

Soft Computing Techniques for Mishaps Prediction

Dinesh Kumar Saini

Faculty of Computing and IT, Sohar

University, Oman

Faculty of Engineering and IT, University of
Queensland, Australia

Jabar H Yousif

Faculty of Computing and IT, Sohar

University, Oman

ABSTRACT

An attempt is made to implement the soft computing techniques in the prediction of the mishaps behaviors. The main objective of this paper is to implement the Multilayer Perceptron (MLP) neural network topology in mishap analysis and prediction. Efforts are made to summarize past research on the road mishaps causes and their analysis techniques and point out their weaknesses. Indicate new ideas in this area and Identify research directions leading to successful development and use of analytical techniques in the area of mishap analysis leading to Road safety

General Terms

Mishaps, Computer Science, Safety System, Soft Computing, ITS Systems,

Keywords

Artificial Intelligence, Neural Networks Multi-layered Perceptron (MLP), Fuzzy ARTMAP, Predictions,

1. INTRODUCTION

Soft computing techniques can be used to study the road safety which is a major area of concern for every country's government and people residing in that country [1]. The need for mishap analysis arrives because of the reason that number of vehicles are increasing day by day and loads on the transportation system in increasing heavy [2,3]. The reasons and types of accidents depend largely upon the traffic density, road network system and average traffic speed on highways, environmental conditions and population of the city/region/country [4].

2. RELATED WORK

There is work going on various aspects of mishaps and accidents, researchers are using various approaches for studying the mishaps [5, 6]. The accident is studied by many researchers and compared the performance of Multi-layered Perceptron (MLP) and Fuzzy ARTMAP, and found that the MLP classification accuracy is higher than the Fuzzy ARTMAP. Neural networks to analyse vehicle accident that occurred at intersections [7].

They chose feed-forward MLP using BP learning. The model had 10 input nodes for eight variables (day or night, traffic flows circulating in the intersection, number of virtual conflict points, and number of real conflict points, type of intersection, accident type, road surface condition, and weather conditions). Shankar, et al. [14] suggested a nested logic formulation for estimating accident severity likelihood conditioned on the occurrence of an accident.

3. SOFT COMPUTING TECHNIQUES IN ANALYSIS

There is an increased use of data collection and storage on every step of life in today's time. There is a sufficient scope of discovering new trends and predicting mishap models to aid the process of reduction of accidents in the Sultanate of Oman [8, 9, and 10].

This can be achieved by making use of the latest analytical techniques in data mining and artificial intelligence. Increased costs of deaths and injuries resulting from traffic accidents have a major impact on society. This prompts the researchers to make great efforts in order to determine the factors that significantly affect the traffic accidents. Several researchers have been addressing this problem of the mishaps using artificial intelligent techniques.

These include neural network, nesting logic formulation, log-linear model, fuzzy ART maps, etc. Applying soft computing techniques to formulate traffic accident data can facilitate to comprehend the characteristics of driver's behaviour, road condition and climate condition that were causally connected with different injury harshness [11, 12, 13 and 14].

This can help decision makers to improve traffic safety control policies. This paper proposed the using of soft computing techniques like MLP neural networks and hybrid techniques to build models that could effectively predicate the number of traffic mishap.

Neural networks are implemented in the areas that traditionally have been used statistical methods. A neural network is a powerful data modelling tool that is able to capture and represent complex input/output relationships both linear and non-linear one [15].

The artificial neural network is a computational model that is motivated by the structure and/or functional aspects of biological neurons. It is considered as one of the most techniques for the learning from the rare data [16].

There are specific features that can stimulate scientist to adopt neural network based solution in solving problems [17]. Their main features are massive parallelism, uniformity, generalization ability, distribution representation and computation, learnability, trainability and adaptivity. Neural approaches have been performed successfully in many aspects of artificial intelligence such as prediction and classification, image processing, NLP, speech recognition, and pattern recognition. [18,19,20]. Figure 1 shows the Pre-accident and Post.

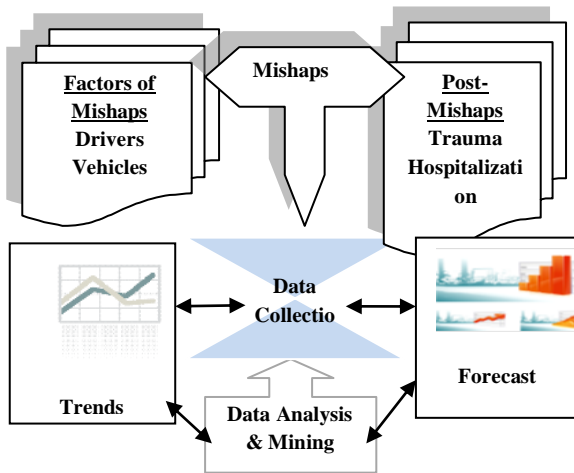


Figure 1 – Pre-accident and Post

4. FACTORS FOR THE MISHAPS

Identify the need and scope of implementation of analytical techniques in mishap analysis and reduction. Propose new methods of analytical of data and the used of soft computing techniques which will be effective for mishap analysis and prediction [21,22, and 23].

➤ ITS systems

Speed cameras play an important role in the road safety measures taken by the transport authorities. It is also important to assess whether the use of speed cameras reduces the incidence of speeding, road traffic crashes, injuries and deaths

➤ Driver behaviour

Driver behaviour is an important aspect of accident occurrence and its causes. It is very important to understand driver needs [7]. This study focuses on the difficulties faced by the drivers when they meet with an accident.

➤ Effect of Drugs on drivers

Effect of drugs/alcohol on drivers which adversely affect their driving skills is one of the serious causes leading to road accidents.

➤ Reduction in severity of injuries due to road accidents

There are many theories which aim to reduce the severity of injuries happening due to road accidents on the highways.

➤ Effect of weather conditions

Effect of weather and changes in climatic conditions within a country has a significant influence on the occurrence of road accidents.

➤ Accident prediction models

Information on the incidents of crashes, traffic flow, geometric design, road environment, and traffic control helps develop crash prediction models such as Poisson and negative binomial regression models, to investigate the associations between crash occurrence and explanatory factors.

➤ Multilevel modelling for the regional effect of enforcement on road accidents:

The use of multilevel models for the analysis of road traffic accident outcome data, and has illustrated the

$$e_i(n) = d_i(n) - y_i(n), \dots (1)$$

With the aim of adaptation of each weight in the network, the theory of gradient learning techniques is used. It corrects the present value of the weight as depicted in formula 2.

application of the methodology to a study of seriously and fatally injured casualties.

5. CASE STUDY OF DATA OF ACCIDENTS

Table 1.0 below shows the statistics of accident data for the last 10 years, obtained from Royal Oman Police. Along with this, a lot of data needs to be collected and analysed by using data mining techniques. It can thus be made useful by trying to find trends and situations which lead to occurrence of road mishaps. Hereby, accident rate can be reduced by providing information and encouraging authorities to take necessary steps.

6. IMPLEMENTATION OF SOFT COMPUTING TECHNIQUES

6.1. MLP Configuration & Design

Supervised learning techniques are used for classification of the given set of data and it is the same set of data used for learning. Multilayer Perceptron (MLP) neural network topology is used in the analysis of behaviour of mishaps. The discriminate functions take shapes which are not definite but it gives the shapes according to the input data set cluster. Normalization of input data set is used to achieve a high performance of the maximum a posteriori receiver. The output classes gives values of 0/1 clusters, which is optimal from a classification point of view [24,25]. Back Propagation learning algorithm is used for purpose of learning and generalization of the network, and it permits to accomplish the adaptation of weights in the layers of PEs. Accident analysis is carried out in order to conclude the causes of an accident which can be helped to prevent further incidents of a similar kind. It may be achieved by a number of experts, including forensic scientists, forensic engineers or health and safety advisers. The analysis includes the coding of the levels of selected variables of the accident data. Besides, specifying the membership value of the data using cluster analysis techniques. Furthermore, determine the main factors causing the accident using ANN and predicting future values. [19]

In this paper, we are used Neuro solution software to implement a multilayer perceptron which have one hidden layer, 11 PEs as input, 4 PEs as output, data set of 200 pairs (year, type of mishap). The maximum number of epochs is 1000. The TanhAxon as transfer function is implemented in hidden and output layer. The TanhAxon function applies a bias and Tanh formula to each neuron in the layer. This will flatten the range of each neuron in the layer to be between [-1 and 1]. Such nonlinear elements provide a network with the ability to make soft decisions.

The back propagation learning algorithm is used to propagate the errors through the network and permits to accomplish the adaptation of weights in the hidden PEs. The error correction learning works in the response at PE i at iteration n , $y_i(n)$, and the preferred result $d_i(n)$ for a given input pattern an instantaneous error $e_i(n)$ is defined by the formula 1.

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n) x_j(n) \dots (2)$$

Where the local error $\delta_i(n)$ is computed from $e_i(n)$ at the output PE or can be computed as a weighted sum of errors at the internal PEs. The constant step size is η .

For the sake of an improvement to the straight gradient descent, the Momentum learning is used which is speed up and stabilize convergence of network. The update of the weights of momentum learning computed as in formula 3.

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n) + x_j(n) + \alpha(w_{ij}(n) - w_{ij}(n-1))$$

The best value of α the momentum value is to be set between 0.1 and 0.9.

6.2. DATASET COLLECTION

This research used datasets which is produced by the Royal Police of Oman (Which can be reached at www.ropoman.org). This datasets are intended to be a nationally representative probability sample from the annual estimated around thousands of accident reported in Oman. The dataset for the study contains traffic accident records of a decade from 2000 to 2010. According to the variable definitions for dataset, this dataset has mishaps cause's records only and does not include passenger's information. It includes labels, which are listed in the table 1.

6.3. Data Preparation and Interpretation

The variables are already categorized and represented by numbers of Mishaps. In fact we are predicting the mishap factors verses the causes. In Figure 2 MSE Epoch is coming very close to the real values which validate our claims. In the figure 3 the number of real collision is coming out very close to the output of our network, which means it gives us the correct values and indicates that we can predicate the number of collision in the future. Figures 4, 5, and 6 suggest that the output of our network is very close to the real data of the run over, overturn and collision with the fixed object, which give clear indicates that we can predicts the number of mishaps happening at any time because of these factors.

6.4. experiments

With the given set of experiments it is observed that in 2004, 2005 the median is going up, in year 2010, 2009, 2008 the average is getting down and in year 2002, 2006, 2005 the average shows it is static, but in year 2000, and 2001 it is observed to be moving very high as depicted in Figures 8, 9, 10, 11, 12, 13.

7. CONCLUSION

The attributes used in this test are collisions, Run over, Overturn and Collision with the fix Objects. The table 2 illustrates the best network results for the training and cross validation phases of the network. It clearly shows that the MSE is very reasonable. The Final MSE is also coming very reasonable. In the paper we analyzed the causes of the mishaps and the data set for the study contains accidents records from year 2000 to 2010. The data of main factors like collisions, Run over, Overturn and Collision with the fix Objects are taken from Royal Oman Police and investigated the performance of Multi-Layer Perceptron Networks for classifying and predicting classification accuracy.

Table-1 the best network results for the training and cross validations

| Best Networks | Training | Cross Validation |
|---------------|-------------|------------------|
| Epoch # | 788 | 934 |
| Minimum MSE | 0.003086226 | 0.004111095 |
| Final MSE | 0.003437341 | 0.004912488 |

Sensitivity analysis is very useful technique to find out the impact of independent variables on dependent variables with some assumptions.

Table 2 Sensitivity with respect to the factors responsible

| Sensitivity | Collision | Run over | Overturn | Collision with Fixed object |
|-------------|-----------|----------|----------|-----------------------------|
| 2010 | 3237.223 | 85.31246 | 115.2820 | 262.8193 |
| 2009 | 4015.667 | 251.4350 | 275.9089 | 566.8022 |
| 2008 | 1285.743 | 18.91464 | 31.02104 | 77.03896 |
| 2007 | 223.832 | 49.75236 | 50.41868 | 85.67295 |
| 2006 | 1618.925 | 302.4082 | 276.0734 | 421.8413 |
| 2005 | 76.8680 | 211.0761 | 202.2462 | 331.7376 |
| 2004 | 283.569 | 71.21879 | 66.30162 | 100.7760 |
| 2003 | 771.3364 | 58.70114 | 49.42318 | 63.40469 |
| 2002 | 114.6104 | 17.50180 | 15.86065 | 23.09528 |
| 2001 | 4486.860 | 421.5736 | 425.7617 | 861.9401 |
| 2000 | 2235.556 | 604.9243 | 555.2740 | 1063.997 |

In order to predict the output under different situation as compared to the expected predictions, the best suited technique is sensitivity analysis. From our analysis about sensitivity of the factors it is concluded that collision is the main factor of mishaps which is affecting the other factors of mishaps. The other finding is that collision with fix object also has greater sensitivity in comparison with other factors that also leads to increase mishaps.

8. FUTURE DIRECTION

The Genetic Algorithm optimization techniques can be used for optimizing the feature selection to reduce the dimensionality of the dataset. This can help to reduce the time of training of the network. The research can be extended to implement the more factors like different cases such as Mishaps, Casualty and Vehicle for finding the cause of mishaps and the severity of accident. The other factors can be considered are like tyre burst and other factors.

9. REFERENCES

- [1] R Andrew P. Jones and Stig H. Jorgensen, "The use of multilevel models for the prediction of road accident outcomes," Accident Analysis and Prevention 35 (2003) 59–69.
- [2] Anna Andersson and Lee Chapman "The use of a temporal analogue to predict future traffic accidents and winter road conditions in Sweden," Meteorol. Appl. 18: 125–136 (2011).

- [3] Becky P.Y. Loo, S.C. Wong , W.T. Hung, Hong K. Lo, "A review of the road safety strategy in Hong Kong," *Journal of Advanced Transportation*, Vol. 41, No. 1, pp. 3-37 (2005).
- [4] Christopher A. Alford, "Sleepiness, countermeasures and the risk of motor vehicle accidents," *Drugs, Driving and Traffic Safety* © 2009 Birkhäuser Verlag/Switzerland.
- [5] Dai, Ying, Basabi Chakraborty and Minghui Shi. "Kansei Engineering and Soft Computing: Theory and Practice." IGI Global, 2011. 1-436. Web. 19 Mar. 2012. doi:10.4018/978-1-61692-797-4
- [6] Eleni Petridou & Maria Moustaki, "Human factors in the causation of road traffic crashes," *European Journal of Epidemiology* 16: 819-826, 2000.
- [7] Erin Kelly, Shane Darke and Joanne Ross, "A review of drug use and driving: epidemiology, impairment, risk factors and risk perceptions," *Drug and Alcohol Review* (September 2004), 23, 319 – 344.
- [8] Florence Conche, Miles Tight, "Use of CCTV to determine road accident factors in urban areas," *Accident Analysis and Prevention* 38 (2006) 1197–1207.
- [9] Guy D. Vitaglione, "Driving under the influence (of mass media): A four-year examination of NASCAR and West Virginia aggressive-driving accidents and injuries," *Journal of Applied Social Psychology* (2006).
- [10] Helai Huang, Hong Chor Chin, "Modeling road traffic crashes with zero-inflation and site-specific random effects," *Stat Methods Appl* (2010) 19:445–46.
- [11] Jabar H. Yousif, "INFORMATION TECHNOLOGY DEVELOPMENT", LAP LAMBERT Academic Publishing, Germany ISBN 9783844316704, 2011.
- [12] Jabar H. Yousif, and Mabruk A. Fekihal, "Neural Approach for Determining Mental Health Problems" , *Journal of Computing*, Volume 4, Issue 1, pp6-11, January 2012.
- [13] Kathleen Beullens, Keith Roe, Jan Van den Bulck, "Excellent gamer, excellent driver? The impact of adolescents' video game playing on driving behavior: A two-wave panel study," *Accident Analysis and Prevention* 43 (2011) 58–65.
- [14] Kenneth M. Bach, "Evaluating driver attention and driving behaviour: comparing controlled driving and simulated driving," Published by the British Computer Society, pg 193-201 (2008).
- [15] Masahiro TADA, Haruo NOMA, Kazumi RENGE, "Evidence-based automated traffic hazard zone mapping using wearable sensors," *ACM* 978-1-4503-0414-6/10/11, Japan (2010).
- [16] Mika Moran, Orna Baron-Epel, Nassim Assi, "Causes of road accidents as perceived by Arabs in Israel: A qualitative study," *Transportation Research Part F* 13 (2010) 377–387.
- [17] Patrick Brézillon and Juliette Brézillon, "Ch. 42: Context-sensitive decision support systems in road safety," part of the ACC project supported by PREDIT and the French Minister of Transportation.
- [18] Paul Atchley, Stephanie Atwood, Aaron Boulton, "The choice to text and drive in younger drivers: Behavior may shape attitude," *Accident Analysis and Prevention* 43 (2011) 134–142
- [19] Poul Greibe, "Accident prediction models for urban roads," *Accident Analysis and Prevention* 35 (2003) 273–285.
- [20] Dinesh Kumar Saini, Lingaraj A. Hadimani and Nirmal Gupta "Software Testing Approach for Detection and Correction of Design Defects in Object Oriented Software" *Journal of Computing*, Volume 3, Issue 4, April 2011, ISSN 2151-9617, Page No. 44-50.
- [21] Samer H. Hamdar and Hani S. Mahmassani, "Colliding particles: Beyond accident-free car following models," research funded by the National Science Foundation's Human and Social Dynamic Systems under Award Number 0527573.
- [22] Dinesh Kumar Saini "A Mathematical Model for the Effect of Malicious Object on Computer Network Immune System" *Applied Mathematical Modeling*, 35(2011) Page No. 3777-3787 USA, doi:10.1016/.2011.02.025.
- [23] Sharon Clarke and Ivan T. Robertson, "A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings," *Journal of Occupational and Organizational Psychology* (2005), 78, 355–376.
- [24] Taimur Usman, Liping Fu, Luis F. Miranda-Moreno, "Quantifying safety benefit of winter road maintenance: Accident frequency modeling," *Accident Analysis and Prevention* 42 (2010) 1878–1887.
- [25] Wilson C, Willis C, Hendrikz JK, Le Brocq R, Bellamy N., "Speed cameras for the prevention of road traffic injuries and deaths", *Cochrane database of systematic reviews* 2010, Issue 11. Art. No.: CD004607.
- [26] Dinesh Kumar Saini "Testing Polymorphism in Object Oriented Systems for improving software Quality" *ACM SIGSOFT Volume 34 Number 2 March 2009*, ISSN: 0163-5948, USA.

Table-3 the statistics of accident data for the last 10 years

| Details | 2010 | 2009 | 2008 | 2007 | 2006 | 2005 | 2004 | 2003 | 2002 | 2001 | 2000 |
|-------------------------------|---------|---------|---------|---------|---------|--------|--------|--------|--------|--------|--------|
| Collision | 3441 | 3485 | 3994 | 5000 | 6230 | 5702 | 5705 | 6361 | 5029 | 7863 | 7444 |
| Run over | 1091 | 1018 | 1100 | 990 | 965 | 961 | 982 | 1003 | 1141 | 1413 | 1514 |
| Overturn | 1160 | 1190 | 1167 | 1149 | 994 | 1023 | 1109 | 1115 | 1214 | 1523 | 1496 |
| Collision with Fixed object | 1879 | 1560 | 1721 | 1677 | 1680 | 1561 | 1664 | 1710 | 1723 | 2502 | 2584 |
| Total of traffic accidents | 7571 | 7253 | 7982 | 8816 | 9869 | 9247 | 9460 | 10187 | 9107 | 13301 | 13038 |
| Deaths | 820 | 953 | 951 | 798 | 681 | 689 | 637 | 578 | 580 | 499 | 492 |
| Injuries | 10066 | 9783 | 10558 | 8531 | 7548 | 6858 | 8636 | 6735 | 7907 | 9625 | 9323 |
| Driving Tests | 410824 | 408721 | 393796 | 336723 | 322800 | 269188 | 250400 | 233401 | 229363 | 206872 | 176827 |
| Learners passed Driving Tests | 62771 | 55630 | 53190 | 45798 | 44207 | 38269 | 35302 | 32166 | 32183 | 31699 | 30993 |
| New driving licenses | 64939 | 56589 | 54246 | 46686 | 45599 | 39878 | 34999 | 32127 | 33023 | 32917 | 31595 |
| Total no. of licenses | 909978 | 840002 | 777741 | 718697 | 667917 | 620025 | 578808 | 567953 | 541752 | 519772 | 493571 |
| New vehicles | 120662 | 127001 | 177441 | 136516 | 104891 | 73421 | 57130 | 42561 | 39376 | 42978 | 48740 |
| Routine Inspections | 345473 | 321260 | 341476 | 296466 | 259174 | 222034 | 217912 | 215225 | 204560 | 208147 | 187611 |
| Accident vehicle inspection | 13446 | 13971 | 21343 | 43229 | 52089 | 41704 | 36923 | 38451 | 38752 | 37146 | 38199 |
| Total vehicles inspected | 358919 | 335231 | 362819 | 339695 | 311263 | 263738 | 254835 | 253676 | 243312 | 245293 | 225810 |
| Val. of vehicles due to acc. | 2792 | 4173 | 2156 | 1430 | 1201 | 1080 | 1017 | 1316 | 1192 | 1529 | 1477 |
| Traffic offences | 2205623 | 2070347 | 2067173 | 1569895 | 1433862 | 476221 | 452267 | 409081 | 266792 | 210859 | 230712 |

Appendix – Graphs

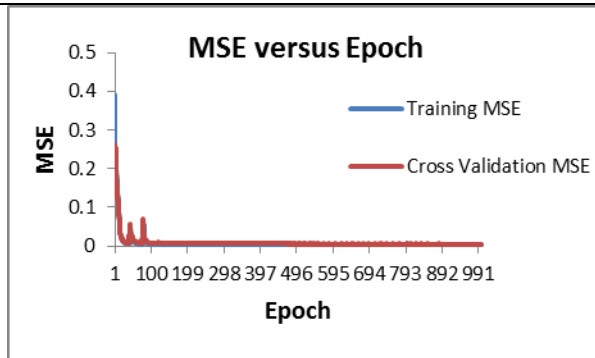


Figure2 MSE versus Epoch

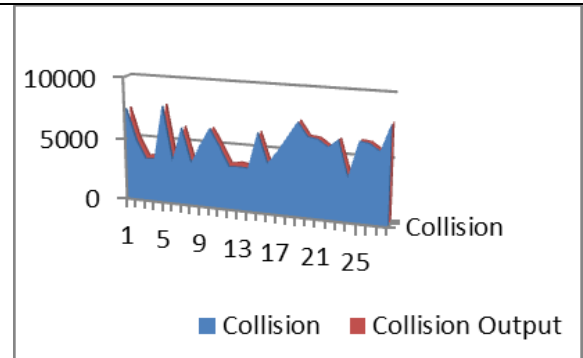


Figure-3 the results of real value of collision and network output

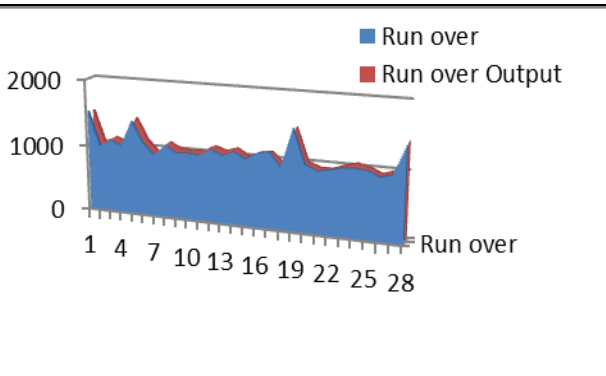


Figure-4 the results of real value of run over and network output

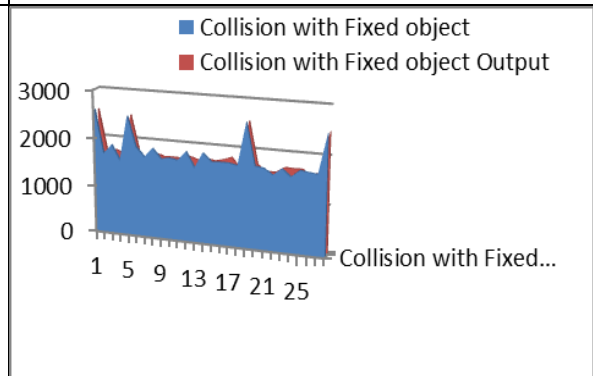


Figure-5 the results of real value of collision with fixed object and network output

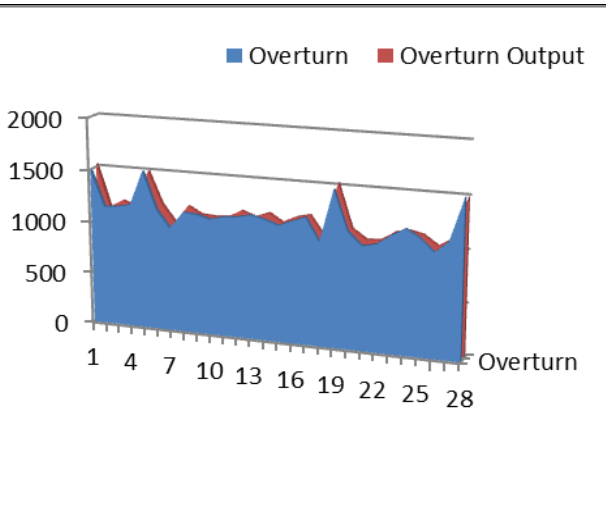


Figure-6 the results of real value of overturn and network output

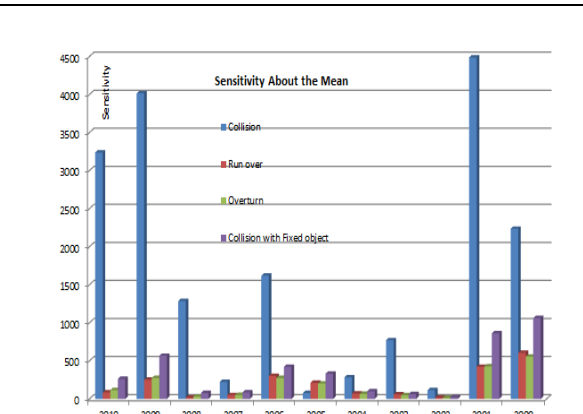


Figure-7 the results of sensitivity about the mean for all four factors.

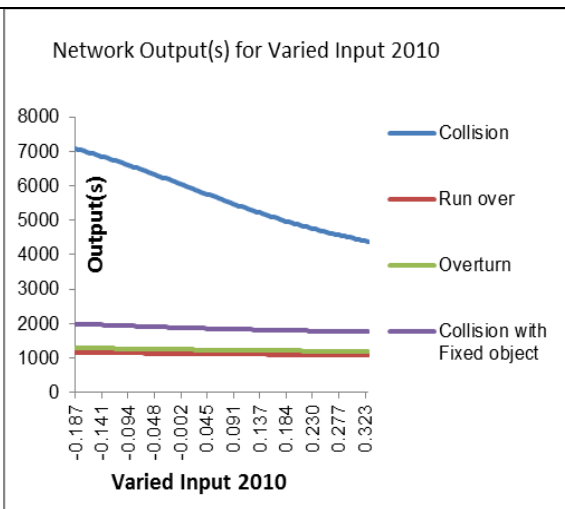


Figure-8 Varied input 2010

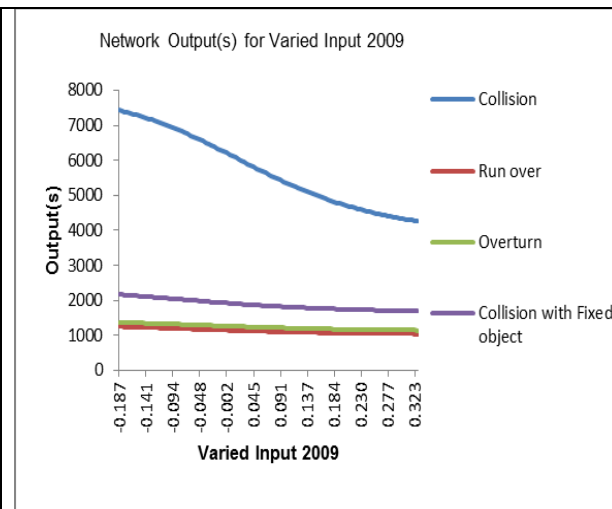


Figure-9 Varied input 2009

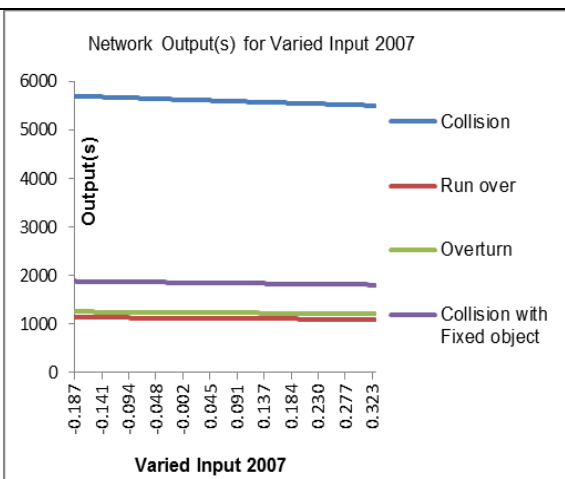


Figure-10 Varied input 2007

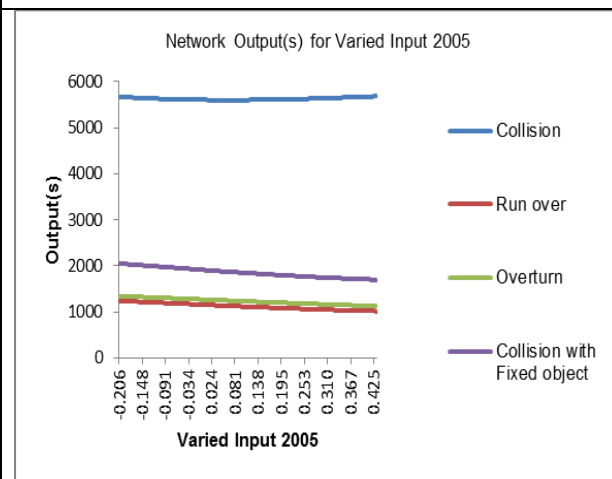


Figure-11 Varied input 2005

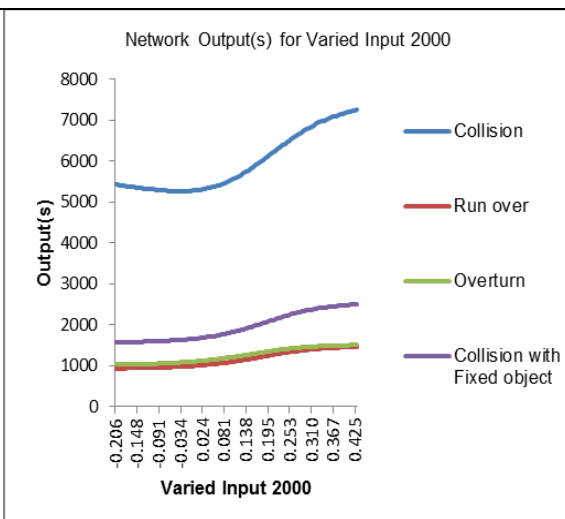


Figure-12 Varied input 2000

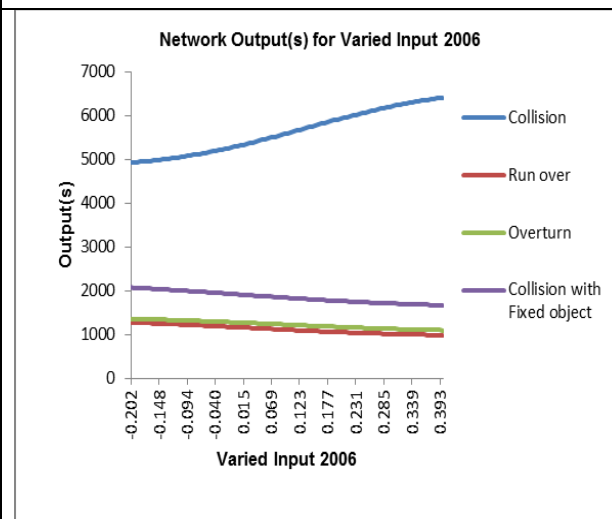


Figure-13 Varied input 2006