

# AI exhibits racial bias in mortgage underwriting decisions, researchers find

August 21 2024

---



Credit: Pixabay/CC0 Public Domain

Putting AI to use in mortgage lending decisions could lead to discrimination against Black applicants, according to new research. But researchers say there may be a surprisingly simple solution to mitigate

this potential bias.

In an experiment using leading commercial large language models (LLMs) to evaluate loan application data, Lehigh researchers found that LLMs consistently recommended denying more loans and charging higher interest rates to Black applicants compared to otherwise identical white applicants.

This discovery is particularly alarming given the historical and ongoing racial disparities in homeownership.

"This finding suggests that LLMs are learning from the data they are trained on, which includes a history of racial disparities in mortgage lending, and potentially incorporating triggers for racial bias from other contexts," said Donald Bowen III, assistant professor of finance in the College of Business and one of the authors of [the study](#), available as a working paper on *SSRN*.

The study used real mortgage application data, drawn from a sample of 1,000 loan applications included in the 2022 Home Mortgage Disclosure Act (HMDA) dataset, to create 6,000 experimental loan applications. In the experiment, researchers manipulated race and credit score variables to determine their effects.

The results were stark: Black applicants consistently faced higher barriers to homeownership, even when their financial profiles were identical to white applicants.

Based on the experimental results using OpenAI's GPT-4 Turbo LLM, Black applicants would, on average, need credit scores approximately 120 points higher than white applicants to receive the same approval rate, and about 30 points higher to receive the same interest rate.

Models also exhibited bias against Hispanic applicants, generally to a lesser extent than against Black applicants.

The bias against minority applicants was highest for "riskier" applications that had a low credit score, high debt-to-income ratio, or high loan-to-value ratio.

Researchers also tested other LLMs, including OpenAI's GPT 3.5 Turbo (2023 and 2024) and GPT 4, as well as Anthropic's Claude 3 Sonnet and Opus, and Meta's Llama 3-8B and 3-70B.

Bias was generally consistent across the spectrum of LLMs in regard to interest rate recommendations. However, researchers found high variation in approval rates produced by different models.

ChatGPT 3.5 Turbo was found to show the highest discrimination, while ChatGPT 4 (2023) exhibited virtually none.

"It's somewhat surprising to see racial bias, given the efforts LLM creators take to reduce bias overall combined with the large amount of regulations relating to fair lending," Bowen said, noting that the training data of these models almost certainly includes federal regulations prohibiting the use of race as a factor in making lending decisions.

But even more surprising was the ability to remove persistent bias in results with a simple solution—instructing the LLM to use no bias in making decisions.

When the LLMs were instructed to ignore race in their decision-making, the [racial bias](#) virtually disappeared.

"It didn't partly reduce the bias, or overcorrect. It almost exactly undid it," Bowen said.

**More information:** Donald E. Bowen III et al, Measuring and Mitigating Racial Bias in Large Language Model Mortgage Underwriting, *SSRN* (2024). [DOI: 10.2139/ssrn.4812158](https://doi.org/10.2139/ssrn.4812158). [papers.ssrn.com/sol3/papers.cfm ... ?abstract\\_id=4812158](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4812158)

Provided by Lehigh University

Citation: AI exhibits racial bias in mortgage underwriting decisions, researchers find (2024, August 21) retrieved 22 August 2024 from <https://phys.org/news/2024-08-ai-racial-bias-mortgage-underwriting.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.