An Improved Particle Swarm Optimization for Prediction of Accident Severity

Swarnima Singh[1] and Vikash Yadav[2]

[1]*ABES Engineering College, Ghaziabad, U.P, India*
[2]*Department of Technical Education, Uttar Pradesh, India*

*Corresponding Author: Vikash Yadav; Email: vikas.yadav.cs@gmail.com

▓ **ABSTRACT:** In recent time, road traffic accidents are increased day by day due to incremental growth in number of vehicles. Traffic accidents are also the crucial and important issue for achieving the sustainable transportation development. One of the tasks of sustainable transportation is to reduce the number of accidents and also design the traffic assessment policy. Many researchers consider the traffic accident issue of sustainable transportation and developed prediction models for measuring severity of accidents. But the accuracy of accident severity is one of the major issues. In this work, an attempt is made to improve the accuracy of accident severity. To achieve the same, a particle swarm optimization-based algorithm is applied for evaluating the accident severity. Prior to implement the PSO, two modification are incorporated into PSO algorithm, called improved PSO. These modifications can be described as mutation operator and trail candidate generation. The performance of improved PSO is examined over accident traffic severity dataset and results are evaluated using accuracy, recall@5 and precision@5 metrics. Several existing techniques are considered for comparing the results of IPSO algorithm. It is revealed that IPSO achieves more accurate results among all techniques.

Keywords: Accident Severity, Particle Swarm Optimization, Road Traffic, Traffic Accident and Assessment

▓ 1. INTRODUCTION

In the era of sustainable transportation development, one of the important issues is traffic safety assessment and accident prediction [1]. Traffic accidents have adverse impact on the economy of countries and also lead to traffic jam, responsible for environmental pollution, causalities and sometimes human death. For achieving the sustainable and healthy transportation, the rate of traffic accident should be reduced as well as design the traffic safety policy. On the other side, it is reported that there is an incremental growth in the field of automated information system and these systems provides optimized solution for many real-life problems. In recent time, several agencies and government sectors try to develop the intelligent transportation system [2]. But, one of crucial and challenging task for sustainable development of transportation is traffic accident prediction. By using the traffic accident prediction, the traffic safety assessment policy can be designed on the basis of existing road accident data. It is also important to determine the frequency of accidents and hotspot location for accidents. It is also ensuring that the timely and accurate prediction of accidents severity can help for achieving better traffic safety, management and control. Recent time, several advance technologies can be inculcated in the vehicles for reducing the accidents and also incorporated some traffic safety assessment. According to the one of the studies conducted by WHO in 2017, more than 1.2 million can loss their life due to accidents and more than 50 million can have non-fatal injuries and traffic accident is ninth cause of death worldwide [3]. The traffic accidents can be occurred at any time in a day, but these can be predicted and prevented. The main task of the researcher worked in the field of traffic accident and assessment is to reduce the traffic accidents and also determine the factors that can responsible for accidents. Several researchers explore the relation among severity of accidents and influential factors [4-5]. It is reported that earlier studies focused on logit and Probit models for evaluating the accident severity [6]. Recent studies also focused on applicability of intelligent classifiers for accident severity prediction [7-9]. Such classifiers can be named as decision tree, SVM, random forest and neural network etc. [10-11].

1.1 Contribution of the Work

The aim of this research work is to predict the accident severity. For prediction the severity of accident, a particle swarm optimization (PSO) based algorithm is adopted. PSO is well known meta-heuristic than can be applied for solving the wide variety of optimization problems [12-13]. Prior to implement the PSO algorithm, two modifications are incorporated in PSO algorithm, called improved PSO to make it more efficient and also balancing the search processes for getting the optimal solutions. These modifications can be described in terms of mutation operator and trial candidate generation. The accident severity data is taken from the kaggle repository and having sixty-two attributes with four classes.

The highlights of the work are mentioned as:
1. Severity of Accidents is predicted using an improved PSO algorithm.

2. Prior to implement the PSO, two modifications in terms of mutation operator and trial candidate generation is proposed.

3. Aim of mutation operator is to balance exploration and exploitation phases of PSO, whereas, trail candidate guides the search towards the global optimal solution.

4. The performance of proposed improved PSO is evaluated using accuracy, recall and precision metrics.

5. The well-known road accident dataset is considered for assessing the performance of improved PSO.

The rest of paper can be structured as. *Section 2* summarizes the related works of accident severity and prediction. The improved PSO algorithm and its modifications are discussed in *section* 3. The simulation results are discussed into *section 4*. Finally, the entire work is concluded into *section 5*.

2. RELATED WORKS
This section presents the related works in the direction of accident prediction and prevention.

Xiong et al. [14] presented a framework for chain of road traffic accident, called CRTI. The proposed framework extracts the features form the movement of vehicles. Further, CRTI contains a two-stage modelling method based on scenario strategy. In first stage, a SVM is applied for classifying the vehicles leaving the own lane and entering into another lane. In second stage, a hidden MARKOV model with Gaussian mixture is developed for determining the accident and non-accident. It is claimed that proposed accident framework obtains promising results.

Zheng et al. [15] developed an accident prevention and risk assessment framework to determine the risk of accident and prevention among intelligent vehicles. The proposed framework implements the hidden MARKOV model. The HMM model is used to predict the probabilities of steering angle and also to compute risk of road traffic. The proposed framework is validated using accuracy parameter. Authors claimed that proposed framework achieves more accurate results.

Accuracy is one of the primary concerns in traffic flow prediction because it is an important factor for improving the traffic capacity. An et al. [16] presented a fuzzy based convolutional neural network, called F-CNN for predicting the flow of traffic more accurately. In F-CNN, fuzzy approach is used to identify the traffic accident features and CNN is applied to measure the uncertain traffic accident information. Further, the F-CNN model can be used to train the internal information. The proposed framework is implemented on Beijing taxi can trajectory. It is noticed that proposed framework provides superior results.

Many existing studies considered the single factor for traffic prediction. Wang and Chen [17] explored multiple factors for predicting the traffic. Authors consider the real time traffic scenario for traffic prediction using onboard camera installed on vehicles. Further, a deep learning model based on spatial temporal method and convolutional network, called VSTGC is adopted for determining the spatial temporal structure and predicting the trends for safety in future. The simulation results showed that proposed VSTGC model provides competitive results using real time traffic scenario.

Prasetijo et al. [18] considered the road accident and hazard location issues and developed a prediction model for predicting the road accidents. The proposed model considers the speed, global positioning system and operating profile for determining the speed and car. Further, these characteristics are considered to design the prediction model. It is revealed that proposed prediction model have significant impact for predicting the road accidents.

Galatioto et al. [19] presented a web-based model for ensuring the road safety and also preventing the road accidents. The proposed web-based model consists of machine learning techniques and econometric method for predicting severity of accidents. The simulation results are compared with several existing statistical and ML techniques. It is found that proposed web-based model provides superior results for predicting road accidents due to parametric approach.

Gianfranco et al. [20] considered the urban road network for predicting road accidents. The motive of work is to correlate the accident frequency with road infrastructure and further, can be used for road safety in term of planning and designing. The proposed model is based on the combination of poissions and negative binomial algorithms. It is also seen that three situations can be considered for urban road network. These situations are roundabout, straight road and 3 to 4 way junction. Results reported that proposed model is capable to determine the accident hot spot.

In sustainable transportation system, traffic accident prediction and safety assessment are an important factor. Gan et al. [21] considered the aforementioned issue of sustainable transportation system and developed a prediction model by considering the deep forest algorithm. Several other ML techniques are also adopted for comparing the simulation results of deep forest technique. The findings show the significance of the work and also beneficial for designing the road safety curricula.

Road accidents can be responsible for high intensity injuries and sometimes lead to threaten the human life. Parvareh et al [22] considered the road accident as one of the research problems and tries to characterize and measure the frequency of road accidents. To conduct the above-mentioned study, a time series data can be collected regarding the road accidents. Further, accident is divided into three classes i.e. motorcyclist, car occupant and pedestrian injuries. The study revealed that pedestrian and motorcyclist injuries can be seasonal due to air temperature.

For the effective management of urban traffic, traffic accident prediction is one of major concern. Zhou [23] stated that

existing studies considered the classical ML approaches for accident prediction and failed due to involvement of cross domain knowledge. In recent time, due availability of cross domain data and also presence of more viable and efficient techniques such as deep learning, it is quite possible to predict the traffic accident more accurate with cross domain knowledge. Hence, author presented an accident framework based on Resnet to correlate the traffic accident with urban traffic data.

The motive of crash prediction model is to measure either severity of crash or unknown severity. Iranitalab and Khattak [24] adopted the four well known statistical and ML methods for evaluating the crash severity. These methods are multinomial logit, nearest neighbor classification, SVM and random forest. The objective of the work is described in 3-fold. First is to predict the severity of traffic crash using the aforementioned methods. Second is to develop a crash predictive model with low cost for predicting traffic crash. Third is to investigate the impact of clustering methods on the traffic crash prediction. The performance of crash severity prediction model is evaluated using accuracy metric. The random forest and SVM are outperform than MNL and NNC methods. Further, it is also noticed that clustering algorithms are not improved the results of SVM method. But, K-means clustering can improve the results of random forest, NNC and MNL methods in significant manner.

Singh et al. [25] presented a deep learning-based prediction model for accident prediction. For this study, eight non-urban highway is selected for collecting the accident related data. The final processed dataset contains sixteen variables with two hundred twenty data points. The results of deep learning model are compared with GEP and RENB. The RMSE metric is used for evaluating the performance of proposed model. Authors claimed that deep learning model achieves better accident frequency rate and determine the appropriate variable for traffic accidents.

The prediction of traffic accident is tough task due to spatial information, dynamic interaction and external factor like heterogeneous data. Yu et al. [26] considered the aforementioned factors and developed an accident prediction model based on spatio-temporal graph and convolutional network. The proposed prediction model consists of three components. These components are spatial learning layer, spatio-convolutional layer and embedding layer. These layers perform the specific tasks like operation on spatial information, dynamic variations and semantic information. Results showed that proposed model provides superior results than classical and other existing techniques.

To reduce the number of accidents and also forecast the future accidents, Jomnonkwao [27] adopted four methods such time series analysis, regression analysis, curve estimation and path analysis for forecasting of accidents. The simulation results are evaluated using mean absolute error. Results showed that regression model achieves lower mean absolute error as compared to other methods.

3. PROPOSED IMPROVED PSO ALGORITHM

This section describes the proposed improved PSO algorithm that can be adopted for detection of accident severity. In literature, it is claimed that performance of meta-heuristic algorithm is highly influenced through exploration and exploitation process [28]. Further, for obtaining the good optimal solutions, these processes should be balanced. In order to make the algorithm more robust and efficient, a mutation operator is inculcated in PSO algorithm, called an improved PSO algorithm. The description of mutation operator is described below.

$$\text{Mut}(X_{i,new}) = X_{i,old} + F(L_{i,best} - X_{i,old}) \tag{1}$$

In above equation, $\text{Mut}(X\_new)$ represents the new mutated position of ith particle, $X\_{(i,old)}$ represents the old position of ith particle, F is an amplitude matrix in the range of [0,1] and $L\_{(i, best)}$ is the local best position in the neighborhood of ith particle.

$$T_i = X_i + A(\beta \times G_{best} - X_i) \tag{2}$$

In above equation, T_i represent the trial position of ith particle, X_i represents the ith position of particle, A can be given as $2 \times rand()$, β is random variable in the range of [0,1] and G_best is the global best position of particle in i^{th} iteration.

The aforementioned equations are used to update the population of PSO algorithm as well as ensure the optimum balance between exploration and exploitation processes. The computational procedure of improved PSO algorithm for accident severity prediction is mentioned in Algorithm 1 and the flowchart of improved PSO mentioned in *Figure 1*.

| Algorithm 1: Computational procedure improved PSO for accident severity prediction. |
| --- |
| **Input:** Accident severity dataset (D), particles population (P), and severity classes(C) |
| **Output:** Severity Prediction |
| **Step 1:** Initialize the accident severity dataset, population of particles (P), initial velocity (v), inertia weight (w), severity classes (C), maximum Iteration, mutation probability (PM) etc. |
| **Step 2:** Randomly initialize the population of particles. |
| **Step 3:** While(termination condition is not met), do following |
| **Step 4:** Compute the fitness of each particle using the equation $$MSE = \sum_{i=1}^{N}(y_i - \hat{y}_i) \tag{3}$$ In equation 3, y_i represents he actual output and $\hat{y}_i$ represents the predicted output. |
| **Step 5:** Determine local best (L_{best}) and global best (G_{best}) position of particles. |
| **Step 6:** Determine the new velocity and position of particles using following equations. $$V_{p,new} = w \times V_{p,old} + r_1(pbest_p - X_{p,old}) + r_2(gbest - X_{p,old}) \tag{4}$$ $$X_{p,new} = X_{p,old} + V_{p,new} \tag{5}$$ |
| **Step 7:** Compute the trial position of particles using equation 2. |
| **Step 8:** If$(rand > PM)$ Update the trial position using equation 1. |
| **Step 9:** Compute the fitness of trail position and update the L_{best} and G_{best}. |
| **Step 10:** Repeat the steps 4-9 until severity classes is not predicted for all instances. |
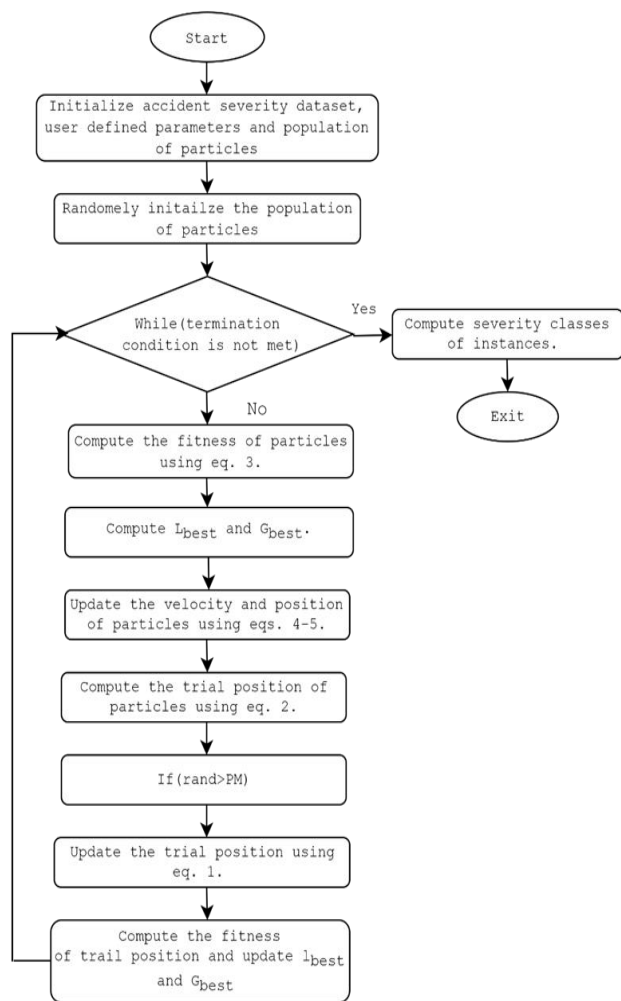
Figure 1: Flowchart of proposed improved PSO

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section discusses the simulation results of improved PSO algorithm using accident severity dataset. The main motive of this work is to predict the severity of accidents. The accident severity dataset contains two lakhs eighty-five thousand three hundred thirty-one and having sixty-two attributes. The dataset is downloaded from Kaggle repository. The accident severity into four classes i.e. "0" stands for no severity, "1" stands for low severity level, "2" stands for medium severity level and "3" stands for high severity level. Further, the performance of improved PSO is examined using accuracy, recall and precision metrics. Moreover, simulation results of IPSO techniques are evaluated using 5-cross fold method. In this method, the entire dataset is divided into five equal parts. Out of five, four parts acts as training data and one-part acts as test data. It is observed that every time, different test data can be fed to model. Further, several existing algorithms are considered for comparing simulation results of improved PSO algorithm. These algorithms are naive bayes (NB), support vector machine (SVM), stacking, genetic algorithm (GA) and particle swarm optimization (PSO). *Table I* depicts the

simulation results of improved PSO and other state of art existing techniques using accuracy, precision and recall metrics. It is noticed that proposed IPSO algorithm achieves higher accuracy rate (84.63) as compared to another existing algorithm. It is also observed that NB exhibits lower accurate rate among all techniques. On the analysis of precision and recall metrics, it is seen that proposed IPSO algorithm higher precision and recall rates as compared to NB, SVM, stacking, GA and PSO techniques. While, NB technique obtains lower precision and recall rates among all technique. Hence, it is stated that proposed IPSO algorithm can detect the severity of accident more accurately than other techniques.

Table 1: Simulation results of severity prediction using proposed IPSO algorithm and existing techniques on accident severity dataset

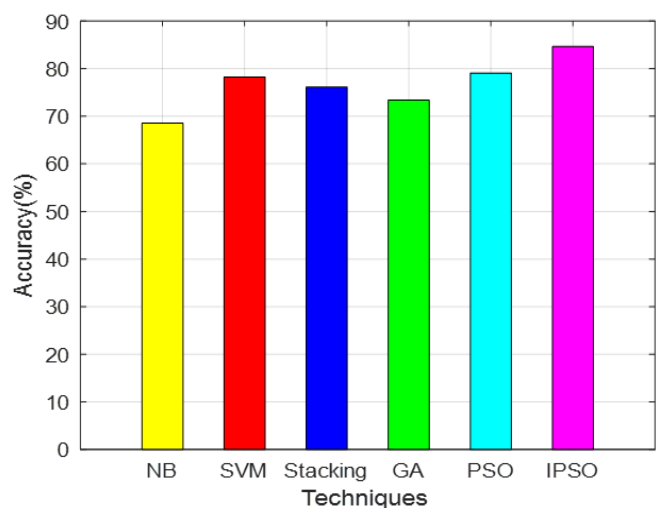| Technique	Performance Metric		
	Accuracy	Precision	Recall
NB	68.53	64.91	65.06
SVM	78.26	73.02	74.29
Stacking	76.14	75.54	74.83
GA	73.39	71.87	72.11
PSO	79.05	77.94	77.68
IPSO	84.63	81.26	83.08

Figure 2: Illustration of accuracy metric for accident severity dataset using different techniques.

Apart to the tabular comparison, the graphical representation of results is also reported using *Figures 2-4*. *Figure 2* illustrates the graphical representation of accuracy metric using accident severity dataset. It is cleared that IPSO having better accuracy results than other techniques. It is also stated that proposed improvements enhance the accuracy rate of PSO technique. The results of precision metric are demonstrated in *Figure 3*. Precision is an important parameter to evaluate the

performance techniques. It is revealed that proposed IPSO technique again achieves higher precision rate than other techniques. From literature, it is concluded that both precision and recall can be used to assess performance of techniques. It is also observed that both metrics should be obtained uniform results i.e. there can be small difference among the results of both parameters. *Figure 4* reports the results of all techniques using recall metric. It is seen that IPSO technique obtains higher recall rate than other techniques. Further, it is also analyzed that both precision and recall metrics having small difference in the simulation results.
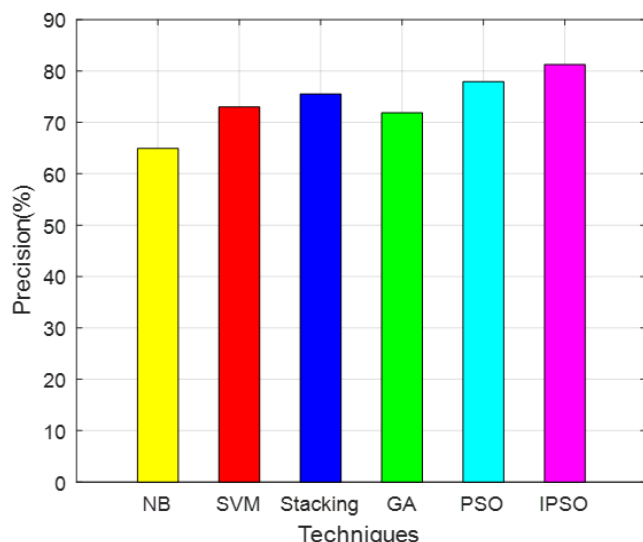

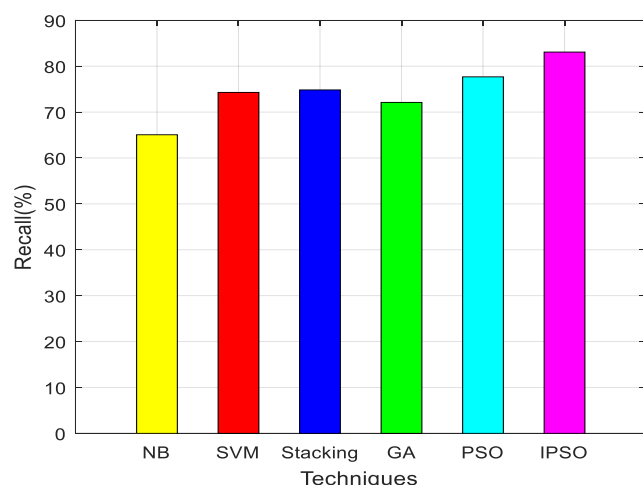**Figure 3**: Illustration of precision metric for accident severity dataset using different techniques.

Figure 4: Illustration of recall metric for accident severity dataset using different techniques.

5. CONCLUSION

This paper presents an improved PSO algorithm for predicting the severity of accidents using the accident severity dataset. Prior to implement the improved PSO algorithm, two modifications are inculcated into PSO algorithm. The motive of these modifications is to balance the exploration and exploitation process. Firstly, a mutation operator is proposed for enhancing the balance among search processes. In second modification, a trial position of particles can be generated. The aim to generate the trial position of particles is directing the search towards the global optimum direction. The performance of IPSO is examined over accident severity dataset and several metrics are considered for evaluating the simulation results. These metrics are accuracy, recall and precision. Moreover, several existing classifiers can be used for comparing the simulation results of proposed IPSO technique. From simulation results, it is concluded that IPSO achieves better results in terms of accuracy, precision and recall. It is also observed that proposed modifications also improve the severity rate of PSO algorithm in significant manner. In future, other meta-heuristic algorithms like ABC, Cuckoo Search BAT will be considered for evaluating the severity of accidents. It is noticed that feature engineering also plays an important role for predicting the severity level. So, feature selection algorithms can be considered for determine relevant features for severity prediction.

REFERENCES

[1] MENG, X. H., ZHENG, L., & QIN, G. M. (2009). Traffic Accidents Prediction and Prominent Influencing Factors Analysis Based on Fuzzy Logic [J]. Journal of transportation systems engineering and information Technology, 2, 015.

[2] Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis & Prevention, 43(5), 1666-1676.

[3] World Health Organization (WHO) (2017), A Road Safety Technical Package, World Health Organization, Geneva, Switzerland.

[4] De Oña, J., López, G., Mujalli, R., & Calvo, F. J. (2013). Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. Accident Analysis & Prevention, 51, 1-10.

[5] Dadashova, B., Ramírez, B. A., McWilliams, J. M., & Izquierdo, F. A. (2016). The identification of patterns of interurban road accident frequency and severity using road geometry and traffic indicators. Transportation research procedia, 14, 4122-4129.

[6] Yan, Z., Lu, X., & Hu, W. (2020). Analysis of factors affecting traffic accident severity based on heteroskedasticity ordinal Logit. In ICTE 2019 (pp. 422-435). Reston, VA: American Society of Civil Engineers.

[7] Mujalli, R. O., López, G., & Garach, L. (2016). Bayes classifiers for imbalanced traffic accidents datasets. Accident Analysis & Prevention, 88, 37-51.

[8] García de Soto, B., Bumbacher, A., Deublein, M., & Adey, B. T. (2018). Predicting road traffic accidents using artificial neural network models. Infrastructure Asset Management, 5(4), 132-144.

[9] Zhang, X. F., & Fan, L. (2013, May). A decision tree approach for traffic accident analysis of Saskatchewan highways. In 2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1-4). IEEE.

[10] Pu, Z., Li, Z., Jiang, Y., & Wang, Y. (2020). Full Bayesian before-after analysis of safety effects of variable speed limit system. IEEE transactions on intelligent transportation systems.

[11] Zhang, J., Li, Z., Pu, Z., & Xu, C. (2018). Comparing prediction performance for crash injury severity among various machine learning and statistical methods. IEEE Access, 6, 60079-60087.

[12] Gupta, I., & Yadav, V., (2019). Medical Data Clustering based on Particle Swarm Optimization and Genetic Algorithm. International Journal of Advanced Intelligence Paradigms, 14(3/4), 345-358.

[13] Kumar, Y., & Sahoo, G. (2015). Hybridization of magnetic charge system search and particle swarm optimization for efficient data clustering using neighborhood search strategy. Soft Computing, 19(12), 3621-3645.

[14] Xiong, X., Chen, L., & Liang, J. (2017). A new framework of vehicle collision prediction by combining SVM and HMM. IEEE Transactions on Intelligent Transportation Systems, 19(3), 699-710.

[15] Zheng, X., Zhang, D., Gao, H., Zhao, Z., Huang, H., & Wang, J. (2018). A novel framework for road traffic risk assessment with HMM-based prediction model. Sensors, 18(12), 4313.

[16] An, J., Fu, L., Hu, M., Chen, W., & Zhan, J. (2019). A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information. Ieee Access, 7, 20708-20722.

[17] Wang, J., & Chen, Q. (2021). A traffic prediction model based on multiple factors. The Journal of Supercomputing, 77(3), 2928-2960.

[18] Prasetijo, J. (2020). Accident prediction model by relating integrated design consistency model. International journal of integrated engineering, 12(1), 260-269.

[19] Galatioto, F., Catalano, M., Shaikh, N., McCormick, E., & Johnston, R. (2018). Advanced accident prediction models and impacts assessment. IET Intelligent Transport Systems, 12(9), 1131-1141.

[20] Gianfranco, F., Soddu, S., & Fadda, P. (2018). An accident prediction model for urban road networks. Journal of Transportation Safety & Security, 10(4), 387-405.

[21] Gan, J., Li, L., Zhang, D., Yi, Z., & Xiang, Q. (2020). An alternative method for traffic accident severity prediction: using deep forests algorithm. Journal of advanced transportation, 2020.

[22] Parvareh, M., Karimi, A., Rezaei, S., Woldemichael, A., Nili, S., Nouri, B., & Nasab, N. E. (2018). Assessment and prediction of road accident injuries trend using time-series models in Kurdistan. Burns & trauma, 6.

[23] Zhou, Z. (2019, June). Attention based stack resnet for citywide traffic accident prediction. In 2019 20th IEEE International Conference on Mobile Data Management (MDM) (pp. 369-370). IEEE.

[24] Iranitalab, A., & Khattak, A. (2017). Comparison of four statistical and machine learning methods for crash severity prediction. Accident Analysis & Prevention, 108, 27-36.

[25] Parveen, H., Yadav, V., (2018). A Particle Swarm Optimization Strategy using QSAR Modeling on the second generation Neural Network. International Conference on Advances in Computing, Communication, Control and Networking, 1188-1199.

[26] Yu, L., Du, B., Hu, X., Sun, L., Han, L., & Lv, W. (2021). Deep spatio-temporal graph convolutional network for traffic accident prediction. Neurocomputing, 423, 135-147.

[27] Jomnonkwao, S., Uttra, S., & Ratanavaraha, V. (2020). Forecasting road traffic deaths in Thailand: Applications of time-series, curve estimation, multiple linear regression, and path analysis models. Sustainability, 12(1), 395.

[28] Houssein, E. H., Gad, A. G., Hussain, K., & Suganthan, P. N. (2021). Major Advances in Particle Swarm Optimization: Theory, Analysis, and Application. Swarm and Evolutionary Computation, 63, 100868.