**Crime Hotspots prediction using ML**

**ABSTRACT: -**

Predicting crime is crucial for shaping effective policing strategies and implementing measures for prevention and control. While machine learning has emerged as the prevailing method for such forecasts, there's a scarcity of comprehensive studies comparing different algorithms. This study addresses this gap by examining historical public property crime data in a sizable coastal urban area in Southeast Asia. Through a systematic evaluation of multiple machine learning techniques. it reveals that LSTM outperforms its counterparts when considering only historical crime data. Furthermore, the incorporation of built environment features, such as points of interest (POIs) and urban road network density, into the LSTM model significantly enhances its predictive capabilities. This underscores the significance of integrating both historical crime data and covariates associated with criminological theories for more accurate crime prediction models.

|  |  |
| --- | --- |
| **EXSISTING SYSTEM** | **PROPOSED SYSTEM** |
| Previous existing systems for crime prediction typically relied on statistical methods or rule-based systems. These systems often utilized historical crime data, demographic information, and geographical features to identify patterns and make predictions. However, these approaches had limitations in terms of accuracy and scalability. Data mining, on the other hand, is a process of discovering patterns, anomalies, and correlations within large datasets. It involves various techniques such as clustering, classification, association rule mining, and anomaly detection. In the context of crime prediction, data mining techniques are used to extract valuable insights from diverse sources of data including crime records, socioeconomic indicators, geographic data, and environmental factors. | Machine learning is a rapidly advancing field that has revolutionized the way we approach problem-solving and decision-making. It is a subset of artificial intelligence that enables systems to learn and improve from experience without being explicitly programmed. By leveraging algorithms and statistical models, machine learning allows computers to perform specific tasks effectively by analyzing data and identifying patterns. This technology has found applications in a wide range of domains, including natural language processing, computer vision, speech recognition, and predictive analytics. Machine learning algorithms can be broadly categorized into supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training models on labeled data to make predictions or decisions, while unsupervised learning aims to uncover hidden patterns and structures in unlabeled data. Reinforcement learning, on the other hand, focuses on training agents to take actions in an environment to maximize a reward. |
| **EXISTING ALGORITHM**   * Data Mining | **PROPOSED ALGORITHM: -**   * Machine Learning Techniques (random forest, svm) |
| Data mining is a rapidly evolving field that involves the process of extracting and discovering patterns in large data sets. It is an interdisciplinary subfield of computer science and statistics that aims to extract valuable information from massive amounts of data. Data mining involves various techniques such as association rule learning, clustering, classification, regression, and summarization to uncover hidden patterns and relationships within the data. These methods are used to identify unusual data records, relationships between variables, and to group similar data points together. Data mining is crucial in various domains, including market research, customer profiling, fraud detection, and predictive analytics. It is also used in credit risk management, spam filtering, and sentiment analysis. | **ALGORITHM DEFINITION: -**  Machine learning is a rapidly advancing field that has revolutionized the way we approach problem-solving and decision-making. It is a subset of artificial intelligence that enables systems to learn and improve from experience without being explicitly programmed. By leveraging algorithms and statistical models, machine learning allows computers to perform specific tasks effectively by analyzing data and identifying patterns. This technology has found applications in a wide range of domains, including natural language processing, computer vision, speech recognition, and predictive analytics. Machine learning algorithms can be broadly categorized into supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training models on labeled data to make predictions or decisions, while unsupervised learning aims to uncover hidden patterns and structures in unlabeled data. Reinforcement learning, on the other hand, focuses on training agents to take actions in an environment to maximize a reward. As machine learning continues to advance, it has the potential to drive innovation and solve complex problems in fields such as healthcare, finance, transportation, and beyond. However, it is crucial to address ethical considerations and ensure that machine learning systems are developed and deployed responsibly, with a focus on fairness, transparency, and accountability. |
| **DRAWBACKS: -**   * Performance Limitations * Limited Cross-Platform Support * Limited Advanced Processing Techniques | **ADVANTAGES: -**   * High accuracy * Time process is less * Training high level dataset and efficient. |

**SYSTEM ARCHITECTURE:**

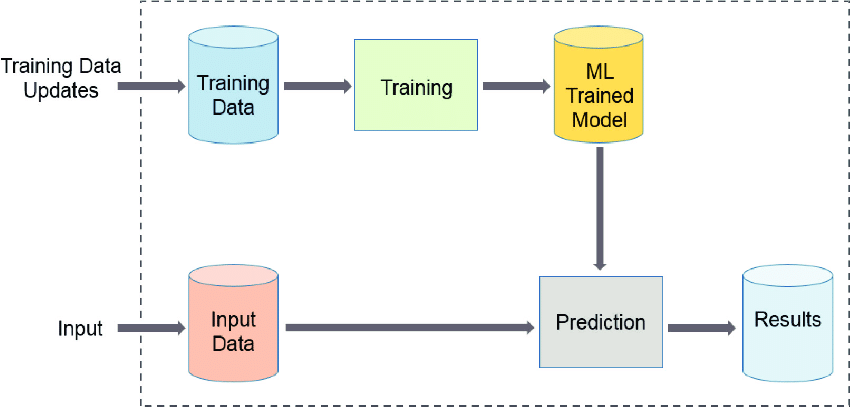


Fig:- proposed model

**MINIMUMSYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS**

* PROCESSOR : Pentium i3 Processor
* RAM : 2GB DD RAM
* HARD DISK : 250 GB

**SOFTWARE REQUIREMENTS**

* BACK END : PYTHON
* OPERATING SYSTEM : WINDOWS 7
* IDE : Spyder3