Crime Case Reasoning Based Knowledge Discovery Using Sentence Case Relative Clustering for Crime Analyses

M. Ramzan Begam[1*], P. Sengottuvelan[2]

[1]Research Scholar, Bharathiar University, Coimbatore.
[2]Associate Professor and Head, Department of Computer Science, PG Extension Centre, Periyar University, Dharmapuri.
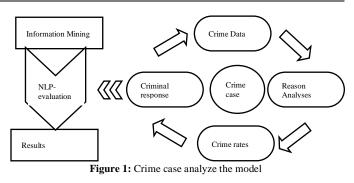**Corresponding author E-mail:Prof.ramzanbegam@gmail.com*

Abstract

Day to day involvement in crime becomes higher statistics for providing information against crime occurrences. A crime committed in different locations, the point of crime occurrence, strategy be analyzed very tedious using only information records. Because information collection in the form of attribute case records with direct crime rates score, so valid factor identification of crime category is a problem. By using the crime cluster in data mining technique to analyze the criminal records to propose a sentence case relative clustering algorithm (SCRCA)with addition classification rule mining algorithm to solve the crime problems. Also to use the sentence case observer technique for knowledge discovery from the crime records for proper case identification from sentence case records and to help increase the predictive accuracy. Crime examines a developing method and identifying the field in law implementation without standard definitions for correct judgment. With the expanding utilization of the clustering automated frameworks to track crimes, information examiners helping the law implementation officers and analysts to accelerate the way toward measuring crimes. The main contribution is to analyses the attribute case with relative sentences of count word analyzes factor for improving the crime prediction for categorizing crime type.

Keywords: Data mining, crime data analysis, clustering, crime-patterns, crime profiling.

1. Introduction

To perform crime analysis for proper information mining using an advanced approach to crime rates which cluster supports an arrangement of crime records. The information arrangements are from crime records is more comparative than those in different crime conferences and included different calculations that very fundamentally in their thought of what constitutes a group. This is how to discover them to find crime rate analysis efficiently. So suitable field need to perform crime analysis, and as information mining is utilized here to process the high volume crime informational index. Crime learning picked up from information mining approaches is helpful and bolster police powers. To perform crime analysis using sentence case information mining approach and cluster role of crime rate analysis with supportive natural language processing. The process analyzes the crime response, reasons and crime rates are easily predict the knowledge discovery of crime prediction.

Figure 1: Crime case analyze the model

. Sentence case observer clustering is the process of combining data objects into groups based on the category of predicting crime rates. The crime intelligence is the activity which generates more specific information about any crime from the large crime data set as shown in figure 1. From the crime data set, the information related to the particular offense has been identified which can be used to perform analysis of the current crimes. The law enforcers have to efficiently meet out challenges of crime control and maintenance of public order. One test for law implementation and knowledge offices is the trouble of investigating large volumes of information engaged with criminal and psychological militant exercises. Hence, the creation of a database for crimes and criminals is needed. Information mining holds the guarantee of making it a simple crime analysis, helpful and pragmatic to investigate enormous databases for associations and crime involvement. Building up a decent crime analysis instrument

to distinguish crime designs rapidly and efficiently for future crime design discovery is the testing field for analysts. Investigative crime analysis includes profiling suspect, and casualties for agents in light of a review of accessible data through cluster ensemble approaches. It is now and again called "criminal investigative analysis" by its group of categories. In high dimensional data centers on speculating about what sort of individual crime is carrying out a specific crime arrangement. Intelligence analysis centers become sorted out crime, psychological oppression, and supporting particular examinations with data analysis using cluster formulation. Investigators can support inspections by turning into the crime rates of data for officers. The waste of testing, the devices of study can be utilized to sort out investigative data and show it as timetables and affiliation interface outlines leads more time complexity.

Relative Clustering depends on discovering connections between various Crime and Criminal qualities having some beforehand obscure regular attributes in crime rates. Association rule mining depends on produce rules from crime dataset in light of frequents event of crime rates to help the leaders of our security society to influence anticipation to activity

Advances in innovation, which permit examines of extensive amounts of information, are the establishment of the generally new field known as crime investigation on the various group of categories. What offense breaks down is a growing field of law requirement without standard definitions. This makes it hard to decide the crime dissects center for offices that are new to the ground. In some police offices, what is called "crime analysis" comprise of mapping crimes for summon staff and delivering crime insights in different organizations, crime analysis may mean concentrating on examining different police reports and suspect data to help specialists in significant crime units.

The intent of new sentence relative cluster ensembles the cluster analyses from crime dataset to make and test the crime rates, and afterward, this information were preprocessed to get perfect and precise information utilizing unique preprocessing systems (cleaning, missing esteems and evacuating irregularity). The preprocessed data were used to discovering distinctive crime and criminal patterns and practices, and crimes and criminals were assembled into clusters as indicated by their critical qualities.

2. Related Works

A massive number of clustering calculations have been distributed from that point forward; K-means is the utilized with an extensive variety of uses utilizing Modified k-means clustering [2]. Different modifications have been proposed in the literature for improving the performance of the k-means clustering algorithm. Proposed method first calculates the initial cluster. They get initial centroid from this initial cluster look at the use of clustering procedure for a statistics quarrying method to help notice the criminalities patterns [3, 4] and speed up the process of solving the crime. This looked at k-means clustering with some enhancements to aid in the process of identification of crime patterns.

The crime pattern applied these techniques to real crime data from a sheriff's office and validated our results optimal local problems of crime rates are detected with k-means which don't carry the cluster forms. The question arises due to improper assignment of cluster weightage of crime novelty with population-based search algorithm called the Bees Algorithm [6] that is capable of locating optimal solutions efficiently.

Document Clustering has been increasingly becoming an essential task for obtaining good results with the unsupervised learning methods [11]. It aims to automatically group similar documents in one cluster using different types of extractions and cluster

algorithms. There are ongoing works done to improve Document Clustering techniques such as Extractions and Clustering in significant data analysis [8,9] approaches to overcome the difficulty in designing a general purpose document clustering for crime investigation [14]and the ill-posed problem of extraction and clustering. In real cases recently linear regression model is used to make future crime prediction of various crimes for Delhi city [15]. With the given information on different crimes for a long time, direct relapse demonstrate is resolved and the offense for a long time ahead is anticipated.

Be that as it may, an absence of reconciliation makes those frameworks less accommodating in practice [16]. A coordinated structure called PerpSearch that will take a given depiction of a crime, including its area, type, and the physical portrayal of suspects (individual attributes or vehicles) as info. To distinguish suspects, the framework will process these contributions through four incorporated segments: geographic profiling, interpersonal organization analysis, crime designs, and physical coordinating.

Finding the classification model for depicting the crime rates [18]. The increasing advent of computerized systems, crime data analysts, can help the Law enforcement officers to speed up the process of solving crimes. To extract the reasoned analysis to the crime pattern which they remove previously unknown, useful information from an unstructured data [19]. Here this have an approach between computer science and criminal justice to develop a data mining procedure that can help solve crimes faster. They record information used to determine and generate course content quality likestudent online usage reports [20]. This intelligence is then directed to the mentors for assessment and incentive purpose for crime enforcement [21].They projected the usage of k-means clustering and self-organizing map to cluster learning objects (learning objects are educational resources.

Agglomerative clustering and Density Based Spatial Clustering with Noise algorithms are used to cluster crime activities based on some predefined cases [22], and the results of these clustering are compared to find the best suitable clustering algorithm for crime detection. It is the identification of events [24], items or observations that do not conform to an expected pattern. -It is also known as anomaly detection.

Today, collection and analysis of crime-related data are imperative to security agencies. The use of a consistent method to classify these data based on the rate and location of occurrence, detection of the hidden pattern among the committed crimes at different times, and prediction of their future relationship is the most critical aspects that have to be addressed. To detect credit card fraud, computational intelligence has been widely used and plays a vital role. The selection of the preeminent classifier algorithm for a fraud detection system is a pervasive problem [25]. The classifier algorithms only don't make that have been most precious for crime detection.

All the methods discussed above have the problem of producing false classification and poor results in crime intelligence.

3. Crime Case Reasoning Based Knowledge Discovery Using Sentence Case Relative Clustering

The crime knowledge remains the attribute observational and readable study observer for a lot of problems due to the high dimensional dataset. The crime analyst may choose an attribute range and one or more types of crime from specific person id involvement and cluster to predict the category. The resulting dataset of data becomes the input source for the data mining processing for cluster evaluation. The relative analyses use the attribute raw score of crime occurrence with knowledge discovery of prediction rates.

These predefined clusters are defined as attribute case of crime record values and the weights. The resulting, groups have the possible crime patterns. These resulting clusters are categorized based on the crime category weightage clusters. The different clusters or the crime patterns resemble the index terms in sentence case analysis in reason analysis. For each group, the legend provides the total number of crimes incidents included in the group along with the essential attributes that characterize the group. This evidence of categorized crime archives is useful for the investigator to the appearance at when reviewing the predicted crime clusters. Finally, the average mean weightage of relative clustering centroid weightage identifies the cluster class labels of crime category.

The crime contributions carry out the following steps

1. To identify the nature of crime and the crime prevention process and Extracting named entities from narrative reports.

2. To sentence case analyzer that supports clustering count attributes and categorization rule dependent to experiment with criminal records.

3. To build a query evaluation attribute case reasoning model. To interpret analyze the results of the model that how strong is the model to extract crime data patterns.

4. To compare the clustering and association rules data mining techniques and select the one which performs the best results.

5. To compare our proposed model with some new working and finally to forward recommendations based on the findings of the study.
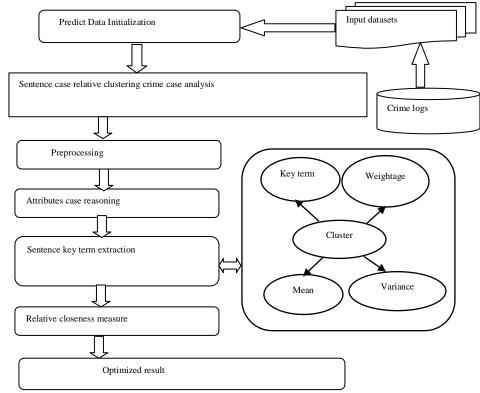

Figure 2: Architectue for sentence case crime reasoning

For this crime reasoning situation, it can help the experts to distinguish crimes rate analyses with average relative mean cluster and help to settle on speedier choices to reviews the attributes weightage to categorize the crime rates. They crime phrasing a cluster is a gathering of crimes in a topographical locale or a problem area of crime. While in information mining wording a group is gathering of comparative information focuses which can be a conceivable crime design as shown in figure 2. Along these lines, proper clusters or a subset of the cluster will have a balanced correspondence to crime designs. In this way clustering calculations in information, mining is proportional to the undertaking of recognizing gatherings of records that are comparable between themselves yet not the same as whatever is left of the information in relative clustering. For our situation, some of these clusters will be valuable for distinguishing a crime blast carried out by one or same gathering of suspects. Given this data, the following test is to discover the factors giving the best clustering.

Data Collection: The data set is the collection of the field in the data from crime analyzer using web resources. The dataset which consists of the text from various attributes related to different case crimes involvement.

A) **Preprocessing**: A data preprocessing is a process that consists of data cleaning, data integration and data transformation of records to reduce dimensionality. It intends to reduce some noises, incomplete and inconsistent data. The results from the preprocessing step can be later proceeding by data mining algorithm.

B) **Clustering:** Crime rates in data clustering process is putting similar crime data into groups. A clustering algorithm partitions the data set with sentence case observer with related Relative model groups crime measure such that the similarity within a group is more substantial than among groups. Based on records are categorized on crime risks weightage.

Data Collection Phase

In this phase, the dataset that used the training and testing data were extracted from the police department. These data contain data about both Crimes and Criminals with the following primary attributes:

- Crime ID:unique crime IDs designate individual Crimes.
- Crime Type: indicates crime type
- Date: report when a crime happened.
- Gender: MALE or FEMALE and total involvement.
- Age: age of Individual Criminal.
- Crime Address: location of the crime.
- Marital status: status of the Criminal.
- Reason: sentence case observation.
- More than 350 crime records were collected to test the proposed model.

Preprocessing

Preprocess holds the collected dataset, Real world data usually have various problems due to Incompleteness, Noisy, and Inconsistency. So, preprocess state reduction proceeds which to get the data suitable for analysis of the weightage values. Purpose of the method reads the crime data point. From the data points retrieved, the method identifies the list of crimes and list of the unique attribute. Then for each data point, the method verifies the presence of all the dimensions. If any of the data points have been identified as incomplete, then it will be removed from the dataset. The preprocessed data set will be used to perform clustering. The preprocessing includes the following tasks.

- Data cleaning: fill in missing values, smooth noisy Data, identify or remove outliers, and resolve Inconsistencies
- Data integration: using multiple databases
- Data Transformation.
- Data reduction and Data discretization

Algorithm

Input: Crime Set Cds
Output: Preprocessed Set Ps
Start
Step 1 :Input dataset cds➔c1, c2,c3,,are the records
        For (c check NULL)
Remove record for the Null field.
              Verify Attribute Set As Fill case.
     Reorder crime records CS←cts(crime test records)
      End for
For (distinct and miss attribute)
Fill CS➔attribute value nearer
End for
Step 2: Analyze the reasoning term index word sentence
 If (CS!=null and attribute fill As)
Step 3: Identify the list of all distinct attributes As.
As = $\int_{i=1}^{size(Cds)} \sum Cds(i).Attribute \ni As \rightarrow$ sentence case index terms
        Remove stop words
        Cs = $\int_{i=1}^{size(Cds)} \sum Cds(i).Crime \ni Cs < --$ stop words
        Step4: For each data point Di contains stop words removes.
            If di contains $cs < --$ attribute count words
            Add count words to processed set Ps attribute
Add to preprocessed set Ps.
                Else
                        Remove the noisy dataset ns
                End if
            End for
 End if
Stop

The above algorithm prepares to process data which is non redundant the original data much reduce the noisy terms of unfilled, distinct case of original records. The preprocess completes from redundant factor holds the crime reasoning records cs, outcomes the preprocessed dataset.

Attribute Case Reasoning Estimation

To take care of the different attributes of different crimes types, this introduced the concept of weighing the characteristics in reason analysis in crime records in sentence case. This allows placing different weights on different attributes dynamically based on the crime types being clustered. This categorization of weightage factor to allow, unlike just the numerical attributes that can be easily scaled for weighting them. Using the integral weights, the categorical attributes can be replicated as redundant columns to increase the effective weight of that variable or feature.

Given a transactional data set Ts to be assigned Ts←ps, the number of attributes Na is identified, and the total number of transaction Tn is computed. Possible patterns identified attributes holds the set Ps is calculated. Unlike generalized frequent pattern mining algorithm the support and count values are processed for the detected models in the pattern set Ps. Using computing support and count values the pattern instance Pi which has support value more significant than support threshold St will be selected for sanitization process, and other will be omitted.

Step 1: start
Step 2: read data set Ts from cs(ps).
Step 3: identify attribute set As.
Step 4: Compute the number of Crime instance per age, weapons, groups Tn, K, S.
Step 5: if (compute a combination of possible crime patterns =valid term)
        Ps = ((Tn * As)/(K (Tn – K) * As)).
k- Specific crime instance related attributes values.
Step 6: for each pattern Pi from Ps
Compute count = ø(Pi£(Ps)).
Ø- Number of crime instance pi contained in ps.
Compute support = count/Tn.
        If support means $\rightarrow$ ST then.
Add Pi to filter dataset Ss.
Ss←Average mean
End.
End for
Step 7: return Ss crime substitution occurrence weightage.
Step 8: End.

The usage of weightage factors analyses using filters to appropriate for clustering elsewhere in average mean value, as upon normalization all attributes assume equal importance in the clustering algorithm. However, rule mining predicts the average mean weight for ordering attribute records based on the category. Based on our weighted clustering attributes, cluster the dataset for crime patterns and then present the results to the detective or the domain expert along with the statistics of the critical characteristics.

Scaling Features on Sentence Case Clustering Analysis

The crime at the sentence case clustering, records are clustered based on the predetermined attributes and the weights. The resulting, clusters have the possible crime patterns the different clusters or the crime patterns are pointed to the relative score. For each group, the legend provides the total number of crimes incidents included in the group along with the critical attributes that characterize the group. The scaling features specify incidence of crime category. To evaluate results for the detected crime patterns by examining the suspect ions of crime were accepted or rejected on category scaling measure, which the measured regarding the essential attributes or

features or variables of crimes are demographic crime pattern, the suspect, the victim, etc.

Algorithm

Input: attribute sentence case records Cr:

Output: cluster groups relative measure

Step 1: Assign initial value for mean depth points

Cr1, Cr2….Crn. Di1, Di2,…

Analyze relative sentence key terms.

For each crime records D_i from Ds.

Text T = Extract Text from Di.

Sentence set Ss= ($\sum_{n=1}^{size(Ds)}$ Text€Di) × Splitby(.)

For each sentence Si from SS

Continue each Si++

End For

End for

Step 2: Assign each item xi individual for the cluster which has the closest mean;

Compute semantic linkage weight Slw. Eccentric measure ELV, Initial tag ILV, Nearest relative Nr

Semantic weightage $Slw = (\frac{ILV}{Nr} \times \frac{ELV}{Nr}) + NIL$

Add to weight set Ws = $\sum$Ws(Di)+Slw

Step 3: Calculate new mean depth for each cluster;

For each crime data instanced from So

Compute the mean value of semantic linkage weight.

Mean Slw = $\frac{\sum Slw(Ws(Di)}{size(Di)}$

End for

Choose the most valued crime data Di.

Class C = max(mean).

Until the cluster was performed;

Step 4: int result=0;

Step 5: for each D in intermediate values

Relative are added;

End for

Step 6: result += (D);

Step 7: Compact crime measures (Risk Crime state);

Subsequently, the cluster the crimes based on sentence case weighing technique, to come crime scores of risk states, which contain the possible crime patterns of crime states. The categorization resembles the attributes in the record for different situations of crime patterns along with the essential characteristics related to particular crime. To quantify these groups attribute weightage determines the crime risks presented to the list of previous crime incidents in separate orders or some predetermined sort order.

Crime Scaling Risk Factors

Scaling factors evaluate cluster resultant to measure and split carrying the crime category of law enforcement of average weightage. Classification ensembles the categorization of high, medium and low factorization of crime risk categorization on the medium which crime factors be availed. The scaling range of crime records finalize the crime patens of the crime rate. This integrates the feature analyses of crime prediction supportive for law enforcement.

4. Results and Discussion

The collected data information are testes with sentence case reasoning proposed system results are carried out through .net framework 4.0 with visual studio front end, UCI real case datasets are carried out criminal datasets for input progress of testing. The proposed SCRCA based clustering algorithm has been implemented and tested for its efficiency with crime category analysis. The implementation provides a higher efficiency result compared with previous test case results,

Table 4.1: Details of a Dataset

Parameter	Value
Number of Records	30000
Number of Crimes attributes	10
Number of Subclasses	5

Table 4.1 Reviews the datasets from the criminal records value contains attributes and information processed in the data collection phase. The performance of SCRCA is evaluated through clustering accuracy (cs), false classification ratio (Fcr), time complexity (Ts) and frequent occurrence (Fs) The resultant figure given below shows the performance,

Clustering accuracy (cs) $= \sum_{k=0}^{k=n} \times \frac{total\ number\ of\ cluster\ group\ dataset(Cds)}{Total\ originate\ data\ records(Tr)}$
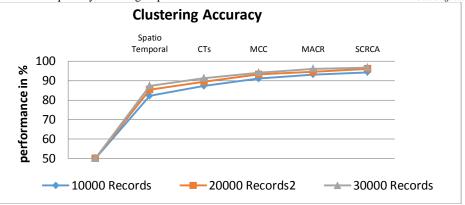

Figure 4.3: Comparison of clustering accuracy

Figure 4.3, shows the comparison of clustering accuracy and shows that the proposed method has produced higher clustering accuracy than other methods.

Table 4.2: Comparison of Clustering Accuracy

Methods/number of records	Spatiotemporal	Cts	MCC	MACR	SCRCA
10000 records	82.2	87.3	91.1	93.1	94.3
20000 records	85.4	89.5	93.2	94.5	95.3
30000 records	87.4	91.3	94.1	96.1	96.7

Table 4.2, shows the comparison of clustering accuracy produced 10000records as 94.3%, 20000records as 95.3% and 30000records as 96.7 % indicates that the proposed approach has produced higher clustering accuracy.

The similar datasets are ignored as unclassified clusters be considered as false classification, the false classification is calculated by

False classification Ratio (Fcr) $= \sum_{k=0}^{k=n} \times \frac{clustering\ Accuracy(cs)}{Total\ no\ of\ failed\ cluster\ rate(Fr)}$
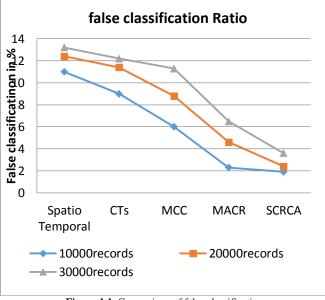
Figure 4.4: Comparison of false classification

The Figure 4.4, shows the comparison of false classification ratio produced by different methods and the proposed plan has generated less false classification ratio than other ways.

Table 4.3: Comparison of False Classification

Methods/number of records	Spatiotemporal	Cts	MCC	MACR	SCRCA
10000 records	11	9	6	2.3	1.9
20000 records	12.4	89.5	8.8	4.6	2.4
30000 records	13.2	12.2	11.3	6.5	3.6

The Table 4.3, shows the comparison of false classification ratio produced 10000records as 1.9%, 20000records as 2.4% and 30000records as 3.6 % indicates that the proposed approach produces less false classification ratio.

The average time taken to calculate the dataset cluster evaluation is calculated by,

Time complexity (Tc) $=\sum_{k=0}^{k=n}\times\frac{clustering\ Accuracy(cs)+false\ classification(Fcr)}{Time\ taken(Ts)}$
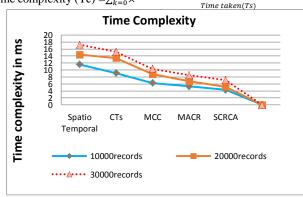

Figure 4.5: Comparison of time complexity

The Figure 4.5, shows the comparison of time complexity produced by different methods and shows that the proposed approach has produced less time complexity than other methods.[26]

Table 4.4: Comparison of Time Complexity

Methods/number of records	Spatiotemporal	Cts	MCC	MACR	SCRCA
10000 records	11.6	9.1	6.3	5.3	4.1
20000 records	14.4	13.4	8.8	6.6	5.2
30000 records	17.2	15.2	10.3	8.5	7.3

The Table 4.4, shows the comparison of time complexity SCRCA produced 10000records as 4.1(ms), 20000records as 5.2(ms) and 30000records as 7.3 (ms) shows that the proposed approach has produced less time complexity.

Frequent occurrence (Fc) $=\sum_{k=0}^{k=n}\times\frac{repeated\ clusters(Rs)+irrelavant\ clusters(Irc)}{Total\ number\ of\ clusters}$
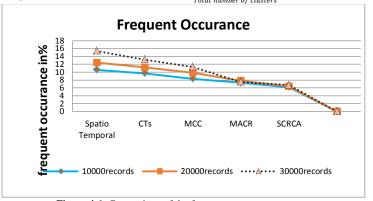

Figure 4.6: Comparison of the frequent occurrence

The Figure 4.6, shows the comparison of frequent occurrence produced by different methods and shows that the proposed approach has produced less frequent than other ways.

Table 4.5: Comparison of Time Complexity

Methods/number of records	Spatiotemporal	Cts	MCC	MACR	SCRCA
10000 records	10.6	9.7	8.3	7.3	6.1
20000 records	12.4	11.2	9.8	7.8	6.4
30000 records	15.4	13.2	11.3	7.5	6.7

The Table 4.5, shows the comparison of frequent occurrence SCRCA produced 10000records as 6.1%, 20000records as 6.4% and 30000records as 6.7 % shows that the proposed approach has produced less regular occurrence.

5. Conclusion

In this paper, an efficient sentence case relative clustering for crime analyses has been presented. The reasons are carried out through datasets with sentence case record with relative closeness point of measurement. Then this method creates the cluster evaluation of record group with a similarly categorized dataset. The categories are split with crime rate support measure be ordered in cluster groups which are high, medium, low crime states. The projected method which produces crime case prediction intelligence generates highly and produces efficient results on clustering accuracy 98.1 % and intelligence generation. This resultant had higher performance with low time complexity in 8.5 milliseconds. Also, the method reduces the frequent occurrences of cluster group evolution 6.7 % as well. The features from clusters that determine higher prediction crime types which is supportive for real-world kinds.

References

[1] Han J & Kamber M, *Data mining: concepts and techniques*, Jim Gray, Series Editor Morgan Kaufmann Publishers, (2000).

[2] Rodriguez JJ, Kuncheva LI & Alonso CJ, "Rotation forest: a new classifier ensemble method", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, (2006), pp.1619-1630.

[3] Bhattacharyya S, Jha S, Tharakunnel K & Westland JC, "Data mining for credit card fraud: a comparative study", *Decision Support Systems*, Vol.50, No.3,(2011), pp. 602-613.

[4] Nath SV, "Crime pattern detection using data mining", *Proceedings of the IEEE/WIC/ACM Web Intelligence and Intelligent Agent Technology international conference*, (2006), pp.41-44.

[5] Ferligo A, "Recent developments in cluster analysis", *Telecommunication Systems*, Vol.1, No.4, (2003), 205-220.

[6] Pham DT, Otri S, Afifty A, Mahmuddin M & Al-Jabbour H, "Data clustering using the Bees algorithm", *proceedings of 40th CRIP International Manufacturing Systems Seminar*, (2006).

[7] Chen H, Chung W, Xu JJ, Wang G, Qin Y & Chau M, "Crime data mining: a general framework and some examples", *computer*, Vol.37, No.4, (2004), pp.50-56.

[8] Arora S & Chana I, "A survey of clustering techniques for big data analysis", *5th International Conference-Confluence The Next Generation Information Technology Summit,* (2014), pp.59-65.

[9] Kaur PJ, "A survey of clustering techniques and algorithms", *2nd International Conference on Computing for Sustainable Global Development (INDIACom),* (2015), pp.304-307.

[10] Perry WL, McInnis B, Price CC, Smith SC & Hollywood JS, *Predictive Policing The Role of Crime Forecasting in Law*, RAND Corporation, (2013).

[11] Bsoul Q, Salim J & Zakaria LQ, "An intelligent document clustering approach to detect crime patterns", *Procedia Technology*, Vol.11, (2013), pp.1181-1187.

[12] Gera P & Vohra R, "City Crime Profiling Using Cluster Analysis", *International Journal of Computer Science and Information Technologies*, Vol.5, No.4,(2014), pp.5145-5148.

[13] Ying K, Chang M, Chiarella AF & Heh JS, "Clustering students based on their annotations of a digital text", *IEEE Fourth International Conference on Technology for Education (T4E),* (2012), pp.20-25.

[14] Joshi S & Nigam B, "Categorizing the document using multi-class classification in data mining", *International Conference on Computational Intelligence and Communication Systems*, (2011).

[15] Gera P & Vohra R, "Predicting Future Trends in City Crime Using Linear Regression", *International Journal of Computer Science & Management Studies*, Vol. 14, No.07, (2014).

[16] Ding L, Steil D, Hudnall M, Dixon B, Smith R, Brown D & Parrish A, "PerpSearch: an integrated crime detection system", *IEEE International Conference on Intelligence and Security Informatics,* (2009), pp.161-163.

[17] Bogahawatte K & Adikari S, "Intelligent criminal identification system", *IEEE 2013 The 8th International Conference on Computer Science & Education*, (2013).

[18] Babakura A, Sulaiman N & Yusuf M, "Improved method of classification algorithms for crime prediction", *International Symposium on Biometrics and Security Technologies,* (2014).

[19] Sathyadevan S & Gangadharan S, "Crime analysis and prediction using data mining", *First International Conference on Networks & Soft Computing (ICNSC),* (2014), pp.406-412.

[20] Saxena R, "Educational Data Mining: Performance Evaluation of Decision Tree and Clustering Techniques Using WEKA Platform" *Business Informatics as International Journal of Computer Science*, Vol.15, No.2, (2015).

[21] James Manoharan J, Hari Ganesh S & LovelinPonnFelcia M, "Discovering Student's Academic Performance Based on GPA using k-Means Clustering Algorithm", *IEEE World Congress on Computing and Communication Technologies*, (2013).

[22] Sivaranjani S, Sivakumar S & Aasha M, "Crime prediction and forecasting in Tamilnadu using clustering approaches", *Emerging Technological Trends (ICETT)*, (2014).

[23] Dutt A, Aghabozrgi S, Ismail MAB & Mahroeian H, "Clustering algorithms applied in educational data mining", *International Journal of Information and Electronics Engineering*, Vol.5, No.2,(2015), pp.112-116.

[24] Kaur S & Singh W, "Systematic Review of Crime Data Mining", *International Journal of Advanced Research in Computer Science,* Vol.8, No.5, (2017).

[25] Mahmud MS, Meesad P & Sodsee S, "An evaluation of computational intelligence in credit card fraud detection", International Computer Science and Engineering Conference (ICSEC), (2016), pp.1-6.

[26] G Mussabekova, S Chakanova, A Boranbayeva, A Utebayeva, K Kazybaeva, K Alshynbaev (2018). Structural conceptual model of forming readiness for innovative activity of future teachers in general education school. Opción, Año 33. 217-240