Predicting Victim Age and Sex Using Crime Data

Problem Statement

The revised problem at hand involves predicting the age and sex of individuals who become victims of crime using crime data and potentially other relevant variables. By analyzing patterns within crime data, we aim to develop predictive models that estimate the age and sex of victims, which can have applications in law enforcement, victim support and aid victim service providers target relevant areas.

Context

Victims of crimes often provide valuable information for law enforcement investigations and victim support services. Predicting the age and sex of victims from crime data can aid law enforcement agencies in identifying potential suspects and allocating resources effectively. Additionally, this information can assist victim support organizations in tailoring their services to meet the specific needs of different demographic groups.

Criteria for Success

The success of this project will be determined by the accuracy of the predictive models in estimating victim age and sex. The model's performance will be evaluated using appropriate metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and F1-score. Achieving a low prediction error will signify the effectiveness and reliability of the model in predicting victim age and sex.

Scope of Solution Space

The solution space involves building predictive models that leverage crime data and potentially other relevant variables to predict victim age and sex accurately. Variables such as the type of crime, location of the crime, time of the crime, and any available contextual information will be considered. The focus will be on developing robust and interpretable models that can be used by law enforcement agencies and victim support organizations to enhance their operations.

Constraints

Data Availability: The quality and availability of crime data may vary by location and jurisdiction. Efforts will be made to obtain the most comprehensive and reliable data possible.

Bias: The goal is to develop predictive models that are equitable and do not reinforce any biases present in the original data.

Stakeholders

- Law Enforcement Agencies: Accurate predictions of victim age and sex can aid in investigations and crime prevention efforts.
- Victim Support Organizations (governmental and nongovernmental): This information can help tailor support services to meet the unique needs of different victim demographics.
- Policy Makers: Insights from the predictive models can inform policies related to crime prevention and victim support.

Data Sources

LAPD crime data containing information about the type of crime, location, time, and any relevant contextual information. Potentially, demographic data from other sources to enhance predictions.

Approach

1. Data Wrangling: Collect, assemble, and organize the victim demographics and crime data obtained from reliable sources. This may involve dealing with inconsistencies and inaccuracies in the data due to its origin from paper-based crime reports.
2. Exploratory Data Analysis: Investigate the data to understand its patterns, distributions, and relationships. Visualize the data through charts and graphs to identify trends and potential outliers.
3. Data Preprocessing: Clean the data by handling missing values, correcting errors, and standardizing formats. Prepare the data for analysis by ensuring it's accurate and complete.
4. Feature Selection: Identify which pieces of information (features) are most relevant for predicting the relevant variable(s). Focus on victim demographics and relevant crime data that could impact the case management duration.
5. Model Selection: Choose an appropriate prediction method based on the characteristics of the data. Consider methods that can capture how victim demographics and crime data relate to the case management duration.
6. Model Training: Train the chosen model using the cleaned and preprocessed data. The model will learn from the relationships between victim demographics, crime data, and case management duration.
7. Model Evaluation: Assess the model's performance by comparing its predictions with the actual details of the victim of crime.
8. Prediction: Utilize the trained model to predict the age and sex of the victim for new cases.
9. Interpretation: Analyze the model's predictions and the factors it considers most important for making those predictions.
10. Documentation: Document the entire process, including data sources, preprocessing steps, model choice, training details, evaluation metrics, and insights gained.

Deliverables

- A well-preprocessed and cleaned dataset for model training and evaluation.
- Predictive models capable of estimating victim age and sex based on crime data.
- Model evaluation results, including classification metrics and potential areas for improvement.
- Insights into the relationship between crime data and victim age and sex.
- A GitHub repo containing the work you complete for each step of the project, including:
    - A slide deck
    - A project report
- Documentation detailing the project's methodology, data sources, and findings for stakeholders and future reference.