evaluating cluster quality with the Davies-Bouldin Index (DBI) [4]. The findings indicated that combining LBP with K-means clustering significantly improved the ability to identify highly connected road segments, leading to more accurate traffic flow predictions using a Support Vector Machine (SVM) model.

Moreover, by comparing LBP-enhanced clustering with other methods like Principal Component Analysis (PCA) with K-means, the study demonstrated that the LBP approach achieved lower DBI scores, indicating better clustering cohesion and separation. Traffic state predictions made using the LBP-K-means-SVM combination outperformed predictions from models that only used K-means or PCA-K-means, thereby validating LBP's effectiveness in refining traffic condition predictions for urban networks.

One study reviewed the Grey Wolf Optimizer (GWO) with title "Review of Grey Wolf Optimizer" by Dhruva Shaw, a bio-inspired metaheuristic algorithm developed to solve complex optimization problems by simulating the hierarchical and cooperative hunting behaviors of grey wolves [5]. GWO employs roles within a virtual pack (alpha, beta, delta, and omega wolves) to balance exploration and exploitation during the search for optimal solutions, making it effective in both diverse and continuous optimization tasks. This model has been widely used in machine learning, engineering, and signal processing, demonstrating strengths such as simplicity, high precision, and fast convergence in finding accurate solutions within search spaces.

Further advancements in GWO include hybrid approaches that integrate other algorithms, like Particle Swarm Optimization, to enhance performance and robustness. Modified versions, such as the Binary Grey Wolf Optimizer (BGWO), have been applied to specific fields like feature selection and hyperparameters tuning in machine learning, resulting in improved accuracy and efficiency. In fields

requiring computationally intensive calculations, GWO's ability to adapt and reduce complexity has enabled effective implementations for electromagnetic field applications and antenna design. These adaptations highlight GWO's utility and potential for broader optimization applications.

A study titled "Breast Cancer Diagnosis Using Support Vector Machine Optimized by Improved Quantum Inspired Grey Wolf Optimization" addresses the urgent need for early and accurate breast cancer diagnosis by optimizing support vector machine (SVM) models with an advanced hybridization of Grey Wolf Optimizer (GWO) techniques [6]. This study combines a Quantum-Inspired Binary Grey Wolf Optimizer (IQI-BGWO) with an SVM to enhance breast cancer classification on the MIAS dataset. By leveraging quantum computing principles in the feature selection process, the study achieved a significant improvement in accuracy, with mean accuracy, sensitivity, and specificity reaching up to 99.25%, 98.96%, and 100%, respectively, using tenfold cross-validation.

This research demonstrates that the hybrid IQI-BGWO-SVM approach surpasses traditional methods like Particle Swarm Optimization and Genetic Algorithm in breast cancer classification, making it a highly effective computational model. By utilizing IQI-BGWO for feature selection, the research not only boosts classification accuracy but also enhances SVM's overall stability in processing complex, high-dimensional medical imaging data. Consequently, this study offers valuable insights and advancements in automated breast cancer detection, promoting early and precise diagnosis through the proposed hybrid model.

One of the studies titled "Optimizing Sentiment Analysis for Bekasi Flood Management Using SVM and Naïve Bayes with Advanced Feature Selection" achieved an accuracy of 92.37% using the SVM algorithm, while the Naïve Bayes algorithm achieved an accuracy of 89.21% [7]. Based on previous research, it has been proven that the Naïve Bayes and SVM algorithms can perform well in predicting diabetes.

In addition to the use of classification, the method that can be implemented in this research is the application of feature selection using the Binary Grey Wolf Optimizer on the data distribution. One of the studies titled "A Novel Hybrid IoT Based IDS Using Binary Grey Wolf Optimizer (BGWO) and Naïve Bayes (NB)" achieved an accuracy of 99.15% using BGWO-NB, while the one that did not use feature selection achieved an accuracy of 90.6% [8]. The study shows that the use of feature selection can improve accuracy in classification.

Therefore, this research is conducted with the title "Comparison of SVM and Naïve Bayes using Binary Grey Wolf Optimizer for Diabetes Mellitus Prediction," where BGWO will be used to select attributes first. The results of the feature selection will be used by SVM and Naïve Bayes to achieve the highest accuracy among the two classifications.

The objective of this research is to improve the accuracy of diabetes disease classification by comparing the SVM and Naïve Bayes algorithms, where a high accuracy value determines the effectiveness of the algorithm in predicting diabetes mellitus. Additionally, the use of BGWO can enhance the performance of SVM and Naïve Bayes in terms

of the stability of classification results.

II. RELATED WORK

This research will examine the optimization approach employed for feature selection via the Binary Grey Wolf Optimizer (BGWO), subsequently utilizing classification techniques including Support Vector Machine (SVM) and Naïve Bayes.

A. Diabetes Mellitus

Diabetes mellitus is a common medical condition, although it has the potential to cause adverse effects, the prevalence of which has increased over the past few decades and has become a major public health challenge in the 21st century [9]. The reason is that this disease not only affects the elderly but has also become a problem for young people. This disease has existed for a long time, with the Ebers papyrus dating back to 1500 BC describing a polyuric condition resembling diabetes [10].

This disease itself has already become rampant. Diabetes patients are twice as likely to be hospitalized, and one in six hospitalized patients has diabetes; in some hospitals, diabetes patients occupy 25% of all inpatient beds [11]. For diabetes-related mortality, it is estimated to account for 11.3% of global deaths among adults aged 20-79. This percentage varies by region, from 6.8% in the African Region to 16.2% in the Eastern Mediterranean and North Africa Region [12]. The percentage is indeed not as high as the percentages of other deadly diseases like cancer or heart attacks, but still, for the following diabetes disease, special and prompt handling is required to prevent further spread.

B. Data Mining

Data mining is a technique used to determine the types of patterns or knowledge that will be discovered in the process [13]. Data mining can be used in data prediction, such as estimating the cause of an event, which can ultimately be studied in the future in the analysis of various matters. The following are types of data mining:

- Supervised Learning

This approach instructs the machine utilizing a labeled dataset including compressed input-output pairs [14]. This diagram is the most commonly used diagram.

- Unsupervised Learning

Unsupervised learning is a machine learning method that works by finding patterns or structures in data that do not have labels or correct answers. In contrast to supervised learning, which relies on labeled data to train the model, unsupervised learning seeks to interpret the data independently without any guidance from a supervisor who provides correct answers or points out errors for each observation [15].

C. Data Collection

Data collection is the initial step in research analysis, where the data must include relevant features. When the

data has been collected, data selection can be performed. Careful data selection becomes crucial, especially when considering the challenges that arise from the presence of very large datasets in the context of machine learning and artificial intelligence.

D. Preprocessing

Preprocessing is the first step in data analysis, consisting of a series of actions or methods to clean, organize, and prepare the raw data before conducting more in-depth analysis. The goal of data preprocessing is to obtain more structured results in the desired format. This process also aims to adjust the data so that it can be used with various types of algorithms during the processing and implementation stages.

E. Binary Grey Wolf Optimizer (BGWO)

Binary Grey Wolf Optimizer (BGWO) is an optimization algorithm inspired by the social behavior of grey wolves. BGWO will select the top three wolves with the best quality: alpha, beta, and delta [16]. BGWO simulates the hunting and prey-tracking behaviors of grey wolves in the wild. Feature selection aims to identify a relevant subset from a broader set of features. The primary objective of this algorithm is to enhance model performance by utilizing only the most informative features, reducing data dimensionality by removing irrelevant or redundant features. This helps enhance model performance, reduce overfitting, and improve interpretability.

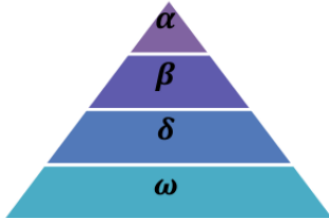


Figure 1. Hierarchy of Grey Wolf

Picture from The Grey Wolf Optimizer for Antenna Optimization Designs: Continuous, binary, single-objective, and multiobjective implementations

The working mechanism of BGWO is through parameter initialization. At this stage, the algorithm parameters such as the number of wolves and the number of iterations are set with appropriate initial values. Next, initialization is performed with random binary values for the initial wolf population. This binary value represents the presence or absence of a feature in the selected feature subset.

After the wolf population is formed, the next stage is fitness evaluation. At this stage, the fitness value of each wolf is calculated using a predetermined fitness function. This fitness function is responsible for measuring the quality of the feature subset selected by the wolves. The wolf with the highest fitness is designated as the alpha, the second-highest as the beta, and the third-highest as the delta. The alpha wolf is responsible for leading the pack. Then, assisted by the beta and delta wolves in making decisions about determining prey at certain times.

Then, the process of updating the wolf positions

continues for several predetermined iterations, where the linear parameter ranges from 2 to 0 to control the level of exploration and exploitation. Here is the formula for calculating the parameter value, which is useful for controlling the scale of change so that the solution is more exploratory at the beginning of the iteration and more exploitative at the end of the iteration.

$$\alpha = 2 - \left(\frac{2t}{Max\ iter} \right) \quad (1)$$

Wolves will detect the location where prey is present and immediately circle around the prey. Each wolf's position is updated based on its distance from the alpha, beta, and delta wolves. Here is the formula to calculate A to introduce a random value controlled by parameter a. Meanwhile, value C is the coefficient matrix used to obtain random variation.

$$C = 2 \cdot rand_2 \quad (2)$$

$$A = 2 \cdot \alpha \cdot rand_1 - a \quad (3)$$

Below is the formula for calculating D, which aims to determine how far the current solution (alpha, beta, delta) is from the desired target or position. Value D helps the wolf determine the direction and distance of movement towards the optimal target.

$$D_\alpha = |C \cdot X_\alpha - X| \quad (4)$$

$$D_\beta = |C \cdot X_\beta - X| \quad (5)$$

$$D_\delta = |C \cdot X_\delta - X| \quad (6)$$

The positions of the alpha, beta, and delta wolves will change during iterations due to the wolf position update process. This position update process is intended to guide the wolves towards better solutions represented by the alpha position. Wolves that are closer to the alpha tend to have a greater chance of influencing the overall solution discovery. Therefore, the alpha wolf is chosen because it has a combination of proximity to the best solution and good fitness value. This makes it the group leader and guides the movement of other wolves in the search for better solutions. Here is the formula for refining the solution by considering the influence of the optimal solutions alpha, beta, and delta, and parameter A.

$$X_1 = X_\alpha - A_1 \cdot D_\alpha \quad (7)$$

$$X_2 = X_\beta - A_2 \cdot D_\beta \quad (8)$$

$$X_3 = X_\delta - A_3 \cdot D_\delta \quad (9)$$

After the iteration is complete, the selected features come from the alpha wolf's features. The features chosen by the alpha wolf in the last iteration are considered the most optimal for the feature selection problem. This is because during the iteration process, the wolves' positions are updated and evaluated periodically to approach the optimal solution.

1. Initialize the value of the variable N grey wolf from the population $X_i (i = 1, 2, \dots, n)$
2. Initialize parameters A, a and C
3. Find the fitness value of the initial population that has been created
4. $\vec{X}_\alpha \sim$ the 1st best search agency
5. $\vec{X}_\beta \sim$ the 2nd best search agency
6. $\vec{X}_\delta \sim$ the 3rd best search agency
7. While $t <$ maximum iteration value
8. For all population fitness
9. Update the position with the value of t using the equation
10. Calculation of the position of each gray wolf
11. End For
12. Update parameters BGWO (A, a and C)
13. Update $\vec{X}_\alpha, \vec{X}_\beta, \vec{X}_\delta$
14. Gray wolf position evaluation
15. Update $\vec{X}_\alpha, \vec{X}_\beta, \vec{X}_\delta$
16. $t = t + 1$
17. End While
18. Return \vec{X}_α

F. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a statistical learning method introduced by Vapnik et al., founded on the principles of structural risk minimization and high-dimensional space theory. It transforms the solution of convex optimization problems into linear programming outcomes, thus simplifying the problem-solving process [18].

SVM is a widely used and effective machine learning algorithm for classification and regression tasks. Its primary focus is on classification, aiming to separate data into two distinct classes using a hyperplane. The objective is to maximize the margin (distance) between the hyperplane and the nearest data points from each class. A hyperplane represents a higher-dimensional space; for instance, in two dimensions, it is a line, while in three dimensions, it is a plane that divides the feature space into two sections containing data points from both classes.

The key concept in SVM is to identify a hyperplane that maximizes the margin, which is the distance between the hyperplane and the nearest data points from each class. These data points are referred to as support vectors, as they play a crucial role in determining the position and orientation of the hyperplane. By maximizing the margin, SVM aims to develop the most generalizable model for classifying new, unseen data. For SVM classification, the model can be represented by the mathematical equation:

$$w \cdot x_i + b = 0 \tag{10}$$

SVM utilizes the weight vector and the bias to identify the hyperplane that separates the two classes with the maximum margin. In general, the SVM training process involves optimizing the model parameters (for example, the weight vector W and the bias b using methods such as gradient optimization or quadratic optimization. The main goal of the training is to find the optimal hyperplane that can

separate the data with the largest margin, so that the SVM model can generalize well to new, unseen data.

Pseudocode SVM Algorithm [19]

1. Input: $D=[X,Y]$; X (array of input with m features), Y (array of class labels)
 2. $Y = \text{array}(C)$ // Class label
 3. Output: Find the performance of the system
 4. function $\text{train_svm}(X, Y, \text{number_of_runs})$
 5. initialize $\text{learning_rate} = \text{Math.random}()$;
 6. for learning_rate in number_of_runs
 7. $\text{error} = 0$;
 8. for i in X
 9. if $Y[i] \cdot (X[i] \cdot w) < i$ then
 10. update $w = w + \text{learning_rate} \left(x[i] \cdot y[i] - \frac{2}{\text{number_of_runs}} \right)$
 11. else
 12. update $w = w + \text{learning_rate} \left(- \frac{2}{\text{number_of_runs}} \right)$
 13. end if
 14. end
 15. end
-

Naïve Bayes

The Naïve Bayes method is one of the classification algorithms based on Bayes' Theorem with the assumption of independence among predictors [20]. The advantage of this algorithm can be used for databases with large data and quick model creation. The formula of this algorithm is stated as follows:

$$P(H|X) = \frac{P(X|H) P(H)}{P(X)} \tag{11}$$

G. Metrics Evaluation

After the calculation process with both algorithms, the next step is evaluation by calculating accuracy, precision, recall, and F1-Score by comparing the predicted values with the actual values. Here is the Confusion Matrix in Matlab for performance evaluation:

Table 1. Confusion Matrix

		Predicted Label	
		0	1
True Label	0	True Negative (TN)	False Positive (FP)
	1	False Negative (FN)	True Positive (TP)

Here is the explanation of Accuracy, Precision, Recall, and F1-Score.

- Accuracy

Accuracy is the ratio of correct predictions (both positive and negative) to the total data. This accuracy answers questions regarding Diabetes. Here is the accuracy formula.

$$Accuracy = \frac{TP+TN}{(TP+TN+FN+FP)} \times 100 \quad (12)$$

- *Precision*

Precision is an evaluation metric used in machine learning to assess the accuracy of a classification model. Specifically, it measures the ratio of true positive predictions to the total number of instances predicted as positive, which includes both true positives and false positives. Precision is important because it reduces false positives, where someone is mistakenly diagnosed with diabetes. In mathematical terms, precision can be calculated using the following formula:

$$Precision = \frac{TP}{(FP+TP)} \quad (13)$$

- *Recall*

Recall is the proportion of true positive predictions relative to the total number of actual positive instances. This metric is very important to show errors where someone actually has diabetes but is classified as negative. (false negatives). To calculate recall, use the following formula:

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

- *F1-Score*

The F1-Score is a metric utilized to evaluate the performance of a machine learning model. The F1-Score combines two important metrics, namely precision and recall, by using the average of both. This makes the F1-Score a useful tool for evaluating models, especially when considering the balance between precision and recall. Here is the formula:

$$F_1 \text{ Score} = \frac{2(Recall \times Precision)}{(Recall+Precision)} \quad (15)$$

III. METHODOLOGY

If This research utilizes secondary data as the main source of information, where the data is not obtained directly, but from internet sites. The dataset utilized is the PIDD, which contains information on 768 women from the Pima Indian tribe. This Pima Indian Diabetes dataset specifically includes data on Pima Indian women aged 21 and older, is a popular benchmark dataset [21]. Here is the research flow for predicting diabetes mellitus.

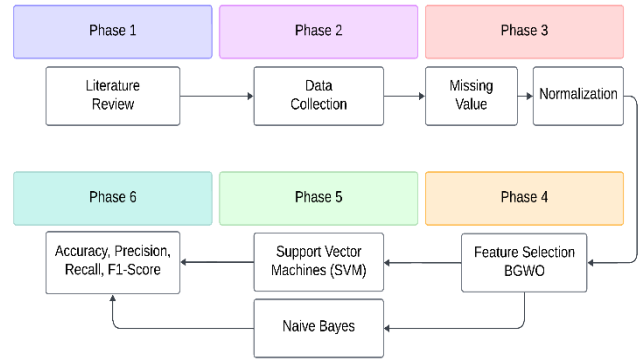


Figure 2. Research Flow Diagram

The first stage in the research is conducting a literature review aimed at identifying the problem in depth and gathering information from previous publications. This step helps in understanding the theories and concepts relevant to the research that will be conducted. The next step is the search and collection of relevant datasets. For this research, the dataset used is the Pima Indians Diabetes Dataset.

Once the dataset is obtained, preprocessing is performed to clean and prepare the data. This preprocessing includes various steps, this involves activities such as eliminating missing values and applying normalization methods like Min-Max Scaling to ensure that all features are scaled uniformly. Then, feature selection was also performed, which involves selecting the most relevant features to improve the model's performance. The BGWO (Binary Grey Wolf Optimizer) algorithm can be used in feature selection to identify which variables have the most influence in predicting diabetes.

The final stage is modeling and evaluation. The SVM and NV algorithms serve as classification techniques for creating prediction models. Once the model is trained, its performance is assessed using metrics like accuracy, precision, recall, and F1-score. This evaluation is crucial to verify that the model can yield results applicable in real-world scenarios.

IV. RESULT

The PIDD dataset was obtained from a public dataset and consists of 768 rows and 9 features. The first thing that must be done is data cleaning to eliminate invalid or inconsistent values, such as zero values that can provide misleading results or NaN (Not a Number) values that indicate missing data. Next, normalization using techniques such as Min-Max Scaling.

A. Feature Selection with BGWO

Once the normalization process is finished, the next step is to proceed to the feature selection stage, which is the final phase of data preprocessing. During this stage, the parameters of the BGWO algorithm will be configured as follows.

Table 2. Parameters BGWO

BGWO Parameters	Value
N	8

Max Iteration	60
---------------	----

Components		
------------	--	--

The value of N represents the number of wolves utilized in the BGWO algorithm, while Max Iteration refers to the maximum number of repetitions allowed in the algorithm. This method aims to determine the best features that contribute to the model's performance. The following are the results of the features selected using the BGWO method.

Table 3. Feature Selection BGWO

Feature	Feature Choice
Pregnancies	✓
Glucose	✓
Blood Pressure	
Skin Thickness	
Insulin	✓
BMI	
Diabetes Pedigree Function	✓
Age	
Outcome	✓

The results of feature selection using BGWO show that five features selected for further analysis are the Number of Pregnancies, Glucose Concentration, Insulin Concentration, Family History of Diabetes, and Outcome. These features can provide important information in predicting the condition to be analyzed. By only using the selected features, the built model becomes more efficient and accurate, and reduces the risk of overfitting that often occurs when using too many features.

B. Classification Algorithm

After the feature selection process, the next step involves modeling the SVM and NV algorithms using 10 fold cross validation. In this approach, the dataset is split into 10 folds, with one fold serving as the test data during each iteration, while the remaining 9 folds are used to train the model. This process is carried out 10 times so that each fold becomes the test data once, and the accuracy results from each iteration are averaged to obtain the overall performance. Here are the parameters used

Table 4. Parameter Classification Algorithm

Parameters	SVM	Naïve Bayes
Kernel Function	Quadratic	
Box Constraint	4	
Kernel Scale	5	
Kernel Type		Gaussian
Support		Unbounded
Component Reduction Criterion		Specify Number of Components
Number of Numeric		1

This is the results of the tests conducted on the classification algorithms.

Table 5. Comparison Between SVM, and Naïve Bayes

Algorithm	Pima Indian Dataset			
	Accuracy	Precision	Recall	F1-Score
SVM-BGWO	73,30%	85,74%	55,90%	67,78%
SVM	72,60%	83,93%	55,90%	67,11%
NV-BGWO	72,60%	84,88%	55,00%	66,75%
NV	70,80%	83,02%	52,30%	64,17%

The research results show that the application of BGWO to the classification algorithm provides a significant performance improvement. In addition, the comparison of results between SVM-BGWO and Naïve Bayes-BGWO (NV-BGWO) on the Pima Indian Diabetes dataset shows that SVM-BGWO outperforms in various evaluation metrics. SVM-BGWO has an accuracy of 73.30%, while NV-BGWO recorded an accuracy of 72.60%. This difference indicates that SVM-BGWO is more effective in producing correct predictions. Additionally, SVM-BGWO also recorded a higher precision of 85.74% compared to the precision of NV-BGWO, which is only 84.88%. This shows that SVM-BGWO is better at avoiding classification errors, thus providing more reliable results in the context of classification.

V. CONCLUSIONS

Large Based on the research results conducted by comparing the Support Vector Machine (SVM) and Naïve Bayes (NB) algorithms, both with and without using the Binary Grey Wolf Optimizer (BGWO) method for diabetes mellitus prediction, The following are the conclusions of the research:

1. The research results indicate that the application of BGWO to the classification algorithms provides a significant performance improvement.
2. The comparison results between SVM-BGWO and Naïve Bayes-BGWO indicate that SVM-BGWO has better performance. In terms of accuracy, SVM-BGWO achieved 73.30%, while NV-BGWO recorded an accuracy of 72.60%. These results show that SVM-BGWO is overall more effective in classification compared to NV-BGWO.

REFERENCES

- [1] E. Gulshan Tokhirova, "RISK FACTORS FOR DEVELOPING TYPE 2 DIABETES MELLITUS," 2024. [Online]. Available: <http://www.newjournal.org/>

- [2] K. Arumugam, M. Naved, P. P. Shinde, O. Leiva-Chauca, A. Huaman-Osorio, and T. Gonzales-Yanac, "Multiple disease prediction using Machine learning algorithms," *Mater Today Proc*, vol. 80, pp. 3682–3685, Jan. 2023, doi: 10.1016/j.matpr.2021.07.361.
- [3] M. I. Akazue, G. A. Nwokolo, O. A. Ejaita, C. O. Ogeh, and E. Ufiofio, "Machine Learning Survival Analysis Model for Diabetes Mellitus," 2023. [Online]. Available: www.ijisrt.com/754
- [4] B. Priambodo, R. A. Kadir, and A. Ahmad, "Clustering Urban Roads Using Local Binary Patterns to Enhance the Accuracy of Traffic Flow Prediction," vol. 14, no. 5, 2024.
- [5] D. Shaw, "Review of Grey Wolf Optimizer," 2024, doi: 10.13140/RG.2.2.14111.57763.
- [6] A. Bilal, A. Imran, T. I. Baig, X. Liu, E. Abouel Nasr, and H. Long, "Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization," *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-61322-w.
- [7] A. Amali, D. Maulana, E. Widodo, A. Firmansyah, and M. Danny, "Optimizing Sentiment Analysis of Bekasi Flood Management Using SVM and Naive Bayes with Advanced Feature Selection," *Brilliance: Research of Artificial Intelligence*, vol. 4, no. 1, pp. 362–371, Jul. 2024, doi: 10.47709/brilliance.v4i1.4268.
- [8] E. ÜLKER and I. M. Nur, "A Novel Hybrid IoT Based IDS Using Binary Grey Wolf Optimizer (BGWO) and Naive Bayes (NB)," *European Journal of Science and Technology*, Oct. 2020, doi: 10.31590/ejosat.804113.
- [9] D. Tomic, J. E. Shaw, and D. J. Magliano, "The burden and risks of emerging complications of diabetes mellitus," Sep. 01, 2022, *Nature Research*. doi: 10.1038/s41574-022-00690-7.
- [10] R. Bilous, R. Donnelly, and I. Idris, "Hanbook of Diabetes," 2021.
- [11] C. A. Whicher, S. O'Neill, and R. I. G. Holt, "Diabetes in the UK: 2019," *Diabetic Medicine*, vol. 37, no. 2, pp. 242–247, Feb. 2020, doi: 10.1111/dme.14225.
- [12] P. Saedi *et al.*, "Mortality attributable to diabetes in 20–79 years old adults, 2019 estimates: Results from the International Diabetes Federation Diabetes Atlas, 9th edition," *Diabetes Res Clin Pract*, vol. 162, Apr. 2020, doi: 10.1016/j.diabres.2020.108086.
- [13] M. K. Gupta and P. Chandra, "A comprehensive survey of data mining," *International Journal of Information Technology (Singapore)*, vol. 12, no. 4, pp. 1243–1257, Dec. 2020, doi: 10.1007/s41870-020-00427-7.
- [14] R. Verma, V. Nagar, and S. Mahapatra, "THE COMMENCEMENT OF MACHINE LEARNING SOLICITATION TO BIOINFORMATICS," 2021.
- [15] D. S. Watson, "On the Philosophy of Unsupervised Learning," *Philos Technol*, vol. 36, no. 2, Jun. 2023, doi: 10.1007/s13347-023-00635-6.
- [16] D. Wang, Y. Ji, H. Wang, and M. Huang, "Binary grey wolf optimizer with a novel population adaptation strategy for feature selection," *IET Control Theory and Applications*, vol. 17, no. 17, pp. 2313–2331, Nov. 2023, doi: 10.1049/cth2.12498.
- [17] P. Hu, J. S. Pan, and S. C. Chu, "Improved Binary Grey Wolf Optimizer and Its application for feature selection," *Knowl Based Syst*, vol. 195, May 2020, doi: 10.1016/j.knosys.2020.105746.
- [18] Y. Yu *et al.*, "Quantitative analysis of multiple components based on support vector machine (SVM)," *Optik (Stuttg)*, vol. 237, Jul. 2021, doi: 10.1016/j.ijleo.2021.166759.
- [19] K. Harimoorthy and M. Thangavelu, "Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system," Mar. 01, 2021, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s12652-019-01652-0.
- [20] M. R. Islam, S. Banik, K. N. Rahman, and M. M. Rahman, "A comparative approach to alleviating the prevalence of diabetes mellitus using machine learning," *Computer Methods and Programs in Biomedicine Update*, vol. 4, Jan. 2023, doi: 10.1016/j.cmpbup.2023.100113.
- [21] V. Chang, J. Bailey, Q. A. Xu, and Z. Sun, "Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms," *Neural Comput Appl*, vol. 35, no. 22, pp. 16157–16173, Aug. 2023, doi: 10.1007/s00521-022-07049-z.

Andi Nugroho

Phd Student in Computer Sciences – Computer Science Department, Mercu Buana University, Jakarta, Indonesia
Email : andi.nugroho@mercubuana.ac.id

Scopus Author ID : 57208427717

ORCID : [orcidID = https://orcid.org/0000-0002-1713-035X](https://orcid.org/0000-0002-1713-035X).

Berliana Fajrina

Computer Science Student - Department of Information Systems, Universitas Mercu Buana, Jakarta, Indonesia
Email : 41821010098@mercubuana.ac.id

Safina Faradilla Hasibuan

Computer Science Student - Department of Information Systems, Universitas Mercu Buana, Jakarta, Indonesia
Email : 41821010110@mercubuana.ac.id