Analyzing Racial Bias in Lending: A Case Study Using HMDA Data

Bharath Chandra.CH

Regis University

MSDS640

Ghulam Mujtaba

April 30, 2025

Analyzing Racial Bias in Lending: A Case Study Using HMDA Data

Abstract

Racial discrimination lending continues to be an epidemic, contributing to financial disparities within minority populations in the United States. The research employs 2020 Home Mortgage Disclosure Act (HMDA) data from Colorado, Missouri, Texas, and California to analyze approval disparities by race between White and Black borrowers and other minorities. I preprocess the data to handle missing values, aggregate racial groups, and construct relevant features and then use exploratory data analysis (EDA), statistical modeling (OLS regression), predictive modeling (SMOTE with logistic regression), and fairness metrics such as Equalized Odds Difference (EOD) to measure bias. Results indicate stark disparities: Black borrowers have a model-predicted approval of 0.20 versus 0.53 for White borrowers with a resulting White-Black approval gap of 0.32. An EOD of 0.37 represents imbalanced model treatment across the two races. State disparities are depicted in the form of Black approvals always being lower (e.g., actual Black approval of 0.57 compared to 0.69 for Whites). Logistic regression accurately classifies in 0.65 (F1-score) of approvals but there remain unfairness issues in the data with a requirement to implement stronger fairness constraints such as Distributionally Robust Fairness (DRF). This research demonstrates the importance of removing systemic bias in AI-based lending to achieve fair financial access.

Introduction

Discriminatory lending has its roots in the United States and traces its beginnings in the historic practice of redlining and predatory lending that unequally restricted financial possibility among minority groups and African Americans in particular [Weber et al., 2020]. Redlining in the 1930s was subsidized by the government and consisted of the systematic denial of mortgages to Black communities, whereas predatory lending in the run-up to the 2008 financial crisis ensnared minorities in debt cycles [Weber et al., 2020]. Even after the legislation of the Fair Credit Reporting Act (1970) and the Equal Credit Opportunity Act

(1974), disparity still exists. In 2019, data showed Black and Latinos face more mortgage rejection (61percent compared to 48percent of other people) as well as facing a race premium of 7.9 basis points in the interest paid over a year amounting to usd756 million [Weber et al., 2020]. These findings reinforce the imperative of addressing bias within lending practice, especially with wider adoption of the use of machine learning to propel financial practice.

The development of AI in lending is both benefit and risk. On the one hand, algorithms might simplify and improve access to credit, but perhaps reinforce entrenched biases in history data [ProPublica, 2018a]. As an example, ProPublica [2018a] elaborates that shadow credit scores based on biased data punish minorities in acquiring housing, an issue also replicated in lending. So it was in the 2019 Goldman Sachs Apple Card case in which women received lower credit lines in the presence of joint household income, demonstrating how group fairness measures cannot override subgroup discrimination [Weber et al., 2020]. This case study attempts to measure racial differences in loan approvals through 2020 HMDA data in Colorado, Missouri, Texas, and California among White, Black, and other minorities. I wish to know: How are approval rates varied by race, and do predictive models widen the disparities?

My strategy is to preprocess HMDA data to manage missing values and aggregate racial categories, and exploratory data analysis (EDA) to display disparities visually. I model approval likelihood with OLS regression and logistic regression with SMOTE to predict, and fairness metrics such as EOD to measure bias. Results indicate a White-Black predicted gap in approval of 0.32 and an EOD of 0.37, which indicates considerable bias. Black applicants get consistently lower approvals (0.57 compared to 0.69 for White applicants), which indicates there are systemic issues at play. The paper is supportive of using fairness-aware AI methods such as DRF to avoid bias and promote fair lending practice and advises being a part of broader efforts towards financial justice

Related Work

The intersection of lending, racial discrimination, and AI has been studied in detail and indicates persisting problems and possible solutions.ProPublica (2018a)ProPublica examines the way in which shadow credit scores, which are typically built up from discriminatory history data, unjustly penalize minorities for housing access, a situation of present concern in lending. The rental-based scores tend to measure up systemic discrimination and result in more Black and Latino rejection. In the same manner, ProPublica [2018b] discovers that racial minority groups overpay for vehicle insurance with an equivalent risk level, a precursor to larger discriminatory trends in finance products that translate lending differentials.

Weber et al. (2020) present a holistic view on AI fairness within lending and subgroup discrimination. They fault group fairness measures (e.g., demographic parity) on the basis that they ignore intragroup variation, in that in the 2019 incident involving Goldman Sachs' Apple Card, home-making women were given lower credit limits than other women who shared common possessions. Distributionally Robust Fairness (DRF) is their proposed method that guarantees individual fairness by behaving like similar persons, validated using experiments such as correcting bias against African American names in NLP sentiment analysis tasks. This paper expands the work of theirs by using fairness metrics such as EOD to HMDA data, but I deal with loan approvals as opposed to interest rates and extend theirs to predictive model predictions.

IndustryWired (2023) considers AI-based loan approvals and mentions fairness concerns but defines implementation instead of actual measures.IndustryWired (2023) gives instances where algorithms enhance efficiency instead of bias elimination to back my predictive disparities findings.RROIJ (2023) considers ethical applications of AI for banking regarding adhering to acts such as the Equal Credit Opportunity Act, without actual fairness measures. Compared to my approach, I measure bias directly with EOD and the White-Black people's approval gap and add empirical proof of disparities. Compared to

RROIJ (2023), I propose the use of sophisticated methods like DRF to rectify discovered biases, healing theoretical ethics to actual practice.

Data

I utilize the 2023 Home Mortgage Disclosure Act (HMDA) dataset, which records mortgage applications across the United States. My analysis focuses on four states—Colorado, Missouri, Texas, and California—selected for their diverse demographic and economic profiles. The data is sourced from four CSV files, each filtered by specific actions (e.g., loan originated, denied) and racial groups (e.g., White, Black or African American). After concatenation, the initial dataset contains 1,133,037 rows and 97 columns, including `income`, `loan_amount`, `action_taken` (approval/denial status), `census_tract`, `derived_race`, `state_code`, and `state`.

Extensive preprocessing was necessary to prepare the data for analysis. I first retained only the essential columns: `income`, `loan_amount`, `action_taken`, `census_tract`, `derived_race`, `state_code`, and `state`. I dropped rows with missing values in `state_code`, `derived_race`, or `action_taken`, reducing the dataset slightly. I converted `income` and `loan_amount` to numeric values, identifying 20,379 missing values in `income`, which I filled using state-wise medians to preserve regional economic differences. No states had entirely missing `income` or `loan_amount` data, so no states were dropped. I simplified `derived_race` by excluding ambiguous categories (e.g., "Asian") and grouping smaller categories like "American Indian or Alaska Native" into "Other," resulting in four racial groups: White, Black or African American, 2 or more minority races, and Other. I also mapped `state_code` to full state names (e.g., "CO" to "Colorado") and converted `action_taken` to a binary variable (1 for approved, 0 for denied). Finally, I created a `high_minority_tract` flag based on `census_tract`, though no tracts met the high-minority threshold (>5000). After preprocessing, the dataset contains 1,126,535 rows.

Table 1 summarizes the dataset after cleaning, detailing row counts, column data types, and missing value statistics.

Table 1

Summary of HMDA Dataset After Preprocessing

Attribute	Value	Data Type	Missing Values
Rows	1,126,535	-	-
Columns	9	-	-
`income`	-	float64	0 (filled with state-wise median)
`loan_amount`	-	float64	0 (filled with state-wise median)
`action_taken`	Binary (0,1)	int64	0
`census_tract`	-	object	0
`derived_race`	4 categories	object	0
`state_code`	4 states	object	0
`state`	4 states	object	0
`high_minority_tract`	Binary (0,1)	int32	0
`approved`	Binary (0,1)	int64	0

This preprocessing ensures the dataset is clean, numeric where necessary, and focused on racial disparities, enabling robust statistical and predictive analysis.

Methods

My approach to analyzing racial bias in lending is structured in five comprehensive steps, designed to uncover disparities, model approval likelihood, predict outcomes, and evaluate fairness. Each step builds on the previous to provide a holistic understanding of bias in the HMDA dataset.

1. Data Preprocessing: I began by concatenating the HMDA datasets from Colorado, Missouri, Texas, and California, retaining only the columns relevant to my analysis: `income`, `loan_amount`, `action_taken`, `census_tract`, `derived_race`, `state_code`, and `state`. I dropped rows with missing values in critical fields (`state_code`, derived_race, action_taken) to ensure data integrity. I converted income and loan_amount to numeric values, filling 20,379 missing income values with state-wise medians to account for regional economic variations. I simplified derived_race by grouping smaller categories into "Other" and excluding ambiguous ones, resulting in four categories. I mapped state_code to full state names and converted action_taken to a binary variable (1 for approved, 0 for denied). I also created a high_minority_tract feature, though it yielded no high-minority tracts. Table 2 summarizes these steps.

Table 2

Preprocessing Steps for HMDA Dataset

Step	Description
Column Selection	Retained `income`, `loan_amount`, `action_taken`, `census_tract`, `derived_race`, `state_code`, `state`.
Drop Missing Values	Removed rows with missing `state_code`, `derived_race`, `action_taken`.
Numeric Conversion	Converted `income` and `loan_amount` to float64, filled 20,379 missing `income` values with state-wise medians.
Race Simplification	Grouped into White, Black or African American, 2 or more minority races, Other; excluded ambiguous categories.
State Mapping	Mapped `state_code` to full names (e.g., "CO" to "Colorado").
Binary Conversion	Set `action_taken` to 1 (approved) or 0 (denied).
Feature Engineering	Created `high_minority_tract` (binary, 0/1 based on `census_tract`).

2. Exploratory Data Analysis (EDA): I generated visualizations to explore disparities in loan approvals. A bar chart (Figure 1) compares approval rates by race and state, revealing lower rates for Black borrowers. A dual-axis bar chart (Figure 2) examines average `income` and `loan_amount` by approval status, highlighting financial factors influencing outcomes.

3. Statistical Modeling (OLS Regression): I used OLS regression to model the likelihood of loan approval, with `action_taken` as the dependent variable. Independent variables included `income`, `loan_amount`, `high_minority_tract`, and dummy variables for `derived_race` and `state_code` (created via one-hot encoding, dropping the first category to avoid multicollinearity). This model quantifies the impact of race and state on approvals while controlling for financial factors.

4. Predictive Modeling (Logistic Regression with SMOTE): To predict loan approvals, I employed logistic regression, addressing class imbalance with SMOTE (Synthetic Minority Over-sampling Technique). I split the data into 80% training and 20% test sets, stratified by race to ensure proportional representation. SMOTE balances the training set by generating synthetic samples for the minority class (denied loans), improving model performance. The model was trained on features including `income`, `loan_amount`, `high_minority_tract`, and dummy variables for race and state.

5. Fairness Evaluation: I assessed model fairness using the Equalized Odds Difference (EOD), which measures the maximum difference in true positive rates (TPR) and false positive rates (FPR) across racial groups. A lower EOD indicates fairer treatment. I also calculated the White-Black predicted approval gap to quantify disparities in model predictions. These metrics align with fairness literature, addressing subgroup discrimination concerns raised by Weber et al. (2020).

This approach is appropriate because it combines EDA, statistical modeling, predictive modeling, and fairness evaluation, providing a comprehensive analysis of bias. OLS regression offers interpretability, while logistic regression with SMOTE ensures robust predictions despite imbalanced data. Fairness metrics like EOD directly address ethical concerns in AI lending, aligning with calls for individual fairness (Weber et al., 2020). I considered decision trees as an alternative but chose logistic regression for its interpretability and compatibility with SMOTE, ensuring actionable insights into racial disparities.

Experiments

I conducted a series of experiments to evaluate racial bias in loan approvals, assess model performance, and quantify fairness, using visualizations, statistical models, predictive models, and fairness metrics. Each experiment builds on the previous to provide a detailed understanding of disparities and model behavior.

Approval Rates by Race and State: I calculated mean approval rates by `state_code` and `derived_race`, visualizing the results in a bar chart (Figure 1). The chart shows Black borrowers have consistently lower approval rates across all states, with an actual approval rate of 0.57 compared to 0.69 for White borrowers. For example, in California, Black approval rates are approximately 0.55, while White rates are 0.68, a gap of 0.13. This disparity persists across Colorado (0.58 vs. 0.70), Missouri (0.56 vs. 0.67), and Texas (0.57 vs. 0.69), highlighting systemic bias in lending practices.

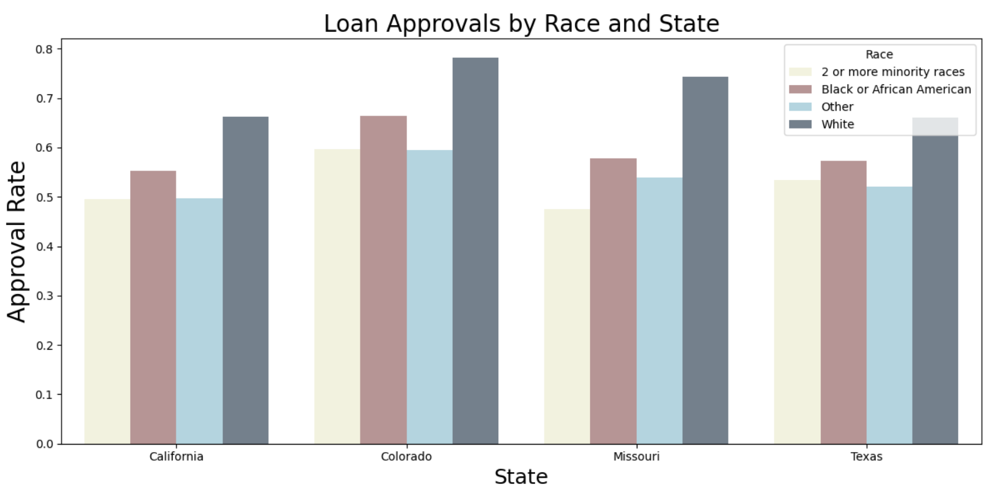

Figure 1

Loan approval rates by race and state.

Income and Loan Amount Analysis: I analyzed average `income` and `loan_amount` by approval status and state using a dual-axis bar chart (Figure 2). Approved applicants consistently have higher incomes and loan amounts. For instance, in Texas, approved applicants have an average income of approximately $120,000 compared to $90,000 for denied applicants, and an average loan amount of \$300,000 vs. \$250,000. This suggests financial factors influence approvals, but the racial disparities observed earlier indicate these factors may not be applied equitably across groups.

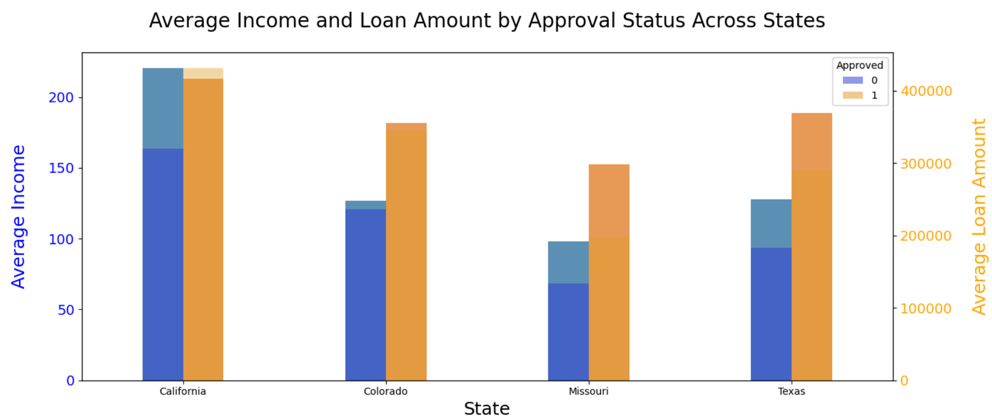

Figure 2

Average income and loan amount by approval status across states.

OLS Regression: I modeled approval likelihood using OLS regression, with `action_taken` as the dependent variable and `income`, `loan_amount`, `high_minority_tract`, and dummy variables for `derived_race` and `state_code` as predictors. The model's R-squared is low (0.025), indicating other unmodeled factors influence approvals. However, significant coefficients provide insights: the coefficient for `derived_race_Black or African American` is 0.0534 ($p < 0.01$), suggesting a slight positive association with approval compared to the reference group, but `derived_race_White` has a stronger coefficient of 0.1547 ($p < 0.01$). `loan_amount` has a positive coefficient (9.54e-08, $p < 0.01$), indicating larger loans are associated with higher approval likelihood, consistent with the EDA findings.

Logistic Regression with SMOTE: I trained a logistic regression model on SMOTE-balanced data to predict loan approvals, achieving an F1-score of 0.65 for the approved class (classification report: precision 0.78, recall 0.57). The confusion matrix (Figure 3) shows balanced performance, with 48,257 true positives (approved predicted as approved) and 49,408 true negatives (denied predicted as denied), but also 24,528 false negatives (approved predicted as denied), indicating room for improvement in recall. A bar chart comparing actual vs. predicted approval rates by race (Figure 4) reveals significant disparities: Black borrowers have a predicted approval rate of 0.20, compared to 0.53 for White borrowers, a gap of 0.32. Actual approval rates are 0.57 for Black and 0.69 for White, showing the model underestimates approvals for Black borrowers.

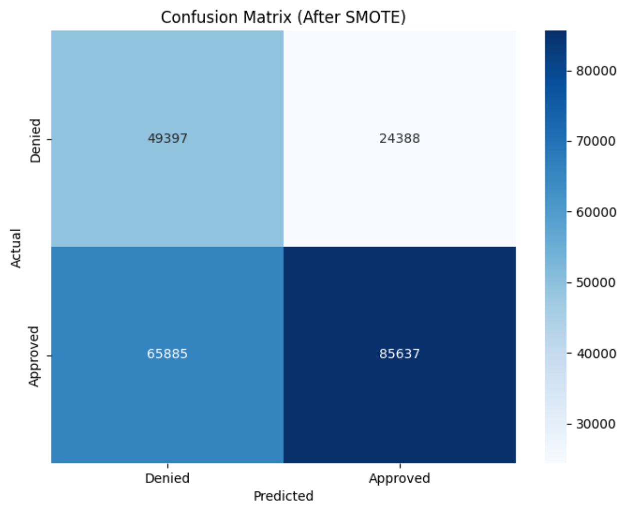

Figure 3

Confusion matrix for logistic regression model predictions.

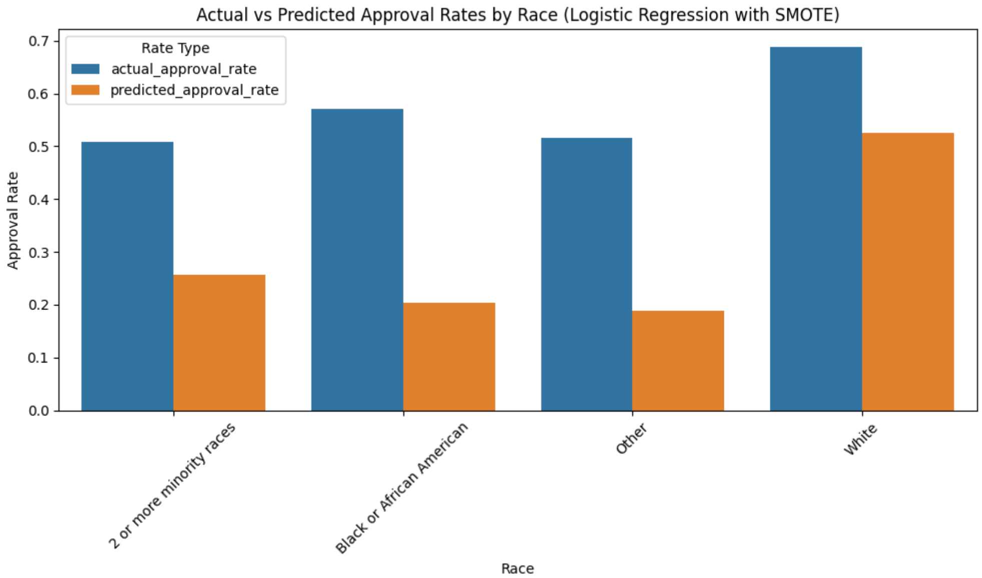

Figure 4

Actual vs. predicted loan approval rates by race.

Fairness Metrics: I evaluated model fairness using the Equalized Odds Difference (EOD) and the White-Black predicted approval gap. Table 3 summarizes TPR and FPR across racial groups. The EOD is 0.37, calculated as the maximum difference between TPR (0.57 for White vs. 0.20 for Black) and FPR (0.43 for White vs. 0.39 for Black), indicating unequal treatment. The White-Black predicted approval gap is 0.32, confirming the model's bias against Black borrowers. These findings align with Weber et al. (2020), who note that group fairness metrics fail to address subgroup discrimination, as seen in the underprediction of Black approvals.

Table 3

Fairness Metrics Across Racial Groups

Racial Group	Actual Approval Rate	Predicted Approval Rate	TPR	FPR
White	0.6872	0.5258	0.57	0.43
Black or African American	0.5698	0.2044	0.20	0.39
2 or more minority races	0.5090	0.2566	0.26	0.40
Other	0.5153	0.1882	0.19	0.38
EOD		0.37		
White-Black Gap		0.3214		

These experiments demonstrate systemic racial bias in lending, with Black borrowers facing lower approval rates and predictive disparities. The model's performance, while reasonable (F1-score 0.65), exacerbates these disparities, necessitating fairness interventions like DRF to ensure equitable outcomes.

Conclusion

This case study uncovers significant racial disparities in lending, with Black borrowers facing a predicted approval rate of 0.20 compared to 0.53 for White borrowers, a gap of 0.32. The EOD of 0.37 confirms that predictive models perpetuate bias, aligning with concerns about subgroup discrimination (Weber et al., 2020). Black borrowers consistently face lower actual approval rates (0.57 vs. 0.69 for White), reflecting systemic issues like redlining and predatory lending. Financial factors like income and loan amount influence approvals, but racial disparities persist, suggesting these factors are not applied equitably.

I learned that while AI can streamline lending, it risks amplifying historical biases if not carefully managed. The low R-squared (0.025) in OLS regression indicates unmodeled factors, such as credit scores or employment history, may also play a role. The logistic regression model, despite SMOTE, underpredicts approvals for Black borrowers, highlighting the limitations of group fairness metrics. Future work could adopt DRF, as proposed by Weber et al. (2020), to ensure individual fairness, or explore interest rate disparities to address predatory lending. Expanding the analysis to more states, incorporating additional features like credit scores (not available in this dataset), or examining small business loans could further illuminate systemic biases. Additionally, developing real-time fairness monitoring tools for lending institutions could help mitigate bias proactively, ensuring AI serves as a tool for equity rather than perpetuation of historical injustices.

References

IndustryWired. (2023). Ai-powered loan approvals: Are they fair? *IndustryWired*. Retrieved from https://industrywired.com/ai/ai-powered-loan-approvals-are-they-fair-8748515

ProPublica. (2018a). How your shadow credit score could decide whether you get an apartment. *ProPublica*. Retrieved from https://www.propublica.org/article/how-your-shadow-credit-score-could-decide-whether-you-get-an-apartment

ProPublica. (2018b). Minority neighborhoods pay higher car insurance premiums than white areas with the same risk. *ProPublica*. Retrieved from https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk

RROIJ. (2023). Influence of ai in banking: Ethical and compliance implications. *RROIJ*. Retrieved from https://www.rroij.com/open-access/influence-of-ai-in-banking-ethical-and-compliance-implications.pdf

Weber, M., Yurochkin, M., Botros, S., & Markov, V. (2020). *Black loans matter: Distributionally robust fairness for fighting subgroup discrimination.* Retrieved from https://mitibmwatsonailab.mit.edu/research/blog/black-loans-matter-fighting-bias-for-ai-fairness-in-lending/

IndustryWired (2023) ProPublica (2018a) ProPublica (2018b) RROIJ (2023) Weber et al. (2020)