ChatGPT

Historical Patterns of Discrimination in Lending Technologies

Lending decisions have long been shaped by social categories. In many countries, it was once legal (and common) for lenders to use **race, gender, class,** etc. to deny credit. For example, mid-20th-century US banks practiced *redlining*: literally drawing "red" zones around predominantly Black or immigrant neighborhoods and systematically refusing mortgages to those residents, regardless of income [1]. Similarly, women were routinely denied credit or required male co-signers before legal bans. Only in the 1960s–70s did civil rights and women's movements drive anti-bias laws (e.g. the US Fair Credit Reporting Act in 1970 and Equal Credit Opportunity Act in 1974) that forbade using race or gender in lending [2] [3]. Credit-scoring technology itself was invented in part to remove human bias: FICO's founders explicitly aimed to make credit decisions *"more objective and efficient"* by eliminating the racial, gender and class biases of manual review [4].

Yet lending technologies have often **amplified** social stratification rather than erased it. Under pre-digital (class-based) systems, individuals were grouped by shared identities: e.g. early insurance and credit rules treated members of a class (say, women or a racial group) identically. By contrast, modern credit scoring is "attribute-based": each person is scored on innumerable factors, fragmenting individuals out of collective groups [5] [6]. The shift means that, in practice, an individual's credit score detaches them from any recognizable social group – making it harder for disfavored groups to organize collective resistance. Sociologists note that in class-based regimes, group membership was explicit and contestable; in attribute-based scoring, social categories (like gender) are often suppressed or hidden, even as they still influence decisions indirectly [5] [6].

Some populations have been **"edge cases"** of these systems. Large groups (e.g. the rural poor, informal workers, indigenous communities) historically fall outside formal data collection. For instance, in Vietnam only 30–40% of adults have bank accounts or credit histories; roughly 60 million people are entirely "under the radar" of traditional credit bureaus [7]. These informal-sector borrowers are often targets for fintechs using new data, but they remain technically "invisible" to legacy systems. Likewise in the US, millions of low-income or immigrant residents have no FICO score at all. Thus, rather than optimizing for everyone, lending algorithms often simply *ignore* or exclude those lacking standard data, effectively cementing them outside the credit market.

Data Representation and Social Context

Lending data do not neutrally mirror society; they are shaped by politics and power. Which categories get recorded (e.g. marital status, number of dependents, area of residence) and which are omitted (e.g. race or class markers) is a social choice with consequences. As Bourdieu observed, the very "categories that make the social world possible" are themselves "the stakes … of political struggle" [5]. In credit, protected classes are barred as inputs, but other attributes (ZIP codes, education level, mobile usage patterns, etc.) can stand in as proxies. Anupam Datta's team notes that when an algorithm *"uses ZIP code… it is strongly correlated with race"* in many countries, the system ends up making de facto race-based decisions even without explicit race data [8] [9]. This is the politics of classification: demographic labels are transformed (and sometimes erased) in the data, shifting power to those who define the categories.

Large gaps in data introduce **strategic ignorance**. For example, about 26 million U.S. adults ("one in ten") are *"credit invisible"* – they have no credit history with any bureau [10]. They include disproportionate numbers of low-income, immigrant, and rural consumers [11]. Without records, these people are often excluded from credit entirely. Globally, it is even starker: roughly **2 billion** adults worldwide are unbanked and have no formal credit history [12]. In Vietnam (as a case study), 61–70% of the population lacked bank accounts as of 2019, leaving some 60–70 million borrowers locked out of formal loans [7]. In effect, the absence of data is structural: informal-economy workers and marginalized groups are made invisible, perpetuating inequality. Only recently have lenders tried to fill these voids with alternate "digital footprints" (mobile usage, online behavior, social media) [12]. Such efforts can boost inclusion – for example, one study found combining online data with traditional credit data vastly improved predictions, potentially granting loans to many previously excluded applicants [12]. But it also raises new biases, as those digital signals often correlate with wealth or social status.

Thus, **classification and bias** are woven into the data itself. Social categories (race, caste, gender, religion) may be hidden or scrambled in the algorithmic age, but their legacies survive in the structure of who gets counted. As Krippner and Hirschman argue, the digital shift "detaches individuals from sociologically meaningful groups," often *suppressing* salient social categories [13] [14]. The resulting "statistical citizens" of credit systems are defined by a profusion of attributes, which can conceal group identities. This fragmentation concentrates power in the hands of designers: if classifications are too opaque or individuated, affected communities struggle to contest unfair patterns.

Ethical Frameworks for Fairness Evaluation

Multiple ethical theories offer different lenses on fairness in lending. **Consequentialist** approaches judge fairness by outcomes. A *utilitarian* framework would aim to maximize total benefit (e.g. overall loan repayment rates or economic welfare), potentially at the cost of unequal treatment. In practice, a utilitarian lender might favor decisions that boost profit or default rates, even if it means accepting more risk for some group if it raises aggregate utility. *Egalitarian* views demand equal distribution – for example, striving for equal loan-approval rates or interest rates across all demographic groups. By contrast, *prioritarian* ethics insist on helping the worst-off: here lenders would give extra weight to disadvantaged applicants' welfare, perhaps by relaxing requirements or offering more generous terms to those with thin files. (In AI terms, social-welfare researchers formalize these ideas via "alpha-fairness" functions that interpolate between purely utilitarian and maximin goals [15] [16].) As one analysis notes, utilitarianism, prioritarianism and egalitarianism respectively "maximize the sum of utility," "give more weight to utility gains for the worst off," and "minimize inequality" [17]. In credit, adopting each yields different design choices: a utilitarian system might endorse strict risk models, an egalitarian one might enforce quotas or equalized odds, and a prioritarian one might bias in favor of underserved borrowers.

Deontological ethics focus on duties and rights. A deontologist would insist that certain practices are categorically impermissible regardless of outcomes. For example, even if excluding a protected attribute (like race) slightly harms predictive accuracy, a deontological perspective might forbid any use of it or any proxies. By Kantian logic, all applicants deserve equal respect and fair treatment as ends-in-themselves. In practice this could translate into strict non-discrimination rules: e.g. never using sensitive categories, and applying exactly the same credit criteria to similar profiles. (NumberAnalytics illustrates this by comparing medical triage debates: a deontologist would refuse to take from one patient even to save many, reflecting a rule-bound stance [18]. In lending, this maps to "I will not penalize someone because of a protected class membership" as a categorical rule.)

Virtue ethics shifts focus to the moral character of agents (e.g. the lender or the system designers). It asks: what kind of *people* (or companies) should make lending decisions? Virtue ethics emphasizes traits like fairness, empathy and justice in the developers and institutions. For instance, a virtue-ethics view might stress that lenders cultivate *justice* by ensuring their criteria are compassionate and transparent, or that engineers build algorithms embodying *honesty* and *prudence* [19]. Rather than fixating on rules or totals, it asks whether the process respects human dignity – a lender acting virtuously would go beyond mere compliance and consider whether their models serve the common good.

These frameworks have concrete implications for AI lending. A utilitarian data scientist might optimize a model for total repayment (even if it means denying many minority applications as long as defaults fall). A prioritarian might add a "bonus" in the scoring for historically marginalized borrowers. An egalitarian might enforce post-hoc parity (e.g. adjust thresholds so approval rates match across groups). A deontologist would push for safeguards: banning use of sensitive attributes or any proxies, and demanding strict audit trails. A virtue ethicist would encourage involving community stakeholders in design, or building in mechanisms for redress. Tensions arise: for example, aiming for equalized error rates can conflict with maximizing utility or vice versa [20]. In practice, no single doctrine suffices. Rather, designing fair credit AI likely requires a blend of values – prioritizing the needy while respecting inviolable rights and cultivating a culture of fairness.

Modern Manifestations of Historical Biases

Historical inequities often resurface in subtle, technical forms today. **Proxy discrimination** is ubiquitous: features permitted by law turn out to stand in for prohibited categories. For example, numerous studies show that ZIP or postal code strongly encodes race or wealth in many countries [8] [21]. An algorithm using geography for "default risk" may thus indirectly charge higher rates to racial minorities – a practice Datta et al. call **"not a defensible proxy."** Indeed, CMU researchers warn that if "ZIP code is encoding race," then using it in credit is effectively race-based decision-making [8] [9]. Similarly, **mobile phone data** have become proxies for income and behavior. Fintech firms in Asia mine call records, airtime top-ups, app usage and even social network connections to score credit [22] [23]. But phone-usage patterns correlate with socioeconomic status and gender (women and the elderly often use less data), so these signals can replicate class or gender bias under the guise of “alternative data.” Likewise, online behavior (shopping choices, email provider, even the smartphones one owns) can correlate with race, age or wealth [24] [25] – so digital footprints, while promising inclusion, risk encoding historical prejudices into new models.

These cycles are often **self-reinforcing**. Research finds that standard credit scores are significantly *less accurate* for disadvantaged groups: one study showed scores for low-income and minority borrowers were 5–10% less predictive of default than for richer or non-minority groups [26] [27]. In other words, biased data make the algorithms noisier for certain people. The result is a vicious feedback loop: if minorities or the poor are denied credit more often (by algorithm or human), they can’t build credit histories; this yields “thin files” that further degrade their scores [27] [28]. The Stanford analysis concludes this “misallocation of credit” *perpetuates* inequality – people who *should* qualify often get rejected and thus lose the chance to build wealth [28]. In effect, the algorithm’s errors become new (automated) redlining: past discrimination generates skewed data, which today’s AI then “learns,” further disadvantaging those very communities [29] [27].

These issues pose thorny ethical dilemmas. For instance, some features have “dual use”: debt-to-income ratio is known to correlate with race, yet it is a legitimate credit factor. Datta et al. note that while zip code is an *unjustified* proxy for race, debt-to-income – even if unevenly distributed – may be defensible if it truly predicts repayment [30]. Designing algorithms thus requires judging when a feature is a harmless risk factor vs. a covert proxy. There are also emergent harms beyond discrimination. Critics warn that if online behavior feeds credit scores, consumers might feel compelled to “game the system.” Imagine someone avoiding useful websites or even buying an iPhone (instead of a cheaper phone) just to appear creditworthy – a pressure on personal freedom [25]. Or consider how black-box scoring can produce puzzling denials, eroding trust in lending. In short, credit algorithms can entrench prejudice (through proxies) and privacy invasions (through digital surveillance) *without anyone intending these harms*.

Summary

In sum, lending systems embed historical bias in both their **technology and data**. Early classification schemes overtly tied credit to race, gender, caste or class – a fact only later outlawed. Modern credit scoring, though formally “blind” to identity, inherits the legacy of those categories via the data itself and the proxies people generate. Data gaps (the unbanked, informal borrowers) create entire blind spots that mirror global inequality. Thus algorithmic fairness in finance is not just a technical fix: it must grapple with centuries of structural discrimination and the uneven social terrain of data.

For this reason, standard AI fairness tools (like enforcing statistical parity) are insufficient in credit. They often miss the *depth* of historical injustice and the *value* of economic inclusion. Instead, fairness in lending demands context-sensitive ethics: frameworks that weigh total welfare against equity for the worst-off, that respect individual rights (no clandestine discrimination), and that cultivate lender accountability. In practice, this might mean blending utilitarian aims (broad access to credit) with prioritarian corrective measures (subsidies or score adjustments for marginalized borrowers), all while upholding deontological bans on explicit bias and virtues of transparency and justice.

Globally, lenders and regulators have begun to recognize these challenges. For example, some countries mandate reporting on outcomes across social groups, and fintech innovators explore *experimental lending* to gather new data (as Stanford suggests) [31]. But the key lesson is that **ethical AI in lending must adapt to its unique context**. Consumer credit is not like generic ranking – it is entwined with socio-economic mobility, historical exclusion, and personal identity. Therefore, fairness evaluations in this domain must go beyond technical parity and incorporate lessons from history, economics and human rights. In the end, a truly fair lending system will be one that explicitly acknowledges and compensates for past stratification, rather than pretending those old categories never mattered.

Sources: Credible studies and analyses of lending discrimination and AI fairness, as cited above [1] [4] [7] [8] [26] [27] [17] [18] [24] [32], among others.

---

[1] [2] [3] How Race Impacts Credit: The History And Effect On Lending Practices | Bankrate
https://www.bankrate.com/personal-finance/credit/does-race-affect-credit-score/

[4] When Did Credit Scores Start? | Capital One
https://www.capitalone.com/learn-grow/money-management/when-did-credit-scores-start/

[5] [6] [13] [14] The person of the category: the pricing of risk and the politics of classification in insurance and credit
https://sites.lsa.umich.edu/krippner/wp-content/uploads/sites/248/2024/06/Krippner-and-Hirschman_Person-of-the-Category_2022.pdf

[7] [22] [23] iseas.edu.sg
https://www.iseas.edu.sg/wp-content/uploads/2021/01/ISEASEWP2021-1Lainez.pdf

[8] [9] [30] Thwarting bias in AI systems - College of Engineering at Carnegie Mellon University
https://engineering.cmu.edu/news-events/news/2018/12/11-datta-proxies.html

[10] [11] [12] [21] [24] [25] [29] When Algorithms Judge Your Credit: Understanding AI Bias in Lending Decisions
https://www.accessiblelaw.untdallas.edu/post/when-algorithms-judge-your-credit-understanding-ai-bias-in-lending-decisions

[15] [16] [17] [20] Reassessing the Fairness of Artificial Intelligence through Social Welfare Optimization - Securities.io
https://www.securities.io/reassessing-the-fairness-of-artificial-intelligence-through-social-welfare-optimization/

[18] [19] Navigating AI Ethics in Philosophy
https://www.numberanalytics.com/blog/ethics-of-ai-in-intro-to-philosophy

[26] [27] [28] [31] How Flawed Data Aggravates Inequality in Credit | Stanford HAI
https://hai.stanford.edu/news/how-flawed-data-aggravates-inequality-credit

[32] gala.gre.ac.uk
https://gala.gre.ac.uk/id/eprint/30306/3/30306%20SANGWAN_3000_Years_Of_Discrimination_And_Counting_%28Draft%29_2020.pdf