© Universiti Tun Hussein Onn Malaysia Publisher's Office

JST

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/jst

ISSN : 2229-8460 e-ISSN : 2600-7924

Journal of Science and Technology

Statistical Clustering Performance in Pavement Condition Prediction as Decision Supporting System Tool

Mohammad Mehdi Khabiri[1*]

[1] Yazd University, Civil Engineering Department

University Blvd, Safeties, Yazd, 98195-741, IRAN

*Corresponding author

DOI: https://doi.org/10.30880/jst.2019.11.02.004

Received 12 August 2019; Accepted 15 December 2019; Available online 29 December 2019

Abstract: Mathematical methods and statistical patterns have always been considered by managers, designers and science and technology expert in order to develop technology and engineering objectives. During the development of data-gathering tools and increment of data-bases, data mining have made suitable tools in management and engineering. The assessment of roads' maintenance is highly important in order to prevent early deterioration of roads and performing maximum road capacity during the service-life. Pavement management of roads has also implemented this tool to make proper decisions and preferences of pavement repair methods, using decision tree. Through engineering management, cluster analysis is one of the basic tools of data mining and knowledge discovery and makes the decision making, easier in engineering. Data categorization is helpful for planning and is important in picking proper methods. This study was performed by using recorded data from other scientific sources considering data mining method and analyzing data with respect to statistical clustering. The results indicate that bitumen content in asphalt mix, pavement age, marshal strength and rate of passing vehicles have the most important effect on decrement of condition index of pavement, relatively. Also, the highest deterioration in asphalt happens in 5.5% and higher values of bitumen content and the progressive deteriorations take place when the pavement age exceeds 35 years.

Keywords: Statistical Clustering, Decision Supporting System, Pavement Condition Prediction, Asphalt Pavement Maintenance Management

1. Introduction

Data clustering or numerical taxonomy include mathematical and statistical technics and is used for classification of existing data into groups and classes with the same condition. Cluster analysis is one of the most useful methods in many majors. Also in civil engineering, statistical methods are implemented to determine data effectiveness and their separation.

Nomenclature	
$S_{(xk,yk)}$	the similarity of two objects
n	objects or record number
x	component of the base class label
$\psi_{(S/t)}$	goodness criterion

The cluster analysis is one of basic data mining and knowledge discovery tools in engineering management. Data classification is helpful in programming and choosing proper repair method and is so important as well.

Data mining process has had successful performance in several civil engineering problems and has solved many complicated problems and has omitted the troubles of expensive tests and data records (Li, and Yang, 2018), having a

*Corresponding author: mkhabiri@yazd.ac.ir

glimpse on literature, it is obvious that data mining implementation in a vast variety of problems such as civil engineering and especially in pavement field, has concluded in successful results (Yuan, 2019, Peng, 2018).

Immense costs of traffic accidents have caused the safety conditions of ways to become one of the most crucial objectives of most of the countries in field of transportation engineering (Ziari, and Khabiri, 2006, Castro-Nuño et. al. 2018). Several researches have revealed that pavement characteristics and its deterioration state is one of significant parameters in accidents' possibility. On the other hand various research activities are performed to improve and rehabilitate the engineering infrastructures, all over the world. Data mining tool is widely being used in civil engineering these years (Chen, et. al., 2018). Transportation engineering also uses this tool to optimally manage the transportation networks (Tafti, et. al.2016). Data mining and its methods in design, performance and management of pavement in recent studies (Khabiri, 2010). Road maintenance operations are highly important for early deterioration of roads and utilizing the highest road capacity during its expected service-life (Khadka, Paz, and Singh, 2018; Ziari, Ameri, and Khabiri 2010) additives and wastes are also used in pavement life (Nabiun and Khabiri 2010). Prioritization and implementation of optimizing decision methods that could predict the pavement state without implementing expensive and complicated equipment, is determined as a requirement, for developing countries by the scientists (Wang, et. al., 2018). Development of traffic data clusters is critical for use of the Mechanistic–Empirical Pavement Design Guide (MEPDG) while site-specific traffic data are not existing and statewide data are too general. However, an ideal method to traffic data clustering is not specified in the MEPDG. A study in the United Statesoffered a new clustering combination method, correlation-based clustering, that considered the effects of traffic inputs on pavement design depths, so that purpose of the amount of clusters is made accurately (Mai, et.al. 2013).

In present study, scientific basis of data clustering method is defined firstly, and then a numerical-practical study in data clustering was performed by utilizing pavement maintenance management model that is known as "Pavement Management System". Prediction of pavement condition is now possible and spread due to creation of data bank. In this research scientific basis of decision tree pattern is introduced and the technological application is also defined through a case study.

2. Scientific Basis of Data Clustering

Implementation of data mining and data clustering tools as powerful engineering tools, causes in complicated analysis time and high costs decrement. Decision tree construction (DTC) is a typical method for classification. A database for DTC involves of a set of data archives that are pre-classified into q (≥ 2) recognized classes. The objective of decision tree structure is to partition the data to discrete the q classes. A decision tree has two types of nodes, decision nodes and leaf nodes. A decision node requires some test on a single characteristic. A leaf node shows the class. From a geometric opinion, a decision tree signifies a dividing of the data space. A serial of tests from the root node to a leaf node represents a hyper-rectangle (Liu, et.al.2000). Data classification or drawing decision is a two stage process in which a model is made in the first stage that determines a set of data classes or concepts (Liu, et.al. 2018). This stage is called "learning" in which an algorithm creates a model by means of a training set as base component. In the second stage, learning is performed through a y=f(x) function that could predict each x component of the base class label. This function could be defined as classification rules, decision tree or mathematical formulas. Data division process in decision tree is presented in fig. (1).

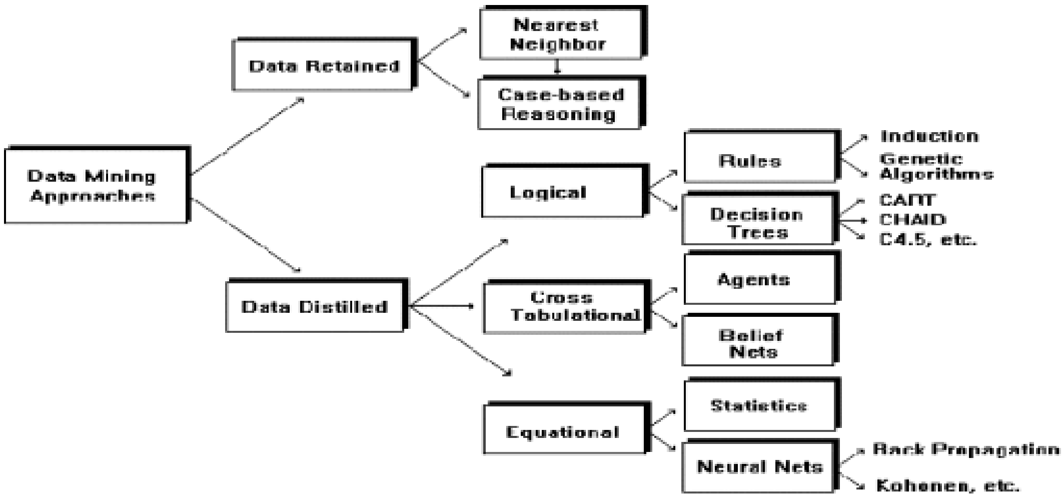

Fig. 1- Decision tree algorithm process in data mining technic (Rygielski, et.al. 2002)

The stages that must be completed in data clustering process are:

- To provide and present data matrix from data bases available in design and monitoring engineering systems,

- To standardize the "Similarity matrix" based on distance and similarity coefficient methods
- To perform clustering method (by means of relations or practical software),
- To calculate the validity criterion and its classification,

As this study deals with scientific applications, mathematical and statistical content is therefore summarized.

2.1. Standardization of Data Matrix

Raw data matrix is introduced as matrix (1) in which each line represents an object or record and each column represents an index, n stands for objects number, d stands for indexes number and x_{ij} represents raw data. It is important to convert this matrix into standard data matrix as z_{ij} in relation (2).

$$x_{ij} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{d1} & \cdots & x_{dn} \end{bmatrix} \tag{1}$$

$$z_{ij} = \begin{bmatrix} z_{11} & \cdots & z_{1n} \\ \vdots & \ddots & \vdots \\ z_{d1} & \cdots & z_{dn} \end{bmatrix} \tag{2}$$

However there are various methods to standardize data matrix including normal standard method, converting the values to scores (Miligan and Cooper method). During the clustering process of data, after the data was gathered and became standard, the distance or data similarity must be calculated in a group. Making proximity matrix, similarity matrix and resemblance matrix is done in this stage. These two matrices are considered in relations (3) and (4).

$$M_d = \begin{bmatrix} \cdot & d_{12} & d_{13} & \cdots & d_{1n} \\ d_{21} & \cdot & d_{23} & \cdots & d_{2n} \\ \vdots & \cdot & \cdot & \cdots & \vdots \\ d_{n1} & d_{n2} & \cdots & \cdots & \cdot \end{bmatrix} \tag{3}$$

$$M_s = \begin{bmatrix} 1 & S_{12} & S_{13} & \cdots & S_{1n} \\ S_{21} & 1 & S_{23} & \cdots & S_{2n} \\ \vdots & \cdot & \cdot & \cdots & \vdots \\ S_{n1} & d_{n2} & \cdots & \cdots & 1 \end{bmatrix} \tag{4}$$

To measure the data similarity Gower general factor is used that is defined in relation (5). In this relation, "d" presents the number of indexes, S (x_k,y_k) is the similarity of two objects from "k" index point of view and $\alpha_{(xk,yk)}$ is a binary factor showing that k component is a factor used to calculate suitably. Pearson's correlation coefficient, follows the relation (6) to define similarity.

$$S_{(x.y)} = \frac{\sum_{k=1}^{d} \alpha\,(x_k.y_k)\,S\,(x_k.y_k)}{\sum_{k=1}^{\alpha} \alpha\,(x_k.y_k)} \tag{5}$$

$$R_{(x.y)} = \frac{\sum_{i=1}^{n}(x_i y_i - n\bar{x}\bar{y})}{\sqrt{\left(\sum_{i=1}^{n} x_i - n\bar{x}^2\right)}\sqrt{\left(\sum_{i=1}^{n} y_i - n\bar{y}^2\right)}} \tag{6}$$

2.2. CART Classification and Regression

Data classification and regression tree method was introduced by Breiman et. al. in 1984.(Momeni, 2010) Decision trees produced by CART technic, are binary. CART algorithm expands the decision making through studying all available criterion and optimal separations of each decision node according to tree base. If ψ(S/t) goodness criterion is an optimal split in node t, relation (7) is true.

$$\varphi_{(s,t)} = 2P_l P \sum_{j=1}^{t} [P_{(j,t_l)} - P_{(j,t_{lr})}] \tag{7}$$

In which:
t_L: Left side successor of t node
t_R: Right side successor of t node

In which,

$$\mathrm{P(J,t_L)} = \frac{\text{Number of records belonging to the class } T_l \textit{ in } J}{\text{Record number in } t} \tag{8}$$

$$P_R = \frac{\textit{The number of records in the TR}}{\textit{Number of records in training series}} \tag{9}$$

The optimum decision is a choice that maximizes the ψ(S/t) criterion on all optimal splits.

2.3. Decision Tree Quality Parameter

In a decision tree when N existing data or records must be arranged in a correct cluster. Relative value of errors means that a ratio of records are incorrectly classified. These could be obtained from relations (10) and (11).

$$P_{err} = \frac{x_{err}}{N} \tag{10}$$

$$x_{er} = \sum_{S=1}^{M} \sum_{i=1}^{K} x_i^{S'} \tag{11}$$

K is the number of classes or groups. Relative variance could be obtained from relations (12) to (15) for a decision tree (Zeinelhamadani, et. al., 2013).

$$P_{cm} = \frac{d_{oc}}{d_o} \tag{12}$$

$$d_{oc} = \frac{1}{N} \sum_{S=1}^{M} \sum (x(s) - x)^2 \tag{13}$$

$$d_o = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x}) \tag{14}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{15}$$

Accuracy parameter is made as a validating tool of decision tree (Pakgohar, and Sadeghikia, 2007),

3. Research Method

Present research is a compound of mathematical science (statistics and data use) utility to implement in civil engineering of transportation field. Firstly, the mathematic scientific fundaments related to it are discussed and then the engineering management of technology -known as PMS- will be introduced.

To describe the pavement condition, an index know as PCI (Pavement Condition Index) is used of which value varied between 0 (for an unusable pavement) and 100 (for a proper pavement). PCI is both calculated from a field visual study and also accurately by considering the type and the severity of the deteriorations by measuring their amount. A schematic image of this technology is apparent in fig. (2), to determine PCI, a segment of pavement from all the pavement classification is selected. Due to deterioration state related to its general condition, the PCI number will be obtained that is saved in Micro Paver software database.

Initial information about this research is provided from Moaiedfar and Taqadosi's research in 2018 that was held in 2nd municipal zone of Tehran, the capital city of Iran (Moaidifar, Taghdosi, 2018). At first each segment of pavement is determined and condition index of it, was also calculated. The asphalt sampling was performed to determine the characteristics of asphalt (Bianchini, and Bell, 2019). According to table (2), descriptive statics and the calculation of central indexes of data is represented in it. All statistical analyzes are performed by means of common SPSS software.

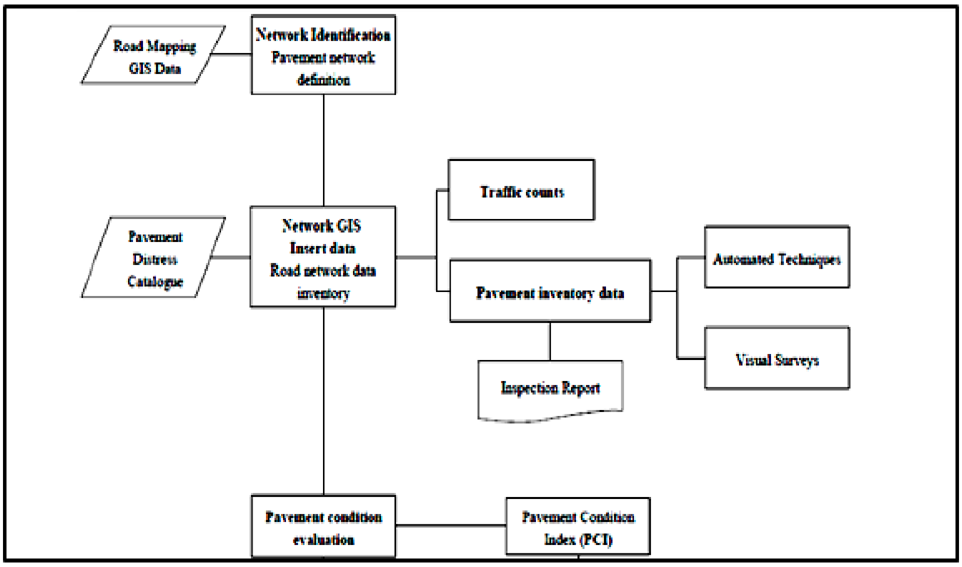

Fig. 2- Schematic graph of calculation steps of PCI (Loprencipe, et.al.2017)

Table 1- Field data records according to previous studies, asphalt pavement of case study

Pavement Condition(PCI)	Bitumen (%)	Marshal Stability(kN)	Annual Average Traffic(Veh/hr)	Age (year)	Pavement Condition(PCI)	Bitumen (%)	Marshal Stability(kN)	Annual Average Traffic(Veh/hr)	Age (year)
93.5	5.9	1011	1351	39	88.8	6.1	1202	127	39
91.5	5.8	1222	1014	39	86.9	5.6	1087	53	38
84	5.5	1269	436	36	87	6.3	925	135	39
94.1	5.6	1334	19	37	90.8	6.1	976	85	39
87.1	5.7	952	47	37	87.5	6.1	869	29	39
87	5.7	974	46	37	91.2	6.2	1078	26	39
86.1	5	974	79	38	89.5	6.3	958	79	39
89.7	5.9	1271	973	37	89.9	5.2	1231	21	30
92	6	887	27	39	87.1	4.8	865	14	30
92.1	5.9	1083	753	38	93.8	4.4	748	19	30
89.7	5.9	856	962	38	86.3	5.2	1162	11	30
88.3	6.3	838	1278	38	92.5	4.4	731	108	30
87.8	6.4	729	942	38	83	5.3	1045	35	30
87.2	6.1	767	1148	38	90.8	4.1	802	1011	31
91	6	832	114	39	91.5	4.5	904	1245	31
92.9	6.3	859	97	39	91.6	4.7	1067	9	31
91.6	5.9	1500	67	39	93	4.8	965	11	31

Table 2- Calculation of central indexes and characteristics summary of subjective statistics of the study data

Index	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
PciPav	37	83.00	94.10	89.7108	2.85334	8.142	-.389	.388	-.651	.759
BitPerc	37	4.10	6.40	5.5270	.68906	.475	-.720	.388	-.741	.759
MarSta	37	729	1500	990.86	183.543	33687.898	.796	.388	.273	.759
AATraf	37	8	1351	349.46	463.786	215097.144	1.082	.388	-.541	.759
AgYea	37	30	39	35.76	3.774	14.245	-.688	.388	-1.442	.759
Valid N (listwise)	37									

Checking the normality of data that include output data of Marshal Test (in kg-f) and calculated amounts for PCI, based on deterioration observations. At first, normality test was performed for PCI results by means of graphical method and its conclusions are apparent in fig. (3).

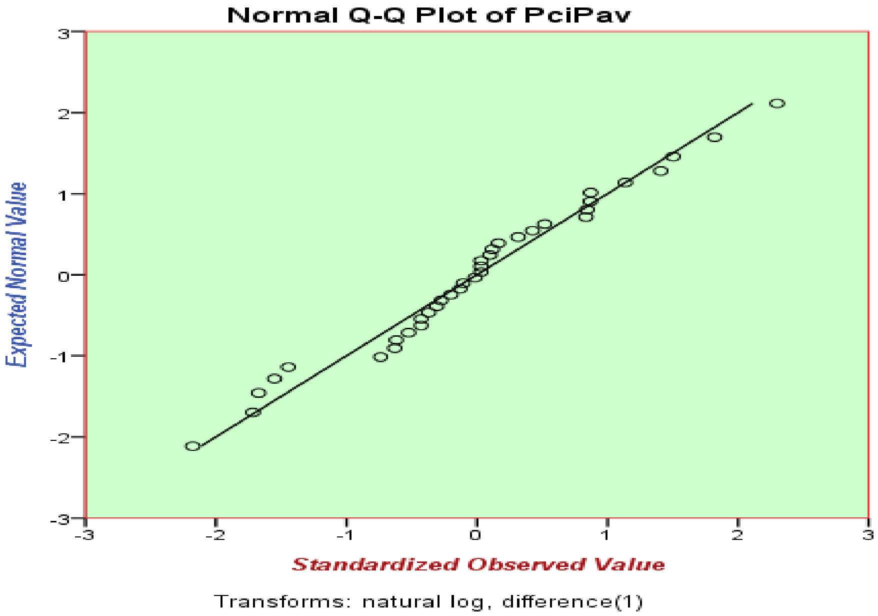

Fig. 3- Normality graph of data implemented in the study, pavement condition index

4. The Results of Numerical Analysis and Discussion

Decision tree was modeled, using SPSS software for data and then fig. (4), was obtained. This tree has got 4 branches and 15 leaves. As it is obvious, in the first leaf, PCI is depended to bitumen content. If the bitumen content exceeds 4.5%, PCI drastically increases. The PCI is depended to pavement age in second branch. If the pavement age passes 38.2, pavement PCI dramatically decreases.

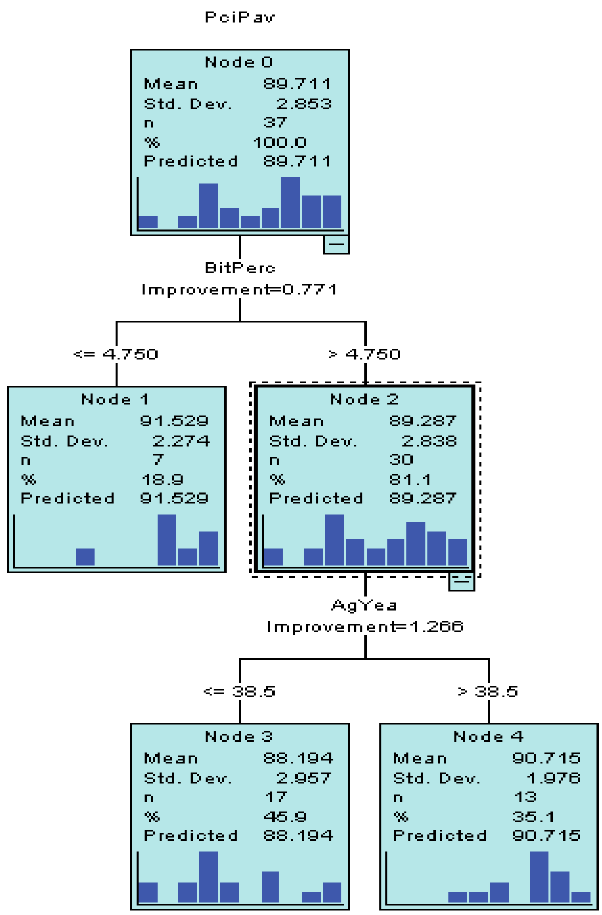

Fig. 4- Decision tree pattern is for predicting the road pavement condition

Decision support system is also implemented as a holistic title that is used almost in all professional and technical fields. Furthermore, it could be used in road pavement and related fields. One of the design procedures of the systems, that aids managers and engineers leading these sort of systems to make rapid, principal and correct decisions is implementing knowledge discovery and association rule based methods. An example of these analyzes is apparent in fig. (5), as data clustering output. As it is obvious, to predict pavement condition, according to described parameters, association rules could be used by software.

```
/* Node 1 */.
DO IF (((VALUE(BitPerc) LE 4.75) OR SYSMIS(BitPerc)  AND  ((VALUE(AATraf) LE 10) OR SYSMIS(AATraf)  AND  ((VALUE(AgYea) LE 33.5) OR SYSMIS(AgYea)  AND
(VALUE(MarSta) LE 825.5))))).
COMPUTE nod_001 = 1.
COMPUTE pre_001 = 91.5285714285714.
END
 IF.
EXECUTE.

/* Node 3 */.
DO IF (((VALUE(BitPerc) GT 4.75) OR SYSMIS(BitPerc)  AND  ((VALUE(AATraf) GT 10) OR SYSMIS(AATraf)  AND  ((VALUE(AgYea) GT 33.5) OR SYSMIS(AgYea)  AND
 (SYSMIS(MarSta) OR (VALUE(MarSta) GT 825.5))))))  AND  (((VALUE(AgYea) LE 38.5) OR SYSMIS(AgYea)  AND
  ((VALUE(BitPerc) LE 5.75) OR SYSMIS(BitPerc)  AND  ((VALUE(MarSta) GT 938.5) OR SYSMIS(MarSta)  AND  (SYSMIS(AATraf) OR (VALUE(AATraf) GT 285.5)))))).
COMPUTE nod_001 = 3.
COMPUTE pre_001 = 88.1941176470588.
END IF.
EXECUTE.

/* Node 4 */.
DO IF (((VALUE(BitPerc) GT 4.75) OR SYSMIS(BitPerc)  AND  ((VALUE(AATraf) GT 10) OR SYSMIS(AATraf)  AND  ((VALUE(AgYea) GT 33.5) OR SYSMIS(AgYea)  AND
 (SYSMIS(MarSta) OR (VALUE(MarSta) GT 825.5))))))  AND  (((VALUE(AgYea) GT 38.5) OR SYSMIS(AgYea)  AND
  ((VALUE(BitPerc) GT 5.75) OR SYSMIS(BitPerc)  AND  ((VALUE(MarSta) LE 938.5) OR SYSMIS(MarSta)  AND  (VALUE(AATraf) LE 285.5))))).
COMPUTE nod_001 = 4.
COMPUTE pre_001 = 90.7153846153846.
END IF.
EXECUTE.
```

Fig. 5- Association rules, derived from clustering procedure of decision tree in pavement deterioration

After that, statistical modeling in decision tree to determine the most important parameter is demonstrated in fig. (6), as it is apparent, bitumen content in asphalt mixture, pavement age, Marshal Strength and vehicle pass are important relatively.

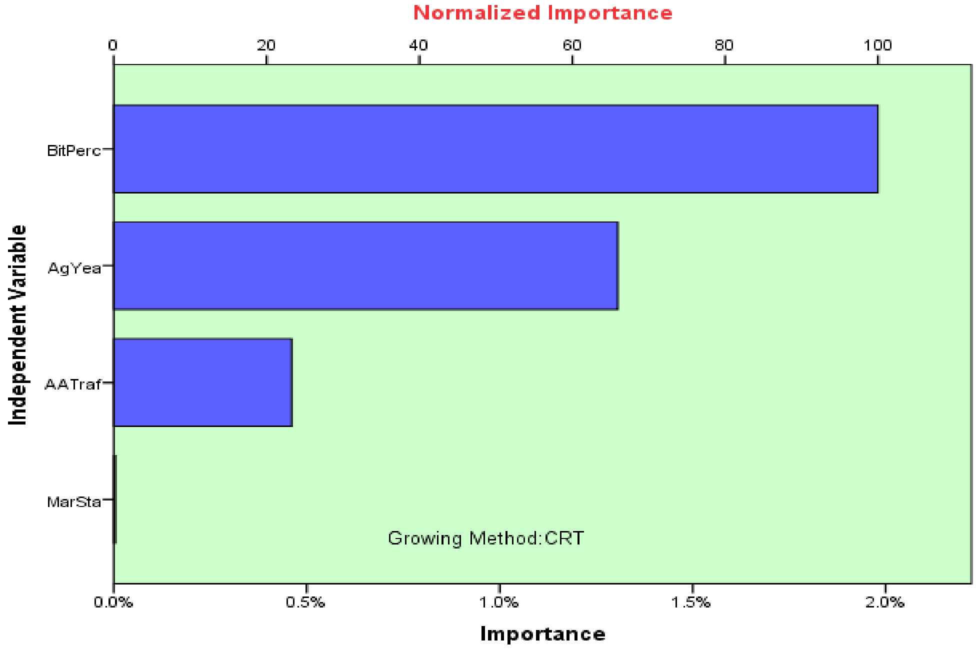

Fig. 6- Prioritization of effective factors in pavement deterioration based on data clustering results

In practical and scientific researches, multiple parameters are usually studied. Thus, to describe the observations, proper multi-criterion graphs must be used. The effect of parameters lead to deterioration of asphalt pavement, are graphically presented in Minitab software in fig. (7), the red and yellow stains represent the maximum deterioration that is related to bitumen content and pavement age. Bitumen content exceeding 5% is the most critical factor of deterioration and the pavement age of more than 35 includes the most portion of deteriorations.

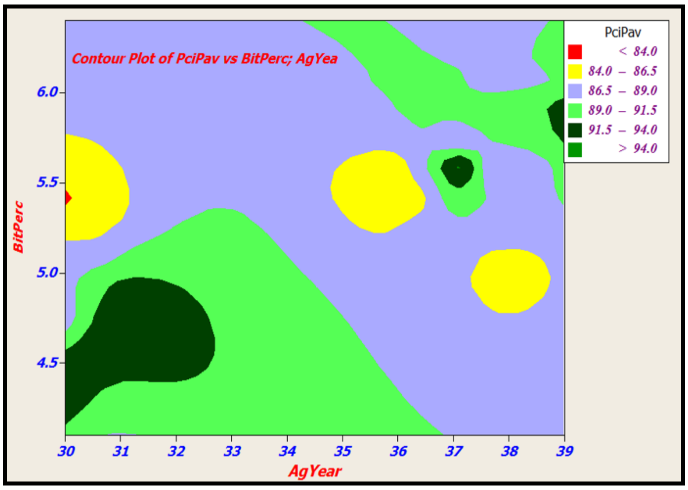

Fig. 7- Histogram of pavement age and bitumen content of asphalt mix parameters' effect on pavement condition

5. Conclusion

The main objective of this research is to study the implementation of mathematical and statistical new technologies in civil engineering and pavement maintenance. Deterioration state of pavement, plays a critical role in programming and management design of road maintenance. Due to this, the relation between effective parameters, association rules and clustering method is found by means of various graphs and diagrams, considering simultaneous effects of bitumen content, pavement age and traffic volume, by using data mining and data clustering tools, a data base and also implementing CART analysis method as a powerful analysis tool in uncertainty-included problems, is studied in this research. Some conclusions are classified as follows:

- Through parameters' importance analysis, the bitumen content, pavement age, Marshal Strength and vehicles' pass are resulted to be important relatively.
- On the first node of obtained decision tree, the PCI is depended on the bitumen content. If it exceeds 4.5%, PCI will extremely increase.
- On the second node of obtained decision tree, the PCI value is related to pavement age. If the pavement age exceeds 38.2, PCI index will dramatically decrease.
- The maximum deterioration of asphalt coat takes place in bitumen contents of 5.5% and higher. Also the progressive deteriorations reveal when the age of the pavement exceeds 35 years.

However, mathematical and statistical science is capable to determine and recommend the priority of parameters affecting optimal maintenance of road pavements as data mining and data clustering algorithms. It could be a beneficiary tool of pavement repair technology by means of utilizing accurate information and using statistical and mathematical functions. Hope that, present research, become one of the steps, towards further studies related to low-cost maintenance and using up-to-date tools and science.

Acknowledgement

The author would like to thank to the University Blvd, Safeties, Yazd Iran for the financial and facilities support throughout this research.

References

Bianchini, A., & Bell, H. P. (2019). Fuzzy cluster approach for area FWD representative basin from deflection measurement spatial variability. International Journal of Pavement Engineering, 20(7), 844-852.

Castro-Nuño, M., Castillo-Manzano, J. I., & Fageda, X. (2018). The role of road safety in a sustainable urban mobility: An econometric assessment of the Spanish NUTS-3 case. International journal of sustainable transportation, 12(3), 205-217.

Chen, C., Wu, Q., Zhang, G., Liu, X. C., & Prevedouros, P. D. (2018). Extracting Arterial Access Density Impacts on Safety Performance Based on Clustering and Computational Analysis. Journal of Transportation Engineering, Part A: Systems, 144(4), 04018008.

Kamarulzaman, N. A., Adnan, S. H., Sari, K. M., Osman, M. H., Jeni, M. A., Abdullah, M. S., & Anuar, M. W. (2018). Properties of Cement Brick Containing Expanded Polystyrene Beads (EPS) and Palm Oil Fuel Ash (POFA). Journal of Science and Technology, 10(4).

Khabiri, M. M. (2010). The effect of stabilized subbase containing waste construction materials on reduction of pavement rutting depth. Electronic Journal of Geotechnical Engineering, 15, 1211-1219.

Khadka, M., Paz, A., & Singh, A. (2018). Generalised clusterwise regression for simultaneous estimation of optimal pavement clusters and performance models. International Journal of Pavement Engineering, 1-13.

Li, T. Z., & Yang, X. L. (2018). Risk assessment model for water and mud inrush in deep and long tunnels based on normal grey cloud clustering method. KSCE Journal of Civil Engineering, 22(5), 1991-2001.

Liu, B., Xia, Y., & Yu, P. S. (2000, November). Clustering through decision tree construction. In Proceedings of the ninth international conference on Information and knowledge management (pp. 20-29). ACM.

Liu, J., Lou, L., Huang, D., Zheng, Y., & Xia, W. (2018, May). Lane Detection Based on Straight Line Model and K-Means Clustering. In 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS), pp. 527-532.

Loprencipe, G., Pantuso, A., & Di Mascio, P. (2017). Sustainable pavement management system in urban areas considering the vehicle operating costs. Sustainability, vol.9, No.3, pp. 453.

Mai, D., Turochy, R. E., & Timm, D. H. (2013). Correlation-Based Clustering of Traffic Data for the Mechanistic–Empirical Pavement Design Guide. Transportation Research Record, 2339(1), 104-111.

Mahmud, S., Islam, M. I., Ali, R. B., Islam, M. M., & Hasan, M. M. (2018). Compressive Behavior of Concrete by Using Bagasse Ash from Sugar Mill. Journal of Science and Technology, 10(3).

Moaidifar, R., Taghdosi, A., (2017) Road Pavement Condition Index (PCI) Optimization Based on Marshall Mix Design Properties (Case Study of Tehran District 2), Article 12, Volume 14, Number 1, Spring 2016, Page 157-169.

Momeni, M.E. (2010) Data Clustering (Cluster Analysis) Daneshgaran Publications, University of Tehran, p. 307.

Nabiun, N., & Khabiri, M. M. (2016). Mechanical and moisture susceptibility properties of HMA containing ferrite for their use in magnetic asphalt. Construction and Building Materials, 113, 691-697.

Pakgohar, A., Sadeghikia, A. (2007), Analysis of Statistical Data of Traffic Accidents by Decision Tree, Journal of Traffic Studies and Management, Third Year, No. 8, pp. 27-46.

Peng, K. (2018). Risk Evaluation for Bridge Engineering Based on Cloud-Clustering Group Decision Method. Journal of Performance of Constructed Facilities, 33(1), 04018105.

Rygielski,c. Wang,Y.C. Yen,D.C. (2002),Data mining techniques for customer relationship management, Technology in Society,Volume 24, Issue 4,2002,Pages 483-502,https://doi.org/10.1016/S0160-791X(02)00038-6.

Tafti, M. F., Khabiri, M. M., & Sanij, H. K. (2016). Experimental investigation of the effect of using different aggregate types on WMA mixtures. International Journal of Pavement Research and Technology, 9(5), 376-386.

Wang, W., Wang, S., Xiao, D., Qiu, S., & Zhang, J. (2018). An Unsupervised Cluster Method for Pavement Grouping Based on Multidimensional Performance Data. Journal of Transportation Engineering, Part B: Pavements, 144(2), 04018005.

Yuan, X. X. (2019). Discussion on Chen, Zhuo and Liu, Xiaoyue Cathy (2019), “Roadway asset inspection sampling using high‐dimensional clustering and locality‐sensitivity hashing,” Computer‐Aided Civil and Infrastructure Engineering, 34: 2, 116–129. Computer‐Aided Civil and Infrastructure Engineering, 34(6), 539-541.

Zeinel Hamadani, AS, Ebrahimian, F, Yaghoubzadeh, H. (2013) Data Knowledge, Introduction to Data Mining, Isfahan University of Technology Publications, 2014, p.

Ziari, H., Ameri, M., & Khabiri, M. M. (2007). Resilient behavior of hot mixed and crack sealed asphalt concrete under repeated loading. Technological and Economic Development of Economy, 13(1), 56-60.

Ziari, H., & Khabiri, M. M. (2006). Analysis characteristics and provide a prediction model of public bus accident in Tehran. J. Journal of Applied Science, 6(2), 247-250.