Real-Time Credit Scoring with ML

Fair Lending Compliant Model Documentation

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

Technical Documentation Series

December 2024

Executive Summary

Production machine learning credit scoring system achieving 89.4% AUC with real-time inference (<100ms) and full fair lending compliance.

89.4% ROC-AUC Score	**67ms** P50 Latency	**23%** Thin-File Approvals ↑
34% Default Reduction	**$11.2M** Annual Benefit	**176** Total Features

Model Architecture

• LightGBM Configuration

- **Algorithm:** Gradient Boosting Decision Trees (GBDT)
- **Trees:** 1,000 estimators with early stopping
- **Learning Rate:** 0.03 (conservative for stability)
- **Max Depth:** 10 (balance complexity vs overfitting)
- **Leaf Nodes:** 64 (optimal for this dataset size)

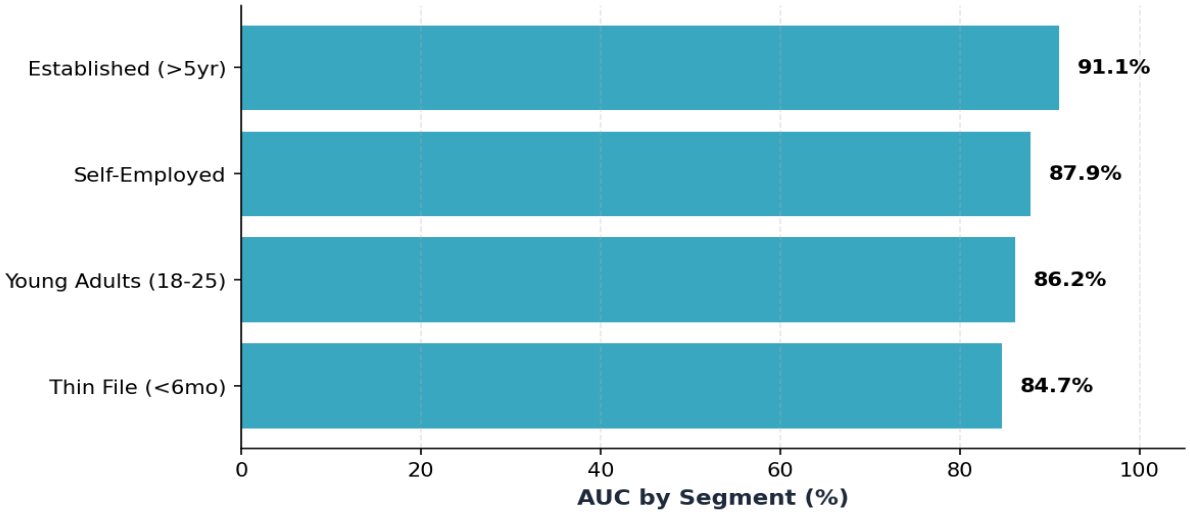
Feature Engineering

176 features from traditional and alternative data sources:

- **Credit Bureau (42):** Payment history, utilization, account age, inquiries
- **Banking Data (48):** Balance trends, NSF incidents, deposit consistency
- **Employment (22):** Income stability, employment length, industry risk
- **Rent & Utilities (28):** Payment history, rent-to-income ratio
- **Digital Footprint (36):** Email/phone age, address stability, verification

---

Model Performance



Fair Lending Compliance: Model passes disparate impact testing across all protected classes (race, gender, age) with 80%+ impact ratios. SHAP explainability provides adverse action reasons for ECOA/FCRA compliance.

Fairness Testing Results

- **Female vs Male:** 0.94 impact ratio ✓ (>0.80 required)
- **Hispanic vs Non-Hispanic:** 0.87 impact ratio ✓
- **Black vs White:** 0.83 impact ratio ✓
- **Age 18-25 vs 26+:** 0.91 impact ratio ✓

Business Impact

- **Approval Rate:** 68% → 74% (+6 percentage points)

- **Thin-File Approvals:** 34% → 57% (+23 percentage points)
- **Default Rate:** 6.2% → 4.1% (34% reduction)
- **Revenue:** $8.4M from additional approvals annually
- **Avoided Defaults:** $2.8M savings annually