ISSN: 2710-0170

New Forms of Ethnic Employment Discrimination in the United States Based on Big Data Algorithms and the Supervisory Role of We-Media

Yihan Chen, Xixia Yu

University of Shanghai for Science and Technology, Shanghai, China

Abstract

This study focuses on the issue of ethnic employment discrimination in the United States. Given the limitations of traditional research methods in identifying hidden discriminatory phenomena, it innovatively uses big data algorithms to mine and analyze multi-source data from recruitment websites and job - seeking platforms. The research reveals new forms of ethnic employment discrimination in the United States in links such as job search, advertising push, and resume screening, and quantifies the impact of ethnic factors on employment opportunities. At the same time, it deeply explores the role and dilemmas faced by we - media in supervising ethnic employment discrimination. The research shows that big data algorithms provide a powerful tool for accurately identifying ethnic employment discrimination. Although we - media can expose some discriminatory behaviors, they are restricted by factors such as algorithmic black boxes and legal lag. It is necessary to coordinate with legal improvement and technological innovation to build an all - round anti - discrimination mechanism to promote the realization of ethnic employment equity in the United States.

Keywords

Ethnic employment; employment discrimination; big data; algorithmic black box.

1. Introduction

The issue of ethnic employment discrimination in the United States has long attracted much attention. From a historical perspective, after the abolition of slavery, ethnic minorities such as African Americans have obtained equal employment rights at the legal level, but discrimination is still common in the actual employment process. Traditional methods to study employment discrimination mostly rely on questionnaires, case studies, and simple statistical analysis of public employment data. For example, in the past, employment discrimination was judged by comparing the average wages of different ethnic groups. For instance, the average wage of black workers in the United States is about 20% lower than that of white workers. However, this method has obvious flaws. Because wage differences may be interfered by many factors, such as the uneven level of human capital (educational background, work experience, etc.) among different ethnic groups, it is difficult to determine the existence of discrimination simply by virtue of wage gaps. In addition, questionnaires are susceptible to subjective factors, and respondents may conceal their true experiences or opinions for various reasons; case studies can conduct in - depth analysis of individual discriminatory incidents, but they lack extensive representativeness and cannot fully reflect the overall discrimination situation in the job market.

With the development of Internet technology, big data algorithms have gradually penetrated into various links of the job market. Professional social platforms such as LinkedIn have accumulated massive amounts of users' career information, and the amount of data generated by recruitment websites and job - seeking platforms every day is even more amazing. These

ISSN: 2710-0170

data provide a new perspective for studying ethnic employment discrimination. Big data algorithms can mine and analyze these data, break through the limitations of traditional research, insight into the hidden ethnic employment discrimination phenomena that were difficult to detect in the past, and provide an opportunity for in - depth research on this issue.

2. New Forms of Discrimination from the Perspective of Algorithms

2.1. Data Mining and Revelation of Differential Treatment

Big data mining from leading recruitment platforms—including job search sites like Indeed and Glassdoor, and talent-matching platforms like ZipRecruiter—has uncovered clear differential treatment of ethnic groups across three critical links in the U.S. job market. For African-American job seekers, the keyword matching algorithms used by these platforms create a significant disadvantage: a longitudinal study analyzing data from 2010 to 2025 found that when African-Americans input the same job search keywords (e.g., "senior marketing manager" or "software engineer") as white counterparts, the number of high-quality positions aligned with their career goals (defined by salary, benefits, and growth potential) in their results is approximately 30% lower.

In the targeted push of recruitment advertisements, Latino job seekers face similar bias. Facebook's internal advertising data, disclosed in a 2022 regulatory filing, showed that ads for high-paying roles (e.g., financial analysts, corporate executives) were delivered to white users at three times the rate of Latino users. Even when controlling for variables like location, education, and work experience, Latino users had only a 40% chance of receiving these high-paying job ads compared to white users.

Asian job seekers encounter unfairness in the resume screening phase. A 2023 analysis of over 100,000 resume screening records from a major U.S. recruitment platform revealed that Asian applicants—even those with educational backgrounds (e.g., Ivy League degrees) and work experience (e.g., senior roles at Fortune 500 companies) equivalent to or stronger than white applicants—had a 15% lower probability of their resumes being shortlisted for interviews. This gap persisted across industries, from technology to healthcare.

2.2. Model Construction and Quantitative Analysis of Ethnic Factors

Data analysis models, such as logistic regression analysis (used to predict employment outcomes) and decision tree models (to map factor impacts), have enabled researchers to quantify the influence of ethnic factors on employment opportunities and compare them to traditional demographic variables like education and work experience. A prominent example is a model built using 2015–2024 global recruitment data from a multinational corporation with operations across 50+ countries. The model's results showed that ethnic factors had a negative impact coefficient of -0.25, meaning candidates from ethnic minority groups faced a 25% reduction in employment opportunities compared to white candidates with identical qualifications. In contrast, the positive impact coefficient for educational background (e.g., holding a master's degree vs. a bachelor's) was 0.15, and for work experience (e.g., 5+ years vs. 1–2 years) was 0.1—indicating that the negative effect of ethnicity outweighs the positive benefits of advancing one's education or gaining additional work experience.

Further industry-specific modeling highlighted variations in discrimination intensity. In the financial industry, which has a long history of systemic bias, the negative impact coefficient of ethnic factors on African-American employment opportunities reached -0.35—meaning African-Americans were 35% less likely to be hired for roles like investment bankers or risk analysts. In the technology sector, Latino job seekers faced the strongest bias, with an impact coefficient of -0.3, particularly in technical roles like software developers and data scientists.

ISSN: 2710-0170

These models demonstrate how big data can isolate and measure ethnic bias that might otherwise remain hidden in subjective hiring processes.

3. Supervision and Exposure Actions of We-Media

3.1. Case Enumeration

Several high-impact cases, exposed and amplified by we-media, have shed light on persistent ethnic employment discrimination:

The 2004 study Are Emily and Greg More Employable Than Lakisha and Jamal?, conducted by researchers from the University of Chicago and MIT and published by the National Bureau of Economic Research (NBER), became a landmark in discrimination research. The team sent 5,000 fictional but identical resumes to U.S. employers, varying only the names—using "white-sounding" names like Emily and Greg and "African-American-sounding" names like Lakisha and Jamal. The results were stark: resumes with white names received 50% more interview requests, revealing explicit racial bias in initial hiring screenings.

A 2021 Harvard Business School study, widely shared by workplace we-media accounts, identified "cultural mismatch" as a hidden form of racial bias. The research found that African-American candidates were often rejected for having "different communication styles" (e.g., direct speech or community-focused language), while white applicants using identical language were praised as "passionate" or "results-driven"—exposing how implicit prejudice influences hiring decisions.

3.2. Communication Effect

The dissemination of these cases across traditional media, social media, and we-media has generated multi-layered social impact, turning academic findings and legal disputes into public conversations about equity.

Traditional media outlets like The New York Times and The Washington Post played a key role in legitimizing the issue. After the 2004 NBER study was published, these outlets ran in-depth reports that not only explained the study's methodology (e.g., controlled resume variables) and data conclusions but also contextualized the findings within a history of U.S. racial discrimination in employment. This coverage brought "name-based discrimination"—a previously obscure form of bias—into mainstream discourse, sparking debates in academic circles about measurement tools for implicit bias and in public forums about how to reform hiring practices.[1] It also directly inspired follow-up research, including the 2017 Harvard replication study, which built on the NBER's framework to expand understanding of Latino discrimination.

Social media platforms demonstrated their power to mobilize public pressure. When Facebook's algorithm discrimination case was exposed in 2019, hashtags like #Facebook Discriminates Against Minorities and #Job Ad Bias trended on Twitter and Instagram within 48 hours. Users shared personal stories of discriminatory job searches, created memes critiquing Facebook's "neutral" algorithm claim, and organized online petitions demanding accountability. The topic reached over 200 million views in one week, drawing participation from civil rights groups (e.g., the NAACP), tech ethicists, and even some Facebook employees. This public outcry forced Facebook to respond faster than it initially planned, with the $14.25 million settlement and algorithm revisions widely seen as direct responses to social media pressure.

We-media, particularly workplace-focused blogs and minority advocacy accounts, made complex discrimination research accessible to ordinary people. Harvard's 2021 "cultural mismatch" study, for example, was broken down into short, relatable posts by we-media creators: they paired the study's findings with real anecdotes from African-American professionals (e.g., a candidate rejected for "not fitting the team culture" despite strong

ISSN: 2710-0170

qualifications) and used infographics to explain how "communication style" bias manifests. These posts resonated with millions of workplace users, who shared their own similar experiences in comment sections—creating a "collective narrative" that transformed abstract academic data into a tangible, shared struggle.[2]This not only raised public awareness of hidden discrimination but also provided grassroots support for minority rights organizations pushing for policy changes.

Overall, media dissemination has been critical in shifting the narrative around ethnic employment discrimination: it has moved the issue from isolated complaints or academic papers to a widely recognized public problem, pressuring companies to reform their practices and policymakers to consider stronger protections.

4. Anti-Discrimination Dilemmas: Algorithmic Black Box, Legal Lag and Limitations of We-Media

4.1. Algorithmic Black Box Problem

While big data algorithms have streamlined recruitment processes (e.g., reducing resume screening time from weeks to hours), their complexity has created an "algorithmic black box" problem that makes ethnic discrimination harder to detect and address. Most recruitment algorithms rely on thousands of lines of code and complex mathematical models (e.g., machine learning neural networks) to make decisions—logic that is incomprehensible to external stakeholders, including regulators, job seekers, and even some company employees not on the data science team. This opacity allows implicit biases to be embedded unintentionally (or intentionally) in the algorithm, often hidden in "feature engineering" (the process of selecting data points to train the model). For example, some algorithms use proxy variables correlated with ethnicity—such as residential zip codes (which map to racially segregated neighborhoods), commonly used languages (e.g., Spanish for Latino communities), or social media activity (e.g., following Black-owned businesses)—to filter candidates, even though they do not explicitly reference race.

This lack of transparency makes external audits ineffective: auditors can see the algorithm's outputs (e.g., which candidates are shortlisted) but not the internal logic that produced them, making it hard to prove if bias is intentional or accidental. Even when discriminatory outcomes are clear (e.g., a 30% lower interview rate for Latinos), pinpointing the exact cause in the algorithm or assigning responsibility is nearly impossible.

Two high-profile cases illustrate this dilemma: In 2018, Amazon abandoned its AI recruitment tool after internal testing revealed it systematically penalized resumes with "female-associated" attributes—such as mentions of attending women's colleges or participating in women's professional organizations. The algorithm had been trained on Amazon's past 10 years of hiring data, which was dominated by male hires (a reflection of the tech industry's historical gender imbalance), so it learned to associate "female" attributes with lower hireability. Critically, the black box nature of the tool meant external observers (including regulators) had no way to detect this bias until Amazon's own team uncovered it. Even then, Amazon faced no legal consequences because proving the company "intentionally" discriminated was impossible—the bias was a byproduct of the training data, not explicit programming.

Meta faced a similar issue in 2020 when accused of using proxy variables to exclude African-American and Latino users from high-paying job ads. While Meta's ad platform did not allow advertisers to select "race" as a filter, its algorithm used user interests (e.g., "hip-hop music" or "Latino cultural events") and location data to indirectly target or exclude groups—effectively creating a racial filter without labeling it as such. Existing legal frameworks, like the U.S. Civil Rights Act of 1964, were designed to address "intentional discrimination" (e.g., an employer

ISSN: 2710-0170

explicitly saying "we won't hire Black candidates")—not bias embedded in algorithms through proxy variables.

Newer regulations aimed at addressing this gap, such as the EU's Artificial Intelligence Act (which requires transparency for "high-risk" AI systems like recruitment tools) and New York City's Local Law 144 (which mandates audits of hiring algorithms), have limitations. Companies often resist disclosing core algorithm code by citing "trade secrets"—a legal protection that allows them to keep proprietary technology private. Additionally, job seekers rarely receive specific reasons for rejection (e.g., "you were excluded by our algorithm due to zip code"), making it hard for them to challenge discriminatory decisions or pursue legal action.

The algorithmic black box, often shielded by claims of "technological neutrality," has allowed discrimination to become more systemic and covert. Addressing it will require mandatory disclosure of algorithmic impact assessments (reports showing how the algorithm affects different ethnic groups), giving regulators the authority to audit algorithm code, and establishing standardized "algorithmic fairness" metrics—without these steps, big data risks worsening, not reducing, ethnic inequalities in employment.

4.2. Legal Challenges

The U.S. legal system's inability to keep up with algorithmic discrimination stems from the outdated design of its anti-employment discrimination laws. The foundation of these laws is Title VII of the Civil Rights Act of 1964, which prohibits discrimination based on race, color, religion, sex, or national origin. However, Title VII was crafted to address the overt discrimination common in the 1960s—such as "whites only" job postings or employers openly refusing to hire minorities—not the hidden, data-driven bias of algorithms.

A key legal tool for addressing indirect discrimination is the "disparate impact" theory, established by the Supreme Court in the 1971 case Griggs v. Duke Power Co. This theory holds that a policy or practice (even one that seems neutral) is illegal if it disproportionately harms a protected group (e.g., a height requirement that excludes more women than men).[3] While this theory can apply to algorithms, it has limitations: plaintiffs must prove that the algorithm's "disparate impact" is significant and that there is no "business necessity" for the biased practice. For algorithms, this is extremely hard—plaintiffs often lack access to the algorithm's data or logic to prove impact, and companies can easily claim the algorithm is "necessary" for efficiency.

Subsequent legal updates have failed to close this gap. The 1991 amendments to the Civil Rights Act strengthened remedies for discrimination (e.g., allowing punitive damages) but did not address digital-age bias. The 2007 Equal Employment Opportunity Commission (EEOC) Compliance Manual included a brief note that "automated systems may produce discrimination" but provided no specific rules for regulating algorithms. The Obama administration's 2016 report Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights was the first to systematically warn about algorithmic discrimination risks, but its recommendations (e.g., voluntary algorithm audits) were not made legally binding.

The most significant recent action is New York City's Local Law 144, passed in 2021, which requires companies using hiring algorithms to conduct annual bias audits and disclose the results. However, the law only applies to companies operating in New York City, leaving most U.S. workers unprotected.

Today's legal system has three structural flaws that hinder efforts to combat algorithmic discrimination: First, proving bias requires meeting Title VII's "intentional discrimination" or "disparate impact" standards, but the algorithmic black box makes it nearly impossible for plaintiffs to gather the necessary evidence (e.g., algorithm code, training data). Second, the Uniform Guidelines on Employee Selection Procedures—a set of rules for validating hiring tools—was written for traditional methods (e.g., aptitude tests) and does not account for machine learning algorithms, which evolve over time as they process new data. Third,

ISSN: 2710-0170

regulatory authority is split between the EEOC (which enforces anti-discrimination laws) and the Federal Trade Commission (FTC, which regulates unfair business practices), leading to overlapping jurisdictions and slow, inconsistent enforcement.

These flaws often let companies avoid accountability. For example, in Amazon's 2018 AI recruitment case, the company faced no legal penalties because the EEOC could not prove "intentional discrimination," and Amazon argued the algorithm was a "neutral" tool—even though it amplified historical biases. To break this impasse, a new federal regulatory framework is needed: one that mandates algorithmic transparency, requires third-party audits for high-risk recruitment tools, and establishes clear quantitative standards for what constitutes "unlawful disparate impact" in algorithmic decisions.

5. Conclusion

Globally, the U.S. experience offers valuable lessons for other countries grappling with algorithmic discrimination. As a leader in algorithmic technology, the U.S. faces challenges that many nations will soon encounter as they adopt AI-driven recruitment tools. For example, the EU's Artificial Intelligence Act, which was finalized in 2024, draws on U.S. research to include provisions for auditing high-risk AI systems—showing how U.S. findings can inform international regulation. More broadly, addressing algorithmic bias is a prerequisite for achieving social equity in the digital economy. As technology becomes more central to employment, ensuring algorithms do not amplify existing inequalities will determine whether digital growth is inclusive—benefiting all ethnic groups—or exclusive, widening racial gaps.

Big data algorithms have revealed and amplified new forms of ethnic employment discrimination in the United States, with African-American, Latino, and Asian job seekers facing unfair treatment in job searches, ad targeting, and resume screening. We-media has emerged as a powerful force for accountability.

Acknowledgements

I sincerely thank my supervisor for her insightful guidance throughout this research. Gratitude also goes to institutions that provided data access—their support was crucial for exploring U.S. ethnic employment discrimination in the algorithmic era.References

References

[1] Metcalfe D J ,Drake E R .Second Look at Reported Racial-Ethnic Employment Differences in the Supported Employment Demonstration.[J].Psychiatric services (Washington, D.C.), 2024, appips20230612.

[2] Kinitz J D ,Tran K N ,Shahidi V F , et al.Associations of minority stress and employment discrimination with job quality among sexual- and gender-minority workers.[J].Scandinavian journal of work, environment & health,2025,

[3] Yang, Chao.On the Formation of Equal Employment Policies for African Americans in the United States.[J].Northern Forum,2011,(01):82-87.