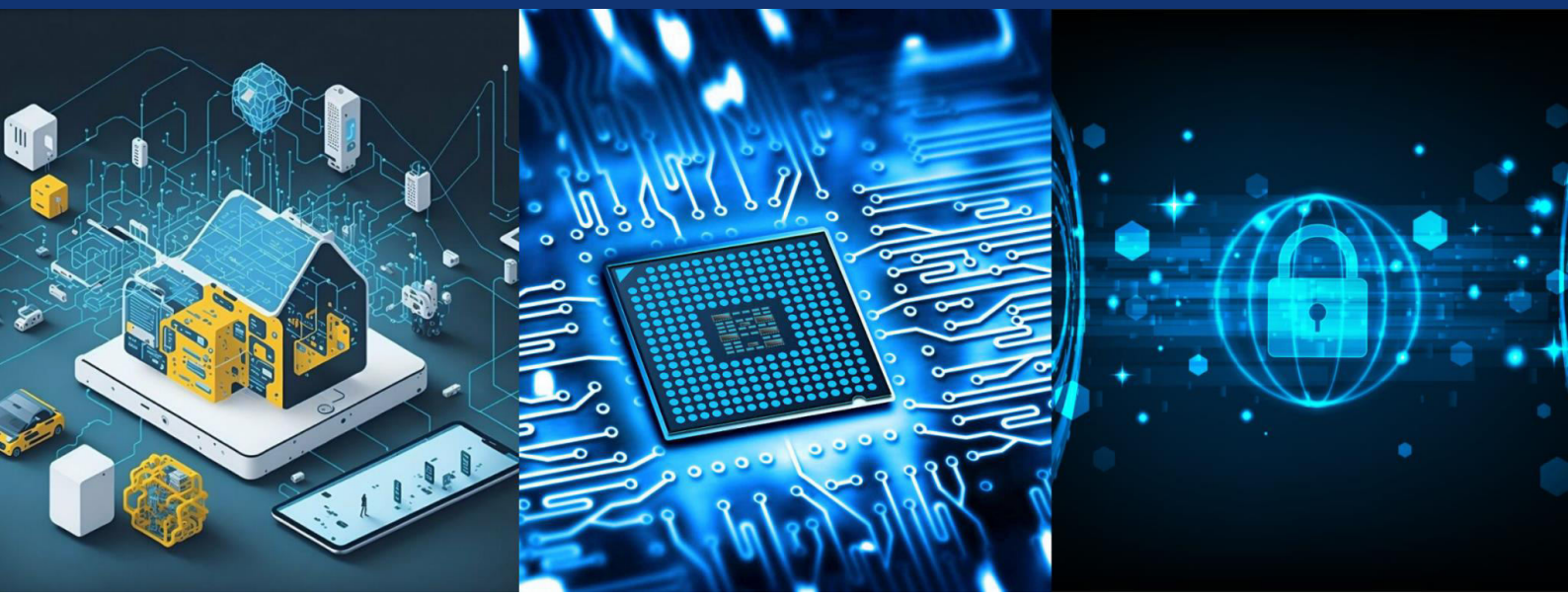
IJIRCCE

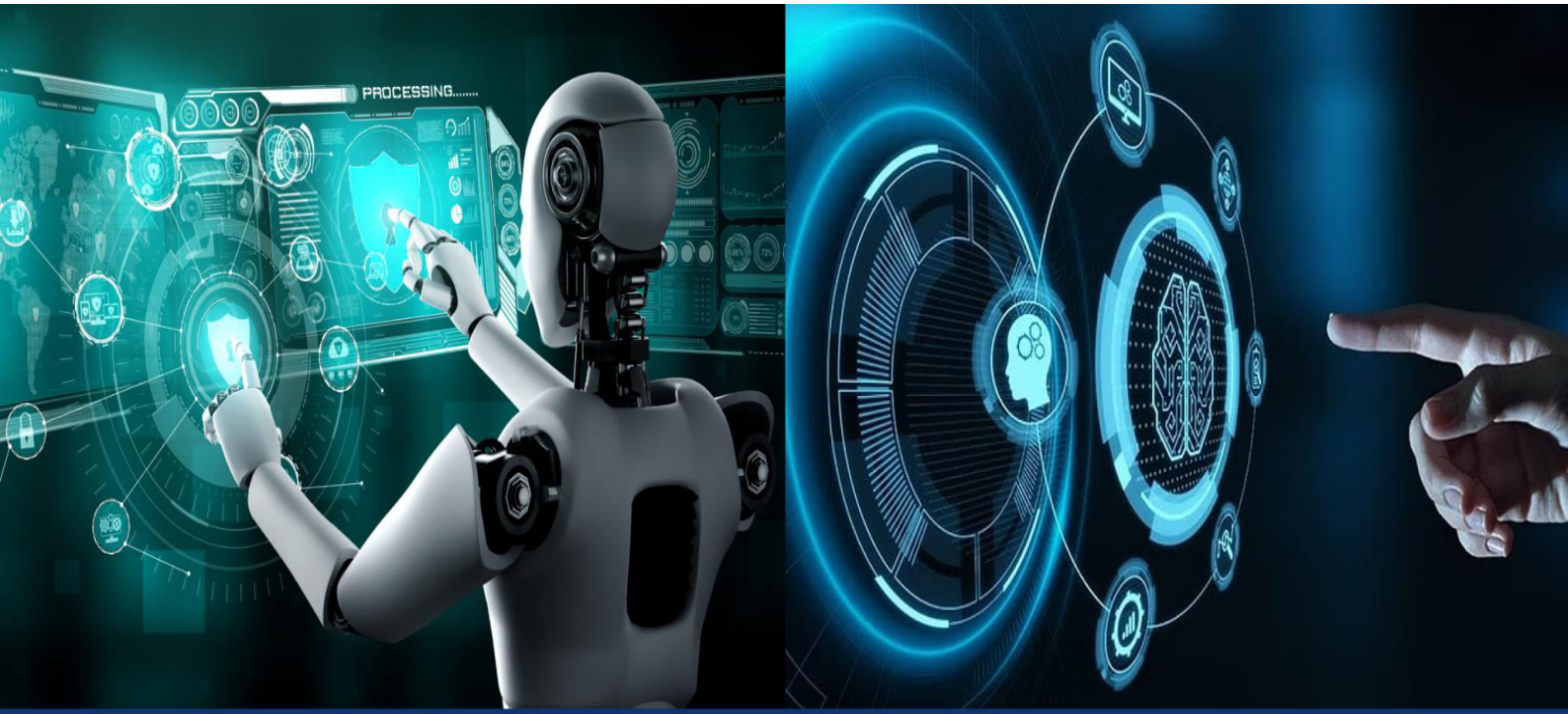
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 6, June 2026

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406092

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Crime Prediction System

Dr. T.V.S Sriram[1], P. Prasad[2], K. Pranay[3]

Sr.Asst. Professor, Dept.of MCA, NSRIT, Visakhapatnam, AP, India[1]

Asst. Professor, Dept.of MCA, NSRIT, Visakhapatnam, AP, India[2]

Dept. of MCA, NSRIT, Visakhapatnam, AP, India[3]

ABSTRACT: Crime Prediction System is a machine learning-based application designed to predict crime categories using historical crime data. The system analyzes crime records, identifies patterns, and predicts the type of crime that is likely to occur based on input parameters. By utilizing advanced machine learning algorithms such as LightGBM, the system provides accurate predictions that assist law enforcement agencies in crime prevention and resource allocation. The application offers a user-friendly interface for entering crime-related information and obtaining prediction results, thereby supporting data-driven decision-making and improving public safety.

KEYWORDS: Crime prediction, Machine learning, LightGBM, Data Analysis, Crime Classification, Public Safety.

I. INTRODUCTION

Crime is a major issue that affects public safety and social stability. With the increase in crime incidents, analyzing crime data manually has become difficult and time-consuming. Machine Learning provides an effective way to analyze historical crime records and predict crime categories based on patterns in the data.

The Crime Prediction System is developed to predict crime types using machine learning algorithms. The system processes crime-related data, identifies patterns, and generates accurate predictions. This helps law enforcement agencies take preventive measures and make better decisions.

Overall, the Crime Prediction System improves crime analysis, supports public safety, and assists authorities in effective crime management.

II. LITERATURE REVIEW

Several studies have explored the use of Machine Learning techniques for crime prediction and analysis. Traditional crime analysis methods mainly rely on historical reports and manual investigation, which can be time-consuming and less effective in identifying future crime patterns.

Recent research has shown that machine learning algorithms such as Decision Trees, Random Forest, Support Vector Machines, and LightGBM can be used to analyze crime data and predict crime categories with high accuracy. These systems help law enforcement agencies identify trends, improve resource allocation, and support crime prevention strategies.

Although many existing systems provide crime analysis and visualization, some lack accurate prediction capabilities and efficient data processing. The proposed Crime Prediction System addresses these limitations by utilizing machine learning models to generate reliable crime predictions based on historical crime records, thereby improving public safety and decision-making.

III. SYSTEM OVERVIEW

The Crime Prediction System is a machine learning-based application designed to analyze historical crime data and predict crime categories accurately. The system helps law enforcement agencies identify crime patterns and make informed decisions for crime prevention and public safety.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The core functionality of the system is based on data analysis and machine learning algorithms. Users provide crime-related information, and the system processes the data to generate crime predictions. The prediction results help authorities understand potential crime trends and take preventive measures accordingly.

The system consists of three main components: users, machine learning models, and administrators. Users can enter crime-related data and view prediction results, while administrators manage datasets and monitor system performance. The machine learning model analyzes the data and provides accurate crime predictions, creating an efficient and intelligent crime analysis environment.

The Crime Prediction System is designed using a modular architecture that ensures scalability, accuracy, and efficient data processing. It consists of components such as the user interface layer, data processing layer, machine learning layer, and database layer. The machine learning layer plays a vital role in analyzing crime data and generating predictions, while the database layer stores crime records and prediction results. This architecture enables seamless interaction between users and the prediction system, providing accurate and reliable crime analysis.

IV. METHODOLOGY

The development of the Crime Prediction System follows a structured approach to ensure accurate crime analysis and prediction. The methodology consists of the following phases.

4.1 Data Collection

Crime-related data such as crime type, location, date, time, and incident details are collected from historical crime datasets for analysis and prediction

4.2 Data Proprocessing

The collected data is cleaned by handling missing values, removing inconsistencies, and normalizing features to improve data quality and model performance.

Feature Extraction

Important attributes influencing crime occurrence are identified and selected to enhance the efficiency and accuracy of the prediction model.

4.4 Model Training

Machine learning algorithms such as LightGBM, Random Forest, K-Means Clustering, and Facebook Prophet are trained using the processed dataset to learn crime patterns and trends.

4.5 Prediction and Classification

The trained model analyzes user input data and classifies crimes into categories such as Violent Crime, Property Crime, and Other Crimes. The system also predicts future crime trends and identifies potential crime hotspots.

4.6 Performance Evaluation

The system performance is evaluated using metrics such as Accuracy, Precision, Recall, and F1-Score to determine the effectiveness and reliability of the proposed framework.

V. SYSTEM ARCHITECTURE

The architecture of the Crime Prediction System follows a layered design that enables efficient processing, analysis, and prediction of crime-related data. The layered architecture improves system performance, scalability, and accuracy by organizing the system into distinct functional layers.

The system consists of four major layers: User Interface Layer, Application Logic Layer, Data Integration Layer, and Database Layer.

The User Interface Layer provides an interactive environment where users can register, log in, enter crime-related information, and view prediction reports through the crime dashboard.

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406092

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The Application Logic Layer handles user authentication, data processing, machine learning model execution, and crime prediction. This layer acts as the core processing unit of the system and ensures accurate prediction generation.

The Data Integration Layer collects and manages crime-related information from historical crime datasets, crime location data, and crime records. It ensures that the required data is available for analysis and prediction.

The Database Layer stores user information, crime records, prediction results, and crime reports. This layer enables efficient data management, retrieval, and reporting for future analysis.

Together, these layers ensure efficient crime data processing, accurate crime prediction, secure data storage, and effective decision-making support for law enforcement agencies and researchers.

VI. MODULES

The Crime Prediction System consists of three primary modules, each designed to perform specific functions within the system Crime Prediction- Module Overview

6.1 User Module

User registrationand secure login Enter crime-related information Viewn prediction crime categories Access crime prediction reports

6.2 Prediction Module Process and clean crime date

Apply feature selection techniques Execute the LightGBM Prediction Generate accurate crime prediction

6.3 Admin Module Manage user accounts

Monitor system performance Maintain crime datasets

Generate operational and analytical reports

VII. IMPLEMENTATION DETAILS

Technology Stack

The Crime Prediction System is implemented using modern technologies and machine learning tools. Python is used as the programming language, while Streamlit is used for developing the interactive web interface. Data preprocessing and analysis are performed using Pandas and NumPy. Machine learning models such as LightGBM, Random Forest, K-Means Clustering, and Facebook Prophet are used for crime prediction, hotspot analysis, and forecasting. MySQL is used for efficient data storage and management, while Plotly and Matplotlib are used for data visualization and report generation.

Frontend Implementation

The frontend is designed to provide a simple and user-friendly interface where users can enter crime-related information, view prediction results, analyze crime trends, and access generated reports. The responsive design ensures accessibility across different devices.

Backend Implementation

The backend handles data preprocessing, machine learning model execution, prediction generation, user authentication, and communication between the frontend and database. It ensures efficient processing and accurate crime prediction.

Database Design

The database is structured to store and manage crime-related information efficiently. Key tables include User Table, Crime Records Table, Prediction Results Table, and Admin Table. Proper database relationships are maintained to ensure data integrity and efficient retrieval.

Security Measures

To ensure system security and reliability, several security measures are implemented in the Crime Prediction System. Secure password hashing and encrypted authentication are used to protect user credentials. Input validation helps prevent malicious data entry and maintains data integrity.

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406092

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Session management provides secure user access and prevents unauthorized usage. Role-based access control restricts administrative functions to authorized users only. Additionally, crime-related data is stored securely to ensure privacy and confidentiality.

VIII. EXPERIMENTAL DETAILS AND RESULTS

Experimental Setup

The proposed Crime Prediction System was implemented and tested using Python, Streamlit, Scikit-learn, NumPy, Pandas, and LightGBM. The experiments were conducted on a system with an Intel Core i5 processor, 8GB RAM, and Windows operating system. Historical crime records, crime categories, location details, and incident information were used for model training and evaluation.

Test Results

Several machine learning models were evaluated for crime prediction. The performance of the models was measured using standard classification metrics.

Test Case Results for Crime Prediction System Performance Evaluation

The system performance was evaluated using Accuracy, Precision, Recall, and F1-Score metrics. Experimental results indicate that the proposed framework effectively predicts crime categories with high reliability and minimal prediction delay

Limitations Observed

The proposed system successfully classified crime categories and identified crime patterns using machine learning techniques. The use of LightGBM, Random Forest, K-Means Clustering, and Facebook Prophet improved prediction accuracy and supported effective crime analysis and forecasting.

IX. APPLICATIONS

Crime Prediction and Prevention:

The system helps predict potential crime categories using historical crime data, enabling proactive crime prevention strategies.

Law Enforcement Support:

Police departments and law enforcement agencies can use the system to identify crime patterns and improve decision-making.

Crime Hotspot Analysis:

The system assists in identifying high-crime areas, helping authorities allocate resources more effectively.

Public Safety Monitoring:

Crime prediction results can be used to enhance public safety and reduce crime risks in vulnerable regions.

Smart City Applications:

The system can be integrated with smart city initiatives to support intelligent crime monitoring and urban security management.

Resource Allocation Planning:

Authorities can optimize the deployment of police personnel and security resources based on predicted crime trends.

Research and Crime Analysis:

Researchers and analysts can use crime data and prediction results to study crime behavior and trends.

Crime Forecasting:

The system supports future crime trend prediction, helping organizations develop effective crime prevention policies

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406092

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

X. LIMITATIONS

Data Dependency:
The accuracy of the system depends on the quality and reliability of the crime dataset used for training.

Limited Dataset Availability:
Insufficient or incomplete crime records may affect prediction performance and accuracy.

Changing Crime Patterns:
Crime trends may change over time, which can reduce the effectiveness of previously trained models.

Prediction Accuracy:
Machine learning predictions may not always be 100% accurate due to unforeseen factors influencing crime.

Real-Time Data Availability:
The system may face challenges in obtaining real-time crime data for instant prediction and analysis.

Geographical Variations:
Crime patterns differ across locations, making it difficult to generalize predictions for all regions. Privacy and Security Concerns:
Crime-related data must be stored securely to prevent unauthorized access and misuse.

XI. FUTURE WORK

The future scope of the Crime Prediction System focuses on improving prediction accuracy, reliability, and usability through advanced technologies and intelligent analytics. Advanced machine learning and deep learning algorithms can be integrated to enhance crime classification, forecasting, and pattern recognition. The system can be extended with real-time crime monitoring capabilities using live crime feeds, IoT devices, and surveillance systems.

Future enhancements may include crime hotspot detection using geospatial analysis, GIS-based crime visualization, and interactive dashboards for better decision-making. Mobile application support can be introduced to provide instant access to crime predictions and reports. AI-powered alert systems can notify law enforcement agencies about potential crime risks and emerging crime patterns.

The system can also be integrated with government and law enforcement databases to improve data availability and prediction accuracy. Furthermore, smart city infrastructure and cloud-based technologies can be utilized to support large-scale crime monitoring and management. These enhancements can significantly improve crime prevention strategies, resource allocation, public safety, and overall law enforcement effectiveness.

XII. CONCLUSION

The proposed Crime Prediction System provides an effective machine learning-based solution for analyzing crime data and predicting crime categories. The system performs data preprocessing, feature extraction, classification, and forecasting to generate accurate crime predictions. By utilizing algorithms such as LightGBM, Random Forest, K-Means Clustering, and Facebook Prophet, the system improves crime analysis and supports data-driven decision-making. Experimental results demonstrate that the proposed framework can effectively identify crime patterns and predict future crime trends with high accuracy. The system can assist law enforcement agencies in crime prevention, resource allocation, and public safety management. With future enhancements such as real-time monitoring and advanced AI techniques, the Crime Prediction System can become a reliable tool for intelligent crime analysis and forecasting.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

[1] T. Mitchell, Machine Learning, McGraw-Hill, 1997.
[2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
[3] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed., O'Reilly Media, 2022.
[4] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed., Morgan Kaufmann, 2012.
[5] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[6] W. McKinney, Python for Data Analysis, 3rd ed., O'Reilly Media, 2022.
[7] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[8] C. Cortes and V. Vapnik, "Support Vector Networks," Machine Learning, vol. 20, pp. 273–297, 1995.
[9] J. R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann, 1993.
[10] T. Cover and P. Hart, "Nearest Neighbor Pattern Classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967.
[11] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
[12] S. Raschka and V. Mirjalili, Python Machine Learning, 3rd ed., Packt Publishing, 2019.
[13] M. Kuhn and K. Johnson, Applied Predictive Modeling, Springer, 2013.
[14] S. Haykin, Neural Networks and Learning Machines, 3rd ed., Pearson, 2009.
[15] R. S. Pressman and B. R. Maxim, Software Engineering: A Practitioner's Approach, 8th ed., McGraw-Hill Education, 2015.
[16] I. Sommerville, Software Engineering, 10th ed., Pearson Education, 2016.
[17] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," NeurIPS, 2017.
[18] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[19] D. Silver et al., "Mastering the Game of Go with Deep Neural Networks and Tree Search," Nature, vol. 529, pp. 484–489, 2016.
[20] J. Brownlee, Machine Learning Mastery with Python, Machine Learning Mastery, 2016.
[21] E. Alpaydin, Introduction to Machine Learning, 4th ed., MIT Press, 2020.
[22] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[23] F. Chollet, Deep Learning with Python, 2nd ed., Manning Publications, 2021.
[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[25] G. Hinton and R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," Science, vol. 313, no. 5786, pp. 504–507, 2006.
[26] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.
[27] D. Jurafsky and J. H. Martin, Speech and Language Processing, Pearson, 2023.
[28] J. Brownlee, Deep Learning for Computer Vision, Machine Learning Mastery, 2019.
[29] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2017.
[30] R. Collobert and J. Weston, "A Unified Architecture for Natural Language Processing," ICML, 2008.
[31] J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," Communications of the ACM, vol. 51, no. 1, pp. 107–113, 2008.
[32] M. Zaharia et al., "Apache Spark: A Unified Engine for Big Data Processing," Communications of the ACM, vol. 59, no. 11, pp. 56–65, 2016.
[33] K. Schwaber and J. Sutherland, The Scrum Guide, Scrum.org, 2020.
[34] R. Elmasri and S. Navathe, Fundamentals of Database Systems, 7th ed., Pearson, 2016.
[35] A. Silberschatz, H. Korth, and S. Sudarshan, Database System Concepts, 7th ed., McGraw-Hill, 2019.
[36] M. Fowler, UML Distilled, 3rd ed., Addison-Wesley, 2003.
[37] T. White, Hadoop: The Definitive Guide, 4th ed., O'Reilly Media, 2015.
[38] I. Witten, E. Frank, and M. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016.
[39] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed., Pearson, 2021.
[40] J. Leskovec, A. Rajaraman, and J. Ullman, Mining of Massive Datasets, Cambridge University Press, 2020.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details