ASEC 2024 Conference

The 5th International Electronic Conference on Applied Sciences

04–06 December 2024 | Online

GEDAI MDPI

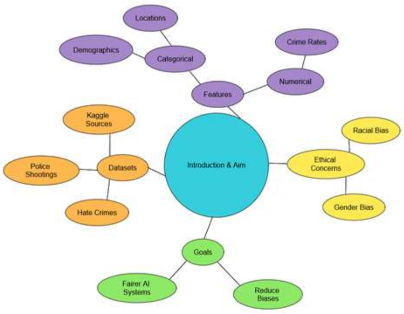
# Ethical Data Engineering for AI in Crime Analysis

Gabriel Souza, Matheus Nascimento, Vagner Silva, Kauã Lima, Ericlécio Araújo, Jean Turet
Federal University of Alagoas
Group of Engineering in Decision-Making and Artificial Intelligence

## INTRODUCTION & AIM

The increasing use of AI in crime analysis has raised concerns about ethical and racial biases in datasets and algorithms. Biases in sensitive attributes, such as race and gender, can lead to unfair outcomes, discrimination, and the marginalization of minority groups. This research focuses on developing an ethical data engineering framework to reduce these biases and ensure fairer AI systems.

Structured datasets on hate crimes and police shootings in the United States, sourced from Kaggle, were selected for this study. These datasets include categorical features (e.g., demographic details, locations) and numerical features (e.g., crime rates), offering a practical basis for evaluating data engineering strategies. The methods developed in this study are intended to pave the way for extending this framework to more complex data types, such as images.



## METHOD

The data engineering framework consists of a structured pipeline designed to address biases at every stage of the data preparation process. The key steps are as follows:

**1. Data Quality Analysis:**

1. Techniques such as outlier detection, correlation analysis, and feature scaling were employed to ensure balanced distributions of sensitive attributes and to minimize distortions in the data.
2. For example, correlation analysis was used to identify relationships between sensitive features (e.g., race or location) and target variables to mitigate bias amplification.

**2. Preprocessing:**

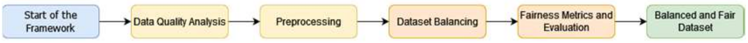
1. Missing values were handled using imputation methods tailored to numerical and categorical data.
2. Incorrect or mislabeled data entries were flagged and corrected.
3. Potentially biased correlations were identified and managed through feature engineering.

**3. Dataset Balancing:**

1. Multiple techniques were applied to address class imbalances in the datasets.
2. Methods included SMOTE (Synthetic Minority Over-sampling Technique), Adaptive Synthetic Sampling, and NearMiss, all of which helped to ensure proportional representation of minority groups in the training data.

**4. Fairness Metrics and Evaluation:**

1. Metrics such as disparate impact and equalized odds were applied to continuously evaluate and refine the fairness of the model outputs.



## RESULTS & DISCUSSION

The results from the preliminary tests underscore the effectiveness of the proposed data engineering framework in mitigating biases in structured datasets used for crime analysis. These tests were conducted using diverse combinations of preprocessing, balancing, and fairness evaluation techniques. Each test was designed to analyze the causal relationship between specific data engineering strategies and their impact on reducing bias in model outputs.

**Detailed Results:**

**1. Reduction in Model Bias:**

1. Models trained on datasets processed with balancing techniques, such as SMOTE, Adaptive Synthetic Sampling, and NearMiss, demonstrated significant improvements in fairness metrics. For example, disparate impact ratios and equalized odds values indicated reduced disparities across sensitive attributes such as race and gender.
2. Handling missing values and correcting label inconsistencies further improved the representativeness of the datasets, contributing to balanced model outputs.

**2. Improvements in Data Representativeness:**

1. By addressing class imbalances, the strategies ensured that minority groups were proportionally represented in the datasets. This led to a noticeable decrease in skewed predictions, which previously favored majority classes.

**3. Fairness Metrics Validation:**

1. The application of fairness metrics, such as disparate impact and equalized odds, allowed a systematic evaluation of how well the models treated various demographic groups. The findings indicated that models processed through the proposed framework had fewer disparities in outcomes, confirming the effectiveness of the methods.

The findings highlight the importance of applying systematic data engineering techniques to structured datasets containing sensitive attributes. Preprocessing steps, such as correlation analysis and feature scaling, minimized the risk of perpetuating biased relationships inherent in the data. Balancing techniques addressed the underrepresentation of minority groups, ensuring a more equitable learning process for AI models.

The causal relationship between balancing sensitive features and improved fairness metrics demonstrates that these methods can address key ethical challenges, including the risk of discrimination and the reinforcement of stereotypes. This framework also provides a scalable approach to integrate fairness into AI systems, especially in public security applications.

Furthermore, the results suggest that ethical data engineering techniques can effectively bridge the gap between technical performance and ethical accountability. Future research should explore how these techniques perform on more complex datasets, such as images, and evaluate their scalability in larger systems. Expanding the use of fairness metrics tailored to specific applications will further enhance the reliability of AI systems in sensitive domains.

## CONCLUSION

The research highlights the importance of ethical data engineering in reducing biases and promoting fairness in AI applications for crime analysis. Implementing data quality checks, preprocessing, balancing techniques, and fairness evaluations showed significant improvements in reducing disparities among sensitive groups. These findings advance the technical and ethical standards of AI systems. Expanding this framework to more complex data types and larger datasets will further validate its utility in real-world applications.

## REFERENCES

1. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society.
2. Banavar, G. (2016). Learning to trust artificial intelligence systems: Accountability, compliance, and ethics in the age of smart machines. Armonk, NY: IBM Research.
3. Bird, S. J., & Housman, D. E. (1995). Trust and the collection, selection, analysis and interpretation of data: A scientist's view. Science and Engineering Ethics.
4. Montomoli, J., Bitondo, M. M., Cascella, M., Rezoagli, E., Romeo, L., Bellini, V., Semeraro, F., Gamberini, E., Frontoni, E., Agnoletti, V., Altini, M., Benanti, P., & Bignami, E. G. (2024). Algor-ethics: Charting the ethical path for AI in critical care. *Journal of Clinical Monitoring and Computing, 38*(4), 931–939.
5. O'Leary, D. E. (2016). Ethics for Big Data and Analytics. *IEEE Intelligent Systems, 31*(4), 81–84.