

Crime Against Women and Girls in Indian states: Analysis, Visualization and Prediction

Seema Aggarwal¹, Geeta Aggarwal^{2*} and Manisha Bansal³

ABSTRACT

Crime against women and girls (CAWG) is an issue of utmost concern affecting women from all levels of society. It is the biggest challenge to government and law enforcing agencies. One quick measure to help today's women could be to understand the existing trends and get insights from reported crimes against women. This will help in not only making the society a safer place, preventing repeat of similar incidents in the future, but also, will help victims fight back and bring the propagators of crime to justice. In this paper we attempt to understand the existing trends of violence against women and girls in Indian states using real life data from the National Commission for Women (NCW) for the last twenty-two years. In addition to the overall attempt at improving women's safety through analysis and recommendations, we also aim to find out nuances using advanced statistical tools to recognize variations in crime against women across various Indian states. The insights from recently reported crimes may be used to frame recommendations to narrow down the gaps that lie in the government's social and safety infrastructure and plug them in to prevent crime against women ameliorating the safety of women in our country. We identify the best states to live in for women and identify "hot states" and "hot crimes" in India. The paper also predicts the crime rate of different crimes in hot states for the next five years.

Keywords: NCW, Linear Regression, Statistical Analysis, Crime data analysis, CAWG.

1. Introduction

Crime against women and girls (CAWG) is an important issue not only in India but all over the world. It is affecting women in all walks of life irrespective of their religion, caste, or socioeconomic status. It is the biggest challenge to gender equality and women empowerment. As per recent UN research an estimated 700 million women have been subjected to partner/non-partner sexual violence at least once in their lives. The situation is no different in India, a country already infamous for sexual exploitation and violence against women. A total of 14220 cases have been reported in India for crime against women in the year 2019 and 16743 have been reported in the year 2020. What is ironic is that a country where women are considered as deities and worshipped, is the place where incidents of violence against women and girls have been persistently increasing bringing shame and a need for developing evidence-based policies and their reinforcement.

In this paper we attempt to understand the existing trends among violence against women and girls in Indian states using real life data. We analyse the official data on crimes

against women in India, from the National Commission for Women (NCW) which summarises the complaints lodged by women in police stations across the country in the last twenty-two years. This analysis focuses on trends and patterns of crimes against women in the last two decades. In addition to the overall attempt at improving women's safety through analysis and recommendations, we also aim to find out nuances using advanced statistical tools and correlation to recognize variations in crime against women across various Indian states. The insights from recently reported crimes may be used to frame our recommendations to predict where the gaps lie in the government's social and safety infrastructure and plug them in and help prevent crime against women ameliorating the safety of women in our country. We identify "hot states" and "hot crimes" in India. We find the best and the worst states to live in India for women. The number of different crimes in identified hot states have been predicted for the next five years.

2. Related Work

Researchers have conducted many studies to understand, analyse crime to develop methods for its prevention. We

1. Miranda House, University of Delhi

2. PGDAV College, University of Delhi

3. Indraprastha College for Women, University of Delhi

* Corresponding Author ✉ geeta.gupta@pgdav.du.ac.in

provide a brief overview of the available research on analyses of both general crime and crime against women. Machine learning techniques have been used to analyse crime data and forecast future trends. Crime is an offence which is punishable by law. Reasons for these crimes are different and varied. A subset of these reasons has been categorised by the Indian Penal Code (IPC). This categorization classifies certain acts as crimes against women which are our prime area of focus. Kher in (Zadgaonkar, 2019) uses data mining algorithms for crime data analysis. Agarwal in (J. Agarwal, 2013) defines "Crime analysis" as an analytical process which provides relevant information relative to crime patterns. This analysis can be used by personnel in planning and deployment of resources for its prevention and suppression. Bengaluru and India newsfeed are used as experimental data by Prathap in (Prathap, 2023) to analyse crime type using kernel density estimation. Clustering is used on Bengaluru and Karnataka datasets by Patil in (Anandhi, 2020) to determine similar and different crime characteristics. Kapoor in (Singh, 2022) uses the dataset of Chicago to predict the place and time of an occurrence of an incident of robbery. They further draw the pattern of race of criminals using visualisation. Vijayalakshmi in (Bandekar, 2020) uses machine learning algorithms to help in reducing crime rate. Gupta in (M. Gupta, 2008) analyses the crime databases maintained by NCRB (National Crime Record Bureau). They use data mining techniques on this official crime data to find crime hot spots. Zubi in (Z. S. Zubi and A. A. Mahummud, 2013) uses data mining techniques to analyse data collected by police departments in Libya. They help in discovering trends and patterns for increasing crime in Libya. Kumar in (Nagpal, 2019) also lays emphasis on finding trends in the reported crimes in order to mitigate crime.

The above work analyses data regarding general crime. However, there are articles (K. Rajeshwari) (P. Prasad, 2021) (Rajeshwari, 2018) (Sharma, 2023) on analysis of data specific to crime against women. They use Huber regression to analyse the data and visualise it using a time series algorithm. They further predict the crime occurring in a particular state. Rajeshwari in (K. Rajeshwari) did the statistical analysis for the crime data for the year 2013. Reddy in (M. Reddy, 2020) uses statistical models like Autoregressive Integrated Moving Average (ARIMA) and Linear regression model to predict crime rate in India on a dataset containing data from 2001 to 2013. A description of the algorithms (including their advantages and disadvantages) that have been used for analysis and prediction of crime against women in India and other nations like Bangladesh is available by Surve in (Upadhyay, 2022). Sharma in (Sharma, 2023) studies the data available from the NCRB from 2001 to 2020. Kumar (Nagpal, 2019) in the paper uses Naïve Bayes Classifier to predict the most likely criminal of a particular crime incident.

We focus on crime against women and girls and that too in the context of Indian women and states for the data of the last twenty-two years. In this paper we attempt to understand the existing trends among crime against women and girls in Indian states using real life data. We analyse the complaints lodged by women in police stations across the country in the last twenty-two years for crimes they faced. In addition to the overall attempt at improving women's safety through analysis and recommendations, we also aim to find out nuances using advanced statistical tools and correlation to recognize variations in crime against women across various Indian states. We then frame our recommendations to help prevent crime against women ameliorating the safety of women in our country. The benefits of identifying these patterns will be two-fold and could help fight a double-edged sword. Not only will they make society and surroundings safer, preventing repeat of similar incidents in the future, but also prepare the government and various NGOs in strengthening the machinery to help victims fight back and bring the propagators of crime to justice. Several sensitization programmes can be taken for the law enforcing agencies.

In this paper data mining techniques have been used to analyse crime against women using real life data. We try to understand the trends among crime against women and using them we frame our recommendation to prevent the same and ameliorate the safety of women in our country. In addition to the overall attempt at improving women's safety we also aim to find out nuances in data using advanced statistical tools to recognize variations in crime against women across various Indian states. The identified patterns could help fight a double-edged war - not only will they make society and surroundings safer, preventing repeat of similar incidents in the future, but also prepare the government and various NGOs in strengthening the machinery to help victims fight back and bring the propagators of crime to justice. Finally, we predict the number of complaints in the next five years. These may help the authorities to devise an action plan to act upon.

3. Methodology

In this paper, we do an exhaustive analysis of the real-life dataset containing information about the crimes against women in India for the last twenty-two years i.e. from 2001 to 2022. The dataset is constructed from the actual number of complaints registered at the National Commission of Women, a government of India enterprise, for acts of violence done against women and girls.

NCW maintains a record of the number of complaints registered at the commission for sixteen different acts of crime done against women in each state and union territory of India. We collated this data for our study. The data thus collated contained the total number of complaints registered for a set of sixteen crimes in each of

the states and union territories of India totalling 36. This was collected for the last twenty-two years. The data was pre-processed and suitably formatted to fulfil our requirement.

Analysis of data was done in two stages. In the first stage, the data was inferred with descriptive statistical analysis, correlation tests and visualisation techniques. We find the best and the worst states to live in India for women along with the most prevalent crimes against women in India. In the second stage we used Linear Regression to study the trend and pattern of crimes in these states. To test the statistical significance of our results, we further used Mann-Kendall statistical test. Finally, we use the linear regression model to predict the number of crimes the authorities may expect in the next five years.

3.1 Data Collection Phase

The dataset that we used was extracted from the official records of the Indian government. We collected the data from the website of the NCW (National commission for women: www.ncw.nic.in). The website has two types of listing available namely State wise and Nature wise. The state wise option lists the number of cases registered for a set of sixteen crimes like: 1. Bigamy/Polygamy 2. Divorce 3. Dowry Death 4. Dowry Harassment /Cruelty to married women 5. Harassment at Workplace 6. Kidnapping /Abduction 7. Maintenance Claim 8. Miscellaneous 9. Murder 10. Outraging Modesty of Women 11. Police Apathy against women 12. Rape 13. Right to live with dignity 14. Sexual harassment including sexual harassment at workplace 15. Shelter & Rehabilitation of Victims 16. Women’s right of custody of children in the event of divorce in every year for each state and union territory.

The Nature wise option lists the total number of complaints registered at the commission for each crime against women every year. There is also an option for listing the data nature wise vs state wise. This format lists the number of cases registered in a state for a particular type of crime. The data was collected for all the years from 2001 to 2022 and then the preprocessing was done on the dataset constructed.

3.2 Data Preprocessing Phase

To draw meaningful inferences from the data we did a preliminary investigation of the data and converted it into a format suitable for our study. Preprocessing began with scans of the dataset followed by changing the data types of a few columns and removing null entries in the data.

Initially we had a set of sixteen crimes out of which, we removed the complaints for miscellaneous crime category as it did not add any value addition to our analysis. We also grouped together a few categories of crime into one category based on their nature. For instance, we summed the complaints received against Divorce, Custody of children after Divorce and maintenance into one single category. Also, all complaints against dowry death and harassment were summed up together. Further, complaints against harassment (Sexual harassment, harassment at workplace and modesty) were summed up together into one category. Lastly, complaints against Rape and rehabilitation and shelter were clubbed into one category. Thus, based on the nature of crime the different categories of crime were now reduced to a total of nine categories.

Finally, a total of three datasets were constructed for our study. Dataset I was constructed to study the pattern of different crimes against women occurring in various states (i.e., crime vs state) in the last twenty-two years. This consisted of a table of 9 rows, one row for each of the final nine categories of crime and had 36 columns, one for all the states and union territories in India, as shown in Table 1 below. The dataset had twenty-two such tables, one for each year from 2001 to 2022. The total of all crimes against women registered in a state for a particular year can easily be calculated from this table by summing across a column.

Dataset II was a time series data constructed to study the total crime in the last twenty-two years in Indian states/union territories (year vs state). This consisted of table of 36 rows, one row for each state/union territory and 22 columns for the last twenty-two years as shown in Table 2.

Dataset III was again a time series data constructed to study the pattern of a particular crime in the last twenty-two

Table 1 : Dataset I (Crime vs State) for the year 2022

Crimes / States	AP	AR	AS	BR	CG	GA	GJ	HR	WB	A&N
Bigamy / Polygamy	1	0	2	21	2	0	1	3	1	0
Divorce + Maintenance Claim + Women's right	0	0	0	0	0	0	0	0	0	0
Dowry Death + Dowry Harassment	16	0	8	137	11	1	14	130	23	1
Harassment at workplace + Sexual harassment	23	1	15	124	46	4	16	158	69	0
Rape + Shelter & Rehabilitation of Victims	3	0	3	33	7	0	5	90	19	2
Kidnapping/Abduction	0	0	0	0	0	0	0	0	0	0
Murder	0	0	0	0	0	0	0	0	0	0
Police Apathy against women	9	0	3	137	11	1	6	88	29	0
Right to live with dignity	65	0	30	424	49	4	83	327	221	0
TOTAL	117	1	61	876	126	10	125	796	362	3

Table 2 : Dataset II (State vs year)

States / Years	2001	2002	2003	2004	2005	2006	2018	2019	2020	2021	2022	Total
AP	22	23	16	8	49	37	102	86	107	156	117	1396
AR	0	2	1	0	1	4	2	1	4	2	1	32
AS	7	14	9	5	19	15	33	35	43	42	61	577
BR	190	309	350	316	507	289	579	509	681	989	876	9902
CG	9	39	26	27	67	53	74	73	111	135	126	1388
GA	1	2	3	0	4	3	11	9	10	88	10	197
GJ	7	10	11	18	48	37	91	80	102	68	125	1421
HR	158	257	228	168	488	469	994	812	942	997	796	15028
WB	31	20	34	15	64	59	255	185	256	455	362	3358
A&N	3	1	1	2	1	0	5	3	3	1	3	53
CH	6	6	6	8	16	24	38	29	39	39	33	552
D&N	0	2	0	1	2	0	2	3	6	1	6	41
D&D	0	0	0	1	0	3	0	0	0	22	0	46
LK	0	0	0	0	0	0	0	0	1	3	0	8
DL	587	901	539	550	1283	1254	1448	1111	1319	979	1572	31145
PC	3	3	0	1	4	4	12	8	7	8	15	161
Total	3311	5630	4547	4868	9495	8090	15873	14220	16743	21244	21904	272457

Table 3 : Dataset III (Year vs Crime) for Delhi

Crime	Bigamy / Polygamy	Divorce + Maintenance Claim + Women's right of custody of children	dowry death+harass	Harassment in all	Kidnapping / Abduction	Murder	Police Apathy against women	Rape+rehabilitation	Right to live with dignity
2001	9	19	162	221	12	16	41	16	91
2002	16	16	253	372	14	13	80	21	116
2003	7	4	139	252	14	7	76	13	27
2004	7	2	153	235	8	11	95	19	20
2016	16	9	504	211	0	0	362	0	828
2017	16	9	193	149	0	0	146	0	734
2018	16	12	222	190	0	0	147	0	861
2019	14	0	246	199	0	0	132	61	459
2020	8	0	269	205	0	0	142	60	635
2021	14	3	358	211	0	0	121	92	180
2022	17	0	405	312	0	0	150	118	570

Table 4: Abbreviations used for crime category

Category	Crime
C1	Bigamy / Polygamy
C2	Divorce + Maintenance Claim + Women's right of custody of children
C3	Dowry Death + Dowry Harassment
C4	Harassment at workplace + Sexual harassment including at workplace + outraging modesty of women
C5	Rape + Shelter & Rehabilitation of Victims
C6	Kidnapping / Abduction
C7	Murder
C8	Police Apathy against women
C9	Right to live with dignity

years in a particular state. This consisted of a table of 9 columns, one for each category of crime and 22 rows for the last twenty-two years as shown in Table 3. This Dataset had four such tables, one for each of the four hot states.

The lists of the abbreviations used for crime categories is given in Table 4 and the abbreviations used for states and union territories in alphabetical order are listed in Table 5.

3.3 Processing Phase

Descriptive statistical analysis was carried on Dataset I and II using Microsoft Excel. We identified the best and the worst state/union territory in India for women. Also, graphs were created to understand the data better. Python code was developed to carry out linear regression modelling on Dataset III. Mann-Kendall test was carried out on Dataset III to further identify the trend of crime in the hot states. Statistical significance of results was also checked.

Time series data set III essentially shows variation of a dependent variable (occurrence of crime in our case) with respect to a unit of time (year). Time series data may exhibit trends over time. We used python code to identify a component of trend in it. Our main objective was to identify the existence of a linear trend in a category of crime occurrence in a particular state in the last twenty-two years. The Dataset III was constructed for this purpose only. We used linear regression to identify trends by fitting a line to the data. We then study the slope of the line. A positive value of the slope indicates an upward trend and a negative indicates a downward trend. Absolute value of the slope indicates the strength of the trend, higher indicates stronger trend.

4. Results and Discussion

4.1 Comparison of crime against women in each state thus finding the best and worst Indian states to live in for women

- About 56 per cent (151874) of the total complaints received (272457) were from the most populous state

of Uttar Pradesh. In fact, Uttar Pradesh is the only state having the registered number of crimes as a six-digit number.

- After UP, there are four states having the number of registered crimes against women between ten thousand and forty thousand. Delhi recorded 31145 complaints, followed by Haryana (15028), Rajasthan (14863) and Madhya Pradesh (10042) as shown in the graph below. The table 6 lists the different categories of crime reported in the four best and worst Indian states. Figure 1 shows the total registered crime in the Indian states.
- States like Goa, Puducherry, Jammu and Kashmir, Chandigarh, Assam, Himachal Pradesh, and Kerala have registered less than one thousand crimes against women.
- On the other hand, Orissa, Telangana, Chhattisgarh, Andhra Pradesh, Gujarat, Karnataka, Tamil Nadu, West Bengal, Jharkhand, Punjab, Uttarakhand, Maharashtra, and Bihar have registered crimes between thousand and ten thousand.
- There are eleven states and union territories like (Mizoram, Nagaland, Sikkim, Manipur, Daman & Diu, Dadar & Nagar Haveli, Arunachal Pradesh, Tripura, Meghalaya, and Andaman & Nicobar Islands) having the number of registered crimes against women less than hundred. Out of these Lakshadweep has registered only 8 crimes against women.
- Kerala with the highest literacy rate has registered only 890 crimes against women in the last twenty-two years.

4.2 Most prevalent and least prevalent crime faced by women in India.

- It is observed that in the last twenty-two years the commission received the maximum number (86430)

Table 5 : Abbreviations used for States and Union Territories.

A&N	Andaman & Nicobar Islands	DL	Delhi	LK	Lakshadweep	PB	Punjab
AP	Andhra Pradesh	GA	Goa	MP	Madhya Pradesh	RJ	Rajasthan
AR	Arunachal Pradesh	GJ	Gujarat	MH	Maharashtra	SK	Sikkim
AS	Assam	HR	Haryana	MN	Manipur	TN	Tamil Nadu
BR	Bihar	HP	Himachal Pradesh	MG	Meghalaya	TL	Telangana
CH	Chandigarh	J&K	Jammu & Kashmir	MZ	Mizoram	TR	Tripura
CG	Chhattisgarh	JH	Jharkhand	NL	Nagaland	UP	Uttar Pradesh
D&N	Dadra & Nagar Haveli	KR	Karnataka	OR	Orissa	UK	Uttarakhand
D&D	Daman and Diu	KE	Kerala	PC	Puducherry	WB	West Bengal

Table 6 : Crime distribution for the worst and the best four states.

States/ Crimes	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
UP	1521	433	35531	22490	8456	2185	805	33164	47289	151874
DL	365	270	6257	6191	978	376	141	5482	11085	31145
HR	122	79	2772	2515	918	203	101	3925	4393	15028
RJ	251	106	3020	2517	1237	212	156	4170	3194	14863
SK	0	1	4	10	1	0	0	1	7	24
MZ	0	1	2	5	1	0	0	4	3	16
NL	0	0	0	5	7	0	0	2	2	16
LK	0	0	1	2	0	0	0	1	4	8

of complaints for the crime “Right to live with dignity”.

- Large number of complaints have also been received for Crimes listed in Police apathy against women (56526) and Dowry harassment (59519), sexual harassment (45865).
- Least number of complaints (less than 500 in the last twenty-two years) have been received for murder of women, Divorce, Shelter and rehabilitation of Victims and Women’s right of custody of children in the event of Divorce. The table 7 and graph in figure 2 show the results.

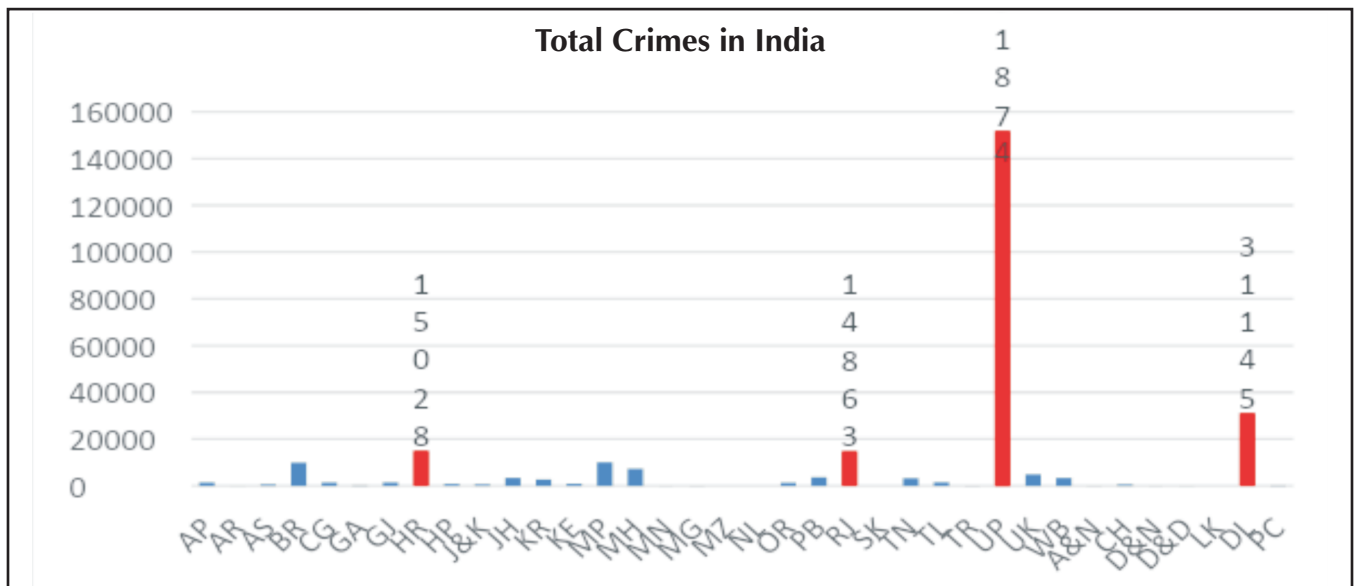


Figure1 : Crime distribution in States of India

Table 7 : Most and least prevalent Crime against women in India.

Crimes / States	AP	AR	AS	BR	CG	GA	GJ	HR	WB	A&N	CH	D&N	D&D	LK	DL	PC	Total	
Bigamy / Polygamy	29	3	14	221	26	3	27	122	48	2	3	1	0	0	0	365	2	3367
Divorce + Maintenance Claim +	22	1	9	70	9	2	25	79	46	1	13	0	0	0	0	270	1	1461
Dowry Death + Dowry Harassm	333	2	99	2386	271	7	274	2772	413	4	84	5	4	1	6257	22	59519	
Harassment at workplace + Sex	285	14	165	1573	311	44	367	2515	706	26	169	8	6	2	6191	52	45865	
Rape + Shelter & Rehabilitation	42	3	31	481	80	6	72	918	119	2	22	4	0	0	978	5	14017	
Kidnapping / Abduction	10	0	8	91	11	0	13	203	27	0	2	0	0	0	376	2	3470	
Murder	16	0	9	147	19	2	13	101	22	0	3	1	0	0	141	1	1802	
Police Apathy against women	170	3	50	2060	241	12	184	3925	395	5	95	7	9	1	5482	30	56526	
Right to live with dignity	489	6	192	2873	420	121	446	4393	1582	13	161	15	27	4	11085	46	86430	
TOTAL	1396	32	577	9902	1388	197	1421	15028	3358	53	552	41	46	8	31145	161	272457	

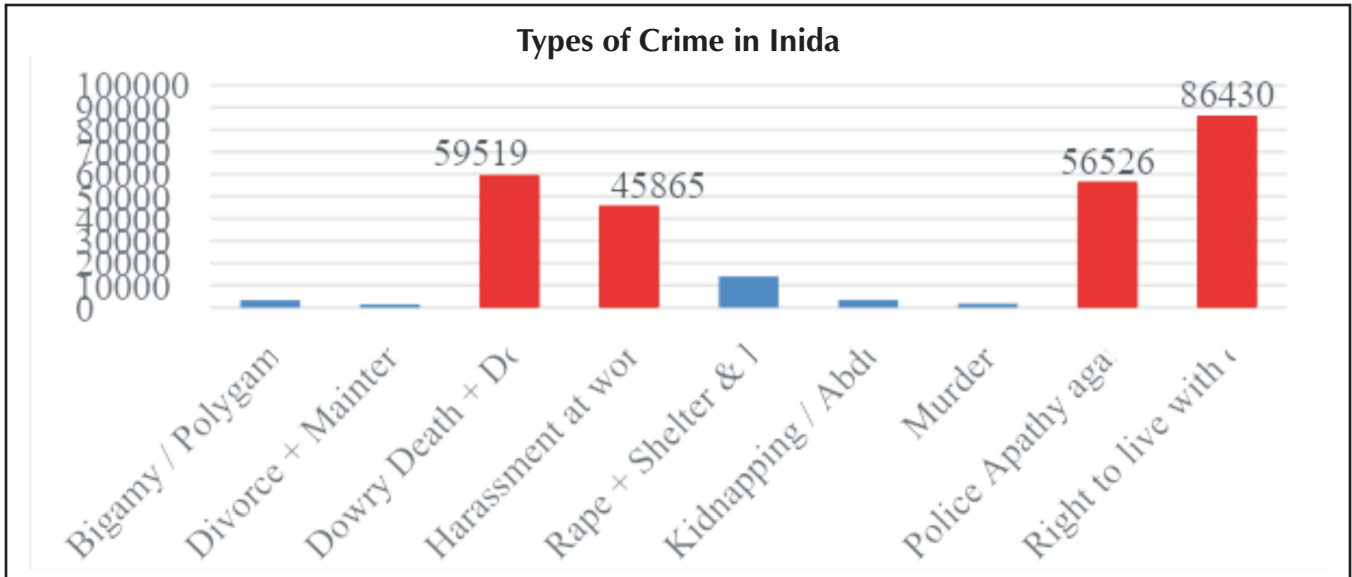


Figure 2 : Most prevalent CAW in India

4.3 Trend analysis of most prevalent crime in four hot states

The hot states in India have been identified as Uttar Pradesh, Delhi, Haryana, and Rajasthan. The crime distribution in these four hot states is shown above in figure 3. As is evident from the figure most prevalent crimes in these four hot states also the most prevalent/hot crimes are again death due to dowry and harassment for

dowry, sexual harassment and harassment at work place, Police apathy against women, and lastly right to live with dignity.

We further study the trend and pattern of these four crimes in these four hot states using linear regression. We created python script for linear regression modelling of the time series data. Timeseries data constructed for all the four hot states was analysed. For all the four hot states

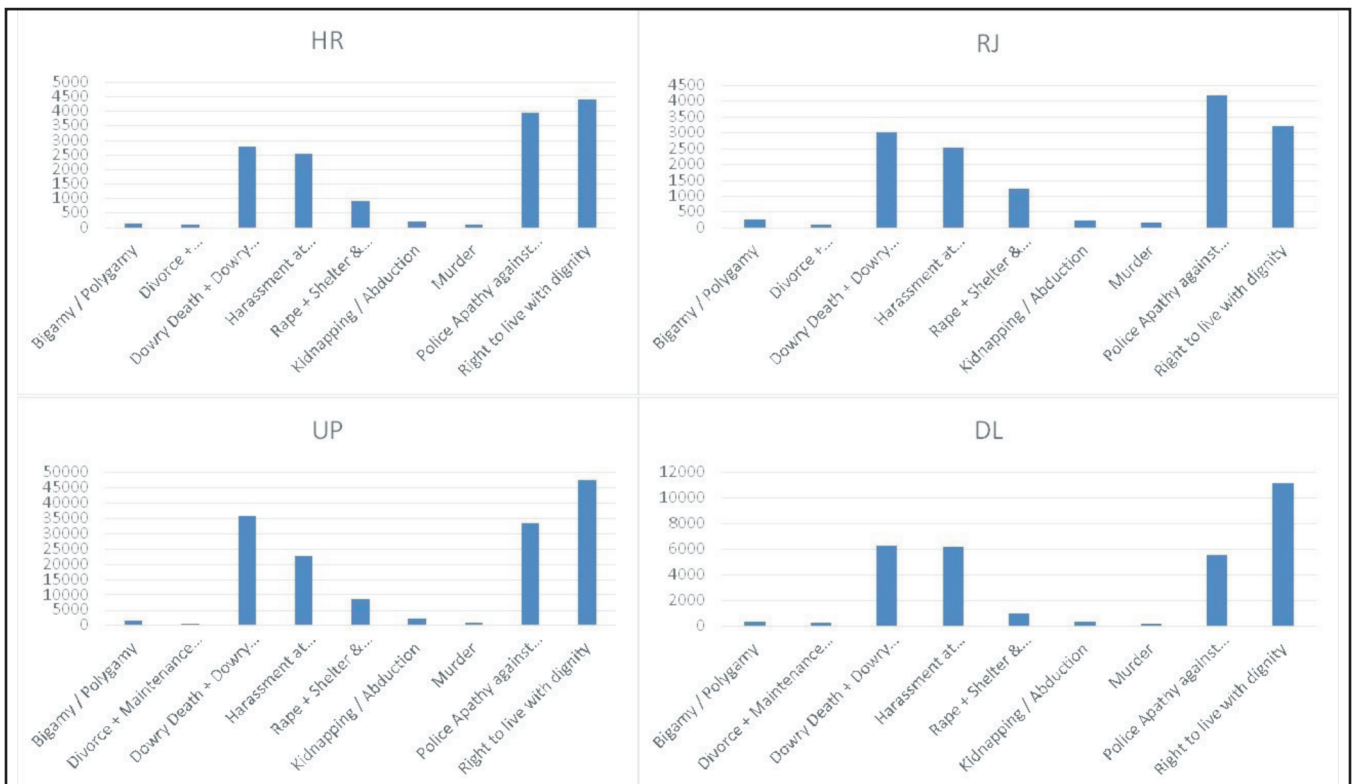


Figure 3 : Crime distribution in the four hot states

and for all the four most prevalent crimes we applied linear regression modelling followed by Mann Kendall test.

We found that the straight line derived after applying the linear regression model had positive slope indicating an increasing trend. Also, the Mann Kendall test when applied on the data reported increasing trend with a p value less than .05. Thus, the positive trend identified by the linear regression model is endorsed by the Mann Kendall test as well This was true for all the four hot crimes analysed for all the four hot states. There was an exception to the positive trend in Delhi and Rajasthan for harassment of women. Delhi and Rajasthan had a negative slope for sexual harassment indicating that the number of complaints received against sexual harassment in both Delhi and Rajasthan have decreased over the years. But the Mann Kendall test indicated no trend and the p value was also greater than 0.05. Therefore, these two trends are ignored.

The trend can be clearly seen in the figures 4-7.

4.4 Prediction of most prevalent crimes in four hot states in the next five years

Linear regression modelling was used to not only study the patterns but also to predict the crime rate in the next five years in these hot states. The numbers as shown in the table 9 are quite high and endorse the increasing trend of crime

rate. The authorities should take clue from the predicted data and act accordingly.

5. Conclusion and Future work

An exhaustive in-depth analysis of the constructed datasets has been carried out using various analytical tools under various categories. Prediction of the crime for the next five years has also been done. The numbers again are quite high and eye opener. Data visualization has also given us a meaningful insight into the data. Results obtained are statistically significant having p-value less than 0.05. Despite getting positive results there is a lot of future scope to improve on the results in the study. Firstly, the study is based on real life data from NCW. This may reflect a lesser picture due to underreporting. Women in India may not be reporting because of lack of awareness and pressure from the family. Socio economic factors and the literacy level of the women will be incorporated in future to bring more relevance to the results obtained.

Our study and analysis clearly show that no matter how advanced the society is, crime against women has been increasing especially for India. The highest number of CAW being reported form Uttar Pradesh followed by Delhi, Rajasthan and Haryana which are the worst states in India for the women to live. On the other than, north eastern states like Nagaland, Mizoram, Arunachal Pradesh, Meghalaya, Tripura have very few incidents of

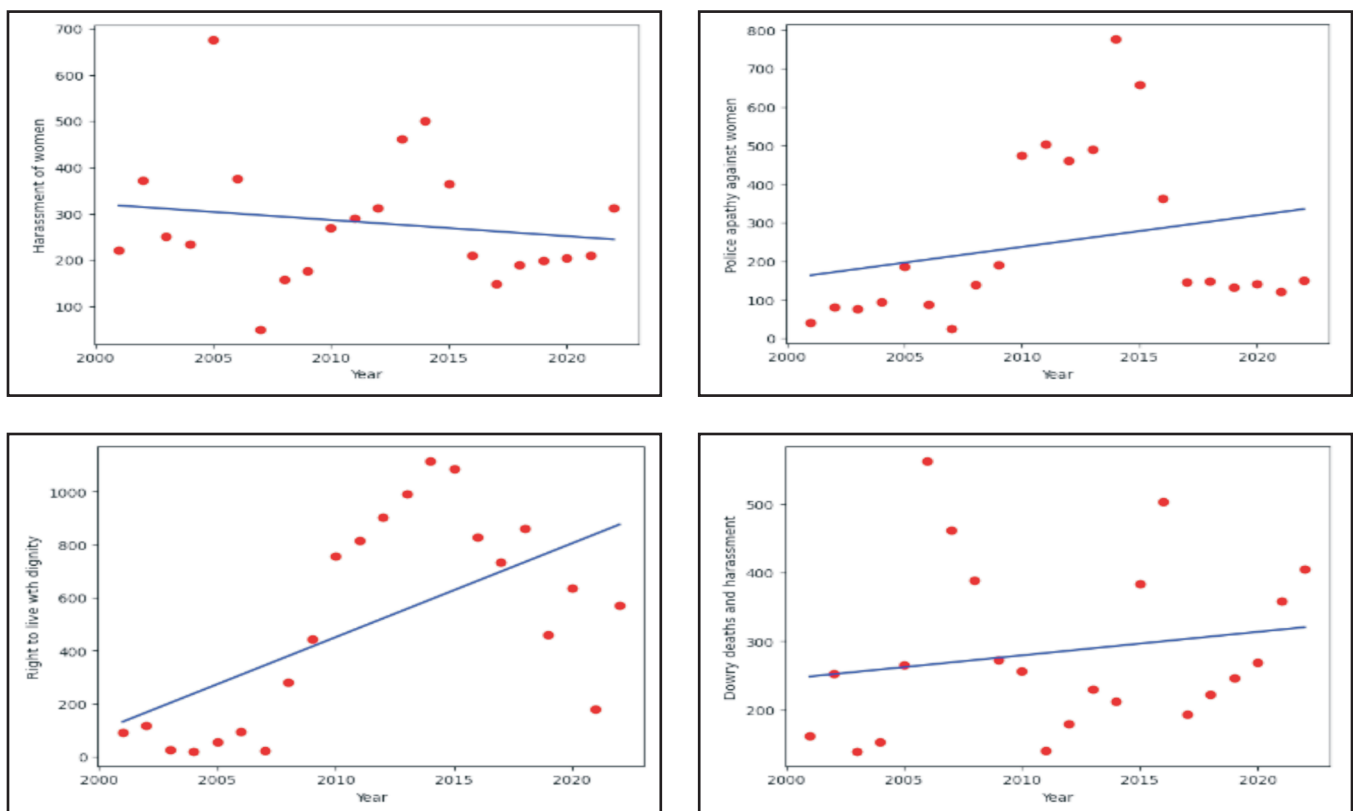


Figure 4 : Trend of Worst crimes in Delhi

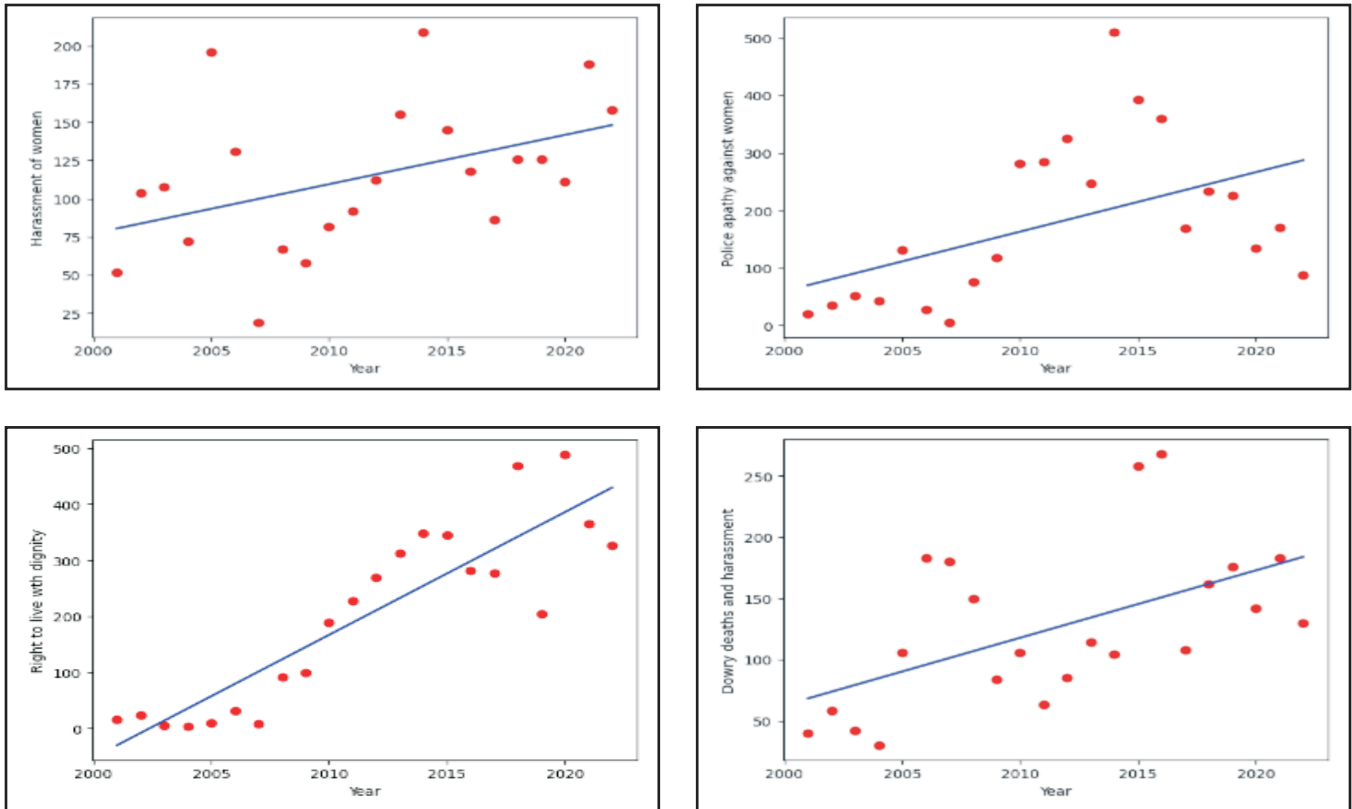


Figure 5 : Trend of Worst crimes in Haryana

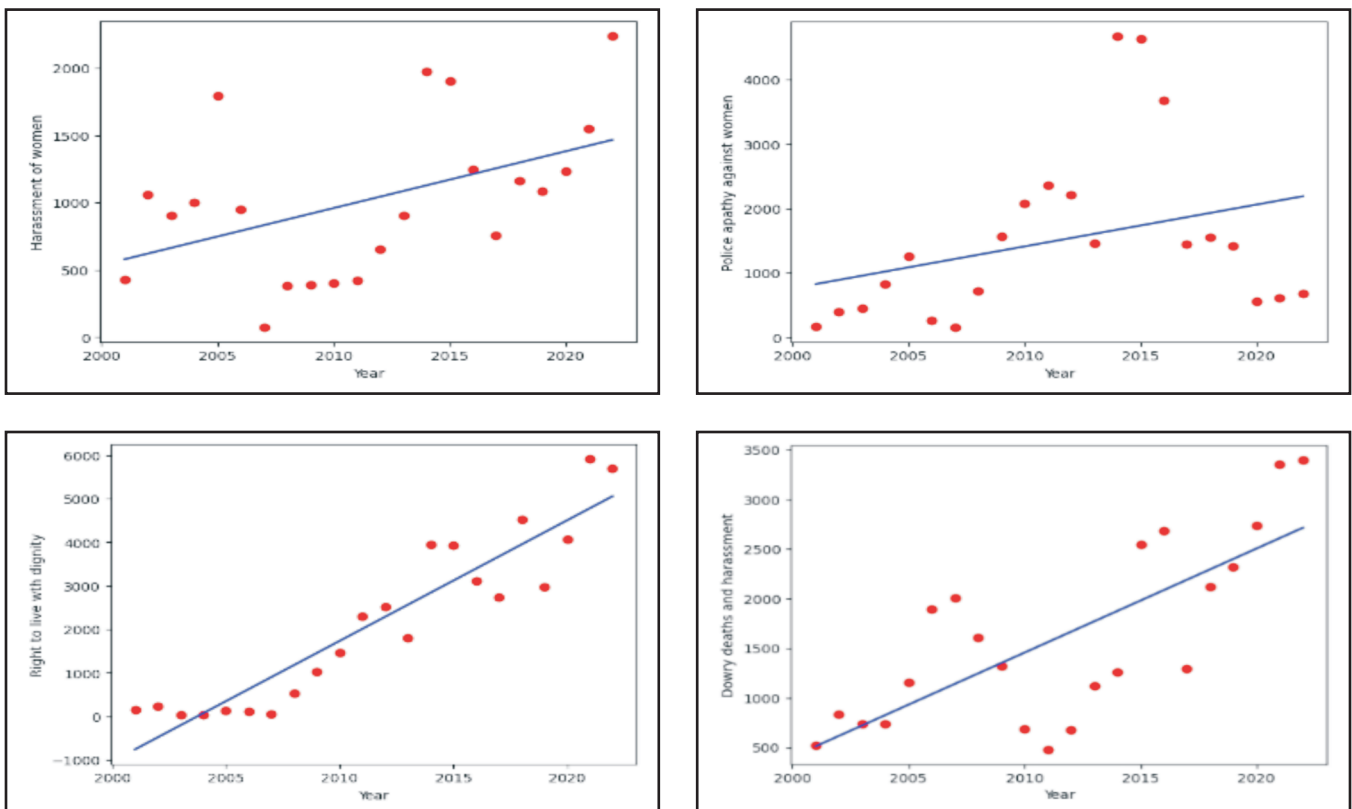


Figure 6 : Trend of Worst crimes in Uttar Pradesh

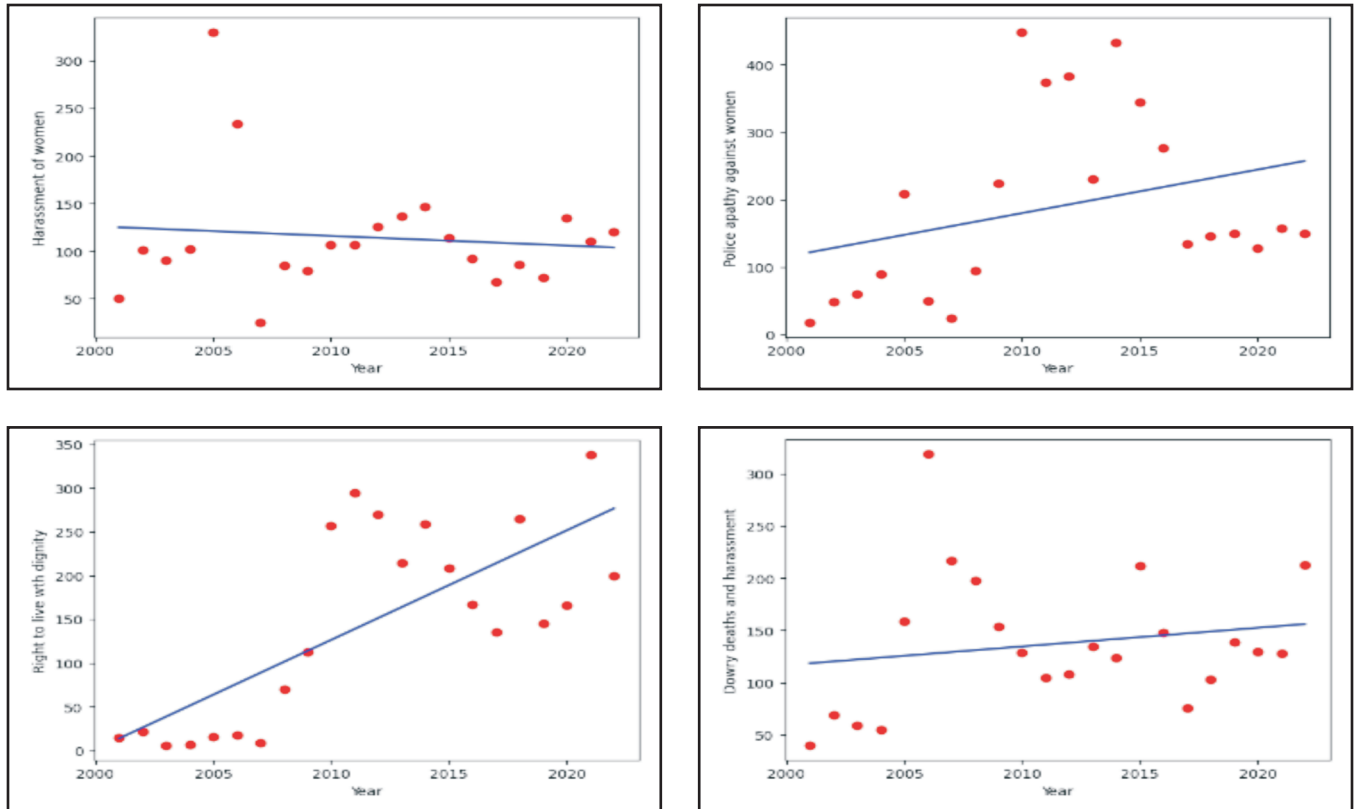


Figure 7 : Trend of Worst crimes in Rajasthan

Table 9 : Prediction of the hot crimes in the hot states for the next five years.

Year	Dowry Death and harassment				Harassment in all				Police apathy against women				Right to live with dignity			
	UP	HR	DL	RJ	UP	HR	DL	RJ	UP	HR	DL	RJ	UP	HR	DL	RJ
2024	2926	194	327	159	1550	154	237	101	2318	307	351	270	5611	473	946	301
2025	3031	200	330	161	1592	157	234	100	2383	318	359	276	5888	495	981	314
2026	3136	205	334	163	1634	161	230	99	2448	328	367	282	6165	517	1017	326
2027	3241	211	337	164	1677	164	227	98	2512	338	376	289	6441	539	1052	339
2028	3346	216	340	166	1719	167	223	97	2577	349	384	295	6718	561	1085	351

such CAW being reported thus can be few best states for the women to live in India. It is recommended that state level appropriate measures be taken to stop the rise in crime against women. The NGO and law enforcement authorities should take policing and corrective actions in the states with a high number of CAW. Right to education at school level increases gender sensitization, strict punishment should be enforced by authorities to bring down the numbers and make India a safer and better place.

References

1. Anandhi, S. P. (2020). Diversity based self-adaptive clusters using PSO clustering for crime data. *International journal information technology*, 319-327.
2. Bandekar, C. V. (2020). Design and Analysis of Machine Learning Algorithms for the reduction of crime rates in India. *Science Direct. Procedia Computer Science*, 172, 122-127.

3. J. Agarwal, R. N. (2013). Crime Analysis using K-Means Clustering. *International Journal of Computer Applications*, 0975-8887.
4. K. Rajeshwari, M. D. (n.d.). Statistical Analysis of Crimes Concerning Women in India using Data Mining Techniques. *Academia.edu*.
5. M. Gupta, B. C. (2008). Crime Data Mining for Indian Police Information System. *Computer Society of India*, pp. 388-397.
6. M. Reddy, K. M. (2020). Forecasting India's Crime rate. *International Journal of Engineering Research & Technology*, 130-132.
7. Nagpal, R. K. (2019). Analysis of Crime pattern using Big Data. *International Journal of Information Technology*, 799-805.
8. P. Prasad, A. N. (2021). Crime against Women analysis and Prediction. *International Journal of Engineering Research & Technology*, 426-429.
9. Prathap, B. R. (2023). Geo-spatial crime density attribution using optimized machine learning algorithms. *International journal information technology*, 1167-1178.
10. Rajeshwari, P. a. (2018). Violence against Women in Haryana Levels and correlate. *Punjab Geographer*.
11. Sharma, U. G. (2023). Analysis of criminal spatial events in india using exploratory data analysis and regression. *Computers and Electrical Engineering*.
12. Singh, P. K. (2022). Robbery pattern analysis (RPA) using the concept of multipolarity and examining the influencing factors. *International Journal of Information Technology*, 1425-1432.
13. Upadhyay, S. S. (2022). A review on Forecasting Crime against women in India using Machine learning Approaches.
14. Z. S. Zubi and A. A. Mahummud, “. D. (2013). *Recent Advances in Image Audio and Signal Processing*.
15. Zadgaonkar, A. K. (2019). Crime Data Analysis using Data Mining. *International Journal of Control and Automation*, pp. 115-120.

