Journal of Information Systems Engineering and Management
2025, 10(30s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Application of Meta-Optimization to Improve Big Mart Dataset Predictions

Ushveen Kaur
[1] Sri Guru Gobind Singh College of Commerce, University of Delhi. ushveenkaur@sggscc.ac.in

ARTICLE INFO

Received: 30 Dec 2024
Revised: 05 Feb 2025
Accepted: 25 Feb 2025

ABSTRACT

Multiple studies have been conducted to get the best classifier machine learning model, using various open source datasets. The dataset used in this study is Bigmart data set with 8524 samples. There are 17 variables in this dataset that can be used for the study. There are a lot of available studies performed on variables like health status variable, item type and store size variables. These studies have performed variations of machine learning models to get the best possible prediction. This study is a follow up study of the one done previously, where nine classifiers were applied on bigmart dataset and OneR classifier performed the best for predicting item type variable giving 84.46% accuracy. In this study, to improve the performance of the classification, four meta-heuristic optimizers, namely Elephant Herding Optimization, Monarch Butterfly Optimization, Harris Hawks Optimization and Slime Mould Algorithm are employed that assist hyper parameter tuning in a systematic way. It was concluded that Elephant herding optimization improved the OneR classification to 95.2181%. This increased accuracy is a significant improvement from the previous values.

Keywords: Elephant herding optimization, Bigmart, Optimizers.

INTRODUCTION

To improve the classification accuracy, the four meta-heuristic optimizers are employed that assists hyper parameter tuning in a systematic way. In the previous study, nine classifiers were applied on the Bigmart dataset to classify the variable item type. In that study, OneR gave the best performance and the required data and graphs are discussed in this paper. In this study, multiple iterations are performed for four optimizers, varying the variables used in each of the optimization models. The improved performance is then discussed in this paper.

LITERATURE SURVEY

The application of meta-optimization techniques in predictive modeling has gained significant attention due to their ability to enhance the performance of machine learning models. Meta-optimizers, such as Genetic Algorithms, Particle Swarm Optimization and Monarch Butterfly Optimization, are widely used to fine-tune hyperparameters and improve model accuracy. These techniques have been successfully applied across various domains, including retail, healthcare, and finance, to optimize prediction outcomes. A study emphasized that using optimize regression models improved sales predictions for e-commerce platforms [1]. Similarly, a study applied Genetic Algorithms to optimize feature selection in sales forecasting models, yielding higher accuracy compared to traditional approaches[2] . The BigMart dataset, commonly used for sales prediction and inventory management, presents an ideal testbed for applying meta-optimization techniques

EXPERIMENTAL MODEL

This section discussesthe four meta-heuristic optimizers in detail, the variables and parameters used in each one of them.

1.1 Meta-heuristic algorithms for optimization

Four optimization models namely, Elephant Herding Optimization, Monarch Butterfly Optimization, Harris Hawks Optimization and Slime Mould Algorithm are explained in the following sections.

3.1.1 Elephant Herding Optimization (EHO)

To create a four-elephant herding optimizer, elephant herding behavior was studied. Because elephants are herd animals, each herd is made up of a number of distinct clans. Each clan of elephants is led by a matriarch, and male elephants live alone until they are old enough to join the family group. To tackle optimization difficulties, elephants are housed in this environment. Once the population size has been established (the number of elephants and the number of clans), the elephant locations are updated using the "clan updating operator" and the "clan separating operator".

Clan Updating Operator

An elephant must be in peak physical condition to be considered a matriarch. The formula for updating the elephants in the clan, aside from the matriarch, is as follows:

$$X_{(newc_i, j)} = X_{(ci, j)} + \alpha \times (X_{(best\ c_i)} - X_{(c_i,\ j)}) \times r \tag{1.1}$$

where $X_{(newc_i,\ j)}$ and $X_{(c_i, j)}$ are the newly updated and old position for elephant j in clan c_i.

Clan Separating Operator

When male elephants reach the age of puberty, they will leave their family group and live on their own for the first time. Consider the elephants with the lowest fitness values and run the separation operator at each generation, as calculated in the equation below:

$$X_{(worstc_i)} = X_{(min)} + (X_{(max)} - X_{(min)} + 1) \times rand \tag{1.2}$$

where X_{max} and X_{min} are upper and lower bounds respectively. In this case, the upper and lower bounds of the location of the elephant individual, which are derived from the initialized parameters.

3.1.2 Monarch Butterfly Optimization (MBO)

Butterfly belongs to the type of Lepidoptera and over and above 18,000 species found all over the realm. Up to now, the three most significant optimizer approaches related to butterflies were anticipated, excluding monarch butterfly optimization. Through imitating search mechanism for earning food and breeding conduct of butterflies, butterfly optimizer was proposed first that was engaged for solving globe optimization and problems from engineering domain. Motivated through a breed-discovery style of a few butterfly classes, an artificial butterfly optimizer was presented then that offers a novel operative computation outline for resolving optimization difficulties. Like- wise, motivated through the breeding conduct of butterfly classes in the environment, a butterfly mating optimizer was proposed that divided into four stages to deploy a hunting course. Among these, the monarch butterfly optimizer turns out to be the greatest. The classic MBO seems to be modest and easier to deploy. In monarch butterfly optimizer, there exist two identical-size sub-populations, called as sub-population_1 and sub-population_2.Consequently, two plans are created in arch butterfly optimizers such as relocation operative and butterfly altering operative. At the end of the iteration, the globe optimum info is preserved, and two sub-populations are restructured into a populace once more. Afterwards, the complete populace is divided into two sub-populations as stated by the fresh fitness rate. This procedure is continued till the end state is encountered. Fig. 1 is the flowchart of the Monarch Butterfly Optimization.

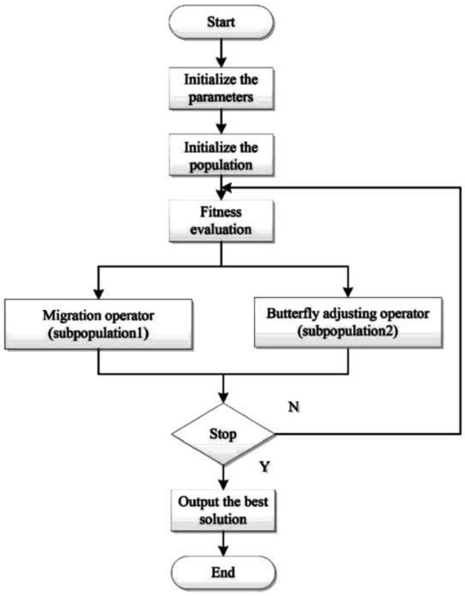

Fig. 1 Flowchart of the monarch butterfly optimization

3.1.3 Harris Hawks Optimization (HHO)

A renowned fowl of quarry - The Harris' Hawk, endures in fairly stable clusters and originate in south part of State Arizona, United States of America. Coordinated scavenging, including numerous faunae for holding and then distributing the killed fauna, have convincingly detected only for specific Mammalia flesh-eaters. The Harris's hawk is unique, owing to its novel supportive scavenging actions composed with another household associates existing in a similar steady cluster; however, another raptor typically bout to determine and hook a prey, unaccompanied.

The avian-desert hunter demonstrates advanced squad racing competencies in locating, enclosing, blushing out, and ultimately provoking the impending prey.

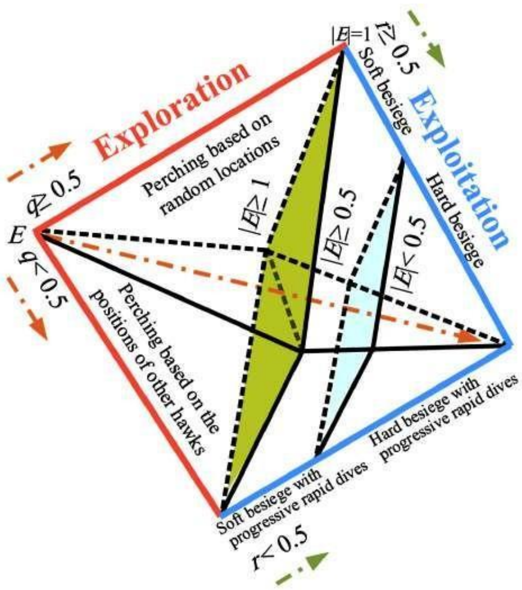

Fig. 2 Different stages of HHO

The smart-birds host banquet gatherings are containing numerous folks in the non-breed period and recognized as really supportive killers in a raptor kingdom. They initiate the squad job at pre-lunch sunset, with parting the rest perches and frequently resting on massive bushes or power-poles in the home kingdom. They are very well aware of their family associates and effort to be conscious of changes throughout the bout.

Once everyone is gathered and festivity starts, some hawks in a row initiate short trips and then place them on somewhat higher roosts. Like so, the hawk infrequently performs a 'leap-frog' gesture every wherein the aim place and rejoins to divide numerous periods to vigorously hunt for the hidden faunae that are typically rabbits. Fig. 2 highlights different stages of HHO.

3.1.4 Slime Mould Algorithm (SMA)

SMA is defined as an optimization technique founded on a nature-inspired population approach and the oscillatory fashion of slime mould. The slime mould optimizer offers an exclusive mathematics modelling that simulates negative and positive feedback of propagator waves of slime mould. The structure is dynamic with a steady equilibrium among local and global searching drift. While modelling the approach style of slime mould mathematically, the subsequent regulation is established to initiate to shrinkage phase.

$$Z(r+1) = Z_b(T) + V_b(Y.Z_A(T)) \quad : r < p (1.3)$$

$$V_c Z_r - x \quad : r \geq p$$

Where $V_b$is factor with the intermission of $| - m, m|$, $V_c$drops from 1 to 0in linear fashion. T signifies the present repetition,$Z_b$expresses the distinct location with the uppermost concentration of odour presently discovered, Z defines the position vector of SM, $Z_(A)$ and $Z_B$signifies two folks which have been arbitrarily designated from a present populace, Y signifies a weight of SM. The value is expressed as:

$$p = tan|s(i)) - DF|(1.4)$$

RESULTS AND DISCUSSIONS

Hyper parameter tuning and optimized classification results

Over the past decade, meta-heuristic algorithms have found of immense use in optimizing machine learning models for solving real-life engineering problems [3]. In a previous study, nine prediction models were applied on BigMart dataset for ten-fold cross validation. There were 16 classes (labels) in item types such as Dairy, Soft Drinks, Meat, Fruits and Vegetables, Household, Baking Goods, Snack Foods, Frozen Foods, Breakfast, Health and Hygiene, Hard Drinks, Canned, Starchy Foods, Breads, Seafood, and Others. Table 1 shows the classification of 16 item types designed using nine classifiers considering 10-fold cross-validation. The designed model using a training set was evaluated using a test set. The total Bigmart training datasets was consisted of 8524 samples. For classification using Training set, only 70% of the total Bigmart dataset was used and classification using Test set, remaining 30% dataset was used.

Table 1 Classification result with respect to item type labels

Classifier	10 fold cross validation	Training set	Test set
J48 decision tree	83.4944	88.8537	78.8983
Hoeffding tree	84.6912	86.9647	81.6023
Random forest	78.4087	99.8475	78.8256
Random Tree	77.8344	100	79.4552
REPTree	85.4538	86.5775	82.1107
Bayes net	76.8391	79.5025	74.2499
Naïve Bayes	66.5439	75.2083	62.0258
OneR	**85.4656**	86.5775	81.876
Decision Table	81.7462	86.1786	74.2108

Figure 3 clearly shows that for 10-fold cross-validation, the OneR classifier has executed the highest accuracy 84.4656%. The lowest accuracy 68.5439% was executed by the Naïve Bayes classifier. The accuracy ranges from 68.5439% to 84.4656%.

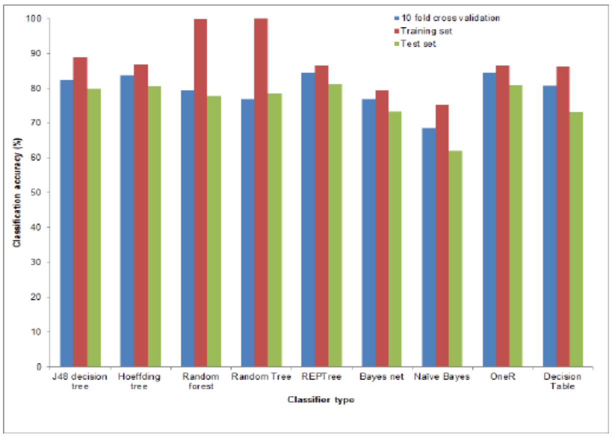

Fig. 3 Classification result with respect to item type labels

The confusion matrix for the classifier, which exhibits the highest accuracy for a particular classification model, is presented to understand classification and misclassification. Fig. 4 depicts the 3D confusion matrix for classification using an OneR classifier.

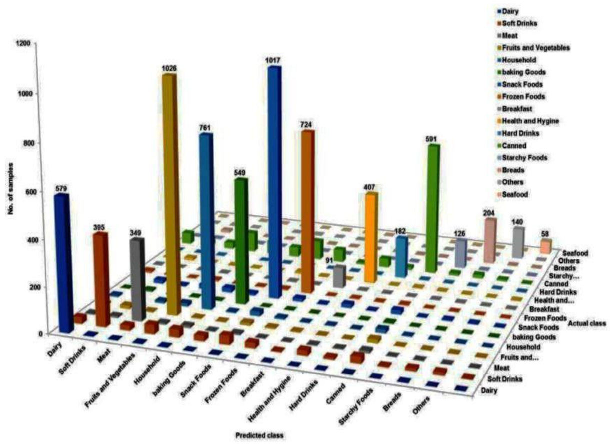

Fig. 4 Confusion matrix for classification using OneR classifier considering 10 fold cross validation

The results for optimized classification are presented here. First of all, classification results using an OneR classifier with respect to item type labels were optimized using Elephant Herding Optimization, Monarch Butterfly Optimization, Harris Hawks Optimization, Slime Mould Algorithm. The variable 'Itr' is the no. of instances initialized, and 'Epo' is the no. of epochs or generations the optimizer was tested on.

Table 2 Optimized classification result with respect to item type labels (OneR classifier)

Optimization algorithm	**Common constraints**	**Algorithm specific constraints**	**10 fold cross validation accuracy**	**Training set accuracy**	**Test set accuracy**
Elephant herding optimization	*Itr*= 42 *Epo*= 50	*clans* = 6 α = 0.5 β = 0.5	**94.2181**	100	85.4518
Monarch butterfly optimization	*Itr*= 16 *Epo*= 50	*p* = 5/12 *peri* = 1.2 *BAR*= 7/12 S_{max}= 50	91.3322	100	87.1212

Harris Hawks Optimization	*Itr*= 200 *Epo*= 50	-	92.8426	99.8151	82.5112
Slime Mould Algorithm	*Itr*= 100 *Epo*= 50	z = 0.1 *peri* = 1.2	92.1278	96.2185	83.4579

The number highlighted in the bold text, in table 2, show the highest accuracy for a particular classification mode.In table 2, the iterations for each optimizer were varied from 10 to 200 and epochs varied from 20 to 100. For elephant herding optimization, it was found that when 'Itr' was 42 and 'Epo' was 50, and clans were 6, alphas were 0.5 and beta was 0.5, accuracy was 95.2181% (10 fold cross-validation), 100% (Training set), 85.4518% (Test set). For monarch butterfly optimization, it was found that when 'Itr' was 16 and 'Epo' was 50, p was 5/12, peri was 1.2, BAR was 7/12 and Smaxwas 50, accuracy was 92.3322% (10 fold cross-validation), 100% (Training set), 87.1212% (Test set). In a comparison of 4 optimizers, elephant herding optimization yields the highest accuracy for a ten- fold cross-validation and training set. Owing to optimization studies, for the variable 'item type' the accuracy has increased from 84.46% to 95.2181%.

CONCLUSION

The 'Bigmart' dataset was considered and provided as an input and classification with respect to item type using nine machine learning classifiers namely J48 decision tree, HoeffdingTree, Random Forest, Random Tree, REPTree, Bayes Net, Naïve Bayes, OneR, and Decision Table. In the previous study, OneR gave the best classification. In this study, four optimization models were used and the accuracy for prediction of 'item type' increased drastically to 95.2181% from 84.46% using elephant herd optimization. The four metaheuristic optimizers were employed to improve the classification accuracy that systematically assists hyper parameter tuning. For future study, the parameters of the optimizers can be further tuned to check the possibility of improvement. Also, optimization models can be applied to other variables too of the Bigmart dataset.

REFERENCES

[1] Mohsine, S., Mzili, I., Riffi, M.E. (2024). Resolving the Quadratic Assignment Problem with the Elephant Herding Optimization. In: Ezziyyani, M., Kacprzyk, J., Balas, V.E. (eds) International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023). AI2SD 2023. Lecture Notes in Networks and Systems, vol 931. Springer, Cham.

[2] S. J. Isabella and S. Srinivasan, "An understanding of machine learning techniques in big data analytics: A survey," *Int. J. Eng. Technol.*, (2018), doi: 10.14419/ijet.v7i3.12.16450.

[3] Odegua, R. (2020). Applied Machine Learning for Supermarket Sales Prediction.https://www.researchgate.net/publication/338681895

[4] Tom, M., Raju, N., Isaac, A., James, J., & R, R. S. (2021). Supermarket Sales Prediction Using Regression. International Journal of Advanced Trends in Computer Science and Engineering, 10(2), 1153–1157.

[5] A. Rath, A. Das Gupta, V. Rohilla, A. Balyan, and S. Mann, "Intelligent Smart Waste Management Using Regression Analysis: An Empirical Study," in *Communications in Computer and Information Science*, (2022). doi: 10.1007/978-3-031-07012-9_12.

[6] A. H. Ali, M. Z. Abdullah, S. N. Abdul-Wahab, and M. Alsajri, "A Brief Review of Big Data Analytics Based on Machine Learning," *Iraqi J. Comput. Sci. Math.*, (2020), doi: 10.52866/ijcsm.2020.01.02.002.