International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Fine-tuning Road Classification Models: Optimization Strategies for Deep Belief Networks in Transportation Big Data

B. Sangeetha[1], Dr. M. Prabakaran[2]
[1]Research Scholar, PG and Research Department of Computer Science, Government Arts College (Autonomous) (Affiliated to Bharathidasan University, Tiruchirappalli), Karur -5, Tamilnadu, India.
[2]Research Advisor& Associate Professor, PG and Research Department of Computer Science, Government Arts College (Autonomous) (Affiliated to Bharathidasan University, Tiruchirappalli), Karur – 5, Tamilnadu, India.

Abstract

Road traffic accidents are very essential for common people, consequential an estimated 1.2 million deaths and 50 million injuries all over the world every year. In this emerging world, the road accidents are among the principal reason of fatality and injury. The concern of traffic safety has heaved immense alarms across the manageable enhancement of contemporary traffic and transportation. The analysis on road traffic accident grounds can detect the major aspects quickly, professionally and afford instructional techniques to the prevention of traffic accidents and reduction of road traffic accident, which might significantly decrease personal victim by means of road traffic accidents. Data Mining techniques are used in the process of knowledge discovery for many domains' problems. Feature Selection plays a vital role for a large number of datasets. In this paper, the classification of road accident in transportation domain was analyzed with the assistance of the proposed Intelligent classification technique. In this proposed technique, the DBN hidden layers weights are optimized by using evolutionary Genetic algorithm. This GA is utilized to enhance the classification accuracy by applying the hidden layers of Restricted Boltzmann Machine (RBM). The comparative results show that the proposed intelligent classifier gives the improved accuracy, specificity, precision, Sensitivity, F-Measure, and reduced false positive rate.

Keywords: Road accident, Deep Belief Network, Evolutionary algorithm, Genetic algorithm, Artificial Neural Network, Restricted Boltzmann Machine.

1. Introduction

India has the second biggest road network in the world. Therefore, road accidents occur rather recurrently and moreover, they assert several lives every year. It is essential to detect the origin for road accidents with the aim of evade them. In the recent years the compilation of data on traffic amount has turned out to be a huge part of the work of road forecasting programs by both value and workforce. The Traffic data is apportioned into various rules by categorizing the breakpoints for traffic factors in the data. In two-regime traffic methods, decisive possession is employed to divide free flow and crammed flow conditions. Traffic clogging roots incredible failure in terms of both time and power worn out. The forecast of Traffic flow is an essential detective subject in an Intelligent Transportation System (ITS) [1], and it can be utilized as an vital measurement to sort the predicament of traffic congestion. Traffic clogging is rooted while the traffic stipulates methods or surpasses the obtainable facility of the traffic network. Fortunately, as a result of systematic sensor instrumentations of road networks in main metropolis in addition to the vast accessibility of supplementary commodity sensors from which traffic data can be copied (e.g., CCTV cameras, GPS devices), a huge amount of instantaneous and chronological traffic data at very high spatial and sequential decisions has grown to be obtainable [2].

The concern of affording the travel safety on the road network in the urban and suburban is one of the primary ethics which leads the engineering, traffic and transportation development. Approximately 3,500 people die by the road accidents all over the world every day. Tens of millions of people are wounded or disabled each year. Pedestrians, children, cyclists, bike, car users and the aged are amongst the most susceptible of road users. WHO functions with partners - governmental and nongovernmental - around the globe to elevate the report of the inevitability of road traffic injuries and support high-quality trainings associated to helmet and seat-belt wearing, drunk and drive, rash riding and being noticeable in traffic [2]. On the other hand, the accident is predictable, specified the definition, "an incident is distinct as the rapid accidental release of or the spotlight to a dangerous substance that marks in or might logically have brought about the injuries, deaths, considerable possessions or ecological harm, evacuation or sheltering in place [3]". The expense of these accidents can be serious trouble to the government. Therefore, the road accidents are a severe threat to public transportation. In particular, the colossal economic toll of road accidents on human societies inflict, recuperating road safety needs awareness to three effective features: human, the road and the vehicle [4] [5]. The bureaucrats of the transport system also require stabling the road safety wants through inadequate resources to lessen the accidents and augment the road conditions. In fact, the p is as primary objective is as conceivable in reducing accidents that can be loomed to this target by managing the traffic engineering, driver training and implementation [6].

2. Related Works

Kong, Xiangjie, et al [7] proposed LoTAD to discover abnormal areas with long-standing poor traffic conditions. In particular, the authors methods pack sourced bus data into TS-segments (Temporal and Spatial segments) to form the traffic condition. In a while, the authors investigate anomalous TS-segments in each bus line by scheming their AI (Anomaly Index). Subsequently, the authors merge anomalous TS-segments identified in various lines to mine anomalous regions. The data of anomalous areas affords recommendations for prospect traffic planning.

Kumar, PriyanMalarvizhi, et al [8] suggested an valuable traffic control method with the support of IoV technology. The proposed technique is established in the study area of Vellore district, Tamil Nadu, India. The street maps are segmented into small number of different maps. Ant colony algorithm is practiced to each map with the intention to discover the optimal route. In addition, Fuzzy logic-based traffic intensity calculation function is proposed in this paper to form the heavy traffic. The planned IoV based route selection technique is contrasted with the active shortest path selection algorithms like Prim's algorithm, Kruskal's algorithm, and Dijikstra algorithm.

Lin, Lu, et al [9] intended to advance the road traffic speed prediction by combining the customary speed sensing data with new-type "sensing" data from cross domain sources, such as tweet sensors from social media and route sensors from map and traffic service platforms. Jointly modelling data from numerous datasets fetches several disputes, including location uncertainty of low-resolution data, language vagueness of traffic depiction in texts, and heterogeneity of cross-domain data. Due to these difficulties, the authors exhibit a unified probabilistic framework, called Topic-Enhanced Gaussian Process Aggregation Model (TEGPAM).

Moriya, Koichi, Shin Matsushima, and Kenji Yamanishi [10] proposed a new method for mining traffic menace from such assorted data. Traffic threat refers to the prospect of event of traffic accidents.

Zhang, Zhenhua, et al [11] occupied deep learning in identifying the traffic accident from social media data. Initially, the authors methodically explore the 1-year over 3 million tweet contents in two metropolitan regions: Northern Virginia and New York City. The paired tokens can incarcerate the association rules inbuilt in the accident-related tweets and enhance the precision of the traffic accident detection. Second, two deep learning methods: Long Short-Term Memory (LSTM) and Deep Belief Network (DBN) are examined and employed on the extorted token.

Even if some of the earlier stated techniques exhibit advancements by forecast accuracy, these developments derived at the rate of intricacy and interpretability of the method. In contrast, the most preceding functions highlight the forecasting power of their methods, without generous adequate consideration to the issue of recognizing the most instructive variables that disturb the blend of clients. Through this research work, an effective classification method has proposed to classify the road according to the road accident in the transportation domian by blending the Genetic algorithm and Deep Belief Network.

3. Genetic Algorithm

Genetic Algorithm is an intelligent probabilistic search algorithm which can be applied to a variety of combinational optimization problems. Theoretical foundations of Genetic Algorithm were initially developed by Holland in 1970's. The inspiration of GA is based on the evolutionary process of biological organisms in nature. During the course of evolution, natural population evolves according to the principle of natural selection and survival of the fittest. Individuals who are easily adaptable to all environmental conditions and have higher fitness are more likely to reproduce and generate offspring while lower fitness individuals are eliminated from population [16]. The combination of good characteristics from highly adaptive ancestors may produce even more fit offspring. In this way, species evolve more and more to become well adapted on environment.

A Genetic Algorithm stimulates these processes by taking an initial population of individuals and applying GA operators in each generation. Each individual is encoded as a chromosome which is a solution to the problem. A chromosome is a collection of genes, means an individual is made up of genes. The fitness of each individual is calculated by objective function. Highly fit individuals are given chances for reproduction, in crossover procedure. Mutation is optional for changing some of genes in individual to avoid duplicity. This evolution, selection, crossover process repeated until the condition is fulfilled.

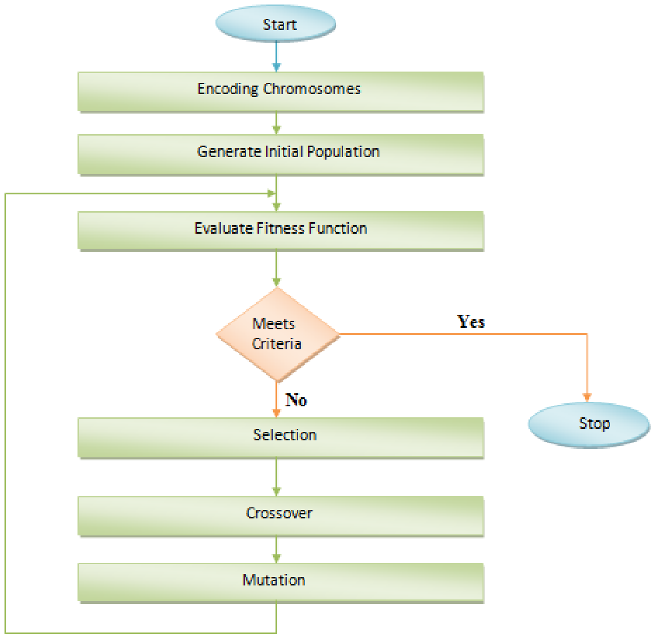

Figure 1: Flow of Genetic Algorithm

3.3.3.1 Encoding of a Chromosome

The chromosome should be encoded in such a way that it must represent information about the solution. The most commonly used way to encode chromosome is in binary string. Each bit represents some information about solution. Every chromosome is a collection of genes where each gene represents each bit of chromosome.

Table 1: Representation of the Chromosome

Chromosome 1	1110001101011010
Chromosome 2	1101010101110100

Pseudo Code of Genetic Algorithm

BEGIN

Initialize population with random candidate Solutions

Evaluate each candidate

Repeat until (Termination condition is satisfied)

DO

Select parents

Recombine pair of parents

Mutate the resulting offspring

Evaluate new candidate

Select individuals for next generation

END DO

END

Here the use of genetic algorithm is divided into two main sections: the pre-calculation section and also the detection section. In that pre-calculation section, collection of chromosome is created by set of training data. That chromosome sets are utilized in the next section for the purpose of comparison. The primary steps in pre-calculation as

Algorithm 1: Initialize chromosomes for evaluation

Input: Reduced Dataset (for training)

Output: A collection of chromosomes

Step 1: Range = 0.125

Step 2: For every training information

Step 3: If it has any nearest neighbor chromosome within Range

Step 4: Combine it with the adjacent chromosome

Step 5: Else

Step 6: Generate a new chromosome

Step 7: End if

Step 8: End for

4. Deep Belief Network

Deep belief networks (DBNs) [17][18] is a feasible multilayer neural system comprised of piling restricted Boltzmann machine for experiencing intricate data patterns. DBN can be distributed into two phases of the learning process. Initially, it is unsupervised learning that is employed by Contrastive Divergence (CD) algorithm called pre-training to acquire the initial weights from stacked RBMs. Secondly, it is supervised learning employed by error back-propagation (BP) algorithm named as fine-tuning that tunes the initial weights to get the final weights.

In Joint dissemination between visible layer x and hidden layer h gained by the energy-based probabilistic framework as follows:

$$p(x, h^1, \dots, h^n) = \left(\prod_{i=1}^{n-2} p(h^i \middle| h^{i+1})\right) . p(h^{n-1}|h^n)$$

In the above equation, 1=h^0, $p(h^i \middle| h^{i+1})$ is a conditional distribution in RBM for hidden-hidden units agreeing to the kth level of DBNs, and hidden joint distribution in top-level RBM is $p(h^{n-1}|h^n)$ in which each layer, the measured output of RBM is used as a subsequent layer input.

4.1 Restricted Boltzmann Machines

Restricted Boltzmann Machines (RBMs) [19] [20] is the fundamental component of a DBNs that is preserved as a generative stochastic graphical and an unsupervised energy-based generative model. RBM is a level of visible units x and a layer of hidden units h, undirected and associated by proportionally weighted networks. Each hidden unit of RBM has the competence to encrypt at best one higher order interaction among inputs.

To epitomize the tricky complication, RBM entails a reduced amount of hidden units when an explicit amount of latent reasons given in the input and this state can be examined by RBM technique with contrastive divergence (CD) unsupervised learning algorithms. The technique elucidates the energy function as follow:

$$E(x, h; \theta) = -\sum_{i=1}^{l} \sum\nolimits_{j=}^{m} x_i W_{ij} h_j - \sum_{i=1}^{l} b_i x_i - \sum_{j=1}^{l} a_j h_j$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Where *l* and *m* are the number of visible and hidden units and θ = {w a b} are the model parameters. Specifically, the joint distribution of visible (x) and hidden (h) units are shown as follow:

$$P(x, h; \theta) = \frac{1}{Z(\theta)} \exp\left(-E(x, h; \theta)\right)$$

In the above equation, Z(θ) is known as a partition function that is employed for stabilizing constant for energy function. And systematically computation of qualified prospects is as follow:

$$p(h_j = 1/x) = \sigma\left(b_j + \sum_{i=1}^{l} x_i w_{ij}\right)$$

$$p(x_i\, 1/h) = \sigma\left(a_i + \sum_{j=1}^{m} h_j w_{ji}\right)$$

Where σ is a sigmoid function.

5. Proposed Classification (Genetic Algorithm + Deep Belief Network) Method

The following figure 2 gives the proposed classification method for categorizing the road accident dataset into Accident Zone (YES), Not an accident zone (NO) group. Deep Belief Network (DBN) has utilized for the arrangement and its hidden layer weights are augmented by using Genetic algorithm. The following phases are employed in this recommended classification method.

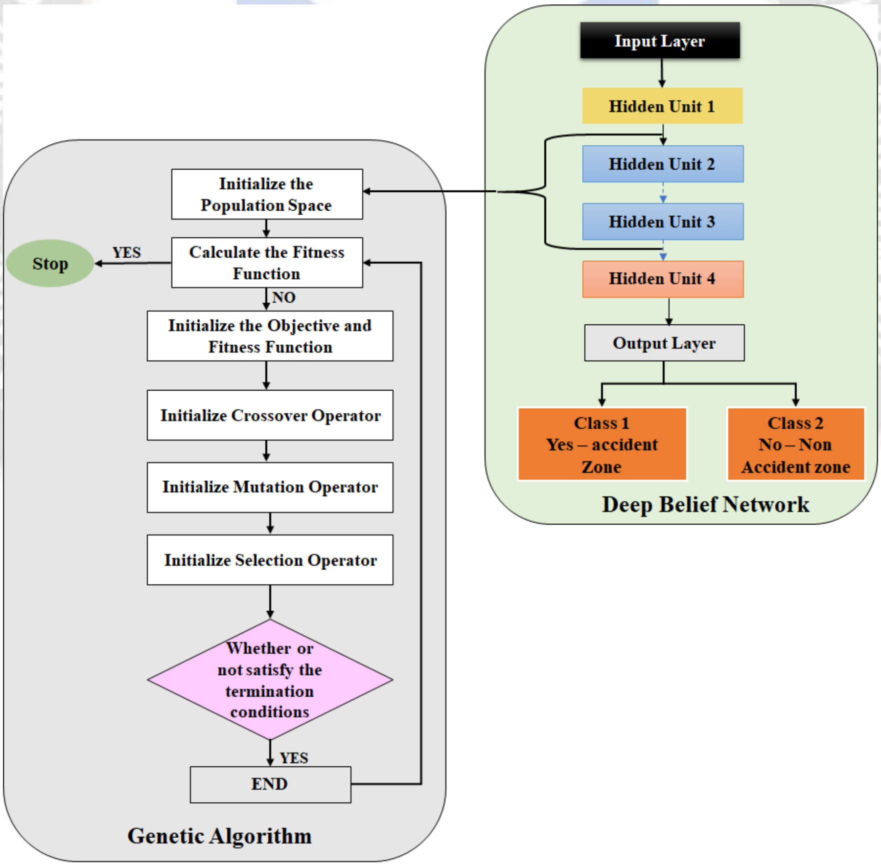

Figure 2: Proposed Classification Method using Deep Belief Network (DBN) and Genetic Algorithm (GA)

Stage 5.1: Population Initialization

The personalities are produced unsystematically. For each genetic factor, the value of that gene is nominated as a random positive integer in the series, and the size of the initial population is 100.

Stage 5.2: Objective function and fitness function

The root means a squared error of the weights is nominated as an objective function. The lesser value of the RMS of the weights means the better classification. The form of the objective function can be defined as follows:

$$g = \sqrt{\frac{1}{N}\sum_{t=1}^{N}|\ddot{x}_2|^2}$$

where N is the number of sampling points. In general, the elucidation epitomized by a chromosome is openly assessed by the fitness function. The fitness function is described by the (1/g). The individual with the greater fitness value displays that the property of control rules is healthier. When the advancement method finishes, the optimum individual can be acquired.

Stage 5.3: Crossover Operation

Crossover operation is the process of generating new individuals along with the boundary rate concluded the random combination of genetic materials opted from the parents. Two-point crossover algorithm is employed in this study. Each pair of parents has a crossover rate expressed by p_c= 0.95, and the offspring can be created by consuming it. Two points have unsystematically opted, and by associating the crossover rate and random rate, if the crossover condition is mollified, the new strings can be attained through consuming swap technique.

Stage 5.4: Mutation Operation

Mutation is the method of arbitrarily varying the values of genes in a chromosome. The objective of mutation is to host a new genetic material into a present individual and add diversity to the genetic features of the population. Mutation is smeared at a specific rate expressed by pm, which is commonly a low value and can make every gene create the mutated offspring. Here, in-order mutation technique is employed. Two mutation points have unsystematically opted and only the strings between these mutation points accomplish mutation. The mutation rate is p_m= 0.5. If the random rate is greater than the mutation rate, the strings between the points are in reciprocal order.

Stage 5.5: Selection Operation

A new population of contestant results is opted at the completion of each generation to assist the population of the next generation. The selection operator confirms that virtuous individuals can endure and be put into the next generation. The major goal of the selection operator is to attain the best solution. Roulette wheel selection is a proportional selection operator. By using this technique, the fitness values are regularized via distributing the fitness value of each individual by the outline of the fitness values of all individuals. As a result, the probability distribution can be viewed as a roulette wheel, and the size of each slice is proportional to the normalized selection possibility of an individual. Selection can be equated to the circum-rotating of a roulette wheel and the slice which finishes up at the maximum will be documented. Subsequently, the resultant individual has opted. Roulette wheel selection method s employed and the selection rate is p_s= 0.95. The probability of the i^{th} individual is pronounced in the following expression:

$$P_i = \frac{f_i}{\sum_{i=1}^{I} f_i}$$

f_i is the fitness value of the ith individual. I is the size of all the individuals in the population space.

Stage 5.6: Fine-tuning stage

The working regulation of this phase hangs on the customary back-propagation algorithm. To identify and categorize the churners, non-churners and hesitant, an output layer is suggested as the maximum point of the Deep Neural Network. Furthermore, there is 'N' number of input neurons (based on the features), and three hidden layers are employed in the present investigation of Deep Neural Network. The optimized weight is calculated through the training stage with the support of a training data set, where back propagation initiates with the weights that were accomplished in the pre-training stage. From the optimal weights, the layer work is revitalized and is presented as follows.

$$T(m_i = 1/n) = \sigma(m_i + \sum opt\_w_{ij} n_j)$$

$$T(n_i = 1/m) = \sigma(n_i + \sum opt\_w_{ij} m_j)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Where m and n represent the bias vector and hidden layers and σ is the sigmoid function with the range of (0,1). Moreover, the training dataset is accomplished in anticipation of the optimized weight is grasped, or determined accuracy is accomplished with the support of the above equation.

6. Result and Discussion

The suggested intelligent categorization method was applied in the operational platform MATLAB 2018 with system configurations like an i7 processor with 16GB RAM.

6.1 Description of the Dataset

The road accident dataset is taken from the Kaggle Repository. Table 2 depicts the features involved in the road accident dataset [R].

Table 2: Description of the Dataset

Feature Number	Feature Name
1	accident_index
2	vehicle_reference
3	vehicle_type
4	towing_and_articulation
5	vehicle_manoeuvre
6	vehicle_location-restricted_lane
7	junction_location
8	skidding_and_overturning
9	hit_object_in_carriageway
10	vehicle_leaving_carriageway
11	hit_object_off_carriageway
12	1st_point_of_impact
13	was_vehicle_left_hand_drive?
14	journey_purpose_of_driver
15	sex_of_driver
16	age_of_driver
17	age_band_of_driver
18	engine_capacity_(cc)
19	propulsion_code
20	age_of_vehicle
21	driver_imd_decile
22	driver_home_area_type
23	vehicle_imd_decile
24	NUmber_of_Casualities_unique_to_accident_index
25	No_of_Vehicles_involved_unique_to_accident_index
26	location_easting_osgr
27	location_northing_osgr
28	longitude
29	latitude
30	police_force
31	accident_severity
32	number_of_vehicles
33	number_of_casualties
34	date
35	day_of_week
36	time

37	local_authority_(district)
38	local_authority_(highway)
39	1st_road_class
40	1st_road_number
41	road_type
42	speed_limit
43	junction_detail
44	junction_control
45	2nd_road_class
46	2nd_road_number
47	pedestrian_crossing-human_control
48	pedestrian_crossing-physical_facilities
49	light_conditions
50	weather_conditions
51	road_surface_conditions
52	special_conditions_at_site
53	carriageway_hazards
54	urban_or_rural_area
55	did_police_officer_attend_scene_of_accident
56	lsoa_of_accident_location
57	casualty_reference
58	casualty_class
59	sex_of_casualty
60	age_of_casualty
61	age_band_of_casualty
62	casualty_severity
63	pedestrian_location
64	pedestrian_movement
65	car_passenger
66	bus_or_coach_passenger
67	pedestrian_road_maintenance_worker
68	casualty_type
69	casualty_home_area_type
70	casualty_imd_decile

6.2 Performance Metrics

Following table 3 portrays the performance metrics. These metrics have measured for assessing the current methods with the recommended method.

Table 3: Performance Metrics for evaluation of the proposed technique

Metrics	Equation
Accuracy	$\frac{TP+TN}{TP+FN+TN+FP}$
True Positive Rate (TPR) (Sensitivity or Recall)	$\frac{TP}{TP+FN}$
False Positive Rate (FPR)	$\frac{FP}{FP+TN}$
Precision	$\frac{TP}{TP+FP}$
True Negative Rate (Specificity)	1- False Positive Rate (FPR)

False Negative Rate	1-True Positive Rate (TPR)
F-Measure	$2.\frac{Precision.Recall}{Precision+Recall}$

From the above table 4,

True Positive (TP) = the number of cases correctly identified as churners

True Negative (TN) = the number of cases correctly identified as Non-churners

False Positive (FS) = the number of cases incorrectly identified as churners

False Negative (FN) = the number of cases incorrectly identified as Non-churners.

6.3 Result and Discussion

In the pre-processing stage, a new feature selection has proposed based on FireFly and Animal Migration Optimization algorithm [12][13][14][15]. It gives the most relevant and reduced the redundant features from the dataset. Table 3 gives the reduced dataset obtained from the pre-processing step.

Table 4: Optimal Dataset from Pre-processing step

Sl.No	**Feature Name**
1	vehicle_type
2	towing_and_articulation
3	vehicle_location-restricted_lane
4	junction_location
5	skidding_and_overturning
6	hit_object_in_carriageway
7	vehicle_leaving_carriageway
8	hit_object_off_carriageway
9	was_vehicle_left_hand_drive?
10	sex_of_driver
11	age_of_driver
12	engine_capacity_(cc)
13	age_of_vehicle
14	NUmber_of_Casualities_unique_to_accident_index
15	No_of_Vehicles_involved_unique_to_accident_index
16	longitude
17	latitude
18	accident_severity
19	number_of_vehicles
20	number_of_casualties
21	date
22	day_of_week
22	time
23	local_authority_(district)
24	local_authority_(highway)
25	road_type
26	speed_limit
27	weather_conditions
28	road_surface_conditions
29	casualty_severity
30	car_passenger
31	bus_or_coach_passenger

Table 5 depicts the performance analysis on accuracy for the original dataset and pre-processed optimal dataset with ANN, DBN and proposed intelligent classification method. From table 5, it is clear that the accuracy of the proposed classification method is 94.74% for a reduced dataset, whereas 92.01% accuracy obtained for original dataset. When comparing with ANN, 86.40% for an original dataset and 91.05% for a reduced dataset. DBN gives the accuracy of 89.16% for an original dataset and 93.30% for the reduced dataset. Figure 3 presents the graphical representation of the performance analysis of ANN, DBN and proposed intelligent classification method.

Table 5: Performance analysis on accuracy for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	86.40%	91.05%
Deep Belief Network	89.16%	93.30%
Proposed Classification method	92.07%	94.74%

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	86.40%	91.05%
Deep Belief Network	89.16%	93.30%
Proposed Classification method	92.07%	94.74%

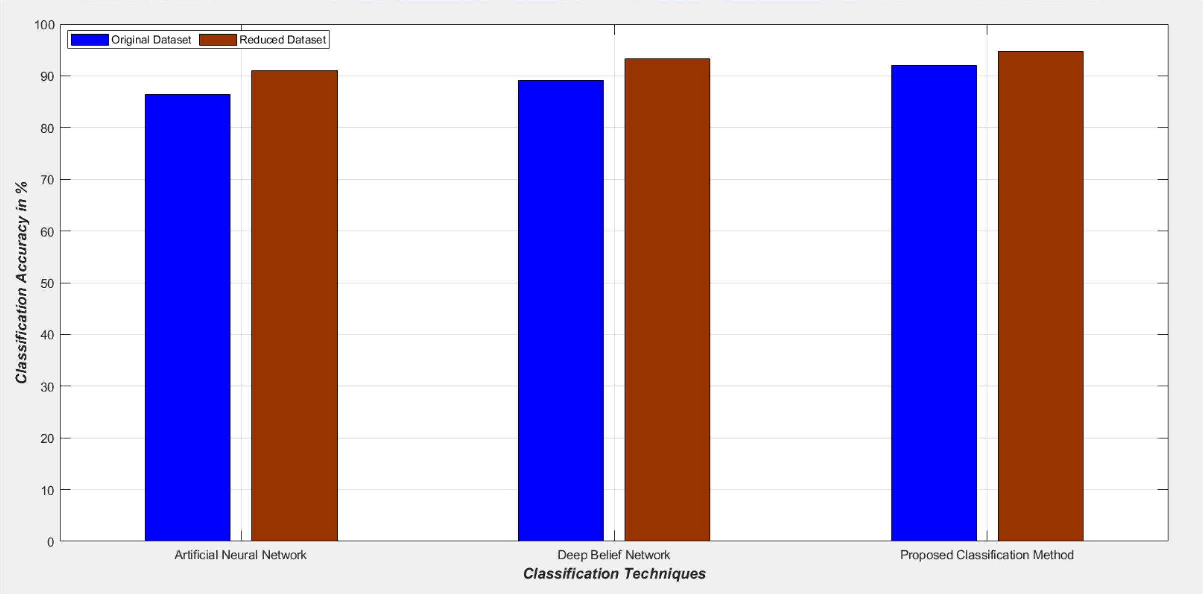

Figure 3: Graphical representation of the Classification Accuracy of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

Table 6 depicts the performance analysis on True Positive Rate (Sensitivity or Recall) obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 6, the proposed method gives 0.8390 (original dataset) and 0.9395 for the reduced dataset. DBN gives the 0.7089 for an original dataset and 0.8754 for the reduced dataset, whereas ANN gives only 0.5781 for an original dataset and 0.9042 for the reduced dataset. The proposed classifier gives the increased TPR for original and reduced dataset than the existing techniques. Figure 4 depicts the graphical representation of the True Positive Rate (Sensitivity or Recall) obtained by ANN, DBN and proposed intelligent classification method.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Table 6: Performance analysis on True Positive Rate (Sensitivity or Recall) for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.5781	0.9042
Deep Belief Network	0.7089	0.8754
Proposed Classification method	0.8390	0.9395

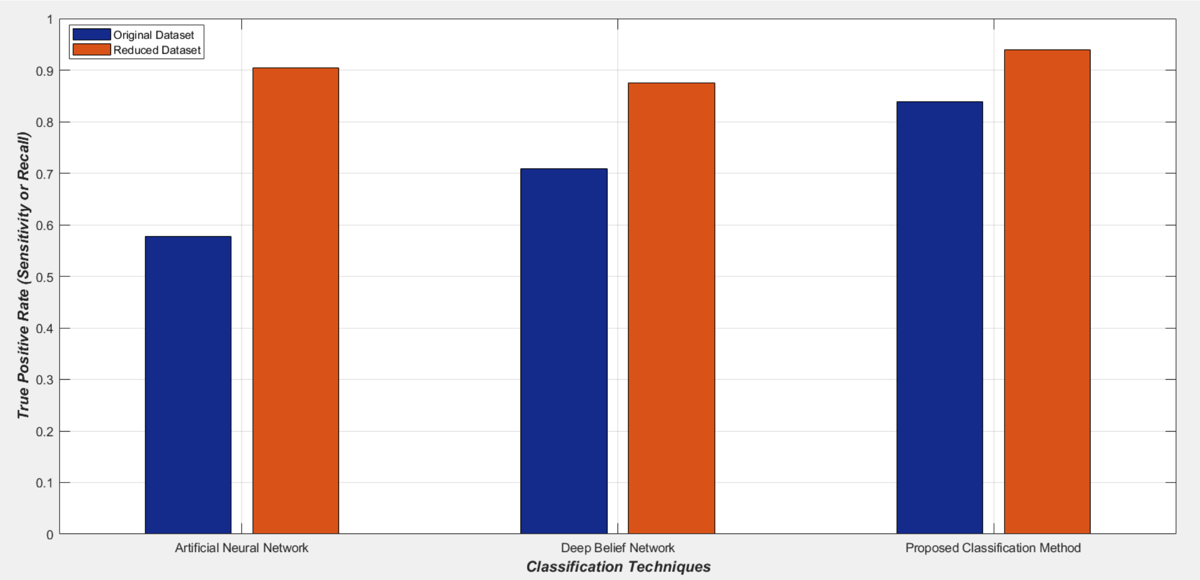

Figure 4: Graphical representation of the True Positive Rate (Sensitivity or Recall) of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

Table 7 depicts the performance analysis on False Positive Rate obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 7, the proposed method gives 0.0704 (original dataset) and 0.0476 for the reduced dataset. DBN gives the 0.0907 for an original dataset and 0.0602 for the reduced dataset, whereas ANN gives only 0.1184 for an original dataset and 0.0910 for the reduced dataset. It is clear that the proposed classifier gives the least value of FPR while comparing with other classification techniques. Figure 5 depicts the graphical representation of the False Positive Rate (FPR) obtained by ANN, DBN and proposed intelligent classification method.

Table 6: Performance analysis on False Positive for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.1184	0.0910
Deep Belief Network	0.0907	0.0602
Proposed Classification method	0.0704	0.0476

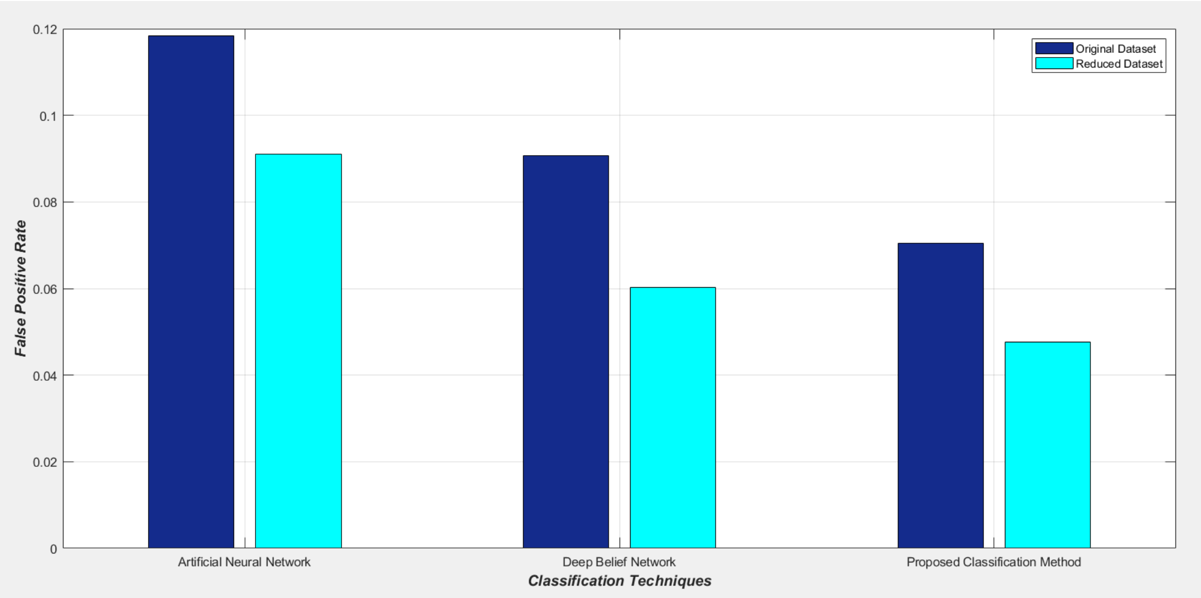

Figure 5: Graphical representation of the False Positive Rate of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

Table 8 depicts the performance analysis on Precision obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 8, the proposed method gives 0.5610 (original dataset) and 0.7080 for the reduced dataset. DBN gives the 0.4285 for an original dataset and 0.6273 for the reduced dataset, whereas ANN gives only 0.2298 for an original dataset and 0.4120 for the reduced dataset. The proposed classifier gives the increased Precision for original and reduced dataset than the existing techniques. Figure 6 depicts the graphical representation of the True Positive Rate (Sensitivity or Recall) obtained by ANN, DBN and proposed intelligent classification method.

Table 8: Performance analysis on Precision for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.2298	0.4120
Deep Belief Network	0.4285	0.6273
Proposed Classification method	0.5610	0.7080

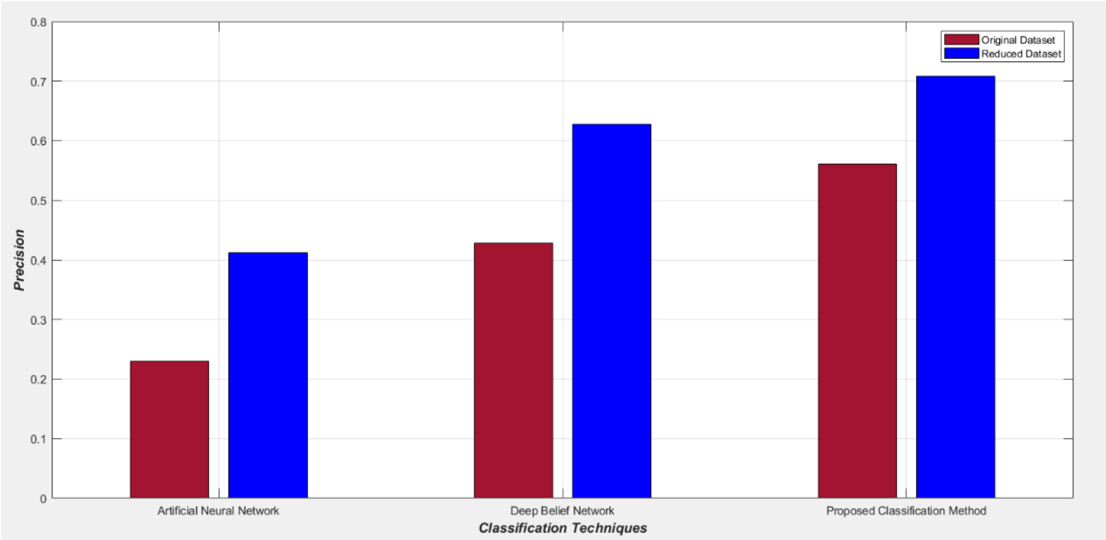

Figure 6: Graphical representation of Precision of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

Table 9 depicts the performance analysis on Specificity obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 9, the proposed method gives 0.8816 (original dataset) and 0.909 for the reduced dataset. DBN gives the 0.9093 for an original dataset and 0.9398 for the reduced dataset, whereas ANN gives only 0.9296 for an original dataset and 0.9524 for the reduced dataset. The proposed classifier gives the increased Specificity for original and reduced dataset than the existing techniques. Figure 7 depicts the graphical representation of the Specificity obtained by ANN, DBN and proposed intelligent classification method.

Table 9: Performance analysis on Specificity for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.8816	0.909
Deep Belief Network	0.9093	0.9398
Proposed Classification method	0.9296	0.9524

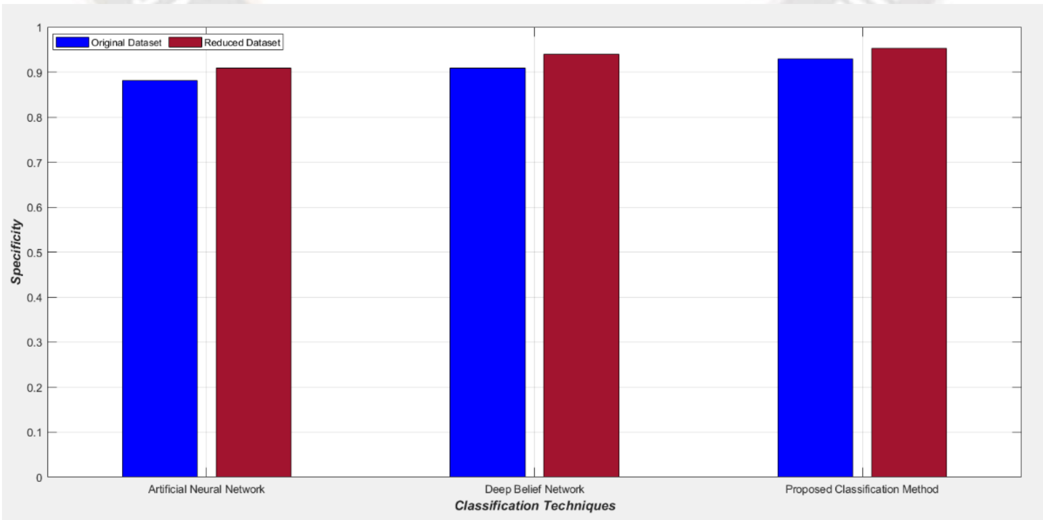

Figure 7: Graphical representation of Specificity of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Table 10 depicts the performance analysis on False Negative Rate (Miss Rate) obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 10, the proposed method gives 0.161 (original dataset) and 0.0905 for the reduced dataset. DBN gives the 0.2911 for an original dataset and 0.1246 for the reduced dataset, whereas ANN gives only 0.4781 for an original dataset and 0.1658 for the reduced dataset. It is clear that the proposed classifier gives the least value of False Negative Rate (Miss Rate) while comparing with other classification techniques. Figure 8 depicts the graphical representation of the False Positive Rate (FPR) obtained by ANN, DBN and proposed intelligent classification method.

Table 10: Performance analysis on False Negative Rate (Miss Rate) for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.4781	0.1658
Deep Belief Network	0.2911	0.1246
Proposed Classification method	0.161	0.0905

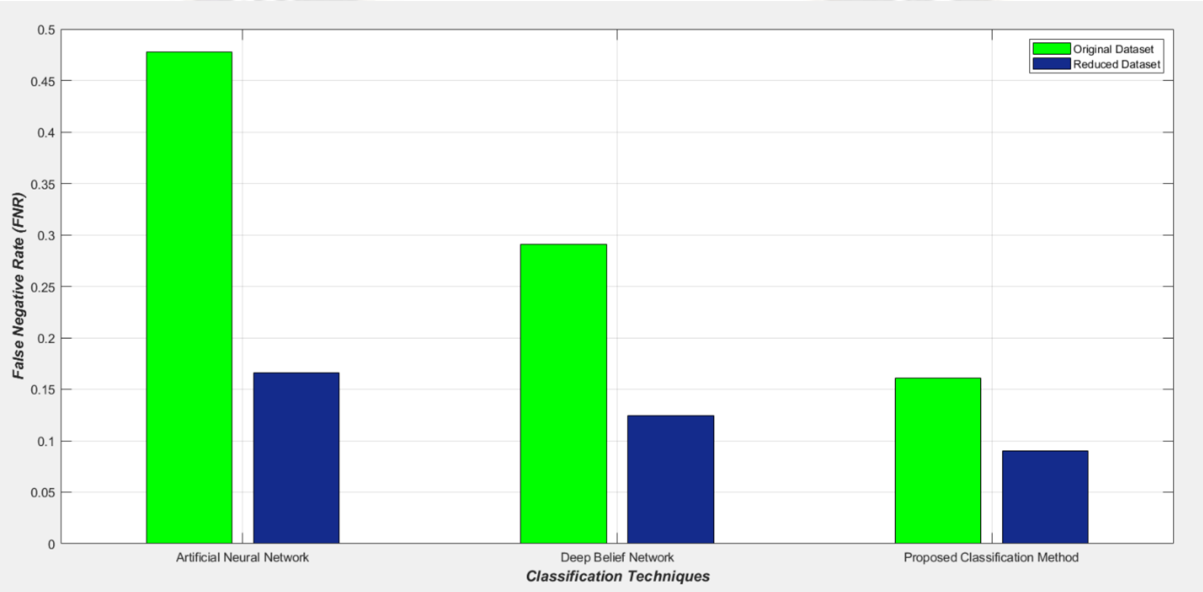

Figure 8: Graphical representation of False Negative Rate (FNR) of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

Table 11 depicts the performance analysis on F-Measure obtained by ANN, DBN and proposed intelligent classification method for the original dataset and reduced dataset. From table 11, the proposed method gives 0.6723 (original dataset) and 0.7961 for the reduced dataset. DBN gives the 0.5344 for an original dataset and 0.7308 for the reduced dataset, whereas ANN gives only 0.3287 for an original dataset and 0.5716 for the reduced dataset. The proposed classifier gives the increased F-Measure value for original and reduced dataset than the existing techniques. Figure 9 depicts the graphical representation of the F-Measure obtained by ANN, DBN and proposed intelligent classification method.

Table 11: Performance analysis on F-Measure for the original dataset and pre-processed optimal dataset with Artificial Neural Network, Deep Belief Network, and Proposed classification method

Classification Techniques	Type of Dataset	
	Original Dataset	Pre-Processed Dataset
Artificial Neural Network	0.3287	0.5716
Deep Belief Network	0.5344	0.7308
Proposed Classification method	0.6723	0.7961

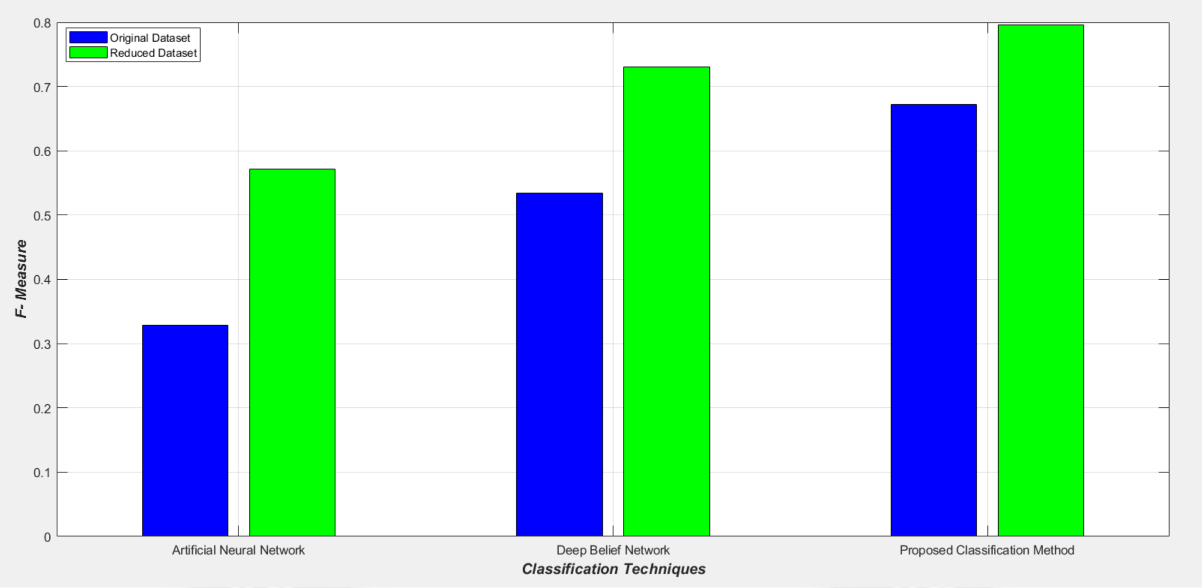

Figure 9: Graphical representation of F- Measure of the proposed Classification method, Artificial Neural Network and Deep Belief Network for the classification of churn customers

7. Conclusion

In brief, the objective of this research study is to forecast the road in the transportation domain by using the proposed intelligent classification model. Through Deep Learning architecture and evolutionary algorithm, we planned to attain better accuracy in recognizing the accident zone and non-accident zone in the road transportation. Along with the experimental outcome, the recommended intelligent classification method is operative for the arrangement of accident zone and non-accident zone in terms of accuracy, specificity, precision, sensitivity, F-measure, False Negative Rate (Miss Rate) with its values. The precision level has visibly evident that the recommended classification technique is deeply proficient in identifying the accident zone in road transportation. The categorization presentations of this study establish the benefits of this technique: it is rapid, modest to operate, non-invasive.

Refrences

[1] Alsolami, Bdoor, Rashid Mehmood, and Aiiad Albeshri. "Hybrid Statistical and Machine Learning Methods for Road Traffic Prediction: A Review and Tutorial." Smart Infrastructure and Applications. Springer, Cham, 2020. 115-133.
[2] Laña, Ibai, et al. "Road Traffic Forecasting Using NeuCube and Dynamic Evolving Spiking Neural Networks." International Symposium on Intelligent and Distributed Computing. Springer, Cham, 2018.
[3] Lee, Jooyoung, Ganghwa Kim, and Kitae Jang. "Estimation of Road Surface Condition during Summer Season Using Machine Learning." The Journal of The Korea Institute of Intelligent Transport Systems 17.6 (2018): 121-132.
[4] Ali Khalil, Mohamad, Khaled Hamad, and Abdallah Shanableh. "Developing Machine Learning Models to Predict Roadway Traffic Noise: An Opportunity to Escape Conventional Techniques." Transportation Research Record (2019): 0361198119838514.
[5] Elfar, Amr, Alireza Talebpour, and Hani S. Mahmassani. "Machine learning approach to short-term traffic congestion prediction in a connected environment." Transportation Research Record 2672.45 (2018): 185-195.
[6] Surya, S., and N. Rakesh. "Traffic Congestion Prediction and Intelligent Signalling Based on Markov Decision Process and Reinforcement Learning." Proceedings of International Conference on Cognition and Recognition. Springer, Singapore, 2018.
[7] Kong, Xiangjie, et al. "LoTAD: Long-term traffic anomaly detection based on crowdsourced bus trajectory data." World Wide Web 21.3 (2018): 825-847.
[8] Kumar, Priyan Malarvizhi, et al. "Ant colony optimization algorithm with Internet of Vehicles for intelligent traffic control system." Computer Networks 144 (2018): 154- 162.
[9] Lin, Lu, et al. "Road traffic speed prediction: a probabilistic model fusing multi- source data." IEEE Transactions on Knowledge and Data Engineering 30.7 (2018): 1310- 1323.
[10] Moriya, Koichi, Shin Matsushima, and Kenji Yamanishi. "Traffic risk mining from heterogeneous road statistics." IEEE Transactions on Intelligent Transportation Systems 99 (2018): 1-14.
[11] Zhang, Zhenhua, et al. "A deep learning approach for detecting traffic accidents from social media data." Transportation research part C: emerging technologies 86 (2018): 580-596.

[12] Poornappriya, T. S., and M. Durairaj. "High relevancy low redundancy vague set based feature selection method for telecom dataset." Journal of Intelligent & Fuzzy Systems 37.5 (2019): 6743-6760.

[13] Durairaj, M., and T. S. Poornappriya. "Choosing a spectacular Feature Selection technique for telecommunication industry using fuzzy TOPSIS MCDM." International Journal of Engineering & Technology 7.4 (2018): 5856-5861.

[14] Durairaj, M., and T. S. Poornappriya. "Importance of MapReduce for Big Data Applications: A Survey." Asian Journal of Computer Science and Technology 7.1 (2018): 112-118.

[15] Durairaj, M., and T. Poornappriya. "A Review on Big Data Analytics Tools for Telecommunication Industry." International Journal of Circuit Theory and Applications 9 (2016): 185-193.

[16] Demidova, L. A., M. M. Egin, and R. V. Tishkin. "A Self-tuning Multiobjective Genetic Algorithm with Application in the SVM Classification." Procedia Computer Science 150 (2019): 503-510.

[17] Verma, Maneesh Kumar, et al. "Phishing Website Detection Using Neural Network and Deep Belief Network." Recent Findings in Intelligent Computing Techniques. Springer, Singapore, 2019. 293-300.

[18] Dan, Qiulin, et al. "A Fault Diagnosis Approach Based on Deep Belief Network and Its Application in Bearing Fault." Proceedings of 2018 Chinese Intelligent Systems Conference. Springer, Singapore, 2019.

[19] Jaworski, Maciej, et al. "Resource-Aware Data Stream Mining Using the Restricted Boltzmann Machine." International Conference on Artificial Intelligence and Soft Computing. Springer, Cham, 2019.

[20] Durairaj, M., and T. S. Poornappriya. "Why feature selection in data mining is prominent? A survey." Proceedings of International Conference on Artificial Intelligence, Smart Grid and Smart City Applications: AISGSC 2019. Springer International Publishing, 2020.