

Mining Traffic Accident Features by Evolutionary Fuzzy Rules

Pavel Krömer^{*†}, Tibebe Beshah[‡], Dejene Ejigu[‡], Václav Snášel^{*†}, Jan Platoš^{*†}, Ajith Abraham^{*†}

^{*}IT4Innovations, VŠB - Technical University of Ostrava

[†]Department of Computer Science, VŠB - Technical University of Ostrava,

Ostrava, Czech Republic

Email: {pavel.kromer, vaclav.snasel, jan.platos}@vsb.cz; ajith.abraham@ieee.org

[‡]IT Doctoral Program

Addis Ababa University, Addis Ababa, Ethiopia

Email: tibebe.beshah@gmail.com; ejigud@yahoo.com

Abstract—Traffic accidents represent a major problem threatening peoples lives, health, and property. Traffic behavior and driving in particular is a social and cultural phenomenon that exhibits significant differences across countries and regions. Therefore, traffic models developed in one country might not be suitable for other countries. Similarly, attributes of importance, dependencies, and patterns found in data describing traffic in one region might not be valid for other regions. All this makes traffic accident analysis and modelling a task suitable for data mining and machine learning approaches that develop models based on actual real-world data. In this study, we investigate a data set describing traffic accidents in Ethiopia and use a machine learning method based on artificial evolution and fuzzy systems to mine symbolic description of selected features of the data set.

Keywords—*machine learning; genetic programming; fuzzy rules; traffic accidents; binary classification; multi-class classification*

I. INTRODUCTION

Data mining and knowledge discovery deals with the analysis of real-world data and aims at construction of models describing the behavior encoded in the data. The main goals of such models include data classification, value prediction, dimension reduction, significant attribute selection and so on. In order to accomplish these goals, various machine learning techniques are employed. Among them, bio-inspired approaches have shown great ability to find models, structures, and patterns that accurately describe the phenomena captured in the data. In this study, a bio-inspired approach (genetic programming) is used to find a fuzzy rule [1]–[3] (i.e. a fuzzy information retrieval-based classifier) that would classify a data set describing traffic accidents in Ethiopia in the course of several years according to selected attributes.

Fuzzy classifiers form a broad family of tools, systems, and models that exploit various aspects of fuzzy set theory to mine, label, and generally process data. There are simple fuzzy classifiers as well as complex rule-based fuzzy classification systems that usually build and maintain sophisticated rule bases. The popularity of fuzzy classifiers can be attributed to their ability to perform soft classification, to assign multiple labels to data samples, and to the ease of their interpretation.

Genetic programming is a bio-inspired search and optimization method that was designed to evolve symbolic tree-

like structures in an automated manner [4]. Because of that, it is a good tool to evolve symbolic expressions such as the fuzzy rule. When compared to a more sophisticated fuzzy classifier systems, the fuzzy rule can be seen as a single expression that maps data features onto a real value from the range [0, 1]. This classification method was in this study applied to learn the description of selected attributes of the data set describing real-world traffic accidents in Ethiopia. The algorithm was used to mine fuzzy rules from the data set after some imperfections (e.g. missing values, unbalanced classes) that affected previous analyses [5] were mitigated. Understanding the magnitude of the problem it is easy to learn the need for multifaceted research and experiments. In line with this the objectives of this specific research were to mine symbolic description of selected features of the data set and compare the ability of evolutionary fuzzy rules to evolve classifiers for binary and multi-class attributes. The use of evolutionary fuzzy rules was motivated by the need for apprehension of the symbolic nature of the data set and need for easily interpretable results that could be used for design of safety measures and traffic regulations.

II. RELATED WORK

In this section, we provide a brief overview of related work in the area of traffic accident analysis and evolutionary synthesis of fuzzy classification tools and systems.

A. Traffic accident data analysis in Ethiopia and worldwide

Previously, several studies dealing with machine learning and other analytical techniques for traffic accident classification and modelling were published. Some of them were directly related to the region of Ethiopia while the others dealt with traffic accidents in other locations. In Ethiopia, the studies by Tesema, Abraham, and Grosan [6], Regassa [7], Kifle [8], Mossie [9], Tabor [10], and Beshah and Hill [11] used various methods and addressed various aspects of the traffic accident/traffic safety domain.

Tesema, Abraham, and Grosan [6] used adaptive regression trees for rule mining and classification. Regassa [7] explored classification algorithms for the study of accident severity and driver characteristics. Taking a different approach, Kifle [8] explored the application of data mining in order to identify dangerous locations in Addis Ababa. Mossie [9]

demonstrated the usefulness of several data mining models for accident severity analysis with the aim to reduce the number of road traffic accidents by identifying and predicting the major vehicles and driver's risk factors (i.e. important attributes) that contributed to road traffic accidents. Bayesian Network power predictor and constructor was employed by Tabor [10] for prediction and modelling purpose respectively in the process of two experiments which were made before and after the elicitation of the domain experiments. Beshah and Hill [11] utilized Decision Tree (J48), Naive Bayes and K-Nearest Neighbors algorithms to explain the role of road related factors for accident severity.

Outside Ethiopia, many flavours of traffic accident analyses were performed in both, developed and developing countries. Daigavane and Bajaj [12] analyzed road traffic accident data and identified important attributes including weather, poor road conditions, age of vehicle, time duration, and mechanical failure. Hongguo, Huiyong and Fang [13] explored the applicability of Bayesian Network in traffic accident causality analysis. Krishnaveni and Hemalatha [14] also conducted perspective analysis of traffic accident data using data mining techniques including selected classification models to predict severity of injury that occurred during traffic accidents. An application of Factor Analysis on Road Traffic Accident is due to Haixia and Zhihong [15]. Li et al. [16] analyzed road accident data by k-means clustering to partition highway roads in order to avoid the occurrence of accidents. Jinlin et al. [17] developed a three-layer analysis system based on spatial data mining of a GIS. Saunier, Mourji and Agard [18] investigated collision factors by mining microscopic data (road user's trajectories) of all traffic events with and without collision.

Nayak et al. [19] presented a data mining methodology using decision trees for modeling the crash proneness of road segments using available road crash data. Pakgohar et al. [20] explored the role of human factors on incidence and severity of road crashes in Iran. Chang and Wang [21] used classification and regression tree (CART) to analyze the 2001 accident data for Taipei, Taiwan. Ona, Mujalli, and Calvo [22] showed the possibility of using Bayesian Networks (BNs) to classify traffic accidents according to their injury severity. Anastasopoulos and Mannering [23] compared fixed and random parameter statistical models in order to analyze 5year data from interstate highways in Indiana, USA.

Another related study is due to Pei Liu [24], who used a self-organizing feature maps and a data mining based decision support system for liability authentications of traffic crashes in Taiwan. Delen, Sharda, and Bessonov [25] used a series of artificial neural networks to model potentially non-linear relationships between the injury severity levels and crash-related factors. Savolainen et al. [26] assessed and summarized the evolution of research and current thinking as it relates to the statistical analysis of motor-vehicle injury severities. Morgan and Mannering [27] used a mixed logic analysis to assess the effects that age, gender and other factors have on severity of crashes occurring under various circumstances. Beshah et al. [28] used CART and RandomForest to analyze the effect of 12 heuristically selected road user related variables on accident severity. In [5], Beshah et al. used the evolutionary fuzzy rules to predict the severity of traffic accidents in Ethiopia.

B. Fuzzy classification systems evolved by evolutionary algorithms

The design of fuzzy classifiers and fuzzy rule-based systems has been successfully aided by nature inspired methods in the recent years. In this section we summarize few examples of such an evolution or more generally nature inspired fuzzy classifier design. For a comprehensive survey on the automated evolution of fuzzy classification tools see e.g. [29].

Multi-objective evolutionary algorithms were used for the evolution of linguistic fuzzy rule-based classification systems in the work of Cordón et al. [30]. Another multi-objective evolutionary approach to the evolution of fuzzy rule-based systems was proposed by Ishibuchi and Nojima [31]. They used a hybrid 2-stage approach that combined an initial heuristic stage to select fuzzy rules and evolutionary stage to optimize and tune the system.

Wang et al. [32] used the genetic algorithms to integrate fuzzy rule sets and membership functions learned from various information sources. In [33], Freischlad et al. used an evolutionary algorithm to generate fuzzy rules for knowledge representation. Zhou and Khotanzad [34] used a genetic algorithm to learn various parameters of a fuzzy classification system from a training data set.

The usage of another nature inspired method - the particle swarm optimization - for fuzzy classification system design was studied recently in [35].

III. EVOLUTIONARY FUZZY RULES

This work employs genetically evolved fuzzy rules [1], [2], [5], i.e. fuzzy classifiers heavily inspired by the area of information retrieval (IR). In the IR, extended Boolean IR model utilizes fuzzy set theory and fuzzy logic to facilitate flexible and accurate search [36]. It uses extended Boolean queries that contain search terms, operators, and weights and evaluates them against an internal representation (index) of a collection of documents. Fuzzy rules use similar data structures, basic concepts, and operations and apply them to general data processing such as classification, prediction, and so forth.

The database used by the fuzzy rule is a real valued matrix. Each row of the matrix corresponds to a single data record which is interpreted as a fuzzy set of features. Such a general real valued matrix D with m rows (data records) and n columns (data attributes, features) can be mapped to an IR index that describes a collection of documents.

The fuzzy rule has the form of a weighted symbolic expression roughly corresponding to an extended Boolean query in the fuzzy IR analogy. The predictor consists of weighted feature names and weighted aggregation operators. The evaluation of such an expression assigns a real value from the range $[0, 1]$ to each data record. Such a valuation can be interpreted as an ordering, labeling, or a fuzzy set induced on the data records.

The fuzzy rule is a symbolic expression that can be parsed into a tree structure. The tree structure consists of nodes and leafs (i.e. terminal nodes). In the fuzzy rule, three types of terminal nodes are recognized:

- a) *feature* node which represents the name of a feature (a search term in the IR analogy). It defines a requirement on a particular feature in the currently processed data record.
- b) *past feature* node which defines a requirement on certain feature in a previous data record. The index of the previous data record (current - 1, current - 2 etc.) is a parameter of the node.
- c) *past output* node which puts a requirement on a previous output of the predictor. The index of the previous output (current - 1, current - 2) is a parameter of the node.

An example of fuzzy rule written down using a simple infix notation is given below:

feature1:0.5 and:0.4 (feature2[1]:0.3 or:0.1 ([1]:0.1 and:0.2 [2]:0.3)

where *feature1:0.5* is a feature node, *feature2[1]:0.3* is a past feature node, and *[1]:0.5* is a past output node. Different node types can be used when dealing with different data sets. For example, the past feature node and past output node are useful for the analysis of time series and data sets where the ordering of the records matters, but their usage is pointless for the analysis of regular data sets. The feature node is the basic building block of predictors developed for any type of data.

The operator nodes supported currently by the fuzzy rule are *and*, *or*, *not*, *prod*, and *sum* nodes but more general or domain specific operators can be defined. Both nodes and leafs are weighted to soften the criteria they represent.

The operators *and*, *or*, *not*, *prod*, and *sum* can be evaluated using fuzzy set operations. Fuzzy set operations are extensions of crisp set operations on fuzzy sets [37]. They are defined using the characteristic functions of operated fuzzy sets [38]. In [37] L. Zadeh defined basic formulas to evaluate the complement, union, and intersection of fuzzy sets but besides these standard fuzzy set operations, whole classes of prescriptions for the complements, intersections, and unions on fuzzy sets were defined [39].

In this study, we used the standard t-norm (1) and t-conorm (2) for the implementation of *and* and *or* operators and fuzzy complement for the evaluation of the *not* operator (3). Product t-norm (4) was used to evaluate the *prod* operator and dual product t-conorm (5) was used to evaluate the *sum* operator.

$$t(x, y) = \min(x, y) \quad (1)$$

$$s(x, y) = \max(x, y) \quad (2)$$

$$c(x) = 1 - x \quad (3)$$

$$t_{prod}(x, y) = xy \quad (4)$$

$$s_{prod}(x, y) = a + b - ab \quad (5)$$

The fuzzy rule is a simple version of a general fuzzy classifier. In contrast to more complex fuzzy rule-based systems that usually constitute traditional fuzzy classifiers, it consists of a single expression that describes soft requirements on data records in terms of data features. The fuzzy rule modeling selected attribute of a data set is in this study created by the means of genetic programming [4]. The tree structures of parsed fuzzy rules are evolved by iterative application of crossover, mutation, and selection operators in order to find an accurate model of the training data.

For more information about fuzzy rules, their structure, weighting, evaluation, and evolution see [1], [2], [5].

IV. PREDICTION OF TRAFFIC ACCIDENT FEATURES

Evolutionary fuzzy rules were in this study used to learn classes of data describing real traffic incidents according to several features. We have used a data set obtained from the Addis Ababa Traffic Office in Ethiopia. It represents car accidents in Addis Ababa, capital city of Ethiopia in a specified time frame. The data was recorded by traffic police. The data set was initially processed by evolutionary fuzzy rules in a study by Beshah et al. [5]. In this work, we use a cleaned and balanced version of the data set and predict more attributes with different characteristics and multi-class nature.

The original data set contained information about more than 12336 traffic accidents that occurred in Ethiopia between the years 1998 and 2000. Each accident record contained information about accident time (e.g. year, month, week, day, hour), involved vehicle (e.g. vehicle type, technical condition), driver of the vehicle (e.g. driver age, driving experience), accident location and local conditions (e.g. road condition, road orientation, weather situation), and accident victim (e.g. type of victim, type of collision, injury severity). Several attributes in the data set were numerical (e.g. accident year) and the rest was categorical (e.g. vehicle type).

There are multiple attributes that classify the records according to accident results and severity:

- a) *accidentresult* divides the data into two classes: injury and non-injury.
- b) *accidentresult2* reveals more details and classifies the records to injury, fatal, and non-injury classes.
- c) *injuryseverity* divides the data into 4 classes: serious injury, slight injury, fatal injury, and property damage.

The original data set was highly imbalanced and incomplete [5] which is not desirable when dealing with a multi-class classification problem. In order to improve the data set and get more representative models, the records with unknown value of target attributes were removed and the data set was balanced according to the classes of the *injuryseverity* attribute. Afterwards, the data set contained 2000 records with 500 records per each *injuryseverity* class. Needless to say that the reduction of the number of records was dramatic.

The goal of this study was to evolve a fuzzy rule for each of the output attributes listed above. The evolution of fuzzy rule was for each output attribute executed 30 times independently in order to verify the ability of the stochastic genetic programming to find a fuzzy rule that would assign the records to appropriate classes. The settings used for the GP are summarized in table I. The results of fuzzy rule evolution were analyzed separately for each of the output variables.

Accidentresult is binary attribute and corresponding classifier can be evaluated in terms of sensitivity and specificity [40]. The results for *accidentresult* are shown in table II. High values of both sensitivity (on average 99.97%) and specificity (on average 97.43%) suggest that the evolution of fuzzy rules managed to find good classifiers for the binary attribute. An example of fuzzy rule evolved for the *accidentresult* attribute is shown in Fig. 1. The figure represents the structure of

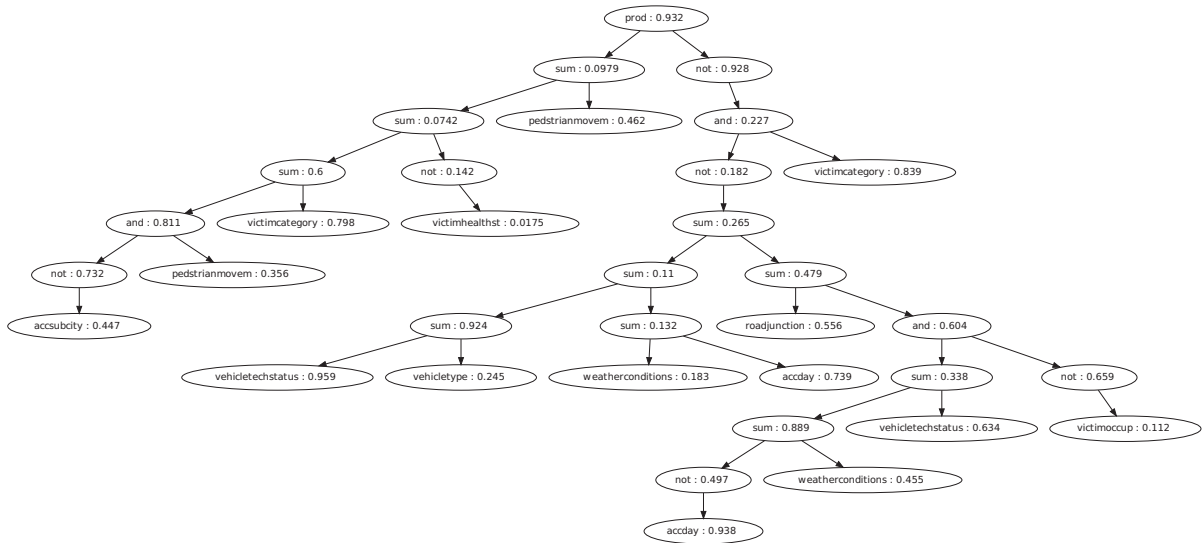


Fig. 1: Example of fuzzy rule for *accidentresult* attribute.

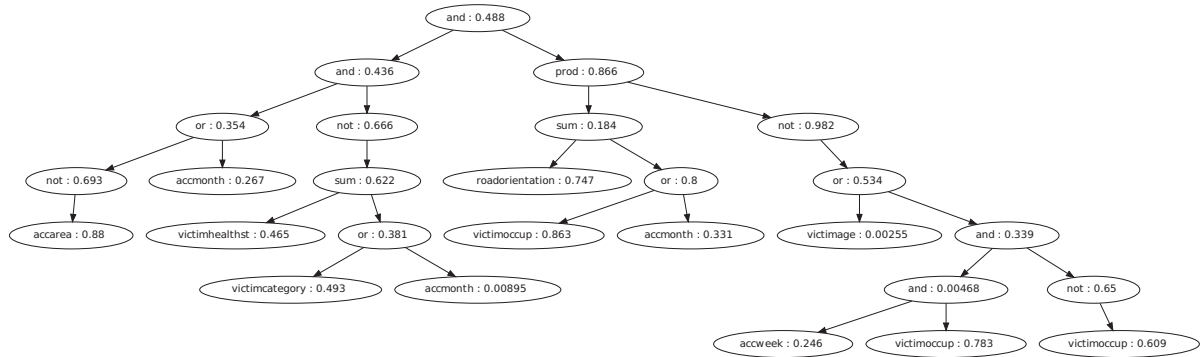


Fig. 2: Example of fuzzy rule for *accidentresult2* attribute.

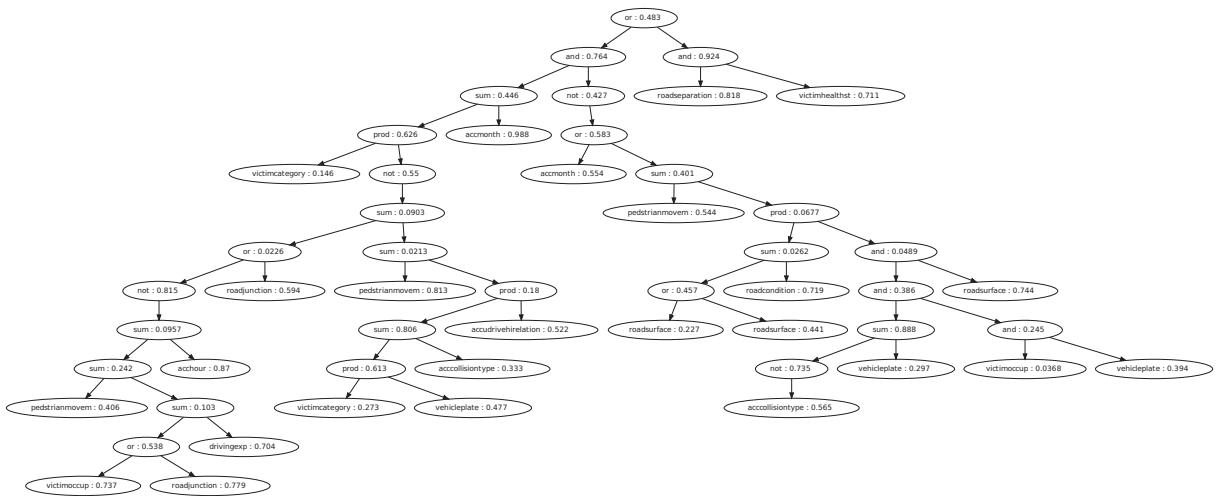


Fig. 3: Example of fuzzy rule for *injuryseverity* attribute.

TABLE I: GP settings.

Parameter	Value
Population size	100
Crossover probability	0.8
Mutation probability	0.2
F-score β	1.0
No. of. generation	5000 or 10000

TABLE II: Specificity and sensitivity of fuzzy rules evolved for *accidentresult*.

Accidentresult	Max.	Avg.	Min
Specificity	100	99.973	99.6
Sensitivity	99.067	97.427	94.333

the classifier as a tree. Leaf nodes represent feature values and other nodes represent operators that combine the features. *Accidentresult2* and *injuryseverity* are multi-class attributes. The quality of evolved fuzzy rules for *accidentresult2* and *accidentseverity* in terms of classification correctness for each class are shown in Tab. IIIa and Tab. IIIb respectively. The average precision of classification by fuzzy rules for *accidentresult2* is lower than the classification of the binary *accidentresult*. We can see that the precision of classification of the injury and no injury classes was good while the fatal class was not recognized very well. The average correctness of classification was 76.02%. The overall precision of classification was even lower for the *accidentseverity* attribute. Even though the classes of *accidentseverity* were perfectly balanced, evolved fuzzy rules were able to correctly assign records only to two classes. The average correctness of classification was 52.69%. Examples of fuzzy rules evolved for the *accidentresult2* and *accidentseverity* attributes are shown in Fig. 2 and Fig. 3 respectively.

TABLE III: Classification accuracy for multi-class attributes *accidentresult2* and *accidentseverity*.(a) Classification results for *accidentresult2* (in %).

Class	Size	Max. correct	Avg. correct	Min. correct
No injury	500	100	98.627	90.8
Injury	1000	96.5	94.607	86.6
Fatal	500	24.4	16.253	8
Total	2000	77.5	76.023	74.4

(b) Classification results for *accidentseverity* (in %).

Class	Size	Max. correct	Avg. correct	Min. correct
Fatal injury	500	21.8	16.1	5.6
Serious injury	500	0.6	0.086	0
Slight injury	500	95.2	94.829	92
Property damage	500	100	99.771	97.2
Total	2000	53.45	52.69	51.8

V. CONCLUSIONS

Genetic programming was in this study used to mine fuzzy rules approximating classes of selected attributes of traffic accident data. A real world data set from Ethiopia was used to evolve fuzzy rules. Thanks to the symbolic nature of the classifier, the attributes that were used for data labeling can be easily extracted and explored. The number of important attributes, their relations and weights were selected automatically and they represent an important knowledge that can be exploited to improve road safety. Moreover, the data mining method can be easily applied to other data sets describing traffic situation in other regions and learn rules and hidden dependencies specific for each region.

This study compared the ability of evolutionary fuzzy rules to evolve classifiers for binary and multi-class attributes. While the rules for a binary attribute were successful, the artificial evolution as implemented in this work was not able to find fuzzy rules that would accurately classify data according to selected multi-class attributes. In our future work, we will study the reason of this behavior and apply investigated method to other data sets describing traffic accidents.

ACKNOWLEDGEMENT

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070) and by the Bio-Inspired Methods: research, development and knowledge transfer project, reg. no. CZ.1.07/2.3.00/20.0073 funded by Operational Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic.

REFERENCES

- [1] V. Snášel, P. Krömer, J. Platos, and A. Abraham, "The evolution of fuzzy classifier for data mining with applications," in *SEAL*, ser. Lecture Notes in Computer Science, K. Deb, A. Bhattacharya, N. Chakraborti, P. Chakraborty, S. Das, J. Dutta, S. K. Gupta, A. Jain, V. Aggarwal, J. Branke, S. J. Louis, and K. C. Tan, Eds., vol. 6457. Springer, 2010, pp. 349–358.
- [2] P. Krömer, J. Platoš, V. Snášel, and A. Abraham, "Fuzzy classification by evolutionary algorithms," in *IEEE International Conference on Systems, Man, and Cybernetics*. IEEE System, Man, and Cybernetics Society, 2011, pp. 313 – 318.
- [3] P. Krömer, J. Platoš, V. Snášel, A. Abraham, L. Prokop, and S. Mišák, "Genetically evolved fuzzy predictor for photovoltaic power output estimation," in *2011 Third International Conference on Intelligent Networking and Collaborative Systems (INCoS)*. IEEE, 2011, pp. 41 – 46.
- [4] M. Affenzeller, S. Winkler, S. Wagner, and A. Beham, *Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications*. Chapman & Hall/CRC, 2009.
- [5] T. Beshah, D. Ejigu, P. Kromer, V. Snasel, J. Platos, and A. Abraham, "Learning the classification of traffic accident types," in *Intelligent Networking and Collaborative Systems (INCoS), 2012 4th International Conference on*, sept. 2012, pp. 463 –468.
- [6] T. B. Tesema, A. Abraham, and C. Grosan, "Rule mining and classification of road traffic accidents using adaptive regression trees," *International Journal of Simulation*, vol. 6, no. 10-11, pp. 80–94, 2005.
- [7] Z. Regassa, "Determining the degree of drivers responsibility for car accident: the case of addis ababa traffic office," Master's thesis, Addis Ababa University, 2009.
- [8] H. Kifle, "Application of data mining technology to support the prioritization of dangerous crash location: the case of addis ababa traffic office," Master's thesis, Addis Ababa University, 2009.

- [9] G. Mossie, "Applying data mining with decision tree and rule induction techniques to identify determinant factors of drivers and vehicles in support of reducing and controlling road traffic," Master's thesis, Addis Ababa University, 2009.
- [10] A. Tabor, "Bayesian approach for analysis of road traffic accidents: The case of addis ababa," Master's thesis, Addis Ababa University, 2009.
- [11] T. Beshah and S. Hill, "Mining road accidents data to improve safety: the role of road related factors on accident severity," in *Proceeding of AAAI Symposium on Artificial Intelligence for Development*. Stanford University, 2010, pp. 14 – 19.
- [12] P. Daigavane and P. Bajaj, "Analysis of selective parameters contributing to road accidents on highways for establishing suggestive precautionary strategies," in *Proceedings of the 2009 Second International Conference on Emerging Trends in Engineering & Technology*, ser. ICETET '09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 576–580. [Online]. Available: <http://dx.doi.org/10.1109/ICETET.2009.31>
- [13] X. Hongguo, Z. Huiyong, and Z. Fang, "Bayesian network-based road traffic accident causality analysis," in *Information Engineering (ICIE), 2010 WASE International Conference on*, vol. 3, aug. 2010, pp. 413–417.
- [14] S.Krishnaveni and Dr.M.Hemalatha, "A perspective analysis of traffic accident using data mining techniques," *International Journal of Computer Applications*, vol. 23, no. 7, pp. 40–48, June 2011, published by Foundation of Computer Science.
- [15] Y. Haixia and N. Zhihong, "An application of factor analysis on road traffic accident," in *Computer Science and Education (ICCSE), 2010 5th International Conference on*, aug. 2010, pp. 1355–1358.
- [16] T. Li, Y. Chen, S. Qin, and N. Li, "Highway road accident analysis based on clustering ensemble," in *Computer Science for Environmental Engineering and Ecolinformatics*, ser. Communications in Computer and Information Science, Y. Yu, Z. Yu, and J. Zhao, Eds. Springer Berlin Heidelberg, 2011, vol. 159, pp. 212–217, 10.1007/978-3-642-22691-5_37. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-22691-5_37
- [17] W. Jinlin, C. Xi, Z. Kefa, W. Wei, and Z. Dan, "Application of spatial data mining in accident analysis system," in *Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing, ETT and GRS 2008. International Workshop on*, vol. 1, dec. 2008, pp. 472–475.
- [18] N. Saunier, N. Mourji, and B. Agard, "Mining microscopic data of vehicle conflicts and collisions to investigate collision factors," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2237, no. -1, pp. 41–50, Dec. 2011. [Online]. Available: <http://dx.doi.org/10.3141/2237-05>
- [19] R. Nayak, D. Emerson, J. Weligamage, and N. Piyatrapoomi, "Road crash proneness prediction using data mining," in *EDBT, A. Ailamaki, S. Amer-Yahia, J. M. Patel, T. Risch, P. Senellart, and J. Stoyanovich*, Eds. ACM, 2011, pp. 521–526.
- [20] A. Pakgohar, R. S. Tabrizi, M. Khalili, and A. Esmaeili, "The role of human factor in incidence and severity of road crashes based on the cart and lr regression: a data mining approach," *Procedia Computer Science*, vol. 3, no. 0, pp. 764 – 769, 2011, ;ce:title;World Conference on Information Technology;ce:title;. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050910005016>
- [21] L.-Y. Chang and H.-W. Wang, "Analysis of traffic injury severity: An application of non-parametric classification tree techniques," *Accident Analysis & Prevention*, vol. 38, no. 5, pp. 1019 – 1027, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0001457506000583>
- [22] J. de Oña, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury severity on spanish rural highways using bayesian networks," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 402 – 411, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0001457510002733>
- [23] P. C. Anastasopoulos and F. L. Mannering, "An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data," *Accident Analysis & Prevention*, vol. 43, no. 3, pp. 1140 – 1147, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S000145751000401X>
- [24] P. Liu, "A self-organizing feature maps and data mining based decision support system for liability authentications of traffic crashes," *Neurocomput.*, vol. 72, no. 13-15, pp. 2902–2908, Aug. 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.neucom.2008.06.032>
- [25] D. Delen, R. Sharda, and M. Bessonov, "Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks," *Accident Analysis & Prevention*, vol. 38, no. 3, pp. 434 – 444, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0001457505001879>
- [26] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1666 – 1676, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0001457511000765>
- [27] A. Morgan and F. L. Mannering, "The effects of road-surface conditions, age, and gender on driver-injury severities," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1852 – 1863, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0001457511001059>
- [28] T. Beshah, D. Ejigu, A. Abraham, V. Snasel, and P. Kromer, "Pattern recognition and knowledge discovery from road traffic accident data in ethiopia: Implications for improving road safety," in *Information and Communication Technologies (WICT), 2011 World Congress on*, dec. 2011, pp. 1241–1246.
- [29] O. Cerdón, "A historical review of evolutionary learning methods for mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 894 – 913, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0888613X11000582>
- [30] O. Cerdón, M. J. del Jesus, and F. Herrera, "Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods," *International Journal of Intelligent Systems*, vol. 13, pp. 1025–1053, 1998.
- [31] H. Ishibuchi and Y. Nojima, "Multiobjective formulations of fuzzy rule-based classification system design," in *EUFLAT Conf.*, E. Montseny and P. Sobrevilla, Eds. Universidad Polytechnica de Catalunya, 2005, pp. 285–290.
- [32] C.-H. Wang, T.-P. Hong, and S.-S. Tseng, "Integrating membership functions and fuzzy rule sets from multiple knowledge sources," *Fuzzy Sets Syst.*, vol. 112, pp. 141–154, May 2000. [Online]. Available: <http://portal.acm.org/citation.cfm?id=346592.346613>
- [33] M. Freischlad, M. Schnellbach-Held, and T. Pullmann, "Evolutionary generation of implicative fuzzy rules for design knowledge representation," in *EG-ICE*, ser. Lecture Notes in Computer Science, I. F. C. Smith, Ed., vol. 4200. Springer, 2006, pp. 222–229.
- [34] E. Zhou and A. Khotanzad, "Fuzzy classifier design using genetic algorithms," *Pattern Recogn.*, vol. 40, pp. 3401–3414, December 2007. [Online]. Available: <http://dx.doi.org/10.1016/j.patcog.2007.03.028>
- [35] R. Prado, S. Garcia-Galán, J. Exposito, and A. Yuste, "Knowledge acquisition in fuzzy-rule-based systems with particle-swarm optimization," *Fuzzy Systems, IEEE Transactions on*, vol. 18, no. 6, pp. 1083–1097, dec. 2010.
- [36] G. Pasi, "Fuzzy sets in information retrieval: State of the art and research trends," in *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, ser. Studies in Fuzziness and Soft Computing, H. Bustince, F. Herrera, and J. Montero, Eds. Springer Berlin / Heidelberg, 2008, vol. 220, pp. 517–535. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-73723-0_26
- [37] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. pp. 338–353, 1965.
- [38] J. Jantzen, "Tutorial On Fuzzy Logic," Technical University of Denmark, Dept. of Automation, Technical Report 98-E-868 (logic), 1998.
- [39] T. Feuring, *Fuzzy-systeme*. Institut fr Informatik, Westflische Wilhelms Universitt, Mnster, 1996.
- [40] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed., ser. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, 2005.