Least-Square Support Vector Machine (LS-SVM) Parameters Optimization using Hybrid Cuckoo Search and Harmony Search Algorithm for Pre Collision Warning on Driver Assistance System (DAS)

A Sumarudin\textsuperscript{1}, A Puspaningrum\textsuperscript{1} and A Suheryadi\textsuperscript{1}
\textsuperscript{1}Politeknik Negeri Indramayu, Indonesia

Abstract. The growth of technology has triggered transportation sector to deliver various kind of advantages. One of them is pre collision warning. There are several steps in warning a collision, namely classifying and predicting the collisions then. Many supervised machine learning algorithms have been conducted, one of well known algorithm is Least Square – Support Vector Machine (LS-SVM). Radial Based Function (RBF) is one of LS-SVM kernels which is well-known method to support reliable performance. However, C and Gamma of its parameters are chosen randomly. This makes the performance of the classifier less optimal. So that, hybrid cuckoo search and harmony search algorithm is conducted to optimize LS-SVM parameters in this research. 8437 transportation records were used in the experiment and evaluated by using accuracy. Furthermore, the proposed method was also evaluated using several metaheuristic optimization algorithms namely Cuckoo Search Algorithm (CS-SVM), Bat Algorithm (BA-SVM), and Firefly Algorithm (FA-SVM). Experimental results show that hybrid Cuckoo Search and Harmony Search Algorithm (CSHS-SVM) successfully enhance the performance by reaching 84.513\% for accuracy, compared to Cuckoo Search Algorithm, Bat Algorithm, and Firefly Algorithm.

1. Introduction
The growth of technology has triggered transportation sector to deliver various kind of advantages. One of them is pre collision warning. It supports several features to reduce collisions such as visual and audio recommendation in order to assist driver to drive safely \cite{1}\cite{2}\cite{3}.

There are several steps in warning a collision, namely classifying and predicting the collisions then. Many supervised machine learning algorithms have been conducted, such as Least Square Support Vector Machine (LS-SVM), Random Forests (RF), and Artificial Neural Networks (ANN) as supervised machine learning algorithms \cite{4}\cite{5}\cite{6}. Beside that, one of well known classifying algorithm is Least Square – Support Vector Machine (LS-SVM). Radial Based Function (RBF) is one of LSSVM kernels which is well-known method to support reliable performance. There are two parameters of LS-SVM, namely: C and Gamma. However, C and Gamma of its parameters are chosen randomly. This makes the performance of the classifier less optimal.
Furthermore, there are many optimization research conducted using various nature-inspired metaheuristic algorithms to optimize the value of LS-SVM parameters, for example Cuckoo Search Algorithm, Particle Swarm Optimization, Genetic Algorithm, and Bat Algorithm [7][8][9]. On the other hand, in 2009, Cuckoo Search Algorithm (CSA) was developed by Yang and Deb to overcome optimization problems[10]. Besides that, Harmony Search (HS) is a metaheuristic algorithm which provide mutation operator which can increase exploitation ability if it is combined by Cuckoo Search Algorithm.

Moreover, hybrid CSA and HS has been conducted in several research before, for instance water distribution networks, software effort estimation, wireless sensor network, contrast enhancement of medical images [11][12][13][14].Thus, to optimize the value of C and Gamma parameters, this paper proposed hybrid cuckoo search algorithm and harmony search to get the optimal performance. The performance is evaluated by Accuracy then. After implementing the method to 8437 transportation records and finding the optimal performance, the parameters can be implemented widely to the classification process in order to mitigate the collisions.

The structure of research paper is organized as follows: Section 1 contains the introduction, Section 2 contains literature review, Section 3 explains the proposed method which has been done to improve the accuracy, result of the experiment is analyzed in section 4, and Section 5 presents the conclusion and future work of this research.

2. Related works
LS-SVM as one of intelligent technique with fast computation time has two random value within, namely: C and Gamma parameter. It can influence performance of the classification [15]. So that, these two parameters can be optimized by using optimization algorithm, such as Genetic Algorithm [8] which is implemented in 11 real world dataset of UCI repository. The experiment shows that the best performance of accuracy achieved in fewer features of classification.

Furthermore, Particle Swarm Optimization (PSO) Algorithm used the same dataset as [8] done. The research shows that Genetic Algorithm yields better performance compared to Particle Swarm Optimization. Beside that, Bat Algorithm were also used to optimize LS-SVM parameters [9]. Experimental result shows that Bat Algorithm can achieve optimal performance and has lower classification error compared to Particle Swarm Optimization and Genetic Algorithm. Moreover, Cuckoo Search Algorithm is also implemented to optimize LS-SVM parameters by using transportation dataset. It is compared to Bat Algorithm and Firefly Algorithm. Experimental result shows that Cuckoo Search Algorithm can achieve best performance compared to benchmarking methods.

From the literature above, most of those researches attempt to achieve best optimization performance. To the best of knowledge, the proposed method of this paper is the only research which attempts to optimize the parameter values of LS-SVM by using hybrid optimization methods and implemented in real transportation dataset. The dataset implemented consists of six attributes as [8] recommended to use fewer features in the previous research.

3. Methodology
To achieve better performance, the proposed method combines Cuckoo Search and Harmony Search. In optimizing LS-SVM parameters, CSA is implemented in the first step to find the optimal performance, and harmony search is used in the next step. CSHS-SVM was implemented in a PC with the following features: Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz, 16G RAM, a Windows 10 Operating System, and Matlab 2014a. There are 8437 records used in this experiment. 4907 is labeled as not danger class, and 3530 is labeled as danger class.

As shown as in Table 1, the label of the classification consists of two label, danger and not danger. The experimental data used in this research consists of six attributes, namely id, speed, distance, period, warning, pair, label. Besides that, accuracy, parameter C and Gamma are the
output of the program. K-fold cross-validation has also been used in all experiments. The value of k was set to 10. The dataset then divided into training and testing data.

![CSHS-SVM block diagram](image)

**Figure 1.** CSHS-SVM block diagram

Proposed method in this research actually consists of two main process, namely: optimization and classification steps as shown in Figure 1. In the optimization steps there are two optimization algorithm used in this research. Cuckoo search is used in the first step to find the best solution ($X_{global}$). Then, harmony search is also used in the second stage to create another solutions ($HM_{best}$). After that, the solution from harmony search is compared to $X_{global}$. If it gets a better fitness, then $HM_{best}$ will substitutes the value of $X_{global}$. The process will be reiterated as many as the number of iteration. After acquiring the best optimization parameters, those parameters will be used in classification process.

| Parameters           | Values          |
|----------------------|-----------------|
| Number of Cuckoo     | 5               |
| Number of Eggs       | 15              |
| Iteration            | 2               |
| Position             | Minimal = 50    |
|                      | Maximal = 125   |
| Discovery Rate (Pa)  | 0.25            |
| Memory Consideration | 0.95            |
| (HMCR)               |                 |
| Pitch Adjustment Rate| 0.5             |
| (PAR)                |                 |
| Pitch Adjustment Bandwith | 0.1        |
| (BW)                 |                 |

Table 1 shows the parameters used in the proposed method. Then, accuracy will be used as evaluation method and fitness value measurement of CSHS-SVM. The accuracy formula is shown in Equation 1

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$ (1)

Table 1. Parameters of CSA-SVM
According to Equation 1, True Positive (TP) illustrates that the danger label which correctly classified as danger. But, when the danger label is classified as not label, it will be classified as False Negative (FN). In addition, when 'not danger' label is correctly classified as 'not danger', the record will be classified as True Negative (TN). However, when 'not danger' is classifies as 'danger', it will be classified as False Positive (FP).

4. Experimental Result

In this section, the proposed method has been implemented for LS-SVM parameters optimization. In addition, the Cuckoo Search and Harmony Search (CSHS-SVM) is then compared to other recent metaheuristic methods such as Cuckoo Search Algorithm (CSA-SVM), Bat Algorithm (BA-SVM) and Firefly Algorithm (FA-SVM) as benchmarking methods. We considered 8437 records to be tested in five experiments. Each experiment tests performance of each methods.

![Figure 2. Average performance result]

According to Figure 2, the average value of CSHS-SVM has reached 84.513% as the best performance compared to CSA-SVM, BA-SVM and FA-SVM. The average performance is achieved based on five experiments which have been done. The detail performance for each experiment is described in Table 2.

| Experiment | BA - SVM | FA - SVM | CSA - SVM | CSHS-SVM |
|------------|----------|----------|-----------|-----------|
| 1          | 77.657   | 83.217   | 82.659    | 84.651    |
| 2          | 78.726   | 83.205   | 83.655    | 84.461    |
| 3          | 76.034   | 83.229   | 83.561    | 84.473    |
| 4          | 81.368   | 83.146   | 84.544    | 84.438    |
| 5          | 79.447   | 83.146   | 84.817    | 84.544    |
| Average    | 78.646   | 83.189   | 83.847    | 84.513    |

Table 2 shows the result of the experiment among four algorithms. CSHS-SVM yields the best performance among other algorithms. It illustrates that the proposed method can generate more optimal parameters. So that, the accuracy value of the classification can be increased compared to cuckoo algorithm performance itself. The parameters optimized in this research are C and Gamma which will be used for the classification step next. The result of the optimization parameters are shown in Table 3.
Table 3. Parameters performance of CSA-SVM

| Experiment | Gamma   | C       | CSA – SVM |
|------------|---------|---------|-----------|
| 1          | 50.230  | 101.078 | 84.651    |
| 2          | 59.524  | 118.503 | 84.461    |
| 3          | 72.049  | 125.000 | 84.473    |
| 4          | 67.552  | 125.000 | 84.438    |
| 5          | 62.410  | 114.971 | 84.544    |

Table 4 shows the best optimized CSA-SVM parameters in traffic records dataset.

Table 4. Optimized parameters of CSA

| Coefficient | Value   |
|-------------|---------|
| C           | 50.230  |
| Gamma       | 101.078 |

Table 4 shows the best optimized CSA-SVM parameters in traffic records dataset. To sum up, there are two parameters optimized in this research namely C and Gamma. The parameters become parameters to change the parameters which is chosen randomly. But when the parameters is configured for all fold validation, the classification cannot give an optimal result even the proposed method got the best performance compared to other benchmarking methods. Furthermore, there are also some influencing parameters in this experiment. So that, the performance is highly controlled by those parameters, for example the quality of the data, optimization algorithm parameters, and number of data.

5. Conclusion and further work
A hybrid Cuckoo Search and Harmony Search Algorithm is proposed for optimizing LS-SVM parameters. Experimental result has done by using 8437 transportation records to verify the proposed method’s accuracy. The proposed method yields the best performance for each of the 10-folds compared to the other methods by reaching 84.513% for accuracy, compared to Cuckoo Search, Bat Algorithm and Firefly Algorithm.
References

[1] Ian J Reagan, Jessica B Cicchino, Laura B Kerfoot, and Rebecca A Weast. Crash avoidance and driver assistance technologies–are they used? *Transportation research part F: traffic psychology and behaviour*, 52:176–190, 2018.

[2] Angela H Eichelberger and Anne T McCartt. Toyota drivers’ experiences with dynamic radar cruise control, pre-collision system, and lane-keeping assist. *Journal of safety research*, 56:67–73, 2016.

[3] Jianqiang Wang, Jian Wu, Xunjia Zheng, Daiheng Ni, and Keqiang Li. Driving safety field theory modeling and its application in pre-collision warning system. *Transportation research part C: emerging technologies*, 72:306–324, 2016.

[4] Nejdet Dogru and Abdulhamit Subasi. Traffic accident detection using random forest classifier. In *2018 15th learning and technology conference (LeT)*, pages 40–45. IEEE, 2018.

[5] Samuel L Gomes, Elizângela de S Rebouç as, Edson Cavalcanti Neto, João P Papa, Victor HC de Albuquerque, Pedro P Rebouç as Filho, and Joao Manuel RS Tavares. Embedded real-time speed limit sign recognition using image processing and machine learning techniques. *Neural Computing and Applications*, 28(1):573–584, 2017.

[6] Hyun Soo Park, Dae Jung Kim, Chang Mook Kang, Seok Cheol Kee, and Chung Choo Chung. Object detection in adaptive cruise control using multi-class support vector machine. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6. IEEE, 2017.

[7] Shih-Wei Lin, Kuo-Ching Ying, Shih-Chieh Chen, and Zue-Jung Lee. Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert systems with applications*, 35(4):1817–1824, 2008.

[8] Cheng-Lung Huang and Chiieh-Jen Wang. A ga-based feature selection and parameters optimization for support vector machines. *Expert Systems with applications*, 31(2):231–240, 2006.

[9] Alaa Tharwat, Aboul Ella Hassanien, and Basem E Elmaghl. A ba-based algorithm for parameter optimization of support vector machine. *Pattern Recognition Letters*, 93:13–22, 2017.

[10] Xin-She Yang and Suash Deb. Cuckoo search: recent advances and applications. *Neural Computing and Applications*, 24(1):169–174, 2014.

[11] Madhuri Naveen Naidu, Pankaj Srimal Boindala, A Vasan, and Murari RR Varma. Optimization of water distribution networks using cuckoo search algorithm. In *Advanced Engineering Optimization Through Intelligent Techniques*, pages 67–74. Springer, 2020.

[12] I Made Mika Parwita, Riyanto Sarno, and Alifia Puspaningrum. Optimization of cocomo ii coefficients using cuckoo optimization algorithm to improve the accuracy of effort estimation. In *2017 11th International Conference on Information & Communication Technology and System (ICTS)*, pages 99–104. IEEE, 2017.

[13] Md Akhtaruzzaman Adnan, MA Razzaque, Md Anowarul Abedin, SM Salim Reza, and Molla Rashied Hussein. A novel cuckoo search based clustering algorithm for wireless sensor networks. In *Advanced Computer and Communication Engineering Technology*, pages 621–634. Springer, 2016.

[14] Ebenezer Daniel and J Anitha. Optimum wavelet based masking for the contrast enhancement of medical images using enhanced cuckoo search algorithm. *Computers in biology and medicine*, 71:149–155, 2016.

[15] Siuly Siuly and Yan Li. Improving the separability of motor imagery eeg signals using a cross correlation-based least square support vector machine for brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(4):526–538, 2012.

Acknowledgments

This research was fully supported by Pusat Penelitian dan Pengabdian Masyarakat, Politeknik Negeri Indramayu as Penelitian Terapan Dana Lokal Polindra 2019.