Project Progress Report

Project Title: Real-Time Crime Forecasting

Date: 20th November 2023

Progress So Far:

Completed Tasks

I have made significant progress in project. I successfully acquired, preprocessed, and analyzed the dataset from the "Real-Time Crime Forecasting Challenge." My completed tasks included:

- Data cleaning: I identified and resolved inconsistencies, missing values, and outliers in the dataset to ensure data quality.
- Data exploration: I performed initial exploratory data analysis (EDA) to gain insights into the distribution of crime data, trends, and patterns.
- Feature engineering: I developed new features to capture temporal and spatial aspects of crime occurrences, which will be crucial for forecasting.

Data:

The dataset I am utilizing for this project can be found at the following link: (https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge-posting#data) and I'm using "NIJ2016_JAN01_JUL31" dataset for this project

Challenges:

Encountered Difficulties:

During the initial stages of the project, I encountered several challenges:

- Data complexity: The dataset contains a wide range of variables, and understanding their relationships required time and effort.
- Imbalanced data: Crime data is inherently imbalanced, with few instances of certain types of crimes, making modeling challenging.
- Model selection: Choosing the most appropriate machine learning algorithms for crime forecasting is still work in progress

How These Challenges were Addressed:

To address these challenges, I took the following steps:

- Data Complexity for Crime Dataset: I regularly engaged in detailed explorations of the crime dataset using Python, utilizing libraries like Pandas and NumPy for in-depth analysis and matplot lib for visualizations, addressing complex data patterns.
- Model selection: Still work in progress

Remaining Challenges:

While I have made substantial progress in overcoming initial obstacles, I still face challenges:

- Model optimization: Fine-tuning hyperparameters to achieve the best predictive performance.
- Interpretability: Interpreting the results and ensuring the model's outputs align with practical law enforcement needs.

I plan to address these challenges through continued research and experimentation.

Next Steps:

Remaining Tasks:

1. Model optimization: Fine-tune hyperparameters and evaluate different model architectures.

Project Progress Report

2. Interpretability: Develop methods to explain model predictions, making them more useful for law enforcement agencies.
3. Finalize the project report: Document my findings, methodology, and results comprehensively.
4. Prepare for the project presentation: Create a clear and engaging presentation to communicate my work effectively.

Plan to Complete Remaining Tasks:

To finalize the remaining tasks with the crime dataset, I will systematically segment the workload into distinct, manageable segments, each with a clear, achievable deadline. This will involve utilizing Python for advanced data processing, applying statistical analyses, and leveraging visualization tools for insights.

Potential Challenges:
I'm anticipating potential challenges in fine-tuning models and ensuring model interpretability. To mitigate these challenges, will seek guidance, review relevant literature.

This extended progress report provides a more detailed overview of project's status, challenges, and plans for the next steps.

Sirisha Kandukuri
kanduksi@mail.gvsu.edu