Journal of Advances in Science and Technology
Vol. 19, Issue No. 1, March-2022, ISSN 2230-9659

Predicting Road Traffic Accident Using Genetic Algorithm

Ms. Nikita Ingawale[1]*, Dr. R.R. Sorte[2]

[1]PG Student, Department of Civil Engineering, TSSM'S Padmabhooshan Vasantdada Patil Institute Technology,Pune
[2]Professor, Department of Civil Engineering, TSSM'S Padmabhooshan Vasantdada Patil Institute Technology,Pune
Pune, India

Abstract - Predicting traffic accidents is a difficult task for everyone involved in motor vehicle traffic since it has the potential to save lives. Predicting the severity of traffic accidents has been attempted with a variety of classification methods, Artificial neural networks, support vector machines, decision trees, and more are included in this category. It is true that neural networks and support vector machines have a huge drawback when it comes to human interpretation (SVMs), whereas low accuracy is a major drawback of classic deep learning techniques like C4.5, ID3, and CART. This paper proposes an alternative to classic DTs based on user preferences in order to address the shortcomings of current methods for predicting traffic accidents. In order to optimise and uncover rules based on the recommended method's support, confidence, and comprehension criteria, we created a new genetic algorithm. Users, such as traffic cops, road and transportation engineers, will be able to apply their expertise while balancing all of the conflicting objectives using the proposed technique. A five-year period (2008–2013) of Tehran Province, Iran, traffic accident data was utilised to evaluate the recommended approach. When it comes to classification measures like accuracy and rule performance metrics, this strategy outperforms classification methods like ANN, SVM and conventional DTS.

Keywords - Road traffic Accidents Prediction, Genetic Algorithm, Highway

----------X----------

1. INTRODUCTION

As the world's population grows and the number of automobiles on the road increases, so does the number of traffic accidents. Improved geometric design, congestion management measures, and more driver education and enforcement are among the most common ways for preventing collisions. In spite of the fact that these methods are usually useful, putting them in place is frequently impossible or extremely expensive. It is hard for transportation safety experts to utilise a single parameter to effectively describe the severity of traffic accidents since several elements play a role and some of them have a considerable impact on one another. Reducing the amount of traffic accidents can be accomplished via the use of mixed current models that consider input and outcome factors as well as their interactions.

Road and traffic accidents are a leading source of death and disability around the world. A collision between a conveyance and another conveyance, a person, or other items is referred to as a road contingency. A road accident not only causes physical damage, but it can also cause partial or complete incapacitation, and in some cases, death. The rising number of traffic accidents is not a healthy indicator for vehicle safety. The only solution is to analyze traffic congestion data in order to discover various causes of road accidents and take preventative actions.

Predictive crash models are a critical tool for assessing the link between collision occurrence and risk variables associated with different traffic entities (also known as the safety performance function).

Iran's roads claim the lives of more than 28000 people per year, resulting in significant economic and social consequences for the country. Accident Driver age and gender, seatbelt use, environmental variables and highway circumstances, as well as vehicle type and safety, are all factors that contribute to the severity of the incident.

For both restricted and unconstrained optimization issues, a genetic method based on natural selection is utilised to solve the model. The third model we look at is one that combines genetic algorithm (GA) and other models. Since GAmodels are still

relatively new in transportation safety research, we'll combine these models to enhance prediction accuracy.

Studies looking at accident frequency have traditionally relied on statistical methods like linear regression models, Poisson model regression models, and negative binomial model regression models, due to the possibility that the frequency of accidents on a highway segment could be considered random.

Genetic algorithm

John Holland created GAs in the 1970s. The following definitions and explanations may help you better understand some of the jargon used in GA. A chromosome is a collection of genes that demonstrates an answer to a specific issue. A distinct attribute of the solution is represented by each of the genes. A population is a group of individuals who share the same set of chromosomes at a certain point in time. A randomly selected population initiates an evolutionary process, which is then repeated repeatedly. During each repetition, a new generation is created. Mutation and crossover operators across separate chromosomes are used in each generation to create this type of evolution. Replacement schemes are used to ensure that the future generation of the population is made up of the best parents or children chromosomes. A population's evolutionary trajectory comes to an end when some preset criteria are met. This process of evolution can occur in a variety of ways. An encoded solution, i.e. a representation of the solution in DNA form, an initial population, mutation and crossover operators, and selection and replacement procedures are all necessary for a GA to work. Algorithm 1 contains the pseudocode for a GA.

On the Road Safety

The availability of transportation is inextricably linked to a country's economic development, and the goal of transportation systems is to make moving freight and passengers from one area to another as efficient and safe as possible. Traffic accidents, which resulted in the loss of lives and property, had become a severe social concern as the number of automobiles on the road increased. According to various accident studies, road accidents are caused by negligence and a lack of road safety standards, rather than natural reasons. Environmental conditions, such as fog in the winter, play a role in the causes of traffic accidents. As a result, road/highway safety has become a need in today's world.

2. METHODOLOGY

Genetic Algorithm

Natural selection, the mechanism that drives biological evolution, may be used to solve both limited and unconstrained optimization problems using a genetic algorithm. Genetic algorithms continuously alter an individual population of solutions. The genetic algorithm randomly selects parents from the existing population at each phase and utilises them to generate the children of the following generation. Evolution is the process through which the population gradually moves toward the best possible answer. There are a number of optimization problems that can't be solved using conventional techniques, such as those with discontinuous or non-differentiable objective functions. Genetic algorithms can be used to solve these difficulties. It was popularised by one of Holland's students, Goldberg, who utilised this approach to tackle a difficult problem for his dissertation on gas-pipeline transmission management in the 1960s and 1970s (Goldberg 1989). The earliest effort to lay a theoretical foundation for Gas was Holland's schema theorem.

De Jong's (1975) work was the first systematic effort to improve GA parameters and established the utility of GAs for function optimization.

GA operators include both cross-over (the swapping of randomly selected chromosomal slices) and mutation (the alteration of a randomly selected chromosome segment).

Figure 1 depicts the GA's genetic cycle, in which the best individuals are continually picked and modified through crossover and mutation.

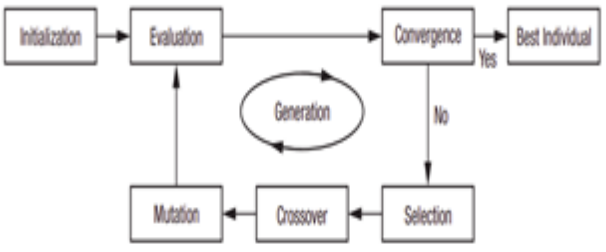

Figure 1: The general structure of Genetic algorithm

3. LITERATURE SURVEY

Seyed Hessam-Allah Hashmienejad & Seyed Mohammad Hossein Hasheminejad (2017) Traffic accident severity prediction using a novel

multi-objective genetic algorithm,International Journal of Crashworthiness,,

In this study, we describe a new rule-based method for predicting the severity of traffic accidents. It was used to build a multi-objective genetic algorithm to optimise and find rules based on the metrics of Support, Confidence, and Comprehensibility. The suggested technique intends to make use of user information, including traffic police, roads and transportation engineers, and the trade-off between all the opposing objectives. Data from 14211 traffic accidents on rural and urban roadways in Tehran Province, Iran, between 2008 and 2013 is used to evaluate the proposed technique. Classification measures such as 88.1 percent accuracy and rule performance indicators such as support and confidence were found to be inferior to the proposed approach (0.79 and 0.74, respectively).

Graphical Prediction of Road Accidents using Data Analysis Swetha P.C,Vaishalli. A. Devika.M,Sneha Balaji,Vol-4 Issue-2 IJARIIE, 2018,

In this research, both aspects are considered. The current study on road accidents typically uses Big Data as a way of data analysis. Another major area of study in the field of traffic accidents is trend analysis. Using trend analysis, you can figure out why the number of accidents is growing or decreasing. There is a lot of data being used in traffic control. It's a big deal, especially when it comes to car accidents. People are more worried about traffic safety when they're driving. As a result, traffic accident data analysis is of paramount importance. Using data from the United Kingdom's Department of Transport, we have developed a method for analysing road accidents in this study. An important goal of this research is the development of an effective method for analysing and predicting the accident zones and their severity. The results suggest that utilising graphical forecasts, the proposed strategy may effectively reduce the occurrence of accidents.

Camilo Gutierrez-Osorio, Cesar Pedraza Departamento de Ingenier, de Sistemas e Industrial, Universidad Nacional de Colombia, Bogota, Colombia,15 May 2020, "Modern Data Sources And Techniques For Analysis And Forecast Of Road Accidents: A Review"Journal of traffic and transportation engineering,

Using machine learning algorithms and cutting-edge information analysis techniques, this study seeks to offer an overview of current state-of-the-art methods for predicting traffic accidents. This article compiles and analyses the most commonly used data sources for road accident forecasting. In addition, a categorization system based on the data's source and features, such as open data, measuring technology, on-board equipment, and social media data, is being considered.

Article ID 7318917, "A Genetic Algorithm Approach For Accelerated Crossing of Emergency Vehicles in Connected and Autonomous Intersection Traffic," by Qiang Lu and Kyoung-Dae Kim, published in the Journal of Advanced Transportation on 10 October 2017.

When an emergency vehicle needs to cross an intersection as quickly as feasible, but other cars' journey times aren't adversely affected, this article provides an intersection control method.

Comparison of artificial neural networks, genetic algorithms, hybrid genetic algorithms, and pattern search approaches for traffic accident severity prediction Iman Aghayan & Nima Noii, Mehmet Metin Kunt, Iman Aghayan & Nima Noii, Mehmet Metin Kunt, Iman Aghayan & Nima Noii, Mehmet Met

A genetic algorithm (GA), pattern search (PS), and artificial neural network (ANN) modelling technique are used in this study to forecast the severity of highway traffic accidents. It was determined that the models should be built using the following inputs: age and gender of the driver; seat belt use; vehicle make and model; road surface; slope; speed to slope ratio; speed to impact ratio; and traffic flow. For our study, just a few thousand collisions happened on the motorway during 2007, thus we used those as our basis for the models. The GA examined a total of eleven to arrive at the best equation. It was then decided to blend the GA and PS techniques using the best GA equation available. The neural network's multi-layer perceptron (MLP) architecture included many layers of feed-forward neurons with hidden sigmoid and linear outputs, which were able to tackle arbitrary multi-dimensional mapping difficulties. 12 inputs, 25 neurons in the hidden layer, and 3 neurons in the output layer were employed in the ANN for testing and validation of the algorithm's performance. R-value, root mean square error (RMSE), mean absolute error (MAE), and the sum of square errors were all used to find the best-fit model (SSE).

An Effective Machine Learning Algorithm and Clustering Method for Predicting the Severity of Crash Injuries NEDAL RATOUTT: Syed Masiur Rahman, Umer Mansoor and Khaled (June 27,

2020); Syed Masiur Rahman, Umer Mansoor and Khaled (June 27, 2020);

Machine learning was used in this work to estimate the extent of accident damage based on 15 crash-related indicators (ML). As a result of the usage of fuzzy c-means, the accuracy of forecasts was increased. Support Vector Machines (SVMs), fuzzy c-means clustering feed-forward neural networks based on SVMs, and fuzzy SVMs based on c-means were the final four machine learning models (SVM-FCM). The trauma centre was able to determine the seriousness of the incident based on the first information supplied by the crash site and make the necessary arrangements for the care of the victims by using features that were easily spotted on accident sites with minimal research. The vehicle and the road conditions are the most important input variables.

The development of a reliable neural network model for estimating the severity of crash-related injuries Qiang Zeng and Helai Huang, 29 May 2019, respectively.

In order to train a neural network model for predicting the severity of accident damage, the study presents a convex combination (CC) strategy and a modified NN pruning methodology (N2PFA). Using a 2006 two-vehicle collision dataset provided by the Florida Department of Highway Safety and Motor Vehicles (DHSMV), the suggested approaches and a traditional back-propagation (BP) model were compared. The CC algorithm surpasses the BP method in terms of convergence and training speed, according to the results. The optimised neural network has the same classification accuracy as a fully connected neural network, although having fewer nodes. Fitting and prediction both exceed the OL, proving once again that the NN has an edge over statistical models when it comes to estimating the degree of accident injury.

A cluster analysis on road traffic accidents using genetic algorithms Sabariah Saharan and Roberto Baragona, 2017,

As the cost of traffic accidents rises and public safety on the road becomes more vital, so does the importance of accident analysis. It is possible to conduct research on factors that influence the frequency and severity of accidents because of the widespread availability of huge datasets. There is a lot of imbalance and overlap in the data. New Zealand's capital city of Christchurch saw a total of 26440 traffic incidents between 2000 and 2009. Road traffic incidents are classified into four levels of severity based on 50 elements. A big imbalanced data set necessitated the use of genetic algorithms, which we chose over more traditional clustering methods like the k-means algorithm. Accidents with varied levels of severity have been studied using the genetic algorithm, based on clustering for unknown K, (GCUK).

Iman Aghayan, Mansour Hadji Hosseinlou, and Mehmet Metin Kunt3, Journal of Civil Engineering and Urbanism, Volume 5, Issue 5: 193-199, September 25, 2018.

Crash damage severity was predicted using a support vector machine and several kernel functions, as detailed in this research. Individual injuries can be estimated using a support vector machine model built from crash data. It was determined that the models should be built using the following inputs: age and gender of the driver; seat belt use; vehicle make and model; road surface; slope; speed to slope ratio; speed to impact ratio; and traffic flow. In addition, this study's output parameters included three different degrees of damage. Multi-layer perceptrons, genetic algorithms combined with genetic algorithms, and pattern searches were evaluated for total prediction accuracy.

Efficient Analysis of Traffic Accident Using Mining Techniques S.Vigneswaran[1]; A.Arun Joseph[2]; E.Rajamanickam[3],march 2014, Two machine-learning algorithms are used in this study to classify injury severity in traffic accidents and forecast how severe the injuries would be. Classification of injury severity in various traffic incidents was evaluated using Nave Bayesian classifier and J48 decision tree Classifier, and the results reveal that J48 beats Nave Bayesian.

ANALYSIS OF ROADWAY FATAL ACCIDENTS USING ENSEMBLE-BASED META-CLASSIFIERS Waheeda Almayyan, July 2020,

A great deal of focus has been placed on improving the safety of public roadways during the last several decades. Road traffic data must be analysed with the use of data mining in order to determine the contributing variables to fatal accidents. The Fatality Analysis Reporting System (FARS) dataset is analysed in this research utilising a variety of data mining techniques. Four meta-classifiers and four data-oriented approaches, all based on Random Forest classifier, are compared here for their performance in handling unbalanced datasets. As part of our research, we examine the impact of a variety of feature selection techniques, including PSO (Poisson sampling), Cuckoo, and Bat, on classification accuracy and efficiency. The

Threshold selection meta-classifier paired with over-sampling approaches yielded extremely promising results. Overall Accuracy is 91% and Balanced Accuracy ranges from 86% to 100% employing 7-15 features instead of the original 50; this is the suggested technique's Accuracy in this regard.

The Naive Bayes Algorithm was used to predict road traffic. RTICCT – 2019, E. Baby Anitha1, R. Aravinth2, S. Deepak3

Predicting the speed of moving traffic is a difficult challenge in an ITS and one that is receiving increasing attention. Both government bodies in charge of transportation and everyday commuters are deeply concerned about the safety of roadway traffic. Road traffic data must be carefully analysed in order to identify factors that are closely associated with deadly accidents. The Topic-Enhanced Gaussian Process Aggregation Model (TEGPAM) is a probabilistic framework that combines three separate components: location disaggregation, traffic topic modelling, and traffic speed modelling. These models attempt to overcome these issues by integrating new-type data with traditional data. Transportation authorities and residents alike are concerned about the safety of roadways. Traffic problems and answers can be found in our project. Traffic data is uploaded and preprocessed. After removing the null values from the dataset, it is saved to SQL as a preprocessed dataset. Using the Naive Bayes algorithm, we are able to pinpoint the specific time and date of traffic incidents and accidents in order to improve our safety training.

Using a Hybrid Genetic Algorithm to Evaluate Accidental Death Records International Conference on Innovative Computing and Communication, Nikhil Sharma, Kaushik, Rajat Rathi, Santosh Kumar, Nikhil Sharma, Kaushik, Rajat Rathi, Rajat Rathi, Rajat Rathi, Rajat Rathi, Rajat Rathi, Rajat Rathi (ICICC-2020),

This study proposes a strategy for optimising outcomes that combines support vector machine regression with linear regression. The Government of India has collected many subcategories of unintended death data due to natural disasters. In order to calculate the rate of error, the mean square error was used, which was assessed on the Python platform. For both male and female applicants, age was used as one of the differentiating criteria.

A comparative study on machine learning based algorithms for prediction of motorcycle crash severity Lukuman WahabID, Haobin Jiang,April 4, 2019,

For the three machine learning algorithms, the findings demonstrate that machine learning algorithms outperform the MNLM in terms of accuracy and effectiveness, and the RF-based algorithms are best for their global optimization and capacity to extrapolate. Study participants were shown to be much more likely to be injured in motorcycle crashes if they were involved in a collision at a specific location and time of day, as well as if their collision partner was also involved.

Overview of traffic incident duration analysis and prediction Ruimin Li, Francisco C. Pereira and Moshe E. Ben-Akiva,European Transport Research Review (2018),

I Some of the data resources and characteristics include different traffic incident time phases, data set sizes, incident classifications, and length time distribution. traffic incident duration prediction and evaluation methods and data resources and characteristics, including significant influence factors and unobserved heterogeneity and randomness; (v) data resources and characteristics, including significant influence factors and unobserved heterogeneity and randomness.

Graphical Prediction of Road Accidents using Data Analysis Swetha P.C,Vaishalli. A,Devika. M,Sneha Balaji,Vol-4 Issue-2 2018 IJARIIE-ISSN(O)-2395-4396,

An important goal of this research is the development of an effective method for analysing and predicting the accident zones and their severity. The results suggest that utilising graphical forecasts, the proposed strategy may effectively reduce the occurrence of accidents.

Journal of Data Analysis and Information Processing: Imran Chowdhury Dipto1 Md Ashiqur Rahman1 Tanzila Islam2 H M Mostafizur Rahman3 In this study, Imran Chowdhury Dipto, Md Ashiqur Rahman, Tanzila Islam, and H M Mostafizur Rahman participated.

This study utilises Python and R to visualise data in preparation for training an artificial neural network. The analytical abilities of Python and R are contrasted in this paper's scope of work. Accident severity may be anticipated using an Artificial Neural Network with Multilayer Perceptron. It was determined which programming language was most suited for data visualisation, data processing,

predictive analytics, and other data-driven applications.

Interregional travel demand data quality analysis: Using matrix decomposition to extract travel patterns Makoto Tsukai, Akimasa Fujiwara, Canh Xuan Do, Makoto Tsukai, Canh Xuan Do, Makoto Tsukai, Makoto Tsukai, Mak

In this work, we employed a nonnegative matrix factorization approach to see if there were any commonalities in the travel patterns found in these different data sets. MOBI travel patterns were found to be significantly different from those of NPTS and WEB in this analysis, however there were some similarities in terms of origin–destination pairings between NPTS and WEB. There are a few issues, but they will improve with time.

The use of a genetic algorithm and machine learning to improve traffic control Prof. Fang Chen, Adriana-Simona Mihait, Senior Member of IEEE, Hai L. Vu, and Prof. In March of 2021, Adriana-Simona Mihait, IEEE Senior Member, Fang Chen, and Hai L. Vu, Greetings, Tuo Mao. Deputy Chairperson Adriana-Simona Mihait,

According to the findings of this study, an innovative approach to enhancing the timing of traffic signals at signalised urban intersections in the face of irregular traffic volumes has been devised. Rapid Machine Learning (ML) techniques are combined with trusted Genetic Algorithms (GA) in a single optimization framework to give fast and reliable options. By employing phase duration as a decision variable and the target function as our objective function, we begin with the traditional GA approach with the purpose of reducing network trip time. In order to acquire the best results, we fine-tune the GA for crossover, mutation, and the calculation of fitness A number of machine learning regression models are used to train and choose the optimal regressor for hypertuning, which we subsequently do. Our next step is to introduce a new method called BGA-ML, which combines GA and the EGDT, which is known to be the best regressor, into a single optimization framework. According to the results of the comparison and assessment, the new BGA-ML approach is substantially faster than the prior GA algorithm and may be used successfully in non-recurrent incident circumstances.

Intelligent Computation Technology and Automation Conference, Macro-prediction model of traffic accident based on neural network and genetic algorithm, QIN Liyan SHAG Chunfu, 2009

In China, the issue of road safety has become more important to the general public. The assessment indexes and primary influencing elements of road traffic safety were examined in order to properly anticipate the future state of road traffic safety. After then, the mortality toll from traffic accidents was used as an evolution index, with three variables serving as parameters: the number of vehicles, highway miles, and GDP per capita. It was decided to use a combination of a neural network and a genetic algorithm to make predictions about outcomes. To train the neural networks, we used traffic accident data from 1978 through 1998. The models were tested using data from 1999 to 2004. Results reveal that a neural network model that incorporates genetic algorithms is more accurate and generalizable than a BP neural network, which predicts accident fatality tolls at 139 thousand and 167 thousand individuals for 2010 and 2020 in the corresponding years.

Machine Learning Approach to a Mortality-Risk Prediction Model from Drunk Drivers' Road Traffic Injuries, Wachiranun Sirikul 1,2, Nida Buawangpong 3, Ratana Sapbamrer 1, and Penprapa Siviroj, all on October 8, 2021.

Deaths from alcohol-related car accidents are common in developing and emerging economies, such as Thailand. Legal limits on the efficacy of screening procedures for driver impairment may be increased by automated machine-learning algorithms. Methods: We employed machine learning methods (Gradient Boosting Classifier; MLP; RF; K-Nearest Neighbor; KNN) and a parsimonious logistic regression model (Logit) to predict the mortality risk of road traffic injury in intoxicated drivers from cross-sectional secondary datasets from 2002–2004. Drinking habits and contextual factors were all included as predictors of alcohol-related crashes. More over 4300 (92%) of the 4974 drivers in the sample survived, while just 429 (8%) succumbed to their injuries. In order to address the class imbalance, the Synthetic Minority Oversampling technique was applied. There was good-to-excellent discrimination in all of the models tested here.

Using in-depth traffic accident data, a study was conducted on the correlation of traffic accident tendency with driver characteristics. Wenguang Wu Hindawi, Lin Hu, Xingqian Bao1, Hequan Wu1,2 and Wenguang Wu Hindawi, Journal of Advanced Transportation, Volume 2020, Article ID 9084245, June 2020.

In many cases, a driver's personality has a significant role in the type of driving he or she engages in. The China In-Depth Accident Study (CIDAS) database was used to investigate the link between driver attributes and the likelihood of being involved in a traffic accident. The age, driving experience, and driving style of the driver were all investigated.

The grey cluster analysis was used to categorise the drivers into four accident risk groups using the number of victims in the incident as an assessment parameter.

The majority of drivers who have a high accident risk propensity are between the ages of 18 and 30.

. Overtaking and stealing traffic priority are the most common causes of accidents, and these occur most frequently during rush hour. Psychologically, young people tend to be aggressive, overconfident, and have an unbalanced view of their own abilities.

As expected, ordinary drivers account for the majority of incidents, which contradicts the popularly held idea that adventure style drivers are more likely to be involved in an accident. Drivers of different ages, styles, and road types are exposed to varied amounts of driving based on their age and driving experience, which is not taken into account in statistical analysis.

4. PROBLEM STATEMENT

At the city bypass Highway, the problem is that there isn't enough highway space for cars to pass through the junction at different times, resulting in traffic jams and accidents at various times of the day, according to Highway survey data and traffic analysis.

5. AIM OF PROJECT

The purpose of this research is to discover the key factors that determine the occurrence of accidents and to develop an accident prediction model based on Genetic Algorithms.

6. OBJECTIVES

It is the intention of this paper to provide a comprehensive evaluation of the current status of genetic algorithms and sophisticated data analysis approaches for predicting vehicle collisions with pedestrians. To better understand the black spot in a length of road by collecting data on age, gender, vehicle, time, place, month, and other characteristics. To utilise this data in a Genetic Algorithm, convert it to binary. Selecting a fit parent, cross-covering for mutation, and creating its offspring are all steps in this process that yield the greatest possible answer. Determine if the GA result is in line with the real data using this data. To put it into action, follow through with it.

7. RESEARCH METHODOLOGY

- Examine and classify the most often used data sources for analysis and predictive research.
- Detection of traffic events in real time
- Describe the steps used to conduct out data analysis on road accidents using genetic algorithms, including their strengths and limits.
- Using GA to predict road accidents.
- Comparing real-time and GA data to determine if they may be used to anticipate accidents.

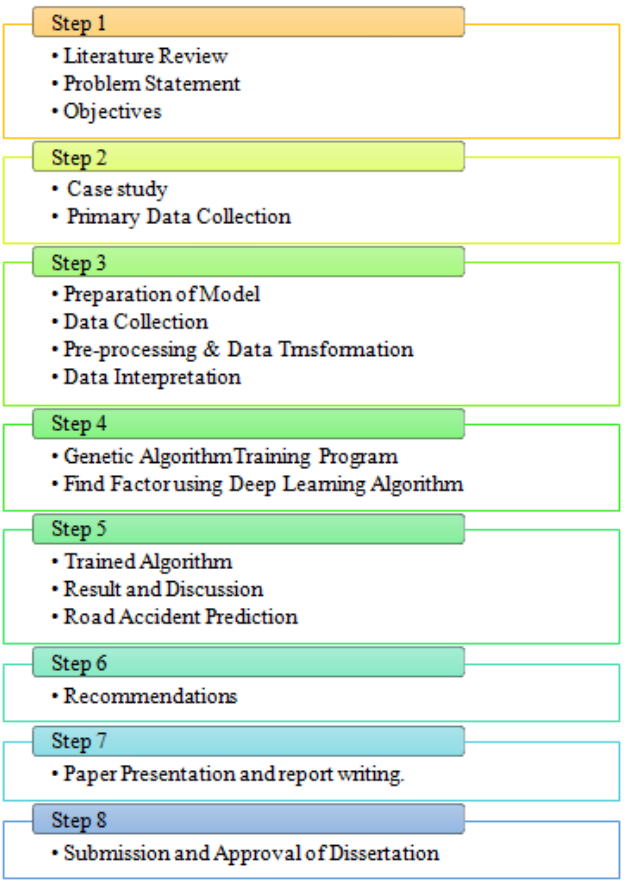

Figure 2: Methodology Flow

8. GENETIC ALGORITHM MODEL DEVELOPMENT AND APPLICATION

What is GA?

John Holland invented GAs in the 1970s. In GA, there are various terms that are explained in the following sections. Chromosomes are collections of genes that provide a specific answer to a problem.. The trait represented by each gene differs from one another. A population is a collection of chromosomes at a specific point in time. A random population is chosen as the beginning point for an iterative evolutionary process; each generation is

referred to as a generation. The application of mutation and crossover operators between separate chromosomes is used to achieve such evolution at each generation. The fittest parents or children's chromosomes are used to create the next generation of population, A replacement method is used to accomplish this. A population's evolutionary process ends when it satisfies a set of predetermined requirements.

Basic Terminology

This course on Genetic Algorithms requires you to get familiar with a few crucial terms.

The population contains all of the encoded solutions to the issue. Instead of actual humans, we have Candidate Solutions that represent human beings in a GA, which has a population equivalent to that of humans.

Problem-Solving Using Chromosomes A solution like chromosomes is an illustration of this.

chromosomal position A gene is one of the chromosome elements.

The value of a gene on a particular chromosome is known as an allele.

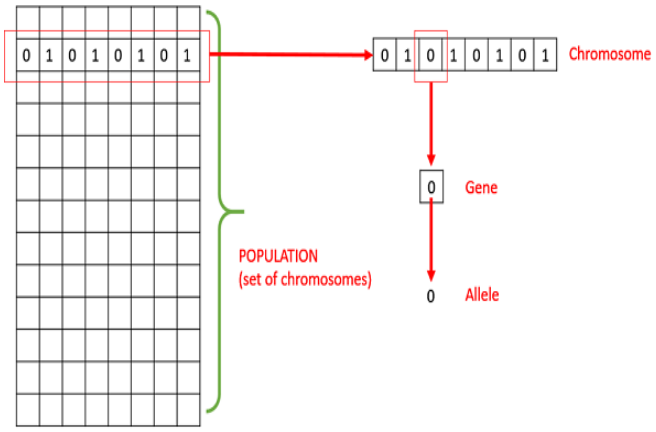

Figure 3: Basic Terminologies

Genotype: In the computing space, a genotype is a population. The solutions are represented in the computation space in a fashion that a computer machine can easily understand and manipulate.

Phenotype: The population in the actual real-world solution space in which solutions are represented as they are in real-world settings.

There is no difference between the phenotype and genotype spaces in basic cases. On the other hand, the phenotypic and genotype spaces are generally distinct in most cases. This is a difference between decoding and encoding: decoding moves a solution from the genotype space back to the phenotype.

Let's look at an example: 0/1 Knapsack Problem. Phenotype solutions consist of solely the item numbers of the items that need to be selected.

genotype-space binary string of length n, on the other hand (where n is the number of items). In this case, the item at position x has been picked, as indicated by a 0; in the other case, a 1. Genotype and phenotypic space are not the same in this circumstance.

Fitness Function: An input solution is accepted, and the appropriateness of the solution is returned as an output. a fitness feature While in some cases, the fitness function and the goal function may be the same, they may also be separate in other cases.

Genetic Operators: These have an impact on the genetic make-up of future generations. Examples of these processes include hybridization, mutation, and selection.

The Foundation: A GA's basic structure is as follows:

Selecting parents for reproduction from an initial population is the first step in the process (which may be generated at random or seeded by other heuristics). Crossover and mutation operators can be used to the parents to produce new offspring. After that, the new generation replaces the old, and the cycle repeats. An attempt is made here to mimic human evolution in certain respects.

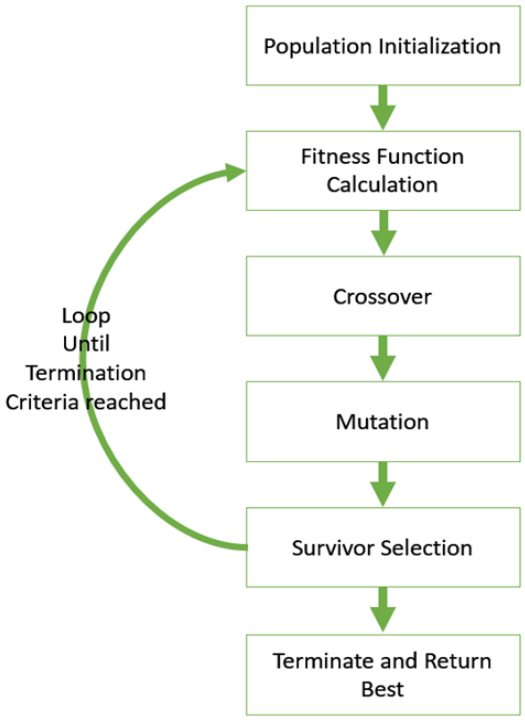

Figure 4: GA Flow Chart

Population Initialization

In a GA, there are two main ways to start a population. Those are.–

- **Random Initialization –** Use fully random solutions to populate the first population.
- **Heuristic initialization –** Use a recognised heuristic for the problem to populate the first population.

Initializing the whole population with one solution can lead to highly similar solutions and very little variety, as has been shown. Experiments have shown that random solutions are the ones that propel the population toward optimality. This means that with heuristic initializations, we seed the population with a small number of high quality solutions, and then fill the balance with random solutions.

Heuristic initialization has also been found to simply alter the population's initial fitness in some instances, but in the end, it is the variety of solutions that leads to optimality.

Fitness Function

Fitting functions are functions that take a possible solution to a problem as an input and output how "fit" or "great" the response is in respect to the circumstance at hand.

A GA's fitness value computation should be fast enough because it is performed many times. A GA can be slowed to a crawl if the fitness value is not computed quickly enough.

When trying to increase or reduce a certain objective function, the fitness function and the objective function are almost always interchangeable terms since they serve the same goal. To solve increasingly complex problems, an Algorithm Designer may utilise a different fitness function.

The following are the requirements for a fitness function:

Fast calculation of how fit an answer is or how well-suited individuals may be created from a given answer is essential to the fitness function's usefulness.

Because of the problem's intrinsic complexity, it may be hard to directly calculate the fitness function. In these circumstances, we use fitness approximation to satisfy our needs.

Selection

As a part of parent selection, it is necessary to choose parents who will mate and reproduce to generate children for the following generation. The GA's convergence rate is influenced by the quality of the parents that each individual has.

In order to avoid a solution dominating the whole population in a few generations, caution should be given to prevent solutions in the solution space from becoming too near to one another, resulting in the loss of variety. To ensure the success of a GA, it is essential to maintain a high level of variety in the population. Premature convergence is a problem in GAs because it occurs when a single, highly well-suited solution consumes the whole population.

Crossover

A Crossover Operator and its many modules will be explained in this chapter. We will also discuss its uses and benefits.

Biological crossover and reproduction can be compared to the operation of the crossover operator. One or more kids are conceived utilising the genetic material of both parents in this situation. Typically, crossover is utilised in a GA with a high possibility of success.

Genetic Algorithm Mutations

There are many different ways to think about mutations, but they all boil down to one thing: random mutations. In order to preserve and add genetic variety, it is often utilised with a low

probability – pm A extremely high probability reduces the GA to a random search.

The "exploration" of the search space is connected with the GA component of mutation. In order for the GA to converge, mutation is necessary, but crossover is not.

Operators of Mutation

In this part, we'll discuss some of the most often used mutation operators. The GA designer may benefit more from a mix of these strategies or a problem-specific mutation operator than the one given here, and this list is far from exhaustive.

Survivor Selection Genetic Algorithms

In accordance with the Survivor Selection Policy, some survivors are evicted while others are retained for the sake of future generations. Maintaining demographic variety while protecting the healthiest members of the population is essential.

Some GAs employ elitism in their rhetoric. According to this, a person's genes are passed down from generation to generation until they die out. To put it another way, the present population's fittest individual can never be replaced.

Even though it is the easiest strategy to just remove individuals of the population at random, this technique usually leads to convergence concerns, hence the following procedures are commonly used:

Condition of Termination

The termination condition of a Genetic Algorithm plays an important role in defining when a GA run will come to a finish. After an initial burst of rapid advancement, when better solutions seem to appear on a seemingly endless stream of iterations, the GA appears to plateau in its later phases, where incremental gains become increasingly inconsequential. Ending with a condition that assures our solution is almost ideal is what most of us are looking for.

- When the population has not improved after X iterations.
- After a certain number of generations have passed.
- A predetermined threshold is reached when the objective function value approaches.

Genetic Algorithms - Application Areas

Aside from optimization issues, Genetic Algorithms are often employed in a wide range of different fields.

Genetic Algorithms are extensively employed in a variety of fields. These are –

- **Optimization −** When attempting to maximise or decrease the value of an objective function while also taking into account a particular set of restrictions, genetic algorithms are frequently employed. Throughout the lesson, the method for solving optimization issues has been discussed.
- **Economics −** There are a number of applications for GAs, such as the cobweb model of equilibrium resolution and asset pricing in game theory.
- **Neural Networks −** In addition to neural networks, GAs may train recurrent neural networks (RNN).
- **Parallelization −** Additionally, GAs have excellent parallel capabilities and have been shown to be highly useful tools for resolving specific issues.
- **Image Processing −** The dense pixel matching task, for example, is carried out using GAs.
- **Vehicle routing problems −** There are several soft times, depots, and a diversified fleet to keep track of.
- **Scheduling applications −** Another difficulty that GAs are used to tackle is the problem of time tabling.
- **Machine Learning −** GBML, or genetics-based machine learning, is an advanced topic in the field of machine learning.
- **Robot Trajectory Generation −** A robot arm's trajectory can be planned using GAs in the past, but this is no longer necessary.
- **Parametric Design of Aircraft −** By altering the settings, GAs have been utilised to develop better solutions for aircraft design.
- **DNA Analysis −** DNA structure has been determined using GAs and spectrometric data from a sample.
- **Multimodal Optimization −** In multimodal optimization, where we must identify several optimal solutions, GAs are unquestionably excellent techniques.
- **Traveling salesman problem and its applications −** This combinatorial problem, known as the TSP, has been solved by GAs employing unique crossover and packing procedures.

9. CONCLUSION

Based on input data and injury severity levels, this GA study will be useful in predicting the severity of traffic accidents. This data will be combined with data gathered along a specific route to determine

the severity of the accident. To see if GA outperforms other methods for calculating accidents;

10. ACKNOWLEDGMENT

We express our sincere thanks to Project Guide Dr. R.R. Sorte for his continuous support. We also thankful to our Head of Department of civil Dr. R.R. Sorte For support

REFERENCES

[1]. Seyed Hessam-Allah Hashmienejad & Seyed Mohammad Hossein Hasheminejad (2017) Traffic accident severity prediction using a novel multi-objective genetic algorithm,International Journal of Crashworthiness.

[2]. Graphical Prediction of Road Accidents using Data Analysis Swetha P.C,Vaishalli.A,Devika.M,Sneha Balaji,Vol-4 Issue-2 2018 IJARIIE.

[3]. Camilo Gutierrez -Osorio*, Cesar Pedraza Departamento de Ingenieri´a de Sistemas e Industrial, Universidad Nacional de Colombia, Bogota, Colombia,15 May 2020“Modern Data Sources And Techniques For Analysis And Forecast Of Road Accidents: A Review”Journal of traffic and transportation engineering,

[4]. QiangLu1 andKyoung-DaeKim2, 11 October 2017,“A Genetic Algorithm Approach For Expedited Crossing Of Emergency Vehicles In Connected And Autonomous Intersection Traffic”Journal of Advanced Transportation Volume 2017, Article ID 7318917, 14 pages.

[5]. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods,Mehmet Metin Kunt , Iman Aghayan & Nima Noii,21 March 2014, At: 06:19 Publisher: Taylor & Francis

[6]. Predicting Crash Injury Severity with Machine Learning Algorithm Synergized with ClusteringTechnique: A Promising Protocol Khaled Assi, Syed Masiur Rahman , Umer Mansoor and Nedal Ratrout ,27 June 2020

[7]. A stable and optimized neural network model for crash injury severity prediction Qiang Zeng, Helai Huang ,Received 29 May 2014

[8]. A cluster analysis on road traffic accidents using genetic algorithms Sabariah Saharan and Roberto Baragona,2017

[9]. Application of Support Vector Machine for Crash Injury Severity Prediction: A Model Comparison Approach Iman Aghayan[1], Mansour Hadji Hosseinlou[2] , Mehmet Metin Kunt[3],Journal of Civil Engineering and Urbanism Volume 5, Issue 5: 193-199; September 25, 2015

[10]. Efficient Analysis of Traffic Accident Using Mining Techniques S.Vigneswaran[1] A.Arun Joseph[2]; E.Rajamanickam[3],march 2014

[11]. ANALYSIS OF ROADWAY FATAL ACCIDENTS USING ENSEMBLE-BASED META-CLASSIFIERS Waheeda Almayyan, July 2020

[12]. Prediction of Road Traffic using Naive Bayes Algorithm. Baby Anitha[1],R. Aravinth[2], S. Deepak[3] ,RTICCT – 2019

[13]. Evaluation of Accidental Death Records Using Hybrid Genetic Algorithm Nikhil Sharma,Ila Kaushik,Rajat Rathi,Santosh Kumar,International Conference on Innovative Computing and Communication (ICICC-2020)

[14]. A comparative study on machine learning based algorithms for prediction of motorcycle crash severity Lukuman WahabID, Haobin Jiang,April 4, 2019

[15]. Overview of traffic incident duration analysis and prediction Ruimin Li, Francisco C. Pereira and Moshe E. Ben-Akiva,European Transport Research Review (2018)

[16]. Graphical Prediction of Road Accidents using Data Analysis Swetha P.C,Vaishalli.A,Devika.M,Sneha Balaji,Vol-4 Issue-2 2018 IJARIIE-ISSN(O)-2395-4396

[17]. Prediction of Accident Severity Using Artificial Neural Network: A Comparison of Analytical Capabilities between Python and R Imran Chowdhury Dipto1, Md Ashiqur Rahman1, Tanzila Islam2, H M Mostafizur Rahman3,Journal of Data Analysis and Information Processing, 2020,August 7, 2020

[18]. Data quality analysis of interregional travel demand: Extracting travel patterns using matrix decomposition☆ Canh Xuan Do, Makoto Tsukai, Akimasa Fujiwara a,2020

[19]. Boosted Genetic Algorithm using Machine Learning for traffic control optimization Tuo Mao, Adriana-Simona Mihait Senior Member, IEEE, Fang Chen, and Hai L. Vu.,march 2021

[20]. Macro prediction model ofroad traffic accident based on neural network and

genetic algorithm,QIN Liyan SHAG Chunfu,2009

[21]. Mortality-Risk Prediction Model from Road-Traffic Injury in Drunk Drivers: Machine Learning Approach,Wachiranun Sirikul, Nida Buawangpong, Ratana Sapbamrer 1 and Penprapa Siviroj,8 October 2021,

Corresponding Author

Ms. Nikita Ingawale*

PG Student, Department of Civil Engineering, TSSM'S Padmabhooshan Vasantdada Patil Institute Technology,Pune