1. **Develop and Compare Machine Learning Models for Crime Prediction:** The core of this objective is to implement and test multiple machine learning models to predict crime patterns. We can use the provided Chicago crime data (from 2001-2023) and compare models such as **XGBoost, Random Forest, and K-Nearest Neighbors (KNN)**. The goal is to determine which model achieves the highest accuracy and performance metrics (e.g., F1-score, AUC-ROC) for predicting a specific crime type like theft, while also considering how to handle common data issues like class imbalance.
2. **Enhance Crime Forecasting with External Data Integration:** This objective focuses on improving crime prediction accuracy by augmenting traditional crime datasets with external information. We can design a system that integrates data on weather, public transportation, and census information to capture a broader range of factors that influence crime. The project should demonstrate how these additional features, when combined with a model like a hybrid RNN+CNN, can lead to more accurate, fine-grained predictions, such as forecasting crime counts for the next day.
3. **Propose a Real-Time Violence Detection System Architecture:** We should design a system for detecting violent activities from video feeds, with a focus on a practical and efficient architecture. The design would involve using a lightweight deep learning model, such as **MobileNetV2 or CNN-LSTM**, to analyze video frames for violent patterns. A key component of the design is a system for sending real-time alerts to a platform like the Telegram app, including critical details like the time, location, and a captured image of the incident.
4. **Implement a Data-Driven System for Identifying Crime Hotspots:** This objective requires we to build a system that analyzes crime data to identify spatial and temporal crime hotspots. Using a tool like Power BI for visualization, we would perform an exploratory data analysis (EDA) to map the geographical distribution of crimes and identify areas with high crime concentrations. The project would also examine temporal patterns, such as hourly, daily, or seasonal trends, to provide a comprehensive understanding of when and where crimes are most likely to occur.
5. **Design a System for Improving Real-Time Computational Efficiency:** This objective addresses a major challenge in real-time video surveillance: the high computational overhead of processing vast amounts of data. We can propose an innovative system design that employs a **priority-based scheduling algorithm** to intelligently allocate computational resources. The system would prioritize video streams that show a higher probability of a violent event occurring, thereby reducing the workload and moving closer to a truly real-time inference capability. This would be a theoretical design and would not require full implementation.
6. **Evaluate Spatiotemporal Features and Model Interpretability:** This objective involves a deep dive into the features that drive crime prediction models. We can focus on the importance of spatiotemporal features—those related to time and location—for predicting "crimes of opportunity" like theft. The project should go beyond a black-box approach by using an interpretability method like **Shapley Additive Explanations (SHAP)** to explain how specific features contribute to a model's final prediction, demonstrating the value of a transparent and interpretable system.
7. **Conduct a Comparative Analysis of Machine Learning and Deep Learning Architectures:** We should perform a critical evaluation of different model architectures to determine their suitability for specific tasks. The goal is to compare deep learning models (e.g., CNN, RNN) with traditional machine learning methods to show their respective strengths and weaknesses. This analysis can also include an ablation study to measure the impact of external features and spatial components on the models' performance, providing insights and recommendations for future model design.