Implementation of Glowworm Swarm Optimization-SVM Method for Predicting Antiproliferative Activity Against MDA-MB-231 Cell Line

Nurfaidzi Ramdhani Arifin
School of Computing
Telkom University
Bandung, Indonesia
faidzir@student.telkomuniversity.ac.id

Angel Metanosa Afinda
School of Electrical Engineering
Telkom University
Bandung, Indonesia
angelmetanosa@telkomuniversity.ac.id

Isman Kurniawan
School of Computing
Telkom University
Bandung, Indonesia
ismankrn@telkomuniversity.ac.id

***Abstract*—Breast cancer remains one of the leading causes of mortality worldwide, particularly among women, with more than 300,000 new cases reported annually. Conventional drug screening for anticancer compounds is costly and time-consuming, motivating the adoption of machine-learning-based in silico approaches. SMILES representations are widely used in predictive molecular modeling. However, the feature space generated from SMILES descriptors is often high-dimensional and not all features contribute meaningfully to classification performance. Therefore, feature selection is essential to reduce dimensionality, improve learning efficiency, and enhance model accuracy. This study develops a Support Vector Machine (SVM) model to classify antiproliferative activity of SMILES-based compounds while integrating Glowworm Swarm Optimization (GSO) as a metaheuristic feature selection strategy. An initial Variance Threshold reduced more than 800 molecular descriptors to fewer than 200, and GSO further refined the subset with reduction rates of 55–57%. All SVM kernels demonstrated performance improvement after feature reduction, with the RBF kernel achieving the most stable generalization results with 0.85 accuracy using GSO-selected features, which increased to 0.86 following hyperparameter tuning. These findings indicate that swarm-based feature optimization can effectively reduce dimensionality while maintaining strong predictive capability. Overall, the proposed GSO–SVM framework shows promising potential for in silico anticancer screening, with future research directed toward expanding evaluation across different cancer cell lines and exploring more advanced molecular representation techniques.**

***Keywords*—breast Cancer, MDA-MB-231, antiproliferative activity, glowworm swarm optimization, support vector machine**

I. Introduction

Cancer remains one of the leading causes of death worldwide and continues to pose a major global health burden, consistently ranking among the third to fourth highest causes of mortality [1]. In 2020, an estimated 2.3 million new breast cancer cases were reported, accounting for 11.7% of all newly diagnosed cancers, with a total of 684,996 recorded deaths [2]. In the United States, breast cancer stands as the second most common cause of cancer-related death among women, following lung cancer, with approximately 316,000 new cases reported annually [3].

The MDA-MB cell line, a representative model of Triple-Negative Breast Cancer (TNBC), one of the most aggressive breast cancer subtypes, is widely used in studies investigating antiproliferative activity of natural products [4,5]. However, conventional drug discovery approaches are often limited by high costs, long development cycles, and relatively low efficacy. These challenges have motivated the adoption of in silico strategies, particularly Machine Learning (ML)-based models, as promising alternatives to improve research efficiency while reducing financial burden and minimizing risks associated with preclinical and clinical testing [6,7].

Several studies have previously explored the use of MDA-MB-231 breast cancer cells as a model for antiproliferative screening and computational drug discovery. In 2017, He *et al*. developed a machine learning model for predicting compound activity in MDA-MB-231 cells, reporting 84% accuracy with an AUC greater than 0.88 [8]. A few years later, Nada *et al*. in 2023 introduced a dual-cell evaluation framework for MCF-7 and MDA-MB-231, where Random Forest outperformed other models with $R^2 > 0.7$ [9]. Another contribution came from Islam *et al*. who conducted a comparative study across five ML classifiers for breast cancer prediction, with Artificial Neural Networks reaching the highest accuracy (98.57%) and Logistic Regression the lowest (95.7%) [11].

Based on the literature survey, numerous studies have implemented conventional ML models for breast cancer detection and drug discovery, achieving relatively high predictive performance. However, most of these works do not incorporate systematic hyperparameter optimization or feature selection during model development, resulting in suboptimal learning efficiency. The implementation of feature selection methods in QSAR modeling is essential to enhance the predictive power and accuracy of the developed models. For instance, Alkady et al. in 2019, successfully reduced dimensional complexity from 1666 features to just 8 significant attributes, achieving

an accuracy of 95% [12]. In 2021, Ewees et al. introduced a novel feature selection method by hybridizing the Slime Mould Algorithm and the Firefly Algorithm (SMAFA). When evaluated on QSAR modeling datasets, this approach achieved accuracies of 0.84 for H1N1 and 0.96 for Hepatitis [13]. Considering that the dataset used in this study contains a large number of molecular descriptors which won't contribute many to classification performance, a metaheuristic Glowworm Swarm Optimization (GSO) approach is adopted to reduce feature dimensionality and identify the most informative subset of features. The GSO algorithm has been widely reported as an effective feature selection strategy, capable of efficiently exploring the search space and determining optimal feature combinations [14].

This study aims to predict the antiproliferative activity of compounds against MDA-MB-231 breast cancer cells using a hybrid Glowworm Swarm Optimization–Support Vector Machine (GSO–SVM) approach. To the best of our knowledge, this is one of the first studies to implement this specific hybrid framework for MDA-MB-231 activity screening. GSO is employed as the optimization method due to its simple mechanism, its ability to avoid local optima, and its effectiveness in identifying optimal feature combinations through adjustable parameters such as the number of agents and the search radius. Meanwhile, SVM serves as the classification algorithm responsible for identifying the optimal hyperplane that accurately separates classes, offering strong generalization capability and robustness against overfitting. By integrating the strengths of both methods, the resulting model is expected to provide more optimal and reliable predictions of antiproliferative activity.

II. Materials and Methods

A. Dataset

This study utilizes a dataset obtained from the PubChem repository maintained by the U.S. National Library of Medicine, consisting of 762 chemical compounds whose antiproliferative activity against MDA-MB breast cancer cells was evaluated using the MTT assay over a 48-hour incubation period [15]. Each compound is represented using the Simplified Molecular Input Line Entry System (SMILES), a textual encoding of molecular structure, and is associated with an IC_{50} value measured in µM. Compounds with $IC_{50} < 1.7\,\mu M$ are categorized as inhibitors, whereas those with $IC_{50} \geq 1.7\,\mu M$ are classified as neutral.

Molecular descriptors were generated from SMILES representations utilizing the PaDEL-Descriptor library. Specifically, the compounds were encoded as PubChem fingerprints, transforming each molecular structure into a high-dimensional binary vector. In this representation, a bit value of 1 signifies the presence of a specific chemical substructure, while a value of 0 indicates its absence. This approach resulted in an initial feature set of over 800 descriptors, which were subsequently refined using a Variance Threshold ($\sigma^2 < 0.1$) to eliminate non-informative or constant attributes. As illustrated in Fig. 1, this preprocessing stage establishes a structured feature space prior to the application of Glowworm Swarm Optimization for final feature selection

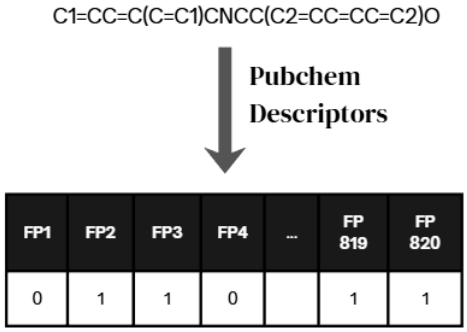

Fig. 1. Conversion of SMILES strings into PubChem binary fingerprints

The dataset was randomly partitioned into two subsets using an 80%–20% split, resulting in 609 samples for model training and 153 samples for independent testing. The overall class distribution is illustrated in Fig. 2.

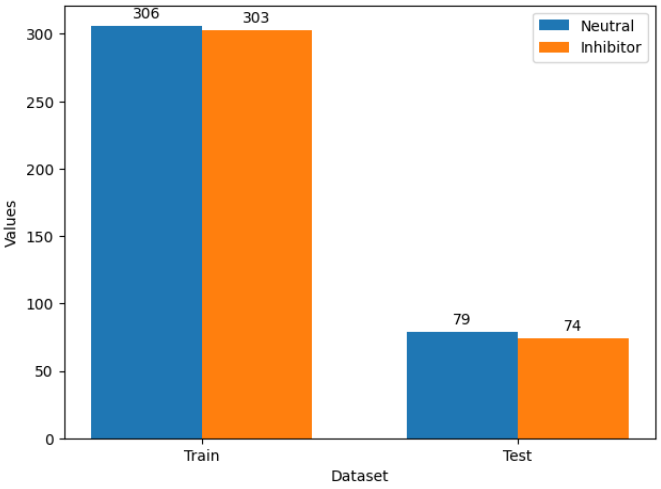

Fig. 2. Dataset distribution

B. Feature Selection

Glowworm Swarm Optimization (GSO) is a swarm-intelligence–based optimization algorithm introduced by Krishnanand and Ghose in 2008 to address multimodal optimization problems, where the objective function contains multiple local optima [16].

In GSO, each agent (referred to as a *glowworm*) maintains a luciferin value that reflects the quality of its position in the search space. This luciferin value is updated iteratively according to Equation (1)

$$L_i(t+1) = (1-\rho)L_i(t) + \gamma J\big(x_i(t)\big) \tag{1}$$

where $L_i(t)$ denotes the luciferin value of agent i at iteration t, ρ is the luciferin decay constant, γ is the luciferin

enhancement constant, and $J(x_i(t))$ represents the objective function value at the agent's current position.

After updating the luciferin values, each agent evaluates its neighboring agents within a local sensing radius and then probabilistically selects one neighbor with a higher luciferin level. The selection probability is defined as shown in Equation (2)

$$p_{ij}(t) = \frac{L_j(t) - L_i(t)}{\sum_{k \in N_i(t)} [L_k(t) - L_i(t)]}, \tag{2}$$

where $p_{ij}(t)$ denotes the probability that agent i selects agent j as its target neighbor, $L_j(t)$ and $L_i(t)$ represent the luciferin values of agents j and i, respectively, and $N_i(t)$ is the set of potential neighboring agents within the decision radius of agent i.

Once the target neighbor is selected, the agent updates its position by moving toward the chosen neighbor. The new position is computed as shown in Equation (3)

$$x_i(t+1) = x_i(t) + s \cdot \frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}, \tag{3}$$

where $x_i(t)$ denotes the position of agent i at iteration t, $x_j(t)$ represents the position of the selected neighboring agent, and s is the fixed step size.

The fitness value for each agent is determined as a weighted combination of classification accuracy and the proportion of utilized features, as formulated by Halladay and Dozier in Equation (4)

$$f_i = \omega \cdot (1 - \text{accuracy}) + (1 - \omega) \cdot \frac{N_{\text{selected}}}{N_{\text{total}}} \tag{4}$$

where f_i denotes the fitness score of the i-th agent, ω represents the feature reduction weight ($0 \leq \omega \leq 1$), N_{selected} is the number of selected features, and N_{total} is the total number of available features [17]. The parameter ω serves as a tuning factor to control the trade-off between maximizing classification accuracy and minimizing feature dimensionality.

To perform feature selection using GSO, several key parameters are configured as listed in Table I.

TABLE I. GLOWWORM SWARM OPTIMIZATION PARAMETERS

Parameter	Description	Value
`population_size`	Number of glowworms (agents) in the population	25
`max_iters`	Maximum number of iterations	80

C. Model Development

Support Vector Machine (SVM) is a widely used machine learning algorithm for various classification tasks. Introduced by Cortes and Vapnik in 1995, SVM aims to identify the optimal separating line or hyperplane that best distinguishes two data classes [18]. This hyperplane is constructed such that it maximizes the margin, i.e., the distance between the hyperplane and the closest data points from each class, thereby improving the model's generalization capability for unseen data.

Figure 3 illustrates the fundamental concept of SVM, where two groups of data points are separated by a hyperplane. The points lying on the boundary of the margin are referred to as support vectors, as they play a crucial role in determining the position and orientation of the separating hyperplane.

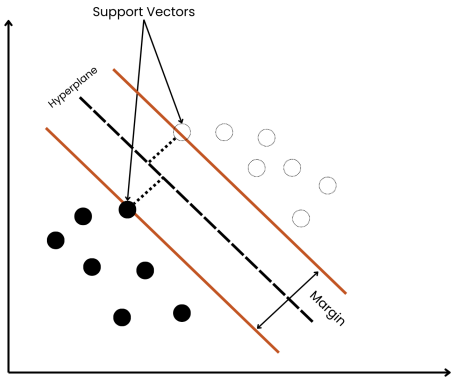

Fig. 3. Support Vector Machine illustration

For linearly separable data, SVM determines the optimal hyperplane by minimizing the norm of the weight vector $\mathbf{w}$, as shown in Equation (5). Each training instance must satisfy the constraint given in Equation (6).

$$\min_{\mathbf{w},b} \frac{1}{2}\|\mathbf{w}\|^2 \tag{5}$$

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \tag{6}$$

The corresponding separating hyperplane is defined in Equation (7),

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \tag{7}$$

and the two margin boundaries can be written in Equation (8).

$$\mathbf{w} \cdot \mathbf{x} + b = \pm 1 \tag{8}$$

Kernel functions enable SVM to handle data that are not linearly separable by implicitly transforming them into a higher-dimensional feature space. Commonly used kernels and their corresponding equations are listed in Table II.

TABLE II. KERNEL FUNCTIONS EQUATIONS

Kernel	Equation
Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^d$
RBF	$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$

Each kernel type possesses specific parameters that necessitate precise adjustment to attain optimal classification performance. To address this, a hyperparameter tuning process is utilized to identify the most effective parameter configuration

within a predefined search space [19]. In this study, GridSearch is employed to systematically evaluate all possible parameter combinations, a strategy aligned with the findings of Afinda *et al.*, who demonstrated that such optimization significantly enhances SVM model performance [20]. The tuning procedure is conducted for Linear, Polynomial, and RBF kernels, with their respective parameter ranges detailed in Table III.

TABLE III. SUPPORT VECTOR MACHINE HYPERPARAMETERS

Parameter	Value Range
C	$[0.001, 0.01, 0.1, 1, 10, 100, 1000]$
Degree	$[1, 2, 3, 4, 5]$
γ	{'auto', 'scale'}
Kernel	{Linear, Polynomial, RBF}

D. Model Validation

Model performance was systematically evaluated using four primary classification metrics, namely *accuracy*, *precision*, *recall*, and *F1-score*. These metrics are computed based on the *confusion matrix*, which summarizes the relationship between the actual class labels and the model's predicted outputs.

Table IV presents the structure of a standard confusion matrix.

TABLE IV. CONFUSION MATRIX

Actual	Predicted	
	Positive	**Negative**
Positive	TP (True Positive)	FN (False Negative)
Negative	FP (False Positive)	TN (True Negative)

The evaluation metrics are computed using Equations (9-12).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \tag{9}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{11}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{12}$$

Using these four metrics provides a comprehensive assessment of the model's classification performance, particularly when dealing with datasets exhibiting class imbalance. Among these metrics, accuracy is used as the primary indicator for selecting the best-performing model [22].

III. RESULTS AND DISCUSSION

A. Feature Selection

The Variance Threshold ($\sigma^2 < 0.1$) method was initially applied to remove low-variance features, as such features typically contribute little to the classification process. This preprocessing step reduced the original feature set from more than 800 to fewer than 200 features. GSO was then employed to further refine the feature subset by evaluating the fitness values associated with different feature combinations. As shown in Fig. 4, the fitness values for each kernel generally decrease after several initial iterations.

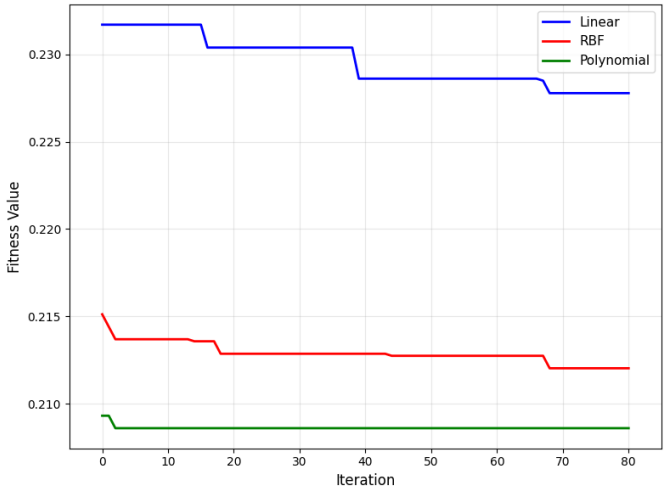

Fig. 4. Convergence Graph

Table V summarizes the number of selected features obtained from each kernel along with the corresponding reduction percentage relative to the original feature set. Both the Linear and RBF kernels produced identical feature subsets. The Polynomial kernel yielded a slightly different number of features, although the overall reduction rate remains comparable to the other kernels.

TABLE V. FEATURE SELECTION REDUCTION

Kernel	Selected Features	Reduction Percentage
Linear	124	55.7%
RBF	124	55.7%
Polynomial	121	56.8%

A comparison of model accuracy using the full feature set and the GSO-selected feature subset for each kernel is presented in Fig. 5. The Linear kernel exhibits the most substantial improvement, with an accuracy increase of nearly 0.2 compared to its full-feature baseline.

B. Model Development

After the feature selection process, the model underwent a hyperparameter tuning stage using Gridsearch Method to further improve its predictive capability by determining the most optimal SVM parameter configuration. This procedure allows

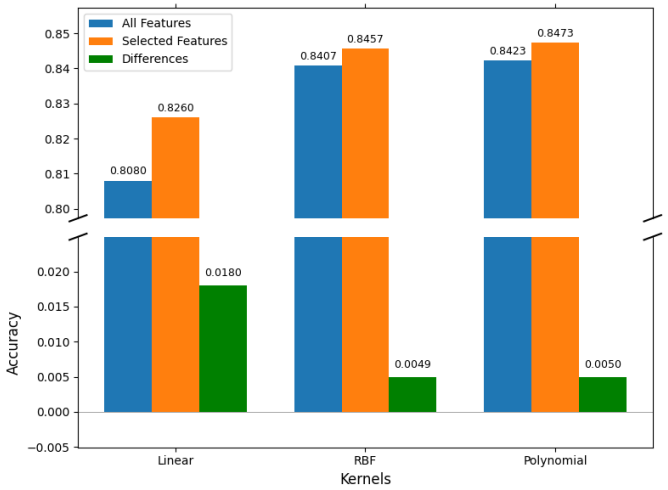

Fig. 5. Accuracy comparison using all features and selected features

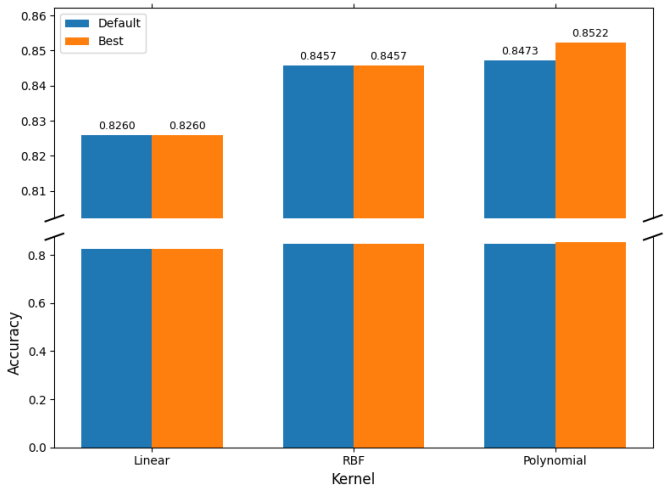

Fig. 6. Accuracy default parameters and best parameters

the model to identify optimal values of C, γ, and polynomial degree, thereby enhancing the decision boundary structure and reducing misclassification rates [20]. The selected hyperparameters for each kernel are presented in Table VI. Notably, the Polynomial kernel required parameter tuning to reach its best performance, resulting in an improvement of 0.6%, while the Linear and RBF kernels performed optimally without altering their default settings.

TABLE VI. HYPERPARAMETER TUNING RESULTS

Kernel	Parameter	Default Value	Best Value
Linear	C	1	1
Polynomial	C	1	100
	degree	3	2
	γ	scale	auto
RBF	C	1	1
	γ	scale	scale

As shown in Fig. 6, the improvement effect is primarily observed in the Polynomial kernel, where tuning reduced the model's sensitivity to feature scaling and adjusted the decision boundary to better capture non-linear class relationships. This finding aligns with the hyperparameter shift in Table VI, where increasing C and lowering the degree from 3 to 2 improved generalization by preventing excessive curve fitting on the training data.

C. Model Validation

Once all schemes were applied, each kernel was evaluated using a confusion matrix, as presented in Table VII. The evaluation considered four basic metrics, True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN), on both the training and testing data. From the training results, the Polynomial kernel achieved the highest TP and TN values, indicating better accuracy in classification. On the other hand, the Linear kernel showed the highest FP and FN, which led to the weakest performance among the kernels.

TABLE VII. CONFUSION MATRIX RESULTS

Model	TP	FP	TN	FN
Train Set				
Linear	266	35	264	44
Polynomial	**288**	13	**280**	28
RBF	269	32	268	40
Test Set				
Linear	64	12	63	14
Polynomial	**65**	11	65	12
RBF	**65**	11	**66**	11

A different trend is observed on the test set. The RBF kernel demonstrates the strongest generalization ability with an accuracy of 0.86, followed closely by the Polynomial kernel at 0.85. This suggests that although Polynomial performs most effectively during training, RBF generalizes better when evaluated on unseen data. The Linear kernel consistently shows the lowest performance in both evaluation stages.

A complete summary of evaluation metrics is provided in Table VIII. On the training set, Polynomial records the highest Accuracy, Precision, Recall, and F1-Score, reflecting not only high predictive capability but also classification stability. On the test set, the RBF kernel maintains the most balanced and consistent scores across all metrics (0.855 for Accuracy, Precision, Recall, and F1-Score). These results indicate that RBF is the most suitable kernel for achieving optimal generalization, particularly when combined with a reduced feature subset.

IV. CONCLUSION

This study demonstrated the effectiveness of Glowworm Swarm Optimization (GSO) as a feature selection method for enhancing Support Vector Machine (SVM) performance in predicting the antiproliferative activity of compounds against MDA-MB-231 breast cancer cells. After the initial Variance Threshold reduction, GSO successfully refined the feature

TABLE VIII. VALIDATION PARAMETER CALCULATION RESULTS

Model	Accuracy	Precision	Recall	F1-Score
Train Set				
Linear	0.870	0.858	0.884	0.871
Polynomial	**0.933**	**0.911**	**0.957**	**0.934**
RBF	0.882	0.871	0.894	0.882
Test Set				
Linear	0.830	0.821	0.842	0.831
Polynomial	0.850	0.844	**0.855**	0.850
RBF	**0.856**	**0.855**	**0.855**	**0.855**

subset further, improving model efficiency without compromising predictive capability. The evaluation of all three kernels revealed consistent performance enhancement after feature selection, with the Linear kernel achieving the most significant increase in accuracy at 0.18.

Hyperparameter tuning played an important role in refining model performance. Although the Linear and RBF kernels retained their default parameters as the optimal configuration, the Polynomial kernel achieved better results when tuned using alternative parameter values achieving 0,6% higher that default parameters. Among the three kernels, the RBF kernel produced the most balanced and consistent classification performance, achieving an Accuracy and F1-Score of 0.855, indicating strong generalization capability on unseen data.

Although the results are promising, the scope of this study remains confined to one breast cancer cell line (MDA-MB-231), and GSO parameter tuning has not been performed. As such, broader conclusions cannot yet be generalized. Future research may consider optimizing GSO parameters to enhance feature selection performance and validating the approach using other breast cancer cell lines.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209–249, 2021.

[2] R. Hong and B. Xu, "Breast cancer: An up-to-date review and future perspectives," Cancer Communications, vol. 42, pp. 913–936, 2022.

[3] K. P. Trayes and S. E. H. Cokenakes, "Breast cancer treatment," American Family Physician, vol. 104, no. 2, pp. 171–178, 2021, retrieved from https://www.aafp.org/afp.

[4] G. Ren, Z. Shi, C. Teng, and Y. Yao, "Antiproliferative activity of combined biochanin a and ginsenoside rh on mda-mb-231 and mcf-7 human breast cancer cells," Molecules, vol. 23, no. 11, p. 2908, Nov 2018.

[5] M. Kusmiyati, A. Rusdin, E. Trinovani, M. I. Prawira-Atmaja, D. B. Purkon, and G. P. E. Mulyo, "Cytotoxicity and antiproliferative activity of ethanol and ethyl acetate fractions from polymeric nanoparticles of green tea leaves (camellia sinensis) in breast cancer cell line mda mb-132," Journal of Advanced Pharmaceutical Technology & Research, vol. 13, no. 4, pp. 301–305, Oct–Dec 2022.

[6] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: machine intelligence approach for drug discovery," Molecular Diversity, vol. 25, no. 3, pp. 1315–1360, Aug 2021.

[7] Y. Wu, L. Ma, X. Li, J. Yang, X. Rao, Y. Hu, J. Xi, L. Tao, J. Wang, L. Du, G. Chen, and S. Liu, "The role of artificial intelligence in drug screening, drug design, and clinical trials," Frontiers in Pharmacology, vol. 15, p. 1459954, Nov 2024.

[8] S. He, D. Zhao, Y. Ling, H. Cai, Y. Cai, J. Zhang, and L. Wang, "Machine learning enables accurate and rapid prediction of active molecules against breast cancer cells," Frontiers in Pharmacology, vol. 12, 2021.

[9] H. Nada, A. R. Gul, A. Elkamhawy, S. Kim, M. Kim, Y. Choi, T. J. Park, and K. Lee, "Machine Learning-Based Approach to Developing Potent EGFR Inhibitors for Breast Cancer-Design, Synthesis, and In Vitro Evaluation," ACS Omega, vol. 8, no. 35, pp. 31784–31800, 2023, doi: 10.1021/acsomega.3c02799.

[10] O. E. Oyeneyin, B. S. Obadawo, D. S. Metibemu, T. O. Owolabi, A. A. Olanrewaju, S. M. Orimoloye, N. Ipinloju, and O. Olusayo, "An explo ration of the antiproliferative potential of chalcones and dihydropyrazole derivatives in prostate cancer via androgen receptor: combined qsar, ma chine learning, and molecular docking techniques," Physical Chemistry Research, vol. 10, no. 2, pp. 211–223, 2022.

[11] M. M. Islam, M. R. Haque, H. Iqbal, M. M. Hasan, M. Hasan, and M. N. Kabir, "Breast cancer prediction: a comparative study using machine learning techniques," SN Computer Science, vol. 1, no. 5, p. 290, 2020.

[12] Alkady, Walaa, Walaa Gad, and Khaled Bahnasy. "Swarm intelligence optimization for feature selection of biomolecules." 2019 14th International Conference on Computer Engineering and Systems (ICCES). IEEE, 2019.

[13] Ewees, Ahmed A., et al. "Improved Slime Mould Algorithm based on Firefly Algorithm for feature selection: A case study on QSAR model." Engineering with Computers 38.Suppl 3 (2022): 2407-2421.

[14] N. Zainal, A. M. Zain, N. H. M. Radzi, and A. Udin, "Glowworm swarm optimization (gso) algorithm for optimization problems: A state-of-the art review," Applied Mechanics and Materials, vol. 421, pp. 507–511, 2013.

[15] M. Torabi, I. Haririan, A. Foroumadi, H. Ghanbari, and F. Ghasemi, "A deep learning model based on the BERT pre-trained model to predict the antiproliferative activity of anti-cancer chemical compounds," SAR and QSAR in Environmental Research, vol. 35, no. 11, pp. 971–992, 2024.

[16] K. Krishnanand and D. Ghose, "Glowworm swarm optimization for si multaneous capture of multiple local optima of multimodal functions," Swarm intelligence, vol. 3, pp. 87–124, 2009.

[17] S. Halladay and G. Dozier, "A comparison of genetic & swarm intelligence-based feature selection algorithms for author identification," in 2020 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2020, pp. 1731–1738.

[18] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273–297, 1995.

[19] I. W. A. S. Putra and I. Kurniawan, "Implementation of gravitational search algorithm-support vector machine in prediction bioactivity of plk1 inhibitor as anti-cancer agent," in 2024 International Conference on Ar tificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD), 2024, pp. 297–302.

[20] A. M. Afinda, A. Aditsania, dan I. Kurniawan, "Implementasi metode Genetic Algorithm-Support Vector Machine pada studi QSAR in-house molecules sebagai inhibitor Papain-like Protease (PLpro) SARS-CoV-2," Jurnal Tugas Akhir Fakultas Informatika, vol. 8, no. 5, Sep. 2021.

[21] R. M. P. Bamba, I. Kurniawan, and W. Astuti, "In silico-based toxicity predicition using camel algorithm-support vector machine: Case study nr-ahr toxicity type," in 2024 International Conference on Data Science and Its Applications (ICoDSA), 2024, pp. 317–321.

[22] A. K. Linanda and I. Kurniawan, "Implementation of cuckoo search support vector machine method in detecting ulcerative colitis based on gene expression data," in 2024 International Conference on Artificial Inte lligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD), 2024, pp. 249–254.