(ISSN: 2582-5526) www.thevoiceofcreativeresearch.com

The Voice of Creative Research

Vol. 7 & Issue 2 (April 2025)




https://doi.org/10.53032/tvcr/2025.v7n2.28

Research Article

Cyber Crime in India: An Exploratory Analysis and Predictive Study

Ms. Anjali Prakashan
Student,
DES's NMITD, Dadar,
anjaliprakashan1809@gmail.com

Mrs. Lavina Mistry
Assistant Professor,
MCA Department, DES's NMITD, Dadar,
lavinajdhv@gmail.com

Abstract
The research paper investigates the increasing trend of cybercrimes in India due to rapid digitization. It aims to analyze historical cybercrime data (2016–2018) using machine learning techniques to identify trends and predict future crime rates. The study utilizes datasets from Kaggle, applying Power BI for data visualization and predictive analysis. Key findings indicate that states like Uttar Pradesh, Karnataka, and Maharashtra report the highest cybercrime cases, with a consistent upward trend. The research highlights the effectiveness of machine learning models in forecasting cybercrime rates and suggests integrating real-time reporting and deep learning for improved accuracy.

Keywords: Digital forensics, Cyber security threats, Law enforcement challenges, Cyber laws in India, Internet governance

1.Introduction

As Indian services are quickly becoming online, cybercrime has become a major concern. More individuals accessing online platforms have led to increased cybercriminal activity, such as financial fraud, identity theft, and ransomware. Cybercrimes victimize people and even impact national security and economic stability. Being aware of the trends and being able to

forecast future trends can enable law enforcement agencies to be proactive. This study seeks to examine previous cybercrime cases in India, determine trends and forecast future criminal rates through machine learning tools.

Research Questions:

- What are the key trends in cybercrime incidents across Indian states?
- Which states report the highest number of cybercrime cases?
- Can machine learning models accurately forecast future cybercrime trends?

Dataset Used: We use a dataset sourced from Kaggle containing cybercrime records across various Indian states from 2016 to 2018. It includes parameters such as crime type, location, and year-wise occurrences.

2. **Literature Review**

The paper reviews how machine learning enhances cybersecurity by improving malware detection, intrusion detection, and attack prevention. It discusses various cyber threats, ML techniques like deep learning and SVM, and cybersecurity frameworks such as NIST. Key challenges include dataset availability, data leakage, and the need for hybrid learning approaches[1]

The paper explores the integration of cybersecurity and machine learning, highlighting how ML enhances threat detection while facing challenges like data poisoning and privacy breaches. It compares various ML-based security techniques and suggests future improvements in accuracy, efficiency, and compatibility[2]

The paper explores how machine learning enhances cybersecurity in cloud computing by improving threat detection, real-time monitoring, and automated responses. It also discusses challenges like data privacy, scalability, and integration with existing security systems while suggesting future research directions[3]

The paper discusses how machine learning enhances cybersecurity by automating threat detection, phishing prevention, malware analysis, and intrusion detection. It highlights challenges like data privacy, adversarial attacks, and evolving cyber threats, emphasizing the need for continuous model updates and integration with existing security frameworks[4]

The paper discusses how machine learning and big data analytics enhance real-time cybersecurity threat detection by identifying anomalies and predicting attacks. It highlights challenges like false positives and data privacy while emphasizing continuous learning models for improved security[5]

The paper explores how machine learning enhances cybersecurity through intelligent data analysis and automation for threat detection and prevention. It highlights challenges like adversarial attacks and data privacy while emphasizing future AI-driven security solutions[6]

The paper explores the role of Machine Learning (ML) in cybersecurity, highlighting its advantages over traditional human-driven detection methods. It discusses ML applications in threat detection, malware analysis, and phishing prevention, along with challenges like data privacy, adversarial attacks, and deployment gaps between research and practice. The study

also presents real-world case studies and emphasizes the need for collaboration among researchers, industry experts, and policymakers for future advancements[7]

The paper provides a comprehensive review of cybersecurity, covering state-of-the-art techniques, challenges, and future directions. It highlights the role of AI and ML in threat detection, discusses security concerns in IoT, cloud computing, and smart cities, and explores the evolving nature of cyber threats and countermeasures. The study emphasizes collaboration, innovation, and automation to strengthen cybersecurity frameworks[8]

The paper examines how AI and Machine Learning improve cybersecurity by enabling early detection and prediction of cyber attacks. It evaluates models like SVM, Decision Trees, and Random Forest, highlighting their effectiveness in threat mitigation while addressing challenges and ethical concerns[9]

The paper explores how AI and ML enhance cybersecurity by improving threat detection accuracy, response time, and prevention success rates. It highlights that ML algorithms achieve a 95.7% detection accuracy, significantly outperforming traditional methods. AI-powered anomaly detection also reduces response times from 45 to 12 minutes, enabling proactive defense against cyber threats. Despite these advancements, challenges such as algorithm bias, adversarial attacks, and data privacy remain. The study emphasizes the need for continuous innovation and ethical considerations in AI-driven cybersecurity solutions[10]

3. Research methods

Data Collection:

The dataset was obtained from Kaggle, containing cybercrime cases from 2016 to 2018 for different states in India. It was preprocessed to remove missing values and normalize the data for accurate analysis.

Power BI for Data Analysis:

Power BI was used to analyze and visualize the dataset. It provided interactive dashboards for trend analysis and pattern recognition.

The key charts used were:

- **Trend Graph:** This graph was used to visualize the growth of cybercrime cases over multiple years. It provided insights into whether cybercrime rates have been increasing, decreasing, or fluctuating, helping to identify patterns and long-term trends.
- **Bar Chart:** The bar chart compared the number of cybercrime cases across different states in India. This allowed us to identify which states had the highest and lowest number of reported cases, helping to analyze regional variations and hotspots of cybercrime activity.
- **Pie Chart:** This chart illustrated the proportion of different types of cybercrimes, such as online fraud, identity theft, and hacking. It helped in understanding which category of cybercrime is most prevalent and required more attention from law enforcement and policymakers.

- **Line Chart:** The line chart depicted the fluctuations in cybercrime rates over time, highlighting periods of sudden spikes or declines. It was useful in analyzing short-term variations and possible seasonal trends in cybercrime activities.

These visualizations helped in identifying key trends and patterns in cybercrime across India.

4. Research Findings

a)How has cybercrime increased from 2016 to 2018 across India?

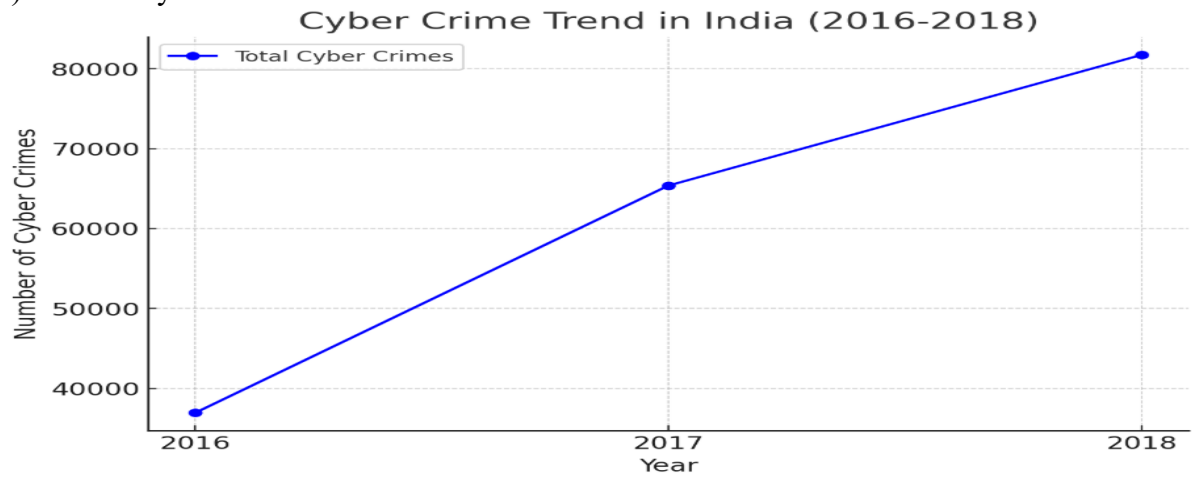

Figure 1Fig.(1) : Cyber Crime Trend Graph

Here is the trend graph showing the increase in cybercrimes across India from 2016 to 2018. The graph clearly indicates a rising trend, with cybercrimes growing each year.

b) Which states had the highest increase in cybercrimes?

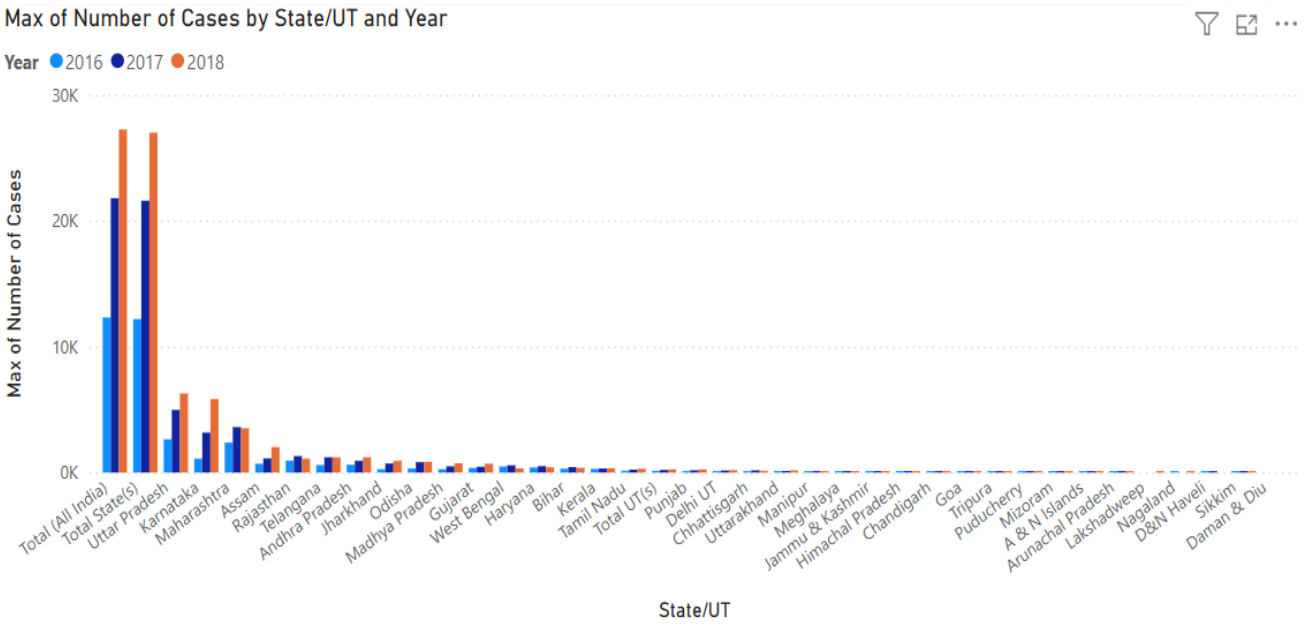

Figure 2Cybercrime Cases by State/UT (2016-2018)

The bar chart shows that Uttar Pradesh, Karnataka, and Maharashtra had the highest cybercrime cases, with a significant increase from 2016 to 2018. The taller bars for 2018 indicate a rising trend, especially in these states. States with shorter bars have relatively lower cybercrime growth.

c)What is the percentage share of cybercrimes per state?

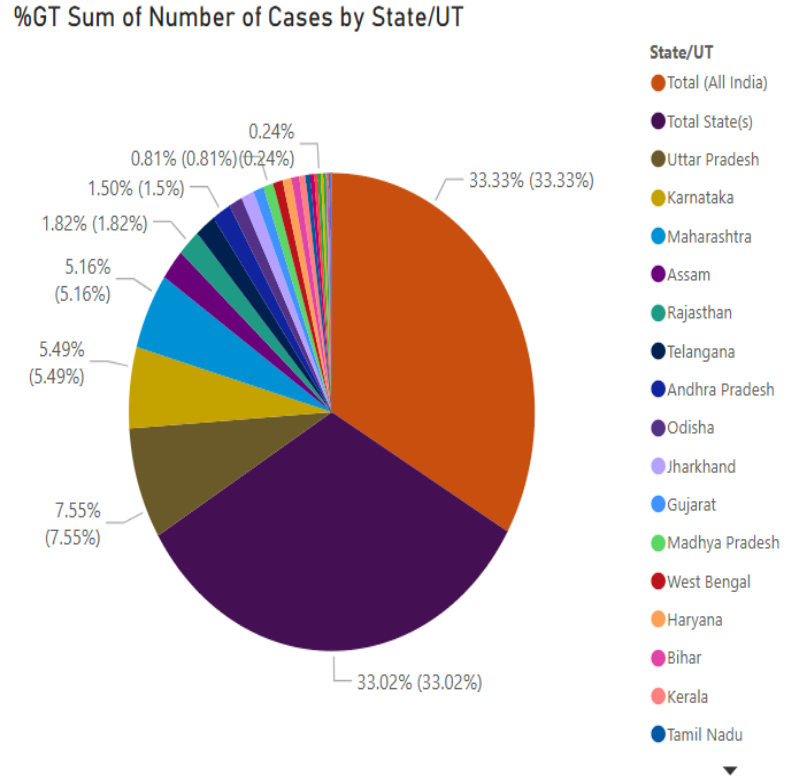

Figure 3 percentage share of cybercrimes per state

This pie chart shows the percentage share of cybercrimes across different states in **2018**. **Uttar Pradesh (7.55%)** recorded the highest share among states, followed by **Karnataka (5.49%)** and **Maharashtra (5.16%)**. Smaller slices represent states with lower cybercrime cases.

d)Which states had the highest & lowest cybercrimes in 2018?

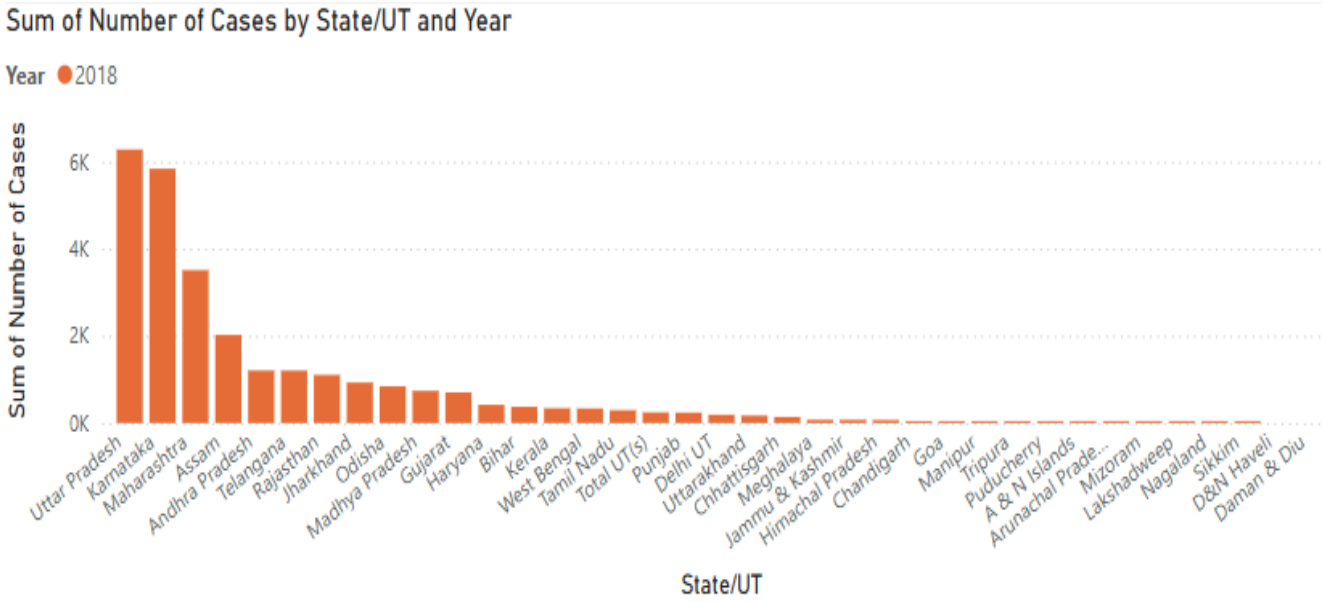

Figure 4 Highest & lowest cybercrimes in 2018

The bar chart shows cybercrime cases across Indian states in 2018. **Uttar Pradesh, Karnataka, and Maharashtra** had the highest number of cases, while **Sikkim, Nagaland, D & N Haveli and Daman & Diu** reported the least. This visualization helps identify the most and least affected regions.

e) Which states have the highest crime rate per lakh population?

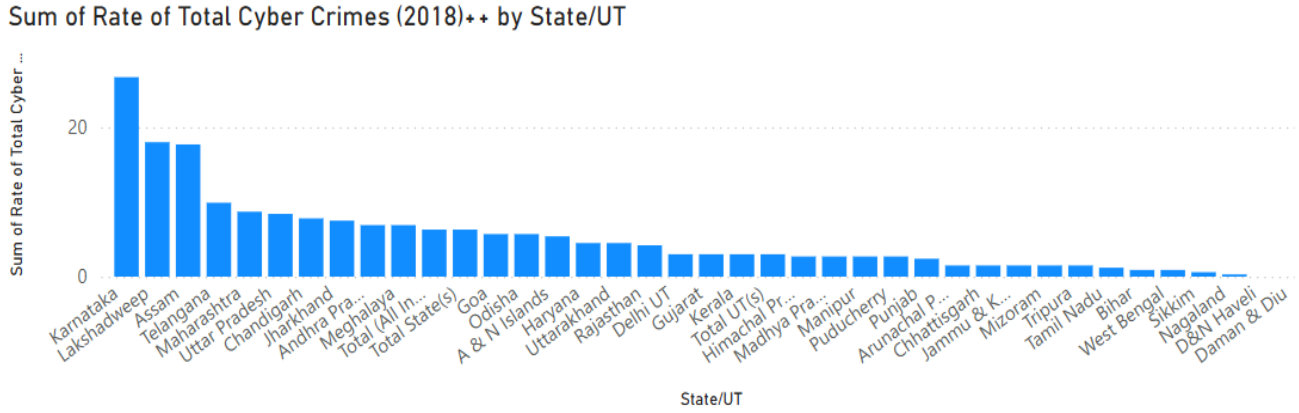

Figure 5Cybercrime Rate by State/UT in 2018

The bar chart shows the **cybercrime rate per lakh population** for different states in 2018. **Karnataka** has the highest crime rate, followed by **Lakshadweep, Assam, Telangana, and Maharashtra**. The states with the lowest cybercrime rates include **Daman & Diu, Nagaland, and West Bengal**.

f) Can we forecast cybercrime trends for 2019,2020 & 2021?

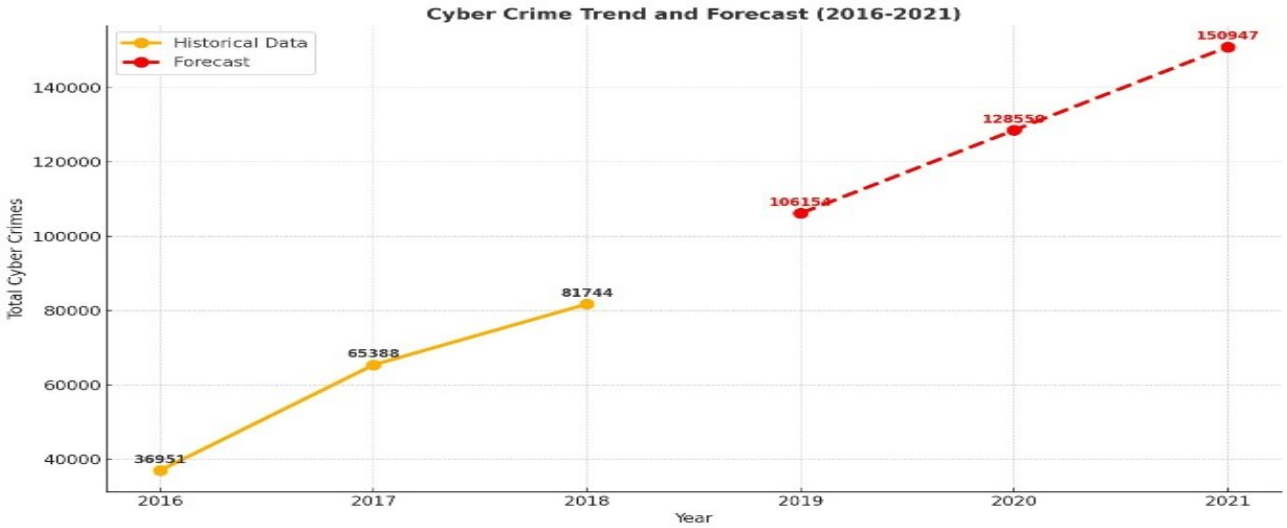

Figure 6 Forecast cybercrime trends for 2019,2020 & 2021

It is the clearer and enhanced graph showing the cybercrime trend (2016-2018) and forecast for 2019, 2020, and 2021. Data labels are added for better readability. The blue line represents historical data. The dashed red line represents the forecasted values.

g)Which states with high cybercrime rates in 2018 continue to rise in 2019,2020 and 2021?

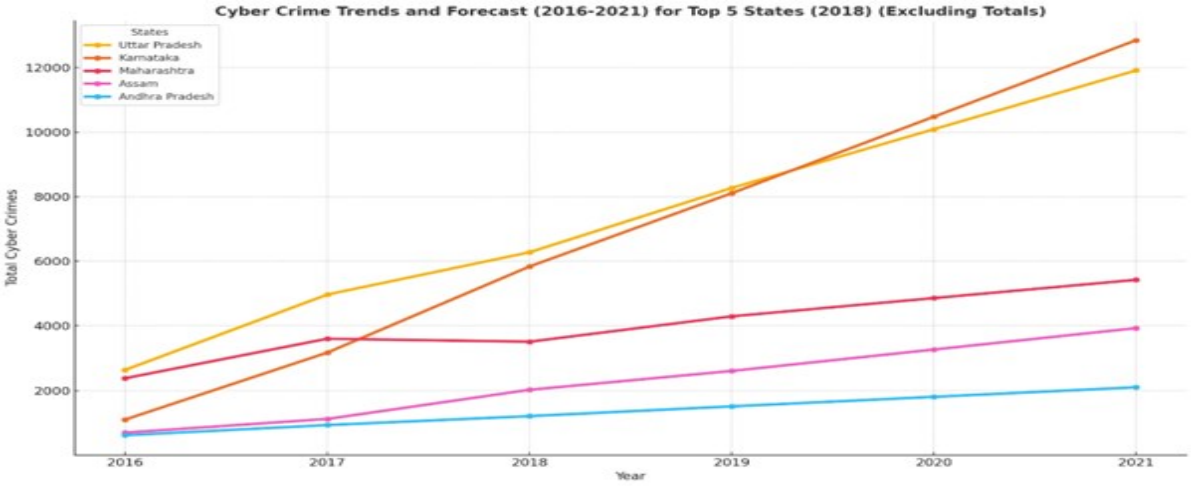

Figure 7 Cybercrime Trends (2016-2021) – Top 5 States

This line chart shows cybercrime trends and forecasts (2016-2021) for the top 5 states with the highest cases in 2018. Uttar Pradesh and Karnataka exhibit the steepest rise, indicating a significant increase in cybercrimes, while other states show a steady upward trend. The forecast suggests a continued rise in cybercrime cases through 2020 and 2021.

h)Which states with lowest cybercrime rates in 2018 continue to rise in 2019,2020 and 2021?

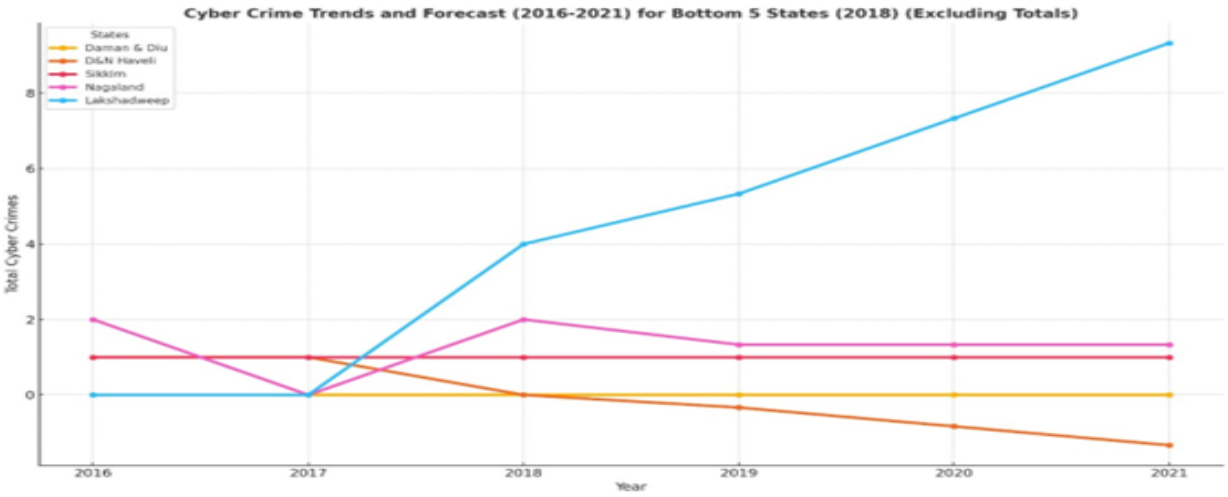

Figure 8 Cybercrime Trends (2016-2021) – Bottom 5 States

This line chart represents cybercrime trends and forecasts (2016-2021) for the bottom 5 states with the lowest cases in 2018. While most states show minimal or stagnant cybercrime activity, Lakshadweep exhibits a noticeable increase in cases from 2018 onwards, suggesting emerging cybercrime concerns in the region.

i)Maximum number of cases in 2016,2017 & 2018.

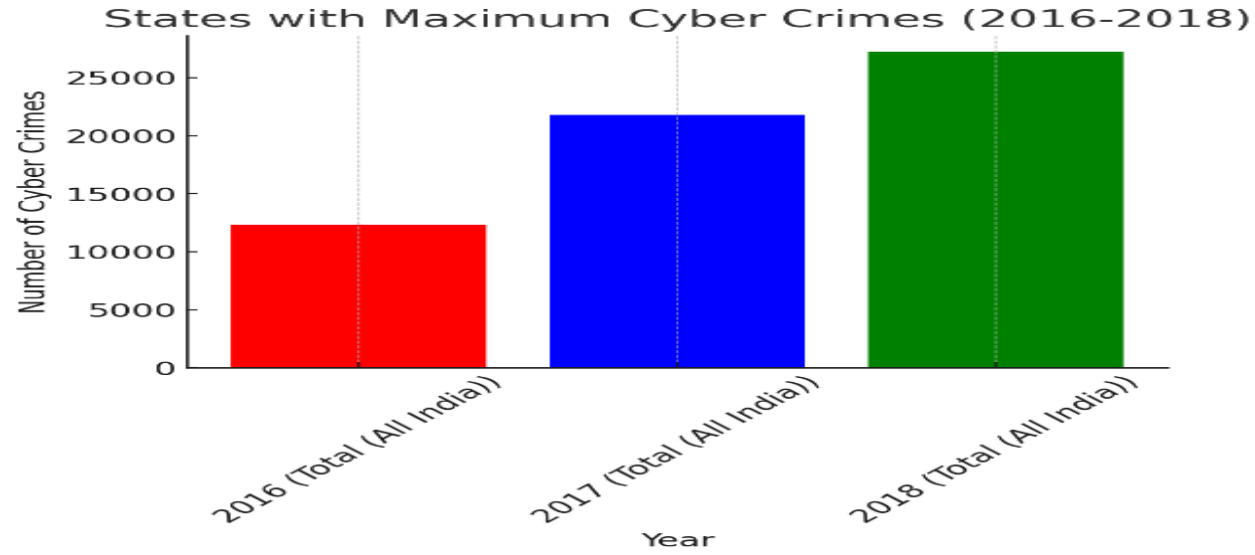

Figure 9 Maximum number of cases in 2016,2017 & 2018

Here is the bar graph showing the states with the maximum number of cybercrime cases in 2016, 2017, and 2018. Each bar represents the highest reported cases for the respective year along with the corresponding state.

j)Maximum number of cases by state in 2016.

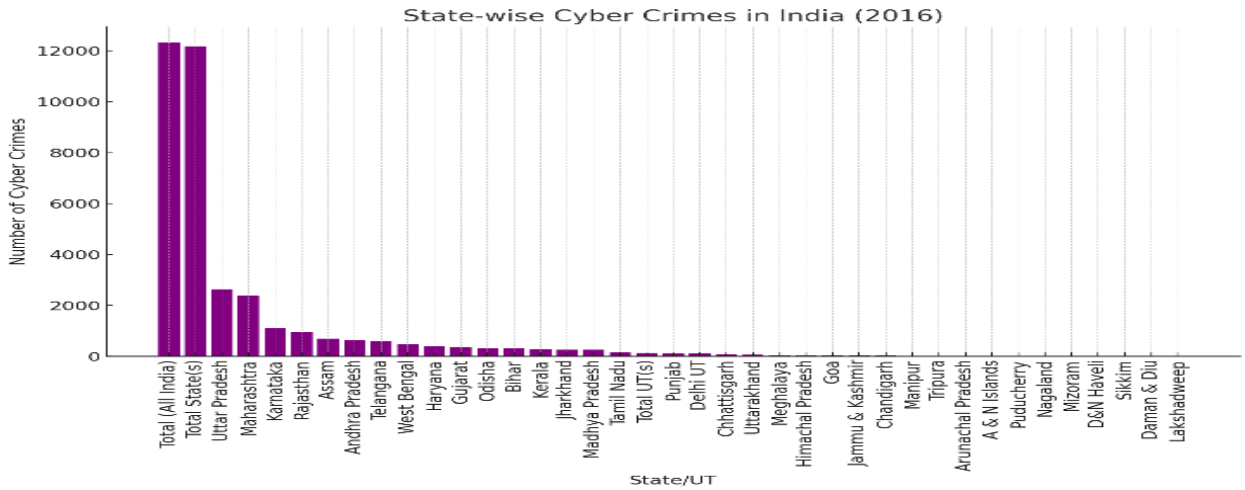

Figure 10 Maximum number of cases by state in 2016

This bar graph represents the state-wise distribution of cybercrime cases in India for 2016. Each bar corresponds to a state or union territory, with its height representing the number of cybercrime cases reported that year.

k)Maximum number of cases statewise in india 2017.

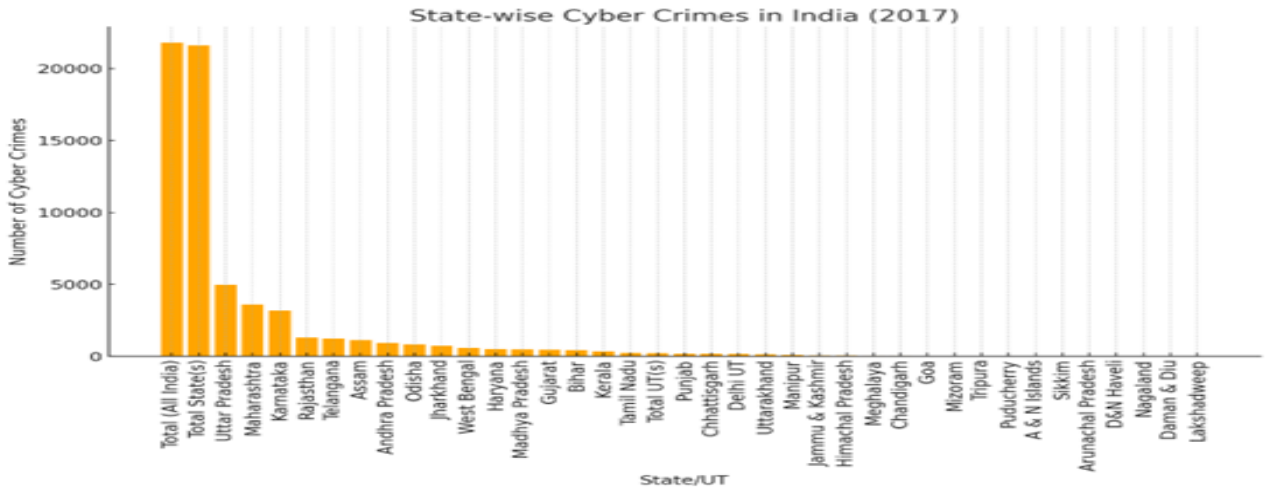

Figure 11Maximum number of cases by state in 2017

Here is the state-wise cybercrime distribution in India for 2017. The states on the left had the highest number of cases, while those on the right had fewer incidents.

l)Maximum number of cases statewise in india 2018.

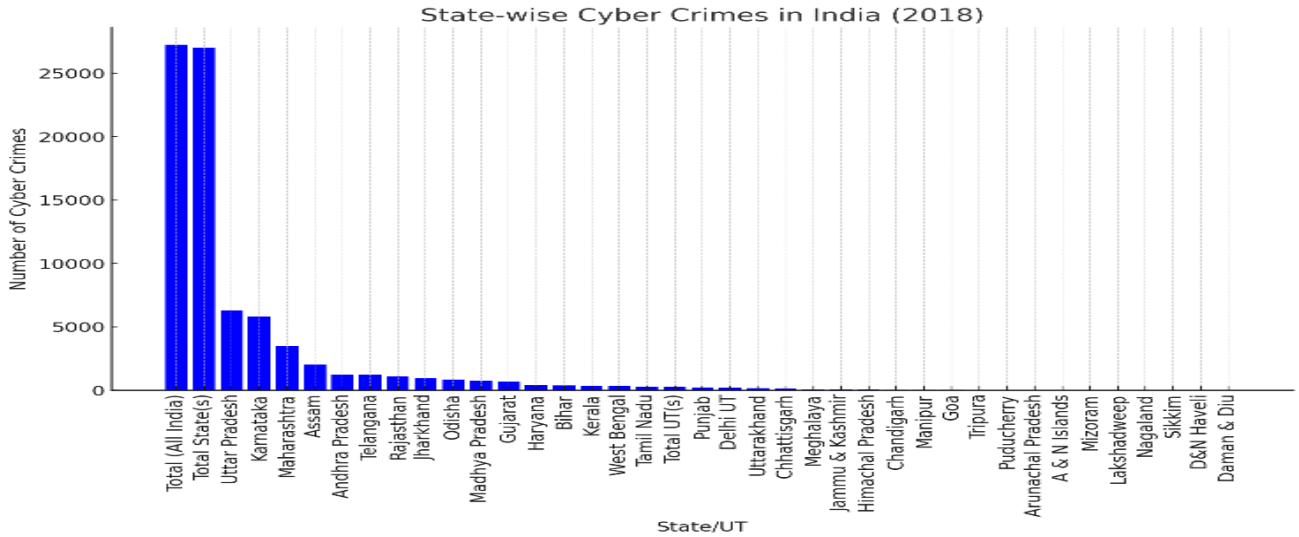

Figure 12Maximum number of cases by state in 2017

Here is the state-wise cybercrime distribution in India for 2018. The states on the left had the highest number of cases, while those on the right had fewer incidents.

m)Comparative analysis of the first five states having highest number of crimes from 2016 to 2018.

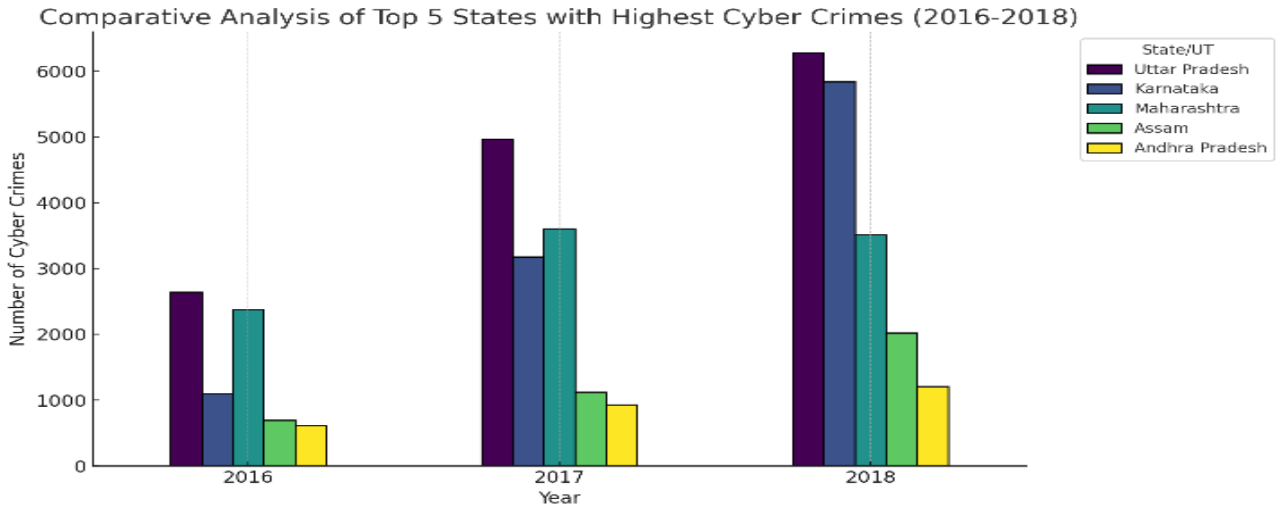

Figure 13Top 5 States – Cybercrime Comparison (2016-2018)

This bar chart compares the top 5 states with the highest cybercrime cases from 2016 to 2018. It shows a rising trend in most states, indicating an increase in cybercrimes over the years. The tallest bars highlight the states with the highest number of cases each year.

n)Comparative analysis of bottom 5 states having atleast one case from 2016 to 2018.

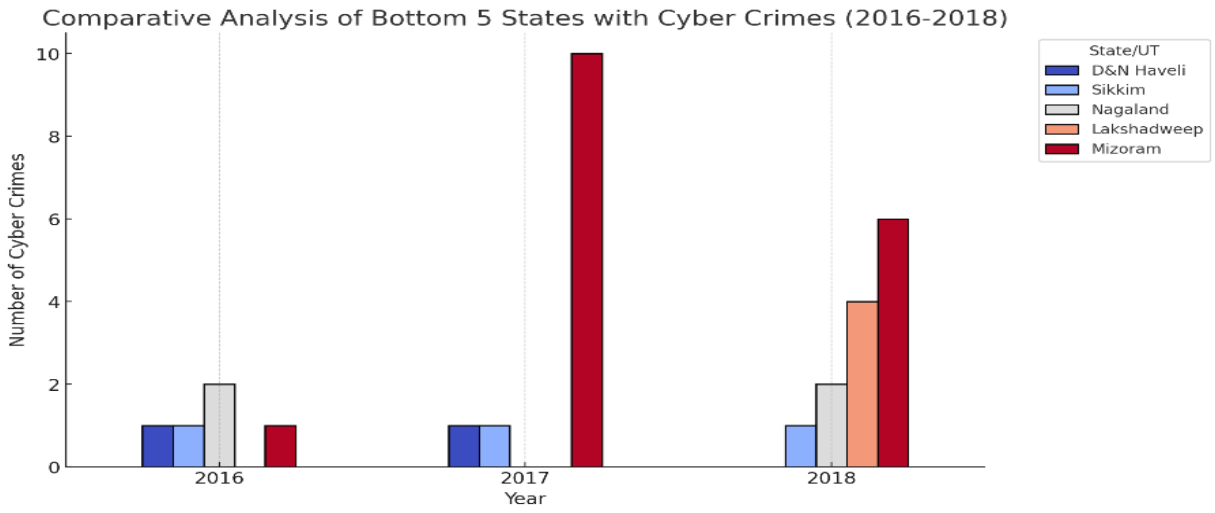

Figure 14Bottom 5 States – Cybercrime Comparison (2016-2018)

This bar chart shows the bottom 5 states with the lowest cybercrime cases (excluding zero-case states) from 2016 to 2018. These states had the least reported incidents, with some showing slight increases or fluctuations over the years. The comparison helps identify regions with minimal cybercrime activity.

Sikkim and D & N Haveli having constant crime rate.

5. Conclusion & Future Work

Conclusion:

This study highlights a rising trend in cybercrimes across India, emphasizing the need for proactive measures. The analysis demonstrates that machine learning models can effectively predict future cybercrime patterns, aiding law enforcement and policymakers in strategic decision-making.

- Future Work

  To enhance prediction accuracy and applicability, future research should focus on: Integrating real-time cybercrime reporting data to improve data reliability. Exploring deep learning models to enhance prediction accuracy and uncover hidden patterns in cybercrime trends.

6. References

1. Ahsan, M., Nygard, K. E., Gomes, R., Chowdhury, M. M., Rifat, N., & Connolly, J. F. (2022). Cybersecurity threats and their mitigation approaches using Machine Learning—A Review. Journal of Cybersecurity and Privacy, 2(3), 527-555.
2. Wazid, M., Das, A. K., Chamola, V., & Park, Y. (2022). Uniting cyber security and machine learning: Advantages, challenges, and future research. ICT Express, 8(3), 313-321.

3. Mukesh, V. (2022). Cloud Computing Cybersecurity Enhanced by Machine Learning Techniques. Journal ID, 1663, 8854.
4. Bharadiya, J. (2023). Machine learning in cybersecurity: Techniques and challenges. European Journal of Technology, 7(2), 1-14.
5. Ofoegbu, K. D. O., Osundare, O. S., Ike, C. S., Fakeyede, O. G., & Ige, A. B. (2024). Real-Time Cybersecurity threat detection using machine learning and big data analytics: A comprehensive approach. Computer Science & IT Research Journal, 4(3).
6. Sarker, I. H. (2023). Machine learning for intelligent data analysis and automation in cybersecurity: Current and future prospects. Annals of Data Science, 10(6), 1473-1498.
7. Apruzzese, G., Laskov, P., Montes de Oca, E., Mallouli, W., Brdalo Rapa, L., Grammatopoulos, A. V., & Di Franco, F. (2023). The role of machine learning in cybersecurity. Digital Threats: Research and Practice, 4(1), 1-38.
8. Admass, W. S., Munaye, Y. Y., & Diro, A. A. (2024). Cyber security: State of the art, challenges, and future directions. Cyber Security and Applications, 2, 100031.
9. Khalaf, M. A., & Steiti, A. (2024). Artificial intelligence predictions in cyber security: Analysis and early detection of cyber attacks. Babylonian Journal of Machine Learning, 2024, 63-68.
10. Hussain, H., Kainat, M., & Ali, T. (2025). Leveraging AI and Machine Learning to Detect and Prevent Cyber Security Threats. Dialogue Social Science Review (DSSR), 3(1), 881-895.