The International Journal of Interdisciplinary Organizational Studies
ISSN: 2324-7649 (Print), ISSN: 2324-7657 (Online)
Volume 19, Issue 1, January-June, 2024

A NOVEL HYBRID K-MEANS AND IGOA-GA APPROACH FOR CRIME DATA CLUSTERING

K. Dheenathayalan[1], K.K. Savitha[2]
[1]Assistant Professor, Department of Computer Science, Kamban College of Arts and Science, Coimbatore, Tamil Nadu, India.
[2]Assistant Professor, Department of Computer Applications, Bharathiar University PG Extension and
Research Centre, Erode, Tamil Nadu, India.
[1]dheenamca08@gmail.com, [2]savitha.gopinath@gmail.com

Abstract: Clustering crime data helps to identify and understand the underlying crime patterns. The k-means algorithm is very familiar and effectively applied in clustering problems, and grieves from a few drawbacks due to its choice of initial cluster centroids. A hybrid method based on combining the k-means algorithm, improved grasshopper optimization algorithm, and genetic algorithm called KM-IGOA-GA is proposed in this research work. KM-IGOA-GA searches for cluster centers of the dataset as does the k-means algorithm and it can effectively find the global optima. The new KM-IGOA-GA is applied to the crime dataset acquired from NCRB, India and its performance is compared with those of KM-IGOA, KM-GOA, KM-FOA, KM-MFO, and K-means clustering. Results show that KM-IGOA-GA is effective and suitable for data clustering with an accuracy of 89%. This proposed work is implemented using R-Studio.
Keywords: clustering, optimization, metaheuristic, improved grasshopper, k-means

1. Introduction

The optimization strategies are proposed to unravel the issues with tremendous numbers of factors, enormous complexity, or having no analytical solutions that were met during the behaviour of investigating, exploiting, and overcoming nature by humans. It is always difficult to discover a widespread effective way for nearly all issues. Therefore, researchers and engineers around the world are still beneath ways to discover more optimization calculations and more appropriate strategies. The stochastic algorithms are found to be capable of achieving global solutions with randomness, thus, stochastic algorithms are being focused and more algorithms are proposed. Nature-inspired algorithms are stochastic algorithms proposed with inspiration of the nature and they have been proven to be effective in optimization. Some of the metaheuristic algorithms are stochastic algorithms and they are swarm intelligence algorithms. Nearly all of these algorithms and their enhancements so far are stimulated directly by the behaviours of the creatures such as searching, hunting, pollinating, and flashing [20,21].

K-means clustering is a center-based algorithm that is very effective in dealing with large and high-dimensional datasets. K-means is a form of hard partitional clustering as each data object is assigned to only one cluster. In K-means, the process of assigning data objects to the disjoint clusters repeats until there is no significant change in objective function values or membership of clusters it is called convergence. The K-means algorithm's advantages include quicker convergence and effectiveness in clustering big data sets. It creates N-dimensional population clusters with the least dissimilarity metric for each cluster. Each cluster in K-means has a mean at its center, and the goal is to minimize the value of its objective function, which is a sum of square error function. K-means reliance on the initial cluster centers, sensitivity to outliers, lack of guaranteed optimal solutions, and tendency to produce unbalanced clusters are some of its drawbacks. The algorithm falls into the local optima because it relies on its initial cluster centers.

The grasshopper optimization algorithm was proposed by Saremi et al. in [25], which is a swarm intelligence algorithm that mimics grasshoppers' natural seeking and swarming behaviours. Grasshoppers are insects, a hazardous pest that affects and harm crop production and agriculture [25]. Their life cycle includes two stages nymph and adulthood. The nymph stage is characterized by small steps and slow movements, while the adulthood stage is characterized by long-range and rapid movements. The movements of nymphs and adulthood constitute the intensification and diversification phases of GOA.

The Firefly Algorithm examines the foraging habits of fireflies by using a population-based strategy to determine the global optima of objective functions based on swarm intelligence [18,19]. Physical elements, such as agents or fireflies, are dispersed randomly throughout the search space by the firefly algorithm. Agents are compared to fireflies that exhibit luciferin, a luminescent property that causes them to emit light in proportion to this value. The stronger brightness of nearby fireflies attracts the attention of individual fireflies. The more spread out apart they are, the less attractive they are. It will move at random if there isn't a firefly that is brighter than it. Initially, all the agents (fireflies) are distributed randomly throughout the search space according to the objective function.

Moths have developed to fly in the evening utilizing the moonlight and they depend on a strategy called cross-over direction for route [26]. In this technique, a moth flies by keeping a decent point concerning the moon (for example the light source). This strategy is viewed as an exceptionally successful procedure for voyaging significant distances in a straight way. Since the moon is far away from the moth, this strategy ensures flying in a straight line. At the point when moths see a human-made counterfeit light, they attempt to keep a comparable point with the light to fly in a straight line. Since such a light is very closely contrasted with the moon, be that as it may, keeping a comparable point to the light source causes a futile or dangerous winding flyway for moths. It very well might be seen that for fake lights the moth at last unites towards the light. Moths and flames are the principal parts of the algorithm. The candidate solutions are moths and the moth's positions in space are the problem's variables. Along these moths can fly in 1-D, 2-D, three-dimensional, or even hyper-layered space (of aspect d) with changing position vectors. Since the MFO algorithm is a population-based algorithm, the arrangement of n moths is utilized as search specialists in the issue space. Flames are the best n places for moths that are acquired up to this point. Hence, every moth looks through around a banner (fire) and updates it in the event of tracking down an improved arrangement.

2. Methodology

2.1 The proposed KM-IGOA-GA clustering approach

Selecting appropriate initial cluster centers is necessary to improve clustering using K-means. Due to K-means' sensitivity to the number of clusters and centroid locations that are relatively confined in the search space, it is unable to fully explore the search space for the global optima.

This work demonstrates the integration of bio-inspired optimization techniques with K-means to solve this K-means limitation. The K-Means clustering algorithm uses the outputs of this optimization technique as its initial cluster centers.

Grasshopper optimization algorithm (GOA), firefly optimization algorithm (FOA), and moth-flame optimization algorithm (MFO) are population-based algorithms. Hence the initial population is generated from given data objects. At first generation, agents are initialized randomly. Each agent represents k clustering centers. In this work, each agent represents one candidate solution to a problem so one agent is k initial clustering centers. The whole population represents a variety of clustering schemes. In this manner, the bio-inspired algorithms aim to find optimal clustering centers according to the objective function value (distance measure). After that, optimal cluster centers from each bio-inspired algorithm will be taken to K-means for clustering the data objects. Finally, a better technique is selected based on statistical analysis [22, 23].

High-crime-intensity areas are known as crime hotspots. Identifications of hotspots help public safety organizations allocate resources for crime prevention activities. This is known as geographical analysis usually made based on maps and reported crime events over a certain period. The k-means algorithm is integrated with bio-inspired algorithms for optimizing cluster centers. The performance of bio-inspired algorithms is measured using some cluster analysis metrics. They are the standard deviation (SD), the within-cluster sum of squares (WCSS), the between-cluster sum of squares (BCSS), and the average silhouette score. The state-wise, and district-wise cluster results are mapped to the respected boundary shapefile. A shapefile is a simple, nontopological format for storing the geometric location and attribute information of geographic features.

2.2 Improved Grasshopper Optimization Algorithm

The improved grasshopper optimization algorithm is proposed by Qin P., et. al. in [1]. The impact of gravity on the updated position of each grasshopper is not considered in basic GOA, which may describe why GOA has a slower convergence rate. Based on this, the two updated methods of each grasshopper's position yield the improved GOA (IGOA). One is that in the basic GOA, every grasshopper's updated position now includes the gravitational force. The other is that each grasshopper's updated position includes velocity, and the new position is determined by adding the velocity to the current position. Then, based on likelihood, each grasshopper accepts the modified position most appropriately. finally, IGOA is performed initially on the 23 classical benchmark functions.

2.3 Genetic algorithm

In GAs, the parameters of the search space are encoded in the form of strings called chromosomes. A collection of such strings is called a population. Initially, a random population is created, which represents different points in the search space. An objective and fitness function are associated with each string that represents the degree of goodness of the string. Based on the principle of survival of the fittest, a few of the strings are selected and each is assigned some copies that go into the mating pool. Biologically inspired operators like crossover and mutation are applied to these strings to yield a new generation. The process of selection, crossover, and mutation continues for a "fixed number of generations or till a termination condition is satisfied [27].

3. KM-IGOA-GA

The KM-IGOA contains three phases namely, initialization phase, fitness calculation phase, and exploration and exploitation phase. In the initialization phase, initial parameters such as, the number of clusters k, population size pop, the maximum number of iterations maxiter, initial position of each grasshopper are initialized. The fitness value WCSS (Within Cluster Sum of Squares) of every grasshopper is calculated using the default distance measure Euclidean distance once the grasshoppers are clustered into groups. Based on the best (minimum) fitness value the target position (best centroid) is selected. The other grasshopper parameters coefficient, gravity force, social force, velocity, distance, and probability are calculated for each grasshopper. The current position of each grasshopper is updated based on the probability value p. If $p < 0.5$ then the updated position of the grasshopper is obtained by eq.1. If $p >= 0.5$ then the updated position of the grasshopper is obtained by eq.2.

$$p < 0.5 , \quad P_i^d = c\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} c \, \frac{ub_d - lb_d}{2} \, S\left(|P_j^d - P_i^d| \right) \frac{P_j - P_i}{d_{ij}} \right) - g \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{P_j - P_i}{d_{ij}} \right) + \hat{T}_d \qquad \text{eq.1}$$

$$p >= 0.5 , \quad P_i^d = c\left(\sum_{\substack{j=1 \\ j\neq i}}^{N} c \, \frac{ub_d - lb_d}{2} \, S\left(|P_j^d - P_i^d| \right) \frac{P_j - P_i}{d_{ij}} \right) - g \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{P_j - P_i}{d_{ij}} \right) + \hat{T}_d + v \qquad \text{eq.2}$$

The output of IGOA is the updated position of each grasshopper. These grasshoppers are the initial population for GA. Each grasshopper is a set of cluster centroids and is considered a single chromosome. Firstly, for each chromosome, the fitness value (WCSS) is calculated, and then the ten chromosomes with the best fitness value are selected for crossover and mutation operations. The selection of the best chromosomes among the total population is done only once to reduce the burden of the genetic algorithm. Since the genetic algorithm is a maximization problem, the fitness value is the inverse of WCSS. After the first crossover and mutation operations, the usual genetic algorithm process continues until it reaches the maximum number of iterations. Finally, the best chromosome (best centroid) is obtained from GA. This best centroid is applied in k-means to test the accuracy.

Figure 1(a) represents the flowchart of the KM-IGOA method for clustering and Figure 1(b) represents the flowchart of the proposed method KM-IGOA-GA which combines improved grasshopper optimization, genetic algorithm, and k-means. Initially, the cluster centroids obtained from k-means are optimized using IGOA, then the top ten of these optimized clusters are once again optimized by genetic algorithm, and finally, the best cluster centroids are obtained. The performance of the proposed method is compared with existing clustering techniques like, simple k-means, KM-GOA, KM-FOA, KM-MFO, and KM-IGOA. the clustering performance of these techniques is measured using a few of the cluster quality metrics such as standard deviation, average silhouette score, within cluster sum of squares, between cluster sum of squares, and accuracy percentage.

Steps in integrated bio-inspired clustering

The integrated clustering technique is divided into four basic steps

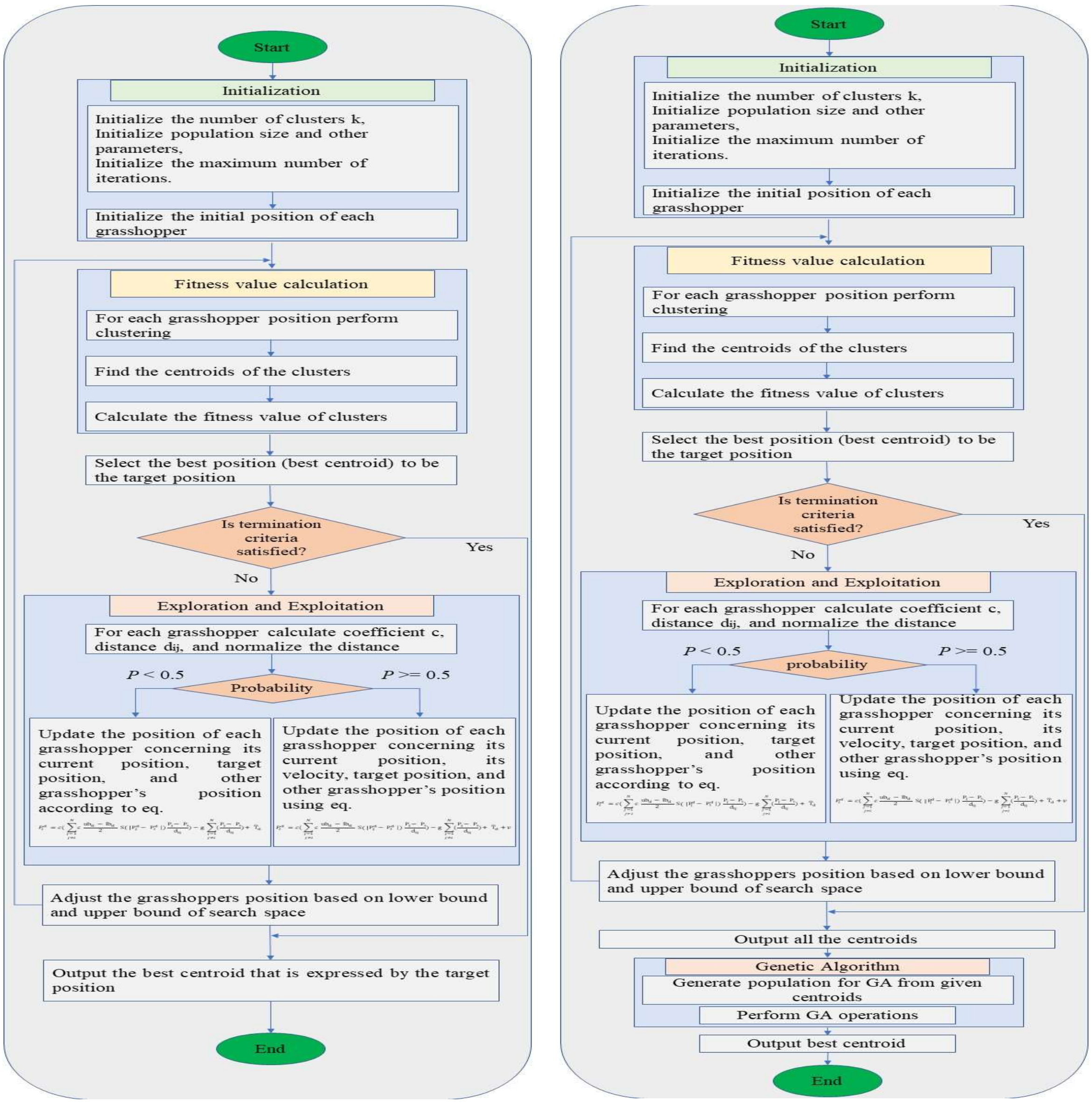
(1) Initialization.

(2) Cluster assignment and fitness calculation.

(3) Exploration and Centroid locations update.

(4) Termination.

Here, all the bio-inspired algorithms that are used in the integration with K-Means employ the same initialization step, cluster assignment and fitness calculation step and termination step but each algorithm has its unique heuristic method to explore the search space to find the optimal solution.



a) KM-IGOA b) KM-IGOA-GA

Figure 1. Flowchart of the proposed technique

3.1 Initialization

Here, all the biologically inspired algorithms used in the integration with K-Means use the same initialization step, cluster assignment, and coverage calculation step as well as the termination step, but each algorithm has its heuristic to explore the search space to find the optimal solution.

3.2 Cluster assignment and fitness calculation

In the proposed method cluster assignment step follows the same rule as the original k-means. Each search agent is represented as a matrix of size k*D where D is the number of attributes of the dataset and therefore it is the dimension of the search space, and the i-th search agent will look like A_i = [c1, c2, c3] where c1, c2 and c3 are vectors of size 1x3 that represent the clusters centroid locations for example, c1 = [x_{i1}, x_{i2}, x_{i3}] where x_i is a data point in the search space. The population is represented as a size N*K*D three-dimensional matrix, where N is the population's total size. To allocate each data point to the nearest cluster centroid, we first compute the Euclidean distance between each data point and each cluster centroid in the search agent. The Euclidean distance, which we employed in the proposed method, is shown as follows:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad \text{eq.3}$$

Our objective function is to find the best group of centroid locations that form clusters with the minimum sum of the squared error, therefore we used SSE to be our objective function value.

3.3 Exploration and Exploitation

In bio-inspired algorithms, the objective function value is the position variable of the centroid. This position variable is relocated many times targeting to catch a position that achieves the best objective value which is the minimum SSE value of the clustering process. In KM-FOA, KM-MFO, KM-GOA, and KM-IGOA methods, each agent updates its position concerning the current position and other important parameters. In KM-IGOA, the probability value p is calculated, if p is less than 0.5 then the position of the grasshopper is updated using eq.1, if p is greater than or equal to 0.5 then the position is updated using eq.2. Each grasshopper updates its position concerning its current position, target position, velocity, and other grasshopper's position. In KM-IGOA-GA, the output centroids from KM-IGOA are optimized using the genetic algorithm. Finally, the best centroid is produced.

3.4 Termination

The optimization process ends when the optimal solution is found or when it reaches the maximum number of iterations. In the proposed method, the algorithm breaks when it hits the maximum number of iterations. After the termination, the output will be the best search agent with the best fitness value

4. Dataset description

The datasets were acquired from the NCRB website of India. These datasets are crime statistics of different states, union territories, and districts in India. Table 1 represents sample data which contains the number of incidents under five different crime heads. The full dataset contains the number of incidents under 17 different crime heads for 28 states and 8 union territories. These crime heads are classified into IPC (Indian Penal Code) and SLL (Special Local Laws) crimes. There are two datasets other than this, which are state-wise and district-wise datasets, they consist year-wise total number of crime incidents from 2001-2020.

Table 1. Sample crime dataset

State	rape/ attempted	kidnap	dowry deaths	assault	insult
Andhra Pradesh	23502	21094	7804	91825	55547
Arunachal Pradesh	1147	1222	5	1561	62
Assam	33488	65723	2655	43809	1026
Bihar	22050	69026	22598	10356	658
Chhattisgarh	24342	15378	1852	33716	2977

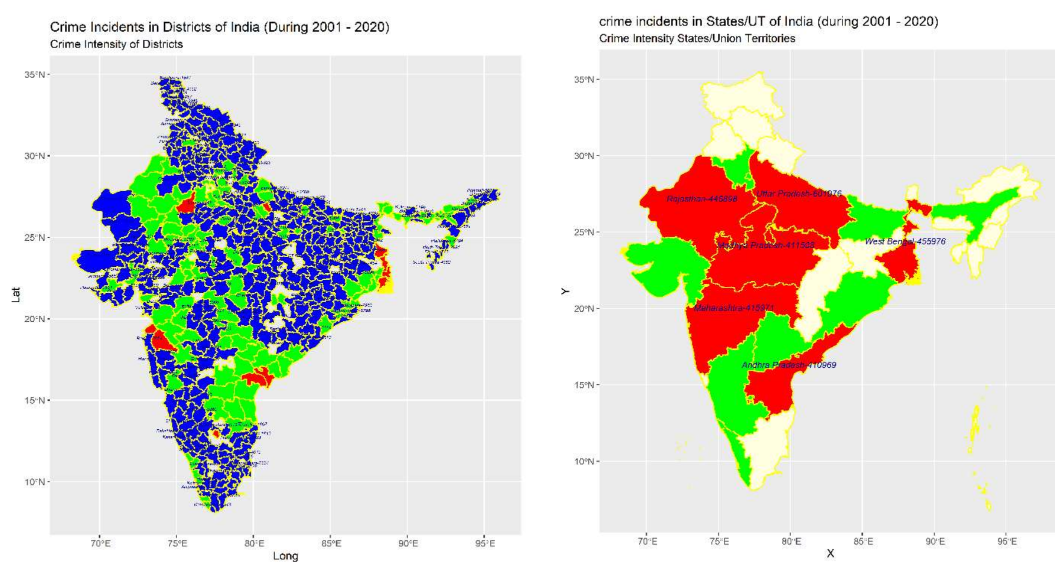
5. Result and discussion

The proposed work is implemented in the R-studio tool. Table1, represents cluster quality metrics of different clustering techniques. The effectiveness and efficiency of the basic grasshopper algorithm are improved by introducing gravity force and velocity to it. The comparison between the quality metrics of KM-GOA and KM-IGOA proves it. To improve the effectiveness of clustering, the IGOA is combined with GA to optimize the cluster centroids, proving the best performance than other techniques. According to the cluster quality metrics, we found that KM-IGOA-GA occupies the first place in the optimization performance comparison for crime datasets and KM-IGOA occupies the second place but the simple k-means occupies the last place

Table 2. Cluster quality metrics

Clustering Algorithm	SD	Avg. Sil	Total WSS	Between SS	Accuracy %
KM-IGOA-GA	**0.63**	**0.68**	**83.40**	**634.90**	**89.0**
KM-IGOA	0.65	0.66	84.10	635.89	87.3
KM-GOA	0.68	0.64	88.94	631.05	86.6
KM-MFO	0.69	0.61	101.02	618.97	84.2
KM-FOA	0.70	0.56	114.94	605.05	82.1
K-Means	0.79	0.33	170.10	549.89	76.4

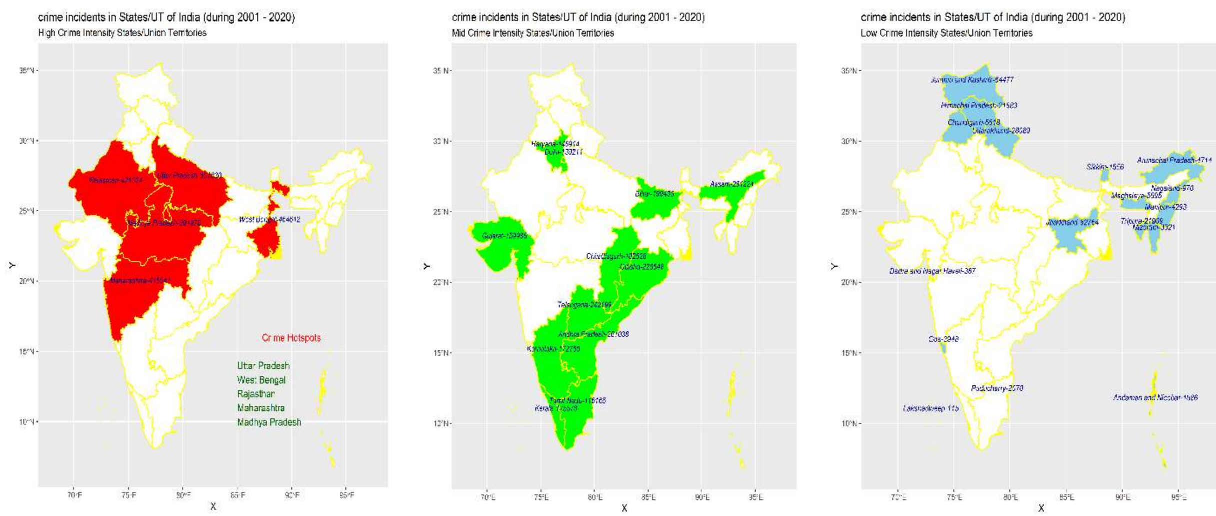
Geospatial clustering is done to predict the state-wise and district-wise crime intensity areas under three different levels namely, high, mid, and low. Figures 2(a), and (b), represent high, mid, and low crime-intensity areas concerning districts and states respectively. The districts and states in red color are high-crime intensity areas, and green in color are mid-crime intensity areas.



(a) (b)

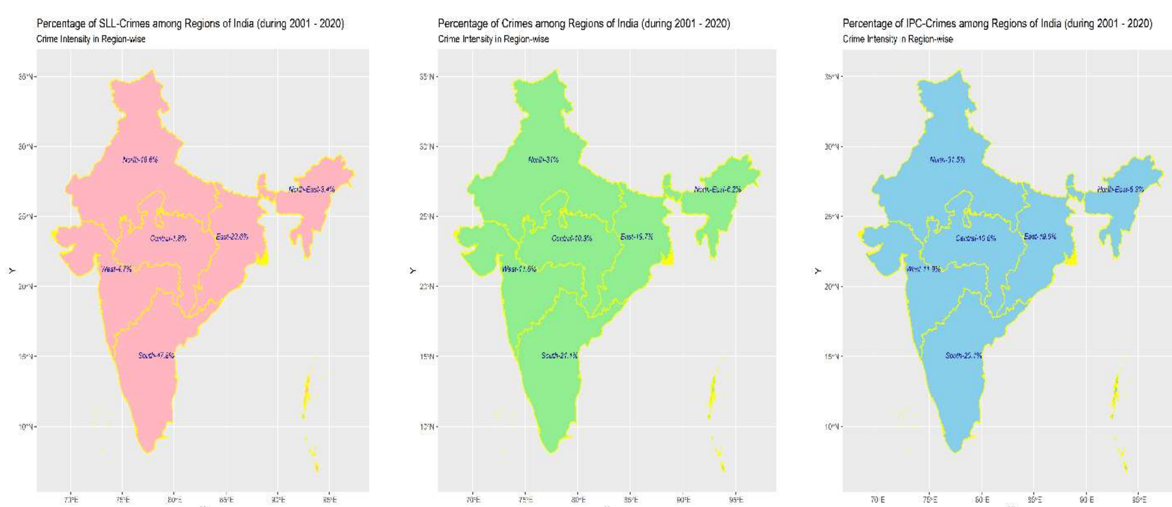
Figure 2. Crime intensity (year 2001 to 2020) – (a) District-wise, (b) State-wise

Figure 3 is the extension of Figure 2(b) and it shows high, mid, and low crime intensity states separately. There are six regions in India, they are East, West, North, South, Central, and North East. Figure 4 (a), represents the state/UT-wise percentage distribution of IPC crimes, Figure 4 (b), represents the state/UT-wise percentage distribution of SLL crimes, and Figure 4 (c), represents the state/UT-wise percentage distribution of total (IPC+SLL) crimes.



(a) (b) (c)

Figure 3. crime prune states/UTs in India – under different crime heads
(a) High, (b) Mid, and (c) Low



(a) (b) (c)

Figure 4. Region-wise crime percentage distribution in India

Conclusion

The main objective of this work is to find geospatial clusters concerning crimes committed against women in India. State/UT-wise and district-wise clusters are found to identify high, mid, and low crime-intensity areas. To find geospatial clusters, we proposed a bio-inspired clustering technique KM-IGOA-GA. This technique is a combination of improved grasshopper optimization, genetic algorithm, and k-means clustering. The basic grasshopper optimization algorithm is improved by introducing two important parameters gravity force and velocity value. The proposed technique is applied to the Indian NCRB crime dataset to find the clusters. The set of cluster centers is given as a population to IGOA, and it produces the optimized cluster centers. These optimized centers are once again optimized by genetic algorithm to improve the clustering accuracy. The proposed technique is compared with other techniques like, simple k-means, KM-GOA, KM-FOA, KM-MFO, and KM-IGOA considering cluster quality metrics such as standard deviation (SD), within cluster sum of squares (WCSS), between cluster sum of squares (BCSS), average silhouette score, and accuracy. It is proven that the proposed technique outperformed other above-mentioned techniques with the best accuracy of 89% followed by KM-IGOA with an accuracy of 87.3%. In the future, the proposed hybrid method can be enhanced by integrating other optimization techniques or replacing existing optimization techniques with other techniques or this same method can be applied to various clustering problems.

Reference

[1] Qin, P., Hu, H. & Yang, Z. The improved grasshopper optimization algorithm and its applications. Sci Rep 11, 23733 (2021). https://doi.org/10.1038/s41598-021-03049-6.

[2] Almaw and Kadam K.. Crime Data Analysis and Prediction Using Ensemble Learning, Second - ICICCS, pp. 1918-1923, doi: 10.1109/ICCONS. (2018).8663186.

[3] Alwee R, Siti Mariyam Hj Shamsuddin, and Roselina Sallehuddin. Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models Improved by Particle Swarm Optimization for Property Crime Rates Forecasting with Economic Indicators, Hindawi Publishing Corporation The Scientific World Journal Volume (2013), Article ID 951475, 11 pages

[4] Bharati A. and Sarvanaguru K.Crime Prediction and Analysis Using Machine Learning, IRJET, Volume: 05 Issue: 09, (2018), e-ISSN: 2395-0056 , p-ISSN: 2395-0072.

[5] Chen P, Hongyong Yuan, Xueming Shu. Forecasting Crime Using the ARIMA Model, Fifth International Conference on Fuzzy Systems and Knowledge Discovery, (2008), pp. 627-630, doi: 10.1109/FSKD.2008.222.

[6] Goyat and Suman. (2019). Crime analysis for Multistate Network using Naive Bayes classifier. International Journal of Computer Science and Mobile Computing, Vol.8 Issue.6, June- (2019), pg. 134-143.

[7] Hossain S, Abtahee A, Kashem I, Hoque M.M and Sarker I.H. Crime Prediction Using Spatio-Temporal Data. Computing Science, Communication and Security. COMS2 (2020). CCIS, vol 1235. Springer, Singapore. https://doi.org/10.1007/978-981-15-6648-6_22

[8] Kauser P. Prediction and Comparison of Crime Rate using Advanced Machine Learning Algorithms and Bio inspired Computing Techniques. IJAST. 29. (2020). 12479-12490

[9] Alif Ridzuan Khairuddin, Razana Alwee and Habibollah Haron. A review on applied statistical and artificial intelligence techniques in crime forecasting. IOP Conference Series: Materials Science and Engineering. (2019) 551. 012030. 10.1088/1757-899X/551/1/012030.

[10] Kim S, Param Joshi, Parminder Singh Kalsi, and Pooya Taheri. Crime Analysis Through Machine Learning," (2018) IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 415-420, doi: 10.1109/IEMCON.2018.8614828.

[11] Kounadi, O, Alina Ristea, Adelson Araujo and Michael Leitner. A systematic review on spatial crime forecasting. Crime Sci 9, 7 (2020). https://doi.org/10.1186/s40163-020-00116-7

[12] Mahmud S, Musfika Nuha and Abdus Sattar. Crime Rate Prediction Using Machine Learning and Data Mining. Soft Computing Techniques and Applications. Advances in Intelligent Systems and Computing, vol 1248. (2021), Springer, Singapore.

[13] Mrinalini J and Shaveta K. Naïve Bayes Approach for the Crime Prediction in Data Mining, (2019), IJCA(0975 – 8887) Volume 178

[14] Shah N, Nandish Bhagat and Manan Shah. Crime forecasting: a machine learning and computer vision approach to crime prediction and prevention. Vis. Comput. Ind. Biomed. (2021) Art 4.

[15] Wibowo A H and Oesman T I, J. Phys.: Conf. Ser. (2020), 1450 012076.

[16] Yassine Meraihi, Asma Benmessaoud Gabis, Seyedali Mirjalili and Amar Ramdane-cherif. Grasshopper Optimization Algorithm: Theory, Variants and Applications. DOI 10.1109 / ACCESS. (2021). 3067597, IEEE Access

[17] Tahereh Hassanzadeh and Mohammad Reza Meybodi . A New Hybrid Approach for Data Clustering using Firefly Algorithm and K-means. The 16th CSI (AISP 2012).

[18] Nurlela Pandiangan, Rahmat Gerowo and Vicensius Gunawan. K-Means Clustering and Firefly Algorithm for Shortest Route solution based on Crime Hotspots. IJCA (0975 – 8887) Volume 180 – No.52, June (2018).

[19] J. Senthilnath, S.N. Omkar and V. Mani. Clustering using firefly algorithm: Performance study. Swarm and Evolutionary Computation 1 (2011) 164–171.

[20] Doaa Abdullah, Hala Abdel-Galil and Ensaf Hussein. Enhancing K-Means Clustering with Bio-Inspired Algorithms. IJCSN - Volume 7, Issue 6, December (2018).

[21] Binitha S and S Siva Sathya. A Survey of Bio inspired Optimization Algorithms. (IJSCE) ISSN: 2231-2307, Volume-2, Issue-2, May (2012).

[22] M. A. El-Shorbagy and A. Y. Ayoub. Integrating Grasshopper Optimization Algorithm with Local Search for Solving Data Clustering Problems. IJCIS Vol. 14(1), (2021), pp. 783–793. ISSN: 1875- 6891.

[23] Simon Fong, Suash Deb, Xin-She Yang and Yan Zhuang. Towards Enhancement of Performance of K-Means Clustering Using Nature-Inspired Optimization Algorithms. The Scientific World Journal Volume (2014), Article ID 564829, 16 pages.

[24] Khushabu A. Bokde, Tiksha P. Kakade, Dnyaneshwari S. Tumsare and Chetan G. Wadhai Crime Analysis Using K-Means Clustering. (IJERT) ISSN: 2278-0181.IJERTV7IS040099 Vol. 7 Issue 04, April-(2018)

[25] Shahrzad Saremi, Seyedali Mirjalili and Andrew Lewis. Grasshopper Optimisation Algorithm: Theory and application. Advances in Engineering Software 105 (2017) 30–47.

[26] Seyedali Mirjalili. Moth-Flame Optimization Algorithm: A Novel Nature-inspired Heuristic Paradigm. Knowledge-Based Systems (2015).

[27] Ujjwal Maulik and Sanghamitra Bandyopadhyay, Genetic algorithm-based clustering technique. Pattern Recognition Society, 33 (2000) 1455-1465. Elsevier Science Ltd.

[28] Dheenathayalan, K. ., Savitha, K. . (2022). Integrating Bio-Inspired Algorithm with K-Means clustering to Analyze and Predict Crime against Women in India, Journal of Propulsion Technology.

[29] S. AshokKumar, K.P. Rajesh (2023) "Hyper-Parameters Activation on Machine Learning Algorithms to Improve the Recognition of Human Activities with IoT Sensor Dataset", Indian Journal of Science and Technology Vol. 16 Issue35 pp: 2856-2867. DOI: 10.17485/IJST/v16i35.882

[30] Arun, Bernard Edward Swamidoss, Venkatesan (2023), Impact of Hospitality Services on Tourism Industry in Coimbatore District, Journal of Namibian Studies - History Politics Culture, Volume 33, Special Issue 3, Pp. 2381-2393.

[31] Vijai, C., Bhuvaneswari, L., Sathyakala, S., Dhinakaran, D. P., Arun, R., & Lakshmi, M. R. (2023). The Effect of Fintech on Customer Satisfaction Level. Journal of Survey in Fisheries Sciences, 10(3S),6628-6634.

[32] R. Arun, M. Umamaheswari, A. Monica, K. Sivaperumal, Sundarapandiyan Natarajan and R. Mythily, "Effectiveness Performance of Bank Credit on the Event Management Firms in Tamilnadu State", In: Satyasai Jagannath Nanda and Rajendra Prasad Yadav (eds), Data Science and Intelligent Computing Techniques, SCRS, India, 2023, pp. 463-470. https://doi.org/10.56155/978-81-955020-2-8-42

[33] Singh, B., Dhinakaran, D. P., Vijai, C., Shajahan, U. S., Arun, R., & Lakshmi, M. R. (2023). Artificial Intelligence in Agriculture. Journal of Survey in Fisheries Sciences, 10(3S), 6601-6611.

[34] Mythili, Udhayakumar, Umamaheswari, Arun (2023) Factors Determining Mutual Fund Investments in Coimbatore City, European Chemical Bulleting, 12(special issue 6), 4719–4727.

[35] Arun, R. "A Study on the Performance of Major Spices in India." Recent Trends in Arts, Science, Engineering and Technology (2018): 149.

[36] K. Rani, Dr. J.Udhayakumar, Dr. M.Umamaheswari, Dr.R.Arun,(2023) "Factors Determining The Purchases of Clothing Products Through Social Media Advertisements in Coimbatore City", European Chemical Bulleting,12(special issue 6), 4728– 4737.

[37] Edson Nirmal Christopher, Sivakumar, Arun ,Umamaheswari (2023) Iiimmunoinformatic Study for a Peptide Basd Vaccine Against Rabies Lyssavirus Rabv Strain Pv, European Chemical Bulleting, 12(special issue 9), 631– 640.

[38] Arun (2019), "Sustainable Green Hotels -Awareness for Travelers", International Journal of Emerging Technologies and Innovative Research ISSN:2349-5162, Vol.6, Issue 4, page no. pp343-347,http://doi.one/10.1729/Journal.20408

[39] Buying behavior of meet's consumption relates to food safety from north and south part of the Coimbatore City. International Journal of Recent Technology and Engineering, 7, 429-433. https://www.ijrte.org/wp-content/uploads/papers/v7i5s/ES2177017519.pdf

[40] Chandramouli Shivaratri, Prakash, Arun, Krishna Mayi, Kavitha, Sivaperumal (2023), Clothing Products Purchases through Social Media Advertisements and the Problems Involved, Remittances Review, Vol. 8, Issue 4, Pp. 3260-3268.

[41] Akkur, S. A., R, R., S, S., P, D. K., Miryala, R. K., & Arun, R. (2023). Leadership Qualities Among Women Leaders in Educational Institutions at Bangalore City. International Journal of Professional Business Review, 8(9), e03772. https://doi.org/10.26668/businessreview/2023.v8i9.3772

[42] P, S., Prakash, K. C., Arun, R., C, N., Kousalya, M., & Sivaperumal, K. (2023). Green HRM Practices and the Factors Forcing it: A Study on Health Care Entities in Chennai. International Journal of Professional Business Review, 8(9), e03773.

[43] K. C. Prakash, R. Arun, Ram Chandra Kalluri, Souvik Banerjee, M R Vanithamani, Biswo Ranjan Mishra(2023), Consumer Confidence Index and Economic Growth- Indian Context after the Covid-19, European Economic Letters, Pp 746-754, DOI: https://doi.org/10.52783/eel.v13i5.824

[44] Arumugam, T., Arun, R., Natarajan, S., Thoti, K. K., Shanthi, P., & Kommuri, U. K. (2024). Unlocking the Power of Artificial Intelligence and Machine Learning in Transforming Marketing as We Know It. In S. Singh, S. Rajest, S. Hadoussa, A. Obaid, & R. Regin (Eds.), Data-Driven Intelligent Business Sustainability (pp. 60-74). IGI Global. https://doi.org/10.4018/979-8-3693-0049-7.ch005