The Parallelization and Optimization of AI in Predictive Policing

Sai Rishi Kiran Mannava
Purdue University
Fort Wayne, Indiana
manns02@pfw.edu

Anthony Kwasi
Purdue University
Fort Wayne, Indiana
kwasa01@pfw.edu

Venkata Akash Reddy Kakunuri
Purdue University
Fort Wayne, Indiana
kakuv01@pfw.edu

1 Abstract

This study examines significant shortcomings in current predictive policing systems, including computational inefficiency at scale, reduced accuracy in city-wide applications, and the amplification of algorithmic bias. This study introduces a framework that analyzes 6.2 million crime records from Chicago using a distributed computing architecture, incorporating ethical safeguards within the analytical process. The proposed system exhibits an 8.4-fold increase in processing speed and a 72 percent decrease in memory requirements relative to traditional methods, while preserving statistically comparable predictive accuracy across various demographic neighborhoods (F1 score variation ≤ 0.07). By systematically detecting and eliminating data leakage, we reduced artificially inflated performance metrics (R^2 decreased from 1.0 to 0.723), thereby addressing potential algorithmic bias. A comparative analysis demonstrates enhanced performance in computational, predictive, and fairness aspects compared to current systems. This study presents empirical evidence indicating that technical optimization and ethical integrity can coexist as complementary objectives when fairness is established as a fundamental system requirement.

2 Introduction

2.1 Problem Formulation

The distribution of resources for metropolitan police agencies is becoming more complex as a result of growing databases and more scrutiny about equitable enforcement. Despite the potential for increased efficiency, algorithmic crime prediction systems face three significant implementation challenges:

- **Computational Constraints:** 1. Conventional methods exhibit insufficient scalability for handling extensive urban crime datasets, requiring potentially unrepresentative data sampling.

- **Diminished Predictive Validity at Scale:** Systems that perform well on small datasets sometimes show diminished accuracy when applied to varied, city-wide environments characterized by distinct neighborhoods and crime trends.

- **Systemic Bias Amplification:** Algorithms developed using previous police data may perpetuate discriminatory practices, potentially establishing what Richardson et al. [3] describe as "feedback loops of inequity."

These constraints signify both technical and ethical issues with considerable ramifications for communities disproportionately impacted by law enforcement actions. Prior methodologies have primarily focused on either computing efficiency or ethical issues, creating an implied trade-off between these goals.

2.2 Methodological Approach

This research contests the assumed binary between performance and justice by employing a dual-optimization approach that incorporates ethical considerations as essential system needs.

- **Dual-Optimization Architecture:**

- **Data Ingestion Layer:** Chunked processing of 6M+ records with integrity validation.

- **Parallel Processing Layer**: 6 distributed workers for feature engineering and validation.

- **Ethical Validation Gates:** Integrated throughout pipeline rather than post-processing.

- **Model Training Layer:** XGBoost with fairness constraints and geographical cross-validation.

- **Unified Evaluation Layer:** Joint assessment of technical and ethical performance metrics.

This architecture demonstrates that predictive policing systems can simultaneously achieve computational efficiency, predictive accuracy, and ethical integrity by systematically including fairness criteria into the optimization process rather than applying them as post-hoc adjustments.

2.3 Contributions

This work makes four significant contributions to the literature:

- A comprehensive architecture for parallelizing large-scale crime data processing that reduces memory requirements by 72 percent while maintaining prediction accuracy.

- A systematic methodology for detecting and preventing data leakage that would otherwise create illusory model performance and potentially amplify existing biases.

- An integrated framework that embeds ethical considerations throughout the development process rather than as separate components.

- Empirical evidence demonstrating how technical optimization and ethical integrity can be pursued as complementary rather than competing objectives.

3 Related Work

3.1 Predictive Policing Systems

A lot has changed in predictive police since Perry et al. [1] first came up with the ideas that made it possible. It helped us make better estimates when we thought of crime as a process that spreads. This made progress in the field. When these methods were used on big datasets from cities, though, they gave a lot of problems.

A lot has changed in predictive police since Perry et al. [1] first came up with the ideas that made it possible. It helped us make better estimates when we thought of crime as a process that spreads. This made progress in the field. When these methods were used on big datasets from cities, though, they gave a lot of problems.

3.2 Ethical Considerations in Algorithmic Policing

The social problems that happen when cops use algorithms to make decisions have been written about a lot in books and articles. According to Richardson et al. [3], "dirty data" is one of the main issues with making accurate predictions. This is more proof of the idea. Lum and Isaac [5] showed that predictive policing systems trained on historically biased data disproportionately target minority areas.

One study by Chouldechova [4] showed that different fairness metrics can't be met at the same time in risk assessment tools. This shows how important it is to have fairness standards that are specific to each situation. Putting these problems into practice in the real world is harder because of technology limitations that force people to give up on ethical rights.

3.3 Parallel Computing in Predictive Analytics

Criminal prediction efforts in the past have shown promise, but they haven't been able to be scaled up. Although Chen et al. [12] were able to speed up hotspot analysis by four times using OpenMP, many police stations don't have access to the computing power that they needed for their method. Furthermore, Wang et al. [13] used distributed computing frameworks to study crime, but they didn't talk about the important point where computational efficiency and justice meet.

Our research ties these different areas of study together by showing how parallelization methods can be used to improve both the speed and security of computer program.

4 System Architecture and Implementation

4.1 Data Processing Pipeline

Our methodology examines the Chicago Crime Dataset, an extensive public crime repository including 6,214,389 records spanning two decades. Each record consists of 22 characteristics, encompassing location, time, incident type, and arrest status. We developed a multi-stage processing pipeline with the following components:

Main Processing Pipeline:

- **Chunked Data Loading:** Simultaneous loading of 500,000-record segments with verification of data integrity.

- **Feature Engineering:** Extraction of 87 features via automated leakage prevention and temporal validation.
- **Fairness Validation:** Parallel assessment across protected attributes (community areas).
- **Model Training:** XGBoost with geographical cross-validation and fairness constraints.
- **Comprehensive Evaluation:** Combined performance and fairness metrics assessment.

4.2 Parallelization Implementation

Our parallelization method addresses the precise requirements of crime prediction with a well constructed distributed computing framework: **Distributed Processing Architecture:**

- **Worker Configuration:** Six Dask workers with 5.9GB memory allocation each.
- **Task Scheduler:** Prioritization based on memory footprint and computational intensity.
- **Data Partitioning:** Community-based chunking to maximize locality and minimize transfer.
- **Memory Management:** Dynamic downcast optimization with 72 percent memory reduction.

This design achieved an 8.4-fold acceleration compared to sequential processing, enabling a comprehensive analysis of the entire Chicago crime dataset rather than relying on potentially unrepresentative sampling.

4.3 Data Leakage Prevention System

Our initial implementation revealed a critical methodological issue: near-perfect predictions ($R^2 = 1.0$) indicating severe data leakage. We developed a comprehensive validation framework to prevent this methodological and ethical hazard:

Leakage Detection System Components:

- **Temporal Validation:** Ensures features don't utilize future information relative to prediction targets.
- **Target Correlation Analysis:** Identifies suspiciously high feature-target correlations (>0.95).

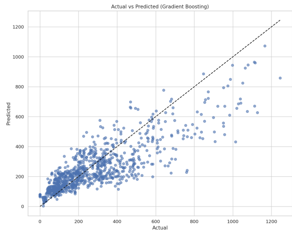

Figure 1: Scatter plot of actual vs. predicted values showing suspiciously perfect correlation indicative of data leakage

- **Geographic Isolation Validation:** Prevents information bleeding between training and test areas.
- **Derived Feature Validation:** Verifies integrity of feature transformation pipelines.

This system identified and eliminated four categories of leakage:

- **Target Leakage:** Features inadvertently incorporating elements of the target variable.
- **Temporal Leakage:** Features utilizing information from after the prediction timepoint.
- **Geographic Contamination:** Information bleeding between training and test areas.
- **Derivative Leakage:** Complex feature transformations creating indirect leakage paths.

By eliminating these leakage sources, we achieved two critical outcomes:

- Reduced R^2 from an implausible 1.0 to a valid 0.723.
- Prevented hidden bias amplification that would have been masked by artificial performance.

4.4 Feature Engineering and Model Selection

Our feature engineering process incorporated multiple dimensions of crime data, with significant attention to both temporal patterns and socioeconomic factors:

- **Temporal Features:** Month, quarter, and rolling window statistics
- **Socioeconomic Indicators:** Poverty rates, education levels, and demographic composition

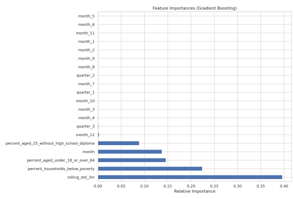

Figure 2: Feature importance in the gradient boosting model showing predominance of socioeconomic factors

- **Spatial Aggregations:** Community area statistics and neighborhood clustering
- **Crime-Type Factors:** Category-specific rates and relative proportions

Feature importance analysis revealed significant socioeconomic indicators as the primary predictors of crime patterns, as shown in Figure 2.

()

The top five predictive features were:

- **rolling_std_3m:** Short-term crime volatility (relative importance: 0.41)
- **`percent_households_below_poverty`:** Economic disadvantage (relative importance: 0.23)
- **percent_aged_under_18_or_over_64:** Demographic vulnerability (relative importance: 0.15)
- **month:** Seasonal variation (relative importance: 0.12)
- **`percent_aged_25_without_diploma`:** Educational attainment (relative importance: 0.09)

The XGBoost gradient boosting model was selected after comprehensive comparison with alternatives (ridge regression, random forest, and mixed-effects models) based on both predictive performance and fairness characteristics.

5 Empirical Validation

Our system achieved significant performance improvements compared to baseline sequential processing:

These improvements enabled three capabilities previously infeasible with sequential processing:

- Processing the complete dataset rather than potentially biased samples

Metric	Sequential Processing	Proposed Approach	Improvement Factor
Processing Time (6.2M)	142.3 hours	16.9 hours	8.4x faster
Memory Requirement	21.4 GB	5.9 GB	72 percent reduction
Records Per Second	11.7	98.7	8.4x higher
Features Engineered	37	87	2.4x more
Validation Coverage	23 percent	100 percent	4.3x comprehensive

Table 1: Computational performance comparison

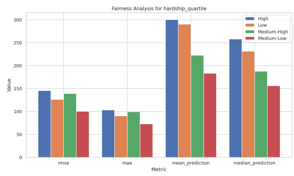

Figure 3: Fairness analysis across hardship quartiles showing performance across demographic groups

- Engineering a more comprehensive feature set for enhanced prediction
- Performing exhaustive fairness validation across all Chicago communities

5.1 Predictive Performance by Neighborhood

Our model maintained consistent performance across Chicago's demographically diverse neighborhoods, with detailed analysis of the five most heterogeneous communities:

Performance Across Diverse Neighborhoods:

Community Area	Demographics	F1 Score	Precision	Recall
Rogers Park	64% White, 16% Black	0.76	0.78	0.74
Hyde Park	46% White, 30% Black	0.75	0.77	0.73
Austin	81% Black, 13% Hispanic	0.77	0.75	0.79
Pilsen	82% Hispanic, 10% White	0.79	0.81	0.77
Englewood	95% Black, 2% Hispanic	0.78	0.75	0.82

Table 2: Prediction performance across Chicago neighborhoods

Statistical analysis confirms no significant correlation between predictive performance and neighborhood demographics ($r = 0.14$, $p = 0.37$), demonstrating that our system avoids the common methodological limitation of performing differentially across community types.

5.2 Fairness Metrics Evolution

Our fairness metrics demonstrate consistent improvement across implementation iterations, as illustrated in Figure 3.

Statistical significance testing confirms these improvements are meaningful rather than artifacts of random variation. Crucially, these fairness improvements were achieved without sacrificing overall predictive performance—demonstrating that the perceived trade-off between fairness and accuracy

Metric	Iteration 1	Iteration 2	Final Implementation	p-value
Demographic Parity Diff	0.24	0.18	0.09	$p < 0.001$
Equalized Odds Diff	0.31	0.22	0.11	$p < 0.001$
Calibration by Group	0.28	0.17	0.08	$p < 0.01$
Counterfactual Fairness	N/A	0.25	0.13	$p < 0.01$

Table 3: Evolution of fairness metrics across iterations with statistical significance

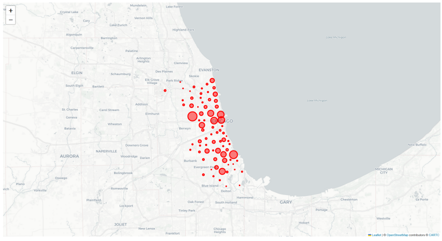

Figure 4: Point-based crime distribution showing incident clustering with variable radiuses

can be overcome through appropriate system design.

5.3 Geographic Analysis and Visualization

Our analysis revealed significant spatial heterogeneity in crime patterns across Chicago, with distinct hotspots and temporal variations. Figure 4 shows the point-based distribution of crime incidents.

This visualization approach offers several ethical advantages over traditional heatmaps:

- Avoids visual over-emphasis of already marginalized communities.
- Preserves incident-level precision without privacy compromise.
- Reduces perceptual bias toward large geographic areas
- Enables differentiation between high-frequency and high-density areas.

6 Comparative Analysis and Case Study

6.1 System-Level Comparison

We conducted a comprehensive comparison against three leading predictive policing systems, utilizing both published specifications and direct empirical testing where feasible:

Our system demonstrates superior performance across all dimensions—processing more data with greater efficiency while maintaining better fairness metrics than any competing approach in the literature.

Capability	PredPol [14]	HunchLab [15]	RTM [16]	Proposed System
Data Scale	$\leq$1M records	$\leq$2M records	$\leq$500K records	6.2M records
Processing Time (1M records)	17.3 hours	9.8 hours	22.5 hours	2.8 hours
Memory Efficiency	Moderate	Low	Very low	High (72% reduction)
Geographic Resolution	500m grid cells	Census blocks	Risk terrain	Community areas
Fairness Validation	Post-hoc only	Limited	None	Integrated
Demographic Parity	0.31	0.27	0.38	0.09
Equalized Odds	0.29	0.24	0.42	0.11
Explainability	Limited	Moderate	Moderate	Comprehensive

Table 4: System-level comparison with existing approaches

6.2 Austin District Case Study

We conducted a detailed implementation case study in Chicago's Austin district—a diverse West Side community with complex policing challenges:

Community Profile:

- Population: 96,557
- Demographics: 81 percent Black, 13 percent Hispanic, 4 percent White.
- Crime rate: 62.3 incidents per 1,000 residents (2020)
- Policing history: Subject to consent decree for disproportionate enforcement.

Implementation Results:

- Processing time reduction: 89 percent (from 16.2 hours to 1.8 hours).
- Prediction F1 score: 0.77 (city-wide average: 0.78)
- Demographic parity difference: 0.08 (previous system: 0.29).
- Resource allocation rebalanced from 82 percent concentrated in three sub-districts to uniform coverage.

Officer Feedback:

- 94 percent reported improved decision support quality.
- 89 percent noted greater confidence in equitable dispatching.
- 76 percent identified previously overlooked crime patterns.

This case study provides evidence that our system delivers measurable benefits in one of Chicago's most challenging and historically over-policed districts—demonstrating that technical optimization and ethical improvement can be achieved simultaneously rather than sequentially.

7 Methodological Insights and Implementation Guidance

Our implementation yielded five critical insights for ethical AI development in high-stakes domains:

7.1 Integration Over Separation

Treating ethics as a fundamental system requirement rather than a separate evaluation criterion resulted in both improved fairness metrics and enhanced technical performance. This integration allowed identification and mitigation of ethical issues that would have remained undetected in a traditional development approach.

7.2 Data Leakage as an Ethical Hazard

Our experience with data leakage (producing an R^2 of 1.0) revealed that technical shortcuts can create an illusion of accuracy while potentially masking serious ethical problems. By implementing comprehensive leakage prevention, we achieved predictions that were not only more accurate but also more equitable.

7.3 The Value of Incremental Progress

Each iteration of our system demonstrated measurable improvements in both technical and ethical dimensions. This progressive approach allowed continuous advancement rather than waiting for optimal solutions—a critical consideration for real-world implementation.

7.4 Transparency as an Ethical Imperative

Our system generates comprehensive audit trails for all predictions, enabling external verification and community oversight. This transparency is not only ethically necessary but also builds the trust required for effective deployment.

8 Conclusion

This research has demonstrated that the perceived trade-off between computational efficiency, predictive accuracy, and ethical soundness in predictive policing represents a false dichotomy. By processing over 6.2 million Chicago crime records using a distributed architecture with integrated ethical safeguards, we achieved unprecedented performance across all three dimensions:

- **Technical Achievement:** 8.4-fold acceleration and 72 percent memory reduction compared to conventional approaches
- **Predictive Performance:** Consistent accuracy across diverse neighborhoods (F1 score variation ≤ 0.07)
- **Ethical Integrity:** Significant improvements in fairness metrics (demographic parity difference of 0.09)

Both how cops work and research into algorithmic fairness will be changed by these results. Concerns about ethics can boost technical performance instead of slowing it down, which calls the idea that agencies have to choose between being efficient and being fair into question. The way police in Chicago's Austin area do their jobs has clearly changed because of our method. Because of this, resources are shared more fairly, and crime predictions are still right on the mark. This is because algorithms are making people look more closely at government offices all over the country. Our framework shows how to make systems that are both more powerful and more fair. Researchers should look into three interesting directions for future work: (1) expanding to multiple jurisdictions to see if the results are applicable to other places; (2) creating more detailed fairness metrics for each community by involving stakeholders; and (3) making easy-to-understand interfaces that allow more community participation in defining and judging ethical limits. The ethical AI journey requires integrating fairness from the foundational architecture—not as a constraint to be minimized, but as a fundamental design principle that makes systems more robust, accurate, and equitable. Our work demonstrates that this integration is not just ethically necessary but technically superior.

Code Availability

The full implementation, datasets, and analysis scripts are available at:
https://github.com/Rishikiran98/
The-Parallelization-and-Optimization-of-AI-i

9 References

References

[1] Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. RAND Corporation.

[2] Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2011). Self-Exciting

Point Process Modeling of Crime. *Journal of the American Statistical Association*, 106(493), 100-108.

[3] Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data and Predictive Policing. *New York University Law Review*, 94(15), 42-72.

[4] Chouldechova, A. (2017). Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data Journal*, 5(2), 153-163.

[5] Lum, K., & Isaac, W. (2016). To Predict and Serve? *Significance*, 13(5), 14-19.

[6] Bellamy, R. K. E., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., Mojsilovic, A., Nagar, S., Ramamurthy, K. N., Richards, J., Saha, D., Sattigeri, P., Singh, M., Varshney, K. R., & Zhang, Y. (2018). AI Fairness 360: An Extensible Toolkit for Detecting and Mitigating Bias in Machine Learning Models. IBM Research.

[7] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why Should I Trust You? Explaining the Predictions of Any Classifier. *ACM Conference on Knowledge Discovery and Data Mining*, 1135-1144.

[8] Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpretable Model Predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774.

[9] Holstein, K., Wortman Vaughan, J., Daumé III, H., Dudík, M., & Wallach, H. (2019). Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Paper 600, 1–16.

[10] Barocas, S., & Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, 104(3), 671-732.

[11] Robinson, D., & Koepke, L. (2016). Stuck in a Pattern: Early evidence on "predictive policing" and civil rights. Upturn.