Title: Gender difference in under-reporting hiring discrimination in South Korea: a machine learning approach

Author: Jaehong Yoon*, Ji-Hwan Kim*, Yeonseung Chung, Jinsu Park, Glorian Sorensen, Seung-Sup Kim

*These authors contributed equally to this work

Objective: To examine the gender difference in under-reporting hiring discrimination by building prediction models for the workers who responded 'Not applicable (NA)' to a question about hiring discrimination, although they were eligible to answer.

Methods: Using the data from 3,576 waged workers in the 7th wave (2004) of the Korea Labor and Income Panel Study, we trained and tested nine machine learning algorithms using 'Yes' or 'No' responses regarding the lifetime experience of hiring discrimination. We then applied the best-performing model to estimate prevalence of experiencing hiring discrimination among those who answered NA. Under-reporting hiring discrimination was calculated by comparing the prevalence of hiring discrimination between 'Yes' or 'No' group and 'NA' group.

Results: Based on the prediction from the random forest model, we found that 58.8% of the 'NA' group were predicted to experience hiring discrimination, while 19.7% of the 'Yes' or 'No' group reported hiring discrimination. Furthermore, we found that female workers were more likely to under-report hiring discrimination than male workers.

Conclusions: This study introduces a methodological strategy for epidemiologic studies to address the under-reporting of discrimination by applying machine learning algorithms.

Keywords: Social Discrimination; Social Epidemiology; Machine Learning

Affiliation and contact information for all authors:

Jaehong Yoon

Department of Public Health Sciences, Graduate School of Korea University, Seoul, South Korea

jhyoon.epik@gmail.com

Ji-Hwan Kim

Department of Public Health Sciences, Graduate School of Korea University, Seoul, South Korea

sprits2004@korea.ac.kr

Yeonseung Chung

Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

dolyura@kaist.edu

Jinsu Park

Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea

starlight7@kaist.ac.kr

Glorian Sorensen

Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Center for Community-Based Research, Dana-Farber Cancer Institute, Boston, Massachusetts, USA

glorian_sorensen@dfci.harvard.edu

Seung-Sup Kim

Department of Public Health Sciences, Graduate School of Korea University, Seoul, South Korea

Department of Social and Behavioral Sciences, Harvard School of Public Health, Boston, MA, USA

Interdisciplinary Program in Precision Public Health, Korea University, Seoul, South Korea

ssk3@korea.ac.kr

Corresponding author: Seung-Sup Kim

Department of Public Health Sciences, Korea University, 368 Hana Science Hall B, Anam-ro 145, Seonbuk-gu, Seoul 02841, Republic of Korea. E-mail: ssk3@korea.ac.kr

Running Title

Gender difference in under-reporting discrimination in South Korea

Author Contributions:

Conceptualization: SSK, JY, JHK.

Data curation: JY, JHK.

Data analysis: SSK, JY, JHK, YC, JP.

Writing – original draft: SSK, JY, JHK.

Writing – review & editing: SSK, JY, JHK, YC, JP, GS.

ORCID:

Jaehong Yoon: <https://orcid.org/0000-0002-5987-8913>

Ji-Hwan Kim: <https://orcid.org/0000-0001-9424-5962>

Yeonseung Chung: <https://orcid.org/0000-0002-1501-2747>

Jinsu Park: <https://orcid.org/0000-0003-4147-4758>

Glorian Sorensen: <https://orcid.org/0000-0001-6625-7931>

Seusup-Sup Kim: <https://orcid.org/0000-0003-1830-0282>

Epub ahead of print

Gender difference in under-reporting hiring discrimination in South Korea: a machine learning approach

Abstract

Objective: To examine the gender difference in under-reporting hiring discrimination by building prediction models for the workers who responded ‘Not applicable (NA)’ to a question about hiring discrimination, although they were eligible to answer.

Methods: Using the data from 3,576 waged workers in the 7th wave (2004) of the Korea Labor and Income Panel Study, we trained and tested nine machine learning algorithms using ‘Yes’ or ‘No’ responses regarding the lifetime experience of hiring discrimination. We then applied the best-performing model to estimate prevalence of experiencing hiring discrimination among those who answered NA. Under-reporting hiring discrimination was calculated by comparing the prevalence of hiring discrimination between ‘Yes’ or ‘No’ group and ‘NA’ group.

Results: Based on the prediction from the random forest model, we found that 58.8% of the ‘NA’ group were predicted to experience hiring discrimination, while 19.7% of the ‘Yes’ or ‘No’ group reported hiring discrimination. Furthermore, we found that female workers were more likely to under-report hiring discrimination than male workers.

Conclusions: This study introduces a methodological strategy for epidemiologic studies to address the under-reporting of discrimination by applying machine learning algorithms.

Keywords: Social Discrimination; Social Epidemiology; Machine Learning

1 Introduction

2 Discrimination harms people's life [1]. It could deprive of socio-economic opportunity
3 in daily life (e.g., education, employment, income), and these disadvantages could be
4 accumulated over the life-course [2-4]. Furthermore, the experience of discrimination can exert
5 adverse health consequences [5]. Mounting evidence implicates discrimination in an increased
6 risk of hypertension [6, 7], cardiovascular disease [8, 9], depressive symptoms [10-12], and
7 suicide [13-15]. Indeed, a meta-analysis of 333 articles published between 1983 and 2013
8 found concrete epidemiologic evidence on the health impacts of discrimination [16].

9 However, an important issue in the field is the potential under-reporting of
10 discriminatory experience. Several factors could lead to the underreporting of discrimination
11 [17, 18]. One factor is social desirability, which is the tendency for survey participants to
12 answer questions in a way that they deem socially acceptable [18, 19]. Another relevant factor
13 is 'internalized oppression,' in which people in subordinate groups accept their social status and
14 internalize negative attitudes towards them, thus perceiving their mistreatment experience as
15 being deserved [18].

16 Indeed, previous studies sought to understand the mechanism of reporting
17 discrimination, especially among socially vulnerable groups. For example, in a study of 400
18 people living in a Boston suburb, females tended to deny their personal experience of gender
19 discrimination despite knowing that female workers generally did not receive the opportunities
20 and rewards they deserved [20]. Using a dataset from South Korea, Kim *et al.* [21] compared
21 the predicted logit score of discrimination between genders and found that female workers
22 tended to under-report their discriminatory experiences. Also, a body of literature has found
23 'personal/group discrimination discrepancy,' which is the tendency for people to under-report
24 their personal experience of discrimination while still acknowledging discrimination against

their group [22, 23].

Machine learning offers effective approaches to improving predictive performance in various fields, including social epidemiology [24-26]. Machine learning algorithms have enabled investigating the role of social factors in predicting individual health conditions. Such factors include community-level socioeconomic status (SES) [27, 28], individual-level SES [25, 29], language skill [19], and insurance coverage [19, 25]. For example, a recent study of 66 low- and middle-income countries explored the ability of SES indicators, including household wealth, educational attainment, and occupation, to predict a female's height, which could be an indicator of human welfare at the population level [29]. Thus, machine learning may facilitate attaining a similarly improved understanding of discrimination.

This study sought to investigate the following research purposes using a nationally representative dataset in South Korea: 1) Applying nine machine learning algorithms to build a prediction model for the experience of hiring discrimination among waged workers; 2) To predict whether workers who responded 'Not applicable (NA)' actually experienced hiring discrimination or not using the best prediction model; 3) To assess under-reporting of discrimination by comparing the prevalence of hiring discrimination between training sample ('Yes' or 'No' group) and prediction sample ('NA' group) ; and 4) To examine gender difference in the under-reporting of discrimination

Methods

Study population

This study analyzed a nationally representative longitudinal dataset obtained from the Korean Labor and Income Panel Study (KLIPS), launched in 1998. Five thousand households (13,321 respondents) participated in the 1st wave of KLIPS. KLIPS data from the 1st (1998)

through 22nd (2019) waves are publicly available (www.kli.re.kr). We utilized data from the 7th wave (2004), the only one in which KLIPS measured the experience of discrimination. It provided a unique chance to investigate the behavior of responding NA among those eligible to answer a question about hiring discrimination. Our analyses included waged workers at the time of the survey to ensure that all survey participants were eligible to answer the question about the experience of hiring discrimination. Figure 1 shows the flow chart of data analysis. Our analyses included waged workers at the time of the survey (n=4,257) to ensure that all survey participants were eligible to answer the question about the experience of hiring discrimination. Exclusion criteria were: participants with missing information on any of the variables used in the prediction model (n=621), and those who responded NA to all questions about discrimination (n=60) because the experience of discrimination in other situations provided critical information to improve the performance of the predictive models. The final sample size was 3,576 (Figure 1). Because KLIPS data are publicly available, this study was exempt from review by the Institutional Review Board of the Office of Human Research Administration (KUIRB-2021-0049-01).

Target variable

Hiring discrimination was measured by the question, “Have you ever experienced discrimination in getting hired?” Respondents could answer ‘No,’ ‘Yes,’ or ‘Not Applicable’ (‘NA’). All included individuals were eligible to answer either ‘Yes’ or ‘No’ to the question because they were waged workers at the time of the survey. However, we found that 97 participants (2.7%) responded ‘NA’ to the experience of hiring discrimination. Machine learning models were built to predict whether those who answered ‘NA’ (prediction sample, n = 97) experienced hiring discrimination, using the information from those who answered ‘Yes’

(n=686) and 'No' (n=2,793) (training sample, n=3,479; Figure 1).

Predictors

We selected variables to be included in the prediction model by referring to South Korean literature on discrimination [30, 31]. Gender, age, education level, marital status, employment status, equivalized household income, birth region, self-rated health condition, disability, residential areas, and experience of discrimination in other situations were included as predictors. All predictors used in the analyses were obtained from the 7th wave of the survey, except for information on disability (only available in Wave 9). Age was categorized into six groups (16–24, 25–34, 35–44, 45–54, 55–64, 65+ years old). Education level was measured and classified into three categories: middle school graduate or less, high school graduate, and college graduate or more. Marital status was assessed in three categories: never married, currently married, previously married. Employment status was assessed using two questions about the status of workers (i.e., permanent, temporary, daily) and the type of working hours (i.e., full-time vs. part-time). Full-time permanent workers were defined as the permanent group, whereas all other workers were classified as the non-permanent group. Equivalized household income was assessed by dividing the total household income by the square root of the number of household members, and was classified into four categories based on the quartiles. Birth region was separated into Jeolla-do, South Korea versus other regions considering the stigma that has existed against people born in Jeolla-do [32]. Self-rated health condition was assessed on a five-point scale to the question, “How would you rate your health?” Participants could choose a response from 'very good' (score 1) to 'very poor' (score 5). Considering the small number of very poor and poor responses, the two groups were merged into poor in the analyses. Disability was measured by the yes/no question, “Do you have any

impairment or disability?” Residential area was measured at the metropolitan and provincial levels (i.e., Seoul, Busan, Daegu, Daejeon, Incheon, Gwangju, Ulsan, Gyeonggi-do, Gangwon-do, Chungcheongbuk-do, Chungcheongnam-do, Jeollabuk-do, Jeollanam-do, Gyeongsangbuk-do, Gyeongsangnam-do). Finally, experiences of discrimination in other situations were measured using a modified version of the “Experience of Discrimination” questionnaire. Participants could answer ‘No’, ‘Yes’, or ‘NA’ for the experience of discrimination in each of the following seven situations: receiving income, training, getting promoted, being fired, obtaining higher education, at home, and general social activities.

Building a prediction model

We considered nine machine learning algorithms to build a prediction model for the experience of hiring discrimination among training sample (‘Yes’ or ‘No’ group): logistic regression, random forest, penalized logistic regression (ridge, lasso, and elastic net), k-nearest neighbor, support vector machine (polynomial and radial basis kernel functions), and single-layer artificial neural network. Table S1 details the machine learning algorithms and tuning parameters. We conducted a 10-fold cross-validation to identify the optimal tuning parameters for each of the nine algorithms. For the cross-validation, we split the data into 10 folds such that the experience of hiring discrimination was equally distributed in each fold to prevent the imbalanced distribution of the target variable.

The predictive performance of machine learning algorithms was assessed by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. We selected the algorithm with the largest AUC as the best-performing algorithm and determined the optimal threshold for the binary prediction based on the ROC curve from the best-performing algorithm

such that the sum of sensitivity and specificity was maximized. We then predicted the binary response of experiencing hiring discrimination among those in the prediction sample, and individuals were classified into two groups: 'NA-Yes' ($>\text{threshold}$) or 'NA-No' ($\leq\text{threshold}$).

Finally, we applied a modified Poisson regression model with a robust error variance to investigate under-reporting of discrimination by comparing the prevalence of hiring discrimination between training sample and prediction sample. The odds ratio estimated in logistic regression can overestimate the prevalence ratio (PR), given the high outcome prevalence ($>10\%$) [33]. Results from the association analyses are presented as PRs with 95% confidence intervals (CIs). All machine learning algorithms were estimated from the *tidymodels* package in R statistical software (version 4.0.2; R Development Core Team).

Sensitivity analyses

We checked whether our results were robust to: (1) excluding/including people who answered NA to experiences of discrimination in seven situations other than hiring ($n = 60$), and (2) different choices of probability threshold used to classify the experience of hiring discrimination. These aspects were considered separately at first and later considered together. Therefore, we conducted three sensitivity analyses (Table S2 of the Supplementary Materials).

Results

Table 1 shows the distribution of the study population and the prevalence of hiring discrimination by different categories for each predictor. The overall prevalence of hiring discrimination was 19.2% ($n = 686$). The prevalence of hiring discrimination tended to be higher among participants who were older, had a lower education level, non-permanent

employment status, lower household income, poor self-rated health conditions, or experiences of discrimination in other situations.

Figure 2 shows the cross-validated AUC for each of the nine machine learning algorithms; corresponding ROC curves are depicted in Supplementary Figure S2. AUCs for each algorithm ranged from 0.821 to 0.891; the random forest algorithm performed the best among all algorithms, showing the largest AUC (AUC: 0.891, 95% CI: 0.878, 0.904). The AUC of other machine learning algorithms, except for K-nearest neighborhood algorithms, was similar to that of the random forest model. Variable importance scores of predictors for classifying hiring discrimination in the random forest algorithm (best-performing algorithm) are illustrated in Figure S1. Experience of discrimination in receiving income had the highest score, followed by discriminatory experiences in general social activities and in getting promoted. In the random forest algorithm, the optimal threshold value was determined as 0.268 (Figure S3).

Based on the prediction from the random forest algorithms, we compared the prevalence of hiring discrimination between the training sample ('Yes' or 'No' group) and prediction sample ('NA' group) to investigate the under-reporting of hiring discrimination (Table 2). We found that 58.8% of the 'NA' group were predicted to experience hiring discrimination, while 19.7% of the 'Yes' or 'No' group reported hiring discrimination. It shows that the predicted prevalence of hiring discrimination among workers who responded 'NA' is about three times higher than those among workers who responded 'Yes' or 'No' (PR: 2.98, 95% CI: 2.49–3.57). Combining observed prevalence from the training sample ('Yes' or 'No' group) and predicted prevalence from the prediction sample ('NA' group), we estimated that the prevalence of hiring discrimination among the total population could be 20.8% (n=743) if we count NA-Yes as those who actually experienced discrimination in addition to Yes group. Considering the prevalence of hiring discrimination among the 'Yes' or 'No' was 19.7%

(n=686), it could imply that under-reporting has occurred during the process of reporting discriminatory experience.

We also found a gender difference in under-reporting of hiring discrimination (Table 2). The predicted prevalence of hiring discrimination among the male and female prediction sample ('NA' group) was 45.3% and 84.8%, respectively, whereas the observed prevalence among the male and female training sample ('Yes' or 'No' group) was 18.2% and 20.6%, respectively. Although the predicted prevalence of hiring discrimination was higher than the observed prevalence among male and female workers, under-reporting of hiring discrimination was more substantial among female workers (PR: 4.02, 95% CI: 3.37–4.79) than male workers (PR: 2.41, 95% CI: 1.82–3.20). The estimated actual prevalence of hiring discrimination was 22.6% among total female workers, while the estimated actual prevalence was 19.6% among total male workers. Therefore, underestimating the prevalence of hiring discrimination could occur unless considering the NA response resulting from underreporting, especially among female workers.

In sensitivity analyses, we evaluated the robustness of our results, considering the selection criteria for the probability threshold used to classify the experience of hiring discrimination and the inclusion criteria for the study population (Table S2). First, we applied a different selection criterion when determining a probability threshold for the classification (Figure S4-A). Results of sensitivity analysis 1 were consistent with those of the main analyses and showed that female workers were more likely to under-report hiring discrimination (Table S3). Second, we included participants who responded 'NA' to all questions about the experience of discrimination in the analysis (sensitivity analysis 2 and sensitivity analysis 3), and the results of the cross-validated AUC were similar to those of the main analyses (Figures S5, S6). Gender differences in under-reporting of hiring discrimination (Table S4, Table S5)

also showed similar trends as the main analyses.

Discussion

The results of this study demonstrate the potential to build a performative prediction model for the experience of discrimination using covariates, including demographic and SES information. We found that the predicted prevalence of hiring discrimination in the prediction sample ('NA' group) was higher than those observed in the training sample ('Yes' or 'No' group), which could imply under-reporting of hiring discrimination among Korean workers. Furthermore, we found that the under-reporting of hiring discrimination differed by gender.

These results are consistent with findings from previous studies investigating personal-group discrimination discrepancy [22, 23]. The authors in those studies postulated that **vulnerable groups** might tend to conceal their personal experiences of discrimination even though they recognize that their group experiences discrimination. Furthermore, Operario and Fiske[34] used a quantitative study design to examine personal-group discrimination discrepancy across ethnic groups and found that ethnic minorities were more likely to report personal-group discrimination discrepancy than White people.

Applying machine learning algorithms, we found that there might be a substantial underreporting of discriminatory experience in hiring situations both among male and female who responded that they were not applicable to the question. Previous studies have indicated that asking sensitive questions in the survey could lead to respondents' underreporting of their experience [35, 36]. For example, Yan argued that survey respondents might misreport the sensitive question, even after they decided to answer it, to avoid embarrassing themselves [36, 37]. Also, people who have experienced hiring discrimination might avoid the question to

1 conceal their discriminatory experience to make their response recognized as socially desirable
2 by the interviewer [36, 38]. This underreporting may have been an issue, especially with the
3 KLIPS, in which data are collected through in-person interviews. Additionally, some victims
4 of discrimination might have felt that it is painful to acknowledge that they had been victimized
5 by unfair treatment, which could be a reason for underreporting the experience [37].

6 Notably, we found that underreporting of hiring discrimination was more prevalent
7 among female than male. Previous studies provide several possible explanations for the
8 observed gender differences in the prevalence of 'NA-Yes.' For example, there may be gender
9 differences in the extent to which internalized oppression is experienced among those who
10 answer NA to the hiring discrimination question. People from vulnerable groups may
11 internalize negative attitudes against themselves and perceive their discriminatory experience
12 as non-discriminatory [17, 18]. Previous studies also indicated that socially vulnerable groups,
13 including female, are more likely to internalize and accept unfair treatment against them [39,
14 40], which might lead them not to report their experiences of discrimination. People who have
15 actually experienced hiring discrimination might avoid the question to conceal their
16 discriminatory experience and make their response socially desirable to the interviewer [18]. It
17 may particularly have been an issue with the KLIPS, in which data are collected through in-
18 person interviews.

19 The gender difference in underreporting hiring discrimination observed in this study
20 could be public health concerns. Although previous studies have indicated that not disclosing
21 a violent event to others can negatively influence the victims' health [41-43], few studies
22 examine the underestimating the health impacts of discrimination due to under-reporting. As a
23 posthoc analysis, we examined the association between hiring discrimination and self-rated
24 health after dividing the response to hiring discrimination among the entire population into four

groups: 'No,' 'Yes,' 'NA-No,' 'NA-Yes' (Table S6). The 'NA-Yes' group was more likely to have poor self-rated health than the 'No' group, and this association was statistically significant among both male and female workers. These results should be interpreted cautiously because we could not adjust for confounding variables due to the small sample size. Therefore, future studies with a larger sample are necessary to assess the potential underestimation in the association between discrimination and health.

This study has several limitations. First, people who did not have a job in the 7th wave of KLIPS were excluded from the analyses because we could not determine their eligibility to answer the question of hiring discrimination. This exclusion may have led to the loss of information on people who experienced severe hiring discrimination and as a result were unemployed, which might have improved the performance of the prediction models. Second, unmeasured predictors may have contributed to increasing the predictive power of our models, although our prediction model showed relatively high performance (AUC of the random forest: 0.891). For example, information on physical appearance and sexual orientation were not measured in the KLIPS, which might be a source of discriminatory experience in getting hired. Third, a relatively small sample (n=3,479) was used to build a prediction model of hiring discrimination. A previous study indicated that machine learning algorithms could need more sample compared with logistic regression to avoid overfitting [44]. Therefore, future studies need to analyze larger sample size to build prediction model for hiring discrimination. Fourth, our results might not be generalized to other populations considering the time-point when data analyzed in our study was collected. For example, the Korean government has applied blind recruitment to national competency standards since 2017 [45]. The probability of hiring discrimination in 2004 could be different from those in 2021 when blind recruitment was applied. However, we could not find other datasets available for replicating our hypothesis.

1 Future studies should investigate whether the gender difference in underreporting of hiring
2 discrimination would be observed in other populations.

3 To the best of our knowledge, this is the first study to assess the magnitude of
4 underreporting of hiring discrimination using machine learning algorithms. Additionally, by
5 applying machine learning algorithms, this study showed overwhelming superiority over a
6 previous study investigating the predictive probability of having experienced hiring
7 discrimination based on the same dataset [21]. The previous study did not consider multiple
8 models for the prediction, did not assess the performance of their prediction model such as
9 AUC, and did not estimate the predictive probability of hiring discrimination for each
10 individual.

12 CONCLUSION

13 This study found that there could be a potential underestimation of hiring
14 discrimination among Korean workers due to under-reporting. Furthermore, female workers
15 were more likely to under-report their experience of hiring discrimination by answering NA.
16 Also, this study introduces a strategy to apply machine learning algorithms for social
17 epidemiology studies addressing the under-reporting of discriminatory experience.

References

1. Rasmussen, K.B., *Harm and Discrimination*. Ethical Theory and Moral Practice, 2019. **22**(4): p. 873-891.
2. Krieger, N., *Epidemiology and the people's health: theory and context*. 2011: Oxford University Press.
3. Krieger, N., *Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination*. Int J Health Serv, 1999. **29**(2): p. 295-352.
4. Inseo, S. and K. Seung-Sup, *A Systemic Review of Research on Perceived Discrimination and Health in South Korea*. Health and Social Welfare Review, 2015. **35**(1): p. 26-57.
5. Williams, D.R., et al., *Understanding how discrimination can affect health*. Health Services Research, 2019. **54**: p. 1374-1388.
6. Williams, D.R. and H. Neighbors, *Racism, discrimination and hypertension: evidence and needed research*. Ethnicity & disease, 2001. **11**(4): p. 800-816.
7. Dolezsar, C.M., et al., *Perceived racial discrimination and hypertension: a comprehensive systematic review*. Health Psychology, 2014. **33**(1): p. 20.
8. Lewis, T.T., et al., *Self-reported experiences of discrimination and cardiovascular disease*. Current cardiovascular risk reports, 2014. **8**(1): p. 365.
9. Chae, D.H., et al., *Do experiences of racial discrimination predict cardiovascular disease among African American men? The moderating role of internalized negative racial group attitudes*. Social science & medicine, 2010. **71**(6): p. 1182-1188.
10. Lee, H., et al., *Association between Discrimination and Depressive Symptoms Among 2,162 Lesbian, Gay, and Bisexual Adults in South Korea: Does Community Connectedness Modify the Association?* Journal of Homosexuality, 2019: p. 1-18.
11. Lee, N., et al., *Association between workplace discrimination and depressive symptoms among firefighters in South Korea*. American Journal of Industrial Medicine, 2018. **61**(9): p. 741-750.
12. Park, G.R., I. Son, and S.-S. Kim, *Perceived ethnic discrimination and depressive symptoms among biethnic adolescents in South Korea*. Journal of Preventive Medicine and Public Health, 2016. **49**(5): p. 301.
13. Kim, G. and M.-A. Lee, *Age discrimination and suicidal ideation among Korean older adults*. The American Journal of Geriatric Psychiatry, 2020. **28**(7): p. 748-754.
14. Kwon, S. and D. Han, *Discrimination, mental disorders, and suicidal ideation in Latino adults: Decomposing the effects of discrimination*. Journal of immigrant and minority health, 2019. **21**(1): p. 143-150.
15. Cheref, S., D. Talavera, and R.L. Walker, *Perceived discrimination and suicide ideation: moderating roles of anxiety symptoms and ethnic identity among Asian American, African American, and Hispanic emerging adults*. Suicide and Life-Threatening Behavior, 2019. **49**(3):

- p. 665-677.
16. Paradies, Y., et al., *Racism as a determinant of health: a systematic review and meta-analysis*. PloS one, 2015. **10**(9): p. e0138511.
 17. Krieger, N. and S. Sidney, *Racial discrimination and blood pressure: the CARDIA Study of young black and white adults*. American journal of public health, 1996. **86**(10): p. 1370-1378.
 18. Krieger, N., *Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination*. International journal of health services, 1999. **29**(2): p. 295-352.
 19. Hung, M., et al., *Exploring the Intersection between Social Determinants of Health and Unmet Dental Care Needs Using Deep Learning*. Int J Environ Res Public Health, 2020. **17**(19).
 20. Crosby, F., *The denial of personal discrimination*. American behavioral scientist, 1984. **27**(3): p. 371-386.
 21. Kim, S.-S., et al., *Measuring discrimination in South Korea: underestimating the prevalence of discriminatory experiences among female and less educated workers?* PloS one, 2012. **7**(3): p. e32872.
 22. Taylor, D.M., et al., *The personal/group discrimination discrepancy: Perceiving my group, but not myself, to be a target for discrimination*. Personality and social psychology bulletin, 1990. **16**(2): p. 254-262.
 23. Barreto, M. and N. Ellemers, *Detecting and Experiencing Prejudice*. 2015. p. 139-219.
 24. KREATSOULAS, C. and S.V. Subramanian, *Machine learning in social epidemiology: Learning from experience*. SSM Popul Health, 2018. **4**: p. 347-349.
 25. Seligman, B., S. Tuljapurkar, and D. Rehkopf, *Machine learning approaches to the social determinants of health in the health and retirement study*. SSM Popul Health, 2018. **4**: p. 95-99.
 26. Wiemken, T.L. and R.R. Kelley, *Machine Learning in Epidemiology and Health Outcomes Research*. Annu Rev Public Health, 2020. **41**: p. 21-36.
 27. Irvin, J.A., et al., *Incorporating machine learning and social determinants of health indicators into prospective risk adjustment for health plan payments*. BMC Public Health, 2020. **20**(1): p. 608.
 28. Chiavegatto Filho, A.D.P., et al., *Overachieving Municipalities in Public Health: A Machine-learning Approach*. Epidemiology, 2018. **29**(6): p. 836-840.
 29. Daoud, A., R. Kim, and S.V. Subramanian, *Predicting women's height from their socioeconomic status: A machine learning approach*. Social Science & Medicine, 2019. **238**: p. 112486.
 30. Kim, S.-S. and D.R. Williams, *Perceived discrimination and self-rated health in South Korea: a nationally representative survey*. PloS one, 2012. **7**(1): p. e30501.
 31. Inseo Son and Seung-Sup Kim, *A Systemic Review of Research on Perceived Discrimination*

- 1 *and Health in South Korea*. Health and Social Welfare Review, 2015. **35**(1): p. 26-57.
- 2 32. Jean, A., *The socio-economic background of the Gwangju Uprising*. New Political Science,
- 3 2003. **25**(2): p. 159-176.
- 4 33. Zocchetti, C., D. Consonni, and P.A. Bertazzi, *Estimation of prevalence rate ratios from cross-*
- 5 *sectional data*. International journal of epidemiology, 1995. **24**(5): p. 1064-1065.
- 6 34. Operario, D. and S.T. Fiske, *Ethnic identity moderates perceptions of prejudice: Judgments*
- 7 *of personal versus group discrimination and subtle versus blatant bias*. Personality and Social
- 8 Psychology Bulletin, 2001. **27**(5): p. 550-561.
- 9 35. Grollman, E.A. and N. Hagiwara, *Measuring self-reported discrimination: Trends in question*
- 10 *wording used in publicly accessible datasets*. Social Currents, 2017. **4**(4): p. 287-305.
- 11 36. Yan, T., *Consequences of Asking Sensitive Questions in Surveys*. Annual Review of Statistics
- 12 and Its Application, 2021. **8**: p. 109-127.
- 13 37. Tourangeau, R. and T. Yan, *Sensitive questions in surveys*. Psychological bulletin, 2007. **133**(5):
- 14 p. 859.
- 15 38. Williams, D.R. and S.A. Mohammed, *Discrimination and racial disparities in health: evidence*
- 16 *and needed research*. Journal of behavioral medicine, 2009. **32**(1): p. 20-47.
- 17 39. Karlsen, S. and J.Y. Nazroo, *Relation between racial discrimination, social class, and health*
- 18 *among ethnic minority groups*. American journal of public health, 2002. **92**(4): p. 624-631.
- 19 40. Krieger, N., *Racial and gender discrimination: risk factors for high blood pressure?* Social
- 20 science & medicine, 1990. **30**(12): p. 1273-1281.
- 21 41. Kim, J.-H., J.Y. Kim, and S.-S. Kim, *School violence, depressive symptoms, and help-seeking*
- 22 *behavior: a gender-stratified analysis of biethnic adolescents in South Korea*. Journal of
- 23 preventive medicine and public health, 2016. **49**(1): p. 61.
- 24 42. Kim, J.-H., et al., *Organizational response to workplace violence, and its association with*
- 25 *depressive symptoms: A nationwide survey of 1966 Korean EMS providers*. Journal of
- 26 occupational health, 2019. **61**(1): p. 101-109.
- 27 43. McLaughlin, K.A., M.L. Hatzenbuehler, and K.M. Keyes, *Responses to discrimination and*
- 28 *psychiatric disorders among Black, Hispanic, female, and lesbian, gay, and bisexual*
- 29 *individuals*. American journal of public health, 2010. **100**(8): p. 1477-1484.
- 30 44. Riley, R.D., et al., *Calculating the sample size required for developing a clinical prediction*
- 31 *model*. Bmj, 2020. **368**.
- 32 45. Korea, H.R.D.S.o. *Blind recruitment guide*. 2017 [cited 2021 Nov, 1]; Available from:
- 33 <https://www.ncs.go.kr/blind/bl01/RH-102-001-01.scdo>.

Figures

Figure 1. Flow chart of data analysis

1

2

3 Figure 2. Cross-validated performance of the machine learning algorithms according to the
 4 AUC

5

1 **Table**

2 Table 1. Distribution of the study population and prevalence of hiring discrimination by
3 predictors among waged workers in South Korea (n = 3576)

	Distribution	Prevalence of hiring discrimination	
	N (%)	N (%)	p-value*
Overall	3576 (100.0)	686 (19.2)	
Gender			0.131
Male	2165 (60.5)	395 (18.2)	
Female	1411 (39.5)	291 (20.6)	
Age			<0.001
16–24	277 (7.7)	54 (19.5)	
25–34	1124 (31.4)	185 (16.5)	
35–44	1025 (28.7)	178 (17.4)	
45–54	744 (20.8)	147 (19.8)	
55–64	306 (8.6)	88 (28.8)	
65+	100 (2.8)	34 (34.0)	
Education			<0.001
Middle school graduate or less	851 (23.8)	262 (30.8)	
High school graduate	1445 (40.4)	270 (18.7)	
College graduate or more	1280 (35.8)	154 (12.0)	
Marital status			<0.001
Never married	869 (24.3)	186 (21.4)	
Currently married	2501 (69.9)	432 (17.3)	
Previously married	206 (5.8)	68 (33.0)	
Employment status			<0.001
Permanent	2728 (76.3)	434 (15.9)	
Non-permanent	848 (23.7)	252 (29.7)	
Household income			<0.001
Less than 1Q	554 (15.5)	164 (29.6)	
1Q-2Q	909 (25.4)	232 (25.5)	
2Q-3Q	1005 (28.1)	165 (16.4)	
3Q-	1108 (31.0)	125 (11.3)	
Birth region			0.082
Other regions	2892 (80.9)	574 (19.8)	
Jeolla-do	684 (19.1)	112 (16.4)	
Self-rated health conditions			<0.001
Very good	161 (4.5)	18 (11.2)	
Good	2061 (57.6)	362 (17.6)	
Fair	1115 (31.2)	236 (21.2)	
Poor and Very Poor	239 (6.7)	70 (29.3)	
Having a disability			<0.001
No	3484 (97.4)	651 (18.7)	
Yes	92 (2.6)	35 (38.0)	

Residential area			<0.001
Seoul	811 (22.7)	100 (12.3)	
Busan	354 (9.9)	123 (34.7)	
Daegu	212 (5.9)	60 (28.3)	
Daejeon	112 (3.1)	22 (19.6)	
Incheon	258 (7.2)	30 (11.6)	
Gwangju	98 (2.7)	21 (21.4)	
Ulsan	115 (3.2)	31 (27.0)	
Gyeonggi-do	769 (21.5)	119 (15.5)	
Gangwon-do	40 (1.1)	9 (22.5)	
Chungcheongbuk-do	80 (2.2)	18 (22.5)	
Chungcheongnam-do	101 (2.8)	18 (17.8)	
Jeollabuk-do	140 (3.9)	24 (17.1)	
Jeollanam-do	80 (2.2)	10 (12.5)	
Gyeongsangbuk-do	157 (4.4)	42 (26.8)	
Gyeongsangnam-do	249 (7.0)	59 (23.7)	
Discriminatory experience (income)			<0.001
No	2980 (83.3)	260 (8.7)	
Yes	545 (15.2)	418 (76.7)	
Not applicable	51 (1.4)	8 (15.7)	
Discriminatory experience (Training)			<0.001
No	3136 (87.7)	470 (15.0)	
Yes	71 (2.0)	53 (74.6)	
Not applicable	369 (10.3)	163 (44.2)	
Discriminatory experience (Promotion)			<0.001
No	2924 (81.8)	398 (13.6)	
Yes	209 (5.8)	108 (51.7)	
Not applicable	443 (12.4)	180 (40.6)	
Discriminatory experience (Fired)			<0.001
No	3098 (86.6)	501 (16.2)	
Yes	65 (1.8)	53 (81.5)	
Not applicable	413 (11.5)	132 (32.0)	
Discriminatory experience (Education)			<0.001
No	3383 (94.6)	597 (17.6)	
Yes	38 (1.1)	21 (55.3)	
Not applicable	155 (4.3)	68 (43.9)	
Discriminatory experience (Home)			<0.001
No	3470 (97.0)	631 (18.2)	
Yes	75 (2.1)	46 (61.3)	
Not applicable	31 (0.9)	9 (29.0)	
Discriminatory experience (Social activities)			<0.001
No	3272 (91.5)	508 (15.5)	
Yes	282 (7.9)	171 (60.6)	
Not applicable	22 (0.6)	7 (31.8)	

*P-value of the Chi-square test comparing the prevalence of discriminatory experiences in getting hired across different categories for each predictor

Table 2. Gender difference in under-reporting hiring discrimination based on the random forest prediction

	Total	Prevalence of hiring discrimination	Prevalence Ratio (95% CI)
	N	N (%)	
Training sample (Yes or No group)	3,479	686 (19.7)*	2.98 (2.49–3.57)
Prediction sample (NA group)	97	57 (58.8)**	
Male (n=2,165)			
Training sample	2,101	395 (18.2)*	2.41 (1.82–3.20)
Prediction sample	64	29 (45.3)**	
Female (n=1,411)			
Training sample	1,378	291 (20.6)*	4.02 (3.37–4.79)
Prediction sample	33	28 (84.8)**	

*Observed value, **Predicted value

Supplementary materials

List of contents

Figure S1. Variable importance scores for the predictors in the best-performing machine learning algorithm (random forest).

Figure S2. ROC curve of cross-validated classification performance of each model

Figure S3. Predictive performance of the best-performing machine learning algorithm (random forest) across different probability threshold values, and the optimal value of probability threshold for classification.

Figure S4. Predictive performance of the best-performing machine learning algorithm (random forest) across different probability threshold values, and the optimal value of probability threshold for classification in the sensitivity analysis.

Figure S5. Cross-validated performance of the machine learning algorithms according to the AUC in sensitivity analyses 2 & 3.

Figure S6. ROC curve of cross-validated classification performance of each model in sensitivity analyses 2 & 3.

Table S1. Machine learning algorithms and tuning parameters for each model

Table S2. Overview of sensitivity analysis

Table S3. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 1

Table S4. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 2

Table S5. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 3

Table S6. Association between the experience of hiring discrimination and poor self-rated health among waged workers in South Korea

Figure S1. Variable importance scores for the predictors in the best-performing machine learning algorithm (random forest).

Table S1. Machine learning algorithms and tuning parameters for each model

Algorithms	Tuning parameters
Logistic regression	
Random Forest	mtry (the number of predictors that will be randomly sampled at each split when creating the tree models), min_n (the minimum number of data points in a node that is required for the node to be split further)
K-nearest neighbor	neighbors (the number of neighbors to consider), weight_func (the type of kernel function used to weight distances between samples), dist_power (the parameter used in calculating Minkowski distance)
Ridge regression	penalty (the total amount of regularization)
LASSO regression	penalty (the total amount of regularization)
Elastic net	penalty (the total amount of regularization), mixture (the proportion of lasso regularization in the model)
Support Vector Machine with radial	cost (the cost of predicting a sample within or on the wrong side of the margin), rbf_sigma (A positive number for radial basis function), margin (epsilon in the SVM insensitive loss function)
Support Vector Machine with polynomial	cost (cost of predicting a sample within or on the wrong side of the margin), scale_factor (polynomial scaling factor),

	margin (epsilon in the SVM insensitive loss function), degree (polynomial degree)
Neural networks	hidden_units (the number of units in the hidden model.), dropout (the proportion of model parameters randomly set to zero during model training), penalty (the amount of weight decay)

Epub ahead of print

Figure S2. ROC curve of the cross-validated classification performance of each model

Figure S3. Predictive performance of the best-performing machine learning algorithm (random forest) across different probability threshold values, and the optimal value of probability threshold for classification.

Table S2. Overview of the sensitivity analyses

Analysis	Study population		Probability threshold for classification	
	N	Inclusion criteria	Value	Selection criteria
Main analysis	35 76	Waged workers in the 7th wave of KLIPS who had complete information for all predictors except those who answered 'Not Applicable' to all questions about the experience of discrimination	0.2678 58	Maximizing the sum of sensitivity and specificity
Sensitivity analysis 1	35 76	Waged workers in the 7th wave of KLIPS who had complete information for all predictors	0.1758 445	Minimizing the difference between sensitivity and specificity
Sensitivity analysis 2	36 36	Waged workers in the 7th wave of KLIPS who had complete information for all predictors	0.2800 794	Maximizing the sum of sensitivity and specificity
Sensitivity analysis 3	36 36	Waged workers in the 7th wave of KLIPS who had complete information for all predictors	0.1757 411	Minimizing the difference between sensitivity and specificity

Figure S4. Predictive performance of the best-performing machine learning algorithm (random forest) across different probability threshold values, and the optimal value of probability threshold for classification in the sensitivity analyses.

Figure S5. Cross-validated performance of the machine learning algorithms according to the AUC in sensitivity analyses 2 & 3.

Figure S6. ROC curve of cross-validated classification performance of each model in sensitivity analyses 2 & 3.

Table S3. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 1

	Total	Prevalence of hiring discrimination	Prevalence Ratio (95% CI)
	N	N (%)	
Training sample (Yes or No group)	3,479	686 (19.7)*	3.29 (2.80–3.87)
Prediction sample (NA group)	97	63 (64.9)**	
Male (n=2,165)			
Training sample	2,101	395 (18.2)*	2.74 (2.13–3.53)
Prediction sample	64	33 (51.6)**	
Female (n=1,411)			
Training sample	1,378	291 (20.6)*	4.30 (3.71–4.99)
Prediction sample	33	30 (90.9)**	

*Observed value, **Predicted value

Table S4. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 2

	Total	Prevalence of hiring discrimination	Prevalence Ratio (95% CI)
	N	N (%)	
Training sample (Yes or No group)	3,482	688 (19.8)*	3.75 (3.34–4.20)
Prediction sample (NA group)	154	114 (74.0)**	
Male (n=2,196)			
Training sample	2,103	397 (18.9)*	3.30 (2.76–3.96)
Prediction sample	93	58 (62.4)**	
Female (n=1,440)			
Training sample	1,379	291 (21.1)*	4.35 (3.83–4.94)
Prediction sample	61	56 (91.8)**	

*Observed value, **Predicted value

Table S5. Gender difference in under-reporting hiring discrimination based on the random forest prediction from sensitivity analysis 3

	Total	Prevalence of hiring discrimination	Prevalence Ratio (95% CI)
	N	N (%)	
Training sample (Yes or No group)	3,482	688 (19.8)*	3.95 (3.54–4.39)
Prediction sample (NA group)	154	120 (77.9)**	
Male (n=2,196)			
Training sample	2,103	397 (18.9)*	3.53 (2.98–4.18)
Prediction sample	93	62 (66.7)**	
Female (n=1,440)			
Training sample	1,379	291 (21.1)*	4.51 (4.01–5.06)
Prediction sample	61	58 (95.1)**	

*Observed value, **Predicted value

Table S6. Association between the experience of hiring discrimination and poor self-rated health among waged workers in South Korea (n = 3,576)

Experience of hiring discrimination	Distribution	Prevalence of poor self-rated health	Prevalence Ratio	95% CI
	N (%)	N (%)		
Total (n=3,576)				
No	2,793 (78.1)	156 (5.6)	Reference	
Yes	686 (19.2)	70 (10.2)	1.83	1.40–2.39
NA-No	40 (1.1)	4 (10.0)	1.79	0.70–4.59
NA-Yes	57 (1.6)	9 (15.8)	2.83	1.52–5.25
Male (n = 2,165)				
No	1,706 (78.8)	86 (5.0)	Reference	
Yes	395 (18.2)	32 (8.1)	1.61	1.09–2.38
NA but predicted as No	35 (1.6)	3 (8.6)	1.70	0.57–5.12
NA but predicted as Yes	29 (1.3)	3 (10.3)	2.05	0.69–6.11
Female (n = 1,411)				
No	1,087 (77.0)	70 (6.4)	Reference	
Yes	291 (20.6)	38 (13.1)	2.03	1.40–2.94
NA but predicted as No	5 (0.4)	1 (20.0)	3.11	0.53–18.19
NA but predicted as Yes	28 (2.0)	6 (21.4)	3.33	1.58–7.01

We examined the association between hiring discrimination and self-rated health by applying a modified Poisson regression model with a robust error variance after dividing the response to hiring discrimination among the entire population into four groups

Epub ahead of print