Original Article

Social media data analysis to predict mental state of users using machine learning techniques

R. Lokeshkumar, Om Ashish Mishra, Shivam Kalra

Abstract:

BACKGROUND: Social media platforms such as Facebook, WhatsApp, and Instagram etc., are becoming very popular now not only for youth but for all walks of life. People are more often seen in busy in tweeting, chatting, or putting selfies. No one actually knows the mental state of a person in the online platform. In this article, we will be focusing on how social media is affecting issues such as road accident, murder, and suicide. The research is done by three parts.

MATERIALS AND METHODS: Google Form analysis, machine learning used for prediction, and by sentimental analysis of what people think in twitter. All the datasets are based in India. From these datasets, the different machine learning algorithm is used to do the analysis. The project strives to bring the real-world solution in the matter of advancement.

RESULTS: The static data analysis and dynamic data analysis shows the various sentimental analysis and predictions and the technique to predict different mental states. Thus we get clearly about the current world is getting into social issues. This research findings helps to bring social awareness among the current generation by understanding the sensitivity of the youths.

CONCLUSION: Thus through this paper we get known clearly how the current world is getting into social issues like victim of murders or road accidents or committing suicide. The paper clearly helps us to understand the sensitivity of the youths. Therefore brings a social awareness among the current generation.

Keywords:

Clustering, mental state, prediction models, sentiment analysis

Introduction

Social media has become the home of the current generation apart from personal life and work. People are giving their precious time into Facebook or Instagram or Snapchat, what people are basically doing on this platform is to see which friend is liking what, staying where, eating what, listening to which type of music, etc., We also see people are showing off or letting unrealistic stories. Thus, psychologists explain social media is a major part of people's lives in spite of which it is detrimental. They found on various research that people have a

great difference in sleep, body language, and social interactions by logging on too much. The sleep of the people is getting hampered and due to not having quality sleep the people are getting agitated. Thus, it is advised to stay away from mobile 1 h before bedtime. Even the blue ray of the electronics is not healthy for them. It sends the signal to our brain that it is still daytime. Then, our brain releases cortisol hormones, which then keep us in an alert, wakeful state. Scrolling and spending lot of time on social media can affect sleep routines. This results in dullness and sluggishness in the morning, thus affecting people mentally. The body image of different people on the Internet with a perfect body, perfect makeup, and

Department of Analytics, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Address for correspondence:
Dr. R. Lokeshkumar, School of Computer Science and Engineering, Department of Analytics, Vellore Institute of Technology, Vellore, Tamil Nadu, India. E-mail: lokeshkumar.r @vit.ac.in, omashish100@gmail.com, shivam.kalra2016@ vitstudent.ac.in

How to cite this article: Lokeshkumar R, Mishra OA, Kalra S. Social media data analysis to predict mental state of users using machine learning techniques. J Edu Health Promot 2021;10:301.

perfect faces can leave with a feeling of quiet down about their perspective about themselves. This misleads to believe that this is a reality. Self-esteem and body image have been well-documented in studies about the vulnerability of young people. The platforms such as Facebook and Snapchat have created an environment based on rating people and earning approval through comments ratings and likes. This has resulted in strong bonding between body image and social belonging. Even the way the social interaction of people is getting affected by social media. Even a phone in a room can take away the attention of the room. Be it meetings, lectures, hospitals etc., people are always on the standby mode for the next notification. Thus, this constant connectivity with others has become a human need. However, excess of connectivity can result in a negative impact on our mental well-being and our lives.

The article has the main concern with reducing the death rate and understands the human causes and problems through analysis. The three methods that shall be used by us are machine learning, sentimental analysis, and Google Form circulation. The topics which are getting affected by these are suicide, road accident, and victims of murder. The data we received are in recent times starting from 1965 to 2018. This will help in the analysis as well as telling where a particular social issue to places at what time. Social Media which has become so easy to access nowadays, the youth is sometimes getting misguided to take wrong actions. The use of social media should be for the benefit of the people and not for the sake of misinterpretation and wrongness. Thus, the article aims not only to give a visual representation of the people getting affected but also aims to give healthy ways of living to the life of the people. The article is made to spread social awareness and is made with a selfless motive.

As discussed above, the article is divided into three parts:

i. Machine learning analysis on the datasets by classification and clustering, thus comparing the best-suited algorithm for that particular search
ii. Sentimental analysis of the dataset using twitter dynamic dataset, thus classifying for neutral, good, or bad
iii. Analysis through Google Forms for the option of the people for real-time data analysis.

Thus, we draw our understanding from them, see the performance in each case, and check the best-suited algorithms and remedies of these problems.

Suicide is not new in human history rather it is as old as humanity itself and its sources reach far back into the beginning of the culture. It is a specifically human problem. Any animal can die by disease and can be destroyed intentionally or accidentally by an outside agency, but as far as we know only man can will his death and kill himself. At some stage of evolution, man must have discovered that he can kill himself. It is the most personal action, which an individual can take. The study on suicide illustrates that human action, however personal, is also interaction with other people and that the individual cannot understand in isolation from his social matrix. Suicide is widely prevalent and no nation and culture has escaped from it, though the toll varies from place to place. The prevalence of suicide in today's world is quite alarming. In the year 2000, about 800,000 suicide deaths occurred worldwide.

In many countries around the world, injuries are the leading cause of death. Approximately 20% of all unintentional deaths worldwide occur in children under 15 years old and are among 10 leading causes of death. Road accidents account for 21% of all death in this age group. 0–14 year children constitute 30.4% of total population in our country. Accidental death of children accounts for 6.7% of total such death, of which 36.3% are due to road accidents. Road traffic injuries are a leading cause of death in children. Pedestrians are 30 times more involved in accidents as compared to cyclists and car occupants. Road accidents accounted for 55% of all accidental death in children and in almost all of these, the unsafe behavior of a child was considered to be at fault. These road accident deaths occur in healthy children who might have been expected to have had productive lives and cause immeasurable distress and guilt to the parents and other parties involved. Hence, the prevention of accidents in children is being increasingly recognized as an important public health issue.

Mutilation (the act of removing or destroying a conspicuous or essential part or organ) of a corpse has always been viewed by society to be a more dreadful crime than the homicide itself. "Mutilation-murder" is an extremely rare crime and is defined as "those homicides where the offender tried to dismember the victim." Pu¨sc hel and Koops have suggested four different kinds of homicidal mutilation: defensive: where the motive is to assist in hiding or moving the body, or getting rid of evidence, or making identification of the victim more difficult; aggressive: where the killing and mutilation is brought about by the same aggressive strong emotions; offensive: where the dismemberment is in fact the real purpose of the murder all along (lust and nacro sadistic murders); and necromantic: mutilation carried out on a dead body with a purpose of using some body part(s) as a trophy, symbol, or fetish.

Thus, a lot of research has gone into the analysis of the social media caused issues. Many prediction models have already been tried out to automatically classify the type of injury severity. Decision trees, hybrid learning approaches, and concurrent hybrid model involving

decision trees and the neural network are used to train the neural networks. Even multi-layered perceptron and Fuzzy ARTMAP were also been used. Critical Analysis Reporting and Environment was developed at the University of Alabama using back propagation neural network. The concept of accident index was used to calculate the ratio between the number of accidents for a given intersection and the number of accidents at the most dangerous intersection. Few things such as use of seatbelt and low speed might prevent accidental fatalities. Thus, the application of support vector machines (SVMs) has also been seen. Data analysis is mostly into head-on collision which has the highest percentage of fatal injury.[1]

Another article has talked about the dataset collected from police stations for Krishna district for the year 2013. Using K-medoids and expectation-maximization algorithms, clusters are formed, which are then analyzed to discover hidden patterns using the apriori algorithm. Machine learning techniques are used and the result obtained shows their effectiveness, the result clearly shows the exact hidden pattern from the data. Accident data visualization is done using the density histograms. The proposed system consisted of the human role, vehicle, and infrastructure-related factors in the accident. The results were seen due to the numerous combination of human and vehicle faults.[2]

Road accidents have always been a big issue in most parts of the world. Traffic-related action analysis language and ID3 algorithms have to be a major contributor in this area of research. Multidisciplinary accident investigation teams, self-report studies using interviews and questionnaires, and finding meaning general patterns have helped in the analysis of the accidents. Genetic algorithms are used to generate predictive rules that would discriminate injury from damage-only accidents. The algorithm focused on the deterministic and nondeterministic approaches. This involved drawing of the decision tree to classify into "onto" accident and "off" accident. Tree-growing algorithm was causing the dataset to grow too big through pruned decision trees are used now. The research mainly seemed to concern about severity and gender differences.[3]

Social media is required for awareness of human factors affecting traffic safety which has been seen. Human error is responsible for about 85% of all car accidents. The study aims to bring to the notice the contribution of social network sites to bring out the awareness of traffic safety among the people. The article emphasizes on the role of social media for the cause. The recommendations were set according to the deficiency and intensity of the identified factors contributing to traffic safety. The aim of these recommendations is to improve the traffic safety in the log

run and solve the current problems. The highest number of Facebook users is in Egypt and Suadi Arabia. Driver behavior is important for road safety. Networking sites have great potential in educating the young generation.[4]

Suicide tendency has being seen occurring in patterns using machine learning algorithms like C4.5 and visualization of data using R language. Seventy percent of the suicide occurred in low-income countries and close to 1 million victims a year have been recorded. The gravity of the problem and the result needs to have psychological analysis. The feeling of shame and guilt and feeling of being a burden to others, the aging of the person, the death of loved ones, the consumption of drugs, illness, etc., have caused many times to quit people's life. To identify and prevent depressive and suicidal behaviors, this research article focused by the simulation and based on the data generated stochastically and adjusted to reflect the reality at Peru. This research article has also shown although the women with less interaction on social media still fall into depressive and suicidal behavior. Through the Viterbi algorithm and electroencephalography, anxiety disorder and stress could be measured. Rotation Forest algorithm was used to calculate the F-measure. R language was used to estimate the C4.5 and naïve Bayes. Weka Tools were used in modeling test and train set. The data was provided from interview-type information. The dataset was used to classify into age categories.[5] Another article mentioned about the machine learning analysis-based algorithms from twitter data. Suicide is one of the top leading causes of death. The methodology was used where SVMs, maximum entropy, and naive Bayes lead to a ranking of term frequency, negotiation, n-grams, and part-of-speech. Semantic analysis of twitter data is a measure based on the methodologies implement with reference available in the wordnet.[6]

Suicide is now even looked after by Facebook though it is suicide prevention protocol. This uses pattern recognition and detection algorithms by community operation team members. These machine learning algorithms are sufficiently high accuracy, but the implementation is not done completely yet. Thus, more data analysis is needed for the time being.[7] Analyzing the connectivity and communication of suicidal users on twitter by checking the content of the tweets of a user was done. Social networking analysis of specific Twitter users connections using mutual graph by was done by taking average of the node degree, graph density and the shortest path length. Thus followed by term frequency-inverse document frequency analysis, n-grams and random sampling of 1000 tweets was sampled from the 4 million data collected over a week. With respect to the connectivity, it was found that the friends and followers graphs of suicidal users did not make much of a difference. Mutual

links of the suicidal people were higher about 78% with respect to nonsuicidal that was 42%.[8]

Machine learning algorithms have also being used to analyze crime data using WEKA, data mining software. The dataset used was based on crime statistical data for the state of Mississippi. The scope of this article is to show how the machine learning algorithms can predict violent crimes by using algorithms used for data mining. The algorithms used were linear regression, additive regression, and decision stump.[9] The evaluation metrics consisted of correlation coefficient, mean absolute error, root mean square method, relative absolute error, and root relative square error. The dataset contained socioeconomic data from 1990 census, law enforcement data from 1990 Law Enforcement and Admin Stats Survey, and crime data from the 1995 FBI UCR Mississippi 2013 Crime dataset. As a result, accuracy is based on error values. Poor performance of the decision stump algorithm could be because of various factors like randomness that exist in the crimes.

Survey of analysis of crime detection techniques has been done using artificial neural networks, decision tree, rule induction, nearest neighbor method, and genetic algorithm; thus, an organized analysis recognizes and determines the pattern of crime. The survey took place on various types of crime such as fraud detection, violent crime, traffic violence, sexual assault, and Cyber Crime. However, this documentation did not give us any stable ways to tackle situations.[10]

Materials and Methods

Study design and setting
First, we went through some records of the past incidence and through the various modes of communication such as newspaper, magazine, social media platforms such as twitter, Facebook, and WhatsApp. We could see that there are my issues and everyday these are getting larger, day by day, and the law is not able to conclude a proper conclusion. Thus, we went through some dataset websites like Kaggle.com, UIC Machine Learning, and data.gov to receive the data to analysis on these current data issues and come to a conclusion for a better understanding of the subject.

Study participants and sampling
The dataset which is chosen for the research is Indian based, this is because we are focusing on road accidents and victims of murder that took place in India. The datasets are divided into state and union territories to an emphasis on the region-wise accident rate. The research articles shown above explained mainly region wise and showed some drawbacks. Therefore, we went through them and found both supervised and unsupervised

learning both play a great role in the analysis of data. Thus, we took this dataset did the analysis to check the performance of the various algorithm to get a deeper insight into the solution as shown in Table 1. The suicide analysis has been done through sentimental analysis by web scraping twitter data using twitter API, thus forming word cloud to see the frequency of the word and impact of it in social media.

The road accident of India has data over month wise for each state and union territories in every month giving us a clear picture of the statistics.

The above two datasets represent the static data analysis through machine learning algorithms of classification and clustering for understanding the better performance. The next dataset is constructed through API calling of twitter for dynamic data analysis. The data analysis of the pre-processed data is done and values are used for further analysis as a sample perspective of it is provided in the Table 2 and Table 3.

This data processing is done to clean the data so that it fits the proper training and testing set. The process of extraction of the dataset using various ways like the removal of missing values, feature scaling, data

Table 1: The dataset features for road accident

Data attributes: Road accident prediction

State/union territories: The regions involved

Year: The specific year of accidents

Months (January-December): This is the monthly road accident details

Total: The number of accident in a year for a particular state

Table 2: The dataset features for victims of murder

Data Attributes: Victims of murder

State/union territories: The regions involved

Year: The specific year of murder

Gender: Male/female

Hours (0-23): The specification time in terms of the time of the day of the murder

Total: The number of accident in a year for a particular state

Table 3: The dataset features for suicide sentimental analysis

Data attributes: Twitter data analysis

ID: Serial number

Tweets: The recently tweeted data

Length: The length of each tweet

ID: The Identification of each tweet

Date: The date of the tweet

Source: The device of input

Likes: Number of likes

RTs: Number of retweets

Sentiment: Classify as neutral or negative or positive

Senti: Represent the classification as 0 or 1 or-1

RTs=Twitter Retweets

integration, and dimension reduction is used before the final algorithms can be run on them. This increases the efficiency of the algorithms and analysis becomes easy.[9,11,12]

Data collection tool and technique

The proposed works compare different machine learning technologies. The data analysis is done on rad accidents dataset first, then victims of murder dataset are analyzed. The following steps are followed as:

1. Importing the required libraries
2. Loading the dataset
3. Removal of unknown or missing values
4. Check through correlation the importance of dataset columns and removal and filtering the dataset.
5. Using feature scaling to change the accessibility of the columns
6. Use the HotOneEncoder function to change from categorical to the numeric dataset.
7. Divide the dataset into train and test set
8. Then fit the different the different classification algorithms to check for the best-supported algorithm.

or

1. Apply clustering algorithms for seeing the computation of the clusters formed
2. Data visualization for better understanding.

The road accident dataset

Figure 1a-d represents different analysis techniques of classification in machine learning.

Figure 2 shows the results of the different clustering analysis under with respect to each year of road accident, as the state-wise representation was very clumsy. The below method represents the optimum number of clusters by checking the critical point on the graph. The most optimum number of the clusters here is 14. The details are given in the Table 4.

The victims of murder dataset

Figure 3a-d represents different analysis techniques of classification in machine learning.

Table 4: Supervised learning algorithms

Supervised learning: Classification	
Techniques	Description
Logistic regression	Predictive learning model. A statistical method of analyzing the dataset.
Decision tree	Incremental development for checking all possible ways to reach the goal
SVM	Closest boundaries are checked for the data classification.
Naive Bayes	Assumes the presence of one attribute with a comparison to other attributes. Generative learning model

SVM=Support vector machine

Figure 4 shows the results of the different clustering analyses under with respect to each year of road accident, as the state-wise representation was very clumsy. The below method represents the optimum number of clusters by checking the critical point on the graph. The most optimum number of the clusters here is 14. The details are given in the Table 5.

Twitter data analysis

Thus, we now go for the sentimental analysis using twitter API. The data get stored in a tabular format as show in the figure given below.

Results and Performance Evaluation

The different methods are used to represent the data apart from these.

Classification and clustering data analysis

Thus, the above-analyzed algorithms provides better results and explored to know the following information such as Confusion matri*x* for logistic regression.

Tweets	len	ID	Date	Source	Likes	RTs	Sentiment	Senti
Did the Mu	140	1.05E+18	########	Twitter We	1265	1020	Neutral	0
In the spac	140	1.05E+18	########	Twitter We	998	304	Neutral	0
Today, we	139	1.05E+18	########	Twitter We	806	363	Neutral	0
@tcmilber	139	1.05E+18	########	Twitter for	0	0	Neutral	0
@schwifty	41	1.05E+18	########	Twitter for	0	0	Neutral	0
@MazinRi:	120	1.05E+18	########	Twitter Lit	0	0	Negative	-1
one day ya	68	1.05E+18	########	Twitter for	0	0	Neutral	0
I cant affo	140	1.05E+18	########	Twitter for	0	0	Neutral	0
These are	82	1.05E+18	########	Twitter We	0	0	Neutral	0
Get ready	67	1.05E+18	########	Twitter for	0	0	Positive	1
Liddle is a	101	1.05E+18	########	Twitter We	0	0	Neutral	0
Today we	140	1.05E+18	########	Twitter for	0	0	Neutral	0

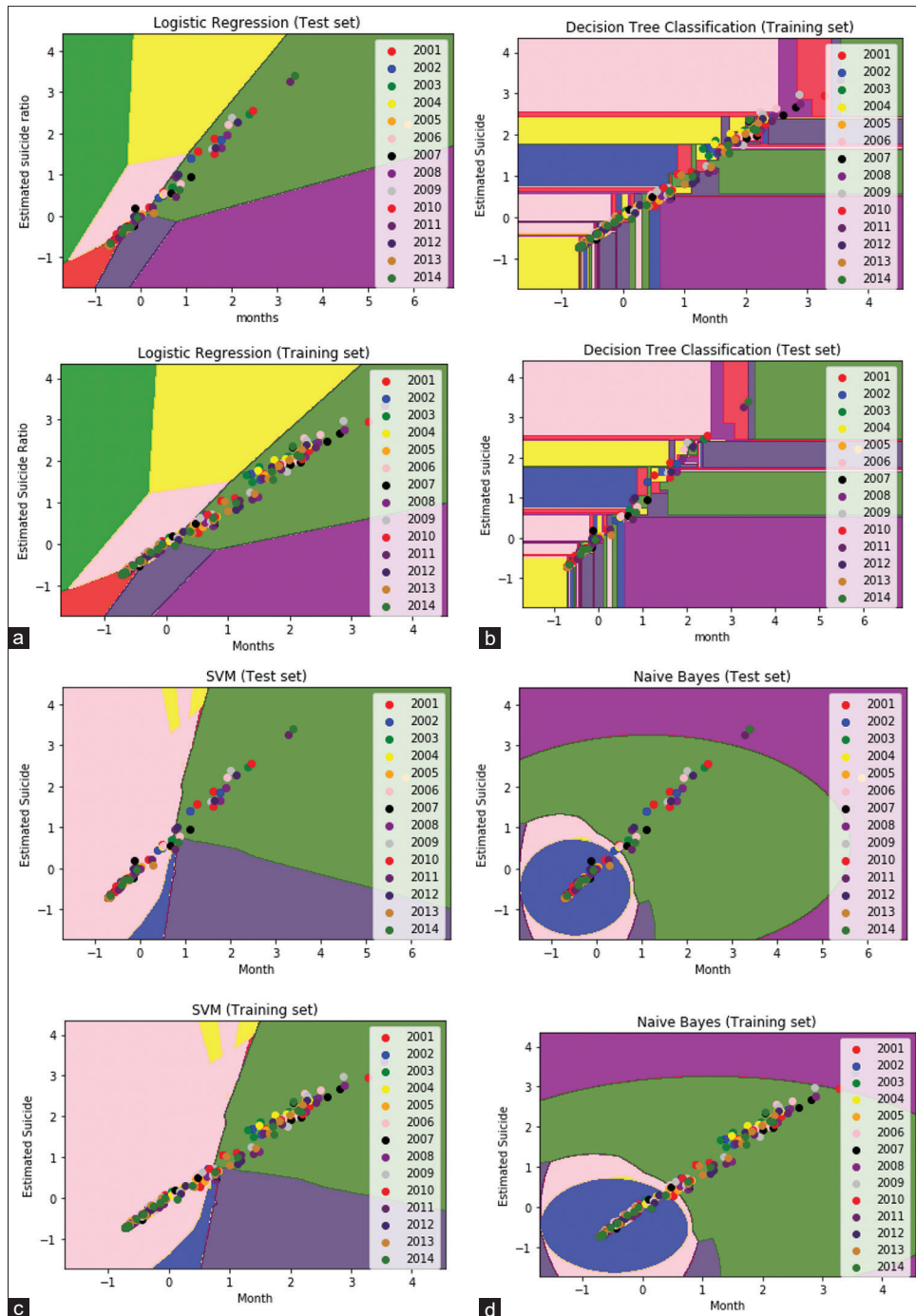

Sentiment of 20 Tweets about suicide
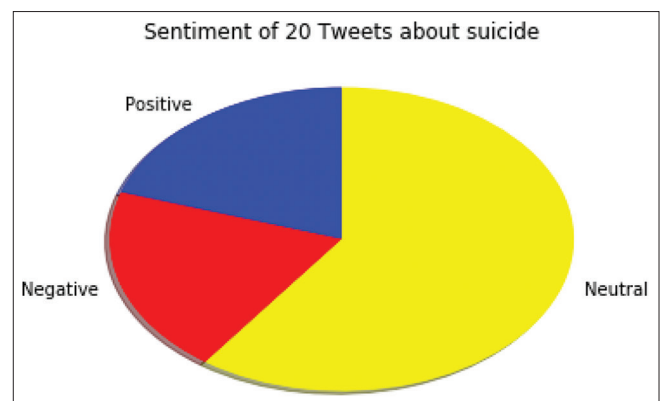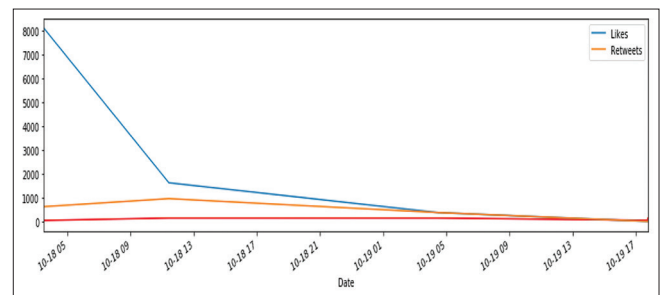
Figure 1: (a) Logistic regression. (b) Decision tree. (c) Support vector machine. (d) Naïve Bayes

Twitter suicide sentimental analysis

Percentage of positive tweets: 20.0.

Percentage of negative tweets: 20.0.

Percentage of neutral tweets: 60.0.

Google Form data analysis

A survey using Google Form is collected to carry out the perspective of general people to have knowledge about various kinds of crimes and their causes with the measures in which normal citizens encounters in their day to day lives.

Some of the results that were obtained through the survey were:

I. Crime rates are increasing in India and social media such as Facebook, twitter, and Whatsapp play a role in instigating such behavior
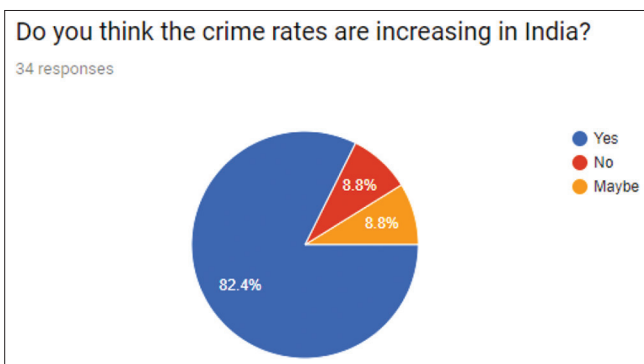
II. Nowadays, people are often depressed or contemplate
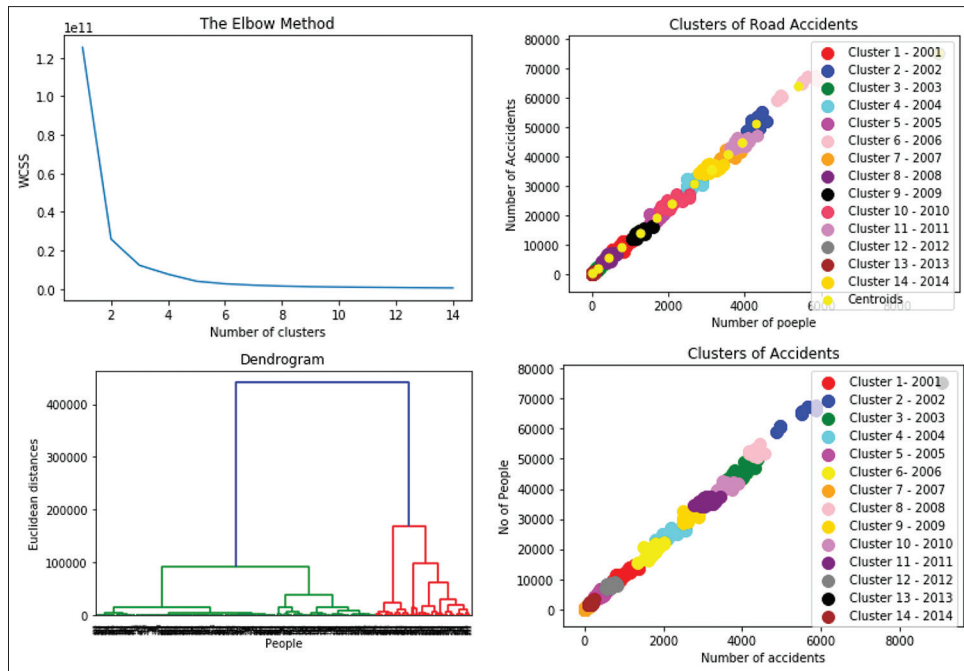
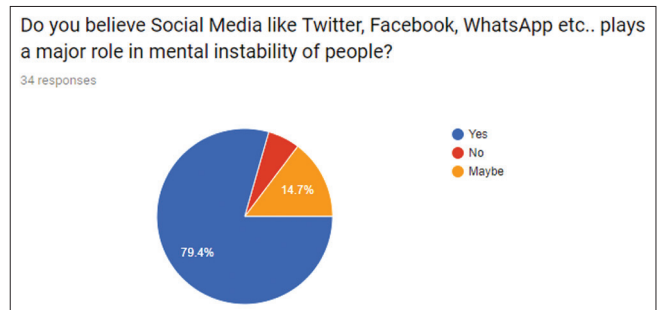**Figure 2:** Clusters of victim

suicide due to a hectic lifestyle, people are mentally fatigued
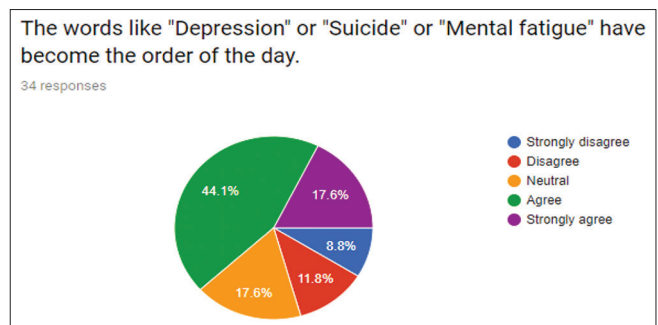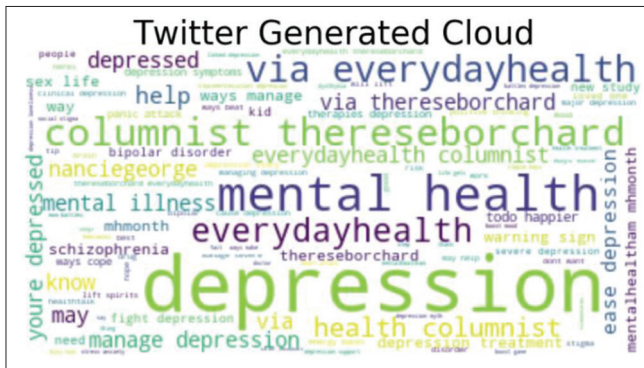
III. Social media is a very powerful tool. However unfortunately, instead of using such a powerful tool for the betterment of society, some people are using it to promote hatred and to exploit somebody else's life. Even people who are not very educated are being brainwashed by the influence of social media

IV. Before, only drunk and driving used to be a major cause of road accidents. However nowadays, using social media while driving, for example, texting while driving or taking a selfie while driving, is also becoming a major reason for road accidents

V. As the saying goes, an ocean is made by adding water drop by drop. Similarly, taking small steps one at a time can lead to recovery and reduction of all these
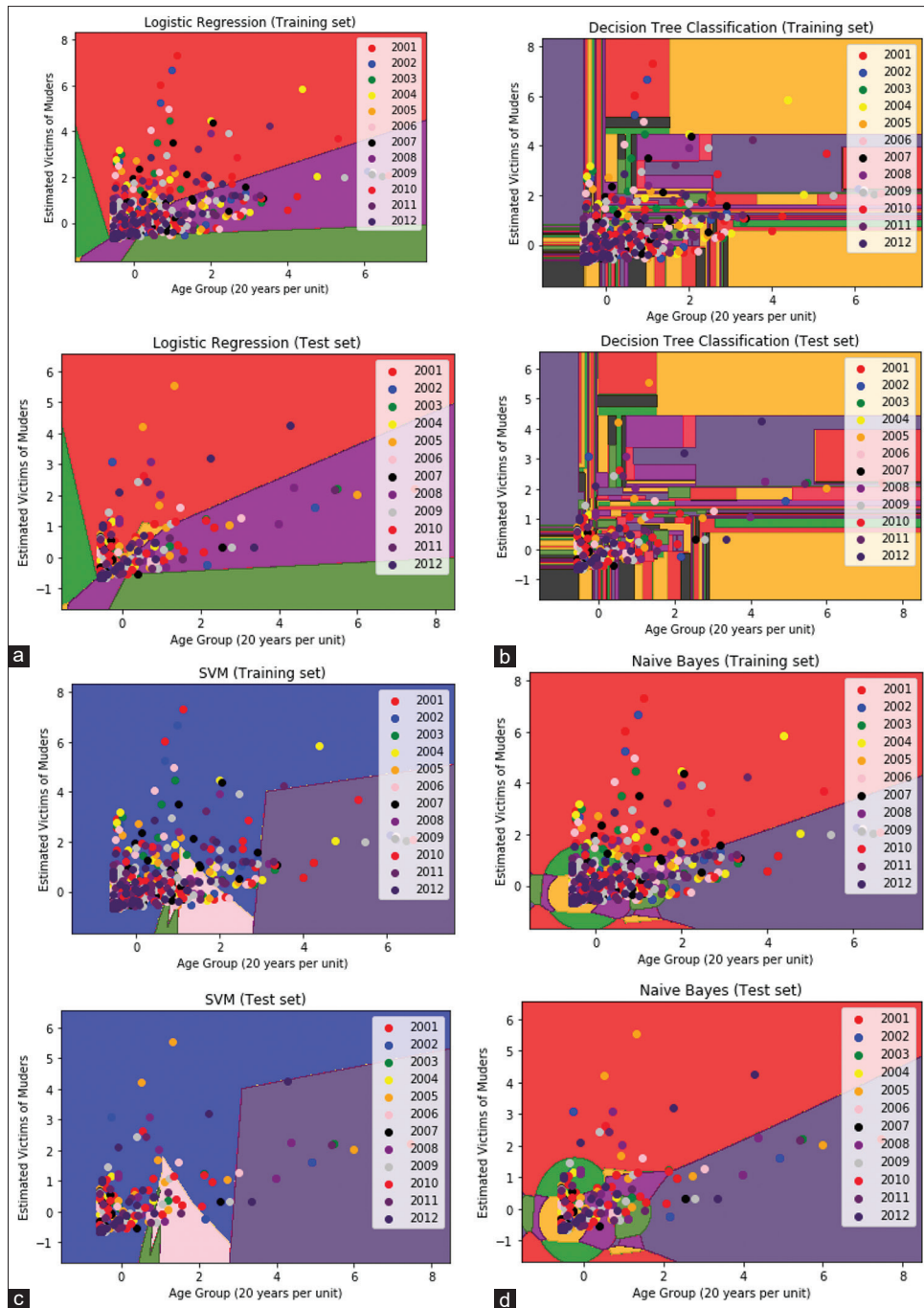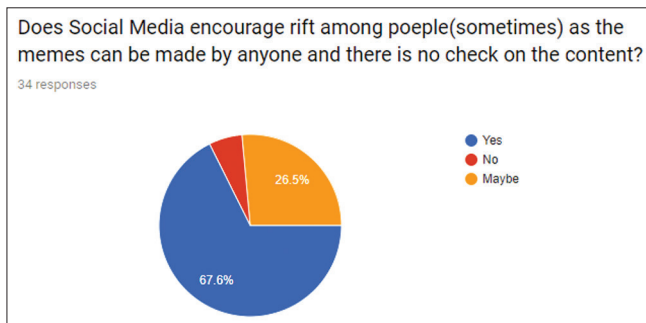
Figure 3: (a) Logistic regression. (b) Decision tree. (c) Support vector machine. (d) Naïve Bayes

dreadful events. People can use social media in a more responsible and disciplined manner. Also, if people just think and respond to a situation rather than reacting furiously to anything that can also be a misinterpretation or a misunderstanding that can lead to lives being saved.

Conclusion
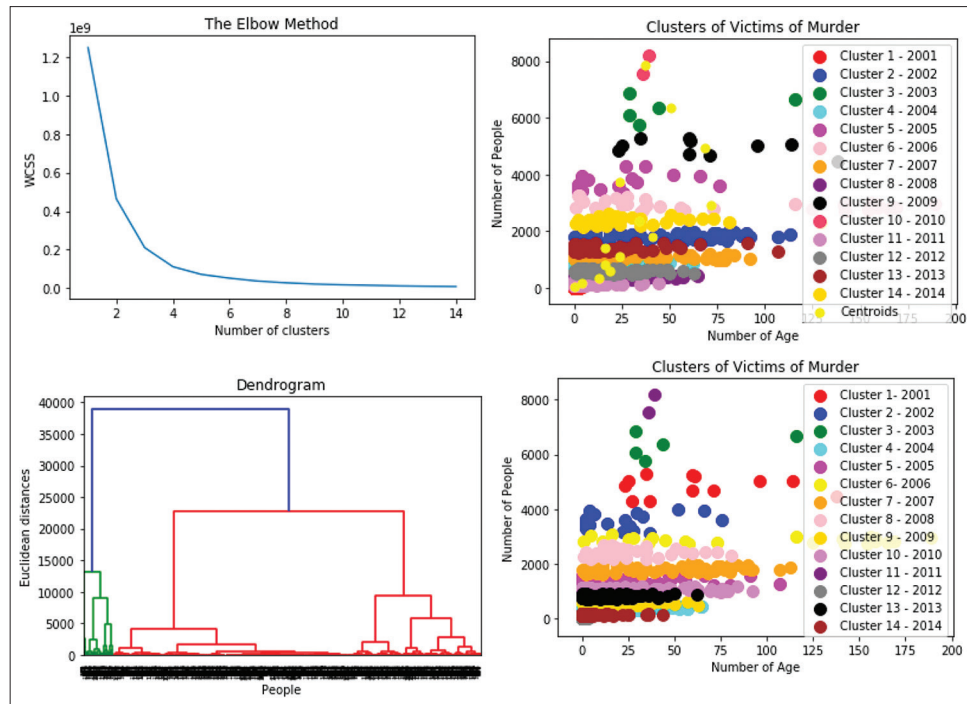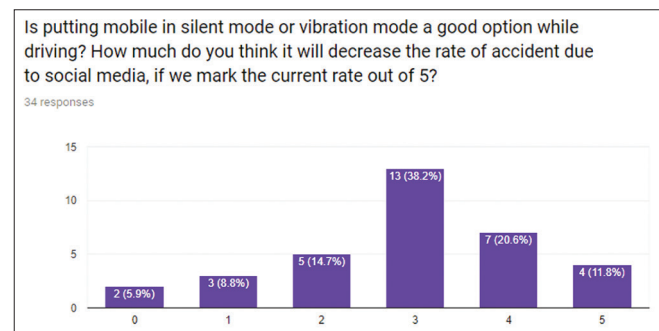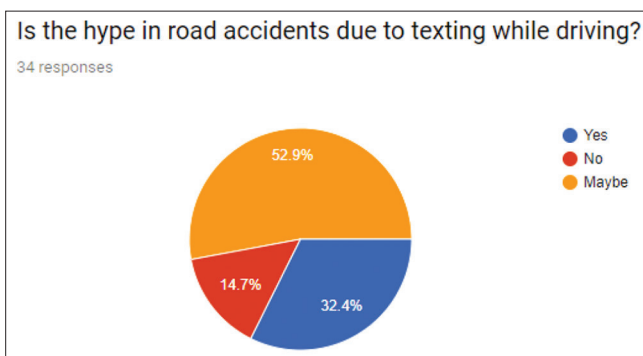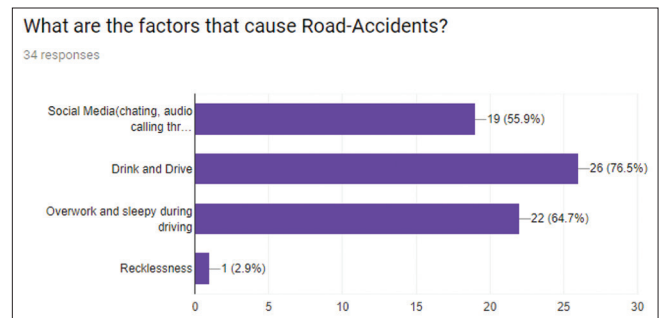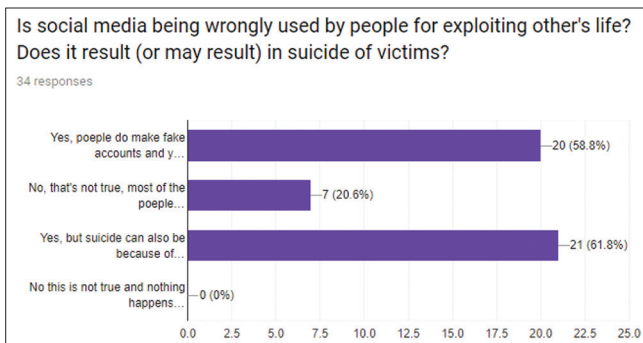
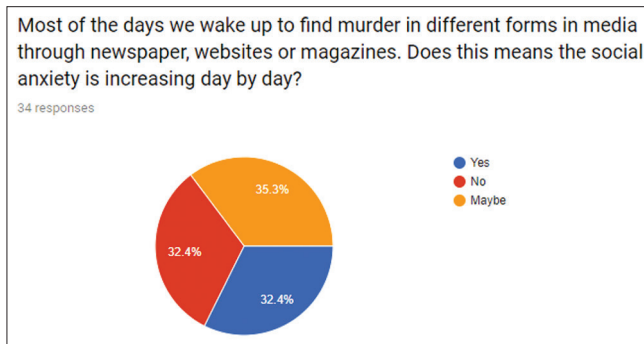- The static data analysis shows the various prediction

Figure 4: Clusters of Victim Murders

and the technique to predict an accident of roads and victims of murder. Dynamic data analysis shows the various sentimental analysis like how most of the people remain aloof to the suicide and keywords are being represented in word clouds. Overall view of the data by use of prediction and machine learning algorithms to predict ups and downs in the rate of each above topic. The Google Form shows the thoughts process of the current mentality of people for the things going around related to social issues. Moreover, we are seeing how the road accident is increasing day by day and the data analysis also show that logistic regression and SVM suits the best for the prediction by data visualization. The clustering algorithm also shows how the data analysis getting a rise in deaths. Thus

Do you drive car and use mobile at the same time?
34 responses

- Yes — 11.8%
- No — 64.7%
- Maybe — 23.5%

Can murder be a cause for social media misuse or mis-interpretion?
34 responses

- Yes — 41.2%
- No — 26.5%
- Maybe — 32.4%

Most of the days we wake up to find murder in different forms in media through newspaper, websites or magazines. Does this means the social anxiety is increasing day by day?
34 responses

- Yes — 32.4%
- No — 32.4%
- Maybe — 35.3%

Can you suggest ways to tackle depression and suicidial thoughts?
25 responses

Share your problems with family friends or seniors
Proper counseling and pessimistic thinging
By yoga and pranayama
Being positive in life...talking with parents about our problems.
Therapy
Talk to people. Socialize in the real world they can help you better than the ones siting behind the screen
Discussing it with friends and family and think of positive aspects of life rather than overthinking about what had happened in past.
Talk to near ones. Don't just think they won't understand. They would. If they won't, seek external help. There's no shame in that.
Consult pshycologist
Getting help and reaching out. Keeping oneself occupied.
Exercise

Do yoga
Find some work you love to do
meet new people and make friends and try meeting them daily.
Look at positive quotes
Just go to sleep
Positive attitude and optimism
Assume responsibility. Build a timetable. Follow a repetitive schedule.
Make good friends
Just do it
Talking helps a lot. Try to not be hard on yourself. It's okay to fail sometimes and it's okay to choose an easy path rather than a complex way.
No
Meditation and counselling

Leave social media
Focus on work rather than results

Can murders be decreased if people become more disciplined and use social media in a much efficient way?
34 responses

- Strongly disagree — 11.8%
- Disagree — 14.7%
- Neutral — 26.5%
- Agree — 41.2%
- Strongly agree

Can social issues decrease if people start being co-operative in nature and rather than just reacting to the situations?
34 responses

- Yes — 67.6%
- No
- Maybe — 29.4%
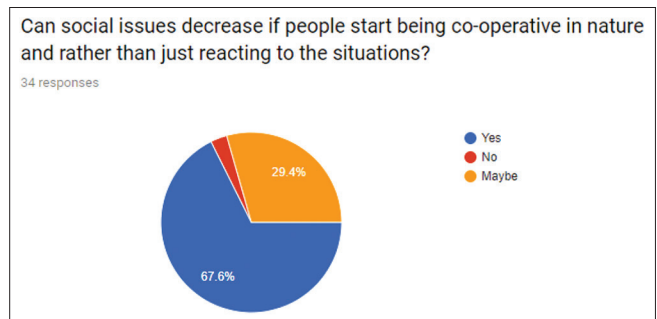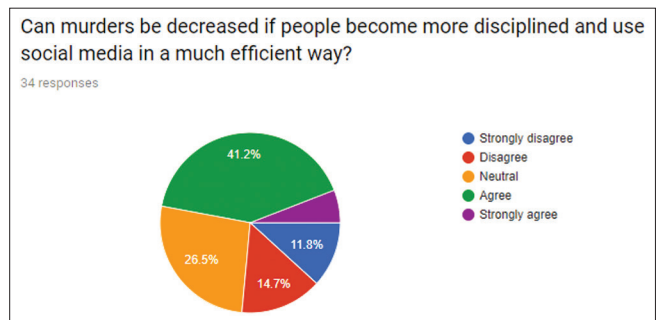
Table 5: Clustering techniques for dataset visualization

Unsupervised learning: Clustering	
Techniques	Description
K-means	Simple clustering algorithm
Hierarchical clustering	The more complex clustering algorithm

by supervised and unsupervised learning, we see the death rate is increasing. The neutrality is more toward the suicide and depression so from Google Survey we could see had the people generally see, the feeling of suicide. Thus from Google Form to

static machine learning prediction to dynamic sentimental analysis, the end results show that we should be humble with the people around us and we should find ways of lessing the social media impact on our generation, thus reducing the use of social media and reducing the impact of social issues.Thus through this article, we get known clearly how the current world is getting into social issues like victim of murders or road accidents or

Give some suggestions, how we, the citizens of India can decrease the social issue to some extent if not fully.

34 responses

By creating awareness among people

With the feeling of togetherness and forgiveness everything is possible

Impersonal problems, for example, shortage of money;
Personal problems, for example, emotional or health problems;
Interpersonal problems, such as disagreements with other people; and.
Community and wider societal problems, such as litter or crime rate.

We shouldn't be addicted to that

We must not support he wrong things that happen in front of us..we must fight back.

Monitored content

Preaching values with logic and Boosting morale. Nowadays due to the competitive world citizens tend to forget basic values and morales

Being cooperative in nature rather than just blindly reacting to situation which can adhere negative image on society and on self.

Have a sense of ownership towards the motherland and everything in it.

Be vigilant

Collapsing the barriers between genders and treating everyone equally. Also by demolishing racial and religious prejudice can play an important role.

Don't believe things blindly

Try therapy

Communicate with people more.

Spreading awareness using social media

Study properly

We can become more tolerant towards hearing everyone's side of the story rather than simply thinking that we are right all the time.

education

Just talk to people.

By eating and sleeping together

Self defense and awareness among people

People should try to be articulate in their speech and support them with plausible explanations whenever required. Misinformed actions can be avoided.

Participating in activities is needed at every level(example society, school). People should also be touch in real life rather than in virtual

Instead of sharing start reacting

Why should I care

Be friendly and helpful to your society

Create Equality Among Citizens

Show compassion

Be the change you want to see. Small steps like avoiding cyber bullying and teasing etc. And maybe using a little less of social media and living a real life will help you a lot

Don't believe in every news that you see in social media most of them are fake

---

No

Counselling

Always choose political party that follows right wing

By involving ourselves more into outside activities rather being oN social media

---

committing suicide. The article clearly helps us to understand the sensitivity of the youths, therefore brings a social awareness among the current generation.

Financial support and sponsorship
Nil

Conflicts of interest
There are no conflicts of interest.

References

1. Chong M, Abraham A, Paprzycki M. Traffic Accident Data Mining Using Machine Learning Paradigms. Informatica (Slovenia), 2005. p. 89-98.
2. Vasavi S. Extracting hidden patterns within road accident data using machine learning techniques. In: Information and Communication Technology. Advances in Intelligent Systems and Computing, Springer, Singapore, Springer; 2018. p. 13-22.
3. Clarke DD, Forsyth R, and Wrights R. Machine learning in road accident research: Desicion trees describing road accidents during croos-flow turns. Processes and countermeasures in overtaking road accidents. Ergonomics, 1999, 42:6, p. 846-867.
4. Nofal EM. Utilization of Social Media for Raising Required Awareness of Human Factors Affecting Traffic Safety; National Partnership and Corporate Social Responsibility"; 2013, Proceedings of The Second Forum "Traffic Safety: National Partnership and Corporate Social Responsibility"Dammam, Kingdom of Saudi Arabia, p. 204 – 215.
5. Calderon-Vilca HD, Wun-Rafael WL, Loarte MR., Machine Learning and Metaheuristics Algorithms, and Applications, First Symposium, SoMMA 2019, Trivandrum, India, December 18–21, 2019.
6. Birjalia M, Hssanea AB, Erritalib M. Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks 2017. Procedia Computer Science, 113, 2, 2017, p. 65-72.
7. Glen C, Leary R, Crutchley P, Fine A. Natural Language Processing of Social Media as Screening for Suicide Risk; Biomedical Informatics Insights, Proceedings from the Digital Mental Health Conference-London 2018.
8. Colomboa GB, Burnap P, Hodorog A, Scourfield J. Analyzing the connectivity and communication of suicidal users on twitter. Computer Communications, 73,1,2015, p. 291-300.
9. McClendon L, Meghanathan N. Using machine learning algorithms to analyze crime data. Machine Learning and Applications: An International Journal (MLAIJ) 2,1,2015, p. 11-21.
10. Prabakaran S, Mitra S. Survey of analysis of crime detection techniques using data mining and machine learning. Journal of Physics: Conference Series, 2018, DOI :10.1088/1742-6596/1000 /1/012046.
11. Paul Elhorst J. Research on the Raw Data Processing Method of the Hydropower Construction Paper; IOP Conference Series: Earth and Environmental Science, 108,5, 2017, p1-5.
12. Vijayarani S, Sharmila S. Research in big data – An overview. Journal of Xi'an University of Architecture & Technology, 2016, p. 1604-1610.