

Fairness and Bias Auditing in Financial Machine Learning Models

Introduction

Algorithmic decision-making systems are increasingly used in the financial sector for credit scoring and loan approvals. While these systems promise efficiency and objectivity, they can inadvertently propagate or amplify existing biases present in historical data. This is especially critical when decisions disproportionately affect disadvantaged or protected groups, such as racial minorities or women.

Algorithmic Bias in Finance

Bias in financial ML models typically arises from:

- Historical data reflecting societal inequalities.
- Feature selection that correlates with protected attributes.
- Optimization objectives focusing solely on predictive accuracy without fairness constraints.

When left unchecked, biased models may unfairly deny loans or assign worse credit scores to certain groups, leading to systemic discrimination.

Fairness Metrics

To address bias, several fairness definitions are used:

- **Demographic Parity:** The decision outcome (e.g., loan approval) should be independent of the protected attribute. Formally, $P(\hat{Y} = 1|A = 0) = P(\hat{Y} = 1|A = 1)$.
- **Equal Opportunity:** The true positive rate should be equal across groups, i.e., $P(\hat{Y} = 1|Y = 1, A = 0) = P(\hat{Y} = 1|Y = 1, A = 1)$.
- **Disparate Impact:** The ratio of positive outcomes between groups should be at least 0.8 (the "80% rule" used in legal contexts).

Bias Auditing Algorithm

A typical bias auditing procedure involves:

1. Train the model on the training data including a protected attribute.
2. Evaluate group-wise performance metrics (e.g., accuracy, precision, recall).
3. Compute fairness metrics (demographic parity difference, equal opportunity difference, disparate impact).
4. If fairness violations are detected, consider mitigation strategies such as re-weighting, adversarial debiasing, or post-processing adjustments.

Conclusion

Continuous auditing of financial ML systems is essential to ensure fair treatment of individuals and compliance with ethical and legal standards. Such frameworks not only improve trust but also mitigate reputational and regulatory risks.