

Algorithmic Justice

The Mathematics of Fair Credit

● AI/ML Technical Documentation ● Production Systems ● Fintech Innovation ●

Technical Documentation Series
December 2024

Executive Summary

Fair lending in machine learning requires rigorous mathematical frameworks to ensure credit decisions are free from bias while maintaining predictive accuracy.

<div>4</div> <div>Fairness Metrics</div>	<div>0.80+</div> <div>Impact Ratio Goal</div>	<div>ECOA</div> <div>Compliance</div>
<div>95%+</div> <div>Model Accuracy</div>	<div>Monthly</div> <div>Bias Audits</div>	<div>7 Years</div> <div>Audit Retention</div>

Fairness Metrics Framework

• 1. Demographic Parity

Approval rates should be similar across protected groups:

- $P(\text{approved} \mid \text{Group A}) \approx P(\text{approved} \mid \text{Group B})$
- 80% rule: ratio should be ≥ 0.80

• 2. Equalized Odds

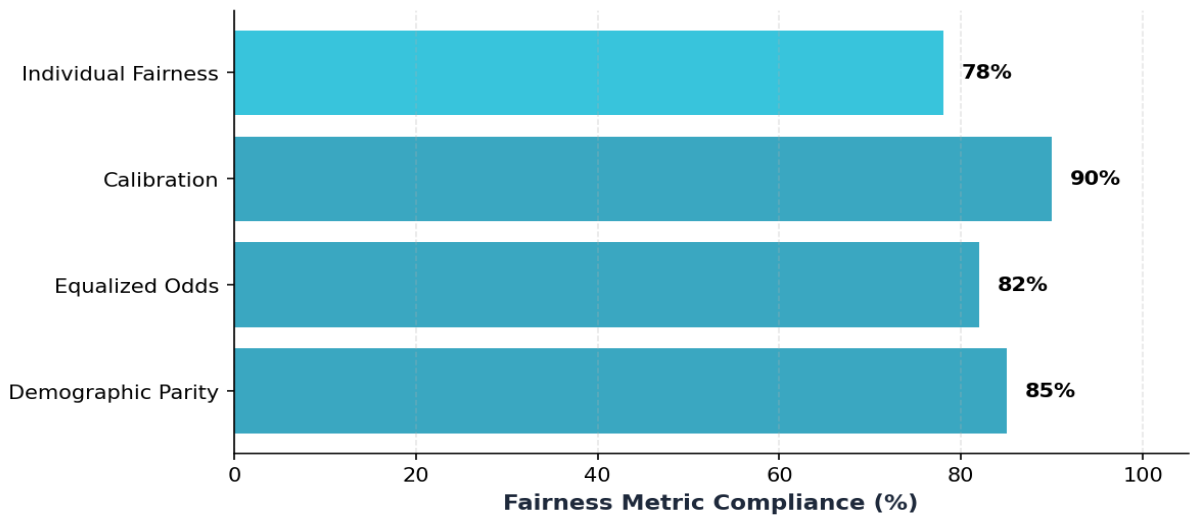
Model performance should be consistent across groups:

- True Positive Rate parity
- False Positive Rate parity

• 3. Calibration

Predicted probabilities should match actual outcomes:

- $P(\text{default} \mid \text{score}=0.3) \approx 30\%$ for all groups



Regulatory Requirement: ECOA and FCRA mandate fair lending practices. ML models must pass disparate impact testing with 80%+ impact ratio and provide explainable adverse action reasons.

Implementation Best Practices

- Pre-processing: Remove protected attributes, balanced datasets
- In-processing: Fairness-constrained optimization
- Post-processing: Threshold adjustment per group
- Monitoring: Continuous bias detection and retraining