biomimetics

MDPI

Article

MSBKA: A Multi-Strategy Improved Black-Winged Kite Algorithm for Feature Selection of Natural Disaster Tweets Classification

Guangyu Mu [1,2,†], Jiaxue Li [1,†], Zhanhui Liu [3,*], Jiaxiu Dai [1], Jiayi Qu [1] and Xiurong Li [4]

[1] School of Management Science and Information Engineering, Jilin University of Finance and Economics, Changchun 130117, China; guangyumu@jlufe.edu.cn (G.M.); 6221192039@s.jlufe.edu.cn (J.L.)
[2] Key Laboratory of Financial Technology of Jilin Province, Changchun 130117, China
[3] Changchun Community Official Staff College of Jilin Province, Changchun 130052, China
[4] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
* Correspondence: zhanhuiliu@163.com
† These authors contributed equally to this work.

Abstract: With the advancement of the Internet, social media platforms have gradually become powerful in spreading crisis-related content. Identifying informative tweets associated with natural disasters is beneficial for the rescue operation. When faced with massive text data, choosing the pivotal features, reducing the calculation expense, and increasing the model classification performance is a significant challenge. Therefore, this study proposes a multi-strategy improved black-winged kite algorithm (MSBKA) for feature selection of natural disaster tweets classification based on the wrapper method's principle. Firstly, BKA is improved by utilizing the enhanced Circle mapping, integrating the hierarchical reverse learning, and introducing the Nelder–Mead method. Then, MSBKA is combined with the excellent classifier SVM (RBF kernel function) to construct a hybrid model. Finally, the MSBKA-SVM model performs feature selection and tweet classification tasks. The empirical analysis of the data from four natural disasters shows that the proposed model has achieved an accuracy of 0.8822. Compared with GA, PSO, SSA, and BKA, the accuracy is increased by 4.34%, 2.13%, 2.94%, and 6.35%, respectively. This research proves that the MSBKA-SVM model can play a supporting role in reducing disaster risk.

Keywords: black-winged kite algorithm; machine learning; social media platform; feature selection; natural disaster tweets; emergency response

check for updates

Academic Editors: Heming Jia and Gang Hu

Received: 4 December 2024
Revised: 3 January 2025
Accepted: 8 January 2025
Published: 10 January 2025

Citation: Mu, G.; Li, J.; Liu, Z.; Dai, J.; Qu, J.; Li, X. MSBKA: A Multi-Strategy Improved Black-Winged Kite Algorithm for Feature Selection of Natural Disaster Tweets Classification. *Biomimetics* **2025**, *10*, 41. https://doi.org/10.3390/biomimetics10010041

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Natural disasters usually refer to events caused by abnormal changes in meteorology, geology, ocean, and ecology [1]. During the calamities, countless users publish posts on social media like Weibo and Twitter to communicate their status or seek assistance [2,3]. So, these platforms have become crucial data sources in emergency management [4]. Because of the complexity of natural disasters, decision-makers need to obtain vital information in the shortest possible time [5]. As shown in Table 1, informative tweets are closely related to disaster events and contribute to humanitarian relief [6], including infrastructure damage, casualties, and urgent demands. Unlike solely distinguishing whether tweets are relevant to the topic [7,8], the research on informative tweet recognition is more valuable for emergency response and situational awareness. In addition, technological progress makes it no longer difficult for researchers to obtain large-scale, real-time data from social media platforms [9]. Nevertheless, the speed of tweet generation is fast, and there is a severe

problem of information overload [10]. Automatic and accurate classification of natural disaster tweets can assist the government and relevant departments in making efficient arrangements [11,12]. Therefore, this paper has specific practical significance.

Table 1. The instances of tweets with two labels.

Tweets	Labels
We visited Puerto Rico today!	Not Informative
Ten nurses and one doctor from @UR_Med are leaving this AM to help in Puerto Rico. Hear from them on @13WHAM.	Informative

As an elementary Natural Language Processing (NLP) task, text classification assigns data to predefined categories according to extracted features [13]. Identifying "Informative" or "Not Informative" tweets is a binary classification issue. Automatic text categorization must utilize classifiers [14], such as Naive Bayes (NB), Random Forest (RF), and Support Vector Machine (SVM). NB is conceived from the Bayes theorem and the independent hypothesis of characteristic conditions [15]. The category with the most significant probability is selected as the prediction result by calculating the conditional probability of each category under the given text features [16]. The NB model has a simple structure and high running efficiency. However, the hypothesis of feature independence is often not valid in reality, which will affect the classification outcome. RF is an integrated learning method based on the decision tree [17]. Each decision tree randomly opts for some data and features from the original data set for training. The final upshot is obtained by voting or averaging [18]. Due to the randomness, RF can decrease the risk of over-fitting and enhance the model's generalization capability. Nevertheless, when the number of decision trees is large, the prediction process is slow, and more computational resources are needed. SVM follows the structural risk minimization principle and seeks the optimal hyperplane by maximizing the interval [19]. SVM is particularly able to solve the practical binary classification problem [20], but the kernel function must be used properly.

One of the prime challenges of text classification is that numerous irrelevant or redundant features impact the classifiers' efficiency and accuracy [21]. Feature selection technology is specially adopted to deal with these difficulties. Choosing features with plentiful information and high correlation can enhance the model classification effect and reduce the computation cost [22,23]. In order to acquire ideal results, many scholars consider applying metaheuristic algorithms to feature selection, such as Genetic Algorithm (GA) [24,25], Particle Swarm Optimization (PSO) [26,27], and Sparrow Search Algorithm (SSA) [28,29]. The metaheuristic algorithm is exceedingly suitable for complicated optimization tasks. The principle is to discover the global optimal solution in the search space. BKA, proposed in 2024, simulates the behavior of black-winged kites in daily life and their formidable adaptability to environmental transformation [30]. With the merits of a straightforward framework, few manual parameters, and speedy convergence, BKA has been widely utilized in engineering [31–33]. Compared with other previous algorithms, BKA shows better competence. Nevertheless, no method is always reliable under the "No Free Lunch" theory [34]. The algorithm's performance is a direct factor affecting feature selection. BKA's global exploration and local exploitation are not balanced enough. The algorithm's convergence accuracy is insufficient, and it is easy to fall into the local optimum. Therefore, making some corresponding improvements to BKA is necessary.

Here, we can obtain the following two motivations for this study:

(1) An applicable classifier must be chosen for the binary classification of natural disaster tweets.

(2) The improved metaheuristic algorithm can be used for feature selection to further increase the classification model's accuracy.

This research's primary contributions are as follows:

- We adopt three tactics to improve BKA. Firstly, we utilize the enhanced Circle mapping for population initialization. Secondly, we integrate the hierarchical reverse learning into BKA's attack behavior. Thirdly, we introduce the Nelder–Mead method to obtain a better solution.
- We contrast the performance of several classification approaches and choose the SVM (Radial Basis Function, RBF kernel function) with the best result as the classifier. Then, we construct an MSBKA-SVM model based on the wrapper method's principle.
- We conduct contradistinctive experiments with other metaheuristic algorithms on the tweet data set of four natural disasters. Empirical analysis demonstrates the proposed model's superiority in feature selection and binary classification tasks.

2. Literature Review

2.1. Feature Selection

Feature selection is a data dimensionality reduction technology. According to the diverse interaction with the model training process, feature selection is split into the filter, embedding, and wrapper methods.

The filter method measures the features' significance on the basis of statistical analysis, such as the Pearson Correlation Coefficient [35,36], Chi-square test [37,38], and Mutual Information [39,40]. This approach is carried out before model training and is independent of any specific algorithm [41]. Its main advantage is its ease of calculation, making it suited for large-scale data sets [42]. However, the filter method usually does not consider the features' interaction and may ignore some essential features.

The embedding method allows the algorithm to decide which features to utilize. In other words, feature selection and model training are executed in simultaneity [43]. The embedding method is generally divided into two types: originating from penalty terms or tree models. The manners stemming from penalty terms include L1 Regularization (Lasso Regression) [44,45], L2 Regularization (Ridge Regression) [46,47], and Elastic-net Regression [48,49]. The approaches derived from the tree models contain Decision Tree [50,51], eXtreme Gradient Boosting (XGBoost) [52,53], and Light Gradient Boosting Machine (LightGBM) [54,55]. The embedding method is efficient, but it is difficult to determine a valid critical value to judge which features are momentous.

The wrapper method evaluates features based on the classifier's performance [56], like accuracy or error. This approach regards the feature selection process as a search problem. Common means are Forward Selection [57,58], Backward Elimination [59,60], and Recursive Feature Elimination [61,62]. The wrapper method can consider the interaction and dependence between features [63]. It must try various feature subsets to train the model, so it is relatively accurate but more time-consuming.

In summary, the filter method ordinarily does not need complicated search strategies because of its independence from algorithms. The embedding method combines feature selection with the model training process, which commonly depends on the models' internal mechanism. Metaheuristic algorithms are especially suitable for the wrapper method, given their high efficiency in dealing with complex optimization problems [64–67]. There are three prime reasons. Firstly, the wrapper method uses the model performance as the evaluation criterion of feature subsets. Metaheuristic algorithms can dynamically adjust tactics according to the feedback of model performance when searching for the optimal solution. The feature subsets with the most remarkable performance improvement are more likely to be found. Secondly, the wrapper approach contemplates the interaction between

features and selects the optimal feature subsets by iteratively adding or removing features. Metaheuristic algorithms are able to handle this combinatorial optimization issue well. Thirdly, the wrapper method's search space and calculation cost are comparatively large, so efficient algorithms are needed to reduce running expenses. Therefore, this research applies the metaheuristic algorithm to the feature selection on the basis of the wrapper method's principle.

2.2. Classification of Natural Disaster Tweets

Social media platforms cover a variety of information [68,69], so scholars classify natural disaster tweets in diverse ways. The related research principally analyzes the content and sentiment expressed by tweets.

During calamities, data transmission and reception are normally mutual [70]. On the one hand, the affected people desire information about shelter, medical care, and food [71]. On the other hand, humanitarian organizations need to inform the current arrangements and provide the corresponding locations [72]. Identifying these specific contents in tweets promotes the rapid deployment of emergency response to natural disasters [73,74]. Related tweets are generally annotated as binary labels, such as urgent and not urgent [75] or high priority and low priority [76]. In addition, tweet data can be divided into fine-grained categories [77–79], including infrastructure damage, the trapped people's situation, or the demand for relief materials.

Natural disasters' unpredictability and severe destructiveness can easily induce negative emotions in netizens, leading to a public opinion crisis [80,81]. Identifying the sentiment tendencies assists relevant departments in evaluating the public's views and attitudes toward emergencies [82,83]. Unlike lengthy articles or reports, tweets are limited by characters. This concise expression is conducive to understanding the textual emotions [84]. Sentiment polarities are classified into positive and negative [85] or positive, neutral, and negative [86]. Negative emotions ordinarily account for a more significant proportion in a particular period. Consequently, some sample synthesis techniques are applied to balance the data sets and ensure the model classification results' accuracy [87]. Additionally, scholars usually concentrate on time and space to show the variational trend of emotional characteristics in diverse stages and regions [88–90].

Actually, only partial tweets are valuable for emergency management and situational awareness [91–93]. Tweets must first be filtered out of irrelevant text [94–96]. Then, the filtered data can be used to dig for beneficial content for disaster analysis [97]. However, being related to the incidents is not equivalent to contributing to humanitarian relief. This research is more specific and practical than merely distinguishing topic relevance.

3. Method

This study proposes an MSBKA for feature selection of natural disaster tweets classification. Figure 1 is the MSBKA-SVM model's architecture, which comprises four modules: data preprocessing, feature extraction, feature selection, and model classification.

3.1. Data Preprocessing

Data preprocessing is the key to data analysis. The purpose is to improve data quality, uniform format, and remove noise. The specific steps are as follows:

- Remove URL links and special characters.
- Remove user names and topic tags.
- Remove punctuation and numbers.
- Convert all text contents to lowercase.
- Divide tweets into word lists.

- Filter out words less than three in length.
- Reconnect the remaining words into a string.

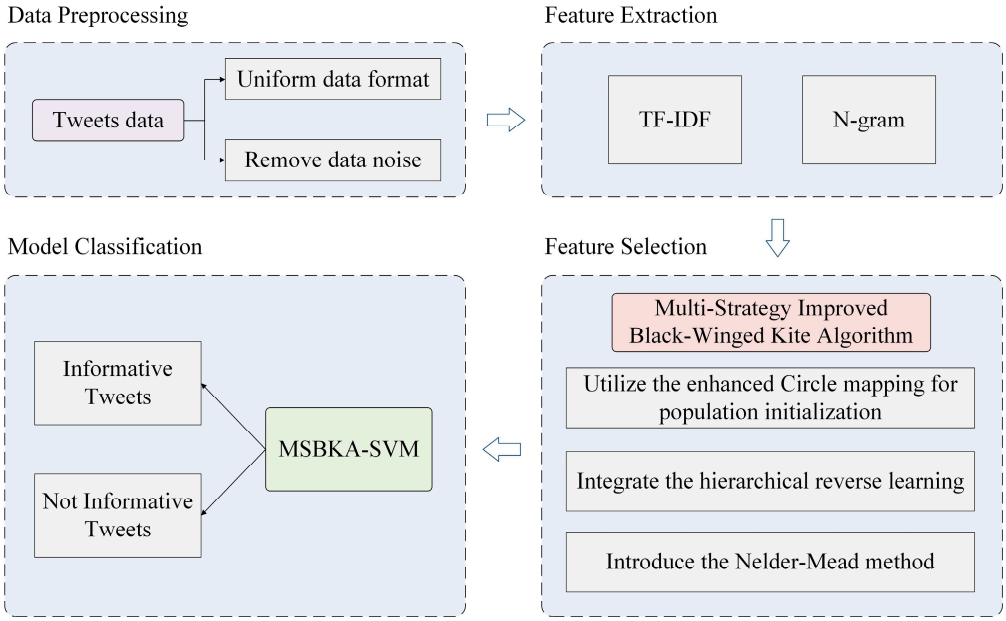


Figure 1. The architecture of the MSBKA-SVM model.

3.2. Feature Extraction

Unstructured text data cannot be immediately utilized for model training. This study adopts Term Frequency–Inverse Document Frequency (TF-IDF) combined with N-gram to convert tweets into numerical forms that classifiers can process. TF-IDF is a technique to measure the importance of words in document collections [98]. The TF-IDF value is directly proportional to the frequency of words emerging in a tweet but inversely proportional to the frequency of words appearing in the overall data set. The calculation process is shown in Equation (3).

$$TF_{i,j} = \frac{s_{i,j}}{\sum_k s_{k,j}} \tag{1}$$

$$IDF_i = log\frac{S}{s_i} \tag{2}$$

$$TF - IDF = TF_{i,j} * IDF_i \tag{3}$$

where $s_{i,j}$ is the quantity of times a certain word emerges in the tweet j. $\sum_k s_{k,j}$ is the total quantity of words in the tweet j. S is the total quantity of tweets in the data set. s_i is the quantity of tweets comprising a certain word.

N-gram selects n adjacent words from the text as features [99]. This study allows the model to consider single words (1-gram) and word pairs (2-gram) to capture more semantic information and contextual relationships. For the sake of improving the training efficiency and prediction performance, a scaling operation is performed to make the features all have an identical magnitude. Moreover, according to the ranking of TF-IDF values, up to 10,000 features are restricted for subsequent feature selection.

3.3. Black-Winged Kite Algorithm

A black-winged kite is a miniature bird with formidable hovering ability. The attack and migration behaviors inspire BKA. The algorithm's population form is displayed in

Matrix (4). The individual's position is expressed as Equation (5). The parameters are described in Table 2.

$$BK = \begin{bmatrix} BK_{1,1} & BK_{1,2} & \cdots & \cdots & BK_{1,dim} \\ BK_{2,1} & BK_{2,2} & \cdots & \cdots & BK_{2,dim} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ BK_{pop,1} & BK_{pop,2} & \cdots & \cdots & BK_{pop,dim} \end{bmatrix} \tag{4}$$

$$BK_{i,\,j} = BK_{lb} + rand(BK_{ub} - BK_{lb}) \tag{5}$$

Table 2. The elaborations of the parameters in Equations (4) and (5).

Parameters	Elaborations
pop	The population size
dim	The optimization problem's dimension
lb	The lower boundary
ub	The upper boundary
$rand$	A random value within the interval [0, 1]

3.3.1. The Attack Behavior

Black-winged kites feed on mammals and insects. During the flight, the bird will adjust the angle of its wings and tail in accordance with the wind speed, calmly hover to survey the prey, and then rapidly dive to onslaught. Two kinds of attack behaviors of BKA are indicated by Equation (6). The parameters are represented in Table 3.

$$y_{t+1}^{i,\,j} = \begin{cases} y_t^{i,\,j} + n(1 + sin(r)) \times y_t^{i,\,j} & p < r \\ y_t^{i,\,j} + n(2r - 1) \times y_t^{i,\,j} & else \end{cases} \tag{6}$$

$$n = 0.05 \times e^{-2\times(\frac{t}{T})^2} \tag{7}$$

Table 3. The elaborations of the parameters in Equations (6) and (7).

Parameters	Elaborations
$y_t^{i,\,j}$, $y_{t+1}^{i,\,j}$	The ith black-winged kite's position information in the jth dimension at the tth and $t+1$th iterations
r	A random digit from 0 to 1
p	A constant of 0.9
n	A dynamically adjusted variable
t	The current iteration quantity
T	The maximum quantity of iterations

3.3.2. The Migration Behavior

In order to adapt to climate and environmental changes, black-winged kites migrate from north to south in winter. Leaders commonly lead bird migration. Their navigation capabilities are pivotal to the team's achievement. BKA hypothesis: If the current individual's fitness value F_i is less than the random individual's fitness value F_{ri} in the population, the leader will abandon the leadership and enter the migration team. Instead, the leader

will lead the population to their destinations. This dynamic selection strategy ensures the migration's success, as Equation (8) expresses. The parameters are indicated in Table 4.

$$y_{t+1}^{i,j} = \begin{cases} y_t^{i,j} + C(0,1) \times \left(y_t^{i,j} - L_t^j \right) & F_i < F_{ri} \\ y_t^{i,j} + C(0,1) \times \left(L_t^j - m \times y_t^{i,j} \right) & else \end{cases} \tag{8}$$

$$m = 2 \times sin(r + \pi/2) \tag{9}$$

Table 4. The elaborations of the parameters in Equations (8) and (9).

Parameters	Elaborations
L_t^j	The leader's location in the jth dimension at the tth iteration
$C(0, 1)$	The Cauchy distribution function
m	A dynamically adjusted variable

The Cauchy distribution is a continuous probability distribution. The probability density function is shown in Equation (10).

$$f(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}, -\infty < x < \infty \tag{10}$$

When the scale parameter $\gamma = 1$ and the location parameter $x_0 = 0$, the probability density function of the standard Cauchy distribution is indicated in Equation (11).

$$f(x) = \frac{1}{\pi(x^2 + 1)}, -\infty < x < \infty \tag{11}$$

3.4. Multi-Strategy Improved Black-Winged Kite Algorithm

Combining the Cauchy mutation and leader tactics enhances the original BKA's search capability and convergence speed. Nevertheless, there is still a disequilibrium between global exploration and local exploitation. The algorithm has deficient optimization accuracy and easily falls into local optimization. Accordingly, MSBKA adopts three improved strategies.

3.4.1. Utilize the Enhanced Circle Mapping for Population Initialization

The incipient population's diversity and quality are the pivotal factors affecting the metaheuristic algorithm's performance [100]. The original BKA adopts a random way for population initialization, leading to the aggregation of generated individuals and low otherness. Chaotic mapping is ergodic and can explore all possible states in space [101]. The prime idea is to generate a chaotic sequence by mapping in the range [0, 1] and then convert the sequence into the population's search space [102,103]. Among several types of chaotic mappings, Circle mapping is relatively steady and has higher coverage of chaotic values [104]. In order to avoid the chaotic values between [0.2, 0.4] densely, the traditional Equation (12) is improved to Equation (13) in this study. The initial solutions' frequency distribution histogram is shown in Figure 2. The solutions' dimension i is taken as 3000 to display the contrast effect intuitively. It is evident that the enhanced Circle mapping has more well-proportioned chaotic values.

$$x_{i+1} = mod\left(x_i + 0.2 - \left(\frac{0.5}{2\pi}\right) sin(2\pi x_i), 1 \right) \tag{12}$$

$$x_{i+1} = mod\left(3.85 \times x_i + 0.4 - \left(\frac{0.4}{3.85\pi}\right) sin(3.85\pi x_i), 1\right) \quad (13)$$

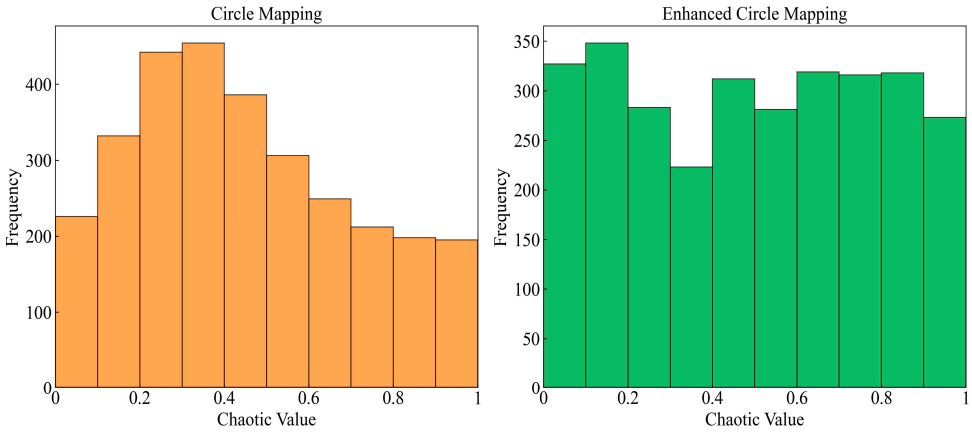


Figure 2. Frequency distribution histogram of initial solutions.

3.4.2. Integrate the Hierarchical Reverse Learning

Individuals tend to gradually gather in some local areas in the iterative process, which makes the algorithm fall into local optimization [105,106]. In order to avoid premature convergence, MSBKA combines the original BKA's attack behavior with the hierarchical reverse learning. This strategy needs to calculate each individual's fitness value in the population and rank it as Rank 1, Rank 2, . . ., and Rank N from tiny to enormous. Then, N individuals are divided into NL levels, each with LS individuals. L_1 represents the individual's best grade. L_{NL} means the worst grade. The hierarchical schematic is indicated in Figure 3.

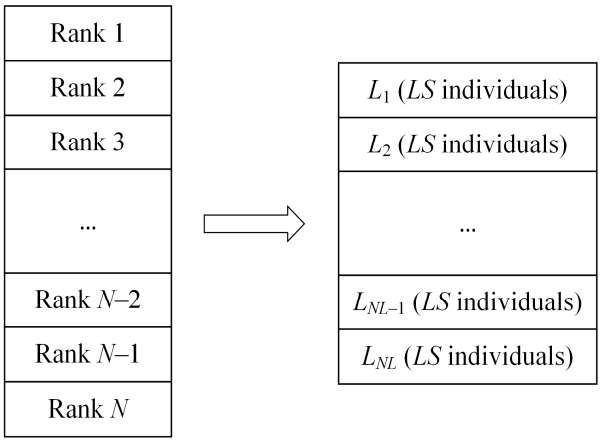


Figure 3. The schematic diagram of hierarchical reverse learning.

The reverse individual's quality is not necessarily better than that of the current individual. There are still some individuals whose quality has declined after reverse learning. Accordingly, in this study, the obtained reverse individual $X_{opposite}$ and the previous level's optimal individual X_j^{best} are convexly combined to obtain the reverse individual tending towards excellence. This process can be expressed as Equation (14).

$$X_j' = \lambda \times X_{opposite} + (1 - \lambda) \times X_j^{best} \quad (14)$$

3.4.3. Introduce the Nelder–Mead Method

When faced with a complicated high-dimensional search space, the original BKA may have the issue of slow convergence or stagnation in some areas. Therefore, the Nelder–Mead method is introduced in this research. The polyhedron formed by connecting $D+1$ vertexes in D-dimensional space is called a simplex. The Nelder–Mead method is a simple and speedy local search algorithm [107]. The core idea is to obtain better points by performing several operations, replace the worst points, and approximate the best points by forming a novel polyhedron. MSBKA sets the Nelder–Mead method's trigger condition as no improved fitness value every ten consecutive iterations. The Nelder–Mead method's search points are shown in Figure 4.

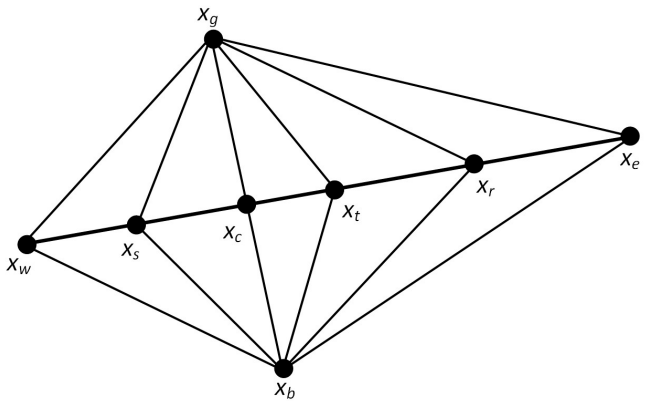


Figure 4. Search points of the Nelder–Mead method.

The detailed operation steps are as follows:

(1) Calculate the fitness values of all vertexes and seek out the best point x_g, the second-best point x_b, and the worst point x_w. The fitness values are $f(x_g)$, $f(x_b)$, and $f(x_w)$. Calculate the center point's position by Equation (15).

$$x_c = \frac{x_g + x_b}{2} \tag{15}$$

(2) The reflection operation is performed by Equation (16). Reflect the worst point x_w according to the center point x_c. α is the reflection coefficient. The reflection point's fitness value is $f(x_r)$.

$$x_r = x_c + \alpha(x_c - x_w) \tag{16}$$

(3) When $f(x_r) > f(x_g)$, the reflection direction is correct. Then, the extension operation is executed by Equation (17). μ is the extension coefficient. The extension point's fitness value is $f(x_e)$. When $f(x_e) > f(x_g)$, displace the worst point x_w with the extension point x_e. Otherwise, displace the worst point x_w with the reflection point x_r.

$$x_e = x_c + \mu(x_r - x_c) \tag{17}$$

(4) When $f(x_r) < f(x_g)$, the reflection direction is false. Then, the compression operation is performed by Equation (18). β is the compression coefficient. The compression point's fitness value is $f(x_t)$. When $f(x_t) > f(x_w)$, replace the worst point x_w with the compression point x_t.

$$x_t = x_c + \beta(x_w - x_c) \tag{18}$$

(5) When $f(x_w) < f(x_r) < f(x_g)$, the shrinkage operation is executed by Equation (19). δ is the shrinkage coefficient. The shrinkage point's fitness value is $f(x_s)$. When

$f(x_s) > f(x_w)$, displace the worst point x_w with the shrinkage point x_s. Otherwise, displace the worst point x_w with the reflection point x_r.

$$x_s = x_c - \delta(x_w - x_c) \tag{19}$$

3.4.4. The Running Steps of MSBKA

The MSBKA's running steps are as follows:

Step 1: Set the related parameters of MSBKA.

Step 2: Initialize the population using the enhanced Circle mapping by Equation (13).

Step 3: Evaluate the fitness value of each individual. The individual with the best fitness value is determined as the leader in the initial population.

Step 4: If a random number r within [0, 1) is less than 0.4, the attack behavior is executed by Equation (6). Otherwise, the migration behavior is performed by Equation (8). In the attack behavior, the hierarchical reverse learning strategy is carried out by Equation (14) with a probability of 0.5.

Step 5: Update the current optimal solution and fitness value. Determine whether to perform the Nelder–Mead method utilizing Equations (15)–(19).

Step 6: Repeat Steps 4 to 5. Output the global optimal solution and fitness value when reaching the maximum quantity of iterations.

3.4.5. The Time Complexity of MSBKA

Time complexity is utilized to evaluate the algorithm's efficiency, which reflects the relationship between the time required for algorithm execution and input data size. Big O notation is a standard method [108]. This approach describes the algorithm's performance in the worst condition. Set the maximum iterations T, the population scale N, and the problem dimension D. The BKA's time complexity is represented as $O(N \times D \times T)$. The influence of three improved strategies adopted by MSBKA on the time complexity is limited. Chaotic initialization does not directly participate in the subsequent iterative process. The extra calculation involved in reverse learning is linear for the individuals. The Nelder–Mead method has a trigger condition. Therefore, MSBKA has the same order of magnitude in time complexity as the original BKA.

3.5. Feature Selection

The essence of feature selection is to let individuals search in the space composed of the data set's feature dimensions. The metaheuristic algorithm's individual represents the potential solution. Each iteration's update of individual positions equals the feature subsets' renewal. The individual with the highest fitness value is the global optimal solution. The corresponding feature subset is considered the optimal feature subset. If the data set has n features, then each individual is a vector of length n. In the initial population generated by MSBKA, each position in the vector is given a random value within [0, 1]. When this value exceeds 0.5, it is set to 1, indicating that the feature is selected. If this value is less than or equal to 0.5, it is set to 0, signifying that the feature is not selected. As shown in Figure 5, assuming that the data set contains five features, only the second and fourth features in the current individual are selected.

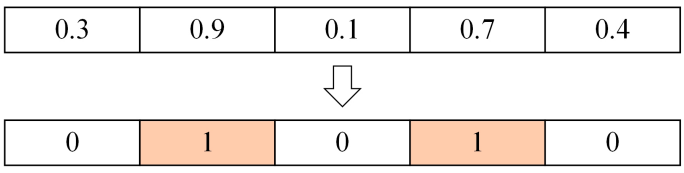


Figure 5. Schematic diagram of the feature selection process.

The feature selection is aimed at seeking out the feature subset with the highest accuracy and the least quantity of features. Consequently, the fitness function is defined as Equation (20).

$$Fitness = \omega \times Accuracy - (1 - \omega) \times \frac{R}{C} \tag{20}$$

R is the number of selected features. C is the total quantity of features in the data set.

3.6. Model Classification

This study utilizes several models to implement the binary classification of tweets, comprising Naive Bayes, Random Forest, and Support Vector Machine. Among them, SVM selects Linear and RBF kernel functions. These two expressions are expressed as Equations (21) and (22).

$$K(x,\ y) = (x \cdot y) \tag{21}$$

$$K(x,\ y) = e^{(\frac{-\|x-y\|^2}{2\sigma^2})} \tag{22}$$

4. Empirical Analysis

4.1. Data Acquisition

CrisisMMD is a Twitter data set about natural disasters [109,110]. Each tweet is marked as "Informative" or "Not Informative", with labels 1 and 0, respectively. This study selects four types of natural disasters: Sri Lanka Floods, Mexico Earthquake, Hurricane Maria, and California Wildfires. Finally, 4370 tweets are retained after filtering and duplicate items processing. The number of labels in the two categories is balanced at 2185. The descriptions of four natural disaster events are indicated in Table 5.

Table 5. Descriptions of four natural disaster events.

Natural Disaster Events	Start Date	End Date	Tweet Counts
Sri Lanka Floods	31 May 2017	3 July 2017	460
Mexico Earthquake	20 September 2017	6 October 2017	580
Hurricane Maria	20 September 2017	13 November 2017	2770
California Wildfires	10 October 2017	27 October 2017	560

4.2. Experimental Details

The training set and the test set are randomly selected in a ratio of eight to two. The other parameter settings are indicated in Table 6.

Table 6. The parameter settings of the whole experiment.

Parameters	Value
T	50
N	50
Chaotic iteration	150
λ	0.5
α	1
μ	2
β	0.5
δ	0.5
ω	0.5

4.3. Evaluation Metrics

For the purpose of comprehensively contrasting diverse classification models, this research adopts four evaluation metrics: *accuracy*, *precision*, *recall*, and *F*1. The normative

binary classification confusion matrix is indicated in Table 7. The higher the values of all evaluation metrics, the better the performance of the classifiers. The calculation processes are expressed by Equations (23)–(26).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{23}$$

$$Precision = \frac{TP}{TP + FP} \tag{24}$$

$$Recall = \frac{TP}{TP + FN} \tag{25}$$

$$F1 = \frac{2 * Recall * Precision}{Precision + Recall} \tag{26}$$

Table 7. The normative binary classification confusion matrix.

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

4.4. Experimental Results

4.4.1. The Selection of Classifiers

Before combining with the metaheuristic algorithms, this research contrasts Naive Bayes, Random Forest, and SVM with two kernel functions. As shown in Table 8, SVM with RBF kernel function has the best performance. The accuracy reaches 0.7426. Compared with Naive Bayes, Random Forest, and SVM with a linear kernel function, the accuracy is improved by 4.68%, 2.05%, and 3.18%, respectively. Consequently, this study selects SVM with RBF kernel function as the feature selection task's classifier.

Table 8. The evaluation metrics contrast of several classifiers.

Classifiers	***Accuracy***	**Labels**	***Precision***	***Recall***	***F1***
Naive Bayes	0.7094	0	0.7798	0.5835	0.6675
		1	0.6673	0.8352	0.7419
Random Forest	0.7277	0	0.6970	0.8055	0.7473
		1	0.7696	0.6499	0.7047
SVM (Linear)	0.7197	0	0.7388	0.6796	0.7080
		1	0.7034	0.7597	0.7305
SVM (RBF)	0.7426	0	0.7409	0.7460	0.7434
		1	0.7442	0.7391	0.7417

4.4.2. The Evaluation Metrics Comparison of Classification Models

This study compares other metaheuristic algorithms to attest to the proposed model's advantage in tweet classification, including GA, PSO, SSA, and BKA. The whole experiment is conducted in identical conditions and parameter settings to guarantee the credibility of the contrast results. The evaluation metrics are indicated in Table 9. The separate evaluation of two data labels can analyze the model performance in more detail.

Table 9. The evaluation metrics comparison of five classification models.

Models	*Accuracy*	Labels	*Precision*	*Recall*	*F1*
GA-SVM	0.8455	0	0.8268	0.8741	0.8498
		1	0.8665	0.8169	0.8410
PSO-SVM	0.8638	0	0.8597	0.8696	0.8646
		1	0.8681	0.8581	0.8631
SSA-SVM	0.8570	0	0.8467	0.8719	0.8591
		1	0.8679	0.8421	0.8548
BKA-SVM	0.8295	0	0.7869	0.9039	0.8413
		1	0.8871	0.7551	0.8158
MSBKA-SVM (Proposed)	0.8822	0	0.8568	0.9176	0.8862
		1	0.9113	0.8467	0.8778

The Receiver Operating Characteristic (ROC) curve and the Area Under the ROC Curve (AUC) are vital indexes for measuring the binary classification model's performance. ROC is realized by plotting the relationship between the True Positive Rate (*TPR*) and False Positive Rate (*FPR*). The range of AUC is from 0.5 (random guess) to 1 (perfect classifier), which quantifies the model's capability to distinguish diverse categories. Figure 6 shows the ROC-AUC diagram of five classification models.

$$TPR = \frac{TP}{TP + FN} \tag{27}$$

$$FPR = \frac{FP}{FP + TN} \tag{28}$$

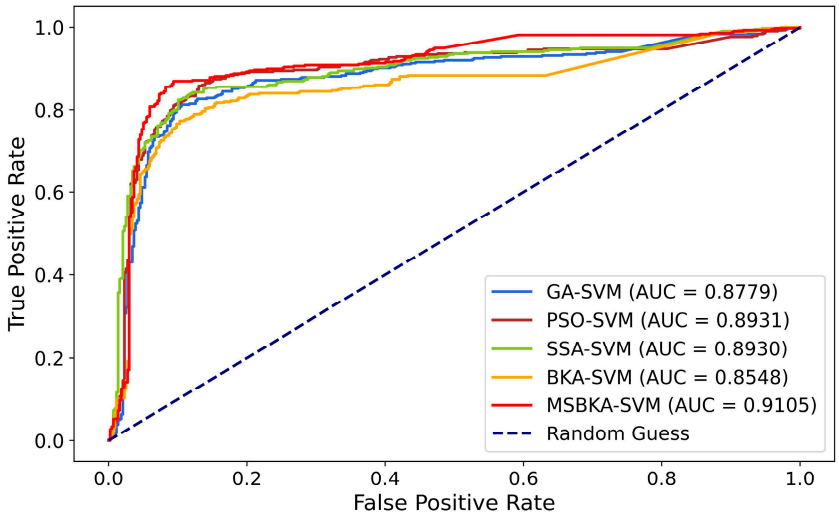


Figure 6. The ROC-AUC diagram of five classification models.

The evaluation metrics of five classification models reveal the following findings:

- The classifier's accuracy is strikingly increased after feature selection combined with the metaheuristic algorithms. Compared with individual SVM (RBF kernel function), the accuracy of GA-SVM, PSO-SVM, SSA-SVM, BKA-SVM, and MSBKA-SVM is improved by 13.86%, 16.32%, 15.41%, 11.70%, and 18.80%, respectively.
- By utilizing the enhanced Circle mapping, integrating the hierarchical reverse learning, and introducing the Nelder–Mead method, MSBKA's performance is better than that of the original BKA. Furthermore, MSBKA shows the most remarkable enhancement among all the models. Its accuracy achieves 0.8822, heightened by 4.34%, 2.13%, 2.94%,

and 6.35% compared to GA, PSO, SSA, and BKA. The main reason is that these original algorithms have some defects, which lead to poor stability.

- MSBKA-SVM has the highest AUC value of 0.9105, which is 3.71%, 1.95%, 1.96%, and 6.52% higher than GA-SVM, PSO-SVM, SSA-SVM, and BKA-SVM. The proposed model's AUC is the closest to 1. In other words, the capability to distinguish two categories is excellent.

4.4.3. The Performance Contrast of Classification Models

In addition to accuracy, the quantity of features selected by classification models and the running time also need to be considered. Furthermore, to show the optimization algorithm's performance changes in the iterative process, the fitness value is utilized as an evaluation criterion to plot the convergence curve. The comparison results of these aspects are indicated in Figures 7–9.

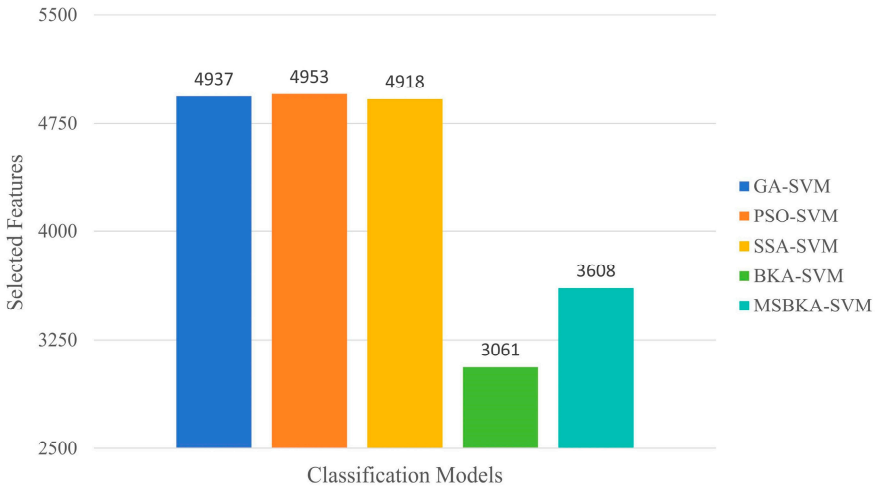


Figure 7. The number comparison of features selected by five classification models.

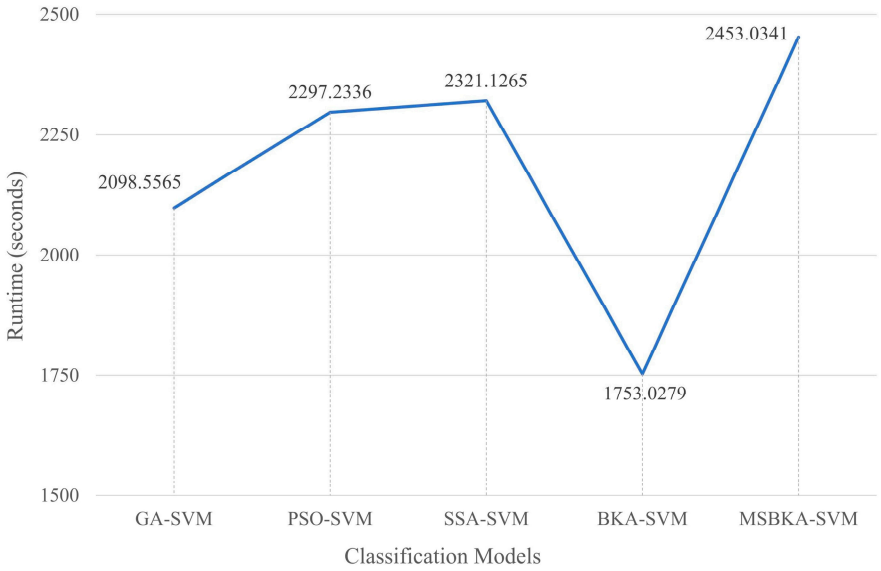


Figure 8. The runtime comparison of five classification models.

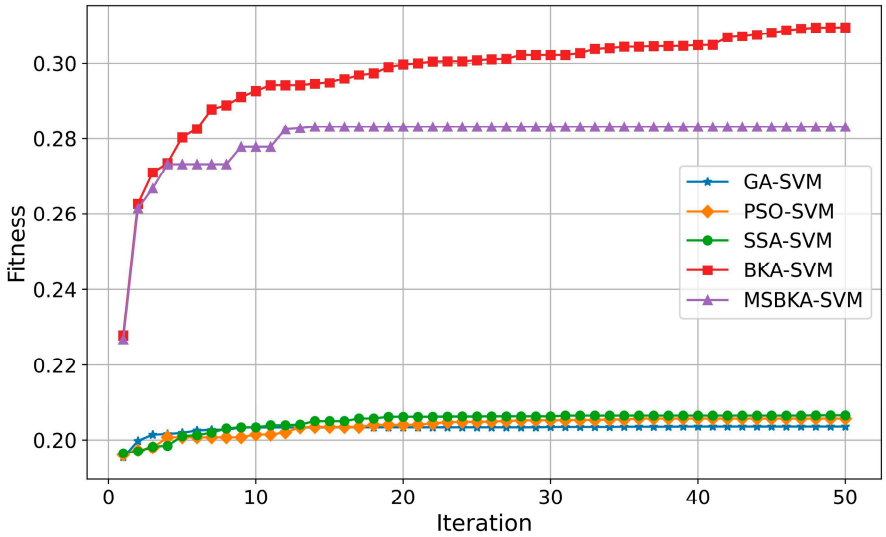


Figure 9. The convergence curves of five classification models.

As shown in Figure 7, in contrast with GA-SVM, PSO-SVM, SSA-SVM, and MSBKA-SVM, the number of selected features for BKA-SVM is minimal. Nevertheless, the accuracy of BKA-SVM is not high enough. This phenomenon is due to the features selected by BKA-SVM that are irrelevant or not helpful to model classification. The quantity of features selected by MSBKA-SVM is slightly more than that of BKA-SVM, but the accuracy is dramatically improved. That is to say, MSBKA-SVM can select the most relevant and critical features from text data. Similarly, Figure 8 shows that BKA-SVM has the minimum running time and MSBKA-SVM has the maximum execution time among all classification models. In Figure 9, the fitness function set in this study takes into account both the classification accuracy and the number of selected features. The fitness value of MSBKA-SVM is lower than that of BKA-SVM. However, this research focuses on practical applications, and accuracy is obviously the most pivotal criterion. In summary, combined with various factors, the proposed model is still optimal.

5. Conclusions and Prospect

5.1. Conclusions

Identifying informative tweets related to natural disasters avails disaster relief assignments. The irrelevant and redundant features in tweets will affect the classification accuracy. These issues can be resolved by using feature selection technology. A novel metaheuristic algorithm, BKA, is suitable for complicated optimization tasks. However, BKA suffers from inadequate search accuracy and easily falls into the local optimum. There is still room for capability advancement.

This study proposes a multi-strategy improved black-winged kite algorithm for feature selection of natural disaster tweets classification. Specifically, three optimization tactics are adopted to handle the BKA's deficiencies. Firstly, the enhanced Circle mapping is utilized for population initialization. Secondly, hierarchical reverse learning is integrated into BKA's attack behavior. Thirdly, the Nelder–Mead method is introduced to obtain a better solution. Among all classifiers, SVM (RBF kernel function) performs best. The hybrid MSBKA-SVM model based on the wrapper method's principle sufficiently combines the metaheuristic algorithm and machine learning.

Experimental results display that feature selection plays a positive role in model classification. Compared with individual SVM (RBF kernel function), the accuracy of GA-SVM, PSO-SVM, SSA-SVM, BKA-SVM, and MSBKA-SVM is improved by 13.86%,

16.32%, 15.41%, 11.70%, and 18.80%, respectively. Furthermore, MSBKA shows the most conspicuous enhancement among all the models. Its accuracy achieves 0.8822, increased by 4.34%, 2.13%, 2.94%, and 6.35% compared to GA, PSO, SSA, and BKA. Meanwhile, considering the quantity of selected features, the model runtime, and the fitness value, the proposed model is still the most ideal scheme.

On the whole, this research has prominent pragmatic implications. It is challenging to discriminate complex tweet content. The MSBKA-SVM's classification capability is more distinguished than that of other methods. Contrast experiments demonstrate the proposed model's advantage in dealing with natural disaster tweets. Accordingly, the government and relevant departments can immediately handle the identified vital information to reduce the damages caused by calamities and prevent a large-scale public opinion crisis.

5.2. Limitation and Future Direction

This research still has certain limitations. Due to the Internet's anonymity, there is much false or unverified information on social media platforms. Especially during emergencies, these forged contents can easily cause public panic and social unrest, even threatening national security. Identifying informative tweets about natural disasters is merely the first phase of emergency management. Organizing and utilizing these tweets effectively and detecting false information simultaneously is an issue that needs further contemplation.

Author Contributions: Conceptualization, G.M. and J.L.; methodology, J.L.; software, J.L.; validation, G.M. and J.L.; formal analysis, J.L.; investigation, J.D. and J.Q.; resources, Z.L. and X.L.; data curation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, G.M. and J.L.; visualization, J.L.; supervision, G.M.; project administration, G.M.; funding acquisition, G.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Social Science Fund of China under Grant No. 19BJY246, the Natural Science Fund Project of the Science and Technology Department of Jilin Province under Grant No. 20240101361JC, the Think Tank Fund Project of the Jilin Science and Technology Association, and the Pioneering Project of Jilin University of Finance and Economics under Grant No. 2024LH009.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The raw data and code supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors are grateful for the financial support from the National Social Science Fund of China under Grant No. 19BJY246, the Natural Science Fund Project of the Science and Technology Department of Jilin Province under Grant No. 20240101361JC, the Think Tank Fund Project of the Jilin Science and Technology Association, and the Pioneering Project of Jilin University of Finance and Economics under Grant No. 2024LH009.

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

References

1. Mu, G.; Li, J.; Li, X.; Chen, C.; Ju, X.; Dai, J. An Enhanced IDBO-CNN-BiLSTM Model for Sentiment Analysis of Natural Disaster Tweets. *Biomimetics* **2024**, *9*, 533. [CrossRef] [PubMed]
2. Ghafarian, S.H.; Yazdi, H.S. Identifying Crisis-Related Informative Tweets Using Learning on Distributions. *Inf. Process. Manag.* **2020**, *57*, 102145. [CrossRef]
3. Ngamassi, L.; Shahriari, H.; Ramakrishnan, T.; Rahman, S. Text Mining Hurricane Harvey Tweet Data: Lessons Learned and Policy Recommendations. *Int. J. Disaster Risk Reduct.* **2022**, *70*, 102753. [CrossRef]
4. Malik, M.S.I.; Younas, M.Z.; Jamjoom, M.M.; Ignatov, D.I. Categorization of Tweets for Damages: Infrastructure and Human Damage Assessment Using Fine-Tuned BERT Model. *PeerJ Comput. Sci.* **2024**, *10*, e1859. [CrossRef] [PubMed]

5. Moussa, A.M.; Abdou, S.; Elsayed, K.M.; Rashwan, M.; Asif, A.; Khatoon, S.; Alshamari, M.A. Enhanced Arabic Disaster Data Classification Using Domain Adaptation. *PLoS ONE* **2024**, *19*, e0301255. [CrossRef] [PubMed]
6. Alam, F.; Ofli, F.; Imran, M. CrisisMMD: Multimodal Twitter Datasets from Natural Disasters. In Proceedings of the International AAAI Conference on Web and Social Media, Palo Alto, CA, USA, 25–28 June 2018; Volume 12. [CrossRef]
7. Olteanu, A.; Castillo, C.; Diaz, F.; Vieweg, S. CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises. In Proceedings of the International AAAI Conference on Web and Social Media, Ann Arbor, MI, USA, 1–4 June 2014; Volume 8, pp. 376–385. [CrossRef]
8. Cobo, A.; Parra, D.; Navón, J. Identifying Relevant Messages in a Twitter-Based Citizen Channel for Natural Disaster Situations. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18 May 2015; ACM: New York, NY, USA, 2015; pp. 1189–1194.
9. Mu, G.; Chen, C.; Li, X.; Li, J.; Ju, X.; Dai, J. Multimodal Sentiment Analysis of Government Information Comments Based on Contrastive Learning and Cross-Attention Fusion Networks. *IEEE Access* **2024**, *12*, 165525–165538. [CrossRef]
10. Wu, X.; Mao, J.; Xie, H.; Li, G. Identifying Humanitarian Information for Emergency Response by Modeling the Correlation and Independence between Text and Images. *Inf. Process. Manag.* **2022**, *59*, 102977. [CrossRef]
11. Ahmad, Z.; Jindal, R.; Mukuntha, N.S.; Ekbal, A.; Bhattachharyya, P. Multi-Modality Helps in Crisis Management: An Attention-Based Deep Learning Approach of Leveraging Text for Image Classification. *Expert Syst. Appl.* **2022**, *195*, 116626. [CrossRef]
12. Hossain, E.; Hoque, M.M.; Hoque, E.; Islam, M.S. A Deep Attentive Multimodal Learning Approach for Disaster Identification from Social Media Posts. *IEEE Access* **2022**, *10*, 46538–46551. [CrossRef]
13. Wahid, J.A.; Shi, L.; Gao, Y.; Yang, B.; Wei, L.; Tao, Y.; Hussain, S.; Ayoub, M.; Yagoub, I. Topic2Labels: A Framework to Annotate and Classify the Social Media Data through LDA Topics and Deep Learning Models for Crisis Response. *Expert Syst. Appl.* **2022**, *195*, 116562. [CrossRef]
14. Deng, X.; Li, Y.; Weng, J.; Zhang, J. Feature Selection for Text Classification: A Review. *Multimed. Tools Appl.* **2018**, *78*, 3797–3816. [CrossRef]
15. Jiang, L.; Li, C.; Wang, S.; Zhang, L. Deep Feature Weighting for Naive Bayes and Its Application to Text Classification. *Eng. Appl. Artif. Intell.* **2016**, *52*, 26–39. [CrossRef]
16. Xu, S. Bayesian Naïve Bayes Classifiers to Text Classification. *J. Inf. Sci.* **2016**, *44*, 48–59. [CrossRef]
17. Jalal, N.; Mehmood, A.; Choi, G.S.; Ashraf, I. A Novel Improved Random Forest for Text Classification Using Feature Ranking and Optimal Number of Trees. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 2733–2742. [CrossRef]
18. Chen, H.; Wu, L.; Chen, J.; Lu, W.; Ding, J. A Comparative Study of Automated Legal Text Classification Using Random Forests and Deep Learning. *Inf. Process. Manag.* **2022**, *59*, 102798. [CrossRef]
19. Luo, X. Efficient English Text Classification Using Selected Machine Learning Techniques. *Alex. Eng. J.* **2021**, *60*, 3401–3409. [CrossRef]
20. Cervantes, J.; Garcia-Lamont, F.; Rodríguez-Mazahua, L.; Lopez, A. A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing* **2020**, *408*, 189–215. [CrossRef]
21. Pintas, J.T.; Fernandes, L.A.F.; Garcia, A.C.B. Feature Selection Methods for Text Classification: A Systematic Literature Review. *Artif. Intell. Rev.* **2021**, *54*, 6149–6200. [CrossRef]
22. Abiodun, E.O.; Alabdulatif, A.; Abiodun, O.I.; Alawida, M.; Alabdulatif, A.; Alkhawaldeh, R.S. A Systematic Review of Emerging Feature Selection Optimization Methods for Optimal Text Classification: The Present State and Prospective Opportunities. *Neural Comput. Appl.* **2021**, *33*, 15091–15118. [CrossRef]
23. Song, X.; Zhang, Y.; Zhang, W.; He, C.; Hu, Y.; Wang, J.; Gong, D. Evolutionary Computation for Feature Selection in Classification: A Comprehensive Survey of Solutions, Applications and Challenges. *Swarm Evol. Comput.* **2024**, *90*, 101661. [CrossRef]
24. Endalie, D.; Haile, G.; Taye Abebe, W. Feature Selection by Integrating Document Frequency with Genetic Algorithm for Amharic News Document Classification. *PeerJ Comput. Sci.* **2022**, *8*, e961. [CrossRef] [PubMed]
25. Ansari, G.J.; Shah, J.H.; Farias, M.C.Q.; Sharif, M.; Qadeer, N.; Khan, H.U. An Optimized Feature Selection Technique in Diversified Natural Scene Text for Classification Using Genetic Algorithm. *IEEE Access* **2021**, *9*, 54923–54937. [CrossRef]
26. Asif, M.; Nagra, A.A.; Ahmad, M.B.; Masood, K. Feature Selection Empowered by Self-Inertia Weight Adaptive Particle Swarm Optimization for Text Classification. *Appl. Artif. Intell.* **2021**, *36*, e2004345. [CrossRef]
27. Lu, Y.; Liang, M.; Ye, Z.; Cao, L. Improved Particle Swarm Optimization Algorithm and Its Application in Text Feature Selection. *Appl. Soft. Comput.* **2015**, *35*, 629–636. [CrossRef]
28. Sun, L.; Si, S.; Ding, W.; Xu, J.; Zhang, Y. BSSFS: Binary Sparrow Search Algorithm for Feature Selection. *Int. J. Mach. Learn. Cybern.* **2023**, *14*, 2633–2657. [CrossRef]
29. Babalik, A.; Babadag, A. A Binary Sparrow Search Algorithm for Feature Selection on Classification of X-Ray Security Images. *Appl. Soft. Comput.* **2024**, *158*, 111546. [CrossRef]
30. Wang, J.; Wang, W.; Hu, X.; Qiu, L.; Zang, H. Black-Winged Kite Algorithm: A Nature-Inspired Meta-Heuristic for Solving Benchmark Functions and Engineering Problems. *Artif. Intell. Rev.* **2024**, *57*, 98. [CrossRef]

31. Zhao, H.; Li, P.; Duan, S.; Gu, J. Inversion of Image-Only Intrinsic Parameters for Steel Fibre Concrete under Combined Rate-Temperature Conditions: An Adaptively Enhanced Machine Learning Approach. *J. Build. Eng.* **2024**, *94*, 109836. [CrossRef]
32. Zhang, Z.; Wang, X.; Yue, Y. Heuristic Optimization Algorithm of Black-Winged Kite Fused with Osprey and Its Engineering Application. *Biomimetics* **2024**, *9*, 595. [CrossRef]
33. Gong, X.; Wang, X.; Xiong, W.; Zhang, H.; Xin, Y. An Effective PID Control Method of Air Conditioning System for Electric Drive Workshop Based on IBK-IFNN Two-Stage Optimization. *J. Build. Eng.* **2024**, *98*, 111028. [CrossRef]
34. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. *IEEE Trans. Evol. Comput.* **1997**, *1*, 67–82. [CrossRef]
35. Gong, H.; Li, Y.; Zhang, J.; Zhang, B.; Wang, X. A New Filter Feature Selection Algorithm for Classification Task by Ensembling Pearson Correlation Coefficient and Mutual Information. *Eng. Appl. Artif. Intell.* **2024**, *131*, 107865. [CrossRef]
36. Liu, Y.; Mu, Y.; Chen, K.; Li, Y.; Guo, J. Daily Activity Feature Selection in Smart Homes Based on Pearson Correlation Coefficient. *Neural Process. Lett.* **2020**, *51*, 1771–1787. [CrossRef]
37. Zhu, Y.; Zhao, L.; Chen, X.; Li, Y.; Wang, J. Identification of Cashmere and Wool Based on LBP and GLCM Texture Feature Selection. *J. Eng. Fiber Fabr.* **2023**, *18*. [CrossRef]
38. Thaseen, I.S.; Kumar, C.A.; Ahmad, A. Integrated Intrusion Detection Model Using Chi-Square Feature Selection and Ensemble of Classifiers. *Arab. J. Sci. Eng.* **2018**, *44*, 3357–3368. [CrossRef]
39. Vergara, J.R.; Estévez, P.A. A Review of Feature Selection Methods Based on Mutual Information. *Neural Comput. Appl.* **2013**, *24*, 175–186. [CrossRef]
40. Zhou, H.; Wang, X.; Zhu, R. Feature Selection Based on Mutual Information with Correlation Coefficient. *Appl. Intell.* **2021**, *52*, 5457–5474. [CrossRef]
41. Bommert, A.; Sun, X.; Bischl, B.; Rahnenführer, J.; Lang, M. Benchmark for Filter Methods for Feature Selection in High-Dimensional Classification Data. *Comput. Stat. Data Anal.* **2020**, *143*, 106839. [CrossRef]
42. Ming, H.; Heyong, W. Filter Feature Selection Methods for Text Classification: A Review. *Multimed. Tools Appl.* **2023**, *83*, 2053–2091. [CrossRef]
43. ElDahshan, K.A.; AlHabshy, A.A.; Mohammed, L.T. Filter and Embedded Feature Selection Methods to Meet Big Data Visualization Challenges. *CMC-Comput. Mat. Contin.* **2023**, *74*, 817–839. [CrossRef]
44. Zhang, H.; Wang, J.; Sun, Z.; Zurada, J.M.; Pal, N.R. Feature Selection for Neural Networks Using Group Lasso Regularization. *IEEE Trans. Knowl. Data Eng.* **2020**, *32*, 659–673. [CrossRef]
45. Li, F.; Lai, L.; Cui, S. On the Adversarial Robustness of LASSO Based Feature Selection. *IEEE Trans. Signal Process.* **2021**, *69*, 5555–5567. [CrossRef]
46. Dupré la Tour, T.; Eickenberg, M.; Nunez-Elizalde, A.O.; Gallant, J.L. Feature-Space Selection with Banded Ridge Regression. *Neuroimage* **2022**, *264*, 119728. [CrossRef]
47. Paul, S.; Drineas, P. Feature Selection for Ridge Regression with Provable Guarantees. *Neural Comput.* **2016**, *28*, 716–742. [CrossRef] [PubMed]
48. Jenul, A.; Schrunner, S.; Liland, K.H.; Indahl, U.G.; Futsaether, C.M.; Tomic, O. RENT—Repeated Elastic Net Technique for Feature Selection. *IEEE Access* **2021**, *9*, 152333–152346. [CrossRef]
49. Amini, F.; Hu, G. A Two-Layer Feature Selection Method Using Genetic Algorithm and Elastic Net. *Expert Syst. Appl.* **2021**, *166*, 114072. [CrossRef]
50. Zhou, H.; Zhang, J.; Zhou, Y.; Guo, X.; Ma, Y. A Feature Selection Algorithm of Decision Tree Based on Feature Weight. *Expert Syst. Appl.* **2021**, *164*, 113842. [CrossRef]
51. Duan, L.; Ge, H.; Ma, W.; Miao, J. EEG Feature Selection Method Based on Decision Tree. *Bio-Med. Mater. Eng.* **2015**, *26*, S1019–S1025. [CrossRef]
52. Chen, C.; Zhang, Q.; Yu, B.; Yu, Z.; Lawrence, P.J.; Ma, Q.; Zhang, Y. Improving Protein-Protein Interactions Prediction Accuracy Using XGBoost Feature Selection and Stacked Ensemble Classifier. *Comput. Biol. Med.* **2020**, *123*, 103899. [CrossRef] [PubMed]
53. Chen, C.; Shi, H.; Jiang, Z.; Salhi, A.; Chen, R.; Cui, X.; Yu, B. DNN-DTIs: Improved Drug-Target Interactions Prediction Using XGBoost Feature Selection and Deep Neural Network. *Comput. Biol. Med.* **2021**, *136*, 104676. [CrossRef]
54. Zhang, Y.; Jiang, Z.; Chen, C.; Wei, Q.; Gu, H.; Yu, B. DeepStack-DTIs: Predicting Drug–Target Interactions Using LightGBM Feature Selection and Deep-Stacked Ensemble Classifier. *Interdiscip. Sci.* **2021**, *14*, 311–330. [CrossRef]
55. Qiu, Y.; Ding, S.; Yao, N.; Gu, D.; Li, X. HFS-LightGBM: A Machine Learning Model Based on Hybrid Feature Selection for Classifying ICU Patient Readmissions. *Expert Syst.* **2020**, *38*, e12658. [CrossRef]
56. Parlak, B. Ensemble Feature Selection for Single-Label Text Classification: A Comprehensive Analytical Study. *Neural Comput. Appl.* **2023**, *35*, 19235–19251. [CrossRef]
57. He, Z.; Li, L.; Huang, Z.; Situ, H. Quantum-Enhanced Feature Selection with Forward Selection and Backward Elimination. *Quantum Inf. Process.* **2018**, *17*, 154. [CrossRef]
58. Reif, M.; Shafait, F. Efficient Feature Size Reduction via Predictive Forward Selection. *Pattern Recognit.* **2014**, *47*, 1664–1673. [CrossRef]

59. Pham, B.T.; Van Dao, D.; Acharya, T.D.; Van Phong, T.; Costache, R.; Van Le, H.; Nguyen, H.B.T.; Prakash, I. Performance Assessment of Artificial Neural Network Using Chi-Square and Backward Elimination Feature Selection Methods for Landslide Susceptibility Analysis. *Environ. Earth Sci.* **2021**, *80*, 686. [CrossRef]
60. Narin, A.; Isler, Y.; Ozer, M. Investigating the Performance Improvement of HRV Indices in CHF Using Feature Selection Methods Based on Backward Elimination and Statistical Significance. *Comput. Biol. Med.* **2014**, *45*, 72–79. [CrossRef]
61. Zhou, S.; Li, T.; Li, Y. Recursive Feature Elimination Based Feature Selection in Modulation Classification for MIMO Systems. *Chin. J. Electron.* **2023**, *32*, 785–792. [CrossRef]
62. Jeon, H.; Oh, S. Hybrid-Recursive Feature Elimination for Efficient Feature Selection. *Appl. Sci.* **2020**, *10*, 3211. [CrossRef]
63. Chaudhuri, A. Search Space Division Method for Wrapper Feature Selection on High-Dimensional Data Classification. *Knowl.-Based Syst.* **2024**, *291*, 111578. [CrossRef]
64. Alyasiri, O.M.; Cheah, Y.-N.; Abasi, A.K.; Al-Janabi, O.M. Wrapper and Hybrid Feature Selection Methods Using Metaheuristic Algorithms for English Text Classification: A Systematic Review. *IEEE Access* **2022**, *10*, 39833–39852. [CrossRef]
65. Pham, T.H.; Raahemi, B. Bio-Inspired Feature Selection Algorithms with Their Applications: A Systematic Literature Review. *IEEE Access* **2023**, *11*, 43733–43758. [CrossRef]
66. Al-Shalif, S.A.; Senan, N.; Saeed, F.; Ghaban, W.; Ibrahim, N.; Aamir, M.; Sharif, W. A Systematic Literature Review on Meta-Heuristic Based Feature Selection Techniques for Text Classification. *PeerJ Comput. Sci.* **2024**, *10*, e2084. [CrossRef]
67. Alyasiri, O.M.; Cheah, Y.-N.; Zhang, H.; Al-Janabi, O.M.; Abasi, A.K. Text Classification Based on Optimization Feature Selection Methods: A Review and Future Directions. *Multimed. Tools Appl.* **2024**. [CrossRef]
68. Lv, J.; Wang, X.; Shao, C. AMAE: Adversarial Multimodal Auto-Encoder for Crisis-Related Tweet Analysis. *Computing* **2022**, *105*, 13–28. [CrossRef]
69. Mu, G.; Li, J.; Liao, Z.; Yang, Z. An Enhanced IHHO-LSTM Model for Predicting Online Public Opinion Trends in Public Health Emergencies. *SAGE Open* **2024**, *14*. [CrossRef]
70. Madichetty, S.; M, S. A Stacked Convolutional Neural Network for Detecting the Resource Tweets during a Disaster. *Multimed. Tools Appl.* **2020**, *80*, 3927–3949. [CrossRef]
71. Giri, K.S.V.; Deepak, G. A Semantic Ontology Infused Deep Learning Model for Disaster Tweet Classification. *Multimed. Tools Appl.* **2023**, *83*, 62257–62285. [CrossRef]
72. Madichetty, S.; M, S. Classifying Informative and Non-Informative Tweets from the Twitter by Adapting Image Features during Disaster. *Multimed. Tools Appl.* **2020**, *79*, 28901–28923. [CrossRef]
73. Zhou, B.; Zou, L.; Mostafavi, A.; Lin, B.; Yang, M.; Gharaibeh, N.; Cai, H.; Abedin, J.; Mandal, D. VictimFinder: Harvesting Rescue Requests in Disaster Response from Social Media with BERT. *Comput. Environ. Urban Syst.* **2022**, *95*, 101824. [CrossRef]
74. Gao, W.; Zheng, C.; Zhu, X.; Deng, H.; Wang, Y.; Hu, G. Knowledge-Injected Prompt Learning for Actionable Information Extraction from Crisis-Related Tweets. *Comput. Electr. Eng.* **2024**, *118*, 109398. [CrossRef]
75. Devaraj, A.; Murthy, D.; Dontula, A. Machine-Learning Methods for Identifying Social Media-Based Requests for Urgent Help during Hurricanes. *Int. J. Disaster Risk Reduct.* **2020**, *51*, 101757. [CrossRef]
76. Singh, J.P.; Dwivedi, Y.K.; Rana, N.P.; Kumar, A.; Kapoor, K.K. Event Classification and Location Prediction from Tweets during Disasters. *Ann. Oper. Res.* **2017**, *283*, 737–757. [CrossRef]
77. Rudra, K.; Goyal, P.; Ganguly, N.; Imran, M.; Mitra, P. Summarizing Situational Tweets in Crisis Scenarios: An Extractive-Abstractive Approach. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 981–993. [CrossRef]
78. Ramachandran, D.; R, P. A Novel Domain and Event Adaptive Tweet Augmentation Approach for Enhancing the Classification of Crisis Related Tweets. *Data Knowl. Eng.* **2021**, *135*, 101913. [CrossRef]
79. Karimiziarani, M.; Moradkhani, H. Social Response and Disaster Management: Insights from Twitter Data Assimilation on Hurricane Ian. *Int. J. Disaster Risk Reduct.* **2023**, *95*, 103865. [CrossRef]
80. Zhang, Y.; Wang, N.; Yaya Hamadou, K.; Su, Y.; Guo, X.; Song, W. Multi-Level Emotion Propagation in Natural Disaster Events: Diverse Leadership of Super-Spreaders in Different Levels of Hierarchy. *Online Inf. Rev.* **2024**, *49*. [CrossRef]
81. Mu, G.; Liao, Z.; Li, J.; Qin, N.; Yang, Z. IPSO-LSTM Hybrid Model for Predicting Online Public Opinion Trends in Emergencies. *PLoS ONE* **2023**, *18*, e0292677. [CrossRef]
82. Mendon, S.; Dutta, P.; Behl, A.; Lessmann, S. A Hybrid Approach of Machine Learning and Lexicons to Sentiment Analysis: Enhanced Insights from Twitter Data of Natural Disasters. *Inf. Syst. Front.* **2021**, *23*, 1145–1168. [CrossRef]
83. Win Myint, P.Y.; Lo, S.L.; Zhang, Y. Unveiling the Dynamics of Crisis Events: Sentiment and Emotion Analysis via Multi-Task Learning with Attention Mechanism and Subject-Based Intent Prediction. *Inf. Process. Manag.* **2024**, *61*, 103695. [CrossRef]
84. Dereli, T.; Eligüzel, N.; Çetinkaya, C. Content Analyses of the International Federation of Red Cross and Red Crescent Societies (Ifrc) Based on Machine Learning Techniques through Twitter. *Nat. Hazards* **2021**, *106*, 2025–2045. [CrossRef]
85. Zhang, X.; Ma, Y. An ALBERT-Based TextCNN-Hatt Hybrid Model Enhanced with Topic Knowledge for Sentiment Analysis of Sudden-Onset Disasters. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106136. [CrossRef]

86. Li, S.; Sun, X. Application of Public Emotion Feature Extraction Algorithm Based on Social Media Communication in Public Opinion Analysis of Natural Disasters. *PeerJ Comput. Sci.* **2023**, *9*, e1417. [CrossRef]
87. Minocha, S.; Singh, B. A Novel Sentimental Analysis for Response to Natural Disaster on Twitter Data. *Int. J. Pattern Recognit. Artif. Intell.* **2023**, *37*. [CrossRef]
88. Karimiziarani, M.; Shao, W.; Mirzaei, M.; Moradkhani, H. Toward Reduction of Detrimental Effects of Hurricanes Using a Social Media Data Analytic Approach: How Climate Change Is Perceived? *Clim. Risk Manag.* **2023**, *39*, 100480. [CrossRef]
89. Ma, M.; Gao, Q.; Xiao, Z.; Hou, X.; Hu, B.; Jia, L.; Song, W. Analysis of Public Emotion on Flood Disasters in Southern China in 2020 Based on Social Media Data. *Nat. Hazards* **2023**, *118*, 1013–1033. [CrossRef]
90. Zander, K.K.; Garnett, S.T.; Ogie, R.; Alazab, M.; Nguyen, D. Trends in Bushfire Related Tweets during the Australian 'Black Summer' of 2019/20. *For. Ecol. Manag.* **2023**, *545*, 121274. [CrossRef]
91. Kruspe, A.; Kersten, J.; Klan, F. Review Article: Detection of Actionable Tweets in Crisis Events. *Nat. Hazards Earth Syst. Sci.* **2021**, *21*, 1825–1845. [CrossRef]
92. Huang, Q.; Xiao, Y. Geographic Situational Awareness: Mining Tweets for Disaster Preparedness, Emergency Response, Impact, and Recovery. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1549–1568. [CrossRef]
93. Zou, L.; He, Z.; Zhou, C.; Zhu, W. Multi-Class Multi-Label Classification of Social Media Texts for Typhoon Damage Assessment: A Two-Stage Model Fully Integrating the Outputs of the Hidden Layers of BERT. *Int. J. Digit. Earth* **2024**, *17*, 2348668. [CrossRef]
94. Kersten, J.; Klan, F. What Happens Where during Disasters? A Workflow for the Multifaceted Characterization of Crisis Events Based on Twitter Data. *J. Cont. Crisis Manag.* **2020**, *28*, 262–280. [CrossRef]
95. Palshikar, G.K.; Apte, M.; Pandita, D. Weakly Supervised and Online Learning of Word Models for Classification to Detect Disaster Reporting Tweets. *Inf. Syst. Front.* **2018**, *20*, 949–959. [CrossRef]
96. Sit, M.A.; Koylu, C.; Demir, I. Identifying Disaster-Related Tweets and Their Semantic, Spatial and Temporal Context Using Deep Learning, Natural Language Processing and Spatial Analysis: A Case Study of Hurricane Irma. *Int. J. Digit. Earth* **2019**, *12*, 1205–1229. [CrossRef]
97. Bhoi, A.; Balabantaray, R.C.; Sahoo, D.; Dhiman, G.; Khare, M.; Narducci, F.; Kaur, A. Mining Social Media Text for Disaster Resource Management Using a Feature Selection Based on Forest Optimization. *Comput. Ind. Eng.* **2022**, *169*, 108280. [CrossRef]
98. Mujahid, M.; Lee, E.; Rustam, F.; Washington, P.B.; Ullah, S.; Reshi, A.A.; Ashraf, I. Sentiment Analysis and Topic Modeling on Tweets about Online Education during COVID-19. *Appl. Sci.* **2021**, *11*, 8438. [CrossRef]
99. Alsubari, S.N. Data Analytics for the Identification of Fake Reviews Using Supervised Learning. *CMC-Comput. Mat. Contin.* **2022**, *70*, 3189–3204. [CrossRef]
100. Wen, X.; Liu, X.; Yu, C.; Gao, H.; Wang, J.; Liang, Y.; Yu, J.; Bai, Y. IOOA: A Multi-Strategy Fusion Improved Osprey Optimization Algorithm for Global Optimization. *Electron. Res. Arch.* **2024**, *32*, 2033–2074. [CrossRef]
101. Zhang, J.; Zheng, J.; Xie, X.; Lin, Z.; Li, H. Mayfly Sparrow Search Hybrid Algorithm for RFID Network Planning. *IEEE Sens. J.* **2022**, *22*, 16673–16686. [CrossRef]
102. Ding, J.; Huang, S.; Hao, Y.; Xiao, M. A Modified Reptile Search Algorithm for Parametric Estimation of Fractional Order Model of Lithium Battery. *Optim. Control. Appl. Methods* **2023**, *44*, 3204–3218. [CrossRef]
103. Liu, W.; Yan, W.; Li, T.; Han, G.; Ren, T. A Multi-Strategy Improved Grasshopper Optimization Algorithm for Solving Global Optimization and Engineering Problems. *Int. J. Comput. Intell. Syst.* **2024**, *17*, 182. [CrossRef]
104. Fan, Z.; Xiao, Z.; Li, X.; Huang, Z.; Zhang, C. MSBWO: A Multi-Strategies Improved Beluga Whale Optimization Algorithm for Feature Selection. *Biomimetics* **2024**, *9*, 572. [CrossRef] [PubMed]
105. Xu, Y.; Yang, Z.; Li, X.; Kang, H.; Yang, X. Dynamic Opposite Learning Enhanced Teaching–Learning-Based Optimization. *Knowl.-Based Syst.* **2020**, *188*, 104966. [CrossRef]
106. Chandran, V.; Mohapatra, P. Enhanced Opposition-Based Grey Wolf Optimizer for Global Optimization and Engineering Design Problems. *Alex. Eng. J.* **2023**, *76*, 429–467. [CrossRef]
107. Nelder, J.A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313. [CrossRef]
108. Shrivastava, P.; Kok Soon, T.; Bin Idris, M.Y.I.; Mekhilef, S.; Adnan, S.B.R.S. Combined State of Charge and State of Energy Estimation of Lithium-Ion Battery Using Dual Forgetting Factor-Based Adaptive Extended Kalman Filter for Electric Vehicle Applications. *IEEE Trans. Veh. Technol.* **2021**, *70*, 1200–1215. [CrossRef]
109. Nguyen, D.T.; Joty, S.; Imran, M.; Sajjad, H.; Mitra, P. Applications of Online Deep Learning for Crisis Response Using Social Media Information. Available online: https://arxiv.org/abs/1610.01030 (accessed on 1 September 2024).
110. Ofli, F.; Alam, F.; Imran, M. Analysis of Social Media Data Using Multimodal Deep Learning for Disaster Response. Available online: https://arxiv.org/abs/2004.11838 (accessed on 1 September 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.